Identification of the Active Site of Vertebrate Oxidosqualene Cyclase[1]

Ikuro Abe and Glenn D. Prestwich*

Department of Chemistry, University at Stony Brook, Stony Brook, New York 11794-3400

ABSTRACT: Active site mapping of rat liver oxidosqualene cyclase (OSC), a 78 kDa membrane-bound enzyme, was carried out using the mechanism-based irreversible inhibitor, [^{3}H]29-methylidene-2,3-oxidosqualene. The amino acid sequence of the radiolabeled CNBr peptide fragment showed unexpectedly high similarity to the yeast OSC, plant OSC, and bacterial squalene cyclases. Further, radio analysis established that the two adjacent Asp residues in the highly conserved region (*Asp-Asp*-Thr-Ala-Glu-Ala, or DDTAEA) were equally labeled by the irreversible inhibitor. This result provided the first information on the structural details of the active site of OSC, and showed for the first time the ancient lineage of this vertebrate enzyme to ancestral eukaryotic and prokaryotic cyclases. Interestingly, the covalently-modified DDXX(D/E) sequence of rat liver OSC showed surprising similarity to the putative allylic diphosphate binding site sequence of other terpene cyclases and prenyl transferases. The Asp-rich motif may act as a point charge to stabilize incipient cationic charge.

Lipids 30, 231–234 (1995).

The enzymatic cyclization of oxidosqualenes is the most remarkable step in the biosynthesis of sterols and triterpenes (1). In a single step, an acyclic polyene is converted by a specific enzyme to either a tetracyclic or pentacyclic triterpene. In the case of lanosterol formation, (3*S*)-oxidosqualene first adopts a prechair–boat–chair conformation. Then protonation of the epoxide initiates a cascade of ring-forming reactions and backbone rearrangements that proceed through rigidly held carbocationic intermediates (2–6). Several oxidosqualene cyclases (OSC) (EC 5.4.99.7) have been purified to homogeneity from vertebrate (7–9), plant (10–13), and yeast sources (14). Recently, three OSC enzymes have been cloned and sequenced from two fungal sources, *Saccharomyces cerevisiae* (15,16) and *Candida albicans* (17–19) and, from one plant, *Arabidoposis thaliana* (20). The predicted molecular masses ranged from 80 to 90 kDa, and the cDNA sequences predicted proteins with significant sequence similarity. In addition, sequence comparisons of these cyclases with bacterial squalene cyclases (SC) revealed an existence of a highly-conserved repetitive motif rich in aromatic amino acids (the QW motif) (21,22).

