Molecular dynamics investigations of BioH protein substrate specificity for biotin synthesis

Qiao Xue \textsuperscript{a,b,*}, Ying-Lu Cui\textsuperscript{b}, Qing-Chuan Zheng \textsuperscript{b,*} and Hong-Xing Zhang \textsuperscript{b}

\textsuperscript{a}State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100000, P.R. China; \textsuperscript{b}Institute of Theoretical Chemistry, Jilin University, Changchun 130023, P.R. China

Communicated by Ramaswamy H. Sarma.

(Received 21 April 2015; accepted 29 June 2015)

BioH, an enzyme of biotin synthesis, plays an important role in fatty acid synthesis which assembles the pimelate moiety. Pimeloyl-acyl carrier protein (ACP) methyl ester, which is long known to be a biotin precursor, is the physiological substrate of BioH. Azeloyl methyl ester, which has a longer chain than pimeloyl methyl ester, conjugated to ACP as substrate in addition to the physiological substrate pimeloyl (C7) methyl ester, whereas the hydrolysis is slower than that of the pimeloyl substrate (Agarwal et al., 2012). Thus, the longer azeloyl chain raises the question of what the molecular origin for the experimentally observed rate changes of hydrolysis by the chain elongation have remained elusive. To this end, we have investigated chain elongation effects on the structures by using the fully atomistic molecular dynamics simulations combined with binding free energy calculations. The results indicate that the substrate specificity is determined by BioH together with ACP. The added two methylene groups would increase the structural flexibility by protein motions at the interface of ACP and BioH, instead of making steric clashes with the side chains of the BioH hydrophobic cavity. On the other hand, the slower hydrolysis of azeloyl substrate is suggested to be associated with the loose of contacts between BioH and ACP, and with the lost electrostatic interactions of two ionic/hydrogen bonding networks at the interface of the two proteins. The present study provides important insights into the structure–function relationships of the complex of BioH with pimeloyl-ACP methyl ester, which could contribute to further understanding about the mechanism of the biotin synthetic pathway, including the catalytic role of BioH.

Keywords: BioH substrate specificity; biotin synthetic pathway; molecular dynamics simulation; MM-GBSA

1. Introduction

Biotin, an essential nutrient for human beings, is necessary for cell growth and the metabolism of fats and amino acids. Biotin consists of two fused heterocyclic rings plus a valeric acid chain. In the biosynthetic pathway of biotin, several enzymes named BioA–H which are encoded by bio genes play important roles (Agarwal, Lin, Lukk, Nair, & Cronan, 2012). The steps of the pathway are starting from the synthesis of pimelic acid, a seven-carbon (C7) dicarboxylic acid (White, Zheng, Zhang, & Rock, 2005). In these steps, BioC converts the \omega-carboxyl group of malonyl to a methyl ester, which provides a methyl carbon that mimics the methyl ends of normal fatty acyl chains. Then two cycles of fatty acid synthesis are performed to give pimeloyl acyl carrier protein (ACP) methyl ester (Me-pimeloyl-ACP) (Lin, Hanson, & Cronan, 2010). Once pimelate is synthesized, BioH acts as a gatekeeper, cleaving the methyl ester of Me-pimeloyl-ACP to prevent further chain elongation. After that, BioF consumes the thioester to provide the product 7-keto-8-aminopelargonic acid (KAPA), which is required for the beginning of subsequent biotin ureido ring formation. The ACP moiety is subsequently released in the BioF reaction (Agarwal et al., 2012).

However, how can a nonspecific enzyme catalyze a specific reaction? In the absence of other factors, Agarwal et al. (2012) have suggested that BioH might derail biotin synthesis by hydrolyzing intermediates before Me-pimeloyl-ACP, and thereby abort the pathway. Previous study (Cronan, 1982) has demonstrated that acyl-ACPs sequester the first 6–8 carbon atoms of the acyl chain from solvent, and only the ester group of C7 species would become fully exposed to the active site of BioH. A hydrophobic tunnel which consists of a bundle of four helices would largely protect the short ester groups. However, BioH also indeed accepts azeloyl (C9) methyl ester conjugated to ACP as substrate in addition to the physiological substrate pimeloyl (C7) methyl ester, whereas the hydrolysis is slower than that of the pimeloyl substrate (Agarwal et al., 2012). Thus, the longer azeloyl chain raises the question of what the molecular origin of the slower catalysis in C9 species is.

