Structural Biology and Chemistry of the Terpenoid Cyclases

David W. Christianson*

Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104-6323

Received October 21, 2005

Contents

1. Introduction 3412
2. Monoterpene Cyclases 3416
3. Sesquiterpene Cyclases 3418
   3.1. Trichodiene Synthase 3418
   3.2. Epi-Aristolochene Synthase 3424
   3.3. Aristolochene Synthase 3425
   3.4. Pentalenene Synthase 3426
4. Diterpene Cyclases 3428
   4.1. Abietadiene Synthase 3428
   4.2. Taxadiene Synthase 3429
5. Triterpene Cyclases 3430
   5.1. Squalene–Hopene Cyclase 3430
   5.2. Oxidosqualene Cyclase 3432
6. Antibody Catalysis of Cationic Cyclization Reactions 3433
   6.1. Catalytic Antibody 19A4 3434
   6.2. Catalytic Antibody 4C6 3437
7. Concluding Remarks 3437
8. Acknowledgments 3441
9. References 3441

Born in Attleboro, Massachusetts, in 1961, David W. Christianson received the A.B. (1983), A.M. (1985), and Ph.D. (1987) degrees in chemistry from Harvard, where he studied in the research group of William N. Lipscomb, Jr. Christianson joined the faculty of the University of Pennsylvania in 1988, where he is currently the Roy and Diana Vagelos Professor in Chemistry and Chemical Biology. Christianson's research focuses on the structures and mechanisms of metal-requiring enzymes.

1. Introduction

Terpenoid cyclases catalyze the most complex chemical reactions occurring in Nature, in that on average, two-thirds of the carbon atoms of a linear polyisoprenoid substrate undergo changes in bonding, hybridization, or configuration during the course of an intricate cyclization cascade initiated by the formation and propagation of highly reactive carbocation intermediates.[1−10] Tens of thousands of terpene natural products have been characterized to date, yet each ultimately derives from one of only a handful of linear isoprenoid substrates (Figure 1). Terpene biosynthesis thus provides an elegant example of Nature's strategy for combinatorial chemical synthesis and diversity.

Most terpenoid cyclases generate a unique hydrocarbon product with remarkable structural and stereochemical precision. For example, consider the reaction catalyzed by oxidosqualene cyclase, also known as lanosterol synthase, in the biosynthesis of cholesterol (Figure 1). Only 1 out of a possible 128 lanosterol stereoisomers is formed in the cyclization of the 30-carbon triterpene (*S*)-2,3-oxidosqualene, so the enzyme active site clearly serves as a high-fidelity template that chaperones the flexible polyisoprenoid substrate and carbocation intermediates through a series of precise conformations leading to one unique cyclization product. In contrast, other terpenoid cyclases are rather promiscuous in their cyclization chemistry: extreme examples include the conifer sesquiterpene cyclases δ-selinene synthase and γ-humulene synthase, which generate a total of 34 and 52 cyclization products, respectively.[11] Possibly, multiple product formation could reflect incomplete evolution of a precise active site template in these cyclases.

The terpene cyclase substrates geranyl diphosphate (C_{10}), farnesyl diphosphate (C_{15}), geranylgeranyl diphosphate (C_{20}), and geranylfarnesyl diphosphate (C_{25}) (Figure 1) each derive from a chain elongation reaction catalyzed by an isoprenyl diphosphate synthase (also known as a prenyltransferase) utilizing the five-carbon precursors dimethylallyl diphosphate and isopentenyl diphosphate.[12] The 30-carbon substrate squalene derives from the head-to-head condensation of two molecules of farnesyl diphosphate in a reaction catalyzed by squalene synthase.[13,14] Remarkably, the enzymes that generate terpene cyclase substrates are themselves structurally related to class I terpene cyclases (Figure 2), indicative of their divergence from a common primordial ancestor early in the evolution of terpene biosynthesis. This structural homology is reminiscent of suggestions that enzymes that catalyze successive steps in biosynthetic pathways may share common evolutionary origins.[15]

Given its prominence as the archetypical prenyltransferase, and given the prominent structural and functional homologies

* Phone: 215-898-5714. Fax: 215-573-2201. E-mail: chris@sas.upenn.edu.

10.1021/cr050286w CCC: $59.00
Published on Web 06/01/2006

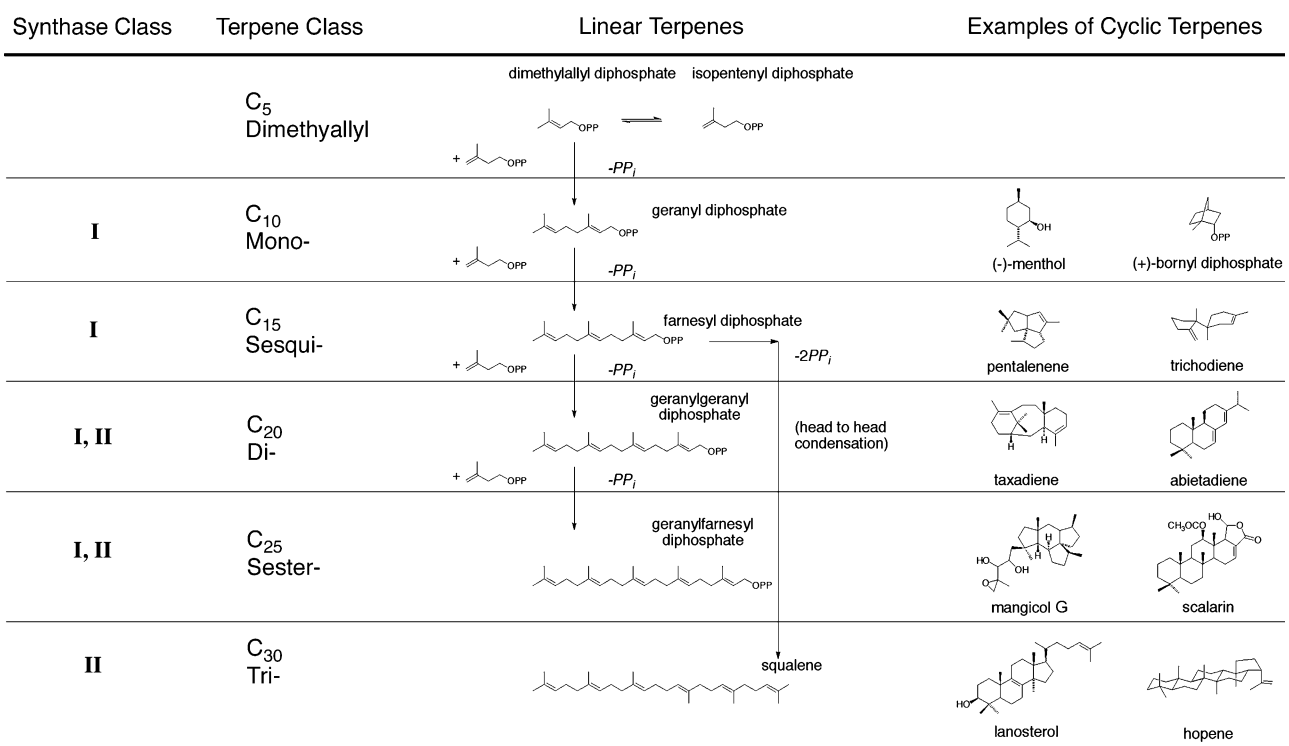


Figure 1. General scheme of terpenoid nomenclature and biosynthesis (OPP = diphosphate, PP_i = inorganic diphosphate).

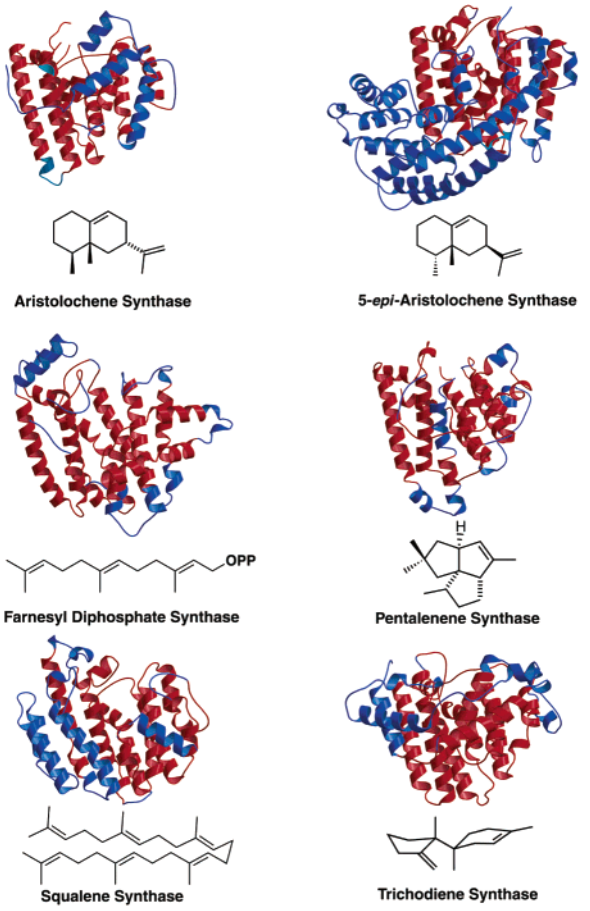
between the prenyltransferases and the terpenoid cyclases, farnesyl diphosphate synthase provides a suitable introduction and foundation for understanding terpenoid cyclase structure and mechanism. Avian farnesyl diphosphate synthase is a homodimer of 44 kD α-helical subunits with a core of 10 α-helices surrounding a hydrophobic active site of approximate volume 2400 Å^3 (Figures 3 and 4).[16] Helices D and H contain so-called "aspartate-rich" motifs, DDXXD, implicated in catalytic function.[17−19] The structure of the unliganded wild-type enzyme[16] and the structure of the Phe-112→Ala/Phe-113→Ser variant complexed with geranyl diphosphate[20] reveal that the first residue in the first aspartate-rich motif, Asp-117, makes a *syn,syn*-bidentate coordination interaction with two Mg^{2+} ions, and one carboxylate oxygen of the last residue in this motif, Asp-121, makes a *syn,anti*-bridging coordination interaction between the two metal ions (Figure 5).

The recent crystal structure determination of farnesyl diphosphate synthase from *Escherichia coli* complexed with substrate isopentenyl diphosphate and the nonreactive substrate analogue dimethylallyl-*S*-thiolodiphosphate[21] reveals the first view of a complete trinuclear magnesium cluster liganded by both aspartate-rich motifs and the diphosphate group of dimethylallyl-*S*-thiolodiphosphate (Figure 6).[22] The $Mg^{2+}{}_1$ and $Mg^{2+}{}_2$ ions each form six-membered ring chelate structures with thiolodiphosphate oxygen atoms. The first aspartate of the first aspartate-rich motif, Asp-105 on helix D, makes *syn,syn*-bidentate coordination interactions with $Mg^{2+}{}_2$ and $Mg^{2+}{}_3$. One carboxylate oxygen of the third aspartate in the motif, Asp-111, bridges $Mg^{2+}{}_2$ and $Mg^{2+}{}_3$ with *syn,anti*-coordination stereochemistry. Notably, the first aspartate-rich motif in the bacterial enzyme is of the form DDXXXXD; this contrasts with the avian enzyme, in which the characteristic DDXXD pattern is observed. Regardless, the first and last aspartates of the first aspartate-rich motif in each enzyme make identical interactions with a pair of magnesium ions (Figures 5 and 6). The first aspartate in the second aspartate-rich motif on helix H, Asp-244, coordinates to $Mg^{2+}{}_1$.

The structure of the ternary enzyme−substrate−substrate analogue complex[22] provides a foundation for understanding the electrophilic alkylation mechanism proposed for isoprenoid chain elongation in which the metal-triggered ionization of dimethylallyl diphosphate yields an allylic carbocation−pyrophosphate ion pair,[23−25] followed by alkylation of the C3−C4 π bond of isopentenyl diphosphate to yield a tertiary C3 carbocation;[26,27] stereospecific elimination of a C2 proton,[28,29] proposed to be mediated by a pyrophosphate oxygen,[22] terminates the chain elongation reaction to yield geranyl diphosphate.[26] A second condensation reaction with the addition of isopentenyl diphosphate to geranyl diphosphate yields product farnesyl diphosphate (Figure 7). Interestingly, farnesyl diphosphate synthase can be engineered to generate longer chain products such as geranylgeranyl diphosphate by site-directed mutagenesis of aromatic residues at the base of the active site cleft; the deeper active site cleft allows for the synthesis of a longer product.[20,30]

Like the prenyltransferase reaction, the terpenoid cyclase reaction is initiated by the formation of a highly reactive carbocation. The terpenoid cyclases utilize two main chemical strategies for initial carbocation formation: (1) metal-triggered departure of a pyrophosphate (PP_i) leaving group, similar to the strategy employed by farnesyl diphosphate synthase, or (2) protonation of an epoxide ring or a carbon−carbon double bond, a strategy employed by oxidosqualene cyclase or squalene−hopene cyclase, respectively. Thus, the initiation step of a terpene cyclase reaction is either ionization-dependent or protonation-dependent.

Ionization-dependent terpenoid cyclases—monoterpene, sesquiterpene, and diterpene cyclases—contain an aspartate-rich sequence, DDXXD/E, on helix D that is topologically identical to the first such sequence occurring in the prenyltransferases on helix D, but the second such sequence on helix H is not "aspartate-rich" in the terpenoid cyclase: it



	FDPS	AS	EAS	PS	SQS
TS	3.9 Å (238)	4.0 Å (205)	3.2 Å (212)	3.3 Å (251)	3.6 Å (192)
FDPS	-	3.7 Å (193)	3.4 Å (244)	2.9 Å (159)	3.5 Å (215)
AS	-	-	3.1 Å (222)	3.0 Å (291)	3.6 Å (191)
EAS	-	-	-	3.2 Å (238)	3.6 Å (227)
PS	-	-	-	-	3.4 Å (191)

Figure 2. Structural similarities among various terpenoid synthases (red) define the core class I terpenoid cyclase fold. The product of each synthase is indicated. At the bottom, rms deviations of Cα atoms and the number of structurally similar residues (in parentheses) are listed. Reprinted with permission from ref 60. Copyright 2001 National Academy of Sciences, U.S.A.

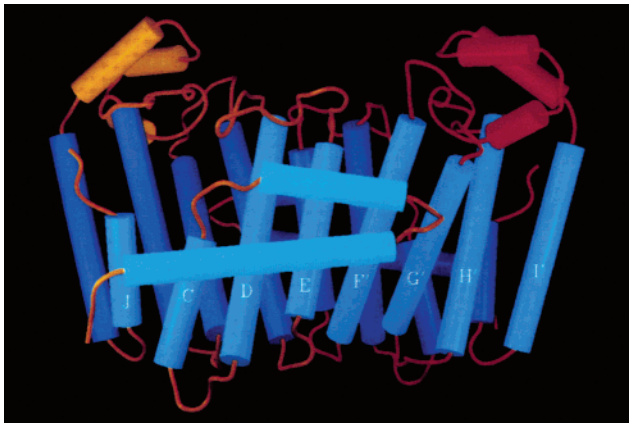


Figure 3. Dimer of avian farnesyl diphosphate synthase. α-Helices are represented by cylinders. Helices of monomer 1 are labeled with capital letters and connected by orange loops, and helices of monomer 2 are labeled with primed capital letters and connected by magenta loops. The 2-fold axis of the dimer is approximately vertical, and the two active site clefts are oriented in parallel fashion and open at the "top" of the molecule. Reprinted with permission from ref 16. Copyright 1994 American Chemical Society.

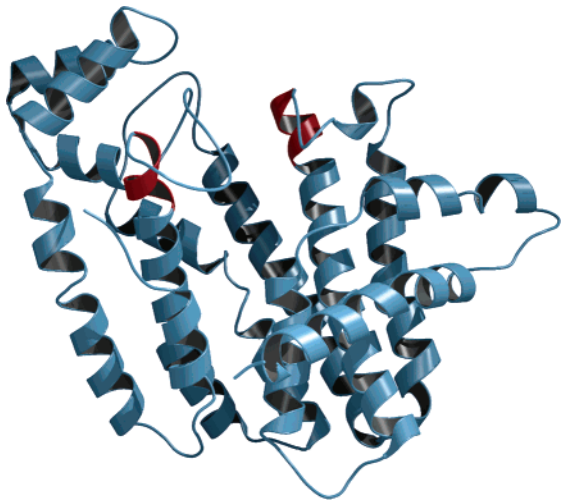


Figure 4. Ribbon plot of the avian farnesyl diphosphate synthase. Two aspartate-rich motifs, Asp-117−Asp-121 and Asp-257−Asp-261 (red), flank the mouth of the active site.[16]

has diverged to a consensus sequence of (L,V)(V,L,A)-(**N**,**D**)D(L,I,V)X(**S**,**T**)XXX**E** (boldface residues are magnesium ligands, so this motif is also designated the "NSE/DTE" motif).[31] Both motifs together bind a trinuclear magnesium cluster that triggers the departure of the substrate diphosphate leaving group, as further outlined in sections 2, 3.1, and 3.2. In contrast, protonation-dependent terpenoid cyclases—certain diterpene cyclases and triterpene cyclases—contain a single signature sequence DXDD, in which the central aspartate residue is implicated as the proton donor that triggers initial carbocation formation.[10]

The common α-helical fold of the ionization-dependent terpenoid cyclases (Figure 2) is designated the class I terpenoid cyclase fold, and the unrelated α-barrel fold of the protonation-dependent cyclases (discussed in section 5) is designated the class II terpenoid cyclase fold.[6] Interestingly, plant cyclases such as (+)-bornyl diphosphate synthase (section 2) or epi-aristolochene synthase (section 3.2) contain both domains, but only the class I domain is the catalytically active terpenoid cyclase. Strikingly, the diterpene cyclase abietadiene synthase from grand fir also contains both domains, and each catalyzes a discrete reaction in the cyclization of geranylgeranyl diphosphate to form abietadiene (section 4.1).

It should be noted that the initiation step of a terpene synthase can result in direct carbocation formation, for example, the departure of the pyrophosphate leaving group from dimethylallyl diphosphate yields an allylic cation that subsequently reacts with the C2−C3 π-bond of isopentenyl diphosphate (Figure 7). Alternatively, the initiation step can be accompanied by direct attack of one of the remaining π bonds at the site of incipient carbocation formation in concert with leaving group departure to form a carbocation at a more distant carbon center (S_N' reaction[32]). The reactive π bond may open in either Markovnikov fashion (forming the more stable carbocation) or anti-Markovnikov fashion (forming the less stable carbocation). Notably, the electrostatic environment of the active site can contribute to the preferential stabilization of carbocations resulting from anti-Markovnikov processes that would ordinarily be disfavored in the absence of enzyme. This expands the potential product diversity for the cyclization of an isoprenoid diphosphate substrate. For

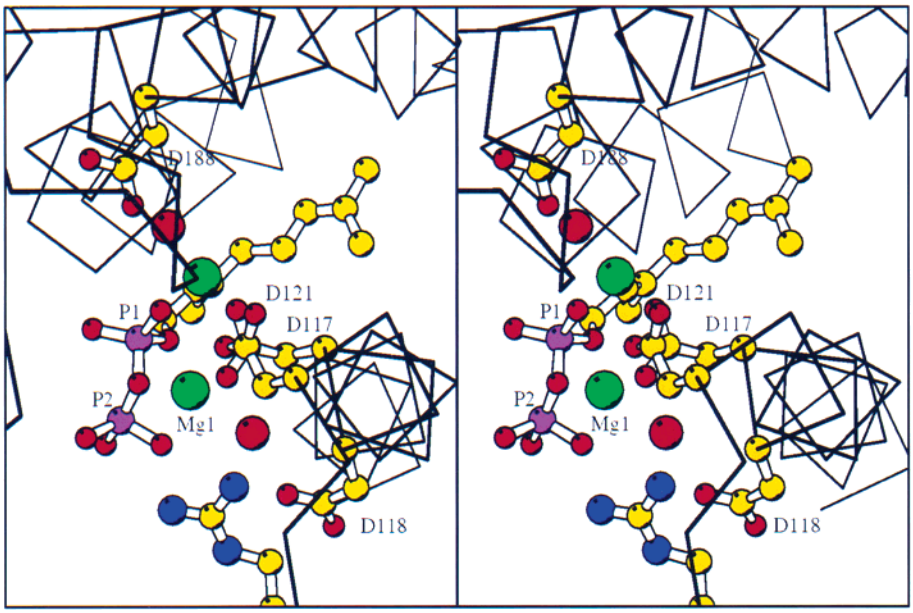


Figure 5. The binding of geranyl diphosphate in the active site of the Phe-112 → Ala/Phe-113 → Ser variant of avian farnesyl diphosphate synthase reveals that Asp-117 and Asp-121 of the first aspartate-rich motif make bridging coordination interactions with two Mg^{2+} ions (green). Ordered water molecules are shown as red spheres. Reprinted with permission from ref 20. Copyright 1996 National Academy of Sciences, U.S.A.

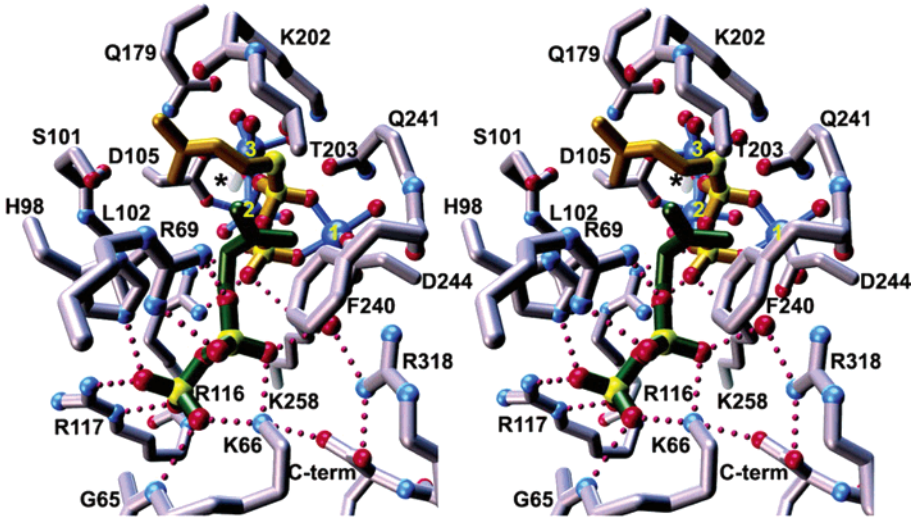


Figure 6. The active site of *E. coli* farnesyl diphosphate synthase complexed with three Mg^{2+} ions (blue spheres 1, 2, and 3), dimethylallyl-*S*-thiolodiphosphate (yellow), and isopentenyl diphosphate (green). For clarity, Asp-111 is indicated by an asterisk. Metal coordination and hydrogen bond interactions are indicated by solid blue and dotted magenta lines, respectively. Reprinted with permission from ref 22. Copyright 2004 The American Society for Biochemistry & Molecular Biology.

example, four possible outcomes for initial carbon−carbon bond formation triggered by farnesyl diphosphate ionization are illustrated in Figure 8.[33] Subsequent reaction of the remaining π bonds with the four different carbocation intermediates in Figure 8, followed by hydride shifts, methyl migrations, etc., can create a potentially vast and diverse product array in the cyclization of farnesyl diphosphate.

With carbocation intermediates thus firmly implicated in the chemical mechanism of isoprenoid biosynthesis and cyclization, how are these highly reactive intermediates handled in the relatively mild milieu of an enzyme active site without alkylating polar residues or being quenched by solvent? In short, the hydrophobic prenyltransferase or terpene cyclase active site serves as an "inert organic solvent" for carbocation-mediated chain elongation or cyclization reactions. Moreover, just as carbocation reactions are facilitated in polar, aprotic organic solvents capable of stabilizing transiently formed carbocation intermediates through dipole−charge interactions, certain fixed and protected dipoles in an enzyme active site can similarly stabilize carbocation intermediates. Additionally, carbocation reaction intermediates can be stabilized by the ring π electrons of aromatic residues such as phenylalanine, tyrosine, and tryptophan through cation−π or charge−quadrupole interactions (Figures 9 and 10).[34,35]

In the remainder of this review, I shall outline the currently available X-ray crystal structures and chemical mechanisms for ionization-dependent monoterpene and sesquiterpene cyclases in sections 2 and 3. Although no diterpene cyclase has yet yielded a crystal structure, important structure−mechanism analogies can be made for this family of terpene cyclases based on the structures of monoterpene and sesquiterpene cyclases, and these are reviewed in section 4. It is especially intriguing that the bifunctional diterpene cyclase

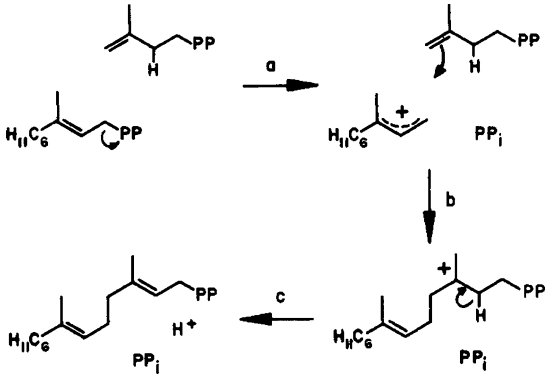


Figure 7. Chain elongation reaction catalyzed by farnesyl diphosphate synthase: ionization of geranyl diphosphate (a) leads to the formation of the corresponding allylic cation, which then undergoes condensation with isopentenyl diphosphate (b) to yield a tertiary C3 carbocation, from which stereospecific elimination of a proton (c) yields farnesyl diphosphate. PP = diphosphate; PP_i = inorganic pyrophosphate. Reprinted with permission from ref 25. Copyright 1981 American Chemical Society.

