Crystal Structure of Human Squalene Synthase
A KEY ENZYME IN CHOLESTEROL BIOSYNTHESIS *

Received for publication, May 15, 2000, and in revised form, July 12, 2000
Published, JBC Papers in Press, July 13, 2000, DOI 10.1074/jbc.M004132200

Squalene synthase catalyzes the biosynthesis of squalene, a key cholesterol precursor, through a reductive dimerization of two farnesyl diphosphate (FPP) molecules. The reaction is unique when compared with those of other FPP-utilizing enzymes and proceeds in two distinct steps, both of which involve the formation of carbocationic reaction intermediates. Because FPP is located at the final branch point in the isoprenoid biosynthesis pathway, its conversion to squalene through the action of squalene synthase represents the first committed step in the formation of cholesterol, making it an attractive target for therapeutic intervention. We have determined, for the first time, the crystal structures of recombinant human squalene synthase complexed with several different inhibitors. The structure shows that SQS is folded as a single domain, with a large channel in the middle of one face. The active sites of the two half-reactions catalyzed by the enzyme are located in the central channel, which is lined on both sides by conserved aspartate and arginine residues, which are known from mutagenesis experiments to be involved in FPP binding. One end of this channel is exposed to solvent, whereas the other end leads to a completely enclosed pocket surrounded by conserved hydrophobic residues. These observations, along with mutagenesis data identifying residues that affect substrate binding and activity, suggest that two molecules of FPP bind at one end of the channel, where the active center of the first half-reaction is located, and then the stable reaction intermediate moves into the deep pocket, where it is sequestered from solvent and the second half-reaction occurs. Five α helices surrounding the active center are structurally homologous to the active core in the three other isoprenoid biosynthetic enzymes whose crystal structures are known, even though there is no detectable sequence homology.

The isoprenoid biosynthetic pathway yields a structurally diverse family of low molecular weight molecules with a variety of physiological functions (1–3). In humans, the pathway produces such critical end-products as cholesterol, bile acids, dolichol, ubiquinone, steroid hormones, and prenylated proteins. Chain length-selective prenyltransferases catalyze the successive head-to-tail addition of the 5-carbon (C5-) isopentenyl diphosphate to the growing isoprene chain to form a series of linear C10-, C15-, C20- and C25-isoprenoid diphosphates. Cyclic terpenes are formed from these linear isoprenoid diphosphates through the actions of numerous class I terpene cyclases (3). The longer 30-carbon linear terpene, squalene, of cholesterol biosynthesis and the 40-carbon linear terpene, phytene, of carotenoid biosynthesis are formed by head-to-head condensations, respectively, of two 15-carbon and 20-carbon isoprenoid diphosphates, through the actions of the prenyltransferases squalene synthase and phytoene synthase (4) and are subsequently converted into their respective cyclic terpenes through the action of specific class II terpene cyclases (3). Both the isoprenoid chain elongation reactions, catalyzed by prenyltransferases, and the isoprenoid cyclization reactions, catalyzed by the terpene cyclases proceed via electrophilic alkyla-

ditions in which a new carbon-carbon single bond is generated through interaction between a highly reactive electron-deficient allylic carbocation and an electron-rich carbon-carbon double bond (2, 3).

Squalene synthase (EC 2.5.1.21) is a 47-kDa membrane-associated enzyme that catalyzes the reductive dimerization of two molecules of farnesyl diphosphate (FPP)1 in a two-step reaction to form squalene (Fig. 1). The uniqueness of the head-to-head coupling of two FPP molecules to form squalene via a stable cyclopropylcarbinyl diphosphate intermediate has elicited much mechanistic speculation over the years. The reaction proceeds in two distinct steps, both of which involve the formation of carbocationic reaction intermediates (4, 5). In the first half-reaction, two molecules of FPP react to form the stable cyclopropylcarbinyl diphosphate intermediate, presqualene diphosphate (PSQPP), with concomitant release of a proton and a molecule of inorganic diphosphate. In the second half-reaction, thought to occur at a second catalytic site within the enzyme active center (6–8), PSQPP undergoes heterolysis, isomerization, and reduction with NADPH to form squalene. It has been postulated that the SQS catalytic machinery consists of two nonidentical FPP binding sites (9, 10), one that binds FPP in a conformation that facilitates its cleavage to yield an allylic carbocation, and one that binds FPP in an orientation that facilitates carbocation insertion into its C2-C3 double bond.

