Quantum Mechanical/Molecular Mechanical Studies of the Triosephosphate Isomerase-Catalyzed Reaction: Verification of Methodology and Analysis of Reaction Mechanisms

Qiang Cui[†] and Martin Karplus*[,†,‡]

Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, and Laboratoire de Chimie Biophysique, ISIS, Universitè Louis Pasteur, 67000 Strasbourg, France

Received: July 11, 2001; In Final Form: November 6, 2001

Three possible mechanisms for the reactions catalyzed by triosephosphate isomerase (TIM) have been studied by the combined quantum mechanical/molecular mechanical (QM/MM) approach at a number of QM levels including AM1, AM1 with specific reaction parameters (SRP), and B3LYP/6-31+G(d,p). The comparison of the various QM levels is used to verify the adequacy of our recent B3LYP/MM analysis of the reaction mechanism (Cui et al. *J. Am. Chem. Soc*. **2001**, *123*, 2284), which showed that the intramolecular proton transfer pathway is ruled out, due largely to the unfavorable interaction between the transition state and His 95. The relative contributions from the two other proposed pathways, however, are difficult to determine at the present level of theory; both pathways are also consistent with available experiments. To obtain information about the role of the enzyme, density functional calculations were made for model systems in the gas phase and in solution; selected models were also studied with ab initio calculations at the levels of MP2 and CCSD to confirm the B3LYP results. Mulliken population analysis of the transition states demonstrates that hydrogen transfers essentially as proton for all the reactions in TIM, with an electron population between +0.33 and +0.44. Adiabatic mapping calculations for path A indicate that the two relevant proton-transfer steps between the substrate and His 95 proceed in a nearly concerted manner. The QM model calculations in solution and a QM/MM perturbation analysis shows that a number of factors combine to yield the rate enhancement by a factor of 10^9 in TIM. These include orienting catalytic groups (e.g., Glu 165, His 95) in good positions for the proton transfers, employing charged and polar groups (e.g., Lys 12, Asn 10) that stabilize the reaction intermediates and permitting flexibility of the catalytic groups (e.g., Glu 165 along path C). Some residues far from the active site, such as the main-chain atoms in Gly 210, as well as certain water molecules, also make significant contributions. For the electrostatic interaction and polarization to function effectively, the active site of TIM has a relatively low effective dielectric "constant", which reflects the structural integrity of the enzyme active site as compared with solution. Short hydrogen bonds occur during the reaction (e.g., between the reactant substrate and Glu 165), but the calculated energetics indicate that they do not have a specific role in catalysis; i.e., no contribution was found from the rather short hydrogen bond between His 95 and the substrate in path A.

I. Introduction

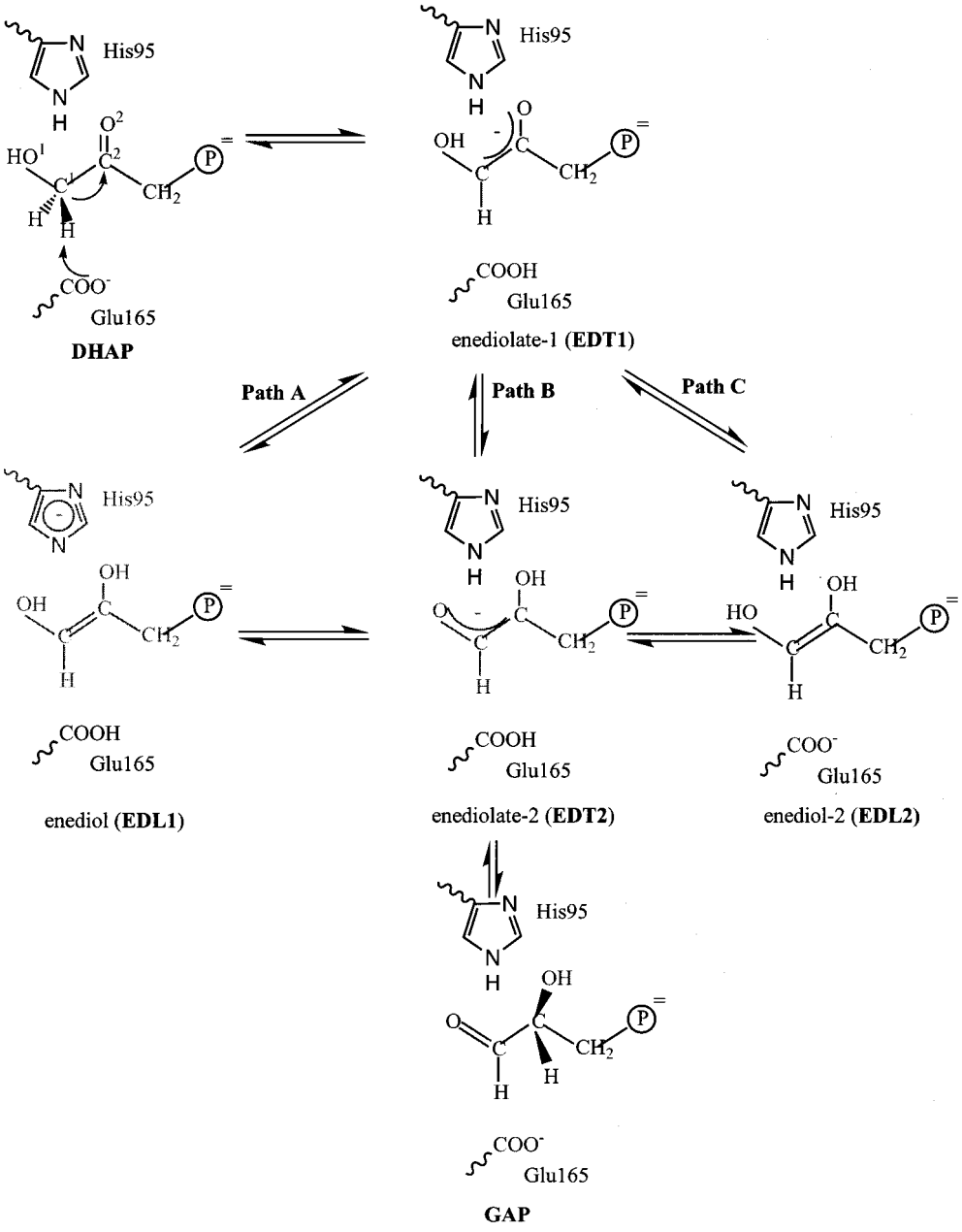
Triosephosphate isomerase (TIM) is a dimeric enzyme that catalyzes the conversion between dihydroxyacetone phosphate (DHAP) and *R*-glyceraldehyde 3-phosphate (GAP), which is an important step in the glycolytic pathway. TIM increases the rate of the chemical step of the reaction by more than 10^9 times relative to that in solution.[1] As a result, diffusion of the product out of the active site is rate-limiting and TIM has been referred to as a "perfect" enzyme.[2] Many experimental techniques[3] including kinetics,[4,5] X-ray crystallography,[6,7] nuclear magnetic resonance (NMR),[8−10] infrared spectroscopy (IR),[11] and site-specific mutagenesis[12,13] have been used to study the enzyme. These have been supplemented by a number of theoretical calculations.[14−20] Nevertheless, the detailed catalytic mechanism is not fully understood and there continue to be alternative proposals for the chemical steps involved in the overall reaction.

To make clear the present situation concerning the mechanism of the TIM-catalyzed reactions, we summarize the different proposals in Scheme 1. It is generally accepted, in accord with experiment and theory, that the first step involves the transfer of a proton from C^1 of the substrate, DHAP, to a carboxyl oxygen of the catalytic base, Glu 165, which yields the enediolate EDT1 with a proton on O^1. However, for the second step, there are several proposals. Two of them involve generation of the enediol, EDL, by transfer of a proton either from His 95 (path A) or from the protonated Glu 165 (path C); the enediol then forms the enediolate EDT2 by transferring the proton O^1 back to His 95 (path A) or Glu 165 (path C). The third proposal (path B) involves internal proton transfer from O^1 to O^2 in the enediolate to form EDT2 directly from EDT1, without any proton transfer from neighboring groups. The enediolate EDT2 is common to all paths and the final step is transfer of a proton from Glu 165 to C^2 of EDT2 to form the product, GAP. The observed proton exchange between solvent and the enzyme with bound substrate indicated that an intermediate is involved in

[†] Harvard University.
[‡] Universitè Louis Pasteur.

10.1021/jp012659c CCC: $22.00
Published on Web 01/29/2002

SCHEME 1: Proposed Catalytic Mechanisms for TIM-Catalyzed Reactions



the catalyzed reaction,[4] although there may be more than one; the identity of the intermediate(s) is not clear from experiments. The experimentally estimated free energy profile suggests that the slowest chemical step is the conversion between DHAP and the intermediate, which has a rate constant of $2 \times 10^3\ s^{-1}$ (i.e., this corresponds to a free energy barrier of 13 kcal/mol). The overall equilibrium constant for the TIM-catalyzed reactions is close to 1 (see footnote 5 in ref 2) and therefore can also be studied from the other direction, i.e., with GAP as the substrate and converting it into DHAP.[4] In this direction, the slowest chemical step is the conversion from the intermediate to DHAP, which has a rate constant of $6 \times 10^3\ s^{-1}$ (i.e., corresponds to a free energy barrier of 12.4 kcal/mol).

Thus, the part of the catalytic mechanism that remains controversial concerns whether an enediol is formed and, if so, whether the source of the proton is His 95 or Glu 165. Mechanism A is the one proposed originally by Knowles and co-workers.[2,4,6b] Although it is based on a range of experiments, there is no direct evidence for intermediate formation of an enediol during the reaction. Nevertheless, the mechanism has been generally accepted. It has been analyzed theoretically and supported by Bash et al.,[15] who used the QM/MM methodology[21] at the AM1[22]/CHARMM19[23] level. The step with the highest barrier was found to be the proton transfer from EDT1 to EDL1; the calculated barrier of about 16 kcal/mol is slightly larger than the experimental estimate of 13 kcal/mol for the overall reaction.[3b] Due to the fact the TIM-catalyzed reactions are at the diffusion limit, direct evaluation of the reaction free energy from kinetic measurement is not possible. The experimental estimates were obtained by use of the elegant isotope exchange−conversion technique; kinetic measurements with isotope-labeled substrate in ^1H_2O and unlabeled substrate in tritiated water were analyzed to yield the free energy profiles of the significant kinetic steps.[4] Bash et al.[15] studied the effect of protonation of His 95 and came to the rather surprising conclusion that a neutral (imidazole) rather than protonated His 95 was more likely to be the proton donor, because the latter would lead to too stable an enediol intermediate. The presence

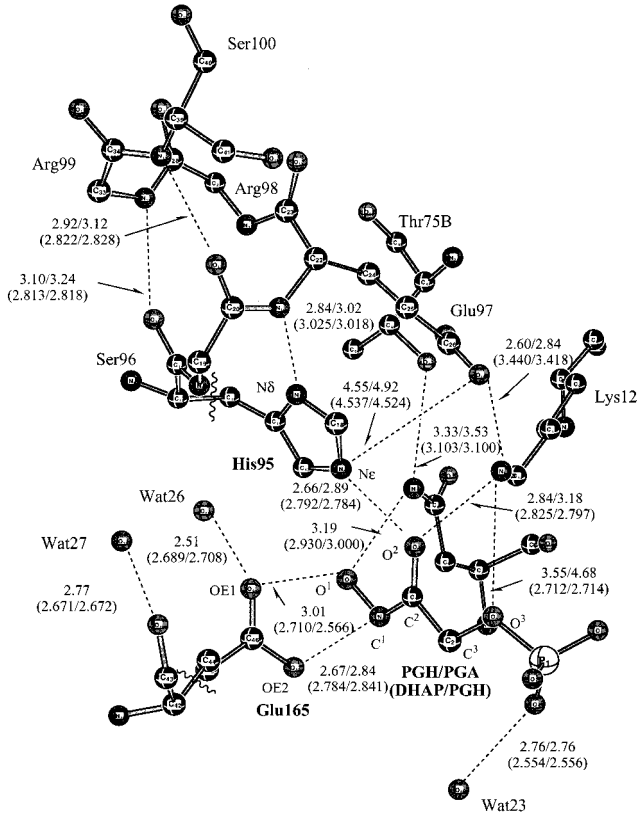


Figure 1. Arrangement of residues in the active site of the yeast triosephosphate isomerase (TIM).[7a] The distances (in angstroms) without parentheses are from X-ray structures; those before the slashes are for the PGH−TIM complex, and those after the slashes are for the PGA−TIM complex. The numbers in parentheses are obtained from B3LYP/6-31+G(d,p)/CHARMM optimizations; the values before and after the slashes are for the DHAP−TIM and PGH−TIM complexes, respectively. The wiggles indicate the dividing bond between the QM and MM region. No hydrogen atoms are shown for clarity. For more details, see section II of the text.

of His 95 in neutral form in the enzyme was subsequently confirmed by NMR[9] and since then has been shown to be involved in analogous enzyme-catalyzed reactions.[24] Through perturbation calculations, Bash et al. concluded that Lys 12, which had been presumed to be important in catalysis as well as binding of the substrate[25] due to its proximity (see Figure 1 for the active-site structure), has the largest effect on the energetics of the reaction. However, other *charged* residues were found to make a significant contribution in enhancing the reaction rate. Path A has also been studied by Åqvist and Fothergill[16] with the effective valence bond (EVB) method and free energy perturbation techniques.[26] The rate-determining chemical step was found to be the proton transfer to Glu 165 (a barrier of 12.5 kcal/mol vs ∼10 kcal/mol for EDT2 to GAP), which is in accord with the result that mutation Glu 165 → Asp 165 leads to a $\sim 10^3$ decrease in k_{cat}.[27] It is of course possible that even if the second step (EDT1 to EDL1) is rate-determining for the wild-type enzyme, the mutation Glu 165 → Asp 165 may make the first step rate-determining. Åqvist and Fothergill found that an active-site water (626 in the PGH−TIM X-ray structure) played an important role; they also suggested it explained certain mutation results28 (e.g., the reduced activity of the Ser 96 to Pro mutant, which might displace the water).13 Certain residues of the other subunit were found to be involved in the reaction, which could explain the low reactivity of a monomeric TIM.[29] No quantitative analysis of these proposed contributions, however, were performed.

Alagona et al.[14c] proposed an alternative mechanism (path B), in which an intramolecular proton transfer leads from EDT1 to EDT2 without proton transfer from His 95. The corresponding barrier height was calculated to be lower than those involved in path A by 5 kcal/mol in MP2/6-31+G(d) quantum mechanical calculations for a model system that did not include environmental effects. The result for the proton transfer from His 95 to EDT1 (path A) was endothermic by 20 kcal/mol, very different from the values obtained by Bash et al.[15] and by Åqvist and Fothergill;[16] see above. To rationalize the ∼200-fold reduction in the activity of H95Q and H95N mutants, they argued, though they did not test this, that Gln 95 and Asn 95 were not in as good positions to stabilize the enediolate as was His 95.

A third mechanism, path C in Scheme 1, has been discussed by a number of authors. In this path, Glu 165 is the only catalytic residue; it transfers protons during the reaction as shown in Scheme 1. According to the X-ray structure of wild-type TIM with bound phosphoglycolohydroxamate (PGH), large displacement of the Glu 165 side chain is necessary along path C, while the position of His 95 is ideal for the proton transfers proposed in path A. Therefore, this mechanism was proposed by Knowles and co-workers[12a] only for the H95Q mutant. Recently Harris et al.[10] carried out NMR experiments on yeast TIM and observed a strongly deshielded proton at 14.9 ppm in the enzyme with the inhibitor PGH in the active site. It is 6.2 ppm downfield relative to the free PGH in solution, which suggests that there is a strong hydrogen bond[30] between the carboxylate of Glu 165 and the N−OH proton of PGH. On the basis of this observation and the measured exchange rate of the His-95 NϵH proton with solvent, they proposed that path C may take place in the wild-type enzyme as well. In a set of isotope experiments, the extent of tritium transfer from the *pro-R* position in DHAP to GAP was studied as a function of the substrate concentration. Analysis of the results suggests that the generally accepted mechanism (path A) contributes at least 3.9% in yeast TIM. The partition between paths A and C was impossible to estimate because of the uncertainty in the efficiency of conservation and transfer of tritium in the enzyme active site. Path B was not discussed by the authors of ref 10, although one expects the isotope transfer to be similar to that along path A. Therefore, the experiments in ref 10 do not rule out path A or B and do not support path C. Prior to these experiments, Peräkylä et al.[18] had analyzed the mechanism of the TIM-catalyzed reaction in a series of ab initio calculations on model systems with emphasis on path C. They found that EDL2 is more stable than EDL1, which occurs in path A, so that they speculated that path C is more likely. However, no transition states were located in their study and the model was limited to the substrate, Glu 165, His 95, Asn 10, and Lys 12; the environmental effect was described with the finite difference Poisson−Boltzmann method. No calculation was performed to explain the observed importance of His 95 during the catalysis.

In a previous report,[33a] we considered the three suggested paths for the chemical reaction catalyzed by TIM and determined their relative importance. We performed theoretical calculations of the alternative pathways with the QM/MM methodology[31,32] at a high enough QM level [DFT at the B3LYP level with a 6-31+G(d,p) basis set] to obtain meaningful results. We found that the two pathways that involve an enediol species (paths A and C) give similar values for the barriers, and the calculated rates are in satisfactory agreement with experiment.[2,4] By contrast, the mechanism that involves intramolecular proton transfer in the enediolate (path B) was found to be energetically

unfavorable due to electrostatic interactions with His 95, a conserved residue in TIM from different organisms. A simple perturbation analysis found that charged (e.g., Lys 12) and polar residues (e.g., His 95, Asn 10, Thr 75B), plus a number of water molecules in the active site, are important for the energetics of the reactions in TIM. In this paper, we provide the results from gas-phase and solution models that lead to insights into the catalytic contribution of the enzyme. We also present the details of the computation methods and the tests that were made to verify that the methods are appropriate. In a separate paper,[33b] we report a reinvestigation of the dynamics of the reactions by going beyond the classical approach of Neria and Karplus[17] and determining the effect of quantum mechanical tunneling on the rate constant.[34,35] It has been suggested that the latter could be significant even at room temperature because several proton-transfer steps are involved.

Section II outlines the computational methods used in the present work; some details, including optimization of the AM1 parameters and comparisons of the results obtained at different QM levels, are given in Appendices. In section III, we describe the results of ab initio calculations on several model systems for the TIM-catalyzed reactions in the gas phase and in solution. The QM/MM calculations for the enzyme are presented and compared with the model system results. A discussion of the contribution of the enzyme to catalysis is presented in section IV.

II. Computational Methods

The computational methods used include semiempirical calculations at the AM1 level with standard parameters, AM1 with parameters optimized for the system of interest, and density functional (DFT) calculations with the B3LYP[36] method. Full quantum-mechanical calculations were done for model systems and compared with QM/MM results. Also, solvation corrections were calculated with PCM and COSMO approaches. For the enzyme, QM/MM calculations with the various QM approaches were performed. The methodology for all of the models investigated is presented in the rest of this section. Only the results obtained at the DFT level are given in the main text; comparisons with the other QM approaches are given in Appendix 1. Details concerning the optimization of the AM1 parameters are presented in Appendix 2.

II.1. Model Systems in the Gas Phase and in Solution. As the first step, a set of calculations for models of the TIM-catalyzed reactions were carried out in the gas phase and in solution. The motivations for these calculations are as follows. First, the B3LYP method,[36] which will be used as the QM level in the QM/MM calculations, has been shown to yield reliable structures and energetics for a variety of systems.[37] However, given the semiempirical nature of the method (i.e., the choice of the exchange-correlation functional and the mixing fraction of exact exchange),[38] it is necessary to evaluate its performance for the specific system under study by comparing the results with high-level ab initio calculations. Because of the computational expense involved, rather small model systems were examined. Second, the results for these model systems can be compared to those for reaction in the enzyme to obtain information concerning the role of the latter. Finally, the B3LYP results for model systems can be used to determine reaction-specific semiempirical parameters of the AM1 type, which can then be used in more elaborate studies of the enzyme, including free energy simulations and rate constant calculations.

Four sets of model systems were investigated (see Table 5 for definitions). The set composed of the smallest compounds, referred as the zero-order set, was subject to QM calculations at the highest level and used to test the DFT treatment (see Figure 24). The set includes models for the different species (substrate, possible intermediates, product) that participate in the reaction; they are models for DHAP, the enediol (EDL), the enediolate (EDT), and the product (GAP, which is equivalent to DHAP for this model). To obtain meaningful gas-phase test results (i.e., a phosphate group with charge -2 has positive energies for the highest occupied orbitals, which suggests that calculations with the phosphate group without including solvation effect would not be meaningful), the $-CH_2OPO_3^{2-}$ group in the various species was replaced by a hydrogen atom (see below). Calculations were also carried out for CH_3COO^-, CH_3COOH, imidazole, and imidazolate as models for the side chains of the catalytic residues Glu 165 and His 95 in different protonation states. A construct was set up to study the proton transfer between His 95 and the substrate in the gas phase; the position of Nϵ in the imidazole (see Figure 23c) was held fixed in position relative to C^1, C^2, and O^2 (see Scheme 1) of the model substrate in accord with the X-ray structure of TIM with bound PGH. All geometries were optimized at the B3LYP/6-31+G(d,p)[39] level. For the models of the species participating in the reaction, the geometries were also optimized at the MP2/6-31+G(d,p) level for comparison. Single-point energy calculations at the MP2/6-311+G(d,p)[40] and CCSD[41]/6-311+G(d,p) levels were also performed for selected structures to provide a further test of the adequacy of the computational method.

In the actual substrate DHAP and the other species participating in the reaction, there is a $-CH_2OPO_3^{2-}$ terminal group. The primary role in the TIM reaction of the phosphate group, $-OPO_3^{2-}$, is to obtain strong binding to the enzyme; i.e., the phosphate group appears not to be directly involved in the reaction, but it interacts strongly with Lys 12[42] and with the positive N-terminal end of helices, as well as some other residues and a number of water molecules.[3] To consider its effect on the reactive part of the system, three subsets of models (see Figure 24) with $-CH_2OCH_3$, $-CH_2PO_3H_2$, and $-CH_2PO_3^{2-}$, as the terminal group were compared; the group $-PO_3H_2$ is expected to behave similarly to $-PO_3^{2-}$ in the enzyme because of the charged and polar groups just mentioned that effectively "neutralize" the phosphate charge. The gas-phase calculations were performed at the B3LYP/6-31+G(d,p) level. In addition, solvated calculations were performed, since solvation is expected to have a large effect on the energetics and geometries of models including the phosphate group, as was found in previous studies of phosphate acids.[43] The geometries for $-OCH_3$ and $-PO_3H_2$ were optimized at the B3LYP-COSMO[44]/6-31+G(d,p) level. The less approximate PCM approach[45] was used to calculate single-point energies for the $-OCH_3$ structures for comparison. For the structures with $CH_2OPO_3^{2-}$, the COSMO calculations were plagued with convergence problems; similar problems with structure optimization of charged groups have been found in other studies.[81] To overcome these difficulties, geometries were optimized with the finite difference Poisson−Boltzmann (FDPB) module implemented in Jaguar,[46] which yielded stable results in most cases. To test the Jaguar results, the geometries of the various species terminated by $-OPO_3H_2$ were optimized with it and with COSMO. In the COSMO calculations, both the standard Pauling atomic radii and the recently proposed united atom (UA) radii[47] were tested. It has been shown[39] that the UA set of radii yields better solvation energies for small molecules, especially for negative ions, than the Pauling set. The calculations also are less expensive due to the fewer number of tessarae required because the hydrogen atoms do not need

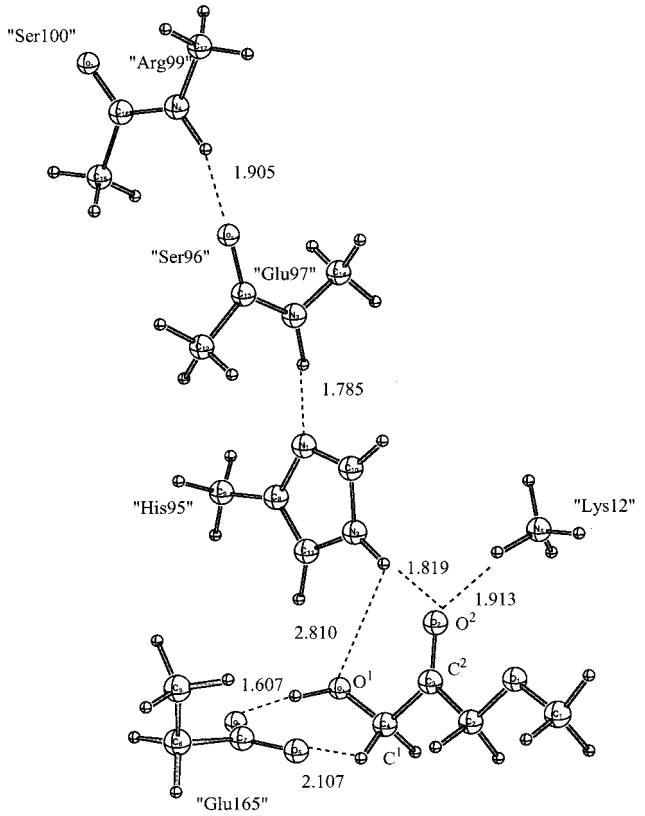


Figure 2. Structure that illustrates the various models studied in the absence of the enzyme (also see Table 5). The basic model consists of the substrate and one catalytic residue (Glu 165 *or* His 95). The active-site model consists of the substrate plus Glu 165 *and* His 95. The positions of the terminal CH moieties in "Glu 165" and "His 95" are fixed relative to the C^1, C^2, and O^2 atoms in the substrate during the geometry optimization to prevent large deviations from the structure in the enzyme. A few single-point energy calculations were also performed including Lys 12 (modeled by NH_4^+) and the "well-aimed" helix formed by the main-chain atoms of Glu 97, Ser 96, Arg 99, and Ser 100 in the active site of TIM (modeled by two NMA). The geometries used are derived from the optimized structures of the active-site model and the PGH−TIM X-ray data (see Figure 1).

to be represented as separate spheres. However, the UA set depends on the hybridization types of atoms, so that its application to transition states is not straightforward. In the present work, we used the averaged value of the radii in the reactant and product for the transition states. For the FD-PB calculations, the Pauling set of radii was used in accord with the default values in the Jaguar program.

