The Plant Cell, Vol. 19: 1994–2005, June 2007, www.plantcell.org

Rational Conversion of Substrate and Product Specificity in a *Salvia* Monoterpene Synthase: Structural Insights into the Evolution of Terpene Synthase Function [W]

Sotirios C. Kampranis,[a,1] Daphne Ioannidis,[a,b,2] Alan Purvis,[b,2] Walid Mahrez,[a] Ederina Ninga,[a] Nikolaos A. Katerelos,[b] Samir Anssour,[a] Jim M. Dunwell,[b] Jörg Degenhardt,[c] Antonios M. Makris,[a] Peter W. Goodenough,[b] and Christopher B. Johnson[a]

[a] Department of Natural Products and Biotechnology, Mediterranean Agronomic Institute of Chania, 73100 Chania, Greece
[b] School of Biological Sciences, University of Reading, Whiteknights, Reading, Berkshire RG6 6AS, United Kingdom
[c] Max Planck Institute for Chemical Ecology, D-07745 Jena, Germany

Terpene synthases are responsible for the biosynthesis of the complex chemical defense arsenal of plants and microorganisms. How do these enzymes, which all appear to share a common terpene synthase fold, specify the many different products made almost entirely from one of only three substrates? Elucidation of the structure of 1,8-cineole synthase from *Salvia fruticosa* (Sf-CinS1) combined with analysis of functional and phylogenetic relationships of enzymes within *Salvia* species identified active-site residues responsible for product specificity. Thus, Sf-CinS1 was successfully converted to a sabinene synthase with a minimum number of rationally predicted substitutions, while identification of the Asn side chain essential for water activation introduced 1,8-cineole and α-terpineol activity to *Salvia pomifera* sabinene synthase. A major contribution to product specificity in Sf-CinS1 appears to come from a local deformation within one of the helices forming the active site. This deformation is observed in all other mono- or sesquiterpene structures available, pointing to a conserved mechanism. Moreover, a single amino acid substitution enlarged the active-site cavity enough to accommodate the larger farnesyl pyrophosphate substrate and led to the efficient synthesis of sesquiterpenes, while alternate single substitutions of this critical amino acid yielded five additional terpene synthases.

INTRODUCTION

Terpenoids, of which nearly 40,000 have now been identified, are major contributors to the chemical arsenal of plants and microorganisms, with profound roles in the defense against enemies, the attraction of pollinators, and signaling to other plants. With a constant war raging between prey and predator, the chemical profile of terpenoids produced by an organism has to evolve and adapt rapidly in what has been called a coevolutionary arms race (Ehrich and Raven, 1964; Harborne, 1993). Thus, terpene synthases, the key enzymes of terpene biosynthesis, provide an attractive model to study the evolution of enzyme function.

The reaction mechanism of terpene synthases begins with the divalent cation-dependent ionization of the prenyl-diphosphate substrate (geranyl pyrophosphate [GPP; C_{10}], farnesyl pyrophosphate [FPP; C_{15}], or geranyl geranyl pyrophosphate [C_{20}] in mono-, sesqui-, and diterpene synthases respectively). The reaction proceeds via the formation of carbocationic intermediates to the synthesis of a large array of structurally diverse products (Figure 1) (Croteau, 1987; Tholl, 2006). This chemical complexity is achieved via a common scaffold, the terpene synthase fold, which is highly conserved from fungi to plants and between mono-, sesqui-, and diterpene synthases (Lesburg et al., 1997; Starks et al., 1997; Caruthers et al., 2000; Rynkiewicz et al., 2001; Christianson, 2006). However, apart from the presence of a short conserved motif related to metal ion binding (DDxxD), the sequence similarities between terpene synthases are dominated by species relationships regardless of substrate or product specificity (Bohlmann et al., 1998). Thus, the gymnosperm synthases represent a discrete group distinct from that of the angiosperms, while within the latter, the sequences related to similar species again group together much more than enzymes of similar function. The information to date suggests that evolution of terpene synthases has been rapid and predominantly divergent but compounded by instances of convergent evolution (Trapp and Croteau, 2001; Aubourg et al., 2002; Sharkey et al., 2005). This has given rise to suggestions that sequence-based identification of terpene synthase function based on even quite closely related species is not likely to succeed and has presented a challenge for sequence-based approaches to elucidate the molecular determinants of product specificity.

Attempts so far to identify functionally important regions of monoterpene synthase genes have involved domain swapping or site-directed mutagenesis. Domain-swapping experiments with synthases of *Salvia officinalis* were important in establishing

[1] To whom correspondence should be addressed. E-mail sotirios@maich.gr; fax 30-28210-35001.
[2] These authors contributed equally to this work.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Sotirios C. Kampranis (sotirios@maich.gr).
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.106.047779

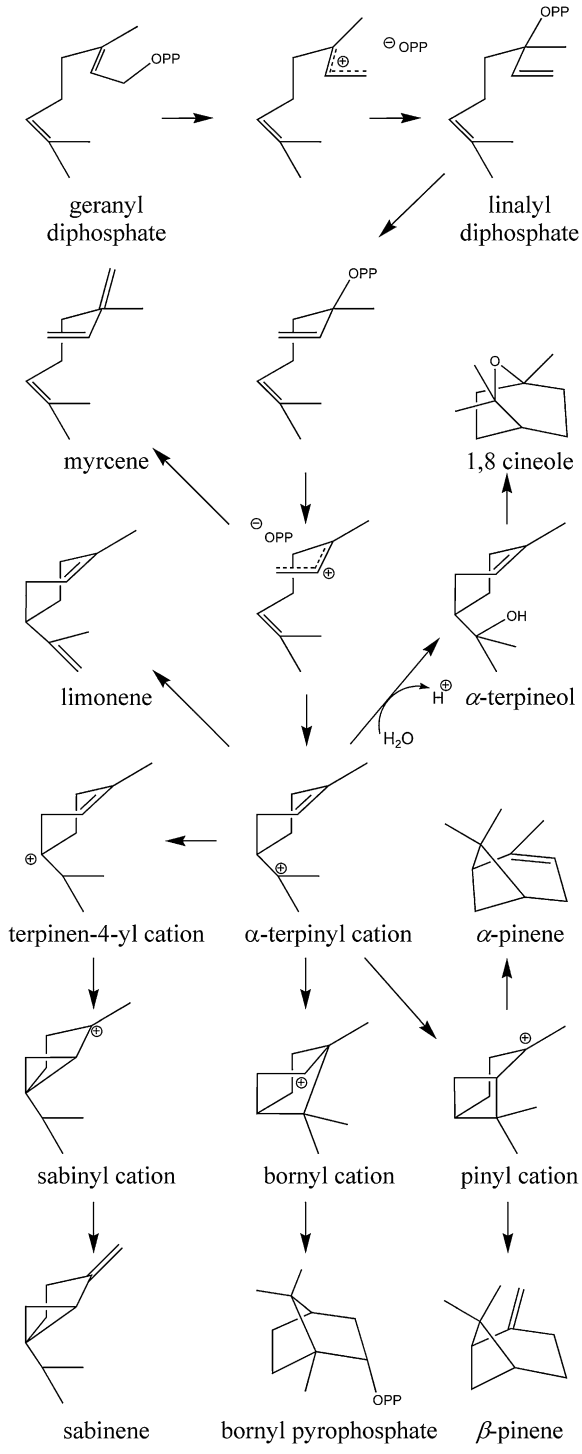

Figure 1. Scheme for Product Formation in Sf-CinS1 and Its Mutants.

that only the C-terminal region plays a role in determining product specificity (Peters and Croteau, 2003). However, substitution of shorter segments within this region was at best partially effective in switching the product spectrum. A similar approach between *Citrus limon* β-pinene synthase and γ-terpinene synthase showed that within the C-terminal domain of these monoterpene synthases, a region comprising 200 amino acids, of which 41 were different, is responsible for determining product specificity (El Tamer et al., 2003). A structural modeling-based approach with the highly homologous pinene synthase and camphene synthase from *Abies grandis* was only partially successful in converting the pinene synthase to a camphene synthase even when 12 different mutations were combined (Hyatt and Croteau, 2005). One reason for this must be the lack of availability of a precise structure for either of the two enzymes, and the other must be a lack of sequences of enzymes of similar function from closely related species. Despite high sequence homology (82%), there were >50 residues in the C-terminal region that could have been potential targets for mutagenesis and no strong rationale for eliminating unimportant or essentially conserved amino acids from the selection. A further degree of complexity came from a recent report by Greenhagen et al. (2006) on a pair of sesquiterpene synthases. Tobacco (*Nicotiana tabacum*) 5-*epi*-aristolochene synthase (TEAS) and *Hyoscyamus muticus* premnaspirodiene synthase (HPS) are two plant sesquiterpene synthases that share 72% amino acid identity. Although these two enzymes make different products, molecular modeling of HPS using the structural coordinates for TEAS indicated that those residues in immediate contact with the substrate FPP were nearly identical. Mutation of residues farther away from the active site was necessary for the interconversion of substrate specificity between the two enzymes, showing that catalytic specificity in terpene synthases may also be modulated by distant structural features. Consequently, the same combination of active-site residues may not lead to the same product spectrum when not combined with very similar scaffolds.