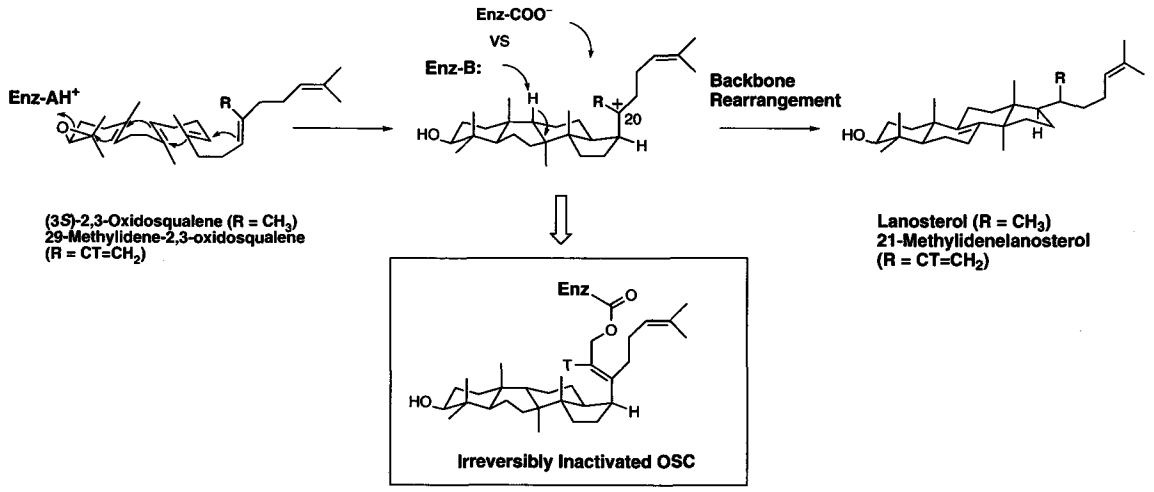
Recent work from our laboratories on active site mapping of rat liver OSC using 29-methylidene-2,3-oxidosqualene (29-MOS), the first mechanism-based irreversible inactivator of vertebrate OSC, will be summarized in this paper. 29-MOS was first prepared in racemic form and showed potent inhibition of OSC. An IC_{50} value of 0.5 μM, an apparent K_i value of 4.4 μM, a k_{inact} value of 221 min^{-1}, and a partition ratio of 3.8 (inactivations per turnover) were obtained for pig liver OSC (23). Incubation of 0.1 μM 29-MOS with pig liver OSC gave a polycyclic product in 30% yield; however, at [29-MOS] > K_i, complete inactivation precluded isolation of this product (23). A new synthesis of the enantiomerically-enriched substrate, [^{3}H](3*S*)-29-MOS, by the Sharpless asymmetric dihydroxylation reaction (24–26) was achieved recently (27). Several other 26- and 29-functionalized 2,3-oxidosqualene derivatives were found to be slightly less active than 29-MOS (28).

Vertebrate OSC enzymes can be efficiently and specifically labeled with tritium-labeled 29-MOS (8). For a typical affinity labeling experiment, enzyme preparations were first incubated with [^{3}H]29-MOS (1.0 μM, 2.3 Ci/mmol) for 15 min at 37°C. Next, the proteins were separated by denaturing polyacrylamide gel electrophoresis (SDS–PAGE), and the covalently-modified protein bands in fluorophore-enhanced, dried gels were visualized on X-ray film by fluorescence autoradiography. For each enzyme source, a single radiolabeled protein band was observed having a molecular mass between 70 and 80 kDa (rat, 78 kDa; pig, 75 kDa; dog, 73 kDa; and human, 73 kDa). The proposed mechanism of inhibition by 29-MOS involves initial cyclization to the 21-methylidene-protosterol cation as proposed for lanosterol formation. However, we have postulated that instead of backbone rearrangement, this allylic cation can be trapped by an active-site nucleophile, resulting in irreversible inactivation of the enzyme (Scheme 1). Moreover, this hypothesis also predicts that the structures of the released and enzyme-bound tetracyclic products are different; 21-methylidene lanosterol should be released, while the covalent moiety attached to the enzyme

[1]Based on a paper presented at the Symposium on the "Regulation of Biosynthesis and Function of Isopentenoids," Atlanta, Georgia, May 1994.

*To whom correspondence should be addressed.

Abbreviations: 29-MOS, 29-methylidene-2,3-oxidosqualene; OSC, oxidosqualene cyclase; SC, squalene cyclase; SDS–PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis.



SCHEME 1

should retain the unrearranged protosterol skeleton. Verification of these structures is in progress.

In contrast, yeast and plant OSC enzymes were not labeled even under forcing conditions, and [^{3}H]29-MOS was recovered unchanged by chromatography (8). Because 29-hydroxy-2,3-oxidosqualene is quantitatively converted to 21-hydroxylanosterol by yeast OSC (29), it seemed most likely that 29-MOS would be accepted as a substrate. The initially formed 21-methylidene protosterol cation could be stabilized by the enzyme, preventing backbone rearrangement to the cyclization products. In this case, instead of covalent bond formation, the stabilization would probably be attained by an electrostatic interaction. As a result, 29-MOS appears to act as a slow tight-binding inhibitor; and an IC_{50} value of 1.5 μM, an apparent K_i value of 1.5 μM, and a k_{inact} value of 0.088 min^{-1} were obtained for yeast OSC (8).

