PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

A QM/MM study of the reaction mechanism of (R)-hydroxynitrile lyases from *Arabidopsis thaliana* (AtHNL)

Wenyou Zhu,[1] Yongjun Liu,[1,2]* and Rui Zhang[3]

[1] School of Chemistry and Chemical Engineering, Shandong University, Jinan, Shandong 250100, China

[2] Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, Qinghai 810001, China

[3] School of Agriculture, Ludong University, Yantai, Shandong 264025, China

ABSTRACT

Hydroxynitrile lyases (HNLs) catalyze the conversion of chiral cyanohydrins to hydrocyanic acid (HCN) and aldehyde or ketone. Hydroxynitrile lyase from *Arabidopsis thaliana* (AtHNL) is the first R-selective HNL enzyme containing an α/β-hydrolases fold. In this article, the catalytic mechanism of AtHNL was theoretically studied by using QM/MM approach based on the recently obtained crystal structure in 2012. Two computational models were constructed, and two possible reaction pathways were considered. In Path A, the calculation results indicate that the proton transfer from the hydroxyl group of cyanohydrin occurs firstly, and then the cleavage of C1-C2 bond and the rotation of the generated cyanide ion (CN^-) follow, afterwards, CN^- abstracts a proton from His236 via Ser81. The C1-C2 bond cleavage and the protonation of CN^- correspond to comparable free energy barriers (12.1 vs. 12.2 kcal mol^{-1}), suggesting that both of the two processes contribute a lot to rate-limiting. In Path B, the deprotonation of the hydroxyl group of cyanohydrin and the cleavage of C1-C2 bond take place in a concerted manner, which corresponds to the highest free energy barrier of 13.2 kcal mol^{-1}. The free energy barriers of Path A and B are very similar and basically agree well with the experimental value of HbHNL, a similar enzyme of AtHNL. Therefore, both of the two pathways are possible. In the reaction, the catalytic triad (His236, Ser81, and Asp208) acts as the general acid/base, and the generated CN^- is stabilized by the hydroxyl group of Ser81 and the main-chain NH-groups of Ala13 and Phe82.

Proteins 2015; 83:66–77.

Key words: R-selective HNL enzyme; asymmetric synthesis; catalytic mechanism; theoretical study; chiral cyanohydrins; hydrocyanic acid.

INTRODUCTION

Hydroxynitrile lyases (HNLs) are important enzymes in a variety of cyanogenic plants,[1] which catalyze the cleavage of chiral cyanohydrin to generate hydrocyanic acid (HCN) and aldehyde or ketone (Scheme 1).[3–5] As a high-toxic compound, the product HCN is a good weapon for defense against microbial and herbivores attack in cyanogenic plants.[5–7] Besides, HCN is also used as a nitrogen source in the biosynthesis of asparagine.[8,9] However, the main concern of biologists and chemists is focused on the reverse of the HNL-catalyzed reaction, because HNL can catalyze the asymmetric synthesis of cyanohydrins (α-hydroxynitriles) using the non-natural chiral substrate HCN and ketone as materials,[10–14] and the enantiomerically enriched cyanohydrin is a valuable intermediate in pharmaceuticals, agrochemicals, and chemicals.[10,15–17] Because chiral cyanohydrins are difficult to synthesize through the classical chemical methods,[18,19] their asymmetric biosynthesis using HNLs has attracted great attentions.[16,17,20,21] Understanding the catalytic mechanism of HNLs at atomistic level may provide useful help for the improvement of biocatalysts.

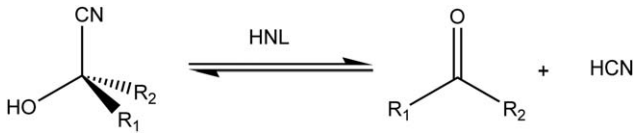
Hydroxynitrile lyases belong to the aldehyde lyases.[4] And HNLs family can be categorized into five diverse

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Natural Science Foundation of China; Grant number: 21173129, 31200048, 21373125.

*Correspondence to: Y. Liu, School of Chemistry and Chemical Engineering, Shandong University, Jinan, Shandong 250100, China.
E-mail: yongjunliu_1@sdu.edu.cn

Received 26 March 2014; Revised 7 July 2014; Accepted 15 July 2014
Published online 22 July 2014 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24648



Scheme 1

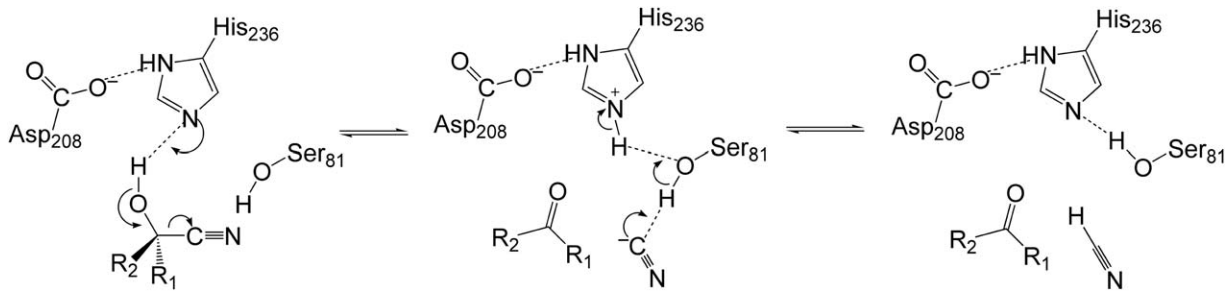
Reaction catalyzed by HNL.[2]

enzymes: MeHNL, HbHNL, SbHNL, PaHNL, and AtHNL,[4] among which MeHNL, HbHNL and SbHNL belong to the S-selective enzymes, whereas PaHNL and AtHNL are R-selective enzymes. As a typical (S)-hydroxynitrile lyase, HbHNL from *Hevea brasiliensis* has been experimentally and theoretically studied in detail.[22–24] According to the crystal structure, HbHNL belongs to the FAD-independent HNL, and is classified into the α/β-hydrolases fold family.[6,25] Like other α/β-hydrolases, the HbHNL catalytic reaction involves a Ser-His-Asp catalytic triad which acts as a general base to deprotonate the cyanohydrin substrate.[2,22,26,27] Besides, a positively charged Lys236 in the active site is required to stabilize the negative charge of cyano group during the reaction.

In 2007, an R-selective HNL enzyme named as AtHNL was discovered from *Arabidopsis thaliana*.[28] AtHNL is structurally similar with the S-selective HbHNL, with 45% identity and 67% similarity in sequence.[2] Before AtHNL was discovered, it was believed that only the S-selective hydroxynitrile lyases of the HNLs family contain an α/β-hydrolases fold.[29] But the discovery of AtHNL broke the accepted rule.[28] AtHNL also belongs to the α/β-hydrolases superfamily, and can use a wide range of substrates usually the non-chiral aldehyde or ketone and HCN to synthesize the (R)-cyanohydrins.[28,30,31] The catalytic triad Ser-His-Asp was also found in the active site of AtHNL, and the important roles of these three residues have been demonstrated by site-directed mutagenesis.[2] Specifically, Ser81, His236, and Asp208 were respectively replaced by the non functional residues Ala, Phe and Asn, and anyone of these variants could make the catalytic activity of AtHNL drastically lost.

