Crystal structures of the monofunctional chorismate mutase from *Bacillus subtilis* and its complex with a transition state analog

(x-ray crystallography/shikimate pathway/claisen rearrangement)

YUH MIN CHOOK*, HENGMING KE†, AND WILLIAM N. LIPSCOMB*

*Gibbs Chemical Laboratory, Harvard University, Cambridge, MA 02138; and †Department of Biochemistry and Biophysics, Medical School, University of North Carolina, Chapel Hill, NC 27599

Contributed by William N. Lipscomb, June 18, 1993

ABSTRACT We have solved the structure of a chorismate mutase (chorismate pyruvatemutase, EC 5.4.99.5), the 1.9-Å crystal structure of the monofunctional enzyme from *Bacillus subtilis*. The structure determination process was an unusual one, involving 12 monomers of the enzyme in the asymmetric unit. This structure was solved by the multiple isomorphous replacement method with partial structure phase combination and molecular averaging. The final model, which includes 1380 residues and 522 water molecules in an asymmetric unit, has been refined at 1.9 Å and the current crystallographic *R* value is 0.201. The *B. subtilis* chorismate mutase is a homotrimer, with β -sheets from each monomer packing to form the core of a pseudo- $\alpha\beta$ -barrel with helices on the outside of the trimer. In addition, the active sites have been located by using data from a complex with an *endo*-oxabicyclic inhibitor that mimics the transition state of the reaction. The structure of this complex has been refined to 2.2 Å with a current *R* value of 0.182 for a model that includes 1388 residues, 12 inhibitor molecules, and 530 water molecules in the asymmetric unit. In each trimer, three equivalent active sites are located at the interfaces of two adjacent subunits.

Chorismate mutase (chorismate pyruvatemutase, EC 5.4.99.5) lies at a branch point in the biosynthetic pathway of aromatic amino acids (1). It is the only characterized enzyme that catalyzes a pericyclic process, the unimolecular rearrangement of chorismate to prephenate (see Fig. 1). Neither the mechanistic details of the enzyme-catalyzed rearrangement nor the origins of the 2×10^6 -fold rate enhancement over the uncatalyzed process (2) are known. The enzymatic and nonenzymatic rearrangement of chorismate to prephenate is a concerted, asynchronous reaction, which proceeds via a transition state with chair-like geometry (3, 5). Strong binding of an *endo*-oxabicyclic transition state analog (4) (Fig. 1) suggests that the enzyme-catalyzed rearrangement also likely proceeds via the chair-like transition state (3, 5) without observed intermediates. However, neither the rate-limiting transition state nor the mechanistic pathway of this reaction has been determined.

Catalytic antibodies with chorismate mutase activities have also been prepared, accelerating the reaction by a factor of 2×10^4 in one case (6) and by a factor of 2×10^2 in another (7, 8). Comparison of the unimolecular rearrangement of chorismate to prephenate in the different proteins should provide insights into the evolution of specificity, catalytic mechanism, and origins of rate enhancement in both enzymes and engineered proteins. Furthermore, as chorismate mutase lies at a branch point in the shikimate pathway, which exists only in bacteria, fungi, and higher plants, the enzyme may be extremely useful in development of herbicides and antibacterial therapeutics. The poor sequence homology between

FIG. 1. Claisen rearrangement of chorismate to prephenate, proceeding via a transition state with chair-like geometry (3, 5). The *endo*-oxabicyclic inhibitor closely mimics the putative transition state (4).

chorismate mutases from different organisms also provides potential for design of specific inhibitors targeted at specific plants and microorganisms.

We report here the structure of a chorismate mutase, the 1.9-Å crystal structure of the monofunctional enzyme from the Marburg strain of *Bacillus subtilis*. This is the smallest natural chorismate mutase known, with 127 residues per subunit (9, 10). Furthermore, this enzyme is particularly amenable to study since, unlike other chorismate mutases (11, 12), it is monofunctional, nonallosteric, unaffected by end-product amino acids and effectors, and follows Michaelis–Menten kinetics (10). We have also located the active sites by using data at 2.2 Å from a complex of the enzyme with an *endo*-oxabicyclic inhibitor that mimics the putative transition state of the reaction. Although the structures of the enzyme and its active site are reported here,† a detailed analysis of its mechanism of catalysis and results of refinement of both the native enzyme and the transition state analog complex at high resolution will be published elsewhere.

