Crystal Structure Determination of Aristolochene Synthase from the Blue Cheese Mold, *Penicillium roqueforti*  

Jonathan M. Caruthers‡§, Igu Kang‡, Michael J. Rynkiewicz‡, David E. Cane¶, and David W. Christianson‡  

From the ‡Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323 and the §Department of Chemistry, Brown University, Providence, Rhode Island 02912  

Received for publication, January 18, 2000, and in revised form, May 22, 2000  
Published, JBC Papers in Press, May 23, 2000, DOI 10.1074/jbc.M000433200

The 2.5-Å resolution crystal structure of recombinant aristolochene synthase from the blue cheese mold, *Penicillium roqueforti*, is the first of a fungal terpenoid cyclase. The structure of the enzyme reveals active site features that participate in the cyclization of the universal sesquiterpene cyclase substrate, farnesyl diphosphate, to form the bicyclic hydrocarbon aristolochene. Metal-triggered carbonation formation initiates the cyclization cascade, which proceeds through multiple complex intermediates to yield one exclusive structural and stereochemical isomer of aristolochene. Structural homology of this fungal cyclase with plant and bacterial terpenoid cyclases, despite minimal amino acid sequence identity, suggests divergence from a common, primordial ancestor in the evolution of terpene biosynthesis.

Aristolochene synthase is a terpenoid cyclase from the blue cheese mold, *Penicillium roqueforti*, that catalyzes the metal-dependent cyclization of farnesyl diphosphate to form the bicyclic hydrocarbon aristolochene (Fig. 1) (1). Farnesyl diphosphate is the universal precursor of myriad cyclic sesquiterpenes, so each sesquiterpene cyclase plays a critical role in governing the structural and stereochemical outcome of its particular cyclation reaction. Accordingly, sesquiterpene cyclase reactions maximize product diversity starting from a minimal substrate pool, indeed, a single substrate, and the structural basis of this catalytic diversity comprises a growing question at the interface of chemistry and biology.

Aristolochene synthase is a 38-kDa monomeric sesquiterpene cyclase that has been cloned (2) and overexpressed (3) in *Escherichia coli*. Numerous enzymological studies of *P. roqueforti* and *Aspergillus terreus* aristolochene synthases using stereospecifically labeled substrates (4–6), a mechanism-based inhibitor (7), and the anomalous substrate (7R)-6,7-dihydrofarnesyl diphosphate (8) indicate a complex cyclization cascade proceeding through at least two discrete intermediates. Aristolochene formation is the first committed step in the biosynthesis of a large group of sesquiterpenoid fungal toxins, the most lethal of which is the novel bis-epoxide PR-toxin (4). Interestingly, the (+)-enantiomer of aristolochene is generated by the fungi *P. roqueforti* and *A. terreus*, but the (−)-enantiomer is generated by the plants *Aristolochia indica* (9) and *Bixa orellana* (10). Accordingly, each aristolochene synthase must provide a different template for binding the flexible polyisoprenoid substrate and subsequent intermediates in productive conformations leading to correct stereoisomer formation.

The diastereomeric sesquiterpene epi-aristolochene (4-epiemerophila-9,11-diene) has been identified in tobacco (*Nicotiana tabacum*) (11–13) and results from the cyclization of farnesyl diphosphate by epi-aristolochene synthase (Fig. 1) (14). This enzyme catalyzes the cyclization of farnesyl diphosphate by a mechanism similar in some respects to that of *P. roqueforti* aristolochene synthase despite only 16% amino acid sequence identity (15). However, the two enzymes must differ in how they precisely chaperone the conformation(s) of cyclization intermediates so as to control the exclusive formation of their respective diastereomeric sesquiterpene products, aristolochene or epi-aristolochene.

