Letter

Transformation-Guided Genome Mining Provides Access to Brominated Lanthipeptides

Nirmal Saha,[▽] F. N. U. Vidya,[▽] Youran Luo, Wilfred A. van der Donk, and Vinayak Agarwal*

Cite This: *Org. Lett.* 2025, 27, 984−988

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Natural product biosynthesis is nature's tinkering ground for developing new enzymes that can achieve chemical transformations that are outside the purview of traditional chemical catalysis. Herein we describe a genome mining approach that leads to the discovery of a halogenase that regioselectively brominates a tryptophan side chain indole for a macrocyclic peptide substrate, enabling downstream chemical arylation by Suzuki−Miyaura coupling. The halogenase was found to prefer a macrocyclic peptide substrate over a linear peptide. The brominase presents a starting point for biocatalytic access to macrocyclic peptides bearing a chemically versatile aryl-bromide reactive handle
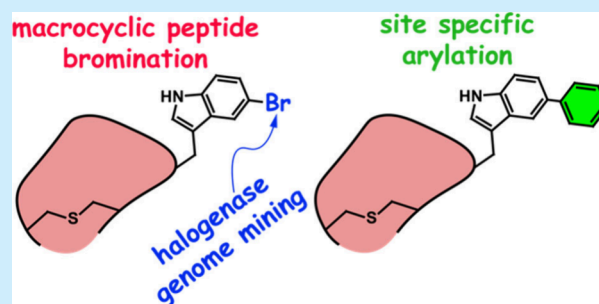
There is a resurgence of intense pharmacological interest in macrocyclic peptides; they are the Goldilocks molecules bridging selective but difficult to maneuver biologics and small molecules that are easier to deliver but struggle in engaging large biomolecular surfaces. Macrocyclization endows peptides with membrane permeability and stability against proteolysis.[1−3] Among the macrocyclization modalities, intramolecular thioether bonds are well represented among bioactive peptides.[4] Lanthipeptides are ribosomally derived and post-translationally modified peptides (RiPPs) that are macro-cyclized by thioether bond formation between a Cys thiol and a Cβ carbon atom of dehydrated Ser and Thr side chains. Chemoenzymatic derivatization of lanthipeptides is the key to fully exploiting their pharmacological potential.

Akin to the prominence of brominated intermediates in organic synthesis, enzymatic bromination of ribosomally derived peptides and proteins allows for their chemical derivatization by transition-metal-assisted C−C bond-forming reactions.[5−7] Enzymatic bromination benefits from the regiospecific delivery of the bromine atom onto the peptide. With a view toward downstream chemical derivatization, bromination is particularly desirable. While not suffering from the instability of aryl iodides, the increased reactivity of aryl bromides as compared to that of aryl chlorides allows for C−C bond-forming reactions to proceed in aqueous solvents under mild reaction conditions. Considering these observations, in this study, we adopted a genome mining strategy directed toward the discovery of a biosynthetic gene cluster (BGC) that could furnish macrocyclic lanthipeptides bearing an aryl bromide reactive handle. We further explored the transition-metal-assisted chemical derivatization of macrocyclic RiPPs.

We have previously described the activity of the flavin-dependent halogenase SrpI that brominates a tryptophan indole of linear peptides.[8] Using SrpI as a query sequence, we mined for halogenases in sequence databases. The ensuing hits were organized into a sequence similarity network (SSN) illustrated in Figure 1A.[9] With the motivation to access macrocyclic halogenated peptides, we mined this SSN for BGCs that possess a SrpI-like halogenase encoding gene in the neighborhood of a lanthionine synthetase (LanM) encoding gene. LanMs catalyze macrocyclizing thioether bond formation using the above-mentioned strategy of ATP-dependent Ser/Thr side chain dehydration and the addition of a Cys side chain thiol at the Cβ of the dehydroalanine/dehydrobutyrine (Dha/Dhb) residue.[10] Using this strategy, four BGCs were detected wherein an SrpI-like halogenase was encoded in the vicinity of a LanM (Figure S1). To bias the search for the discovery of brominases that do not suffer from contaminating chlorination activity, we chose the marine cyanobacterium *Moorena producens*-derived RiPP BGC (henceforth abbreviated as the *mpp* BGC) for further experimental evaluation; marine cyanobacterial halogenases have a propensity to be obligate brominases.[11,12]

