

## RESEARCH ARTICLE

# Insight into substrate-assisted catalytic mechanism and stereoselectivity of bifunctional nocardicin thioesterase

Qian Yu<sup>1</sup> | Wei Lu<sup>2</sup> | Daixi Li<sup>3,4</sup> | Ting Shi<sup>1</sup> 

<sup>1</sup>State Key Laboratory of Microbial Metabolism, Joint International Research Laboratory of Metabolic and Developmental Sciences, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

<sup>2</sup>Department of Food Science and Engineering, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China

<sup>3</sup>Institute of Biothermal Science and Technology, University of Shanghai for Science and Technology, Shanghai, China

<sup>4</sup>AI Research Center, Peng Cheng National Laboratory, Shenzhen, Guangdong, China

## Correspondence

Ting Shi, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, No. 800 Dongchuan Rd., Shanghai 200240, China.

Email: [tshi@sjtu.edu.cn](mailto:tshi@sjtu.edu.cn)

## Funding information

National Key R&D Program of China, Grant/Award Numbers: 2019YFA0905400, 2021YFC2100600; National Natural Science Foundation of China, Grant/Award Number: 32070041

## Abstract

The inversion from L- to D-stereochemistry endows peptides improved bioactivity and enhanced resistance to many proteases and peptidases. To strengthen the biostability and bioavailability of peptide drugs, enzymatic epimerization becomes an important way to incorporate D-amino acid into peptide backbones. Recently, a bifunctional thioesterase NocTE, which is responsible for the epimerization and hydrolysis of the C-terminal (*p*-hydroxyphenyl)glycine residue of  $\beta$ -lactam antibiotic nocardicin A, exclusively directs to the generation of D-diastereomers. Different from other epimerases, NocTE exhibits unique stereochemical selectivity. Herein, we investigated the catalytic mechanism of NocTE via molecular dynamic (MD) simulations and quantum mechanical/molecular mechanics (QM/MM) calculations. Through structural analyses, two key water molecules around the reaction site were found to serve as proton mediators in epimerization. The structural characteristics inspired us to propose a substrate-assisted mechanism for the epimerization, where multi-step proton transfers were mediated by water molecules and  $\beta$ -lactam ring, and the free energy barrier was calculated to be 20.3 kcal/mol. After that, the hydrolysis of D-configured substrate was energetically feasible with the energy barrier of 14.3 kcal/mol. As a comparison, the energy barrier for the direct hydrolysis of L-configured substrate was obtained to be 24.0 kcal/mol. Our study provides mechanistic insights into catalytic activities of bifunctional thioesterase NocTE, uncovers more clues to the molecular basis for stereochemical selectivity and paves the way for the directed biosynthesis of novel peptide drugs with various stereostructural characteristics by enzyme rational design.

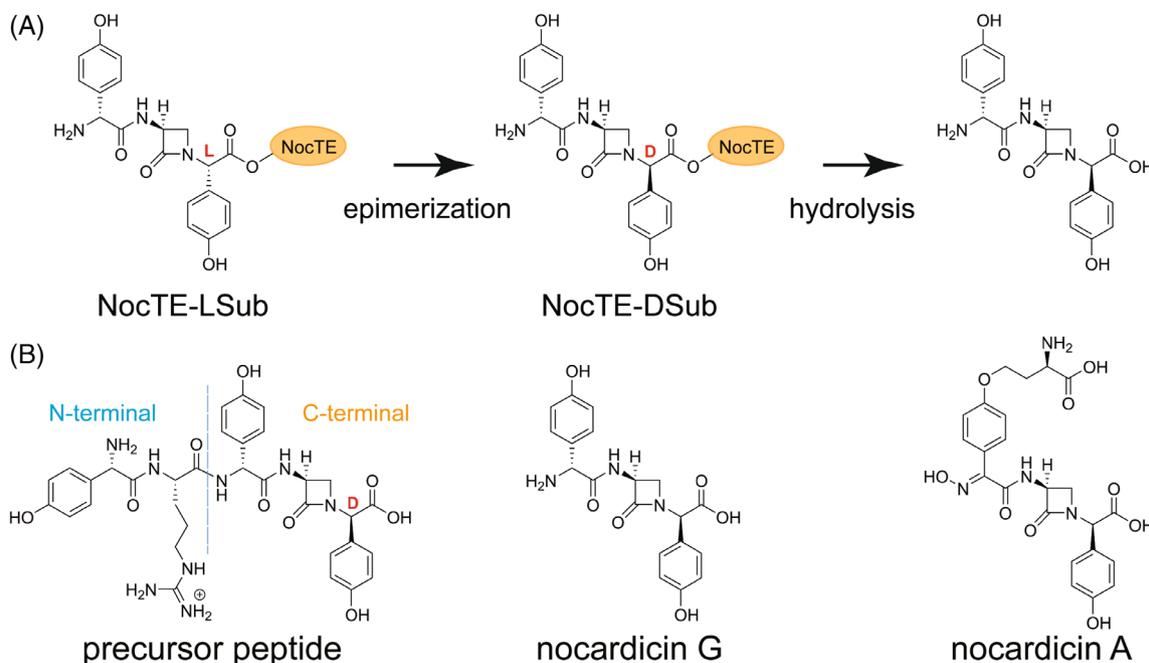
## KEYWORDS

epimerization, MD simulations, peptide drug, QM/MM, thioesterase

## 1 | INTRODUCTION

Due to remarkable advantages over many small molecule drugs, such as lower toxicity and enhanced selectivity, peptide drugs have been applied to a wide range of diseases including cancer, bacillosis, and immunological rejection.<sup>1,2</sup> Nowadays, peptide backbones originating from microbial biosynthesis become one of important sources for drug development. However, L-amino acid as the common peptide unit is

very vulnerable to proteases and peptidases, which limits the medical application of peptide compounds. Therefore, the introduction of D-amino acids is regarded as an effective approach to enhancing the biostability and bioavailability of peptide drugs, which can withstand the hydrolysis activity of most endogenous enzymes and endow peptides stereochemical properties.<sup>3</sup> In general, the enzyme responsible for epimerization always produces an equilibrating pool of L-antipode and D-antipode or diastereomers, from which the downstream



**FIGURE 1** (A) The epimerization-hydrolysis catalytic activities of bifunctional thioesterase NocTE; (B) the structures of the precursor peptide, nocardicin G and nocardicin A

enzyme exclusively selects D-configured peptide chain for the following process.<sup>4</sup>

Uniquely, nocardicin thioesterase NocTE, which develops additional epimerization activity and stereochemical selectivity, prefers D-Hpg to L-Hpg in the C-terminal of products with high diastereomeric purity (Figure 1A).<sup>5,6</sup> The bifunctional thioesterase participates in the biosynthesis of β-lactam antibiotic nocardicin A as an essential domain of the nonribosomal synthetase NocB.<sup>7</sup> Nocardicin A is active against various gram-negative bacteria and possesses β-lactamase resistance.<sup>5,8,9</sup> During its biosynthesis, nonribosomal peptide synthetases NocA and NocB assemble a precursor peptide using nonproteinogenic D-(p-hydroxyphenyl)glycine (D-Hpg) residues and some common proteinogenic residues (L-Hpg-L-Arg-D-Hpg-L-β-lactam-D-Hpg; Figure 1B). The precursor peptide contains a C-terminal core peptide and an N-terminal follower peptide that has been proven to determine substrate selectivity for NocTE in previous studies.<sup>10</sup> After the cleavage of N-terminal follower peptide, the tripeptide nocardicin G is generated as the core structure of nocardicin A. Nocardicin G contains two D-Hpg residues flanking a β-lactam ring (D-Hpg-L-β-lactam-D-Hpg). The N-terminal D-Hpg is produced by the typical epimerization domain, while the C-terminal D-Hpg formation is catalyzed by bifunctional thioesterase NocTE.

