

The proficiency of a thermophilic chorismate mutase enzyme is solely through an entropic advantage in the enzyme reaction

Xiaohua Zhang and Thomas C. Bruice[†]

Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA 93106

Contributed by Thomas C. Bruice, October 21, 2005

A study of the *Thermus thermophilus* chorismate mutase (TtCM) is described by using quantum mechanics (self-consistent-charge density-functional tight binding)/molecular mechanics, umbrella sampling, and the weighted histogram analysis method. The computed free energies of activation for the reactions in water and TtCM are comparable to the experimental values. The free energies for formation of near attack conformer have been determined to be 8.06 and 0.05 kcal/mol in water and TtCM, respectively. The near attack conformer stabilization contributes $\approx 90\%$ to the proficiency of the enzymatic reaction compared with the reaction in water. The transition state (TS) structures and partial atom charges are much the same in the enzymatic and water reactions. The difference in the electrostatic interactions of Arg-89 with O13 in the enzyme-substrate complex and enzyme-TS complex provides the latter with but 0.55 kcal/mol of 1.92 kcal/mol total TS stabilization. Differences in electrostatic interactions between components at the active site in the enzyme-substrate complex and enzyme-TS complex are barely significant, such that TS stabilization is of minor importance and the enzymatic catalysis is through an entropic advantage.

potential of mean force | self-consistent-charge density-functional tight binding | quantum mechanics/molecular mechanics

The contributions of ground-state (GS) conformations and transition-state (TS) stabilization to the proficiency of an enzymatic reaction can be more easily understood by studying a simple one-substrate enzyme reaction that involves the intramolecular rearrangement of a substrate to product without formation of a covalent intermediate (1, 2). The Claisen rearrangement of chorismate to prephenate is a well studied example (3) as shown in Fig. 1 where the reactive conformation near attack conformer (NAC) is defined.

Our choice to research the *Thermus thermophilus* chorismate mutase (TtCM) relates to our current interest in the chorismate \rightarrow prephenate reaction (4–6) and the questions of the temperature specificity of thermophilic enzymes (7). TtCM belongs to a rare group of the AroH-type monofunctional CMs that are principally found in Gram-positive bacteria and have been discovered as a thermophilic enzyme in the Gram-negative *T. thermophilus* (8, 9). The experimental free energies of activation $\Delta G_{\text{exp}}^{\ddagger}$ are 15.2 and 24.2 kcal/mol for the reactions in TtCM (343 K, the bacterium's optimal growth temperature) and water (298 K) (10), respectively. In this article, we describe a quantum mechanics (QM)/molecular mechanics (MM) (11) and umbrella sampling (12) study of the reactions at 343 K for TtCM and 298 K for water.

Methods

The initial coordinates of TtCM were modified from the x-ray structure of the F55S mutant of the TtCM trimer (Protein Data Bank ID code 1ODE; resolution, 1.65 Å) (8). The QM/MM simulations were carried out with CHARMM version 31b1 by the use of the self-consistent-charge density functional tight binding (SCCDFTB) module (13–15). For the simulations in the active

site of TtCM, the chorismate was treated as the QM region. The enzyme-substrate complex (E-S) was solvated with a water cap of 25 Å radius centered at the QM region and consisted of 1,542 TIP3P water molecules (16). Stochastic boundary conditions were applied to the system (17). Umbrella sampling (18) was used to sample the conformational ensembles of the reactions. NAC definition of C5-C16 distance was chosen as the reaction coordinate in umbrella sampling. The same procedure was used for the simulations of the chorismate in water. The free energies of activation for the reactions in water [$\Delta G_{\text{calc}}^{\ddagger(w)}$] and TtCM [$\Delta G_{\text{calc}}^{\ddagger(E)}$] were determined by the potential of mean force (PMF). The free energies for NAC formation in water and enzyme were calculated from the mole fraction of the chorismate in the GS present as NAC. Detailed methods of initial coordinate preparation, QM/MM simulation setup, umbrella sampling, and calculation of free energy for NAC formation are available in *Supporting Text*, which is published as supporting information on the PNAS web site.