An X-ray crystal structure of the complex of Me-pimeloyl-ACP with BioH S82A (PDB ID: 4ETW), which lacks the catalytic serine nucleophile in catalytic triad, was determined by Agarwal et al. (2012) to test the model of putative gatekeeper reaction. They demonstrated that BioH is the gatekeeper and acts before BioF,
Me-pimeloyl-ACP is the physiological substrate of BioH. In this study, we built a model structure mutated back from A82 to S82 to compare the differences between the crystal structure and real catalytic triad in Me-pimeloyl-ACP with BioH complex. In addition, based on the built Me-pimeloyl-ACP with BioH S82 structure, we extended the pimeloyl moiety to azelayl methyl ester to explore the substrate specificity for BioH generated by the dimensions of the hydrophobic cavity. All-atom explicit water molecular dynamics (MD) simulations reveal to what extent C9 species alters the complex structure as well as the molecular origin responsible for the slower hydrolysis of azelayl substrate. On the basis of MM-GBSA method, we show that the binding affinity changes originate mainly from hydrophilic residues rather than from hydrophobic ones. Moreover, the structural characterizations and free energy analysis allow us to bridge the gap between the theory and the experimentally observed rate changes of hydrolysis.

2. Methods

2.1. Preparation of the structure

Initial structure for the complex was taken from Protein Data Bank (PDB ID: 4ETW) (Agarwal et al., 2012), which lacks the catalytic serine nucleophile. We built a model structure mutated back from A82 to S82, and also extended the pimeloyl moiety to azelayl methyl ester by using Discovery Studio 2.5 (Studio, 2007; Xuan-yu, Zhuo, Rui-juan, Hong-xing, & Qing-chuan, 2012). Hydrogen atoms were added to three systems with the t-Leap procedure of AMBER11 (Case et al., 2010). All crystal water molecules were kept during the MD simulations. The force field parameters of pimeloyl moiety or azelayl moiety linked with S35 were supplied by AMBER11 with electrostatic potential fitting procedure and Gaussian09 software (Frish et al., 2009). Structural optimizations of pimeloyl moiety or azelayl moiety linked with S35 were conducted using B3LYP combined with 6–31G* (Krański, Cypryk, Kwiatkowska, Mikołajczyk, & Kielbasinski, 2012; Xu, Meher, Eustache, & Wang, 2014) basis set using the Gaussian09 software. RESP fitting procedure was used for charge derivation based on the optimal geometry.

2.2. MD simulations

MD simulations for three systems including energy minimization were carried out using AMBER11 (Case et al., 2010) software package and the ff99SB force field (González-Mendiöroz, Álvarez-Vázquez, & Rubio-Martínez, 2012; Shinde & Sobhia, 2013). To keep the whole systems neutral, sodium ions (Na⁺) were added. Each system was then solvated with the TIP3P water model in a truncated octahedron box with a 10 Å distance around the solute. Whole systems were subjected to 2000 steps of steepest descent minimization followed by 3000 steps of conjugate gradient minimization, and then heated from 0 to 300 K in 300 ps by a Langevin dynamics with the collision frequency of 1 ps⁻¹. Then, the systems went through 500 ps equilibrium MD simulations. Finally, a total of 60 ns were simulated for each system under NPT ensemble condition using periodic boundary conditions and particle-mesh Ewald (PME). Short-range interactions were cut off at 12 Å, and bonds involving hydrogen were held fixed using SHAKE. The time step was set to 2 fs. The cluster analysis of the protein conformations was performed using the average linkage as the clustering algorithm, and backbone atom RMSD as the distance metric. The representative structures of the clusters with highest occurrences were chosen to present the structural information. The hydrogen bond is defined if the distance between donor and acceptor is shorter than 3.5 Å and the angle is less than 60°. The ionic bond is defined if the distance is less than 5 Å. The ionic bond is considered to be formed between the negative- and positive-charged amino acids. All of the figures were created with PyMOL and Chimera (DeLano, 2002; Pettersen et al., 2004).