The hydrolytic activity of NocTE is mediated by catalytic triad S1779-H1901-D1806. Once the substrate is loaded on the serine (S1779) of NocTE, the histidine (H1901) captures a proton from a water molecule near the active site and promotes the remaining nucleophile to attack the substrate, forming a tetrahedral hydrolytic intermediate. Then, the covalent bond between the substrate and S1779 breaks and the product gradually escapes from the active pocket. At the same time, H1901 returns the proton toward S1779 to

restore the initial enzymatic environment, ready for the next catalysis cycle. However, the epimerization catalyzed by thioesterases has not been studied so far. As we know, two catalytic residues are usually the requisite for epimerization reaction, one of which acts as the base to abstract the proton (deprotonation) and then the other (maybe a water molecule) complements a proton (re-protonation).<sup>4,11–13</sup> For the mysterious epimerization activity of NocTE, Gulick et al.<sup>6</sup> has proposed that the histidine H1901 of the catalytic triad deprotonates the substrate according to site-directed mutagenesis results and crystal architecture characteristics, while the detailed catalytic mechanism is still ambiguous. Although the catalytic activities of S1779 mutants and H1901 mutants are completely lost, it is undistinguishable whether these mutations mainly effect the hydrolysis or epimerization reaction. In addition, H1808 mutant can catalyze the generation of a few L-diastereomers (*epi*-nocardicin G, D-Hpg-L-β-lactam-L-Hpg), influencing the stereochemical selectivity of NocTE. More detailed information about NocTE needs further investigations, such as how the epimerization reaction proceeds, how the epimerization and hydrolysis combine, and why *epi*-nocardicin G is unable to hydrolyze efficiently before the stereochemical inversion.

In this study, we investigated the catalytic mechanism and stereochemical selectivity of the thioesterase NocTE through a series of computational methods. The pre-reaction structures for epimerization were sampled by analyzing the distributions of deprotonation conformations. Combining structural characteristics abstracted from MD simulations, the molecular mechanism of the substrate-assisted epimerization was proposed and confirmed by QM/MM calculations. Besides, the energy barrier difference between hydrolysis reactions for *epi*-nocardicin G and nocardicin G further explained the stereochemical selectivity for NocTE catalysis. Our studies uncover the

catalytic mechanism of bifunctional thioesterase NocTE and generate a spark of appreciation for the research and development of novel peptide drugs.

## 2 | MATERIALS AND METHODS

### 2.1 | System construction

Based on the crystal structure of NocTE bound to the mimic structure of the hydrolytic tetrahedral intermediate of nocardicin G (PDB ID: 6OJD),<sup>6</sup> two catalytic systems of NocTE covalently bound to *epi*-nocardicin G (NocTE-LSub) and nocardicin G (NocTE-DSub) were modeled, respectively. The protonation state for each residue was determined via  $pK_a$  calculations on PDB2PQR web serve.<sup>14</sup> For parameter preparations, the substrates covalently bound to NocTE were capped with  $-\text{CO}-\text{CH}_3$  group in the N-terminal and  $-\text{NH}-\text{CH}_3$  group in the C-terminal. Then, the capped geometric structure was optimized at M062X<sup>15</sup>/6-31G\*<sup>16</sup> level and electrostatic surface potential (ESP) charge calculations were carried out at HF<sup>17</sup>/6-31G\* level using Gaussian 09 program.<sup>18</sup> The charge distribution information used for system construction was calculated by a two-step restrained electrostatic potential (RESP)<sup>19</sup> method through Multiwfn<sup>20</sup> program and the antechamber<sup>21</sup> package implemented in AMBER 18.<sup>22</sup> These systems were set up within an octahedral box of TIP3P water molecules. Besides, 15 sodium ions were added to maintain the electroneutrality in each system.

### 2.2 | Molecular dynamic simulations

Through AMBER 18 program suite, NocTE-LSub and NocTE-DSub systems were subjected to molecular dynamic (MD) simulations using ff14SB force field.<sup>23</sup> First, energy minimizations were carried out for these complexes to rationalize crystal structures. Afterwards, each system was heated from 0 to 300 K and equilibrated for 50 ps under the isothermal–isobaric ensemble. Starting from this preprocessed conformation, six replicas of MD simulation trajectories were collected. Long-range electrostatic interactions and bond lengths in simulations were described using Particle Mesh Ewald (PME) method<sup>24</sup> and the SHAKE algorithm.<sup>25</sup> Trajectory analyses, like root mean square deviation (RMSD) fluctuations and conformation distributions, were performed via the *cpptraj* module<sup>26</sup> of AMBER 18.

For epimerization, the deprotonation distance  $d(\text{H1901}_{\text{N}_\epsilon} - \text{LSub}_{\text{H}_\alpha})$  was constrained between 3.00 and 3.50 Å using a harmonic vibrational potential with a force constant of 200 kcal/(mol Å<sup>2</sup>) in the equilibration step for NocTE-LSub system, catching the critical pre-reaction structure for epimerization. Then, the system was re-equilibrated without constraints and six replicas of 100 ns MD simulation trajectories were obtained. According to the distribution of the deprotonation distance and angle, two optimal trajectories md2 and md6 were selected and lengthened to 310 ns for the following mechanism investigations.

### 2.3 | QM/MM calculations

In order to balance the computational cost and quality, cluster models were constructed to study the reaction energy profile for thioesterase NocTE via a two-layered QM/MM ONIOM scheme<sup>27,28</sup> implemented in Gaussian 09 program. The initial conformations originated from the most dominant cluster and pre-reaction states of MD simulations. In the cluster model, the residues within 5.00 Å of the substrate were reserved, containing the substrate (LSub or DSub), 25 residues and several water molecules near the catalytic center (Figure S1). The atom number for the cluster model was 444 in total. The backbone of all residues was frozen to preserve the original protein architecture. The QM region consisted of the substrate covalently bound to NocTE, the side chains of D1806 and H1901, the whole residue H1808 and 3 water molecules, 93 atoms in total.