To establish the structural basis for the formation of different product diastereomers by the fungal and plant cyclases, we now report the crystal structure of native recombinant *P. roqueforti* aristolochene synthase at 2.5-Å resolution and compare this structure with that of *N. tabacum* epi-aristolochene synthase (16). Importantly, both enzymes contain conserved “aspartate-rich” DDXV(D/E) signature sequences that presumably bind catalytically obligatory Mg$^{2+}$ ion(s) (2, 17), and additional signature segments are indicated in structural comparisons. With more than 300 cyclic sesquiterpenes identified to date, all deriving from the universal precursor farnesyl diphosphate (18–20), the protein structural basis for the precise control of product structure and stereochemistry in farnesyl diphosphate cyclization remains a critical issue at the interface of biosynthetic chemistry and structural biology (21–24).

**EXPERIMENTAL PROCEDURES**

Expression and Purification of Aristolochene Synthase—From a glycerol stock of *E. coli* BL21(DE3) carrying pZW04, 5 ml of seed culture in LB-ampicillin was cultured at 37 °C overnight. 500 ml of prewarmed (30 °C) LB-ampicillin media in a 3-liter flask was inoculated with this overnight seed, giving an initial A at 600 nm of 0.05, and further incubated at 30 °C until the A reached 0.7–1.0. Isopropyl-1-thio-β-D-galactopyranoside was added to 0.5 mM final concentration to induce production of recombinant aristolochene synthase. Cells were harvested after a 4-h induction period, resuspended in 50 ml of cell lysis buffer (5 mM EDTA, 5 mM β-mercaptoethanol, 20 mM Tris (pH 7.5)), and sonicated to disrupt cells. After three cycles
Five heavy atom derivatives were used to phase the initial electron density map (Table I). Initial heavy atom sites were determined in subsequently calculated difference Fourier maps; heavy atom positions were refined with the program PHASES (27). The model was fit into an electron density map calculated with solvent-flattened non-crystallographic symmetry averaged phases (28) at 3.3-Å resolution. Subsequent refinement and rebuilding of the native model was achieved with CNS (29) and O (30), respectively. Individual B-factors were refined, and a bulk-solvent correction was applied. This model was then refined against the 2.5-Å resolution synchrotron data, and a total of 275 well ordered solvent molecules were located in the asymmetric unit. The quality of the model was improved by gradually releasing noncrystallographic symmetry constraints into appropriately weighted restraints as judged by R_{free}. Refinement statistics are recorded in Table I. The final model has excellent stereochemistry with only 1 of 298 total residues (Ala-186) adopting a disallowed backbone conformation.

RESULTS AND DISCUSSION

Structure and Evolutionary Relationships—Aristolochene synthase from P. roqueforti adopts the class I α-helical terpenoid cyclase fold (21, 31), in which 6 α-helices (of 11 total helices) surround a large active site cleft approximately 15 Å wide by 20 Å deep (Fig. 2). Polypeptide loops connecting these helices are relatively short (average 5 residues) on the side of the protein distal to the active site and relatively long (average 12 residues) on the side of the protein proximal to (that is, surrounding) the active site cavity. This feature is a consequence of the helix packing arrangement necessary to form a roughly conical cleft. The longest loops surrounding the active site cavity are the “180°” loop (Asp-184—Ile-190), which forms one wall of the active site cleft, and a disordered “250°” loop (Asp-252—Leu-265) that flanks the mouth of the active site. Additionally, a 40-residue segment is disordered at the N terminus (Met-1—Thr-40), and a 5-residue segment (Thr-336—Asn-341) is disordered at the C terminus.

The conserved aspartate-rich segment (DDVIE, beginning with Asp-115) is located on the upper wall of the active site cleft on helix C (Figs. 2 and 3). This is the coordination site for the Mg²⁺ ion(s) necessary to trigger initial carbocation formation. Notably, a single Sm³⁺ ion binds to Asp-115 in the SmCl₃ derivative used for phasing the initial electron density map. As found in the structure determination of avian farnesyl diphosphate synthase, Sm³⁺ is a good analogue for Mg²⁺, one Sm³⁺ ion binds to each of two aspartate-rich segments in this syn- thase (32). One of the aspartate-rich segments of farnesyl diphosphate synthase (DDIMD, beginning with Asp-117) corresponds to the DDVIE segment of aristolochene synthase (Fig. 4), so this segment in aristolochene synthase is similarly implicated in binding the catalytically obligatory metal ion(s). In general, a trivalent metal ion binds ~10³-fold more tightly than a divalent metal ion to a carboxylate cluster, which is defined as 3–4 carboxylate groups located in a shallow cleft on a protein surface (33). Because carboxylate clusters discriminate more on the basis of ionic charge rather than ionic radius (33), this rationalizes the facile binding of Sm³⁺ to the Mg²⁺-binding aspartate-rich segment of aristolochene synthase.

