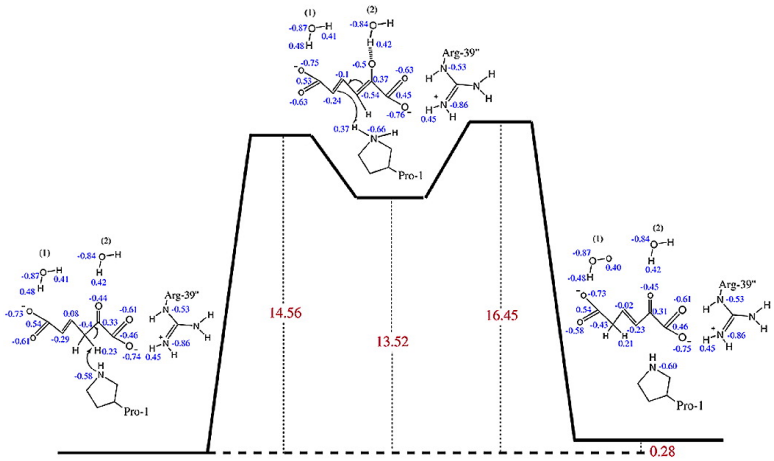
Article

Ab Initio QM/MM Study Shows There Is No General Acid in the Reaction Catalyzed by 4-Oxalocrotonate Tautomerase

G. Andrs Cisneros, Haiyan Liu, Yingkai Zhang, and Weitao Yang

J. Am. Chem. Soc., **2003**, 125 (34), 10384-10393• DOI: 10.1021/ja029672a • Publication Date (Web): 05 August 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 18 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 08/05/2003

Ab Initio QM/MM Study Shows There Is No General Acid in the Reaction Catalyzed by 4-Oxalocrotonate Tautomerase

G. Andrés Cisneros,[†] Haiyan Liu,[†,‡] Yingkai Zhang,[†,§] and Weitao Yang*,[†]

Contribution from the Department of Chemistry, Duke University, Box 90346, Durham, North Carolina 27708-0346, School of Life Science, University of Science and Technology of China, Hefei, Anhui, 230026, China, and Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, 9500 Gillman Drive, La Jolla, California 92093-0365

Received December 10, 2002; E-mail: weitao.yang@duke.edu

Abstract: The mechanism for the reaction catalyzed by the 4-oxalocrotonate tautomerase (4-OT) enzyme has been studied using a quantum mechanical/molecular mechanical (QM/MM) method developed in our laboratory. Total free energy barriers were obtained for the two steps involved in this reaction. In the first step, Pro-1 acts as a general base to abstract a proton from the third carbon of the substrate, 2-oxo-4-hexenedioate, creating a negative charge on the oxygen at C-2 of this substrate. In the second step, the same hydrogen abstracted by the N-terminal Pro-1 is shuttled back to the fifth carbon of the substrate to form the product, 2-oxo-3-hexenedioate. The calculated total free energy barriers are 14.54 and 16.45 kcal/mol for the first and second steps, respectively. Our calculations clearly show that there is no general acid in the reaction. Arg-39″, which is hydrogen bonded to the carboxylate group of the substrate, and an ordered water, which moves closer to the site of the charge formed in the transition state and intermediate, play the main role in transition state/intermediate stabilization without acting as general acids in the reaction.

1. Introduction

4-Oxalocrotonate tautomerase (4-OT) catalyzes the isomerization of unconjugated α-keto acids such as 2-oxo-4-hexenedioate to its conjugated isomer, 2-oxo-3-hexenedioate, through the dienol intermediate 2-hydroxy-2,4-hexadienedioate (also named 2-hydroxymuconate) (Figure 1). This enzyme is produced by *Pseudomonas putida mt-2*, a soil bacterium. It is part of a degradative metabolic pathway that converts various aromatic hydrocarbons to intermediates in the citric acid cycle (Krebs cycle). The plasmid that encodes 4-OT has been determined to be the TOL plasmid pWW0,[1] which encodes the entire degradative pathway and enables the bacteria to utilize various aromatic hydrocarbons as their sole sources of carbon and energy.[2]

4-OT is an enzyme composed of six subunits arranged in dimers, that is, a "trimer of dimers" with 62 residues per subunit. The subunit secondary structure consists of an α-helix, two β-strands, a β-hairpin, two loops, and two turns.[3] There are a total of six active sites in this protein, located in hydrophobic regions around the Pro-1 of each subunit. Several additional features of 4-OT also make it a particularly interesting system to study: the amino terminal Pro-1 functions as a general base,

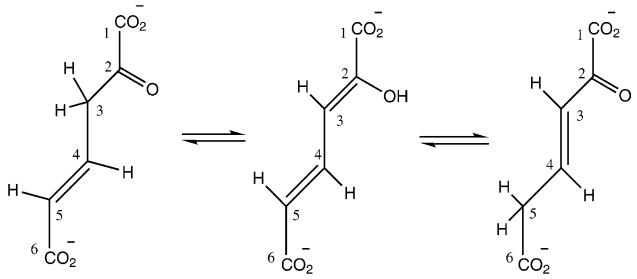


Figure 1. Isomerization catalyzed by 4-OT.

the enzyme has a very small monomer size (62 residues), and 4-OT catalyzes a proton-transfer reaction without the aid of any metal ion or cofactor.

Recent X-ray studies[4] have shown that, in addition to Pro-1, there are a number of other active site residues belonging to other subunits, such as Arg-11′, Arg-39″, and Arg-61′, which might be involved in the isomerization reaction. This same study has shown that several other residues are in close proximity to the active site and may interact with the substrate. In our notation, following that of ref 4, the unprimed and single-primed residues belong to two different monomers from the same dimer, and the double-primed residue belongs to a third monomer.

Several experiments have been carried out to determine the mechanism by which 4-OT isomerizes the substrate. Experimental results suggest that this is a general acid−base mechanism, with Pro-1 acting as the general base. However, the

[†] Duke University.
[‡] University of Science and Technology of China.
[§] Howard Hughes Medical Institute.

(1) Chen, L. H.; Kenyon, G. L.; Curtin, F.; Harayama, S.; Bembenek, M. E.; Hajipour, G.; Whitman, C. P. *J. Biol. Chem.* **1992**, *267*, 17716.
(2) Harayama, S.; M. Rekik, K.-L. N.; Ornston, L. *J. Bacteriol.* **1989**, *171*, 6251.
(3) Stivers, J. T.; Abeygunawardana, C.; Whitman, C. P.; Mildvan, A. S. *Protein Sci.* **1996**, *5*, 729.
(4) Taylor, A. B.; Czerwinski, R. M.; Johnson, W. H., Jr.; Whitman, C. P.; Hackert, M. L. *Biochemistry* **1998**, *37*, 14692.