Although AtHNL has been identified for several years, the detailed reaction mechanism of AtHNL is still not clear. In some enzymes involving the Ser-His-Asp catalytic triad, it was suggested that the Ser residue usually acts as a nucleophile. But this serine nucleophilic attack mechanism was not supported by the recently homology model and docking studies.[2] By comparing with HbHNL, a similar general acid/base mechanism of AtHNL has been proposed by Andexer et al. (Scheme 2).[2] It was suggested that residue His236 firstly abstracts a proton from the hydroxyl group of cyanohydrins, and then the C—C bond is cleaved to yield CN^-. One difference between the two proposed mechanisms of AtHNL and HbHNL is that, in HbHNL the generated intermediate CN^- is stabilized by a positively charged residue (Lys236), while in AtHNL the negative charge of CN^- is stabilized by an "oxyanion hole" and an α-helix dipole. Another difference is the two proton transfer processes take place in an opposite manner. In the cyanohydrin cleavage reaction of AtHNL, His236 directly abstracts a proton from the substrate, and then protonates CN^- via a mediator Ser81, whereas in the reaction of HbHNL, His235 abstracts a proton from the substrate via the mediator Ser80, and then directly denotes a proton to CN^-. Of course, the catalytic mechanism of AtHNL is only proposed on the basis of site-directed mutagenesis and a homology model, there is still no theoretical study addressed on the detailed catalytic mechanism of AtHNL.

Liu *et al.* have studied the catalytic mechanism of HbHNL by using density functional theory (DFT) with two simplified models,[22] which provided a strong theoretical support to the previously proposed mechanism, and gave the detailed reaction pathways and energetics of HbHNL. In this article, to understand the mechanism of AtHNL, the quantum mechanical/molecular mechanical (QM/MM) method[32,33] was employed to study all the intermediates and transition states involving in the reaction cycle. QM/MM method is becoming a popular approach for investigating enzymatic reactions.[34–36] On the basis of our calculations, the details of reaction



Scheme 2

Proposed mechanism for AtHNL.[2]

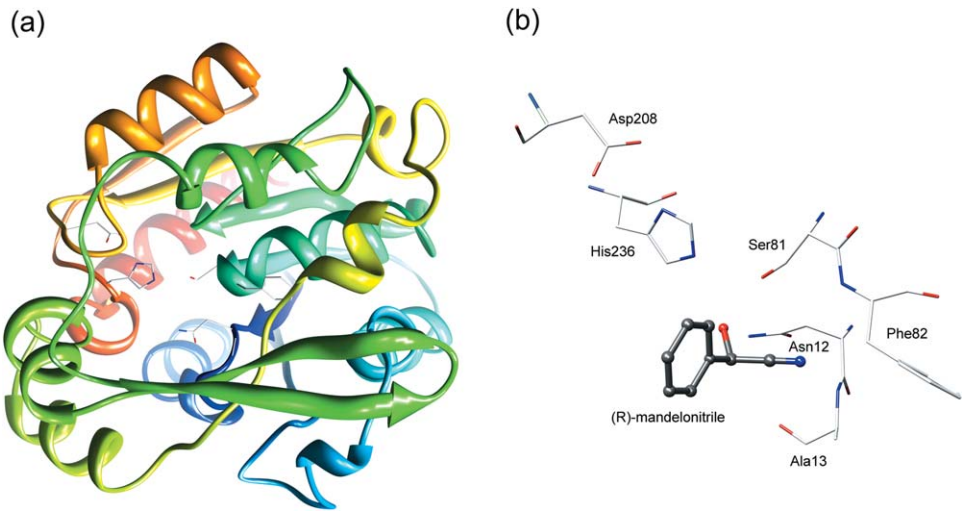


Figure 1

The crystal structure of AtHNL (**a**) and the active site structure derived from docking calculations (**b**). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

pathways and the roles of key pocket residues have been illuminated at atomistic level.

METHODS

Computational models

The crystal structure of AtHNL at 2.50 Å resolution (PDB code: 3DQZ)[2] was used to construct the computational model [Fig. 1(a)]. To obtain the structure of AtHNL in complex with its biological substrate, a (R)-mandelonitrile molecule was firstly docked into the active site of AtHNL using AutoDock Tools 4.0.[37] In the docking calculation, the grid map was set to 64 Å × 50 Å × 58 Å with the grid spacing of 0.375 Å, and 100 independent runs were performed, in which the structure of enzyme was kept rigid and the substrate was kept free. Docking results were analyzed at the criterion of RMSD (root-mean-square deviation) of 2.0 Å, and only one cluster was observed, which contains 100 very similar conformations. Thus, one representative docking structure was selected as the computational model [Fig. 1(b)]. Then the protonation states of all charged residues were checked using the PROPKA3.1[38–41] on the basis of experimental condition. As a residue of the catalytic triad, His236 plays important roles in the catalytic reaction, and the protonation state of His236

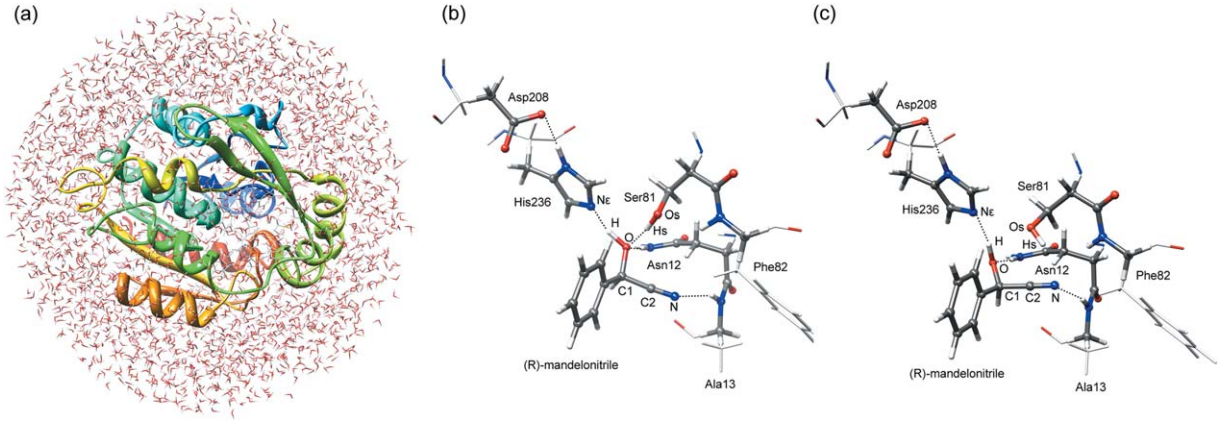


Figure 2

Optimized structures of solvated model (**a**) and the active sites of R_A (**b**) and R_B (**c**). The atoms in QM region are displayed in ball and stick model. Atoms mentioned in the text are labeled. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

