

Insights into Specificity and Catalytic Mechanism of Amphotericin B/Nystatin Thioesterase

Rufan Wang¹, Wentao Tao², Lei Liu², Chen Li², Linquan Bai², Yilei Zhao², and Ting Shi¹

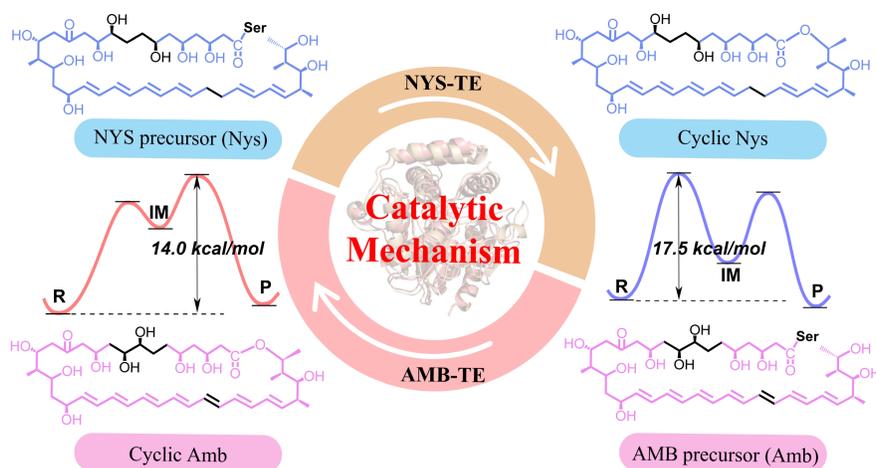
¹Shanghai Jiao Tong University - Minhang Campus

²Shanghai Jiao Tong University

August 23, 2020

Abstract

Polyene polyketides amphotericin B (AMB) and nystatin (NYS) are important antifungal drugs. Thioesterases (TEs), located at the last module of PKS, control the release of polyketides by cyclization or hydrolysis. Intrigued by the tiny structural difference between AMB and NYS, as well as the high sequence identity between AMB TE and NYS TE, we constructed four systems to study the structural characteristics, catalytic mechanism, and product release of AMB TE and NYS TE with combined MD simulations and QM/MM calculations. The results indicated that compared with AMB TE, NYS TE shows higher specificity on its natural substrate and R26 as well as D186 were proposed to a key role in substrate recognition. The energy barrier of macrocyclization in AMB-TE-Amb and AMB-TE-Nys systems were calculated to be 14.0 and 22.7 kcal/mol, while in NYS-TE-Nys and NYS-TE-Amb systems, their energy barriers were 17.5 and 25.7 kcal/mol, suggesting the cyclization with their natural substrates were more favorable than that with exchanged substrates. At last, the binding free energy obtained with the MM-PBSA.py program suggested that it was easier for natural products to leave TE enzymes after cyclization. And key residues to the departure of polyketide product from the active site were highlighted. We provided a catalytic overview of AMB TE and NYS TE including substrate recognition, catalytic mechanism and product release. These will improve the comprehension of polyene polyketide TEs and benefit for broadening the substrate flexibility of polyketide TEs.



Keywords

Amphotericin B, Nystatin, Thioesterase, Macrocyclization, MD simulations, QM/MM calculations

1. Introduction

Polyene polyketides [1,2] consisting of macrolactone cores with 3-8 conjugated double bonds are mainly synthesized by *streptomyces* and related *actinobacteria*. They are widely used as anti-fungal and anti-parasitic drugs, especially for the treatment of human systemic fungal infection [3]. The infection, as one of the major complications of immunodeficiency [4], has become the main cause of death from AIDS and cancer [5]. Besides, polyenes are uncovered with anti-retroviral, anti-prion, and immunomodulating activities, which increasingly attract the attention of researchers.

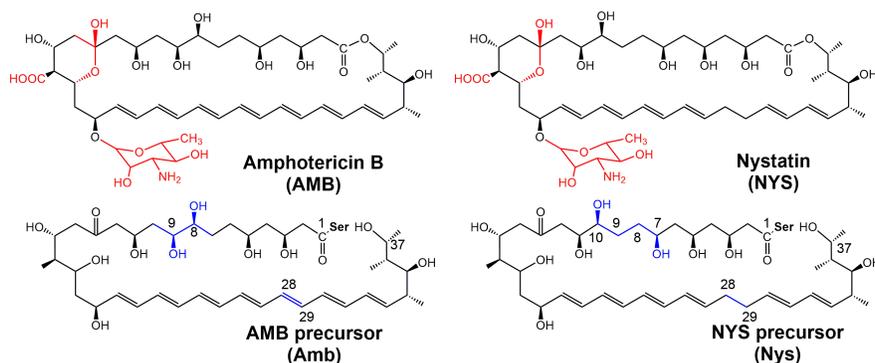


Fig. 1. The structures of amphotericin B (AMB), nystatin (NYS) and their precursors linked to serine of TE (Amb and Nys).

Amphotericin B (AMB) [6] and nystatin (NYS) are typical polyene polyketides. They are very similar in structure except for the position of hydroxyl groups and the number of double bonds, shown in Fig. 1. AMB, produced by *Streptomyces nodosus*, is a prominent heptaene-containing polyketide [7,8]. Benefited from broad spectrum of potent antifungal activity and rare drug resistance [9-11], AMB has become one of the indispensable drugs for the treatment of deep fungal infections [12,13]. In addition, NYS, another famous tetraene-containing polyketide [14-16], is documented in the World Health Organization’s Essential Medicines List as one of the most effective and safest drugs required for health systems [17]. Although the outstanding pharmacological effects are exerted [18], polyene drugs have certain toxic and side effects on human cells (such as severe hemolytic toxicity, nephrotoxicity and liver toxicity), thus severely limiting their clinical application for antifungal therapy [19]. Considerable efforts have been made by chemists, biophysicists and biologists to study the structure-activity relationships of polyenes, which is of great significance for continuously exploring and developing new types of polyketide drugs with the need for non-toxic antifungals.

Both AMB and NYS are biosynthesized by Modular type I polyketide synthases (PKSs). Considered as “molecular plant”, PKSs contain a plurality of functional proteins linearly assembled into modules to be responsible for the synthesis of polyenes through the initiation, elongation, reduction, and extension of the construction of polyketide scaffolds [20]. The biosynthesis process is like a pipeline: (1) the substrate covalently bound to acyl carrier protein ACP is transferred to the active site of the next enzyme [21]; (2) the repeated decarboxylation condensation reaction causes the polyketone chain to grow; (3) the thioesterase (TE) domain at the end of the PKS catalyzes the cyclization and release of the polyketone chain [22]. Eventually, the lactone ring product is formed and further catalyzed by a series of post-modification enzymes (e.g., glycosyltransferases and P450 oxidases) to form the final product of the polyketide.

