JCTC Journal of Chemical Theory and Computation

Effects of Arg90 Neutralization on the Enzyme-Catalyzed Rearrangement of Chorismate to Prephenate

Cristiano Ruch Werneck Guimarães,* Marina Udier-Blagović, Ivan Tubert-Brohman, and William L. Jorgensen*

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

Received March 28, 2005

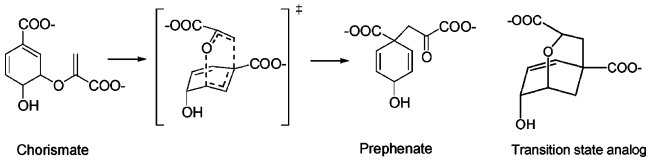
Abstract: Chorismate mutase (CM) is an enzyme that catalyzes the Claisen rearrangement of chorismate to prephenate. In a recent effort to understand the basis for catalysis by CM, Kienhöfer and co-workers (*J. Am. Chem. Soc.* **2003**, *125*, 3206−3207) reported results on the mutation of Arg90 in *Bacillus subtilis* CM (BsCM) to citrulline (Cit), an isosteric but neutral arginine analogue. An ca. 10^4-fold decrease in k_{cat} or 5.9 kcal/mol increase in the free-energy barrier ($\Delta G^{\ddagger}$) for the overall catalysis was observed upon mutation. In this work, attention is turned to determining the key factors that contribute to the reduced catalytic efficiency of Arg90Cit BsCM. Using a combined QM/MM Monte Carlo/Free-Energy Perturbation method, a $\Delta\Delta G^{\ddagger}$ value of 3.3 kcal/mol is obtained. The higher free-energy barrier for the mutant is exclusively related to inferior stabilization of the TS, particularly one of its carboxylate groups, by neutral Cit. In addition, the reaction becomes 2.0 kcal/mol more exergonic. As BsCM is limited by product release, this step contributes to the remainder of the 10^4-fold decrease in the rate constant in going from Arg90 to Cit.

Introduction

Chorismate mutase (CM) is an enzyme that catalyzes the Claisen rearrangement of chorismate to prephenate, a key step in the shikimate pathway for generating aromatic amino acids in plant, fungal, and bacterial systems.[1−3] The reaction, which proceeds via a chairlike transition state (TS) with the C−O cleavage preceding the C−C bond formation (Scheme 1), is a relatively rare example of a chemical transformation that occurs via an identical mechanism in various solvents as well as within the enzymatic environment provided by CM.[4]

This system has been the subject of a variety of experimental[5] and theoretical[6] investigations over the last 20 years. Despite the great attention, the origins of the rate enhancement by CM still remain controversial. Preferential stabilization of the highly polarized TS in the enzyme environment,[7] favorable formation of near attack conformations (NACs)[8] in the active site,[9] and destabilization of the reactant caused by a protein-driven conformational compression[6g,10] have all been presented as possibilities.

Scheme 1. The Chorismate to Prephenate Rearrangement and the Transition State Analogue

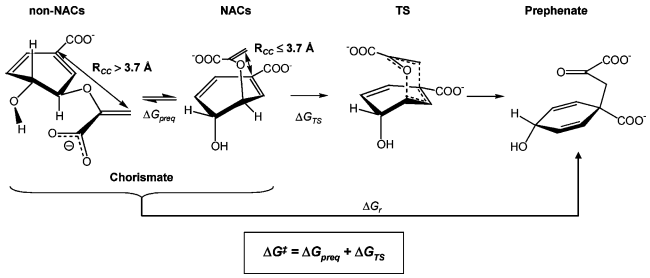


In two recent works,[11] the origins of solvent effects and rate enhancement provided by the *Bacillus subtilis* CM (BsCM) for the rearrangement of chorismate to prephenate were examined using mixed quantum mechanics and molecular mechanics (QM/MM)[12] combined with Monte Carlo[13]/Free-Energy Perturbation[14] (MC/FEP) simulations and the NAC concept.[9] Scheme 2 illustrates the important steps in the NAC-based model for the chorismate to prephenate rearrangement. The first component of the free energy of activation ($\Delta G^{\ddagger}$), the preequilibrium free energy (ΔG_{preq}), is the free energy associated with the process of orienting

* Corresponding author e-mail: cris@ramana.chem.yale.edu (C.R.W.G.) william.jorgensen@yale.edu (W.L.J.).

10.1021/ct0500803 CCC: $30.25
Published on Web 06/04/2005

Scheme 2. NAC-Based Model for the Chorismate to Prephenate Rearrangement



chorismate in appropriate conformations, NACs, that are well disposed to lead directly to the TS of the reaction. This step substitutes the former pseudodiequatorial/pseudodiaxial preequilibrium, which provides a less general and more restrictive definition of a reactive conformer.[15] Chorismate conformers with the forming C−C bond (R_{CC}) less than or equal to 3.70 Å, a distance shorter than the van der Waals distance between vinyl carbons, are considered as NACs. The second component of $\Delta G^{\ddagger}$ is the free energy associated with the conversion of the NACs into the TS (ΔG_{TS}), leading ultimately to the formation of prephenate, and ΔG_r is the free-energy change for the reaction. The QM/MM MC/FEP results suggested that the rate enhancement by CM over the aqueous phase arises primarily from conformational compression of NACs by the enzyme, in agreement with the work of Lyne et al.[6g] and Khanjin et al.[10]

In one of the latest efforts to understand the basis for catalysis by CM, Kienhöfer and co-workers[16] reported results on the mutation of Arg90 in BsCM to citrulline (Cit), an isosteric but neutral arginine analogue. In doing that, the authors aimed at investigating the effects of neutralization of a critical cationic group in the selective stabilization of the TS. An ca. 10^4-fold decrease in k_{cat} or a 5.9 kcal/mol increase in the free-energy barrier for the overall catalysis was observed. A more modest 2.7-fold increase in the K_m value, which corresponds to a 0.6 kcal/mol less favorable free energy of binding for chorismate, indicated only minor perturbation of the ground-state Michaelis complex upon mutation. Therefore, the results show that the binding of the TS to BsCM is considerably affected by the Arg90Cit mutation, becoming 6.5 kcal/mol less favorable. Interestingly, the inhibition constant for Bartlett's transition state analogue (TSA)[17] (Scheme 1) increases only 5.7-fold upon mutation, which corresponds to a 1.1 kcal/mol less favorable free energy of binding; this contrasts dramatically with the 6.5 kcal/mol for the TS. Kienhöfer and co-workers[16] suggest that the TS is more affected by the neutralization of Arg90 due to greater anionic character for the ether oxygen in the TS than in the TSA.

In this work, further attention is turned to determining the key factors that contribute to the reduced catalytic efficiency of Arg90Cit BsCM. To accomplish that, QM/MM MC/FEP simulations were carried out to compute the effects of the Arg90→Cit transformation on ΔG_{preq}, ΔG_{TS}, and ΔG_r. This work also aims at understanding the different decrease in affinity upon mutation for the binding of the TS and TSA to BsCM.

