

- March 1986 [M. De Beer, *Ann. Mus. R. Afr. Cent. Sci. Zool.* **257**, 57 (1989)] to 1.2 and 1.6 m in February to March 1995 and February to March 1996 (O. Seehausen, J. J. M. van Alphen, F. Witte, unpublished data).
12. In an experiment that was started in 1994, we crossed two females of *Platytaenioides degeni* with one male *H. nyererei*. A phylogenetic analysis confirmed that they belong to two different genera (E. Lippitsch, O. Seehausen, N. Bouton, unpublished data). The F_1 's were fully viable and did not suffer decreased fertility: nine females of the two F_1 's produced clutch sizes of 2.64 ± 1.77 ($n = 5$) and 4.84 ± 2.57 ($n = 4$) fry per gram of body weight, versus 1.14 ± 1.29 ($n = 10$) in the *P. degeni* F_1 and 3.47 ± 2.43 ($n = 7$) in the *H. nyererei* F_1 . The hybrid F_2 was also fully viable and fertile. Two females produced clutch sizes of 3.44 ± 3.34 fry per gram of body weight, and the first F_3 nest is now about to reach maturity. Sex ratios were 2:1 female biased in F_1 and F_2 hybrids as well as in both parental species. M. D. Cragon de Caprona and B. Fritsch [*Neth. J. Zool.* **34**, 503 (1984)] found no difference in fertility between *Astatotilapia nubila* \times *Haplochromis* "black lividus" hybrids and the parent species. Three further laboratory hybridizations that we did and three done elsewhere [M. D. Cragon de Caprona, *Ann. Mus. R. Afr. Cent. Sci. Zool.* **251**, 117 (1996)] yielded viable and, where tested (two intergeneric crosses in our lab), fertile offspring. Successful hybridization between genera of haplochromine cichlids, resulting in viable offspring, has also been reported from Lake Malawi [J. R. Stauffer, N. J. Bowers, T. D. Kocher, K. R. McKaye, *Copeia* (1996), p. 203].
 13. All haplochromines, the retina of which has been studied, have one absorbance peak on blue and one on yellow to red [H. J. van der Meer and J. K. Bowmaker, *Brain Behav. Evol.* **45**, 232 (1995)].
 14. We have recorded seven sympatric blue/yellow-red sibling species pairs in three genera of rock-restricted haplochromines (7) and many more in eight genera that inhabit other habitats [for example, R. J. C. Hoogerhoud, F. Witte, C. D. N. Barel, *Neth. J. Zool.* **33**, 283 (1983)]. We have also found blue and yellow-red sympatric male color morphs in 11 rock-restricted species of five genera (7).
 15. R. A. Fisher, *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930); A. Zahavi, *J. Theor. Biol.* **53**, 205 (1975); R. Lande, *Proc. Nat. Acad. Sci. U.S.A.* **78**, 3721 (1981); W. D. Hamilton and M. Zuk, *Science* **218**, 384 (1982); M. Kirkpatrick, *Am. Nat.* **36**, 1 (1982); A. L. Basolo, *Science* **250**, 808 (1990); A. Grafen, *J. Theor. Biol.* **144**, 517 (1990); T. Goldschmidt, *Ethology* **88**, 177 (1991); M. Kirkpatrick and M. J. Ryan, *Nature* **350**, 33 (1991); M. J. Ryan, *Oxf. Surv. Evol. Biol.* **7**, 157 (1991); J. A. Endler, *Am. Nat.* **139**, s125 (1992).
 16. S. N. Archer, J. A. Endler, J. N. Lythgoe, J. C. Patriidge, *Vision Res.* **28**, 1243 (1987); A. L. Basolo and J. A. Endler, *Trends Ecol. Evol.* **10**, 489 (1995); G. F. Turner and M. T. Burrows, *Proc. R. Soc. London. Ser. B.* **260**, 287 (1995); interspecific differences in photopic sensitivity to red and blue light have been demonstrated for haplochromine cichlids [S. A. Smit and G. CH. Anker, *Neth. J. Zool.* **47**, 9 (1997)].
 17. J. A. Endler and A.E. Houde, *Evolution* **49**, 456 (1995).
 18. O. Seehausen and J. J. M. van Alphen, *Behav. Ecol. Sociobiol.*, in press.
 19. In one color morph, both sexes predominantly courted and responded to mates of their own morph (paired two-tailed *t* test: $P = 0.00005$ for females, $n = 12$, $t = -5.03$; $P = 0.003$ for males, $n = 5$, $t = -4.31$). In the other morph, females were unselective ($P = 0.24$, $n = 12$, $t = 1.21$) but males again courted predominantly females of their own morph ($P = 0.021$, $n = 4$, $t = 2.85$). In *N.* "blue scraper" both sexes are polymorphic. The morphs do not differ in ecology, nor in anatomy [O. Seehausen and N. Bouton, *Ecol. Freshw. Fish* **6**, 161 (1997)].
 20. N. G. Jerlov, *Marine Optics* (Elsevier, Amsterdam, 1976).
 21. E. Trewavas, *Ann. Mag. Nat. Hist.* **11**, 435 (1938).
 22. J. A. Endler, *Evol. Biol.* **11**, 319 (1978).
 23. ———, *Evolution* **34**, 76 (1980).
 24. A. Zahavi, *J. Theor. Biol.* **67**, 603 (1977); A. J. Pomiankowski, *ibid.* **128**, 195 (1987); A. P. Møller, *Nature* **339**, 132 (1989); P. H. Luyten and N. R. Liley, *Behaviour* **95**, 164 (1985); *Behav. Ecol. Sociobiol.* **28**, 329 (1991).
 25. We compared the male coloration of the *H. nyererei* population at Nyegezi rocks (Mwanza Gulf) in 1986 with that of the same population in 1993 through 1995. The red/blue ratio [610 nm/515 nm (31)] was significantly higher in the earlier year (1.18, SE = 0.14, $n = 6$ versus 0.75, SE = 0.17, $n = 11$; $Z = -1.96$, $P = 0.05$). Two males photographed in 1978 had an even higher red/blue ratio than those from 1986. Between 1986 and 1995 the water transparency at Nyegezi rocks had dropped from 3.2 to 1.2 m (17).
 26. Minimum adequate regression model (and sign of correlation) for species number: area size (+) $r^2 = 0.62$, $F = 19.48$, spectral width (+) $r^2 = 0.19$, $F = 6.72$; model $r^2 = 0.81$, $P = 0.007$; model for color morph number: water transparency (+) $r^2 = 0.69$, $F = 28.68$, slope (-) $r^2 = 0.13$, $F = 6.45$; model $r^2 = 0.82$, $P = 0.002$.
 27. The Shannon-Wiener index of species diversity is calculated from the number of species and the equitability of allotment of individuals among the species [C. J. Krebs, *Ecology, the Experimental Analysis of Distribution and Abundance* (Harper & Row, New York, 1972)]. The absence of a correlation with light conditions is explained by an inverse relation between species number and abundance equitability, indicating that resources are more equally allotted among species where species number is lower. This is supported by results of a study on resource utilization, which showed that niche width and interspecific niche overlap are larger where species number is lower (N. Bouton, O. Seehausen, J. J. M. van Alphen, *Ecol. Freshw. Fish*, in press).
 28. O. Seehausen, *Ecol. Freshw. Fish* **6**, 59 (1997).
 29. T. E. Reimchen, *Evolution* **43**, 450 (1989); J. A. Endler, *Vision Res.* **31**, 587 (1991).
 30. T. Goldschmidt, F. Witte, J. Wanink, *Conserv. Biol.* **7**, 686 (1993).
 31. Fish color and light transmission were measured with a microspectrometer (Ocean Optics PS 1000) and Ocean Optics Acquisition software on a portable computer. Color was measured on the flanks of live fishes photographed (Kodak Elite 100 ISO) under daylight immediately after capture (reference: gray card). Light penetration was measured as the bandwidth of the transmission spectrum at 2 m water depth (reference: surface light). Data points for color are means from 10 spectral scans; those for light are means from 2×10 spectral scans, taken with an interval of 1 year.
 32. M. J. Crawley, *GLIM for Ecologists* (Blackwell, London, 1993).
 33. E. Lippitsch, O. Seehausen, N. Bouton, unpublished data.
 34. We thank the Tanzanian Government and P. O. J. Bwathondi for research permits; the Mwanza center of TAFIRI (E. F. B. Katunzi) for hospitality and support; M. Kayeba, R. Enoka, A. Samwel Terry, and M. Haluna for assistance; L. Assembe for technical realization; L. Schadhauer, A. Samwel Terry, H. Nielsen, R. Kiteri, H. G. Mbinyi, C. C. Schaefer, G. De Leeuw, P. Swelderswaard, R. Stawikowski, and R. R. Mapunda for logistic support; T. Reimchen for discussion; K. Barel, N. Bouton, P. Brakefield, S. Collins, G. Fryer, F. Galis, J. Deutsch, T. Goldschmidt, I. Hardy, R. Lande, P. Mayhew, H. Proehl, U. Schliewen, J. Sevenster, H. Slabbekoorn, J. de Visser, and an anonymous reviewer for comments on the manuscript; and M. Brittin for the figures. We particularly acknowledge J. Yoshimura for stimulating discussions. Supported by grants from the Netherlands Foundation for the Advancement of Tropical Research, German Cichlid Association, TETRA, EHEIM, and with diving equipment from SCUBAPRO and SEAWAY.

