Ribosomally derived lipopeptides containing distinct fatty acyl moieties

Florian Hubrich[a,1], Nina M. Bösch[a,1], Clara Chepkirui[a], Brandon I. Morinaka[a,b], Michael Rust[a], Muriel Gugger[c], Serina L. Robinson[a,d], Anna L. Vagstad[a,2], and Jörn Piel[a,2]

[a]Institute of Microbiology, Eidgenössische Technische Hochschule (ETH) Zürich, 8093 Zürich, Switzerland; [b]Department of Pharmacy, National University of Singapore, 117543 Singapore, Singapore; [c]Collection of Cyanobacteria, Institut Pasteur, 75724 Paris, France; and [d]Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland

Edited by Chaitan Khosla, Departments of Chemistry and Chemical Engineering, Stanford University, Stanford, CA; received July 17, 2021; accepted November 30, 2021

Lipopeptides represent a large group of microbial natural products that include important antibacterial and antifungal drugs and some of the most-powerful known biosurfactants. The vast majority of lipopeptides comprise cyclic peptide backbones N-terminally equipped with various fatty acyl moieties. The known compounds of this type are biosynthesized by nonribosomal peptide synthetases, giant enzyme complexes that assemble their products in a non–gene-encoded manner. Here, we report the genome-guided discovery of ribosomally derived, fatty-acylated lipopeptides, termed selidamides. Heterologous reconstitution of three pathways, two from cyanobacteria and one from an arctic, ocean-derived alphaproteobacterium, allowed structural characterization of the probable natural products and suggest that selidamides are widespread over various bacterial phyla. The identified representatives feature cyclic peptide moieties and fatty acyl units attached to (hydroxy)ornithine or lysine side chains by maturases of the GCN5-related *N*-acetyltransferase superfamily. In contrast to nonribosomal lipopeptides that are usually produced as congener mixtures, the three selidamides are selectively fatty acylated with C_{10}, C_{12}, or C_{16} fatty acids, respectively. These results highlight the ability of ribosomal pathways to emulate products with diverse, nonribosomal-like features and add to the biocatalytic toolbox for peptide drug improvement and targeted discovery.

RiPPs | natural products | peptides | N-acylation | biosynthesis

Lipopeptides are a widespread class of bioactive natural products in bacteria and fungi (1–3). They consist of a hydrophobic moiety comprising predominantly fatty acyl units of various lengths and functionalization, linked to a hydrophilic peptide portion that is often cyclized and/or contains nonproteinogenic amino acids. This structural composition provides lipopeptides with amphiphilic properties that confer versatile functionalities. Lipopeptides fulfill diverse ecological roles [e.g., as biosurfactants, in biofilm formation, and to access nutrients (4, 5) and as mediators of interspecies interactions that include defense against predators, competition, pathogenicity, and symbiosis (1, 6, 7)]. Relevant examples are the surfactin-type biosurfactants, orfamides and fengycins that induce systemic resistance in plants, and the antifungal echinocandins (1, 3, 8, 9). Lipopeptides have important medical and industrial applications as antimicrobial, antitumor, drug-delivery, and vaccine agents (2, 8, 9), among which daptomycin, polymyxin, and echinocandin are prominent representatives of antibiotics in clinical use.

Two major, fundamentally different biosynthetic systems exist for most-known peptide natural products: 1) nonribosomal peptide synthetases (NRPSs) that form giant multienzyme complexes and can assemble a wide range of amino acids and other building blocks by thiotemplate biosynthesis and 2) ribosomally synthesized and posttranslationally modified peptides (RiPPs) biosynthesis, in which maturase enzymes introduce structural modifications after peptide production from the standard amino acids (10–14). All known fatty-acylated lipopeptides derive from NRPS pathways (1). The fatty acid moieties are usually attached to the peptide N terminus and derived from the primary metabolic pool, requiring no additional biosynthetic pathways.

Within the last two decades, genomic and biosynthetic studies on RiPPs have revealed an unexpected richness in maturases and structural modifications that remain only partially explored (13). Examples for more-recently discovered RiPP families are the proteusin and Nif11-type peptides (15–19), of which only few representatives have been characterized. These are defined by gene-encoded peptide precursors that carry an N-terminal nitrile hydratase-like leader peptide (NHLP) or a nitrogen fixation protein 11-like region. While the majority of products remain only genomically predicted, the few characterized members exhibit a remarkable richness in maturases catalyzing unusual posttranslational modifications: The cytotoxic polytheonamides and aeronamides are exceptionally complex β-helical proteusins containing an unparalleled density of modifications, including a large number of D-configured residues,

Significance

Fatty-acylated lipopeptides are metabolites of high medical and biotechnological importance, exemplified by the clinically used daptomycin, polymyxin, and echinocandin antiinfectives. The known representatives are produced by giant nonribosomal enzymes that usually generate heterogeneously acylated peptides requiring elaborate procedures to obtain defined compounds. While a few interesting acylated, ribosomally synthesized peptides have been described, none of them use fatty acids in their biosynthesis. Computational and synthetic biology methods identified a cryptic family of ribosomal lipopeptides present in diverse bacteria. A multifaceted structure elucidation approach revealed specific and distinct fatty acylations of amino acid side chains for three pathways. The results bring the production of tailored, gene-encoded lipopeptides using minimalistic gene sets into reach and provide guidelines for synthetic biology–based discovery.

Author contributions: F.H., N.M.B., C.C., B.I.M., M.R., M.G., S.L.R., A.L.V., and J.P. designed research; F.H., N.M.B., C.C., B.I.M., M.R., and S.L.R. performed research; F.H., N.M.B., C.C., S.L.R., A.L.V., and J.P. analyzed data; and F.H., N.M.B., C.C., B.I.M., M.G., S.L.R., A.L.V., and J.P. wrote the manuscript.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]F.H. and N.M.B. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: avagstad@ethz.ch or jpiel@ethz.ch.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2113120119/-/DCSupplemental.

Published January 13, 2022.

C- and *N*-methylations, and hydroxylations (16, 17, 20, 21). The antiviral landornamide proteusins harbor, beside lanthionine (Lan) bridges and D-amino acid residues, two ornithine residues installed by a peptide arginase as a new modification (18, 19, 22). For Nif11 RiPPs, a highly promiscuous lanthionine synthetase was identified that generates natural libraries of diversely cyclized peptides in cyanobacterial prochlorosins (23). In the spliceotide family, an extraordinary peptide-splicing reaction catalyzed by radical *S*-adenosylmethionine maturases generates β-amino acids with various side chains (24).

The abundance of new maturation reactions and the low number of known compounds suggest considerable untapped potential for peptide diversification. Here, we conducted a bioinformatic search for uncharacterized maturase genes that occur in proteusin- and Nif11-type RiPP biosynthetic gene clusters (BGCs) with high frequency. Heterologous expression of three selected gene clusters from an arctic, ocean-derived Rhodobacteraceae bacterium and two cyanobacteria revealed natural products with specific but distinct fatty acids attached to N-terminal ornithine or lysine side chains. The modification is catalyzed by members of the GCN5-related N-acetyltransferase (GNAT) superfamily that fall into a previously uncharacterized RiPP-associated clade (25). To our knowledge, the products that we identified comprise a previously unreported family of ribosomal lipopeptides carrying fatty acyl moieties. The advantages of gene-encoded rather than nonribosomal biosynthesis of the peptide portion and high fatty acyl selectivity provide potential for streamlined biosynthetic engineering of lipopeptides for drug discovery.

Results

Genome-mining Expands the Proteusin and Nif11 BGC Diversity. To assess the maturase diversity of bacterial NHLP and Nif11 RiPPs, we performed large-scale, targeted genome-mining (*SI Appendix*). We explored the sequence space using position-specific iterative basic local alignment search to detect remote homologs of characterized NHLP and Nif11 precursors (26). To differentiate candidate BGCs from genome regions not likely related to RiPP biosynthesis, precursor sequences were filtered to only include RiPP precursor sequences containing a characteristic substrate recognition motif frequently used for core release by proteolytic cleavage in various RiPP precursor types (27). After additional filtering steps, a resulting set of 2,346 candidate RiPP BGCs were found in 1,921 genomes spanning 378 microbial genera in more than 20 different phyla (Fig. 1*A*). Regarding their taxonomy, ∼41.6% of candidate BGCs were assigned to Cyanobacteria, followed by Firmicutes (18.6%) and Proteobacteria (16.7%). Since there are no shared protein domains conserved across all clusters in the dataset, we used a sequence alignment–independent approach to construct a distance matrix of co-occurring protein domains in BGCs and visualized their distribution using principal coordinates analysis (PCoA). BGCs for the characterized proteusins to date (polytheonamides, aeronamides, and landornamides) were identified, revealing large sections of the PCoA plot containing Nif11 and NHLP BGCs lacking any known natural products (Fig. 1*B*) (16, 18, 21). Protein domains that were best able to explain the variation between BGCs include Nif11 and NHLP precursor type (Fig. 1*D*) and ABC transporters (Fig. 1*E*), the latter often co-occurring with peptidase C39 domains (Fig. 1*F*) (27). During the course of our analysis, we noticed that members of the GNAT superfamily (PF00538) are broadly distributed among NHLP and Nif11-like BGCs (Fig. 1*G*). GNAT superfamily proteins occurred in 184 diverse BGCs or about 8% of candidate RiPP BGCs identified in this study. After common transporters and regulatory proteins were filtered out to focus on RiPP maturases, GNATs were the third-most-commonly occurring candidate maturase (Fig. 1*C*). Proteins in the GNAT superfamily are reported to have a versatile repertoire of substrates with respect to both the identity of the acyl donor and target molecule and are involved in basic cellular functions as well as specialized metabolism (25, 28). Therefore, we decided to explore their as-yet unknown role in proteusin and Nif11 RiPP biosynthesis.

