JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

A Collective Coordinate to Obtain Free Energy Profiles for Complex Reactions in Condensed Phases

Kirill Zinovjev,[‡] Sergi Martí,[§] and Iñaki Tuñón*[,‡]

[‡]Departament de Química Física, Universitat de València, 46100 Burjassot, Spain

[§]Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castellón, Spain

Supporting Information

ABSTRACT: Exploration of chemical reactions in complex explicit environments has become an affordable task with the use of hybrid quantum mechanics/molecular mechanics potentials which allow calculating free energy profiles of chemical reactions under the influence of the surroundings. Tracing these free energy profiles requires the selection of a reaction coordinate, which can be cumbersome for those processes involving more than a single chemical event in a concerted step. We here propose a collective coordinate to be used in the calculation of free energy profiles for complex reactions in condensed phases. This coordinate is based in the definition of the advance along a path introduced by Branduardi et al. (*J. Chem. Phys.* **2007**, *126*, 054103) but modified to use internal coordinates which are more adequate for the description of chemical reactions. The coordinate is tested with the analysis of the isochorismate transformation to pyruvate and salycilate in aqueous solution and in the active site of PchB, a reaction that involves a CO bond breaking simultaneously with a proton transfer between two carbon atoms. The coordinate introduced here allows obtaining smooth and meaningful free energy profiles of the reaction.

1. INTRODUCTION

Free energy changes (Gibbs and Helmholtz) determine the sense of a spontaneous process carried out at constant pressure and temperature or constant volume and temperature, respectively.[1] Thus, determination of this magnitude is one of the main goals during computational analysis of chemical transformations. In principle, the Helmholtz free energy can be evaluated (up to a constant which is irrelevant for free energy differences) as[2]

$$A = -kT\cdot\ln\int \mathrm{e}^{-E(\mathbf{r}^N)/kT}\mathrm{d}\mathbf{r}^N \tag{1}$$

where k is the Boltzmann constant, T is the temperature, and $E(\mathbf{r}^N)$ is the potential energy function of the system (a similar expression, including the expansion work holds for the Gibbs free energy). Direct evaluation of the multidimensional integral is unaffordable except for very simple ideal systems. However, calculation of free energy differences is more affordable, and many efforts have been devoted to the development of methods for the computation of free energy differences in real systems, such as free energy perturbation, thermodynamic integration, slow growth, and others.[3,4] Applicability of these methods covers many areas of chemistry, including materials and drug design, where the prediction of the free energy change associated to a particular structural modification can be of relevance to guide the preparation of new relevant compounds with targeted properties.

Transition-state theory (TST)[5,6] extended the applicability of the free energy functions not only to the prediction of equilibrium properties but also to kinetics. Introduction of the transition-state (TS) concept allows the determination of the rate constant of a given chemical transformation from the free energy difference between this and the reactants state. In particular, the rate constant can be formulated in terms of the free energy change associated to a particular reaction coordinate or potential of mean force (PMF).[7] This magnitude is related to the probability density of finding the system at a given value of the selected coordinate:

$$A(s_0) = -kT\cdot\ln\rho(s_0) = -kT\cdot\ln\frac{\int \mathrm{e}^{-E(\mathbf{r}^N)/kT}\delta(s_0 - s)\mathrm{d}\mathbf{r}^N}{\int \mathrm{e}^{-E(\mathbf{r}^N)/kT}\mathrm{d}\mathbf{r}^N} \tag{2}$$

where δ represents the Dirac's delta function. In many cases the range of interest of the coordinate (s) covers energy differences much larger than kT, and then special sampling techniques are required, such as umbrella sampling,[8] steered dynamics,[9] adaptative force bias[10], or metadynamics.[11] In addition to sampling, the other important problem for the correct evaluation of the PMF associated to a particular transformation is the selection of an appropriate s coordinate allowing a smooth and meaningful transition from the initial to the final state and providing a good guess of the TS. Identifying reaction coordinates that satisfy the above requirements is not obvious. In some cases, a simple geometrical coordinate can be enough for this purpose. For example the typical S_N2 reaction can be followed using the antisymmetric combination of the distances of the substituted carbon atom to the nucleophile and the leaving group, while a simple S_N1 process can be described using the distance associated to the broken bond of the leaving group. More generally, an energy gap coordinate can be used in conjunction with a valence bond description of the reacting system.[12] Any of these strategies faces limitations to describe complex chemical reaction that involve TSs where many relevant valence coordinates change simultaneously or that

