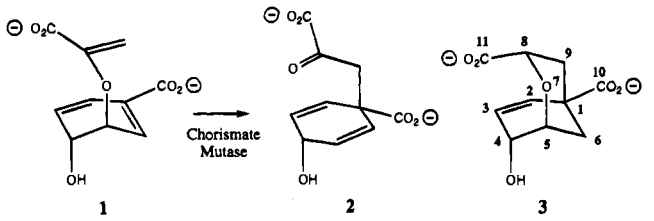
Atomic Structure of the Buried Catalytic Pocket of *Escherichia coli* Chorismate Mutase

Angela Y. Lee,† P. Andrew Karplus,‡ Bruce Ganem,† and Jon Clardy*,†

Departments of Chemistry and Biochemistry Cornell University, Ithaca, New York 14853

Received November 29, 1994

The first committed step in the biosynthesis of the aromatic amino acids phenylalanine and tyrosine is the chorismate mutase-catalyzed conversion of chorismate (**1**) to prephenate (**2**).[1] How the enzyme accelerates this pericyclic rearrangement



1 **2** **3**

by at least 10^6-fold is not well understood,[2,3] and analysis has been complicated by the existence of many different chorismate mutases with low sequence identity and unknown structural similarity. Recently reported structures of the monofunctional chorismate mutase from *Bacillus subtilis* (BsCM),[4,5] and the monofunctional allosteric chorismate mutase from yeast (*Saccharomyces cerevisiae*, ScCM),[6] presented important insights, but the extent of mechanistic and structural similarity among chorismate mutases remained unanswered.

The biosynthesis of phenylalanine in *E. coli* is initiated by the bifunctional P-protein, whose N-terminal 109 residues constitute a fully functional chorismate mutase domain.[7] The X-ray crystal structure at 2.2 Å resolution shows that the engineered chorismate mutase from *E. coli* (EcCM) forms a homodimer (Figure 1). Each monomer consists of a single polypeptide chain folded into three α-helices connected by two loops: α-helix 1 (H1), α-helix 2 (H2), and α-helix 3 (H3) (Figure 1A). The three α-helices, two long and one short, are arranged to give the monomer a novel fold resembling the figure "4". The two monomers are related by a noncrystallographic twofold axis where the helices of one monomer cross over those of the other monomer (Figure 1A). Dimer formation buries almost one-third (roughly 2600 Å^2) of the monomer's solvent-accessible surface area and gives an elongated dimer with approximate overall dimensions of 60 Å × 35 Å × 20 Å (Figure 1). The twofold molecular symmetry generates two antiparallel helix pair interactions between H1–H1′ and H3–H3′ (Figure 1A). The unusually long H1–H1′ pair forms an antiparallel coiledcoil[8,9] involving numerous interhelical leucine–leucine interactions with the heptad repeat Leu 10, Leu 17, Leu 24, Leu 31 in the δ-position as well as Leu 7 and Leu 21 in the α-position (Figure 1C). The H3–H3′ helices are essentially antiparallel. The H1–H1′ and H3–H3′ pairs pack together with a 45° crossing angle to form an antiparallel four helix bundle motif[10] that is quite hydrophobic and well packed. The dimer has two other regions where two pairs of helices interact in a parallel fashion: H1–H1′ with H2′–H3′ and the twofold related H1′–H1 with H2–H3 (Figure 1A). Inhibitor **3**, which is based on the putative transition state for the rearrangement,[11] binds near the middle of this relatively open helical bundle (Figure 1A,B). The fold of EcCM is novel, but the central region of the dimer is related to the fold of ScCM.[6] In the monomer of ScCM, helices H2, H7, H8, and H12 correspond to H1, H3, H1′, and H3′ of EcCM, respectively.

The clear electron density of inhibitor **3** identifies two equivalent active sites in the dimer, each with contributions from both monomers (Figure 2A). Overall there are 12 hydrogen bonds or electrostatic interactions between the enzyme and inhibitor **3**, and every oxygen atom of **3** is involved in two polar interactions (Figures 2A and 3). All but one interaction involves side chain, not main chain, atoms, and an extensive network of interactions holds the side chains in the appropriate conformation for binding. Key residues playing this supporting role include Gln 88, Arg 47, Asp 18′, Ser 15′, and Asp 48 (data not shown). The inhibitor is completely buried and has no solvent-accessible surface area in the complex (Figure 1B,C).

