Title	Heterochiral coupling in non-ribosomal peptide macrolactamization
Author(s)	Matsuda, Kenichi; Zhai, Rui; Mori, Takahiro; Kobayashi, Masakazu; Sano, Ayae; Abe, Ikuro; Wakimoto, Toshiyuki
Citation	Nature Catalysis, 3, 507-515 https://doi.org/10.1038/s41929-020-0456-7
Issue Date	2020-05-04
Doc URL	http://hdl.handle.net/2115/79676
Type	article (author version)
File Information	WoS_94265_Wakimoto.pdf

Heterochiral coupling in non-ribosomal peptide macrolactamization

Kenichi Matsuda,[1] Rui Zhai,[2] Takahiro Mori,[2] Masakazu Kobayashi,[1] Ayae Sano,[1] Ikuro Abe,[2,3]* Toshiyuki Wakimoto[1]*

[1]Faculty of Pharmaceutical Sciences, Hokkaido University, Kita 12, Nishi 6, Kita-ku, Sapporo 060-0812, Japan

[2]Graduate School of Pharmaceutical Sciences, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

[3]Collaborative Research Institute for Innovative Microbiology, The University of Tokyo, Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657, Japan

Abstract

Heterochiral coupling is favored in abiotic peptide bond formation, whereas biotic peptide bond formation is dominated by homochiral coupling. Herein, we report that heterochiral coupling is a rather general paradigm in the head-to-tail macrolactamization of non-ribosomal peptide biosynthesis. The canonical *cis*-acting offloading cyclases, such as the type I thioesterase (TE) and terminal condensation-like (C_T) domains, catalyze the head-to-tail macrolactamization between *N*- and *C*-terminal residues with the D- and L-configurations, respectively. In contrast, the penicillin-binding protein (PBP)-type TEs, a recently identified family of *trans*-acting cyclases, couple heterochiral residues with complementary stereoselectivity to the canonical one. Thus, a suite of *cis*- and *trans*-TE non-ribosomal peptide synthetases (NRPSs) could overcome the stereochemical constraints present in the heterochiral head-to-tail macrolactam formation in bacterial NRPS biosynthesis. Furthermore, we provide the structural rationale for the *C*-terminal stereoselectivity of the non-canonical offloading cyclases. The PBP-type TEs with broad substrate specificity are potentially applicable as biocatalysts and genetic tools for synthetic biology.

Introduction

Homochiral coupling in biotic peptide bond formation performed by ribosomes is a fundamental principle of life. In contrast, abiotic peptide bond formation exhibits a significant preference for heterochiral coupling. For example, competitive coupling reactions between the *C*-terminus of the *N*-acyl group and the *N*-terminus of amido derivatives of amino acids resulted in the selective accumulation of the heterochiral product (mostly >80%).[1] The preference for heterochiral coupling is also seen in chemical macrolactamization; therefore, the proper pair of *N*- and *C*-terminal residues should be taken into account when considering the cyclization position for efficient macrolactamization.[2,3]

Many naturally occurring macrolactams are biosynthesized by non-ribosomal peptide synthetases (NRPSs).[4] In macrolactam biosynthesis, an offloading cyclase, such as a type I thioesterase (TE) or terminal condensation-like (C_T) domain that is typically fused at the *C*-terminus of the synthetase, catalyzes the intramolecular cyclization of the grown linear peptide.[5,6] These *cis*-acting elements are crucial, as they confer the structural rigidity to the peptide that is important for the exertion of its specific biological activity.[7]

We previously reported a new family of offloading cyclases, penicillin-binding protein (PBP) type TEs, that lack sequence homology to the canonical TEs, but are rather similar to the PBPs, a group of enzymes responsible for transpeptidation in peptide glycan biosynthesis.[8] A striking feature of the PBP-type TEs is that they are physically discrete from the NRPSs, and thus act *in trans*. This new family of TEs was originally found in the biosynthetic gene cluster of surugamides, which are two groups of simultaneously synthesized but structurally unrelated non-ribosomal peptides (NRPs), isolated from *Streptomyces* sp. JAMM992 and several other strains.[8,9,10,11] Surugamide A, with cathepsin B inhibitory activity, and its derivatives surugamides B-E are cyclic-octapeptides, while cyclosurugamide F and its linear derivative are decapeptides with unknown biological functions.[12,13]

The NRPS genes for surugamide biosynthesis are clustered in a single genomic locus, and none of them possess canonical offloading domains at their *C*-termini.[12] Uniquely, a single PBP-type TE, named SurE, acts *in trans* on two distinct NRPS assembly lines to catalyze the macrolactamization of two unrelated peptides, highlighting its wide substrate tolerance.[13] The genes encoding homologs of SurE are also found in the biosynthetic gene clusters of antimicrobial cyclic peptides (Figure 1a and b), as well as in various NRPS gene clusters with biosynthetic products that are yet to be identified.[14,15,16] After our first description of a PBP-type TE in 2018, two other groups published reports on the same enzyme,[17,18] indicating the growing attention to this family of cyclases. Hereafter, we refer to the NRPS system that lacks the canonical *cis*-acting offloading domain but contains the PBP-type TE as the *trans*-TE NRPS (Figure 1c).

Considering the fact that the chemo- and regio-specific cyclizations of peptides remain challenging tasks,[19,20] the use of a PBP-type TE family as the biocatalyst for macrolactamization is quite interesting. However, its substrate scope is largely unexplored. In this study, we examined the substrate scope of SurE using synthetic substrate derivatives, revealing its strict requirement for the heterochirality between the *N*- and *C*-terminal residues with L- and D-amino acids, respectively. This in turn evoked our interest toward the enantiomeric pair of the terminal residues, in the case of the head-to-tail macrolactamization catalyzed by other typical offloading cyclases. We found that the macrolactamization mediated by the canonical *cis*-acting offloading domains (type I TE and C_T) is dominated by heterochiral coupling with the opposite configuration (D-amino acid at *N*-terminus and L-amino acid at *C*-terminus). This observation provides a perspective for the use of the PBP-type TEs, as a complementary strategy to the alternative heterochirality in the biosynthesis of non-ribosomal head-to-tail cyclopeptides. Our structural analysis, together with the model of the acyl-*O*-SurE complex, provide the structural basis for SurE's stereoselective recognition of the substrate *C*-terminal residue. Moreover, genetic manipulation of the PBP-type TE containing gene cluster yielded non-natural macrolactams, highlighting the remarkable plasticity of this family of stand-alone macrocyclases.
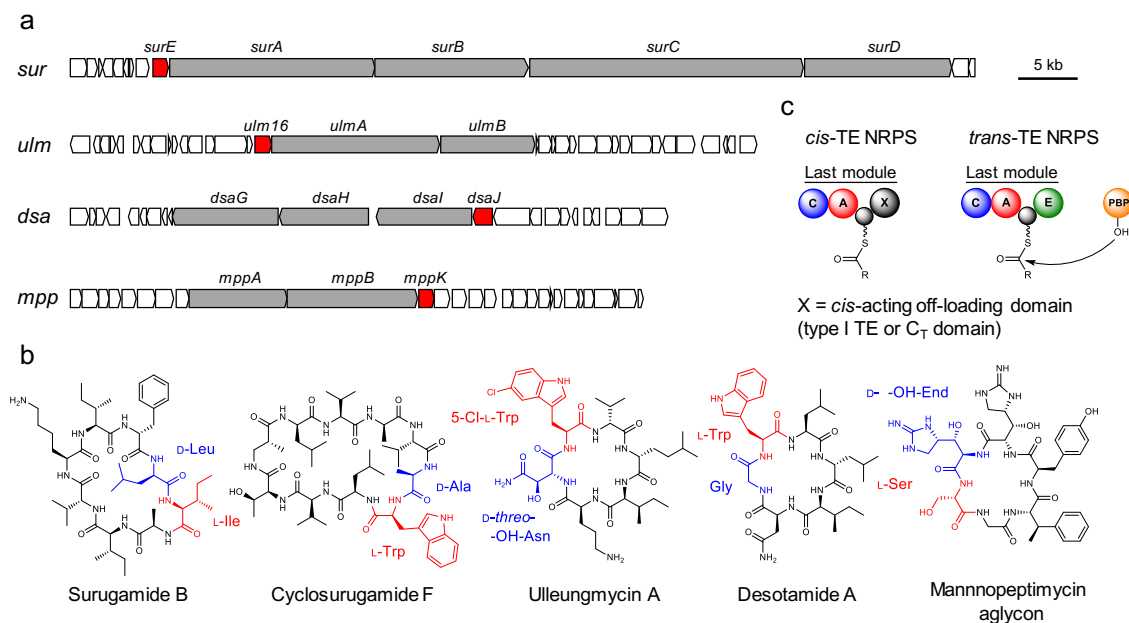
4