Received: January 31, 2012
Published: March 29, 2012

   dx.doi.org/10.1021/ct300070b | *J. Chem. Theory Comput.* 2012, 8, 1795−1801

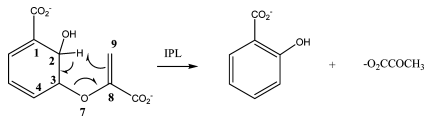
imply more than two valence bond states. This is normally the situation of a chemical reaction in which two or more chemical events take place simultaneously, as, for example, in concerted hydride and proton transfers.[13]

Several strategies have been proposed to trace free energy paths in terms of collective variables. Transition path sampling[14] and the string method[15] are viable approaches to obtain free energy information between an initial (reactants) and a final (products) state. The string method has been recently applied to study complex processes, such as the conversion between two different conformations of myosin VI.[16] Some related variants of the method have been introduced recently.[17−19] In this work we propose an adaptation of one of these strategies to be used in the context of chemical reactivity in complex environments, specifically in solution or enzymatic sites. We suggest as an adequate coordinate to be employed in this studies the advance along a minimum energy path (MEP) obtained for the reaction under study. The MEP is defined using a subset of coordinates of the system $\mathbf{q}(t)$. For a particular structure (t) with coordinates $\mathbf{q}(t)$, the advance along the MEP is then defined using the function introduced by Branduardi et al.:[18]

$$s(\mathbf{q}) = \lim_{\lambda\to\infty} \frac{\int_0^1 t\cdot e^{-\lambda(\mathbf{q}-\mathbf{q}(t))^2} dt}{\int_0^1 e^{-\lambda(\mathbf{q}-\mathbf{q}(t))^2} dt} \tag{3}$$

If the path is defined in such a way that $\mathbf{q}(t = 0)$ and $\mathbf{q}(t = 1)$ define the initial and the final reference structures, respectively, then $s(\mathbf{q})$ ranges from 0 to 1 in the sense of the advance of the chemical process. We here show that for the study of chemical reactions $\mathbf{q}(t)$ can be conveniently selected as a set of internal coordinates taken from structures lying on the MEP traced from a transition structure to the reactant and product valleys. With such a choice one can avoid the exploration of multidimensional free energy surfaces even in chemical transformations that involve TSs where more than two valence coordinates change simultaneously. The method will be tested with the study of the transformation of isochorismate into salycilate and pyruvate (see Scheme 1), in two different

Scheme 1.



environments, aqueous solution and the active site of *Escherichia coli* PchB, an enzyme with isochorismate pyruvate lyase (IPL) activity.[20,21]

This reaction was shown to proceed through a pericyclic mechanism in which C3O7 bond breaking and proton transfer from C2 to C9 take place in a concerted step,[22,23] with a single TS shown in Figure 1.

In a previous work theoretical exploration of the free energy landscape for the pyruvate lyase activity and the counterpart reaction in aqueous solution required of the use of at two geometrical coordinates the C3O7 distance (d_1, see Figure 1) and the antisymmetric combination of the distances of the transferred proton to C2 and C9 (d_5-d_4).[24] From the two-dimensional free energy surface the minimum free energy path was obtained, a TS structure was located, and the reaction mechanism was fully characterized. We here show that with a proper definition of a generalized reaction coordinate, this can be done using a monodimensional free energy profile or potential of mean force.

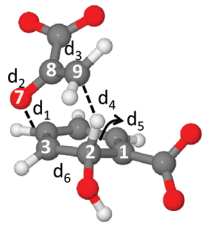


Figure 1. Transition structure for the IPL activity. Atom and bond distance numbering is also provided.

2. METHODOLOGY

2.1. Definition of the Coordinate. In order to be used as a collective coordinate in molecular dynamic simulations the $s(\mathbf{q})$ coordinate defined according to eq 1 must be discretized as

$$s(\mathbf{q}) = \frac{1}{P-1} \frac{\sum_{i=1}^{P} (i-1) e^{-\lambda(\mathbf{q}-\mathbf{q}(i))^2}}{\sum_{i=1}^{P} e^{-\lambda(\mathbf{q}-\mathbf{q}(i))^2}} \tag{4}$$

where P is the number of structures considered to represent the path. As pointed out by Branduardi et al.,[18] the structures $\mathbf{q}(i)$ should be as equidistant as possible, and λ should be comparable to the inverse value of the mean square displacement between successive structures:

