Synthetic biology of non-ribosomal cyclopeptide

Toshiyuki WAKIMOTO

Faculty of Pharmaceutical Sciences, Hokkaido University

Research objective

The macrocyclization of medium-sized molecules is considered one of the most challenging reactions to achieve in organic synthesis. It is difficult to bring both ends of the linear substrate close together, and in the case of peptides, epimerization of the C-terminal amino acid can lead to the formation of byproducts. In the presence of multiple reactive functional groups on the substrate, protecting groups are necessary to achieve the desired regioselectivity. Furthermore, high-dilution conditions are essential owing to the competition between the desired intramolecular reactions and undesired intermolecular reactions. This can pose a serious problem for large-scale synthesis processes. Considering natural products, several compounds with macrocyclic skeletons have been biosynthesized efficiently by biosynthetic enzymes. By discovering naturally occurring cyclases and attempting to modify their functions using enzyme engineering and directed evolution, we could develop cyclases with diverse substrate specificities that can be employed for producing versatile macrocyclic medium-sized molecules.

In the present study, we modified penicillin-binding protein-type thioesterase (PBP-type TE), a novel cyclase originally discovered by our group from an actinomycete, to develop a biocatalyst that can easily synthesize macrocyclic skeletons and realize a biosynthetic platform for producing a variety of cyclic peptides. Given that macrocyclization of peptides and polyketides can confer advantageous physical properties, such as fixation of the active conformation and resistance to digestive enzymes, compared with linear structures, this study provides a fundamental technology for potentially modifying medium-sized drug leads.

Methods

We discovered a new peptide cyclase, SurE, involved in the biosynthesis of surugamide, a non-ribosomal cyclic peptide produced by actinomycetes, and elucidated its substrate selectivity and crystal structure. SurE is a 50 kDa "stand-alone" enzyme independent of non-ribosomal peptide synthetase (NRPS), whereas cyclases of the NRPS pathway are typically fused to the C-terminus of a giant enzyme of several mega Da (MDa). Owing to its homology to PBP, we classified SurE as a novel peptide cyclase family, PBP-type TE. SurE can be easily prepared as a recombinant enzyme and has high potential as a biocatalyst, given its wide substrate selectivity and catalytic efficiency, comparable with that of commercially available peptide cyclization biocatalysts.

PBP-type TE-mediated non-ribosomal peptide biosynthesis is a universal biosynthetic

mechanism distributed in more than 10% of actinomycetes and is responsible for the biosynthesis of cyclic peptides of various structures. These PBP-type TE homologs exhibit high amino acid sequence homology with each other but display completely different substrate selectivities from that of SurE. Herein, we elucidated the molecular basis of the substrate recognition mechanism by biochemically and structurally comparing multiple PBP-type TEs that exhibit high sequence homology but distinct substrate selectivities. Accordingly, we can obtain fundamental knowledge for freely controlling selectivity to design improved biocatalysts with extended substrate and reaction selectivity, as well as develop novel efficient synthetic methods for cyclic peptides.

Results

More than 700 candidate genes for PBP-type TE are available in the public genome databases. The identified genes were subjected to network analysis based on amino acid sequence similarity and found to be classified into clusters of approximately 40 types according to substrate selectivity. This analysis enabled the efficient identification of novel homologous enzymes. Therefore, we selected 10 homologous enzymes representing each clade and attempted to clone and express them heterologously. Plasmids for recombinant protein expression were prepared, and the recombinant enzymes expressed in *Escherichia coli* were purified by affinity chromatography with Ni resin. Subsequently, 8 of the 10 homologous enzymes were obtained as insoluble enzymes. Two proteins were identified as soluble enzymes (Nsm16 and SurE14988). Furthermore, PenA, which is responsible for the cyclization of the cyclic pentapeptide pentaminomycin, was cloned from *Streptomyces cacaoi* NBRC 12748, and heterologous expression was attempted in *E. coli*. Given that PenA provides a cyclic peptide with a smaller ring size than SurE, we performed an in-depth assessment of its substrate specificity. Based on the observed results, although PenA is unable to catalyze the cyclization of surugamide B, it was found to specialize in short-chain substrates (Fig. 1). SurE could cyclize peptides of 5–11 amino acid residues; for tetrapeptide generation, dimerization also occurred in addition to cyclization. In contrast, PenA cyclized peptides with 4–5 amino acid residues and selectively catalyzed the cyclization of tetrapeptides without dimerization (Fig. 1). The PenA model structure revealed a loop structure that may be responsible for its unique selectivity, confirming that PenA could cyclize shorter peptide substrates than SurE.[1] To structurally elucidate the mechanism underlying this unique selectivity, we generated a model structure of PenA using the SurE structure (PDB:6KSU) as a template and compared it with the model structure of SurE. The results revealed that a portion of the loop constituting the C-terminal lipocalin domain of PenA was markedly elongated, protruding toward the substrate-binding pocket (Fig. 1). These loops narrow the substrate pocket volume, which is likely to be more suitable for short peptides.

