Theor Chem Acc (2011) 129:781–791
DOI 10.1007/s00214-011-0935-7

REGULAR ARTICLE

QM/MM study on catalytic mechanism of aspartate racemase from *Pyrococcus horikoshii* OT3

Chenghua Zhang · Yong Guo · Ying Xue

Received: 7 January 2011 / Accepted: 29 March 2011 / Published online: 16 April 2011

Abstract The enzyme aspartate racemase from *Pyrococcus horikoshii* OT3 catalyzes the interconversion between L- and D-Asp. In this work, we employed the hybrid QM/MM approach with the self-consistent charge-density functional tight binding (SCC-DFTB) model to study the catalytic mechanism for the conversion of L-Asp into D-Asp. The molecular dynamics simulation showed that the substrate L-Asp forms an extensive network of interactions with the active-site residues of the aspartate racemase through its side chain carboxylate, ammonium group, and α-carboxylate. The potential of mean force calculations confirmed that the racemization reaction involves two proton transfers (from the α-carbon to Cys194 and from Cys82 to the α-carbon), which occurs in a concerted way, although highly asynchronous. The calculated free energy of activation is 17.5 kcal/mol, which is consistent with the reaction rate measured from experiment. An electrostatic interaction analysis was performed to estimate the key role played by individual residues in stabilizing the transition state. The docking study on the binding of L-Asp and D-Asp to aspartate racemase indicates that this enzyme employs a "two-base" mechanism not a "one-base" mechanism.

Keywords Aspartate racemase · Molecular dynamics simulation · QM/MM · Potential of mean force · Docking

C. Zhang · Y. Guo · Y. Xue (✉)
College of Chemistry, Key Laboratory of Green Chemistry and Technology in Ministry of Education, Sichuan University, 610064 Chengdu, People's Republic of China
e-mail: yxue@scu.edu.cn

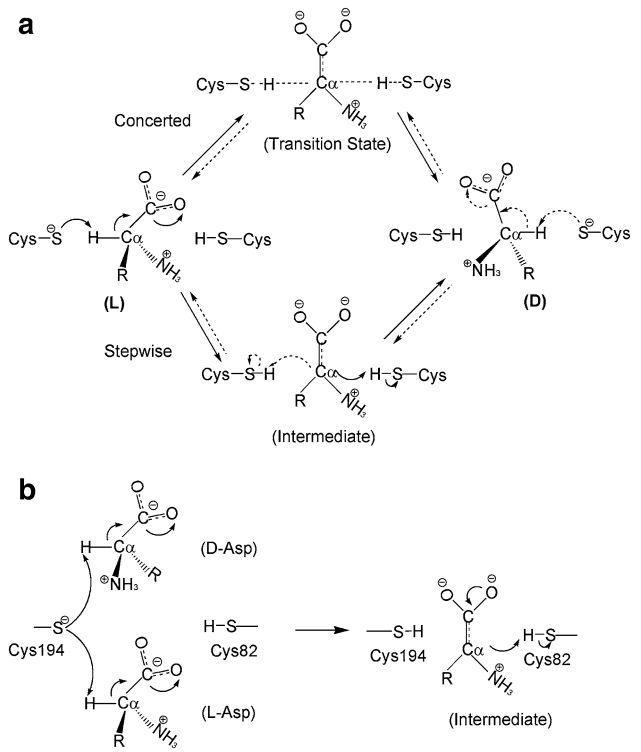
Y. Xue
State Key Laboratory of Biotherapy, Sichuan University, 610064 Chengdu, People's Republic of China

1 Introduction

L-amino acids are predominant in living organisms, and the D-amino acids are also distributed ubiquitously in eubacteria as components of peptidoglycan in cell walls [1]. D-amino acids play crucial roles in rigidifying the bacterial cell walls and in regulation of the specific activities as mediators or transmitters in mammalian nervous and endocrine systems. In particular, D-aspartic acid is of interest because it has been found at substantial levels in bacteria, archaea, and mammals [2–5].

The enzyme aspartate racemase (AspR, EC 5.1.1.13) catalyzes the interconversion between L- and D-Asp [6] and it belongs to an interesting subgroup of the large amino acid racemase (AAR) family. AARs are reported to employ a mechanism that involves deprotonation of the α-carbon of the substrate amino acid followed by reprotonation of the resulting anionic intermediate from the reverse side and inversion of the configuration [7]. AARs can be classified into two groups: the pyridoxal 5'-phosphate (PLP)-dependent and PLP-independent groups [8]. AspR belongs to the PLP-independent group along with glutamate racemase (GluR, EC 5.1.1.3) [9], proline racemase (ProR, EC5.1.1.4) [10], diaminopimelate epimerase (DapE, EC 5.1.1.7) [11], and 4-hydroxyproline racemase (EC5.1.1.8) [12]. The racemization catalyzed by the PLP-independent AARs was thought to take place through a "two-base" mechanism and use cysteine residues as the conjugated catalytic acid and base [13–17]. The thiolate group of one cysteine acts as a catalytic base by deprotonation of the substrate to form an anionic intermediate, and subsequently the thiol group of the other cysteine residue acts as a catalytic acid by reprotonation of the intermediate (Scheme 1).

The first crystal structure of AspR from *Pyrococcus horikoshii* OT3 (PhAspR) at a resolution of 1.9 Å was



Scheme 1 The proposed "two-base" racemization mechanisms of PLP-independent AARs (**a**), and the hypothetical "one-base" mechanism of PhAspR from Ref 20 (**b**)

determined by Liu and his co-workers [18]. This enzyme has a preferred temperature up to 90 °C, and it does not exhibit any activity at room temperature [3]. PhAspR adopts a dimeric form in crystals and in solution, the monomer consists of two compact α/β domains, which are highly homologous. Analysis of the dimeric structure of PhAspR revealed that the putative catalytic acid and base (Cys82 and Cys194 residues of the active site) are not located in the dimerization interface but on both sides of a cleft between the two domains. The two active sites in the same dimer of AspR are independent. At each of the active sites of PhAspR, two cysteine residues (Cys82 and Cys194) face each other. These two catalytic cysteine residues and the other surrounding residues, such as Asn83, Thr84, Gly193, and Thr195, are strictly conserved among all AspRs and other PLP-independent racemases, including GluRs. These residues are arranged in a pseudo mirror-symmetry. Particularly, Cys82 and Cys194 are located in the active site with the pseudo mirror-symmetrical orientations of their thiol groups.

The experimentally observed distance between the two sulfur atoms of the active cysteine residues (Cys82 and Cys194) is 9.6 Å [18], which is beyond the distance for cooperative action of them. So, Yoshida et al. examined the molecular mechanism of the PhAspR by mutational analyses and molecular dynamics (MD) simulations [19]. The kinetic parameters of some mutants PhAspR were determined, and the mutational analyses revealed that Arg48 and Lys164 were essential for catalysis in addition to the putative catalytic cysteine residues [19]. The MD simulations revealed that the distance between the two active sulfur atoms of cysteine residues oscillate to periodically become shorter than the predicted cooperative distance at high temperature. The Tyr160 and Lys164 residues were the entrance gate of the substrate L-Asp. In the docking molecular dynamics simulations, Yoshida et al. focused on the entrance process of the substrate (L-Asp) into the active site [19].

