ACS Synthetic Biology

pubs.acs.org/synthbio
Research Article

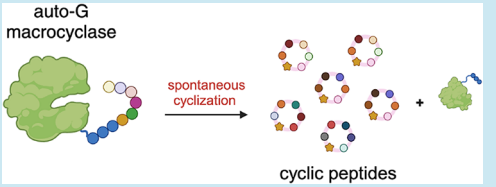
An Autocatalytic Peptide Cyclase Improves Fidelity and Yield of Circular Peptides In Vivo and In Vitro

Noel Lacerna, II, Ying Cong, and Eric W. Schmidt*

Cite This: ACS Synth. Biol. 2024, 13, 394−401
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Peptide cyclization improves conformational rigidity, providing favorable pharmacological properties, such as proteolytic resistance, target specificity, and membrane permeability. Thus, many synthetic and biosynthetic peptide circularization strategies have been developed. PatG and related natural macrocyclases process diverse peptide sequences, generating millions of cyclic derivatives. However, the application of these cyclases is limited by low yields and the potential presence of unwanted intermediates. Here, we designed a covalently fused G macrocyclase with substrates that efficiently and spontaneously release cyclic peptides. To increase the fidelity of synthesis, we developed an orthogonal control mechanism enabling precision synthesis in *Escherichia coli*. As a result, a library comprising 4.8 million cyclic derivatives was constructed, producing an estimated 2.6 million distinct cyclic peptides with an improved yield and fidelity.



KEYWORDS: *cyanobactins, cyclic peptide library, autocatalysis*

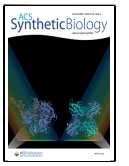
INTRODUCTION

Cyclic peptides have attracted attention as drug candidates because their properties bridge those of biologics and small molecules. As a result of their therapeutic advantages, cyclic peptides are growing rapidly as FDA-approved drug and clinical trial candidates.[1,2] Numerous cyclization strategies have been developed, including synthetic and biosynthetic methods, each of which has advantages in specific applications. For example, the biosynthetic approach enables genetic randomization to construct peptide libraries, which are compatible with selection- or phenotype-based screens directly within the producing organisms.[3−6] For that reason, many genetic methods have been developed. Arguably, the most successful of these has been split-intein circular ligation of peptides and proteins (SICLOPPS).[6] Other tools have been developed, such as modified thioesterase domains.[7,8] Genetic code reprogramming supplements these methods by bringing in nonproteinogenic amino acids.[9,10] Each of these methods has strengths and limitations in the scope of substrate tolerance, production of unwanted mixtures, and relative yield produced in vivo.

The naturally relaxed substrate promiscuity of ribosomal and post-translationally modified peptide (RiPP) pathways has enabled their application to library generation.[11−14] For example, cyanobactin-family macrocyclases, the G enzymes, show exceptionally broad substrate tolerance and high fidelity, efficiently generating a single cyclic peptide without observable linear side products. G enzymes are subtilisin-like proteases that, instead of hydrolysis, perform transpeptidation to form macrocycles. The native substrates for PatG and PagG macrocyclases are short peptides composed of 6−10 amino acids in the hypervariable "core peptide" region that encodes the final circular product and 3−5 amino acids in the C-terminus that comprise the conserved "recognition sequence". The core peptide is directed to the enzyme's active site through interactions with the recognition sequence. Once positioned, G enzymes cleave the peptide to form an enzyme−core-peptide covalent intermediate. The core peptide's free N-terminus attacks the ester intermediate, releasing a circular peptide product. In both PagG and PatG, a heterocyclic residue is required at the C-terminus of the core peptide, while other positions in the core are broadly substrate tolerant. In PagG, the preferred heterocycle is proline, while in PatG, an azoline is preferred; this azoline is usually produced enzymatically.[15−18]

Because of their unusual fidelity and substrate tolerance, G enzymes have been widely applied to synthesizing circular peptide derivatives and large libraries.[12,19] Experiment has demonstrated that G enzymes broadly accept many different substrates and are often capable of making millions of compounds, by just considering proteinogenic amino acids.[11,17] Adding to this diversity, G enzymes are compatible with D-amino acids[18] and nonproteinogenic amino acids,[10] making them amenable to generating compounds and libraries that bridge synthetic and genetic strategies.

Received: October 24, 2023
Revised: December 27, 2023
Accepted: January 2, 2024
Published: January 9, 2024



