

Theoretical Studies on the Catalytic Mechanism and Substrate Diversity for Macrocyclization of Pikromycin Thioesterase

Ting Shi, Lanxuan Liu, Wentao Tao, Shenggan Luo, Shuobing Fan, Xiao-Lei Wang, Linquan Bai, and Yi-Lei Zhao*®

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 200240, People's Republic of China

Supporting Information

ABSTRACT: Polyketide synthases (PKSs) share a subset of biosynthetic steps in construction of a polyketide, and the offload from the PKS main module of specific product release is most often catalyzed by a thioesterase (TE). In spite of the fact that various PKS systems have been discovered in polyketide biosynthesis, the molecular basis of TE-catalyzed macrocyclization remains challenging. In this study, MD simulations and QM/MM methods were combined to investigate the catalytic mechanism and substrate diversity of pikromycin (PIK) TE with two systems (PIK-TE-1 and PIK-TE-2), where substrates 1 and 2 correspond to TE-catalyzed precursors of 10-deoxymethynolide and narbonolide, respectively. The results showed that, in comparison with PIK-TE-2, system PIK-TE-1 exhibited a greater tendency to form a stable prereaction state, which is critical to macrocyclization. In addition, the structural



characteristics of prereaction states were uncovered through analyses of hydrogen-bonding and hydrophobic interactions, which were found to play a key role in substrate recognition and product release. Furthermore, potential energy surfaces were calculated to study the molecular mechanism of macrocyclization, including the formation of tetrahedral intermediates from *re*- and *si*-face nucleophilic attacks and the release of products. The energy barrier of macrocyclization from *re*-face attack was calculated to be 16.3 kcal/mol in PIK-TE-1, 3.6 kcal/mol lower than that from *si*-face attack and 4.1 kcal/mol lower than that from *re*-face attack in PIK-TE-2. These results are in agreement with experimental observations that the yield of 10-deoxymethynolide is superior to that of narbonolide in PIK TE catalyzed macrocyclization. Our findings elucidate the catalytic mechanism of PIK TE and provide a better understanding of type I PKS TEs in protein engineering.

KEYWORDS: thioesterase, macrocyclization, catalytic mechanism, substrate diversity, prereaction state

INTRODUCTION

Polyketides and analogues have been developed as macrolide antibiotics, antifungal agents, immunosuppressants, and anticancer agents,^{1–5} owing to their significant medicinal properties. They are efficiently assembled by type I polyketide synthase (PKS),⁶ which are comprised of multiple requisite domains (i.e. "modules"). Each module is responsible for a particular function in the construction of polyketide scaffolds, including initiation, elongation, reduction, and extension. Finally, when the polyketide chain reaches the desired size, the product can be offloaded from the PKSs with the assistance of a thioesterase (TE) via hydrolysis or macrocyclization.^{7–10}

Pikromycin, a naturally occurring ketolide antibiotic, is biosynthesized by a type I PKS in *Streptomyces venezuelae* ATCC 15439.¹¹ Unlike other type I PKS systems, the pikromycin PKS demonstrates the unique ability to efficiently generate two types of macrolactone products: methymycin and pikromycin. As we all know, pikromycin thioesterase (PIK TE) could serve to cyclize both linear hexaketide and heptaketide chain elongation intermediates toward the 12- and 14membered macrolactones 10-deoxymethynolide and narbonolide, respectively (Figure 1). Further processing of 10deoxymethynolide and narbonolide by post-PKS tailoring enzymes, including a glycosyltransferase and a cytochrome P450 hydroxylase, completes the biosynthesis of methymycin and pikromycin.

PIK TE adopts an α , β -hydrolase fold, and it contains a central seven-stranded β -sheet connected by α -helices, with $\beta 2$ antiparallel to the remaining strands. Similar to the substratebinding region of 6-deoxyerythronolide B synthase (DEBS) TE, the lid region is composed of two helices (L1 and L2) and a lid loop (Figure 1). Crystal structures of the two homologous enzymes show many similarities, including a hydrophobic interface for dimerization, an open substrate channel, and a characteristic Ser-His-Asp catalytic triad (Ser148, His268, and Asp176 in PIK TE and Ser142, His259, and Asp169 in DEBS

Received:March 22, 2018Revised:April 10, 2018Published:April 10, 2018



Figure 1. Stucture of PIK TE and macrocyclization catalyzed by PIK TE with substrates 1 and 2.

TE).¹² The triad is located at the center of a long open substrate channel.¹³ In vivo, the polyketide substrate is initially bound to the phosphopantetheinyl arm of an acyl carrier protein (ACP) and undergoes transesterification to the active site Ser148 of PIK TE. Intramolecular attack by the distal hydroxyl group of the substrate on the acyl-enzyme intermediate affords the macrolactone, with simultaneous release of the product from the enzyme. Although the biochemical function of these PKSs has been unveiled,^{14,15} a detailed description of how protein dynamics determines their biological role at the atomic and molecular level remains a scientific challenge.

As we know, PKS TEs have been shown to be highly specific in catalytic processes. Both acyl-enzyme intermediate formation and macrocyclization are substrate specific. For example, the DEBS TE appears to require a carbonyl function group connected to the carbon atom of the substrate for the nucleophilic cyclization. Replacement of the carbonyl group with a hydroxyl group results in the exclusive occurrence of hydrolysis.^{16,17} On the other hand, DEBS TE demonstrates broad tolerance for the ring size of the polyketide generated. In more detail, although the natural product of DEBS TE is a 14membered ring, this enzyme has been shown to support the formation of alternatively functionalized 6-, 8-, 12-, 14-, and 16membered-ring systems. This study was expected to provide an understanding of the mechanism and substrate selectivity of these type I PKS TEs, which could present new opportunities for chemoenzymatic synthesis of polyketides and analogues, and thereby accelerate the development of new medicines in the future.