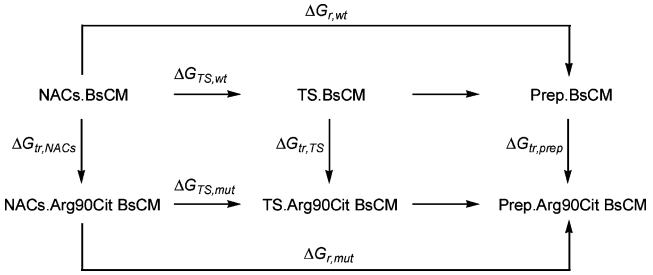


Figure 1. Thermodynamic cycles used for the calculation of relative free energies between the mutant and wild-type BsCMs for the conversion of NACs into the TS ($\Delta\Delta G_{TS}$) and for the reaction ($\Delta\Delta G_r$). $\Delta G_{tr,TS}$, $\Delta G_{tr,NACs}$, and $\Delta G_{tr,prep}$ are the free-energy changes associated with the Arg90→Cit transformation for the TS, NACs, and prephenate bound to the enzyme environments.

Computational Details

NACs vs non-NACs (ΔG_{preq}). Since non-NACs are in equilibrium with NACs, ΔG_{preq} can be estimated by counting the number of NACs and non-NACs sampled in molecular dynamics or Monte Carlo simulations using eq 1, where P_{NACs} and $P_{non-NACs}$ are the probabilities of sampling NACs and non-NACs.

$$\Delta G_{preq} = -RT \ln \frac{P_{NACs}}{P_{non-NACs}} \quad (1)$$

Alternatively, ΔG_{preq} can be derived from a potential of mean force (pmf)[14b] that is computed for the reaction coordinate that converts non-NACs into NACs (R_{CC}) in Scheme 2. The latter approach was followed in this work, and ΔG_{preq} was obtained by computing pmf curves for R_{CC}, $G(R_{CC})$, in BsCM and Arg90Cit BsCM. The lower limit for R_{CC} was taken as the value of R_{CC} in which the ether C−O linkage begins to break in order to form the transition state. The upper limit for R_{CC} was determined as the value of R_{CC} in which the chorismate structure begins to deform from a normal geometry for a ground state. The radial distribution function for R_{CC} ($g(R_{CC})$) in each environment was obtained from $G(R_{CC})$. The probability of sampling R_{CC} ($P(R_{CC})$) was obtained as a product of the respective $g(R_{CC})$ and the volume element of the configuration space corresponding to the coordinate R_{CC}. The volume element can be interpreted as a quantity proportional to the probability of sampling R_{CC} with the potential function set to zero.[14b] In the present case, the volume element is $4\pi R_{CC}^2$.

$\Delta\Delta G_{TS}$, $\Delta\Delta G_r$, and TS vs TSA. Since the free energy is a thermodynamic state function, the cycles shown in Figure 1 give the relative free energies between the mutant and wild-type BsCMs for the conversion of NACs into the TS ($\Delta\Delta G_{TS}$) (eq 2) and for the reaction ($\Delta\Delta G_r$) (eq 3). In eqs 2 and 3, $\Delta G_{TS,mut}$ and $\Delta G_{r,mut}$ are the free-energy changes for the conversion of NACs into the TS and prephenate in the mutant BsCM, while $\Delta G_{TS,wt}$ and $\Delta G_{r,wt}$ are the corresponding quantities for the wild-type BsCM. $\Delta G_{tr,TS}$, $\Delta G_{tr,NACs}$, and $\Delta G_{tr,prep}$ are the free-energy changes associated with the Arg90→Cit transformation for the TS, chorismate NACs, and prephenate bound to the active site of the enzyme, respectively. It is important to note that formally ΔG_r is the

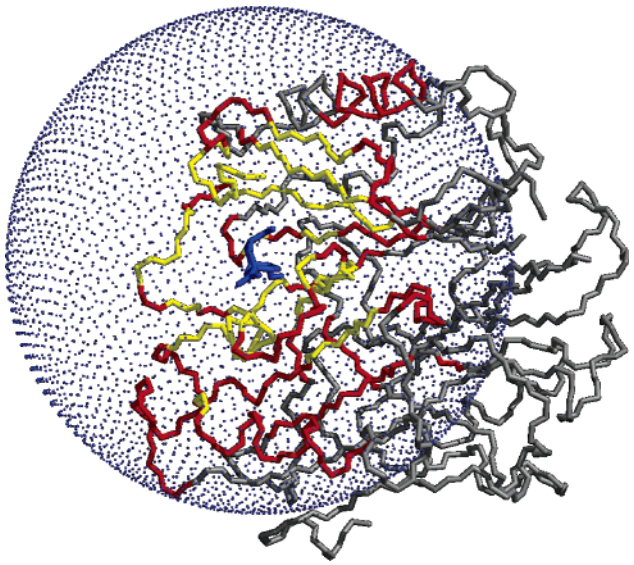


Figure 2. Chorismate mutase with Bartlett's inhibitor (shown in blue). The blue dots represent the 22-Å radius water cap of 766 molecules. Residues with any atom within 15 Å of the center of the reacting system were retained in the simulations (residues in yellow and red).

free-energy change for the conversion of chorismate to prephenate. However, as almost 100% of the chorismate conformers in the active site of both the wild-type and mutant BsCMs are NACs, as described in the Results section, $\Delta\Delta G_r$ can be written as the difference between $\Delta G_{tr,prep}$ and $\Delta G_{tr,NACs}$.

$$\Delta\Delta G_{TS} = \Delta G_{TS,mut} - \Delta G_{TS,wt} = \Delta G_{tr,TS} - \Delta G_{tr,NACs} \quad (2)$$

$$\Delta\Delta G_r = \Delta G_{r,mut} - \Delta G_{r,wt} = \Delta G_{tr,prep} - \Delta G_{tr,NACs} \quad (3)$$

The different decrease in affinity upon mutation for the binding of the TS and TSA to BsCM was also investigated. To accomplish that, QM/MM MC/FEP simulations for the Arg90→Cit transformation with the TSA bound to the active site of the enzyme were also performed.