10 June 1997; accepted 6 August 1997

Structure and Function of a Squalene Cyclase

K. Ulrich Wendt, Karl Poralla, Georg E. Schulz*

The crystal structure of squalene-hopene cyclase from *Alicyclobacillus acidocaldarius* was determined at 2.9 angstrom resolution. The mechanism and sequence of this cyclase are closely related to those of 2,3-oxidosqualene cyclases that catalyze the cyclization step in cholesterol biosynthesis. The structure reveals a membrane protein with membrane-binding characteristics similar to those of prostaglandin- H_2 synthase, the only other reported protein of this type. The active site of the enzyme is located in a large central cavity that is of suitable size to bind squalene in its required conformation and that is lined by aromatic residues. The structure supports a mechanism in which the acid starting the reaction by protonating a carbon-carbon double bond is an aspartate that is coupled to a histidine. Numerous surface α helices are connected by characteristic QW-motifs (Q is glutamine and W is tryptophan) that tighten the protein structure, possibly for absorbing the reaction energy without structural damage.

The cyclization reactions catalyzed by squalene cyclases (S-cyclases) and 2,3-oxidosqualene cyclases (OS-cyclases) are highly complex (1) and give rise to various

fused-ring compounds (2, 3). Although these enzymes have been well studied earlier (2, 4), present-day recombinant techniques have contributed much to the understanding of their function and reactivity (1, 5). Early suggestions for the reaction mechanism favored a concerted process (2), whereas current hypotheses (6, 7) dissect a series of carbocationic intermediates (Fig. 1). Recently, the OS-cyclases became targets for the development of antifungal and anticholesteremic drugs (8). The integral

K. U. Wendt and G. E. Schulz, Institut für Organische Chemie und Biochemie, Albertstrasse 21, D-79104 Freiburg im Breisgau, Germany.

