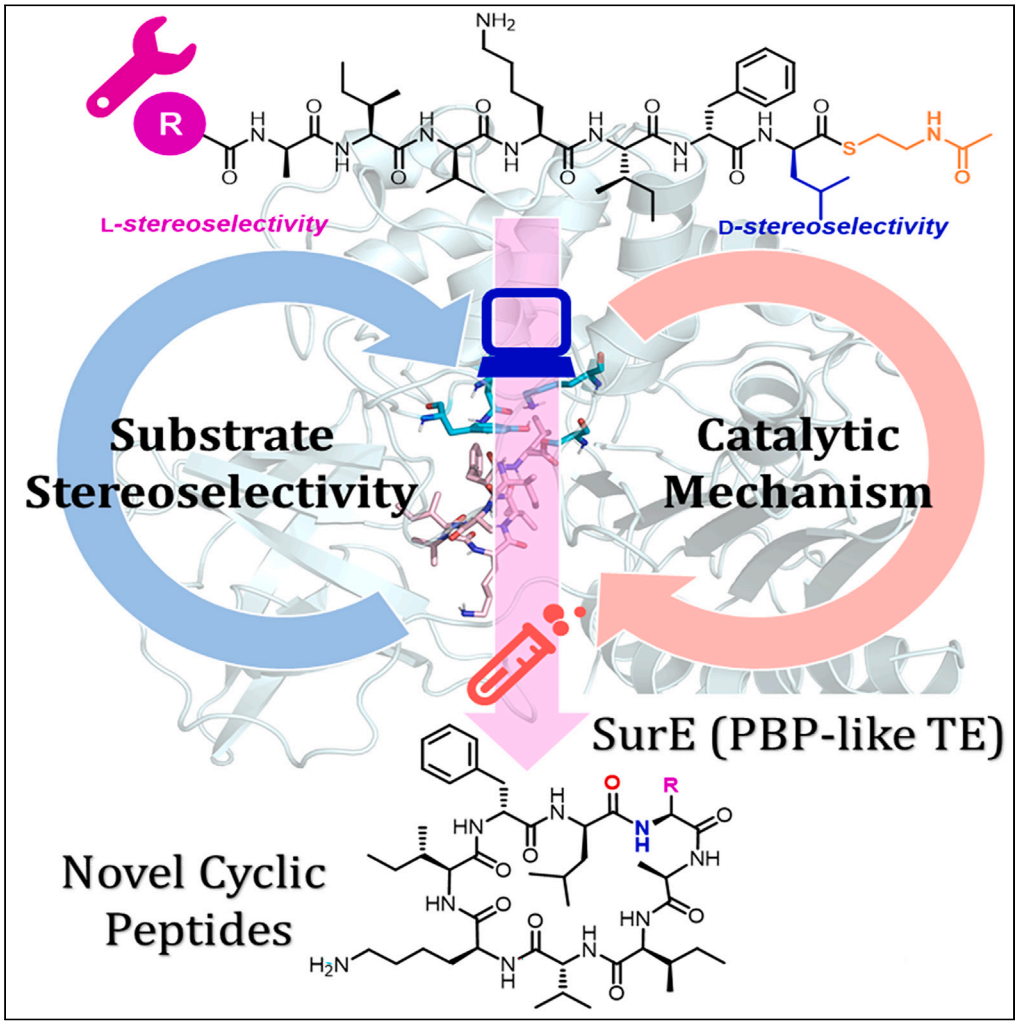
iScience

CellPress
OPEN ACCESS

Article

Exploring the substrate stereoselectivity and catalytic mechanism of nonribosomal peptide macrocyclization in surugamides biosynthesis



Zeqian Du, Yinhao Ma, Yaoyao Shen, Xuefeng Jiang, Yongjun Zhou, Ting Shi

xfjiang@chem.ecnu.edu.cn (X.J.)
zhouyongjun66@163.com (Y.Z.)
tshi@sjtu.edu.cn (T.S.)

Highlights

Exploring the molecular mechanism was critical for heterochiral coupling in NRPs

An anion hole played a pivotal role in facilitating *C*-terminal recognition

The hydrogen-bond network was instrumental in *N*-terminal recognition and cyclization

This work will guide engineering PBP-like TEs and designing a variety of novel NRPs

Du et al., iScience *27*, 108876
February 16, 2024
https://doi.org/10.1016/j.isci.2024.108876

iScience

CellPress
OPEN ACCESS

Article

Exploring the substrate stereoselectivity and catalytic mechanism of nonribosomal peptide macrocyclization in surugamides biosynthesis

Zeqian Du,[1,4] Yinhao Ma,[2,4] Yaoyao Shen,[3,4] Xuefeng Jiang,[2,*] Yongjun Zhou,[3,*] and Ting Shi[1,5,*]

SUMMARY

SurE, the first reported penicillin-binding protein-like thioesterase (PBP-like TE), is known as a new off-loading cyclase, which catalyzes heterochiral coupling in nonribosomal peptides (NRPs). However, the structural rationale for substrate stereoselectivity and enzymatic mechanism remains mysterious. Here, computational models, integrating MD simulations and QM/MM methods, unveiled SurE's substrate recognition and catalytic process. An oxyanion hole stabilized the *C*-terminal D-residue during recognition. Residue R446 anchored the substrate for macrocyclization. A vital hydrogen-bonding network (Y154, K66, N156), verified by mutation results, was responsible for the recognition of *N*-terminal L-residue and involvement in catalytic process with a calculated 19.4 kcal/mol energy barrier. Four novel-designed peptide precursors were effectively cyclized into cyclopeptides by SurE based on computational analysis. Our results provide a comprehensive understanding of SurE's catalytic mechanism and guiding design of versatile PBP-like TEs for novel macrocyclic NRPs.

INTRODUCTION

Nonribosomal peptides (NRPs) exhibit diverse biological activities and promising therapeutic potential.[1–3] Compared with linear peptides, cyclic peptides display additional desirable properties such as proteolytic stability, membrane permeability, and binding affinities.[4–6] Therefore, cyclic peptide backbones are widely used in pharmacology, from antibiotics (daptomycin, vancomycin) and immunosuppressants (cyclosporin A) to anticancer agents (bleomycin A2) and anthelmintics (avermectin).[3,7]

Generally, cyclic peptides are biosynthesized by multi-modular enzyme complexes known as NRP synthases (NRPSs). The NRPS modules mainly consist of the adenylation (A), peptidyl carrier protein (PCP), and condensation (C) domains.[8] Besides, some non-essential tailoring domains are responsible for various structural modifications for NRPs, like methyltransferase, epimerization, and oxidation/reduction domains.[9] An off-loading domain, usually a type I thioesterase (TE-I) or a condensation-like (C_T) domain, embedded in the *C* terminus of most NRPS complexes,[10,11] catalyzes the release of the final peptide via cyclization or hydrolysis reaction.[12] Because of their key roles in substrate macrocyclization and product release, engineering TEs for the production of non-natural macrocyclic NRPs is a promising approach toward preparing novel bioactive materials.

Penicillin-binding protein-like thioesterases (PBP-like TEs) are a new family of off-loading cyclases that catalyze the head-to-tail macrocyclization of a series of NRPs ranging from cyclic pentapeptides to decapeptides,[13–15] such as pentaminomycin, surugamides, desotamides, ulleungmycins, noursamycins, and mannopeptimycins (Figure S1). Distinguishable from commonly embedded cyclases like TE-I and C_T domains, which offload products by *cis*-acting, PBP-like TEs release peptides by *trans*-acting as discrete domains[16] (Figure 1A). In addition, PBP-like TEs only specifically catalyze the macrocyclization between *N*- and *C*-terminal residues with L- and D-configurations, respectively, while the *cis*-acting offloading domains accomplish heterochiral coupling with the opposite configuration (D-amino acid at the *N*-terminus and L-amino acid at the *C*-terminus).[17] Therefore, these specific characteristics of PBP-like TEs can be exploited for various applications, which are certainly not limited to combining module swapping with enzymatic transformation to produce novel non-native cyclic peptides.

