JCTC Journal of Chemical Theory and Computation

Catalytic Mechanism of Diaminopimelate Epimerase: A QM/MM Investigation

Marco Stenta,*,† Matteo Calvaresi,† Piero Altoè,† Domenico Spinelli,‡
Marco Garavelli,† Roberta Galeazzi,§ and Andrea Bottoni*,†

Dipartimento di Chimica "G. Ciamician", Università di Bologna, Via Selmi 2, 40126 Bologna, Italy, Dipartimento di Chimica Organica "A. Mangini", Università di Bologna, Via S. Giacomo 11, 40126 Bologna, Italy, and Dipartimento di Scienze e Tecnologie Chimiche, Università Politecnica delle Marche, via Brecce Bianche, 60131, Ancona, Italy

Abstract: A QM/MM investigation, based on a DFT(B3LYP)//Amber-ff99 potential, has been carried out to elucidate the mechanism of diaminopimelate epimerase. This enzyme catalyzes the reversible stereoconversion of one of the two stereocenters of diaminopimelate and represents a promising target for rational drug design aimed to develop new selective antibacterial therapeutic agents. The QM/MM computations show that the reaction proceeds through a highly asynchronous mechanism where the side-chain of a negatively charged Cys-73 (thiolate) deprotonates the α-carbon substrate. Simultaneously, the Cys-217 thiolic proton moves toward the same carbon atom on the opposite face, thus determining the configuration inversion. A fingerprint analysis provides a detailed description of the influence of the various residues surrounding the active site and clearly shows the electrostatic nature of the most important contributions to the catalysis.

I. Introduction

During the past decade hybrid Quantum Mechanics/Molecular Mechanics (QM/MM)[1−8] methods have been successfully used to investigate large molecular systems. The study of enzymatic reactivity[8] certainly represents a field where QM/MM methods have been most widely applied. These hybrid methods are particularly suitable to deal with this class of problems since the enzyme can be easily and almost "naturally" partitioned into two regions: one (described at the QM level) approximately corresponding to the active site and the other (described at the MM level) that includes the remaining part of the enzyme and, in some cases, the solvent molecules.

In the present paper we use a QM/MM approach to provide a complete and exhaustive analysis of an interesting enzymatic system (the diaminopimelate epimerase), and we present a general strategy for the study of similar problems involving enzymatic systems. In particular, we suggest general criteria to build a reliable model-system (using various dynamics techniques), and we show how, after having computed the potential energy surface (PES), we can use various tools to analyze the results and identify the key-factors that control the catalytic mechanism.

In section II we introduce the basic principles of the computational techniques employed in our study. These techniques concern the features of our QM/MM[1−8] approach and the type of analysis used to rationalize in details the enzyme catalytic effects. Our QM/MM code has been developed according to the hybrid approach described in a previous paper[12] and is included in the **COBRAMM**[9] suite of programs together with the analysis tools described in the following sections.

The catalytic mechanism of diaminopimelate (DAP) epimerase (E.C. 5.1.1.7)[10−12] is the case study discussed in

* Corresponding author phone: ++39-051-2099477; fax: ++39-051-2099456; e-mail: andrea.bottoni@unibo.it.

† Dipartimento di Chimica "G. Ciamician", Università di Bologna.

‡ Dipartimento di Chimica Organica "A. Mangini", Università di Bologna.
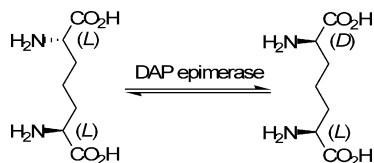
§ Università Politecnica delle Marche.

1916 *J. Chem. Theory Comput., Vol. 5, No. 7, 2009*

Stenta et al.

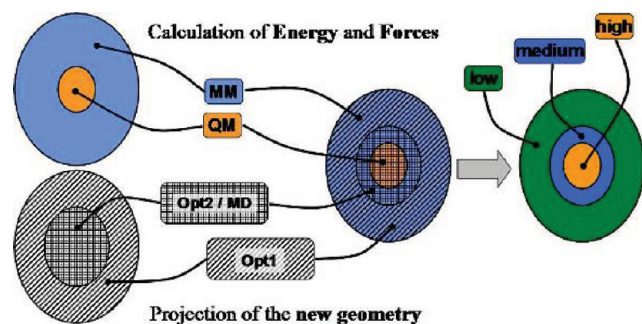**Figure 1.** The reaction catalyzed by the enzyme diaminopimelate epimerase.

Figure 2. A schematic representation of the partition scheme of the whole system. The two main regions (**MM** and **QM**) are subdivided into three layers (**H**, **M**, and **L**) when computing energy and forces. "Microiterations" (**Opt1**) and "macroiterations" (**Opt2**) refer to the optimization process applied to the **L** and **H**+**M** regions, respectively.

section III. This enzyme belongs to the group of lyases which are capable of inverting the absolute configuration of a carbon atom in substrates containing one (racemases) or more stereocenters (epimerases).[13] While many enzymes belonging to this group (for instance alanine,[14,15] serine, threonine racemases) need the presence of a molecule of pyridoxal-5′-phosphate (PLP)[16] behaving as a cofactor, diaminopimelate epimerase,[10,11] and a few other members of the family, like glutamate[17−19] and proline[20−22] racemases, can exert their catalytic action without the participation of additional molecules. The enzymes belonging to this class of racemases and epimerases (i.e., those not requiring a cofactor), despite a substantial difference in the active site structure, which is due to the specificity for different substrates, share a pair of catalytically important cysteine residues. In particular, diaminopimelate epimerase, that catalyzes the epimerization of L,L- to D,L-*meso*-diaminopimelate (see Figure 1), without assistance of PLP, operates via a "two base" mechanism involving one active-site cysteine thiolate (Cys-73) acting as a base that deprotonates the α-carbon of the substrate and a second cysteine thiol group (Cys-217) that protonates the opposite side *via* a general-acid catalysis.

To stress the importance of understanding the catalytic mechanism of DAP epimerase and its potential effects on pharmacological research, a short discussion on the involvement of this enzyme in some important metabolic pathways can be helpful. It is well-known that bacteria, plants, and fungi metabolize aspartic acid to produce four different amino acids (lysine, threonine, methionine, and isoleucine) through a sequence of reactions known as the "aspartate pathway".[23] These reactions produce several important metabolic intermediates such as diaminopimelic acid, an essential component of the bacterial cell wall biosynthesis.[24] Members of the animal kingdom do not possess this pathway and must

therefore acquire these essential amino acids through their diet. Since the enzymes[23] involved in this pathway are not present in animals, inhibitors of them are promising targets for the development of novel antibiotics, herbicides, and fungicides. The recent emergence of bacterial resistance to currently available antibiotics has determined a renewed interest in the search for novel antibacterial compounds. Since DAP epimerase is the key enzyme of the diaminopimelic acid/lysine branch of the aspartate pathway, it represents an optimal target for developing new selective drugs capable of blocking the synthesis of Gram-positive bacterial cell walls (by interrupting the lysine synthesis) or able to interfere with the building of the peptidoglycan layer of Gram-negative and mycobacterial cell walls (by stopping the D,L-diaminopimelate[25] synthesis) without interfering with host cell metabolism.

Thus, detailed information on this enzymatic mechanism (concerted or stepwise process, role of the various residues in the vicinity of the active site dyad (Cys-73 and Cys-217), electrostatic interactions controlling the substrate selectivity) is of primary importance and opens the way to important applications in drug design. In principle, a new class of antibiotics could originate from the discovery of specific inhibitors of the diaminopimelate epimerase.

In the present paper we have chosen to investigate the *Haemophilus influenzae* DAP epimerase because, for this enzyme, kinetics[10] and crystallographic[11] studies are available in literature. The kinetic[10] and the structural X-ray data[11] have been very helpful in detecting the two catalytic cysteine residues (Cys-73 and Cys-217), which are responsible for the observed pH dependency of the reaction velocity (V/K profiles in ref 10). These are characterized by pK values of 6.7 and 8.5 and must be unprotonated and protonated, respectively, to allow the L,L- to D,L-*meso*-diaminopimelate conversion (forward reaction). The protonation state must be inverted for the reverse reaction (D,L-*meso*- to L,L-diaminopimelate). The primary deuterium isotope effects suggest that substrate epimerization together with the double proton transfer strongly affect the reaction rate (rate-determining step). However, the experiments indicate that another slow step, which follows the product dissociation, could be partially rate determining. This step, which corresponds to the back-proton transfer involving, after product release, the two catalytic Cys residues, restores the original protonation state of the active site for subsequent turnovers. On the basis of this experimental evidence the experimental k_{cat} value (128 s^{-1}) for the forward (L,L → D,L) reaction and the corresponding barrier of 15.6 kcal mol^{-1} (obtained from k_{cat} by applying the Eyring equivalence[26]) can be associated, in principle, either to the substrate epimerization step or the back-proton transfer occurring in the substrate-free enzyme. Thus, this experimental value should be considered as an upper limit for the activation energy of the epimerization step.

II. Computational Details

II.A. Setting-up the System. To build a reliable model-system of the diaminopimelate (DAP) epimerase, we used the recently obtained crystal structures[11] of this enzyme (from

Catalytic Mechanism of Diaminopimelate Epimerase

J. Chem. Theory Comput., Vol. 5, No. 7, 2009 **1917**

Haemophilus influenzae) binding two isomers of the irreversible inhibitor aziridino-DAP,[27] which mimics the natural substrate [PDB[28] codes: 2GKE(L,L-AziDAP), resolution 1.35 Å and 2GKJ (D,L-AziDAP), resolution 1.70 Å]. The methylene carbon of the aziridine ring of the two diastereomeric inhibitors is covalently bonded to the sulfur atom of Cys-73 or Cys-217 after the nucleophilic attack of the sulfur on the aziridine ring that irreversibly inhibits the enzyme. The DAP epimerase backbones obtained from the two crystal structures show negligible differences.[11] The structural features of the covalent bond between the cysteine sulfur and the inhibitor (L,L-AziDAP or D,L-AziDAP) provides information to model the mechanism of approach of the thiol base during the α-deprotonation/protonation process of the DAP substrate. We decided to use the coordinates from the 2GKE PDB file to build the model-system since 2GKE has a better resolution than 2GKJ. We modified the L,L-AziDAP into the natural substrate L,L-DAP, and we retained only the A-conformer where multiple conformations of the amino acids were available in the crystal structure. The model-system obtained from these crystallographic data was protonated with the **H++**[29] software, using the default parameters available in that package. This code employs an automatic algorithm that computes pK_a values for the various ionizable groups in macromolecules and adds missing hydrogen atoms according to the specified pH value of the environment. The positions of the added hydrogen atoms are also optimized by this algorithm. The protonation state of all titratable residues was carefully checked by visual inspection of the **H++** output structure. It followed that all Asp and Glu residues were unprotonated (negatively charged), while all Lys and Arg residues were protonated (positively charged). This result is consistent with the solvent exposure of the side chain of all these titratable residues, with two important exceptions: Glu-208 and Arg-209 which, however, interact with the zwitterionic substrate by their charged side chains. Furthermore, all His residues were found to be neutral with the proton on the ϵ nitrogen atom. The only exception was the double protonated (thus, positively charged) His-50 whose side chain is exposed to the solvent but is also interacting with the hydroxyl group of the near Tyr-98 residue through the δ hydrogen of the side chain imidazole ring. Since the accurate computation of the pK values of the active site residue was beyond the purposes of our research - and adequate experimental data were available - we did not calculate the pK value of the catalytic Cys residues. The protonation state of these two residues was decided on the basis of the kinetics studies,[10] (suggesting the presence of one unprotonated Cys side chain, as outlined in the Introduction) and by inspection of the crystallographic structure of the enzyme bound to a substrate mimic,[27] indicating that the thiolate moiety should correspond to Cys-73 to allow the reaction to take place in the forward direction (L,L $\rightarrow$ D,L). Identical results were obtained using the Propka2.0 software.[30]

The L,L-DAP and D,L-*meso*-DAP molecules were parametrized using the Generalized Amber Force Field (GAFF).[31] Partial atomic charges were assigned to atoms using the AM1-BCC method[32,33] as implemented in the *antechamber* module of the AMBER8.0 package.[34]

The initial model-system geometry was fully minimized at the MM level using the *sander* module of AMBER8.0. The minimization was carried out with the Amber Force Field (Amber-*ff*99)[35] until the root-mean-square deviation (rmsd) of the Cartesian elements of the gradient was less than 0.0001 kcal mol^{-1}. A full conjugate gradient minimization approach and the General Born (GB) model[36] to simulate the aqueous environment (as implemented in the *sander* module of the AMBER8.0 code) were used.