1 The abbreviations used are: FPP, farnesyl diphosphate; PSQPP, presqualene diphosphate; SQS, squalene synthase; FPP, farnesyl-diphosphate synthase; EAS, 5-epi-aristolochene synthase.
Crystal Structure of Human Squalene Synthase

(4, 11). It has also been suggested that the two FPP molecules bind sequentially with the donor FPP binding first (9, 12) and that translocation of PSQPP from the first to the second reaction center occurs without its release from the enzyme (12, 13). Several residues thought to be involved in the SQS catalytic machinery have been identified through site-specific mutagenesis (14).

SQS catalyzes the first committed step in cholesterol formation in mammals and is an attractive site of therapeutic intervention (15–18). We undertook the structure determination of this key enzyme to aid in the design of cholesterol synthesis inhibitors of potential therapeutic importance.

**EXPERIMENTAL PROCEDURES**

**Protein Purification and Crystallization**—The protocols for expression, purification, and crystallization of the recombinant truncated form of SQS have been described previously (19). Briefly, recombinant, doubly truncated human SQS was overexpressed in *Escherichia coli* and purified by ion-exchange chromatography. Crystals of the protein-inhibitor complex were obtained by the hanging drop vapor diffusion method. Initially, small, single but irregular crystals were grown from drops containing protein at 1–2 mg ml⁻¹, inhibitor at a 1:1 molar ratio and half-strength well solution equilibrated over wells containing 25–100 mM potassium acetate, pH 5.6, and 25% polyethylene glycol 4000, 0.1 M sodium citrate, pH 5.8, and half-strength well solution equilibrated over wells containing 25–1-fold molar excess of inhibitor and protein at 3–5 mg ml⁻¹. There are three polypeptide chains/asymmetric unit, giving a *V*ₐₚ(20) of about 2.3 Å³ Da⁻¹.

**Data Collection and Heavy Atom Screening**—Crystals were flash frozen at 100 K in a gaseous nitrogen stream following transfer to a cryostabilization solution consisting of well solution made with 27% polyethylene glycol 4000 and 5% (v/v) 2-methyl-2,4-pentanediol. Diffraction data (Table I) were collected on a RAXIS IIc imaging plate detector using CuKα x-rays from a Rigaku RU200 generator operated at 40 kV, 100 mA. Data sets were processed with DENZO and SCALEPACK (21). Crystals were soaked in 1–5-fold molar excess of inhibitor and protein at 3–5 mg ml⁻¹.

**Overall Structure**—The structure of the CP-458003 complex structure. The missing residues (315–327) were not seen, so they were not included in the final model.

**RESULTS**

**Overall Structure**—The structure of SQS² is entirely α-helical, with the axes of all the helices somewhat aligned and arranged in three layers (Fig. 2, a and b). The first layer is formed by helices A, B, and K; the second layer contains helices E, C, J, and L; and the third layer is formed by D, F, G, H, I, and M. The protein is folded as a single domain, with a large channel running through the center, surrounded by helices C, F, G, H, and J.