Recently, Ohtaki and his co-workers have determined the crystal structure of an inactive mutant PhAspR (Cys82 → Ala82, C82A) complexed with a citric acid (Cit) at a resolution of 2.0 Å [20]. In this structure, the L-Asp moiety of Cit is likely to take the substrate position of the PhAspR-substrate complex. Based on the "two-base" racemization mechanism [13–17], they also carried out the modeling studies of the PhAspR/Cit complex structure, in which Ala82 of the mutant C82A PhAspR was replaced by Cys82 that indicates the catalytic residue. And the results showed that the sulfur atom of Cys194 is in a favorable orientation to deprotonate the α-carbon of L-aspartate, and the Cys82 is located on the other side of the active site relative to Cys194 to reprotonate the α-carbon of L-Asp. But the binding mode of PhAspR-Cit indicated that Cys194 also plays the role of the base for the deprotonation from α-carbon of D-Asp and PhAspR adopts a "one-base" mechanism [20]. So,the binding of PhAspR with its real substrate and the detailed reaction mechanism of aspartate racemase need to be further studied.

The catalytic mechanisms of the glutamate racemase [21–23], proline racemase [24, 25], and diaminopimelate epimerase [26] have been studied theoretically, and results showed that there are numerous hydrogen bond donors surrounding the substrates and they all perform a highly asynchronous concerted process (Scheme 1). Up to now, to the best of our knowledge, there is no theoretical report on the catalytic mechanism of the aspartate racemase. In this work, we use the hybrid quantum mechanical/molecular mechanical (QM/MM) [27] method to investigate the substrate (L-Asp) binding in the active site of the PhAspR and the mechanism of the L-Asp → D-Asp reaction catalyzed by PhAspR.

2 Computational methods

The combined QM/MM approach is considered as a useful tool to understand characteristics of the enzyme systems [27–30], and it is a simulation approach which combines accuracy of QM and efficiency of MM. The QM part

consists of the atoms which are involved in bond forming and bond breaking processes, and it should be treated with an accurate QM method, such as ab initio methods. However, the numerical costs of the ab initio QM/MM approach restricted its application to systems involving small QM regions [31]. In this work, we employed the recently developed approximate density functional method, namely the self-consistent charge-density functional tight binding (SCC-DFTB) method [32] to treat the QM region. The SCC-DFTB method has been extensively tested [33–36] and applied successfully to several enzymatic systems [37–41]. The MM region was described by the CHARMM22 all-atom force field [42].

The starting structure for molecular simulation was based on the recently reported X-ray structure of an inactive mutant PhAspR (Cys82 → Ala82; C82A) complex with a citric acid (Cit) at a resolution of 2.0 Å (PDB code 2DX7) [20]. The substrate (L-Asp) was located at the position of the L-Asp moiety of Cit in the mutant PhAspR-Cit (C82A/Cit) complex and was modeled in the amino acid zwitterionic form with the protonated amino group and the deprotonated carboxylic group. The residue Ala82 was mutated back to Cys82, and the Cys194 residue was modeled in its deprotonated state and ready to accept a proton from the substrate. All the coordinates of hydrogen atoms of the substrate, the protein, and the crystallographic water molecules were determined using the HBUILD facility in the program CHARMM [43]. Care was taken in determining the protonation state of titrable residues at pH 7. As for the histidine residues, the δ-N is protonated for taking possible hydrogen bonding with the environment into consideration. Then a very short steepest descent minimization of the substrate and residue Cys82 were performed to remove bad contacts.

Once the model enzyme-substrate (ES) complex was constructed, it was solvated by a pre-equilibrated sphere of TIP3P waters [44] with a 25 Å radius centered at the Cα atom of substrate. Any water molecule found within a 2.8 Å radius of a heavy protein/substrate atom was deleted, followed by a 30 ps molecular dynamics applied with all protein, substrate and X-ray water molecules fixed at their original positions to relax the added solvent molecules. Then, to reduce the computational cost, stochastic boundary conditions [45] were applied to the solvated system. The system is governed by Newtonian dynamics on a classical potential in the reaction region ($r < 22$ Å). The atoms in the reservoir zone outside the 25 Å were deleted. In between, the atoms in the buffer zone ($22 < r < 25$ Å) were simulated with the Langevin dynamics with friction and random forces stemming from the bulk solvent that were not explicitly included in the simulation. A group-based switching scheme was used for nonbonded interactions [46]. The final model of the complex consisted of 7,356 atoms, including substrate L-Asp (15 atoms) and 1,035 water molecules.

In this study, the QM region comprises the substrate L-Asp and the side chains of the conjugated catalytic acid Cys82 and base Cys194. The interface between the QM and MM regions was approximated using link atoms [27], which were added to C_β atoms of the two protein residues. The CHARMM van der Waals parameters were used for all QM atoms. The SCC-DFTB/MM MD simulation was carried out for 1 ns after 100 ps of equilibration, in which the temperature was slowly brought to 375 K that was close to the physiological optimal temperature for PhaspR. The integration step for all of MD simulations was 1 fs, and the SHAKE algorithm [47] was applied to restrain the covalent bonds involving hydrogen atoms.

The racemization is composed of two elementary proton transfers. To obtain the two-dimensional potential energy surface (2D-PES), which was used to check whether the racemization process is concerted or step-wise, the two coordinates are used in this study to define this 2D-PES. *RC*1 is the antisymmetric combination of the distances involving the H_α deprotonation from the C_α by the Cys194 sulfur.

$$RC1 = r_{C_\alpha - H_\alpha} - r_{H_\alpha - S_{C194}} \quad (1)$$

*RC*2 is the antisymmetric combination of distances involving the protonation of the C_α by the Cys82 H_γ.

$$RC2 = r_{S_{C82} - H_{\gamma(C82)}} - r_{H_{\gamma(C82)} - C_\alpha} \quad (2)$$

Fig. 1 shows the results obtained using the combined SCC-DFTB/MM potential. We can see from Fig. 1 that the SCC-DFTB/MM optimizations yield a concerted mechanism where the two proton transfer via a single transition structure. This result is similar to the case in the

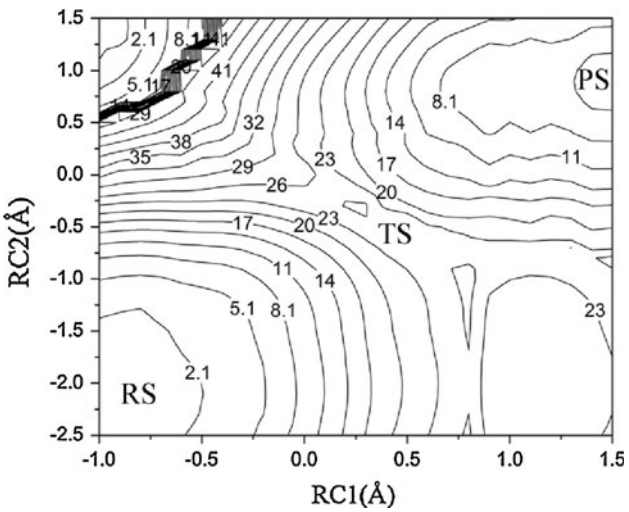


Fig. 1 Bidimensional SCC-DFTB/MM PES of the racemization reaction. The contour gap is 3.0 kcal/mol

previous calculations on the glutamate racemase [23], proline racemase [24, 25], and diaminopimelate epimerase [26].