In this study, PIK TE is employed to develop a highresolution model detailing enzyme–substrate interactions to rationalize macrocyclization of 10-deoxymethonilide and narbonylide. We combined molecular dynamics (MD) simulations and quantum mechanics/molecular mechanics (QM/MM) calculations on PIK-TE-1 and PIK-TE-2 systems in an aqueous environment. In addition, we report the formation of a substrate–enzyme prereaction state, which is critial to mutual recognition and catalytic preparation and description of its structural characteristics with hydrogenbonding and hydrophobic interactions. Furthermore, two-layer ONIOM-based QM/MM calculations were employed to investigate the catalytic mechanism in detail. Transition states and key intermediates were located in the calculated potential energy surface. A possible molecular mechanism of the macrocyclization process was proposed. These results are in good agreement with biochemical and structural studies. Our study elucidates the catalytic mechanism of PIK-TE with its substrates, provides insight into the macrocyclization process of PIK-TE, and increases a better understanding of type I PKS TEs in an engineered PKS pathway.

MATERIALS AND METHODS

System Preparation. The crystal structure of PIK TE (PDB entry: 2HFK)¹⁵ was used as the starting structure in the preparation of MD simulations. Substrates 1 and 2 were both covalently bonded to residue Ser148 of PIK TE. Conformation of the substrates was determined through the following procedure. (1) We blocked the terminal of the substrates by adding an N-terminal cap $(-CO-CH_3)$ and a C-terminal cap $(-NH-CH_3)$. (2) We conducted a classical conformational search with a systematic search method (CAESAR)¹⁸ encoded in Discovery Studio 3.5, and 300 ligand conformations were generated. (3) The top 20 low-energy conformations were picked and further optimized with the PM3 method¹⁹⁻²¹ in Gaussian09.²² (4) The conformation with the lowest energy was then placed into the PIK TE catalytic pocket and connected with the protein through a covalent bond, and the cap was removed from the substrate. Following the steps above, two complex systems, PIK-TE-1 and PIK-TE-2, were set up. Files for complexes were prepared with the *tleap* module of AMBER 12.²³ After preparation of the complex structures, parameters for substrates 1 and 2 were generated. We first performed an conformational optimization with the Gaussian 09 program at the level of $HF/6-31G(d)^{24-26}$ and then computed its electrostatic surface potential (ESP) charge. A two-step restrained electrostatic potential (RESP)²⁷ chargefitting procedure was then carried out on the substrates. Finally, missing parameters such as bond and dihedral information were generated by the Antechamber package.

MD Simulation and Trajectory Analysis. MD simulations were performed on PIK-TE-1 and PIK-TE-2 complexes using the AMBER ff03.r1 force field. The proteins were solvated in a cubic box of TIP3P water molecules, with the



Figure 2. Prereaction states in PIK-TE-1 and PIK-TE-2. Substrate 1 is shown in salmon and 2 in magenta.

thickness of the external water layer exceeding 10 Å. Sodium ions were then added to achieve charge neutralization in the system.

Both solvated systems were subjected to a minimization and heating cycle, with the temperature gradually being raised to 298 K. After equilibration for 50 ps, a 50 ns molecular dynamics (MD) simulation was conducted on the complexes without any restraints under NPT conditions. Afterward, a simulation for each complex was repeated six times with a different random number, and one of them was extended to 300 ns. The Particle Mesh Ewald (PME) method²⁸ was employed to account for long-range electrostatic interactions, and the SHAKE algorithm in its matrix form was used to fix bonds and angles involving hydrogen atoms.²⁹ The cutoff for van der Waals interactions was set to 10.0 Å. This protocol was applied to all MD simulations for both systems using the parallel version of PMEMD.cuda in the AMBER12 suite.

QM and QM/MM Calculations. QM/MM calculations were started from one snapshot closest to the ensemble average of the most dominant cluster from MD trajectories. All of the QM calculations were performed with the M062x³⁰/6-31G(d) method. QM/MM calculations at the ONIOM (M062x/6-31G(d):Amber) level were performed using a two-layered ONIOM method^{31,32} encoded in the Gaussian09 program. We selected substrate 1 or 2, which contains the group $(-O_{ser}-CH_2-)$ of residue Ser148, the anionic carboxymethyl group $(-CH_2-COO^-)$ of Asp176, and the side chain of His268 as our QM region. A total of 74 atoms were included in the QM system of substrate 1 and 82 atoms in the system of substrate 2. This caused the QM level to bear 1 negative charge while the whole system had 17 negative charges, due to the removal of Na⁺ ions from the structure.

The QM region was described with density functional theory, and geometries were optimized with the M062x exchangecorrelation functional and 6-31G(d) basis set. Single-point energy calculations were performed on the basis of the optimized structures using larger basis sets, including 6-311+G(d), 6-311+G(d,p), and 6-311+G(2df,2p). The remainder of the system (MM region) was treated with the AMBER Parm99 force field. QM/MM calculations of PIK-TE-1 and PIK-TE-2 consisted of 8649 atoms and 8657 atoms, respectively. The electrostatic interactions between the QM and MM regions were calculated using the electronic embedding method, which provides a better description of the electrostatic interaction between the QM and MM regions and allows the QM wave function to be polarized.