Recently, Yoshikuni et al. (2006) performed exhaustive site-directed mutagenesis to 19 residues indicated by homology modeling to surround the active site of the sesquiterpene γ-humulene synthase. This enabled the identification of several plasticity residues that were systematically recombined, on the basis of a mathematical model, so as to construct novel terpene synthases. Using this approach, seven different specific and active synthases were successfully constructed, demonstrating the feasibility of exploiting the underlying evolvability of the terpene scaffold and providing evidence that protein engineering approaches can be successfully applied to the design of terpene synthase function.

In the literature, sequences of enzymes of similar function are rare and virtually nonexistent for species sufficiently closely related to one another to be useful for primary structure comparisons. There is, perhaps understandably, a similar lack of crystal structures for two closely related proteins in the same or closely related species. It seemed to us that an approach combining precise structural information with amino acid sequence comparisons between enzymes of similar activity from closely related species would greatly aid our understanding of terpene synthase specificity. Such an approach would enable us, first, to distinguish those amino acids that differ between different enzymes from those shared and, second, to highlight residues that are common between two enzymes of similar activity and that may be responsible for product specificity. This made a potent combination that we have successfully exploited here to predict

the molecular determinants of substrate and product specificity in *Salvia* monoterpene synthases.

We have solved the structure of 1,8-cineole synthase from *Salvia fruticosa* and employed a primary structure comparison between two 1,8-cineole synthases (one from *S. officinalis* [So-CinS1] and one from *S. fruticosa* [Sf-CinS1]), two sabinene synthases (from *S. officinalis* [So-SabS1] and *S. pomifera* [Sp-SabS1]), and *S. officinalis* bornyl pyrophosphate synthase (So-BPPS) (Wise et al., 1998) (whose three-dimensional structure is also available [Whittington et al., 2002]) to identify amino acids located in the active-site region that might be functionally significant. Using this approach, we have asked the following questions. (1) What are the main structural elements contributing to substrate and product specificity in a synthase? (2) Can we easily and predictably change the products formed? (3) Could this give us clues as to how these enzymes have evolved?

RESULTS

Structure of Sf-CinS1

The structure of Sf-CinS1 was solved at 1.95-Å resolution (a summary of crystal statistics is given in Table 1). The secondary structure of Sf-CinS1 is that of an α-helical protein comprised of 23 α-helices and eight 3_{10} helices (Figure 2). These are arranged with remarkable similarity to So-BPPS (PDB: 1N1B) and tobacco TEAS (PDB: 5EAS) with Cα RMS deviation of 1.56 and 2.17 Å, respectively (>484 residues compared with 1N1B and >474 residues compared with 5EAS). (Cα RMS deviation, a common statistical measure of the differences between two structures, is the root mean squared deviation of the distance between two corresponding atoms when these two structures are superimposed. Here, we are calculating the root mean squared deviation between α-carbons [Cα RMS]. A very small number [i.e., 1 Å or below] means that two structures are exceptionally similar; larger numbers indicate structures that are more different.) Helices were labeled from 2 to 4 and 6 to 25 in accordance with the So-BPPS structure. The small α-helix (α1), which lies over the α11 and α12 helices near the active site of So-BPPS, is absent in the Sf-CinS1 structure due to some disorder of the N-terminal region (helix α1 includes residues 67 to 71 of So-BPPS and 69 to 73 in Sf-CinS1), while Sf-CinS1 also lacks helix α5 of So-BPPS due to a five-residue deletion between residues 141 and 142 (also the case in TEAS). The Sf-CinS1 monomer is split into two α-helical domains (Figure 2). The N-terminal domain consists of eight α-helices (numbered 2 to 4 and 6 to 9) arranged in an α-α barrel and includes residues up to and including Leu-270. The domain is well defined, with only minor structural differences in comparison to the So-BPPS and TEAS structures. The larger C-terminal domain consists of 15 α-helices and two 3_{10} helices arranged in an orthogonal bundle. The domain is well conserved with a Cα RMS fit of 1.32 Å compared with So-BPPS and 1.92 Å compared with TEAS. TEAS shows some significant differences within the C-terminal region compared with the two monoterpene synthases, where minor shifts in the nearby α17, α19, and α20 helices result in a significant 2.4-Å shift of the α21 helix away from the active site, thereby increasing the size of the cavity.

The metal ion binding DDxxD (345 to 349) motif, common to all terpene synthases, is located on the C-terminal region of helix α14. The active site is located within a large cavity created between six α-helices (numbers 13, 14, 18, 19, 21, and 24) arranged with α-α barrel architecture and is capped by the α24-α25 loop, which includes a short 3_{10} helix that preceeds the α24 helix. The active site is remarkably well conserved with a Cα RMS fit of 0.99 Å compared with So-BPPS.

Table 1. Data Collection and Refinement Statistics (Molecular Replacement)

Data Collection	
Space Group	C2221
Cell dimensions	
a, *b*, *c* (Å)	124.55, 171.15, 123.81
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	30.0 (1.95)[a]
R_{merge}	5.5 (30.1)
$I/\sigma I$	14.3 (4.2)
Completeness (%)	99.6 (98.8)
Redundancy	3.7 (3.7)
Refinement	
Resolution (Å)	1.95
No. measured reflections	354,544
No. unique reflections	95,749
R_{work}/R_{free}	21.8 / 23.5
No. atoms	
Protein	8120
Ligand/ion	16
Water	459
B-factors	
Protein (Å^2)	32.2
Ligand/ion (Å^2)	55.1
Water (Å^2)	35.6
RMS deviations	
Bond lengths (Å)	0.026
Bond angles (°)	2.2

[a] Highest-resolution shell of 1.95 to 2.06 Å is shown in parenthesis.

Deciphering the Molecular Determinants of Product Specificity

Monoterpene synthases initiate their respective reaction sequences by Mg^{2+}-dependent ionization of GPP, isomerization to linalyl diphosphate, and reionization with cyclization to the α-terpinyl cation (Figure 1). This intermediate can have different fates. In 1,8-cineole synthase, water capture of the α-terpinyl cation yields α-terpineol, which undergoes protonation of the endocyclic double bond and internal addition to produce 1,8-cineole (Croteau, 1987; Croteau et al., 1994). Examination of the Sf-CinS1 active-site region compared with the So-BPPS crystal structure and the amino acid alignment of the different *Salvia* monoterpene synthases (Figure 3) revealed a number of residues that were conserved between the different synthases and others that varied and could play a role in product specificity. The conserved residues are Trp-317, Ile-337, Thr-342, Tyr-420, Ser-445, Ile-451, Leu-485, and Tyr-564 (Figure 2 and denoted by

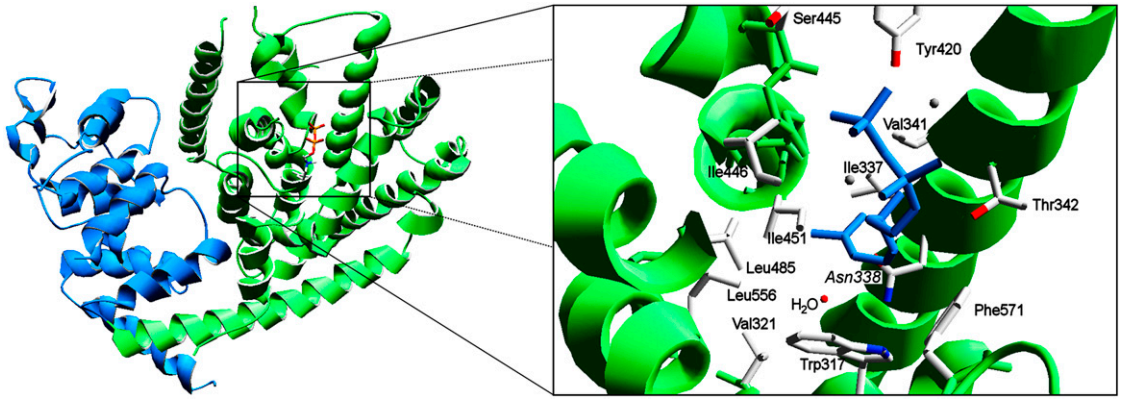

Figure 2. The Structure of Sf-CinS1.

The N-terminal domain is shown in blue and the C-terminal domain in green. The inset shows the active-site region of Sf-CinS1 with the residues forming the active-site contour indicated together with the water molecule at Asn-338. The structure of So-BPPS (PDB: 1N23) is superimposed on that of Sf-CinS1, and the 3-aza-2,3-dihydrogeranyl diphosphate ligand of the former is shown. The image was produced by DeepView (Swiss-PdbViewer).

an asterisk in Figure 3). The variable residues appear to be clustered in two regions. Region 1 is located on helix α14 at the bottom of the active-site cavity and comprises residues 338 to 341 (Figure 3). This region includes Asn-338, whose side chain appears to be hydrogen bonding to a water molecule at the active site (Figure 4) and is conserved between the *S. officinalis* and *S. fruticosa* cineole synthases. The corresponding residue is also conserved in the *Arabidopsis thaliana* 1,8-cineole synthase (Chen et al., 2004), suggesting that this Asn may be critical for the cineole synthase catalytic mechanism and that it is most likely involved in the deprotonation of the water molecule facilitating the attack on the α-terpinyl cation that results in the formation of α-terpineol (Figure 1).