**Evolutionary Fold Conservation**—Despite the lack of any sequence homology, the SQS core structure is similar to that of the other three class I isoprenoid biosynthetic enzymes whose crystal structures are known (Fig. 3). Of these, one is avian farnesyl-diphosphate synthase (27) (FPS), which catalyzes the synthesis of farnesyl diphosphate (the substrate of SQS) from dimethyl allyl phosphate and isopentenyl diphosphate, and the other two are cyclases that use farnesyl diphosphate as a substrate to catalyze the synthesis of cyclic sesquiterpenes: pentalenene synthase (28) (PLS) from *Streptomyces* and 5-epi-
FIG. 2. Overall structure of human squalene synthase. a, stereo Ca trace of SQS, with the NH₂ and COOH termini and every 10th Ca atom numbered. The inhibitor CP-458003 is shown in ball-and-stick representation, with carbon atoms are colored magenta, nitrogen are blue, oxygen...
aristolochene synthase (29) (EAS) from tobacco. The structural homology seen here is consistent with suggestions made earlier that all class I isoprenoid biosynthetic enzymes may have evolved with similar structures through divergence regardless of the degree of amino acid sequence identity (3, 28). The conserved feature in all the structures is an \( \alpha \)-helical core surrounding a central active site cavity, of which one end is predominantly hydrophobic, and the other end is more hydrophilic and contains a signature “aspartate-rich sequence” (30).

**Inhibitor Binding Site**—In squalene synthase, the central channel is solvent exposed, except at one end, which is covered by a “flap” formed by residues 50–54, which connect helices A and B (Fig. 2b). The side chain of Phe\(^{54} \) forms one wall of a large hydrophobic cavity under the flap, into which the inhibitors bind (Figs. 4a and 6). This hydrophobic pocket leads to the opening of the solvent exposed part of the channel, which runs through the middle of the protein. The other end of the channel has two conserved aspartate-containing sequence motifs

---

**Fig. 3.** SQS is structurally homologous to other class I isoprenoid biosynthetic enzymes. Each protein structure was individually superimposed on SQS, and all four are shown in exactly the same orientation. The percentage of identical residues in the superimposed parts is less than 16%. Five helices surrounding the active center in each enzyme are shown as magenta cylinders, with the rest of the protein backbone in yellow. Highly conserved aspartates, which are expected to be involved in binding the diphosphate moiety of the substrates via Mg\(^{2+} \) ions (2, 11, 36) are indicated in ball-and-endor representation. Of the five helices forming the active core, a kink in the helix forming the base of active center appears to be a conserved feature in all the structures, which may play a role in stabilizing cationic intermediates (details under “Discussion”). This figure was drawn using MOLSCRIPT(34) and RASTER3D(35).

---

are red, and chlorine are green. b, ribbon diagram (34, 35) of SQS bound to the inhibitor CP-458003. \( \alpha \)-Helices are shown as spirals and are colored from blue to red going from the NH\(_2 \) to the COOH terminus. The helices are numbered as below: A:38–50, B:52–59, C:65–84, D:90–98, E:126–134, F:137–156, G1:165–174, G2:177–189, H1:196–199, H2:201–218, H3:220–225, I1:252–256, I2:270–278, J:283–303, K:318–327, L:331–346, M:356–368. The inhibitor is shown in space-filling representation, with the atoms colored as described above. c, stereo view of the original electron density map calculated from solvent flattened (DM (24)) experimental MIR phases to 2.4 \( \AA \). The density around CP-320473 is shown, contoured at 1\( \sigma \).
The central cleft in SQS is lined by the most highly conserved residues in the squalene and phytoene synthase family and is the inhibitor binding site. a, close up view of the central cleft in SQS. Colored patches on the molecular surface (calculated using GRASP (37)) correspond to residues conserved in the squalene and phytoene synthases, colored similarly in Fig. 4. The loop between helices A and B (residues 50–54) forms a flap over the inhibitor binding site and is excluded from the surface calculation in this figure, shown instead as a tube. Phe84, which forms one wall of the hydrophobic pocket in which the inhibitors bind, is shown in stick form. Residues 116–119 are also excluded from the surface calculation to better show the conserved residues from helix C and helix F, which form one wall of the central cleft. b, alignment of squalene and phytoene synthase sequences. The first three sequences are squalene synthases from human (P36268), rat (Q02769), and Saccharomyces cerevisiae (P29704). The next three are phytoene synthases from tomato (P37273), Erwinia uredovora (P22872), and Rhodobacter capsulatus (P17056). The numbers in parentheses are the accession numbers of these sequences in the SWISS-PROT data base. Like SQS, phytoene synthases also catalyze a head-to-head condensation of two isoprenyl diphosphates, in this case geranyl geranyl diphosphate, to form phytoene, a 40-carbon terpene (4, 31). Two regions in the sequence that are highly conserved in the squalene synthases and not at all in the phytoene synthases are colored magenta and pale green. Five other colored regions identify residues that are highly conserved in both enzymes. The arrows mark the amino and carboxyl termini of the recombinant construct of human SQS used for crystallization.

