Reaction Mechanism of the Mandelate Anion Racemization Catalyzed by Mandelate Racemase Enzyme: A QM/MM Molecular Dynamics Free Energy Study

Xavier Prat-Resina,[†,‡] Àngels González-Lafont,[†,§] and José M. Lluch*[,†,§]

Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain, and Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain

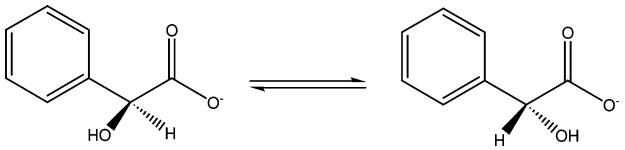
Received: April 29, 2005; In Final Form: September 2, 2005

The present work studies the reaction mechanism of the racemization of mandelate substrate by mandelate racemase enzyme. The reaction has some intriguing aspects such as the deprotonation of a nonacid hydrogen and the achievement of the pseudosymmetry necessary to obtain the racemic mixture. We will make use of a QM/MM potential energy surface to compute the free energy profiles associated with the reaction. The most favorable reaction mechanism consists of two proton transfers and the configuration inversion of the stereogenic carbon taking place in a concerted manner. We have also designed a suitable reaction coordinate to compute the free energy profiles for this rather complicated reaction. In addition, analysis of the electrostatic effects and bond distances along the reaction will explain how the enzyme accomplishes the catalysis. Finally, the enzymatic reaction will be compared to a model of the uncatalyzed reaction and the catalytic effect of mandelate racemase will be evaluated.

1. Introduction

1.1. Mandelate Racemase. Mandelate racemase (E.C. 5.1.2.2.; MR) from the soil bacteria *Pseudomonas putida* catalyzes the reversible interconversion of the (*S*) and (*R*) enantiomers of mandelate anion.[1] Experimentally, it has been studied as a paradigm for enzymes which catalyze rapid carbon−hydrogen bond cleavage.[2−4]

SCHEME 1



The flexibility of its hydrophobic pocket[5] permits the racemization of many α-hydroxy-*β*-*γ*-unsaturated carboxylate substrates.[6,7] For this reason, MR has been suggested as a useful and remarkably stable biocatalyst with important industrial and biotechnological applications,[8] mainly when combined with an enantioselective transformation to obtain a deracemization process.

In addition, from the mechanistic point of view, the reaction is characterized by several intriguing questions. The first is the rapid proton exchange involving carbon−hydrogen bond cleavage of carbon acids with relatively high pK_a values. The estimated pK_a in solution of the α-hydrogen is ∼22 or ∼29 for the mandelic acid or its mandelate anion, respectively.[2] A comparison between the reaction kinetics of enzymatic and nonenzymatic processes has concluded that mandelate racemase produces a rate enhancement under neutral conditions at 25 °C of 1.7×10^{15}-fold.[4,9]

The second intriguing question is the origin of the symmetry of the reaction.[10] The racemase enzyme family catalyzes the interconversion of both substrate enantiomers (the epimerase family is closely related, since it interconverts the two diastereoisomers). Although most enzymes are famous for being exquisitely asymmetric, by definition, racemases process both enantiomers at similar rates. Thus, the following question arises: how can these enzymes, which are inherently asymmetric, deal with both enantiomers with at least approximately equal facility? The logical answer to this question is that racemases must have evolved a functional and structural *pseudosymmetry* in their active sites; however, a detailed and microscopic explanation to this fact has not been given yet.

Another interesting aspect is that MR belongs to the enolase superfamily[11] which is characterized by the catalysis of Mg^{2+} assisted abstraction of the α-protons of carboxylic acids. Although the different members of the family catalyze very different types of reaction, they share a common catalytic strategy and all of them have a highly conserved active site. This fact will help us to look at certain conserved residues and understand their role and why they are conserved.

A large amount of experimental data has been described for MR enzyme. The X-ray structure[12] and several kinetic measurements, isotopic effects, and site-directed mutagenesis (K166R,[13] H297N,[14] E317Q,[3] D270N,[15] and N197A[16]) help to create an understanding of the basic reaction mechanism (a representation of the active site is depicted in Figure 1). It is generally assumed that mandelate racemase proceeds by a two-base mechanism. The ϵ-amino group of Lys166 is the general base catalyst that abstracts the α-proton from (*S*)-mandelate, whereas the imidazole group of His297 acts as the general base catalyst that removes the corresponding α-proton from (*R*)-mandelate. The

* Corresponding author.
† Departament de Química.
‡ Present address: Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Ave, Madison, WI 53706.
§ Institut de Biotecnologia i de Biomedicina.

10.1021/jp052239d CCC: $30.25
Published on Web 10/15/2005

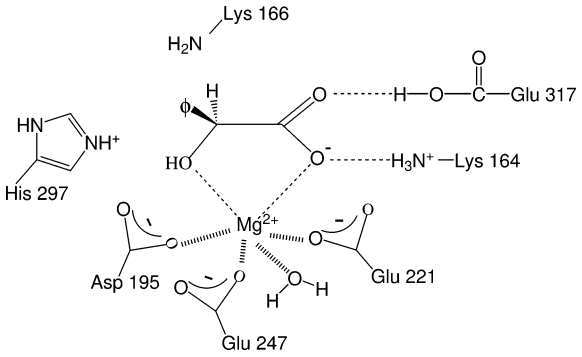


Figure 1. Schematic representation of the active site of mandelate racemase.

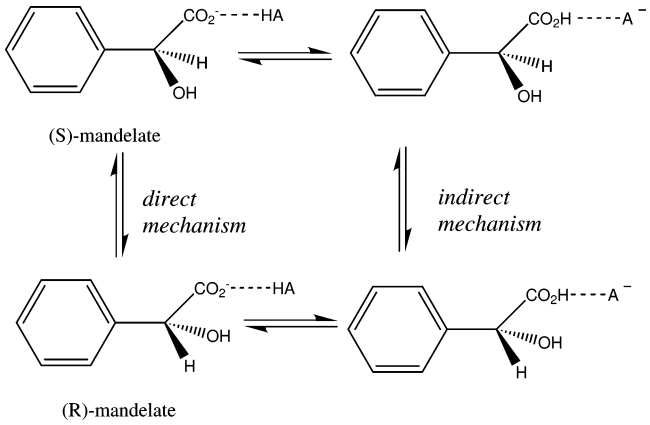


Figure 2. Three possible reaction mechanisms have been studied in our previous reaction path optimization study. Two indirect mechanisms that need the previous catalytic protonation (through Lys164 or Glu317) of the mandelate carboxylic group before configuration inversion and only one direct mechanism where *S* to *R* transition takes place in a concerted manner.

conjugate acids of Lys166 and His297 serve as the proton donors in the formation of (*S*)- and (*R*)-mandelate, respectively.

In addition, Lys164, Glu317, and magnesium cation help in the ligand binding at the first stage of the reaction and the electron density withdrawal to stabilize the enolic intermediate when proton abstraction takes place.

The experimental design of alternative substrates has helped the important factors in the active site for the binding process as well as for the reaction to be discovered. For example, related to the binding process, the α-OH group seems to be crucial, and while meta- and para-substituted phenyl rings bind the active site, the ortho-substituted rings provoke a remarkable steric hindrance.[7] Since the amide derivative of mandelate, although with a significantly lower rate, can racemize,[6] the presence of a carboxyl or negatively charged oxygen on the substrate does not seem to be essential for binding. This new experimental result seems to be against the hypothesis that the formation of a strong and short hydrogen bond between mandelate and Glu317 is essential for the catalysis.[17]

The detailed substrate spectrum also gives information about the possible mechanism, for example, electron-withdrawing phenyl substituents enhance the enzyme activity, which means that a stabilization of the negative charge is needed for the racemization.[8] An aromatic system must be present in the *β*-position being the vinyl-glycolate the minimal conjugated system.[18,19] In the absence of *π*-electrons in this position, such as in the case of lactate, no racemization occurs.

The experimental results suggest that the mechanism takes place through an enolic intermediate. The term *enolic intermediate* is preferred rather than carbanion, enolate, or enol to avoid specifying the extent to which the proton is transferred from the general acid catalyst (Glu317 or Lys164) to the oxygen atom of the intermediate. In any case, any isolation of this intermediate has not been reported.

In this paper, we carry out a theoretical study that will shed some light on some issues where experimental results lack microscopic detail. We already studied the MR reaction mechanism exploring the potential energy surface (PES),[19−21] and we optimized minimum and saddle point structures[22] on a quantum mechanics/molecular mechanics (QM/MM) system that included the whole enzyme. We studied the racemization of different substrates, reproducing the observed experimental tendency in their catalytic rate and explaining the detailed reaction mechanism.

For all of the different substrates, we saw that there were three possible mechanisms for racemization. As shown in Figure 2, two stepwise indirect mechanisms and a direct one exist. The two indirect mechanisms consist of three basic steps: first, there is a catalytic protonation of the carboxylate group in mandelate by the hydrogen bonded residues, either Lys164 or Glu317; then, the configuration inversion takes place through deprotonation (from (*S*) to (*R*) enantiomers) of mandelic acid by Lys166 and back-donation by His297; finally, the catalytic proton at the carboxylic group is back-transferred to Lys164 or Glu317. The direct mechanism does not need the protonation of the mandelate carboxylate group prior to the proton abstraction from the C_α; in this case, Lys166 pulls up the alpha proton from the (*S*)-mandelate substrate, while the protonation by His297 converts the enolic intermediate into (*R*)-mandelate. On the basis of the reaction path optimization, we concluded that the direct mechanism was the most favorable and took place rather asynchronously, where at the first stage of the reaction Lys166 played the major role while the doubly protonated His297 is the most relevant residue during the second half of the reaction. We saw that the enol intermediate was rather a transition state (TS) than a stable species.

The aim of the present work is twofold. First, we try to define an appropriate reaction coordinate along which a reliable free energy profile for the mandelate anion racemization catalyzed by MR can be calculated. Then, we analyze the averaged and long range effects and the overall enzyme contribution to catalysis in order to shed light on the reaction mechanism. Despite the fact that our previous studies already explained the basic intriguing points related to the reaction mechanism, the PES optimization method used before does not take into account the temperature effects and the essential geometric flexibility at the active site.