Finally we checked if the crystal structures of 2GKE and 2GKJ mimic precisely the natural binding mode of the substrates or if the covalent bond between the enzyme and the L,L-AziDAP and D,L-AziDAP inhibitors can perturb the natural binding mode of L,L-DAP and D,L-*meso*-DAP. To this purpose we carried out a conformational study of the binding mode of L,L-DAP and D,L-*meso*-DAP within the protein environment using different methods i.e. Cluster Analysis, Simulated Annealing and Docking.

Cluster Analysis. We carried out high temperature Molecular Dynamics starting from the optimum structure obtained for the complex. A region of 5 Å around the substrate was free to move during the MD simulation. The system was heated from 0 to 800 K in 100 ps, and then a trajectory of 2 ns was computed at constant temperature (800 K). The integration step of 2 fs was used in conjunction with the SHAKE algorithm[37] to constrain the stretching of bonds involving hydrogen atoms. The coordinates of the system were saved on a trajectory file every 2 ps, giving a total of 1000 structures. Solvation effects were taken into account using the GB model with a dielectric constant of 78.5. To determine the most populated conformations of L,L-DAP and D,L-*meso*-DAP within the protein binding pocket, we performed a Cluster Analysis on the different conformations visited by the two molecules during the simulation. To this purpose we used the MMTSB toolset,[38] and we clustered different conformations of the substrates on the basis of structural similarity; we carried out our analysis employing the *kclust* module with a fixed radius of 1.0 Å on the Cartesian coordinate rmsd computed for heavy atoms. Then, we determined the centroid of each cluster. For each cluster we chose the structure closest to the corresponding centroid as representative of the cluster itself (this structure is characterized by the smallest rmsd value with respect to the centroid).

Simulated Annealing. We performed 10 cycles of simulated annealing for the two complexes formed by the protein and the substrate molecules L,L-DAP and D,L-*meso*-DAP. We used the same parameters of the previous simulation, and we heated the system from 0 to 1000 K in 30 ps, holding at 1000 K for equilibration for 10 ps. Then, we cooled from 1000 to 0 K in 60 ps. The heat bath coupling for the system was tight for heating and equilibration (0.1 ps). The cooling phase was divided in three periods: during the first 48 ps the cooling was very slow (coupling of 5.0 ps); this was followed by a cooling phase of 6 ps (with coupling of 1.0 ps) and a last cooling phase of 6 ps (with a coupling changing from 0.1 to 0.05 ps). At the end of the simulated annealing

1918 *J. Chem. Theory Comput., Vol. 5, No. 7, 2009*

Stenta et al.

a complete minimization was carried out, and the final coordinates of the complexes were retained.

Docking. The previous model obtained from the PDB structure 2GKE was used in the docking calculations after having deleted the ligand from the cavity site. The orientation sampling of the substrate into the cavity was carried out using spheres calculated by the *sphgen* module of DOCK6 [39,40] within 10.0 Å of root-mean-square deviation (rmsd) from every atom of the minimized structure of the ligand. Partial atomic charges for the substrates were obtained with the AM1-BCC method. We carried out flexible ligand docking using the Anchor-and-Grow algorithm implemented in DOCK6 with the Grid-Based Score function as primary scoring function. The results were then rescored with the new Amber score as secondary scoring function allowing a minimization of the ligand and residues within 5 Å (the same mobile residues of the previous calculations) for 100 steps. The best 10 poses obtained in the ranking of L,L-DAP and D,L-*meso*-DAP were considered for subsequent QM/MM calculations.

II.B. QM/MM Details. The QM/MM potential[9] is based on a subtractive scheme[2,9,41,42] (see Figure 1). The boundary zone between the QM and MM regions is handled by means of a hydrogen atom link approach,[2] and a charge shifting scheme within the electrostatic embedding method[2] is adopted to avoid hyper-polarization of the QM wave function. A special and particularly important feature of our QM/MM approach[9] is the partition of the system into three layers. The innermost layer called "high" (**H**) is treated at the QM level, while the outermost one, named "low" (**L**), is treated at the MM level. The presence of an intermediate layer, denoted as "medium" (**M**), improves the efficiency of the geometry optimization procedure. It has been shown that the decoupling[43,44] of the QM an MM regions during the optimization process can significantly improve the geometry convergence (faster optimization) and the accuracy of the results.[43] This approach allows a full relaxation of the MM region ("microiteration" phase[44,45] carried out with a cheap and fast optimization algorithm, like "steepest descent", indicated as **Opt1** in Figure 1) at each optimization cycle of the QM region ("macroiteration", based on an accurate algorithm, like BFGS[46−49] and denoted as **Opt2**). It is valuable to notice that, in the "macroiteration" step, the new geometry projection task can be, alternatively, carried out by a molecular dynamics code (MD), to obtain a molecular dynamics simulation on the **HM** region (**H** layer + **M** layer). During "microiterations" the **HM** region is kept frozen, and its electrostatic potential is taken into account by means of atomic point charges coming from the MM force field for **M** atoms or from the QM calculations for **H** atoms (CHELPG[50] charges have been used in this work). This approach, as pointed out by many authors and demonstrated in a previous paper,[22] gives good results in terms of the simulation cost/efficiency ratio. The obtained results are also in good agreement with more expensive methods. A possible pitfall of this approach is the transition state search procedure when it involves the simultaneous rearrangement of MM and QM atoms. This can be solved by expanding the QM subregion, but this causes, of course, an increase of the computational cost. The introduction of the intermediate (or "buffer") layer (**M**) between the QM (**H**) and MM (**L**) subregions partially overcomes this problem, because the **M** layer is treated at the MM level but is optimized together with the **H** region. This strategy allows a detailed description of large molecular motions involving several tens of atoms with a minor increase in the computational cost.

The QM/MM potential adopted throughout this work is based on DFT[51,52]/B3LYP[53] calculations using the double-ζ DZVP[54] basis set for all atoms of the QM region, while the Amber-*ff99*[35] force field has been employed for the MM atoms (GAFF[31] parameters have been adopted for the DAP substrate). In the following discussion this potential will be referred to as DFT(B3LYP/DZVP)/Amber-*ff99* potential.

The nature of the critical points on the PES can be determined by means of QM/MM numerical frequency calculations on the whole enzyme. In these computations we change only the geometry of the **HM** region in the presence of the MM potential determined by the frozen **L** region. Thus, a complete numerical frequency run (denoted as "**fullfreq**" calculation) would require a total of $1 + 6N^{HM}$ QM/MM energy evaluations, N^{HM} being the number of atoms of the **HM** region. A different and stronger level of approximation for frequency computations (simply denoted as "**freq**" calculation) has been tested. Within this approximation we hypothesize that the small motions of an MM atom have only a tiny effect on the wave function. Under this assumption it becomes possible to save $6N^M$ QM computation (N^M being the number of M atoms) by summing the current MM energy value to the reference (initial) QM energy when an MM atom is moved. Then, a new QM computation is carried out only when an atom of the **H** region is moved. In this way only $1 + 6N^H$ wave function evaluations are required. Both these approximations have been tested, and the results obtained at the two levels show a good qualitative agreement.

All the QM/MM computations described in this paper have been carried out using the general-purpose package **CO-BRAMM**,[9] which interfaces many commercially available QM and MM codes as well as some analysis routines. For the present work we used the GAUSSIAN03[55] (C02 version) and AMBER8.0[34] packages to perform QM and MM calculations, respectively. In the geometry optimization (to locate minima and saddle-points) we applied for the **HM** region ("macroiteration") the BFGS optimization algorithm implemented in the Gaussian code.[45,56,57] For the **L** region ("microiteration") we used the "steepest descent" method from the *sander* tool of AMBER8.0.

II.C. Fingerprint Analysis. In our QM/MM scheme all contributions to energy within the **H** region are computed by means of single point computations on a molecular system (denoted as *model-H*) formed by the **H** region where properly placed hydrogen atoms saturate the dangling bonds at the QM-MM boundary (according to the hydrogen atom link scheme).[2,22] Bonding and nonbonding terms of the **M** and **L** regions are computed at the MM level. A special caution is required to take into account the QM-MM cross terms, and many recipes have been proposed in literature to handle this problem. We chose a general and rather popular approach

Catalytic Mechanism of Diaminopimelate Epimerase

J. Chem. Theory Comput., Vol. 5, No. 7, 2009 **1919**

that describes all cross terms (i.e., van der Waals, bonding, bending, torsions), except the electrostatic ones, at the MM level. We adopted the electrostatic embedding scheme (EES) to describe the electrostatic contributions.[22] This consists in the computation of the QM wave function in the presence of the atomic point charges of the **M** and **L** layers. We assume that the polarization of the wave function determined by this EES computations accounts for the electrostatic cross-term interactions. Under this assumption it becomes easy to derive a procedure that splits the cross-terms energy contributions into single-residue contributions.

In this section we provide a short description of two general procedures, named Direct and Reverse Finger Print analysis (**DFP** and **RFP**, respectively), that allow to rank the electrostatic effects of the single residues and a third procedure (**vdWFP**) able to evaluate the van der Waals contributions. These analyses can provide semiquantitative information about the role of a given residue (or group of residues) in determining the relative stabilization/destabilization of two critical points. If, for instance, a transition state is compared to the nearest minimum, then we can rank the effects of each residue on the entity of the barrier for the corresponding process. It is worth remembering that, since the original publication by Karplus,[58] the **DFP** approach has been used by several others authors[59−66] (in many cases the term "charge perturbation method" has been used), and the method has been demonstrated to be a valuable approach to obtain precious information and new insight into enzyme catalysis.