Results

PMF Profiles Obtained from Umbrella Sampling. The free-energy profiles for the chorismate \rightarrow prephenate reaction along the C5-C16 reaction coordinate in both water and the active site of TtCM, as obtained by SCCDFTB/MM and umbrella sampling, are shown in Fig. 2. The PMF profile for the reaction in water is plotted as a blue solid line, and the free-energy profile of NAC formation is plotted in a red dashed line. The activation barriers (TS) are present at reaction coordinate $R_{\text{C5-C16}}$ of 2.17 and 2.37 Å in water and TtCM, respectively. The minima of the free-energy profile of NAC formation in water and TtCM are present at $R_{\text{C5-C16}}$ of 3.03 and 3.09 Å. The most favorable distance of C5-C16 for NAC formation in enzyme is comparable to that in water. Thus, the substrate forms similar NAC species regardless of the environments.

The free energies of activation in water and TtCM ($\Delta G_{\text{calc}}^{\ddagger}$) are 22.31 and 12.38 kcal/mol, which are in reasonable agreement with the experimental values of 24.5 and 15.4 kcal/mol, respectively (Table 1). The free energy of activation in the enzymatic reaction is 9.93 kcal/mol lower than the free energy of activation in water. The minimum of the free-energy profile of NAC formation in water [$G_{\text{NAC}}^{\text{min}}(\xi')$] is 8.06 kcal/mol, which is in good agreement with previous results (8.1 kcal/mol) obtained from thermodynamic integration (4). In the enzymatic reaction, free energy for NAC formation is 0.73 kcal/mol when $R_{\text{C5-C16}}$ is the distance of the NAC criteria (3.7 Å). The minimum of the free-energy profile of NAC formation in

Conflict of interest statement: No conflicts declared.

Abbreviations: CM, chorismate mutase; TtCM, *Thermus thermophilus* CM; EcCM, *Escherichia coli* CM; BsCM, *Bacillus subtilis* CM; NAC, near attack conformer; E-S, enzyme-substrate complex; E-TS, enzyme-TS complex; TS, transition state; GS, ground state; QM, quantum mechanics; MM, molecular mechanics; SCCDFTB, self-consistent-charge density-functional tight binding; PMF, potential of mean force.

[†]To whom correspondence should be addressed. E-mail: tcbuice@chem.ucsb.edu.

© 2005 by The National Academy of Sciences of the USA

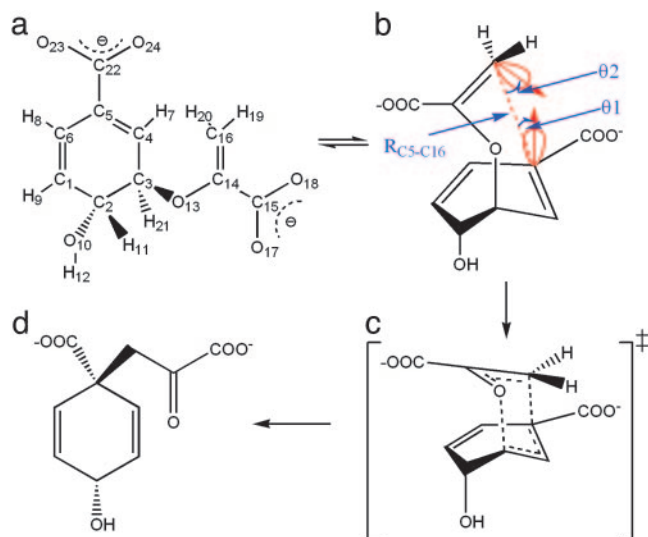


Fig. 1. Diagram of the chorismate \rightarrow prephenate reaction. (a) Chorismate. (b) NAC of the chorismate is defined by three criteria: R_{C5-C16} , $\theta 1$, and $\theta 2$. $\theta 1$ is the angle of the C16 approach to the C5 π -orbital. $\theta 2$ is the C16 π -orbital direction relative to the C5 atom. In the TS, $\theta 1 = 8.2^\circ$ and $\theta 2 = 18.6^\circ$. For the NACs, the bonding angles $\theta 1$ and $\theta 2$ are allowed $\pm 20^\circ$ deviations from the TS structure. The distance $R_{C5-C16} \leq 3.7 \text{ \AA}$. (c) TS. (d) Prephenate.