QM/MM System. Coordinates for the 2.2 Å *Bacillus subtilis* crystal structure complexed with the TSA (2cht from the Brookhaven Protein Data Bank) were employed.[18] Of the four homotrimers of chorismate mutase in the 2cht structure, only the first was retained (Figure 2). A single active site was taken as the center of the reacting system. Residues with any atom within 15 Å of the center of the reacting system were retained in the simulations (residues in yellow and red), 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 substrate or the TSA (shown in blue). In the QM/MM calculations, degrees of freedom for the protein backbone atoms were not sampled. Only the bond angles and dihedral angles for the side chains of residues with any atom within 10 Å from the center of the reacting system were varied (residues in yellow). The substrate and TSA are fully flexible except where specified below. Charge neutrality for the wild-type BsCM was imposed by having a total protein charge of 2 e; charged residues near the active

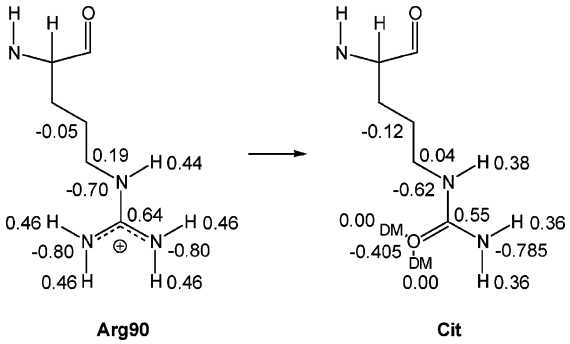


Figure 3. The Arg90→Cit transformation and the change in the charge distribution. DM is a dummy atom.

site were assigned normal protonation states for physiological pH, and the adjustments for neutrality were made to the most distant residues. Each 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. The preparation of the enzymatic systems was much facilitated by use of the chop delegate and MidasPlus 2.1.[19]

Arg90→Cit Transformations. Initial geometries for the Arg90→Cit simulations with the TS, chorismate, and prephenate bound to the BsCM active site were obtained from the previous study with the wild-type enzyme.[11a] The R_{CO} (1.86 Å) and R_{CC} (2.20 Å) distances obtained by the TS search in BsCM were assumed to be the same for the mutant. This is reasonable based on previous findings that showed that the TS geometry in the gas-phase, water, methanol, and in BsCM are essentially the same, being little affected by the environment.[11] The R_{CO} and R_{CC} distances were fixed during the simulations with the TS bound to the active site; however, they were fully sampled for chorismate and prephenate in the active site. The initial geometry for the Arg90→Cit simulations with the TSA bound to BsCM was built by adding two hydrogen atoms to the initial geometry for the TS and by reducing R_{CO} and R_{CC} to the values in the 2cht crystal structure. The R_{CO} and R_{CC} distances were sampled with the TSA bound in the active site.

The Arg90→Cit transformations were performed using the single topology approach by melding the force field parameters for bond angles, torsions, and nonbonded interactions. To keep the number of atoms fixed, two guanidinium hydrogens were mutated to dummy atoms (DM) (Figure 3). The bond angles involving dummy atoms have the same parameters as their counterparts in the initial state. Associated unphysical contributions to the free-energy differences cancel in a thermodynamic cycle.[20] Since the bond lengths for the side chains were not sampled in the QM/MM simulations, they were treated as geometrical variables in the Arg90→Cit transformation; the C−N equilibrium distance in Arg90 was perturbed from 1.34 to 1.23 Å, the C═O equilibrium distance in Cit. The N−H distances were perturbed from 1.01 Å to the value of 0.5 Å for the O−DM distances. Shrinking the bond length for dummy atoms is a common practice to improve convergence in FEP simulations.[20−22] Finally, Figure 3 illustrates how the charge distribution is changed by the transformation.

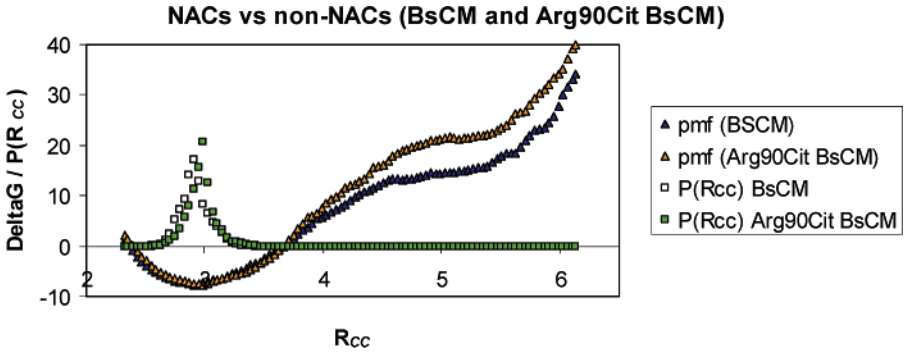


Figure 4. Plots of the one-dimensional potentials of mean force following R_{CC} (the reaction coordinate that converts non-NACs into NACs) and the probability of sampling it in the wild-type and mutant BsCMs.

MC Simulations Protocol. Pmf curves following R_{CC}, computed using increments of 0.02 Å, were obtained via MC statistical mechanics in conjunction with FEP calculations at 25 °C using double-wide sampling. To ensure convergence, each MC/FEP simulation was extensive. Initial reorganization of the solvent was performed for 5×10^6 configurations. This was followed by 30×10^6 configurations of full equilibration and 30×10^6 configurations of averaging for each window. The Arg90→Cit transformations were also performed via MC/FEP calculations at 25 °C using double-wide sampling. The initial and final states were coupled using 20 windows with values for the coupling parameter (λ) evenly distributed between 0 and 1 (0.025, 0.075, 0.125, ..., 0.875, 0.925, 0.975). In this case, initial reorganization of the solvent was performed for 5×10^6 configurations, followed by 30×10^6 configurations of full equilibration and 100×10^6 configurations of averaging for each window.

The combined QM/MM method as implemented in MCPRO 2.1[23] was used to perform all calculations. Established procedures including Metropolis and preferential sampling and the Zwanzig equation were employed.[13,24] Statistical uncertainties were obtained from the batch means procedure with batch sizes of 1×10^6 configurations.[13] The present QM/MM calculations used AM1[25] for the solute intramolecular energy. Computation of the QM energy and atomic charges is required for every attempted move of the QM region, which occurred every 200 configurations. Intermolecular interactions ($E_{QM/MM}$) of the QM solute with the MM region are given by eq 4 as the sum of all interactions between the QM and MM atoms, where q_i and q_j are the partial atomic charges and σ_{ij} and ϵ_{ij} are Lennard-Jones (LJ) parameters.

$$E_{QM/MM} = \sum_{i}^{\text{on QM}} \sum_{j}^{\text{on MM}} [q_i q_j e^2/r_{ij} + 4\epsilon_{ij}(\sigma_{ij}^{12}/r_{ij}^{12} - \sigma_{ij}^{6}/r_{ij}^{6})] \quad (4)$$

Partial atomic charges for the solutes were obtained from the CM1A procedure,[26] while standard Lennard-Jones parameters were assigned for solute atoms [for C, $\sigma = 3.550$ Å, $\epsilon = 0.070$ kcal/mol; for H on C, $\sigma = 2.460$ Å, $\epsilon = 0.030$ kcal/mol; for hydroxyl H, $\sigma = \epsilon = 0.000$; for O, $\sigma = 3.000$ Å, $\epsilon = 0.170$ kcal/mol]. The choice of partial charges is important in the present context, and the AM1-based CM1A charges have been shown to perform well in studies of medium effects on reaction rates and for free energies of hydration.[11,27]

The protein was represented with the OPLS-AA force field,[28] and the TIP4P model was used for water.[29] Residue-based cutoffs of 10 Å were employed. However, as the Arg90→Cit transformation involves a charge mutation, no cutoffs were applied for the mutating residue. In this manner, the missing long-range electrostatic interactions between Arg90 and the residues and solvent molecules outside the finite spherical system would likely cancel when computing double free-energy differences.