1.2. Theoretical Methods in Enzyme Catalysis. Theoretical methods are currently used to elucidate some challenging problems related to enzymatic catalysis.[23,24] QM/MM methods,[25−27] combined with molecular dynamics (MD) or Monte Carlo (MC) methods to ensure enough sampling over the phase space, are the most adequate strategies for exploring the enzymatic reactivity. Although state of the art methodology has still some lack of accuracy that does not permit one to go far beyond the semiquantitative description, the information that theoretical methods provide is extremely useful for understanding the enzymatic complexity at the atomic detail and to complement the experimental results.

In the present work, we compute free energy profiles and averaged contributions with a QM/MM model of MR. We will be able to elucidate some clue factors in the catalysis and in the mechanism. In particular, we describe the nature of the enolate intermediate, the contribution of some residues to the

mechanism, how the enzyme stabilizes the unstable carbanionic species, and how the *pseudosymmetry* of a racemase is achieved in an asymmetric environment. In addition, some discussion on the methodology is given, with special attention to the reaction coordinate selection. The reaction coordinate is the degree of freedom (geometrical or any other mapping property) capable of describing the reaction process. While in the gas phase the minimum energy path (MEP) is usually a good descriptor for the reaction,[28,29] in the condensed phase, there are many possible reaction pathways and it is more convenient to compute the free energy profile of the reaction through statistical sampling rather than to use a single MEP. In some cases, when the multiplicity of pathways comes from the multiple conformation of the solvent or the environment, some combined schemes can be applied,[30−32] where the MEP is computed on a core and statistical sampling is performed on the environment. However, these methods do not account for the flexibility and the conformation sampling in the active site, which in some cases, such as in the case of MR, is very important. Therefore, a reaction coordinate parameter must be defined prior to any free energy computation being made. Although some methods exists that avoid this selection and only need the specification of reactants and products,[33] they are still too expensive, and in most cases of enzymatic reactivity, the reaction coordinate is chosen on the basis of chemical intuition. In our present case, we will take profit of the information provided by the previous location of minima and saddle points that already give a very useful idea of the reaction mechanism.

The paper is organized as follows. In the next section, a description of the methods is given. Section 3.1 first reviews some previous conclusions on MR reactivity that will help you know what kind of chemical reaction we are studying; then, in section 3.2, potential of mean force (PMF) calculations on several reaction coordinates are performed. In section 3.3, a geometric and electrostatic analysis at different stages of the reaction will explain the role of some residues, and in section 3.4, the results from the enzyme model are compared to those from a model of the uncatalyzed reaction. Finally, some conclusions on the reactivity and methodology will be given.

2. Methods and Computational Details

2.1. The Structure and the Model. When building a molecular model to study enzymatic catalysis, we want the inclusion of the whole enzyme molecule through a QM/MM scheme for two main reasons. The first one is the energy, because the typical electronic embedding QM/MM scheme has discrete point charges polarizing the quantum wave function which describes adequately the anisotropic electrostatic effect of the enzyme. The second reason is geometric, even if we do not expect large scale movements in our reaction, the flexible nature of the active site implies the inclusion of the whole environment to avoid artificial and ambiguous frozen atoms. In this sense, QM/MM methods are not only a consistent manner to include the effect of the environment but also a reliable way to reproduce the particular mobility of the active site. In the case of MR, no fruitful ab initio calculations based on gas phase models have been reported due to the flexibility of the active site[34] which emphasized the necessity to include a large number of atoms in the model to obtain reliable optimized structures.[22]

The starting geometry coordinates have been taken from the 2.0 Å resolution X-ray crystal structure of the complex mandelate racemase−inhibitor (*R*)-α-phenylglycidate[12] (PDB code 1MNS). Since mandelate substrate is a rather rigid structure and the binding of the inhibitor is very well-known, the substrate has been added in the active site matching the maximum number of atoms with the inhibitor found in the X-ray structure.

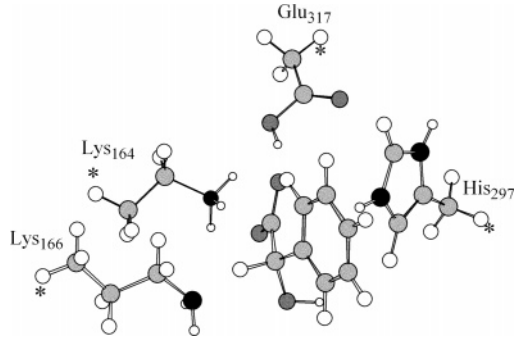


Figure 3. Atoms included in the QM part. The asterisks indicate the boundary atoms treated with the GHO method in such a way that Lys164 is modeled as an ethylamine, Lys166 as a propylamine, His297 as a methylimidazole, and Glu317 as acetic acid.

The nonpolar hydrogen atoms have been added to the enzyme crystal structure by the HBUILD facility of the CHARMM package.[35] The protonation state corresponding to all of the titrable residues has been set according to pH 7. The exception is Glu317 which is protonated to reproduce the hydrogen bond with the carboxylate of the substrate as appeared in the X-ray. As for the neutral histidine residues, the ϵ-N or the δ-N are protonated depending on the possibility to form hydrogen bonds. A remarkable issue is that when the active site binds the (*S*)-mandelate (from now on the reactants), Lys166 is unprotonated to permit the interaction with the α-hydrogen observed in the X-ray structure, and His297 is positively charged. No further calculations, such as pK_a computations, have been done to justify this protonation state, since Lys166 and His297 have been generally accepted as the two acid/base catalyst residues in MR and similar enzymes in the whole enolase superfamily.

2.2. Potential Energy Surface. In our QM/MM scheme, we have used the same semiempirical PM3[36] QM level we employed in our previous works. Although calculations using ab initio and density functional theory (DFT) levels are currently common, they are still limited to using optimization methods[37] or one-valley approaches[31,32] because they are too expensive to permit an extensive sampling of the core and the environment of the enzyme. The selection of the QM part is done according to the residues that potentially can exchange a proton with the substrate. In this case, part of the lateral chains of Lys164, Lys166, His297, and Glu317 are included in the quantum part (see Figure 3) along with mandelate. The length of the lateral chains selected as QM is done according to previous calibration studies.[19] As noted before,[19,34] the proton abstraction of mandelate by an amine or imidazole group does not take place in simplified gas phase models, neither at the semiempirical level nor at the ab initio level. This makes our PM3/MM level difficult to calibrate. However, we saw that in our case PM3 performs significantly better than AM1 in terms of energies and geometries. In addition, proton affinities (PAs) have been computed at the PM3 level and at the higher B3LYP/6-31+G(d,p) level for mandelate dianion, methylimidazole, and propylamine, and they are shown in Table 1. On the basis of these results, although PAs computed at the PM3 level are lower than the DFT ones, in the three computed cases, the error is very similar (∼17 kcal/mol; note that this value is less than 4% of the PA of mandelate dianion). Since our enzymatic reaction involves essentially proton transfers between these three species, it is expected that PM3 will not change the mechanism nor the relative stabilization of the intermediates if compared to an eventual DFT level. Therefore, it is our opinion that the PM3/MM method will give

TABLE 1: Proton Affinities (kcal/mol) Computed at the DFT and Semiempirical Levels for Gas Phase Models of the Substrate, Lys166, and His297

	DFT(B3LYP/ 6-31+G(d,p))	PM3
mandelate(Hα)$^-$/mandelate anion^{-2}	458.95	442.47
propylamine(H)$^+$/propylamine	228.04	210.82
methylimidazole(H)$^+$/methylimidazole	238.41	221.80

a good qualitative picture for the chemical events and the chemical analysis made here.

The partition of covalent bonds in the QM/MM frontier has been treated with the generalized hybrid orbital framework,[38] in particular using the PM3(GHO) method published recently.[39] The CHARMM22 force field[40] is used for the MM part as well as for the QM/MM van der Waals and electrostatic interactions. In our QM/MM partition, the only critical nonbonding interaction is between mandelate substrate and magnesium cation. However, after some tests, we have seen that the nonbonded QM/MM interaction mandelate(PM3)−magnesium(CHARMM22) using the standard charge and van der Waals parameters from the CHARMM22 force field is very similar in energy and structure to the full PM3-SRP results used in our previous studies.[19,20] The rest of the classical ligands bound to the magnesium cation, including water, are stable enough to keep its coordination over all of our 2 ns simulation.

The nonbonded interactions have been calculated with the following characteristics. The pair list is built on the basis of a group based cutoff using a distance of 13.5 Å (both QM and MM regions). The pair list is updated every 35 steps during the dynamics simulation. The van der Waals interactions are calculated using a shifting function with a cutoff at 13 Å. The electrostatic interactions have been calculated using a switching function activated at 12 Å at which the smoothing function begins to reduce the potential and is completely eliminated at 13 Å.

2.3. Molecular Dynamics. To mimic the aqueous environment, the stochastic boundary molecular dynamics (SBMD) method has been used.[41−43] It has been shown that SBMD reproduces the mobility of localized processes such as the MR catalytic reaction. We have used a sphere of pre-equilibrated waters with a radius of 24 Å to solvate the system. This sphere was centered on the stereogenic carbon of the substrate (C_α). All of the crystallographic waters beyond the sphere were removed along with any water in which oxygen is closer than 2.5 Å to any heavy atom of the protein. A soft and deformable boundary potential is applied at the edge of the sphere of waters to mimic the effect of the inexistent bulk solvent.

Classical molecular dynamics are carried out partitioning the system into three zones: the dynamics region which consists of atoms within a distance of 20 Å from the center, the buffer region which contains the atoms surrounding the dynamics region from 20 up to 24 Å, and the reservoir region which includes the remainder of the system and is excluded in the explicit dynamics simulation. The sizes of the dynamics and buffer region are large enough to ensure that the active site and all of the possible rearrangements in its surroundings during the reaction are adequately modeled.

The constants for boundary forces and friction coefficients needed in the SBMD framework are taken from the original publications[44] which have been tested thoroughly in more recent works.[45,46] The trajectory in the dynamics region is propagated using Newton's equation of motion, while in the buffer region the Langevin equations with a stochastic term are used. Both friction coefficients and boundary forces applied to protein atoms are scaled by a screening function that depends on the distance from the center.

The leapfrog algorithm to integrate the MD equations is used in all cases with a time step of 1 fs. The hydrogen atoms are constrained using the SHAKE algorithm.[47] Before taking any data, the system is heated and finally equilibrated at 298 K for 100 ps.