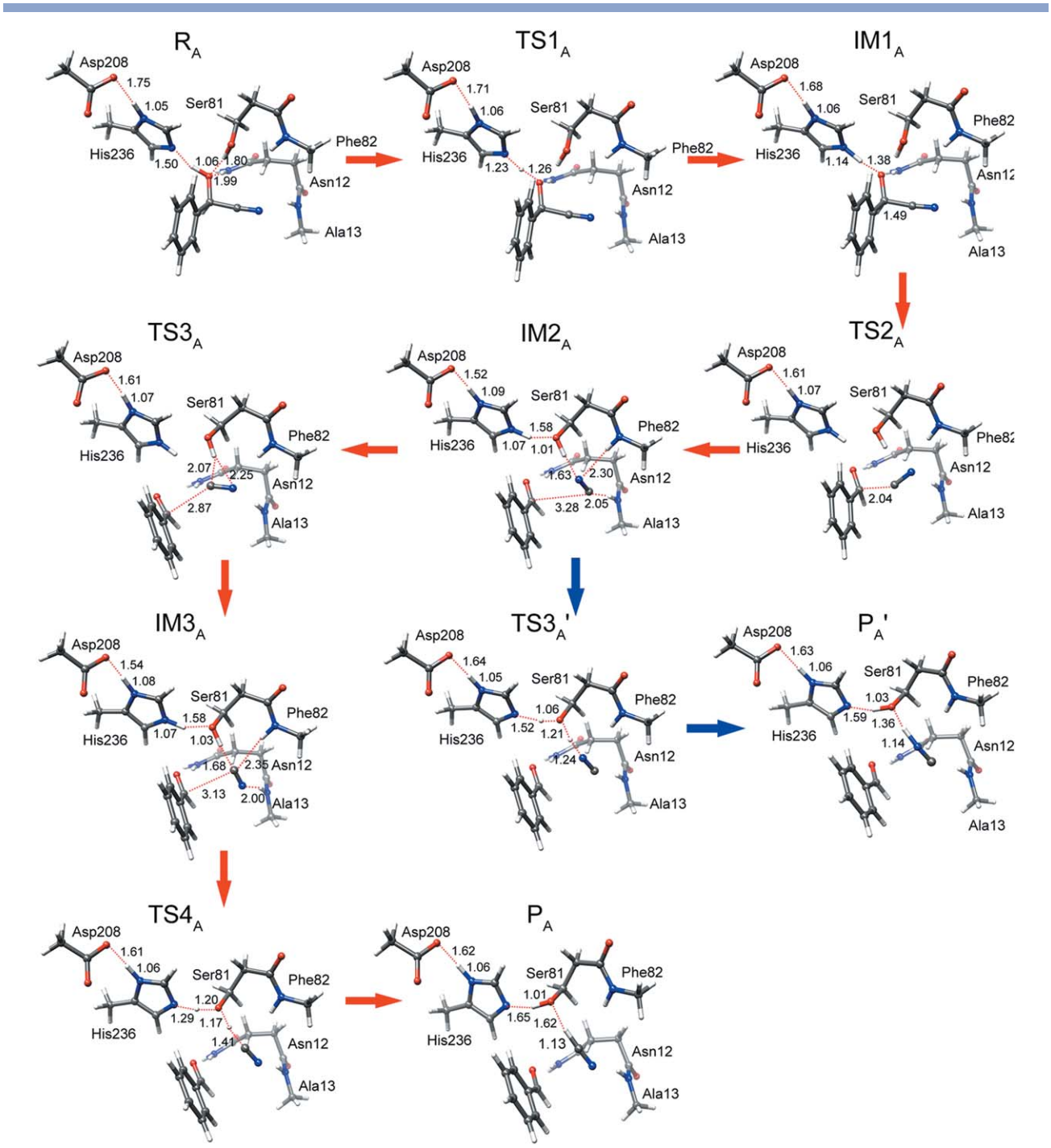


Figure 3

Optimized structures of the reactant, intermediates, transition states and products in Path A. The distances are given in angstrom. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

is crucial for the AtHNL catalytic reaction. Because His residue can adopt anyone of the three possible protonation states (on Nδ, Nϵ, or on both). Therefore, the protonation state of His236 needs to be specified. Based on the calculation results of PROPKA3.1 and the orientation of His236 in the crystal structure, His236 adopted its deprotonation state on the Nϵ atom. Besides, the other 9 His residues beyond the active site of this protein were also checked. According to the pK_a values and hydrogen bond networks around these residues, the same protonation states as His236 were assigned to these His residues. Other titratable residues were set to their normal protonation states in the following MD simulation and QM/

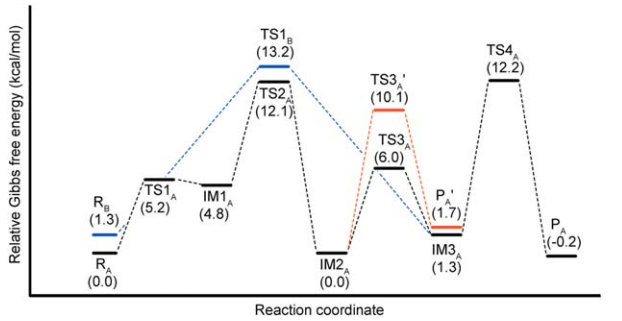


Figure 4

Relative Gibbs free energy profiles including dispersion corrections for Paths A and B. The free energy of reactant R_A was set to zero. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

MM calculations. The missing hydrogens of the crystal structure were added by the HBUILD program of the CHARMM package.[42] Subsequently, 3059 TIP3P[43] water molecules were placed into the enzyme model and 7 sodium cations were added to neutralize the system. Afterward, following a series of minimizations, 10 ns MD simulation was carried out using the CHARMM22/CMAP force field.[42,44,45] The obtained root-mean-squared deviations (RMSDs) of the whole system is shown in Supporting Information Figure S1. One can see that the backbone of the protein was basically equilibrated after 7 ns, which indicates that a 10 ns MD simulation is sufficient to equilibrate the protein. Thus, the last typical snapshot at 10 ns was selected for the following QM/MM calculations. To test whether the QM/MM energies derived from this local minima are stable and reliable, two other snapshots taken from the MD trajectory at 8 and 9 ns were also used in the calculations of pathway $R_A \rightarrow P_{A'}$.

QM/MM calculations

The reaction detail was studied by using QM/MM approach. In the calculations, the whole system was divided into QM and MM regions. Hydrogen link atoms were used to saturate and simulate the valencies and covalent bonds at the QM/MM interfaces.[46] The QM calculations were performed by density functional theory (DFT) with the B3LYP functional[47,48] in Turbomole,[49] which was built into ChemShell.[50] The MM part was characterized using the CHARMM22 force field[45] in DL-POLY.[51]

In the present work, the QM subsystem was composed by a complete (R)-mandelonitrile molecule, the side chains of Ser81, His236, Asp208, and Asn12, and part of residues Ala13 and Phe82. All other residues were included in the MM region. In the following QM/MM calculations, only the QM region and partial MM region within 13 Å of the (R)-mandelonitrile were allowed to move. All the rest atoms were frozen. The geometry optimizations of the QM region were calculated at 6–31G(d,p) level. To get the Gibbs free energies at 298.15 K, the single point energy calculations were performed at 6–311++G(3df,3pd) level,[47,48] whereas the frequencies were calculated with the 6–31G(d,p) basis set at the optimized structures. Besides, to correct the B3LYP energy for dispersion, the DFT-D3 program[52,53] was used to calculate the empirical dispersion corrections. Dispersion was found to have only a minor effect on the relative energies. Thus, all the reported energy items are single point energies at 6–311++G(3df,3pd) level including dispersion corrections. The NBO (natural bond orbital) analysis was also performed at 6–31G(d,p) level to obtain NPA (natural population atomic) charges.[54–56] The calculated structures of reactants and intermediates were served as initial structures to scan transition states and next intermediates or products. During the QM/MM calculations, the hybrid delocalized internal coordinates (HDLC) optimizer[57] was used for geometry optimization. Local minima were searched by a quasi-Newton L-BFGS algorithm (limited memory Broyden–Fletcher–Goldfarb–Shanno).[58,59] The L-BFGS algorithm is a second-order optimizer that takes an optimization step towards the minima, and it is particularly useful for optimization problems. The highest energy point between two local minima was used to search transition state by partitioned rational function optimization (P-RFO) algorithm[60] via following eigenmodes of the Hessian. The obtained transition states were described by a single negative eigenvalue. In our calculations, the QM/MM boundary was treated by the hydrogen link atoms with charge shift model,[61] and the