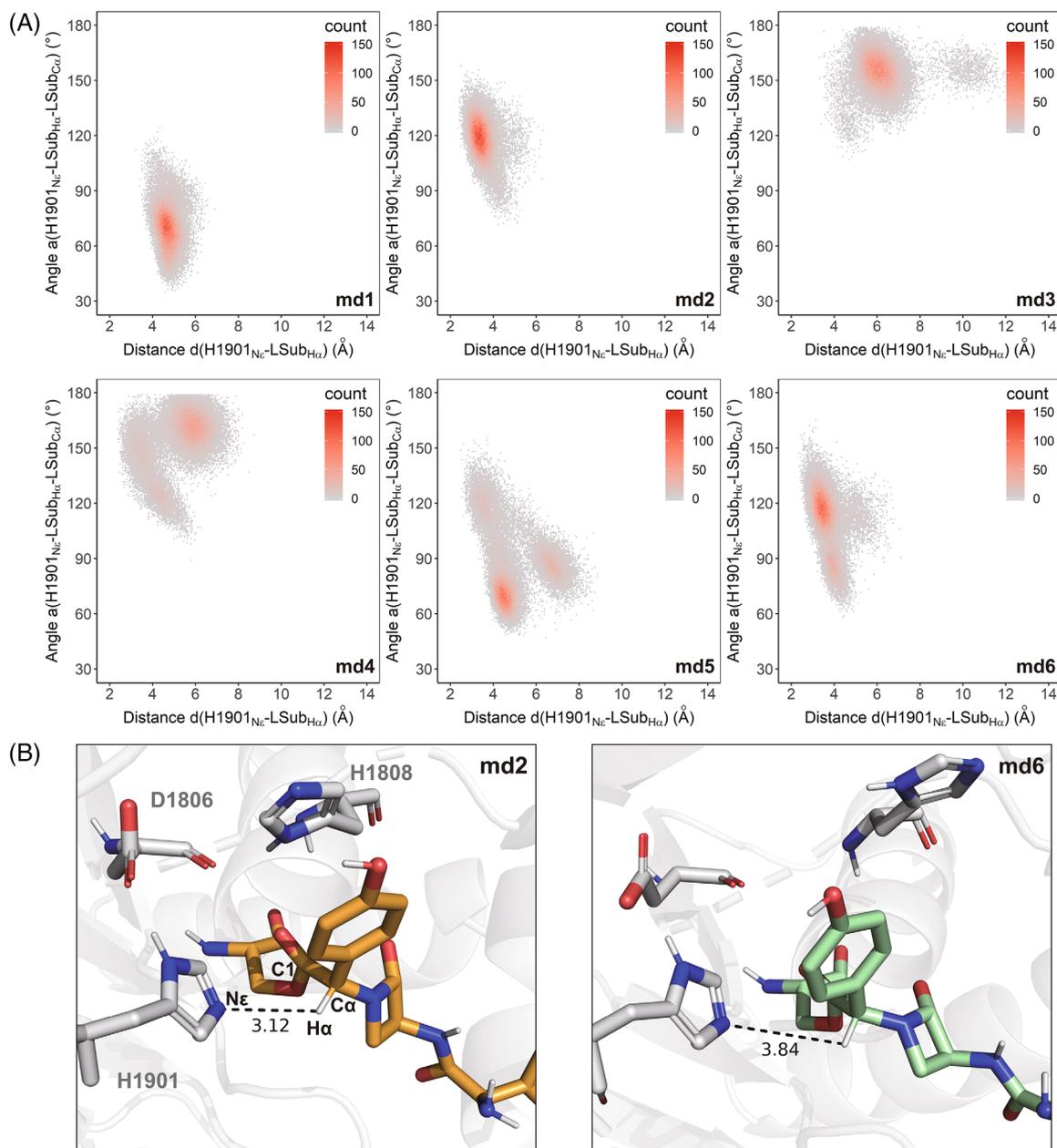
The geometry optimizations for transition states (TSs) and other minima, and intrinsic reaction coordinate (IRC) calculations were performed at the ONIOM (M062X/6-31G\*:Amber) level. Each transition state structure was confirmed with a sole imaginary frequency and reasonable movement tendency. After that, single point energy calculations for each minimum were carried out using 6-311+G\*\* basis set and solvation model based on density (SMD) solvation correction<sup>29</sup> to improve the accuracy of energy profiles. The interaction between QM and MM layers was treated with the electrostatic embedding formalism.<sup>30</sup>

## 3 | RESULTS AND DISCUSSION

### 3.1 | Deprotonated conformational distribution

Both pentapeptides and tripeptides containing the  $\beta$ -lactam ring can be catalyzed efficiently by NocTE, and more importantly, tripeptides have simpler structure and more detailed experimental observations.<sup>5</sup> Therefore, *epi*-nocardicin G covalently bound to NocTE (NocTE-LSub) system was constructed to explore the catalytic mechanism of NocTE. According to the hypothesis from Andrew M. Gulick's studies,<sup>6</sup> the proton acceptor H1901 may be involved in not only hydrolysis but also epimerization. Considering the conformation difference of surrounding residues between the epimerization reaction and hydrolysis reaction, the deprotonation distance  $d(\text{H1901}_{\text{N}_\epsilon} - \text{LSub}_{\text{H}_\alpha})$  was restrained between 3.00 and 3.50 Å to gain an appropriate initial structure, and then six replicas of 100 ns MD simulations without this restraint were performed to capture deprotonation conformations for epimerization (Figure S2).

For proton transfer step, reasonable deprotonation distances and angles are of great importance, so the reaction distance  $d(\text{H1901}_{\text{N}_\epsilon} - \text{LSub}_{\text{H}_\alpha})$  and angle  $\alpha(\text{H1901}_{\text{N}_\epsilon} - \text{LSub}_{\text{H}_\alpha} - \text{LSub}_{\text{C}_\alpha})$  were analyzed (Figure 2). Among these MD simulation trajectories, only md2 and md6 exhibited abundant proper conformations for deprotonation. For these two trajectories, the averages of the distance  $d(\text{H1901}_{\text{N}_\epsilon} - \text{LSub}_{\text{H}_\alpha})$  were 3.45 and 3.65 Å, while above 4.50 Å in other trajectories. Besides, the angle  $\alpha(\text{H1901}_{\text{N}_\epsilon} - \text{LSub}_{\text{H}_\alpha} - \text{LSub}_{\text{C}_\alpha})$  was mainly concentrated around 120° that is beneficial for the proton