The third set of structures, referred as the "basic model", treat the elementary proton-transfer processes. They include the substrate, intermediate, etc., and the catalytic residue (Glu 165 or His 95) directly involved (see Figure 2). To study the proton transfer between the substrate and His 95 in path A, the model includes the substrate with an $-OCH_3$ termination group and a 4-methylimidazole that mimics His 95. During the geometry optimization, one CH moiety in the CH_3 group of 4-methylimidazole was fixed relative to C^1, C^2, and O^2 (see Scheme 1) of the substrate in accord with the PGH−TIM X-ray structure. Such constraints prevent large displacements of the two fragments but give enough flexibility for adjustment of the relative positions. To study the other proton-transfer steps, the model includes the same substrate model and a $CH_3CH_2CO_2^-$ that represents Glu 165. The position of one CH moiety in the CH_3 group of $CH_3CH_2CO_2^-$ was fixed relative to C^1, C^2, and O^2 in the substrate. For simplicity, the backbone of the substrate (i.e., the $C^1-C^2-C^3-O^3$ dihedral) in most calculations was kept planar, which is a good approximation according to test calculations without any geometrical constraints. No such approximation was made for the proton transfer from EDT2 to GAP since C^2 is sp^3-hybridized in GAP, and the corresponding EDT2 structure is labeled np (nonplanar). For path B, although the intrinsic barrier for the intramolecular proton transfer in the enediolate form of the substrate (EDT1 to EDT2) was calculated with the zero-order model, it was also studied with the two models just described; i.e., the intramolecular protein transfer was calculated in the presence of 4-methylimidazole or $CH_3CH_2CO_2^-$. For comparison, AM1 results and the single-point B3LYP/6-31+G(d,p) energies at AM1 optimized structures were also determined. Single-point calculations at the MP2/6-31+G(d,p) level were made to further test the accuracy of the B3LYP/6-31+G(d,p) approach.

The most complete "active-site" model, in the absence of the full enzyme (see below) includes the substrate with the $-CH_2-OCH_3$ terminal group, $CH_3CH_2CO_2^-$, that mimics Glu 165 and 4-methylimidazole that mimics His 95 (see Figure 2). The same geometrical constraints described for the basic model were applied to the CH moieties of $CH_3CH_2CO_2^-$ and 4-methylimidazole. Two sets of structures, referred to as the syn and anti set, were considered. They are defined according to the orientation of the carboxylate group of Glu 165 relative to the substrate. It is believed that the syn orbital is more basic than the anti orbital,[48] although some controversy exists on this point.[49] Glu 165 has an approximately syn orientation relative to PGH in the wild-type TIM−PGH but an anti orientation in the H95Q-PGH[11b] and E165D structures;[25] the latter was suggested as one factor that makes these mutants less efficient than wild-type TIM. In this work we compare the two orientations to determine if there are significant differences between them. Also, the effect of the difference between His 95 and Gln 95 on the TIM reactions was studied with the same model by replacing His 95 by a Gln residue (see Figure 18). Because we are interested in the "intrinsic" difference between the contributions of His and Gln to the reaction energetics, we assume a similar orientation of the model Glu 165 in the calculations, despite the X-ray result that shows Glu 165 adopts a geometry in H95Q TIM that is significantly different from that in the wild type.[11b,19] Due to the large size of the active-site models, HF/3-21+G calculations were employed for the geometry optimization. Single-point energies for the HF/3-21+G geometries were performed at the B3LYP/6-31+G(d,p) level. Test calculations comparing geometry optimization with B3LYP/6-31+G(d,p) and HF/3-21+G for several species (i.e., DHAP, EDT, and EDT2) indicate that such an approach gives results very similar to full B3LYP/6-31+G(d,p) optimization. Single-point PB calculations were carried out at the B3-PB/6-31+G-(d,p) level with Jaguar for the more important syn structures. In addition, a number of structures with NH_4^+ mimicking Lys 12 in the active site of TIM were examined (see Figure 2). Guo and Salahub[50] have pointed out the existence of hydrogen-bonding networks in a number of enzymes including serine protease, Δ^5-ketosteroid isomerase, and TIM. Some of the networks involve active-site helices (e.g., TIM, subtilisin), while others do not (e.g., chymotrypsin-related serine proteases). Model calculations were carried out to illustrate the importance of the differential binding effect due to the hydrogen-bonding network. Here we examine the effect by introducing two *N*-methylacetamide (NMA) molecules that mimic a half-turn of the helix (residues 96−100) near the substrate (see Figure 2). In the active site of TIM, the main-chain atoms from these

residues form an extended hydrogen-bond network with the side chain of His 95, the substrate, and Glu 165 (see also Figure 1). Only single-point energies were calculated for these structures, whose geometries were constructed by adding NH_4^+ and NMA molecules to the optimized model system in accord with the X-ray structure of TIM−PGH. A more detailed study on the effect of hydrogen-bond networks on "charge-relay" type reactions (e.g., the TIM reactions) in enzymes with a new multilayer QM/MM approach will be reported separately.[51]

Most calculations were performed with Gaussian94,[52] supplemented with the latest routines for PCM and COSMO.[44b,45b] The FD-PB calculations were done with Jaguar.[46] Three basis sets, 3-21+G, 6-31+G(d,p), and 6-311+G(d,p), were used in the various calculations.

The results obtained by analyzing the reaction paths at the B3LYP/6-31+G(d,p) level are presented in section III. The results with other basis sets used to validate the methodology are described in Appendix 1.

II.2. Enzyme Model. To study the TIM-catalyzed reaction in the enzyme active site, the QM/MM method was employed with three QM levels. They are AM1, AM1-SRP (AM1 with specific reaction parameters), and B3LYP/6-31+G(d,p); different optimized QM/MM van der Waals parameters were used for AM1, AM1-SRP, and B3LYP (see Appendix section A1.2). The parameters used in AM1-SRP were derived with the B3LYP results for the model systems in the gas phase; details are given in the Appendix section A1.1. We note that different sets of specific reaction parameters were used for different reaction steps to achieve reliable AM1-SRP/CHARMM results; as indicated, it was found to be important to include AM1-SRP/MM calculations for model compounds during parameter fittings. The QM region includes the substrate and the side chain of Glu 165 in most calculations; the exception is the calculations for studying the process from EDT1 through EDL1 to EDT2 (path A in Scheme 1), for which the substrate and the side chain of His 95 are treated with QM. The enzymatic environment is described with the CHARMM22 all-atom force field for proteins.[53] Single-point calculations at the B3LYP/CHARMM level with the substrate and both Glu 165 and His 95 treated as QM were performed to obtain the best estimates of energetics in the present work; these are the values referred to in the discussions.

The enzyme system used the stochastic boundary model,[54] similar to that employed in the previous AM1/MM study of the TIM reaction.[15] However, there are some important differences, as noted below. Starting from the 1.90 Å resolution X-ray structure of yeast TIM complexed with the PGH transition-state analogue,[7a] the amide NH in PGH was replaced by a CH_2 group to obtain the initial position of the bound DHAP. A 16 Å spherical region centered on DHAP C^2 was selected, and the protein atoms outside this region were deleted; due to the group construct of the CHARMM force field, the deleted region (and therefore the remaining region) has an integer charge. The structure includes 118 residues of the subunit that contains the active site under study and nine residues (K69′, S71′, G72′, A73′, F74′, T75′, G76′, E77′, and Q82′) from the other subunit of the dimer. Explicit water (TIP3P) molecules were added to solvate the active site by overlaying a 16 Å water sphere; another three overlays were made after the water molecules were equilibrated by short (20 ps) molecular dynamics simulations and the stochastic boundary conditions (harmonic constraints on the protein boundary atoms and boundary forces[54c] on the water molecules) were applied. The harmonic force constants are derived from the temperature factors in the X-ray data and are scaled to zero at the buffer zone/reactive zone boundary with a smoothing function.[54a] The complete system consists of 1947 protein atoms and 131 water molecules. Test calculations at the AM1/CHARMM level with the atoms beyond the spherical region held fixed, rather than being deleted, were carried out and very similar results were obtained. Figure 1 shows the active-site structure, with selected geometrical parameters from both X-ray results of the PGH−TIM complex and the present B3LYP/6-31+G(d,p)/CHARMM optimization with either PGH or PGH replaced by DHAP. Geometrical parameters from the PGA−TIM structure[55] (PDB code 2YPI, 2.5 Å resolution) are included for comparison to show the variations found in the X-ray structure. The agreement between the four structures is quite good. The largest difference is found for the position of the Lys 12 side chain, which swings away from Glu 97 and comes significantly closer to O^3 of the phosphate group of the substrate in the B3LYP/CHARMM structure than in the X-ray structure. This is likely to be due to the scaling of the partial charges on Glu 97, which is discussed below. However, the position of Lys 12 relative to the chemically important parts of the substrate is well described; the $N_4^{K12}-O_2^{subs}$ distance is 2.84 and 2.83 Å from the X-ray data and our calculation, respectively, for the PGH structure.

A major difference between the current computational protocol and that used in the previous AM1/CHARMM study[15] is the treatment of electrostatics from charged residues. In the previous study, it was found that contributions from charged residues, some of which are more than 10 Å from the center of the active site, were significant. However, as many of the charged residues are exposed to the solvent, dielectric shielding can have a strong influence that is not adequately accounted for by the explicit water shell around the active site. To include such electrostatic shielding effects in the simulations, we employed the protocol described by Simonson et al.,[56] which was introduced to achieve faster convergence and more realistic structures in free energy simulations involving a change in the total charge of the system. Although the overall charge of the system does not change in the present calculations, it is appropriate to use the same approach because charge redistribution plays an important role. For each charged residue, a Poisson−Boltzmann (PB)[57] calculation is first carried out with the truncated enzyme ($\epsilon^{prot} = 4$) in a continuum solvent ($\epsilon^{solv} = 80$), keeping only the partial charges on this residue. Then a similar PB calculation is carried out in the vacuum. A scaling factor (greater than 1) is then introduced for each charged residue; it is equal to the ratio of the electrostatic potentials at the DHAP site from the two PB calculations. In the QM/MM calculations, the partial charges on the charged residues are reduced by the scale factors. The scale factors obtained from the UHBD program[58] for both the truncated (16 Å) and the full TIM systems are shown in Table 1. There are some differences in the exact values of the scale factors, because the enzymatic atoms beyond 16 Å of the substrate have been replaced by the continuum solvent in the truncated model. Nevertheless, the qualitative results are the same. The scaling factors are large in magnitude, and the effect of scaling is nearly equivalent to neutralizing all the charged residues on the surface. One exception is Glu 97, which has a relatively small scale factor (6.3) because it is buried inside the enzyme.

As described in ref 56, the effect of scaling is corrected after the simulations by two additional PB calculations. First, the partial charges of the scaled residues are changed back to normal and a PB calculation is carried out in a vacuum to obtain the change in the electrostatic free energy (ΔG_{scale}^{vac}) due to scaling.

TABLE 1: Scale Factors for Charged Residues in the Truncated TIM System

residue ID[a]	distance[b] (Å)	α[c]	residue ID[a]	distance[b] (Å)	α[c]
Asp 180A	17.6	74.6 (57.5)	Arg 26A	17.1	52.9 (37.2)
Asp 227A	17.1	67.9 (23.5)	Arg 98A	14.8	29.3 (14.2)
Asp 242A	17.1	75.9 (59.6)	Arg 99A	13.7	25.8 (13.9)
Glu 37A	18.2	58.6 (36.3)	Arg 205A	17.9	68.8 (28.9)
Glu 97A	7.0	6.3 (5.9)	Arg 247A	16.3	40.0 (22.4)
Glu 104A	16.0	45.6 (16.9)	Lys 134A	15.7	56.8 (45.3)
Glu 129A	13.4	21.1 (15.9)	Lys 237A	11.4	14.7 (13.5)
Glu 239A	13.3	25.4 (18.0)	Lys 69B	15.4	52.7 (40.8)
Glu 77B	16.1	61.9 (25.7)			

[a] The last letter in the residue ID indicates to which subunit the residue belongs. Subunit A contains the active site used in the simulation. [b] Measured from the Cγ in DHAP to Cζ of Arg, Cγ of Asp, Nζ of Lys, and Cδ of Glu, respectively. [c] The numbers without parentheses are obtained with the 16 Å spherical enzymatic model (which was used to obtain the results in the present paper), and those in parentheses are calculated with the full enzyme. In the Poisson−Boltzmann calculations, three focusing steps were used with grid sizes of 2.0, 1.5, and 1.0 Å, respectively, and a border spacing of 20.0, 10.0, and 5.0 Å, respectively. Note that these coarse grids were used only in determining the scale factors for the changes; finer sets of grids were used in the Poisson−Boltzmann calculations used in eq 1. The ionic strength and probe radius were set to zero. The interior of the protein is set to $\epsilon = 4$; the water and vacuum are described with $\epsilon = 80$ and 1, respectively.

The solvation free energy (ΔG_{solv}) of the system with the full charge is calculated with the second PB calculation. The total energy of the system is then estimated according to

$$E^{tot} = E_{QM/MM}{}^{vac} + \Delta G_{solv} + \Delta G_{scale}{}^{vac} \quad (1)$$

where $E_{QM/MM}{}^{vac}$ is the QM/MM potential energy calculated with the stochastic boundary system and the scaled charges. It should be noted that in eq 1 a free energy is added to the QM/MM potential energy so that E^{tot} corresponds to the effective energy of a given configuration of the enzyme in the presence of the canonically averaged interaction with the solvent.[59] In the PB calculations, the QM atoms are replaced by fixed Mulliken charges from the QM/MM calculations with the scaled partial charges. This is appropriate because, in the system we try to model (i.e., a fully solvated enzyme−substrate system), the wave function corresponds to the QM region not polarized by the full set of partial charges. Calculations also show that although $\Delta G_{scale}{}^{vac}$ and ΔG_{solv} are large in absolute magnitude, their contributions cancel nearly perfectly if relative energies for structures on the reaction paths are considered. For example, for the conversion from DHAP to EDT1, the differences in $\Delta G_{scale}{}^{vac}$ and ΔG_{solv} are 0.3 and −0.4 kcal/mol, respectively, with $\epsilon^{prot} = 4.0$, while the individual values for DHAP and EDT1 are (−1337.9, −170.2) and (−1337.6, −169.8) kcal/mol, respectively. This is reasonable because the total charge of the system is conserved during the TIM reactions, and only the charge distribution of the substrate and a few key residues in the active site change significantly. Given such a small effect, $E_{QM/MM}{}^{vac}$ is used directly in the analysis.

The paths for the various possible TIM-catalyzed reactions were studied by the following procedure. First, the antisymmetric stretching mode, involving the proton being transferred and the two end atoms (without mass-weighting), was used as the approximate reaction coordinate; it has the form $s = r_{D-H} - r_{A-H}$, where D and A denote the donor and acceptor atoms, respectively, and H is the proton being transferred. The adiabatic potential energy along this approximate reaction coordinate was obtained by a series of geometry optimizations of the entire structure in the presence of the harmonic constraint:

$$V^{RESD} = {}^1/_2 k^{RESD}(s - s_i^{REF})^2 \quad (2)$$

with the RESD module (REStrained Distance) in CHARMM. In eq 2, k^{RESD} is the constraining force constant, which was set equal to 2000 kcal/(mol·Å^2). The quantity s_i^{REF} is the reference value of the approximate reaction coordinate, which is close to zero at the saddle point; i.e., the proton is nearly equidistant from D and A. A step size of 0.1 was used during the reaction path scan, and for each value of s_i^{REF}, 1500 steps of adapted basis Newton−Rhapson steps were done at the AM1/CHARMM or AM1-SRP/CHARMM level, and 200−300 steps were done at the B3LYP/CHARMM level. The true saddle point for each reaction was then determined with the conjugate peak refinement algorithm[60] by use of the TRAVEL module in CHARMM with the RESD path determined from eq 2 as input. From calculations at the AM1/CHARMM and AM1-SRP/CHARMM levels, we found that the points optimized with s_i^{REF} are very close to the true saddle points obtained with the TRAVEL module, which requires more computation time. For example, the intramolecular proton-transfer barrier height (path B) obtained with TRAVEL at the AM1-SRP/CHARMM level is 14.2 kcal/mol at $s = -0.002$ ($r_{D-H} = 1.237$ Å, $r_{A-H} = 1.239$ Å), differing by only 0.2 kcal/mol from the value obtained with RESD at $s_i^{REF} = 0.0$. Therefore, only approximate saddle points were optimized at the B3LYP/CHARMM level with RESD, starting with the corresponding structure and s_i^{REF} values obtained from the AM1-SRP/CHARMM calculations.

As described in previous work,[33a] a perturbation analysis was carried out to elucidate the role of individual residues in variety steps of the TIM reactions.[15,33a] The electrostatic terms make the dominant contribution to the relative stability of the stationary points; van der Waals terms are small in the minimized structures. Consequently, the perturbation analysis is restricted to the former. Due to the large number of QM/MM calculations that are required, the analysis was carried out at the AM1-SRP/CHARMM level, which was found to give semiquantitative results as compared to B3LYP/CHARMM calculations. Although the results were already presented in ref 33a, we included the similar data for discussion.

III. Results

Having outlined the methods (section II) and established that they give meaningful results for the species involved in the reactions catalyzed by TIM (see Appendix 1), we present the results at the B3LYP/6-31+G(d,p) level with QM alone for the simple models and QM/MM for the enzyme. We note that AM1-SRP Mulliken population analysis suggests that the transferring protons at the transition states typically have a charge of +0.4; i.e., approximately +0.2 to +0.3 larger than those in the stable structures. Therefore, it is clear that all the H transfer reactions are indeed proton transfers, not hydrogen atom transfers followed by electron transfers from the acceptor to the donor atoms.

Each chemical step of the TIM-catalyzed reaction is treated separately. For each step, we first describe the "basic" model, which includes only the groups required for the reaction; in most steps, this is the substrate and Glu 165. Basic model calculations were performed both in the gas phase and in solution with the solvent treated by a continuum model; the geometry optimization was carried out in the gas phase. To investigate the effect of the active-site residues that experiment and earlier theoretical

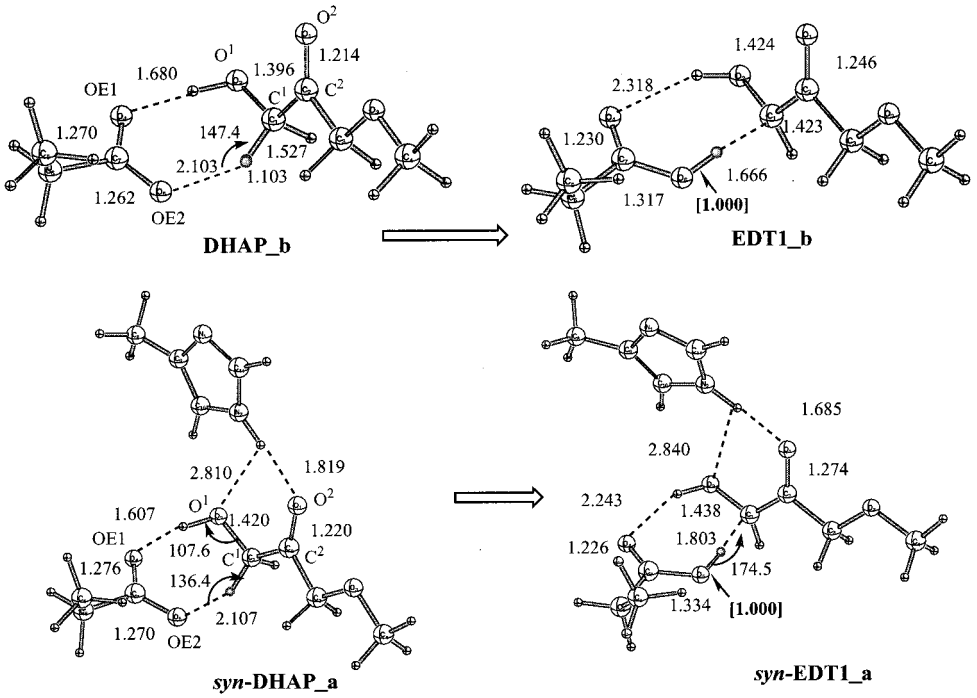


Figure 3. Optimized geometries for the first proton-transfer step (DHAP → EDT1) in the TIM reactions with the basic model (B3LYP/6-31+G-(d,p)), labeled b, and the active-site model (HF/3-21+G), labeled a (see Figure 2 and Table 5). The filled circles indicate the transferred protons. The boldface number in brackets indicates that the bond was kept frozen during optimization (see text). The orientation of the "Glu 165" is such that the syn orbital of the carboxylic group is involved in the proton transfer. Distances are given in angstroms.

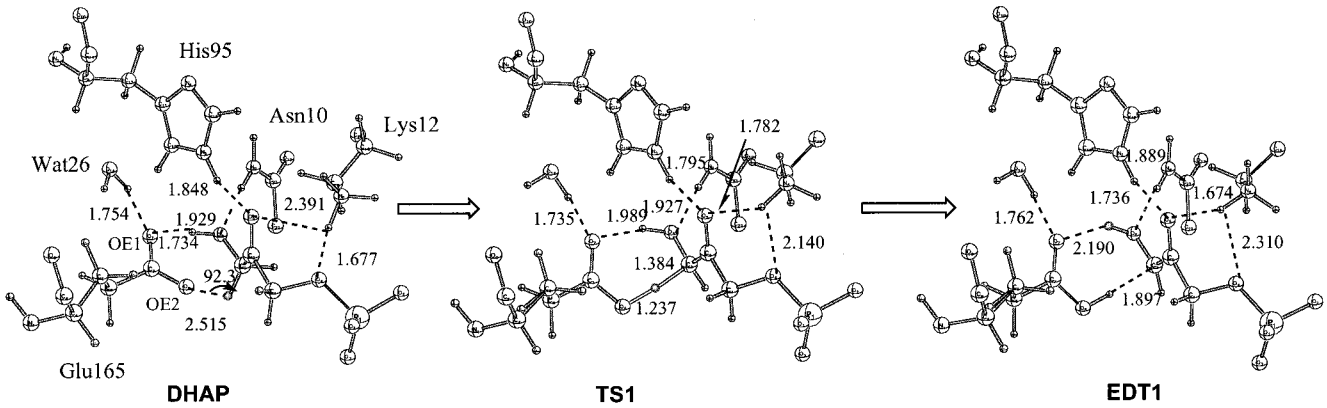


Figure 4. Active-site geometries at the B3LYP/6-31+G(d,p)/CHARMM level for important configurations in the first proton-transfer step in the TIM reactions (DHAP → EDT1). The substrate and the side chain of Glu 165 are treated with QM in these calculations (see text). Distances are given in angstroms.

work indicate to be most important in catalyzing the reactions, results obtained with an "active-site" model are described. To determine the contributions of the rest of the enzyme environment, the stochastic boundary simulation results are reported. By comparing the various results we are able to determine the important contributions to the catalysis, which are discussed in section IV. We also examine the effect of the orbital orientation of Glu 165 (syn vs anti) and the mutation from His 95 to Gln 95. Since zero-point energy (ZPE) was not included in most gas-phase models, energetics without ZPE are used in the analysis. For the overall reaction results in the enzyme and the effect of ZPE and vibrational contributions, see ref 33a.

III.1. Common Proton-Transfer Steps: DHAP → EDT1 and EDT2 → GAP. We first describe the results obtained with the various models for the first and last chemical steps of the TIM-catalyzed reaction (see Scheme 1). The mechanisms for these steps are supported by all the available experimental and theoretical evidence, so there is general agreement that they are correct.

III.1.a. DHAP → EDT1. Figure 3 shows the structures involved in the proton transfer to Glu 165 from C^1 of the substrate obtained with the various simplified models: the basic model (only $CH_3CH_2CO_2^-$ for Glu 165 from the protein is included) and the active-site model ($CH_3CH_2CO_2^-$ for Glu 165 and 4-methylimidazole for His 95 are included). The corresponding structures in the full enzyme model are in Figure 4. In the absence of the full enzyme model, an OCH_3 terminal group replaces the phosphate (see Figures 2 and 3), as described in the Methods section, and the phosphate group is included in the full active-site calculation. The energies obtained with the various models are summarized in Table 2 and Figure 7a. The proton transfer from the model substrate to $CH_3CH_2CO_2^-$ is an energetically unfavorable process and is endothermic by 32.9 kcal/mol in solution when the two groups are at infinite separation. When the two groups are placed in positions similar to that in TIM (i.e., in the basic model), the process becomes less endothermic but still very unfavorable energetically: it requires 24.4 kcal/mol in the gas phase and 23.4 kcal/mol in

TABLE 2: Energetics for the Common Proton Transfer Steps in the TIM-Catalyzed Reactions with Different Models[a]

	basic model			active-site model						
	isolated[b]	complex[b]		*syn*-OCH_3		*syn*-PO_3H_2	*syn*-OCH_3 + Lys		*syn* + Lys + NMA	
reaction	PB	vac	PB	vac	PB	vac	vac	PB	vac	enzyme
					DHAP → EDT1					
DHAP	*−650.90485*	*−650.79608*	*−650.90219*	*−916.37263*	*−916.47350*	*−1444.80264*	*−973.39894*	*−973.51477*	*−1470.50417*	*−1180.59902*
TS1	nc	→ EDT1	nc	→ EDT1	nc	→ EDT1	nc	nc	nc	11.8 (11.3)
EDT1	32.9	24.4	23.4	17.6	18.3	14.9	0.1	16.1	−1.1	7.4 (6.2)
					EDT2 → GAP*c*					
EDT2_np	*−650.86911*	*−650.77001*	*−650.87113*	*−905.89865*	*−916.43929*	nc	nc	nc	nc	*−1180.58643*
TS2	nc	0.5	2.6	→ EDT2	nc	nc	nc	nc	nc	1.2 (2.9)
GAP	−18.1	−11.5	−9.8	−10.1	−10.2	nc	nc	nc	nc	−4.1 (−2.0)

[a] The total energies for reference structures (DHAP, EDT2), in hartrees, are given in italic type; relative energies are given in Roman text, in kilocalories per mole. The energetics were obtained with B3LYP/6-31+G(d,p) for the basic and active-site models and B3LYP/6-31+G(d,p)/CHARMM for the enzyme. The geometries were optimized in a vacuum with B3LYP/6-31+G(d,p) for the basic model and HF/3-21+G* for the syn structures in the active-site model. No geometry optimizations were performed for syn + Lys or syn + Lys + NMA (see text). In the Poisson−Boltzmann (PB) calculations, the Pauling set of atomic radii was used. For the enzyme, geometries were optimized with the small QM partition including the substrate and Glu 165. Numbers in parentheses are single-point energetics with a larger QM partition that also includes His 95. The notation → X means that the transition state does not exist (or was very difficult to locate due to the fact that the potential energy surface is very flat) and all searches collapsed to structure X; nc indicates that the item was not calculated. [b] The "isolated" values were obtained with isolated Glu 165 and substrate model. [c] The EDT2_np structures are used as reference for the basic and active-site models (see text).