Table I

Natural Population Atomic Charges of Species in Pathway at B3LYP/6–31 G(d, p) Level

Species	R_A	$TS1_A$	$IM1_A$	$TS2_A$	$IM2_A$	$TS3_A$	$IM3_A$	$TS4_A$	P_A
N	−0.39	−0.41	−0.42	−0.65	−0.72	−0.68	−0.61	−0.55	−0.45
C1	0.01	0.02	0.02	0.28	0.42	0.42	0.42	0.42	0.41
C2	0.34	0.34	0.34	0.11	−0.06	−0.16	−0.11	−0.03	0.11
O	−0.83	−0.86	−0.88	−0.73	−0.59	−0.58	−0.59	−0.58	−0.57
H	0.52	0.50	0.49	0.49	0.49	0.48	0.49	0.51	0.53
Nε	−0.60	−0.57	−0.56	−0.54	−0.56	−0.54	−0.56	−0.60	−0.62
Os	−0.81	−0.82	−0.83	−0.80	−0.83	−0.81	−0.83	−0.81	−0.82
Hs	0.50	0.51	0.51	0.51	0.52	0.51	0.47	0.39	0.28

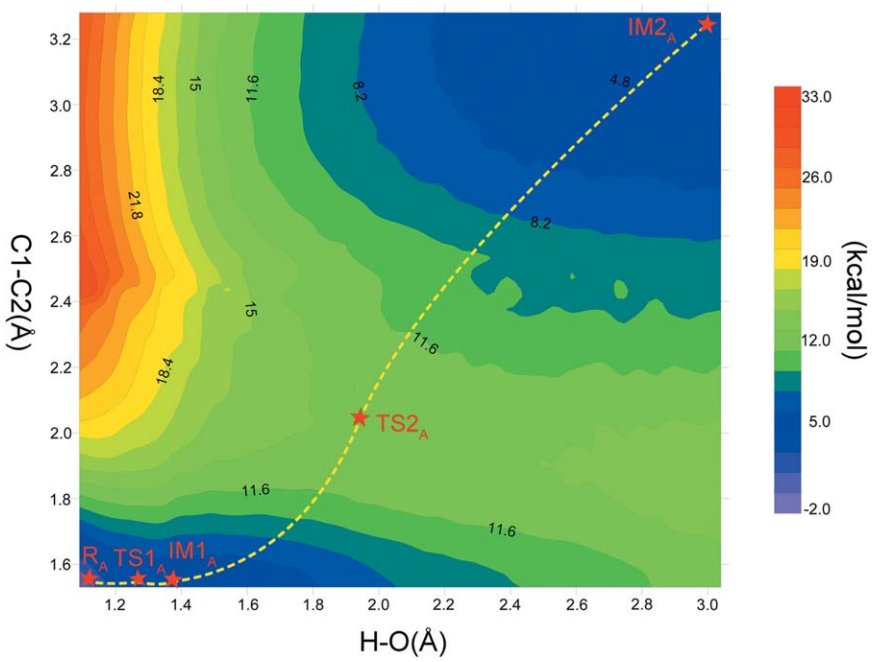


Figure 5

PES calculated at the B3LYP/6–31G(d) level using distances C1—C2 and H—O as reaction coordinates. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

electronic embedding scheme[62] was used for energy calculations. In the electronic embedding scheme, the MM region charges were included into one-electron Hamiltonian of QM calculations.[62] This scheme can avoid the hyperpolarization of the QM wave function and can describe the effect of MM region on QM region. As the potential energy surface (PES) can afford a visually intuitive way to study the reaction mechanism, the PESs for some crucial steps were also calculated. In this work, the PESs were characterized by two reaction coordinates related to the breaking or formation of chemical bonds. To obtain the three PESs, a total of 3504 structures were optimized, and the basis set of 6–31G(d) was used to save the computing resource.

RESULTS AND DISCUSSION

To construct the reaction model, the (R)-mandelonitrile molecule was placed into the active site by using AutoDockTools 4.0. The docking result is shown in Figure 1(b). One can see that the α-hydroxy group of the substrate is located in the vicinity of the side chains of Asn12, Ser81, and His236, and the cyano group is close to the peptide chain NH-groups of Ala13 and Phe82. After a series of minimizations, and subsequent MD simulation and QM/MM optimizations, the enzyme-substrate complexes were obtained. Figure 2(b,c) show the two main reactant states [denoted as R_A and R_B, as shown in Fig. 2(b,c), respectively] of the system, which differ only in the orientations of hydroxyl group of Ser81. According to our MD simulation results, the state R_A is the dominate conformation. In R_A, the hydroxyl group of (R)-mandelonitrile forms three hydrogen bonds with the Nε atom of His236, the hydroxyl group of Ser81 and amino group of Asn12, respectively. But in R_B, the hydrogen bond between the hydroxyl group of substrate and Ser81 no longer exists. Because of the minor conformational change of active site, R_A is more stable than R_B by 1.3 kcal mol^{-1} in free energy.

Using R_A and R_B as reactants, two possible pathways (Path A and B) were calculated, which will be discussed in detail in the following sections.

Path A

Starting from R_A, the optimized structures of intermediates, transition states and products are shown in Figure 3, and the corresponding relative Gibbs free energy profiles are shown in Figure 4. In R_A, a strong hydrogen bond network is formed between the substrate and the pocket residues. In this pathway, the reaction starts by a proton transfer from the hydroxyl group of (R)-mandelonitrile to His236. Via $TS1_A$, $IM1_A$ is formed by overcoming a low free energy barrier of 5.2 kcal mol^{-1} (Fig. 4), meaning this proton transfer is very easy to occur. It is noted that both the structures and free energies of $TS1_A$ and $IM1_A$ are very similar, because in $IM1_A$ the proton on the Nϵ atom of

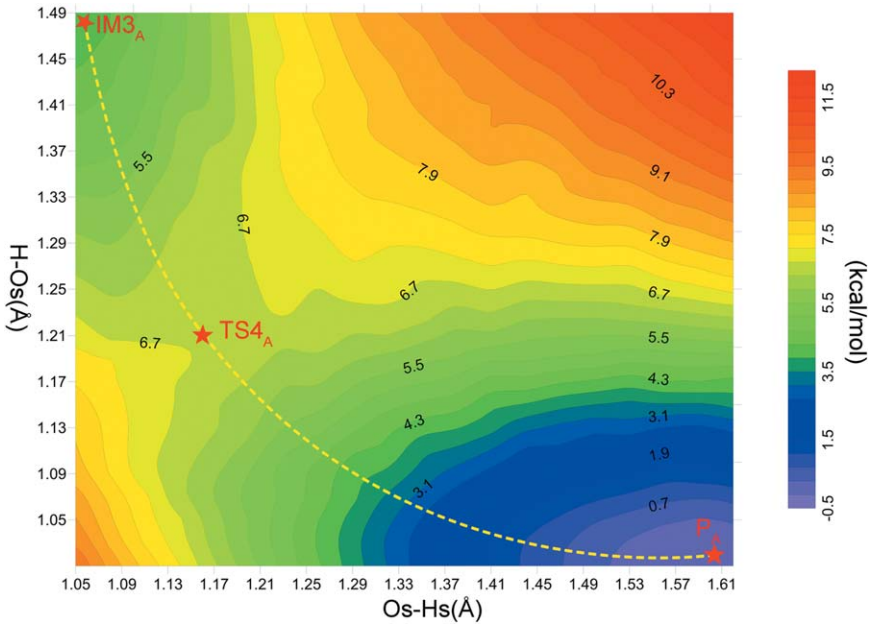


Figure 6

PES calculated at the B3LYP/6–31G(d) level using distances H—Os and Os—Hs as reaction coordinates. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

