BIOMOLECULAR NMR STUDIES OF THE STRUCTURE AND BIOSYNTHESIS OF RIPP NATURAL PRODUCTS

BY

ADAM DICAPRIO

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois Urbana-Champaign, 2021

Urbana, Illinois

Doctoral Committee:

Professor Douglas A. Mitchell, Chair
Professor Paul J. Hergenrother
Professor William W. Metcalf
Professor Huimin Zhao

ABSTRACT

Natural products have been a rich source of new chemical matter with clinical applications. These compounds are broadly defined as small molecules generated by organisms from multiple kingdoms including plants, fungi, and prokaryotes. The outsized impact natural products have had on modern medicine has inspired ongoing efforts to find and characterize new natural products.

The vast increase in sequenced genomes has allowed natural product discovery to move from a blind screening approach to a bioinformatics guided approach. Screening of extracts for bioactivity is not only laborious, but also leads to the rediscovery of known compounds. With a genome-mining approach, bioinformatic analysis of natural product gene clusters empowers researchers to identify producer strains and novel products. This eliminates the rediscovery problem and reduces the throughput necessary for novel natural product isolation.

Within the large family of natural products, the Ribosomally-produced and Post-translationally modified Peptides (RiPPs) are a particularly intriguing class. These compounds begin as a genetically encoded precursor, which is translated into a short (~50) residue peptide. This precursor peptide is then modified by genetically co-located enzymes to generate a mature natural product. The diverse, enzymatically-derived modifications to RiPP precursors include both sidechain and backbone modifications and generate a large and chemically diverse class of natural products.

RiPPs are a potentially fruitful source of new chemical matter for clinical application for several reasons. First, the fact that RiPP precursors are genetically encoded means that there is a direct link between genetic sequence and chemical structure. Because of this, bioinformatic analysis allows relatively easy prediction of final chemical structure, and opens the possibility

for analog generation through relatively simple genetic modifications. Secondly, recent advances in genome sequencing have greatly increased the known space of organismal genomes. As far as RiPPs are concerned, this offers plentiful opportunity for genome mining to identify novel RiPP biosynthetic gene clusters with potentially novel bioactivities.

In addition to leveraging bioinformatics to discover novel RiPPs, this class of natural product has an increasingly diverse range of bioactivities. The wealth of information regarding known mechanisms of action of RiPPs serves as further justification for the discovery and characterization of novel RiPPs. In Chapter 1, I discuss the current state of understanding of the mechanisms of action of several classes of RiPP. This introductory chapter is an excerpt from a comprehensive review of the mechanisms of action of all known RiPPs.

In Chapter 2, we report the bioinformatics-guided discovery and biosynthetic reconstitution of the lasso peptide fusilassin. This biosynthetic gene cluster was shown to possess exceptional plasticity, and was used to generate a number of non-native lasso peptides, including lasso peptides from separate pathways. The enzymes in this pathway also exhibited the greatest *in vitro* stability of known lasso peptides and served as a platform for a number of additional efforts aimed at characterizing interactions of biosynthetic enzymes and generating lasso peptide analog libraries.

In Chapter3, we report the NMR-based characterization of protein-protein interactions governing the biosynthesis of the lasso peptide fusilassin. We accomplished the full backbone assignment of the FusE RiPP recognition element and utilized chemical shift perturbations to characterize the precursor peptide-FusE interactions. Solvent paramagnetic relaxation enhancements were used to characterize the interactions between FusE and the leader peptidase FusB.

In Chapter 4, we described the discovery of two novel thioether-containing RiPPs discovered through genome mining. The sactipeptide huazacin was characterized by HRMS/MS and NMR spectroscopy and was shown to have bioactivity against *Clostridia* sp. Utilizing sequence homology and HRMS/MS, the natural product freyrasin was hypothesized to contain unprecedented alpha-to-beta carbon thioether linkages. This novel structure was confirmed by the complete assignment of freyrasin and TOCSY and NOESY-based experimental characterization of this novel post-translational modification. Based on the NMR structure elucidation, this natural product served as the first fully characterized member of a novel class of RiPPs termed the ranthipeptides.

In Chapter 5 we utilize reactivity-based screening to characterize the post-translational modification of the citrulassin lasso peptides which contain a non-proteinogenic citrulline residue. Comparative genomic analyses and complementation experiments identified the arginine deiminase responsible for the installation of this functional groups. Solution NMR analysis generated a 3D solution NMR ensemble structure for this peptide.

In Chapter 6 we report the discovery of two novel thiopeptide RiPPs. Genome-mining identified a minimal thiopeptide biosynthetic gene cluster with two predicted precursor peptides. These compounds were produced synthetically and chemoenzymatically and fully characterized by HRMS/MS and NMR. The structure elucidation of these compounds led to the reclassification of the thiopeptides as pyritides.

In Chapter 7 genome mining approaches were used to identify and isolate a novel graspetide termed thatisin A. Results from HPLC purifications demonstrated a thermally-dependent atropisomerization. Temperature-controlled NMR structure elucidation accomplished the structure elucidation of thatisin A and demonstrated the presence of a *cis*-proline residue

located within a constrained macrocycle. Computational and MS-based methanolysis experiments further confirmed our proline isomerization conformational exchange hypothesis.

Table of Contents

Chapter 1: Mechanism of Action of Ribosomally-Produced and Post-Translationally Modified Peptide Natural Products[1]

1.1 Introduction

Natural products have served as a valuable source for bioactive compounds with clinical application. The major classes of natural products include the terpenoids, alkaloids, polyketides, non-ribosomal peptides and the ribosomally produced and post-translationally modified peptides (RiPPs). The biosynthesis of RiPPs begins as a linear, translated peptide which is then modified by biosynthetic enzymes to produce a mature natural product. The precursor peptide and biosynthetic enzymes are generally co-located within a biosynthetic gene cluster, which allows the leveraging of bioinformatic analyses to identify and characterize novel RiPP natural products.

The RiPP class of natural products is very diverse, and includes over 40 separate families. In addition to their structural diversity, they also collectively possess broad bioactivities. Comprehensive reviews of the different classes of RiPP have been reported,[1,2] though a full understanding of their mechanisms of action have yet to be reported. In this chapter, we introduce several classes of RiPP and describe the current state of understanding of their mechanisms of action.

1.2 Lasso Peptides[1]

The lasso peptides are RiPPs that possess a macrolactam ring formed between the N-terminus and an internal Asp or Glu residue. This ring encircles the C-terminus, forming a lariat knot conformation. The C-terminus is held within the loop by steric clashing between bulky

[1] This chapter is an excerpt from a draft manuscript in preparation for submission to *Chem. Rev.* Text is the original work of A.J.D. Chayanid Ongpipattanakul assisted in generation of protein structure figures.

sidechains and the macrolactam or by disulfide bonds which endows some lasso peptides with resistance to thermal and proteolytic denaturation. Lasso peptides are divided into four classes based on this topology. This conformation is known colloquially as the "lasso fold," as opposed to a branched-cyclic conformation, and is subdivided into the tail, loop and ring regions. Known lasso peptides exhibit a diverse range of bioactivities, including enzymatic inhibition and receptor antagonism. The sequence diversity of the lasso peptides and their attendant mechanisms of action are broad and represent a promising class of RiPP natural product for further discovery and development.

The cohort of lasso peptides originally discovered were primarily active against cell-surface receptors. The first discovered lasso peptide, anantin, possessed antagonist activity towards atrial natriuretic factor receptor, with a reported K_d of 18.9 pM and *in vitro* IC_{50} of 1.0 μM.[3,4] Currently, no knowledge concerning the molecular level details of the anantin mechanism of action exist. Subsequent discovery and evaluation of another lasso peptide, RES-701-1, demonstrated selective antagonism of the endothelin type B receptor with an *in vitro* IC_{50} of 10 nM with no effect on endothelin type A receptors.[5] Related lasso peptides RES-701-2, -3, and -4 were also shown to be selective antagonists of this receptor.[6] Structural studies of RES-701-1 conclusively demonstrated this compound possesses a threaded conformation.[7] Chemically synthesized RES-701-1 existed in the branched-cyclic conformation, wherein the macrolactam is formed but the C-terminal tail is not threaded through the ring. The branched-cyclic RES-701 lacked the selectivity/activity of naturally isolated RES-701-1, demonstrating the critical importance of a threaded conformation for bioactivity.[8,9] Structure-activity relationship (SAR) studies demonstrated that the C-terminal Trp of RES-701-1 is directly involved in the selectivity of this compound, wherein the modification of the C-terminal carboxylate to a methyl ester or

conversion of the C-terminal Trp to tryptamine decreased the IC_{50} against the endothelin type A receptor.[10] Endothelin B receptor IC_{50} was only slightly increased by the removal of the C-terminal residue or by substitution with either Phe or Tyr.[11] Substitution with Gly or Ala substantially increased the *in vitro* IC_{50} against endothelin B, from 10 nM for to 120 and 540 nM for the Gly and Ala variants, respectively.[11] Most interestingly, a C-terminal substitution with a β-2-naphthyl-alanine decreased the IC_{50} from 10 nM to 6.3 nM.[11] These SAR studies suggest that larger/aromatic sidechains at the C-terminus stabilize the binding conformation of RES-701-1, as smaller residues increase the flexibility of the C-terminus and destabilize the interaction of RES-701 and the endothelin B receptor interacting conformation. Biological activity of RES-701-1 was also demonstrated *ex vivo* in isolated rat blood vessels and small intestines as well as *in vivo*.[12,13] Atomic resolution structures of RES-701-1 and endothelin B have not been reported, which would be valuable information for further optimization of this inhibitor.

Among the most clinically relevant receptor antagonist lasso peptides identified, BI-32169 was demonstrated to inhibit glucagon-dependent cAMP production by the human glucagon receptor with an IC_{50} of 440 nM.[14,15] The activity of BI-32169 was improved with the conversion of the C-terminal carboxylate to a methyl ester.[14] Groundbreaking work accomplished the total synthesis of the threaded conformation of BI-32169. In this work, anti-glucagon receptor antagonism was enhanced from 520 nM to 160 nM against BHK-21 cells expressing the human glucagon receptor with the D-enantiomer of BI-32169.[16] To date, no detailed SAR or structural studies elucidating the molecular-level details of glucagon receptor inhibition for BI-32169 have been performed.

Known lasso peptides also display protease inhibitory activity. Propeptin was initially shown to have potent activity against human, bovine, and flavobacterial prolyl endopeptidases, as well as

antibacterial activity against *Mycobacterium phlei*.[17,18] Isolation of propeptin-2, which lacked the two Ser-Pro C-terminal residues of propeptin retained anti-prolyl endopeptidase activity but lost all antibacterial activity against *M. phlei*.[19]

The lasso peptide lassomycin is potently active against *M. tuberculosis,* with MICs of 0.21-1.65 μM against multiple clinical and drug-resistant strains.[20] Lassomycin-resistant bacteria possessed mutations to the ClpC1 ATPase, an essential protein that interacts with the also essential ClpP1P2 protease responsible for protein recycling in *M. tuberculosis*.[20–22] *In vitro*, purified ClpC1 exhibited increased ATPase activity in the presence of lassomycin but a decreased ability to degrade casein.[20] This was presumed to be by lassomycin-dependent inhibition of translocation of a bound protein substrate to the proteolytic active site. Structural modeling studies indicated lassomycin likely interacts with the N-terminus of ClpC1, potentially interfering with the conformational changes necessary for proper function, thereby decoupling ATP hydrolysis from inducing conformational changes in ClpP1P2 necessary for proteolytic activity.[20] Efforts to chemically synthesize lassomycin generated a compound that was isobaric to the naturally occurring compound, but exhibited markedly different NMR chemical shifts.[23] Synthetically prepared lassomycin also lacked any anti-*M. tuberculosis* bioactivity up to 100 μg/mL.[23] These results are consistent with the synthesis of the branched-cyclic form of lassomycin, which demonstrates the critical role that the threaded conformation has on bioactivity.

Antiviral activities of lasso peptides are exemplified by siamycins I and II, which were originally identified as inhibitors of HIV and HSV, in addition to some Gram-positive antibacterial activity.[24] The discovery of siamycins I and II was accomplished using bioactivity screening, specifically the syncytia inhibition assay in which a cell line expressing the HIV envelope

proteins gp120 and gp41 were cocultured with HeLa cells expressing the CD4 antigen protein.[25] Fusion of these cells generates syncytia, or polynuclear cells, which can be prevented in the presence of potential inhibitors of HIV. The bioassay-guided isolation of siamycins I and II utilized the HIV gp160 protein instead of the gp120 protein.[24,26] In cell monoculture, anti-HIV IC_{50} values for siamycin I and II were 7 and 9 µg/mL, respectively, while anti-HSV IC_{50} for siamycin I was 48 µg/mL and siamycin II was 27 µg/mL.[24] Structural studies revealed that the siamycins possess a tricyclic Class I structure, wherein the traditional lasso fold also contains an intrapeptide disulfide linkage.[27] In studies of siamycin-I-resistant HIV particles, the interaction site was localized to gp160, the HIV glycosylated preprotein responsible for cell fusion, which contained several point mutations.[26] Utilizing the syncytia inhibition assay, it was demonstrated that preincubation with siamycin I, followed by removal by washing restored syncytia formation, suggesting this compound acts as a non-covalent protein-protein interaction inhibitor.[26] A third anti-HIV lasso peptide, termed RP 71955, was discovered utilizing a fluorometric assay to screen for inhibitors of HIV-1 protease.[28] Structural work on this peptide illuminated it also shared a class I lasso topology and led to the renaming of RP 71955 to siamycin III.[29] No detailed mechanistic analyses on the anti-HIV activities of these peptides have been performed.

Concerning the antibiotic activity of siamycin I, several studies elucidating the mechanism of action have been performed. Several siamycin-I resistant mutants exhibited cross-resistance to both vancomycin and the RiPP nisin (see Lanthipeptides), two antibiotics known to target the cell wall.[30] Treatment with siamycin I induced the *liaI* (lipid II interacting antibiotic) cell wall stress response, a hallmark of antibiotics targeting the cell wall.[30] Inhibition of the transpeptidation reaction of lipid II by penicillin-binding protein 2 was also identical to that of vancomycin, which binds lipid II directly.[30] This, taken with the fact that siamycin I can prevent

the dephosphorylation and recycling of lipid II but does not influence the processing of cytosolic cell wall precursors, provided strong evidence that siamycin I most likely directly binds lipid II at the cell surface.[30] Siamycin I has also shown inhibitory activity (IC_{50}= 8 μM) against myosin light chain kinase, which further highlights the diversity of lasso peptide bioactivities.[31,32] Bioactivity against cell wall regulation has also been suggested in the case of the lasso peptide streptomonomicin. *In vitro* MICs of purified streptomonomicin were <10 μM against multiple *Bacillus* sp. with little to no activity against fungi or Gram-negative bacteria.[33] Whole-genome sequencing of spontaneously resistant *Bacillus anthracis* colonies demonstrated five mutations to the *walR* gene and a single mutation in the upstream region of this gene. Characterization of WalR has shown this protein to be a response regulator which functions with WalK, a histidine kinase, to form a two-component regulator that controls the expression of proteins involved in cell wall metabolism and division.[34,35] Of the isolated WalR mutants, three mutations occurred within the DNA binding domain, while two mutations occurred within the receiver domain.[33] Phenotypically, *B. anthracis* mutants grew in long chains, indicative of a septal cleavage defect.[33] This is rationalized as the WalRK system influences the expression of LytE an autolysin involved in septal cleavage in *Bacillus* sp.[33,36] The downregulation of *lytE* was demonstrated for streptomonomicin-resistant mutants by qRT-PCR, suggesting that streptomonomicin targets cell wall biosynthesis through the WalR protein.[33]

In addition to anti-lipid II activity, siamycin I also attenuates the *fsr* quorum-sensing system of *Enterococcus faecalis.* This quorum-sensing system is auto induced by the peptide hormone gelatinase biosynthesis-activating pheromone (GBAP) and leads to the expression of two virulence factors: the gelatinase GelE and a serine protease SerE. Screening efforts identified siamycin I as an inhibitor of GelE expression by attenuation of GBAP production at sublethal

concentrations.[37] The inhibition of GBAP by siamycin was demonstrated by inhibition of the *fsr* two-component signal transduction proteins FsrC-FsrA.[37] Subsequent work showed siamycin I inhibited the autophosphorylation activity of FsrC.[38] Furthermore, siamycin I also possessed *in vitro* inhibitory activity against several ATP-dependent enzymes, including membrane sensor kinases, protein kinases, and ATPases of both prokaryotes and eukaryotes.[38] Despite these *in vitro* activities, siamycin I lacks anti-Gram negative activity *in vivo*, suggesting that this lasso peptide cannot cross the bacterial outer membrane.[24,38] Follow up studies using circular dichroism spectroscopy also supported a direct interaction between siamycin I and the FsrC kinase occurring at a separate site from the native GBAP ligand.[39]

The lasso peptide sviceucin shares the class I topology with the siamycins and has also demonstrated anti-*fsr* activity.[40] Sviceucin has a MIC of 10 µM against *E. faecalis*, but the treatment of this bacteria with as little as one µM reduces GelE production by 50-70%.[40] The shared bioactivity and topology between siamycin I and sviceucin suggest that these two lasso peptides share a mechanism of action against the *fsr* pathway.

Several lasso peptides have a demonstrated activity against Gram-negative RNA polymerase. The best-studied of these, and lasso peptides general, is microcin J25 (MccJ25), a bactericidal lasso peptide active against clinical *E. coli, Salmonella,* and *Shigella* strains.[41–43] Studies of spontaneously resistant mutants possessed a T391I substitution in RpoC, the β' subunit of *E. coli* RNA polymerase (RNAP).[44] MccJ25 targeting of RNA polymerase was further supported with complementation and *in vitro* and *in vivo* transcription assays.[44] The potential site of MccJ25 binding was proposed to be in the *E. coli* RNAP secondary channel, which was further supported by the generation of several other resistant mutants containing point mutations within this channel.[45] Biophysical characterizations demonstrated that MccJ25 binds the secondary channel

of *E. coli* RNAP, which prevents entry of nucleotides to the enzymatic active site.[46,47] Biophysical studies showed MccJ25 "locked" *E. coli* RNAP into an inactive conformation.[46] Saturation mutagenesis of this secondary channel also demonstrated that substitutions at multiple sites within this channel could endow resistance to MccJ25 (Fig. 1).[47] Recent studies on the cocrystal structure of MccJ25 and *E. coli* RNAP confirmed these findings, as MccJ25 bound within the secondary channel (Fig. 1).[48]

Transposon mutagenesis endowing MccJ25-resistance identified insertions in the *fhuA* gene MccJ25, which encodes a multifunctional outer membrane protein largely involved in siderophore transport.[49] A complete model of MccJ25 internalization was established when additional spontaneously resistant *E. coli* possessed mutations in *tonB*, *exbB/D*, and *smbA* genes.[50] These genes encode components of the TonB membrane transport pathway, wherein complexation of TonB, ExbB, ExbD, and SmbA to an unidentified proton channel induce a conformational change in FhuA, causing it to open and import MccJ25.[50] Complementation of *E. coli* with the *Salmonella enteritidis* or *Pantoea agglomerans* FhuA protein rendered

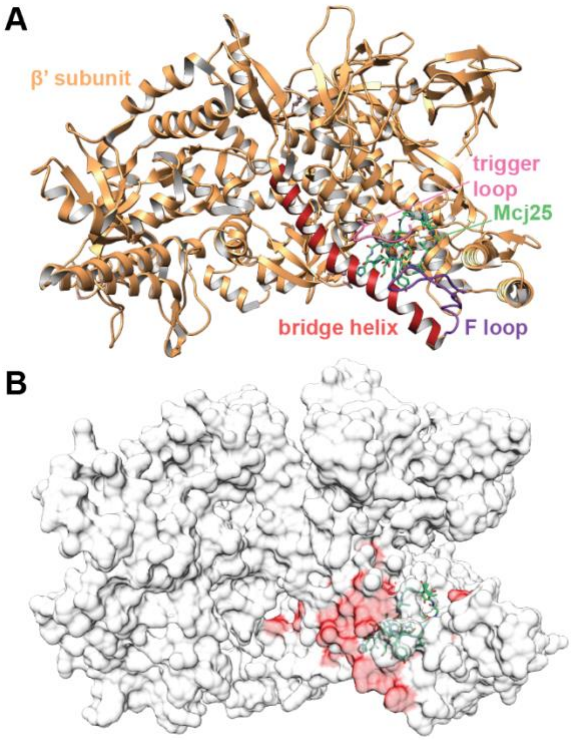


Figure 1.1. Interaction sites of Microcin J25 and *E. coli* RNA Polymerase β' subunit. The cocrystal structure of MccJ25 and *E. coli* RNAP demonstrated binding within the secondary channel (A). Residues identified through mutagenesis which contribute to MccJ25 resistance occur within the secondary channel (red, B).

E. coli resistant to MccJ25. In contrast, *Salmonella enteritidis* typhimurium could be rendered sensitive to MccJ25 with *E. coli* FhuA complementation.[51,52] Deletion studies of the *fhuA* gene demonstrated MccJ25 bound the exterior loops of the protein, as deletions of several of these loops abrogated MccJ25 bioactivity.[53] Direct binding of MccJ25 to FhuA *in vitro* demonstrated a K_d of 1.2 μM and stoichiometry of 2:1 MccJ25 to FhuA.[54] More recent structural studies showed MccJ25 directly bound the exterior cavity of FhuA in a similar binding mode to the native ligand ferrochrome (Fig. 2).[55] This interaction induces MccJ25 uptake via the TonB-dependent membrane transport system.[55] Contrary to extensive polar contacts formed between ferrichrome and FhuA, MccJ25 only forms three between the MccJ25-H5 imidazole, the MccJ25-A3 backbone carbonyl and the MccJ25-V6 backbone carbonyl.[55]

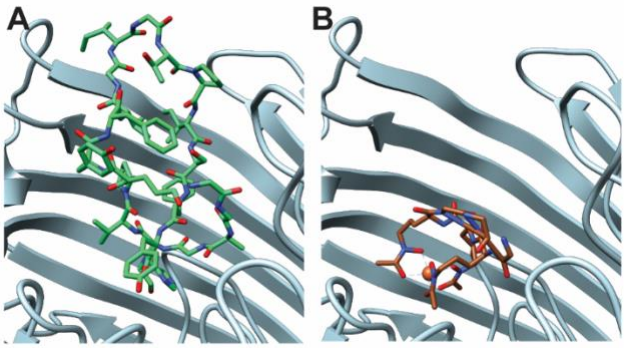


Figure 1.2. Cocrystal structures of FhuA-MccJ25 and FhuA-ferrichrome. The cocrystal structure of FhuA and MccJ25 demonstrated this lasso peptide bound to the exterior cavity of FhuA (A). The native FhuA ligand, ferrichrom, also binds to the exterior cavity of FhuA.

Ligand-centric studies of MccJ25 originally established the importance of the threaded topology to bioactivity, as thermolysin cleavage of the β-hairpin loop region of MccJ25 generated an interlocked dual-chain version (t-MccJ25).[56,57] In growth inhibition assays, t-MccJ25 had MICs of 80-150 nM and 300-600 nM against *S. enteritidis* (WT MccJ25 = <2 nM) and *E. coli* (WT MccJ25 = 2-5 nM), respectively.[56,57] *In vitro* transcription studies of demonstrated both MccJ25 and t-MccJ25 inhibited *E. coli* RNAP and that losses in bioactivity

result from impaired FhuA-dependent uptake.[58,59] Modification of MccJ25-His5 by carbethoxylation also reduced biological activity, likely by interfering with peptide uptake.[60] Additional studies on the role of this residue suggested MccJ25-His5 was important for internalization by the SmbA transporter.[61] Bypass of the FhuA receptor by osmotic shock did not restore antibacterial activity to MccJ25-H5A or H5R variants, supporting the importance of SmbA on the transport of MccJ25.[61] Modification of the MccJ25-H5 residue is also consistent with structural studies, which demonstrate the importance of the MccJ25-H5 imidazole in FhuA interactions.[55] Amidation of the C-terminal Gly of MccJ25 was also shown to inhibit RNAP interaction *in vitro* but did not abolish peptide uptake.[62] Taken together, these results indicate the ring and loop regions are critical for uptake. Simultaneously, the C-terminal tail is necessary for RNAP interaction and inhibition, which was confirmed by structural studies.[55]

Secondary mechanisms of action for MccJ25 have been proposed, including dissipation of the cytoplasmic membrane gradient, superoxide production, and antimitochondrial activity. In studies of *Salmonella enterica* serovar *newport*, MccJ25 was shown to inhibit respiration, as measured by oxygen consumption in the presence of 10 mM D-glucose.[63] Assays with the tetrazolium dye MTT demonstrated a marked decrease in the activities of NADH and succinate dehydrogenases in isolated membrane fractions of MccJ25-treated *S. newport*.[63] Both of these enzymes are integral membrane proteins, and effects on peripheral membrane enzymes were not observed, nor were similar effects noted in assays against *E. coli.*[63] Similar effects on oxygen consumption have been observed on the oxygen consumption of *S. typhimurium*, *Salmonella derby,* and *Salmonella enteritidis* strains transformed with the *E. coli* FhuA protein.[52,59] This MOA is further supported by studies of MccJ25 effects on lipid vesicles *in vitro*, where MccJ25 not only inserted into but also increased the fluidity and permeability of membranes, including

bacteriomimetic vesicles in a FhuA-independent manner.[64,65] This secondary MOA has also been supported by findings that 2,4-dinitrophenol-dependent dissipation of the proton motive force, which FhuA uses for active transport of MccJ25, did not drastically impair the antibacterial activity of this lasso peptide, suggesting RNAP is not the only target.[66]

Reactive oxygen species (ROS) production has also been observed in *E. coli* under FhuA overexpression.[67] Utilizing a C-terminally amidated analog of MccJ25, which does not interact with RNAP, *E. coli* FhuA-overexpressing strains were sensitized to MccJ25 through the production of ROS.[68] *E. coli* strains expressing FhuA with a deletion of either cytochrome *bd*-I or bo_3 remained resistant to MccJ25, implicating these two cytochrome enzymes as potential targets of MccJ25 which effect superoxide production.[68] MccJ25 was further shown to weakly inhibit purified cytochrome *bd*-I and induce ROS production by this enzyme.[69] Follow up studies confirmed the role of Tyr_9 in superoxide production, and electron paramagnetic resonance studies demonstrated the formation of a tyrosyl radical and its participation in the redox reaction cycle.[70,71] Antimitochondrial activities of MccJ25 have also been observed in rat heart mitochondria, wherein MccJ25 inserted into the membrane, dissipated the proton motive force, reduced intracellular ATP, and inhibited complex III of the cytochrome c reductase respiratory chain.[72] These studies involved concentrations of MccJ25 as low as one µM. However, this is significantly higher than bioactive concentrations observed against *S. enteritidis* and *E. coli,* suggesting that this is not a primary MOA of MccJ25.[56,72] Through ROS production, MccJ25 was shown to activate the mitochondrial transition pore, leading to mitochondrial swelling and cytochrome c release.[73] Furthermore, MccJ25 was shown to induce carbonylation of mitochondrial proteins and oxidation of mitochondrial lipids, both of which would further induce the release of cytochrome c.[74] As previously, the MccJ25-dependent effects on the mitochondrial

pore and membrane were only observed in the presence of 20 μM MccJ25, which is significantly higher than antibacterial concentrations suggesting that this is not a physiologically relevant mechanism of action.[73,74]

Other known lasso peptide inhibitors of RNAP include capistruin,[75] acinetodin,[76] klebsidin,[76] citrocin,[77] and ubonodin.[78] Structural characterization of capistruin demonstrated a class I topology with bioactivity against Gram-negative *Burkholderia* sp. and *Pseudomonas* sp. bacteria, inspired the hypothesis that capistruin also targets RNAP.[79] This hypothesis was supported by findings that capistruin inhibited transcription activity of *E. coli*, and MccJ25-resistant RNAP secondary channel mutants exhibited cross-resistance to capistruin *in vitro*.[75] Structural work on RNAP and capistruin demonstrated that this lasso peptide bound within the secondary channel much like MccJ25.[48] This activity is contrary to the findings that capistruin is only weakly active (MIC = 25 μM) against *E. coli*, suggesting that capistruin is not able to cross the *E. coli* outer membrane as effectively as MccJ25.[79] Structural studies of the *E. coli* FhuA receptor demonstrated that capistruin could not be modeled into the exterior cavity without steric clashing.[55] Interestingly, capistruin is active against *Burkholderia* strains which is thought to be a consequence of the low sequence homology of *E. coli* and *Burkholderia* FhuA receptors, allowing capistruin to enter Gram-negative *Burkholderia* cells but not *E. coli*.[55] This proposal is speculative, as there is no current data suggesting capistruin is imported by the FhuA receptor.

Like capistruin, acinetodin and klebsidin are active against *E. coli* RNAP, and MccJ25-resistant mutants exhibited cross-resistance to both of these lasso peptides.[76] Additionally, while these compounds were inactive against *E. coli*, klebsidin was active against *K. pneumonia*.[76] *E. coli* could be rendered sensitive to klebsidin through expression of the *K. pneumonia* FhuA protein, which possesses 60% identity with the *E. coli* FhuA.[76] These findings support the mechanism by

which these lasso peptides are imported by the FhuA/TonB-dependent transport system and rationalize the relatively narrow ranges of bioactivity, as these lassos are particularly sensitive to the primary sequence of FhuA receptors.

The lasso peptides citrocin and ubonodin have also demonstrated anti-*E. coli* RNAP activity *in vitro.*[77] At one μM, citrocin reduced transcription levels of RNAP to 15% of control, a reduction that was matched by 100 μM MccJ25.[77] Spontaneously resistant *E. coli* mutants contained mutations in the *sbmA* but didn't contain mutations in the *fhuA* gene.[77] This suggests a similar MOA to MccJ25, though citrocin has a much more potent RNAP activity despite a higher MIC against *E. coli* (Citrocin = 16-125 μM vs. MccJ25 = 1 μM) likely due to a reduced ability for citrocin to cross the Gram-negative outer membrane.[77] Ubonodin has demonstrated potency against members of the *Burkholderia* genus, with the most potent activity against *Burkholderia cepacia* (MIC = 4 μM).[78] Due to similarity in structure to MccJ25, ubonodin was evaluated for *in vitro* transcription abortive activity and demonstrated slightly lower potency than MccJ25.[78] At 100 μM, ubonodin reduced transcription levels by ~60%, compared to ~85% with MccJ25.[77,78] This reduction suggests that ubonodin also targets RNAP. Its lower antibacterial activity against *E. coli* correlates with outer membranetransport. In the case of ubonodin, susceptible *Burkholderia* strains were not found to contain highly homologous FhuA proteins (>40% sequence identity), indicating that ubonodin could have a unique internalization mechanism relative to MccJ25.[78]

Lariatin A and B were initially observed to be growth-suppressive against *M. smegmatis* and also possessed activity against *M. tuberculosis*.[80] Structure-activity studies of these peptides have been performed highlighting positions necessary for bioactivity, but no detailed studies regarding the mechanism of action have been performed.[81,82]

Streptomonomicin is a lasso peptide with proven bioactivity and a suggested mechanism of action. *In vitro* MICs of purified streptomonomicin were <10 μM against multiple *Bacillus* sp., little to no activity against fungi or Gram-negative bacteria.[33] Whole-genome sequencing of spontaneously resistant *B. anthracis* colonies demonstrated five mutations to the *walR* gene and a single mutation in the upstream region of this gene. Characterization of WalR has shown this protein to be a response regulator which functions with WalK, a histidine kinase, to form a two-component regulator that controls the expression of proteins involved in cell wall metabolism and division.[34,35] Of the isolated WalR mutants, three mutations occurred within the DNA binding domain, while two mutations occurred within the receiver domain.[33] Phenotypically, *B. anthracis* mutants grew in long chains, indicative of a septal cleavage defect.[33] This is rationalized as the WalRK system influences the expression of LytE, and autolysin involved in septal cleavage in *Bacillus* sp.[33,36] The downregulation of *lytE* was demonstrated for streptomonomicin-resistant mutants by qRT-PCR, suggesting that streptomonomicin targets cell wall biosynthesis through the WalR protein.[33]

Antifungal lasso peptides have also been reported. Humidimycin was identified by high-throughput screening for potentiators of capsofungin activity against pathogenic *Aspergillus fumigatus* and *Candida albicans*.[83] Humidimycin, when introduced with capsofungin, was found to prevent phosphorylation/activation of SakA, the primary controller kinase of the high osmolarity glycerol pathway.[83] Deletion mutants of SakA also exhibited higher sensitivity to humidimycin, suggesting this pathway is the target of this lasso peptide, though currently no exact mechanism of action is known.[83] These findings are of interest due to the previously observed kinase inhibitory activity of the closely related lasso peptide siamycin I.[38]

1.3 Sactipeptides

Sactipeptides are a RiPP class defined by the presence of intrapeptide sulfur to alpha carbon thioether crosslinks installed by a radical S-adenosyl methionine enzyme. Published structures of the sactipeptides have illuminated their stable hairpin structure, which in the case of subtilosin A and sporulation killing factor, includes head-to-tail cyclization.[84–89] All characterized sactipeptides possess antibacterial activities. The majority possess intragenus bioactivity and anti-biofilm and bactericidal activity against the human pathogens *Clostridium difficile, Listeria monocytogenes, Staphylococcus sp.* and *Gardnerella vaginalis*.[87–98] Studies of the mode of action of the sactipeptides indicate that these compounds likely target the membrane.[89,94,99–101] Original structural studies of subtilosin A led to the observation that this compound is hydrophobic with discrete basic and acidic regions, contributing to the compounds amphipathicity and suggesting it may target the membrane.[99,102] Using the intrinsic fluorescence of the single Trp in Subtilosin A, it was demonstrated that this RiPP doesn't form multimeric complexes at concentrations below 16.4 µM.[99] Furthermore, the intrinsic fluorescence of subtilosin A shifted to lower wavelengths in the presence of POPC and POPG small unilamellar vesicles (SUVs), indicating the Trp residue is located within a low polarity environment in these artificial membranes.[99] These SUVs were formulated to contain carboxyfluorescein and exposed to subtilosin A, which induced dye leakage in POPC and 7:3 POPE/POPG (a bacterial mimetic SUV) in a subtilosin A concentration-dependent manner.[99] No dye leakage was observed for SUVs composed of 4.5:4.5:1:1 POPC/SM/POPE/Chl, a mammalian membrane mimic, which indicates the selectivity of subtilosin A is at least partially dependent on membrane composition.[99] NMR studies of ^{15}N-labelled subtilosin A in aligned POPC lipid bilayers demonstrated that subtilosin A exists in two conformations, one perpendicular to the membrane

surface and a smaller population aligned near parallel to the membrane surface.[99] Utilizing ^{31}P NMR to study lipid conformation, upfield shifts of ~2-4 ppm of POPC, POPG, and DMPC bilayers in the presence of subtilosin A indicate an altered lipid head group orientation.[99] This conformational change was further confirmed in studies of quadrupolar splitting of the α- and β-methylenes of choline in DMPC membranes, indicating a subtilosin A-dependent conformational change as well as a modest increase in lipid acyl chain conformational exchange.[99] As the concentrations necessary to induce permeability were higher than the MIC of subtilosin A, it was theorized that this compound's MOA includes interaction with a surface receptor *in situ.*[99,103]

General mechanisms of membrane interaction were supported by the isolation of a hemolytic variant of subtilosin A possessing a single T6I mutation.[103] This variant also had vastly decreased MIC values against Gram-positive and -negative bacteria (Table 1).[103] The finding that a single amino acid substitution considerably altered bioactivity was proposed to occur through modified interaction with phospholipids, as a Thr to Ile mutation increases the peptide's hydrophobicity.[103]

Table 1.1. Comparative Minimum Inhibitory Concentrations (MICs) of Subtilosin A and the hemolytic variant Subtilosin A T6I. *Hemolysis was assessed using rabbit blood agar, as opposed to erythrocytes..

Species	Subtilosin A MIC (μM)	Subtilosin A T6I MIC (μM)
Bacillus subtilis	250	100
Bacillus anthracis	16	2.56
Bacillus cereus	40	6.4
Bacillus thuringiensis	40	6.4
Enterococcus faecalis	100	16
Staphylococcus carnosus	100	16
Listeria monocytogenes	40	6.4
Streptococcus pyogenes	100	6.4
Rabbit blood agar*	>250	16

Subtilosin A has also demonstrated spermicidal activity and inhibition of late-stage replication of Herpes-Simplex Virus-1 (HSV-1).[104–106] Subtilosin A has shown an *in vitro* potency against human spermatozoa with an IC_{50} of 64.5 µg/mL, and treatment of ectocervical cell cultures with subtilosin A only reduced cell viability by ~5-10% over 24 hrs.[105] In a dose-dependent fashion, subtilosin A inhibited the forward progression of spermatozoa and induced tail coiling, though no detailed mechanism for this observation has been proposed.[105] Concerning HSV-1, subtilosin inhibited viral replication *in vitro* by 91.7% at concentrations as low as 12 µg/mL.[106] This inhibition was shown to be through direct interaction with HSV-1, as preincubation with 200 µg/mL reduced the viral titer by 99%, though concentrations from 6-100 µg/mL did not reduce viral titer.[106] Inactivation of HSV-1 was shown to occur not in the early stages of viral replication but instead likely inhibited viral assembly or release.[106]

Studies of subtilosin A against the bacterial pathogen *G. vaginalis* demonstrated the compound's ability to dissipate the transmembrane pH gradient and induce efflux of intracellular ATP.[94] Subtilosin A treatment of *G. vaginalis* at two µg/mL facilitated ATP release, causing extracellular ATP content to increase to 40% of the total intracellular + extracellular ATP, fully 20% above control.[94] Utilizing the fluorophore 3,3'-dipropylthiadicarbocyanine and the K^+/H^+ ionophore exchanger nigericin subtilosin A was did not affect the transmembrane potential ($\Delta\Psi$).[94] The lanthipeptide nisin (See: Lanthipeptides) induced an immediate increase in fluorescence, implying dissipation of $\Delta\Psi$.[94] Though subtilosin A didn't deplete $\Delta\Psi$, it was shown to completely deplete the transmembrane pH gradient (ΔpH), as did nisin.[94] The fact that subtilosin A only dissipated the ΔpH component of the proton motive force suggests that this RiPP forms transient pores in the bacterial membrane.[94] Additional studies against *L.*

monocytogenes demonstrated minimal effect on the pH gradient and ATP efflux, which further suggests that subtilosin A activity is impacted by the composition of the cell membrane.[95] Additional MOA studies of subtilosin A indicate this compound can also inhibit quorum sensing-dependent biofilm formation.[107] At sub-MIC concentrations subtilosin A was shown to inhibit biofilm formation of *G. vaginalis* (0.78 µg/mL, >90% reduction), *L. monocytogenes* (15.1 µg/mL, 80% reduction) and *E. coli* (15.1 µg/mL, 60% reduction).[107] Subtilosin A was also implicated as a potential inhibitor of quorum sensing using the indicator bacteria *Chromobacterium violaceum*. The production of the pigment violacein by *C. violaceum* has been used previously to detect the production of N-acyl homoserine lactone quorum sensing compounds.[108] Subtilosin A at sub MIC concentrations of 7.8-125 µg/mL reduced *C. violaceum* violacein production by >70%.[107] Similarly, Autoinducer-2 (AI-2) is a quorum-sensing molecule found in a wide variety of bacteria.[109] Production of AI-2 in *G. vaginalis* was inhibited by subtilosin A at 3 and 4 µg/mL with no influence on bacterial growth. However, the same inhibition was not observed for *L. monocytogenes* at 0.95-15.1 µg/mL subtilosin A, which reduced cell viability >50% at 15.1 µg/mL.[107] Combining the findings of subtilosin A illuminates its ability to interact with and disrupt the membrane and highlights the complicated MOA of select RiPPs, which necessitates further study to determine the full mechanistic function of these natural products.

Mechanism of action studies have also been undertaking for thuricin CD, thuricin Z/huazacin, and thurincin H. Thuricin CD is a two-component antibiotic composed of the separate sactipeptides thuricin α and β, both of which are necessary for full activity.[84,92] Studies of thuricin CD have demonstrated induction of irreversible membrane depolarization and pore-formation in *Bacillus firmus* with a MIC of 125 nM.[101] Membrane polarization was studied

using the dye 3,3'-Diethyloxacarbocyanine iodide, which loses fluorescence as bacterial membranes depolarize. In real-time flow cytometry assays, both thuricin α and thuricin β at three μM induced depolarization and reduced cell size in isolation.[101] At higher concentrations (9 μM), thuricin CD elicited a faster depolarization to a greater degree than the individual thuricin peptides.[101] Light microscopy showed *B. firmus* contained low-density regions of the plasma membrane after thuricin CD treatment, indicating that thuricin CD targets the *B. firmus* membrane.[101] The individual components of thuricin CD seem to act synergistically to disrupt the bacterial membrane, though the potential role of a native cell surface receptor has not been ruled out.[101]

Similarly, thuricin Z/huazacin was also demonstrated to interact directly with the bacterial cell membrane and induced ion permeability of bacteria mimetic vesicles.[89] Confocal fluorescence microscopy demonstrated *B. cereus* uptake of propidium iodide upon thuricin Z treatment, indicating membrane permeabilization.[89] Transmission electron microscopy of treated *B. cereus* cells intact cell envelope with partial lysis of the inner bilayer.[89] Scanning electron micrographs additionally demonstrated a collapsed bacterial cell surface and membrane perforations.[89] Thuricin Z possesses a relatively narrow spectrum of activity, with MICs of 2-8 μM against several strains of *B. cereus* but no activity against many other Gram-positive and -negative bacteria.[89] This narrow spectrum of a membrane-targeting antibiotic is likely a consequence of membrane consequence. Preparation of bacteriomimetic large unilamellar vesicles (LUVs) (9:1 1,2-dipalmitoyl-*sn*-glycero-3-phosphatidyl glycerol to 1,2-dipalmitoyl-sn-glycero-3-phosphatidylethanolamine) was permeabilized to K/H ions upon thuricin Z treatment. Additionally, fungi-like LUVs (4:1 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine to ergosterol) did not demonstrate permeabilization in a thuricin Z-dependent manner.[89] The

sensitivity of thuricin Z activity to membrane composition explains its relatively narrow spectrum activity and, by extrapolation, the activity of thuricin CD.