$$\lambda \approx \frac{1}{(\mathbf{q}(i+1) - \mathbf{q}(i))^2} \tag{5}$$

In our implementation $\mathbf{q}(i)$ is a subset of coordinates of structures picked from the MEP of the reaction under consideration. The procedure employed to obtain this MEP is detailed below. In the original implementation $\mathbf{q}(i)$ was defined in terms of the Cartesian coordinates.[18] This definition was already used to analyze the free energy profile of a chemical reaction in a continuum solvent model.[25] However, this choice can be problematic when applied to the study of chemical reactions in discrete environments. In such cases one is usually interested to follow the change in a reduced subset of atoms, those defining the reacting system (in QM/MM studies this could usually match the QM subsystem). Hindered rotations and translational motions of the reacting subsystem could then lead to changes in $s(\mathbf{q})$ that do not reflect any advance in the reaction progress. We then propose that $\mathbf{q}(i)$ can be better defined as a subset of internal coordinates (distances, valence angles, and/or dihedral angles) reflecting the most relevant changes taking place during the chemical reaction of interest.

$$\mathbf{q}(i) = \{\mathbf{d}_j(i)\} \tag{6}$$

To be able to evaluate the free energy barrier from the PMF obtained using such a coordinate, the isosurfaces obtained for a given $s(\mathbf{q})$ value should keep perpendicular to the MEP. This can be assured by using a sufficiently large number of reference structures from the MEP and, subsequently, a large value for λ. However, in this case problems arise because the exponential terms may become extremely small, arising then numerical instabilities. To fix this, we propose to rewrite the initial eq 4 as

$$s(\mathbf{q}) = \frac{1}{P-1}\frac{\sum_{i=1}^{P}(i-1)e^{\gamma-\lambda(\mathbf{q}-\mathbf{q}(i))^2}}{\sum_{i=1}^{P}e^{\gamma-\lambda(\mathbf{q}-\mathbf{q}(i))^2}} \quad (7)$$

The modified equation is mathematically equivalent to the original one. Since γ can be any number, it can be chosen as follows:

$$\gamma = \min[\lambda(\mathbf{q}-\mathbf{q}(i))^2] \quad (8)$$

In this way, the largest exponential term under the summation will be equal to unity, avoiding the aforementioned numerical problem.

The definition of the $s(\mathbf{q})$ coordinate was implemented as a module in the fDYNAMO library[26] (details are provided in the Supporting Information). The PMF as a function of the $s(\mathbf{q})$ coordinate was obtained using the umbrella sampling technique.[27] The range of interest of values of the $s(\mathbf{q})$ coordinate was covered in a series of simulation windows centered at a particular value of the coordinate s_0 by means of a parabolic potential:

$$V_s = \frac{1}{2}K_s(s(\mathbf{q})-s_0)^2 \quad (9)$$

The value of the force constant, K_s, was chosen as a compromise between the control on the behavior of the coordinate and the efficiency of the simulations to sample the desired values. Note that being $s(\mathbf{q})$ dimensionless, the force constant here has energy units.

2.2. The Minimum Energy Path. In order to trace a MEP to be used for the definition of the $s(\mathbf{q})$ coordinate, we first localized a transition structure using the micro/macro iterations method.[28,29] In this method the coordinates of the system are divided into two subsets defined as the control and complementary spaces. The control space contains the more relevant coordinates of the system and usually contains, at least, the coordinates of the QM subsystem. For the control space, the Hessian is calculated, and the stationary structure search is carried out using a BAKER algorithm[30] keeping the complementary space frozen. At each step, this complementary space is minimized using a gradient-based algorithm.

The transition structures obtained in the enzymatic active site and in aqueous solution were used to trace the reaction paths up to the products and reactant valleys. A relaxed MEP was obtained following the gradient vector of the control space, while the complementary step is fully optimized at each new step. For the initial step around the transition structure, considering the low value of the gradient, the transition vector was followed instead.

Reactant and product structures were obtained after optimization from the end points of the MEP. In order to have better defined reactant and product valleys in the PMFs, some structures were extrapolated following the direction of the MEP in the vicinity of the minimum energy structures, so the initial and final structures in the set representing the MEP do not correspond to the minimum energy structures.