https://doi.org/10.1021/acssynbio.3c00645
ACS Synth. Biol. 2024, 13, 394−401

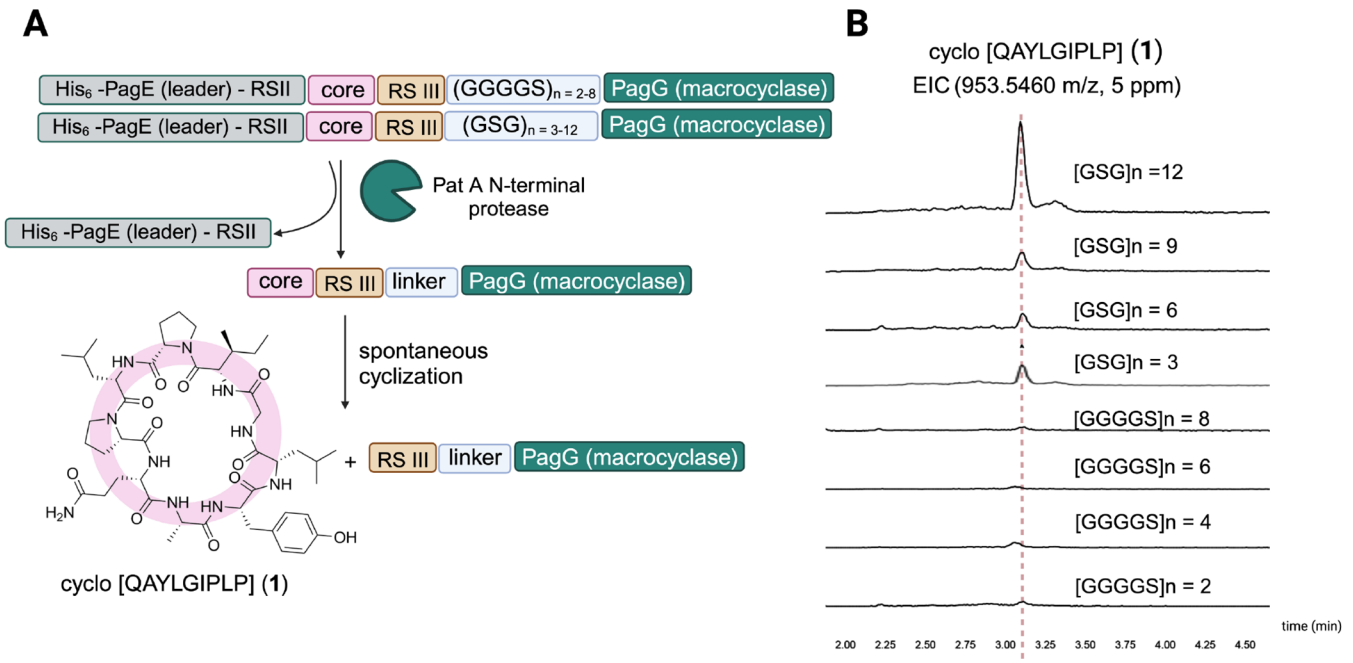


Figure 1. Developing covalently fused PagG macrocyclase enzymes with a prenylagaramide core substrate. (A) Protein constructs containing the artificial PagE precursor peptide were fused with the PagG macrocyclase enzyme through flexible linkers with varying sizes. Treatment of these proteins with PatA protease cleaves the leader sequences, liberating the N-terminus of the core sequences, which then ultimately leads to the rapid release of circular products. (B) LC−MS analysis of the reaction mixture of designed constructs shows cyclic prenylagaramide (**1**) formation at t_R = 3.15 min, with the longest-size GSG linker providing the most efficient production. Scale is constant in all traces and represents total counts.

However, several factors have limited their success. In *Escherichia coli*-based expression libraries, instead of a simple core-recognition sequence substrate, the precursor peptides also contain N-terminal elements that must be processed and cleaved prior to the action of G enzymes, and the reactions are quite slow, creating several difficulties.[20] In vitro, the purified G enzymes and substrates are combined, but reactions are sometimes low-yield and do not always go to completion. This leaves a mixture that is not ideal for screening.[18] We hypothesized that the in vivo and in vitro limitations were largely caused by the substrates and not the inherent efficiencies of the enzymes. For example, in vivo different linear substrates may be readily hydrolyzed in *E. coli* in a sequence-dependent manner. Similarly, in vitro short peptide substrates often have unfavorable solubility, oligomerization, or folding terms, which are also sequence dependent.

If this hypothesis was correct, then we reasoned that covalent fusion of the substrate to the N-terminus of a G protein would significantly increase efficiency, yield, and fidelity. Furthermore, we reasoned that developing orthogonal control of cyclization activity would provide several advantages for in vivo and in vitro peptide circularization. Here, we report the design and implementation of the covalent fusion strategy, leading to the efficient synthesis of cyclic peptides in vivo and in vitro. We generated a mutant library with a high circularization success rate. We also performed a large-scale culture and isolated one of the successful mutants, yielding enough material for full spectral analysis.

RESULTS AND DISCUSSION

Development of Autocatalytic G Macrocyclase. We first used PagG because its substrates do not require enzymatic modification prior to the circularization step.[17] The precursor peptide, PagE,[21] was fused through its C-terminus to different linkers, which in turn were fused with the PagG protease domain. The native core peptide (QAYLGIPLP) and C-terminal recognition sequence ("RSIII") were used, but the N-terminal PagA protease recognition sequence was replaced with the PatA element ("RSII") since the PatA enzyme has been much more widely used (Figure S1).[22] Cleavage of RSII would liberate a free N-terminus that is ready for circularization (Figure 1a). Thus, in principle, after PatA action, autocatalysis would afford the macrocycle. The cyclic peptide is the native PagG product and the precursor of the natural product, prenylagaramide from cyanobacteria.[23]

We tested constructs with variable GlySer linkers (GSG_n n = 3−12 and $GGGGS_n$ n = 2−8) that range between 3 and 40 amino acids.[24,25] Based upon available crystal structures,[15,26] the linker size ranges easily spanned the distance between the substrate and the active site of the enzymes. Eight enzyme constructs with variable linkers were expressed in *E. coli* and purified (Figure S3). To confirm that the PagG enzyme constructs retained activity, we performed an *in-trans* enzyme assay using the PagG native substrate: a short peptide consisting of QAYLGIPLP fused to RSIII. The cyclic peptide was produced by each of the PagG constructs, indicating that the additional N-terminal sequences in the protein do not interfere to its catalytic activity (Figure S4).

The purified proteins were next treated with PatA protease (Figure S5), which cleaves the leader peptide and frees the N-terminus of the core peptide. In principle, this should lead to rapid circularization and the release of cyclic products. In the event, we found that circularized peptides were detected in experiments using constructs with GSG linkers but not those with GGGGS linkers. Circularization was most efficient with a 36 amino acid linker comprising 12 GSG repeats (Figure 1b). Therefore, this linker length was used in all of the further experiments.