RESULTS AND DISCUSSION

To understand the catalytic mechanism and substrate diversity of the macrocyclization process catalyzed by PIK TE, the two complex systems (PIK-TE-1 and PIK-TE-2) were each constructed six times in 50 ns molecular dynamic (MD) simulations, where substrates 1 and 2 covalently bonded with the residue Ser148 in the active site of PIK TE. After that, two additional 300 ns simulations of PIK-TE-1 and PIK-TE-2 were performed to investigate the structural characteristics of prereaction states. Furthermore, potential energy surfaces were calculated with the QM/MM method to study the molecular mechanism of macrocyclization including (1) formation of a tetrahedral intermediate from re- or si-face nucleophilic attack of the C1 carbonyl group and (2) release of the product 10-deoxymethynolide or narbonolide. Pivotal transition states were located, and the energy barriers were obtained. Finally, a detailed mechanism of PIK TE catalyzed macrocyclization was proposed.

Observation of More "Active" States in PIK-TE-1 in Comparison to PIK-TE-2. In our previous studies of DEBS TE,³³ we observed a prereaction state during MD simulations, which is not only critical to the precise recognition of enzyme– substrate but also significant in the macrocyclization process. In this special conformation, His259 and Asp169 in the catalytic triad constitute a proton transfer chain to facilitate the deprotonation of the distal hydroxyl group of the substrate and further promote the intramolecular nuclear attack to obtain a tetrahedral intermediate. According to our study, whether the enzyme–substrate could develop to the prereaction state seemed to be decisive of whether macrocyclization or hydrolysis occurred.

To identify the existence of prereaction states in PIK TE systems, six 50 ns MD simulations were performed in PIK-TE-1 and PIK-TE-2, respectively (Figure S1). A clustering algorithm was utilized to collect all typical substrate conformations on the basis of 12 simulation trajectories. Five representative conformations were observed to be prereaction states, including three in the PIK-TE-1 system and two in the PIK-TE-2 system (Figure 2). Furthermore, we used two distances, which were highlighted in the DEBS TE prereaction state, to evaluate the reliability of the prereaction state in PIK TE. One is defined by the distance between the N_e atom of His268 in the catalytic triad and the O atom of a distal hydroxyl group of the substrate, designated as $d(N_e-O11)$ in PIK-TE-1 and $d(N_e-O13)$ in PIK-TE-2, indicating the formation of hydrogen bonding and further deprotonation of the substrate hydroxyl group. The other distance is defined between the C1



Figure 3. Conformer populations with $d(N_{e}-O11)/d(N_{e}-O13)$ and d(C1-O11)/d(C1-O13) distances obtained from 6×50 ns MD simulations in PIK-TE-1 and PIK-TE-2. $d(N_{e}-O11)/d(N_{e}-O13) \leq 3.0$ Å and d(C1-O11)/d(C1-O13) distances ≤ 4.5 Å are highlighted in cyan and pink.



Figure 4. $d(N_e - O11)/d(N_e - O13)$ and d(C1 - O11)/d(C1 - O13) distances in 300 ns MD simulations in PIK-TE-1 and PIK-TE-2. The different distance proportions are indicated by circles.

and the O atom of the distal hydroxyl group, designated as d(C1-O11) in PIK-TE-1 and d(C1-O13) in PIK-TE-2, demonstrating the nucleophilic attack on the acyl-enzyme intermediate.

It has been shown by QM calculations^{33,34} that the prereaction state adopts a reactive conformation, where the former distance is around 3.0 Å and the latter is close to 4.5 Å. Therefore, conformations with restraints of both $d(N_{\varepsilon}-O11)/d(N_{\varepsilon}-O13) \leq 3.0$ Å and $d(C1-O11)/d(C1-O13) \leq 4.5$ Å will be an indication of a prereaction state. The two distances

were calculated from all conformations of MD simulations, and their frequency distributions were analyzed to evaluate the formation of a prereaction state in PIK-TE-1 and PIK-TE-2. In Figure 3, the average $d(N_e-O11)$ distance in all 6×50 ns simulations is 4.32 Å, less than 4.73 Å for $d(N_e-O13)$ in PIK-TE-2. On comparison of $d(N_e-O11)$ with $d(N_e-O13)$, it is found that the population of $d(N_e-O11) \leq 3.0$ Å in PIK-TE-1 is about 24.3%, larger than the 14.2% value of $d(N_e-O13)$ in PIK-TE-2. These results suggest that the formation of a hydrogen bond between the N_e atom of His268 and the O

atom of the distal hydroxyl group in PIK-TE-1 might be preferable to that in PIK-TE-2. On the other hand, although the average d(C1-O11) distance is 4.94 Å, similar to 5.03 Å of d(C1-O13), the population of distance $d(C1-O11) \leq 4.5$ Å in PIK-TE-1 increased by 10.3% in comparison with PIK-TE-2, suggesting the preference in formation of favorable conformations for macrocyclization in PIK-TE-1. Taken together, these results demonstrated that, in comparison to PIK-TE-2, PIK-TE-1 could generate more "active" conformations, which might be easily transformed to prereaction states.