Figure 1. PBP-type TEs for non-ribosomal macrolactam biosynthesis. a) The *trans*-TE NRPS gene clusters. The genes encoding PBP-type TEs and those encoding cognate NRPSs are colored red and gray, respectively. In surugamide biosynthesis, the octapeptides (i.e.; surugamide A and its derivatives) are synthesized by SurA/SurD, whereas the decapeptides (i.e.; cyclosurugamide F and its linear derivative) are synthesized by SurB/SurC. A single *trans*-acting PBP-type TE, SurE, offloads the products from both assembly lines. The cognate NRPSs lack the canonical *C*-terminal offloading domain. b) Selected macrolactams biosynthesized by PBP-type TEs. The residues at the *N*- and *C*-termini of the linear precursors are colored red and blue, respectively. End stands for enduracididine. c) Schematic comparison of the offloading steps in *cis*-TE NRPSs (left) and *trans*-TE NRPSs (right). In the *trans*-TE NRPSs, the epimerization (E) domain is located at the *C*-terminus of the NRPS, whereas the E domain is rarely found in the last modules of the *cis*-TE NRPSs. The colored balls represent functional domains: blue, condensation (C) domain; red, adenylation (A) domain; gray, peptidyl-carrier protein (PCP); green, E domain; orange, *trans*-acting PBP-type TE.

Results

5

Substrate scope of SurE.

To evaluate the substrate scope of SurE, we utilized an *N*-acetyl cysteamine (NAC) derivative (peptide-SNAC) to mimic the PCP-tethered precursor. The kinetic parameters for each substrate variant are summarized in Table 1. SurE is selective to a peptidyl substrate with D-amino acid at its *C*-terminus,[8] consistent with its sequence homology to PBP whose substrates are D-Ala-D-Ala dipeptide motif in the cell wall. On the other hand, SurE's selectivity against the *N*-terminal configuration remained unknown. When the mutated substrate with the substitution of L-Ile1 to D-*allo*-Ile was incubated with SurE, it was not cyclized but hydrolyzed to give a linear octapeptide, showing that SurE specifically catalyzes the heterochiral coupling between the L-amino acid at the *N*-terminus and the D-amino acid at the *C*-terminus (Supplementary Figure 1). Meanwhile, the derivative with L-Ile1 substituted to L-Trp was cyclized more efficiently than the native sequence. The relaxed recognition of SurE for the side-chain of the *N*-terminus is in stark contrast to that of TycC-TE, a well characterized macrolactam forming-type I TE in tyrocidine biosynthesis, which strictly recognizes both the stereochemistry and side-chain of the *N*-terminal residue (D-Phe1).[5] Slow hydrolysis was observed with a substrate with L-Pro at the *N*-terminus. Its over three-fold improved K_M value, as compared to the native sequence, suggests that a substrate with a hydrophobic *N*-terminus is more favored for the formation of the peptide-*O*-SurE intermediate.

Next, SurE's specificity for the side-chain in the middle of the peptide chain was evaluated, using alanine-substituted derivatives. The configuration of each residue was not altered in these derivatives. These substitutions caused up to five-fold increases in K_M and three-fold decreases in k_{cat}, but all were converted to cyclic peptides quantitatively without detectable formation of the hydrolytic linear peptides, showing that these substitutions had relatively little influence on the cyclizing activity. In the previous report, the substitution of L-Lys5 to L-Ala resulted in a severe reduction of the cyclization rate, probably due to poor

6

water solubility.[8] Therefore, to evaluate the importance of L-Lys5, it was substituted to L-α-amino adipate (LαAA) possessing the opposite charge. Although this mutation negatively affected both the K_m and k_{cat} values, this mutated substrate was cyclized without detectable hydrolysis, indicating that the L-Lys5 residue is dispensable for cyclization by SurE. The side-chain and configuration of the *C*-terminal residue exhibited the largest influence on SurE-mediated cyclization. The substitution of D-Leu8 to D-Ala or achiral Gly gave trace amounts of cyclic products, while the substitution to D-Phe conferred higher affinity than the native substrate (two-fold reduction in K_M).
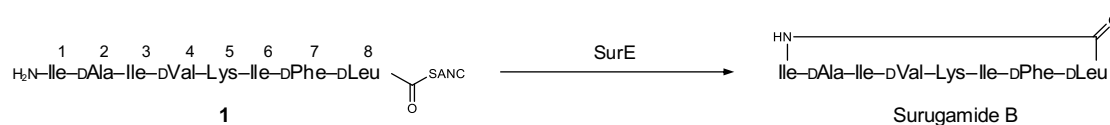
Besides the SNAC-activated substrates, SurE showed cyclization activity against the methylester substrate, which is much easier to synthesize (Supplementary Figure 3). Although its efficiency remains low, this trait makes SurE a fascinating starting point for engineering to improve catalytic efficiency.

The specificity of the heterochiral coupling in the SurE-mediated macrolactamization prompted us to examine the enantioselectivities of the known *cis*-acting offloading cyclases. From over 300 NRP gene clusters in the MiBiG database[21] supplemented with our personal updates, we sought *cis*-acting offloading cyclases whose biosynthetic products are head-to-tail macrolactams cyclized between two α-amino acid residues. During this search, cyclase ligating residues other than α-amino acids were excluded. In total, thirteen offloading domains, including eight type I TEs and five C_T domains, were identified as cyclases that catalyze macrolactamization in a head-to-tail manner (Supplementary Table 1, Supplementary Figure 6). Then, the structures of the pre-cyclized intermediates were predicted, based on the domain architecture of each assembly line. The results revealed that the head-to-tail macrolactam-forming type I TEs and C_T domains exclusively couple a D-amino acid at the *N*-terminus and an L-amino acid at the *C*-terminus (in the biosynthesis of ferrichrome-type cyclic peptides, the *N*-terminal residue is achiral Gly). Rare examples against this trend are the type I TEs in the biosyntheses of cyclic heptapeptides, cyclomarines and ilamycins, which ligate L-amino acids at both termini of their acyclic substrates.[22, 23] In contrast to these canonical offloading

7

domains, the PBP-type TE catalyzes heterochiral coupling between an L-amino acid at the *N*-terminus and a D-amino acid at the *C*-terminus. (in the biosynthesis of desotamides, the *C*-terminal residue is achiral Gly). This contrast in the heterochiral pair provided a new view of the *trans*-TE NRPSs, as a strategy to complement the alternative heterochirality and to diversify the biosynthetic routes of NRP macrolactams.