Then, we used a one-dimensional (1D) reaction coordinate *RC*4 that takes into account both processes (deprotonation and protonation) at the same time to describe the 1D potential energy profile (1D-PEP).

$$RC4 = RC1 + RC2 \tag{3}$$

The 2D-PES and 1D-PEP have been obtained by a series of geometry optimizations in the presence of harmonic restraints applied on the reaction coordinates (*RC*1 and *RC*2 for the 2D-PES and *RC*4 for the 1D-PEP, respectively) used to drive the racemization process. Some snapshots were taken from the MD trajectories as the initial structures for reaction path calculations after minimization. The Adopted Basis Newton–Raphson (ABNR) method [43] was then used for the geometry optimization, in which the restraining force constant was set to equal to 2,000.0 kcal/(mol Å^2), the convergence condition for the rms gradient was 0.001 kcal/(mol ·Å), and the step size was 0.1 Å.

Both the 2D-PES presented in Fig. 1 and the 1D-PEP displayed in Fig. 2a suggest that the racemization catalyzed by the hAspR takes place in a concerted mechanism. From the 1D-PEP shown in Fig. 2a, one can see that the energy changes smoothly during the reaction. Consequently, the potential of mean force (PMF) can be determined by following the 1D reaction coordinate *RC*4. In this work, the PMF for the racemization was computed using umbrella sampling [48] with the structures along the putative reaction coordinate *RC*4 as the initial structures. The harmonic force constant ranged from 100 to 200 kcal/(mol Å^2). A total 27 windows were employed to cover the whole range of the reaction coordinates. In each window, a 100 ps constrained MD simulation was carried out, with 50 ps heating and equilibration to bring the system to 375 K and 50 ps data collection. The final PMF was obtained using the weighted histogram analysis method (WHAM) [49].

To estimate the contribution of individual residues to transition state stabilization, an electrostatic interaction analysis was performed according to the procedure used by Rubinstein et al. [25]. For residue *I*, initially, the QM and QM/MM energy of the fully charged system was determined, and then, the QM and QM/MM energy was determined after the MM charge of residue *I* zeroed, the energy difference corresponds to the interaction energy between this residue and the QM subsystem:

$$\Delta E(I) = (E_{\mathrm{QM}} + E_{\mathrm{QM/MM}}) - (E_{\mathrm{QM}}(I) + E_{\mathrm{QM/MM}}(I)) \tag{4}$$

Where $E_{\mathrm{QM}}(I)$ and $E_{\mathrm{QM/MM}}(I)$ stand for the QM and QM/MM energies calculated with the MM charges of residue *I* annihilated.

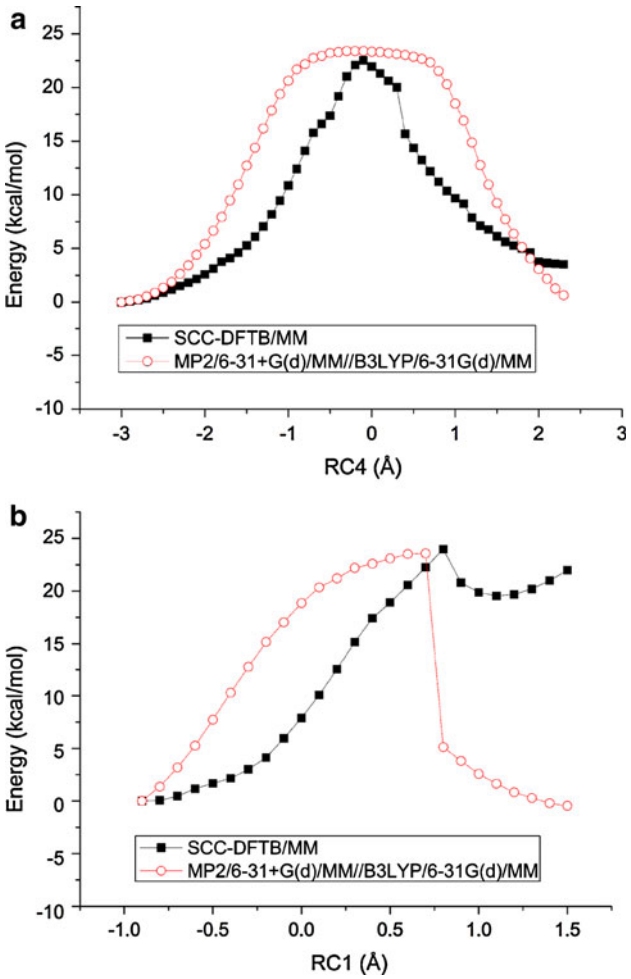


Fig. 2 One-dimensional potential energy profile along the reaction coordinate RC4 (**a**) and along the reaction coordinate RC1 (**b**) using the SCC-DFTB/MM method. For comparison, energies obtained from MP2/6-31+G(d)/MM//B3LYP/6-31G(d)/MM single point calculations are also shown

From the umbrella sampling MD simulations, we selected 400–600 snapshots with the reaction coordinate (±0.1 Å) value corresponding to stationary points, namely reactant state (RS) and transition state (TS). At each structure, the interaction energy has been computed and the results have been averaged over the whole analysis of 400–600 snapshots. The order in which we selected the residue to zero out their charges is based on the distance between the Cα of the residue and the Cα of the substrate up to 20 Å.

The binding mode of PhAspR-Cit indicated that Cys194 also plays the role of the base for the deprotonation from α-carbon of D-Asp and PhAspR seemed adopt a "one-base" mechanism [20]. To further characterize the binding modes of L-Asp and D-Asp, we first obtained the apo-PhAspR by deleting the substrate and water molecules of the pre-constructed ES complex, and then docked the L-Asp and D-Asp into the apo-PhAspR using

Autodock 4.2 [50], respectively. Initial modes for L-Asp and D-Asp were generated from PubChem Compound Database through the unique chemical structure identifier CID: 5460294 [51] and 5460295 [52], respectively. The PRODRG server [53] was used to convert structure of L-Asp and D-Asp to PDB format. Polar hydrogens were added and Kollman charges [54] were assigned to all atoms. Of the two identical active sites from the PhAspR homo-dimer, the active site of chain A was chosen for the docking site. The grid map with 61 × 61 × 61 points and a grid-point spacing of 0.375 Å centered at the midpoint between the sulfur atoms of Cys82 and Cys194 by use of Autogrid4.2 [50]. Then Autodock4.2 [50] was used for the docking simulation. We selected the Lamarckian Genetic Algorithm (LGA), which combines a genetic algorithm and an adaptative local-search algorithm. For L-Asp and D-Asp, the docking parameters were as follows: trial of 100 dockings, population size of 150, translation step size of 0.2, torsion step of 5°, a maximum of 2.5×10^6 energy evaluations, a maximum of 27,000 generations, an elitism value of 1, a mutation rate of 0.02, a crossover rate of 0.8, and local search rate of 0.06.