10.1021/ja029672a CCC: $25.00

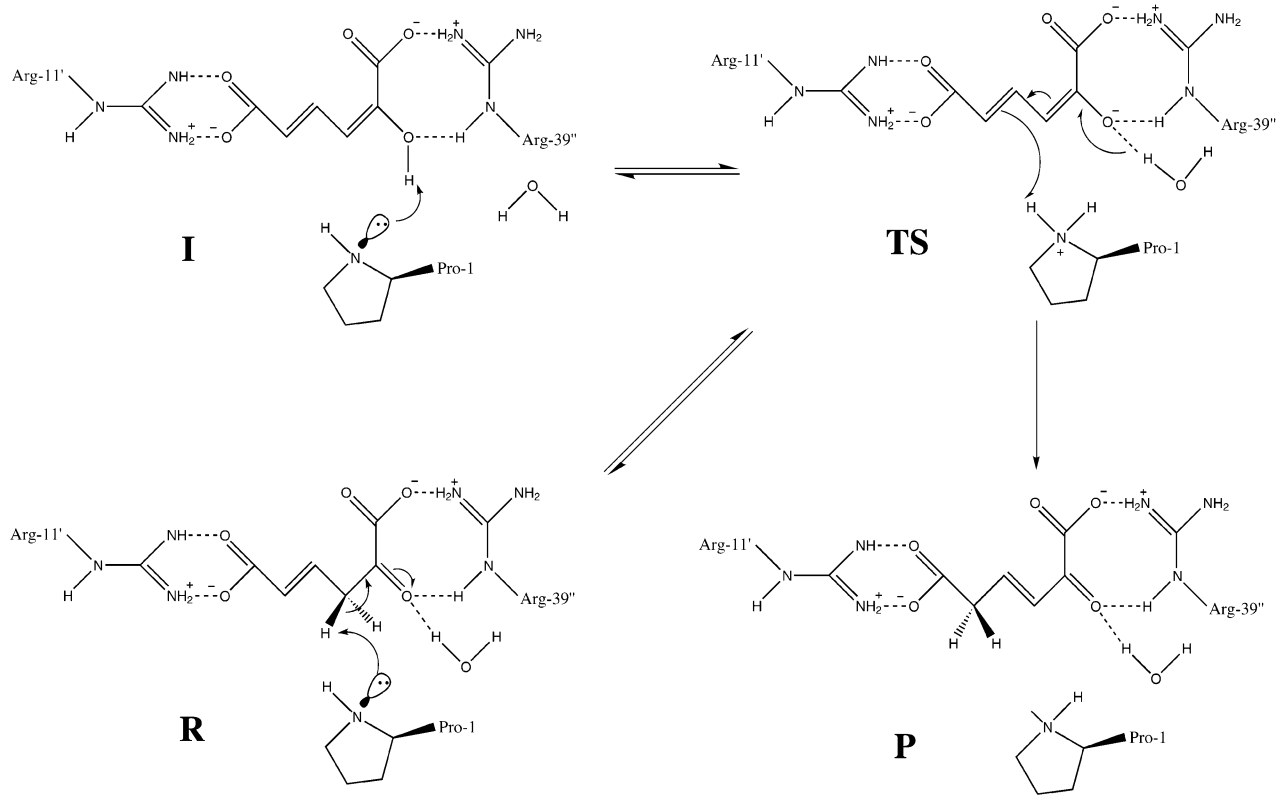


Figure 2. Reaction scheme proposed by Harris et al.[9] Reactant (R), dienol intermediate (I), transition state (TS), and product (P) states are presented.

general acid in this reaction has remained elusive.[5−10] Arg-39″ has been suggested to act as a general acid in the reaction because theoretical calculations yielded pK_a values of 9.8 to 10.3 for this residue.[11] However, it has been found experimentally that even though Arg-39″ might be involved in the reaction mechanism (see Figure 2), none of the three arginine residues in the active site have the pK_a of 9.0 determined experimentally to be the essential acidic group in free 4-OT.[12]

A reaction scheme has been proposed by Harris et al.[9,10] based on experimental data in which Arg-39″ along with a water in the active site helps stabilize the negative charge formed on the carbonyl oxygen of the substrates in the transition states (TSs) and intermediate structures (see Figure 2). In particular, "... Arg-39″ and an ordered water molecule each provide a hydrogen bond to the C-2 oxygen of substrate. Arg-39″ plays an additional role in the binding of the C-1 carboxylate group. Arg-11′ participates both in substrate binding and catalysis [catalytic role explained as an electron sink] ...".[10]

In the present research, we have carried out calculations on 4-OT using a combined quantum mechanical/molecular mechanical (QM/MM)[13] method developed recently in our laboratory.[14−16] Our computational modeling provides further elucidation of the reaction mechanism in 4-OT and addresses the issue of the catalytic acid which is yet to be identified.

2. Computational Methods

2.1. QM/MM Free Energy of Activation. Our approach uses the pseudobond model of the QM/MM interface developed by Zhang et al.[14] The reaction paths are determined by an iterative energy minimization procedure. The free energies along the reaction path are determined by free energy perturbation calculations and the harmonic approximation for the fluctuation of the QM subsystem.

All calculations were performed using QM/MM methodology[14−16] that has been implemented in a modified version of Gaussian 98,[17] which interfaces to a modified version of TINKER.[18] The AMBER95 all-atom force field parameter set[19] and the TIP3P model[20] for water were used.

A very important part of this QM/MM implementation is the use of the pseudobond model for the QM/MM boundary as developed in ref 14, which provides a smooth connection between the QM and the MM subsystems as well as an integrated expression for the potential energy of the overall system.

(5) Lian, H.; Whitman, C. P. *J. Am. Chem. Soc.* **1993**, *115*, 7978.
(6) Stivers, J. T.; Abeygunawardana, C.; Mildvan, A. S.; Hajipour, G.; Whitman, C. P.; Chen, L. H. *Biochemistry* **1996**, *35*, 803.
(7) Fitzgerald, M. C.; Chernushevich, I.; Standing, K. G.; Whitman, C. P.; Kent, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6851.
(8) Czerwinski, R. M.; Johnson, W. H., Jr.; Whitman, C. P.; Harris, T. K.; Abeygunawardana, C.; Mildvan, A. S. *Biochemistry* **1997**, *36*, 14551.
(9) Harris, T. K.; Czerwinski, R. M.; Johnson, W. H., Jr.; Legler, P. M.; Abeygunawardana, C.; Massiah, M. A.; Stivers, J. T.; Whitman, C. P.; Mildvan, A. S. *Biochemistry* **1999**, *38*, 12343.
(10) Whitman, C. *Arch. Biochem. Biophys.* **2002**, *402*, 1.
(11) Soares, T.; Goddsell, D.; Briggs, J.; Ferreira, R.; Olson, A. *Biopolymers* **1999**, *50*, 319.
(12) Czerwinski, R. M.; Harris, T. K.; Johnson, W. H., Jr.; Legler, P. M.; Stivers, J. T.; Mildvan, A. S.; Whitman, C. P. *Biochemistry* **1999**, *38*, 12358.
(13) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227.
(14) Zhang, Y.; Lee, T.; Yang, W. *J. Chem. Phys.* **1999**, *110*, 46.
(15) Zhang, Y.; Liu, H.; Yang, W. *J. Chem. Phys.* **2000**, *112*, 3483.
(16) Zhang, Y.; Liu, H.; Yang, W. Ab Initio QM/MM and Free Energy Calculations of Enzyme Reactions. In *Computational Methods for Macromolecules-Challenges and Applications*; Springer-Verlag: Heidelberg, Germany, 2002.
(17) Frisch, M. J.; et al. *Gaussian 98*, revision A.8; Gaussian, Inc.: Pittsburgh, PA, 1998.
(18) Ponder, J. *TINKER, Software Tools for Molecular Design, Version 3.6: the most updated version for the TINKER program can be obtained from J. W. Ponder's WWW site at http://dasher.wustl.edu/tinker.*; Washington University: St. Louis, MO, 1998.
(19) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(20) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.