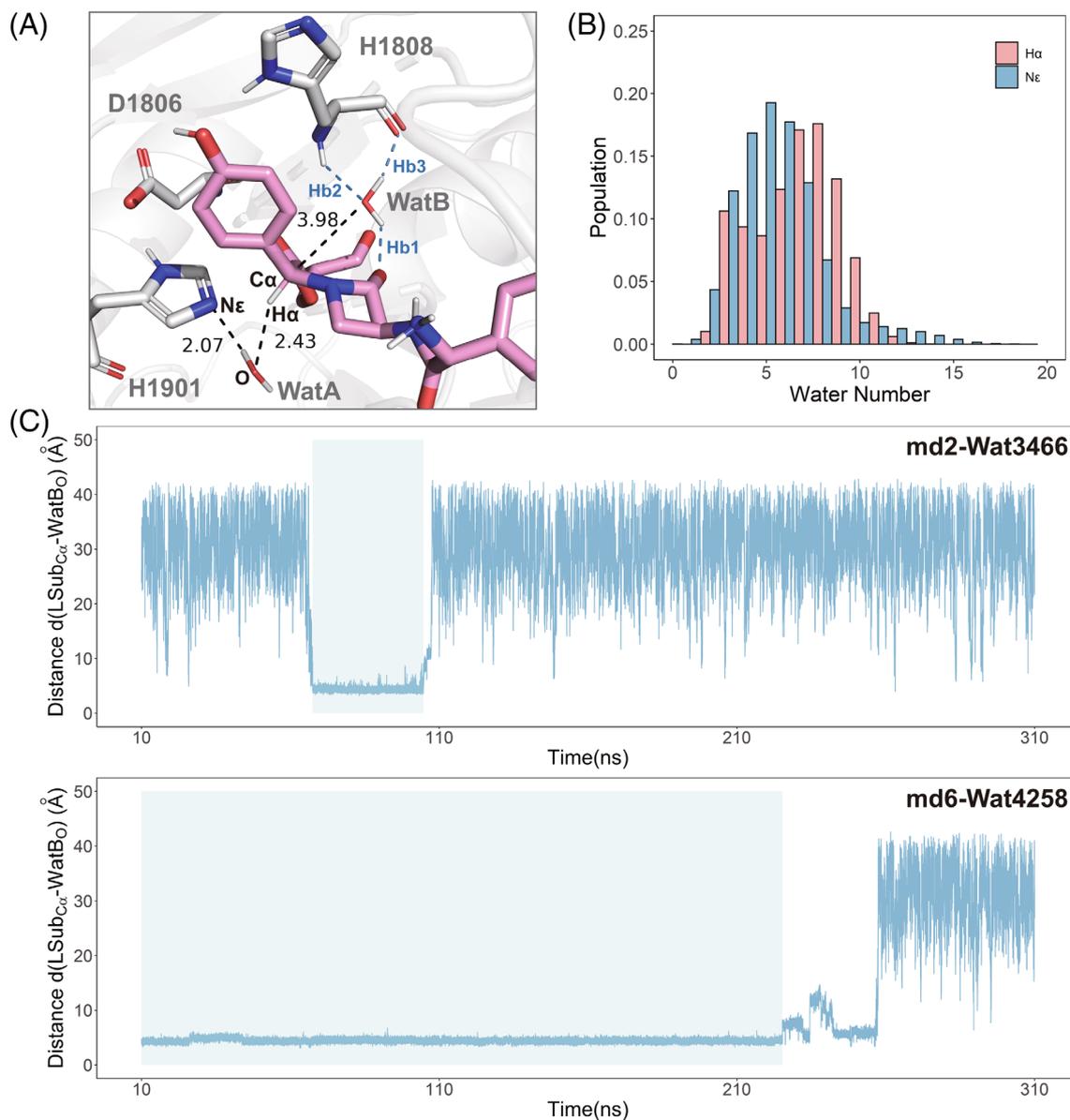
**FIGURE 2** The deprotonation distance and angle in the epimerization reaction catalyzed by NocTE. (A) The distributions of the distance  $d$  ( $\text{H1901}_{\text{Ne}} - \text{LSub}_{\text{H}\alpha}$ ) and the angle  $a$  ( $\text{H1901}_{\text{Ne}} - \text{LSub}_{\text{H}\alpha} - \text{LSub}_{\text{C}\alpha}$ ) in six MD simulation trajectories; (B) the most representative structures for md2 and md6 MD simulation trajectories

transfer. Hence, md2 and md6 trajectories were rich in pre-reaction conformations and selected to be lengthened to 310 ns for further study (Figure S3).

### 3.2 | Two key water molecules in epimerization reaction

To explore the epimerization mechanism of NocTE, structural clustering was carried out based on the coordinates of the backbone atoms originating from all residues and LSub (Table S1).

Amazingly, two key water molecules were found in the representative structure of the most dominant cluster from md2 and md6 MD simulation trajectories (Figure 3A). One of the water molecules WatA stayed between the  $\text{N}_{\epsilon}$  atom of H1901 and the  $\text{H}_{\alpha}$  atom of LSub. Because the distance  $d(\text{H1901}_{\text{Ne}} - \text{WatA}_{\text{O}})$  and  $d(\text{LSub}_{\text{H}\alpha} - \text{WatA}_{\text{O}})$  was small enough, two possible deprotonation paths were proposed, including direct deprotonation by H1901 and indirect deprotonation mediated by WatA. In order to confirm the feasibility of indirect deprotonation path, the total number of water molecules near the reaction site was counted (Figure 3B). The average numbers of water molecules within 5.00 Å of  $\text{N}_{\epsilon}$  and  $\text{H}_{\alpha}$  atoms