2.3. Free energy calculations

The MM-GB/SA methods (Du et al., 2010; Mulakala & Viswanadhan, 2013) implemented in AMBER11 (Case et al., 2010) was applied to calculate the binding free energy between BioH and ACP. The binding free energy (ΔGbind) in MM-GB/SA between a ligand and a receptor to form a complex was calculated as

$$
\Delta G_{\text{bind}} = G_{\text{complex}} - (G_{\text{receptor}} + G_{\text{ligand}})
$$

$$
G = E_{\text{MM}} + G_{\text{sol}} - TS
$$

$$
E_{\text{MM}} = E_{\text{int}} + E_{\text{ele}} + E_{\text{vdw}}
$$

$$
G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}}
$$

In Equation 2, the $E_{\text{MM}}$, $G_{\text{sol}}$, and $TS$ represent molecular mechanics component in gas phase, the stabilization energy due to salvation, and a vibrational entropy term, respectively. $E_{\text{MM}}$ is given as a sum of $E_{\text{int}}$, $E_{\text{ele}}$, and $E_{\text{vdw}}$ which are internal, Coulomb and van der Waals interaction terms, respectively. Solvation energy, $G_{\text{sol}}$, is separated into an electrostatic solvation free energy ($G_{\text{GB}}$) and a nonpolar solvation free energy ($G_{\text{SA}}$). The former can be obtained from the generalized Born (GB) method. The latter is considered to be proportional to the molecular solvent accessible surface area (SASA). The binding free energies were obtained by...
averaging over the values calculated for 20,000 snapshots from the last 40 ns of the trajectories at 2 ps intervals for the complex structures.

To obtain a detailed view of BioH and Me-pimeloyl-ACP interactions, MM-GBSA method was also used to calculate the binding free energy of each residue at the interface between two proteins. Each residue can be partitioned into two groups: backbone and side chain. The snapshots used in the binding free energy decomposition are the same as those used in the binding free energy calculation.

2.4. SMD simulation

In the study, a total of 10 steered molecular dynamics (SMD) simulations of each system were performed using Amber11. For each simulation, the reaction coordinate was specified as the distance between the side chain O atom (OG) of Ser82 and the hydroxyl oxygen atom (OAX) of pimeloyl methyl ester (C7) as well as OG of Ser82 and the OAX of azelayl methyl ester (C9). The pulling velocity was set as 1 Å/ns. C7 and C9 were pulled from about 3.6–26 Å on the reaction coordinate. The position at 26 Å was absolutely out of the binding site. A large spring constant \( k = 5000 \text{ kcal mol}^{-1} \text{ Å}^{-2} \) was used to ensure that the chosen atoms follow the constraint closely, which is required for the calculation of Jarzynski’s identity (Park, Khalili-Araghi, Tajkhorshid, & Schulten, 2003). Jarzynski’s equality was used to establish a connection between equilibrium free energies calculation and nonequilibrium processes (Jarzynski, 1997). Based on the result, potential of mean force (PMF) was calculated.

3. Results and discussion

3.1. Comparison between the crystal structure and the real catalytic triad in Me-pimeloyl-ACP with BioH complex

Initial structure for the complex of Me-pimeloyl-ACP with BioH S82A was taken from Protein Data Bank (PDB ID: 4ETW) (Agarwal et al., 2012), which lacks the catalytic serine nucleophile. To mimic the real environment of the esterase reaction mediated by BioH, we built a model structure mutated back from A82 to S82. Note that this paper refers to Me-pimeloyl-ACP with BioH S82A as M7-AB-S82A and the model enzyme as M7-AB.

ACP consists of four well-defined helices bundle. Three major helices (α-1, 2, 4) run largely parallel to each other, and α-3 markedly shorter and almost perpendicular to the bundle. BioH consists of a core domain of a seven-member central β-sheet flanked on either side by α helices and a four helical capping domain present at the top of the loops emanating from the central β-sheet. The catalytic triad (S82-H235-D207) locates at the interface of the core domain and the capping domain.

The distributions of the Cu root-mean-square deviations (RMSD) of both M7-AB-S82A and M7-AB with respect to their initial coordinates are displayed in Figure S1. It is clear that M7-AB-S82A has much larger fluctuations than that of M7-AB. The result indicates that the catalytic triad (S82-H235-D207) might have a stabilizing effect on the Me-pimeloyl-ACP with BioH complex. The representative structures of both complexes based on the clustering analysis were superimposed with each other to obtain the comparable local structural features of the protein (Figure 1).