The sactipeptide thurincin H also possesses potent activity against several *Bacillus* species and *Listeria* sp., including the pathogen *L. monocytogenes*.[110] Minimum inhibitory concentrations against *B. cereus* strains have been reported as 0.5-1 nM.[111] Preliminary studies of thurincin H against *B. cereus* demonstrated a reduction in cell viability of ~3-log CFU/mL at 16xMIC, but no decrease in O.D. even at 256xMIC, suggesting that thurincin H does not act in a lytic fashion.[100] *B. cereus* cells treated at large excesses of thurincin H demonstrated a collapsed rod-shaped morphology. Still, they did not exhibit any membrane punctures, as was the case when cells were treated with nisin (See: Lanthipeptides).[100] In assays utilizing propidium iodide, treatment with thurincin H did not permeabilize the membrane to this dye, further supporting the finding that thurincin H does not permeabilize the *B. cereus* membrane.[100] Whether or not thurincin H acts in concert with a cellular receptor remains to be determined.

1.4 Graspetides

The graspetides are a RiPP class defined by the presence of intramolecular lactones between the ω-carboxy groups of glutamate or aspartate and the hydroxy groups of serine and threonine. Graspetides are also known to possess lactam bonds between the ω-carboxy and ε-amino groups of glutamate/aspartate and lysine. The first characterized member of this class, microviridin A, is an inhibitor (IC_{50} = 330 μM) of tyrosinase.[112] In subsequent years, many other graspetides were isolated, all of which possessed potent inhibitory activity towards serine proteases (Table 2). An additional compound, marinostatin, is an inhibitor of several serine and cysteine proteases and pronase metalloprotease.[113,114] Given the universal antiproteolytic activity of the graspetides, an

obvious mechanism of action is as a peptidic substrate mimic. This mechanism was further supported in studies of microviridin J's *in situ* function, which induced a lethal molting defect in the cyanobacterial predatory grazer *Daphnia pulicaria.*[115] This bioactivity is consistent with previous findings demonstrating the very potent inhibition of *Daphnia*-specific trypsin proteases by microviridin J.[116]

Table 1.2. IC_{50} values of characterized graspetides

Name	Trypsin	Plasmin	Elastase	Thrombin	Chymotrypsin	Papain	Tyrosinase	*Daphnia* trypsin-like	Subtilisin	Ref.
Microviridin A	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	.57	N.D.	N.D.	Ishitsuka et al. 1990[112]
Microviridin B	58	>100	0.044	>100 / 7.9	2.5	N.D.	N.D.	N.D.	N.D.	Okino et al. 1995, Anas et al. 2017[126,127]
Microviridin C	32	>100	0.084	>100	4.9	N.D.	N.D.	N.D.	N.D.	Okino et al. 1995[126]
Microviridin D	>100	>100	0.7	>100	1.2	>100	N.D.	N.D.	N.D.	Shin et al. 1996[128]
Microviridin E	>100	>100	0.6	>100	1.1	>100	N.D.	N.D.	N.D.	Shin et al. 1996[128]
Microviridin F	>100	>100	5.8	>100	>100	>100	N.D.	N.D.	N.D.	Shin et al. 1996[128]
Microviridin G	>100	>100	0.019	>100	1.4	>100	N.D.	N.D.	N.D.	Murakami et al. 1997[129]
Microviridin H	>100	>100	0.031	>100	2.9	>100	N.D.	N.D.	N.D.	Murakami et al. 1997[129]
Microviridin I	26.2	N.D.	0.34	N.D.	21.7	N.D.	N.D.	N.D.	N.D.	Fujii et al. 2000[130]
Microviridin J	0.034	N.D.	>10	N.D.	2.8	N.D.	N.D.	0.0039	N.D.	Rohrlack et al. 2003[116]
Microviridin SD1634	13.4	N.D.		N.D.	25.7	N.D.	N.D.	N.D.	N.D.	Reshef et al. 2006[131]
Microviridin L	100	N.D.	>100	N.D.	72	N.D.	N.D.	N.D.	14.9	Ziemert et al. 2010[132]
Microviridin M	>50	N.D.	4.2	N.D.	>50	N.D.	N.D.	N.D.	10.0	Gatte-Picchi et al. 2014[133]
Microviridin N	>50	N.D.	4.2	N.D.	>50	N.D.	N.D.	N.D.	3.3	Gatte-Picchi et al. 2014[133]
Microviridin A3[a]	0.047	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	Ahmed et al. 2017[123]
Microviridin A6[a]	0.036	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	Ahmed et al. 2017[123]
Microviridin A7[a]	N.D.	N.D.	1.9	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	Ahmed et al. 2017[123]
Microviridin A8[a]	N.D.	N.D.	32	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	Ahmed et al. 2017[123]

Table 1.2 (cont.)

Microviridin 1777	>17	N.D.	0.28	N.D.	0.18	N.D.	N.D.	N.D.	N.D.	Sieber et al. 2020[134]
Plesiocin[b]	>134	N.D.	0.071	N.D.	0.0335	N.D.	N.D.	N.D.	N.D.	Lee et al. 2017[124]
Marinostatin[c]	>69	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	0.00082	Taichi et al. 2010[135]

Values reported as IC_{50} µg/mL unless otherwise indicated

[a]Performed using chemoenzymatic synthesized compound; Microviridin A7 and A8 were incompletely cyclized and only contained two heterocycles.

[b]Produced by *E. coli* heterologous expression

[c] Reported as K_i; Performed using chemically synthesized compound

Molecular-level detail of the MOA of microviridins was recently established using mutagenesis and X-ray cocrystal structures of microviridin J and bovine trypsin (Fig. 3).[117] In these structural studies, it was demonstrated that microviridin J residues 4-6 (Thr-Arg-Lys; P2, P1 and P1', respectively) adopted a trypsin substrate-like conformation, which is rigidified by intramolecular lactone bonds (Fig. 3). It is proposed that the rigid, compact structure of the microviridin prevents cleavage of the target amide bond despite the proximity of Arg5 to the catalytic Ser195 nucleophile.[117]

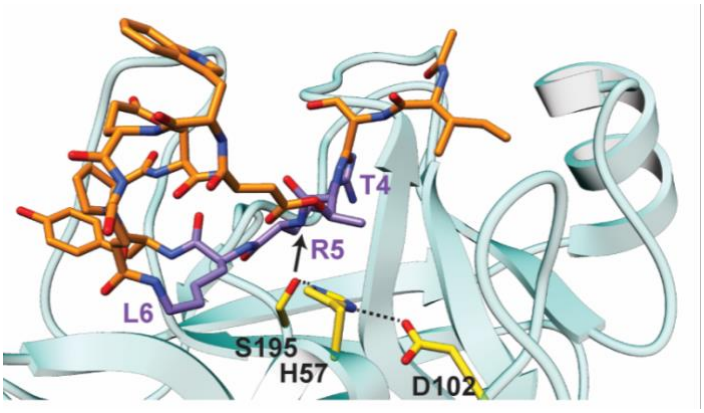


Figure 1.3. Cocrystal structure of bovine trypsin and microviridin J. Tryptic substrate mimicry is accomplished by residues Thr4, Arg5 and Leu6 on microviridin J (purple). The serine nucleophile of trypsin (Ser195) is proximal to the scissile bond (arrow), but cleavage is prevented by the rigidified microviridin J structure.

Through mutagenesis studies of the related microviridin L, it was demonstrated that the P1 site residue contributed to the protease inhibition specificity of microviridin L (Table 3). Introduction of the basic residues Arg and Lys decreased IC_{50} values against trypsin. In contrast, the bulky aliphatic residues Leu and Val caused a loss of potent inhibition of trypsin and a gain of inhibition against elastase.[117] Introduction of basic residues also

led to an increase in inhibition of plasmin, a serine protease with a preference for basic residues, and biases Lys over Arg.[118]

Table 1.3. Protease inhibition by microviridin K variants. The microviridin was varied at position 5 (the P1 site) and was either acetylated (like WT) or deacetylated. Values were determined using a *p*-nitroanilide chromogenic assay *in vitro.* Table adapted from Ref. 8.

Microviridin L Variant	**Elastase IC_{50} (μM)**	**Chymotryspin IC_{50} (μM)**	**Trypsin IC_{50} (μM)**	**Subtilisin IC_{50} (μM)**	**Plasmin IC_{50} (μM)**
F5L	0.65	14	>65	0.78	>65
F5I	5.2	>33	>65	>33	>65
F5V	0.16	66	>66	0.4	>66
F5K	>65	>65	2.5	0.32	2.0
F5R	64	>64	1.8	0.64	47

These findings were reinforced by subsequent studies demonstrating the rational design of microviridin K via mutagenesis of the P1 site residue.[119] These microviridins possess a tricyclic structure arising from two intramolecular lactones and one intramolecular lactam linkage and also contain an acetylated N-terminus.[119] The P1 site of WT microviridin K is Met, a hydrophobic residue that strongly biases inhibition towards subtilosin over trypsin (Table 4). Microviridin B contains a P1 Leu, a relatively bulky residue, which prevents inhibition of

subtilosin or trypsin, likely due to steric clashing with the protease active site. Furthermore, acetylation of the N-terminus had a minimal impact on bioactivity, though seemed to increase potency slightly.

Structural studies of the graspetide marinostatin, a potent subtilisin inhibitor, by solution NMR demonstrated similar structural features to the microviridins, which includes a structurally stable P2-P2' (Thr3-Met4-Arg5-Tyr6) motif which functions as an uncleavable substrate mimic.[120] Replacement of the native ester linkages with disulfide bonds also generated a marinostatin with a stable P2-P2' protease substrate motif. The P2-P1' backbone amide residues exhibited similar rates of hydrogen-deuterium exchange between WT and disulfide marinostatins.[120] Though the disulfide variant was not tested for protease inhibition, these findings suggest that conformational restriction arising from intramolecular linkages plays a significant role in stabilizing an uncleavable conformation of marinostatins and graspetides generally. Furthermore,

Table 1.4. Protease inhibition by microviridin L variants. The microviridin was varied at position 5 (the P1 site), and also contained a G2A mutation, but was fully cyclized and N-terminally acetylated, as was WT microviridin L. Values were determined using a *p*-nitroanilide chromogenic assay *in vitro.*

Microviridin	**Trypsin IC_{50} (μM)**	**Subtilosin IC_{50} (μM)**
MdnK (WT)	>100	5.2 ± 0.32
MdnK (WT, deacetylated)	-	3.9±0.04
MdnK (M5R, deacetylated)	0.41±0.05	-
MdnK (M5R, acetylated)	0.2±0.02	-
MdnK (M5K, deacetylated)	0.48±0.11	-
MdnK (M5K, acetylated)	0.17±0.01	-
MdnB (WT)	>100	>100
MdnB (L5R, deacetylated)	0.09±0.002	-
MdnB (L5R, acetylated)	0.05±0.01	-
MdnB (L5K, deacetylated)	0.36±0.01	-
MdnB (L5K, acetylated)	0.35±0.03	-
MdnB (L5Q, deacetylated	-	10.4±1.1
MdnB (L5Q, acetylated)	-	4.8±0.39
MdnB (L5M, deacetylated)	-	4.3±0.14
MdnB (L5M, acetylated)	-	4.8±0.14

this P2-P2' conformational motif was demonstrated to be similar in two other peptidic protease inhibitors, SSI and OMTKY3, by both 3D alignment and phi/psi bond angles.[120] Additionally, the protease inhibitors SSI and OMTKY3 P2-P2' regions are flanked by a C-terminal β-sheet and an N-terminal disulfide bond, which serve to stabilize the protease substrate region.[121,122] In marinostatins and microviridins, this is accomplished by the intramolecular lactone bonds, which serve to stabilize this region in a similar conformation (Fig. 4). The importance of these stabilizing bonds has been demonstrated, as generally tricyclic microviridins are more potent protease inhibitors than variants containing fewer rings.[117,119,123]

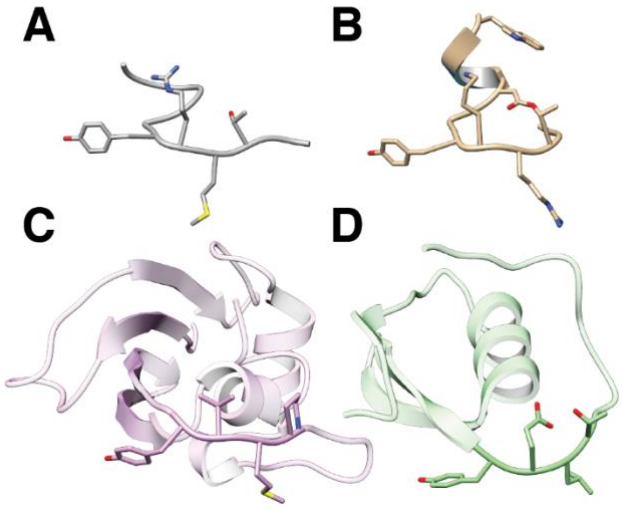


Figure 1.4. P2-P2' site of four peptidic protease inhibitors: Marinostatin (A), Microviridin J (B), SSI (C) and OMTKY3 (D). All four inhibitors adopt very similar structural conformations surrounding the scissile bond.

A third recently reported group of graspetides are named the plesiocins characterized by bicyclic hairpin structures formed by intramolecular ester linkages.[124] The motif contained between the ester linages was found to be Leu-Ala-Ile-Gly, which endowed plesiocin with inhibitory activity against the P1-hydrophobic selective proteases elastase and chymotrypsin.[124] Each plesiocin molecule contains multiple tandem hairpin repeats, and it has been demonstrated that the number of these repeats directly influences the inhibitory ability of these compounds.[124] The *in vitro* plesiocin molecule contains four hairpin motifs and possesses K_i values of 16 nM against

elastase and 7.5 nM against chymotrypsin. In contrast, inhibition assays with only the first of four hairpin motifs demonstrated K_i values of 380 nM against elastase and 63 nM against chymotrypsin.[124] The four hairpin structure was obtained from the heterologously expressed compound, and so it is currently unknown if the native graspetide is cleaved into smaller fragments *in situ* or contains multiple hairpins. The Leu within the bicyclic motif was critical for protease inhibition, as mutation to Ala increased the K_i from 53 nM for WT to >10,000 nM against chymotrypsin.[125] Subsequent work on the stoichiometry of plesiocin inhibition has demonstrated that these compounds can act as a super stoichiometric inhibitor.[125] The full-length plesiocin demonstrated near-complete inhibition of chymotrypsin at a plesiocin:chymotrypsin ration of ~0.16, where a ratio of ~1.0 would indicate a 1:1 stoichiometry.[125] A 1:1 inhibitor to chymotrypsin ratio was demonstrated for a single bicyclic hairpin motif, further demonstrating superstoichiometric inhibition of the four-hairpin plesiocin.[125] Protease inhibition plasticity similar to that of the microviridins was also shown for plesiocin, which tends to inhibit elastase/chymotrypsin over trypsin due to the hydrophobicity of the bicyclic motif, while the introduction of a basic residue to the scissile site of the hairpin can bias inhibition towards trypsin over chymotrypsin.[124,125]

1.5 Proteusins

The proteusins are a class of RiPP defined by the presence of a nitrile-hydratase homologous leader peptide.[136] Originally discovered in the early 1990s, the founding family of the proteusins were the polytheonamides, which are extensively modified and contain N-acylation, multiple *tert*-Leu, N- and C-methylated residues, and alternating D- and L- residue stereochemistry.[137–139] These compounds have demonstrated potent cytotoxicity of 98 pM against murine cells *in vitro*.[139] Modern mechanistic work was performed on polytheonamide B (pTB, Fig. 5), which

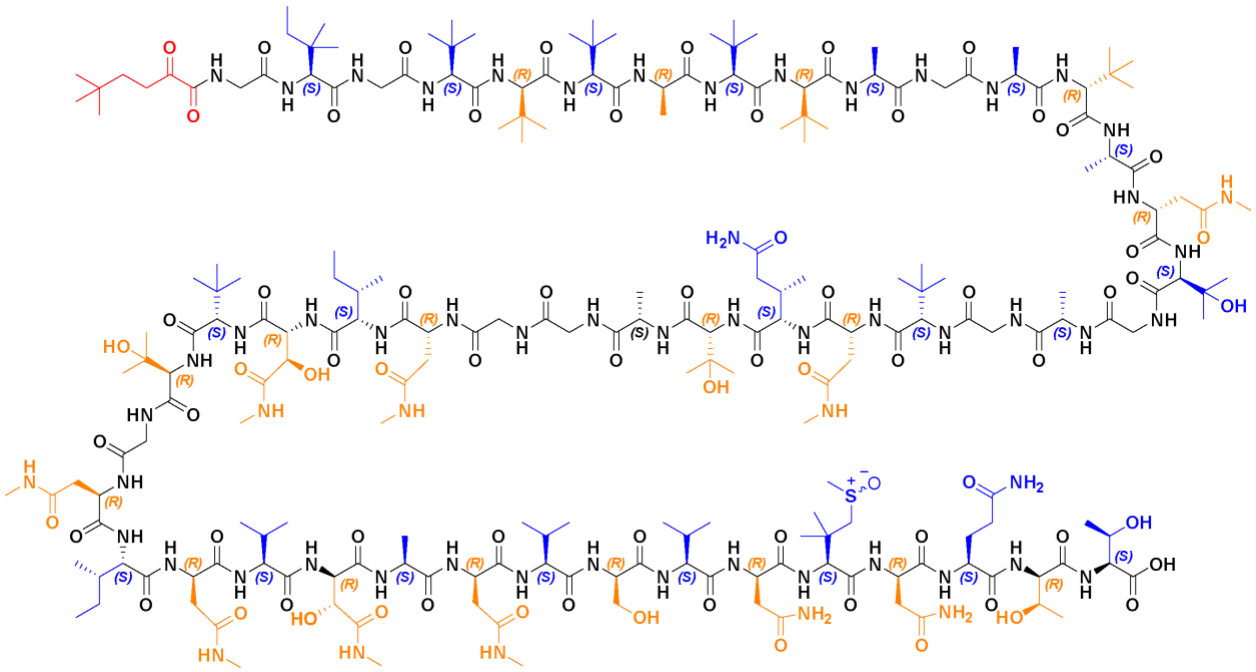


Figure 1.5. Structure of Polytheonamide B. The N-terminal 5,5-dimethyl-2-oxohexanoate is indicated in red. Alternating (R) and (S) stereochemistry are indicated in orange and blue, respectively. Non-proteinogenic amino acids composing polytheonamide B include eight γ-*N*-methyl aparagines, 13 β-methylated valine, threonine, isoleucine and glutamines as well as one β,β-dimethyl methionine oxide.

shares structural homology to the other polytheonamides.

The MOA of pTB was established early on when it was found that this peptide could act as a cationic membrane channel.[140] Detailed structural studies using solution NMR shed light on this mechanism as pTB adopts a single ~45 Å β-helical channel with a central cavity of ~4 Å, which is long enough to span the entire eukaryotic cell membrane.[141] This β-helix is stabilized by five sidechain N-methylated Asn and one Asn, which occur in a "stack" and are likely hydrogen bonded.[141] The N-terminal half of pTB is largely nonpolar while the C-terminus is relatively polar, leading to the hypothesis that pTB insertion into the membrane is facilitated by this

hydrophobic N-terminus (Fig 6). This structure was generated in a 1:1 methanol to chloroform solvent, though because the relative hydrophobicity of this solvent mixture is a reasonable facsimile to that of a cell membrane, this conformation is likely that of native pTB.[141] In synthetic planar membranes, pTB demonstrated insertion and permeabilization to K^+ and Cs^+ atoms, which occurred cyclically as voltage studies showed defined opening and closing events as measured by ammeter readings on either side of the membrane containing pTB.[142] The probability of an opening event was lower at -200 mV than at +200 mV, indicating a voltage-dependent gating mechanism.[142] These findings suggest an asymmetric pTB structure that orients in the membrane, as ion permeability was unidirectional, i.e. permeabilization of K^+ and Cs^+ ions only occurred in one direction across the membrane.[142] Transmembrane current amplitude was also directly proportional to pTB concentration, indicating that this RiPP

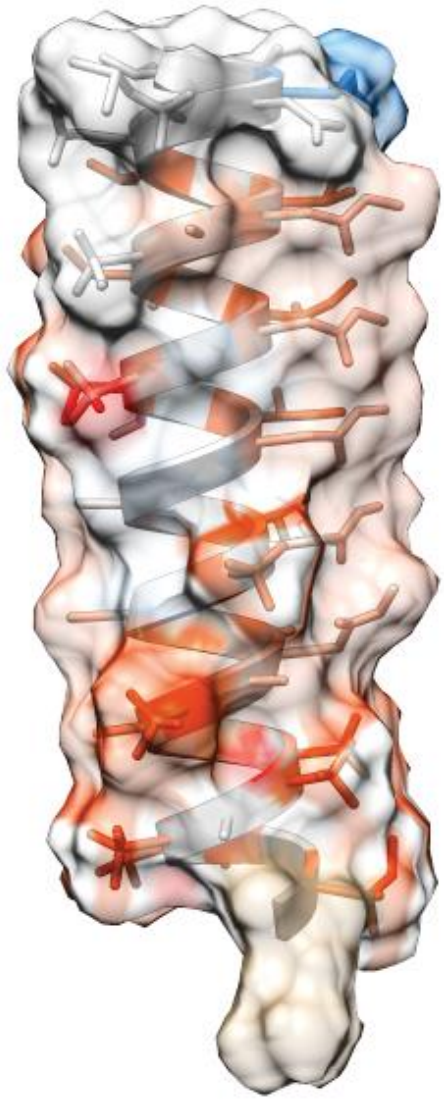


Figure 1.6. Surface hydrophobicity of solution NMR structure of polytheonamide B. Hydrophobic residues are colored orange and hydrophilic residues are colored blue. Polytheonamide B exhibits a relatively polar C-terminus (top) and a more nonpolar N-terminus (bottom). Residues with intermediate hydrophilicity are colored white.

forms a unimolecular ion channel instead of a channel composed of multiple pTB subunits.[142] These current recordings also demonstrated ion selectivity for monovalent cations of $Cs^+ > Rb^+ > K^+ > Na^+ > Li^+$, with no permeability to divalent cations and that pTB is not released from the membrane once it is incorporated.[142] The selective orientation of pTB is attributed to its asymmetric polarity, as the N-terminus is nonpolar relative to the hydrophilic C-terminus.[142] As planar membranes are not close models of cell membranes, liposomal studies were used to support the ion permeability mechanism of pTB *in situ*.[143] Liposomes were generated containing carboxyfluorescein as a pH indicator at pH 6.5 in a bulk solution of pH 5.5, and introduction of pTB dissipated this pH gradient leading to liposomal fluorescence despite the finding that pTB did not disrupt the membrane sufficiently to allow transmembrane passage of carboxyfluorescein, which has a molecular radius of ~5 Å.[143]

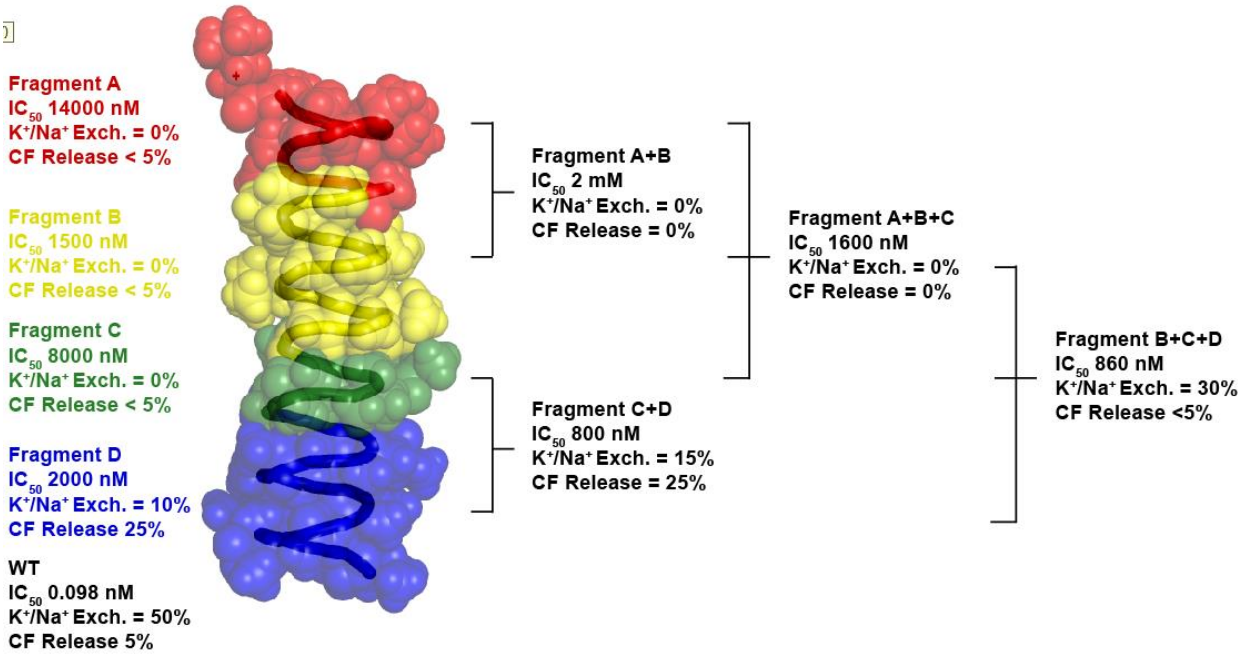


Figure 1.7 Bioactivity of Polytheonamide B fragments. The C-terminal Fragment D exhibited the highest membrane perturbation measured by carboxyfluorescein (CF) release, but lacked ion transport ability measured by K^+/Na^+ exchange. Conversely, N-terminal fragments lacked membrane perturbation activity. Additions of N-terminal fragments to Fragment D yielded a compound that lost membrane disruption activity but gained ion transport activity. Fragment IC_{50} values were determined *in vitro* against murine p388 leukemia cells. Ion exchange and carboxyfluorescein (CF) release values were estimated from graphical representations.

Synthetic studies have supported this mechanism and localized the membrane disruption activity to the C-terminus and membrane insertion/orientation to the hydrophobic N-terminus.[143–146] Synthetic truncants allowed localization of pTB function. C-terminal fragments possessed membrane disruption properties as treatment of liposome-encapsulated carboxyfluorescein induced carboxyfluorescein release upon treatment with pTB fragments D and C+D (Fig. 7).[143] These fragments also exhibited far lower H^+/Na^+ exchange across liposomal membranes. This

rate was increased as fragments were added from the C towards the N-terminus (Fig. 7).[143] In planar membrane channel current assays, The B+C+D fragment demonstrated a similar open/close gating mechanism to WT pTB, though with a lower voltage dependency.[143] These data support the general mechanism that the C-terminus functions as a membrane disruption domain, which is endowed with ion channel properties by the N-terminal fragments, which add sufficient length to span the entire lipid bilayer.[143] Conversion of the N-terminal acyl group (5,5-dimethyl-2-oxohexanamide) to the more polar amine, acetylamide, or trimethyl ammonium increased IC_{50} by 240-, 480-, or 2,500-fold respectively.[144] Conversely, introducing a hydrophobic octanamide functionality only decreased toxicity by 5-fold while a palmitamide decreased IC_{50} by a factor of three.[144] Current measurements in planar bilayers demonstrated ion channel activity for all variants except the trimethyl ammonium analog, which was proposed to be through a channel-blocking mechanism by the tertiary cation within this N-terminal modification.[144]

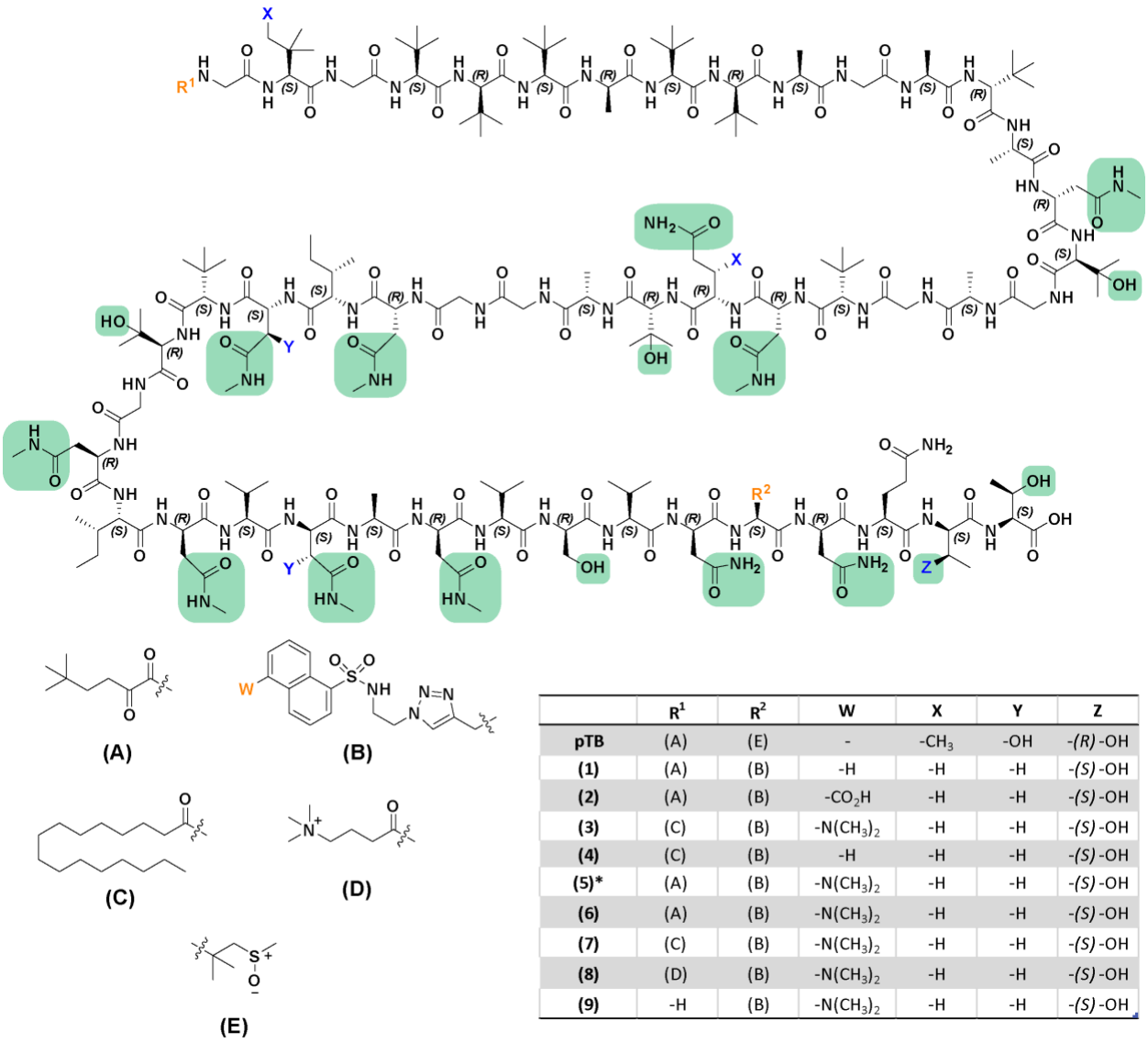
Derivitization studies have also been performed to elucidate SAR of pTB. Based on the solution NMR structures, a network of hydrogen bonding residues were identified as being critical for pore formation of pTB (Fig. 8). In efforts to generate a more synthetically accessible analog of pTB possessing a fluorescent reporter group, non-hydrogen bonding functional groups were manipulated. A total of six modifications were introduced into pTB, which resulted in analog that possessed an IC_{50} of 12 nM, approximately 100-fold higher than pTB (Fig. 8, Analog **(6)**).[147] The introduction of a C-terminal dansyl group served as a solvatochromic handle, and based on the UV-Vis excitation/emission Stokes shift of this group, this pTB analog was shown to properly insert into liposomal membranes.[147] *In vitro* analyses of liposomal H^+/Na^+ exchange

also demonstrated this mimic can efficiently permeabilize phospholipid bilayers to small cations with a gating mechanism similar to that of pTB.[147]

Truncation studies of the dansylated pTB analog demonstrated removal of 12 N-terminal residues yielded an analog with an IC_{50} of 3.7 nM against murine P388 mouse leukemia cells, compared to 12 nM against the full length dansylated pTB (Fig. 8, **(5)**).[148] Addition or removal of a single amino acid to this truncant led to compounds with IC_{50} values of >420 and 81 nM, respectively.[148] Ion transport assays across lipid bilayers in pH gradient liposomes demonstrated a significant reduction (~5-fold) in this truncants ability to permeabilize membranes to H^+/Na^+ cations.[148] Utilizing pH-indicator dyes, this analog was also unable to induce a pH shift in murine P388 leukemia cells, further supporting this compounds inability to induce membrane transport. This finding is consistent with truncation studies on native pTB, where C-terminal truncations reduced membrane disruption and cation transport ability (Fig. 8).[143] The potent cytotoxicity of the truncated dansyl-pTB despite its inability to efficiently transport cations suggest a novel mechanism of action for this pTB analog.[148]

Modification of the N-terminus of dansyl pTB further increased its potency. Introduction of a palmitamide further decreased the IC_{50} from 12 to 1.0 nM (Fig. 8, **(7)**).[146] This was attributed to the increased logP of this analog, which was determined to be 4.9, compared to 4.5 for pTB.[146] Introduction of the more polar N-terminal functional groups tetramethylammonium (Fig. 8, **(8)**) or the free amine (Fig. 8 **(9)**) increased the IC_{50} to 83 and 25 nM, respectively.[146] As before, this is likely due to the decreased logP of these compounds, which were determined to be 3.2 for the free amine and 2.6 for the trimethylammonium derivatives.[146] In addition to the hydrogen bonding network, the relative hydrophobicity of the N-terminus is also critical to the bioactivity of pTB.

Manipulation of the N-terminus and dasnyl moiety in combination further altered ion transport activity and cytotoxicity of pTB. Conversion of the dimethylaminonaphthyl to a hydrophilic naphthoic acid (Fix. 8, **(2)**) or a hydrophobic naphthyl (Fig. 8, **(1)**) group led to a decrease in IC_{50} to 6.3 and 0.54 nM, respectively.[145] Conversion of the N-terminal 5,5-dimethyl-2-oxohexanoate to a hydrophobic palmitate coupled with the naphthyl dansyl derivative (Fig. 8 **(4)**) yielded a pTB analog with an IC_{50} of 0.16 nM.[145] This derivative possessed a greater degree of hydrophobicity and exhibited a liposomal H^+/Na^+ exchange of 55%, compared to 48% for pTB.[145] Results from these studies demonstrate the cytotoxicity is largely dictated by the overall hydrophobicity of pTB, which facilitates its insertion into cell membranes.



	R^1	R^2	W	X	Y	Z
pTB	(A)	(E)	-	$-CH_3$	-OH	*-(R)*-OH
(1)	(A)	(B)	-H	-H	-H	*-(S)*-OH
(2)	(A)	(B)	$-CO_2H$	-H	-H	*-(S)*-OH
(3)	(C)	(B)	$-N(CH_3)_2$	-H	-H	*-(S)*-OH
(4)	(C)	(B)	-H	-H	-H	*-(S)*-OH
(5)*	(A)	(B)	$-N(CH_3)_2$	-H	-H	*-(S)*-OH
(6)	(A)	(B)	$-N(CH_3)_2$	-H	-H	*-(S)*-OH
(7)	(C)	(B)	$-N(CH_3)_2$	-H	-H	*-(S)*-OH
(8)	(D)	(B)	$-N(CH_3)_2$	-H	-H	*-(S)*-OH
(9)	-H	(B)	$-N(CH_3)_2$	-H	-H	*-(S)*-OH

Figure 1.8. Chemical structure of polytheonamide B. Residues contributing to the network of stabilizing hydrogen bonds are highlighted in green. Sites not involved in structure stabilization were selected for derivatization. *Analog (5) refers to the indicated modifications in addition to removal of the 11 N-terminal residues.

Structural features of pTB have been further studied using a combination of the published 3D NMR structure and molecular dynamics. A network of hydrogen bonds between Asn and

methyl-Asn sidechains serves to stabilize the channel structure, allowing polytheonamide B to be "trapped" in the membrane and retain channel activities over long periods, contributing to this compounds potency.[141,142,149] Ion transport through the channel is proposed to occur through coordination of cations by backbone carbonyls and accompanying conformational change and is facilitated by the movement of water molecules through the channel.[150,151] The waters residing in pTB also function as a water wire, which serves to transport H^+ cations during membrane depolarization.[152] This ion transport mechanism was demonstrated to be reversible depending on the voltage potential, and proton transport is at equilibrium at a voltage potential of ~0mV, confirming pTB's ability to depolarize cell membranes regardless of potential direction.[152] Recent work has further expanded the MOA of pTB to explain its exceptional cytotoxicity. In addition to membrane depolarization, cell localization studies also demonstrated the ability of pTB to neutralize lysosomes through a micropinocytosis pathway.[153,154] A synthetic pTB compound was generated with an N-terminal BODIPY functional group, facilitating studies of pTB in live cells.[153] The membrane potential dye Bis-(1,3-dibutylbarbituric acid) trimethine oxonol ($DiBAC_4(3)$) is used to study changes in membrane potential, as the dye can enter cells and fluoresce upon membrane depolarization. Treatment of $DiBAC_4(3)$ exposed cells with pTB led to complete membrane depolarization in MCF-7 cells within 1hr at 5 nM pTB.[153] Imaging studies of the BODIPY-functionalized pTB demonstrated that this compound did not localize to the plasma membrane but instead fluoresced within smaller organelles.[153] Colocalization studies with LysoTracker (which accumulates in acidic lysosomes) demonstrated BODIPY-pTB was located within the lysosomal membranes.[153] The lysosomal pH-indicator fluorescein-tetramethylrhodamine-tagged dextran (FRD) demonstrated a pH increase from 4 to 7 in lysosomes after treatment with pTB, presumably as a consequence of proton transport by pTB.[153]

Internalization of pTB was shown to be reduced by 5-(*N*-ethyl-*N*-isopropyl)-amiloride which is an inhibitor of the micropinocytosis, suggesting pTB is actively internalized by a native cellular pathway.[153] Flow cytometry studies with pTB demonstrated induction of the apoptotic pathway due to membrane depolarization and dissipation of the lysosomal pH gradient.[153] of This secondary effect was observed at concentrations near the measured IC_{50} values, supporting the hypothesis that this secondary MOA is not artifactual.

Recent work aimed at expanding the diversity of the proteusins has led to the discovery of several new families. The aeronamides and polygeonamides share gene cluster and precursor homology, as well as cytotoxicity, with the polytheonamides and may thus share an MOA.[155] A very recent report details the discovery of the landornamides, which possess a nitrile hydratase homologous leader peptide but have a markedly different architecture than the other proteusins.[156] The landornamides also have demonstrated anti-arenaviral bioactivity, thus suggesting a completely novel MOA for the proteusin class and necessitating further study.[156]

1.6 Bottromycins

The bottromycin class is defined by the presence of a macrocyclic amidine, decarboxylated C-terminal thiazole as well as several C- and O-methylated amino acids. Potent antibacterial activity has been demonstrated against *Mycoplasma* and Gram-positive bacteria, including MRSA and VRE.[157–159] Mechanistic work on the bottromycins was carried out primarily by studying bottromycin A_2 (hereby referred to as bottromycin), which was among the first two isolated members of this family.[160] The MOA of bottromycin was elucidated using ^{14}C and ^{3}H labelling studies, and was originally demonstrated to be by inhibition of protein synthesis.[161] Interestingly, significant translation inhibition was observed only for poly-C polynucleotides encoding polyproline. This effect was not observed for poly-A and poly-U polynucleotides, and

it was demonstrated that bottromycin did not affect amino acid acylation, tRNA activation or ribosomal assembly. The first proposed mechanism of bottromycin inhibition was by a direct interaction with ribosome-bound mRNA in a nucleotide specific manner.[161] Further studies demonstrated inhibition of *in vitro* translation of poly-UC and poly-UG polynucleotides. This inhibition was irrespective of mRNA concentration, but dependent on ribosome concentration.[162] This study also found no influence of puromycin-dependent peptide release, refining the MOA to be inhibition of translocation, and not transpeptidation, by directly binding the ribosome.[162] This theory was supported by a concurrent study which demonstrated poly-A translation inhibition by bottromycin at low ribosome concentration, and also elaborated that puromycin-dependent peptide release could be partially inhibited by bottromycin in the presence of EF-G and GTP.[163] As bottromycin was suspected to inhibit translocation during translation, inclusion of bottromycin would prevent this step in the presence of EF-G and GTP, thus trapping the polypeptide in the A site and inhibiting puromycin release from the P site. This mechanism was subsequently supported by findings that bottromycin bound the 50S ribosomal subunit.[164]

A series of publications by Sydney Pestka investigating the mechanisms of ribosomal translation served to complicate the accepted mechanism of bottromycin inhibition. Building on earlier studies demonstrating bottromycin failed to inhibit puromycin-dependent peptide release in protopasts,[165] Pestka claimed that the peptidyl transferase reaction was not the target of bottromycin as it failed to substantially inhibit puromycin-dependent peptide release.[166] Concurrent studies also demonstrated bottromycin inhibited di- and triphenylalanine translation equally in the presence and absence of EF-G and GTP, though was able to nearly completely inhibit polyphenylalanine (>3 Phe) translation at a lower concentration.[167] Inhibition of phenyl-

puromycin and oligophenylalanine formation by bottromycin was also shown to be equivalent, which was used to argue that translocation was perhaps not the only step of inhibition.[168] Inhibition of translocation was further supported in *in vitro* translation studies demonstrating bottromycin partially inhibited the puromycin reaction with both N-formylmethionine and polylysine.[169] This was supported by a subsequent study demonstrating bottromycin did not inhibit the puromycin reaction with native *E. coli* polysomes.[170] These findings are consistent with the previously proposed mechanism, where bottromycin prevents translocation to the P site, and thus blocks puromycin activity. Studies by Pestka involving native polysomes exhibited no inhibition by bottromycin, as these preparations would include any number of ribosomes with polypeptidyl tRNA in either the A or P site.[170]

Significant refinement of the MOA of bottromycin was provided in a series of publications by Otaka & Kaji. In these studies, it was demonstrated that bottromycin released peptidyl- and aminoacylated-tRNA from the ribosome A site, and N-acetylphenylalanine-tRNA was insensitive to bottromycin under non-enzymatic translocation conditions. Furthermore, aminoacyated tRNA release was equivalent to general translation inhibition induced by micrococcin.[171] Conflicting results regarding the incomplete inhibition of polylysyl- and f-Met-puromycin by bottromycin[163,169] with complete inhibition of N-acetylphenylalanine-puromycin formation[168] were resolved by demonstration that bottromycin reduces the affinity between the A site and puromycin.[172] Finally, bottromycin's seeming inability to inhibit protein synthesis in nascent polysomes[170] was explained by demonstrating that bottromycin prevented proper binding of charged tRNAs to the A site, but was unable to release large oligopeptidyl-tRNAs from the 50S subunit due to an increased binding affinity.[173] The conclusion of these ~20 years of research established the MOA of bottromycin to be by binding the ribosomal A site to prevent

recognition of charged tRNAs. Furthermore, no structural insights into this interaction exists, nor do studies of other members of this class both of which represent avenues for future research.