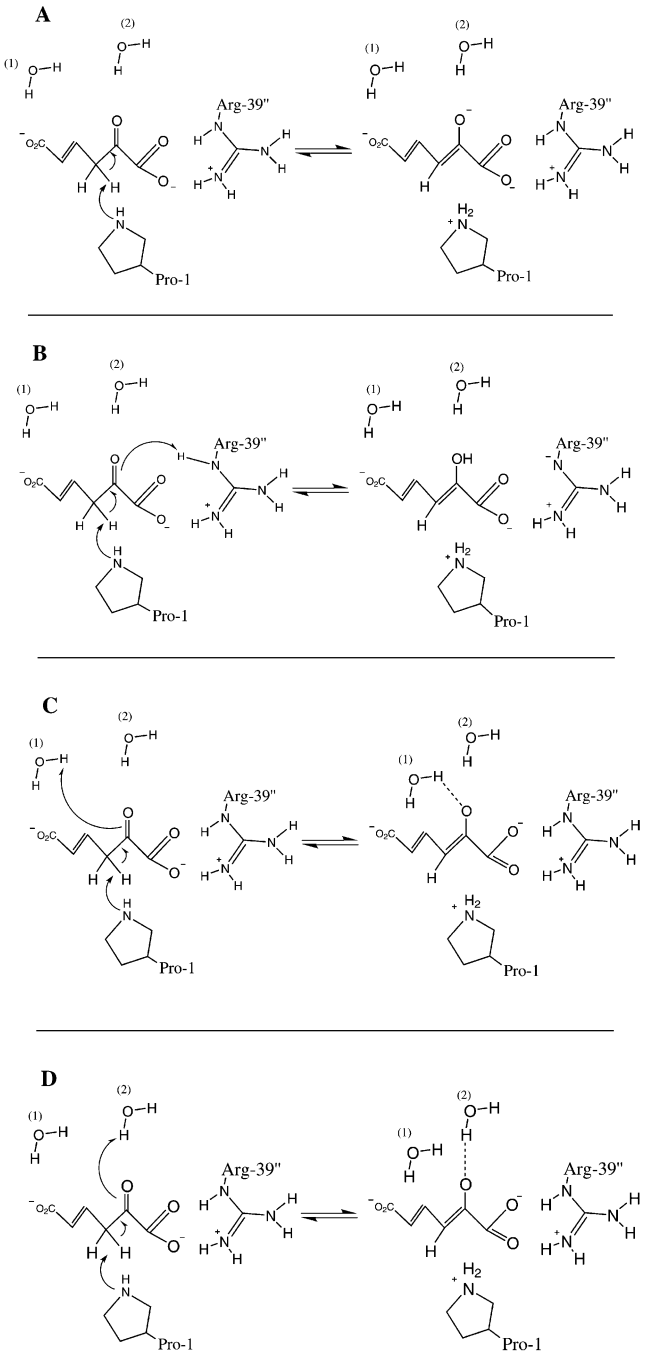


Figure 3. Tested reaction schemes for the 4-OT catalyzed reaction.

In the QM/MM potential energy model, the total energy of the system is

$$E_{\text{Total}} = E_{\text{MM}} + E_{\text{QM}} + E_{\text{QM/MM}} \tag{1}$$

The QM/MM interactions ($E_{\text{QM/MM}}$) are taken to include bonded and nonbonded interactions. For the nonbonded interactions, the subsystems interact with each other through Lennard−Jones and point charge interaction potential. When the electronic structure is determined for the QM subsystem, the charges in the MM subsystem are included as a collection of fixed point charges in an effective Hamiltonian which describes the QM subsystem; that is, in the calculation of the QM subsystem, we determine the contributions from the QM subsystem (E_{QM}) as well as the electrostatic contributions from the interaction between the QM and MM subsystems as explained by Zhang et al.[15]

Geometry optimizations are carried out by an iterative minimization procedure as described by Zhang et al.[15] In this procedure, one iteration consists of a complete optimization of the QM subsystem, followed by a complete optimization of the MM subsystem. At each point, the subsystem not being optimized is held fixed at the geometry obtained from the previous iteration; QM/MM interactions are also included at each iteration. The iterations are continued until the geometries of both systems no longer change.

When the MM subsystem is being optimized, or a molecular dynamics simulation is being carried out on the MM subsystem, the QM/MM electrostatic interactions are approximated with fixed point charges on the QM atoms which are fitted to reproduce the electrostatic potential (ESP) of the QM subsystem.[21]

The reaction paths were calculated using the reaction coordinate driving method (RCDM).[22] This method introduces a harmonic restraint on the reaction coordinate, which is a linear combination of the distances between the atoms involved in the reaction to perform an optimization along a proposed reaction path. The reaction coordinate is given by the expression:

$$R = \sum_{i=1}^{n} a_i r_i \tag{2}$$

where r_i values are the distances between atoms; a_i is constant 1 for the distance that increases and -1 for the distance that decreases. The sum over i includes all the distances that change throughout the course of the reaction. R is included in the following energy expression:

$$E_{\text{Restrain}} = k(R - s)^2 \tag{3}$$

where R is given by eq 2, s is an adjustable parameter corresponding to the value of the reaction coordinate which is varied in a stepwise manner by 0.1 Å at each point on the PES, and k is a force constant; in this case, the value of k was set to 2000 kcal/mol for all points. This energy is included in the total energy expression in the process of the optimization.

All the calculated reaction paths were determined by stepping forward (from initial state to final state for that particular step) and backward (from final state to initial state for that particular step) along the path several times until there was no change between the forward and backward paths.

Vibrational frequency calculations were performed on the structures obtained for the maxima and minima along the paths (reactant, product, intermediates and transition states) to characterize the stationary points on the PES. Stationary points with one and only one imaginary (negative) vibrational frequency were characterized as transition states. Reactant, product, and stable intermediates were characterized as having no imaginary frequencies. All vibrational frequencies were calculated at the HF/3-21G level, with a scaling factor of 0.9409.[23]

After the reaction paths were determined, the free energy perturbation method (FEP)[24] was employed to determine the free energy profiles associated with the calculated reaction paths for scheme A (see Figure 3). These calculations were carried out in the following manner: In each molecular dynamics (MD) simulation, the QM subsystem was fixed to a given state along the reaction path. At each of these states, the reaction coordinate had a particular value, and the QM subsystem had a fixed geometry and charge distribution as obtained from the reaction path calculation; this state is called a simulated state. The free energy changes associated with perturbing the simulated state "forward" and "backward" to neighboring states along the reaction path were calculated according to the FEP theory.[24]

(21) Besler, B.; Merz, K. M., Jr.; Kollman, P. *J. Comput. Chem.* **1990**, *11*, 431.
(22) Williams, I.; Maggiora, G. *J. Mol. Struct.* **1982**, *89*, 365.
(23) Foresman, J.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*; Gaussian Inc.: Pittsburgh, PA, 1996.
(24) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.

The free energy perturbation calculations at the stationary points were further improved by calculating the contributions from fluctuations of the QM subsystem to the free energy difference.[15,16] These contributions were determined by calculating the Hessian matrixes of the stationary points for the degrees of freedom involving atoms in the QM subsystem and the subsequent calculation of the vibrational frequencies. By using the quantum mechanical harmonic approximation, the change in contribution from the fluctuations of the QM subsystem for all stationary points was determined.

2.2. Initial Structure Determination for 4-OT. The initial structure of 4-OT inactivated with 2-oxo-3-pentynoate was taken from the Brookhaven protein data bank (EC 5,3,2; 4-OT, 1BJP).[4,25] We chose the inactivated crystal structure as opposed to the native one in order to superpose the substrate to the inhibitors located in the active site. This provided us with an initial guess of the orientation of the substrates in the active site based on experimental results. It is important to point out that 2-oxo-3-pentynoate is an irreversible inhibitor, and it is not clear if the covalent binding of this inhibitor could affect the active site structure. However, since there is no known 4-OT crystal structure with any substrate or competitive inhibitor in the active site, we feel that this inactivated crystal structure presents the best available initial structure for our calculations.

We decided to perform our calculations with *trans*-2-oxo-4-hexenedioate as the substrate, based on experimental findings of Whitman et al.[26] who have suggested that *trans*-2-oxo-4-hexenedioate is a better substrate than the dienol intermediate, 2-hydroxymuconate.

Following Cornell et al.,[19] an MP2 calculation using the 6-31G* basis set was carried out on a neutral proline residue with Gaussian 98[17] to determine the partial ESP charges, which were added to the AMBER95 parameter file. This procedure was also performed for the substrate because five substrate molecules were included in the MM subsystem.

It is known that the AMBER95 force field sets some Lennard−Jones (LJ) parameters for hydrogens to zero. In the case of 4-OT, there are only two types of residues that present these parameters, serine (Ser 12, 24, 28, 30, 37, and 58) and threonine (Thr 18, 36, and 43). From the crystal structure, we determined that none of the hydrogen atoms with these LJ parameters from these residues are closer than 5 Å from the active site. Therefore, we conclude that this will have no effect on our calculations.

After superposing the six substrate molecules (*trans*-2-oxo-4-hexenedioate) to the inhibitor molecules and removing the inhibitor molecules, we carried out an initial minimization. The rest of the system was restrained to the crystal structure coordinates, while the newly added substrates were allowed to move inside the active site to find the minimum energy structure within the restricted enzyme residues. Protonation states of the ionizable residues were assigned based on the theoretical pK_a calculations by Soares et al.[11]

Based on the converged structure, one of the active sites was selected to be used as the center on which the relaxation would be carried out, using the distances between the substrate's oxygens and the enzyme's arginine residues as references. The selected active site was solvated using a water sphere around C-3 of the chosen substrate. The radius of the water sphere was 30 Å, and water molecules overlapping with the crystal structure atoms were removed. This water sphere was added to model the system in solution with a finite system. The water sphere was comprised of 2407 water molecules, which, added to the enzyme's 5977 atoms, produces a total of 13197 atoms.