The A82 to serine mutation adds an oxygen atom of S82 side chain to provide the nucleophile for the reaction. As shown in Figure 1, the pimeloyl carboxylate carbon, which would be the site of nucleophilic addition, deviates from the initial position and comes closer to the OG atom of S82 in M7-AB complex. The distance between the A/S82 and the pimeloyl carboxylate carbons also proves this, see Figure S2. With the movement of the pimeloyl methyl ester, the long interaction between the carboxylate oxygen atom of the pimeloyl moiety and the side chain of Q147 makes a rotation of the Q147 side chain by nearly 120°. Such reorientation of Q147 side chain can also be observed in M9-AB system, which will be discussed in the following section. The results imply an important interaction between Q147 and the carboxylate oxygen atom of the pimeloyl moiety to stabilize the pimeloyl methyl ester, which were not discussed in the crystal structure. In addition, it is suggested that Q147 together...
with H235, L83, W22, collectively make interactions with the carboxylate oxygen atoms to constitute the "oxyanion hole" for stabilization of the tetrahedral hemiacetal intermediate (Jaeger, Dijkstra, & Reetz, 1999) by NH groups (Figure 2). The results also emphasize the importance of dynamic structural analysis for investigating real conditions of the binding of ACP with BioH because the catalytic triad could not be completed in the crystal structure (PDB ID: 4ETW).

### 3.2. Elongation of pimeloyl carbon chain to C9 species would not make steric clashes with the BioH hydrophobic cavity

Previous work demonstrated that only upon chain elongation to the C7 species would the ester group enter the active site and become fully exposed to BioH (Roujeinikova et al., 2007). Although azeloyl (C9) methyl esters conjugated to ACP would be accepted by BioH, the hydrolysis of this substrate is slower than that of C7 substrate (Agarwal et al., 2012). Thus based on the built M7-AB structure, we extended the pimeloyl moiety to azeloyl methyl ester to explore the molecular origin for the substrate specificity for BioH generated by the dimensions of the hydrophobic cavity. The built Me-azeloyl-ACP with BioH structure is abbreviated as M9-AB.

In the 60 ns MD simulations, RMSD of all Cα atoms for both complex structures were calculated to provide an overall measure of the departure of the structures from the initial coordinates (Figure S3). It is clear that the RMSD values are convergent and the systems remain in equilibrium during the last 40 ns. All subsequent conformational characteristic analysis of M7-AB and M9-AB were performed on the last 40 ns of the simulation trajectories. During the last 40 ns simulations, the Cα RMSD values for both structures stay fairly low with the average RMSD values of 1.75 Å (standard deviation: .15 Å) and 1.97 Å (.19 Å) for C7-AB and C9-AB, respectively.

The representative structures of C7-AB and C9-AB based on the clustering analysis are displayed in Figure 3.

The hydrophobic channel cavity allows the pimeloyl moiety to efficiently reach the BioH catalytic triad, whereas the longer chain has no difficulty in traversing such cavity. As shown in Figure 3, the added two carbon chain methylenes would not make steric clashes with the side chains of the hydrophobic cavity. Instead, the pantetheine moiety of azeloyl methyl ester curls up towards α-4 helix of BioH, leading to a little movement of the α4 helix of BioH, see Figure 4. Thus, we can infer that the elongation of the carbon chain would not make steric clashes to BioH but induce the shift of BioH local structures, especially for α-4 helix.

The intermolecular interaction between C7/C9 and the catalytic triad (S82-H235-D207), see Figure S4. The distances between C7-S82, C7-H235 were 3.660, 3.938 Å, respectively. Compared with C7, the distances between C9-S82, C9-H235 were 3.761, 4.507 Å, respectively. The distances between S82 and H235 were 3.237, 4.234 for M7-AB and M9-AB, respectively. Obviously, C7 was more close with the catalytic triad than C9. The intermolecular interaction between C7 and the catalytic was stronger than C9. Thus, when the interaction started, the close structure would be helpful for the faster hydrolysis rate.

![Figure 2](image1.png) Q147, W22, L83, H235 constitute the “oxyanion hole” in M7-AB (yellow). The side chain of Q147 makes a rotation in M7-AB-S82A (blue).

