HHS Public Access

Author manuscript

Nat Chem. Author manuscript; available in PMC 2024 August 14.

Published in final edited form as:

Nat Chem. 2024 August ; 16(8): 1320–1329. doi:10.1038/s41557-024-01491-3.

Non-modular Fatty Acid Synthases Yield Unique Acylation in Ribosomal Peptides

Hengqian Ren[1,2,7], **Chunshuai Huang**[2,7], **Yuwei Pan**[3], **Shravan R. Dommaraju**[2,4], **Haiyang Cui**[2], **Maolin Li**[2], **Mayuresh G. Gadgil**[2,4], **Douglas A. Mitchell**[2,4,5], **Huimin Zhao**[1,2,4,6,*]

[1]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[2]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[3]Department of Molecular and Cellular Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[4]Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[5]Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[6]Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA.

[7]These authors contributed equally to this work: Hengqian Ren, Chunshuai Huang.

Abstract

Recent efforts in genome mining of ribosomally synthesized and post-translationally modified peptides (RiPPs) have expanded the diversity of post-translational modification chemistries. However, RiPPs are rarely reported as hybrid molecules incorporating biosynthetic machinery from other natural product families. Here, we report lipoavitides, a class of RiPP/fatty acid hybrid lipopeptides that display a unique, putatively membrane-targeting 4-hydroxy-2,4-dimethylpentanoyl (HMP)-modified *N*-terminus. The HMP is formed via condensation of isobutyryl-coenzyme A (isobutyryl-CoA) and methylmalonyl-CoA catalyzed by a 3-ketoacyl-acyl carrier protein (3-ketoacyl-ACP) synthase III enzyme, followed by successive tailoring reactions in the fatty acid biosynthetic pathway. The HMP and RiPP substructures are then connected by an acyltransferase exhibiting promiscuous activity towards the fatty acyl and RiPP substrates. Overall, the discovery of lipoavitides contributes a prototype of RiPP/fatty acid hybrids and provides possible enzymatic tools for lipopeptide bioengineering.

*Correspondence to: zhao5@illinois.edu.

Author contributions H.R. and C.H. contributed equally to this work. H.R. performed bioinformatics analysis, direct cloning, heterologous expression, HR-MS/MS analysis, stereochemistry determination, bioactivity assays, mutational analysis, and enzymatic reactions. C.H. performed compound purification, isotopic feeding study, and intermediate characterization. C.H. also designed, performed, and analyzed NMR experiments. Y.P. assisted in compound purification and genetic manipulation. H.C. performed AlphaFold2 and AutoDock analysis. D.A.M. contributed to the bioinformatics analysis of *lpv* genes. S.R.D. performed Marfey's analysis, reductive desulfurization, and interpretation of NMR data to assign stereochemistry with oversight from D.A.M. M.L. designed and assisted in synthesizing racemic and (*S*)-DPD standards. M.G.G. performed and analyzed ^{1}H-^{1}H ROESY NMR experiments with oversight from D.A.M. H.R. and C.H. wrote the manuscript with editorial oversight from H.Z., along with input from all other authors. H.Z. conceived of and supervised the overall project.

Competing interests The authors declare no competing interests.

Introduction

Recent advances in genome sequencing, bioinformatics, synthetic biology, and analytical chemistry have enabled the genome mining of ribosomally synthesized and post-translationally modified peptides (RiPPs)[1, 2]. RiPP biosynthesis starts with the expression of genetically encoded precursor peptides, followed by an organized post-translational modification (PTM) process. While only starting from 20 proteinogenic amino acids, RiPPs reveal vast structural diversity and hence a broad range of biological functions endowed by the disparate PTMs[3, 4]. To date, 49 class-defining PTMs have been identified[2], most of which are derived from the precursor peptide backbone and side chains. For example, the oxazol(in)e and thiazol(in)e heterocycles, which are found in the linear azol(in)e-containing peptide (LAP)[5], bottromycin[6], thiopeptide[7], and cyanobactin[8] RiPP classes, are formed through multistep reactions between the hydroxyl or thiol group of serine, threonine, or cysteine and the preceding backbone carbonyl[9]. RiPPs can also be modified with structures that are not produced by their biosynthetic gene clusters (BGCs), as exemplified by glycocins, which are synthesized using host endogenous nucleoside diphosphate (NDP)-monosaccharides as glycosyl donors[10]. In some cases, modifying enzymes from different RiPP classes are also found in a single BGC to produce RiPPs with combined class-defining PTMs[11]. However, RiPP BGCs comprising non-RiPP biosynthetic genes/pathways for synthesizing unconventional PTMs are rarely reported.

Hybrid BGCs, encoding biosynthetic machinery from multiple natural product families, are widely distributed in nature, often resulting in products with merged structural features[12–14]. Notable examples include polyketide synthase–nonribosomal peptide synthetase hybrid BGCs, in which modules of type I polyketide synthases and nonribosomal peptide synthetases work together to produce compounds sharing structural features from both families[15]. Only a few RiPPs, including the *e*-series thiopeptides[16], pheganomycins[17], cacaoidin[18], lymphostin[19], lipolanthines[20, 21], and goadvionins[22], are known to be synthesized by hybrid machinery (Fig. 1). To synthesize the non-RiPP-derived components of these molecules, the corresponding BGCs encode enzymes of other natural product families, such as the polyketide synthases in the goadvionin BGC. Investigation into these hybrid BGCs has unveiled biosynthetic modes for the synthesis and fusion of RiPP and non-RiPP parts, providing potential routes for peptide natural product bioengineering.

In this work, we uncovered a class of hybrid BGCs that produce ribosomally derived lipopeptides. In addition to RiPP biosynthetic enzymes, these BGCs ubiquitously encode a 3-ketoacyl-acyl carrier protein (3-ketoacyl-ACP) synthase III (FabH) and other enzymes resembling type II fatty acid synthases. Direct cloning and heterologous expression of a representative BGC from *Streptomyces* sp. NRRL S-1521 yielded an aminovinylmethyl-cysteine (AviMeCys)-containing peptide with a 4-hydroxy-2,4-dimethylpentanoyl (HMP)-modified *N*-terminus. This natural product thus represents a member of a RiPP class termed the lipoavitides (AviMeCys-containing lipopeptides). The isolated lipoavitide exhibited hemolytic activity, and the subsequent structure-activity relationship study suggested an essential role of the *N*-terminal fatty acyl moiety. The HMP moiety is synthesized through condensation of isobutyryl-coenzyme A (isobutyryl-CoA) and methylmalonyl-CoA by the FabH, followed by successive oxidation and reduction reactions, yielding HMP-CoA. An

acyltransferase then utilizes HMP-CoA as the acyl donor and appends HMP to the RiPP portion of lipoavitide. Characterization of the acyltransferase showed a broad substrate tolerance, which implies future potential in lipopeptide engineering. Overall, the discovery of lipoavitide serves as a prototype for RiPP/fatty acid hybrids.

Results

RiPP BGCs encoding enzymes for fatty acid biosynthesis

Among 439 Actinomycetota species sequenced in a previous work[23], genome mining by AntiSMASH[24] revealed a putative BGC (*lpv*) in *Streptomyces* sp. NRRL S-1521, consisting of genes for synthesizing multiple families of natural products (Fig. 2, Supplementary Table 1, Supplementary Table 2). Encoded by *lpvBCD* are a HopA1-phosphotransferase pair and a flavin-dependent decarboxylase, a set of enzymes commonly known for synthesizing aminovinyl(methyl)-cysteine [Avi(Me)Cys] moieties[25] in RiPPs, exemplified by the thioamitides[26] and class V lanthipeptides[18, 27, 28]. A plethora of other known RiPP biosynthetic enzymes, such as dehydroalanine reductase (LpvJ), methyltransferase (LpvM1), and metalloproteases (LpvP1 and LpvP2), are also encoded in *lpv* (Supplementary Fig. 1, Supplementary Fig. 2). The *lpv* pathway further encodes a putative NDP-monosaccharide pathway (LpvLNOQM2) and a glycosyltransferase LpvI, suggesting the BGC produces a glycosylated RiPP(s). Of note, several biosynthetic enzymes (LpvFGHVS) in *lpv* are predicted to be involved in fatty acid/polyketide cyclization, elongation, keto/enoyl reduction, and oxidation[29] (Supplementary Table 3, Supplementary Fig. 3, Supplementary Fig. 4). These enzymes are uncommon in RiPP BGCs and might produce an acyl group for usage by the acyltransferase LpvE.

To explore the distribution of *lpv* BGCs, we performed PSI-BLAST searches of the NCBI non-redundant database using LpvB (phosphotransferase), LpvC (HopA1), LpvE (acyltransferase), and LpvG (FabH) as queries, yielding 15 BGCs that have local homologs of all four enzymes and encompass seven unique precursor peptide sequences (Fig. 2, Supplementary Fig. 5). Alignment of these precursor peptides shows a conserved TxxxHC motif, commonly found in thioviridamide-like thioamitides (Supplementary Table 4). We therefore hypothesized that the threonine and cysteine would form the AviMeCys moiety and the histidine would be transformed into β-hydroxy-*N,N*-dimethyl-L-histidine (hdmHis) through methylation and hydroxylation by LpvM1 and LpvK, respectively[30, 31] (Supplementary Fig. 1). Sequence alignment also identified a conserved serine residue, which is predicted to be converted into dehydroalanine by LpvB and LpvC, and subsequently reduced to D-alanine by the dehydroalanine reductase LpvJ[27]. Despite being absent in one of the seven precursor peptides, the tyrosine next to the TxxxHC motif is a likely glycosylation site, as observed in cacaoidin[18] (Fig. 1). The replacement of tyrosine by tryptophan in the precursor peptide StsA concurred with the absence of the NDP-monosaccharide pathway and glycosyltransferase in the corresponding BGC, which supports our hypothesis (Fig. 2, Supplementary Fig. 5). Through bioinformatics analysis, we tentatively assigned functions to many of the encoded genes; however, no indications were found for residue(s) that may undergo acylation.

Discovery of lipoavitide with an acylated *N*-terminus

To investigate the function of the acyltransferase and enzymes for fatty acid/polyketide synthesis, we next experimentally characterized the *lpv* BGC. The Cas12a-assisted precise targeted cloning using *in vivo* Cre-*lox* recombination (CAPTURE) method was used to directly clone *lpv* for heterologous expression[32]. The cloned *lpv* and the pBE45 empty vector were individually conjugated into *Streptomyces lividans* TK24 and *Streptomyces albus* J1074. After 5 d of cultivation, colonies were picked, extracted by methanol, and analyzed by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF-MS). Upon transformation with *lpv*, *S. lividans* TK24 showed no significant metabolic differences to the pBE45 negative control on mass spectra. However, a prominent ion at *m/z* 1636 was observed from the *lpv*-expressing *S. albus* J1074 (Fig. 3, Supplementary Fig. 6). A 10 L production culture was then prepared to afford ~10 mg of material for structural characterization (Supplementary Fig. 7). The high-performance liquid chromatography (HPLC)-purified product was first analyzed by high-resolution mass spectrometry and tandem mass spectrometry (HR-MS/MS) (Supplementary Fig. 8). The hdmHis was corroborated by the formation of 4-formyl-1,3-dimethyl-1*H*-imidazolium fragment through a retro aldol reaction[26, 30, 33]. We also assigned a series of y-ions to the precursor peptide, with molecular weight differences implicating serine conversion into alanine and tyrosine glycosylation.