MATERIALS AND METHODS

Purification and Crystallization. Purified chorismate mutase from *B. subtilis* (10) was a gift from Jeremy R. Knowles (Harvard University). The *B. subtilis* chorismate mutase native crystals were grown with protein solution (12 mg/ml) in microdialysis wells dialyzed against solution containing 5 mM Tris-HCl, 1 mM 2-mercaptoethanol, 0.5 mM NaN₃, 0.1

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: MIR, multiple isomorphous replacement.

†The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (references 1CHS, 1CHT).

Table 1. Heavy atom derivative preparation

Compound	Concentration, mM	pH	Time, days
Ethyl mercury phosphate	0.2	5.3	2
Iridium hexachloride	0.2	5.3	2
Osmium hexachloride	0.2	5.3	2
Osmium hexachloride	1.0	5.3	2

mM EDTA, and 11–12% polyethylene glycol (PEG) 3350 at pH 5.3. Crystals grew in ≈ 7 days to dimensions of $0.4 \times 0.5 \times 0.6$ mm³ and diffracted to 1.9-Å resolution. These crystals belong to space group P2₁ ($a = 102.4$ Å, $b = 68.3$ Å, $c = 102.8$ Å, $\beta = 105.6^\circ$) and have the unusual number of 12 monomers of the enzyme in its asymmetric unit.

Structure Determination. Heavy atom derivatives were prepared by soaking crystals in a stabilizing solution (20 mM Bicine/40% PEG 3350/0.1 mM NaN₃, pH 5.3) containing heavy atom reagents (see Table 1 for details of the soaking conditions). All data sets were collected at the Resource for Crystallography located in the laboratory of N. h. Xuong (University of California, San Diego) with the use of Mark III, two multiwire proportional chambers (13), and graphite-monochromated CuK α radiation from a Rigaku RU200 rotating anode (see Table 2 for data collection statistics).

Difference Patterson maps were used to locate the heavy atom sites. Twelve mercury sites were located by manual interpretation of the difference Patterson map of the ethyl mercury phosphate derivative and single isomorphous replacement phases calculated with these sites were used in difference and crossed-phase Fourier maps to locate the remainder of heavy atom binding sites in the mercury derivative and in the other derivatives. Heavy atom parameters were then refined with programs HEAVY (14) and PHASES.[§] Anomalous scattering was used to determine the set of heavy atom sites with the correct handedness.

The multiple isomorphous replacement (MIR) phase determination was particularly challenging due to the large numbers of heavy atom sites in various derivatives (Table 2). Low-concentration heavy atom solutions were required to minimize nonisomorphism, resulting in low occupancy heavy atom sites, weak MIR phases, and uninterpretable electron density maps. However, this problem was solved by biosynthetic preparation of a selenomethionine substituted enzyme. The selenomethionine enzyme was prepared by growing *Escherichia coli* strain BL21 transformed with pBSCM2 (plasmid containing the chorismate mutase gene) in minimal medium supplemented with 19 amino acids and selenome-

thionine. Data from this selenomethionine protein were also collected at the Resource for Crystallography located in the laboratory of N. h. Xuong. MIR phases from refinement of the heavy atom sites of the soaked derivatives were used in crossed-phase Fourier maps to obtain the selenium sites. Seventy-two selenium sites were found and they correspond to six highly occupied selenium sites in each monomer. Incorporation of this selenomethionine derivative into MIR phase refinement finally produced an interpretable electron density map (see Table 2 for summary of multiple isomorphous replacement data).

The electron density map calculated with these MIR phases allowed a partial model polyaniline trace of all the helices and β -strands to be made using FRODO (15, 16). Phase combination of the MIR phases and the partial polyaniline model phases (17) further improved the electron density and allowed identification of side chains and loop regions to complete the trace. Then, 144-fold improper noncrystallographic symmetry averaging was done, using hand drawn masks for each monomer, electron density calculated from MIR phases, and Bricogne's averaging programs (18) to provide an independent crystallographic check of the model.

The final model consisting of 1380 residues and 522 water molecules in the asymmetric unit was refined by using XPLOR (19) to an R value of 0.194 for 71,847 reflections from 10.0–1.9 Å. The root mean square deviations from ideal values are 0.014 Å for bond lengths and 2.92° for bond angles.