![Figure 3](image2.png) Superposition of M7-AB (yellow) and M9-AB (pink). Helices 4 and 5 of BioH and helices of ACP are labeled. The squared region refers to the evident different structural regions in the two complexes.

![Figure 4](image3.png) The positions of α-4 helix residues in M7-AB (yellow) and M9-AB (pink).
3.3. Longer chain causes conformational changes at the interface of ACP and BioH

The distributions of the radius of gyration (Rg) as a function of time are shown in Figure 5. The higher Rg values of C9-AB, compared with C7-AB, may suggest the decreased conformational stability of C9-AB complex. Another look at the plots in Figure S5 points out the structural stability of the individual elements of secondary structure. As shown in Figure S5, on average, the global secondary structure motifs for BioH are well maintained over the last 40 ns simulations of two complexes. Figure S5(B) and (D) demonstrate that the C-terminal of α1 helix in ACP (residues 10–15) of C9-AB is almost lost, whereas that of C7-AB is normally maintained until the end of the simulation. In addition, the N-terminal of α2 helix in ACP (residues 35–40) of C9-AB is more unstable than that of C7-AB. The compared root-mean-square fluctuations (RMSF) of all Cα atoms for two systems were calculated to find out which regions of the proteins are more flexible and are undergoing larger structural changes. (Figure 6) Both simulations show similar fluctuations throughout the complexes, expect for several regions. As shown in Figure 6, C9-AB shows obviously larger fluctuations for the α8 helix for BioH as well as for the C-terminal of α1 helix and α1–α2 loop for ACP. C9-AB also shows slightly larger fluctuations around α4, α5 helices for BioH, and α2, α4 helices for ACP.

The superimposed representative structures in Figure 3 reveal conformational changes observed at the interface of ACP and BioH (α4 to α8 helices of BioH as well as α1, α2 helices and α1–α2 loop of ACP). This observation explains the larger RMSF values of the corresponding locations for C9-AB in Figure 6. Noticeably, the residues in α4 helix of BioH make movements to acclimate to the curled up chain of azelayl pantetheine moiety, leading to the concomitant conformational changes accordingly. These conformational changes would strongly affect the intermolecular interactions between ACP and BioH, which would be discussed in the following section. In addition, the ACP serine residue (Ser35) that anchors the phosphopantetheine arm has minor shift due to the different positions of phosphopantetheine arm for M9-AB, which might have impact on the motion of α1–α2 loop. Overall, chain elongation to the C9 species causes conformational changes at the interface of ACP and BioH as well as decreases thereby complex stability of native conformation.

![Figure 5. Rg as a function of simulation time. M7-AB (black), M9-AB (red).](image)

![Figure 6. RMSF relative to initial structure for backbone atoms of M7-AB (black) and M9-AB (red). Secondary structures are labeled.](image)
3.4. Longer chain reduces the intermolecular interactions between ACP and BioH

To further characterize the effect of C9 species on the interaction between ACP and BioH, we analyzed the hydrogen bonds as well as ionic bonds formed at the interface of ACP and BioH (Table S1). The intermolecular contacts between BioH and ACP could be concluded as two ionic/hydrogen bonding networks. (Figures 7 and 8) One ionic/hydrogen bonding network involves D34 and D37 in ACP α2 helix as well as R138 and R142 in BioH α5 helix. Another network involves E46, E47, and E52 in ACP α2 helix as well as R155 and R159 in BioH α6 helix. As shown in Table S2, it is clear that the hydrogen bonds between ACP D37 and BioH R138 for M9-AB decrease by half, whereas the hydrogen bonds between D37 and R142, E52 and R159 disappear for M9-AB. The ionic bonds between D37 and R138, as well as between D37 and R142, also decrease for M9-AB. From Figure 7, we could demonstrate that the reduced interactions between ACP D37 and two BioH Arg residues may be attributed to the movement of ACP α2 helix including the motion of D37. Similarly, the reduced interactions between E52 and R159 may be resulted from the motion of E40 in ACP α2 helix. In addition, the curled-up pantetheine moiety occupies the original position of Q131, directly leading to the departure of Q131. The hydrogen bonds between Q131 and R138 break accordingly, thus the guanidine side chain of R138 makes a rotation toward two carboxylate groups (D37 and E40) which could explain the formation of hydrogen bonds and ionic bonds between ACP E40 and BioH R138 for M9-AB. Moreover, the movement of ACP α2 helix also has an impact on the interactions involved in the second ionic bonding network. The