Table 1. Kinetic parameters of SurE against SNAC substrates. Substituted residues are highlighted in gray.

s.e.m.: standard error of the mean.

$$\text{H}_2\text{N–Ile–DAla–Ile–DVal–Lys–Ile–DPhe–DLeu–SANC} \quad \xrightarrow{\text{SurE}} \quad \text{Ile–DAla–Ile–DVal–Lys–Ile–DPhe–DLeu (Surugamide B)}$$

substrate	Sequence								K_m (mM) (± s.e.m.)	k_{cat} (min^{-1}) (± s.e.m.)
	1	2	3	4	5	6	7	8		
SB-SNAC (**1**)	Ile	DAla	Ile	DVal	Lys	Ile	DPhe	DLeu	152.4±4.5	15.7±0.6
I1DalloI (**2**)	**DalloI**	DAla	Ile	DVal	Lys	Ile	DPhe	DLeu	55.5±10.0[a]	1.5±0.1[a]
I1W (**3**)	**Trp**	DAla	Ile	DVal	Lys	Ile	DPhe	DLeu	58.8±1.2	17.0±0.3
I1P (**4**)	**Pro**	DAla	Ile	DVal	Lys	Ile	DPhe	DLeu	41.1±1.2[a]	1.6±0.1[a]
I3A (**5**)	Ile	DAla	**Ala**	DVal	Lys	Ile	DPhe	DLeu	157.2±7.0	7.7±0.5
V4A (**6**)	Ile	DAla	Ile	**DAla**	Lys	Ile	DPhe	DLeu	159.4±11.0	9.3±0.9
K5aAA (**7**)	Ile	DAla	Ile	DVal	**aAA**	Ile	DPhe	DLeu	794.0±51.4	6.4±0.5
I6A (**8**)	Ile	DAla	Ile	DVal	Lys	**Ala**	DPhe	DLeu	393.1±35.3	13.7±2.4
F7A (**9**)	Ile	DAla	Ile	DVal	Lys	Ile	**DAla**	DLeu	714.3±8.9	8.9±1.3
L8A (**10**)	Ile	DAla	Ile	DVal	Lys	Ile	DPhe	**DAla**	n.d.[b]	n.d.[b]
L8F (**11**)	Ile	DAla	Ile	DVal	Lys	Ile	DPhe	**DPhe**	78.0±1.4	18.7±0.5
L8G (**12**)	Ile	DAla	Ile	DVal	Lys	Ile	DPhe	**Gly**	n.d.[c]	n.d.[c]

[a] : kinetic parameters for hydrolysis are shown.

[b] : low V_{max} precluded accurate estimation.

[c] : non-regioselective cyclization was observed under the reaction conditions.

9

Crystal structure of SurE.

To obtain mechanistic insights into SurE catalysis, we solved the crystal structure of SurE at 2.2 Å resolution (Figure 2a). The asymmetric unit contained two SurE monomers. The final crystal data and intensity statistics are summarized in Supplementary Table 2. The monomers are nearly identical to each other, with a root-mean-square deviation (RMSD) of 0.2 Å. The overall structure of SurE consists of two subdomains: a *N*-terminal domain (1-350) with a typical alpha-beta hydrolase fold and a *C*-terminal lipocalin-like domain (368-451) containing eight antiparallel beta strands. These two subdomains are linked with a long, unfolded linker region (351-367). A DALI server[24] search revealed that SurE exhibits structure similarity with the AmpH-type class C PBPs, such as AmpC β-lactamase[25] (PDB: 5F1F, RMSD for Cα-atoms: 2.8 Å, AAid: 23%) from *Klebsiella aerogenes*, the Class C β-lactamase[26] (PDB: 4GZB, RMSD for Cα-atoms: 2.8 Å, AAid: 23%) from *Pseudomonas aeruginosa* PAO1, and the alkaline D-peptidase ADP[27] (PDB: 4Y7P, RMSD for Cα-atoms: 3.0 Å, AAid: 27%) from *Bacillus cereus*. The residues and positions of the catalytic tetrad, Ser63, Lys66, Asn156, and His305, were well conserved in the putative active site of the PBP domain, as compared to other PBPs (Figure 2b). The oxyanion hole, formed by the backbone amides of Ser63 and Ala308, was occupied by a water molecule. The significant differences between the PBP domain of SurE and ADP are four loop regions (121-133, 187-228, 295-301, and 331-333 of SurE) that are predicted to play an important role in the conformational change upon substrate binding in ADP[27] (Figure 2c and 2d). The different lengths and conformations of the loops significantly altered the size of the active site cleft. SurE possesses a relatively wide and shallow cleft at the active site, as compared to that of ADP.

Additionally, the *C*-terminal lipocalin domain covers the PBP domain to form a large cleft between the PBP and lipocalin domains (Supplementary Figure 7a). Pab87 (PDB: 2QMI, RMSD 5.5 Å for Cα-atoms, AAid: 27%), a D-stereospecific self-compartmentalizing peptidase from archaea, also possesses domains homologous to PBP at the *N*-terminus and a lipocalin domain at the *C*-terminus, similar to SurE[28]. In Pab87,

however, these two domains are connected with an alpha-helical linker, while in SurE they are connected with a flexible loop region (Supplementary Figure 7b and c). Although Pab87-lipocalin plays a role in oligomerization to form a cube-shaped octamer, no obvious interactions are observed between the individual SurE-lipocalins or lipocalin and the PBP domains. Indeed, a gel-filtration analysis showed that recombinant SurE exists as monomer in solution (Supplementary Figure 8). Taken together, these observations suggest that the lipocalin domain of SurE participates in substrate binding, rather than oligomerization.

To gain structural insight into the stereospecificity of SurE against the *C*-terminus of the substrate, we constructed a model of the peptide-*O*-SurE complex based on the *apo*-SurE structure, and calculated the conformation of the peptidyl moiety using the residues located within 8 Å from the active site serine (Ser63). During the calculation, *N*-formyl D-Leu, mimicking the *C*-terminus of a linear substrate, was used to reduce the computational cost. As a result, side chain of the tethered D-Leu was accommodated in the neighboring hydrophobic pocket, consisting of Leu231, Ala234, Gly235, and Val309 (Figure 2e). This pocket is large enough to accommodate bulky side chains, such as that of Phe. The formyl group was recognized by the N^{ω} atom of Arg446, which is conserved among the PBP-type TE lipocalin domains. On the other hand, the calculation using an antipodal ligand, *N*-formyl L-Leu, showed that the L-Leu side chain was excluded from the pocket. The potential energy of the resulting structure with *N*-formyl L-Leu is 30 kJ/mol higher than that of *N*-formyl D-Leu. This suggests that in SurE, the spatial arrangement of the oxyanion hole, the side chain-recognizing-hydrophobic pocket, and the carbonyl recognizing-Arg446 confer the D-stereoselectivity toward the *C*-terminus of the substrate.

To understand further intimate structural details of the substrate recognition, the crystal of *apo*-SurE was soaked in buffer containing I1P (**4**), a substrate derivative that cannot be cyclized by SurE, to obtain the crystal of the peptide-*O*-SurE complex. As a result, an electron density of D-Leu-SNAC that is a *C*-terminal portion of soaked **4** was observed in the active site of one of the monomers, and in the other monomer, a large electron

11

density connected to the side chain of Ser63 was observed.   Considering the enzyme reaction, the *C*-terminal D-Leu seemed to be tethered to the Ser63 residue (Figure 2f). Although the other residues of the peptide substrate were not clearly observed, weak extra densities also remained close to the amino group of the D-Leu residue (Supplementary Figure 9a). At the same time, the loop region (211-224) was highly disordered in the complex structure (Supplementary Figure 9b). Notably, the His225 and Met226 residues on the loop protruded toward the active site, and occupied the hydrophobic pocket that was predicted to accommodate the side chain of D-Leu tethered on Ser63 (Figure 2f and Supplementary Figure 9c), thus the tethered D-Leu residue in the complex structure was pointing in the opposite direction from the previous simulation result. This suggested that the conformational rearrangement of the flexible loop is induced by the formation of the peptidyl-*O*-SurE complex. Our computational model of SurE tethering entire **4** suggested that the loop and lipocalin domain interact with tethered peptide (Supplementary Figure 10). To investigate the role of this loop, His225 was substituted to Ala. This mutant retained the cyclization activity; however, the kinetic analysis showed that the k_{cat} was decreased by three-fold as compared to the wild type (5.5 ± 0.4 min^{-1}), although it exhibited a similar K_m value (214.1 ± 10.5 μM). This suggests that the loop is involved in cyclization, rather than the formation of the peptidyl-*O*-SurE complex.