The α-barrel fold of P. roqueforti aristolochene synthase bears a striking resemblance to that of N. tabacum epi-aristolochene synthase despite only 16% sequence identity. The N. tabacum synthase consists of a 281-residue C-terminal catalytic domain and a 240-residue N-terminal domain of unknown function (16). Here too, the catalytic domain consists of 11 major helices arranged with identical topology and surrounding the active site cleft of similar diameter and hydrophobicity, as found in the P. roqueforti cyclase. The cyclization domains of the fungal and plant cyclases are of the same approximate shape and dimensions, approximately 50 × 60 × 30 Å (Fig. 4).

Despite minimal sequence identities (16–21%), the α-helical folds of a bacterial sesquiterpene cyclase, penatalenene synthase

---

1 The abbreviation used is: PAGE, polyacrylamide gel electrophoresis.
and avian farnesyl diphosphate synthase (32), are also highly similar to the 
α-helical fold of P. roqueforti aristolochene synthase (Fig. 4). This structural homology is reminiscent of suggestions that enzymes catalyzing successive steps in a biosynthetic pathway may evolve with similar structure through divergence, regardless of negligible amino acid identity; function evolves more rapidly than sequence, and sequence evolves more rapidly than tertiary structure (34). Despite the catalysis of strikingly different cyclization cascades, further structural homology between the fungal aristolochene synthase and terpenoid synthases from bacterial, plant, and animal species suggests a common primordial ancestor early in the evolution of terpene biosynthetic pathways prior to the divergence of plants, animals, bacteria, and fungi (Fig. 4).

### Structural Inferences on the Cyclization Mechanism

The active site of P. roqueforti aristolochene synthase is a 20-Å deep cleft lined with many hydrophobic aliphatic and aromatic residues that define a unique contour. Accordingly, this cleft serves as a template to bind the flexible farnesyl diphosphate substrate in the unique productive conformation leading to aristolochene formation as outlined by Cane and colleagues (4–8) (Fig. 5). Enzyme crystallized in the presence of 16 mM farnesol reveals electron density in the active site consistent with disordered binding of the substrate analogue (data not shown), indicating that coordination of the substrate diphosphate group to metal(s) bound by the aspartate-rich sequence plays an important role in anchoring the substrate conformation (35). This is consistent with site-directed mutagenesis

### Table I

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Resolution</th>
<th>Reflections (measured/unique)</th>
<th>Completeness (overall/outer shell)</th>
<th>( R_{\text{merge}}^{a} ) (overall/outer shell)</th>
<th>( R_{\text{iso}}^{b} )</th>
<th>Number of sites</th>
<th>Phasing power</th>
<th>Figure of merit (3.3 Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>2.5</td>
<td>166088/46972</td>
<td>90.1/86.0</td>
<td>0.074/0.395</td>
<td></td>
<td></td>
<td></td>
<td>0.527</td>
</tr>
<tr>
<td>SmCl₃</td>
<td>3.8</td>
<td>58384/14731</td>
<td>97.4/91.5</td>
<td>0.101/0.212</td>
<td>0.117</td>
<td>2</td>
<td>1.61</td>
<td></td>
</tr>
<tr>
<td>TMLA</td>
<td>4.0</td>
<td>30162/12923</td>
<td>88.6/90.9</td>
<td>0.119/0.232</td>
<td>0.117</td>
<td>2</td>
<td>1.35</td>
<td></td>
</tr>
<tr>
<td>AuCl₃</td>
<td>4.4</td>
<td>52123/7563</td>
<td>83.3/85.4</td>
<td>0.105/0.294</td>
<td>0.169</td>
<td>2</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td>TERPT&lt;sup&gt;d&lt;/sup&gt;</td>
<td>5.0</td>
<td>32344/5678</td>
<td>84.7/88.0</td>
<td>0.108/0.309</td>
<td>0.139</td>
<td>2</td>
<td>1.43</td>
<td></td>
</tr>
<tr>
<td>PtCl₄</td>
<td>3.8</td>
<td>40698/12620</td>
<td>84.0/85.7</td>
<td>0.113/0.300</td>
<td>0.130</td>
<td>1</td>
<td>1.39</td>
<td></td>
</tr>
</tbody>
</table>