Region 2 comprises residues 446 to 450, which are part of the loop connecting helices α18 and α19 (Figure 3). Disruption of the hydrogen bonding network by the presence of a Pro at position 450 contributes to the formation of a kink between helices α18 and α19 and the exposure of the carbonyl oxygens of Ile-446 and Gly-447 to the active-site cavity. It has been proposed that an important aspect of the catalytic mechanism of terpene synthases may be the stabilization of the unstable carbocationic intermediates by local partial charges in the catalytic pocket (Lesburg et al., 1997; Starks et al., 1997; Rynkiewicz et al., 2001). It is therefore possible that the observed deformation of helix α18 is an essential structural feature of these enzymes. In support of this, a similar deformation is observed in the structure of

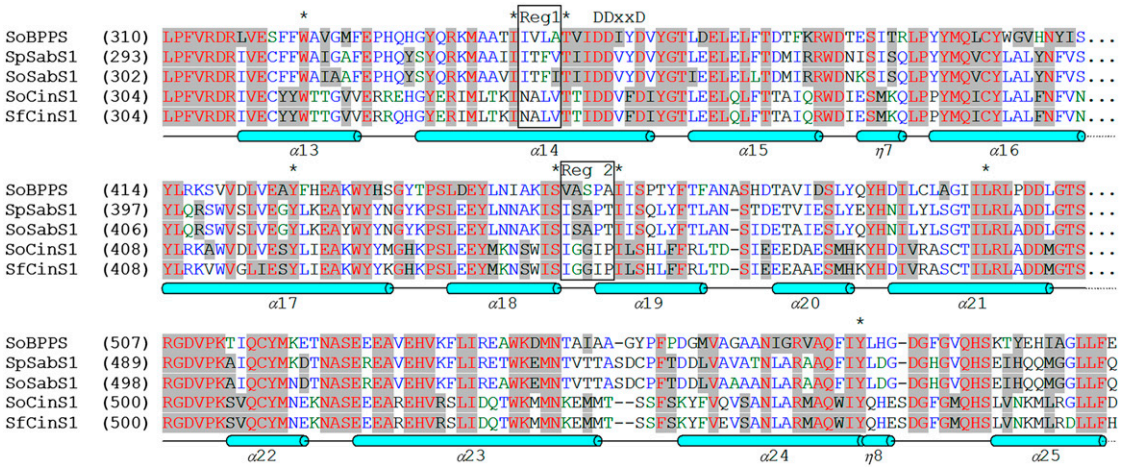

Figure 3. Amino Acid Sequence Alignment of the C-Terminal Domain of Monoterpene Synthases from *Salvia* Species.

Alignment of So-BPPS, So-SabS1, and So-CinS1 (Wise et al., 1998) with Sf-CinS1 and Sp-SabS1. Residues lining the active site that are common between these enzymes are denoted by an asterisk, while the two boxes highlight the two regions of the active site where variability is observed. The conserved DDxxD motif is also shown, together with a diagrammatic representation of the secondary structure of Sf-CinS1 (cylinders labeled with α represent α-helices, while 3_{10} helices are indicated with an η).

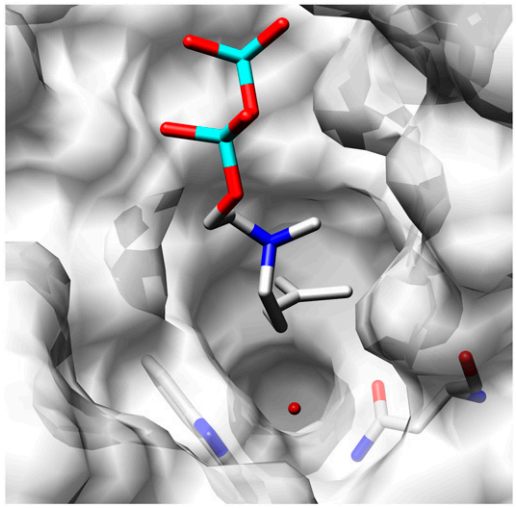

Figure 4. Asn-338, Trp-317, and the Active Site of Sf-CinS1.

The water molecule (oxygen in red) likely involved in the hydroxylation of the α-terpinyl cation. The structure of So-BPPS (PDB: 1N23) is superimposed on that of Sf-CinS1, and the 3-aza-2,3-dihydrogeranyl diphosphate ligand of the former is shown. The image was produced by UCSF Chimera.

S. officinalis BPPS (Whittington et al., 2002) supported by the presence of another Pro residue, this time Pro-455 (Figures 3 and 5A). Notably, when the two structures are superimposed, the position of the disrupting Pro in BPPS differs by one amino acid from that of Sf-CinS1, possibly contributing to the conformation of the resulting loop and the orientation of the exposed carbonyls (Figure 5A). Small alterations in the structure of this region may have a significant impact in the product specificity of the terpene synthases. In sesquiterpene synthases, a kink also appears to be present in helix G of tobacco TEAS (Starks et al., 1997) (Figure 5B; PDB: 5EAS) and the corresponding helices of *Streptomyces* pentalenene synthase (Lesburg et al., 1997) (PDB: 1PS1), *Fusarium sporotrichoides* trichodiene synthase (Rynkiewicz et al., 2001) (PDB: 1JFA), and *Penicillium roqueforti* aristolochene synthase (Caruthers et al., 2000) (PDB: 1DI1), further supporting the conservation of this deformation and its potential functional importance in both mono- and sesquiterpene synthases.

Conversion of Sf-CinS1 to a Sabinene Synthase

To address the relation between the variability observed in these two regions and the products formed by terpene synthases, we attempted the conversion of Sf-CinS1 to a sabinene synthase based on an amino acid comparison of this enzyme with So-SabS1 and Sp-SabS1. In sabinene synthase, the α-terpinyl cation does not undergo water addition, as is the case in cineole synthase. Instead, a 1,2-hydride shift within the α-terpinyl cation, followed by secondary closure to the cyclopropane ring and deprotonation from the methyl group, provides sabinene (Figure 1). Thus, the attempted conversion from cineole to sabinene synthase was initiated by the most obvious substitution, that of Asn-338 to Ile (the corresponding residue in Sp-SabS1 and So-SabS1), so as to abolish water capture and to direct the reaction toward the steps leading to sabinene synthesis (Figure 1). Whereas in Sf-CinS1, 1,8-cineole accounts for 72.4% of the volatiles produced, accompanied by 7.1% α-terpineol, 9.1% β-pinene, 4.6% α-pinene, 3.6% sabinene, 2.2% myrcene, and <1% limonene ($k_{cat} = 3.18 \pm 0.3$ min^{-1}, $K_m = 65.4 \pm 18.4$ μM), the N338I mutation resulted in the production of 48.3% sabinene and 37% limonene but no α-terpineol or 1,8-cineole ($k_{cat} = 5.7 \pm 1.4$ min^{-1}, $K_m = 56.6 \pm 26.6$ μM; Figure 6, Table 2). This clearly demonstrates the importance of Asn-338 in the activation of the water molecule and the hydroxylation of the α-terpinyl cation. The production of significant amounts of limonene is the result of the failure to hydroxylate the α-terpinyl cation, which is instead converted to limonene by proton elimination (Figure 1). The residue next to Asn-338 in Sf-CinS1 is Ala-339, but in Sp-SabS1 and So-SabS1, the corresponding residue is a Thr. To address whether the presence of the bulkier Thr side chain is necessary for the optimal architecture of the active site, Ala-339 was

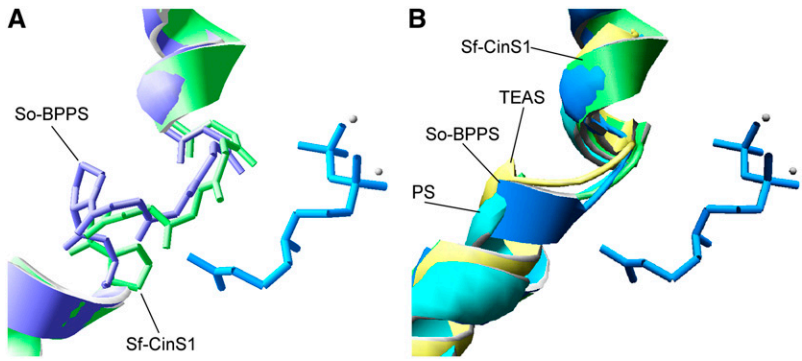

Figure 5. Structure of the Conserved Kink.

(A) Region 2 superimposition of Sf-CinS1 on So-BPPS to show the difference in the conformation of the kink in helix α18. The conformation of the backbone and the two Pro residues are shown (Sf-CinS1 in green). The 3-aza-2,3-dihydrogeranyl diphosphate ligand of So-BPPS is also shown.
(B) Superposition of Sf-CinS1 (green), So-BPPS (blue), TEAS (beige), and *Streptomyces* pentalenene synthase (PS; turquoise) to show the conservation of the kink between mono- and sesquiterpene synthases.

