THE JOURNAL OF PHYSICAL CHEMISTRY B LETTERS

**2005,** *109,* 3707−3710
Published on Web 02/10/2005



# Computing Kinetic Isotope Effects for Chorismate Mutase with High Accuracy. A New DFT/MM Strategy

**Sergio Martí,[†] Vicent Moliner,*[,†] Iñaki Tuñón,*[,‡] and Ian H. Williams[§]**

*Departament de Ciències Experimentals, Universitat Jaume I, Box 224, 12080 Castellón, Spain, Departamento de Química Física, Universidad de Valencia, 46100 Burjassot, Spain, and Department of Chemistry, University of Bath, BA2 7AY Bath, United Kingdom*

*Received: December 9, 2004; In Final Form: January 27, 2005*

A novel procedure has been applied to compute experimentally unobserved intrinsic kinetic isotope effects upon the rearrangement of chorismate to prephenate catalyzed by *B. subtilis* chorismate mutase. In this modified QM/MM approach, the "low-level" QM description of the quantum region is corrected during the optimization procedure by means of a "high-level" calculation in vacuo, keeping the QM-MM interaction contribution at a quantum "low-level". This allows computation of energies, gradients, and Hessians including the polarization of the QM subsystem and its interaction with the MM environment, both terms calculated using the low-level method at a reasonable computational cost. New information on an important enzymatic transformation is provided with greater reliability than has previously been possible. The predicted kinetic isotope effects on $V_{max}/K_m$ are 1.33 and 0.86 (at 30 °C) for 5-$^3$H and 9-$^3H_2$ substitutions, respectively, and 1.011 and 1.055 (at 22 °C) for 1-$^{13}$C and 7-$^{18}$O substitutions, respectively.

We have applied a novel quantum-mechanical/molecular-mechanical (QM/MM) procedure to compute estimates for experimentally unobserved intrinsic kinetic isotope effects (KIEs) upon the rearrangement of chorismate **1** to prephenate **2** catalyzed by chorismate mutase (Scheme 1). This study provides new information on an important enzymatic transformation with greater reliability than has previously been possible.

Claisen rearrangements proceed by a concerted mechanism involving a chairlike transition state (TS) with asynchronous carbon−oxygen bond cleavage and carbon−carbon bond formation.[1,2] However, the rate of conversion of chorismate to prephenate catalyzed by chorismate mutase from *B. subtilis* (BsCM)[3] or from *E. coli* (EcCM)[4] is only partially limited by the chemical rearrangement step, with the consequence that observed primary 1-$^{13}$C and 7-$^{18}$O and secondary 5-$^3$H and 9-$^3H_2$ KIEs for substitutions at the C1−C9 bond-making and C5−O7 bond-breaking positions do not reflect their intrinsic values and do not provide insight into the nature of the TS for the enzymic Claisen rearrangement.

We recently showed the feasibility of KIE calculations for this reaction within a fully flexible QM/MM description of the protein environment of BsCM.[5] However, the accuracy of the results was prejudiced by use of the semiempirical AM1 Hamiltonian for the QM region, because it is known that this method predicts TS structures for gas-phase Claisen rearrangements that are too tight with respect to those obtained by more reliable ab initio Hartree−Fock or density functional theory (DFT) methods.[2] We now present DFT/MM results by employing a new strategy to reduce the expense of the calculations.

All structures are determined using a modified QM/MM approach in which a "low-level" (LL) QM contribution is corrected during the geometry optimization procedure by means of a "high-level" (HL) calculation in vacuo. The total energy of the full system comprising the QM core within the MM environment may be written as

$$E_{total} = E^{HL}_{core} + (E^{LL}_{core/env} - E^{LL}_{core}) + E^{MM}_{env} \quad (1)$$

where $E^{HL}_{core}$ is the gas-phase energy for the core atoms alone using a high-level QM method (in this case B3LYP/6-31G*), $E^{LL}_{core}$ is the corresponding term obtained by a low-level method (in this case AM1), $E^{LL}_{core/env}$ is the energy of the QM region within the MM environment as obtained using the low-level method, and $E^{MM}_{env}$ is the energy of the MM region. Alternatively, the potential energy function can be expressed as

$$E_{total} = E^{HL}_{core} + (E^{LL}_{core/env:pol} + E^{LL}_{core/env:int}) + E^{MM}_{env} \quad (2)$$