A 20 ps MD simulation was carried out keeping all atoms in the enzyme and substrates fixed to equilibrate the water sphere. After this procedure, the resulting structure was again minimized with the same constraints. The new system was subjected to a 100 ps MD simulation.

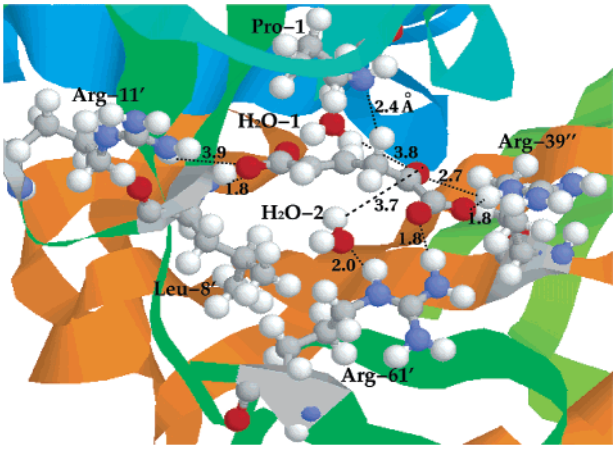


Figure 4. Initial structure for the 4-OT reaction with significant distances from substrate atoms to atoms in residues in the active site.

All atoms in the active sphere (see section 2.3) were restrained to their original coordinates for the first 10 ps. The restraints were gradually reduced every 10 ps until no restraints were left. When the backbone of the equilibrated structure was superposed to the backbone of the crystal structure (Supporting Information), no significant change in the overall backbone structure was observed. The maximum RMS deviation between the backbones was 0.74 Å.

After this minimization and MD procedures, two ordered water molecules from the calculated structure were found in the active site forming hydrogen bonds (H-bonds) with the substrate and Arg-61′. These water molecules are labeled 1 and 2, respectively; see Figure 4. X-ray studies of inactivated 4-OT have shown that two water molecules are conserved in the active site, one of which forms an H-bond to Ser-37 in the crystal structure.[4] There is also a significant change in the position of Arg-11′ from the inactivated crystal structure to the equilibrated structure: while the distance between N_ϵ of Arg-11′ and the C-5 in the inhibitor is approximately 9 Å in the inactivated crystal structure, the distance between the same nitrogen atom and C-5 of the substrate is now 6.0 Å. Figure 4 shows that the N_ϵ of Arg-11′ is now close to the carboxylate oxygen of the substrate. This indicates that Arg-11′ may play a role in the catalysis, as the analysis in section 3.3 confirms.

Based on these results, the QM subsystems were chosen as follows. For schemes A and C (see Figure 3), the QM subsystem was chosen to include the substrate, Pro-1, and H_2O-1. H_2O-1 was included in the QM subsystem because of its proximity to the substrate after the MD and minimizations. This gave a total of 36 atoms in the QM subsystem including the pseudobond atom for the calculation of these reaction schemes as shown in Figure 4. Note that interactions between the substrate and the enzyme include an H-bond to Leu-8′ through the backbone N atom in this residue, as well as H-bonds between the substrate and Arg-39″ and Arg-61′. Arg-11′ is close to the substrate but does not form an H-bond with the substrate. For scheme B, both the substrate and Pro-1 remained in the QM subsystem. The water molecule was replaced by Arg-39″ resulting in a QM subsystem with 55 atoms including 2 pseudoatoms. Finally, for scheme D, for the calculation of the PES, H_2O-1 was replaced by H_2O-2 in the QM subsystem, giving the same number of atoms in the QM subsystem as that for the case of the QM subsystems for the calculation of schemes A and C.

The nature of this particular system, namely, the N-terminal proline as a general acid, forced us to set the pseudobond[14,15] atom for Pro-1 on the backbone of the enzyme at the C_α for Ile-2, producing an N−C_{ps} bond, where C_{ps} is the pseudoatom. All the reported results in the present work were calculated with the available C−C_{ps} parameters previously calculated by Zhang et al.[14,15]

Test calculations were performed with N−C_{ps} parameters, which were only recently available, in the enzyme environment at the HF/3-

(25) Johnson, W. H.; Czerwinski, R. M.; Fitzgerald, M. C.; Whitman, C. P. *Biochemistry* **1997**, *36*, 15724.
(26) Whitman, C. P.; Aird, B.; Gillespie, W.; Stolowich, N. *J. Am. Chem. Soc.* **1991**, *113*, 3154.

21G level. The ΔE between the barriers calculated with $C-C_{ps}$ and $N-C_{ps}$ parameters was 0.14 and 0.23 kcal/mol for the first (with scheme A) and second steps, respectively. Structural parameters were not significantly altered as can be seen in the Supporting Information, where we also include results for a small test molecule. These results show that in this case the difference in parameters has a relatively small effect in calculated energetic and structural properties.

2.3. Computational Procedure for 4-OT. We now turn to the details of the selection of the reaction coordinates for the RCDM calculations for 4-OT. As explained above, it has been suggested that the reaction catalyzed by 4-OT is a general acid−base reaction, where the general acid is still unknown. For this reason, several reaction coordinates which involve different general acids were tested, as shown in Figure 3.

This reaction takes place in two steps. In all cases, the first step consists of a proton abstraction from C-3 of the substrate by the N atom in Pro-1. For scheme A in Figure 3, the chosen reaction coordinate (R_1) was set as the difference between the distance of one of the protons attached to C-3 of the substrate and the distance between this same proton and the N atom from Pro-1. For scheme B, in addition to this distance difference, we included the difference between the distance of N_ϵ to H_ϵ in Arg-39″ and this same hydrogen to the carbonyl oxygen on C-2 of the substrate. In the case of schemes C and D, we replaced the arginine as the general acid with H_2O-1 and H_2O-2, respectively, including the distance difference between the O atom and one H atom in the water molecule and this H and the carbonyl oxygen on C-2 of the substrate in the expression of R.

In the second step (R_2), Pro-1 returns the abstracted proton to C-5 of the substrate. The possible mechanism for the second step depends on whether there is a general acid in the first step. For this step R_2 comprises the distance from the leaving proton to the N of Pro-1 and the distance from this proton to C-5 of the substrate. The reason we did not include any other molecule in the reaction coordinate is that we did not find any general acid in the first step, as explained in section 3.1.

Geometry optimizations for the QM subsystem along both reaction paths were calculated at the HF/3-21G level of theory with interactions from the MM subsystem treated by the QM/MM method. Single point calculations at the B3LYP/6-31G* level were performed for each point calculated from the HF/3-21G QM/MM geometries using the QM/MM method to include the MM interactions with the QM subsystem. For the QM subsystems, the convergence criteria are set to be the default for Gaussian 98. In the case of the MM subsystems, an RMS below 0.1 kcal/mol Å was employed.

For all calculations involving the MM subsystem (MM optimizations and MD simulations), because we do not simulate an infinite system, the conformational fluctuations of atoms near the boundary will not be realistic and thus were ignored. An active sphere with a radius of 20 Å centered around C-3 of the substrate in the QM subsystem was selected. All atoms outside this active sphere were constrained to their crystal structure positions. The twin range cutoff method[27] was used for nonbonded interactions with a long-range cutoff distance of 15 Å and a short-range cutoff of 8 Å.

All MD simulations including FEP were carried out with a time step of 1.0 fs, maintaining a constant temperature of 300 K. In all cases, the SHAKE[28] algorithm was employed to constrain all bond lengths of bonds involving hydrogen atoms. In the case of the FEP simulations, full MM simulations were employed with the QM part held fixed at a particular state along the reaction path, as implemented by Zhang et al. (see above).[15] All free energy simulations were calculated using the ESP charges previously obtained on the QM atoms fitted to the B3LYP/6-31G* calculations.

