JCTC Journal of Chemical Theory and Computation

Letter

pubs.acs.org/JCTC

A Novel Strategy to Study Electrostatic Effects in Chemical Reactions: Differences between the Role of Solvent and the Active Site of Chalcone Isomerase in a Michael Addition

J. Javier Ruiz-Pernía,[†] Sergio Martí,[†] Vicent Moliner,*[,†] and Iñaki Tuñón*[,‡]

[†]Departament de Química Física i Analítica, Universitat Jaume I, 12071 Castelló, Spain

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

Supporting Information

ABSTRACT: The electrostatic behavior of active site residues in enzyme catalysis is quite different from that of water molecules in solution. To highlight the electrostatic differences between both environments, we propose a QM/MM strategy to study the role of the environment in chemical reactions. The novelty of the present communication is that free energy surfaces are generated by means of two distinguished reaction coordinates: a solute coordinate and the electrostatic potential created by the environment. This is applied to analyze the origin of catalysis in the transformation of a chalcone into a flavanone, a Michael addition that requires the desolvation of the nucleophile.

The physical properties of a chemical reaction can be dramatically affected by the environment. However, there is much that remains to be understood about the influence of the solvent on reaction rates. Recent experimental measurements of the time-resolved vibrational spectra of the ketoesteroid isomerase site suggest that the electrostatic response of the active site residues is quite restricted, a behavior completely different from that found in typical solvents.[1] This experiment would agree with our simulations showing that the electric field acting on the reacting subsystem in enzymatic active sites remains essentially constant, while the counterpart reaction in aqueous solution would take place under the influence of a fluctuating reaction field.[2]

Understanding the origin of enzyme catalysis requires dissection of the role of the environment for the catalyzed reaction and the counterpart process in aqueous solution by means of appropriately designed simulations.[3] The most relevant kinetic and thermodynamic quantities for any chemical reaction, in solution or in an enzymatic environment, are the activation and the reaction free energies. These magnitudes can be derived from the corresponding free energy profile or potential of mean force (PMF). Usually, the effect of the environment is added into the simulations assuming that the movements of the solvent degrees of freedom are fast enough to adapt to any instantaneous variation in the chosen reaction coordinate. Obviously, this equilibrium approximation is questionable because the redistribution of the electron density associated with any chemical reaction must be coupled to variations in the interactions established with the surroundings of the chemical system.[2−4] In order to take into account the coupling between the solute and the environment, one can follow the reaction progress using a collective reaction coordinate involving not only the reacting system but also the solvent degrees of freedom. The energy gap between the reactant and the product states of a given reaction has been used as a single collective coordinate to drive a particular reaction and then to obtain the corresponding free energy profile.[5] In addition, N-dimensional free energy surfaces (FES), including both solute and solvent coordinates in an equal footing treatment, can be generated. In its simplest and more affordable version, one can trace a two-dimensional surface (2D-FES) using a solute coordinate (r) and a solvent coordinate (s).[5] This 2D-FES can be expressed as

$$W(r, s) = C' - kT \ln \int \rho(x^N)\, \delta(r(x^N) - r)\, \delta(s(x^N) - s)\, dx^N$$

where $\rho(x^N)$ is the probability density of finding the system at a particular configuration (x^N). The Dirac δ functions allow this to be evaluated only at the desired values of the selected coordinates, and C' is an integration constant. 2D-FESs, obtained using for example the solute and solvent energy gaps[6] or related coordinates,[7] allow for addressing explicitly the question of the participation and the timing of the solvent motions in the reaction progress. This is especially important in order to understand the differences between the free energy barriers of a chemical reaction in solution and those in an enzymatic active site and thus to rationalize the origin of catalysis.