https://doi.org/10.1021/acssynbio.3c00645
ACS Synth. Biol. 2024, 13, 394−401

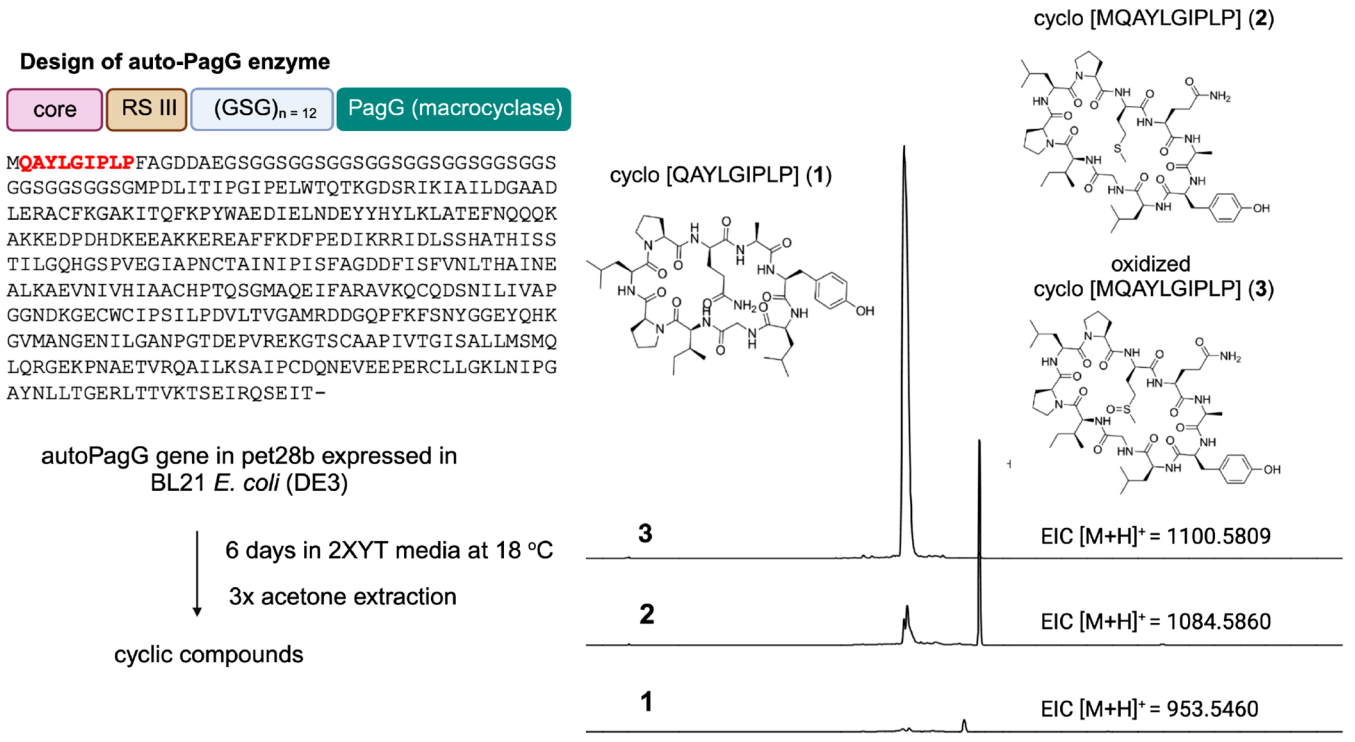


Figure 2. A simplified construct dubbed as "auto-PagG" was designed to produce cyclic peptides in*E. coli* in vivo. A mixture of cyclic peptides containing the core and core with additional methionine was observed from organic extracts in this simplified protein construct, as shown by EIC peaks from LC−MS analysis. Scale is constant in all traces, showing the relative abundance of each peak in the extract.

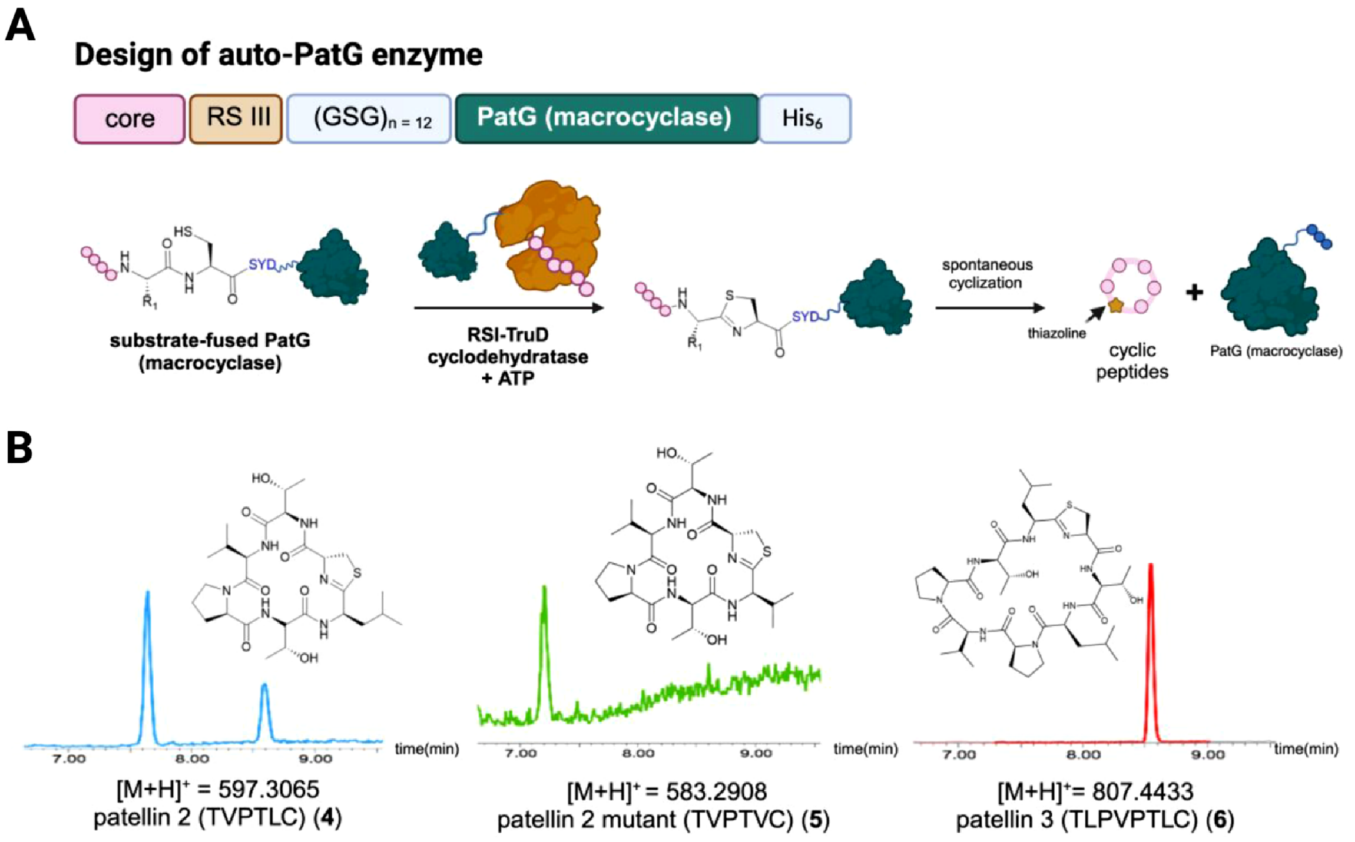


Figure 3. Utilizing auto-PatG and RSI-TruD as a control for peptide cyclization. (A) Protein constructs containing patellin cores and RSIII (SYD) were fused with the PatG macrocyclase enzyme through the optimized GSG flexible linker. Treatment of these constructs with RSI-TruD converts the C-term cysteine residue of the core into thiazoline ring, cascading into peptide cyclization. (B) LC−MS analysis of the reaction mixture of designed protein constructs showing EIC peaks of cyclized products.

