2-24-2012

Insights into Diterpene Cyclization from Structure of Bifunctional Abietadiene Synthase from Abies grandis

Ke Zhou
Iowa State University

Yang Gao
Iowa State University

Julie A. Hoy
Iowa State University

Francis M. Mann
Iowa State University

Richard B. Honzatko
Iowa State University, honzatko@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/bbmb_ag_pubs

Part of the Biochemistry Commons, Bioinformatics Commons, Molecular Biology Commons, and the Structural Biology Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/bbmb_ag_pubs/115. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.

This Article is brought to you for free and open access by the Biochemistry, Biophysics and Molecular Biology at Iowa State University Digital Repository. It has been accepted for inclusion in Biochemistry, Biophysics and Molecular Biology Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Insights into Diterpene Cyclization from Structure of Bifunctional Abietadiene Synthase from Abies grandis*

Received for publication, December 23, 2011 Published, JBC Papers in Press, January 4, 2012, DOI 10.1074/jbc.M111.337592

Ke Zhou, Yang Gao, Julie A. Hoy1, Francis M. Mann, Richard B. Honzatko, and Reuben J. Peters2
From the Department of Biochemistry, Biophysics, and Molecular Biology, Iowa State University, Ames, Iowa 50011

Background: Class I and II diterpene synthases, although poorly understood, generate diverse products. Results: Reported here is the structure of the bifunctional abietadiene synthase and supporting experimental/computational work. Conclusion: Visualization of the class I and II active sites confirms known and implicates new determinants of product formation. Significance: Residues, previously unrecognized, are assigned specific roles in substrate binding and catalysis.

Abietadiene synthase from Abies grandis (AgAS) is a model system for diterpene synthase activity, catalyzing class I (ionization-initiated) and class II (protonation-initiated) cyclization reactions. Reported here is the crystal structure of AgAS at 2.3 Å resolution and molecular dynamics simulations of that structure with and without active site ligands. AgAS has three domains (α, β, and γ). The class I active site is within the C-terminal α domain, and the class II active site is between the N-terminal γ and β domains. The domain organization resembles that of monofunctional diterpene synthases and is consistent with proposed evolutionary origins of terpene synthases. Molecular dynamics simulations were carried out to determine the effect of substrate binding on enzymatic structure. Although such studies of the class I active site do lead to an enclosed substrate-Mg2+ complex similar to that observed in crystal structures of related plant enzymes, it does not enforce a single substrate conformation consistent with the known product stereochemistry. Simulations of the class II active site were more informative, with observation of a well ordered external loop migration. This “loop-in” conformation not only limits solvent access but also greatly increases the number of conformational states accessible to the substrate while destabilizing the nonproductive substrate conformation present in the “loop-out” conformation. Moreover, these conformational changes at the class II active site drive the substrate toward the proposed transition state. Docked substrate complexes were further assessed with regard to the effects of site-directed mutations on class I and II activities.

Terpenoids form the largest group of natural products, with some 50,000 known (1). Underlying the observed diversity are the manifold hydrocarbon skeletal backbones produced by the cyclization and/or rearrangement of acyclic precursors catalyzed by the relevant terpene synthases (2). They are first subdivided based on the number of five-carbon isoprenoid repeats; with the ten-carbon monoterpenoids generally derived from geranyl diphosphate, the fifteen-carbon sesquiterpenoids from farnesyl diphosphate, the twenty-carbon diterpenoids from geranylgeranyl diphosphate (GGPP),3 and the thirty-carbon triterpenoids from squalene (3).

As might be suspected from the composition of their precursors, the skeletal backbone structures of the lower (mono-, sesqui, and di-) terpenes are generally produced by cyclization and/or rearrangement reactions initiated by ionization of the substrate allylic diphosphate ester bond, whereas triterpenes are cyclized via a protonation-initiated reaction (i.e. of a terminal C=C double bond or the oxido-ring of the derived oxido-squalene). The corresponding enzymes have been termed classes I and II, respectively (4). Notably, biosynthesis of the large superfamily of labdane-related diterpenoids (~7,000 known) combines both types of reactions (5), in which GGPP is bicycled by a class II diterpene cyclase (EC 5.5.1.x) prior to further cyclization and/or rearrangement catalyzed by a class I diterpene synthase (EC 4.2.3.x).

The labdane-related diterpenoids include many important plant natural products, such as the gibberellin phytohormones and others that act in defense (e.g. the ubiquitous conifer resin acids (6)). Indeed, it has been suggested that the diterpene synthases involved in gibberellin biosynthesis are the ancestral progenitors to the extensive family of lower (i.e. class I) terpene synthases found in plants (7). This is possible because, while catalyzing mechanistically distinct reactions, the corresponding class II ent-copalyl diphosphate synthase and subsequently acting class I ent-kaurene synthase are clearly homologous. Nevertheless, the vast majority of the plant lower terpene synthases are significantly smaller than these diterpene synthases (8).