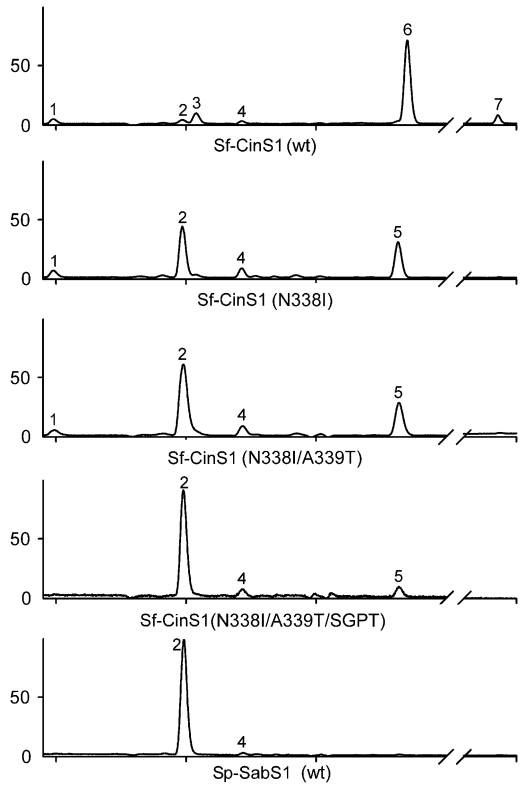

Figure 6. Gas Chromatography Traces Showing the Conversion of Sf-CinS1 to a Sabinene Synthase.

Compounds, confirmed by mass spectrometry analysis and by comparisons with standards, are as follows: α-pinene (1), sabinene (2), β-pinene (3), myrcene (4), limonene (5), 1,8-cineole (6), and α-terpineol (7). A trace showing the product distribution of wild-type Sp-SabS1 is also shown. SGPT denotes G447S/I449P/P450T.

mutated to Thr in Sf-CinS1 (N338I), resulting in an increase of sabinene production to 62.1% of total products in the double mutant (Figure 6, Table 2).

In an attempt to further increase the specificity of the mutated enzyme, we embarked on substituting the residues in region 2 with those found in common in Sp-SabS1 and So-SabS1. The first mutation introduced was that of Gly-447 to Ser to assess whether the presence of a less inert side chain in this region contributes to the isomerization steps resulting in sabinene. No increase in the level of sabinene was observed nor was any significant change in any other product (Table 2), suggesting that despite the fact that the hydroxyl group of Ser is expected to be pointing into the catalytic site, neither its chemical properties nor its size appear to affect the cyclization cascade. To assess the role of the backbone conformation in region 2, we attempted to shift the helix-breaking Pro by one as is the case in Sp-SabS1 (Figure 3). This was achieved in two steps, first by substituting Ile-449 with Pro, creating a double Pro mutant that was essentially inactive as a terpene synthase (data not shown), and then by removing the second Pro (Pro-450) and replacing it with Thr (the corresponding residue in Sp-SabS1). This not only restored enzyme activity but also resulted in an enzyme that produced 86.8% sabinene ($k_{cat} = 0.26 \pm 0.02$ min^{-1}, $K_m = 12.9 \pm 2.8$ μM), clearly confirming the role of kink conformation in product specificity (Figure 6, Table 2).

To further confirm the role of the conserved Asn in cineole synthesis, we attempted the reverse conversion (i.e., to introduce cineole synthase capacity into Sp-SabS1). Wild-type Sp-SabS1 makes almost 100% sabinene with a trace of myrcene. Substitution of Ile-327 with Asn resulted in the production of 7.7% α-terpineol and 2.4% cineole (Table 2), suggesting that a small but detectable level of hydroxylation of the α-terpinyl cation is afforded by the introduced Asn and that despite the possible lack of a dedicated cavity for the activated water molecule, like the one present in the Sf-CinS1 structure (Figure 4), there is some space in the active-site cavity to accommodate it. Further substitution of Thr-328 to Ala appears to improve active-site architecture, resulting in an enzyme that produces 10.9% α-terpineol and 3.4% cineole, a total of almost 15% hydroxylated products (Table 2). Alterations in region 2 and in particular replacement of Ser-436 with Gly further increased production of α-terpineol to 18.1% but abolished production of cineole (Table 2). Likely, the presence of Gly at this position and the effect it might have on the backbone conformation may be detrimental for the subsequent protonation of the endocyclic double bond and the internal addition steps that give rise to cineole. The low levels of hydroxylation observed, compared with wild-type cineole synthase, may be attributed to the fact that the sabinene synthase active site cannot be easily made to efficiently accommodate the water molecule required. It will be difficult to introduce a dedicated cavity in Sp-SabS1, like the one present in the cineole synthase structure (Figure 4), without precise structural information specific to the sabinene synthase.

Conversion of Sf-CinS1 to a Sesquiterpene Synthase

With very few exceptions, monoterpene synthases use GPP as a substrate, and only a few have been reported to be able to also use the five-carbon-longer FPP, thus catalyzing the formation of sesquiterpenes (Schnee et al., 2002; Aharoni et al., 2004). Sf-CinS1 does not make sesquiterpenes when presented with FPP as substrate. Enlargement of the active-site cavity by mutation of the essential Asn to Ala was employed in an attempt to enable Sf-CinS1 to accommodate the FPP substrate into its active-site cavity. This resulted in the formation of 49% *trans*-α-bergamotene together with 16.8% *E*-β-farnesene, 15.1% β-selinene, 9.2% β-bisabolene, 6% β-sesquiphellandrene, 4% *cis*-α-bergamotene, and 2.6% *Z*-β-farnesene from FPP (Figure 7, Table 3), with a moderate increase in the K_m ($k_{cat}^{FPP} = 0.14 \pm 0.04$ min^{-1}, $K_m^{FPP} = 72.7 \pm 6.9$ μM; Table 3). Molecular modeling of the active site of this mutant revealed that removal of the larger Asn side chain and the concomitant loss of the associated water allow enough additional space to accommodate the five-carbon-longer substrate in an extended conformation (Figure 7). Such a conformation is consistent with the structure of the products formed. The structural analogy between pinene and bergamotene, limonene and bisabolene, or myrcene and farnesene suggests that cyclization likely proceeds mainly in the upper part of the active-site cavity, similar to the reaction of the wild-type enzyme with GPP, while the additional five-carbon atoms of FPP

Table 2. Volatile Composition of Sf-CinS1, Sp-SabS1, and Their Mutants

	Total Monoterpenes (%)							Kinetic Parameters		
Enzyme Variant	α-Pinene	Sabinene	β-Pinene	Myrcene	Limonene	1,8-Cineole	α-Terpineol	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1} \cdot min^{-1}$)
Sf-CinS1 (wild type)	4.6	3.6	9.1	2.2	<1.0	72.4	7.1	3.18	65.4	0.049
Sf-CinS1 (N338I)	6.5	48.3	–	8.2	37.0	–	–	5.78	56.6	0.102
Sf-CinS1 (N338I/A339T)	5.4	62.1	–	6.4	26.1	–	–	1.42	135.8	0.011
Sf-CinS1 (N338I/A339T/G447S)	5.7	59.6	–	5.0	29.7	–	–	0.26	95.5	0.027
Sf-CinS1 (N338I/A339T/SGPT$^{\alpha}$)	–	86.8	–	5.5	7.7	–	–	0.26	12.9	0.020
Sp-SabS1 (wild type)	–	98.4	–	1.6	–	–	–			
Sp-SabS1 (I327N)	–	83.1	–	4.1	2.7	2.4	7.7			
Sp-SabS1 (I327N/T329A)	–	79.2	–	1.8	4.7	3.4	10.9			
Sp-SabS1 (I327N/T329A/S436G)	–	81.9	–	–	–	–	18.1			
Sf-CinS1 (N338I)	6.5	48.3	–	8.2	37.0	–	–	5.78	56.6	0.102
Sf-CinS1 (N338L)	4.7	5.6	13.2	27.2	49.3	–	–	0.54	26.5	0.020
Sf-CinS1 (N338A)	11.7	33.1	33.2	16.9	5.1	–	–	4.15	11.9	0.348
Sf-CinS1 (N338S)	9.8	34.5	34.2	6.8	2.6	5.6	6.5	1.30	14.4	0.090
Sf-CinS1 (N338V)	–	61.2	–	8.0	30.8	–	–	0.73	55.1	0.013
Sf-CinS1 (N338C)	16.1	41.0	32.1	4.7	6.1	–	–	3.13	201.0	0.016

SGPT$^{\alpha}$ denotes G447S/I449P/P450T.

remain tightly restrained at the bottom of the cavity (Figure 7). Conversion of a monoterpene to a sesquiterpene synthase by a single amino acid alteration highlights the thin line separating the two classes of terpene synthases and provides insights on how the one may have evolved from the other.