To further uncover the stability of prereaction states, additional 300 ns MD simulations were performed separately on the basis of two 50 ns simulations of PIK-TE-1 and PIK-TE-2. Both $d(N_e - O_{11})/d(N_e - O_{13})$ and $d(C_{1} - O_{11})/d(C_{1} - O_{13})$ O13) were monitored (Figure 4). According to the 300 ns root-mean-square deviation (RMSD) values, the PIK-TE-1 and PIK-TE-2 systems reached their respective equilibriums after 10 ns simulations (Figure S2). To ensure the accuracy of our results, the first 10 ns simulations were not considered in our analysis. The MD trajectories revealed that the most prevalent conformations adopted prereaction states in either PIK-TE-1 or PIK-TE-2. To be precise, distances $d(N_e - O11)$ in almost half of the conformers (46.0%) are less than 3.0 Å in PIK-TE-1, while the proportion was reduced to 27.5% in PIK-TE-2, indicating more appropriate orientations of the distal hydroxyl group toward His268 of PIK TE. What is more, once the hydrogen bond had formed in PIK-TE-1, it maintained the bond for simulations of more than 90 ns, reflecting the stability of the hydrogen bond. On the other hand, although the average d(C1-O11) and d(C1-O13) distances were essentially in accord with each other (4.63 Å in PIK-TE-1 and 4.59 Å in PIK-TE-2), the proportion of $d(C1-O11) \le 4.5$ Å in PIK-TE-1 was 24.3%, larger than the 14.2% observed in PIK-TE-2. Remarkably, the proportions of "active" conformations with both $d(N_{\varepsilon}-O11) \leq 3.0$ Å and $d(C1-O11) \leq 4.5$ Å in PIK-TE-1 were 11.2%, approximately twice that observed in PIK-TE-2 (5.7%). In conclusion, it was observed again that, in comparison with PIK-TE-2, PIK-TE-1 was more favorable for the formation of a prereaction state, and once it emerged, the prereaction state would remain for a relatively long time in PIK-TE-1.

Structural Characteristics of Prereaction States. To further probe into the structural characteristics of prereaction states, we analyzed the trajectories of PIK-TE-1 and PIK-TE-2 systems. Hydrogen-bonding and hydrophobic interactions between PIK TE and substrates 1 and 2 were carefully investigated. More specifically, the critical hydrogen-bonding interactions between the Ne atom of His268 and the active distal hydroxyl group were observed in both PIK-TE-1 and PIK-TE-2 (Figure 5). Additionally, a hydrogen bond between the C9 carbonyl and the side chain of Thr77, which is located in the loop connecting strand β 3 and helix α 3, was found in PIK-TE-2. In contrast, a water molecule bridge between the substrate and side chain of Thr77 was captured in PIK-TE-1 (Figure S3). These were in accord with previous reports that Thr77 plays a key role in stabilization of the oxygenion intermediate¹³ and the cyclic reactive conformation.³⁴ Furthermore, two dominant hydrophobic interactions between a binding cavity residue (Ala78 and Gly150) and substrate 1 were found in PIK-TE-1 and three hydrophobic interactions of Ala78, Gly150, and Tyr178 with substrate 2 existed in PIK-TE-2. The mutation of Y171F in DEBS TE, corresponding to Tyr178 in PIK TE, were found to be catalytically inactive,



Figure 5. Key interactions between substrates and PIK TE.

indicating the key role of Tyr171 in DEBS TE catalysis. However, theoretical studies on the role of Y178 in PIK TE are quite rare. According to our simulations, Y178 has hydrophobic interactions with substrate **2**, suggesting a sort of correlation with substrate recognition. On the other hand, Y178 in PIK-TE-**1** was observed to leave away from active site, expanding the exit gate and consequently promoting product release (Figure S4). Therefore, it was proposed that Y178 of PIK-TE might play a different role in substrate recognition and product release. In short, these results demonstrated that a hydrogenbonding interaction with His268 is essential to the formation of the prereaction state and hydrophobic interactions may play a significant role in substrate recognition and product release.

For further study, we focused on the active site pockets in PIK-TE-1 and PIK-TE-2 in 5 ns MD simulations (obtained from the dominant clusters of MD simulations) utilizing POVME 2.0, ^{36,37} a computational tool that characterizes the shape and size of protein pockets. Although the ring sizes of substrates 1 and 2 are slightly different, their binding pockets have comparable volumes of 335 Å³ in PIK-TE-1 and 355 Å³ in PIK-TE-2, indicating the tolerance of PIK TE (Figure 6). We speculated that PIK TE has sufficient space to accommodate different substrates in its pockets and that whether the substrate could transform to its prereaction state seemed to be decisive of whether macrocyclization occurred.

Next, we carefully compared the active sites of prereaction states in PIK-TE-1 and PIK-TE-2. It was observed that the major interactions between PIK TE and substrates were similar in both systems, especially at the end of the substrates, where the N_e atom of His268 grabbed the distal hydroxyl group (O11 and O13) by a hydrogen bond and further stabilized the substrates. However, the major difference lay in the reactive center, where the orientation of the C1 carbonyl group looked different, if not reverse, in PIK-TE-1 and PIK-TE-2 (Figure 7). As a result, the nucleophilic attack, accompanied by a deprotonated hydroxyl group, would target the *si*- or *re*-face of the C1 carbonyl group. The different conformations were supported by the optimization calculated by the QM/MM method, where the C1 carbonyl group was found either in the same direction with the distal hydroxyl group, accompanied by



Figure 6. Active binding pocket and exit of PIK TE. Substrate 1 is shown in salmon and 2 in magenta.



Figure 7. Scheme of nucleophilic attacks of substrate 1 from the *re-* or *si-*face of the C1 carbonyl group to form tetrahedral intermediates and the optimized products of 10-deoxymethynolide and narbonolide.

si-face nucleophilic attack, or in the reverse orientation, followed by *re*-face attack. In addition, the conformer corresponding to *re*-face attack was found to be 7.1 kcal/mol more stable than that of the *si*-face, demonstrating the preponderance of conformer taken by *re*-face attack. Meanwhile, we optimized 10-deoxymethynolide and narbonolide at the M062x/6-31g(d) level and compared their stability. Products obtained from *si*-face nucleophilic attack were less stable than those from *re*-face by 2.7 kcal/mol (10-deoxymethynolide) and 3.0 kcal/mol (narbonolide). These results suggested that macrocyclization from *re*-face nucleophilic attack may be more favorable than that from *si*-face attack in thermodynamics.