Despite their low amino acid sequence similarity—17% identity using standard alignment methods—EcCM and BsCM have been viewed similarly since they are roughly the same size (109 vs 127 residues, respectively) and have similar values of k_{cat}/k_{uncat} and K_i for inhibitor **3**. It is now clear that their secondary structures and folds are completely different. BsCM monomers form a five-stranded mixed β-sheet with an 18 residue α-helix and two 3_{10}-helices.[4,5] Three BsCM monomers form a symmetric trimer, with two monomers contributing to each active site.[4,5]

Given the disparate protein structures, the bound inhibitor has been used to align the active sites. Binding chorismate in its reactive pseudodiaxial conformation[12] is central to catalysis,[13,14] and both enzymes bind the reactive conformer **1** through multiple, albeit different, interactions (Figure 2). The two enzymes show the greatest similarities in binding the left side of **3** (C11 carboxylate and C4 hydroxyl) and the greatest differences in binding the right side (C10 carboxylate and O7). The C11 carboxylate interacts with Arg 11′ in EcCM and Arg 7 in BsCM, while the C4 hydroxyl interacts with Glu 52 in EcCM or Glu 78 in BsCM (Figures 2 and 3). These charged residues approach the inhibitor from different directions in the two structures (Figure 2). The Lys 39 residue that interacts with carboxylate C11 and the ether oxygen O7 in EcCM has an equivalent position and identical charge with the Arg 90 residue in BsCM. These residues, which share the greatest degree of structural similarity, create a highly charged active site. In EcCM, the tightly bound water bridging the two carboxylates and Arg 51 finds no simple analogy in the Tyr 108 residue of BsCM. In the EcCM binding pocket, the C10 carboxylate interacts strongly with Arg 28 and Ser 84, while the ether oxygen O7 interacts with the side chain of Gln 88 (Figures 2A and 3). There are no equivalent interactions in BsCM (Figure 2B). The hydrophobic interactions in the two structures are roughly equivalent with Val 85, Val 35, Leu 55,

† Department of Chemistry.
‡ Department of Biochemistry.

(1) Haslam, E. *Shikimic Acid Metabolism and Metabolites*; John Wiley and Sons: New York, 1993.
(2) Andrews, P. R.; Smith, G. D.; Young, I. G. *Biochemistry* **1973**, *18*, 3492–3498.
(3) Gorisch, H. *Biochemistry* **1978**, *17*, 3700–3705.
(4) Chook, Y. M.; Ke, H.; Lipscomb, W. N.; *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8600–8603.
(5) Chook, Y. M.; Gray, J. V.; Ke, H.; Lipscomb, W. N. *J. Mol. Biol.* **1994**, *240*, 476–500.
(6) Xue, Y.; Lipscomb, W. N.; Graf, R.; Schnappauf, G.; Braus, G. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10814–10818.
(7) Stewart, J.; Wilson, D. B.; Ganem, B. *J. Am. Chem. Soc.* **1990**, *112*, 4582–4584.
(8) Cusack, S.; Berthet-Colominas, C.; Härtlein, M.; Nassar, N.; Leberman, R. *Nature* **1990**, *347*, 249–254.
(9) Banner, D. W.; Kokkinidis, M.; Tsernoglou, D. *J. Mol. Biol.* **1987**, *196*, 657–675.
(10) Harris, N. L.; Presnell, S. R.; Cohen, F. E. *J. Mol. Biol.* **1994**, *236*, 1356–1368.
(11) Bartlett, P. A.; Johnson, C. R. *J. Am. Chem. Soc.* **1985**, *107*, 7792–7793.
(12) Guilford, W. J.; Copley, S. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1987**, *109*, 5013–5019.
(13) Turnbull, J.; Cleland, W. W.; Morrison, J. F. *Biochemistry* **1991**, *30*, 7777–7782.
(14) Haynes, M. R.; Stura, E. A.; Hilvert, D.; Wilson, I. A. *Science* **1994**, *263*, 646–652.