To elucidate the structure of lipoavitide **1**, nuclear magnetic resonance (NMR) spectroscopy was performed, confirming most of the proposed PTMs (Fig. 3, Supplementary Table 5, Supplementary Fig. 9–15). The heteronuclear multiple bond correlation (HMBC) confirmed the conversion of serine into alanine, with Marfey's analysis establishing the D-configuration[34] (Supplementary Fig. 16, Supplementary Fig. 17). The tyrosine modification was uncovered as an *O*-glycosylation by 2-*O*-methyl-β−6-deoxygulose putatively synthesized by the NDP-monosaccharide pathway (LpvLNOQM2) and glycosyltransferase LpvI (Supplementary Fig. 18, Supplementary Fig. 19). The formation of hdmHis was confirmed by HMBC correlations (Supplementary Fig. 20). The AviMeCys moiety was formed at the *C*-terminus, with a *Z*-alkene between C-5 and C-6 suggested by the coupling constant between H-5 and H-6 ($\delta_{H\text{-}5}$ = 5.61, $\delta_{H\text{-}6}$ = 7.29, $^{3}J_{H\text{-}5,H\text{-}6}$ = 6.2 Hz) (Supplementary Fig. 21). To elucidate the stereochemistry at C-2 and C-3, the AviMeCys was linearized via reductive desulfurization, followed by hydrolysis releasing the original threonine residue as aminobutyric acid (Abu). The peptide hydrolysate was derivatized by Marfey's method[34] for liquid chromatography-mass spectrometry (LC-MS) analysis which detected L-Abu and D-Abu. We attributed L-Abu to reductively desulfurized Met9, indicating D-Abu arises from AviMeCys with an (*S*)-configuration at C-2 (Supplementary Fig. 22). Indicative rotating frame overhauser enhancement spectroscopy (ROESY) correlations within AviMeCys between H-3 and H-2, H-4 and H-2, and H-4 and NH-2 ($\delta_{NH\text{-}2}$ = 8.75, $\delta_{H\text{-}2}$ = 4.04, $\delta_{H\text{-}3}$ = 3.05, $\delta_{H\text{-}4}$ = 1.51) suggested an (*S*)-configuration at C-3, which was further supported by the lack of a ROESY correlation between H-3 and NH-2 (Supplementary Fig. 23). Altogether, these data suggested the formation of *S*-[(*Z*)-2-aminovinyl]-(3*S*)-3-methyl-D-cysteine during macrocyclization.

Notably, we found a 4-hydroxy-2,4-dimethylpentanoyl (HMP) moiety attached to the *N*-terminus of **1**, corroborating our initial hypothesis that *lpv* produces an acylated peptide (Supplementary Fig. 24). To investigate the stereochemistry at C-2, we first hydrolyzed lipoavitide under mildly acidic conditions[35], yielding 2,4,4-trimethyl-γ-butyrolactone derived from HMP and the corresponding peptide with a free *N*-terminus (**2**) (Supplementary Fig. 25–28). Of note, the tyrosine was also partially deglycosylated after hydrolysis, forming the aglycone of **2** (**3**). Next, we reduced 2,4,4-trimethyl-γ-butyrolactone into 2,4-dimethyl-1,4-pentanediol (DPD) using $LiAlH_4$ and chemically synthesized both racemic and (*S*)-DPD standards (Supplementary Fig. 29–31). Chiral gas chromatography afforded sufficient separation of the DPD enantiomers, and co-injection of DPD from **1** with the synthesized DPD standards revealed the (*S*)-configuration for C-2 in the HMP moiety of **1** (Supplementary Fig. 32). To explore the uniqueness of HMP and 2-*O*-methyl-β−6-deoxygulose, we performed substructure searches in SciFinder and found no reported natural products consisting of these units. As the product can be structurally classified as a lipopeptide, we therefore termed this RiPP class lipoavitides (AviMeCys-containing lipopeptides).

Isolated lipoavitide is a hemolysin

We first tested lipoavitide **1** for antimicrobial activity using a standard agar disk diffusion assay. No growth inhibition was observed for the tested Bacillota, Actinomycetota, Pseudomonadota, and Ascomycota strains (Supplementary Table 6). Owing to their amphiphilicity, many lipopeptides insert into the lipid bilayer and disrupt the integrity of cellular membrane[36]. To evaluate hemolytic activity, lipoavitide **1** was added to bovine erythrocytes. We observed cell lysis at low micromolar concentrations (Fig. 3). To investigate if the HMP-modified *N*-terminus contributed to the hemolytic activity, peptides **2** and **3** obtained via hydrolysis of **1**, and racemic potassium 4-hydroxy-2,4-dimethylpentanoate **4**, were tested for hemolytic activity with none causing significant cell lysis, suggesting the essentiality of HMP to the hemolytic activity of **1**.

Production of HMP-CoA via a fatty acid biosynthetic pathway

Bioinformatic analysis suggested that the HMP moiety is produced by the *lpv* fatty acid/polyketide-like biosynthetic enzymes. Considering the predicted functional annotations of enzymes in *lpv* and the structure of **1**, we hypothesized that HMP biosynthesis starts with condensation of isobutyryl-CoA **5** and methylmalonyl-CoA **6** catalyzed by LpvG, resulting in 3-oxo-2,4-dimethylpentanoyl-CoA **7**. Subsequently, **7** would be successively converted into possible intermediates such as **8** to yield HMP-CoA **10**. LpvE then utilizes **10** as the acyl donor to generate the HMP-modified *N*-terminus (Fig. 4). To confirm or refute this hypothesis, we first disrupted *lpvE* and *lpvG*, respectively, using the RecET recombination system[37]. The resulting BGCs were expressed in *S. albus* J1074, and both produced intermediate **2** with an unmodified *N*-terminus (Fig. 5). We next performed isotopic-labeling experiments to gain insight into HMP biosynthesis. *S. albus* J1074 containing the *lpv*-locus was cultured in ISP4 solid medium supplemented with either [$^{13}C_5$]-L-valine or [2-^{13}C]-propionate to obtain ^{13}C-labeled **5** and **6**, respectively[38, 39] (Supplementary Fig. 33). Isotopically labeled **1** was produced from both substrates, as confirmed by MALDI-TOF-

MS (Supplementary Fig. 34). Analysis by ^{13}C-NMR confirmed isotopic incorporation at C-3, C-4, C-5 and C-7 in HMP when supplementing [$^{13}C_5$]-L-valine, and at C-2 in HMP when supplementing [2-^{13}C]-propionate (Fig. 5, Supplementary Fig. 35).

Next, we characterized LpvG *in vitro*. Unlike FabH from canonical fatty acid biosynthesis, which requires an acyl carrier protein (ACP)-bound extender unit, we predicted LpvG would directly utilize acyl-CoA because *lpv* lacks ACP and malonyl-CoA-ACP transacylase (FabD) homologs (Supplementary Fig. 3). Such ACP-independent activities have been reported in the FabH family[40, 41]. Unfortunately, *E. coli* expression failed to produce soluble LpvG (Supplementary Table 7), and thus we examined other FabH-like enzymes from putative orthologous *lpv* BGCs. StsG from *Streptomyces* sp. NRRL S-920 is highly sequence similar to LpvG (70% identity over 94% coverage) and was expressible and purifiable from *E. coli* in a soluble form (Supplementary Fig. 36). Using isobutyryl-CoA **5** and racemic methylmalonyl-CoA **6** as substrates, we reconstituted the activity of StsG *in vitro*, leading to the formation of 3-oxo-2,4-dimethylpentanoyl-CoA **7**, as confirmed by HPLC and high-resolution mass spectrometry (HR-MS) (Supplementary Fig. 37, Supplementary Fig. 38). We observed that upon extended time and usage of higher enzyme concentrations, only around half of the substrates were consumed. Therefore, we suspected that StsG is selectively acting on one diastereomer of **6**. We then conducted the reaction at varying substrate ratios, and the ratio of product **7** to unreacted **6** remained about 1:1, even when **5** was supplied in excess. However, when a methylmalonyl-CoA epimerase[42] was added into the reaction that interconverts the diastereomers of **6**, both **5** and **6** were consumed to near completeness (Fig. 5, Supplementary Fig. 39).

LpvV, a predicted short-chain dehydrogenase/reductase, was next characterized (Supplementary Table 1, Supplementary Table 2). Disruption of *lpvV* did not alter the production of **1**, suggesting compensation by endogenous *S. albus* enzymes (Supplementary Fig. 40). LpvV was expressed and purified from *E. coli* for *in vitro* characterization (Supplementary Fig. 36). When LpvV was supplied to the StsG reaction with NADH or NADPH, the 3-oxo-2,4-dimethylpentanoyl-CoA **7** was converted to 3-hydroxy-2,4-dimethylpentanoyl (3-HMP)-CoA **8**, as monitored by HPLC and HR-MS (Fig. 5, Supplementary Fig. 41). LpvV prefers NADH, as demonstrated by the incomplete conversion of **7** even when NADPH was supplied at a concentration ten times higher than used for NADH.

Conversion of 3-HMP-CoA **8** to HMP-CoA **10** putatively occurs through dehydration, reduction, and hydroxylation catalyzed by LpvF, LpvH, and LpvS, respectively (Fig.4, Supplementary Fig. 42). Disruption of *lpvF* and *lpvH* did not provide insights into this conversion, yielding HPLC profiles similar to the wild type BGC or complete abolishment of **1** biosynthesis without any detectable intermediates. However, the *lpvS*-disrupted BGC produced **11**, which is isobaric to **1** (Fig. 5, Supplementary Fig. 43). We isolated the product and elucidated its structure using NMR, which shares all structural features of **1** but contains a 3-HMP-modified *N*-terminus (Supplementary Table 8, Supplementary Fig. 44–46). Next, we sought to characterize LpvF, LpvH, and LpvS *in vitro*. Expression of *lpvF*, *lpvH*, or their *sts* orthologous genes in *E. coli* failed to produce any soluble proteins (Supplementary Table 7), but LpvS was successfully expressed in *E. coli* and purified in a soluble form

(Supplementary Fig. 36). Owing to the formation of 3-HMP-modified peptide **11** by the *lpvS*-disrupted BGC, we first examined LpvS using **8** as the substrate, yet no activity was detected (Supplementary Fig. 47). In contrast, 2,4-dimethylpentanoyl-CoA **9** synthesized via a carbonyldiimidazole-activation strategy[43] could be consumed by LpvS (Supplementary Fig. 48). HPLC analysis of the reaction mixture detected free CoA, which was suspected from the spontaneous hydrolysis of **10** (Fig. 5). In fact, subsequent LC-MS analysis indeed confirmed the presence of 2,4,4-trimethyl-γ-butyrolactone, corroborating **10** formation.

