PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

PCCP

RSCPublishing

PAPER

A Full Picture of Enzymatic Catalysis by Hydroxynitrile Lyases from *Hevea Brasiliensis*: Protonation Dependent Reaction Steps and Residue-Gated Movement of Substrate and Product

Cite this: DOI: 10.1039/xoxxooooox

Received ooth September 2014,
Accepted ooth September 2014

DOI: 10.1039/xoxxooooox

www.rsc.org/

Yuan Zhao,[a] Nanhao Chen,[b] Yirong Mo[c] and Zexing Cao*[a]

Hydroxynitrile lyases (HNLs) defend plants from herbivores and microbial attack by releasing cyanide from hydroxynitriles. The reverse process has been productively applied to bioorganic syntheses of pharmaceuticals and agrochemicals. To improve our understanding of the catalytic mechanism of HNLs, extensive ab initio QM/MM and classical MM molecular dynamics simulations have been performed to explore the catalytic conversion of cyanohydrins into aldehyde (or ketone) and HCN by hydroxynitrile lyases from *Hevea brasiliensis* (*Hb*HNLs). It was found that the catalytic reaction approximately follows a two-stage mechanism. The first stage involves two fast processes including the proton abstraction of substrate through a double-proton transfer and the C-CN bond cleavage, while the second stage concerns the HCN formation and is rate-determining. The complete free energy profile exhibits a peak of ~18 kcal/mol. Interestingly, the protonation state of Lys236 only influences the efficiency of the enzyme to some extent, but it changes the entire catalytic mechanism. The dynamical behaviors of the substrate delivery and HCN release are basically modulated by the gate movement of Trp128. The remarkable exothermicity of substrate binding and the facile release of HCN may drive the enzyme-catalyzed reaction to proceed along the substrate decomposition efficiently. Computational mutagenesis reveals the key residues which play an important role in the catalytic process.

1 Introduction

Hydroxynitrile lyases (HNLs) are important members of the α/β-hydrolase superfamily, and they catalyze the cleavage of cyanohydrins into aldehyde (or ketone) and hydrocyanic acid (HCN). The release of HCN not only can protect plant system from herbivores and microbial attack[1-3] but also provides a nitrogen source for the biosynthesis of asparagine.[4-5] The reverse in vivo reactions may also occur at different conditions, with the efficient enantioselectivity for the synthesis of chiral compounds.[6-12] In recent years, HNLs have been successfully utilized in the production of pharmaceuticals and agrochemicals owing to their importance in biocatalytic retrosynthesis.[13-20] Certainly, an improved understanding of the enzymatic catalysis mechanism of HNLs potentially can further optimize the production processes in industry and help the rational design of biocatalysts.

Herein, the hydroxynitrile lyase from *Hevea brasiliensis* (*Hb*HNL) has been considered. This enzyme can catalyze the formation of carbon-carbon bond with high (S)-stereoselectivity during the chiral molecule synthesis.[19,22] Up to now, a number of *Hb*HNL crystal structures, including the apo state and the complex state, have been determined (see Table S1 in the † Electronic Supplementary Information (ESI)).[21,23-28] These static structures provide an opportunity for subsequent computational simulations which can reveal the dynamical conformational evolution along the enzyme-catalyzed reactions.

Based on the crystal structures of the wild-type enzyme and its complexes with inhibitor or substrate in combination with the activity analyses of site-directed mutants, a general acid/base catalysis has been experimentally proposed as the most probable mechanism (see Figure 1),[21] where Ser80, His235 and Asp207 serve as the catalytic triad to initiate the reaction. It should be noted that all members of the α/β hydrolases contain a conserved catalytic triad (nucleophile-histidine-aspartate),[29] and in HNLs the nucleophile is serine. In consideration of the enzymatic environment, it was hypothesized that the deprotonation and the C-C bond cleavage

of substrate occur with a concerted mechanism, followed by the HCN formation and release. We note that previous density functional calculations with an active-site model proposed a three-step mechanism (where the deprotonation and the C-C bond cleavage occur separately. See Figure 1) and the predicted relative energies for the reaction severely depend on the choice of cluster models.[30]

The kinetic characterization on the enzyme mutants indicates that the protonation state of Lys236 plays a critical role in substrate binding, together with Thr11 and Ser80.[21] Moreover, the positively-charged -NH_3^+ group of the protonated Lys236 can stabilize the nascent cyanide ion which acts as a better leaving group. The fact that the cyanide ion is a leaving group is very important in the recycling of the enzyme back to its initial state. Based on the X-ray determined crystal structures of *Hb*HNL complexs with non-natural chiral substrates, Gartler et al.[25] proposed that Lys236 may be involved in determining the enantioselectivity of the enzyme. They also claimed that the binding modes of the chiral substrates were identical with the acetone cyanohydrin. This means that chiral and achiral substrates may undergo the very same mechanism. However, the detailed elucidation of the effects of key residues on both substrate binding and catalytic process is still elusive.

It has been well known that the accessibility of the active site for the substrate binding and the product release makes important contribution to the overall enzymatic efficiency. The crystal structure of *Hb*HNL indicates that the active site is deeply buried inside the enzyme and there is only one narrow channel for the substrate delivery.[24] In the channel for the ligand transportation, the flexibility of the side-chain residue Trp128 can be expected to play a nontrivial role.[25] However, plausible mechanisms and dynamical properties of acetone cyanohydrin delivery and hydrocyanic acid release also remain unknown.

It is thus expected that computational studies can provide a microscopic description of the catalytic process and shed light on the exact roles of individual residues and the details of substrate delivery and product release. Our computational investigations will basically focus on the key issues related to the whole enzymatic catalysis accordingly. The questions to be addressed include, (i) what is the most likely catalytic mechanism for the enzymatic reaction in the protein environment? Are the proposed reaction steps of double-proton transfer, C-C bond cleavage, and HCN formation stepwise or concerted? And which is the rate-determining step? (ii) What is the protonation state of Lys236 on these catalytic processes and how does this protonation state influence the overall catalytic reaction? (iii) What are the precise roles of the key residues in the whole catalytic process? (iv) What are the transportation mechanisms for the substrate acetone cyanohydrin and the product hydrocyanic acid? In an attempt to clarify these issues, we have conducted both classical molecular dynamics (MD) and combined quantum mechanics and molecular mechanics (QM/MM) MD simulations, which are expected to generate a full picture for the overall enzymatic catalysis and provide detailed information in enzyme engineering for biosynthesis of organic molecules.

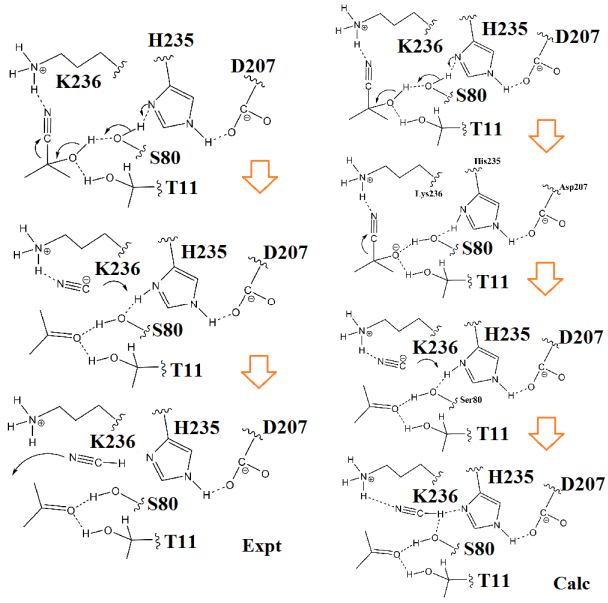

Fig. 1 The proposed catalytic mechanisms by experiment[21] and by the QM calculation.[30]

2 Computational Methods

Since the achiral and chiral substrates share the same binding mode and follow similar catalytic mechanisms, the initial computational model was built based on the X-ray crystal structures of *Hb*HNL complexed with the achiral acetone cyanohydrin (PDB: 1SC9).[21] The protonation states of ionizable residues were determined at pH = 7.5 via the program PROPKA 3.1[32-34] as well as the previous Poisson-Boltzmann calculations.[21] The Amber99SB force field[35-37] and TIP3P model[38] were employed for the protein and water molecules, respectively. As to the ligand, the force field parameters were obtained by Amber GAFF force field (GAFF),[39] and the charge parameters were determined by the restrained electrostatic potential (RESP) method[40] at the HF/6-31G(d) level with Gaussian 03 package.[41] The whole system was solvated into a ~82×83×75 Å cubic water box with 15 Å buffer solvents on each side, and neutralized by sodium ions. Then, multistep optimizations were performed to remove bad interatomic contacts. After that, the system was heated up gradually from 0 to 300 K by 50 ps, and another 50 ps MD simulation was carried out for further relaxation of the system. Finally, 10 ns MD simulation for equilibration was performed. All the MD simulations were employed under the NVT ensemble by using the periodic boundary condition with 12 Å cutoff distances for van der Waals and electrostatic interaction calculations. The Langevin method was utilized to control the temperature at 300 K. The bonds involved hydrogen atoms were constrained by the SHAKE scheme.[42,43] The last 3 ns trajectories were used for ligand-residue interaction decomposition and binding free

energy analysis by employing the MM-GBSA method.[44,45] All the molecular dynamic simulations are performed by AMBER12 software.[46]