TtCM is 0.05 kcal/mol. Therefore, the GS conformations of chorismate in the active site of TtCM are essentially NACs. Using the experimental value of the free energy of reaction in

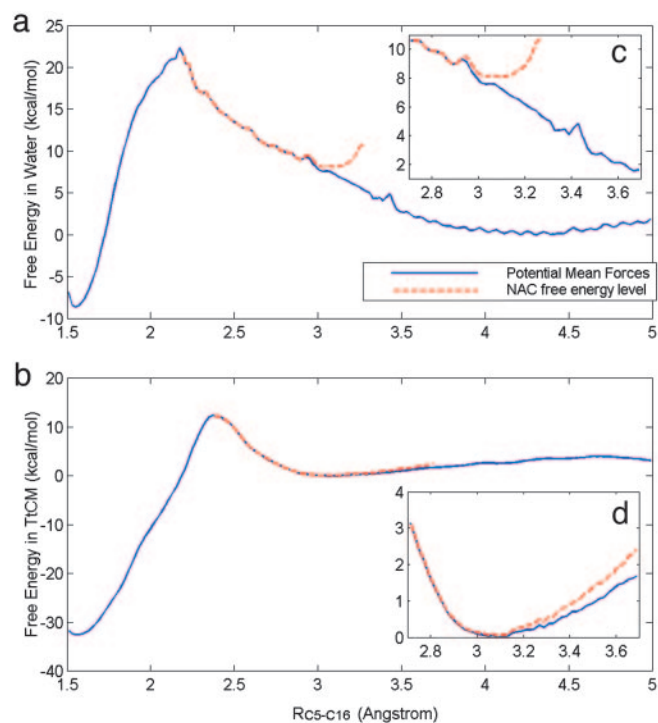


Fig. 2. Free energy profiles. (a and b) Plots of the free energy profiles of the chorismate \rightarrow prephenate reactions in water (a) and TtCM (b). The blue solid lines are PMF profiles and red dashed lines are the free energy profiles of NAC formation. (c and d) Within the region of $2.7 \leq R_{C5-C16} \leq 3.7 \text{ \AA}$, the free energy profiles were plotted in high resolution. Both PMF and NAC free energy profiles are offset by the values of free energy minima for the chorismate in GSs in PMF.

Table 1. The free energies for the chorismate \rightarrow prephenate reaction in both water and TtCM (kcal/mol)

Free energies	Water	TtCM
$G_g^{\min}(\xi'')$	0.00	0.00
$G_{NAC}^{\min}(\xi')$	8.06	0.05
$\Delta G_{NAC}^o = G_{NAC}^{\min}(\xi') - G_g^{\min}(\xi'')$	8.06	0.05
$\Delta G_{calc}^\ddagger = G^\ddagger(\xi) - G_g^{\min}(\xi'')$	22.31	12.38
ΔG_{exp}^\ddagger	24.5	15.4
$\Delta \Delta G_{NAC}^o = \Delta G_{NAC}^{o(w)} - \Delta G_{NAC}^{o(E)}$		8.01
$\Delta \Delta G_{calc}^\ddagger - \Delta \Delta G_{calc}^\ddagger(w) - \Delta G_{calc}^\ddagger(E)$		9.93
$\Delta \Delta G_{exp}^\ddagger$		9.1
$\Delta \Delta G_{TS}^\ddagger$		1.92
$\Delta \Delta G_{NAC}^o / \Delta \Delta G_{calc}^\ddagger \times 100 \%$		80.6%
$\Delta \Delta G_{NAC}^o / \Delta \Delta G_{exp}^\ddagger \times 100 \%$		88.0%