Macrocyclization in PIK-TE-1. To investigate the energetic distinction between macrocyclization reacted from *re-* or *si-*face nucleophilic attack, system PIK-TE-1 was carefully studied and four potential energy surfaces demonstrating the formation of a tetrahedral intermediate and the release of product 10-deoxymethynolide were plotted. These computational results demonstrated that the activated energy barrier via *re-*face

nucleophilic attack is 16.3 kcal/mol, 3.6 kcal/mol lower than that via *si*-face nucleophilic attack.

Macrocyclization from re-Face. Two two-dimensional potential energy surfaces were calculated with the QM/MM method to aid in understanding the macrocylization catalyzed by PIK TE from *re*-face nucleophilic attack. First, a potential energy surface corresponding to formation of a tetrahedral intermediate was calculated by defining the distances $d_2(O11-H)$ and $d_1(O11-C1)$ as the reaction coordinates, which represents the deprotonation of a hydroxyl group and *re*-face nucleophilic attack, respectively. The potential energy surface with the key structures along the reaction pathway, including 1 TS1 and 1 TS2, are shown in Figure 8. The d_1 distance was



Figure 8. Two-dimensional potential energy surfaces in the PIK-TE-1 system including the formation of tetrahedral intermediates from *re*face nucleophilic attack and the release of product 10-deoxymethyno-lide, shown with optimized structures of transition states.

found to be 2.67 Å in the optimized structure $\mathbf{1}_R$ and 1.60 Å in the optimized intermediate $\mathbf{1}_I$ IM1, and the d_2 distance is 1.01 Å in $\mathbf{1}_R$ and 1.65 Å in $\mathbf{1}_I$ IM1. The transition state $\mathbf{1}_I$ TS1 was located at $d_1 = 1.60$ Å and $d_2 = 1.60$ Å. The calculated potential energy barrier of formation of the tetrahedral intermediate was 10.2 kcal/mol, indicating that the formation of a tetrahedral intermediate could proceed spontaneously.

Next, another potential energy surface was calculated to uncover the energy barrier of product release step by defining the distance $d_3(O_{ser}-C1)$ and $d_4(O_{ser}-H)$ as the reaction coordinates, which represented the released product and the refreshed Ser148 hydroxyl group. The d_3 distance was found to be 1.46 Å in the optimized intermediate 1_IM1 and 2.57 Å in the optimized product 1_P, and the d_4 distance was 2.36 Å in 1_IM1 and 1.00 Å in 1_P. The transition state 1_TS2 was located at $d_3 = 1.80$ Å and $d_4 = 1.30$ Å. The calculated potential energy barrier of formation of the tetrahedral intermediate was 6.9 kcal/mol. Taken together, the energy barrier of the macrocylization catalyzed by PIK TE from *re*-face nucleophilic attack was 16.3 kcal/mol.

Macrocyclization from si-Face. We also studied the macrocyclization reaction from *si*-face nucleophilic attack. Similar to that from the *re*-face, the potential energy surfaces were calculated by defining the $d_2(O11-H)$ and $d_1(O11-C1)$ distances as the reaction coordinates, representing the deprotonation of the hydroxyl group and *si*-face nucleophilic



Figure 9. Two-dimensional potential energy surfaces in the PIK-TE-1 system including the formation of tetrahedral intermediates from *si*-face nucleophilic attack and the release of product 10-deoxymethynolide are shown with optimized structures of the transition state.

distance was found to be 2.77 Å in the optimized structure $\mathbf{1}_{R'}$ and 1.58 Å in the optimized intermediate $\mathbf{1}_{IM1'}$, and the d_2 distance was 1.03 Å in $\mathbf{1}_{R'}$ and 1.86 Å in $\mathbf{1}_{IM1'}$. The transition state $\mathbf{1}_{TS1'}$ was located at $d_1 = 1.90$ Å and $d_2 = 1.60$ Å. The calculated potential energy barrier of formation of the tetrahedral intermediate from the *si*-face was 12.3 kcal/mol, which was 2.1 kcal/mol higher than that from *re*-face nucleophilic attack, suggesting macrocyclization proceeding from *re*-face nucleophilic attack.

Next, we carefully examined the structure of $1_IM1'$, whose energy was found to be slightly lower than that of the transiton state $1_TS1'$ by 1.0 kcal/mol, suggesting that dissociation of the intermediate state could occur easily.^{38,39} We proposed a conformational transition through the movement of imidazole in His268 to transfer its proton to Ser148, since the distance between N_e of His268 and O of Ser148 is larger than 4.61 Å in $1_IM1'$. The conformational transition was supported through further optimizations. $1_IM2'$ and $1_IM3'$ were located along the reaction pathway (Figure 10), which was 1.4 and 2.4 kcal/ mol higher than $1_IM1'$ in energy, suggesting that the conformational transition could proceed spontaneously. This structural reorganization made the O atom of Ser148 more inclined to accept the proton of His268 and promote the release of products.

Then the energy barrier of the product-releasing step was calculated with a potential energy surface by defining the $d_3(O_{ser}-C1)$ and $d_4(O_{ser}-H)$ distances to represent the departure of the product and the refresh of the Ser148 hydroxyl group. The d_3 distance was found to be 1.50 Å in the optimized intermediate **1**_IM3' and 2.79 Å in the optimized product **1**_P', and the d_4 distance was 2.49 Å in **1**_IM3' and 1.04 Å in **1**_P'. The transition state **1**_TS2' was located at $d_3 = 1.90$ Å and $d_4 = 1.70$ Å. The calculated potential energy barrier of product release was 6.2 kcal/mol, suggesting that the release step should proceed readily.