3.5. Binding free energy calculations

To get greater insights into the interface interactions between BioH and ACP, we performed binding free energy calculations using MM-GBSA method (Cui et al., 2013; Xue et al., 2013). The binding free energies for both complexes are displayed in Table 1. In order to clearly discern the contributions of binding free energy of each residue, the energies were also decomposed into individual residue contributions. A large fraction of the binding affinity for both two ligands can be attributed to a common set of residues. The contributions of these residues are listed in Table S3.

Overall, for C7-AB complex, there are about $-229.14$ kJ/mol of overall affinity, with $-422.84$ kJ/mol from van der Waals contacts, $-3856.42$ kJ/mol from intermolecular electrostatic contributions, $3930.03$ kJ/mol from electrostatic solvation interactions, and $-65.79$ kJ/mol from nonpolar solvation contributions. For C9-AB complex, there are about $-134.97$ kJ/mol of overall affinity, with about $-371.05$ kJ/mol from van der Waals contacts, $-3611.89$ kJ/mol from intermolecular electrostatic contributions, $3682.53$ kJ/mol from electrostatic solvation interactions, and $-60.86$ kJ/mol from nonpolar solvation contributions. Consistent with previous analysis, the calculations exhibit the most profound change for C9-AB. The binding free energy for C9-AB is decreased by 94.17 kJ/mol, which is in accord with experimentally observed slower hydrolysis of azelayl substrate (Agarwal et al., 2012). By decomposing the binding free energy into different items, it is clear that

Figure 7. Ionic bonds at the interface of BioH and ACP. (A) and (C) present the ionic bonds of M7-AB (BioH, yellow; ACP, orange). (B) and (D) present the ionic bonds of M9-AB (BioH, pink; ACP, blue).
the most important contributions of binding affinity are coming from the intermolecular electrostatic interactions which are generally presented as hydrogen bonds or ionic interactions between two proteins. The intermolecular electrostatic contributions for C7-AB are much lower than that for C9-AB by \(-244.53\) kJ/mol. The contribution of van der Waals interactions of C9-AB also decreases significantly. This reduction is mainly due to the untight contacts of Me-azelalyl-ACP with BioH.

When the binding free energies are decomposed into individual residues, the lost binding affinity is suggested to be associated with several residues. The differences of binding free energies for each residue for two systems are displayed in Figure 9. Noticeably, the overall free energy change is found to originate mainly from the ACP \(\alpha2\) helix and BioH \(\alpha5, \alpha6\) helices regions. Furthermore, we found that the main contributors to the intermolecular electrostatic interactions changes are hydrophilic residues, which are mainly involved in the bonding networks. Moreover, those changes mainly originate from side chains rather than main chains. Such results of decreased binding free energies are consistent with the reduced intermolecular interactions of these residues

| System | \(\Delta E_{\text{vdw}}\) | \(\Delta E_{\text{ele}}\) | \(\Delta G_{\text{GB}}\) | \(\Delta G_{\text{SA}}\) | \(\Delta H^a\) | \(\Delta S^a\) | \(\Delta G^b\) |
|--------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| M7-AB  | -422.84 ± 26.58 | -3856.42 ± 165.98 | 3930.03 ± 154.49 | -65.79 ± 2.54 | -414.82 ± 33.48 | 185.67 ± 7.23 | -229.14 ± 14.29 |
| M9-AB  | -371.05 ± 23.24 | -3611.89 ± 122.30 | 3682.53 ± 112.06 | -60.86 ± 1.67 | -361.27 ± 26.37 | 226.30 ± 17.63 | -134.97 ± 4.81 |

\(^a\Delta H = \Delta E_{\text{ele}} + \Delta E_{\text{vdw}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}}.\)

\(^b\Delta G = \Delta H - \Delta S.\)
residues, which could cause an overall decrease in the complex binding affinity.

In addition, F111, M149, and H235 are found to decrease their contributions to the azelalyl substrate binding. M111 and M149 lie on the hydrophobic channel, and M149 has interaction with the phosphate group. Remarkably, the binding energy change of H235, which is involved in the catalytic triad, may have direct impact on the esterase activity.