### Refinement

<table>
<thead>
<tr>
<th>Structure</th>
<th>Atoms</th>
<th>Solvent molecules</th>
<th>Resolution</th>
<th>No. reflections (work/free)</th>
<th>( R_{\text{work}}^{c} )</th>
<th>Root mean square deviations</th>
<th>Phasing power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Native</td>
<td>4692</td>
<td>275</td>
<td>2.5</td>
<td>39475/4420</td>
<td>0.247/0.273</td>
<td>1.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>18.8&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> \( R_{\text{merge}}^{a} = \frac{\sum|I_i - \langle I_i \rangle|}{\sum |I_i|} \), where \( I_i \) is the intensity measurement for reflection \( i \), and \( \langle I_i \rangle \) is the mean intensity for reflection \( i \) from replicate data.

<sup>b</sup> \( R_{\text{iso}}^{b} = \frac{\sum |F_{PH} - |F_p||}{\sum |F_{PH}|} \), where \( F_{PH} \) and \( F_p \) are the derivative and native structure factor amplitudes, respectively.

<sup>c</sup> TMLA, trimethyllead(II) acetate.

<sup>d</sup> TERPT, chloro(2,2',6,2''-terpyridine)platinum(II).

---

**Fig. 2.** Electron density map of the carboxylate cluster. Simulated annealing omit map of Phe-112—Met-121, which contains the aspartate-rich segment on the upper wall of the active site cleft. The map is contoured at 3σ, and residues in the carboxylate cluster are indicated.

**Fig. 3.** Ribbon plot of P. roqueforti aristolochene synthase. The “aspartate-rich” segment DDVIE, beginning with Asp-115, is located on the upper wall of the active site cleft (red). N and C termini are indicated. Figure was prepared with MOLSCRIPT (41) and Raster 3D (42, 43).
studies of trichodiene synthase, where alteration of metal binding residues resulted in the formation of aberrant cyclization products reflecting alternative substrate conformations (36).

Enzymological studies with inhibitors and isotopically labeled substrates by Cane and colleagues (4–8) demonstrate that a unique conformation of the farnesyl diphosphate substrate is required to form the correct aristolochene stereoisomer, with compelling evidence for intermediates germacrene A and the eudesmane cation; the resulting mechanistic sequence is outlined in Fig. 5. Because the enzyme active site is a template that enforces the correct substrate conformation, we have modeled the pathway of aristolochene biosynthesis that incorporates this mechanistic information within the context of the enzyme active site (Fig. 6). Models of enzyme-substrate, -intermediate, and -product complexes were created by optimizing individual substrate and intermediate conformations using energy minimization routines in the program MacroModel (37).