https://doi.org/10.1021/acssynbio.3c00645
ACS Synth. Biol. 2024, 13, 394−401

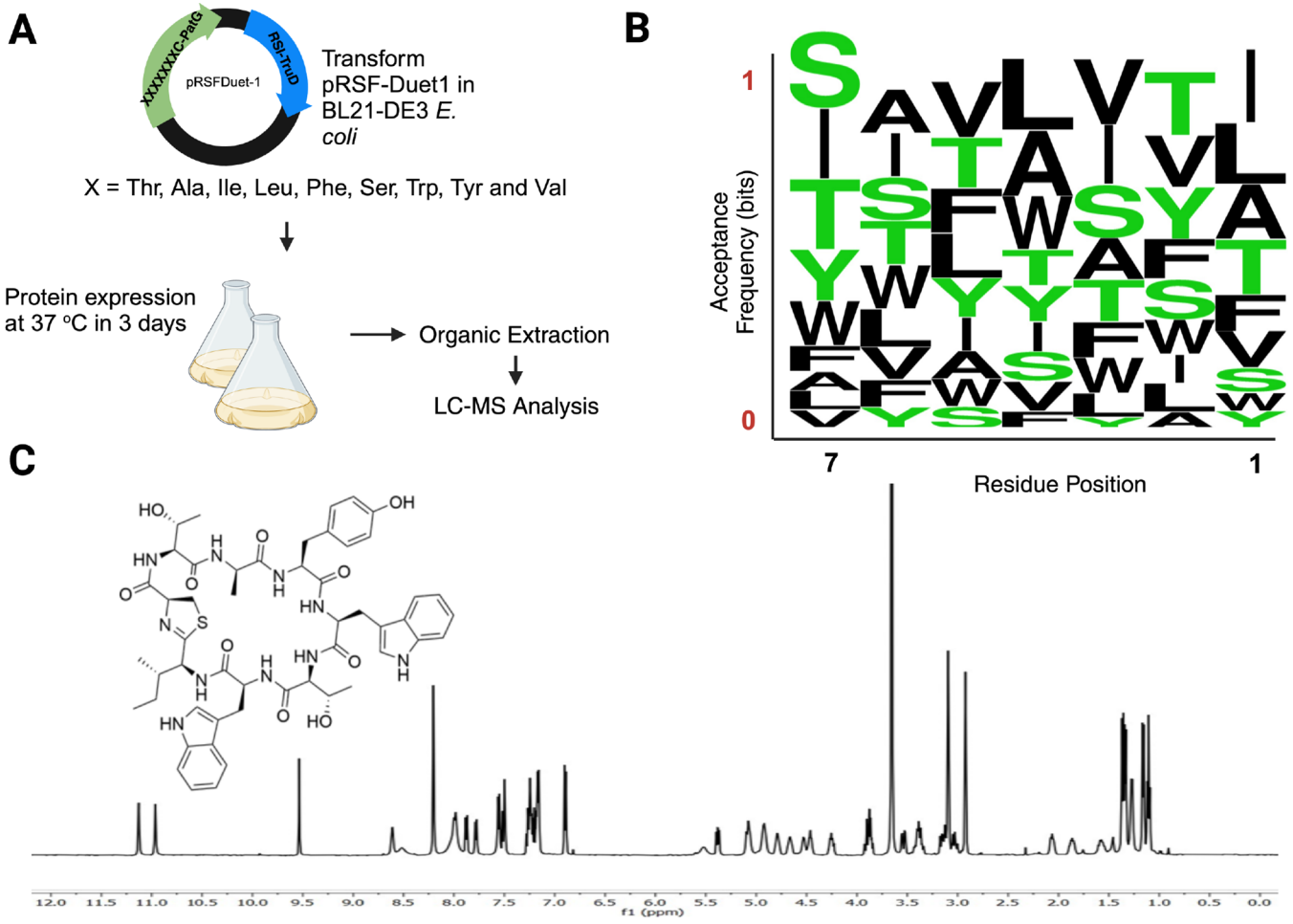


Figure 4. Utilizing auto-PatG and RSI-TruD in making an octapeptide library. (A) Core sequences of substrate-fused PatG were randomized from the P1−7 position, allowing codons specific to certain amino acids. Coexpression with RSI-TruD enables the conversion of cysteine into a thiazoline ring, which permits peptide cyclization. Mutant plasmids were transformed in BL21(DE3) *E. coli* cells and grown for 3 days at 37 °C. Cells were harvested and treated with acetone, which is then used for LC−MS analysis. (B) Acceptance rate of selected amino acids in positions P1−P7 for peptide cyclization, generated by WebLogo.[31] (C) ^{1}H NMR of cyclic TAYWTWIC (**7**) in DMF- d_7 (500 MHz) at 25 °C.

We aimed to efficiently produce the cyclic products in vivo without any additional PatA or related protein. Leader sequences were removed through mutagenesis, leaving a free N-terminus that was primed for immediate circularization. This PagG protein construct, dubbed auto-PagG (Figures S6 and S7), was expressed in *E. coli*, and products were extracted from the resulting cell pellet using acetone. Abundant cyclic peptides were observed by UPLC−MS (Figure 2). Interestingly, while the desired product cyclo[QAYLGIPLP] (**1**) was present in the mixture, by far the major product consisted of the cyclic peptide containing additional Met, cyclo-[MQAYLGIPLP] (**2**), and the derivative with a spontaneously oxidized Met sulfoxide (**3**) (Figures S8 and S9). The incorporation of Met in the final product suggests a competing reaction between the methionyl endopeptidase and PagG macrocyclization, indicating unexpectedly rapid cyclization kinetics. This increased rate is analogous to what has been observed in some engineered protein kinases, in which phosphorylation is dramatically enhanced when the substrate is covalently linked to the enzyme, and the enhancement strongly depends on the linker length.[27]

Applying RSI-TruD as an Orthogonal Control for Autocyclization. A possible way to prevent unwanted Met incorporation would be to include PatA and the PatA RSII sequence in vivo. A disadvantage is that PatA releases the linear N-terminal sequence, which is potentially a complicating factor. This issue led us to apply the same methodology to our previously described PatG macrocyclase domain as an alternative.[18] PatG native substrates contain Cys (instead of Pro) at the C-terminus; Cys is not a substrate for circularization, but instead, it must be heterocyclized to thiazoline by the action of D enzymes. Among these, TruD has been engineered to be constitutively active by fusing the enzyme N-terminus with a leader peptide, producing RSI-TruD.[21,28,29] This would enable orthogonal control of enzyme action in which a macrocycle would be observed only in the presence of active RSI-TruD.

