THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Enzymatic Minimum Free Energy Path Calculations Using Swarms of Trajectories

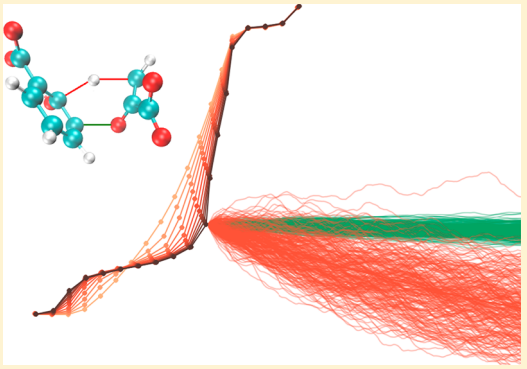
Melchor Sanchez-Martinez,[†] Martin Field,[‡] and Ramon Crehuet*[,†]

[†]Institute of Advanced Chemistry of Catalonia (IQAC), CSIC, Jordi Girona 18-26, 08034, Barcelona, Spain

[‡]Institut de Biologie Structurale (CEA, CNRS UMR5075, Université Joseph Fourier - Grenoble I), 71 Avenue des Martyrs, CS 10090, 38044 Grenoble Cedex 9, France

Supporting Information

ABSTRACT: The development of approaches for simulating rare events in complex molecular systems is a central concern in chemical physics. In recent work, Roux and co-workers proposed a novel, swarms of trajectories (SoT) method for determining the transition paths of such events. It consists of the dynamical refinement on the system's free energy surface of a putative transition path that is parametrized in terms of a set of collective variables (CVs) that are identified as being important for the transition. In this work, we have implemented the SoT method and used it to investigate the catalytic mechanisms of two enzymatic reactions using hybrid QM/MM potentials. Our aim has been to test the performance of SoT for enzyme systems and to devise robust simulation protocols that can be employed in future studies of this type. We identify the conditions under which converged results can be obtained using inertial and Brownian dynamical evolutions of the CVs, show that the inclusion of several CVs can give significant additional insight into the mechanisms of the reactions, and show that the use of minimum energy paths as starting guesses can greatly accelerate path refinement.



INTRODUCTION

The theoretical description of enzymatic mechanisms is based on free energy profiles, and the calculation of these profiles has become an important problem in computational biochemistry.[1−4] The free energy profile describes the chemical mechanism, and the resulting energy barrier allows the estimation of the rate of the process.[5] These profiles are defined along a hypothetical reaction coordinate whose finding is highly nontrivial.

There have been different approaches to the description of these profiles. On the one hand, one can define a set of presumably relevant collective variables (CVs) and calculate a free energy surface (aka potential of mean force) depending on these variables. Once a free energy surface is determined, it is usually projected in two dimensions and visually inspected to determine minimum free energy paths (MFEPs) connecting the different basins. These paths give a one-dimensional representation of the surface, and the value of the free energy along these paths produces a free energy profile (see the Methods section for a mathematical definition).[6] Methods such as adaptive biased force (ABF)[7] sampling, metadynamics,[8,9] and umbrella sampling[10] require a precise choice of a few CVs. As these methods describe the full free energy surface, they scale exponentially with the number of variables and rapidly become impractical due to the computational expense and difficulty of exploring multidimensional energy surfaces. Unfortunately, enzymatic reactions are complex and often need many CVs to be completely described.

On the other hand, the computational burden would be highly reduced if one could directly trace the paths in the free energy surface without having to fully determine that surface. Chain-of-states methods, that include the zero-temperature string[11,12] and nudged elastic band (NEB)[13,14] methods, can directly determine reaction paths. However, in their basic versions, these methods produce only minimum (potential) energy paths (MEPs), as they omit sampling and entropic contributions.[15] Nevertheless, they can be extended or generalized so that the determination of free energies is, in principle, possible.[11,16,17]

In one example of this type, the string method was modified to produce MFEPs by permitting sampling among a set of CVs.[15] In a related development, Roux and co-workers[18] proposed a novel method that employed swarms of trajectories (SoT) to evolve the string and to estimate its average displacement in CV space. Subsequently, they applied it to the study of a large biomolecular transformation.[19] Chemical reactions catalyzed by enzymes, however, have very different

Special Issue: William L. Jorgensen Festschrift

Received: July 2, 2014
Revised: October 1, 2014
Published: October 6, 2014

   dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

dynamics than the latter types of processes, as they take place on the time scale of bond vibrations and involve only small-amplitude motions of a few atoms. All of these effects make the relaxation from the transition state an inertial process, far from the diffusive process that was assumed in the original SoT formulation. This would suggest that SoT is not a good framework to study enzyme catalysis, but in a recent comprehensive study, Maragliano et al. showed that SoT gives the correct converged MFEP *independent* of the dynamics of the system, as long as the time scale chosen to calculate the changes of the CVs is short enough.[20]

Alternative approaches to calculate free energies based on CVs have also been proposed,[21] but their application to large systems such as enzymes is rare. An exception is the work of Tuñon and co-workers who adapted the method of Branduardi et al.[21] to study enzyme catalysis and applied it to the mechanism of isochorismate pyruvate lyase (IPL).[22] This method does not seek to optimize a MFEP but instead calculates the free energy profile associated with the curve followed by the minimum potential energy path.

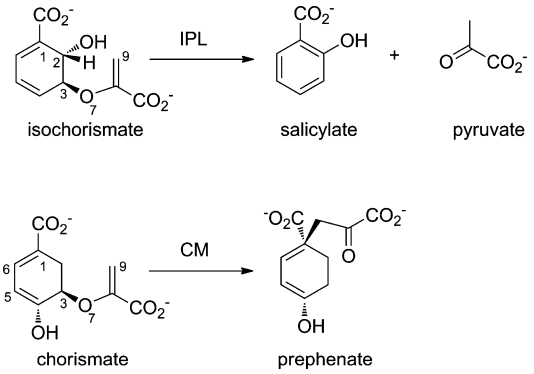
In this work, we have implemented the SoT method in the pDynamo molecular modeling library[23] and applied it to the study of the IPL and chorismate mutase (CM) enzymatic reaction mechanisms. These are both realistic test cases about which much information concerning the mechanisms is already known.[22,24−32] Our aim has been to investigate the performance of SoT in enzymes and to evaluate the conditions under which it works. We show that using several CVs can give significant additional insight into enzymatic mechanisms and also find that large increases in performance can be obtained if one initiates the SoT calculation from MEPs calculated with a hybrid string/NEB chain-of-states method.[4,33,34]

METHODS

Computational Details. All methodological development and simulations, including system setup, were done with the pDynamo[23] program, version 1.8.0. As preliminary work, we implemented a version of the SoT method in pDynamo and ensured that it reproduced the results of tests, including those on the blocked alanine dipeptide, that were published in the original papers.[15,18] Subsequently, to test the validity of our implementation of the SoT method for the investigation of enzymatic reactions, we chose two different enzyme systems, IPL and CM. We studied the reactions using hybrid quantum mechanical (QM) / molecular mechanical (MM) potentials in combination with chain-of-states reaction path calculations. As the QM method in our QM/MM potentials, we employed the AM1[35] semiempirical Hamiltonian. Although the latter is less precise than, say, density functional theory (DFT) methods, it is much less computationally demanding and thus makes possible a thorough analysis of the SoT approach to our test systems. In any case, we intend to apply these higher level potentials in our future studies using the SoT method.

Isochorismate Pyruvate Lyase. IPL transforms isochorismate into salicylate and pyruvate in a pericyclic reaction (Scheme 1). Our simulation model of IPL was derived from the X-ray crystallographic structure with PDB entry 2H9D.[36] This contains the pyruvate-bound IPL from *Pseudomonas aeruginosa* (PchB) with two pyruvate molecules in the active site. The latter were removed and replaced by an isochorismate molecule which represents the reactant state. The positions of hydrogens were then built and the whole system solvated in an orthorhombic water box of dimension 68 × 46 × 36 Å^3, with

Scheme 1. Schemes of the Reactions Catalyzed by Isochorismate Pyruvate Lyase (IPL) and Chorismate Mutase (CM)



an appropriate number of Na^+ counterions added to neutralize the overall charge of the system, giving ∼11000 atoms in total. We used the OPLS/AA[37] force field to describe the protein, together with the TIP3P[38] model for the water solvent, and periodic boundary conditions for the long-range interactions. To generate a suitable starting structure for the simulation, the structure of the solvated protein was first energy minimized, followed by an equilibrating molecular dynamics (MD) simulation at a temperature of 300 K. For the QM/MM simulations, the QM region contained just the substrate, isochorismate, numbering 24 atoms. In the QM/MM calculations, only atoms within 20 Å from the O7 oxygen of the substrate were allowed to move (Figure 1).

