

Crystal Structure of Pentalenene Synthase: Mechanistic Insights on Terpenoid Cyclization Reactions in Biology

Charles A. Lesburg, Guangzhi Zhai, David E. Cane, David W. Christianson*

- modeling were carried out against all data between 20.0 and 2.25 Å. Simulated-annealing omit maps were used to confirm model building. Water molecules were added with ARP [V. X. Lamzin and K. S. Wilson, *Acta Crystallogr. D* **49**, 127 (1993)], and X-PLOR (29). Our model consists of residues 36 to 521 and 533 to 548, two Mg²⁺ ions, and 270 water molecules; 95% of the residues are in the most favored regions of the Ramachandran plot, and none are in disallowed regions.
13. A. E. Aleshin, C. Hoffman, L. M. Firsov, R. B. Honzatko, *J. Mol. Biol.* **238**, 575 (1994).
 14. M. Juy *et al.*, *Nature* **357**, 89 (1992).
 15. L. C. Tarshis, M. Yan, C. D. Poulter, J. C. Sacchettini, *Biochemistry* **33**, 10871 (1994); L. C. Tarshis, P. J. Proteau, B. A. Kellogg, J. C. Sacchettini, C. D. Poulter, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15018 (1996).
 16. D. E. Cane, Q. Xue, B. C. Fitzsimons, *Biochemistry* **35**, 12369 (1996).
 17. The uncomplexed TEAS structure was initially refined to 2.8 Å against data collected from a crystal grown in the presence of 2 mM FHP. Electron density at the active site allowed unambiguous modeling of FHP, the A-C and J-K loops, and nine additional residues at the NH₂-terminus. The refined TEAS·FHP model consists of residues 17 to 548, three Mg²⁺ ions, 150 water molecules, and one FHP molecule.
 18. The uncomplexed TEAS structure was refined against data collected from a crystal grown in the presence of 1 mM F₃-FPP. Strong electron density initially appeared for the diphosphate moiety of F₃-FPP; as refinement proceeded, density also appeared for additional residues at the NH₂-terminus and in the J-K loop. The refined TEAS·F₃-FPP model consists of residues 21 to 523 and 529 to 548, 150 water molecules, three Mg²⁺ ions, and one F₃-FPP molecule.
 19. K. Back, C. M. Starks, J. P. Noel, J. Chappell, in preparation.
 20. D. A. Dougherty, *Science* **271**, 163 (1996).
 21. K. M. Biswas and A. H. Jackson, *J. Chem. Soc. Perkin Trans. I* **11**, 1981 (1989) and M. I. Abdullah, A. H. Jackson, P. P. Lynch, K. A. F. Record, *Heterocycles* **30**, 317 (1990).
 22. M. R. Wildung and R. Croteau, *J. Biol. Chem.* **271**, 9201 (1996).
 23. X. Lin, M. Hezari, A. E. Koepf, H. G. Floss, R. Croteau, *Biochemistry* **35**, 2968 (1996).
 24. RIBBONS, M. Carson and C. E. Bugg, *J. Mol. Graphics* **4**, 121 (1986); MOLSCRIPT, P. Kraulis, *J. App. Crystallogr.* **24**, 946 (1991); RASTER3D, D. Bacon and W. F. Anderson, *J. Mol. Graphics* **6**, 219 (1988); GRASP, A. Nicholls, K. A. Sharp, B. Honig, *Proteins Struct. Funct. Genet.* **11**, 281 (1991).
 25. Z. Otwinowski, in *Data Collection and Processing*, L. Sawyer, N. Isaacs, S. Bailey, Eds. (CCP4 Study Weekend, SERC Daresbury Laboratory, Warrington, UK, 1993), pp. 56–62.
 26. ———, *ML-PHARE* (CCP4, SERC Daresbury Laboratory, Warrington, UK, 1991).
 27. K. Cowtan, *DM* (Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography **31**, 34 (1994).
 28. T. A. Jones, J. Y. Zou, S. W. Cowan, M. Kjeldgaard, *Acta Crystallogr.* **A47**, 110 (1991).
 29. A. T. Brünger, *X-PLOR Version 3.1—A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992), pp. 187–207.
 30. We thank D. W. Christianson for communicating results prior to publication; C. D. Poulter for providing the FHP and F₃-FPP; J. Kyte for discussion; M. Bowman, M. Harrington, and D. Wild for technical assistance; A. Bilwes, M. Bowman, R. Ranganathan, and S. Redford for assistance during data collection at beamline 7-1 at the Stanford Synchrotron Radiation Laboratory (SSRL); the staff at SSRL for assistance and advice. Supported in part by the Lucille P. Markey Charitable trust and NIH grant GM54029 (J.P.N.); NSF grant IBW-9408152 (J.C.); the Chapman Foundation and NIH training grant GM07240, administered through the University of California, San Diego (C.M.S.). Work performed at SSRL was supported by the NIH and DOE. Coordinates for uncomplexed TEAS have been deposited in the Brookhaven Protein Data Bank, code 5EAS.