As the first member of the PBP-like TE family, SurE was discovered in the biosynthetic gene cluster of surugamides (A-F) and isolated from *Streptomyces* (sp. JAMM992).[18] SurE has been validated in charge of the biosynthesis of cyclic octapeptide surugamides A-E and linear decapeptide surugamide F through two discrete NRPS assembly lines and generate cyclic peptides with heterochiral coupling[19] (Figure 1B). Matsuda et al. recently reported the crystal structures of open *apo*-SurE (6KSU) and closed *holo*-SurE with substrate derivative (6KSV), where

[1]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, China
[2]Shanghai Key Laboratory of Green Chemistry and Chemical Process, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200062, China
[3]Research Center for Marine Drugs, State Key Laboratory of Oncogenes and Related Genes, Department of Pharmacy, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China
[4]These authors contributed equally
[5]Lead contact
*Correspondence: xfjiang@chem.ecnu.edu.cn (X.J.), zhouyongjun66@163.com (Y.Z.), tshi@sjtu.edu.cn (T.S.)
https://doi.org/10.1016/j.isci.2024.108876

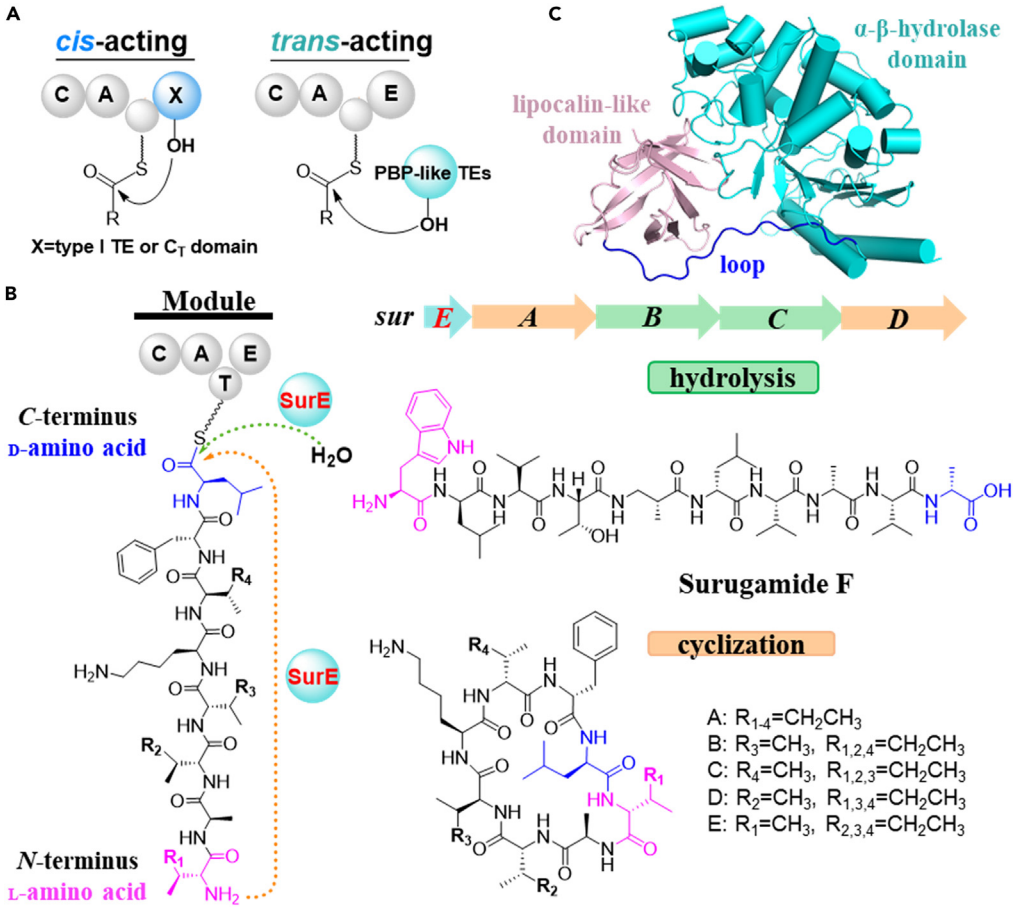


Figure 1. PBP-like TE SurE in surugamides biosynthesis

(A) Off-loading patterns in NRPS via *cis*-acting (TE-I or CT domains) or *trans*-acting (PBP-like TEs).

(B) The *surE* gene cluster encodes two discrete NRPS assembly lines where the SurE enzyme is responsible for the biosynthesis of cyclic octapeptide surugamides A-E (cyclization) and linear decapeptide surugamide F (hydrolysis).

(C) The crystal structure of SurE (PDB: 6KSU). Lipocalin-like, α-β-hydrolase domain and loop are colored pink, cyan, and blue, respectively.

a critical loop (residues 211–224) requires significant conformational changes from open to closed form for performance of peptide macrocyclization[20] (Figure S2). SurE is composed of three main parts including a typical α-β-hydrolase domain (residue 1–350), a lipocalin-like domain (residue 368–451), and a long-flexible loop (residue 351–367) between these two domains[21] (Figure 1C). Extensive experimental data both *in vivo* and *in vitro* were obtained by Zhou and Matsuda et al.,[21,22] demonstrating that SurE not only exhibits a strict heterochiral preference for the *N*- and *C*-terminal residues of substrates but also possesses broad substrate tolerance. Different cyclic peptides can be generated by changing the number or sequence of residues in substrates. Consequently, SurE is expected to be utilized as biocatalysts and genetic tools for the biosynthesis of a series of new cyclic peptides and their derivatives with desired pharmaceutical bioactivities.[23]

Although a two-step mechanism (loading and releasing step) has been generally accepted for TE-catalyzed product release,[24] the mechanistic insights of macrocyclization reaction catalyzed by PBP-like TEs remain obscure. Due to substrate diversity and stereoselectivity, the catalytic mechanism of specific heterochiral coupling will be more complicated than traditional TE-mediated macrocyclization. What's more, the conserved catalytic tetrad S/Y/H/K in PBP-like TEs has been assumed to be involved with the catalysis process[13,22] (Figure S3); nevertheless, their exact biological functions have not been identified. The strict heterochiral requirement for the configuration of residues in both termini of substrates during SurE catalysis deserves further detailed investigations. However, it is difficult to ensure that macrocyclization occurs at the designated residue, because side-chain-to-tail cyclization may be more favorable than head-to-tail cyclization, and intermolecular oligomerizations usually occur in competition with intramolecular macrocyclization.[25] As a result, it is still a challenging task to facilitate the design, prediction, and achievement of highly specific new drugs via PBP-like TEs.

In this work, we built a series of computational models for substrate-SurE complexes to understand substrate stereoselectivity and molecular mechanism of the macrocyclization catalyzed by PBP-like TEs. Based on the computational results, the mechanism of SurE-catalyzed macrocyclization was put forward, and the roles of key residues were confirmed by mutation experiments. Finally, the designed precursor peptides were successfully utilized to generate novel cyclic peptides. Our work identifies the important residues involved in both substrate

Table 1. Computational models with peptidyl-SNAC substrates

No	Model	The sequence of peptidyl-SNAC substrates
1	SNAC-S	L-Ile$_1$-D-Ala$_2$-L-Ile$_3$-D-Val$_4$-L-Lys$_5$-L-Ile$_6$-D-Phe$_7$-D-Leu$_8$-SNAC
2	SNAC-D$_1$	**D-*allo*-Ile$_1$**-D-Ala$_2$-L-Ile$_3$-D-Val$_4$-L-Lys$_5$-L-Ile$_6$-D-Phe$_7$-D-Leu$_8$-SNAC
3	SNAC-L$_8$	L-Ile$_1$-D-Ala$_2$-L-Ile$_3$-D-Val$_4$-L-Lys$_5$-L-Ile$_6$-D-Phe$_7$-**L-Leu$_8$**-SNAC

The bolded residue is different from that in the SNAC-S model.

recognition and macrocyclization, provides a comprehensive overview of the catalytic mechanism of SurE, and sheds light on the design of versatile PBP-like TEs or TE domains for the development of diverse macrocyclic NRP analogs.

RESULTS

Pivotal *C*-terminal stereoselectivity in substrate loading process

As a discrete cyclase, SurE has been identified to catalyze the release of surugamides by a two-step process like common TE domains. Therefore, our study starts from the loading step, where the nucleophilic active site Ser63 of SurE attacks the carbonyl group of the thioester of the peptidyl-*S*-SNAC (N-acetyl cysteamine thioesters), which was utilized to mimic the PCP-tethered precursor, to form peptidyl-*O*-SurE and release the SNAC.

Herein, three complex models were constructed to investigate the substrate stereoselectivity in *C*- and *N*-terminal residues at the loading step. They were named as SNAC-S, SNAC-D$_1$, and SNAC-L$_8$ where SurE adopted an open form. The linear peptide precursor of surugamide B was utilized for the SNAC-S model and the SurE-D$_1$ variant with *N*-terminal L-Ile$_1$ substituted by D-*allo*-Ile$_1$ as well as SurE-L$_8$ variant with *C*-terminal D-Leu$_8$ substituted by L-Leu$_8$ (as shown in Table 1 and Figure 2A).