The set of internal coordinates defining $s(\mathbf{q})$ is obtained from the structures of the MEP. The coordinates are selected starting from the transition structure and moving to both the reactants and products sides. For each side of the MEP, we select the first structure (i) along the sequence in which the mean square deviation of the selected internal coordinates with respect to the transition structure values is larger than $1/\lambda$. Then a linear interpolation between the coordinates of structures i and $i-1$ of the MEP is used to obtain equispaced values that enter in the definition of the $s(\mathbf{q})$ coordinate. The procedure is repeated starting from this new set of coordinates and continuing the search along the MEP toward reactants and products.

2.3. QM/MM Simulations. The simulations carried out in this work were based in the models used in our previous analysis on the pyruvate lyase activity of PchB.[24] Briefly, the enzymatic reaction was modeled starting from the X-ray structure of PchB with PDB code 2H9D.[20] The hydrogen atoms were incorporated into the structure using fDYNAMO.[26] After that, sodium counterions were added to neutralize the total charge. The resulting system was placed in a simulation water box of 79.5 Å of side, removing any water molecule with an oxygen atom lying within a radius of 2.8 Å of any nonhydrogen atom. The substrate, see Scheme 1, was described using quantum mechanics at the AM1 level,[31] while for the rest of the system we used the OPLS-AA[32] and TIP3P[33] force fields. Periodic boundary conditions were employed combined with a force switched cutoff radius defined between 14.5 and 16 Å for the nonbonded interactions. The system was equilibrated at 300 K using the NVT ensemble and the Langevin–Verlet integrator with a time-step of 1.0 fs during a total time of 500 ps. To study the chemical reaction in aqueous solution, we used a similar protocol, placing the QM subsystem in a box of 55.8 Å of side of TIP3P water molecules. These models were employed to obtain the free energy surface corresponding to the reaction under analysis using as distinguished coordinates the distance of the C–O bond to be broken (d_1 in Figure 1) and the antisymmetric combination of the distances of the transferred proton to the donor and acceptor carbon atoms (d_5–d_4).[24] The TSs structures localized on the potential energy surfaces obtained in both PchB and aqueous solution were the starting points for this work.

Free energy profiles along the $s(\mathbf{q})$ coordinate were obtained using 70 simulation windows using umbrella sampling with a parabolic restraining potential (eq 9) with a force constant of 10^5 kJ·mol^{-1}. With this value of the force constant, the oscillations of the coordinate can be kept centered around the target value. In addition this value is a good compromise between the number of windows and the time per window required to sample adequately the whole range of interest of the coordinate. Figure S1, Supporting Information, shows the behavior of $s(\mathbf{q})$ with different values of the force constant. Each simulation window consisted in 5 ps of equilibration followed by 20 ps of production. The whole probability distribution of the coordinate was finally recovered using WHAM.[34] An example of the distribution of values of the $s(\mathbf{q})$ coordinate obtained along the PMF is provided in Figure S2, Supporting Information.

3. RESULTS

3.1. The Enzymatic Reaction. We first investigated the PMF for the enzymatic reaction employing two different definitions of the coordinate. In the first one we included all the distances involved in the cycle formed in the TS of the reaction (see Figure 1):

$$s_1(\mathbf{q}) = s(d_1, d_2, d_3, d_4, d_5, d_6) \quad (10)$$

We also tested a second definition where we only included the distances of those bonds broken or formed during the process:

$$s_2(\mathbf{q}) = s(d_1, d_4, d_5) \quad (11)$$

The behavior of the discretized version of the $s(\mathbf{q})$ coordinate was monitored using different values of λ, which in turns determines the number of structures (P in eq 4) included in the definition of the reference path. We obtained the corresponding PMFs for λ = 50, 200, 500, and 1000 Å^{-2} that roughly corresponds to P = 20, 45, 70, and 100, respectively.

Figure 2 shows the PMFs obtained using $s_1(\mathbf{q})$ together with the MEP used as reference. The curves obtained using $s_2(\mathbf{q})$ are

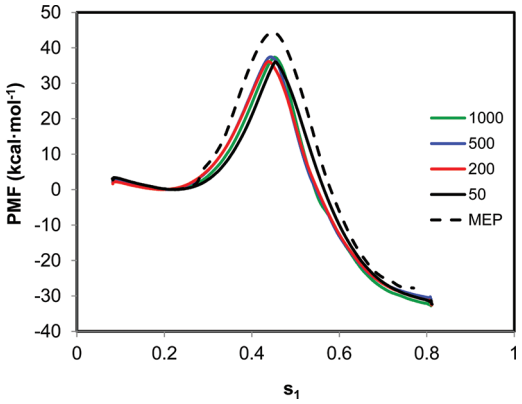


Figure 2. PMFs and MEP for the isochorismate transformation into pyruvate and salycilate in PchB as a function of the $s_1(\mathbf{q})$ coordinate and different values of λ (in Å^{-2}).

almost identical and are given in Figure S3, Supporting Information. The free energy barriers directly estimated as the difference between the PMF values at the maximum and the reactants minimum are provided in Table 1 for different λ values. In any case, the results are quite close and within the expected statistical error.