K. Poralla, Biologisches Institut, Mikrobiologie/Biotechnologie, Auf der Morgenstelle 28, D-72076 Tübingen, Germany.

*To whom correspondence should be addressed. E-mail: schulz@bio5.chemie.uni-freiburg.de

membrane protein character of these enzymes is generally accepted (9, 10). As shown below, the S- and OS-cyclases have related sequences and should therefore have similar spatial structures. We now present the crystal structure of a bacterial S-cyclase at 2.9 Å resolution.

Recombinant squalene-hopene cyclase from *Alicyclobacillus acidocaldarius* has been previously crystallized in three distinct forms (11). For phase determination by multiple isomorphous replacement (MIR), we used the crystal form A', which contains three polypeptide chains in the asymmetric unit that show only small displacements from the threefold symmetry of crystal form A (Tables 1 and 2). A molecular model was derived from the A' crystal MIR electron

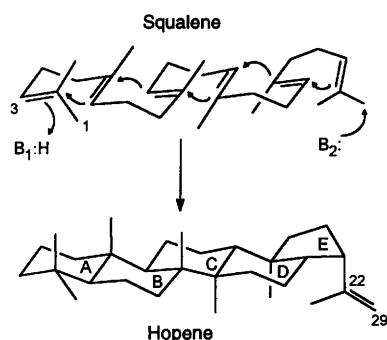


Fig. 1. The proposed reaction steps in squalene-hopene cyclases involving carbocationic intermediates. The side chains of Q and W are stacked, forming hydrogen bonds with the amino end of the adjacent outer barrel helix and with the carbonyl end of the preceding outer barrel helix, respectively (Fig. 3). The QW-motifs connect all outer helices of the

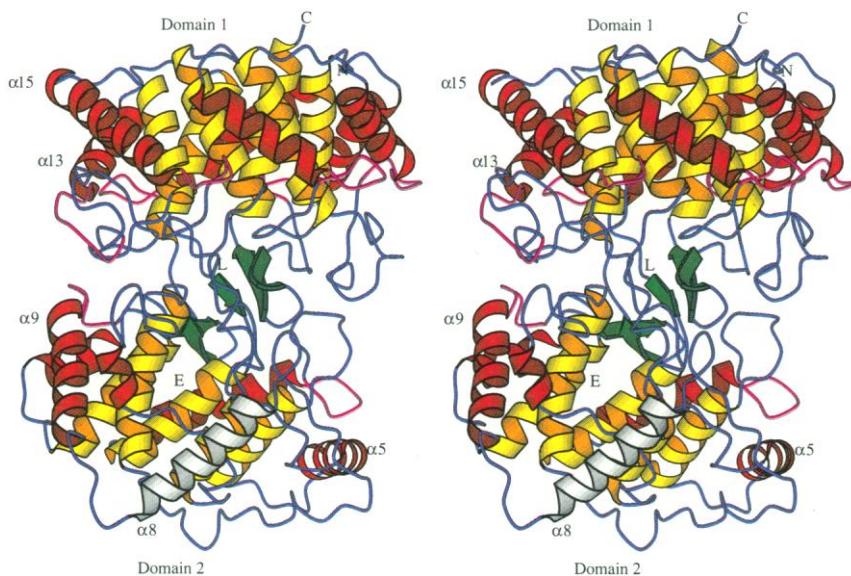


Fig. 2. Stereoview of squalene-hopene cyclase chain fold with labeled NH₂- and COOH-termini (N and C), inhibitor position (L), and channel entrance (E). Color code: internal (yellow) and external (red) barrel helices, β structure (green), QW-motifs (purple), and $\alpha 8$ in the nonpolar plateau (white).

density. This model was then refined in crystal form A, which shows exact threefold symmetry with only one chain in the asymmetric unit (Table 2). In agreement with our gel permeation chromatography results and those of others (12), the crystal structure reveals the presence of a dimeric enzyme. Both subunits are related by a twofold axis and form a polar interface of 850 Å². Each subunit consists of 631 amino acid residues and has a molecular mass of 71,569 daltons; residues 1 to 9 and 629 to 631 could not be located in electron density maps. The chain is organized in two domains (Fig. 2), forming a dumbbell shape. Domain 1 is an α_6 - α_6 barrel of two concentric rings of parallel α helices (Fig. 3) that resembles the chain fold of glucanases and a farnesyl transferase (13). Domain 2 is inserted into domain 1 and contains an α - α barrel, which appears to be an evolved version of the α_6 - α_6 barrel. Both barrels point with the amino ends of their inner α helices toward the molecular center, which consists of long loops from both domains forming a small β structure and enclosing a large cavity of about 1200 Å³ (14).

The aligned sequences of S-cyclase and human OS-cyclase are presented in Fig. 4. The S-cyclase contains eight QW-sequence motifs (15), five of which are also present in the OS-cyclases. Seven of the eight motifs assume virtually identical polypeptide conformations. The side chains of Q and W are stacked, forming hydrogen bonds with the amino end of the adjacent outer barrel helix and with the carbonyl end of the preceding outer barrel helix, respectively (Fig. 3). The QW-motifs connect all outer helices of the

α_6 - α_6 barrel and several outer helices of the other α - α barrel, thus stabilizing the whole protein (Figs. 2 and 3).