To evaluate the magnitude of the enzymatic catalysis, we built an uncatalyzed model which consists of the substrate and a model for Lys166 or His297 (see the Results and Discussion section for more details on the molecular models). This minimal reaction model is surrounded by a 38 Å cube of equilibrated waters. This system interacts under periodic boundary conditions with the same cutoff pattern and the same MD conditions described above. However, in this case, the Newton molecular dynamics are run all over the model without any buffer or excluding region, using the Nose−Hoover thermostat.

2.4. Potential of Mean Force Calculations. The molecular dynamics simulation described above has as a principal goal the computation of the PMF ($\Delta G(Rc)$) of the reaction.

The design of new methods for free energy calculations along a predefined degree of freedom (reaction coordinate) for the simulation of rare events is a very active field. Standard techniques still require a large amount of sampling over the phase space and need the selection of a predefined reaction coordinate which is crucial to obtain good results. Current trends are devoted, for example, to improve the methodology for free energy computations[48] to accelerate the sampling[49,50] or to avoid the a priori definition of a restrictive and usually unknown reaction coordinate.[33] Processes such as conductivity through ion channels, protein/protein interactions, or mechanisms of protein folding are specially delicate in this matter. However, the PMF for the MR mechanism is not expected to be problematic because we already know, on the basis of the PES optimization study, the information to design an adequate reaction coordinate and in our case no exceptionally large sampling is needed (no large scale conformational changes are expected during the course of the reaction). Taking this into account, the PMF calculation has been performed using the standard umbrella sampling technique.[51] In the umbrella sampling method, at every simulation window, the reaction coordinate is restrained within a limited range by imposing a bias potential. A harmonic restraining potential will be used as a biasing function. In some cases, to obtain an efficient sampling, an additional biasing potential will be employed[52] which consists of cubic spline functions fitted to a number of points along the reaction coordinate.

$$U_{tot} = U_{QM/MM} + k(Rc - Rc_0)^2 + U_{spline} \qquad (1)$$

This strategy obtains an enhanced exploration of the low probability zones without increasing the harmonic constant. The drawback is that this additional biasing function should be ideally equal to the negative of the unknown PMF ($-\Delta G(Rc)$) that is actually what we want to calculate. Usually, the points to which the spline is fitted must be determined by trial and error until obtaining the expected sampling. Some techniques exist to determine the bias potential self-consistently,[53] but they will not be used here, since the sampling is not a major problem in our case.

The most important issue will be the selection of an adequate reaction coordinate. We have used several geometric distance combinations to compute the PMF, and their adequacy will be discussed.

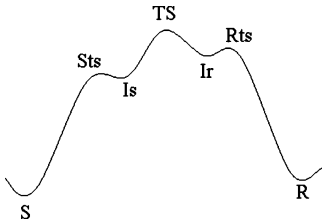
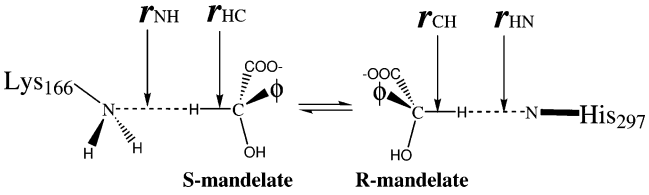


Figure 4. Potential energy profile found for the direct mechanism in the racemization of mandelate substrate by means of optimization methods.

SCHEME 2



The simulations are carried out by overlapping regions (windows) until covering the entire reaction coordinate. For every window, we run 15 ps for equilibration and 50 ps for sampling. The reaction coordinate is divided in finite bins with a length of 0.01 Å. The weighted histogram analysis method (WHAM)[54] has been used to join the different windows of simulations and build the free energy profile. In addition, for the sake of comparison, the statistical analysis has also been performed by the traditional technique of directly matching the windows, where the unknown free energy constants are determined by adjusting the various adjacent windows in the region until they match.[51] The WHAM results have been compared with those obtained by the matching technique, and few differences have been encountered. Unless indicated explicitly, the free energy profile using the WHAM will be shown.

3. Results and Discussion

3.1. Optimization of the Structures. In our precedent work on the MR mechanism,[20,22] as we commented in the Introduction, we optimized the potential energy surface and localized with a microiterative scheme the different stationary points. On the basis of those results, we decided that the most favorable mechanism, and the one studied in the present work, is the direct one which takes place without the previous catalytic protonation of the carboxylate substrate group by Lys164 or Glu317. On the PES, we saw that this mechanism is quite asynchronous. The abstraction of the proton formed an extremely shallow intermediate, I_s (see Figure 4), which is characterized by an ion pair between the carbanion and the positive Lys166. I_s undergoes the configuration inversion through the central and highest energy point (TS). At this point, the doubly protonated His297 approaches the carbanion where another very shallow minimum, I_r, is found with the same ion pair characteristics between positive His297 and the substrate. Then, through a very small energy barrier, the product (*R*) is reached.

From the stationary points found along the reaction path for the direct mechanism, we can design an appropriate reaction coordinate. Despite the fact that there are some works that used an energetic or solvent coordinate[55] which has been shown to be useful, we will use a geometrical coordinate which is a magnitude easier to define in our QM/MM scheme.

The reaction coordinate must be a combination of the main geometrical parameters that participate in the reaction, namely, the distances between heavy atoms and transferred hydrogens, r_{NH}, r_{HC}, r_{CH}, and r_{NH}, as indicated in Scheme 2. The most intuitive combinations are shown in Table 2 along with the values obtained optimizing the stationary points.

TABLE 2: Possible Combinations of Geometrical Parameters To Be Used as Distinguished Reaction Coordinates in the PMF Calculation[a]

def.	R_{NHC} $r_{HC}-r_{NH}$	R_{CHN} $r_{HN}-r_{CH}$	R_4 $r_{HC}-r_{NH}+r_{HN}-r_{CH}$	R_{HCH} $r_{HC}-r_{CH}$	R_{NCN} $r_{NC}-r_{CN}$	$\Delta U_{QM/MM}$ (kcal/mol)
S	−0.646	−1.932	−2.578	−1.769	−0.960	0.0
Sts	0.328	−1.869	−1.541	−1.343	−1.120	13.81
I_s	0.419	−1.850	−1.431	−1.275	−1.106	13.79
TS	1.247	−1.080	0.167	0.165	0.008	19.50
I_r	1.607	−0.490	1.117	1.005	0.719	16.45
Rts	1.732	−0.309	1.423	1.227	0.860	16.75
R	1.864	0.607	2.471	1.703	0.782	4.63

[a] The labels for the distances are shown in Scheme 2. In the last column, the QM/MM potential energy is shown. The results correspond to the optimized structures found in our previous study.[20]

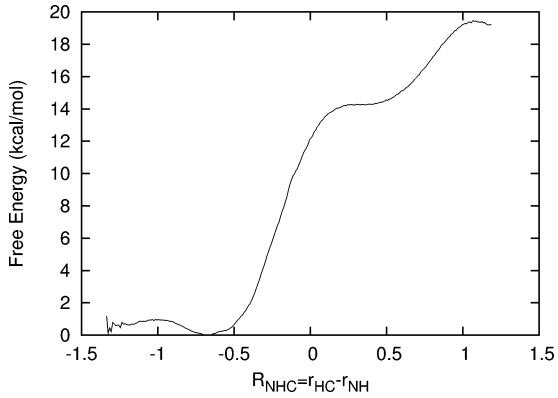


Figure 5. PMF profile using the R_{NHC} reaction coordinate.

While the asymmetric stretch coordinates R_{NHC} and R_{CHN} would be the most appropriate reaction coordinate for the proton transfer step between substrate and Lys166 and His297, respectively, we already see that they only describe partially the evolution from *S* to *R* structures. Alternatively, the combination of two asymmetric stretches could describe the whole process as long as the two magnitudes do not interfere with the other half of the reaction.

In what follows, some of these reaction coordinates are employed in the monodimensional PMF calculation to evaluate their performance.

3.2. Potential of Mean Force: Comparison between Different Reaction Coordinates. *Combining Two Bond Distances.* Starting from the equilibrated *S* structure, the reaction coordinate R_{NHC}, as defined in Table 2, is scanned by the different windows in the $S \rightarrow R$ direction. After this simulation, the final equilibrated *R* structure is taken to scan the R_{CHN} reaction coordinate in the $R \rightarrow S$ direction. Unless a major sampling is needed, the reference reaction coordinate Rc_0 used in the umbrella potential is increased by units of 0.2 Å.

The PMF using the reaction coordinate R_{NHC} is shown in Figure 5. It reaches a zone after the proton transfer from substrate to Lys166 ($R_{NHC} \sim 1.0$ Å) where this coordinate is unable to describe the carbon configuration inversion. After this region, the simulation explores a zone that does not belong to the reaction path anymore. At this point, Lys166 moves away from the substrate, leaving the stereogenic carbon in the *S* configuration.

The PMF calculation using R_{CHN} (shown in Figure 6) has different behavior but similar consequences. Starting from the

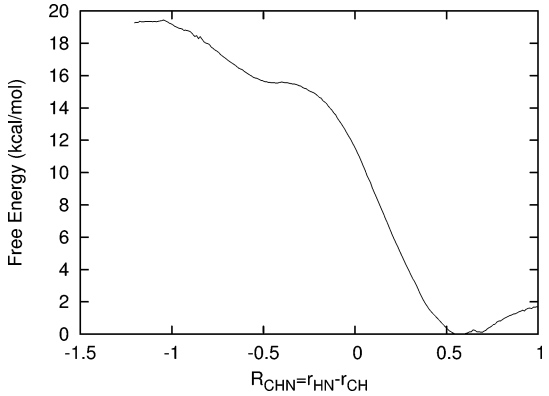


Figure 6. PMF profile using the R_{CHN} reaction coordinate.

TABLE 3: Energetics in kcal/mol Corresponding to the PMF Using R_{NHC} and R_{CHN} (the Corresponding Reaction Coordinate Bin (±0.005 Å) is in Parentheses)

reaction coordinate	R_{NHC}	R_{CHN}
minimum	*S*: 0.00 (−0.675)	*R*: 0.00 (0.575)
inflection barrier	14.29 (0.305)	15.59 (−0.405)
inflection minimum	14.25 (0.365)	15.53 (−0.425)
highest point	19.46 (1.075)	19.44 (−1.045)

R structure, R_{CHN} is not able to describe the whole reaction either. In the $R \rightarrow S$ direction, after His297 abstracts the hydrogen from the substrate ($R_{CHN} \sim -1.0$Å) and after some progress in the configuration inversion, there is a simulation window where the molecular dynamics falls into the steep valley directly to the reactant *S* region without describing the other half of the reaction. This different behavior in the case of R_{NHC} can be attributed to the fact that the central TS (as we know from the optimized structures) is rather productlike (*R*-like). Consequently, the reaction coordinate R_{CHN} is able to reach some configurations that belong to the TS and after that falls down directly to the *S* configuration.