Based on the MD simulations, two QM/MM models (A and B) were prepared by deleting the water molecules beyond 24 Å radiuses from the sulfur atom of Cys81. With respect to Model A, the Lys236 residue took a protonated state, while for Model B, the Lys236 residue was neutral (unprotonated). As shown in Figure 2, the QM subsystem includes the substrate, and the side chains of Ser80, His235, Thr11, Asp207 and Lys236. The QM region was treated by the DFT method with the B3LYP functional[47-50] and the 6-31G(d) basis set, which has been successfully used in many enzymatic catalytic systems.[51-60] As for the MM subsystem, Amber99SB force field[35-37] was employed as in the above classical MD simulations. The QM/MM boundary was handled by the improved pseudo-bond approach.[61-63] The spherical boundary condition was applied, and the atoms beyond 20 Å from the spherical center were fixed. The cutoff values of 18 and 12 Å were utilized for electrostatic and van der Waals interactions among MM atoms, respectively.

The QM/MM optimization was carried out to derive a minimum energy path with the reaction coordinate driving (RCD) method,[64] and for the electronic structures along the reaction path, the MM region was equilibrated for 500 ps by MM MD simulations. The snapshots abstracted from these MM MD simulations were served as the initial structures for subsequent QM(DFT)/MM MD simulations with the umbrella sampling. Each window was calculated for 20 ps with the time step of 1 fs by Beeman algorithm.[65] The Berendsen themostat66 was employed to maintain the system temperature at 300 K. Afterwards, free energy profiles were determined by the weighted histogram analysis method (WHAM)[67-70] with the probability distribution for each window of last 15 ps umbrella sampling. All ab initio QM/MM calculations were performed with the modified Q-Chem 4.0[71] and Tinker 4.2[72] programs.

To figure out the plausible entry and exit channels for acetone cyanohydrin and hydrocyanic acid transportation, the combined random acceleration molecular dynamics (RAMD) and MD simulations (RAMD-MD)[73,74] have been carried out with the NAMD 2.9 software.[75] The Amber99SB force field[35-37] and GAFF have been utilized for the protein and ligand, respectively. For the complex of enzyme with achiral acetone cyanohydrin, the initial structure was obtained from the MM MD simulations. For the complex of enzyme with hydrocyanic acid, the initial structure was set up by the sphere model of *Hb*HNL-product state (containing the acetone molecule and hydrocyanic acid) from QM/MM MD simulations by removing the acetone molecule.

In RAMD simulations, an additional force with random orientation is added to the center of mass (or other pre-defined point) of the ligand to identify the possible pathways in the binding pocket for the ligand fleeing away from the protein in a computationally feasible time. Within a certain period of time, when the ligand moves away beyond the threshold distance, the direction is maintained. Otherwise, a new random direction will be chosen. However, once the ligand escapes from the initial position, it may move towards the wrong channels because the random force is generally higher than the resistance of the protein. This issue can be avoided by using the combined RAMD-MD approach. Upon the escape of ligand from the initial position, the conventional MD simulations would be switched on and the equilibration sampling is recovered. Herein, the random accelerations of 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15 kcal $Å^{-1}$ g^{-1} are applied to the C1 atom of achiral acetone cyanohydrin and the C atom of HCN. 24 RAMD MD trajectories for each model have been explored. In what follows, one or two most favorable channels will be discussed in details by mapping out their free energy profiles along dynamic pathways with the umbrella sampling technique. A series of biasing harmonic potentials along the defined reaction coordinate have been tested. Based on the most appropriate biasing harmonic potential, 8 ns MD simulations at least are performed for each window. The free energy profiles are generated by WHAM[67-70] with the probability distribution for all windows.

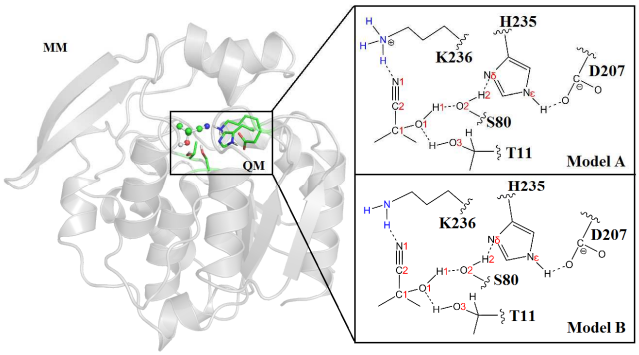

Fig. 2 QM/MM models for the *Hb*HNL enzyme with a protonated Lys236 (Model A) and an unprotonated Lys236 (Model B).

3 Results and Discussions

3.1 Catalytic mechanism by *Hb*HNL

The catalytic mechanism would be explored based on Model A at first, in which the Lys236 takes a protonated state as suggested from experiments for the real enzymatic environment. Herein, $d_{C1\text{-}C2}+d_{O1\text{-}H1}$ (RC1) and $d_{H2\text{-}N\delta}-d_{C2\text{-}H2}$ (RC2) are chosen as reaction coordinates. Whereas RC2 obviously refers to the proton transfer from Ser80 to His235, the choice of RC1 is intended to examine whether the C-C breaking within acetone cyanohydrin and the proton transfer from acetone cyanohydrin to Ser80 are stepwise or concerted. As we can see from the relative free energy profiles (see Figure 3a) and the structures of corresponding key states (see Figure 4), the whole reaction can be envisioned with quasi-two-step one; the first is the substrate deprotonation mediated by Ser80 and His235, which occurs approximately synchronously with the subsequent C1-C2 bond cleavage (**R→TS1→IM1→TS2→IM2**) owing to the presence of metastable proton-transfer

intermediate (**IM1**), and the second is the protonation of cyanide anion by His235 to yield HCN (**IM2→TS3→IM3**).

During the process from the reactant state **R** to **IM1**, the H1 atom of substrate is abstracted by the O2 atom of Ser80, which is concomitant with the proton transfer from Ser80 to the Nδ-position of His235. In other words, the two proton transfers are concerted. At the same time, C1-C2 bond is stretched slightly, and the single C1-O1 bond has the notable tendency to form a polarized $C1^{\delta+}$-$O1^{\delta-}$ bond (see the population analyses in Table S2). Thus, **IM2** mostly corresponds to the structure with protons transfers via substrate→Ser80→His235, but the C1-C2 bond retained. In the following reaction step from **IM1** to **IM2**, the double protons have completely transferred together with the C1-C2 bond cleavage. Simultaneously, the single C1-O1 bond is shortened from 1.36 ± 0.02 Å to 1.23 ± 0.01 Å and thus is converted to a typical carbonyl group. Moreover, the Mulliken charge distributions complied in Table S2 show that the negative charges on the C2 atom remarkably increase from 0.21 ± 0.09 to -0.53 ± 0.13 a.u., suggesting that the cyanide anion intermediate is formed.

At this first stage, the His235–Asp207 dyad accommodates one proton and its charged configuration is changed from His–Asp^- to His^+–Asp^-. Clearly, the hydrogen bond between them becomes stronger due to the favorable electrostatic attraction, implying that the two residues play a crucial role in stabilizing the nascent intermediate in catalysis. Here they serve as a general base to accept the proton from Ser80 to facilitate the deprotonation of the substrate by Ser80. Most noteworthy is the distance between the -NH_3^+ group of Lys236 and the N1 atom, which is significantly shortened from 3.04 ± 0.33 to 1.75 ± 0.13 Å. Such strong hydrogen bond interactions can stabilize the labile cyanide anion, showing that the protonated Lys236 can assist the cleavage of C-C bond through strong electrostatic (field) interactions as the driving force. The energy barriers double-proton transfer and the C-C bond cleavage are 6.7 and 6.6 kcal/mol, respectively, suggesting that both processes are extremely fast and experimentally difficult to differentiate. This may be the reason why experiments support the hypothesis that the deprotonation and the C-C bond cleavage of substrate are concerted. Overall, for the first stage from the reactant state to the stable **IM2**, the free energy change ΔG is 4.8 kcal/mol.