TE is uncovered to be a selective switch for controlling the cyclization or hydrolysis release of polyketides, and the precise identification of substrates and catalytic efficiency by thioesterase determines the yield of polyketides to some extent [23,24]. The catalytic mechanism of thioesterase is generally considered to be two steps. The first step is the loading process of a substrate: the serine of catalytic triad attacks the thioester bond of the substrate to form a covalent enzyme-substrate complex. The second step is the release of the product: the complex is attacked by a nucleophilic group of substrates or a water molecule, thus releasing

the cyclized polyketide [25] or the hydrolysate. It is worth noticing that the sequence identity between AMB TE and NYS TE is up to 86.12 %.

Our previous studies [26] have shown that the thioesterase-substrate system will reach Pre-Reaction States (PRSs) before the cyclization reaction occurs. In this special state, histidine and aspartic acid 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 nucleophilic attack to obtain a tetrahedral intermediate. The formation of PRS will play a decisive role in the occurrence of the cyclization reaction.

Intrigued by structural similarity between AMB/AMB precursor (Amb) and NYS/NYS precursor (Nys), as well as high sequence identity between AMB TE and NYS TE, we built four systems as AMB-TE-Amb, AMB-TE-Nys, NYS-TE-Nys and NYS-TE-Amb to investigate their respective structural characteristics and the catalytic differences between them with combined MD simulations and QM/MM calculations. By exploring the details in substrate recognition, catalytic mechanism as well as product release, we found some unique amino acid residues in AMB and NYS TEs and elucidated the catalytic mechanisms of AMB and NYS TEs. Our results may provide a better understanding in protein engineering of AMB and NYS TEs, as well as type I PKS TEs.

2. Materials and Methods

2.1. System Preparation

There is no available crystallization data of AMB TE and NYS TE in PDB Bank, so pikromycin TE [27] (PDB code: 2H7Y) was used as a template to build AMB TE (sequence similarity: 67.8 %) and NYS TE (sequence similarity: 67.5 %) structures by homology modeling (<http://swissmodel.expasy.org/>). Since the substrate is transferred to TE by forming a covalent enzyme-substrate complex without lactonic ring and exocyclic mycosamine, the linear Amb and Nys molecules (precursors of AMB and NYS) were utilized as substrates in our computational models. Classic conformation search method CAESAR [28] was used to obtain total of 60 conformations of substrates Amb and Nys. After that, the molecules were optimized with Gaussian 09 [29] AM1 method [30]. Then the energetically stabilized substrates were covalently bonded to active site Ser136 on AMB TE and Ser134 on NYS TE, with a hydrogen bond forming between their terminal hydroxyl and Ne of active site His254 in AMB TE and His252 in NYS TE. Furthermore, the PROPKA program (<http://propka.ki.ku.dk/>), one of the most commonly used empirical pKa predictors, was introduced to evaluate the protonated state of residues in AMB TE and NYS TE systems. All ionizable side chains were maintained in their standard protonation states at pH 7.0. To rationalize the reaction process, the protonation state of His254 (AMB TE) and His252 (NYS TE) were defined as HID.

To obtain the parameters of linear Amb and Nys substrates, an N-terminal cap (-CO-CH₃) and a C-terminal cap (-NH-CH₃) were added onto the disconnected Amb and Nys from Ser to block its ends. Then conformational optimization was performed at the level of HF/6-31G(d) on the molecules and calculated the electrostatic surface potential (ESP) charge of them. A two-step restrained electrostatic potential (RESP) model was used to determine charge distribution on the substrates. Finally, topology files for protein-substrate complexes were prepared with *tleap* module in AMBER 14 [31]. The AMB-TE-substrate systems (AMB-TE-Amb/ AMB-TE-Nys) and NYS-TE-substrate systems (NYS-TE-Nys/ NYS-TE-Amb) were solvated in an octahedral TIP3P water box [32] with a water thickness extending at least 10 Å away from the protein surface. Sixteen sodium ions were added to the AMB-TE systems and fourteen sodium ions were added to the NYS-TE systems to maintain charge neutralization.

2.2. Molecular Dynamics Simulation

The prepared structures mentioned above were taken as the starting points for MD simulations. Four systems including AMB-TE-Amb, AMB-TE-Nys, NYS-TE-Nys, and NYS-TE-Amb were constructed and 5×50 ns MD simulations for each system were performed. After that, two additional 300 ns simulations of AMB-TE-Amb and NYS-TE-Nys were performed to investigate the structural characteristics of TEs with their natural substrates. RMSD (root-mean-square deviation) and RMSF (root-mean-square fluctuation) analyses

revealed that all trajectories attained their equilibriums (Fig. S1-S2+).

Each MD simulation was carried out using the AMBER program suite with the ff03.r1 force field. First, the systems were subjected to 10000 steps of steepest descent energy minimization, to avoid any instability that might occur during the MD simulations. After that, systems changed into 1000 cycles of conjugate gradient minimization with bonds involving hydrogen constrained by SHAKE algorithm [33] followed by another 10000 steps of steepest descent energy minimization and 5000 cycles of conjugate gradient minimization with no constraint exerted. The systems were then heated from 0 to 300K through 25000 iterations. After a 200 ps-equilibrium in NPT ensemble, five 50-ns simulations (300 K, 1 atm) with different random seeds were conducted and Particle Mesh Ewald (PME) method [34] was used to calculate long-range electrostatic interactions. The van der Waals (VDW) interactions were cut off at 15 Å. Analyses of trajectories were performed using *cpptraj* in Ambertools14. The mutant systems were built by Discovery studio 3.5.

2.3. Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations.

QM/MM calculations were performed in the Gaussian 09 program using a two-layered ONIOM method [35,36]. Initial structures corresponding to pre-reaction state were obtained from the dominant MD clusters and further subjected to geometry optimization. The QM region consisted of side chains of active site triad (Ser136, Asp164, and His254 of AMB TE and Ser 134, Asp162, and His252 of NYS TE) and polyketide chain (substrate Amb or Nys), which added up to 139 atoms in Amb systems and 141 atoms in Nys systems. The QM layer bore one negative charge and the total system bore sixteen negative charges in AMB-TE system (fourteen negative charges in NYS-TE system), since all Na^+ ions were extracted from the structures. The optimizations in QM region were carried out under M06-2X [37] functional with 6-31G(d) basis set [38] and the calculations in MM region were performed with AMBER force field.

2.4. Preparation for Production Release Simulation.

After conformational optimization and parameter preparation for products (cyclized Amb and Nys), four enzyme-product systems were prepared for MD simulations. Each model was performed with 3x50 ns MD simulations. After that, the last 10 ns of each 50 ns MD simulation was used for MM-GBSA and MM-PBSA calculations. Parameter topology files of ligand, receptor, and complex were prepared with ante-MMPBSA.py and calculation was performed with MMPBSA.py

3. Results and Discussion

3.1. Conformational Analyses in AMB-TE and NYS-TE Systems

According to the general consensus, the macrocyclization easily proceeds when a hydrogen bond forms between terminal hydroxyl ($\text{O}_{\text{sub}}\text{H}$) of substrates and N_ϵ atom of catalytic histidine in TE, and the same O_{sub} atom falls exactly in a conducive range for the nucleophilic attack to carbonyl C_1 . These conformations in TE with $d(\text{O}_{\text{sub}}-\text{N}_\epsilon)$ [?] 3.0 Å and $d(\text{C}_1-\text{O}_{\text{sub}})$ [?] 4.5Å were defined as Pre-Reaction States (PRSs) in our previous works [25,26]. The PRS has been widely used in structural comparisons and mutant analyses [39-41]. To detect the stability of PRSs, our simulations were started from the low-energy structures where a hydrogen bond was formed between the terminal hydroxyl $\text{O}_{\text{sub}}\text{H}$ and N_ϵ atom of His254 in AMB TE or His252 in NYS TE.