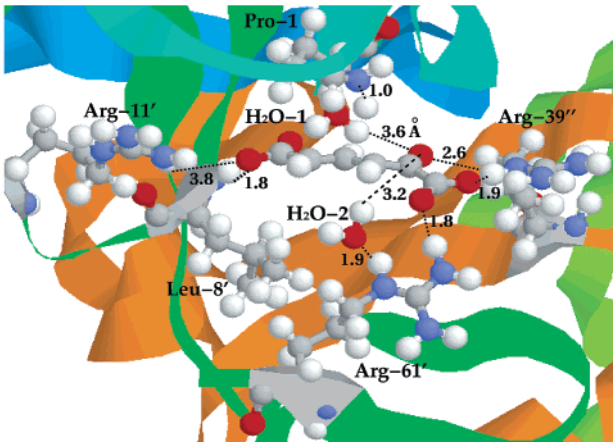


Figure 5. Calculated intermediate for the 4-OT reaction. Significant distances from substrate atoms to residue atoms are detailed; all distances are in Å.

3. Results and Discussion

3.1. Potential and Free Energies of Activation. Based on the acid−base mechanism suggested from experimental studies, several reaction coordinates were tested in this study to determine which residue or water molecule (if any) acts as the general acid in the reaction catalyzed by 4-OT. Figure 3 shows the reaction schemes for which PES scans were performed. For this calculation, only the first step was examined. The calculated potential energy barriers are 17.78, 22.28, 27.83, and 24.20 kcal/mol for schemes A, B, C, and D, respectively. These energy barriers were obtained using HF/3-21G geometry optimizations for each point on the PES, followed by a single point B3LYP/6-31G* calculation on each point. The smallest energy difference between the reaction scheme with no general acid (scheme A, Figure 3) and a scheme with an explicit general acid is 4.5 kcal/mol. Based on these results, we conclude that scheme A is the most likely reaction mechanism. That is, our energy barrier calculations show that there is no general acid involved in the reaction.

For the case of the second step, only one reaction scheme was calculated. After scheme A was determined as the most likely reaction path, there is only one reaction path possible for the second step in this scheme. The reason is that because there is no general acid in the first step, we only need to calculate the proton transfer from Pro-1 to the substrate to form the product in the second step, as mentioned in section 2.3 when we discussed the reaction driving for the second step of the reaction. Figures 4−6 show the calculated structures for the reactant, intermediate, and product states.

The potential energy barrier for the first step is 17.78 kcal/mol and is centered around $R_1 = 0.24$ Å (see Figure 7A). In the second step, the calculated potential energy barrier is 18.71 kcal/mol relative to the reactant state as shown in Figure 7B, centered around $R_2 = -0.26$ Å. As in the potential energy surface (PES) of the first step, Figure 7B shows the B3LYP/6-31G* single point PES obtained from the HF/3-21G geometries. These results suggest that both reaction steps may be rate determining because of the small difference between energy barriers. Experimental results by Stivers et al.[26,29] have shown that the second step is the rate determining step.

(27) van Gunsteren, W.; Berendsen, H.; Colonna, F.; Perahia, D.; Hollenberg, J.; Lellouch, D. *J. Comput. Chem.* **1984**, *5*, 272.
(28) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327.
(29) Stivers, J. T.; Abeygunawardana, C.; Mildvan, A. S.; Hajipour, G.; Whitman, C. P. *Biochemistry* **1996**, *35*, 814.

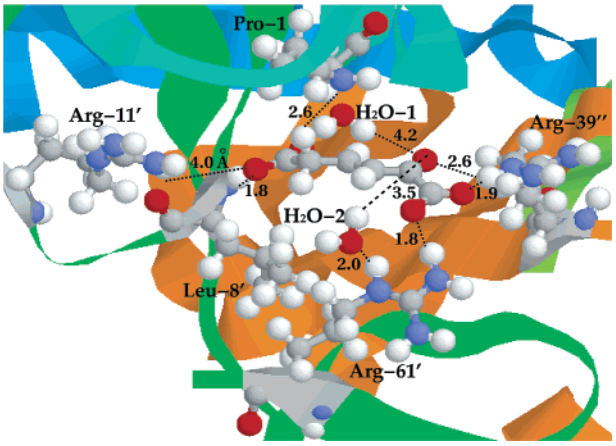


Figure 6. Product for the calculated 4-OT reaction path with significant distances from substrate atoms to atoms in residues in the active site.

To obtain a more accurate estimate to the reaction rate, we have used the FEP method to calculate the free energy surfaces for both steps of the reaction for scheme A. In both cases, the approximate free energy surfaces were obtained by combining the QM part of the potential energy from the B3LYP/6-31G*//HF/3-21G QM/MM calculations and the free energy perturbation calculation. Figure 7A and B also shows the obtained free energy surfaces for the first and second steps of the reaction catalyzed by 4-OT, respectively. The overall surfaces are similar to the respective PESs, suggesting that there is no significant contribution from the MM fluctuations to the overall reaction. The vibrational frequencies of the stationary points in these PES were determined, and they reveal that all minima (reactant, product, and intermediate) have no imaginary frequencies. In the case of the TSs, the vibrational frequency calculations reveal a single imaginary frequency for each TS.

After the contribution of the QM subsystem fluctuation was added to the free energy (see Table 1), the calculated free energy barrier for the first step is 14.54 kcal/mol, 3.22 kcal/mol lower than the calculated potential energy barrier. In the second step, the barrier is 16.45 kcal/mol with respect to the reactant (ES complex). These free energies of activation are in agreement with experimental findings that the second step is rate determining.

The calculated overall free energy barrier is 16.45 kcal/mol which is approximately 3.4 kcal/mol higher than the experimental determination of ∼13 kcal/mol.[8,9] However, note that this experimental value is only approximate because it was calculated from the reported experimental k_{cat} using the transition-state theory.

3.2. Analysis of the Mechanism. For the first step, the change in the reaction coordinate (R_1) is dominated by the proton transfer from C-3 to Pro-1. Although there is no general acid per se in the chosen reaction path, two ordered water molecules in the active site were observed to get closer to the carbonyl oxygen where a negative charge is formed in the intermediate (Figure 5). In the case of H_2O-2 the decrease in distance is 0.5 Å, whereas for H_2O-1 the decrease is 0.2 Å, in agreement with the proposed reaction mechanism.[9] The hydrogen atom bound to N_ϵ of Arg-39″ also changes position relative to the oxygen. The rest of the interactions to the different residues (Leu-8′, Arg-11′, and Arg-61′) remain unchanged. We thus conclude that these two water molecules and Arg-39″ help stabilize the charge on the oxygen via electrostatic interactions.

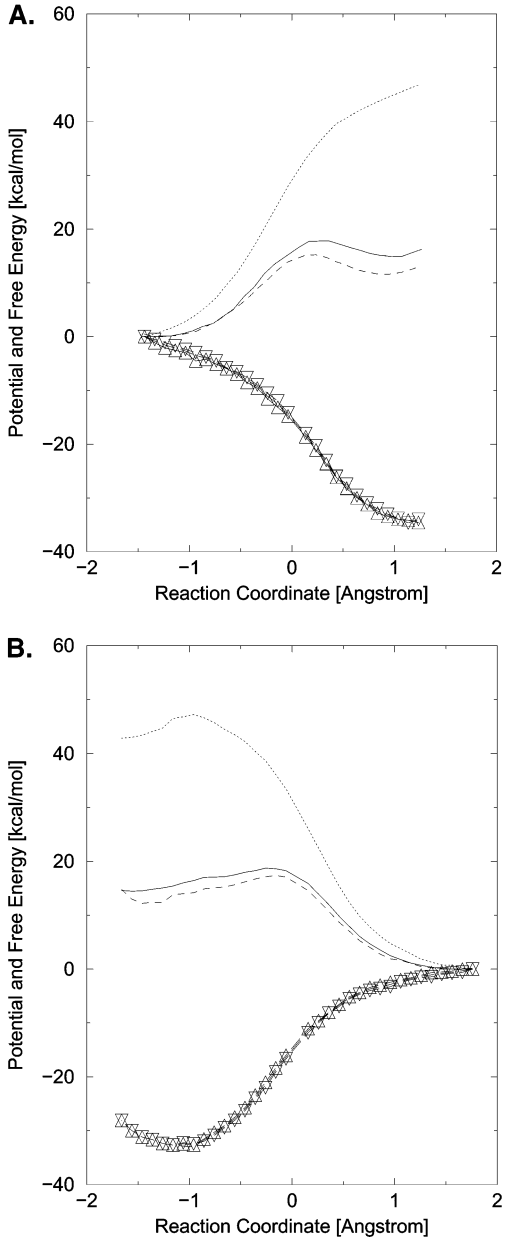


Figure 7. Relative potential and free energy surfaces for the first and second steps, A and B, respectively. (—) Total potential energy surface; (- - -) total free energy surface; (— — —) free energy associated with interactions with MM subsystem; (···) QM potential energy. The forward (△) and backward (▽) FEP results are also shown.