At the second stage, the cyanide anion gradually approaches His235 to form cyanic acid. We note that the charge on the C2 atom remarkably changes from -0.53 ± 0.13 to 0.35 ± 0.09 a.u. while the charge on the Nδ atom goes towards the opposite direction, lending support to the proton transfer from His235 to cyanide anion. Meanwhile, the distance between the -NH_3^+ group of Lys236 and CN^- is stretched from 1.75 ± 0.13 to 3.05 ± 0.59 Å, and the hydrogen bond between Ser80 and His235 dissolves. The hydrogen bond distance between His235 and Asp207 slightly increases from 1.62 ± 0.12 to 1.83 ± 0.12 Å. All these structural changes point to the formation of HCN at this stage.

What is noteworthy is that the relatively high free energy barrier (13.1 kcal/mol) for the HCN formation, indicating that this step in the protein environment is rate-determining and crucial for the overall enzymatic catalysis. This finding is quite different from the previous study without considering the protein environment and its dynamics,[30] which claimed the cleavage of the C-C bond as the rate-limiting step. The comparison of the equilibrium configurations of **IM2** and **IM3** suggests that the loss of strong ion-pair bonding interaction of $(CN)^-$ ••• $(NH_3)^+$–Lys236 and the N-H bond cleavage in the

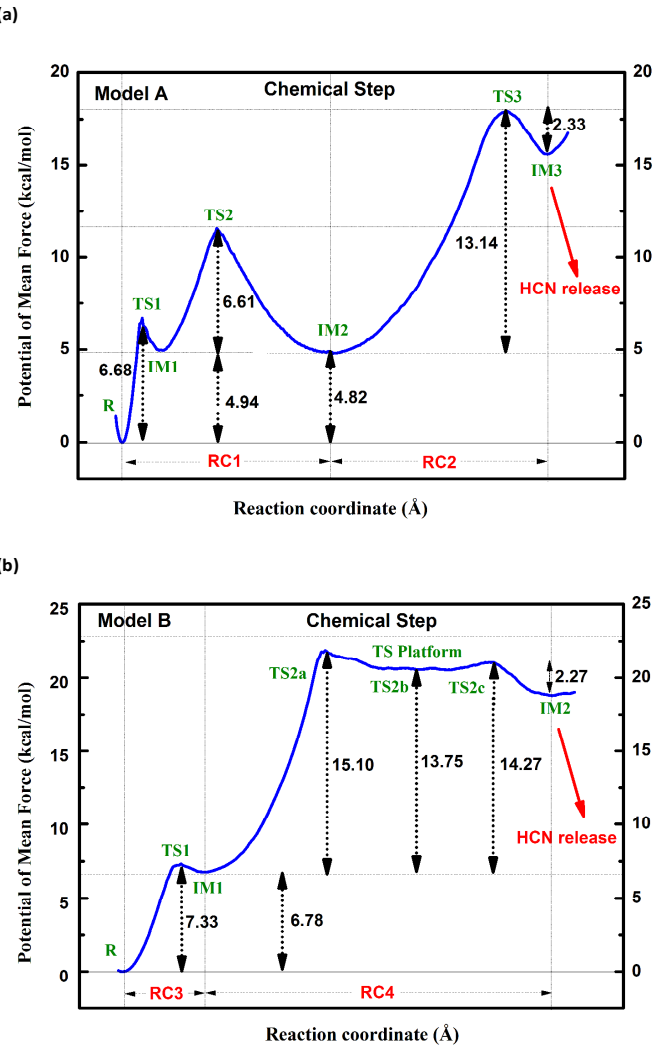

Fig. 3 Predicted free energy profiles of the whole catalytic reaction by ab initio QM/MM MD simulations for Model A **(a)** and Model B **(b)**

process of the HCN formation cannot be sufficiently compensated by the H-C bond formation, resulting in a relatively high energy barrier and a metastable intermediate configuration **IM3**. Furthermore, the remarkable configuration evolution and the environmental effect of proteins during the proton abstraction of His235 by CN^- are also responsible for these thermodynamic and dynamical properties to a certain extent.

Seeing that the rate-determining step has a barrier of 13.1 kcal/mol, and two earlier steps have the barriers of 6.7 and 6.6 kcal/mol, we anticipate that the complex reaction barrier would

be very close to the value of 16.1 kcal/mol estimated by the transition state theory from the experimental data (k_{cat} = 10 s^{-1} at 303 K. Though the kinetic parameters were measured with partially purified enzyme, they can still provide useful information for the kinetic property of the enzyme).[76] It should be pointed out that the **IM3** intermediate is very unstable compared to **IM2** and thus the forward catalytic reaction (rate-determining step) seems to be quite unfavorable. How does then the reaction take place? Our further simulations (see following) demonstrate that the facile release of the newly-formed HCN plays a key role in facilitating this forward reaction, and the remarkable energy release in the substrate binding may further compensate this energy requirement (vide infra).

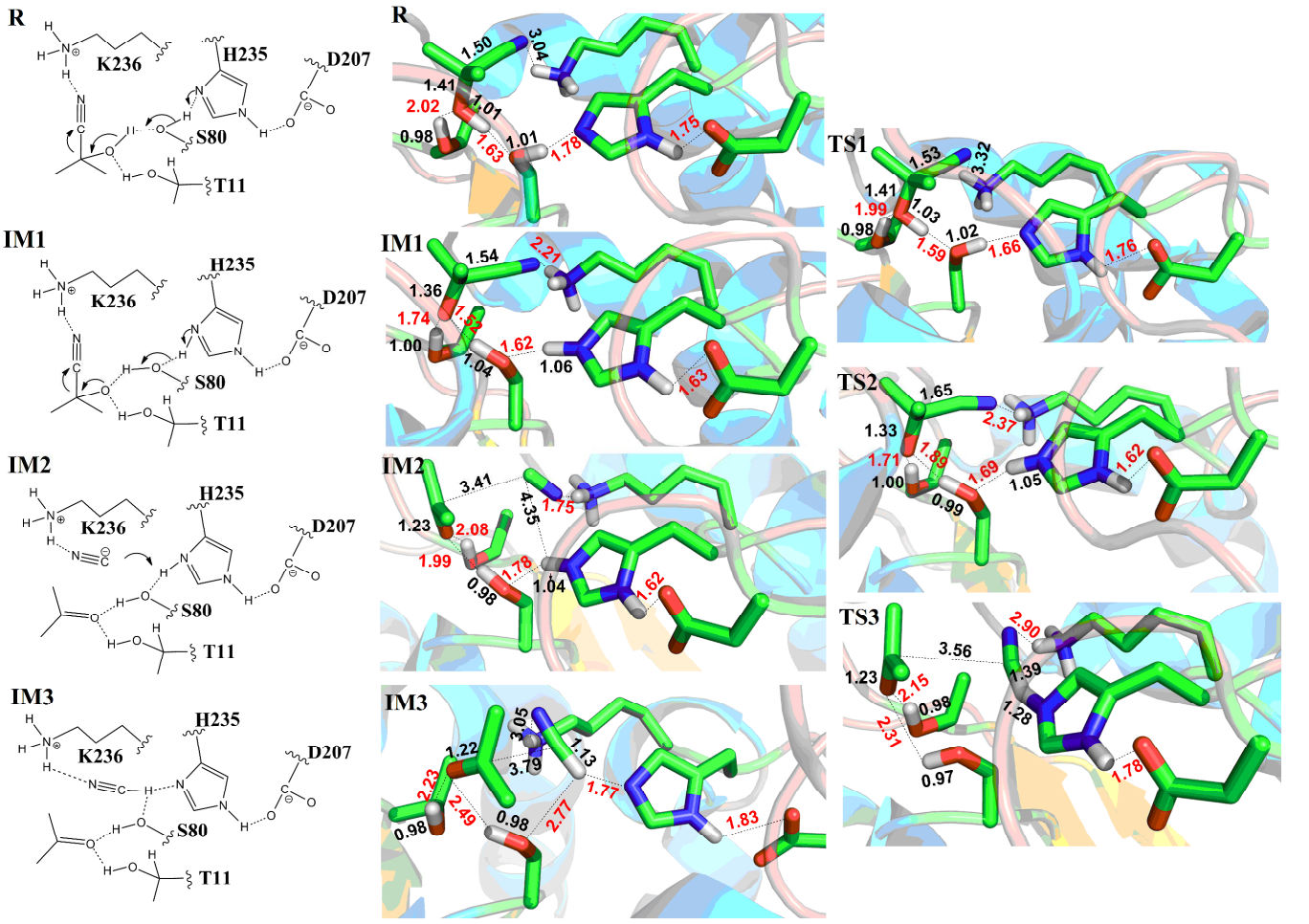

Fig. 4 The structures of the reactant, transition states, intermediates involved in the catalytic reaction for Model A (the average distances for selected bonds in Å) from the QM/MM-MD sampling.