To identify and study the active site of chorismate mutase, we also determined the three-dimensional structure of its complex with an *endo*-oxabicyclic transition state analog (4). This inhibitor was a gift from Paul A. Bartlett (University of California, Berkeley). Crystals of the enzyme-inhibitor complex were prepared by soaking native crystals in a stabilizing solution (5 mM Tris-HCl/40% PEG 3350, pH 5.3) containing 0.5 mM inhibitor for 8 days. Data to 2.2 Å with $R_{\text{sym}} = 11.0\%$ were collected on the Hamlin Multiwire Systems Mark II (13) at Rutgers University. $F_o - F_c$ difference Fourier density (F_o is observed structure factor of the complex of chorismate mutase with the *endo*-oxabicyclic inhibitor; F_c and phases are from the refined model of the unbound enzyme) clearly revealed the inhibitor binding in the active site of the enzyme (see Fig. 3). The refined model of the native enzyme was used as the starting model for refinement of this complex. All 12 inhibitors in the asymmetric unit were subsequently incorporated into the model by using $F_o - F_c$ and $2F_o - F_c$ maps (F_o is observed structure factor of the complex and F_c is structure factor of the refined model of the complex without its inhibitors). This model, with 1388 residues, 12 inhibitor molecules, and 530 water molecules, has been refined to 2.2 Å with R value currently at 0.182.

[§]Furey, W. & Swaminathan, S., American Crystallographic Association Meeting, April 8–13, 1990, New Orleans.

Table 2. Crystallographic and MIR data

Parameter	Native	HG	IR	OS1	OS2	Se-Met
Resolution, Å	1.9	2.7	2.5	2.5	2.5	2.2
Reflections, observed	208,894	168,648	213,528	233,898	216,981	182,472
Reflections, unique	80,528	26,102	41,120	36,665	37,547	46,010
Completeness, %	75.0	71.0	86.9	76.8	77.8	75.8
R_{sym} , %	5.5	5.6	6.3	4.7	6.8	5.3
MIR statistics						
Number of sites		24	24	24	24	72
R_{iso} , %		9.7	15.2	12.2	21.0	15.5
R_{cullis} (15.0–3.0 Å)		0.67	0.98	0.98	1.00	0.60
Phasing power (15.0–3.0 Å)		0.77	1.05	1.15	1.23	1.11

R_{sym} is the R factor on F between symmetry-related reflections; MIR statistics were calculated by using HEAVY (14); $R_{\text{iso}} = \Sigma (|F_{\text{der}}| - |F_{\text{nat}}|) / \Sigma |F_{\text{nat}}|$; $R_{\text{cullis}} = \Sigma (|F_{\text{H}}| - (|F_{\text{PH}}| - |F_{\text{P}}|)) / \Sigma |F_{\text{H}}|$; phasing power = $(F_{\text{H}}) / (\text{closure error})$; figure of merit (15.0–3.0 Å) = 0.66. HG, ethyl mercury phosphate; IR, iridium hexachloride; OS1 and OS2, osmium hexachloride; Se-Met, selenomethionine.

RESULTS AND DISCUSSION

Monomer Structure. The chorismate mutase monomer consists of a single domain (Fig. 2a). Its N-terminal methionine and the 12 C-terminal residues are not observed in the electron density map because of structural disorder. The remainder of the monomer folds into a five-stranded mixed β -sheet packed against an 18-residue α -helix and a two-turn 3_{10} helix. The C-terminal tail consists of one turn of 3_{10} helix (residues 112–115) that protrudes away from the β -sheet on the face opposite the helices into the solvent region.

The three inner strands, I, II, and IV, of the β -sheet (composed of residues 2–11, 44–50, and 88–94) are antiparallel and the two outer strands, III and V (composed of residues 73–76 and 107–109), are parallel. This β -sheet is characterized by $+2x, +1x, -2, -2$ topology (21). The 18-residue α -helix H1 (residues 18–34) is connected to strands I and II by 8- and 9-residue loops, respectively. The 6-residue

FIG. 2. (a) Ribbon drawing of chorismate mutase monomer generated by using MOLSCRIPT (20). β -Strands (I–V) and helices (H1 and H2) are numbered in order from the N terminus. N and C termini are labeled N and C, respectively. (b) Ribbon drawing (20) of chorismate mutase trimer in a view looking down the barrel/pseudo-threefold axis.

3_{10} helix H2 is parallel to the α -helix and is connected to strands II and III by 9- and 8-residue loops, respectively (see Fig. 2a).