The structures of three different inhibitor complexes are described. The two more potent compounds, CP-320473 and CP-424677 (IC50 values of 56 and 32 nM, respectively, measured as described previously (15)), fill the cavity under the flap with bulky hydrophobic groups (naphthyl and chlorophenyl group in CP-320473 and biphenyl and dichlorophenyl group in CP-424677) (Fig. 6). The third compound, CP-458003, which is an analog of CP-320473 lacking the naphthyl group, is significantly less potent (IC50 = 30 μM) and fills only a part of
aspartates and arginines and Tyr171 (discussed under “Results”) are close-up of the solvent exposed, lower end of the central cleft. Conserved directed mutagenesis experiments using rat SQS (14). Any diphosphates of two substrate FPP molecules via bridging ions that stabilize binding of diphosphate groups in the substrates. Based on this superposition, it is highly likely that the second reaction activity, although retaining first reaction activity. The structure shows that the Phe<sup>298</sup> side chain forms one wall of the hydrophobic cavity and could stabilize one of the carbocationic intermediates in the second half-reaction. Cation-π interactions for the stabilization of carbocation intermediate have been proposed for several terpenoid polyene cyclases (28, 29, 33).

**DISCUSSION**

In this work, by determining the three-dimensional structure of squalene synthase complexed with various inhibitors, we have identified some of the key residues involved in catalysis and provided a structural framework for building a de-
tailed mechanistic model of the two-step reaction catalyzed by this enzyme. Comparison with structures of other isoprenoid biosynthetic enzymes reveals a common folding architecture at the catalytic core and a conserved aspartate-rich sequence motif at the active site designed to bind prenyl phosphates via bridging Mg$^{2+}$ ions. Another common feature of all class I isoprenoid biosynthetic enzymes is the formation of a relatively stable allylic carbocation species by the release of the phosphate group. Another conserved structural feature suggests a common mechanism by which the primary carbocation could be stabilized in all the enzymes of this family. Superposition of SQS on FPS, EAS, and pentalenene synthase reveals a kink in the helix (helix G in SQS) forming the base of the catalytic cleft in each of the enzymes (residues Val175-Ala176 in SQS, Lys214-Thr215 in FPS, Thr401-Thr402 in EAS, and Ile177-Gly178 in pentalenene synthase), suggesting a common functional role for this feature (Fig. 3). One consequence of such a break in a helix is that backbone amide nitrogens and carbonyl oxygens are available for hydrogen bonding to other ligands, as they are not involved in making hydrogen bonds within the helix. The kink in helix G forms a shallow depression at the mouth of the hydrophobic pockets that are formed under the flap in squalene synthase. The backbone carbonyl atoms of residues 175 and 176 point into the central cavity, forming H-bonds with bound water molecules. In the proposed mechanism for the reaction catalyzed by EAS (29), the allylic carbocationic intermediate is positioned near the main chain carbonyl oxygens of residues 401 and 402, at the bend in helix G. A similar bend at an analogous position suggests that the same role may be played by the backbone carbonyl oxygens of residues 175 and 176 in SQS. The side chain of Arg77, which is completely conserved in all the squalene and plant phytoene synthase sequences (31) (Fig. 4), points into the same pocket from the opposite direction, suggesting that it may play a role in directing the released phosphate group away from the carbocation.

