A Transition Path Sampling Study of the Reaction Catalyzed by the Enzyme Chorismate Mutase

Ramon Crehuet*,† and Martin J. Field*,‡

Departament de Química Orgànica Biològica Institut de Investigaciones Químiques i Ambientals de Barcelona, Consejo Superior de Investigaciones Cientificas, Jordi Girona 18, 08034 Barcelona, Catalonia, Spain, and Laboratoire de Dynamique Moléculaire Institut de Biologie Structurale—Jean-Pierre Ebel, CEA/CNRS/UJF, UMR 9075, 41 Rue Jules Horowitz, 38027 Grenoble Cedex 1, France

Received: November 16, 2006; In Final Form: February 12, 2007

The study of the chemical steps in enzyme-catalyzed reactions represents a challenge for molecular simulation techniques. One concern is how to calculate paths for the reaction. Common techniques include the definition of a reaction coordinate in terms of a small set of (normally) geometrical variables or the determination of minimum energy paths on the potential energy surface of the reacting system. Both have disadvantages, the former because it presupposes knowledge of which variables are likely to be important for reaction and the latter because it provides a static picture and dynamical effects are ignored. In this paper, we employ the transition path sampling method developed by Chandler and co-workers, which overcomes some of these limitations. The reaction that we have chosen is the chorismate-mutase-catalyzed conversion of chorismate into prephenate, which has become something of a test case for simulation studies of enzyme mechanisms. We generated an ensemble of ∼1000 independent transition paths for the reaction in the enzyme and another ∼500 for the corresponding reaction in solution. A large variety of analyses of these paths was performed, but we have concentrated on characterizing the transition state ensemble, particularly the flexibility of its structures with respect to other ligands of the enzyme and the time evolution of various geometrical and energetic properties as the reaction proceeds. We have also devised an approximate technique for locating transition state structures along the paths.

1. Introduction

A great many computational as well as experimental approaches have been used to explore enzymatic catalytic reactions.[1−7] A particularly fruitful pairing of techniques has been the use of hybrid quantum mechanical (QM)/molecular mechanical (MM) or empirical valence bond (EVB) potentials in conjunction with transition state theory (TST). The former permit the chemical steps of catalysis to be simulated without neglecting any of the enzyme or its environment whereas the latter provides a general framework for the evaluation of the mechanisms, the free-energy profiles, and the rates of reaction.

Many TST studies of reactions employ the following sequence of steps:

(i) Selection of a reaction coordinate (RC) that will describe the progress of the reaction. The RC is most usually a function of a small set of geometrical parameters, such as interatomic distances and angles, but it can also employ other variables such as the energy gap between the reactant and the product electronic potential energy surfaces.[2]

(ii) Evaluation of the free-energy profile for the reaction as a function of the RC using a method such as umbrella sampling.

(iii) Calculation of the TST rate constant by equating the transition state (TS) for the reaction to the value of the RC with the highest free energy.

(iv) correction of the TST rate by computation of dynamical factors which, in classical TST, requires determination of the transmission coefficient.

The above procedure is a powerful one, but it has limitations. One of the most important of these is that it is based upon a predefined RC. Although, in principle, the dynamically corrected TST rate is exact, the form of the RC will affect the free-energy profile for the reaction and will determine how efficiently the transmission coefficient can be obtained (or, indeed, if it can be obtained at all). The choice of the optimum RC is difficult because, in many complex systems, it may be impossible to express the RC as a simple combination of variables, such as internal coordinates.

Because of these problems, other schemes have been developed that do not make a priori judgements about which degrees of freedom will be important in the transition process. One class of these comprise the transition path sampling (TPS) algorithms, developed by Chandler and co-workers,[8−12] which generate unbiased ensembles of short-duration transition path trajectories that pass between reactant and product states. Analysis of the trajectories allows the mechanism of the reaction to be determined and provides a stepping stone for the evaluation of other quantities, such as rate constants.

The work in this paper applies the TPS method to the study of an enyzme reaction. As far as we are aware, there is only one previous study of this type, by Basner and Schwartz,[13] although the focus of the presentation in this paper is different. Our aim has been to generate a transition path ensemble (TPE) and to analyze it geometrically and energetically for information about the reaction. We have been particularly interested by the

* Authors to whom correspondence should be addressed. E-mail: rcsqtc@iiqab.csic.es; mjfield@ibs.fr.

† Institut de Investigacions Químiques i Ambientals de Barcelona.

‡ Laboratoire de Dynamique Moléculaire Institut de Biologie Structurale—Jean-Pierre Ebel.

10.1021/jp067629u CCC: $37.00
Published on Web 05/03/2007

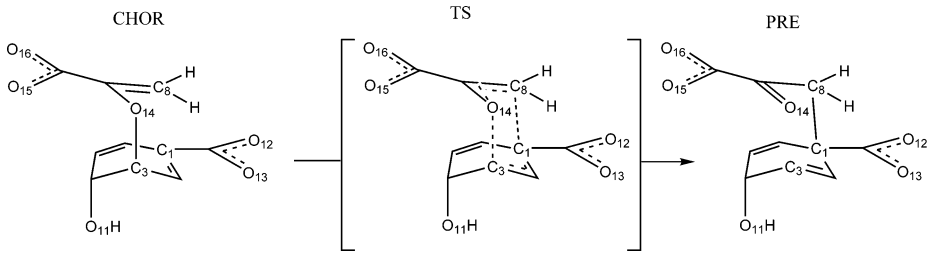

Figure 1. Scheme of the reaction catalyzed by chorismate mutase with the atom numbering employed in the text.

identification of TS structures in the TPE and by a comparison of some of their properties with those of the enzyme complexed with stable species, namely, reactants, products, and transition state analogues. We have not, at present, attempted to calculate rate constants because the TPS scheme of Chandler and co-workers requires expensive supplementary simulations in addition to the generation of the TPE.

The reaction that we have modeled is that catalyzed by the enzyme chorismate mutase (CM). This represents something of a model system for theoretical studies of enzyme reactions and has been investigated by many groups (see, for example, refs 14−29), including our own.[30] It is a relatively simple system (for an enzyme), and computational methods have been shown to give results that are in quantitative agreement with experiment. Nevertheless, the interpretation of these results by different groups remains controversial.

The CM reaction is a Claisen rearrangement and is one of the few examples of a biochemical sigmatropic rearrangement. A scheme of the reaction along with appropriate atom numbering is given in Figure 1. An image of the active site of the enzyme, which is found in a pocket formed by two different chains, is shown in Figure 2. To differentiate residues of different chains, the residues of one chain have been marked with asterisks.

The outline of this paper is as follows. Section 2 describes the CM system and the theoretical techniques that we employed to study it, section 3 discusses the results of the simulations, and section 4 concludes.

2. Methods

2.1. The System. There are four published structures of the *Bacillus subtilis* wild-type chorismate mutase (CM) with Protein Data Bank (PDB) codes 1COM, 2CHS, 2CHT, and 1DBF.[31−33] The earliest structures were 2CHS, resolved at a resolution of 1.9 Å with no bound ligand, and 2CHT with a resolution of 2.2 Å and a transition state analogue in the active site.[31] These were followed by 1COM, also at 2.2 Å resolution, but with bound prephenate[32] and 1DBF with a resolution of 1.3 Å and a glycerol and a sulfate ion in the active site.[33] Most computational studies have employed the structures 1COM and 2CHT as these have substrate or substrate-like ligands bound to the active site. In accordance with previous experience in our group and with other studies, we also started with a structure derived from 1COM. However, we found the structure of Gilliland and co-workers (PDB code 1DBF)[33] valuable during our analyses as it has a higher resolution and resolves some questions arising from the other structures, including the position of Arg63 and the structure of the protein from residue 115 onward. It has the disadvantage, though, that there are disordered residues, one of which, Val73 of chain B, belongs to the active site.