Therefore, by analogy to PagG, we synthesized PatG variants containing native PatG core peptides, the PatG RSIII, and the GSG_{12} sequence linked to the PatG protease domain (C-terminal His tag) (Figure 3a). This construct was dubbed "auto-PatG". Core peptide sequences used included the precursors to the natural products patellin 2 (TVPTLC) (**4**) and patellin 3 (TLPVPTLC) (**6**), as well as a mutant of the patellin 2 sequence (TVPTVC) (**5**). The RSIII sequence was truncated to SYD, in line with recent reports of a minimum sequence requirement. The three auto-PatG protein variants were expressed in *E. coli* and purified (Figure S10). The

https://doi.org/10.1021/acssynbio.3c00645
ACS Synth. Biol. 2024, 13, 394−401

purified RSI-TruD enzyme was added, and the reaction mixtures were analyzed by UPLC−MS. No reaction products were observed unless RSI-TruD was included, demonstrating the requirement for thiazoline synthesis prior to cyclization (Figure 3a). Efficient cyclic peptide formation was observed when RSI-TruD and ATP were included in the reaction mixture (Figure 3b) (Figures S11 and S12). Following these in vitro experiments, the same experiments were performed in *E. coli*, in which the auto-PatG and RSI-TruD constructs were simultaneously expressed from a single plasmid. The cell pellets were isolated, and their acetone extracts contained the expected cyclic peptides (Figure S13).

Interestingly, the patellin 2 core peptide led to two products, one of which was twice the expected mass of the TVPTLC core (Figure S11). This dimer was observed in both in vitro and in vivo expression experiments. The presence of the dimeric patellin 2 derivative was highly sequence-dependent since it was not observed in a Leu-Val mutant TVPTVC, nor was it observed in the octapeptide patellin 3 product (Figure 3b). In addition, dimers were not observed in larger libraries (see below). To investigate how these dimers are formed, the molar ratio of auto-PatG and RSI-TruD varied from 80:1 to 1:1 (Figure S14). Noticeably, as both enzymes reach equal concentration, dimer product increased indicating that formation could be dependent upon the concentration of activated G protein available for cyclization. This implies that the product was synthesized by cross-reaction between two auto-PatG constructs and not by more complex possibilities involving already-synthesized macrocycles. To test this latter idea, preformed macrocycles were added to enzyme reaction mixtures, but they had no effect on the dimer: monomer ratio. Moreover, since the dimer formation was restricted only to the patellin 2 sequence, we could not test mechanistic hypotheses by mixing enzymes and observing unsymmetrical dimer formation. Further work is required to elucidate the mechanism of dimer formation, but since it was restricted to a single compound, it did not prevent further technology development.

An Efficient Cyclic Peptide Library. The cyclic peptide library was constructed in the duet vector encoding RSI-TruD and auto-PatG (Figure 4). We synthesized an octapeptide library in which the seven N-terminal amino acids in the core peptide were randomized. The eighth amino acid Cys was left intact to direct cyclization. The library was prepared using trimer phosphoramidites encoding Ala, Val, Ile, Leu, Phe, Ser, Trp, Tyr, and Thr; these were selected because, in previous work, other amino acids exhibited greater positional dependence.[11] The resulting library of 9^7 (4.8×10^6) theoretical derivatives was transformed into *E. coli*, and 100 individual colonies were picked randomly and sequenced. One of these anomalously contained Lys and was thus discounted in further experiments. The 99 remaining plasmids had the correct elements in place, achieving a relatively even incorporation of all codons: Ala (67), Val (66), Ile (81), Leu (77), Phe (73), Ser (77), Trp (84), Tyr (64), and Thr (69).

The colonies were used in expression experiments at a 25 mL scale. After 3 days of incubation, cultures were pelleted, extracted with acetone, and enriched by passage over a solid phase extraction resin. Analysis of the extract by UPLC−MS revealed that of the 99 mutants considered, 55 produced the desired cyclic peptide. Five of these were duplicated sequences; therefore, 51 out of 94 peptide sequences were successfully turned over into products. This analysis was robust because every chromatogram served as a negative control for every other chromatogram analyzed. Some extracts contained two distinct, isobaric peaks, suggesting possible spontaneous epimerization of the C−H protons adjacent to the thiazoline double bond, as previously observed.[30]

Every amino acid appeared multiple times in each of the seven mutated positions. In comparison to the natural product sequence TLPVPTLC (patellin 3), highly different sequences such as YSVFWAVC, TWASIIIC, SAYTLSLC, TTTLIFVC, and VVLSSYIC were efficiently processed (Tables S4 and S5). Amino acids were not evenly accepted, with overall acceptance rates between 45 and 76% (Figure S15). Thr is the preferred amino acid (76%). In contrast, aromatic amino acids are notably less accepted, with specific positions less so. For example, Trp has an average acceptance rate of 45%, with the lowest at P1 at 27% (Figure S15). Similarly, Phe has an average of 49% but the lowest acceptance rate at P7 at 19%. Examination of the 43 peptide sequences that failed to cyclize produced a logo plot similar to the trend observed in cyclized sequences (Figure S16), suggesting that while aromatic amino acids are tolerated, their presence, specifically at the ends of the core sequences (P1 and P7), reduces cyclization efficiency.

Based upon MS areas under the curve (AUCs), the apparent yield of cyclic peptides varies depending on the core sequences (Tables S4 and S5). AUC depends upon many factors that are both compound and context dependent and therefore is at best semiquantitative. Nonetheless, this analysis revealed that many cyclic peptides are present and noticeably abundant in the 25 mL cultures. To confirm the abundance and identity of cyclic peptides, we selected a plasmid encoding a sequence TAYWTWIC whose product would be readily isolatable due to aromatic residues that would make the product obvious even as a minor component in rich media. The proteins were expressed at 12 L scale for 3 days. After extraction, the predicted peptide was clearly observed in the total ion chromatogram (not just in the mass-filtered chromatogram), indicating abundant production (Figure S17). After purification, cyclo[TAYWTWIC] (**7**) was obtained (4.9 mg), and its structure was confirmed using full spectroscopic analysis (Figure S19). Using this peptide as a standard, we determined that the crude extract contained 17 mg/L of peptide before purification (Figure S27).