Figure 2. Crystal structure of SurE. (a) The overall structure of SurE. The PBP domain, linker loop, and lipocalin domain are colored blue, pink, and green, respectively. (b) The active site of SurE. Catalytic tetrad
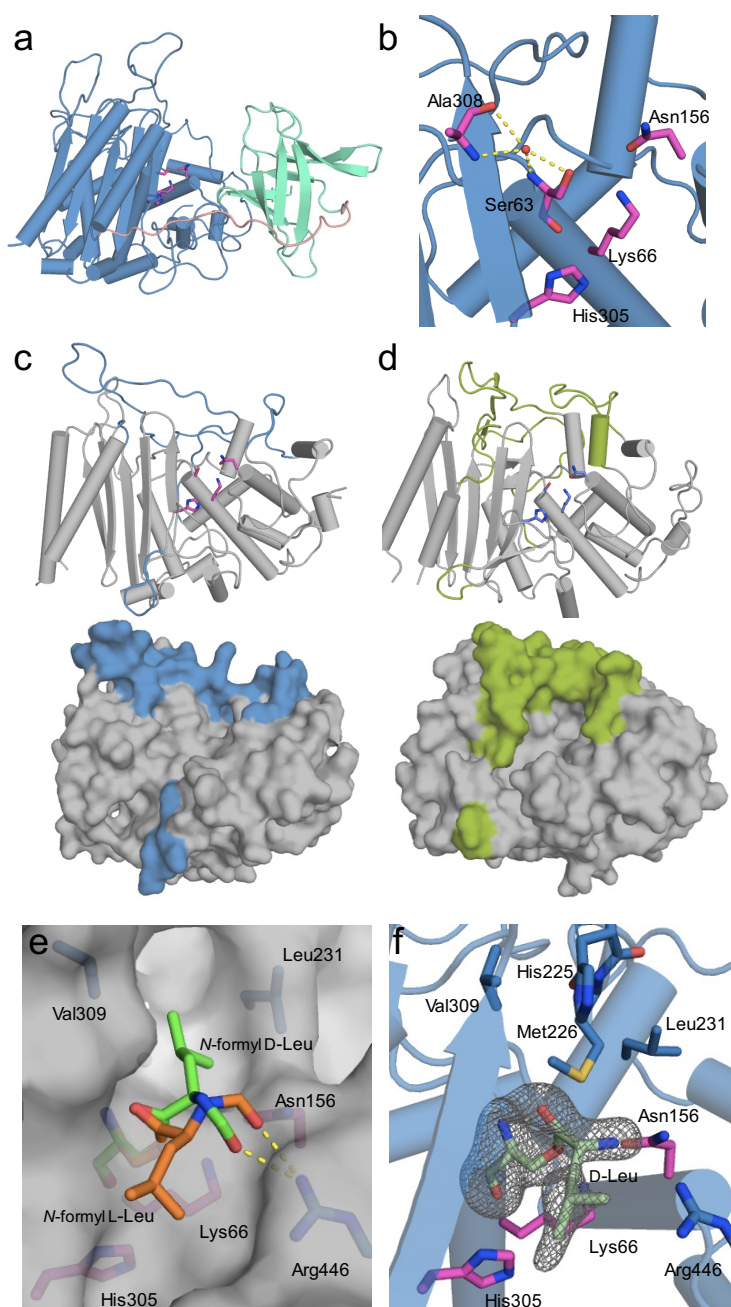
and oxyanion hole residues are shown as magenta sticks. A water molecule is depicted as a red sphere. The hydrogen bond networks are shown as yellow dots. (c and d) Comparison of the PBP domains between SurE

and ADP. The different loop regions in SurE and ADP are colored blue and green, respectively. (e) The result of the conformational calculations of *N*-formyl L-Leu and *N*-formyl D-Leu in the active site of SurE. The hydrogen bond networks are shown as yellow dots. (f) Close-up views of the active site of SurE in complex with D-Leu. The omit |Fo-Fc| electron density map of the ligands is represented as a gray mesh, contoured at +3.0 sigma.

SurE as a genetic tool for engineering the NRP assembly line.

Taking advantage of the considerable promiscuity exhibited by SurE, we investigated its applicability to metabolic pathway engineering. Reprogramming the multi-modular assembly line by altering the specificity of domains or swapping/combining modules is a promising strategy to diversify the chemical structures of NRPs. However, the manipulation of an NRPS often results in an impaired or non-functional assembly line.[29] Although recent efforts are opening the door to reliable methodologies for flexible engineering,[30, 31, 32] the strict specificity of the *cis*-acting offloading domain still remains as an obstacle.

To examine whether SurE is functional in an engineered assembly line, we disrupted modules 5 and 6 of SurC, which are responsible for the extension of the *C*- terminal dipeptidyl unit L-Val9-D-Ala10 of decapeptides (cyclosurugamide F and surugamide F) (Supplementary Figure 11a and b). An LC-MS analysis of the genetically engineered strain revealed that, while the production of cyclosurugamide F and its linear derivative were abolished, a new cyclooctapeptide **13**, and its linear derivative **14** accumulated (Figure 3, Supplementary Figure 11c, d and e). Structure of **13** was validated by comparison with synthetic standard (Supplementary Figure 11f). This result indicates that SurE interacts with a non-cognate PCP at module 4 of SurC to offload the non-native substrate. Together with our biochemical analysis, SurE should be capable of efficiently cyclizing PCP-bound peptide ranging from octa- to decapeptide with proper heterochiral-, hydrophobic

14

residues in both termini, *in vivo*. The PBP-type TEs will provide valuable opportunities to realize the flexible

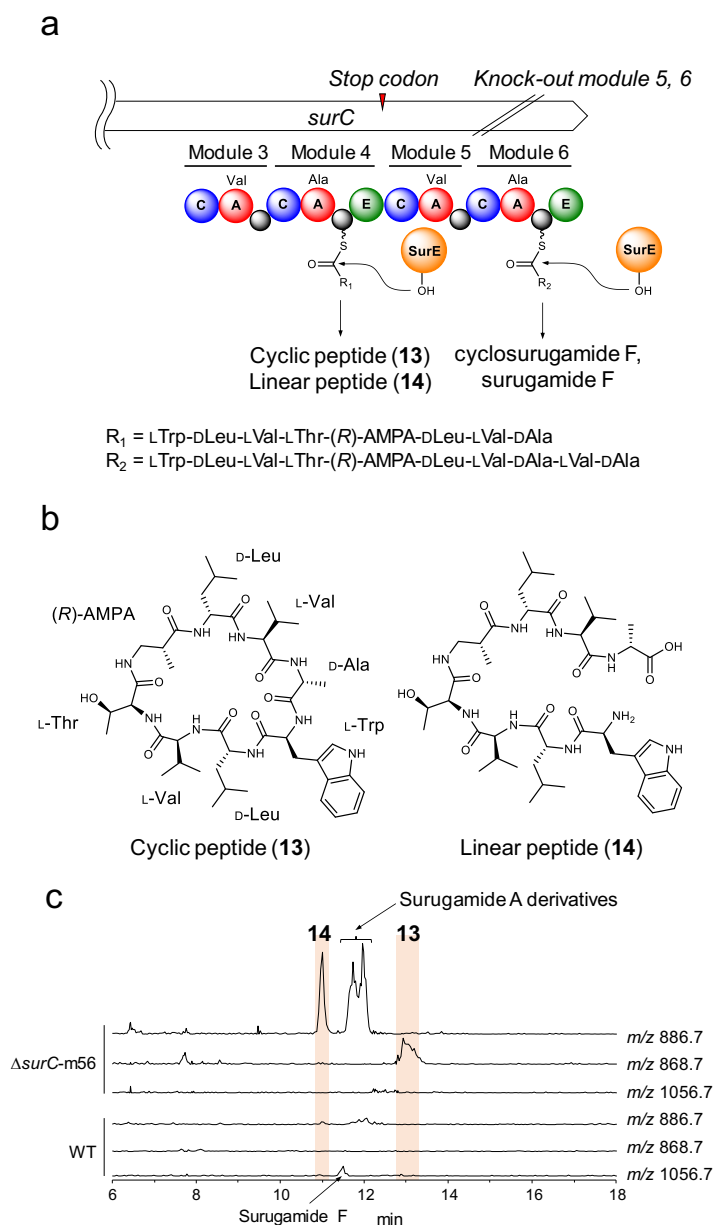engineering of NRPSs for the production of non-natural macrolactams.

Figure 3. Engineering *sur*NRPS revealed the plasticity of SurE. a) Offloading of cyclosurugamide F,

surugamide F from SurC (wt), and **13**, **14** from engineered SurC lacking modules 5 and 6. AMPA

stands for 3-amino-2-methylpropionic acid. b) Structures of the newly biosynthesized NRPs **13** and

14. c) LC-MS analyses of Δ*surC*-m56. Extracted ion chromatographs for **13**, **14** and surugamide F are depicted.