3.2 Effect of the protonation state of Lys236 on catalysis

To evaluate the role of Lys236 with different protonation states in catalysis, we performed computational studies similar to the above to probe the possible catalytic mechanism based on Model B where Lys236 is unprotonated and neutral in *Hb*HNL. In contrast with Model A, the energy scanning calculations along RC1 (d_{C1-C2}+d_{O1-H1}) were unsuccessful with very high energy barriers. Alternatively, new reaction coordinates of d_{O1-H1}-d_{O2-H1} (RC3) and d_{H2-N}-d_{O2-H2} (RC4) were defined to follow the proton transfers. Figure 3b shows the energy profile with the Model B, where the largest energy barrier is 15.1 kcal/mol, higher than the rate-limiting step with Model A (13.1 kcal/mol). Present results suggest that, while a neutral Lys236 does slow down the enzyme-catalyzed decomposition of cyanohydrins, the process is still efficient enough as the overall complex barrier is less than 22.4 kcal/mol. What we see from the comparison between Figures 3a and 3b is the dramatic difference in the reaction mechanisms.

As Figure 3b and Figure S1 in Supporting Information show, the reaction mechanism with Model B essentially comprises only two steps; one is the deprotonation of substrate through the double-proton transfer mediated by Ser80 and His235 (**R→TS1→IM1**), and the other is the concurrent carbon-carbon bond cleavage and the HCN formation (**IM1→TS platform→IM2**). This mechanism is totally different from Model A. Here no C-C bond cleavage

intermediate with the CN^- moiety (**IM2** in Figure 3a) was identified and a quite flat plateau of transition state appears in the relative free energy profile. Therefore, the protonation state of Lys236 controls the catalytic mechanism.

In Model A, the $-NH_3^+$ group of the protonated Lys236 significantly stabilizes the nascent CN^- through strong ion-pair bonding interactions and thereby can drive the C-C bond cleavage, while there is no such strong electrostatic interactions in Model B with the unprotonated Lys236. Without the electrostatic field from the protonated Lys236, the intermediate configuration with a CN^- species is unlikely formed in Model B. The structures of possible reaction configurations are presented in Figures S1 and S2 in †Electronic Supplementary Information (ESI)

We note that here the C-C bond cleavage and the HCN formation require a long journey in the active domain, which may account for occurrence of the flat transition-state plateau in Model B. As a consequence, in Model B transition-state configurations may be very different during our sampling simulations, though they have comparable energies. Here two kinds of reaction tendencies (see Figure S2 in †ESI) have been discussed. One may proceed towards the reactant-analogue state while the other shows a tendency to the final HCN formation, where the hydrogen atom in the newly-formed HCN may come from either of the two transferred protons.

3.3 The role of key residues in catalytic process

Experimentally,[21] it has been assumed that the residues Ser80-His235-Asp207 form a catalytic triad as a general base to take part in the whole reaction. Among them, Ser80 and His235 mediate the double-proton transfer directly, and Asp207 retains the hydrogen bonding with the Nδ atom of His235 to keep the latter at a proper position towards the hydrogen abstraction and strengthen the basicity of Ser80-His235. Apart from the catalytic triad, other important residues include Lys236 and Thr11. As mentioned above, the protonation state of Lys236 plays a role in the catalytic process. The residue Thr11, however, is involved in the hydrogen bonding network around the substrate.

To further elucidate the roles of these key residues, QM(B3LYP)/MM calculations have been performed on the Asp207Ala, Lys236Ala, and Thr11Ala mutant systems. Based on the QM/MM optimized structures, QM/MM energy scans along the reaction coordinate of RC1 were performed. As we can see from Figure 5, the barrier for the carbon-carbon bond cleavage in the Asp207Ala mutant system is higher than that of the wild system by 4.4 kcal/mol. Quite interestingly, the T11A mutant system exhibits a much lower barrier of ~10 kcal/mol, compared to the wild *Hb*HNL enzyme. However, the T11A mutation may destroy the hydrogen-bonding network around the substrate and makes the substrate less likely for the initial proton transfer, resulting in a loss of the enzymatic activity as observed experimentally.[21] As for the Lys236Ala system, the C-C bond breaking cannot be achieved as it needs to couple other bond cleavage and thus experiences much high barriers which make the reaction unrealistic at ambient condition. In conclusion, mutations of the key residues of Asp207, Lys236, and Thr11 may have a remarkable effect on the C1-C2 bond cleavage.

The QM/MM-optimized configurations indicate that the hydrogen bond distances between N1 of substrate and $-NH_3^+$ of Lys236 are 2.49, 2.74, and 1.99 Å for wild-type *Hb*HNL, mutant systems Asp207Ala and Thr11Ala, respectively. These results show that the lack of hydrogen bonds between the substrate and Thr11 may remarkably enhance the hydrogen bond between the substrate and (NH_3^+)–Lys236, which will facilitate the C-CN bond cleavage through strong electrostatic interactions, although this near-attack conformer (NAC) is unlikely accessible in such a situation.

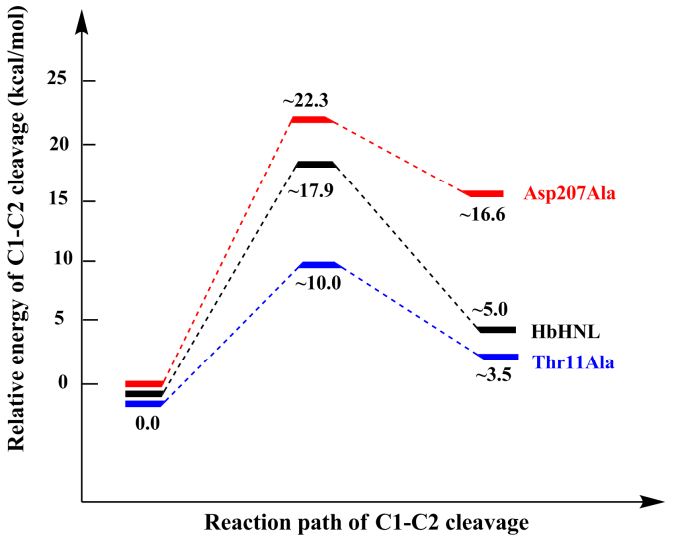

Fig. 5 The predicted relative energies for the carbon-carbon cleavage in *Hb*HNL, ASP207ALA, and THR11ALA systems by the QM/MM scan.

3.4 The role of key residues in substrate binding.

In previous studies, Gruber et al.[21] proposed that the residues of Lys236, Ser80, and Thr11 play an important role in the substrate binding. To get a deeper insight into the individual contributions of these three residues on substrate binding, the binding free energies were calculated for the Lys236Ala, Ser80Ala, Thr11Ala, and original *Hb*HNL systems. Major results were summarized in Figure 6a and Table S3. We note that the van der Waals, electrostatic, non-polar solvation interactions are the major driving forces for substrate binding, while the polar solvation energy and the entropy effect are unfavorable (i.e., positive) terms. Although strong electrostatic interactions approximately counteract the polar solvation energy here, they should be important for the C-CN bond activation and breaking. In comparison with the wild *Hb*HNL system, the binding free energies for the Lys236Ala, Ser80Ala, Thr11Ala mutant systems decrease by 2.96, 4.74, and 1.23 kcal/mol, respectively, which mainly arise from the

electrostatic interactions. On the basis of per-residue type, the free energy decomposition shows that the residues Ile12, Cys81, Leu157, and His235 in the active site also have notable contributions to the substrate binding, as shown in Figure 6b, in which the van der Waals and non-polar solvation energies play a primary role in binding the substrate.