The active site is located in the large central cavity (Fig. 5A), as indicated by the bound competitive inhibitor (16) *N,N*-dimethyldodecylamine-*N*-oxide (LDAO) (inhibition constant $K_i = 0.14 \mu\text{M}$). This location is corroborated by a putative mechanism-based inhibitor that labeled the equivalent of Asp³⁷⁶ at the top of the cavity (Fig. 5A) (17) and similar experiments were done that labeled segments containing residues lining the cavity (18). Moreover, mutation of Asp³⁷⁶ or Asp³⁷⁷ inactivated S-cyclase (19), and mutation of the equivalent of Asp³⁷⁶ (18) inactivated OS-cyclase. Finally, the active site location at the amino ends of the inner helices of the α_6 - α_6 barrel corresponds to those of the four enzymes with similar barrel structures (13).

The central cavity can be accessed through a nonpolar channel between the helices of domain 2 (Fig. 5A). This domain

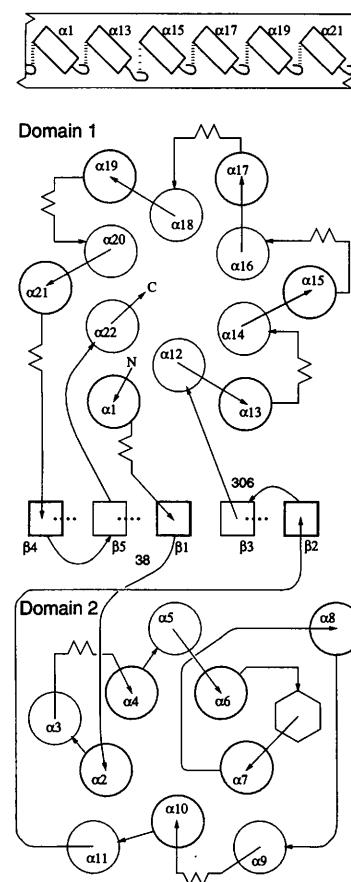


Fig. 3. Chain fold topology showing α helices (circles), one 3_{10} -helix (hexagon), β strands (squares), and QW-motifs (zigzag) with marked chain directions upward (thin) and downward (thick). Domain border positions are given. The top sketch shows how the outer ring of the α_6 - α_6 barrel is stabilized by the QW-motifs, one of which deviates from the others.

consists of an α - α barrel including a 3_{10} -helix (Fig. 3) that leaves space for the crossing channel. Helix $\alpha 8$ with the following loop, the segment between $\alpha 6$ and $\alpha 7$ as well as loop $\alpha 15$ - $\alpha 16$ (Figs. 2 and 3), form a large and rather mobile nonpolar plateau on the protein surface that surrounds the channel entrance (Fig. 5B).

This plateau has a solvent-accessible surface of about 1600 \AA^2 (14) and is encircled by a ring of positively charged residues (Fig. 5B), which suggests that the plateau plunges into the nonpolar center of a membrane, whereas the ring forms salt bridges with the displaced phospholipid and sulfolipid head groups. Such a model characterizes a monotopic membrane protein (20) and explains how the substrate squalene diffuses from the membrane interior, where it is dissolved (21) into the central cavity.

The dimer interface contains, in large part, residues from domain 2 (Fig. 6A). It orients the nonpolar plateaus in parallel (Fig. 6B). Such nonpolar plateau twins also occur with the other structurally known monotopic membrane protein (Fig. 6C), prostaglandin- H_2 -synthase (22). This enzyme has a completely different chain fold but also takes up its substrate from the membrane. The resemblance of these arrangements indicates a functional role as, for instance, a doubling of the membrane interaction per particle.

In S-cyclase the channel between the nonpolar plateau and active site contains a constriction formed by Phe¹⁶⁶, Val¹⁷⁴, Phe⁴³⁴, and Cys⁴³⁵ (Fig. 7), which appears to block it (Fig. 5A). Residues 434 and 435, however, are on a mobile loop that contacts a surface loop around position 210 with even higher mobility (23). We therefore suggest that the constriction behaves as a gate that permits substrate passage.

Fig. 4. Sequence alignment of squalene-hopene cyclase from *A. acidocaldarius* (top row) and human liver 2,3-oxidosqualene-cyclase (bottom row) from program CLUSTAL. Secondary structures (29) and QW-motifs are marked in the color code of Fig. 2. Asterisks indicate identical residues. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