Our simulation system was very similar to one that we used in a previous study, and so only a brief description will be given here.[30] We started from the PDB 1COM structure[31,32] and, of the four homotrimers in the enzyme, retained and solvated only one. We built three different enzyme−ligand complexes, one

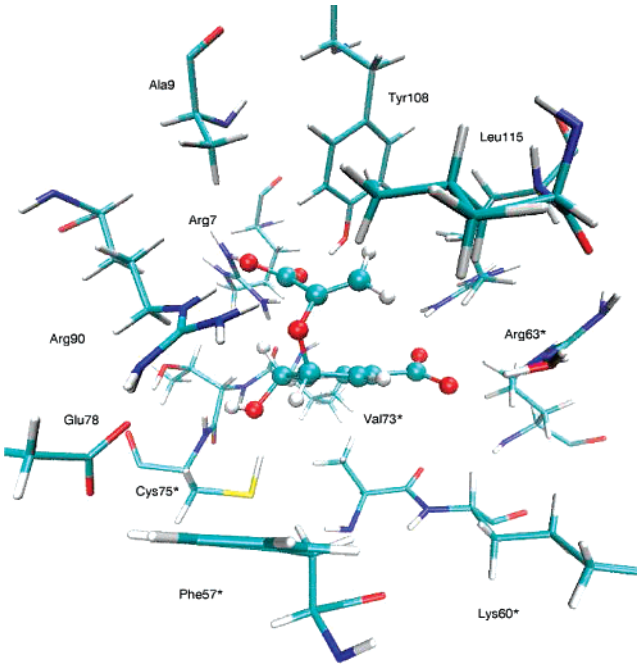

Figure 2. Active site of chorismate mutase. The chorismate substrate is drawn with balls and sticks, and the protein residues are drawn with sticks only. This figure was prepared with VMD.[67]

with the reactant chorismate (CHOR), one with the product prephenate (PRE), and one with the transition state analog (TSA), 8-hydroxy-2-oxa-bicyclo[3.3.1]non-6-ene-3,5-dicarboxylic acid. In all three complexes, the structure of the protein was such that both Arg63 and Arg116* could hydrogen-bond to the ligands.

The simulation system contained 17 183 atoms with CHOR and PRE as ligands and 17 185 atoms with the TSA ligand. In contrast to our previous study, in which all atoms were free to move, we fixed 13 532 atoms in our simulations—equivalent to those that were more than 12 Å away from the substrate. All QM/MM calculations were performed with the DYNAMO library.[34,35] The QM region comprised the substrate whereas the remaining atoms of the protein and solvent were in the MM region and were described with the OPLS-AA force field[36] for the protein and the TIP3P[37] model for water. None of the protein was treated quantum mechanically, and so there were no covalent bonds between the atoms of the QM and the MM regions. Semiempirical QM methods reproduce the correct mechanism and approximate geometry for the chorismate mutase reaction, but they overestimate the energy barrier.[23,27,38] Because our goal was not to match experimental results and we needed large sampling times, we have used the AM1 semiempirical QM method.[39] Noncovalent interactions were truncated with an atom-based force-switching approximation with 8 and 12 Å inner and outer cutoffs, respectively.

For comparison purposes, we also studied the chorismate reaction in water. For this reaction our simulation system had

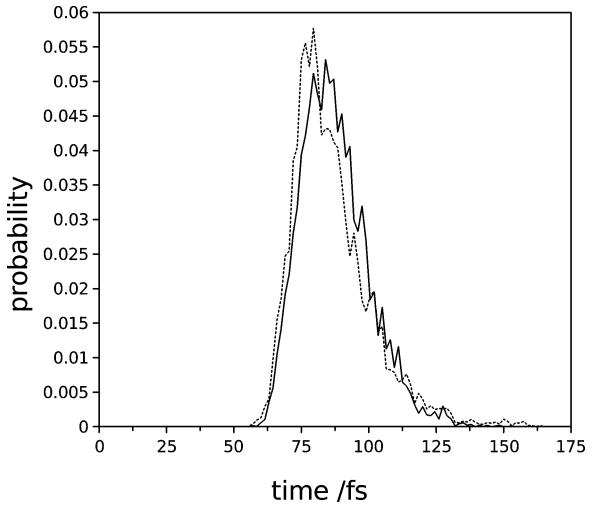

Figure 3. Trajectory transition times for the reaction in the enzyme (solid line) and in water (dashed line). These are calculated as the minimum times required to go between structures in the reactant and product basins.

chorismate or prephenate solvated by 710 water molecules in a periodic cubic box with a dimension of about 28 Å.

2.2. Transition Path Sampling. To generate an ensemble of trajectories for the CHOR to PRE reaction we employed the transition path sampling (TPS) algorithm whose implementation is clearly described in ref 12. The algorithm creates a set of trajectories that start in the reactant region, A, and end in the product region, B. In our case the trajectories were propagated with Newtonian dynamics using the velocity Verlet integrator. This means that the individual trajectories conserve total energy, although the set of trajectories was produced so that it belonged to the canonical (NVT) ensemble. Even though the total energy of the system is conserved, the reacting atoms will experience energy dissipation because they are coupled to the surrounding mobile atoms that act as a thermal bath.

The part of the TPS algorithm used to generate new trajectories from an existing one normally requires randomly displacing the momenta of the mobile atoms in the system. However, due to the large number of atoms, this procedure led to increases in the kinetic energy that were always too large. Consequently we adopted the two-step mechanism, also suggested in ref 12, in which a random change in the norm of the momenta was followed by a random change in direction. In detail, the procedure is as follows: (i) choose a frame from the trajectory at random; (ii) generate a random change in the kinetic energy from a Gaussian distribution; (iii) accept or reject the change using the Boltzmann factor for an NVT ensemble; (iv) change the direction of the momenta without changing their total norm; and (v) integrate the trajectory forward and backward. Overall, the size of the kinetic energy changes in step ii was adjusted so that the acceptance ratio in step iii was about 40%.

The standard TPS algorithm generates trajectories of fixed lengths, and so we had to choose simulation times that were long enough to capture transition events. After some testing we chose 160 fs as this was long enough to describe essentially all reactive events in both the enzyme and water. Figure 3 shows the distributions of transition times for the TPS trajectories that we calculated in water and the enzyme. Their shapes are very similar to those calculated by Zuckerman and Woolf for the isomerization of butane.[40]

For the enzyme, an ensemble of 194 000 trajectories was produced, but only 1 in 200 was saved for analysis to ensure that they were uncorrelated. This gave 970 trajectories in total, each consisting of 110 structures as the integration time step was 0.5 fs and only 1 structure in 3 was kept. For water we adopted a similar procedure and generated, in all, 467 trajectories for analysis.

The TPS algorithm (and analysis of the trajectories) requires a criterion for identifying when a substrate is in reactant (CHOR) or product (PRE) conformations. To enable direct comparison, we employed the same criterion for both water and the enzyme reactions and defined the reactant (product) region as having C3−O14 distances shorter (longer) than 1.7 Å (2.9 Å) and C1−C8 distances longer (shorter) than 2.9 Å (1.8 Å).

2.3. Location of Transition States. The definition of transition states (TS) in the TPS ensemble has been the subject of some debate. Chandler and co-workers define them as those structures that have the same probability of ending at the reactant region as the product region.[10] However, Hummer and co-workers define TS's as "those points in configuration space with the highest probability that equilibrium trajectories passing through them are reactive".[41] We have preferred the former definition in this work and search for TS structures by throwing multiple trajectories, with random initial momenta, from structures along the trajectories in the TPS ensemble. We then identify TS's as those trajectory structures for which half of the thrown trajectories go to reactants and half to products. This is the same procedure as in ref 10.

Determination of TS's in this way is very costly if trajectories are thrown from every structure of every transition trajectory. Instead we adopted the following strategy. In practice, we are only interested in those structures at or close to the TS. Because the distances C1−C8 and C3−O14 define the reactant and product basins, we use them as an indicator of how close we are to the TS region. For every point in every saved trajectory we calculate the distances C1−C8 and C3−O14. If they are close to the reactant or product region, then we do nothing. If they are in the TS region, then we throw 10 trajectories. If all of them go to the same basin, then we assume the probability to be close to 1 or 0. Otherwise, we calculate how many trajectories need to be thrown to obtain an error of ± 0.03 for a binomial distribution. Because the standard deviation of this distribution depends on the value of the probability, the number of attempts is generated self-consistently. Unfortunately the standard deviation is highest at $p = 0.5$, and at that point we generate about 200 trajectories. Each of them is followed until it reaches the reactant or the product basins. We performed the calculation of probabilities, known as committor probabilities, for 60 trajectories from the enzyme TPE but for none of those from the water TPE.

The calculation of committor probabilities is expensive, even with the procedure that we describe above, and so we decided to see if there were alternative ways of identifying TS structures. One possibility is to see if a particular structure is in the geometrical basin of reactants or products, i.e., assuming no dynamical effects. A minimization of the structure is not completely reliable because there is no guarantee that the optimized structure belongs to the same basin as the starting one. A better choice is to calculate a steepest-descent path (SDP) for several steps and then refine the structure by minimization if needed. Transition states were identified as occurring between structures, one of whose SDPs led to the reactant basin, A, and the other's SDP to the product basin, B. To define a unique structure for each TS, the average structure of the two structures bracketing the TS was calculated. This proved to be a reasonable approximation given the similarity of neighboring structures.