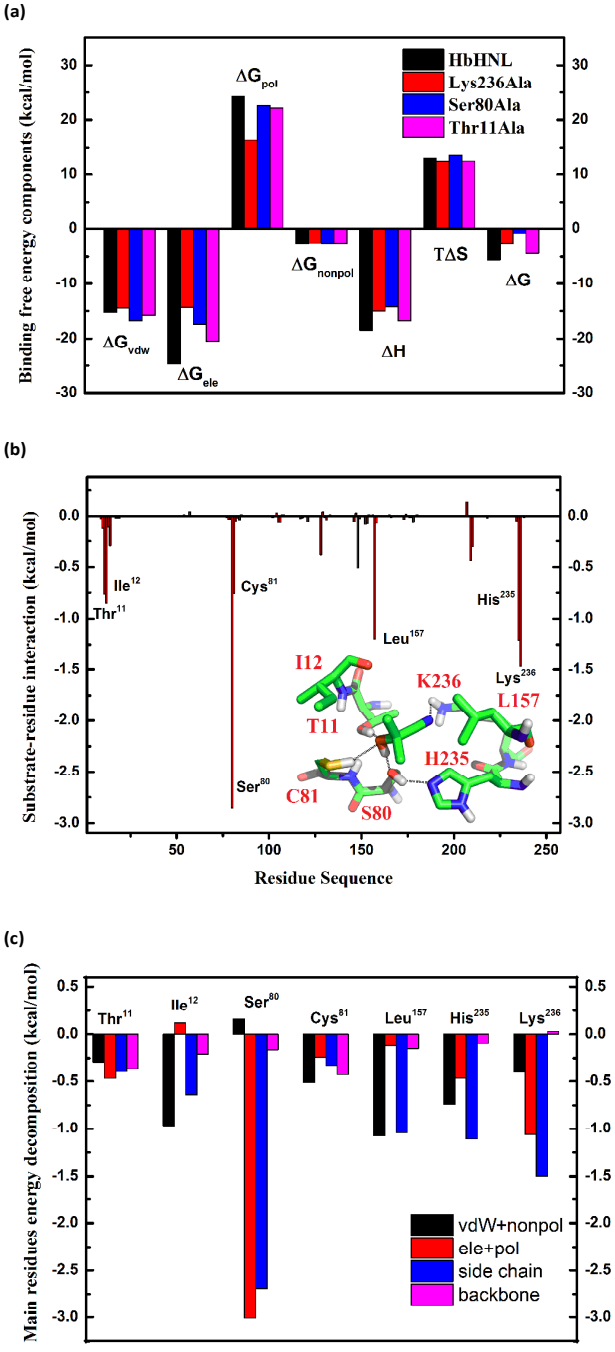

Fig. 6 (a) The binding free energy components for *Hb*HNL, Lys236Ala, Ser80Ala, and Thr11Ala systems. **(b)** The free energy decomposition on the basis of per-residue type along with the binding mode. **(c)** The energy decomposition for the key residues.

3.5 The acetone cyanohydrine access to the active site

Most of studies on enzymatic catalysis focus on chemical reaction processes. However, it is equally important to investigate the delivery processes of substrate and product particularly when the active site of an enzyme is buried inside and both substrate and product need go through a channel. To have an insight into the thermodynamic and dynamical properties of *Hb*HNL to accommodate the substrate, we have explored the possible pathways for the substrate delivery to the active domain. However, in view of the computational efficiency and convenience, its reverse process, i.e., the release of acetone cyanohydrine, in the complex of enzyme with acetone cyanohydrin, has actually been used to characterize the substrate access to the active site of *Hb*HNL.

Figure 7 shows four possible pathways (Pa-1, Pa-2, Pb, and Pc) for the acetone cyanohydrin release identified by RAMD-MD simulations, which are defined as Pa-1 (between helix D and Trp128), Pa-2 (between Trp128 and helix D1'), Pb (between helix D1' and helix E), and Pc (between helix D2' and D3'). 24 RAMD-MD trajectories are summarized in Table 1. We note that 17 trajectories among them follow the Pa-1 channel and Pa-1 is thus the predominant pathway for the substrate release (or the substrate docking). The possibilities of the other three channels are almost the same. For Pa-1, according to the escaping direction, the distance between the Cδ atom of Lys236 and the C1 atom of acetone cyanohydrin is chosen as the release coordinate (named as RC5, see Figure S3a). MD simulations along this coordinate (RC5) from 6.5 to 16.0 Å with biasing harmonic potential of 30 kcal/mol have been performed. The potential of the mean force (PMF) profile is depicted in Figure 8a. Clearly, the relative free energy profiles for different sampling time durations are very similar, showing reliable convergence of MD simulations for the PMF profiles. The predicted relative free energies in Figure 8a reveal that the access of substrate to the active site is quite favorable thermodynamically with a negligible barrier (~ 1 kcal/mol), and the binding free energies of ΔG are about -15 kcal/mol. Such remarkable exothermicity may drive the subsequent catalytic process.

Table 1. Statistics of 24 trajectories for the substrate delivery and the HCN release*

substrate	share	possibility	HCN	share	Possibility
Pa-1	17	70.8%	Pa	18	75.0%
Pa-2	3	12.5%	Pb	2	8.0%
Pb	2	8.3%	Pc	4	17.0%
Pc	2	8.3%			

* Random accelerations of 0.50, 0.45, 0.40, 0.35, 0.30, 0.25, 0.20, 0.15 kcal $Å^{-1}$ g^{-1} are applied to the C1 atom of achiral acetone cyanohydrin and the C atom of hydrocyanic acid, respectively.

The PMF profiles for the channel Pa-2 have been also characterized by RAMD-MD simulations (see Figure S4). The distance between the CB atom of Thr11 and the C1 atom of substrate has been taken as the release coordinate (named as RC6, see Figure S3a). MD simulations suggest that the free

energies of release ΔG are about 20 kcal/mol, thus it should be facile for the transport of substrate to the active site through the channel Pa-2 energetically. For the reverse process, i.e., the substrate release, such relatively higher energy requirement may result in the occurrence of only few trajectories for the channel Pa-2 during MD simulations. Presumably, the steric effects along the channels are also important for the substrate delivery. In what follows, we will discuss the substrate transportation along the most probable channel Pa-1 in detail, in order to study the dynamical features and key residue effects on the substrate movement.

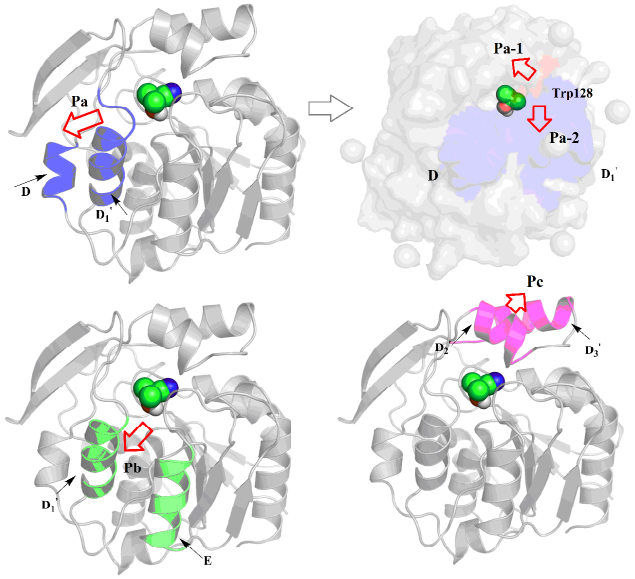

Fig. 7 The possible channels for acetone cyanohydrin delivery predicted by RAMD-MD simulations. The acetone cyanohydrine in the active site is shown as sphere, and the red part of the surface in the upper right figure refers to Trp128.

We approximately classify the process of acetone cyanohydrine release into four stages based on the conformational evolutions along the reaction coordinate RC5 (Figure 8a). At the first stage (7.0 Å ⩽ RC < 9.0 Å), the channel is totally closed, and Trp128 blocks the release of acetone cyanohydrine. Thr11 and Ser80 may stabilize the acetone cyanohydrin at the original active site via the hydrogen bonding interactions. At the second stage (9.0 Å ⩽ RC ⩽ 11.0 Å), the residue Trp128 turns on one side gradually as a switch, and the "door" (indole ring) swings between an opening state and a closing state. The acetone cyanohydrine tends to move away from the original active site, subsequently inducing a conformational change of helix-D from an α-helix to a loop structure. At the third stage (11.0 Å < RC ⩽ 14.0 Å), the system overcomes a barrier of ~16.0 kcal/mol. Trp128 no longer blocks the ligand from moving out of the channel, as the pocket almost keeps open, and the helix-D recovers to its original α-helix structure gradually. Moreover, residues Tyr133, Trp128, and Gln180 provide the driving force for the release of cyanohydrin through the direct and stabilizing hydrogen bonding, the σ-π interaction, and the hydrogen-bond network