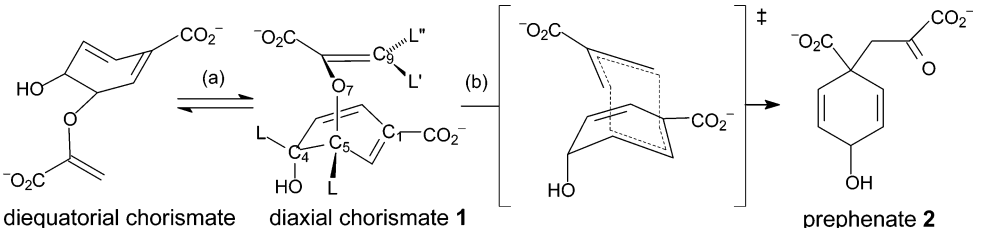
where the second and third terms are the polarization of the QM core and the interaction of this with the MM environment, with both terms being evaluated using the low-level QM method. This equation stresses the fact that the core is electronically embedded in the environment through the low-level method. Calculations using this potential energy function are denoted as HL:LL/MM.

It is important to stress how the present method differs from the popular ONIOM scheme of Morokuma and co-workers.[6] Following our notation, the total energy for an ONIOM3

* Corresponding authors. V.M.: fax, +34964728066; tel, +34964728084; e-mail, moliner@nuvol.uji.es. I.T.: fax, +34964728066; tel, +34964728084; e-mail, ignacio.tunon@uv.es.
[†] Universitat Jaume I.
[‡] Universidad de Valencia.
[§] University of Bath.

10.1021/jp044387u CCC: $30.25 

**SCHEME 1**




treatment of a system involving three spatial regions (core, shell, and environment)[7] and three methods (high-level QM, low-level QM, and MM) would be written as

$$E_{\mathrm{ONIOM3}} = E^{\mathrm{HL}}{}_{\mathrm{core}} + E^{\mathrm{LL}}{}_{\mathrm{core+shell}} + E^{\mathrm{MM}}{}_{\mathrm{core+shell+env}} - E^{\mathrm{LL}}{}_{\mathrm{core}} - E^{\mathrm{MM}}{}_{\mathrm{core+shell}} \quad (3)$$

where $E^{\mathrm{MM}}{}_{\mathrm{core+shell+env}}$ denotes the MM energy of all the atoms of the full system, and so on. In our method there is no "shell" region between the core and environment, for the HL and LL QM methods are both used to describe the core region. Thus eq 3 simplifies to

$$E_{\mathrm{ONIOM3}} = E^{\mathrm{HL}}{}_{\mathrm{core}} + E^{\mathrm{MM}}{}_{\mathrm{core+env}} - E^{\mathrm{MM}}{}_{\mathrm{core}} \quad (4)$$

which could be expressed alternatively as

$$E_{\mathrm{ONIOM3}} = E^{\mathrm{HL}}{}_{\mathrm{core}} + E^{\mathrm{MM}}{}_{\mathrm{core/env:int}} + E^{\mathrm{MM}}{}_{\mathrm{env}} \quad (5)$$

Equation 5 differs significantly from eq 2 inasmuch as the interaction between the core and environment is evaluated at the MM level, whereas in our method it is determined by means of the LL/QM method. Furthermore, our method includes polarization of the QM core by the MM environment, which the ONIOM method does not include.

The present scheme, which can be easily extended to obtain gradients and Hessians, is computationally advantageous only if, during the course of geometry optimization, the HL computations may be carried out much less often than the LL ones. This is the purpose of the algorithm used to explore the potential energy surface (PES), which utilizes a "control" space of the coordinates of the atoms directly related with the chemical process—the QM atoms plus their immediate surroundings—and a "complementary" space including the rest of the coordinates of the system. At each optimization step in the control space, the complementary space is fully relaxed (for a frozen control-space configuration); because the QM/MM interaction term is based on a low-level description, these energy minimization cycles take place relatively fast.[8] Afterward, high- and low-level calculations for the QM atoms defining the control space are performed, from which a Hessian is obtained. Thus the method provides all the necessary information to carry out TS structure searches and characterization of the chemical region with a high-level description at a reasonable computational cost, because the number of optimization cycles on the control space is usually much smaller than on the complementary space.