In order to localize the 29-MOS binding site region of rat liver OSC, purified enzyme was labeled with [^{3}H]29-MOS and then selectively cleaved adjacent to Met residues with CNBr (30). A single radioactive 6 kDa peptide fragment was separated by Tricine SDS–PAGE. After electrophoretic transfer to a polyvinylidene difluoride membrane, the labeled peptide was sequenced by Edman degradation (30). The sequence of the first 30 amino acids of the peptide showed unexpectedly high similarity to the other eukaryotic sterol and terpene cyclases: 50% identity with Arg^{439}–Val^{468} of yeast OSC (15,16), 50% identity with Arg^{433}–Val^{462} of *C. albicans* OSC (18,19), 63% identity with Ser^{466}–Leu^{495} of plant *A. thaliana* OSC (cycloartenol synthase) (20), and to the bacterial SC enzymes, 33% identity with Lys^{356}–Leu^{384} of *Alicyclobacillus* (formerly *Bacillus*) *acidocaldarius* SC (31), 27% identity with Arg^{375}–Ala^{404} of *Zymomonas mobilis* SC (32) (Fig. 1). This result provides the first information on the structural details of the active site of OSC and demonstrates the ancient lineage of this vertebrate enzyme with respect to the ancestral eukaryotic and prokaryotic cyclases.

Furthermore, radioanalysis clearly established that the two adjacent Asp residues in the highly-conserved region (*Asp*-Thr-Ala-Glu-Ala, or DDTAEA) were equally labeled by the irreversible inhibitor (30). The initially formed 21-methylidene-protosterol cation would be trapped by nucleophilic attack by the carboxyl group of one of the two adjacent Asp residues, resulting in ester bond formation. The two Asp residues appear to be equally accessible to the methylidene-extended protosterol cation. In the case of yeast or plant OSC enzymes, the appropriate nucleophilic residues seem to either be absent or not to have access to the allylic cation. Indeed, each of these enzymes has a Cys residue in the second position, giving a characteristic DCTAEA motif. Recently, rat liver OSC was cloned and sequenced in our laboratory (Abe, I., and Prestwich, G.D., unpublished data). According to the deduced amino acid sequence of this cDNA, the 29-MOS binding site has the DCTAEA sequence instead of DDTAEA. Isozymes of this enzyme with the DDTAEA sequence may exist, or there may be sequencing anomalies associated with a lipid-modified amino acid. We are exploring these two possibilities.

Interestingly, the covalently-modified DDTAEA sequence of rat liver OSC showed surprising similarity to the known consensus sequence, Asp-Asp-Xaa-Xaa-Asp (the "DDXXD" motif) found in diterpene cyclases (33), sesquiterpene cyclases (34–37), and monoterpene cyclases (33), as well as in several prenyl transferases (38–40) (Fig. 1). Each of these enzymes uses an allylic diphosphate as a substrate and divalent metal ions as a co-factor. Thus, this aspartate-rich domain has been implicated in catalyzing the ionization of the allylic diphosphate to the reactive allylic cation, presumably by coordination *via* an Mg^{2+}-diphosphate bridge. Recent site-directed mutagenesis experiments of the conserved aspartate residues of farnesyl–diphosphate synthase suggest that the first Asp residue in this domain is important in catalysis (41–43). The high degree of sequence conservation of the 29-MOS binding site in triterpene and sterol cyclases supports a catalytic or structural function for this domain. Stabilization of tertiary or allylic carbocations may be the true raison d'être for the DDXX(D/E) motif. We present this hypothesis in