2.4. Calculation of *B*-Factors. As part of our analysis we wanted to determine *B*-factors for the CM enzyme when bound to the substrate in its various states, CHOR, TS, and PRE as well as to the TSA. The *B*-factors for the TS states could be calculated with those structures obtained from the TPS calculations, but for the other states this is not possible because the CHOR and PRE structures sampled in the transition trajectories are not fully equilibrated within their respective basins. Thus, we performed molecular dynamics simulations for the three enzyme−ligand complexes (CHOR, PRE, and TSA) in addition to the TPS calculations that we described above. We used a velocity Verlet Langevin dynamics algorithm with a simulation time step of 1 fs and a total simulation time for each species of 5.5 ns. The collision frequency was 100 ps^{-1} and was chosen to ensure that the structures generated by the dynamics decorrelated as quickly as possible. Structures were saved every 600 fs giving ∼9100 structures in total per simulation for later analysis.

The structures from the CHOR, PRE, and TSA molecular dynamics simulations (∼9100 each) and those of the TS from the TPS calculations (∼1000) were employed to determine *B*-factors. For each atom, *i*, they were calculated as[42]

$$B_i = \frac{8\pi^2}{3M}\sum_{i=1}^{M}(\mathbf{r}_i - \bar{\mathbf{r}}_i)^2 \quad (1)$$

where M is the number of structures, $\mathbf{r}_i$ is the position of the ith atom, and $\bar{\mathbf{r}}_i$ the average position, previously calculated as $\bar{\mathbf{r}}_i = (\sum_{i=1}^{M} \mathbf{r}_i)/M$. Superposition of different structures from the dynamics and TPS calculations was not required because of the presence of fixed atoms.

2.5. Other Analyses. We have performed a large number of analyses on the transition trajectories, not all of which we describe here. One important technique is principal component analysis (PCA), which has been used quite extensively to analyze structures coming from molecular dynamics simulations. It forms, for example, the basis of the essential dynamics method.[43] We employed PCA on the TS structures to try to identify characteristics of the ensemble, such as the transition vector. Other approaches that we tried were quasiharmonic analysis,[44] conformational clustering,[45] and the determination of various autocorrelation functions and their power spectra, including the projection of the force on the RC.[46,47] Many of these analyses and others were carried out with the Modular Toolkit for Data Processing.[48]

In a second group of analyses, we investigated the variation of different properties for trajectories from the TPS ensemble after they had been centered at their TS structures. This sort of analysis has been carried out before (see, for example, refs 49−51) but in a context different from that of a TPS study of an enzyme. As part of this analysis, we also wanted to investigate more fully what happens to the transition trajectories when they reach the reactant and product basins and so required trajectories longer than the 160 fs that were in the ensemble. What we did was to select trajectories randomly from the TPS ensemble and extend them backward and forward in time so that there was ∼1.0 ps of dynamics on either side of a trajectory's TS structure. For analysis, trajectories were then aligned so that their TS structures occurred at the same time (i.e., at $t = 0$). A very few selected trajectories had multiple TS structures, usually three but sometimes more, although they were always grouped within a short time period (5−20 fs) of each other. In these cases, we chose the central TS structure (their number is always odd) to put at the time origin. In all we centered and extended 70 randomly chosen trajectories for the enzyme system and 70 for the water system. This number seemed to be sufficient to give converged results for most properties.

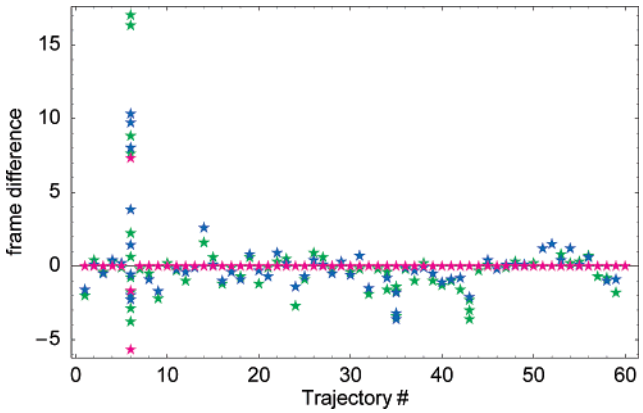

Figure 4. Comparison of TS structures obtained by determination of the committor probabilities in two independent calculations (green and blue) and the steepest-descent method (purple). Exactly 60 trajectories from the enzyme TPS ensemble were analyzed. The abscissa gives the number of the TS (arbitrary) whereas the ordinate gives the number of the structure in the trajectory where the TS is located with respect to the number of the TS structure from the steepest-descent method. For trajectory 6 the steepest-descent method found three TS structures, in which case we place the mean of the three structure numbers at the origin.

3. Results

3.1. Transition State Structures. We employed the method of determining TS structures by committor probabilities on a set of 60 trajectories for the enzyme system. For each trajectory, we performed two independent probability calculations. We compared these structures with those obtained on the same trajectories with our alternative steepest-descent method and found a high correlation between the two sets of results. These are shown in Figure 4. For trajectories 35 and 43, the steepest-descent method found a single transition state whereas at least one of the committor probability calculations found three. This is probably due to statistical imprecision in the probability calculations and would reduce to one TS if more trajectories were thrown from the relevant path structures. A more extreme example of this occurs in trajectory 6 although here the steepest-descent method also finds multiple TS structures (in this case three). Overall, the agreement between the two methods is reasonable, and because the steepest-descent method is much cheaper we could use it to find TS structures for all of the trajectories in the enzyme and water TPEs.

There was usually one transition structure for each trajectory (whichever way they were determined) but in a very few trajectories (∼30) there were more. This may be surprising as recrossings in condensed media are important and other studies with enzymes have obtained transmission coefficients that can be significantly less than 1. The reason for this is that recrossings in these studies take place on reduced dimensionality surfaces as the TS's are defined by means of only one or two RC variables. The larger the dimensionality of the TS surface, the less recrossings there are (see page 262 in ref 52 for a discussion). Our TS uses the full dimensionality, so trajectories very rarely recross the surface dividing the reactant and product basins. It should also be noted that there is no clear relation between the transmission factor and the TPE because the definition of the transmission factor comes from TST and requires a dividing TS surface with a Boltzmann population.

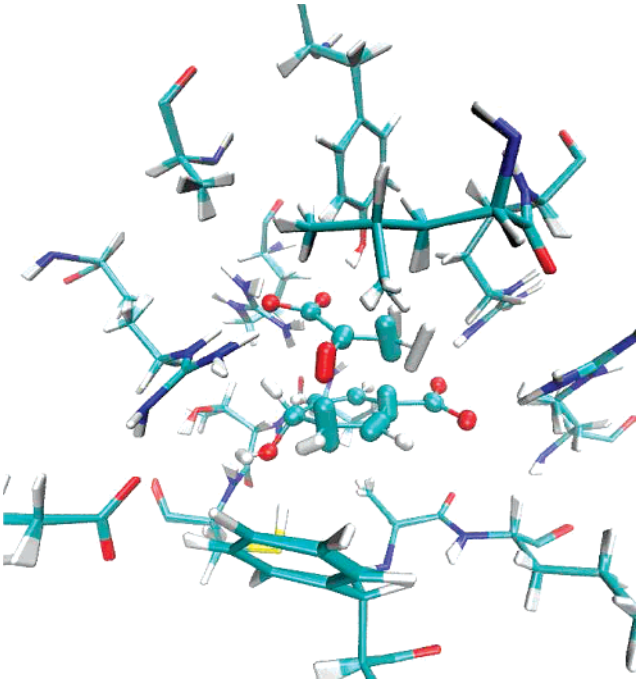

Figure 5. Average structure obtained from the TS ensemble and the transition vector obtained by averaging the TS structures' velocities. The magnitude and direction of the displacements are represented by cylinders for the substrate atoms and by wedges for the atoms of the protein.