The potential energy surface for the second reaction step presents a shoulder close to the beginning of the reaction. This shoulder corresponds to the substrate's rearrangement in the active site for the hydrogen transfer from Pro-1 to C-5 of the substrate. After the substrate molecule has shifted, the hydrogen transfer takes place from the N-terminal proline to the substrate to form the final product.

In the process of the product formation, the two ordered water molecules that helped stabilize the negative charge on the oxygen of the C-2 carbonyl group move away (see Figure 6). This is more noticeable in H_2O-1 with an increased distance of 0.6 Å. All the residues interacting with the substrate in the active site remain the same except for Arg-11′, which moves away by 0.2 Å.

Table 1. Calculated Potential and Free Energy Differences (in kcal/mol) between the Determined Structure and the Reactant (ES complex), Where ΔE_a Is the Total HF Potential Energy Difference and ΔE_b Is the Total B3LYP Potential Energy Difference[a]

structure	ΔE_a	ΔE_b	ΔE_{QM}	$\Delta F_{QM/MM}$	$\Delta F^{fluct.}_{QM/MM}$	$\Delta E_{QM} + \Delta F_{QM/MM}$	ΔF
TS1	22.02	17.78	35.47	−20.29	−0.64	15.18	14.54
I	17.27	14.84	44.47	−32.87	1.92	11.60	13.52
TS2	22.50	18.71	36.08	−18.75	−0.88	17.33	16.45
P	0.63	1.12	0.66	−0.04	−0.34	0.62	0.28

[a] ΔE_{QM} refers to the QM energy difference between two QM subsystems. $\Delta F_{QM/MM}$ is the free energy change in the QM/MM interaction. $\Delta F^{fluct}_{QM/MM}$ refers to the contribution of the QM subsystem fluctuations to the free energy, and $\Delta F = \Delta E_{QM} + \Delta F_{QM/MM} + \Delta F^{fluct.}_{QM/MM}$.

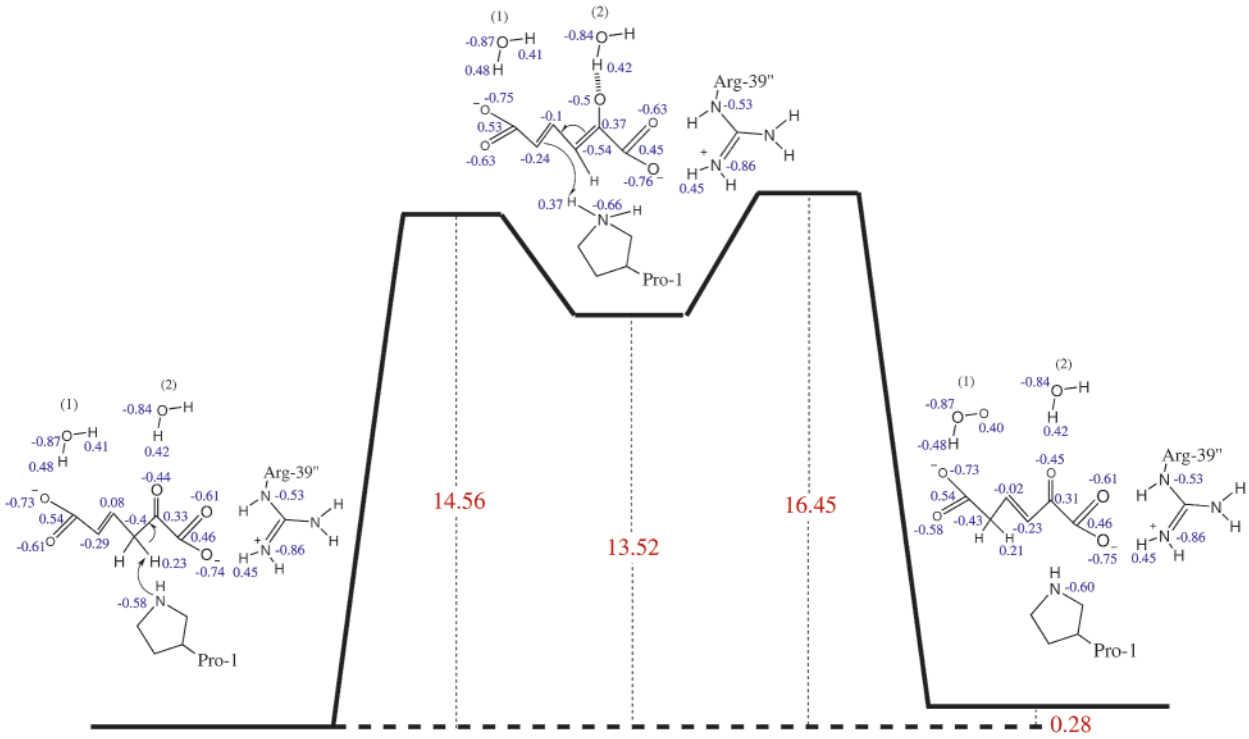


Figure 8. Calculated reaction scheme for reaction step one (scheme A) and two. Numbers in red indicate total free energy barriers (in kcal/mol), numbers in blue indicate Mülliken charges for molecules in the QM subsystem (substrate, Pro-1, and H_2O-1) and EPS charges for molecules in the MM subsystem (in a.u.).

The free energy surface for the second step is also essentially the same as the corresponding potential energy surface for this portion of the reaction scheme (see Figure 6B). The shoulder at the beginning of the reaction was also observed to be centered around the same distance in the reaction path (R_2). Close examination of the potential and free energy surfaces for the second step show that the initial part of the reaction, $R_2 = -0.9$ to -0.6 Å, is dominated by the substrate rearrangement in the active site to accept the proton from the N-terminal proline. After this shift in the substrate has taken place, the reaction is dominated by the proton transfer to form the product, 2-oxo-3-hexenedioate. As the bond between C-5 and the incoming hydrogen atom forms, the water molecules that helped stabilize the negative charge in the carbonyl oxygen move further away from it.

Figure 8 shows both steps of the reaction as well as selected charges on the substrate and certain residues. The charges corresponding to atoms in the QM subsystem were obtained from Mülliken charge analysis, whereas, in the case of atoms in the MM subsystem, all reported numbers correspond to MM charges in the force field.

A significant change in charge is observed on the carbonyl oxygen at C-2 of the substrate, going from reactant to intermediate and then to products. In the first step, the charge on this oxygen decreases by −0.06, while, in the second step, it increases by 0.05, producing a final charge almost equal to the initial one. This explains why H_2O-2 moves closer to this oxygen in the intermediate state and subsequently moves further away as the reaction progresses to products. C-3 shows a significant negative charge in the reactant state of −0.4, which decreases in the intermediate to −0.54 and increases significantly in the product state giving a final charge of −0.23. The leaving proton presents a positive charge that increases as the reaction goes to the intermediate and decreases in the product state. These results are consistent with a proton abstraction mechanism.

Note that the charge in C-5 remains basically unchanged in the first step of the reaction (−0.29 units); however, the charge in this carbon becomes more negative in the product state (−0.43 units) as the H atom binds to it. This decrease in charge is balanced by the increase in charge on C-3 in the final state, where the proton was abstracted from in the first step. A charge redistribution was also observed in the remaining oxygens that form the carboxylate moieties in the substrate. All four oxygen atoms experience a slight decrease in charge, followed by a small increase again going from reactant to intermediate and

subsequently to product. This charge redistribution is also noticeable in the Pro-1 nitrogen.