N-terminus acylation by LpvE with substrate promiscuity

From gene disruption studies, we propose LpvE forms **1** via condensation of **2** and **10** (Fig. 4 and Fig. 5). To characterize its acyltransferase activity *in vitro*, LpvE was expressed in *E. coli* and purified (Supplementary Fig. 36). We first supplied LpvE and **2** to the LpvS reaction, resulting in a 2,4-dimethylpentanoyl-modified peptide **12** (Supplementary Fig. 49). This supports the acyltransferase activity of LpvE, although it may require other proteins to effectively use the unstable **10** as a substrate. Isolation of **11** suggests that LpvE can also use 3-HMP-CoA **8** as a substrate. To provide more insights into the function of LpvE, we introduced LpvE and **2** to the StsG/LpvV reaction. MS analysis of the reaction products showed a prominent peak at *m/z* 1636, consistent with the formation of **11** (Fig. 6). The identity of *m/z* 1636 was confirmed through HPLC co-elution of the reaction mixture and a standard of **11** produced from the *lpvS*-disrupted BGC (Fig. 6, Supplementary Fig. 43). Omission of StsG or LpvV from the reaction abolished production of **11**. However, mass spectral analysis revealed the addition of isobutyryl or 3-oxo-2,4-dimethylpentanoyl moieties to **2** yielded **13** and **14** with significantly reduced conversions. This result implies that isobutyryl-CoA **5** and 3-oxo-2,4-dimethylpentanoyl-CoA **7** are not native substrates but are still tolerated by LpvE (Supplementary Fig. 50). Despite the ample supply of **5** through primary metabolism, **13** was not detected from the *S. albus* containing the *lpvG*-disrupted BGC, indicating its low *in vivo* productivity which is likely attributed to the limited activity of LpvE towards **5**. The inability of LpvE to use methylmalonyl-CoA **6** also suggested that LpvG-catalyzed condensation occurs prior to acylation. Interestingly, LpvE can also utilize acyl-CoAs with straight or cyclized carbon chains (**15**-**19**), aromatic (**20**), and even alkynyl moieties (**21**), albeit with reduced efficiency, implied its tolerance to different acyl donors (Fig. 6, Supplementary Fig. 51–52).

Since LpvE tolerated various acyl donors, we next investigated tolerance towards various peptide substrates. We first asked if modified residues in the linear region of **1** (region ii, iii, and iv) would affect acylation (Fig. 6). Val1 of the core peptide was replaced with alanine using RecET-assisted recombination, yielding **22** at a comparable titer to **1**. The molecular weight of **22** is 28 Da smaller than **1**, suggesting the variant was HMP-modified at the *N*-terminus (Fig. 6, Supplementary Fig. 53). However, when Val1 was substituted with lysine, no expected product was detected by MS, possibly due to interference in acylation or leader peptide removal. Acylated products **23**-**25** were also observed upon disruption of LpvJ (dehydroalanine reductase) or substitution of the core residue Ser3 (Fig. 6, Supplementary Fig. 54). Disrupting LpvI (glycosyltransferase) led to the formation of acylated product **26** lacking glycosylated tyrosine. However, no acylated products with or without glycosylation were detected by MALDI-TOF-MS when this residue was substituted by phenylalanine or

tryptophan (Fig. 6, Supplementary Fig. 55). Although tolerated by LpvE, most structural changes in region ii and iv led to a significant decrease in productivity of the corresponding peptide variants, suggesting acylation follows the glycosylation and D-alanine formation. The rate of *in vitro* acylation of **3** was much slower than for **2**, which also corroborates this hypothesis (Supplementary Fig. 56).

We next investigated the LpvE tolerance to the PTMs in the macrocyclic region. The hdmHis moiety exerts little impact on acylation, as acylated products **27** and **28** consisting of unmodified and dimethylated histidine were observed with a comparable titer to **1** upon disruption of LpvM1 and LpvK[30, 31] (Supplementary Fig. 57). Nevertheless, *in vitro* acylation of linear peptide VG^{D}AIIYTFMIHC with D-alanine at position 3 was unsuccessful, which revealed that the AviMeCys macrocyclization was required for LpvE-catalyzed acylation (Supplementary Fig. 58).

To shed more light on the promiscuous activity of LpvE, we predicted the protein structure by AlphaFold2[44] and performed molecular docking studies with AutoDock[45] (Supplementary Fig. 59). The *N*-terminus of **2** and the HMP moiety of **10** were directed into the putative active site for condensation. Most of the other residues in **2** are located outside the substrate-binding pocket of LpvE, possibly explaining the tolerance to structural variations at these sites. In contrast, an analog of **2** without the AviMeCys was randomly docked with LpvE, indicating the structural rigidity afforded by the macrocycle in **2** might be essential (Supplementary Fig. 60). Phylogenetic analysis revealed that lipoavitide acyltransferases form a distinct clade within the GCN5-related *N*-acetyltransferase (GNAT) superfamily (Supplementary Fig. 61, Supplementary Table 9). LpvE possesses an overall structure similar to that of other GNAT acyltransferases, such as goadvionin acyltransferase GdvG[22] and chloramphenicol acyltransferase PA2578[46] (Supplementary Fig. 62). Nevertheless, structural superimposition reveals differences in amino acid composition within the predicted acyl-CoA binding sites, which could affect their preference for acyl-CoA substrates. To further investigate the substrate tolerance of lipoavitide acyltransferases, we successfully reconstituted the activity of StsE *in vitro* using both **2** and **3** as substrates, even though 8 of the 12 residues in the StsA core peptide differ from those in LpvA (Fig. 2, Supplementary Fig. 36, Supplementary Fig. 63). However, LpvE is unable to accommodate other RiPP products harboring unmodified *N*-termini, such as thatisin[47] and darobactin[48] (Supplementary Fig. 64). We anticipate that lipoavitide acyltransferases will use a broad range of substrates, yet further characterization is needed to explore the key factors determining their substrate selectivity.

Discussion

Advances in DNA sequencing and bioinformatics have enabled the detection of hybrid BGCs with increasing frequency. These pathways often produce natural products consisting of structural characteristics that span multiple families, of which genome mining can significantly expand natural product structural diversity and biosynthetic routes[49–51]. However, only a few RiPP/non-RiPP hybrid BGCs have been discovered. Here, we leveraged the CAPTURE method to enable rapid characterization of BGCs in high complexity. In doing so, we uncovered a class of RiPP/fatty acid hybrid BGCs that produce

lipoavitides characterized by a unique fatty acyl moiety (HMP), an AviMeCys macrocycle, and other modified amino acids. Further experiments have confirmed the role of hybrid biosynthesis in forming the fatty acyl and peptide moieties.

Lipoavitides are structurally categorized as lipopeptides, a class of natural products known for their amphiphilic properties[52]. Lipopeptides typically contain a hydrophobic moiety composed of fatty acyl units varying in length and functionalities, connected to a hydrophilic peptide that often contains macrocycles and non-proteinogenic amino acids. These molecules are associated with a myriad of biological functions and have shown tremendous potential in medical applications, as shown by the clinical usage of lipopeptide antibiotics including daptomycin[53], polymyxin[54], and echinocandin[55]. Although nature has developed multiple pathways to synthesize peptidic natural products, known lipopeptides are predominantly synthesized by nonribosomal peptide synthetases, which function as megaenzymatic assembly lines capable of incorporating a wide range of amino acid building blocks. To date, ribosomal synthesis has only been found to produce a small subset of lipopeptides, such as lipolanthines[20, 21], goadvionins[22], and selidamides[56]. In this study, we characterized the lipoavitides as ribosomally synthesized lipopeptides. Structure-activity relationship studies demonstrated the essential role of the fatty acyl moiety in hemolytic activity, suggesting a direct interaction with cellular membranes. Due to the versatility of lipoavitide biosynthesis, the characterization of lipoavitide BGCs may also offer valuable tools in the engineering of lipopeptides for drug development purposes.

Among ribosomal peptides characterized with Avi(Me)Cys moieties, thioviridamide-like thioamitides are most structurally related to lipoavitides[26]. In addition to the presence of hdmHis, thioamitides resemble lipoavitides with an acylated *N*-terminus (Supplementary Table 4). However, the acyl moieties of thioamitides, specifically pyruvyl and lactyl, are formed via dehydration and tautomerization of *N*-terminal serine residues[33], whereas lipoavitides employ fatty acid biosynthetic enzymes to produce fatty acyl-CoAs, which can be transferred to the *N*-terminus. In addition, lipoavitides contain modified residues including D-alanine and glycosylated tyrosine, while thioamitides contain thioamides in the *N*-terminal region. These structural differences of lipoavitides and thioamitides may confer disparate biological functions which deserve further investigation.

Our examination of the *lpv* biosynthetic enzymes showed LpvG as a FabH-like enzyme, which initiates the synthesis of HMP via condensation of isobutyryl-CoA **5** and methylmalonyl-CoA **6**. FabH enzymes possessing similar activities were also reported to produce other fatty acylated natural products. One such example is trehangelin, an oligosaccharide natural product harboring a unique fatty acyl group named angelyl[41]. ThgI, the FabH enzyme encoded in the trehangelin BGC, condenses acetyl-CoA and methylmalonyl-CoA, yielding 2-methyl-acetoacetyl-CoA that is further converted to angelyl-CoA by the ketoreductase ThgK and enoyl-CoA dehydratase ThgH successively. Similar biosynthetic enzymes were also identified in a type II polyketide BGC, producing the tetracycline SF2575 that also comprise an angelyl moiety[57]. With the continued efforts toward mining genomes, we envision that FabH enzymes will be increasingly implicated in the biosynthesis of natural products with unique fatty acyl moieties.

Acyltransferases of the GNAT family are widely distributed in RiPP BGCs. While previously known for *N*-acetylation of LAPs[58], lasso peptides[59], and graspetides[60], the recent discoveries of lipolanthines[20, 21], goadvionins[22], and selidamides[56], have shown their ability to incorporate relatively long chain fatty acyl units into RiPPs. GdvG, the acyltransferase involved in goadvionin biosynthesis, can append a fatty acyl group of 32 carbons to an avionin-containing peptide. However, GdvG only accepts ACP-bound fatty acyl units, which somewhat restricts its application as a biocatalyst. In this study, the lipoavitide acyltransferase LpvE accepts fatty acyl-CoA substrates of varying length and functionality. In addition, LpvE and its homologous enzyme StsE revealed high degree of substrate tolerance. Such promiscuity renders these enzymes as promising biocatalysts, offering valuable opportunities for preparing lipopeptide libraries and engineering new-to-nature lipopeptides.