The superscripts min, o, and \ddagger designate the minimum of the free energy profile, standard free energy, and free energy of the activation for the reaction, respectively. The superscripts (E) and (W) stand for the free energy change in enzyme and water, respectively. The subscripts calc and exp represent the computed and experimental values. ξ , ξ' , and ξ'' denote the positions of TS, NAC, and GS along the reaction coordinates.

enzyme (9.1 kcal/mol), the kinetic advantage of the enzymatic reaction, compared with the reaction in water, is caused 88% (8.01/9.1) by favorable NAC formation. Strajbl *et al.* (19), in an earlier empirical valence bond study, confirmed the importance of NAC formation in determining the proficiency of the *Bacillus subtilis* CM (BsCM) enzymatic reaction.

Reaction Coordinates. Because chorismate NAC is defined by three criteria, the distance R_{C5-C6} , and the angles $\theta 1$ and $\theta 2$, 1D PMF profiles were projected into 2D PMF profiles by the weighted histogram analysis method (20, 21) along the R_{C5-C6} and $\theta 1$, and R_{C5-C6} and $\theta 2$, respectively (Fig. 3). The TS is identified from the highest saddle points in the contour plots along the reaction coordinates, labeled with blue asterisks. For the reaction in water (Fig. 3 a and b), $\theta 1$ is 14° and $\theta 2$ is 19° in the TS. For the reaction in TtCM (Fig. 3 c and d), $\theta 1$ is 12° and $\theta 2$ is 20° . Thus, $\theta 1$ and $\theta 2$ differ only by 1–2°. The TS structures in water and TtCM are very similar. With the contracting of the C5-C16 distance (starting from the GS), the reaction pathway becomes narrow. The conformations of the chorismate are all NACs when $R_{C5-C16} < 2.83 \text{ \AA}$ in water and $R_{C5-C16} < 2.81 \text{ \AA}$ in TtCM where the “bottleneck” shapes occur. It has also been identified from Fig. 2 that the free-energy profiles nearly overlap with the free-energy profiles of the NAC formation when R_{C5-C16} is $< 2.83 \text{ \AA}$ in water and 2.81 \AA in TtCM.

Comparison of Conformations in Different Reaction Stages. The conformations of the TtCM-bound GS, NAC, TS, and product were extracted from trajectories and averaged. The interaction distances are defined as the distances between the geometric center of chorismate and that of the side chains of nearby residues (within 6 \AA), which are calculated from the four averaged structures as shown in Table 2. The interaction distances do not change much in the course of the reaction. For example, the interaction distance of the chorismate and Leu-114 are 5.98, 6.00, 6.18, and 6.07 \AA in the GS, NAC, TS, and product, respectively. The calculated radii of gyration for the heavy atoms of the chorismate and residues within 6 \AA of the chorismate are 7.20, 7.22, 7.23, and 7.20 \AA in the GS, NAC, TS, and product, respectively. Therefore, the radii of gyration also show no significant differences. The species housed in the active site were removed and the volumes of the empty cavities at various distances along the reaction coordinate were calculated by using the VOIDOO program (22) with a grid spacing and probe radius of 1.0 and 1.2 \AA , respectively. The cavity volumes of the active