11 April 1997; accepted 26 August 1997

The crystal structure of pentalenene synthase at 2.6 angstrom resolution reveals critical active site features responsible for the cyclization of farnesyl diphosphate into the tricyclic hydrocarbon pentalenene. Metal-triggered substrate ionization initiates catalysis, and the α -barrel active site serves as a template to channel and stabilize the conformations of reactive carbocation intermediates through a complex cyclization cascade. The core active site structure of the enzyme may be preserved among the greater family of terpenoid synthases, possibly implying divergence from a common ancestral synthase to satisfy biological requirements for increasingly diverse natural products.

Sesquiterpenes comprise a group of natural products secreted by marine and terrestrial plants, fungi, and certain microorganisms. The structural diversity and stereochemical complexity of the C₁₅-isoprenoid skeletons of these metabolites are remarkable. Indeed, of the more than 300 cyclic sesquiterpenes that have been characterized to date, each is derived from a common acyclic precursor, farnesyl diphosphate (1), in a reaction catalyzed by a sesquiterpene cyclase (2). Many cyclic sesquiterpenes exhibit useful medicinal properties and have been essential components of the pharmacopoeia since times of antiquity. For example, the sesquiterpenes furanoeudesma-1,3-diene and curzarene are responsible for the analgesic effects of myrrh by interacting with brain opioid receptors, thereby explaining the prescription of myrrh

for use as a pain killer by Pliny the Elder, Hippocrates, and their predecessors (3).

In addition to the 15-carbon sesquiterpenes, the greater family of terpenoids includes the 10-carbon monoterpenes derived from geranyl diphosphate, the 20-carbon diterpenes derived from geranylgeranyl diphosphate, the 25-carbon sesterterpenes derived from geranylarnesyl diphosphate, and the 30-carbon triterpenes (sterols) derived from two farnesyl diphosphate molecules. Terpenoid cyclases (also known as terpenoid synthases) perform critical biosynthetic tasks in metabolic pathways as diverse as cholesterol biosynthesis in mam-

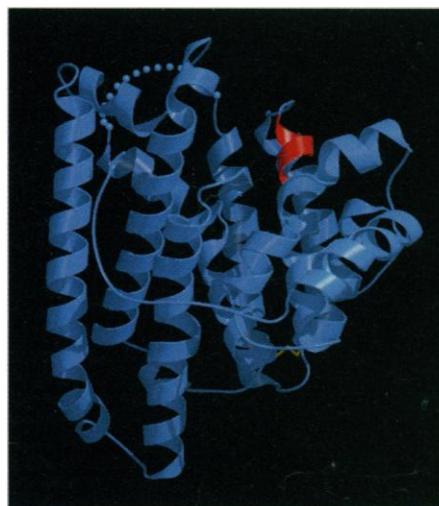


Fig. 1. Ribbon plot (25) of pentalenene synthase. The mouth of the active site cavity opens toward the top of the figure, and the aspartate-rich segment beginning with Asp⁸⁰ is red. The dotted line is the disordered Phe¹⁵⁸-Asp¹⁶⁴ loop, and the Cys¹²⁸-Cys¹³⁶ disulfide linkage is yellow.

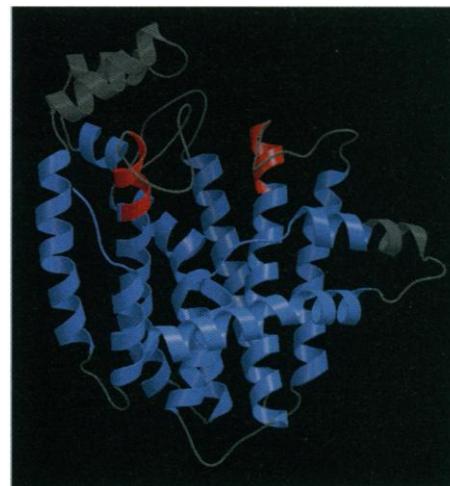


Fig. 2. Ribbon plot (25) of farnesyl diphosphate synthase (8) (PDB accession code 1FPS). The core terpenoid synthase structure shared with pentalenene synthase is blue; the two aspartate-rich segments [beginning with Asp²⁵⁷ (left) and Asp¹¹⁷ (right)] are red. The Asp¹¹⁷ segment of farnesyl diphosphate synthase aligns with the Asp⁸⁰ segment of pentalenene synthase (Fig. 1). Both substrates of this enzyme, isopentenyl diphosphate and dimethylallyl diphosphate, are believed to bind to the two aspartate-rich segments through bridging magnesium ions.

Table 1. Summary of x-ray crystal structure determination. Crystals of pentalenene synthase form as short hexagonal rods and belong to space group $P6_3$ with hexagonal unit cell dimensions of $a = b = 179.8 \text{ \AA}$, $c = 56.6 \text{ \AA}$; two 38-kD monomers reside in the asymmetric unit and are related by two-fold noncrystallographic symmetry (NCS) (7). Diffraction data from pentalenene synthase crystals were collected at room temperature on an R-AXIS IIc image plate area detector, and intensity data integration and reduction were performed with MOSFLM (19) and CCP4 (20), respectively. For phase determination by multiple isomorphous replacement (MIR), initial heavy atom positions were determined in difference Patterson maps and refined with the program MLPHARE (20, 21). The model was fit into an electron density map calculated with solvent-flipped and NCS-averaged MIR phases extended to 3.3 Å resolution with SOLOMON (20, 22). Subsequent refinement and rebuilding of the native model was performed with X-PLOR (23) and O (24), respectively. Group B factors (one main chain and one side chain B factor per residue) were refined and a bulk solvent correction was applied. Strict NCS constraints were maintained as judged by R_{free} ; refinement at 3.3 Å resolution converged to a crystallographic R factor of 0.215 ($R_{\text{free}} = 0.277$). A crystal derivatized with 1.0