Structural Basis for Rapid Evolution

To further examine the role of region 1 in product specificity, we mutated Asn-338 to the different amino acids present in this position in other characterized monoterpene synthases. Among 61 sequences examined, the corresponding residue was found to be Asn in 9.8%, Ile in 27.9%, Val in 19.7%, Leu in 2%, Ala in 3.5%, Ser in 13.1%, and Cys in 16.4% of the cases. All the remaining five substitutions were made and resulted in the creation of practically five novel terpene synthases. Mutation to Val results in an enzyme that produces 61.2% sabinene, 30.8% limonene, and 8% myrcene ($k_{cat} = 0.73 \pm 0.1\ min^{-1}$, $K_m = 55.1 \pm 12.6$ μM; Table 2), an enzyme almost identical in product specificity to the double mutant N338I/A339T discussed above. When the same residue was converted to Leu, limonene comprised almost 50% of the products, followed by 27% myrcene and 13.3% β-pinene ($k_{cat} = 0.54 \pm 0.1\ min^{-1}$, $K_m = 26.5 \pm 5.8$ μM; Table 2). Enlargement of the active-site cavity by the introduction of an Ala resulted not only in the ability to use FPP as substrate but also in a promiscuous monoterpene synthase producing 33.1% sabinene, 33.2% β-pinene, 16.9% myrcene, 11.7% α-pinene, and 5.1% limonene ($k_{cat} = 4.15 \pm 0.39\ min^{-1}$, $K_m = 11.9 \pm 3.5$ μM; Table 2). Cys substitution results in the formation of 41% sabinene, 32.1% β-pinene, 16.1% α-pinene, 4.7% myrcene, and 6.1% limonene from GPP (Table 2) as well as 45.8% *trans*-α-bergamotene together with 17.7% *E*-β-farnesene, 13.5% β-selinene, 10% β-bisabolene, 7% β-sesquiphellandrene, 5.9% *cis*-α-bergamotene, and 3.1% *Z*-β-farnesene from FPP (Table 3). It is noteworthy that substitution of the Asn essential for cineole synthesis with aliphatic amino acids completely abolishes both cineole and α-terpineol formation, underlining the essential role of the Asn in the activation of the water molecule. Introduction of Ser, which appears to be partially capable of coordinating the water molecule, results in the production of 5.6% cineole and 6.5% α-terpineol, together with 34.5% sabinene, 34.2% β-pinene, 9.8% α-pinene, 6.8% myrcene, and 2.6% limonene ($k_{cat} = 1.30 \pm 0.1\ min^{-1}$, $K_m = 14.4 \pm 5.0$ μM; Table 2). This enzyme is also active as a sesquiterpene synthase producing 45.1% β-bisabolene, 39.4% *E*-β-farnesene, and 15.4% β-sesquiphellandrene ($k_{cat} = 0.2 \pm 0.1\ min^{-1}$, $K_m = 31.0 \pm 89.6$ μM; Table 3). The kinetic characteristics of the mutants suggest that the mutations introduced have a significant effect in the active-site contour, in a way that both the binding of the substrate and the rate-limiting step in the reaction mechanism are affected. This is corroborated from molecular modeling analysis of the active-site contour that suggests significant alterations in the active-site surface (see Supplemental Figure 1 online). This is in agreement with a mechanism by which product specificity in terpene synthases is dictated by a combination of the contour and the dynamics of the active site in a way that some isomerization steps are facilitated, while others are hindered, and not by the precise positioning of certain reactive groups.

In these experiments, substitution of a single amino acid has resulted in the production of five additional terpene synthases, some with remarkably different product profiles. The identification of such a high plasticity residue in the Sf-CinS1 active site provides the structural basis for a potential mechanism of rapid evolution, by which a single mutation can drastically modify the

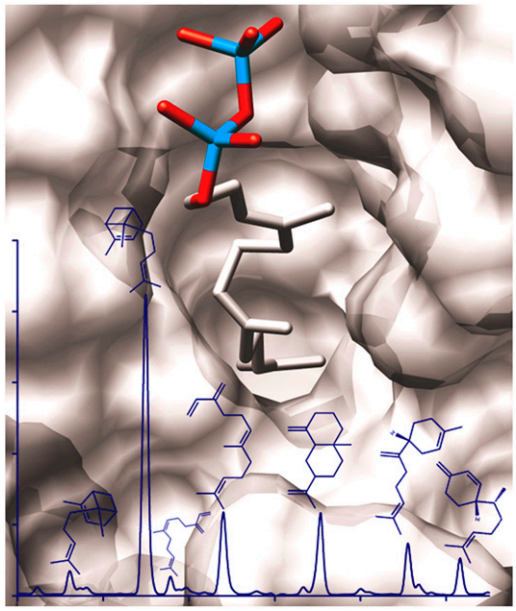

Figure 7. Gas Chromatography Trace Showing the Sesquiterpene Products Formed by Sf-CinS1 (N338A).

From left to right, *cis*-α-bergamotene, *trans*-α-bergamotene, *Z*-β-farnesene, *E*-β-farnesene, β-selinene, β-bisabolene, and β-sesquiphellandrene. In the background is a model of the active-site region of this mutant with the substrate (FPP) in place. In Sf-CinS1(N338A), cyclization appears to proceed mainly in the upper part of the active-site cavity, similar to the reaction of the wild-type enzyme with GPP, while the additional five-carbon atoms of FPP seem to remain tightly restrained at the bottom of the cavity. This is clearly evidenced by the structural analogy between pinene and bergamotene, limonene and bisabolene, or myrcene and farnesene. The image was produced by UCSF Chimera.

terpene profile produced. Analysis of an amino acid sequence alignment of 46 published monoterpene synthases from both gymnosperms and angiosperms revealed that both region 1 (N338) and region 2 fall into pronounced similarity minima. All other regions of low similarity observed correspond to either the divergent N-terminal part of the enzymes or, as indicated by the structure of Sf-CinS1, in surface-exposed residues (see Supplemental Figure 2 online). Moreover, sequence analysis of Sf-CinS1 and Sp-SabS1 indicated a high accumulation of nucleotide changes, which cause alterations of the amino acid sequence in region 2. The proportion of these nonsynonymous changes ($K_a = 1.08$) is much higher in region 2 than in the overall coding sequence ($K_a = 0.37$), suggesting a reduced amount of purifying selection in region 2 (data not shown). Evidence supporting the rapid evolution of terpene synthases comes from genomic information now available for model plants. Analysis of the *Arabidopsis* terpene synthase gene family, which contains >30 different members, clearly indicates rapid divergence of catalytic activities and tissue-specific expression patterns resulting from cycles of gene duplication and multiple mutations (Aubourg et al., 2002; Chen et al., 2004; Ro et al., 2006).

DISCUSSION

We used a rational approach based on the combination of structural information with functional and phylogenetic relationships to achieve the predictive conversion of both substrate and product specificity in a monoterpene synthase. By solving the structure of Sf-CinS1 and combining structural and sequence information from only very closely related species, we highlighted two regions in the active site responsible for product specificity. Region 1, which is located at the bottom of the active-site pocket, contains an Asn side chain that is essential for the activation and stabilization of the water molecule that attacks the α-terpinyl intermediate, and its mutation completely abolishes production of hydroxylated products. This residue appears to be a site of significant plasticity, as its alternate substitution to a number of other residues inferred by phylogenetic comparisons gives rise to enzymes with very different product spectra.

A water molecule was also found in the crystal structure of So-BPPS (Whittington et al., 2002). This water molecule (namely water 110) was found to hydrogen bond to the diphosphate group of the substrate analog, the backbone carbonyl of Ser-451 (corresponding to Ser-445 in Sf-CinS1), and the side chain of Tyr-426 (Tyr-420 in Sf-CinS1). It is located at a distinctly different position than the water associated with Asn-338, closer to the top of the active-site cavity and near the diphosphate binding region. It was postulated that water 110 could serve as a diphosphate-assisted general base to account for the generation by this enzyme of cyclic olefin coproducts (pinenes, camphene, and limonene) derived by direct deprotonation of carbocation intermediates (Whittington et al., 2002). However, no further attempt to identify its role has since been reported.

In this report, we also identify a conserved structural feature, a kink in one of the helices forming the active site, that appears to be a key determinant of product specificity. The residues present in region 2 appear to have a significant contribution to the conformation of the kink, whose alteration by the shift of a Pro residue is necessary for the complete conversion of Sf-CinS1 to a

Table 3. Percentage of Sesquiterpene Products Produced by Sf-CinS1 Mutants

	Total Sesquiterpenes (%)							Kinetic Parameters		
Enzyme Variant	*cis*-α-bergamotene	*trans*-α-bergamotene	*Z*-β-farnesene	*E*-β-farnesene	β-Selinene	β-Bisabolene	β-Sesquiphellandrene	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($\mu M^{-1} \cdot min^{-1}$)
Sf-CinS1 (N338A)	4.0	49.0	2.6	16.8	15.1	9.2	6.0	0.14	72.9	0.0019
Sf-CinS1 (N338S)	–	–	–	39.4	–	45.1	15.4	0.02	31.0	0.0008
Sf-CinS1 (N338C)	5.9	45.8	3.1	17.7	13.5	10.0	7.0	0.05	101.0	0.0005

sabinene synthase. Further support for the role of this conserved kink in substrate selectivity in terpene synthases comes from the work of Köllner et al. (2004), who identified two maize (*Zea mays*) terpene synthases, TPS4 and TPS5, each forming the same complex mixture of sesquiterpenes but with different stereospecificity of products. The differences in the stereoselectivity of TPS4 and TPS5 were found to be determined solely by four amino acid substitutions in a stretch of five residues (407 to 411). Primary structure comparison of the maize TPS with Sf-CinS1 reveals that these four residues correspond exactly to Sf-CinS1 region 2, while modeling of the maize enzymes on the TEAS structure locates these residues in the G-helix kink.