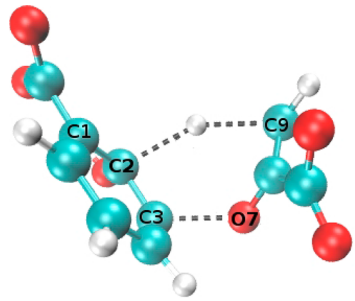


Figure 1. Transition state structure for the IPL transformation. The bonds that are being broken or formed are represented with dotted lines.

Chorismate Mutase. CM catalyzes the Claisen rearrangement from chorismate to pyruvate (Scheme 1). We modeled our CM simulation system from the X-ray crystallographic structure of the *Bacillus subtilis* enzyme with PDB entry 1COM.[39] The latter contained four homotrimers and one prephenate (PRE) of which we retained one homotrimer together with the PRE molecule. The remaining setup was similar to that we employed for the IPL system, except that we used K^+ and Cl^- counterions to neutralize the overall system charge. The final system had ∼17700 atoms with box dimension 60 × 50 × 60 Å^3. In the QM/MM calculations, the QM region contained only the substrate, PRE, numbering 24 atoms, and in the reaction path calculations, only atoms within 12 Å from the C1 carbon of the PRE were allowed to move (Figure 2).

Initial Reaction Paths. Initial paths are required to perform SoT calculations. We employed MEPs that were obtained by

dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

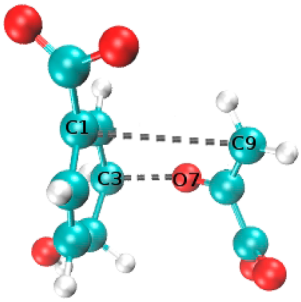


Figure 2. Transition state structure for the CM transformation. The bonds that are being broken or formed are represented with dotted lines.

carrying out chain-of-states reaction path calculations using the hybrid NEB/string method that is implemented in the pDynamo program.[4,33,34] For each of the test cases, we started from a small number of path structures and gradually increased them until the energy profile converged. The number of structures per path depended on the path length and ruggedness, ranging from 19 to 37 in IPL and from 18 to 28 in CM.

The MFEP and SoT Methods. Here we present a brief summary of the theoretical background behind the MFEP and SoT calculations. Readers are referred to the original papers for a full discussion.[15,18,20]

The Minimum Free Energy Path. Consider a system described by the Cartesian coordinates $\mathbf{x} \in R^n$ with a standard equilibrium distribution

$$p(\mathbf{x}) = Z^{-1} e^{-\beta V(\mathbf{x})} \tag{1}$$

where $\beta = 1/k_B T$ is the inverse temperature, $V(\mathbf{x})$ is the potential energy, and $Z = \int d\mathbf{x} e^{-\beta V(\mathbf{x})}$. We assume that there are no constraints in the system, and that the part of the density arising from the momenta has been integrated out.[15] We now introduce M CVs that are functions of $\mathbf{x}$ and that can distinguish distinct reacting configurations of the system

$$\tilde{\mathbf{z}}(\mathbf{x}) = \{\tilde{z}_1(\mathbf{x}), \tilde{z}_2(\mathbf{x}), ..., \tilde{z}_M(\mathbf{x})\} \tag{2}$$

The free energy, also known as the potential of mean force, associated with $\tilde{z}(\mathbf{x})$ is a function that depends on $z = (z_1, ..., z_M)$ and is calculated as

$$F(\mathbf{z}) = -k_B T \ln(Z^{-1} \int_{R^n} e^{-\beta V(\mathbf{x})} \delta(z_1 - \tilde{z}_1(\mathbf{x})), ..., \delta(z_M - \tilde{z}_M(\mathbf{x}))\, d\mathbf{x}) \tag{3}$$

This is the M-dimensional free energy surface described in the Introduction. A MFEP is defined on a free energy surface of eq 3 in the same way as a MEP is defined on a potential energy surface. Thus, it is the path between two minima on the surface such that the following condition holds:

$$[\mathbf{M}(\mathbf{z}) \cdot \nabla F(\mathbf{z})]^{\perp} = 0 \tag{4}$$

In this equation, $\nabla F(\mathbf{z})$ is the gradient of the free energy, $\mathbf{M}(\mathbf{z})$ is a metric tensor that accounts for the curvilinear nature of the CVs, and $\perp$ indicates projection in the direction perpendicular to the curve. Full details, including an expression for the metric tensor, may be found in the work of Maragliano et al.[15,20] We note that the metric tensor also appears when calculating minimum potential energy paths in terms of CVs, so it is not exclusive to the free energy. The main advantage of the SoT method is that this tensor is never calculated, as its influence is implicitly taken into account with the swarms of trajectories (see below).

As long as the energy barriers are high compared to the thermal energy and that the reaction trajectories cluster around reaction tubes, the MFEP corresponds to the path with highest likelihood, and its maximum corresponds to structures that have the same chance of falling into either the reactant or product basins.[15]

The SoT Reaction Path: Evolution and Reparameterization. Equation 3 can be used to determine the MFEP if a convenient representation of the path as a function of the variables, z, is available. In the SoT formalism, this is done by parametrizing the path as $z(\alpha)$, with $\alpha \in [0, 1]$, and where $\alpha = 0$ represents the initial (reactant) state and $\alpha = 1$ the final (product) conformation.[15,18] It is then assumed that the CVs evolve according to a noninertial Brownian dynamics over some time step, Δt, according to

$$z_i(\Delta t) = z_i(0) + \sum_j \left(-\beta D_{ij}[\mathbf{z}(0)] \partial_j F[\mathbf{z}(0)] + \partial_{z_j} D_{ij}[\mathbf{z}(0)]\right) \Delta t + R_i(0) \tag{5}$$

where D_{ij} is the diffusion tensor, which is equal to $k_\beta T \mathbf{M}(\mathbf{z})$ and $R_i(t)$ is a Gaussian thermal noise with a mean of zero. Once eq 5 has been defined, it can be employed to locate the most probable transition path,[15,18] which is the path such that a system anywhere along it will have the highest probability of remaining on it as it evolves. This is so because the most probable value of the Gaussian noise is zero. The key finding of ref 20 was to prove that when δT is small eq 5 also defines a path that satisfies

$$[-\mathbf{M}(\mathbf{z}) \nabla F(\mathbf{z}) + k_B T \nabla M(\mathbf{z})]^{\perp} = 0 \tag{6}$$

The extra term in this equation compared to eq 3 defining the MFEP was shown to be negligible (at least for the molecular system studied),[22] and from now on, we will consider that the SoT method converges to the MFEP. The equivalence of the SoT method—originally defined to locate the most probable transition paths—and the MFEP can be understood and expected because the MFEP corresponds to the path with the maximum likelihood for a system described with a given set of CVs.[15]

To evolve an initial path toward the MPTP, an approximation to eq 5 is needed. A way to accomplish this is using the so-called average drift[40] (or average displacement) evaluated from an ensemble of unbiased trajectories of length Δt initiated from each image of the path

$$\overline{\Delta z_i(\Delta t)} = \overline{z_i(\Delta t) - z_i(0)} \equiv \sum_j \left(-\beta D_{ij}[\mathbf{z}(0)] \partial_j F[\mathbf{z}(0)] + \partial_{z_j} D_{ij}[\mathbf{z}(0)]\right) \Delta t \tag{7}$$

where the thermal noise in eq 5 is averaged to zero. Thus, by calculating how the CVs evolve as a function of Δt, we need not calculate the terms on the right-hand side of eq 7.

An important insight of the string method is that of reparameterization,[11,12,15,16] which consists of the imposition of a constraint, normally a Euclidean distance,[15] between neighboring path images after each iteration. In practice, this is done by interpolating a curve through the path image structures and then redistributing them along the interpolated path. This is essential because it avoids the problem of the path

dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

images congregating in regions of low free energy after repeated application of eq 7.

Implementation of the SoT Method. Implementation of the SoT method is quite straightforward and consists of the following steps:

(i) Generate a path of N images with M CVs that describes the reacting system.

(ii) Perform thermalized molecular dynamics simulation for each image with the values of the CVs for each image restrained about their starting (reference) values from the preceding step—either (i) or (v). This step consists of a short equilibrium trajectory followed by the generation of a larger production trajectory for each image. We employed Langevin dynamics for these simulations using a collision frequency of 25 ps^{-1}. The CVs were restrained using potentials of harmonic form, with force constants of 8000 kJ mol^{-1} $Å^{-2}$, when a single distance was being restrained, and of 4000 kJ mol^{-1} $Å^{-2}$ for the sum or difference of two distances. In the case of IPL, we also tested a larger single-distance force constant of 12000 kJ mol^{-1} $Å^{-2}$ but obtained the same results as with the smaller value (data not shown).