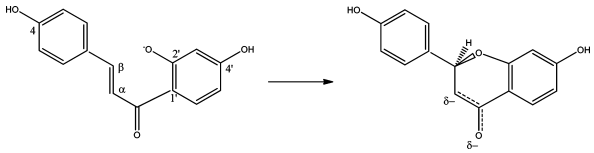
In this work, we propose a strategy to generate 2D-FESs by means of two distinguished reaction coordinates: a solute coordinate and a solvent coordinate. For the latter, the electrostatic potential created by the environment over one particular atom or group of atoms involved directly in the chemical reaction is proposed as an adequate generalized solvent coordinate. This coordinate can be used to highlight the different behavior, from the electrostatic point of view, between the aqueous environment and an enzyme. This definition of the solvent coordinate can be implemented in any QM/MM code and working at any QM level. Note that while the energy gap is

Received: January 30, 2012
Published: April 2, 2012

adequate for valence bond treatments, it is not easily adapted for working with molecular orbital methods, which are the most popular ones in the study of chemical reactions in condensed media (proteins or solvents). Obviously, the only limitation is the cost of the QM/MM simulations, which greatly depends on the CPU time needed for each QM cycle (note that a 2D-FES may require an order of magnitude of 10^8 energy and gradient evaluations).

In this work, we present an application of this strategy to the chemical transformation of chalcones into isoflavanones catalyzed by chalcone isomerase (CHI).[8−11] The rate-limiting step of this process consists of an intramolecular Michael addition of the nucleophilic oxygen (O2′) of the substrate (6′-deoxychalcone) on the β-carbon of the double bond (see Scheme 1).[8−10] We explore the catalyzed and uncatalyzed

Scheme 1



processes using the electrostatic potential on the O2′ atom as a generalized solvent coordinate (s):

$$s = \sum_{j=1}^{j=M} \frac{q_{\mathrm{MM}}^{j}}{|r_j - r_{\mathrm{O2'}}|}$$

where j runs over the M sites of the environment with charge q, while the O2′−C$^{\beta}$ distance is an adequate solute coordinate (r). It is important to point out that, while it is well-known that point charges may incorrectly polarize the wave function of the QM subsystem at medium and short distances,[12−14] this is the basic approximation for the treatment of the MM subsystem in most of the currently used QM/MM codes, and then we will keep it in our definition of the solvent coordinate. In principle, other definitions of the solvent coordinate based on the electrostatic potential, such as a combination of the potential over a subset of atoms, could be designed to study this reaction. However, the present strategy allows focusing on the role of the environment on the nucleophilic atom along the full chemical reaction.

Molecular dynamics simulations are carried out using hybrid QM/MM potentials where the substrate is described at the AM1 level,[15] the protein using the OPLS-AA force field,[16] and water molecules by means of the TIP3P model.[17] The umbrella sampling technique is employed to explore the relevant configurations of the system at the desired values of the r and s coordinates during the simulations. Computational details are provided as Supporting Information.

The 2D-FESs obtained for the transformation of 6′-deoxychalcone into (2S)-5-deoxyflavanone are presented in Figure 1. As observed, the processes in aqueous solution (Figure 1a) and in the active site of CHI (Figure 1b) are completely different. The reactant state in aqueous solution is found at long O2′−C$^{\beta}$ distances ($r \approx 3.2$ Å) and large positive electrostatic potentials on the O2′ atom ($s \approx 150$ kcal·mol^{-1}·e^{-1}). The nucleophile is a negatively charged oxygen atom, and then the charge on this atom is stabilized by means of strong hydrogen bonds with solvent molecules at equilibrium. The product state is found at a O2′−C$^{\beta}$ distance

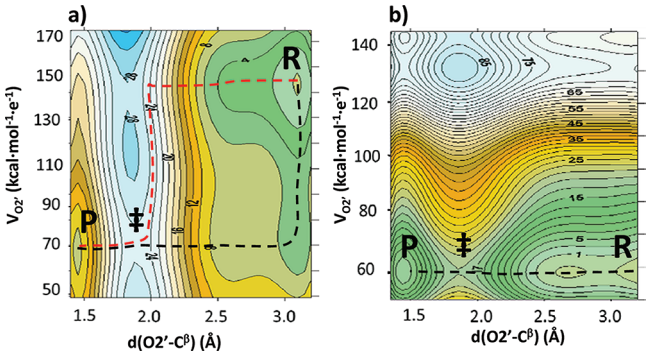


Figure 1. Free energy surfaces corresponding to the transformation of chalcone to flavanone (a) in aqueous solution and (b) in CHI. The distinguished reaction coordinates are the O2′−C$^{\beta}$ distance and the electrostatic potential created by the environment on the nucleophile, $V_{O2'}$. Each isoenergetic line represents a 2 kcal mol^{-1} increase in free energy. The black dashed lines represent the minimum free energy path on the free energy surfaces, while the red dashed line in part a is the path obtained in aqueous solution assuming equilibrium solvation along the solute coordinate. R, ‡, and P refer to reactants, transition state, and products, respectively.