To estimate the binding affinity of C7 and C9 on the catalytic domain, as well as the accessibility for C7 and C9 migrating along the hydrophobic cavity, the migration free energy profiles were evaluated in Figure 10. Ten parallel trajectories for each system were performed to obtain the free energy information. The reaction coordinate was specified as Ser82 (OG) and C7 (OAX) as well as Ser82 (OG) and C9 (OAX). In the SMD process, both C7 and C9 were pulled out of the active site along the hydrophobic tunnel. In the dissociation process, the energies of C7 and C9 changed to 181 and 163 kJ/mol, respectively. The PMF values of M7-AB and M9-AB had about 18 kJ/mol difference. The dissociation of C7 needed more energy than C9. This result reflected the binding affinity and stability of C7 was better than C9. The stable catalytic structure would be helpful for the catalytic interaction. The SMD simulations and the MM-GBSA calculations give an agreement with the experimental result of the catalytic rate difference caused by the chain length.

Overall, the slower hydrolysis of azelalyl substrate is suggested to be associated with the lost binding affinity between ACP and BioH. The result of MM-GBSA shows the importance of ACP in the recognition of substrate. As the pimeloyl methyl ester carrier protein, ACP offers most of the binding free energies to BioH. Without this protein, pimeloyl methyl ester is hard to access the hydrophobic cavity. This is consistent with the experimental result (Agarwal et al., 2012) that pimeloyl methyl ester is hydrolyzed first and then releases ACP moiety. Elongation of pimeloyl carbon chain to C9 species would increase the structural flexibility by protein motions at the interface of ACP and BioH. The contacts of ACP and BioH become loose which induce the loss of van der Waals interactions and the disruption of the ionic/hydrogen bonding network. From the previous researches (Agarwal et al., 2012), we know that the short carbon chain (C5 or C6) fails to make interaction with the catalytic triad (S82-H235-D207) even the ACP binds to BioH. In the present work, the results show that although the long chain (C9) occupies the binding site of BioH, the long chain makes it difficult for ACP to bind with BioH. Thus, we can suggest that BioH together with ACP determine the substrate specificity.

4. Conclusion

By performing fully hydrated, all-atom MD simulations with MM-GBSA calculations, we have investigated the atomic-level structural variations for Me-pimeloyl-ACP with BioH, and Me-azelalyl-ACP with BioH of experimental interests to understand the molecular origin for the substrate specificity for a nonspecific enzyme. Although upon chain elongation to C9 species would not make steric clashes with the side chains of the BioH hydrophobic cavity, the added two carbon chain methylenes cause conformational changes at the interface of ACP and BioH. The elongation of the carbon chain not only relaxes the contacts of BioH and ACP, but also disrupts the two intermolecular ionic/hydrogen bonding networks. Moreover, the decreased binding affinity for Me-azelalyl-ACP with BioH observed by MM-GBSA is in accord with the experimental observations that the hydrolysis of this azelalyl substrate is slower than that of the pimeloyl substrate, and that ACP α2 helix plays a critical role in determining the binding affinity. The decomposition energy analysis highlights the fact that the changes in free energy mainly originate from hydrophilic residues, rather than from the hydrophobic residues. Furthermore, ACP plays an indispensable role in stabilizing the substrate binding within the BioH active site. BioH together with ACP determine the substrate specificity. In summary, the MD simulation results with MM-GBSA calculations render the structural motif and molecular origin to rationalize the experimentally observed rate changes of hydrolysis and enhance the understanding of the substrate specificity in biotin synthesis.

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2015.1068223.
Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This work was supported by Chinese Academy of Sciences [grant number KZCX2-YW-BR25]; National Natural Science Foundation of China [grant number 21273095].

References
Agarwal, V., Lin, S., Lukk, T., Nair, S. K., & Cronan, J. E. (2012). Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. Proceedings of the National Academy of Sciences, 109, 17406–17411.

Case, D., Darden, T., Cheatham, T., III, Simmerling, C., Wang, J., Duke, R., ... Merz, K. (2010). AMBER 11. San Francisco: University of California. p. 142.