Table 1. PMF Differences between the Maximum and the Reactants Minimum (in kcal·mol^{-1}) Obtained Using Different Values of λ (in Å^{-2})

λ	$\Delta G^{\ddagger}$ (s_1)
50	36.0
200	36.1
500	37.4
1000	37.3

The free energy barriers provided in Table 1 compare well with the value derived from the aforementioned free energy surface traced as a function of d_1 and d_5-d_4 using the same AM1/MM potential, 38.4 kcal·mol^{-1}.[24] Note that free energy barriers derived from differences in PMFs traced using different coordinates are not necessarily identical.[7] The activation free energy derived from the experimental rate constant is significantly smaller, 17.7 kcal·mol^{-1}.[35] The discrepancy between the theoretical and experimental estimations can be completely attributed to the use of the AM1 Hamiltonian, and single point corrections at higher energy levels are enough to bring the calculated value to much closer agreement with experimental estimations.[24]

Figure 3 shows the evolution of the most relevant bond distances (d_1, d_4, and d_5) along the PMFs obtained for $s_1(\mathbf{q})$ and λ = 50, 200, 500, and 1000 Å^{-2}. The data represented in this figure correspond to values averaged at each simulation window for the different PMFs. In all cases, the plots of the distances versus the reaction coordinate described a concerted process where the proton transfer from C2 to C9 slightly precedes the C3−O7 bond breaking. At the TS, which appears at s_2 = 0.44−0.45 and is marked with a vertical dashed line in Figure 3, the proton transfer is nearly halfway, while the C−O bond breaking is in the first stages. It must be noted that the chemical description of the process depends on the QM level employed. As showed in our previous work on this reaction, the AM1 Hamiltonian favors a concerted process, while at the B3LYP level, the reaction is clearly asynchronous. Using B3LYP corrections, the heterolytic cleavage of the carbon−oxygen bond takes place first by generating a negatively charged enolpyruvyl side chain and a positively charged ring, which in turn favors the proton transfer from C2 to C9.[24] However, the problems associated to the QM level selected for the description of the substrate are out of the scope of this paper, which is focused on the definition of an adequate coordinate for a monodimensional description of the reaction process.

With respect to the effect of the environment, the reaction is assisted by the presence of a positively charged residue in the active site, Lys42, which stabilizes the charge developed on the O7 atom. A representative snapshot of a TS structure is given in Figure 4. The process is also assisted by other residues of the active site (Arg14, Arg31, and Arg53) that facilitate the correct positioning of the substrate by means of interactions with the carboxylate groups of the substrate. These residues also play an important role in catalysis allowing the approach of the two carboxylic groups during the reaction process.

The influence of the selected value for λ can be observed comparing the plots appearing in Figure 3. While in all cases the three key distances involved in the bond breaking and forming processes follow a similar evolution, there are some differences near the TS region. In particular, the behavior of C2−H (d_5) and C9−H (d_4) distances for λ = 50 and 200 Å^{-2} is anomalous at the TS region, with the proton coming back to the donor atom (C2) during a short range of s_1 values. This evolution of the proton transfer with the reaction coordinate is not observed for larger λ values (500 and 1000 Å^{-2}). The change in the C3−O7 distance is also smoother when λ increases. So, while the free energy barrier is quite well reproduced even for small values of λ (see Table 1), a correct geometrical control of the process seems to require larger λ values. Similar trends were observed for the $s_2(\mathbf{q})$ coordinate, and the corresponding graphics are given in Figure S4, Supporting Information.