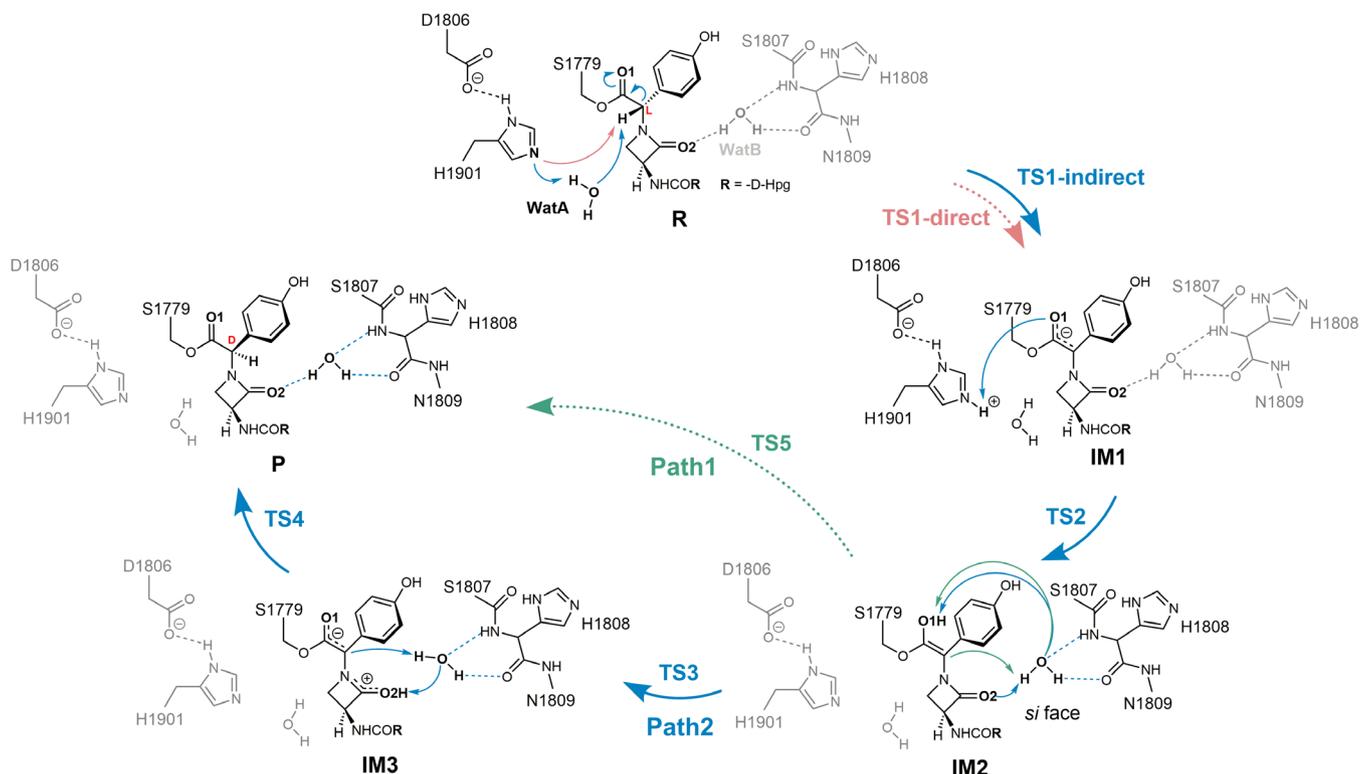


**FIGURE 3** Two key water molecules obtained by structural clustering analyses. (A) The representative structure of the most dominant cluster from md2 and md6 MD simulation trajectories; (B) the number distributions of water molecules within 5.00 Å of the reaction site H1901<sub>Ne</sub> and LSub<sub>Hα</sub>, respectively; (C) the distributions of the distances between the C $\alpha$  atom of LSub and the O atom of the water molecules (Wat3466 and Wat4258) in md2 and md6 trajectories, respectively. The light blue-shaded area indicated the water molecule stayed within 5.00 Å of the C $\alpha$  atom for a period of time

were 5.70 and 6.70, respectively, indicating there were abundant water molecules around the active site for indirect deprotonation. In addition, two water molecules Wat380 and Wat5671 were selected randomly from md2 and md6 trajectories to verify they could retain between H1901<sub>Ne</sub> and LSub<sub>Hα</sub> for a certain time in 4 ns simulations (Figure S4). Apart from two key water molecules, the other water molecules around the active site were responsible for conformation stabilization and ready to replace WatA to mediate the proton transfer as candidates.

In the opposite direction of the deprotonation site, the water molecule WatB was 3.98 Å away from the C $\alpha$  atom of LSub, and formed hydrogen bonds (Hb1, Hb2, and Hb3) with the backbone of

H1808 and the  $\beta$ -lactam ring of LSub. Owing to no other residue as the proton donor, the water molecule WatB was thought to provide the proton for the C $\alpha$  atom from *si* face, finishing the L to D stereochemical inversion of the C-terminal Hpg. Similarly, two water molecules Wat3466 and Wat4258 were selected randomly. The distances between the C $\alpha$  atom of LSub and the O atom of these water molecules and the populations of three important hydrogen bonds were analyzed (Figure 3C, Table S2). As these data shown, Wat3466 and Wat4258 could stayed within 5.00 Å of the C $\alpha$  atom for a long time through these hydrogen bond interactions, getting ready for the subsequent re-protonation step. These similar structural characteristics concerning key water molecules were likewise observed in the



**FIGURE 4** The proposed molecular mechanism for epimerization reaction catalyzed by NocTE. The direct deprotonation path and re-protonation path Path1 were colored in pink and green. The abbreviations R, TS, IM, and P represented the reactant, transition state, intermediate, and product

pentapeptide covalently bound to NocTE system (Figures S5–S7 and Table S3).

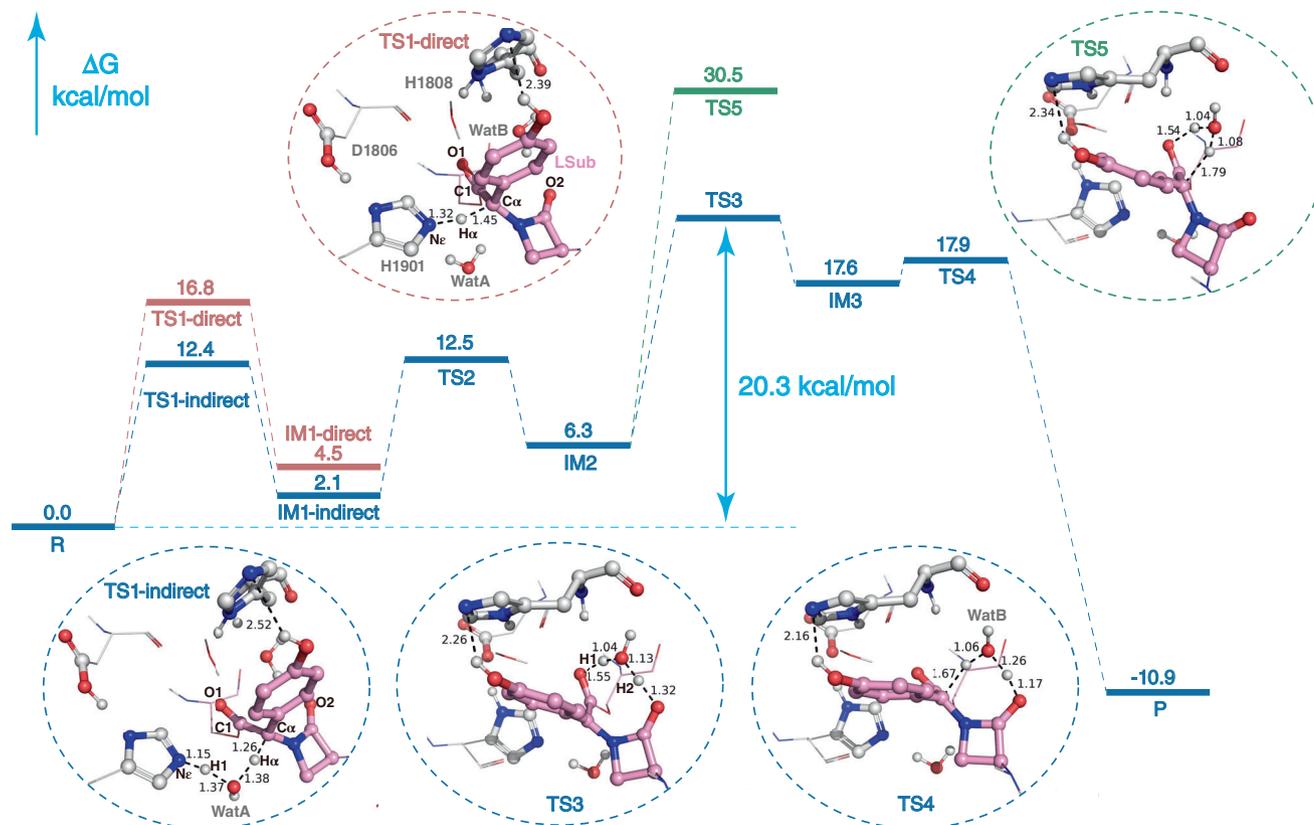
### 3.3 | The mechanism of epimerization reaction

Distinguishable from other epimerases, the main product pool generated by thioesterase NocTE is D-configured peptides (nocardicin G) and no obvious proton donor exists around the active site for epimerization. According to the structural characteristics abstracted from MD simulations, two possible reaction paths were put forward for deprotonation and re-protonation steps, respectively. For the deprotonation step, the proton acceptor H1901 can extract the proton  $H_{\alpha}$  from LSub directly (direct path), or the proton is transferred via the water molecule WatA from LSub to H1901 (indirect path; Figure 4). Then, this proton moves from H1901 to the *si* face with the assistance of the O1 atom of LSub, prepared for the following step. For the re-protonation step, we come up with two possible reaction paths, Path1 and Path2. In Path1, the proton is transferred from LSub to WatB in the *si* face and the  $C_{\alpha}$  atom obtains a proton from WatB simultaneously, generating the epimerization product DSub. In Path2, the  $\beta$ -lactam ring facilitates the proton transfer. The proton transfer from O1 atom to the O2 atom of  $\beta$ -lactam ring is mediated by WatB first, and then this proton is transferred via WatB from the  $\beta$ -lactam ring toward the  $C_{\alpha}$  atom again.