mM trimethyllead acetate diffracted to 2.6 Å resolution when flash-frozen at the Cornell High Energy Synchrotron Source (CHESS, beamline A-1, $\lambda = 0.91 \text{ \AA}$). Although this crystal was nonisomorphous with native crystals (at room temperature), the 3.3 Å resolution model of pentalenene synthase served as the starting point for rigid-body refinement, followed by iterative rounds of simulated annealing refinement and rebuilding against the 2.6 Å resolution data with X-PLOR (23) and O (24), respectively. Restrained individual B factors were refined and a bulk solvent correction was applied. The quality of the model was improved in the final stages of refinement by releasing the NCS constraints into appropriately weighted restraints as judged by R_{free} . Refinement converged smoothly to a final crystallographic R factor of 0.198 ($R_{\text{free}} = 0.273$). Disordered segments in the final model include Pro²-Gln³, Phe¹⁵⁸-Asp¹⁶⁴, and Arg³¹⁴-His³³⁷ at the COOH-terminus. The final model has excellent stereochemistry with no residues adopting unfavorable backbone conformations. Since there are no significant structural differences between the 3.3 Å resolution and the 2.6 Å resolution models, and since the lead binding site is removed from the active site (lead makes an interlattice contact between two protein molecules), active site features of the 2.6 Å resolution structure represent those of the native enzyme.

	Native	TMLA-1*	UO ₂ (NO ₃) ₂	TMLA-2*	PHMPS*	TERPT*	K ₂ OsCl ₆	Hg(OAc) ₂
Data collection								
Resolution (Å)	3.3	2.6	3.5	3.5	3.4	5.0	5.0	5.0
Reflections								
Total (N)	38086	70683	19460	18734	41827	9470	6530	9287
Unique (N)	14197	28128	12093	11831	13848	4272	3671	4057
Completeness (%)	89.8	87.5	91.5	89.0	96.7	93.0	80.2	86.7
Number of sites		1	5	5	1	1	1	2
R_{sym}^{\dagger}								
Overall	0.087	0.056	0.065	0.132	0.159	0.066	0.059	0.117
Outer shell	0.245	0.158	0.231	0.449	0.581	0.116	0.140	0.272
Phasing (15 – 3.5 Å)								
$R_{\text{iso}}^{\ddagger}$			0.232	0.185	0.372	0.168	0.255	0.210
Phasing power §			1.11	1.03	0.60	0.71	0.80	1.17
Figure of merit	0.506							
Refinement								
Resolution (Å)	20 – 3.3	20 – 2.6						
Protein atoms (N)	2177	2425						
Solvent atoms (N)	0	33						
Reflections								
work (N)	12498	25158						
free (N)	636	1342						
$R_{\text{crist}}^{\parallel}$	0.215	0.198						
$R_{\text{free}}^{\parallel}$	0.277	0.273						
rms deviations								
Bonds (Å)	0.014	0.016						
Angles (°)	1.7	1.8						

*TMLA, trimethyllead acetate; PHMPS, *p*-hydroxymercuri(II)phenylsulfonate; TERPT, chloro(2,2':6',2''-terpyridine)platinum(II). $\dagger R_{\text{sym}} = \sum_i |I_i - \langle I_i \rangle| / \sum_i \langle I_i \rangle$, where I_i is the intensity measurement for reflection i , and $\langle I_i \rangle$ is the mean intensity calculated for reflection i from replicate data. $\ddagger R_{\text{iso}} = \sum_i |F_{\text{PH}} - F_{\text{P}}| / \sum_i |F_{\text{PH}}|$, where F_{PH} and F_{P} are the derivative and native structure factors, respectively. \S Phasing power = $\langle F_{\text{H}} \rangle / E$ where $\langle F_{\text{H}} \rangle$ is the root-mean-square heavy-atom structure factor and E is the residual lack of closure error. $\parallel R = \sum_i |F_{\text{O}}| - |F_{\text{C}}| / \sum_i |F_{\text{O}}|$, where R_{crist} and R_{free} are calculated with the working and test reflection sets, respectively. The test set reflections were held aside throughout refinement.

mals (4) and paclitaxel (TaxolTM) synthesis in the Pacific yew (5). These complex intramolecular cyclizations are unified by the biogenetic isoprene rule (1) and require precise conformational and stereochemical control (2, 6).

The biogenetic isoprene rule postulates that most sesquiterpene cyclization reactions occur through variations of a common mechanism involving (i) ionization of farnesyl diphosphate and electrophilic attack of the resultant allylic cation on one of the

remaining π bonds of the substrate, (ii) subsequent cationic transformations (additional electrophilic cyclizations, hydride transfers, and Wagner-Meerwein rearrangements), culminated by (iii) quenching of the positive charge by deprotonation or capture of an exogenous nucleophile such as water. In the absence of a crystal structure it is not clear how a terpenoid cyclase mediates such complex reactions. For instance, how does the enzyme select and enforce the correct precatalytic substrate conformation? The starting conformation of farnesyl diphosphate is known to be a critical determinant of the ultimate cyclization product (2). In addition, how does the enzyme trigger substrate ionization and stabilize positive charge in the resulting cascade of highly reactive, elec-

trophilic intermediates? Finally, how does the enzyme manage proper charge quenching to terminate the cyclization cascade?