1.7 Thiopeptides

Thiopeptides are defined by the presence of a six-membered nitrogenous heterocycle forming a peptidic macrocycle which can also bear several thioazole rings and dehydrated amino acids. Isolated thiopeptides exhibit potent antibacterial activity against Gram positive bacteria, as well as activity against *Plasmodium falciparum* and the eukaryotic FOXM1 transcription factor. Broadly speaking, the mechanisms of action of the thiopeptides can be divided into four main categories: 1) Ribosomal RNA (rRNA) L11 binding domain (L11BD) binding, 2) Elongation Factor Thermo unstable (EF-Tu) binding, 3) TipAL transcription factor binding and 4) FOXM1 transcription factor inhibition. Several anti-eukaryotic mechanisms exist as well, including antimalarial activity and anti-transcription factor activity.

The prototypical thiopeptide thiostrepton was known early on as an inhibitor of protein synthesis through interaction with the 50S ribosomal subunit and subsequent prevention of amino-acylated tRNA, Elongation Factor G (EF-G) and GTP binding to the ribosome.[174–178] Thiostrepton acts as a translocation inhibitor by preventing the binding of EF-G and GTP to the ribosome and inhibiting the dissociation of the translation complex.[175,179–182] The exact site of thiostrepton binding to the 50S ribosomal subunit was identified using the nascent resistance mechanism of the producer *S. azureus* as well as the mechanism of spontaneous resistance mutants. Identification of a methyltransferase responsible for methylation of an adenine ribose, which granted resistance the thiostrepton producer, led to the hypothesis that this compound targeted an rRNA component of the 50S subunit.[183,184] This thiostrepton resistance mechanism was also observed in several other streptomycetes, as well as two producers of the structurally related

thiopeptides siomycin and sporangiomycin.[185] In both *B. subtilis* and *B. megaterium* resistant mutants, bacteria were missing the L11 protein, a critical component of the 50S subunit, which was also shown to be necessary for thiostrepton inhibition of the *E. coli* translational machinery in vitro.[186–190] *B. subtilis* resistant mutants have also been generated containing mutations to A1067 of the 23S rRNA and Pro 25 of L11.[191] Taken in isolation, these results suggested that thiostrepton interacted with both proteinaceous and RNA components of the ribosome, either together or in isolation.

These seemingly conflicting targets were rationalized with the findings that L11 bound the 23S rRNA along the same sequence that is targeted for methylation or mutation in resistant organisms.[192–194] It was subsequently demonstrated that thiostrepton also bound within this rRNA sequence, and so it was hypothesized that thiostrepton interfered with this native interaction.[195] This mechanism was proposed to be through allosteric cooperativity, as thiostrepton inhibited EF-G dependent and independent translocation and had no effect on the P site binding of tRNAs.[196] An allosteric effect was further supported as micrococcin P, a related thiopeptide, bound the same region of the 23S rRNA but stimulated EF-G GTPase activity; opposite the effect of thiostrepton.[195,197] Structural and mutational studies of the 23S rRNA subunit demonstrated significant conformational change in the 3D structure upon binding L11, stabilizing select conformations and allowing for the cooperative binding of thiostrepton.[197–200] Conformational studies of the L11BD highlighted that L11 binding brought the TMBD loops in close proximity, forming a "slot" in which thiostrepton and micrococcin insert to further stabilize the bound conformation.[201–204] Binding of thio/micro within this slot, at mid- to low-pM affinity, prevented conformational transitions of L11 necessary for translation.[205–209] Furthermore, thiostrepton was shown to stabilize the same 23S rRNA L11BD conformation as L11 itself, and

additional interactions between the L11 N-terminal domain (NTD) and thio/micro contribute to further stabilization of this complex and prevention of downstream conformational changes and interactions with EF-G and RRF necessary for translation.[205,210–214] Prevention of these conformational changes has been shown to abrogate complexation of EF-G with amino-acylated tRNA during elongation.[215] Studies on resistant *Thermus thermophilus* L11 mutants demonstrated that amino acid substitutions served to increase the flexibility of the L11 protein, potentially overcoming the conformational stabilizing effect of thiostrepton.[216] These findings were further supported by NMR studies of *Thermotoga maritima* L11 mutants, demonstrating a compact, stable L11-thio-rRNA structure and mutations at conserved Pro sites served to increase L11 flexibility to function in the presence of thio.[217]

Partial elucidation of the disparate activities of thio/micro on the GTPase center highlighted that thio increases the dissociation rate of GTP, thus retarding GTP hydrolysis and micro increases the dissociation of GDP+P_i from EF-G, subsequently increasing the rate of GTP hydrolysis.[211] The additional effect of altered GTP hydrolysis is that thio prevents EF-G turnover by the ribosome, while micro increases it, which is a consequence of the presence (thio) or absence (micro) of a quinaldic acid heterocycle, which directly interacts with the L11BD rRNA at A1067.[218,219] In addition to elongation steps, thio/micro also possesses inhibitory activity toward translation initiation by destabilizing the interactions between the ribosome, tRNA and initiation factor 2.[220–225]

Thiostrepton and related thiopeptides were shown to induce the expression of several proteins in *Streptomyces*, including the *tipA* locus.[226] The $TipA_L$ protein C-terminus directly interacts with thiostrepton to induce expression by binding the *tipA* promoter through the N-terminal DNA binding domain which endows *S. lividans* with resistance to a diverse set of antibiotics including

daunorubicin, sparsomycin and tetranactin.[227–229] The $TipA_L$ thiostrepton binding domain is expressed as a separate protein, termed $TipA_S$ which is expressed at a ~20-fold excess.[227] Induction of *tipA* has also been used to screen extracts for several novel thiopeptides.[230–233] Detailed studies of TipAS demonstrated the formation of a covalent linkage between thio and TipAS/AL Cys214, which was tentatively assigned a role in irreversible phenotyptic developments e.g. sporulation and mycelial autolysis.[234–236] Subsequent study demonstrated both TipAS and TipAL covalently bind thio, but TipAL induces the *tipA* promoter and TipAS functions as an antibiotic sequestration resistance protein.[237–239] Binding of thiostrepton induces a disorder-order transition within the antibiotic binding cleft of TipAS.[240] The disordered nature of TipAS allows the tipA system to recognize and sequester a diverse set of thiopeptides.[240] Homologous *tipA* operons have also been found in distantly related thiopeptide non-producers, implicating this operon in low-level/transient resistance.[241–243]

The discovery that extrachromosomal organellar DNA in *Plasmodium sp.* plastids encode rRNAs inspired efforts to target *Plasmodium falciparum* organellar protein synthesis with the thiopeptide thiostrepton.[244,245] Conservation of the nucleotide A1067 in the plasmodial 23S rRNA allows the binding of thiostrepton to this plastid rRNA, terminating organellar, but not nuclear, protein translation and arresting life stage transition from trophozite to schizont.[246–248] Thiostrepton was demonstrated to induce tertiary conformational change in plastid rRNA fragments analogous to those occurring in bacterial rRNA, despite only 60% sequence identity between *P. falciparum* and *E. coli* rRNA.[249] Analogously to thiostrepton resistance in bacteria, a single nucleotide substitution at the A1067-equivalent position in the plastid rRNA abrogated the binding of thiostrepton.[249] Similar anti-*Plasmodium* bioactivity has been observed for micrococcin P, and is presumed to occur by a similar mechanism to thiostrepton.[250] Bioactivity

of thiostrepton against *Babesia* parasites, which also possess a plastid organelle, has also been demonstrated.[251] Further developments on the anti-malarial activity of thiostrepton demonstrated in addition to apicoplast translation inhibition, this thiopeptide also inhibited the *P. falciparum* 20S proteasome, explaining the rapid killing effect of thiostrepton.[252]

The ability of thiopeptides to target the translation apparatus was further expanded with the discovery of GE2270.[253] Mechanistic studies on this compound demonstrated it bound EF-Tu, theoretically by mimicking an aminoacylated tRNA, locking EF-Tu in a "non-productive" conformation which couldn't bind GTP or tRNA.[254,255] The site of action of GE2270 was further shown to be unique among EF-Tu targeting antibiotics, e.g. kirromycin, as both could simultaneously bind EF-Tu.[256] Subsequent studies of the native resistance mechanism of the GE2270 producer *Planobispora rosea* demonstrated this organism possessed an alternate EF-Tu (EF-Tu1) which was completely resistant to several classes of antibiotic, including thiopeptides, and contained only a modest mutation of seven residues (of 393 total).[257] Structurally related thiopeptides, termed amythiamicins, also inhibited protein synthesis and resistance mutants possessed a single valine-to-alanine mutation in EF-Tu.[258] As with thiostrepton and micrococcin, GE2270 also possessed bioactivity against *P. falciparum*, which contains an EF-Tu protein encoded by the aforementioned plastid organelle.[259]

Molecular level details dictating the GE2270 mechanism were illuminated by the cocrystal structure of GE2270:EF-Tu.[260] Conformation comparisons demonstrated that GE2270A-occupied space caused steric clashes with the GTP-bound EF-Tu conformation.[260] These steric clashes prevented EF-Tu from transitioning between conformations necessary for translation.[260] Subsequent structural work also demonstrated EF-Tu-bound GE2270A generated steric clashes with the 3'-aminoacyl group and acceptor stem of charged tRNA.[261]

Bioactivity of the thiopeptides has also been observed against human cancer cell lines.[262–266] Sensitization of cancer cells was originally shown to be through down-regulation of cyclooxegenase I which induced mitochondrial oxidative stress.[263] Additional anticancer MOA studies of thiostrepton was shown to be by inhibition of bovine mitochondrial protein translation *in vitro*.[267] Subsequent studies highlighted another potential target of thiostrepton to be the transcription factor FOXM1.[265,266] Conclusive studies demonstrated direct binding of thiostrepton to FOXM1, which prevented interaction of FOXM1 with target promoters.[268] Thioamides are compounds that contain one or more backbone thioxoamides; these modifications are found in several classes of RiPPs, hereby referred to as the thioamitides. The first thioamitides discovered were the thioviridamides[269] The structural diversity of the thioamitides is reflected in their mechanisms of action, which impact both primary metabolism and antibiotic activities.

1.8 Thioviridamides

Of the thioviridamides, isolated members have bioactivity against several human cell lines, especially E1A transformed cells.[269–275] Recent studies have confirmed that thioviridamide is an acetone adduct of the true natural product prethioviridamide (pTvr), though both compounds have comparable cytotoxic activities.[273,275] Preliminary mechanistic studies on the thioamitide thioalbamide (Tab) demonstrate Tab's ability to arrest the cell cycle at G1 phase and induce apoptosis.[274] Furthermore, Tab activates intrinsic and extrinsic apoptotic pathway, as observed through DNA fragmentation, externalization of phosphatidylserine to recruit macrophages, loss of mitochondrial membrane potential, ejection of cytochrome c from mitochondria and activation of initiator caspases 8 and 9.[274] The activity of Tab was further localized to the mitochondria, where it was shown to effect the production of reactive oxygen species and induces

overexpression of the mitochondrial superoxide dismutase isoform SOD2.[274] Exposure to Tab *in vitro* led to a >60% reduction of glycolysis and oxidative respiration in breast cancer cell lines.[274] A general mechanism of metabolic disruption is further supported by the fact that Tab induced similar metabolic deficiencies in several biologically heterogenous breast cancer cell lines as well as cancer stem cells.[274]

In a complementary study, Takase et al. demonstrated the cytotoxicity of pTvr is attributed to activation of the integrated stress response (ISR).[276] Cellular morphology studies confirmed pTvr induced significant morphological changes and dissipation of mitochondrial membrane potential, a time-dependent reduction in O_2 consumption as well as global translation inhibition.[276] Screening of shRNA libraries further illuminated pTvr activation of the GCN2-ATF4 metabolic stress response pathway, which is central to the ISR and apoptotic induction.[276,277] Photoaffinity pulldowns identified the target of pTvr as the ATP5B subunit of the mitochondrial F_1 ATP synthase.[276] Inhibition of ATPase activity was confirmed *in vitro,* and it was further demonstrated that inhibition of F_1F_o-ATPase was responsible for the upregulation of the GCN2-ATF4 pathway.[276]

Incorporating mechanistic findings of both Tab and pTvr indicates perturbation of oxidative respiration in cancer cells, though a comprehensive understanding of the mechanism is lacking. The exact molecular-level details regarding pTvr-ATP5B interaction is unknown. Recent derivatization studies have identified pTvr analogs with improved cytotoxicity, a course of study which could be vastly enhanced by an understanding of pTvr's structural interactions.[278] Additionally, how both Tab and pTvr can induce apoptosis through disruption of mitochondrial respiration requires additional analyses. Furthermore, the higher sensitivity to E1A-transformed

cancer cells to thioviridamides is thought to occur as E1A can amplify GCN2-ATF4 signaling, but a detailed understanding of these complex interactions requires further study.[276]

1.9 References

(1) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; Donk, W. A. van der. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2012**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(2) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; Heel, A. J. van; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Burgos, M. J. G.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; Wezel, G. P. van; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; Donk, W. A. van der. New Developments in RiPP Discovery, Enzymology and Engineering. *Nat. Prod. Rep.* **2021**, *38* (1), 130–239. https://doi.org/10.1039/D0NP00027B.

(3) Weber, W.; Fischli, W.; Hochuli, E.; Kupfer, E.; Weibel, E. K. ANANTIN - A PEPTIDE ANTAGONIST OF THE ATRIAL NATRIURETIC FACTOR (ANF). *J. Antibiot.* **1991**, *44* (2), 164–171. https://doi.org/10.7164/antibiotics.44.164.

(4) Wyss, D. F.; Lahm, H.-W.; Manneberg, M.; Labhardt, A. M. ANANTIN - A PEPTIDE ANTAGONIST OF THE ATRIAL NATRIURETIC FACTOR (ANF). *J. Antibiot.* **1991**, *44* (2), 172–180. https://doi.org/10.7164/antibiotics.44.172.

(5) Morishita, Y.; Chiba, S.; Tsukuda, E.; Tanaka, T.; Ogawa, T.; Yamasaki, M.; Yoshida, M.; Kawamoto, I.; Matsuda, Y. RES-701-1, A NOVEL AND SELECTIVE ENDOTHELIN TYPE B RECEPTOR ANTAGONIST PRODUCED BY Streptomyces Sp. RE-701. *J. Antibiot.* **1994**, *47* (3), 269–275. https://doi.org/10.7164/antibiotics.47.269.

(6) Ogawa, T.; Ochiai, K.; Tanaka, T.; Tsukuda, E.; Chiba, S.; Yano, K.; Yamasaki, M.; Yoshida, M.; Matsuda, Y. RES-701-2, -3 and -4, Novel and Selective Endothelin Type B Receptor Antagonists Produced by Streptomyces Sp. *J. Antibiot.* **1995**, *48* (11), 1213–1220. https://doi.org/10.7164/antibiotics.48.1213.

(7) Katahira, R.; Shibata, K.; Yamasaki, M.; Matsuda, Y.; Yoshida, M. Solution Structure of Endothelin B Receptor Selective Antagonist RES-701-1 Determined by 1H NMR Spectroscopy. *Bioorganic & Medicinal Chemistry* **1995**, *3* (9), 1273–1280. https://doi.org/10.1016/0968-0896(95)00122-W.

(8) He, J. X.; Cody, W. L.; Flynn, M. A.; Welch, K. M.; Reynolds, E. E.; Doherty, A. M. Res-701-1, Synthesis and a Reevaluation of Its Effects on the Endothelin Receptors. *Bioorganic & Medicinal Chemistry Letters* **1995**, *5* (6), 621–626. https://doi.org/10.1016/0960-894X(95)00084-7.

(9) Katahira, R.; Shibata, K.; Yamasaki, M.; Matsuda, Y.; Yoshida, M. RES-701-1, Comparative Study of the Synthetic and the Microbial-Origin Compounds. *Bioorganic & Medicinal Chemistry Letters* **1995**, *5* (15), 1595–1600. https://doi.org/10.1016/0960-894X(95)00264-T.

(10) Shibata, K.; Yano, K.; Tanaka, T.; Matsuda, Y.; Yamasaki, M. C-Terminal Modifications of an Endothelin Antagonist RES-701-1: Production of ETA/ETB Dual Selective Analogs. *Letters in Peptide Science* **1997**, *4* (3), 167–170. https://doi.org/10.1023/A:1008887108419.

(11) Shibata, K.; Yano, K.; Tanaka, T.; Matsuda, Y.; Yamasaki, M. Analogs of an Endothelin Antagonist RES-701-1: Substitutions of C-Terminal Amino Acid. *Bioorganic & Medicinal Chemistry Letters* **1996**, *6* (7), 775–778. https://doi.org/10.1016/0960-894X(96)00127-8.

(12) Endothelin ETB Receptor Antagonist, RES-701-1: Effects on Isolated Blood Vessels and Small Intestine. *European Journal of Pharmacology* **1994**, *262* (3), 255–259. https://doi.org/10.1016/0014-2999(94)90739-0.

(13) Ikemura, T.; Ohmori, K.; Tanaka, T.; Matsuda, Y.; Kitamura, S. Effects of Endothelin B Antagonist RES−701−1 on Endothelin-Induced Contractile Responses In-Vivo and In-Vitro in Guinea-Pigs. *Journal of Pharmacy and Pharmacology* **1996**, *48* (1), 100–105. https://doi.org/10.1111/j.2042-7158.1996.tb05886.x.

(14) Potterat, O.; Wagner, K.; Gemmecker, G.; Mack, J.; Puder, C.; Vettermann, R.; Streicher, R. BI-32169, a Bicyclic 19-Peptide with Strong Glucagon Receptor Antagonist Activity from Streptomyces Sp. *J. Nat. Prod.* **2004**, *67* (9), 1528–1531. https://doi.org/10.1021/np040093o.

(15) Knappe, T. A.; Linne, U.; Xie, X.; Marahiel, M. A. The Glucagon Receptor Antagonist BI-32169 Constitutes a New Class of Lasso Peptides. *FEBS Letters* **2010**, *584* (4), 785–789. https://doi.org/10.1016/j.febslet.2009.12.046.

(16) Chen, M.; Wang, S.; Yu, X. Cryptand-Imidazolium Supported Total Synthesis of the Lasso Peptide BI-32169 and Its d -Enantiomer. *Chemical Communications* **2019**, *55* (23), 3323–3326. https://doi.org/10.1039/C8CC10301A.

(17) Esumi, Y.; Suzuki, Y.; Itoh, Y.; Uramoto, M.; Kimura, K.-I.; Goto, M.; Yoshihama, M.; Ichikawa, T. Propeptin, a New Inhibitor of Prolyl Endopeptidase Produced by Microbispora. *J. Antibiot.* **2002**, *55* (3), 296–300. https://doi.org/10.7164/antibiotics.55.296.

(18) Kimura, K.; Kanou, F.; Yamashita, Y.; Yoshimoto, T.; Yoshihama, M. Prolyl Endopeptidase Inhibitors Derived from Actinomycetes. *Bioscience, Biotechnology, and Biochemistry* **1997**, *61* (10), 1754–1756. https://doi.org/10.1271/bbb.61.1754.

(19) Kimura, K.; Yamazaki, M.; Sasaki, N.; Yamashita, T.; Negishi, S.; Nakamura, T.; Koshino, H. Novel Propeptin Analog, Propeptin-2, Missing Two Amino Acid Residues from the Propeptin C -Terminus Loses Antibiotic Potency. *The Journal of Antibiotics* **2007**, *60* (8), 519–523. https://doi.org/10.1038/ja.2007.66.

(20) Gavrish, E.; Sit, C. S.; Cao, S.; Kandror, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K. Lassomycin, a Ribosomally Synthesized Cyclic Peptide, Kills Mycobacterium Tuberculosis by Targeting the ATP-Dependent Protease ClpC1P1P2. *Chemistry & Biology* **2014**, *21* (4), 509–518. https://doi.org/10.1016/j.chembiol.2014.01.014.

(21) Leodolter, J.; Warweg, J.; Weber-Ban, E. The Mycobacterium Tuberculosis ClpP1P2 Protease Interacts Asymmetrically with Its ATPase Partners ClpX and ClpC1. *PLOS ONE* **2015**, *10* (5), e0125345. https://doi.org/10.1371/journal.pone.0125345.

(22) T. Malik, I.; Brötz-Oesterhelt, H. Conformational Control of the Bacterial Clp Protease by Natural Product Antibiotics. *Natural Product Reports* **2017**, *34* (7), 815–831. https://doi.org/10.1039/C6NP00125D.

(23) Lear, S.; Munshi, T.; S. Hudson, A.; Hatton, C.; Clardy, J.; A. Mosely, J.; J. Bull, T.; S. Sit, C.; L. Cobb, S. Total Chemical Synthesis of Lassomycin and Lassomycin-Amide. *Organic & Biomolecular Chemistry* **2016**, *14* (19), 4534–4541. https://doi.org/10.1039/C6OB00631K.

(24) Tsunakawa, M.; Hu, S.-L.; Hoshino, Y.; Detlefson, D. J.; Hill, S. E.; Furumai, T.; White, R. J.; Nishio, M.; Kawano, K.; Yamamoto, S.; Fukagawa, Y.; Oki, T. Siamycins I and II, New Anti-HIV Peptides: I. Fermentation, Isolation, Biological Activity and Initial Characterization. *J. Antibiot.* **1995**, *48* (5), 433–434. https://doi.org/10.7164/antibiotics.48.433.

(25) Oka, M.; Iimura, S.; Tenmyo, O.; Sawada, Y.; Sugawara, M.; Ohkusa, N.; Yamamoto, H.; Kawano, K.; Hu, S. L.; Fukagawa, Y. Terpestacin, a New Syncytium Formation Inhibitor from Arthrinium Sp. *J Antibiot (Tokyo)* **1993**, *46* (3), 367–373. https://doi.org/10.7164/antibiotics.46.367.

(26) Lin, P. F.; Samanta, H.; Bechtold, C. M.; Deminie, C. A.; Patick, A. K.; Alam, M.; Riccardi, K.; Rose, R. E.; White, R. J.; Colonno, R. J. Characterization of Siamycin I, a Human Immunodeficiency Virus Fusion Inhibitor. *Antimicrob. Agents Chemother.* **1996**, *40* (1), 133–138.

(27) Nar, H.; Schmid, A.; Puder, C.; Potterat, O. High-Resolution Crystal Structure of a Lasso Peptide. *ChemMedChem* **2010**, *5* (10), 1689–1692. https://doi.org/10.1002/cmdc.201000264.

(28) Helynck, G.; Dubertret, C.; Mayaux, J.-F.; Leboul, J. ISOLATION OF RP 71955, A NEW ANTI-HIV-1 PEPTIDE SECONDARY METABOLITE. *J. Antibiot.* **1993**, *46* (11), 1756–1757. https://doi.org/10.7164/antibiotics.46.1756.

(29) Frechet, D.; Guitton, J. D.; Herman, F.; Faucher, D.; Helynck, G.; Monegier du Sorbier, B.; Ridoux, J. P.; James-Surcouf, E.; Vuilhorgne, M. Solution Structure of RP 71955, a New 21 Amino Acid Tricyclic Peptide Active against HIV-1 Virus. *Biochemistry* **1994**, *33* (1), 42–50. https://doi.org/10.1021/bi00167a006.

(30) Tan, S.; Ludwig, K. C.; Müller, A.; Schneider, T.; Nodwell, J. R. The Lasso Peptide Siamycin-I Targets Lipid II at the Gram-Positive Cell Surface. *ACS Chem. Biol.* **2019**, *14* (5), 966–974. https://doi.org/10.1021/acschembio.9b00157.

(31) Katahira, R.; Yamasaki, M.; Matsuda, Y.; Yoshida, M. MS-271, A Novel Inhibitor of Calmodulin-Activated Myosin Light Chain Kinase from Streptomyces Sp.—II. Solution Structure of MS-271: Characteristic Features of the 'Lasso' Structure. *Bioorganic & Medicinal Chemistry* **1996**, *4* (1), 121–129. https://doi.org/10.1016/0968-0896(95)00176-X.

(32) Yano, K.; Toki, S.; Nakanishi, S.; Ochiai, K.; Ando, K.; Yoshida, M.; Matsuda, Y.; Yamasaki, M. MS-271, a Novel Inhibitor of Calmodulin-Activated Myosin Light Chain Kinase from Streptomyces Sp.—I. Isolation, Structural Determination and Biological Properties of MS-271. *Bioorganic & Medicinal Chemistry* **1996**, *4* (1), 115–120. https://doi.org/10.1016/0968-0896(95)00175-1.

(33) Metelev, M.; Tietz, J. I.; Melby, J. O.; Blair, P. M.; Zhu, L.; Livnat, I.; Severinov, K.; Mitchell, D. A. Structure, Bioactivity, and Resistance Mechanism of Streptomonomicin, an Unusual Lasso Peptide from an Understudied Halophilic Actinomycete. *Chemistry & Biology* **2015**, *22* (2), 241–250. https://doi.org/10.1016/j.chembiol.2014.11.017.

(34) Dhiman, A.; Bhatnagar, S.; Kulshreshtha, P.; Bhatnagar, R. Functional Characterization of WalRK: A Two-Component Signal Transduction System from Bacillus Anthracis. *FEBS Open Bio* **2014**, *4*, 65–76. https://doi.org/10.1016/j.fob.2013.12.005.

(35) Dubrac, S.; Msadek, T. Tearing Down the Wall: Peptidoglycan Metabolism and the WalK/WalR (YycG/YycF) Essential Two-Component System. In *Bacterial Signal Transduction: Networks and Drug Targets*; Utsumi, R., Ed.; Advances in Experimental Medicine and Biology; Springer: New York, NY, 2008; pp 214–228. https://doi.org/10.1007/978-0-387-78885-2_15.

(36) Hashimoto, M.; Ooiwa, S.; Sekiguchi, J. Synthetic Lethality of the LytE CwlO Genotype in Bacillus Subtilis Is Caused by Lack of d,l-Endopeptidase Activity at the Lateral Cell Wall. *Journal of Bacteriology* **2012**, *194* (4), 796–803. https://doi.org/10.1128/JB.05569-11.

(37) Nakayama, J.; Tanaka, E.; Kariyama, R.; Nagata, K.; Nishiguchi, K.; Mitsuhata, R.; Uemura, Y.; Tanokura, M.; Kumon, H.; Sonomoto, K. Siamycin Attenuates Fsr Quorum Sensing Mediated by a Gelatinase Biosynthesis-Activating Pheromone in Enterococcus Faecalis. *Journal of Bacteriology* **2007**, *189* (4), 1358–1365. https://doi.org/10.1128/JB.00969-06.

(38) Ma, P.; Nishiguchi, K.; Yuille, H. M.; Davis, L. M.; Nakayama, J.; Phillips-Jones, M. K. Anti-HIV Siamycin I Directly Inhibits Autophosphorylation Activity of the Bacterial FsrC Quorum Sensor and Other ATP-Dependent Enzyme Activities. *FEBS Letters* **2011**, *585* (17), 2660–2664. https://doi.org/10.1016/j.febslet.2011.07.026.

(39) K. Phillips-Jones, M.; G. Patching, S.; Edara, S.; Nakayama, J.; Hussain, R.; Siligardi, G. Interactions of the Intact FsrC Membrane Histidine Kinase with the Tricyclic Peptide Inhibitor Siamycin I Revealed through Synchrotron Radiation Circular Dichroism. *Physical Chemistry Chemical Physics* **2013**, *15* (2), 444–447. https://doi.org/10.1039/C2CP43722H.

(40) Li, Y.; Ducasse, R.; Zirah, S.; Blond, A.; Goulard, C.; Lescop, E.; Giraud, C.; Hartke, A.; Guittet, E.; Pernodet, J.-L.; Rebuffat, S. Characterization of Sviceucin from Streptomyces Provides Insight into Enzyme Exchangeability and Disulfide Bond Formation in Lasso Peptides. *ACS Chem. Biol.* **2015**, *10* (11), 2641–2649. https://doi.org/10.1021/acschembio.5b00584.

(41) Salomón, R. A.; Farías, R. N. Microcin 25, a Novel Antimicrobial Peptide Produced by Escherichia Coli. *J. Bacteriol.* **1992**, *174* (22), 7428–7435. https://doi.org/10.1128/jb.174.22.7428-7435.1992.

(42) Wilson, K.-A.; Kalkum, M.; Ottesen, J.; Yuzenkova, J.; Chait, B. T.; Landick, R.; Muir, T.; Severinov, K.; Darst, S. A. Structure of Microcin J25, a Peptide Inhibitor of Bacterial RNA Polymerase, Is a Lassoed Tail. *J. Am. Chem. Soc.* **2003**, *125* (41), 12475–12483. https://doi.org/10.1021/ja036756q.

(43) Bayro, M. J.; Mukhopadhyay, J.; Swapna, G. V. T.; Huang, J. Y.; Ma, L.-C.; Sineva, E.; Dawson, P. E.; Montelione, G. T.; Ebright, R. H. Structure of Antibacterial Peptide Microcin J25: A 21-Residue Lariat Protoknot. *J. Am. Chem. Soc.* **2003**, *125* (41), 12382–12383. https://doi.org/10.1021/ja036677e.

(44) Delgado, M. A.; Rintoul, M. R.; Farías, R. N.; Salomón, R. A. Escherichia Coli RNA Polymerase Is the Target of the Cyclopeptide Antibiotic Microcin J25. *J. Bacteriol.* **2001**, *183* (15), 4543–4550. https://doi.org/10.1128/JB.183.15.4543-4550.2001.

(45) Yuzenkova, J.; Delgado, M.; Nechaev, S.; Savalia, D.; Epshtein, V.; Artsimovitch, I.; Mooney, R. A.; Landick, R.; Farias, R. N.; Salomon, R.; Severinov, K. Mutations of Bacterial RNA Polymerase Leading to Resistance to Microcin J25. *J. Biol. Chem.* **2002**, *277* (52), 50867–50875. https://doi.org/10.1074/jbc.M209425200.

(46) Adelman, K.; Yuzenkova, J.; La Porta, A.; Zenkin, N.; Lee, J.; Lis, J. T.; Borukhov, S.; Wang, M. D.; Severinov, K. Molecular Mechanism of Transcription Inhibition by Peptide Antibiotic Microcin J25. *Molecular Cell* **2004**, *14* (6), 753–762. https://doi.org/10.1016/j.molcel.2004.05.017.

(47) Mukhopadhyay, J.; Sineva, E.; Knight, J.; Levy, R. M.; Ebright, R. H. Antibacterial Peptide Microcin J25 Inhibits Transcription by Binding within and Obstructing the RNA Polymerase Secondary Channel. *Molecular Cell* **2004**, *14* (6), 739–751. https://doi.org/10.1016/j.molcel.2004.06.010.

(48) Braffman, N. R.; Piscotta, F. J.; Hauver, J.; Campbell, E. A.; Link, A. J.; Darst, S. A. Structural Mechanism of Transcription Inhibition by Lasso Peptides Microcin J25 and Capistruin. *PNAS* **2019**, *116* (4), 1273–1278. https://doi.org/10.1073/pnas.1817352116.

(49) Salomón, R. A.; Farías, R. N. The FhuA Protein Is Involved in Microcin 25 Uptake. *Journal of Bacteriology* **1993**, *175* (23), 7741–7742. https://doi.org/10.1128/jb.175.23.7741-7742.1993.

(50) Salomón, R. A.; Farías, R. N. The Peptide Antibiotic Microcin 25 Is Imported through the TonB Pathway and the SbmA Protein. *Journal of Bacteriology* **1995**, *177* (11), 3323–3325. https://doi.org/10.1128/jb.177.11.3323-3325.1995.

(51) Killmann, H.; Braun, M.; Herrmann, C.; Braun, V. FhuA Barrel-Cork Hybrids Are Active Transporters and Receptors. *Journal of Bacteriology* **2001**, *183* (11), 3476–3487. https://doi.org/10.1128/JB.183.11.3476-3487.2001.

(52) Vincent, P. A.; Delgado, M. A.; Farías, R. N.; Salomón, R. A. Inhibition of Salmonella Enterica Serovars by Microcin J25. *FEMS Microbiol Lett* **2004**, *236* (1), 103–107. https://doi.org/10.1111/j.1574-6968.2004.tb09634.x.

(53) Endriß, F.; Braun, V. Loop Deletions Indicate Regions Important for FhuA Transport and Receptor Functions in Escherichia Coli. *Journal of Bacteriology* **2004**, *186* (14), 4818–4823. https://doi.org/10.1128/JB.186.14.4818-4823.2004.

(54) Destoumieux-Garzón, D.; Duquesne, S.; Peduzzi, J.; Goulard, C.; Desmadril, M.; Letellier, L.; Rebuffat, S.; Boulanger, P. The Iron–Siderophore Transporter FhuA Is the Receptor for the Antimicrobial Peptide Microcin J25: Role of the Microcin Val11–Pro16 β-Hairpin Region in the Recognition Mechanism. *Biochemical Journal* **2005**, *389* (3), 869–876. https://doi.org/10.1042/BJ20042107.

(55) Mathavan, I.; Zirah, S.; Mehmood, S.; Choudhury, H. G.; Goulard, C.; Li, Y.; Robinson, C. V.; Rebuffat, S.; Beis, K. Structural Basis for Hijacking Siderophore Receptors by Antimicrobial Lasso Peptides. *Nature Chemical Biology* **2014**, *10* (5), 340–342. https://doi.org/10.1038/nchembio.1499.

(56) A, B.; M, C.; D, D.-G.; I, S.-M.; J, P.; C, G.; S, R. Thermolysin-Linearized Microcin J25 Retains the Structured Core of the Native Macrocyclic Peptide and Displays Antimicrobial Activity. *Eur J Biochem* **2002**, *269* (24), 6212–6222. https://doi.org/10.1046/j.1432-1033.2002.03340.x.

(57) Rosengren, K. J.; Blond, A.; Afonso, C.; Tabet, J.-C.; Rebuffat, S.; Craik, D. J. Structure of Thermolysin Cleaved Microcin J25: Extreme Stability of a Two-Chain Antimicrobial Peptide Devoid of Covalent Links,. *Biochemistry* **2004**, *43* (16), 4696–4702. https://doi.org/10.1021/bi0361261.

(58) Semenova, E.; Yuzenkova, Y.; Peduzzi, J.; Rebuffat, S.; Severinov, K. Structure-Activity Analysis of Microcin J25: Distinct Parts of the Threaded Lasso Molecule Are Responsible for

Interaction with Bacterial RNA Polymerase. *Journal of Bacteriology* **2005**, *187* (11), 3859–3863. https://doi.org/10.1128/JB.187.11.3859-3863.2005.

(59) Bellomio, A.; Vincent, P. A.; de Arcuri, B. F.; Salomón, R. A.; Morero, R. D.; Farías, R. N. The Microcin J25 β-Hairpin Region Is Important for Antibiotic Uptake but Not for RNA Polymerase and Respiration Inhibition. *Biochemical and Biophysical Research Communications* **2004**, *325* (4), 1454–1458. https://doi.org/10.1016/j.bbrc.2004.10.186.

(60) Bellomio, A.; Rintoul, M. R.; Morero, R. D. Chemical Modification of Microcin J25 with Diethylpyrocarbonate and Carbodiimide: Evidence for Essential Histidyl and Carboxyl Residues. *Biochemical and Biophysical Research Communications* **2003**, *303* (2), 458–462. https://doi.org/10.1016/S0006-291X(03)00373-5.

(61) Cristóbal, R. E. de; Solbiati, J. O.; Zenoff, A. M.; Vincent, P. A.; Salomón, R. A.; Yuzenkova, J.; Severinov, K.; Farías, R. N. Microcin J25 Uptake: His5 of the MccJ25 Lariat Ring Is Involved in Interaction with the Inner Membrane MccJ25 Transporter Protein SbmA. *Journal of Bacteriology* **2006**, *188* (9), 3324–3328. https://doi.org/10.1128/JB.188.9.3324-3328.2006.

(62) Vincent, P. A.; Bellomio, A.; de Arcuri, B. F.; Farías, R. N.; Morero, R. D. MccJ25 C-Terminal Is Involved in RNA-Polymerase Inhibition but Not in Respiration Inhibition. *Biochemical and Biophysical Research Communications* **2005**, *331* (2), 549–551. https://doi.org/10.1016/j.bbrc.2005.03.220.

(63) Rintoul, M. R.; de Arcuri, B. F.; Salomón, R. A.; Farías, R. N.; Morero, R. D. The Antibacterial Action of Microcin J25: Evidence for Disruption of Cytoplasmic Membrane Energization in Salmonella Newport. *FEMS Microbiol Lett* **2001**, *204* (2), 265–270. https://doi.org/10.1111/j.1574-6968.2001.tb10895.x.

(64) Rintoul, M. R.; de Arcuri, B. F.; Morero, R. D. Effects of the Antibiotic Peptide Microcin J25 on Liposomes: Role of Acyl Chain Length and Negatively Charged Phospholipid. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2000**, *1509* (1), 65–72. https://doi.org/10.1016/S0005-2736(00)00249-2.

(65) Bellomio, A.; Oliveira, R. G.; Maggio, B.; Morero, R. D. Penetration and Interactions of the Antimicrobial Peptide, Microcin J25, into Uncharged Phospholipid Monolayers. *Journal of Colloid and Interface Science* **2005**, *285* (1), 118–124. https://doi.org/10.1016/j.jcis.2004.11.025.

(66) Dupuy, F. G.; Chirou, M. V. N.; de Arcuri, B. F.; Minahk, C. J.; Morero, R. D. Proton Motive Force Dissipation Precludes Interaction of Microcin J25 with RNA Polymerase, but Enhances Reactive Oxygen Species Overproduction. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2009**, *1790* (10), 1307–1313. https://doi.org/10.1016/j.bbagen.2009.07.006.

(67) Bellomio, A.; Vincent, P. A.; Arcuri, B. F. de; Farías, R. N.; Morero, R. D. Microcin J25 Has Dual and Independent Mechanisms of Action in Escherichia Coli: RNA Polymerase Inhibition and Increased Superoxide Production. *Journal of Bacteriology* **2007**, *189* (11), 4180–4186. https://doi.org/10.1128/JB.00206-07.

(68) Galván, A. E.; Chalón, M. C.; Schurig-Briccio, L. A.; Salomón, R. A.; Minahk, C. J.; Gennis, R. B.; Bellomio, A. Cytochromes Bd-I and Bo3 Are Essential for the Bactericidal Effect of Microcin J25 on Escherichia Coli Cells. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2018**, *1859* (2), 110–118. https://doi.org/10.1016/j.bbabio.2017.10.006.

(69) Galván, A. E.; Chalón, M. C.; Ríos Colombo, N. S.; Schurig-Briccio, L. A.; Sosa-Padilla, B.; Gennis, R. B.; Bellomio, A. Microcin J25 Inhibits Ubiquinol Oxidase Activity of Purified Cytochrome Bd-I from Escherichia Coli. *Biochimie* **2019**, *160*, 141–147. https://doi.org/10.1016/j.biochi.2019.02.007.

(70) Chalon, M. C.; Bellomio, A.; Solbiati, J. O.; Morero, R. D.; Farias, R. N.; Vincent, P. A. Tyrosine 9 Is the Key Amino Acid in Microcin J25 Superoxide Overproduction. *FEMS Microbiol Lett* **2009**, *300* (1), 90–96. https://doi.org/10.1111/j.1574-6968.2009.01770.x.

(71) Chalón, M. C.; Wilke, N.; Pedersen, J.; Rufini, S.; Morero, R. D.; Cortez, L.; Chehín, R. N.; Farias, R. N.; Vincent, P. A. Redox-Active Tyrosine Residue in the Microcin J25 Molecule. *Biochemical and Biophysical Research Communications* **2011**, *406* (3), 366–370. https://doi.org/10.1016/j.bbrc.2011.02.047.

(72) Niklison Chirou, M. V.; Minahk, C. J.; Morero, R. D. Antimitochondrial Activity Displayed by the Antimicrobial Peptide Microcin J25. *Biochemical and Biophysical Research Communications* **2004**, *317* (3), 882–886. https://doi.org/10.1016/j.bbrc.2004.03.127.

(73) Chirou, M. N.; Bellomio, A.; Dupuy, F.; Arcuri, B.; Minahk, C.; Morero, R. Microcin J25 Induces the Opening of the Mitochondrial Transition Pore and Cytochrome c Release through Superoxide Generation. *The FEBS Journal* **2008**, *275* (16), 4088–4096. https://doi.org/10.1111/j.1742-4658.2008.06550.x.

(74) Niklison-Chirou, M. V.; Dupuy, F.; Pena, L. B.; Gallego, S. M.; Barreiro-Arcos, M. L.; Avila, C.; Torres-Bugeau, C.; Arcuri, B. E.; Bellomio, A.; Minahk, C.; Morero, R. D. Microcin J25 Triggers Cytochrome c Release through Irreversible Damage of Mitochondrial Proteins and Lipids. *The International Journal of Biochemistry & Cell Biology* **2010**, *42* (2), 273–281. https://doi.org/10.1016/j.biocel.2009.11.002.

(75) Kuznedelov, K.; Semenova, E.; Knappe, T. A.; Mukhamedyarov, D.; Srivastava, A.; Chatterjee, S.; Ebright, R. H.; Marahiel, M. A.; Severinov, K. The Antibacterial Threaded-Lasso Peptide Capistruin Inhibits Bacterial RNA Polymerase. *Journal of Molecular Biology* **2011**, *412* (5), 842–848. https://doi.org/10.1016/j.jmb.2011.02.060.

(76) Metelev, M.; Arseniev, A.; Bushin, L. B.; Kuznedelov, K.; Artamonova, T. O.; Kondratenko, R.; Khodorkovskii, M.; Seyedsayamdost, M. R.; Severinov, K. Acinetodin and Klebsidin, RNA Polymerase Targeting Lasso Peptides Produced by Human Isolates of Acinetobacter Gyllenbergii and Klebsiella Pneumoniae. *ACS Chem. Biol.* **2017**, *12* (3), 814–824. https://doi.org/10.1021/acschembio.6b01154.