Results and Discussion

NACs vs non-NACs. Figure 4 shows the computed pmf curves ($G(R_{CC})$) and the probability of sampling R_{CC} ($P(R_{CC})$) in the mutant and wild-type BsCMs. All chorismate conformations with R_{CC} less than or equal to 3.70 Å are characterized as NACs. It should be noted that ΔG_{preq} only contributes to $\Delta G^{\ddagger}$ in the case of an unfavorable preequilibrium step; the number of non-NACs is greater than the number of NACs. Figure 4 shows that when chorismate is bound to either the mutant or wild-type BsCM, the population of NACs is almost 100%, which makes ΔG_{preq} very favorable, −9.0 kcal/mol in BsCM and −9.3 kcal/mol in Arg90Cit BsCM. In other words, the preequilibrium step is unimportant to the kinetics for both enzymes and thus is not responsible for the 10^4-fold decrease in the rate constant in going from Arg90 to Cit.

Average R_{CC} distances of 3.08 and 3.04 Å were obtained for the complexes between fully flexible chorismate and the wild-type and mutant BsCMs. The average distances agree with the R_{CC} distribution profiles extracted from the pmf curves obtained in both environments and show once again that the chorismate conformers in the active site with either Arg90 or Cit are essentially NACs.

Arg90→Cit Transformations. The Arg90→Cit transformation was performed for the TS, TSA, chorismate (or NACs), and prephenate bound to the active site. Long MC/FEP simulations generated smooth free-energy curves in all cases (Figure 5). The neutralization of the cationic Arg90 is uniformly unfavorable as the solutes have a charge of −2; $\Delta G_{tr,TS}$, $\Delta G_{tr,TSA}$, $\Delta G_{tr,NACs}$, and $\Delta G_{tr,prep}$ were computed to be 68.2 ± 0.1, 70.0 ± 0.1, 64.9 ± 0.1, and 62.9 ± 0.1 kcal/mol, respectively.

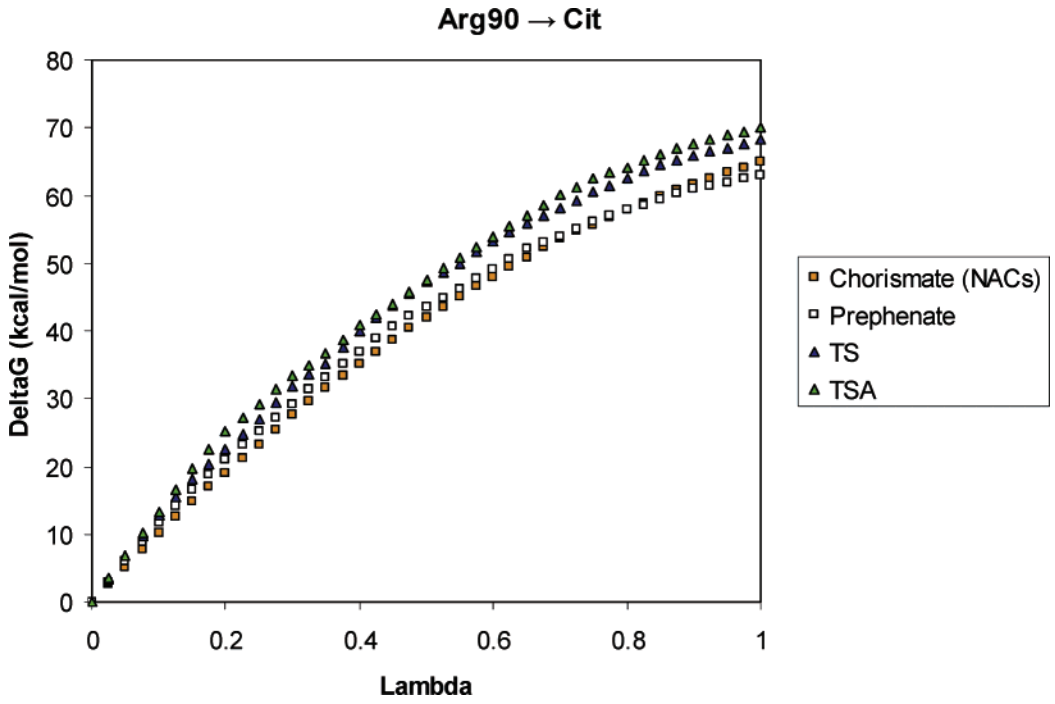


Figure 5. Free-energy curves associated with the Arg90→Cit transformation for the TS, TSA, chorismate, and prephenate bound to the active site. $\Delta G_{tr,TS}$, $\Delta G_{tr,TSA}$, $\Delta G_{tr,NACs}$, and $\Delta G_{tr,prep}$ values of 68.2, 70.0, 64.9, and 62.9 kcal/mol were obtained.