0002-7863/95/1517-3627$09.00/0

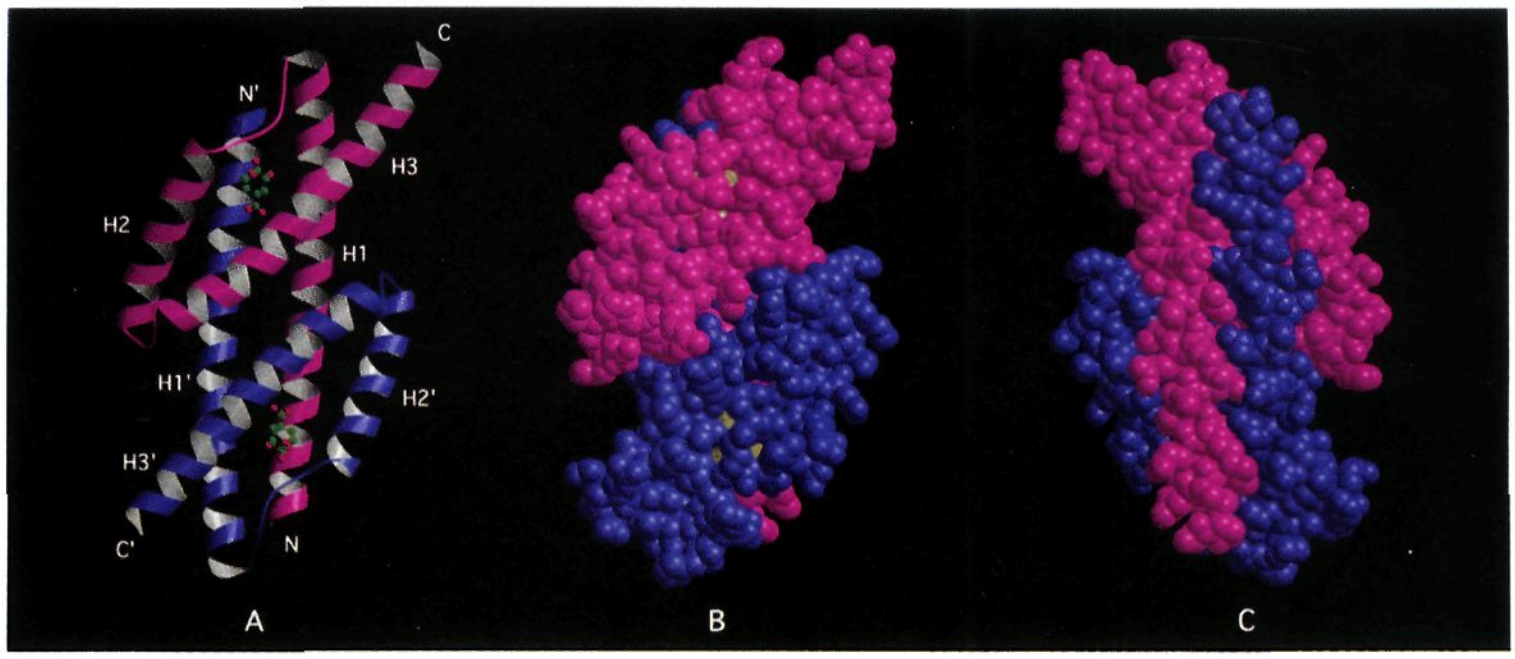


Figure 1. Dimeric chorismate mutase domain on *E. coli* P-protein[7] complexed with a transition state analog inhibitor[11] as defined by a 2.2 Å resolution X-ray diffraction analysis. (A) Ribbon diagram with the different monomer chains shown in purple and magenta, and the inhibitor shown as a ball and stick model. (B) The same view as A but with a space-filling rendering. Note the barely visible inhibitor (**3**) shown in yellow. (C) The view in B rotated by 180° around a horizontal axis to show the back. Note the "knobs in holes" interaction of the coiled coil of H1−H1′. From this side the inhibitor is not visible. The inhibitor is buried by a single layer of side chains on each side of the roughly 20 Å thick enzyme so that access to the active site may be possible from either side.