Cronan, J. (1982). Molecular properties of short chain acyl thioesters of acyl carrier protein. Journal of Biological Chemistry, 257, 5013–5017.

Cui, Y. L., Zhang, J. L., Zheng, Q. C., Niu, R. J., Xu, Y., Zhang, H. X., & Sun, C. C. (2013). Structural and dynamic basis of human cytochrome P450 7B1: A survey of substrate selectivity and major active site access channels. Chemistry-A European Journal, 19, 549–557.

DeLano, W. L. (2002). The PyMOL molecular graphics system (Version 1.1). Portland, OR: Schrödinger, LLC.

Discovery Studio. (2007). version 2.0. San Diego, CA: Accelrys.

Du, J., Xi, L., Lei, B., Lu, J., Li, J., Liu, H., & Yao, X. (2010). Structure-based quantitative structure-activity relationship studies of checkpoint kinase 1 inhibitors. Journal of Computational Chemistry, 31, 2783–2793.

Frisch, M., GWT, H., Schlegel, G., Scuseria, M., Robb, J., Cheeseman, G. S., ... Nakatsuji, H. (2009). Gaussian 09. Wallingford, CT: Gaussian Inc.

González-Mendióroz, M., Álvarez-Vázquez, A. B., & Rubio-Martínez, J. (2012). Structural analysis of the inhibition of APRIL by TACI and BCMA through molecular dynamics simulations. Journal of Molecular Graphics and Modelling, 39, 13–22.

Jaeger, K., Dijkstra, B., & Reetz, M. (1999). Bacterial biocatalysts: Molecular biology, three-dimensional structures, and biotechnological applications of lipases. Annual Review of Microbiology, 53, 315–351.

Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. Physical Review Letters, 78, 2690–2693.

Krasinski, G., Cypryk, M., Kwiatkowska, M., Mikolajczyk, M., & Kielbasiński, P. (2012). Molecular modeling of the lipase-catalyzed hydrolysis of acetoxymethyl (t-propoxy) phenylphosphine oxide and its P-borane analogue. Journal of Molecular Graphics and Modelling, 38, 290–297.

Lin, S., Hanson, R. E., & Cronan, J. E. (2010). Biotin synthesis begins by hijacking the fatty acid synthetic pathway. Nature Chemical Biology, 6, 682–688.

Mulakala, C., & Viswanadhan, V. N. (2013). Could MM-GBSA be accurate enough for calculation of absolute protein/ligand binding free energies? Journal of Molecular Graphics and Modelling, 46, 41–51.

Park, S., Khalili-Araghi, F., Tajkhorshid, E., & Schulten, K. (2003). Free energy calculation from steered molecular dynamics simulations using Jarzynski’s equality. The Journal of Chemical Physics, 119, 3559–3566.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera? A visualization system for exploratory research and analysis. Journal of Computational Chemistry, 25, 1605–1612.

Roujeinitkova, A., Simon, W. J., Gilroy, J., Rice, D. W., Rafferty, J. B., & Slabas, A. R. (2007). Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. Journal of Molecular Biology, 365, 135–145.

Shinde, R. N., & Sobhia, M. E. (2013). Binding and discerning interactions of PTP1B allosteric inhibitors: Novel insights from molecular dynamics simulations. Journal of Molecular Graphics and Modelling, 45, 98–110.

White, S. W., Zheng, J., Zhang, Y.-M., & Rock, C. O. (2005). The structural biology of type II fatty acid biosynthesis. Annual Review of Biochemistry, 74, 791–831.

Xu, Z., Meher, B. R., Eustache, D., & Wang, Y. (2014). Insight into the interaction between DNA bases and defective graphs: Covalent or non-covalent. Journal of Molecular Graphics and Modelling, 47, 8–17.

Xuan-yu, M., Zhuo, L., Rui-juan, N., Hong-xing, Z., & Qing-chuan, Z. (2012). Molecular docking of 3-methyldiene-containing drugs binding into CYP3A4. Chemical Research in Chinese Universities, 28, 137–141.

Xue, Q., Zhang, J.-L., Zheng, Q.-C., Cui, Y.-L., Chen, L., Chu, W.-T., & Zhang, H.-X. (2013). Exploring the molecular basis of dsRNA recognition by Mss116p using molecular dynamics simulations and free-energy calculations. Langmuir, 29, 11135–11144.