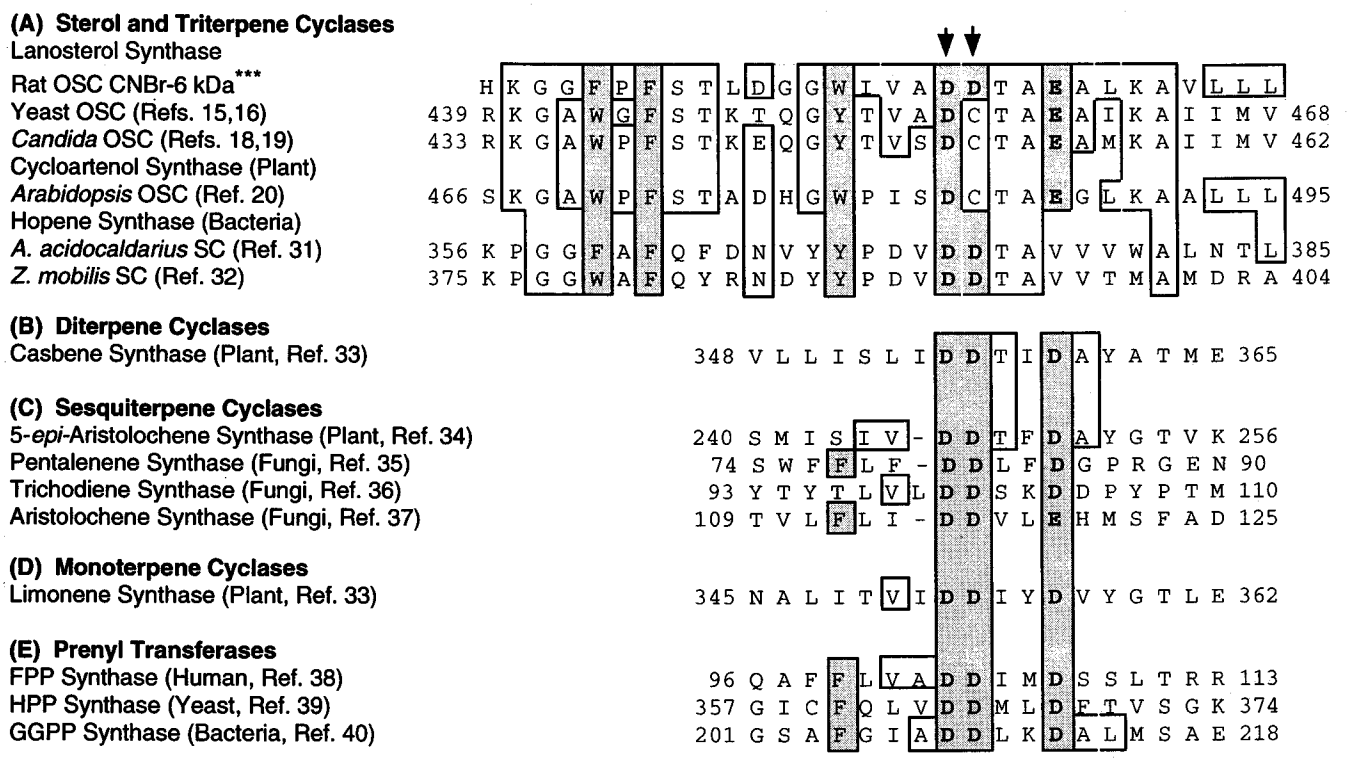
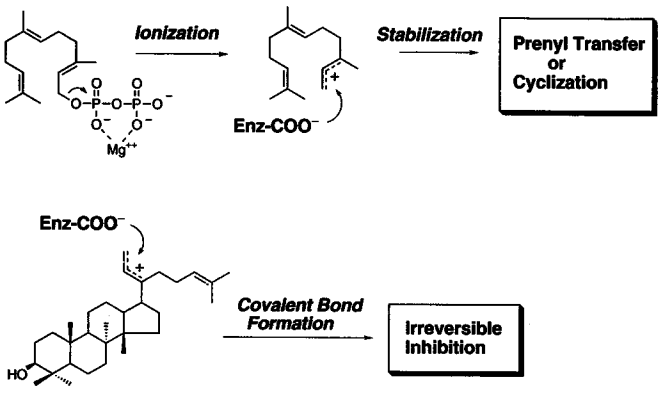


FIG. 1. Comparison of the DDXX(D/E) motif in the predicted amino acid sequences of terpene biosynthetic enzymes. OSC, oxidosqualene cyclase; SC, squalene cyclase.