To illustrate the specific features of our analysis, we consider two critical points A and B located on the QM/MM Potential Energy Surface. The overall electrostatic contribution can be easily computed as follows. QM calculations *in vacuo* on *model-H* (i.e., the QM region after hydrogen addition) provide for the two points the corresponding energy values E_0^A and E_0^B. QM calculations in the presence of all atomic point charges give the two energy values E_t^A and E_t^B. From these values, after subtraction of the corresponding charge self-energies (e_t^A and e_t^B), we obtain (eqs 1 and 2) the two quantities E_{QM}^A and E_{QM}^B (the charge self-energy corresponds to the coulomb term describing the interaction between point charges). E_{QM}^A and E_{QM}^B represent, for A and B, respectively, the sum of pure QM and electrostatic cross terms, as inserted in our QM/MM potential.

$$E_{QM}^A = E_t^A - e_t^A \qquad (1)$$

$$E_{QM}^B = E_t^B - e_t^B \qquad (2)$$

The net electrostatic effects of the MM regions on the QM wave function E_{pol}^A and E_{pol}^B can be estimated by means of eqs 3 and 4.

$$E_{pol}^A = E_{QM}^A - E_0^A \qquad (3)$$

$$E_{pol}^B = E_{QM}^B - E_0^B \qquad (4)$$

It is important to notice that the more negative these values are, the greater is the charge stabilization effect (more precisely, values smaller or greater than zero indicate that the charge polarization contribution is stabilizing or destabilizing, respectively).

A first comparison between A and B can be carried out using the terms computed with eqs 1−4. In particular, $\Delta E_0(A,B)$ and $\Delta E_{QM}(A,B)$ (eqs 5 and 6) represent the QM energy difference between A and B, in the absence and in the presence of the MM atomic point charges, respectively. Equation 7 outlines the connection between the differential stabilization effect of charges on A and B. The stability factor $S_{tot}(A,B)$ computed by eq 7 represents the magnitude of the point charge effect in promoting or discouraging the passage from A to B.

$$\Delta E_0(A, B) = E_0^B - E_0^A \qquad (5)$$

$$\Delta E_{QM}(A, B) = E_{QM}^B - E_{QM}^A \qquad (6)$$

$$S_{tot}(A, B) = \Delta E_{QM}(A, B) - \Delta E_0(A, B) = E_{QM}^B - E_{QM}^A - E_0^B + E_0^A = E_{pol}^B - E_{pol}^A \qquad (7)$$

This analysis can provide important information on the influence of the protein environment (described at the MM level) on the rate of the reaction occurring within the small QM region. An estimate, even if qualitative, of the contribution coming from each single residue can make this approach particularly useful. If A and B are, for instance, a stable species (minimum) and the near transition state (saddle point), it becomes possible to detect the residues that most significantly affect the barrier height, thus playing the main catalytic effect.

We adopted two different decomposition schemes to rank the influence of the various enzyme residues on the rate of reaction A → B.

The first scheme, denoted here as Direct Finger Print (**DFP**) analysis, requires a series of single point QM calculations (SPc) on the QM region (i.e., *model-H*) for both structures A and B (using the optimized QM/MM geometry).

After the evaluation of all terms needed to compute $S_{tot}(A,B)$ (eqs 1 to 7), we perform N SPc's (where N is the total number of residues to analyze). In each calculation *model-H* is surrounded by the atomic point charges of the **i**th residue only (see Figure S1a in the Supporting Information), with the charges placed according to the atomic coordinates of the **i**th residue itself. This procedure provides NE_i^A and NE_i^B energy values together with the corresponding charge self-energy values e_i^A and e_i^B (this term represents the pure electrostatic contribution among the point charges of the **i**th residue only). We can easily compute the electrostatic (Coulomb) effect of the **i**th residue on the QM region ($E_{pol,i}^A$ and $E_{pol,i}^B$ from eqs 8 and 9) using eqs 3 and 4 (proposed for the total electrostatic effect). This effect is stabilizing or destabilizing if the corresponding value is lesser or greater than zero, respectively.

$$E_{pol,i}^A = (E_i^A - e_i^A) - E_0^A \qquad (8)$$

$$E_{pol,i}^B = (E_i^B - e_i^B) - E_0^B \qquad (9)$$

We obtain the stability parameter S_i (eq 10) for the **i**th residue by comparing the values of $E_{pol,i}^A$ and $E_{pol,i}^B$. If $S_i < 0$,

then the ith residue favors the transition from A to B. On the contrary, if $S_i > 0$ the ith residue slows down the process.

$$S_i = E^B_{pol,i} - E^A_{pol,i} \tag{10}$$

A further concern is important to establish the accuracy of our method in decomposing the total electrostatic effect into smaller components. An ideal decomposition scheme should provide an exact equivalence between the total electrostatic effects E^A_{pol} and E^B_{pol} and the terms $E^A_{polsumm}$ and $E^B_{polsumm}$ obtained by summing up all single contributions E^A_{pol} and E^B_{pol} values) as stated in eqs 11 and 12.

$$E^A_{polsumm} = \sum_{i=1}^{N} E^A_{pol,i} \tag{11}$$

$$E^B_{polsumm} = \sum_{i=1}^{N} E^B_{pol,i} \tag{12}$$

It is evident that, under the adopted approximations, the terms $E^A_{polsumm}$ and $E^B_{polsumm}$ are not perfectly equivalent to E^A_{pol} and E^B_{pol}. To estimate the error we introduce the following three error parameters.

$$X^A_{add,S} = E^A_{pol} - \sum_{i=1}^{N} E^A_{pol,i} \tag{13}$$

$$X^B_{add,S} = E^B_{pol} - \sum_{i=1}^{N} E^B_{pol,i} \tag{14}$$

$$X_{add,S}(A, B) = X^B_{add,S} - X^B_{add,S} \tag{15}$$

If the errors made on the two critical points A and B ($X_{add,S}{}^A$ and $X_{add,S}{}^B$) differ significantly, $X_{add,S}(A,B)$ is an important evaluation of the reliability of the adopted decomposition scheme.

An immediate outlook of the whole enzyme influence on the A → B transformation can be obtained by plotting each S_i value against the corresponding residue number. The relative magnitude of the stability parameter S_i can be used to rank the residues according to their importance in the catalytic process. The results of the DFP analysis can be affected by the basic approximations of our approach: the S_i factor is estimated by comparing the unperturbed *in vacuo* system (to obtain E^A_{pol} and E^B_{pol}) and the system perturbed by a single-residue (to obtain $E^A_{pol,i}$ and $E^B_{pol,i}$). However, in principle, the latter situation can be rather different, in terms of wave function polarization, with respect to the QM region fully embedded into the MM point charge cloud (according to the electrostatic embedding scheme).

An opposite scheme, referred to here as Reverse Finger Print (**RFP**) analysis, can be used to improve the previous description. We compute again the E^A_{QM}, E^B_{QM}, E^A_0, and E^B_0 terms by means of single point calculations (SPc) on the *model-H* system in the presence and absence of the whole set of atomic point charges. We define from these calculations the two terms E^A_{dest} and E^B_{dest} (eqs 16 and 17).

$$E^A_{dest} = -E^A_{pol} = E^A_0 - E^A_{QM} \tag{16}$$

$$E^B_{dest} = -E^B_{pol} = E^B_0 - E^B_{QM} \tag{17}$$

We perform N SP computations (N is the total number of residues to analyze) where the system *model-H* is surrounded by all atomic point charges except those corresponding to the ith residue (see Figure S1b in the Supporting Information). Charges are placed, as previously described for DFP, according to the QM/MM optimized geometry. This procedure provides N $E^A_{h,i}$ and N $E^B_{h,i}$ energy values and the corresponding charge self-energies $e^A_{h,i}$ and $e^B_{h,i}$, which represent the pure electrostatic contributions of all point charges except those corresponding to the "hole" of the missed i^{th} residue. We can now compute the electrostatic effects $E^A_{dest,i}$ and $E^B_{dest,i}$ on the QM region, which are due to the absence of the ith residue (eqs 18 and 19). These terms represent a destabilization or stabilization if they are lesser or greater than zero, respectively.

$$E^A_{dest,i} = (E^A_{h,i} - e^A_i) - E^A_{QM} \tag{18}$$

$$E^B_{dest,i} = (E^B_{h,i} - e^B_i) - E^B_{QM} \tag{19}$$

A destabilization parameter D_i (eq 20) can be obtained for each residue i^{th}. $D_i < 0$ or $D_i > 0$ indicate that the ith residue has the effect of reducing or enhancing the rate of the A → B transformation, respectively.

$$D_i = E^B_{dest,i} - E^A_{dest,i} \tag{20}$$

Again, as previously outlined for the **DFP** analysis, the summation over the single contributions E^A_{dest} and E^B_{dest} does not perfectly correspond to the total electrostatic effects $E^A_{dest,i}$ and $E^B_{dest,i}$. In other words eqs 21 and 22 do not hold.

$$E^A_{destsumm} = \sum_{i=1}^{N} E^A_{dest,i} \tag{21}$$

$$E^B_{destsumm} = \sum_{i=1}^{N} E^B_{dest,i} \tag{22}$$

The error can again be estimated by the following three error parameters (see eqs 23, 24, and 25).

$$X^A_{add,D} = E^A_{dest} - \sum_{i=1}^{N} E^A_{dest,i} \tag{23}$$

$$X^B_{add,D} = E^B_{dest} - \sum_{i=1}^{N} E^B_{dest,i} \tag{24}$$

$$X_{add,D}(A, B) = X^B_{add,D} - X^A_{add,D} \tag{25}$$

Thus, $X_{add,D}(A,B)$ is a good estimate of the reliability of the adopted decomposition procedure.

Plotting the D_i values as a function of a residue index provides information which is similar to that obtained from the S_i diagram and allows a ranking of the importance of each residue in favoring/disfavoring the A → B transformation. The two diagrams are only apparently different, since a residue which favors the process has a negative S factor but a positive D factor.

Catalytic Mechanism of Diaminopimelate Epimerase

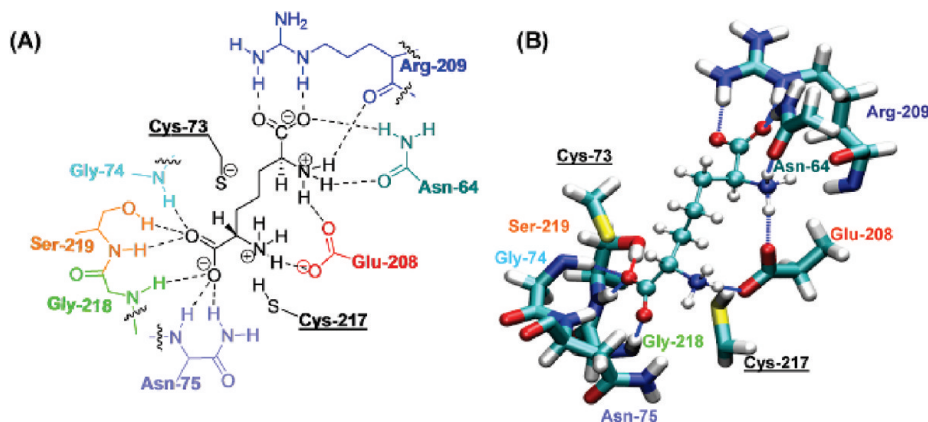*J. Chem. Theory Comput., Vol. 5, No. 7, 2009* **1921**

Figure 3. Two-dimensional (A) and three-dimensional representation of L,L-DAP within the active site. The most important residues involved in the H-bond network are shown (Asn-11 and Gln-44 are omitted for clarity).