Because the ensemble of TS structures that we calculated is generated from a set of trajectories with physical meaning, as opposed to the constrained trajectories needed for free-energy perturbation or integration, we can extract dynamical properties from it. The most relevant one is probably the transition vector. One approximation to this can be obtained by averaging the velocity sets of the TS structures. Figure 5 shows this transition vector from which it can be seen that the substrate atoms are the most mobile and the active site residues are relatively static. We also tried other ways of determining an average transition vector, such as the difference between the structure that is the first to enter the product well and the structure that is the last to leave the reactant well, but the results are almost indistinguishable.

By means of a quasi-harmonic analysis, we can obtain force constants for the C1−C8 and C3−O14 bonds from the variance of their bond-length distributions in the TS structures. The variances of the C1−C8 and C3−O14 distances are 0.006679 and 0.003484 Å^2, respectively, which give force constants of 62 and 119 N m^{-1}. These values are in qualitative agreement with the work of Ruggiero et al.[53] (30 and 210 N m^{-1}), but the C1−C8 value is larger and the C3−O14 value is smaller. We believe though that our analysis is less biased because it does not assume a set of internal coordinates for the RC and because it is based on a quasi-harmonic approximation that is more general than the harmonic description that arises from analysis of a Hessian. The anharmonicity effect is probably also the root of the slightly longer average distances for the C1−C8 and C3−O14 bonds that we obtain, 2.19 and 1.91 Å, respectively, compared to 2.16 and 1.88 Å.[53]

We also analyzed the global properties of the TS ensemble, including a cluster analysis of the TS structures. This showed that there were not distinct classes of TS structures rather that there was a gradation of structures drawn from a single, continuous distribution.

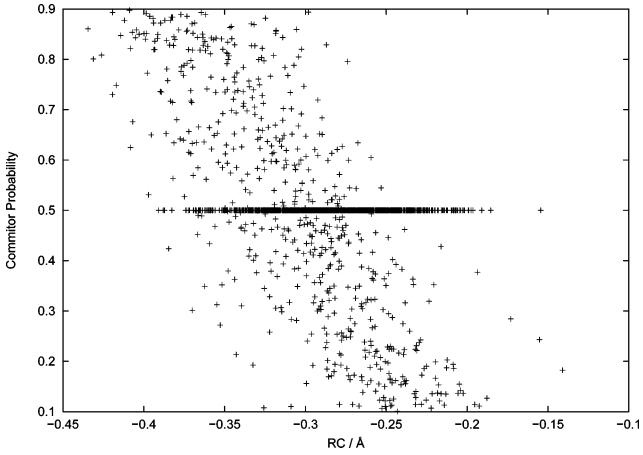

Figure 6. Committor probability versus values of the RC, defined as $s = d_{C1-C8} - d_{C3-O14}$, in Å, for all enzyme−substrate structures for which committor probabilities were determined.

3.2. Determining a Reaction Coordinate. Many computational studies of enzyme (and other) reactions in which free energies are being calculated employ a predefined RC. In principle, the best RC is the committor probability because it indicates directly the likelihood that a structure will go to reactants and to products.[54] In practice, however, it is not much help because it requires a full statistical study of the reaction (as we are doing here) to determine. The question arises then of how well RCs employed in previous studies correlate with the committor probability or, alternatively, whether it is possible to obtain a mapping between the committor probabilities and a (simple) function of geometrical or energetic variables.

A common RC that has been chosen for the CM reaction is the difference of two distances, $s = d_{C1-C8} - d_{C3-O14}$. If this RC provided a perfect description of the reaction, then there would be a direct correspondence between reaction progress and s. In particular, all TS structures (i.e., all those structures with equal probabilities of going to reactants and products) would have identical values of s. To test this, we plot the value of s for all those TPS structures for which we determined committor probabilities in Figure 6. It can be seen that there is a correlation between s and the probability but also that the dispersion is high. We note that the TS as determined from umbrella-sampling free-energy calculations with our enzyme−substrate system occurs when $s \approx -0.27$, but at this value, there are many conformations with probabilites far from 0.5, and likewise, when the probability is 0.5, there are structures with s values in the range from −0.40 to −0.15.

Having shown that the RC, s, is not ideal for discriminating TS structures, we can ask whether a better one exists by analyzing the committor probabilities directly. Our search was unsuccessful, and we summarize our results, many of them negative, here:

•Other definitions of the RC that have been employed in simulation studies of the CM reaction involve dihedral angles in the substrate (see, for example, ref 24). We investigated some of these and found much worse correlation between them and the committor probabilities than for the difference of two distances, s. We show none of this data here.

•The committor probabilities for the reaction in water were not calculated, but an analysis of the s values of the TS structures shows that their range of values is not significantly different from that in the enzyme.

•We considered whether a linear model of the form $\mathbf{p} = \mathbf{AX} + \mathbf{c}$ is able to descibe the committor probabilities, $\mathbf{p}$, where $\mathbf{X}$

are the 3*N* Cartesian coordinates of the *N* atoms in the model. As this method requires the committor probabilities, this limits the number of data available because we only calculated them for a few trajectories. By contrast, our steepest-descent method is a faster way to locate TS structures, but it does not give committor probabilities.

To derive the model, we used PCA in which only atoms in residues closer than 10 Å from the substrate were included. Water molecules were also excluded, because they freely diffuse, as were some hydrogens, notably those of methyl groups. Unfortunately, a robust model was not found in the sense that we could reproduce the committor probabilities of structures employed in the PCA but not those of structures that were excluded. This is due to the limited number of structures for which committor probabilities were available and the fact that they display a huge variance that is (probably) mostly unrelated to the committor probability.

•We also tried correlating the committor probabilities with an RC defined by a TS transition vector. One transition vector was that obtained as the average of the velocities of the TS structures discussed in the previous section. Another was determined by generating a representative minimum energy path for the CM reaction using the nudged elastic band (NEB) method[55] that was recently implemented in the DYNAMO program.[56] We then identified the transition vector as the vector tangent to the path at the structure of highest potential energy. However, it became clear that in neither case could a single vector correctly describe the behavior of the committor probabilities throughout the reaction. Indeed, we think it unlikely that any simple linear model is capable of characterizating the TSE (as a plane) and would require a curved hypersurface.

3.3. Near-Attack Conformations. The proposition of Bruice and co-workers is that the enzyme stabilizes near-attack conformations (NACs) more than water and that this is the origin of catalysis. In CM, these workers define NACs as those structures for which the reacting atoms C1 and C8 are at distances of ~3.7 Å or less (approximately equal to the sum of the atoms' van der Waals radii). They further limit the NAC definition by specifying that the angles between the p-orbitals that are to form the new bond, on atoms C1 and C8, and the C1−C8 vector are within ±20° of their TS values. Denoting these angles as θ_1 and θ_2, respectively, the allowed ranges are 8.2° ± 20° and 18.6° ± 20°.[24,28,57] Bruice and co-workers emphasized that a proper definition of the angles is necessary to exclude conformations that do not represent near-attack ones, as was not done in the work of Jorgensen and co-workers.[17,18]

Figure 7 shows the average C1−C8 distance for the transition trajectories in water and in enzyme. In the enzyme all structures fulfill the distance criterion for NACs. The results in water are very similar although the distances for CHOR-like structures at times of −600 fs and less are slightly larger. It is probable that these distances would increase further if the trajectories were integrated backward for long enough as this would enable the most stable CHOR conformations in water to be sampled. In contrast to the distances, the angles (Figure 8) show larger differences, particularly for θ_1 for which the values at the TS are quite different in water and the enzyme and for which the values before the TS increase steadily in water but remain at about 35° in the enzyme. Unlike the distances most of the transition structures are not NACs according to the angle criteria given in the preceding paragraph.

Conformational compression has also been suggested as a possible source of catalysis by Jorgensen and co-workers.[17,18] To test this we plot in Figure 9 the average distances between

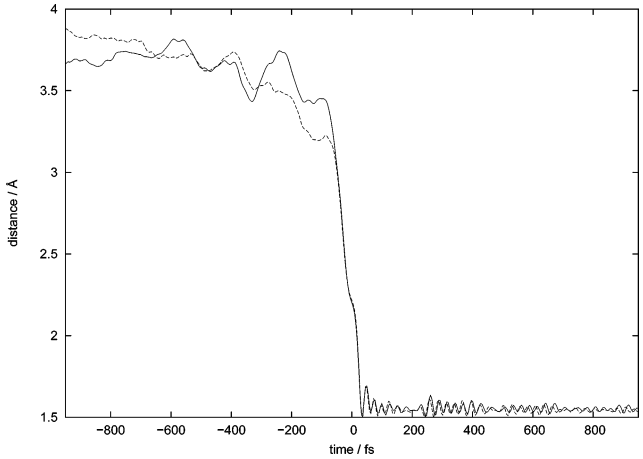

Figure 7. Average distances in Å between the reacting carbons C1−C8 for the transition trajectories. The TS occurs at time zero with reactant (CHOR-like) structures to the left and product (PRE-like) structures to the right: solid line, enzyme; dashed line, water.