To shed light on the detailed mechanism of this special epimerization, QM/MM calculations were performed using the cluster model on the basis of three structure replicas from the dominant cluster of MD simulations. Geometry optimizations, TS search and IRC calculations were carried out at the ONIOM (M062X/6-31G\*: Amber) level. Then, single point energies were calculated with larger basis set 6-311+G\*\* and SMD solvation model to obtain more accurate energy profile. The free energy profile and important TS structures from the representative replica were displayed in Figure 5. Other minimum structures and two replicas were provided in Figure S8, Tables S4 and S5.

In the deprotonation step, LSub provided the proton  $H_{\alpha}$  for H1901 straightway with the energy barrier of 16.8 kcal/mol, while the energy barrier for the indirect path with the assistance of the water molecule WatA was 12.4 and 4.4 kcal/mol lower than the former. Combining abundant water molecules observed in MD simulations and energy barrier superiority, the indirect path was considered as the optimal reaction path for the deprotonation. After that, the proton accepted by H1901 was transferred to the O1 atom of LSub to form an enol intermediate (IM2), facilitating the proton movement toward the *si* face. This process required 10.4 kcal/mol relative to the IM1.

In the re-protonation step, the  $\beta$ -lactam ring played a crucial role in the proton transfer. WatB was anchored around the reaction site by the backbone of H1808 and  $\beta$ -lactam ring in advance. In IM2,



**FIGURE 5** The free energy profile and optimized structures of important transition states (TSs) of the epimerization reaction catalyzed by thioesterase NocTE. The direct deprotonation path and re-protonation path Path1 were colored in pink and green, respectively, while the optimal reaction pathway was shown in blue. The important residues (D1806, H1808, and H1901) and LSub were colored in gray and pink, respectively

WatB formed hydrogen bonds with the  $\beta$ -lactam ring and the enol group of LSub (Figure S8), implying the existence of Path1 and Path2. For Path1, the proton deposited on the LSub was transferred to WatB, synchronizing with the proton transfer from WatB to the C $\alpha$  atom of LSub, inverting the stereochemistry of the C-terminal Hpg. The energy barrier for this path was 24.2 kcal/mol relative to IM2, suggesting that Path1 was unfavorable for re-protonation due to the prohibitively high barrier. For Path2, the proton transfer from the O1 atom to the O2 atom of the  $\beta$ -lactam ring took place, followed by the proton transfer from the  $\beta$ -lactam ring to the atom C $\alpha$ , which were both mediated by WatB. The former as the rate-determining step needed 14.0 kcal/mol relative to IM2, and the energy barrier for the latter was 0.3 kcal/mol relative to IM3. Although Path2 seemed more complicated than Path1, the remarkable energy superiority made Path2 more suitable for the re-protonation.

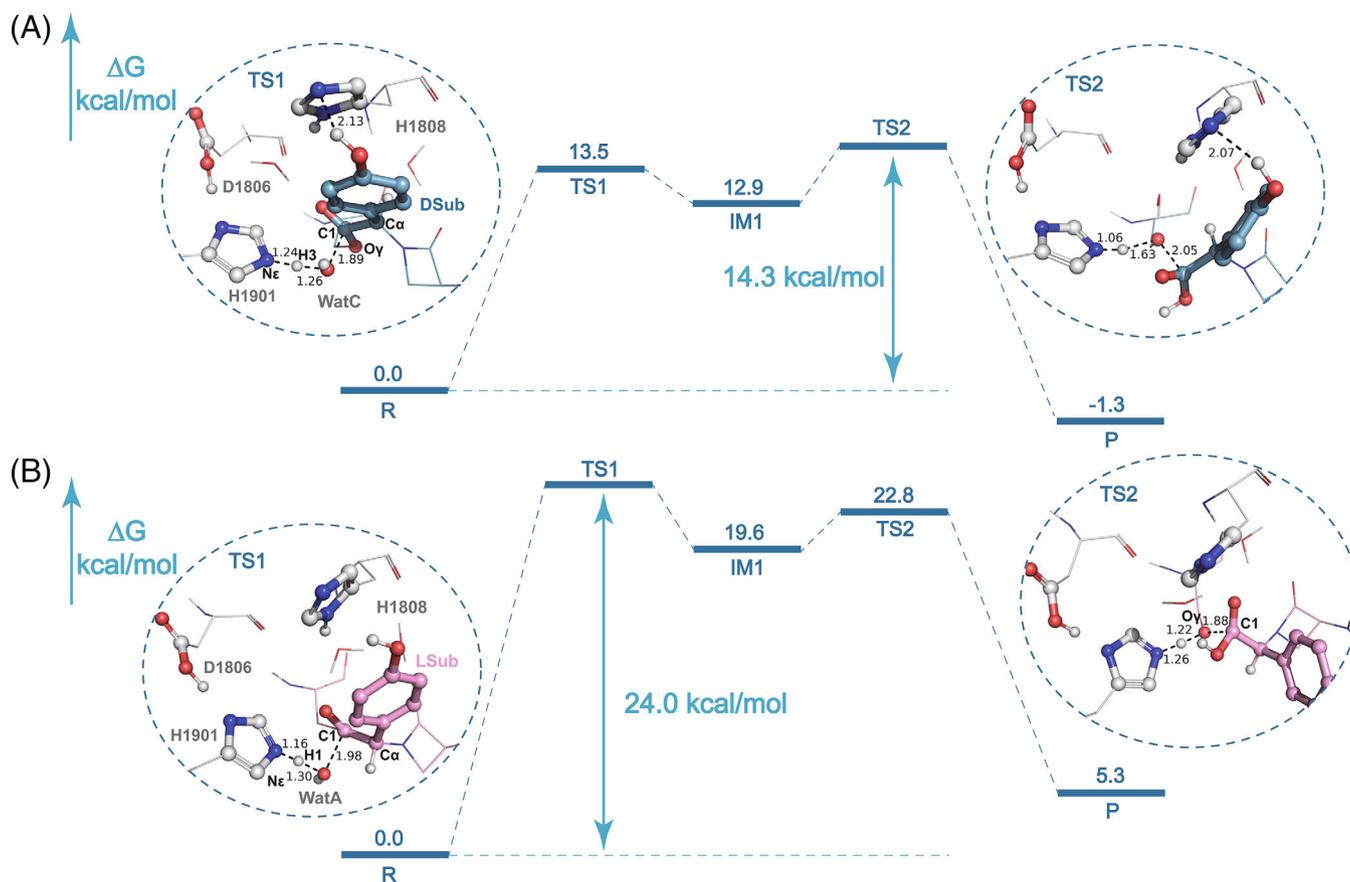
From the view of structural characteristics, the O2 atom of the  $\beta$ -lactam ring got closer to the proton H2 of WatB than the C $\alpha$  atom in IM2. The C-terminal Hpg and WatB formed a tortuous six-membered structure in the transition state TS5 for Path1, resulting from the steric hindrance of the hydroxyphenyl group and the vertical plane conformation formed by the carbonyl group of the C-terminal Hpg. The O1 atom and the C $\alpha$  atom stayed too close, leading to not enough space for two proton transfers. Though the similar structure