His 236 still forms a strong hydrogen bond with the hydroxyl oxygen anion. Besides, both in R_A and $IM1_A$, the amino group of Asn12 always forms a strong hydrogen bond with the hydroxyl group of substrate. Therefore, Asn12 is suggested to play an important role in stabilizing and fixing the substrate. To further understand this proton transfer process, NBO was also calculated at 6–31G(d,p) level. The NPA charges of all species in Path A are listed in Table I. One can see that the NPA charge on Nϵ atom of His236 changes from −0.60 au in R_A to −0.56 au in $IM1_A$, implying the negative charge on Nϵ atom is becoming smaller. In contrast, the charge of hydroxyl O atom of the substrate changes from −0.83 au in R_A to −0.88 au in $IM1_A$, suggesting the hydroxyl O atom is getting more negative.

The next step is the cleavage of C1—C2 bond of the substrate leading to the remove of cyano group. In $TS2_A$, the C1—C2 bond is greatly weakened with a distance of 2.04 Å. In $IM2_A$, C1—C2 bond is completely broken with

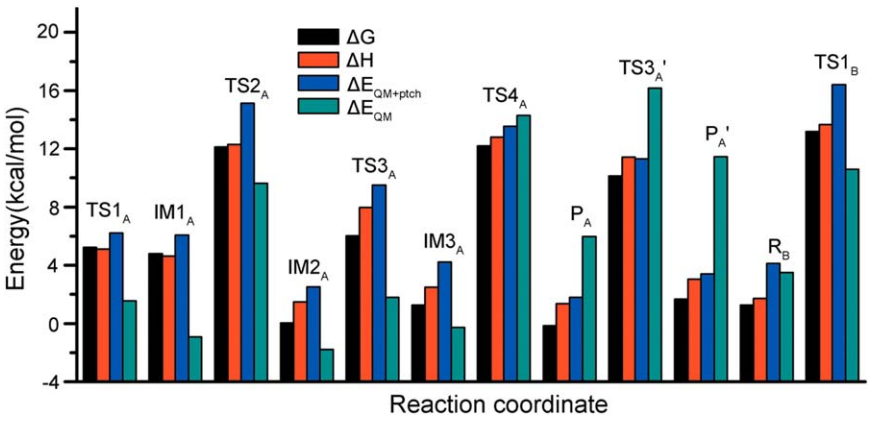


Figure 7

The relative free energies (ΔG) (black), enthalpies (ΔH) (red), QM region energies calculated at the field of MM point charges ($\Delta E_{QM+ptch}$) (blue) and QM region energies without point charges (ΔE_{QM}) (green). All these energies include dispersion corrections. In this plot, the energy of reactant R_A was set to zero. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

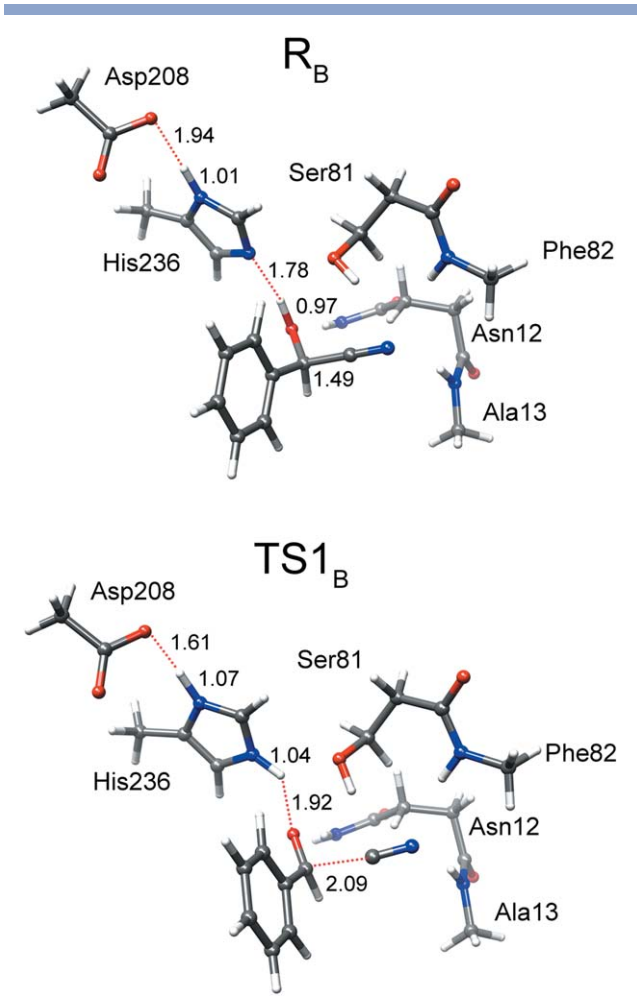


Figure 8

Optimized structures of R_B and $TS1_B$. The distances are given in angstrom. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

a distance of 3.28 Å, yielding the benzaldehyde and the negative ion (CN^-). This process corresponds to a free energy barrier of 7.3 kcal mol^{-1}. Companying this process, the hydroxyl group of Ser81 switches its orientation from towards the O atom in $IM1_A$ to the N atom of CN^-. Besides, the C1—C2 bond cleavage causes some obvious changes of NPA charges. For example, the charge on C1 atom changes from 0.02 au in $IM1_A$ to 0.42 au $IM2_A$, while the charge on C2 atom changes from 0.34 au in $IM1_A$ to −0.06 au $IM2_A$. In particularly, the charge on the N atom of the released CN^- changes to −0.72 au. This negative ion forms two hydrogen bonds with Ser81 and Ala13 with distances of 1.63 and 2.05 Å, respectively. Thus, the released CN^- is stabilized by Ser81 and the so called "oxyanion hole" (main-chain NH-groups of Ala13 and Phe82). To further understand this C1—C2 bond cleavage, a PES including the proton transfer from the substrate to His236 and the cleavage of C1—C2 bond to generate CN^- was calculated, which is shown in Figure 5. In Figure 5, the proton transfer coordinate is described in horizontal axis, and the C1—C2 bond cleavage is represented in vertical axis. The variation range of reaction coordinate for proton transfer was from 1.09 to 3.04 Å, and that for C1—C2 bond cleavage was from 1.53 to 3.28 Å. The increments of both reaction coordinates were assigned to 0.05 Å. Thus, 1440 (40 × 36) structures were obtained in this PES. As shown in Figure 5, two low-energy regions were recognized in the lower left corner and upper right corner, respectively. According to the reaction coordinates, the O—H and C1—C2 bonds of the structure in the upper right corner have been completely cleaved. Thus, the low-energy region in the upper right corner represents the structure of $IM2_A$, and the low-energy region in the lower left corner represents the structure of R_A.