The estimate of the van der Waals contributions (**vdWFP** analysis) is an additional useful tool to analyze the enzyme catalytic effect. This analysis is rather straightforward, being these contributions included in the QM/MM potential at the MM level. Moreover the single-residue contributions are additive for the MM force-field definition. We use a procedure similar to **DFP** and **RFP** analysis to obtain information on the role played by van der Waals interactions of each residue on the A → B transformation. We compute separately the van der Waals interaction energy between each residue and the **H** layer (note that the system here is not *model-H*). We obtain N (the total number of residues) energy values ($E_{vdW,i}^A$ and $E_{vdW,i}^B$) for both A and B. The lower $E_{vdW,i}^B$ is (relative to $E_{vdW,i}^A$), the greater is the stabilization for the critical point under examination. It is easy to compare the values obtained for the two structures and derive a stabilization factor W_i (eq 26). In this case the contribution of each component is perfectly additive (this follows from the definition of the adopted force field).

$$W_i = E_{vdW,i}^B - E_{vdW,i}^A \tag{26}$$

$$E_{vdW,i}^B > E_{vdW,i}^A \Rightarrow W_i > 0 \tag{27}$$

$$E_{vdW,i}^B < E_{vdW,i}^A \Rightarrow W_i < 0 \tag{28}$$

The **vdWFP** analysis has been performed by means of several calculations carried out with the *anal* module from the AMBER8.0 package. The **vdWFP** results can be easily represented using a plot of W_i versus i, similar to the diagrams obtained for **DFP** and **RFP**.

The three methods of analysis previously described have been implemented in the **COBRAMM** package. Useful discussions on this type of analysis can be found in previous works by Karplus and co-workers.[58-63,65] These papers have been very helpful and informative to develop our **DFP** and **RFP** approaches. These computational tools have been used here to obtain a per-residue analysis, but they can be easily exploited for a generic A → B transformation to investigate the effects arising from different groups of atoms or single atoms. Thus, in principle, we can easily provide a per-atom analysis, or alternatively, we can perform our analysis to understand the role played in the catalysis by secondary

structures (α-helices or β-sheets or combinations of them) present in the enzyme under examination.

III. Results and Discussion

III.A. Structure of the Initial Complex. The conformational studies of the binding mode of L,L-DAP and D,L-*meso*-DAP using the three different methods described in section II.A point to enzyme−substrate complexes which are very similar in structure. The geometry of these complexes is also very close to the crystallographic structures of 2GKE and 2GKJ, that strictly mimic the natural binding mode of the substrates but are characterized by a covalent bond between the enzyme and L,L-AziDAP and D,L-AziDAP. In the following discussion we refer to the structure obtained in the full MM optimization of the final point provided by the simulated annealing procedure (no cutoff concerning long-range interactions was used in this procedure). This geometry has been used to construct the starting point of the QM/MM study.

The structural features of these complexes show why DAP epimerase can bind only DAP isomers characterized by configuration L at the distal ε-carbon. The binding pocket has an asymmetric arrangement of residues that can form (as a donor or acceptor) hydrogen bonds strictly suitable to bind the L isomer at the distal site. The carboxyl group forms a salt bridge with the positively charged side chain of Arg-209 and three H-bonds with the side chains of Asn-64 and Asn-190. At the same time the positively charged amino group is hydrogen bonded to the side chains of Asn-64 and Glu-208 and the carbonyl oxygen of Arg-209. These structures show that the substrates enter the active site as zwitterion.

When the L,L-DAP and D,L-*meso*-DAP are bound to the enzyme the α-carboxyl group is bonded with the amidic hydrogen of Gly-74, Asn-75, Gly-218, and Ser-219 and also with the side chain of Ser-219. The charged amino group forms hydrogen bonds with the side chains of the Asn-11, Gln-44, and Glu-208 residues. These interactions are schematically represented in Figure 3.

III.B. Model1: the Simplest Model System. The construction of the model system is a crucial point in QM/MM computations if we wish to obtain the most convenient cost/
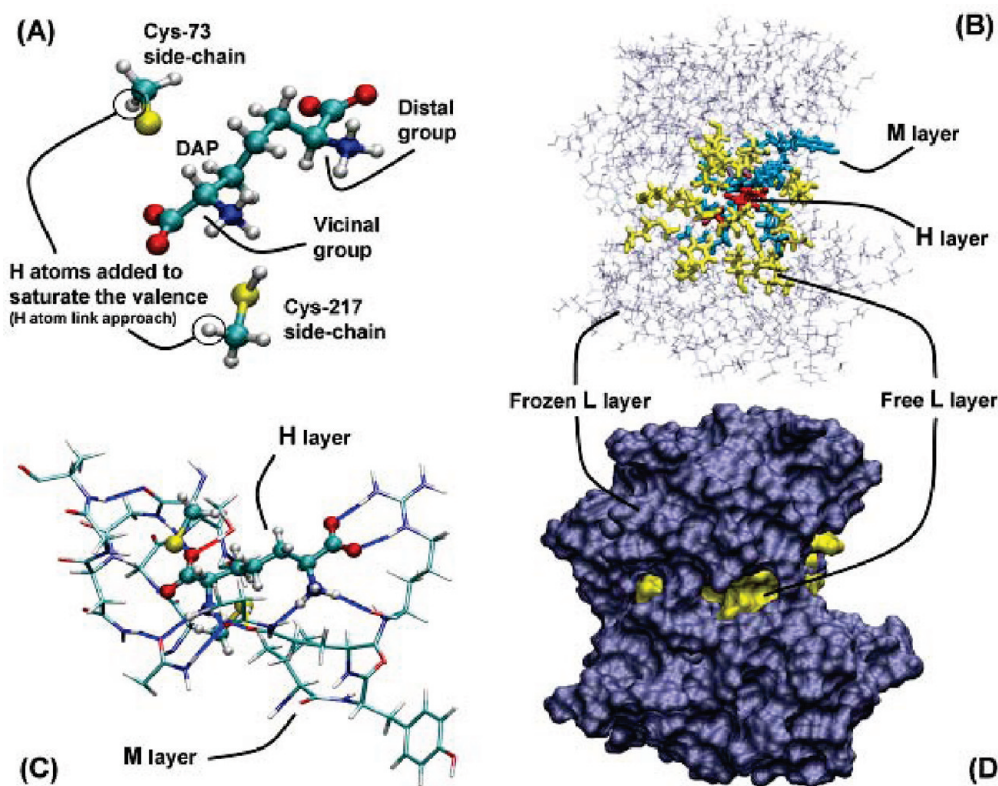
Figure 4. Partitioning scheme adopted for **Model1**. (A) QM subsystem. (B) **H** layer (red), **M** layer (blue), free **L** layer (yellow), frozen **L** layer (dark blue). (C) Hydrogen-bond network in the active site (**H** layer: ball and stick, **M** layer: stick). (D) Solvent exposed surface: free **L** layer (yellow), frozen **L** layer (dark blue).

efficiency ratio for the adopted hybrid potential. Thus, the atom selection to define the various layers (**H**, **M**, and **L** regions) is of primary importance. In the present case we have reduced the **H** layer to the smallest possible set of atoms (see Figure 4A). Since only two cysteine residues are directly involved in the enzyme-catalyzed stereoinversion process, we have included in the **H** layer the side chains of Cys-73 and Cys-217, after saturation of the dangling bonds with hydrogen atoms (atom link approach) and the entire diaminopimelate substrate. Some additional residues (see Table S1 in the Supporting Information and Figure 4B,C) surround the reacting core and are hypothesized to have an important effect on the catalytic process. These residues have been included in the **M** layer to improve the description of the system without increasing the computational demand. All remaining residues form the **L** region. No solvent effects (either explicitly or implicitly) have been taken into account in the PES computation. This can be considered a satisfactory approximation because the active site is a deep pocket far beneath the enzyme surface exposed to solvent. Moreover, a careful comparison of the crystallographic structures of L,L-AziDAP and D,L-AziDAP (the covalent complexes between enzyme and reactant-like and product-like inhibitor, respectively) shows that the substrate → product conversion does not cause important changes in the enzyme structure outside the active site. For this reason, to avoid unrealistic deformations of the structure, due to the lack of solvent, a few residues of the **L** layer in the vicinity of the **M** border were free to move during the "microiteration" steps of geometry optimization, while all other residues belonging

to **L** have been kept "frozen" at the initial positions. The "free" residues are not exposed (or exposed only to a negligible extent) to the external enzyme surface (see Figure 4D). In the following discussion we will refer to this model system as **Model1**.

The investigation of the Potential Energy Surface (PES) for **Model1** has demonstrated the existence of two critical points **M1** and **M2** that describe the enzyme bound to the reactant (L,L-DAP) and product (D,L-*meso*-DAP) molecule, respectively. An opposite protonation state of the catalytic cysteine dyad features **M1** and **M2**. In particular, in **M1** Cys-217 is protonated and Cys-73 is unprotonated and also oriented in such a way to easily grab a proton from substrate. In the hypothesized reaction mechanism the side-chain of the negatively charged Cys-73 captures a proton from the carbon substrate, while the Cys-217 thiolic proton moves toward the same carbon atom on the opposite face. It is not evident from the experimental results if the stereoinversion of the carbon atom is a concerted or stepwise process. In the second case the PES should be characterized by an intermediate species between two transition states. Any attempt to locate the intermediate of the hypothetical stepwise process has failed, and we have located only one Transition State (**TSC**) where the two cysteine residues are almost completely protonated, being the Sγ-H distance 1.43 and 1.36 Å for Cys-73 and Cys-217, respectively. This finding accounts for a concerted but highly asynchronous process. In **TSC** the substrate is deprotonated and planar. This structure allows a delocalization of the partial negative charge over the extended π orbital system. A measure of charge
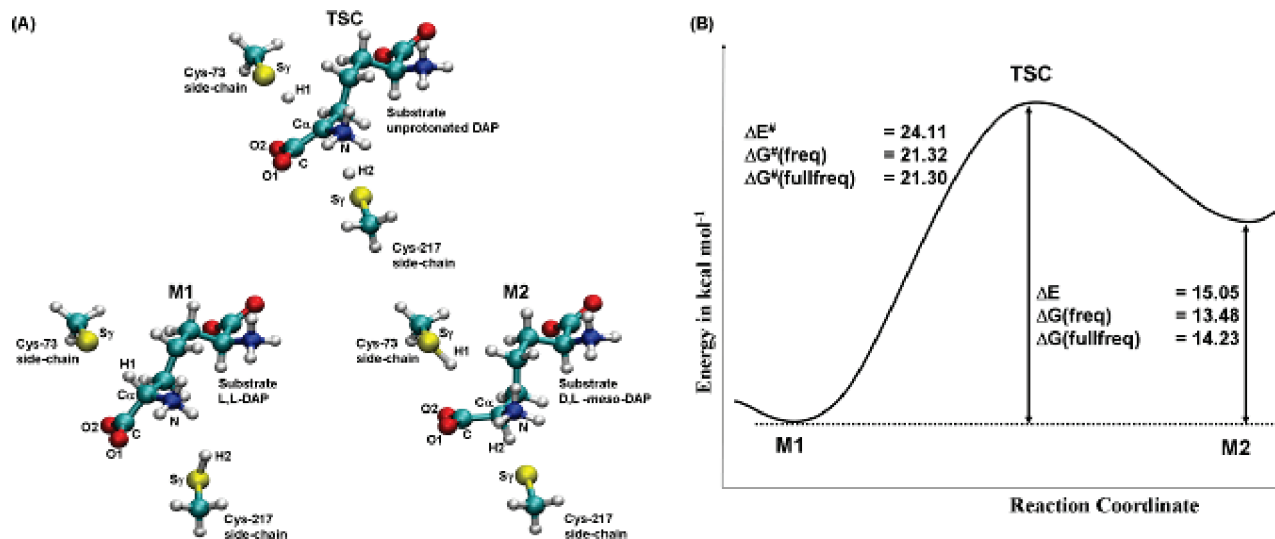
Figure 5. A) Schematic representation of the *model-H* subsystem for the three critical points **M1**, **TSC**, and **M2**. B) Reaction profile (E and G denote total and Gibbs free energy, respectively).