Two successful attempts have recently been made to rationally dictate the products formed by a sesquiterpene synthase. Greenhagen et al. (2006) applied a concentric contact mapping approach to identify second-tier residues that would enable the conversion of TEAS to HPS. Their work highlighted the importance of outer-tier residues in the structure of the active site and the difficulties involved when applying rational approaches to enzymes whose scaffold might, to some extent, have been affected by evolution. In this case, a large number of mutations may be required to cover variation from first and outer-tier residues without any means of determining the minimum and specific number of substitutions required. In our approach, the availability of sequence and structural information from very closely related species was essential in the identification of specific structural determinants minimizing the number of candidates and the contribution of outer-tier residues.

Yoshikuni et al. (2006) performed exhaustive site-directed mutagenesis to 19 residues indicated by homology modeling to surround the active site of the sesquiterpene γ-humulene synthase to demonstrate the presence of plasticity residues in the terpene synthase active site and their importance in the catalytic outcome. The high plasticity of the terpene synthase active site is consistent with the notion that product specificity in these enzymes is dictated not so much by the precise positioning of reactive groups but rather by a combination of the contour and the dynamics of the active site in a way that facilitates certain isomerization steps and disallows others. The results presented here provide further evidence for the plasticity of the active site and indicate that minor alterations of the active-site contour, sometimes not more dramatic than the addition or removal of a methyl group or the shift of a backbone atom by a fraction of an Ångstrom, can have profound effects on the activity of these enzymes (Tables 2 and 3). Our work extends the findings of Yoshikuni et al. (2006) by providing precise structural information on the role of active-site residues in product specificity. Both reports confirm how evolvable the terpene scaffold is and show that protein engineering can successfully be applied to the design of terpene synthase function.

The extent to which substitutions of critical residues could affect reaction outcome is further underlined by the conversion of Sf-CinS1 to a sesquiterpene synthase by the alteration of a single residue. Phylogenetic analysis has shown that within each of several species, monoterpene and sesquiterpene genes cluster closer together than genes encoding for enzymes of similar activity in different species (Bohlmann et al., 1998; Tholl, 2006). Although no immediate evolutionary advantage may result from an enzyme of such dual functionality because of the likely absence or inefficiency of sesquiterpene production in the plastids, a subsequent event resulting in the loss of the monoterpene-specific transit peptide could make this gene product a functional sesquiterpene synthase. Evidence that such an event can take place in nature is reported by Aharoni et al. (2004), who identified a bifunctional linalool/nerolidol synthase from strawberry (*Fragaria* spp) (Fa-NES1) that has acquired cytosolic localization due to translation initiation from a downstream ATG after insertion of a stop codon in its transit peptide region.

Analysis of the *Arabidopsis* terpene synthase gene family revealed >30 different members, while preliminary evidence indicates an even higher number in *S. fruticosa* (E. Ninga and S.C. Kampranis, unpublished data). Analysis of the *Arabidopsis* genes indicates that these are the result of several cycles of gene duplication and mutagenesis (Aubourg et al., 2002; Chen et al., 2004; Ro et al., 2006). Consistent with the demands of a changing environment, terpene biosynthesis has to adapt rapidly by drastic alterations in the terpenoid profile produced. Our results provide the structural evidence for a mechanism where plants, instead of using a different dedicated gene for the production of each different terpene, may employ a highly flexible simple scaffold that can, with only very few substitutions, provide the complete spectrum of activities required. It may be noted that the point mutations introduced here into Sf-CinS1 resulted in enzymes with a product spectrum sufficient to account qualitatively and quantitatively for the content and variation found between the monoterpenes (or their volatile precursors) in the three *Salvia* species examined in our phylogenetic comparison and within their populations (Skoula et al., 2000), indicating the relative simplicity of the evolutionary development of product diversity in these species. The combination of specificity and promiscuity observed in terpene synthases may be a snapshot of an evolutionary mechanism that maintains a dynamic state between desirable specific activities and a capacity for rapid change.

METHODS

Cloning of the Sf-*CinS1* and Sp-*SabS1* Genes into Appropriate Expression Vectors

The genes encoding for the *Salvia fruticosa* 1,8-cineole synthase and the *Salvia pomifera* sabinene synthase were isolated through an EST sequencing approach using two tissue-specific glandular trichome-derived cDNA libraries (S.C. Kampranis, C.B. Johnson, A.M. Makris, and J. Degenhardt, unpublished data). For expression in *Escherichia coli*, Sf-CinS1 and Sp-SabS1 were subcloned in the pRSETa vector as follows: the open reading frame of Sp-SabS1 was amplified using primers 5′Sbs(NdeI) (5′-CATATGCGACGCTCTGGGGATTACCAA-3′) and 3′Sbs(SalI) (5′-GTCGACTCAGACATAAGGCTGGAATAGCAG-3′), removing the chloroplastic transit peptide and introducing *Nde*I and *Sal*I restriction sites in the 5′ and 3′, respectively. The PCR product was initially cloned into the pCR2.1 TOPO TA vector (Invitrogen) and then transferred into the pRSETa (Invitrogen) vector's *Nde*I and *Xho*I restriction sites, exploiting the compatibility of *Sal*I and *Xho*I overhangs. Sf-CinS1 was amplified using the gene-specific primer 5′CsR(NdeI) (5′-CATATGCGACGAACTGGAGGCTACCAGCCT-3′) and primer pDNRLrev (5′-CCAAACGAATGGTCTAGAAAGCTTCTCGAC-3′), which is specific for the library plasmid. The insert was cloned into the pRSET vector as above.

Large-Scale Protein Production and Purification

E. coli BL21 [DE3, pLANT(3)/RIL; Finkelstein et al., 2003] cells carrying the cineole synthase plasmid were grown using a New Brunswick BIOFLOW 4500 aerated fermenter at 35°C and 300 rpm in 20 liters of terrific broth containing equal amounts (50 μg/mL) of kanamycin, ampicillin, and chloramphenicol antibiotics. Protein expression was induced at an OD_{600} of 0.5 to 0.7 with 1 mM isopropylthio-β-galactoside, after which cell growth was continued for 20 h at 19°C before the cells were harvested. The cells were isolated at 4°C using a Sartorius benchtop concentrator and centrifugation at 2800*g*, prior to resuspension in lysis buffer containing 60 mM Tris, pH 8.8, 5 mM β-mercaptoethanol, 0.5 M NaCl, 10 mM imidazole, 8% (w/v) glycerol, 0.8% Tween 20, 1 mM PMSF, 1 mg/mL of lysozyme, and 1.5 ng/mL of benzonase. Cell debris was removed via ultracentrifugation at 35,000*g* and 4°C before lowering the pH to 8.0 using 1 M Tris. The His-tagged cineole synthase was extracted using a Ni-column preequilibrated with 60 mM Tris, pH 8.8, 5 mM β-mercaptoethanol, 0.5 M NaCl, 10 mM imidazole, and 8% glycerol. The protein was washed off the column using 500 mM imidazole. The sample was then dialyzed against 20 mM Tris, pH 8.0, 5 mM β-mercaptoethanol, 150 mM NaCl, and 4% glycerol at 4°C prior to gel filtration using a Pharmacia Biotech Superdex 200 HR 10/30 column using the same equilibration buffer. Finally, the protein was polished by ion exchange with a Pharmacia Biotech Mono-Q HR 5/5 column using a 0.15 to 1 M NaCl gradient. The pure cineole synthase was then desalted and stored at 4 mg/mL in 20 mM Tris, pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol, and 4% glycerol at −80°C until ready for use.

Crystallization

Cineole synthase crystals were grown using the hanging drop method of vapor diffusion with a 3-μL drop containing equal amounts of precipitant and protein solution at 20°C. Crystals were grown using 100 mM Tris, pH 8.0, 5 mM β-mercaptoethanol, 1.0 M sodium formate, and 10% (w/v) polyethylene glycol 4000 as the precipitant. The crystals grew in an orthorhombic system (space group $C222_1$) with unit cell dimensions of $a = 124.6$ Å, $b = 171.1$ Å, and $c = 123.8$ Å.

Data Collection, Processing, and Phasing

The x-ray diffraction data were collected at cryogenic temperatures from a single crystal at the Daresbury Synchrotron Radiation Source on beamline 10.1 ($\lambda = 1.380$ Å) using a MAR CCD165 detector (Cianci et al., 2005). The crystals were transferred to a small drop of precipitant solution containing 19% glycerol for ~5 s before flash freezing in a gaseous nitrogen stream (100K) prior to data collection. The data were collected in a single batch of 300 images to a resolution of 1.9 Å using an oscillation range of 0.3° and an exposure time of 2 s. The data set was processed using MOSFLM (Leslie, 1998) before scaling and merging using SCLALA (Collaborative Computational Project, Number 4, 1994), at which point the high-resolution data were cut back to 1.95 Å. A summary of the data statistics is shown in Table 1. The data were then phased by molecular replacement using MOLREP (Vagin and Teplyakov, 1997) and solved using a So-BPPS monomer (PDB: 1N1B), excluding water molecules and metal ions, as the search model. The crystal was found to contain two molecules in the asymmetric unit with solvent contents of 52.4%.

