Appendix B

Computationally designed variants of *Escherichia coli* chorismate mutase show altered catalytic activity

The text of this chapter has been adapted from a manuscript coauthored with Jonathan Kyle Lassila, Peter Oelschlaeger, and Stephen L. Mayo.

Lassila, J.K., Keeffe, J.R., Oelschlaeger, P., and Mayo, S.L. (2005) Computationally designed variants of *Escherichia coli* chorismate mutase show altered catalytic activity *Protein Engineering, Design, and Selection 18*, 161-163.

Abstract

Computational protein design methods were used to predict five variants of monofunctional *E. coli* chorismate mutase expected to maintain catalytic activity. The variants were tested experimentally and three active site mutations exhibited catalytic activity similar to or greater than the wild type enzyme. One mutation, Ala32Ser, showed increased catalytic efficiency.

Introduction

The Claisen rearrangement of chorismate to prephenate (Figure B-1) is a rare enzyme-catalyzed pericyclic reaction that proceeds through the same mechanism uncatalyzed in solution. Chorismate mutases from various organisms provide rate enhancements of around 10^6 despite strong dissimilarities in three-dimensional fold.[1-3] The metabolic importance of chorismate as the key branch point in the shikimate pathway has prompted extensive experimental investigation of the chorismate-prephenate rearrangement since the 1960s[4] and has driven complementation experiments to probe the structural determinants of enzyme catalysis.[5] The concerted, unimolecular nature of the rearrangement and the lack of covalent protein contacts have encouraged numerous theoretical studies of the catalyzed and uncatalyzed reactions. The question of how chorismate mutases achieve rate enhancement has been actively discussed in recent years.[6-10]

We used computational protein design techniques to identify mutations within the active site of the chorismate mutase domain (EcCM) of *Escherichia coli* chorismate mutase-prephenate dehydratase ("P-protein")[11] consistent with catalytic activity. The objective of the study was to evaluate the viability of a rotamer-based approach with an empirical mechanics force field in modeling the active site. We report three active site mutations that permit catalytic activity at or above the level of the wild type enzyme. One of these appears to have enhanced catalytic efficiency relative to wild type.

Materials and Methods

Computational design. An *ab initio* calculated transition state structure[12] was modeled at the position of a transition state analog in the EcCM crystal structure, PDB code 1ECM.[3] Translation (± 0.2 Å each axis) and rotation (± 5° each axis) of the transition state structure was allowed and all amino acids except glycine, proline, cysteine, and methionine were permitted in the following positions: 28, 32, 35, 39, 46, 47, 48, 51, 52, 55, 81, 84, 85, 88 (chain A); 7, 11, 14, 18 (chain B). Residues from both chains compose each of the two active sites in this symmetric homodimer. A backbone-dependent side chain rotamer library[13] was used with expansion by one standard deviation about $\chi 1$ and $\chi 2$ for all amino acids except arginine and lysine. The energy function used in the ORBIT protein design software[14] was based on the DREIDING force field[15] and includes a scaled van der Waals term,[16] hydrogen bonding and electrostatic terms,[17] and a surface-area based solvation potential.[18] Transition state partial atomic charges were obtained as previously reported.[12] An additional energy penalty was applied to effectively eliminate from consideration all sequences that could not maintain key contacts between the transition state structure and Arg 11 and Arg 28. These contacts were previously demonstrated to be necessary for catalysis.[19] The HERO rotamer optimization method was used to obtain the minimum energy amino acid sequence and conformations.[20] Six mutations were predicted in the optimized structure: Leu7Ile, Ala32Ser, Val35Ile, Asp48Ile, Ile81Leu, and Val85Ile. Visual inspection and subsequent calculations indicated that four of the mutations were independent, so they were treated separately.