delocalization can be obtained from the comparative analysis of some relevant atomic distances and atomic point charges (see Table S2 in the Supporting Information, where CHELPG charges are reported). It is evident that the negative charge is mainly localized on the two oxygen atoms of the carboxyl group. Also, the carboxylic and -NH$_3$ groups become more negative and less positive, respectively, on passing from the minima to the transition state. On the contrary, the deprotonated α carbon atom does not show a significant charge variation. The charge delocalization is also proved by a bond order decrease of both carboxylic C–O bonds and a simultaneous bond order increase of the Cα–C bond.

The computed reaction profile for DAP epimerization is shown in Figure 5B (see also Table S3 in the Supporting Information). Energy (ΔE and ΔE^*) and Gibbs free energy (ΔG and ΔG^*) values relative to **M1** are given in the figure. Thermal energy corrections to compute Gibbs free energy have been obtained by numerical frequency calculations on the **HM** layer with a frozen **L** layer. Results obtained using either the approximated (**freq**) or complete (**fullfreq**) frequency computation approach are reported.

It is evident from the reaction profile that the product complex (Enzyme/D,L-*meso*-DAP) is less stable than the reactant complex (Enzyme/L,L-DAP) by 15.05 kcal mol^{-1}. This value becomes 13.48 and 14.23 kcal mol^{-1} when we consider the free energy computed with the **freq** and **fullfreq** procedure, respectively. The computed barrier for the stereoinversion is about 24.11 kcal mol^{-1}, while the corresponding free activation energies obtained with the **freq** and **fullfreq** procedure are almost identical (21.32 and 21.30 kcal mol^{-1}, respectively). The normal mode corresponding to the imaginary frequency obtained for **TSC** describes the protonation/deprotonation process of the planar substrate. The computation of numerical frequencies using the approximated procedure **freq** gives results comparable to those obtained by the **fullfreq** approach but with a significant saving of CPU time. Also, the shape of the transition vector is very similar in the two cases. Here the **H** and **M** layers are composed by 36 and 170 atoms, respectively. Thus, the **fullfreq** procedure

Table 1. Effects (kcal mol^{-1}) of the Electrostatic Interactions on the **M1/M2** and **M1/TSC** Energy Difference, As Obtained for **Model1**

$\Delta E_{QM/MM}$(**M1,M2**)	15.05
$\Delta E_{QM/MM}$(**M1,TSC**)	24.11
ΔE_0(**M1,M2**)[a]	−1.22
ΔE_0(**M1,TSC**)[a]	20.21
ΔE_{qm}(**M1,M2**)[b]	16.50
ΔE_{qm}(**M1,TSC**)[b]	26.42

[a] See eq 5. [b] See eq 6.

requires 1+(36 + 170)*6=1+(206)*6=1237 wave function evaluations, while the **freq** procedure requires only 1+(36)*6= 217, which saves about the 80% of computation time.

We have carried out the fingerprint analysis for the three critical points **M1, M2**, and **TSC** to obtain a detailed description of the influence of the various residues surrounding the reacting core. This analysis should be also helpful to ascertain the reliability of the adopted model system. In Table 1 we have collected the values of the **M1/M2** and **M1/TSC** energy differences as obtained from QM calculations (single point) on the *model-H* system (obtained from the QM/MM optimized geometries) *in vacuo* and in the presence of the atomic point charges of the whole enzyme.

The two minima are almost isoenergetic *in vacuo*: ΔE_0(**M1,M2**) is only −1.22 kcal mol^{-1}. However, in the presence of the atomic point charges the energy difference (ΔE_{qm}(**M1,M2**)) significantly increases and becomes 16.50 kcal mol^{-1}, a value which is close to the QM/MM value of 15.05 kcal mol^{-1}. This finding demonstrates the importance on the **M1/M2** equilibrium of the electrostatic interactions due to the protein environment. In particular, the negatively charged Glu-208 side chain (described at the MM level) seems to play an important role in destabilizing **M2** with respect to **M1**. This is caused by the unfavorable electrostatic interaction, occurring in **M2**, with the negatively charged Cys-217 (see discussion below). A similar effect, even less significant, is evident for the **M1/TSC** pair. The computed barrier changes from 20.21 (ΔE_0(**M1,TSC**)) to 26.42 kcal
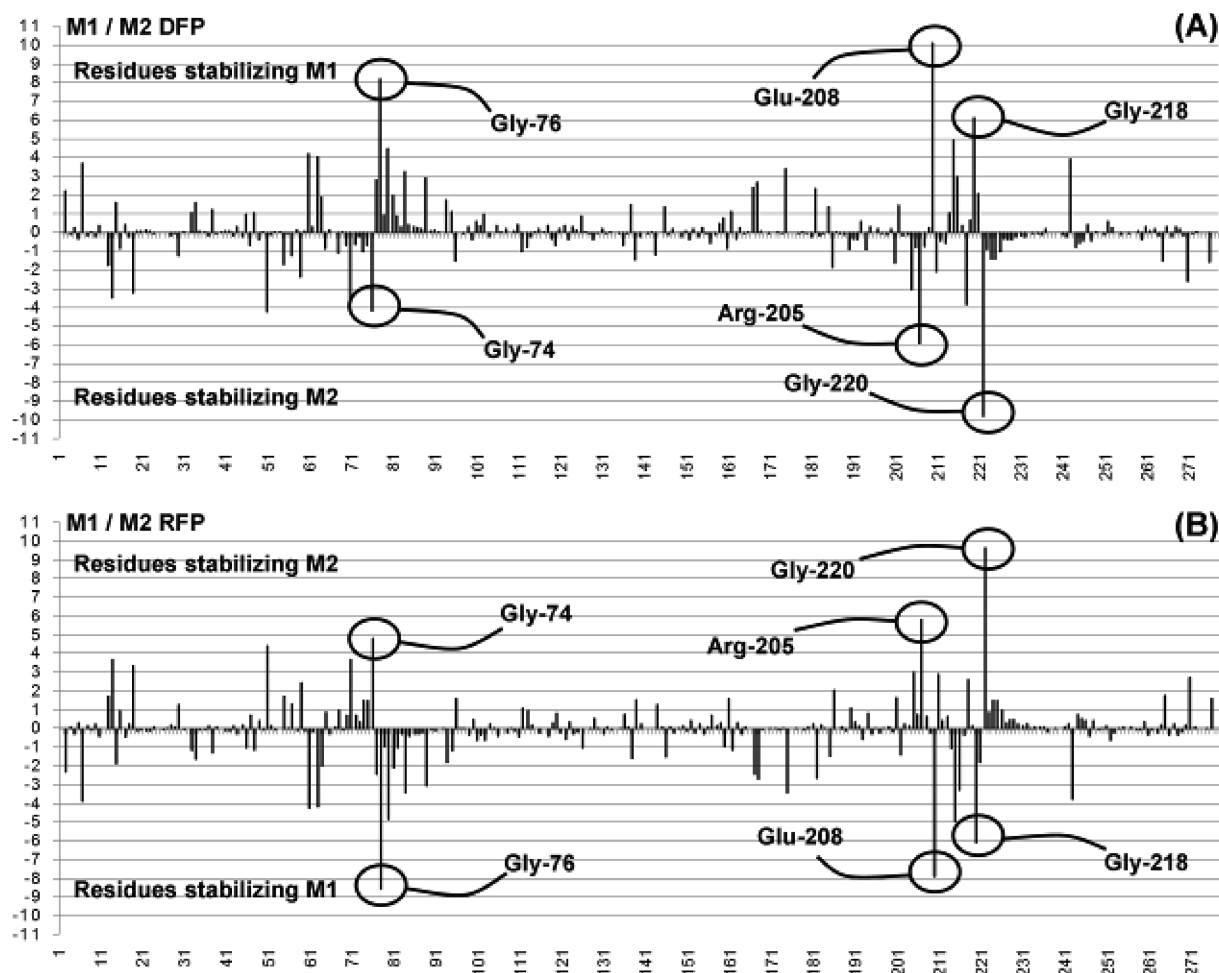
Figure 6. **DFP** and **RFP** diagrams obtained for the **M1/M2** pair in **Model1** (see Table S4 in the Supporting Information).
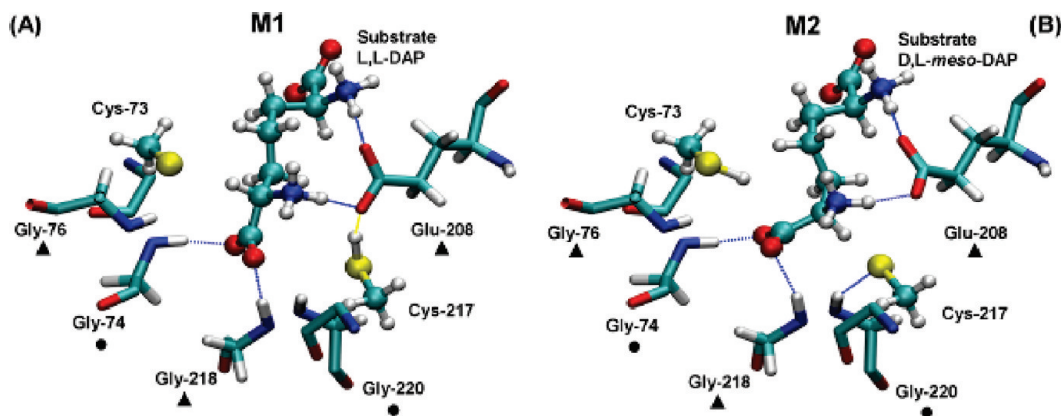

Figure 7. Schematic representation of the residues playing a key-role in the stabilization/destabilization of the **M1/M2** pair as found in **Model1**. (•) Residues stabilizing **M2** over **M1**; (▲) Residues stabilizing **M1** over **M2**.

mol^{-1} (ΔE_{qm}(**M1,TSC**)) when atomic point charges are added to the bare QM core.