appeared in the transition state TS4 for Path2, the  $\beta$ -lactam ring facilitated the formation of more reasonable horizontal plane for electron delocalization, lowering the energy barrier. Besides, the O2 atom of the  $\beta$ -lactam ring that served as the proton donor got away from the C $\alpha$  atom, making room for proton transfers. In addition, the imidazole ring of H1808 was found to form hydrogen bonds and  $\pi$ - $\pi$  stacking interactions with the C-terminal Hpg of LSub to maintain the substrate conformation during the overall reaction, partially elucidating the adverse effect of H1808 mutation on the diastereomer purity of products.

In brief, the indirect protonation step and the substrate-assisted re-protonation step (Path2) composed the optimal reaction pathway for the epimerization catalyzed by NocTE with the total energy barrier of 20.3 kcal/mol, in harmony with experiment results.<sup>5</sup> The water molecules involved in epimerization were a hint that D-configured product was hard inverted into L-configured products once again, resulting in the stereochemical selectivity of NocTE.

### 3.4 | The mechanism of hydrolysis reaction

After the epimerization, the substrate with the D-Hpg in the C-terminal (DSub) should be released via hydrolysis reaction. The system



**FIGURE 6** The free energy profile and optimized structures of important transition states (TSs) for hydrolysis reaction in NocTE-DSub (A) and NocTE-LSub (B) systems. The important residues (D1806, H1808, and H1901), DSub and LSub were colored in gray, blue, and pink, respectively

of nocardicin G covalently bound to NocTE (NocTE-DSub) was modeled to show the energy profile of hydrolysis. The starting conformation for QM/MM calculation was the representative structure from the pre-reaction states in six replicas of 50 ns MD simulation trajectories (Figure S9). The free energy profile and key minimum structures for hydrolysis were shown in Figure 6A, Figure S10 and Table S6. Some important distances involved in the reaction were listed in Table S7.

As we can see, H1808 still helped DSub to maintain the proper hydrolytic conformation through hydrogen bonding and  $\pi$ - $\pi$  stacking interactions. First, H1901 abstracted a proton from the water molecule WatC to promote the nucleophilic attack on the C1 atom, generating a tetrahedral intermediate IM. The energy barrier for this step was 13.5 kcal/mol relative to the reactant R. Subsequently, the O $\gamma$ -C1 bond broke and H1901 provided the proton for S1779, resulting in the product generation and enzymatic environment recovery for the next catalysis cycle. The energy barrier for the whole hydrolysis process was 14.3 kcal/mol, basically consistent with the fact that the rate of hydrolysis was faster than that of epimerization.<sup>8</sup>

The special spatial location of WatA revealed the competition between epimerization and hydrolysis in NocTE-LSub system. WatA was not only able to mediate the proton transfer between H1901 and the C $\alpha$  atom, but also had the potential to directly participate in the

hydrolysis of LSub. When H1901 captured the proton H1 of WatA, the generated nucleophile could choose to reacquire a proton from C $\alpha$  atom to complete the deprotonation step, or attack the C1 atom to form the hydrolytic intermediate. However, the distance between the O atom of WatA and the H $\alpha$  atom (2.42 Å) was closer than that for the C1 atom (3.29 Å), so WatA is more advantageous in acquiring the proton H $\alpha$  (Figure S11). In addition, the Hpg residue in the C-terminal of LSub caused certain steric hindrance and the unreasonable orientation in the hydrolysis process prevented WatA from forming a stable hydrogen bond with H1808, unbeneficial for maintaining low-energy conformations. These structural disadvantages seriously hindered WatA from getting close to C1 atom directly, making it difficult to promote the hydrolysis reaction.

Aimed to assess the difficulty of direct hydrolysis of LSub, we obtained the energy profile of hydrolysis reaction starting from the initial conformation for previous epimerization study (Figure 6B and Table S8). Different from the energy profile of DSub hydrolysis, the rate-limiting step for LSub hydrolysis was the nucleophilic attack step of WatA toward the C1 atom rather than the step of breaking the covalent bond between S1779 and substrate, which could be well explained by structural disadvantages mentioned above. In addition, the energy barrier for LSub hydrolysis (24.0 kcal/mol) was 11.6 kcal/mol higher than that for the deprotonation step of epimerization

(12.4 kcal/mol), indicating the early-stage epimerization was essential for NocTE catalysis and further accounting for stereochemical selectivity.

## 4 | CONCLUSION

In this work, we presented detailed descriptions of catalytic mechanism for the thioesterase NocTE, a promising insight into substrate-assisted epimerization with stereochemical selectivity for directed drug design. Through restrained MD simulations, we sampled enough pre-reaction conformations for epimerization reaction. From the main structural cluster, two water molecules were inferred to promote the proton transfer. Then, a substrate-assisted molecular mechanism for epimerization activity was explored and confirmed by QM/MM calculations. The overall free energy barrier of the epimerization mediated by water molecules and  $\beta$ -lactam-ring was obtained to be 20.3 kcal/mol and the energy barrier of the subsequent hydrolysis was calculated to be 14.3 kcal/mol. In addition, we assessed the hydrolysis difficulty of L-configured substrate without epimerization (LSub) and found the energy profile for the nucleophilic attack of water molecules toward LSub was up to 24.0 kcal/mol. The remarkable barrier disparity for DSub and LSub hydrolysis provided more clues to the stereochemical selectivity of NocTE catalysis, in addition to the comparison of product release in our previous studies.<sup>10</sup>

Two water molecules became the key to promote epimerization, contributing to the preference reflected in the stereochemical inversion. The pathway choice of water molecule WatA attached importance upon the epimerization activity for NocTE catalysis. During the deprotonation step, the direct hydrolysis of LSub was hindered because the hydrolytic site C1 atom was relatively far from WatA in comparison with the epimerization site C $\alpha$  atom, preventing WatA from approaching C1 atom to participate in hydrolysis. Therefore, after H1901 abstracted a proton from WatA, the proton H $\alpha$  from the C $\alpha$  atom could be easily caught by the deprotonated WatA, resulting in the regeneration of WatA and the deprotonation of LSub. Another water molecule WatB was involved in the proton cyclization for overall reaction, making the D to L stereochemical inversion more difficult. WatB was found to stay around the reaction site C $\alpha$  atom by forming three stable hydrogen bonds with H1808 and the  $\beta$ -lactam ring of LSub for re-protonation. With the assistance of  $\beta$ -lactam ring, WatB delivered a proton to the C $\alpha$  atom, completing the L to D stereochemical inversion of the C-terminal Hpg.