We now report the x-ray crystal structure of recombinant pentalenene synthase (E.C. 4.6.1.5). This terpenoid cyclase has been isolated from *Streptomyces* UC5319, cloned, expressed in *Escherichia coli*, and crystallized (7). Pentalenene synthase catalyzes the cyclization of farnesyl diphosphate into pentalenene, a tricyclic sesquiterpene that is the hydrocarbon precursor of the pentalenolactone family of antibiotics. The three-dimensional structure of pentalenene synthase reveals active site features responsible for the conformational and stereochemical control of farnesyl diphosphate cyclization and clarifies the structural basis of the catalytic metal requirement.

C. A. Lesburg and D. W. Christianson, Roy and Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104-6323. G. Zhai and D. E. Cane, Department of Chemistry, Brown University, Providence, RI 02912.

*To whom correspondence should be addressed. E-mail: chris@xtal.chem.upenn.edu

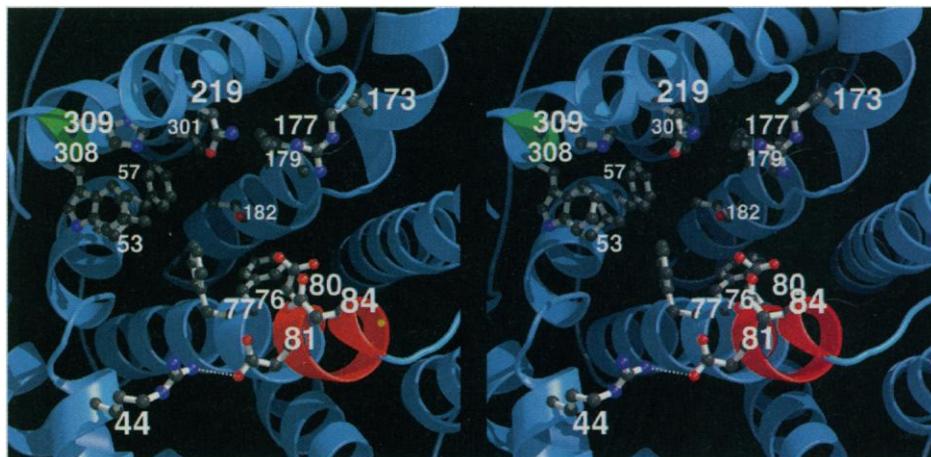
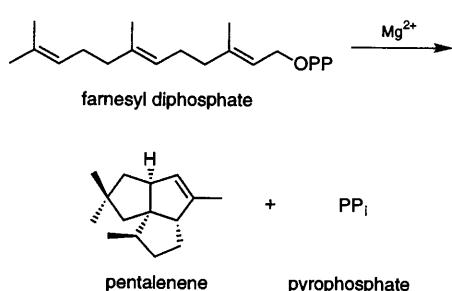


Fig. 3. Close-up view of active site residues in pentalenene synthase; selected residues are indicated by sequence number and discussed in the text. The aspartate-rich segments are red and the proposed catalytic base His³⁰⁹ is green.



The pentalenene synthase crystal structure was determined by multiple isomorphous replacement and refined at 2.6 Å resolution ($R = 0.198$, $R_{\text{free}} = 0.273$; Table 1). The enzyme has a globular structure (Fig. 1) of approximate dimensions 60 Å by

50 Å by 40 Å defined by an aggregation of 11 α helices (designated A to K). Helices B, C, G, H, and K surround an active site cavity approximately 15 Å deep and 9 Å wide. Polypeptide loops connecting these helices are relatively short (average three residues) on the side of the protein distal to the active site, and relatively long (average ten residues) on the side of the protein proximal (that is, surrounding) to the active site cavity. This feature is a consequence of the helix packing arrangement necessary to enclose a central, well-defined cavity. A disulfide linkage is found between Cys¹²⁸ and Cys¹³⁶.

A similar, but not identical, fold to that of pentalenene synthase is found in farnesyl diphosphate synthase (Fig. 2), which cata-

lyzes the synthesis of farnesyl diphosphate (the substrate of a sesquiterpene cyclase) from isopentenyl diphosphate and dimethylallyl diphosphate (8). Despite only 15% sequence identity between the two enzymes, the different substrates utilized by each enzyme, and the different outcomes of the electrophilic condensation reactions catalyzed, general structural features of their active sites are conserved. This structural homology is reminiscent of suggestions that enzymes catalyzing successive steps in a metabolic pathway may evolve with similar structures through divergence, regardless of the degree of amino acid sequence identity; function evolves more rapidly than sequence and sequence evolves more rapidly than tertiary structure (9). This expectation is certainly consistent with the unexpected structural homology observed between pentalenene synthase and the catalytic core of *epi*-aristolochene synthase despite only 16% sequence identity (10). In anticipation of broad structural relationships among this family of biosynthetic enzymes, we designate the α -helical topology observed for pentalenene synthase as a minimal "terpenoid synthase fold."