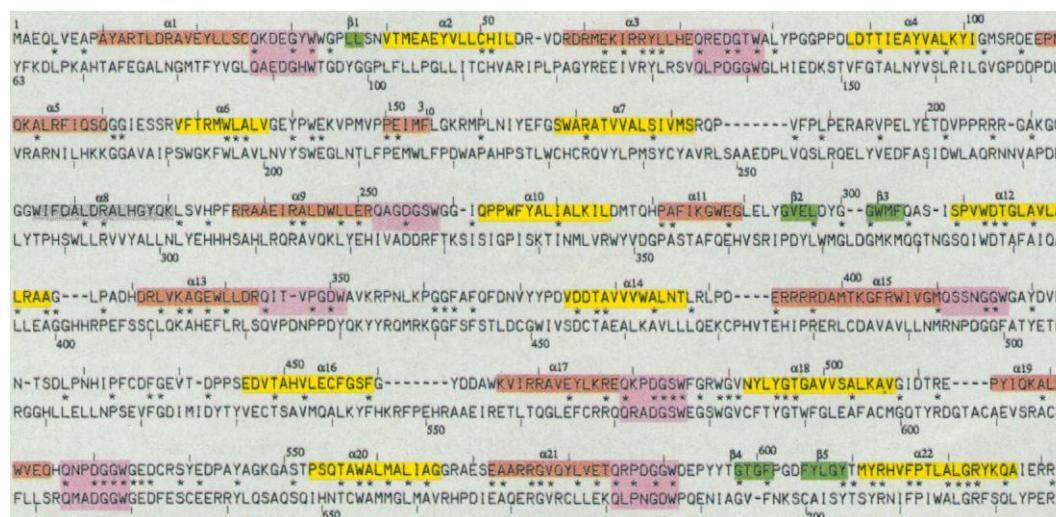
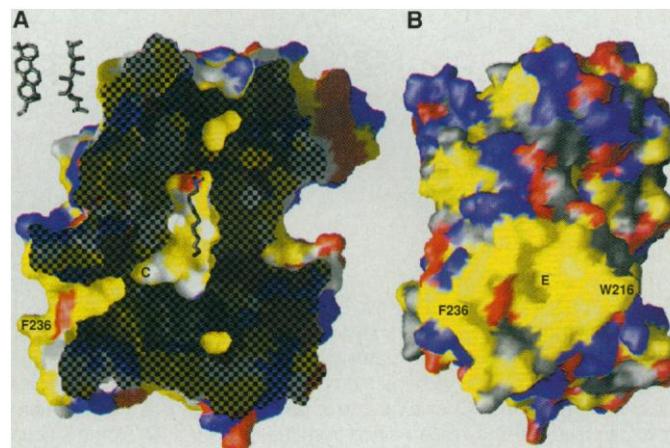


Table 1. MIR phasing in crystal form A' ($P3_221$, $a = b = 140.9 \text{ \AA}$, $c = 243.8 \text{ \AA}$), which is closely related to form A ($P321$, c axis one-third) with one asymmetric chain (11). For the "native" data (NAT) and all soaks we used single and double mutants containing D376C (Asp at position 376 mutated to Cys). The soaks were as follows: M1 and M2, methylmercury acetate; T1, T2, and T3, tetrakis(acetomercury)methane; Cl, chloromercury ferrocene; and KHg, $K_2Hg(CN)_4$. Phasing was initiated with the difference-Patterson of T1 and performed with MLPHARE (25), resulting in a figure of merit of 0.424 at 3.2 \AA resolution. T3 was from double-mutant C50S-D376C and lacked the major site of T1 and T2. Because of the close relation with form A, all sites are triplicate.

Item	NAT	M1	M2	T1	T2	T3	Cl	KHg
Soak (mM)	—	1	10	1	5	5	10	2
Soak time (days)	—	0.25	2.5	0.1	2	1	3	30
Resolution (\AA)	2.8	3.4	4.2	5.4	3.9	4.0	5.4	3.0
Completeness (%)	83	75	91	94	93	88	91	83
(last shell)	44	29	66	78	85	72	58	65
Redundancy	3.9	2.4	2.3	2.6	2.0	2.0	2.6	1.7
Average I/σ	5.9	10.0	15.4	17.2	10.0	8.3	17.9	9.3
R_{sym} (%)*	8.3	5.9	3.7	3.2	6.4	6.7	3.2	4.8
(last shell)	17.0	16.2	6.8	7.5	11.5	10.3	8.2	12.0
R_{iso} (%)†	—	17.6	25.7	18.1	20.1	18.2	16.1	16.1
Sites	—	12	12	9	9	3	9	9
R_{cullis} (%)‡	—	57	75	51	73	86	58	79
Phasing power	—	1.63	1.29	2.57	1.23	0.86	2.08	0.76

* $R_{sym} = \sum_j |I_j - \langle I \rangle| / \sum_j I_j$ with $j = (hkl, i)$, where I_j is the intensity for the reflection j , and $\langle I \rangle$ is the mean intensity for multiply measured reflections. † $R_{iso} = \sum_{hkl} |F_{PH} - F_P| / \sum_{hkl} F_P$, where F_{PH} and F_P are the derivative and native structure factor amplitudes, respectively. ‡ $R_{cullis} = \sum_j | |F_{PH} - F_P| - F_H | / \sum_j F_H$ with $j = (\text{centric } hkl)$, where F_H is the calculated heavy atom structure factor.

Fig. 5. The color-coded surface representations (30) with nonpolar (yellow), positive (blue), and negative (red) areas. (A) View similar to Fig. 2 but rotated around a vertical axis and sliced. The cutting plane (checked) opens the large internal cavity with the bound inhibitor LDAO. The nonpolar channel runs to the left, opening into a nonpolar plateau. The channel constriction (C) appears closed, but it is mobile enough to be readily opened. At the upper left, hopane (two views) is shown at scale. (B) View similar to Fig. 2 directly onto the 1600 \AA^2 nonpolar plateau with the channel entrance (E) at its center and two nonpolar side chains pointing to the outside. This is the only large nonpolar region on the surface.



The central 1200 Å³ active site cavity is mainly nonpolar, but it has a highly polar patch at the top (Figs. 5A and 7). It is lined by numerous aromatic residues that could stabilize the carbocationic intermediates of the cyclization reaction by their π-electrons (7). The residues lining the cavity are well conserved but show a gradient with highest conservation at the top and lowest at the bottom. This gradient indicates that the first reaction step common for S- and OS-cyclases occurs at the polar top of the cavity and that the variable features are at the bottom. Accordingly, the initially protonating acid B₁:H (Fig. 1) should be at the top; base B₂ of S-cyclase should accept a proton at the bottom end of the cavity; and B₂ of the

OS-cyclases should be near to the center of the substrate where proton uptake from the lanosteryl cation is expected (4). Our proposal appears to be consistent with the experiments of (18) where a histidine equivalent to Trp¹⁶⁹ near to the cavity center (Fig. 7) was labeled in later reaction steps, but it appears to disagree with the suggestion of (17) that Asp³⁷⁶ is near a position equivalent to ring E of hopene.