SCHEME 2

Scheme 2, and we suggest that rather than assisting the departure of the pyrophosphate, the Asp-rich motif may act as a point charge to stabilize incipient cationic charge. For transferases and other cyclases, this corresponds to stabilization of the transition state to a high-energy intermediate cation at the initiation of the reaction, or perhaps during cyclization. For OSC, the Asp-rich motif would stabilize a "stopping point" in the cyclization pathway, i.e., after tetracycle formation but prior to backbone rearrangement. Very recently, Poralla *et al.* (44) reported that a point mutation of the first Asp residue of the DDTAVV motif of bacterial SC (*A. acidocaldarius*) can cause almost complete loss of the cyclase activity. Additional single mutations in cyclases will provide further tests for this hypothesis. Finally, it should be noted that the existence of such a consensus sequence in the terpene and sterol biogenesis enzymes is very interesting, considering the molecular evolution of these enzymes.

ACKNOWLEDGMENTS

The authors are indebted to T. Fischer at the Center of Analysis and Synthesis of Macromolecules (Stony Brook) for Edman peptide sequencing and Dr. X.-y. Xiao for synthesis of the [^{3}H]29-MOS used in this study. This work was supported by the Center for Biotechnology/New York State Science and Technology Foundation, Sandoz Research Institute and the National Institutes of Health (GM 44836).

REFERENCES

1. Abe, I., Rohmer, M., and Prestwich, G.D. (1993) *Chem. Rev. 93*, 2189–2206.
2. Eschenmoser, A., Ruzicka, L., Jeger, O., and Arigoni, D. (1955) *Helv. Chim. Acta 38*, 1890–1904.
3. Stork, G., and Burgstahler, A.W. (1955) *J. Am. Chem. Soc. 77*, 5068–5077.
4. van Tamelen, E.E. (1982) *J. Am. Chem. Soc. 104*, 6480–6481.
5. Corey, E.J., Virgil, S.S., and Sarshar, S. (1991) *J. Am. Chem. Soc. 113*, 8171–8172.
6. Corey, E.J., and Virgil, S.C. (1991) *J. Am. Chem. Soc. 113*, 4025–4026.
7. Kusano, M., Abe, I., Sankawa, U., and Ebizuka, Y. (1991) *Chem. Pharm. Bull. 39*, 239–241.