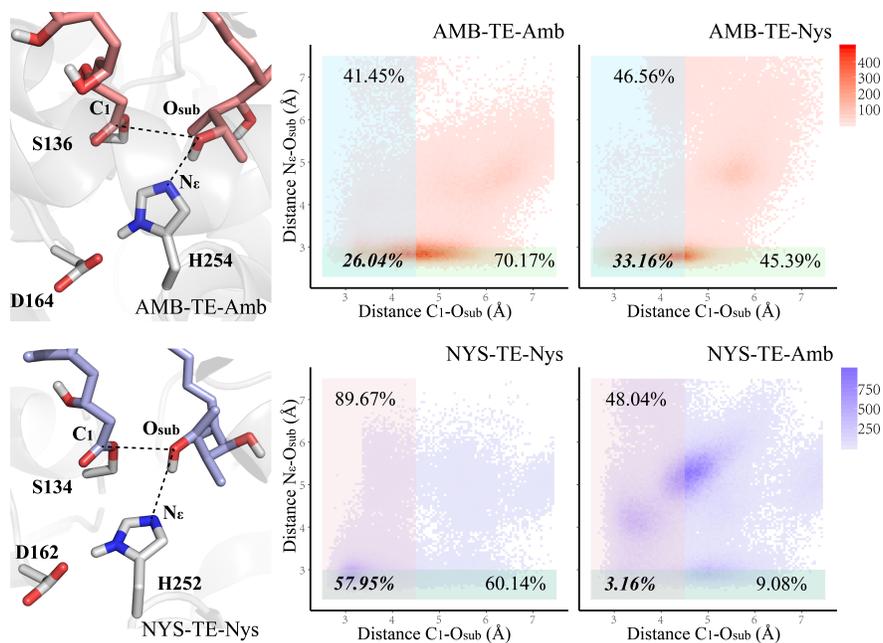


Fig. 2. Structures of catalytic centers in AMB-TE-Amb and NYS-TE-Nys systems and PRS proportions in 5×50 ns MD simulations.

According to our simulations, the proportions of $d(C_1-O_{sub})$ [?] 4.5Å was similar in AMB-TE-Amb (41.45 %) and AMB-TE-Nys (46.56 %) systems, while it rose up to 89.67 % in NYS-TE-Nys system, which strongly suggested the preference for nucleophilic attack of Nys. On the other hand, both TEs indicated the superiority of hydrogen-bond formation between O_{sub} and N_{ϵ} of H254/H252 to its natural substrate, since the proportion of $d(O_{sub}-N_{\epsilon})$ [?] 3.0 Å was found to be 70.17 % in the AMB-TE-Amb system and 60.14 % in the AMB-TE-Nys system, while it decreased to 45.39 % in AMB-TE-Nys and 9.08 % in NYS-TE-Amb system. Combined the two parts, the proportions of PRSs in consideration of both distances $d(O_{sub}-N_{\epsilon})$ and $d(C_1-O_{sub})$ were found to be 26.04 % in AMB-TE-Amb and 33.16 % in AMB-TE-Nys. The similar conformation proportion of PRSs suggested the good tolerance of AMB TE to different substrates. On the other hand, the proportion of PRS in NYS-TE-Nys system was 57.95 % (Fig. 2), much higher than 3.16 % in NYS-TE-Amb system, suggested that the natural substrate Nys has more favorable conformations in mutual recognition process that will eventually facilitate the following macrocyclization. In conclusion, NYS TE showed better recognition towards its natural substrate Nys, while AMB TE presented inconspicuous differences to Amb and Nys.

3.2. Hydrogen Bonding and Hydrophobic Interactions in NYS-TE Systems

Intrigued by obvious differences in conformational statistical analyses between NYS-TE-Nys and NYS-TE-Amb systems, where the PRSs maintained stably in the former and became collapsed in the latter, we carefully investigated the interactions between the substrates (Nys and Amb) and NYS TE in both systems. As shown in Fig. 3, besides the H254 of the catalytic triad, three hydrogen bonds between D186, R26, T179 and hydroxyl groups of its natural substrate Nys were detected. Their proportions of forming hydrogen bonds in MD simulations reached to 75.97 %, 59.13 %, and 52.03 % in the NYS-TE-Nys system, while they were sharply reduced to 4.37 %, 13.78 %, and 24.93 % in the NYS-TE-Amb system, respectively. Meanwhile, the top three were substituted to P68 (48.70 %), D175 (41.03 %), and T135 (39.99 %) in the NYS-TE-Amb system. These data were collected in Table S1+.

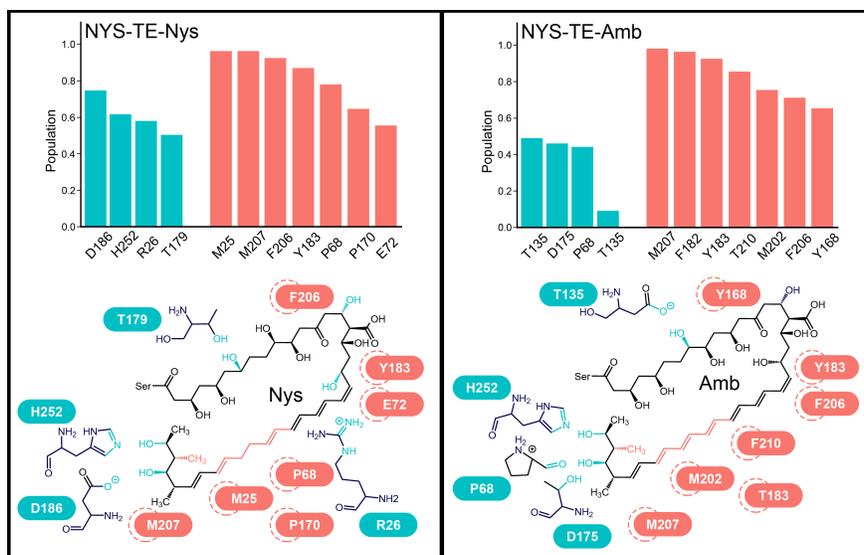


Fig. 3. Key residues of hydrogen bonds and hydrophobic interaction in NYS-TE systems.