Three times 50 ns molecular dynamic (MD) simulations were carried out and the loading distance d(C1-O$^\gamma$), which represents the nucleophilic attack by O$^\gamma$ of Ser63 onto C1 of peptidyl-*S*-SNAC, was measured in all trajectories. As exhibited in Figure 2B, the distance distributions were highly concentrated in SNAC-S and SNAC-D$_1$ models, indicating the conformations with D-configured Leu$_8$ were stable and suitable for the catalytic pocket, that is, the conformations were competent for substrate recognition at the loading step. Conversely, the distance distribution in the SNAC-L$_8$ model with L-Leu$_8$ was significantly scattered, even up to 9.0 Å, suggesting the peptidyl-SNAC substrate with *C*-terminal L-configuration was unstable and unattainable to approaching the Ser63 of SurE. The substrate seemed to be trapped in unfavorable conformations due to the steric hindrance caused by the L-Leu$_8$ residue. These results indicated that the *C*-terminal stereoselectivity, rather than the *N*-terminal, played a key role in the substrate loading process.

Stabilization of *C*-terminal D-residue by oxyanion hole in substrate recognition

Apart from the conformational differences at the substrate loading step, we also explored the impact of *C*-terminal stereoselectivity in substrate recognition. Here, the peptide substrates in the peptidyl-*O*-SurE adduct were covalently tethered to Ser63 after substrate loading, where a nearby flexible loop of SurE (residues 211–224) induced fit to covering the substrate (closed form).

Two catalytic models, SurE-S and SurE-L$_8$, were established, where SurE-S represents the natural enzyme-substrate complex compared with the artificial SurE-L$_8$ (Table 2, Row 1–2). The substitution of D-Leu$_8$ to L-Leu$_8$ in the peptidyl-SNAC substrate resulted in no detectable reaction activity in the previous SurE assay, providing a valid reference to the SurE-L$_8$ model established in this study.[22,23]

To ensure sufficient conformational sampling, four separate independent simulations of 100 ns each were performed for both the SurE-S and SurE-L$_8$ models. In both models, the typical oxyanion hole of TE families was formed through two backbone amide N–H bonds of Ser63 and Ala308 (Figure 3A and Table S1). The SurE-S model showed that H-bond1 was maintained for 61.9% in all trajectories, while H-bond2 was maintained for 88.1%. However, in the SurE-L$_8$ model, the proportion of H-bond1 and H-bond2 decreased sharply to 12.8% and 52.7%, respectively, indicating better performance of the oxyanion hole in stabilizing the *C*-terminal D-residue after substrate loading (Figure 3B). The oxyanion hole is expected to stabilize the anionic peptidyl-*O*-SurE tetrahedral adduct and facilitate proper orientation for nucleophilic attack during macrocyclization. Besides, the MD simulations of SNAC-S and SurE-S models demonstrated that the frequency of water occurrence within 5 Å of O$^\gamma$ from S63 accounted for 57.6% and 5.0%, respectively, indicating the water molecules barely existed in the catalytic pocket after the substrate loading.

Crucial role of R446 for formation of pre-organization states

After the linear substrate was loaded onto SurE, it could happen to conformationally pre-organize into the ring-like structure with the assistance of key residues. Here, the ring-like conformations with attacking distance between head to tail (N8$_{Ile1}$-C1$_{Leu8}$) within 4.5 Å were defined as pre-organization states (POSs)[26–28] in our models, which is a prerequisite for the liner substrate to macrocyclization.

To identify the key residues responsible for maintaining the POSs, MM-GBSA binding free energy decomposition was introduced based on the trajectories of the SurE-S model. The residues contributed more than 1.0 kcal/mol were listed. Among these, R446 ranked first (up to 4.1 kcal/mol). Additionally, residues K66, N156, Y157, A308, and V309 contributed 3.1, 1.3, 1.3, 1.1, and 1.1 kcal/mol to the substrate binding, respectively (Figure S4). R446 was found to lock the tail of the substrate at a favorable conformation for macrocyclization through the hydrogen

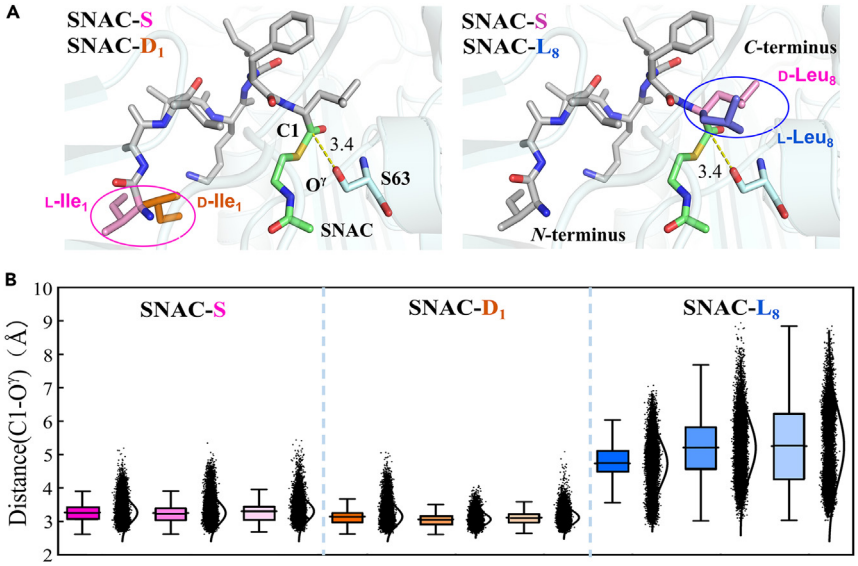


Figure 2. The loading conformations with different peptidyl-SNAC substrates

(A) The comparison of three different substrates with the same initial loading distance of 3.4 Å.

(B) The boxplot of substrate loading distance distributions obtained by each MD simulation.

bonds connecting its guanidyl group with the main-chain carbonyl groups ($O7_{sub}$ of D-Ala_2 and $O8_{sub}$ of L-Ile_1) in the substrate's *N* terminus (Figure 4A). The total proportions of these hydrogen bonds between R446 and substrate were 85.7%, 45.7%, and 15.4% in SurE-S (Figures 4B and Table S2). What's more, the total proportion of POSs in all SurE-S trajectories was calculated to be 97.3%, indicating the high stability of ring-like POSs (Figures 4C and S5–S7).

To further validate the significance of R446, we conducted comprehensive MD simulations for the R446A model, as depicted in Figure 4. The hydrogen bonds between R446A and the substrate were totally destroyed, and the proportion of POSs in trajectories of R446A was only 9.6% (Figures 4C and S5–S7). These findings revealed that in the absence of R446, the distance between the N8 of L-Ile_1 (head) and the C1 of D-Leu_8 (tail) in the peptide substrate exhibited significant and erratic fluctuations, leading to the instability and disappearance of POSs, which critically dependent on the hydrogen bonds formed between the guanidyl group of R446 and the carbonyls of both D-Ala_2 and L-Ile_1.

Precise hydrogen-bonding network for *N*-terminal stereoselectivity

Subsequently, we established the SurE-D_1 model, in which a sole modification was implemented: the natural L-Ile_1 in the SurE-S model was substituted with D-*allo*-Ile (as indicated in Table 2, Row 3). This modification was undertaken with the specific objective of probing the stereoselective preferences associated with the *N*-terminal substrate residue. We conducted four independent sets of MD simulations for the SurE-D_1 model to elucidate the structural divergence arising during the catalytic process of SurE via comparative analysis with SurE-S. It should be noted that the loading distance and the oxyanion hole were remarkably similar in both SurE-D_1 and SurE-S models because of the same *C*-terminal D-Leu_8 residue (Figure S8).

The determination of key interatomic distances often serves as a valuable metric to assess the possibility of catalytic reactions within the enzyme-substrate complex. To reveal the influence of chiral residues at the *N*-terminus of the substrate on the catalytic cyclization reaction mediated by the SurE enzyme, two key reaction distances, namely d1(C1-N8) and d2(N8-O_{Y154}), in the SurE-S and SurE-D_1 model were selected for comparison and utilized to evaluate the possibility of reactions (Figure 5A). The results indicated that, in the SurE-S model, the proportion of distances both d1(C1-N8) and d2(N8-O_{Y154}) less than 3.5 Å was 80.3%; however, in the SurE-D_1 model, the proportion dramatically decreased to zero (Figure 5B), which was consistent well with the experimental evidence that SurE was able to cyclize the substrate with L-residue at *N*-terminal while failed to cyclize the substrate with D-residue at *N*-terminal.

Finally, structural analysis showed that a precise closed hydrogen-bond network played an important role in the recognition of *N*-terminal L-residue, which included H-bond6 to H-bond9 (Figures 5C and Table S3). H-bond6, connecting $N8_{sub}$ and O_{Y154}, could maintain 29.6% during the simulations in the SurE-S model, while it barely existed in the SurE-D_1 model. H-bond7 appeared more stable in SurE-S (98.1%) than in the SurE-D_1 (66.0%) model. Different from H-bond8, which had a similar proportion in SurE-S (48.5%) and SurE-D_1 (42.3%) model, the proportion of H-bond9, connecting $O8_{sub}$ and N156, was calculated to be 62.4% in SurE-S model, while only 4.4% in SurE-D_1 model (Figure 5D). Residues involved in forming the aforementioned hydrogen bonds (K66, Y154, and N156) were also highlighted by free energy calculation. These findings illustrated that the presence of a D-residue at the *N*-terminus significantly disrupted the hydrogen-bonding network. Conversely, in the case of an *N*-terminal L-residue, the precise and stable hydrogen bonds were observed to play an essential role in the preparation of pre-reaction states, facilitating subsequent macrocyclization (Figures S9 and S10).