The polar top of the cavity has the sequence motif DXDD (X, any amino acid) of the S-cyclases (residues 374 to 377), which is important for catalysis (19). The OS-cyclases contain only the second aspartate residue, and this is crucial for catalysis (18). A similar catalytic motif, DDXXD, is present in farnesyl diphos-

phate synthase, which runs also through a cationic intermediate but appears to use this motif for metal binding (24). Because the chain folds are different, there is no evolutionary connection.

In S-cyclases Asp³⁷⁴ and Asp³⁷⁷ of the DXDD motif are hydrogen-bonded (2.6 Å distance) and close to the hydrogen-bonded (2.7 Å distance) pair Asp³⁷⁴:His⁴⁵¹ (Fig. 7). We propose that the pair Asp³⁷⁴:Asp³⁷⁷ carries a negative charge that stabilizes a positive charge on Asp³⁷⁶:His⁴⁵¹ and thus Asp³⁷⁶ in its protonated form. For the reaction, squalene diffuses into the central cavity where it assumes the required conformation (Fig. 1) with its C3 atom near to the putative proton on Asp³⁷⁶. We propose that Asp³⁷⁶ is the acid B₁:H which protonates the first C-C double bond, creating a carbocation that steps through the squalene chain and converts it to hopene (Fig. 1). The protonation leaves Asp³⁷⁶:His⁴⁵¹ as a stable salt bridge. The pair Asp³⁷⁴:Asp³⁷⁷ might support the initial protonation step. It is absent in the OS-cyclases, presumably because epoxides are more readily protonated than C-C double bonds.

Base B₂ of S-cyclase should be located at C29 of the hopenyl cation, which is

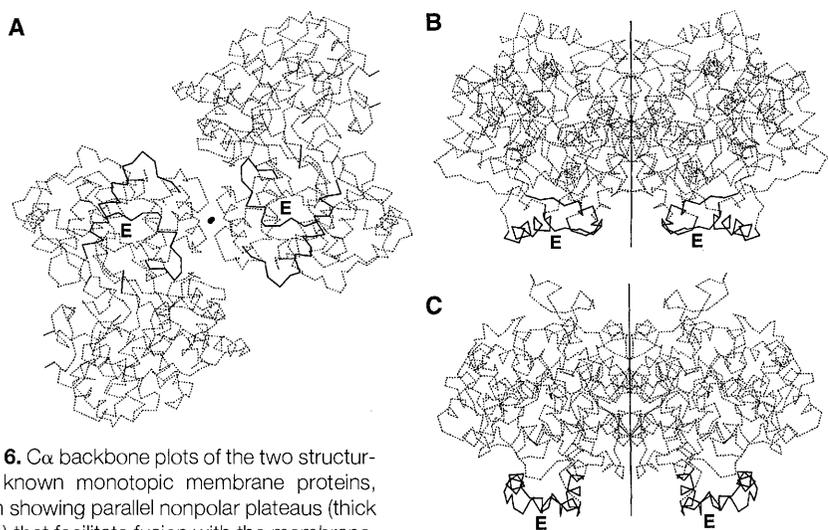


Fig. 6. α backbone plots of the two structurally known monotopic membrane proteins, both showing parallel nonpolar plateaus (thick lines) that facilitate fusion with the membrane. The channel entrances (E) for substrate uptake are labeled. (A) Homodimeric S-cyclase viewed along the molecular diad. (B) S-cyclase viewed from the top of (A). (C) Homodimeric prostaglandin-H₂-synthase (22) oriented like the S-cyclase in (B).

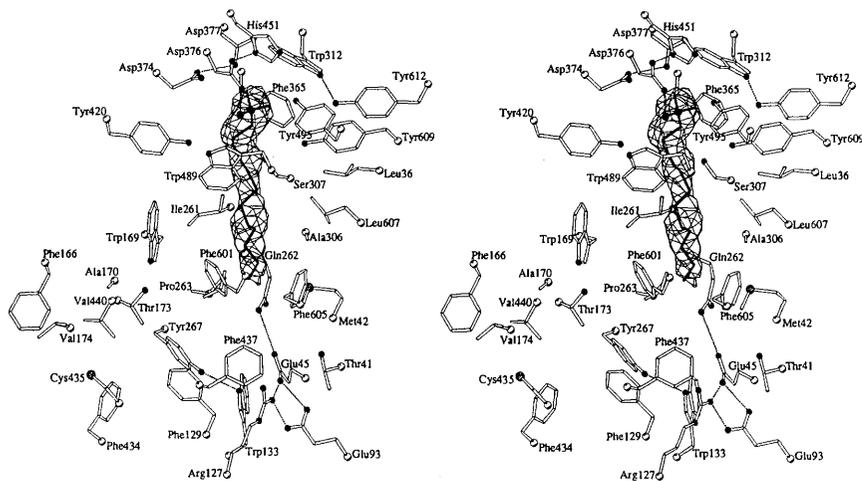


Fig. 7. Stereo view of the active site cavity with the inhibitor LDAO in its $(2F_o - F_c)$ electron density at a level of 1.3σ . The view is similar to Fig. 5A. Among the aromatic residues shown, Trp³¹², Phe³⁶⁵, Trp⁴⁸⁹, Tyr⁴⁹⁵, Phe⁶⁰¹, Tyr⁶⁰⁹, and Tyr⁶¹² are conserved in all known sequences of the family. The top end of the cavity is highly polar, and there is a polar net at the bottom.