In the NYS-TE-Nys system, R26 was observed form a hydrogen bond with the Nys's O_3 and drag Nys's head up (Fig. 4). With head pulling up by R26, the Nys's tail had more chance to enter reaction active site, which was similar to a lever, due to the rigidity of conjugate double bonds in Nys. With the help of R26, the substrate's tail was easily fixed by D186, which was located at helix $\alpha 1$. To verify the function of residues R26, 3×50 ns MD simulations were performed for R26A mutant in the NYS-TE-Nys system (Fig. S3+). In R26A mutant system, the substrate lost traction on its head and tail part detached from helix $\alpha 1$, which eventually obstructed the formation of PRS. Comparing to the high proportion of PRS (57.95 %) in wild type (WT), the proportion decreased to 29.01 % in the R26A mutant (Fig. S4+). From these observations, we proposed that R26 played a critical role in adjusting the substrate position to favorable PRS. Without the hydrogen bond of R26, the substrate was more likely to rotate randomly, making it difficult to lock the hydroxyl group tail with the active histidine of NYS TE and making it hard for nucleophilic attack at C1 after deprotonation.

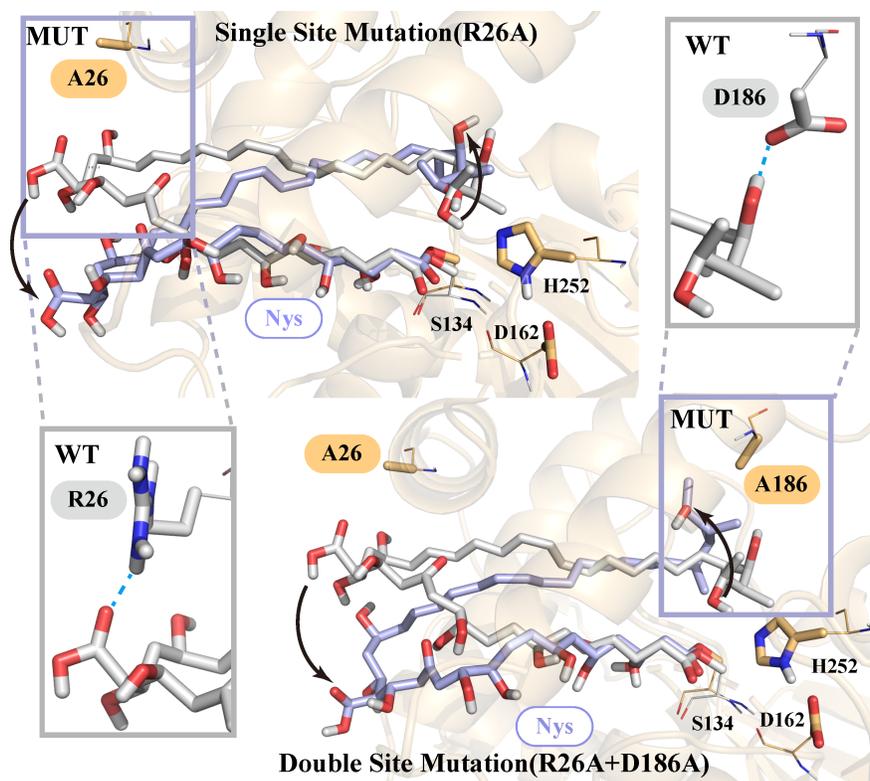


Fig. 4. Alignments of mutant R26A and R26A/D186A with wild type in NYS-TE-Nys system. Nys is shown in gray in wild type and is shown in light blue in mutants.

Considering the strong and stable interactions of a hydrogen bond between residue D186 and Nys, 3×50 ns MD simulation were performed for the combined mutant of R26A/D186A (Fig. S5+). As expected, the proportion of PRS in D186A/R26A system sharply decreased to 19.88 %, while it was 75.97 % in WT. By losing the interactions with the two critical residues R26 and D186, the structures of the substrate in the simulations were no longer stable. The head flipped down and the tail turned upside down. These results indicated that R26 and D186 played a critical role in the recognition process.

Different from Nys, where a carbon-carbon single bond in its hydrophobic side enhanced the flexibility of the polyene substrate, the conjugated heptaene of Amb made it more difficult to form hydrogen bonds between R26 and H252. The important hydrogen bonds were collected in Table S2+. Losing the locking effect of R26 and D186, Amb became unstable and few PRSs were observed in the Nys-TE-Amb system. Therefore, it could be concluded that R26 and D186 should be the specific recognition sites for NYS TE.

On the other hand, hydrophobic interactions were also remarkable in both systems. To be precise, hydrophobic amino acid Y183, F206, and M207 were detected in both systems. Besides that, M25 and P68 highlighted in the NYS-TE-Nys system and F182 and T210 stood out in the NYS-TE-Amb system. Similar hydrophobic interactions were reported by Akey [42]. With the assistance of hydrophobic amino acids, the linear and hydrophobic substrates can be induced to the substrate channel into conformations suitable to form macrocycles.

3.3. Key residues in AMB-TE Systems

Corresponding to R26 in NYS TE, in AMB TE, R28 could partly form a hydrogen bond with the hydroxyl oxygen atom of Amb. However, owing to the missing of a carbon-carbon single bond in its conjugated structure, it became difficult for Amb to form a stable hydrogen bond with R28. On the other hand,

although Q78, which was corresponding to D186 in Nys TE, could also form a stable hydrogen bond with Amb, it was far away from residue Ser132 of the active site (Fig. 5). Therefore, the strong interaction pulled the hydroxyl group of Amb away from C₁, resulting in fewer PRSs in the AMB-TE-Amb system.

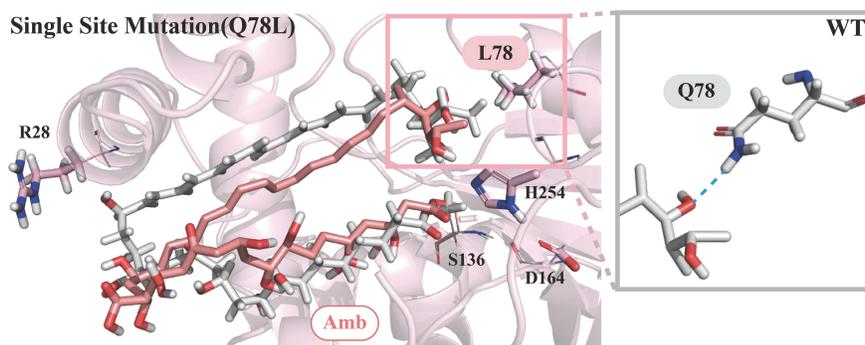


Fig. 5. Alignments of mutant Q78L with wild type. Amb is shown in gray in wild type and is shown in salmon in mutant.

What's more, 5×50 ns mutation MD simulations were carried out for Q78L in the AMB-TE-Amb system (Fig. S6-7+). As is shown in Fig. 5, while residue L78 failed to form a hydrogen bond with the substrate, substrate's tail could move freely, which made it more possible for the substrate to enter the active site. Comparing to the PRS proportion in WT (26.04 %), the PRS proportion of Q78L reached 41.74 %, increasing by 11.86 % (Fig. S8+). These results indicated the obstruction of Q78 in forming PRSs and mutation Q78L might help to improve the efficiency of macrocyclization. The wet experiment is in processing.