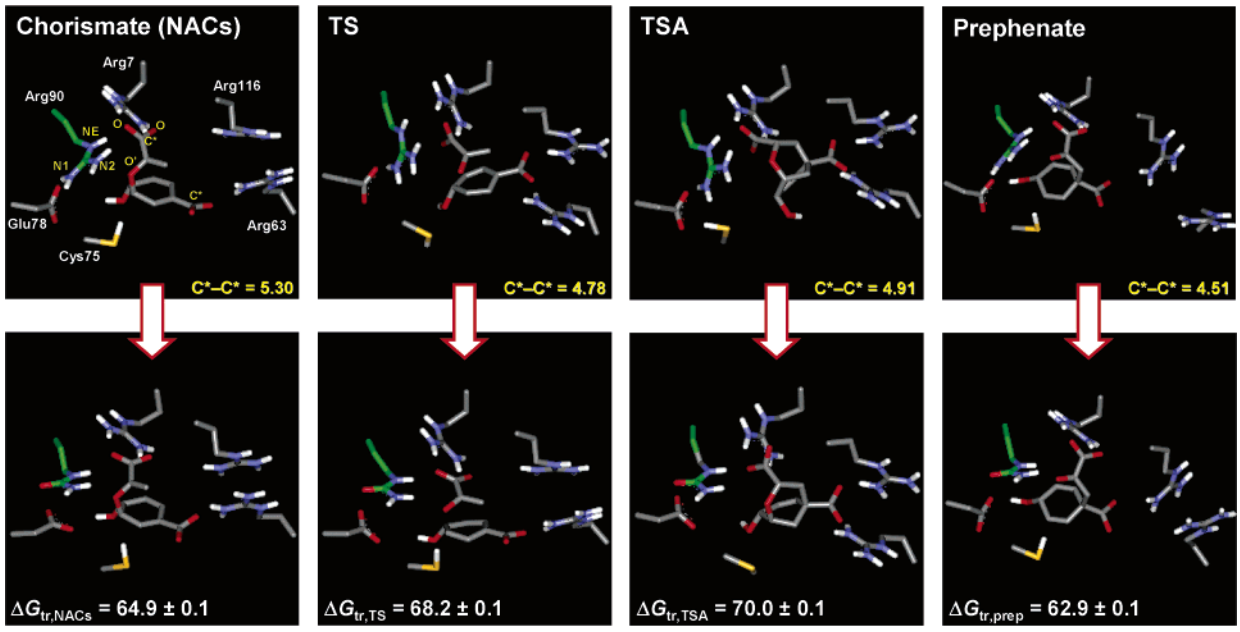


Figure 6. Snapshots illustrating key interactions between each solute and selected residues from the active site. The carbon atoms of Arg90 and Cit are colored green. C*−C* is the average distance between the carbon atoms of the carboxylate groups.

To investigate the effects of the Arg90→Cit transformation for all complexes, the interaction mode between each compound and selected residues from the active site was analyzed (Figure 6). It can be seen that the carboxylate groups of the TS and TSA closely interact with all arginine residues (Arg7, Arg63, Arg90, and Arg116) in the pocket for the wild-type enzyme, whereas chorismate and prephenate interact with only three; chorismate loses the interaction with Arg116 and prephenate with Arg63. The TS and TSA appear to have optimal arrangement of the carboxylate groups to maximize their interactions with the cationic residues. The average distance between the carbon atoms of the carboxylate groups (C*−C*) for the TS and TSA is very similar, 4.78 Å for the TS and 4.91 Å for the TSA. However, the average C*−C* distance for chorismate is too long (5.30 Å), while for prephenate it is too short (4.51 Å).

In the specific case of the interactions with Arg90, Table 1 shows that the average distances between its guanidinium nitrogens (NE, N1, and N2) and the oxygens (O) of the carboxylate group closely interacting with the cationic residue are shorter for the TS and TSA than for chorismate and prephenate. On the other hand, the average distances between the ether oxygen (O′; carbonyl oxygen for prephenate) and the Arg90 guanidinium nitrogens are shorter overall for

Table 1. Average Distances between the Arg90 Guanidinium Nitrogens and Pyruvyl and Ether Oxygens for All Solutes[a]

	chorismate	TS	TS*[b]	TSA	prephenate
O−NE	3.65	3.22	2.94	2.87	3.78
O−N1	5.86	4.15	4.18	4.22	5.94
O−N2	4.56	3.46	3.23	3.16	4.98
O′-NE	3.31	4.75	4.66	4.59	3.03
O′-N1	4.61	4.47	4.73	4.86	4.24
O′-N2	2.96	3.29	3.15	3.12	2.99

[a] Obtained with the Mok program.[30] [b] TS* has R_{CO} and R_{CC} increased to 2.16 and 2.50 Å from the values of 1.86 and 2.20 Å for the TS.

Table 2. Computed and Experimental Free-Energy Changes for the Claisen Rearrangement of Chorismate Catalyzed by the Wild-Type and Mutant BsCMs[a]

	BsCM	Arg90Cit BsCM
ΔG_{preq}	−9.0	−9.3
ΔG_{TS}[b]	26.9 ± 0.5	30.2 ± 0.5
$\Delta G^{\ddagger}_{calc}$	26.9 ± 0.5	30.2 ± 0.5
$\Delta G^{\ddagger}_{exp}$	15.4[c]	21.3[d]
$\Delta\Delta G^{\ddagger}_{calc}$	0.0	3.3 ± 0.1
$\Delta\Delta G^{\ddagger}_{exp}$	0.0	5.9
ΔG_r[b]	−40.7 ± 0.6	−42.7 ± 0.6

[a] Values in kcal/mol. [b] The error in ΔG_{TS} and ΔG_r were calculated by propagating the standard deviation (σ_i) on the individual ΔG_i used to obtain the pmf curve. The equation

$$\sqrt{\sum_{i}^{N}\sigma_i^2}$$

was used, where N is the number of ΔG_i values. [c] Reference 31. [d] Reference 16.

chorismate and prephenate than for the TS and TSA. The latter pattern is at odds with the ΔG_{tr} values. Thus, the results indicate that the ΔG_{tr} values are dominated by the proximity of Arg90 and the carboxylate group; the complex between the TSA and BsCM, which has the closest proximity to the carboxylate group (Table 1), is the most affected by the neutralization followed by the TS, chorismate, and prephenate.

$\Delta\Delta G_{TS}$ and $\Delta\Delta G_r$. As ΔG_{preq} is negative for the wild-type and mutant BsCMs, the preequilibrium step does not contribute to the free energies of activation in both environments; $\Delta G^{\ddagger}$ is equal to ΔG_{TS}. In the prior work,[11a] ΔG_{TS} and, consequently, $\Delta G^{\ddagger}$ were computed to be 26.9 ± 0.5 kcal/mol for BsCM (Table 2). Then, insertion of the computed $\Delta G_{tr,TS}$ and $\Delta G_{tr,NACs}$ values in eq 2 leads to a $\Delta\Delta G_{TS}$ value of 3.3 ± 0.1 kcal/mol. Thus, a $\Delta G^{\ddagger}$ value of 30.2 ± 0.5 kcal/mol is obtained for the mutant. The calculated barriers are larger than the experimental $\Delta G^{\ddagger}$ of 15.4 and 21.3 kcal/mol for BsCM and Arg90Cit BsCM, respectively.[16,31] This difference arises from the previously described deficiency in the AM1 activation barriers.[11b] However, in examining the difference in the free energy of activation between the mutant and wild-type BsCMs, $\Delta\Delta G^{\ddagger}$, errors from the QM method are mostly expected to cancel. Indeed, the calculated $\Delta\Delta G^{\ddagger}$ (3.3 kcal/mol) compares favorably to the experimental value (5.9 kcal/mol).