The initial structures of the putative stationary points were selected from hybrid QM/MM molecular dynamical simulations[9] using the DYNAMO program,[10] and were refined by means of AM1/OPLS-AA and B3LYP/6-31G*:AM1/OPLS-AA calculations with full gradient relaxation of the position of 17159 atoms (protein and water molecules). Periodic boundary conditions have been applied in all calculations. Once the stationary structures on the PES were located and characterized, the rigid-rotor/harmonic-oscillator approximations were used with the CAMVIB/CAMISO programs[11] to calculate semiclassical KIEs without scaling of vibrational frequencies, as explained and applied in previous papers.[5,12] A core of 70 atoms was used to define the Hessian for these KIE calculations, consistent with the "cut-off rule" and the local nature of isotope effects.[12] This subset of atoms was equivalent to the QM region substrate plus those atoms of the amino acids, explicitly depicted in Scheme 2, which are interacting with the substrate by means of hydrogen bonds.

**SCHEME 2**

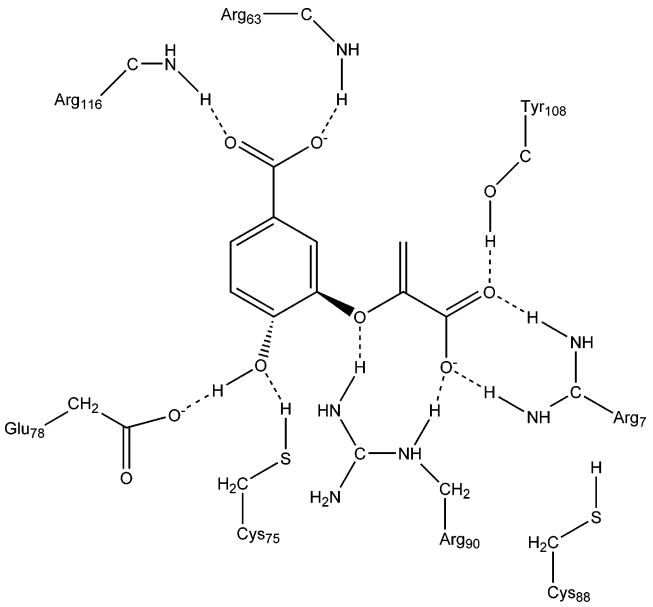



**TABLE 1: Selected Interatomic Distances (Å), Activation Enthalpies (kcal mol$^{-1}$), and Transition Frequencies (cm$^{-1}$) for the Optimized Geometries of the Transition State Obtained by Means of AM1/MM and B3LYP/6-31G*:AM1/MM Calculations**

| | AM1/MM | B3LYP/6-31G*:AM1/MM | experimental |
|---|---|---|---|
| C5···O7 | 1.912 | 2.106 | |
| C1···C9 | 2.202 | 2.790 | |
| $\Delta H^{\ddagger}$ | 37.4 | 16.4 | 12.7[a] |
| $\nu^{\ddagger}$ | 743*i* | 183*i* | |

[a] Reference 15.

Selected interatomic distances for the TS located in the enzyme at AM1/MM and B3LYP/6-31G*:AM1/MM levels, as well as the activation enthalpies and the transition frequencies are presented in Table 1. Both methods predict a chairlike TS but, as previously reported,[2] the AM1 structure has 1,4-diyl character, whereas B3LYP gives a more dissociative bis-allyl-like structure. The B3LYP/6-31G*:AM1/MM method gives an activation enthalpy $\Delta H^{\ddagger}$ much closer to the experimental value than does AM1/MM. The imaginary frequencies $\nu^{\ddagger}$, corresponding to the transition vector of each TS, differ greatly between the two methods.

**TABLE 2: Isotope Effects $EIE_a$ for the (dieq-chorismate)$_{water}$ = (diax-chorismate)$_{enzyme}$ Preequilibrium Step (a) and $KIE_b$ for the Rearrangement Step (b) and the Overall Rearrangement from Aqueous Diequatorial Chorismate to Prephenate, Calculated as $EIE_a \times KIE_b$, in the BsCM Enzyme Active Site**