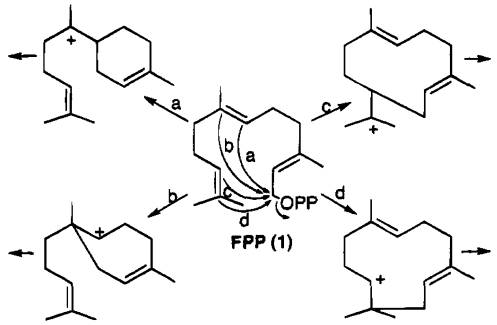


Figure 8. Four possible initial ring closure reactions for farnesyl diphosphate (FPP), the substrate of a sesquiterpene cyclase. By enforcing a particular conformation of the flexible isoprenoid substrate, the enzyme active site plays a critical role in directing initial carbon−carbon bond formation to generate tertiary carbocation intermediates (Markovnikov addition to a double bond, pathways a and c) or less stable secondary carbocation intermediates (anti-Markovnikov addition to a double bond, pathways b and d). Pathways a and b require initial ionization and isomerization to form nerolidyl diphosphate, reionization of which allows for the incorporation of a cis double bond in the six-membered or seven-membered ring, respectively. Reprinted with permission from ref 33. Copyright 1995 American Chemical Society.

abietadiene synthase contains separate ionization-dependent and protonation-dependent active sites. The triterpene cyclases are exclusively protonation-dependent, and squalene−hopene cyclase and oxidosqualene cyclase are reviewed in section 5. In section 6, progress in the generation of novel antibody catalysts of isoprenoid/polyene cyclization reactions is discussed, and I conclude in section 7 with comments on emerging structure−mechanism relationships common to the greater family of terpenoid cyclases.

2. Monoterpene Cyclases

Each member of the family of monoterpene cyclases catalyzes the metal-dependent ionization and cyclization of the linear 10-carbon isoprenoid precursor geranyl diphosphate (Figure 1). Although monoterpene cyclases catalyze the simplest terpene cyclization cascades in that they utilize the shortest linear isoprenoid cyclization substrate, these enzymes nevertheless achieve remarkable structural and stereochemical diversity in their product arrays (Figure 11).[9] The volatile

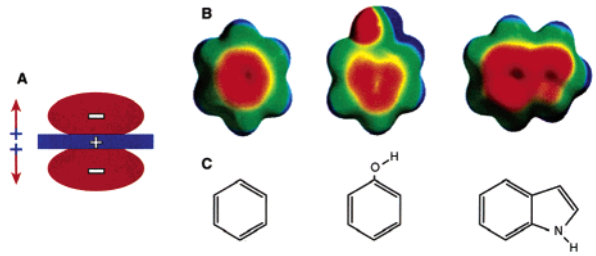


Figure 9. (a) The side chain of an aromatic residue can be described as an electronic quadrupole, the characteristics of which are no net charge and opposed dipoles. Electrostatic surface potentials for benzene, phenol, and indole (b, c) show that the faces of these aromatic rings bear significant positive electrostatic surface potential, and the same is expected for the corresponding side chains of the amino acids phenylalanine, tyrosine, and tryptophan, respectively (red = negative, blue = positive). Aromatic residues in a terpene synthase active site may stabilize reactive carbocation intermediates without the danger of reacting with such intermediates. Reprinted with permission from *Science* (http://www.aaas.org), ref 34. Copyright 1996 American Association for the Advancement of Science.

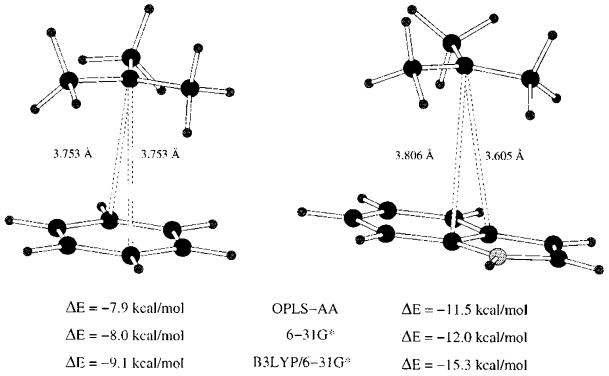


Figure 10. Optimized geometries calculated with the OPLS-AA force field for cation−π interactions between the 2-methyl-2-propyl cation and benzene (left) and indole (right). Stabilization energies for complex formation from the separated cation and molecule are thermodynamically favorable and are calculated using three different computational procedures. The optimal geometries of these theoretical cation−π interactions can easily be achieved in a terpenoid cyclase active site for cation interactions with the π systems of the side chains of phenylalanine, tyrosine, or tryptophan. Reprinted with permission from ref 35. Copyright 1997 American Chemical Society.

monoterpenes and their derivatives are largely responsible for the pleasing fragrances and flavors of aromatic plants.

(+)-Bornyl diphosphate synthase, a monoterpene cyclase from *Salvia officinalis* (culinary sage), is a 128 kD homodimer that catalyzes an unusual cyclization cascade in that the PP_i leaving group of geranyl diphosphate is reincorporated into the final cyclization product (Figure 12).[36] Positional isotope exchange studies with ^{18}O-labeled substrate demonstrate that the same phosphoester oxygen atom is contained in the prenyl diphosphate linkages of substrate geranyl diphosphate and product (+)-bornyl diphosphate,[37,38] suggesting very precise positioning and orientation of the diphosphate moiety throughout the cyclization cascade as it is covalently bound or ion-paired with substrate, intermediates, and product. Mechanistic features of the cyclization cascade have been carefully delineated in enzymological studies (Figure 12).[37−42]

The X-ray crystal structure of (+)-bornyl diphosphate synthase[43] reveals a C-terminal domain that exhibits the characteristic class I terpene cyclase fold: this is where the

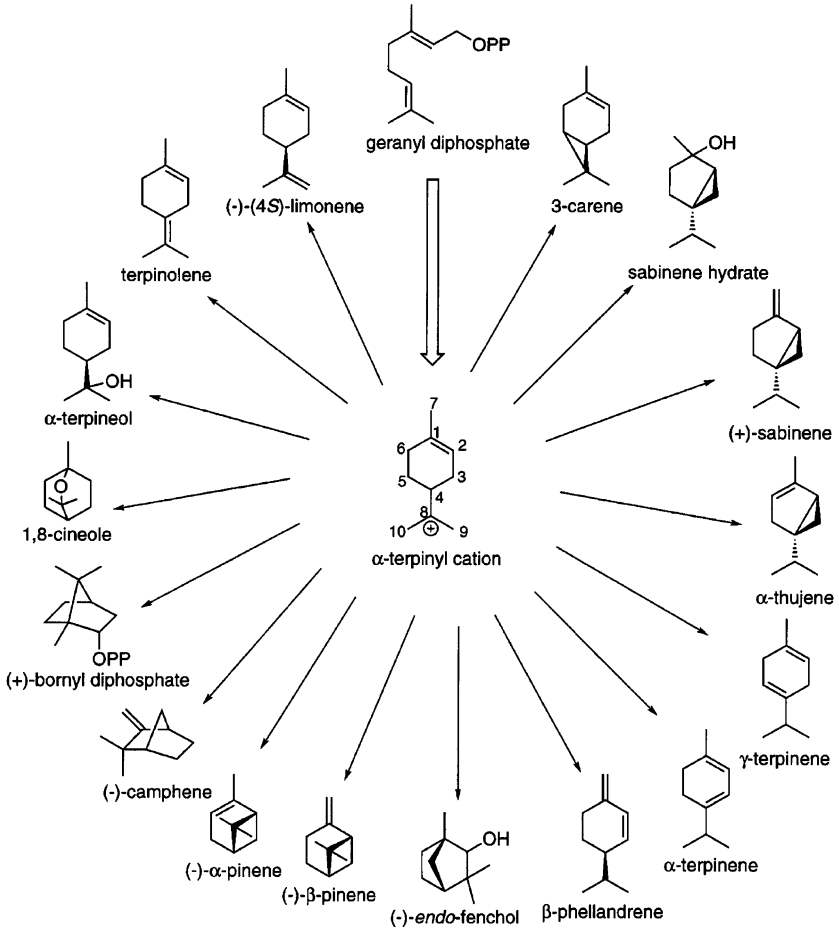


Figure 11. The cyclization of geranyl diphosphate by monoterpene cyclases yields a diverse array of monocyclic and bicyclic products emanating from the α-terpinyl carbocation intermediate. Reprinted with kind permission of Springer Science and Business Media from ref 9. Copyright 2000 Springer-Verlag.

metal-dependent cyclization of geranyl diphosphate is catalyzed (Figure 13). The N-terminal glycosyl hydrolase-like domain has no well-defined function, although the N-terminal polypeptide caps the active site in the C-terminal domain upon the binding of the diphosphate group of the substrate analogue 3-aza-2,3-dihydrogeranyl diphosphate, the PP_i group in ternary complex with 7-aza-7,8-dihydrolimonene or 2-azabornane, or product (+)-bornyl diphosphate (Figure 13).[43] The N-terminal polypeptide segment is involved in numerous hydrogen bond interactions with the C-terminal domain, including two with the aspartate-rich motif (Arg-56−Asp-355 carbonyl and Tyr-60−Asp352). Interestingly, tandem arginines Arg-55−Arg-56 of (+)-bornyl diphosphate synthase are reminiscent of tandem arginines found in the N-terminal segments of several (but not all) plant terpene cyclases. Truncation studies with limonene synthase implicate the corresponding tandem arginines in the isomerization step of the mechanism,[44] and it is conceivable that the tandem arginines in this monoterpene cyclase serve a similar function as Arg-55−Arg-56 of (+)-bornyl diphosphate synthase. A similar proposal is considered for diterpene cyclases containing an Arg−Arg or Lys−Arg pair in the N-terminus (reviewed in section 4).

The structures of six (+)-bornyl diphosphate synthase complexes with prenyl diphosphates or PP_i and aza analogues of carbocation intermediates reveal essentially identical molecular recognition of the diphosphate moiety involving a trinuclear magnesium cluster and several hydrogen bond interactions (Figure 14).[43] The first aspartate residue in the aspartate-rich motif, Asp-351, bridges Mg_A^{2+} and Mg_C^{2+} with *syn*,*syn*-bidentate coordination stereochemistry, whereas one oxygen atom of the third aspartate, Asp-355, bridges Mg_C^{2+} and Mg_A^{2+} with *syn*,*anti*-coordination stereochemistry. The NSE/DTE motif appears as Asp-496, Thr-500, and Glu-504, and these residues chelate Mg_B^{2+} on the opposite side of the active site cleft. One oxygen of the β-phosphate of the substrate analogue and product bridges Mg_A^{2+} and Mg_C^{2+}, and a second oxygen coordinates to Mg_B^{2+}. One oxygen atom of the α-phosphate coordinates to Mg_C^{2+} and a second oxygen coordinates to Mg_B^{2+} such that two six-membered ring chelate interactions are formed between the diphosphate group, Mg_B^{2+}, and Mg_C^{2+} (Figure 14).

Intriguingly, the constellation of diphosphate−metal interactions observed in (+)-bornyl diphosphate synthase is identical to that observed in diphosphate complexes with *E. coli* farnesyl diphosphate synthase (Figure 6).[22] Accordingly, it is tempting to speculate that this diphosphate molecular recognition scheme corresponds to the universal molecular recognition of prenyl diphosphates by class I terpenoid cyclases. However, the binding of the trinuclear magnesium cluster to the sesquiterpene cyclase trichodiene synthase is slightly divergent, in that only one residue in the aspartate-rich segment participates in metal binding (section 3.1); metal binding to epi-aristolochene synthase exhibits additional differences (section 3.2). Possibly, the divergence of metal binding and prenyl diphosphate recognition by the trinuclear

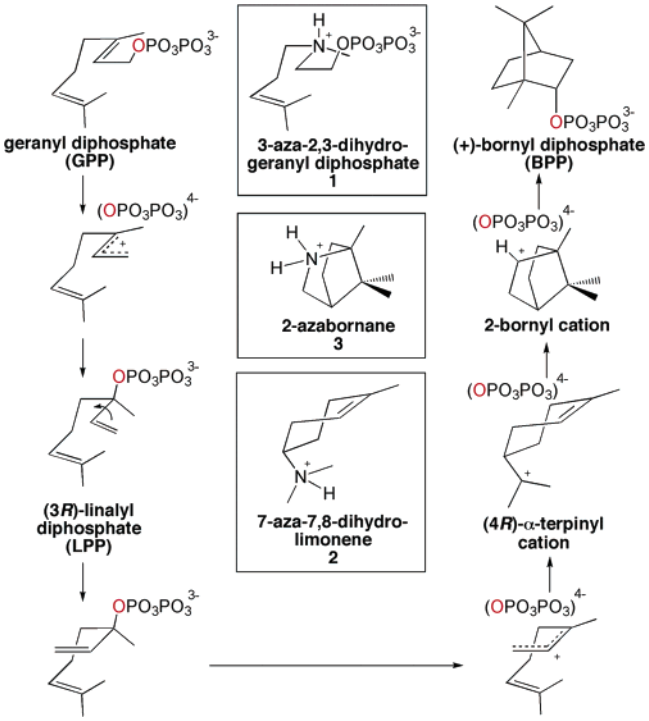


Figure 12. The cyclization of geranyl diphosphate to form (+)-bornyl diphosphate proceeds through intermediate (3*R*)-linalyl diphosphate and the (4*R*)-α-terpinyl and 2-bornyl carbocation intermediates. The diphosphate ester oxygen of the substrate (red) is retained in the diphosphate ester of the product. Aza analogues of carbocation intermediates are illustrated in boxes. Reprinted with permission from ref 43. Copyright 2002 National Academy of Sciences, U.S.A.

magnesium cluster reflects an additional strategy in the evolution of diverse terpenoid cyclases.

Comparisons of experimentally determined structures of complexes between (+)-bornyl diphosphate synthase and aza analogues of carbocation intermediates (compounds **1**−**3** in Figure 12), as well as product, indicate that the hydrophobic active site cleft serves as a template to chaperone reactive substrate and intermediate conformations through an intricate cyclization cascade while the diphosphate group is rigidly bound.[43] A variety of aliphatic and aromatic residues lining the hydrophobic active site cavity, including Trp-323, Ile-344, Val-452, and Phe-578, ensure the proper binding conformation of the substrate and carbocation intermediates. Notably, a single water molecule, water no. 110, remains trapped in the active site cavity in all enzyme−analogue and enzyme−product complexes (Figure 15). Water no. 110 makes hydrogen bond interactions with diphosphate or the PP_i anion, the backbone carbonyl of Ser-451, and the side chain of Tyr-426, so this water molecule appears to be firmly anchored in the active site, sufficiently so that it cannot easily react with carbocation intermediates to prematurely quench the cyclization cascade. Possibly, water no. 110 could serve as a diphosphate-assisted general base to account for the generation of cyclic olefin side products such as pinenes, camphene, and limonene.[36] That the reaction of noncyclizable analogues of geranyl diphosphate generates significant amounts of acyclic terpenol products may reflect the presence of a trapped water molecule in the enzyme active site.[45] It is notable that a trapped water molecule can contribute to the contour of the active site template of a terpenoid cyclase, and it is also notable that such a water molecule could play an essential role in catalysis.

For any cyclization reaction involving geranyl diphosphate, it is clear that diphosphate ionization must be followed by isomerization to a cisoid linalyl intermediate to facilitate cyclization to form the α-terpinyl cation (the trans C2−C3 double bond of geranyl diphosphate would otherwise prevent formation of the six-membered ring) (Figure 12). Following metal-activated diphosphate departure, isomerization to (3*R*)-linalyl diphosphate allows rotation about the C2−C3 bond to the cisoid conformer, which re-ionizes to facilitate S_N' cyclization by C1−C6 ring closure to form the (4*R*)-α-terpinyl cation−PP_i ion pair (Figure 16). Cation−π interactions with Phe-578 and Trp-323 appear to stabilize the C8 carbocation. Subsequent anti-Markovnikov cyclization yields the secondary 2-bornyl cation, which is neutralized by stereospecific C−O bond formation with PP_i on the endo face to yield (+)-bornyl diphosphate.

The binding orientations and conformations of substrate analogue 3-aza-2,3-dihydrogeranyl diphosphate and (4*R*)-α-terpinyl cation analogue 7-aza-7,8-dihydrolimonene do not properly mimic the productive conformations required for catalysis, possibly due in part to overwhelmingly favorable electrostatic and hydrogen bond interactions involving the positively charged aza nitrogen atoms. However, the binding conformation of 2-azabornane in the ternary complex with PP_i nearly matches that of the product (+)-bornyl diphosphate. Taken together, these results seem to imply that the more product-like the carbocation analogue is, the more likely it is to be bound in a catalytically productive conformation. This hypothesis is also consistent with the anomalous binding modes of aza-analogues complexed with trichodiene synthase, discussed in section 3.1.

3. Sesquiterpene Cyclases

Sesquiterpene cyclases catalyze the metal-dependent cyclization of farnesyl diphosphate in the generation of more than 300 known monocyclic, bicyclic, and tricyclic products.[1,2,7,8] Bacterial and fungal enzymes such as pentalenene synthase from *Streptomyces* UC5319 and aristolochene synthase from *Penicillium roqueforti* are usually active as ~40 kD single-domain monomers, but there are also examples of homodimeric enzymes such as trichodiene synthase from *Fusarium sporotrichioides*. As for plant monoterpene cyclases, plant sesquiterpene cyclases such as epi-aristolochene synthase from *Nicotiana tabacum* contain an additional N-terminal domain exhibiting a fold most closely related to that of the glycosyl hydrolases and weakly homologous to the class II terpenoid cyclase fold.[6] Although there is no specific catalytic chemistry established for the N-terminal domain of a plant sesquiterpene cyclase, it nevertheless plays a role in capping the active site of the C-terminal class I cyclase domain. There is generally very low amino acid sequence identity among sesquiterpene cyclases from bacteria, fungi, and plants, but these enzymes adopt the characteristic class I terpenoid cyclase fold and contain the signature aspartate-rich and NSE/DTE motifs implicated in binding the trinuclear magnesium cluster required for catalysis. In the remainder of this section, structure−mechanism relationships in four sesquiterpene cyclases are reviewed that provide interesting and important comparisons with the monoterpene cyclase (+)-bornyl diphosphate synthase discussed in the previous section.

3.1. Trichodiene Synthase

Trichodiene synthase is a sesquiterpene cyclase that catalyzes the first committed step in the biosynthesis of

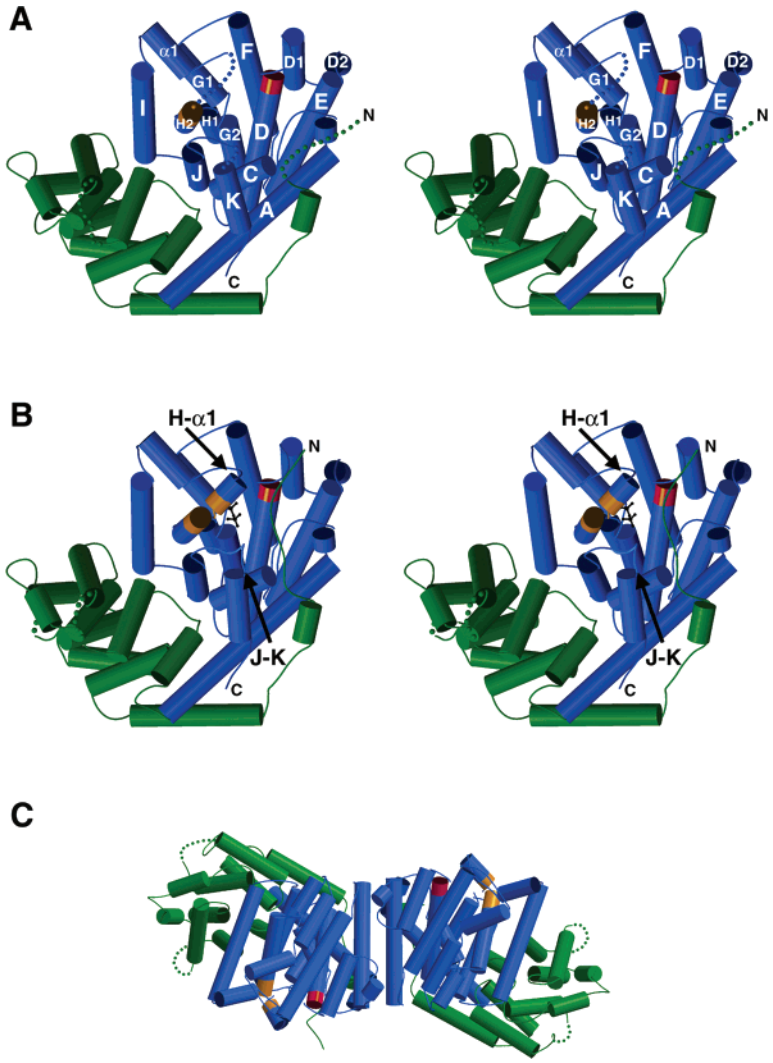


Figure 13. (+)-Bornyl diphosphate synthase. Panel a shows the unliganded monomer, oriented such that the viewer is looking directly into the active site in the C-terminal domain (blue). Helical segments are labeled using the convention established for farnesyl diphosphate synthase in Figure 3.[16] The aspartate-rich motif is red, and the NSE/DTE motif is orange: both motifs are involved in binding the trinuclear magnesium cluster required for catalysis. The N-terminal domain is green. Disordered segments (Glu-50−Ala-63, Glu-228−Ile-233, Thr-500−Asp-509, and Gly-579−Ser-583) are indicated by dotted lines. Panel b shows the monomer complex with the PP_i anion (black), oriented the same as the structure in panel a. The PP_i anion, as well as the diphosphate groups of substrate analogue 3-aza-2,3-dihydrogeranyl diphosphate or product (+)-bornyl diphosphate, trigger conformational changes that stabilize previously disordered polypeptide segments that include most of the N-terminus (Ile-54−Ala-63), the C-terminus of helix H and the H−α1 loop (Thr-500−Asp-509), and part of the J−K loop (Gly-579−Ser-583). These ordered segments help cap the active site. Panel c shows the homodimer with the active sites of each monomer oriented in antiparallel fashion. The dimer interface is formed by the C-terminus of helix A and helices D1, D2, and E. Reprinted with permission from ref 43. Copyright 2002 National Academy of Sciences, U.S.A.

nearly 100 different trichothecene mycotoxins such as T-2 toxin and vomitoxin, which are frequent contaminants of grain-based agricultural products.[46,47] The enzyme purified from *Fusarium sporotrichioides* is a dimer of 45 kD subunits as determined by SDS−PAGE,[48] and the enzyme has been cloned,[49] expressed,[50] and highly overexpressed[51] in *Escherichia coli*. Trichodiene synthase is the best-studied sesquiterpene cyclase in terms of mechanistic enzymology and structural biology, and the mechanism initially proposed for trichodiene formation based on studies with the first enzyme isolated from the apple mold fungus, *Trichothecium roseum*, remains generally accepted (Figure 17).[52]

The metal-dependent ionization of farnesyl diphosphate **1** (Figure 1) is immediately followed by recapture of the pyrophosphate leaving group to form (3*R*)-nerolidyl diphosphate **2**. This isomerization step is required to circumvent the geometric barrier, the trans C2−C3 double bond, that would otherwise thwart the subsequent C1−C6 ring closure reaction. (3*R*)-Nerolidyl diphosphate is conclusively demonstrated to be a catalytic intermediate in enzymological studies with isotopically labelled farnesyl diphosphate and nerolidyl diphosphate.[53,54] Additionally, experiments with the modified substrates (7*S*)-*trans*-6,7-dihydrofarnesyl diphosphate and (7*R*)-*trans*-6,7-dihydrofarnesyl diphosphate yield product arrays consistent with the formation of the corresponding 6,7-dihydronerolidyl diphosphate intermediate, which is unable to cyclize further due to the absence of the C6−C7 double bond.[55] Since farnesyl diphosphate ionization is the rate-limiting chemical step of catalysis,[56] alternative product arrays derived from modified substrates provide a powerful demonstration of the existence of this otherwise undetectable intermediate.[55]

Rotation about the C2−C3 bond of (3*R*)-nerolidyl diphosphate **2** from a transoid to a cisoid conformation facilitates S_N' cyclization by C1−C6 closure with PP_i departure to form the corresponding bisabolyl carbocation **4** (Figure 17). The

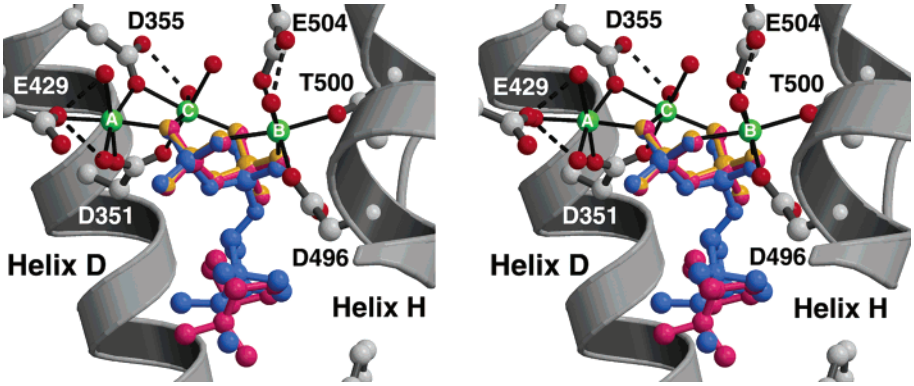


Figure 14. Molecular recognition of diphosphate or the PP_i anion by (+)-bornyl diphosphate synthase is governed by numerous metal coordination and hydrogen bond interactions. The diphosphate position and orientation is nearly invariant in enzyme complexes with PP_i (yellow), 3-aza-2,3-dihydrogeranyl diphosphate (not shown), 7-aza-7,8-dihydrolimonene (not shown), 2-azabornane−PP_i (magenta), or (+)-bornyl diphosphate (blue). This rigid molecular recognition is consistent with the observation that the same oxygen atom comprises the prenyl diphosphate ester linkages of both substrate and product.[37,38] Reprinted with permission from ref 43. Copyright 2002 National Academy of Sciences, U.S.A.