Protein expression and kinetic characterization. The gene encoding EcCM residues 1-109 of the bifunctional chorismate mutase-prephenate dehydratase (P protein) was amplified from genomic DNA (ATCC 700926D) and inserted into the pTYB11 vector from the IMPACT-CN intein fusion system (New England Biolabs) between the *Sap*I and *Xho*I sites. Inverse PCR mutagenesis[21] was used to construct five variants: Leu7Ile, Ala32Ser, Val35Ile, Asp48Ile, and Ile81Leu/Val85Ile. Constructs were verified by DNA sequencing. Variant and wild type proteins were expressed in *E. coli* BL21(DE3). Expression was induced with 1 mM IPTG at OD_{600} = 0.6 and cells were grown for 18 hours at 22 °C. Cells were harvested and lysed by French press in 20 mM Tris-Cl pH 8.0, 500 mM NaCl, 1 mM EDTA, 10 mM $MgCl_2$, 1 mM PMSF, with DNase and RNase. Chitin affinity purification was conducted with a column buffer containing 20 mM Tris-Cl, pH 8.0, 500 mM NaCl, and 1 mM EDTA. On-column cleavage proceeded in 50 mM DTT at 25 °C for 18 hours. Mass spectrometry and SDS-PAGE verified that this protocol results in successful cleavage and yields the expected product. Wild type and variant proteins were further purified by gel filtration on a HiPrep Sephacryl S-100 high-resolution column (Amersham Biosciences) with 20 mM Tris-Cl pH 7.8, 100 mM NaCl as the running buffer and final storage buffer. Protein characterization followed procedures recently reported for the same construct.[22] Protein concentration was determined by Bradford assay using BSA as a standard. Chorismate mutase activity was determined by following the disappearance of chorismate with UV absorbance at 275 nm. Activity assays were conducted at 37 °C and contained 20 nM protein in 50 mM Tris pH 7.8, 2.5 mM EDTA, 20 mM β-mercaptoethanol, and 0.01 % BSA. Initial velocities were buffer-corrected and were determined with less than 6 %

depletion of initial substrate concentration. All proteins were initially tested using a substrate concentration range of about 50-500 μM. The wild type and the Ala32Ser mutant were further assayed with substrate ranges of approximately 20-2000 μM and a minimum of five trials including two separate protein preparations each. Kinetic parameters were determined by non-linear fitting to the Michaelis-Menten equation.

Molecular dynamics simulations. Molecular dynamics simulations of EcCM and mutant Asp48Ile as free enzymes and in complex with the oxabicyclic transition state analog[23] were carried out using the *Sander* Program of the *AMBER 7.0* software package.[24] For symmetry, both chains were truncated to residues 6 through 95. The transition state analog was minimized and the electrostatic potential surface was calculated using DFT with B3LYP and the lacvp** basis set in the Jaguar 5.5 package (Schrödinger LLC). Atomic charges were obtained with *RESP*. Parameters for the transition state analog were defined according to similar structures in the *AMBER* libraries. The binding sites of the free enzymes were filled with water by solvating the molecules with a TIP3P[25] water shell using a closeness parameter of 0.4 in *xLEaP*. All systems were then solvated by a truncated octahedron of TIP3P water 8 Å around the protein and neutralized with counter ions. The ff02EP version of the ff99 force field[26,27] was used to represent the protein and the ligand. After minimization, all systems were heated to 300 K with constant volume, starting at 10 K with three different initial velocities. Molecular dynamics simulations were prolonged at 300 K and constant pressure until the backbone RMSD reached a constant value of about 1.5 Å compared to

the first frame; this was observed after 500 to 800 ps. Average structures over 10 ps of the equilibrated systems were generated and analyzed visually.

Results and Discussion

As seen in Table B-1, three of the five variants showed similar or greater catalytic efficiency relative to the wild type enzyme. The Ala32Ser mutation results in a slightly more efficient catalyst than wild type due to both a decrease in K_M and an increase in k_{cat}. The fact that substrate binding is enhanced in addition to catalysis is consistent with the suggestion that factors stabilizing the transition state may be likely to contribute to ground state stabilization in the catalyzed rearrangement.[28] The rate enhancement corresponds to a change in activation energy of less than 1 kcal/mol, making a detailed structural explanation unwarranted. However, it should be noted that in the predicted structure of this mutant, Ser 32 is capable of hydrogen bonding with Gln 88 (Figure B-2), a residue that makes an essential contact to the ether oxygen of the breaking bond.[19]