The two flat zones encountered after the two corresponding proton transfers in Figures 5 and 6 are around $R_{NHC} \sim 0.4$ and $R_{CHN} \sim -0.5$ and before the configuration inversion steps are not thermally stable at 298 K. Therefore, although these two structures were minima with very low barrier saddle points, they cannot be considered as intermediates in terms of free energy. As a consequence, the mechanism cannot be considered stepwise, but concerted, and therefore, a unique reaction coordinate should be used to describe the whole step.

The energetics for the free energy profiles are collected in Table 3. Despite the fact that the highest energy point is meaningless, the energy corresponding to the inflection zones is well described and they will be valuable as a useful reference in the following sections.

Other possible combinations were tested such as R_{HCH} (as defined in Table 2) or a combination of Pauling's bond order[56] with contributions from the solid angle in the C_α sp^3 configurations. However, none of these alternatives were able to describe smoothly and continuously the whole reaction process.

Combining Four Bond Distances. Intuitively, if R_{NHC} reproduced adequately the *S side* of the reaction and R_{CHN} the *R side*, it indicates that combining the two of them in R_4, as it is defined in Table 2, could be an adequate way to describe the whole process. In addition, this coordinate gives freedom to the variation of the four distances, since the biasing potential only penalizes the whole combination (R_4). Therefore, this would permit the observation of a hypothetical stepwise mechanism.

Reaction coordinate R_4 is scanned from negative values by increments of 0.2 Å until ∼2.5 Å. In Figure 7, the PMF profile

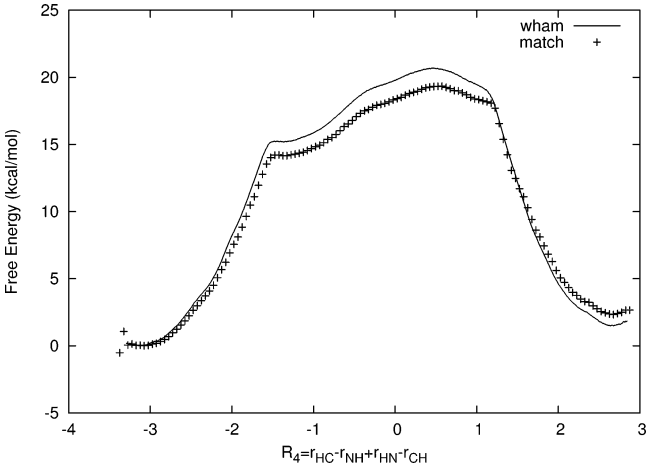


Figure 7. PMF profile using the R_4 reaction coordinate. The two profiles correspond to the PMF built by the WHAM or by a direct match of the overlapping windows.

TABLE 4: Free Energy in kcal/mol for the PMF Using R_4 as a Reaction Coordinate (the Corresponding Reaction Coordinate Bin (±0.005 Å) is in Parentheses)

PMF technique	matching	WHAM
minimum *S*	0 (−3.095)	0 (−3.095)
inflection barrier (Sts)	14.20 (−1.395)	15.25 (−1.475)
inflection minimum (I_s)	14.12 (−1.355)	15.20 (−1.415)
TS	19.38 (0.575)	20.70 (0.455)
inflection (Rts)	18.13 (1.135)	19.00 (1.125)
R	2.32 (2.675)	1.49 (2.675)

is shown. The two free energy profiles correspond to the one obtained by the WHAM technique and the other by matching the different windows. The free energy values corresponding to the most important points are displayed in Table 4. Despite some difficulties that we comment on below, the PMF computed using this reaction coordinate is able to draw the free energy profile all along the process. The shape of the profile is not the same as the one we encountered using R_{NHC} in the *S side* and R_{CHN} in the *R side* (Figures 5 and 6). The main difference is that the inflection zone in the *R side* is now less detailed. However, the energies for the early small barriers representing the proton transfer, comparing values in Tables 3 and 4, are approximately the same. In addition, R_4 can describe the central step, where the configuration inversion takes place and which is the highest point in the PMF, along with the approximation of Lys166 and His297 for the two proton transfers.

There are some differences between the results coming from WHAM analysis and by adjusting automatically the several adjacent windows by a simple match criterion. The free energy profile is in some points up to ∼1 kcal/mol different. The WHAM technique is a more sophisticated iterative procedure that takes into account all of the data without discarding the overlapping regions in the simulation, and therefore, it is a more accurate technique than directly matching. However, we cannot accept the WHAM results without a deeper analysis and trying to explain the differences. With this aim, we compare the two free energy profiles obtained using R_4 with those obtained previously using R_{NHC} and R_{CHN}. In Table 4, the energetics of the different points are shown. Note that since R_4 is unable to describe the *R side* as accurately as R_{CHN}, we only give an inflection point. A comparison between Tables 4 and 3 can be made at the inflection points. If we accept that R_{NHC} and R_{CHN} can reproduce adequately the *S side* and *R side*, respectively, we must conclude that the better free energy profile computed with R_4 will come from the matching technique. The energies

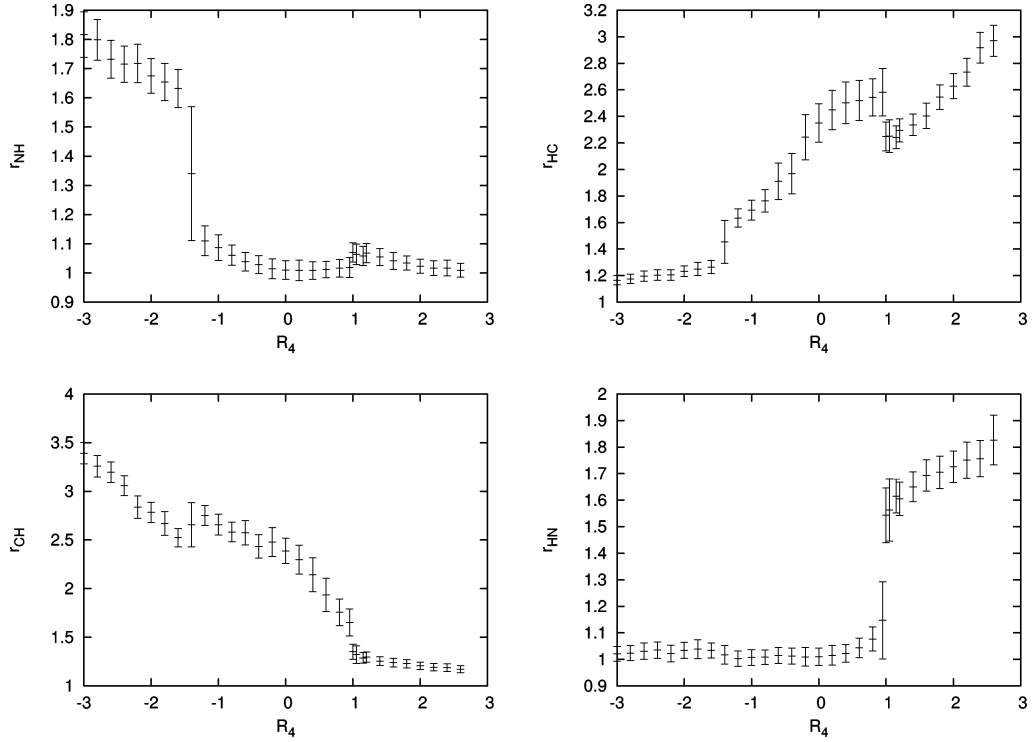


Figure 8. Evolution of the four bond distances as a function of R_4. The average value for every window is plotted along with its standard deviation.

corresponding to the inflection points in Table 3 (14.29 and 15.59 kcal/mol) are closer to the points of the first column in Table 4, Sts and Rts (14.20 kcal/mol and 15.81 (18.13−2.32) kcal/mol) using the matching technique. The reason for this disagreement between the two techniques is some hysteresis problems in the Sts and Rts inflection points. This fact does not imply that the WHAM is not accurate; on the contrary, while WHAM detects the problem and has some difficulties in converging, the less accurate matching technique directly discards the problematic zone.

A further analysis comes from monitoring the evolution of the four bond distances during the reaction. The plots displayed in Figure 8 show the average of the four distances that combine in R_4 at every simulation window.

The distances r_{HC} and r_{CH} do not change brusquely, but the approximation of Lys166 and His297 represented by the distances r_{NH} and r_{HN}, respectively, is in some zones not well reproduced. An analysis of the geometries obtained during the reaction shows that when the approximation of one residue is not energetically favorable, it is compensated in the R_4 reaction coordinate combination by the approximation of the conjugated residue whose movement is more labile. For example, when the first proton abstraction by Lys166 to the substrate must take place, His297 which is not coordinated to any residue and whose movement is rather free tends to approximate to the (*S*) substrate in order to compensate in the four-distance combination (R_4) the more expensive proton abstraction by Lys166. After the proton transfer has taken place, His297 goes slightly back forward to its original position.

This situation is repeated more remarkably in the *R side* with the Lys166 free movement where this residue artificially approaches the substrate in order to compensate the energetically more expensive proton transfer between the (*R*)-mandelate substrate and His297. This is why residues Lys166 and His297 go forward and backward during the complementary proton transfer in which they do not participate actively, and the corresponding proton transfer occurs suddenly.

Switching Functions. To solve the above problem, we introduced two switching functions to the R_4 reaction coordinate to activate/deactivate the appropriate asymmetric stretch combination.

$$R_{\text{switch}} = f_S(R_{\text{NHC}} + R^S_{\text{CHN}}) + f_R(R_{\text{CHN}} + R^R_{\text{NHC}}) \qquad (2)$$

where $R^S_{\text{CHN}} = -1.93$ and $R^R_{\text{NHC}} = 1.86$ are the constant values of the combination of distances at the reactants (*S*) and products (*R*), respectively, and they are taken from the stationary points shown in Table 2.