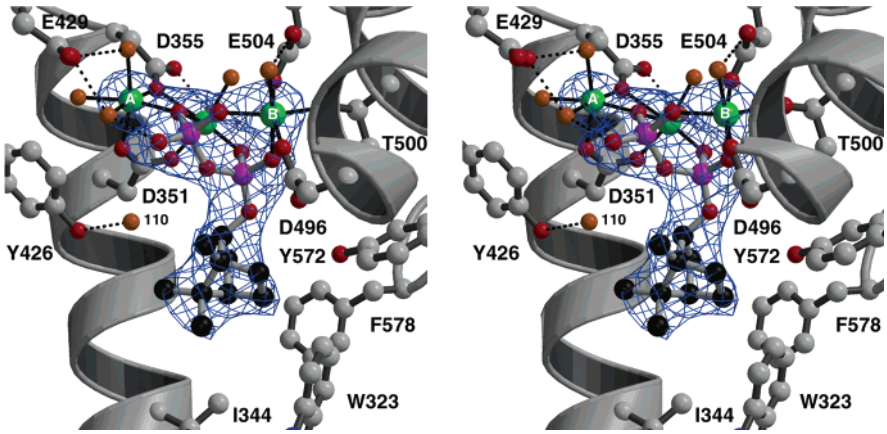


Figure 15. Simulated annealing omit electron density map of (+)-bornyl diphosphate synthase complexed with product (+)-bornyl diphosphate (contoured at 4σ). Metal coordination and hydrogen bond interactions are indicated by solid and dashed black lines, respectively. Note that water molecule no. 110 remains trapped in the active site cavity along with the cyclization product. Reprinted with permission from ref 43. Copyright 2002 National Academy of Sciences, U.S.A.

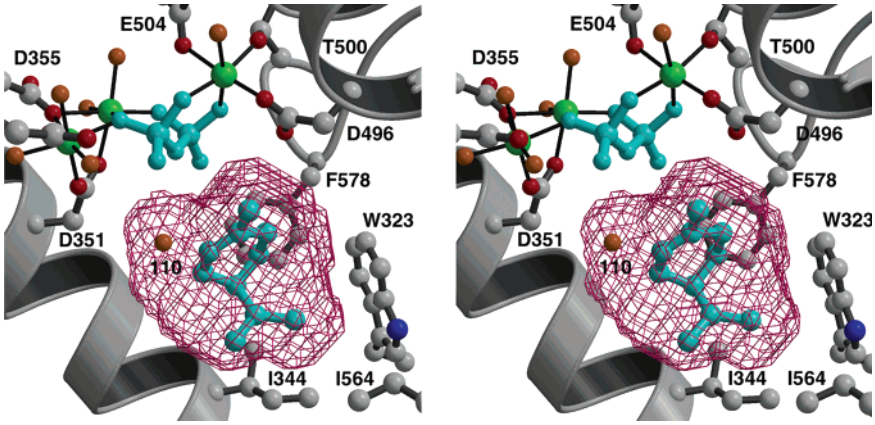


Figure 16. Model of the (4*R*)-α-terpinyl cation bound to the active site of (+)-bornyl diphosphate synthase based on the binding mode of 2-azabornane−PP_i and (+)-bornyl diphosphate. The tertiary carbocation may be stabilized by cation−π interactions with Trp-323 and Phe-578. Reprinted with permission from ref 43. Copyright 2002 National Academy of Sciences, U.S.A.

cationic analogue of bisabolyl carbocation **4**, (4*R*)-7-azabisabolene (Figure 18), exhibits strong synergistic inhibition with PP_i (K_i = 0.51 ± 0.07 μM), but since (4*S*)-7-azabisabolene exhibits comparable behavior (K_i = 0.47 ± 0.07 μM), it is possible that stereochemical discrimination is weak at the corresponding step of catalysis.[57] In fact, the X-ray crystal structures of the Asp-100 → Glu, Tyr-305 → Phe, and Arg-304 → Lys variants of trichodiene synthase complexed with PP_i and azabisabolene inhibitors reveal multiple or disordered azabisabolene binding conformations, some of which appear to be driven by favorable electrostatic interactions between the negatively charged PP_i anion and the positively charged alkylammonium moiety of the azabisabolene inhibitor (Figure 19).[58,59] Aromatic residues in the active

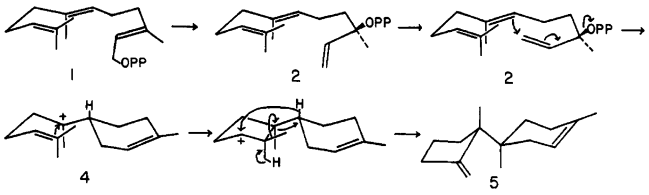


Figure 17. Cyclization of farnesyl diphosphate **1** by trichodiene synthase requires the ionization−isomerization−reionization sequence via intermediate (3*R*)-nerolidyl diphosphate **2** to facilitate closure of the six-membered ring forming the tertiary bisabolyl carbocation intermediate **4**. This is analogous to the ionization−isomerization−reionization sequence encountered in the mechanism of (+)-bornyl diphosphate synthase in Figure 12. Reprinted with permission from ref 52. Copyright 1981 American Chemical Society.

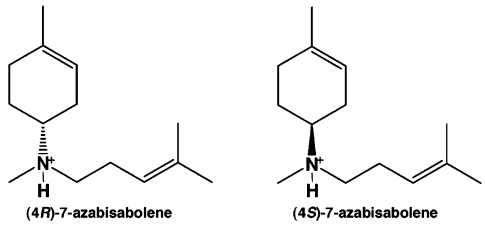


Figure 18. (4*R*)-7-Azabisabolene is a cationic analogue of the tertiary bisabolyl carbocation in the trichodiene synthase mechanism. Its enantiomer, (4*S*)-7-azabisabolene, binds with essentially equal affinity in synergy with PP_i binding.[57]

site also make electrostatic interactions with these inhibitors, and modeling studies suggest that Tyr-93 may stabilize the bisabolyl intermediate through cation−π interactions.[60]

Notably, the thermodynamically favorable binding modes observed for azabisabolene inhibitors do not mimic the catalytically productive orientation and conformation expected for the actual bisabolyl carbocation intermediate. Presuming that these nonproductive orientations and conformations are accessible during catalysis, it would appear that the lifetime of the bisabolyl carbocation is too short to allow it to adopt comparable orientations and conformations. The same conclusion can be drawn with regard to the anomalous binding modes observed for certain aza analogues in complex with (+)-bornyl diphosphate synthase (section 2). These binding phenomena suggest that the formation of transient carbocation intermediates in terpene cyclization reactions is under kinetic rather than thermodynamic control.[58]

Further cyclization of the bisabolyl carbocation followed by a 1,4-hydride shift, tandem 1,2-methyl migrations, and final deprotonation completes the biosynthesis of trichodiene **5** (Figure 17). An enzyme-bound base may mediate the final deprotonation step. Based on crystal structure analysis, Rynkiewicz and colleagues consider either the carboxylate group of Asp-100 or the PP_i anion as possible proton acceptors due to their active site locations.[60] Since Asp-100 is a ligand to Mg^{2+}_A and Mg^{2+}_C, this residue is less likely to serve as a proton acceptor unless significant active site structural changes accompany the final step of catalysis. Thus, the PP_i anion itself is the more likely proton acceptor, and it is also considered as a possible proton acceptor in the generation of certain aberrant products by D100E trichodiene synthase.[61]

The X-ray crystal structure of wild-type trichodiene synthase reveals that each subunit of the dimer adopts the class I terpenoid cyclase fold (Figure 20).[60] The quaternary structure is similar to that of the monoterpene cyclase (+)-bornyl diphosphate synthase (Figure 13) in that the two active sites of the trichodiene synthase dimer are oriented in antiparallel fashion. However, there are differences with regard to the specific helices that form the dimerization interface. It is interesting that the dimerization modes of these two cyclases further contrast with that of farnesyl diphosphate synthase, in which active sites are oriented in parallel fashion (Figure 3).

The X-ray crystal structure of trichodiene synthase complexed with $Mg^{2+}{}_3$−PP_i provides critical details regarding the molecular recognition of PP_i.[60] The structure of this complex reveals significant PP_i-induced structural changes involving the N-terminus, helices 1, D, H, J, K, and L, and loops 1−A, D−D1, H−α-1, J−K, and K−L (Figure 21). These conformational changes are also likely to be triggered by the substrate diphosphate group during catalysis since they serve to cap the active site cleft and sequester it from bulk solvent.

The binding of the trinuclear magnesium cluster in the trichodiene synthase active site (Figure 21) is quite similar, but not identical, to that observed in the active site of (+)-bornyl diphosphate synthase (Figure 14). In the $Mg^{2+}{}_3$−PP_i

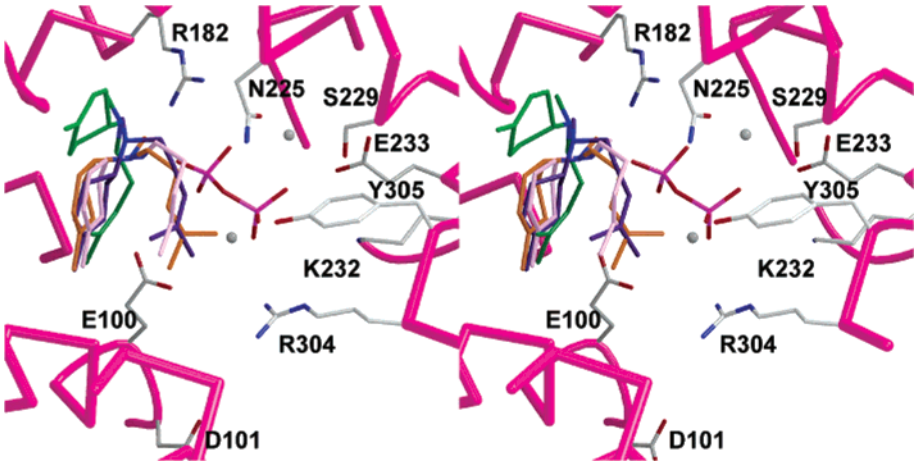


Figure 19. Variable binding orientations and conformations observed for (4*R*)-7-azabisabolene (green, violet) and (4*S*)-7-azabisabolene (orange, pink) in ternary complexes with PP_i and the Asp-100 → Glu and Tyr-305 → Phe variants of trichodiene synthase. Although azabisabolene inhibitors mimic the tertiary bisabolyl carbocation intermediate, these inhibitors are thought to bind with nonproductive orientations and conformations. Reprinted with permission from ref 58. Copyright 2005 American Chemical Society.

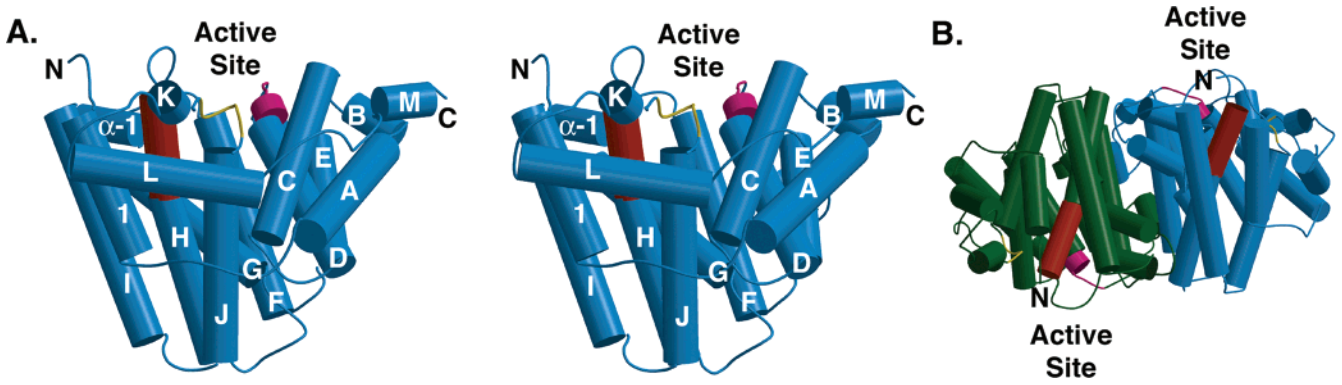


Figure 20. Panel a shows a stereoplot of the trichodiene synthase monomer. The aspartate-rich motif beginning with Asp-100 on helix D is magenta, and the NSE/DTE motif beginning with Asn-225 on helix H is orange. A conserved basic motif beginning with Arg-303 in the J−K loop is yellow. Helices are labeled according to the convention established for farnesyl diphosphate synthase.[16] Panel b shows the structure of the trichodiene synthase dimer. Reprinted with permission from ref 60. Copyright 2001 National Academy of Sciences, U.S.A.

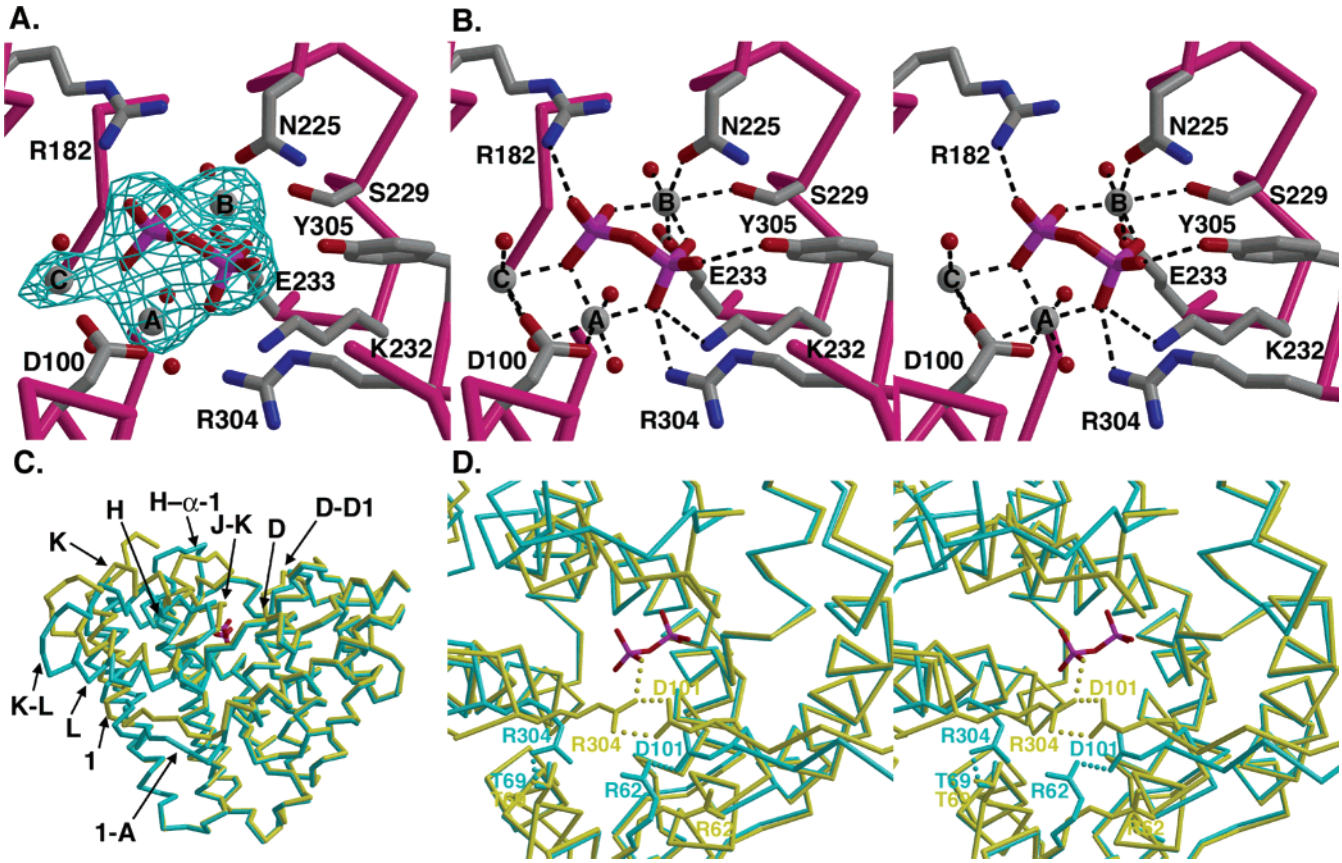


Figure 21. Panel a shows the simulated annealing omit map of the trichodiene synthase−$Mg^{2+}{}_3$−PP_i complex contoured at 5σ. Refined atomic coordinates for PP_i, Mg^{2+} ions (gray spheres), and metal-coordinated water molecules (red spheres) are superimposed. Panel b shows the hydrogen bond and metal coordination interactions in the trichodiene synthase−$Mg^{2+}{}_3$−PP_i complex viewed from the base of the active site. In panel c, superposition of unliganded trichodiene synthase (cyan) and the trichodiene synthase−$Mg^{2+}{}_3$−PP_i complex (yellow) reveals structural changes triggered in indicated segments that cap the active site cleft. The PP_i anion is magenta. In panel d, a view looking directly into the active sites of unliganded trichodiene synthase (cyan) and the trichodiene synthase−$Mg^{2+}{}_3$−PP_i complex (yellow) reveals that Asp-101 (the second residue in the conserved aspartate-rich motif) and Arg-304 (the second arginine in the conserved basic motif) form a double hydrogen bonded salt link that stabilizes the ligand-bound conformation. Because Arg-304 also donates a hydrogen bond to PP_i, this interaction reveals how the diphosphate moiety of the farnesyl diphosphate substrate likely triggers a conformational change to sequester the active site from bulk solvent. Reprinted with permission from ref 60. Copyright 2001 National Academy of Sciences, U.S.A.

complex with trichodiene synthase, the carboxylate group of Asp-100 (the first residue in the aspartate-rich motif) coordinates to Mg_A^{2+} and Mg_C^{2+} with *syn*,*syn*-bidentate geometry. The corresponding aspartate also coordinates to Mg_A^{2+} and Mg_C^{2+} in farnesyl diphosphate synthase (Figures 5 and 6),[20,22] (+)-bornyl diphosphate synthase (Figure 14),[43] and epi-aristolochene synthase (section 3.2). However, in contrast with these terpene synthases, no other aspartate in the aspartate-rich segment of trichodiene synthase coordinates to a metal ion. The NSE/DTE motif of trichodiene synthase (Asn-225, Ser-229, and Glu-233) chelates the Mg_B^{2+} ion (Figure 21) in a fashion similar to that observed in (+)-bornyl diphosphate synthase (Figure 14) and epi-aristolochene synthase (section 3.2).

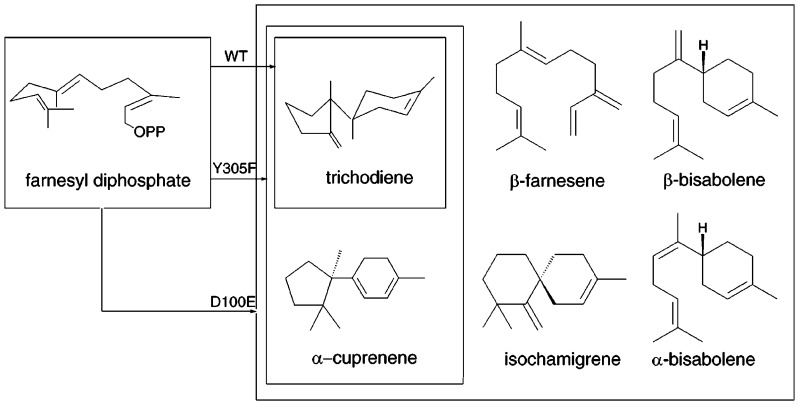


Figure 22. Amino acid substitutions for residues that coordinate to catalytic metal ions (Asp-100 → Glu) or residues that hydrogen bond with PP_i (Tyr-305 → Phe) result in substantially broader product arrays reflecting incomplete cyclization or cyclization to alternative bicyclic products.[62−64] Reprinted with permission from ref 58. Copyright 2005 American Chemical Society.

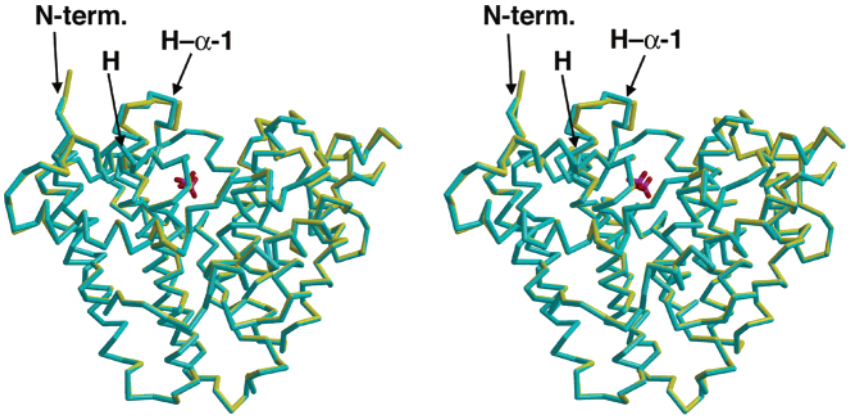


Figure 23. Structure of Asp-100 → Glu trichodiene synthase unliganded (cyan) and complexed with $PP_i{-}Mg^{2+}{}_2$ (yellow). PP_i is red. The PP_i-triggered active site closure is significantly attenuated in comparison with that observed for the wild-type enzyme in Figure 21c. Reprinted with permission from ref 61. Copyright 2002 American Chemical Society.

Divergent metal binding interactions among the terpenoid cyclases implicate subtle differences in $Mg^{2+}{}_3$−substrate diphosphate interactions as a possible strategy for the evolution of product diversity. Consider, for example, site-specific variants of trichodiene synthase in which amino acid substitutions are made in the aspartate-rich motif. The Asp-100 → Glu and Asp-101 → Glu variants exhibit modest differences in K_M and k_{cat} values, and the Asp-104 → Glu variant exhibits wild-type kinetic parameters.[62] However, all three variants generate aberrant cyclization products indicative of alternative substrate orientations and conformations in the enzyme active site. Thus, amino acid substitutions for a residue that coordinates to a catalytically obligatory magnesium ion, even with another residue capable of metal coordination (Asp-100 → Glu), result in the significantly broadened product array shown in Figure 22. The crystal structure determination of Asp-100 → Glu trichodiene synthase complexed with PP_i reveals an incomplete metal cluster with only Mg_A^{2+} and Mg_B^{2+} ions bound; moreover, Glu-100 makes a lengthy 3.4 Å interaction with Mg_A^{2+}, which is too long to be considered inner-sphere coordination.[61] The incomplete metal cluster results in incomplete diphosphate-triggered active site closure (Figure 23). The enclosed active site volume is 12% greater than that of the closed conformation of the wild-type enzyme, and the compromised active site template allows for more conformational degrees of freedom for the flexible substrate, resulting in the formation of alternative cyclization products. The proportion of aberrant products increases when Mn^{2+} is substituted for Mg^{2+}, so it is clear that the metal coordination polyhedra play a crucial role in stabilizing the proper active site contour required for productive catalysis.[62]

Amino acid substitutions for residues that hydrogen bond with PP_i, such as Tyr-305 (Figure 21b), similarly yield diverse product arrays. The Tyr-305 → Phe substitution has little effect on k_{cat} and a minor effect on K_M, yet this variant generates 25% α-cuprenene and 75% trichodiene (Figure 22).[63,64] The X-ray crystal structure of this variant complexed with $Mg^{2+}{}_3{-}PP_i$ and (4*R*)-7-azabisabolene reveals a slightly larger enclosed active site cavity that facilitates alternative product formation.[58]

Diverse product arrays result from amino acid substitutions for Asp-101 and Arg-304, which break hydrogen bond interactions in the unliganded state of the enzyme and form a double hydrogen bonded salt link in $Mg^{2+}{}_3{-}PP_i$ complexes that is likely formed in the enzyme−substrate complex as well (Arg-304 also donates a hydrogen bond to PP_i, see Figure 21d).[60] The Asp-101 → Glu substitution results in a 5-fold decrease in k_{cat}/K_M and the generation of five aberrant products in addition to trichodiene,[62] and the Arg-304 → Lys substitution results in a 5000-fold decrease in k_{cat}/K_M, as well as the generation of at least two aberrant products in addition to trichodiene.[63] The X-ray crystal structure of Arg-304 → Lys trichodiene synthase complexed with $Mg^{2+}{}_3{-}PP_i$ and (4*R*)-7-azabisabolene shows that the altered hydrogen bond array shifts the PP_i anion ~0.7 Å and weakens metal coordination interactions.[59] These structural changes result in an enclosed active site cavity volume 17% greater than

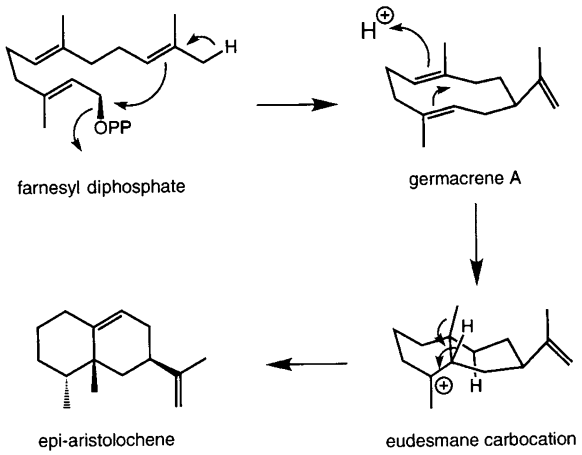


Figure 24. Proposed cyclization mechanism catalyzed by epi-aristolochene synthase. Adapted with permission from ref 71. Copyright 2000 American Chemical Society.

that of the wild-type enzyme, thereby allowing for the generation of alternative cyclization products. Weakened $Mg^{2+}{}_3-PP_i$ interactions implicate weakened $Mg^{2+}{}_3-$substrate diphosphate interactions in catalysis, suggesting that weaker activation of the diphosphate leaving group accounts for the dramatic activity loss measured for this variant. Clearly, even subtle perturbations in the trinuclear magnesium cluster or the constellation of hydrogen bond interactions responsible for substrate diphosphate recognition can lead to the formation of alternative cyclization products. It is conceivable that this strategy similarly underlies the evolution of new terpenoid cyclases in nature.