(iii) Run multiple short unbiased trajectories for each image using configurations from the trajectories generated in step (ii). We explored different lengths and types of dynamics for these unbiased trajectories, the details of which are discussed later.

(iv) Calculate the average displacements of the CVs for each image arising from the unbiased trajectories using eq 7.

(v) Determine if the differences between the current average displacements and those of the previous iteration fall below a certain tolerance level. If so, convergence of the SoT calculation has been achieved and the simulation stops. If not, the path is reparameterized by ensuring that the images are redistributed in CV space and the simulation returns to step (ii).

Free Energy Calculations. As we discussed in the Introduction, the advantage of finding a MFEP is that we can directly calculate a free energy profile from it. Thanks to the metric tensor present in eq 4 and using eq 3, we calculate the free energy profile from[15]

$$F(\mathbf{z}(\alpha)) - F(\mathbf{z}(0)) = \int_0^{\alpha} \sum_{i=0}^{M} \frac{\mathrm{d}z_i(\alpha')}{\mathrm{d}\alpha'} \frac{\partial F(\mathbf{z}(\alpha'))}{\partial z_i} \,\mathrm{d}\alpha' \tag{8}$$

where M is the number of CVs and α is a scalar that parametrizes the path curve. In our implementation, we have parametrized $z(\alpha)$ as a cubic spline which means that we can calculate its derivatives with respect to α analytically. The derivatives of the free energy with respect to the CVs are obtained from the constrained dynamics simulations of step (ii) by averaging over the constraint forces applied to each of the CVs. Once these averages have been determined, we also parametrize them with a spline so that the integral of eq 8 and, hence, the free energy profile can be evaluated with an arbitrary number of points. We calculated confidence intervals for the profiles using a bootstrap method in which 1000 resamples of the raw data were generated for each constrained dynamics.

Evolution of the SoT and Dynamics of the Unbiased Trajectories. The free energy is a thermodynamic property of a system that does not depend on its dynamical evolution (eq 8). However, the definition of the MFEP is based on the evolution of the CVs as a function of a time increment (eq 7).

$$\begin{aligned}\tilde{\mathbf{z}}(\mathbf{x}(\Delta t)) = \tilde{\mathbf{z}}(\mathbf{x}(0)) &+ \Delta t \sum_i v_i \frac{\partial \tilde{z}_\alpha(\mathbf{x}(0))}{\partial x_i} \\ &+ \frac{1}{2}\Delta t^2 \sum_i \frac{1}{m_i}\left(-\frac{\partial V}{\partial x_i} - \gamma_i v_i\right)\frac{\partial \tilde{z}_\alpha(\mathbf{x}(0))}{\partial x_i} \\ &+ \frac{1}{2}\Delta t^2 \sum_{i,j} v_i v_j \frac{\partial^2 \tilde{z}_\alpha(\mathbf{x}(0))}{\partial x_i \partial x_j} + O(\Delta t^3)\end{aligned} \tag{9}$$

This formula was deduced assuming that z evolves in the Brownian regime.[18] Later, Maragliano et al. showed that this formula is valid irrespective of the dynamics of the system,[20] as long as it is in the limit of short-time evolution. In that regime, the evolution of the CVs depends on Δt^2 because the velocities average to zero. This result is based on Langevin dynamics for a system (which include inertial dynamics for zero friction). However, in the overdamped regime (Brownian dynamics), the velocity does not average to zero but is proportional to the force:

$$\langle v_i \rangle = \left\langle -\frac{\partial V}{\gamma_i \partial x_i} \right\rangle \tag{10}$$

where γ_i is the friction coefficient. Thus, in a Brownian, noninertial regime, the evolution of the CV is linear with time.

For the IPL system, we have explored three collision frequencies for the evolution of the unbiased trajectories. A very low collision frequency of 25 ps^{-1}, close to the inertial regime, a high collision frequency of 2500 ps^{-1}, close to the Brownian regime, and an intermediate frequency of 250 ps^{-1}.

RESULTS AND DISCUSSION

Minimum Energy Path for IPL. We calculated MEPs with different numbers of structures for IPL, and these are shown in Figure S1 (Supporting Information). The potential energy barrier is high and narrow, which needs a fairly high number of points to be well-defined. The barrier height is 38 kcal/mol, which agrees with the value obtained by Tuñon and co-workers.[22]

The MEPs in Figure S1 (Supporting Information) are functions of the coordinates of all the movable atoms in the system. This is wasteful, as an important part of the path is devoted to describe relaxations that do not involve any energy change or motions relevant to reaction. Instead, it is advantageous to reparameterize these paths in terms of pertinent CVs, as this permits a more compact description and the use of less points because the transformation is better defined. IPL catalyzes the transformation of isochorismate into salycilate and pyruvate in a one-step process in which a proton is transferred from C2 to C9 and the C3−O7 bond is cleaved (Scheme 1, Figure 1, and Figure S2, Supporting Information). In this section, we have chosen the three most obvious CVs, namely, the C2−H, C9−H, and C3−O7 distances. In later sections, we explore the effects of using different sets of reaction coordinates. Figure S2 (Supporting Information) shows the energy profile for the reparameterized path and the evolution of these three CVs along it. It can be seen that there is an error of approximately 2.5 kcal/mol in the barrier height for the path with 19 points in comparison to that of 37 points, whereas the latter is in almost exact agreement with the 73-point barrier of Figure S1 (Supporting Information).

We started the SoT simulations from the reparameterized MEP. This has two advantages. First, the MEP is much closer

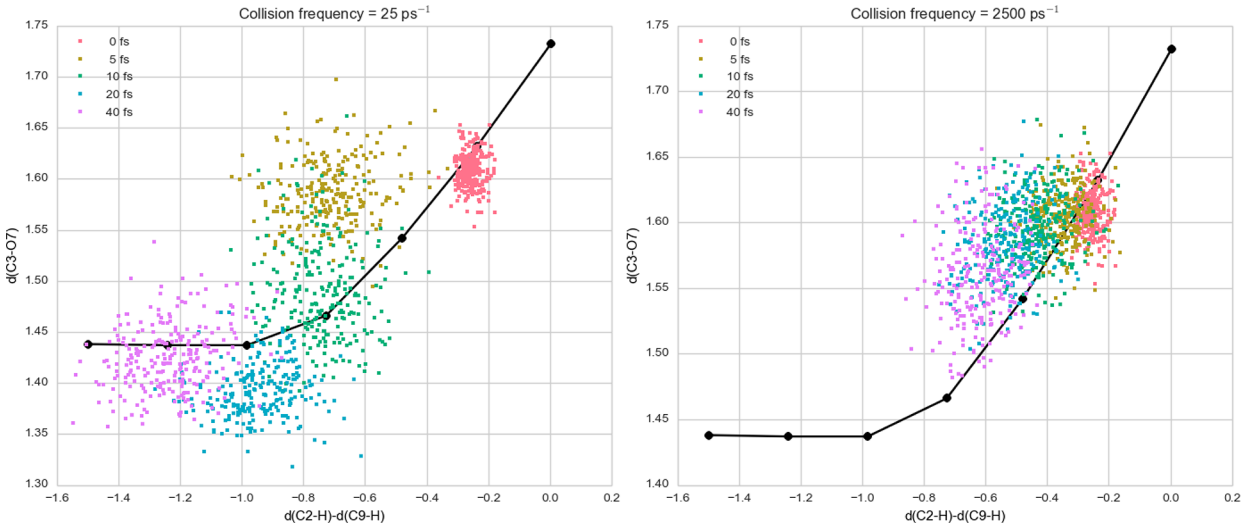


Figure 3. Positions of 250 unbiased trajectories after different simulation times for two different collision frequencies using the IPL system. For the low collision frequency (left), structures are almost reactant-like after only 40 fs of simulation. The MEP structures (used as a starting curve for the evolution of the MFEP) are plotted as black dots.

to the MFEP than any arbitrary initial guess. And second, we avoid the generation of strained geometries that could result in convergence problems. In fact, for this system, we wanted to compare convergence starting from a guess obtained by linear interpolation between the reactant and product structures. This, however, proved impossible, as sampling one of the frames of the linear guess resulted in a proton transfer from C2 to C1, which is both unrealistic and precluded any further optimization of the path.

As a final remark, we note that, around the barrier, the values of the three CVs change smoothly along the MEP with no oscillations. We expect the MFEP to display similar behavior, although oscillations have been observed in free energy paths based on the definition of Branduardi et al.[21] for many combinations of the algorithm's parameters.[22]

The Dynamics of the IPL System. Figure 3 shows that, when the friction is low, the CVs relax very fast, whereas, with high friction, as expected, the evolution of the system is slower. Thus, even though the ensemble average of eq 9 is not friction dependent,[20] the validity of the expression does depend on the friction, as the limiting case of eq 10 shows. Even with high frictions, the trajectories do not fall along the MEP, which indicates that the MEP and the MFEP will be different, as the results in the following section show.