(77) Cheung-Lee, W. L.; Parry, M. E.; Jaramillo Cartagena, A.; Darst, S. A.; Link, A. J. Discovery and Structure of the Antimicrobial Lasso Peptide Citrocin. *J Biol Chem* **2019**, *294* (17), 6822–6830. https://doi.org/10.1074/jbc.RA118.006494.

(78) Cheung-Lee, W. L.; Parry, M. E.; Zong, C.; Cartagena, A. J.; Darst, S. A.; Connell, N. D.; Russo, R.; Link, A. J. Discovery of Ubonodin, an Antimicrobial Lasso Peptide Active against Members of the Burkholderia Cepacia Complex. *ChemBioChem* **2020**, *21* (9), 1335–1340. https://doi.org/10.1002/cbic.201900707.

(79) Knappe, T. A.; Linne, U.; Zirah, S.; Rebuffat, S.; Xie, X.; Marahiel, M. A. Isolation and Structural Characterization of Capistruin, a Lasso Peptide Predicted from the Genome Sequence of Burkholderia Thailandensis E264. *J. Am. Chem. Soc.* **2008**, *130* (34), 11446–11454. https://doi.org/10.1021/ja802966g.

(80) Iwatsuki, M.; Uchida, R.; Takakusagi, Y.; Matsumoto, A.; Jiang, C.-L.; Takahashi, Y.; Arai, M.; Kobayashi, S.; Matsumoto, M.; Inokoshi, J.; Tomoda, H.; Ōmura, S. Lariatins, Novel Anti-Mycobacterial Peptides with a Lasso Structure, Produced by *Rhodococcus Jostii* K01-B0171. *The Journal of Antibiotics* **2007**, *60* (6), 357–363. https://doi.org/10.1038/ja.2007.48.

(81) Iwatsuki, M.; Koizumi, Y.; Gouda, H.; Hirono, S.; Tomoda, H.; Ōmura, S. Lys17 in the 'Lasso' Peptide Lariatin A Is Responsible for Anti-Mycobacterial Activity. *Bioorganic &*

Medicinal Chemistry Letters **2009**, *19* (10), 2888–2890. https://doi.org/10.1016/j.bmcl.2009.03.033.

(82) Inokoshi, J.; Koyama, N.; Miyake, M.; Shimizu, Y.; Tomoda, H. Structure-Activity Analysis of Gram-Positive Bacterium-Producing Lasso Peptides with Anti-Mycobacterial Activity. *Scientific Reports* **2016**, *6* (1), 30375. https://doi.org/10.1038/srep30375.

(83) Valiante, V.; Monteiro, M. C.; Martín, J.; Altwasser, R.; Aouad, N. E.; González, I.; Kniemeyer, O.; Mellado, E.; Palomo, S.; Pedro, N. de; Pérez-Victoria, I.; Tormo, J. R.; Vicente, F.; Reyes, F.; Genilloud, O.; Brakhage, A. A. Hitting the Caspofungin Salvage Pathway of Human-Pathogenic Fungi with the Novel Lasso Peptide Humidimycin (MDN-0010). *Antimicrobial Agents and Chemotherapy* **2015**, *59* (9), 5145–5153. https://doi.org/10.1128/AAC.00683-15.

(84) Sit, C. S.; McKay, R. T.; Hill, C.; Ross, R. P.; Vederas, J. C. The 3D Structure of Thuricin CD, a Two-Component Bacteriocin with Cysteine Sulfur to α-Carbon Cross-Links. *J. Am. Chem. Soc.* **2011**, *133* (20), 7680–7683. https://doi.org/10.1021/ja201802f.

(85) Sit, C. S.; van Belkum, M. J.; McKay, R. T.; Worobo, R. W.; Vederas, J. C. The 3D Solution Structure of Thurincin H, a Bacteriocin with Four Sulfur to α-Carbon Crosslinks. *Angewandte Chemie International Edition* **2011**, *50* (37), 8718–8721. https://doi.org/10.1002/anie.201102527.

(86) Kawulka, K. E.; Sprules, T.; Diaper, C. M.; Whittal, R. M.; McKay, R. T.; Mercier, P.; Zuber, P.; Vederas, J. C. Structure of Subtilosin A, a Cyclic Antimicrobial Peptide from Bacillus Subtilis with Unusual Sulfur to α-Carbon Cross-Links: Formation and Reduction of α-Thio-α-Amino Acid Derivatives,. *Biochemistry* **2004**, *43* (12), 3385–3395. https://doi.org/10.1021/bi0359527.

(87) Liu, W.-T.; Yang, Y.-L.; Xu, Y.; Lamsa, A.; Haste, N. M.; Yang, J. Y.; Ng, J.; Gonzalez, D.; Ellermeier, C. D.; Straight, P. D.; Pevzner, P. A.; Pogliano, J.; Nizet, V.; Pogliano, K.; Dorrestein, P. C. Imaging Mass Spectrometry of Intraspecies Metabolic Exchange Revealed the Cannibalistic Factors of Bacillus Subtilis. *PNAS* **2010**, *107* (37), 16286–16290. https://doi.org/10.1073/pnas.1008368107.

(88) Hudson, G. A.; Burkhart, B. J.; DiCaprio, A. J.; Schwalen, C. J.; Kille, B.; Pogorelov, T. V.; Mitchell, D. A. Bioinformatic Mapping of Radical S-Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. *J. Am. Chem. Soc.* **2019**, *141* (20), 8228–8238. https://doi.org/10.1021/jacs.9b01519.

(89) Mo, T.; Ji, X.; Yuan, W.; Mandalapu, D.; Wang, F.; Zhong, Y.; Li, F.; Chen, Q.; Ding, W.; Deng, Z.; Yu, S.; Zhang, Q. Thuricin Z: A Narrow-Spectrum Sactibiotic That Targets the Cell Membrane. *Angewandte Chemie* **2019**, *131* (52), 18969–18973. https://doi.org/10.1002/ange.201908490.

(90) Babasaki, K.; Takao, T.; Shimonishi, Y.; Kurahashi, K. Subtilosin A, a New Antibiotic Peptide Produced by Bacillus Subtilis 168: Isolation, Structural Analysis, and Biogenesis. *J Biochem* **1985**, *98* (3), 585–603. https://doi.org/10.1093/oxfordjournals.jbchem.a135315.

(91) Zheng, G.; Slavik, M. F. Isolation, Partial Purification and Characterization of a Bacteriocin Produced by a Newly Isolated Bacillus Subtilis Strain. *Letters in Applied Microbiology* **1999**, *28* (5), 363–367. https://doi.org/10.1046/j.1365-2672.1999.00545.x.

(92) Rea, M. C.; Sit, C. S.; Clayton, E.; O'Connor, P. M.; Whittal, R. M.; Zheng, J.; Vederas, J. C.; Ross, R. P.; Hill, C. Thuricin CD, a Posttranslationally Modified Bacteriocin with a

Narrow Spectrum of Activity against Clostridium Difficile. *PNAS* **2010**, *107* (20), 9352–9357. https://doi.org/10.1073/pnas.0913554107.

(93) Rea, M. C.; Dobson, A.; O'Sullivan, O.; Crispie, F.; Fouhy, F.; Cotter, P. D.; Shanahan, F.; Kiely, B.; Hill, C.; Ross, R. P. Effect of Broad- and Narrow-Spectrum Antimicrobials on Clostridium Difficile and Microbial Diversity in a Model of the Distal Colon. *PNAS* **2011**, *108* (Supplement 1), 4639–4644. https://doi.org/10.1073/pnas.1001224107.

(94) Sutyak Noll, K.; Sinko, P. J.; Chikindas, M. L. Elucidation of the Molecular Mechanisms of Action of the Natural Antimicrobial Peptide Subtilosin Against the Bacterial Vaginosis-Associated Pathogen Gardnerella Vaginalis. *Probiotics & Antimicro. Prot.* **2011**, *3* (1), 41–47. https://doi.org/10.1007/s12602-010-9061-4.

(95) Kuijk, S. van; Noll, K. S.; Chikindas, M. L. The Species-Specific Mode of Action of the Antimicrobial Peptide Subtilosin against Listeria Monocytogenes Scott A. *Letters in Applied Microbiology* **2012**, *54* (1), 52–58. https://doi.org/10.1111/j.1472-765X.2011.03170.x.

(96) Duarte, A. F. de S.; Ceotto-Vigoder, H.; Barrias, E. S.; Souto-Padrón, T. C. B. S.; Nes, I. F.; Bastos, M. do C. de F. Hyicin 4244, the First Sactibiotic Described in Staphylococci, Exhibits an Anti-Staphylococcal Biofilm Activity. *International Journal of Antimicrobial Agents* **2018**, *51* (3), 349–356. https://doi.org/10.1016/j.ijantimicag.2017.06.025.

(97) Sutyak, K. E.; Wirawan, R. E.; Aroutcheva, A. A.; Chikindas, M. L. Isolation of the Bacillus Subtilis Antimicrobial Peptide Subtilosin from the Dairy Product-Derived Bacillus Amyloliquefaciens. *Journal of Applied Microbiology* **2008**, *104* (4), 1067–1074. https://doi.org/10.1111/j.1365-2672.2007.03626.x.

(98) Amrouche, T.; Sutyak Noll, K.; Wang, Y.; Huang, Q.; Chikindas, M. L. Antibacterial Activity of Subtilosin Alone and Combined with Curcumin, Poly-Lysine and Zinc Lactate

Against Listeria Monocytogenes Strains. *Probiotics & Antimicro. Prot.* **2010**, *2* (4), 250–257. https://doi.org/10.1007/s12602-010-9042-7.

(99) Thennarasu, S.; Lee, D.-K.; Poon, A.; Kawulka, K. E.; Vederas, J. C.; Ramamoorthy, A. Membrane Permeabilization, Orientation, and Antimicrobial Mechanism of Subtilosin A. *Chemistry and Physics of Lipids* **2005**, *137* (1), 38–51. https://doi.org/10.1016/j.chemphyslip.2005.06.003.

(100) Wang, G.; Feng, G.; Snyder, A. B.; Manns, D. C.; Churey, J. J.; Worobo, R. W. Bactericidal Thurincin H Causes Unique Morphological Changes in Bacillus Cereus F4552 without Affecting Membrane Permeability. *FEMS Microbiol Lett* **2014**, *357* (1), 69–76. https://doi.org/10.1111/1574-6968.12486.

(101) Mathur, H.; Fallico, V.; O'Connor, P. M.; Rea, M. C.; Cotter, P. D.; Hill, C.; Ross, R. P. Insights into the Mode of Action of the Sactibiotic Thuricin CD. *Front. Microbiol.* **2017**, *8*. https://doi.org/10.3389/fmicb.2017.00696.

(102) Marx, R.; Stein, T.; Entian, K.-D.; Glaser, S. J. Structure of the Bacillus Subtilis Peptide Antibiotic Subtilosin A Determined by 1H-NMR and Matrix Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *J Protein Chem* **2001**, *20* (6), 501–506. https://doi.org/10.1023/A:1012562631268.

(103) Huang, T.; Geng, H.; Miyyapuram, V. R.; Sit, C. S.; Vederas, J. C.; Nakano, M. M. Isolation of a Variant of Subtilosin A with Hemolytic Activity. *Journal of Bacteriology* **2009**, *191* (18), 5690–5696. https://doi.org/10.1128/JB.00541-09.

(104) Silkin, L.; Hamza, S.; Kaufman, S.; Cobb, S. L.; Vederas, J. C. Spermicidal Bacteriocins: Lacticin 3147 and Subtilosin A. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18* (10), 3103–3106. https://doi.org/10.1016/j.bmcl.2007.11.024.

(105) Sutyak, K. E.; Anderson, R. A.; Dover, S. E.; Feathergill, K. A.; Aroutcheva, A. A.; Faro, S.; Chikindas, M. L. Spermicidal Activity of the Safe Natural Antimicrobial Peptide Subtilosin https://www.hindawi.com/journals/idog/2008/540758/ (accessed Apr 9, 2020). https://doi.org/10.1155/2008/540758.

(106) Torres, N. I.; Noll, K. S.; Xu, S.; Li, J.; Huang, Q.; Sinko, P. J.; Wachsman, M. B.; Chikindas, M. L. Safety, Formulation and In Vitro Antiviral Activity of the Antimicrobial Peptide Subtilosin Against Herpes Simplex Virus Type 1. *Probiotics & Antimicro. Prot.* **2013**, *5* (1), 26–35. https://doi.org/10.1007/s12602-012-9123-x.

(107) Algburi, A.; Zehm, S.; Netrebov, V.; Bren, A. B.; Chistyakov, V.; Chikindas, M. L. Subtilosin Prevents Biofilm Formation by Inhibiting Bacterial Quorum Sensing. *Probiotics & Antimicro. Prot.* **2017**, *9* (1), 81–90. https://doi.org/10.1007/s12602-016-9242-x.

(108) McClean, K. H.; Winson, M. K.; Fish, L.; Taylor, A.; Chhabra, S. R.; Camara, M.; Daykin, M.; Lamb, J. H.; Swift, S.; Bycroft, B. W.; Stewart, G. S. A. B.; Williams, P. 1997. Quorum Sensing and Chromobacterium Violaceum: Exploitation of Violacein Production and Inhibition for the Detection of N-Acylhomoserine Lactones. *Microbiology 143* (12), 3703–3711. https://doi.org/10.1099/00221287-143-12-3703.

(109) Waters, C. M.; Bassler, B. L. QUORUM SENSING: Cell-to-Cell Communication in Bacteria. *Annu. Rev. Cell Dev. Biol.* **2005**, *21* (1), 319–346. https://doi.org/10.1146/annurev.cellbio.21.012704.131001.

(110) Lee, H.; Churey, J. J.; Worobo, R. W. Biosynthesis and Transcriptional Analysis of Thurincin H, a Tandem Repeated Bacteriocin Genetic Locus, Produced by Bacillus Thuringiensis SF361. *FEMS Microbiol Lett* **2009**, *299* (2), 205–213. https://doi.org/10.1111/j.1574-6968.2009.01749.x.

(111) Wang, G.; Manns, D. C.; Guron, G. K.; Churey, John. J.; Worobo, R. W. Large-Scale Purification, Characterization, and Spore Outgrowth Inhibitory Effect of Thurincin H, a Bacteriocin Produced by Bacillus Thuringiensis SF361. *Probiotics & Antimicro. Prot.* **2014**, *6* (2), 105–113. https://doi.org/10.1007/s12602-014-9159-1.

(112) Ishitsuka, M. O.; Kusumi, T.; Kakisawa, H.; Kaya, K.; Watanabe, M. M. Microviridin. A Novel Tricyclic Depsipeptide from the Toxic Cyanobacterium Microcystis Viridis. *J. Am. Chem. Soc.* **1990**, *112* (22), 8180–8182. https://doi.org/10.1021/ja00178a060.

(113) Imada, C.; Simidu, U.; Taga, N. Isolation and Characterization of Marine Bacteria Producing Alkaline Protease Inhibitor. *Nippon Suisan Gakkaishi* **1985**, *51* (5), 799–803. https://doi.org/10.2331/suisan.51.799.

(114) Imada, C.; Taga, N.; Maeda, M. Cultivation Conditions for Subtilisin Inhibitor-Producing Bacterium and General Properties of the Inhibitor "Marinostatin." *Nippon Suisan Gakkaishi* **1985**, *51* (5), 805–810. https://doi.org/10.2331/suisan.51.805.

(115) Rohrlack, T.; Christoffersen, K.; Kaebernick, M.; Neilan, B. A. Cyanobacterial Protease Inhibitor Microviridin J Causes a Lethal Molting Disruption in Daphnia Pulicaria. *Appl. Environ. Microbiol.* **2004**, *70* (8), 5047–5050. https://doi.org/10.1128/AEM.70.8.5047-5050.2004.

(116) Rohrlack, T.; Christoffersen, K.; Hansen, P. E.; Zhang, W.; Czarnecki, O.; Henning, M.; Fastner, J.; Erhard, M.; Neilan, B. A.; Kaebernick, M. Isolation, Characterization, and Quantitative Analysis of Microviridin J, a New Microcystis Metabolite Toxic to Daphnia. *J Chem Ecol* **2003**, *29* (8), 1757–1770. https://doi.org/10.1023/A:1024889925732.

(117) Weiz, A. R.; Ishida, K.; Quitterer, F.; Meyer, S.; Kehr, J.-C.; Müller, K. M.; Groll, M.; Hertweck, C.; Dittmann, E. Harnessing the Evolvability of Tricyclic Microviridins To Dissect

Protease–Inhibitor Interactions. *Angewandte Chemie International Edition* **2014**, *53* (14), 3735–3738. https://doi.org/10.1002/anie.201309721.

(118) Harris, J. L.; Backes, B. J.; Leonetti, F.; Mahrus, S.; Ellman, J. A.; Craik, C. S. Rapid and General Profiling of Protease Specificity by Using Combinatorial Fluorogenic Substrate Libraries. *PNAS* **2000**, *97* (14), 7754–7759. https://doi.org/10.1073/pnas.140132697.

(119) Reyna-González, E.; Schmid, B.; Petras, D.; Süssmuth, R. D.; Dittmann, E. Leader Peptide-Free In Vitro Reconstitution of Microviridin Biosynthesis Enables Design of Synthetic Protease-Targeted Libraries. *Angewandte Chemie International Edition* **2016**, *55* (32), 9398–9401. https://doi.org/10.1002/anie.201604345.

(120) Kanaori, K.; Kamei, K.; Taniguchi, M.; Koyama, T.; Yasui, T.; Takano, R.; Imada, C.; Tajima, K.; Hara, S. Solution Structure of Marinostatin, a Natural Ester-Linked Protein Protease Inhibitor. *Biochemistry* **2005**, *44* (7), 2462–2468. https://doi.org/10.1021/bi048034x.

(121) Hiromi, K. *Protein Protease Inhibitor: The Case of Streptomyces Subtilisin Inhibitor (SSI)*; Elsevier, 1985.

(122) Fujinaga, M.; Sielecki, A. R.; Read, R. J.; Ardelt, W.; Laskowski, M.; James, M. N. G. Crystal and Molecular Structures of the Complex of α-Chymotrypsin with Its Inhibitor Turkey Ovomucoid Third Domain at 1.8 Å Resolution. *Journal of Molecular Biology* **1987**, *195* (2), 397–418. https://doi.org/10.1016/0022-2836(87)90659-0.

(123) Ahmed, M. N.; Reyna-González, E.; Schmid, B.; Wiebach, V.; Süssmuth, R. D.; Dittmann, E.; Fewer, D. P. Phylogenomic Analysis of the Microviridin Biosynthetic Pathway Coupled with Targeted Chemo-Enzymatic Synthesis Yields Potent Protease Inhibitors. *ACS Chem. Biol.* **2017**, *12* (6), 1538–1546. https://doi.org/10.1021/acschembio.7b00124.

(124) Lee, H.; Park, Y.; Kim, S. Enzymatic Cross-Linking of Side Chains Generates a Modified Peptide with Four Hairpin-like Bicyclic Repeats. *Biochemistry* **2017**, *56* (37), 4927–4930. https://doi.org/10.1021/acs.biochem.7b00808.

(125) Lee, C.; Lee, H.; Park, J.-U.; Kim, S. Introduction of Bifunctionality into the Multidomain Architecture of the ω-Ester-Containing Peptide Plesiocin. *Biochemistry* **2020**, *59* (3), 285–289. https://doi.org/10.1021/acs.biochem.9b00803.

(126) Okino, T.; Matsuda, H.; Murakami, M.; Yamaguchi, K. New Microviridins, Elastase Inhibitors from the Blue-Green Alga Microcystis Aeruginosa. *Tetrahedron* **1995**, *51* (39), 10679–10686. https://doi.org/10.1016/0040-4020(95)00645-O.

(127) Anas, A. R. J.; Mori, A.; Tone, M.; Naruse, C.; Nakajima, A.; Asukabe, H.; Takaya, Y.; Imanishi, S. Y.; Nishizawa, T.; Shirai, M.; Harada, K. FVIIa-STF and Thrombin Inhibitory Activities of Compounds Isolated from Microcystis Aeruginosa K-139. *Marine Drugs* **2017**, *15* (9), 275. https://doi.org/10.3390/md15090275.

(128) Shin, H. J.; Murakami, M.; Matsuda, H.; Yamaguchi, K. Microviridins D-F, Serine Protease Inhibitors from the Cyanobacterium Oscillatoria Agardhii (NIES-204). *Tetrahedron* **1996**, *52* (24), 8159–8168. https://doi.org/10.1016/0040-4020(96)00377-8.

(129) Murakami, M.; Sun, Q.; Ishida, K.; Matsuda, H.; Okino, T.; Yamaguchi, K. Microviridins, Elastase Inhibitors from the Cyanobacterium Nostoc Minutum (NIES-26). *Phytochemistry* **1997**, *45* (6), 1197–1202. https://doi.org/10.1016/S0031-9422(97)00131-3.

(130) Fujii, K.; Sivonen, K.; Naganawa, E.; Harada, K. Non-Toxic Peptides from Toxic Cyanobacteria, Oscillatoria Agardhii. *Tetrahedron* **2000**, *56* (5), 725–733. https://doi.org/10.1016/S0040-4020(99)01017-0.

(131) Reshef, V.; Carmeli, S. New Microviridins from a Water Bloom of the Cyanobacterium Microcystis Aeruginosa. *Tetrahedron* **2006**, *62* (31), 7361–7369. https://doi.org/10.1016/j.tet.2006.05.028.

(132) Ziemert, N.; Ishida, K.; Weiz, A.; Hertweck, C.; Dittmann, E. Exploiting the Natural Diversity of Microviridin Gene Clusters for Discovery of Novel Tricyclic Depsipeptides. *Appl. Environ. Microbiol.* **2010**, *76* (11), 3568–3574. https://doi.org/10.1128/AEM.02858-09.

(133) Gatte-Picchi, D.; Weiz, A.; Ishida, K.; Hertweck, C.; Dittmann, E. Functional Analysis of Environmental DNA-Derived Microviridins Provides New Insights into the Diversity of the Tricyclic Peptide Family. *Appl. Environ. Microbiol.* **2014**, *80* (4), 1380–1387. https://doi.org/10.1128/AEM.03502-13.

(134) Sieber, S.; Grendelmeier, S. M.; Harris, L. A.; Mitchell, D. A.; Gademann, K. Microviridin 1777: A Toxic Chymotrypsin Inhibitor Discovered by a Metabologenomic Approach. *J. Nat. Prod.* **2020**, *83* (2), 438–446. https://doi.org/10.1021/acs.jnatprod.9b00986.

(135) Taichi, M.; Yamazaki, T.; Kawahara, K.; Motooka, D.; Nakamura, S.; Harada, S.; Teshima, T.; Ohkubo, T.; Kobayashi, Y.; Nishiuchi, Y. Structure–Activity Relationship of Marinostatin, a Serine Protease Inhibitor Isolated from a Marine Organism. *Journal of Peptide Science* **2010**, *16* (7), 329–336. https://doi.org/10.1002/psc.1244.

(136) Freeman, M. F.; Gurgui, C.; Helf, M. J.; Morinaka, B. I.; Uria, A. R.; Oldham, N. J.; Sahl, H.-G.; Matsunaga, S.; Piel, J. Metagenome Mining Reveals Polytheonamides as Posttranslationally Modified Ribosomal Peptides. *Science* **2012**, *338* (6105), 387–390. https://doi.org/10.1126/science.1226121.

(137) Hamada, T.; Sugawara, T.; Matsunaga, S.; Fusetani, N. Polytheonamides, Unprecedented Highly Cytotoxic Polypeptides, from the Marine Sponge Theonella Swinhoei: 1. Isolation and

Component Amino Acids. *Tetrahedron Letters* **1994**, *35* (5), 719–720. https://doi.org/10.1016/S0040-4039(00)75799-6.

(138) Hamada, T.; Sugawara, T.; Matsunaga, S.; Fusetani, N. Polytheonamides, Unprecedented Highly Cytotoxic Polypeptides from the Marine Sponge Theonella Swinhoei 2. Structure Elucidation. *Tetrahedron Letters* **1994**, *35* (4), 609–612. https://doi.org/10.1016/S0040-4039(00)75851-5.

(139) Hamada, T.; Matsunaga, S.; Yano, G.; Fusetani, N. Polytheonamides A and B, Highly Cytotoxic, Linear Polypeptides with Unprecedented Structural Features, from the Marine Sponge, Theonella Swinhoei. *J. Am. Chem. Soc.* **2005**, *127* (1), 110–118. https://doi.org/10.1021/ja045749e.

(140) S, O.; I, M.; S, M.; N, F. [A Channel-Forming Peptide Toxin: Polytheonamide from Marine Sponge (Theonella Swinhoei)]. *Nihon Yakurigaku Zasshi* **1997**, *110 Suppl 1*, 195P-198P. https://doi.org/10.1254/fpj.110.supplement_195.

(141) Hamada, T.; Matsunaga, S.; Fujiwara, M.; Fujita, K.; Hirota, H.; Schmucki, R.; Güntert, P.; Fusetani, N. Solution Structure of Polytheonamide B, a Highly Cytotoxic Nonribosomal Polypeptide from Marine Sponge. *J. Am. Chem. Soc.* **2010**, *132* (37), 12941–12945. https://doi.org/10.1021/ja104616z.

(142) Iwamoto, M.; Shimizu, H.; Muramatsu, I.; Oiki, S. A Cytotoxic Peptide from a Marine Sponge Exhibits Ion Channel Activity through Vectorial-Insertion into the Membrane. *FEBS Letters* **2010**, *584* (18), 3995–3999. https://doi.org/10.1016/j.febslet.2010.08.007.

(143) Matsuoka, S.; Shinohara, N.; Takahashi, T.; Iida, M.; Inoue, M. Functional Analysis of Synthetic Substructures of Polytheonamide B: A Transmembrane Channel-Forming Peptide. *Angewandte Chemie* **2011**, *123* (21), 4981–4985. https://doi.org/10.1002/ange.201101533.

(144) Shinohara, N.; Itoh, H.; Matsuoka, S.; Inoue, M. Selective Modification of the N-Terminal Structure of Polytheonamide B Significantly Changes Its Cytotoxicity and Activity as an Ion Channel. *ChemMedChem* **2012**, *7* (10), 1770–1773. https://doi.org/10.1002/cmdc.201200142.

(145) Hayata, A.; Itoh, H.; Matsutaka, S.; Inoue, M. Dual Chemical Modification of a Polytheonamide Mimic: Rational Design and Synthesis of Ion-Channel-Forming 48-Mer Peptides with Potent Cytotoxicity. *Chemistry – A European Journal* **2016**, *22* (10), 3370–3377. https://doi.org/10.1002/chem.201504632.

(146) Itoh, H.; Matsutaka, S.; Kuranaga, T.; Inoue, M. Control of the Cytotoxicity of Dansylated Polytheonamide Mimic, an Artificial Peptide Ion Channel, by Modification of the N-Terminal Structure. *Tetrahedron Letters* **2014**, *55* (3), 728–731. https://doi.org/10.1016/j.tetlet.2013.12.007.

(147) Itoh, H.; Matsuoka, S.; Kreir, M.; Inoue, M. Design, Synthesis and Functional Analysis of Dansylated Polytheonamide Mimic: An Artificial Peptide Ion Channel. *J. Am. Chem. Soc.* **2012**, *134* (34), 14011–14018. https://doi.org/10.1021/ja303831a.

(148) Itoh, H.; Inoue, M. Structural Permutation of Potent Cytotoxin, Polytheonamide B: Discovery of Cytotoxic Peptide with Altered Activity. *ACS Med. Chem. Lett.* **2013**, *4* (1), 52–56. https://doi.org/10.1021/ml300264c.

(149) Renevey, A.; Riniker, S. The Importance of N-Methylations for the Stability of the $$\beta ^{6.3}$$β6.3-Helical Conformation of Polytheonamide B. *Eur Biophys J* **2017**, *46* (4), 363–374. https://doi.org/10.1007/s00249-016-1179-1.

(150) Iwamoto, M.; Matsunaga, S.; Oiki, S. Paradoxical One-Ion Pore Behavior of the Long β-Helical Peptide of Marine Cytotoxic Polytheonamide B. *Sci Rep* **2014**, *4* (1), 1–7. https://doi.org/10.1038/srep03636.

(151) Kalathingal, M.; Sumikama, T.; Mori, T.; Oiki, S.; Saito, S. Structure and Dynamics of Solvent Molecules inside the Polytheonamide B Channel in Different Environments: A Molecular Dynamics Study. *Physical Chemistry Chemical Physics* **2018**, *20* (5), 3334–3348. https://doi.org/10.1039/C7CP06299K.

(152) Matsuki, Y.; Iwamoto, M.; Mita, K.; Shigemi, K.; Matsunaga, S.; Oiki, S. Rectified Proton Grotthuss Conduction Across a Long Water-Wire in the Test Nanotube of the Polytheonamide B Channel. *J. Am. Chem. Soc.* **2016**, *138* (12), 4168–4177. https://doi.org/10.1021/jacs.5b13377.

(153) Hayata, A.; Itoh, H.; Inoue, M. Solid-Phase Total Synthesis and Dual Mechanism of Action of the Channel-Forming 48-Mer Peptide Polytheonamide B. *J. Am. Chem. Soc.* **2018**, *140* (33), 10602–10611. https://doi.org/10.1021/jacs.8b06755.

(154) Xue, Y.-W.; Hayata, A.; Itoh, H.; Inoue, M. Biological Effects of a Simplified Synthetic Analogue of Ion-Channel-Forming Polytheonamide B on Plasma Membrane and Lysosomes. *Chemistry – A European Journal* **2019**, *25* (66), 15198–15204. https://doi.org/10.1002/chem.201903974.

(155) Bhushan, A.; Egli, P. J.; Peters, E. E.; Freeman, M. F.; Piel, J. Genome Mining- and Synthetic Biology-Enabled Production of Hypermodified Peptides. *Nat. Chem.* **2019**, *11* (10), 931–939. https://doi.org/10.1038/s41557-019-0323-9.

(156) Bösch, N.; Mariana, B.; Greczmiel, U.; Morinaka, B.; Gugger, M.; Oxenius, A.; Vagstad, A. L.; Piel, J. Landornamides, Antiviral Ornithine-Containing Ribosomal Peptides Discovered by Proteusin Mining. *Angewandte Chemie n/a* (n/a). https://doi.org/10.1002/ange.201916321.

(157) Shimamura, H.; Gouda, H.; Nagai, K.; Hirose, T.; Ichioka, M.; Furuya, Y.; Kobayashi, Y.; Hirono, S.; Sunazuka, T.; Ōmura, S. Structure Determination and Total Synthesis of Bottromycin A2: A Potent Antibiotic against MRSA and VRE. *Angewandte Chemie International Edition* **2009**, *48* (5), 914–917. https://doi.org/10.1002/anie.200804138.

(158) Tanaka, N.; Nishimura, T.; Nakamura, S.; Umezawa, H. Biological Studies on Bottromycin A and Its Hydrazide. *J. Antibiot.* **1966**, *19* (4), 149–154. https://doi.org/10.7164/antibiotics.19.149.

(159) Tanaka, N.; Nishimura, T.; Nakamura, S.; Umezawa, H.; Hayami, T. ACTIVITY OF BOTTROMYGIN AGAINST MYCOPLASMA GALLISEPTICUM. *J. Antibiot.* **1968**, *21* (1), 75–76. https://doi.org/10.7164/antibiotics.21.75.

(160) Waisvisz, J. M.; van der Hoeven, M. G.; van Peppen, J.; Zwennis, W. C. M. Bottromycin. I. A New Sulfur-Containing Antibiotic. *J. Am. Chem. Soc.* **1957**, *79* (16), 4520–4521. https://doi.org/10.1021/ja01573a072.

(161) Tanaka, N.; Sashikata, K.; Yamaguchi, H.; Umezawa, H. Inhibition of Protein Synthesis by Bottromycin A2 and Its Hydrazide. *J Biochem* **1966**, *60* (4), 405–410. https://doi.org/10.1093/oxfordjournals.jbchem.a128451.

(162) Lin, Y.-C.; Tanaka, N. Mechanism of Action of Bottromycin in Polypeptide Biosynthesis. *J Biochem* **1968**, *63* (1), 1–7. https://doi.org/10.1093/oxfordjournals.jbchem.a128735.

(163) Lin, Y.-C.; Kinoshita, T.; Tanaka, N. MECHANISM OF PROTEIN SYNTHESIS INHIBITION BY BOTTROMYCIN A2 : STUDIES WITH PUROMYCIN. *J. Antibiot.* **1968**, *21* (8), 471–476. https://doi.org/10.7164/antibiotics.21.471.

(164) Kinoshita, T.; Tanaka, N. ON THE SITE OF ACTION OF BOTTROMYGIN A2. *J. Antibiot.* **1970**, *23* (6), 311–312. https://doi.org/10.7164/antibiotics.23.311.

(165) Cundliffe, E.; McQuillen, K. Bacterial Protein Synthesis: The Effects of Antibiotics. *Journal of Molecular Biology* **1967**, *30* (1), 137–146. https://doi.org/10.1016/0022-2836(67)90249-5.

(166) Pestka, S. Studies on the Formation of Transfer Ribonucleic Acid-Ribosome Complexes: VIII. Survey of the Effect of Antibiotics on N-Acetyl-Phenylalanyl-Puromycin Formation: Possible Mechanism of Chloramphenicol Action. *Archives of Biochemistry and Biophysics* **1970**, *136* (1), 80–88. https://doi.org/10.1016/0003-9861(70)90329-2.

(167) Pestka, S. Studies on the Formation of Transfer Ribonucleic Acid-Ribosome Complexes: IX. Effect of Antibiotics on Translocation and Peptide Bond Formation. *Archives of Biochemistry and Biophysics* **1970**, *136* (1), 89–96. https://doi.org/10.1016/0003-9861(70)90330-9.

(168) Pestka, S.; Brot, N. Studies on the Formation of Transfer Ribonucleic Acid-Ribosome Complexes XV. EFFECT OF ANTIBIOTICS ON STEPS OF BACTERIAL PROTEIN SYNTHESIS: SOME NEW RIBOSOMAL INHIBITORS OF TRANSLOCATION. *J. Biol. Chem.* **1971**, *246* (24), 7715–7722.

(169) Tanaka, N.; Lin, Y.-C.; Okuyama, A. Studies on Translocation of F-Met-TRNA and Peptidyl-TRNA with Antibiotics. *Biochemical and Biophysical Research Communications* **1971**, *44* (2), 477–483. https://doi.org/10.1016/0006-291X(71)90626-7.

(170) Pestka, S. Studies on Transfer Ribonucleic Acid-Ribosome Complexes XIX. EFFECT OF ANTIBIOTICS ON PEPTIDYL PUROMYCIN SYNTHESIS ON POLYRIBOSOMES FROM ESCHERICHIA COLI. *J. Biol. Chem.* **1972**, *247* (14), 4669–4678.

(171) Otaka, T.; Kaji, A. Mode of Action of Bottromycin A2. Release of Aminoacyl- or Peptidyl-TRNA from Ribosomes. *J. Biol. Chem.* **1976**, *251* (8), 2299–2306.

(172) Otaka, T.; Kaji, A. Mode of Action of Bottromycin A2: Effect on Peptide Bond Formation. *FEBS Letters* **1981**, *123* (2), 173–176. https://doi.org/10.1016/0014-5793(81)80280-3.

(173) Otaka, T.; Kaji, A. Mode of Action of Bottromycin A2: Effect of Bottromycin A2 on Polysomes. *FEBS Letters* **1983**, *153* (1), 53–59. https://doi.org/10.1016/0014-5793(83)80118-5.

(174) Pestka, S. Thiostrepton: A Ribosomal Inhibitor of Translocation. *Biochemical and Biophysical Research Communications* **1970**, *40* (3), 667–674. https://doi.org/10.1016/0006-291X(70)90956-3.

(175) Sopori, M. L.; Lengyel, P. Components of the 50S Ribosomal Subunit Involved in GTP Cleavage. *Biochemical and Biophysical Research Communications* **1972**, *46* (1), 238–244. https://doi.org/10.1016/0006-291X(72)90655-9.

(176) Cundliffe, E. The Mode of Action of Thiostrepton in Vivo. *Biochemical and Biophysical Research Communications* **1971**, *44* (4), 912–917. https://doi.org/10.1016/0006-291X(71)90798-4.

(177) Weisblum, B.; Demohn, V. Thiostrepton, an Inhibitor of 50S Ribosome Subunit Function. *Journal of Bacteriology* **1970**, *101* (3), 1073–1075.

(178) Ll, J. M.; Cabrer, B.; Parmeggiani, A.; Azquez, D. V. Inhibition by Siomycin and Thiostrepton of Both Aminoacyl-TRNA and Factor G Binding to Ribosomes. *PNAS* **1971**, *68* (8), 1796–1800. https://doi.org/10.1073/pnas.68.8.1796.

(179) Bodley, J. W.; Lin, L.; Highland, J. H. Studies on Translocation VI: Thiostrepton Prevents the Formation of a Ribosome-G Factor-Guanine Nucleotide Complex. *Biochemical and Biophysical Research Communications* **1970**, *41* (6), 1406–1411. https://doi.org/10.1016/0006-291X(70)90543-7.

(180) Weisblum, B.; Demohn, V. Inhibition by Thiostrepton of the Formation of a Ribosome-Bound Guanine Nucleotide Complex. *FEBS Letters* **1970**, *11* (3), 149–152. https://doi.org/10.1016/0014-5793(70)80515-4.

(181) Pestka, S.; Hintikka, H. Studies on the Formation of Ribonucleic Acid-Ribosome Complexes XVI. EFFECT OF RIBOSOMAL TRANSLOCATION INHIBITORS ON POLYRIBOSOMES. *J. Biol. Chem.* **1971**, *246* (24), 7723–7730.

(182) Highland, J. H.; Lin, L.; Bodley, J. W. Translocation. VIII. Protection of Ribosomes Frm Thiostrepton Inactivation by the Binding of G Factor and Guanosine Diphosphate. *Biochemistry* **1971**, *10* (24), 4404–4409. https://doi.org/10.1021/bi00800a009.

(183) Cundliffe, E. Mechanism of Resistance to Thiostrepton in the Producing-Organism Streptomyces Azureus. *Nature* **1978**, *272* (5656), 792–795. https://doi.org/10.1038/272792a0.

(184) Cundliffe, E.; Thompson, J. Ribose Methylation and Resistance to Thiostrepton. *Nature* **1979**, *278* (5707), 859–861. https://doi.org/10.1038/278859a0.

(185) Thompson, J.; Cundliffe, E. Resistance to Thiostrepton, Siomycin, and Sporangiomycin in Actinomycetes That Produce Them. *Journal of Bacteriology* **1980**, *142* (2), 455–461.

(186) Cundliffe, E.; Dixon, P.; Stark, M.; Stöffler, G.; Ehrlich, R.; Stöffler-Meilicke, M.; Cannon, M. Ribosomes in Thiostrepton-Resistant Mutants of Bacillus Megaterium Lacking a Single 50 S Subunit Protein. *Journal of Molecular Biology* **1979**, *132* (2), 235–252. https://doi.org/10.1016/0022-2836(79)90393-0.

(187) Wienen, B.; Ehrlich, R.; Stöffler-Meilicke, M.; Stöffler, G.; Smith, I.; Weiss, D.; Vince, R.; Pestka, S. Ribosomal Protein Alterations in Thiostrepton- and Micrococcin-Resistant Mutants of Bacillus Subtilis. *J. Biol. Chem.* **1979**, *254* (16), 8031–8041.

(188) Pestka, S.; Weiss, D.; Vince, R.; Wienen, B.; Stöffler, G.; Smith, I. Thiostrepton-Resistant Mutants of Bacillus Subtilis: Localization of Resistance to the 50S Subunit. *Molec. gen. Genet.* **1976**, *144* (3), 235–241. https://doi.org/10.1007/BF00341721.

(189) Smith, I.; Weiss, D.; Pestka, S. A Micrococcin-Resistant Mutant of Bacillus Subtilis: Localization of Resistance to the 50S Subunit. *Molec. gen. Genet.* **1976**, *144* (3), 231–233. https://doi.org/10.1007/BF00341720.

(190) Highland, J. H.; Howard, G. A.; Ochsner, E.; Hasenbank, R.; Gordon, J.; Stöffler, G. Identification of a Ribosomal Protein Necessary for Thiostrepton Binding to Escherichia Coli Ribosomes. *J. Biol. Chem.* **1975**, *250* (3), 1141–1145.

(191) Baumann, S.; Schoof, S.; Bolten, M.; Haering, C.; Takagi, M.; Shin-ya, K.; Arndt, H.-D. Molecular Determinants of Microbial Resistance to Thiopeptide Antibiotics. *J. Am. Chem. Soc.* **2010**, *132* (20), 6973–6981. https://doi.org/10.1021/ja909317n.

(192) Thompson, J.; Schmidt, F.; Cundliffe, E. Site of Action of a Ribosomal RNA Methylase Conferring Resistance to Thiostrepton. *J. Biol. Chem.* **1982**, *257* (14), 7915–7917.

(193) Hummel, H.; Böck, A. Thiostrepton Resistance Mutations in the Gene for 23S Ribosomal RNA of Halobacteria. *Biochimie* **1987**, *69* (8), 857–861. https://doi.org/10.1016/0300-9084(87)90212-4.

(194) Thompson, J.; Cundliffe, E.; Dahlberg, A. E. Site-Directed Mutagenesis of Escherichia Coli 23 S Ribosomal RNA at Position 1067 within the GTP Hydrolysis Centre. *Journal of Molecular Biology* **1988**, *203* (2), 457–465. https://doi.org/10.1016/0022-2836(88)90012-5.

(195) Egebjerg, J.; Douthwaite, S.; Garrett, R. A. Antibiotic Interactions at the GTPase-Associated Centre within Escherichia Coli 23S RRNA. *The EMBO Journal* **1989**, *8* (2), 607–611. https://doi.org/10.1002/j.1460-2075.1989.tb03415.x.

(196) Hausner, T. P.; Geigenmüller, U.; Nierhaus, K. H. The Allosteric Three-Site Model for the Ribosomal Elongation Cycle. New Insights into the Inhibition Mechanisms of Aminoglycosides, Thiostrepton, and Viomycin. *J. Biol. Chem.* **1988**, *263* (26), 13103–13111.