Prioritization and Analysis of Proteusin BGCs Encoding GNATs. Based on the genome analyses, we prioritized two proteusin BGCs for our studies. One locus (termed *ksp* BGC; Fig. 1*B* and *SI Appendix*, Fig. S2) belongs to *Kamptonema* sp. PCC 6506, a cyanobacterium that is also the source of landornamides originating from a different (*osp*) cluster (18) and one (termed *pha*) from the marine alphaproteobacterium *Pseudophaeobacter arcticus* DSM 23566. The *pha* and *ksp* clusters encode members of the as-yet-uncharacterized, RiPP-associated GNAT family detected in our in silico analysis as well as members of known maturase families: both pathways contain an OspR-like (22) peptide arginase (PF12640; KspR and PhaR) and a LanM-type lanthionine synthetase with dehydratase and cyclization domains (PF05147 and PF13575; KspM and PhaM) expected to modify a single, colocalized NHLP-family precursor (KspA and PhaA). Unique to the *pha* cluster is a further gene encoding a predicted Fe(II)/α-ketoglutarate–dependent hydroxylase homolog (PF08007; PhaI) supported by a Phyre2 (29) analysis.

Based on the precursor core sequence and the sets of maturases, we hypothesized that the key structural features of these two proteusins consist of at least one (methyl)lanthionine (MeLan) bridge installed by the M proteins, an ornithine unit derived from arginine (R proteins), acyl moieties (N proteins), and for PhaI, a hydroxylation (Fig. 2 *A* and *D*). For both RiPP precursors, the region with NHLP homology ends with a Gly–Gly motif that commonly marks the leader-core junction. While the *pha* cluster encodes a predicted peptidase as part of the membrane-embedded export machinery (27), the *ksp* cluster lacks obvious protease or transporter genes but may utilize those of the landornamide (*osp*) BGC (18). Sequence analysis of the GNAT-type N proteins did not reveal any export signal motifs, suggesting that they act intracellularly prior to peptide cleavage. Although known GNAT homologs commonly transfer an acetyl moiety from acetyl-CoA to their substrates, some family members use other acyl units (25, 28, 30).

Pathway Reconstruction Using Heterologous Expression and Cell-Free Assays. To obtain insights into the products of the two prioritized pathways, we reconstructed them individually in *Escherichia coli* by cloning the precursor and complete sets of maturase genes into Duet Expression System vectors. Additionally, we cloned the full *pha* cluster with its native intergenic regions into pET28a, since its gene arrangement suggested a single-operon architecture with the precursor gene *phaA* at the upstream end. In all constructs, precursor genes were modified to encode an additional N-terminal His_6-tag. Precursor-only expressions were used as controls. The His-tagged precursors from KspA+KspNRM or PhaA+PhaNRIM coproductions in *E. coli* BL21(DE3) were purified using nickel affinity chromatography. Proteolytic digests with the endoproteinase GluC, and subsequent analyses of the digested precursor samples by high-performance liquid chromatography coupled to high-resolution mass spectrometry (HPLC-HRMS) revealed ions corresponding to a mixture of unmodified and matured core fragments (*SI Appendix*, Figs. S16 and S17). HPLC-HRMS analysis of the *ksp* coexpression (KspA+KspNRM) yielded GluC fragments with masses corresponding to a single dehydration and/or cyclization (**1**) (31), the conversion of four arginines to ornithines (**2**) (18, 22), and a combination of both modification types (**3**) (Fig. 2 and *SI Appendix*, Figs. S16 and S97). The *pha* coexpression (PhaA+PhaNRIM) products yielded the

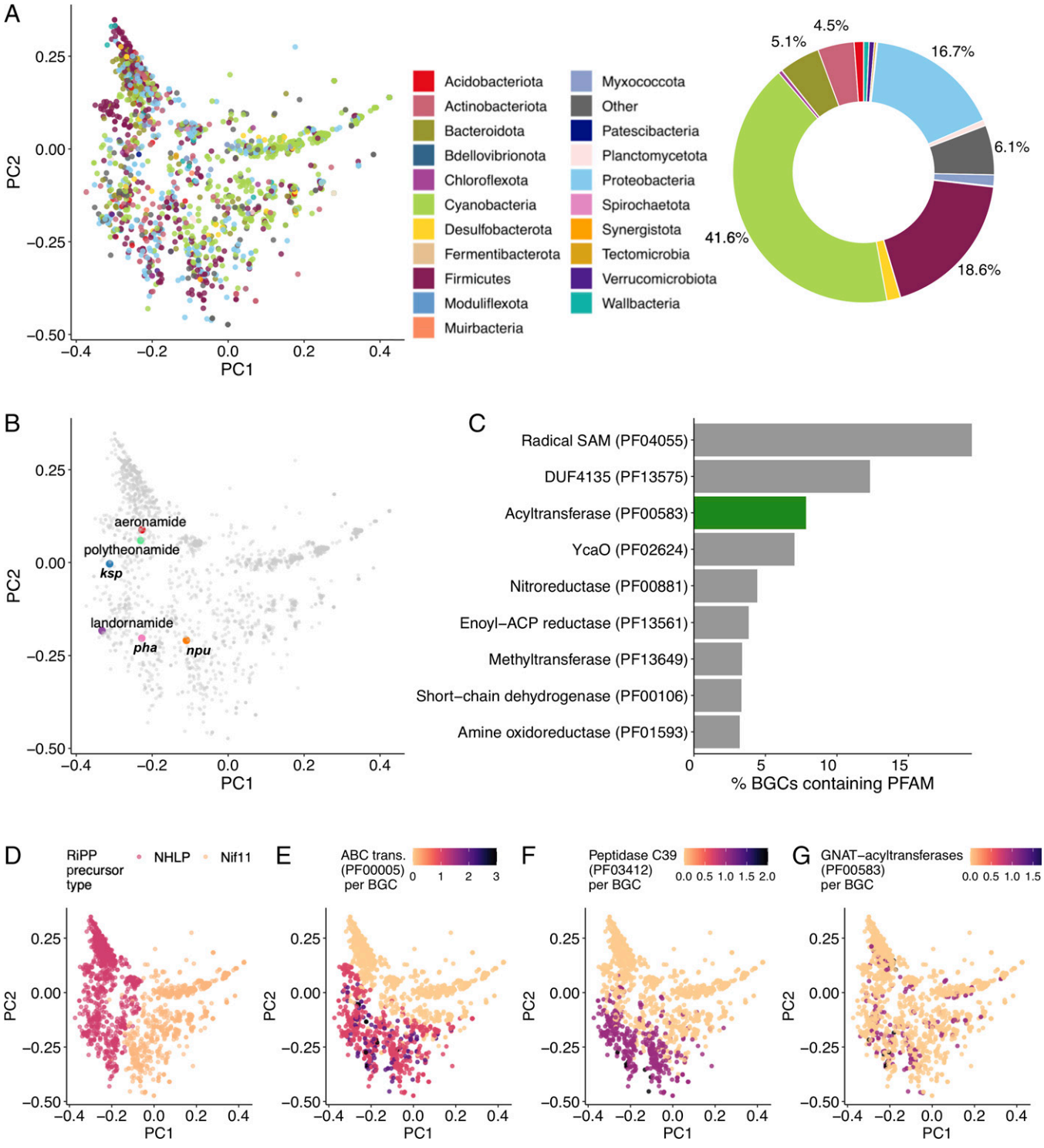

Fig. 1. (*A*) PCoA of 2,346 candidate BGCs containing one or more NHLP or Nif11-type RiPP precursors from 1,921 distinct bacterial genomes reveals wide taxonomic distribution. Jaccard distances between BGCs were calculated using a protein domain presence–absence matrix. Donut chart shows the relative proportions of different bacterial phyla containing candidate BGCs with one or more NHLP or Nif11-type RiPP precursors. (*B*) Characterized NHLP and Nif11-type BGCs are depicted with the following colored points: green: polytheonamide (16, 17), red: aeronamide (21), purple: landornamide (18), blue: ***ksp***, kamptornamide BGC (this study) from *Kamptonema* sp. PCC 6506, pink: ***pha***, phaeornamide from *P. arcticus* DSM 23566 (this study), and orange: ***npu***, nostolysamides from *N. punctiforme* PCC 73102 (this study). (*C*) Barplot of the most-abundant candidate maturases found in NHLP and Nif11-like RiPP BGCs. For clarity of visualization, only candidate protein families (PFAMs) likely to be involved in RiPP maturation are displayed. PFAMs associated with natural product transport, resistance, and regulation are displayed in *SI Appendix*, Fig. S1. The GNAT superfamily proteins (green, PF00583) highlighted in this study are the third-most-abundant maturase type. (*D*) The main driver of BGC separation along the *x*-axis (PC1) is the RiPP precursor type (NHLP or Nif11). BGC separation in the *y*-axis (PC2) is mainly driven by abundance of (*E*) ABC transporters and (*F*) peptidase C39 domains, which frequently co-occur. (*G*) GNAT homologs are broadly distributed across BGCs with different RiPP precursors and diverse protein domain combinations, suggesting that these maturases are involved in the modification of a wide array of RiPPs.