3.2. Epi-Aristolochene Synthase

Epi-aristolochene synthase from *N. tabacum* (tobacco) catalyzes the metal-dependent cyclization of farnesyl diphosphate to form the bicyclic hydrocarbon epi-aristolochene (Figure 24), which comprises the first committed step in the biosynthesis of the antifungal phytoalexin capsidiol.[65−67] The X-ray crystal structure of this enzyme was one of the first of a sesquiterpene cyclase[68] and revealed two α-helical domains (Figure 25): a C-terminal catalytic domain exhibiting the class I terpenoid synthase fold analogous to that first observed for avian farnesyl diphosphate synthase,[16] and an N-terminal domain of unknown function exhibiting structural homology to glucoamylase[69] and endoglucanase CelD,[70] subsequently designated the class II terpenoid synthase fold.[6] The overall structure of this plant cyclase is homologous to that of (+)-bornyl diphosphate synthase (Figure 13).[43]

The enzyme mechanism proceeds by initial ionization of the substrate diphosphate group, which is triggered by metal coordination interactions and hydrogen bond interactions. X-ray crystal structures of epi-aristolochene synthase complexes with farnesyl hydroxyphosphonate and trifluorofarnesyl diphosphate reveal the binding of three Mg^{2+} ions: Mg^{2+}_A and Mg^{2+}_C are liganded by Asp-301 and Asp-305 in the signature "aspartate-rich" sequence, and Mg^{2+}_B is liganded by Asp-444, Thr-448, and Glu-452 in the NSE/DTE motif (Figure 26). The phosphonate and diphosphate groups of the substrate analogues make coordination interactions with Mg^{2+}_A and Mg^{2+}_C and the diphosphate group of trifluorofarnesyl diphosphate interacts with Mg^{2+}_B through a metal-bound water molecule (Figure 26). This contrasts with diphosphate/PP_i interactions with Mg^{2+}_B in (+)-bornyl diphosphate synthase (Figure 14)[43] and trichodiene synthase (Figure 21b),[60] where diphosphate/PP_i makes inner-sphere coordination interactions with Mg^{2+}_B. The divergence of metal coordination polyhedra among these terpenoid cyclases could reflect an evolutionary strategy to achieve divergent cyclization chemistry in each enzyme active site. That single amino acid substitutions of metal-binding residues in the aspartate-rich or NSE/DTE motifs diversify the terpenoid cyclase product array, for example, as discussed for trichodiene synthase in section 3.1, is consistent with this notion.

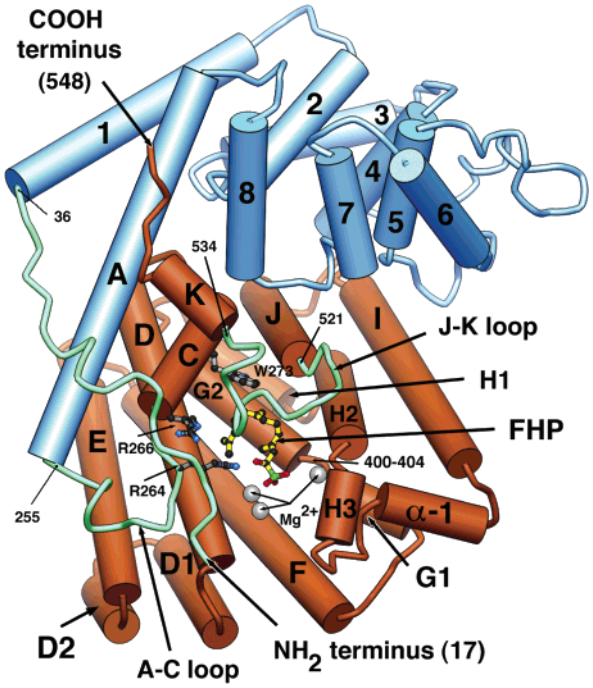


Figure 25. Structure of epi-aristolochene synthase complexed with farnesyl hydroxyphosphonate (FHP). The C-terminal class I cyclase domain (orange) contains the active site for the cyclization reaction. The N-terminal domain (blue) has no known catalytic function but nevertheless adopts a class II terpenoid cyclase fold. The binding of substrate analogues in the C-terminal domain is accompanied by the ordering of 36 residues in the N-terminal segment, and the A−C and J−K loops. These structural changes help to cap the active site cavity and are likely triggered by the diphosphate group of substrate farnesyl diphosphate. Reprinted with permission from *Science* (http://www.aaas.org), ref 68. Copyright 1997 American Association for the Advancement of Science.

In the epi-aristolochene synthase mechanism, substrate ionization invites attack of the C10−C11 π bond at C1 with proton elimination at C12 to yield the stable, neutral intermediate, germacrene A (Figure 24). Enzymological studies with the Tyr-520 → Phe variant of epi-aristolochene synthase demonstrate that germacrene A is formed as the sole cyclization product.[71] Although Tyr-520 was initially proposed as the general acid residue responsible for protonating germacrene A to form the eudesmane cation[68] (Figure 24) and its deletion results in the premature termination of the cyclization cascade with germacrene A formation,[71] Starks and colleagues[68] note that Asp-444 and Asp-525 make hydrogen bond interactions with Tyr-520 and contemplate the possibility of Asp-444 as a general acid[71] by analogy with germacrene C synthase, in which an aspartic acid residue is similarly implicated.[72] This possibility is especially interesting in view of the fact that an aspartic acid is implicated as a proton donor in protonation-dependent cyclases such as abietadiene synthase (section 4.1), squalene−

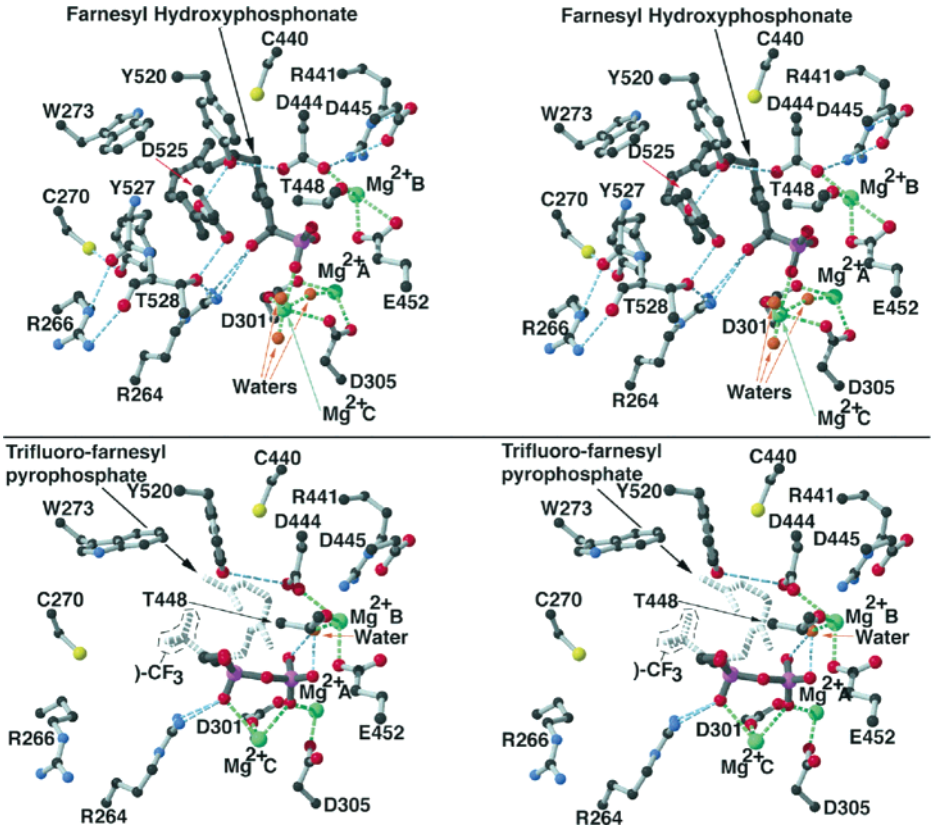


Figure 26. Binding of farnesyl hydroxyphosphonate (top) and trifluorofarnesyl diphosphate (bottom) to the active site of epi-aristolochene synthase. Metal coordination and hydrogen bond interactions are designated by green and blue dashed lines, respectively. Inhibitors are indicated with darker bonds than surrounding protein residues; the isoprenoid portion of trifluorofarnesyl diphosphate is said to be characterized by weak electron density and is indicated with dashed bonds.[68] Reprinted with permission from *Science* (http://www.aaas.org), ref 68. Copyright 1997 American Association for the Advancement of Science.

hopene cyclase (section 5.1), and oxidosqualene cyclase (section 5.2).

Protonation of germacrene A results in formation of the bicyclic eudesmane cation, which then undergoes tandem 1,2-methyl and 1,2-hydride migrations with final deprotonation to generate epi-aristolochene (Figure 24). Clearly, the hydrophobic contour of the enzyme active site must serve as a high-fidelity template to properly chaperone reactive intermediates leading to epi-aristolochene formation. The active site is lined by numerous aliphatic and aromatic residues, and the side chain of Trp-273 is proposed to stabilize carbocation intermediates through cation−π interactions.[68] Interestingly, chimeric cyclases in which selected subdomains of epi-aristolochene synthase and vetispiradiene synthase from *Hyoscyamus muticus* are swapped generate mixtures of the reaction products of both enzymes, indicating that the active site contour of terpenoid cyclases can be altered to accommodate alternative cyclization pathways emanating from the eudesmane carbocation intermediate common to both cyclases.[73] This work demonstrates the great promise of engineering diversity in terpenoid biosynthesis by design.

3.3. Aristolochene Synthase

Aristolochene synthase catalyzes the Mg^{2+}-dependent cyclization of farnesyl diphosphate to form the bicyclic hydrocarbon (+)-aristolochene **4** (Figure 27), which is a stereochemical isomer of epi-aristolochene discussed in section 3.2. Mechanistic studies using stereospecifically labelled farnesyl diphosphate,[74,75] the anomalous substrate

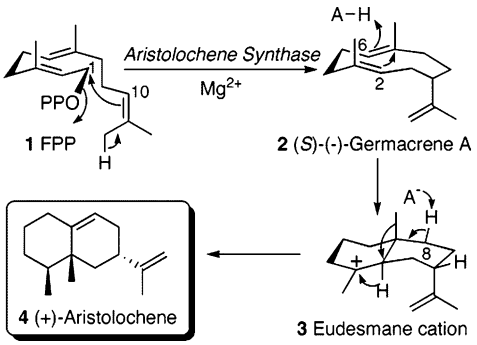


Figure 27. Cyclization mechanism of aristolochene synthase. Mechanistic details are summarized in the text. Reprinted with permission from ref 87. Copyright 2004 American Chemical Society.

dihydrofarnesyl diphosphate,[76] and a mechanism-based inhibitor[77] indicate a cyclization mechanism in which substrate ionization is accompanied by attack of the C10−C11 π bond at C1 with inversion of configuration to form the neutral intermediate **2**, (*S*)-(−)-germacrene A; protonation at C6 and attack of the C2−C3 π bond at the resulting C7 carbocation yields eudesmane cation **3**, which yields (+)-aristolochene **4** following methyl migration, hydride transfer, and deprotonation.

(+)-Aristolochene synthases have been purified from *P. roqueforti*[78] and *Aspergillus terreus*,[31] and both enzymes have been cloned and overexpressed in *E. coli*.[31,79,80] Both enzymes are monomeric and their deduced amino acid sequences share ~60% identity; although neither enzyme shares significant sequence identity with other terpenoid cyclases, the X-ray

Figure 28. Ribbon plot of (+)-aristolochene synthase from *P. roqueforti*.[81] The aspartate-rich motif (red) and the NSE/DTE motif (orange) flank the mouth of the active site.

crystal structure of each enzyme reveals conservation of the characteristic class I terpenoid cyclase fold.[81,82] The ribbon plot of (+)-aristolochene synthase from *P. roqueforti* in Figure 28 clearly shows the α-helical topology of the terpenoid cyclase fold and conservation of the aspartate-rich and NSE/DTE metal-binding motifs.[81] Interestingly, Asp-115, which is the first residue in the aspartate-rich motif, coordinates to a single Sm^{3+} ion in one of the heavy atom derivatives prepared in the X-ray crystal structure determination.[81] This feature is reminiscent of Sm^{3+} binding to each of the two aspartate-rich motifs in farnesyl diphosphate synthase.[16] The Sm^{3+} ion is a reasonable analogue of the biologically relevant Mg^{2+} ion, and each readily binds to carboxylate clusters on protein surfaces, which tend to discriminate more on the basis of ionic charge rather than ionic radius.[83]

The lower region of the active site of aristolochene synthase is quite hydrophobic in nature and its contour is defined by several aromatic and aliphatic amino acid side chains. Molecular modeling of substrate, intermediates, and product in the aristolochene synthase active site suggests that the side chains of Phe-178 and Phe-112 are suitably oriented to stabilize the developing partial positive charge at C1 during the initial cyclization reaction, and the side chain of Trp-333 may stabilize the eudesmane cation through cation−π interactions.[81] Substitution of this residue with the smaller aliphatic side chains of valine or leucine results in catalytically compromised enzyme variants that essentially halt at germacrene A formation, consistent with the loss of a residue critical for the subsequent step of eudesmane cation formation.[84]

The identification of basic residues in the enzyme active site capable of performing the initial deprotonation at C12 or the final deprotonation of the eudesmane cation **3** to yield aristolochene **4** (Figure 27) is ambiguous. Inspection of the aristolochene synthase active site does not reveal any residue that could assist with the first deprotonation step other than the diphosphate leaving group itself.[81] Molecular modeling of the enzyme complexed with the germacrene A or eudesmane cation intermediates suggested that Tyr-92 could serve as a general acid to protonate the germacrene A intermediate and subsequently as a general base to deprotonate the eudesmane cation,[81] but the Tyr-92 → Phe variant

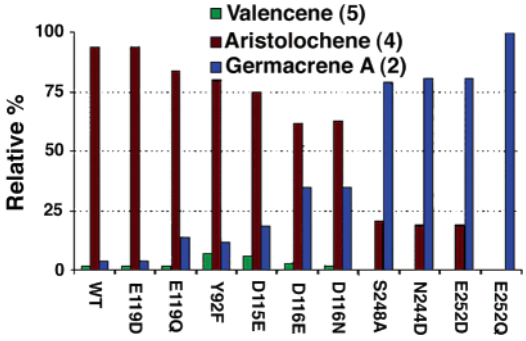


Figure 29. Product distribution observed for site-specific variants of *P. roqueforti* aristolochene synthase. Reprinted with permission from ref 87. Copyright 2004 American Chemical Society.

generates appreciable quantities of (+)-aristolochene.[85−87] Furthermore, proper kinetic analysis of this variant reveals only a minor 2-fold decrease in k_{cat}/K_M, so Tyr-92 is clearly ruled out as a catalytically obligatory general acid/base.[87] Felicetti and Cane advance the intriguing notion that a firmly bound solvent molecule trapped in the active site might perform a general acid/base function.[87] The identification of a single trapped water molecule in the active site of (+)-bornyl diphosphate synthase complexed with the product as well as stable analogues of carbocation intermediates shows that a water molecule can be sequestered in the enzyme active site and, if firmly anchored by hydrogen bond interactions, would not prematurely quench carbocation intermediates in the cyclization cascade.[43] The X-ray crystal structure determinations of aristolochene synthase complexes with substrate and intermediate analogues may illuminate the possible role of solvent in the cyclization mechanism.

It is interesting that wild-type aristolochene synthase from *P. roqueforti* generates small amounts of germacrene A as well as (−)-valencene, a bicyclic regioisomer of (+)-aristolochene (Figure 29),[85,87] whereas wild-type aristolochene synthase from *A. terreus* generates 100% (+)-aristolochene.[87] The product distribution of each enzyme is modulated by amino acid substitutions in the metal-binding aspartate-rich and NSE/DTE motifs, as illustrated for the *P. roqueforti* enzyme in Figure 29.[87] It is clear that perturbations of these motifs, which govern the active site closure mechanism, result in compromised templates for ionization and cyclization, and it is notable that even the wild-type cyclase from *P. roqueforti* does not appear to be perfectly evolved to yield a single cyclization product like the corresponding enzyme from *A. terreus*. Comparison of the X-ray crystal structures of these enzymes may illuminate the structural basis for the more perfect evolution of (+)-aristolochene biosynthesis by the *A. terreus* enzyme.[82]

3.4. Pentalenene Synthase

Pentalenene synthase catalyzes the Mg^{2+}-dependent cyclization of farnesyl diphosphate to form the tricyclic hydrocarbon pentalenene (**2**, Figure 30),[88,89] a biosynthetic precursor of the pentalenolactone family of antibiotics.[90] The cyclization mechanism summarized in Figure 30 was conclusively delineated by Cane and colleagues using stereospecifically labelled farnesyl diphosphate as a substrate, which allowed for the precise dissection of individual steps in the reaction sequence.[88,89,91−93]

Pentalenene synthase from *Streptomyces* UC5319 has been cloned, expressed in *E. coli*,[94] and crystallized,[95] and crystals of the 42-kD recombinant wild-type enzyme yielded one of

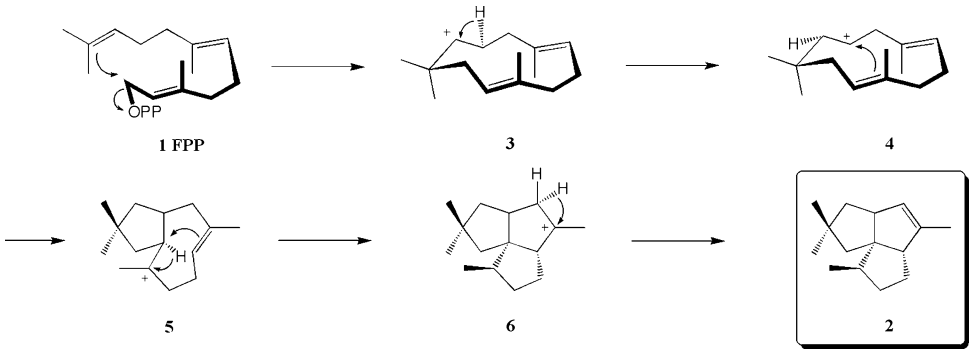


Figure 30. Cyclization mechanism of pentalenene synthase. Briefly, Mg^{2+}-dependent ionization of farnesyl diphosphate (FPP, **1**) is accompanied by attack of the C10−C11 π bond in anti-Markovnikov fashion to yield the cyclic humulyl cation **3**, which isomerizes to **4** through either a deprotonation−reprotonation sequence or a 1,2-hydride shift. A second cyclization step yields the seco-illudyl cation **5**. The final cyclization requires a hydride transfer and stereospecific deprotonation of the H-7*re* proton from intermediate **6** to yield pentalenene **2**. Reprinted with permission from ref 97. Copyright 2002 American Chemical Society.

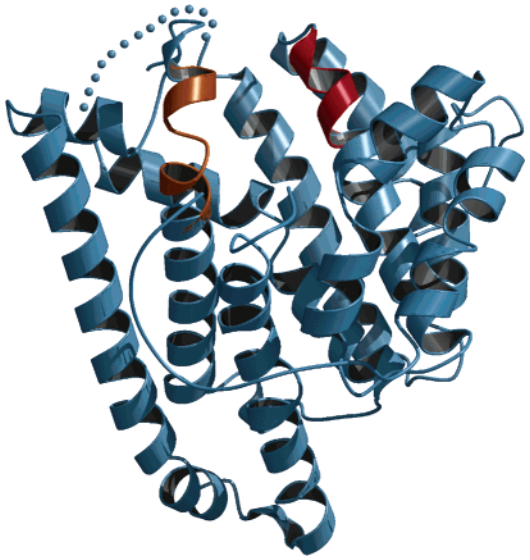


Figure 31. Ribbon plot of pentalenene synthase.[96] The disordered Phe-158−Asp-164 loop (dotted line), the aspartate-rich motif (red), and the NSE/DTE motif (orange) flank the mouth of the active site.

the first class I terpenoid cyclase structures (Figure 31).[96] The lower region of the enzyme active site cleft is predominantly hydrophobic in nature and contains several aromatic and aliphatic residues that define the overall contour of the cavity. The upper region of the active site cavity is more hydrophilic in nature and includes polar side chains required for magnesium binding (the aspartate-rich motif (Asp-80−Asp-84) on helix D, and the NSE/DTE motif (Asn-219−Glu-227) on the opposite wall of the active site on helix H. Although no X-ray crystal structures have been determined to date for pentalenene synthase complexed with Mg^{2+} ions, the corresponding motifs bind trinuclear magnesium clusters in farnesyl diphosphate synthase,[22] (+)-bornyl diphosphate synthase,[43] trichodiene synthase,[60] and epi-aristolochene synthase.[68] Additionally, site-directed mutagenesis experiments with pentalenene synthase confirm the catalytic importance of selected residues in the two metal-binding motifs:[97] in the aspartate-rich motif, glutamate substitutions for Asp-80 and Asp-81 result in severely compromised catalytic efficiency, whereas the substitution of glutamate for Asp-84 yields only minor effects on catalysis; in the NSE/DTE motif, alanine and leucine substitutions for Asn-219 abolish catalysis. Interestingly, Asn-219 → Asp pentalenene synthase retains some catalytic activity, with a 60-fold decrease in k_{cat} and a 55-fold increase in K_M, and also produces 9% of an aberrant cyclization product, β-caryophyllene.[97] As noted by Seemann and colleagues,[97] the first residue in the NSE/DTE motif is aspartate instead of asparagine in the majority of plant monoterpene and sesquiterpene cyclases, and the retention of aspartate or the evolution of asparagine for residue 219 in pentalenene synthase ought to support Mg^{2+} binding.

Upon the initial structure determination of pentalenene synthase,[96] Asn-219 was proposed to govern the substrate binding conformation and stabilize carbocation intermediates, but structural data subsequently acquired from the related cyclases 5-epi-aristolochene synthase,[68] trichodiene synthase,[60] and (+)-bornyl diphosphate synthase[43] instead suggested a conserved metal-binding function for Asn-219, Ser-223, and Glu-227 in the NSE/DTE motif of pentalenene synthase.[97] Also considered in the pentalenene synthase mechanism was His-309, which was proposed to mediate a possible deprotonation−reprotonation sequence in the isomerization of humulyl cations **3** and **4**, as well as the final deprotonation of intermediate **6** shown in Figure 30. However, subsequent site-directed mutagenesis experiments demonstrated that the His-309 → Ala, His-309 → Ser, and His-309 → Phe variants of pentalenene synthase exhibit only minor increases in K_M values and 3−17-fold decreases in k_{cat} values, with up to 22% aberrant cyclization products formed.[98] Clearly, His-309 is not required for cyclization activity, but it is required to maintain the fidelity of pentalenene cyclization. Possibly, slight alterations of the active site contour caused by His-309 substitutions compromise the cyclization template by conferring excessive conformational freedom to reactive carbocation intermediates, thereby allowing formation of alternative sesquiterpene hydrocarbons. That His-309 does not serve as a catalytically obligatory general base for pentalenene biosynthesis points toward a 1,2-hydride transfer instead of a deprotonation−reprotonation sequence to achieve the isomerization of humulyl cations **3** and **4** in Figure 30.[98]

Even so, the penultimate carbocation intermediate **6** in Figure 30 must undergo a final stereospecific deprotonation of H-7*re* to yield pentalenene **2**. Analysis of the native enzyme active site does not point to any alternative basic residues that could mediate this deprotonation step. Noting that this stereospecific step cannot be a spontaneous chemical event and presumably must be mediated by a suitably oriented active site base, Seemann and colleagues[97] suggest three possibilities for a catalytic base: (1) enzyme-bound

water molecule(s), (2) the carbonyl oxygen of the polypeptide backbone, or (3) the PP_i anion coproduct. The pK_a values of the conjugate acid forms of all three species are sufficiently higher (greater than −4) than the pK_a of a tertiary carbocation (−10) to mediate a thermodynamically favorable deprotonation to yield a neutral hydrocarbon product. With regard to an enzyme-bound water molecule as a possible base, it is intriguing to note that a single water molecule is firmly trapped in the active site of (+)-bornyl diphosphate synthase during the monoterpene cyclization cascade,[43] so if a water molecule were similarly trapped in the pentalenene synthase active site during catalysis, it might be able to mediate the final deprotonation step. With regard to PP_i as a possible base, it is intriguing that this moiety is considered as a possible base in the mechanisms of trichodiene synthase[60,61] and farnesyl diphosphate synthase.[22] Clarification of these mechanistic aspects for pentalenene synthase must await the X-ray crystallographic structure determinations of enzyme complexes with analogues of the substrate and carbocation intermediates.