The switching functions f_S and f_R should vanish in the reciprocal zone (*R* and *S*, respectively) and should approximate to 1 near their corresponding zone. A possible expression for these switching functions is

$$f_S(R_4) = \frac{1}{2}(1 + \tanh(-p_s(R_4 - c_s))) \qquad (3)$$

$$f_R(R_4) = \frac{1}{2}(1 + \tanh(p_r(R_4 - c_r))) \qquad (4)$$

To normalize the combination ($f_R + f_S = 1$), we set the constants $c_s = c_r$ and $p_s = p_r$. From the chemical point of view, the crossing point ($c_{(s,r)}$) is the value in R_4 where one of the two asymmetric combinations ($R_{(\text{NHC,CHN})}$) becomes more important than the other. The slope ($p_{s,r}$) describes how fast the switch must take place. Using the stationary point data in Table 2 and after some tests, we saw that an adequate choice is $c_s = 0.0$ and $p_s = 0.8$.

Then, the PMF was recomputed using R_{switch} as the reaction coordinate. The simulation conditions used before were also employed here. A total of 14 windows were run between a range from $R_{\text{switch}} = -2.4$ to $R_{\text{switch}} = 2.4$. In Figure 9, the PMF using the reaction coordinate R_{switch} is displayed. With this new reaction coordinate, we observe some relevant changes with respect to PMF(R_4). The first one is that the two proton transfers are now described smoothly; mainly, the proton transfer between

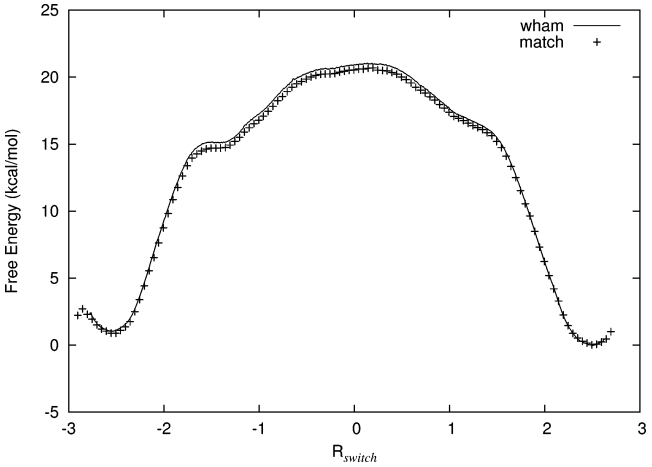
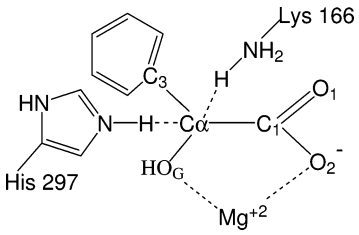


Figure 9. Free energy profile using R_{switch} as defined in eq 2.

the substrate and His297 is now more detailed. As we can see in Figure 9, the free energy profile coming from the WHAM analysis coincides with the results obtained matching the different windows. This indicates that we do not have the hysteresis problem observed using R_4. As a result of this improvement, the reaction becomes more symmetric, in the sense that both reactants and products have the same free energy. This is exactly what we should expect from a racemase, where the two reactions have similar rates. In addition, using the switching functions, there is no artificial compensation in the combination of distances and the evolution of bond distances does not change brusquely. The two proton transfers take place in separate ranges of the reaction coordinate where only their corresponding asymmetric stretch contributes to R_{switch}, while in the configuration inversion zone, around zero values of R_{switch}, the four bond distances are involved.

3.3. Reaction Mechanism. At this point, once the free energy is adequately computed, it permits us to look into the chemistry of the reaction. For the reaction mechanism, the first thing that must be pointed out is that when comparing the PMF profile and the previous results from the PES optimization, the energetics and the evolution of the main bond distances are very similar. On the free energy basis, the mechanism is concerted but asynchronous. Lys166 and His297 play their acid/base catalytic role at different stages of the reaction. The nature of the intermediate species is even less relevant in the PMF profile than in the potential energy profile. However, this similarity between PMF and PES must not guide the erroneous conclusion that the present dynamic study provides the same information as did our previous work; both techniques must be regarded as complementary rather than competitive. The nature of enzymatic catalysis is inherently dynamic, and from both energetic and structural points of view, some relevant differences have sometimes been encountered.[57,58] In addition, the following dynamic analysis permits the study of averaged residue contributions that in the PES optimization study would not have been possible.

SCHEME 3



At this point, we want to analyze the individual residue contributions to the reaction mechanism. With this objective, we carried out two strategies that should provide complementary information. The first strategy is geometry analysis, and the second one is the electrostatic energy perturbation analysis.

Geometry Analysis. To analyze the residue contributions through a geometric analysis, we ran molecular dynamics restrained on the critical points of our reaction using the same bias potential that has been used during the PMF calculation. In particular, different 50 ps runs of MD simulations were performed at the reactants, products, the transition state, and the I_s and I_r intermediate point regions. A test run of 100 ps showed that a longer MD did not change the results substantially. Every 10 fs, the coordinates were saved, and the relevant bond distances have been analyzed. A more restricted analysis was performed taking only those structures with a reaction coordinate that corresponded more closely (±0.05 Å) to the value of the PMF. However, the results obtained in this latter case were qualitatively the same as those obtained analyzing all of the structures in a simulation.

In Table 5, we show the averaged bond distances obtained from the simulation at different zones. For the hydrogen bond interactions, the percentages of occupancy (%, in brackets) are also shown. The criteria used to compute the occupancy is when the distance between the hydrogen and the acceptor atom is lower than 2.5 Å. The labels used for the substrate atoms are displayed in Scheme 3, while the labels for the residue atoms are the standard ones.

In some cases, there is a large root-mean-square deviation with respect to the averaged values. These are some loose hydrogen bonds that form and break several times during the dynamics. This is why we find some hydrogen bond interaction with a longer distance and with lower occupancy.

The set of distances given in Table 5 can be classified depending on which zone of the reaction they stabilize the most (the distance becomes shorter). Thus, in general, the distances that involve His297 stabilize the reactants, while the distances that involve the substrate stabilize the TS. Distances involving Lys166 stabilize both the TS and the products. This observed tendency should be expected taking into account that His297 has a positive charge in the reactants, and therefore, the

TABLE 5: Averaged Bond Distances at Different Points of the Reaction (for the Hydrogen Bond Interactions, the Values in Parentheses Specify the Occupancy in %)

pair of interaction	reactant (*S*)	I_s	TS	I_r	product (*R*)
His297(H_ϵ)−Glu247(O_ϵ)	2.55 (63)	2.84 (18)	2.88 (6)	2.93 (4)	3.15 (0.4)
His297(H_δ)−Asp270(O_ϵ)	1.78 (100)	1.72 (100)	1.87 (99)	1.84 (99)	2.27 (84)
O_1−Glu317(H_ϵ)	1.76 (100)	1.73 (100)	1.72 (100)	1.72 (100)	1.77 (100)
O_2−Lys164(H_z)	1.88 (99)	1.79 (100)	1.78 (100)	1.81 (100)	1.79 (100)
O_1−Ser139(H_γ)	1.98 (99)	1.90 (100)	1.79 (100)	1.79 (100)	2.10 (96)
O_2−Mg	1.87	1.85	1.83	1.84	1.86
O_G−Mg	2.20	2.16	2.10	2.13	2.15
Lys166(H_z)−Asp195(O_δ)	3.05 (1)	2.77 (17)	2.69 (41)	2.75 (32)	2.85 (18)
Lys166(H_z)−Asn197(O_δ)	3.88 (0.4)	2.05 (99)	2.26 (90)	2.81 (61)	2.40 (73)

negatively charged residues near His297 will stabilize it at this stage of the reaction. This is the case for residues Glu247(O_ϵ) and Asp270(O_ϵ) which have a hydrogen bond interaction with His297 that becomes shorter at reactants.

In addition, Asp270 has been proposed to form a catalytic dyad with His297 that should modulate its pK_a,[15] but in the present QM/MM model, we cannot observe this possibility. The same tendency can be observed with Lys166. This residue has a positive charge at products, and the residues nearby capable of interacting with it will have a shorter distance at the products than in the reactants. This is the case for Asp195(O_δ) and Asn197(O_δ). Some mutagenesis experiments[16] indicate the important role of Asn197 in the catalysis. We carried out molecular dynamics on a model of the N197A mutant, and we saw that on average Lys166 without the interaction with Asn197 changes its conformation and does not point to the substrate anymore.

As a result of these tendencies, the function of Asp195 and Asn197 at products and Glu247 and Asp270 at reactants seems to be twofold. The first is to keep the catalytic residues Lys166 and His297 at the appropriate orientation to abstract the nonacid proton from the substrate, and the second function is to modulate their pK_a through hydrogen bond interactions. As a result of these two facts, the enzyme modulates adequately the two-direction reaction to make them at a similar rate and therefore to be symmetric and obtain the racemic mixture.

The bond distances shown in Table 5 that stabilize the TS are due to different types of interactions. In general, at the TS region, the active site has a maximum of negative charge concentrated on the substrate that the enzyme has to withdraw. A possibility of withdrawing charge is through the hydrogen bond interaction between the carboxylic group of the substrate and Glu317, Lys164, and the unexpected but conserved in the enolase family Ser139. The stabilization of Glu317 and Lys164 should not be confused with easy proton donation or the so-called low barrier hydrogen bond. In a previous article,[20] we saw that proton transfer between these two residues and the substrate was not easy and therefore not relevant to the catalysis.

A second zone in the active site where the TS is stabilized is around the magnesium cation coordination sphere. In particular, magnesium is bicoordinated to the substrate, and during the reaction ($S \rightarrow R$), it approaches 0.04 Å to the carboxylic oxygen O_2 and 0.1 Å to the hydroxyl oxygen O_G. Although the Mg shift does not seem important, we will see the relevance of its effect from the electrostatic point of view when compared to an uncatalyzed reaction.

The ligands coordinated to the magnesium cation (Asp195, Glu221, Glu247, and a water molecule; see Figure 1) do not modify significantly their interaction distance during the chemical reaction. In our previous studies,[19,20] we selected a bigger quantum part where the sphere of ligands was included and the distances to the Mg did not change either. Therefore, this cannot be attributed to a lack of electronic charge redistribution in this new QM/MM selection.