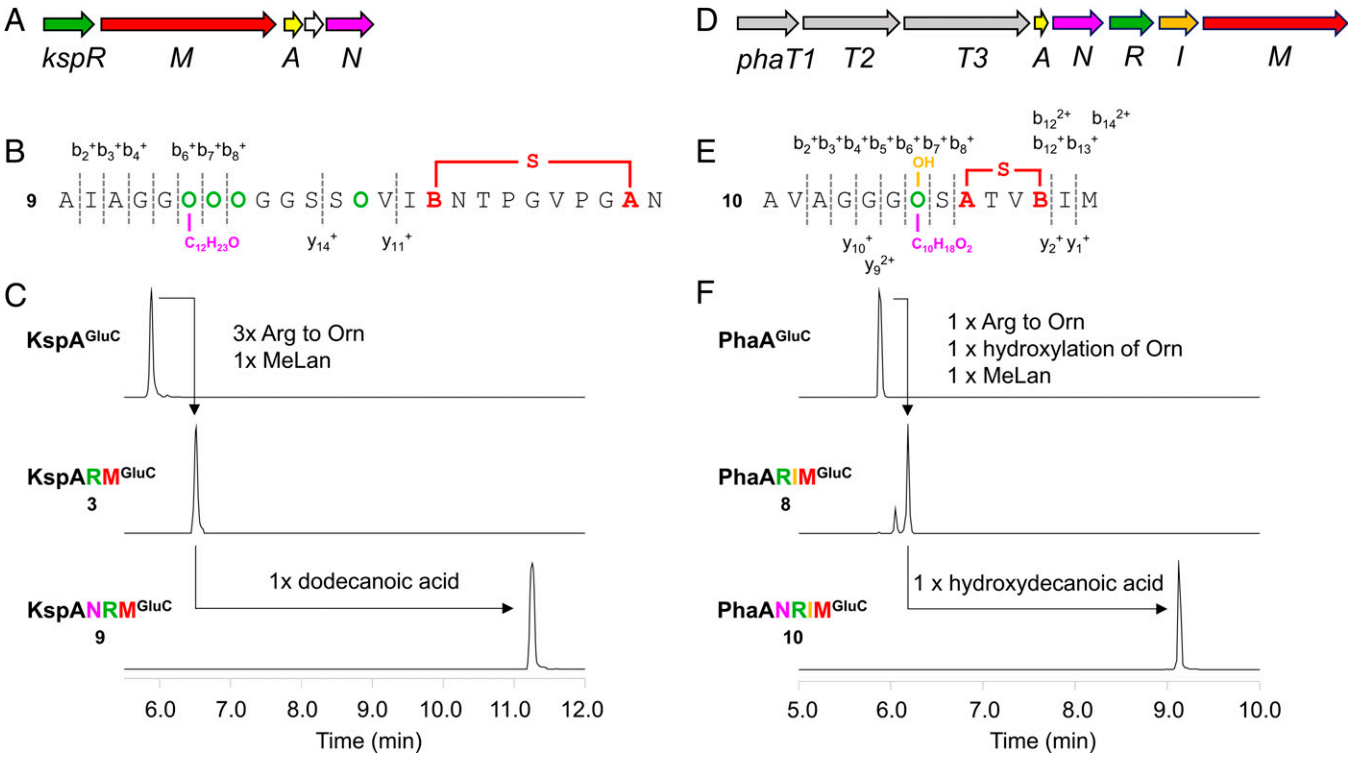

Fig. 2. (*A*) Architecture of the *ksp* cluster from *Kamptonema* sp. PCC 6506. (*B*) MS^2 fragmentation of the fully matured GluC fragment **9**. (*C*) HPLC-HRMS extracted ion chromatogram (EIC) of GluC fragments KspA, **3** and **9** from *kspANRM* coexpressions in *E. coli*. The arrows highlight extracted GluC fragments that are maturated based on MS^2 fragmentation. (*D*) Architecture of the *pha* cluster from *P. arcticus* DSM 23566. (*E*) MS^2 fragmentation of the fully matured GluC fragment **10**. (*F*) HPLC-HRMS EIC of GluC fragments PhaA, **8** and **10** from *phaANRIM* coexpressions in *E. coli*. The arrows highlight extracted GluC fragments that are maturated based on MS^2 fragmentation.

equivalent fragments **4**, **5**, and **6** with masses corresponding to one dehydration/cyclization (**4**), one ornithine residue (**5**), or both (**6**), respectively (*SI Appendix*, Figs. S17 and S97). Additionally, two fragments were observed, namely **7** with a mass corresponding to one ornithine residue in combination with one hydroxylation and **8** corresponding to the combination of one ornithine, one dehydration, and one hydroxylation (Fig. 2 and *SI Appendix*, Figs. S17 and S97). Tandem MS (MS^2) fragmentation of the tentatively most-matured fragment **8** indicated that the detected masses correspond to the expected modifications (*SI Appendix*, Fig. S19). Moreover, hydroxylation was exclusively observed for fragments containing the ornithine moiety and MS^2 fragmentation of **5**, **7**, and **8** suggest that it occurs at this residue (*SI Appendix*, Figs. S17, S19–S22).

In addition to these peptides, we also detected species **9** (for *ksp* expressions) and **10** (for *pha*) with unexpectedly large retention time and mass gains when compared to fragments **3** and **8** (Fig. 2 and *SI Appendix*, Fig. S97). The same shifts were observed for undigested versions of the fully matured precursor peptides. The formation of these peptides was dependent on the presence of the N proteins, since they were not observed in control experiments lacking these GNAT homologs (*SI Appendix*, Figs. S10 and S40). MS^2 fragmentation pinpointed the new modification to the ornithine residue at core position 1 of **9** and to the hydroxyornithine at core position 2 of **10**. Compared to **3** and **8**, the GluC-released products from *kspANRM* (**9**) and *phaANRIM* (**10**) coexpressions showed mass gains of 182.1672 Da and 170.1298 Da, respectively (Fig. 2). These suggested the molecular formula $C_{12}H_{22}O$ (*m/z* calculated = 182.1665 Da) for **9**, corresponding to a C_{12} acylation, and for **10** a C_{10} unit with the formula $C_{10}H_{18}O_2$ (*m/z* = 170.1301 Da). No further products with different mass gains attributed to KspN or PhaN were observed, suggesting that the maturases are highly selective for their respective distinct acyl units (*SI Appendix*, Figs. S16 and S17).

To test whether some of the observed modifications might be artifacts of *E. coli*, we performed cell-free assays using His_6-PhaA precursor purified from *E. coli* and incubated with lysate of *P. arcticus*. Subsequent nickel affinity purification and HPLC-HRMS analysis of the matured products revealed the same species as observed from heterologous whole-cluster expressions (*SI Appendix*, Figs. S54 and S55). This result suggests that the same modifications, including PhaN-dependent acylation, are also installed in the natural product. Again, no other acyl modifications were detectable for wild-type lysates, showing that PhaN is highly selective for the $C_{10}H_{18}O_2$ moiety in both bacteria. Moreover, the cell-free assays indicate that the pathway enzymes were expressed under the cultivation conditions. After extensive trials, we were also able to detect the expected mature compound in small amounts in cell pellet extracts of *P. arcticus* cultures grown under the same conditions (*SI Appendix*, Figs. S98 and S99).

Proteolytic Release and Purification of the Natural Products. To release the final natural products from their leaders in vitro, we applied the highly promiscuous peptidase domain of LahT, which had previously been shown to cleave a variety of precursor peptides following Gly–Gly motifs (27). HPLC-HRMS analyses of the proteolytic digests revealed the LahT-released fully matured cores **11** and **12** (Fig. 3) tantamount with the expected natural products that we named kamptornamide (**11**) and phaeornamide (**12**). A species with an identical HPLC-HRMS signature to **12** was also detected directly in pellet extracts from *P. arcticus* cultivations (*SI Appendix*, Figs. S98 and S99), highlighting that the heterologous product phaeornamide **12** is an authentic natural product.