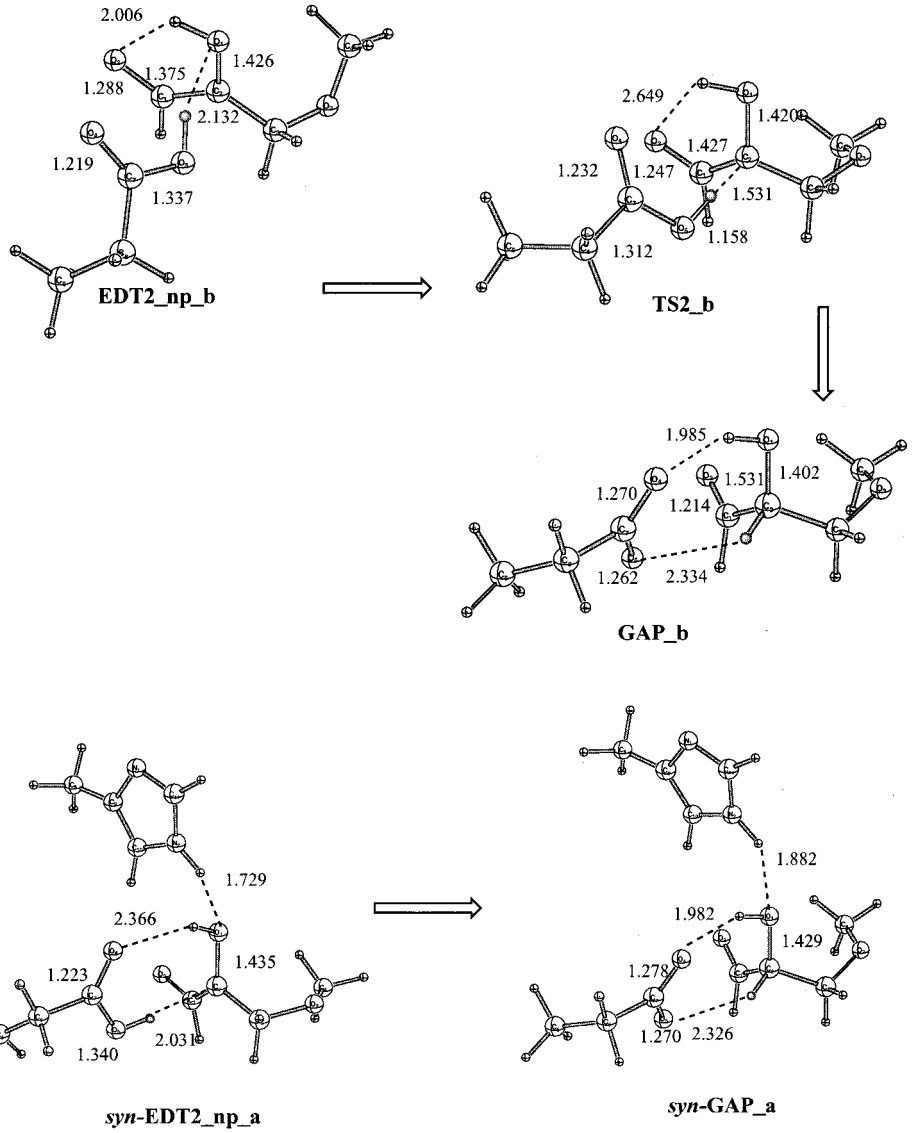


Figure 5. Optimized geometries for the final proton-transfer step (EDT2 → GAP) in the TIM reactions with the basic model (B3LYP/6-31+G-(d,p)), labeled b, and the active-site model (HF/3-21+G), labeled a (see Figure 2 and Table 5). The filled circles indicate the transferred protons. Distances are given in angstroms.

solution;(the similarity of these two values is due to the use of an OCH_3 group in place of the phosphate). No saddle point was found for this proton transfer reaction; in fact, geometry optimization led from EDT1 directly to DHAP if no constraint

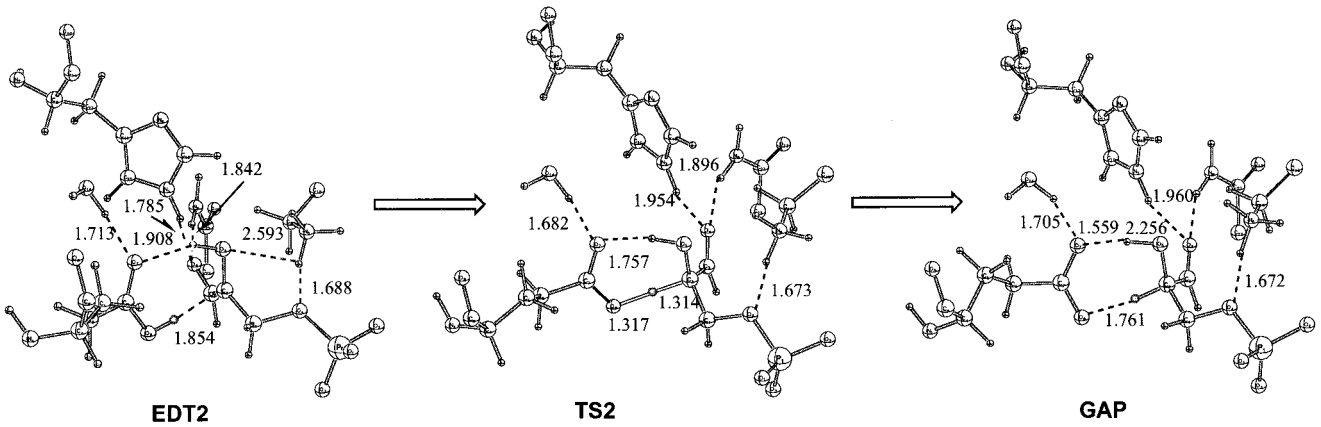


Figure 6. Active-site geometries at the B3LYP/6-31+G(d,p)/CHARMM level for important configurations in the final proton-transfer step in the TIM reactions (EDT2 → GAP). The substrate and the side chain of Glu 165 are treated with QM in these calculations. Distances are given in angstroms.

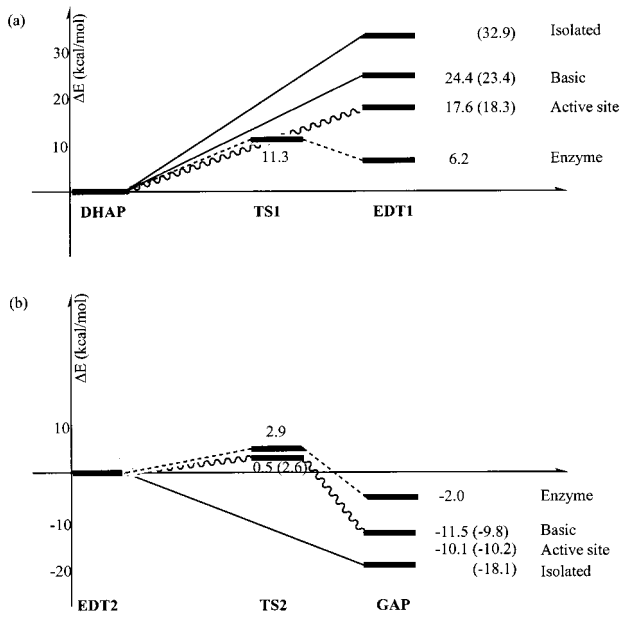


Figure 7. Comparison of energetics for the reactions (a) DHAP → EDT1 and (b) EDT2 → GAP in the gas phase, in solution (in parentheses), and in TIM; see text (Table 5) and Figure 2 for a description of the various models.

was used for the OE2−H bond. The process is less endothermic in the presence of His 95 in the active-site model, 17.6 kcal/mol in the gas phase and 18.3 kcal/mol in solution (see Figure 7a); thus, the electrostatic interaction due to His 95 is equal to 6.8 and 5.1 kcal/mol in the gas phase and in solution, respectively. By putting a NH_4^+ at the position of Lys 12 in the TIM−PGH X-ray structure,[7a] EDT1 is stabilized in the gas phase by 17.5 kcal/mol relative to DHAP, making the first proton transfer nearly thermoneutral; this result is due to the fact that the negative charge formed on O^2 in EDT1 interacts strongly with the Lys model (see Figure 2). However, this large electrostatic effect is nearly completely quenched, from 17.5 kcal/mol in the gas phase to 1.5 kcal/mol in solution, when the solvation is included with a continuum model. The effect of the hydrogen-bonding network modeled by two NMA molecules (see Figure 2) was found to be 1.2 kcal/mol on the first proton transfer, stabilizing EDT1 relative to DHAP.

In the enzyme (Figure 4), Glu 165 forms a rather strong hydrogen bond with the OH group of the substrate DHAP; the proton−acceptor distance is 1.734 Å, slightly longer than that (1.680 Å) in the basic model. This is significantly longer than the value of 1.57 ± 0.05 Å for the hydrogen bond distance between the OH of PGH and Glu 165, deduced in an NMR study[10] from an empirical relationship between the proton chemical shift and hydrogen-bond distance;[61] the X-ray structure of the TIM−PGH complex has a donor−acceptor distance of 3.08 Å (corresponds to a hydrogen−acceptor distance of 2.08 Å) and minimization of this structure yields a value somewhat closer to the NMR estimate. Thus, it is not clear whether the NMR results for the PGH complex is applicable to the DHAP substrate. The two carboxylate oxygen atoms in Glu 165 are 2.689 and 2.784 Å away, respectively, from the C^1 of in the substrate; these are close to the X-ray values of 2.505 and 2.665 Å, respectively. In contrast to the suggestion[3,7] that the value of 2.665 Å implies a short distance between OE2 in Glu 165 and the substrate Cα proton, the minimized QM/MM structure gives a rather long OE2−proton distance of 2.515 Å because the OE2···H···Cα angle is small (92.3°); i.e., the carboxyl group in Glu 165 is oriented such that the hydrogen bond to the more acidic OH group in DHAP is maximized. The Lys 12 interacts strongly with the phosphate group, with an $-NH_3^+$ proton−ester oxygen distance of 1.677 Å; the $-NH_3^+$ proton is quite far, 2.391 Å, from the carbonyl oxygen in the substrate, which is substantially less charged than the ester oxygen. In EDT1, the hydrogen bond between the carboxyl oxygen of Glu 165 and the hydroxyl in the substrate is elongated to 2.190 Å but shorter (2.318 Å) than in the basic model. By contrast, the distance between OE2 and the C^1 in the substrate is 2.897 Å, longer than that (2.721 Å) in the basic model. These differences presumably are mainly due to Asn 10, which polarizes the substrate such that the O^1 atom becomes more negatively charged in the enzyme. The side chain of Lys 12 is seen to move about 0.72 Å, toward the carbonyl oxygen of the substrate upon the proton abstraction; the distances between the $-NH_3^+$ proton and the carbonyl oxygen are 2.391 and 1.674 Å before and after the proton has moved from DHAP to Glu 165. The proton-transfer saddle point is rather "late" judging from the OE2···H distance of 1.237 Å. However, it is much earlier than in the gas-phase models, since there the saddle points simply collapsed onto the products due to the large endothermicities. Lys 12 has already moved toward the carbonyl oxygen of the substrate by 0.61 Å at the transition state.

Comparison of the various energetic results in Table 2 (see also Figure 7a) shows that the changes in the reaction endothermicity and barrier height are coupled as expected; the reverse barrier (i.e., measured from EDT1) becomes nonvanishing in the enzyme due to the stabilization of EDT1. The proton transfer

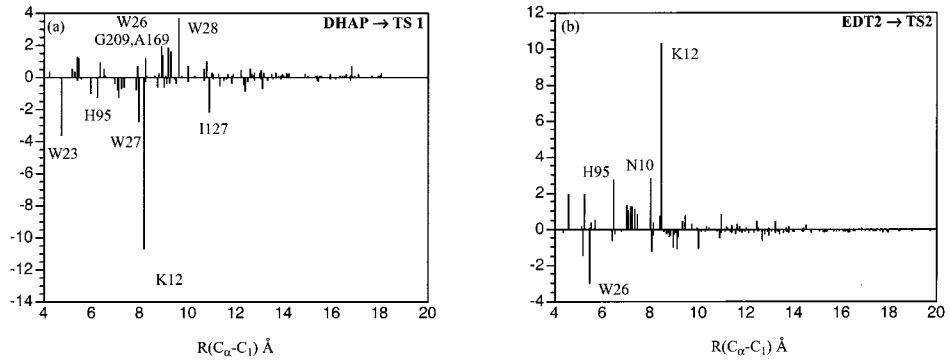


Figure 8. Perturbation analysis for energetic contributions from residues and X-ray observed water molecules in TIM at the AM1-SRP/CHARMM level. Positive values indicate unfavorable contributions. Shown are the results for the two common proton-transfer steps.

is less endothermic when the two reacting groups are placed together in the basic model relative to when they are infinitely separated (Figure 7a), suggesting that the spatial arrangement in TIM favors the interaction between EDT1 and "GluH 165" over DHAP and "Glu 165". The difference between the basic model and the active-site model indicates that His 95 has a substantial electrostatic effect and stabilizes the enediolate species. The perturbation analysis, as shown in Figure 8a, demonstrates that it is mainly the favorable electrostatic interaction between EDT1 and Lys 12 that puts EDT1 just 6.2 kcal/mol above DHAP, compared to 24.4 kcal/mol in the basic model and 17.6 kcal/mol in the active-site model without NH_4^+. The effect on the first proton transfer barrier is 10.7 kcal/mol in the enzyme; the value is intermediate between that in the gas phase (17.5 kcal/mol) and that in solution (1.5 kcal/mol). Perturbation analysis also suggests that Asn 10 lowers the barrier by 1.3 kcal/mol. In addition, other residues also contribute substantially to lowering the barrier. The list includes Ser 211, Ile 127, and two active-site water molecules, Wat23 and Wat27; the latter two are bound to the phosphate group of the substrate and the main-chain carbonyl of Glu 165, respectively. The Wat27 is also hydrogen-bonded with Wat26, which interacts strongly with the side chain of Glu 165 throughout the reaction. There are also a few residues that contribute unfavorably to the proton transfer, i.e., the active-site waters Wat26 and Wat28, Gly 209, and Ala 169 (through main-chain atoms), which are close to the substrate. As discussed in the previous work,[33a] the contribution of the three active-site water molecules (Wat26, Wat27, and Wat28) is qualitatively consistent with the recent experimental findings of Petsko and co-workers.[13]

III.1.b. EDT2 → GAP. The structures involved in the conversion from EDT2 to GAP with the basic and active-site models are shown in Figure 5. The energetics are summarized in Table 2 and Figure 7b. The process is exothermic because an acidic proton is transferred from GluH165 to the carbon atom in EDT2; the exothermicity is 18.1 kcal/mol in solution when the two fragments are at infinite separation. Although this proton-transfer step is very similar in character to the first proton-transfer just discussed (DHAP → EDT1), the energy difference between EDT2 and GAP (18.1 kcal/mol in the gas phase) is much smaller than that between EDT1 and DHAP (32.9 kcal/mol). In the basic model, the exothermicity is calculated to be 11.5 and 9.8 kcal/mol, respectively, in the gas phase and in solution. The saddle point is very early, as seen from Figure 5, and the barrier height is very small (0.5 kcal/mol) in the gas phase and slightly higher (2.6 kcal/mol) in solution. The influence of His 95 on the step from EDT2 to GAP was found to be rather small (i.e., the difference between the basic and active-site models); the exothermicity is 10.1 and 10.2 kcal/mol in the gas phase and in solution, respectively. The proton transfer was so facile that no transition state was located between EDT2 and GAP in the active-site model.

In the enzyme, the last proton-transfer barrier was found to be only 2.9 kcal/mol, similar to that found in the basic model. The small barrier height seems reasonable as judged from the structure of EDT2 in Figure 6. Clearly, the proton on OE2 in Glu 165 is well positioned for transfer. The GAP structure is only 2.0 kcal/mol lower in energy than EDT2, compared to the value of 11.5 kcal/mol in the basic model. The decrease in the stability of GAP relative to EDT2 in the enzyme is presumably due to the weakened hydrogen bonding between the substrate and nearby residues such as His 95 and Asn 10. This is reflected in the increases in these hydrogen-bond distances in GAP relative to those in EDT2; the values are 0.43 and 0.12 Å for His 95 and Asn 10, respectively. Perturbation analysis (Figure 8b) verified that His 95 and Asn 10 favor EDT2 over the proton-transfer transition state by 2.8 and 2.7 kcal/mol, respectively. Other residues such as Lys 12, plus active-site water molecules (e.g., Wat26), also favor EDT2 substantially. Overall, the enzyme stabilizes EDT2 over GAP relative to the basic model by ~14 kcal/mol, a value similar to that found for the stabilization of EDT1 relative to DHAP.

III.2. Alternative Paths between EDT1 and EDT2. The three alternative pathways between EDT1 and EDT2 shown in Scheme 1 are now described. The calculated energetics are summarized in Table 3. Comparisons of different models in the gas phase, in solution, and in the enzyme are shown in Figure 18.

III.2.a. Path A with His 95 Acting as a General Acid. In Figure 9, structures involved along path A with various models are shown. In this path, His 95 first delivers its proton to the O^2 of enediolate EDT1 to yield the enediol EDL1; then it abstracts the O^1 proton from the substrate, forming another enediolate species, EDT2. When the reacting groups are infinitely apart in solution, the enediol species is 5.4 kcal/mol lower in energy relative to EDT1, while EDT2 is 2.0 kcal/mol higher in energy. With the basic model, the first proton transfer becomes nearly thermoneutral, with EDL1 0.2 kcal/mol lower than EDT1 in the gas phase but 1.9 kcal/mol higher in solution. Therefore, the interaction between the enediol and deprotonated His 95 is not as favorable as that between the enediolate (EDT1) and neutral His 95, although strong hydrogen bonds involving the donor−acceptor atoms occur in both EDT1 and EDL1 with proton−acceptor distances of 1.570 and 1.534 Å, respectively. The EDL1 is destabilized by solvation relative to EDT1 because there is one more hydrogen bond between the substrate and "His 95" in the former, which becomes weaker upon solvation. The barrier height for this step is very small; it is 1.9 kcal/mol in the gas phase and 3.1 kcal/mol in solution. The second proton transfer is endothermic by 11.0 kcal/mol in the gas phase and

TABLE 3: Energetics for the Different Paths in the TIM-Catalyzed Reactions with Different Models[a]

	basic model			active-site model			
	isolated[b]	complex[b]		*syn*-OCH_3		*syn*-OP_3H_2	
reaction	PB	vac	PB	vac	PB	vac	enzyme
				Path A			
EDT1_A	*−647.99148*	*−647.90798*	*−648.00589*	*−916.34458*	*−916.44434*	*−1444.77890*	*−1177.72082*
TSA	nc	1.9	3.1	→ EDL1	nc	→ EDL1	6.7 (8.4)
EDL1	−5.4	−0.2	1.9	7.1	7.7	7.6	→ EDT2
TSA2	nc	→ EDT2	nc	→ EDL1	nc	→ EDL1	nc
EDT2_A	2.0	10.8	5.3	7.7	5.3	9.5	−4.1 (−2.6)
				Path B			
EDT1	*−650.85250*	*−650.75720*	*−650.86490*	*−916.34458*	*−916.44434*	*−1444.77890*	*−1180.58723*
TSB	8.8	4.8 [13.0][c]	7.5 [10.7][c]	13.9	13.7	13.9	14.0 (14.0)
EDT2_B	2.0	−5.2 [10.8][c]	−2.7 [5.3][c]	7.7	5.3	9.5	0.5 (−1.7)
				Path C			
EDT1	*−650.85250*	*−650.75720*	*−650.86490*	*−916.34458*	*−916.44434*	*−1444.77890*	*−1180.58723*
EDT1_2	0.0	−3.8	0.9	−2.4	2.5	1.0	nc
TSC1	nc	0.6	7.3	8.5	11.2	7.9	6.6 (6.4)
EDL2	−20.5	−13.3	−4.2	1.0	3.7	4.6	−4.4 (−4.1)
TSC2	nc	→ EDT2_C	nc	→ EDT2_C	nc	→ EDT2_C	6.3 (6.0)
EDT2_C	2.0	−5.2	−2.7	7.7	5.3	9.5	0.5 (−1.7)

[a] The EDT1 structure is used as the energy reference because different paths diverge from there. See footnote *a* in Table 2 for the format. [b] The isolated values were obtained with isolated Glu 165 and substrate model. [c] The numbers without brackets were calculated with the basic model consists of the substrate and model Glu 165 (first row in Figure 9); those in brackets were values for the basic model consisting of the substrate and model His 95 (second row in Figure 9).

3.4 kcal/mol in solution. EDT2 is high in energy in the gas phase because the negatively charged O^1 in the substrate is not hydrogen-bonded due to the geometrical constraints applied on the "His 95" in our model. The effect is partly compensated by the solvent in solution; the energy difference between EDT2 and EDL1 is decreased by 7.6 kcal/mol upon solvation. The saddle point for the second proton transfer step was located at the AM1 level but does not exist at the B3LYP/6-31+G(d,p) level. Single-point B3LYP/6-31+G(d,p) energy calculations at the AM1 geometries confirm that the AM1 saddle point actually has a lower energy than EDT2.

With the active-site model, the EDL1 is less stable than EDT1 by 7.1 and 7.7 kcal/mol in the gas phase and in solution, respectively. These are considerably higher than the corresponding values for the basic model, which are −0.2 and 1.9 kcal/mol in the gas phase and in solution, respectively. It is not immediately clear why the additional "Glu 165" in the active-site model destabilizes the enediol form relative to the enediolate species. Searches for the saddle points between EDL1 and EDT1/EDT2 collapsed to local minima in all cases. As discussed below, this is due to the fact the potential energy curve is very flat around EDL1 in the active-site model. EDL1 is similar in energy to EDT2, while EDT2 is less stable than EDL1 in the basic model, by 11.0 kcal/mol in the gas phase and 3.4 kcal/mol in solution. This results from the fact that the GluH 165 in the active-site model can stabilize the negatively charged O^1 atom in EDT2.

The structures involved in path A in the enzyme are shown in Figure 10. From Figures 4 and 10, it is evident that the important geometrical parameters for the active site of EDT1_A are similar to those of EDT1. Given that different fragments have been treated with QM during the optimizations of the two structures (substrate + His 95 in EDT1_A, substrate + Glu 165 in EDT1), the similarity of the structures indicates that the approach and the QM/MM parameters used here are satisfactory.

There are some differences in the details of path A between the present B3LYP/6-31+G(d,p)/CHARMM calculations and the AM1/CHARMM study of Bash et al.;[15] AM1/CHARMM calculations with the present enzymatic model found results similar to those in ref 15. In the previous study,[15] starting with EDT1_A, a proton transfer from His 95 to the carbonyl oxygen led to the enediol species, EDL1_A, which is calculated to be 11 kcal/mol above EDT1_A; the barrier for this step is 16 kcal/mol.[15] Next, His 95 abstracted a proton from a different oxygen atom in the substrate, leading to EDT2_A, which is 3 kcal/mol lower than EDT1_A; the barrier for this step is rather small, only 5 kcal/mol. At the B3LYP/6-31+G(d,p)/CHARMM level, however, the EDL1_A species was not found to be a local minimum. Optimizations starting from the AM1/CHARMM EDL1_A structure leads directly to EDT2_A. An intermediate structure obtained during the optimization is shown in Figure 10 to illustrate the behavior.