*This work was supported, in whole or in part, by National Institutes of Health Grants GM076324 (to R. J. P.) and NS010546 (to R. B. H.). This work was also supported in part by Department of Energy Grants W-31–109-ENG-38 and DE-AC02–06CH11357 and by the National Science Foundation. The atomic coordinates and structure factors (code 3S9V) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1Present address: California Institute of Technology, Division of Chemistry and Chemical Engineering, Pasadena, CA 91125.
2To whom correspondence should be addressed: Iowa State University, 4216 Mol. Biol. Bldg., Ames, IA 50011. Fax: 515-294-8580; E-mail: rjpeters@iastate.edu.

3The abbreviations used are: GGPP, (E,E)-geranylgeranyl diphosphate; AgAS, abietadiene synthase from A. grandis; AtCPS, ent-CPP synthase from Arabidopsis thaliana; CPP, copalyl diphosphate; TbT5, taxadiene synthase from Taxus brevifolia.
Structural characterization of plant and microbial lower terpene synthases demonstrated that the class I reaction is catalyzed in a conserved α-helical bundle domain (9–18) that has been termed the α domain (19). This includes the placement of two highly conserved acidic motifs, DDXXE and (N/D)DX(S/T)XXXE (hereafter called NTE), which coordinate the trio of Mg$^{2+}$ required for catalysis (21, 22). This invokes similarly allylic double-bond isomers (38). Subsequent oxidation reactions transform the resulting abietadienes into resin acids (6).

**FIGURE 1.** Class II and class I reactions catalyzed by AgAS. Isopimar-15-en-8-yl$^+$ is a key intermediate of the class I reaction, with formation of a mixture of allylic double-bond isomers (38). Subsequent oxidation reactions transform the resulting abietadienes into resin acids (6).

Experimental Procedures

Cloning, Expression, and Purification of AgAS—For crystallization the Δ84 pseudo-mature form of AgAS was recombinantly expressed using a previously described pSETB construct (38), typically in the C41 OverExpress strain of *Escherichia coli* (Lucigen) grown in NZY medium. For selenomethionine labeling, AgAS was expressed in B834(DE3) methionine auxotrophic *E. coli* (Novagen) grown in selenomethionine medium base plus nutrient mix (Molecular Dimensions) with the addition of L-selenomethionine (Fisher). AgAS was then purified much as previously described (38). Briefly, clarified bacterial lysates were initially fractionated over type II ceramic hydroxyapatite, and AgAS was purified over Mono Q and type II ceramic hydroxyapatite again using a BioLogic Logic system (GE Healthcare). The resulting AgAS (>98% pure by SDS-PAGE) was then dialyzed against protein storage buffer (10 mM Bis-Tris, pH 6.8, 10% (v/v) glycerol, 150 mM KCl, 10 mM MgCl$_2$, and 5 mM DTT), following which it could be stored at −80 °C for several months without significant loss of activity.

Crystallization of AgAS—Initial crystallization trials were carried out at the Hauptman-Woodward Institute. Promising conditions were replicated in our lab by the hanging drop vapor diffusion methods at both 4 and 18 °C. Typically, a 2-µL drop of AgAS in storage buffer was mixed with a 1.6-µL drop of precipitant solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), as well as with 0.4 µL of additional 0.1 M L-proline or 0.1 M phenol from an additive solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), before a final polishing purification step over a Mono Q column on an AKTA Fplc system (GE Healthcare). The resulting AgAS (>98% pure by SDS-PAGE) was then dialyzed against protein storage buffer (10 mM Bis-Tris, pH 6.8, 10% (v/v) glycerol, 150 mM KCl, 10 mM MgCl$_2$, and 5 mM DTT), following which it could be stored at −80 °C for several months without significant loss of activity.

Crystallization of AgAS—Initial crystallization trials were carried out at the Hauptman-Woodward Institute. Promising conditions were replicated in our lab by the hanging drop vapor diffusion methods at both 4 and 18 °C. Typically, a 2-µL drop of AgAS in storage buffer was mixed with a 1.6-µL drop of precipitant solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), as well as with 0.4 µL of additional 0.1 M L-proline or 0.1 M phenol from an additive solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), before a final polishing purification step over a Mono Q column on an AKTA Fplc system (GE Healthcare). The resulting AgAS (>98% pure by SDS-PAGE) was then dialyzed against protein storage buffer (10 mM Bis-Tris, pH 6.8, 10% (v/v) glycerol, 150 mM KCl, 10 mM MgCl$_2$, and 5 mM DTT), following which it could be stored at −80 °C for several months without significant loss of activity.

Crystallization of AgAS—Initial crystallization trials were carried out at the Hauptman-Woodward Institute. Promising conditions were replicated in our lab by the hanging drop vapor diffusion methods at both 4 and 18 °C. Typically, a 2-µL drop of AgAS in storage buffer was mixed with a 1.6-µL drop of precipitant solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), as well as with 0.4 µL of additional 0.1 M L-proline or 0.1 M phenol from an additive solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), before a final polishing purification step over a Mono Q column on an AKTA Fplc system (GE Healthcare). The resulting AgAS (>98% pure by SDS-PAGE) was then dialyzed against protein storage buffer (10 mM Bis-Tris, pH 6.8, 10% (v/v) glycerol, 150 mM KCl, 10 mM MgCl$_2$, and 5 mM DTT), following which it could be stored at −80 °C for several months without significant loss of activity.