In the Supporting Information we have reported the computed values of the various terms occurring in eqs 1 to 36 for **DFP** and **RFP** analysis. Here we discuss the most important stabilization (**S**) and destabilization (**D**) values obtained with the two approaches for the two pairs **M1/M2** and **M1/TSC**. **S** and **D** values for a few selected residues for the pair **M1/M2** are reported in Table S4 of the Supporting Information. The effects of the various residues are represented in the two diagrams of Figure 6.

It is evident from Figure 6 that Gly-76 plays a fundamental role in stabilizing **M1** with respect to **M2**. This stabilization (Figure 7) is due the hydrogen bond involving the Sγ atom of Cys-73. This interaction is stronger in **M1** than in **M2** because of the different protonation state of the sulfur atom: in **M2** this atom grabs the proton from the substrate and loses part of its negative charge (see Table S2 in the Supporting Information). This trend is evidenced by the (Gly-76)-NH---Sγ-(Cys-73) distance that increases from 2.34 Å to 2.39 Å on passing from **M1** to **M2**. A major role in stabilizing **M1** with respect to **M2** is also played by Glu-208. In **M1** this
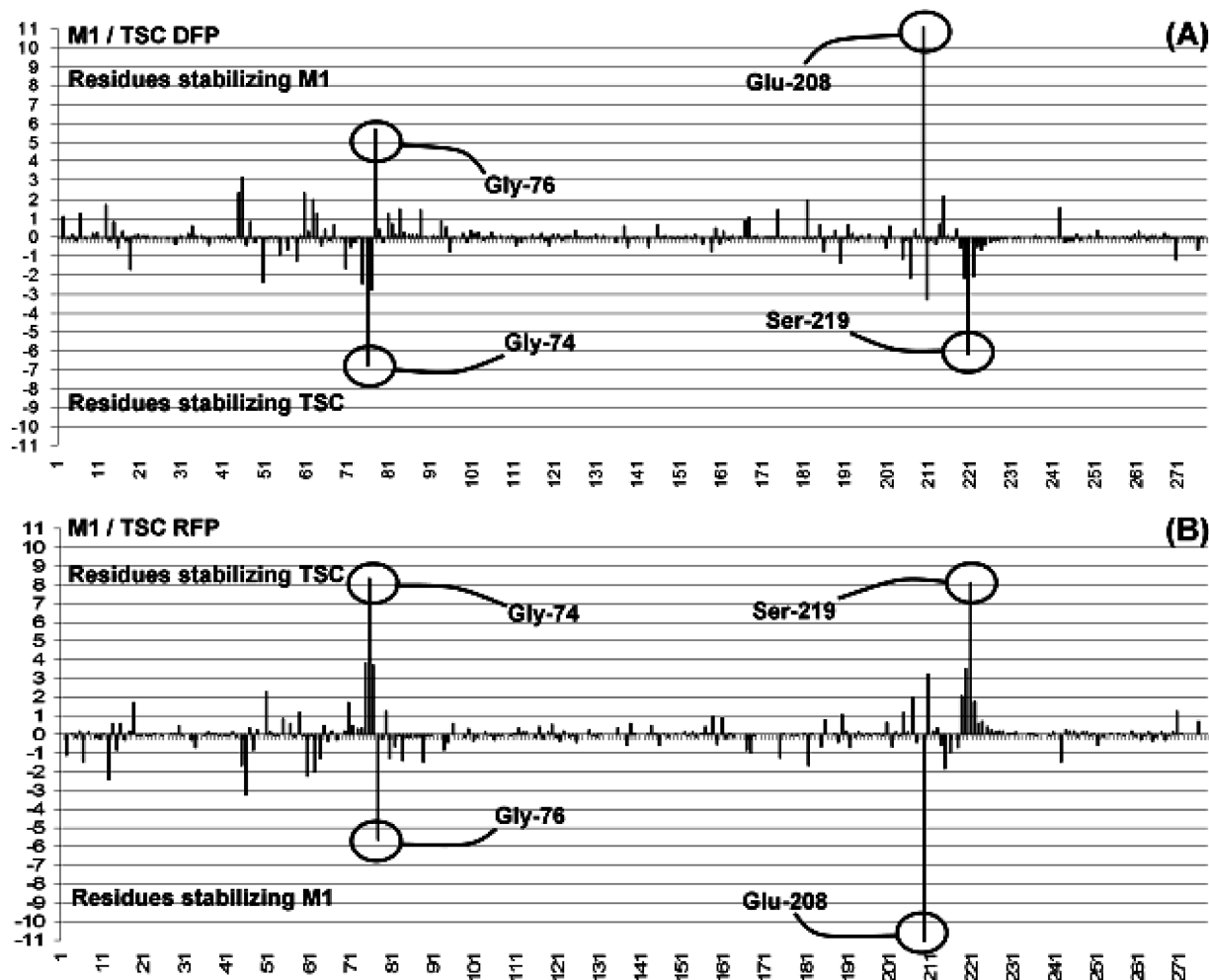
Figure 8. DFP and RFP diagrams for the **M1/TSC** pair in **Model1** (see Table S5 in the Supporting Information).

residue behaves as a triple hydrogen-bond acceptor: the unprotonated carboxylic side chain interacts with the two NH_3 substrate groups and with the thiolic hydrogen atom of Cys-217. Since in **M2** this hydrogen atom has been transferred to the substrate, only two hydrogen bonds remain active. Thus, the loss of one hydrogen bond and the negative charge on the Cys-217 $S\gamma$ atom explain the strong destabilization due to Glu-208 on **M2**. Gly-218 behaves as an H-bond donor toward the substrate carboxylic group adjacent to the α carbon undergoing the stereoinversion. The best structural arrangement is found in **M1** where the (Gly-218)-NH---$^{(-)}$OCO-(DAP) distance is 1.81 Å. This value increases to 2.09 Å in **M2** because of the stereo inversion, thus causing a decrease of stabilization in **M2**. On the contrary, some residues play an important role in stabilizing **M2** with respect to **M1.** For instance, the effect of Gly-74 is similar, but opposite, to that of Gly-218. It behaves as an H-bond donor toward the same substrate carboxylic group, but the interaction favors the D,L-*meso*-DAP (**M2**). This is confirmed by the (Gly-74)-NH---$^{(-)}$OCO-(DAP) distance which becomes shorter (from 1.77 to 1.70 Å) on passing from **M1** to **M2**. In a similar way Ala-216 stabilizes **M2** because of a more favorable structural arrangement of the NH_3 group of D,L-*meso*-DAP. The most important stabilizing effect for **M2** is due to Gly-220. This residue is a hydrogen-bond donor toward the $S\gamma$ atom of Cys-217: when this atom is not

protonated (as in **M2**) this interaction becomes much stronger and **M2** is stabilized.

A list of selected residues (with the corresponding **S** and **D** values) that play an important role in the electrostatic catalysis by stabilizing or destabilizing **TSC** with respect to **M1** is reported in Table S5 of the Supporting Information. The effects of the various residues are indicated in the two diagrams of Figure 8, and the key interactions are shown in Figure 9. The effect of Gly-76 and Gly-74 is similar to that found for the **M1/M2** pair: these residues stabilize and destabilize **M1** with respect to **TSC**, respectively, but to a smaller extent. Glu-208 again stabilizes **M1**, because in **TSC** one of the three H-bonds (involving the substrate NH_3 group) is almost completely lacking. The substrate carboxylic group is a double hydrogen-bond acceptor toward Ser-219, and the better structural arrangement favoring these interactions is found in **TSC**. However, while the distance that features the H-bond between the carboxylic group and (Ser-219)-OH varies from 1.60 Å in **M1** to 1.53 Å in **TSC** and 1.61 Å in **M2**, the distance between the carboxylic group and (Ser-219)-NH is almost constant in **M1** and **TSC** (1.76 Å and 1.77 Å, respectively) and becomes 1.87 Å in **M2**. Thus, the latter hydrogen bond does not play a key role in stabilizing the transition state.
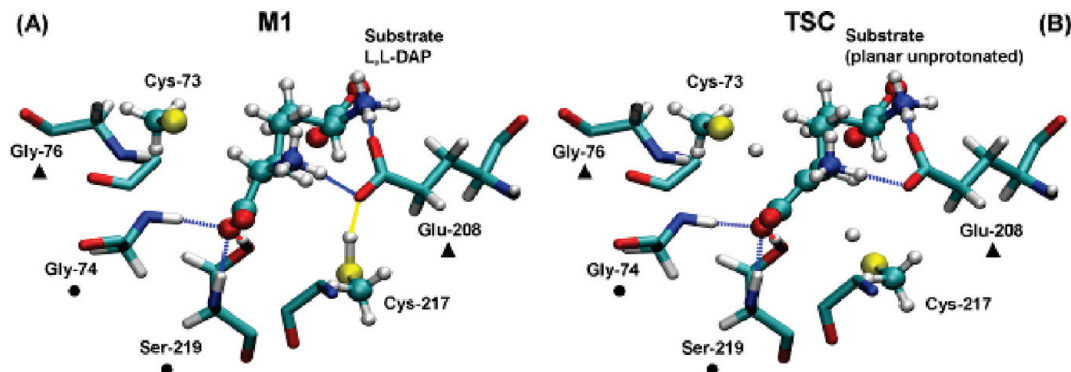
Figure 9. Schematic representation of the residues playing a key-role in the stabilization/destabilization of the **M1/TSC** pair as found in **Model1**. (•) Residues stabilizing **TSC** over **M1**; (▲) Residues stabilizing **M1** over **TSC**.

The analysis of the contributions of the various residues demonstrates the general agreement of the **DFP** and **RFP** analysis.

In Tables S6 and S7 of the Supporting Information we have collected the van der Waals contributions of a few selected residues. It is evident that these contributions are significantly less important than the electrostatic contributions (evidenced by **DFP** and **RFP** analysis) in determining the relative stabilization/destabilization of the **M1/TSC** and **M1/TSC** pairs. This suggests that, in the present case, it is reasonable to consider only the electrostatic contributions to rationalize the enzyme catalytic effect.

III.C. Model2: A More Accurate Model System. The **DFP** and **RFP** analysis on **Model1** has clearly shown the role of single residues and their importance in determining the shape of the potential surface. Since **DFP** and **RFP** can identify the residues exerting the strongest effects on the catalysis, they can help to improve the features of the model-system and obtain more reliable results.