C8 and the atom CB of Leu115 and between C3 and the atom CD1 of Phe57 and indeed see that these distances are at a minimum approximately 150 fs before and 50 fs after the TS, respectively. This would indicate that fluctuations in the active site leading to these "minimum distance" conformations are correlated with catalysis. We also tried other ways of defining the average distances for these interactions, such as between C3 and all aromatic carbons of Phe57, and obtained very similar results.

3.4. Energetic Analysis of the Transition Trajectories. A large variety of analyses concerning the energies (and forces) of the trajectory structures are possible, but we show only a few of them here. A simple analysis is to investigate how the total potential energies of the enzyme and water systems change along the trajectories. We do not show them here as they are very similar in the enzyme and in water, except on the reactant side for which the potential energy decreases more in water. This implies that the barrier is higher in water (by reversing the path). An analysis of the individual terms making up the potential energy shows that by far the most important contribution to this difference arises from the sum of the substrate's internal (QM) energy and its electrostatic interaction (QM/MM) energy with the environment.

Figure 10 shows the electrostatic interaction energies between the substrate and the environment along the trajectories. These were calculated by subtracting the energies of QM calculations for the substrate in the environment and in vacuum and so take into account the electrostatic interaction of the substrate with the environment and the polarization that the substrate's charge distribution undergoes in adapting to the environment. The profiles are similar in both environments with the greatest stabilization occurring in the TS region. The largest differences occur for product-like structures that are less stabilized in water. It is perhaps surprising that the electrostatic stabilization energy is of a similar magnitude at the TS for the enzyme and water. However, it should be emphasized that we are looking at transition trajectories for which the environment's structure is adapted for reaction to occur. Our calculations do not permit a rigorous estimate of the reorganization energy required to achieve these reactive environmental configurations, although presumably it would be larger in water than in the enzyme.

The polarization contribution to the total electrostatic interaction energy between substrate and environment can be determined as the difference between the vacuum QM energy of the

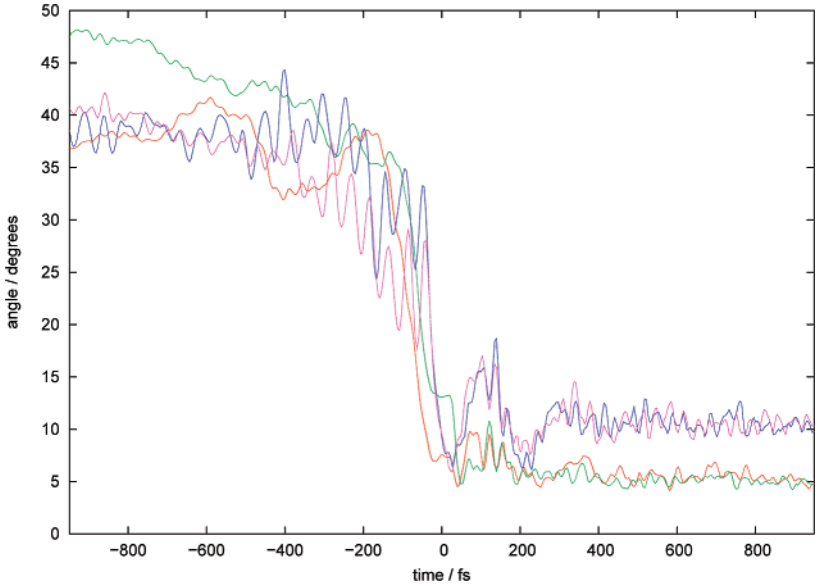

Figure 8. Average attack angles θ_1 and θ_2 in degrees for the transition trajectories: red, θ_1 in the enzyme; green, θ_1 in water; blue, θ_2 in the enzyme; purple, θ_2 in water.

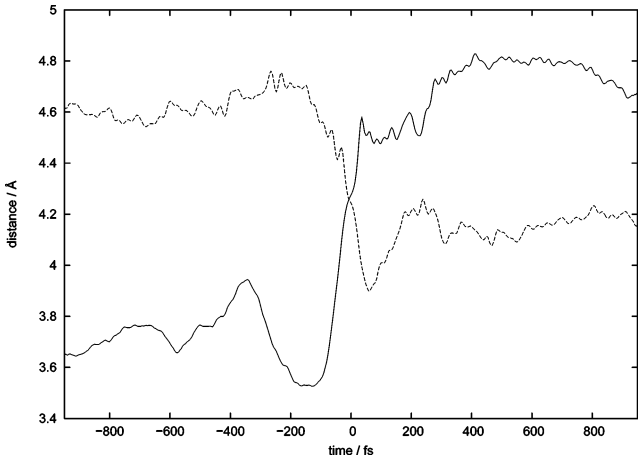

Figure 9. Average C8−Leu116 (solid line) and C3−Phe57 (dashed line) distances in Å for the transition trajectories.

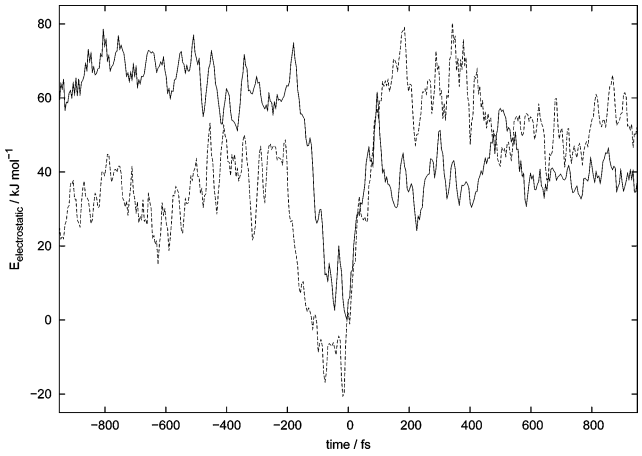

Figure 10. Average substrate−environment electrostatic interaction energies in kJ mol^{-1} for the transition trajectories: solid line, enzyme; dashed line, water.

substrate calculated with the wavefunction determined for it in the presence of the environment and its vacuum QM energy with the vacuum-optimized wavefunction. These energies are shown in Figure 11 and indicate that the substrate is polarized much more in water than in the enzyme and also that the

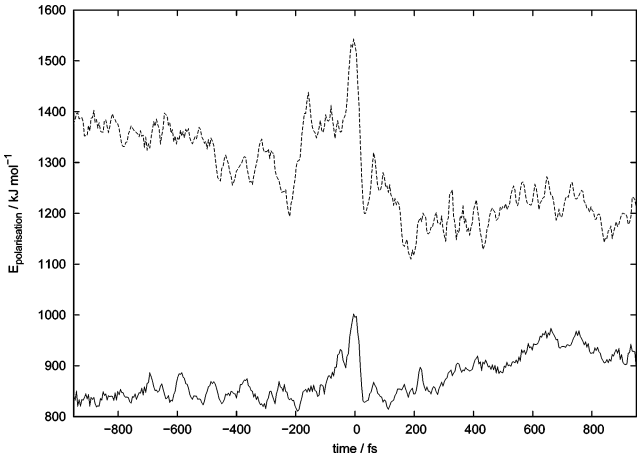

Figure 11. Average substrate polarization energies in kJ mol^{-1} for the transition trajectories: solid line, enzyme; dashed line, water.

variations in the polarization energy are much larger in water. In both environments the polarization energies of the structures in the vicinity of the TS are largest although the energies of the product-like enzyme structures approach those of the TS. In contrast, the reactant-like enzyme structures have polarization energies that vary relatively little whereas those for water increase significantly as they get further from the TS.

We also have the velocities for the structures along the trajectory, so we can examine how the kinetic and total (kinetic plus potential) energies of the substrate change. We note that the trajectories were calculated using Newtonian dynamics, so their total energies are constant. Plots of the kinetic energies in water and the enzyme are very similar, but those for the total energies are more instructive. These are shown in Figure 12, which indicates that the transfer of energy between the substrate and the environment is less in the enzyme than in water. The difference is most marked for reactant-like structures.