It is clear that with a collision frequency of 25 ps^{-1} we have to take a much shorter time increment, Δt, than with a collision frequency of 2500 ps^{-1}. A short Δt is important not only to remain in the quadratic regime but also to remain in a region where the curvature of the underlying free energy surface is negligible compared to the evolution of the CV. Considering that at low frequency the CVs relax to the reactant values in less than 100 fs, this time has to be very short. At these short time scales, the evolution of the CVs cannot be considered Brownian or diffusive. Figure 4 shows that inertial oscillations remain after 100 fs. Therefore, it is obvious that one cannot find a Δt where the CVs evolve in a Brownian regime and, at the same time, the curvature of the free energy surface is not apparent. To alleviate this, we have two possibilities, both of which we explore. First, increase the friction coefficient in an artificial manner, or second, remain in an inertial regime using a short Δt. Although Vanden-Eijnden showed that the validity of the SoT evolution

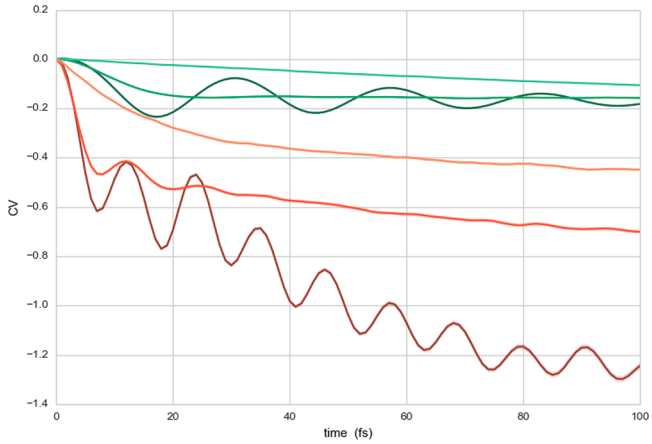


Figure 4. Time evolution of two CVs with different friction coefficients for the IPL system: d(C3−O7) with friction coefficients of 25, 250, and 2500 ps^{-1} (green from dark to light) and d(C2−H) − d(C9−H) with coefficients of 25, 250, and 2500 ps^{-1} (orange, from dark to light). Although we plot the difference of d(C2−H) and d(C9−H), they were treated as independent variables.

does not require diffusive dynamics, in the numerical example he used, the situation was diffusion-like (see Figure 3 in ref 20), and thus, our work is the first that studies SoT for inertial systems.

Figure 4 shows the values of the CVs as a function of time after averaging over 250 trajectories. It can be seen that the fast evolution of the CVs is also linked to an inertial evolution, as oscillations for the low collision frequency remain even after averaging. These oscillations are almost absent when the collision frequency is 250 ps^{-1} and disappear completely for a frequency of 2500 ps^{-1}. These results confirm that our simulations with different collision frequencies cover both the inertial and Brownian regimes.

To confirm these findings, Figure 5 shows the initial evolution of the CVs in the low and high collision frequency cases. As expected, at low friction, the evolution is quadratic for short time scales, so that for times <20 fs the dominant term in eq 9 is quadratic. By contrast, at high friction, the evolution is

dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

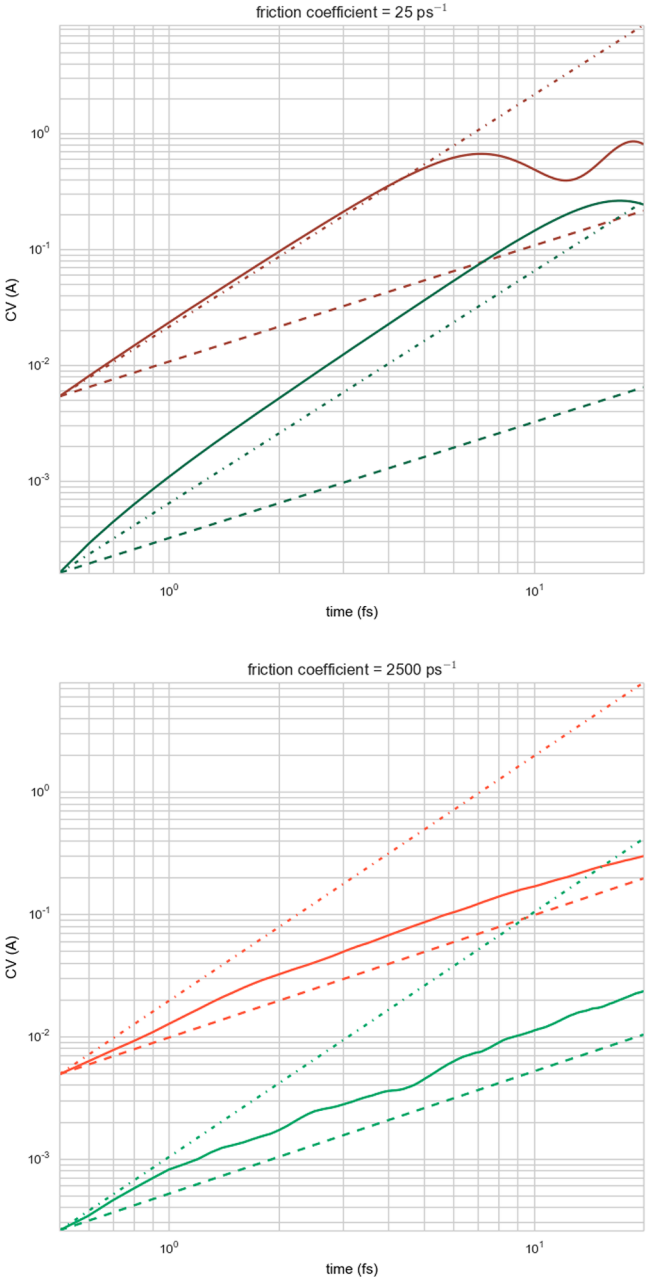


Figure 5. Short time evolution of the CVs for the IPL system (log scale) for the lowest and highest collision frequency dynamics. The color code is the same as in Figure 4. The dashed line corresponds to a linear regime and the dashed−dotted line to a quadratic regime.

initially quadratic but reverts quickly to a near linear form, as expected for Brownian motion.

On the basis of these results, we decided to try two different evolutions of the system. For the low collision frequency of 25 ps^{-1}, the unbiased trajectories were performed with 10 steps and a time step of 0.1 fs, resulting in Δt = 1 fs which assured we remained in the quadratic regime.

Because the change in **z** was small, we scaled $\langle \Delta z(\Delta t) \rangle_{\tilde{z}(\mathbf{x}(0))=z}$ in eq 7 by a factor of 4 so as to obtain similar displacements as those in the Brownian regime. The second setting corresponded to the Brownian regime, with a collision frequency of 2500 ps^{-1}, for which we performed 10 steps with a time step of 1 fs, which resulted in Δt = 10 fs.

Minimum Free Energy Paths. Figure 6 shows the evolution of the MFEPs using three CVs and the two