Crystallization of AgAS—Initial crystallization trials were carried out at the Hauptman-Woodward Institute. Promising conditions were replicated in our lab by the hanging drop vapor diffusion methods at both 4 and 18 °C. Typically, a 2-µL drop of AgAS in storage buffer was mixed with a 1.6-µL drop of precipitant solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), as well as with 0.4 µL of additional 0.1 M L-proline or 0.1 M phenol from an additive screen (Hampton Research), and equilibrated against 1 ml of precipitant solution in the well. Plate-shaped crystals with maximal dimensions of 0.3 × 0.3 × 0.05 mm appeared within a week. Crystals were briefly transferred to a cryoprotectant solution (25% PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate), as well as with 0.4 µL of additional 0.1 M L-proline or 0.1 M phenol from an additive solution (24% (v/v) PEG 8000, 0.1 M sodium citrate, pH 5.1, 0.1 M dibasic ammonium phosphate, 10% (v/v) glycerol) and then flash-frozen. One selenomethionine labeled crystal (crystal I) was soaked in 5.5 mM tantalum bromide (Ta$_5$Br$_{12}^{3+}$; Jena Bioscience) prior to flash-freezing. Although this derivative was not successfully used for phasing, it appears to have altered the crystal lattice (from P2$_1$ to P2$_2$2$_1$2) which facilitated localization of the selenium for phasing. This crystal, which diffracted to 2.5 Å resolution, belonged to space group P2$_1$2$_1$2 with unit cell dimensions $a = 100.14$ Å, $b = 190.175$ Å, and $c = 92.932$ Å (two molecules in the asymmetric unit). The data also were
collected for another selenomethionine-labeled crystal (crystal II), which diffracted to 2.3 Å, and belonged to space group P2₁ with unit cell parameters \(a = 92.84\,\text{Å}, b = 189.09\,\text{Å}, \text{and } c = 99.90\,\text{Å}, \beta = 91.45°\) (four molecules in the asymmetric unit).

**X-ray Diffraction Data Collection, Processing, Phasing, and Model Refinement**—X-ray diffraction data were collected at the Advanced Photon Source, that for crystal I on Beamline 24-ID-D, and that for crystal II on Beamline 19-ID, using \(\lambda = 0.9792\,\text{Å},\) corresponding to the absorption peak of selenium. All of the diffraction data were indexed and merged using HKL-3000 (44). Selenium sites in crystal I (located by the Solve routine in Phenix (45)), led to the automatic generation of several peptide fragments by Autosol. A complete model \((R_{free} = 0.29, \text{resolution of } 2.8 \text{Å})\) for crystal I resulted from cycles of model building (Xtalview) (46) and restrained refinement by Refmac (47) and CNS (48). The model from crystal I enabled a molecular replacement structure determination of crystal II using Phaser (49). CNS refinement of the structure from crystal II was refined against data to 2.3 Å resolution using standard restraints on stereochemistry and thermal parameters, and auxiliary restraints on donor-acceptor backbone distances in regions of regular secondary structure. Ramachandran plot statistics came from PROCHECK (50). Atomic coordinates and crystallographic structure factors are in the Protein Data Bank. The figures of protein structure were generated using PyMOL.

**Site-directed Mutagenesis and Analysis**—AgAS R356A, N451A, and V727T mutants were constructed via overlapping PCR of previously described AgAS pENTR (Invitrogen) constructs (51), either wild type (all three) or an existing D621A mutant (R356A and N451A only). The latter double mutants (AgAS:R356A/D621A and AgAS:N451A/D621A) enabled separate analysis of class II activity, as the D621A mutation selectively abrogates AgAS class I activity (21). The resulting pENTR mutant constructs were verified by complete gene sequencing and then transferred to pDEST expression vectors by directional recombination, specifically to pDEST17 for expression as His\(_6\)-tagged fusion proteins. The product outcome mediated by the single mutants was characterized by co-expression with a GGPP synthase in a previously described modular metabolic engineering system (51). The R356A and N451A mutants also were expressed, purified, and kinetically characterized, as previously described, separately analyzing either class II (AgAS:R356A/D621A and AgAS:N451A/D621A) (43) or class I (AgAS:R356A and AgAS:N451A) (42) activity.

**Molecular Dynamics Simulations and Structural Comparisons**—Normal and accelerated molecular dynamics simulations, the threshold energy \(E\) and acceleration factor \(\alpha\) were estimated as \(E = V + 4'N\) and \(\alpha = 4''N/5,\) respectively, where \(V\) is the average dihedral potential from a normal molecular dynamics simulation, and \(N\) is the number of residues in protein. For every simulation, a 100-picosecond minimization step with gradient temperature increasing and a 100-picosecond equilibrium step with NPT condition preceded the productive 10-ns run. Results from molecular dynamics simulations are analyzed primarily with VMD (60) and carma (61). Root mean square deviation-based structural clustering was done according to Ref. 62 with a cutoff of 1 Å.

### RESULTS

**Crystal Structure of AgAS (Protein Data Bank Code 359V)**—The initial 84 amino acid residues coded by the gene for AgAS represent a plastid-targeting segment not present in the mature protein. Hence, by the convention established in the literature, the first residue (methionine) is enumerated as 84 for the truncated recombinant protein used here in the structure determination.