3 Results and discussions

3.1 Validation of SCC-DFTB

The SCC-DFTB method could give good agreement with B3LYP in describing the geometry and relative energetic ordering of various conformations of model peptides [55], and it has been used to investigate proton transfer in some enzymes, e.g., triose phosphate isomerase [33], carbonic anhidrase II [56], yeast cytosine deaminase [57], and cytidine deaminase [39]. Comparison with higher-level methods indicated that SCC-DFTB gives a reasonable description of structures and energetics of proton transfer in these systems. In the triose phosphate isomerase, the SCC-DFTB energetics agreed with B3LYP/6- 31+G(d,p) within 2–4 kcal/mol (rms values) [33].

For the particular system in this work, we performed the minimization at the B3LYP/6-31+G(d)/MM level. The results are compared with those obtained from the SCC-DFTB/MM minimization in Fig. 3 with the convergence criterion that the rms gradient is 0.001 kcal/(mol Å). We can see that the geometry differences between the SCC-DFTB/MM and B3LYP/6-31+G(d)/MM results are very small. This indicates that the SCC-DFTB method is reliable for describing our system. The B3LYP/MM calculation was performed using CHARMM interfaced with GAMESS-US package [58], which was also used to calculate the MP2/6-31+G(d)/MM single point energy and Mulliken charge (see blow).

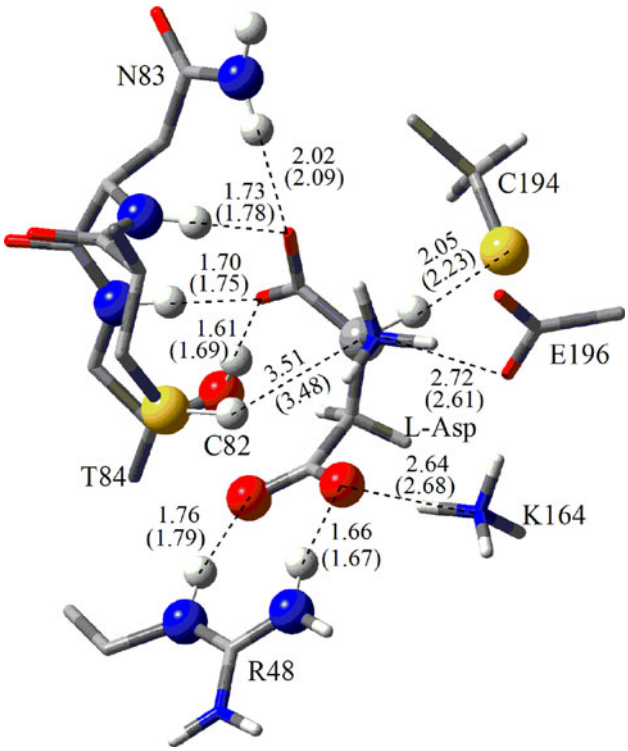


Fig. 3 Comparison of the active site between SCC-DFTB/MM and B3LYP/6-31+G(d)/MM (*parentheses*) minimization. Unit for distances is given in angstrom

3.2 QM/MM molecular dynamics simulation of the enzyme-substrate complex

During the simulation, the substrate L-Asp forms an extensive network of interactions with the active-site residues through its α-carboxylate, ammonium group, and side chain carboxylate. The network of hydrogen bond is depicted in Fig. 4. One can see that the conserved Asn83 residue interacts through its side chain $H_{\delta 22}$ atom with the atom O_2 of substrate with a distance of 2.22 ± 0.29 Å. Furthermore, the backbone H_N of Asn83 and Thr195 can interact with O_2 of the α-carboxylate of substrate, as implied by the $O_2{-}H_{N(N83)}$ distance of 1.80 ± 0.14 Å and the $O_2{-}H_{N(T195)}$ distance of 2.13 ± 0.25 Å, respectively. The next conserved residue Thr84 interacts with O_1 atom of the substrate through its side chain hydroxyl hydrogen Hγ and its back bone H_N, with the $O_1{-}H_{\gamma(T84)}$ distance of 1.66 ± 0.11 Å and the $O_1{-}H_{N(T84)}$ distance of 1.78 ± 0.12 Å, respectively. In the simulation, we found a crystallographic water molecular (cw507) in the active site, and it interacts with the O_1 atom of the substrate, as evidenced by the $O_1{-}O_{(cw507)}$ distance of 2.83 ± 0.26 Å. There are some residues interacting with the substrate ammonium group. The distance of $N{-}O_{\gamma(T195)}$ was found to be 2.86 ± 0.13 Å. There is another hydrogen bond between the $O_{\varepsilon 1}$ atom of Glu196 and the NH_3^+ group with the distance of

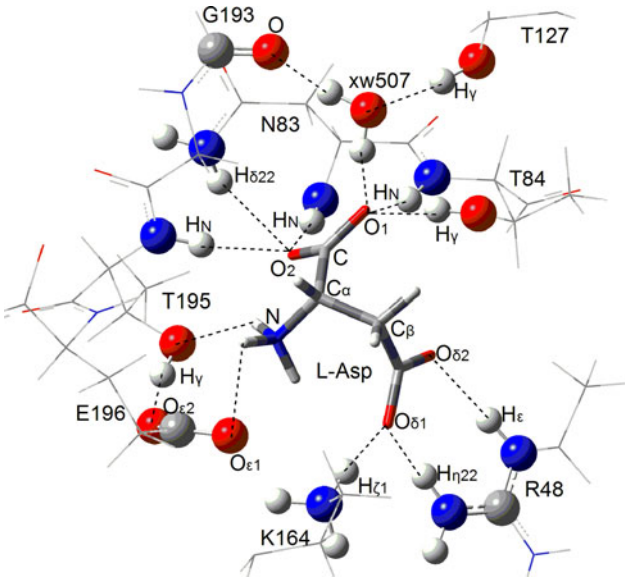


Fig. 4 Hydrogen bond interactions of the substrate in the active site of the PhAspR

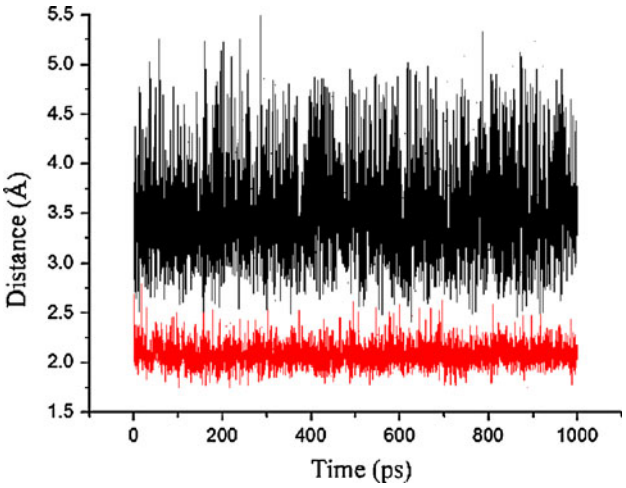


Fig. 5 Time evolution of the distance between atom $C_{\alpha B}$ and $H_{\gamma\ (C82)}$ (*black*, *top*) and the distance between atom H_α and $S_{\gamma(C194)}$ (*red*, *bottom*) in the ES complex

$N–O_{\varepsilon 1(E196)}$ being 2.79 ± 0.16 Å. Furthermore, the residue Lys164 interacts with the substrate side chain carboxylate, as evidenced by the $O_{\delta 1}–N_{\zeta(K164)}$ distance of 2.67 ± 0.10 Å. The indispensable residue Arg48 interacts with the substrate side chain carboxylate by strong hydrogen bonds, with the $O_{\delta 2}–H_{\varepsilon(R48)}$ distance of 1.76 ± 0.12 Å and the $O_{\delta 1}–H_{\eta 22(R48)}$ distance of 1.69 ± 0.13 Å, respectively. These residues in the active site orient and bind to the substrate well, similar to the binding of the glutamate racemase to D-glutamate [23].