The negatively charged CN^- has a strong tendency to abstract a proton from the nearby residues. By checking the pocket residues, Ser81 is the most possible proton donor. But to form HCN, the free CN^- should firstly undergo a rotation in the active site. Therefore, the rotation of CN^- was calculated, and the corresponding transition state ($TS3_A$) and intermediate ($IM3_A$) are shown in Figure 3, and their relative free energies are shown in Figure 4. Figure 4 shows that the rotation of CN^- corresponds to an energy barrier of 6.0 kcal mol^{-1}. In $TS3_A$, the hydroxyl Hs atom of Ser81 forms a triangular structure with the C2 and N atoms of CN^- with distances of 2.07 and 2.25 Å, respectively. Like in the structure of $IM2_A$, in $IM3_A$ the CN^- forms a hydrogen bond network with residues Ser81, Ala13, and Phe82. In addition, in $IM3_A$ a hydrogen bond chain is formed by Asp208, His236, Ser81 and CN^-, which greatly facilitates the protonation of CN^- in the next step.

As shown in Figures 3 and 4, the protonation of CN^- to produce the hydrogen cyanide (HCN) is easy. The H atom at the Nϵ atom of His236 transfers to CN^- by a mediator Ser81, corresponding to a free energy barrier of 10.9 kcal mol^{-1}. In the structures of $IM3_A$, $TS4_A$, and P_A, Ser81 always form strong hydrogen bonds with His236 and CN^- or HCN, acting as an important bridge for proton transfer. The NPA charges in Table I show that the charge distributions on C2 and N atoms only change slightly from $IM3_A$ to $TS4_A$, while they increase by 0.15 and 0.10 au from $TS4_A$ to P_A. To further understand this protonation process, the PES of this step was calculated at the level of 6–31G(d), which is shown in Figure 6. The horizontal axis describes the change of distance between Os and Hs atoms of Ser81, and the vertical axis characterizes the change of distance between H and Os atoms. The increment was designated 0.03 Å for Os-Hs and 0.02 Å for H-Os. Therefore, 500 (20 × 25) structures were optimized in this PES. As shown in Figure 6, the upper left corner and lower right corner represent the structures of $IM3_A$ and P_A, respectively, which

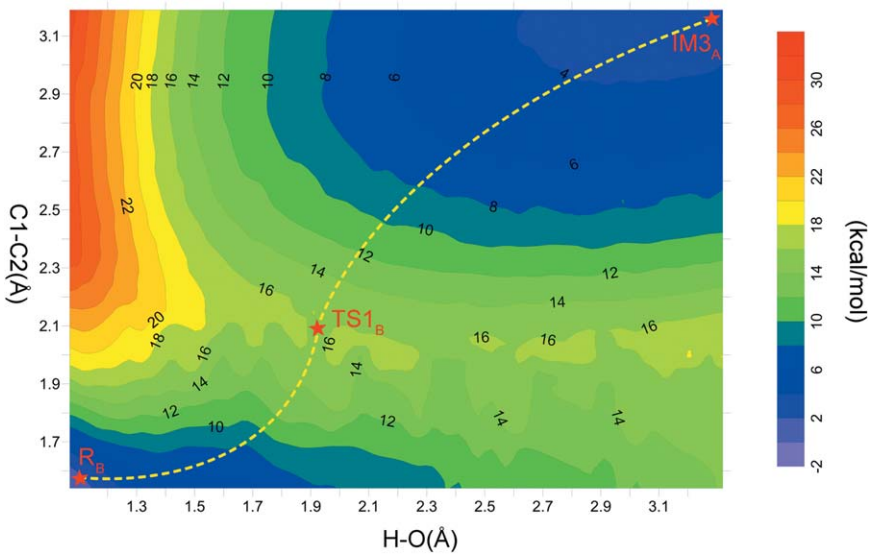


Figure 9

PES calculated at the B3LYP/6–31G(d) level using distances C1—C2 and H—O as reaction coordinates. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

clearly show that the protonation of CN^- by His236 via a Ser81 mediator is a concerted reaction.

According to the structure of $IM2_A$, one can see that the hydrogen bond network in the active site is also favorable for the proton transfer from the hydroxyl group of Ser81 to the N atom of CN^- to generate hydrogen isocyanide (HNC) without undergoing the rotation of CN^-. It is indicated that HNC is another product of cyanohydrin cleavage catalyzed by AtHNL. HNC may be isomerized to HCN in water solution after it is released from active site. As shown in Figure 4, this step corresponds to a free energy barrier of 10.1 kcal mol^{-1}.

As we know, the QM/MM energy of the whole system may exhibit some instability owing to the multiple local-minima of enzyme-substrate complex. Therefore, to test whether the obtained QM/MM energies are stable and reliable, other two snapshots taken from the MD trajectory at 8 and 9 ns were also used to optimize the local minima of the reactant, which were further used in the calculations of pathway of $R_A \rightarrow P_{A'}$. Superimposition of the optimized structures of intermediates, transition states and products in pathway of $R_A \rightarrow P_{A'}$ is shown in Supporting Information Figure S2, and the corresponding energy profiles are shown in Supporting Information Figure S3. It can be seen that all the structures derived from the three snapshots are superimposed well, and the energy profiles along the pathway $R_A \rightarrow P_{A'}$ are basically kept stable, which indicate that our calculations are reliable.

In Path A, the C1—C2 bond cleavage step corresponds to a free energy barrier of 12.1 kcal mol^{-1}, while the protonation of CN^- to produce the hydrogen cyanide (HCN) corresponds to a free energy barrier of 12.2 kcal mol^{-1}. Therefore, both of the two processes contribute a lot to rate-limiting.

To evaluate how the electrostatic effect of MM part influences the energies of QM region, the QM/MM energies were further divided into E_{MM} and $E_{QM+pcth}$ components. The E_{MM} component represents the effect of surroundings, and $E_{QM+pcth}$ component is a QM energy calculation in the field of MM point charges, which can be derived from our QM/MM calculations. Thus, the difference between $E_{QM+pcth}$ and the E_{QM} (the same calculation without point charges at the same level by Gaussian 03 package[63]) can describe the electrostatic effect of the surroundings. Figure 7 shows the comparison of relative Gibbs free energies (ΔG), enthalpies (ΔH), $\Delta E_{QM+pcth}$ and ΔE (the data of these items are shown in Supporting Information Table SI). The differences between ΔG and ΔH are minor (<2.0 kcal mol^{-1}), which means the entropy effects are very small. Whereas the differences between $\Delta E_{QM+pcth}$ and ΔE are large, indicating the electrostatic effects of MM part have great influence on the QM region.