Table 2. Structure refinement in crystal form A (*P*321, $a = b = 140.7$ Å, $c = 81.9$ Å) with one asymmetric chain (11). After MIR phasing (Table 1), solvent-flattening, and density averaging over the three asymmetric chains of A' (25), about 90% of the sequence could be placed (26). The model was then transferred to crystal form A, where genuine native data were available, and refined with X-PLOR (27) without any cutoff of the data. At the end, B factors were refined as grouped atoms (two groups per residue). The model had cysteines at all heavy atom sites. Residual density in the central cavity was modeled as LDAO (11). The model contains residues 10 to 628 and 38 water molecules, and 90% of the residues were in the most favored regions of the Ramachandran plot (28). Values in parentheses are for the last shell (2.95 to 2.85 Å), rmsd, root-mean-square deviation.

Item	Value
Resolution (Å)	20.0 to 2.9
Total reflections	53,288 (1579)
Unique reflections	18,914 (1017)
Completeness (%)	86 (49)
Redundancy	2.8 (1.6)
Average I/σ	11.9 (2.0)
R_{sym} (%)	5.6 (36.4)
All atoms	5,017
R factor (%)*	16.7
R_{free} of a 5% set (%)†	24.3
rmsd bond lengths (Å)	0.007
rmsd bond angles (degrees)	1.2

* $R_{\text{factor}} = \sum |F_o - F_c| / \sum F_o$, where F_o and F_c are the observed and calculated structure factors, respectively.
 † R_{free} is the cross-validation R factor computed for the test set of reflections (5% of the total), which are omitted in the refinement process.

near to the last carbon of LDAO (Fig. 7) where the protein offers no suitable residue. Because the reported S-cyclase shows a side reaction (10) resulting in about 10% diplopterol (hopan-22-ol), we suggest that B₂ is a water molecule polarized by other waters that are in contact with the hydrogen-bonding network of Gln²⁶²:Glu⁴⁵:Glu⁹³:Arg¹²⁷, which could store a proton (Fig. 7). Diplopterol is formed if the front water adds as hydroxyl to the last carbocation instead of accepting the proton.

Taken together, squalene enters the central cavity through the constriction, which acts as a gate, and is forced into the unfavorable conformation necessary for catalysis. The scale of the required actions corresponds to the low turnover number of 0.3 s⁻¹ (12, 16). Hopene formation releases ~200 kJ/mol at the bottom of the α₆-α₆ barrel, exceeding by far the usual protein stabilization energy of ~50 kJ/mol. The barrel does not disintegrate because its α helices at the surface are connected and thus stabilized by the QW-motifs, characteristic for this enzyme family. The resulting excitation then facilitates the return of hopene to the membrane.

REFERENCES AND NOTES

- E. J. Corey, S. P. T. Matsuda, B. Bartel, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2211 (1994).
- A. Eschenmoser, L. Ruzicka, O. Jeger, D. Arigoni, *Helv. Chim. Acta* **38**, 1890 (1955).
- G. Ourisson and M. Rohmer, *Acc. Chem. Res.* **25**, 403 (1992).
- R. B. Woodward and K. Bloch, *J. Am. Chem. Soc.* **75**, 2023 (1953); J. W. Cornforth *et al.*, *ibid.* **87**, 3224 (1965); I. Abe, M. Rohmer, G. D. Prestwich, *Chem. Rev.* **93**, 2189 (1993).
- E. J. Corey *et al.*, *Biochem. Biophys. Res. Commun.* **219**, 327 (1996); C. H. Baker, S. P. T. Matsuda, D. R. Liu, E. J. Corey, *ibid.* **213**, 154 (1995); I. Abe and G. D. Prestwich, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9274 (1995); Z. Shi, C. J. Buntel, J. H. Griffin, *ibid.* **91**, 7370 (1994); E. J. Corey, S. P. T. Matsuda, B. Bartel, *ibid.* **90**, 11628 (1993); C. A. Roessner *et al.*, *Gene* **127**, 149 (1993); T. Kaneko *et al.*, *DNA Res.* **3**, 109 (1996); I. G. Reipen, K. Poralla, H. Sahm, G. A. Sprenger, *Microbiology* **141**, 155 (1995); D. Ochs *et al.*, *J. Bacteriol.* **174**, 298 (1992).
- E. E. van Tamelen, *J. Am. Chem. Soc.* **104**, 6480 (1982); R. B. Boar, L. A. Couchman, A. J. Jaques, M. J. Perkins, *ibid.* **106**, 2476 (1984).
- W. S. Johnson *et al.*, *ibid.* **109**, 5852 (1987); C. J. Buntel and J. H. Griffin, *ibid.* **114**, 9711 (1992).
- Y. F. Zheng *et al.*, *J. Am. Chem. Soc.* **117**, 670 (1995).
- G. Balliano, F. Viola, M. Ceruti, L. Cattel, *Arch. Biochem. Biophys.* **293**, 122 (1992).
- B. Seckler and K. Poralla, *Biochim. Biophys. Acta.* **881**, 356 (1986).
- Crystal forms A, A', and A'' grew under identical conditions in hanging drops starting at a protein concentration of 8 mg/ml, 0.3% (w/v) C₈E₄ (n-octyl-tetraoxyethylene), 50 mM sodium citrate, pH 4.8, with a reservoir at 100 mM sodium citrate, pH 4.8 [K.-U. Wendt *et al.*, *Protein Sci.* **6**, 722 (1997)]. For the refinement we used a form A crystal that grew with 0.01% (w/v) LDAO as an additive in the starting drop. Diffraction data were collected with an area detector (Siemens X-1000) on a rotating anode (Rigaku RU200B) and processed with XDS [W. Kabsch, *J. Appl. Crystallogr.* **21**, 916 (1988)].
- S. Neumann and H. Simon, *Biol. Chem. Hoppe Seyler* **367**, 723 (1986).
- P. M. Alzari, H. Souchon, R. Dominguez, *Structure* **4**, 265 (1996); M. Juy *et al.*, *Nature* **357**, 89 (1992); A. Aleshin *et al.*, *J. Biol. Chem.* **267**, 19291 (1992); H.-W. Park *et al.*, *Science* **275**, 1800 (1997).
- Cavity volume (probe included) and solvent-accessible surface areas were calculated with a 1.4 Å radius probe [M. L. Connolly, *Science* **221**, 709 (1983)].
- K. Poralla *et al.*, *Trends Biochem. Sci.* **19**, 157 (1994).
- D. Ochs *et al.*, *Eur. J. Biochem.* **194**, 75 (1990).
- I. Abe and G. D. Prestwich, *J. Biol. Chem.* **269**, 802 (1994).
- E. J. Corey *et al.*, *J. Am. Chem. Soc.* **119**, 1277 (1997); E. J. Corey *et al.*, *ibid.*, p. 1289.
- C. Feil, R. Stüssmuth, G. Jung, K. Poralla, *Eur. J. Biochem.* **242**, 51 (1996).
- G. Blobel, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 1496 (1980).
- S. A. Simon *et al.*, *Biophys. J.* **19**, 83 (1977).
- D. Picot *et al.*, *Nature* **367**, 243 (1994).
- Mobilities are equated with crystallographic B factors.
- L. C. Tarshis, M. Yan, C. D. Poulter, J. C. Sacchettini, *Biochemistry* **33**, 10871 (1994).
- CCP4, Collaborative Computing Project Number 4, *Acta Crystallogr.* **D50**, 760 (1994).
- T. A. Jones *et al.*, *ibid.* **A47**, 110 (1991).
- A. T. Brünger *et al.*, *Science* **235**, 458 (1987).
- R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J. Appl. Crystallogr.* **26**, 283 (1993).
- D. Frishman and P. Argos, *Proteins Struct. Funct. Genet.* **23**, 566 (1995).
- A. Nicholls *et al.*, *ibid.* **11**, 281 (1991).
- We thank A. Lenhart for several measurements, C. Feil for the clone of mutant D376C, and R. Schwesinger for discussions. Supported by the Deutsche Forschungsgemeinschaft under contracts SFB-323 and SFB-388. The atomic model and the structure factors are deposited in the Protein Data Bank under the accession code 1SQC.