The racemization involves the deprotonation of the substrate H_α by the $S_{\gamma(C194)}$ and the reprotonation of the C_α by the $H_{\gamma(C82)}$. In the ES complex, the evolutions of the corresponding distances along the SCC-DFTB/MM MD are showed in Fig. 5. Our QM/MM MD simulation of the ES complex showed that the average distance between the two γ-sulfur atoms is 7 Å, which is shorter than that in the structure of PhAspR (9.6 Å). This distance is almost the same as in the modeling study by Ohtaki et al. (7.4 Å) [20]. The sulfur atom of Cys194 interacts strongly with the α-hydrogen of substrate, as evidenced by the $H_\alpha–S_{\gamma(C194)}$ distance of 2.08 ± 0.12 Å, the $S_{\gamma(C194)}$ points to the H_α well. The average distance between the H_γ of Cys82 and C_α of substrate was found to be 3.52 Å, although this distance fluctuates mostly from 3.0 Å to 4.0 Å, it is expected to be easily reduced when the deprotonation of the substrate begins and the C_α becomes negatively charged.

3.3 Racemization reaction

Since this racemization reaction is composed of two elementary proton transfers, the first point to be considered is that the deprotonation and reprotonation steps are concerted or step-wise. On the basis of our SCC-DFTB/MM calculations, the 2D SCC-DFTB/MM PES built as a function of reaction coordinates *RC*1 and *RC*2 is shown in Fig. 1. The reaction coordinates evolve from negative to positive values as the deprotonation and reprotonation processes, respectively. From Fig. 1, we can locate two lower energy regions at the bottom left corner and at the top right corner, respectively. These two structures were further minimized with no restraints to obtain structures corresponding to the reactant and the product of the enantioneric inversion reaction. We can locate the reactant state (RS) corresponding to negative values of *RC*1 (−0.8 Å) and *RC*2 (−2.2 Å) and the product state (PS) with the values of *RC*1 (1.4 Å) and *RC*4 (0.9 Å), respectively. The 2D-PES exploration also leads to a corresponding transition state (TS) in a concerted way, and at the values of *RC*1 (0.5 Å) and *RC*2 (−0.6 Å) (see Fig. 1).

Then, we obtained a 1D-PEP using an adiabatic mapping approach as a function of *RC*4. Interestingly, as shown in Fig. 2a, the 1D-PEP is very smooth and there is only one maximum, which is further investigated by a MP2/6-31+G(d)/MM//B3LYP/6-31G(d)/MM single point calculation. The results are consistent with the no formation of an intermediate. The difference of energy barrier is very small, indicating that the SCC-DFTB method is reliable for describing the system. The computed 1D path reveals a highly asynchronous concerted process, similar to the glutamate racemase [23], proline racemase [24, 25], and diaminopimelate epimerase [26]. The structures of the QM/MM stationary states are displayed in Fig. 6. In the TS, both Cys194 and Cys82 are nearly protonated, as evidenced by the $H_\alpha–S_{\gamma(C194)}$ distance of 1.40 Å and the $H_{\gamma(C82)}–S_{\gamma(C82)}$ distance of 1.37 Å, somewhat longer than

the covalent S–H distance of a cysteine (1.33 Å). These values clearly indicate that at the TS the deprotonation process has nearly finished when the reprotonation process has just began. Furthermore, in the TS, the unprotonated substrate is almost planar with the dihedral angle N–C_α–C_β–C of −160.70 degree. Such TS structure allows a delocalization of the partial negative charge over the extended π orbital system. We find that the negative charge is mainly localized on the two oxygen atoms of the α-carboxyl group when the proton on the α-carbon is transferred to the basic, catalytic residue Cys194. It is observed that the Mulliken charge on the O_1 and O_2 atoms is −0.81 au and −0.74 au (these calculations were performed using B3LYP/6-31+G(d)/MM method, see above) in RS, respectively. In the TS structure, the Mulliken charge on these oxygen atoms is −0.94 au and −0.93 au, respectively. Also, the carboxylic and $–NH_3$ groups become more negative and less positive, respectively, on passing from RS to TS.

The 2D-PES from Fig. 1 also shows there might be a stepwise path passing through the bottom right corner. To characterize this stepwise mechanism, we obtained the potential energy profile as a function of *RC*1 and the result was shown in Fig. 2b. There is one maximum in this potential energy profile at the value of *RC*1 0.8 Å which indicates that the H_α atom has transferred to the $S_{\gamma(C194)}$ atom. The potential energy decreases sharply from the maximum point. We checked these structures near to the maximum point and found that the new forming hydrogen bond between the transferred H_α atom and the O_1 atom results in the decreasing of potential energy. We also calculated the MP2/6-31+G(d)//MM single point energy at the optimized structures by B3LYP/6-31+G(d)//MM, and the results showed that the sharp decline of energy in the potential energy profile is caused by the forming of the product not the carbanion intermediate. The above results indicate that the PhAspR is not likely to undergo a stepwise mechanism.

3.4 Potential of mean force

The calculated PMF is depicted in Fig. 7a, and it is consistent with our previous reaction path calculations, there only one transition state can be identified for the racemization reaction. To further check whether there is a metastable state between the RS and PS, we have carried out the unrestrained SCC-DFTB/MM molecular dynamics simulations with 50 random snapshots around the transition state region TS. Among all simulations, there is no other stable state observed, except for the reactant state RS and