The active site cavity of pentalenene synthase (Fig. 3) is identified by structural homology with farnesyl diphosphate synthase (8) as well as the binding of a nonreactive substrate analogue (11). The bottom of this cavity is predominantly hydrophobic in nature and contains aromatic residues Phe⁵⁷, Phe⁷⁶, Phe⁷⁷, and Trp³⁰⁸, and aliphatic residues Leu⁵³, Val¹⁷⁷, Val¹⁷⁹, Thr¹⁸², and Val³⁰¹. These residues confer an overall shape to the cavity, which serves as a tem-

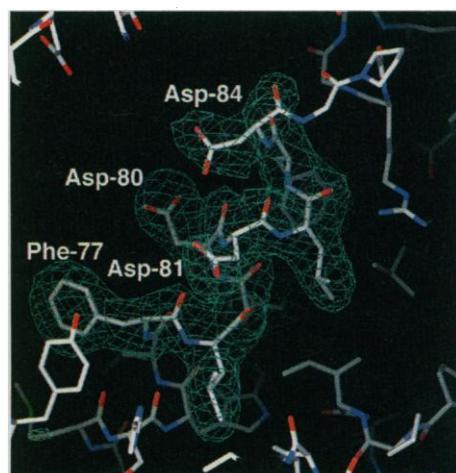


Fig. 4. Averaged omit map of the aspartate-rich segment calculated with Fourier coefficients $|F_o| - |F_c|$ and phases derived from the final model less the atoms of Phe⁷⁷-Asp⁸⁴ (contoured at 4.5σ). Given that the carboxylate side chains of Asp⁸⁰ and Asp⁸⁴ are oriented toward the active site cavity, these residues most likely bind Mg^{2+} in the enzyme-substrate complex.

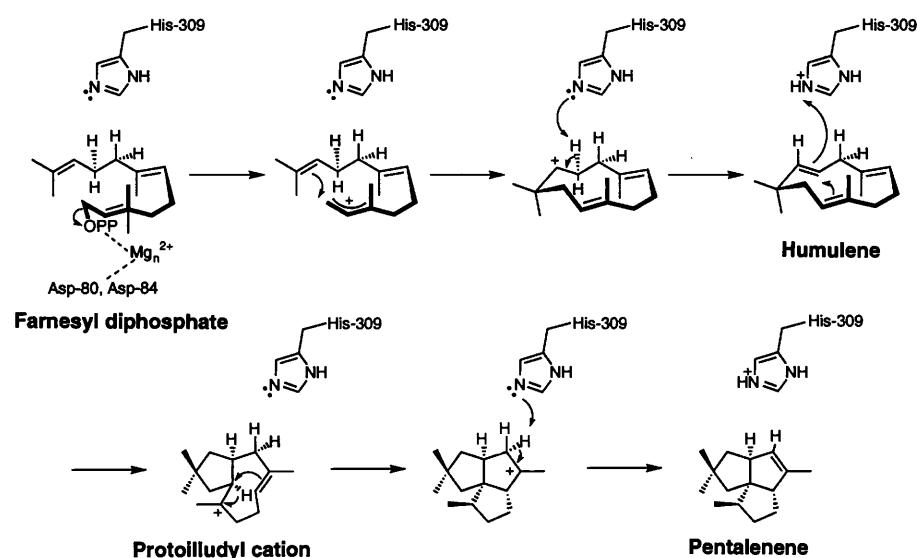


Fig. 5. Proposed mechanism for cyclization of farnesyl pyrophosphate to intermediates humulene and protoilludyl cation, with subsequent rearrangement into pentalenene (7, 14). The pyrophosphate leaving group is omitted for clarity; however, it may remain bound in the enzyme active site during the cyclization cascade and contribute to the electrostatic stabilization of carbocation intermediates.

plate for the binding of farnesyl diphosphate in the correct conformation for cyclization to pentalenene. The upper region of the active site cavity is somewhat more hydrophilic in nature and includes the polar or charged side chains of His³⁰⁹, Asn²¹⁹, Arg⁴⁴, Arg¹⁵⁷, Arg¹⁷³, Lys²²⁶, and Arg²³⁰. In addition, the carboxylate side chains of Asp⁸⁰ and Asp⁸⁴ (and to a lesser degree, Asp⁸¹, which salt links with Arg⁴⁴) protrude into the upper active site cavity (Fig. 4). These three aspartate residues comprise a signature "aspartate-rich segment" (12), which indicates a postulated metal-binding site (7); Mg²⁺ is required to facilitate pyrophosphate departure in the first step of the cyclization reaction (13). Basic residues Arg¹⁵⁷, Arg¹⁷³, Lys²²⁶, and Arg²³⁰ may also stabilize the pyrophosphate leaving group.