3.4. Catalytic mechanisms in AMB-TE and NYS-TE Systems

The macrocyclization takes place with the deprotonation of a hydroxyl group of polyketides, accompanied by the anionic oxygen nucleophile attacking C₁ and leading to the formation of a tetrahedral intermediate. Then, following the collapse of the tetrahedral intermediate, the cyclized polyketide product is released and the catalytic Serine is refreshed (Fig. 6). Herein, we utilized the ONIOM(M06-2X/6-31G(d):Amber) method to calculate two potential energy surfaces of the whole pathway of macrocyclization.

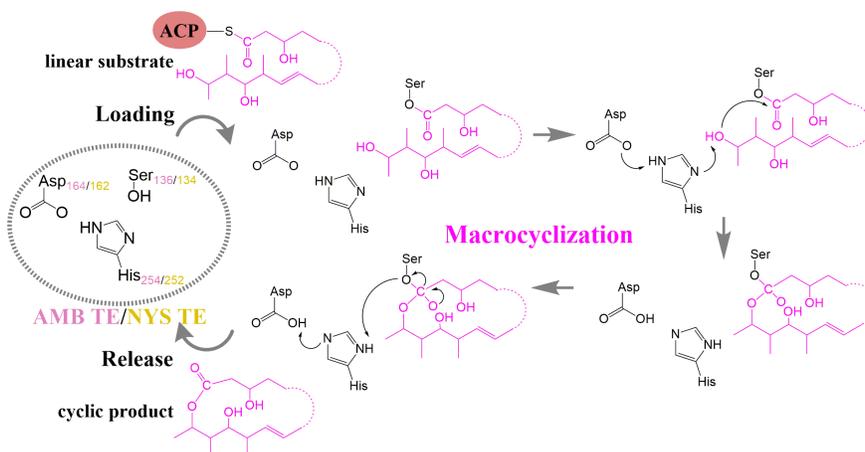


Fig. 6. The proposed mechanism of macrocyclization catalyzed by AMB TE and NYS TE.

3.4.1. AMB-TE system

First, we defined the distances $d_1(O_{11}-C_1)$ and $d_2(O_{11}-H)$ as the reaction coordinates, which corresponds to the formation of a tetrahedral intermediate. The potential energy surface with the key structures (TS1 and TS2) along the reaction pathway in the AMB-TE system are shown in Fig. 7.

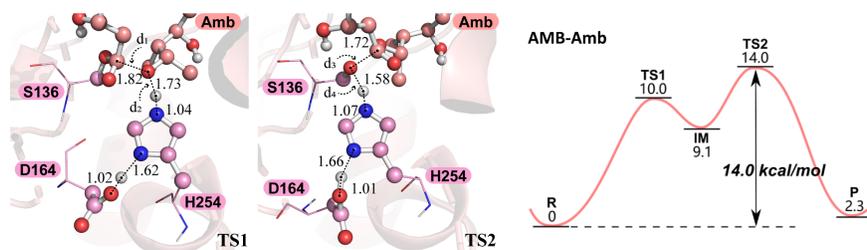


Fig. 7. Transition states and energy barriers of macrocyclization in AMB-TE-Amb system. Key residues in AMB TE are shown in pink and Amb is in salmon.

It was found that d_1 distance and d_2 distance were 2.82 Å and 1.00 Å in the reactant (**R**) of AMB-TE-Amb while they were 1.59 Å and 1.70 Å in the tetrahedral intermediate (**IM**) of AMB-TE-Amb (Fig. S9+). The transition state **TS1** was located at the $d_1=1.82$ Å and $d_2=1.73$ Å. The calculated potential energy barrier of formation of the tetrahedral intermediate was 10.0 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 the product release step by defining the distance $d_3(O_{ser}-C_1)$ and $d_4(O_{ser}-H)$ as the reaction coordinates. This represented the process of product releasing and Ser136 hydroxyl group refreshing. The d_3 distance was found to be 1.46 Å in the optimized intermediate **IM** and 2.61 Å in the optimized product (**P**) of AMB-TE-Amb, and the d_4 distance was 3.68 Å in **IM** and 1.02 Å in **P**. The transition state **TS2** was located at $d_3=1.72$ Å and $d_4=1.58$ Å. The calculated release potential energy barrier of the tetrahedral intermediate was 4.9 kcal/mol.

To sum up, the whole energy barrier of the cyclization was 14.0 kcal/mol, which is similar to the barrier (16.3 kcal/mol) of 10-deoxymethynolide macrocyclization catalyzed by picromycin TE[25]. These results indicated that AMB-TE can easily catalyze this macrocyclization with its natural substrate Amb.

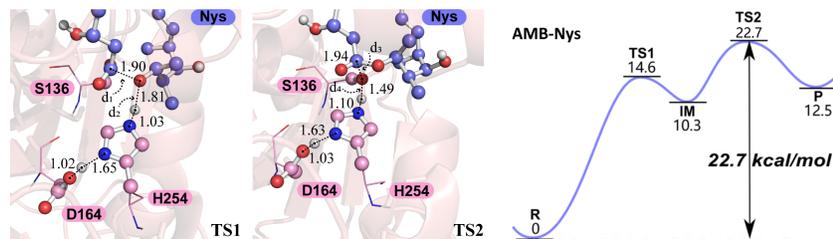


Fig. 8. Transition states and energy barriers of macrocyclization in AMB-TE-Nys systems. Key residues in AMB TE are shown in pink and Nys is in blue.

Similar to AMB-TE-Amb system, four reaction coordinates including distances $d_1(O_{11}-C_1)$ and $d_2(O_{11}-H)$, as well as $d_3(O_{ser}-C_1)$ and $d_4(O_{ser}-H)$ were defined in AMB-TE-Nys system (Fig. 8), where Nys was utilized as a substrate for the calculations of the two potential energy surfaces. The transition state **TS1** of AMB-TE-Nys was located at the $d_1=1.90$ Å and $d_2=1.81$ Å. The optimized structures of **R**, **IM**, and **P** were collected in Fig. S10+. The calculated formation potential energy barrier of the tetrahedral intermediate was 14.6 kcal/mol, which was 4.6 kcal/mol higher than the Amb substrate. On the other hand, the **TS2** was located at $d_3=1.94$ Å and $d_4=1.49$ Å. The calculated release potential energy barrier of the tetrahedral intermediate was 12.6 kcal/mol.

Compared with the AMB-TE-Amb system, the energy barrier of the macrocyclization catalyzed by AMB-TE in AMB-TE-Nys was much higher, which was calculated to be 22.7 kcal/mol. Besides, the product seemed quite unstable, indicating the reversible reaction (*i.e.*, the reconnection of the intermediate state) could easily occur.