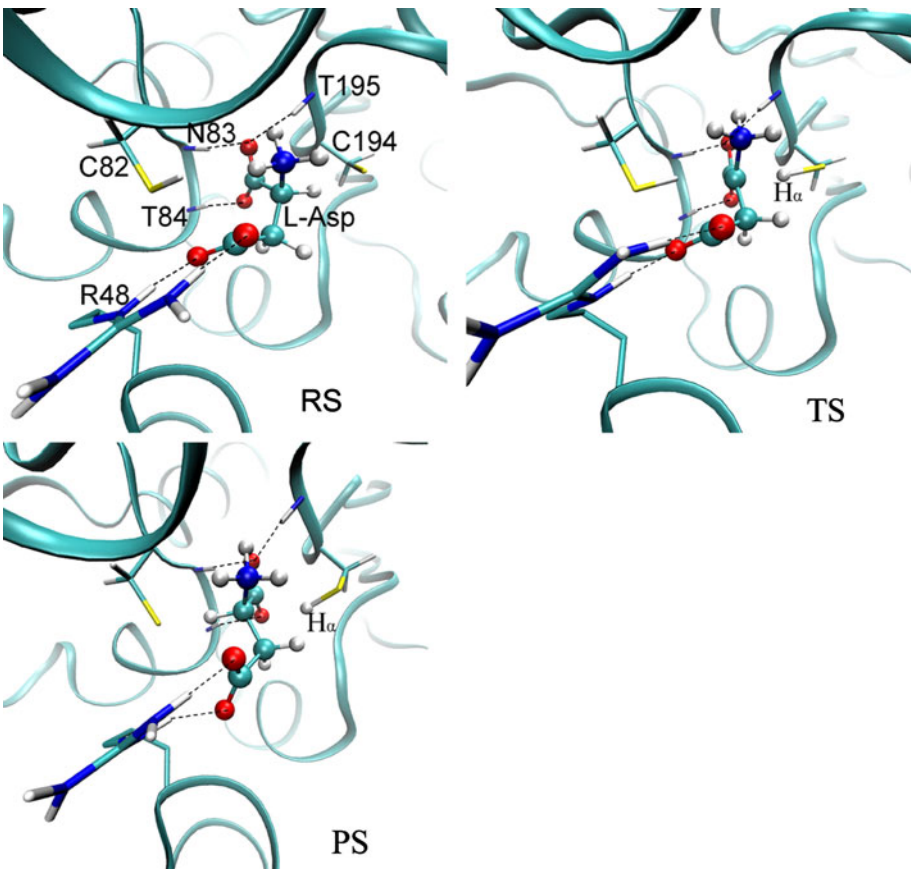


Fig. 6 Snapshots of the QM/MM stationary states along the reaction path for the PhAspR-catalyzed racemization. *Black dashed lines* represent the hydrogen bonds, and for the sake of clarity, some hydrogen bonds are not displayed

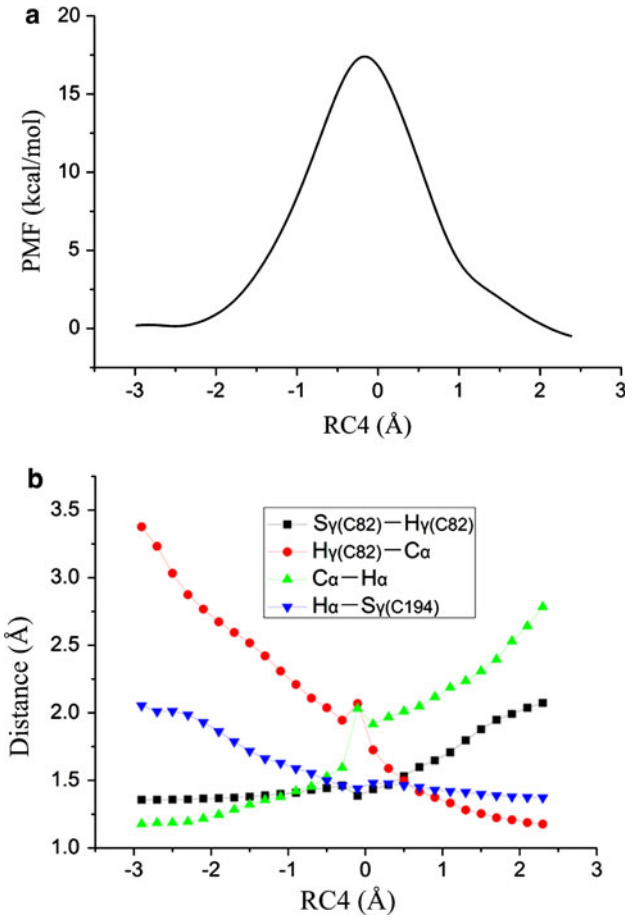


Fig. 7 Calculated potential of mean force along the reaction coordinate RC4 (**a**) and average distance of R($S_{\gamma(C82)}$-$H_{\gamma(C82)}$) (*black*), R($H_{\gamma(C82)}$-C_α) (*red*), R(C_α-H_α) (*green*), R(H_α-$S_{\gamma(C194)}$) (*blue*) as a function of RC4 obtained from the free energy simulations (**b**)

product state PS. The free energy barrier is 17.5 kcal/mol for this reaction. The experimental k_{cat} value for the L → D racemization is 1.9×10^4 min^{-1} [19], corresponding to the free energy barrier of 16.3 kcal/mol obtained by the transition-stated theory. It is a really low energy barrier taking into account that a cysteine is deprotonating a very low acidic proton such as the $H_{\alpha B}$ of L-Asp. To obtain more quantitative values for reaction barriers, it might be desirable to carry out further free energy computations using higher-level method like ab initio/MM free energy computations.

Fig. 7b shows the average distances of corresponding breaking and forming bonds along *RC*4 from the umbrella sampling. In the beginning, *RC*4 from −3 to −2 Å, both the deprotonation and reprotonation do not begin, as the distances C_α–H_α and $H_{\gamma(C82)}$–$S_{\gamma(C82)}$ do not change. The increase in free energy correlates with the decrease in distance between the $H_{\gamma(C82)}$ and the C_α. After this region, the deprotonation process begins, as evidenced by the increased distance C_α–H_α and the decreased distance H_α–$S_{\gamma(C194)}$. Before the saddle point region, the distance between $H_{\gamma(C82)}$ and $S_{\gamma(C82)}$ does not change, indicating that the reprotonation has not begun even the deprotonation process has nearly finished. In the saddle point, both Cys194 and Cys82 are nearly protonated. After the saddle point, the reprotonation begins by the Cys82. This racemization is an asynchronous concerted process and the deprotonation precedes the reprotonation.

3.5 Electrostatic interaction analysis

An electrostatic interaction analysis was performed as described in the computational details section to analyze the contribution of individual residue to transition state stabilization. The analysis of TS versus RS is shown in Fig. 8.

In the TS, the Asp α-carboxylate is bound in a *carboxylate hole* stabilized by a hydrogen bonds network, similar to that suggested for GluR [23] and ProR [25]. The side chains of Asn83 and Thr84 donate hydrogen bonds to the α-carboxylate, while the Asn83, Thr84, and Thr195 backbone amides donate additional hydrogen bonds. These hydrogen bonds become shorter when the system moves from RS to TS (Table 1) and these residues stabilize the TS by 7.6, 6.0, and 5.1 kcal/mol relative to RS, respectively. Replacing Thr84 by Ala causes significant loss of activity [19]. The electrostatic interaction between the side chain carboxylate of Asp and Arg48 stabilizes the TS relative to RS by 17.6 kcal/mol, and another important residue Lys164 contributes 11.6 kcal/mol to TS stabilization. It is not surprising that these two positive residues involved in binding Asp have a strong stabilizing effect; they all help to

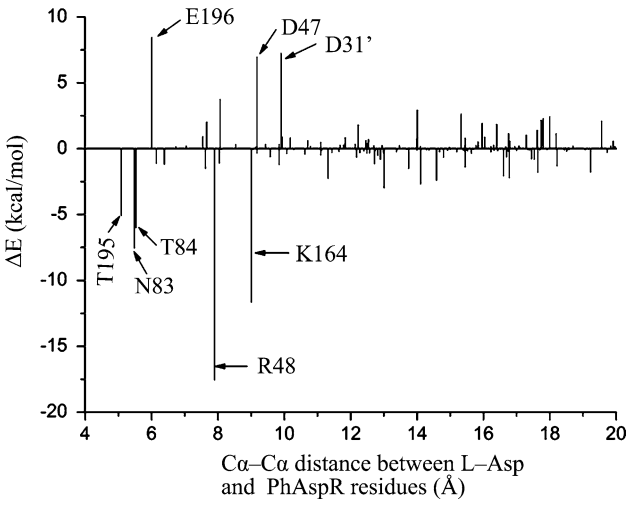


Fig. 8 Individual residue contribution to stabilization or destabilization of the TS relative to the RS as a function of the C_α–C_α distance between the L–Asp and PhAspR residues