Initiation of a cyclization cascade is triggered by the ionization of a substrate molecule bound in the enzyme active site with precisely controlled folding, shape, and size (2, 4, 6). Mechanistic studies of pentalenene synthase by stereospecifically labeled substrates are consistent with the

mechanism and stereochemistry (Fig. 5) wherein farnesyl diphosphate undergoes ionization and electrophilic attack of the incipient allylic cation-pyrophosphate pair on the distal π bond (7, 14). The carboxylate side chains of Asp⁸⁰ and Asp⁸⁴ are oriented most favorably for Mg²⁺ complexation; the coordination of pyrophosphate to Mg²⁺, as well as hydrogen bond interactions with nearby basic residues, trigger leaving group departure. Deprotonation of the resultant humulyl cation is thought to result in the formation of the intermediate humulene. Stereospecific reprotonation of humulene yields the protoilludyl cation, which in turn undergoes rearrangement and cyclization to the product pentalenene. Inspection of the pentalenene synthase active site reveals that the likely catalytic base is His³⁰⁹. The N δ -H group of this residue donates a hydrogen bond to the backbone carbonyl of Ser³⁰⁵, leaving the lone electron pair at Ne oriented toward the active site cavity and poised for catalysis at the pH optimum of 8.2 to 8.4. Notably, a catalytically essential histidine residue has been

identified in the humulene synthase from sage leaf (15).

The precatalytic binding conformation of farnesyl diphosphate in the pentalenene synthase active site can be modeled unambiguously on the basis of the well-determined stereochemical details of the cyclization reaction (16). A properly folded substrate molecule will contact the catalytic base His³⁰⁹, and the diphosphate moiety will contact the aspartate-rich sequence of the enzyme through a bridging magnesium ion (or ions) (Fig. 6). Approximately 80% of the substrate surface area is buried in the enzyme-substrate complex, helping to prevent premature quenching of carbocation intermediates by solvent (16).

The U-shaped conformation of substrate farnesyl diphosphate is centered about an axis defined by Phe⁷⁷ and Asn²¹⁹ (Fig. 6). We propose that these two residues in particular, and the remaining active site residues in general, provide a template that channels reactive conformations along the exclusive reaction coordinate leading to pentalenene formation. As such, these residues destabilize incorrect substrate and intermediate conformations that would otherwise lead to spurious cyclization products; alternative substrate starting conformations do not fit well in the enzyme active site. We therefore propose that Phe⁷⁷ and Asn²¹⁹ are optimally located to stabilize highly reactive carbocation intermediates through favorable quadrupole-charge (17) and dipole-charge interactions, respectively:

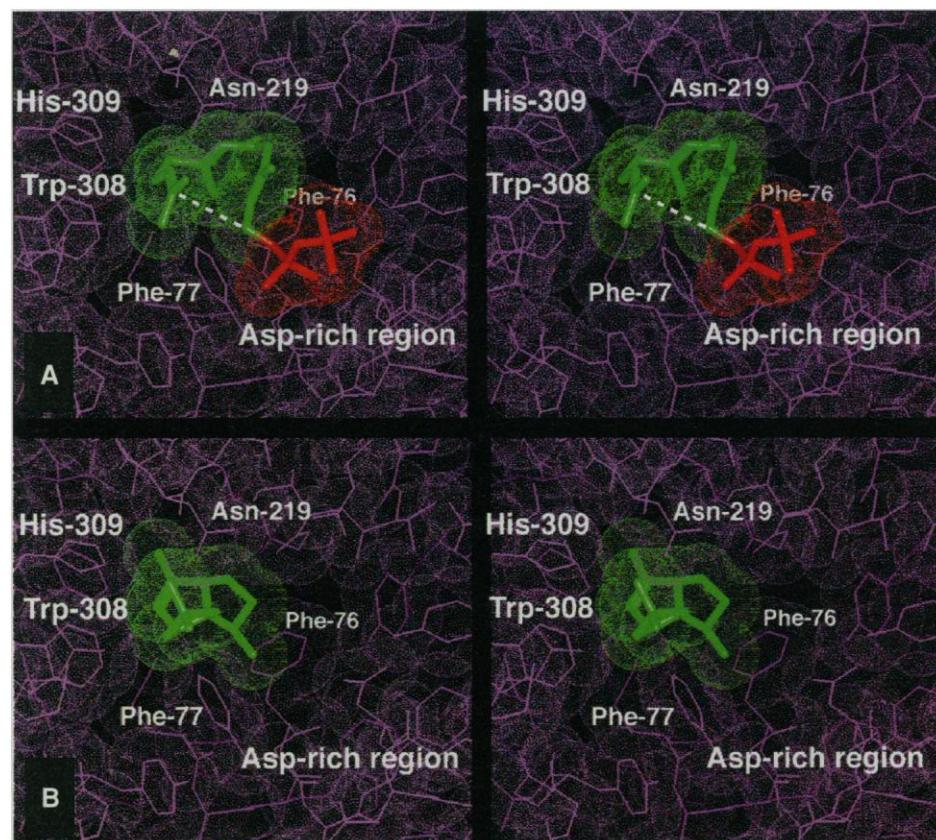
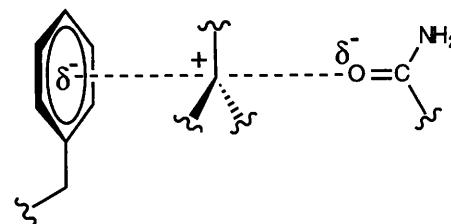


Fig. 6. (A) Farnesyl diphosphate must bind in the active site cleft with the correct conformation required for cyclization to pentalenene, with the C9-H oriented toward catalytic base His³⁰⁹ and the pyrophosphate moiety oriented toward the aspartate-rich segment through bridging Mg²⁺ (metal ions not shown for the sake of clarity). The trajectory of C-C bond formation leading to humulene synthesis in the initial cyclization step is indicated by a dashed line. (B) Pentalenene (as well as the preceding intermediates) is complementary in shape to the active site cavity. Comparison of substrate and product binding suggests that the active site is a template for correct conformation and stereochemistry in the cyclization cascade.