4. Diterpene Cyclases

Diterpene cyclases catalyze the cyclization of the linear 20-carbon substrate geranylgeranyl diphosphate to form a variety of cyclic and polycyclic products. Although no crystal structure of a diterpene cyclase has yet been reported, amino acid sequence similarities among plant terpenoid cyclases provide valuable inferences regarding structure−mechanism relationships. Amino acid sequences of certain plant diterpene cyclases indicate the presence of both class I and class II terpenoid synthase folds containing the signature sequence motifs DDXXD/E and DXDD, respectively (Figure 32).[99,100] Unique to such plant cyclases is a ∼240-residue "insertional" domain at the N-terminus. Studies of abietadiene synthase truncation variants indicate that residues in the insertional domain affect catalysis in the C-terminal domain.[101]

Given the presence of catalytically active class I and class II terpenoid cyclase domains in the family of diterpene cyclases, these enzymes comprise an evolutionary bridge of sorts between the single-domain class I bacterial terpenoid cyclases such as pentalenene synthase and the double-domain class II terpenoid cyclases such as the triterpene cyclases discussed in section 5. In the remainder of this section, structural and mechanistic inferences regarding abietadiene synthase are discussed with regard to the distinct cyclization reactions catalyzed by its class I and class II cyclase domains. A brief comparison with taxadiene synthase completes an overview of the diterpene cyclases.

4.1. Abietadiene Synthase

In response to a surface wound by a bark beetle, the grand fir tree (*Abies grandis*) secretes oleoresin, which is a cocktail of ∼50% volatile monoterpene olefins (turpentine) and ∼50% diterpene resin acids (rosin), plus small amounts of certain sesquiterpenes.[102] The volatile turpentine component evaporates, leaving the hardened rosin to seal the wound. The chemical defense mechanism is further enhanced by insecticidal and fungicidal terpene natural products.[103] (−)-Abietic acid is the principal resin acid component of the wound response in the grand fir. The hydrocarbon precursor of this resin acid is (−)-abieta-7(8),13(14)-diene, or simply abietadiene.

Abietadiene is generated by the cyclization of geranylgeranyl diphosphate[104] in two consecutive reactions catalyzed by a single protein, abietadiene synthase (Figure 33).[105] Protonation at the substrate C14 atom triggers attack by the C10−C11 π bond, which triggers attack by the C6−C7 π bond at C11 and deprotonation at C19 to yield the stable bicyclic intermediate (+)-copalyl diphosphate **2**.[106] Ionization of the allylic diphosphate ester of **2** and anti-S_N' cyclization[107] yields the tricyclic C8-sandaracopimerenyl cation $\mathbf{3a}^+$, which then undergoes intramolecular proton transfer to form the reactive secondary carbocation intermediate $\mathbf{3b}^+$.[106−108] This drives a 1,2-methyl migration to generate the tertiary carbocation $\mathbf{4}^+$, deprotonation of which yields abietadiene **5** or regioisomers levopimaradene **6** and neoabietadiene **7**.[109] Subsequent oxidation of the C18 methyl group of **5** yields (−)-abietic acid.

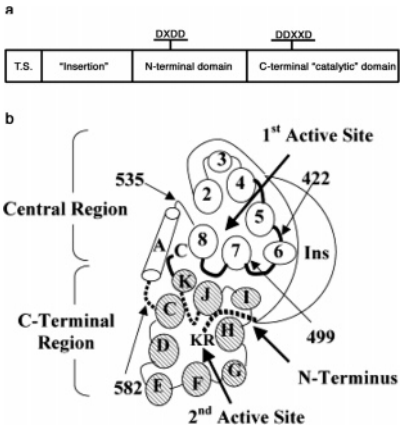


Figure 32. Panel a shows a general scheme of plant diterpene synthase domain structure based on abietadiene synthase as modeled using epi-aristolochene synthase as a template. The ∼9 kD targeting sequence (T.S.) is cleaved in plastids to yield a mature protein consisting of a ∼30 kD "insertional" domain of largely unknown function (although a pair of charged residues, Lys-86 and Arg-87, affect catalysis in the C-terminal domain), a ∼29 kD N-terminal domain containing the signature class II cyclase sequence, and a ∼32 kD domain containing the signature class I cyclase sequence DDXXD that is designated the C-terminal catalytic domain. In abietadiene synthase, the N-terminal and C-terminal domains each catalyze a distinct cyclization reaction. In other plant terpenoid cyclases that share this domain structure, such as taxadiene synthase, catalytic activity is demonstrated only for the C-terminal domain. Reprinted with permission from ref 110. Copyright 2001 American Chemical Society. Panel b shows a topographical projection of pseudomature abietadiene synthase from which the targeting sequence has been cleaved, leaving the basic Lys-86−Arg-87 pair ("KR") to interact with the active site in the C-terminal domain ("2nd Active Site", the site of ionization-induced cyclization). The insertional domain ("Ins") links the basic pair to the N-terminal domain ("1st Active Site", designated here as the "Central Region", the site of protonation-induced cyclization). Reprinted with permission from ref 101. Copyright 2003 American Chemical Society.

Kinetic studies utilizing geranylgeranyl diphosphate and (+)-copalyl diphosphate indicate that the two distinct cyclization reactions are catalyzed by two distinct active sites,[109] making abietadiene synthase a bifunctional enzyme. Site-directed mutagenesis experiments focusing on the DXDD motif in the class II cyclase domain or the DDXXD motif in the class I cyclase domain specifically abolish the geranylgeranyl diphosphate cyclization reaction or the (+)-copalyl diphosphate cyclization reaction, respectively.[110] Additionally, the binding of an aza analogue of geranylgeranyl diphosphate, 15-aza-GGPP,[111] preferentially inhibits the cyclization of geranylgeranyl diphosphate but not the cyclization of (+)-copalyl diphosphate.[110] Free (+)-copalyl diphosphate is detected in steady-state assays, consistent with

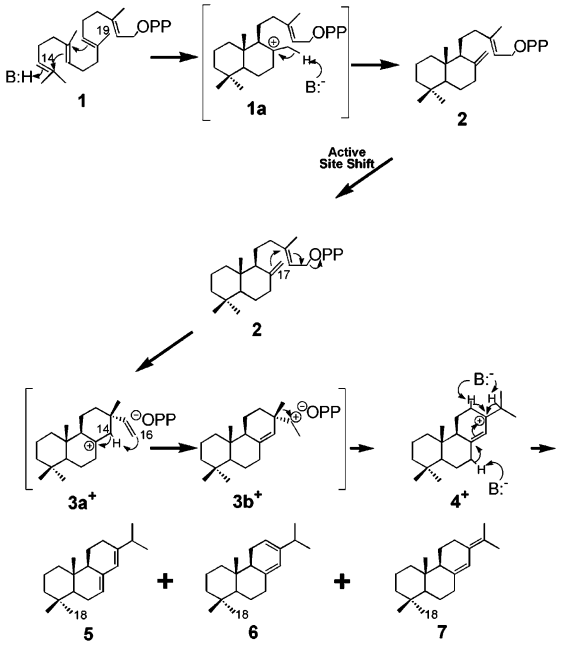


Figure 33. Cyclization of geranylgeranyl diphosphate by abietadiene synthase. The first cyclization reaction, triggered by protonation of geranylgeranyl diphosphate (**1**) and terminated by formation of (+)-copalyl diphosphate (**2**), is catalyzed by the N-terminal class II cyclase domain. Asp-404 is the likely general acid B:H. The second cyclization reaction is catalyzed by the C-terminal class I cyclase domain, which requires that (+)-copalyl diphosphate **2** must shift from one active site to the other. Metal-triggered ionization of **2** initiates the second cyclization cascade, which is terminated by proton elimination to form abietadiene (**5**) or regioisomeric elimination products levopimaradiene (**6**) or neoabietadiene (**7**). Reprinted with permission from ref 101. Copyright 2003 American Chemical Society.

the free diffusion transfer of (+)-copalyl diphosphate out of the active site of the class II cyclase domain and into the active site of the class I cyclase domain, so there is no evidence for a more complex substrate channeling mechanism.[110]

Mutagenesis of amino acid residues in the class II cyclase active site of abietadiene synthase and analysis of a homology model of the enzyme constructed using epi-aristolochene synthase as a template suggests that the central aspartate of the DXDD motif, Asp-404, is the general acid B:H that initiates the first cyclization sequence by protonating geranylgeranyl diphosphate (Figure 2).[110,112] Mutagenesis of residues in the class I cyclase active site indicates that amino acid side chains in the DDXXD motif and the NSE/DTE motif (which appears as Asn-765, Thr-679, and Glu-773), both of which are implicated in binding essential Mg^{2+} ions, are critical for catalysis.[113] Intriguingly, Peters and Croteau find that the distribution of products abietadiene **5**, levopimaradiene **6**, and neoabietadiene **7** (Figure 33) is significantly altered in certain variants, suggesting that the three possibilities for the final deprotonation of tertiary carbocation $\mathbf{4}^+$ (Figure 33) are highly sensitive to slight changes in the orientation of intermediate $\mathbf{4}^+$ in the enzyme active site.[113] This feature is reminiscent of recent structural studies of site-specific variants of trichodiene synthase in which altered product arrays are believed to arise from altered molecular recognition of PP_i.[58,59]

The study of abietadiene synthase truncation variants implicates a pair of charged residues in the N-terminal domain, Lys-86 and Arg-87, in the catalytic activity of the class II cyclase active site at the C-terminal domain.[101] Interestingly, (+)-bornyl diphosphate synthase also contains a pair of charged residues, Arg-55 and Arg-56, in its N-terminal domain, but it does not contain an entire "insertional" domain comparable to that of the plant diterpene cyclases. Arg-56 of (+)-bornyl diphosphate synthase donates a hydrogen bond to the carbonyl of Asp-355 in the aspartate-rich motif in the complex with (+)-bornyl diphosphate or ternary complexes with PP_i and analogues of carbocation intermediates, thereby facilitating the "capping" of the active site by the N-terminal polypeptide.[43] If Arg-55 and Arg-56 of (+)-bornyl diphosphate synthase correspond to Lys-86 and Arg-87 of abietadiene synthase, it is interesting to speculate that these abietadiene synthase residues may likewise interact with the diphosphate group of (+)-copalyl diphosphate, or at least they may be involved in facilitating the contact between the insertional domain and the class I cyclase domain, which could possibly serve to cap the active site cavity.[101] The X-ray crystal structure determination of abietadiene synthase complexed with substrate analogues should help to pinpoint features of substrate recognition and catalysis in both active sites of this novel bifunctional cyclase.

4.2. Taxadiene Synthase

Paclitaxel (Taxol) is a taxane diterpene isolated from the bark of the Pacific yew (*Taxus brevifolia*)[114] that exhibits potent antitumor properties by stabilizing microtubules and arresting the cell cycle.[115−117] Intriguingly, recent reports show that paclitaxel may also be useful in treating neurodegenerative diseases such as Alzheimer's disease by compensating for the loss of function of the microtubule-binding protein tau.[118−120] Taxadiene synthase catalyzes the first committed step of paclitaxel biosynthesis through the metal-dependent cyclization of geranylgeranyl diphosphate to form taxa-4(5),11(12)-diene (Figure 34).

Taxadiene synthase has been isolated and purified from the Pacific yew,[121] and the full-length, 862-residue enzyme (98 kD) has been cloned and expressed in *Escherichia coli*.[122] The amino acid sequence displays 46% identity with that of abietadiene synthase,[122] so taxadiene synthase shares the general domain structure outlined in Figure 32a. The preparation of a truncation variant in which 60 residues of the plastidial targeting sequence are deleted from the N-terminus results in a "pseudomature" form of the enzyme that facilitates overexpression in *Escherichia coli*, purification, and assay.[123] This recombinant enzyme generates taxa-4(5),11(12)-diene as the major cyclization product (94%), but minor amounts of side products, taxa-4(20),11(12)-diene (~5%) and verticillene (~1%), are also detected.

Given that taxadiene synthase contains both class I and class II cyclase folds, as indicated by its significant amino acid sequence identity with abietadiene synthase, the first question regarding structure−mechanism relationships focuses on which domain(s) catalyze the cyclization cascade. Amino acid sequence alignments of taxadiene synthase and abietadiene synthase reveal that the signature DXDD motif of the class II cyclase domain of abietadiene synthase appears as DLNT in taxadiene synthase; since the central aspartate of this domain is typically the general acid that triggers the protonation-induced cyclization cascade of abietadiene synthase[112] and the triterpene cyclases (section 5), and since the central aspartate is absent in taxadiene synthase, it appears that the class II cyclase domain is not equipped for protonation-induced carbocation formation. Instead, the entire

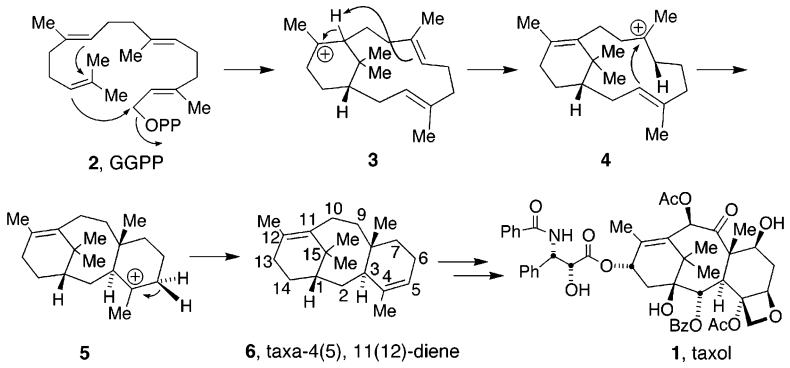


Figure 34. The cyclization of geranylgeranyl diphosphate (GGPP, **2**) catalyzed by taxadiene synthase requires Mg^{2+} to trigger the departure of the pyrophosphate leaving group (OPP). The cyclization cascade terminates with the formation of taxa-4(5),11(12)-diene **6**, which undergoes further biosynthetic derivatization in the Pacific yew to yield paclitaxel (Taxol, **1**). Reprinted from ref 126, Copyright (2000), with permission from Elsevier.

cyclization reaction occurs in the C-terminal class I cyclase domain and is triggered by the Mg^{2+}-dependent departure of the geranygeranyl diphosphate leaving group. Studies with isotopically labelled substrates indicate that the CH_2OPP group and the C14−C15 and C10−C11 π bonds of geranylgeranyl diphosphate are favorably aligned for the possibly concerted alkylation and π-cyclization to form the verticillyl carbocation intermediate **3** (Figure 34) and confirm that the C1 atom of the substrate undergoes inversion of configuration upon A ring formation.[123−125]

Additional studies with specifically deuterated substrates indicate an intramolecular proton transfer of the C11 proton to the *re* face of C7 to generate tertiary carbocation intermediate **4** (Figure 34).[124,126] This proton transfer is somewhat unusual in that a π-bond of the substrate, rather than an enzyme-bound residue, serves as a "general base" to accept the C11 proton. This step is analogous to the intramolecular proton transfer proposed for the sandaracopimerenyl cation in the abietadiene synthase mechanism (**3a**$^+$ → **3b**$^+$ in Figure 33). When 6-fluoro-geranylgeranyl diphosphate is used as a substrate with taxadiene synthase, the intramolecular C11 → C7 proton transfer is slower due to the electron-withdrawing effects of fluorine and B/C ring closure is blocked.[127] The resultant isolation and characterization of partially cyclized fluoroverticillenes shows that the stereochemistry at C11 of the verticillenyl carbocation intermediate is 11*R*.[127]

Transannular B/C ring closure of carbocation intermediate **4** leads to formation of the taxenyl carbocation **5** (Figure 34). Enzymological studies utilizing specifically deuterated geranylgeranyl diphosphate indicate that isomeric side products derive from this intermediate;[123] elimination of the H5β proton terminates the cyclization cascade to form taxa-4(5),11(12)-diene **6** (Figure 3).[126] Although the role of the enzyme in catalyzing this reaction is not clearly defined (e.g., does an enzyme residue mediate the final deprotonation of the taxenyl cation?), the future X-ray crystallographic structure determination of taxadiene synthase will allow such detailed structure−mechanism relationships to be conclusively established.[128]

5. Triterpene Cyclases

The triterpene cyclases are designated as class II terpenoid cyclases because their double α-barrel fold is topologically distinct from the characteristic α-helical fold of a class I terpenoid cyclase.[6] Triterpene cyclases catalyze the cyclization of the linear 30-carbon isoprene substrates squalene or squalene oxide to generate a variety of polycyclic products. Initial carbocation formation is achieved by protonation of a carbon−carbon double bond in squalene or by epoxide protonation and ring opening in squalene oxide. The central aspartate in the signature DXDD amino acid sequence motif is implicated as the requisite proton donor. Interestingly, squalene derives from two molecules of farnesyl diphosphate connected in "head-to-head" fashion by squalene synthase, the structure of which reveals topological similarities with class I terpenoid synthases.[14]

Despite the size of their substrates, the triterpene cyclases can achieve the same degree of structural and stereochemical precision in catalysis as previously noted for class I terpenoid cyclases. The X-ray crystal structures of two triterpene cyclases, squalene−hopene cyclase[129,130] and oxidosqualene cyclase (also known as lanosterol synthase),[131] now provide a critical foundation for understanding structure−mechanism relationships in this enzyme family. Given that an exhaustive and outstanding review of triterpene cyclization mechanisms has been presented,[10] here I focus primarily on structural aspects of triterpene cyclase structure−mechanism relationships discerned from experimentally determined crystal structures reported within the past two years.

5.1. Squalene–Hopene Cyclase

Squalene−hopene cyclase catalyzes the cyclization of squalene to form the pentacyclic hydrocarbon hopene (Figure 35). Hopene is a biosynthetic precursor of the bacterial hopanoids, which modulate membrane fluidity. The X-ray crystal structure of squalene−hopene cyclase from *Alicyclobacillus acidocaldarius* reveals a dimeric enzyme:[129,130] each monomer contains two domains, and each domain adopts a double α-barrel fold. The two domains assemble so as to enclose a central hydrophobic active site cavity (Figure 36).

Squalene−hopene cyclase is a monotopic membrane protein, meaning that it penetrates but does not completely pass through the bacterial membrane. Substrate squalene, dissolved in the membrane, enters the cyclase active site through a hydrophobic channel open to the membrane surface. A nonpolar "plateau" flanks the channel entrance in the vicinity of α-helix 8, and this likely comprises the monotopic membrane association motif (Figures 36 and 37). As with the class I terpenoid cyclases, the active site of squalene−hopene cyclase is very hydrophobic in nature and a variety of aromatic and aliphatic amino acid side chains define a precise active site contour that serves as a template

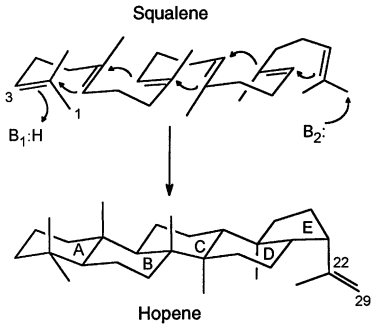


Figure 35. Squalene−hopene cyclase reaction. Various data indicate that general acid B_1:H is Asp-376, which protonates squalene at C3 and triggers the cyclization cascade. The crystal structure reveals no specific residue that could serve as general base B_2: for the final deprotonation step at C29. This could account for the generation of 10% hopan-22-ol as a side product resulting from quenching of the final C22 carbocation by solvent. Reprinted with permission from *Science* (http://www.aaas.org), ref 129. Copyright 1997 American Association for the Advancement of Science.

to enforce the catalytically productive conformation of the large, flexible polyisoprene substrate. The recent X-ray crystallographic structure determination of the complex with the nonreactive substrate analogue 2-azasqualene provides an elegant illustration of this template function (Figure 38), since the substrate analogue binds with a conformation very close to that required for the chair conformations of hopene rings A−D.[132]

The short amino acid sequence DXDD/E has been identified as a consensus motif in squalene cyclases,[133,134] and site-directed mutagenesis has been used to probe the catalytic function of aspartate residues in the squalene cyclase reaction. The central acidic residue, Asp-376, is the likely general acid responsible for protonating the C3 atom of the squalene substrate. The Asp-376 → Glu variant retains 10.3% relative catalytic activity compared with the wild-type enzyme, whereas the Asp-376 → Gly and Asp-376 → Arg variants are essentially inactive.[134] Similar amino acid substitutions of Asp-377 result in virtually inactive enzyme variants, presumably due to the disruption of the Asp-376 environment. Consistent with these mutagenesis results, the side chain of Asp-376 is oriented toward the N−C3 bond of 2-azasqualene in Figure 38, which corresponds to the face of the C2−C3 π bond of squalene (Figure 39).[132]

On the basis of the crystal structure of the complex between squalene−hopene cyclase and the polyisoprenoid substrate analogue 2-azasqualene (Figures 38 and 39), a mechanism for hopene formation is proposed[132] that involves low-barrier Markovnikov-type cation−olefin additions to form the six-membered A and B rings in the chair conformation, with a tertiary carbocation intermediate at C10. The binding conformation of 2-azasqualene suggests that although it is possible that the C10 carbocation could react with the C14−C15 π-bond in Markovnikov fashion to yield a five-membered C ring intermediate, it appears more likely that a six-membered C-ring is formed directly by anti-Markovnikov reaction of the C14−C15 π bond with the C10 carbocation in concert with Markovnikov reaction of the incipient secondary C14 carbocation with the C18−C19 double bond to yield a five-membered D ring with a resulting tertiary carbocation ion at C19. This intermediate is designated the 6−6−6−5 tetracycle with rings A−C adopting the chair−chair−chair conformation of hopene. Approach of the C22−C23 π bond toward the C19 carbocation appears to be sterically hindered, leading Reinert and colleagues to suggest that the 6−6−6−5 carbocation is a relatively long-lived intermediate.[132] That derivatives of the 6−6−6−5 carbocation are observed as minor byproducts in the squalene−hopene cyclase reaction[135] and as significant byproducts generated by site-specific variants of squalene−hopene cyclase[136] is consistent with a 6−6−6−5 intermediate.[132] Subsequent D-ring expansion and Markovnikov E-ring formation yields the tertiary hopenyl carbocation, which undergoes proton elimination to yield hopene. Detailed aspects of the chemical mechanisms leading to the formation

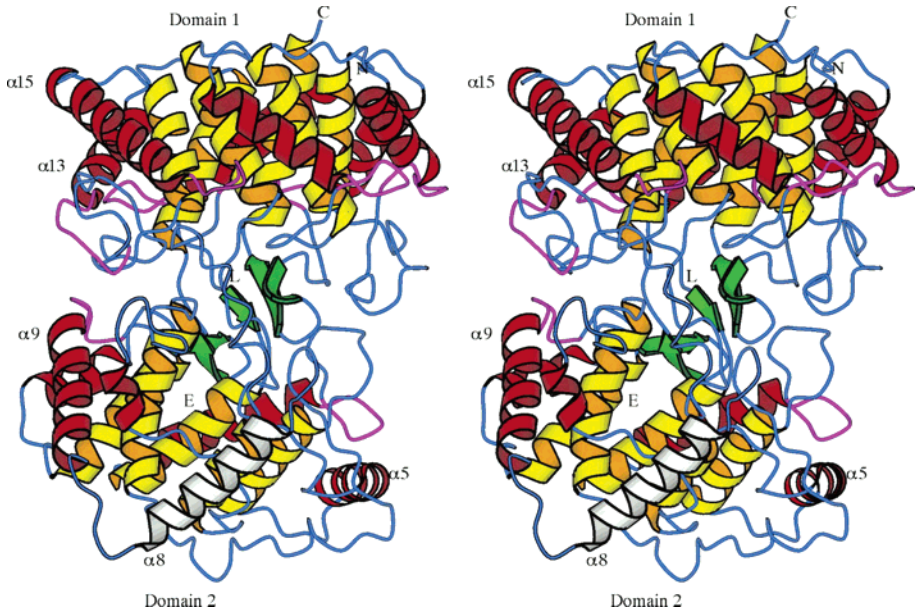


Figure 36. Structure of the squalene−hopene cyclase monomer. Amino (N) and carboxyl (C) terminal residues are located in domain 1. Internal and external α-helices of the two α/α-barrel domains are yellow and red, respectively. A small β-sheet (green) comprises one wall of the interior active site cavity (L), which is accessed through the hydrophobic channel entrance (E). Loop segments containing characteristic glutamine−tryptophan motifs are magenta; α-helix 8 in domain 2 (white) anchors the enzyme in the membrane. Reprinted with permission from *Science* (http://www.aaas.org), ref 129. Copyright 1997 American Association for the Advancement of Science.