SurE homologs are widely distributed in bacteria.

PBP-type TEs are widely distributed among *Actinobacteria* and some bacteria from different phyla. A phylogenetic analysis classified this family into two large classes, according to the domain architecture (Supplementary Figure 12). Those belonging to the class I PBP-type TEs harbor a lipocalin domain at the *C*-terminus. This class consists of stand-alone cyclases and a few PBP-type TEs fused at the *C*-terminus of the cognate NRPS. All PBP-type TEs, such as SurE, DsaJ, MppK, and Ulm16, which are linked to the biosynthetic product, belong to this class I. They are further subdivided into several clades. The module architecture of the neighboring NRPSs revealed that the class I cyclases form distinct clades correlated to the substrate sequence and length. This fact emphasizes the chemical diversity of the peptide substrates that could be cyclized by this family and suggests that this family could be exploited not only as genetic probes for mining new cyclic NRPs, but also as catalytic tools for synthesizing various cyclic peptides. Indeed, two recombinant SurE homologs with biosynthetic products that are currently unknown exhibited cyclization activities against the surugamide type-substrate **1** and its derivatives **10**, **11** (Supplementary Figure 13). Noteworthy, head-to-side-chain cyclization was observed when **11** was incubated with $SurE_{15452}$, indicating that PBP-type TE catalysis is not limited to head-to-tail cyclization. This result indicates the considerable substrate promiscuity of this family of cyclases. Class II is a group of PBP-type TEs lacking the lipocalin domain. All of the class II PBP-type TEs are fused to the *C*-terminus of the cognate NRPS. Although no natural product has been associated with a class II PBP-type TE yet, some are integrated into the PKS/NRPS hybrid systems, suggesting that the natural products synthesized by the PBP-type TE pathway are not limited to peptides.

Discussion

The PBP-type TEs are a newly identified family of NRP offloading cyclases that act in *trans* to peptide assembly lines. More than 200 SurE homologs are associated with the bacterial *trans*-TE NRPSs. Two of them indeed catalyzed the head-to-tail macrolactamization of surugamide B, suggesting that these homologs are mainly offloading macrocyclases. Among the known PBP-type TEs, SurE is particularity notable because it interacts with two assembly lines to offload and cyclize a structurally distinct acyclic substrate, suggesting its wide substrate tolerance. In this study, our detailed analysis on the substrate scope of SurE, using an array of synthetic substrate derivatives, revealed that both termini of the substrate are important for SurE recognition, while it exhibits promiscuity with the residues in the middle. Furthermore, the relaxed substrate specificity of SurE enabled us to produce non-natural NRPs by manipulating the *sur* assembly line. This clearly highlights SurE's potential as a genetic tool for synthetic biology, providing a significant opportunity to develop a new methodology of NRPS engineering.

Notably, SurE specifically catalyzed the head-to-tail macrolactamization with L- and D-configured residues at the *N*- and *C*-terminus, respectively, which is consistent with the fact that the all *trans*-TE NRPS assembly lines are terminated by the E domain. On the other hand, the *cis*-TE NRPS assembly lines responsible for head-to-tail macrocyclic peptide biosyntheses usually start with the first module containing the E domain, indicating that the canonical macrolactam-forming enzymes, type I TEs and C_{TS}, catalyze heterochiral coupling with the opposite stereochemistry to the PBP-type TEs, as confirmed by the *in vitro* functional analysis of tyrocidine TE[5]. Thus, a suite of *cis*- and *trans*-TE NRPSs can overcome the topological constraints present in heterochiral head-to-tail macrolactam formation in bacterial NRPS biosynthesis and diversify the architecture of the assembly lines.

The crystallographic analysis of SurE provided structural data for this family of enzymes, and together with the conformational calculations, revealed a structural rationale for the stereoselectivity exerted by SurE.

17

Although the source of the heterochiral preference in the enzymatic head-to-tail macrolactamization remains obscure, similar phenomena have also been reported in chemical head-to-tail macrolactamization.[2, 3] Hill *et al.* demonstrated that the condensation between substrates with opposite configurations is favored in competitive activated amide coupling using racemic amino acid derivatives.[1] They postulated that the substrates form a six-membered transitional state with a chair conformation, and that D/L-configured substrates would enable the more favored di-equatorial disposition of the side chains, leading to a significant preference for heterochiral coupling. A similar mechanism may also work in the enzymatic head-to-tail macrolactamization of NRP biosynthesis. This idea requires structural validations in the future.

In conclusion, our results have demonstrated the notable potentials of the PBP-type TEs as peptide-cyclizing biocatalysts, genetic tools for synthetic biology, and genetic probes for mining novel macrolactams or PKS/NRPS hybrid natural products, and set the stage for further investigations of this fruitful family of enzymes.

Methods

General remarks. ^{1}H and ^{13}C NMR spectra were recorded on a JEOL ECA 500 (500 MHz for ^{1}H NMR) or a JEOL ECX 400P (400 MHz for ^{1}H NMR) spectrometer. Chemical shifts are denoted in δ (ppm) relative to residual solvent peaks as internal standard (CDCl$_3$, [1]H δ 7.24, [13]C δ 77.0, DMSO-d_6, [1]H δ 2.50, [13]C δ 39.5). ESI-MS spectra were recorded on a Thermo Scientific Exactive mass spectrometer or a SHIMADZU LCMS-2020 spectrometer. Optical rotations were recorded on a JASCO P-1030 polarimeter. High performance liquid chromatography (HPLC) experiments were performed with a SHIMADZU HPLC system equipped with a LC-20AD intelligent pump and an Agilent Technologies HPLC system equipped with a 1120 Compact LC intelligent pump. LC-MS/MS experiments were performed with amaZon SL-NPC (Bruker Daltonics). Fragmentation of precursor ions were also performed with amaZon SL-NPC using helium gas with amplitude value 1.0 V. All reagents were used as supplied unless otherwise stated. Analytical thin-layer chromatography (TLC) was performed using E. Merck Silica gel 60 F254 pre-coated plates. Column chromatography was performed using 40-50 μm Silica Gel 60N (Kanto Chemical Co., Inc.). *E. coli* DH5α was used as a host for general cloning. Oligonucleotides used for genetic manipulation were purchased from Fasmac Co., Ltd. (Kanagawa, Japan).

Peptide-SNAC synthesis by solid-phase peptide synthesis (SPPS). Fmoc-peptide-2-chlorotrityl resin (0.0500 mmol) in Libra tube was swelled in CH$_2$Cl$_2$ for 1 h, which was subjected to the following SPPS protocol (Steps 1–4). Step 1: Fmoc group of the solid supported peptide was removed by using 20% piperidine/DMF solution (10 min, room temperature). Step 2: The resin in the reaction vessel was washed with DMF ($\times$ 3) and CH$_2$Cl$_2$ ($\times$ 3). Step 3: To the solution of carboxylic acid (4 eq.) were added DIC (4 eq., 0.50 M in NMP) and Oxyma (4 eq., 0.50 M in DMF). After 2-3 min of pre-activation, the mixture was injected to the reaction vessel. The resulting mixture was stirred for 30 min at 37 °C. Step 4: The excess unreacted amino acid and other reagents were filtered-off and the resin in the reaction vessel was washed with DMF ($\times$

3) and CH$_2$Cl$_2$ ($\times$ 3). Amino acids were condensed onto the solid support by repeating Steps 1–4. For cleavage of the peptide from 2-chlorotrityl chloride resin, a cleaving cocktail made of 30% HFIP in CH$_2$Cl$_2$ (0.5 mL) was introduced into the libra tube. After being stirred for 30 min at room temperature, the reaction mixture was filtered and washed with 30% HFIP in CH$_2$Cl$_2$ (0.5 mL). Following this, the filtrate was collected in round bottom flask and concentrated and azeotroped with toluene ($\times$ 3) to afford crude peptide, which was used in the next reaction without further purification. To the peptides in DMF (1 mL) were added *N*-acetylcysteamine (2 eq.), 2,4,6-collidine (1 eq.), HCTU (1 eq.). After being stirred overnight at room temperature, the reaction mixture was quenched with saturated aqueous NH$_4$Cl. The resulting solution was extracted with EtOAc ($\times$ 3). The organic layer was washed with brine, dried over MgSO$_4$, and filtered. The solution was concentrated and azeotroped with toluene ($\times$ 3). To the residue was added a cleaving cocktail of TFA/H$_2$O/^{i}Pr$_3$SiH = 95:2.5:2.5 (1.00 mL) to remove the protecting group. After complete of reaction, the reaction mixture was filtered, and diluted with Et$_2$O (24 mL), centrifuged at 3,500 $\times$ *g* for 5 min at 4 °C, and Et$_2$O layer was removed by decantation. This procedure was repeated twice. The crude peptide was purified by reversed-phase HPLC to afford peptides. Detailed characterization of synthetic products were given in supplementary information.