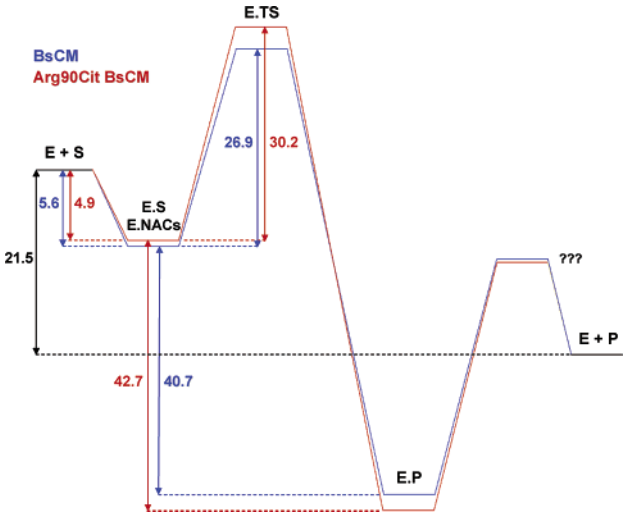


Figure 7. Free-energy diagram for the wild-type and mutant BsCMs. The free energies of binding between chorismate and the enzymes were estimated using the K_m values reported by Kienhöfer and co-workers.[16] The TS for the product release step was assumed not to be affected by mutation (see text for details). The more stable complex between prephenate and Arg90Cit BsCM would increase the barrier and retard product release for the mutant.

Mattei et al. have shown that BsCM is limited by product release at high substrate concentrations.[32] Analysis of the kinetic data indicated that the diffusive transition state for product release is 40% rate-determining in this concentration regime. Calculation of $\Delta\Delta G_r$ (eq 3) shows that the reaction becomes 2.0 ± 0.1 kcal/mol more exergonic with the Arg90→Cit transformation (Table 2), which should make the mutant less efficient in catalyzing the rearrangement due to slower product release. This, obviously, assumes that the TS for the product release step is not affected by mutation. This assumption is supported by the experimental work of Mattei et al.,[32] which indicates that the product release step is controlled by conformational changes observed on the C-terminal tail of BsCM; this region does not contain Arg90. Finally, if the computed $\Delta\Delta G^{\ddagger}$ included not only the chemical step but also the product release step, the calculated value should be more positive than 3.3 kcal/mol. This would bring the experimental and calculated values into closer agreement.

Figure 7 summarizes the free-energy results for the wild-type and mutant BsCMs. The free energies of binding of chorismate to the wild-type and mutant BsCMs were taken from the experimental work of Kienhöfer et al.[16] The calculated ΔG_r value of −21.5 ± 0.6 kcal/mol obtained in water[11b] is significantly more negative than the experimental value of −13.4 ± 2.2 kcal/mol,[5c] while it is close to other theoretical predictions ranging from −20.5 to −22.0 kcal/mol.[7f] However, ΔG_r in BsCM (Table 2) was computed to be more favorable than other theoretical estimations.[6b,7f]

Our previous QM/MM MC/FEP calculations indicated that the rate enhancement by BsCM over the aqueous phase resulted primarily from conformational compression of NACs by the enzyme and that the selective stabilization of the TS in the enzyme environment relative to water played a

secondary role in the catalysis.[11a] More specifically, conformational compression was estimated to lower the free-energy barrier by 7.9 kcal/mol of a total $\Delta\Delta G^{\ddagger}$ of 11.2 kcal/mol, while the remaining 3.3 kcal/mol is due to preferential TS stabilization. This conflicts with other QM/MM calculations which showed significant TS stabilization by the enzyme.[7a,7d] One possible explanation for this is the use of a different thermodynamic analysis to estimate the contributions for the lowering of the barrier in BsCM. While in our work[11a] conformational compression (or substrate preorganization) was defined as unfavorable intramolecular contributions due to geometrical changes of the substrate upon binding to the enzyme, other QM/MM calculations defined substrate preorganization as the free energy cost of forming the same distribution in solution as observed in the enzyme. An alternative source of discrepancy might be associated with energetic deficiencies of AM1 in reproducing the activation barrier.[11b] This could potentially lead to an overestimation of the conformational compression contributions for the binding of NACs to BsCM and consequently for the lowering of the free-energy barrier in BsCM relative to water.

The calculations show that the free energy of activation for the mutant is increased by 3.3 kcal/mol relative to the wild-type BsCM; the slower product release step should contribute to the remaining of the 10^4-fold decrease in the rate constant in going from Arg90 to Cit. Also, the average R_{CC} distance for chorismate changes from 3.5 Å in water[11a] to ca. 3.0 Å in the complex with either Arg90 or Cit in the active site. These results suggest that the mutation completely abolishes the preferential TS stabilization in the enzyme, particularly due to loss of the salt-bridge with the pyruvyl carboxylate group by neutral Cit, while giving similar conformational compression of NACs. In other words, the reaction in the mutant is faster than in water due exclusively to conformational compression of the substrate.

TS Geometry. The R_{CO} (1.86 Å) and R_{CC} (2.20 Å) distances obtained previously by the TS search in BsCM[11a] using AM1 for the solute were found to be more compact than those obtained by higher levels of QM theory.[4a,6a,6c,7b,7f,7g] This is consistent with a previous study which examined the Claisen rearrangements of allyl vinyl ether and allyl phenyl ether, where AM1 transition structures were found to be tighter with greater 1,4-diyl character than various ab initio structures.[33] Thus, the effects of different geometries for the TS on $\Delta G_{tr,TS}$ were examined by perturbing R_{CO} and R_{CC} from 1.86 and 2.20 Å to 2.16 and 2.50 Å, respectively, with increments of 0.05 Å for Arg90 and Cit in the active site. The elongated TS is referred to here as the TS*, and its values for the reaction coordinate were estimated by averaging the R_{CO} and R_{CC} values obtained with higher levels of QM theory.[4a,6a,6c,7b,7f,7g]

Figure 8 shows that the neutralization of Arg90 to Cit becomes increasingly unfavorable as R_{CO} and R_{CC} are elongated; $\Delta G_{tr,TS}$ changes from 68.2 to 70.9 kcal/mol when the TS is transformed to the TS* in the active site. The results in Table 1 once again suggest that the ΔG_{tr} values are largely dominated by the proximity of Arg90 to the pyruvyl carboxylate group; the complexes between the TSA and BsCM and the TS* and BsCM, which have similar values

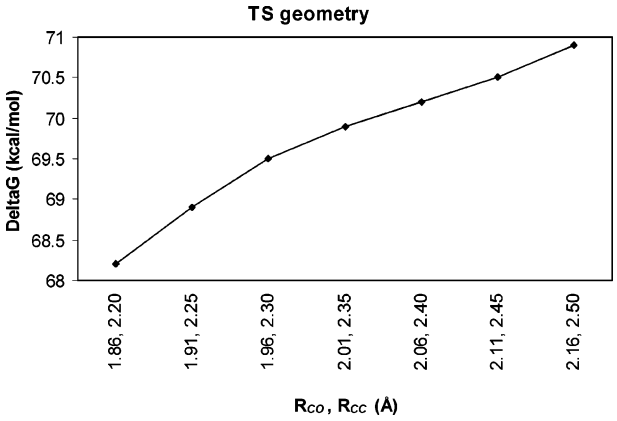


Figure 8. Effects of modification of the TS geometry on $\Delta G_{tr,TS}$ from MC/FEP calculations. The neutralization of Arg90 to Cit becomes increasingly unfavorable as R_{CO} and R_{CC} are elongated.

for the average distances between Arg90 and the carboxylate oxygens (O), are expected to be the most affected by the neutralization followed by the TS. Contributions from the interaction between Arg90 and the ether oxygen (O′) should also be relevant. Figure 9 shows that the average charge distributions for the TSA, TS, and TS* are very similar, except that the latter has more negative charge on the ether oxygen (O′). Finally, if the more dissociated TS* is considered, the increased $\Delta G_{tr,TS}$ value in eq 2 yields a $\Delta\Delta G^{\ddagger}$ value of 6.0 ± 0.1 kcal/mol. This is the same as the experimental value (5.9 kcal/mol) without consideration of the effect of the mutation on product release.