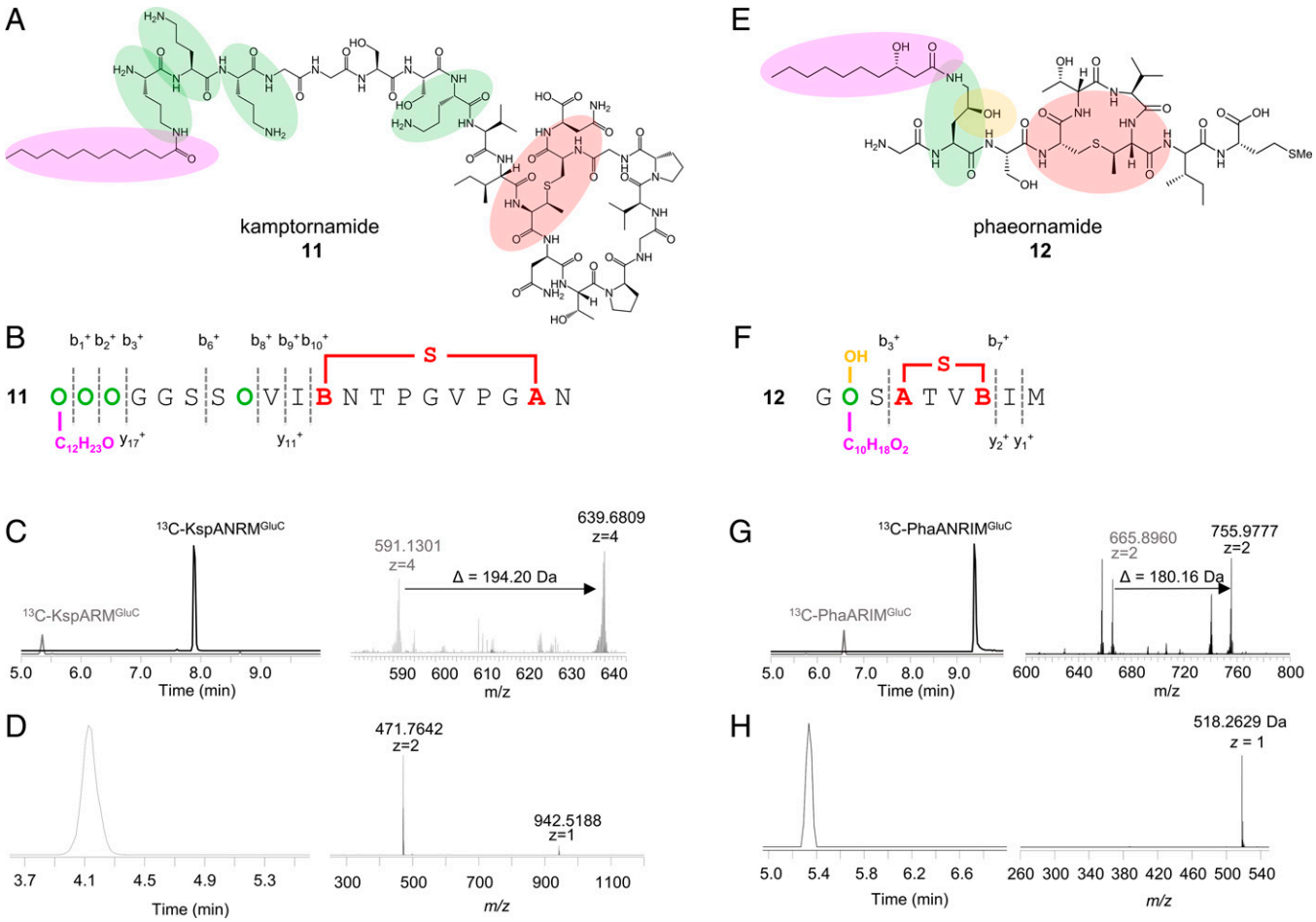

Fig. 3. (*A*) Structure of the putative natural product kamptornamide (**11**) of the *ksp* cluster. The key structural features installed by posttranslational modifications are highlighted in colors that correspond to the respective genes in Fig. 2*A*. (*B*) MS^2 fragmentation of the LahT fragment **11**. (*C*) EICs (left) and mass spectra (right) of the fully ^{13}C-labeled GluC fragments **3** and **9** from *kspANRM* coexpression using ^{13}C-Isogro broth. The mass shift between fully ^{13}C-labeled GluC fragments **3** and **9** is 194.20 Da. (*D*) EIC (left) and mass spectrum (right) of the released MeLan ring from triple digest of the fully matured precursor peptide KspANRM using the endopeptidase GluC and the exopeptidases aminopeptidase M and carboxypeptidase Y. The masses of 471.7642 Da (*m/z* = 2) and 942.5188 Da (*m/z* = 1) correspond to a nine-membered MeLan ring formed between Thr11 and Cys19 with an Asn20 extension. (*E*) Structure of the final natural product phaeornamide (**12**) of the *pha* cluster. The key structural features installed by posttranslational modifications are highlighted in colors that correspond to the respective genes in Fig. 2*D*. (*F*) MS^2 fragmentation of the LahT fragment **12**. (*G*) EICs (left) and mass spectra (right) of the fully ^{13}C-labeled GluC fragments **8** and **10** from *phaANRIM* coexpression using ^{13}C-Isogro broth. The mass shift between fully ^{13}C-labeled GluC fragments **8** and **10** is 180.16 Da. (*H*) EIC (left) and mass spectrum (right) of the released MeLan ring from triple digest of the fully matured precursor peptide PhaANRIM using the endopeptidase GluC and the exopeptidases aminopeptidase M and carboxypeptidase Y. The mass of 518.2629 Da (*m/z* = 1) corresponds to a four-membered MeLan ring formed between Cys4 and Thr7 with an Ile8 extension.

Extensive attempts failed to purify the heterologously produced natural products at larger scale after LahT cleavage using different combinations of methods such as foam extraction, preparative HPLC, and polymeric adsorbent beads (XAD; *SI Appendix*, *Supplementary Text*) (32). The lipopeptides largely resisted resolubilization in a variety of tested solvents after evaporation. Small amounts of pure **11** and semipure **12** were eventually recovered (*SI Appendix*, Figs. S62 and S63), but we were not able to obtain sufficient quantities for further experiments, such as bioactivity screenings and NMR spectroscopy; therefore, we resorted to other methods.

Elucidation of Key Structural Features. Full structure elucidation was achieved with a multifaceted approach using a combination of chemical derivatization, labeling studies, and peptide deconstruction. The mature, leader-bound peptides were degraded through chemical or proteolytic hydrolysis to release key moieties for individual NMR experiments or for comparison to authentic standards. This enabled us to infer the final structures by piecing together the resulting fragments, guided by the known core sequences and accompanying MS^2 data.

First, we determined the carbon content of the acyl moieties via heterologous expression of the two clusters in *E. coli* using ^{13}C-Isogro broth, nickel affinity purification, proteolytic digest, and HPLC-HRMS analysis. Cultivation in ^{13}C-Isogro broth, containing solely ^{13}C-labeled carbon sources, allowed us to express fully ^{13}C-labeled peptides to determine the carbon content of the fatty acyl moieties by comparison with the nonlabeled and nonacylated precursor peptides. This revealed a C_{12} moiety attached to kamptornamide (**11**; Fig. 3 and *SI Appendix*, Fig. S68), while an oxygen-containing C_{10} moiety is present in phaeornamide (**12**; Fig. 3 and *SI Appendix*, Fig. S67). To elucidate the nature and the regio- and stereochemistry of the oxygen-containing acylated hydroxyornithine, we liberated it from the matured, full-length phaeornamide precursors by triple-protease digests using a combination of the endopeptidase GluC and the promiscuous exoproteases aminopeptidase M and carboxypeptidase Y. The proposed nature of the

attached fatty acid was supported by detection of the expected masses and MS^2 fragments (Fig. 3 and *SI Appendix*, Figs. S69 and Fig. S70). Subsequent preparative HPLC purification and MS analysis supported the expected fatty-acylated amino acid (**13**) and a compound with a mass loss of 18.01 Da (**14**) that were separated from each other by another HPLC purification step (*SI Appendix*, Figs. S82, S83, and S97). One-dimensional and two-dimensional NMR spectroscopy of **13** and **14** showed that the mass loss in **14** was due to lactonization. We localized the hydroxyl moieties to the γ-position of ornithine (lactonized in **14**) and the β-position of an n-decanoyl moiety (*SI Appendix*, Table S10 and Figs. S84, S87–S94). By Mosher's derivatization of the lactone **14**, we established the absolute configuration of the hydroxylated fatty acid carbon as *S* (33). Subsequent circular dichroism (CD) measurement of the underivatized nonlactonized **13** assigned the absolute configuration of the hydroxylated carbon in the ornithine side chain as *S* as well (Fig. 3 and *SI Appendix*, Table S10 and Figs. S85, S86, S95, and S96). Similarly, acylated ornithine (**15**) was released by triple digest of the matured kamptornamide precursor (*SI Appendix*, Fig. S97). MS^2 fragments and the expected masses supported the nature of the fatty acyl moiety attached to ornithine (Fig. 3 and *SI Appendix*, Fig. S70).