Table 2. Computational models with different peptidyl-SurE substrates

No	Model	The sequence of peptidyl-SurE substrates
1	SurE-L_8	L-Ile$_1$-D-Ala$_2$-L-Ile$_3$-D-Val$_4$-L-Lys$_5$-L-Ile$_6$-D-Phe$_7$-**L-Leu$_8$**-Ser63
2	SurE-S	L-Ile$_1$-D-Ala$_2$-L-Ile$_3$-D-Val$_4$-L-Lys$_5$-L-Ile$_6$-D-Phe$_7$-D-Leu$_8$-Ser63
3	SurE-D_1	**D-*allo*-Ile$_1$**-D-Ala$_2$-L-Ile$_3$-D-Val$_4$-L-Lys$_5$-L-Ile$_6$-D-Phe$_7$-D-Leu$_8$-Ser63

The bolded residue is different from that in the SNAC-S model.

As reported by Matsuda et al., SurE could cyclize all AEEA-containing (2-[2-[2-(Fmoc-amino) ethoxy] ethoxy] acetic acid (AEEA)) variants with one to seven AEEAs as long as the *N*- and *C*-termini were L-Ile and D-Leu, respectively. To explore whether short chain length could be cyclized by SurE with computational methods, we constructed the model containing a short substrate (only one AEEA as the internal residue, AEEA1, with L-Ile and D-Leu in the *N*- and *C*-termini, respectively) covalently binding with SurE according to the molecular docking results. Three times 50 ns MD simulations were performed. The results revealed that the proportion of distances d1(C1-N8) and d2(N8-O_{Y154}) both less than 3.5 Å was 53.3% in the SurE-AEEA1 model, indicating the abundance of favorable pre-reaction states, which was critical for the liner substrate to macrocyclization. The precise and stable hydrogen-bonding network also existed to recognize and anchor *N*-terminal residue, which was important for the cyclization of short peptide fragments (Figure S11).

Molecular mechanism of SurE-catalyzed macrocyclization

From the understanding of structural characteristics in the heterochiral coupling of substrate *N*- and *C*-terminal residues, we proposed that (1) H305 abstracted the hydrogen from the hydroxyl group of S63, accompanied by the nucleophilic attack of the hydroxyl oxygen on the carbonyl carbon of the substrate (Figure S3); (2) the conventional oxyanion hole could lock the *C*-terminus of the peptide substrate; (3) R446 could anchor substrate *N*-terminus via hydrogen bonds; (4) a stable hydrogen-bonding network could recognize and anchor L-residue at *N*-terminus of the substrate then further to facilitate a proton-transfer chain from the substrate to K66 via Y154; and (5) K66 was speculated to act as a general base to accept a proton. Therefore, a possible catalytic process of SurE was proposed (Figure 6).

It started from the conducive POSs, where the attacking distance between head to tail (N8$_{Ile1}$-C1$_{Leu8}$) was less than 4.5 Å. Then the N8 of L-Ile$_1$ gradually approached the carbonyl C1 of D-Leu$_8$. At a certain distance, the amino with its lone pair attacked the C1 carbonyl to form the tetrahedral intermediate, accompanied by its deprotonation with the assistance of residue Y154 and K66 via proton-transfer chain. Along with the restoration of S63 acquiring hydrogen from K66, collapse of the tetrahedral intermediate released the cyclic peptide product and the active site of the SurE was reactivated. The catalytic cycle completes. Since the catalytic residue H305 was on the opposite of the catalytic residues (K66-Y154-N156), we anticipated that H305 mainly took part in the loading of the substrates.

To better understand the molecular basis of the proposed catalytic mechanism of SurE, QM/MM calculations were performed using a two-layered ONIOM method, and the key structures including transition states and intermediates were optimized (Figure S12 and Table S4) and the energy profile for SurE-catalyzed macrocyclization was finally obtained. Entirely consistent with our proposal, as the cyclization started, the

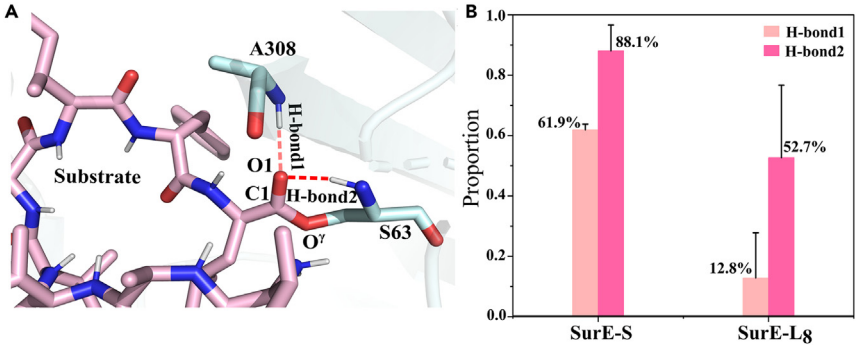


Figure 3. The oxyanion hole in SurE models

(A) The oxyanion hole formed by S63 and A308 in the clustered structure of the SurE-S model. H-bond1 represents the hydrogen bond between O1 and the amide of A308. H-bond2 represents the hydrogen bond between O1 and the amide of S63.

(B) The frequency of hydrogen bonds within the oxyanion hole during molecular dynamics (MD) simulations in the SurE-S and SurE-L_8 systems.

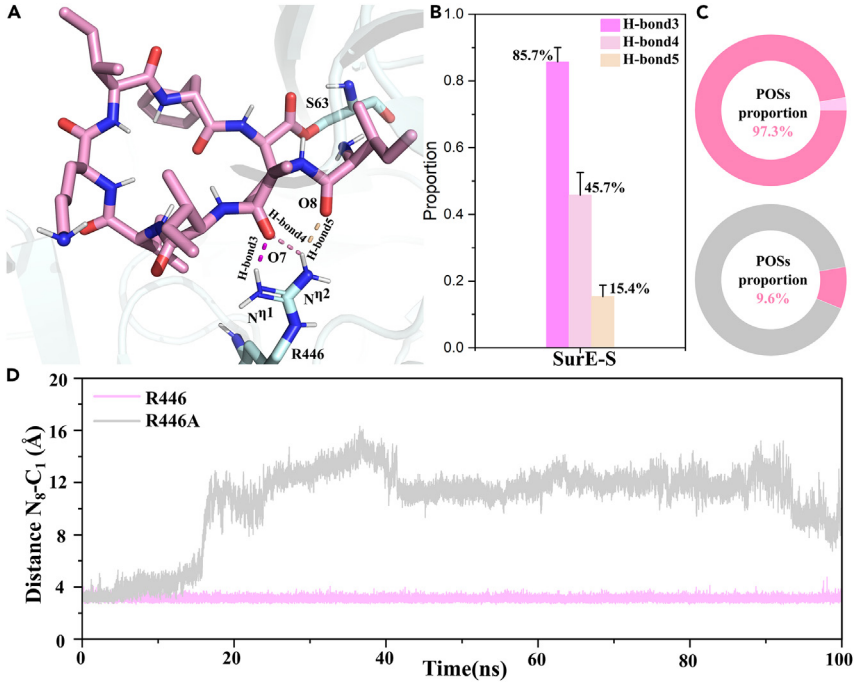


Figure 4. The crucial role of R446

(A) The hydrogen bonds formed by R446 and substrate in the SurE-S.

(B) The proportion of each hydrogen bond between R446 and substrate *N* terminus in MD simulations.

(C) The proportion of POSs (colored by deep pink) in SurE-S (upper) and R446A (lower) models, respectively.

(D) The attacking distance between head to tail ($N8_{Ile1}$-$C1_{Leu8}$) was estimated for the wild type (R446) in pink and mutant (R446A) in gray.

amino group of the substrate *N*-terminus gradually drew near the carbonyl group (C1). The attacking distance of d(N8-C1) decreased from 2.60 Å in the reactant (R) to 1.61 Å in the transition state (TS1), where the deprotonation of the amino group took place through a proton-transfer chain, and the distance of N–H broke at 1.37 Å. Meanwhile, the distance between carbonyl C1 and O^{γ} of S63 increased from 1.35 Å (R) to 1.53 Å (TS1), suggesting the formation of the tetrahedral intermediate (IM1). The energy barrier of the nucleophilic attack in conjunction with the deprotonation process was calculated to be 16.9 kcal/mol. Next, the tetrahedral intermediate collapsed in TS2 with the distance d(C1-O^{γ}) of 1.78 Å, where accompanied by the return of proton on S63 from K66, the cyclized product started to release from SurE catalytic pocket, indicating the occurrence of ester elimination reaction (Figure S12). Likewise, the breakup of tetrahedral intermediate and the restoration of S63 are concerted processes. The energy barrier was calculated to be 19.4 kcal/mol as the rate-limiting step. Furthermore, the cyclic peptide was more stable than the substrate by 2.8 kcal/mol. These results indicated that the macrocyclization catalyzed by SurE was favorable in energy, taking both thermodynamic and kinetic factors into consideration.