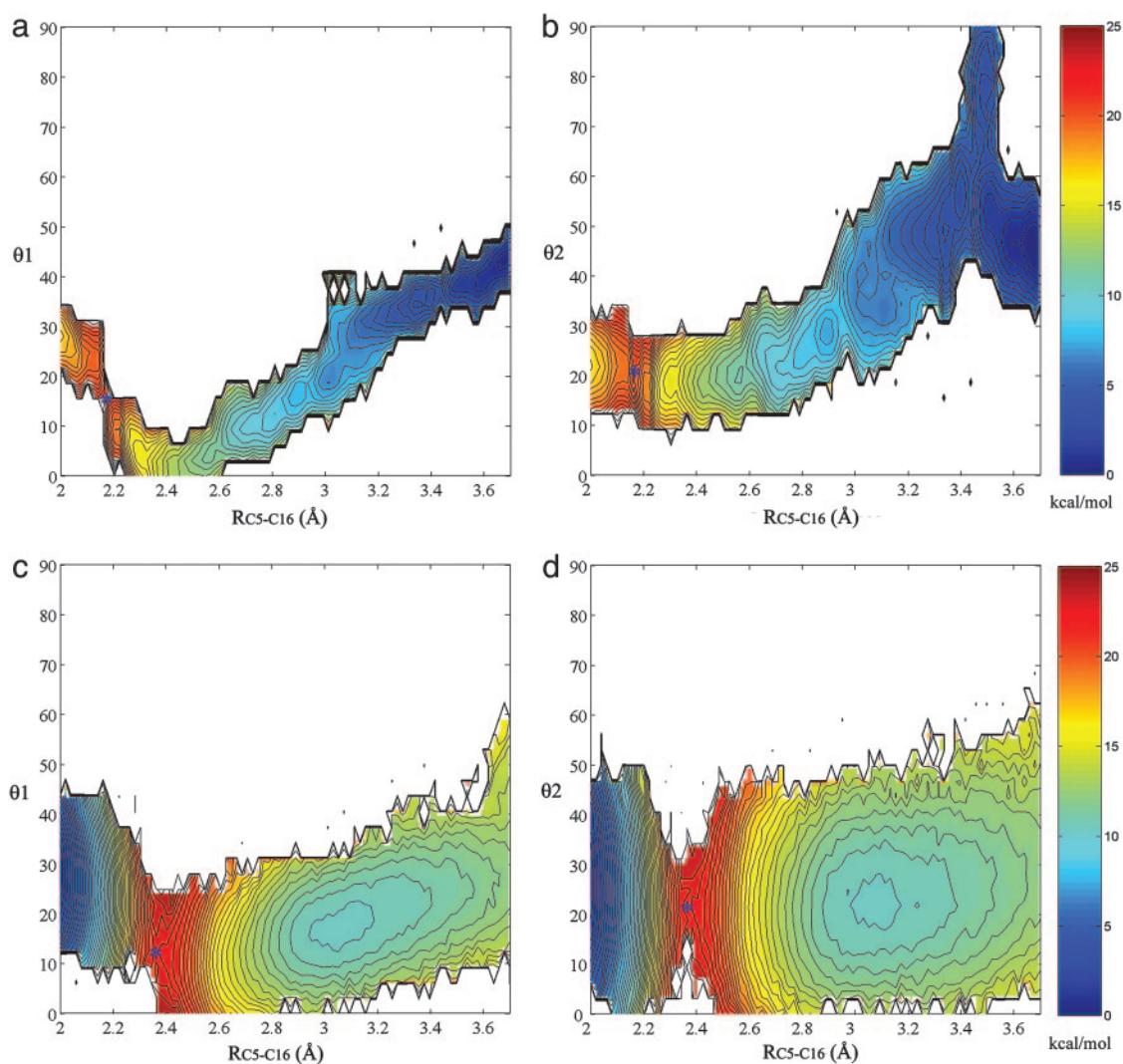


Fig. 3. NAC conformation characterization. The RC_{5-C16} ranges from 2.0 to 3.7 Å, θ_1 and θ_2 range from 0° to 90°, and the increment of the contour line is 0.5 kcal/mol. 1D PMFs along the reaction coordinate RC_{5-C16} are projected to the θ_1 and θ_2 , respectively. The TS is labeled by blue asterisks. (a and b) PMF of the reaction in water projected to θ_1 (a) and θ_2 (b). (c and d) PMF of the reaction in TtCM projected to θ_1 (c) and θ_2 (d).

site are 178 and 163 Å³ in the E·S and enzyme–TS complex (E·TS). The similarity of the cavity volumes of the active site indicates that the chorismate binds no tighter in the TS than GS. Superimposing TtCM-bound E·S and E·TS shows that the geometry of the active site structure remains unchanged (Fig. 5, which is published as supporting information on the PNAS web site).