| | | AM1/MM | | | B3LYP/6-31G*:AM1/MM | | | | $KIE_b$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *T*/°C | $EIE_a$ | $KIE_b$ | KIE | $EIE_a$ | $KIE_b$ | KIE | experimental | HF/EFP[d] BsCM | $KIE_b$B3LYP/6-31G*[e] in vacuo |
| [4-$^2$H] $k_H/k_D$ | 30 | 0.9504 | 1.0203 | 0.9697 | 0.8158 | 1.0303 | 0.8405 | | 1.163 | |
| [4-$^3$H] $k_H/k_T$ | 30 | 0.9317 | 1.0284 | 0.9582 | 0.7494 | 1.0435 | 0.7820 | | 1.225 | |
| [5-$^2$H] $k_H/k_D$ | 30 | 1.0085 | 1.1165 | 1.1260 | 1.0531 | 1.1653 | 1.2272 | | 1.475 | |
| [5-$^3$H] $k_H/k_T$ | 30 | 1.0112 | 1.1673 | 1.1804 | 1.0756 | 1.2386 | 1.3322 | $1.003 \pm 0.020$[a] | | 1.763 |
| [9-$^2$H′] $k_H/k_D$ | 30 | 0.9932 | 0.9410 | 0.9346 | 0.9555 | 0.9922 | 0.9480 | | 0.884 | |
| [9-$^3$H′] $k_H/k_T$ | 30 | 0.9904 | 0.9207 | 0.9119 | 0.9380 | 0.9892 | 0.9279 | | | |
| [9-$^2$H″] $k_H/k_D$ | 30 | 0.9706 | 0.9825 | 0.9536 | 0.9322 | 1.0131 | 0.9444 | | 1.028 | |
| [9-$^3$H″] $k_H/k_T$ | 30 | 0.9592 | 0.9766 | 0.9367 | 0.9062 | 1.0216 | 0.9258 | | | |
| [9-$^3$H$_2$] $k_H/k_T$ | 30 | 0.9511 | 0.9006 | 0.8566 | 0.8515 | 1.0102 | 0.8602 | $1.012 \pm 0.004$[a] | | |
| [1-$^{13}$C] $k_{12C}/k_{13C}$ | 22 | 1.0017 | 1.0193 | 1.0210 | 1.00089 | 1.00999 | 1.0109 | $1.0043 \pm 0.0002$[b]<br>$1.0057 \pm 0.0002$[c] | 1.012 | 1.014 |
| [7-$^{18}$O] $k_{16O}/k_{18O}$ | 22 | 1.0013 | 1.0504 | 1.0518 | 1.0099 | 1.0449 | 1.0552 | $1.045 \pm 0.003$[b]<br>$1.053 \pm 0.002$[c] | 1.071 | 1.043 |

[a] Reference 4, wild-type EcCM. [b] Reference 13, wild-type BsCM. [c] Reference 13, BsCM C75S mutant. [d] Reference 14. [e] Reference 13.

Inspection of the KIEs (Table 2) for the diaxial chorismate to prephenate rearrangement (step (b) in Scheme) within the BsCM active site shows a normal ($k_{light} > k_{heavy}$) value for substitution of tritium for protium attached to C5, consistent with the change from tetrahedral toward trigonal geometry about this center as C5−O7 bond cleavage proceeds; the larger value obtained with B3LYP/6-31G*:AM1/MM than with AM1/MM reflects the more dissociative character of the TS as predicted by the DFT method. Whereas AM1/MM predicts an inverse ($k_{light} < k_{heavy}$) effect for substitution of two tritiums attached to C9, consistent with progress from trigonal toward tetrahedral geometry about this center in the TS, B3LYP/6-31G*:AM1/MM predicts a very small normal effect, implying that there is no significant degree of bond-making between C1 and C9 in the more dissociative TS. Both methods predict small, normal heavy-atom effects for 1-$^{13}$C and 7-$^{18}$O substitutions, as expected. B3LYP/6-31G* calculations[13] for step (b) in a vacuum yield KIEs of 1.014 and 1.043 for 1-$^{13}$C and 7-$^{18}$O, respectively, values not greatly different from the present estimates of 1.0099 and 1.0449; this may suggest that the protein environment does not have much influence upon the heavy-atom KIEs for this step. Worthington et al. have recently reported KIEs calculated for step (b) using a QM/MM method combining HF/4-31G SBK with effective fragment potentials to represent the protein environment.[14] In every case their KIEs are larger in magnitude than our B3LYP/6-31G*:AM1/MM values (see Table 2). In particular, their values for deuterium substitution at C4, C5, and C9 appear to be grossly overestimated.