3.5. *B*-Factor Calculations. In a recent report, Noonan et al. analyzed the experimental *B*-factors of the enzyme cytidine deaminase and concluded that the TS (or, in their case, a TS analogue) was more tightly bound than reactants and products.[58] This was because the *B*-factors of the enzyme−TSA complex for the ligand and residues in the vicinity of the active site were smaller. To see this qualitatively, consider a harmonic potential

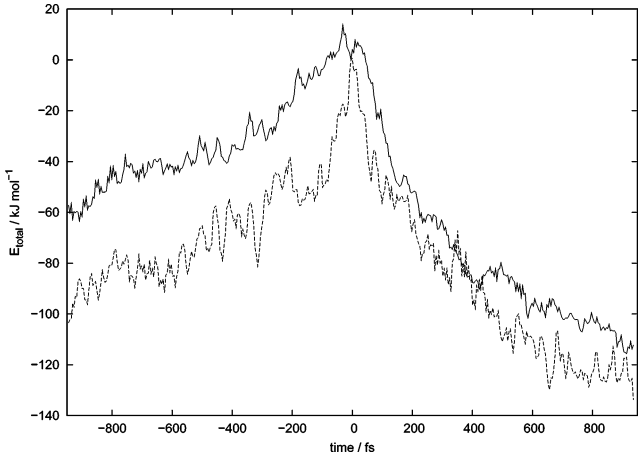

Figure 12. Average total energies (potential plus kinetic) of the substrate in kJ mol^{-1} for the transition trajectories: solid line, enzyme; dashed line, water.

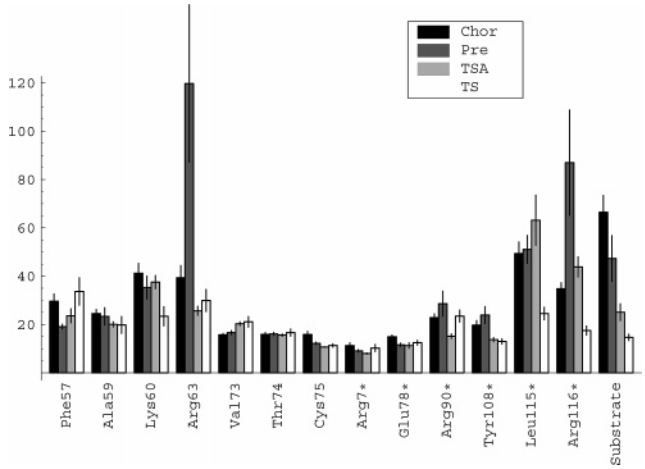

Figure 13. Calculated *B*-factors and error bars for the substrate and active site residues from the CHOR, PRE, TS, and TSA simulations. The error bars were determined as the variance of the block-average mean using data block sizes of 2^{10} points.

with force constant, *k*, for which the *B*-factor is proportional to *T*/*k*, where *T* is the absolute temperature. Thus, smaller *B*-factors mean higher force constants and tighter binding. This is a potentially important result because little is known experimentally about TS binding. However, there are two caveats. First, the conclusions were based on a TSA, so it is unknown whether they apply to the true TS structures (which, of course, are inaccessible experimentally), and second, it is unknown whether the result holds for other enzymes.

If low *B*-factors (tight-binding) are correlated with a preferential stabilization of TS structures, then this would manifest itself in a larger value of the catalytic rate constant, k_{cat}, for the reaction. By contrast, low *B*-factors for reactant and product structures would have a tendency to manifest themselves in smaller values of the Michealis−Menten constants, K_M, for the forward and reverse reactions.

We first analyzed the *B*-factors for the four crystallographic CM structures. It should be emphasized that this is a delicate matter because experimental *B*-factors contain contributions from factors other than thermal motion, such as static disorder, and that *B*-factors from different structures are not directly comparable as they depend upon such things as experimental conditions and structure resolution. To mitigate these effects, we determined the ratio of the average *B*-factor for selected active site residues (Phe57, Ala59, Lys60, Arg63, Val73, Thr74, Cys75, Arg7*, Glu78*, Arg90*, Tyr108*, and Leu115*) divided by the average *B*-factor for all residues of a given structure. We assume by doing this that many of the effects that are specific to a particular structure and that are included in the *B*-factors will cancel out. We find the following values: 1.14 for 2CHS which has no bound ligand; 0.7 for 1DBF which has a glycerol and a sulfate ion; 1.01 for 1COM which has PRE; and 0.96 for 2CHT which has a TSA. With the exception of 1DBF, this is the trend that would be expected if the mobility of an active site is reduced more when the ligands it binds more closely resemble TS structures. We should emphasize that the 1DBF results are probably not comparable, in any case, with the others because the 1DBF structure was obtained by a different group than the others using different crystallographic conditions and models.

To calculate the *B*-factors for the different ligands, we employed structures from the TS ensemble and from the molecular dynamics simulations. Analysis of the latter showed that the binding of PRE and TSA in the active site was stable throughout the dynamics. In the CHOR simulation, by contrast, at ∼3.0 ns, CHOR and Arg116* separate, and the hydrogen bonds between them break. The CHOR conformation also changes as the C1−C8 distance (where the bond is to form) increases from ∼3 Å to more than 4 Å by the end. The loss of interaction between CHOR and Arg116* was also reported by Guimaraes et al.[29] They used a similar QM/MM potential to ours but a different sampling scheme (Monte Carlo) and starting crystal structure (2CHT). However, this change in conformation was not observed in another study with a similar QM/MM potential.[26]

Figure 13 shows the calculated *B*-factors of the ligands and various active site residues from the CHOR, PRE, TS, and TSA simulations. The following residues have been considered important in catalysis because they are close to and interact with the substrate: Arg63, Cys75, Arg7*, Arg90*, Tyr108*, and Arg116*. Glu78* is also needed for catalysis because of its interaction with Arg90*. For the ligand, CHOR is the most mobile, then PRE, then TSA, and finally TS. This is the trend that should be observed if the hypothesis that TS-like ligand structures bind more tightly to the enzyme is correct. As the enzyme is presumably optimized to bind the TS, the higher mobility of the TSA is to be expected although this is still less than the mobilities of either the reactant or product forms of the substrate.

For the protein, the picture is more complicated. For one set of residues, comprising Ala59, Val73, Thr74, Cys75, Arg7*, and Glu 78*, the *B*-factors do not change significantly with the ligand. Then there are the residues, Phe57, Lys60, Arg90*, and Tyr108*, whose *B*-factors show some variation and, finally, Arg63, Leu115*, and Arg116*, whose *B*-factors change markedly with ligand.

It is interesting to correlate the calculated *B*-factors with interactions, particularly hydrogen bonds, at the active site. The hydrogen-bond distance distributions between the ligand and pertinent active site residues are illustrated in Figure 14. For certain interactions, the identity of the donor and acceptor atoms change during the course of the simulation (by, for example, swapping between the oxygens of a carboxylate group). In these cases, the distributions are calculated using the best hydrogen-bond interaction that is available in a given frame.

Lys60. This residue has no specific interactions with the ligand, so the fact that it has a lower *B*-factor in the TS structure could possibly indicate a more static active site when the enzyme is complexed with the TS.