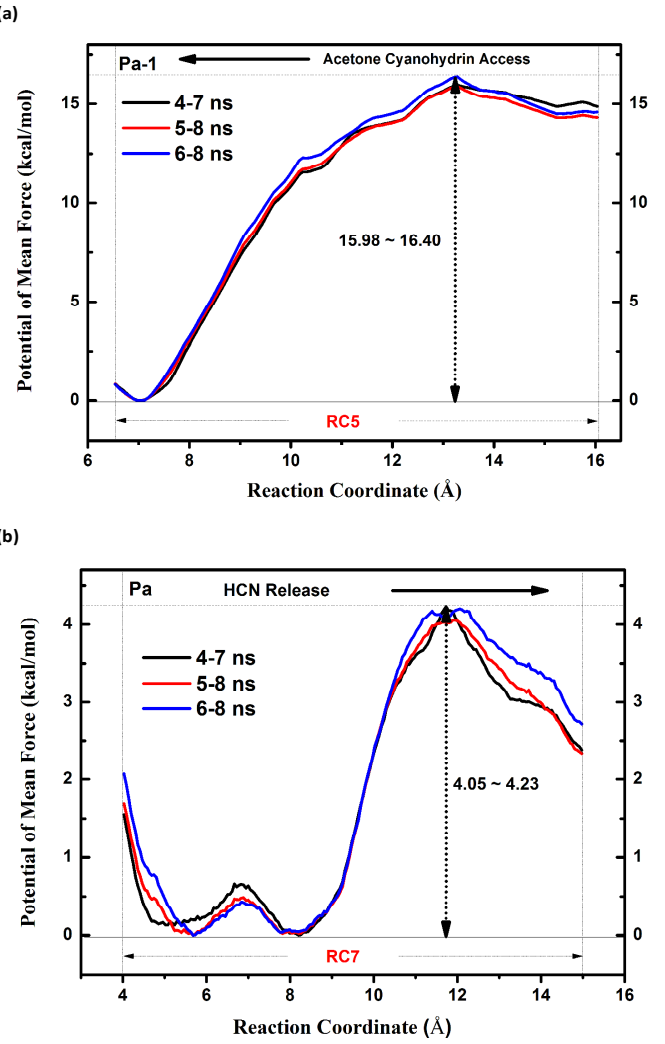

Fig. 8 (a). Free energy profiles for the release of acetone cyanohydrin along the RC5 (The distance between the Cδ atom of Lys236 and the C1 atom of acetone cyanohydrin). **(b)**. Free energy profiles for the release of HCN along the RC7 (The distance between the CZ atom of Tyr158 and the C atom of HCN).

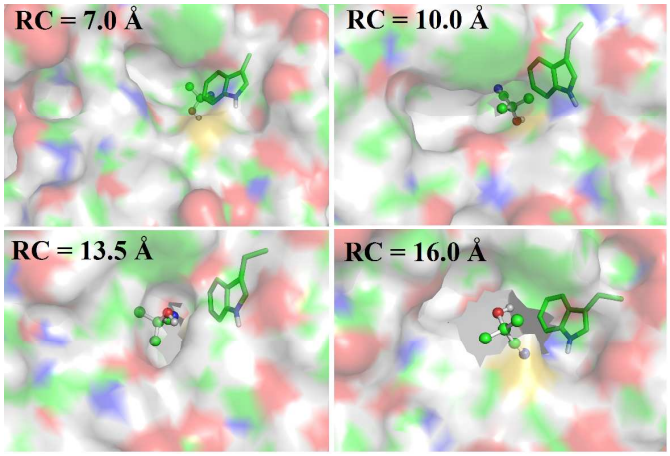

Fig. 9 (a). Surfaces of the binding pocket from different windows along the RC5 (The distance between the Cδ atom of Lys236 and the C1 atom of ligand) for the release of acetone cyanohydrin.

with adjacent water molecules, respectively. At the last stage (RC > 14.0 Å), the acetone cyanohydrin leaves from the protein completely, and the Trp128 returns to the original "closed" state once again. The surfaces of binding pocket from different windows are shown in Figure 9a, and the secondary structures and key residues relative to the release process are shown in Figure 9b. The flexibility of Trp128 agrees well with the previous experimental study,[25] and similar dynamical behaviours have been found in other biosystems.[77]

It is worth mentioning that the HNL from *Manihot esculenta* (*Me*HNL) has a similar structure to *Hb*HNL. In *Me*HNL, Trp128 was mutated to Ala, resulting in an improved conversion for larger aromatic cyanohydrines.[78] In addition, the residue also influences the stereoselectivity.[79] Despite these testified roles of Trp128 for the catalytic process in *Me*HNL, they are still required to be elucidated for *Hb*HNL. Therefore, we mutated Trp128 to alanine in silico to quantify its function in the substrate release (or binding) process. We obtained the free energy profile for the Trp128Ala mutant system by the umbrella sampling approach (see Figure S5). The energy required for the substrate release in the Trp128Ala mutant is about 8 kcal/mol, only half of the value (~16 kcal/mol) in the wild system. Clearly, the steric effect from the large indole ring and the conformational dynamics of Trp128 may dominate the substrate delivery, though other hydrogen-bond networks and hydrophobic residues are also involved in this process to some extent.

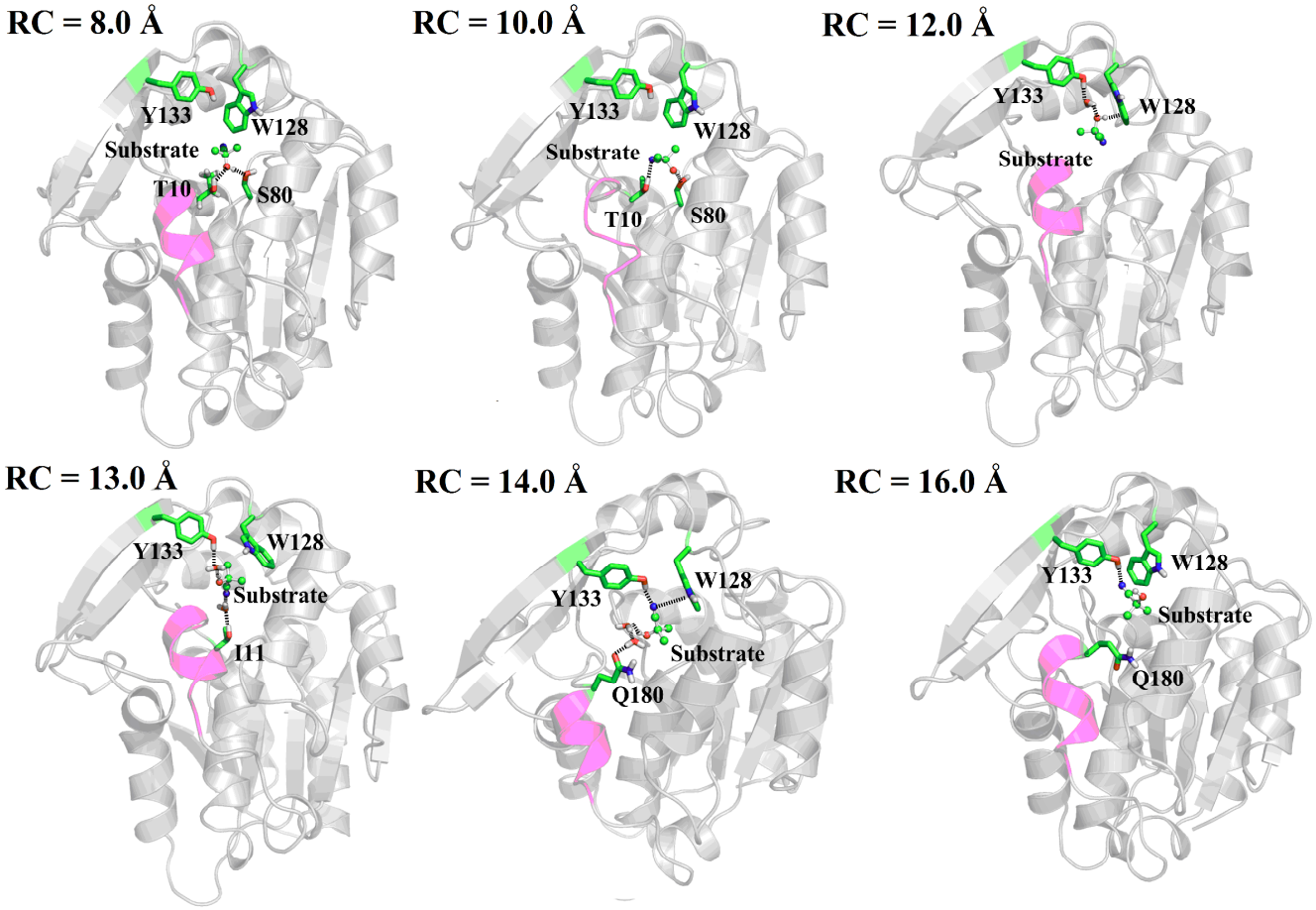

Fig. 9 (b). The key residues residues in the active site of *Hb*HNL from different windows along the RC5 (The distance between the Cδ atom of Lys236 and the C1 atom of ligand) for the release of acetone cyanohydrin. The purplish red part represents helix-D, and obviously its conformer changes along with the flip of Trp128. The Trp128 is shown in the stick presentation and the reaction path corresponds to Pa-1 in Fig. 7.