In addition to the halogenase and LanM encoding genes *mppI* and *mppM*, respectively, the *mpp* BGC possesses genes encoding the RiPP precursor peptide (*mppE*), two genes encoding arginases (*mppA* and *mppB*), a prenyltransferase and a terpene synthase (*mppC* and *mppD*, respectively), and an
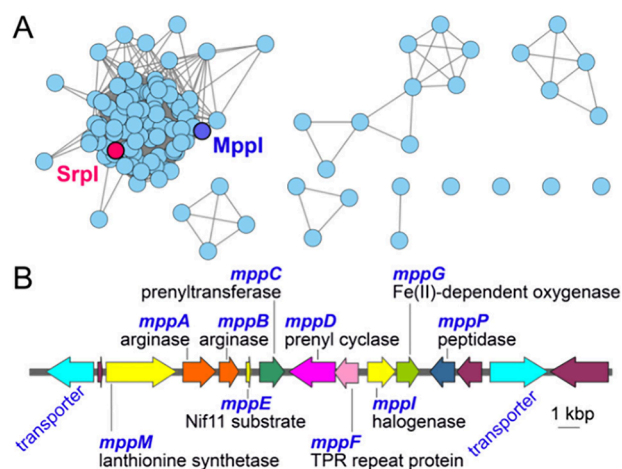
Figure 1. (A) SSN with 124 sequences generated using SrpI as the query demonstrating the clustering of nodes corresponding to SrpI and MppI halogenases. (B) The mpp BGC with predicted activities of encoded enzymes annotated. Genes colored purple are hypothetical genes for which we could not predict activity or function.

oxygenase (*mppG*) (Figure 1B). RiPP precursor peptides comprise an N-terminal leader and a C-terminal core; the leader guides the post-translational modification of the core peptide. Consistent with this architecture, the MppE precursor peptide possess a Nif11-like 77 amino acid long leader appended to a 7 amino acid ACWRWSG core (Table S1).[13]

Eventual proteolytic release of the modified core is postulated to be catalyzed by a peptidase encoded by the *mppP* gene; even though the MppE leader terminates in a Gly-Gly motif, there is no peptidase domain embedded within either of the two transporters encoded in the vicinity of the *mpp* BGC (Figure 1B).[14] The RiPPs encoded by the *mpp* BGC are as yet cryptic; this study is directed toward describing the activity of the MppI halogenase in conjunction with the macrocyclizing lanthionine synthetase MppM.

Coexpression of the precursor peptide encoding gene *mppE* together with the LanM encoding gene *mppM* in *Escherichia coli* led to lanthionine ring formation between the MppE Cys79 and Ser83 side chains (Figure 2, Figures S2−S4). Degradation experiments established the configuration of the lanthionine ring as DL-lanthionine (Figure S5).[15] When *mppE* was coexpressed with *mppM* and *mppI*, we observed the product mass corresponding to the addition of a bromine atom upon the macrocyclic RiPP; no chlorination was observed (Figure 2, Figure S6). Flavin-dependent halogenases bear a conserved catalytic Lys residue in their active sites that is implicated in bridging the site for halide oxidation proximal to the flavin cofactor and the site where the aryl substrate is bound.[16−18] Consistent with this mechanistic hypothesis, the MppI Lys73 → Ala mutation abolished bromination (Figure S7).