Table 1 Selected geometric parameters for stationary states along the reaction path of the PhAspR-catalyzed racemization of L-Asp calculated using the SCC-DFTB/MM method and for the docked models

Distances (Å), angle (deg)	RS	TS	PS	MD (ES)[a]	L-Asp-docked	D-Asp-docked
O_2–$H_{\delta 22(N83)}$	2.00	1.80	1.97	2.22 ± 0.29	2.03	1.87
O_2–$H_{N(N83)}$	1.74	1.71	1.82	1.80 ± 0.14	2.04	1.81
O_2–$H_{N(T195)}$	2.00	1.76	1.80	2.13 ± 0.25	1.66	1.68
O_1–$H_{\gamma(T84)}$	1.60	1.60	1.65	1.66 ± 0.11	1.74	1.89
O_1–$H_{N(T84)}$	1.70	1.69	2.04	1.78 ± 0.12	1.68	2.04
O_1–$O_{(cw507)}$	2.66	2.58	2.61	2.83 ± 0.26	–	–
N–$O_{\gamma(T195)}$	2.81	2.80	2.78	2.86 ± 0.13	2.81	2.79
N–$O_{\varepsilon 1(E196)}$	2.69	2.77	2.79	2.79 ± 0.16	2.60	2.65
$O_{\delta 1}$–$H_{\eta 22(R48)}$	1.58	1.54	1.57	1.69 ± 0.13	1.90	1.65
$O_{\delta 1}$–$N_{\zeta(K164)}$	2.62	2.58	2.57	2.67 ± 0.10	2.71	2.68
$O_{\delta 2}$–$H_{\varepsilon(R48)}$	1.71	1.67	1.76	1.76 ± 0.12	2.18	1.95
$S_{\gamma(C82)}$–$H_{\gamma(C82)}$	1.35	1.37	2.01		–	–
$H_{\gamma(C82)}$–C_α	3.48	2.03	1.18	3.52 ± 0.51	4.59[b]	3.42[b]
C_α–H_α	1.17	1.96	2.84	–	3.25[c]	4.36[c]
H_α–$S_{\gamma(C194)}$	2.05	1.40	1.37	2.08 ± 0.12	–	–
N–C_α–C_β–C	124.63	−160.70	−125.97	122.28 ± 5.74	119.82	−121.04

[a] The average geometries from our QM/MM MD simulation for ES

[b] Distances between atom $S_{\gamma(C82)}$ and atom C_α

[c] Distances between atom C_α and atom $S_{\gamma(C194)}$

neutralize the negative charge of the carboxylate of Asp, which is increasingly important during the abstraction of a proton from Asp. The mutation of either Arg48 or Lys164 forms hydrogen bond to the side chain carboxylate of substrate, resulting in a dramatic decrease in activity [19]. In contrast to the above residues, Glu196 interacts with the Asp NHB^+_{3B} group and destabilizes the TS for repulsive electrostatic interaction with the Asp carbanion. Asp31' (adjacent monomer) destabilizes the TS because of the electrostatic interaction with Arg48 and Lys164. Asp47 can interact with Asp31' (adjacent monomer) through hydrogen bond intermediated by a water molecule, and it destabilizes the TS. It is interesting to note that most of the residues discussed in the above are conserved in the GluR [22, 23] and ProR [25]. These residues are very important for the transition state stabilization.

3.6 Docking study the binding of L-Asp and D-Asp

Two models have been built to represent the PhAspR-Cit complex. The so-called L-Asp-docked model and D-Asp-docked model have been built from the coordinates obtained after docking the L-Asp and D-Asp residue into the active site of PhAspR by using a docking computational method, respectively. The configuration with the lowest binding free energy was selected and depicted in Fig. 9. Figure 9 and the distances given in the last two columns of Table 1 indicate that in both the L-Asp-docked model and D-Asp-docked model, the hydrogen bonds are formed between residue Asn83, Thr84, Thr195, and the α-carboxylate group of the ligand. The ligand is also bound by its side chain carboxylate to residue Arg48 and Lys164, and there is an electrostatic interaction between its ammonium group and Glu196. The binding model of L-Asp-docked is very similar to that of our constructed ES complex from PhAspR-Cit and the sulfur atom of Cys194 is in a favorable orientation to deprotonate the α-carbon of L-Asp, as shown in Fig. 9. The binding model of D-Asp-docked model indicates that the orientation of the α-carbon is directed to the Cys82 residue. Therefore, PhAspR employs a "two-base" mechanism not a "one-base" mechanism.

4 Conclusions

In this work, we reported QM/MM studies on the substrate binding and the mechanism of the racemization of L-aspartate to D-aspartate by the enzyme aspartate racemase from *Pyrococcus horikoshii* OT3. The QM/MM Molecular Dynamics simulation to study the substrate (L-Asp) binding in the active site shows that the substrate forms an extensive network of interactions with the active-site residues through its side chain carboxylate, ammonium group, and α-carboxylate. The interconversion reaction involves two proton transfers, one is the deprotonation of the α-hydrogen by the $S_{\gamma(C194)}$ and the other is the reprotonation of the

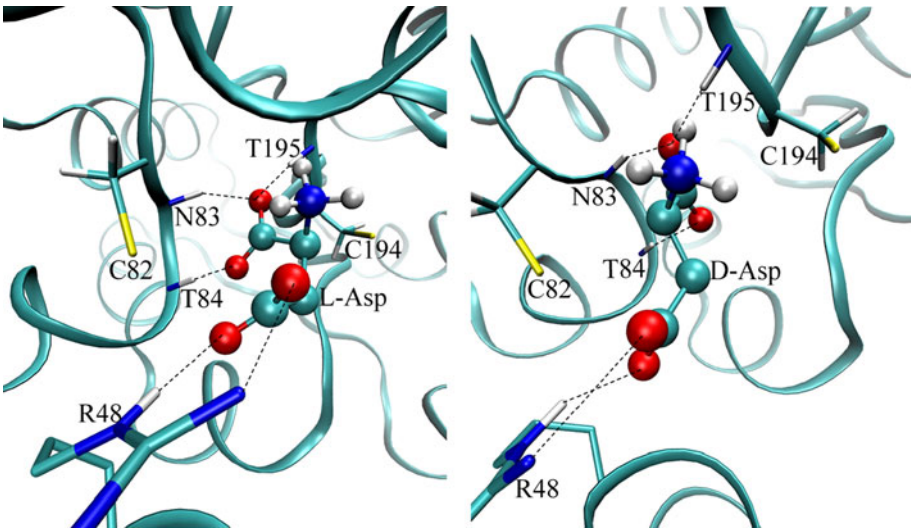


Fig. 9 The configurations of L-Asp-docked model and D-Asp-docked model. *Black dashed lines* represent the hydrogen bonds, and for the sake of clarity, some hydrogen bonds are not displayed

α-carbon by the $H_{\gamma(C82)}$. The QM/MM potential energy profile and potential of mean force suggest that the reaction proceeds via a highly asynchronic concerted mechanism with a single transition state. In addition, we perform an electrostatic interaction analysis to estimate the key role played by individual residues in stabilizing the transition state and conclude that the enzyme stabilizes the transition state through local hydrogen bond and short range electrostatic contribution. The docking study indicates that Cys194 is responsible for the deprotonation of the α-carbon of L-Asp, and Cys82 is responsible for the deprotonation of the α-carbon of D-Asp, the enzyme aspartate racemase from *Pyrococcus horikoshii* OT3 employs a "two-base" mechanism not a "one-base" mechanism. The results together with the mechanism of the glutamate racemase [21–23], proline racemase [24, 25], and diaminopimelate epimerase [26] indicate a common mechanism for the PLP-independent enzymes.