Val35Ile shows increased k_{cat} but also increased K_M, resulting in k_{cat}/K_M similar to wild type. The Leu7Ile mutation also does not have a significant effect on catalytic efficiency. Both mutations create slightly different hydrophobic packing environments near the essential residue Arg 11 in the predicted structures. The double mutant, Ile81Leu/Val85Ile, was predicted to alter packing against the hydrophobic ring face of the reacting molecule. This rearrangement of hydrophobic amino acids does not result in substantial change in k_{cat}. However, the mutations result in increased K_M and reduced catalytic efficiency. The relative insensitivity of the enzyme to changes in some amino

acids is not surprising given the recent finding that the reaction is efficiently catalyzed by a protein exhibiting all the characteristics of a molten globule.[29]

The Asp48Ile mutation abolished measurable catalysis under the conditions tested. This position makes backbone contacts to the hydroxyl group of the transition state analog in the EcCM crystal structure. Hydroxyl contacts to a negatively charged residue were suggested to create a favorable electrostatic gradient in the *Bacillus subtilis* enzyme.[30] However, in the *E. coli* structure the Asp 48 side chain points away from the active site and is distant from the transition state analog. Molecular dynamics simulations provide some insight into a possible function of Asp 48 that this mutant lacks. Averaged structures from equilibrated systems show Asp 48 hydrogen bonding to Arg 11 in both the unliganded and inhibitor-bound enzyme. In simulations of unliganded Asp48Ile mutants, however, Arg 11 is dramatically displaced into solution, suggesting that one role of Asp 48 may be to stabilize this key active site side chain in a conformation compatible with substrate binding and catalysis.

While the choice of mutations in this experiment was based solely on computational modeling and no sequence alignment information was used in the process, a BLAST search[31] using the EcCM sequence as the query showed that sequence variations corresponding to the Ala32Ser, Val35Ile, Val85Ile, and Ile81Leu mutations were observed in chorismate mutases from related organisms.

Our design procedure stabilizes a static active site configuration with a bound transition state structure. Although the substrate of the reaction is not explicitly considered, we expect that modeling interactions using the structure and charges of the transition state should promote some degree of differential stabilization of the transition

state relative to substrate. The present study demonstrates that this approach can be used effectively to represent the active site of a natural enzyme. In this case, the predicted mutations were in residues not directly contacting the reacting molecule. The favorable result from the Ala32Ser mutation suggests that such secondary contacts are important in the enzyme design process. The complete loss of catalytic activity from the Asp48Ile mutation implies that improved treatment of electrostatics and consideration of the unbound enzyme could offer some benefit in future design efforts.

References

1. Chook, Y.M., Hengming, K. and Lipscomb, W.N. (1993) Crystal-structures of the monofunctional chorismate mutase from *Bacillus-subtilis* and its complex with a transition-state analog. *Proc. Natl Acad. Sci. USA,* **90**, 8600-8603.

2. Xue, Y., Lipscomb, W.N., Graf, R., Schnappauf, G. and Braus, G. (1994) The crystal-structure of allosteric chorismate mutase at 2.2-Ångstrom resolution. *Proc. Natl Acad. Sci. USA*, **91**, 10814-10818.

3. Lee, A.Y., Karplus, P.A., Ganem, B. and Clardy, J. (1995) Atomic-structure of the buried catalytic pocket of *Escherichia coli* chorismate mutase. *J. Am. Chem. Soc.,* **117**, 3627-3628.

4. Gibson, M.I. and Gibson, F. (1964) Preliminary studies on isolation + metabolism of intermediate in aromatic biosynthesis - chorismic acid. *Biochem. J*., **90**, 248-256.

5. Woycechowsky, K.J. and Hilvert, D. (2004) Deciphering enzymes – Genetic selection as a probe of structure and mechanism. *Eur. J. Biochem*., **271**, 1630-1637.

6. Guimarães, C.R.W., Repasky, M.P., Chandrasekhar, J., Tirado-Rives, J. and Jorgensen, W.L. (2003) Contributions of conformational compression and preferential transition state stabilization to the rate enhancement by chorismate mutase. *J. Am. Chem. Soc.*, **125**, 6892-6899.

7. Strajbl, M., Shurki, A., Kato, M. and Warshel, A. (2003) Apparent NAC effect in chorismate mutase reflects electrostatic transition state stabilization. *J. Am. Chem. Soc*., **125**, 10228-10237.