Prior to energy minimization, molecular conformations were moderately restrained such that atoms involved in the reaction coordinate of carbon-carbon bond-making and bond-breaking steps (as illustrated in Fig. 5) were properly oriented to accommodate the required stereochemistry for each step. For example, the precatalytic substrate conformation was calculated by restraining the C-1 and C-10 atoms to be 3.3 ± 0.2 Å apart with the proper orientation of the π-electrons at C-10 toward the C-1–O σ* antibonding orbital in anticipation of the initial C-1–C-10 bond-making step to yield the germacrene A intermediate with S stereochemistry at the C-10 atom (the substrate numbering scheme is indicated in Fig. 5). The energy-minimized substrate, intermediate, and product conformations were then manually docked into the enzyme active site cleft. The substrate orientation was fixed so that the diphosphate leaving group could interact with metal ion(s) bound to the aspartate-rich sequence (as indicated by Sm³⁺ binding in the SmCl₃ derivative). This particular interaction defined a unique precatalytic orientation of the substrate in the enzyme active site, there was only one way to fit the unique precatalytic substrate conformation required for the correct cyclization stereochemistry into the active site while constrained by specific diphosphate-metal-protein interactions. Indeed, the binding of the required substrate conformation in the active site was highly complementary to the contour of the active site, consistent with the role of the terpenoid cyclase as a template to enforce the proper starting conformation of the flexible isoprenoid substrate. The geometries of enzyme-substrate, -intermediate, and -product complexes were optimized by subjecting the manually docked complexes to conjugate gradient energy minimization with no experimental energy term in CNS (29). Two hundred cycles of minimization were performed with nonbonded interaction cutoff = 13 Å and dielectric constant = 1.

In our model of the cyclization reaction (Fig. 6), the substrate isopropenyl group is positioned within a hydrophobic surface formed by Phe-112 and Phe-178 and also Leu-108, Leu-111, and Gly-205. Farnesyl diphosphate is positioned such that the C-10–C-11 π bond is oriented for attack at C-1 subsequent to (or concurrent with) metal-triggered diphosphate departure; metal ion(s) are liganded by residues in the aspartate-rich motif. After loss of the H12 proton, the germacrene A intermediate results with S stereochemistry. Enzymological studies with isotopically labeled substrates indicate that it is exclusively the cis-methyl group (C-12) that undergoes deprotonation (6); based on analysis of the enzyme active site, there is no apparent candidate for an enzyme general base to assist this deprotonation step. Possibly, the diphosphate leaving group could capture this proton.
Aristolochene Synthase Structure

Subsequent protonation of the C-6 atom of germacrene A yields the tertiary C-7 carbocation, which is then attacked by the C-2–C-3 π-bond to form the eudesmane cation. Potentially, this could occur in concerted fashion. The conformation of germacrene A is enforced by Phe-112, ensuring that the proper face of the C-2–C-3 π-bond undergoes reaction to yield the proper stereochemistry of the eudesmane cation. Accordingly, the enzyme serves as a template throughout the cyclization cascade and chaperones the conformations of reactive intermediates once the reaction is underway. Inspection of the model suggests that the phenolic hydroxyl group of Tyr-92 is suitably oriented to accept this proton, which completes the proposed role of Tyr-92 as a general acid/base.

In conclusion, the four key features of the aristolochene synthase structure and mechanism are representative of those features of the eudesmane synthase (right) (OPP = diphosphate leaving group). The atomic numbering scheme of farnesyl diphosphate is adopted for all structures depicted in the P. roqueforti aristolochene synthase mechanism, and specific atoms discussed in the text are numbered accordingly. Analysis of the aristolochene synthase structure suggests that Tyr-92 is the putative general acid/base.

Active site aromatic residues appear to be optimally positioned to stabilize carbocation intermediates through cation-π interactions (39). Specifically, Phe-178 and Phe-112 appear to be appropriately oriented (40) to stabilize developing partial positive charge on C-1 in the first step of catalysis, and Trp-333 is appropriately oriented to stabilize positively charged C-3 in the eudesmane cation. Cation-π interactions comprise a highly effective chemical strategy for catalysis: electrostatic stabilization of high energy carbocation intermediates can be achieved without risk of quenching the carbocation and annihilating the protein catalyst, as would be the case if stabilization were achieved with nucleophilic protein residues. Additional electrostatic stabilization may result from the negative electrostatic surface potential present in the active site cleft (data not shown).