Statistics of data collection and refinement of the AgAS structure are in Table 1. The four monomers of AgAS in the asymmetric unit are nearly identical, exhibiting root mean square deviations of 0.25 Å in pairwise superpositions of C\(_\alpha\) atoms. Electron density for residues 84–109 is absent for each monomer, although Edman sequencing (performed by the Iowa State University Protein Facility) indicated a mixture of polypeptides beginning with residues 84 or 85. Electron density for residues 846–849 is absent in chains A and D, but present, albeit weakly, in chains B and C. All of the monomers have

### Table 1

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Crystal I</th>
<th>Crystal II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (selenium peak)</td>
<td>0.979</td>
<td>0.979</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>50.00–2.50</td>
<td>50.00–2.30</td>
</tr>
<tr>
<td>No. of reflections (total/unique)</td>
<td>271,627/61,432</td>
<td>474,752/148,465</td>
</tr>
<tr>
<td>Completeness (%)(^a)</td>
<td>99.2 (98.4)</td>
<td>97.7 (96.6)</td>
</tr>
<tr>
<td>I/σ(I)(^b)</td>
<td>25.1 (2.4)</td>
<td>18.5 (3.3)</td>
</tr>
<tr>
<td>R(_{merge}),(^c)</td>
<td>0.104 (0.739)</td>
<td>0.092 (0.360)</td>
</tr>
<tr>
<td>R(_{free})(^d)</td>
<td>0.208/0.287</td>
<td>0.196/0.250</td>
</tr>
</tbody>
</table>

\(^a\) Numbers in parentheses refer to the outer 0.04 Å shell of data.

\(^b\) R\(_{merge}\) = \(\Sigma |I–c|/\Sigma I\), where \(I\) is the observed intensity and \(<c>\) is the average intensity calculated from replicate data.

\(^c\) FOM, figure of merit; SAD, single wavelength anomalous dispersion.

\(^d\) R\(_{free}\) = \(\Sigma |F_o| – |F_c|/\Sigma |F_o|\) for reflections contained in the test set held aside during refinement.

\(^e\) Per asymmetric unit.
electron density in the class II active site cleft, which may represent a tripeptide (it is not a component of the crystallization buffer), but this is not included in the reported coordinates, nor in further analyses.

AgAS has three helical domains (termed $\gamma\beta\alpha$; Fig. 2), consistent with prediction (19), and resembling recently reported structures of the class I taxadiene synthase from *Taxus brevifolia* (TbTS) (63) and class II ent-copalyl diphosphate synthase from *Arabidopsis thaliana* (AtCPS) (64). AgAS differs from TbTS and AtCPS in having functional class I and class II active sites rather than a single active site. Putative active sites in AgAS are clefts that include relevant catalytic motifs and correspond to binding sites for substrate analogs in the TbTS and AtCPS structures. The absence of a channel connecting the two active sites in AgAS is consistent with the diffusive transfer of the CPP intermediate (21).

For the most part, the $\gamma$, $\beta$, and $\alpha$-domains include residues from the first, middle, and last thirds of the amino acid sequence, respectively; however, the N-terminal segment is part of all three domains. Lys$^{86}$ and Arg$^{87}$, for instance, are important for class I, but not class II, activity (40), indicating that some of the first 26 residues (electron density absent in the AgAS structure) are part of the $\alpha$-domain. Moreover, the N-terminal segment also makes up the first helix of the $\beta$-domain, before the $\gamma$-domain begins. Hence, similar to the triterpene cyclases (26, 27), as well as TbTS and AtCPS (63, 64), the AgAS $\gamma$-domain falls between the first and second helices of the $\beta$-domain.

The AgAS class II active site exhibits a conformation distinct from other class II systems, specifically loop 482–492 is at the exterior of the protein (hereafter the loop-out conformation), whereas in ligated complexes of AtCPS (64) the corresponding loop penetrates deeply into the active site crevice (the loop-in conformer) (Fig. 3). Despite significantly different environments for conserved residues, the respective loop in the AgAS and AtCPS structures are well ordered, with thermal parameters at or below average values for their respective structures. Residues 402–405, corresponding to the class II $DXXD$ catalytic motif, also are well ordered and in strong electron density.

The class I active site of AgAS has no bound ligand. Residues 846–849, which have weak or absent electron density in each of the four AgAS molecules in the asymmetric unit, correspond to a loop that covers ligated active sites of other class I structures. Residues belonging to the class I catalytic motifs (DDXXD, residues 621–625, and NTE, residues 765–773) are ordered and well placed in electron density.

**Mutagenesis of AgAS**—The identification of the class II active site of AgAS rests largely on similarities in sequence and struc-
that cross over from the γ- to β-domain and could conceivably influence the size of the active site crevice. Our original hypothesis that the basic residues at this position would be directly involved in class II catalysis was largely based on the >1,000-fold reduction in activity observed with mutation of the corresponding histidine in AtCPS to alanine (43). To verify a similar effect in AgAS, alanine was substituted for Arg1756, which also led to a ~1,000-fold decrease in class II catalytic activity, with a relatively small loss of class I catalytic activity again indicative of retention of overall structural integrity (Table 2). Consistent with the hypothesis that GGPP substrate inhibition is synergistic with inhibitory Mg2+ binding (67) and a role for Arg1756 in this (43), AgAS:R356A exhibited no substrate inhibition.