Taken together, our computational results revealed that although the formation of a tetrahedral intermediate is



Figure 10. Conformational transformations in the PIK-TE-1 system.

advantageous in energy (10.2 and 12.3 kcal/mol), the dissociation of the intermediate could occur easily, and macrocyclization proceeding from *re*-face nucleophilic attack was more favorable than that from *si*-face attack (Figure 11). Therefore, we will focus on the macrocyclization arising from *re*-face nucleophilic attack pathway in the following study.

Macrocyclization from re-Face in PIK-TE-2. Similar to PIK-TE-1 system, two potential energy surfaces were calculated to understand the formation of the tetrahedral intermediate and release of product narbonolide in PIK-TE-2. According to calculations presented above, only the re-face nucleophilic attack pathway was involved in further study. The $d_1(O13-C1)$ and $d_2(O13-H)$ distances were used to represent nucleophilic attack and deprotonation of the hydroxyl group, while $d_3(O_{ser})$ C1) and $d_4(O_{ser}-H)$ represented the released products and the refreshed Ser148 hydroxyl group (Figure 12). The d_1 distance was found to be 3.33 Å in optimized structure 2 R and 1.50 Å in optimized intermediate 2 IM, and the d_2 distance was 1.00 Å in 2 R and 1.82 Å in 2 IM. The transition state 2 TS1 was located at $d_1 = 1.50$ Å and $d_2 = 1.70$ Å. The potential energy barrier for the formation of the tetrahedral intermediate was calculated to be 15.4 kcal/mol, 3.1 kcal/mol higher than that in PIK-TE-1. Furthermore, the d_3 distance was found to be 1.46 Å in optimized intermediate 2_IM and 2.66 Å in optimized product 2_P, and the d_4 distance was 2.10 Å in 2_IM and 0.99 Å in 2_P. The transition state 2_TS2 was located at $d_3 = 1.70$ Å and $d_4 = 1.30$ Å. The calculated potential energy barrier of product release was 5.1 kcal/mol, 1.1 kcal/mol lower than that in PIK-TE-1. To sum up, the energy barrier for the PIK-TEcatalyzed macrocylization from re-face nucleophilic attack in PIK-TE-2 was 20.4 kcal/mol (Figure S5).

Catalytic Mechanism. These data together suggest that nucleophilic attack, when accompanied by the deprotonation of prereaction state, can proceed from either the *re*-face or *si*-face to result in the formation of a charged tetrahedral intermediate. In the system PIK-TE-1, the energy barrier was calculated to be 10.2 kcal/mol in the former and 12.3 kcal/mol in the latter. Unlike in the *re*-face nucleophilic attack, the conformational reorganization was found to be indispensable in *si*-face attack to refreshing Ser148. Finally, following the collapse of the transient intermediate, the final macrocyclic product was



Figure 11. Energy profile of PIK-TE-1 system.



Figure 12. Two-dimensional potential energy surfaces and key structures in the PIK-TE-2 system.

Table 1. Energetic Corrections at Different Levels with M	M06-2x Method
---	---------------

	M062x/6-31g(d)	M062x/6-311+g(d)	M062x/6-311+g(d,p)	M062x/6-311+g(2df,2p)
1	-1618.101628 hartree	-1618.550480 hartree	-1618.607062 hartree	-1618.697213 hartree
1_TS1	-1618.085385 hartree	-1618.533414 hartree	-1618.587198 hartree	-1618.676407 hartree
	10.2 kcal/mol	10.7 kcal/mol	12.5 kcal/mol	13.1 kcal/mol
1_IM	-1618.086696 hartree	-1618.534606 hartree	-1618.588215 hartree	-1618.677421 hartree
1_TS2	-1618.075711 hartree	-1618.522130 hartree	-1618.578242 hartree	-1618.667330 hartree
	6.9 kcal/mol	7.8 kcal/mol	6.3 kcal/mol	6.3 kcal/mol
EB ^a	16.3 kcal/mol	17.8 kcal/mol	18.1 kcal/mol	18.8 kcal/mol
2	-1770.724558 hartree	-1771.217906 hartree	-1771.276472 hartree	-1771.374591 hartree
2_ TS1	-1770.700096 hartree	-1771.191112 hartree	-1771.246252 hartree	-1771.343941 hartree
	15.4 kcal/mol	16.8 kcal/mol	19.0 kcal/mol	19.2 kcal/mol
2_ IM	-1770.691987 hartree	-1771.191075 hartree	-1771.246289 hartree	-1771.343915 hartree
2 _TS2	-1770.728631 hartree	-1771.182171 hartree	-1771.239703 hartree	-1771.336885 hartree
	5.1 kcal/mol	5.6 kcal/mol	4.1 kcal/mol	4.4 kcal/mol
EB	20.4 kcal/mol	22.4 kcal/mol	23.1 kcal/mol	23.7 kcal/mol

^{*a*}EB denotes energy barrier.

released with a relatively low barrier of 6.9(6.2) kcal/mol in PIK-TE-1 and 5.1 kcal/mol in PIK-TE-2. The energy barrier of the whole pathway was 16.3 kcal/mol in PIK-TE-1 and 20.5 kcal/mol in PIK-TE-2. Moreover, energetic corrections of key structures at different levels were performed and the results consistently suggest the favored preference for the product 10-deoxymethynolide (Table 1). As we all know, the distribution of polyketide products is governed by a number of programmed events that are controlled by polyketide synthases. Herein, our calculated results were coincidentally in agreement with the experimental observation,⁴⁰ where the catalysis of the hexaketide *N*-acetylcysteamine thioester (SNAC) connected substrate under PIK module 6+TE (in the presence of cosubstrate methylmalonyl CoA) gave a 4:1 mixture of 10-deoxymethynolide and narbonolide.