8. Abe, I., Bai, M., Xiao, X.-y., and Prestwich, G.D. (1992) *Biochem. Biophys. Res. Commun. 187*, 32–38.
9. Moore, W.R., and Schatzman, G.L. (1992) *J. Biol. Chem. 267*, 22003–22006.
10. Abe, I., Ebizuka, Y., and Sankawa, U. (1988) *Chem. Pharm. Bull. 36*, 5031–5034.
11. Abe, I., Sankawa, U., and Ebizuka, Y. (1989) *Chem. Phar. Bull. 37*, 536–538.
12. Abe, I., Ebizuka, Y., Seo, S., and Sankawa, U. (1989) *FEBS Lett. 249*, 100–104.
13. Abe, I., Sankawa, U., and Ebizuka, Y. (1992) *Chem. Pharm Bull. 40*, 1755–1760.
14. Corey, E.J., and Matsuda, S.P. (1991) *J. Am. Chem. Soc. 113*, 8172–8174.
15. Corey, E.J., Matsuda, S.P.T., and Bartel, B. (1994) *Proc. Natl. Acad. Sci. USA 91*, 2211–2215.
16. Shi, Z., Buntel, C., and Griffin, J.H. (1994) *Proc. Natl. Acad. Sci. USA 91*, 7370–7374.
17. Kelly, R., Miller, S.M., Lai, M.H., and Kirsch, D.R. (1990) *Gene 87*, 177–183.
18. Buntel, C.J., and Griffen, J.H. (1992) *J. Am. Chem. Soc. 114*, 9711–9713.
19. Roessner, C.A., Min, C., Hardin, S.H., Harris-Haller, L.W., McCollum, J.C., and Scott, A.I. (1993) *Gene 127*, 149–150.
20. Corey, E.J., Matsuda, S.P.T., and Bartel, B. (1993) *Proc. Natl. Acad. Sci. USA 90*, 11628–11632.
21. Poralla, K., Hewelt, A., Prestwich, G.D., Abe, I., Reipen, I., and Sprenger, G. (1994) *Trends Biochem. Sci. 19*, 157–158.
22. Poralla, K. (1994) *Bioorg. Med. Chem. Lett. 4*, 285–290.
23. Xiao, X.-y., and Prestwich, G.D. (1991) *J. Am. Chem. Soc. 13*, 9673–9674.
24. Sharpless, K.B., Amberg, W., Bennani, Y.L., Crispino, G.A., Hartung, J., Jeong, K.-S., Kwong, H.-L., Morikawa, K., Wang, Z.-M., Xu, D., and Zhang, X.-L. (1992) *J. Org. Chem. 57*, 2768–2771.
25. Crispino, G.A., and Sharpless, K.B. (1992) *Tetrahedron Lett. 33*, 4273–4274.
26. Crispino, G.A., Ho, P.T., and Sharpless, K.B. (1993) *Science 259*, 64–66.
27. Madden, B.A., and Prestwich, G.D. (1994) *J. Org. Chem. 59*, 5488–5489.
28. Abe, I., Tomesch, J.C., Wattanasin, S., and Prestwich, G.D. (1994) *Nat. Prod. Rep. 11*, 279–302.
29. Xiao, X.-y., and Prestwich, G.D. (1991) *Tetrahedron Lett. 32*, 6843–6846.
30. Abe, I., and Prestwich, G.D. (1994) *J. Biol. Chem. 269*, 802–804.
31. Ochs, D., Kaletta, C., Entian, K.-D., Beck-Sickinger, A., and Poralla, K. (1992) *J. Bacteriol. 174*, 298–302.
32. Reipen, I., Sprenger, G., and Poralla, K. (1993) *GenBank Data Base*, accession no. X73561.
33. Colby, S.M., Alonso, W.R., Katahira, E.J., McGarvey, D.J., and Croteau, R. (1993) *J. Biol. Chem. 268*, 23016–23024.
34. Facchini, P.J., and Chappel, J. (1992) *Proc. Natl. Acad. Sci. USA 89*, 11088–11092.
35. Cane, D.E., Sohng, J.-K., Lamberson, C.R., Rudnicki, S.M., Wu, Z., Lloyd, M.D., Oliver, J.S., and Hubbard, B.R. (1994) *Biochemistry 33*, 5846–5857.
36. Hohn, T.M., and Beremand, P.D. (1989) *Gene 79*, 131–138.
37. Proctor, R.H., and Hohn, T.M. (1993) *J. Biol Chem. 268*, 4543–4548.
38. Sheares, B.T., White, S.S., Molowa, D.T., Chan, K., Ding, V.D.H., Kroon, P.A., Bostedor, R.G., and Karkas, J.D. (1989) *Biochemistry 28*, 8129–8135.
39. Ashby, M.N., and Edwards, P.A. (1990) *J. Biol. Chem. 265*, 13157–13164.
40. Math, S.K., Hearst, J.E., and Poulter, C.D. (1992) *Proc. Natl. Acad. Sci. USA 89*, 6761–6764.
41. Marrero, P.F., Poulter, C.D., and Edwards, P.A. (1992) *J. Biol. Chem. 267*, 21873–21878.
42. Joly, A., and Edwards, P.A. (1993) *J. Biol. Chem. 268*, 26983–26989.
43. Song, L., and Poulter, C.D. (1994) *Proc. Natl. Acad. Sci. USA 91*, 3044–3048.
44. Poralla, K., Ochs, D., and Feil, C. (1994) *INFORM 5*, 470.

[Received July 14, 1994, and in revised form December 23, 1994; Revision accepted December 27, 1994]