A single-residue switch that substantially alters product outcome was first identified in ent-kaurene synthases, which react with enantiomeric (ent-)CPP (68). In short, replacement of an ethyl group with a hydroxyl (a threonine for isoleucine substitution) redirects ent-kaurene synthases to form pimaradienes (68). A similar effect has been observed in AgAS, where substitution of Ala203 with serine also leads to pimaradiene production. However, this position is four residues N-terminal to that corresponding to the Ile → Thr product switching position in ent-kaurene synthases (42), which is a valine in AgAS. The Val727 side chain also protrudes into the class I active site, so to investigate its effect on product outcome of hydroxyl group introduction at this position, we substituted in threonine. This AgAS:V727T mutant enzyme produces the same mixture of abietadiene double-bond isomers as wild-type AgAS, with no significant increase in pimaradiene content.

**Molecular Dynamics Simulations: Class II Active Site**—Given the large change in conformation of loop 482–492 in our AgAS crystal structure (loop-out) relative to that of the corresponding loop in the substrate analog containing AtCPS crystal structure (loop-in), we were interested in the role of this loop. In particular, we wanted to discover whether the loop-in conformer is accessible in the absence of substrate and what the effect of the loop-in conformer is on the class II active site. This was examined by selected molecular dynamics simulations of our AgAS structure.

Loop 482–492 remains in its starting (loop-out) conformation in normal molecular dynamics simulations. Asp327 (γ-domain) is in contact with Trp400 (β-domain) and blocks loop 482–492 from entering the active-site crevice (Fig. 3). In accelerated simulations, however, loop 482–492 transiently enters the active-site bringing Asp488 in proximity to Val446 and the N terminus of helix 446–457 (Figs. 4 and 5). Note that the lone hydrogen bond between this loop and the rest of the protein in the AtCPS loop-in structure is between the side chain of Asp462 and the backbone amide of Val420, residues that correspond to

### Table 2: Kinetic Analysis

<table>
<thead>
<tr>
<th></th>
<th>Class II activity</th>
<th>Class I activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_m$ (µM)</td>
<td>$k_{cat}$ $s^{-1}$</td>
</tr>
<tr>
<td>WT</td>
<td>0.5 ± 0.1</td>
<td>1.1 ± 0.1</td>
</tr>
<tr>
<td>R356A</td>
<td>6 ± 3</td>
<td>1.1×10$^{-3}$</td>
</tr>
<tr>
<td>N451A</td>
<td>6 ± 4</td>
<td>0.016</td>
</tr>
</tbody>
</table>

* Determined in D621A background to abrogate class I activity.
Asp\textsuperscript{488} and Val\textsuperscript{446} of AgAS, respectively. Phi and psi angles of Phe\textsuperscript{487} undergo abrupt and significant change during the simulation, indicating points of articulation. In addition, movement of Asp\textsuperscript{227} correlates strongly with that of loop 482–492. Hence, Asp\textsuperscript{227} appears to act as a gate in providing loop 482–492 access to the active site.

The binding locus for the pyrophosphate moiety of the substrate (as suggested by the bound analog in the AtCPS structure) is near residues 223–225, suggesting a substrate-induced displacement of Asp\textsuperscript{227} that enables movement of loop 482–492 into the active site. Alternatively, as the accelerated simulation reveals movement toward the loop-in conformation in the absence of substrate, the two conformational states of loop 482–492 may exist at equilibrium. The loop-in conformer, however, puts Lys\textsuperscript{489} proximal to the pyrophosphate binding locus. Hence, substrate binding could shift the equilibrium population in favor of the loop-in conformer.

Normal simulations of the AgAS enzyme-substrate complex employ the substrate (GGPP) with loop 482–492 in its out-and-in conformations. (The loop-in conformation of AgAS was modeled after the Ca coordinates of the corresponding loop of the AtCPS structure and the starting conformation of GGPP after the coordinates of the substrate analog bound to the class II active site of the same structure). Atoms of the substrate exhibit low root mean square fluctuations in the loop-out conformation and high fluctuations in the loop-in conformation (Fig. 6). The isoprenoid moiety of the substrate, however, is hydrated to a greater extent in the loop-out relative to the loop-in conformation (Fig. 7). Cluster analysis reveals one dominant conformational type for the substrate in the loop-out simulation but many conformational types of equal weight in the loop-in simulation. Moreover, distances between the atoms that participate in covalent bond formation or proton transfer during the catalyzed reaction approach those expected for the occurrence of such chemistry among the loop-in conformers but are largely absent among loop-out conformers (Fig. 7). Given that the starting coordinates for the GGPP substrate were derived from the analog bound in the AtCPS crystal structure, which is not appropriately positioned for cyclization, this was a striking result that we further investigated.
As a means of equilibrating the active site about the putative transition state, force restraints were applied to pairs of atoms that undergo covalent bond formation or proton transfer (Fig. 8). The docked transition state complex reveals contacts between planar isoprenoid groups and the side chains of Tyr287 and Phe354, with the side chains of Tyr287 and/or Tyr536 in position to accept a proton from atom C19 after, or in concert with the cyclization reaction. Lys269, Lys489, and Arg272 in the model interact with the pyrophosphate moiety.