Trimer Structure. Prior to structure determination by x-ray crystallography, the monofunctional chorismate mutase from *B. subtilis* was reported to be a homodimer in one study (10) and a homotrimer in another (22). Interpretation of the electron density map clearly revealed the enzyme to be a homotrimer with the three monomers packing to form a pseudo- $\alpha\beta$ -barrel. The subunits of the trimer are related by pseudo-3-fold symmetry (Fig. 2b). The core of this trimer consists of large portions of the β -sheet (strands I–IV) from each monomer, together forming a closed barrel of β -strands. Side chains protruding into the core of the barrel are almost entirely hydrophobic (Ile-3, Val-43, Leu-46, Pro-72, Met-76, and Met-92). This core of β -sheets is surrounded by three α -helices and three 3_{10} helices.

Near the surface of the trimer, the interfaces between adjacent subunits form three equivalent clefts (Fig. 3a). The walls of each cleft are composed of the 3_{10} helix (residues 59, 60, 62, and 63), a loop (residues 66–71), and β -strand III (residues 72–76) from one subunit, and of β -strand V (residues 104–108), parts of β -strands I and IV (residues 7–10 and 89–91), and the C-terminal tail (residues 111–115) from the adjacent subunit. This cleft is capped with a loop (residues 78–89) from the adjacent subunit.

Active Site Structure. The active site of chorismate mutase has been located by using data from a complex of the enzyme with an *endo*-oxabicyclic transition state analog (4), $K_i = 3 \mu\text{M}$ (10). Binding of this inhibitor in the cleft described above, creating a shared active site, is clearly indicated in the difference electron density (Fig. 3a). Refinement of this model to 2.2 Å indicates little change in the conformation of the enzyme upon ligation in the active-site region.

Numerous interactions are present between enzyme and inhibitor (for details, see legend to Fig. 3b). These interactions include hydrophobic contacts (between the inhibitor and side chains Phe-57 and Ala-59 from one monomer and Leu-115 from the adjacent monomer), ionic contacts (between a carboxylate group on the inhibitor and Arg-7 and Arg-90 and between the other carboxylate group and Arg-116), and polar contacts (between the hydroxyl of the inhibitor and side chains of Glu-78, Cys-75, and main chain amide of Cys-75 and between the ether oxygen of the inhibitor and side chain of Arg-90). However, there are no functional groups capable of proton transfer to the ether oxygen of the inhibitor. This observation is consistent with the flat pH profile of k_{cat}/K_m between pH 4 and pH 9 and the lack of solvent isotope effect for the *B. subtilis* chorismate mutase (23), strongly suggesting that the enzymatic rearrangement of chorismate to prephenate is a pericyclic process similar to the uncatalyzed reaction. In this case, the chemistry of the rearrangement is an inherent property of chorismate itself and chorismate mutase serves as a template for accelerating the reaction and may serve to stabilize the transition state by specific interactions—e.g., of the ether oxygen with Arg-90 or the carboxylate anions with the groups mentioned above. Detailed analysis of the active site structure of this transition state analog complex and the complexes of chorismate mutase with prephenate and with a substrate analog will provide more information on the origins of enzymatic rate acceleration. Furthermore, the active site structure will allow for rational design of better binding inhibitors.

Note Added in Proof. We have also solved the structure of the complex between chorismate mutase and the product, prephenate, which binds in 8 of the 12 active sites of the asymmetric unit. The carboxylate groups of prephenate are in the same positions as those of the transition state analog, and the carbonyl group of prephenate is bound to Arg-90.

FIG. 3. (a) Stereoview displaying the location of the active site in the trimer. The *endo*-oxabicyclic inhibitor binds in the cleft at the interface of adjacent subunits. Difference Fourier density map ($F_o - F_c$ map; F_o is structure factor of the complex of chorismate mutase with the *endo*-oxabicyclic inhibitor; F_c and phases are from the refined model of the unbound enzyme) contoured at 3σ is shown with C^α trace of the refined native enzyme. (b) Stereoview showing details of the active site. The *endo*-oxabicyclic inhibitor is shown with active site residues. Residues B57 and B75 are from one monomer; residues C7, C78, C90, C108, and C115 are from the adjacent monomer. Interactions between the inhibitor and the enzyme are as follows (see Fig. 1 for numbering of atoms on the inhibitor): C3_{inh} and C5_{inh}, side chain of Phe-57 (3.5–3.8 Å); C8_{inh}, Ala-59 (3.5 Å); O4_{inh}, amide N of Lys-60 (3.4 Å); C' _{inh}, C α of Thr-74 (3.5 Å); O1_{inh}, amide N of Cys-75 (3.1 Å); O1_{inh}, side chain of Cys-75 (3.3 Å); O1_{inh}, side chain of Glu-78 (2.9 Å); O2_{inh}, side chain of Arg-7 (2.8 Å); O2_{inh}, side chain of Arg-90 (3.1 Å); O' _{inh}, side chain of Arg-90 (3.1 Å); O3_{inh}, side chain of Tyr-108 (3.0 Å); C7_{inh} and C9_{inh}, side chain of Leu-115 (3.7–3.8 Å); O4_{inh}, side chain of Arg-116 (3.1–4.1 Å).