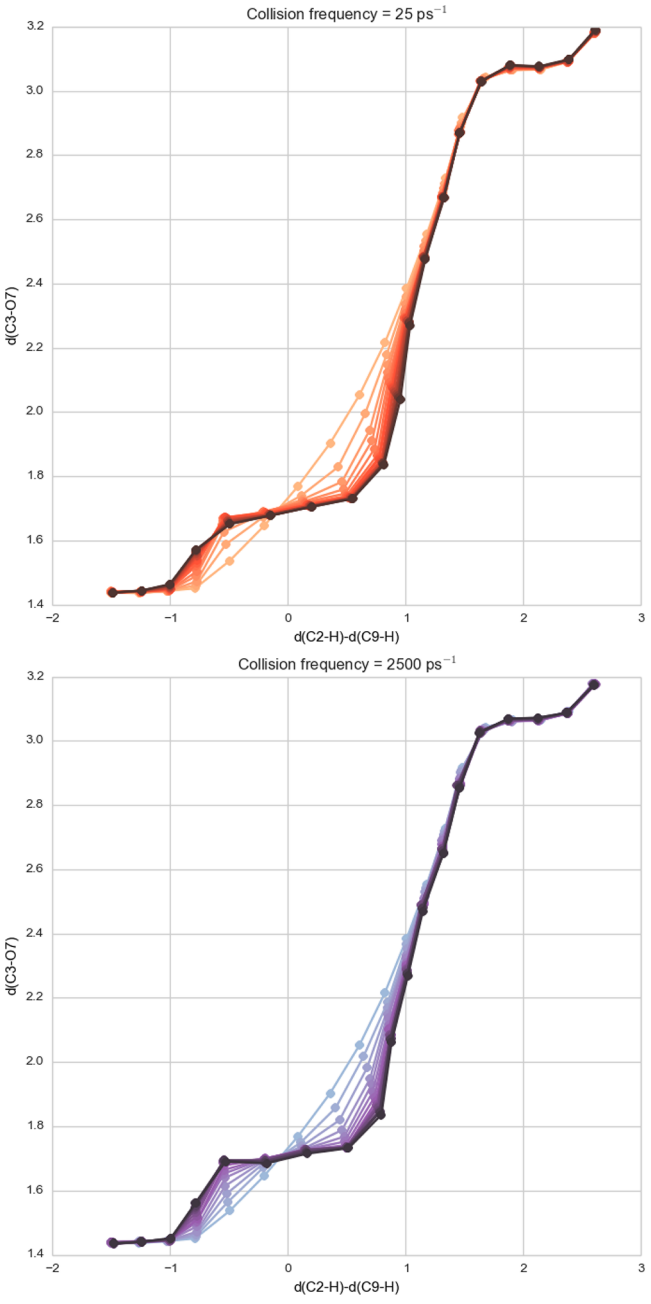


Figure 6. Evolution of the MFEP CVs for the IPL system calculated with SoT and in two different dynamical regimes (low friction, top; high friction, bottom). In both cases, the evolution is from light to dark colors. Although we plot the difference of d(C2−H) and d(C9−H), they were both treated as independent variables.

dynamical regimes described in the previous section. One can see that the MFEP is close to the MEP and that the evolution for both dynamical regimes is also very similar. We plot the difference of two of the CVs because of the difficulty of plotting three CVs separately—nevertheless, in the simulations, they were treated independently. The values of the CVs along the paths are compared in Figure 7. These paths are slightly different from the ones found by Tuñon and co-workers,[22] as ours are smoother and closer to the MEP. In addition, ours

dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

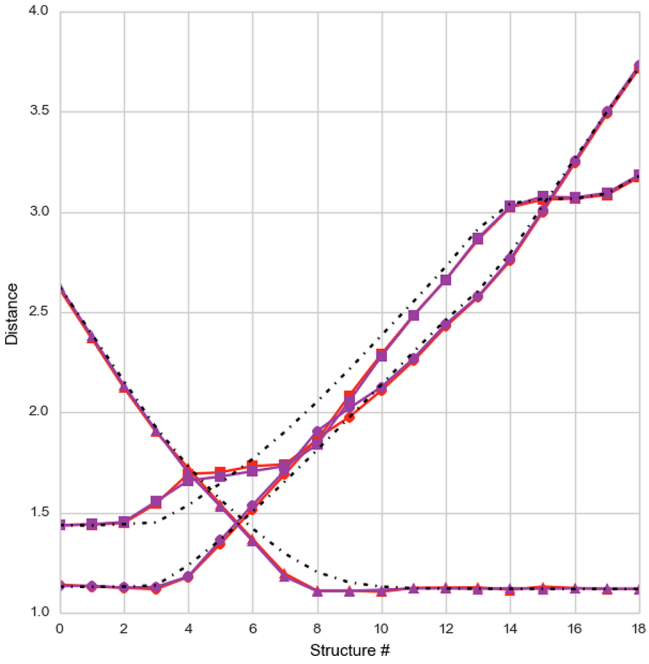


Figure 7. Values of the three CVs for the IPL system in the final path (same color code as Figure 6). The dashed black values are the MEP values. d(C3−O7), squares; d(C2−H), circles; d(C9−H), triangles.

seem to better characterize the products as they extend further and we find that the products have a larger C2−H distance than C3−O7 distance.

As the final paths are similar in both regimes, one expects that the free energies will also be similar. The free energy calculation is independent of the way the path is calculated, and is determined uniquely by the values of the CVs. As detailed in the Methods, we calculate the mean force contribution to the free energy integral with the same settings and the same underlying dynamics. The resulting profiles are shown in Figure 8, from which it can be seen that the barrier increases slightly during path optimization. As we start from the MEP, this seems a reasonable result, because the MFEP will have a better transition vector and thus a higher free energy.[15] Although Tuñon and co-workers find a free energy barrier lower than the potential energy barrier, our free energy barrier is above the potential energy barriers. In both cases, the differences are only 1−2 kcal/mol and thus the discrepancies are small. There are two reasons to expect a higher free energy barrier than a potential energy barrier. First, the interactions with the transition state are stronger than with reactants or products which will tend to rigidify the ensemble of transition state structures compared to reactants or products,[24] although this effect is often small.[41] Second, the transition state of this reaction is a cyclic species that is more constrained than either reactants or products. Figure S4 (Supporting Information) plots the confidence intervals of the final profile of Figure 8 (bottom). We can see that the oscillations of the path and the confidence interval are of the same order. The errors tend to accumulate at the end of the path because the profile is calculated by integrating the mean force, via eq 8.

Number of Degrees of Freedom. An important difference between the MEP and the MFEP is that the MFEP involves sampling along all the nonconstrained degrees of freedom. Some sets of reaction coordinates that are adequate for MEPs are not sufficient for MFEPs, as they do not include

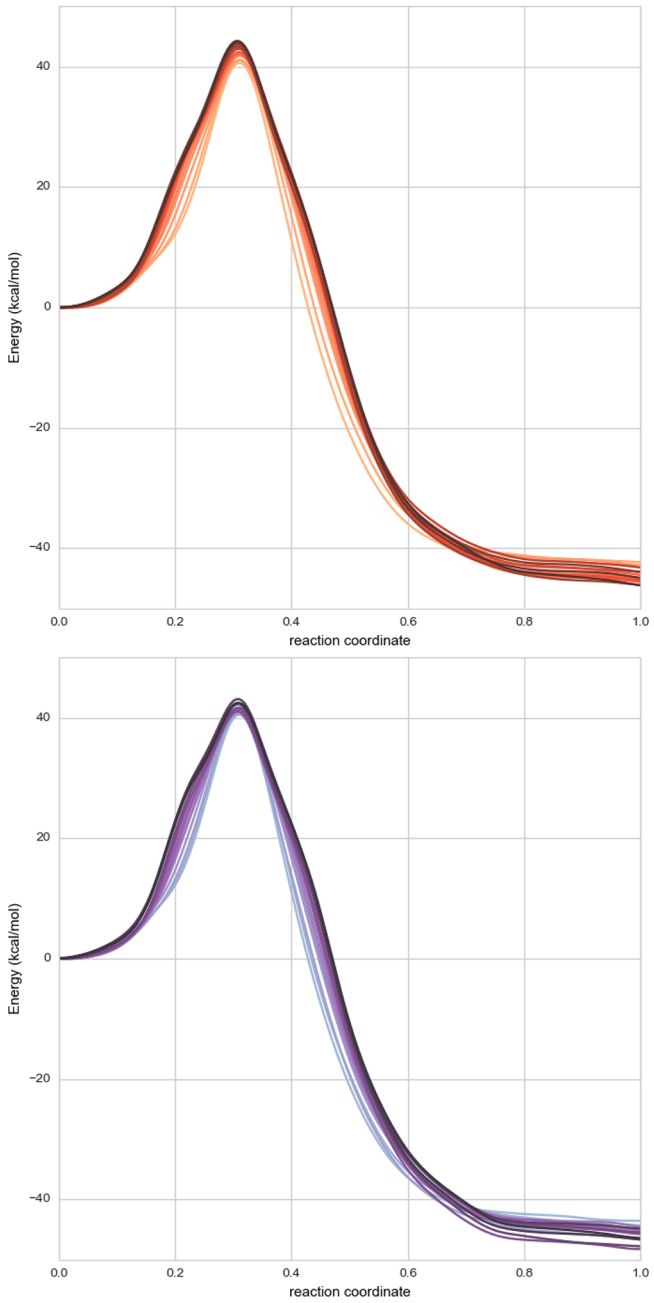


Figure 8. Evolution of the free energy profiles calculated from the MFEP for the IPL system calculated with SoT and in two different dynamical regimes (same color code as Figure 6).

all the variables involved in the transition vector at the transition state hypersurface. When this is the case, the free-energy barrier is too low. Thus, finding a lower free energy barrier does not necessarily indicate a more favorable path; it could also be due to a poor choice of reaction coordinates.[15,42,43] The use of SoT frees us from the use of a small number of degrees of freedom to determine the free energy profile. This is in contrast to more traditional methods that calculate free energies as a function of only one or two coordinates. Employing more coordinates in these schemes involves an exponential increase in computational cost and also produces surfaces in more than two dimensions, from which the extraction of the MFEP is nontrivial because gradients of these surfaces are not available.

dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

A typical approximation that is made when an atom is transferred is to use the difference of bond distances between donor and acceptor atoms, as was done in previous work on the IPL system.[25] We recalculated the IPL MFEP with two CVs defined as $z_1 = d(\mathrm{C2-H}) - d(\mathrm{C9-H})$ and $z_2 = d(\mathrm{C3-O7})$. The resulting path and its evolution is shown in Figure 9, and is similar to the three-CV results plotted in Figure 6, thereby confirming that this is a reasonable simplification in this system. The MFEP (Figure 10) gives also a barrier equal to the one obtained with three CVs.