Model Building and Refinement

Refinement was performed progressively using higher-resolution data until the highest possible resolution was obtained. CNS (Brunger et al., 1998) and or ARP/wARP (Perrakis et al., 1999, 2001) and REFMAC5 (Collaborative Computational Project, Number 4, 1994) were used in the refinement between progressive rounds of manual model building using the program Quanta (Cerius2 Modeling Environment, release 4.10; Accelrys Software). During the course of the refinement, the stereochemical properties were checked using WHAT_CHECK (Vriend, 1990) and PROCHECK (Laskowski et al., 1993). The final R-factor for crystal was 21.8% (23.5% R-free). The final models showed no residues within the disallowed regions of the Ramachandran plot. Disordered regions include molecule A (N terminus-83, 224 to 227, 306 to 307, 494 to 503, 511 to 516, and 573) and molecule B (N terminus-86, 226 to 227, 493 to 502, and 572 to 573). Thirty-one side chains had undetermined orientations and were removed from the structure (127 atoms in total). A summary of the refinement statistics is shown in Table 1.

Site-directed mutagenesis was performed using the Quickchange method (Stratagene) according to the manufacturer's instructions. The primers used are listed below: CinS1 (N338I)5, 5′-GGATAATGCTCACCAAAATAATTGCTCTTGTTACAACAATAGACG-3′; CinS1 (N338I)3, 5′-CGTCTATTGTTGTAACAAGAGCAATTATTTTGGTGAGCATTATCC-3′; CinS1 (N338I-A339T)5, 5′-GGATAATGCTCACCAAAATAATTACTCTTGTTACAACAATAGACG; CinS1 (N338I-A339T)3, 5′-CGTCTATTGTTGTAACAAGAGTAATTATTTTGGTGAGCATTATCC-3′; CinS1 (G447S)5; 5′-GAAGAATAGTTGGATATCAATCAGCGGCATCCCCATTCTATCTC-3′; CinS1 (G447S)3, 5′-GAGATAGAATGGGGATGCCGCTGATTGATATCCAACTATTCTTC-3′; SabS1 (I327N)5, 5′-GAGAAAAATGGCCGCCATTATTAATACTTTCGTAACAATTATCG-3′; SabS1 (I327N)3, 5′-CGATAATTGTTACGAAAGTATTAATAATGGCGGCCATTTTTCTC-3′; SabS1 (I327N-T328A)5, 5′-GAGAAAAATGGCCGCCATTATTAATGCTTTCGTAACAATTATCG-3′; SabS1 (I327N-T328A)3, 5′-CGATAATTGTTACGAAAGCATTAATAATGGCGGCCATTTTTCTC-3′; SabS1 (S436G)5, 5′-CTCAACAACGCCAAGATTTCAATAGGGGCTCCTACAATCATATCC-3′; SabS1 (S436G)3, 5′-GGATATGATTGTAGGAGCCCCTATTGAAATCTTGGCGTTGTTGAG-3′; CinS (N338L)5, 5′-GAGGATAATGCTCACCAAAATACTTGCTCTTGTTACAACAATAGACG; CinS (N338L)3, 5′-CGTCTATTGTTGTAACAAGAGCAAGTATTTTGGTGAGCATTATCCTC-3′; CinS (N338A)5, 5′-GGATAATGCTCACCAAAATAGCTGCTCTTGTTACAACAATAGACG-3′; CinS (N338A)3, 5′-CGTCTATTGTTGTAACAAGAGCAGCTATTTTGGTGAGCATTATCC-3′; CinS (N338S)5, 5′-GAGGATAATGCTCACCAAAATATCTGCTCTTGTTACAACAATAGACG-3′; CinS (N338S)3, 5′-CGTCTATTGTTGTAACAAGAGCAGATATTTTGGTGAGCATTATCCTC-3′; CinS (N338V)5, 5′-GAGGATAATGCTCACCAAAATAGTTGCTCTTGTTACAACAATAGACG-3′; CinS (N338V)3, 5′-CGTCTATTGTTGTAACAAGAGCAACTATTTTGGTGAGCATTATCCTC-3′; CinS (N338C)5, 5′-GAGGATAATGCTCACCAAAATATGTGCTCTTGTTACAACAATAGACG-3′; CinS (N338C)3, 5′-CGTCTATTGTTGTAACAAGAGCACATATTTTGGTGAGCATTATCCTC-3′; CinS (SGPP)5, 5′-GTTGGATATCAATCAGCGGCCCCCCCATTCTATCTCATCTATTTTTC-3′; CinS (SGPP)3, 5′-GAAAAATAGATGAGATAGAATGGGGGGGCCGCTGATTGATATCCAC-3′; CinS (SGPT)5, 5′-GGATATCAATCAGCGGCCCCACCATTCTATCTCATCTATTTTTCC-3′; CinS (SGPT)3, 5′-GGAAAAATAGATGAGATAGAATGGTGGGGCCGCTGATTGATATCC-3′. All mutations were verified by automated nucleotide sequencing.

Small-Scale Protein Expression in Bacteria

Protein expression was induced by the addition of isopropylthio-β-galactoside in 250 mL of *E. coli* BL21 [DE3, pLANT(3)/RIL; (Finkelstein et al., 2003) cultures growing at 37°C. After growth for 3 to 4 h at 30°C, the cells were collected, resuspended in lysis buffer (50 mM NaH_2PO_4, pH 8, 300 mM NaCl, and 10 mM imidazole), and kept in ice for 30 min after the addition of protease inhibitor cocktail (Roche) and lysozyme (0.1 mg/mL). The cells were disrupted by sonication, and the lysate was then centrifuged at 13,000*g* at 4°C for 20 min. The supernatant was collected and assayed immediately. The His-tagged proteins were then purified by Ni^{2+} affinity chromatography as described above.

Enzyme Assays

Monoterpene synthase activities were assayed by adding the protein extract in a 500-μL reaction containing 10 mM MOPS, pH 7.0, 20 mM $MgCl_2$, 2 mM $MnCl_2$, 10 mM DTT, and 5% (w/v) glycerol. They were mixed into a 2.5-mL glass vial and then overlaid with 1 mL of pentane to trap the volatile products. The reaction was initiated by the addition of 27 μM final concentration of GPP (Echelon Biosciences), with incubation at room temperature for 1 h up to overnight, depending on the enzyme. The pentane layer was then removed and passed through a short column of Na_2SO_4 in a Pasteur pipette to remove the water. The volume of the organic phase was adjusted to 1 mL and then concentrated to 300 μL for gas chromatography analysis. For the detection of sesquiterpene synthase activity, 54 μM FPP (Echelon Biosciences) was used instead, and the products were extracted by a mixture of pentane-diethyl ether. A Hewlett Packard 5890 II gas chromatograph equipped with a flame ionization detector was used for the analysis. The DB5 column used (30 m long and 0.25 mm in diameter), and the split ratio was 1:50. The temperatures of the injector and detector were 230 and 260°C, respectively. The initial oven temperature was 45°C, increasing at 1.5°C/min until 150°C, then at 40°C/min until 220°C, being held at this temperature for the last 10 min. For the structural identification of the produced volatiles, solid-phase microextraction (SPME) sampling coupled to gas chromatography–mass spectrometry analysis was performed at VIORYL.

SPME

The SPME holder and fibers used (7-mm bonded polydimethylsiloxane coating) were obtained from Supelco. Sampling of a 0.5-mL reaction contained in a 5-mL sealed Supelco vial was performed by inserting the syringe needle of the SPME assembly, through the septum cap, into the headspace above the sample. Volatiles were adsorbed by extending the fiber into the headspace. After 30 min at 30°C, the fiber was withdrawn into the outer septum-piercing needle, removed from the vial, and inserted in the heated injection port of the gas chromatography. The products were identified by comparing retention times and mass spectra with authentic reference compounds.

Determination of Kinetic Parameters

The determination of the kinetic parameters of the reaction of the *Salvia* terpene synthases or their mutants with GPP or FPP was performed as follows: 0.8 to 40 μg of purified enzyme (depending on the activity of the mutant form) was added in an enzymatic assay (described above) while varying the substrate concentration in the range 1 to 81 μM for GPP and 3 to 108 μM for FPP (at least 11 different concentrations tested in each determination, while each concentration was assayed in duplicate). The reaction was overlaid with 500 μL of myrcene-containing pentane (for monoterpenes) or nerolidol-containing pentane-diethyl ether mixture (for sesquiterpenes) and allowed to proceed for 1 h at 25°C. Myrcene and nerolidol served as internal standards for the accurate determination of volatile products produced. Subsequent extraction steps and gas chromatography analysis were as described above. k_{cat} numbers were determined on the basis of the total amount of mono- or sesquiterpenes formed. Kinetic analysis was performed using the computer software SigmaPlot (SPSS).

Molecular modeling was initiated by the introduction of the desired mutation and followed by energy minimization of the residues in a 6-Å radius of the mutated side chain using the Gromos96 force field implemented in DeepView (Swiss-PdbViewer) software (Guex and Peitsch, 1997). The FPP substrate was then fitted manually into the active-site cavity, maintaining the superposition of the diphosphate part with that of the 3-aza-2,3-dihydrogeranyl diphosphate ligand of the So-BPPS structure (PDB: 1N23). Graphics were produced by DeepView (Guex and Peitsch, 1997) and UCSF Chimera (Sanner et al., 1996; Pettersen et al., 2004). The analysis of synonymous and nonsynonymous nucleotide changes between terpene synthases was conducted with DnaSP software by Universitat de Barcelona, Spain.