To further analyze the proton-transfer mechanism along path A, a two-dimensional adiabatic map was calculated at the AM1-SRP/CHARMM level. The two coordinates used to describe the reaction are the antisymmetric stretch involving the transferring protons (H_1 in the substrate and Hϵ2 in His 95) and associated donor and acceptor atoms (the two substrate oxygen atoms involved and the Nϵ in His 95); all degrees of freedoms other than the two mapping coordinates were optimized in the two-dimensional grid scan used to generate the adiabatic map. As shown in Figure 11, there is only one saddle point in the conversion between EDT1 and EDT2 along path A. The character of this saddle point involves mainly the motion of H_1, i.e., the first proton to be transferred in path A. The motion of the second proton, Hϵ2 in His 95, becomes significantly involved immediately after the saddle point (i.e. $\delta_1 > 0.2$ Å). As expected, the change from the neutral His 95 to a negative His 95^- is not favored energetically, so that the second proton transfer is nearly concerted with the first. The barrier height for the first proton transfer is calculated to be 8.4 kcal/mol. The product EDT2_A is 2.6 kcal/mol more stable than EDT1_A. Compared to the situation in the active-site model, in which EDT2 is 7.7 kcal/mol higher than EDT1, it is seen that the enzyme has managed to considerably stabilize the EDT2 species. This is due to the presence of Asn 10 in the active site, which interacts favorably with the negatively charged O^1 atom in EDT2 (see below). In the active-site model, the O^1 is exposed because the "His 95" has to interact with the hydroxyl O^2 due to limited

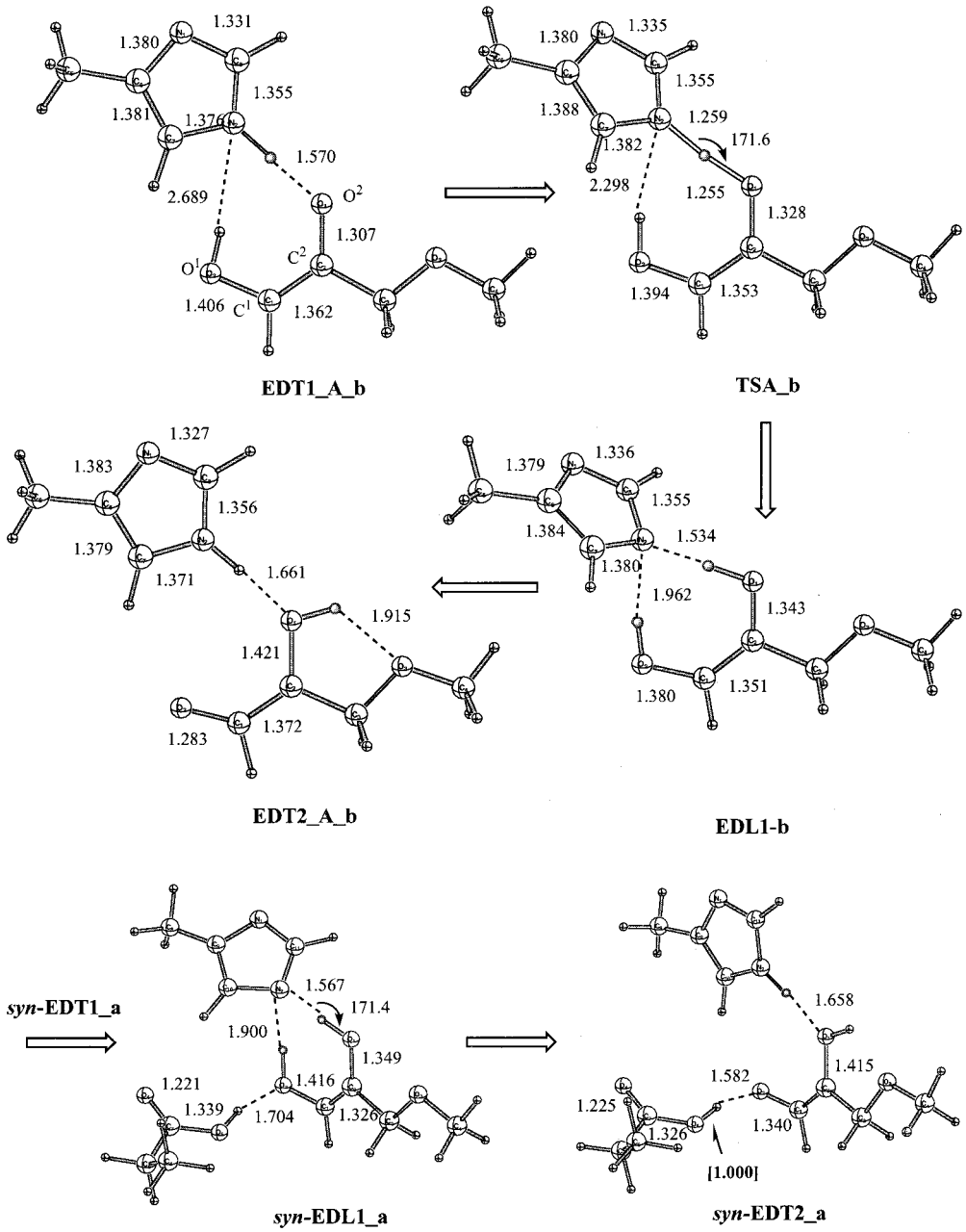


Figure 9. Optimized geometries for structures on path A with the basic model [B3LYP/6-31+G(d,p)], labeled b, and active-site model (HF/3-21+G), labeled a (see Figure 2 and Table 5). The filled circles indicate the transferred protons. The boldface number in brackets indicates that the bond was kept frozen during optimization (see text). Distances are given in angstroms.

flexibility as a result of constraints employed in the geometry optimization.

Perturbation analysis (Figure 17a) shows that one residue from the other subunit in the dimeric TIM, Thr 75B, contributes favorably to the barrier height (at TSA) by about 2 kcal/mol. A similar effect was observed in the simulation of Åqvist et al.[14] They related this to the mutation results of Borchert et al.,[29] who found that shortening the loop around Thr 75B yielded a stable monomeric species that has a 1000-fold lower value of k_{cat}. Lys 12 and Glu 165 are the other two important residues that do not participate the proton transfer explicitly. They contribute unfavorably to the barrier height for the first proton transfer in path A but favorably to the stability of EDT2_A relative to EDT1_A (Figure 17); Asn 10 also stabilizes EDT2_A by about 1.3 kcal/mol. The three residues (Thr 75B, Lys 12, and Glu 165) together yield a value of 5.6 kcal/mol, relative to the total barrier of 8.0 kcal/mol; other significant residues are active-site water molecules Wat27 and Wat28, which contribute to raise the barrier, and His 95 (main chain), Glu 97, and Arg 98, which lower the barrier. Thus, the barrier for this step is actually increased in the enzyme (also see Figure 18a), which is consistent with the fact that EDL1 is not stabilized; as suggested in ref 15 and confirmed in ref 9, a neutral His 95 is involved in the proton transfer to avoid overstabilization of the enediol, EDL1.

III.2.b. Path B Involving an Internal Proton Transfer. The structures involved in the intramolecular proton-transfer step (path B) with various models are shown in Figure 12. The intrinsic energetics for this step can be obtained from the model substrate with the phosphate group replaced by $-CH_2OCH_3$, as discussed in Appendix 2. In this model, EDT2 is higher than EDT1 by only 2 kcal/mol, and the barrier for the intramolecular proton transfer is about 9 kcal/mol. With the basic model that includes interaction with "Glu 165", EDT2 is lower in energy

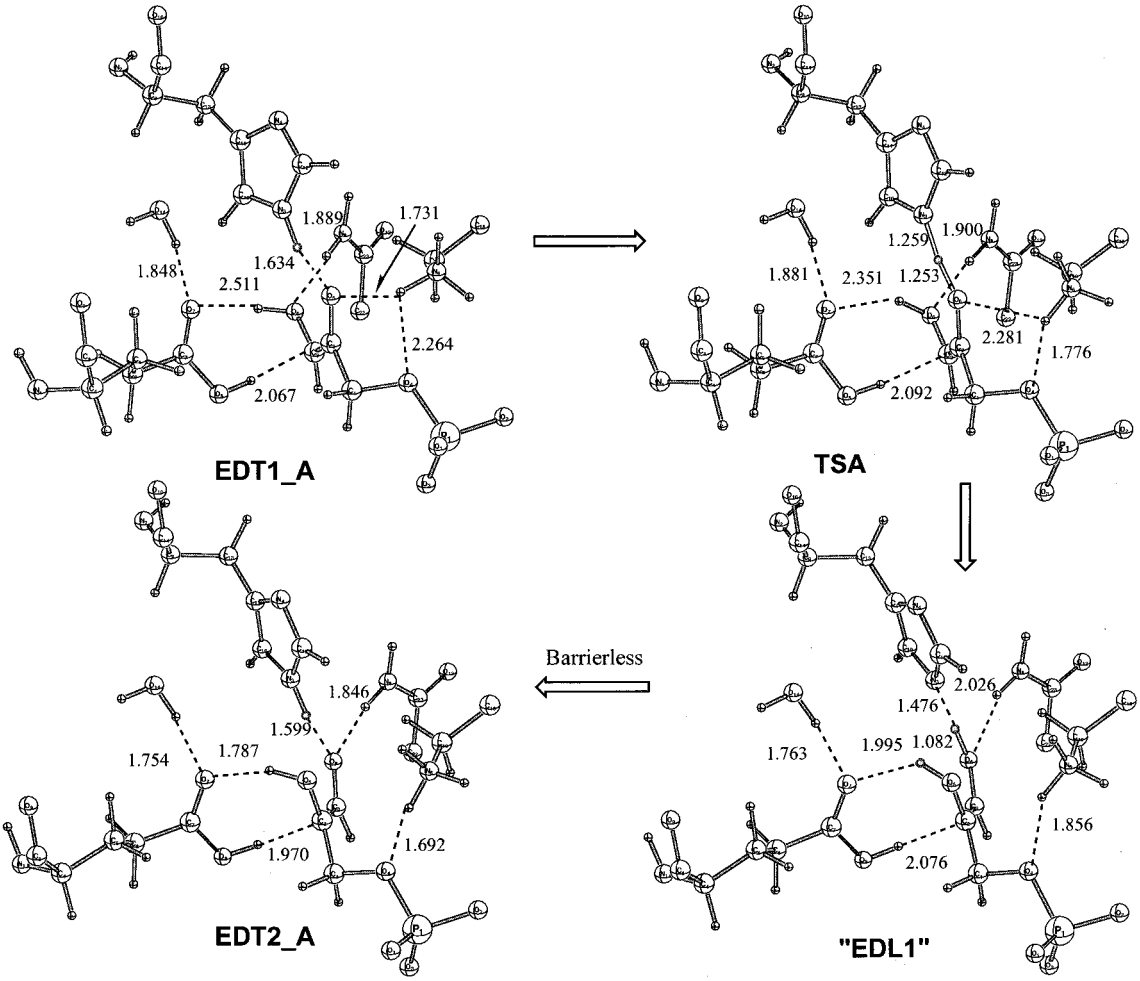


Figure 10. Active-site geometries at the B3LYP/6-31+G(d,p)/CHARMM level for important configurations along path A in the TIM reactions. The substrate and the side chain of His 95 are treated with QM in these calculations. Note that "EDL1" was not found to be a stationary point on the potential energy surface in the current work (see section III.2a in the text). Distances are given in angstroms.

than EDT1 by 5.2 kcal/mol in the gas phase and 2.7 kcal/mol in solution. This results from the fact that a strong hydrogen bond (with an O···H distance of 1.550 Å) is formed in EDT2 between the acidic proton of "GluH 165" and the negatively charged O_1 of the substrate. The barrier height for the intramolecular proton transfer is also lowered by the same effect, judging from the existence of this hydrogen bond in TSB in Figure 12. The barrier height is 4.8 kcal/mol in the gas phase and 7.5 kcal/mol in solution. As expected, screening from the solvent decreases the magnitude of the hydrogen bonding effect, which is largely electrostatic in nature.[62] Figure 12 also shows the transition-state structure (TSB′) for the intramolecular proton transfer with the basic model involving His 95. The barrier height for this step is 13.0 and 10.7 kcal/mol in the gas phase and in solution, respectively, which are higher by 3.9 and 1.9 kcal/mol than the corresponding values in the simple model system without "His 95" in the gas phase and in solution, respectively. This difference is due to the unfavorable interaction between the transferred proton and "His 95". With the active-site model that includes both Glu 165 and His 95, the barrier height is rather similar, 13.9 and 12.7 kcal/mol in the gas phase and in solution, respectively.

The active-site structures involved along path B in the enzyme are shown in Figure 13. In EDT2, the hydroxyl group in the substrate and OE1 in Glu 165 form a hydrogen bond, and the OE2·H group also interacts favorably with the negatively charged C^1 atom in the substrate. Both His 95 and Asn 10 move closer to the O^1 atom in the substrate to stabilize its negative charge. Lys 12 moves further from the substrate O^2 atom, which is neutralized by the transferred proton. The EDT2 is calculated to be 1.7 kcal/mol lower in energy than EDT1, slightly less stable compared to the corresponding value of −5.2 kcal/mol in the basic model. EDT2 in the enzyme is more stable than the corresponding structure in the active-site model, in which EDT2 is 7.7 kcal/mol higher in energy than EDT1. As discussed above, this is due to the greater flexibility of His 95 in enzyme, which allows it to move to form a hydrogen bond with O^1 in the substrate and therefore stabilize the EDT2 species. The saddle point for the intramolecular proton transfer has similar critical geometrical parameters to those in the gas-phase model, with a barrier height of 14.0 kcal/mol, which is also very similar to that in the gas-phase models including His 95. In all cases, the existence of His 95 increases the barrier by 4−5 kcal/mol for this process compared to that in the isolated enediolate (see Appendix 2). The perturbation analysis (Figure 17b) clearly demonstrated that His 95 contributes unfavorably to both the barrier heights. The fact that electrostatic perturbation has reproduced the effect indicates that the unfavorable effect from His 95 is not steric in nature. Two active-site water molecules, Wat26 and Wat27, contribute favorably by about 2−3 kcal/mol to the stability of EDT2. Lys 12 and Gly 209 also make contributions on the order of 2 kcal/mol.

III.2.c. Path C with GluH 165 as the Generalized Acid. When the proton donor (GluH 165 in path C and His 95 in

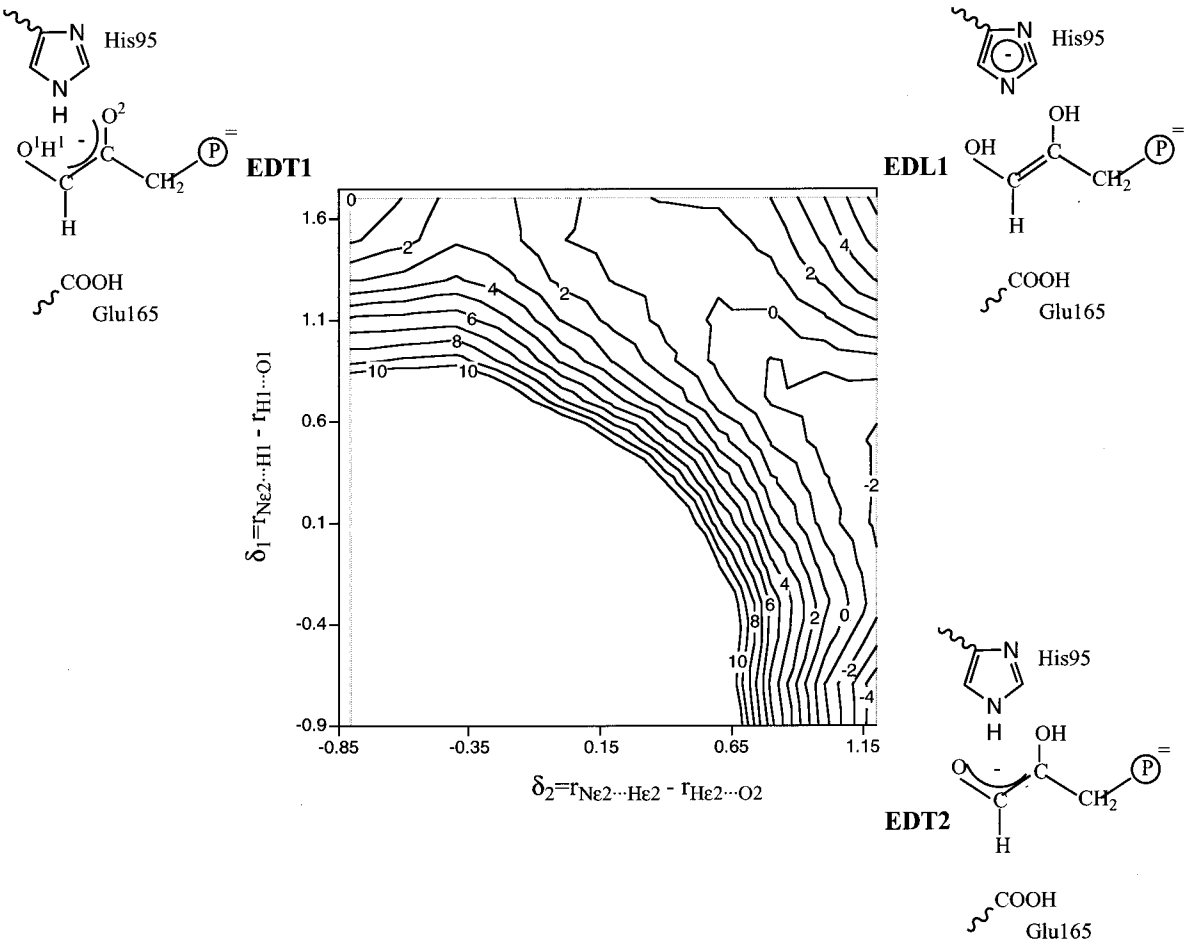


Figure 11. Two-dimensional potential energy surface for path A obtained from adiabatic mapping; the reaction coordinates are defined in the figure (also see text). The energy is in kilocalories per mole, with that for EDT1 set to zero. The calculations were performed at the AM1-SRP/CHARMM level.

path A) and the substrate are infinitely separated in solution, the first proton transfer along path C is energetically more favorable than path A (compare Figure 18 panels a and c) by nearly 20 kcal/mol. This is consistent with the fact that "GluH 165" is much more acidic than "His 95". In Figure 14, we show the basic model structures involved along path C in Scheme 1. The starting structure along this path is EDT1_2, which was obtained by manually altering the orientation of GluH 165; the same set of geometrical constraint as for EDT1 was used. In EDT1_2, the acidic proton of "GluH 165" forms a hydrogen bond with O^2 of the substrate rather than with C^1 of EDT1. In addition, the hydroxyl group of the substrate forms an intramolecular hydrogen bond with O^2 instead of the intermolecular hydrogen bond with the carboxylate group of "GluH 165". These changes in the hydrogen-bonding pattern stabilize EDT1_2 by 3.8 kcal/mol, relative to EDT1 in the gas phase, but increase the energy by 0.9 kcal/mol in solution. As the acidic proton gets transferred to the substrate, negative charge develops on the oxygen atoms of "GluH 165" in EDL2 and the intramolecular substrate hydrogen bond is broken while the intermolecular hydrogen between O^1H and "Glu 165" is re-formed. The process of going from the enediolate EDT1_2 to the EDL2 is exothermic by 9.5 kcal/mol in the gas phase and 5.1 kcal/mol in solution. The barrier height is calculated to be 4.4 and 6.4 kcal/mol in the gas phase and in solution, respectively. The next step in the reaction is proton abstraction from O^1H in the substrate by "Glu 165" to form EDT2. The enediol species is very much favored in the gas phase so that optimization without any constraint on the OE2−H bond led from EDT2 directly back to EDL2. This is consistent with the proposal[82] that the carboxylic group ("GluH 165") is more acidic than an enediol. Although a stable EDT2 might be obtained by minimization in the solution, this was not examined in the present work. Instead, we simply constrained the OE2−H bond during the optimization (see Figure 13). The process from EDL2 to EDT2 was found to be endothermic by 8.1 and 1.5 kcal/mol in the gas phase and in solution, respectively. Due to the uphill nature of the potential energy surface associated with this proton-transfer step in the gas phase, no saddle point was located.

The structures involved along path C with the active-site model are shown in Figure 15. The barrier height measured for TSC1 relative to EDT1_2 is 6.1 kcal/mol; the corresponding value is 4.4 kcal/mol in the basic model. EDL2 is more stable than EDT1_2 by only 1.4 kcal/mol, in contrast to the corresponding value of −9.5 kcal/mol in the basic model. This is due to the fact that the additional "His 95" in the active-site model favors the enediolate species relative to the enediol form, as reflected by the corresponding hydrogen-bond distances of 1.643 and 1.744 Å in EDT1_2 and EDL2, respectively. Similar to the situation in the basic model, the EDT2 structure was not stable without the constraint on the OE2−H bond; unconstrained optimization led to EDL2 due to the greater acidity of the carboxylic group relative to enediol. With the bond constraint (see Figure 15), EDT2 is 6.7 kcal/mol higher in energy than EDL2, which is similar to the value of 8.1 kcal/mol in the basic

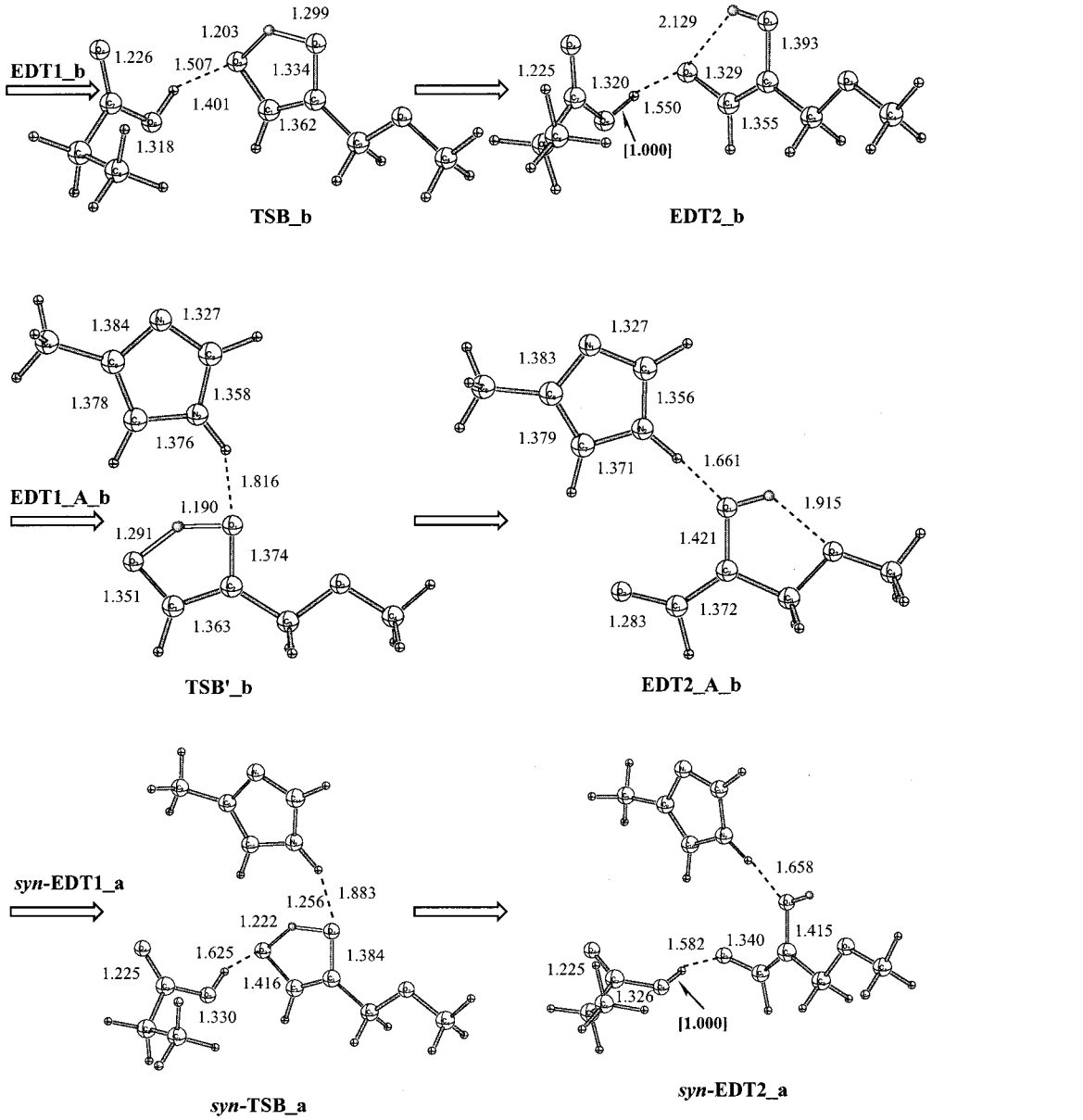


Figure 12. Optimized geometries for structures along path B with the basic model [B3LYP/6-31+G(d,p)], labeled b, and active-site model (HF/3-21+G), labeled a. The filled circles indicate the transferred protons. The boldface number in brackets indicates that the bond was kept frozen during optimization (see text). Distances are given in angstroms.

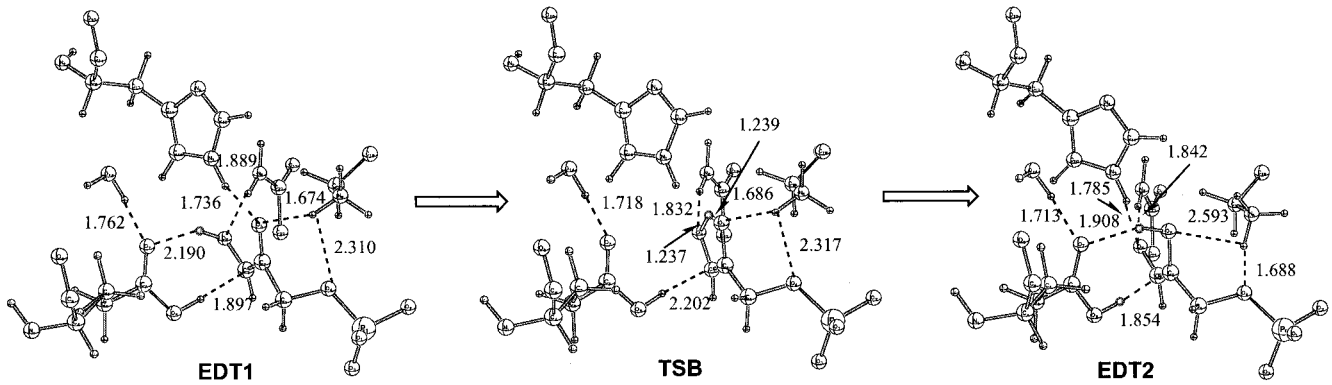


Figure 13. Active-site geometries at the B3LYP/6-31+G(d,p)/CHARMM level for important configurations along path B in the TIM reactions. The substrate and the side chain of Glu 165 are treated with QM in these calculations. Distances are given in angstroms.

model. The saddle point between the EDL2 and EDT2 species does not exist on the potential energy surfaces in either the basic model or the active-site model due to the uphill nature in this region for the gas-phase calculations.