Methods

Direct cloning of the *lpv* BGC from *Streptomyces* sp. NRRL S-1521.

Direct cloning of *lpv* was achieved by the CAPTURE method[32]. Briefly, the *Streptomyces* sp. NRRL S-1521 was recovered on ISP2 agar medium (malt extract 10 g/L, yeast extract 4 g/L, glucose 4 g/L, agar 20 g/L, pH 7.2–7.4) at 37 °C until colony appears (about 3 d). A single colony was inoculated into 5 ml ISP2 liquid medium and grown at 37 °C with 250 rpm shaking until saturation (3 d). An aliquot of 1 mL cell culture was then transferred into 50 mL fresh ISP2 liquid medium and cultivated for 18–20 h. Cells were then harvested by centrifugation at 3000 × g for 15 min. All the rest of the direct cloning experiments, including the genomic DNA isolation, *in vitro* gRNA transcription, DNA receivers preparation, genomic DNA digestion and Cre-lox *in vivo* recombination, were performed following the CAPTURE method as described elsewhere[32]. The *lpv* BGC sequence was analyzed by SnapGene software (version 4.1.9). Sequences of the DNA templates for the *in vitro* gRNA transcription were shown in Supplementary Table 10 and synthesized by Integrated DNA Technologies. The cloned BGC was then transformed into WM6026 (supplemented with 2,6-diaminopimelic acid at the final concentration 40 μg/mL for growth) and conjugated into *S. lividans* TK24 and *S. albus* J1074 for expression by the protocol described elsewhere[61].

Product analysis from colony extracts.

The exconjugants of *S. lividans* TK24 and *S. albus* J1074 containing *lpv* or the empty pBE45 vector were selected and grew on fresh MS (mannitol 20 g/L, soybean flour 20 g/L, agar 20 g/L) and ISP4 (soluble starch 10 g/L, K_2HPO_4 1 g/L, $MgSO_4{\cdot}7H_2O$ 1 g/L, NaCl 1 g/L, $(NH_4)_2SO_4$ 2 g/L, $CaCO_3$ 2 g/L, $FeSO_4$ 1 mg/L, $MnCl_2$ 1 mg/L, $ZnSO_4$ 1 mg/L, agar 20 g/L) agar medium supplied with apramycin at a final concentration of 50 μg/mL and incubated under 30 °C for 5 d. A pinhead-sized portion of cell mass was picked from the plate, placed in 20 μL methanol and incubated at room temperature for 1 h. An aliquot of 1 μL methanol extract was then mixed with an equal volume of 15 mg/mL 70% aq. acetonitrile solution of α-cyano-4-hydroxycinnamic acid with 0.1% (*v/v*) trifluoroacetic acid (TFA) on a ground steel MALDI target, and the droplet was dried under ambient conditions. Samples were analyzed using a Bruker UltrafleXtreme MALDI-TOF-MS (Bruker, MA) for

reflector positive mode. The mass spectra were visualized by flexAnalysis (version 3.4, Bruker).

Heterologous expression and product isolation.

Freshly obtained exconjugants of *S. albus* J1074 containing wild type or mutated *lpv* were individually grew on MS plates with apramycin and incubated under 30 °C for 5 d. Colonies were verified for producing the lipoavitide by MALDI-TOF-MS by the method described above. The colonies were then scratched individually from the plate by sterile cotton swabs, spread on fresh ISP4 plates, and allowed to grow for another 5 d. The ISP4 plates were then extracted by methanol under 4 °C overnight. Methanol extracts were then harvested by filtration, mixed with an equal volume of water and loaded onto an Agilent Bond Elut C18 Solid Phase Extraction column (bed mass, 10 g; volume, 60 mL; particle size 120 μm), which was pre-equilibrated by 50 mL 5% B (solvent A = 0.1% (*v/v*) TFA in water; solvent B = 0.1% (*v/v*) TFA in acetonitrile). The compounds were then eluted using a step gradient with increasing percentage of solvent B in 150 mL volumes: 5%, 20%, 40%, 60%, 80% and 100% B. The eluted compounds were monitored by MALDI-TOF-MS. Fractions containing the compounds (60% B) were lyophilized to dryness and powders were redissolved into methanol. Semi-preparative HPLC purification was performed using an Agilent 1290 Infinity II Preparative LC System equipped with a Phenomenex Luna C5 column (5 μm, 100 Å, 250 × 10 mm) equilibrated in 5% B. Compounds were eluted by an increase to 100% B over 20 min with a flow rate of 3 mL/min. Collected fractions were analyzed by MALDI-TOF-MS, lyophilized to dryness, and stored at −80 °C until further use. Regarding to the feeding study, fermentation and product purification of **1** followed the same methods except that [$^{13}C_5$]-L-valine and [2-^{13}C]-propionate were supplemented into the ISP4 solid medium to a final concentration of 2 mg/mL.

Antimicrobial activity assay.

Lipoavitide **1** was dissolved in methanol to achieve a concentration of 10 μM. Agar plates were prepared by mixing 200 μL of stationary phase overnight cell culture with 20 mL of melted solid medium (cooled to 42 °C for 5 min). The seeded agar was poured into a sterile 100-mm round dish (VWR) and allowed to solidify at 25 °C for 10 min. Lipoavitide **1** was directly spotted on the solidified agar. Plates were incubated at temperatures according to Supplementary Table 6 for 16 h, and the antimicrobial activity was determined by the presence or absence of inhibition zones.

Hemolytic Assay.

Fresh defibrinated whole bovine blood was purchased from Hemostat Laboratories. Whole blood was washed three times, diluted to a final concentration of 1:25 (*v/v*) using phosphate-buffered saline, and then split into 50 μL aliquots in individual 1.7 mL Eppendorf tubes. Next, stock solutions of concentrations at 1 mM, 200 μM, and 40 μM were prepared for **1**, **2**, **3**, and **4** in methanol. Aliquots of 2.5 μL each stock solution were mixed with the blood, yielding final concentrations at 50 μM, 10 μM, and 2 μM. An equal volume of methanol and Triton X-100 were used as negative and positive controls. The mixtures were then incubated in Eppendorf ThermoMixer C with a heated lid for 18 h at 37 °C. After incubation, the

samples were centrifuged at 500 × g for 10 min, and the supernatants were measured for hemoglobin absorbance at 410 nm on a ThermoFisher NanoDrop spectrophotometer. Each measurement was performed in biological triplicate, and the percentage of hemolysis was calculated using the following equation:

$$\%\,\text{Hemolysis} = \frac{\text{Abs of test sample} - \text{Abs of methanol}}{\text{Abs of Triton X} - 100 - \text{Abs of methanol}} \quad (1)$$

Expression and purification of recombinant proteins.

Genes were PCR amplified from the genomic DNA of *Streptomyces* sp. NRRL S-1521 and *Streptomyces* sp. NRRL S-920 using primers listed in Supplementary Table 10, which were cloned into pET28a to produce proteins with *N*-terminal His-tag using NEBuilder HiFi DNA assembly master mix. Plasmids were transformed into *E. coli* BL21(DE3) for heterologous expression. Cells were grown overnight on LB agar plates (5 g/L yeast extract, 10 g/L tryptone, 10 g/L NaCl, 20 g/L agar) containing 50 μg/mL kanamycin at 37 °C. Single colonies were used to inoculate 5 mL of LB liquid medium containing 50 μg/mL kanamycin and grown at 37 °C for 14–18 h. This culture was used to inoculate 500 mL Terrific Broth (TB; 24 g/L yeast extract, 20 g/L tryptone, 0.4% (*v/v*) glycerol, 17 mM KH_2PO_4, and 72 mM K_2HPO_4) containing 50 μg/mL kanamycin and grown to an optical density at 600 nm (OD_{600}) of 0.6–0.8. Protein expression was induced by addition of isopropyl *β*-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM for 18 h at 18 °C. Cells were harvested by centrifugation and stored at −80 °C until further use.

Cell pellets were resuspended in 20 mL buffer A (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 7.5) and treated with 20 mg lysozyme and 5 μL benzolase at 4 °C for 1 h. Cells were then lysed on ice by sonication using a Vibra Cell sonicator with the following settings: 60% amplitude, 5 min total sonication time, alternating between 2 s on/5 s off-pulse. Lysates were centrifuged at 4 °C for 30 min at 30,000 × g, and supernatants were filtered through a syringe filter (0.45 μm). Ni-NTA purification was performed using a 5 mL Ni-NTA HisTrap column (GE Healthcare). After loading, the column was washed with 50 mL buffer A and eluted using a step gradient with an increasing percentage of buffer B (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, pH 7.5) each in 15 mL volumes: 20%, 40%, 60%, 80% and 100% B. Fractions containing the recombinant protein were analyzed on SDS-PAGE and concentrated with Amicon centrifugal filter units (at suitable molecular weight cut-offs) to around 2 mL. Concentrated proteins were loaded to Cytiva PD-10 columns and eluted by storage buffer (50 mM Tris-HCl, 5 % (*v/v*) glycerol, pH 7.5). Proteins were flash frozen by liquid nitrogen and stored at −80 °C until further use, except for LpvS which was stored at 4 °C.

Characterization of the *in vitro* activity of StsG.

The reaction of StsG contains 50 mM Tris-HCl (pH 7.5), 2 mM tris(2-carboxyethyl)phosphine hydrochloride, 1 mM isobutyryl-CoA **5**, 1 mM methylmalonyl-CoA **6**, and 50 μM StsG in a total volume of 20 μL. Boiled StsG was used as negative control. Either 2 mM or 0.5 mM **5** were used to set up reactions with substrate molar ratios of **5**:**6**

at 2:1 or 1:2, respectively. When using the methylmalonyl-CoA epimerase, the enzyme was added to a final concentration of 50 μM. All reactions were performed at 30 °C for 10 min, quenched by adding 2 μL formic acid, stored at −20 °C, and diluted 5-fold by water right before analysis. A 5 μL aliquot of each sample was analyzed by HPLC using an Agilent 1260 system equipped with Phenomenex Kinetex SB-C18 column (5 μm, 100 Å, 250 × 10 mm) under the following conditions: mobile phase (A = 10 mM ammonium acetate, B = methanol), 5% to 100 % B over 15 min, 100% to 5% B for 1 min, 5% B for 15 min; flow rate 1.0 mL/min. The UV absorbance was monitored at 260 nm.

Characterization of the *in vitro* activity of LpvV.

The two-enzyme reaction of StsG/LpvV was set up by adding 50 μM LpvV, along with either 1 mM NADH or 10 mM NADPH, to the StsG reaction in a total volume of 20 μL. Boiled LpvV was used as negative control. The reaction was performed at 30 °C for 10 min and quenched by adding 2 μL formic acid. Subsequent sample preparation and HPLC analysis followed the same methods as described for the StsG reaction.