8. Hur, S. and Bruice, T.C. (2003) The near attack conformation approach to the study of the chorismate to prephenate reaction. *Proc. Natl Acad. Sci. USA*, **100**, 12015-12020.

9. Crespo, A., Scherlis, D.A., Martí, M.A., Ordejón, P., Roitberg, A.E. and Estrin, D.A. (2003) A DFT-based QM-MM approach designed for the treatment of large molecular systems: Application to chorismate mutase. *J. Phys. Chem. B*, **107**, 13728-13736.

10. Ranaghan, K.E., Ridder, L., Szefczyk, B., Sokalski, W.A., Hermann, J.C. and Mulholland, A.J. (2004) Transition state stabilization and substrate strain in enzyme catalysis: *Ab initio* QM/MM modeling of the chorismate mutase reaction. *Org. Biomol. Chem*., **2**, 968-980.

11. Stewart, J., Wilson, D.B. and Ganem, B. (1990) A genetically engineered monofunctional chorismate mutase. *J. Am. Chem. Soc*., **112**, 4582-4584.

12. Wiest, O. and Houk, K.N. (1994) On the transition-state of the chorismate-prephenate rearrangement. *J. Org. Chem*., **59**, 7582-7584.

13. Dunbrack, R.L. and Cohen, F.E. (1997) Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein Sci*., **6**, 1661-1681.

14. Dahiyat, B.I. and Mayo, S.L. (1997) De novo protein design: Fully automated sequence selection. *Science,* **278**, 82-87.

15. Mayo, S.L., Olafson, B.D. and Goddard, W.A. (1990) DREIDING – A generic force-field for molecular simulations. *J. Phys. Chem*., **94**, 8897-8909.

16. Dahiyat, B.I. and Mayo, S.L. (1997) Probing the role of packing specificity in protein design. *Proc. Natl Acad. Sci. USA*, **94**, 10172-10177.

17. Dahiyat, B.I., Gordon, D.B. and Mayo, S.L. (1997) Automated design of the surface positions of protein helices. *Protein Sci*., **6**, 1333-1337.

18. Street, A.G. and Mayo, S.L. (1998) Pairwise calculation of protein solvent-accessible surface areas. *Fold. Des*., **3**, 253-258.

19. Liu, D.R., Cload, S.T., Pastor, R.M. and Schultz, P.G. (1996) Analysis of active site residues in Escherichia coli chorismate mutase by site-directed mutagenesis. *J. Am. Chem. Soc*., **118**, 1789-1790.

20. Gordon, D.B., Hom, G.K., Mayo, S.L. and Pierce, N.A. (2002) Exact rotamer optimization for protein design. *J. Comput. Chem*., **24**, 232-243.

21. Hemsley, A., Arnheim, N., Toney, M.D., Cortopassi, G. and Galas, D.J. (1989) A simple method for site-directed mutagenesis using the polymerase chain-reaction. *Nucleic Acids Res.,* **17**, 6545-6551.

22. Zhang, S., Wilson, D.B. and Ganem, B. (2003) An engineered chorismate mutase with allosteric regulation. *Bioorg. Med. Chem.*, **11**, 3109-3114.

23. Bartlett, P.A., Johnson, C.R. (1985) An inhibitor of chorismate mutase resembling the transition-state conformation. *J. Am. Chem. Soc.*, **107***,* 7792-7793.

24. Case, C.A., Pearlman, D.A., Caldwell, J.W., Cheatham, T.E., Wang, J., Ross, W.S., Simmerling, C.L., Darden, T.A., Merz, K. M. Jr., Stanton, R.V., Cheng, A.L., Vincent, J.J.; Crowley, M., Tsui, V., Gohlke, H., Radmer, R.J., Duan, J., Pitera, J., Massova, I., Seibel, G.L., Singh, U.C., Weiner, P.K., Kollman, P.A. AMBER 7, University of California, San Francisco, 2002.

25. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*., **79**, 926-935.

26. Wang, J., Cieplak, P., Kollman, P.A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.*, **21***,* 1049-1074.

27. Cornell, W.D., Cieplak, J.W., Bayly, C.I., Gould, I.R., Merz, K.M. Jr., Ferguson, D. M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., Kollman, P.A. (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.,* **117***,* 5179-5197.