The active-site structures involved along path C in the enzyme are shown in Figure 16. Interestingly, the geometry of the substrate Glu 165 moiety in the high-energy structure involved in the first proton-transfer step along path C is similar to that

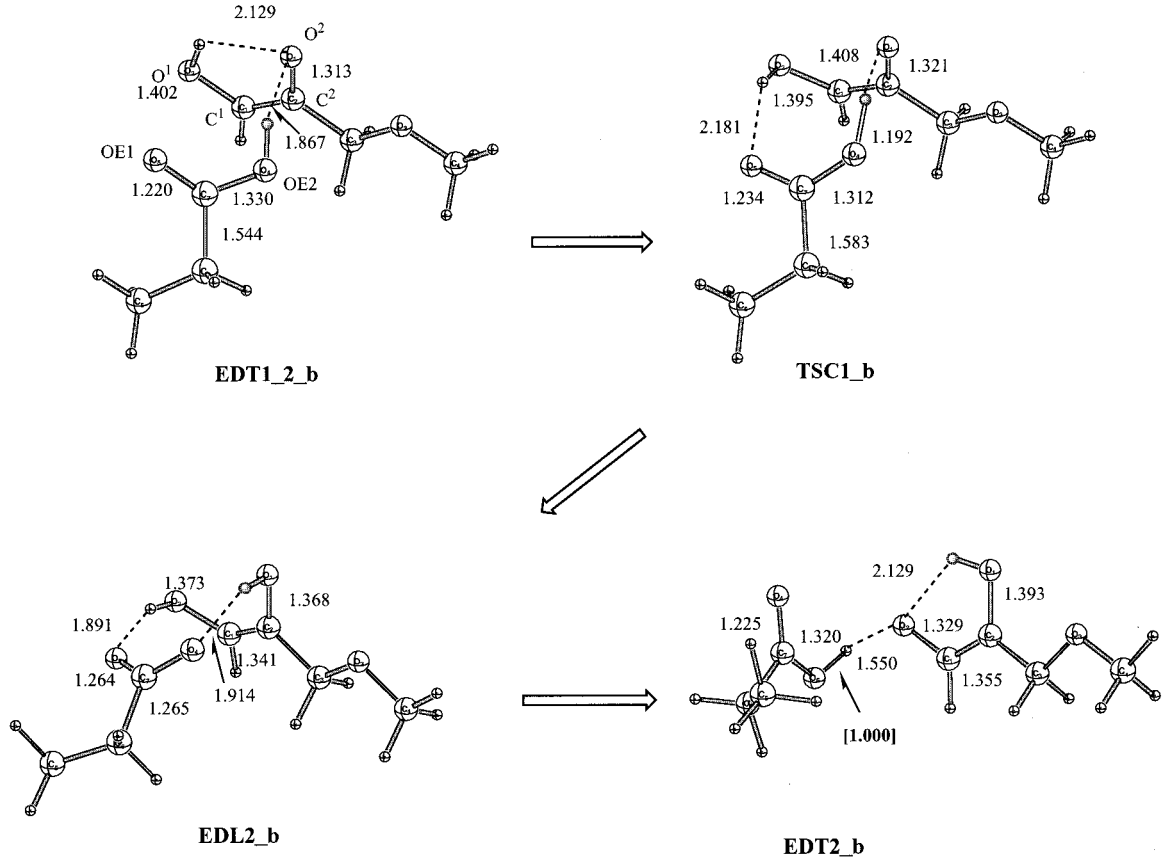


Figure 14. Optimized geometries for structures along path B with the basic model, labeled b, at the B3LYP/6-31+G(d,p) level. The filled circles indicate the transferred protons. The boldface number in brackets indicates that the bond was kept frozen during optimization (see text). Distances are given in angstroms.

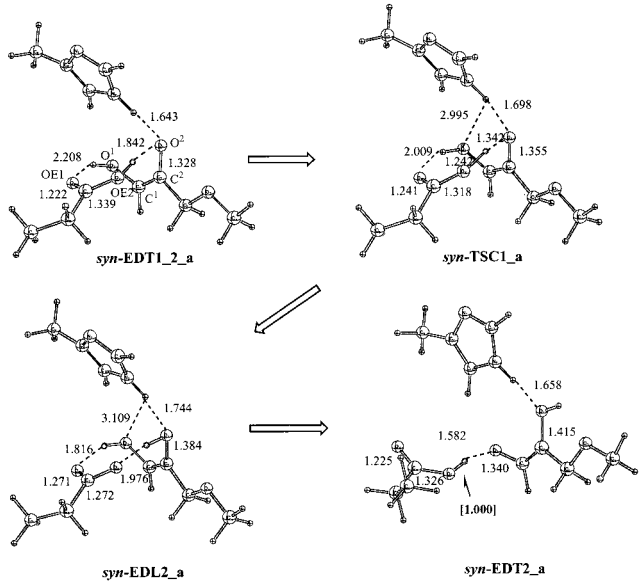


Figure 15. Optimized geometries for structures along path B with the active-site model, labeled a, at the HF/3-21+G level. The filled circles indicate the transferred protons. The boldface number in brackets indicates that the bond was kept frozen during optimization (see text). Distances are given in angstroms.

of EDT1_2 in the basic model; the proton to be transferred is about 2.155 Å from the acceptor oxygen atom in the substrate. In the proton-transfer product, EDL2, a very short hydrogen bond is formed between the donor and acceptor oxygen atoms; the OE2···H distance is only 1.383 Å. It has been reported in the literature that DFT overestimates the strength and underestimates the length of negatively charged hydrogen-bonded species.[83] To test this, we have carried out a geometry optimization at the HF/3-21+G*/CHARMM level; the resultant distance is 1.490 Å. It is clear, therefore, that this hydrogen bond in EDL2 is extremely short. As a result, the energy-limiting step for the initial step in path C is actually the displacement of the side chain of Glu 165 toward the substrate carbonyl oxygen. The cost in energy is 6.4 kcal/mol relative to EDT1, which is similar to the values of 3.6 and 8.7 kcal/mol in the basic and active-site models, respectively, in solution. As soon as the side chain is in position, the proton transfer is facile in the enzyme and leads to the formation of a short hydrogen bond in EDL2, which is 4.1 kcal/mol lower than EDT1. The value is similar but a little larger than that in the active-site model (1.4 kcal/mol).

In the next proton-transfer step along path C, the short hydrogen bond formed in ELD2 has to break. In the transition state, TSC2, this hydrogen bond increases in length to the more standard value of 1.614 Å. Substantial movement of the active-site water molecule Wat26 was observed during this proton transfer because the side chain of Glu 165 has to rotate significantly (compare TSC2 and EDT2). The barrier is 10.1 kcal/mol and EDT2 lies 2.4 kcal/mol above EDL2. It is interesting to note that EDT2 is a stable intermediate, in contrast to the situation in the basic and active-site models where the EDT2 species collapses to the enediol without any bond constraint. This occurs in the gas phase because the carboxylic group is more acidic than the enediol. In the enzymatic environment, the negative charge on O^1 in the substrate is stabilized by the surrounding residues such as Asn 10 and His 95. Perturbation analysis demonstrated that Lys 12 makes the

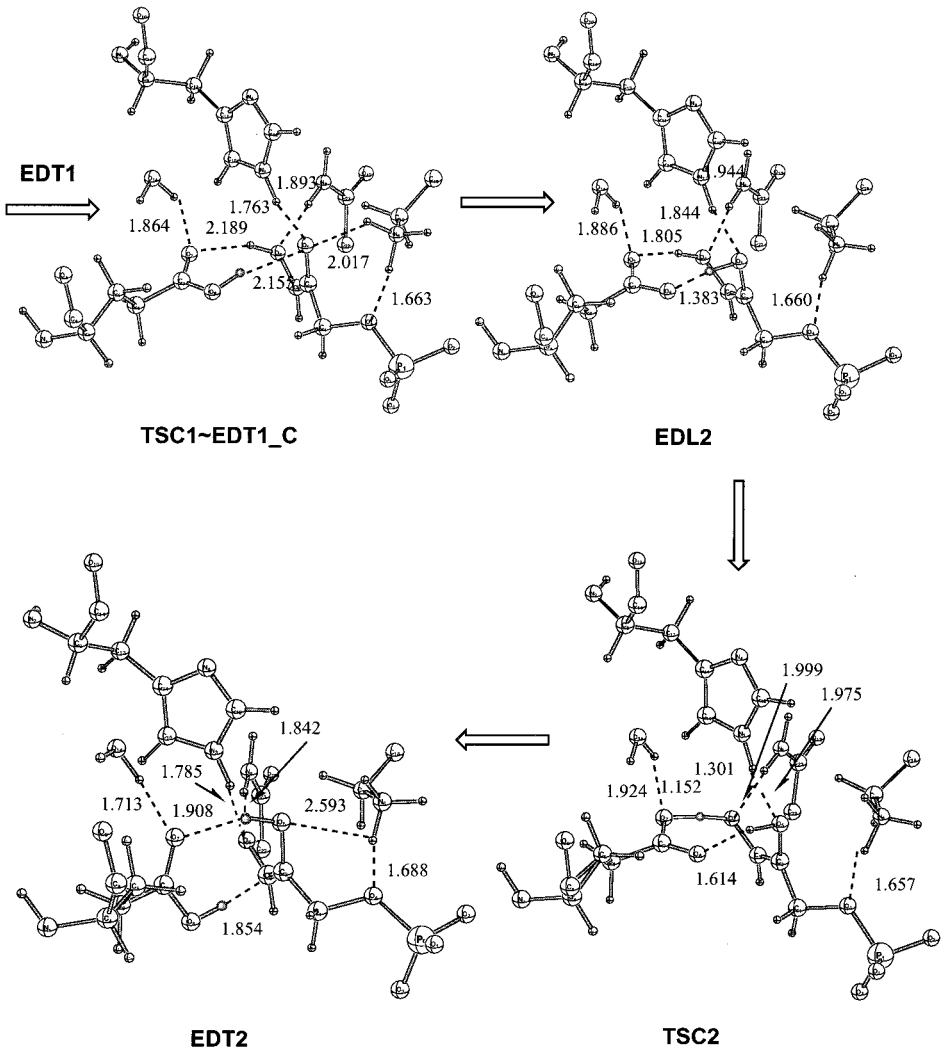


Figure 16. Active-site geometries at the B3LYP/6-31+G(d,p)/CHARMM level for important configurations along path C for the TIM reactions. The substrate and the side chain of Glu 165 are treated with QM in these calculations. Distances are given in angstroms.

largest contribution (see Figure 17c). It strongly destabilizes EDL2 relative to both EDT1 and EDT2 because the substrate is more negatively charged in the latter two structures. A few other residues, including Gly 210, Ile 127, and the active-site water Wat28, also favor the enediolate species over the enediol. Two other active-site water molecules, on the other hand, favor EDL2 over EDT1 and EDT2. His 95 also favors EDL2 over EDT1 by nearly 6 kcal/mol.

III.3. Other Aspects of Catalysis: III.3.a. Models with Glu 165 Lone Pair in Syn and Anti Orientations. It has been suggested that the syn orbital of the carboxylate group is more basic than the anti orbital.[3,48,49] From Table 4, we see that the anti-EDT1 structure is 13.6 kcal/mol higher in energy than *anti*-DHAP, which is lower than the value of 17.6 kcal/mol for the reaction involving the syn orbital. This suggests that it is easier (by 4.0 kcal/mol) to abstract a Cα proton with the anti carboxylate orbital than with the syn orbital. This, however, does not imply that the anti orbital is more basic than the syn orbital because other effects such as intermolecular hydrogen bonding have to be considered. From Figures 3 and 19, it is seen that a strong hydrogen bond exists between the substrate hydroxyl and "Glu 165" in the syn conformation; this hydrogen bond is absent in the anti structure. As a result, the *syn*-DHAP is 10.4 kcal/mol more stable than the *anti*-DHAP. This provides an alternative explanation for the fact that Glu 165 adopts the syn conformation in the substrate-bound state. After the proton abstraction, this hydrogen bond is significantly weakened, as reflected by the change in distance from 1.607 Å in *syn*-DHAP to 2.243 Å in *syn*-EDT1. Therefore, there is the additional energy penalty for proton abstraction in the syn conformation, which could be larger than 4.0 kcal/mol (the difference between the barriers in the syn and anti conformations). Taking this into account, we see that "intrinsically", proton abstraction with the syn orbital is not necessarily more difficult than the anti orbital. In the work of Joseph-McCarthy et al.,[27] the Glu 165 was

TABLE 4: Effect of Glu 165 Orbital Orientation and the Difference between His 95 and Gln 95 in the Active-Site Model[a]

species[b]	syn/His 95	anti/His 95	syn/Gln 95
DHAP_ a	*−916.37263*	*−916.35601*	*−938.68291*
EDT1_a	17.6	13.6	20.8
TSB_a	31.5	27.0	27.3
EDT2_a	25.3	24.4	22.0
TSA_a	*c*	17.7	nc[d]
EDL1_a	24.7	17.1	nc
TSA2_a	*c*	24.6	nc

[a] Syn/anti refers to the orbital orientation of the carboxylic group of the model Glu 165 (see text). The energies were calculated at the B3LYP/6-31+G(d,p) level at HF/3-21+G optimized structures. [b] Label a indicates the active-site model. [c] The transition state does not exist (or was very difficult to locate due to the fact that the potential energy surface is very flat) and all searches collapsed to structure EDL1. [d] The item was not calculated.

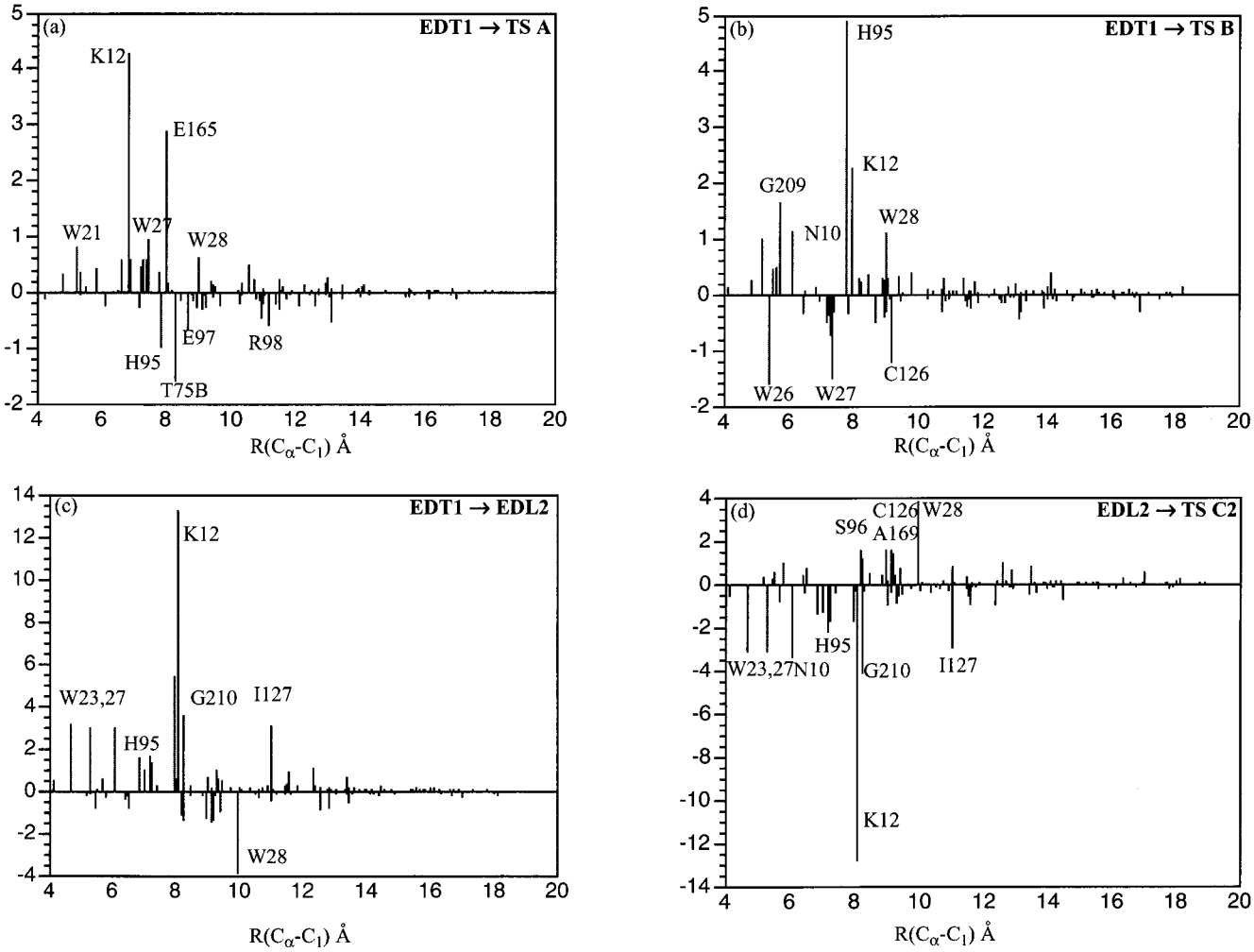


Figure 17. Perturbation analysis for energetic contributions from residues and X-ray observed water molecules in TIM at the AM1-SRP/CHARMM level. Positive values indicate unfavorable contributions. Shown are the results for selected processes along paths A, B, and C: (a) EDT1 → TSA; (b) EDT1 → TSB; (c) EDT1 → EDL2; (d) EDL2 → TSC2.

mutated to an Asp and the catalytic efficiency was observed to drop by about 1000-fold. According to the X-ray structure of the E165D mutant, the side chain of Asp 165 is about 1 Å further from the transition-state analogue. Further, the orientation of the side chain is such that Asp 165 has to employ the anti carboxylate orbital to perform the proton abstraction. Therefore, they hypothesized that the loss of catalytic efficiency of the E165D mutant is due either to the use of the less basic anti orbital or to the greater distance between the base and the substrate. According to the present calculations, we believe that the longer distance has a more significant effect on the catalytic efficiency and that the basicity of the anti orbital is less important.

For the other steps of the TIM reactions, the effects of the carboxylate group orientation are not expected to be large, so that only selected structures were examined. In Figure 19, we showed the structures involved along path A. One of the reasons for showing these structures is that the transition states have been located successfully at the HF/3-21+G level, while the corresponding ones in the syn set model have been found only at the AM1 level. The TSA is very early in character with a small barrier height of 0.6 kcal/mol, while TSA2 is very late, with a reverse barrier height of only 0.2 kcal/mol. Therefore, it is not surprising that the corresponding transition states in the syn set model are very difficult to locate if they exist at all.

III.3.b. His 95 vs Gln 95. In the presence of His 95, the steps with the higher barriers are the first proton transfer (DHAP → EDT1) and the intramolecular proton transfer (path B). We chose, therefore, to study those two steps with the neutral His 95 replaced by Gln 95 to compare with the mutational measurements on the H95Q mutant TIM.[12a] It is useful to recall that His 95 stabilizes the first process by 5.1 kcal/mol in solution (Figure 7a), while it increases the intramolecular proton-transfer barrier by 6.1 kcal/mol (Figure 18b), as compared with the values obtained in the basic model without His 95.

The electrostatic effect of Gln 95 is smaller than that of His 95, as expected. This is apparent from their influence on the energetics associated with the conversion from DHAP to EDT1; the value is 24.4 kcal/mol in the basic model and 20.8 (17.6) kcal/mol in the presence of Gln 95 (His 95). In other words, the electrostatic effect is equal to 3.6 and 6.8 kcal/mol for Gln 95 and His 95, respectively. A similar conclusion was drawn from earlier studies on the carbonyl stretching frequency in TIM.[19] One would expect, therefore, that the first proton-transfer step becomes more difficult when His 95 is replaced by Gln in the enzyme, even without invoking the factor that Glu 165 is further away from the substrate.[11b]

The effect of Gln 95 and His 95 on the intramolecular proton transfer barrier is also different. In the presence of Gln 95, the barrier height is 6.5 kcal/mol, only 1.8 kcal/mol higher than the value in the basic model in the gas phase. By contrast, His 95 increases this barrier by 9.1 kcal/mol. The unfavorable interaction between His 95 and the saddle-point structure is revealed by the longer $HN\epsilon-O^2$ distance, 1.883 Å, compared

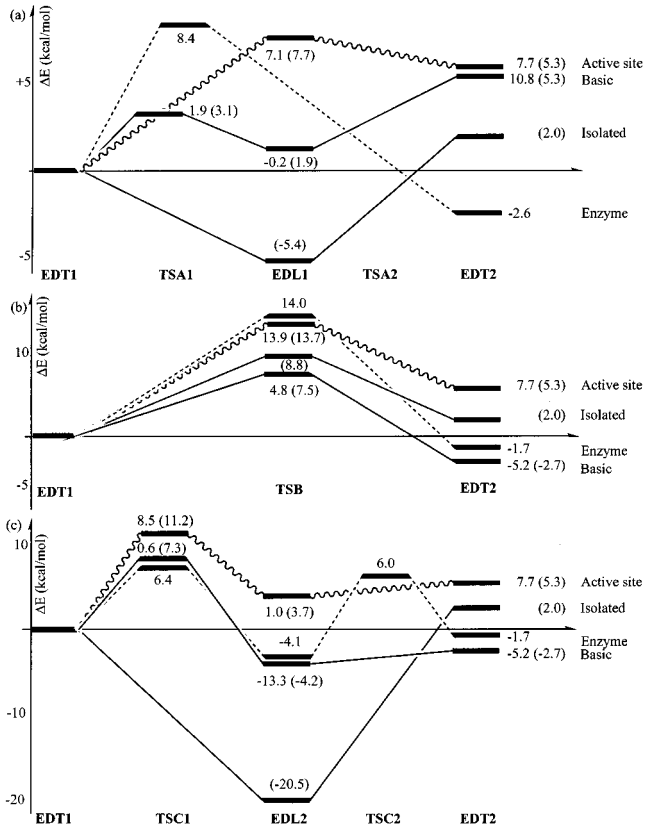


Figure 18. Relative energies for different models (see text) along the alternative proton-transfer steps; (a) path A; (b) path B; (c) path C.

to the value in *syn*-EDT1, 1.685 Å. By contrast, the HN−O^2 distance is about 1.75 Å in the two corresponding Gln 95 structures (Figure 20). An interesting implication of this result is that path B appears to be feasible for the H95Q mutant. However, path B appears to be inconsistent with the isotope transfer experiments of Knowles and co-workers.[12] Two important observations from that work are that no transfer (<1%) of tritium from labeled substrate (at C^1 position) to product was detectable (i.e., complete loss of tritium to solvent) and no solvent-derived tritium appeared in the remaining substrate (i.e., no gain of tritium from solvent). The first observation would require the C^1 proton to be transferred to a position that can exchange rapidly with the solvent. If the reaction follows path B, this would suggest that GluH 165 has to exchange rapidly with the solvent. This, however, would imply that the remaining substrate may pick up tritium from the solvent in the reverse reaction, which is in disagreement with the second observation. The fact that the H95Q mutant enzyme apparently selects path C over path B could be investigated by calculations with the proper H95Q active-site geometry.[11b]

IV. Concluding Discussion

In the following discussion, we consider the validation of the methodology and the role of the enzyme in catalysis as determined by the comparison with solution results and by the perturbation analysis of the reaction in the enzyme.

IV.1. Validation of Methodology. In the present work, the energetics of critical structures along the reaction paths were obtained from adiabatic mapping at the level of B3LYP/6-31+G(d,p)/CHARMM. To save computation time, the AM1-SRP/CHARMM approach was used as a guide; i.e., more stringent energy minimizations and saddle point optimization (based on the conjugate peak refinement)[60] were done first at the AM1-SRP/CHARMM level, and the resulting structures were reoptimized with the B3LYP/CHARMM approach. For this protocol to yield meaningful results, the AM1-SRP/CHARMM and B3LYP/CHARMM optima have to be rather similar, so that fast convergence can be achieved at the latter level starting from the structures obtained with the former. As shown in the Appendix, the AM1-SRP derived properties for gas-phase models are satisfactory. Accurate structures, energies and dipole moments for the gas-phase active-site models are obtained, as compared to the B3LYP/6-31+G(d,p) calculations; the RMS error in the energies and dipole moments are 3.9 kcal/mol and 0.87 D, respectively. For comparison, the RMS errors in the energies and dipole moments with the standard set of AM1 parameters are 11.1 kcal/mol and 2.08 D, respectively. Moreover, the AM1-SRP/CHARMM calculations also agree well with B3LYP/CHARMM calculations for both gas-phase models and actual reactions in the enzyme; for example, the RMS errors are 1.5 and 4.5 kcal/mol,[75b] respectively. These are significantly better than the values obtained with the standard set of AM1 parameters, which are 9.7 and 8.4 kcal/mol, respectively.

Another issue of concern is the reliability of the B3LYP approach associated with the 6-31+G(d,p) basis set for the proton-transfer reactions in TIM. Although B3LYP has been shown to be superior to Hartree−Fock and nearly as accurate as MP2 with large basis sets in many cases, its accuracy in proton-transfer processes and hydrogen-bonded systems has been questioned.[83] As shown by the comparisons with high-level ab initio methods such as MP2/6-311+G(d,p) and CCSD/6-311+G(d,p) presented in the Appendix, the B3LYP/6-31+G(d,p) method sometimes gives substantial errors on the absolute proton affinities [e.g., 8.5 kcal/mol for DHAP relative to CCSD/6-311+G(d,p)], but error cancellation usually leads to rather satisfactory results for proton-transfer reactions. For example, the exothermicity for the proton exchange reaction between DHAP and CH_3COO^- was found to be 13.4, 13.8, and 15.4 kcal/mol at the B3YP/6-31+G(d,p), MP2/6-311+G(d,p), and CCSD/6-311+G(d,p) levels, respectively. Overall, B3LYP performed rather well and gave results very similar to the MP2/6-311+G(d,p) calculations.