The bromine atom could be added by MppI to either of two indole side chains in the MppE cores, namely, Trp80 and Trp82. To determine which of these two Trp residues was
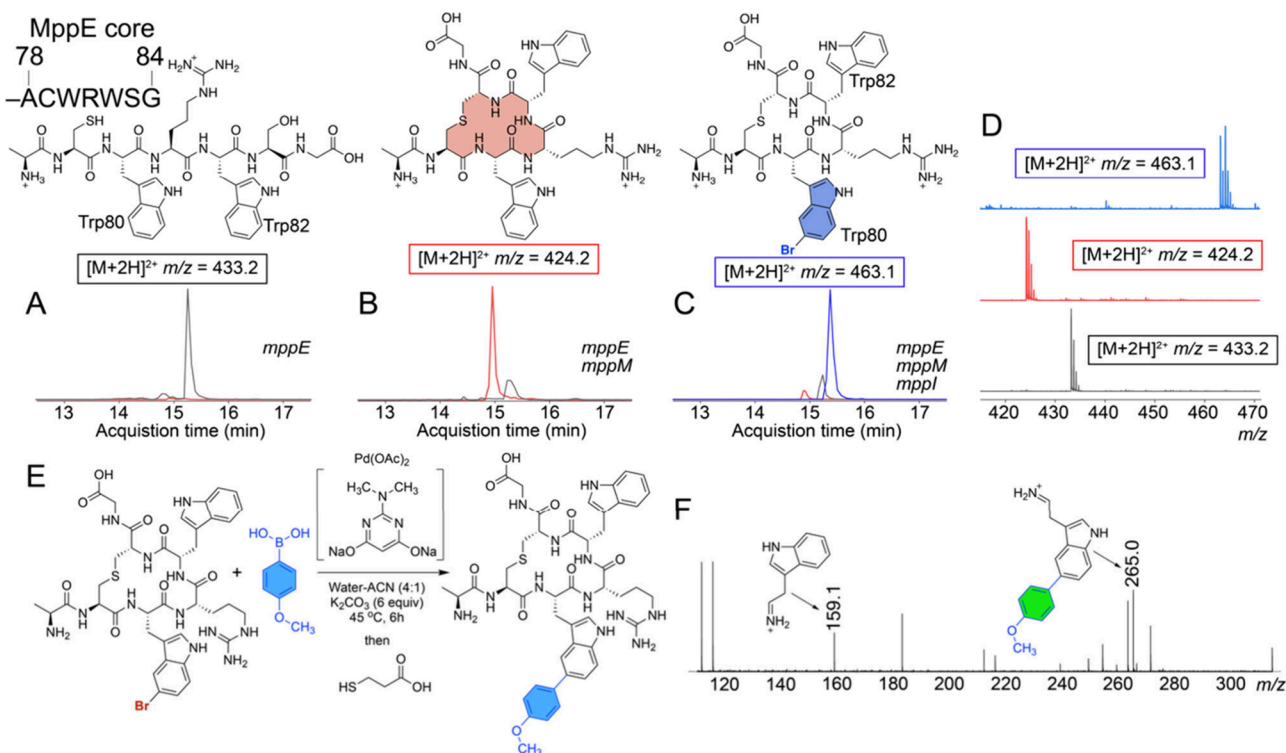

Figure 2. Extracted ion chromatograms (EICs) in black, red, and blue, respectively, demonstrating the presence of linear MppE, macrocyclized MppE, and macrocyclized and brominated MppE core peptides when (A) *mppE* was expressed in *E. coli*, (B) *mppE* was coexpressed with *mppM*, and (C) *mppE* was coexpressed with *mppM* and *mppI*. The EICs were generated by using the [M + 2H]²⁺ ions. The MppE core was excised from the leader using the LahT150 protease.[14] (D) [M + 2H]²⁺ MS¹ spectra for (from bottom to top) the linear, macrocyclized, and macrocyclized and brominated MppE core peptide. (E) Scheme for SMCC of 4-methoxyphenyl boronic acid to the brominated lanthipeptide. The reaction was quenched by the addition of 3-mercaptopropionic acid. (F) MS² spectra with the key fragment ions annotated that demonstrate C−C bond formation to be affected by the Trp side chain indole.