3.2. The Reaction in Aqueous Solution. We also tested the performance of the implementation of the $s(\mathbf{q})$ coordinate in terms of internal coordinates to study the same reaction in aqueous solution. In this case the PMF was traced only for λ = 500 Å^{-2} with both definitions of the coordinate, including 6 or 3 bond distances. Only the results for the first case, $s_1(\mathbf{q})$, will be presented in the paper, while results for the second, $s_2(\mathbf{q})$, are given in Figure S5, Supporting Information. The MEP used to construct the reaction coordinate and the free energy profile obtained is presented in Figure 5. The free energy barrier is 42.1 kcal·mol^{-1}, a value in good agreement with the free energy barrier derived from the two-dimensional free energy surface traced as a function of d_1 and d_5-d_4, 43.1 kcal·mol^{-1}.[24] Again these free energy barriers are too high when compared to the value derived from the experimental rate constant in aqueous solution, 26.2 kcal·mol^{-1}.[23] Differences are due to the use of the AM1 semiempirical Hamiltonian. However, as was already mentioned, the purpose of this work is not to describe

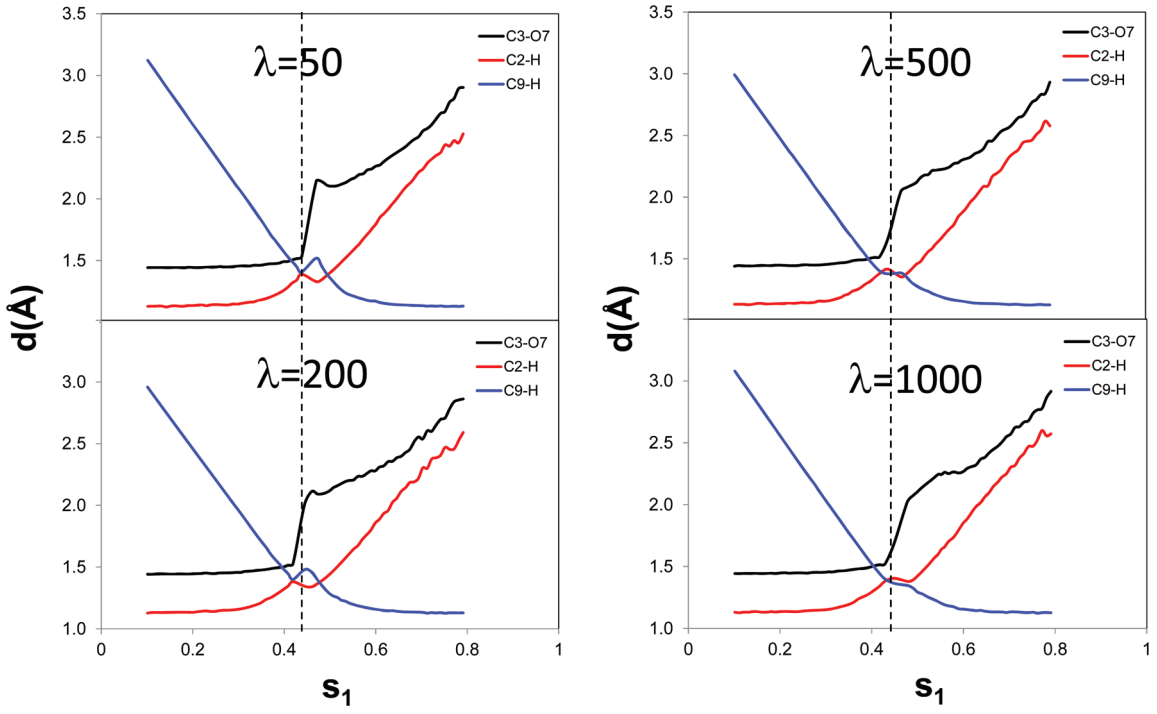


Figure 3. Averaged values of distances corresponding to bonds broken or formed during the reaction as a function of the reaction coordinate. Each plot corresponds to different values of λ (in Å^{-2}). Position of the TS is shown with a vertical dashed line.

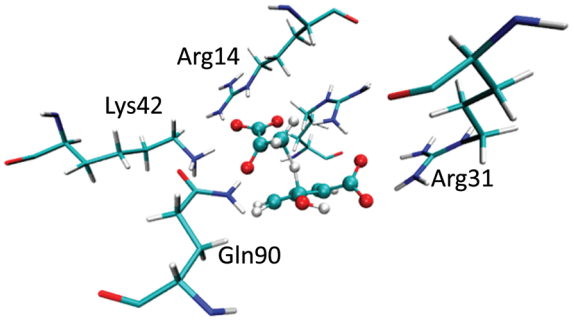


Figure 4. Representative snapshot of the TS structure for IPL activity in the active site of PchB.

accurately the energetics of the reaction but to show the performance of our definition of the $s(\mathbf{q})$ coordinate to obtain good estimations of the activation free energies using only monodimensional PMFs for different environments (enzymatic and aqueous solution), even in those cases that usually would require two-dimensional treatments because they involve more than a single chemical event.

