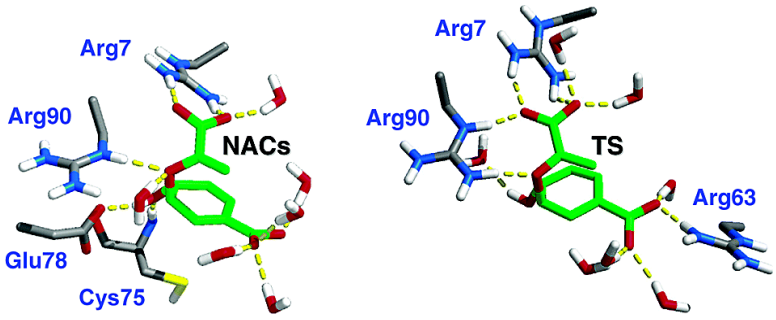
Article

Contributions of Conformational Compression and Preferential Transition State Stabilization to the Rate Enhancement by Chorismate Mutase

Cristiano Ruch Werneck Guimares, Matthew P. Repasky, Jayaraman Chandrasekhar, Julian Tirado-Rives, and William L. Jorgensen

J. Am. Chem. Soc., **2003**, 125 (23), 6892-6899• DOI: 10.1021/ja021424r • Publication Date (Web): 10 May 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 05/10/2003

Contributions of Conformational Compression and Preferential Transition State Stabilization to the Rate Enhancement by Chorismate Mutase

Cristiano Ruch Werneck Guimarães, Matthew P. Repasky,[†] Jayaraman Chandrasekhar,[‡] Julian Tirado-Rives, and William L. Jorgensen*

Contribution from the Department of Chemistry, Yale University, 225 Prospect Street, New Haven, Connecticut 06520-8107

Received December 9, 2002; E-mail: william.jorgensen@yale.edu

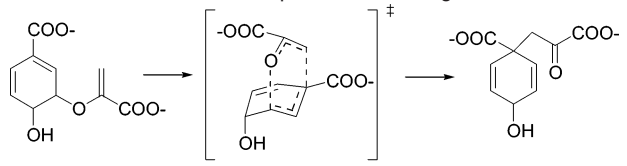
Abstract: The rate enhancement provided by the chorismate mutase (CM) enzyme for the Claisen rearrangement of chorismate to prephenate has been investigated by application of the concept of near attack conformations (NACs). Using a combined QM/MM Monte Carlo/free-energy perturbation (MC/FEP) method, 82% and 100% of chorismate conformers were found to be NAC structures in water and in the CM active site, respectively. Consequently, the conversion of non-NACs to NACs does not contribute to the free energy of activation from preorganization of the substrate into NACs. The FEP calculations yielded differences in free energies of activation that well reproduce the experimental data. Additional calculations indicate that the rate enhancement by CM over the aqueous phase results primarily from conformational compression of NACs by the enzyme and that this process is enthalpically controlled. This suggests that preferential stabilization of the transition state in the enzyme environment relative to water plays a secondary role in the catalysis by CM.

Introduction

The reliable prediction of enzymatic mechanisms and rate enhancements remains a great challenge.[1−3] A complicating issue is the relative dearth of enzymes that catalyze reactions which occur via an identical mechanism in both aqueous and enzyme environments. An important exception is chorismate mutase (CM), which catalyzes the Claisen rearrangement of chorismate to prephenate (Scheme 1), a key step in the Shikimate pathway for generating aromatic amino acids in plant, fungal, and bacterial systems.[4−6]

CM holds many advantages enabling quantitative elaboration of details responsible for the observed rate enhancements. First, the catalyzed reaction is an intramolecular rearrangement with no substrate−protein covalent linkages formed or broken over the course of the transformation. This simplifies the application of combined quantum and molecular mechanics (QM/MM) methods. Second, much experimental kinetic data and crystal structures exist for several CM enzymes including those from

Scheme 1. Chorismate to Prephenate Rearrangement



Chorismate Prephenate

Escherichia coli,[7] *Bacillus subtilis*,[8,9] and yeast.[10−12] Finally, because of the relatively small size of the natural substrate, detailed investigations with ab initio QM methods have revealed important structural and energetic details of the reaction.[13−15]

In another paper,[16] the origins of solvent effects on the rate of the Claisen rearrangement of chorismate to prephenate were examined in the gas phase, methanol, and water using the concept of near attack conformations (NACs).[17−19] Following the methodology of this previous study,[16] we now turned our

[†] Current address: Schrödinger Inc., 120 West Forty-Fifth Street, New York, NY 10036.

[‡] Current address: Neurogen Corporation, 35 Northeast Industrial Rd., Brandford, CT 06405.

(1) Blow, D. *Structure* **2000**, *8*, R77−R81.
(2) Kollman, P. A.; Kuhn, B.; Perakyla, M. *J. Phys. Chem. B* **2002**, *106*, 1537−1542.
(3) Shurki, A.; Strajbl, M.; Villa, J.; Warshel, A. *J. Am. Chem. Soc.* **2001**, *124*, 4097−4107.
(4) Ganem, B. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 936−945.
(5) Haslam, E. *Shikimic Acid: Metabolism and Metabolites*; John Wiley & Sons: New York, 1993.
(6) Conn, E. E. In *The Shikimic Acid Pathway*; Conn, E. E., Ed.; Recent Advances in Phytochemistry; Plenum Press: New York, 1986.
(7) Lee, A. Y.; Karplus, P. A.; Ganem, B.; Clardy, J. *J. Am. Chem. Soc.* **1995**, *117*, 3627−3628.
(8) Chook, Y. M.; Gray, J. V.; Ke, H.; Lipscomb, W. N. *J. Mol. Biol.* **1994**, *240*, 476−500.
(9) Ladner, J. E.; Reddy, P.; Davis, A.; Tordova, M.; Howard, A. J.; Gilliland, G. L. *Acta Crystallogr.* **2000**, *D56*, 673−683.
(10) Strater, N.; Hakansson, K.; Schnappauf, G.; Braus, G.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 3330−3334.
(11) Strater, N.; Schnappauf, G.; Braus, G.; Lipscomb, W. N. *Structure* **1997**, *5*, 1437−1452.
(12) Xue, Y.; Lipscomb, W. N.; Gray, R.; Schnappauf, G.; Braus, G. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10814−10818.
(13) Wiest, O.; Houk, K. N. *J. Org. Chem.* **1994**, *59*, 7582−7584.
(14) Wiest, O.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 11628−11639.
(15) Davidson, M. M.; Guest, J. M.; Craw, J. S.; Hillier, I. H.; Vincent, M. A. *J. Chem. Soc., Perkin Trans. 2* **1997**, *7*, 1395−1400.
(16) Repasky, M. P.; Guimarães, C. R. W.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* In press.