Mutation of key residues

In our previous study, the S63/K66/N156 was identified to be crucial for the activity of SurE.[22] Herein, with the guidance of computational analysis, we investigated the function of Y154, N156, A308, and R446 in the catalytic process of SurE. In the end, the site-directed mutations of Y154 to E/F, N156 to D, A308 to G/L/T (A308 located at α-helix, glycine can make α-helix instability, thus destroying the oxyanion hole formed by the amide of A308 and S63), and R446 to A/K were generated and the soluble protein of Y154F, N156D, A308 G/L/T, and R446K was obtained from *E. coli* (Figure S13). Same as the mutation of Y154A in our previous work, the Y154E encountered insoluble protein again, implying the importance of the Y154 for maintaining the active conformation of SurE. HPLC-MS analysis of the enzymatic assays revealed that the Y154F mutation completely abolished the activity of SurE. According to the production levels of cyclopeptide c1, the N156D, A308G, and R446K mutations led to the activity of SurE reduced to around 1%, 10%, and 5%, respectively (Figure S14). Therefore, the site-mutagenetic analysis of SurE supports the computational analysis related to the crucial role of R446 in the formation of POSs. Y154 mediates the proton transfer from the substrate to K66. All mutational data revealed that the key residues involved in the hydrogen-bonding network (Y154-K66-N156-substrate) and oxyanion hole formation (S63 and A308) are indispensable for the activity of SurE.

Design of novel cyclic peptides

Finally, guided by our proposed catalytic mechanism of SurE, five novel cyclic peptides were generated by altering the sequence of peptidyl-SNAC substrate with canonical and non-canonical amino acids (Figure 7A).

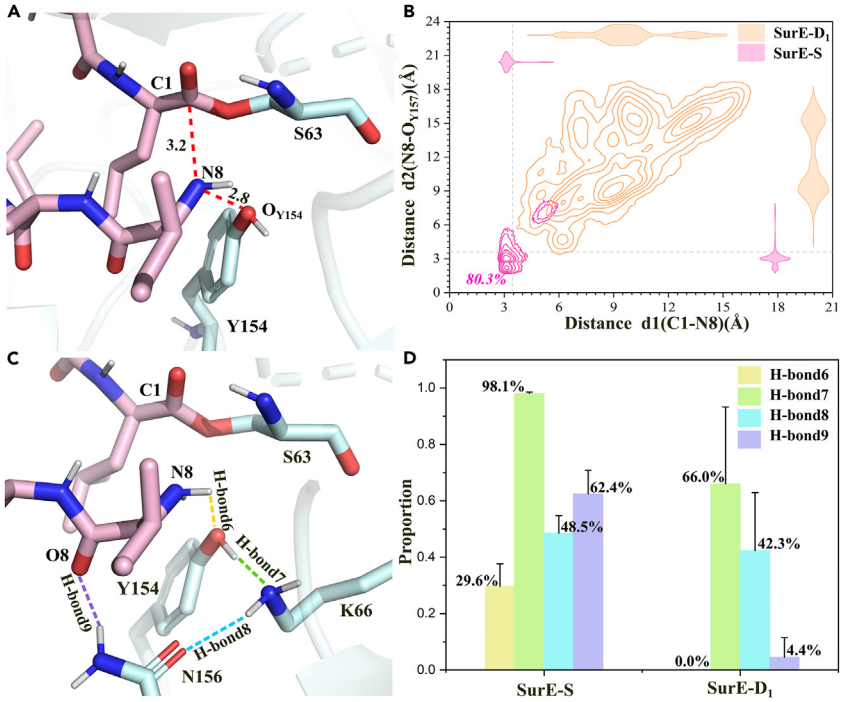


Figure 5. The key distances and important hydrogen bond interactions at the *N* terminus

(A) Two key distances d(C1-N8) and d(N8-O_{Y154}) in representative clustered structure in the SurE-S model.

(B) The distribution of two key distances d(C1-N8) and d(N8-O_{Y154}) in MD simulations.

(C) The precise hydrogen bond network for N-terminal stereoselectivity in the SurE-S model.

(D) The fraction of hydrogen bonds responsible for N-terminal recognition during MD simulations in SurE-S and SurE-D1, respectively.

Considering the essential role of rigorous stereoselectivity of *C*- and *N*-termini in substrate recognition and conformational organization for macrocyclization, we retained their stereoselective configurations and focused on exploring the variation of non-canonical amino acids in *N*-terminus. Besides, three factors were considered, including (1) the high electronegativity of the amide of *N*-terminus could enhance its nucleophilicity to cyclization; (2) the inherent hydrophobicity of *N*-terminal residue could be suitable for the hydrophobic catalytic pocket; and (3) the substituted *N*-terminal residue should be similar to the native one in shape to fit the catalytic pocket well. Therefore, the substrate *N*-terminal residue was screened with different hydrophobic groups as shown in the substrates **2**–**6**, with the native substrate **1** as positive control. Fortunately, all expected cyclic peptides were successfully generated by SurE with a nearly 100% conversion ratio (Figures S15 and S16), indicating the high substrate tolerance of SurE. It should be noted that all substrates **2**–**6** showed similar van der Waals surfaces and electrostatic potentials on the amide group to **1** at M06-2X/def2-TZVP level (Figure S17). Significantly, **4** has stronger electronegativity (−41.2 kcal/mol) than the others (Figure 7B), and it coincidentally presented higher k_{cat}/K_m values, about 2.5-fold, than the native substrate **1** in the SurE enzymatic analysis (Tables 3 and S18). These experimental results were consistent with our proposed substrate stereoselectivity and catalytic mechanism.

DISCUSSION

SurE, as one of the representatives of the PBP-like TE family, can catalyze head-to-tail macrocyclization between the *N*-terminal L-configurated residue and the *C*-terminal D-configurated residue. However, the specific heterochiral preference in SurE-catalyzed macrocyclization remains unclear, which limits a deep comprehension of the catalytic properties of PBP-like TEs. Herein, an approach combining theoretical calculations with mutational experiments was pursued to uncover the structural rationale for the substrate stereoselectivity and catalytic mechanism of SurE.

According to our computations, *C*-terminal D-configurated substrate exhibited better performance at the loading step. Compared with SNAC-L_8 model (L-Leu_8 substrate), conformations in SNAC-S and SNAC-D_1 model with D-Leu_8 were more stable in the catalytic pocket, and easily attainable to approaching the Ser63 of SurE. Also, the carbonyl oxygen of the *C*-terminal D-residue was able to be stabilized by the oxyanion hole. Residue R446, located in the lipocalin-like domain, formed hydrogen bonds with the *N*-terminal O7 and O8 of the substrate. These hydrogen bonds were critical to stabilizing the *N*-terminus in favorable POS conformations for macrocyclization. The proportion of POSs sharply reduced in MD simulations of R446A mutant, verifying the critical role of R446 in locking the *N*-terminus of substrate. In addition, a vital closed hydrogen bonding network was discovered (Substrates-Y154-K66-N156-Substrates). Apart from the catalytic residue Y154 and

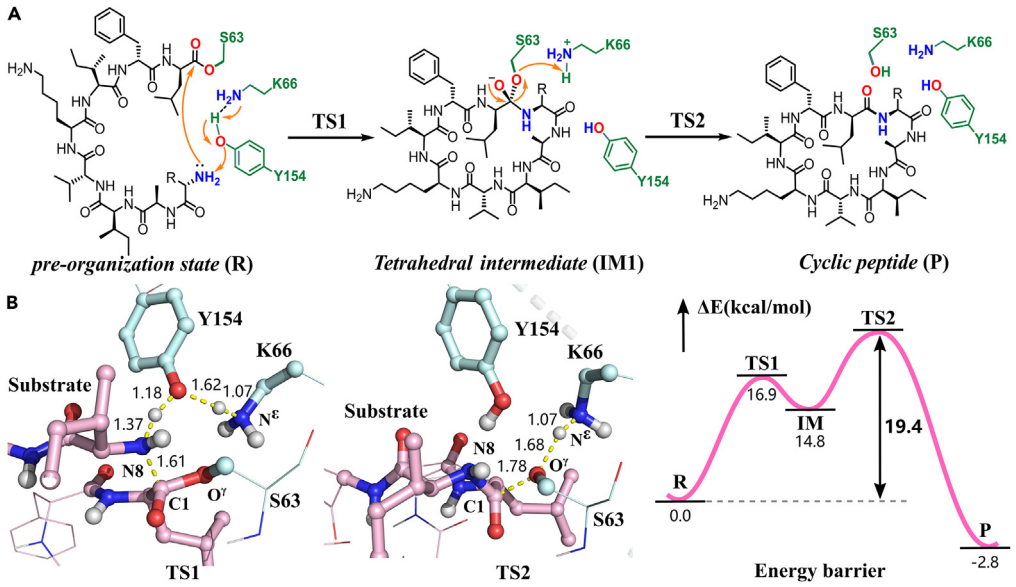


Figure 6. Macrocyclization mechanism catalyzed by SurE

(A) Proposed reaction pathway of SurE-catalyzed macrocyclization.