(197) Ryan, P. C.; Lu, M.; Draper, D. E. Recognition of the Highly Conserved GTPase Center of 23 S Ribosomal RNA by Ribosomal Protein L11 and the Antibiotic Thiostrepton. *Journal of Molecular Biology* **1991**, *221* (4), 1257–1268. https://doi.org/10.1016/0022-2836(91)90932-V.

(198) Egebjerg, J.; Douthwaite, S. R.; Liljas, A.; Garrett, R. A. Characterization of the Binding Sites of Protein L11 and the L10.(L12)4 Pentameric Complex in the GTPase Domain of 23 S Ribosomal RNA from Escherichia Coli. *Journal of Molecular Biology* **1990**, *213* (2), 275–288. https://doi.org/10.1016/S0022-2836(05)80190-1.

(199) Thompson, J.; Cundliffe, E. The Binding of Thiostrepton to 23S Ribosomal RNA. *Biochimie* **1991**, *73* (7), 1131–1135. https://doi.org/10.1016/0300-9084(91)90156-U.

(200) Thompson, J.; Cundliffe, E.; Stark, M. Binding of Thiostrepton to a Complex of 23-S RRNA with Ribosomal Protein L11. *European Journal of Biochemistry* **1979**, *98* (1), 261–265. https://doi.org/10.1111/j.1432-1033.1979.tb13184.x.

(201) Rosendahl, G.; Douthwaite, S. The Antibiotics Micrococcin and Thiostrepton Interact Directly with 23S RRNA Nucleotides 1067A and 1095A. *Nucleic Acids Res* **1994**, *22* (3), 357–363. https://doi.org/10.1093/nar/22.3.357.

(202) Rosendahl, G.; Douthwaite, S. Ribosomal Proteins L11 and L10.(L12)4 and the Antibiotic Thiostrepton Interact with Overlapping Regions of the 23 S RRNA Backbone in the Ribosomal GTPase Centre. *Journal of Molecular Biology* **1993**, *234* (4), 1013–1020. https://doi.org/10.1006/jmbi.1993.1655.

(203) Xing, Y.; Draper, D. E. Cooperative Interactions of RNA and Thiostrepton Antibiotic with Two Domains of Ribosomal Protein L11. *Biochemistry* **1996**, *35* (5), 1581–1588. https://doi.org/10.1021/bi952132o.

(204) Conn, G. L.; Draper, D. E.; Lattman, E. E.; Gittis, A. G. Crystal Structure of a Conserved Ribosomal Protein-RNA Complex. *Science* **1999**, *284* (5417), 1171–1174. https://doi.org/10.1126/science.284.5417.1171.

(205) Porse, B. T.; Leviev, I.; Mankin, A. S.; Garrett, R. A. The Antibiotic Thiostrepton Inhibits a Functional Transition within Protein L11 at the Ribosomal GTPase Centre11Edited by D. E. Draper. *Journal of Molecular Biology* **1998**, *276* (2), 391–404. https://doi.org/10.1006/jmbi.1997.1541.

(206) Wimberly, B. T.; Guymon, R.; McCutcheon, J. P.; White, S. W.; Ramakrishnan, V. A Detailed View of a Ribosomal Active Site: The Structure of the L11–RNA Complex. *Cell* **1999**, *97* (4), 491–502. https://doi.org/10.1016/S0092-8674(00)80759-X.

(207) Schoof, S.; Baumann, S.; Ellinger, B.; Arndt, H.-D. A Fluorescent Probe for the 70 S-Ribosomal GTPase-Associated Center. *ChemBioChem* **2009**, *10* (2), 242–245. https://doi.org/10.1002/cbic.200800642.

(208) Ilin, S.; Hoskins, A.; Ohlenschläger, O.; Jonker, H. R. A.; Schwalbe, H.; Wöhnert, J. Domain Reorientation and Induced Fit upon RNA Binding: Solution Structure and Dynamics of Ribosomal Protein L11 from Thermotoga Maritima. *ChemBioChem* **2005**, *6* (9), 1611–1618. https://doi.org/10.1002/cbic.200500091.

(209) Baumann, S.; Schoof, S.; Harkal, S. D.; Arndt, H.-D. Mapping the Binding Site of Thiopeptide Antibiotics by Proximity-Induced Covalent Capture. *J. Am. Chem. Soc.* **2008**, *130* (17), 5664–5666. https://doi.org/10.1021/ja710608w.

(210) Blyn, L. B.; Risen, L. M.; Griffey, R. H.; Draper, D. E. The RNA-Binding Domain of Ribosomal Protein L11 Recognizes an RRNA Tertiary Structure Stabilized by Both Thiostrepton and Magnesium Ion. *Nucleic Acids Res* **2000**, *28* (8), 1778–1784. https://doi.org/10.1093/nar/28.8.1778.

(211) Porse, B. T.; Cundliffe, E.; Garrett, R. A. The Antibiotic Micrococcin Acts on Protein L11 at the Ribosomal GTPase Centre11Edited by D. E. Draper. *Journal of Molecular Biology* **1999**, *287* (1), 33–45. https://doi.org/10.1006/jmbi.1999.2600.

(212) Bausch, S. L.; Poliakova, E.; Draper, D. E. Interactions of the N-Terminal Domain of Ribosomal Protein L11 with Thiostrepton and RRNA. *J. Biol. Chem.* **2005**, *280* (33), 29956–29963. https://doi.org/10.1074/jbc.M504182200.

(213) Seo, H.-S.; Kiel, M.; Pan, D.; Raj, V. S.; Kaji, A.; Cooperman, B. S. Kinetics and Thermodynamics of RRF, EF-G, and Thiostrepton Interaction on the Escherichia Coli Ribosome. *Biochemistry* **2004**, *43* (40), 12728–12740. https://doi.org/10.1021/bi048927p.

(214) Bowen, W. S.; Dyke, N. V.; Murgola, E. J.; Lodmell, J. S.; Hill, W. E. Interaction of Thiostrepton and Elongation Factor-G with the Ribosomal Protein L11-Binding Domain. *J. Biol. Chem.* **2005**, *280* (4), 2934–2943. https://doi.org/10.1074/jbc.M407008200.

(215) Gonzalez, R. L.; Chu, S.; Puglisi, J. D. Thiostrepton Inhibition of TRNA Delivery to the Ribosome. *RNA* **2007**, *13* (12), 2091–2097. https://doi.org/10.1261/rna.499407.

(216) Cameron, D. M.; Thompson, J.; Gregory, S. T.; March, P. E.; Dahlberg, A. E. Thiostrepton-Resistant Mutants of Thermus Thermophilus. *Nucleic Acids Res* **2004**, *32* (10), 3220–3227. https://doi.org/10.1093/nar/gkh644.

(217) Jonker, H. R. A.; Ilin, S.; Grimm, S. K.; Wöhnert, J.; Schwalbe, H. L11 Domain Rearrangement upon Binding to RNA and Thiostrepton Studied by NMR Spectroscopy. *Nucleic Acids Res* **2007**, *35* (2), 441–454. https://doi.org/10.1093/nar/gkl1066.

(218) Lentzen, G.; Klinck, R.; Matassova, N.; Aboul-ela, F.; Murchie, A. I. H. Structural Basis for Contrasting Activities of Ribosome Binding Thiazole Antibiotics. *Chemistry & Biology* **2003**, *10* (8), 769–778. https://doi.org/10.1016/S1074-5521(03)00173-X.

(219) Harms, J. M.; Wilson, D. N.; Schluenzen, F.; Connell, S. R.; Stachelhaus, T.; Zaborowska, Z.; Spahn, C. M. T.; Fucini, P. Translational Regulation via L11: Molecular Switches on the Ribosome Turned On and Off by Thiostrepton and Micrococcin. *Molecular Cell* **2008**, *30* (1), 26–38. https://doi.org/10.1016/j.molcel.2008.01.009.

(220) Cameron, D. M.; Thompson, J.; March, P. E.; Dahlberg, A. E. Initiation Factor IF2, Thiostrepton and Micrococcin Prevent the Binding of Elongation Factor G to the Escherichia Coli Ribosome. *Journal of Molecular Biology* **2002**, *319* (1), 27–35. https://doi.org/10.1016/S0022-2836(02)00235-8.

(221) Brandi, L.; Marzi, S.; Fabbretti, A.; Fleischer, C.; Hill, W. E.; Gualerzi, C. O.; Stephen Lodmell, J. The Translation Initiation Functions of IF2: Targets for Thiostrepton Inhibition. *Journal of Molecular Biology* **2004**, *335* (4), 881–894. https://doi.org/10.1016/j.jmb.2003.10.067.

(222) Lockwood, A. H.; Sarkar, P.; Maitra, U.; Brot, N.; Weissbach, H. Effect of Thiostrepton on Polypeptide Chain Initiation in Escherichia Coli. *J. Biol. Chem.* **1974**, *249* (18), 5831–5834.

(223) Sarkar, P.; Stringer, E. A.; Maitra, U. Thiostrepton Inhibition of Initiation Factor 1 Activity in Polypeptide Chain Initiation in Escherichia Coli. *Proc Natl Acad Sci U S A* **1974**, *71* (12), 4986–4990.

(224) Naaktgeboren, N.; Roobol, K.; Gubbens, J.; Voorma, H. O. The Mode of Action of Thiostrepton in the Initiation of Protein Synthesis. *European Journal of Biochemistry* **1976**, *70* (1), 39–47. https://doi.org/10.1111/j.1432-1033.1976.tb10953.x.

(225) Masuda, T.; Petrov, A. N.; Iizuka, R.; Funatsu, T.; Puglisi, J. D.; Uemura, S. Initiation Factor 2, TRNA, and 50S Subunits Cooperatively Stabilize MRNAs on the Ribosome during Initiation. *PNAS* **2012**, *109* (13), 4881–4885. https://doi.org/10.1073/pnas.1118452109.

(226) Murakami, T.; Holt, T. G.; Thompson, C. J. Thiostrepton-Induced Gene Expression in Streptomyces Lividans. *Journal of Bacteriology* **1989**, *171* (3), 1459–1466. https://doi.org/10.1128/jb.171.3.1459-1466.1989.

(227) Holmes, D. j.; Caso, J. l.; Thompson, C. j. Autogenous Transcriptional Activation of a Thiostrepton-Induced Gene in Streptomyces Lividans. *The EMBO Journal* **1993**, *12* (8), 3183–3191. https://doi.org/10.1002/j.1460-2075.1993.tb05987.x.

(228) Guilfoile, P. G.; Hutchinson, C. R. A Bacterial Analog of the Mdr Gene of Mammalian Tumor Cells Is Present in Streptomyces Peucetius, the Producer of Daunorubicin and Doxorubicin. *PNAS* **1991**, *88* (19), 8553–8557. https://doi.org/10.1073/pnas.88.19.8553.

(229) Plater, R.; Robinson, J. A. Cloning and Sequence of a Gene Encoding Macrotetrolide Antibiotic Resistance from Streptomyces Griseus. *Gene* **1992**, *112* (1), 117–122. https://doi.org/10.1016/0378-1119(92)90312-D.

(230) Yun, B.-S.; Seto, H. Promoinducin, a Novel Thiopeptide Produced by Streptomyces Sp. SF2741. *Bioscience, Biotechnology, and Biochemistry* **1995**, *59* (5), 876–880. https://doi.org/10.1271/bbb.59.876.

(231) Yun, B.-S.; Hidaka, T.; Furihata, K.; Seto, H. MICROBIAL METABOLITES WITH Tip A PROMOTER INDUCING ACTIVITY. *J. Antibiot.* **1994**, *47* (9), 969–975. https://doi.org/10.7164/antibiotics.47.969.

(232) Yun, B.-S.; Hidaka, T.; Furihata, K.; Seto, H. PROMOTHIOCINS A AND B, NOVEL THIOPEPTIDES WITH A Tip A PROMOTER INDUCING ACTIVITY PRODUCED BY Streptomyces Sp. SF2741. *J. Antibiot.* **1994**, *47* (4), 510–514. https://doi.org/10.7164/antibiotics.47.510.

(233) Yun, B.-S.; Hidaka, T.; Furihata, K.; Seto, H. Thiotipin, a Novel Thiopeptide with a TipA Promoter Inducing Activity Produced by Streptomyces Sp. DT31. *Tetrahedron* **1994**, *50* (40), 11659–11664. https://doi.org/10.1016/S0040-4020(01)85660-X.

(234) Chiu, M. L.; Folcher, M.; Griffin, P.; Holt, T.; Klatt, T.; Thompson, C. J. Characterization of the Covalent Binding of Thiostrepton to a Thiostrepton-Induced Protein from Streptomyces Lividans. *Biochemistry* **1996**, *35* (7), 2332–2341. https://doi.org/10.1021/bi952073e.

(235) Ali, N.; Herron, P. R.; Evans, M. C.; Dyson, P. J. Osmotic Regulation of the Streptomyces Lividans Thiostrepton-Inducible Promoter, PtipA. *Microbiology*, **2002**, *148* (2), 381–390. https://doi.org/10.1099/00221287-148-2-381.

(236) Myers, C. L.; Harris, J.; Yeung, J. C. K.; Honek, J. F. Molecular Interactions between Thiostrepton and the TipAS Protein from Streptomyces Lividans. *ChemBioChem* **2014**, *15* (5), 681–687. https://doi.org/10.1002/cbic.201300724.

(237) Chiu, M. L.; Folcher, M.; Katoh, T.; Puglia, A. M.; Vohradsky, J.; Yun, B.-S.; Seto, H.; Thompson, C. J. Broad Spectrum Thiopeptide Recognition Specificity of TheStreptomyces Lividans TipAL Protein and Its Role in Regulating Gene Expression. *J. Biol. Chem.* **1999**, *274* (29), 20578–20586. https://doi.org/10.1074/jbc.274.29.20578.

(238) Chiu, M. L.; Viollier, P. H.; Katoh, T.; Ramsden, J. J.; Thompson, C. J. Ligand-Induced Changes in the Streptomyces Lividans TipAL Protein Imply an Alternative Mechanism of Transcriptional Activation for MerR-Like Proteins. *Biochemistry* **2001**, *40* (43), 12950–12958. https://doi.org/10.1021/bi010328k.

(239) Kahmann, J. D.; Sass, H.-J.; Allan, M. G.; Seto, H.; Thompson, C. J.; Grzesiek, S. Structural Basis for Antibiotic Recognition by the TipA Class of Multidrug-Resistance Transcriptional Regulators. *The EMBO Journal* **2003**, *22* (8), 1824–1834. https://doi.org/10.1093/emboj/cdg181.

(240) Habazettl, J.; Allan, M.; Jensen, P. R.; Sass, H.-J.; Thompson, C. J.; Grzesiek, S. Structural Basis and Dynamics of Multidrug Recognition in a Minimal Bacterial Multidrug Resistance System. *PNAS* **2014**, *111* (51), E5498–E5507. https://doi.org/10.1073/pnas.1412070111.

(241) Yun, B.-S.; Hidaka, T.; Kuzuyama, T.; Seto, H. Thiopeptide Non-Producing Streptomyces Species Carry the TipA Gene: A Clue to Its Function. *J. Antibiot.* **2001**, *54* (4), 375–378. https://doi.org/10.7164/antibiotics.54.375.

(242) Vasant Kumar, C.; Martín, J. F. Thiostrepton Induced Proteins in Streptomyces, Amycolatopsis and Nocardia Species. *FEMS Microbiol Lett* **1994**, *118* (1–2), 107–111. https://doi.org/10.1111/j.1574-6968.1994.tb06811.x.

(243) Dong, L.; Nakashima, N.; Tamura, N.; Tamura, T. Isolation and Characterization of the Rhodococcus Opacus Thiostrepton-Inducible Genes TipAL and TipAS: Application for Recombinant Protein Expression in Rhodococcus. *FEMS Microbiol Lett* **2004**, *237* (1), 35–40. https://doi.org/10.1111/j.1574-6968.2004.tb09675.x.

(244) McConkey, G. A.; Rogers, M. J.; McCutchan, T. F. Inhibition of Plasmodium Falciparum Protein Synthesis TARGETING THE PLASTID-LIKE ORGANELLE WITH THIOSTREPTON. *J. Biol. Chem.* **1997**, *272* (4), 2046–2049. https://doi.org/10.1074/jbc.272.4.2046.

(245) Chaubey, S.; Kumar, A.; Singh, D.; Habib, S. The Apicoplast of Plasmodium Falciparum Is Translationally Active. *Molecular Microbiology* **2005**, *56* (1), 81–89. https://doi.org/10.1111/j.1365-2958.2005.04538.x.

(246) Clough, B.; Strath, M.; Preiser, P.; Denny, P.; Wilson, I. (R J. M.). Thiostrepton Binds to Malarial Plastid RRNA. *FEBS Letters* **1997**, *406* (1–2), 123–125. https://doi.org/10.1016/S0014-5793(97)00241-X.

(247) Tarr, S. J.; Nisbet, R. E. R.; Howe, C. J. Transcript-Level Responses of Plasmodium Falciparum to Thiostrepton. *Molecular and Biochemical Parasitology* **2011**, *179* (1), 37–41. https://doi.org/10.1016/j.molbiopara.2011.05.004.

(248) Goodman, C. D.; Su, V.; McFadden, G. I. The Effects of Anti-Bacterials on the Malaria Parasite Plasmodium Falciparum. *Molecular and Biochemical Parasitology* **2007**, *152* (2), 181–191. https://doi.org/10.1016/j.molbiopara.2007.01.005.

(249) Rogers, M. J.; Bukhman, Y. V.; McCutchan, T. F.; Draper, D. E. Interaction of Thiostrepton with an RNA Fragment Derived from the Plastid-Encoded Ribosomal RNA of the Malaria Parasite. *RNA* **1997**, *3* (8), 815–820.

(250) Rogers, M. J.; Cundliffe, E.; McCutchan, T. F. The Antibiotic Micrococcin Is a Potent Inhibitor of Growth and Protein Synthesis in the Malaria Parasite. *Antimicrobial Agents and Chemotherapy* **1998**, *42* (3), 715–716. https://doi.org/10.1128/AAC.42.3.715.

(251) AbouLaila, M.; Munkhjargal, T.; Sivakumar, T.; Ueno, A.; Nakano, Y.; Yokoyama, M.; Yoshinari, T.; Nagano, D.; Katayama, K.; El-Bahy, N.; Yokoyama, N.; Igarashi, I. Apicoplast-Targeting Antibacterials Inhibit the Growth of Babesia Parasites. *Antimicrobial Agents and Chemotherapy* **2012**, *56* (6), 3196–3206. https://doi.org/10.1128/AAC.05488-11.

(252) Aminake, M. N.; Schoof, S.; Sologub, L.; Leubner, M.; Kirschner, M.; Arndt, H.-D.; Pradel, G. Thiostrepton and Derivatives Exhibit Antimalarial and Gametocytocidal Activity by Dually Targeting Parasite Proteasome and Apicoplast. *Antimicrobial Agents and Chemotherapy* **2011**, *55* (4), 1338–1348. https://doi.org/10.1128/AAC.01096-10.

(253) Selva, E.; Beretta, G.; Montanini, N.; Saddler, G. S.; Gastaldo, L.; Ferrari, P.; Lorenzetti, R.; Landini, P.; Ripamonti, F.; Goldstein, B. P.; Berti, M.; Montanaro, L.; Denaro, M. ANTIBIOTIC GE2270 A: A NOVEL INHIBITOR OF BACTERIAL PROTEIN SYNTHESIS. *J. Antibiot.* **1991**, *44* (7), 693–701. https://doi.org/10.7164/antibiotics.44.693.

(254) Anborgh, P. H.; Parmeggiani, A. New Antibiotic That Acts Specifically on the GTP-Bound Form of Elongation Factor Tu. *The EMBO Journal* **1991**, *10* (4), 779–784. https://doi.org/10.1002/j.1460-2075.1991.tb08009.x.

(255) Anborgh, P. H.; Parmeggiani, A. Probing the Reactivity of the GTP- and GDP-Bound Conformations of Elongation Factor Tu in Complex with the Antibiotic GE2270 A. *J. Biol. Chem.* **1993**, *268* (33), 24622–24628.

(256) Landini, P.; Soffientini, A.; Monti, F.; Lociuro, S.; Marzorati, E.; Islam, K. Antibiotics MDL 62,879 and Kirromycin Bind to Distinct and Independent Sites on Elongation Factor Tu (EF-Tu). *Biochemistry* **1996**, *35* (48), 15288–15294. https://doi.org/10.1021/bi9610818.

(257) Möhrle, V. G.; Tieleman, L. N.; Kraal, B. Elongation Factor Tu1 of the Antibiotic GE2270A ProducerPlanobispora RoseaHas an Unexpected Resistance Profile against EF-Tu Targeted Antibiotics. *Biochemical and Biophysical Research Communications* **1997**, *230* (2), 320–326. https://doi.org/10.1006/bbrc.1996.5947.

(258) Shimanaka, K.; Iinuma, H.; Hamada, M.; Ikeno, S.; Tsuchiya, K. S.; Arita, M.; Hori, M. Novel Antibiotics, Amythiamicins. *J. Antibiot.* **1995**, *48* (2), 182–184. https://doi.org/10.7164/antibiotics.48.182.

(259) Clough, B.; Rangachari, K.; Strath, M.; Preiser, P. R.; Iain Wilson, R. J. M. Antibiotic Inhibitors of Organellar Protein Synthesis in Plasmodium Falciparum. *Protist* **1999**, *150* (2), 189–195. https://doi.org/10.1016/S1434-4610(99)70021-0.

(260) Heffron, S. E.; Jurnak, F. Structure of an EF-Tu Complex with a Thiazolyl Peptide Antibiotic Determined at 2.35 Å Resolution: Atomic Basis for GE2270A Inhibition of EF-Tu,. *Biochemistry* **2000**, *39* (1), 37–45. https://doi.org/10.1021/bi9913597.

(261) Parmeggiani, A.; Krab, I. M.; Okamura, S.; Nielsen, R. C.; Nyborg, J.; Nissen, P. Structural Basis of the Action of Pulvomycin and GE2270 A on Elongation Factor Tu,. *Biochemistry* **2006**, *45* (22), 6846–6857. https://doi.org/10.1021/bi0525122.

(262) Bhat, U. G.; Zipfel, P. A.; Tyler, D. S.; Gartel, A. L. Novel Anticancer Compounds Induce Apoptosis in Melanoma Cells. *Cell Cycle* **2008**, *7* (12), 1851–1855. https://doi.org/10.4161/cc.7.12.6032.

(263) Bowling, B. D.; Doudican, N.; Manga, P.; Orlow, S. J. Inhibition of Mitochondrial Protein Translation Sensitizes Melanoma Cells to Arsenic Trioxide Cytotoxicity via a Reactive Oxygen Species Dependent Mechanism. *Cancer Chemother Pharmacol* **2008**, *63* (1), 37–43. https://doi.org/10.1007/s00280-008-0705-y.

(264) Kwok, J. M.-M.; Myatt, S. S.; Marson, C. M.; Coombes, R. C.; Constantinidou, D.; Lam, E. W.-F. Thiostrepton Selectively Targets Breast Cancer Cells through Inhibition of Forkhead Box M1 Expression. *Mol Cancer Ther* **2008**, *7* (7), 2022–2032. https://doi.org/10.1158/1535-7163.MCT-08-0188.

(265) Bhat, U. G.; Halasi, M.; Gartel, A. L. Thiazole Antibiotics Target FoxM1 and Induce Apoptosis in Human Cancer Cells. *PLoS One* **2009**, *4* (5). https://doi.org/10.1371/journal.pone.0005592.

(266) Bhat, U. G.; Halasi, M.; Gartel, A. L. FoxM1 Is a General Target for Proteasome Inhibitors. *PLoS One* **2009**, *4* (8). https://doi.org/10.1371/journal.pone.0006593.

(267) Zhang, L.; Ging, N. C.; Komoda, T.; Hanada, T.; Suzuki, T.; Watanabe, K. Antibiotic Susceptibility of Mammalian Mitochondrial Translation. *FEBS Letters* **2005**, *579* (28), 6423–6427. https://doi.org/10.1016/j.febslet.2005.09.103.

(268) Hegde, N. S.; Sanders, D. A.; Rodriguez, R.; Balasubramanian, S. The Transcription Factor FOXM1 Is a Cellular Target of the Natural Product Thiostrepton. *Nature Chem* **2011**, *3* (9), 725–731. https://doi.org/10.1038/nchem.1114.

(269) Hayakawa, Y.; Sasaki, K.; Nagai, K.; Shin-ya, K.; Furihata, K. Structure of Thioviridamide, a Novel Apoptosis Inducer from Streptomyces Olivoviridis. *J Antibiot* **2006**, *59* (1), 6–10. https://doi.org/10.1038/ja.2006.2.

(270) Kjaerulff, L.; Sikandar, A.; Zaburannyi, N.; Adam, S.; Herrmann, J.; Koehnke, J.; Müller, R. Thioholgamides: Thioamide-Containing Cytotoxic RiPP Natural Products. *ACS Chem. Biol.* **2017**, *12* (11), 2837–2841. https://doi.org/10.1021/acschembio.7b00676.

(271) Izumikawa, M.; Kozone, I.; Hashimoto, J.; Kagaya, N.; Takagi, M.; Koiwai, H.; Komatsu, M.; Fujie, M.; Satoh, N.; Ikeda, H.; Shin-ya, K. Novel Thioviridamide Derivative—JBIR-140: Heterologous Expression of the Gene Cluster for Thioviridamide Biosynthesis. *J Antibiot* **2015**, *68* (8), 533–536. https://doi.org/10.1038/ja.2015.20.

(272) Kawahara, T.; Izumikawa, M.; Kozone, I.; Hashimoto, J.; Kagaya, N.; Koiwai, H.; Komatsu, M.; Fujie, M.; Sato, N.; Ikeda, H.; Shin-ya, K. Neothioviridamide, a Polythioamide Compound Produced by Heterologous Expression of a Streptomyces Sp. Cryptic RiPP Biosynthetic Gene Cluster. *J. Nat. Prod.* **2018**, *81* (2), 264–269. https://doi.org/10.1021/acs.jnatprod.7b00607.

(273) Frattaruolo, L.; Lacret, R.; Cappello, A. R.; Truman, A. W. A Genomics-Based Approach Identifies a Thioviridamide-Like Compound with Selective Anticancer Activity. *ACS Chem. Biol.* **2017**, *12* (11), 2815–2822. https://doi.org/10.1021/acschembio.7b00677.

(274) Frattaruolo, L.; Fiorillo, M.; Brindisi, M.; Curcio, R.; Dolce, V.; Lacret, R.; Truman, A. W.; Sotgia, F.; Lisanti, M. P.; Cappello, A. R. Thioalbamide, A Thioamidated Peptide from

Amycolatopsis Alba, Affects Tumor Growth and Stemness by Inducing Metabolic Dysfunction and Oxidative Stress. *Cells* **2019**, *8* (11), 1408. https://doi.org/10.3390/cells8111408.

(275) Izawa, M.; Nagamine, S.; Aoki, H.; Hayakawa, Y. Identification of Essential Biosynthetic Genes and a True Biosynthetic Product for Thioviridamide. *The Journal of General and Applied Microbiology* **2018**, *64* (1), 50–53. https://doi.org/10.2323/jgam.2017.05.002.

(276) Takase, S.; Kurokawa, R.; Kondoh, Y.; Honda, K.; Suzuki, T.; Kawahara, T.; Ikeda, H.; Dohmae, N.; Osada, H.; Shin-ya, K.; Kushiro, T.; Yoshida, M.; Matsumoto, K. Mechanism of Action of Prethioviridamide, an Anticancer Ribosomally Synthesized and Post-Translationally Modified Peptide with a Polythioamide Structure. *ACS Chem. Biol.* **2019**, *14* (8), 1819–1828. https://doi.org/10.1021/acschembio.9b00410.

(277) Pakos-Zebrucka, K.; Koryga, I.; Mnich, K.; Ljujic, M.; Samali, A.; Gorman, A. M. The Integrated Stress Response. *EMBO reports* **2016**, *17* (10), 1374–1395. https://doi.org/10.15252/embr.201642195.

(278) Kudo, K.; Koiwai, H.; Kagaya, N.; Nishiyama, M.; Kuzuyama, T.; Shin-ya, K.; Ikeda, H. Comprehensive Derivatization of Thioviridamides by Heterologous Expression. *ACS Chem. Biol.* **2019**, *14* (6), 1135–1140. https://doi.org/10.1021/acschembio.9b00330.

Chapter 2: Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin

Adam J. DiCaprio, Arash Firouzbakht, Graham A. Hudson, Douglas A. Mitchell

2.1 Introduction

The ribosomally synthesized and post-translationally modified peptides (RiPPs) are a natural product class comprised of no fewer than two dozen structural subclasses encoded by all three domains of life.[1] RiPP biosynthesis begins with the translation of a ribosomally synthesized precursor peptide which, for the majority of RiPPs, is composed of an N-terminal "leader peptide" and a C-terminal "core peptide." A specific motif within the leader peptide, the recognition sequence, is bound by the RiPP recognition element (RRE)[2] which acts to recruit the cognate biosynthetic enzymes to modify the core peptide. During lasso peptide biosynthesis, the leader peptidase (a homolog of the enzyme transglutaminase) site-specifically removes the leader peptide in an RRE-dependent fashion.[3] Subsequently, a lasso cyclase enzyme (homologous to asparagine synthetase) presumably adenylates an Asp or Glu residue of the core peptide while folding the substrate into a pre-lasso conformation.[4] The liberated N-terminus then attacks the adenylated residue, forming a macrolactam linkage and a sterically locked lasso peptide. Some lasso peptides display additional modification that include C-terminal methyl ester formation (e.g. lassomycin[5]), arginine deimination (e.g. citrulassin[3]), or epimerization (e.g. MS-271[6]). Steric constraints within the three-dimensional structure endow lasso peptides with extraordinary resistance to heat denaturation and proteolytic degradation. On rare occasions, some lasso peptides are known to unthread after prolonged periods at elevated temperatures.[7,8]

Nearly 50 lasso peptides have been isolated to date although less than half report a biological activity. Known bioactivities include cell-surface receptor antagonism, HIV fusion inhibition, RNA polymerase inhibition, and mycobacterial Clp protease inhibition.[9] Beyond activities

possessed by the native lasso peptides, the scaffold is highly amenable to alteration, as previously characterized lasso peptides have demonstrated tolerance to core peptide variation. A major exception to this flexibility is an intolerance to variation for the two residues involved in macrolactam formation. Introduction of non-native epitopes has conferred new functions to lasso peptides. Specifically, epitope grafting has been used to endow lasso peptides with activity towards the human integrin receptor.[10,11] The rational engineering of lasso peptides is especially relevant when considering their translational potential. For instance, lasso peptides have demonstrated good stability under simulated gastrointestinal conditions[12] and possess efficacy in mouse models of *Helicobacter pylori*[13] and *Salmonella enterica*[14] infections.

Given the potential of lasso peptides for future therapeutic development, there have been significant efforts from a number of research groups to isolate and characterize novel lasso peptides. Given their disparate biological activities, lasso peptides discovered by bioassay-guided isolation are serendipitous and low throughput. By leveraging recent advances in bioinformatics, rapid progress has been made in the identification and isolation of novel natural products. Frustratingly, the compactness of lasso peptide biosynthetic gene clusters (BGCs, sometimes less than 3 kb total) and the high sequence similarity of these enzymes to counterparts in primary metabolism (i.e. transglutaminase and asparagine synthetase) complicate genome-mining efforts. Additionally, the short length and hypervariability of the lasso precursor peptides typically means the coding sequences are not found by automated gene-finders and only nearly identical sequences can be found by sequence similarity search tools such as BLAST.

To address these longstanding issues, we developed the bioinformatic algorithm RODEO (Rapid ORF Description & Evaluation Online: www.ripprodeo.org), which automates the identification of lasso peptide[3] and other RiPP BGCs[15] from genomes available in GenBank. RODEO uses

profile Hidden Markov Models and supervised machine learning to locate RiPP BGCs and predict the most likely precursor peptide, respectively. Previously, we leveraged RODEO to define the extent of lasso peptides encoded by all sequenced bacteria and archaea, which expanded the number of putative lasso peptide BGCs by two orders of magnitude and revealed significant new insights into the natural product family as a whole. Using the data set generated during this study, we selected several candidate BGCs from thermophilic bacteria for enzymatic reconstitution. We surmised that thermophilic proteins would exhibit higher general stability *in vitro*, a strategy that we have successfully applied to the study of other RiPP biosynthetic pathways.[16,17]

2.2 Results and Discussion

2.2.1 Heterologous expression of fusilassin in E. coli.

Thermobifida fusca, a mesothermophilic actinomycete, was found to encode a lasso peptide BGC with a putative precursor peptide (FusA, coding sequence not annotated as a gene on NC_007333.1) adjacent to a lasso cyclase (FusC, WP_011291592.1), RRE (FusE, WP_011291591.1), leader peptidase (FusB, WP_011291590.1), and an ABC transporter (FusD, WP_011291589.1) (Fig. 1A). Attempts to detect "fusilassin" from the native producer by mass spectrometry (MS) were not successful. We thus cloned FusA into a modified pET28 vector that provides an N-terminal fusion to maltose-binding protein (MBP, see Supporting Methods, Table S1). FusB, FusC, and FusE were cloned into a separate pACYC expression vector, and the two plasmids were co-transformed into *E. coli*. After a 3 h expression, fusilassin was readily detected by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS in the crude methanolic extracts of the induced cells (Fig. 1B).

Fusilassin represents a new lasso peptide of novel composition and is the first observed lasso peptide containing tryptophan at position one of the core peptide. It was previously believed that only small residues would be biosynthetically tolerated at this position.[9] Given that Trp is the largest proteinogenic residue, we hypothesized that fusilassin would possess intrinsic tolerance to substitution at position 1, a site that has been historically recalcitrant to substitution.[23–28] Fusilassin analogs were successfully produced via *E. coli* expression for all classes of amino acid (Fig. S1), an ability not demonstrated by any previously characterized lasso peptide. Only the Pro1 variant failed to generate a detectable lasso peptide. To discriminate if FusB, FusC, or both, were incapable of tolerating Pro1, we first sought to reconstitute fusilassin biosynthesis *in vitro* using purified proteins.

2.2.2 Enzymatic reconstitution and mutational analysis.

The genes encoding FusA, B, C, and E were individually cloned into pET28-MBP vectors, expressed in *E. coli*, and purified by amylose-affinity chromatography (Supporting Methods, Fig. S2). Production of a dehydrated product (*m/z* 2269 Da) was readily observed in reactions that included ATP and all four proteins (Fig. 2A). The 2269 Da species was further analyzed by high-resolution and tandem MS (HRMS-MS/MS), which confirmed a macrolactam linkage between Trp1 and Glu9 of the FusA core peptide (Fig. S3). Omission of ATP or FusC resulted in the formation of a mass corresponding to the linear FusA core peptide (*m/z* 2287). Generation of the core peptide was dependent upon the presence of both FusE and FusB, as omission of either protein led to no observed mass changes.

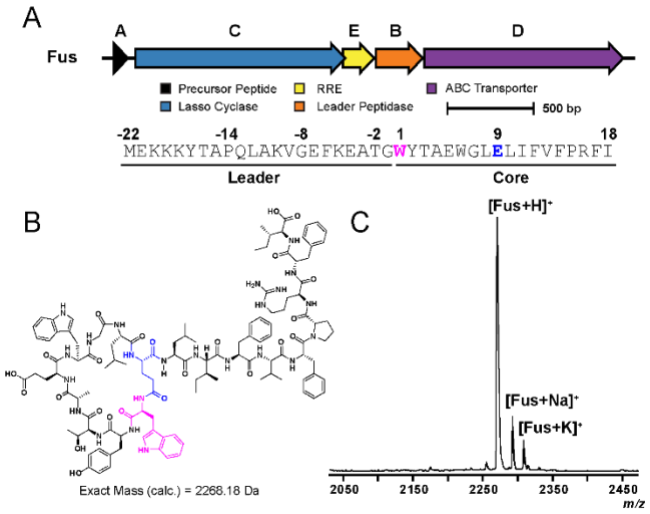


Figure 2.1. *E. coli*-based expression of fusilassin. (A) Biosynthetic gene cluster diagram and primary sequence of FusA. (B) Predicted structure of lasso peptide fusilassin shown in an unthreaded conformation for clarity. The residues involved in macrolactam formation (Trp1 and Glu9) are highlighted pink and blue, respectively. (C) MALDI-TOF-MS spectrum demonstrating the expression of fusilassin (Fus) in *E. coli*.

To probe the importance of highly conserved residues of the leader peptidase and lasso cyclase, we generated separate multiple sequence alignments for FusB and FusC using 11 homologous proteins from previously characterized lasso peptide BGCs. Six residues of FusB were targeted for replacement with Ala, including the predicted catalytic Cys79 and His112 residues, a nearby highly conserved Trp, and three Glu residues (Fig. S4). The catalytic activity of the FusB variants was then assessed using the purified enzymes. Ala-substitution of Cys79 and His112 abolished catalytic activity while the E116A variant was severely impaired (Fig. S5). FusB-W114A exhibited extensive proteolytic degradation upon expression in *E. coli*, leading us to conclude this residue is of structural significance (Fig. S2). Ala-replacement of FusB-E123 and -E124 did not significantly

perturb proteolytic efficiency under the reaction conditions employed. Four residues of FusC predicted to be involved in ATP-binding were targeted for replacement with Ala, based on the known structure of *E. coli* asparagine synthetase B (Protein DataBank entry 1CT9, Fig. S6).[21] The catalytic activity of FusC-D238A and -D340A was abolished, even when assaying at high concentration (5 mM) of ATP. The effects were less severe for FusC-S239A and G336A, which displayed catalytic impairment only when supplied with 500 μM ATP (Fig. S7). The concentration-dependent activity of these variants likely arises from an increased K_M for ATP, which corroborate the predicted ATP-binding residues of FusC.

To address enzymatic processing of FusA variants at core position 1, each variant was prepared as above and subjected to identical reconstitution reactions using wild-type FusB, FusC, and FusE. In good agreement with *E. coli* expression, all variants, except Pro1, were tolerated by the Fus biosynthetic enzymes although the ion intensity for the Glu1 variant was at the limit of detection (Figs. 2, S8). We next evaluated if the leader peptidase, the lasso cyclase, or both were intolerant of Pro1. While leader proteolysis was efficient on wild-type FusA treated with FusB and FusE (Fig. 2A), no reaction was observed with FusA-W1P as the substrate (Fig. S8). The intolerance of Pro at the P1' position of many proteases, including cysteine hydrolases, is well known.[22] To evaluate if FusC could tolerate Pro1, we first determined if the Fus biosynthetic proteins were capable of processing the linear core of wild-type FusA in a leader peptide-independent manner. While FusC was incapable of forming product on its own, the presence of FusB and FusE in the reaction led to a low but detectable level of product formation (Fig. S9) in the absence of leader peptide. Similar results were obtained when repeating the reaction with the linear core of FusA-W1P; thus, the leader peptidase, not the lasso cyclase, was intolerant of Pro1.

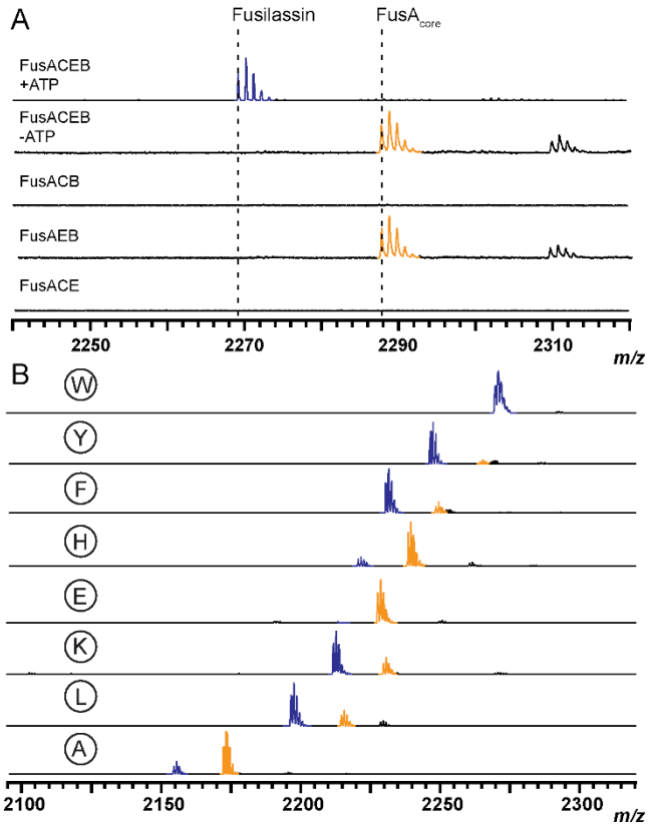


Figure 2.2. Enzymatic reconstitution of fusilassin. (A) MALDI-TOF-MS of all requisite biosynthetic components (top) shown alongside parallel reactions with individual components omitted from the reaction. (B) Relative processing of core position 1 variants from purified enzyme assays with wild-type Trp1 as the top spectrum. W1E product formation was at the limit of detection (Fig. S8). Spectra represent end-point assays after 3 h reactions . Macrolactam-containing products are blue and linear core (uncyclized) peptides are orange. Low-intensity peaks at +22 Da represent sodiated ions of either the linear core or mature lasso peptide.

In addition to resolving biosynthetic ambiguities like core position one tolerance, purified enzymatic reconstitution allows for deeper views into a biosynthetic pathway. In contrast with the *E. coli*-based expression system, where only mature metabolites are easily extracted and detected, reactions using purified FusB, FusC, and FusE provide straightforward assessments of conversion

efficiency for FusA variants. For example, larger residues at core position 1 (e.g. Tyr, Phe, Lys, and Leu) were most efficiently processed. Smaller and charged residues (e.g. His, Glu, and Ala) at position 1 gave comparatively less product by enzymatic reconstitution, which was not discernable by *E. coli*-expression (Fig. S1). As previously documented with plantazolicin and thiomuracin[17,23–25], which derive from two distinct RiPP classes, a greater level of biosynthetic granularity is achieved through enzymatic reconstitution.