10.1021/ja021424r CCC: $25.00

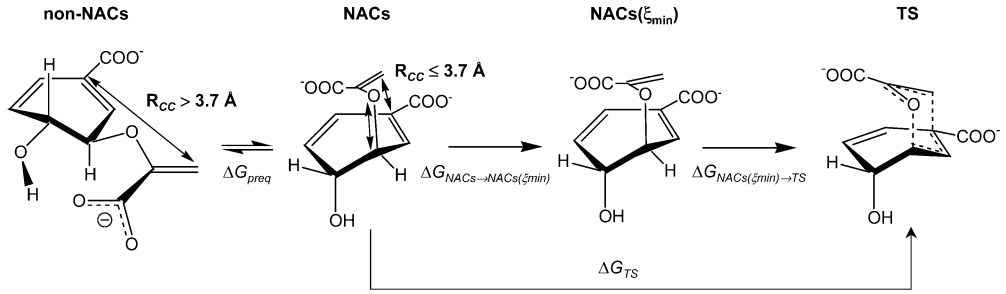


Figure 1. Important steps in the NAC-based model for the chorismate to prephenate rearrangement. Chorismate conformers with an $R_{CC} \leq 3.7$ Å are considered as NACs.

attention to determining the detailed basis of catalysis by the CM enzyme. The goal is to identify the key factors and to estimate their quantitative contributions to the overall rate enhancement in CM relative to an aqueous environment.

This system has been the subject of a variety of experimental[20−26] and theoretical[17,21,27−36] investigations over the last 20 years, and a number of potential mechanisms for the observed rate enhancement have been suggested. Preferential transition state (TS) stabilization in the enzyme environment has been assigned the majority of the rate enhancement in the recent theoretical studies of Hall et al.[30] However, Hur and Bruice[17] suggested, against the objections of Shurki et al.,[3] that the rate enhancement in CM may arise from the favorable formation of NACs. In another study, upon application of the spatiotemporal theory to CM, Menger and co-workers suggested that a protein-driven conformational compression of the reactant provided sufficient destabilization to account for the observed rate enhancement.[21]

In this work, the Claisen rearrangement occurring in *Bacillus subtilis* CM was studied using a QM/MM method with Monte Carlo (MC) statistical mechanics and free-energy perturbation (FEP) calculations. Upon demonstration of the applicability of the theoretical model for the enzymatic reaction, the basis for the observed increase in the rate constant relative to that in aqueous solution (3×10^6) is addressed. As the contributions to the free energy of activation ($\Delta G^{\ddagger}$) from preorganization of the substrate as NACs, from conformational compression of the reactant, and from preferential TS stabilization are thermodynamically separable, estimates of their relative magnitudes are included in the analyses.

For a unimolecular reaction such as the Claisen rearrangement, $\Delta G^{\ddagger}$ either in water or in the protein environment can be divided into two components (Figure 1). The first component, the preequilibrium free energy (ΔG_{preq}), is the free energy associated with the process of orienting chorismate in the aqueous phase or in the Michaelis complex in appropriate conformations, NACs, that are well juxtaposed to lead directly to the TS of the reaction. The second component of $\Delta G^{\ddagger}$ is the free energy associated with the conversion of NACs into the TS (ΔG_{TS}). As described in another paper,[16] to facilitate computations of potentials of mean force (pmf's),[37] ΔG_{TS} is broken into two quantities, the free energy associated with the conversion of all NACs into NACs(ξ_{min}) ($\Delta G_{\text{NACs} \rightarrow \text{NACs}(\xi_{\text{min}})}$) and the free energy associated with the conversion of NACs(ξ_{min}) into the TS ($\Delta G_{\text{NACs}(\xi_{\text{min}}) \rightarrow \text{TS}}$). NACs($\xi_{\text{min}}$) is a subset of NACs, which displays the lowest free energy along the reaction coordinate ξ that generates the TS. ξ is taken here as the difference in the lengths of the forming C−C and breaking C−O bonds. Thus, the reaction in CM can be faster than the reaction in water by increasing the population of NACs and/or by binding the TS more favorably than NACs.

Computational Details

To treat the Claisen reaction in CM and obtain ΔG_{preq}, $\Delta G_{\text{NACs} \rightarrow \text{NACs}(\xi_{\text{min}})}$, and $\Delta G_{\text{NACs}(\xi_{\text{min}}) \rightarrow \text{TS}}$ (Figure 1), the procedure described in another paper was employed.[16] The gas-phase MERP (minimum energy reaction path) from AM1 calculations was used to determine "1D" pmf's for the reaction in the gas phase and in solution with complete solute flexibility except for the reaction coordinate. The reaction coordinate was subsequently relaxed through additional FEP calculations off the MERP to locate the exact critical points in the different media. In practice, the 1D pmf's closely identified the reaction coordinate values for the true free-energy minimum and maximum in each case.[16] This was again found to be true for the TS of the enzymatic reaction. In view of these results and the high computational demands of the off-MERP FEP calculations, the structure with the lowest free energy along the 1D pmf was taken as the true minimum for chorismate in CM. The populations of NACs and non-NACs were computed as before by obtaining a pmf for chorismate as a function of the forming C−C bond, $G(R_{\text{CC}})$.[16]

Cartesian coordinates for the 2.2 Å *Bacillus subtilis* crystal structure complexed with Bartlett's transition state analogue (2cht from the Brookhaven Protein Data Bank) were employed (Figure 2). Of the four homotrimers of chorismate mutase identified in the 2cht structure, only