CONCLUSIONS

TE-catalyzed macrocyclization of a linear polyketide acyl chain is known to be the offload step in polyketide synthase (PKS) mediated biosynthesis of macrocyclic polyketides. Although substantial effort has been undertaken to elucidate the structural characteristics of TEs, the molecular mechanism remains unclear. Here we combined MD simulations with QM/ MM calculations on the PIK-TE-1 and PIK-TE-2 complexes to develop a high-resolution model detailing enzyme-substrate interactions to rationalize macrocyclization of 10-deoxymethonilide and narbonylide by the pikromycin (PIK) TE. In comparison with PIK-TE-2, "active" conformers were more frequently observed in PIK-TE-1 with the potential to easily transform into the prereaction state, which appeared to be critical to macrocyclization. In addition, a hydrogen-bonding interaction with His268 was pivotal to the formation of the prereaction state and the hydrophobic interactions had potentially an essential role in substrate recognition and product release. In particular, Tyr178 was proposed to widen the exit and promote product release in PIK-TE-1. Calculations on volumes of binding pockets indicated that PIK TE had sufficient space to accommodate both substrates 1 and 2, exhibiting the tolerance of PIK TE. Moreover, potential energy surfaces were calculated with the QM/MM method to obtain the transition states and the energy barriers. Our computational results indicated that the macrocylization catalyzed by PIK TE from *re*-face nucleophilic attack is more favorable than that from si-face attack in both thermodynamic and kinetic aspects. Although the formation of a tetrahedral intermediate was advantageous in energy, the reverse dissociation of the intermediate occurred easily. The energy barrier of the whole pathway was calculated to be 16.3 kcal/mol in PIK-TE-1 and 20.4 kcal/mol in PIK-TE-2. This is in good agreement with biochemical and structural studies. Our study provides insight into TE-catalyzed macrocyclization and increases the understanding of type I PKS TEs in the engineered PKS pathway.

ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.8b01156.

Root-mean-square deviation (RMSD) and root-meansquare fluctuation (RMSF) values of MD simulations of PIK-TE-1 and PIK-TE-2 systems, hydrogen-bonding interactions between Thr77 and substrates, the Tyr178 in PIK-TE-1, and energy profile in the PIK-TE-2 system (PDF)

AUTHOR INFORMATION

Corresponding Author

*Y.-L.Z.: e-mail, yileizhao@sjtu.edu.cn; tel/fax, +86-21-34207190.

ORCID [©]

Yi-Lei Zhao: 0000-0003-4687-7847

Author Contributions

T.S., L.B., and Y.-L.Z. conceived and designed the investigation. T.S., L.L., W.T., S.L., and X.-L.W. performed calculations and analyses. T.S., S.F., and Y.-L.Z. wrote up the paper.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank the National Basic Research Program of China "973" (2012CB721005) and the National High-tech R&D Program of China "863" (2012AA020403), National Science Foundation of China (21377085 and 31770070), SJTU-YG2016MS42, and the SJTU-HPC computing facility award for financial support and computational hours. The authors acknowledge Dr. Timo Törmäkangas for his helpful revision on language. This work is dedicated to Professors Zhenyi Wen, Kendall N. Houk, and Yundong Wu on the occasion of their 80th, 75th, and 60th birthdays, respectively.

REFERENCES

(1) Sieber, S. A.; Marahiel, M. A. Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. *Chem. Rev.* **2005**, *105*, 715–738.

(2) Meier, J. L.; Burkart, M. D. The chemical biology of modular biosynthetic enzymes. *Chem. Soc. Rev.* 2009, 38, 2012–2045.

(3) Williams, G. J. Engineering polyketide synthases and nonribosomal peptide synthetases. *Curr. Opin. Struct. Biol.* **2013**, *23*, 603– 612.

(4) Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J. Nat. Prod.* **2012**, 75, 311–335.

(5) Larsen, E. M.; Wilson, M. R.; Taylor, R. E. Conformation-activity relationships of polyketide natural products. *Nat. Prod. Rep.* 2015, *32*, 1183–1206.

(6) Khosla, C. Structures and mechanisms of polyketide synthases. J. Org. Chem. 2009, 74, 6416–6420.

(7) Wang, M.; Zhou, H.; Wirz, M.; Tang, Y.; Boddy, C. N. A thioesterase from an iterative fungal polyketide synthase shows macrocyclization and cross coupling activity and may play a role in controlling iterative cycling through product offloading. *Biochemistry* **2009**, *48*, 6288–6290.

(8) Du, L.; Lou, L. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255–278.

(9) Pinto, M.; Wang, M.; Horsman, M.; Boddy, C. N. 6-Deoxyerythronolide B synthase thioesterase-catalyzed macrocyclization is highly stereoselective. *Org. Lett.* **2012**, *14*, 2278–2281.

(10) Xu, Y.; Zhou, T.; Zhang, S.; Xuan, L.-J.; Zhan, J.; Molnar, I. Thioesterase domains of fungal nonreducing polyketide synthases act as decision gates during combinatorial biosynthesis. *J. Am. Chem. Soc.* **2013**, *135*, 10783–10791.

(11) Xue, Y.; Zhao, L.; Liu, H.-W.; Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in Streptomyces venezuelae: architecture of metabolic diversity. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95, 12111–12116.

(12) Lu, H.; Tsai, S.-C.; Khosla, C.; Cane, D. E. Expression, sitedirected mutagenesis, and steady state kinetic analysis of the terminal thioesterase domain of the methymycin/picromycin polyketide synthase. *Biochemistry* **2002**, *41*, 12590–12597.