brominated, Trp80 and Trp82 were individually mutated to Phe, and the genes encoding the mutant MppE precursor peptides were coexpressed with *mppM* and *mppI* (Table S1). The Trp80 → Phe mutation greatly reduced the production of a brominated product, while robust brominated product formation was observed for the Trp82 → Phe mutant (Figure S8). These observations allowed us to posit that bromination is affected by the Trp80 side chain (Figure 2). Enzymatic degradation of the brominated product followed by a retention time comparison with authentic synthetic standards demonstrated regiospecific indole-5 bromination (Figure S9). Taken together, these data led to the structural characterization of the brominated lanthipeptide furnished by MppM and MppI (Figure 2).

To exploit the utility of the aryl bromide, we screened for reaction conditions for palladium-assisted Suzuki−Miyaura cross coupling (SMCC) of 4-methoxyphenyl boronic acid to the brominated lanthipeptide furnished by MppM and MppI. As peptides containing tryptophan residues are prone to oxidative degradation, as we had observed previously, we were mindful of the reaction conditions not to necessitate high temperatures (Table S2).[6] Various Pd ligands were evaluated. Of these, the previously described 2-dimethylamino-4,6-dihydroxypyrimidine demonstrated the requisite product formation (Figure 2E).[19,20] Product identity was confirmed by high-resolution mass spectrometry; fragmentation of the $[M + H]^{1+}$ parent ion led to detection of characteristic immonium product ions corresponding to the unmodified Trp82 side chain indole as well as the Trp80 side chain indole, which now bears the methoxyphenyl adduct (Figure 2F). Note that immonium ions for the Arg side chain are usually not observed in this experiment, and the immonium ions for Ala and Gly residues are too small to be detected.[21] Elimination of the catalyst from the reaction abolished the formation of the methoxyphenyl adduct (Figure S10).

Along with the macrocyclic brominated lanthipeptide, coexpression of *mppE* with *mppM* and *mppI* led to the detection of a linear brominated product as well; here, bromination proceeds without macrocyclization (Figure S11). Hence, the order of MppM- and MppI-catalyzed transformations and the substrate preference of MppI (macrocyclic or linear MppE) was unclear. The recombinant MppI enzyme was purified, and bromination activity was reconstituted *in vitro* for the linear MppE and macrocyclic MppE substrates (Figure 3A). The *in vitro* assays involved the flavin reductase RebF and the NADH-generating phosphite dehydrogenase PTDH as reaction partners.[22,23] In a competition experiment, wherein equimolar linear MppE and macrocyclic MppE were provided as substrates to MppI, we detected the preferential bromination of the macrocyclized MppE substrate (Figure 3B, Figures S12 and S13). Furthermore, when substrate conversions were evaluated for these two substrates in separate individual *in vitro* bromination reactions, the conversion for the macrocyclized MppE substrate was greater (Figure 3C, Figures S14 and S15). These data allow us to posit that the physiological substrate for MppI is the macrocyclized MppE peptide and that macrocyclization precedes bromination.

We next evaluated the leader peptide dependence for MppI. RiPP precursor peptides, such as MppE, are divided into a C-terminal core sequence that is post-translationally modified and an N-terminal leader sequence that is not modified but serves to facilitate peptide−enzyme interaction. When the
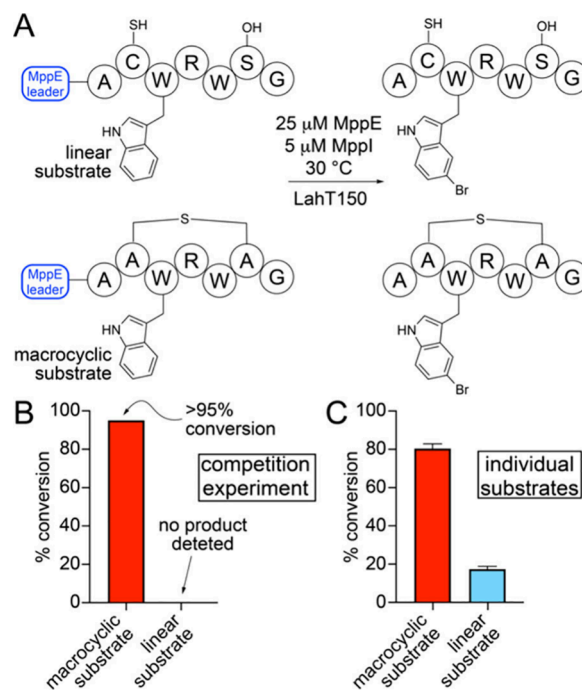