corresponding to a formed bond ($r \approx 1.4$ Å) and significantly smaller values of the electrostatic potential on the O2′ atom ($s \approx 70$ kcal·mol^{-1}·e^{-1}), in accordance with the fact that the negative charge has been now delocalized on the C$^{\alpha}$ and the carbonyl oxygen atoms (see Scheme 1). The timing between the approach of the nucleophile to the C$^{\beta}$ atom and the desolvation of O2′ can be deduced from the minimum free energy path (black dashed line in Figure 1a), concluding that the nucleophilic O2′ atom must be desolvated prior to the intramolecular addition. We have analyzed the solvation around the O2′ atom for reactant structures located at high and low values of the electrostatic potential. Simulations were carried out on the 2D-FES with reference values ($r = 3.2$ Å and $s = 140$ kcal·mol^{-1}·e^{-1}; $r = 3.2$ Å and $s = 70$ kcal·mol^{-1}·e^{-1}). The radial distribution functions (RDFs) of solvent molecules around the O2′ atom in these two cases are significantly different (see the Supporting Information, Figure S1). At high values of the electrostatic potential, we found a well-defined solvation shell around the negatively charged O2′ atom with a sharp peak centered at an O2′−O$_{\mathrm{W}}$ distance of 2.8 Å that corresponds to the formation of strong hydrogen bonds between the nucleophile and water molecules. At low values of the electrostatic potential, we find a broader and lower peak centered at longer distances (3.5 Å), reflecting the desolvation of this atom. Representative snapshots of these two different environments are also provided as Supporting Information (Figure S2).

The activation free energy in solution obtained after integration on the $W(r,s)$ surface (see Computational Details in Supporting Information) is 25.7 kcal·mol^{-1}. According to the minimum free energy path traced on $W(r,s)$, the activation free energy in solution can be decomposed into two contributions: the free energy associated with the change in the solvent coordinate (s) and the free energy cost due to the change in the solute coordinate (r):

$$\Delta G_{\mathrm{aq}}^{\ddagger} = \Delta G_{s,\mathrm{aq}}(s_{\mathrm{aq}}^{\mathrm{R}} \rightarrow s_{\mathrm{aq}}^{\ddagger}) + \Delta G_{r,\mathrm{aq}}(r_{\mathrm{aq}}^{\mathrm{R}} \rightarrow r_{\mathrm{aq}}^{\ddagger})$$

The first term makes a contribution of about 6 kcal·mol^{-1} to the free energy barrier. That is, desolvation of the nucleophile is

dx.doi.org/10.1021/ct300064f | J. Chem. Theory Comput. 2012, 8, 1532−1535

responsible for approximately 24% of the free energy barrier of this reaction in solution.

It is worthy to note that if the reaction were studied tracing the free energy profile only as a function of the solute coordinate (these profiles are shown in Figure S3 in the Supporting Information), the description obtained would be quite different. Running the simulations with just the O2′−C^{β} distance as a reaction coordinate would imply that the solvent is equilibrated at each step. The equilibrium path (red dashed line in Figure 1a) is then quite different from the minimum free energy path traced on $W(r,s)$ (black dashed line in Figure 1a), which is an example of the difficulties associated with the use of one single distinguished coordinate.[18] The approach of the nucleophile to the C^{β} atom precedes its desolvation in the equilibrium description.