In the previous section we have shown that Glu-208 has the most important effect in stabilizing **M1** with respect to **TSC** and **M2**. In **Model1** Glu-208 is completely described by an MM potential, while their interactions with the QM core are taken into account by electrostatic and van der Waals QM-MM cross terms. We focus here on the electrostatic effects, being that the van der Waals terms are much smaller. To check the reliability of **Model1** and to establish if the interactions between the residue Glu-208 and the substrate are correctly described, it is essential to build a different and more accurate model system, that we denote as **Model2** (see Figure S2 in the Supporting Information). The important feature of **Model2** is the inclusion of the side-chain of Glu-208 in the **H** layer (see Figure S3). In this way the negatively charged Glu-208 side-chain and its interactions with the substrate molecule and the Cys-217 residue is fully described at the QM level. We have determined for **Model2** the three critical points **M1**, **M2**, and **TSC** starting from the corresponding structures obtained for **Model1**, and we have carried out for each critical point **freq** and **fullfreq** computations and **DFP** and **RFP** analysis. Some important parameters (structural features and charge distribution) are reported in Table S8 in the Supporting Information.

The reaction energetics, as obtained for **Model2**, is reported in Figure S3 in the Supporting Information. The energy difference between reactants (Enz/L,L-DAP) and products (Enz/D,L-*meso*-DAP) decreases and becomes 11.74 kcal mol^{-1} (the corresponding free energy values computed with **freq** and **fullfreq** are 11.56 and 10.80 kcal mol^{-1}, respectively). The activation barrier only slightly changes and with respect to **Model1** becoming 25.31 kcal mol^{-1} (23.00 and 22.48 kcal mol^{-1} are the free energy values obtained from **freq** and **fullfreq** procedure, respectively). Again, the normal mode associated with the **TSC** imaginary frequency is very similar when computed with the two approaches (**freq** and **fullfreq**) and describes the protonation/deprotonation process of the planar substrate.

Thus, a comparison of the results obtained for the two model systems (Figures 5 and S3 and Tables S3 and S9 in the Supporting Information) shows that the barrier associated with the stereoinversion only slightly changes on passing from **Model1** (24.11 kcal mol^{-1}) to **Model2** (25.31 kcal mol^{-1}), while the inclusion of Glu-208 in the QM layer has a significant effect on the **M1/M2** relative energies (15.05 and 11.74 kcal mol^{-1} for **Model1** and **Model2**, respectively). The negligible variation of the barrier height suggests that the description of the Glu-208 residue is reliable enough at the MM level. On the contrary, since the inclusion of Glu-208 in the model-system has a significant effect on the **M1/M2** energy difference, a QM description of Glu-208 seems to be essential in the **M1**-**M2** comparison. This finding can be understood if we consider the particular geometrical arrangement of **M2** where the Glu-208 carboxylate is rather close to the negatively charged sulfur atom of Cys-217. The localized point charges placed on the carboxylate oxygen atoms in the MM treatment are evidently not able to correctly account for the interaction between the Cys-217 and Glu-208 residues in the **M2** structure.

A further understanding of the factors controlling the energetics of the process is provided by the values reported in Table 1 and 2 where we have collected the results of QM calculations *in vacuo* (ΔE_0) and in the presence of the enzyme atomic point charges (ΔE_{qm}) on the *model-H* system. Interestingly, for both **Model1** and **Model2** the **M1/TSC** and **M1/M2** energy differences significantly vary with respect to the QM/MM values when the "naked" *model-H* system is considered (see (ΔE_0(**M1**,**TSC**) and ΔE_0(**M1**,**M2**)). However, these terms become very similar to the QM/MM values after inclusion of the enzyme

Catalytic Mechanism of Diaminopimelate Epimerase

J. Chem. Theory Comput., Vol. 5, No. 7, 2009 **1927**

Table 2. Effects (kcal mol^{-1}) of the Electrostatic Interactions on the **M1/M2** and **M1/TSC** Energy Difference, As Obtained for **Model2**

$\Delta E_{QM/MM}$(M1,M2)	11.74
$\Delta E_{QM/MM}$(M1,TSC)	25.31
ΔE_0(M1,M2)^a	7.17
ΔE_0(M1,TSC)^a	37.81
ΔE_{qm}(M1,M2)^b	10.20
ΔE_{qm}(M1,TSC)^b	24.45

^{a} See eq 5. ^{b} See eq 6.

Table 3. Reaction Energies ($\Delta E_{QM/MM}$(M1,M2)) and Activation Energies ($\Delta E_{QM/MM}$(M1,TSC)) Obtained for **Model3**^{b}

	$\Delta E_{QM/MM}$(M1,M2)	$\Delta E_{QM/MM}$(M1,TSC)
Model3 (DZVP)	10.9	18.7
Model3 (TZVP)	10.2	16.6
experimental^a		15.6

^{a} The experimental activation energy corresponds to a free energy change ΔG, as obtained by applying the Eyring equation to the experimental k_{cat} value. ^{b} Values are in kcal mol^{-1}.

atomic point charges (compare ΔE_{qm}(**M1,TSC**) and ΔE_{qm}(**M1,M2**) to $\Delta E_{QM/MM}$(**M1,TSC**) and $\Delta E_{QM/MM}$(**M1,M2**), respectively). This clearly indicates that the most important contribution to the catalysis is electrostatic.[67]

The results of **DFP** and **RFP** analysis (see Tables S10 and S11 in the Supporting Information) are comparable to those obtained for **Model1**: the ranking order for the various residues is almost identical in the two cases, except, of course, for Glu-208 which has been included in the QM region (see Figures S4-S7 in the Supporting Information).

III.D. Model3: Toward a Reliable Description of the Chemical Transformation. We have shown in the previous sections that **DFP** and **RFP** are extremely useful in detecting the residues that play a key role in the catalysis. In both **Model1** and **Model2** these important residues have not been included in the QM region but belong to the MM shell, and, consequently, their interactions are described by the Amber potential. Thus, it is not surprising that the reaction barriers computed for **Model1** and **Model2** (21.3 and 22.5 kcal mol^{-1}, respectively) differ from the experimental value (15.6 kcal mol^{-1}, as reported in the Introduction). This is probably due to the failure of the MM force field to properly describe the changes (on passing from **M1** to **TSC**) of the charge distribution that features the hydrogen bond network involving the substrate and the above-mentioned residues. The following important point, which stems from the previous discussion, must be outlined. Even if small model systems such as **Model1** and **Model2** are not able to quantitatively reproduce the experimental data, they provide, when coupled to **DFP** and **RFP,** a valuable tool to build more reliable model systems, by making possible the inclusion in the QM region of the residues indicated as important for the catalysis. Following this approach we have defined the new model system **Model3** by adding to the QM shell all residues detected by the **DFP** and **RFP** analysis carried out for the three critical points **M1**, **M2**, and **TS**. In particular, we have considered Gly-74, Gly-76, the backbone of Asn-75, Gly-218, part of Ser-219 and Gly-220. We have carried out single point calculations on **Model3** at the geometry previously obtained for **Model2** (see section III.C). These new computations have been performed at two different levels: DFT(B3LYP/DZVP)/Amber-ff99 and DFT(B3LYP/TZVP)/Amber-ff99 levels to examine the effect of increasing the basis set accuracy.

The results confirm the capability of **DFP** and **RFP**, applied to small systems like **Model1** and **Model2**, of identifying the catalytically important residues. It is evident

from the results reported in Table 3 that the inclusion of the new residues in the QM layer determines a significant improvement in the computed reaction barrier (18.7 kcal mol^{-1}), even if single point calculations, rather than full geometry optimizations, have been carried out. A further improvement of the computed barrier (16.6 kcal mol^{-1}) is observed when the more accurate basis set TZVP is employed.

IV. Biological Insights into a Class of Enzymes: Reaction Mechanism and Catalysis

According to sequence and structural similarity, glutamate and aspartate racemases belong to one homologous family of enzymes,[11] while DAP epimerase and proline racemase are usually collected in a different group. To elucidate the general catalytic mechanism of this second family we investigated in a previous paper[22] the reaction surface of proline racemase (TcPRAC).[21] In particular we considered the enzyme found in the eukaryotic parasite *Trypanosoma cruzi*[68] because it represents a promising target for drug design against Chagas' disease[69] and can be considered a reliable model for prokaryotic proline racemases.[20,70,71] In that study we used the same computational approach of the present paper (i.e., a combination of a DFT(B3LYP/DZVP)/Amber-*ff99* potential[9] and fingerprint analysis on the critical points), and we described the mechanism of stereoinversion of the proline α carbon to afford the un-natural D-proline from the more abundant L-proline. We explained the enzyme catalysis in term of electrostatic stabilization of the transition state.

Even if TcPRAC and DAP epimerase belong to the same family, they show very poor similarities in the active site region (except for the conserved catalytically active cysteine pair). This strong structural dissimilarity is due to the need of accommodating in the reacting region two substrates which significantly differ in the shape and electronic properties. In both enzymes the substrate within the active site is present as a zwitterionic ion and is bound to the various residues by a tight network of hydrogen bonds. These interactions determine the right orientation of the substrate by anchoring the amino and the carboxylic groups and, in the case of DAP epimerase, the distal site of the molecule. In particular, the carboxylic group seems to be highly relevant to the catalytic process although not directly involved in the chemical reaction (breaking and forming of chemical bonds). To better understand this aspect, we must focus our attention on the nature of the transition state that describes for both

enzymes a concerted but highly asynchronous mechanism. This transition state corresponds to an almost fully deprotonated substrate where the α carbon (originally the sp[3] carbon) and the carboxylic group form an extended planar π system that delocalizes the partial negative charge left on the substrate after proton abstraction. This charge delocalization decreases the carbanionic character of the transition state (as suggested by charge analysis) and increases the negative charge on the carboxylic group (especially the two O atoms). The final effect on the transition state is an increase of the strength (and stabilization) of the hydrogen bonds between the carboxylic group and some residues in the vicinity of the cysteine pair.

The fingerprint analysis (DFP for TcPRAC[9] and both DFP and RFP for DAP epimerase) has evidenced the residues which are responsible for the electrostatic catalysis. These residues are hydrogen bond donor to the substrate carboxylic group. It is worth pointing out that, even if the stabilization of the transition state is determined by different residues in the two enzymes, the nature of the overall catalytic effect remains the same and can be mainly ascribed to specific electrostatic contributions of "pre-organized" active sites where some residues are in the optimum positions to emphasize the effect of the stabilizing hydrogen-bonds. This finding is in very good agreement with the conclusions reached by Warshel.[6,67,72−77]

The presented results, beside helping to obtain a general picture of the catalytic mechanism for an important class of enzymes, allow the identification of common features shared within the same enzyme family. This information should be highly useful in the future for the design and development of new drugs targeting this group of PLP-independent racemases/epimerases.

V. Conclusions

In this paper we have provided a detailed description of the reaction mechanism of the enzyme diaminopimelate (DAP) epimerase, a promising target for rational drug design aimed at developing new selective antibacterial therapeutic agents. This enzyme represents a model for the PLP-independent racemases/epimerases acting *via* a two-base mechanism involving a pair of cysteine residues (thiol/thiolate pair at neutral pH[11,27]).