Molecular Dynamics Simulations: Class I Active Site—The TbTS complex with a substrate analog (FGG) (63) was the basis for modeling the CPP substrate in the class I active site of AgAS. A normal molecular dynamics simulation of ligand-free and CPP-ligated AgAS reveals substantial reductions in root mean square fluctuations of several regions throughout the entire AgAS structure with ligand bound (Fig. 9). The NTE motif, which is in contact with the pyrophosphate-Mg$^{2+}$/H$^{+}$ moiety of CPP, exhibits less movement than in the absence of CPP. Notable is the reduction in movement of loop 482–492 and residues neighboring and including Asp227. Further analysis, however, reveals no evidence for correlated (or anti-correlated) motions between structural elements of the class I active site and loop 482–492. Hence, the enhanced stability of loop 482–492 may result by chance alone (the absence of a relatively rare and low frequency movement of loop 482–492 in the simulation of the ligated structure).

The simulation of the AgAS-CPP complex did not result in significant conformational changes in the protein, consistent with crystallographic studies of other ligand-free and ligated class I terpene synthases from plants. Ligated and ligand-free forms of bornyl diphosphate synthase (Protein Data Bank codes 1N1B and 1N1Z), isoprene synthase (Protein Data Bank codes 1N1B and 1N1Z), and sesquiterpene synthase (Protein Data Bank codes 1N1B and 1N1Z) reveal substantial differences in conformation.

As a means of equilibrating the active site about the putative transition state, force restraints were applied to pairs of atoms that undergo covalent bond formation or proton transfer (Fig. 8). The docked transition state complex reveals contacts between planar isoprenoid groups and the side chains of Tyr287 and Phe354, with the side chains of Tyr287 and/or Tyr536 in position to accept a proton from atom C19 after, or in concert with the cyclization reaction. Lys269, Lys489, and Arg272 in the model interact with the pyrophosphate moiety.

Molecular Dynamics Simulations: Class I Active Site—The TbTS complex with a substrate analog (FGG) (63) was the basis for modeling the CPP substrate in the class I active site of AgAS. A normal molecular dynamics simulation of ligand-free and CPP-ligated AgAS reveals substantial reductions in root mean square fluctuations of several regions throughout the entire AgAS structure with ligand bound (Fig. 9). The NTE motif, which is in contact with the pyrophosphate-Mg$^{2+}$/H$^{+}$ moiety of CPP, exhibits less movement than in the absence of CPP. Notable is the reduction in movement of loop 482–492 and residues neighboring and including Asp227. Further analysis, however, reveals no evidence for correlated (or anti-correlated) motions between structural elements of the class I active site and loop 482–492. Hence, the enhanced stability of loop 482–492 may result by chance alone (the absence of a relatively rare and low frequency movement of loop 482–492 in the simulation of the ligated structure).

The simulation of the AgAS-CPP complex did not result in significant conformational changes in the protein, consistent with crystallographic studies of other ligand-free and ligated class I terpene synthases from plants. Ligated and ligand-free forms of bornyl diphosphate synthase (Protein Data Bank codes 1N1B and 1N1Z), isoprene synthase (Protein Data Bank codes 1N1B and 1N1Z), and sesquiterpene synthase (Protein Data Bank codes 1N1B and 1N1Z) reveal substantial differences in conformation.

As a means of equilibrating the active site about the putative transition state, force restraints were applied to pairs of atoms that undergo covalent bond formation or proton transfer (Fig. 8). The docked transition state complex reveals contacts between planar isoprenoid groups and the side chains of Tyr287 and Phe354, with the side chains of Tyr287 and/or Tyr536 in position to accept a proton from atom C19 after, or in concert with the cyclization reaction. Lys269, Lys489, and Arg272 in the model interact with the pyrophosphate moiety.

Molecular Dynamics Simulations: Class I Active Site—The TbTS complex with a substrate analog (FGG) (63) was the basis for modeling the CPP substrate in the class I active site of AgAS. A normal molecular dynamics simulation of ligand-free and CPP-ligated AgAS reveals substantial reductions in root mean square fluctuations of several regions throughout the entire AgAS structure with ligand bound (Fig. 9). The NTE motif, which is in contact with the pyrophosphate-Mg$^{2+}$/H$^{+}$ moiety of CPP, exhibits less movement than in the absence of CPP. Notable is the reduction in movement of loop 482–492 and residues neighboring and including Asp227. Further analysis, however, reveals no evidence for correlated (or anti-correlated) motions between structural elements of the class I active site and loop 482–492. Hence, the enhanced stability of loop 482–492 may result by chance alone (the absence of a relatively rare and low frequency movement of loop 482–492 in the simulation of the ligated structure).