The reaction in the enzyme, Figure 1b, proceeds at a nearly constant value of the electrostatic potential. The enzyme provides an adequate environment for the reaction to progress, and the reactant is already found at a value of the electrostatic potential that is the value required to reach the transition state. Note that the value obtained at the transition state in both environments is almost equivalent. On the contrary, the smaller value found for the potential on the O2′ atom in the enzyme active site cannot be directly interpreted as a substrate desolvation. While the electrostatic potential on the nucleophilic oxygen atom is reduced at the enzymatic reactant state, the electrostatic potentials on other atoms, such as C^{α} and the carbonyl oxygen atoms, are increased with respect to the aqueous solution. A water molecule and a positively charged lysine (Lys94) provide this larger electrostatic potential.[19]

According to Figure 1b, a reactant-like structure similar to the one obtained in aqueous solution (long O2′−C^{β} distances and high values of the electrostatic potential on the O2′ atom, $s \approx 140$ kcal·mol^{-1}·e^{-1}) would have a free energy more than 80 kcal·mol^{-1} higher than the one located at low values of the potential ($s \approx 60$ kcal·mol^{-1}·e^{-1}). Figure 2 compares the structures of the active sites (substrate plus closer residues and water molecules) found at high and low values of the electrostatic potential in the active site of the enzyme. Higher electrostatic potentials on the nucleophilic oxygen atom can be attained by means of a rotation around the C(O)−C1′ single bond, approaching the O2′ atom to Lys94. This rotation also implies the displacement of some side-chains of residues found in the active site, such as Met188. The Asn110 side-chain must also be displaced to allow some additional water molecules to come into the active site. This deformation of the active site implies a significant free energy penalty as observed in Figure 1b.

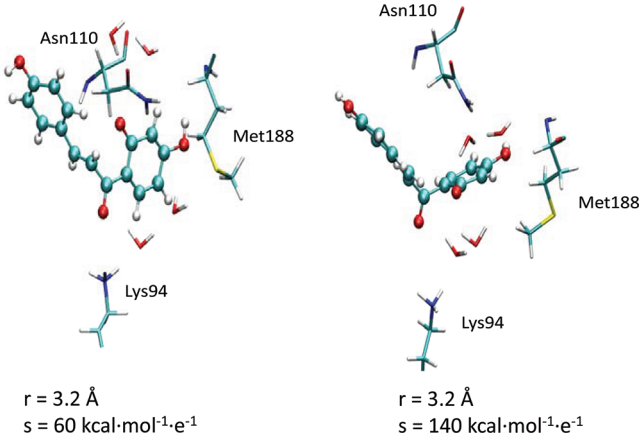


Figure 2. Representative snapshots of CHI active site corresponding to low (left) and high (right) values of the electrostatic potential on the nucleophile.

The activation free energy for the catalyzed process obtained from the $W(r,s)$ surface is 21.5 kcal·mol^{-1}. This value slightly overestimates the activation free energy value derived from the experimental rate constant, 16.0 kcal·mol^{-1}.[10] The difference can be attributed to deficiencies in the AM1 treatment, as demonstrated in our previous work[19] and by means of single point energy calculations at the MP2 level on the structures obtained at the AM1 level (see Table S2 in the Supporting Information). In fact, the predicted catalytic effect, where these kinds of errors are partially canceled out since the effect is computed as the difference between the catalyzed and noncatalyzed reaction, is 4.2 kcal·mol^{-1}, close to that derived from experimental data, 5.7 kcal·mol^{-1}.[8,10] It is very interesting to note that the enzymatic and the aqueous solution free energy barriers are very similar if the previously mentioned 6 kcal·mol^{-1} contribution due to the change in the solvent coordinate is removed from the estimation in solution. Then, catalysis can be mostly assigned to the energetic cost associated with the reorganization of the environment in solution. The enzymatic reaction takes place within a protein structure that provides the adequate electrostatic conditions. The cost associated with the creation of such an environment would be included in the folding process.[20]

■ ASSOCIATED CONTENT

Supporting Information

Computational details, radial distribution functions, representative snapshots of chalcone reactants in aqueous solution, PMFs for the Michael addition, fluctuations of the electrostatic potential, distribution of electrostatic potential values, Cartesian coordinates of a transition and a reactant structure, and gas phase energy barriers. 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, moliner@uji.es.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from MICINN through project CTQ2009-14541-C02, by Generalitat Valenciana, *Prometeo*/2009/053 and ACOMP/2011/038, and by Universitat Jaume I-Bancaixa, projects P1·1A2010-08 and P1·1B2011-23. J.J.R-P. thanks Ministerio Ciencia e Innovación for a "Juan de la Cierva" contract. The authors acknowledge computational facilities of the Servei d'Informàtica de la Universitat de València in the "Tirant" supercomputer.