Characterization of the *in vitro* activity of LpvS.

The *in vitro* electron transfer to LpvS was achieved by either (a) the peroxide shunt pathway or (b) the spinach reductase and ferredoxin. The reaction condition (a) contains 50 mM Tris-HCl (pH 7.5), 100 μM 2,4-dimethylpentanoyl-CoA **9**, 1 mM H_2O_2, and 1 μM LpvS in a total volume of 50 μL. The reaction condition (b) contains 50 mM Tris-HCl (pH 7.5), 100 μM **9**, 3.5 μM spinach ferredoxin, 0.1 U/mL spinach ferredoxin reductase, 1 mM NADPH, and 1 μM LpvS in a total volume of 50 μL. Boiled LpvS was used as negative control. The reactions were performed at room temperature for 20 h. Upon completion, the reaction mixtures were frozen at −80 °C for at least 1 h, dissolved, and centrifuged to remove precipitates before HPLC and LC-MS analysis. A total of 40 μL reaction mixture was analyzed by HPLC using the same method as described for StsG. The leftover 10 μL reaction mixture was then diluted 5-times by water and analyzed by the Thermo Scientific Liquid Chromatography Mass Spectrometry (LC-MS) equipped with a Hypersil GOLD™ VANQUISH™ PFP UHPLC column (1.9 μm, 2.1 mm × 100 mm), under the following condition: mobile phase (A = 0.1 % (*v/v*) formic acid in water, B = 0.1% (*v/v*) formic acid in acetonitrile); 5% B for 1.0 min, 5% to 100 % B over 5.0 min, 100% B for 1 min, 100% to 5% B over 0.1 min, 5% B for 0.5 min; flow rate 0.4 mL/min. The Full MS-SIM was operated with the following parameters: 70,000 resolution, scan range 50 to 750 *m/z*, AGC target T = 3e6, maximum IT 200 ms, and positive polarity. Data analysis was then conducted using the Qual browser application within Xcalibur software (version 4.1.31.9, ThermoFisher Scientific).

Characterization of the *in vitro* activity of acyltransferases.

The three-enzyme reaction of StsG, LpvV, and LpvE was set up by adding 50 μM LpvE and 0.5 mM peptide substrate (**2** or **3**) to the described two-enzyme reaction of StsG and LpvV in a total volume of 20 μL. Boiled LpvE was used as negative control. The reaction was performed at room temperature for 20 h and quenched by adding 2 μL formic acid. A total of 10 μL of the reaction mixture was then analyzed by HPLC using an Agilent 1260

system equipped with Phenomenex Kinetex SB-C18 column (5 μm, 100 Å, 250 × 10 mm) under the following conditions: mobile phase (A = 0.1 % (*v/v*) TFA in water, B = 0.1% (*v/v*) TFA in acetonitrile); 30% to 70 % B over 15 min, 70% to 100% B over 3 min, 100% B for 5 min, 100% to 5% B over 1 min, 5% B for 5 min; flow rate 1.0 mL/min. The UV absorbance was monitored at 280 nm. In addition, the reaction mixture was also analyzed by MALDI-TOF-MS after desalted by Millipore C18 ZipTip. The reaction of LpvS and LpvE contained 50 mM Tris-HCl (pH 7.5), 100 μM 2,4-dimethylpentanoyl-CoA **9**, 3.5 μM spinach ferredoxin, 0.1 U/mL spinach ferredoxin reductase, 1 mM NADPH, 100 μM **2**, 1 μM LpvS, and 1 μM LpvE in a total volume of 50 μL. Boiled LpvE was used as negative control. The reaction was then performed and analyzed by MALDI-TOF-MS as described above. The reactions of LpvE using other acyl-CoA and peptide substrates were assayed in 50 mM Tris-HCl (pH 7.5), 0.5 mM acyl-CoA, 0.5 mM peptide substrate, and 50 μM LpvE. The reactions were performed at room temperature for up to 72 h, quenched by adding 2 μL formic acid, desalted by C18 ZipTip, and analyzed by MALDI-TOF-MS. The reactions of StsE were performed and analyzed using the same method as described for LpvE.

Gene disruption and LpvA point mutation by RecET recombineering.

The genetic manipulation of *lpv* was achieved using the RecET recombination system. Donor DNA fragments containing an ampicillin resistance marker flanked by *Nsi*I restriction sites were amplified by PCR with primers listed in Supplementary Table 10. The amplified DNA fragment contains homologous arms to the target site that enables recombination. When introducing point mutations to *lpvA*, sites of mutagenesis were also introduced into the donor DNA by primers. The donor DNA and *lpv* were co-transformed into *E. coli* GB05-dir that contains an inducible RecET recombination system as described previously[37]. Colonies were picked and grown overnight in 5 mL LB medium supplemented with 50 μg/mL apramycin at 37 °C. The plasmid DNA was purified from the cultures using Qiaprep Spin Miniprep Kit (Qiagen), digested by appropriate restriction enzymes, and analyzed by agarose gel electrophoresis. Correct plasmids were digested by NsiI-HF and re-circularized by T4 DNA ligase to remove the ampicillin resistance marker. Each digestion reaction contains 100 ng plasmid, 0.5 μL NsiI-HF (10 U), and 2 μL rCutSmart Buffer within a total volume of 20 μL. The reaction was performed under 37 °C for 1 h, followed by 80 °C for 20 min to inactivate the enzyme. The reaction was then cooled to room temperature, supplemented by 0.2 μL (400 U) T4 DNA ligase, 0.8 μL 25 mM ATP, and 1 μL 200 mM DTT, and incubated at 20 °C for another 1 h. The reaction mixture was then transformed into *E. coli* NEB10β. Plasmids were purified and analyzed by restriction enzyme digestion as described above. Correct plasmids were transformed into *E. coli* WM6026 and conjugated into *S. albus* J1074 for expression.

Generation of acyltransferase phylogenetic tree.

The acyltransferases for lipoavitide, lipolanthine, and goadvionin biosynthesis, along with other characterized members in the GNAT family, were aligned by Mega Software (version 11.0.10, Pennsylvania State University) using the ClustalW method[62]. A maximum-likelihood phylogenetic tree was then generated based on this alignment with the Jones-Taylor-Thornton (JTT) model using 1000 bootstrap replications. The tree was visualized using the Interactive Tree of Life (iTOL) website (https://itol.embl.de/).

Structure simulation molecular docking of LpvE.

The structures of **2** and **10** were optimized with semi-empirical quantum mechanics MOPAC[63] and energy minimization. AlphaFold2 (DeepMind) was used for predicting the three-dimensional structure of LpvE, with each of the five trained model parameters from CASP14 models[44]. The multiple sequence alignment generation, AlphaFold predictions, and Amber structural relaxation were run on local server with GPU. A full database (updated to 2021-09-17) was applied for the structure predictions. For molecular docking, a grid box of 5 Å was created around the active site within LpvE. The docking study of **2** was performed after obtaining the LpvE-**10** docking complex. Autodock (version 4.2.6, Scripps Research) plug-in[45] within YASARA (version 20.4.24, YASARA Biosciences)[64, 65] was used to perform molecular docking calculations with a fixed protein backbone. The protein residues were treated with AMBER ff99 force field[66], and the ligand atoms were treated using GAFF[67, 68] with AM1-BCC partial charges[69]. A total of 100 docking runs were calculated, and the obtained docking poses were clustered using an RMSD cutoff of 0.5 Å within the YASARA dock_run macro file. Chimera (version 1.16, University of California San Francisco)[70] was used to visualize and analyze the proteins. Cavity volumes were calculated by CAVER Analyst (version 2.0, Masaryk University) with default parameters[71].

Statistics and reproducibility.

The bioactivity assay was conducted with three biological replicates in addition to 2–3 technical replicates, yielding consistent results. The statistical details are provided in the figure legend. Protein expression and purification experiments were successfully conducted 2–3 times for each enzyme characterized *in vitro*. All the other experiments were performed at least three times with independent samples. All experiments were reproduced with similar results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We wish to thank Justine Arrington from the Roy J. Carver Biotechnology Center for HR-MS/MS assistance. This work was supported by a grant from the National Institutes of Health (AI144967 to D.A.M. and H.Z.). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Data availability

We declare that all data supporting the findings of this study are presented in the main text and supplementary information. NCBI (https://www.ncbi.nlm.nih.gov/), PDB (https://www.rcsb.org/), and Uniprot (https://www.uniprot.org/) accessions are referenced in the supplementary information, and these accessions are publicly accessible on the respective NCBI, RCSB PDB, and Uniprot websites. Source data are provided with this paper.

References

1. Arnison PG et al. Ribosomally synthesized and post-translationally modified peptide natural products: Overview and recommendations for a universal nomenclature. Nat Prod Rep 30, 108–160 (2013). [PubMed: 23165928]
2. Montalban-Lopez M et al. New developments in RiPP discovery, enzymology and engineering. Nat Prod Rep 38, 130–239 (2021). [PubMed: 32935693]
3. Cao L, Do T & Link AJ Mechanisms of action of ribosomally synthesized and posttranslationally modified peptides (RiPPs). J Ind Microbiol Biotechnol 48, kuab005 (2021).
4. Ongpipattanakul C et al. Mechanism of action of ribosomally synthesized and post-translationally modified peptides. Chem Rev 122, 14722–14814 (2022). [PubMed: 36049139]
5. Melby JO, Nard NJ & Mitchell DA Thiazole/oxazole-modified microcins: Complex natural products from ribosomal templates. Curr Opin Chem Biol 15, 369–378 (2011). [PubMed: 21429787]
6. Franz L, Kazmaier U, Truman AW & Koehnke J Bottromycins-biosynthesis, synthesis and activity. Nat Prod Rep 38, 1659–1683 (2021). [PubMed: 33621290]
7. Vinogradov AA & Suga H Introduction to thiopeptides: Biological activity, biosynthesis, and strategies for functional reprogramming. Cell Chem Biol 27, 1032–1051 (2020). [PubMed: 32698017]
8. McIntosh JA, Donia MS & Schmidt EW Insights into heterocyclization from two highly similar enzymes. J Am Chem Soc 132, 4089–4091 (2010). [PubMed: 20210311]
9. Burkhart BJ, Schwalen CJ, Mann G, Naismith JH & Mitchell DA YcaO-dependent posttranslational amide activation: Biosynthesis, structure, and function. Chem Rev 117, 5389–5456 (2017). [PubMed: 28256131]
10. Norris GE & Patchett ML The glycocins: In a class of their own. Curr Opin Struc Biol 40, 112–119 (2016).
11. Saad H et al. Nocathioamides, uncovered by a tunable metabologenomic approach, define a novel class of chimeric lanthipeptides. Angew Chem Int Ed 60, 16472–16479 (2021).
12. Medema MH, Cimermancic P, Sali A, Takano E & Fischbach MA A systematic computational analysis of biosynthetic gene cluster evolution: Lessons for engineering biosynthesis. Plos Comput Biol 10, e1004016 (2014). [PubMed: 25474254]
13. Robey MT, Caesar LK, Drott MT, Keller NP & Kelleher NL An interpreted atlas of biosynthetic gene clusters from 1,000 fungal genomes. Proc Natl Acad Sci USA 118, e2020230118 (2021). [PubMed: 33941694]
14. Gavriilidou A et al. Compendium of specialized metabolite biosynthetic diversity encoded in bacterial genomes. Nat Microbiol 7, 726–735 (2022). [PubMed: 35505244]
15. Walsh CT, Brien RVO & Khosla C Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polyketide scaffolds. Angew Chem Int Ed 52, 7098–7124 (2013).
16. Just-Baringo X, Albericio F & Alvarez M Thiopeptide antibiotics: Retrospective and recent advances. Mar Drugs 12, 317–351 (2014). [PubMed: 24445304]
17. Noike M et al. A peptide ligase and the ribosome cooperate to synthesize the peptide pheganomycin. Nat Chem Biol 11, 71–76 (2015). [PubMed: 25402768]
18. Ortiz-Lopez FJ et al. Cacaoidin, first member of the new lanthidin ripp family. Angew Chem Int Ed 59, 12654–12658 (2020).
19. Jordan PA & Moore BS Biosynthetic pathway connects cryptic ribosomally synthesized posttranslationally modified peptide genes with pyrroloquinoline alkaloids. Cell Chem Biol 23, 1504–1514 (2016). [PubMed: 27866908]
20. Wiebach V et al. The anti-staphylococcal lipolanthines are ribosomally synthesized lipopeptides. Nat Chem Biol 14, 652–654 (2018). [PubMed: 29915235]
21. Wiebach V et al. An amphipathic alpha-helix guides maturation of the ribosomally-synthesized lipolanthines. Angew Chem Int Ed 59, 16777–16785 (2020).
22. Kozakai R et al. Acyltransferase that catalyses the condensation of polyketide and peptide moieties of goadvionin hybrid lipopeptides. Nat Chem 12, 869–877 (2020). [PubMed: 32719482]