(17) Hur, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1176−1181.
(18) Lightstone, F. C.; Bruice, T. C. *J. Am. Chem. Soc.* **1996**, *118*, 2595−2605.
(19) Bruice, T. C.; Lightstone, F. C. *Acc. Chem. Res.* **1999**, *32*, 127−136.
(20) Kast, P. M.; Ullah, A.; Hilvert, D. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.
(21) Khanjin, N. A.; Snyder, J. P.; Menger, F. M. *J. Am. Chem. Soc.* **1999**, *121*, 11831−11846.
(22) Meyer, M. P.; DelMonte, A. J.; Singleton, D. A. *J. Am. Chem. Soc.* **1999**, *121*, 10865−10874.
(23) Mattei, P.; Kast, P.; Hilvert, D. *Eur. J. Biochem.* **1999**, *261*, 25−32.
(24) Kast, P.; Grisostomi, C.; Chen, I. A.; Li, S.; Krengel, U.; Xue, Y.; Hilvert, D. *J. Biol. Chem.* **2000**, *275*, 36832−36838.
(25) Copley, S. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1987**, *109*, 5008−5013.
(26) Andrews, P. R.; Smith, G. D.; Young, I. G. *Biochemistry* **1973**, *12*, 3492−3498.
(27) Marti, S.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *J. Phys. Chem. B* **2000**, *104*, 11308−11315.
(28) Marti, S.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *Theor. Chem. Acc.* **2001**, *105*, 207−212.
(29) Marti, S.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J.; Field, M. J. *J. Am. Chem. Soc.* **2001**, *123*, 1709−1712.
(30) Hall, R. J.; Hindle, S. A.; Burton, N. A.; Hillier, I. H. *J. Comput. Chem.* **2000**, *21*, 1433−1441.
(31) Worthington, S. E.; Krauss, M. *J. Phys. Chem. B* **2001**, *105*, 7096−7098.
(32) Worthington, S. E.; Roitberg, A. E.; Krauss, M. *J. Phys. Chem. B* **2001**, *105*, 7087−7095.
(33) Davidson, M. M.; Gould, I. R.; Hillier, I. H. *J. Chem. Soc., Perkins Trans. 2* **1996**, *4*, 525−532.
(34) Kangas, E.; Tidor, B. *J. Phys. Chem. B* **2001**, *105*, 880−888.
(35) Lyne, P. D.; Mulholland, A. J.; Richards, W. G. *J. Am. Chem. Soc.* **1995**, *117*, 11345−11350.
(36) Carlson, H. A.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1996**, *118*, 8475−8484.
(37) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.

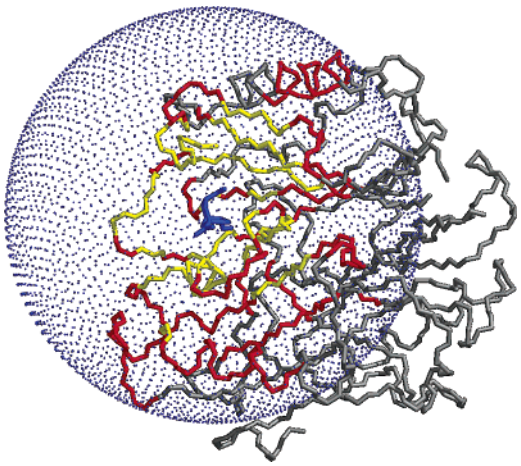
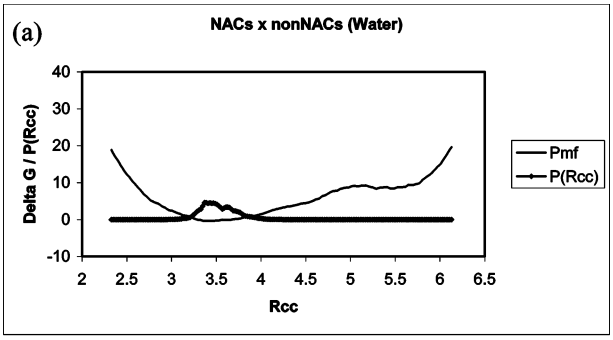


Figure 2. Chorismate mutase with Bartlett's inhibitor shown in blue. The neglected residues are colored gray with all yellow and red residues explicitly included in the simulations. Amino acid side chains are flexible only for yellow residues. The magenta dots represent the 22 Å radius water cap of 766 molecules.

the first was retained. A single active site was taken as the center of the reacting system. Residues with any atom within 15 Å from the center of the reacting system were retained in the simulations, and any clipped residues were capped with acetyl and *N*-methylamine groups. The enzymatic system then has 2249 atoms, consisting of 149 amino acid residues and a single solute. In the QM/MM calculations, degrees of freedom for the protein backbone atoms were not sampled. Only side chains of residues with any atom within 10 Å from the center of the solute were varied. Charge neutrality was imposed by having a total protein charge of 2 e; charged residues near the active site were assigned normal protonation states for physiological pH, and the adjustments for neutrality were made to the most distant residues. The entire protein−solute system was solvated with a 22 Å radius water cap of 766 molecules, and a half-harmonic potential with a 1.5 kcal/mol·Å^2 force constant was applied to water molecules at distances greater than 22 Å from the center of the solute. To insert the reacting system into the protein environment, the TS identified from the gas-phase MERP[16] was optimally overlaid by matching heavy atoms with the structure of Bartlett's inhibitor. Successive structures leading from the TS to the reactant and from the TS to the product were also given initial geometries by similar overlaying in the enzyme active site. The preparation of the enzymatic system was much facilitated by use of the chop delegate and MidasPlus 2.1.[38]

Changes in free energy in the protein environment were computed via MC statistical mechanics in conjunction with FEP calculations at 25 °C using double-wide sampling. The combined QM/MM method as implemented in MCPRO 2.0[39] was used to perform all MC/FEP calculations. Established procedures including Metropolis and preferential sampling and the Zwanzig equation were employed.[40,41] Statistical uncertainties were obtained from the batch means procedure with batch sizes of 1×10^6 configurations.[41] The solute intramolecular energy, Lennard−Jones parameters and partial atomic charges, the intermolecular energy between QM and MM components, and the frequency of the QM energy and charge computations have been previously detailed for the present QM/MM MC/FEP implementation.[16,42] The protein was represented with the OPLS-AA force field,[43] the TIP4P model was used for water,[44] and residue-based cutoffs of 10 Å were employed. To ensure convergence, each MC simulation was extensive. Initial reorganization of the solvent was performed for 2×10^6 configurations. This was followed by 10×10^6 configurations of full equilibration and 20×10^6 configurations of averaging for each window in each MC/FEP calculation.



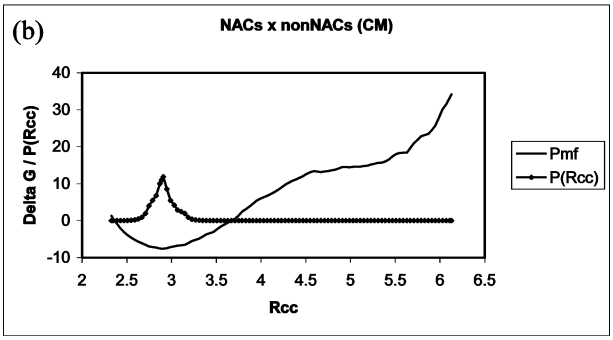


Figure 3. Plots of the one-dimensional potentials of mean force for the change in R_{CC} (the forming C−C bond length) and the probability of sampling it in (a) water and (b) CM.