(13) Tsai, S.-C.; Lu, H.; Cane, D. E.; Khosla, C.; Stroud, R. M. Insights into channel architecture and substrate specificity from crystal structures of two macrocycle-forming thioesterases of modular polyketide synthases. *Biochemistry* **2002**, *41*, 12598–12606.

(14) Giraldes, J. W.; Akey, D. L.; Kittendorf, J. D.; Sherman, D. H.; Smith, J. L.; Fecik, R. A. Structural and mechanistic insights into polyketide macrolactonization from polyketide-based affinity labels. *Nat. Chem. Biol.* **2006**, *2*, 531–536.

(15) Akey, D. L.; Kittendorf, J. D.; Giraldes, J. W.; Fecik, R. A.; Sherman, D. H.; Smith, J. L. Structural basis for macrolactonization by the pikromycin thioesterase. *Nat. Chem. Biol.* **2006**, *2*, 537–542.

(16) Aldrich, C. C.; Venkatraman, L.; Sherman, D. H.; Fecik, R. A. Chemoenzymatic synthesis of the polyketide macrolactone 10deoxymethynolide. *J. Am. Chem. Soc.* **2005**, *127*, 8910–8911.

(17) He, W.; Wu, J.; Khosla, C.; Cane, D. E. Macrolactonization to 10-deoxymethynolide catalyzed by the recombinant thioesterase of the picromycin/methymycin polyketide synthase. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 391–394.

(18) Li, J.; Ehlers, T.; Sutter, J.; Varma-O'Brien, S.; Kirchmair, J. CAESAR: a new conformer generation algorithm based on recursive buildup and local rotational symmetry consideration. *J. Chem. Inf. Model.* **2007**, *47*, 1923–1932.

(19) Stewart, J. J. P. Optimization of parameters for semiempirical methods I: Method. J. Comput. Chem. 1989, 10, 209-220.

(20) Stewart, J. J. P. Optimization of parameters for semiempirical methods II: Applications. J. Comput. Chem. **1989**, 10, 221–264.

(21) Anders, E.; Koch, R.; Freunscht, P. Optimization and application of lithium parameters for PM3. *J. Comput. Chem.* **1993**, *14*, 1301–1312.

(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09; Gaussian, Inc., Wallingford, CT, 2013.

(23) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X., Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California, San Francisco, CA, 2012.

(24) Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-consistent molecular orbital methods. IX. An extended Gaussian-type basis for molecular orbital studies of organic molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.

(25) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

(26) Payne, P. W. The Hartree-Fock theory of local regions in molecules. J. Am. Chem. Soc. 1977, 99, 3787–3794.

(27) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of

Proteins, Nucleic Acids, and Organic Molecules. J. Am. Chem. Soc. 1995, 117, 5179–5197.

(28) 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, 10089–10092.

(29) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, 23, 327–341.

(30) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent 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*, 215–241.

(31) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. A new ONIOM implementation in Gaussian 98. Part 1. The calculation of energies, gradients and vibrational frequencies and electric field derivatives. *J. Mol. Struct.: THEOCHEM* **1999**, 461-462, 1–21.

(32) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.

(33) Chen, X.-P.; Shi, T.; Wang, X.-L.; Wang, J.; Chen, Q.; Bai, L.; Zhao, Y.-L. Theoretical Studies on the Mechanism of Thioesterase-Catalyzed Macrocyclization in Erythromycin Biosynthesis. *ACS Catal.* **2016**, *6*, 4369–4378.

(34) Koch, A. A.; Hansen, D. A.; Shende, V. V.; Furan, L. R.; Houk, K. N.; Jiménez-Osés, G.; Sherman, D. H. A Single Active Site Mutation in the Pikromycin Thioesterase Generates a More Effective Macro-cyclization Catalyst. *J. Am. Chem. Soc.* **2017**, *139*, 13456–13465.

(35) Wang, M.; Boddy, C. N. Examining the role of hydrogen bonding interactions in the substrate specificity for the loading step of polyketide synthase thioesterase domains. *Biochemistry* **2008**, *47*, 11793–11803.

(36) Durrant, J. D.; Votapka, L.; Sørensen, J.; Amaro, R. E. POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics. *J. Chem. Theory Comput.* **2014**, *10*, 5047–5056.

(37) Wagner, J. R.; Sørensen, J.; Hensley, N.; Wong, C.; Zhu, C.; Perison, T.; Amaro, R. E. POVME 3.0: Software for Mapping Binding Pocket Flexibility. J. Chem. Theory Comput. 2017, 13, 4584–4592.

(38) Świderek, K.; Martí, S.; Moliner, V. Theoretical Study of Primary Reaction of Pseudozyma antarctica Lipase B as the Starting Point To Understand Its Promiscuity. ACS Catal. 2014, 4, 426–434.

(39) Martínez-González, J. Á.; González, M.; Masgrau, L.; Martínez, R. Theoretical Study of the Free Energy Surface and Kinetics of the Hepatitis C Virus NS3/NS4A Serine Protease Reaction with the NS5A/5B Substrate. Does the Generally Accepted Tetrahedral Intermediate Really Exist? ACS Catal. 2015, 5, 246–255.

(40) Wu, J.; He, W.; Khosla, C.; Cane, D. C. Chain elongation, macrolactonization, and hydrolysis of natural and reduced hexaketide substrates by the picromycin/methymycin polyketide synthase. *Angew. Chem., Int. Ed.* **2005**, *44*, 7557–7560.

NOTE ADDED AFTER ASAP PUBLICATION

After this paper was published ASAP April 13, 2018, the use of *si-* and *re-* were reversed throughout the text and in Figures 7, 8, 9, and 11. The revised version was reposted April 16, 2018