28. Strajbl, M., Shurki, A., Kato, M. and Warshel, A. (2003) Apparent NAC effect in chorismate mutase reflects electrostatic transition state stabilization. *J. Am. Chem. Soc*., **125**, 10228-10237.

29. Vamvaca, K., Vögeli, B., Kast, P., Pervushin, K. and Hilvert, D. (2004) An enzymatic molten globule: Efficient coupling of folding and catalysis. *Proc. Natl Acad. Sci. USA*, **101**, 12860-12864.

30. Kast, P., Hartgerink, J.D., Asif-Ullah, M. and Hilvert, D. (1996) Electrostatic catalysis of the Claisen rearrangement: probing the role of Glu78 in *Bacillus subtilis* chorismate mutase by genetic selection. *J. Am. Chem. Soc*., **118**, 3069-3070.

31. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*., **25**, 3389-3402.

Table B-1. Kinetic parameters of wild type and mutant EcCM[a]

	k_{cat}	K_M	k_{cat}/K_M	%WT
	min^{-1}	µM	$min^{-1}µM^{-1}$	k_{cat}/K_M
WT	2332 ± 306	304 ± 52	7.8 ± 1.0	100
Ala32Ser	2708 ± 364	220 ± 29	12.4 ± 1.0	159
Val35Ile[b]	3046 ± 172	365 ± 59	8.5 ± 1.1	109
Leu7Ile[b]	2193 ± 291	249 ± 54	9.1 ± 2.4	117
Ile81Leu/Val85Ile[b]	2004 ± 241	669 ± 165	3.1 ± 0.6	40
Asp48Ile[c]	-	-	-	-

[a] Reported errors are standard deviations from at least three trials.
[b] Three mutants were assayed with a limited substrate concentration range, see text.
[c] Reaction rates with the Asp48Ile mutant were within error of the uncatalyzed solution reaction.

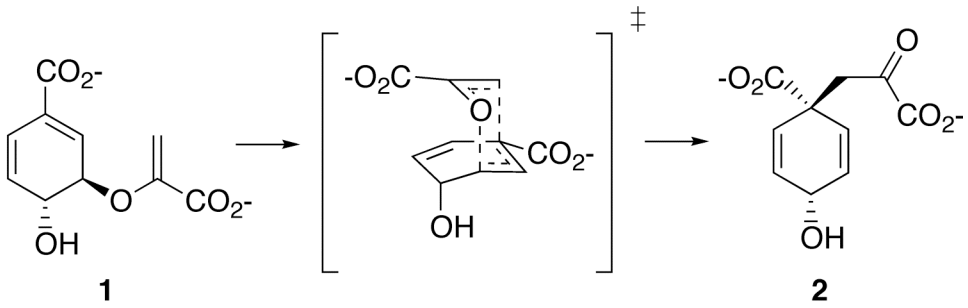

Figure B-1. Claisen rearrangement of chorismate (1) to prephenate (2).

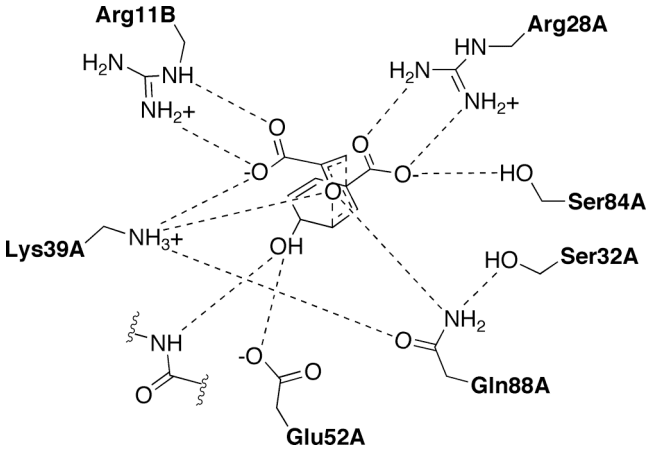

Figure B-2. Predicted hydrogen bonding in the Ala32Ser chorismate mutase variant. Interactions in the wild type crystal structure are the same except for position 32A.