In the context of the peptide library, the AUC of **7** was compared with AUCs for other detected cyclic peptides (Figure S29; Table S6). About 30% of compounds had roughly similar AUCs in comparison to **7** (average AUC ∼3.5× that of **7**), while 70% had lower AUCs in comparison to **7** (average AUC ∼0.1× that of **7**). No sequence trends stood out strongly in comparing peptides with different AUCs. The yield of **7**, which could be precisely quantified with the purified standard, was lower in the library method (estimated prepurification amount 2 mg/L, in comparison to 17 mg/L in the scaled up expression), indicating that modification or optimization of conditions might be required for robust library synthesis.

CONCLUSION

Here, we report the development of a simple cyanobactin-based method to synthesize cyclic peptides and cyclic peptide libraries in vivo and in vitro. Despite its simplicity, the method relies on 18 years of research to understand the biochemistry of the cyanobactin pathway enzymes. The speed and efficiency of this approach, as represented by the incorporation of methionine in a cyclic product, strongly support our

https://doi.org/10.1021/acssynbio.3c00645
ACS Synth. Biol. 2024, 13, 394−401

hypotheses about factors limiting cyclization in previous approaches. Notably, both the amount of product purified and the percent of peptides successfully produced in vivo greatly exceeded our previous capabilities. The method is also relatively straightforward, involving only a single plasmid and no special conditions for fermentation, in contrast to previously reported methods.[20] Finally, the process is compatible with other post-translational modifications, enabling combinatorial RiPP biosynthesis using multiple enzymes.

There are several other genetic methods to produce cyclic peptides and cyclic peptide libraries, dominated by intein-mediated cyclization and RiPP-based cyclization as found in broad-substrate pathways such as some of those to lanthipeptides, thiopeptides, cyclotides, and others.[32−34] Previous studies used engineering approaches to improve the yield and scope of RiPP modifications. In these cases, RiPP enzymes from cyanobactins, microviridins, and lanthipeptides were covalently attached to their leader peptides, allowing efficient leaderless processing.[29,35,36] Instead of this approach, we redesigned the G enzyme to directly link to its substrate, which then independently released cyclic peptides. By doing so, universal issues, such as substrate solubility, concentration, and proteolytic stability, are minimized. We expect this approach to be applied to other ribosomal pathways since many operate under the same biosynthetic logic, allowing the RiPP enzymes to be pliably engineered according to biotechnological needs.

Compared to other RiPP cyclases that usually make multiple products (both linear and cyclic) under the artificial conditions used to create libraries, PatG and TruG are thought to produce almost solely circular peptides.[18,37] This work demonstrates that G family enzymes are indeed exceptionally flexible in substrate tolerance, while simultaneously exhibiting high fidelity for the RSIII requirement and for synthesizing only one circular product. However, the enzyme itself certainly has preferences for specific core peptides in a natural setting. As evidence of this, perhaps the only reasonable way a dimer of patellin 2 could form is if the patellin 2 core peptide itself has an especially stable interaction with the enzyme active site.

Despite this advance, the method has several limitations. Chief among these, it would be desirable to create a system that does not require enzymatic modification before cyclization in *E. coli* and yet still avoids undesired Met incorporation. Obviously, this can be done by modulating the catalytic rate of the G enzymes. In addition, each G enzyme has inherent limitations in substrate scope; therefore, only a subset of sequence space will be accessible. This limitation is clearly discernible in this study, in which despite using a preferred subset of amino acids, only a 54% success rate was achieved. Previously, we proposed that this limitation could be addressed by using multiple cyclases with different substrate selectivities.[17] The same principle used here, in which the substrate is covalently tethered to the enzyme, will be applicable in most foreseeable cases.

This work was inspired by natural autocatalytic enzymes, such as the BURP cyclases, which connect aromatic amino acids in plant peptides, and the *N*-methyltransferases from omphalotin biosynthesis.[38−41] In addition, while we were completing this work, an engineered autocatalytic cyclase was reported that uses sortase as the cyclization enzyme.[42] Sortase has several advantages in comparison with auto-G enzymes. First, it is capable of cyclizing many larger proteins in addition to shorter oligopeptides.[43,44] While G enzymes might be capable of this activity, it has not been experimentally established. Second, the substrate scope for these enzymes is relatively wide, although the substrate scope was not thoroughly established in the autocyclase, and the experiments were largely in vitro. However, sortase has several limitations compared with G enzymes that are crucial in small molecule cyclic peptide design. Foremost among the limitations is that sortase incorporates a five amino acid recognition sequence, LPXTG, into its products.[43,44] By contrast, the G enzyme recognition sequences are external to the products, with the exception of an essential heterocycle. Thus, G enzymes are more suitable for synthesizing 6−8 amino acid, drug-like circular peptide libraries, since 5−7 of those amino acids can be randomized instead of just 2−4 in sortase products.

Another limitation was that sortase required a low concentration of autocatalyst because dimer formation was universal in sortase products, whereas only the best native substrate (patellin 2) was dimerized in the auto-G constructs, and dimerization was otherwise not observed.[42] We propose that dimerization is facilitated when, as found in sortase, a part of the substrate is also a recognition sequence for cyclization. In the case of sortase, this LPXTG recognition sequence is found in the product. However, in the case of PatG, the recognition sequence is external to the product. The patellin 2 sequence itself may have evolved higher affinity for this pocket and, therefore, forms dimers more efficiently than other compounds. As a result, the products of auto-G enzymes are also cleaner than those expected from autosortase enzymes.

Overall, our engineered enzymes remove several practical limitations, enabling compatibility with numerous synthetic biology approaches to designed organisms or in vitro synthetic methods. Here, we demonstrate several of these applications by showing the advantages of the system in vivo and in vitro, such as the generation of high-quality cyclic peptide libraries that are compatible with cell-based functional screens. We also demonstrated that the process is robust enough to produce enough material needed for applications such as pharmacological evaluation. Thus, we present a practical approach that not only robustly and predictably produces compounds and libraries but also does so at a practical scale.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.3c00645.

Supporting materials and methods, theoretical and observed ion adducts of cyclic peptides from both in vivo and in vitro methods, and NMR spectra of the isolated product (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Eric W. Schmidt − *Department of Medicinal Chemistry, College of Pharmacy, University of Utah, Salt Lake City, Utah 84112, United States*; Email: ews1@utah.edu

Authors

Noel Lacerna, II − *Department of Medicinal Chemistry, College of Pharmacy, University of Utah, Salt Lake City, Utah 84112, United States*

https://doi.org/10.1021/acssynbio.3c00645
ACS Synth. Biol. 2024, 13, 394−401

Ying Cong − *Department of Medicinal Chemistry, College of Pharmacy, University of Utah, Salt Lake City, Utah 84112, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.3c00645

Author Contributions

N.L and Y.C. performed cloning, protein expression and purification, enzyme reaction and analysis, compound extraction and isolation, structural analysis, and manuscript preparation. E.W.S. contributed to project oversight and manuscript preparation.