TS vs TSA. The experimental results of Kienhöfer and co-workers[16] show that the binding of the TS to BsCM is significantly affected by the Arg90Cit mutation, becoming 6.5 kcal/mol less favorable. This contrasts dramatically with the effects of mutation on the binding of the TSA, which becomes less favorable by only 1.1 kcal/mol. In other words, the mutation diminishes the binding of the TS to BsCM by 5.4 kcal/mol more than it does for the binding of the TSA. Kienhöfer and co-workers[16] suggest that this is due to greater anionic character for the ether oxygen in the TS. This pattern is not well reproduced here. If the TS and TSA are compared, the TS is less destabilized by the Arg90Cit mutation than the TSA by 1.8 kcal/mol (Figure 5). However, if the TS* is used, it is more destabilized than the TSA but by only 0.9 kcal/mol.

As mentioned above, the ether oxygen is more negative by 0.05 and 0.11 e in the TS and TS* than in the TSA (Figure 9); however, the FEP calculations predict similar effects for the mutation on binding of all three. The more dominant interaction in the computations is between Arg90 and the proximal carboxylate group, which has a similar disposition and charge distribution for the TS, TS*, and TSA (Figures 6 and 9). A possible source of the discrepancy between the computations and the experiment for the TSA vs TS or TS* is the lack of backbone sampling in the calculations. The TSA is more compact and might benefit more from relaxation of the protein backbone upon Arg90 neutralization. Another explanation might be related to the much closer proximity between Arg90 and the pyruvyl carboxylate group of the

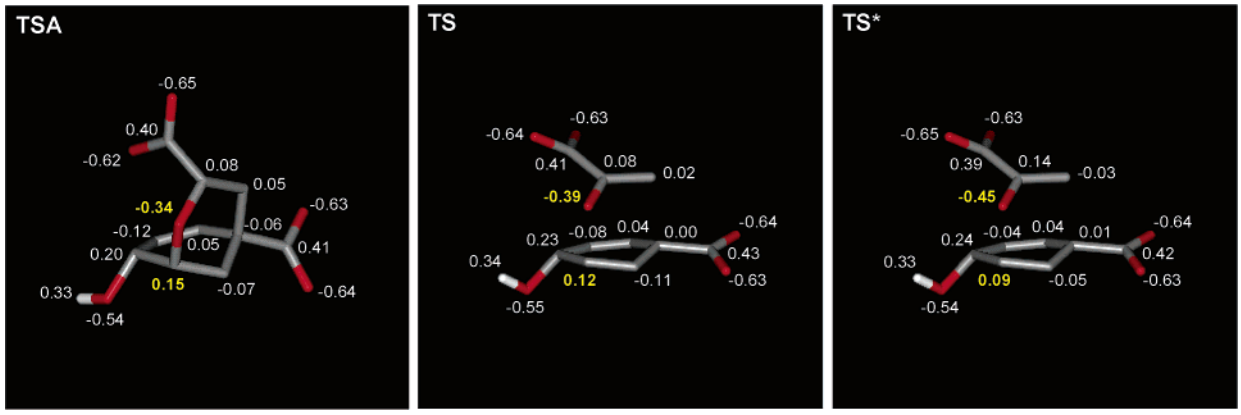


Figure 9. Charge distribution for the TSA, TS, and TS*. The atomic charges with hydrogens were summed onto the carbon atoms.

TSA obtained by the MC simulations. The calculated O−NE, O−N1, and O−N2 distances of 2.87, 4.22, and 3.16 Å (Table 1) compare to experimental values of 4.43, 5.31, and 5.53 Å, obtained by averaging the distances for 12 independent active sites in the BsCM crystal structure complexed with the TSA. However, the experimental distances should be taken with caution as they have considerable noise; the O−NE, O−N1, and O−N2 distances range from 3.84 to 5.66, 3.83 to 6.33, and 4.60 to 7.01 Å with standard deviations of 0.66, 1.02, and 0.75 Å, respectively. This contrasts dramatically with the good agreement between experimental and calculated distances obtained for the complex between prephenate and BsCM. The calculated O−NE, O−N1, and O−N2 distances of 3.78, 5.94, and 4.98 Å (Table 1) compare to experimental values of 3.91, 6.16, and 5.05 Å, obtained by averaging the distances for 9 independent active sites in the BsCM crystal structure complexed with prephenate (1com from the Brookhaven Protein Data Bank).[34] In this case, the experimental distances are much more precise; O−NE, O−N1, and O−N2 range from 3.70 to 4.13, 5.90 to 6.43, and 4.64 to 5.53 Å with standard deviations of 0.14, 0.19, and 0.27 Å, respectively.

Conclusions

Bacillus subtilis chorismate mutase (BsCM) catalyzes the Claisen rearrangement of chorismate to prephenate. In this work, the key factors that contribute to the reduced catalytic efficiency upon mutation of Arg90 to citrulline (Cit), an isosteric but neutral arginine analogue, were investigated. The differential effects on affinity upon Arg90 neutralization for the binding of the transtition state (TS) and the transition state analogue (TSA) to BsCM were also studied. Using a combined QM/MM Monte Carlo/Free-Energy Perturbation method, an increase in the free-energy barrier upon mutation ($\Delta\Delta G^{\ddagger}$) of 3.3 kcal/mol was obtained. The computations have provided extensive structural details, as in Table 1 and Figure 6, and the higher free-energy barrier for the mutant is ascribed to inferior stabilization of the TS, particularly the pyruvyl carboxylate group, by neutral Cit. The calculated $\Delta\Delta G^{\ddagger}$ can be compared to an experimental value of 5.9 kcal/mol for the overall catalysis. However, in addition to the increase in the free-energy barrier, the reaction is computed to be 2.0 kcal/mol more exergonic for the mutant according to the QM/MM simulations. Consideration of the associated retardation of the product release should bring the experimental and calculated values into even closer agreement. Alternatively, it was also shown that if a more dissociative transition state (TS*) is considered, the calculated $\Delta\Delta G^{\ddagger}$ value increases to 6.0 kcal/mol.