Synthesis of peptide methylester. To the linear octapeptide prepared by SPPS approach in DMF (1 mL) were added DIPEA (20 eq.) and iodomethane (20 eq.). After being stirred for 15 min at 100 °C with microwave, the reaction mixture was quenched with saturated aqueous NH$_4$Cl. The resulting solution was extracted with EtOAc ($\times$ 3). The organic layer was washed with brine, dried over MgSO$_4$, and filtered. The solution was concentrated and azeotroped with toluene ($\times$ 3). To the residue was added a cleaving cocktail of TFA/H$_2$O/^{i}Pr$_3$SiH = 95:2.5:2.5 (1.00 mL) to remove the protecting group. After being stirred for 1 min, the reaction mixture was filtered, and the reaction mixture was diluted with Et$_2$O (24 mL), centrifuged at 3500 $\times$ *g* for 5 min at 4 °C, and Et$_2$O layer was removed by decantation. This procedure was repeated twice. The crude peptide was purified by reversed-phase HPLC to afford peptides as a white solid.

Preparation of recombinant SurE for assay. The recombinant SurE (GENBANK protein id: BBZ90014.1) was prepared following the same procedure previously described.[8] Briefly, DNA fragment cording for SurE was amplified by KOD Fx Neo (Toyobo) using a set of primer SurE_Fw/SurE_Rv (SurE_Fw: CCGGAATTCCATATGGGTGCCGAGGGGGCG, SurE_Rv: CCCAAGCTTTCAGAGCCGGTGCATGGC, restriction enzyme sites were underlined). The genomic DNA of *Streptomyces albidoflavus* NBRC12854 was used as template. Amplified fragments were digested by *Eco*RI and *Hind*III and inserted into multicloning site of pUC19 (Takara). After sequence confirmation, the fragment was transferred to *Nde*I/*Hind*III site in multicloning site of expression vector pET28a (EMD Millipore) to generate SurE-pET28a. Expression of SurE-pET28a in *E. coli* BL21(DE3) followed by Ni-NTA affinity column chromatography yielded recombinant SurE with His-tag fused at its *N*-terminus. Imidazole in eluted fraction was removed by Amicon® Ultra 0.5mL filter (Merck Millipore). The concentration of proteins were measured by Bio-Rad protein assay kit (Bio-Rad).

Kinetic assay. To determine the kinetic parameters of SurE, 50 μL of reaction buffer containing 20 mM Tris-HCl (pH 8.0), 500 nM SurE and 12.5-800 μM substrates were incubated at 30 °C for 2 min. For the reaction mixture containing I1D*allo*I (**2**) and I1P (**4**), the mixtures were incubated at 30 °C for 10 min. The reaction was quenched by adding 50 μL of 0.1% TFA. A volume of 20 μL was loaded onto an analytical column J-Pak core C_{18} 2.7μm, 2.1x100 mm (JASCO, Japan) and samples were separated by HPLC with monitoring at 215 nm. H_2O + 0.05% TFA and acetonitrile + 0.05% TFA were used as mobile phase A and B, respectively. For the reactions with I1D*allo*I (**2**) and I1P (**4**), samples were eluted at 0.4 ml/min by isocratic mode: 30% for mobile phase B. For the reactions with other substrates, samples were eluted by gradient mode: 30 to 70% for mobile phase B in 5 min. The concentrations of peptides were estimated based on the extinction coefficient e (215 nm) that were experimentally determined. E (215 nm) values for all SNAC variant were assumed to be equal except for I1W (**3**), F7A (**9**), L8F (**11**). The initial velocities were calculated assuming equal e (215nm) value for SNAC-substrates and corresponding cyclic products. All reaction was carried out in triplicates.

Kinetic parameters were estimated using Michaelis-Menten equation and curve-fitting program Kaleidagraph (Supplementary Figure 5). All enzymatic products shown in Supplementary Figure 2 were characterized by ESI-MS/MS analysis as shown in Supplementary Figure 4.

Enzyme expression and purification for crystallization. The cells harboring the expression plasmid was cultured at 37 ˚C to an OD_{600} of 0.6 in LB medium, containing 50 μg/ml kanamycin. IPTG was then added to 0.1 mM (final concentration), to express the target proteins, and the cultures were continued for 16 h at 16 ˚C. All of the following procedures were conducted at 4 ˚C. The cultured cells were harvested by centrifugation at $6,000 \times g$ and resuspended in 50 mM Tris–HCl buffer, pH 8.0, containing 10% (v/v) glycerol, 300 mM NaCl, and 10 mM imidazole (buffer A). The cells were disrupted by sonication, and the insoluble debris was removed by centrifugation at $12,000 \times g$ for 45 min. The supernatant was loaded onto a HisPurTM Ni-NTA Resin (Thermo Fisher Scientific) column. The resin was washed with buffer A containing 20 mM imidazole, and then the recombinant SurE was eluted with buffer D containing 300 mM imidazole. The protein solution was further purified by gel-filtration chromatography on a HiLoad 16/60 Superdex 200 pg column (GE Healthcare). The eluted protein was concentrated to 10 mg/ml in 25 mM Tris–HCl buffer, pH 8.0, containing 5% (v/v) glycerol, 50 mM NaCl, and 1 mM DTT. The concentration of each protein was calculated by measuring the ultraviolet absorption at A280[33].

Crystallization and structure determination of SurE. The crystals of wild-type and selenomethionine-labeled SurE were obtained at 100 mM HEPES (pH 7.5), containing 1260 mM ammonium sulfate, 10 mM Yttrium(III) Chloride and 10 mM tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, with 15 mg/ml of the purified SurE solution, by using the sitting-drop vapor-diffusion method at 4 ˚C. The crystals were transferred into the cryoprotectant solution (reservoir solution with 20% (v/v) glycerol), and then flash cooled at -173 ˚C in a nitrogen-gas stream. We used BL-1A at the Photon Factory (Tsukuba, Japan) and the beamlines TPS 05 at NSRRC (Taiwan, ROC) to collect X-ray diffraction data sets. The complex structure

containing substrate was prepared by incubating SurE crystals for 2 h at 20 ˚C with I1P (**4**). We used wavelength of 0.9788 Å for SurE (SeMet), based on the fluorescence spectrum of the Se K absorption edge[34], and 1.0000 Å was used for data collections of wild-type SurE and complex structure. The diffraction data sets for SurE were processed and scaled using the XDS program package[35] and Aimless[36]. The determination of Se sites and generation of initial model were performed with Crank2 in CCP4[37]. The initial phases of the SurE apo structure and binary complex structure were determined by molecular replacement, using the SurE SeMet as the search model. Molecular replacement was performed with Phaser in PHENIX[38]. The initial phases were calculated with AutoBuild in PHENIX[39]. The structure was modified manually with Coot[40] and refined with PHENIX[41]. The final crystal data and intensity statistics are summarized in Supplementary table 1. The Ramachandran statistics are as followed: 95.3% favored, 4.7% allowed for SurE apo; 94.4% favored, 5.2% allowed, and 0.4% outliners for SurE complexed with **4**; A structural similarity search was performed, using the Dali program. All crystallographic figures were prepared with PyMOL (DeLano Scientific, http://www.pymol.org). The parameters of the intermediates for the energy minimization calculation were obtained by the PRODRG server[42].