Given the strong interaction of H_2O-2 observed in the reaction process, we investigated the effect of including H_2O-2 in the QM subsystem. The calculation of the PES including this water molecule in the QM subsystem produced a very small variation in the barrier height. The difference in energy between the ΔE for the barrier calculated with H_2O-2 in the QM subsystem and without H_2O-2 in the QM part is -0.38 and 0.73 kcal/mol for the first and second steps, respectively. This shows that the effect of this water molecule is mainly electrostatic and it can be described with the MM force field.

3.3. Effects of Individual Residues. Understanding the role of individual residues in the reaction steps may provide us with further insight into the mechanism of 4-OT. To help comprehend these roles, an analysis of the overall effects of the MM environment is necessary. Note that the information gained from this "breakdown" is only obtained in an approximate manner, to facilitate our understanding of the system of certain experimental results because the energy decomposition is performed through perturbative calculations.

Decomposition of the free energy barrier associated with the two reaction steps may not be possible. However, breaking the energy barrier into components is possible. Here we followed the procedure of Liu et al.[30] Considering the changes in electrostatic interaction energies between individual residues and the QM subsystem when the system goes from reactants to transition states,

$$\Delta E = \langle E_{\mathrm{i/QM}}^{\mathrm{electrostatic}} \rangle_{\mathrm{MM,TS}} - \langle E_{\mathrm{i/QM}}^{\mathrm{electrostatic}} \rangle_{\mathrm{MM,reactant}}$$

where i represents an individual residue (or water molecule) in the MM subsystem, $E_{\mathrm{i/QM}}^{\mathrm{electrostatic}}$ represents the electrostatic interaction energies between residue i and the QM subsystem, and $\langle \cdots \rangle_{\mathrm{MM,TS}}$ and $\langle \cdots \rangle_{\mathrm{MM,reactant}}$ represent the averages over the ensembles where the conformational space of the MM part is sampled with the QM part corresponding to the reactant (or intermediate) and the TS, respectively. ΔE_{i} can only be viewed as a first-order approximation to the contribution of residue i to the stabilization of the TS, which means that other factors such as conformational changes or dielectric screening are absent. Negative ΔE_{i} indicates that the residue lowers the energy barrier and a positive ΔE_{i} indicates the opposite.[30]

The calculation for the electrostatic contributions per residue was carried out using only the conformations sampled by the free energy simulations. The ensemble averages were approximated by averages over 50 conformations for both the first and second steps sampled by the free energy simulations. After this analysis, we observed that only a few residues contribute significantly to either step of the reaction ($|\Delta E_{\mathrm{i}}| \geq 1$ kcal/mol). These residues (and H_2O-2) are shown in Table 2. The sums of all residue contributions give -11.41 for the first step and -12.65 kcal/mol for the second step. These results justify the qualitative contributions from individual residues because these are relatively good approximations to the FEP results of -14.54 and -16.45 kcal/mol for the first and second steps, respectively. Note however that the sum of all residue contributions only includes the electrostatic interaction between the QM subsystem and the residues and H_2O-2; the remaining contributions come from the solvent molecules, which account for the approximately 4 kcal/mol discrepancy for both steps.

Table 2. Individual Residue Contributions to TS Stabilization[a]

	ΔE (kcal/mol)	
residues	first step	second step
Arg-39″	−3.24	−3.39
Arg-11′	−0.30	−1.29
Phe-50′	−1.27	−1.10
Ile-2	−0.69	−1.16
Pro-34	−0.79	−1.19
Ser-37	−3.60	−4.26
Val-38	1.52	1.65
Arg-39	−2.98	−2.07
Ile-7′	1.21	0.26
Leu-8′	−1.37	−0.28
Gly-48′	−0.83	−1.57
Gly-51′	1.18	1.08
Ile-52′	−1.36	−1.47
H_2O-2	−2.08	−2.02

[a] Negative values mean that that particular residue helps stabilize the TS, while a positive value means that the residue destabilizes the TS.

In addition, we have also calculated the Lennard−Jones contributions to the energetics of the TSs. The sum of the contributions for all residues are 0.9 and -0.6 kcal/mol for the first and second steps, respectively, which is less than 10% of the total electrostatic contributions per residue.

From Table 2, it can be seen that several residues contribute to TS stabilization, In particular, Arg-39″, Arg-11′, and Phe-50 are important contributors, in agreement with experimental findings.[9,10,31] H_2O-2 also contributes significantly to the stabilization of the TS in both steps. Note that the comparisons between our calculations and experimental mutagenesis results only relate approximate experimental $\Delta\Delta G$'s obtained by using the transition state theory (TST) to our calculated electrostatic contributions. Our results provide no information on structural change and/or electrostatic contributions from the new residue in the mutant structure which could play a significant role.

Arg-39″: Site directed mutagenesis studies in which mutations of Arg-39 to Ala and Gln were carried out[9,10] suggest a catalytic and structural role for these residues. Mutation of Arg-39 to Ala produces an increase in the overall barrier ($\Delta\Delta G$) of approximately -2.85 kcal/mol. This $\Delta\Delta G$ is determined by the TST using the experimentally determined k_{cat} for the native and mutant structures. In the case of the Ala39Gln mutant, structural changes were also observed from which it was concluded that Arg-39″ helps to stabilize the β-hairpin that covers the active site.[9] Table 2 from our calculations shows that both Arg-39″ and Arg-39 are major contributors to TS stabilization in the free energy picture for both steps of the reaction. The addition of the contributions from all Arg residues gives a total TS stabilization of -6.28 and -5.50 kcal/mol for the first and second steps, respectively, which qualitatively agrees with experimental findings. Note that the calculated values are only approximate because of the perturbative nature of the calculation and the limitation of only including electrostatic components.

Arg-11′: It was also determined that this residue acts as a binding site for the substrate as well as an electron sink.[9,10] Mutation of Arg-11′ to Ala produces an approximate $\Delta\Delta G = -2.64$ kcal/mol,[9,10] as determined from the TST. Our calcula-

(30) Liu, H.; Zhang, Y.; Yang, W. *J. Am. Chem. Soc.* **2000**, *122*, 6560.

(31) Czerwinski, R. M.; Harris, T. K.; Massiah, M. A.; Mildvan, A. S.; Whitman, C. P. *Biochemistry* **2001**, *40*, 1984.

tions show that this residue does not significantly contribute to the first step of the reaction and acts to stabilize the TS in the second step (see Table 2) by −1.29 kcal/mol.

Phe-50′: Residues as far away as 9 Å from Pro-1 have been linked to the catalytic properties of 4-OT by enabling the general base to remove the proton from the substrate by lowering its pK_a. Phe-50 has recently been suggested to have an important effect in lowering the dielectric constant of the active site, thus lowering the pK_a of the general base (Pro-1).[12,31] In the case of the Phe50Ala mutant, an experimental $\Delta\Delta G = -3.02$ kcal/mol is observed.[31] Our calculations show that several apolar residues, some within 9 Å of the active site, have strong contributions to the stabilization of the TS. These residues include Phe-50′, Ile-2, Leu-8′, Gly-48′, and Ile-52′. Yet, some other residues with these characteristics (Val-38, Ile-7′, and Gly-51′) destabilize the TS (see Table 2). For Phe-50′, the contribution is a stabilizing one of −1.27 and −1.10 for the first and second steps, respectively.

It is also interesting to point out that Phe-50′ and Gly-51′ have been determined to play an important structural role.[31] This was determined by the loss of structure of the β-hairpin from residues 50 to 57 in the Phe50Ala mutant. In the native structure, this hairpin covers the active site.

Ser-37: An interesting result from our calculation shows that Ser-37 is a major contributor to the stabilization of the TS. The major contributing atoms for this residue are the hydroxyl oxygen of the side chain and the carbonyl oxygen of the backbone, −2.6 and −0.78 kcal/mol, respectively, for the first step and −3.26 and −1.44 kcal/mol, respectively, for the second step. These two atoms account for the majority of the contribution from this residue. Both oxygen atoms of this residue are located within 4 Å of the N atom in Pro-1; however, the hydroxyl group does not form an H-bond with the N of Pro-1 or the substrate. As explained previously, this residue forms an H-bond to one of the conserved waters in the active site of the inactivated crystal structure. Ser-37 would be a very interesting target for mutation studies.