Evolution of distances associated to the bonds to be broken/formed during the reaction is shown in Figure 6. The geometric

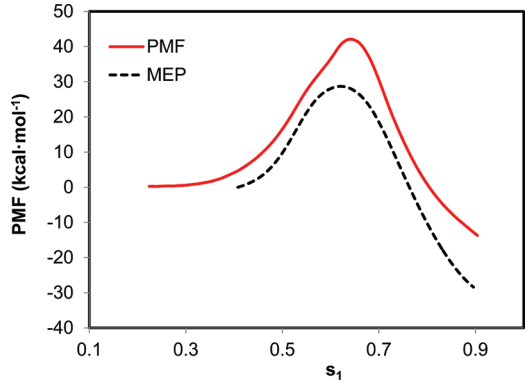


Figure 5. PMF and MEP for the isochorismate transformation into pyruvate and salycilate in aqueous solution as a function of the $s_1(\mathbf{q})$ coordinate with λ = 500 Å^{-2}.

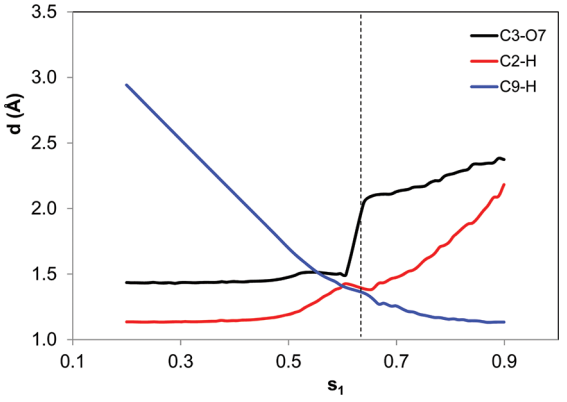


Figure 6. Evolution of the averaged values of relevant distances corresponding to bonds broken or formed during the reaction in aqueous solution as a function of the $s_1(\mathbf{q})$ reaction coordinate with λ = 500 Å^{-2}. Position of the TS is indicated by a vertical dashed line.

description of the reaction at the AM1/MM level in aqueous solution is similar to that obtained in the active site of PchB. At the TS, both the proton transfer between C2 and C9 and the C3−O7 bond breaking are slightly more advanced than for the enzymatic process. So, the TS in solution appears later than for the enzymatic process, which is consistent, in the sense of the Hammond postulate, with the higher free energy barrier observed in the former environment. A very similar evolution is observed when the coordinate is defined using only three

dx.doi.org/10.1021/ct300070b | J. Chem. Theory Comput. 2012, 8, 1795−1801

distances (see Figure S6, Supporting Information). As in the case of the enzymatic reaction, the evolution of these distances is smooth for the selected value of λ, and the methodology seems also adequate to describe the reaction in a very different environment.

4. CONCLUSIONS

The choice of a proper coordinate is a crucial problem to properly sample the relevant configurations of a system and then to obtain free energy profiles. Chemical reactions often imply important changes in more than a single valence coordinate or combination of coordinates. In such cases, PMFs traced as a function of a single distinguished coordinate can lead to severe hysteresis problems. An obvious solution is to obtain free energy surfaces as a function of two coordinates, which obviously increases significantly the computational time required. For that reason we have explored in this work the implementation of a generalized coordinate that can be used to obtain monodimensional free energy profiles even in those cases where more than a chemical event is taking place in a single step. This coordinate is defined as the advance along a path expressed as a function of a subset of internal coordinates. The price to be paid is the exploration of the potential energy surface from which a reasonable path can be proposed. The procedure presented in this work is based on the use of the reaction MEP, which requires the localization and characterization of a transition structure. This is made using the micro/macro iteration algorithm. From a set of structures obtained from this MEP, a reaction coordinate is then defined in terms of the mean square deviation with respect to a subset of relevant internal coordinates. For reactions taking place in condensed environments, this definition is more convenient than the use of Cartesian coordinates because problems associated to pseudotranslations and pseudorotations of the reacting subsystem are avoided.