We have used a QM/MM computational approach based on a DFT(B3LYP/DZVP)//Amber-*ff99* potential.[9] This approach is similar to that employed in a previous paper where we have investigated the mechanism of proline racemase (TcPRAC).[21] Two different model-systems have been investigated. In one case (**Model1**) the entire substrate (DAP molecule) and the side-chains of Cys-73 and Cys-217 (after saturation of the dangling bonds with hydrogen atoms) have been included in the **H** layer described at the QM level. The remaining part of the enzyme has been treated at the MM level (**M** and **L** layers). In the second model-system (**Model2**) the side-chain of Glu-208 has been included in the **H** region. Thus, in this case the negatively charged Glu-208 side-chain and its interactions with the substrate molecule and the Cys-217 residue have been completely described at the QM level. Both model-systems have provided the same

mechanistic picture: the reaction proceeds through a highly asynchronous mechanism where the side-chain of the negatively charged Cys-73 (thiolate) captures a proton from the carbon substrate. Simultaneously, the Cys-217 thiolic proton moves toward the same carbon atom on the opposite face. In the transition state the substrate is essentially unprotonated and planar.

Direct and inverse fingerprint analysis (**DFP** and **RFP** analysis) on the three critical points **M1, M2**, and **TSC** for both **Model1** and **Model2**, have provided a detailed description of the influence of the various residues surrounding the active site and have clearly indicated that the most important contribution to the catalysis is electrostatic. DFP and FFP analysis carried out on **Model1** have pointed out that Glu-208 has the most important effect in stabilizing reactants (**M1**) with respect to transition state (**TSC**) and products (**M2**). The indication of the fingerprint analysis has suggested including this residue in the QM region of the model. The aim of this choice was to establish if the MM potential was reliable to describe the interactions of Glu208 with the substrate and Cys-217. A comparison of the energetics obtained for the two model-systems has shown that, while the stereoinversion barrier does not significantly change (24.11 and 25.31 kcal mol^{-1} for **Model1** and **Model2** respectively), the inclusion of Glu-208 in the QM layer has a stronger effect on the **M1/M2** relative energy, which is 15.05 kcal mol^{-1} in **Model1** and becomes 11.74 kcal mol^{-1} in **Model2**. This finding suggests that the MM description of the Glu-208 residue is reliable when comparing reactants (**M1**) and transition state (**TSC**), while a QM description of this residue seems to be essential in the **M1**-**M2** comparison.

Using the results of the fingerprint analysis on **Model2** we have built a larger (and more reliable) model system, **Model3**, where all important residues detected by **DFP** and **RFP** have been included in the QM region. Single point computations on **Model3**, using the **Model2** structures and the two basis sets DZVP and TZVP for the QM shell, have provided activation energies in good agreement with the experimental value of 15.6 kcal mol^{-1}: 18.7 and 16.6 kcal mol^{-1} at the DZVP and TZVP levels, respectively. These results confirm the validity of our approach and the possibility of using **DFP** and **RFP** analysis to identify the catalytically important residues and, thus, build reliable model systems.

Supporting Information Available: Computational details, additional results in the form of tables and figures in text format, molecular geometries in PDB format and parameters in Amber format for nonstandard residues in a ZIP archive. This material is available free of charge via the Internet at http://pubs.acs.org.

Catalytic Mechanism of Diaminopimelate Epimerase

J. Chem. Theory Comput., Vol. 5, No. 7, 2009 **1929**

References

(1) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VHC Publishers: New York, 1995; Vol. 7, pp 119–185.

(2) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2006**, *117*, 185–199.

(3) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

(4) Hillier, I. H. *J. Mol. Struct.* **1999**, *463*, 45–52.

(5) Sherwood, P. NIC series 2000, 3, 285–305. http://www.fz-juelich.de/nic-series/NIC-Series-e.html (accessed March 03, 2009).

(6) Klahn, M.; Braun-Sand, S.; Rosta, E.; Warshel, A. *J. Phys. Chem. B* **2005**, *109*, 15645–15650.

(7) Senn, H. M.; Thiel, W. In *Atomistic Approaches in Modern Biology: from Quantum Chemistry to Molecular Simulations*; 2007; Vol. 268, pp 173–290.

(8) Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.

(9) Altoe, P.; Stenta, M.; Bottoni, A.; Garavelli, M. *Theor. Chem. Acc.* **2007**, *118*, 219–240.

(10) Koo, C. W.; Blanchard, J. S. *Biochemistry* **1999**, *38*, 4416–4422.

(11) Pillai, B.; Cherney, M. M.; Diaper, C. M.; Sutherland, A.; Blanchard, J. S.; Vederas, J. C.; James, M. N. G. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 8668–8673.

(12) Brunetti, L.; Galeazzi, R.; Orena, M.; Bottoni, A. *J. Mol. Graphics Modell.* **2008**, *26*, 1082–1090.

(13) Yoshimura, T.; Esak, N. *J. Biosci. Bioeng.* **2003**, *96*, 103–109.

(14) Major, D. T.; Gao, J. L. *J. Am. Chem. Soc.* **2006**, *128*, 16345–16357.

(15) Major, D. T.; Nam, K.; Gao, J. *J. Am. Chem. Soc.* **2006**, *128*, 8114–8115.

(16) Amadasi, A.; Bertoldi, M.; Contestabile, R.; Bettati, S.; Cellini, B.; Luigi di Salvo, M.; Borri-Voltattorni, C.; Bossa, F.; Mozzarelli, A. *Curr. Med. Chem.* **2007**, *14*, 1291–1324.

(17) Glavas, S.; Tanner, M. E. *Biochemistry* **2001**, *40*, 6199–6204.

(18) Puig, E.; Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M.; Field, M. J. *J. Phys. Chem. B* **2007**, *111*, 2385–2397.

(19) Puig, E.; Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Phys. Chem. A* **2006**, *110*, 717–725.

(20) Rudnick, G.; Abeles, R. H. *Biochemistry* **1975**, *14*, 4515–4522.

(21) Buschiazzo, A.; Goytia, M.; Schaeffer, F.; Degrave, W.; Shepard, W.; Gregoire, C.; Chamond, N.; Cosson, A.; Berneman, A.; Coatnoan, N.; Alzari, P. M.; Minoprio, P. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 1705–1710.

(22) Stenta, M.; Calvaresi, M.; Altoe, P.; Spinelli, D.; Garavelli, M.; Bottoni, A. *J. Phys. Chem. B* **2008**, *112*, 1057–1059.

(23) Viola, R. E. *Acc. Chem. Res.* **2001**, *34*, 339–349.

(24) Born, T. L.; Blanchard, J. S. *Curr. Opin. Chem. Biol.* **1999**, *3*, 607–613.

(25) Work, E. *Nature* **1950**, *165*, 74–75.

(26) Eyring, H. *Chem. Rev.* **1935**, *17*, 65–77.

(27) Diaper, C. M.; Sutherland, A.; Pillai, B.; James, M. N. G.; Semchuk, P.; Blanchard, J. S.; Vederas, J. C. *Org. Biomol. Chem.* **2005**, *3*, 4402–4411.

(28) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(29) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, 368–371.

(30) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 765–783.

(31) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(32) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132–146.

(33) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623–1641.

(34) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(35) Ponder, J. W.; Case, D. A.; Valerie, D. In *Advances in Protein Chemistry*; Academic Press: 2003; Vol. 66, pp 27–85.

(36) Onufriev, A.; Bashford, D.; Case, D. A. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.

(37) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.

(38) Feig, M.; Karanicolas, J.; Brooks, C. L. *MMTSB Tool Set (2001)*; MMTSB NIH Research Resource, The Scripps Research Institute: 2001.

(39) Moustakas, D. T.; Lang, P. T.; Pegg, S.; Pettersen, E.; Kuntz, I. D.; Brooijmans, N.; Rizzo, R. C. *J. Comput.-Aided Mol. Des.* **2006**, *20*, 601–619.

(40) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. *J. Mol. Biol.* **1982**, *161*, 269–288.

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

(42) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.

(43) Prat-Resina, X.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Mol. Struct. - Theochem* **2003**, *632*, 297–307.

(44) Prat-Resina, X.; Bofill, J. M.; Gonzalez-Lafont, A.; Lluch, J. M. *Int. J. Quantum Chem.* **2004**, *98*, 367–377.

(45) Vreven, T.; Morokuma, K.; Farkas, 951 > O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760–769.

(46) Broyden, C. G. *J. Inst. Math. Appl.* **1970**, *6*, 76–90.

(47) Fletcher, R. *Comput. J.* **1970**, *13*, 317–322.

(48) Goldfarb, D. *Math. Comput* **1970**, *24*, 23–26.

(49) Shanno, D. F. *Math. Comput* **1970**, *24*, 647–656.

(50) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.

(51) Geerlings, P.; De Proft, F.; Langenaeker, W. *Chem. Rev.* **2003**, *103*, 1793–1873.

(52) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Science* **2008**, *321*, 792–794.

(53) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(54) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560–571.

(55) Frisch, M. J., Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004 (http://www.gaussian.com/).

(56) Odon, F.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111*, 10806–10814.

(57) Schlegel, H. B. *J. Comput. Chem.* **2003**, *24*, 1514–1527.

(58) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826–5832.

(59) Cui, Q.; Karplus, M. *J. Am. Chem. Soc.* **2001**, *123*, 2284–2290.

(60) Cui, Q.; Elstner, M.; Karplus, M. *J. Phys. Chem. B* **2002**, *106*, 2721–2740.

(61) Cui, Q.; Karplus, M. *J. Phys. Chem. B* **2002**, *106*, 1768–1798.

(62) Cui, Q.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 3093–3124.

(63) Zhang, X.; Harrison, D. H. T.; Cui, Q. *J. Am. Chem. Soc.* **2002**, *124*, 14871–14878.

(64) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. *Org. Biomol. Chem.* **2004**, *2*, 968–980.

(65) Banerjee, A.; Yang, W.; Karplus, M.; Verdine, G. L. *Nature* **2005**, *434*, 612–618.

(66) Zhang, X.; Bruice, T. C. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 6148–6153.

(67) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. *Chem. Rev.* **2006**, *106*, 3210–3235.

(68) Chamond, N.; Gregoire, C.; Coatnoan, N.; Rougeot, C.; Freitas, L. H.; da Silveira, J. F.; Degrave, W. M.; Minoprio, P. *J. Biol. Chem.* **2003**, *278*, 15484–15494.

(69) Chamond, N.; Goytia, M.; Coatnoan, N.; Barale, J. C.; Cosson, A.; Degrave, W. M.; Minoprio, P. *Mol. Microbiol.* **2005**, *58*, 46–60.

(70) Cardinale, G. J.; Abeles, R. H. *Biochemistry* **1968**, *7*, 3970–3978.

(71) Fisher, L. M.; Albery, W. J.; Knowles, J. R. *Biochemistry* **1986**, *25*, 2529–2537.

(72) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227–249.

(73) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035–27038.

(74) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425–443.

(75) Olsson, M. H. M.; Parson, W. W.; Warshel, A. *Chem. Rev. (Washington, DC, U. S.)* **2006**, *106*, 1737–1756.

(76) Rosta, E.; Klahn, M.; Warshel, A. *J. Phys. Chem. B* **2006**, *110*, 2934–2941.

(77) Warshel, A.; K, S. P.; Mitsunori, K.; Parson, W. W. *Biochim. Biophys. Acta* **2006**, *1764*, 1647–1676.