Results and Discussion

NACs versus Non-NACs. Figure 3 plots the pmf curves ($G(R_{CC})$) and the probability of sampling R_{CC} ($P(R_{CC})$) in water and CM. All conformations with an R_{CC} less than or equal to 3.70 Å are characterized as NACs.[16] The population of NACs in water is 81.9%, and the population of non-NACs is 18.1%, which results in a ΔG_{preq} of −0.9 kcal/mol (Figure 3a). Figure 3b shows that, when chorismate is bound to CM, the population of NACs increases to almost 100%, which makes ΔG_{preq} in CM very negative (−9.0 kcal/mol). Although the population of NACs is increased in CM relative to the aqueous phase, the preequilibrium does not contribute to $\Delta G^{\ddagger}$, as ΔG_{preq} is negative in both environments. Therefore, the 3×10^6-fold increase in the rate constant of the Claisen rearrangement in going from water to CM should arise solely from the free-energy difference associated with the conversion of NACs to the transition state (ΔG_{TS}) in each environment. In other words, the reaction in CM is probably faster than the reaction in water because CM binds the TS more favorably than NACs.

Pmf's Using the Gas-Phase MERP. Figure 4 shows the changes in free energy from the 1D pmf calculations in the aqueous and enzymatic environments. The smooth profiles and

(38) (a) Tirado-Rives, J. *Chop*; Yale University: New Haven, CT, 2002. (b) Huang, C.; Pettersen, E.; Couch, G.; Ferrin, T. *MidasPlus 2.1*; University of California: San Francisco, CA, 1994.
(39) Jorgensen, W. L. *MCPRO*, version 2.0; Yale University: New Haven, CT, 2002.
(40) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(41) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Clarendon Press: Oxford, U.K., 1987.
(42) Kaminski, G. A.; Jorgensen, W. L. *J. Phys. Chem. B* **1998**, *102*, 1787−1796.
(43) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

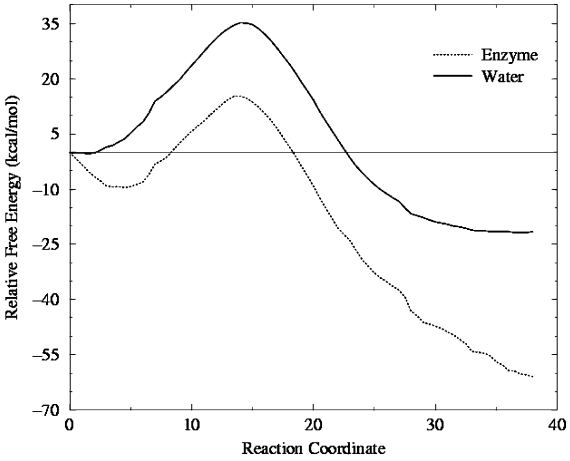


Figure 4. Plots of the one-dimensional potentials of mean force in water and CM for the chorismate to prephenate rearrangement.

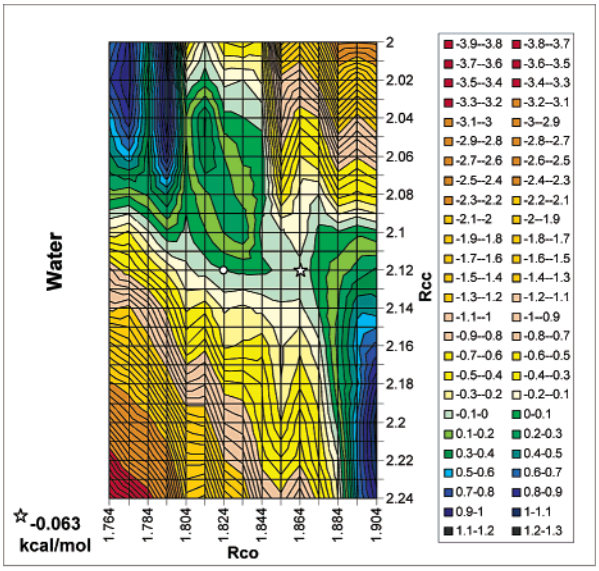
Table 1. R_{CO} and R_{CC} Distances Obtained for the Reactant and the TS in Water and in CM[a]

	water	
	R_{CO}	R_{CC}
reactant		
1D pmf free-energy minimum	1.442	3.448
NACs(ξ_{min})	1.44	3.53
TS		
1D pmf TS	1.824	2.120
optimized TS	1.86	2.12
	CM	
	R_{CO}	R_{CC}
reactant		
1D pmf free-energy minimum	1.470	2.870
NACs(ξ_{min})		
Marti et al.[b]	1.453	3.047
Lyne et al.[c]	1.469	2.849
TS		
1D pmf TS	1.824	2.120
optimized TS	1.86	2.20
Marti et al.[b]	1.944	2.181
Lyne et al.[c]	1.816	2.119

[a] Values in Å. [b] Reference 29. [c] Reference 35.

the good agreement with previous QM/MM studies of Lyne and co-workers[35] suggest that the choice of window sizes and the number of configurations sampled were appropriate. The reaction coordinate was followed from the gas-phase geometry for ξ in the potential energy minimum for chorismate, with R_{CO} and R_{CC} distances of 1.431 and 3.668 Å, to the gas-phase minimum energy structure for prephenate. In water, a slightly compressed geometry was identified as the free-energy minimum in the 1D pmf curve, with R_{CO} and R_{CC} distances of 1.442 and 3.448 Å. The geometry identified in the protein environment lies along a broad minimum in the 1D pmf and is significantly more compressed than the geometry in water with $R_{CO} = 1.470$ Å and $R_{CC} = 2.870$ Å. Formation of these structures is favorable by -0.4 ± 0.1 and -9.5 ± 0.3 kcal/mol in water and in the enzyme, respectively. Comparison of the NACs(ξ_{min}) in water[16] to the free-energy minimum identified by the 1D pmf shows that the 1D pmf closely identified the true free-energy minimum (Table 1).

Refinement of Reactant and TS. As mentioned in the Computational Details section, the identification of the true free-energy minimum in CM is computationally taxing. Since the



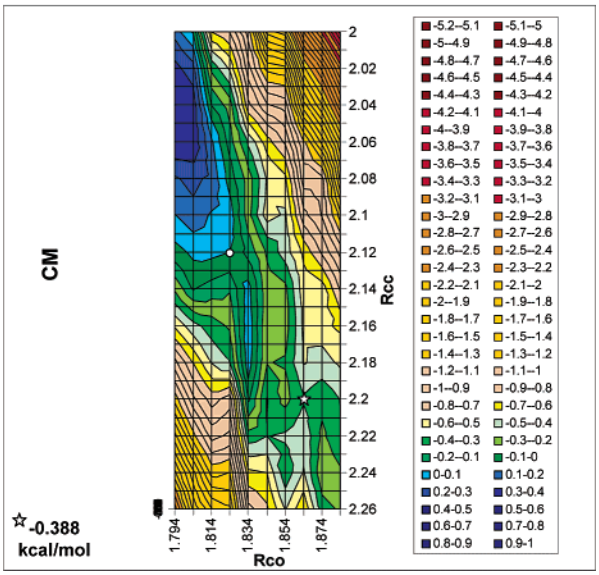


Figure 5. Plots of the two-dimensional potentials of mean force in water and CM, which identify the free-energy maximum with respect to ξ in each environment. (○) indicates the TS obtained from the gas-phase MERP calculation and 1D pmf's. (☆) indicates the TS identified by the map.