DsaJ is a PBP-type TE that is responsible for the biosynthesis of cyclic hexapeptide

desotamides. While common PBP-type TEs selectively cyclize substrates with a D-amino acid residue at the C-terminal end, DsaJ cyclizes substrates with an achiral Gly residue at the C-terminal end. To gain insight into the structural basis for this unique selectivity, a model DsaJ structure was constructed using AlphaFold 2 and compared with that of SurE, suggesting that the hydrophobic pocket surrounding the catalytic serine residue is critical for the C-terminal D-amino acid selectivity of the SurE substrate. In the model DsaJ structure, the residues constituting the hydrophobic pocket are replaced with bulky residues. The unique substrate selectivity of DsaJ could be attributed to the pocket structure, and the residues comprising the pocket were replaced by bulky residues in SurE. Accordingly, the mutant enzyme lost its activity against the original substrate (C-terminal residue D-Leu) but was able to cyclize a new substrate with a C-terminal Gly residue. These results indicated that the hydrophobic pocket is pivotal for recognizing the C-terminal residue of the substrate, and the selectivity of the C-terminal residues of SurE and DsaJ was completely switched.

Conclusion

Focusing on the substrate specificity of naturally occurring PBP-type TE, we attempted to perform a detailed functional analysis and functional modification by enzyme engineering, successfully developing new cyclases compatible with various substrates. Currently, we are developing more efficient methods for synthesizing cyclic peptides using these cyclases.

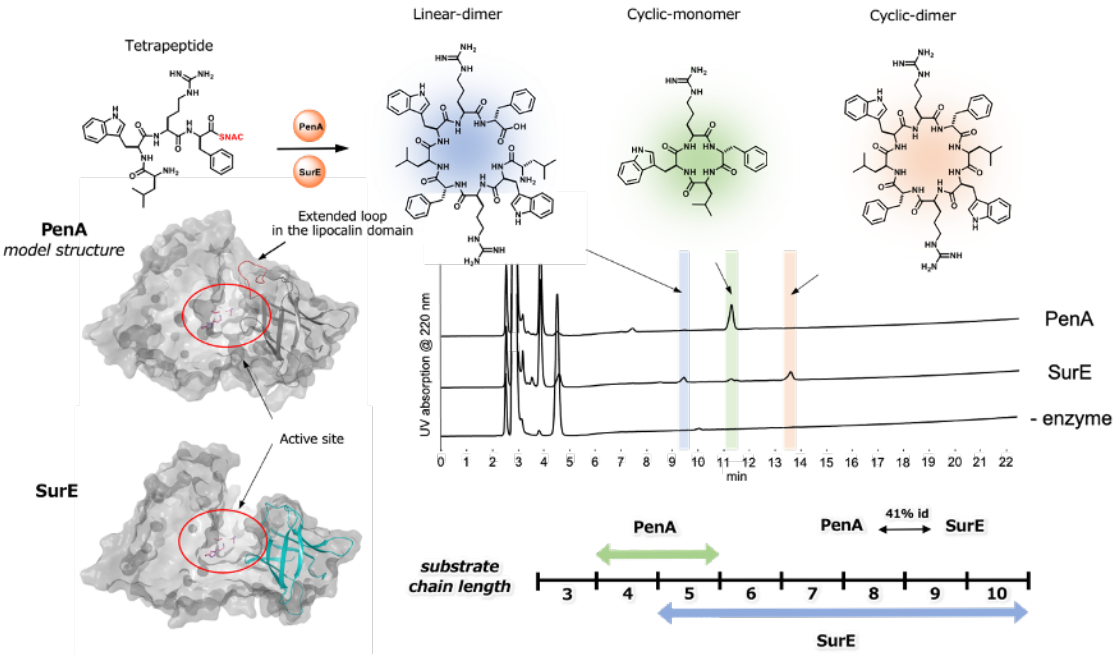

Fig. 1. Functional analysis of PenA

References

1) Matsuda, K., Fujita, K., and Wakimoto, T. (2021) PenA, a penicillin-binding protein-type thioesterase specialized for small peptide cyclization. *J. Ind. Microbiol. Biotechnol.* **48**: kuab023.