Path B

By the rotation of hydroxyl of Ser81, the reactant state R_A can be converted to R_B via an energy barrier of about 8 kcal mol^{-1}. Using R_B as the reactant, the reaction mechanism of AtHNL was also investigated by using QM/MM method. The optimized structures of R_B and $TS1_B$ are shown in Figure 8. It is found that the proton transfer from the hydroxyl group of substrate and the cleavage of C1—C2 bond occur in a concerted but

asynchronous manner. In $TS1_B$, the proton of hydroxyl group of the substrate has been completely transferred to His236, but the distance between C1 and C2 atoms is only changed to 2.09 Å, indicating the C1—C2 bond still on the way to be broken. As shown in Figure 4, the C1—C2 bond cleavage corresponds to a free energy barrier of 11.9 kcal mol^{-1}. To check whether the proton transfer and C1—C2 bond cleavage are concerted, the PES from R_B to $IM3_A$ was also calculated, as shown in Figure 9. The horizontal axis describes the distance between H and O atoms, and the vertical axis designates the corresponding C1—C2 bond cleavage. The reaction coordinate of H—O ranges from 1.01 to 3.32 Å and that of C1—C2 ranges from 1.54 to 3.19 Å by the same increment of 0.05 Å. Therefore, 1564 (46 × 34) optimized structures were achieved. The structure located at the upper right corner corresponds to $IM3_A$, in which C1—C2 bond has been broken with a distance of 3.19 Å. The structure located at the lower left corner represents the structure of R_B. The transition state $TS1_B$ is situated in the center left, in which the reaction coordinate of H—O is ~1.9 Å and C1—C2 is ~2.1 Å. From Figure 9, we can also see that when the C1—C2 bond falls in the range of 2.0–2.1 Å, the potential energy surface is very flat along the horizontal axis with the distance of H—O changing from 1.7 to 3.0 Å. It means the transition state $TS1_B$ may correspond to a large area but not a single point. Not surprisingly, based on our calculations, another transition state ($TS1_B'$) was recognized around H—O = 2.36 Å and C1—C2 = 1.94 Å. The relative energies of this two transition states are very close (16.2 vs. 17.0 kcal mol^{-1} without including the dispersion corrections). Figure 9 clearly shows that the C1—C2 bond cleavage and the proton transfer are concerted.

From R_B via $TS1_B$, the intermediate $IM3_A$ is formed (Fig. 4), which then undergoes a proton transfer to generate HCN and benzaldehyde as the same as in Path A. Thus, only two elementary steps are included in Path B, and the cleavage of C1—C2 bond is the rate limiting step.

CONCLUSION

The reaction mechanism of cyanohydrin cleavage catalyzed by AtHNL has been studied by using QM/MM method. Two optimized structures (R_A and R_B) were used as reactants and two possible reaction pathways were obtained. In Path A, the hydroxyl group of the cyanohydrin was firstly deprotonated by His236, and then the C1—C2 bond was broken to generate the CN^- anion. To form HCN, the CN^- anion should experiences a rotation, and subsequently abstracts a proton from His236 via Ser81. The cleavage C1—C2 bond corresponds to a free energy barrier of 12.1 kcal mol^{-1}, and the protonation of CN^- corresponds to a free energy barrier of 12.2 kcal mol^{-1}, suggesting both of the two processes contribute a lot to rate limiting. Besides, another compound HNC can also be produced in Path A, which could be isomerized to HCN in water solution. In Path B, the deprotonation of the hydroxyl group of cyanohydrin and the cleavage of C1—C2 bond occur in a concerted but asynchronous manner, and this step corresponds to a free energy barrier of 13.2 kcal mol^{-1} relative to R_A. Subsequently, Path B follows the same proton transfer process as Path A. Based on the experimental data of HbHNL,[64] the estimated free energy barrier is 17.1 kcal mol^{-1}, and the theoretical values calculated from the two models are 13.5 and 14.9 kcal mol^{-1},[22] respectively. Therefore, we can conclude that the energy barriers of Path A and Path B basically agree well with the experimental and theoretical results of HbHNL, a similar enzyme of AtHNL. Thus, both Path A and B are suggested to be possible. Our calculations also reveal that His236, Ser81, and Asp208 compose the catalytic triad and act as the general acid/base. Ala13, Phe82, and Ser81 play important roles in stabilizing the negative ion CN^-.

REFERENCES

1. Poulton JE. Localization and catabolism of cyanogenic glycosides. In: Evered D, Harnett S, editors. Cyanide compounds in biology. Chichester: Wiley; 1988. pp 67–91.
2. Andexer JN, Staunig N, Eggert T, Kratky C, Martina Pohl, Gruber K. Hydroxynitrile lyases with α/β-hydrolase fold: two enzymes with almost identical 3D structures but opposite enantioselectivities and different reaction mechanisms. ChemBioChem 2012;13:1932–1939.
3. Hanefeld U. Immobilisation of hydroxynitrile lyases. Chem Soc Rev 2013;42:6308–6321.
4. Dadashipour M, Asano Y. Hydroxynitrile lyases: insights into biochemistry, discovery, and engineering. ACS Catal 2011;1:1121–1149.
5. Conn EE. Cyanogenic glycosides. In: Stumpf PK, Conn EE, editors. The biochemistry of plants: a comprehensive treatise. New York: Academic Press; 1981. pp 479–500.
6. Hickel A, Hasslacher M, Griengl H. Hydroxynitrile lyases: functions and properties. Physiol Plant 1996;98:891–898.
7. Wajant H, Effenberger F. Hydroxynitrile lyases of higher plants. Biol Chem 1996;377:611–617.
8. Lieberei R, Selmar D, Biehl B. Metabolization of cyanogenic glucosides in hevea brasiliensis. Plant Syst Evol 1985;150:49–63.
9. Selmar D. Apoplastic occurrence of cyanogenic beta-glucosidases and consequences for the metabolism of cyanogenic glucosides. ACS Symp Ser 1993;13:191–204.
10. Effenberger F. Synthesis and reactions of optically active cyanohydrins. Angew Chem Int Ed Engl 1994;33:1555–1564.
11. Griengl H, Hickel A, Johnson DV, Kratky C, Schmidt M, Schwab H. Enzymatic cleavage and formation of cyanohydrins: a reaction of biological and synthetic relevance. Chem Commun 1997;20:1933–1940.
12. Klempier N, Pichler U, Griengl H. Synthesis of α/β-unsaturated (S)-cyanohydrins using the oxynitrilase from *Hevea brasiliensis*. Tetrahedron Asymmetry 1995;6:845–848.
13. Müller M. Recent developments in enzymatic asymmetric C-C bond formation. Adv Synth Catal 2012;354:3161–3174.
14. Purkarthofer T, Gruber K, Fechter MH, Griengl H. Asymmetric biocatalytic hydrocyanation of pyrrole carboxaldehydes. Tetrahedron 2005;61:7661–7668.