Acknowledgments This project has been supported by the National Natural Science Foundation of China (Grant Nos. 20773089 and 21075083) and National Basic Research Program of China (973 Program) (2011CB201202). We appreciate Prof. Dingguo Xu for his many stimulating discussions. The CHARMM calculations have been carried out in Wuhan Institute of Physics and Mathematics, the Chinese Academy of Science.

References

1. Walsh CT (1989) J Biol Chem 264:2393–2396
2. Yohda M, Endo I, Abe Y, Ohta T, Iida T, Maruyama T, Kagawa Y (1996) J Biol Chem 271:22017–22021
3. Matsumoto M, Homma H, Long Z, Imai K, Iida T, Maruyama T, Aikawa Y, Endo I, Yohda M (1999) J Bacteriol 181:6560–6563
4. Dunlop DS, Neidle A, AcHale D, Dunlop DM, Lajtha A (1986) Biochem Biophys Res Commun 141:27–32
5. Wolosker H, D'Aniello A, Synder SH (2000) Neuroscience 100:183–189
6. Johnston MM, Diven WF (1969) J Biol Chem 244:5414–5420
7. Gerlt JA, Kenyon GL, Kozarich JW, Neidhart DJ, Petsko GA, Powers VM (1992) Curr Opin Struct Biol 2:736–742
8. Hayashi H, Wada H, Yoshimura T, Esaki N, Soda K (1990) Annu Rev Biochem 59:87–110
9. Nakajima N, Tanizawa K, Tanaka H, Soda K (1986) Agric Biol Chem 50:2823–2830
10. Albery WJ, Knowles JR (1986) Biochemistry 25:2572–2577
11. Higgins W, Tardif C, Richaud C, Krivanek MA, Cardin A (1989) Eur J Biochem 186:137–143
12. Finlay TH, Adams EJ (1970) J Biol Chem 245:5248–5260
13. Gallo KA, Tanner ME, Knowles JR (1993) Biochemistry 32:3991–3997
14. Tanner ME, Gallo KA, Knowles JR (1993) Biochemistry 32:3998–4006
15. Koo CW, Blanchard JS (1999) Biochemistry 38:4416–4422
16. Fisher LM, Belasco JG, Bruice TW, Albery WJ, Knowles JR (1986) Biochemistry 25:2543–2551
17. Yamauchi T, Choi SY, Okada H, Yohda M, Kumagai H, Esaki N, Soda K (1992) J Biol Chem 267:18361–18364
18. Liu L, Iwata K, Kita A, Kawarabayasi Y, Yohda M, Miki K (2002) J Mol Biol 319:479–489
19. Yoshida T, Seko T, Okada O, Iwata K, Liu L, Miki K, Yohda M (2006) Proteins 64:502–512
20. Ohtaki A, Nakano Y, Iizuka R, Arakawa T, Yamada K, Odaka M, Yohda M (2008) Proteins 70:1167–1174
21. Puig E, Garcia-Viloca M, Gonzalez-Lafont A, Lluch JM, Field MJ (2007) J Phys Chem B 111:2385–2397
22. Puig E, Mixcoha E, Garcia-Viloca M, Gonzalez-Lafont A, Lluch JM (2009) J Am Chem Soc 131:3509–3521
23. Spies MA, Reese JG, Dodd D, Pankow KL, Blanke SR, Baudry J (2009) J Am Chem Soc 131:5274–5284
24. Stenta M, Calvaresi M, Alto P, Spinelli D, Garavelli M, Bottoni A (2008) J Phys Chem B 112:1057–1059
25. Rubinstein A, Major DT (2009) J Am Chem Soc 131:8513–8521
26. Stenta M, Calvaresi M, Altoe P, Spinelli D, Garavelli M, Galeazzi R, Bottoni A (2009) J Chem Theory Comput 5:1915–1930
27. Field MJ, Bash PA, Karplus M (1990) J Comput Chem 11:700–733
28. Gao J (1996) Acc Chem Res 29:298–305
29. Warshel A (2003) Annu Rev Biophys Biomol Struct 32:425–443
30. Riccardi D, Schaefer P, Yang Y, Yu H, Ghosh N, Prat-Resina X, Konig P, Li G, Xu D, Guo H, Elstner M, Cui Q (2006) J Phys Chem B 110:6458–6469

31. Friesner RA, Guallar V (2005) Annu Rev Phys Chem 56:389–427
32. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seigert G (1998) Phys Rev B58:7260–7268
33. Cui Q, Elstner M, Kaxiras E, Frauenheim T, Karplus M (2001) J Phys Chem B 105:569–585
34. Elstner M, Hobza P, Frauenheim T, Suhai S, Kaxiras E (2001) J Chem Phys 114:5149–5155
35. Pu J, Gao J, Truhlar DG (2004) J Phys Chem A 108:5454–5463
36. Witek HA, Morokuma K (2004) J Comput Chem 25:1858–1864
37. Cui Q, Elstner M, Karplus M (2002) J Phys Chem B 106:2721–2740
38. Zhang X, Harrison DH, Cui Q (2002) J Am Chem Soc 124:14871–14878
39. Guo H, Rao N, Xu Q, Guo H (2005) J Am Chem Soc 127:3191–3197
40. Liu J, Wang X, Xu D (2010) J Phys Chem B 114:1462–1470
41. Xu Q, Li L, Guo H (2010) J Phys Chem B 114:10594–10600
42. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M (1998) J Phys Chem B 102:3586–3616
43. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comput Chem 4:187–217
44. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926–935
45. Brooks CL III, Brunger A, Karplus M (1985) Biopolymers 24:843–865
46. Steinbach PJ, Brooks BR (1994) J Comput Chem 15:667–683
47. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327–341
48. Torrie GM, Valleau JP (1977) J Comput Phys 23:187–199
49. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM (1992) J Comput Chem 13:1011–1021
50. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639–1662
51. National Center for Biotechnology Information. PubChem Compound Database (2011) CID = 5460294, http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=5460294&loc=ec_rcs accessed Mar 11 2011
52. National Center for Biotechnology Information. PubChem Compound Database (2011); CID = 5460295, http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=5460295&loc=ec_rcs accessed Mar 11 2011
53. Schuettelkopf AW, van Aalten DMF (2004) Acta Cryst D 60:1355–1363
54. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S, Weiner P (1984) J Am Chem Soc 106:765–784
55. Elstner M, Frauenheim T, Kaxiras E, Seifert G, Suhai S (2000) Phys Status Solidi B 217:357–376
56. Riccardi D, Konig P, Guo H, Cui Q (2008) Biochemistry 47:2369–2378
57. Xu Q, Guo H, Gorin A, Guo H (2007) J Phys Chem B 111:6501–6506
58. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su S, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347–1363