3.6 Release of HCN

Needless to say, the release of HCN is significant and the whole purpose of HNLs for the plant defense system and biosynthesis of asparagine. From the perspective of kinetics, the removal of HCN from the active site may also drive the catalytic reaction to move along the forward reaction. Thus, by employing the RAMD-MD method, we similarly explored its possible release channels. The distance between the CZ atom of Tyr128 and the C atom of HCN was chosen as the reaction coordinate (named as RC7, see Figure S3b), and 23 windows were adopted from 4.0 Å to 15.0 Å. Our simulations revealed that HCN may

release through three pathways, including Pa (between helix D and Trp128), Pb (between helix D_1' and helix E), and Pc (between helix D_2' and D_3'), with possibilities of 75%, 8%, and 17%, respectively (see Table 1). Based on the umbrella sampling method, the PMF profiles for Pa have been obtained and shown in Figure 8b. The predicted free energy barrier is only ~4.0 kcal/mol, and thus HCN is easy to escape from the protein. This is understandable due to its much small molecular size. The facile release of HCN can drive the catalytic reaction to proceed efficiently, though the product state is higher in energy than the reactant state in the lytic reaction of acetone cyanohydrin. As mentioned above, the remarkable energy release in the substrate binding may compensate the energy requirement in the catalytic step.

The key residues and secondary structures from different windows are shown in Figure S6. In the beginning period (4.0 Å ≤ RC ≤ 8.0 Å), HCN moves dynamically among the residues His14, Ile12, Lys236, Thr11, Ser80, and His235, with a negligible barrier of about 0.5 kcal/mol. Afterwards (8.0 Å < RC ≤ 12.0 Å), the energy increases gradually, probably due to the flip of Trp128. We can see from Figure S6, HCN sways on the both sides of the residue, which may make the barrier of release rise. At last (RC > 12.0 Å), HCN moves away from the enzyme completely. At this point, Trp128 returns to the closed state as well, and its gating mechanism plays a role in the HCN release, but overall its steric effect is limited due to the small size of HCN.

4 Conclusions

On the basis of extensive QM/MM MD and RAMD MD simulations, the catalytic mechanism of *Hb*HNL and its substrate delivery and product (HCN) release have been fully explored. Approximately the chemical step in the low-energy enzymatic process undergoes two stages. The first stage includes the substrate deprotonation by Ser80 and His235, and the C-CN bond cleavage. These two processes are fast and often considered together. The second stage is the cyanic acid formation which is the rate-determining step with a barrier of 13.1 kcal/mol. Combining these three chemical steps together, we predict that the complex reaction barrier would be pretty much close to the experimental estimate 16.1 kcal/mol. Calculations show that the catalytic mechanism strongly depends on the protonation state of Lys236, as the protonated Lys236 can facilitate the C-CN bond breaking and stabilize the nascent CN^-, resulting in relatively favorable mechanism dynamically. With a protonated Lys236, the rate-determining step appears at the HCN formation, owing to the loss of strong ion-pair bonding interactions of $(CN)^-$ ••• $(NH_3)^+$ – Lys236, the N-H bond cleavage, and remarkable conformational change. In contrast, with a neutral Lys236, the rate-limiting step is the concurrent C-CN bond cleavage and the HCN formation. Energetically, however, hydroxynitrile lyases with either protonated or neutral Lys236 can efficiently catalyze the decomposition of cyanohydrins.

Based on RAMD MD simulations, plausible channels for the substrate delivery and the HCN release have been explored. In the most probable channels, the key residue Trp128 shows a dynamical gating mechanism, which may dominate the transportation of substrate and HCN. The binding of acetone cyanohydrin is predicted to be exothermic remarkably, while the release of HCN requires an energy of only 4.0 kcal/mol. The roles of residues in catalysis and delivery of substrate and HCN are also analyzed based on the MD simulations and free energy decomposition. The mutations of key residues can alter the enzymatic activity remarkably. In comparison with the achiral acetone cyanohydrin, a chiral substrate such as mandelonitile most probably behaves differently. The aromatic residue may be bound to the hydrophobic pocket and thus has a preferred orientation. However, as was deduced from the X-ray structural analysis, the *Hb*HNL complex with symmetric substrates and asymmetric substrates probably follows the similar catalytic mechanisms, binding modes, and release channels.[25] Thus we expect that a thorough understanding of the whole enzymatic catalysis as presented in this work can provide meaningful clues for the biocatalytic retrosynthesis of chiral compounds such as mandelonitile and so on.

Acknowledgements

This work was supported by the Ministry of Science and Technology (2011CB808504 and 2012CB214900), the National Science Foundation of China (21133007 and 21373164), and the US National Science Foundation under the grant CHE-1055310 and Western Michigan University (FRACAA). We thank the National Supercomputing Center in Shenzhen and Guangzhou for providing the computational resource. We also thank Prof. Ruibo Wu at Sun Yat-sen University, Dr. Shenglong Wang at New York University and Dr. Vlad Cojocaru at Max Planck Institute for Molecular Biomedicine for their kind help.

Notes and references

[a] State Key Laboratory of Physical Chemistry of Solid Surfaces and Fujian Provincial Key Laboratory of Theoretical and Computational Chemistry, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, P. R. China

[b] School of Pharmaceutical Sciences, Sun Yat-sen University, Guangzhou 510006, P. R. China

[c] Department of Chemistry, Western Michigan University, Kalamazoo, Michigan 49008, United States

E-mail: *zxcao@xmu.edu.cn

† Electronic Supplementary Information (ESI) available: The *Hb*HNLs crystal structures, Mulliken charge populations of selected atoms that relevant to reaction, The binding free energy and the difference of binding free energy for *Hb*HNL and mutant systems, The structures of reactant, transition state, intermedia for model B along the reaction path, wherein the HCN formed by the deprotonation of His235, The different structures of transition state during the "platform" for model B, The defined reaction coordinate for release of acetone cyanohydrin and HCN, Free energy profiles of along the RC6 for release of acetone cyanohydrin and RC5 for release of acetone cyanohydrin in Trp128Ala system, The

key residues and the second structures from different windows along the RC7 for release of HCN. See DOI: 10.1039/b000000x/