2.2.3 Verification of the threaded lasso conformation.

It has remained unanswered if fusilassin prepared enzymatically *ex vivo* would be in a threaded conformation or in the isobaric "branched-cyclic" state. To address this question, we first performed a MALDI-TOF MS-based hydrogen-deuterium exchange assay on enzymatically prepared, wild-type fusilassin (Fig. S10). Under our H-D exchange conditions, the linear core of FusA was fully deuterated in <1 min while fusilassin prepared enzymatically required ~60 min to obtain the same level of deuteration. To provide additional support for a threaded conformation, we subjected the linear core of FusA and fusilassin to carboxypeptidase Y digestion for 18 h (Fig. S11). Analysis by MS showed the linear core of FusA was completely consumed while the vast majority of the fusilassin peptide remained intact. Indeed, only a single, low intensity peak consistent with proteolytic removal of the C-terminal residue, Ile18, was observed. Lastly, we tested thermal resistance to denaturation of wild-type fusilassin and the W1H and W1L variants. Even after a 2 h, 95 °C heat treatment, all variants exhibited similar resistance to carboxypeptidase Y degradation (Fig. S12). Unfortunately, the poor solubility characteristics and relatively low expression yields of fusilassin by both *E. coli* expression and enzymatic reconstitution prevented assessment of threadedness by less-sensitive methods, such as NMR.

2.2.4 Altered ring sizes, tail lengths, and biosynthetic insights.

Given the robustness and superior substrate tolerance of the Fus proteins, we sought to determine if FusA variants with expanded and contracted macrolactam rings could be produced. With the exception of the caulosegnin pathway[26], which naturally encodes lasso precursor peptides of differing ring sizes, all reported expansion or contraction of the macrolactam ring has abolished biosynthetic processing.[27,28] Fusilassin naturally contains a 9-mer macrolactam (Fig. 1A), and readily accepts a 10-mer macrolactam generated by insertion of Ala within the macrocycle ($_{1}$WYTA**A**EWGLE$_{10}$…FI$_{18}$). Not only was this 10-mer variant easily produced by both *E. coli* expression and enzymatic reconstitution (Fig. S13), but HRMS-MS/MS confirmed formation of an isopeptide bond at the predicted acceptor site (Fig. S14). Proteolytic treatment of this compound with carboxypeptidase Y supported threadedness (Fig. S15). Contracted fusilassin variants with theoretical 8-mer ($_{1}$WYTEWGLE$_{8}$…FI$_{17}$; DA4) and 7-mer ($_{1}$WYEWGLE$_{7}$…FI$_{16}$; DT3/A4) macrolactam ring sizes were then prepared and subsequently evaluated for production. While the 8-mer was readily detected from *E. coli* expression, we note that the same variant was considerably less efficiently produced using purified enzymes (Fig. S13). Continuing the trend, the 7-mer was not be detected by enzymatic reconstitution, and although a mass consistent with the expected product (*m/z* 2097 Da) was detected from *E. coli* expression, the intensity was too low to conclusively verify the identity of the product. We attribute the declining efficiency of production to an increased level of steric clashing between residues comprising the fusilassin ring and the C-terminal tail.

The apparent inability of the fusilassin biosynthetic enzymes to produce a reliable 7-mer macrolactam product, whether it be in a threaded (lasso) or unthreaded (branched-cyclic) conformation, suggested that FusC might be incapable of forming the latter. This topic has attracted

some debate in the lasso peptide field, especially since the initial report on lassomycin suggested the peptide might naturally possess a branched-cyclic topology.[5] A subsequent study from the same group corrected that finding by demonstrating fundamental differences between naturally produced lassomycin and a synthetically prepared branched-cyclic version.[29] To date, every reported lasso peptide produced by enzymatic reconstitution, heterologous expression, or native isolation has been in the threaded conformation. To further evaluate the (in)ability of FusC to generate a branched-cyclic peptide, we prepared C-terminal truncations of FusA. The first tested variant contained only the macrolactam ring ($_{1}$WYTAEWGLE$_{9}$) and failed to form detectable levels of product under *E. coli* expression, thus demonstrating the biosynthetic necessity of the loop and/or tail regions (Fig. S15). Truncation of as few as 3 residues from the C-terminus of FusA (DRFI$_{18}$) also failed to generate observable product through *E. coli* expression. This suggested that the loop region of the FusA core peptide was insufficient for FusC catalysis while the tail region was catalytically necessary. We dissected these findings further using our enzymatic reconstitution platform with Gly-replaced FusA variants at core positions 16-18. All tail variants yielded the expected macrolactam products, albeit formation of the R16G product was impaired relative to the other tested positions (Fig. S15). We postulate that the C-terminal tail of FusA, especially Arg16, is a critical contact point for the lasso cyclase to "pre-fold" the peptide substrate into a lasso-like conformation. Resolving the finer details of this interaction will require extensive additional characterization of the biosynthetic enzymes.

2.2.5 Acceptor specificity and non-cognate lasso peptide formation.

The above data show that the fusilassin biosynthetic enzymes are unusual in that they readily tolerate substitution of the first core position, as well as ring size variants (±1 residue). All previous studies have shown that lasso peptide biosynthetic pathways are intolerant to changes in the

macrolactam acceptor site (either Asp or Glu) and location (core positions 7, 8, or 9).[27,28] Therefore, we generated a FusA-E9D variant, and found that FusC was also incapable of producing the W1-D9 macrolactam by biosynthetic reconstitution (Fig. 3A).

We next applied our new understanding of fusilassin biosynthetic tolerance to the generation of a non-cognate lasso peptide. We generated a chimeric substrate by fusing the leader peptide of FusA (residues -22 thru -1, Fig. 1A) to the core peptide of citrulassin A (CitA), a lasso peptide we characterized previously[3] which has no obvious sequence relationship to fusilassin. Notably, CitA possesses a Leu_1-Asp_8 macrolactam and a core peptide that is 3 residues shorter than that of fusilassin. As expected, purified FusB, FusC, and FusD did not process the chimeric substrate, given the presence of an Asp acceptor (Fig. 3B). The Fus proteins also did not process a Cit-L1W variant, showing the overriding effect of acceptor versus donor site identity. Introduction of D8E to CitA led to a dramatic increase in processing by the Fus proteins. The formation and location of the macrolactam linkage was confirmed by HRMS-MS/MS (Fig. S14). Threadedness of this peptide was supported by recalcitrance to carboxypeptidase Y digestion although the resulting lasso peptide was not stable to prolonged heat treatment (Fig. S16). CitA variants with contracted (W1-E7) and expanded (W1-E9) macrolactam linkages were also readily processed by the Fus proteins (Fig. 3B). The predicted macrolactam for the CitA (W1-E7) macrolactam was demonstrated by HRMS-MS/MS (Fig. S14). Similar to the CitA-D8E variant, we found that the variant bearing an expanded W1-E9 macrolactam ring exhibited resistance to carboxypeptidase Y digestion, which is highly suggestive of a threaded conformation (Fig. S17). Unfortunately, we did not observe product formation from the chimeric substrates using the *E. coli* expression system. One possible reason could be degradation by endogenous *E. coli* proteases outcompeted the rate of lasso peptide formation (Fig. S18). Further characterization would be needed to identify the

origin of these observations. Nevertheless, our data underscore a remarkable substrate tolerance for the fusilassin pathway, especially with purified proteins.

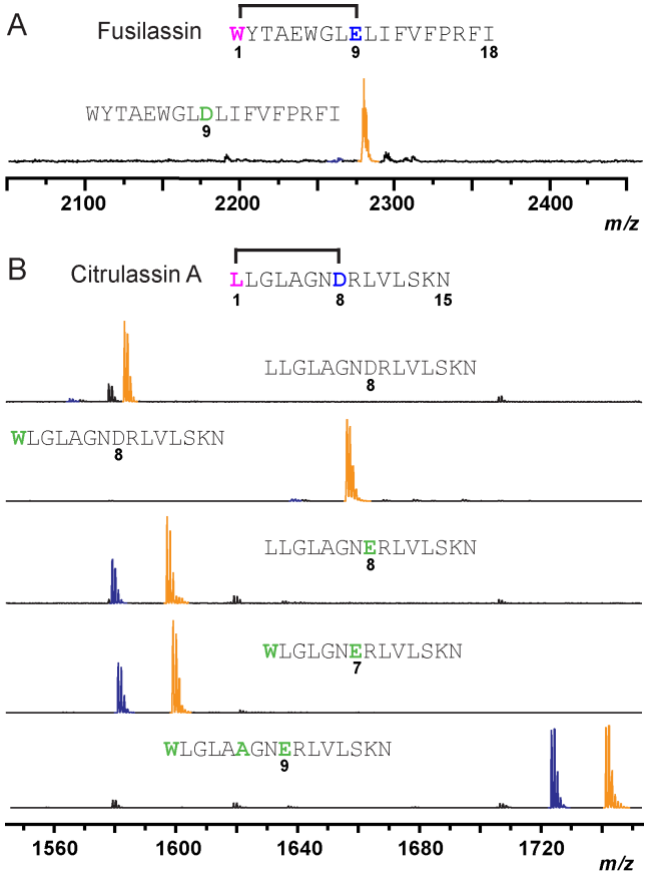


Figure 2.3. Acceptor site specificity. (A) MALDI-TOF-MS of the FusA-E9D variant from enzymatic reconstitution. (B) Evaluation of purified Fus enzymes in processing chimeric substrates bearing wild-type and macrolactam variants of citrulassin A (CitA). Macrolactam-containing products are blue while the linear core (uncyclized) peptides are orange. Substitutions to wild-type FusA in panel A, and CitA in panel B, respectively, are green.

2.2.6 Genomics-guided characterization of protein-protein interactions.

Given the unprecedented performance of FusB, FusC, and FusE *in vitro*, we sought to elucidate residues critical to formation of the fusilassin biosynthetic protein complex. Previously,

retrospective analysis of RODEO data sets uncovered a conserved YxxP motif in predicted lasso leader peptides.[3] Recent work on the lasso peptides paeninodin[18] and chaxapeptin[19] have confirmed that the YxxP motif functions as the recognition sequence for the RRE, as substitution with Ala caused severe reductions in binding affinity and product formation. To confirm their role in fusilassin biosynthesis, and to investigate other modestly conserved residues within the leader peptide, we prepared a brief panel of Ala-substituted FusA variants. We then employed a fluorescence polarization assay with wild-type FusE as the receptor and a fluorescein-labeled, wild-type $FusA_{leader}$ as a competing ligand. Of the five FusA positions surveyed, only Y(-17)A and P(-14)A were found to decrease affinity for FusE with the Tyr(-17) being critically important (Fig. S19, Table S3).

Since the above data only pertain to the substrate (FusA) side of the interaction, we next sought to probe the RRE-portion (FusE) of this interaction by targeted mutagenesis. To rationally direct our mutagenesis to likely sites of interaction, we employed GREMLIN[30], a computational algorithm that predicts interaction sites based on the principle of compensatory mutation. This strategy requires a large amount of diverse sequences, thus necessitating an updated survey of lasso peptide BGCs. We therefore queried RODEO to update the observable lasso peptide genomic space. Our last published survey of this kind was conducted in October 2016.[3] Since then, GenBank has grown by 56% and we have made algorithmic improvements that enhance the prediction of leader peptide cleavage sites.[15] As of August 2018, RODEO identified 2,993 non-redundant, high-scoring, manually curated lasso precursor peptides in publicly available bacteria and archaeal genomes

deposited in GenBank (Supplemental Dataset 1). Our updated RODEO data set provided ample sequence diversity for the application of this algorithm.

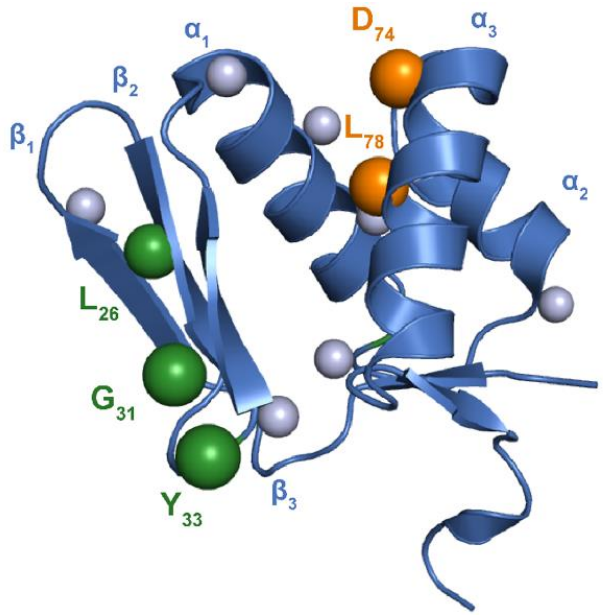


Figure 2.4. Residues critical for FusA/E/B ternary complex. The RRE domain of MccB (residues M_1-S_{86}, PDB entry 3H9J, blue) was aligned with FusE. Spheres indicate residues of FusE probed by mutagenesis. Mutagenesis at sites along α_3 drastically reduced binding to $FusA_{leader}$ (orange). Ala-substitutions along β_2 and β_3 did not influence $FusA_{leader}$ binding but did abolish leader peptide cleavage (green). The validity of these predicted protein-protein interaction sites is underscored by the fact mutagenesis elsewhere (gray) did not reduce leader peptide binding or leader peptidase activity.

Our analysis identified 7 residues of lasso peptide RRE proteins that are under co-evolutionary pressure with the lasso peptide precursors (Fig. S20). As we noted previously[3], all sites within the precursor peptide that were under statistically relevant co-evolution with the RRE were confined to the leader peptide. The receptor residues predicted to interact with the leader peptide were

overrepresented in the predicted b3 and a3 secondary structures of the RRE. Ala-substitution of four of these positions led to a loss of ~2–fold affinity towards $FusA_{leader}$. FusE-D74A exhibited a more dramatic reduction (8–fold) (Fig. S21, Table S4). Based on several available crystal structures of RRE proteins, we mapped Asp74 to the start of helix a3. Presumably, Leu78 would reside on the same face of this helix (Fig. 4). In support of this prediction, the FusE-L78A variant displayed a 10–fold reduction in $FusA_{leader}$ affinity.

Sequence analysis of FusE shows that the protein has an abnormally acidic isoelectric point (pI = 4.3). Given that Asp74 was a critical residue for engaging a relatively basic leader peptide (Fig. 1A), we hypothesized that additional acidic residues of FusE may contribute to this interaction. Single Ala-replacement of select residues also caused a ~2–fold loss in leader peptide affinity. It is possible that the difference in pI provides an electrostatic attraction between the RRE and the leader peptide over longer distances. However, once within proximity, we predict that specific hydrophobic interactions, governed by the YxxP motif, will dominate the orientation and specificity of the two interacting proteins. Investigations that further define this interaction are on-going.

2.2.7 Interaction between RRE and leader peptidase.

Sequence analysis of FusB, which collaborates with the RRE (FusE) during leader peptide cleavage, yielded a considerably basic isoelectric point (pI = 10.7). Thus, we propose interaction of these two proteins is also driven, at least initially, by electrostatic attraction. To determine if this phenomenon was more widespread, we calculated the pI's for all discretely encoded (i.e., non-fused) RRE proteins, leader peptidases, and lasso cyclases identified by RODEO (Fig. S22). While lasso cyclases exhibit neutral pI's (7.4 ± 1.1), RRE proteins are notably acidic (4.7 ± 1.2) and the leader peptidases are strongly basic (10.5 ± 1.1). Within a given lasso peptide BGC, the difference

in pI between the two interacting proteins was considerable (5.7 ± 1.8). This phenomenon is widespread across nearly 2,000 examined lasso peptide BGCs and thus certainly not specific to the fusilassin pathway.

By leveraging the 2018 RODEO-derived dataset, we had a sufficient number of lasso peptide biosynthetic protein sequences to probe co-evolutionary relationships between FusB and FusE in a statistically relevant fashion. GREMLIN analysis on the concatenated protein sequences identified 5-6 residues within FusB and E that were under predicted co-evolutionary pressure (Fig. S20). To evaluate the relative contribution of these residues, we conducted an MS-based leader peptidase cleavage assay that requires both FusB and FusE (Fig. 2A). Under the conditions employed, wild-type FusE orchestrated the efficient delivery of FusA to FusB (Fig. 5). The most statistically probable interaction identified by GREMLIN with FusB involved FusE-Tyr33. Substitution with Ala led to no observed leader peptide cleavage. It is notable that FusE-Y33A has affinity for FusA equal to wild-type FusE (Table S3), which is supported by structure homology arguments as this site does not reside on either the b3 or a3 secondary structure of FusE (Fig. 4).

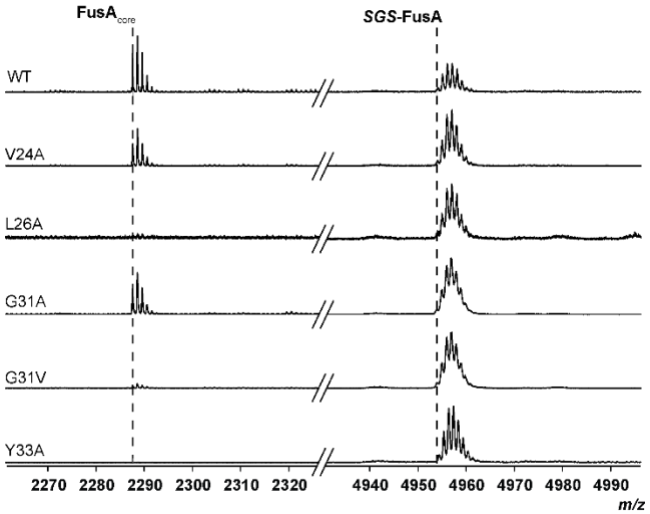


Figure 2.5. RRE:Leader peptidase interaction. MALDI-TOF-MS was used to monitor leader peptidase activity of FusB and a panel of GREMLIN- and structure-derived FusE variants. Variants demonstrated differing degrees of enzymatic processing, supporting our GREMLIN predictions. SGS refers to the N-terminal Ser-Gly-Ser motif remaining after TEV protease cleavage of MBP-FusA.

The second most strongly FusB-correlated residue of FusE was Gly31, which did not influence leader peptide cleavage upon Ala-substitution. However, replacement of Gly31 with Val nearly abolished leader peptidase activity. This variant only exhibited a minor reduction in leader peptide-binding (2–fold); therefore, we predict it also plays a direct role in FusB-binding (Table S3; Fig. 5). In proximity to the $_{31}$GxY$_{33}$ motif was FusE-Leu26, which we had previously substituted with Ala to evaluate FusA$_{leader}$-binding (Fig. S21). This variant abolished leader peptidase activity, but the same was not true for the nearby V24A variant, which would theoretically occupy the same face of the b2 strand (Fig. 4).

2.2.8 Bioinformatic insight into lasso cyclases.

The bioinformatically mined lasso cyclases are visualized in the form of maximum likelihood trees and a sequence similarity network (Figs. S23-24). As noted previously, many lasso peptides are found within the actinobacteria, proteobacteria, and firmicutes phyla with minimal evidence of horizontal gene transfer. Other bacterial and archaeal phyla are represented at lower frequencies. Upon viewing the tree color-coded by the identity of the macrolactam acceptor residue (Asp or Glu), we first noted that many clades display a dominant acceptor identity while a few contain a more balanced representation of the two types. Further, Asp- and Glu-specific lasso cyclases are interspersed throughout the tree, meaning that there is no global sequence distinction between enzymes that prefer one acceptor type over the other.

Given the macrolactam size-flexibility of the fusilassin pathway, we conducted a retrospective analysis of the 2018 lasso peptide dataset to determine how many naturally encoded lasso peptides would be 7, 8, or 9-mers. Upon removing ambiguous linkages from the data set, the proportions were found to be 5.5%, 36.6%, and 57.9%, respectively (Fig. S25). Given the acceptor-intolerance of the fusilassin pathway, we also analyzed the same data and found that 29.5% of RODEO-identified, unambiguous lasso core peptides would use Glu as the macrolactam acceptor site and the remaining 69.5% would contain Asp. The plurality (45%) of lasso peptides are predicted to be Asp-containing 9-mers.

2.3 Conclusions

The data presented herein describe the discovery, reconstitution, and initial enzymatic and biophysical characterization of fusilassin, an unusual lasso peptide from a mesothermophilic actinomycete. Our results confirm that the fusilassin pathway can serve as the first lasso peptide biosynthetic platform amenable to detailed mechanistic characterization. The fusilassin enzymes

exhibit an unprecedented substrate promiscuity with respect to the identity of the first core position, the size of the macrolactam linkage, and core peptide sequence. We have successfully dissected the role of leader peptide from core processing using purified enzymes and synthetic core peptides, which have delineated the enzymatic preferences of the leader peptidase and lasso cyclase proteins. Our data support an inability of the lasso cyclase to generate the isobaric branched-cyclic peptides, as illustrated by employing a series of truncated precursor peptides, proteolytic degradation studies, and hydrogen-deuterium exchange assays. Moreover, the enzymes described have the exceptional capability to produce non-cognate lasso peptides or simple derivatives thereof, presenting unprecedented access to the ~3,000 RODEO-identified lasso peptides found in prokaryotic genomes. Finally, bioinformatic analysis of this dataset enabled the co-evolutionary and biochemical interrogation of additional sites of interaction between the leader peptide, RRE domain, and leader peptidase. The fusilassin BGC represents the most promising pathway yet described that can be utilized to elucidate molecular-level details of lasso peptide biosynthesis.

2.4 References

(1) Arnison, P.G.; Bibb, M.J.; Bierbaum, G.; Bowers, A.A.; Bugni, T.S.; Bulaj, G.; Camarero, J.A.; Campopiano, D.J.; Challis, G.L.; Clardy, J.; Cotter, P.D.; Craik, D.J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P.C.; Entian, K.D.; Fischbach, M.A.; Garavelli, J.S.; Göransson, U.; Gruber, C.W.; Haft, D.H.; Hemscheidt, T.K.; Hertweck, C.; Hill, C.; Horswill, A.R.; Jaspars, M.; Kelly, W.L.; Klinman, J.P.; Kuipers, O.P.; Link, A.J.; Liu, W.; Marahiel, M.A.; Mitchell, D.A.; Moll, G.N.; Moore, B.S.; Müller, R.; Nair, S.K.; Nes, I.F.; Norris, G.E.; Olivera, B.M.; Onaka, H.; Patchett, M.L.; Piel, J.; Reaney, M.J.; Rebuffat, S.; Ross, R.P.; Sahl, H.G.; Schmidt, E.W.; Selsted, M.E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.;

Süssmuth, R.D.; Tagg, J.R.; Tang, G.L.; Truman, A.W.; Vederas, J.C.; Walsh, C.T.; Walton, J.D.; Wenzel, S.C.; Willey, J.M.; van der Donk, W.A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2012**, *30* (1), 108–160.

(2) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A Prevalent Peptide-Binding Domain Guides Ribosomal Natural Product Biosynthesis. *Nat. Chem. Biol.* **2015**, *11* (8), 564–570.

(3) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A New Genome-Mining Tool Redefines the Lasso Peptide Biosynthetic Landscape. *Nat. Chem. Biol.* **2017**, *13* (5), 470–478.

(4) Ortega, M. A.; van der Donk, W. A. New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products. *Cell Chem. Biol.* **2016**, *23* (1), 31–44.

(5) Gavrish, E.; Sit, C. S.; Cao, S.; Kandror, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; et al. Lassomycin, a Ribosomally Synthesized Cyclic Peptide, Kills Mycobacterium Tuberculosis by Targeting the ATP-Dependent Protease ClpC1P1P2. *Chem. Biol.* **2014**, *21* (4), 509–518.

(6) Feng, Z.; Ogasawara, Y.; Nomura, S.; Dairi, T. Biosynthetic Gene Cluster of a D-Tryptophan-Containing Lasso Peptide, MS-271. *ChemBioChem.*, *19* (19), 2045-2048.

(7) Allen, C. D.; Chen, M. Y.; Trick, A. Y.; Le, D. T.; Ferguson, A. L.; Link, A. J. Thermal Unthreading of the Lasso Peptides Astexin-2 and Astexin-3. *ACS Chem. Biol.* **2016**, *11* (11), 3043-3051.

(8) Hegemann, J. D.; Fage, C. D.; Zhu, S.; Harms, K.; Leva, F. S. D.; Novellino, E.; Marinelli, L.; Marahiel, M. A. The Ring Residue Proline 8 Is Crucial for the Thermal Stability of the Lasso Peptide Caulosegnin II. *Mol. BioSyst.* **2016**, *12* (4), 1106–1109.

(9) Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Lasso Peptides: An Intriguing Class of Bacterial Natural Products. *Acc. Chem. Res.* **2015**, *48* (7), 1909–1919.

(10) Hegemann, J. D.; De Simone, M.; Zimmermann, M.; Knappe, T. A.; Xie, X.; Di Leva, F. S.; Marinelli, L.; Novellino, E.; Zahler, S.; Kessler, H.; et al. Rational Improvement of the Affinity and Selectivity of Integrin Binding of Grafted Lasso Peptides. *J. Med. Chem.* **2014**, *57* (13), 5829–5834.

(11) Knappe Thomas A.; Manzenrieder Florian; Mas-Moruno Carlos; Linne Uwe; Sasse Florenz; Kessler Horst; Xie Xiulan; Marahiel Mohamed A. Introducing Lasso Peptides as Molecular Scaffolds for Drug Design: Engineering of an Integrin Antagonist. *Angew. Chem. Int. Ed.* **2011**, *50* (37), 8714–8717.

(12) Naimi, S.; Zirah, S.; Hammami, R.; Fernandez, B.; Rebuffat, S.; Fliss, I. Fate and Biological Activity of the Antimicrobial Lasso Peptide Microcin J25 Under Gastrointestinal Tract Conditions. *Front. Microbiol.* **2018**, *9*, 1764-1777.

(13) Yamamoto, T.; Matsui, H.; Yamaji, K.; Takahashi, T.; Øverby, A.; Nakamura, M.; Matsumoto, A.; Nonaka, K.; Sunazuka, T.; Ōmura, S.; et al. Narrow-Spectrum Inhibitors Targeting an Alternative Menaquinone Biosynthetic Pathway of Helicobacter Pylori. *J. Inf. Chemother.* **2016**, *22* (9), 587–592.

(14) Lopez, F. E.; Vincent, P. A.; Zenoff, A. M.; Salomón, R. A.; Farías, R. N. Efficacy of Microcin J25 in Biomatrices and in a Mouse Model of Salmonella Infection. *J Antimicrob. Chemother.* **2007**, *59* (4), 676–680.

(15) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. Bioinformatic Expansion and Discovery of Thiopeptide Antibiotics. *J. Am. Chem. Soc.* **2018**, *140* (30), 9494–9501.

(16) Mahanta, N.; Liu, A.; Dong, S.; Nair, S. K.; Mitchell, D. A. Enzymatic Reconstitution of Ribosomal Peptide Backbone Thioamidation. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (12), 3030-3035.

(17) Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A. In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *J. Am. Chem. Soc.* **2015**, *137* (51), 16012–16015.

(18) D. Hegemann, J.; J. Schwalen, C.; A. Mitchell, D.; Donk, W. A. van der. Elucidation of the Roles of Conserved Residues in the Biosynthesis of the Lasso Peptide Paeninodin. *Chem. Commun.* **2018**, *54* (65), 9007–9010.

(19) Martin-Gómez, H.; Linne, U.; Albericio, F.; Tulla-Puche, J.; Hegemann, J. D. Investigation of the Biosynthesis of the Lasso Peptide Chaxapeptin Using an E. Coli-Based Production System. *J. Nat. Prod.* **2018**, *81* (9), 2050-2056.

(20) Pavlova, O.; Mukhopadhyay, J.; Sineva, E.; Ebright, R. H.; Severinov, K. Systematic Structure-Activity Analysis of Microcin J25. *J. Biol. Chem.* **2008**, *283* (37), 25589–25595.

(21) Larsen, T. M.; Boehlein, S. K.; Schuster, S. M.; Richards, N. G. J.; Thoden, J. B.; Holden, H. M.; Rayment, I. Three-Dimensional Structure of Escherichia Coli Asparagine Synthetase B: A Short Journey from Substrate to Product. *Biochemistry* **1999**, *38* (49), 16146–16157.

(22) Evnin, L. B.; Vásquez, J. R.; Craik, C. S. Substrate Specificity of Trypsin Investigated by Using a Genetic Selection. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87* (17), 6659–6663.

(23) Deane, C. D.; Burkhart, B. J.; Blair, P. M.; Tietz, J. I.; Lin, A.; Mitchell, D. A. In Vitro Biosynthesis and Substrate Tolerance of the Plantazolicin Family of Natural Products. *ACS Chem. Biol.* **2016**, *11* (8), 2232–2243.

(24) Deane, C. D.; Melby, J. O.; Molohon, K. J.; Susarrey, A. R.; Mitchell, D. A. Engineering Unnatural Variants of Plantazolicin through Codon Reprogramming. *ACS Chem. Biol.* **2013**, *8* (9), 1998–2008.

(25) Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. Biosynthetic Timing and Substrate Specificity for the Thiopeptide Thiomuracin. *J. Am. Chem. Soc.* **2016**, *138* (48), 15511–15514.

(26) Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Caulosegnins I–III: A Highly Diverse Group of Lasso Peptides Derived from a Single Biosynthetic Gene Cluster. *J. Am. Chem. Soc.* **2013**, *135* (1), 210–222.

(27) Knappe, T. A.; Linne, U.; Robbel, L.; Marahiel, M. A. Insights into the Biosynthesis and Stability of the Lasso Peptide Capistruin. *Chem. Biol.* **2009**, *16* (12), 1290–1298.

(28) Ducasse, R.; Yan, K.-P.; Goulard, C.; Blond, A.; Li, Y.; Lescop, E.; Guittet, E.; Rebuffat, S.; Zirah, S. Sequence Determinants Governing the Topology and Biological Activity of a Lasso Peptide, Microcin J25. *ChemBioChem* **2012**, *13* (3), 371–380.

(29) Lear, S.; Munshi, T.; S. Hudson, A.; Hatton, C.; Clardy, J.; A. Mosely, J.; J. Bull, T.; S. Sit, C.; L. Cobb, S. Total Chemical Synthesis of Lassomycin and Lassomycin-Amide. *Org. Biomol. Chem.* **2016**, *14* (19), 4534–4541.

(30) Ovchinnikov, S.; Kamisetty, H.; Baker, D. Robust and accurate prediction of residue–residue interactions across protein interfaces using evolutionary information. *eLife*. **2014**, *3*, e02030.

Chapter 3: Protein-Protein Interactions of the RiPP Recognition Element and Leader Peptidase Involved in Fusilassin Biosynthesis

3.1 Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a natural product class comprised of a myriad of structural subclasses possessing a broad range of bioactivities.[1] RiPP biosynthesis begins with the translation of a precursor peptide which contains bipartite structure with an N-terminal "leader peptide" and a C-terminal "core peptide." The leader peptide contains a recognition motif which is bound by the RiPP recognition element (RRE) which subsequently recruits the cognate biosynthetic enzymes to modify the core peptide.[2] During lasso peptide biosynthesis, the leader peptidase cleaves the leader peptide from the core peptide in an RRE-dependent fashion. Subsequently, a lasso cyclase enzyme folds the core peptide into a pre-lasso conformation before generating an isopeptide bond between the free N-terminus and the carboxylate sidechain of an internal Asp or Glu residue. This conformation traps the C-terminal tail of the core peptide, which is held in place by steric interactions. Additional modifications have been reported, including C-terminal methyl ester formation,[3] arginine deimination,[4] or epimerization.[5] Steric constraints between the lasso peptide tail and macrolactam ring endow lasso peptides with resistance to heat denaturation and proteolytic degradation.

Of the dozens of discovered lasso peptides, less than half possess reported biological activity. Known bioactivities include cell-surface receptor antagonism, HIV fusion inhibition, RNA polymerase inhibition, and mycobacterial Clp protease inhibition.[6] Because RiPP modifications occur in an RRE-dependent fashion, they have great potential for the generation of analog libraries. Previously characterized lasso peptides have demonstrated tolerance to core peptide variation, making them an attractive target for high throughput screening efforts.[7–9] In combination with the

diverse bioactivities of lasso peptides, this highlights the high potential for the discovery of bioactive non-native lasso peptides.

Efforts to identify and characterize novel lasso peptides are ongoing. Complemented with recent advances in bioinformatics, the discovery and characterization of novel lasso peptides has been greatly accelerated.[10] Be leveraging our custom bioinformatic algorithm RODEO (Rapid ORF Description & Evaluation Online: www.ripprodeo.org), we previously identified and reconstituted the lasso peptide BGC from *Thermobifida fusca* and named the resultant lasso peptide fusilassin

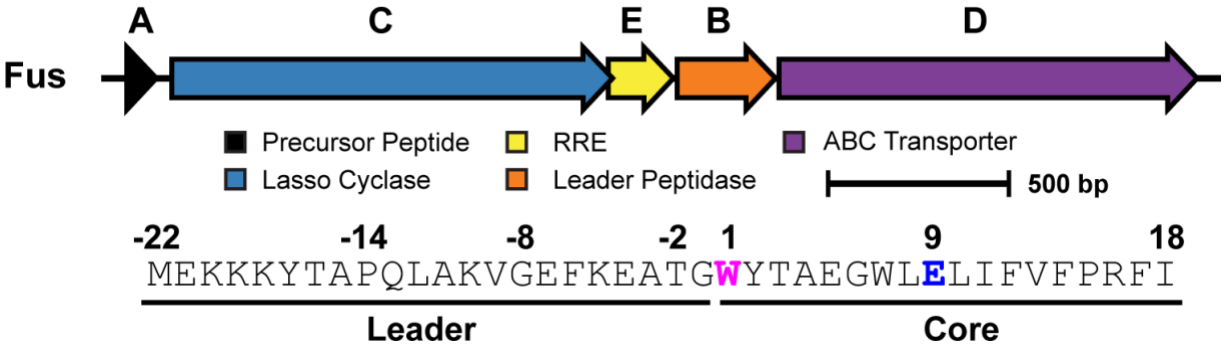


Figure 3.1. Fusilassin biosynthetic gene cluster. Individual proteins from this pathway were successfully expressed as MBP fusion proteins. The first residue of the core peptide is indicated in purple, while the macrolactam acceptor residue is indicated in blue.

(Fig. 3.1).[9] In contrast with previously isolated lasso peptide BGCs, the constituent biosynthetic enzymes from this pathway exhibited exceptional stability upon heterologous expression in *Escherichia coli*. Furthermore, this pathway demonstrated extraordinary substrate tolerance, and was even capable of generating an analog of the lasso peptide citrulassin A from an unrelated lasso BGC.[9]

The molecular level details governing the protein-protein interactions underpinning lasso peptide biosynthesis remain almost exclusively unknown. In our previous report, GREMLIN-based

coevolutionary analyses were used to identify several residues involved in RRE-Leader peptidase interactions.[9] Given the *in vitro* stability of the fusilassin proteins, we chose this cluster to begin investigations of these interactions with additional analytical methods.

Solution Nuclear Magnetic Resonance spectroscopy (NMR) has long been an effective tool in elucidating the molecular-level structures and interactions of proteins. In the context of lasso peptides generally, NMR has been used to generate 3D structures of a number of lasso peptides,[6] but has never been applied to study the biosynthetic enzyme interactions responsible for the generation of these compounds. In this report, we elected to study the solution phase RRE-Leader peptidase interactions by biomolecular NMR.

3.2 Results and Discussion

3.2.1 FusE NMR Assignment

The fusilassin BGC is composed of 3 proteins: The RRE domain FusE (~10 kDa), the leader peptidase FusB (~16 kDa) and the lasso cyclase FusC (~66 kDa) (Fig. 1). With respect to NMR, high protein molecular weight/residue count can severely complicate NMR analyses, so we elected to study the RRE domain. Additionally, FusB and FusC exhibit solution phase instability, and precipitate at concentrations required for NMR studies. FusE was expressed in M9 minimal media containing $^{15}NH_4Cl$ and was analyzed by two-dimensional ^{15}N-^{1}H NMR HMQC. In phosphate buffer at pH 6.5, the two-dimensional spectrum was of high quality, with broad chemical shift dispersion and narrow line shape, indicative of a properly folded protein. To assign the backbone amide resonances, uniformly labelled $^{13}C/^{15}N$ FusE was prepared and analyzed by triple resonance NMR (HNCO, HNCA, HN(CO)CA, CBCA(CO)NH, HNCACB). These experiments allowed complete assignment of the FusE backbone (Fig. 3.2A). Attempts to fully assign residue sidechains

and generate an NOE-refined 3D solution structure were unsuccessful due to broadened side chain resonances and insufficient long-range NOEs.

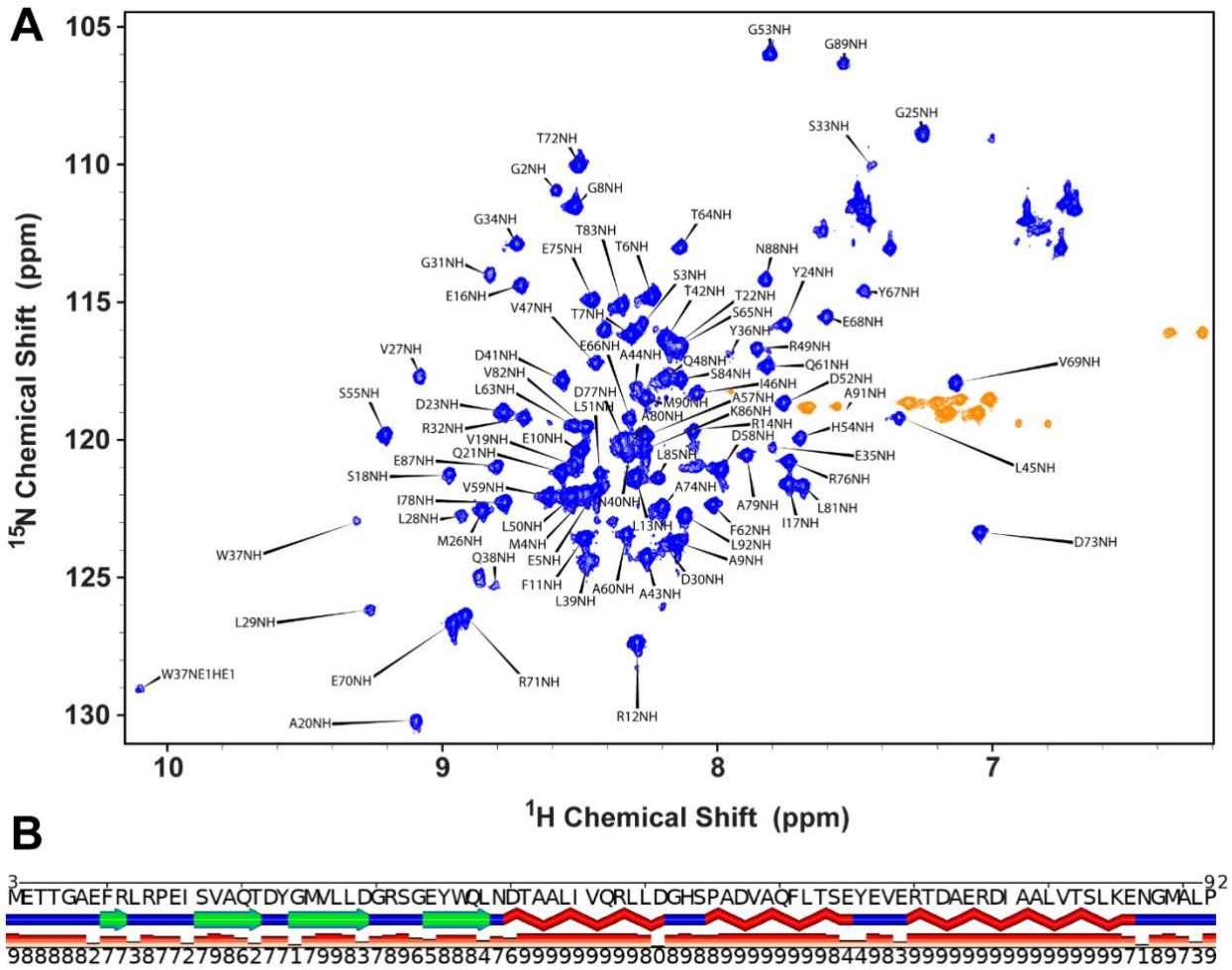


Figure 3.2. Backbone assignment and secondary structure of FusE. All residues were successfully assigned based on 3-dimensional triple-resonance NMR experiments. Aliased arginine sidechains are colored in orange (A). Asparagine and glutamine sidechains remain unassigned but were observed at ~7ppm (^{1}H) and ~112.5ppm (^{15}N). The wide chemical shift dispersion Is indicative of a stable, well-folded protein. Alanine 91 was observable at low threshold but was partially occluded by its proximity to an inverse phased Arg sidechain resonance. Predicted secondary structure of FusE based on TALOS-N are shown in B. Prediction confidence is indicated on a 1 (least confident) to 9 (most confident).

During our studies, an X-Ray structure of FusE bound to the FusA leader peptide ($FusA_{LP}$) was reported (PDB: 6JX3), allowing us a comparative reference four our assignments.[11] The TALOS-N-based secondary structure prediction derived from α-C/H, β- C/H, backbone N/H and carbonyl C chemical shifts generated a secondary structure in near perfect agreement with the published structure (Fig. 3.2B).[12] Interestingly, the backbone resonances along the β_3 sheet exhibited weaker, broadened signals in the analysis of apo-FusE. Given that this β-sheet is directly involved in leader peptide binding, we hypothesized that this region may be undergoing conformational exchange. Analysis of the longitudinal (R_1), transverse (R_2) and heteronuclear NOE relaxation rates demonstrated that this sheet has a slightly increased R_2 (Fig. 3.3). This relaxation rate is consistent with fast conformational exchange, which could indicate flexibility within this region is implicated in leader peptide recognition. [13,14]

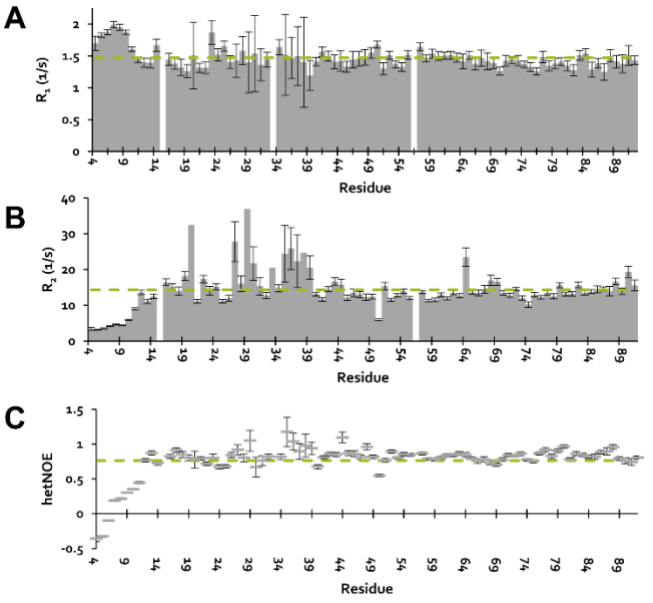


Figure 3.3. Backbone relaxation of FusE. TfusE exhibited a largely consistent T_1 relaxation of ~1.5 s, indicative of a single globular domain (A). Residues located within the beta-sheet region of FusE exhibited faster T_2 (B) and heteronuclear NOE relaxation (C), indicating shorter time scale relaxation. Average relaxation times are indicated by dashed lines. Large errors are reported for beta-sheet residues likely due to relatively weak peak intensities.