Experiments using the competitive method yield KIEs on $V_{max}/K_m$ for the enzyme-catalyzed reaction relative to free substrate and enzyme in aqueous solution,[4,15] and measurements of the temperature dependence of $^1$H NMR coupling constants indicates that the predominant conformer of chorismate in water is the pseudo-diequatorial form.[16] We have computed the equilibrium isotope effects (EIEs) (see Table 2) for the binding step (a) of aqueous diequatorial chorismate going to the diaxial conformer in the BsCM active site. Small normal effects are found with both theoretical methods for the 5-$^3$H and heavy-atom substitutions, but surprisingly large inverse EIEs are predicted for the 4-$^3$H and 9-$^3$H$_2$ substitutions.

The products $EIE_a^{\bullet} \times KIE_b$ give the overall KIEs (Table 2) that represent the best available estimates for the intrinsic KIEs for BsCM. Both theoretical methods predict a [7-$^{18}$O] $k_{16}/k_{18}$ effect within the experimental error of the value observed for the viscosity-insensitive BsCM C75S mutant for which the chemical rearrangement step (b) is probably rate limiting. Both methods appear to overestimate the [1-$^{13}$C] $k_{12}/k_{13}$ effect, although the DFT:AM1/MM result is closer to the experimental value. The two observed secondary tritium KIEs, which are both close to unity within experimental uncertainty, were determined for EcCM;[4] however, because both this and BsCM have comparable activites, inhibition profiles and similarly functionalized active sites,[13] their mechanisms and intrinsic KIEs may also be similar. Finally, it is important to stress that the experimental results do not necessarily reflect intrinsic KIEs because probably the rate-limiting TS precedes the isotopically sensitive rearrangement step (b). It would be most interesting to compare our predicted deuterium or tritium KIEs at C4, C5, and C9 with future experimental values for the BsCM C75S mutant. This new method is now being applied to other enzymatic systems where the chemical reaction has been demonstrated to be the rate-limiting step of the full process.

**Acknowledgment.** We thank the Royal Society for a Joint Project Grant 15666, DGI for project DGI BQU2003-04168-C03, BANCAIXA for project P1A99-03 and Generalitat Valenciana for projects GRUPOS04/28, GV04B-21 and GV04B-131.

## References and Notes

(1) Gajewski, J. J. *Acc. Chem. Res.* **1997**, *30*, 219−225.

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

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

(4) Addadi, L.; Jaffe, E. K.; Knowles, J. R. *Biochemistry* **1983**, *22*, 4494−4501.

(5) Martí, S.; Moliner, V.; Tuñón, I.; Williams, I. H. *Org. Biomol. Chem.* **2003**, *1*, 483−487.

(6) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma K. *J. Phys. Chem.* **1996**, *100*, 19357−19363.

(7) In the notation of Morokuma and co-workers: SModel = core, IModel = core+shell, and Real = core+shell+env.

(8) As a note of efficiency, during the relaxation of the complementary space (the key difference between the standard DFT/MM method and our new approach DFT:AM1/MM) a gradient cycle of calculation of our molecular model requires, on average, ca. 164.5 s at B3LYP(6-31G*)/MM level, whereas it only takes 0.4 s at the B3LYP(6-31G*):AM1/MM level. Obviously, a Hessian cycle takes more or less the same computational cost in both methods. The test has been carried out in a SGI−Altix Itanium-2 (1.5 GHz).

(9) (a) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. *Chem. Eur. J.* **2003**, *9*, 984−991. (b) Martí, S.; Andrés, J.; Moliner, V.;

Silla, E.; Tuñón, I.; Bertrán, J.; Field, M. J. *J. Am. Chem. Soc.* **2001**, *123*, 1709−1712.

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

(11) (a) Williams, I. H. *Chem. Phys. Lett.* **1982**, *88*, 462−466. (b) Williams, I. H. *J. Mol. Struct. THEOCHEM* **1983**, *11*, 275−284.

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

(13) Gustin, D. J.; Mattei, P.; Kast, P.; Wiest, O.; Lee, L.; Cleland, W. W.; Hilvert, D. *J. Am. Chem. Soc.* **1999**, *121*, 1756−1757.

(14) Worthington, S. E.; Roitberg, A. E.; Krauss, M. *Int. J. Quantum Chem.* **2003**, *94*, 287−292.

(15) Kast, P.; Asif-Ullah, M.; Hilvert, D. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.

(16) Copley, S. D.; Knowles, J. R. *J. Am. Chem. Soc.* **1987**, *109*, 5008−5013.