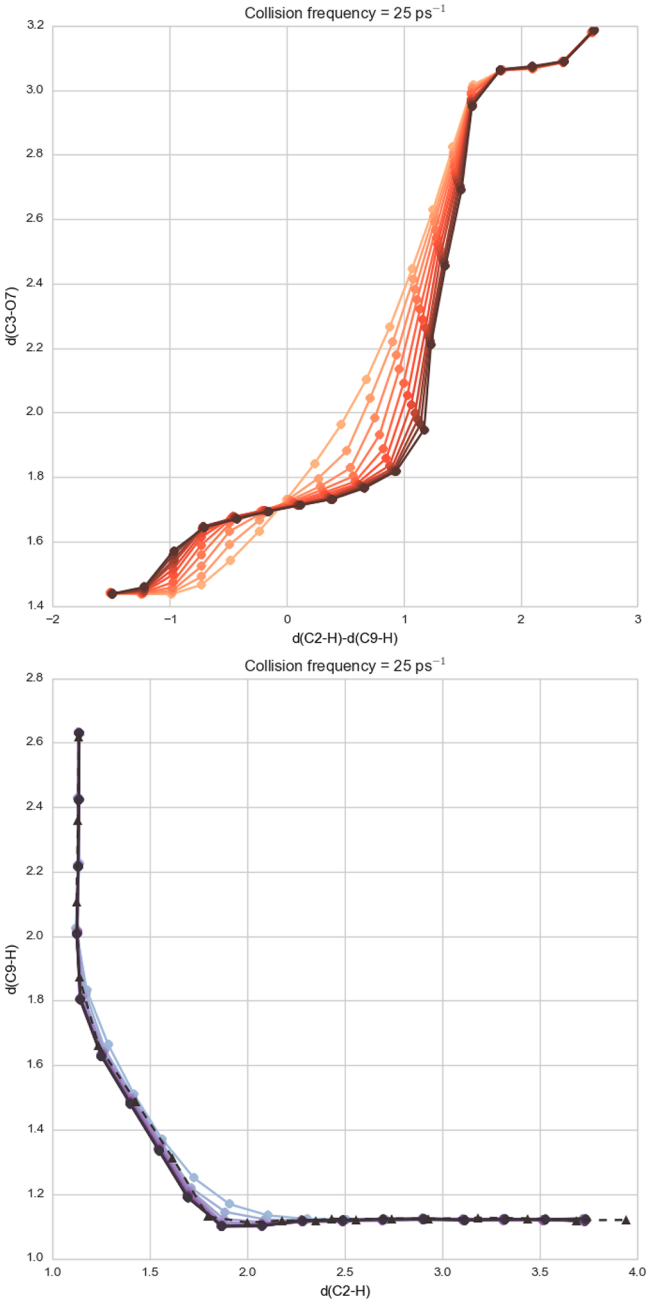


Figure 9. Evolution of the CVs when using two CVs for the IPL system. Top: a difference of distances, $d(\mathrm{C2-H}) - d(\mathrm{C9-H})$, and a distance, $d(\mathrm{C3-O7})$. Bottom: the distances $d(\mathrm{C2-H})$ and $d(\mathrm{C9-H})$ but with an apparently important CV, $d(\mathrm{C3-O7})$, neglected. The optimized values of $d(\mathrm{C2-H})$ and $d(\mathrm{C9-H})$ when using the full three CVs are plotted with red triangles.

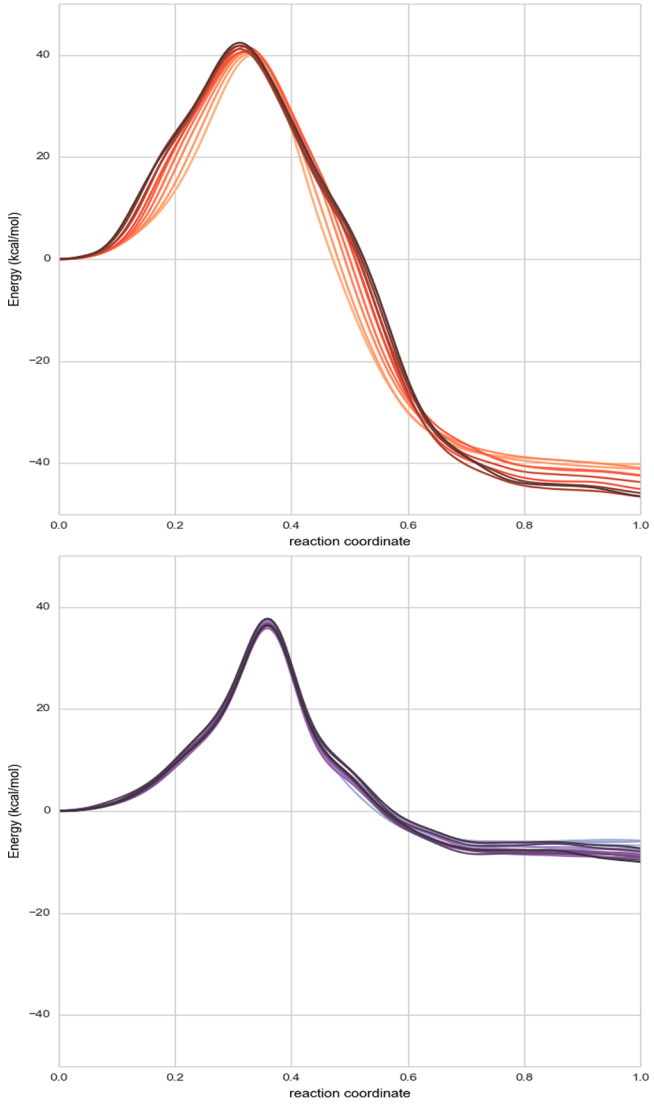


Figure 10. Free energy profiles calculated from the MFEP when using two CVs for the IPL system. Top: $d(\mathrm{C2-H}) - d(\mathrm{C9-H})$ and $d(\mathrm{C3-O7})$ correctly represent the transition vector, and thus give a correct energy profile. Bottom: $d(\mathrm{C2-H})$ and $d(\mathrm{C9-H})$. The relevant $d(\mathrm{C3-O7})$ CV is underestimated, so the free energy barrier is underestimated.

As a next test case, we disregard $d(\mathrm{C3-O7})$ and use only $d(\mathrm{C2-H})$ and $d(\mathrm{C9-H})$ as our CVs. Although unrealistic here, this is a situation that could arise because we inadvertently miss a relevant CV or because we do a traditional free energy calculation method and cannot afford to include one more reaction coordinate. Figure 9 shows that the evolution of the path is correct, and that both distances end with values equivalent to those of the three-CV optimized path. In that sense, the path is correctly described. However, if the missing reaction coordinate is necessary to describe the transition vector—as we expect—this will show up in the resulting free energy profile, which is precisely what Figure 10 shows. When we miss a relevant CV, the free energy barrier is lower than expected. This missing CV can be found with an analysis of the committor probability at the top of the barrier,[15,42] but this is an expensive calculation. When in doubt, one can always include an extra CV to the total set of CVs at no extra cost, and if that CV is not relevant, the results will not be affected. In the

dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

CM sections that follow, we test the case of including an irrelevant CV, and show that this is indeed the case.

Lowering the Cost of the SoT. The systems on which SoT has been used so far[18,19] have energies that are computationally fast to calculate, and thus, there has been little effort in analyzing the performance of SoT when sampling is expensive. For QM/MM calculation, the cost of each calculation is considerable when using semiempirical methods, as we do here, and would be much higher if DFT or other *ab initio* methods were to be employed. In the simulations that we have reported, we performed 250 × 10 = 2500 steps for the unbiased dynamics and 1000 + 5000 = 6000 steps for the constrained dynamics. We have also tried different numbers of molecular dynamics steps that are displayed in Table 1.

Table 1. Different Settings Used in the SoT Simulations[a]

equilibration steps	sampling steps	unbiased steps
1000	5000	250 × 10
1000	1000	250 × 10
500	500	250 × 10
500	500	250 × 5
200	200	100 × 5
100	100	50 × 5

[a]As the time step is 1 fs, these values also correspond to the total time length span of the simulations in fs.