1D pmf calculations in the gas phase, methanol, and water closely identified the true free-energy minimum in each case,[16] the NACs(ξ_{min}) in CM was assumed to be the same as that from the 1D pmf. This is further supported by the accord between the geometry identified by the 1D pmf curve in the protein environment and the computed results of Marti et al.[29] and Lyne et al.[35] (Table 1).

The transition state structures in water and in CM were identified from the 1D pmf's as having R_{CO} and R_{CC} bond lengths of 1.824 and 2.120 Å, which is the same ξ geometry for the transition state as that obtained from the gas-phase MERP. To locate the true transition structures, the region surrounding the free-energy maximum in the 1D pmf curves was mapped (Figure 5). Table 1 shows the resultant optimized TS structures in water and in CM; they are more stable than the TS structures obtained from the 1D pmf's by 0.06 and 0.39

Table 2. Computed and Experimental Free-Energy Changes for the Claisen Rearrangement of Chorismate[a]

	water	CM (*Bacillus subtilis*)
ΔG_{preq}[b]	−0.9	−9.0
$\Delta G_{NACs\rightarrow NACs(\xi_{min})}$[c]	2.7	2.5
$\Delta G_{NACs(\xi_{min})\rightarrow TS}$[d]	35.4 ± 0.5	24.4 ± 0.5
ΔG_{TS}[e]	38.1 ± 0.5	26.9 ± 0.5
$\Delta G^{\ddagger}_{calc}$[f]	38.1 ± 0.5	26.9 ± 0.5
$\Delta G^{\ddagger}_{exp}$	24.5[g]	15.4[g]
$\Delta\Delta G^{\ddagger}_{calc}$	0.0	−11.2 ± 0.7
$\Delta\Delta G^{\ddagger}_{exp}$	0.0[g]	−9.1[g]

[a] Values in kcal/mol. [b] Obtained by eq 1, ref 16. [c] Obtained by eq 2, ref 16. [d] The error in $\Delta G_{NACs(\xi_{min})\rightarrow TS}$ was calculated by propagating the standard deviation (σ_i) on the individual ΔG_i used to obtain the pmf curve. The equation $[\sum_i^N \sigma_i^2]^{1/2}$ was used, where N is the number of ΔG_i values. [e] $\Delta G_{TS} = \Delta G_{NACs\rightarrow NACs(\xi_{min})} + \Delta G_{NACs(\xi_{min})\rightarrow TS}$. [f] When ΔG_{preq} is positive, $\Delta G^{\ddagger}_{calc} = \Delta G_{preq} + \Delta G_{TS}$. When it is negative, $\Delta G^{\ddagger}_{calc} = \Delta G_{TS}$. [g] Reference 45.

kcal/mol, respectively. Comparison of the TS geometries obtained in water and CM to the MERP and 1D pmf results indicates that the environment plays a minor role in the TS structures. The TS obtained here in CM is also very similar to the findings of the previous theoretical studies of Marti et al.[29] and Lyne et al.[35] (Table 1).

After identifying NACs(ξ_{min}) and the TS in each environment, the activation barriers obtained from the 1D pmf's were corrected. The results for the free energy of conversion of NACs(ξ_{min}) into the TS, $\Delta G_{NACs(\xi_{min})\rightarrow TS}$, are 35.4 ± 0.5 and 24.4 ± 0.5 kcal/mol in water and in the enzyme, respectively (Table 2). These values must be added to $\Delta G_{NACs\rightarrow NACs(\xi_{min})}$ to compute ΔG_{TS}. $\Delta G_{NACs\rightarrow NACs(\xi_{min})}$ was calculated by estimating the probability of sampling the subset NACs(ξ_{min}) from all NACs in water and CM,[16] which amounts to 1.0% and 1.5%, respectively. These probabilities give $\Delta G_{NACs\rightarrow NACs(\xi_{min})}$ values of 2.7 and 2.5 kcal/mol. As ΔG_{preq} is negative in water and CM, ΔG_{TS} is equal to $\Delta G^{\ddagger}$. Thus, the computed values of $\Delta G^{\ddagger}$ in water and CM are 38.1 ± 0.5 kcal/mol and 26.9 ± 0.5 kcal/mol, respectively (Table 2). These values compare well to 38.0 and 29.3 kcal/mol from the QM/MM pmf study of Marti et al.,[29] which also employed AM1 as the QM Hamiltonian. However, these barriers are larger than the experimental $\Delta G^{\ddagger}$ values of 24.5 and 15.4 kcal/mol for water and the enzyme, respectively.[45] This difference arises from the previously described deficiency in the AM1 activation barriers.[16]

For the chorismate to prephenate reaction, which exhibits electron delocalization in the transition state, inclusion of electron correlation is probably required. Since semiempirical methods have been parametrized to describe ground states, a poor energetic description should be expected only for the TS. Fortunately, as the water and protein environments provide similar TS structures, in examining the difference in the free energy of activation between CM and water, $\Delta\Delta G^{\ddagger}$, errors from the semiempirical QM method should largely cancel. Indeed, the calculated $\Delta\Delta G^{\ddagger}$ (−11.2 ± 0.7 kcal/mol) compares favorably to the experimental value (−9.1 kcal/mol). Moreover, Mattei et al. have shown that the *Bacillus subtilis* CM is partially limited by diffusive processes, which makes the enzyme less efficient in catalyzing the rearrangement.[23] In other words, by considering only the reaction step, the experimental $\Delta\Delta G^{\ddagger}$ should be more negative than −9.1 kcal/mol. This suggests that the current QM/MM MC/FEP method provides a reliable technique for examining the basis of the rate enhancement by chorismate mutase relative to water.

Table 3. Average Solute−Solvent/Protein Intermolecular Interaction Energies[a]

	average intermolecular energies (kcal/mol)		
	1D pmf starting point[b]	NACs(ξ_{min})$_{CM}$[c]	TS$_{CM}$[d]
	Lennard−Jones		
solvent	2.0	6.5	4.7
protein	−15.2	−18.4	−16.6
total	−13.2	−11.9	−11.9
	Electrostatics		
solvent	−31.3	−111.5	−121.5
protein	−382.2	−396.0	−417.9
total	−413.5	−507.5	−539.4

[a] Values in kcal/mol. [b] Gas-phase geometry for ξ in the potential energy minimum. [c] NACs(ξ_{min})$_{CM}$ is NACs(ξ_{min}) obtained in the protein. [d] TS$_{CM}$ is TS obtained in the protein.