(B)The optimized transition state structures (including TS1 and TS2) and the energy profile of macrocyclization. The energy barrier for macrocyclization is 19.4 kcal/mol.

K66, the mutation of N156D nearly abolished the activity of SurE, supporting that N156 should involve in anchoring the *N* terminus of the substrate for the formation of pre-reaction states. Although these residues were conserved in most PBP-like TEs, their functions were not interpreted before. Based on our computational results, the stable oxyanion hole, the role of R446 for the formation of POSs, and the hydrogen-bonding network were crucial to the heterochiral coupling catalyzed by SurE. According to the sequence alignment analysis of 18 *trans*-acting PBP-like TEs, important residues involved with the hydrogen-bonding network and catalytic process (S63/K66/Y154/N156/H305/R466) were more conserved than A308.

Besides, a possible catalytic pathway via proton transfer was proposed to complete the head-to-tail macrocyclization during SurE catalysis. The reaction took place like general TEs as a two-step mechanism. The energy barriers for the formation and decomposition of the peptidyl-*O*-SurE tetrahedral intermediate were calculated to be 16.9 kcal/mol and 19.4 kcal/mol, respectively. The former was accompanied by the deprotonation of the substrate with the assistance of residue Y154 and K66 via the proton-transfer chain and the latter was 305 accompanied by the restoration of S63, gaining a hydrogen from K66. The rate-limiting step for the SurE catalytic process was the decomposition of the

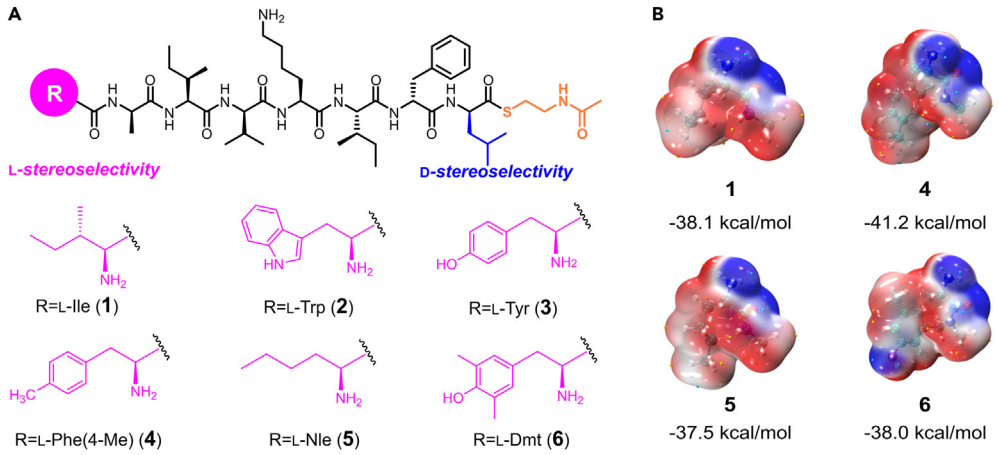


Figure 7. Designed peptides with different N-terminal residues (different ESP)

(A) Designed linear peptide substrates. Here, only the residue at the *N* terminus (colored pink) was replaced by 1–6.

(B) The calculated electrostatic potential.

Table 3. Kinetic analysis of SurE with mimic substrate derivatives

Substrate	The sequence of peptidyl-SNAC	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m($\times 10^{-3}$)
1	L-Ile$_1$-AA$_{2\text{-}7}$-D-Leu$_8$-SNAC	2.30 ± 0.07	301.12 ± 20.17	7.64
2	**L-Trp$_1$**-AA$_{2\text{-}7}$-D-Leu$_8$-SNAC	1.88 ± 0.12	379.20 ± 50.23	4.96
3	**L-Tyr$_1$**-AA$_{2\text{-}7}$-D-Leu$_8$-SNAC	1.55 ± 0.07	252.36 ± 23.00	6.17
4	**L-Phe (4-Me)$_1$**-AA$_{2\text{-}7}$-D-Leu$_8$-SNAC	1.39 ± 0.06	74.01 ± 10.99	18.78
5	**L-Nle$_1$**-AA$_{2\text{-}7}$-D-Leu$_8$-SNAC	2.91 ± 0.30	249.10 ± 62.17	11.68
6	**L-Dmt$_1$**-AA$_{2\text{-}7}$-D-Leu$_8$-SNAC	1.63 ± 0.05	146.38 ± 14.58	11.14

AA$_{2\text{-}7}$ represents polypeptide chain of D-Ala$_2$-L-Ile$_3$-D-Val$_4$-L-Lys$_5$-L-Ile$_6$-D-Phe$_7$. The k_{cat} and K_m values are obtained as means ± S.E. calculated by Prism9 with three replicates. The bolded residue is different from that in the SNAC-S model.

intermediate. Our results indicated that the macrocyclization catalyzed by SurE was advantageous thermodynamically and kinetically, which agreed with the experimental results.

Finally, we created four novel cyclic peptides by considering three factors in the processes of substrate recognition and conformational preparation of SurE-catalyzed macrocyclization.

In summary, a series of computational models were established to uncover the substrate stereoselectivity and catalytic mechanism of SurE-catalyzed macrocyclization. The stable oxyanion hole and a precise hydrogen-bonding network were highlighted, which were critical for substrate recognition. Key residue R446 was important for maintaining POSs. The source of the heterochiral preference of SurE was finally demonstrated from structural analysis. Besides, a catalytic pathway was proposed and proved by our calculations. These results are instructive for understanding the catalytic mechanism of the PBP-like TEs, which are widely spread in the bacterial NRP biosynthetic pathways. In addition, four designed cyclic peptides were efficiently obtained through an enzymatic reaction, guided by the molecular mechanism of SurE on substrate stereoselectivity, which will be beneficial for the development of novel macrocyclic NRPs.

Limitations of the study

A critical loop (residues 211–224), which is missing in the co-crystal structure of SurE, is reconstructed by homology modeling. The initial loop conformation may be not accurate enough. How does the loop interact with linear peptides by induced fit and affect the catalytic efficiency of SurE? This problem remains to be explored.

Although we successfully obtained new cyclic peptides by SurE, the catalytic activity of SurE on these designed linear peptides did not exhibit a significant enhancement. Further investigations should delve into designing linear polypeptide substrates more effectively catalyzed by SurE. Additionally, evaluating the practical impact of the newly designed cyclic peptides in the field of biomedical research is essential.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Computational models
  - Parameters preparation
  - MD simulations
  - Binding free energy calculations
  - QM/MM calculations
  - DNA manipulation and chemicals
  - Protein overexpression and enzyme assays of the wild-type and mutant SurE
  - Kinetic analysis of SurE and its mutants with substrate derivatives
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.108876.

ACKNOWLEDGMENTS

The authors thank the National Natural Science Foundation of China (32270038, 31929001, and 32070041), the National Key R&D Program of China (2019YFA0905400, 2021YFC2100600), and the SJTU-HPC computing facility award for financial support and computational time.

AUTHOR CONTRIBUTIONS

T.S., Y.Z., and X.J. conceived the idea and supervised the whole project. Y.M. and Z.D. performed calculations and analyses. Y.S. carried out the experiments. T.S., Z.D., and Y.M. wrote up the paper. All authors approved the final version of the manuscript for submission.

DECLARATION OF INTERESTS

The authors declare no competing financial interest.