On the basis of these results, we believe that the B3LYP/6-31+G(d,p)/CHARMM method used in the present work is satisfactory for the purpose of distinguishing the energetics of the three proposed pathways. Further, in the companion paper by Lennartz et al.,[85] similar results for path A and B (they did not consider path C) were obtained from different computational models.

IV.2. Role of Enzyme in Catalysis. In a previous publication,[33a] we reported the results of QM/MM calculations on the energetics of TIM-catalyzed reactions with B3LYP/6-31+G(d,p) as the QM level (see Scheme 1 and Figure 1 in ref 33a). It was shown that paths A and C have similar rate-limiting barriers, while the intramolecular proton transfer in path B has a substantially higher barrier. This led to the conclusion that paths A and C are likely to be the catalytic mechanisms in wild-type TIM, in agreement with the recent NMR analysis of Mildvan and co-workers.[10] Perturbation analysis showed that His 95, a conserved residue in the TIM family, is the group that makes path B energetically unfavorable in the enzyme. In a separate study,[33b] the effect of proton tunneling was calculated for the first common proton transfer step and the intramolecular proton transfer reaction in TIM. It was found that the contribution of tunneling is generally rather small; it is less than a factor

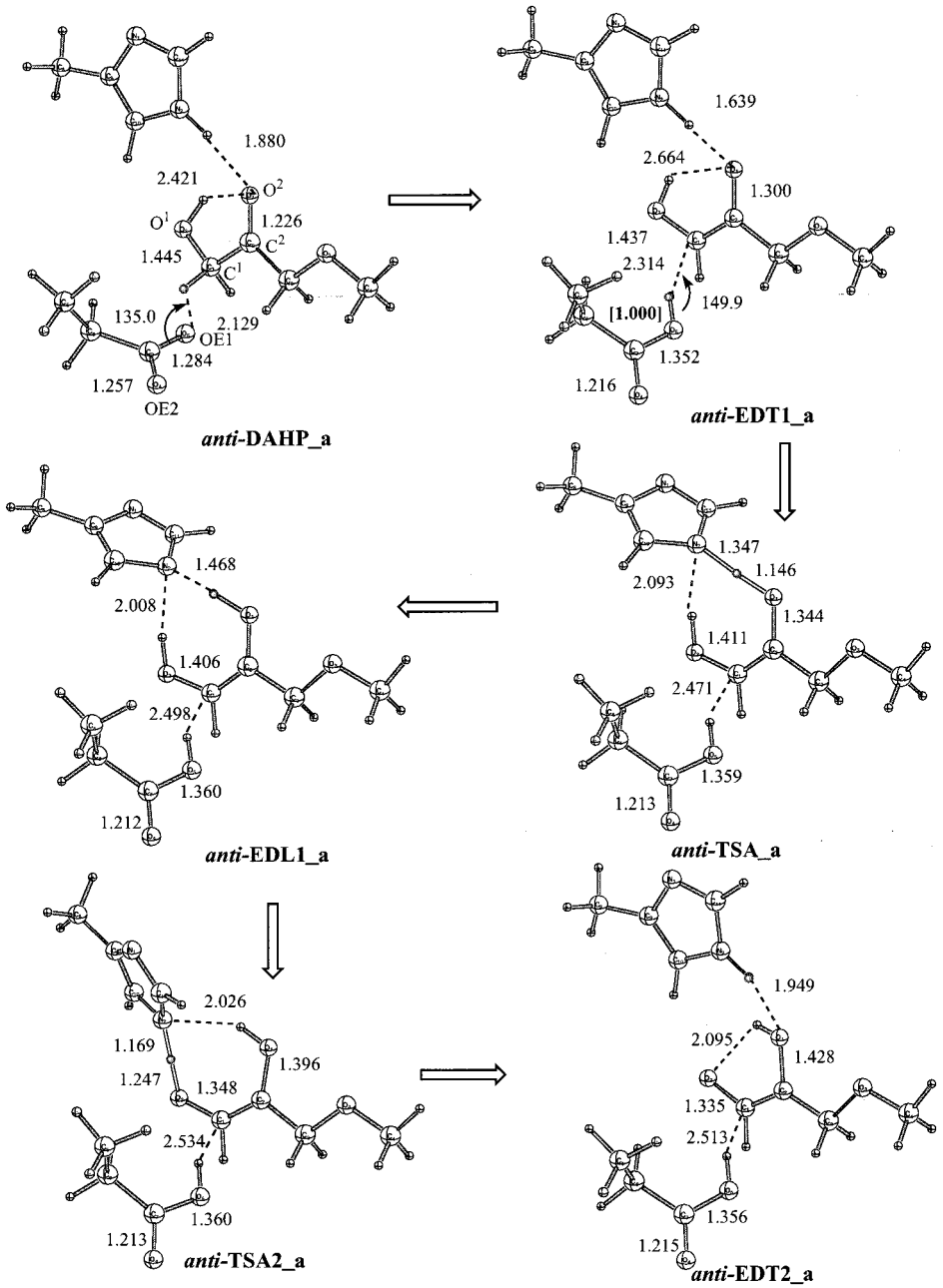


Figure 19. Selected structures in the active-site model (HF/3-21+G) with the $EtCOO^-$ in such a conformation that the anti orbital in the carboxylate group is facing the substrate (see text). Distances are given in angstroms.

of 10 for the rate constant at room temperature. Therefore, the overall conclusions of ref 33a are valid even if proton tunneling is included.

In the present work, we report more complete results on the corresponding reactions in the gas phase and in solution. Comparison with the results in the enzyme allows us to comment further on the origin of the catalytic power of TIM. The results for the common steps are shown in Figure 7 and for the alternative mechanisms in Figure 18.

The overall energetics in solution are significantly different from those in the gas phase (see Figures 7 and 18); note that a neutral $-OCH_3$ terminal group was used for the substrate to obtain meaningful gas-phase results. The major solvation effect is to decrease the strength of the hydrogen-bonding interaction involving ionic species due to solvent screening. This results in the destabilization of enediol intermediates in the basic models (i.e., models involving only the proton donor and acceptor groups arranged in the enzymelike orientation; see Figure 18c). The solvent also stabilizes the uncoordinated ionic species and leads to the stabilization of structures such as EDT2 in the active-site models (i.e., models involving the substrate, His 95, and Glu 165). Nevertheless, in the nonenzyme environments (gas phase and solution), the steps with the highest barriers are those associated with proton transfer from and to the carbon atoms of the substrate, i.e., DHAP → EDT1 and EDT2 → GAP. The calculations indicated that the uncatalyzed proton transfer in solution has a barrier of *at least* 24 kcal/mol (because the donor and acceptor groups were arranged in the favorable orientation found in the enzyme), which is close to the experimental estimate of 26 kcal/mol for the uncatalyzed reaction in solution ($k = 6 \times 10^{-6}$ s^{-1}).[1] With the apparent barrier of 13 kcal/mol for the reactions in enzyme estimated from calculations and experiments,[2−4] there is a rate acceleration of 10^8 for TIM. This is close to the estimate of 10^9 from kinetic measurements,[3] which indicates that both the enzyme and solution model calculations are consistent and reasonable. It

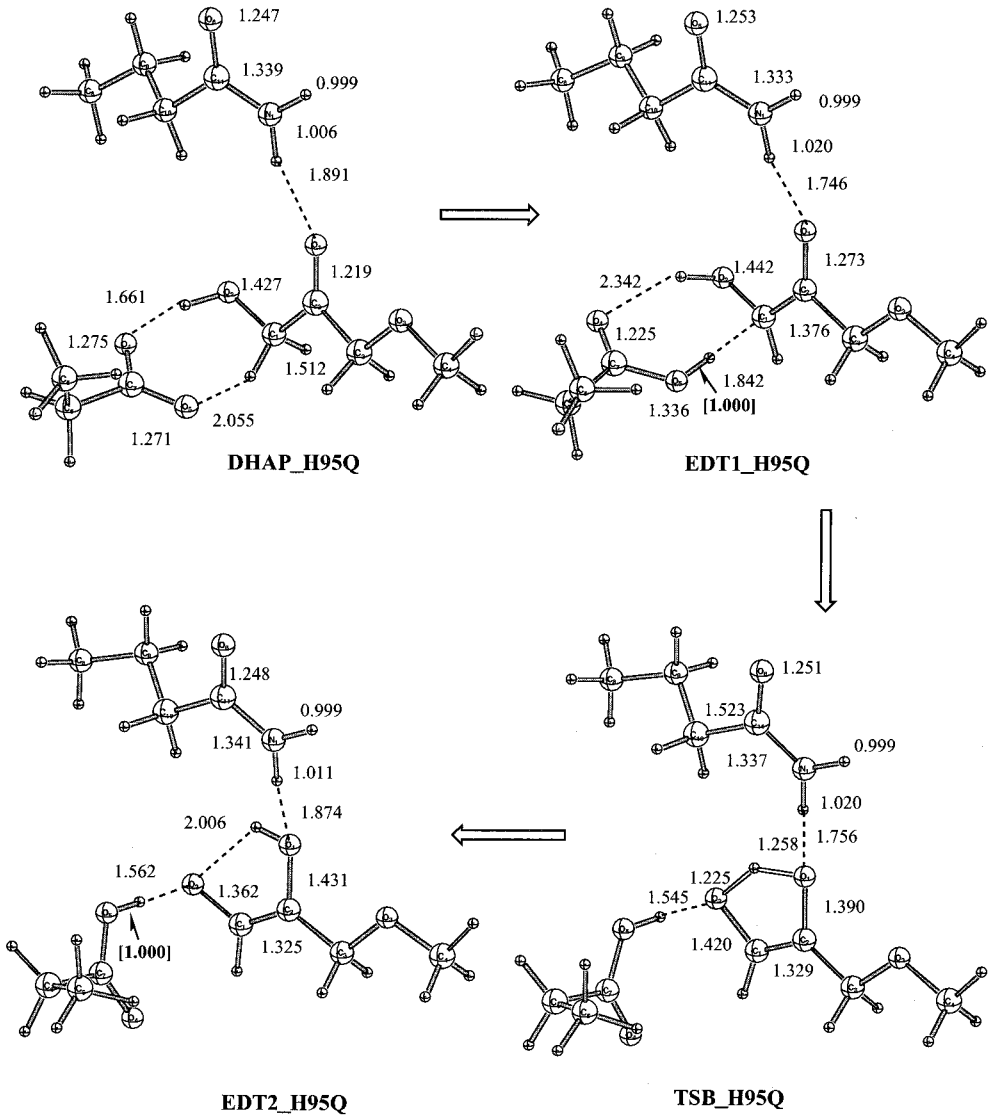


Figure 20. Selected structures for the active-site model (HF/3-21+G) with the His 95 replaced by a model Gln residue. Distances are given in angstroms.

should be noted that because the chemical steps are not rate-limiting in wild-type TIM, the experimental value for the barrier height (and therefore the value for the rate acceleration) was estimated by use of the elegant, but indirect, isotope exchange−conversion technique, which assumed that one, and only one, kinetically significant intermediate occurred during the reaction.[4] The proton transfer processes that do not involve carbon atoms have much lower barriers. Therefore, the primary issue for the enzyme is to stabilize the enediolate species, EDT1 and EDT2, which is exactly what TIM has achieved. As can be seen from Figure 18, most of the other proton transfer steps become more difficult in the enzyme compared to solution. For example, the proton transfer from the enediolate to His 95 along path A has a barrier of only 3.1 kcal/mol in solution, compared to the value of 8.4 kcal/mol in the enzyme (Figure 18a); the proton transfer barrier along path B also is higher in enzyme than in solution by 6.5 kcal/mol (Figure 18b). The reaction of DHAP to give EDT1 becomes more facile in the enzyme due largely to the strong stabilization of EDT1 by Lys 12 (but see below). Moreover, along path A, EDT2 becomes very close in energy (−2.6 kcal/mol) to EDT1 in the enzyme, while the former is less stable by 14.6 and 4.4 kcal/mol in the basic model than in the gas phase and in solution, respectively.

The perturbation analysis at the AM1-SRP/CHARMM level presented in section III gives more detailed information on the energetic contribution from individual residues in variety steps of the TIM reaction. It shows that Lys 12 and certain polar residues, such as His 95, Asn 10, and Ser 211, plus a number of water molecules in the active site, make the most significant contributions to the energetics of the TIM reactions. However, a few residues that are relatively far from the active site, such as Gly 210 and Ile 127, also make significant contributions for certain steps; i.e., the main-chain atoms on these nonpolar residues have a substantial polarizing effect on the reacting groups in proton transfers along path C (see Figure 17c,d). Distant charged residues were found to have a smaller contribution than in the original work of Bash et al.[15] because the dielectric shielding effect of the solvent was treated more accurately; that is, their partial charges were scaled down in the QM/MM calculations based on the solvent screening effect estimated from a continuum solvation (Poisson−Boltzmann) calculation.[56]

Since the barrier for proton transfer reactions depends sensitively on the distance between the donor and acceptor atoms, the first task of the enzyme is to provide an active-site geometry with the key atoms of the catalytic residues in good

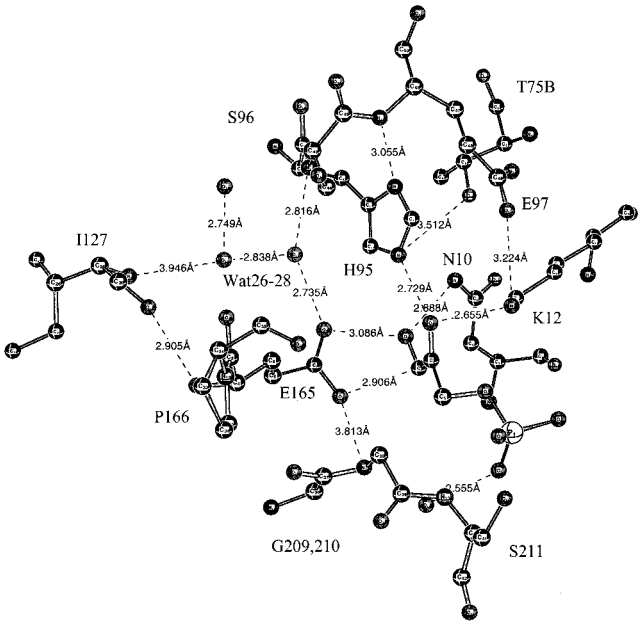


Figure 21. Position of residues and active-site water molecules that were found to be important in the perturbation analysis (see Figures 8 and 17). The structure for EDT1 optimized at the B3LYP/6-31+G-(d,p)/CHARMM level is shown.

positions. The arrangements of the side chains of Glu 165 and His 95 in TIM clearly illustrates this point. Further, as discussed in section III.1.a, placing the reacting groups (i.e., CH_3COO^- and a model DHAP) in the X-ray positions in a solution environment makes the proton transfer between DHAP and Glu 165 less endothermic (Figure 7a). However, this is not sufficient; e.g., the proton transfer from DHAP to EDT1 is still very endothermic (~24 kcal/mol!) in the basic model (with only the substrate and model Glu 165). The protein provides electrostatic interaction and polarization by specific residues to stabilize EDT1; this is achieved by Lys 12 and other polar residues as described above. Furthermore, for the electrostatic polarization to be effective, the protein has to provide an active site that has a low effective dielectric constant. The importance of this effect in enzymatic catalysis has been recognized for a long time.[63] As shown in section III.2, the large electrostatic effect (~17 kcal/mol) from a nearby (~2 Å) NH_4^+ is almost completely quenched if the reacting fragments are immersed in a high dielectric medium corresponding to aqueous solution. The perturbation analysis shows that Lys 12 has a large electrostatic effect (on the order of 10 kcal/mol) for several steps, which requires that the effective dielectric constant in the active site is rather low. This is supported by previous simulations, which indicated that the dielectric shielding in the interior of proteins is rather heterogeneous and that the effective dielectric "constant" is rather low (in the range of 2−6; a value of 4 was used in the Poisson−Boltzmann calculations to estimate the charge-scaling factors) in most regions except close to the protein−water interface.[64,65] The flexibility of the active-site residues, which allows them to effectively stabilize certain intermediates or to get involved explicitly in chemical reactions, is also crucial for the catalysis. A good example in TIM is Glu 165, which has to reorients its side chain considerably (with an RMS of 0.90 Å for the heavy atoms) to shuttle protons back and forth along path C. The positions of residues and active-site water molecules that were found to be important in the perturbation analysis are summarized in Figure 21.

Not all the proton-transfer barriers have to be lower in the enzyme than in solution to achieve the overall rate acceleration. For example, the barrier along path A for the transfer from EDT1 to His 95, which is not the rate-determining chemical step, is raised by residues such as Lys 12 and Glu 165 (see section III.2.a), consistent with the fact that EDL1 is not stabilized in the enzyme. Further, as suggested first in the calculations reported in ref 15 and confirmed by experiment in ref 9, a neutral His 95 is involved in the proton transfer to avoid overstabilization of the enediol. Since the increase in the barrier is due mainly to Lys 12 and Glu 165 (see section III.2.a), the two residues that are critical for the first proton transfer (DHAP to Glu 165), it is likely this is simply a (passive) consequence of the (active) stabilization of EDT1 by the enzyme. The major evolutionary pressure on TIM was to lower the barrier between DHAP and EDT1, in part by stabilizing the latter relative to DHAP (or GAP), since this corresponds to the slow step in the uncatalyzed reaction in solution.

Finally, it is interesting to use the results from present calculations to rationalize the fact that His 95 is a conserved residue among all known TIM sequences. It plays an essential role along path A (as proton donor and acceptor) and also makes favorable contributions through electrostatic interaction to the first proton transfer and in path C (see Figures 8 and 17, respectively). Although the electrostatic effect is relatively small (on the order of 2−3 kcal/mol), the fact that the stabilization is essential for the first proton-transfer step (which is the most difficult elementary step during the reaction cycle) makes it favorable to conserve His 95. A simple active-site model calculation (see section III.3.b) indicated that if the His 95 is mutated to a Gln, the enediolate species will be destabilized by about 3 kcal/mol. This is in accord with mutagenesis experiments, which found that the H95Q mutant has a lower (~200-fold) catalytic efficiency as compared with the wild type.[12]

IV.3. General Considerations. Many different mechanisms have been proposed to explain catalysis by enzymes, and it is likely that different factors play the primary role in different enzymes.[66] They include electrostatic interaction and polarization, proximal and appropriately oriented catalytic groups, preorganization of dipoles, [68] low-dielectric constant of the active site,[69] and metal ion chemistry. Also, much attention has recently been focused on the effect of proton tunneling[34,35] and enzyme dynamics[84] in catalysis. In a separate study,[33b] we show that tunneling is not a major factor in catalysis; at room temperature, it is less than a factor of 10 for the first proton transfer between DHAP and Glu 165 and for the intramolecular proton transfer.

Other suggestions focus on low-barrier hydrogen bonds (LBHB).[30,67] The LBHB model proposes that short, very strong, partially covalent hydrogen bonds are formed between pK_a-matched groups, which makes proton transfer easier. Indicative of LBHB are the short bond length, a very low-field NMR chemical shift, low fractionation factors for solvent exchange, and positive values for the deuterium isotope effects on the chemical shift.[67] TIM is a system exhibiting LBHB between the reactant substrate and Glu 165, based on the NMR experiment of Mildvan and co-workers;[10] other enzyme candidates are serine proteases, ketosteroid isomerase, citrate synthase, and mandelate racemase.[67] In citrate synthase, our earlier work[24] has shown that LBHB are not involved in catalysis. In the present work, the QM/MM calculations indicates that there exists a very short hydrogen bond between Glu 165 and the substrate DHAP, in agreement with experiment. In a separate study,[70] chemical shifts calculations in a QM/MM framework[71] qualitatively reproduced the chemical shifts to low field. However, we did not find any evidence that this short hydrogen bond plays

an essential role in catalysis; i.e., the hydrogen bond does not contribute to lowering the barrier for the first proton transfer, which is the rate-limiting elementary step. It might play a role in increasing the binding energy of the substrate to the active site. It was also proposed that a LBHB forms between His 95 and the enediol or enediolate forms of the possible intermediates.[67c] In the current calculation, we did observe that the hydrogen bond between His 95 and the substrate is rather short in EDT1 for path A; the donor−acceptor distance is 2.6 Å. However, the barrier for this step is not lower (8.4 vs 3.2 kcal/mol) than that in other models with a longer (normal) hydrogen-bond distance (e.g., 2.8 Å in EDT_M1 in Appendix 2). Therefore, our calculations do not support the catalytic role for this LBHB. We also observed a very short hydrogen bond between Glu 165 and the enediol species in EDL2 along path C; the donor−acceptor distance is about 2.4 Å. The existence of such a LBHB indeed makes the proton transfer from GluH 165 to the enediolate easier, so that the energy-determining structural change for this step of the reaction is the displacement of GluH 165 (see section IV.1). However, we have already pointed out that this proton-transfer step does not involve a rate-limiting barrier in the gas phase or in solution. Moreover, since the LBHB has to break for the subsequent proton transfer to take place, the product EDT2 becomes destabilized relative to EDL2 in the enzyme compared to solution (see section IV.2), such that the step from EDL2 to EDT2 becomes more difficult in enzyme.

In summary, the present calculations suggest that TIM, probably like many other enzymes, achieves its catalytic power by a combination of factors. First, the active-site residues are arranged with the key atoms positioned such that their distances and orientations facilitate the proton transfers. This has been referred to by Warshel and Florián[68a] as a "trivial" cage effect, but to us it appears to be an important aspect of the catalysis; i.e., it is not generally true that there is a high probability in solution to satisfy this condition. The enzyme also has to provide charged and polar residues to stabilize the (charged) reaction intermediates (e.g., the enediolate). For the electrostatics to function effectively, the active site of the enzyme has to have a relatively low, though not necessarily gaslike, effective dielectric constant. Rather short hydrogen bonds were observed in several structures during the reaction, which is not surprising given that ionic groups (either the catalytic residue or substrate) are involved. However, the calculations suggest that they do not play an active role in catalysis, though they may be important for binding.

Although the calculations reported here are based on a high-level QM/MM treatment, a limitation is that only one typical configuration of the enzyme was used in the reaction path optimizations. The dynamics of the protein environment can influence the precise values of the barriers, although it is unlikely that such effects will change the relative results for the energetics along different paths. For example, potential-of-mean-force calculations with molecular dynamics gave the same trend for the difference between path A and path B. These results, as well as an exploration of the relation between the barrier heights and the structure of the active site, will be published separately.

Acknowledgment. We thank Professor Barone for providing the latest PCM code in Gaussian94 during the early stage of the work and Professor Tomasi for discussions on solvation effect. The genetic algorithm code made available by Professor D. L. Carroll of the University of Illinois is gratefully acknowledged; it was used for optimizing the reaction-dependent AM1 parameters (AM1-SRP). We thank Dr. X. Lopez, Dr. H. Guo, Dr. S. Fisher, and Dr. S. So for helpful discussions. Discussions with Professor W. Thiel and his collaborators during the final stages of this work, after we learned that they were doing parallel studies, are gratefully acknowledged. Part of the computations have been done on the J90 machines at NERSC, SGI Origin2000 machines at NCSA, and the IBM SP2 machines at Argonne National Lab. The work was supported in part by grants from the Department of Energy and the National Institute of Health.

Appendix 1: Parameters for QM/MM

A1.1. AM1-SRP Parameters. Despite significant progress in linear-scaling DFT and ab initio methodologies[76] over the past decade, it is still unrealistic to perform long DFT/MM molecular dynamics simulations for biological systems. Semiempirical methods such as AM1 and PM3 are much faster ($\sim 10^3$ times) and can be used in long simulations, but they are in general less accurate, especially for the calculation of barriers involved in reactions, including proton transfers. Recently, several authors[77] have used an approach in which the parameters in the semiempirical Hamiltonian are adjusted for the specific system under study. The essential point is that the semiempirical Hamiltonian, which has overall reasonable characteristics for representing bond-formation and bond-breaking phenomena, is used as the fitting functional, instead of polynomials or other analytical mathematical functions used traditionally to fit potential energy surfaces to the results of high-level calculations.[78] In this way, the speed of the semiempirical method is preserved, while the required accuracy for the system of interest is achieved; this is particularly useful for systems with many degrees of freedom. A promising alternative is to develop semiempirical methods that have improved accuracy overall.[75]

In the current work, we adjusted the AM1 parameters for H, N, C, and O atoms; this includes 49 parameters in all. The optimized parameter set is referred to as AM1-specific reaction parameters (AM1-SRP). The parameters for the P atom were not optimized because the PO_3^{2-} group in the substrate is a spectator during the reaction. The data set used for the adjustment includes the gas-phase energetics of the various reactions, norms of gradients for the geometries, and dipole moments of the "basic model" set of structures (see Methods). Since the parameters were designed to be used in calculations in the enzymatic environment, one has to verify that AM1-SRP behaves well in the presence of external charges. To this end, we also included a number of "active-site model" structures with the fragments that are not directly involved in the proton transfer (such as His 95 during DHAP → EDT1) treated with CHARMM. Without these QM/MM data points, we found that the fitted AM1-SRP produces spurious results for calculations in the enzyme. The energies, dipole moments, and norms of the gradient of these structures (i.e., the gradient at the AM1-SRP level for the B3LYP-optimized structures is small, so that the optimized structures at AM1-SRP would be close to the B3LYP ones; this avoids performing geometry optimization at the AM1-SRP level during the parameter fitting calculations) were considered in the fit. The target function subject to minimization, c2, is defined as

$$\chi^2 = w_g \sum_{i=1}^{N_{tot}} |g_i|^2 + w_E \sum_{i=1}^{N_{tot}} |\Delta E_i^{AM1-SRP} - \Delta E_i^{B3LYP}|^2 + w_\mu \sum_{i=1}^{N_{mini}} |\mu_i^{AM1-SRP} - \mu_i^{B3LYP}|^2 \quad (3)$$

TABLE 5: Definition of Model Systems Considered in This Work

notation	name	definition
M_0	zero-set[a]	simplified substrate without -$CH_2OPO_3^{2-}$
M_1	minimal set 1[a]	model His 95 and a small enediolate model
M_{2X}	minimal set 2[b]	substrate with different terminal groups (X) for -$CH^2OPO_3^{2-}$
M_b	basic model[c]	model substrate plus one catalytic group (Glu 165 or His 95)
M_a	active-site model[c]	model substrate plus both catalytic groups (Glu 165 and His 95)

[a] For structures, see Figure 23. [b] For structures, see Figure 24. [c] For structures, see Figures 2, 3 (DHAP → EDT1), 5 (EDT2 → GAP), 9 (path A), 12 (path B), and 14 (path C).