Figure 3. (A) Scheme for the bromination of linear and macrocyclic MppE substrates by MppI, followed by leader peptide excision by LahT150. (B) Macrocyclic and linear substrate conversion when these substrates were provided to MppI in an equimolar amount in a competition experiment. All biochemical experiments were performed in triplicate, and mean and standard deviations from independent measurements are plotted. (C) Conversions when macrocyclic and linear substrates were provided in separate experiments to MppI.

macrocyclized MppE core alone without the MppE leader was provided as a substrate to MppI, no brominated product formation was observed (Figure S16). This observation implies that MppI binds to the MppE leader, which is in contrast to the activity of the flavin-dependent halogenase MibH that chlorinates a macrocyclized lanthipeptide only after the leader is proteolytically removed from the modified core peptide.[24]

The precursor peptides MppE and SrpE both possess long leaders (SrpE is the precursor peptide for SrpI).[8] Much of the long SrpE leader peptide was dispensable for binding to SrpI as long as the two Leu residues comprising of the $L(X)_4L$ motif were preserved (Figure 4A).[25] While SrpE belongs to the nitrile hydratase leader peptide (NHLP) family and MppE belongs to the Nif11-like peptide family, the $L(X)_4L$ motif is conserved in the MppE leader as well.[13] Mutation of the two Leu residues comprised of the $L(X)_4L$ motif in MppE (Leu66 and Leu71) to Ala only had a modest deleterious effect on MppI activity (Figure 4B, Table S1, Figure S17); the corresponding $L(X)_4L \rightarrow A(X)_4A$ mutation in SrpE had abolished SrpI activity.[25]

Next, guided by the sequence alignment between SrpE and MppE, the MppE leader peptide was truncated (Table S1); a similarly truncated SrpE leader that preserved the $L(X)_4L$ motif was a competent substrate for SrpI.[25] Evaluating the truncated MppE substrate peptide at different substrate-to-enzyme stoichiometric ratios revealed that leader truncation led to a loss of MppI activity (Figure 4C). Thus, while MppI and SrpI are both leader-dependent, there is dissonance in the mode of leader engagement between the two enzymes. These biochemical observations are in alignment with AlphaFold3-predicted models for MppI/MppE-leader and SrpI/SrpE-
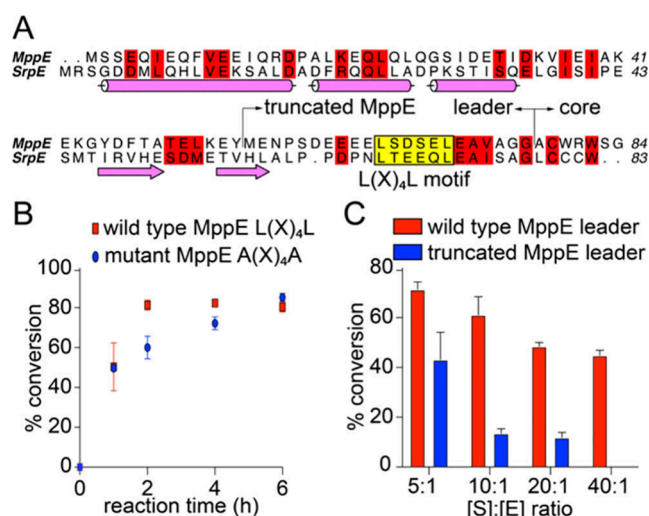
Figure 4. (A) Sequence alignment between MppE and SrpE peptides. The secondary structural elements determined for the SrpE leader are annotated. The conservation of the $L(X)_4L$ motif is highlighted. (B) Time-dependent halogenation of the macrocyclic MppE substrate by MppI bearing the wild-type MppE leader and the when the $L(X)_4L$ motif in MppE leader was mutated to $A(X)_4A$. All biochemical experiments were performed in triplicate, and mean and standard deviations from three independent measurements are plotted. (C) Truncation of the MppE leader (site of truncation denoted in panel A) compromises MppI activity at high $[S]/[E]$ stoichiometric ratios.