The simulation of the AgAS-CPP complex did not result in significant conformational changes in the protein, consistent with crystallographic studies of other ligand-free and ligated class I terpene synthases from plants. Ligated and ligand-free forms of bornyl diphosphate synthase (Protein Data Bank codes 1N1B and 1N1Z), isoprene synthase (Protein Data Bank codes 1N1B and 1N1Z), and sesquiterpene synthase (Protein Data Bank codes 1N1B and 1N1Z) reveal substantial differences in conformation.
3N0F and 3N0G), and 5-epi-aristolochene synthase (Protein Data Bank codes 5EAS and 5EAT) superimpose with overall root mean square deviations of 0.6 Å or less based on at least 525 corresponding Cα atoms. Atoms of CPP had root mean square fluctuations ranging from 0.5–0.9 Å, similar in magnitude to those of GGPP in the loop-out conformation of the class II active site. Atom C16 of CPP exhibits the greatest fluctuation because of the “flipping” of the isoprenoid unit proximal to the pyrophosphate group of CPP. Flipping of this isoprenoid unit infers a mixture of isopimar-15-en-8-yl1 intermediate (Fig. 1) that is not supported by the single dominant stereochemistry of the reaction product. Evidently, the CPP molecule during simulation has too much conformational freedom in the AgAS class I active site. That active site, however, is incomplete in two respects: N-terminal residues (Lys86 and Arg87) critical for catalysis are absent, and loop 846–849, which covers ligated active sites in other class I terpene synthases, is not in place. Given the effects of loop 482–492 on the conformational behavior of GGPP in the class II active site, the missing elements of the AgAS class I active site may be responsible in part for defining a single productive conformer of CPP.

**DISCUSSION**

AgAS has served as a model diterpene synthase for over a decade, yet despite extensive previous mechanistic and mutational analysis, the determination of its structure reported here has further elucidated its activity. Although no substrate analogs are present in this structure, it has been noted that such analogs in the known co-crystal structures often are not in a catalytically relevant conformation (2, 69), including those in the only two other reported diterpene synthase crystal structures (63, 64). Thus, the AgAS structure reported here was followed up by both site-directed mutagenesis and molecular dynamics simulations to probe the inferred structure-function relationships underlying the class I and class II cyclization reactions catalyzed by this bifunctional enzyme.

The wild-type class I active site favors proton and methyl migration from the initial isopimar-15-en-8-yl1 intermediate, resulting primarily in the formation of abietadienes (Fig. 1). The pyrophosphate co-product, presumably still bound to the active site, should favor such carbocation migration through electrostatic attraction (69). We have previously hypothesized that substitution of Ala723 with serine stabilizes the isopimar-15-en-8-yl1 intermediate, allowing deprotonation and the formation of isopimara-7,15-diene as the major product (42). The AgAS structure reported here is consistent with this hypothesis. Moreover, mutation of Val727 to threonine, to put a nearby hydroxyl group in the active site, has no effect on product distribution. Hence, the single-residue switch behavior of position 723 is localized and specific, although difficult to understand without complete knowledge of a fully assembled class I active site.

The AgAS structure forces a revision in the proposed arrangement of catalytic residues in the class II active site (43). Specifically, a conserved asparagine (Asn551 in AgAS) rather than the previously proposed Arg356 activates the catalytic aspartic acid, as not only suggested by the AgAS and AtCPS structures but further demonstrated here by site-directed mutagenesis. On the other hand, our results reaffirm the importance of Arg356 in class II catalysis (Table 2). Inspection of the crystal structure reveals a stacking interaction between the guanidine group of Arg356 and the aromatic ring of Tyr382. The side chain of Arg356 is buried, as is the corresponding histidine in AtCPS. The mutation of Arg356 to alanine likely creates a destabilizing void in the protein, the consequence of which is conformational change to minimize the volume of the void.
One half-turn away from Arg<sup>356</sup> on the same α-helix is a highly conserved aromatic residue, Phe<sup>354</sup> in AgAS, which by simulation stacks against the terminal isoprenoid unit of the substrate (Fig. 8). The relationship of Phe<sup>354</sup> to GGPP is consistent with carbocation stabilization via the aromatic π quadrupoles, despite such stabilization not being built into the molecular dynamics simulation. In addition, such π quadrupole interaction with the π orbitals of the sp<sup>2</sup> hybridized atom C15 of GGPP may bias the double-bonded atom C14 toward accepting a proton from Asp<sup>404</sup> to initiate the reaction.

Although Arg<sup>356</sup> in AgAS and the corresponding His<sup>331</sup> in AtCPS fulfill similar structural roles (planar group stacking with tyrosine), these residues have distinctly different and interchangeable effects on Mg<sup>2+</sup>-dependent inhibition of class II activity. AtCPS is susceptible to Mg<sup>2+</sup>-dependent inhibition, whereas AgAS is not (67). Interchanging residue types also swap susceptibility to Mg<sup>2+</sup>-dependent inhibition, with little to no change in catalytic activity (43). Indeed, this position is conserved as a histidine in all plant class II diterpene cyclases and involved in gibberellin phytohormone biosynthesis, such as AtCPS, but is an arginine in all enzymes dedicated to specialized or secondary metabolism, such as AgAS. The mechanism by which arginine at position 356 abrogates Mg<sup>2+</sup>-dependent inhibition is unclear; however, the electrostatic charge on a buried arginine (as opposed to a neutral histidine) can have long range effects through electrostatic interactions.