23. Doroghazi JR et al. A roadmap for natural product discovery based on large-scale genomics and metabolomics. Nat Chem Biol 10, 963–968 (2014). [PubMed: 25262415]
24. Blin K et al. AntiSMASH 6.0: Improving cluster detection and comparison capabilities. Nucleic Acids Res 49, W29–W35 (2021). [PubMed: 33978755]
25. Grant-Mackie ES, Williams ET, Harris PWR & Brimble MA Aminovinyl cysteine containing peptides: A unique motif that imparts key biological activity. Jacs Au 1, 1527–1540 (2021). [PubMed: 34723257]
26. Eyles TH, Vior NM, Lacret R & Truman AW Understanding thioamitide biosynthesis using pathway engineering and untargeted metabolomics. Chem Sci 12, 7138–7150 (2021). [PubMed: 34123341]
27. Xu M et al. Functional genome mining reveals a class V lanthipeptide containing a d-amino acid introduced by an F420H2-dependent reductase. Angew Chem Int Ed 59, 18029–18035 (2020).
28. Kloosterman AM et al. Expansion of RiPP biosynthetic space through integration of pan-genomics and machine learning uncovers a novel class of lantibiotics. Plos Biol 18, e3001026 (2020). [PubMed: 33351797]
29. Schujman GE & de Mendoza D Regulation of type II fatty acid synthase in Gram-positive bacteria. Curr Opin Microbiol 11, 148–152 (2008). [PubMed: 18372209]
30. Hu L et al. Characterization of histidine functionalization and its timing in the biosynthesis of ribosomally synthesized and posttranslationally modified thioamitides. J Am Chem Soc 144, 4431–4438 (2022). [PubMed: 35230829]
31. Sikandar A, Lopatniuk M, Luzhetskyy A, Muller R & Koehnke J Total in vitro biosynthesis of the thioamitide thioholgamide andinvestigation of the pathway. J Am Chem Soc 144, 5136–5144 (2022). [PubMed: 35263083]
32. Enghiad B et al. Cas12a-assisted precise targeted cloning using in vivo Cre-lox recombination. Nat Commun 12, 1171 (2021). [PubMed: 33608525]
33. Frattaruolo L, Lacret R, Cappello AR & Truman AW A genomics-based approach identifies a thioviridamide-like compound with selective anticancer activity. ACS Chem Biol 12, 2815–2822 (2017). [PubMed: 28968491]
34. Bhushan R & Bruckner H Use of Marfey's reagent and analogs for chiral amino acid analysis: Assessment and applications to natural products and biological systems. J Chromatogr B 879, 3148–3161 (2011).
35. Nilsson J et al. Enrichment of glycopeptides for glycan structure and attachment site identification. Nat Methods 6, 809–811 (2009). [PubMed: 19838169]
36. Bender CL, Alarcon-Chaidez F & Gross DC Pseudomonas syringae phytotoxins: Mode of action, regulation, and biosynthesis by peptide and polyketide synthetases. Microbiol Mol Biol R 63, 266–292 (1999).
37. Fu J et al. Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. Nat Biotechnol 30, 440–446 (2012). [PubMed: 22544021]
38. Tang L, Zhang YX & Hutchinson CR Amino acid catabolism and antibiotic synthesis: Valine is a source of precursors for macrolide biosynthesis in Streptomyces ambofaciens and Streptomyces fradiae. J Bacteriol 176, 6107–6119 (1994). [PubMed: 7928973]
39. Dayem LC et al. Metabolic engineering of a methylmalonyl-CoA mutase-epimerase pathway for complex polyketide biosynthesis in Escherichia coli. Biochemistry 41, 5193–5201 (2002). [PubMed: 11955068]
40. Okamura E, Tomita T, Sawa R, Nishiyama M & Kuzuyama T Unprecedented acetoacetyl-coenzyme A synthesizing enzyme of the thiolase superfamily involved in the mevalonate pathway. Proc Natl Acad Sci USA 107, 11265–11270 (2010). [PubMed: 20534558]
41. Inahashi Y et al. Biosynthesis of trehangelin in polymorphospora rubra k07–0510: Identification of metabolic pathway to angelyl-CoA. Chembiochem 17, 1442–1447 (2016). [PubMed: 27311629]
42. Bobik TA & Rasche ME Identification of the human methylmalonyl-CoA racemase gene based on the analysis of prokaryotic gene arrangements-implications for decoding the human genome. J Biol Chem 276, 37194–37198 (2001). [PubMed: 11481338]
43. Peter DM, Vogeli B, Cortina NS & Erb TJ A chemo-enzymatic road map to the synthesis of CoA esters. Molecules 21, 517 (2016). [PubMed: 27104508]

44. Jumper J et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). [PubMed: 34265844]
45. Morris GM, Goodsell DS, Huey R & Olson AJ Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. J Comput Aided Mol Des 10, 293–304 (1996). [PubMed: 8877701]
46. Kuhn ML, Majorek KA, Minor W & Anderson WF Broad-substrate screen as a tool to identify substrates for bacterial Gcn5-related N-acetyltransferases with unknown substrate specificity. Protein Sci 22, 222–230 (2013). [PubMed: 23184347]
47. Ramesh S et al. Bioinformatics-guided expansion and discovery of graspetides. Acs Chemical Biology 16, 2787–2797 (2021). [PubMed: 34766760]
48. Imai Y et al. A new antibiotic selectively kills Gram-negative pathogens. Nature 576, 459–464 (2019). [PubMed: 31747680]
49. Park HB, Perez CE, Barber KW, Rinehart J & Crawford JM Genome mining unearths a hybrid nonribosomal peptide synthetase-like-pteridine synthase biosynthetic gene cluster. Elife 6, e25229 (2017). [PubMed: 28431213]
50. Schor R, Schotte C, Wibberg D, Kalinowski J & Cox RJ Three previously unrecognised classes of biosynthetic enzymes revealed during the production of xenovulene A. Nat Commun 9, 1963 (2018). [PubMed: 29773797]
51. Yee DA et al. Genome mining of alkaloidal terpenoids from a hybrid terpene and nonribosomal peptide biosynthetic pathway. J Am Chem Soc 142, 710–714 (2020). [PubMed: 31885262]
52. Gotze S & Stallforth P Structure elucidation of bacterial nonribosomal lipopeptides. Org Biomol Chem 18, 1710–1727 (2020). [PubMed: 32052002]
53. Robbel L & Marahiel MA Daptomycin, a bacterial lipopeptide synthesized by a nonribosomal machinery. J Biol Chem 285, 27501–27508 (2010). [PubMed: 20522545]
54. Poirel L, Jayol A & Nordmann P Polymyxins: Antibacterial activity, susceptibility testing, and resistance mechanisms encoded by plasmids or chromosomes. Clin Microbiol Rev 30, 557–596 (2017). [PubMed: 28275006]
55. Huttel W Echinocandins: Structural diversity, biosynthesis, and development of antimycotics. Appl Microbiol Biotechnol 105, 55–66 (2021). [PubMed: 33270153]
56. Hubrich F et al. Ribosomally derived lipopeptides containing distinct fatty acyl moieties. Proc Natl Acad Sci USA 119, e2113120119 (2022). [PubMed: 35027450]
57. Pickens LB et al. Biochemical analysis of the biosynthetic pathway of an anticancer tetracycline SF2575. J Am Chem Soc 131, 17677–17689 (2009). [PubMed: 19908837]
58. Ozaki T et al. Dissection of goadsporin biosynthesis by in vitro reconstitution leading to designer analogues expressed in vivo. Nat Commun 8, 14207 (2017). [PubMed: 28165449]
59. Zong C, Cheung-Lee WL, Elashal HE, Raj M & Link AJ Albusnodin: An acetylated lasso peptide from Streptomyces albus. Chem Commun (Camb) 54, 1339–1342 (2018). [PubMed: 29350227]
60. Ziemert N, Ishida K, Liaimer A, Hertweck C & Dittmann E Ribosomal synthesis of tricyclic depsipeptides in bloom-forming cyanobacteria. Angew Chem Int Ed 47, 7756–7759 (2008).
61. Kieser T, Bibb MJ, Buttner MJ, Chater KF & Hopwood DA Practical streptomyces genetics (John Innes Foundation, 2000).
62. Tamura K, Stecher G & Kumar S MEGA11: Molecular evolutionary genetics analysis version 11. Mol Biol Evol 38, 3022–3027 (2021). [PubMed: 33892491]
63. Stewart JJP MOPAC2016, Stewart Computational Chemistry, Colorado Springs, CO, USA.
64. Krieger E & Vriend G YASARA View-molecular graphics for all devices-from smartphones to workstations. Bioinformatics 30, 2981–2982 (2014). [PubMed: 24996895]
65. Trott O & Olson AJ Software news and update autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31, 455–461 (2010). [PubMed: 19499576]
66. Wang JM, Cieplak P & Kollman PA How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem 21, 1049–1074 (2000).