There is no evidence of a characteristic nucleotide-binding motif in the structure, which would help define the NADPH-binding site, but it is tempting to consider the J-K loop (314VKIRK318) and part of helix K as making up part of the nucleotide binding site. This domain is conserved in the squalene synthases but not in the plant and bacterial phytoene synthases, which do not require a nucleotide co-factor (31) (Fig. 3). It is also the most flexible part of the structure, indeed residues 315–327 are so disordered that they were not seen in the electron density maps calculated for the CP-320473 structure. Co-crystal structure with the inhibitor CP-458003 showed that these residues formed a loop and a helix, which runs over the top of the central catalytic cleft. It is conceivable that this part of the structure is inherently flexible and is stabilized by NADPH binding.

Native SQS is membrane-bound, but the recombinant construct used for crystallization was truncated at both NH$_2$ and COOH termini to generate soluble, active protein (19). Our structure shows that the NH$_2$- and COOH-terminal ends of the truncated protein are on the same face of the protein (the top as viewed in Fig. 2, a and b), which suggests that the membrane is closer to the “top” end of the protein. This would be consistent with the general directionality of the reaction, in which the hydrophilic substrates (FPP) bind to the “lower” end of the central cleft, and the lipophilic product squalene, leaves from the upper end, closer to the membrane.

The structure described here provides evidence that SQS has the same chain fold and overall topology as other class I prenyltransferases and terpene cyclases, suggesting an evolutionary relationship, despite the lack of sequence conservation. However, in addition to these structural similarities, the struc-

![Fig. 6. The inhibitor binding pockets. GRASP (37) surface close-up of the inhibitor binding site for a, CP-320473; b, CP-458003; and c, CP-424677. Side chains of Phe$^54$ and Tyr$^{73}$ form two walls of the inhibitor-binding cavity and are excluded from the surface calculation, shown instead as stick figures. All the flap residues (50–54) are similarly excluded from the surface calculation and shown as a backbone tube. The labeled residues in a are all in van der Waal’s contact with the inhibitor and are conserved in all SQS sequences. The size and shape of the hydrophobic pockets changes to accommodate different ligands, mainly through rotations of the side chains Phe$^{54}$ and Tyr$^{73}$, and through backbone rotations in the flap residues.](image-url)
tecture described here also demonstrates key differences between the active center regions of SQS and other class I prenyltrans-
ferases and terpene cyclases, consistent with the unique reaction mechanism catalyzed by SQS. These observations form the
basis for further evaluation, using site-directed mutagenesis together with modified substrates and substrate analogs of the
dynamics of the unique reaction mechanism catalyzed by hu-
man SQS to allow more detailed mechanistic models to be
inferred. Such studies are currently underway. In addition,
information obtained from these and future explorations into
the structural requirements for SQS-mediated conversion of
FPP to squalene will open up the possibility of modifying the
substrate specificity of SQS, through site-directed mutagenes-
sis, to create novel terpene products by engineering mutations
in the catalytic cleft. Finally, the structural information pre-