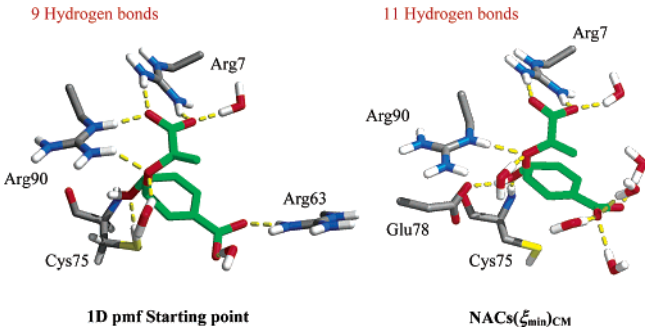


Figure 6. Snapshot of hydrogen-bonding interactions with the 1D pmf starting point and NACs(ξ_{min})$_{CM}$, which reproduce the average number of hydrogen bonds. Both structures were the final configuration from MC runs.

Protein-Driven Compression of NACs(ξ_{min}). Figure 4 shows that NACs(ξ_{min}) from the 1D pmf is compressed and care must be taken to ensure that the smaller calculated R_{CC} distance is not an artifact arising from poor initial protein-solute interactions or from details of the system setup. As mentioned previously, Marti and co-workers[29] have identified a similar minimum energy structure in their QM/MM pmf study. Their lowest-energy structure for the reactant has an R_{CC} only 0.2 Å longer than that of the structure identified here from the 1D pmf. As they employed a fully flexible protein backbone in the MD simulations, the small added compression found here does not likely arise from relief of overly poor protein-solute contacts. To pursue this point, the average solute−protein/solvent interaction energies over 30 × 10^6 MC configurations were decomposed into their Lennard−Jones (LJ) and Coulombic components, as summarized in Table 3. In going from the gas-phase geometry for ξ in the potential energy minimum for chorismate (starting point of the 1D pmf) to NACs(ξ_{min}) in the protein, NACs(ξ_{min})$_{CM}$, the LJ interactions are computed to actually become less favorable by a small amount, 1.3 kcal/mol. If there was a significant steric problem, the LJ interactions would be expected to become significantly more favorable. The shift to the structure for NACs(ξ_{min})$_{CM}$ arises, in fact, from improvement in the Coloumbic interactions with the solvent and the protein. There is a very large increase in the favorable solute−solvent energy of −80.2 kcal/mol. This is combined with an increase in the favorable solute−protein energy of −13.8 kcal/mol to

(45) Galopin, C. C.; Zhang, S.; Wilson, D. B.; Ganem, B. *Tetrahedron Lett.* **1996**, *37*, 8675−8678.

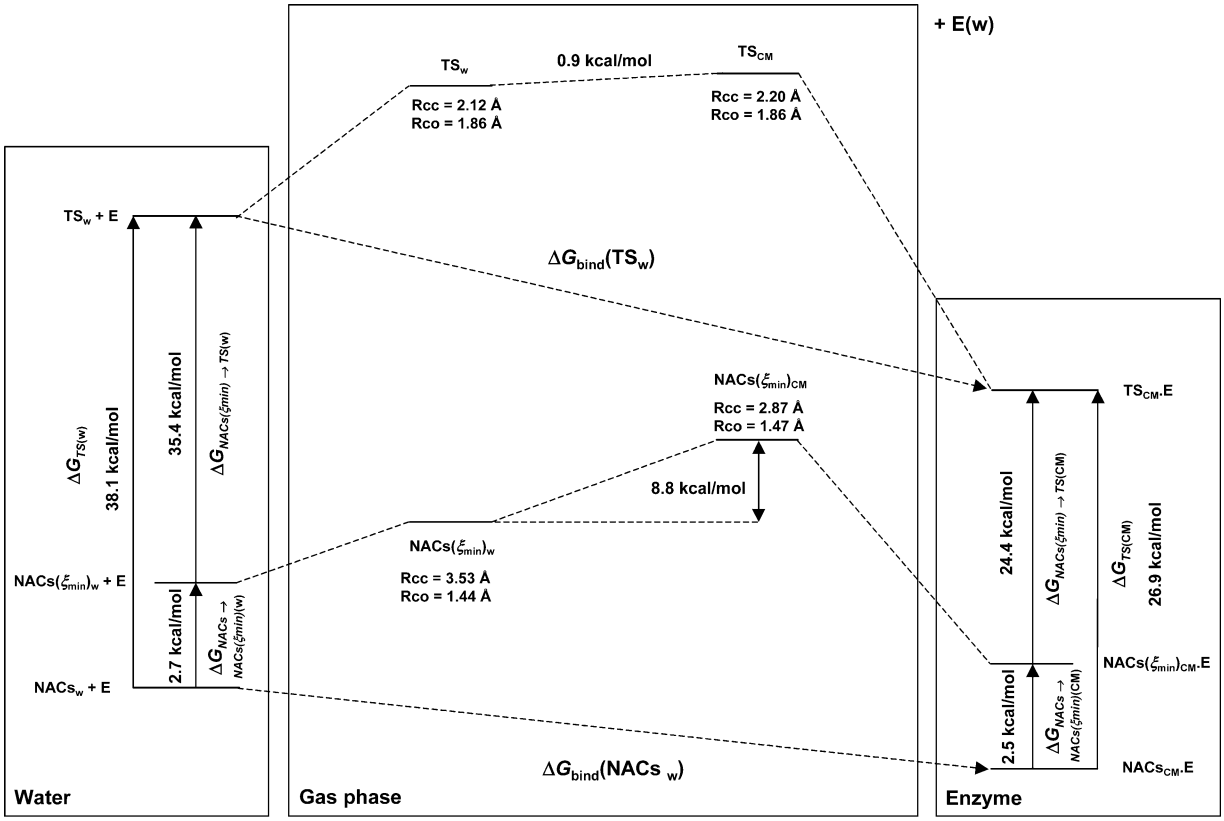


Figure 7. Thermodynamic cycle for the relative binding to CM by the TS and NACs.

give a total gain in electrostatic interaction energy of −94.0 kcal/mol.