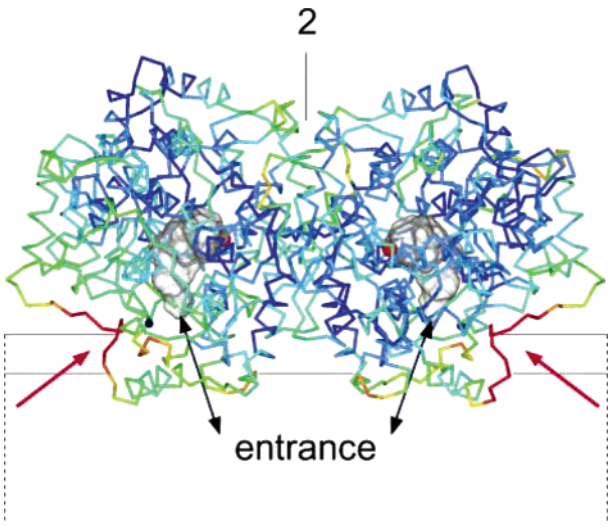


Figure 37. Model of dimeric squalene−hopene cyclase in the membrane. The hydrophobic active sites and channel entrances are indicated, and general acid Asp-376, which initiates the cyclization cascade by protonating C3 of squalene, is marked as a red dot near the "top" of each active site. Recent X-ray crystallographic studies indicate that the outer channel wall, including Phe-434 (black dot), is mobile (red arrows). Such mobility presumably facilitates substrate binding and product dissociation from the enzyme active site by allowing the channel to open fully. Reprinted from ref 132, Copyright (2004), with permission from Elsevier.

of hopene as well as alternative cyclization products have been reviewed by Wendt and colleagues.[10]

When the structure of squalene−hopene cyclase was initially reported,[129] it was noted that there is more conservation of amino acid residues in the upper region of the active site cavity (i.e., the polar, Asp-376 region) than in the lower region of the active site as squalene−hopene cyclase is compared with other triterpene cyclases, suggesting that the first reaction step common to squalene and oxidosqualene cyclases, substrate protonation, occurs in the conserved region of the cavity. The divergence of amino acid composition in the lower region of the active site cavity among triterpene cyclases thus allows for divergent biosynthetic reactions.[129]

5.2. Oxidosqualene Cyclase

Oxidosqualene cyclase catalyzes the exclusive cyclization of oxidosqualene to form lanosterol, a critical step in the biosynthesis of cholesterol (Figure 40). Human oxidosqualene cyclase functions as a monomer and exhibits 25% amino acid sequence identity with squalene−hopene cyclase.[137] Oxidosqualene cyclases usually contain only the central aspartate corresponding to the DXDD motif found in squalene cyclases: in human oxidosqualene cyclase this corresponds to Asp-455, and in oxidosqualene cyclase from *Saccharomyces cerevisiae* this corresponds to Asp-456. Site-directed mutagenesis with the enzyme from *S. cerevisiae* and the binding of irreversible inhibitors implicate Asp-456 as the general acid that protonates the substrate epoxide ring to initiate the cyclization cascade.[138,139]

The X-ray crystal structure of human oxidosqualene cyclase reveals dual double α-barrel domains connected by loops and a smaller *β*-sheet structure with identical topology to that observed in squalene−hopene cyclase (Figure 41).[131] As proposed for squalene−hopene cyclase,[129] a hydrophobic plateau flanking α-helix 8 likely serves as the monotopic membrane insertion motif. The binding of substrate oxidosqualene and the release of product lanosterol occurs through a hydrophobic channel adjacent to α-helix 8 that connects the active site cavity to the membrane surface. However, a *second* active site channel is observed in oxidosqualene cyclase: a polar, solvent-accessible channel leads to the "top" (i.e., the Asp-455 region) of the active site cavity, and the side chain of Glu-459 is located at the midpoint of this channel (Figure 41b).

The structure of human oxidosqualene cyclase complexed with product lanosterol interpreted in view of enzymological studies allows for the conclusive determination of active site

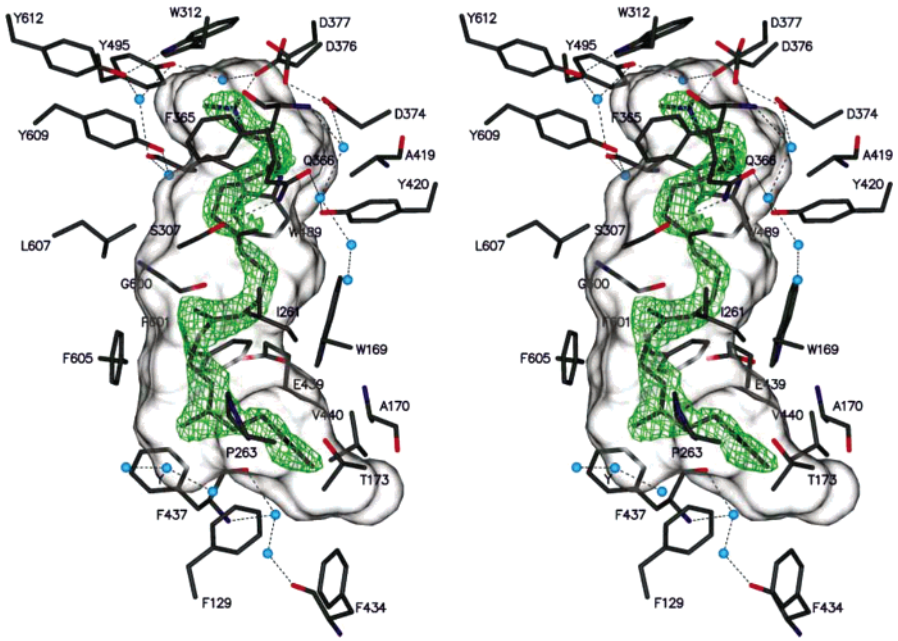


Figure 38. Electron density map showing the precise binding conformation of the polyisoprenoid substrate analogue inhibitor 2-azasqualene in the active site of squalene−hopene cyclase. The inhibitor conformation corresponds to the productive squalene conformation required for hopene biosynthesis, thereby exemplifying the precise template function of the terpenoid synthase active site. Reprinted from ref 132, Copyright (2004), with permission from Elsevier.

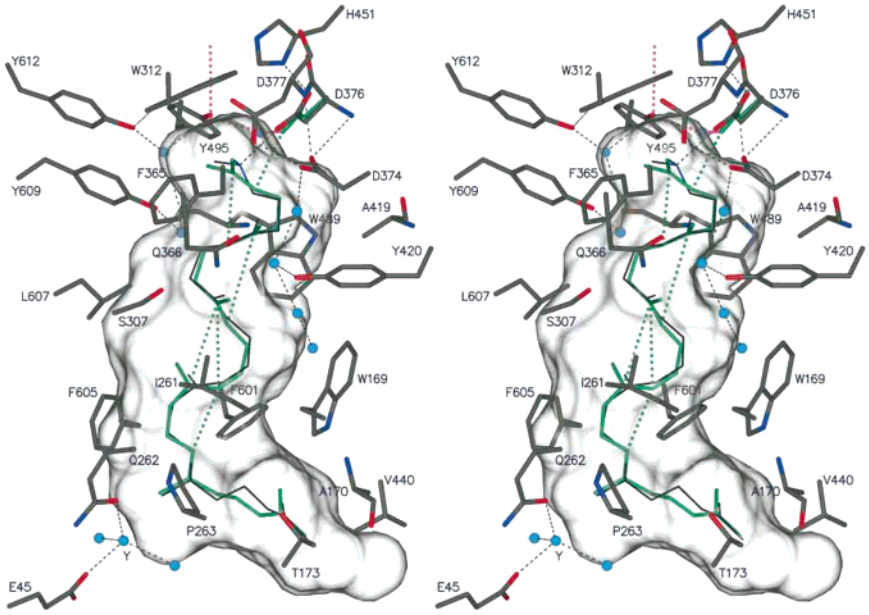


Figure 39. Model of squalene (green) based on the experimentally determined binding conformation of the nonreactive substrate analogue 2-azasqualene (thin dark bonds). Dotted green lines indicate the first four carbocation addition reactions at C2, C6, C10, C14−C15, and C19 that yield the 6−6−6−5 tetracycle. Reprinted from ref 132, Copyright (2004), with permission from Elsevier.

residues that participate in the cationic cyclization cascade.[131] At the "top" of the cavity, the side chain of Asp-455 hydrogen bonds with the lanosterol hydroxyl group, consistent with the proposed role of Asp-455 as the general acid that donates a proton to the epoxide oxygen of the substrate to initiate the cyclization cascade (Figure 42a,b).[138] The side chain of Asp-455 also accepts hydrogen bonds from Cys-456 and Cys-533, which may enhance the acidity of Asp-455. Additionally, ordered water molecules are observed bound in the upper active site, one of which hydrogen bonds with Asp-455 (Figure 42a,b). Intriguingly, Thoma and colleagues suggest that reprotonation of Asp-455 could be achieved by bulk solvent through a hydrogen bonded solvent network with Glu-459 in the polar channel leading to Asp-455 (Figure 41b).[131]

Several high-energy carbocation intermediates occur in the cyclization sequence proposed in Figure 40, and Thoma and colleagues advance that conserved aromatic side chains are appropriately positioned to stabilize these intermediates through cation−π interactions.[131] Specifically, Trp-387, Phe-444, and Trp-581 can stabilize tertiary carbocation intermediates at C6 and C10 (Figure 42b). Additionally, Tyr-98 protrudes into the active site cavity in such a manner that may preferentially stabilize the substrate in the unfavorable boat conformation required for lanosterol B-ring formation. Thus, as with the class I terpenoid cyclases, active site aromatic residues not only define the active site contour that serves as a template for binding the flexible polyisoprenoid substrate, but these residues can also stabilize high-energy cationic intermediates during catalysis without danger of prematurely quenching the cyclization cascade.

Since the oxidosqualene cyclase mechanism has been thoroughly reviewed,[10] it is only briefly recounted here in view of the X-ray crystal structure of the enzyme. The anti-Markovnikov cyclization proposed to form the lanosterol C-ring (Figure 40) results in an unstable secondary carbocation at C14, but the side chains of His-232 and Phe-696 are appropriately oriented to stabilize this intermediate through cation−π interactions (Figure 42c).[131] Although the imidazole side chain of His-232 could conceivably be alkylated by such a reactive carbocation, a hydrogen bond with Tyr-503 appears to be sufficient to prevent His-232 from adopting a conformation that would allow it to react with a C14 carbocation. The closure of the D-ring yields the protosterol C20 cation, which is subsequently converted into lanosterol by tandem 1,2-hydride transfer and 1,2-methyl migration reactions (Figure 40). Thoma and colleagues[131] suggest that the side chain of His-232, an essential and strictly conserved residue among the oxidosqualene cyclases,[139] is the only possible candidate for the base that accepts the C9 proton ultimately eliminated to terminate the cyclization cascade (Figure 40). However, inspection of the active site also reveals that Tyr-503 (also conserved among the oxidosqualene cyclases) could possibly accept this proton, and it is closer and better-oriented with regard to the C9 proton that is ultimately eliminated to form lanosterol (Figure 42). Now that a three-dimensional crystal structure of oxidosqualene cyclase is available, the catalytic functions of Tyr-503 and other active site residues can be probed in further detail to clarify structure−mechanism relationships for this triterpene cyclase.

6. Antibody Catalysis of Cationic Cyclization Reactions

To date, several catalytic antibodies have been generated that catalyze cationic cyclization reactions.[140−146] Especially notable among these novel biological catalysts is antibody 19A4, which catalyzes the tandem cationic cyclization of a linear substrate to form a *trans*-decalin ring system analogous to that of the steroid A and B rings (Figure 43).[144] In comparison, antibody 4C6 catalyzes a slightly simpler reaction, the cyclization of a linear substrate to form a cyclohexane ring system (Figure 44). A key biosynthetic advantage of catalytic antibody cyclases is that they are not limited to the handful of naturally occurring isoprenoid polyene substrates illustrated in Figure 1. Therefore, the potential product diversity of antibody cyclases is even greater than that of the naturally occurring terpenoid cyclases.

Although the characteristic β-fold of the immunoglobulin Fab fragment is topologically distinct from the α-folds of

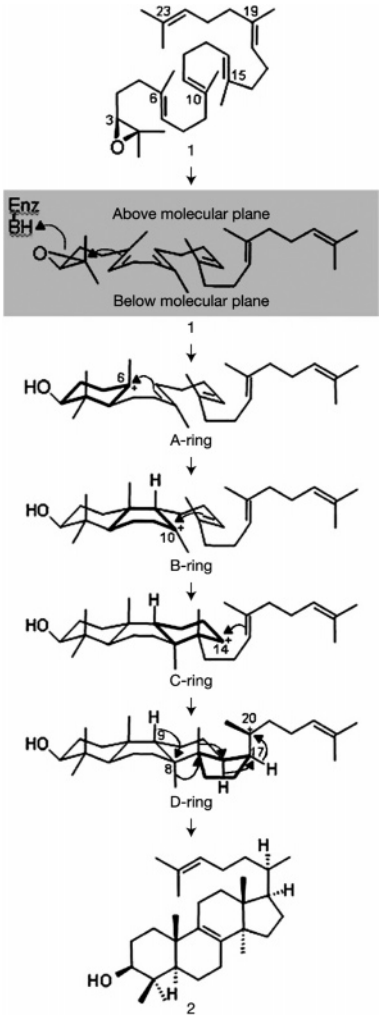


Figure 40. Oxidosqualene cyclase, also known as lanosterol synthase, catalyzes the cyclization of (*S*)-2,3-oxidosqualene to form lanosterol, a key intermediate in cholesterol biosynthesis. The initial substrate chair−boat−chair conformation, proposed conformations following A-, B-, and C-ring closure, and rearrangement of the protosterol cation via 1,2-hydride and 1,2-methyl transfers are indicated. Reprinted with permission from *Nature* (http://www.nature.com), ref 131. Copyright 2004 Nature Publishing Group.

the naturally occurring terpenoid cyclases, the X-ray crystal structures of catalytic antibody cyclases 19A4 and 4C6 reveal important features in the antibody combining sites indicating convergent structure−mechanism relationships with the terpenoid cyclases for the management and manipulation of reactive carbocation intermediates.[147,148] Active site features shared by the "unnaturally-evolved" antibody cyclases and the "naturally-evolved" terpenoid cyclases include (1) a hydrophobic active site cavity capable of sequestering the substrate from bulk solvent, (2) a precisely formed active site contour, that is, a three-dimensional template, that selects for the productive conformation (and selects against unproductive conformations) of the flexible polyene substrate, (3) a "carbocation trigger", that is, polar residue(s) that trigger initial carbocation formation once the substrate is bound in a productive conformation, and (4) aromatic residues that may stabilize high-energy carbocation intermediates through cation−π interactions. In the remainder of this section, structure−mechanism relationships for catalytic antibody cyclases 19A4 and 4C6 are summarized to illustrate these convergent features in the catalysis of carbocation cyclization cascades.

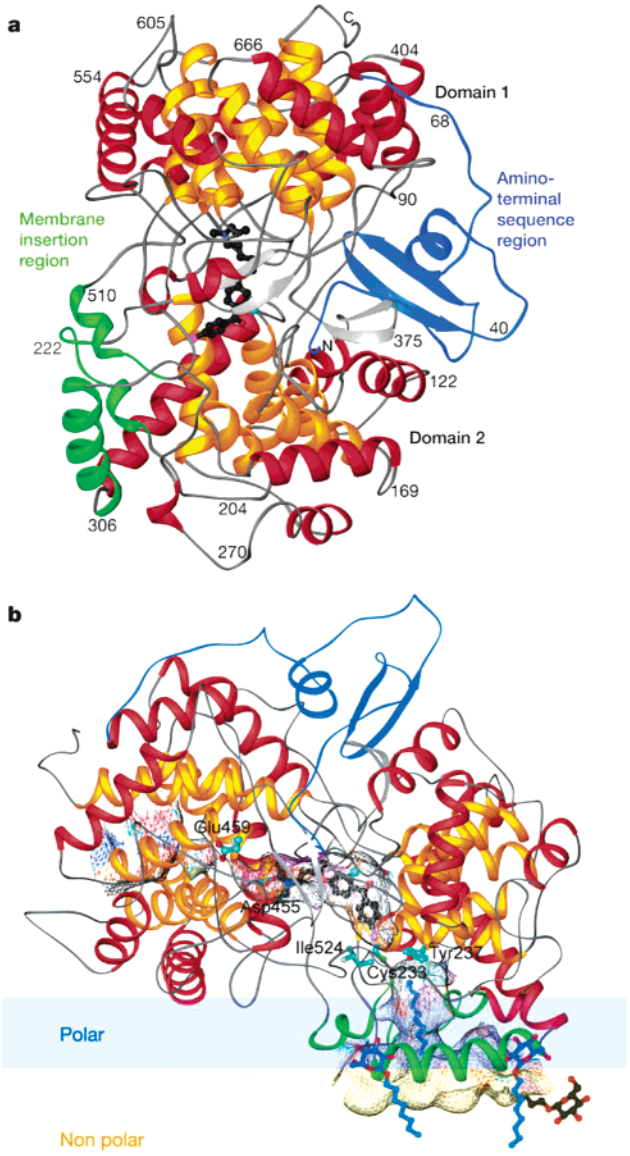


Figure 41. Panel a shows a ribbon diagram of human oxidosqualene cyclase. Inner helices of each α/α-barrel are yellow, and the inhibitor Ro 48-8071 (black) indicates the location of the active site cavity. Panel b shows a model of human oxidosqualene cyclase associated with a membrane leaflet (polar and nonpolar regions of the membrane are colored light blue and light yellow, respectively). Internal surfaces are colored as follows: blue, positive; red, negative; cyan, hydrogen bond donor; magenta, other polar. Ordered detergent molecules or lipid fragments observed in the crystal structure (blue and black stick figures) cluster around the membrane insertion region. Reprinted with permission from *Nature* (http://www.nature.com), ref 131. Copyright 2004 Nature Publishing Group.

6.1. Catalytic Antibody 19A4

The X-ray crystal structure of the Fab fragment of catalytic antibody 19A4 complexed with the bicyclic hapten **5** shown in Figure 43 reveals that the hapten is buried deeply within the antibody combining site and makes numerous van der Waals contacts with aromatic and aliphatic residues that mediate a highly complementary fit (Figure 45).[147] Thus, the precise contour of the antibody combining site serves as a three-dimensional template that binds the flexible polyene substrate in the productive chair−chair conformation required

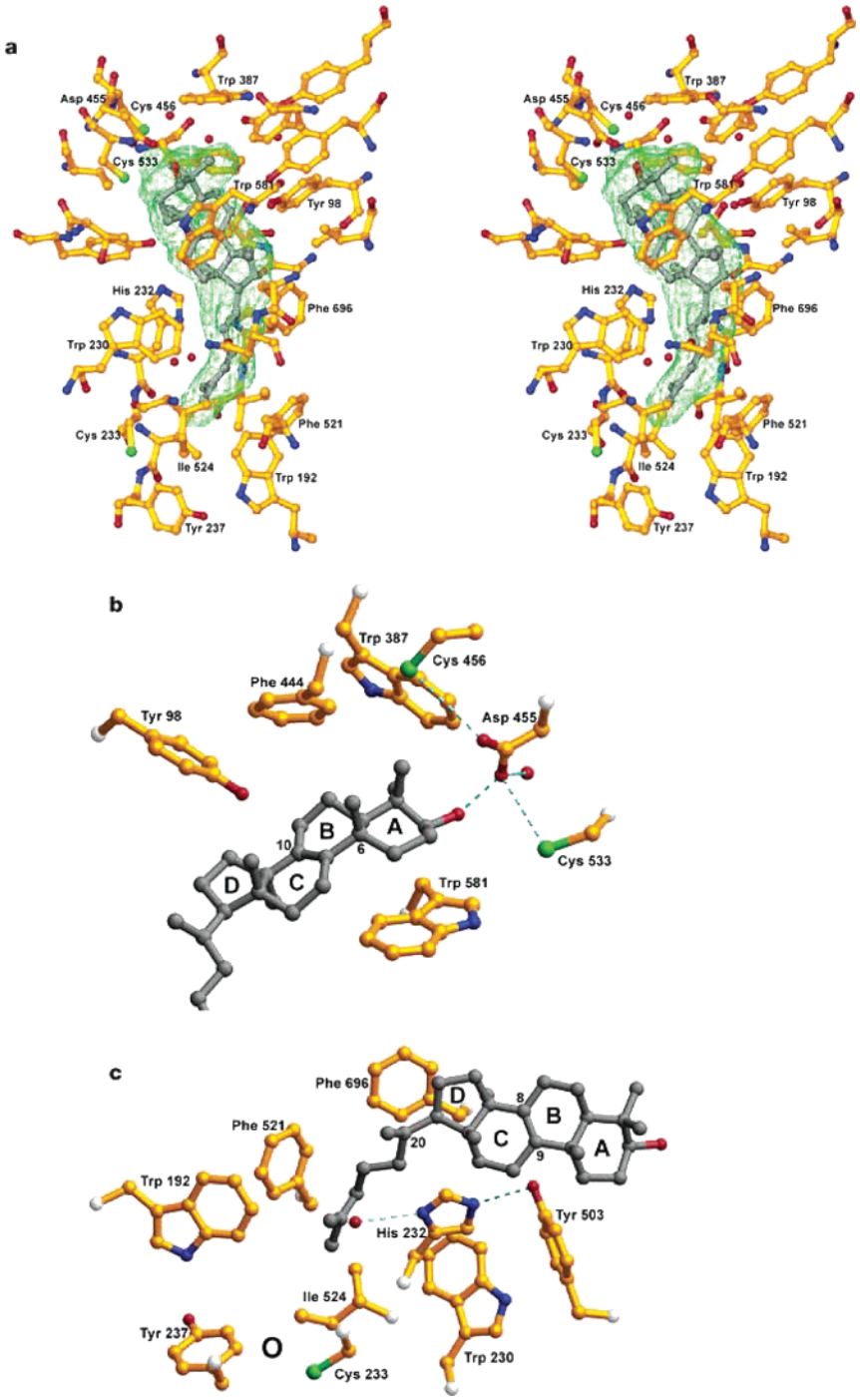


Figure 42. Panel a shows the omit electron density map showing the binding of product lanosterol in the active site of human oxidosqualene cyclase. Panel b shows that the hydroxyl group of lanosterol donates a hydrogen bond to the active site general acid, Asp-455. The aromatic side chains of Trp-387, Phe-444, and Trp-581 are suitably oriented to stabilize carbocation intermediates formed at C6 and C10 during the cyclization sequence through cation$-\pi$ interactions. In panel c, the aromatic side chains of Phe-696 and His-232 are suitably oriented to stabilize the penultimate carbocation intermediate at C20 (protosterol cation) through cation$-\pi$ interactions. Either His-232 or Tyr-503 is suitably oriented to accept the C9 proton from the lanosterol cation in the final elimination step yielding lanosterol. Reprinted with permission from *Nature* (http://www.nature.com), ref 131. Copyright 2004 Nature Publishing Group.

for *trans*-decalin formation. The crystal structure of the native 19A4 Fab and comparison with the hapten-bound structure reveals only limited structural changes triggered by hapten binding, the most significant of which appears to be the $\sim$2$-$3 Å movement of complementarity determining region (CDR) loop H3 toward the bound hapten.[149] This slight structural change sequesters the antibody combining site from bulk solvent and thereby helps to protect high-energy carbocation intermediates from premature quenching by water. It is interesting to note that, as for the naturally occurring terpenoid cyclases, the cyclization template provided by the active site contour does not appear to achieve its final form until substrate binds.

The crystal structure of the 19A4$-$hapten **5** complex[147] confirms the "bait-and-switch" strategy[150,151] underlying the incorporation of specific structural features in the hapten to elicit complementary functional groups to participate in catalysis. In the design of hapten **5**, Hasserodt and colleagues reasoned that in view of the anticipated cyclization mechanism for *trans*-decalin formation, a key feature of a functional

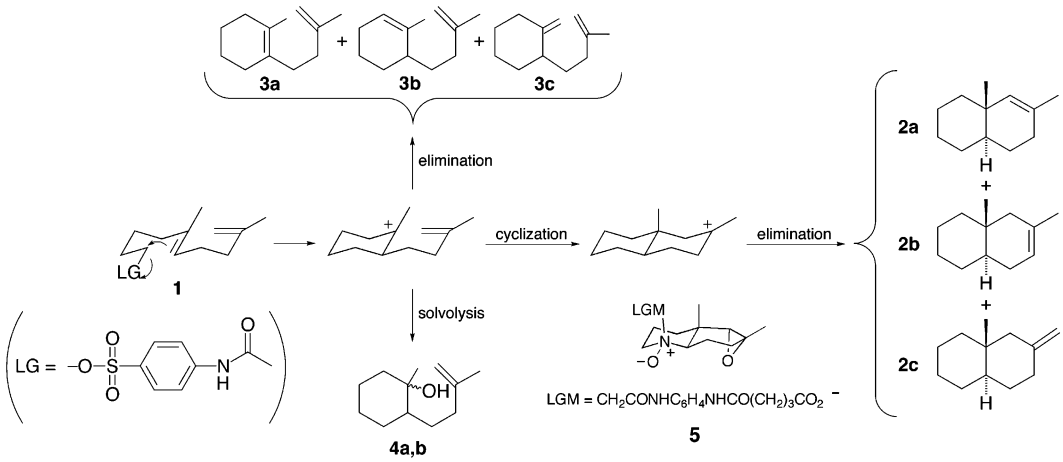


Figure 43. Catalytic antibody 19A4 catalyzes the formation of *trans*-decalin regioisomers **2a**−**c** as the major cyclization products from substrate **1**; premature quenching of the first carbocation intermediate yields olefins **3a**−**c** or alcohols **4a**,**b**.[144] Kinetic characterization based on the rate of sulfonic acid release yields $k_{cat} = 0.021$ min^{-1} and $K_M = 320$ μM; for solvolysis, $k_{uncat} = 9.2 \times 10^{-6}$ min^{-1}, so the rate acceleration for leaving group departure is 2.3×10^3. Bicyclic hapten **5** was used to elicit the antibody and partially mimics the productive chair−chair conformation required for cyclization to form the *trans*-decalin; additionally, the leaving group mimic (LGM)/*N*-oxide moiety was incorporated into the hapten design to elicit functional groups in the antibody combining site that might facilitate leaving group (LG) departure.