1 E. E. Conn, In *The Biochemistry of Plants: A Comprehensive Treatise*, ed. P. K. Stumpf and E. E. Conn, Academic Press, New York, 1981, vol. 7, pp. 479-500.
2 A. Hickel, M. Hasslacher and H. Griengl, *Physiol. Plant.*, 1996, **98**, 891-898.
3 H. Wajant and F. Effenberger, *Biol. Chem.*, 1996, **377**, 611-617.
4 R. Lieberei, D. Selmar and B. Biehl, *Plant Syst. Evol.*, 1985, **150**, 49-63.
5 D. Selmar, *ACS Symp. Ser.*, 1993, **13**, 191-204.
6 M. Gruber-Khadjawi, T. Purkarthofer, W. Skranc and H. Griengl, *Adv. Synth. Catal.*, 2007, **349**, 1445-1450.
7 J. von Langermann, J. -K. Guterl, M. Pohl, H. Wajant and U. Kragl, *Bioproc. Biosyst. Eng.*, 2008, **31**, 155-161.
8 M. Paravidino, M. J. Sorgedrager, V. A. Romano and U. Hanefeld, *Chem. Eur. J.*, 2010, **16**, 7596-7604.
9 S. E. Milner, T. S. Moody and A. R. Maguire, *Eur. J. Org. Chem.*, 2012, **2012**, 3059-3067.
10 I. Oroz-Guinea and E. García-Junceda, *Curr. Opin. Chem. Biol.*, 2013, **17**, 236-249.
11 M. Müller, *Chem. Bio. Eng. Reviews*, 2014, **1**, 14-26.
12 J. Diebler, J. von Langermann, A. Mell, M. Hein, P. Langer and U. Kragl, *Chem. Cat. Chem.*, 2014, **6**, 987-991.
13 M. Fechter and H. Griengl, In *Enzyme catalysis in organic synthesis*, 2nd ed., ed. K. Drauz and H. Waldmann, Wiley-VCH Verlag GmbH, Weinheim, 2002; vol. II, pp. 974-989.
14 T. Purkarthofer, W. Skranc, C. Schuster and H. Griengl, *Appl. Microbiol. Biotechnol.*, 2007, **76**, 309-320.
15 J. Holt and U. Hanefeld, *Curr. Org. Synth.*, 2009, **6**, 15-37.
16 N. J. Turner and E. O'Reilly, *Nat. Chem. Biol.*, 2013, **9**, 285-288.
17 U. Hanefeld, *Chem. Soc. Rev.*, 2013, **42**, 6308-6321.
18 M. Sharma, N. N. Sharma and T. C. Bhalla, *Enzyme Microb. Tech.*, 2005, **37**, 279-294.
19 M. Dadashipour and Y. Asano, *ACS Catal.*, 2011, **1**, 1121-1149.
20 I. Dreveny, A. S. Andryushkova, A. Glieder, K. Gruber and C. Kratky, *Biochemistry*, 2009, **48**, 3370-3377.
21 K. Gruber, G. Gartler, B. Krammer, H. Schwab and C. Kratky, *J. Biol. Chem.*, 2004, **279**, 20501-20510.
22 N. Andexer, N. Staunig, T. Eggert, C. Kratky, M. Pohl and K. Gruber, *Chem. Bio. Chem.*, 2012, **13**, 1932-1939.
23 K. Gruber, M. Gugganig, U. G. Wagner and C. Kratky, *Biol. Chem.*, 1999, **380**, 993-1000.
24 U. G. Wagner, M. Hasslacher, H. Griengl, H. Schwab and C. Kratky, *Structure*, 1996, **4**, 811-822.
25 G. Gartler, C. Kratky and K. Gruber, *J. Biotechnol.*, 2007, **129**, 87-97.
26 C. Mueller-Dieckmann, S. Panjikar, A. Schmidt, S. Mueller, J. Kuper, A. Geerlof, M. Wilmanns, R. K. Singh, P. A. Tucker and M. S. Weiss, *Acta. Crystallogr., Sect. D: Biol. Crystallogr.*, 2007, **63**, 366-380.
27 J. Zuegg, K. Gruber1, M. Gugganig, U. G. Wagner and C. Kratky, *Protein Sci.*, 1999, **8**, 1990-2000.
28 A. Schmidt, K. Gruber, C. Kratky and V. S. Lamzin, *J. Biol. Chem.*, 2008, **283**, 21827-21836.
29 D. L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren and A. Goldman, *Protein Eng.*, 1992, **5**, 197-211.
30 F. Cui, X. Pan and J. Liu, *J. Phys. Chem. B*, 2010, **114**, 9622-9628.
31 H. Li, A. D Robertson and J. H. Jensen, *Proteins*, 2005, **61**, 704-721.
32 D. C. Bas, D. M. Rogers and J. H. Jensen, *Proteins*, 2008, **73**, 765-783.
33 H. M. Olsson, C. R. Søndergard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525-537.
34 C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 2284-2295.
35 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179-5197.
36 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712-725.
37 J. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049-1074.
38 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
39 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
40 D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel and P. Kollman, *Comput. Phys. Commun.*, 1995, **91**, 1-41.
41 J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision A.1; Gaussian Inc.: Wallingford CT*, 2009.
42 J. -P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comp. Phys.*, 1977, **23**, 327-341.
43 S. Miyamoto and P. A. Kollman, *J. Comput. Phys.*, 1992, **13**, 952-962.
44 A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889-897.
45 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H.

Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 12*, *University of California: San Francisco, CA,* 2012.

46 T. Hou, W. Zhang, D. A. Case and W. Wang, *J. Mol. Biol.*, 2008, **376**, 1201-1214.

47 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.

48 C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.

49 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200-1211.

50 J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623-11627.

51 D. W. Rooklin, M. Lu and Y. Zhang, *J. Am. Chem. Soc.*, 2012, **134**, 15595-15603.

52 Z. Ke, G. K. Smith, Y. Zhang and H. Guo, *J. Am. Chem. Soc.*, 2011, **133**, 11103-11105.

53 R. Wu, Z. Lu, Z. Cao and Y. Zhang, *J. Am. Chem. Soc.*, 2011, **133**, 6110-6113.

54 R. Wu, S. Wang, N. Zhou, Z. Cao and Y. Zhang, *J. Am. Chem. Soc.*, 2010, **132**, 9471-9479.

55 R. Wu, H. Xie, Z. Cao and Y. Mo, *J. Am. Chem. Soc.*, 2008, **130**, 7022-7031.

56 Y. Zhou and Y. Zhang, *Chem. Comm.*, 2011, **47**, 1577-1579.

57 Y. Shi, Y. Zhou, S. Wang and Y. Zhang, *J. Phys. Chem. Lett.*, 2013, **4**, 491-495.

58 R. Wu, W. Gong, T. Liu, Y. Zhang and Z. Cao, *J. Phys. Chem. B*, 2012, **116**, 1984-1991.

59 Z. Ke, H. Guo, D. Xie, S. Wang and Y. Zhang, *J. Phys. Chem. B*, 2011, **115**, 3725-3733.

60 N. Chen, J. Zhou, J. Li, J. Xu and R. Wu, *J. Chem. Theory Comput.*, 2014, **10**, 1109-1120.

61 Y. Zhang, *J. Chem. Phys.*, 2005, **122**, 024114.

62 Y. Zhang, *Theor. Chem. Acc.*, 2006, **116**, 43-50.

63 Y. Zhang, T. S. Lee and W. Yang, *J. Chem. Phys.*, 1999, **110**, 46-54.

64 Y. Zhang, H. Liu and W. Yang, *J. Chem. Phys.*, 2000, **112**, 3483-3492.

65 D. Beeman, *J. Comput. Phys.*, 1976, **20**, 131-139.

66 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.

67 S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman and J. M. Rosenberg, *J. Comput. Chem.*, 1992, **13**, 1011-1021.

68 M. Souaille and B. Roux, *Comput. Phys. Commun.*, 2001, **135**, 40-57.

69 A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, 1988, **61**, 2635-2638.

70 A. Grossfield, WHAM: the Weighted Histogram Analysis Method, Version 2.0.6, http://membrane.urmc.rochester.edu/content/wham.

71 Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, R. A. DiStasio, Jr., R. C. Lochan, T. Wang, G. J. Beran, N. A. Besley, J. M. Herbert, C. Y. Lin, T. Van Voorhis, S. H. Chien, A. Sodt, R. P. Steele, V. A. Rassolov, P. E. Maslen, P. P. Korambath, R. D. Adamson, B. Austin, J. Baker, E. F. Byrd, H. Dachsel, R. J. Doerksen, A. Dreuw, B. D. Dunietz, A. D. Dutoi, T. R. Furlani, S. R. Gwaltney, A. Heyden, S. Hirata, C. P. Hsu, G. Kedziora, R. Z. Khalliulin, P. Klunzinger, A. M. Lee, M. S. Lee, W. Liang, I. Lotan, N. Nair, B. Peters, E. I. Proynov, P. A. Pieniazek, Y. M. Rhee, J. Ritchie, E. Rosta, C. D. Sherrill, A. C. Simmonett, J. E. Subotnik, H. L. Woodcock, 3rd, W. Zhang, A. T. Bell, A. K. Chakraborty, D. M. Chipman, F. J. Keil, A. Warshel, W. J. Hehre, H. F. Schaefer, 3rd, J. Kong, A. I. Krylov, P. M. Gill and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2006, **8**, 3172-3191.

72 J. W. Ponder, *TINKER*, *Software Tools for Molecular Design, Version 4.2*, *Washington University: Saint Louis, MO*, 2004.

73 S. K. Lüdemann, V. Lounnas and R. C. Wade, *J. Mol. Biol.*, 2000, **303**, 797-811.

74 H. Vashisth and C. F. Abrams, *Biophys. J.*, 2008, **95**, 4193-4204.

75 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781-1802.

76 R. Selmar, R. Lieberei, B. Biehl and E. E. Conn, *Physiol. Plant.*, 1989, **75**, 97-101.

77 Y. Lin, Z. Cao and Y. Mo, J. Am. Chem. Soc., 2006, 128, 10876-10884.

78 H. Bühler, F. Effenberger, S. Förster, J. Roos and H. Wajant, *Chem. Bio. Chem.*, 2003, **4**, 211-216.

79 H. Bühler, B. Miehlich and F. Effenberger, *Chem. Bio. Chem.*, 2005, **6**, 711-717.