TS Structure Analysis. The hydrophobic and electrostatic interactions of TS at the active site are shown in Fig. 4. As previously mentioned, the active site changes marginally when the chorismate goes from GS to the TS. Thus, the interactions of substrate and residues in E·TS and E·S are generally alike. The most important interactions are the hydrogen bonds between Arg-6 and Arg-63* and the two carboxylates. The residues labeled with * are from the adjacent subunit. These interactions hold the substrate in a di-axial conformation required for the reaction and remain in position to bind the TS (Fig. 4b).

Located beneath the TS are the nonpolar residues of the active site: Phe-57*, Ala-59*, Leu-72*, and Leu-73* (Fig. 4a), which form a basin to support the substrate. Another nonpolar residue, Leu-114, hangs over the TS, protruding from the other subunit.

These two sections of the residues constrain the TS by hydrophobic interactions. The interaction distances between these nonpolar residues and the TS are all ≈ 6 Å, as shown in Table 2. This finding indicates the TS to be in the center of the active site and with no preference to be closer to any of the nonpolar residues. The Phe-57*, Ala-59*, Leu-72*, Leu-73*, and Leu-114 nonpolar amino acids function as “bookends” to prevent the sideways motion of the C16 and C5 atoms in the chorismate from drifting away from each other.

Charge Distributions for E·S and E·TS. The Mulliken charge distributions of the structures, averaged from the trajectories of the GS, NAC, and TS in water and TtCM, were obtained by using single-point energy calculations at the SCCDFTB/MM level (Table 3, which is published as supporting information on the PNAS web site). We have found that the charge distributions do not change much from the GS to TS. The total charges of the six atoms, C3, C4, C5, O13, C14, and C16, which are involved in the pericyclic reaction, show remarkable similarity (−0.54, −0.52, and −0.53 for the GS, NAC, and TS, respectively, in water; −0.50, −0.51, and −0.56, respectively, in TtCM). The charge distributions of the chorismate atoms involved in hydrogen

identifying NAC. The term NAC to recognize conformations, in which reacting atoms are at van der Waals distance regardless of the attacking angle, is tantamount to the neglect of required orbital overlaps for reaction. Contributing to the confusion is a QM/MM study of a BsCM mutant (28) in which positively charged Arg-90 is replaced by a neutral citrulline without sampling of the degrees of freedom for the protein backbone. However, it had been shown that the mutation is accompanied by rearrangements at the active site (5).

In a recent report by Hilvert and coworkers (29), a concerted, albeit asynchronous, enzyme mechanism is supported by kinetic isotope effect studies of BsCM enzymatic reaction. The secondary tritium isotope effects on thermal reaction indicate a significant cleavage of C3-O13 bond but little formation of the C5-C16 at TS. Hilvert and coworkers' density functional theory calculations at the B3LYP/6-31+G* level predict R_{C5-C16} of 2.66 Å and R_{C3-O13} of 2.13 Å *in vacuo*. In our SCCDFT/MM approach, R_{C5-C16} and R_{C3-O13} have been determined to be 2.55 and 2.00 Å *in vacuo* (6) and 2.37 and 1.94 Å in the active site of TtCM, respectively. The differences of bond breaking and making, obtained by Hilvert and coworkers (-0.55 Å *in vacuo*) and our studies (-0.53 Å *in vacuo* and -0.43 Å in TtCM) are similar. Such a concerted mechanism has also been identified by our previous results with EcCM with the 2D QM/MM energy surface method (6).