Finally, the experimental results of Kienhöfer and coworkers[16] indicate that the mutation diminishes the binding of the TS to BsCM by 5.4 kcal/mol more than it does for the binding of the TSA, and they suggested that this is due to greater anionic character for the ether oxygen in the TS. The charge shift is supported here; however, similar effects of the mutation are predicted for the TS and TSA. Further computational exploration on this point is warranted using alternative QM methodology and allowing relaxation of the protein backbone.

Acknowledgment. Gratitude is expressed to the National Institutes of Health (GM032136) for support.

References

(1) Ganem, B. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 936−945.

(2) Haslam, E. *Shikimic Acid: Metabolism and Metabolites*; John Wiley & Sons: New York, 1993.

(3) Conn, E. E. In *The Shikimic Acid Pathway*; Conn, E. E., Ed.; Recent Advances in Phytochemistry; Plenum Press: New York, 1986.

(4) (a) Wiest, O.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 11628−11639. (b) Wiest, O.; Houk, K. N. *J. Org. Chem.* **1994**, *59*, 7582−7584. (c) Copley, S. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1985**, *107*, 5306−5308. (d) Addadi, L.; Jaffe, E. K.; Knowles, J. R. *Biochemistry* **1983**, *22*, 4494−4501.

(5) (a) Kast, P.; Grisostomi, C.; Chen, I. A.; Li, S.; Krengel, U.; Xue, Y.; Hilvert, D. *J. Biol. Chem.* **2000**, *275*, 36832−36838. (b) Meyer, M. P.; DelMonte, A. J.; Singleton, D. A. *J. Am. Chem. Soc.* **1999**, *121*, 10865−10874. (c) Kast, P.; Tewari, Y. B.; Wiest, O.; Hilvert, D.; Houk, K. N.; Goldberg, R. N. *J. Phys. Chem. B* **1997**, *101*, 10976−10982. (d) Kast, P. M.; Ullah, A.; Hilvert, D. *Tetrahedron Lett.* **1996**, *37*, 2691−2694. (e) Copley, S. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1987**, *109*, 5008−5013.

(6) (a) Ruiz-Pernia, J. J.; Silla, E.; Tunon, I.; Marti, S.; Moliner, V. *J. Phys. Chem. B* **2004**, *108*, 8427−8433. (b) Woodcock, H. L.; Hodoscek, M.; Sherwood, P.; Lee, Y. S.; Schaefer, H. F., III; Brooks, B. R. *Theor. Chem. Acc.* **2003**, *109*, 140−148. (c) Worthington, S. E.; Roitberg, A. E.; Krauss, M. *J. Phys. Chem. B* **2001**, *105*, 7087−7095. (d) Marti, S.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *Theor. Chem. Acc.* **2001**, *105*, 207−212. (e) Marti, S.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J.; Field, M. J. *J. Am. Chem. Soc.* **2001**, *123*, 1709−1712. (f) Marti, S.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *J. Phys. Chem. B* **2000**, *104*, 11308−11315. (g) Lyne, P. D.; Mulholland, A. J.; Richards, W. G. *J. Am. Chem. Soc.* **1995**, *117*, 11345−11350.

(7) (a) Ranaghan, K. E.; Mulholland, A. J. *Chem. Commun.* **2004**, *10*, 1238−1239. (b) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. *Org. Biomol. Chem.* **2004**, *2*, 968−980. (c) Marti, S.; Andres, J.; Moliner, V.; Silla, E.; Tunon, I.; Bertran, J. *J. Am. Chem. Soc.* **2004**, *126*, 311−319. (d) Štrajbl, M.; Shurki, A.; Kato, M.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 10228−10237. (e) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. *Mol. Phys.* **2003**, *101*, 2695−2714. (f) Crespo, A.; Scherlis, D. A.; Marti, M. A.; Ordejon, P.; Roitberg, A. E.; Estrin, D. A. *J. Phys. Chem. B* **2003**, *107*, 13728−13736. (g) Hall, R. J.; Hindle, S. A.; Burton, N. A.; Hillier, I. H. *J. Comput. Chem.* **2000**, *21*, 1433−1441.

(8) NACs possess the proper geometric juxtaposition of groups for a reaction to proceed directly to a transition state without significant bond stretching/contraction, angle bending, or torsional motion.

(9) (a) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 10540−10542. (b) Hur, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12015−12020. (c) Hur, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1176−1181.

(10) Khanjin, N. A.; Snyder, J. P.; Menger, F. M. *J. Am. Chem. Soc.* **1999**, *121*, 11831−11846.

(11) (a) Guimarães, C. R. W.; Repasky, M. P.; Tirado-Rives, J.; Chandrasekhar, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6892−6899. (b) Repasky, M. P.; Guimarães, C. R. W.; Tirado-Rives, J.; Chandrasekhar, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6663−6672.

(12) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.

(13) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*; Clarendon Press: Oxford, U.K., 1987.

(14) (a) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395−2417. (b) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Biochem.* **1989**, *18*, 431−492. (c) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184−189.

(15) Carlson, H. A.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1996**, *118*, 8475−8484.

(16) Kienhöfer, A.; Kast, P.; Hilvert, D. *J. Am. Chem. Soc.* **2003**, *125*, 3206−3207.

(17) Bartlett, P. A.; Nakagawa, Y.; Johnson, C. R.; Reich, S. H.; Luis, A. *J. Org. Chem.* **1988**, *53*, 3195−3210.

(18) Chook, Y. M.; Ke, H.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8600−8603.

(19) (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.

(20) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 103−118.

(21) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532−4545.

(22) Price, D. J.; Jorgensen, W. L. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 681−695.

(23) Jorgensen, W. L.; Tirado-Rives, J. *MCPRO, Version 2.1*, Yale University, New Haven, CT, 2004.

(24) Zwanzig, R. *J. Chem. Phys.* **1954**, *22*(8), 1420−1426.

(25) Dewar, M. J. S., Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(26) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 87−109.

(27) (a) Chandrasekhar, J.; Shariffskul, S.; Jorgensen, W. L. *J. Phys. Chem. B* **2002**, *106*, 8078−8085. (b) Udier-Blagović, M.; Morales de Tirado, P.; Pearlman, S. A.; Jorgensen, W. L. *J. Comput. Chem.* **2004**, *25*, 1322−1332.

(28) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(30) Average values for the distances were obtained with the Mok program (www.perlmol.org/pod/mok.html).

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

(32) Mattei, P.; Kast, P.; Hilvert, D. *Eur. J. Biochem.* **1999**, *261*, 25−32.

(33) Meyer, M. P.; DelMonte, A. J.; Singleton, D. A. *J. Am. Chem. Soc.* **1999**, *121*, 10865−10874.

(34) Chook, Y. M.; Gray, J. V.; Ke, H.; Lipscomb, W. N. *J. Mol. Biol.* **1994**, *240*, 476−500.

CT0500803