67. Duan Y et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24, 1999–2012 (2003). [PubMed: 14531054]

68. Wang JM, Wolf RM, Caldwell JW, Kollman PA & Case DA Development and testing of a general amber force field. J Comput Chem 25, 1157–1174 (2004). [PubMed: 15116359]

69. Jakalian A, Jack DB & Bayly CI Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. J Comput Chem 23, 1623–1641 (2002). [PubMed: 12395429]

70. Pettersen EF et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci 30, 70–82 (2021). [PubMed: 32881101]

71. Jurcik A et al. CAVER Analyst 2.0: Analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. Bioinformatics 34, 3586–3588 (2018). [PubMed: 29741570]

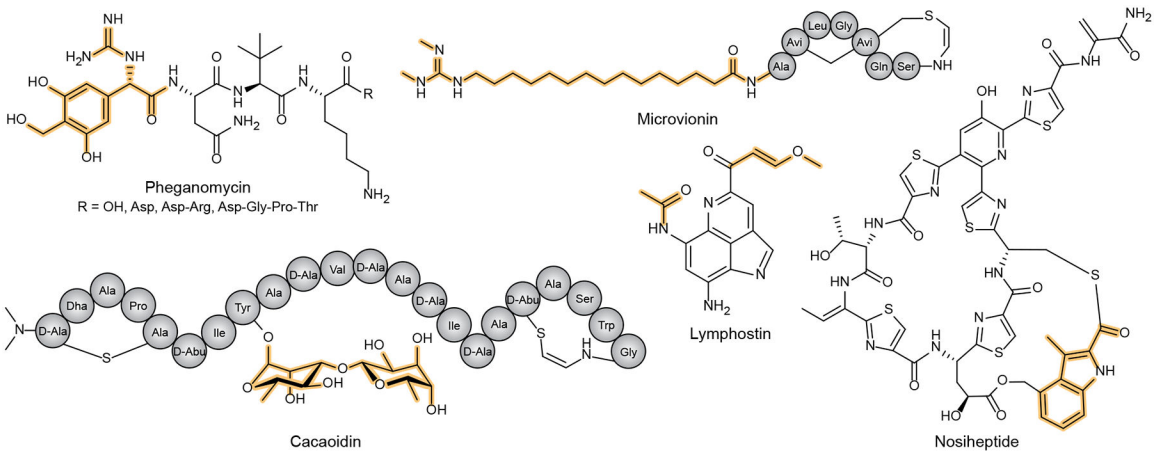

Fig. 1 |. Representative RiPPs produced by RiPP/non-RiPP hybrid BGCs.

Moieties synthesized by the non-RiPP pathways are highlighted in yellow. Microvionin and nosiheptide are listed as class representatives for lipolanthines and *e*-series thiopeptides, respectively.

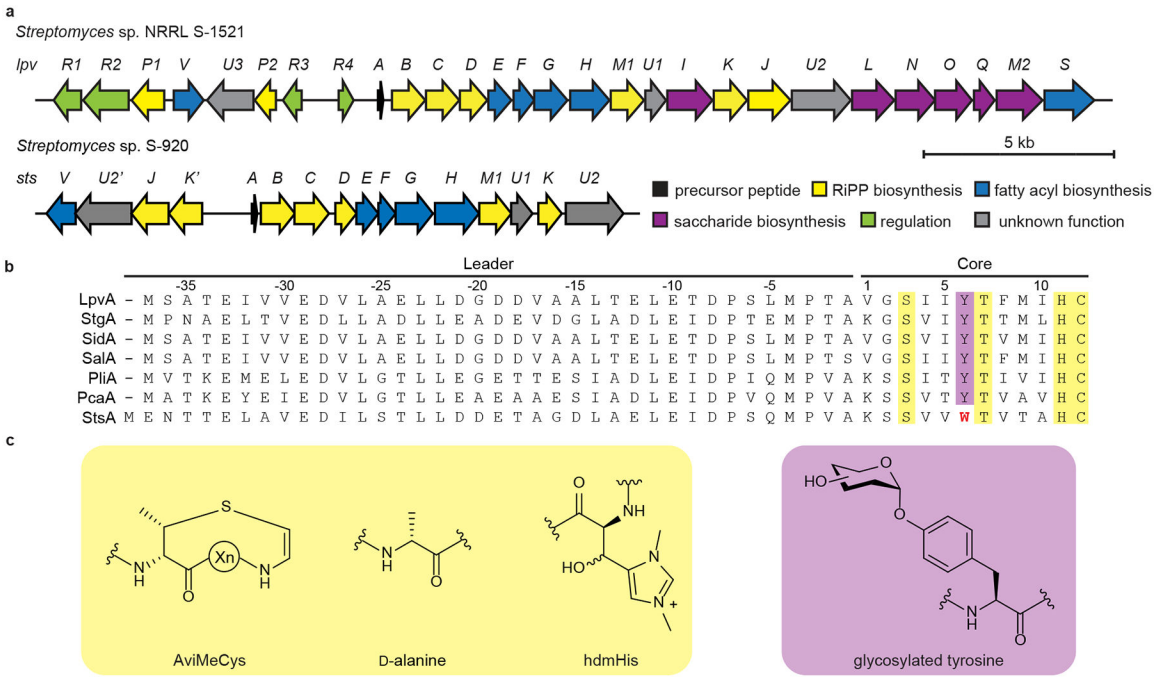

Fig. 2 |. Representative BGCs identified in bioinformatics study.

a, Gene organization of the *lpv* and *sts* BGC. Genes in *sts* are named according to their orthologous genes in *lpv*. Names for genes sharing the same orthology are distinguished by apostrophes (*e.g. stsK* and *stsK'*). **b**, Alignment of LpvA, StsA, and other putative precursor peptides from *lpv* orthologous BGCs. Residues predicted to be modified by RiPP and saccharide biosynthetic enzymes are highlighted in yellow and purple, respectively. The tryptophan residue of StsA aligned to the putative site of glycosylation is highlighted in red. **c**, Proposed PTMs in the final product.

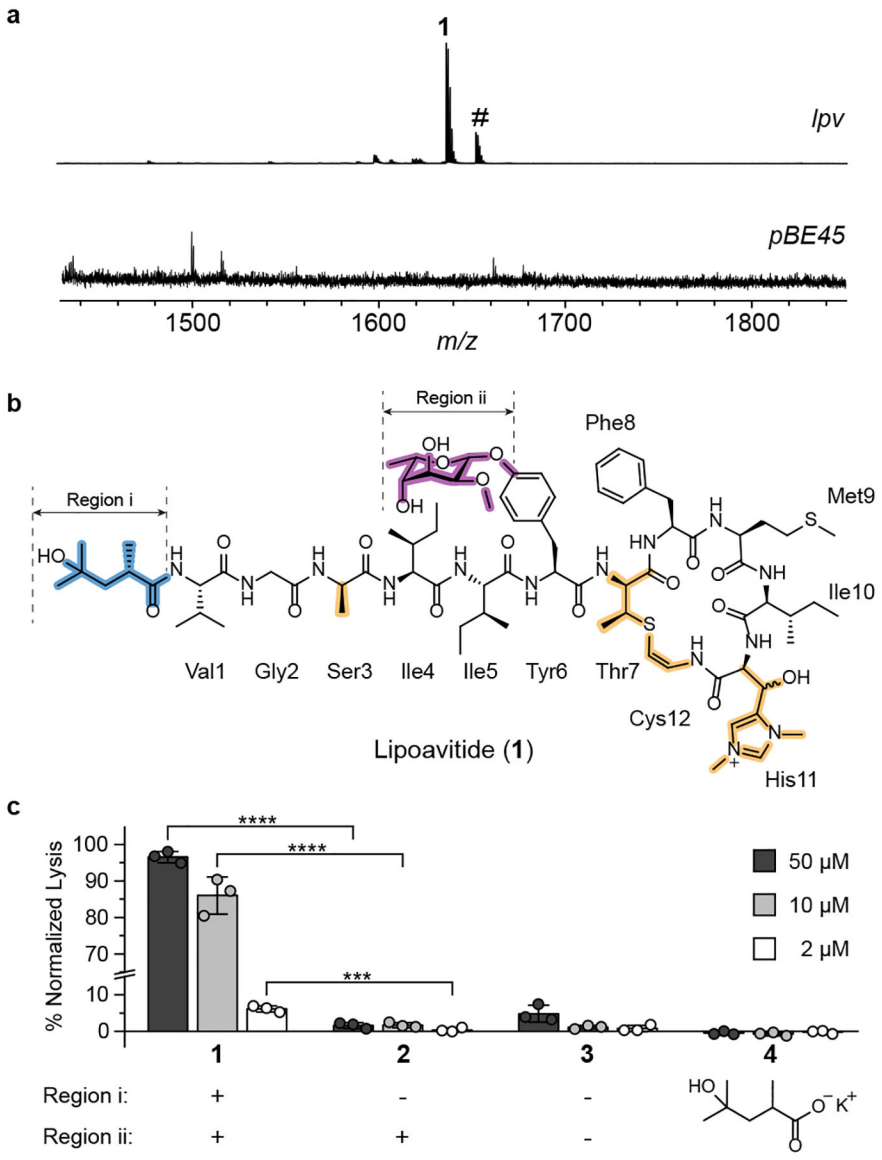

Fig. 3 |. Heterologous expression and product characterization of *lpv*.

a, MALDI-TOF mass spectra of *S. albus* J1074 containing *lpv* (**1**, *m/z* 1636) and the pBE45 empty vector. The signal denoted by '#' (*m/z* 1652) is attributed to spontaneous oxidation to the methionine sulfur. **b**, Proposed structure of lipoavitide **1**. NMR spectra of **1** and assigned signals are shown in Supplementary Fig. 9–15 and Supplementary Table 5, respectively. **c**, Hemolytic activity of **1–4**. The presence and absence of Region i and Region ii in **1-3** are indicated by "'+" and"-". All data represent the mean of n=3 biologically independent samples and error bars show standard deviation. Statistical significance of the difference between results of **1** and **2** are shown (two-tailed t-test: ****$P < 0.0001$; ***$P < 0.001$). P values: 50 μM, 7.9×10^{-8}; 10 μM, 9.0×10^{-6}; 2 μM, 6.2×10^{-4}. **Source data**.