Hilvert and coworkers (29) have proposed that Arg-90 in BsCM stabilizes the ether oxygen of chorismate in asynchronous TS (29). In current simulations of TtCM, the average distances of O13 of chorismate and NH1 of Arg-89 are 3.06 and 2.83 Å for E·S and E·TS, respectively, and the charges on O13 are -0.44 and -0.50 for E·S and E·TS, respectively. The interactions of Arg-89 and chorismate were evaluated from single-point SCCDFTB/MM energy calculation, where the only side chain of the Arg-89 and chorismate were included in the calculations. A link atom was added onto the C β atom of the Arg-89 side chain to preserve the integrated electron and structure of the Arg-89 side chain. The average electrostatic interactions between the Arg-89 side chain and chorismate (ΔE_{elec}) are -197.97 and -198.52 kcal/mol for E·S and E·TS, respectively. Thus, electrostatic interactions between the Arg-89 side chain and chorismate contribute 0.55 kcal/mol ($\Delta\Delta E_{elec}$) to the TS stabilization. The average van der Waals interactions between the Arg-89 side chain and chorismate (ΔE_{vdw}) are -1.96 and -1.99 kcal/mol for E·S and E·TS, respectively, which are almost the same. The total contributions of interactions between the Arg-89 side chain and chorismate to the TS stabilization are 0.58 kcal/mol.

The role of the Arg-90 in BsCM has also been studied by the Arg-90citrulline mutant. The free energy of activation for the reaction in the BsCM Arg-90citrulline mutant increases by 5.8 kcal/mol after mutation (30). The difference in free energies for NAC formation in the active site of the mutant and WT BsCM is 3.8 kcal/mol (4.1 and 0.3 kcal/mol for the mutant and WT BsCM, respectively) (4). Thus, Arg-90 was estimated to contribute 2.0 (5.8–3.8) kcal/mol to TS stabilization in BsCM. The values of the Arg-90 contribution to the TS stabilization in BsCM obtained by the mutations are larger than the direct calculation of the interactions (0.58 kcal/mol) between Arg-89 and chorismate in TtCM, which is possibly because of the rearrangement of the active site after mutations.

Conclusion

We have obtained the PMF profiles for the chorismate \rightarrow prephenate reaction both in water and a thermophilic enzyme by using the combination of SCCDFTB/MM, umbrella sampling, and the weighted histogram analysis method. The free energies of activation were determined from the PMF profile, and the free energies for NAC formation were calculated from the mole fraction of conformations in the GS present as NAC. Small numbers of the free energies for NAC formation in different CMs suggest that the GS is essentially represented by the NAC structure in the active site of the CM. The contribution of NAC formation to the lowering of the activation barrier for the reaction in the thermophilic enzyme is the same (8.01 kcal/mol at the optimum temperature, 343 K) as in the mesophilic one (8.0 and 7.8 kcal/mol for EcCM and BsCM, respectively, at 300 K). The TS structures in water and TtCM are similar. No significant changes in the electrostatic interactions and the geometry of the active site take place during the chorismate \rightarrow prephenate reaction. The Mulliken charges of atoms of the GS and TS species are very similar. The electrostatic and van der Waals interactions between the ether oxygen of chorismate and Arg-90 of 0.58 kcal/mol provide minimal TS stabilization. Only a small electrostatic stabilization contribution to the proficiency of the enzymatic reaction was identified based on the charge analysis. In essence, the enzyme proficiency originates from an entropic advantage of enzyme in forming the reactive GS conformations compared with water.

We thank Dr. Xiaodong Zhang and Joe Toporowski for helpful discussion and advice and the National Partnership for Advanced Computational Infrastructure for its generous allocation of computational resources at DataStar in the University of California at San Diego Supercomputing Center. This work was supported by National Institutes of Health Grant 5R37DK9174-41.