Gly-48′ and Gly-51′: Two glycine residues were found to contribute to (de)stabilization of the TS for both steps. In the case of Gly-48′, the atoms that are mostly responsible for this contribution are C_α, the carbonyl C and O atoms, and the H atom bound to the backbone N of this residue for both steps. The contribution from these atoms accounts for −0.73 and −1.3 kcal/mol of the total stabilization from this residue (see Table 2). For Gly-51′, which has a destabilizing effect on the TS, the atoms that contribute to this stabilization according to our calculation are the backbone N, C_α, and the carbonyl C and O atoms. They account for 0.96 and 1.21 kcal/mol for the first and second steps, respectively. The reason these two residues play opposite roles is that in the case of Gly-51′ the carbonyl atoms effectively cancel their contributions in turn making the N and H atoms responsible for the overall contribution. This effect does not occur for Gly-48′ for which the contribution from the carbonyl C is much less than that of the oxygen, providing an overall stabilization along with the other atoms.

The overall calculated effect from the apolar residues within 9 Å of the active site suggests a small stabilizing role from these amino acids. Note that the calculated effects from the present study only account for electrostatic contributions. It would seem counterintuitive for several of the nonpolar neutral residues such as Phe-50′, Gly-51′, and Ile-52′ to interact with the TS. However, these residues are within 5 Å of the QM region and hold slight partial charges. Other nonpolar neutral groups such as Leu-8′, Gly-48′, and Gly-51′ interact with the substrate via backbone atoms of these residues. In the case of Leu-8′, an H-bond is formed between the backbone N−H and a carboxylate oxygen of the substrate.

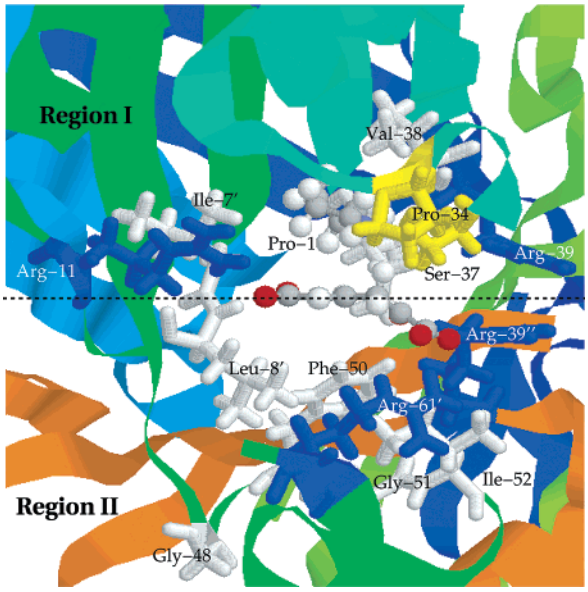


Figure 9. Front view of the 4-OT active site with important residues from Table 1: substrate and Pro-1 are represented as CPK in ball-and-stick form; apolar residues are represented in white; polar residues are represented in yellow; and arginines 39, 11′, 61′, and 39″ are represented in blue.

In the case of Arg-61′, it has been determined from mutagenesis studies[9] that it does not play a role in the catalysis or in substrate binding. Our calculations show that this residue does not significantly contribute to the stabilization of the TS, (−0.13 and −0.40 kcal/mol in the first and second steps, respectively), in agreement with experimental catalytic results. Our results also show that this residue H-bonds to the substrate to provide anchoring during the reaction. Note that since we did not perform binding energy calculations, we cannot relate this result to experimental binding results.

An analysis of the positions of polar and apolar residues around the active site can also provide further insight into the mechanism of 4-OT. Figures 9 and 10 show front and side views of the active site, respectively, with polar and nonpolar interacting residues. In Figure 9, region I is defined as the region above the plane through the substrate in the active site. Region II is defined as the region below this plane. In Figure 10, the enzyme and surface regions are approximately divided by a vertical plane through the substrate.

Most of the important nonpolar residues are in the intersection of the enzyme part and the region II. (see Figures 9 and 10). These include Phe-50′, Val-38, Leu-8′, Gly-51′, and Ile-52′ and are located within 8 Å of Pro-1.

Arg-39″ is also located on the enzyme side of the active site. It functions as an anchor to the substrate and stabilizes the negative charge formed in the intermediate during the reaction. The only other polar residue located on the enzyme side is Arg-39. The rest of the polar residues are situated on the surface of the enzyme (Figure 10). This may allow water molecules to

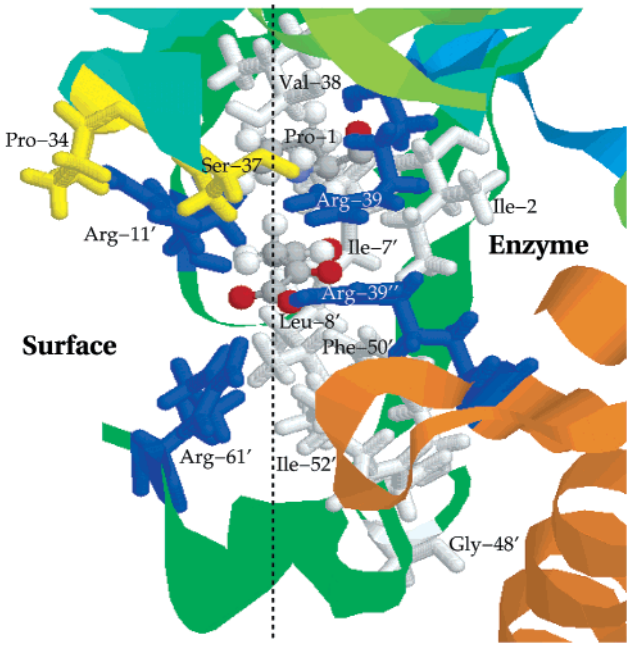


Figure 10. Side view of the 4-OT active site with important residues from Table 1: substrate and Pro-1 are represented as CPK in ball-and-stick form; apolar residues are represented in white; polar residues are represented in yellow; and arginines 39, 11′, 61′, and 39″ are represented in blue.

enter to the active site. In turn, the water molecules, along with Arg-39″, contribute to stabilize the negative charge in the intermediate state.

4. Conclusions

The reaction path for 4-OT was calculated by means of a combined QM/MM method developed in our laboratory. Potential energy and free energy surfaces and profiles were determined for both steps involved in the reaction. For the first step, the calculated barriers are of 17.78 and 14.54 kcal/mol for the potential and free energy calculations, respectively. In the second step, the barriers are 18.71 and 16.45 kcal/mol. These free energy barriers are only slightly higher than the experimental estimation of 13 kcal/mol.[9]

The calculated reaction path confirms the previously proposed mechanism based on experimental findings by Harris et al.[9] This mechanism involves two water molecules in the active site that help stabilize the negative charge formed in the intermediate species. Arg-39″ is also involved with charge stabilization. Our calculations indicate that no residue or water molecule acts as an explicit general acid; Instead, two ordered water molecules and Arg-39″ help stabilize the negative charge formed in the substrate molecule via electrostatic interactions. Our calculations also suggest that Arg-11′ has a predominantly anchoring role in this reaction. These results shed light on the search for a general acid for the reaction catalyzed by 4-OT, which had proved inconclusive in previous studies.[9,11]

Energy decomposition analyses were performed on both steps of the reaction to qualitatively explain individual residue contributions. Our calculations show that several nonpolar residues in the active site have a slightly stabilizing effect on the TS. Arg-39″, Arg-39, and Arg-11′ along with other polar residues contribute strongly to the stabilization of the TS, which produces an overall stabilizing effect in agreement with experimental results.

Acknowledgment. Support from the National Institute of Health is gratefully acknowledged. Computing time from the North Carolina Supercomputing Center was used for the calculations. G.A.C. also wishes to thank CONACyT for financial support. The authors would like to thank Prof. M. Fitzgerald and Mr. P. Silinski for enlightening discussions. The authors wish to thank the reviewers for valuable comments.

Supporting Information Available: Selected geometry parameters for the $C-C_{ps}$ and $N-C_{ps}$ are given as well as a schematic representation of the test molecules used to obtain these results. Also shown are the schematic representation of the isomerization catalyzed by 4-OT and a figure of the superposition of the backbones of the crystal structure 4-OT with our calculated initial structure. This material is available free of charge via the Internet at http://pubs.acs.org.

JA029672A