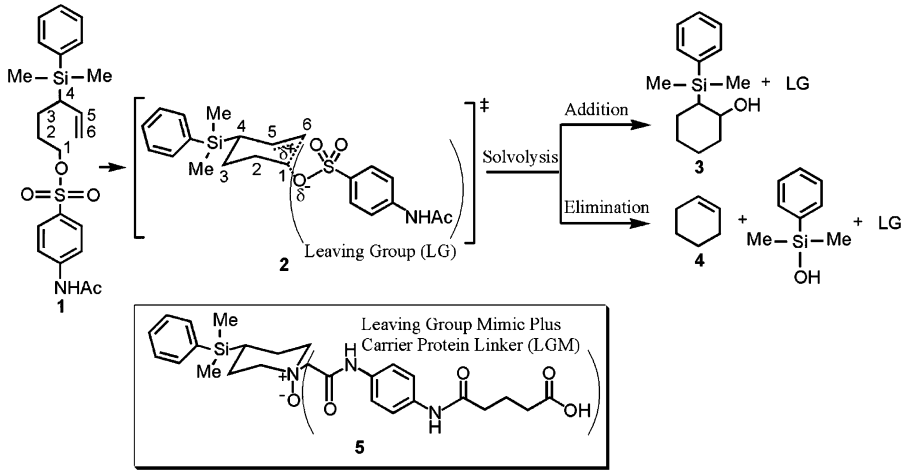


Figure 44. Cationic monocyclization reaction catalyzed by antibody 4C6 affords products *trans*-2-(dimethylphenylsilyl)-cyclohexanol 3 (98%) and cyclohexene 4 (2%).[140,148] Hapten **5** was used to elicit the catalytic antibody. Reprinted from ref 148, Copyright (2003), with permission from Elsevier.

antibody cyclase would be the initiation and stabilization of the first carbocation intermediate; accordingly, the *N*-oxide moiety of hapten **5** would be appropriately located to elicit complementary residues in the antibody combining site that could trigger leaving group departure and stabilize the first carbocation intermediate.[144] The structure of the 19A4−hapten **5** complex reveals a hydrogen bond between the negatively charged oxygen atom of the *N*-oxide moiety and the side chain carboxamide group of Asn-H35A, thereby suggesting that this residue helps trigger the departure of the sulfonate leaving group in catalysis (Figure 45b).[147] Additionally, the positively charged nitrogen atom of the *N*-oxide moiety interacts predominantly with the π-electron clouds of Trp-L91 and Tyr-L96. Since this nitrogen atom corresponds to the substrate C1 atom (Figure 43), which develops partial positive charge during leaving group departure, it appears that the incorporation of the *N*-oxide moiety in hapten **5** successfully elicited residues capable of stabilizing the developing partial positive charge at C1 (Figure 46a). This feature, too, may facilitate leaving group departure in catalysis. Thus, a hapten used to elicit a catalytic antibody may not only incorporate structure features anticipated in the transition state,[152,153] but it may also include additional features capable of eliciting complementary functional groups in the antibody combining site to perform specific chemical steps in catalysis.[150,151]

The proposed cyclization mechanism catalyzed in the active site of antibody 19A4 is summarized in Figure 46.[147,149] In concert with leaving group departure triggered by Asn-H35A, Trp-L91, and Tyr-L96 (Figure 46a), concerted attack of the C5−C6 π bond at C1 forms the six-membered A-ring of the *trans*-decalin with the tertiary C5 carbocation stabilized by cation−π interactions with Trp-L91 and Tyr-L36 (Figure 46b). The generation of minor olefin side products **3a**−**c** and alcohol side products **4a**,**b** (Figure 43)[144] is consistent with formation of a carbocation intermediate at C5. The C5 carbocation is subsequently attacked by the C9−C10 π bond to yield the six-membered B-ring of the *trans*-decalin (Figure 46c). The resulting carbocation intermediate at C9 is stabilized by cation−π interactions with Tyr-L36 and Trp-H103. It is interesting to note that the constellation of aromatic residues in the antibody combining site is such that

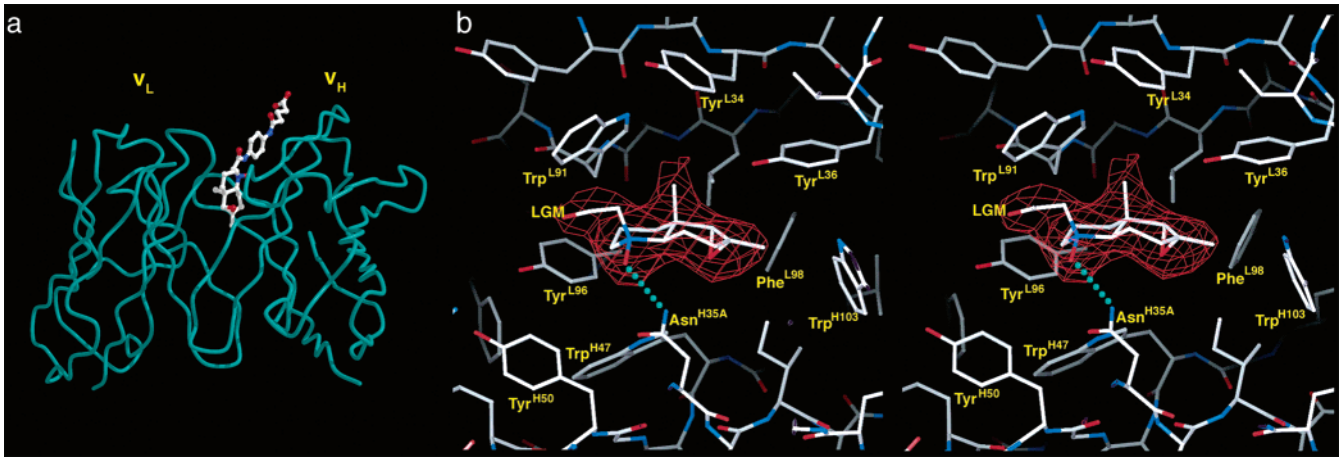


Figure 45. Panel a shows the backbone Cα trace of the variable region of antibody HA5−19A4 complexed with hapten **5** (Figure 43). In panel b, omit electron density map of hapten **5** bound to antibody HA5−19A4 reveals that numerous aromatic residues surround the hapten to form a highly complementary surface, and this surface contour must accordingly be highly complementary for the corresponding tandem bicyclic reaction cascade yielding *trans*-decalin regioisomers **2a**−**c** shown in Figure 43 (for clarity, only a portion of the leaving group mimic (LGM) is shown). A hydrogen bond between the N^+-O^- moiety and Asn-H35A suggests that this residue triggers leaving group departure in catalysis. Reprinted with permission from ref 147. Copyright 1999 Wiley-VCH.

each site of developing partial or full positive charge on the substrate—C1, C5, and C9—is potentially stabilized by two cation−π interactions.

Termination of the cyclization cascade is achieved by proton elimination to yield *trans*-decalin regioisomers **2a**−**c** (Figure 46d). Although the oxirane moiety was incorporated into the design of hapten **5** to elicit a polar residue that could facilitate a specific termination reaction for the cyclization cascade, *trans*-decalins **2a**−**c** are generated with product ratios reflecting their expected thermodynamic stabilities.[144] The crystal structure of the HA5−19A4 complex with hapten **5** reveals that there is no catalytic base appropriately positioned in the antibody combining site to direct the regiochemistry of the final elimination step to yield one exclusive elimination product.[147] Thus, in this system the oxirane moiety is not as successful as the *N*-oxide moiety in the "bait-and-switch" strategy aimed at eliciting a complementary functional group in the antibody combining site.[144]

6.2. Catalytic Antibody 4C6

The cationic monocyclization reaction catalyzed by antibody 4C6 is shown in Figure 44.[140] Here, too, an *N*-oxide moiety was incorporated into the hapten as part of a "bait-and-switch" strategy:[150,151] the negatively charged oxygen atom of *N*-oxide hapten **5** in Figure 44 was intended to elicit residues that could potentially stabilize the developing negative charge on the sulfonate leaving group, and the positively charged nitrogen atom was intended to elicit residues capable of stabilizing the developing partial positive charge at the corresponding substrate C1 atom as well as the subsequently formed C5 carbocation.

The X-ray crystal structure of the 4C6 Fab complexed with hapten **5** illustrated in Figure 44 reveals that the phenolic side chain of Tyr-L96 donates a hydrogen bond to the oxygen atom of the *N*-oxide moiety (Figure 47),[148] thereby confirming the successful "bait-and-switch" strategy underlying hapten design.[140] As expected, the three-dimensional contour of the antibody combining site is quite complementary in shape to the hapten. The binding pocket is primarily hydrophobic in nature and contains several aromatic residues that potentially participate in catalysis (Figure 47). Structural analysis of the 4C6 Fab−hapten complex leads to a mechanistic proposal in which the active site stabilizes the binding of substrate in a quasi-chair-like conformation to accommodate the expected transition state structure with the sulfonate leaving group in a pseudoequatorial position (Figure 48).[140,148] The aromatic rings of both Tyr-L96 and Tyr-H50 are oriented such that they could donate hydrogen bonds to the sulfonate ester oxygen and thereby trigger leaving group departure. The aromatic rings of Tyr-L96, Trp-H97, and Phe-L32 could stabilize the developing partial positive charge at C1 that accompanies leaving group departure in concert with nucleophilic attack by the C5−C6 π bond, and Tyr-H50 and Trp-H97 may stabilize the C5 carbocation intermediate through cation−π interactions. The reaction is terminated primarily by addition of a water molecule to the C5 carbocation to form *trans*-2-(dimethylphenylsilyl)-cyclohexanol, although minor amounts of cyclohexene are formed by elimination of the dimethylphenylsilyl group (Figure 5).[140]

A detailed structural comparison of catalytic antibodies 19A4 and 4C6 is presented by Zhu and colleagues.[148] In general, the combining site of 19A4 extends more deeply into the interface between the variable regions of the heavy and light chains than that for 4C6. It is also notable that the molecular recognition of the *N*-oxide moiety of the hapten differs for each catalytic antibody: 19A4 utilizes an asparagine side chain, whereas 4C6 utilizes a tyrosine side chain, to donate a hydrogen bond to the negatively charged oxygen of the *N*-oxide moiety. Finally, numerous aromatic residues are found in each antibody combining site and may stabilize high-energy carbocation intermediates through cation−π interactions.

7. Concluding Remarks

A total of nine X-ray crystal structures of terpenoid cyclases and catalytic antibody cyclases have been reported in the past eight years, and it is instructive to compare these enzymes to understand the structural basis for their biosynthetic diversity. In terms of their structural biology, the naturally occurring terpenoid synthases adopt one of two topologically distinct α-helical folds, designated the class I and class II terpenoid synthase folds.[6] The class I terpenoid cyclase fold achieves initial carbocation formation by

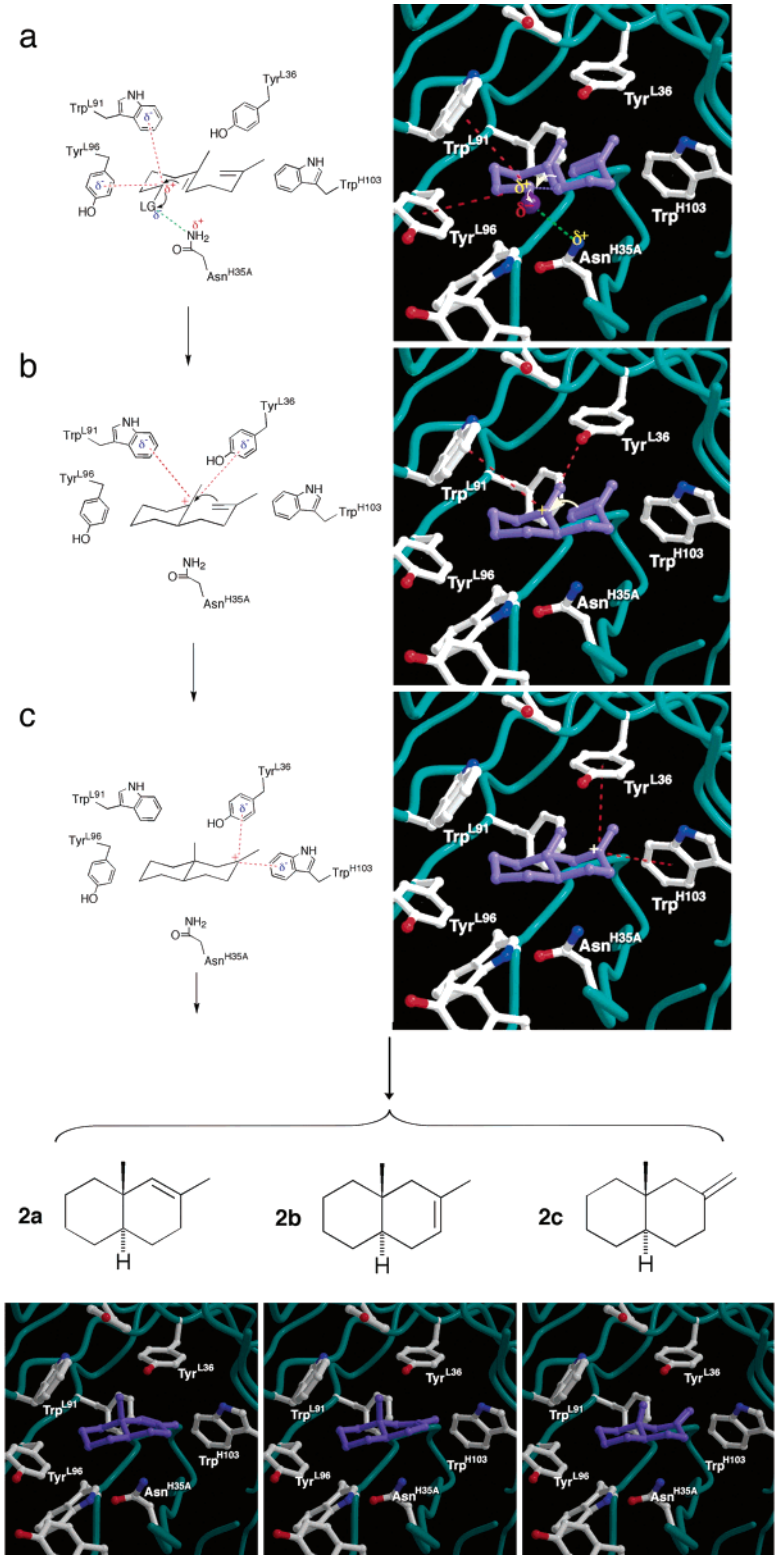


Figure 46. Proposed cyclization mechanism of catalytic antibody HA5−19A4 based on enzymological and X-ray crystallographic analysis.[144,147,149] Structural chemical aspects of this mechanism are discussed in the text. Reprinted with permission from ref 149. Copyright 2000 Chiana M. Paschall.

substrate ionization, and the class II terpenoid cyclase fold achieves initial carbocation formation by substrate protonation.

In class I and class II terpenoid cyclases, amino acid side chains critical for catalysis are located on or adjacent to α-helices that form a barrel-type superstructure in which active site is nested. The active sites of class I terpenoid cyclases (monoterpene, sesquiterpene, and diterpene cyclases) are formed by single α-barrel structures homologous to farnesyl diphosphate synthase (Figure 3);[16] the active site of a class II diterpene cyclase, the N-terminal domain of the bifunctional cyclase abietadiene synthase, is formed by a single-domain α-barrel structure based on amino acid sequence analysis (Figure 32);[101] and two α-barrel domains associate in head-to-head fashion to form the active site cavity of a triterpene cyclase with substrate access through

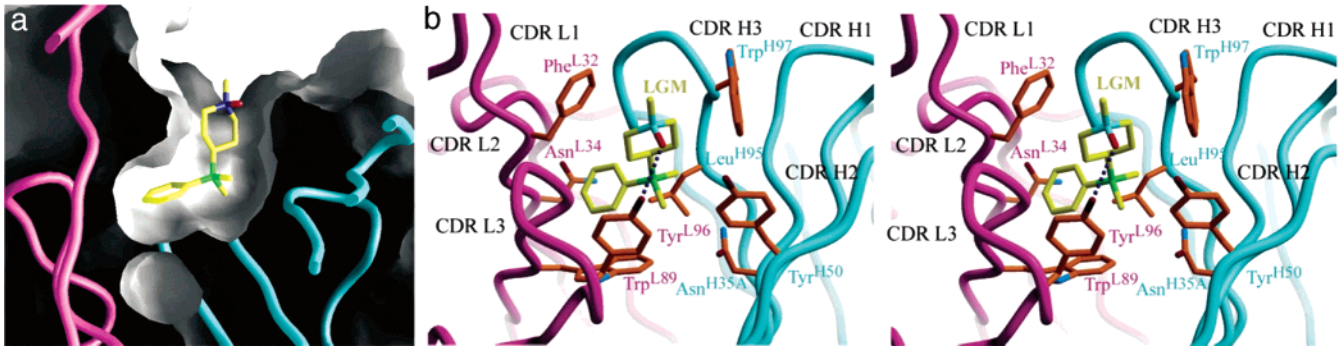


Figure 47. Panel a shows the shape complementarity between the combining site of catalytic antibody 4C6 and hapten **5** illustrated in Figure 44. Panel b presents a stereoview of the 4C6−hapten **5** complex showing antibody combining site residues that may participate in catalysis. Reprinted from ref 148, Copyright (2003), with permission from Elsevier.

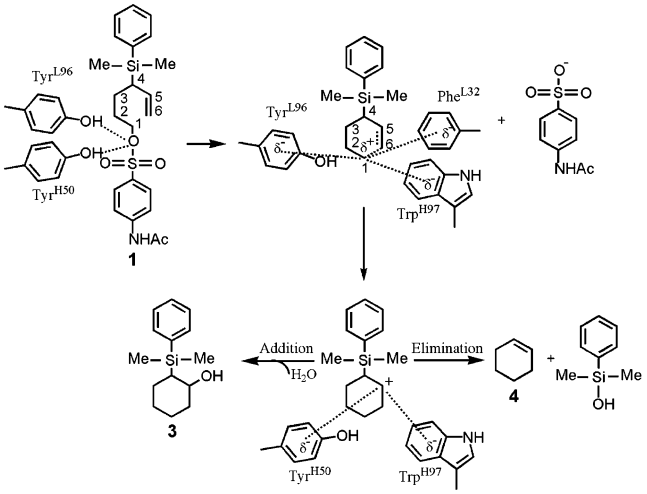


Figure 48. Proposed cyclization mechanism of catalytic antibody 4C6 based on enzymological and X-ray crystallographic analysis.[140,148] Structural chemical aspects of this mechanism are discussed in the text. Reprinted from ref 148, Copyright (2003), with permission from Elsevier.

a narrow tunnel at the interdomain interface. The generation of antibody catalysts that cyclize isoprenoid-like substrates demonstrates that carbocation cyclization cascades are readily catalyzed within β-sheet frameworks as well, so long as there is a suitable mechanism for triggering carbocation formation once the substrate is bound in a productive conformation.

The "aspartate-rich" motif[17−19] DDXXD/E and the "NSE/DTE" motif[31,60] (N/D)DXX(S/T)XXXE are located on opposite sides of the active site cleft of a class I terpenoid cyclase (e.g., see Figure 13), and these motifs signal the binding of a trinuclear magnesium cluster that triggers substrate ionization and initial carbocation formation. Crystal structures of farnesyl diphosphate synthase (Figures 5 and 6), (+)-bornyl diphosphate synthase (Figure 14), trichodiene synthase (Figure 21b), and epi-aristolochene synthase (Figure 26) reveal that the first aspartate in the aspartate-rich motif uniformly coordinates to two metal ions, Mg_A^{2+} and Mg_C^{2+}, with *syn,syn*-bidentate coordination stereochemistry. The second aspartate in this motif does not make any direct interactions with metal ions, but it does appear to engage in other important functions in some enzymes, for example, the Asp-101−Arg-204 salt link stabilizes the closed active site conformation of trichodiene synthase (Figure 21d). The third aspartate in this motif bridges Mg_A^{2+} and Mg_C^{2+} through a single carboxylate oxygen with *syn,anti* coordination stereochemistry in the active sites of farnesyl diphosphate synthase (Figures 5 and 6) and (+)-bornyl diphosphate synthase (Figure 14); it coordinates to Mg_A^{2+} and only poorly, if at all, to Mg_C^{2+} in the active site of epi-aristolochene synthase (Figure 26); and it does not coordinate to any metal ions in the active site of trichodiene synthase (Figure 21b). Thus, the first aspartate in the aspartate-rich motif has retained a common metal coordination function in the evolution of class I terpenoid synthases of known structure, whereas the metal coordination geometry and function of the third aspartate is more divergent. Divergent function is also implied by divergent amino acid sequence as well, for example, the third aspartate in the DDXXD motif appears as glutamate in certain terpenoid cyclases such as aristolochene synthase (section 3.3), and two additional residues are inserted in one of the DDXXD motifs to yield a DDXXXXD motif in farnesyl diphosphate synthase from *E. coli*.[22]

The second metal-binding motif of the terpenoid cyclase, (L,V)(V,L,A)(**N**,**D**)D(L,I,V)X(**S**,**T**)XXX**E** (the "NSE/DTE" motif; boldface residues form a complete Mg_B^{2+} binding site) appears to have diverged from the second aspartate-rich motif of farnesyl diphosphate synthase.[31,60] The NSE/DTE motif uniformly chelates a single metal ion, Mg_B^{2+}, in (+)-bornyl diphosphate synthase (Figure 14), trichodiene synthase (Figure 21b), and epi-aristolochene synthase (Figure 26). This motif therefore distinguishes the amino acid sequence of a terpenoid cyclase from that of a chain-elongating terpenoid synthase.

X-ray crystal structures of class I terpenoid cyclase complexes with PP_i, substrate analogues, and product reveal that the complete trinuclear magnesium cluster is stabilized by the substrate diphosphate group, the product PP_i anion, or both, which make numerous metal coordination and hydrogen bond interactions (e.g., see Figures 5, 6, 14, 21b, and 26). These interactions trigger structural changes in helices and loops flanking the mouth of the active site such that the active site cavity is closed and sequestered from bulk solvent. This ensures that the productive active site contour, the template for the cyclization cascade, is formed only after the substrate is fully bound and ready for ionization.

Significantly, even subtle changes in the metal coordination polyhedra dramatically alter cyclization product diversity, even while sustaining appreciable catalytic activity as described for site-specific variants of trichodiene synthase.[61,62] This suggests that one evolutionary strategy for terpene product diversity may be the introduction of amino acid substitutions in the metal coordination polyhedra, or residues that abut the trinuclear magnesium cluster, which can then

result in altered cyclization reaction. That the generation of alternative cyclization products increases when Mn^{2+} is substituted for Mg^{2+} in assays with Asp-100 → Glu trichodiene synthase clearly demonstrates that the metal coordination polyhedra and substrate diphosphate/PP_i interactions play a crucial role in governing cyclization product outcome.[62]

In contrast with the class I terpenoid cyclases, the class II terpenoid cyclases trigger initial carbocation formation by substrate protonation rather than ionization. Class II terpenoid cyclases contain the DXDD/E motif,[133,134] in which the central aspartic acid serves as a proton donor to either a carbon−carbon double bond (e.g., as in the N-terminal domain of abietadiene synthase (section 4.1) or squalene−hopene cyclase (section 5.1)) or the epoxide oxygen of squalene oxide in the reaction catalyzed by oxidosqualene cyclase (section 5.2). It is possible that conserved hydrogen bond interactions with the aspartic acid proton donor are important for function in squalene−hopene cyclase and oxidosqualene synthase, and it is especially intriguing that a polar tunnel with Glu-459 at its midpoint connects Asp-455 to bulk solvent in oxidosqualene cyclase (Figure 41). In terms of its structural and chemical nature, this tunnel is reminiscent of certain proton channels, and it could define a trajectory for reprotonating the active site general acid at the end of the cyclization cascade.[131] Given the ubiquitous occurrence of aspartic acid as a proton donor in the active sites of class II terpenoid cyclases, it is intriguing to speculate that aspartic acid residues may participate in the protonation of certain catalytic intermediates in the active sites of class I terpenoid cyclases.[71,72]

In protonation-dependent class II terpenoid cyclases such as squalene−hopene cyclase and oxidosqualene cyclase,[154] crystal structures of enzyme−ligand complexes suggest that the active site contour appears to be preformed in the productive conformation required for the cyclization template. This also appears to be the case for the antibody catalysts of polyene cyclization reactions discussed in section 6. In contrast, crystal structures of the ionization-dependent class I terpenoid cyclases complexed with PP_i or diphosphate-containing substrate analogues reveal that the productive active site contour is achieved only after the substrate is bound. Comparison of $Mg^{2+}{}_3$−PP_i-induced structural changes in (+)-bornyl diphosphate synthase (section 2), trichodiene synthase (section 3.1), and epi-aristolochene synthase (section 3.2) reveal some common regions of structure that are consistently involved in ligand-induced structural changes, for example, the J−K loop, but there are more differences than similarities. As noted by Whittington and colleagues,[43] examples include the H−α1 loop, which participates in active site closure mechanisms in (+)-bornyl diphosphate synthase and trichodiene synthase, but not epi-aristolochene synthase, and the N-terminus, which participates in active site closure mechanisms in (+)-bornyl diphosphate synthase and epi-aristolochene synthase, but not trichodiene synthase. Diversity in substrate recognition and active site closure mechanisms may correlate with terpenoid cyclase product diversity, but additional structural data must be acquired with other terpenoid cyclases to assess this possibility.

Regardless of whether the productive active site contour, the template for the isoprenoid cyclization reaction, is induced by substrate binding, as evident for class I cyclases such as (+)-bornyl diphosphate synthase and trichodiene synthase, or whether it is mostly preformed, as evident for class II cyclases such as squalene−hopene cyclase and oxidosqualene cyclase, the active site contour plays a key role as a template throughout the cyclization cascade. The template must not only guide the folding of the linear isoprenoid substrate but also guide the folding of all subsequently formed intermediates that can differ significantly from the substrate in terms of size, shape, conformational flexibility, and charge distribution. Thus, the active site contour of a high-fidelity terpenoid cyclase is optimized for complementarity to multiple structures along the reaction coordinate of cyclization (such complementarity could conceivably require flexibility in the template itself). In contrast, the active site contour of a terpenoid cyclase that generates multiple products is more promiscuous in its complementarity to one or more structures along the reaction coordinate of cyclization. It is interesting to note that fidelity and promiscuity are similarly balanced to accommodate multiple thioether cyclization reactions catalyzed by nisin cyclase, which unexpectedly adopts a double α-barrel fold comparable to that of a class II terpenoid cyclase.[155,156]

Class I and class II terpenoid cyclases and the catalytic antibody cyclases utilize the bulky aromatic residues phenylalanine, tyrosine, and tryptophan to form their active site contours, and in each cyclase these residues also appear to be oriented for the stabilization of carbocation intermediates. Interestingly, structural studies with the catalytic antibody cyclase 19A4 suggest that two aromatic residues are positioned to stabilize each site of developing positive or partial-positive charge as the polyene substrate undergoes cyclization (Figure 46),[147] and this also appears to be the case for the stabilization of certain carbocation intermediates in the oxidosqualene cyclase mechanism (Figure 42).[131] As previously discussed in section 1, aromatic residues can provide electrostatic stabilization of highly reactive carbocation intermediates without danger of alkylation. It is notable that both the naturally evolved terpene cyclases and the "unnaturally-evolved" antibody cyclases have converged to this chemical strategy for carbocation stabilization.