Proteolytic treatments also provided information on the MeLan ring topologies. For kamptornamide (**11**), an intact cyclic fragment was obtained consisting of a nine-membered MeLan ring derived from Thr11 to Cys19 with an Asn20 extension. The phaeornamide (**12**) experiment yielded a four-membered MeLan ring of Cys4 to Thr7 with an Ile8 extension, as detected by HPLC-HRMS and supported by MS^2 fragmentation (Fig. 3 and *SI Appendix*, Figs. S69 and S70). The presence and number of MeLan rings were further supported by iodoacetamide (IAA) treatment of the matured, cleaved precursors, which acylates free thiols of unmodified cysteines (34). This showed that no free cysteines were present in the fully matured compounds. These results established the presence of a single MeLan ring in **11** and **12** (Fig. 3 and *SI Appendix*, Figs. S72–S76).

We further investigated the amino acid composition and configuration of the MeLan rings by acid hydrolysis of **11** and **12** and Marfey's analysis of the resulting monomers (18, 35). All expected single L-amino acids, including L-ornithine in **11** and hydroxyornithine in **12** (Fig. 3 and *SI Appendix*, Table S9 and Figs. S79–S81), were detected. However, hydroxyornithine was solely confirmed by MS, since no authentic standard was commercially available for this amino acid. Liberated MeLan was confirmed by comparison to a hydrolyzed and derivatized sample of commercially available nisin, a lanthipeptide comprising DL-MeLan modifications (35). Unexpectedly, the MeLan stereoisomers isolated from **11** and **12** exhibited slightly different retention times using an optimized HPLC gradient (*SI Appendix*, Table S9 and Figs. S79–S81). Whereas the presence of DL-MeLan in **11** was confirmed by coelution with the authentic standard, a proposed rare LL-MeLan in **12** was inferred by the isobaric mass and different retention time, since no standard was available (Fig. 3 and *SI Appendix*, Fig. S81 and *Supplementary Text*).

In summary, structure elucidation suggested final structures for kamptornamide (**11**) and phaeornamide (**12**) as shown in Fig. 3 and revealed them as fatty acid–containing ribosomal lipopeptides. We suggest the name selidamides for this structural type of RiPPs (from ancient Greek selida: side, based on the unprecedented fatty acyl side chains). As no congeners with different fatty acid chains were detected, the GNATs KspN and PhaN seem to select medium chain-length (hydroxy)fatty acids from the cellular pool with high specificity.

Implications for Biosynthesis from Maturase Coexpressions and Site-Directed Mutagenesis. For a more-detailed picture of how the peptide modifications are installed during biosynthesis, we coproduced the His-tagged precursors KspA and PhaA individually with their respective modifying enzymes or in variable sets and analyzed the distribution of products (*SI Appendix*, Figs. S3–S15 and S25–S46). For the *ksp* cluster, additional flanking genes of unknown function were screened for their ability to further modify KspA; however, no additional products were observed, suggesting the BGC boundaries shown in Fig. 2. HPLC-HRMS analysis of the coexpressions were in line with the product distributions (core peptide variants **1** through **8**) from full-cluster expressions (*SI Appendix*, Figs. S16, S17, and S32). Across all experiments, fatty acylation of the respective precursors by the GNATs was dependent on the conversion of the target arginine residue to ornithine by the arginases as a first step, as no acylated species were detected in expressions lacking KspR or PhaR but with KspN or PhaN present (*SI Appendix*, Figs. S5 and S28). In phaeornamide biosynthesis, although the fatty-acylated PhaA species were present in expressions lacking PhaI, acylation of ornithine was considerably more efficient when PhaI hydroxylase acted first to produce hydroxyornithine as the substrate residue of GNAT PhaN (e.g., compare PhaANR to PhaANRI; *SI Appendix*, Figs. S42 and S44). These results suggest the following preferred maturation order: 1) installation of ornithine by the arginases KspR and PhaR (*SI Appendix*, Figs. S8 and S29), 2) ornithine hydroxylation by the hydroxylase PhaI (not relevant to *ksp*) (*SI Appendix*, Figs. S32 and S37), and 3) acylation of (hydroxy)ornithine by the GNATs KspN and PhaN. MeLan formation by the lanthionine synthetase homologs KspM and PhaM appears to be somewhat independent of all other posttranslational modifications, since corresponding intermediates were present at every stage of arginine processing (*SI Appendix*, Fig. S61 and *Supplementary Text*).

The global BGC analysis (Fig. 1) suggested that many GNAT-associated precursors contain lysine but not arginine in their cores. To further explore the biocatalytic potential and substrate range of the maturases, we introduced point mutations into the *kspA* and *phaA* precursor cores. Analysis of the lysine-containing precursor variants KspA_R1K (**16**) and PhaA_R2K (**17**) coproduced with KspNRM and PhaNRIM, respectively, resulted in fully modified GluC fragments **18** (*ksp*) and **19** (*pha*) (*SI Appendix*, Figs. S56–S60 and S97). MeLan formation as well as hydroxylation and acylation of lysine residues were supported by MS^2 fragmentation. Furthermore, all observed products from these experiments support the proposed order of the PTM enzymes during biosynthesis (*SI Appendix*, Fig. S61 and *Supplementary Text*).

A Family of GNATs Generating Distinctly Acylated Lipopeptides. The transfer of specific, medium chain-length fatty acids to the (hydroxy)ornithine δ-N-moiety in kamptornamide (**11**) and phaeornamide (**12**) in distinct bacterial phyla encouraged us to further explore the distribution of selidamides and their variation regarding the attached fatty acyl moieties. Therefore, we performed a genome neighborhood analysis using the Enzyme Function Initiative Genome Neighborhood Tool (EFI-GNT) (36) with the respective GNAT sequences KspN and PhaN as templates. Manual filtering was performed for BGCs encoding predicted precursors, lanthionine synthetases, and optionally arginases within a window of 10 genes upstream and downstream of the acyltransferase gene hit (*SI Appendix*, Figs. S64 and S65). Precursor peptides associated with these BGCs belong to either the NHLP or the Nif11 families, except one from *Herpetosiphon aurantiacus* DSM 785 that contains precursor genes with a shorter leader (37). Precursor genes encoded in the selidamide clusters encode peptides with accumulation of arginine (largely in clusters with arginase genes) and lysine residues predominantly at the N and/or C termini of the putative cores as potential GNAT acylation sites (*SI Appendix*, Table S7).

The role and distribution of the selidamide acyltransferases within the larger GNAT superfamily was evaluated using a maximum-likelihood phylogenetic tree containing characterized representatives from each GNAT family, including those of goadvionins, lipolanthines, and the putative selidamide GNATs filtered from our EFI-GNT analysis (*SI Appendix*, Table S8) (25, 28, 30, 36, 38, 39). The two characterized selidamide GNATs were located in a well-separated clade containing members from RiPP BGCs (*SI Appendix*, Fig. S66). Within this larger clade, two subclades exist, of which subclade *a* contains solely cyanobacterial homologs, while subclade *b* harbors proteins of diverse bacterial origin (Cyanobacteria, Acidobacteria, and one Chloroflexi), including the cyanobacterial KspN. Regarding these two subclades, PhaN and GNAT homologs from Proteobacteria and Firmicutes branch earlier from the tree.

To further test the biochemical homogeneity of the putative selidamide GNAT clade, we selected another member from a distinct phylogenetic region, subclade *a*. The BGC, located in the cyanobacterium *Nostoc punctiforme* PCC 73102 (*npu* cluster), is distinct to the NHLP- and arginase-encoding *ksp* and *pha* clusters and exhibits a minimalistic architecture with genes predicted to encode a Nif11-type precursor, a GNAT, and a lanthionine synthetase (*SI Appendix*, Fig. S2 and Table S1). Based on the absence of an arginase homolog and the previous mutational experiments, we expected fatty acylation to occur at a lysine ε-*N*-moiety. The *npu* pathway was reconstructed in *E. coli* using the procedures for the *ksp* and *pha* clusters. For coproduction experiments of NpuA with the lanthionine synthetase NpuM and the GNAT NpuN, the HPLC trace showed a new peak relative to precursor-only controls, indicating successful modification by at least one enzyme (*SI Appendix*, Figs. S47, S48, and S50). Coproduction of the precursor NpuA with only either NpuM or NpuN showed that both enzymes are catalytically active independent of each other. For the *npuAM* coexpression, LahT fragments **20a** through **20d** were detected with masses corresponding to the loss of up to four waters and the respective phosphorylated intermediates, suggesting formation of up to four MeLan units, which is consistent with four Cys and five Ser or Thr residues in the 25-mer core (*SI Appendix*, Figs. S48 and S97). MS^2 fragmentation indicated that the dehydrations and cyclizations occurred in the C-terminal part of the core region. Moreover, the lack of fragmentation between Cys14 and Cys25 of the core supported at least two Lan rings involving Ser18 and Ser23 as the only hydroxylated residues in this region. Based on the presence of three Cys residues in this section (Cys14, Cys20, and Cys25), the connectivity remained elusive (*SI Appendix*, Fig. S49). LahT fragments **21** and **22** obtained from *npuAN* coexpression showed a mass gain of 236.2084 Da and 252.2160 Da, respectively, consistent with the attachment by the GNAT of a $C_{16}H_{28}O$ unit (m/z calculated = 236.2135 Da) for **21** and either a $C_{16}H_{28}O_2$ (m/z calculated = 252.2084 Da) or $C_{17}H_{32}O$ (m/z calculated = 252.2448 Da) unit for **22** in equal amounts (*SI Appendix*, Figs. S51 and S97). Consequently, for *npuANM* coexpressions, we observed the expected LahT fragments **23** and **24** with masses (m/z = 722.8614 Da and 726.8633 Da) corresponding to a combination of modifications from the single-maturase experiments (*SI Appendix*, Fig. S97). Subsequent MS^2 fragmentation of the products revealed the fatty acyl attachment site as Lys1 of the core. In addition, the fragmentation pattern supports at least two Lan rings (*SI Appendix*, Figs. S52 and S53). In accordance with the other described family members, we named the putative final proteusin natural products nostolysamide A (**23**) and B (**24**). Furthermore, the observed MS^2 fragments suggested hexadecenoyl and hydroxyhexadecenoyl moieties in **23** and **24**, respectively (*SI Appendix*, Figs. S52 and S53). The two compounds were subjected to similar, more-detailed structure elucidation experiments as used above for kamptornamide **11** and phaeornamide **12** but due to low expression yields, only a partial characterization of the modified Lys monomers released by proteolysis was possible. This, however, supported the presence of C_{16} fatty acid moieties in both final products (*SI Appendix*, Fig. S71). IAA treatment and MS^2 fragmentation also revealed that the major quadruply dehydrated species contains two Lan rings of unknown connectivity involving Ser18 and Ser23, as inferred above from MS^2 data, and two dehydrobutyrates converted from Thr9 and Thr12 (*SI Appendix*, Figs. S77 and S78). Beside this main species, MS ions of LahT-released core variants support the formation of one, three, and four MeLan rings as minor products (*SI Appendix*, Figs. S77 and S78).