Accession Numbers

The nucleotide sequences of the genes encoding for the *S. fruticosa* 1,8-cineole synthase and the *S. pomifera* sabinene synthase can be found in the GenBank/EMBL data libraries under accession numbers DQ785793 for Sf-CinS1 and DQ785794 for Sp-SabS1. The atomic model coordinates for the structure described in this article have been deposited with the Protein Data Bank with accession code 2j5c.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Molecular Model of the Active Site of Some Sf-CinS1 Mutants.

Supplemental Figure 2. Plot of Amino Acid Similarity across the C-Terminal Domain of an Alignment of 46 Monoterpene Synthases.

ACKNOWLEDGMENTS

We acknowledge the financial support of a research grant from European Union Framework V (Contract QLK3-CT-2002-1930). The North West Structural Genomics Centre MAD10 beamline at Synchrotron Radiation Source (UK) is funded by Biotechnology and Biological Science Research Council grants (719/B15474 and 719/REI20571) and a Northwest Development Agency project award (N0002170). We thank Melpo Skoula for plant material and analysis of compounds, Christos Petrakis for assistance with gas chromatography analysis, and Georgia Tsotsou and Nikitas Ragoussis for sesquiterpene analysis. We acknowledge the assistance of Panagiotis Kanellopoulos with the structural modeling of the terpene synthases.

Received September 27, 2006; revised May 9, 2007; accepted May 21, 2007; published June 8, 2007.

REFERENCES

Aharoni, A., Giri, A.P., Verstappen, F.W., Bertea, C.M., Sevenier, R., Sun, Z., Jongsma, M.A., Schwab, W., and Bouwmeester, H.J. (2004). Gain and loss of fruit flavor compounds produced by wild and cultivated strawberry species. Plant Cell **16:** 3110–3131.

Aubourg, S., Lecharny, A., and Bohlmann, J. (2002). Genomic analysis of the terpenoid synthase (AtTPS) gene family of *Arabidopsis thaliana*. Mol. Genet. Genomics **267:** 730–745.

Bohlmann, J., Meyer-Gauen, G., and Croteau, R. (1998). Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. USA **95:** 4126–4133.

Brunger, A.T., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. **54:** 905–921.

Caruthers, J.M., Kang, I., Rynkiewicz, M.J., Cane, D.E., and Christianson, D.W. (2000). Crystal structure determination of aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*. J. Biol. Chem. **275:** 25533–25539.

Chen, F., Ro, D.K., Petri, J., Gershenzon, J., Bohlmann, J., Pichersky, E., and Tholl, D. (2004). Characterization of a root-specific *Arabidopsis*

terpene synthase responsible for the formation of the volatile monoterpene 1,8-cineole. Plant Physiol. **135:** 1956–1966.

Christianson, D.W. (2006). Structural biology and chemistry of the terpenoid cyclases. Chem. Rev. **106:** 3412–3442.

Cianci, M., et al. (2005). A high-throughput structural biology/proteomics beamline at the SRS on a new multipole wiggler. J. Synchrotron Radiat. **12:** 455–466.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: Programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50:** 760–763.

Croteau, R. (1987). Biosynthesis and catabolism of monoterpenoids. Chem. Rev. **87:** 929–954.

Croteau, R., Alonso, W.R., Koepp, A.E., and Johnson, M.A. (1994). Biosynthesis of monoterpenes: Partial purification, characterization, and mechanism of action of 1,8-cineole synthase. Arch. Biochem. Biophys. **309:** 184–192.

Ehrich, P.R., and Raven, P.H. (1964). Butterflies and plants. A study in co-evolution. Evolution Int. J. Org. Evolution **18:** 586–608.

El Tamer, M.K., Lucker, J., Bosch, D., Verhoeven, H.A., Verstappen, F.W., Schwab, W., van Tunen, A.J., Voragen, A.G., de Maagd, R.A., and Bouwmeester, H.J. (2003). Domain swapping of *Citrus limon* monoterpene synthases: Impact on enzymatic activity and product specificity. Arch. Biochem. Biophys. **411:** 196–203.

Finkelstein, J., Antony, E., Hingorani, M.M., and O'Donnell, M. (2003). Overproduction and analysis of eukaryotic multiprotein complexes in *Escherichia coli* using a dual-vector strategy. Anal. Biochem. **319:** 78–87.

Greenhagen, B.T., O'Maille, P.E., Noel, J.P., and Chappell, J. (2006). Identifying and manipulating structural determinates linking catalytic specificities in terpene synthases. Proc. Natl. Acad. Sci. USA **103:** 9826–9831.

Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis **18:** 2714–2723.

Harborne, J.B. (1993). Introduction to Ecological Biochemistry, 4th ed. (London: Academic Press).

Hyatt, D.C., and Croteau, R. (2005). Mutational analysis of a monoterpene synthase reaction: Altered catalysis through directed mutagenesis of (-)-pinene synthase from *Abies grandis*. Arch. Biochem. Biophys. **439:** 222–233.

Köllner, T.G., Schnee, C., Gershenzon, J., and Degenhardt, J. (2004). The variability of sesquiterpenes emitted from two *Zea mays* cultivars is controlled by allelic variation of two terpene synthase genes encoding stereoselective multiple product enzymes. Plant Cell **16:** 1115–1131.

Laskowski, R.A., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). Procheck - A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26:** 283–291.

Lesburg, C.A., Zhai, G., Cane, D.E., and Christianson, D.W. (1997). Crystal structure of pentalenene synthase: Mechanistic insights on terpenoid cyclization reactions in biology. Science **277:** 1820–1824.

Leslie, A.G.W. (1998). New auto-indexing using DPS due to Ingo Steller Robert Bolotovsky and Michael Rossmann. J. Appl. Crystallogr. **30:** 1036–1040.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. Acta Crystallogr. D Biol. Crystallogr. **57:** 1445–1450.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. **6:** 458–463.

Peters, R.J., and Croteau, R.B. (2003). Alternative termination chemistries utilized by monoterpene cyclases: Chimeric analysis of bornyl diphosphate, 1,8-cineole, and sabinene synthases. Arch. Biochem. Biophys. **417:** 203–211.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera – A visualization system for exploratory research and analysis. J. Comput. Chem. **25:** 1605–1612.

Ro, D.K., Ehlting, J., Keeling, C.I., Lin, R., Mattheus, N., and Bohlmann, J. (2006). Microarray expression profiling and functional characterization of AtTPS genes: Duplicated *Arabidopsis thaliana* sesquiterpene synthase genes At4g13280 and At4g13300 encode root-specific and wound-inducible (Z)-gamma-bisabolene synthases. Arch. Biochem. Biophys. **448:** 104–116.

Rynkiewicz, M.J., Cane, D.E., and Christianson, D.W. (2001). Structure of trichodiene synthase from *Fusarium sporotrichioides* provides mechanistic inferences on the terpene cyclization cascade. Proc. Natl. Acad. Sci. USA **98:** 13543–13548.

Sanner, M.F., Olson, A.J., and Spehner, J.C. (1996). Reduced surface: An efficient way to compute molecular surfaces. Biopolymers **38:** 305–320.

Schnee, C., Kollner, T.G., Gershenzon, J., and Degenhardt, J. (2002). The maize gene terpene synthase 1 encodes a sesquiterpene synthase catalyzing the formation of (E)-beta-farnesene, (E)-nerolidol, and (E,E)-farnesol after herbivore damage. Plant Physiol. **130:** 2049–2060.

Sharkey, T.D., Yeh, S., Wiberley, A.E., Falbel, T.G., Gong, D., and Fernandez, D.E. (2005). Evolution of the isoprene biosynthetic pathway in kudzu. Plant Physiol. **137:** 700–712.

Skoula, M., Abbes, J.E., and Johnson, C.B. (2000). Genetic variation of volatiles and rosmarinic acid in populations of *Salvia fruticosa* mill growing in Crete. Biochem. Syst. Ecol. **28:** 551–561.

Starks, C.M., Back, K., Chappell, J., and Noel, J.P. (1997). Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. Science **277:** 1815–1820.

Tholl, D. (2006). Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr. Opin. Plant Biol. **9:** 297–304.

Trapp, S.C., and Croteau, R.B. (2001). Genomic organization of plant terpene synthases and molecular evolutionary implications. Genetics **158:** 811–832.

Vagin, A., and Teplyakov, A. (1997). MOLREP: An automated program for molecular replacement. J. Appl. Crystallogr. **30:** 1022–1025.

Vriend, G. (1990). WHAT IF: A molecular modeling and drug design program. J. Mol. Graph. **8:** 52–56, 29.

Whittington, D.A., Wise, M.L., Urbansky, M., Coates, R.M., Croteau, R.B., and Christianson, D.W. (2002). Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. Proc. Natl. Acad. Sci. USA **99:** 15375–15380.

Wise, M.L., Savage, T.J., Katahira, E., and Croteau, R. (1998). Monoterpene synthases from common sage (*Salvia officinalis*). cDNA isolation, characterization, and functional expression of (+)-sabinene synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase. J. Biol. Chem. **273:** 14891–14899.

Yoshikuni, Y., Ferrin, T.E., and Keasling, J.D. (2006). Designed divergent evolution of enzyme function. Nature **440:** 1078–1082.