3.4.2. NYS-TE system

Beside the AMB-TE system, we calculated potential energy surfaces of the NYS-TE system. Similar to that in AMB-TE system, the potential energy surfaces were calculated through (1) defining the $d_1(\text{O}_{11}-\text{C}_1)$ and $d_2(\text{O}_{11}-\text{H})$ distances as the reaction coordinates to obtain the energy barrier of formation of the tetrahedral intermediate; and (2) treating the distance $d_3(\text{O}_{\text{ser}}-\text{C}_1)$ and $d_4(\text{O}_{\text{ser}}-\text{H})$ as the reaction coordinates to the product release and the refresh of Serine. The key structures (**TS1** and **TS2**) and the reaction pathways were shown in Fig. 9 and Fig. 10.

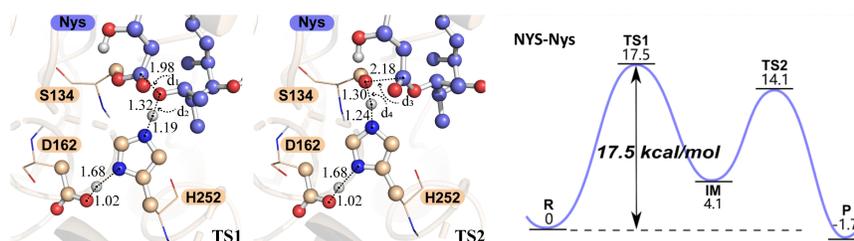


Fig. 9. Transition states and energy barriers of macrocyclization in NYS-TE-Nys systems. Key residues in NYS TE are shown in yellow and Nys is in blue.

In the NYS-TE-Nys system, the transition state **TS1** was located at $d_1=1.98 \text{ \AA}$ and $d_2=1.32 \text{ \AA}$, and **TS2** was obtained at $d_3=2.18 \text{ \AA}$ and $d_4=1.30 \text{ \AA}$. The optimized structures of **R**, **IM**, and **P** were shown in Fig. S11+. The energy barrier for the first step was calculated to be 17.5 kcal/mol, indicating that the formation of a tetrahedral intermediate could proceed spontaneously. Also, the energy barrier for the second step was 9.9 kcal/mol, suggesting the release of cyclized product was favorable.

In short, the nucleophilic attack to the formation of tetrahedral intermediate was determined to be the rate-determining step, and the energy barrier was also in good agreement with that reported energy barrier of picromycin TE-catalyzed macrocyclization[25]. These results indicated that the macrocyclization catalyzed by NYS TE with its natural substrate Nys could favorably take place in both thermodynamic and kinetic aspects.

Next, in NYS-TE-Amb systems, the transition state **TS1** was located at the $d_1=1.90 \text{ \AA}$ and $d_2=1.51 \text{ \AA}$. The optimized structures of **R**, **IM**, and **P** were shown in Fig. S12+. The calculated potential energy barrier of formation of the tetrahedral intermediate was 18.9 kcal/mol, which was 1.4 kcal/mol higher than Nys substrate. Also, the **TS2** was located at $d_3=2.14 \text{ \AA}$ and $d_4=1.41 \text{ \AA}$. The release energy barrier of the cyclized product was 8.4 kcal/mol and the whole energy barrier of the macrocyclization catalyzed by NYS-TE was 25.7 kcal/mol.

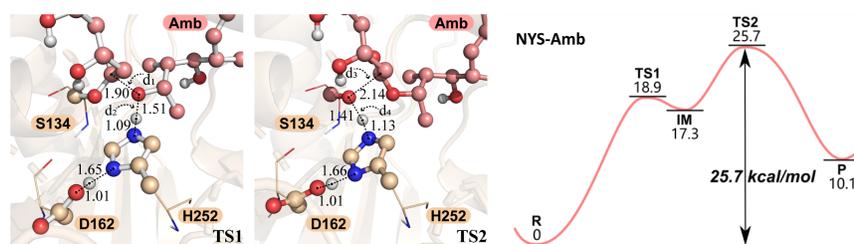


Fig. 10. Transition states and energy barriers of macrocyclization in NYS-TE-Amb systems. Key residues in NYS TE are shown in yellow and Amb is in salmon.

Taken together, the calculated energy barriers indicated that the macrocyclization with its natural substrate proceeded more easily than with its unnatural substrate in both AMB-TE and NYS-TE systems. Besides, the unnatural cyclized products (either in AMB-TE-Nys or NYS-TE-Amb system) were more favorable for the reversible reaction, indicating they were energetic disadvantages in the formation of unnatural products.

3.5. Product Release in AMB-TE and NYS-TE systems

Three times of MD simulations were performed with AMB-TE/NYS-TE and their macrolactone products. The initial structures were received from QM/MM calculations. The binding energies for the four systems were calculated with both MM-GBSA and MM-PBSA. The average binding energies obtained by MM-PBSA were -30.1 kcal/mol in AMB-TE-Amb, -37.3 kcal/mol in AMB-TE-Nys, -27.0 kcal/mol in NYS-TE-Nys and -37.6 kcal/mol in NYS-TE-Amb. This energy trend was as same as that calculated by MM-GBSA (Table S3). These data indicated that both AMB TE and NYS TE had a stronger binding ability to unnatural products than their natural products. In summary, the calculations of binding energy with both MM-GBSA and MM-PBSA methods demonstrated that in comparison to unnatural products, natural substrates are easier to leave TE enzymes after macrocyclization, because of their uncomfortable binding state and unfavorable binding free energy in TE pockets.

Careful analyses were conducted on the disengagement of product from the active site in the four systems. After cyclization, the product would stay in the vicinity of the active site for a while due to VDW (Van der Waals' force) and electrostatic interactions from peripheral residues.