Discussion

Nonribosomally biosynthesized peptides have long been regarded as structurally more diverse than their ribosomal counterparts as they are able to incorporate a wide range of noncanonical amino acids and other acyl units. However, more-recent discoveries have revealed a wide range of "nonribosomal-like" posttranslational modifications in RiPPs resulting in polyketide–peptide hybrids (38, 39), D-amino acids (16), β-amino acids (24, 40), amide *N*-methylations (41, 42), side-chain cross links (43, 44), and many other variants (13) that have blurred the structural boundaries between the two peptide classes. Here, we show that RiPP analogs also exist for nonribosomal lipopeptides, a large, biomedically important family of natural products from a wide range of microorganisms. Selidamides, as defined by RiPPs carrying fatty acyl modifications introduced by GNAT-like maturases ("N enzymes"), share several structural properties with classical lipopeptides exemplified by the clinical antibiotics daptomycin and echinocandins: the representatives characterized in this study contain central metabolism-derived fatty acid moieties attached to nitrogen atoms and linear and cyclic peptide portions (Fig. 4). A contrast to nonribosomal products is that fatty acid acylation occurs at side-chain amino groups rather than the N terminus and that cyclization is accomplished by lanthionine bridges instead of ester or amide bonds. Although fatty acid acylation of lysine side chains by palmitic acid (C_{16}) and myristic acid (C_{14}) is known for proteins from eukaryotes and bacteria, such as the RTX toxins, these structural features in RiPPs have, to our knowledge, not been reported (45–47). The known RiPP lipopeptide families lipolanthines (38) and goadvionins (39) carry N-terminal acyl moieties of polyketide origin with polar urea or ammonium head groups as key structural features produced by dedicated biosynthetic machinery from their respective gene clusters (Fig. 4). For goadvionins, members of the GNAT family transfer an acyl carrier protein–bound polyketide unit (39), while the acylation mechanism of lipolanthines is not yet characterized. Another prominent example of RiPPs with lipophilic moieties are cyanobactins containing one to several prenyl moieties derived from primary metabolism (48).

A remarkable feature of the characterized kamptornamide (**11**), phaeornamide (**12**), and nostolysamides (**23** and **24**) is the high selectivity with which fatty acylation occurred in our experiments. Products of each BGC were modified by a fatty acyl unit of distinct length (C_{12}, C_{10}, and C_{16}) and, for **11** and **12**, even specific functionalization (saturated versus β-hydroxy fatty acid). These fatty acyl units were selected by the respective GNATs from a diverse pool of available fatty acid chain lengths in *E. coli* (49) during heterologous expressions, and for phaeornamide (**12**), the selectivity of PhaN was recapitulated in the native producer *P. arcticus* (*SI Appendix*, Figs. S54, S55, and S98). Although we expect the observed acyl chain-length specificity for *ksp* and *npu* GNATs to be identical in their native cyanobacterial producers, we cannot rule out that the observed products are an artifact from expression in *E. coli*. In contrast,

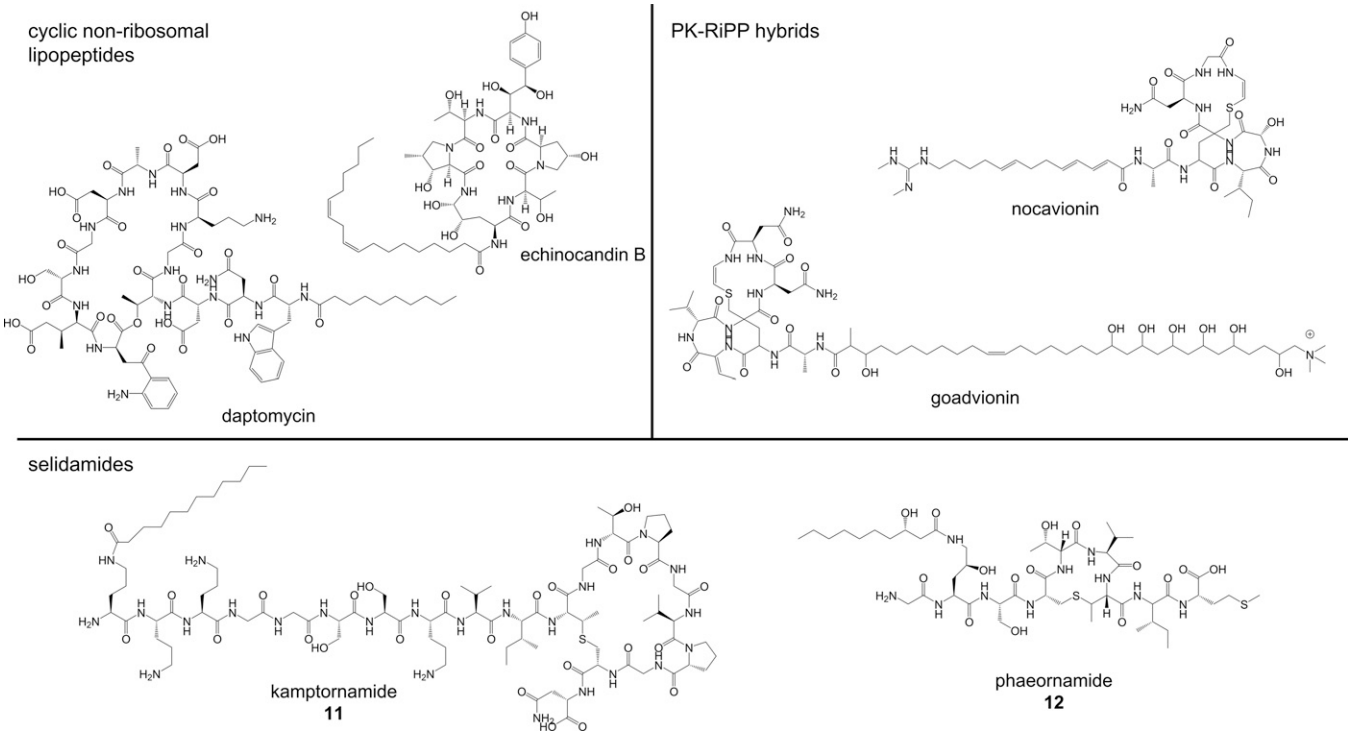

Fig. 4. Lipopeptide structures from different biosynthetic origins. In the *Upper Left* panel, the cyclic nonribosomal peptides daptomycin and echinocandin B are shown. The cyclic nonribosomal peptides are N-terminally acylated with fatty acids from primary metabolism: daptomycin, decanoic acid and echinocandin B, 9,12-octadecadienoic acid. In the *Upper Right* panel, the polyketide-RiPP (PK-RiPP) hybrids nocavionin (lipolanthines) and goadvionin A are shown. The ribosomally synthesized peptide chains containing an avionin ring structure are N-terminally attached with long acyl moieties from polyketide biosynthetic origin. The *Lower* panel shows representatives from the RiPP lipopeptide family selidamides. These ribosomal peptides containing a MeLan ring are *N*-acylated at amino acid side chains with fatty acids from primary metabolism: kamptornamide (**11**), ornithine with dodecanoic acid and phaeornamide (**12**), 4-(*S*)-hydroxyornithine with 3-(*S*)-hydroxydecanoic acid.

nonribosomal lipopeptides including clinical antibiotics are usually produced as diversely fatty-acylated substance mixtures (50). For such compounds, the chain length can considerably influence the therapeutic value (51, 52). For example, for the heterogeneously fatty-acylated A21978C complex produced by *Streptomyces roseosporus*, only the least-toxic decanoyl congener is obtained using an elaborate procedure for use as the therapeutic daptomycin (51). Another example is the variation of fatty acid chain length in surfactins that results, for example, in lower hemolytic activity (52). In ongoing work, we aim to investigate the biocatalytic value of selidamide GNATs and determine the structural features that allow acylation with such exquisitely tuned selectivity. This property, as well as the gene-encoded nature and minimalistic BGC architecture, suggests potential applications for generating diverse "pseudo-nonribosomal" lipopeptides in a streamlined way based on nonribosomal lead structures (9, 53). The suggested potential is supported by our initial site-directed mutagenesis experiments that showed for the ornithine-modifying enzymes KspN and PhaN promiscuity for lysine.