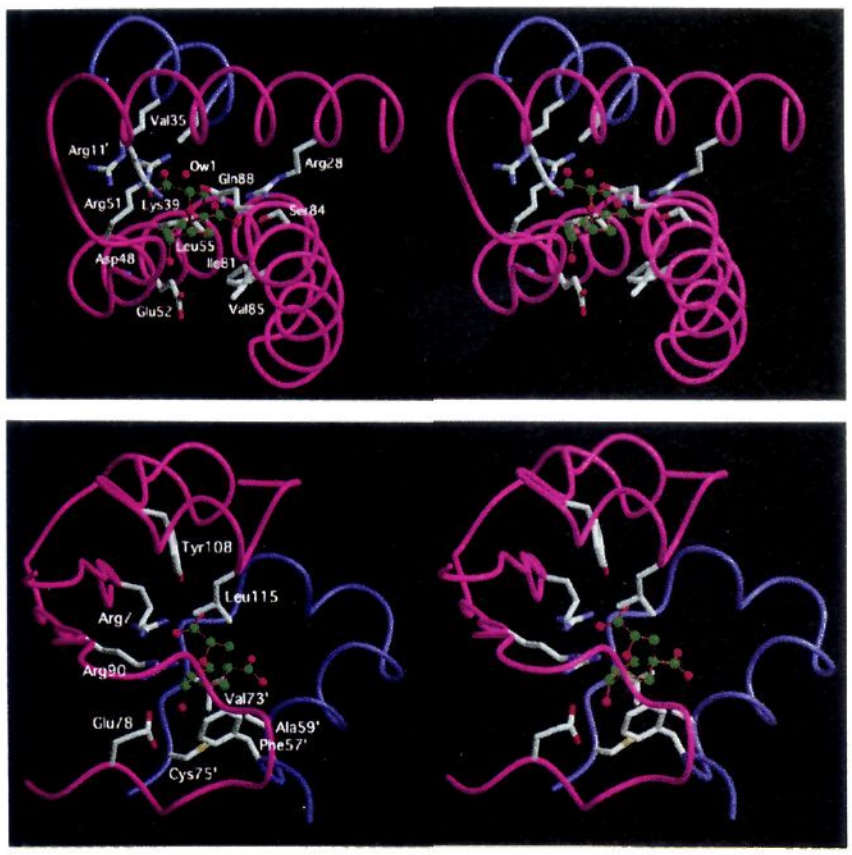


Figure 2. (A, top) Stereoview of the active site of EcCM showing **3** and important side chains. Residues Arg 28, Val 35, Lys 39, Asp 48, Arg 51, Glu 52, Leu 55, Ile 81, Ser 84, Val 85, and Gln 88 from one monomer and Arg 11′ from the other contribute to the active site. (B, bottom) Stereoview of the active site of BsCM, showing inhibitor **3** and important side chains.[4]

and Ile 81 of EcCM, corresponding to Phe 57, Leu 115, Val 73, and Ala 59 of BsCM (Figure 2). A major structural difference is a completely enclosed inhibitor **3** in EcCM and a partially exposed **3** in BsCM. As a consequence, 14 EcCM residues are within 4 Å of an atom of **3**, while 11 BsCM residues make similar contacts.

In addition to binding the reactive conformation, an efficient catalyst might also stabilize the polar pericyclic transition state,[5,14,15] and EcCM and BsCM show interesting differences in this regard. Stabilizing an incipient negative charge on the ether oxygen O7 may be important to catalysis.[13,16] In both EcCM and BsCM, a positively charged side chain interacts with the ether oxygen from one side, but in EcCM there is an additional interaction in the hydrogen bond from Gln 88 on the other side (Figure 3), so that both lone pairs of the ether oxygen are hydrogen bonded to active site residues. By controlling the C1−C10 torsional angle using Arg28 in EcCM, the enzyme

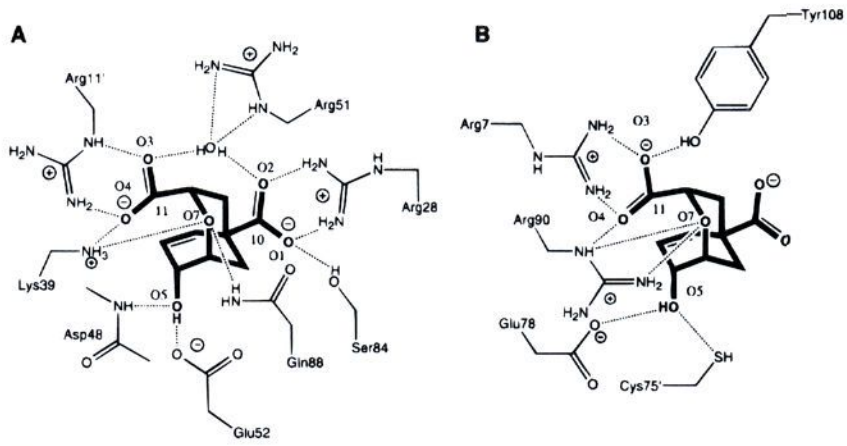


Figure 3. Schematic diagram comparing the hydrogen-bonding and electrostatic interactions of transition state inhibitor **3** with relevant side chains of EcCM (A) and BsCM (B).

may also stabilize the cyclohexadienyl cation fragment and thus enhance the rearrangement rate. There is no equivalent interaction in BsCM.

Activation parameters for the catalyzed and uncatalyzed reactions suggest that chorismate mutases exert their effect through a combination of conformational control and enthalpic lowering. That two quite different folds display mutase activity suggests that the number of catalytic motifs for this pericyclic process may not be exhausted.

Acknowledgment. This work was supported by NIH CA24487 (J.C.), a Molecular Biophysics Training Grant GM08267 (A.Y.L.), and GM24054 (B.G.). We are indebted to W. N. Lipscomb and Y. M. Chook for providing coordinates of PDB entry 1CHT prior to release and for helpful discussions and to P. A. Bartlett for the gift of inhibitor **3**. We also thank G. Schulte, D. Irwin, L. Brinen, and J. Stewart.

Supplementary Material Available: Experimental procedures for the crystallization, data collection, analysis, and refinement of the EcCM-**3** complex (7 pages). This material is contained in many libraries on microfiche, immediately follows this article in the microfilm version of the journal, can be ordered from the ACS, and can be downloaded from the Internet; see any current masthead page for ordering information and Internet access instructions. Coordinates for EcCM-**3** have been deposited with the Brookhaven Protein Data Bank.

JA943859N

(15) Gray, J. V.; Golinelli-Pimpaneau, B.; Knowles, J. R. *Biochemistry* **1990**, *29*, 376−383.
(16) Severance, D. L.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1992**, *114*, 10966−10968.