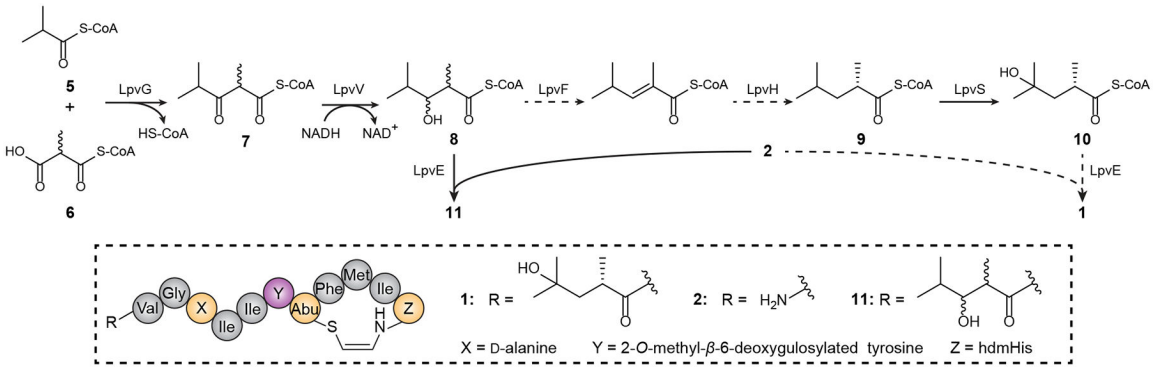

Fig. 4 |. Proposed biosynthetic pathway for the HMP moiety.

Experimental validated and hypothetical steps are shown by solid and dashed arrows, respectively. LpvG, 3-ketoacyl-acyl carrier protein synthase III (FabH); LpvV, short-chain dehydrogenase/reductase; LpvF, dehydratase; LpvH, enoylreductase; LpvS, P450 monooxygenase; LpvE, acyltranferase; Abu, aminobutyric acid; hdmHis, *β*-hydroxy-*N,N*-dimethyl-L-histidine.

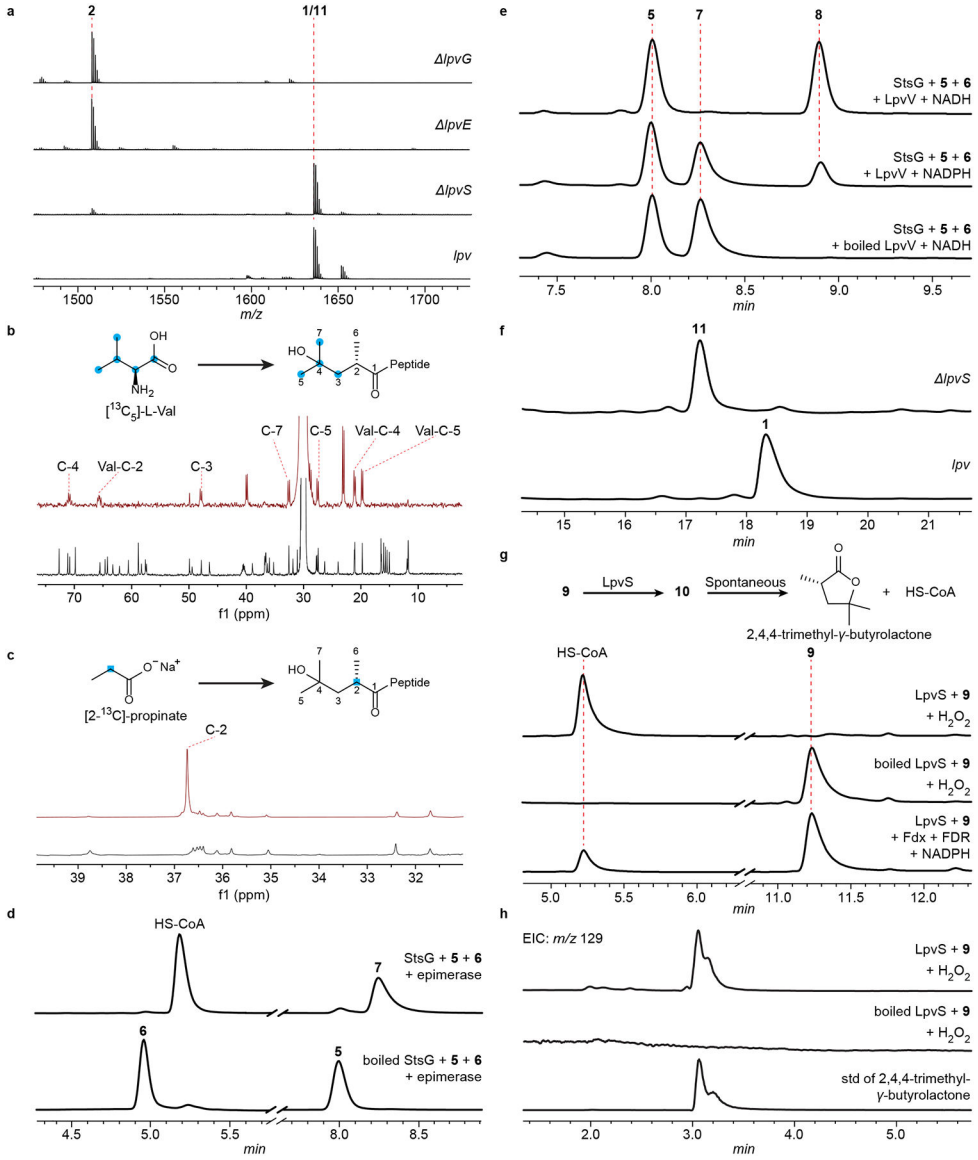

Fig. 5 |. Biosynthesis of the HMP moiety.

a, MALDI-TOF mass spectra of methanol extracts of *S. albus* J1074 containing *lpv* and disruption mutants thereof. **b**, Isotopic-labeling of **1** by [$^{13}C_5$]-L-valine. Top, ^{13}C NMR spectrum of isotopically labeled **1**. Bottom, ^{13}C NMR spectrum of unlabeled **1**. Enriched ^{13}C signals are indicated in the spectra and assigned to the positions in HMP. Of note, Val1 of **1** is also labeled by [$^{13}C_5$]-L-valine and signals are assigned in the spectra accordingly. Full spectra are shown in Supplementary Fig. 35. **c**, Isotopic-labeling of **1** by [2-^{13}C]-propionate. Top, ^{13}C NMR spectrum of isotopically labeled **1**. Bottom, ^{13}C NMR spectrum of unlabeled **1**. Enriched ^{13}C signals are indicated in the spectrum and assigned to the positions in HMP. Full spectra are shown in Supplementary Fig. 35. **d**, HPLC traces (260 nm) of the StsG reaction using **5** and **6** as substrates. **e**, HPLC traces (260 nm) of the StsG/LpvV two-enzyme reaction. **f**, HPLC traces (220 nm) of fractionated methanol extracts of *S. albus* J1074 containing *lpv* and *lpvS*-disrupted constructs. **g**, HPLC traces (260 nm) of

the LpvS reaction using **9** as the substrate. Fdx, spinach ferredoxin; FDR, spinach ferredoxin reductase. **h**, LC-MS analysis (extracted ion chromatogram, EIC) of the LpvS reaction using **9** as the substrate.

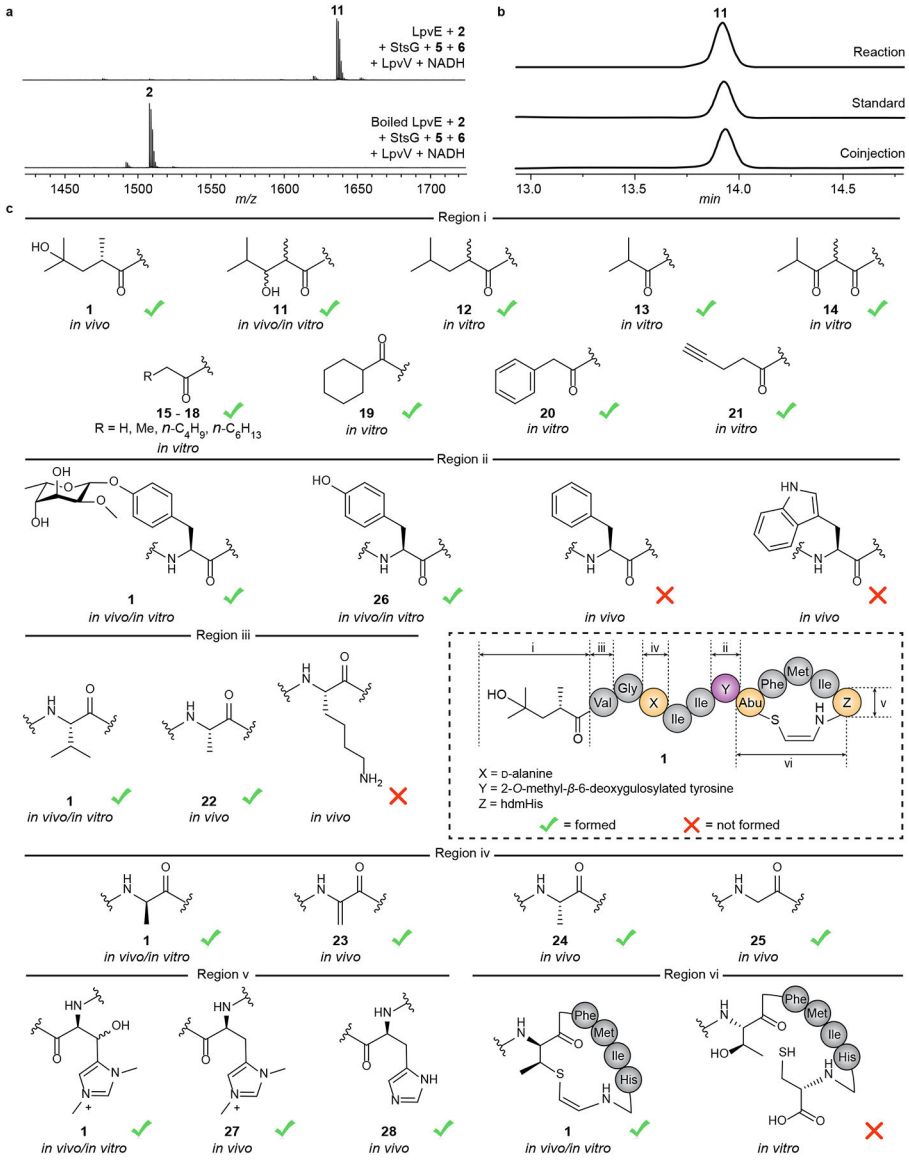

Fig. 6 |. Characterization of the acyltransferase LpvE.

a, MALDI-TOF mass spectra of the StsG/LpvV/LpvE reaction. **b**, HPLC analysis of the StsG/LpvV/LpvE three-enzyme reaction. Shown are chromatograms (280 nm) for the reaction product (top), standard isolated from *S. albus* J1074 containing the *lpvS*-disrupted BGC (middle), and a 1:1 co-injection of these two samples (bottom). **c**, Summary of structural variants of **1** generated by *in vivo* and *in vitro* approaches. The structure of **1** is shown in the dashed box for reference.