The three-dimensional structure of P. roqueforti aristolochene synthase provides some clues regarding the stereospecificity of the cyclization cascade that distinguish this cyclase from N. tabacum epi-aristolochene synthase: why can’t the P. roqueforti cyclase make epi-aristolochene? According to our model, the key determinant of aristolochene stereochemistry at C-10 is the orientation of the C-10–C-11 π bond relative to the C-1 diphosphate leaving group in the first step of catalysis. The precatalytic farnesyl diphosphate conformation required by N. tabacum epi-aristolochene synthase (Fig. 5) is not compatible with the contour of the P. roqueforti aristolochene synthase active site, as inferred from the model of the aristolochene synthase-farnesyl diphosphate complex in Fig. 6. The substrate isopropenyl group is nestled within a hydrophobic niche formed by Phe-112 and Phe-178 (these residues are not conserved in N. tabacum epi-aristolochene synthase). The key determinant of product stereochemistry at C-2 and C-3 is the orientation of the C-2–C-3 π bond in the cyclization of the germacrene A intermediate: in P. roqueforti aristolochene synthase this orientation is fixed by Phe-112 to yield the bicyclic eudesmane cation with a trans-decalin configuration, whereas in N. tabacum epi-aristolochene synthase this orientation must be fixed to yield the bicyclic eudesmane cation with a cis-decalin configuration. A cis-decalin conformation for this intermediate is incompatible with the contour of the P. roqueforti enzyme active site. Therefore, although the two enzymes cyclize farnesyl diphosphate to form hydrocarbon products of identical connectivity, the two enzymes produce very different active site contours in order to properly template and chaperone reactive conformations in catalysis.

In conclusion, the four key features of the aristolochene synthase structure and mechanism are representative of those

![Cyclization sequences. Mechanisms of P. roqueforti aristolochene synthase (left) and N. tabacum epi-aristolochene synthase (right) (OPP = diphosphate leaving group). The atomic numbering scheme of farnesyl diphosphate is adopted for all structures depicted in the P. roqueforti aristolochene synthase mechanism, and specific atoms discussed in the text are numbered accordingly. Analysis of the aristolochene synthase structure suggests that Tyr-92 is the putative general acid/base.](http://www.jbc.org/)

Downloaded from http://www.jbc.org/ by guest on September 24, 2017
that must characterize the greater family of terpenoid cyclases: (a) the terpenoid cyclase provides a hydrophobic template with the proper contour for binding the flexible substrate in a catalytically productive conformation; (b) initial carbocation formation is triggered by coordination of the diphosphate leaving group to Mg$^{2+}$ or Mn$^{2+}$; (c) the conformations of reactive carbocation intermediates are chaperoned through a complex reaction sequence by the active site template; and (d) the hydrophobic active site cleft sequesters these intermediates from bulk solvent, and aromatic residues stabilize carbocation intermediates through local cation-π interactions as well as overall negative electrostatic surface potentials.

**Fig. 6.** Structure-based mechanism of *P. roqueforti* aristolochene synthase. Models of the enzyme complexed with substrate, intermediates, and product are shown; salient mechanistic details are outlined in the text and appear schematically in Fig. 5. Briefly, farnesyl diphosphate binds in the unique productive conformation prior to the departure of the diphosphate leaving group (A). The initial cyclization yields the germacrene A intermediate through formation of the C-1–C-10 bond (B) (the diphosphate leaving group is not shown for clarity). Protonation of C-6 by Tyr-92 accompanied by C-2–C-7 bond formation closes the 10-membered ring of germacrene A to form the bicyclic eudesmane cation intermediate (C). A 1,2-hydride transfer, accompanied by a C-14 methyl migration and the elimination of Hβ8, yield aristolochene (D). Figure prepared with AVS (44).
Acknowledgments—We thank M. Capel for help with data collection at Brookhaven National Laboratory, beamline X12B, and we thank J. D. Cox, C. A. Lesburg, C. M. Paschall, and T. Stams for helpful discussions.

REFERENCES
Crystal Structure Determination of Aristolochene Synthase from the Blue Cheese Mold, *Penicillium roqueforti*
Jonathan M. Caruthers, Ilgu Kang, Michael J. Rynkiewicz, David E. Cane and David W. Christianson

doi: 10.1074/jbc.M000433200 originally published online May 23, 2000

Access the most updated version of this article at doi: 10.1074/jbc.M000433200

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 42 references, 7 of which can be accessed free at http://www.jbc.org/content/275/33/25533.full.html#ref-list-1