Further stabilization may result from a quadrupole-charge interaction with Phe⁷⁶, which is well positioned to stabilize the protoilludyl cation and the derived rearrangement product, which is generated by the 1,2-hydride shift. Similarly, Trp³⁰⁸ may also be well positioned to stabilize positive charge at sites corresponding to the original C10, C9, and C8 atoms of farnesyl diphosphate.

It is interesting to consider that at least two-thirds of the carbon atoms of the farnesyl diphosphate backbone undergo substantial changes in hybridization, configuration, and bonding during the cyclization cascade catalyzed by a typical sesquiterpene cyclase (2). The pentalenene synthase structure shows that the enzyme serves as a template to channel conforma-

tion and stereochemistry in the chemistry of cyclization, and the enzyme elegantly protects and stabilizes reactive carbocation intermediates. These features are likely to be common to all terpenoid cyclases. Alteration of active site residues in a terpenoid cyclase can result in the formation of aberrant cyclization products, so that the architecture and chemical nature of the substrate binding pocket are critical for channeling the correct cyclization chemistry (18). Now that the three-dimensional structure of a terpenoid synthase is available at atomic resolution, we are in a position to test the specific catalytic function of individual amino acids, as well as entire helical subdomains, in structure-based redesign experiments.

REFERENCES AND NOTES

1. L. Ruzicka, *Pure Appl. Chem.* **6**, 493 (1963).
2. D. E. Cane, *Acc. Chem. Res.* **18**, 220 (1985); R. Croteau and D. E. Cane, *Methods Enzymol.* **110**, 383 (1985); D. E. Cane, *Chem. Rev.* **90**, 1089 (1990).
3. Hippocrates, ΠΙΠΙ ΝΟΥΣΩΝ Β (P. Potter, translator) (Harvard Univ. Press, Cambridge, 1988); Pliny, *Naturalis Historiae* (W. H. S. Jones, translator) (Harvard Univ. Press, Cambridge, 1969); C. A. Michie and E. Cooper, *J. R. Soc. Med.* **84**, 602 (1991); P. Dolara *et al.*, *Nature* **379**, 29 (1996).
4. I. Abe, M. Rohmer, G. D. Prestwich, *Chem. Rev.* **93**, 2189 (1993); E. J. Corey *et al.*, *J. Am. Chem. Soc.* **117**, 11819 (1995); E. J. Corey and H. B. Wood, *ibid.* **118**, 11982 (1996); E. J. Corey *et al.*, *ibid.* **119**, 1277 (1997); E. J. Corey *et al.*, *ibid.*, p. 1289.
5. The committed step of paclitaxel (Taxol™) biosynthesis is the generation of taxadiene from geranylgeranyl diphosphate in a reaction catalyzed by taxadiene synthase; X. Y. Lin, M. Hezari, A. E. Knoepp, H. G. Floss, R. Croteau, *Biochemistry* **35**, 2968 (1996).
6. R. Croteau, *Chem. Rev.* **87**, 929 (1987).
7. D. E. Cane *et al.*, *Biochemistry* **33**, 5846 (1994); C. A. Lesburg, M. D. Lloyd, D. E. Cane, D. W. Christianson, *Protein Sci.* **4**, 2436 (1995).
8. L. C. Tarshis, M. Yan, C. D. Poulter, J. C. Sacchettini, *Biochemistry* **33**, 10871 (1994); L. C. Tarshis, P. J. Proteau, B. A. Kellogg, J. C. Sacchettini, C. D. Poulter, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 15018 (1996).
9. D. Reardon and G. K. Farber, *FASEB J.* **9**, 497 (1995).
10. C. M. Starks, K. Back, J. Chappell, J. P. Noel, *Science* **277**, 1815 (1997).
11. C. A. Lesburg, G. Zhai, D. E. Cane, D. W. Christianson, unpublished data.
12. M. N. Ashby and P. A. Edwards, *J. Biol. Chem.* **265**, 13157 (1990).
13. Aqueous Mg²⁺ is known to catalyze the solvolysis of allylic diphosphates even at pH 7.0; L. Chayet, M. C. Rojas, O. Cori, C. A. Buntun, D. C. McKenzie, *Bioorg. Chem.* **12**, 329 (1984).
14. D. E. Cane, C. Abell, A. M. Tillman, *ibid.*, p. 312; D. E. Cane *et al.*, *J. Am. Chem. Soc.* **112**, 4513 (1990); D. E. Cane *et al.*, *Philos. Trans. R. Soc. B.* **332**, 123 (1991); D. E. Cane and S. W. Weiner, *Can. J. Chem.* **72**, 118 (1994).
15. S. S. Dehal and R. Croteau, *Arch. Biochem. Biophys.* **261**, 346 (1988); J. I. M. Rajaonarivony, J. Gershenzon, J. Miyazaki, R. Croteau, *ibid.* **299**, 77 (1992).
16. Models of enzyme-substrate, -intermediate, and -product complexes were prepared by calculating minimum energy conformations along the reaction coordinate of cyclization with MacroModel [F. Mohamadi *et al.*, *J. Comput. Chem.* **11**, 440 (1990)]. Subsequently, these conformers were docked into the pentalenene synthase active site as guided by the strict stereochemical requirements for each step