■ REFERENCES

(1) Jha, S. K.; Ji, M.; Gaffney, K. J.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16612−16617.

(2) (a) Soriano, A.; Silla, E.; Tuñón, I.; Ruiz-López, M. F. *J. Am. Chem. Soc.* **2005**, *127*, 1946−1957. (b) Roca, M.; Andres, J.; Moliner, V.; Tuñón, I.; Bertran, J. *J. Am. Chem. Soc.* **2005**, *127*, 10648−10655. (c) Ruiz-Pernia, J. J.; Tuñón, I.; Moliner, V.; Hynes, J. T.; Roca, M. *J. Am. Chem. Soc.* **2008**, *130*, 7477−7488.

(3) (a) Warshel, A.; Sussman, F.; Hwang, J. K. *J. Mol. Biol.* **1988**, *201*, 139−159. (b) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. *Chem. Rev.* **2006**, *106*, 3210−3235. (c) Marti, S.; Roca, M.; Andres, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertran, J. *Chem. Soc. Rev.* **2004**, *33*, 98−107. (d) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.
(4) (a) Hwang, J. K.; King, G.; Creighton, S.; Warshel, A. *J. Am. Chem. Soc.* **1988**, *110*, 5297−5311. (b) Olsson, M. H. M.; Parson, W.; Warshel, A. *Chem. Rev.* **2006**, *106*, 1737−1756. (c) Gertner, B. J.; Bergsma, J. P.; Wilson, K. R.; Lee, S. Y.; Hynes, J. T. *J. Chem. Phys.* **1987**, *86*, 1377−1386. (d) Poulsen, T. D.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. B* **2003**, *107*, 9567−9578. (e) Roca, M.; Moliner, V.; Tuñón, I.; Hynes, J. T. *J. Am. Chem. Soc.* **2006**, *128*, 6186−6193.
(5) (a) Warshel, A. *J. Phys. Chem.* **1982**, *86*, 2218−2224. (b) Warshel, A.; Hwang, J. K. *J. Chem. Phys.* **1986**, *84*, 4938−4946.
(6) Olsson, M. H. M.; Siegbahn, P. E. M.; Warshel, A. *J. Am. Chem. Soc.* **2004**, *126*, 2820−2825.
(7) Hu, H.; Yang, W. T. *J. Phys. Chem. B* **2010**, *114*, 2755−2759.
(8) Furlong, J. J. P.; Nudelman, N. S. *J. Chem. Soc., Perkin Trans. 2* **1985**, 633−639.
(9) Jez, J. M.; Bowman, M. E.; Dixon, R. A.; Noel, J. P. *Nat. Struct. Biol.* **2000**, *7*, 786−791.
(10) Jez, J. M.; Noel, J. P. *J. Biol. Chem.* **2002**, *277*, 1361−1369.
(11) (a) Hur, S.; Bruice, T. C. *J. Am. Chem. Soc.* **2003**, *125*, 1472−1473. (b) Hur, S.; Newby, Z. E. R.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 2730−2735.
(12) Cisneros, G. A.; Piquemal, J. P.; Darden, T. A. *J. Phys. Chem. B* **2006**, *110*, 13682−13684.
(13) Biswas, P. K.; Gogonea, V. *J. Chem. Phys.* **2008**, *129*, 154108−154113.
(14) (a) Zhang, Y.; Lin, H.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 1378−1398. (b) Wang, B.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2011**, *13*, 10556−10564.
(15) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(16) (a) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666. (b) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474−6487.
(17) Jorgensen, W. L.; Chandrasekhar, J; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(18) Williams, I. H.; Maggiora, G. M. *THEOCHEM* **1982**, *89*, 365−378.
(19) (a) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I. *J. Am. Chem. Soc.* **2007**, *129*, 9117−9124. (b) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I. *J. Phys. Chem. B* **2006**, *110*, 20686−20692.
(20) (a) Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250−5254. (b) Roca, M.; Messer, B.; Hilvert, D.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 13877−13882.