1. Guilford, W. J., Copley, S. D. & Knowles, J. R. (1987) *J. Am. Chem. Soc.* **109**, 5013–5019.
2. Bruice, T. C. (2002) *Acc. Chem. Res.* **35**, 139–148.
3. Ganem, B. (1996) *Angew. Chem. Int. Ed.* **35**, 937–945.
4. Hur, S. & Bruice, T. C. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 12015–12020.
5. Hur, S. & Bruice, T. C. (2003) *J. Am. Chem. Soc.* **125**, 5964–5972.
6. Zhang, X. D., Zhang, X. H. & Bruice, T. C. (2005) *Biochemistry* **44**, 10443–10448.
7. Mazumder-Shivakumar, D. & Bruice, T. C. (2005) *Biochemistry* **44**, 7805–7817.
8. Helmstaedt, K., Heinrich, G., Merkl, R. & Braus, G. H. (2004) *Arch. Microbiol.* **181**, 195–203.
9. Tahirov, T. H., Inagaki, E., Ohshima, N., Kitao, T., Kuroishi, C., Ukita, Y., Takio, K., Kobayashi, M., Kuramitsu, S., Yokoyama, S. & Miyano, M. (2004) *J. Mol. Biol.* **337**, 1149–1160.
10. Mattei, P., Kast, P. & Hilvert, D. (1999) *Eur. J. Biochem.* **261**, 25–32.
11. Warshel, A. & Levitt, M. (1976) *J. Mol. Biol.* **103**, 227–249.
12. Lazaridis, T., Tobias, D. J., Brooks, C. L. & Paulaitis, M. E. (1991) *J. Chem. Phys.* **95**, 7612–7625.
13. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem.* **4**, 187–217.
14. Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S. & Seifert, G. (1998) *Phys. Rev. B* **58**, 7260–7268.
15. Cui, Q. & Karplus, M. (2002) *J. Am. Chem. Soc.* **124**, 3093–3124.
16. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983) *J. Chem. Phys.* **79**, 926–935.
17. Brooks, C. L. & Karplus, M. (1989) *J. Mol. Biol.* **208**, 159–181.
18. Beveridge, D. L. & Dicapua, F. M. (1989) *Annu. Rev. Biophys. Biophys. Chem.* **18**, 431–492.
19. Strajbl, M., Shurki, A., Kato, M. & Warshel, A. (2003) *J. Am. Chem. Soc.* **125**, 10228–10237.
20. Ferrenberg, A. M. & Swendsen, R. H. (1989) *Phys. Rev. Lett.* **63**, 1195–1198.
21. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. (1992) *J. Comput. Chem.* **13**, 1011–1021.
22. Kleywegt, G. J. & Jones, T. A. (1994) *Acta Crystallogr. D* **50**, 178–185.
23. Gustin, D. J., Mattei, P., Kast, P., Wiest, O., Lee, L., Cleland, W. W. & Hilvert, D. (1999) *J. Am. Chem. Soc.* **121**, 1756–1757.
24. Guo, H., Cui, Q., Lipscomb, W. N. & Karplus, M. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 9032–9037.
25. Guo, H., Cui, Q., Lipscomb, W. N. & Karplus, M. (2003) *Angew. Chem. Int. Ed.* **42**, 1508–1511.
26. Repasky, M. P., Guimaraes, C. R. W., Chandrasekhar, J., Tirado-Rives, J. & Jorgensen, W. L. (2003) *J. Am. Chem. Soc.* **125**, 6663–6672.
27. Guimaraes, C. R. W., Repasky, M. P., Chandrasekhar, J., Tirado-Rives, J. & Jorgensen, W. L. (2003) *J. Am. Chem. Soc.* **125**, 6892–6899.
28. Guimaraes, C. R. W., Udier-Blagovic, M., Tubert-Brohman, I. & Jorgensen, W. L. (2005) *J. Chem. Theory Comput.* **1**, 617–625.
29. Wright, S. K., DeClue, M. S., Mandal, A., Lee, L., Wiest, O., Cleland, W. W. & Hilvert, D. (2005) *J. Am. Chem. Soc.* **127**, 12957–12964.
30. Kienhofer, A., Kast, P. & Hilvert, D. (2003) *J. Am. Chem. Soc.* **125**, 3206–3207.