- of the cyclization reaction. Buried surface areas were calculated with the use of GRASP [A. Nicholls, K. A. Sharp, B. Honig, *Prot. Struct. Funct. Gen.* **11**, 281 (1991)].
17. S. K. Burley and G. A. Petsko, *Adv. Prot. Chem.* **39**, 125 (1988); D. A. Dougherty, *Science* **271**, 163 (1996).
18. D. E. Cane and Q. Xue, *J. Am. Chem. Soc.* **118**, 1563 (1996); K. Back and J. Chappell, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6841 (1996).
19. J. Nyborg and A. J. Wonacott, in *The Rotation Method in Crystallography*, U. W. Arndt and A. J. Wonacott, Eds. (North-Holland, Amsterdam, 1977), p. 139.
20. Collaborative Computational Project, Number 4, *Acta Crystallogr.* **D50**, 760 (1994).
21. Z. Otwinowski and W. Minor, *Methods Enzymol.* **276**, 307 (1997).
22. J. P. Abrahams and A. G. W. Leslie, *Acta Crystallogr.* **D52**, 30 (1996).
23. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987).
24. T. A. Jones, J.-Y. Zou, S. W. Cowan, M. Kjeldgaard,

Acta Crystallogr. **A47**, 110 (1991).

25. Figures 1 and 2 were prepared with (i) MOLSCRIPT: P. Kraulis, *J. Appl. Crystallogr.* **24**, 946 (1991); and (ii) Raster3D: D. J. Bacon and W. F. Anderson, *J. Mol. Graph.* **6**, 219 (1988); E. A. Merritt and M. E. P. Murphy, *Acta Crystallogr.* **D50**, 869 (1994).
26. We thank G. Farber, A. Jain, Z. Kanyo, W. N. Lipscomb, M. D. Lloyd, P. Lu, P. Sprengler, T. Stams, and G. M. Whitesides for helpful discussions during the course of this investigation and J. Noel for sharing data on the structure of *epi*-aristolochene synthase prior to publication. Supported by grants from the National Institutes of Health. This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation under award DMR-9311772, with the use of the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by award RR-01646 from the National Institutes of Health. Atomic coordinates have been deposited in the Brookhaven Protein Data Bank with accession code 1PS1.

21 March 1997; accepted 1 August 1997

Toroidal Structure of λ-Exonuclease

Rhett Kovall and Brian W. Matthews

Structure determination at 2.4 angstrom resolution shows that λ-exonuclease consists of three subunits that form a toroid. The central channel is funnel shaped, tapering from an inner diameter of about 30 angstroms at the wider end to 15 angstroms at the narrow end. This is adequate to accommodate the DNA substrate and thus provides a structural basis for the ability of the enzyme to sequentially hydrolyze thousands of nucleotides in a highly processive manner. The results also suggest the locations of the active sites and the constraints that limit cleavage to a single strand.

DNA exonucleases participate in DNA replication, repair, and recombination (1). In particular, processive 5' to 3' single-stranded DNA exonucleases are essential for generating early DNA intermediates in many pathways of prokaryotic and eukaryotic homologous recombination (2). Bacteriophage λ encodes its own exonuclease, λ-exonuclease, which facilitates phage DNA recombination through the double-strand break repair (DSBR) and single-strand annealing pathways (3, 4).

λ-Exonuclease binds a free end of double-stranded DNA and degrades one of those strands in the 5' to 3' direction, releasing 5' mononucleotides at a rate of ~12 nucleotides per second (5). The protein requires Mg²⁺ and the 5'-terminal phosphate for activity (5). λ-Exonuclease is highly processive; it remains bound to DNA while it sequentially cleaves ~3000 nucleotides (6).

In the DSBR pathway of recombination, the long 3' single-stranded overhangs created by λ-exonuclease are bound by the *Escherichia coli* recombination protein RecA. These single-stranded DNA RecA

Institute of Molecular Biology, Howard Hughes Medical Institute, and Department of Physics, University of Oregon, Eugene, OR 97403, USA.

filaments then undergo strand exchange with a homologous piece of double-stranded DNA (3, 7). In the case of single-strand annealing recombination, the creation of 3' overhangs exposes two homologous single-stranded regions of DNA that anneal to form the recombinant double-stranded molecule independently of RecA (4, 7). Analogous pathways of recombination in yeast (8), *Xenopus* oocytes (9), and mammalian cells (10) also require a processive 5' to 3' single-stranded DNA exonuclease.

Here we describe the crystal structure of λ-exonuclease determined at 2.4 Å resolution with the use of multiple isomorphous replacement and anomalous dispersion (Table 1). The asymmetric unit of the crystal structure shows that λ-exonuclease, known to be a multimer (11), is indeed a trimer with noncrystallographic threefold symmetry (Fig. 1A). The three protein subunits form a toroid, with a tapered channel passing through the middle. At the wide end of the channel, the diameter is about 30 Å, and it decreases to about 15 Å at the narrow end. Although the crystal structure determined does not include DNA, it appears that the tapered channel is large enough to accommodate double-stranded DNA at the wide end but