The presence of additional putative selidamide BGCs, as supported by their gene compositions and by GNAT phylogenetic analysis, indicates a larger, unmined diversity of natural lipopeptides. Although phaeornamide **12** was observed in small amounts from *P. arcticus*, attempts to detect the putative natural products in extracts of the wild-type cyanobacteria producers have failed so far. The absence of detectable terminally processed compounds in the cyanobacterial wild-type producers is an observation that we reported for other proteusin BGCs (21), suggesting that these compounds are produced in low quantities or are sequestered in some way that makes their isolation challenging. A recent transcriptomic analysis of BGCs in the nostolysamides producer *N. punctiforme* PCC 73102 shows that many pathways are transcribed below a constitutive level and that screening alternative cultivation conditions, including high-density cultures, only improved transcription levels in approximately half of the studied BGCs (54). These results support that unknown bottlenecks often block attempts to directly discover natural products from the host. However, the synthetic biology-based discovery method demonstrated here should also be applicable to further selidamides, since it has been successful for all three tested BGCs. Considering the high selectivity of selidamide GNATs, this might require an expanded set of expression hosts in case *E. coli* lacks the preferred fatty acids of the natural source (e.g., branched fatty acids) (49, 55, 56).

In summary, the identification of selidamides by a combined genomic mining–synthetic biology approach underlines an unusually high maturase diversity and discovery potential of proteusin- and Nif11-type RiPPs. These insights could pave the way toward a pharmacological and biocatalytic toolbox for gene-encoded lipopeptides using transferases with desired fatty acid specificities.

Materials and Methods

For further details, refer to *SI Appendix, Materials and Methods*.

Construction and Site-Directed Mutagenesis of Expression Vectors. Construction of vectors for heterologous expression was performed as described before (18, 19). Genes encoding precursor peptides and maturase enzymes were amplified by PCR using primers designed for ligation cloning (*SI Appendix*, Table S2*A*) from genomic DNA of *K.* sp. PCC 6506, *N. punctiforme* PCC 73102, and *P. arcticus* DSM 23566, respectively. Site-directed mutagenesis of the precursor peptides pACYC-*kspA_R1K*, pACYC-*phaA_R2K*, and pET28a-*phaA_R2K_*

NRIM was performed using QuikChange primers (*SI Appendix*, Table S2*B*). All obtained constructs were sequence verified (Microsynth AG, Zürich).

(Co)expression and Purification of (Matured) Precursor Peptides. For heterologous expression, the constructs containing precursor and maturase genes were introduced into *E. coli* BL21(DE3). The expression cultures in baffled or Ultra-Yield flasks containing terrific broth or Isogro broth, and the appropriate antibiotics were incubated at 37 °C and 200 rpm up to an optical density at 600 nm of 1.2 to 1.4. Protein expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside, and the cultures were transferred to 16 °C and 180 rpm for 20 h (*pha* and *npu* cluster) or 3 d (*ksp* cluster) before harvested by centrifugation. Purification was conducted by nickel affinity chromatography as described before (18, 19, 22). All fractions were analyzed using sodium dodecyl sulfate—polyacrylamide gel electrophoresis (SDS-PAGE) with a 15% acrylamide gel.

Proteolytic Digests of (Matured) Precursor Peptides. For all proteolytic digests in small scale, ∼10 μg of matured or unmodified precursor peptide was cleaved with the commercially available endopeptidase GluC (New England Biolabs [NEB]), porcine kidney aminopeptidase M (Merck), and 1 μg baker's yeast carboxypeptidase Y (Sigma-Aldrich) in 1× GluC buffer (NEB) with 1 mM Tris(2-carboxyethyl)phosphine (TCEP) or dithiothreitol (DTT). For digests with the promiscuous peptidase LahT, 100 mM Tris·HCl, pH 8.0, with 1 mM TCEP or DTT was used.

For large-scale proteolytic digests, 5 to 15 mg of fully matured precursor peptides were cleaved. All other ingredients were adapted accordingly.

HPLC-HRMS Analysis. Analytical HPLC-MS samples (5- to 20-μL injections) were separated on a Dionex UltiMate 3000 RS UHPLC equipped with a Phenomenex Kinetex C18 (2.6 μm, 100 Å, 150 × 4.6 mm) column heated to 50 °C. HPLC separation used H_2O +0.1% formic acid (FA) and acetonitrile (ACN) +0.1% FA as solvent. The HPLC was coupled to a Thermo Fisher Scientific Q Exactive Hybrid Quadrupole-Orbitrap Mass Spectrometer using heated electrospray ionization in positive ion mode for full MS and MS^2 fragmentation.

HPLC Purification of Acylated Hydroxyornithine. Purification of hydroxyornithine acylated with hydroxydecanoic acid released by large-scale triple digests was conducted on an Agilent 1260 HPLC equipped with a Phenomenex Luna C18 column at ambient temperature. HPLC separation was performed with H_2O +0.1% FA and ACN +0.1% FA as solvents.

Fractions containing the target molecule were pooled, evaporated, and dissolved in 50% ACN. A second round of purification was performed. Fractions containing the purified compound were pooled, evaporated, and dissolved in DMSO-d6 for NMR spectroscopy.

Chemical Modifications of (Matured) Precursor Peptides. Marfey's derivatization for amino acid analysis was performed as described before (19, 21). IAA treatment of free thiol residues was performed according to the method reviewed by Lohans and Vederas (34). Mosher's derivatization was performed after solvent exchange of hydroxyornithine acylated with hydroxydecanoic acid from DMSO-d6 to pyridine-d5 by adding the respective Mosher enantiomer (33). Circular dichroism (CD) spectroscopy was measured in ethanol as solvent.

Data Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We want to thank Prof. Dr. Wilfred van der Donk (University of Illinois Urbana–Champaign) for kindly sharing the plasmid pET-Duet-LahT150. We want to thank Dr. Alexander O. Brachmann for helpful discussions on experimental design and data analysis and Lester Frei and Stefan Leopold-Messer for experimental assistance (Frei for labeling experiments and Messer for NMR spectroscopy experiments and evaluation). We are grateful for financial support (F.H. for financial support by the Vontobel Foundation, S.L.R. for financial support by the ETH Zürich Postdoctoral Fellowship [20-1 FEL-07], and J.P. for financial support by the Swiss National Science Foundation (SNSF) [National Research Program 72 "Antimicrobial resistance," 407240_167051]).