15. Fechter MH, Griengl H. Hydroxynitrile lyases: biological sources and application as biocatalysts. Food Technol Biotech 2004;42:287–294.
16. Purkarthofer T, Skranc W, Schuster C, Griengl H. Potential and capabilities of hydroxynitrile lyases as biocatalysts in the chemical industry. Appl Microbiol Biotechnol 2007;76:309–320.
17. Sharma M, Sharma NN, Bhalla TC. Hydroxynitrile lyases: at the interface of biology and chemistry. Enzyme Microb Technol 2005; 37:279–294.
18. Qian C, Zhu C, Huang T. Enantioselective trimethylsilylcyanation of aldehydes catalyzed by chiral lanthanoid alkoxides. J Chem Soc Perkin Trans 1998;1:2131–2132.
19. Kogut EF, Thoen JC, Lipton MA. Examination and enhancement of enantioselective autoinduction in cyanohydrin formation by cyclo[(R)-His-(R)-Phe]. J Org Chem 1998;63:4604–4610.
20. Johnson DV, Zabelinskaja-Mackova AA, Griengl H. Oxynitrilases for asymmetric C–C bond formation. Curr Opin Chem Biol 2000;4: 103–109.
21. Andexer JN, Langermann JV, Kragl U, Pohl M. How to overcome limitations in biotechnological processes-examples from hydroxynitrile lyase applications. Trends Biotechnol 2009;27:599–607.
22. Cui FC, Pan XL, Liu JY. Catalytic mechanism of hydroxynitrile lyase from hevea brasiliensis: a theoretical investigation. J Phys Chem B 2010;114:9622–9628.
23. Wagner UG, Hasslacher M, Griengl H, Schwab H, Kratky C. Mechanism of cyanogenesis: the crystal structure of hydroxynitrile lyase from *Hevea brasiliensis*. Structure 1996;4:811–822.
24. Hasslacher M, Schall M, Hayn M, Griengl H, Kohlwein SD, Schwab H. Molecular cloning of the full-length cDNA of (S)-hydroxynitrile lyase from *Hevea brasiliensis*. J Biol Chem 1996;271:5884–5891.
25. Hanefeld U, Stranzl G, Straathof A J, Heijnen JJ, Bergmann A, Mittelbach R, Glatter O, Kratky C. Electrospray ionization mass spectrometry, circular dichroism and SAXS studies of the (S)-hydroxynitrile lyase from *Hevea brasiliensis*. Biochim Biophys Acta 2001;1544:133–142.
26. Gruber K, Gartler G, Krammer B, Schwab H, Kratky C. Reaction mechanism of hydroxynitrile lyases of the α/β-hydrolase superfamily. J Biol Chem 2004;279:20501–20510.
27. Zuegg J, Gruber K, Gugganig M, Wagner UG, Kratky C. Three-dimensional structures of enzyme-substrate complexes of the hydroxynitrile lyase from *Hevea brasiliensis*. Protein Sci 1999;8: 1990–2000.
28. Andexer J, Langermann von J, Mell A, Bocola M, Kragl U, Eggert T, Pohl M. An R-selective hydroxynitrile lyase from arabidopsis thaliana with an α/β-hydrolase fold. Angew Chem Int Ed 2007;46:8679–8681.
29. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J. The alpha/beta hydrolase fold. Protein Eng 1992;5:197–211.
30. Okrob D, Paravidino M, Orru RVA, Wiechert W, Hanefeld U, Pohl M. Hydroxynitril lyase from arabidopsis thaliana: identification of reaction parameters for enantiopure cyanohydrin synthesis by pure and immobilized catalyst. Adv Synth Catal 2011;353:2399–2408.
31. Guterl JK, Andexer JN, Sehl T, von Langermann J, Frindi-Wosch I, Rosenkranz T, Fitter J, Gruber K, Kragl U, Eggert T, Pohl M. Uneven twins: comparison of two enantiocomplementary hydroxynitrile lyases with alpha/beta-hydrolase fold. J Biotechnol 2009;141: 166–173.
32. Warshel A, Karplus M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. formulation and parametrization. J Am Chem Soc 1972;94:5612–5625.
33. Warshel A, Levitt M. Theoretical studies of enzymatic reactions: dielectric electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. J Mol Biol 1976;103:227–249.
34. Lin H, Truhlar DG. QM/MM: what have we learned, where are we, and where do we go from here? Theor Chem Acc 2007;117:185–199.
35. Friesner RA, Guallar V. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. Annu Rev Phys Chem 2005;56:389–427.
36. Zhang Y, Liu L, Wu L, Li S, Li F, Li Z. Theoretical investigation on the diatomic ligand migration process and ligand binding properties in non-O2-binding H-NOX domain. Proteins 2013;81:1363–1376.
37. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 1998;19:1639–1662.
38. Li H, Robertson AD, Jensen JH. Very fast empirical pre-diction and rationalization of protein pKa values. Proteins Struct Funct Bioinf 2005;61:704–721.
39. Bas DC, Rogers DM, Jensen JH. Very fast prediction and rationalization of pKa values for protein-ligand complexes. Proteins Struct Funct Bioinf 2008;73:765–783.
40. Olsson MHM, Søndergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. J Chem Theory Comput 2011;7:525–537.
41. Søndergaard CR, Olsson MHM, Rostkowski M, Jensen JH. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values. J Chem Theory Comput 2011;7:2284–2295.
42. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 1983;4:187–217.
43. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
44. Chen JH, Im WP, Brooks CL. Balancing solvation and intramolecular interactions: toward a consistent generalized born force field. J Am Chem Soc 2006;128:3728–3736.
45. 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. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
46. de Vries AH, Sherwood P, Collins SJ, Rigby AM, Rigutto M, Kramer GJ. Zeolite structure and reactivity by combined quantum-chemical-classical calculations. J Phys Chem B 1999;103:6133–6141.
47. Lee C, Yang W, Parr RG. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 1988;37:785–789.
48. Becke AD. Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 1993;98:5648–5652.
49. Ahlrichs R, Bar M, Haser M, Horn H, Kolmel C. Electronic structure calculations on workstation computers: the program system turbomole. Chem Phys Lett 1989;162:165–169.
50. Sherwood P, de Vries AH, Guest MF, Schreckenbach G, Catlow CRA, French SA, Sokol AA, Bromley ST, Thiel W, Turner AJ, Billeter S, Terstegen F, Thiel S, Kendrick J, Rogers SC, Casci J, Watson M, King F, Karlsen E, Sjøvoll M, Fahmi A, Schäfer A, Lennartz C. QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis. J Mol Struct (Theochem) 2003;632:1–28.
51. Smith W, Forester TR. DL_POLY_2.0: a general-purpose parallel molecular dynamics simulation package. J Mol Graph 1996;14:136–141.
52. Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 2010;132:154104–154123.

53. Grimme S, Ehrlich S, Goerigk L. Effect of the damping function in dispersion corrected density functional theory. J Comp Chem 2011; 32:1456–1465.
54. Reed AE, Weinhold F. Natural localized molecular orbitals. J Chem Phys 1985;83:1736–1740.
55. Reed AE, Weinstock RB, Weinhold F. Natural population analysis. J Chem Phys 1985;83:735–746.
56. Reed AE, Curtiss LA, Weinhold F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. Chem Rev 1988; 88:899–926.
57. Billeter SR, Turner AJ, Thiel W. Linear scaling geometry optimisation and transition state search in hybrid delocalised internal coordinates. Phys Chem Chem Phys 2000;2:2177–2186.
58. Nocedal J. Updating quasi-Newton matrices with limited storage. Math Comput 1980;35:773–782.
59. Liu DC, Nocedal J. On the limited memory method for large scale optimization. Math Prog 1989;45:503–528.
60. Banerjee A, Adams N, Simons J, Shepard R. Search for stationary-points on surface. J Phys Chem 1985;89:52–57.
61. de Vries AH, Sherwood P, Collins SJ, Rigby AM, Rigutto M, Kramer GJ. Zeolite structure and reactivity by combined quantum-chemical-classical calculations. J Phys Chem B 1999;103:6133–6141.
62. Bakowies D, Thiel W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. J Phys Chem 1996;100:10580–10594.
63. Frisch MJ, Trucks GW, Schlegel H B, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Jr., Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian 03 (Revision D.01). Wallingford, CT: Gaussian, Inc., 2004.
64. Selmar R, Lieberei R, Biehl B, Conn EE. R-hydroxynitrile lyase in *Hevea brasiliensis* and its significance for rapid cyanogenesis. Physiol Plant 1989;75:97–101.