Received: October 2, 2023
Revised: December 12, 2023
Accepted: January 8, 2024
Published: January 11, 2024

REFERENCES

1. Schwarzer, D., Finking, R., and Marahiel, M.A. (2003). Nonribosomal peptides: from genes to products. Nat. Prod. Rep. *20*, 275–287.
2. Kopp, F., and Marahiel, M.A. (2007). Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. Nat. Prod. Rep. *24*, 735–749.
3. Süssmuth, R.D., and Mainz, A. (2017). Nonribosomal Peptide Synthesis-Principles and Prospects. Angew. Chem., Int. Ed. Engl. *56*, 3770–3821.
4. Ojeda, P.G., Cardoso, M.H., and Franco, O.L. (2019). Pharmaceutical applications of cyclotides. Drug Discov. Today *24*, 2152–2161.
5. Zorzi, A., Deyle, K., and Heinis, C. (2017). Cyclic peptide therapeutics: past, present and future. Curr. Opin. Chem. Biol. *38*, 24–29.
6. Muttenthaler, M., King, G.F., Adams, D.J., and Alewood, P.F. (2021). Trends in peptide drug discovery. Nat. Rev. Drug Discov. *20*, 309–325.
7. Sieber, S.A., and Marahiel, M.A. (2005). Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem. Rev. *105*, 715–738.
8. Fischbach, M.A., and Walsh, C.T. (2006). Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. Chem. Rev. *106*, 3468–3496.
9. Hur, G.H., Vickery, C.R., and Burkart, M.D. (2012). Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. Nat. Prod. Rep. *29*, 1074–1098.
10. Du, L., and Lou, L. (2010). PKS and NRPS release mechanisms. Nat. Prod. Rep. *27*, 255–278.
11. Horsman, M.E., Hari, T.P.A., and Boddy, C.N. (2016). Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? Nat. Prod. Rep. *33*, 183–202.
12. Rauwerdink, A., and Kazlauskas, R.J. (2015). How the Same Core Catalytic Machinery Catalyzes 17 Different Reactions: The Serine-Histidine-Aspartate Catalytic Triad of α/β-Hydrolase Fold Enzymes. ACS Catal. *5*, 6153–6176.
13. Kuranaga, T., Matsuda, K., Sano, A., Kobayashi, M., Ninomiya, A., Takada, K., Matsunaga, S., and Wakimoto, T. (2018). Total Synthesis of the Nonribosomal Peptide Surugamide B and Identification of a New Offloading Cyclase Family. Angew. Chem., Int. Ed. Engl. *57*, 9447–9451.
14. Adrover-Castellano, M.L., Schmidt, J.J., and Sherman, D.H. (2021). Biosynthetic Cyclization Catalysts for the Assembly of Peptide and Polyketide Natural Products. ChemCatChem *13*, 2095–2116.
15. Fazal, A., Webb, M.E., and Seipke, R.F. (2020). The Desotamide Family of Antibiotics. Antibiotics *9*, 452.
16. Thankachan, D., Fazal, A., Francis, D., Song, L., Webb, M.E., and Seipke, R.F. (2019). A trans-Acting Cyclase Offloading Strategy for Nonribosomal Peptide Synthetases. ACS Chem. Biol. *14*, 845–849.
17. Ninomiya, A., Katsuyama, Y., Kuranaga, T., Miyazaki, M., Nogi, Y., Okada, S., Wakimoto, T., Ohnishi, Y., Matsunaga, S., and Takada, K. (2016). Biosynthetic Gene Cluster for Surugamide A Encompasses an Unrelated Decapeptide. Chembiochem *17*, 1709–1712.
18. Takada, K., Ninomiya, A., Naruse, M., Sun, Y., Miyazaki, M., Nogi, Y., Okada, S., and Matsunaga, S. (2013). Surugamides A-E, cyclic octapeptides with four D-amino acid residues, from a marine Streptomyces sp.: LC-MS-aided inspection of partial hydrolysates for the distinction of D- and L-amino acid residues in the sequence. J. Org. Chem. *78*, 6746–6750.
19. Matsuda, K., Kobayashi, M., Kuranaga, T., Takada, K., Ikeda, H., Matsunaga, S., and Wakimoto, T. (2019). SurE is a trans-acting thioesterase cyclizing two distinct non-ribosomal peptides. Org. Biomol. Chem. *17*, 1058–1061.
20. Kobayashi, M., Fujita, K., Matsuda, K., and Wakimoto, T. (2023). Streamlined Chemoenzymatic Synthesis of Cyclic Peptides by Non-ribosomal Peptide Cyclases. J. Am. Chem. Soc. *145*, 3270–3275.
21. Matsuda, K., Zhai, R., Mori, T., Kobayashi, M., Sano, A., Abe, I., and Wakimoto, T. (2020). Heterochiral coupling in non-ribosomal peptide macrolactamization. Nat. Catal. *3*, 507–515.
22. Zhou, Y., Lin, X., Xu, C., Shen, Y., Wang, S.P., Liao, H., Li, L., Deng, H., and Lin, H.W. (2019). Investigation of Penicillin Binding Protein (PBP)-like Peptide Cyclase and Hydrolase in Surugamide Non-ribosomal Peptide Biosynthesis. Cell Chem. Biol. *26*, 737–744.e4.
23. Attia, M.A., and Thibodeaux, C.J. (2020). Cycling back to biocatalysis. Nat. Catal. *3*, 476–477.
24. Chen, X.P., Shi, T., Wang, X.L., Wang, J., Chen, Q., Bai, L., and Zhao, Y.L. (2016). Theoretical Studies on the Mechanism of Thioesterase-Catalyzed Macrocyclization in Erythromycin Biosynthesis. ACS Catal. *6*, 4369–4378.
25. White, C.J., and Yudin, A.K. (2011). Contemporary strategies for peptide macrocyclization. Nat. Chem. *3*, 509–524.
26. Blankenstein, J., and Zhu, J. (2005). Conformation-Directed Macrocyclization Reactions. Eur. J. Org Chem. *2005*, 1949–1964.
27. Chow, H.Y., Zhang, Y., Matheson, E., and Li, X. (2019). Ligation Technologies for the Synthesis of Cyclic Peptides. Chem. Rev. *119*, 9971–10001.
28. Liu, L., Tao, W., Bai, L., Kim, E.-S., Zhao, Y.-L., and Shi, T. (2019). Why does tautomycetin thioesterase prefer hydrolysis to macrocyclization? Theoretical study on its catalytic mechanism. Catal. Sci. Technol. *9*, 6391–6403.
29. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Petersson, G.A., Nakatsuji, H., et al. (2016). Gaussian 16 Rev. C.01.
30. Lu, T., and Chen, F. (2012). Multiwfn: A multifunctional wavefunction analyzer. J. Comput. Chem. *33*, 580–592.
31. Case, D., Ben-Shalom, I., Brozell, S.R., Cerutti, D.S., Cheatham, T., Cruzeiro, V.W.D., Darden, T., Duke, R., Ghoreishi, D., Gilson, M., et al. (2018). Amber 2018.
32. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. *46*, W296–W303.
33. Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. *32*, W665–W667.
34. Ravindranath, P.A., Forli, S., Goodsell, D.S., Olson, A.J., and Sanner, M.F. (2015). AutoDockFR: Advances in Protein-Ligand

Docking with Explicitly Specified Binding Site Flexibility. PLoS Comput. Biol. *11*, e1004586.
35. Stephens, P.J., Devlin, F.J., Chabalowski, C.F., and Frisch, M.J. (1994). Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. J. Phys. Chem. *98*, 11623–11627.
36. 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., and Kollman, P.A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. J. Am. Chem. Soc. *117*, 5179–5197.
37. Zhang, J., and Lu, T. (2021). Efficient evaluation of electrostatic potential with computerized optimized code. Phys. Chem. Chem. Phys. *23*, 20323–20328.
38. Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J. Chem. Theor. Comput. *11*, 3696–3713.
39. Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J. Chem. Phys. *98*, 10089–10092.
40. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. *23*, 327–341.
41. Miller, B.R., 3rd, McGee, T.D., Jr., Swails, J.M., Homeyer, N., Gohlke, H., and Roitberg, A.E. (2012). MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. J. Chem. Theor. Comput. *8*, 3314–3321.
42. Vreven, T., Frisch, M.J., Kudin, K.N., Schlegel, H.B., and Morokuma, K. (2006). Geometry optimization with QM/MM methods II: Explicit quadratic coupling. Mol. Phys. *104*, 701–714.
43. Vreven, T., Byun, K.S., Komáromi, I., Dapprich, S., Montgomery, J.A., Morokuma, K., and Frisch, M.J. (2006). Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM. J. Chem. Theor. Comput. *2*, 815–826.
44. Grimme, S., Antony, J., Ehrlich, S., and Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys. *132*, 154104.
45. Zhao, Y., and Truhlar, D.G. (2008). 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 functionals and 12 other functionals. Theor. Chem. Acc. *119*, 525.
46. Marenich, A.V., Cramer, C.J., and Truhlar, D.G. (2009). Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. J. Phys. Chem. B *113*, 6378–6396.
47. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Gaussian 16C01	Frisch et al.[29]	https://gaussian.com/gaussian16/
Multiwfn 3.8	Tian et al.[30]	http://sobereva.com/multiwfn/
AMBER18	Case et al.[31]	https://ambermd.org/
Other		
Protein structure	Matsuda et al.[20]	6KSU;6KSV
SWISS-MODEL	Waterhouse et al.[32]	https://swissmodel.expasy.org/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the Lead Contact, prof. T. S. (tshi@sjtu.edu.cn).