Finally, in view of the exquisite chemical complexity of the cyclization cascade catalyzed by a terpenoid cyclase, it is notable that given all the structural information on this family of enzymes acquired in the past eight years, the enzyme appears to plays surprisingly little chemical role once the cyclization cascade begins. While initial structure-based hypotheses suggested roles for active site residues as general bases or general acids, enzymological studies increasingly suggest that the substrate itself can provide chemical functionality required for specific general base/acid steps in the cyclization cascade. For example, a carbon−carbon double bond serves as a general base in the taxadiene synthase reaction (Figure 34),[124,126] and the product PP_i anion (or an associated water molecule) is considered to be a general base in cyclization reactions catalyzed by (+)-bornyl diphosphate synthase,[36,43] trichodiene synthase,[60,61] and aristolochene synthase[87] and the chain elongation reaction catalyzed by farnesyl diphosphate synthase.[22] Thus, the terpenoid cyclase itself may do little in the terpenoid cyclization reaction after initial carbocation formation other than to provide a productive template for the cyclization cascade and to stabilize reactive carbocation intermediates. Ironically, that these enzymes catalyze the most complex reactions in nature may be a consequence of the fact that these enzymes do not extensively participate in the cyclization chemistry! Future structural and enzymological studies of terpenoid

cyclases will no doubt allow us to elaborate upon the role of the enzyme in the chemistry and biology of Nature's most remarkable organic transformations.

8. Acknowledgments

I must first thank all the students and faculty collaborators with whom I've worked over the years; their contributions are noted in the citations. Additionally, I thank Luigi Di Costanzo, German Gomez, Ekaterina Shishova, and Sangeetha Vedula for helpful scientific discussions, and I especially thank German Gomez for his superb assistance with the figures. Finally, I thank Professors David E. Cane, Rod Croteau, Joseph P. Noel, and C. Dale Poulter for insightful discussions and comments on the manuscript. This work was supported by NIH Grant GM56838.

9. References

(1) Cane, D. E. *Acc. Chem. Res*. **1985**, *18*, 220.
(2) Cane, D. E. *Chem. Rev*. **1990**, *90*, 1089.
(3) Abe, I.; Rohmer, M.; Prestwich, G. D. *Chem. Rev.* **1993**, *93*, 2189.
(4) McCaskill, D.; Croteau, R. *Adv. Biochem. Eng. Biotechnol*. **1997**, *55*, 107.
(5) Lesburg, C. A.; Caruthers, J. M.; Paschall, C. M.; Christianson, D. W. *Curr. Opin. Struct. Biol*. **1998**, *8*, 695.
(6) Wendt, K. U.; Schulz, G. E. *Structure* **1998**, *6*, 127.
(7) Wise, M. L.; Croteau, R. In *Isoprenoids Including Carotenoids and Steroids*; Cane, D. E., Ed.; Comprehensive Natural Products Chemistry, Volume 2; Elsevier: Oxford, U.K., 1999; p 97.
(8) Cane, D. E. In *Isoprenoids Including Carotenoids and Steroids*; Cane, D. E., Ed.; Comprehensive Natural Products Chemistry, Volume 2; Elsevier: Oxford, U.K., 1999; p 155.
(9) Davis, E. M.; Croteau, R. *Top. Curr. Chem*. **2000**, *209*, 53.
(10) Wendt, K. U.; Schulz, G. E.; Corey, E. J.; Liu, D. R. *Angew. Chem., Int. Ed.* **2000**, *39*, 2812.
(11) Steele, C. L.; Crock, J.; Bohlmann, J.; Croteau, R. *J. Biol. Chem*. **1998**, *273*, 2078.
(12) Kellogg, B. A.; Poulter, C. D. *Curr. Opin. Chem. Biol*. **1997**, *1*, 570.
(13) Poulter, C. D. *Acc. Chem. Res.* **1990**, *23*, 70.
(14) Pandit, J.; Danley, D. E.; Schulte, G. K.; Mazzalupo, S.; Pauly, T. A.; Hayward, C. M.; Hamanaka, E. S.; Thompson, J. F.; Harwood, H. J., Jr. *J. Biol. Chem*. **2000**, *275*, 30610.
(15) Reardon, D.; Farber, G. K. *FASEB J.* **1995**, *9*, 497.
(16) Tarshis, L. C.; Yan, M.; Poulter, C. D.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 10871.
(17) Ashby, M. N.; Edwards, P. A. *J. Biol. Chem*. **1990**, *265*, 13157.
(18) Joly, A.; Edwards, P. A. *J. Biol. Chem*. **1993**, *268*, 26983.
(19) Song, L.; Poulter, C. D. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3044.
(20) Tarshis, L. C.; Proteau, P. J.; Kellogg, B. A.; Sacchettini, J. C.; Poulter, C. D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 15018.
(21) Phan, R. M.; Poulter, C. D. *Org. Lett*. **2000**, *2*, 2287.
(22) Hosfield, D. J.; Zhang, Y.; Dougan, D. R.; Broun, A.; Tari, L. W.; Swanson, R. V.; Finn, J. *J. Biol. Chem.* **2004**, *279*, 8526.
(23) Poulter, C. D.; Satterwhite, D. M. *Biochemistry* **1977**, *16*, 5470.
(24) Poulter, C. D.; Argyle, J. C.; Mash, E. A. *J. Biol. Chem*. **1978**, *253*, 7227.
(25) Poulter, C. D.; Wiggins, P. L.; Le, A. T. *J. Am. Chem. Soc*. **1981**, *103*, 3926.
(26) Davisson, V. J.; Neal, T. R.; Poulter, C. D. *J. Am. Chem. Soc*. **1993**, *115*, 1235.
(27) Davisson, V. J.; Poulter, C. D. *J. Am. Chem. Soc*. **1993**, *115*, 1245.
(28) Cornforth, J. W.; Cornforth, R. H.; Donninger, C.; Popják, G. *Proc. R. Soc. London, Ser. B: Biol. Sci.* **1966**, *163*, 492.
(29) Poulter, C. D.; Rilling, H. C. *Acc. Chem. Res.* **1978**, *11*, 307.
(30) Ohnuma, S.; Nakazawa, T.; Hemmi, H.; Hallberg, A. M.; Koyama, T.; Ogura, K.; Nishino, T. *J. Biol. Chem.* **1996**, *271*, 10087.
(31) Cane, D. E.; Kang, I. *Arch. Biochem. Biophys.* **2000**, *376*, 354.
(32) Bordwell, F. G. *Acc. Chem. Res.* **1970**, *3*, 281.
(33) Cane, D. E.; Yang, G.; Xue, Q.; Shim, J. H. *Biochemistry* **1995**, *34*, 2471.
(34) Dougherty, D. A. *Science* **1996** *271*, 163.
(35) Jenson, C.; Jorgensen, W. L. *J. Am. Chem. Soc*. **1997**, *119*, 10846.
(36) Wise, M. L.; Savage, T. J.; Katahira, E.; Croteau, R. *J. Biol. Chem*. **1998**, *273*, 14891.
(37) Cane, D. E.; Saito, A.; Croteau, R. Shaskus, J.; Felton, M. *J. Am. Chem. Soc*. **1982**, *104*, 5831.
(38) Croteau, R. B.; Shaskus, J. J.; Renstrøm, B.; Felton, N. M.; Cane, D. E.; Saito, A.; Chang, C. *Biochemistry* **1985**, *24*, 7077.
(39) Croteau, R.; Felton, M. *Arch. Biochem. Biophys.* **1981**, *207*, 460.
(40) Croteau, R.; Felton, N. M.; Wheeler, C. J. *J. Biol. Chem*. **1985**, *260*, 5956.
(41) Croteau, R.; Satterwhite, D. M.; Wheeler, C. J.; Felton, N. M. *J. Biol. Chem*. **1989**, *264*, 2075.
(42) Wise, M. L.; Pyun, H.-J.; Helms, G.; Assink, B.; Coates, R. M.; Croteau, R. B. *Tetrahedron* **2001**, *57*, 5327.
(43) Whittington, D. A.; Wise, M. L.; Urbansky, M.; Coates, R. M.; Croteau, R. B.; Christianson, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15375.
(44) Williams, D. C.; McGarvey, D. J.; Katahira, E. J.; Croteau, R. *Biochemistry* **1998**, *37*, 12213.
(45) Schwab, W.; Williams, D. C.; Davis, E. M.; Croteau, R. *Arch. Biochem. Biophys.* **2001**, *392*, 123.
(46) Vesonder, R. F.; Ciegler, A.; Jensen, A. H.; Rohwedder, W. K.; Weisleder, D. *Appl. Environ. Microbiol.* **1976**, *31*, 280.
(47) Tanaka, T.; Hasegawa, A.; Yamamoto, S.; Lee, U.-S.; Sugiura, Y.; Ueno, Y. *J. Agric. Food Chem.* **1988**, *36*, 979.
(48) Hohn, T. M.; Vanmiddlesworth, F. *Arch. Biochem. Biophys*. **1986**, *251*, 756.
(49) Hohn, T. M.; Beremand, P. D. *Gene* **1989**, *79*, 131.
(50) Hohn, T. M.; Plattner, R. D. *Arch. Biochem. Biophys*. **1989**, *275*, 92.
(51) Cane, D. E.; Wu, Z.; Oliver, J. S.; Hohn, T. M. *Arch. Biochem. Biophys*. **1993**, *300*, 416.
(52) Cane, D. E.; Swanson, S.; Murthy, P. P. N. *J. Am. Chem. Soc*. **1981**, *103*, 2136.
(53) Cane, D. E.; Ha, H.-J. *J. Am. Chem. Soc.* **1986**, *108*, 3097.
(54) Cane, D. E.; Ha, H.-J. *J. Am. Chem. Soc*. **1988**, *110*, 6865.
(55) Cane, D. E.; Pawlak, J. L.; Horak, R. M.; Hohn, T. M. *Biochemistry* **1990**, *29*, 5476.
(56) Cane, D. E.; Chiu, H.-T.; Liang, P.-H.; Anderson, K. S. *Biochemistry* **1997**, *36*, 8332.
(57) Cane, D. E.; Yang, G.; Coates, R. M.; Pyun, H.-J.; Hohn, T. M. *J. Org. Chem*. **1992**, *57*, 3454.
(58) Vedula, L. S.; Rynkiewicz, M. J.; Pyun, H.-J.; Coates, R. M.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2005**, *44*, 6153.
(59) Vedula, L. S.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2005**, *44*, 12719.
(60) Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13543.
(61) Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *Biochemistry* **2002**, *41*, 1732.
(62) Cane, D. E.; Xue, Q.; Fitzsimons, B. C. *Biochemistry* **1996**, *35*, 12369.
(63) Cane, D. E.; Shim, J. H.; Xue, Q.; Fitzsimons, B. C.; Hohn, T. M. *Biochemistry* **1995**, *34*, 2480.
(64) Cane, D. E.; Xue, Q. *J. Am. Chem. Soc.* **1996**, *118*, 1563.
(65) Whitehead, I. M.; Threlfall, D. R.; Ewing, D. F. *Phytochemistry* **1989**, *28*, 775.
(66) Whitehead, I. M.; Ewing, D. F.; Threlfall, D. R.; Cane, D. E.; Prabhakaran, P. C. *Phytochemistry* **1990**, *29*, 479.
(67) Back, K. W.; Yin, S. H.; Chappell, J. *Arch. Biochem. Biophys*. **1994**, *315*, 527.
(68) Starks, C. M.; Back, K.; Chappell, J.; Noel, J. P. *Science* **1997**, *277*, 1815.
(69) Aleshin, A. E.; Hoffman, C.; Firsov, L. M.; Honzatko, R. B. *J. Mol. Biol.* **1994**, *238*, 575.
(70) Juy, M.; Amrt, A. G.; Alzari, P. M.; Poljak, R. J.; Claeyssens, M.; Béguin, P.; Aubert, J.-P. *Nature* **1992**, *357*, 89.
(71) Rising, K. A.; Starks, C. M.; Noel, J. P.; Chappell, J. *J. Am. Chem. Soc.* **2000**, *122*, 1861.
(72) Colby, S. M.; Crock, J.; Dowdle-Rizzo, B.; Lemaux, P. G.; Croteau, R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2216.
(73) Back, K.; Chappell, J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6841.
(74) Cane, D. E.; Prabhakaran, P. C.; Salaski, E. J.; Harrison, P. H. M.; Noguchi, H.; Rawlings, B. J. *J. Am. Chem. Soc.* **1989**, *111*, 8914.
(75) Cane, D. E.; Prabhakaran, P. C.; Oliver, J. S.; McIlwaine, D. B. *J. Am. Chem. Soc.* **1990**, *112*, 3209.
(76) Cane, D. E.; Tsantrizos, Y. S. *J. Am. Chem. Soc*. **1996**, *118*, 10037.
(77) Cane, D. E.; Bryant, C. *J. Am. Chem. Soc.* **1994**, *116*, 12063.
(78) Hohn, T. M.; Plattner, R. D. *Arch. Biochem. Biophys.* **1989**, *272*, 137.
(79) Proctor, R. H.; Hohn, T. M. *J. Biol. Chem.* **1993**, *268*, 4543.
(80) Cane, D. E.; Wu, Z.; Proctor, R. H.; Hohn, T. M. *Arch. Biochem. Biophys.* **1993**, *304*, 415.
(81) Caruthers, J. M.; Kang, I.; Rynkiewicz, M. J.; Cane, D. E.; Christianson, D. W. *J. Biol. Chem.* **2000**, *275*, 25533.
(82) Shishova, E. Y.; Di Costanzo, L.; Cane, D. E.; Christianson, D. W. *J. Biol. Chem.*, submitted for publication.
(83) Needham, J. V.; Chen, T. Y.; Falke, J. J. *Biochemistry* **1993**, *32*, 3363.

(84) Deligeorgopoulou, A.; Taylor, S. E.; Forcat, S.; Allemann, R. K. *Chem. Commun.* **2003**, 2162.
(85) Calvert, M. J.; Ashton, P. R.; Allemann, R. K. *J. Am. Chem. Soc.* **2002**, *124*, 11636.
(86) Deligeorgopoulou, A.; Allemann, R. K. *Biochemistry* **2003**, *42*, 7741.
(87) Felicetti, B.; Cane, D. E. *J. Am. Chem. Soc.* **2004**, *126*, 7212.
(88) Cane, D. E.; Tillman, A. M. *J. Am. Chem. Soc.* **1983**, *105*, 122.
(89) Cane, D. E.; Abell, C.; Tillman, A. M. *Bioorg. Chem.* **1984**, *12*, 312.
(90) Seto, H.; Yonehara, H. *J. Antibiot.* **1980**, *33*, 92.
(91) Cane, D. E.; Oliver, J. S.; Harrison, P. H. M.; Abell, C.; Hubbard, B. R.; Kane, C. T.; Lattman, R. *J. Am. Chem. Soc.* **1990**, *112*, 4513.
(92) Cane, D. E.; Weiner, S. W. *Can. J. Chem.* **1994**, *72*, 118.
(93) Cane, D. E.; Abell, C.; Harrison, P. H. M.; Hubbard, B. R.; Kane, C. T.; Lattman, R.; Oliver, J. S.; Weiner, S. W. *Philos. Trans. R. Soc. London, Ser. B* **1991**, *332*, 123.
(94) Cane, D. E.; Sohng, J. K.; Lamberson, C. R.; Rudnicki, S. M.; Wu, Z.; Lloyd, M. D.; Oliver, J. S.; Hubbard, B. R. *Biochemistry* **1994**, *33*, 5846.
(95) Lesburg, C. A.; Lloyd, M. D.; Cane, D. E.; Christianson, D. W. *Protein Sci.* **1995**, *4*, 2436.
(96) Lesburg, C. A.; Zhai, G.; Cane, D. E.; Christianson, D. W. *Science* **1997**, *277*, 1820.
(97) Seemann, M.; Zhai, G.; de Kraker, J.-W.; Paschall, C. M.; Christianson, D. W.; Cane, D. E. *J. Am. Chem. Soc.* **2002**, *124*, 7681.
(98) Seemann, M.; Zhai, G.; Umezawa, K.; Cane, D. *J. Am. Chem. Soc.* **1999**, *121*, 591.
(99) Bohlmann, J.; Meyer-Gauen, G.; Croteau, R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4126.
(100) Trapp, S. C.; Croteau, R. B. *Genetics* **2001**, *158*, 811.
(101) Peters, R. J.; Carter, O. A.; Zhang, Y.; Matthews, B. W.; Croteau, R. B. *Biochemistry* **2003**, *42*, 2700.
(102) Phillips, M. A.; Croteau, R. B. *Trends Plant Sci.* **1999**, *4*, 184.
(103) Raffa, K. F.; Berryman, A. A.; Simasko, J.; Teal, W.; Wong, B. L. *Environ. Entomol.* **1985**, *14*, 552.
(104) LaFever, R. E.; Vogel, B. S.; Croteau, R. *Arch. Biochem. Biophys.* **1994**, *313*, 139.
(105) Vogel, B. S.; Wildung, M. R.; Vogel, G.; Croteau, R. *J. Biol. Chem.* **1996**, *271*, 23262.
(106) Ravn, M. M.; Coates, R. M.; Jetter, R.; Croteau, R. B. *Chem. Commun.* **1998**, 21.
(107) Ravn, M. M.; Coates, R. M.; Flory, J. E.; Peters, R. J.; Croteau, R. *Org. Lett.* **2000**, *2*, 573.
(108) Ravn, M. M.; Peters, R. J.; Coates, R. M.; Croteau, R. *J. Am. Chem. Soc.* **2002**, *124*, 6998.
(109) Peters, R. J.; Flory, J. E.; Jetter, R.; Ravn, M. M.; Lee, H.-J.; Coates, R. M.; Croteau, R. B. *Biochemistry* **2000**, *39*, 15592.
(110) Peters, R. J.; Ravn, M. M.; Coates, R. M.; Croteau, R. B. *J. Am. Chem. Soc.* **2001**, *123*, 8974.
(111) Ravn, M. M.; Jin, Q.; Coates, R. M. *Eur. J. Org. Chem.* **2000**, 1401.
(112) Peters, R. J.; Croteau, R. B. *Biochemistry* **2002**, *41*, 1836.
(113) Peters, R. J.; Croteau, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 580.
(114) Wani, M. C.; Taylor, H. L.; Wall, M. E.; Coggon, P.; McPhail, A. T. *J. Am. Chem. Soc.* **1971**, *93*, 2325.
(115) Horwitz, S. B. *Trends Pharmacol. Sci.* **1992**, *13*, 134.
(116) Hamel, E. *Med. Res. Rev.* **1996**, *16*, 207.
(117) Goldspiel, B. R. *Pharmacotherapy* **1997**, *17*, 110S.
(118) Zhang, B.; Maiti, A.; Shively, S.; Lakhani, F.; McDonald-Jones, G.; Bruce, J.; Lee, E. B.; Xie, S. X.; Joyce, S.; Li, C.; Toleikis, P. M.; Lee, V. M.-Y.; Trojanowski, J. Q. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 227.
(119) Michaelis, M. L.; Ansar, S.; Chen, Y.; Reiff, E. R.; Seyb, K. I.; Himes, R. H.; Audus, K. L.; Georg, G. I. *J. Pharmacol. Exp. Ther.* **2005**, *312*, 659.
(120) Michaelis, M. L.; Ranciat, N.; Chen, Y.; Bechtel, M.; Ragan, R.; Hepperle, M.; Liu, Y.; Georg, G. *J. Neurochem.* **1998**, *70*, 1623.
(121) Hezari, M.; Lewis, N. G.; Croteau, R. *Arch. Biochem. Biophys.* **1995**, *322*, 437.
(122) Wildung, M. R.; Croteau, R. *J. Biol. Chem.* **1996**, *271*, 9201.
(123) Williams, D. C.; Wildung, M. R.; Jin, A. Q.; Dalal, D.; Oliver, J. S.; Coates, R. M.; Croteau, R. *Arch. Biochem. Biophys.* **2000**, *379*, 137.
(124) Lin, X.; Hezari, M.; Koepp, A. E.; Floss, H. G.; Croteau, R. *Biochemistry* **1996**, *35*, 2968.
(125) Jin, Q.; Williams, D. C.; Hezari, M.; Croteau, R.; Coates, R. M. *J. Org. Chem.* **2005**, *70*, 4667.
(126) Williams, D. C.; Carroll, B. J.; Jin, Q.; Rithner, C. D.; Lenger, S. R.; Floss, H. G.; Coates, R. M.; Williams, R. M.; Croteau, R. *Chem. Biol.* **2000**, *7*, 969.
(127) Jin, Y.; Williams, D. C.; Croteau, R.; Coates, R. M. *J. Am. Chem. Soc.* **2005**, *127*, 7834.
(128) Gomez, G.; Di Costanzo, L.; Williams, D. C.; Croteau, R. B.; Christianson, D. W. Study in progress.
(129) Wendt, K. U.; Poralla, K.; Schulz, G. E. *Science* **1997**, *277*, 1811.
(130) Wendt, K. U.; Lenhart, A.; Schulz, G. E. *J. Mol. Biol.* **1999**, *286*, 175.
(131) Thoma, R.; Schulz-Gasch, T.; D'Arcy, B.; Benz, J.; Aebi, J.; Dehmlow, H.; Hennig, M.; Stihle, M.; Ruf, A. *Nature* **2004**, *432*, 118.
(132) Reinert, D. J.; Balliano, G.; Schulz, G. E. *Chem. Biol.* **2004**, *11*, 121.
(133) Abe, I.; Prestwich, G. D. *Lipids* **1995**, *30*, 231.
(134) Feil, C.; Süssmuth, R.; Jung, G.; Poralla, K. *Eur. J. Biochem.* **1996**, *242*, 51.
(135) Pale-Grosdemange, C.; Feil, C.; Rohmer, M.; Poralla, K. *Angew. Chem., Int. Ed.* **1998**, *37*, 2237.
(136) Hoshino, T.; Sato, T. *Chem. Commun.* **2002**, *4*, 291.
(137) Ruf, A.; Müller, F.; D'Arcy, B.; Stihle, M.; Kusznir, E.; Handschin, C.; Morand, O. H.; Thoma, R. *Biochem. Biophys. Res. Commun.* **2004**, *315*, 247.
(138) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1277.
(139) Corey, E. J.; Cheng, H.; Baker, C. H.; Matsuda, S. P. T.; Li, D.; Song, X. *J. Am. Chem. Soc.* **1997**, *119*, 1289.
(140) Li, T.; Janda, K. D.; Ashley, J. A.; Lerner, R. A. *Science* **1994**, *264*, 1289.
(141) Li, T.; Hilton, S.; Janda, K. D. *J. Am. Chem. Soc.* **1995**, *117*, 3308.
(142) Li, T.; Janda, K. D.; Lerner, R. A. *Nature* **1996**, *379*, 326.
(143) Hasserodt, J.; Janda, K. D.; Lerner, R. A. *J. Am. Chem. Soc.* **1996**, *118*, 11654.
(144) Hasserodt, J.; Janda, K. D.; Lerner, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 5993.
(145) Li, T.; Lerner, R. A.; Janda, K. D. *Acc. Chem. Res.* **1997**, *30*, 115.
(146) Hasserodt, J.; Janda, K. D.; Lerner, R. A. *J. Am. Chem. Soc.* **2000**, *122*, 40.
(147) Paschall, C. M.; Hasserodt, J.; Jones, T.; Lerner, R. A.; Janda, K. D.; Christianson, D. W. *Angew. Chem., Int. Ed.* **1999**, *38*, 1743.
(148) Zhu, X.; Heine, A.; Monnat, F.; Houk, K. N.; Janda, K. D.; Wilson, I. A. *J. Mol. Biol.* **2003**, *329*, 69.
(149) Paschall, C. M. Ph.D. Dissertation, University of Pennsylvania, Philadelphia, PA, 2000.
(150) Janda, K. D.; Weinhouse, M. I.; Schloeder, D. M.; Lerner, R. A.; Benkovic, S. J. *J. Am. Chem. Soc.* **1990**, *112*, 1274.
(151) Janda, K. D.; Weinhouse, M. I.; Danon, T.; Pacelli, K. A.; Schloeder, D. M. *J. Am. Chem. Soc.* **1991**, *113*, 5427.
(152) Tramontano, A.; Janda, K. D.; Lerner, R. A. *Science* **1986**, *234*, 1566.
(153) Pollack, S. J.; Jacobs, J. W.; Schultz, P. G. *Science* **1986**, *234*, 1570.
(154) Wendt, K. U. *Angew Chem., Int. Ed.* **2005**, *44*, 3966.
(155) Li, B.; Yu, J. P. J.; Brunzelle, J. S.; Moll, G. N.; van der Donk, W. A.; Nair, S. K. *Science* **2006**, *311*, 1436.
(156) Christianson, D. W. *Science* **2006**, *311*, 1382.

CR050286W