TABLE 6: Optimized Specific Reaction Parameters for the TIM-Catalyzed Reactions[a]

	set 1			set 2			set 3			
	C	O	H	C	O	H	C	O	N	H
U_{SS}	−46.9528	−99.7751	−11.2522	−51.2992	−105.5675	−12.5250	−52.0363	−98.0929	−78.1543	−10.4839
U_{PP}	−39.2794	−78.4345		−42.5926	−83.4430		−41.4701	−80.5754	−60.0860	
β_S	−14.2486	−29.7662	−6.7640	−14.5941	−26.4056	−6.0981	−14.8521	−28.9053	−20.8356	−6.1132
β_P	−6.9677	−27.8642		−6.9549	−31.2877		−7.6993	−29.4058	−17.5963	
ζ_S	1.9256	3.3274	1.0862	1.7614	3.4152	1.1451	1.7807	3.2943	2.1748	1.1524
ζ_P	1.5951	2.4218		1.6067	2.3471		1.5332	2.4477	2.0558	
α	2.6572	4.7895	3.0534	2.7154	4.7877	3.1424	2.7439	4.5757	3.1827	2.9816
FN11	0.0111	0.2648	0.1320	0.0119	0.2585	0.1158	0.0112	0.3062	0.0243	0.1145
FN12	0.0501	0.0846	0.0048	0.0415	0.0826	0.0051	0.0415	0.0889	0.0275	0.0051
FN13	−0.0192		−0.0183	−0.0219		−0.0181	−0.0218		−0.0060	−0.01763
FN14	−0.0013			−0.0014			−0.0012			
FN31	1.4472	0.7671	1.1383	1.4625	0.8253	1.1488	1.4813	0.7944	1.5105	1.3109
FN32	1.6659	1.3079	1.9724	1.9866	1.4460	1.8402	2.0200	1.4255	2.1508	1.6745
FN33	1.9494		2.2919	1.9673		2.1990	2.0457		2.1895	1.9300
FN34	2.6040			2.8021			2.4521			

[a] Set 1 is fitted for the common proton-transfer steps (DHAP → EDT1; EDT2 → GAP) and path C; set 2 is fitted for the intramolecular proton-transfer step, path B; set 3 is fitted for the proton-transfer path A. The parameters are allowed to change within 10% of the standard AM1 set of parameters (see text). For the standard AM1 parameters and their units, see ref 22.

where $w_g = 0.1$, $w_E = 1.0$, and $w_\mu = 10.0$; this set of weights was chosen such that the three contributions have comparable numerical values. N_{mini} and N_{tot} are the number of basic model structures and the total number of structures used in the fit, respectively. The optimizations are carried out with a genetic algorithm[79] with a population of 50 and rates of 0.01, 0.5, and 0.02, respectively, for a single mutation, a crossover, and a creep mutation, respectively. The parameters are allowed to change by ±10% relative to the standard AM1 values. About 200 iterations were required to achieve convergence. We were unable to obtain one set of parameters that accurately described the reactions involved for all three paths; therefore, three different sets of parameters were used for paths A, B, and C. The optimized parameters are listed in Table 6.

The AM1-SRP results for the energetics for the basic model systems are shown together with the B3LYP and standard AM1 data in Table 7. The correlations between the AM1-SRP and B3LYP energetics and dipole moments are illustrated in Figure 22. The AM1 results for intermediates are in quite good agreement with the B3LYP/6-31+G(d,p) results, but the barriers are too high by about 10.0 kcal/mol. An extreme case is the intramolecular proton-transfer transition state, where AM1 overestimates the barrier by 20 kcal/mol. The AM1-SRP results are in much better agreement with the B3LYP results; e.g., the barriers for the basic model (TSB) are 7.1 and 4.8 kcal/mol, respectively. The RMS differences between AM1-SRP and B3LYP/6-31+G(d,p) and between AM1 and B3LYP/6-31+G-(d,p) are 1.84 and 11.9 kcal/mol, respectively, for the energies and 0.87 and 2.08 D, respectively, for the dipole moments. The AM1-SRP parameters derived here have been used in the rate constant calculations for the proton-transfer reactions with tunneling in a separate study.[33b]

A1.2. QM/MM van der Waals Parameters. In the standard QM/MM formulation, the total energy of the system is given by

TABLE 7: Comparison of the Energetics at the AM1, AM1-SRP, and B3LYP/6-31+G(d,p) Levels for the Basic Model Set of Structures[a]

species[b]	AM1	AM1-SRP	B3LYP/6-31+G(d,p)
DHAP_b	*−0.42932*	*−1.43170*	*−650.79608*
TS1_b	*c*	*c*	*c*
EDT1_b	18.0	26.3	24.4
TSB_b	52.2	33.4	29.2
EDT1_2_b	19.9	23.6	20.6
TSC1_b	39.0	34.4	25.0
EDL2_b	13.1	14.1	11.1
TSC2_b	23.8	18.3	*d*
EDT2_b	18.5	18.3	19.2
EDT1_3_b	*−0.17402*	*−2.30015*	*−647.90798*
TSA_b	10.3	5.1	1.9
EDL1_b	3.4	4.4	−0.2
TSA2_b	15.9	11.5	*e*
TSB_b	36.2	14.3	13.0
EDT2_2_b	−1.3	5.9	10.8
EDT2_np_b	*−0.41324*	*−1.41302*	*−650.77001*
TS2_b	5.5	0.3	0.5
GAP_b	−11.4	−12.4	−11.5
RMS difference[f]	11.1	3.9	

[a] The total energies for reference structures are given in hartrees in italic type; the relative energetics are given in kilocalories per mole. [b] Label b indicates the basic model. [c] The transition state does not exist (or was very difficult to locate due to the fact that the potential energy surface is very flat); in this case all searches collapsed to structure EDT1_b. [d] As in footnote *c*; in this case all searches collapsed to structure EDT2_b. [e] As in footnote *c*; in this case all searches collapsed to structure EDT2_2_b. [f] The RMS difference in energies compared to the B3LYP/6-31+G(d,p) results, given in kilocalories per mole.

$$E = \langle \Psi(\vec{\mathbf{r}};\vec{\mathbf{R}}_{QM},\vec{\mathbf{R}}_{MM}) | H^{QM} + H_{el}^{QM/MM} | \Psi(\vec{\mathbf{r}};\vec{\mathbf{R}}_{QM},\vec{\mathbf{R}}_{MM}) \rangle + E^{MM}(\vec{\mathbf{R}}_{MM}) + E_{van}^{QM/MM}(\vec{\mathbf{R}}_{QM},\vec{\mathbf{R}}_{MM}) \quad (4)$$

where H^{QM} and $H_{el}^{QM/MM}$ are the Hamiltonian of the QM atoms and that of the QM/MM electrostatic interactions, respectively.

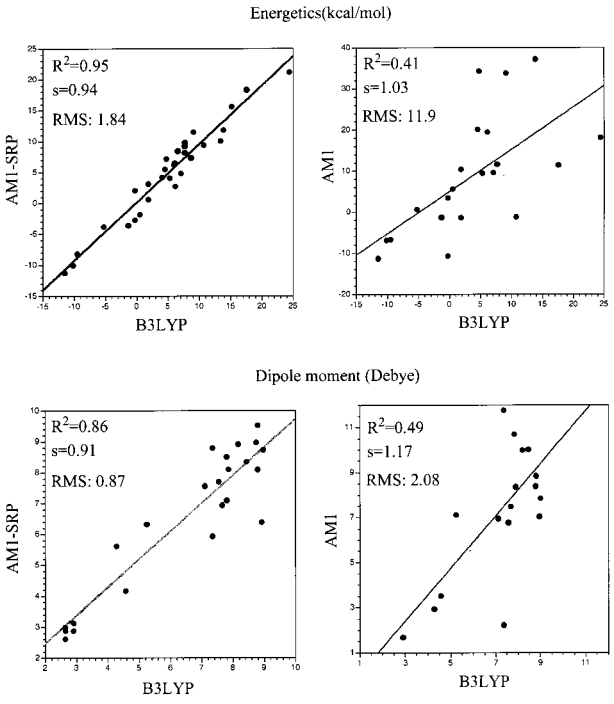


Figure 22. Comparison of AM1, AM1-SRP, and B3LYP/6-31+G-(d,p) energetics (in kilocalories per mole) and dipole moments (in debyes) for model structures constructed for the TIM reactions. The data set includes full QM calculations for the basic model and QM/MM calculations for the active-site model. The comparisons are characterized by root-mean-square deviation (RMS), coefficient of determination (R^2), and slope (s) from a least-squares regression analysis.

Other effects such as the Pauli repulsion and dispersion between the QM and MM atoms are approximated by van der Waals terms. To obtain accurate QM/MM geometries and energetics, an optimized set of van der Waals parameters for the QM atoms has to be used.[80]

Here we choose to optimize the van der Waals parameters of only four atoms in the substrate, C^1, C^2, O^1, and O^2 (see Scheme 1), which are explicitly involved in the reaction. For all other atoms, we used the standard CHARMM parameters. A set of van der Waals parameters were developed for the QM atoms in the side chains of Glu 165 and His 95 (not shown). However, test calculations showed that the standard CHARMM parameters give similar relative energetics. The strategy of parameter optimization used here is similar to that employed by Ho et al.,[80] in which a rigid water molecule was used as a probe around the atom of interest in the model substrate. The van der Waals parameters are adjusted such that the QM/TIP3P calculation that treats the water with the TIP3P model gives a satisfactory water−substrate distance and interaction energy, as compared with full B3LYP/6-31+G(d,p) results. Three different sets of QM van der Waals parameters were fitted for AM1, AM1-SRP, and B3LYP/6-31+G(d,p), by use of a genetic algorithm similar to that employed in the AM1-SRP optimizations. Due to the small number of parameters involved, the optimization converged within about 50 iterations. The optimized van der Waals parameters are shown in Table 8, along with parameters for similar atoms in the CHARMM 22 force field for proteins.

A1.3. Test of QM/MM Models. To examine the accuracy of the QM/MM approach and also the transferability of the AM1-SRP parameters derived above, we carried out test QM/MM calculations for the structures of the active-site model (see Methods and Table 5) with the fragments not explicitly involved in the proton transfer treated by the CHARMM22 force field. Three different QM levels were tested: they are AM1, AM1-SRP, and B3LYP/6-31+G(d,p). For simplicity, only single-point QM/MM energetics at the HF/3-21+G optimized structures were calculated, and the values were compared with full B3LYP/6-31+G(d,p) energies at these geometries. The results are collected in Table 9.

Encouragingly, both AM1-SRP/CHARMM and B3LYP/6-31+G(d,p)/CHARMM energies are in satisfactory agreement with the full B3LYP/6-31+G(d,p) results, with RMS errors of 1.5 and 1.3 kcal/mol, respectively. The AM1/CHARMM results, by contrast, deviate considerably more, especially for EDL2_a and TSB_a, and the RMS error compared to the full B3LYP/6-31+(d,p) energies is 9.1 kcal/mol. The large error comes from the AM1 itself, not from the AM1/CHARMM treatment. This is made clear by comparing AM1/CHARMM to full AM1 single-point energies, for which the RMS difference is only 3.5 kcal/mol.

TABLE 8: Optimized QM/MM van der Waals Parameters[a]

parameters	CHARMM 22	AM1	AM1-SRP[b]	B3LYP/ 6-31+G(d,p)
ϵ_{C1}	−0.11	−0.1957	−0.1960	−0.0900
R_{C1}	4.00	3.94	3.80	4.39
ϵ_{C2}	−0.11	−0.1672	−0.1870	−0.1567
R_{C2}	4.00	3.95	3.60	3.59
ϵ_{O1}	−0.152	−0.1263	−0.1000	−0.0635
R_{O1}	3.48	2.92	3.20	3.60
ϵ_{O2}	−0.152	−0.0518	−0.1750	−0.0516
R_{O2}	3.48	2.88	2.88	3.65

[a] The well depths (ϵ) are given in kilocalories per mole; the van der Waals radii (R) are given in angstroms. Parameters for similar atoms in the CHARMM 22 force field for proteins have been included for comparison. [b] Set 1 of AM1-SRP parameters has been used in the fitting; for other sets, the same set of van der Waals radii was used.

TABLE 9: Comparison of QM/MM and QM Energetics for the Syn Set of the Active-Site Model[a]

species[b]	B3LYP/ 6-31+G(d,p)	B3LYP/ CHARMM	AM1-SRP/ CHARMM[c]
DHAP_a	*−916.37263*	*−650.47769*	0.11670
EDT1_a	17.6	20.1	18.3 (11.3)
EDT2_a	25.3	25.0	26.4 (26.2)
EDT1_2_a	20.0	22.2	20.7 (16.9)
EDL2_a	18.6	17.3	17.9 (6.1)
TSB_a	31.5	30.4	33.5 (54.0)
TSC1_a	26.1	25.9	27.1 (36.1)
EDT2_np_ a	*−916.34198*	*−650.44637*	*−1.43286*
GAP_a	−10.1	−11.9	−10.1 (−7.0)
EDT1_a	*−916.34454*	*−647.56106*	*−2.31278*
EDL1_a	7.1	5.1	4.8 (9.5)
EDT2_a	7.7	6.8	9.2 (11.5)
RMS error[d]		1.5	1.3 (9.7/3.5[e])

[a] Single-point energetics at HF/3-21+G optimized structures. The total energies for reference structures are given in hartrees in italic type; the relative energetics are given in kilocalories per mole. For structures from DHAP_a to GAP_a, the QM part includes the substrate and the model Glu 165. The model His95 is treated with the CHARMM22 force field. For the last three structures, the substrate and model His95 are treated with QM, and the model Glu 165 is described with the CHARMM22 force field. [b] Label b indicates the basic model. [c] The numbers in parentheses are the AM1/CHARMM single-point energetics. [d] RMS difference compared to the full B3LYP/6-31+G(d,p) energetics. [e] The value after the slash is the RMS difference (in kilocalories per mole) between AM1/CHARMM and full AM1 single-point energies at AM1 optimized geometries.

Appendix 2: Tests of Calculational Methods

A2.1. Minimal Model (see Table 5): Comparison of B3LYP with MP2 and CCSD. The optimized structures are

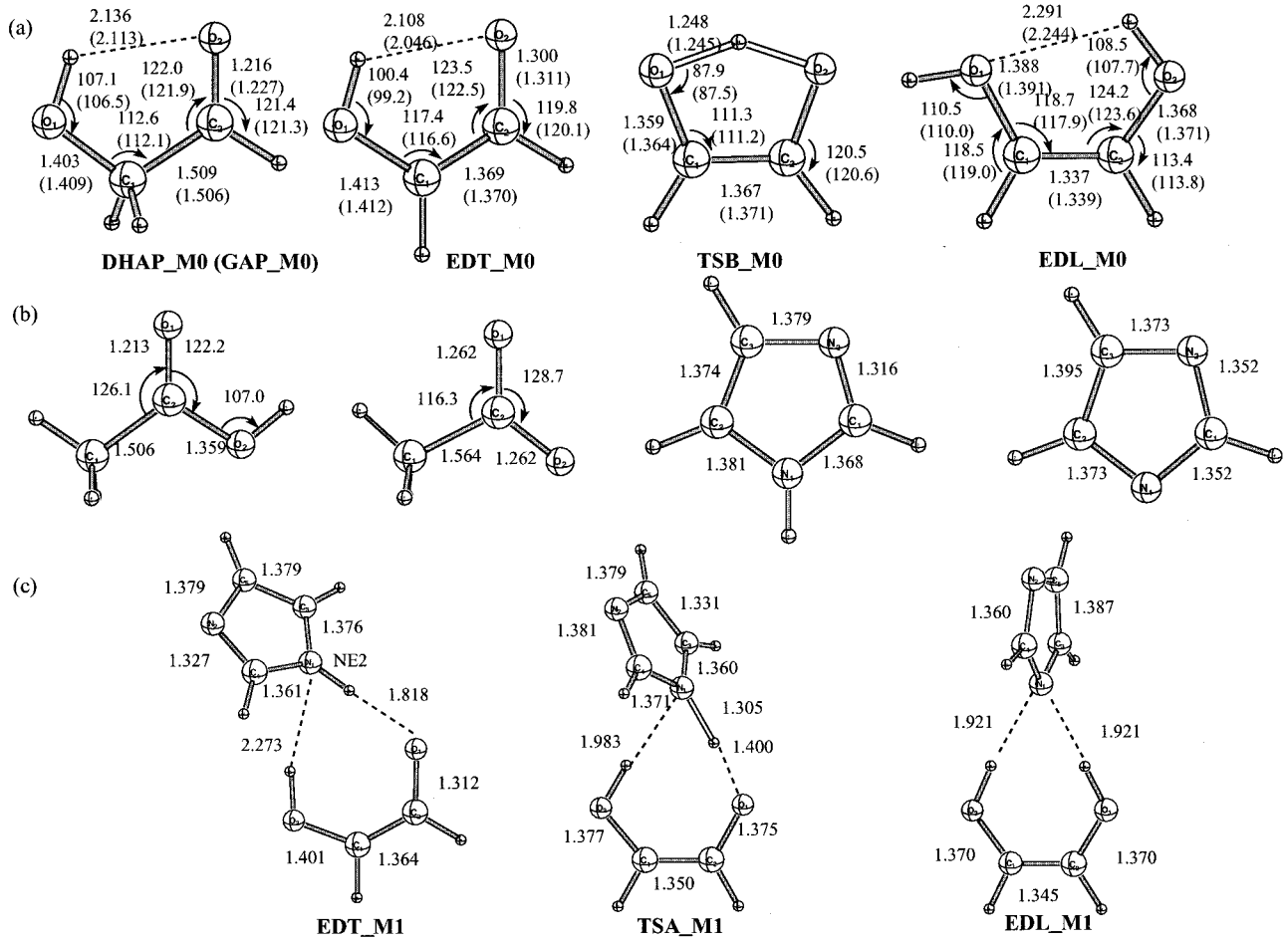


Figure 23. (a) Optimized structures of substrate at different stages of reaction, plus the transition state for the intramolecular proton transfer in the enediolate. The numbers without parentheses are obtained with B3LYP/6-31+G(d,p), and those in parentheses are optimized with MP2/6-31+G-(d,p). (b) Optimized structures at the B3LYP/6-31+G(d,p) level for the models of the side chains of Glu 165 and His 95 in different protonation states. (c) Optimized reactant, transition state, and product for the proton transfer between imidazole and model enediolate. The Nϵ2 atom has been fixed during the optimization according to the position of His 95 in the X-ray structure.

TABLE 10: Comparison of B3LYP, MP2, and CCSD Results for Simple Models of the Substrate and Catalytic Residues[a]

species	B3LYP/ 6-31+G(d,p)	MP2/ 6-311+G(d,p)[b]	CCSD/ 6-311+G(d,p)[b]
DHAP_M0	*−229.06152*	*−228.52206*	*−228.54424*
EDT_M0	366.3	368.9	374.8
TSB_M0	374.8 (8.5)	376.1 (7.2)	384.9 (10.1)
EDL_M0	8.3	9.4	9.4
CH3COOH	*−229.10583*	*−228.56784*	*−228.58734*
$CH3COO^-$	352.9 [13.4]	355.1 [13.8]	359.4 [15.4]
imidazole	*−226.23539*	*−225.63919*	*−225.65714*
imidazolate	356.7 [1.4]	353.4 [6.2]	359.2 [6.3]
EDT_M1	*−454.74383*	*−453.80959*	
TSA_M1	3.2	3.0	
EDL_M1	−11.1	−13.4	

[a] The total energies for reference structures are given in hartrees in italic type. The numbers in parentheses are the barrier height for intramolecular proton transfer measured from EDT_M0. The numbers in brackets are exothermicities for the proton exchange reaction: DHAP_M0/EDL_M0 + CH_3COO^-/imidazolate → EDT_M0 + CH_3COOH/imidazole. [b] Calculated at the B3LYP/6-31+G(d,p) optimized structures. No zero-point correction (ZPC) has been added in the relative energetics, which applies for all the following tables.

shown in Figure 23, and the energetics are given in Table 10. The geometries from the B3LYP and MP2 calculations are very similar, with the largest differences about 0.05 Å in bond lengths and 1.2° in bond angles. The B3LYP/6-31+(d,p) method gives good relative energetics compared to MP2/6-311+G(d,p) and CCSD/6-311+G(d,p) results for most cases. For instance, the EDL_M_0 is calculated to be 8.3, 9.4, and 9.4 kcal/mol, respectively, higher than DHAP_M_0 at the B3LYP/6-31+G-(d,p), MP2/6-311+G(d,p), and CCSD/6-311+G(d,p) levels. The barrier heights for the intramolecular proton transfer in EDT_M_0 are calculated to be 8.5, 7.2, and 10.1 kcal/mol, respectively, at the three levels. There are larger discrepancies in the calculated proton affinities. For instance, the proton affinities of DHAP are 366.3, 368.9, and 374.8 kcal/mol at the B3LYP/6-31+G-(d,p), MP2/6-311+G(d,p), and CCSD/6-311+G(d,p) levels, respectively. Because of some cancellation of errors, the agreement between different methods becomes better if a reaction is considered. As listed in Table 10, the exothermicity values for the proton exchange processes DHAP_M_0 + CH_3COO^- → EDT_M_0 + CH_3COOH are 13.4, 13.8, and 15.4 kcal/mol at the B3LYP/6-31+G(d,p), MP2/6-311+G(d,p), and CCSD/6-311+G(d,p) levels, respectively. For the reaction EDL_M_0 + imidazolate → EDT_M_0 + imidazole, the values are 1.4, 6.2, and 6.3 kcal/mol at the three levels. Therefore, it appears that B3LYP/6-31+G(d,p) has relatively large error for the proton exchange between imidazole and EDL_M_0. However, error cancellation has a favorable effect when the two reactants are put closer, as in Figure 23c. In this case, the B3LYP/6-31+G(d,p) calculations give 3.2 kcal/mol and −11.1 kcal/mol for the proton-transfer barrier height and exothermicity, which are in good agreement with the MP2/6-311+G(d,p) results, 3.0 kcal/mol and −13.4 kcal/mol, respectively.

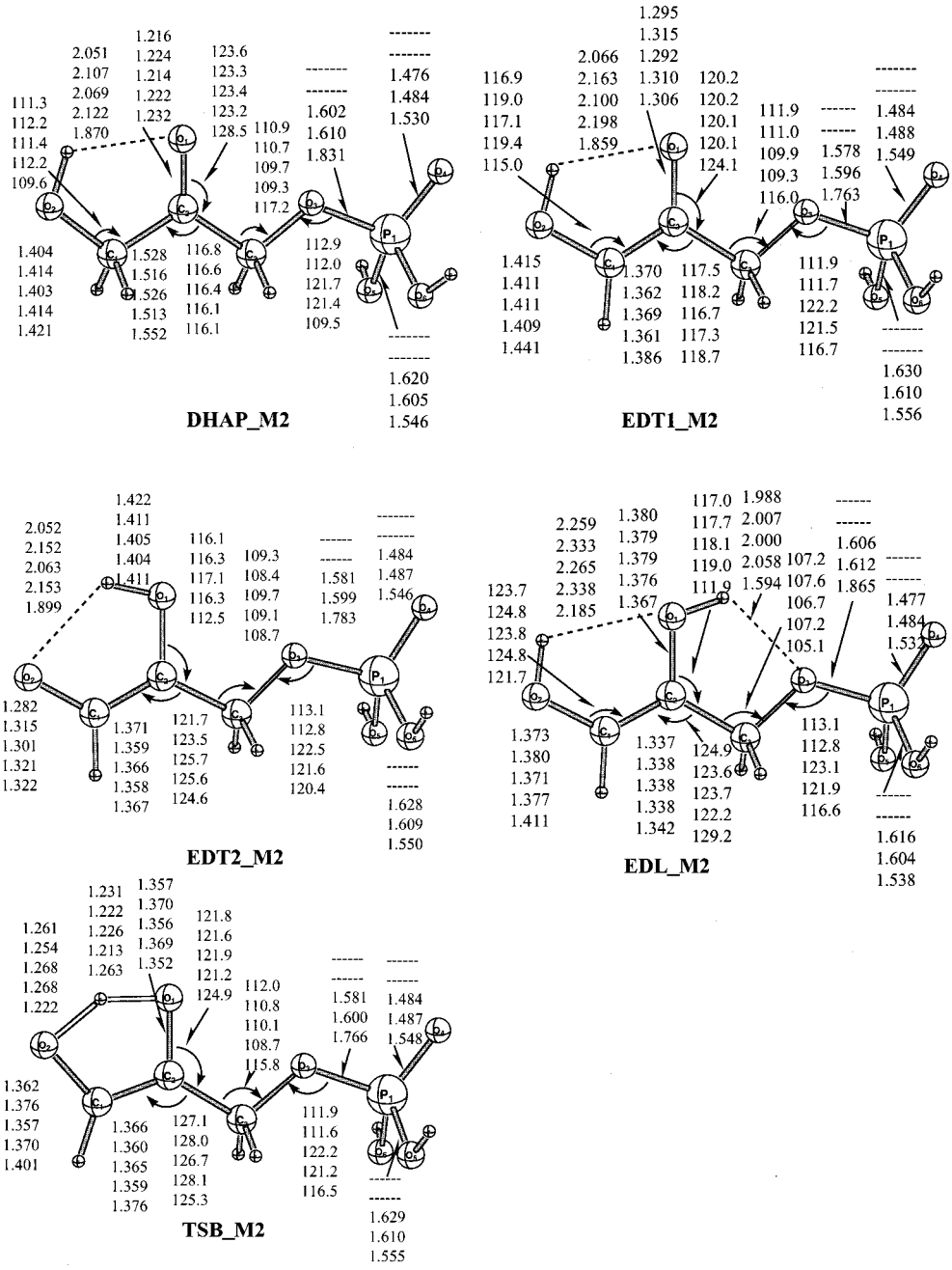


Figure 24. Optimized structures of substrate at different stages of the reaction with different terminal groups in the gas phase and in solution. The numbers in the first and second rows are obtained with the -OCH_3 terminal group in the gas phase at the B3LYP/6-31+G(d,p) level and in solution at the B3LYP-COSMO/6-31+G(d,p) level, respectively. The numbers in the third and fourth rows are obtained with the -OPO_3H_2 terminal group in the gas phase at the B3LYP/6-31+G(d,p) level and in solution at the B3-PB/6-31+G(d,p) level, respectively. The fifth row is obtained with the -PO_3^{2-} terminal group in the gas phase at the B3LYP/6-31+G(d,p) level and in solution at the B3−PB/6-31+G(d,p) level, respectively.