Although further work is necessary to fully understand the significance of Arg<sup>356</sup>, the current effort here has revealed a significant role for loop 482–492 in class II activity. First, Lys<sup>489</sup> and Trp<sup>490</sup> are conserved residues, and as is evident from Fig. 8, these residues make substantial contacts with GGPP. Moreover, the loop-in conformer plays a counterintuitive role by increasing the conformational flexibility of GGPP, even as it reduces the volume of the active site crevice. In the loop-out conformation, the substrate adopts a favored conformation, which evidently is much lower in energy than alternative substrate conformations. The loop-in conformation raises the energy of this conformer so that it is on par with alternative conformations. These alternative conformers bring pairs of atoms into contact consistent with the proposed mechanism of cyclization. In principle, given sufficiently long simulation times, we are more likely to observe the productive substrate conformation in the loop-in conformer.

Trp<sup>358</sup> does not interact with GGPP in the class II complex that arises from constrained simulation. Nevertheless, this is an absolutely conserved residue, and the AgAS-W358A mutant enzyme has been shown to exhibit 1000-fold less class II activity and decreased affinity for the 14,15-dihydro-15-azagere-nylgeranyl-5-thiolodiphosphate analog that approximates the initial carbocation of the proposed reaction mechanism (23). Such tryptophan to alanine substitution will leave the mutant enzyme with a substantial internal void and the concomitant loss of energy-favorable nonbonded contacts. Conformational change to reduce the volume of the void is likely, including perturbation of the adjacent catalytic pairing of Asp<sup>404</sup> and Asn<sup>453</sup>.

As noted for TbTS and AtCPS (63, 64), the AgAS structure is consistent with the previously advanced hypothesis that class II diterpene cyclases evolved from the mechanistically similar triterpene cyclases (19), because they all share structural homology (i.e., across the γβ domains). The diterpene synthase structures also are consistent with the previous suggestion that plant diterpene synthases originated from fusion of bacterial class II and I diterpene synthases (28). This ancestral bifunctional diterpene synthase was presumably involved in gibberellin biosynthesis and underwent gene duplication and subfunctionalization to give rise to the separate class II and class I enzymes that carry out such biosynthesis in higher plants (35).

More speculatively, given the ancestral role for diterpene synthases in the extensive plant sterpene synthase family (7), the overall arrangement of the γβα domain structure (Fig. 2) provides a rationale for the observed βα domain structure of most plant class I terpene synthases (70). In particular, only the β domain makes extensive contacts with the α domain, such that loss of the γ domain might have relatively little effect in monofunctional class I terpene synthases. Consistent with such speculation, the γ domain appears to have been independently lost multiple times in plant terpene synthase evolution (71). By contrast, the βα domain interface appears to be important, because it has been shown for AgAS that class I activity for separately expressed α domain requires the presence of the γβ domains (40).

In addition, the presence of the N terminus in the α domain and its passage through the β domain provides a rationale for the observed conservation pattern of the majority of βα domain plant terpene synthases, wherein loss of the γ domain occurred with retention of the corresponding N-terminal sequence, leading to the original “internal” nomenclature for the sequence element corresponding to the γ domain (7). Moreover, the presence of a pair of basic residues at the N terminus important for class I catalysis in AgAS (40) draws some interesting parallels to a similar N-terminal motif found in many βα domain terpene synthases (15, 70) and also is consistent with retention of the N-terminal sequence in γ-domain loss. Thus, coupled to the extensive previous studies of this model diterpene synthase, the AgAS structure not only provides insights into the class II and class I diterpene cyclization reactions catalyzed by this bifunctional enzyme, but also into the evolution of plant terpene synthases more generally.

**Acknowledgment**—We thank Prof. Gaya Amarasinghe (Iowa State University) for collecting diffraction data on crystal 1.

**REFERENCES**


Structural insights into the evolution of terpene synthase function.

Starks, C. M., Back, K., Chappell, J., and Noel, J. P. (1997) Structural basis of the active sites for protonation-initiated and ionization-initiated cycl- 


dyne synthase from the fungus Physcomitrella patens. FEBS Lett. 580, 6175–6181


cyclase. J. Biol. Chem. 275, 2276–2280


Biochemistry 24, 2141–2423


nyl cyclization reactions in biology. Science 277, 1820–1824


Gerrard, J. P., and Croteau, R. (1987) Stereochemistry of the cyclization-rearrangement of geranylgeranyl diphosphate to (−)-copalyl diphosphate and (+)-copalyl diphosphate from the fungus Phaeosphaeria sp. 1.487 and a comparison with cyclases from higher plants. J. Biol. Chem. 275, 2573–2577


copalyl diphosphate and ent-kaurene synthases in Bradyrhizobium japonicum. FEBS Lett. 583, 475–480

Structure of Abietadiene Synthase

129, 15736–15737
Insights into Diterpene Cyclization from Structure of Bifunctional Abietadiene Synthase from *Abies grandis*

Ke Zhou, Yang Gao, Julie A. Hoy, Francis M. Mann, Richard B. Honzatko and Reuben J. Peters


*doi: 10.1074/jbc.M111.337592 originally published online January 4, 2012*

Access the most updated version of this article at doi: 10.1074/jbc.M111.337592

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 69 references, 19 of which can be accessed free at http://www.jbc.org/content/287/9/6840.full.html#ref-list-1