The coordinate definition has been tested for the reaction leading from isochorismate to pyruvate. The reaction occurs by means of a pericyclic mechanism in which C−O bond breaking and proton transfer between carbon atoms take place in a single step. The free energy barriers obtained both in the active site of PchB and in aqueous solution are in good agreement with previous estimations based in the construction of two-dimensional free energy surfaces, a procedure much more computationally demanding. We believe that the coordinate defined here can be useful to explore complex reaction processes in solution or enzymatic environments.

■ ASSOCIATED CONTENT

Supporting Information

Figures showing the evolution of the s coordinate for different choices of the umbrella force constant, the distribution of s values used to obtain a PMF, figures presenting the PMFs and evolution of distances obtained for the enzymatic and the uncatalyzed reaction using the s_2 definition of the coordinate, and a short description of the implementation of the procedure in fDynamo. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ignacio.tunon@uv.es.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by Ministerio de Ciencia e Innovación, project CTQ2009-14541-C02 and by Generalitat Valenciana ACOMP/2011/028. K.Z. acknowledges an Erasmus Mundus fellowship of the European Commission (European Master in Theoretical and Computational Chemistry).

■ REFERENCES

(1) Callen, H. B. In *Thermodynamics and an Introduction to Thermostatistics*; John Wiley & Sons: Hoboken, NJ, 1985.
(2) Chandler, D. In *Introduction to Modern Statistical Mechanics*; Oxford University Press: New York, 1987.
(3) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(4) Rosta, E; Woodcock, H. L.; Brooks, B. R.; Hummer, G. *J. Comput. Chem.* **2009**, *30*, 1634−1641.
(5) Laidler, K. J.; King, M. C. *J. Phys. Chem.* **1983**, *87*, 2657−2664.
(6) Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 12771−12800.
(7) Schenter, G. K.; Garrett, B. C.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*, 5828−5833.
(8) Patey, G. N.; Valleau, J. P. *J. Chem. Phys.* **1975**, *63*, 2334−2339.
(9) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.
(10) Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, *115*, 9169−9183.
(11) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *20*, 12562−12566.
(12) Warshel., A. *J. Phys. Chem.* **1982**, *86*, 2218−2224.
(13) Ferrer, S.; Tuñón, I.; Martí, S.; Moliner, V.; Garcia-Viloca, M; Gonzalez-Lafont, A.; Lluch, J. M. *J. Am. Chem. Soc.* **2006**, *128*, 16851−16863.
(14) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Faraday Discuss.* **1998**, *110*, 421−436.
(15) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 024106.
(16) Ovchinnikov, V.; Karplus, M.; Vanden-Eijnden, E. *J. Chem. Phys.* **2011**, *134*, 085103.
(17) Khavrutskii, I. V.; McCammon, J. A. *J. Chem. Phys.* **2007**, *127*, 124901.
(18) Branduardi, D.; Gervasio, F. L.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 054103.
(19) Pan, A. C.; Sezer, D.; Roux, B. *J. Chem. Phys. B* **2008**, *112*, 3432−3440.
(20) Zaitseva, J.; Lu, J.; Olechoski, K. L.; Lamb, A. L. *J. Biol. Chem.* **2006**, *281*, 33441−33449.
(21) Luo, Q.; Olucha, J.; Lamb, A. L. *Biochemistry* **2009**, *48*, 5239−5245.
(22) Künzler, D. E.; Sasso, S.; Gamper, M.; Hilvert, D.; Kast, P. *J. Biol. Chem.* **2005**, *280*, 32827−32834.
(23) DeClue, M. S.; Baldridge, K. K.; Kast, P.; Hilvert, D. *J. Am. Chem. Soc.* **2006**, *128*, 2043−2051.
(24) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *J. Am. Chem. Soc.* **2009**, *131*, 16156−16161.
(25) Branduardi, D.; De Vivo, M.; Rega, N.; Barone, V.; Cavalli, A. *J. Chem. Theory Comput.* **2011**, *7*, 539−543.
(26) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088−1100.
(27) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(28) Moliner, V.; Turner, A. J.; Williams, I. H. *J. Chem. Soc. Chem. Comun.* **1997**, 1271−1272.
(29) Martí, S.; Moliner, V.; Tuñón, I. *J. Chem. Theory Comput.* **2005**, *1*, 1008−1016.
(30) Baker, J. *J. Comput. Chem.* **1986**, *7*, 385−395.
(31) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(32) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(34) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(35) Gaille, C.; Kast, P.; Haas, D. *J. Biol. Chem.* **2002**, *277*, 21768−21775.