The role of  $\beta$ -lactam ring was also displayed in the catalytic process. In deprotonation step, the  $\beta$ -lactam ring of LSub helped to maintain the water molecule WatB around the catalytic center by forming a strong hydrogen bond. Besides, the  $\beta$ -lactam ring directly took part in the proton transfer in the following re-protonation step and its coplanar structure was beneficial for electron delocalization to lower the energy barrier, suggesting the significance of the formation of  $\beta$ -lactam ring in advance distinguishable from that in the biosynthesis of isopenicillin N.<sup>31</sup> Another key residue that should be noticed was H1808. H1808 formed the hydrogen bond and  $\pi$ - $\pi$  stacking

interactions with the hydroxyphenyl group of the C-terminal Hpg to stable the substrate conformation, revealing the importance of H1808 in the control of diastereomeric purity.

Our findings draw a clear picture of the stereochemical selectivity for bifunctional thioesterase NocTE and provide a reasonable molecular mechanism as an example of substrate-assisted catalysis. These results contribute to the diversification of the thioesterase catalytic activities and encourage the stereochemical control for peptide drugs through rational designs.

## AUTHOR CONTRIBUTIONS

Ting Shi, Wei Lu, and Daixi Li conceived and designed the investigation. Qian Yu performed calculations and analyses. Ting Shi and Qian Yu wrote up the paper.

## ACKNOWLEDGMENT

The authors thank the National Key R&D Program of China (2019YFA0905400, 2021YFC2100600), the National Science Foundation of China (32070041), and acknowledge 19ZR1427300, SL2020MS024, SJTU-ZH2018ZDA26, and the SJTU-HPC computing facility award for financial support and computational time.

## CONFLICT OF INTEREST

The authors declare no competing financial interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

Ting Shi  <https://orcid.org/0000-0003-3921-4412>

## REFERENCES

1. Sussmuth RD, Mainz A. Nonribosomal peptide synthesis-principles and prospects. *Angew Chem Int Ed Engl*. 2017;56(14):3770-3821.
2. Muttenthaler M, King GF, Adams DJ, Alewood PF. Trends in peptide drug discovery. *Nat Rev Drug Discov*. 2021;20:309-325.
3. Feng Z, Xu B. Inspiration from the mirror: D-amino acid containing peptides in biomedical approaches. *Biomol Concepts*. 2016;7:179-187.
4. Ogasawara Y, Dairi T. Peptide epimerization machineries found in microorganisms. *Front Microbiol*. 2018;9:156.
5. Gaudelli NM, Townsend CA. Epimerization and substrate gating by a TE domain in beta-lactam antibiotic biosynthesis. *Nat Chem Biol*. 2014;10(4):251-258.
6. Patel KD, d'Andrea FB, Gaudelli NM, Buller AR, Townsend CA, Gulick AM. Structure of a bound peptide phosphonate reveals the mechanism of nocardicin bifunctional thioesterase epimerase-hydrolyase half-reactions. *Nat Commun*. 2019;10(1):3868.
7. Gunsior M, Breazeale SD, Lind AJ, Ravel J, Janc JW, Townsend CA. The biosynthetic gene cluster for a monocyclic beta-lactam antibiotic, nocardicin A. *Chem Biol*. 2004;11(7):927-938.
8. Aoki H, Sakai H, Kohsaka M, Konomi T, Hosoda J. Nocardicin A, a new monocyclic beta-lactam antibiotic. I. Discovery, isolation and characterization. *J Antibiot (Tokyo)*. 1976;29(5):492-500.
9. Nishida M, Mine Y, Nonoyama S, Kojo H. Nocardicin A, a new monocyclic beta-lactam antibiotic III. In vitro evaluation. *J Antibiot (Tokyo)*. 1977;30(11):917-925.

10. Yu Q, Xie LF, Li YL, et al. Exploring the molecular basis of substrate and product selectivities of nocardicin bifunctional thioesterase. *Interdiscip Sci.* 2022;14(1):233-244.
11. Tanner ME. Understanding nature's strategies for enzyme-catalyzed racemization and epimerization. *Acc Chem Res.* 2002;35(4):237-246.
12. Liu B, Hou Y, Wang X, et al. Structural basis of the mechanism of beta-methyl epimerization by enzyme MarH. *Org Biomol Chem.* 2019;17(44):9605-9614.
13. Zhang YL, Zheng QC, Zhang JL, Zhang HX. Insights into the epimerization activities of RaCE and pAGE: the quantum mechanics/molecular mechanics simulations. *RSC Adv.* 2015;5(124):102284-102293.
14. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 2004;32:W665-W667.
15. Zhao Y, Truhlar DG. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor Chem Acc.* 2008;120(1-3):215-241.
16. Rassolov VA, Ratner MA, Pople JA, Redfern PC, Curtiss LA. 6-31G\* basis set for third-row atoms. *J Comput Chem.* 2001;22(9):976-984.
17. Payne PW. The hartree-fock theory of local regions in molecules. *J Am Chem Soc.* 1977;100(11):7742-7743.
18. Frisch MJ, Trucks GW, Schlegel HB, et al. *Gaussian 09*. Gaussian, Inc; 2009.
19. Cornell WD, Cieplak P, Bayly CI, et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc.* 1995;117(19):5179-5197.
20. Lu T, Chen F. Multiwfn: a multifunctional wavefunction analyzer. *J Comput Chem.* 2012;33(5):580-592.
21. Wang JM, Wang W, Kollman PA. Antechamber: an accessory software package for molecular mechanical calculations. *Abstr Pap Am Chem Soc.* 2001;222:U403.
22. Case DA, Ben-Shalom IY, Rozell SR, et al. *AMBER 2018*. University of California; 2018.
23. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput.* 2015;11(8):3696-3713.
24. Darden T, York D, Pedersen L. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J Chem Phys.* 1993;98(12):10089-10092.
25. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys.* 1977;23(3):327-341.
26. Roe DR, Cheatham TE. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput.* 2013;9(7):3084-3095.
27. Vreven T, Frisch MJ, Kudin KN, Schlegel HB, Morokuma K. Geometry optimization with QM/MM methods II: explicit quadratic coupling. *Mol Phys.* 2006;104(5-7):701-714.
28. Vreven T, Byun KS, Komáromi I, et al. Combining quantum mechanics methods with molecular mechanics methods in ONIOM. *J Chem Theory Comput.* 2006;2(3):815-826.
29. Marenich AV, Cramer CJ, Truhlar DG. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J Phys Chem B.* 2009;113(18):6378-6396.
30. Cisneros GA, Piquemal JP, Darden TA. Quantum mechanics/molecular mechanics electrostatic embedding with continuous and discrete functions. *J Phys Chem B.* 2006;110(28):13682-13684.
31. Roach PL, Clifton IJ, Hensgens CM, et al. Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature.* 1997;387(6635):827-830.

#### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Yu Q, Lu W, Li D, Shi T. Insight into substrate-assisted catalytic mechanism and stereoselectivity of bifunctional nocardicin thioesterase. *Proteins.* 2022;1-10. doi:10.1002/prot.26395