We were surprised to find that even the least expensive settings gave a good convergence of the path, as is shown in Figure S3 (Supporting Information). The free energy profiles had errors of several kcal/mol for the total exothermicity but had energy barriers within 1 kcal/mol of the converged value. In any case, an optimized MFEP can always be recalculated with larger equilibration and sampling times in a final single iteration. These results suggest that SoT would be a good approach for calculating MFEPs in enzymes using DFT methods, as one can limit the number of steps that are required. In addition, the method is readily parallelizable, as the calculations for each point along the path are independent.

Chorismate Mutase. CM catalyzes a Claisen rearrangement from chorismate to prephanate (Scheme 1). It is probably the most studied enzyme with QM/MM methods,[24,27−32,44] but we will use it to explore two aspects of the SoT method, namely, the convergence of SoT when starting far from the MFEP and how linear combinations of CVs behave.

We first compare the evolution of the MFEP using two different sets of CVs and starting from an initial linear guess for the pathway structures in which intermediate images are linearly interpolated from the reactant and product structures. We note that there were no convergence problems using a linear guess for this system in contrast to the problems we experienced with the equivalent IPL simulations. The first CV set used $z_1 = d(\mathrm{C1-C9})$ and $z_2 = d(\mathrm{C3-O7})$ (see Figure 2). The second used the sum and difference of these two CVs: $z_1 = d(\mathrm{C1-C9}) - d(\mathrm{C3-O7})$ and $z_2 = d(\mathrm{C1-C9}) + d(\mathrm{C3-O7})$. Because they describe exactly the same subspace, they should produce the same free energy profile and they should have the same evolution. Figure 11 shows that this is indeed the case. Although previous work has shown that $z_1 = d(\mathrm{C1-C9}) - d(\mathrm{C3-O7})$ is a sufficient CV to obtain a good free energy profile, this information is usually only known *a posteriori* after the transition state structures or MEP have been determined and free energy calculations carried out. Being able to include both distances as independent CVs from the start gives the SoT approach a great deal of flexibility.

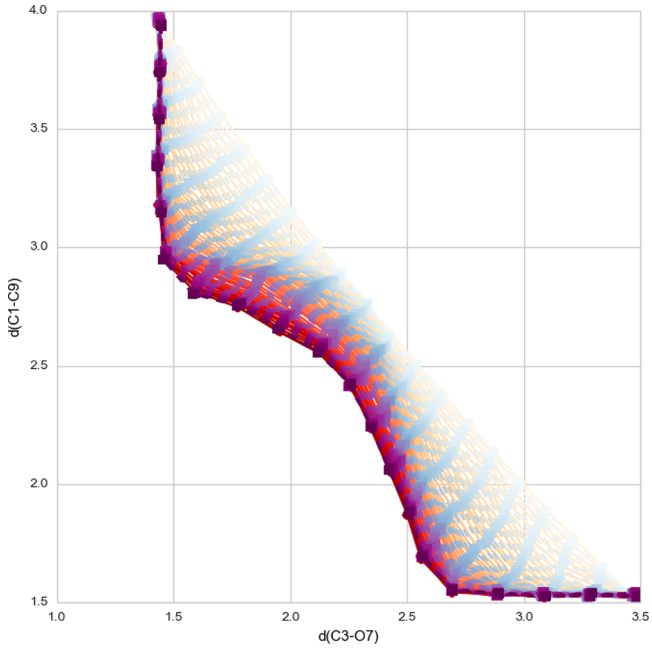


Figure 11. Evolution of two distances when using two different sets of CVs for the CM system. Red (light to dark), the $d(\mathrm{C3-O7})$ and $d(\mathrm{C1-C})$ set; purple (light to dark), the $d(\mathrm{C3-O7}) - d(\mathrm{C1-C9})$ and $d(\mathrm{C3-O7}) + d(\mathrm{C1-C9})$ set. One can see that the evolution and final path shape for both sets of CVs are essentially equivalent.

The calculation of the free energy from eq 8 gives the contribution to the profile as a sum for each of the CVs. In the previous section, we indicated that the use of irrelevant CVs did not affect the calculation of the MFEP or the profile. We have calculated the profile for the CM reaction using the three CVs, $z_1 = d(\mathrm{C1-C9})$, $z_2 = d(\mathrm{C3-O7})$, and $z_3 = d(\mathrm{C5-C6})$. Figure 12 indeed shows that z_3 does not affect the free energy profile, as was to be expected since this distance is not involved in the reaction. We note that the cost of including an extra CV is essentially zero, both in the dynamics and in the free energy calculation. It is possible, though, that extra CVs could slow convergence to the final MFEP, although starting with a

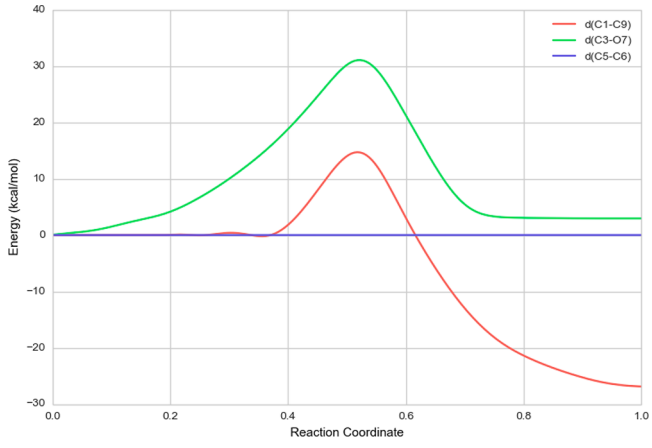


Figure 12. Decomposition of the free energy profile for the CM system into the contributions of the three CVs used. As $d(\mathrm{C5-C6})$ is not involved in this reaction, its contribution is zero.

dx.doi.org/10.1021/jp506593t | J. Phys. Chem. B 2015, 119, 1103−1113

reasonable guess, such as the MEP, should alleviate this problem. Returning to Figure 12, we see that the free energy decomposition in terms of CVs provides insight into their contributions to the MFEP. Thus, the initial free energy cost arises from stretching the C3−O7 bond, which is to be cleaved. In the TS region, both C3−O7 and C1−C9 contribute, whereas the exothermicity that arises in the descent to products comes largely from the formation of the new C1−C9 bond. On the other hand, as the C5−C6 bond remains in its equilibrium position along the MFEP, its contribution is essentially zero. Stretching this bond has a free energy cost, but in this particular reaction, it is not removed from its minimum value at any point, no mean force acts on it, and, thus, its contribution is null. This decomposition of the free energy is mathematically sound in an area where different approaches have proven complex and controversial.[45−48]

CONCLUSIONS

In this work, we have implemented the SoT method[18] and used it to study two enzyme catalytic mechanisms with QM/MM potentials. We have devised a suitable SoT simulation protocol for these types of systems and have shown that the results do not depend on the dynamical evolution of the CVs used to describe the reaction: both inertial and Brownian regimes lead to the same MFEP evolution and final curve. We have also suggested the use of the MEP as an initial starting guess to accelerate the optimization and to reduce convergence problems.

The study of enzyme mechanisms via MFEPs obtained as a function of a set of CVs has several advantages over studies based on more traditional methods. First, the computational cost does not increase with the number of CVs used. Second, the path can be easily visualized and the variation of several CVs analyzed independently. Finally, the calculation of free energies based on these CVs is robust with respect to the set of CVs and can give insights into their respective contributions to the free energy barrier and reaction exo- or endothermicity.

ASSOCIATED CONTENT

Supporting Information

Figures showing minimum energy paths, the evolution of the IPL MFEP, and a free energy profile. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ramon.crehuet@iqac.csic.es. Phone: +0034 934006116. Fax: +0034 932 045 904.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge financial support from the Ministerio de Innovación y Competitividad (CTQ2012-33324) and the Generalitat de Catalunya (2009SGR01472). We also acknowledge access to supercomputing time from the CSUC. M.S-M. thanks the Ministerio de Economia y Competitividad for a predoctoral fellowship. We also thank Sergi Martí for kindly sharing IPL structures with us.