leader complexes wherein MppI contacts the entire MppE leader sequence, while SrpI only contacts a small C-terminal section of the SrpE leader peptide (Figure S18).

While the potential of MppI to furnish diverse brominated macrocyclic peptides remains to be evaluated, findings from this study support the assertion that transformation-guided genome mining can lead to the discovery of new biocatalytic modalities. The Trp indole is the site for numerous enzymatic transformations that are critical for the bioactivities of the corresponding natural products.[26] However, unlike bioorthogonal manipulation of Cys and Lys side chains that offer a singular reactive site, regiospecific transformation of indole side chains in peptides and proteins is hard to reach.[27] While strategies to modify the indole-2 position have been developed, at the moment, regiospecific access to the indole-4−7 positions seems be largely restricted to enzymes.[28−31] With the goal of developing bioorthogonal chemistries that can interface with peptides and proteins, RiPP biosynthetic enzymes are a viable starting point.

ASSOCIATED CONTENT

Data Availability Statement

The data underlying this study are available in the published article and its Supporting Information.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.orglett.4c04529.

> Comprehensive description of materials and methods used in this study, synthetic schemes, compound characterization data, and descriptions of enzyme reaction outcomes (PDF)

AUTHOR INFORMATION

Corresponding Author

Vinayak Agarwal − *School of Chemistry and Biochemistry and School of Biological Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332, United States;* orcid.org/0000-0002-2517-589X; Email: vagarwal@gatech.edu

Authors

Nirmal Saha − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*

F. N. U. Vidya − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332, United States*

Youran Luo − *Department of Chemistry, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States;* orcid.org/0000-0002-8033-2472

Wilfred A. van der Donk − *Department of Chemistry, Carl R. Woese Institute for Genomic Biology, and Howard Hughes Medical Institute, University of Illinois at Urbana−Champaign, Urbana, Illinois 61801, United States;* orcid.org/0000-0002-5467-7071

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.orglett.4c04529

Author Contributions

∇N.S. and F.N.U.V. contributed equally.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Ji, X.; Nielsen, A. L.; Heinis, C. Cyclic Peptides for Drug Development. *Angew. Chem., Int. Ed.* **2024**, *63*, No. e202308251.

(2) You, S.; McIntyre, G.; Passioura, T. The Coming of Age of Cyclic Peptide Drugs: An Update on Discovery Technologies. *Expert Opin. Drug Discovery* **2024**, *19*, 961−973.

(3) Brudy, C.; Walz, C.; Spiske, M.; Dreizler, J. K.; Hausch, F. The Missing Link(er): A Roadmap to Macrocyclization in Drug Discovery. *J. Med. Chem.* **2024**, *67*, 14768−14785.

(4) Golosov, A. A.; Flyer, A. N.; Amin, J.; Babu, C.; Gampe, C.; Li, J.; Liu, E.; Nakajima, K.; Nettleton, D.; Patel, T. J.; Reid, P. C.; Yang, L.; Monovich, L. G. Design of Thioether Cyclic Peptide Scaffolds with Passive Permeability and Oral Exposure. *J. Med. Chem.* **2021**, *64*, 2622−2633.