3.2.2 FusE-FusA$_{LP}$ Chemical Shift Perturbation

To characterize the interaction between FusE and FusA$_{LP}$, we utilized chemical shift perturbations (CSP). In systems exhibiting fast exchange, increasing molar equivalents of a ligand will induce gradual shifts in the chemical shift values of the binding partner, whereby the greatest changes in chemical shift are indicative of strongly interacting residues.[15] In preliminary CSP experiments, the FusE-FusA$_{LP}$ system exhibited slow exchange behavior, where peaks corresponding to the unbound protein gradually decrease in intensity and their counterparts in the ligand-bound protein increase in intensity without undergoing a sequential change in chemical shift value. These results are consistent with reported fluorescence polarization data demonstrating FusE and FusA$_{LP}$ possess a K_d of 9 nM.[9] Attempts to reassign the FusA$_{LP}$-bound FusE backbone by triple resonance NMR were unsuccessful due to line-broadening of resonances of β_1-β_3, preventing analysis by CSP. A previously characterized variant of FusA$_{LP}$ possessing a Y(-17)A mutation exhibited a K_d of >25 μM, an affinity which is more amenable to CSP analysis.[9] Experiments utilizing ^{15}N-FusE and natural abundance (NA) TfusA$_{LP}$(Y-17)A generated results consistent with the FusE-FusA$_{LP}$ structure, as the residues along β_3 exhibited the highest chemical shift change (Fig. 5A). These results are also consistent with mutagenesis studies, which demonstrated mutations along this β-sheet significantly reduced FusA$_{LP}$ binding.[9] In addition to β_3, several residues between the alpha helixes α_2 and α_3 also exhibited large chemical shift changes, indicating FusA$_{LP}$ binds along β_3 in combination with the pocket formed along these α-helices (Fig. 3.4B).

3.2.3 FusE-FusB Protein-Protein Interactions

With the assigned ^{15}N FusE-NA FusA$_{LP}$Y-17A, we next attempted to perform CSP with an excess of NA MBP-FusB. Unfortunately, the poor solution-phase stability led to significant precipitation, even at a conservative ratio of 1:1 MBP-FusB:^{15}N FusE (data not shown). This precipitation led

to complete dissipation of nearly all ^{15}N-^{1}H peaks. Given the requirement of a stoichiometric excess of binding partner for successful CSP analyses, we sought additional methods to determine the protein-protein interactions at sub-stoichiometric equivalents of FusB.

We therefore devised a strategy utilizing paramagnetic relaxation enhancement (PRE). The PRE effect is enabled by utilizing a probe containing unpaired electrons, which enhances the relaxation rates of proximate NMR-active nuclei through magnetic dipolar coupling.[16,17] This strategy has been used with radical[18] and lanthanide probes covalently attached to proteins[19] to generate PRE-restraints used in structure calculations. In these studies, nuclei closest to the PRE probe exhibit the greatest relaxation enhancement, while nuclei farther away do not.

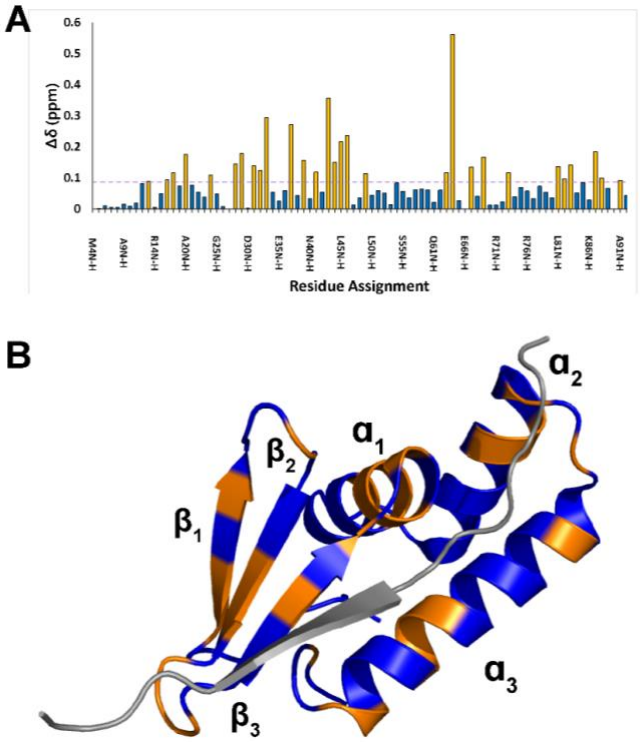


Figure 3.4. Chemical shift perturbation of FusE and $FusA_{LP}$. Calculated chemical shift perturbations of FusE and $FusA_{LP}$(Y-17A) are shown in (A). Chemical shifts greater than one standard deviation of all chemical shift perturbations (dashed, purple line) are indicated in orange, and presumed to be directly interacting with the leader peptide. Proline residues are not displayed. Interacting residues are highlighted in orange on the crystal structure of FusE (PDB: 6JX3, B). Alpha- and beta-sheet assignments are labelled and $FusA_{LP}$ is shown in gray.

To avoid the difficulties associated with generating covalently derivatized proteins, we elected to approach this system utilizing solvent PREs.[20] In this strategy, a lanthanide chelate is dissolved in the sample containing the protein of interest, thus generating a paramagnetic bulk solution. Relaxation enhancements are greatest on the solvent-exposed surfaces of the protein. This methodology can be used to determine the solvent-exposed surface of a protein of interest, with or without residue assignments. Solvent PREs can also illuminate sites of protein-protein interaction, as there would be a dulling effect of the PRE due to occlusion of a surface exposed residue by a protein binding partner.

We selected the PRE probe gadodiamide, known as the commercial MRI contrast agent Omniscan, for two main reasons. First, this probe has been shown to possess very low amounts of non-specific protein association, which can lead to artifactual PRE measurements.[21] Second, the Gd^{3+} nucleus in gadodiamide possesses seven unpaired electrons and exhibits the largest PRE effect radius, making it a more sensitive solvent PRE probe.[22]

To determine the interaction sites of FusE and FusB, ^{15}N-FusE was reconstituted with three equivalents of $FusA_{LP}$Y-17A and a final stoichiometry of 1:8 NA-MBP-FusB to ^{15}N-FusE. Peak intensities were measured with and without 2mM gadodiamide. This methodology serves a s a more simple method to determine PRE effects, compared with traditional methods of using full relaxation analyses.[23] Residues exhibiting the smallest decrease in peak intensity were assigned as FusB interaction sites (Fig. 3.5). Small perturbations in peak height were observed for residues along α_1 and α_2 which occupy largely solvent-inaccessible conformations, as well as residues along

α_3 which are proximal to bound $FusA_{LP}$. Most importantly, a number of residues located along β_1- β_3 exhibited small perturbations. Given that these residues are largely solvent exposed, they were proposed to be shielded from gadodiamide upon FusB binding. These results are consistent with previous studies, which identified several residues within the β-sheet region of FusE.

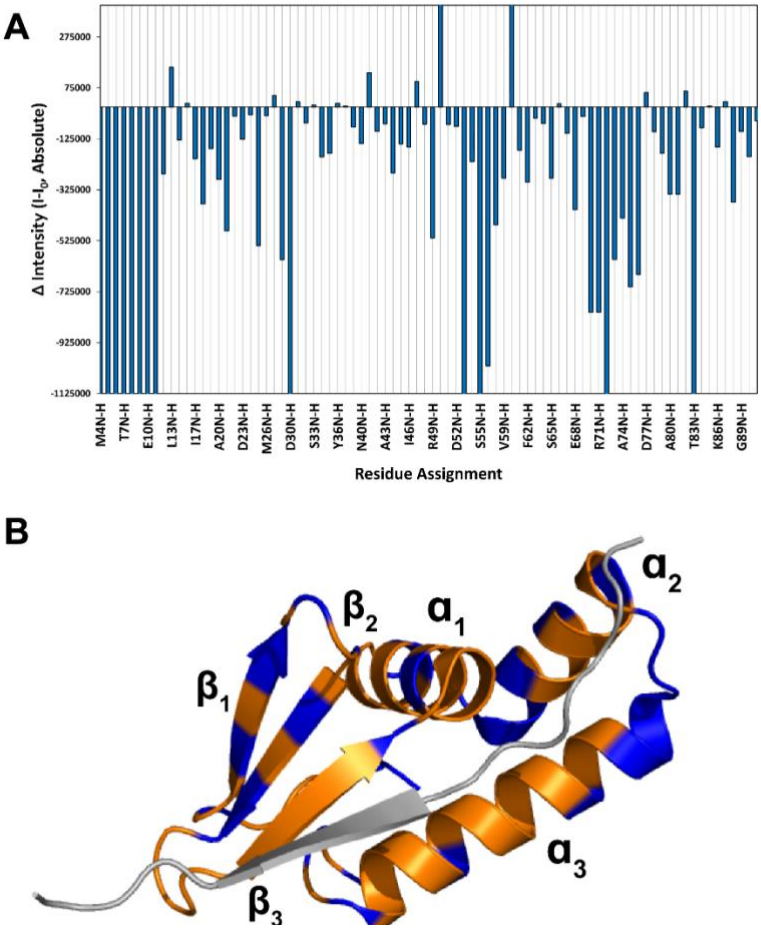


Figure 3.5. Paramagnetic relaxation enhancement of FusE complexed with $FusA_{LP}$Y-17A and FusB. Intensity-based PRE effects of FusE and FusB are shown in (A). Large decreases in peak intensity are indicative of solvent exposed residues, while smaller decreases indicate solvent shielding by interacting partners. Shielded residues were mapped to the X-Ray crystal structure are highlighted in orange on the crystal structure of FusE (PDB: 6JX3, B). Alpha- and beta-sheet assignments are labelled and $FusA_{LP}$ is shown in gray.

3.3 Conclusion

In summary we introduce the full backbone assignment of FusE. These assignments were used to determine the interactions occurring between FusE and the leader peptide as well as the interactions of FusE and

FusB. In addition to CSP analyses, we propose the PRE as a method for further discovery of protein-protein interactions governing RiPP biosynthesis. This method is particularly valuable in the cases where biosynthetic enzymes exhibit poor solution stability. The PRE method allows the use of sub-stoichiometric amounts of binding partners, as the NMR intensity measurements offer finer granularity relative to CSP.

3.4 References

(1) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; Heel, A. J. van; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Burgos, M. J. G.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; Wezel, G. P. van; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; Donk, W. A. van der. New Developments in RiPP Discovery, Enzymology and Engineering. *Nat. Prod. Rep.* **2021**, *38* (1), 130–239. https://doi.org/10.1039/D0NP00027B.

(2) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A Prevalent Peptide-Binding Domain Guides Ribosomal Natural Product Biosynthesis. *Nature Chemical Biology* **2015**, *11* (8), 564–570. https://doi.org/10.1038/nchembio.1856.

(3) Gavrish, E.; Sit, C. S.; Cao, S.; Kandror, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K. Lassomycin, a Ribosomally Synthesized Cyclic Peptide, Kills Mycobacterium Tuberculosis by Targeting the ATP-Dependent Protease ClpC1P1P2. *Chemistry & Biology* **2014**, *21* (4), 509–518. https://doi.org/10.1016/j.chembiol.2014.01.014.

(4) Harris, L. A.; Saint-Vincent, P. M. B.; Guo, X.; Hudson, G. A.; Mitchell, D. A. Reactivity-Based Screening for Citrulline-Containing Natural Products Reveals a Family of Bacterial Peptidyl

Arginine Deiminases. *bioRxiv* **2020**, 2020.07.16.207027. https://doi.org/10.1101/2020.07.16.207027.

(5) Feng, Z.; Ogasawara, Y.; Nomura, S.; Dairi, T. Biosynthetic Gene Cluster of a D-Tryptophan-Containing Lasso Peptide, MS-271. *ChemBioChem* **2018**, *19* (19), 2045–2048. https://doi.org/10.1002/cbic.201800315.

(6) Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Lasso Peptides: An Intriguing Class of Bacterial Natural Products. *Acc. Chem. Res.* **2015**, *48* (7), 1909–1919. https://doi.org/10.1021/acs.accounts.5b00156.

(7) Hegemann, J. D.; De Simone, M.; Zimmermann, M.; Knappe, T. A.; Xie, X.; Di Leva, F. S.; Marinelli, L.; Novellino, E.; Zahler, S.; Kessler, H.; Marahiel, M. A. Rational Improvement of the Affinity and Selectivity of Integrin Binding of Grafted Lasso Peptides. *J. Med. Chem.* **2014**, *57* (13), 5829–5834. https://doi.org/10.1021/jm5004478.

(8) Knappe, T. A.; Manzenrieder, F.; Mas-Moruno, C.; Linne, U.; Sasse, F.; Kessler, H.; Xie, X.; Marahiel, M. A. Introducing Lasso Peptides as Molecular Scaffolds for Drug Design: Engineering of an Integrin Antagonist. *Angewandte Chemie International Edition* **2011**, *50* (37), 8714–8717. https://doi.org/10.1002/anie.201102190.

(9) DiCaprio, A. J.; Firouzbakht, A.; Hudson, G. A.; Mitchell, D. A. Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin. *J. Am. Chem. Soc.* **2019**, *141* (1), 290–297. https://doi.org/10.1021/jacs.8b09928.

(10) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A New Genome-Mining Tool Redefines the Lasso Peptide Biosynthetic Landscape. *Nature Chemical Biology* **2017**, *13* (5), 470–478. https://doi.org/10.1038/nchembio.2319.

(11) Sumida, T.; Dubiley, S.; Wilcox, B.; Severinov, K.; Tagami, S. Structural Basis of Leader Peptide Recognition in Lasso Peptide Biosynthesis Pathway. *ACS Chem. Biol.* **2019**, *14* (7), 1619–1627. https://doi.org/10.1021/acschembio.9b00348.

(12) Shen, Y.; Bax, A. Protein Backbone and Sidechain Torsion Angles Predicted from NMR Chemical Shifts Using Artificial Neural Networks. *J Biomol NMR* **2013**, *56* (3), 227–241. https://doi.org/10.1007/s10858-013-9741-y.

(13) Bibow, S.; Hiller, S. A Guide to Quantifying Membrane Protein Dynamics in Lipids and Other Native-like Environments by Solution-State NMR Spectroscopy. *The FEBS Journal* **2019**, *286* (9), 1610–1623. https://doi.org/10.1111/febs.14639.

(14) Babu, M. M.; Kriwacki, R. W.; Pappu, R. V. Versatility from Protein Disorder. *Science* **2012**, *337* (6101), 1460–1461. https://doi.org/10.1126/science.1228775.

(15) Williamson, M. P. Chemical Shift Perturbation. In *Modern Magnetic Resonance*; Webb, G. A., Ed.; Springer International Publishing: Cham, 2018; pp 995–1012. https://doi.org/10.1007/978-3-319-28388-3_76.

(16) Clore, G. M.; Iwahara, J. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. *Chem. Rev.* **2009**, *109* (9), 4108–4139. https://doi.org/10.1021/cr900033p.

(17) Clore, G. M.; Tang, C.; Iwahara, J. Elucidating Transient Macromolecular Interactions Using Paramagnetic Relaxation Enhancement. *Current Opinion in Structural Biology* **2007**, *17* (5), 603–616. https://doi.org/10.1016/j.sbi.2007.08.013.

(18) Battiste, J. L.; Wagner, G. Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with

Limited Nuclear Overhauser Effect Data. *Biochemistry* **2000**, *39* (18), 5355–5365. https://doi.org/10.1021/bi000060h.

(19) Keizers, P. H. J.; Desreux, J. F.; Overhand, M.; Ubbink, M. Increased Paramagnetic Effect of a Lanthanide Protein Probe by Two-Point Attachment. *J. Am. Chem. Soc.* **2007**, *129* (30), 9292–9293. https://doi.org/10.1021/ja0725201.

(20) Gong, Z.; Schwieters, C. D.; Tang, C. Theory and Practice of Using Solvent Paramagnetic Relaxation Enhancement to Characterize Protein Conformational Dynamics. *Methods* **2018**, *148*, 48–56. https://doi.org/10.1016/j.ymeth.2018.04.006.

(21) Pintacuda, G.; Otting, G. Identification of Protein Surfaces by NMR Measurements with a Paramagnetic Gd(III) Chelate. *J. Am. Chem. Soc.* **2002**, *124* (3), 372–373. https://doi.org/10.1021/ja016985h.

(22) Softley, C. A.; Bostock, M. J.; Popowicz, G. M.; Sattler, M. Paramagnetic NMR in Drug Discovery. *J Biomol NMR* **2020**, *74* (6), 287–309. https://doi.org/10.1007/s10858-020-00322-0.

(23) Huang, S.; Umemoto, R.; Tamura, Y.; Kofuku, Y.; Uyeda, T. Q. P.; Nishida, N.; Shimada, I. Utilization of Paramagnetic Relaxation Enhancements for Structural Analysis of Actin-Binding Proteins in Complex with Actin. *Scientific Reports* **2016**, *6* (1), 33690. https://doi.org/10.1038/srep33690.

Chapter 4: Bioinformatic Mapping of Radical SAM-Dependent RiPPs Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides[2]

Graham A. Hudson, Brandon J. Burkhart, Adam J. DiCaprio, Christopher J. Schwalen, Bryce Kille, Taras V. Pogorelov , Douglas A. Mitchell

4.1 Introduction

With the growing availability of genome sequences, bioinformatics has become an increasingly popular and powerful technique for natural product discovery.[1,2] The biosynthetic gene clusters (BGCs) of known natural product classes are readily identified by the presence of conserved genes and the structure of their products can be predicted to varying degrees of accuracy based on the number and type of enzymes locally encoded.[3] This approach has been particularly useful in the discovery of ribosomally synthesized and posttranslationally modified peptides (RiPPs).[4] RiPPs have no universally conserved gene but instead are united by a common biosynthetic logic: with few known exceptions, the biosynthetic enzymes bind their respective precursor peptides through a recognition sequence in the N-terminal region of the peptide, referred to as the leader peptide,[5] and the posttranslational modifications are installed on the C-terminal region, referred to as the core peptide. During biosynthetic maturation, the leader peptide is eventually removed, and frequently additional enzymatic tailoring of the peptide occurs to yield the final RiPP product. Several posttranslational chemistries are known and are used to categorize RiPPs into their respective classes.[4,6]

Despite the utility of genome mining to discover new natural products, the identification of RiPP BGCs remains challenging. First, RiPPs require analysis of a local genomic context for the presence of encoded regulators and transporters, or better yet, clear supporting genetic markers of RiPP biosynthetic pathways because RiPP enzymes are often members of larger protein

[2] A.J.D. contributions: All NMR data acquisition, processing and interpretation. Generation of all NMR figures and assistance with manuscript preparation

superfamilies that bear homology to proteins not associated with natural product biosynthesis. Second, RiPP precursor peptides are often short and hypervariable, which means their coding sequences are often unannotated when deposited into public databases and they are often not identified by sequence homology search tools (e.g., BLAST[7]). To address these challenges, we developed a bioinformatics program, Rapid ORF Description & Evaluation Online (RODEO), which automates the genome mining process.[8] A typical RODEO input will be a list of NCBI gene accession identifiers for a biosynthetic protein of interest. The program then fetches genomic records from GenBank and uses the PFAM[9] and/or TIGRFAM[10] databases to predict the function of neighboring genes. Next, RODEO identifies all possible precursor peptides and scores their likelihood of being a true RiPP precursor using a scoring function based on motif analysis and supervised machine learning. Given the disparate nature of RiPP precursor peptides, the scoring function is optimized for each distinct class. In our initial report, RODEO was leveraged to identify and classify lasso peptides which led to an order of magnitude expansion of the RiPP class. RODEO has also been reconfigured to aid in the discovery of thiopeptides[11] and was incorporated into antiSMASH (version 4.0),[12] a tool which mines single genome inputs for natural products. Once the requisite sequence attributes and machine learning classifications are established, performing updates for a RiPP class via RODEO is straightforward, as demonstrated very recently with the lasso peptides.[13]

Given the success of RODEO in defining other RiPP classes, we sought to expand its capabilities to aid in the discovery of new sactipeptides. Sactipeptides have garnered interest due to their unique hairpin structures and potent, narrow-spectrum activity towards several human pathogenic bacteria.[14-16] They are defined by a radical SAM (rSAM) enzyme that forms "sactionine" crosslinks between a donor Cys sulfur and the alpha carbon of an acceptor amino acid (**Figure**

1).[15,17,18] Prevalent in RiPP pathways, rSAM enzymes are versatile catalysts that employ [4Fe-4S] clusters to reductively cleave SAM to generate a 5′-deoxyadenosyl radical.[19-23] This reactive intermediate often will abstract a hydrogen atom from the substrate, which leads to a variety of different outcomes based on the specifics of a given biosynthetic pathway. The precise mechanistic details of sactipeptide crosslink formation remains poorly characterized, although several groups are taking on this challenge.[24,25] Studies of the sactipeptide rSAM enzyme itself have revealed that sactisynthases contain three [4Fe-4S] clusters: one which reductively cleaves SAM and two auxiliary clusters in what has been called dubbed a SPASM domain.[26-29] Only five peptides with the hallmark S–Cα crosslinks have been isolated thus far: subtilosin A,[30, 31] thurincin H,[30] thuricin CD (two peptide products),[16] and sporulation killing factor.[31,32] A sixth sactipeptide, hyicin 4244, was recently reported but was not isolated for detailed characterization.[33] All structurally characterized sactipeptides contain ring-within-a-ring topologies and thus form rigid, hairpin-like structures (**Figure 1**). However, in 2011, Haft reported the bioinformatic discovery of a related class of putative thioether-crosslinked peptides, which he referred to as the six Cys in forty-five residue peptides (SCIFFs).[34] No bona fide SCIFFs have been isolated, but in vitro enzymatic reactions with the purified rSAM and precursor peptide indicated that a crosslink is formed between a donor Cys and an acceptor Thr residue.[24,25]

Similar to our earlier work that redefined the genomic landscape of the lasso peptides and thiopeptides, we herein conducted a comprehensive analysis of the sactipeptide RiPP class using a re-written and improved edition of RODEO, version 2.0. This release of RODEO features numerous improvements, including reduction of redundant calculations, more robust handling of internet connectivity and genomic record formatting errors, and enhanced multithreading that enables parallelization of data fetching from GenBank and subsequent processing, collectively

leading to a large increase in processing speed. Although previous bioinformatic surveys have been conducted on the sactipeptide class,[35-39] none have employed machine learning to cast a wider net for identifying candidate precursor peptides. In this study, we have conducted a comprehensive analysis and connected 3,156 rSAM enzymes to their cognate RiPP precursor peptide and found new varieties of thioether-linked natural products. Guided by this dataset, we selected a new sactipeptide, huazacin, for characterization that contains a thioether ring architecture larger than any sactipeptide previously characterized. Additionally, our bioinformatic analysis indicated that SCIFF-associated rSAM maturases were much more closely related to QhpD proteins than characterized sactisynthases. QhpD is a rSAM enzyme which is known to form β- and γ-linked thioethers,[40] thus calling into question the linkage architecture within a SCIFF natural product. We pursued this line of inquiry resulting in the bioinformatic identification and structural characterization of two SCIFF natural products. The first, freyrasin, features six S-Cβ thioether linkages. The second, thermocellin, has previously been reported in the literature and characterized only after partial in vitro enzymatic reconstitution. We have confirmed thermocellin to feature a Cys-Thr thioether as reported, but rather than harboring a S–Cα linkage, the peptide contains a S–Cγ thioether linkage. To distinguish this group from the structurally distinct sactipeptides, we propose the renaming of this RiPP class as the ranthipeptides, for radical non-alpha thioether peptides.

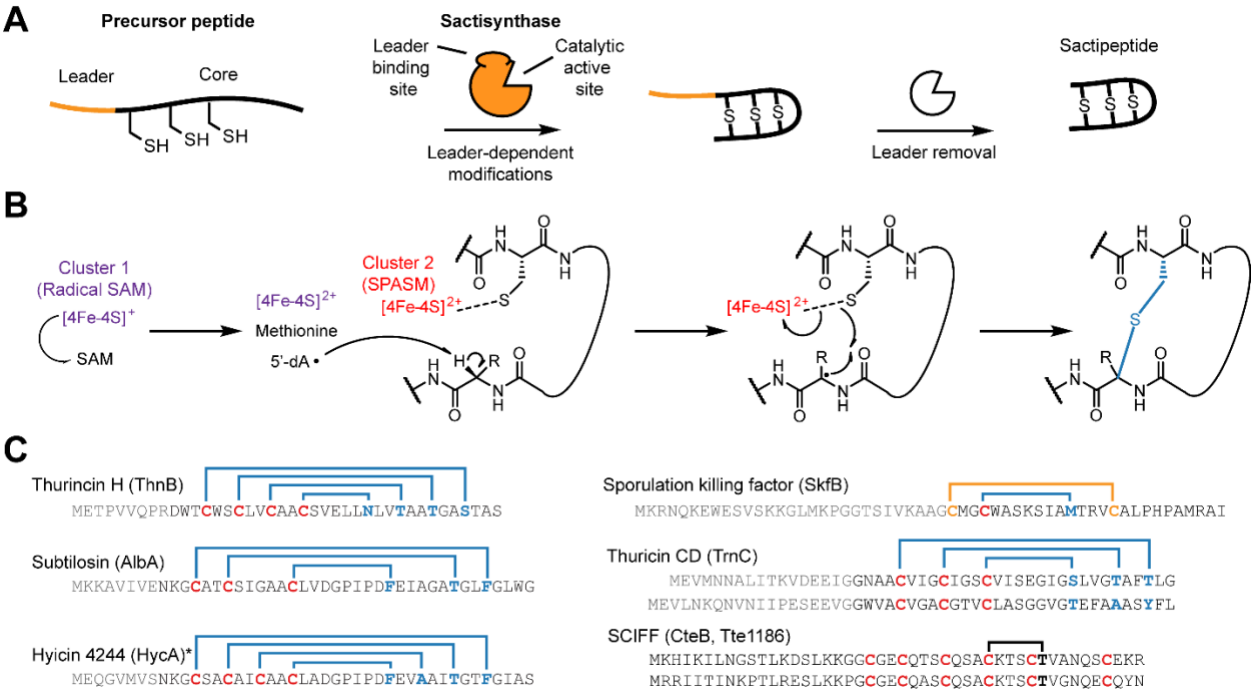


Figure 4.1. Sactipeptide biosynthesis and ring topologies. (**A**) Schematic overview of sactipeptide biosynthesis. (**B**) Proposed mechanism for the rSAM-dependent formation of the sactipeptide S–Cα thioether crosslink. (**C**) Precursor peptide sequence and ring topologies of characterized sactipeptides (blue, S–Cα linkage; yellow, S-S linkage; black, uncharacterized). The requisite rSAM enzyme name is given in parentheses. Characterization of the six Cys in forty-five residues (SCIFFs) is limited to in vitro enzymatic assay for which only one thioether linkage was observed. Data shown later in this manuscript support a Cys-Thr linkage different to what has been previously reported and further that SCIFFs are not sactipeptides. * Hyicin 4244 has not been structurally characterized, thus the linkages indicated are speculative.

4.2 Results and Discussion

4.2.1 RODEO2.0-enabled sactipeptide discovery

With >450,000 rSAM proteins listed in the InterPro 72.0 database (IPR007197),[41] a more focused bioinformatic search was performed to winnow down candidates more likely to be involved in

sactipeptide biosynthesis. Therefore, four rounds of iterative PSI-BLAST (Position-Specific Iterated Basic Local Alignment Search Tool[42]) searches using rSAM sequences from four known sactipeptides (AlbA, SkfB, ThnB, and TrnC)[15] and two partially characterized SCIFFs (CteB[25] and Tte1186[24]) were performed to identify candidate proteins. A stringent expectation value of 1e-70 was used as a cut-off to assist in keeping the retrieved candidate proteins more likely to be involved in RiPP biosynthesis in addition to filtering sequences <300 amino acids. This procedure yielded ~4,600 protein sequences, which were subsequently analyzed by RODEO2.0 (see **Supplementary Methods**; output available in **Supplementary Dataset 1**), to annotate the local genomic context and score potential sactipeptide precursor sequences (**Supplementary Dataset 2**).

Sactipeptide precursor peptides were identified using a series of heuristics in conjunction with support vector machine classification as previously reported resulting in 504 candidate precursor sequences.[12] The top scoring sequences were then manually inspected for similarity to known sactipeptides. In general, a peptide was considered to be a strong candidate sactipeptide if it had 3 or more Cys separated by equal spacing of 1 to 3 residues. Additionally, peptides were scored more favorably if the Cys residues were clustered within the N-terminal half of the core peptide, similar to that observed for other sactipeptides (i.e. the donor residues line one side of the hairpin while the receptor residues line the other side, **Figure 4.1**). Precursor peptides were further analyzed by a sequence similarity network (SSN) generated using the Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST)[43] which revealed that clustering of precursors appears to be largely driven by conservation of donor Cys residues (**Figure S1**), which is predicted to govern the location of crosslinks and thus the overall structure of the sactipeptide. This bioinformatic analysis identified four significantly populated groups of uncharacterized sactipeptides. Two groups, named hypervariable groups 1 and 2, possess precursor peptides that are more than double

the length of any known sactipeptide. While the core sequences within these two groups are overall "hypervariable", there are 6-8 locations that display highly conserved Cys residues. A third uncharacterized group of sactipeptides are concentrated in the *Lachnospiraceae* family, thus we have termed these predicted compounds the "lachnocins" (**Figure S1**).

We elected, based on the above bioinformatics results, to pursue the fourth uncharacterized sactipeptide group. The target producer chosen was *Bacillus thuringiensis* serovar *huazhongensis* (**Figure 4.2**), owing to strain availability from the Bacillus Genetic Stock Center (BGSC 4BD1) and the fact that the predicted structure contained a unique crosslinked structure (4 crosslinks separated by 3 residues each). Although bioinformatically predicted in a previous study,[39] the mature natural product has never been reported. Upon screening various culture extracts by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), we observed a mass consistent with the predicted structure (**Figure S2**). The metabolite, hereafter huazacin, was then subjected to high-resolution and tandem MS (HR-MS/MS) analysis using an Orbitrap instrument. The observed *m/z* value ($[M+3H]^{3+}$ = 1241.5396) agreed with the theoretical mass of the expected molecular formula of $C_{160}H_{241}N_{43}O_{50}S_5^{3+}$ ($[M+3H]^{3+}$ = 1241.5393; error = 0.2 ppm, **Figures S2 and S3**). This formula contains four thioether linkages (8 Da lighter than the unmodified core peptide). Upon subjecting the parent ion to collision-induced dissociation (CID) conditions, we observed robust fragmentation of the amide and thioether linkages (**Figure S3**). Similar to other characterized sactipeptides, this gas-phase dissociation of thioethers enabled the assignment of acceptor residues and was consistent with the thioether linkages of huazacin being S-Cα linked (i.e., sactionines), as predicted.[44]

Multidimensional NMR spectroscopy (^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY) were collected on HPLC-purified huazacin, which corroborated the sactionine (S-Cα) linkages of huazacin (**Figures S4-S5, Table S3**). We observed NOESY correlations for the Cys-CαH (and Cys-CβH) to the recipient site backbone NH (**Figure S6**). Further, no TOCSY correlations were observed from NH to CαH, which was consistent with a quaternary carbon at the alpha position as expected for a sactionine linkage (**Figure S7**). Ring topology was also established by NOESY analysis of the Cys12-Thr30 and Cys8-Tyr34 rings.

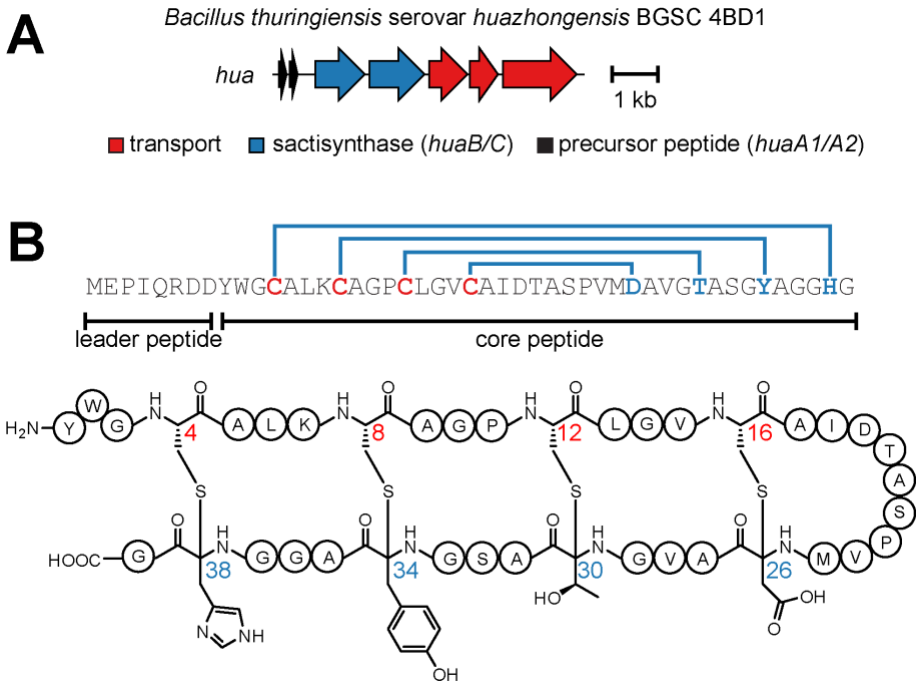


Figure 4.2. Huazacin from *Bacillus thuringiensis* serovar *huazhongensis*. (**A**) Biosynthetic gene cluster. The sactipeptide synthases HuaB and HuaC (NCBI accession identifiers: EEM79976.1 and EEM79977.1, respectively) feature an N-terminal RRE domain and C-terminal SPASM domain similar to other sactisynthases.[27,45] (**B**) Precursor peptide sequence and chemical structure of huazacin. The *hua* gene cluster contains two identical copies of the precursor peptide shown.

We next assessed the antibacterial activity of huazacin by testing against a diverse panel of bacteria for growth suppression using the microbroth dilution method. Reported sactipeptides typically have a narrow spectrum of activity towards Gram-positive organisms, most notably *Clostridium difficile*, *Listeria monocytogenes*, and *Bacillus cereus*.[16,17,46-49] Consistent with previously characterized sactipeptides, huazacin inhibited the growth of *L. monocytogenes* 4b F2365 with a minimum inhibitory concentration (MIC) of 4 μM (**Table S4**). These results with huazacin reinforce the concept that RODEO is a useful tool to accelerate RiPP discovery. While many additional, predicted sactipeptides are found in less populated clusters (with many being singletons, **Figure S1**), these data provide a roadmap to guide for future sactipeptide discovery efforts.

4.2.2 Mining the rSAM-modified RiPP thioether genomic landscape

Having shown the utility of RODEO for the discovery of a new sactipeptide, we sought to more broadly survey the genomic landscape of rSAM-modified RiPPs that were likely to contain thioether linkages. We began this effort by visualizing the sequence similarity of our list of rSAM enzymes (from PSI-BLAST) in the form of a SSN (**Figure 4.3**). Sequences with >70% amino acid identity were conflated to a single node and connected by an edge if significant sequence similarity was shared between the two sequences (expectation value < 1e-50). Thus, only highly similar rSAM proteins form clusters by this analysis. Given that the similarly of a RiPP modifying enzyme strongly correlates with the similarity of the cognate precursor peptides (i.e. similar rSAM enzymes produce similar products),[8] a properly annotated SSN provides a simplified but informative view of the peptide substrate diversity.

To provide context for the rSAM SSN, the nodes were colored based on the sequence characteristics of the cognate precursor peptide identified by RODEO (**Figure 4.3**). Several of the clusters were readily annotated by similarity to known sactipeptide precursor peptides, such as subtilosin (SboA), thurincin H (ThnA), thuricin CD (TrnA), and sporulation killing factor (SkfA). The novel sactipeptide huazacin (HuaA) described above is also indicated on the network. Two outlier sequences, the thermocellins (CteA and Tte1186a), belong to the so-called SCIFF family and are contained in a much larger separate group. Additionally, two large groupings of uncharacterized rSAM proteins with low-level similarity to TrnC were readily identified. The RODEO-identified precursor peptides encoded next to these rSAM proteins comprise the above-mentioned hypervariable groups 1 and 2 (**Supplementary Dataset 2**).

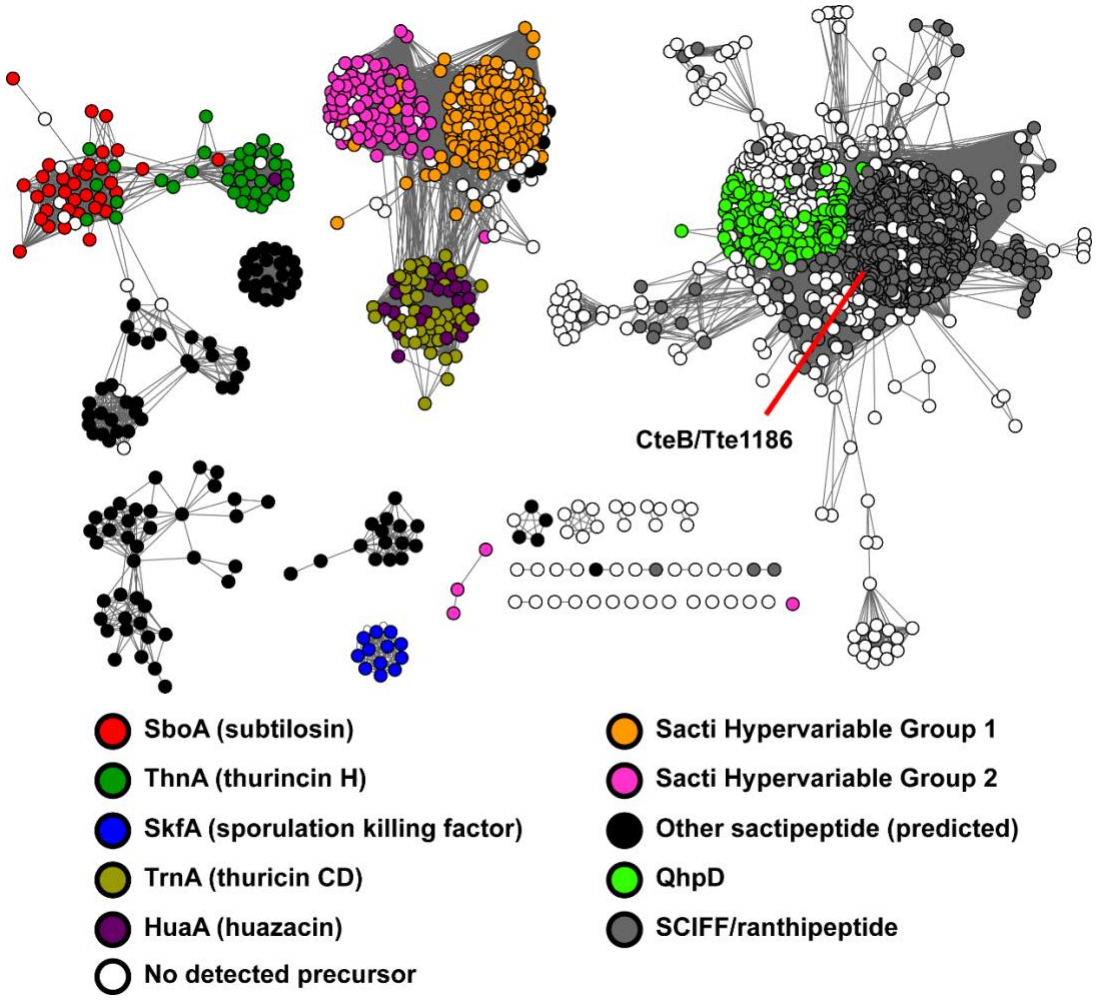


Figure 4.3. Sequence similarity network of rSAM enzymes. rSAM enzymes involved in sactipeptide and SCIFF biosynthesis were used as BLAST inputs to gather additional rSAM enzymes predicted to form RiPP thioether linkages (*n* = 4,693). Sequences were analyzed by EFI-EST[43] and visualized using Cytoscape.[50] Nodes are colored based on the category of its associated precursor peptide (see legend). Sequences sharing >70% identity are conflated as a single node, and connected nodes indicate a similarity score of < 1e-50.[43,50,51] Sequence logos for core peptides associated with these rSAMs are shown in **Figures S1** and **S9**. QhpD, non-RiPP rSAM involved in modifying quinohemoprotein amine dehydrogenase.[40]

Several trends for the number and position of Cys residues within the peptides were further analyzed to gain deeper insight into the RiPP thioether landscape (**Figure S8**). The positioning of the Cys residues indicated that many groups had the potential to form the traditional sactipeptide hairpin-like topology wherein Cys residues are concentrated near the N-terminus and crosslink to acceptor residues concentrated near the C-terminus. Only the SkfA and SCIFF groups seemed capable of adopting an alternative structural topology based on their Cys residues being more evenly distributed (**Figure S8**). However, SKF is known to be N-to-C macrocyclized and it is the only known sactipeptide to contain an internal disulfide bond. No mature SCIFFs have been reported in the literature, thus their topologies have remained undetermined. The positioning of the Cys residues would suggest SCIFFs display non-hairpin topologies (**Figures 4.1** and **S9**).