Hydrogen bond analyses were also performed using 250 configurations that were saved at constant intervals during the 20×10^6 averaging steps in the protein environment. Differences in hydrogen bonding between the 1D pmf starting point and NACs$(\xi_{min})_{CM}$ were analyzed using geometric criteria to define a hydrogen bond.[46] Typical structures are shown in Figure 6. The 1D pmf starting point was found to make on average three hydrogen bonds to water, with one solvent molecule interacting with the enol pyruvate carboxylate, the ether oxygen, and the cyclohexadienyl carboxylate. Six hydrogen bonds were found formed to the protein on average, two each with Arg7 and Arg90 and one each from Arg63 and Cys75. For NACs$(\xi_{min})_{CM}$, six hydrogen bonds were formed to the solvent with the loss of one interaction to the protein. One solute−protein interaction is lost to Arg90, as is the interaction with Arg63, while an additional interaction is gained to Glu78. Thus, on average two hydrogen bonds were gained in going from the 1D pmf starting point to NACs$(\xi_{min})_{CM}$ in the protein environment.

Even with the loss of one solute−protein hydrogen bond, the process is found to lower the protein−solute Coulombic interaction energy by −13.8 kcal/mol. Furthermore, the three additional solvent−solute interactions cannot alone account for the −80.2 kcal/mol favorable stabilization in going from the 1D pmf starting point to NACs$(\xi_{min})_{CM}$. This discrepancy between the increase in the number of hydrogen bonds and the computed lowering of the Coulombic energy most likely arises from bringing the two charged carboxylates into closer proximity. By the Born equation,[47] this process should enhance favorable solute−environment interactions of the NACs$(\xi_{min})_{CM}$ to a greater degree than those for the 1D pmf starting point due to the shorter intercarboxylate distance.

Basis of the Rate Enhancement by Chorismate Mutase. As discussed previously, since ΔG_{preq} is negative in both water and CM, there are no contributions to $\Delta G^{\ddagger}$ from preorganization of the substrate as NACs ($\Delta G_{TS} = \Delta G^{\ddagger}$). Thus, the reaction in CM is faster than the reaction in water because CM binds the TS relatively more favorably than the NACs. Increases in internal energy of the NACs from conformational compression and preferential TS stabilization in the enzyme environment relative to water have been suggested to play important roles.[3,21,30] To estimate their contributions to the rate enhancement, the thermodynamic cycle shown in Figure 7 was used. The vertical arrows give the free energies required to convert NACs to the TS in the aqueous and enzymatic environments. Diagonal arrows provide the absolute binding free energies of NACs and TS to CM. As free energy is a state function, $\Delta\Delta G_{bind}$ between the TS and NACs must be equal to $\Delta\Delta G^{\ddagger}$. ΔG_{bind} in going from water to the protein environment for NACs may be decomposed into five steps: conversion of all NACs into NACs(ξ_{min}) in water, NACs$(\xi_{min})_w$, desolvation of NACs$(\xi_{min})_w$, conversion of NACs$(\xi_{min})_w$ to NACs$(\xi_{min})_{CM}$ in the absence of any environmental effects (gas phase), turning on interactions of NACs$(\xi_{min})_{CM}$ in the binding pocket of CM, and, finally, conversion of NACs$(\xi_{min})_{CM}$ into NACs in the enzyme. On the other hand, ΔG_{bind} in going from water to the protein environment for TS is decomposed into only three steps; the first and

(46) A hydrogen bond is required to have an X−H···Y bond length less than 2.5 Å and an X−H···Y angle greater than 120°, where X, Y = N or O.

(47) Rashin, A.; Honig, B. *J. Phys. Chem.* **1985**, *89*, 5588−5593.

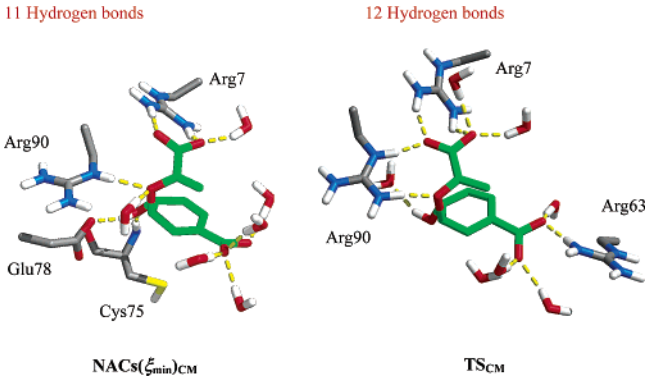
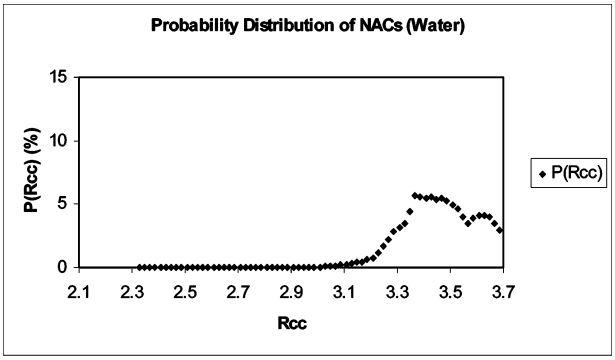


Figure 8. Snapshot of hydrogen-bonding interactions with NACs$(\xi_{min})_{CM}$ and the TS_{CM}, which reproduce the average number of hydrogen bonds. Both structures were the final configuration from MC runs.

the last steps described previously are not necessary (Figure 7). In the thermodynamic cycle, the free energies associated with the desolvation of NACs$(\xi_{min})_w$ and TS_w were estimated by using solvent−solute energy pair distributions (EPD) and hydrogen bond analysis.[16] The free energies associated with the process of turning on interactions of NACs$(\xi_{min})_{CM}$ and TS_{CM} (the transition state in CM) in the binding pocket of CM were estimated by averaging solute−protein/solvent interaction energies (Table 3) and, also, hydrogen bond analysis.

In the previous study for the solvent effects in the chorismate to prephenate rearrangement,[16] it has been shown that the initial desolvation penalty paid by TS_w is greater than that paid by NACs$(\xi_{min})_w$ by at least two solvent−solute interactions. Moreover, by the Born equation, having the two carboxylates closer together leads to a decreased ΔG_{Born} of up to two times that of the infinitely separated case.[47] Transition states in water are better solvated with this mechanism. Thus, in water, TS_w experiences both favorable hydrogen bonds to first shell solvent molecules and a decreased ΔG_{Born}. In the enzyme environment, no gain in the average favorable Lennard−Jones interaction energy was identified between NACs$(\xi_{min})_{CM}$ and TS_{CM} (Table 3). However, an increase in the favorable electrostatic interaction energy of −31.9 kcal/mol, with −10.0 kcal/mol from solute−solvent interactions and −21.9 kcal/mol from solute−protein interactions, was found. On average, there is a gain of one hydrogen bond at the TS_{CM} relative to NACs$(\xi_{min})_{CM}$, as illustrated by the typical structures in Figure 8. While the salt bridge between Arg90 and a carboxylate oxygen atom has been reformed in the TS_{CM}, as has the interaction with Arg63, hydrogen bonds to Glu78 and Cys75 are lost. There is a gain of one solvent−solute interaction in the TS_{CM}. Thus, the overall difference between NACs$(\xi_{min})_{CM}$ and TS_{CM} in CM amounts, on average, to a single solute−solvent interaction. Comparison of NACs$(\xi_{min})_{CM}$ to TS_{CM} in CM indicates that the −10.0 kcal/mol gain in the average solute−solvent interaction energy is more negative than that expected to arise from adding a single solvent−solute interaction. The average solute−protein interaction energy was found to become more favorable by −21.9 kcal/mol, suggesting a more favorable interaction between TS_{CM} and the enzyme occurs even with no change in the number of solute−protein hydrogen bonds. These discrepancies may be accounted for by two effects, the slight change in charges, which leads to stronger hydrogen bonds at the TS_{CM} than those for the NACs$(\xi_{min})_{CM}$, and more favorable electrostatic interactions