(5) Nguyen, N. A.; Agarwal, V. A Leader-Guided Substrate Tolerant RiPP Brominase Allows Suzuki−Miyaura Cross-Coupling Reactions for Peptides and Proteins. *Biochemistry.* **2023**, *62*, 1838−1843.

(6) Saha, N.; Vidya, F. N. U.; Xie, R.; Agarwal, V. Halogenase-Assisted Alkyne/Aryl Bromide Sonogashira Coupling for Ribosomally Synthesized Peptides. *J. Am. Chem. Soc.* **2024**, *146*, 30009−30013.

(7) Montua, N.; Thye, P.; Hartwig, P.; Kühle, M.; Sewald, N. Enzymatic Peptide and Protein Bromination: The BromoTrp Tag. *Angew. Chem., Int. Ed.* **2024**, *63*, No. e202314961.

(8) Nguyen, N. A.; Lin, Z.; Mohanty, I.; Garg, N.; Schmidt, E. W.; Agarwal, V. An Obligate Peptidyl Brominase Underlies the Discovery of Highly Distributed Biosynthetic Gene Clusters in Marine Sponge Microbiomes. *J. Am. Chem. Soc.* **2021**, *143*, 10221−10231.

(9) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and

Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry.* **2019**, *58*, 4169−4182.

(10) Repka, L. M.; Chekan, J. R.; Nair, S. K.; van der Donk, W. A. Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes. *Chem. Rev.* **2017**, *117*, 5457−5520.

(11) Lukowski, A. L.; Hubert, F. M.; Ngo, T.-E.; Avalon, N. E.; Gerwick, W. H.; Moore, B. S. Enzymatic Halogenation of Terminal Alkynes. *J. Am. Chem. Soc.* **2023**, *145*, 18716−18721.

(12) Thapa, H. R.; Agarwal, V. Obligate Brominating Enzymes Underlie Bromoform Production by Marine Cyanobacteria. *J. Phycol.* **2021**, *57*, 1131−1139.

(13) Haft, D. H.; Basu, M. K.; Mitchell, D. A. Expansion of ribosomally produced natural products: a nitrile hydratase- and Nif11-related precursor family. *BMC Biol.* **2010**, *8*, 70.

(14) Bobeica, S. C.; Dong, S.-H.; Huo, L.; Mazo, N.; McLaughlin, M. I.; Jiménez-Osés, G.; Nair, S. K.; van der Donk, W. A. Insights into AMS/PCAT transporters from biochemical and structural character-ization of a double Glycine motif protease. *eLife* **2019**, *8*, No. e42305.

(15) Luo, Y.; Xu, S.; Frerk, A. M.; van der Donk, W. A. Facile Method for Determining Lanthipeptide Stereochemistry. *Anal. Chem.* **2024**, *96*, 1767−1773.

(16) Dong, C.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pee, K. H.; Naismith, J. H. Tryptophan 7-halogenase (PrnA) structure suggests a mechanism for regioselective chlorination. *Science.* **2005**, *309*, 2216−2219.

(17) Yeh, E.; Cole, L. J.; Barr, E. W.; Bollinger, J. M.; Ballou, D. P.; Walsh, C. T. Flavin Redox Chemistry Precedes Substrate Chlorina-tion during the Reaction of the Flavin-Dependent Halogenase RebH. *Biochemistry.* **2006**, *45*, 7904−7912.

(18) Phintha, A.; Prakinee, K.; Jaruwat, A.; Lawan, N.; Visitsatthawong, S.; Kantiwiriyawanitch, C.; Songsungthong, W.; Trisrivirat, D.; Chenprakhon, P.; Mulholland, A.; van Pée, K.-H.; Chitnumsub, P.; Chaiyen, P. Dissecting the low catalytic capability of flavin-dependent halogenases. *J. Biol. Chem.* **2021**, *296*, No. 100068.