REFERENCES
1. Ourisson, G., and Nakatani, Y. (1994) Chem. Biol. 1, 11–23
2. Sacchettini, J. C., and Poulter, C. D. (1997) Science 277, 1788–1789
3. Wendt, K. U., and Schulz, G. E. (1998) Structure 6, 127–133
4. Poulter, C. D. (1990) Acc. Chem. Res. 23, 70–77
5. Jarstfer, M. B., Blagg, B. S. J., Rogers, D. H., and Poulter, C. D. (1996) J. Am.
Chem. Soc. 118, 13098–13099
6. Poulter, C. D., Capson, T. L., Thompson, M. D., and Bard, R. S. (1996) J. Am.
Chem. Soc. 111, 3734–3739
7. Corey, E. J., and Volante, R. P. (1976) J. Am. Chem. Soc. 98, 1291–1293
8. Petras, S. F., Lindsey, S., and Harwood, H. J., Jr. (1999) J. Lipid Res. 40, 24–38
9. Washburn, W. N., and Kow, R. (1977) Tetrahedron Lett. 18, 1555–1558
10. Ortiz de Montellano, P. R., Wei, J. S., Vinson, W. A., Castillo, R., and Boparai,
A. S. (1977) Biochemistry 16, 2680–2685
11. Poulter, C. D., andBilling, H. C. (1981) in Biosynthesis of Isoprenoid
Compounds (Spurgeon, S. L., and Porter, J. W., eds) Vol. I, pp. 413–441,
12. Moskhtiar, K. A., Kalinowski, S. S., Zhang, D., and Poulter, C. D. (1994)
J. Biol. Chem. 269, 11201–11207
13. Popják, G., and Agnew, W. S. (1979) Mol. Cell. Biochem. 27, 97–116
14. Wu, P., Ishii, Y., Spencer, T. A., and Schechter, I. (1998) J. Biol. Chem. 273,
12515–12523
15. Harwood, H. J., Jr., Bardac-Toh, E. G., Petras, S. F., Lindsey, S., and
Pellarin, L. D. (1997) Biochem. Pharmacol. 53, 839–864
16. Lindsey, S., and Harwood, H. J., Jr. (1995) J. Biol. Chem. 270, 9083–9096
17. Bergstrom, J. D., Dufrènes, C., Bills, G. F., Nallin-Onstead, M., and Byrne, R.
(1995) Annu. Rev. Microbiology 49, 607–639
18. Biller, S. A., Neuschwander, K., Ponnipom, M. M., and Poulter, C. D. (1996)
Curr. Pharm. Des. 2, 1–40
19. Thompson, J. P., Danley, D. E., Mazzalupo, S., Milos, P. M., Lira, M. E., and
Harwood, H. J., Jr. (1998) Arch. Biochem. Biophys. 358, 283–290
20. Matthews, B. W. (1968) J. Mol. Biol. 33, 491–497
21. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
22. Tong, L., and Rossmann, M. G. (1993) J. Appl. Crystallogr. 26, 15–21
23. Bailey, S. (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
24. Cowtan, K. D. (1994) in Joint CCP4 and ESRF-EACBM Newsletter on Protein
Crystallography 31, 34–38
25. Brunger, A. T. (1996) X-PLOR (online) A System for X-ray Crystallography and
NMR, Version 3.851, Yale University Press, New Haven, CT
26. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta
Crystallogr. Sect. A 47, 110–119
27. Tarshis, L. C., Yan, M., Poulter, C. D., and Sacchettini, J. C. (1994) Biochem-
istry 33, 10871–10877
28. Lesburg, C. A., Zhai, G., Can, D. E., and Christianson, D. W. (1997) Science
277, 1820–1824
29. Starks, C. M., Back, K., Chappell, J., and Noell, J. P. (1997) Science 277,
1815–1820
30. Ashby, M. N., and Edwards, P. A. (1990) J. Biol. Chem. 265, 13157–13164
31. Robinson, G. W., Tsay, Y. H., Kienzle, B. K., Smith-Monroy, C. A., and Bishop,
R. W. (1995) Mol. Cell. Biol. 13, 2706–2717
32. Tarshis, L. C., Proteau, P. J., Kellogg, B. A., Sacchettini, J. C., and Poulter,
C. D. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 15018–15023
33. Wendt, K. U., Poralla, K., and Schulz, G. E. (1997) Science 277, 1811–1815
34. Kruulis, P. J. (1994) J. Appl. Crystallogr. 24, 945–949
35. Merritt, E. A., and Bacon, D. J. (1997) Methods Enzymol. 277, 505–524
36. Ashby, M. N., Kutsunai, S. Y., Ackerman, S., Tzagoloff, A., and Edwards, P. A.
(1997) J. Biol. Chem. 272, 4128–4136
37. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Proteins Struct. Funct. Genet.
11, 281–296