Perturbation Analysis. The stabilization interactions discussed above are based on the geometry of local bond distances, but this will not give a picture of the long range electrostatic contributions. Electrostatics has already been pointed out as the major factor that contributes to the enzymatic catalysis.[23,24] In this sense, a strategy to identify the electrostatic contribution of every residue is the electrostatic energy perturbation analysis.[59,60] This approach has been used several times to identify the major electrostatic contributions that stabilize the TS with respect to the reactants or the products. The process consists of

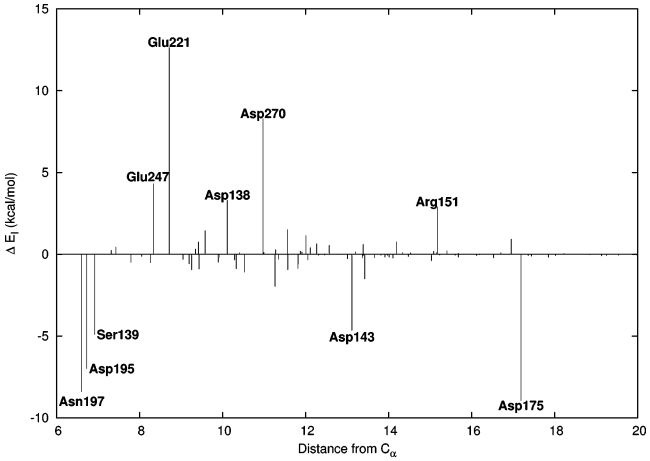


Figure 10. Electrostatic perturbation analysis of the TS structures with respect to the reactants.

progressively annihilating the partial charges of each residue starting from those closest to the active site and then converging again the QM wave function and calculating the QM/MM energy difference with respect to the previous annihilation. Thus, in a system of M residues, for residue I, we have annihilated the 1, ... , I residue charges and only the $I + 1$, ... , M residues contribute to the polarization of the wave function in such a way that we can define

$$E_I = \langle \Phi | \hat{H}_{el} | \Phi \rangle - \sum_i^{\text{electr}} \sum_m^{\text{classical}} \left\langle \Phi \left| \frac{q_m}{r_{im}} \right| \Phi \right\rangle + \sum_j^{\text{nuclei}} \sum_m^{\text{classical}} \frac{Z_j q_m}{r_{jm}} \quad (5)$$

where q_m are the classical point charges belonging to the residues from $I + 1$ to M and i and j correspond to the electrons and nuclei, respectively, of the QM part. The energy difference is evaluated, for example, at the TS and at reactants. The comparison will give us the electrostatic contribution of residue I in the TS with respect to the reactants:

$$\Delta E_I(\text{TS vs reactants}) = (E_I - E_{I-1})_{\text{TS}} - (E_I - E_{I-1})_{\text{reactants}} \quad (6)$$

The residues with negative ΔE_I(TS vs reactants) values will stabilize the TS in a major contribution as compared with what they do in the reactants. Since we eliminate progressively the screening effect of the closest residues, the magnitude of the interaction will be artificially amplified; otherwise, we would only obtain a smooth decay that would not give any useful information.

From the MD runs used above, we selected 40 snapshots with a reaction coordinate value corresponding to the reactants, TS, and products. At each structure, the electrostatic perturbation analysis has been performed and the results have been averaged over the whole analysis of 40 snapshots. In particular, the order in which we selected the residues to zero out their charges is based on the distance between the C_α of the residue and the C_α of the stereogenic center in mandelate substrate until 20 Å.

In Figures 10 and 11, the analysis of the TS versus reactants and the products versus reactants, respectively, are shown. (Note that we did not include the Mg cation in the perturbation analysis.)

Looking at the closest residues to the active site in Figure 11, we obtain an assessment of what has been said in the bond distance analysis; that is, Asp195 and Asn197 stabilize the products through the interaction with the positively charged

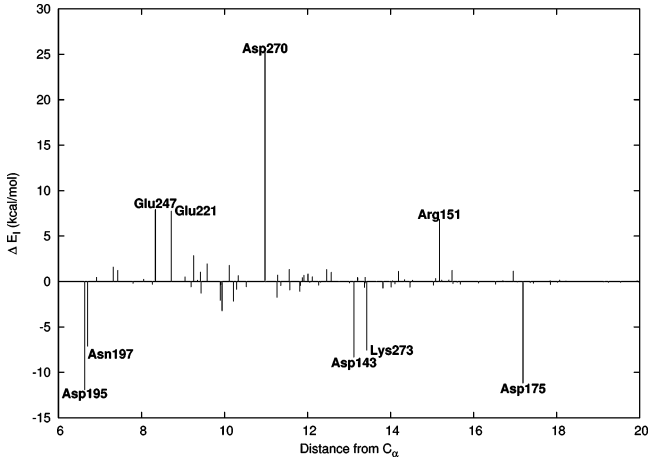


Figure 11. Electrostatic perturbation analysis of the product structures with respect to the reactants.

Lys166, and Glu247 and Asp270 stabilize the reactants interacting with the positive His297 (note the change of scale in Figures 10 and 11 to remark the stabilization of the reactants due to Asp270). In addition, it is necessary to comment on the contributions from some other residues.

Coordinated to the Mg and hydrogen bonded to Lys164, there is Glu221 which stabilizes remarkably the reactants with respect to the TS and, to a minor extent, the products. To understand the Glu221 effect, we must point out that this residue due to its coordination to Lys164 and Mg is relatively close to the negative carboxylate group of the substrate. As we saw in Table 5, Mg approaches the substrate at the TS. As a result of that, Glu221 follows Mg, approaches the substrate, and interacts unfavorably with the carboxylate group at the TS and with less intensity at the products.

Finally, there are some residues considerably far away from the active site whose electrostatic effect is notable. This is the case of Asp143, Arg151, Asp175, and Lys273. To understand their effect, we must only see their natural charge (positive or negative) and where they are located (in the Lys166/pro-*S* side or in the His297/pro-*R* side). Since the $S \rightarrow R$ reaction has as a net yield the translocation of a positive charge from the pro-*R* side to the pro-*S* side, the negative residues located at the pro-*S* side (Asp143 and Asp175) will stabilize the TS and the product (recall the productlike character of the TS). In turn, positive Arg151 at the pro-*S* side will stabilize the reactants. The same argument can be used for Lys273 located at the pro-*R* side whose positive charge stabilizes the product.

As a result of the above discussion, the electrostatic interaction achieves a balanced stabilization of reactants and products that has as a consequence obtaining the racemic mixture, that is, a similar rate when going in the $S \rightarrow R$ and $R \rightarrow S$ directions.

3.4. Contributions to Catalysis. So far, we have discussed the reaction mechanism of mandelate racemase through qualitative analysis that helps one to understand the different residue contributions to the reaction. However, there is still a lack of information referring to the quantitative participation of the enzyme in the reaction. In other words, we want to know how much the enzyme contributes to the catalysis with respect to an uncatalyzed reaction.

The first problem in doing such an analysis is to design a molecular model of the uncatalyzed reaction that keeps the essential factors of the chemical reaction and removes the particular enzyme environment. Building an uncatalyzed model is a very common strategy,[61,62] and it serves as a reference reaction, rather than to reproduce an aqueous reaction, to quantify the enzyme participation in the chemical reaction.

When the enzyme only participates in solvating the substrate without any bond formation, the design of an uncatalyzed model is straightforward: it is built by replacing the enzyme by water molecules. Conversely, there are enzymatic reactions where the enzyme participates actively in the mechanism forming and breaking bonds. In these cases, the model for an uncatalyzed reaction is not unique. This is the case in mandelate racemase where Lys166 and His297 remove and provide essential protons to the substrate. Therefore, if we want to consider the same reaction without the enzyme, we cannot remove the amino group of Lys166 nor the imidazole group of His297 and replace it by water; otherwise, the essence of the reaction would change.

In the present case, we must consider two separate uncatalyzed models accounting for the two halves of the reaction. The reason to split the system into two halves is because we do not have a reaction coordinate that could describe the whole process with the substrate and the two well oriented residue models without freezing additional degrees of freedom.

The first model, labeled as model A, is neutral propylamine (modeling Lys166) and (*S*)-mandelate substrate in a cubic box of waters of 38.0 Å. This model stands for emulating the uncatalyzed first half of the enzymatic reaction where Lys166 removes the pro-*S* proton from the substrate. The system is set up and heated with the same protocol as the enzymatic case with the exception that now periodic bound conditions and a Nose−Hoover thermostat are substituted for the stochastic boundary conditions. The second uncatalyzed model, labeled as model B, stands for the second half of the enzymatic reaction where His297 protonates the pro-*R* side of the dianion substrate. It is modeled by a neutral methyl imidazole and (*R*)-mandelate substrate surrounded by a cubic box of waters and by performing the same simulation procedure that has been used in the other uncatalyzed model.

Free Energies. We have computed the PMFs for the two uncatalyzed model reactions with the same umbrella sampling method described above. The reaction coordinate R_{NHC} as defined in Table 2 has been used for the model A reaction, while R_{CHN} has been used for the model B reaction.

In Figure 12, the free energy profiles for model A (left) and model B (right) are shown and compared with the same profile in the enzyme. As discussed above, the reaction coordinates used are only able to describe from the minimum until the flat regions (I_s and I_r). The configuration inversion corresponding to the highest energy point is only partially described, and therefore, its energy value must be analyzed with care.

The comparison gives a catalytic effect of 6.1 kcal/mol ($R_{NHC} = 0.435$ Å) for model A at the I_s region and 5.9 kcal/mol ($R_{CHN} = -0.425$ Å) for model B at the I_r region. We can see that the enzyme catalyzes the proton abstraction substantially and makes the proton abstraction of a nonacid carbon easier. The significant difference in the free energies shows us that models A and B are adequate models for removing the catalytic effect and therefore are valid for the following comparisons.