8 April 1997; accepted 13 August 1997

Structural Basis for Cyclic Terpene Biosynthesis by Tobacco 5-Epi-Aristolochene Synthase

Courtney M. Starks, Kyoungwhan Back, Joseph Chappell, Joseph P. Noel*

Terpene cyclases catalyze the synthesis of cyclic terpenes with 10-, 15-, and 20-carbon acyclic isoprenoid diphosphates as substrates. Plants have been a source of these natural products by providing a homologous set of terpene synthases. The crystal structures of 5-epi-aristolochene synthase, a sesquiterpene cyclase from tobacco, alone and complexed separately with two farnesyl diphosphate analogs were analyzed. These structures reveal an unexpected enzymatic mechanism for the synthesis of the bicyclic product, 5-epi-aristolochene, and provide a basis for understanding the stereochemical selectivity displayed by other cyclases in the biosynthesis of pharmacologically important cyclic terpenes. As such, these structures provide templates for the engineering of novel terpene cyclases.

Terpene cyclases control the synthesis of cyclic terpenoids including flavors and fragrances such as menthol and camphor, plant defense chemicals like capsidiol and lubimin (1), and more common compounds like steroids and lipid-soluble vitamins. Several cyclic terpenoids have pharmacological activity; for example, limonene can inhibit tumorigenesis induced in mice by particular carcinogens (2), and the diterpenoid taxol has antitumor activity (3). Numerous terpene cyclases from plant and microbial sources have been characterized (4, 5). Although the plant cyclases exhibit a significant degree of similarity in amino acid sequence, very

little similarity is observed between the bacterial, fungal, and plant terpene cyclases (6). These soluble enzymes convert the acyclic isoprenoid diphosphates geranyl diphosphate (GPP, 10 carbon), farnesyl diphosphate (FPP, 15 carbon), and geranylgeranyl diphosphate (GGPP, 20 carbon) into cyclic monoterpenes, sesquiterpenes, and diterpenes, respectively. In most cases, loss of diphosphate (pyrophosphate) from the enzyme-bound acyclic substrate results in an allylic carbocation that electrophilically attacks a double bond further down the terpene chain to effect the first ring closure. Additional rearrangements involving transient carbocations can include proton abstractions, hydride and alkyl migrations, and additional electrophilic attacks.

TEAS (tobacco 5-epi-aristolochene synthase) (7) from *Nicotiana tabacum* converts farnesyl diphosphate (FPP) to 5-epi-aristolochene (Fig. 1) (8), a precursor of the antifungal phytoalexin capsidiol. TEAS shares 77% amino acid identity with *Hyoscyamus muticus* vetispiradiene

C. M. Starks and J. P. Noel, Structural Biology Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA, and Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0301, USA.

K. Back and J. Chappell, Plant Physiology, Biochemistry, and Molecular Biology Program, University of Kentucky, Lexington, KY 40546-0091, USA.

*To whom correspondence should be addressed. E-mail: noel@sbl.salk.edu