REFERENCES

(1) Field, M. J. Simulating Enzyme Reactions: Challenges and Perspectives. *J. Comput. Chem.* **2002**, *23*, 48−58.
(2) Field, M. J. *A Practical Introduction to the Simulation of Molecular Systems*; Cambridge University Press: Cambridge, U.K., 2007.
(3) E, W.; Vanden-Eijnden, E. Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events. *Annu. Rev. Phys. Chem.* **2010**, *61*, 391−420.
(4) Aleksandrov, A.; Field, M. A Hybrid Elastic Band String Algorithm for Studies of Enzymatic Reactions. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12544−12553.
(5) Sheppard, D.; Terrell, R.; Henkelman, G. Optimization Methods for Finding Minimum Energy Paths. *J. Chem. Phys.* **2008**, *128*, 134106.
(6) Weinan, E.; Vanden-Eijnden, E. Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events. *Annu. Rev. Phys. Chem.* **2010**, *61*, 391−420.
(7) Darve, E.; Pohorille, A. Calculating Free Energies Using Average Force. *J. Chem. Phys.* **2001**, *115*, 9169.
(8) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(9) Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. Assessing the Accuracy of Metadynamics. *J. Phys. Chem. B* **2005**, *109*, 6714−6721.
(10) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(11) E, W.; Ren, W.; Vanden-Eijnden, E. String Method for the Study of Rare Events. *Phys. Rev. B* **2002**, *66*, 052301.
(12) E, W.; Ren, W.; Vanden-Eijnden, E. Simplified and Improved String Method for Computing the Minimum Energy Paths in Barrier-Crossing Events. *J. Chem. Phys.* **2007**, *126*, 164103.
(13) Johnson, M. E.; Hummer, G. Characterization of a Dynamic String Method for the Construction of Transition Pathways in Molecular Reactions. *J. Phys. Chem. B* **2012**, *116*, 8573−8583.
(14) Henkelman, G.; Uberuaga, B.; Jónsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113*, 9978−9985.
(15) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. String Method in Collective Variables: Minimum Free Energy Paths and Isocommittor Surfaces. *J. Chem. Phys.* **2006**, *125*, 24106.
(16) Ren, W.; Vanden-Eijnden, E.; Maragakis, P.; E, W. Transition Pathways in Complex Systems: Application of the Finite-Temperature String Method to the Alanine Dipeptide. *J. Chem. Phys.* **2005**, *123*, 134109.
(17) Vanden-Eijnden, E.; Venturoli, M. Revisiting the Finite Temperature String Method for the Calculation of Reaction Tubes and Free Energies. *J. Chem. Phys.* **2009**, *130*, 194103.
(18) Pan, A. C.; Sezer, D.; Roux, B. Finding Transition Pathways Using the String Method with Swarms of Trajectories. *J. Phys. Chem. B* **2008**, *112*, 3432−3440.
(19) Gan, W.; Yang, S.; Roux, B. Atomistic View of the Conformational Activation of Src Kinase Using the String Method with Swarms-of-Trajectories. *Biophys. J.* **2009**, *97*, L8−L10.
(20) Maragliano, L.; Roux, B.; Vanden-Eijnden, E. A Comparison between Mean Forces and Swarms-of-Trajectories String Methods. *J. Chem. Theory Comput.* **2014**, *10*, 524−533.
(21) Branduardi, D.; Gervasio, F. L.; Parrinello, M. From A to B in Free Energy Space. *J. Chem. Phys.* **2007**, *126*, 054103.
(22) Zinovjev, K.; Martí, S.; Tuñon, I. A Collective Coordinate to Obtain Free Energy Profiles for Complex Reactions in Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8*, 1795−1801.
(23) Field, M. J. The pDynamo Program for Molecular Simulations Using Hybrid Quantum Chemical and Molecular Mechanical Potentials. *J. Chem. Theory Comput.* **2008**, *4*, 1151−1161.
(24) Crehuet, R.; Field, M. J. A Transition Path Sampling Study of the Reaction Catalyzed by the Enzyme Chorismate Mutase. *J. Phys. Chem. B* **2007**, *111*, 5708−5718.

(25) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Mechanism and Plasticity of Isochorismate Pyruvate Lyase: A Computational Study. *J. Am. Chem. Soc.* **2009**, *131*, 16156−16161.
(26) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J.; Field, M. J. A Hybrid Potential Reaction Path and Free Energy Study of the Chorismate Mutase Reaction. *J. Am. Chem. Soc.* **2001**, *123*, 1709−1712.
(27) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. A Comparative Study of Claisen and Cope Rearrangements Catalyzed by Chorismate Mutase. An Insight into Enzymatic Efficiency: Transition State Stabilization or Substrate Preorganization? *J. Am. Chem. Soc.* **2004**, *126*, 311−319.
(28) Szefczyk, B.; Claeyssens, F.; Mulholland, A. J.; Sokalski, W. A. Quantum Chemical Analysis of Reaction Paths in Chorismate Mutase: Conformational Effects and Electrostatic Stabilization. *Int. J. Quantum Chem.* **2007**, *107*, 2274−2285.
(29) Claeyssens, F.; Ranaghan, K. E.; Lawan, N.; Macrae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Analysis of Chorismate Mutase Catalysis by QM/MM Modelling of Enzyme-Catalysed and Uncatalysed Reactions. *Org. Biomol. Chem.* **2011**, *9*, 1578−1590.
(30) Repasky, M. P.; Guimarães, C. R. W.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. Investigation of Solvent Effects for the Claisen Rearrangement of Chorismate to Prephenate: Mechanistic Interpretation via near Attack Conformations. *J. Am. Chem. Soc.* **2003**, *125*, 6663−6672.
(31) Guimarães, C. R. W.; Repasky, M. P.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. Contributions of Conformational Compression and Preferential Transition State Stabilisation to the Rate Enhancement by Chorismate Mutase. *J. Am. Chem. Soc.* **2003**, *125*, 6892−6899.
(32) Guimarães, C. R. W.; Udier-Blagović, M.; Tubert-Brohman, I.; Jorgensen, W. L. Effects of Arg90 Neutralization on the Enzyme-Catalyzed Rearrangement of Chorismate to Prephanate. *J. Chem. Theory Comput.* **2005**, *1*, 617−625.
(33) Crehuet, R.; Thomas, A.; Field, M. J. An Implementation of the Nudged Elastic Band Algorithm and Application to the Reaction Mechanism of HGXPRTase from Plasmodium Falciparum. *J. Mol. Graphics Modell.* **2005**, *24*, 102−110.
(34) Galvan, I. F.; Field, M. J.; Galván, I. Improving the Efficiency of the NEB Reaction Path Finding Algorithm. *J. Comput. Chem.* **2008**, *29*, 134−143.
(35) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(36) Zaitseva, J.; Lu, J.; Olechoski, K. L.; Lamb, A. L. Two Crystal Structures of the Isochorismate Pyruvate Lyase from Pseudomonas Aeruginosa. *J. Biol. Chem.* **2006**, *281*, 33441−33449.
(37) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(38) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(39) Chook, Y. M.; Gray, J. V.; Hengming, K.; Lipscomb, W. N. The Monofunctional Chorismate Mutase from Bacillus Subtilis. Structure Determination of Chorismate Mutase and Its Complexes with a Transition State Analog and Prephenate, and Implications for the Mechanism of the Enzymatic Reaction. *J. Mol. Biol.* **1994**, *240*, 476−500.
(40) Hummer, G.; Kevrekidis, I. G. Coarse Molecular Dynamics of a Peptide Fragment: Free Energy, Kinetics, and Long-Time Dynamics Computations. *J. Chem. Phys.* **2003**, *118*, 10762.
(41) Villa, J.; Strajbl, M.; Glennon, T. M.; Sham, Y. Y.; Chu, Z. T.; Warshel, A. How Important Are Entropic Contributions to Enzyme Catalysis? *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 11899−11904.
(42) Bolhuis, P. G.; Dellago, C.; Chandler, D. Reaction Coordinates of Biomolecular Isomerization. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5877−5882.
(43) Bonella, S.; Meloni, S.; Ciccotti, G. Theory and Methods for Rare Events. *Eur. Phys. J. B* **2012**, *85*, 1−19.
(44) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. Transition State Stabilisation and Substrate Strain in Enzyme Catalysis: Ab Initio QM/MM Modelling of the Chorismate Mutase Reaction. *Org. Biomol. Chem.* **2004**, *2*, 968−980.
(45) Brady, G. P.; Szabo, A.; Sharp, K. A. On the Decomposition of Free Energies. *J. Mol. Biol.* **1996**, *263*, 123−125.
(46) Brady, G. P.; Sharp, K. A. Decomposition of Interaction Free Energies in Proteins and Other Complex Systems. *J. Mol. Biol.* **1995**, *254*, 77−85.
(47) Boresch, S.; Karplus, M. The Meaning of Component Analysis: Decomposition of the Free Energy in Terms of Specific Interactions. *J. Mol. Biol.* **1995**, *254*, 801−807.
(48) Mark, A. E.; van Gunsteren, W. F. Decomposition of the Free Energy of a System in Terms of Specific Interactions. Implications for Theoretical and Experimental Studies. *J. Mol. Biol.* **1994**, *240*, 167−176.