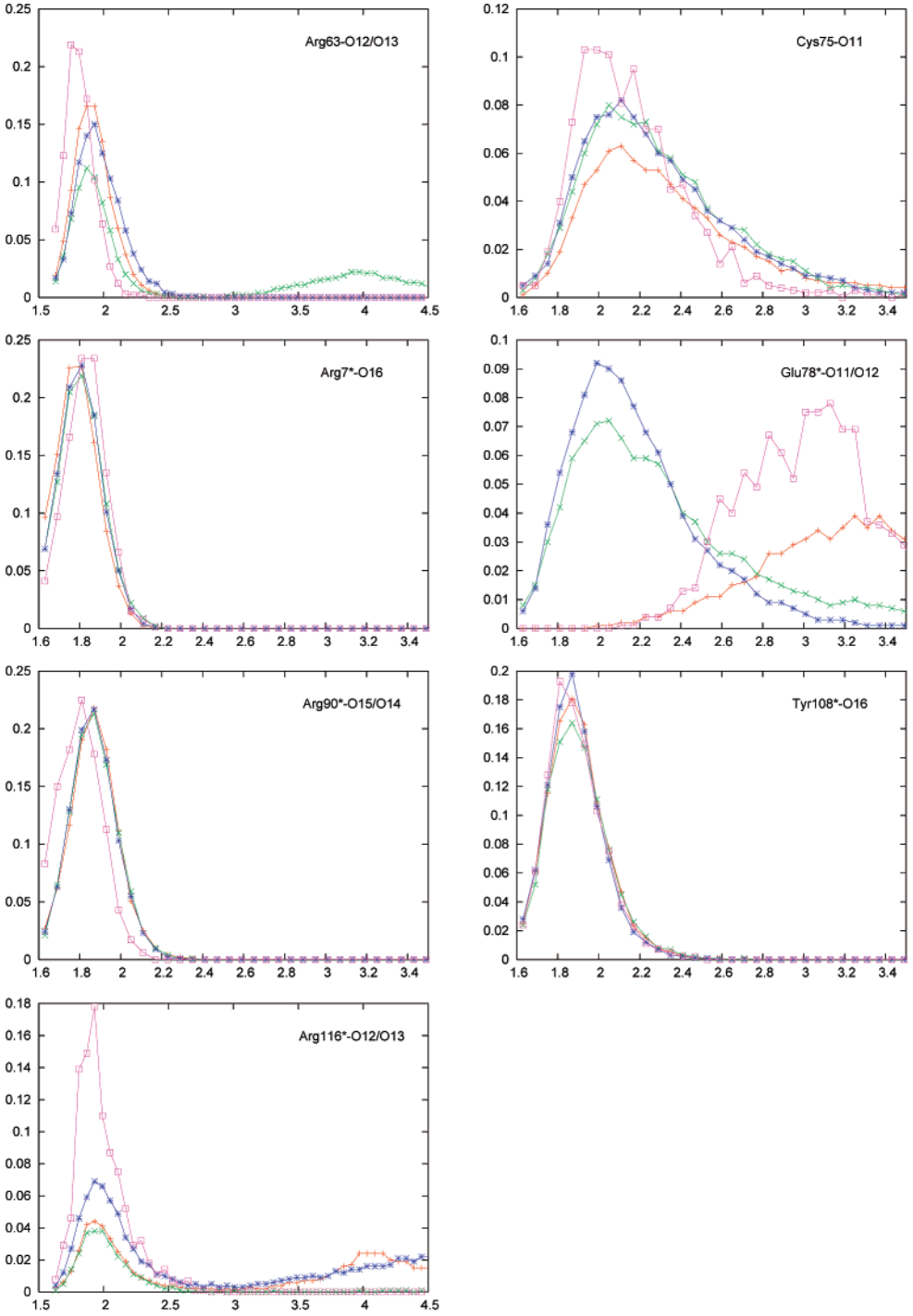

Figure 14. Hydrogen-bond-length distributions for certain ligand−active site residue interactions: CHOR, red; PRE, green; TSA, blue; TS, magenta.

Arg63. Experimentally, this residue has a high mobility. It is one of those that is most strongly affected upon binding[59] and has unreliable or different conformations in the crystallographic structures.[32,33]

There is correlation between the calculated *B*-factor and the hydrogen-bond distribution for the PRE complex for which the *B*-factor is very high and the distribution very broad. By contrast, the *B*-factors for the other residues are similar whereas the hydrogen-bond distribution for the TS complex is sharply peaked, which indicates that this interaction is important for TS binding.

Cys75 has very similar *B*-factors for all ligands although those for the TS and TSA are slightly smaller. This correlates reasonably with the hydrogen-bond distributions, which are most peaked for the TS structure.

Arg7.* Brooks et al.[16] and Mulholland et al.[27] state that this residue stabilizes the TS although this is difficult to ascertain from the *B*-factors and hydrogen-bond distributions, which are very similar for all ligands.

Glu78.* This residue has been identified as important experimentally,[60] but our results, like those of ref 27, find that Glu78* hydrogen-bonds to the substrate for product-like conformations only. The distribution tends toward shorter distances in TS than in CHOR structures, but in neither case can a hydrogen bond be said to exist nor a single conformation describe the TS or the reactant ensemble. For this residue, the

TSA behaves like the product and not like the TS. Likewise, the *B*-factors are similar for all ligand forms and so do not correlate with the changes in this important interaction.

*Arg90**. This residue is the one whose role in catalysis is most debated. It has been claimed that it should bind the TS best because the TS bears the highest degree of negative charge on the carbonyl−ether oxygen (O14) during the reaction.[60] This has also been observed in simulation studies.[26,27,29] According to Lipscomb and co-workers, Arg 90* binds the carbonyl group of PRE (O15) and the ether oxygen of TSA (O14) and, in the TS complex, should contribute more to substrate specificity than to catalysis.[31] By contrast, spectroscopic evidence points to a weak interaction between O14 and Arg90* for the PRE complex.[61]

Overall our simulations do not contradict the finding that Arg90* is crucial for catalysis.[60] The hydrogen-bond distribution is most peaked for the TS, although the effect is not large, and is almost identical for the remaining ligand forms. The *B*-factors are similar for all ligand forms although it is lowest for the TSA, something which cannot be deduced from the hydrogen-bonding distributions.

*Tyr108**. The hydrogen-bond distributions for this residue are very similar for all ligand forms whereas its *B*-factors are slightly lower in the TS and TSA complexes. Our results are in agreement with the NMR experiments of ref 59 and the simulation results of ref 26, which have Tyr108* binding substrate, but not with those of ref 27 in which it points to a water molecule. As the hydrogen-bond interaction is very similar for all complexes, it is uncertain whether it contributes significantly to catalysis, in contrast to what was found in ref 62.

*Arg116**. This residue belongs to a mobile tail (residues 110 upward) whose position varies in the available crystal structures. In a few sites, there is a stabilizing interaction between Arg116* and PRE or TSA, but in others there is no observable interaction, or the residue is oriented toward the solvent. It has been proposed that the crystallographic positions of this residue are strongly influenced by the crystallographic packing.[33] Experiments indicate that movement of the tail plays an important role for the entrance and exit of molecules to the active site in CM,[59,61,63] but kinetic measurements with mutated enzymes have been interpreted as showing that it is not essential for catalysis.[64] Nevertheless, Leu115* is highly conserved (mutations to other large and hydrophobic residues are the only ones that preserve activity), and a preference for the wild-type Arg116* has also been observed.

The *B*-factors for the TS complex and its very peaked hydrogen-bond distribution agree for this residue. Interestingly, the TSA distribution is also quite peaked, but its *B*-factors are higher than those for the CHOR complex. For the CHOR and PRE complexes, a hydrogen bond (at ∼2.0 Å) is observed in some of the structures, but there are many structures in which it is absent.

4. Conclusions

In this paper we have applied the TPS method of Chandler and co-workers[8−12] to study the chorismate to prephenate reaction catalyzed by the enzyme chorismate mutase and in water. We generated ∼1000 uncorrelated transition paths in the enzyme and ∼500 in water for analysis. The most salient conclusions of our work are as follows.

•Using the method proposed in ref 10 for calculating committor probabilities, we determined a set of TS structures from 60 randomly picked transition paths that we generated for the reaction in the enzyme. Although this set is limited in size, it is representative of the TSE for the enzyme-catalyzed reaction and provides information that is inaccessible experimentally.

•Due to the expense of locating TS structures by determining committor probabilities, we devised a faster, energy-minimization-based method, which although approximate gives TS structures that agree well with those from the more exact approach. The new method enabled us to find TS structures for all of the transition paths that we saved for analysis (∼1000 in the enzyme and ∼500 in water).

•We were unable to find any geometrical RC that mirrored the committor probabilities for the transition path structures that we calculated. One commonly used RC, the difference of two distances ($s = d_{C1-C8} - d_{C3-O14}$), showed some correlation, whereas others, such as those based upon dihedral angles, were much worse.

We should note that although the committor probability is, in principle, the exact (non-geometrical) RC, its value is significant only in regions close to the TS and decreases very rapidly away from it. This makes it impractical to employ the committor probability for distinguishing structures far from the TS.

•Many of the structures in the TSE are not NACs as defined by Bruice and co-workers.[24]

•An analysis of the transition paths indicates that there is conformational compression during the reaction in the enyzme.

•Electrostatic stabilization is significantly greater for TS structures and is similar in water and in the enzyme for our transition paths. Presumably, however, the reorganization energy needed to generate reactive conformations is greater in water than in the enzyme, but we did not determine this.

•Analysis of *B*-factors calculated for enyzme−ligands complexes shows that the ligand mobility increases in the order TS < TSA < reactant ∼ product. This agrees with recent experimental work of Noonan et al.[58] who observed similar effects for the enzyme cytidine deaminase. The picture for residues in the active site of the enzyme is more complicated as their mobilities are not always smallest for the TS and TSA ligand forms.