1. S. Götze, P. Stallforth, Structure elucidation of bacterial nonribosomal lipopeptides. *Org. Biomol. Chem.* **18**, 1710–1727 (2020).
2. J. A. Hutchinson, S. Burholt, I. W. Hamley, Peptide hormones and lipopeptides: From self-assembly to therapeutic applications. *J. Pept. Sci.* **23**, 82–94 (2017).
3. W. Hüttel, Echinocandins: Structural diversity, biosynthesis, and development of antimycotics. *Appl. Microbiol. Biotechnol.* **105**, 55–66 (2021).
4. C. Gutiérrez-Chávez, N. Benaud, B. C. Ferrari, The ecological roles of microbial lipopeptides: Where are we going? *Comput. Struct. Biotechnol. J.* **19**, 1400–1413 (2021).
5. S. Götze, P. Stallforth, Structure, properties, and biological functions of nonribosomal lipopeptides from pseudomonads. *Nat. Prod. Rep.* **37**, 29–54 (2020).
6. L. Girard, M. Höfte, R. De Mot, Lipopeptide families at the interface between pathogenic and beneficial *Pseudomonas*-plant interactions. *Crit. Rev. Microbiol.* **46**, 397–419 (2020).
7. S. Zhang, R. Mukherji, S. Chowdhury, L. Reimer, P. Stallforth, Lipopeptide-mediated bacterial interaction enables cooperative predator defense. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2013759118 (2021).
8. M. Muttenthaler, G. F. King, D. J. Adams, P. F. Alewood, Trends in peptide drug discovery. *Nat. Rev. Drug Discov.* **20**, 309–325 (2021).
9. P. J. Naughton, R. Marchant, V. Naughton, I. M. Banat, Microbial biosurfactants: Current trends and applications in agricultural and biomedical industries. *J. Appl. Microbiol.* **127**, 12–28 (2019).
10. H. Kries, Biosynthetic engineering of nonribosomal peptide synthetases. *J. Pept. Sci.* **22**, 564–570 (2016).
11. J. A. McIntosh, M. S. Donia, E. W. Schmidt, Ribosomal peptide natural products: Bridging the ribosomal and nonribosomal worlds. *Nat. Prod. Rep.* **26**, 537–559 (2009).
12. T. A. Scott, J. Piel, The hidden enzymology of bacterial natural product biosynthesis. *Nat. Rev. Chem.* **3**, 404–425 (2019).
13. M. Montalbán-López *et al.*, New developments in RiPP discovery, enzymology and engineering. *Nat. Prod. Rep.* **38**, 130–239 (2021).
14. F. Ruijne, O. P. Kuipers, Combinatorial biosynthesis for the generation of new-to-nature peptide antimicrobials. *Biochem. Soc. Trans.* **49**, 203–215 (2021).
15. D. H. Haft, M. K. Basu, D. A. Mitchell, Expansion of ribosomally produced natural products: A nitrile hydratase- and Nif11-related precursor family. *BMC Biol.* **8**, 70 (2010).
16. M. F. Freeman *et al.*, Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* **338**, 387–390 (2012).
17. M. F. Freeman, M. J. Helf, A. Bhushan, B. I. Morinaka, J. Piel, Seven enzymes create extraordinary molecular complexity in an uncultivated bacterium. *Nat. Chem.* **9**, 387–395 (2017).
18. N. M. Bösch *et al.*, Landornamides: Antiviral ornithine-containing ribosomal peptides discovered through genome mining. *Angew. Chem. Int. Ed. Engl.* **59**, 11763–11768 (2020).
19. B. I. Morinaka *et al.*, Radical S-adenosyl methionine epimerases: Regioselective introduction of diverse D-amino acid patterns into peptide natural products. *Angew. Chem. Int. Ed. Engl.* **53**, 8503–8507 (2014).
20. M. C. Wilson *et al.*, An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506**, 58–62 (2014).
21. A. Bhushan, P. J. Egli, E. E. Peters, M. F. Freeman, J. Piel, Genome mining- and synthetic biology-enabled production of hypermodified peptides. *Nat. Chem.* **11**, 931–939 (2019).
22. S. Mordhorst, B. I. Morinaka, A. L. Vagstad, J. Piel, Posttranslationally acting arginases provide a ribosomal route to non-proteinogenic ornithine residues in diverse peptide sequences. *Angew. Chem. Int. Ed. Engl.* **59**, 21442–21447 (2020).
23. B. Li *et al.*, Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10430–10435 (2010).
24. B. I. Morinaka *et al.*, Natural noncanonical protein splicing yields products with diverse β-amino acid residues. *Science* **359**, 779–782 (2018).
25. A. I. M. Salah Ud-Din, A. Tikhomirova, A. Roujeinikova, Structure and functional diversity of GCN5-related N-acetyltransferases (GNAT). *Int. J. Mol. Sci.* **17**, 1018 (2016).
26. S. F. Altschul *et al.*, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
27. S. C. Bobeica *et al.*, Insights into AMS/PCAT transporters from biochemical and structural characterization of a double Glycine motif protease. *eLife* **8**, e42305 (2019).
28. M. W. Vetting *et al.*, Structure and functions of the GNAT superfamily of acetyltransferases. *Arch. Biochem. Biophys.* **433**, 212–226 (2005).
29. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
30. L. Favrot, J. S. Blanchard, O. Vergnolle, Bacterial GCN5-related N-acetyltransferases: From resistance to regulation. *Biochemistry* **55**, 989–1002 (2016).
31. S. Mukherjee, W. A. van der Donk, Mechanistic studies on the substrate-tolerant lanthipeptide synthetase ProcM. *J. Am. Chem. Soc.* **136**, 10450–10459 (2014).
32. P. Biniarz, M. Łukaszewicz, T. Janek, Screening concepts, characterization and structural analysis of microbial-derived bioactive lipopeptides: A review. *Crit. Rev. Biotechnol.* **37**, 393–410 (2017).
33. T. R. Hoye, C. S. Jeffrey, F. Shao, Mosher ester analysis for the determination of absolute configuration of stereogenic (chiral) carbinol carbons. *Nat. Protoc.* **2**, 2451–2458 (2007).
34. C. T. Lohans, J. C. Vederas, Structural characterization of thioether-bridged bacteriocins. *J. Antibiot. (Tokyo)* **67**, 23–30 (2014).
35. L. M. Repka, J. R. Chekan, S. K. Nair, W. A. van der Donk, Mechanistic understanding of lanthipeptide biosynthetic enzymes. *Chem. Rev.* **117**, 5457–5520 (2017).

36. R. Zallot, N. Oberg, J. A. Gerlt, The EFI web resource for genomic enzymology tools: Leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **58**, 4169–4182 (2019).
37. K. Blin *et al.*, antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47** (W1), W81–W87 (2019).
38. V. Wiebach *et al.*, The anti-staphylococcal lipolanthines are ribosomally synthesized lipopeptides. *Nat. Chem. Biol.* **14**, 652–654 (2018).
39. R. Kozakai *et al.*, Acyltransferase that catalyses the condensation of polyketide and peptide moieties of goadvionin hybrid lipopeptides. *Nat. Chem.* **12**, 869–877 (2020).
40. J. Z. Acedo *et al.*, O-methyltransferase-mediated incorporation of a β-amino acid in lanthipeptides. *J. Am. Chem. Soc.* **141**, 16790–16801 (2019).
41. M. R. Quijano *et al.*, Distinct autocatalytic α-N-Methylating Precursors Expand the Borosin RiPP Family of Peptide Natural Products. *J. Am. Chem. Soc.* **141**, 9637–9644 (2019).
42. S. Ramm *et al.*, A self-sacrificing N-methyltransferase is the precursor of the fungal natural product omphalotin. *Angew. Chem. Int. Ed. Engl.* **56**, 9994–9997 (2017).
43. A. Caruso, R. J. Martinie, L. B. Bushin, M. R. Seyedsayamdost, Macrocyclization via an arginine-tyrosine crosslink broadens the reaction scope of radical *S*-adenosylmethionine enzymes. *J. Am. Chem. Soc.* **141**, 16610–16614 (2019).
44. T. Q. N. Nguyen *et al.*, Post-translational formation of strained cyclophanes in bacteria. *Nat. Chem.* **12**, 1042–1053 (2020).
45. M. D. Resh, Fatty acylation of proteins: The long and the short of it. *Prog. Lipid Res.* **63**, 120–131 (2016).
46. A. Osickova *et al.*, Acyltransferase-mediated selection of the length of the fatty acyl chain and of the acylation site governs activation of bacterial RTX toxins. *J. Biol. Chem.* **295**, 9268–9280 (2020).
47. B. Chen, Y. Sun, J. Niu, G. K. Jarugumilli, X. Wu, Protein lipidation in cell signaling and diseases: Function, regulation and therapeutic opportunities. *Cell Chem. Biol.* **25**, 817–831 (2018).
48. J. A. McIntosh, M. S. Donia, S. K. Nair, E. W. Schmidt, Enzymatic basis of ribosomal peptide prenylation in cyanobacteria. *J. Am. Chem. Soc.* **133**, 13698–13705 (2011).
49. Y. Li *et al.*, Differentiation of bacteria using fatty acid profiles from gas chromatography-tandem mass spectrometry. *J. Sci. Food Agric.* **90**, 1380–1383 (2010).
50. G. Pirri, A. Giuliani, S. Nicoletto, L. Pizzuto, A. Rinaldi, Lipopeptides as anti-infectives: A practical perspective. *Open Life Sci.* **4**, 258–273 (2009).
51. L. Robbel, M. A. Marahiel, Daptomycin, a bacterial lipopeptide synthesized by a nonribosomal machinery. *J. Biol. Chem.* **285**, 27501–27508 (2010).
52. A. Théatre *et al.*, The surfactin-like lipopeptides from *Bacillus* spp.: Natural biodiversity and synthetic biology for a broader application range. *Front. Bioeng. Biotechnol.* **9**, 623701 (2021).
53. T. Dang, R. D. Süssmuth, Bioactive peptide natural products as lead structures for medicinal use. *Acc. Chem. Res.* **50**, 1566–1576 (2017).
54. D. Dehm *et al.*, Unlocking the spatial control of secondary metabolism uncovers hidden natural product diversity in *Nostoc punctiforme*. *ACS Chem. Biol.* **14**, 1271–1279 (2019).
55. D.-C. Zhang *et al.*, *Phaeobacter arcticus* sp. nov., a psychrophilic bacterium isolated from the Arctic. *Int. J. Syst. Evol. Microbiol.* **58**, 1384–1387 (2008).
56. M. Gugger *et al.*, Cellular fatty acids as chemotaxonomic markers of the genera *Anabaena*, *Aphanizomenon*, *Microcystis*, *Nostoc*, and *Planktothrix* (cyanobacteria). *Int. J. Syst. Evol. Microbiol.* **52**, 1007–1015 (2002).