The above results demonstrate that the B3LYP/6-31+G(d,p) method gives reliable geometries and sufficiently accurate energetics for model systems, indicating that this level of calculations is a satisfactory choice for present studies. However, there are some quantitative deviations between the B3LYP/6-31+G(d,p) and the most accurate methods used in the present tests, suggesting that care in interpreting the results is required.

A2.2. Effect of the Phosphate Group in the Substrate. The optimized geometries are shown in Figure 24, and the energetics are summarized in Table 11 for the systems of interest. First, we comment on the different methods of treating solvation effects. From Table 11, we see that COSMO gives results very similar to those from PCM and FD-PB, if the same set of radii is used. For instance, the EDL_M_{2M} and EDT1_M_{2M} are 12.4 and 319.6 kcal/mol, respectively, above DHAP_M_{2M} at the B3LYP-COSMO/6-31+G(d,p) level, in close agreement with the B3LYP-PCM/6-31+G(d,p) results, which are 12.0 and 319.3 kcal/mol, respectively. This is expected since the conductorlike boundary condition used in COSMO is a good approximation for water, which has a dielectric constant of 80. The effect of atomic radii on the energetics is significant when the charge of the molecule changes. For instance, the relative energetics of EDT1_M_{2M} and DHAP_M_{2M} are calculated for the gas-phase geometry to be 319.9 and 311.9 kcal/mol, with the UAHF and Pauling set of atomic radii, respectively. However, the effect of radii is much smaller on the relative energetics of molecules with similar charges. For instance, the energy of EDT2_M_{2M} and TSB_M_{2M}, relative to EDT1_M_{2M}, calculated at the gas-

TABLE 11: Comparison of Different Models for the Phosphate Group in the Substrate in the Gas Phase and in Water[a]

	X = M			X = NP			X = P	
species[b]	B3LYP	B3COSMO[c]	B3PCM[d]	B3LYP	B3COSMO	B3PB[e]	B3LYP	B3PB
DHAP_M2X	*−382.90517*	*−382.92714*	*−382.92601*	*−911.33651*	*−911.37205*	*−911.37774*	*−910.11414*	*−910.48485*
EDL_M2X	8.3	12.4 (12.4/11.6)	12.0	8.8	12.2 (12.5)	13.5	4.1	13.7 (14.6)
EDT1_M2X	371.9	319.6 (319.9/311.9)	319.3	362.5	317.3 (317.2)	308.4	510.4	313.4 (310.1)
EDT2_M2X	373.8	321.6 (322.0/314.1)	322.2	365.4	319.8 (319.7)	310.5	501.4	314.2 (310.1)
TSB_M2X	381.0	328.4 (328.5/320.4)	328.0	372.5	326.6 (326.4)	318.1	515.0	321.2 (317.4)
ΔE_EDT12	1.9	2.0 (2.1/2.2)	2.9	2.9	2.5 (2.5)	2.1	9.0	0.8 (0.0)
E_a	9.1	8.8 (8.6/8.5)	8.7	10.0	9.4	9.7	4.5	7.8 (7.3)

[a] The total energies of DHAP_M2X are given in hartrees in italic type. Relative energies are given in kilocalories per mole. The numbers in parentheses are obtained by single-point calculations at the gas-phase geometries. The basis set used is 6-31+G(d,p). [b] The quantity ΔE_EDT1−2 is the energy difference between EDT1_M2X and EDT2_M2X. The quantity E_a is the barrier height for the intramolecular proton-transfer process between EDT1_M2X and EDT2_M2X, given by the difference between TSB_M2X and EDT1_M2X. [c] The values after the slashes are obtained with the Pauling radii, and the others are obtained with the UAHF radii. [d] The PCM calculations are carried out at the COSMO optimized geometries. [e] In the PB calculations, the Pauling set of atomic radii are used by the Jaguar program.

phase geometries with the Pauling radii is 2.1 and 8.6 kcal/mol, respectively, which are similar to the UAHF results of 2.2 and 8.5 kcal/mol, respectively. The same observation applies to the results for M_{2NP} with COSMO and PB, which use the UAHF and Pauling radii, respectively. The proton affinities of DHAP_M_{2NP} from the two calculations differ by 8.9 kcal/mol, but the energy of EDL_M_{2NP}, measured from DHAP_M_{2NP}, and the energies of EDT2_M_{2NP} and TSB_M_{2NP} relative to EDT1_M_{2NP} from the two sets of calculations are very similar (see Table 11).

Next, we discuss the effect of the terminal phosphate group and solvation. In the gas phase, M_{2M} and M_{2NP} are rather similar in terms of both geometries and energetics (see Table 11 and Figure 24). However, they are different from M_{2P} with the charged phosphate group. For instance, the intramolecular hydrogen bonds are shorter by more than 0.2 Å in M_{2P}, and the phosphate−ester oxygen distance is much longer in M_{2P} than that in M_{2NP}. The proton affinity of DHAP_M_{2P} is almost 150 kcal/mol greater than those of DHAP_M_{2M} and DHAP_M_{2NP}. The EDT2 species is lower in energy than the EDT1 form by 9.0 kcal/mol in M_{2P}, but the former is higher in energy by 1.9 and 2.9 kcal/mol in M_{2M} and M_{2NP}, respectively. Upon solvation, the geometries of M_{2M} and M_{2NP} do not undergo large changes, except for the hydrogen-bond distances, which become a bit longer (by 0.2−0.3 Å) in solution due to the solvent screening effect. This is also evident from the similarity of the COSMO energetics obtained with the gas-phase and solution geometries; the results are shown in Table 11 with and without parentheses, respectively. The largest difference between the two sets of results is 0.4 kcal/mol. Not surprisingly, the structures with the charged phosphate group undergo larger changes upon solvation. The largest changes in geometry take place in the phosphate group, with the phosphate−ester oxygen distance decreasing by nearly 0.1 Å from the gas-phase value. The intramolecular hydrogen-bond distances also become much closer to those in the neutral models. Upon solvation, the energetics for M_{2P} structures become closer to the results for M_{2M} and M_{2NP}. For instance, the proton affinity of DHAP_M_{2P} in solution is 313.4 kcal/mol at the B3-PB/6-31+G(d,p) level, which is rather close to the value of 308.4 kcal/mol for DHAP_M_{2NP}. The energy difference between EDT2_M_{2P} and EDT1_M_{2P} decreases from −9.0 kcal/mol in the gas phase to nearly 0.0 kcal/mol in solution, which is very close to the values of 2.9 and 2.1 kcal/mol, for M_{2M} and M_{2NP}, respectively.

These comparisons confirm that meaningful gas-phase results for the reactive species can be obtained by a neutral group such as -OCH_3, which has little effect on the relative energetics of various forms of the substrate. A number of solvation models have also been tested, and it was found that they give consistent results. The effect of solvation is small when a neutral terminal group is used. Therefore, in the model calculations, we were able to use -OCH_3 as the terminal group and introduce the solvation effect by carrying out single-point calculations at gas-phase optimized geometries.

TABLE 12: Comparison of Energetics at Different Levels for the Basic Model[a]

species[b]	B3LYP/ 6-31+G(d,p)	MP2/6-31+G(d,p)// B3LYP/ 6-31+G(d,p)	B3LYP/ 6-31+G(d,p)// HF/3-21+G
DHAP_b	*−650.79608*	*−648.94022*	*−650.79100*
TS1_b	*c*	*d*	*c*
EDT1_b	24.4	27.2	24.4
TSB_b	29.2	30.0	30.3
EDT1_2_b	20.6	24.0	*d*
TSC1_b	25.0	28.6	*d*
EDL2_b	11.1	13.4	*d*
TSC2_b	*e*	*d*	*d*
EDT2_b	19.2	20.8	20.2
EDT1_3_b	*−647.90798*	*−646.02136*	*−647.90185*
TSA_b	1.9	1.6	1.6
EDL1_b	−0.2	−1.5	−0.1
TSA2_b	*f*	*d*	*f*
TSB_b	13.0	11.5	13.3
EDT2_2_b	10.8	12.2	10.4
EDT2_np_b	*−650.77001*	*−648.91451*	*−650.76329*
TS2_b	0.5	0.1	0.2
GAP_b	−11.5	−15.3	−12.0
RMS difference[g]	*d*	2.2	1.0

[a] The total energies for reference structures are given in hartrees in italic type; the relative energetics are given in kilocalories per mole. [b] Label b indicates the basic model. [c] The transition state does not exist (or was very difficult to locate due to the fact that the potential energy surface is very flat); in this case all searches collapsed to structure EDT1_b. [d] The item was not calculated. [e] As in footnote *c*; in this case all searches collapsed to structure EDT2_b. [f] As in footnote *c*; in this case all searches collapsed to structure EDT2_2_b. [g] RMS difference in energies (in kilocalories per mole) compared to the B3LYP/6-31+G(d,p) results.

A2.3. Tests for Methods Used for the Basic and Active-Site Models (See Table 5). To further test the B3LYP/6-31+G-(d,p) method, we carried out MP2/6-31+G(d,p) single-point calculations for the structures in the basic model optimized at the B3LYP/6-31+G(d,p) level. The results are summarized in Table 12. To validate the B3LYP/6-31+G(d,p)//HF/3-21+G approach used for the active-site model due to its relatively large size, we carried out such calculations for a number of structures in the basic model; the results are also included in Table 12.

The MP2/6-31+G(d,p)//B3LYP/6-31+G(d,p) energetics are in general very close to the B3LYP/6-31+G(d,p) results, with

a RMS difference of 2.2 kcal/mol. The B3LYP/6-31+G(d,p)//HF/3-21+G relative energetics are also very similar, with an RMS difference of 1.0 kcal/mol with respect to the full B3LYP/6-31+G(d,p) values.

References and Notes

(1) (a) Richard, J. P. *J. Am. Chem. Soc*. **1984**, *106*, 4926. (b) Hall, A.; Knowles, J. R. *Biochemistry* **1975**, *14*, 4348.

(2) Knowles, J. R.; Albery, W. J. *Acc. Chem. Res*. **1977**, *10*, 105.

(3) For a recent review: Knowles, J. R. *Nature* **1991**, *350*, 121.

(4) (a) Albery, W. J.; Knowles, J. R. *Biochemistry* **1976**, *15*, 5588. (b) Herlihy, J. M.; Maister, S. G.; Albery, W. J.; Knowles, J. R. *Biochemistry* **1976**, *15*, 5601. (c) Albery, W. J.; Knowles, J. R. *Biochemistry* **1976**, *15*, 5627.

(5) Alston, W. C., II; Kanska, M.; Murray, C. J. *Biochemistry* **1996**, *35*, 12873.

(6) (a) Banner, D. W.; Bloomer, A. C.; Petsko, G. A.; Phillips, D.; Pogson, C. I.; Wilson, I. A. *Nature* **1975**, *255*, 609. (b) Alber, T.; Bloomer, A. C.; Petsko, G. A.; Phillips, D.; Rivers, P. S.; Wilson, I. A. *Philos. Trans. R. Soc. London B* **1981**, *293*, 159.

(7) (a) Davenport, R. C.; Bash, P. A.; Seaton, B. A.; Karplus, M.; Petsko, G. A.; Ringe, D. *Biochemistry* **1990**, *30*, 5821. (b) Lolis, E.; Alber, T.; Davenport, R. C.; Rose, D.; Hartman, F. C.; Petsko, G. A. *Biochemistry* **1990**, *29*, 6609.

(8) Browne, C. A.; Campbell, I. D.; Kiener, P. A.; Phillips, D. C.; Waley, S. G.; Wilson, I. A. *J. Mol. Biol*. **1976**, *100*, 319.

(9) Lodi, P. J.; Knowles, J. R. *Biochemistry* **1991**, *30*, 6948.

(10) (a) Harris, T. K.; Abeygunawardana, C.; Mildvan, A. S. *Biochemistry* **1997**, *36*, 14661. (b) Harris, T. K.; Cole, R. N.; Comer, F. I.; Mildvan, A. S. *Biochemistry* **1998**, *37*, 16828.

(11) (a) Belasco, J. G.; Knowles, J. R. *Biochemistry* **1980**, *19*, 472. (b) Komives, E. A.; Chang, L. C.; Lolis, E.; Tilton, R. F.; Petsko, G. A.; Knowles, J. R. *Biochemistry* **1991**, *30*, 3011.

(12) (a) Nickbarg, E. B.; Davenport, R. C.; Petsko, G. A.; Knowles, J. R. *Biochemistry* **1988**, *27*, 5948. (b) Pompliano, D. L.; Peyman, A.; Knowles, J. R. *Biochemistry* **1990**, *29*, 3186.

(13) Zhang, Z.; Komives, E. A.; Sugio, S.; Blacklow, S. C.; Narayana, N.; Xuong, N. H.; Stock, A. M.; Petsko, G. A.; Ringe, D. *Biochemistry* **1999**, *38*, 4389.

(14) (a) Alagona, G.; Desmeules, P.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc*. **1984**, *106*, 3623. (b) Alagona, G.; Ghio, C.; Kollman, P. A. *J. Mol. Biol*. **1986**, *191*, 23. (c) Alagona, G.; Ghio, C.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 9855.

(15) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826.

(16) Åqvist, J.; Fothergill, M. *J. Biol. Chem.* **1996**, *271*, 10010.

(17) Neria, E.; Karplus, M. *Chem. Phys. Lett.* **1997**, *267*, 23.

(18) (a) Peräkylä, M. *J. Chem. Soc., Perkin Trans*. **1997**, 2185. (b) Peräkylä, M.; Pakkanen, T. A. *Proteins: Struct., Funct., Genet.* **1996**, *25*, 225.

(19) Karplus, M.; Evanseck, J. D.; Joseph, D.; Bash, P.; Field, M. J. *Faraday Discuss*. **1992**, *93*, 1.

(20) Joseph-McCarthy, D.; Petsko, G. A.; Karplus, M. *Science* **1990**, *249*, 1425.

(21) (a) Bash, P. A.; Field, M. J.; Karplus, M. *J. Am. Chem. Soc*. **1987**, *109*, 8092. (b) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem*. **1990**, *11*, 700.

(22) (a) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc*. **1977**, *99*, 4899. (b) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. A.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

(23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187.

(24) Mulholland, A. J.; Lyne, P.; Karplus, M. *J. Am. Chem. Soc*. **2000**, *122*, 534.

(25) Joseph-McCarthy, D.; Lolis, E.; Komives, E. A.; Petsko, G. A. *Biochemistry* **1994**, *33*, 2815.

(26) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523.

(27) Joseph-McCarthy, D.; Rost, L. E.; Komives, E. A.; Petsko, G. A. *Biochemistry* **1994**, *33*, 2824.

(28) Blacklow, S. C.; Knowles, J. R. *Biochemistry* **1990**, *29*, 4099.

(29) Borchert, T. V.; Abagyan, R.; Jaenicke, R.; Wierenga, R. K. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1515.

(30) See, for example, (a) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem*. **1998**, *273*, 25529. (b) Gerlt, J. A.; Gassman, P. G. *Biochemistry* **1993**, *32*, 11934.

(31) (a) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitza, K. B., Boyd, D. B., Eds.; VCH: New York, 1996; Vol. 7, p 119. (b) Sighn, U. C.; Kollman, P. A. *J. Comput. Chem*. **1996**, *7*, 718. (c) Field, M. J. In *Computer Simulations of Biomolecular Systems: Theoretical and Experimental Applications*; van Gunsteren, W. F., Winer, P. K., Wilkinson, A. J., Eds.; ESCOM: Leiden, The Netherlands, 1993; Vol. 2, p 82. (d) Warshel, A.; Karplus, M. *J. Am. Chem. Soc*. **1972**, *94*, 5612. (e) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem*. **1990**, *11*, 700. (f) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys*. **1996**, *105*, 1968.

(32) (a) Maseras, F.; Morokuma, K. *J. Comput. Chem*. **1995**, *16*, 1170. (b) Matsubara, T.; Maseras, F.; Koga, N.; Morokuma, K. *J. Phys. Chem*. **1996**, *100*, 2573. (c) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys*. **1996**, *105*, 1959. (d) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys*. **1996**, *105*, 3654.

(33) (a) Cui, Q.; Karplus, M. *J. Am. Chem. Soc*. **2001**, *123*, 2284. (b) Q. Cui and M. Karplus, in press.

(34) For discussion based on recent experimental results, see for example (a) Alston, W. C., II; Klanska, M.; Murray, C. J. *Biochemistry* **1996**, *35*, 12873. (b) Bahnson, B. J.; Colby, T. D.; Chin, J. K.; Goldstein, B. M.; Klinman, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12797. (c) Kohen, A.; Jonsson, T.; Klinman, J. P. *Biochemistry* **1997**, *36*, 2603.

(35) For discussion on biological system based on theoretical results, see for example (a) Hwang, J. K.; Warshel, A. *J. Phys. Chem*. **1996**, *118*, 11745. (b) Hwang, A. J. K.; Warshel, A. *J. Phys. Chem*. **1993**, *97*, 10053.

(36) (a) Becke, A. D. *Phys. Rev. A* **1998**, *38*, 3098. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. (c) Becke, A. D. *J. Chem. Phys*. **1993**, *98*, 5648.

(37) (a) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *J. Chem. Phys*. **1993**, *98*, 5612. (b) *Chemical Applications of Density Functional Theory*; Laird, B. B., Ross, R. B., Ziegler, T., Eds.; ACS Symposium Series 629; American Chemical Society: Washington, DC, 1996.

(38) (a) Handy, N. C.; Boese, A. D. *Phys. Chem. Chem. Phys*. **1999**, *1*, 3939. (b) Tozer, D. J.; Handy, N. C. *J. Chem. Phys*. **1998**, *109*, 10180. (c) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Chem. Phys*. **1998**, *109*, 6264.

(39) (a) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys*. **1971**, *54*, 724. (b) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys*. **1972**, *56*, 2257. (c) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta*. **1973**, *28*, 213.

(40) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(41) Bartlett, R. J. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995.

(42) Straus, D.; Raines, R.; Kawashima, E.; Knowles, J. R.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2272.

(43) Colvin, M. E.; Evleth, E.; Akacem, Y. *J. Am. Chem. Soc*. **1995**, *117*, 4357.

(44) (a) Klamt, A.; Schüürman, G. *J. Chem. Soc., Perkin Trans*. **1993**, *2*, 799. (b) Barone, V.; Cossi, M. *J. Phys. Chem*. **1998**, *102*, 1995.

(45) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys*. **1981**, *55*, 117. (b) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett*. **1996**, *255*, 327. (c) Froesman, J. B.; Keith, T. A.; Weinberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem*. **1996**, *100*, 16098.

(46) Jaguar 3.0, Schrödinger, Inc., 1997.

(47) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys*. **1997**, *107*, 3210.

(48) Tadayoni, B. M.; Parris, K.; Rebek, J. *J. Am. Chem. Soc*. **1989**, *111*, 4503.

(49) Crammer, K. D.; Zimmerman, S. C. *J. Am. Chem. Soc*. **1990**, *112*, 3680.

(50) Guo, H.; Salahub, D. R. *Angew. Chem*. **1998**, *37*, 2985.

(51) Q. Cui, H. Guo, and M. Karplus, manuscript in preparation.

(52) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian 94, Revision B.3; Gaussian, Inc.: Pittsburgh, PA, 1995.

(53) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem.* **1998**, *102*, 3586.

(54) (a) Brooks, C. L., III; Karplus, M. *J. Mol. Biol*. **1989**, *208*, 159. (b) Brooks, C. L., III; Brünger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843. (c) Brooks, C. L., III; Karplus, M. *J. Chem. Phys*. **1983**, *79*, 6312.

(55) Lolis, E.; Petsko, G. A. *Biochemistry* **1990**, *29*, 6019.

(56) Simonson, T.; Archontis, G.; Karplus, M. *J. Phys. Chem*. **1997**, *101*, 8349.

(57) (a) Bashford, D.; Karplus, M. *Biochemistry* **1990**, *29*, 10219. (b) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144.

(58) (a) Briggs, J. M.; Davis, M. E.; Desai, B. H.; Gilson, M. K.; Ilin, A.; Luty, B. A.; McCammon, J. A.; Madura, J. D.; Tan, R. C.; Wade, R. C. University of Houston Brownian Dynamics Program; Department of

Chemistry, University of Houston: Houston, TX, 1993. (b) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Commun.* **1991**, *62*, 187.

(59) Karplus, M.; Shakhnovich, E. Protein Folding: Theoretical Studies of Thermodynamics and Dynamics. In *Protein Folding*; Creighton, T., Ed.; W. H. Freeman and Company: New York, 1992; p 127.

(60) Fischer, S.; Karplus, M. *Chem. Phys. Lett*. **1992**, *194*, 252.

(61) McDermott, A.; Ridenour, C. F. In *Encyclopedia of NMR*; John Wiley & Sons Ltd.: Chichester, England, 1996; p 3820.

(62) (a) Morokuma, K.; Kitaura, K. *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Politzer, P., Truhlar, D. G., Eds.; Plenum: New York, 1981. (b) Morokuma, K. *Acc. Chem. Res.* **1977**, *10*, 294.

(63) See for example (a) Perutz, M. F. *Proc. R. Soc. London B* **1967**, *167*, 448. (b) Vernon, C. A. *Proc. R. Soc. London B* **1967**, *167*, 389.

(64) Simonson, T.; Brooks, C. L., III *J. Am. Chem. Soc*. **1996**, *118*, 8452.

(65) (a) Simonson, T. *J. Am. Chem. Soc*. **1998**, *120*, 4875. (b) Simonson, T.; Perahia, D. *J. Am. Chem. Soc*. **1995**, *117*, 7987.

(66) Fersht, A. *Structure and mechanism in protein science*; W. H. Freeman and Company: New York, 1999.

(67) (a) Cleland, W. W.; Frey, P. A.; Gerlt, J. A. *J. Biol. Chem*. **1998**, *273*, 25529. (b) Lin, J.; Westler, W. M.; Cleland, W. W.; Markley, J. L.; Frey, F. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14664. (c) Cleland, W. W.; Kreevoy, M. M. *Science* **1994**, *264*, 1887. (d) Frey, P. A.; Whitt, S.; Tobin, J. *Science* **1994**, *264*, 1927.

(68) (a) Warshel, A.; Florián, J. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5950. (b) Warshel, A. *J. Biol. Chem*. **1998**, *273*, 27035.

(69) Cannon, W. R.; Benkovic, S. J. *J. Biol. Chem.* **1998**, *273*, 26257.

(70) Q. Cui and M. Karplus, unpublished results.

(71) Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2000**, *104*, 3721.

(72) Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5250.

(73) (a) Cohen, S. G.; Vaidya, V. M.; Schultz, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *66*, 249. (b) Crosby, J.; Stone, R.; Lienhard, G. E. *J. Am. Chem. Soc*. **1970**, *92*, 2891.

(74) Fife, T. H.; Jaffe, S. H.; Natarajan, R. *J. Am. Chem. Soc*. **1991**, *113*, 7646.

(75) (a) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T. Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260. (b) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569.

(76) (a) White, C. A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M. *Chem. Phys. Lett.* **1994**, *230*, 8. (b) Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996**, *271*, 51. (c) Lee, T.-S.; York, D. M.; Yang, W. *J. Chem. Phys*. **1996**, *105*, 2744.

(77) (a) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett*. **1995**, *233*, 231. (b) Gonzalez-Lafont, A.; Troung, T. N.; Truhlar, D. G. *J. Phys. Chem*. **1991**, *95*, 4618. (c) Bash, P.; Ho, L. L.; MacKerell, A. D., Jr.; Levine, D.; Hallstrom, P. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 3698.

(78) Murrel, J. N. *Molecular Potential Energy Functions*; Wiley: New York, 1984.

(79) Goldberg, D. E. *Genetic algorithms in search, optimization and machine learning*; Addison-Wesley: Reading, MA, 1989.

(80) Ho, L. L.; MacKerell, A. D., Jr.; Bash, P. A. *J. Phys. Chem*. **1996**, *100*, 4466.

(81) York, D.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 11060.

(82) (a) Yagiol, G. *Tetrahedron* **1967**, *23*, 2855 (b) Chiang, Y.; Kresge, A. J. *Science* **1991**, *253*, 395.

(83) For a recent discussion, see, for example, Guo, H.; Sirois, S.; Proynov, E. I.; Salahub, D. R. In *Theoretical Treatments of Hydrogen Bonding*; Hadzi, D., Ed.; John Wiley and Sons Ltd.: Chichester, England, 1997; p 49.

(84) (a) For a recent report, see Wilson, E. K. *Chem. Eng. News* **2000**, *78*, 42. (b) Kohen, A.; Cannio, R.; Bartolucci, S. Klinman, J. *Nature* **2000**, *399*, 496. (c) Bruice, T. C.; Benkovic, S. J. *Biochemistry* **2000**, *39*, 6267.

(85) Lennartz, C.; Schäfer, A.; Terstegen, F.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 1758.