We thank J. R. Knowles for initiating this project, J. V. Gray for extensive and critical research on protein purification, P. A. Bartlett for the generous gift of the inhibitor, N. h. Xuong for the data collection facility, Pittsburgh Supercomputing Center Grant DMB 900069P for computation facility, M. K. Rosen and N. S. Sampson for assistance in selenomethionine protein preparation, R. C. Stevens for assistance in data collection, Y. P. Zhang and R. C. Stevens for assistance in computation, S. K. Burley and K. M. Reinisch for discussion, and M. K. Rosen for critical reading of the manuscript. This work is supported by National Institutes of Health Grant GM06920 to W.N.L.

- Weiss, U. & Edwards, J. M. (1980) *The Biosynthesis of Aromatic Amino Compounds* (Wiley, New York), pp. 134–184.
- Andrews, P. R., Smith, G. D. & Young, I. G. (1973) *Biochemistry* 18, 3492–3498.
- Copley, S. D. & Knowles, J. R. (1985) *J. Am. Chem. Soc.* 107, 5306–5308.
- Bartlett, P. A. & Johnson, C. R. (1985) *J. Am. Chem. Soc.* 107, 7792–7793.
- Sogo, S. G., Widlanski, T. S., Hoare, J. H., Grimshaw, C. E., Berchtold, G. A. & Knowles, J. R. (1984) *J. Am. Chem. Soc.* 106, 2701–2703.
- Jackson, D. Y., Jacobs, J. W., Sugawara, R., Reich, S. H., Bartlett, P. A. & Schultz, P. G. (1988) *J. Am. Chem. Soc.* 110, 4841–4842.
- Bowditch, K., Tang, Y., Hicks, J. B. & Hilvert, D. (1991) *J. Biol. Chem.* 266, 11901–11908.
- Hilvert, D., Carpenter, S. H., Nared, K. D. & Auditor, M.-T. M. (1988) *Proc. Natl. Acad. Sci. USA* 85, 4953–4955.
- Lorence, J. H. & Nester, E. W. (1967) *Biochemistry* 6, 1541–1553.
- Gray, J. V., Golinelli-Pimpaneau, B. & Knowles, J. R. (1990) *Biochemistry* 29, 376–383.
- Koch, G. L. E., Shaw, D. C. & Gibson, F. (1971) *Biochim. Biophys. Acta* 229, 795–804.
- Schmidheini, T., Mosch, H.-U., Evans, J. N. S. & Braus, G. (1990) *Biochemistry* 29, 3660–3668.
- Hamlin, R., Cork, C., Howard, A., Nielsen, C., Vernon, W., Matthews, D. & Xuong, N. h. (1981) *J. Appl. Crystallogr.* 14, 85–93.
- Terwilliger, T. & Eisenberg, D. (1983) *Acta Crystallogr. Sect. A: Found. Crystallogr.* 39, 813–817.
- Jones, T. A. (1985) *Methods Enzymol.* 115, 359–380.
- Jones, T. A. & Thirup, S. (1986) *EMBO J.* 5, 819–822.
- Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. (1990) *Nature (London)* 347, 37–44.
- Bricogne, G. (1976) *Acta Crystallogr. Sect. A: Found. Crystallogr.* 32, 832–847.
- Brunger, A. T. (1992) *X-PLOR version 3.0 Manual* (Yale University, New Haven, CT).
- Lindqvist, Y. (1989) *J. Mol. Biol.* 209, 151–166.
- Richardson, J. S. (1977) *Nature (London)* 268, 495–499.
- Rajagopalan, J. S., Taylor, K. M. & Jaffe, E. K. (1993) *Biochemistry* 32, 3965–3972.
- Gray, J. V., Eren, D. & Knowles, J. R. (1990) *Biochemistry* 29, 8872–8878.