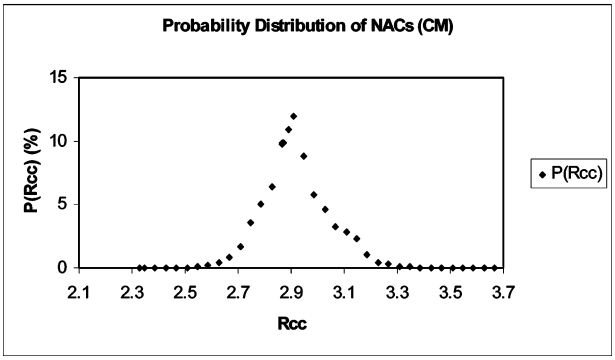


Figure 9. Probability distributions of NACs in (a) water and (b) CM.

with the protein environment, which arise from bringing the charged carboxylate groups into closer proximity.

Using MC/FEP calculations, conversion in the gas phase of the TS identified in the aqueous phase (R_{CO} = 1.86 Å and R_{CC} = 2.12 Å) to the TS in the protein (R_{CO} = 1.86 Å and R_{CC} = 2.20 Å) environment was found to require only +0.9 kcal/mol, as the TS structures are very similar (Figure 7). However, conversion of NACs$(\xi_{min})_w$ (R_{CO} = 1.44 Å and R_{CC} = 3.53 Å) to NACs$(\xi_{min})_{CM}$ (R_{CO} = 1.47 Å and R_{CC} = 2.87 Å) in the gas phase requires +8.8 kcal/mol. This 8.8 kcal/mol energy requirement is mainly a function of the compression of chorismate from an R_{CC} distance of 3.53 Å to 2.87 Å. As the energy required for conformational compression is directly proportional to $\Delta\Delta G_{bind}$ and, consequently, to the overall $\Delta\Delta G^{\ddagger}$, up to −7.9 kcal/mol or ∼70% of the calculated lowering of the free-energy barrier arises from conformational compression of chorismate by CM. As $\Delta\Delta G_{NACs\rightarrow NACs(\xi min)}$ in going from water to CM contributes only −0.2 kcal/mol (Figure 7) out of −11.2 kcal/mol to $\Delta\Delta G_{bind}$, the remaining 3.1 kcal/mol come from the relative desolvation between TS_w and NACs$(\xi_{min})_w$ and from the process of turning on interactions of TS_{CM} and NACs$(\xi_{min})_{CM}$ in the binding pocket of CM. These last results suggest that preferential TS stabilization in the enzyme environment relative to water plays a secondary role in the rate enhancement for the Claisen rearrangement proportioned by CM and strongly support the work of Menger and co-workers.[21]

The compression process is enthalpically controlled. As CM enhances the rate of the chorismate to prephenate rearrangement by compressing NACs when in the binding pocket, it is interesting to see if this process is enthalpically or entropically controlled. Figure 9 plots the probability distributions of NACs in water and CM. It can be viewed that when NACs are bound to CM, the distribution of R_{CC} distances is

narrowed and shifted to smaller R_{CC} values, indicating conformational restriction and compression of NACs, respectively, in the binding pocket. The entropy associated with a particular probability distribution of NACs can be estimated by the following equation:

$$S = -k_B \sum_i P(R_{CC})_i \ln P(R_{CC})_i \quad (1)$$

In eq 1, k_B is the Boltzmann constant and $P(R_{CC})_i$ is the probability of sampling a given R_{CC} bond length less than or equal to 3.70 Å. In this manner, the reduction in entropy associated with restriction of R_{CC} in the binding pocket amounts to −1.1 cal/mol·K, which at 298 K decreases the binding of NACs relative to the TS and makes $\Delta\Delta G_{bind}$ more negative by only 0.3 kcal/mol. One could argue that these entropic contributions do not include contributions from other degrees of freedom of NACs that might affect $\Delta\Delta G_{bind}$ and $\Delta\Delta G^{\ddagger}$ when restricted in CM. However, the experimental $\Delta\Delta S^{\ddagger}$ value[45] between the *Bacillus subtilis* CM and the uncatalyzed reaction in water, which includes also degrees of freedom of water molecules and the protein, is +3.8 cal/mol·K. This gives a contribution of −1.1 kcal/mol to $\Delta\Delta G^{\ddagger}$ at 298 K and, consequently, to the relative binding to CM between the TS and NACs. This indicates that, despite being compressed by CM, NACs still have some flexibility in the binding pocket. Thus, the compression process is under enthalpic control and suggests that the *Bacillus subtilis* CM is able to achieve catalysis without functioning as an entropy trap.

Conclusions

The origin of the catalysis provided by CM for the rearrangement of chorismate to prephenate has been examined by application of the concept of near attack conformations (NACs) in conjunction with MC/FEP calculations. The calculated $\Delta\Delta G^{\ddagger}$ reproduces well the experimental data and supports the applicability of the present methodology in examining this system. It has been shown that as ΔG_{preq} is negative in both water and CM, there are no contributions to $\Delta G^{\ddagger}$ from preorganization of the substrate as NACs. Thermodynamic analyses indicate that conformational compression of the NACs contributes approximately 70% of the calculated lowering of the free-energy barrier by the enzyme over aqueous solution and that this process is enthalpically controlled. This suggests that preferential TS stabilization in the enzyme environment relative to water plays a secondary role in the rate enhancement. The protein merely acts as a scaffold where through artful location of residues the 3×10^6-fold observed increase in the rate constant is generated largely by conformational compression of NACs.

Acknowledgment. Thanks are extended to Dr. Modesto Orozco for valuable discussions. Gratitude is expressed to the National Science Foundation and National Institutes of Health (GM32136) for support of this research. C.R.W.G. acknowledges CNPq/Brazil (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for a postdoctoral fellowship.

JA021424R