Conformational calculation of SurE-tethered *N*-formyl D/L-Leu and I1P (4). The model of *N*-formyl D-Leu-tethering SurE was constructed in Maestro (Schrödinger), using residues located within 8.0 Å from catalytic residue Ser63 in apo structure of SurE. For the construction of the model of **4**-*O*-SurE, the structure of SurE tethering D-Leu was used. Conformational search was performed by Monte Carlo multiple minimum (MCMM) method (1000 steps) with forcefield OPLS3, followed by minimization with Polak-Ribiere conjugate gradient (PRCG) method. Residues other than tethered *N*-formyl D/L-Leu or *N*-terminal seven residues of **4** were fixed during calculation. Also, the distance between the carbonyl *O* atom of D-Leu and amide protons of oxyanion-forming residues (Ser63, Ala308) were constrained at 2.0 Å with force constant at 100. Other parameters were used as default. The computational models for SurE-tethered D-Leu, L-Leu or entire I1P (**4**) together with residues surrounding them were provided as Supplementary Data.

23

Construction of SurE variant H225A. The DNA fragment coding for SurE variant H225A was amplified with a set of primers; forward: 5'-GCCGAGCACGTCGACGCCATGGCCTCGCTCTCG-3', reverse: 5'-CGAGAGCGAGGCCATGGCGTCGACGTGCTCGGC-3' using SurE-pUC19 as a template and KOD One® PCR Master Mix -Blue- (TOYOBO) as polymerase. The amplified fragment was introduced into *E. coli* DH5α and a circular plasmid was extracted from cell. After validating nucleotide sequence, the insert coding for SurE H225A was transferred to multi-cloning site of pET28a using a pair of restriction enzymes *Eco*RI/*Hind*III, to give an expression vector for *N*-terminal His tag-fused SurE H225A. Procedures for expressing, purifying enzyme, and assay conditions are same as that for wild type enzyme.

Genetic manipulation of *surC* and metabolite analysis of manipulated strain. To disrupt the function of module 5 and 6 of SurC, the corresponding genetic region was replaced with neomycin resistance gene (*aphII*), and stop codon was inserted in the linker region between module 4 epimerization domain and module 5 condensation domain (Supplementary Figure 11a). PKS/NRPS analysis web tool[43] was used to analyze the domain architecture of SurC. The approx. 1.5 kb upstream with additional artificial stop codon and downstream of module 5 and 6 of surC were amplified from genomic DNA of *Streptomyces albidoflavus* NBRC12854 using KOD One® PCR Master Mix -Blue-. The upstream region containing artificial stop codon was amplified using following primers; forward: 5'-gccctcgcgccggccgtc-3' and reverse: 5'-tcactggccgtcagaccgtgcggccgtcccc-3' (underlined characters indicates inserted stop codon). The downstream region was amplified using following primers; forward: 5'-cggtacccggtagccggcccccgcttcc-3' and reverse: 5'-gctcctcggggctgagcatc-3'. Also, the fragment coding for *aphII* were amplified from pKU479[44] using following primers; forward: 5'-cacggtctgacggccagtgaattcgagc-3' and reverse: 5'-gggccggctaccgggtaccgagcgaacg-3'. Finally, vector backbone was amplified from pRED[45] using following primers; forward: 5'-gatgctcagccccgaggagcaagcttatctgccaggaagatacttaac-3' (underlined characters indicate *Hind*III site) and reverse: 5'-tggacggccggcgcgagggcaagcttatcccattcatccgcttattatc-3' (underlined characters indicate *Hind*III site). The four fragments were assembled using Gibson Assembly Master Mix (New England Biolabs), then

24

introduced into *E. coli* DH5α. The approx. 4.5 kb insert (upstream-*aphII*-downstream) were transferred to *Hind*III site of pGM160Δ*aac1::oriT::pheS(A339G/T278S)::sacB::aph*[13] to yield disruption plasmid, pGM160-Δ*surC*-m56. This plasmid was introduced into *E. coli* GM2929 hsdS::Tn*10*/pUB307::Tn7[44] and further transferred to *S. albidoflavus* NBRC12854 by conjugation. The desired strain (Δ*surC*-m56) with disruption of module 5 and 6 of *surC* was obtained by positive selection using neomycin, and counter-selection using 4-chloro-DL-phenylalanine. The detailed procedure for counter-selection is described on previous report[13]. The disruption of module 5 and 6 of *surC* was confirmed based on PCR amplification using following primers; forward: 5'-caccccctccgacttcccgctg-3' / reverse: 5'-ctgggcgaacgaggcgggc-3'. Δ*surC*-m56 was inoculated into 10 mL of TSB liquid media containing 25 µg/ml neomycin and cultivated for 2-4 days at 30 ˚C as pre-culture. 1 mL of the pre-culture was then transferred to a flask containing 100 mL K media (2.5% soluble starch, 1.5% soybean meal, 0.2% dry yeast, 0.4% $CaCO_3$ and the pH was adjusted to 6.2) and cultivated at 30 ˚C with shaking at 140 rpm for 5 days. The broth was extracted with double volume of acetone, then filtered to remove debris and evaporated to remove solvent. The residue was dissolved into DMSO and subjected to LC-MS (amaZon SL-NPC) operated in positive mode, coupled with a SHIMADZU HPLC system. Separation was accomplished by COSMOSIL $5C_{18}$-MS-II 2.0×150 mm column (nacalai tesque). H_2O + 0.05 % TFA and acetonitrile + 0.05 % TFA were used as mobile phase A and B, respectively. Samples were eluted with the gradient mode: 10 to 90% for mobile phase B in 20 min with flow rate at 0.2 mL/min.

Cloning and protein expression of PBP-type TEs. The DNA fragment coding for PBP-type TE from *Goodfellowiella coeruleoviolacea* NBRC14988 (Gene ID of $SurE_{14988}$ in JGI database: 2596754607) was amplified from genomic DNA of *G. coeruleoviolacea* NBRC14988 using following primers; forward: 5'-ccggaattcGTGCCCAACGAGCAGGATCTCGGG-3' (underlined characters indicate *Eco*RI site), reverse: 5'-cccaagcttTCATCCGGTCACCTGCCGCCGC-3' (underlined characters indicate *Hind*III site) and KOD One® PCR Master Mix -Blue-. The fragment was transferred to *Eco*RI/*Hind*III sites of expression vector pET28a (EMD Millipore) to generate $SurE_{14988}$-pET28a for the expression of *N*-terminal His-tag fused

SurE$_{14988}$. The DNA fragment coding for PBP-type TE from *Streptomyces noursei* NBRC15452 (Gene ID of SurE$_{15452}$ in JGI database: 2603776488) was amplified from genomic DNA of *S. noursei* NBRC15452 using following primers; foward: 5'- ccggaattcGTGCACGGGGACTCAGCGGATCC-3' (underlined characters indicate *EcoR*I site), reverse: 5'-cccaagcttTTAGTGCGGCCGTGCGCCGTGG-3' (underlined characters indicate *Hind*III site) and KOD One° PCR Master Mix -Blue-. The fragment was transferred to *EcoR*I/*Hind*III sites of pET28a to generate SurE$_{15452}$-pET28a for the expression of *N*-terminal His-tag fused SurE$_{15452}$. Both SurE$_{14988}$ and SurE$_{15452}$ were expressed and purified in the same manner as SurE described in previous section.

Cyclization assays of SurE, SurE$_{14988}$ and SurE$_{15452}$. 50 μL of reaction mixture containing 20 mM Tris-HCl (pH 8.0), 1.0 μM PBP-type TE (SurE or SurE$_{14988}$ or SurE$_{15452}$) and 20 μM substrates were incubated at 30 ˚C for 2 h. The reaction mixtures were quenched with equal volume of 0.1 % TFA and centrifuged at 20,400 × *g* for 10 min. The analyses of the resultant supernatants were performed by LC-MS (amaZon SL-NPC) operated in positive mode, coupled with a SHIMADZU HPLC system. Separation was accomplished by COSMOSIL 5C$_{18}$-MS-II 2.0×150 mm column (nacalai tesque). H$_2$O + 0.05 % TFA and acetonitrile + 0.05 % TFA were used as mobile phase A and B, respectively. Samples were eluted with the gradient mode: 10 to 90% for mobile phase B in 20 min with flow rate at 0.2 mL/min.

Data availability.