Funding

This work was funded by NIH R35GM148283.

Notes

The authors declare the following competing financial interest(s): E.W. Schmidt is a co-owner of Synthetic Biodesign, Inc., which produces in vivo libraries using pat/tru pathway biosynthetic enzymes.

■ ACKNOWLEDGMENTS

We thank Drs. Jack Skalicky and Paul Scesa for their assistance with NMR spectroscopy.

■ REFERENCES

(1) Zhang, H.; Chen, S. Cyclic Peptide Drugs Approved in the Last Two Decades(2001−2021). *RSC Chem. Biol.* **2021**, *3*, 18−31.
(2) Costa, L.; Sousa, E.; Fernandes, C. Cyclic Peptides in Pipeline: What Future for These Great Molecules? *Pharmaceuticals* **2023**, *16*, 996.
(3) Young, T.-S.; Young, D.-D.; Ahmad, I.; Louis, J.-M.; Benkovic, S.-J.; Schultz, P.-G. Evolution of Cyclic Peptide Protease Inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 11052−11056.
(4) Horswill, A.-R.; Savinov, S.-N.; Benkovic, S.-J. A Systematic Method for Identifying Small-Molecule Modulators of Protein−Protein Interactions. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15591−15596.
(5) Tavassoli, A. SICLOPPS Cyclic Peptide Libraries in Drug Discovery. *Curr. Opin. Chem. Biol.* **2017**, *38*, 30−35.
(6) Scott, C.-P.; Abel-Santos, E.; Wall, M.; Wahnon, D.-C.; Benkovic, S.-J. Production of Cyclic Peptides and Proteins in Vivo. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 13638−13643.
(7) Kobayashi, M.; Fujita, K.; Matsuda, K.; Wakimoto, T. Streamlined Chemoenzymatic Synthesis of Cyclic Peptides by Non-Ribosomal Peptide Cyclases. *J. Am. Chem. Soc.* **2023**, *145*, 3270−3275.
(8) Kohli, R.-M.; Walsh, C.-T.; Burkart, M.-D. Biomimetic Synthesis and Optimization of Cyclic Peptide Antibiotics. *Nature* **2002**, *418*, 658−661.
(9) Hipolito, C. J.; Suga, H. Ribosomal Production and in Vitro Selection of Natural Product-like Peptidomimetics: The FIT and RaPID Systems. *Curr. Opin. Chem. Biol.* **2012**, *16*, 196−203.
(10) Tianero, M.-B.-D.; Donia, M.-S.; Young, T.-S.; Schultz, P.-G.; Schmidt, E.-W. Ribosomal Route to Small-Molecule Diversity. *J. Am. Chem. Soc.* **2012**, *134*, 418−425.
(11) Ruffner, D.-E.; Schmidt, E.-W.; Heemstra, J.-R. Assessing the Combinatorial Potential of the RiPP Cyanobactin Tru Pathway. *ACS Synth. Biol.* **2015**, *4*, 482−492.
(12) Donia, M.-S.; Hathaway, B.-J.; Sudek, S.; Haygood, M.-G.; Rosovitz, M.-J.; Ravel, J.; Schmidt, E.-W. Natural Combinatorial Peptide Libraries in Cyanobacterial Symbionts of Marine Ascidians. *Nat. Chem. Biol.* **2006**, *2*, 729−735.
(13) Pan, S.-J.; Link, A.-J. Sequence Diversity in the Lasso Peptide Framework: Discovery of Functional Microcin J25 Variants with Multiple Amino Acid Substitutions. *J. Am. Chem. Soc.* **2011**, *133*, 5016−5023.
(14) Young, T.-S.; Dorrestein, P.-C.; Walsh, C.-T. Codon Randomization for Rapid Exploration of Chemical Space in Thiopeptide Antibiotic Variants. *Chem. Biol.* **2012**, *19*, 1600−1610.
(15) Agarwal, V.; Pierce, E.; McIntosh, J.; Schmidt, E.-W.; Nair, S.-K. Structures of Cyanobactin Maturation Enzymes Define a Family of Transamidating Proteases. *Chem. Biol.* **2012**, *19*, 1411−1422.
(16) Lee, J.; McIntosh, J.; Hathaway, B.-J.; Schmidt, E.-W. Using Marine Natural Products to Discover a Protease That Catalyzes Peptide Macrocyclization of Diverse Substrates. *J. Am. Chem. Soc.* **2009**, *131*, 2122−2124.
(17) Sarkar, S.; Gu, W.; Schmidt, E.-W. Expanding the Chemical Space of Synthetic Cyclic Peptides Using a Promiscuous Macrocyclase from Prenylagaramide Biosynthesis. *ACS Catal.* **2020**, *10*, 7146−7153.
(18) McIntosh, J.-A.; Robertson, C.-R.; Agarwal, V.; Nair, S.-K.; Bulaj, G.-W.; Schmidt, E.-W. Circular Logic: Nonribosomal Peptide-like Macrocyclization with a Ribosomal Peptide Catalyst. *J. Am. Chem. Soc.* **2010**, *132*, 15499−15501.
(19) Gu, W.; Dong, S.-H.; Sarkar, S.; Nair, S.-K.; Schmidt, E.-W. The Biochemistry and Structural Biology of Cyanobactin Pathways: Enabling Combinatorial Biosynthesis. *Methods in Enzymology* **2018**, *604*, 113−163.
(20) Tianero, M.-D.; Pierce, E.; Raghuraman, S.; Sardar, D.; McIntosh, J.-A.; Heemstra, J.-R.; Schonrock, Z.; Covington, B.-C.; Maschek, J.-A.; Cox, J.-E.; Bachmann, B.-O.; Olivera, B.-M.; Ruffner, D.-E.; Schmidt, E.-W. Metabolic Model for Diversity-Generating Biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 1772−1777.
(21) Donia, M.-S.; Schmidt, E.-W. Linking Chemistry and Genetics in the Growing Cyanobactin Natural Products Family. *Chem. Biol.* **2011**, *18*, 508−519.
(22) Sardar, D.; Schmidt, E.-W. Combinatorial Biosynthesis of RiPPs: Docking with Marine Life. *Curr. Opin. Chem. Biol.* **2016**, *31*, 15−21.
(23) Murakami, M.; Itou, Y.; Ishida, K.; Shin, H.-J. Prenylagaramides A and B, New Cyclic Peptides from Two Strains of Oscillatoria Agardhii. *J. Nat. Prod* **1999**, *62*, 752−755.
(24) Chen, X.; Zaro, J.-L.; Shen, W.-C. Fusion Protein Linkers: Property, Design and Functionality. *Adv. Drug Delivery Rev.* **2013**, *65*, 1357−1369.
(25) Reddy Chichili, V.-P.; Kumar, V.; Sivaraman, J. Linkers in the Structural Biology of Protein−Protein Interactions. *Protein Sci.* **2013**, *22*, 153−167.
(26) Koehnke, J.; Bent, A.; Houssen, W.-E.; Zollman, D.; Morawitz, F.; Shirran, S.; Vendome, J.; Nneoyiegbe, A.-F.; Trembleau, L.; Botting, C.-H.; Smith, M.-C.-M.; Jaspars, M.; Naismith, J.-H. The Mechanism of Patellamide Macrocyclization Revealed by the Characterization of the PatG Macrocyclase Domain. *Nat. Struct. Mol. Biol.* **2012**, *19*, 767−772.
(27) Dyla, M.; Kjaergaard, M. Intrinsically Disordered Linkers Control Tethered Kinases via Effective Concentration. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 21413−21419.
(28) Sardar, D.; Hao, Y.; Lin, Z.; Morita, M.; Nair, S.-K.; Schmidt, E.-W. Enzymatic N- and C-Protection in Cyanobactin RiPP Natural Products. *J. Am. Chem. Soc.* **2017**, *139*, 2884−2887.
(29) Koehnke, J.; Mann, G.; Bent, A.-F.; Ludewig, H.; Shirran, S.; Botting, C.; Lebl, T.; Houssen, W.-E.; Jaspars, M.; Naismith, J.-H. Structural Analysis of Leader Peptide Binding Enables Leader-Free Cyanobactin Processing. *Nat. Chem. Biol.* **2015**, *11*, 558−563.
(30) Donia, M.-S.; Ravel, J.; Schmidt, E.-W. A Global Assembly Line for Cyanobactins. *Nat. Chem. Biol.* **2008**, *4*, 341−343.
(31) Crooks, G.-E.; Hon, G.; Chandonia, J.-M.; Brenner, S.-E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188−1190.
(32) Knerr, P.-J.; Van Der Donk, W.-A. Discovery, Biosynthesis, and Engineering of Lantipeptides. *Annu. Rev. Biochem.* **2012**, *81*, 479−505.
(33) Chan, D.-C.-K.; Burrows, L.-L. Thiopeptides: Antibiotics with Unique Chemical Structures and Diverse Biological Activities. *J. Antibiot. (Tokyo)* **2021**, *74*, 161−175.