•An analysis of the simulated ligand−residue hydrogen-bond distributions gave useful insights into interactions that were likely to be important in binding and catalysis but also indicated that there was not necessarily any correlation between the presence or absence of a hydrogen-bond interaction and residue *B*-factor.

To finish, we note that there is a lot of information to be gleaned from the TPE and that many possibilities exist for extending the work that we have pursued here. In the future, we intend to develop methods for obtaining average paths and for determining concise descriptions of RCs that span the full reaction from reactants to products. We will also investigate the relation of TPS to other statistical methods, such as the string method,[65,66] and the possiblity of obtaining free energies from the generated ensembles.

Acknowledgment. R.C. thanks X. Carpena and A. Manich for fruitful discussions, the Spanish Ramón y Cajal Program, the DGYCIT (Grant No. CTQ2006-01345/BQU), and the Generalitat de Catalunya (Grant No. 2005SGR00111) for current support, and the Marie Curie Program of the European Union for a postdoctoral fellowship. The authors also acknowledge the support of the Institut de Biologie Structurale−Jean-Pierre Ebel, the Commissariat à l'Energie Atomique, and the Centre National de la Recherche Scientifique in France.

References and Notes

(1) Benkovic, S. J.; Hammes-Schiffer, S. *Science* **2003**, *301*, 1196−1202.

(2) Warshel, A. *Acc. Chem. Res.* **2002**, *35* (6), 385−395.

(3) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106* (8), 3210−3235.

(4) Bruice, T. C. *Chem. Rev.* **2006**, *106* (8), 3119−3139.

(5) Antoniou, D.; Basner, J.; nez, S. N.; Schwartz, S. D. *Chem. Rev.* **2006**, *106* (8), 3170−3187.

(6) Garcia-Viloca, M.; Gao, J.; M, M. K.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.

(7) Gao, J.; Ma, S.; Major, D. T.; Nam, K.; Pu, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106* (8), 3188−3209.

(8) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998**, *108* (5), 1964−1977.

(9) Geissler, P. L.; Dellago, C.; Chandler, D. *J. Phys. Chem. B* **1999**, *103*, 3706−3710.

(10) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Nat. Acad. Sci. U.S.A.* **2000**, *97* (11), 5877−5882.

(11) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.

(12) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. *Adv. Chem. Phys.* **2002**, *123*, 1−86.

(13) Basner, J. E.; Schwartz, S. D. *J. Am. Chem. Soc.* **2005**, *127*, 13822−13831.

(14) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (16), 9032−9037.

(15) Hur, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (3), 1176−1181.

(16) Lee, Y. S.; Worthington, S. E.; Krauss, M.; Brooks, B. R. *J. Phys. Chem B* **2002**, *106*, 12059−12065.

(17) Guimaraes, C. R. W.; Repasky, M. P.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6892−6899.

(18) Repasky, M. P.; Guimaraes, C. R. W.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2003**, *125*, 6663−6672.

(19) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 10540−10542.

(20) Ranaghan, K. A.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. *Mol. Phys.* **2003**, *101* (17), 2695−2714.

(21) Štrajbl, M.; Shurki, A.; Kato, M.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 10228−10237.

(22) Guo, H.; Cui, Q.; Lipscomb, W. N.; Karplus, M. *Angew. Chem., Int. Ed.* **2003**, *42*, 1508−1511.

(23) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 197−206.

(24) Hur, S.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (21), 12015−12020.

(25) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 1472−1473.

(26) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *J. Am. Chem. Soc.* **2004**, *126*, 311−319.

(27) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. *Org. Biomol. Chem.* **2004**, *2* (7), 968−980.

(28) Zhang, X.; Zhang, X.; Bruice, T. C. *Biochemistry* **2005**, *44* (31), 10443−10448.

(29) Guimarães, C. R. W.; Udier-Blagović, M.; Tubert-Brohman, I.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2005**, *1*, 617−625.

(30) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J.; Field, M. J. *J. Am. Chem. Soc.* **2001**, *123*, 1709−1712.

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

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

(33) Ladner, J.; Reddy, P.; Davis, A.; Tordova, M.; Howard, A.; Gilliland, G. *Acta Crystallogr., Sect. D* **2000**, *56*, 673−683.

(34) Field, M. J. *A Practical Introducction to the Simulation of Molecular Systems*; Cambridge University Press: Cambridge, U. K., 1999.

(35) Field, M. J.; Albe, M.; Bret, C.; Proust-de Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21* (12), 1088−1100.

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

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

(38) Szefczyk, B.; Mulholland, A. J.; Ranaghan, K. E.; Sokalski, W. A. *J. Am. Chem. Soc.* **2004**, *126*, 16148−16159.

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

(40) Zuckerman, D. M.; Woolf, T. B. *J. Chem. Phys.* **2002**, *116*, 2586−2591.

(41) Hummer, G. *J. Chem. Phys.* **2004**, *120* (2), 516−523.

(42) Giacovazzo, C.; Monaco, H.; Artioli, G.; Viterbo, D.; Ferraris, G.; Gilli, G.; Zanotti, G.; Catti, M. *Fundamentals of Crystallography*, 2nd ed.; Giacovazzo, C., Ed.; International Union of Crystallography Texts on Crystallography 7; Oxford University Press: New York, 2002.

(43) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins* **1993**, *17*, 412−425.

(44) Wheeler, R. A.; Dong, H.; Boesch, S. E. *ChemPhysChem* **2003**, *4*, 382−384.

(45) Feher, M.; Schmidt, J. M. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 346−353.

(46) Roca, M.; Moliner, V.; Tuñón, I.; Hynes, J. T. *J. Am. Chem. Soc.* **2004**, *126*, 1369−1376.

(47) Nam, K.; Prat-Resina, X.; Garcia-Viloca, M.; Devi-Kesavan, L. S.; Gao, J. *J. Am. Chem. Soc.* **2004**, *126*, 1369−1376.

(48) Berkes, P.; Zito, T. *Modular Toolkit for Data Processing*, version 1.1.0; 2005.

(49) Chalmet, S.; Harb, W.; Ruiz-López, M. F. *J. Phys. Chem. A* **2001**, *105* (51), 11574−11581.

(50) Käb, G.; Schröder, C.; Schwarzer, D. *Phys. Chem. Chem. Phys.* **2002**, *4*, 271.

(51) Neufeld, A. A.; Schwarzer, D.; Schröeder, J.; Troe, J. *J. Chem. Phys.* **2003**, *119*, 2502−2512.

(52) Hänggi, P. *Rev. Mod. Phys.* **1990**, *62* (2), 251−341.

(53) Ruggeiro, G. D.; Guy, S. J.; Martí, S.; Moliner, V.; Williams, I. H. *J. Phys. Org. Chem.* **2004**, *17*, 592−601.

(54) Rhee, Y. M.; Pande, V. S. *J. Phys. Chem. B* **2005**, *109* (14), 6780−6786.

(55) Henkelman, G.; Jóhannesson, G.; Jónsson, H. In *Theoretical Methods in Condensed Phase Chemistry*; Schwartz, S. D., Ed.; Progress in Theoretical Chemistry and Physics 5; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2000; pp 269−300.

(56) Crehuet, R.; Thomas, A.; Field, M. J. *J. Mol. Graphics Modell.* **2005**, *24*, 102−110.

(57) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 5964−5972.

(58) Noonan, R. C.; Carter, C. W.; Bagdassarian, C. K. *Protein Sci.* **2002**, *11*, 1424−1434.

(59) Eletsky, A.; Kienhöfer, A.; Hilvert, D.; Pervushin, K. *Biochemistry* **2005**, *44*, 6788−6799.

(60) Kast, P.; Hartgerink, J. D.; Asif-Ullah, M.; Hilvert, D. *J. Am. Chem. Soc.* **1996**, *118*, 3069−3070.

(61) Gray, J. V.; Knowles, J. R. *Biochemistry* **1994**, *33*, 9953−9959.

(62) Worthington, S. E.; Roitberg, A. E.; Krauss, M. *J. Phys. Chem. B* **2001**, *105*, 7087−7095.

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

(64) Gamper, M.; Hilvert, D.; Kast, P. *Biochemistry* **2000**, *39*, 14087−14094.

(65) Vanden-Eijnden, E.; Tal, F. A. *J. Chem. Phys.* **2005**, *123*, 184103.

(66) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 024106.

(67) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.