Electronic Indices. Charge and Bond Order. We want to take advantage of the polarized QM/MM model used here to analyze the quantum wave function and compare it to the uncatalyzed reaction. To do so, we have computed the Mulliken net charges and the bond order of some relevant atoms and bonds in the reactive part treated with quantum mechanics. The very well-known expressions for these magnitudes are the following.

$$q_A = Z_A - \sum_{\mu \in A} (PS)_\mu; \quad B_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} (PS)_{\mu\nu}(PS)_{\nu\mu} \quad (7)$$

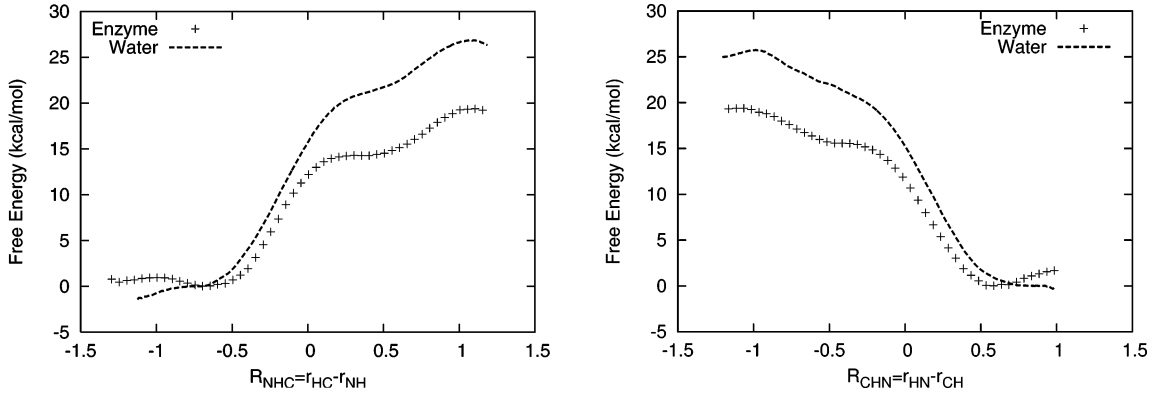


Figure 12. Potential of mean force for the uncatalyzed and enzyme reactions. On the left, model A (propylamine + mandelate) emulates the abstraction by Lys166. On the right, model B (methylimidazole + mandelate) emulates the abstraction by His297.

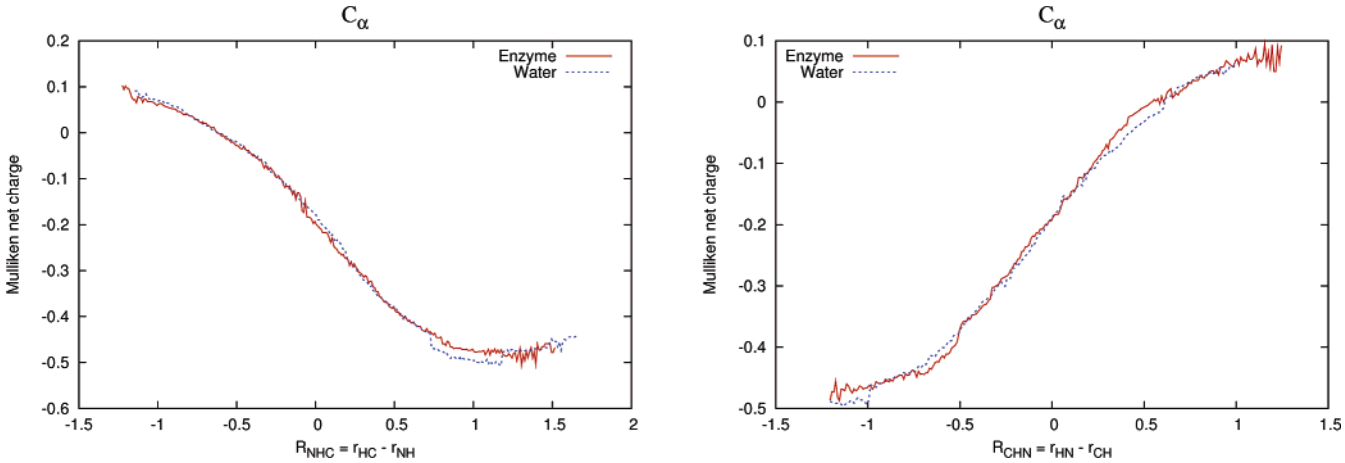


Figure 13. Comparison of the evolution of the Mulliken net charge in C_α in the enzyme and the uncatalyzed reactions model A (left) and model B (right).

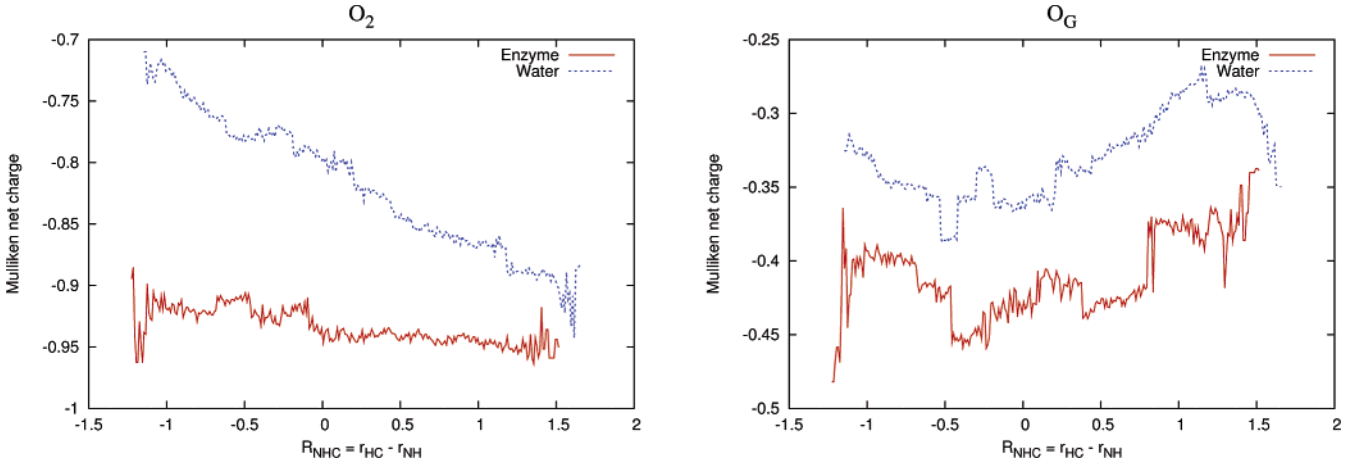


Figure 14. Comparison of the evolution of the Mulliken net charges in the enzyme and the uncatalyzed reactions model A for oxygen atom O_2 (left) and O_G (right).

where q_A is the Mulliken net charge of atom A, B_{AB} is the bond order between atoms A and B, P is the density matrix, and S is the atomic overlap matrix.

The Mulliken net charges and bond orders are calculated at the same PM3 semiempirical level and therefore considering that the overlap matrix (S) is equal to the identity matrix. Despite the lack of rigor of these calculations, the following analysis already gives a useful value of each magnitude and permits an interesting comparison between the enzyme and model A and B reactions.

In Figures 13 and 14, we display the evolution of some Mulliken net charges all along the reaction. The labels used for the substrate atoms are indicated in Scheme 3. The plot shows a continuous representation where there is a charge at every reaction coordinate value (±0.01 Å). The simulation is done running 5 ps of every window used in the PMF calculation printing out the charges, and the analysis has been performed averaging the charges of every structure that fall in a bin in the total density distribution ($\rho(Rc)$). When the evolution of the charge defines a tendency, it must be understood that there is a

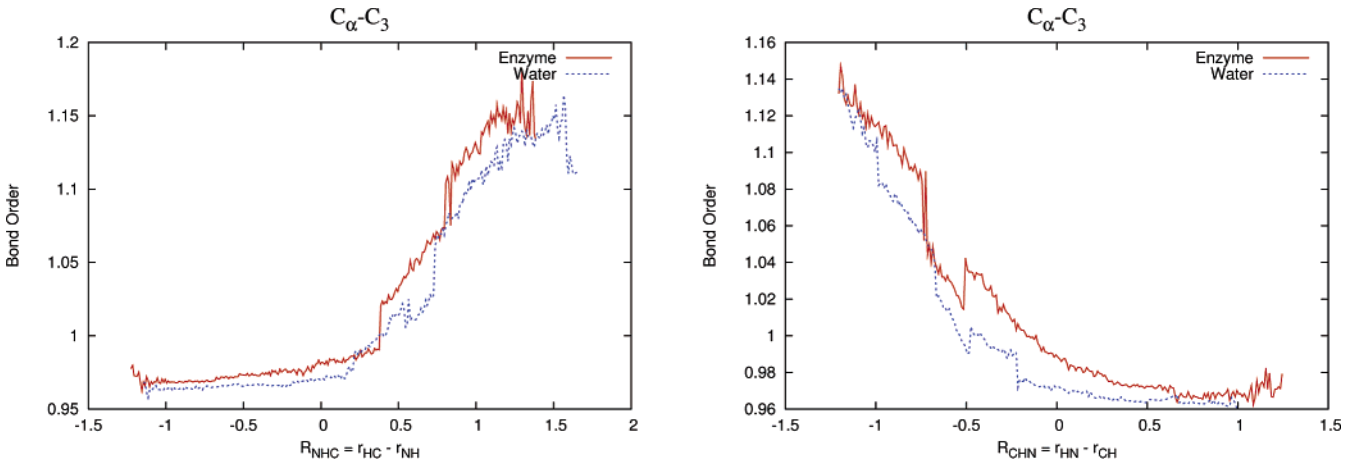


Figure 15. Comparison of the evolution of the $C_\alpha-C_3$ bond order in the enzyme and the uncatalyzed reactions model A (left) and model B (right).

correlation between the charge and the reaction coordinate (the rest of the plots not displayed here can be found in the Supporting Information).

As can be seen in Figure 13, and contrary to what we should expect, the net charge on the C_α during the reaction is the same as that in the uncatalyzed reaction. The net charges on the nitrogen heteroatoms that abstract the proton from the substrate (Figures 9 and 10 in the Supporting Information) have also insignificant differences between the enzyme and the uncatalyzed reaction. Therefore, the catalytic effect is not based on the act of withdrawing electronic charge from the atoms directly involved in the proton transfers.

Definitely, the most remarkable difference from this charge analysis is the net charge on O_2 and O_G which are coordinated to the magnesium in the enzyme and have no coordination element in the uncatalyzed models (see Figure 14). While in the uncatalyzed reaction the carboxylate group is symmetric and the net charge on O_2 becomes more negative as the α-proton is abstracted, in the enzymatic reaction, the net charge is approximately constant and much more negative (~ -0.9). Although we only display the results for model A, the same behavior is observed for model B. This means that Mg^{2+} cation attracts the negative charge during the whole reaction process.

The same study can be applied on bond orders. This property can illustrate how the electronic charge is redistributed in a highly resonant system such as our substrate. In Figure 15, the bond order between α-carbon and C_3 is displayed for the enzyme and the uncatalyzed reactions. We see that, in most parts of the reaction, mainly after the substrate undergoes proton abstraction, the $C_\alpha-C_3$ bond order is higher in the enzyme, indicating a major delocalization in the phenyl ring in the case of the enzyme. A similar tendency in the $C_\alpha-C_1$ bond order is observed (Figures 17 and 18 in the Supporting Information). These cases of major resonance must be interpreted as the capacity of the enzyme through both electronic and geometric effects to delocalize the excess of electronic charge all throughout the substrate.