(34) De Veer, S.-J.; Kan, M.-W.; Craik, D.-J. Cyclotides: From Structure to Function. *Chem. Rev.* **2019**, *119*, 12375−12421.
(35) Reyna-González, E.; Schmid, B.; Petras, D.; Süssmuth, R.-D.; Dittmann, E. Leader Peptide-Free In Vitro Reconstitution of Microviridin Biosynthesis Enables Design of Synthetic Protease-Targeted Libraries. *Angew. Chem.* **2016**, *128*, 9544−9547.
(36) Oman, T.-J.; Knerr, P.-J.; Bindman, N.-A.; Velásquez, J.-E.; Van Der Donk, W.-A. An Engineered Lantibiotic Synthetase That Does Not Require a Leader Peptide on Its Substrate. *J. Am. Chem. Soc.* **2012**, *134*, 6952−6955.
(37) Nguyen, N.-A.; Cong, Y.; Hurrell, R.-C.; Arias, N.; Garg, N.; Puri, A.-W.; Schmidt, E.-W.; Agarwal, V. A Silent Biosynthetic Gene Cluster from a Methanotrophic Bacterium Potentiates Discovery of a Substrate Promiscuous Proteusin Cyclodehydratase. *ACS Chem. Biol.* **2022**, *17*, 1577−1585.
(38) Chigumba, D.-N.; Mydy, L.-S.; De Waal, F.; Li, W.; Shafiq, K.; Wotring, J.-W.; Mohamed, O.-G.; Mladenovic, T.; Tripathi, A.; Sexton, J.-Z.; Kautsar, S.; Medema, M.-H.; Kersten, R.-D. Discovery and Biosynthesis of Cyclic Plant Peptides via Autocatalytic Cyclases. *Nat. Chem. Biol.* **2022**, *18*, 18−28.
(39) Kersten, R.-D.; Weng, J.-K. Gene-Guided Discovery and Engineering of Branched Cyclic Peptides in Plants. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E10961−E10969.
(40) van der Velden, N.-S.; Kälin, N.; Helf, M.-J.; Piel, J.; Freeman, M.-F.; Künzler, M. Autocatalytic backbone N-methylation in a family of ribosomal peptide natural products. *Nat. Chem. Biol.* **2017**, *13*, 833−835.
(41) Ramm, S.; Krawczyk, B.; Mühlenweg, A.; Poch, A.; Mösker, E.; Süssmuth, R.-D. A Self-Sacrificing N -Methyltransferase Is the Precursor of the Fungal Natural Product Omphalotin. *Angew. Chem. Int. Ed* **2017**, *56*, 9994−9997.
(42) Jia, X.; Chin, Y.-K.-Y.; Zhang, A.-H.; Crawford, T.; Zhu, Y.; Fletcher, N.-L.; Zhou, Z.; Hamilton, B.-R.; Stroet, M.; Thurecht, K.-J.; Mobli, M. Self-Cyclisation as a General and Efficient Platform for Peptide and Protein Macrocyclisation. *Commun. Chem.* **2023**, *6*, 48.
(43) Mazmanian, S.-K.; Liu, G.; Ton-That, H.; Schneewind, O. Staphylococcus Aureus Sortase, an Enzyme That Anchors Surface Proteins to the Cell Wall. *Science* **1999**, *285*, 760−763.
(44) Mao, H.; Hart, S.-A.; Schink, A.; Pollok, B.-A. Sortase-Mediated Protein Ligation: A New Method for Protein Engineering. *J. Am. Chem. Soc.* **2004**, *126*, 2670−2671.

https://doi.org/10.1021/acssynbio.3c00645
ACS Synth. Biol. 2024, 13, 394−401