(19) Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q. Copper-Free Sonogashira Cross-Coupling for Functionalization of Alkyne-Encoded Proteins in Aqueous Medium and in Bacterial Cells. *J. Am. Chem. Soc.* **2011**, *133*, 15316−15319.

(20) Gao, Z.; Gouverneur, V.; Davis, B. G. Enhanced Aqueous Suzuki−Miyaura Coupling Allows Site-Specific Polypeptide 18F-Labeling. *J. Am. Chem. Soc.* **2013**, *135*, 13612−13615.

(21) Mohanty, I.; Nguyen, N. A.; Moore, S. G.; Biggs, J. S.; Gaul, D. A.; Garg, N.; Agarwal, V. Enzymatic Synthesis Assisted Discovery of Proline-Rich Macrocyclic Peptides in Marine Sponges. *Chembiochem.* **2021**, *22*, 2614−2618.

(22) Yeh, E.; Garneau, S.; Walsh, C. T. Robust in vitro activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3960−3965.

(23) Johannes, T. W.; Woodyer, R. D.; Zhao, H. Efficient Regeneration of NADPH Using an Engineered Phosphite Dehydro-genase. *Biotechnol. Bioeng.* **2007**, *96*, 18−26.

(24) Ortega, M. A.; Cogan, D. P.; Mukherjee, S.; Garg, N.; Li, B.; Thibodeaux, G. N.; Maffioli, S. I.; Donadio, S.; Sosio, M.; Escano, J.; Smith, L.; Nair, S. K.; van der Donk, W. A. Two Flavoenzymes Catalyze the Post-Translational Generation of 5-Chlorotryptophan and 2-Aminovinyl-Cysteine during NAI-107 Biosynthesis. *ACS Chem. Biol.* **2017**, *12*, 548−557.

(25) Nguyen, N. A.; Vidya, F. N. U.; Yennawar, N. H.; Wu, H.; McShan, A. C.; Agarwal, V. Disordered Regions in Proteusin Peptides Guide Post-Translational Modification by a Flavin-Dependent RiPP Brominase. *Nat. Commun.* **2024**, *15*, 1265.

(26) Duan, S.-F.; Song, L.; Guo, H.-Y.; Deng, H.; Huang, X.; Shen, Q.-K.; Quan, Z.-S.; Yin, X.-M. Research Status of Indole-Modified Natural Products. *RSC Med. Chem.* **2023**, *14*, 2535−2563.

(27) Zhang, C.; Vinogradova, E. V.; Spokoyny, A. M.; Buchwald, S. L.; Pentelute, B. L. Arylation Chemistry for Bioconjugation. *Angew. Chem., Int. Ed.* **2019**, *58*, 4810−4839.

(28) Lee, J. C.; Cuthbertson, J. D.; Mitchell, N. J. Chemoselective Late-Stage Functionalization of Peptides via Photocatalytic C2-Alkylation of Tryptophan. *Org. Lett.* **2023**, *25*, 5459−5464.

(29) Kaplaneris, N.; Puet, A.; Kallert, F.; Pöhlmann, J.; Ackermann, L. Late-stage C−H Functionalization of Tryptophan-Containing Peptides with Thianthrenium Salts: Conjugation and Ligation. *Angew. Chem., Int. Ed.* **2023**, *62*, No. e202216661.

(30) Ruiz-Rodríguez, J.; Albericio, F.; Lavilla, R. Postsynthetic Modification of Peptides: Chemoselective C-Arylation of Tryptophan Residues. *Chem.-Eur. J.* **2010**, *16*, 1124−1127.

(31) Reay, A. J.; Williams, T. J.; Fairlamb, I. J. S. Unified Mild Reaction Conditions for C2-Selective Pd-Catalysed Tryptophan Arylation. *Including Tryptophan-Containing Peptides. Org. Biomol. Chem.* **2015**, *13*, 8298−8309.