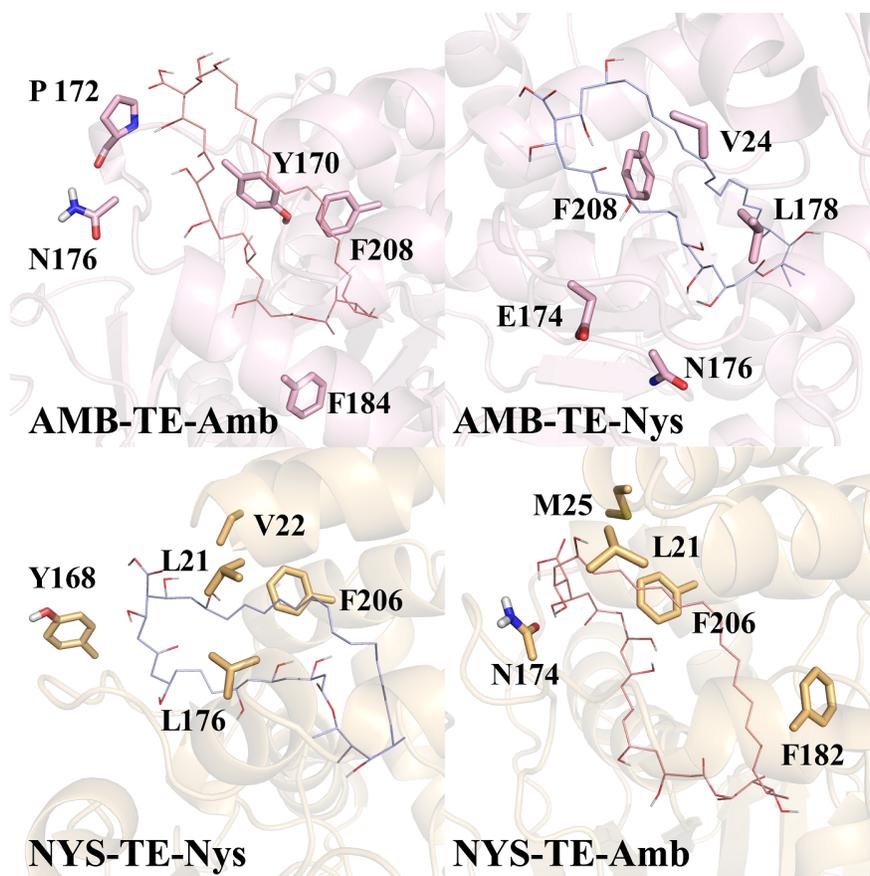


Fig. 11. Key residues in TE pocket in enzyme-product systems. AMB TE is shown in pink and NYS TE is shown in yellow.

In AMB-TE system (Fig. 11), the cyclized Amb product was deformed into a slim ring by F208, P172, E174, N176 and so on (Table S4), while Nys product was more flexible due to a carbon-carbon single bond in its hydrophobic side, which made it easier to get combined with F208 and L178. Therefore, Amb product left AMB TE more easily than Nys product after macrocyclization. On the other hand, things changed to a different situation in NYS TE. There were more hydrophobic amino acids close to product in NYS TE's active site than those in AMB TE. In this case, the flexibility of Nys product became its advantage for leaving TE. With the help of L21, Y168, L176, and F206, Nys product transformed to the suitable conformations, which is helpful for leaving NYS TE. Similar hydrophobic residues were reported in pikromycin TE [42].

4. Conclusion

In this study, we investigated the catalytic overview of AMB TE and NYS TE including substrate recognition, catalytic mechanism, and product release. Firstly, 5×50 ns MD simulations on four systems (AMB-TE-Amb, AMB-TE-Nys, NYS-TE-Nys, and NYS-TE-Amb) were performed for conformational search. Based on the Pre-reaction states (PRSs) calculations, we found that NYS TE showed better recognition towards its natural substrate Nys than Amb, while AMB TE showed good tolerance to both Amb and Nys. Furthermore, R26 of NYS TE was discovered to be responsible for locking the head of substrate Nys, while D186 prevented the substrate from turning the tail part upside down. They played a critical role in substrate recognition. Then the representative PRS structures from the dominant cluster obtained from MD simulations were optimized with QM/MM calculations to obtain the catalytic mechanism of macrocyclization. The energy barrier of cyclization in AMB-TE-Amb was calculated to be 14.0 kcal/mol, and in AMB-TE-Nys was 22.7 kcal/mol. Also, in NYS-TE-Nys and NYS-TE-Amb, their energy barriers were 17.5 kcal/mol and 25.7 kcal/mol, respectively. Our computational results indicated that the macrocyclization catalyzed by both TEs with their natural substrates were more favorable than that with exchanged substrates. Ultimately, the MM-PBSA program was employed to unveil residues mediating product release. The average MM-PBSA and MM-GBSA energies suggested that it was easier for natural products to leave TE enzymes than unnatural products after macrocyclization. And key residues playing a part in the departure of polyketide products were uncovered. We provided an overview of AMB TE and NYS TE catalysis with computational methods, improved the comprehension of polyene polyketide TEs, and offered conducive opinions for highly-productive drug design.

Corresponding Author

* Tel/Fax: +86-21-34207347. Email: tshi@sjtu.edu.cn.

Author Contributions

TS, LQB and YLZ conceived and designed the investigation. WRF, TWT, LC and LL performed calculations and analyses. WRF, and TS wrote up the paper.

Conflicts Notes

The authors declare no competing financial interest.

Acknowledgments

The authors thank the National Science Foundation of China (21377085, 31970041 and 31770070), The National Key R&D Program of China (2019YFA0905400, 2018YFA0901200), 19ZR1427300, SQ2018YFA090024, ZH2018ZDA26, and the SJTU-HPC computing facility award for financial support and computational time.

References

1. Caffrey, P., Aparicio, J. F., Malpartida, F., Zotchev, S. B. *Biosynthetic engineering of polyene macrolides towards generation of improved antifungal and antiparasitic agents*. Current topics in medicinal chemistry, 2008.8 : p. 639-653.