Materials availability

This study does not generate any novel reagents and all materials used in this study are reported either in the main text or in the supplemental information.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report the original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

METHOD DETAILS

Computational models

Using the reported crystal structure of SurE (PDB: 6KSU, 6KSV) as template,[21] we constructed the structures of open SNAC-model and closed SurE-model by homology modelling[32] (Figure S19). The protonated state (pH 8.0) for each residue of SurE was determined via pK_a calculations on PDB2PQR web serve.[33] Open SNAC-model included peptidyl-*S*-SNAC substrate with SurE at open form, named SNAC-S, SNAC-D_1, and SNAC-L_8, which were utilized to investigate the substrate stereoselectivity in the loading step. Closed SurE-model possessed peptidyl-*O*-SurE adduct with SurE at closed form, named SurE-S, SurE-D_1, and SurE-L_8, was established to uncover the substrate stereoselectivity. Closed SurE-model possessed AEEA1-O-SurE adduct with SurE at closed form, named SurE-AEEA1, was established to validate whether SurE can cyclize substrate with different chain lengths. The complexes of SNAC-S, SurE-AEEA1 and the SurE-S models were acquired by AutoDockFR[34] through non-convalent/convalent docking and then constrained molecular dynamics (Figures S20 and S21). SNAC-L_8, SNAC-D_1, SurE-D_1, and SurE-L_8 models were produced by fragment replacement based on SNAC-S and SurE-S models (Figure S22).

Parameters preparation

To obtain the parameters for the complex model, optimizations were performed at the level of B3LYP[35]/6-31G* using Gaussian 16.[29] A two-step restrained electrostatic potential (RESP)[36] charge was generated by Multiwfn and parameters of the bonds, angles, dihedral angles, and van der Waals radii for the substrate were obtained by using the Antechamber package.[30,37] After tleap process, these models were set up within an octahedral box of TIP3P water molecules and 22 sodium ions were added to maintain the electroneutrality in each model.

MD simulations

Amber 18 program packages[31] were utilized to perform molecular dynamics in each model using the ff14SB[38] force field. In the MD simulations, we considered long-range electrostatic interactions via the Particle Mesh Ewald (PME) method.[39] To fix the bonds and angles concerning the hydrogen atoms, we introduced the SHAKE algorithm.[40] We set the non-bonding cutoff value to 10.0 Å. To avoid inappropriate atomic collisions, we performed a two-step minimization process for the system, with the first step directed to all water molecules and the second

step to prevent inappropriate collisions in the rest. Then we gradually heated the system from 0 K to 300 K within 100 ps. After heating the system, we switched it to constant pressure and temperature mode (NPT). In each mode, we performed a 200 ps balance process to ensure that the system was at the correct density. Finally, using the last frame of the NPT processed as the initial structure to perform four times independent MD simulations (100 ns) for each model.

For peptidyl-SNAC models, the loading distance between substrate (C1) and S63 (O^{γ}) was constrained at 3.4 Å using a harmonic vibrational potential with a force constant of 20 kcal $mol^{-1}Å^{-2}$ in the equilibration step for catching the favorable loading conformation. Similarly, the attacking distance between the head (N8) and the tail (C1) was also constrained at 3.5 Å for favorable pre-organized states in peptidyl-SurE models. Then the system was re-equilibrated without constraints and four replicas of 100 ns MD simulation trajectories were obtained.

Binding free energy calculations

The binding free energy between the peptide substrates and SurE was calculated using the molecular dynamics generalized Born surface area (MM-GBSA) method with the python program MMGBSA.py.[41] The last 10 ns of the trajectories were extracted to calculate the binding free energy and the decompositions of the energy were assigned to specific residues whose contributions to substrate-enzyme binding interaction was greater than 1.0 kcal/mol in each model.

QM/MM calculations

In this study, the QM/MM calculations were performed by employing a two-layered ONIOM method[42,43] embedded in the Gaussian16 program. The representative structure from the dominant clusters obtained from simulations was extracted as the original structure for QM/MM calculations. All residues around the substrate within 8 Å were included in MM region. The QM region consisted of the side chain of Y154, K66, and N156, which were essential in catalytic process as we mentioned. Besides, the N- terminus and the C-terminus of the substrate were included in QM region (Figure S23). The QM region contains 81 atoms and the whole model contains 467 atoms in total.

The local minima and transition states (TSs) were optimized at ONIOM (B3LYP-D3[44]/6-31G*:Amber) level, and all TSs were confirmed by vibrational frequency analysis and the intrinsic reaction coordinate (IRC) calculations. Finally, for refining the energy profile, single-point energy calculations based on the optimized structure were completed at the level of ONIOM (M06-2X[45]/6-311+G(d,p): Amber) with the SMD solvation correction.[46]

DNA manipulation and chemicals

Oligonucleotides (Table S5) and gene were synthesized by Shanghai Sangon Biotech (China). Polymerase Phusion High-Fidelity PCR Master Mix (New England Biolabs) was used for DNA cloning, 2 × Flash PCR MasterMix (Dye) (CWBIO, Beijing, China) was used for screening. Restriction endonucleases were purchased from New England Biolabs. Gibson assembly solution is 2 × Ezmax-Muli CloneMix Plus purchased from Shanghai Tolo Biotech (China). Plasmid DNA extraction Kit was obtained from Shanghai Generay Biotech (China). Mimic peptidyl-SNAC thioester precursor was synthesized by our previous work.[22] Briefly, using a Symphony® peptide synthesizer (0.2 mmol scale), the protected linear peptides were synthesized and then eluted from the 2-Cl Trt resin of Symphony® by mixed solvents of trifluoroethano: CH_3COOH: CH_2Cl_2 (1:1:3). The resulting peptides were coupled to N-acetylcysteamine (2.5 equiv) in 1 mL THF containing dicyclohexylcarbodiimide (DCC) (1.2 equiv) and 1-hydroxybenzotriazole hydrate (HOBt) (1.2 equiv). Finally, the peptidyl-SNAC thioesters were deprotected with 16:3:1 trifluoroacetic acid (TFA)/CH_2Cl_2/N-acetylcysteamine, and purified as white solid by reverse-phase (C18) HPLC with the elution condition of 20-100% MeCN/H_2O (0.1% formic acid, v/v) over 50 min.

Protein overexpression and enzyme assays of the wild-type and mutant SurE

For protein overexpression of SurE (PDB: 6KSU), the gene *surE* was synthesized according to the codon from *E. coli* and introduced into pET28a (NdeI, HindIII) by Gibson assembly[47] to obtain the plasmid pRJ244. All the mutational genes of *sur*E were generated based on the PCR template pRJ244 with the primers containing the target mutations. Then, the mutational genes were introduced into pET28a (NdeI, HindIII) respectively to generate pRJ245 to 249 and pRJ429, or into pRJ244 (MscI, HindIII) to generate pRJ561 to 563 (Table S6).

The recombinant SurE proteins were prepared by following the reported procedure. All enzyme assays were performed in a 20 μL reaction system consisting of 50 mM Tris-HCl (pH 8.0), 5% dimethyl sulfoxide (DMSO, v/v), 10 μM enzyme, and 0.1 mM substrate incubated at 30 °C for 2 hours. The reaction was quenched by addition of 500 μL acetonitrile followed by centrifugation (14,000 xg) for 5 min. The 20 μL supernatant was then subjected to HPLC-MS equipped with a Waters Xbridge C18 column (250 mm × 4.6 mm, 5 μm) and eluted with a linear gradient of 30 to 62 % of MeCN-H_2O (0.1% formic acid, v/v), 0.8 mL min-1 flow rate over 15 min.

Kinetic analysis of SurE and its mutants with substrate derivatives

The kinetic analysis of SurE and its mutants was performed using compounds **1**–**6** as substrates using 5, 5′-dithio-2-nitrobenzoic acid (DTNB) reduction method. Each assay was performed at 50 μL system containing 50 mM Tris-HCl buffer (pH 8.0), 0.2 mM DTNB, 2 μM protein (4 μM protein was used for A308T), 10% DMSO (v/v), and stepwise gradient substrate (from 0.05 to 1.5 mM). The reactions were initiated by adding

substrates to the reaction system and monitored at 412 nm by a SpectraMax 190 plate reader. Michaelis-Menten kinetics were obtained employing the value ε = 13,600 $M^{-1}cm^{-1}$ by GraphPad Prime 7. Each reaction was performed in triplicate.

QUANTIFICATION AND STATISTICAL ANALYSIS

We performed three paralleled MD trajectories for each system. All Hydrogen bond occupancy was represented by its average and standard derivation. What's more, we perform a t-test to reveal the significant differences between different systems with 95% confidence intervals.