The crystallographic data that support the findings of this study are available from the Protein Data Bank (http://www.rcsb.org). The coordinates and the structure factor amplitudes for the structures of SurE apo and SurE complexed with **4** were deposited under accession codes 6KSU and 6KSV, respectively. All other data supporting the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

Author Contributions

K.M., T.M., I.A. and T.W. designed the experiments. R.Z. and T.M. performed crystallization experiment. K.M., R.Z., T.M. and M.K. performed in vitro analysis. K.M. measured pre-steady-state kinetic values. M.K., and A.S. synthesized substrate analogs. K.M., M.K. and T.W. performed structure determination of enzyme reaction products. M.K. performed genetic manipulation of *Streptomyces*. K.M., R.Z., T.M., M.K., A.S., I.A. and T.W. analyzed the data. K.M., T.M., I.A. and T.W. wrote the paper.

Competing interests

The authors declare no competing interests

REFERENCES

(1) Hill, R.R.; Birch; D.; Jeff, G.E.; North, M. Enantioselection in peptide bond formation. *Org. Biomol. Chem.* **1**, 965–972 (2003).

(2) Brady, S.F. et al. Practical synthesis of cyclic peptides, with an example of dependence of cyclization yield upon linear sequence. *J. Org. Chem.* **44**, 3101–3105 (1979).

(3) Rich, D.H.; Bhatnagar, P.; Mathiaparanam, P.; Grant, J.A.; Tam, J.P. Synthesis of tentoxin and related dehydro cyclic tetrapeptides. *J. Org. Chem.* **43**, 296–302 (1978).

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

(5) Trauger, J.W.; Kohli, R.M.; Mootz, H.D.; Marahiel, M.A.; Walsh, C.T. Peptidecyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **2000**, *407*, 215–218.

(6) Gao, X. et al. Cyclization of fungal nonribosomal peptides by a terminal condensation-like domain. *Nat. Chem. Biol.* **8**, 823–830 (2012).

(7) Tsomaia, N. Peptide therapeutics: targeting the undruggable space. *Eur. J. Med. Chem.* **94**, 459–470 (2015).

(8) Kuranaga, T. et al. Total synthesis of the nonribosomal peptide surugamide B and identification of a new offloading cyclase family. *Angew. Chem., Int. Ed.* **57**, 9447–9451 (2018).

(9) Takada, K. et al. 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 (2013).

(10) Mohiman, H. et al. Dereplication of peptidic natural products through database search of mass spectra. *Nat. Chem. Biol.* **13**, 30–37 (2016).

(11) Xu, F.; Nazari, B.; Moon, K.; Bushin, L.B.; Seyedsayamdost, M.R. Discovery of a cryptic antifungal compound from *Streptomyces albus* J1074 using high-throughput elicitor screens. *J. Am. Chem. Soc.* **139**, 9203–9212 (2017).

(12) Ninomiya, A. et al. Biosynthetic gene cluster for surugamide A encompasses an unrelated decapeptide, surugamide F. *ChemBioChem* **17**, 1709–1712 (2016).

(13) Matsuda, K. et al. SurE is a *trans*-acting thioesterase cyclizing two distinct non-ribosomal peptides. *Org. Biomol. Chem.* **17**, 1058–1061 (2019).

(14) Magarvey, N.A.; Haltli, B.; He, M.; Greenstein, M.; Hucul, J.A. Biosynthetic pathway for mannopeptimycins, lipoglycopeptide antibiotics active against drug-resistant gram-positive pathogens. *Antimicrob. Agents Chemother.* **50**, 2167–2177 (2006).

(15) Li, Q. et al. Identification of the biosynthetic gene cluster for the anti-infective desotamides and production of a new analogue in a heterologous host. *J. Nat. Prod.* **78**, 944–948 (2015).

(16) Son, S. et al. Genomics-driven discovery of chlorinated cyclic hexapeptides ulleungmycins A and B from a *Streptomyces* species. *J. Nat. Prod.* **80**, 3025–3031 (2017).

(17) Zhou, Y. et al. Investigation of penicillin binding protein (PBP)-like peptide cyclase and hydrolase in surugamide non-ribosomal peptide biosynthesis. *Cell Chem. Biol.* **26**, 737–744 (2019).

(18) Thankachan, D. et al. A *trans*-acting cyclase offloading strategy for nonribosomal peptide synthetases. *ACS Chem. Biol.* **14**, 845–849 (2019).

(19) White, C.J.; Yudin, A.K. Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **3**, 509–524 (2011).

(20) Yu, X.; Sun, D. Macrocyclic drugs and synthetic methodologies toward macrocycles. *Molecules* **18**, 6230–6268 (2013).

(21) Medema, M.H. *et al.* Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* **11**, 625–631 (2015).

(22) Schultz, A.W. et al. Biosynthesis and structures of cyclomarins and cyclomarazines, prenylated cyclic peptides of marine actinobacterial origin. *J. Am. Chem. Soc.* **130**, 4507–4516 (2008).

(23) Ma, J. et al. Biosynthesis of ilamycins featuring unusual building blocks and engineered production of enhanced anti-tuberculosis agents. *Nat. Commun.* **8**, 391 (2017).

(24) Holm, L.; Rosenstrom, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–549 (2010).

(25) An, Y.J.; Kim, M.K.; Na, J.H.; Cha, S.S. Crystal structure of a class C beta lactamase. RCSB Protein Data Bank DOI: 10.2210/pdb5F1F/pdb (2016).

(26) Lahiri, S.D. et al. Structural insight into potent broad-spectrum inhibition with reversible recyclization mechanism: avibactam in complex with CTX-M-15 and *Pseudomonas aeruginosa* AmpC *β*-lactamases. *Antimicrob. Agents Chemother.* **57**, 2496–2505 (2013).

(27) Nakano, S. et al. Structural and computational analysis of peptide recognition mechanism of class-C type penicillin binding protein, alkaline D-peptidase from *Bacillus cereus* DF4-B. *Sci. Rep.* **5**, 13836–13836 (2015).

(28) Delfosse, V. et al. Structure of the archaeal Pab87 peptidase reveals a novel self-compartmentalizing protease family. *PLoS One* **4**, e4712–e4712 (2009).

(29) Calcott, M.J.; Ackerley, D.F. Genetic manipulation of non-ribosomal peptide synthetases to generate novel bioactive peptide products. *Biotechnol. Lett.* **36**, 2407–2416 (2014).

(30) Bozhüyük, K.A.J. et al. *De novo* design and engineering of non-ribosomal peptide synthetases. *Nat. Chem.* **10**, 275–281 (2017).

(31) Niquille, D.L. et al. Nonribosomal biosynthesis of backbone-modified peptides. *Nat. Chem.* **10**, 282–287 (2018).

(32) Bozhüyük, K.A. J. et al. Modification and *de novo* design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nat. Chem.* **11**, 653–661 (2019).

(33) Pace, C.N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **4**, 2411–2423 (1995).

(34) Rice, L.M., Earnest, T.N. & Brunger, A.T. Single-wavelength anomalous diffraction phasing revisited. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56**, 1413–1420 (2000).

(35) Kabsch, W. Xds. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).

(36) Evans, P.R. An introduction to data reduction: space-group determination, scaling and intensity statistic. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 282–292 (2011).

(37) Skubak, P. & Pannu, N.S. Automatic protein structure solution from weak X-ray data. *Nat. Commun.* **4**, 2777 (2013)

(38) McCoy, A.J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

(39) Terwilliger, T.C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 61–69 (2007).

(40) Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

(41) Afonine, P.V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).

(42) Schüttelkopf, A.W., & van Aalten, D.M. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 1355–1363 (2004).

(43) Bachmann, B.O. & Ravel, J. Methods for in silico prediction of microbial polyketide and nonribosomal peptide biosynthetic pathways from DNA sequence data. *Methods Enzymol.* **458**, 181–217 (2009).

(44) Kim, J.H., Komatsu, M., Shin-ya, K., Omura, S. & Ikeda, H. Distribution and functional analysis of the phosphopantetheinyl transferase superfamily in Actinomycetales microorganisms. *Proc. Natl. Acad. Sci. U. S. A.* **115,** 6828–6833 (2018).

(45) Komatsu, M., Uchiyama, T., Omura, S., Cane, D.E. & Ikeda, H. Genome-minimized Streptomyces host for the heterologous expression of secondary metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2646–2651 (2010).