Once again, in accordance with the charge analysis, the most notorious effect is the role played by the magnesium cation. In particular, the carboxylate group in the enzyme is not symmetric (Figures 13−16 in the Supporting Information). For example, while in the uncatalyzed reaction both C_1-O_1 and C_1-O_2 bond orders have a similar value of 1.2 for the anionic substrate and 1.4 for the dianionic one, in the enzyme, C_1-O_1 is 1.6 and C_1-O_2 is 1.2 in the anionic substrate and 1.4 and 1.2, respectively, for the dianionic one. Therefore, the rather single character of the C_1-O_2 bond in the enzyme does not change during the enzymatic reaction, confirming the constant electrostatic interaction of Mg cation.

4. Conclusions

In this article, we carry out molecular dynamics simulations on a reliable QM/MM model of mandelate racemase enzyme. Although the reversible interconversion of the (*S*) and (*R*) enantiomers of mandelate anion consists basically of two proton transfers and a configuration inversion of a stereogenic carbon atom, the selection of a reaction coordinate needed to compute the potential of mean force is not straightforward. The reaction mechanism consists of a single step where several residues are involved at different stages. Using as a guide our previous work where the mechanism was studied with optimization methods, we try several reaction coordinates to compute the free energy profile and design a final combination of bond distances where the role of the relevant residues is activated/deactivated by means of switching functions.

The reaction mechanism is further studied by analyzing the different residue contributions along the reaction through averaged bond distances and electrostatic perturbation analysis. We conclude that the enzyme stabilizes the reactants and the products through local hydrogen bond interactions and short and long range electrostatic contributions. The balanced stabilization of reactants and products has as a consequence the design of a symmetry reaction where both (*S*)-mandelate and (*R*)-mandelate substrates react at a similar rate and therefore the racemic mixture is achieved.

In addition, we examine the catalytic effect by building two uncatalyzed models. The free energy profiles for the enzyme and uncatalyzed reaction are compared. In addition, the evolution of Mulliken atomic net charges and bond orders on the active site permits the electronic properties of the substrate during the catalyzed and uncatalyzed reactions to be displayed. The role of the enzyme in stabilizing the transition state consists of removing electronic charge from the substrate. This is accomplished mainly by the electrostatic interaction of the magnesium dication and by the stabilization of the hydrogen bonded residues Glu317 and Lys164.

Acknowledgment. We are grateful for the financial support from the Spanish "Ministerio de Educación y Ciencia" and the "Fondo Europeo de Desarrollo Regional" through Project No. BQU2002-00301 and the use of computational facilities of the CESCA.

Supporting Information Available: Plots showing the evolution of the Mulliken net charges and bond orders in the

enzyme when compared to an uncatalyzed reaction. This material is available free of charge via the Internet at http://pubs.acs.org.

Note Added after ASAP Publication. The left panel of Figure 12 has been corrected. This paper was published ASAP on 10/15/05. The corrected version was reposted on 10/19/05.

References and Notes

(1) Kenyon, G. L.; Gerlt, J. A.; Petsko, G. A.; Kozarich, J. W. *Acc. Chem. Res.* **1995**, *28*, 178−186.
(2) Gerlt, J. A.; Kozarich, J. W.; Kenyon, G. L.; Gassman, P. G. *J. Am. Chem. Soc.* **1991**, *113* (25), 9667−9669.
(3) Mitra, B.; Kallarakal, A. T.; Kozarich, J. W.; Gerlt, J. A.; Clifton, J. R.; Petsko, G. A.; Kenyon, G. L. *Biochemistry* **1995**, *34* (9), 2777−2787.
(4) Bearne, S. L.; Wolfenden, R. *Biochemistry* **1997**, *36* (7), 1646−1656.
(5) Maurice, M. S.; Bearne, S. L. *Biochemistry* **2004**, *43* (9), 2524−2532.
(6) Goriup, M.; Strauss, U. T.; Felfer, U.; Kroutil, W.; Faber, K. *J. Mol. Catal. B* **2001**, *15*, 207−212.
(7) Goriup, M.; Strauss, U. T.; Felfer, U.; Kroutil, W.; Faber, K. *J. Mol. Catal. B* **2001**, *15*, 213−222.
(8) Schnell, B.; Faber, K.; Kroutil, W. *Adv. Synth. Catal.* **2003**, *345*, 653−666.
(9) Wolfenden, R.; Snider, M. J. *Acc. Chem. Res.* **2001**, *34*, 938−945.
(10) Whitman, C. P.; Hegeman, G. D.; Cleland, W. W.; Kenyon, G. L. *Biochemistry* **1985**, *24* (15), 3936−3942.
(11) Gerlt, J. A.; Babbitt, P. C.; Rayment, I. *Arch. Biochem. Biophys.* **2005**, *433*, 59−70.
(12) Landro, J. A.; Gerlt, J. A.; Kozarich, J. W.; Koo, C. W.; Shah, V. J.; Kenyon, G. L.; Neidhart, D. J.; Fujita, S.; Petsko, G. A. *Biochemsitry* **1994**, *33*, 635−643.
(13) Kallarakal, A. T.; Mitra, B.; Kozarich, J. W.; Gerlt, J. A.; Clifton, J. R.; Petsko, G. A.; Kenyon, G. L. *Biochemistry* **1995**, *34* (9), 2788−2797.
(14) Neidhart, D. J.; Howell, P. L.; Petsko, G. A.; Powers, V. M.; Li, R.; Kenyon, G. L.; Gerlt, J. A. *Biochemistry* **1991**, *30* (38), 9264−9273.
(15) Schafer, S. L.; Barrett, W. C.; Kallarakal, A. T.; Mitra, B.; Kozarich, J. W.; Gerlt, J. A. *Biochemistry* **1996**, *35* (18), 5662−5669.
(16) Maurice, M. S.; Bearne, S. L. *Biochemistry* **2000**, *39* (44), 13324−13335.
(17) Cleland, W.; Frey, P.; Gerlt, J. *J. Biol. Chem*. **1998**, *273* (40), 25529−25532.
(18) Li, R.; Powers, V. M.; Kozarich, J. W.; Kenyon, G. L. *J. Org. Chem.* **1995**, *60* (11), 3347−3351.
(19) Garcia-Viloca, M.; González-Lafont, A.; Lluch, J. M. *J. Am. Chem. Soc.* **2001**, *123*, 709−721.
(20) Prat-Resina, X.; Garcia-Viloca, M.; González-Lafont, A.; Lluch, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5365−5371.
(21) Prat-Resina, X.; González-Lafont, A.; Lluch, J. M. *THEOCHEM* **2003**, *632*, 297−307.
(22) Prat-Resina, X.; Bofill, J. M.; González-Lafont, A.; Lluch, J. M. *Int. J. Quantum Chem.* **2004**, *98* (4), 367−377.
(23) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.
(24) Villà, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 7887−7907.
(25) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(26) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11* (6), 700−733.
(27) Field, M. J. *J. Comput. Chem.* **2002**, *23* (1), 48−58.
(28) Fukui, K. *Acc. Chem. Res.* **1981**, *14* (12), 363−368.
(29) Schlegel, B. *J. Comput. Chem.* **2003**, *24* (12), 1515−1527.
(30) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184−189.
(31) Kollman, P.; Kuhn, B.; Donini, O.; Perakyla, M.; Stanton, R.; Bakowies, D. *Acc. Chem. Res.* **2001**, *34* (1), 72−79.
(32) Lu, Z.; Yang, W. *J. Chem. Phys.* **2004**, *121*, 89−100.
(33) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(34) Alagona, G.; Ghio, C.; Kollman, P. A. *THEOCHEM* **1997**, *390*, 217−223.
(35) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4* (2), 187−217.
(36) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220.
(37) Guallar, V.; Friesner, R. A. *J. Am. Chem. Soc.* **2004**, *126* (27), 8501−8508.
(38) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102* (24), 4714−4721.
(39) Garcia-Viloca, M.; Gao, J. *Theor. Chem. Acc.* **2003**, *111*, 280−286.
(40) 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. B* **1998**, *102* (18), 3586−3616.
(41) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79* (12), 6312−6325.
(42) Brunger, A.; Brooks, C. L., III; Karplus, M. *Chem. Phys. Lett.* **1984**, *105* (5), 495−500.
(43) Brooks, C. L., III; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843−865.
(44) Brooks, C. L., III; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159−181.
(45) Alhambra, C.; Gao, J. *J. Comput. Chem.* **2000**, *21*, 1192−1203.
(46) Garcia-Viloca, M.; Alhambra, C.; G.Truhlar, D.; Gao, J. *J. Comput. Chem.* **2003**, *24*, 177−190.
(47) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(48) Rodriguez-Gomez, D.; Darve, E.; Pohorille, A. *J. Chem. Phys.* **2004**, *120*, 3563−3578.
(49) Berne, B.; Straub, J. *Curr. Opin. Struct. Biol.* **1997**, *7* (2), 181−189.
(50) Tai, K. *Biophys. Chem.* **2004**, *107* (3), 213−220.
(51) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(52) Gao, J. *J. Am. Chem. Soc.* **1991**, *113* (20), 7796−7797.
(53) Bartels, C.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 865−880.
(54) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13* (8), 1011−1021.
(55) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; John Wiley & Sons: New York, 1991.
(56) Pauling, L. *J. Am. Chem. Soc.* **1947**, *69*, 542.
(57) Gao, J.; Truhlar, D. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.
(58) Billeter, S. R.; Hanser, C. F. W.; Mordasini, T. Z.; Scholten, M.; Thiel, W.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2001**, *3*, 688−695.
(59) Bash, P.; Field, M.; Davenport, R.; Petsko, G.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30* (24), 5826−5832.
(60) Cui, Q.; Karplus, M. *J. Am. Chem. Soc.* **2001**, *123* (10), 2284−2290.
(61) Soriano, A.; Silla, E.; Tuñón, I.; Ruiz-Lopez, M. F. *J. Am. Chem. Soc.* **2005**, *127*, 1946−1957.
(62) Nam, K.; Prat-Resina, X.; Garcia-Viloca, M.; Devi-Kesavan, L. S.; Gao, J. *J. Am. Chem. Soc.* **2004**, *126*, 1369−1376.