2. Caffrey, P., De Poire, E., Sheehan, J., Sweeney, P. *Polyene macrolide biosynthesis in streptomycetes and related bacteria: recent advances from genome sequencing and experimental studies*. Applied microbiology and biotechnology., 2016. **100** : p. 3893-3908.
3. Belakhov, V. V., Garabadzhiu, A. V. *Polyene macrolide antibiotics: Mechanisms of inactivation, ways of stabilization, and methods of disposal of unusable drugs*. Russian Journal of General Chemistry, 2015. **85** : p. 2985-3001.
4. Ruhnke, M. Schwartz, S. *Recent developments in the management of invasive fungal infections in patients with oncohematological diseases*. Ther Adv Hematol., 2016. **7** : p. 345-359.
5. Donovan, R., Gold, W., Pagano, J. F., Stout, H. A. *Amphotericins A and B, antifungal antibiotics produced by a streptomycete. I. in vitro studies*. Antibiotics annual, 1955. **3** : p. 579-586.
6. Chudzik, B., Koselski, M., Czuryło, A., Trebacz, K., Gagoś, M. *A new look at the antibiotic amphotericin B effect on Candida albicans plasma membrane permeability and cell viability functions*. European Biophysics Journal., 2015. **44** : p. 77-90.
7. Moen, M. D., Lyseng-Williamson, K. A., Scott, L. J. *Liposomal amphotericin B*. J. Drugs., 2009. **69** : p. 361-392.
8. Davis, S. A., Vincent, B. M., Endo, M. M., Whitesell, L., Marchillo, K., Andes, D. R., Burke, M. D. *Nontoxic antimicrobials that evade drug resistance*. Nature chemical biology, 2015. **11** : p. 481.
9. Mazerski, J., Grzybowska, J., Borowski, E. *Influence of net charge on the aggregation and solubility behaviour of amphotericin B and its derivatives in aqueous media*. European Biophysics Journal, 1990. **18** : p. 159-164.
10. Baginski M, Resat H, Borowski E. *Comparative molecular dynamics simulations of amphotericin B-cholesterol/ergosterol membrane channels*. Biochim Biophys Acta., 2002. **1567** : p. 63–78.
11. Ostrosky-Zeichner, L., Marr, K. A., Rex, J. H., Cohen, S. H. *Amphotericin B: time for a new “gold standard”*. Clinical infectious diseases, 2013. **37**: p. 415-425.
12. Gray KC, Palacios DS, Dailey I, Endo MM, Uno BE, Wilcock BC, Burke MD. *Amphotericin primarily kills yeast by simply binding ergosterol*. Proc Natl Acad Sci USA, 2010. **109** : p. 2234–2239.
13. Chéron M, Cybulska B, Mazerski J, Grzybowska J, Czerwin ´ski A, Borowski E. *Quantitative structure-activity relationships in amphotericin B derivatives*. Biochem Pharmacol, 1988. **37** : p. 827–836.
14. Larson JL, Wallace TL, Tyl RW, Marr MC, Mayers CB, Cossum PA. *The reproductive and development toxicity of the antifungal drug Nyotran (liposomal Nystatin) in rats and rabbits*. Toxicol Sci, 2000. **53** : p. 421–429.
15. Pierce CG, Lopez-Ribot JL. *Candidiasis drug discovery and development: new approaches targeting virulence for discovering and identifying new drugs*. Expert Opin Drug Discov, 2013. **9** : p. 1117–1126.
16. Boros-Majewska, J., Salewska, N., Borowski, E., Milewski, S., Malic, S., Wei, X.Q., Hayes, A.J., Wilson, M.J., Williams, D.W. *Novel nystatin A1 derivatives exhibiting low host cell toxicity and antifungal activity in an in vitro model of oral candidosis*. Med Microbiol Immunol, 2014. **203** : p. 341-355.
17. Semis E, Nili SS, Munitz A, Zaslavsky Z, Polacheck I, Segal E. *Pharmacokinetics, tissue distribution and immunomodulatory effect of intralipid formulation of Nystatin in mice*. J Antimicrob Chemother, 2012. **67** : p. 1716–1721.
18. Semis R, Kagan S, Berdicevsky I, Polacheck I, Segal E. *Mechanism of activity and toxicity of Nystatin-intralipid*. Med Mycol, 2013. **51** : p. 422–431.
19. Kaminski DM, Czernel G, Murphy B, Runge B, Magnussen OM, Gagos ´ M. *Effect of cholesterol and ergosterol on the antibiotic amphotericin B interactions with dipalmitoylphosphatidylcholine monolayers: X-ray reflectivity study*. Biochim Biophys Acta, 2014. **1838** : p. 2947–2953.
20. 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** : p. 6288-6290.
21. Meier, J. L., Burkart, M. D. *The chemical biology of modular biosynthetic enzymes*. Chem. Soc. Rev, 2009. **38** : p. 2012-2045.
22. Williams, G. J. *Engineering polyketide synthases and nonribosomal peptide synthetases*. Curr. Opin. Struct. Biol, 2013. **23** : p. 603- 612.

23. Hansen D A, Koch A A, Sherman D H. *Identification of a thioesterase bottleneck in the pikromycin pathway through full-module processing of unnatural pentaketides*. Journal of the American Chemical Society, 2017. **139** : p. 13450-13455.
24. Du, L., Lou, L. *PKS and NRPS release mechanisms*. Nat. Prod. Rep. 2010. **27** : p. 255-278
25. Shi, T., Liu, L., Tao, W., Luo, S., Fan, S., Wang, X. L., Bai, L. Q., Zhao, Y. L. *Theoretical studies on the catalytic mechanism and substrate diversity for macrocyclization of pikromycin thioesterase*. ACS Catalysis., 2018. **8** : p. 4323-4332.
26. Chen, X. P., Shi, T., Wang, X. L., Wang, J., Chen, Q., Bai, L. Q., Zhao, Y. L. *Theoretical studies on the mechanism of thioesterase-catalyzed macrocyclization in erythromycin biosynthesis*. ACS Catalysis, 2016. **6** : p. 4369-4378.
27. 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*. Nature Chemical Biology, 2006. **2**: p. 531-536.
28. 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*. Journal of chemical information and modeling, 2007. **38** : p. 1923-1932.
29. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox. *GAUSSIAN09. Revision E. 01*. Gaussian Inc., Wallingford, CT, USA, 2009.
30. Jakalian, A., Bush, B. L., Jack, D. B., Bayly, C. I. *Fast, efficient generation of high-quality atomic Charges. AM1-BCC model: I. Method*. Journal of Computational Chemistry, 2000. **21** : p. 132-146.
31. Case, D. A., Babin, V., Berryman, J., Betz, R. M., Cai, Q., Cerutti, D. S., Cheatham III, T. E., Darden, T. A., Duke, R. E., Gohlke, H., Goetz, A. W., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossvary, I., Kovalenko, A., Lee, T. S., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, K. M., Paesani, F. Roe, D. R., Roitberg, A., Sagui, C., Ferrer, R. S., Seabra, G., Simmerling, C. L., Smith, W., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., Kollman, P. A. *AMBER 14*. University of California, San Francisco, CA, 2014.
32. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., Klein, M. L. *Comparison of simple potential functions for simulating liquid water*. Journal of Chemical Physics, 1983.**79** : p. 926-935.
33. 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** : p. 327-341.
34. Darden, T., York, D., Pedersen, L. *Particle Mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems*. Journal of Chemical Physics, 1993. **98** : p. 10089-10092.
35. Vreven, T., Byun, K. S., Komaromi, I., Dapprich, S., Montgomery Jr, J. A., Morokuma, K., Frisch, M. J. *Combining quantum mechanics methods with molecular mechanics methods in ONIOM*. Journal of Chemical Theory & Computation, 2006. **2** : p. 815-826.
36. Vreven, T., Frisch, M., Kudin, K., Schlegel, H., Morokuma, K. *Geometry optimization with QM/MM methods II: Explicit quadratic coupling*. Molecular Physics, 2006. **104** : p. 701-714.
37. 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*. Theoretical Chemistry Accounts, 2008.**120** : p. 215-41.

38. Rassolov, V. A., Ratner, M. A., Pople, J. A., Redfern, P. C., Curtiss, L. A. *6-31G* Basis Set for Third-Row Atoms*. Journal of Computational Chemistry, 2001. **22** : p. 976-84.
39. Bing Mei, S., Shao, Z. H., Li, A. P., Naeem, M., Lin, J., Ye, L. D., Yu, H. W. *Rational design of dehydrogenase/reductases based on comparative structural analysis of prereaction-state and free-state simulations for efficient asymmetric reduction of bulky aryl ketones*. ACS Catalysis, 2020. **10** : p. 864-876.
40. Koch, A. A., Hansen, D. A., Shende, V. V., Furan, L. R., Houk, K. N., Jimenez-Oses, G., Sherman, D. H. *A single active site mutation in the pikromycin thioesterase generates a more effective macrocyclization catalyst*. Journal of the American Chemical Society, 2017. **139** : p. 13456-13465.
41. Liu, L., Tao, W. T., Bai, L. Q., Kim, E. S., Zhao, Y. L., Shi, T. *Why does tautomycin thioesterase prefer hydrolysis to macrocyclization? Theoretical study on its catalytic mechanism*. Catalysis Science & Technology, 2019. **9** : p. 6391-6403.
42. 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*. Nature chemical biology, 2006. **2** : p. 537-542.