The majority of the identified thioether-forming rSAM enzymes reside within the large SCIFF cluster. Despite their greater numbers, the cognate precursor peptides display low sequence diversity. The RiPP products of these BGCs have been termed the thermocellins, although the final structure of thermocellin has not been reported. We identified ~2600 SCIFF precursor peptides (1,023 unique sequences), which outnumber the RODEO-identified sactipeptides by three-fold (**Figure S9** and **Supplemental Dataset 1)**. While the majority of these precursor peptides are virtually identical to thermocellin (CteA/Tte1186a), numerous smaller groupings display distinct core peptides (**Figure S9**). Several of these smaller groups display conserved Asp residues along with the putative Cys donors, indicating that the Asp residues may often serve as an acceptor for thioether formation.

To uncover additional phylogenetic trends, we inspected the taxonomic distribution of all rSAMs predicted form thioether-containing RiPPs. The vast majority (~95%) are encoded by Firmicutes, consistent with *Clostridium* sequencing bias as well as the fact that all previously isolated sactipeptides originate from *Bacillus* (**Figure S10**). Nonetheless, divergent examples occur in numerous other phyla, including Proteobacteria, Actinobacteria, and Bacteroidetes, albeit with much lower frequency. We note that some of the genes encoding the rSAM enzyme from *Clostridium* showed significantly higher % GC content when compared to the entire genome, which may be indicative of horizontal gene transfer (**Figure S11**). Additionally, based on local genomic context (**Table S1**), many unusual gene cluster architectures were identified including cases where the RiPP may be further elaborated by locally encoded enzymes (**Figure S12**). Co-occurrence with genes related to tRNA metabolism were also highly represented in the genomic context of SCIFFs, which may hint towards a function for these RiPPs, although this will require additional investigation (**Table S2** and **Figure S12**).

Not every rSAM enzyme in our dataset could be confidently assigned a substrate peptide. Therefore, the precursor peptides in such cases remain "unidentified" due to one of several possible reasons, such as *(i)* poor quality genome assembly, *(ii)* the precursor peptide is encoded elsewhere in the genome, *(iii)* the BGC has been deactivated (a relic) and there is no longer a coding sequence for the precursor peptide, or finally, (*iv*) the rSAM is not involved in RiPP biosynthesis. Indeed, some members within the unidentified group do not appear to be involved in RiPP biosynthesis; however, our investigation found that only a few examples could be confidently placed in "non-RiPP" category based on local genomic context. A final group of rSAM enzymes were identified as orthologs of QhpD. These proteins are not involved in RiPP biosynthesis but rather they are involved in the posttranslational modification of a large protein.[52-54] Nonetheless, this indicates

that our targeted BLAST searching was sufficient in identifying many thioether-installing rSAM enzymes involved in RiPP biosynthesis.

4.2.3 Insights from relatedness to QhpD

Given our targeted BLAST searching, it was initially surprising to retrieve such a large number of enzymes annotated as QhpD (n = 674, **Supplemental Dataset 2**). QhpD is not involved in RiPP biosynthesis but rather the posttranslational maturation of quinohemoprotein amine dehydrogenase (QHNDH) which catalyzes the oxidative deamination of aliphatic and aromatic amines.[55-57] QHNDH is composed of three subunits and contains several unusual posttranslational modifications, including a cysteinyl tryptophylquinone cofactor as well as three catalytically essential thioether linkages. Two are between Cys-Asp (S-Cβ) while the third thioether linkage is between Cys and Glu (S-Cg).[52-54] As illustrated on the SSN, QhpD and orthologous rSAM enzymes share considerable sequence similarity to rSAM enzymes involved in SCIFF biosynthesis (**Figure 4.3**). This finding is significant because sactipeptides are defined based on a thioether crosslink to the alpha carbon (S-Cα) of an acceptor residue. Therefore, in addition to modifying a protein instead of a RiPP precursor peptide, the QhpD group should be carefully distinguished from rSAM enzymes that form sactipeptides. However, considering the stringency used during our bioinformatics survey and the similarity score threshold in the SSN, we predicted that any rSAM enzyme connected to the QhpD (i.e. the SCIFF cluster) would be more likely to catalyze S-Cβ and/or S-Cg thioether linkges rather than a sactionine linkage. While an SSN provides a useful and simplified illustration of sequence similarity, the relatedness of the separate clusters is not depicted. Thus, the same set of rSAM enzymes were analyzed after constructing a maximum likelihood tree (**Figure S13**). Corroborating the previous analysis, rSAM enzymes implicated in SCIFF biosynthesis are more related to QhpD than those involved in sactionine formation (**Table**

S7). A subset of SCIFFs with precursor peptides distinct from those previously reported (Tte1186a and CteA) were also identified that are more similar to NxxcB, a rSAM enzyme recently described to form a S-Cβ thioether linkage to an Asn acceptor in a RiPP that is neither a sactipeptide nor a SCIFF.[58] It should be noted that this target dataset was aggregated using BLAST queries of biosynthetic enzymes known to be involved in sactipeptide and SCIFF biosynthesis, rather than using an all-inclusive list of rSAM proteins. This resulted in a more manageable dataset size but one that also lacks many proteins that may or may not be involved in peptidic thioether formation. Thus, the relationship of the rSAM proteins involved in sactipeptide versus SCIFF biosynthesis appear here artificially more related to one another. To illustrate this, a BLAST query using AlbA (subtilosin, sacipeptide) will return nearly 40,000 protein sequences before retrieving the less-similar CteB (thermocellin, SCIFF).

4.2.4 Freyrasin is a S–Cβ crosslinked peptide: MS evidence

With bioinformatic evidence that rSAM proteins annotated as SCIFF maturases might not contain S-Cα thioether linkages, and that previous studies never confirmed the nature of the SCIFF crosslink in the SCIFFs,[24,25] we next sought to determine the connectivity of the SCIFF thioether crosslink. To obtain the milligram quantities of peptide that would be required for spectroscopic characterization, we first targeted the predicted SCIFF from *Paenibacillus polymyxa* ATCC 842 (**Figure 4.4**). This BGC was selected because it displayed a unique sequence variation of six CX_3D repeats which suggested that each of the six Cys donors would be linked to an Asp acceptor (similar to QhpD, **Figure S9**). After MS-based screening of various culturing conditions, we did not detect any ions consistent with the expected natural product. Therefore, we heterologously coexpressed the precursor peptide (PapA) and rSAM enzyme (PapB) in *Escherichia coli* following a previously published method.[44,59] After purification, the peptide product obtained was analyzed

by MALDI-TOF-MS, which indicated that the precursor peptide had undergone processing with a mass loss of 12 Da (consistent with six thioether linkages, **Figure S14**). Controls where the precursor peptide was expressed in the absence of the rSAM enzyme resulted in a product with a mass loss of 6 Da from the starting material, consistent with the formation of three disulfide bonds). In contrast to the single crosslinks reported from in vitro studies on Tte1186a and CteA (thermocellin),[24,25] we observed a single product where all of the six Cys were modified. We have termed this group the freyrasins.

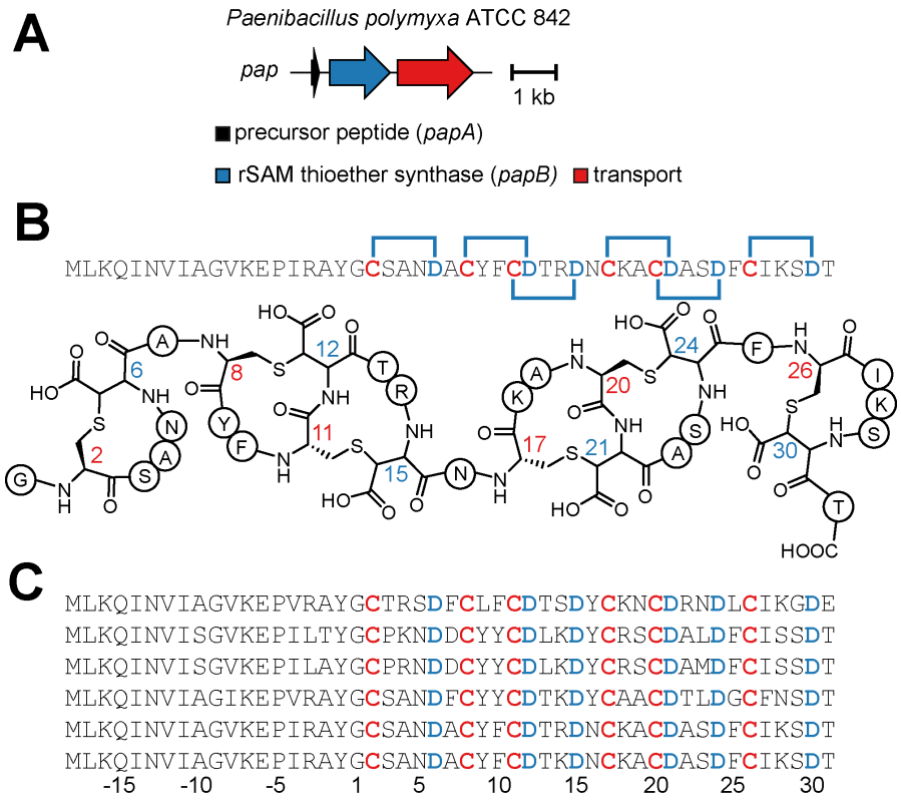


Figure 4.4. Freyrasin biosynthetic gene cluster and related peptides. (**A**) Biosynthetic gene cluster diagram. (**B**) Freyrasin precursor peptide sequence with six CX_3D motifs and thioether topologies highlighted. (**C**) Alignment of additional freyrasin members. Leader peptide residues are negatively numbered relative to the core peptide, which we predict begins with Gly.

The structure of freyrasin obtained through heterologous expression was further examined by HR-MS/MS (**Figure S15**). The observed *m/z* value ($[M+4H]^{4+}$ = 1397.6113) agreed with the molecular formula for the predicted structure ($C_{233}H_{262}N_{66}O_{80}S_7^{4+}$; $[M+4H]^{4+}$ = 1397.6094; error = 1.4 ppm). The tandem MS data were also consistent with the freyrasin precursor peptide, although the fragmentation pattern observed showed a limited number of ions for a peptide of this size. Significant ions were only detected for sites that dissociated either directly before or after non-overlapping CX_3D motifs. Such fragmentation behavior was reminiscent of a lanthipeptide-like structure as opposed to a sactipeptide-like structure.[60-62] Indeed, prior work from our lab and others have shown that the sactionine linkage itself (S-Cα thioether) readily dissociate under standard CID condition whereas S-Cb thioether linkages do not, as is well-documented for lanthipeptides.[16,59,63] These observations are further consistent with the lack of inter-ring fragmentation data recently reported from in vitro work on SCIFF peptides that contain a single thioether.[24,25] By virtue of not undergoing this characteristic fragmentation, we speculated that freyrasin might not contain S-Cα linkages.

To better understand this difference in CID fragmentation behavior, we initiated quantum mechanical calculations of two model tripeptides (Gly-Cys-Gly and Gly-Asp-Gly) that were Cys-Asp thioether linked to the Cα or Cβ of Asp (**Supplemental Methods**). These calculations revealed that the difference in zero-point energy between the two model peptides was comparable to thermal energies (1 kcal/mol) and thus negligible. However, the electronic energy of the S-Cβ linked model peptide was found to be 12 kcal/mol more stable than the corresponding S-Cα linked isomer (**Figure S16**). We attribute the weaker bond energy of the S-Cα linkage to the greater stability of the alpha-centered radical that arises after homolytic bond dissociation. Taken with the

bioinformatic predictions, these calculations provide an explanation for the differential CID behavior between the two linkage types and further suggests that SCIFF thioethers are not alpha linked.

4.2.5 Freyrasin is a S–Cβ crosslinked peptide: NMR evidence

To rigorously establish the atomic connectivity of the freyrasin thioether linkages, various multidimensional NMR experiments were performed. Proteolytic digestion yielded a peptide cleaved between Tyr/Gly that was employed for NMR analysis (**Figure S17**). Using the known precursor peptide sequence (**Figure 4.4**) and ^{1}H-^{1}H TOCSY spectrum, residues were assigned based on the cross-peak patterns observed in the amide region from 7.70–8.60 ppm (**Figure S18**). The ^{1}H-^{1}H NOESY spectrum was used to confirm the linear freyrasin sequence (**Figure S19**). All residues were confidently assigned with the exception of Cys17 (**Table S5**). Through-space correlations were visible for all residues with the exception of Asn5 and Asn16. This is likely due to a flexible structure, as demonstrated by a narrow amide chemical shift dispersion.[64] As revealed in the TOCSY spectrum, freyrasin displayed NH-CαH-CβH correlations for five of the six Asp residues believed to be thioether crosslinked, confirming a continuous spin system. This correlation would not be possible for a S-Cα thioether (i.e. sactionine). Additionally, each of the six Asp residues generated NH-CαH-CβH TOCSY correlations similar to that of Ser, with two deshielded resonances occurring in the CαH region (3.6–4.7 ppm), again supportive of a Cys-Asp (S-Cβ) linked thioether (**Figure S20).** To assign the ring topology of freyrasin, NOESY correlations between Cys-Asp were assigned based on the chemical shift values obtained by TOCSY (**Figure 4.5**). Two of the six thioether rings were assigned unambiguously. The redundancy of the freyrasin primary sequence unfortunately led to extensive chemical shift overlap compounded with line broadening, presumably arising from the flexibility of the peptide, so

complete NOESY correlations were not evident for the remaining rings in freyrasin. Nonetheless, when coupled with the bioinformatics, MS data, and the definitive NMR assignments, freyrasin has been determined to contain exclusively Cys-Asp (S-Cβ) linkages. With the structure confirmed, we next tested for antibiotic activity of the heterologously obtained freyrasin, but it was not significantly active against any strain tested (MIC ≥ 32 μM). It is possible that additional modifications are present in the native product, but also possible are that freyrasin harbors very narrow-spectrum activity towards an untested strain or instead functions in a capacity unrelated to growth suppression (**Table S4**).

4.2.6 Thermocellin contains a Cys-Thr (S-Cγ) linked thioether

Given that freyrasin contains exclusively S-Cβ linkages in agreement with our bioinformatic predictions, we sought to determine if other "SCIFFs" also lacked sactionine linkages. The only two (partially) characterized SCIFFs thus far reported (i.e. Tte1186a and thermocellin) were previously investigated after in vitro enzymatic reconstitution, which resulted in the formation of

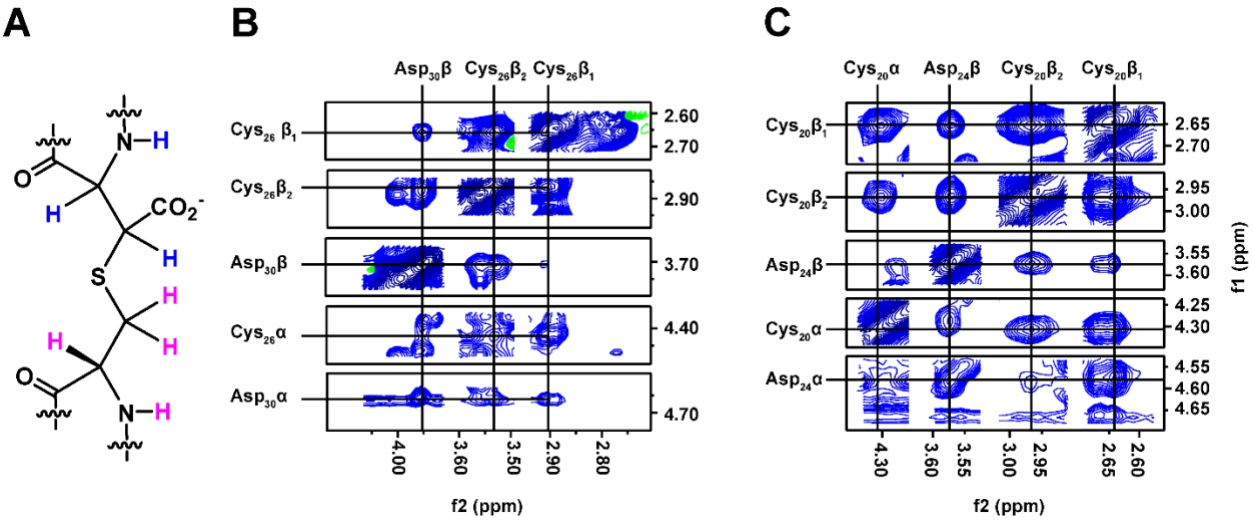


Figure 4.5. Freyrasin contains beta-linked thioethers. (**A**) Cys-Asp (S–Cβ) connectivity with hydrogens colored pink (Cys) or blue (Asp). (**B**) ^{1}H-^{1}H NOESY correlations confirming a S–Cβ thioether linkage between Cys26-Asp30. (**C**) Same as panel B but for Cys20-Asp24.

only one Cys-Thr thioether.[24,25] Owing to the precursor containing six cysteines, the single thioether-containing species is likely a biosynthetic intermediate and not reflective of fully mature thermocellin. Despite these studies drawing on homology to sactionine-forming rSAM enzymes to predict linkage type, the presumed thioether installed by these enzymes did not undergo the gas-phase dissociation reaction that is characteristic of sactionine linkages. When further considering *(i)* our tandem MS data that show huazacin, a Cys-Thr (S-Cα)-containing sactipeptide, behaved as expected under CID conditions (**Figure S3**) and *(ii)* the thermocellin-forming rSAM enzyme, CteB, is more related to QhpD than any known sactionine-forming rSAM enzyme (**Figure S13**), we questioned whether thermocellin actually contains S-Cα linkages, as previously reported.

The possible thioethers for the Cys-Thr crosslink in thermocellin could theoretically include either the alpha, beta, or gamma carbon of Thr. We believed that in addition to an unlikely Cys-Thr (S-Cα) linkage for the above-mentioned reasons, a (S-Cβ) linkage would also be unlikely given the instability of resulting thiohemiacetal. To confirm or refute if thermocellin possessed a Cys-Thr (S-Cγ) crosslink, we prepared an isotopically labeled version of the thermocellin precursor peptide, CteA, that included ^{2}H at the α and β positions of Thr while retaining ^{1}H at the γ-methyl group. Isotopically labeled Thr was incorporated by overexpression of CteA in a Thr auxotroph (Δ*thrC*) of *E. coli* in defined minimal media where the only source of Thr was the supplied L-[4-^{13}C,2,3-$^{2}H_2$]-Thr. Initial protein preparations led to scrambling of the ^{13}C and ^{2}H labels into other amino acids, resulting in a complex series of masses (**Figure S21**). Isotopic scrambling was readily suppressed by the addition of natural abundance Ile and Gly to the expression media, resulting in production of fully labeled CteA with no detectable under- or over-labeling (**Figure S21**).[65] The location of the isotopic labels was verified to be confined to Thr residues by HR-MS/MS (**Figure S22**).

Having devised a means to probe the chemical nature of the thioether linkage in thermocellin, we next coexpressed CteA with CteB (the rSAM enzyme) using the same conditions that led to exclusive Thr isotope labeling. Consistent with a Cys-Thr (S-Cγ) linkage, a two Da mass loss was observed relative to the starting material. This result is consistent only with the abstraction of ^{1}H from Thr and the loss of the thiol hydrogen during thioether formation. Abstraction of ^{2}H from either the α or β carbons of Thr would have yielded a three Da mass loss (**Figure S23**). HR-MS/MS analysis of the product confirmed a Cys-Thr (S-Cγ) thioether species that did not undergo gas-phase dissociation associated with S-Cα thioethers. Unexpectedly, the tandem MS data unequivocally showed that the Cys-Thr linkage is to the Thr that resides two positions away to the donor Cys, which is in contrast to previous reports that show localization to a Thr five residues away from the donor Cys (**Figure 4.6**). This localization was further corroborated by the observation of fragmentation within the new assignment of the macrocycle in unmodified peptide reacted with *N*-ethylmaleimide (**Figure S24**). To evaluate if the incorporation of stable isotopes altered the in vivo activity of CteB, we analyzed the coexpression product with natural abundance Thr and verified the same location for the thioether linkage (**Figure S25**). While it cannot be ruled out that the difference in localization is due to previous work being done in vitro with purified enzymes, our data conclusively show that ^{1}H was abstracted from the γ-methyl group of Thr during thioether formation which is consistent only with a (S-Cγ) linkage.

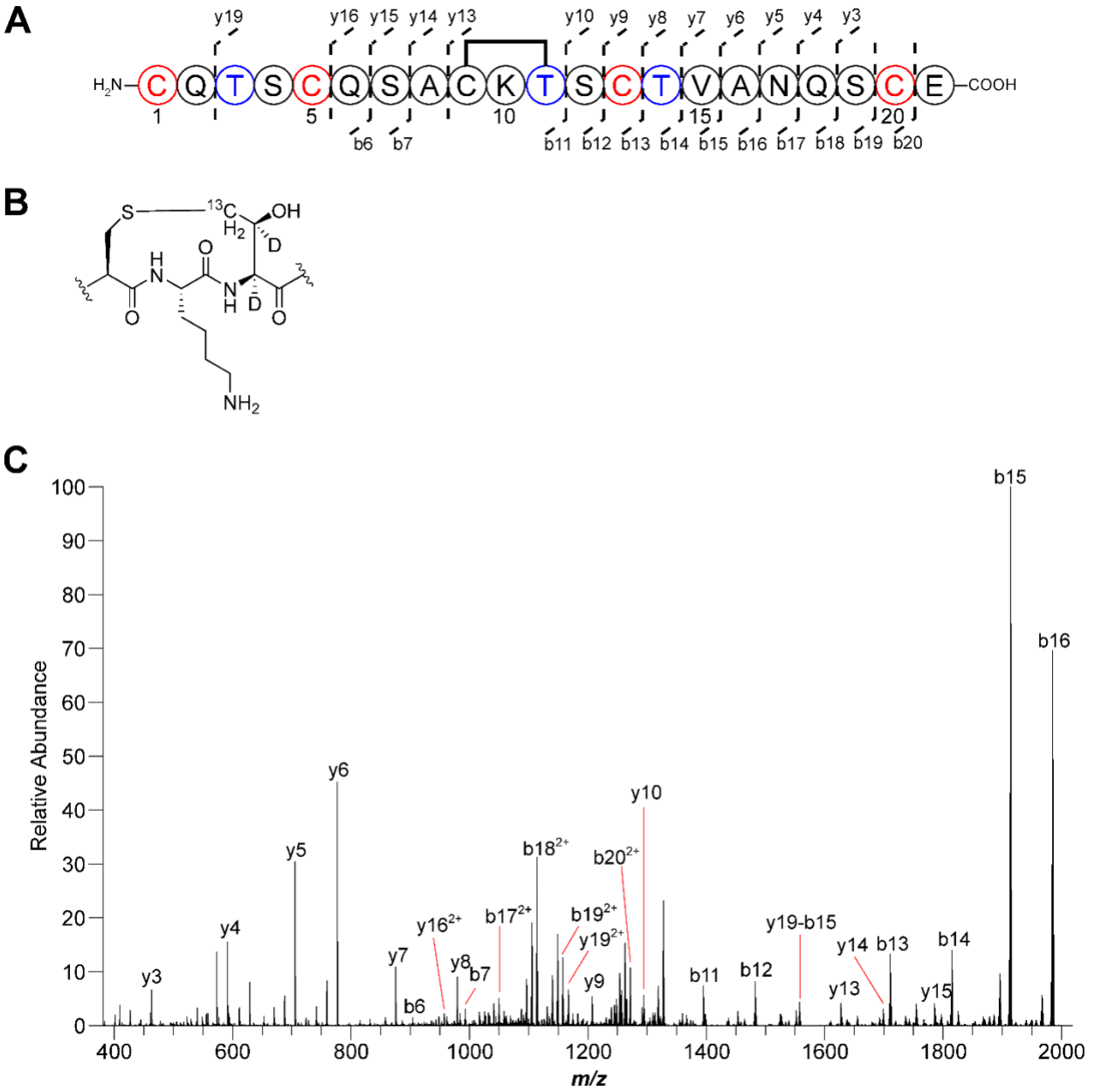


Figure 4.6. Thermocellin contains a Cys-Thr (S-Cγ) thioether. (**A**) Coexpression of the thermocellin precursor containing labeled Thr (blue) features a -2 Da shift relative to unmodified peptide, demonstrating a S-Cγ linkage was formed. The modified peptide was treated with endoproteinase GluC and unmodified Cys were alkylated with *N*-ethylmaleimide (red) Residues are numbered based on the sequence after GluC treatment. (**B**) Structure of the isotopically labeled thioether linkage. (**C**) MS/MS analysis yielded fragmentation consistent with a macrocycle

Figure 4.6 (cont.)formed between Cys9 and Thr11. Error tables for assigned ions are available in **Table S6**.

4.2.7 Proposal to alter nomenclature for thioether-containing natural products

Both freyrasin and thermocellin belong to the "SCIFF" family owing to both containing six Cys residues in a span of ~45 total precursor residues. As we have shown that freyrasin and thermocellin lack S-Cα thioethers but instead harbor S-Cβ and S-Cγ linkages, respectively. Therefore, we propose to rename this class the ranthipeptides (radical non-alpha thioether peptides) to distinguish them both structurally and biosynthetically from other thioether-containing RiPP classes, namely the sactipeptides and lanthipeptides.

4.3 Conclusion

In this work, we revealed that SCIFFs are not sactipeptides but instead feature rSAM-installed thioether crosslinks between Cys donor residues and the β- or γ-carbon of an acceptor residue. To better reflect their rSAM-dependent biosynthesis and to draw a correlation to lanthipeptides, which are S-Cβ linked, we propose the name ranthipeptides for this RiPP class. Our assignment of ranthipeptides as a RiPP class distinct from sactipeptides is supported by several independent lines of evidence including: (*i*) rSAM enzymes associated with ranthipeptide precursor peptides are considerably more sequence similar to QhpD, a known S-Cβ and S-Cγ thioether-forming enzyme, than any sactipeptide-associated rSAM enzyme, (*ii*) unlike sactionines, ranthipeptide thioethers do not fragment under mild CID conditions, an observation now supported by quantum mechanical bond dissociation energy calculations, (*iii*) multidimensional NMR spectroscopy, including ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY that support a Cys-Asp S-Cβ linkage for freyrasin, (*iv*) direct tandem MS and NMR comparison of freyrasin to the sactipeptide huazacin, which contains a Cys-Asp S-

Cα linkage, and (*v*) CteB (thermocellin-forming rSAM enzyme) abstracting hydrogen from the γ-carbon of the acceptor Thr during thioether formation. Collectively, these data provide conclusive evidence that both freyrasin and thermocellin are not alpha-linked. With previously reported beta-linked thioether-containing RiPPs being restricted to the lanthipeptides and the recently reported NxxcA, the ranthipeptide linkage chemistry is now considerably much richer as it is theoretically compatible with any non-Gly acceptor residue. This feature renders ranthipeptides an attractive target for engineering conformationally restrained peptides.

Over the past few years, several new peptide reaction chemistries catalyzed by RiPP rSAM enzymes[23] have been discovered and include the proteusins,[66] epipeptides,[67] streptide and structurally similar doubly-crosslinked RiPPs,[68,69] mycofactocin,[70,71] and the alpha-keto beta-amino acid-containing peptides.[45] The discovery of S-Cβ and S-Cγ linked thioethers adds ranthipeptides to this growing list of unique chemistry found in RiPP biosynthetic pathways. We propose that a combination of sequence similarity and tandem MS analysis will provide valuable insight in discerning the various linkage chemistries of yet-uncharacterized ranthipeptide products.

4.4 References

(1) Ziemert, N.; Alanjary, M.; Weber, T. The evolution of genome mining in microbes - a review. *Nat. Prod. Rep.* **2016**, *33* (8), 988.

(2) Medema, M. H.; Fischbach, M. A. Computational approaches to natural product discovery. *Nat. Chem. Biol.* **2015**, *11* (9), 639.

(3) Tietz, J. I.; Mitchell, D. A. Using genomics for natural product structure elucidation. *Curr. Top. Med. Chem.* **2016**, *16* (15), 1645.

(4) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.;

Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K. D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J.; Rebuffat, S.; Ross, R. P.; Sahl, H. G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G. L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108.

(5) Oman, T. J.; van der Donk, W. A. Follow the leader: The use of leader peptides to guide natural product biosynthesis. *Nat. Chem. Biol.* **2010**, *6* (1), 9.

(6) Hudson, G. A.; Mitchell, D. A. RiPP antibiotics: biosynthesis and engineering potential. *Curr. Opin. Microbiol.* **2018**, *45*, 61.

(7) Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; McGinnis, S.; Madden, T. L. NCBI BLAST: a better web interface. *Nucleic Acids Res.* **2008**, *36*, W5.

(8) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. *Nat. Chem. Biol.* **2017**, *13* (5), 470.

(9) Finn, R. D.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Mistry, J.; Mitchell, A. L.; Potter, S. C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; Salazar, G. A.; Tate, J.; Bateman, A. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44* (D1), D279.

(10) Haft, D. H.; Selengut, J. D.; Richter, R. A.; Harkins, D.; Basu, M. K.; Beck, E. TIGRFAMs and Genome Properties in 2013. *Nucleic Acids Res.* **2013**, *41* (D1), D387.

(11) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. Bioinformatic expansion and discovery of thiopeptide antibiotics. *J. Am. Chem. Soc.* **2018**, *140* (30), 9494.

(12) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45* (W1), W36.

(13) DiCaprio, A. J.; Firouzbakht, A.; Hudson, G. A.; Mitchell, D. A. Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin. *J. Am. Chem. Soc.* **2019**, *141* (1), 290.

(14) Mathur, H.; Fallico, V.; O'Connor, P. M.; Rea, M. C.; Cotter, P. D.; Hill, C.; Ross, R. P. Insights into the Mode of Action of the Sactibiotic Thuricin CD. *Front. Microbiol.* **2017**, *8*, 696.

(15) Flühe, L.; Marahiel, M. A. Radical S-adenosylmethionine enzyme catalyzed thioether bond formation in sactipeptide biosynthesis. *Curr. Opin. Chem. Biol.* **2013**, *17* (4), 605.

(16) Rea, M. C.; Sit, C. S.; Clayton, E.; O'Connor, P. M.; Whittal, R. M.; Zheng, J.; Vederas, J. C.; Ross, R. P.; Hill, C. Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against Clostridium difficile. *Proc. Natl. Acad. Sci. USA* **2010**, *107* (20), 9352.

(17) Babasaki, K.; Takao, T.; Shimonishi, Y.; Kurahashi, K. Subtilosin A, a new antibiotic peptide produced by Bacillus subtilis 168: isolation, structural analysis, and biogenesis. *J. Biochem.* **1985**, *98* (3), 585.

(18) Flühe, L.; Knappe, T. A.; Gattner, M. J.; Schafer, A.; Burghaus, O.; Linne, U.; Marahiel, M. A. The radical SAM enzyme AlbA catalyzes thioether bond formation in subtilosin A. *Nat. Chem. Biol.* **2012**, *8* (4), 350.

(19) Broderick, J. B.; Duffus, B. R.; Duschene, K. S.; Shepard, E. M. Radical S-adenosylmethionine enzymes. *Chem. Rev.* **2014**, *114* (8), 4229.

(20) Bauerle, M. R.; Schwalm, E. L.; Booker, S. J. Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *J. Biol. Chem.* **2015**, *290* (7), 3995.

(21) Mahanta, N.; Hudson, G. A.; Mitchell, D. A. Radical S-adenosylmethionine enzymes involved in RiPP biosynthesis. *Biochemistry* **2017**, *56* (40), 5229.

(22) Mahanta, N.; Zhang, Z.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A. Reconstitution and Substrate Specificity of the Radical S-Adenosyl-methionine Thiazole C-Methyltransferase in Thiomuracin Biosynthesis. *J. Am. Chem. Soc.* **2017**, *139* (12), 4310.

(23) Benjdia, A.; Balty, C.; Berteau, O. Radical SAM enzymes in the biosynthesis of ribosomally synthesized and post-translationally modified peptides (RiPPs). *Front. Chem.* **2017**, *5* (87), 87.

(24) Bruender, N. A.; Wilcoxen, J.; Britt, R. D.; Bandarian, V. Biochemical and spectroscopic characterization of a radical S-adenosyl-l-methionine enzyme involved in the formation of a peptide thioether cross-link. *Biochemistry* **2016**, *55* (14), 2122.

(25) Grove, T. L.; Himes, P. M.; Hwang, S.; Yumerefendi, H.; Bonanno, J. B.; Kuhlman, B.; Almo, S. C.; Bowers, A. A. Structural insights into thioether bond formation in the biosynthesis of sactipeptides. *J. Am. Chem. Soc.* **2017**, *139* (34), 11734.

(26) Benjdia, A.; Guillot, A.; Lefranc, B.; Vaudry, H.; Leprince, J.; Berteau, O. Thioether bond formation by SPASM domain radical SAM enzymes: Cα H-atom abstraction in subtilosin A biosynthesis. *Chem. Commun.* **2016**, *52* (37), 6249.

(27) Grell, T. A. J.; Goldman, P. J.; Drennan, C. L. SPASM and Twitch Domains in S-Adenosylmethionine (SAM) Radical Enzymes. *J. Biol. Chem.* **2015**, *290* (7), 3964.

(28) Barr, I.; Stich, T. A.; Gizzi, A. S.; Grove, T. L.; Bonanno, J. B.; Latham, J. A.; Chung, T.; Wilmot, C. M.; Britt, R. D.; Almo, S. C.; Klinman, J. P. X-ray and EPR characterization of the auxiliary Fe-S clusters in the radical SAM enzyme PqqE. *Biochemistry* **2018**, *57* (8), 1306.

(29) Grell, T. A. J.; Kincannon, W. M.; Bruender, N. A.; Blaesi, E. J.; Krebs, C.; Bandarian, V.; Drennan, C. L. Structural and spectroscopic analyses of the sporulation killing factor biosynthetic enzyme SkfB, a bacterial AdoMet radical sactisynthase. *J. Biol. Chem.* **2018**, ASAP.

(30) Sit, C. S.; van Belkum, M. J.; McKay, R. T.; Worobo, R. W.; Vederas, J. C. The 3D solution structure of thurincin H, a bacteriocin with four sulfur to α-carbon crosslinks. *Angew. Chem., Int. Ed.* **2011**, *50* (37), 8718.

(31) Liu, W.-T.; Yang, Y.-L.; Xu, Y.; Lamsa, A.; Haste, N. M.; Yang, J. Y.; Ng, J.; Gonzalez, D.; Ellermeier, C. D.; Straight, P. D.; Pevzner, P. A.; Pogliano, J.; Nizet, V.; Pogliano, K.; Dorrestein, P. C. Imaging mass spectrometry of intraspecies metabolic exchange revealed the cannibalistic factors of Bacillus subtilis. *Proc. Natl. Acad. Sci. USA* **2010**, *107* (37), 16286.

(32) Flühe, L.; Burghaus, O.; Wieckowski, B. M.; Giessen, T. W.; Linne, U.; Marahiel, M. A. Two [4Fe-4S] clusters containing radical SAM enzyme SkfB catalyze thioether bond formation during the maturation of the sporulation killing factor. *J. Am. Chem. Soc.* **2013**, *135* (3), 959.

(33) Duarte, A. F. S.; Ceotto-Vigoder, H.; Barrias, E. S.; Souto-Padrón, T.; Nes, I. F.; Bastos, M. Hyicin 4244, the first sactibiotic described in staphylococci, exhibits an anti-staphylococcal biofilm activity. *Int J Antimicrob. Agents* **2018**, *51* (3), 349.

(34) Haft, D. H.; Basu, M. K. Biological systems discovery in silico: radical S-adenosylmethionine protein families and their target peptides for posttranslational modification. *J. Bacteriol.* **2011**, *193* (11), 2745.

(35) Azevedo, A. C.; Bento, C. B. P.; Ruiz, J. C.; Queiroz, M. V.; Mantovani, H. C. Distribution and genetic diversity of bacteriocin gene clusters in rumen microbial genomes. *Appl. Environ. Microbiol.* **2015**, *81* (20), 7290.

(36) Letzel, A.-C.; Pidot, S. J.; Hertweck, C. Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. *BMC Genomics* **2014**, *15*, 983.

(37) Walsh, C. J.; Guinane, C. M.; Hill, C.; Ross, R. P.; O'Toole, P. W.; Cotter, P. D. In silico identification of bacteriocin gene clusters in the gastrointestinal tract, based on the Human Microbiome Project's reference genome database. *BMC Microbiol.* **2015**, *15*, 183.

(38) Skinnider, M. A.; Johnston, C. W.; Edgar, R. E.; Dejong, C. A.; Merwin, N. J.; Rees, P. N.; Magarvey, N. A. Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining. *Proc. Natl. Acad. Sci. USA* **2016**, *113* (42), E6343.

(39) Murphy, K.; O'Sullivan, O.; Rea, M. C.; Cotter, P. D.; Ross, R. P.; Hill, C. Genome mining for radical SAM protein determinants reveals multiple sactibiotic-like gene clusters. *PLoS One* **2011**, *6* (7), e20852.

(40) Nakai, T.; Ito, H.; Kobayashi, K.; Takahashi, Y.; Hori, H.; Tsubaki, M.; Tanizawa, K.; Okajima, T. The radical S-adenosyl-L-methionine enzyme QhpD catalyzes sequential formation of intra-protein sulfur-to-methylene carbon thioether bonds. *J. Biol. Chem.* **2015**, *290* (17), 11144.

(41) Finn, R. D.; Attwood, T. K.; Babbitt, P. C.; Bateman, A.; Bork, P.; Bridge, A. J.; Chang, H.-Y.; Dosztányi, Z.; El-Gebali, S.; Fraser, M.; Gough, J.; Haft, D.; Holliday, G. L.; Huang, H.;

Huang, X.; Letunic, I.; Lopez, R.; Lu, S.; Marchler-Bauer, A.; Mi, H.; Mistry, J.; Natale, D. A.; Necci, M.; Nuka, G.; Orengo, C. A.; Park, Y.; Pesseat, S.; Piovesan, D.; Potter, S. C.; Rawlings, N. D.; Redaschi, N.; Richardson, L.; Rivoire, C.; Sangrador-Vegas, A.; Sigrist, C.; Sillitoe, I.; Smithers, B.; Squizzato, S.; Sutton, G.; Thanki, N.; Thomas, P. D.; Tosatto, Silvio C E.; Wu, C. H.; Xenarios, I.; Yeh, L.-S.; Young, S.-Y.; Mitchell, A. L. InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* **2017**, *45* (D1), D190.

(42) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25* (17), 3389.

(43) Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta* **2015**, *1854* (8), 1019.

(44) Himes, P. M.; Allen, S. E.; Hwang, S.; Bowers, A. A. Production of sactipeptides in Escherichia coli: Probing the substrate promiscuity of subtilosin A biosynthesis. *ACS Chem. Biol.* **2016**, *11* (6), 1737.

(45) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. *Nat. Chem. Biol.* **2015**, *11* (8), 564.

(46) Mathur, H.; Rea, M. C.; Cotter, P. D.; Hill, C.; Ross, R. P. The sactibiotic subclass of bacteriocins: an update. *Curr. Protein Pept. Sci.* **2015**, *16* (6), 549.

(47) Shelburne, C. E.; An, F. Y.; Dholpe, V.; Ramamoorthy, A.; Lopatin, D. E.; Lantz, M. S. The spectrum of antimicrobial activity of the bacteriocin subtilosin A. *J. Antimicrob. Chemother.* **2007**, *59* (2), 297.

(48) Wang, G.; Manns, D. C.; Guron, G. K.; Churey, J. J.; Worobo, R. W. Large-Scale Purification, Characterization, and Spore Outgrowth Inhibitory Effect of Thurincin H, a Bacteriocin Produced by Bacillus thuringiensis SF361. *Probiotics Antimicrob. Proteins* **2014**, *6* (2), 105.

(49) Lee, H.; Churey, J. J.; Worobo, R. W. Biosynthesis and transcriptional analysis of thurincin H, a tandem repeated bacteriocin genetic locus, produced by Bacillus thuringiensis SF361. *FEMS Microbiol. Lett.* **2009**, *299* (2), 205.

(50) Kohl, M.; Wiese, S.; Warscheid, B. Cytoscape: software for visualization and analysis of biological networks. *Methods Mol. Biol.* **2011**, *696*, 291.

(51) Su, G.; Morris, J. H.; Demchak, B.; Bader, G. D. Biological network exploration with Cytoscape 3. *Curr. Protoc. Bioinformatics* **2014**, *47*, 8 13 1.

(52) Satoh, A.; Kim, J. K.; Miyahara, I.; Devreese, B.; Vandenberghe, I.; Hacisalihoglu, A.; Okajima, T.; Kuroda, S.; Adachi, O.; Duine, J. A.; Van Beeumen, J.; Tanizawa, K.; Hirotsu, K. Crystal structure of quinohemoprotein amine dehydrogenase from Pseudomonas putida. Identification of a novel quinone cofactor encaged by multiple thioether cross-bridges. *J. Biol. Chem.* **2002**, *277* (4), 2830.

(53) Datta, S.; Mori, Y.; Takagi, K.; Kawaguchi, K.; Chen, Z.-W.; Okajima, T.; Kuroda, S. i.; Ikeda, T.; Kano, K.; Tanizawa, K.; Mathews, F. S. Structure of a quinohemoprotein amine dehydrogenase with an uncommon redox cofactor and highly unusual crosslinking. *Proc. Natl. Acad. Sci. USA* **2001**, *98* (25), 14268.

(54) Vandenberghe, I.; Kim, J. K.; Devreese, B.; Hacisalihoglu, A.; Iwabuki, H.; Okajima, T.; Kuroda, S.; Adachi, O.; Jongejan, J. A.; Duine, J. A.; Tanizawa, K.; Van Beeumen, J. The covalent structure of the small subunit from Pseudomonas putida amine dehydrogenase reveals the presence

of three novel types of internal cross-linkages, all involving cysteine in a thioether bond. *J. Biol. Chem.* **2001**, *276* (46), 42923.

(55) Takagi, K.; Yamamoto, K.; Kano, K.; Ikeda, T. New pathway of amine oxidation respiratory chain of Paracoccus denitrificans IFO 12442. *Eur. J. Biochem.* **2001**, *268* (2), 470.

(56) Takagi, K.; Torimura, M.; Kawaguchi, K.; Kano, K.; Ikeda, T. Biochemical and Electrochemical Characterization of Quinohemoprotein Amine Dehydrogenase from Paracoccus denitrificans. *Biochemistry* **1999**, *38* (21), 6935.

(57) Adachi, O.; Kubota, T.; Hacisalihoglu, A.; Toyama, H.; Shinagawa, E.; Duine, J. A.; Matsushita, K. Characterization of Quinohemoprotein Amine Dehydrogenase from Pseudomonas putida. *Biosci. Biotechnol. Biochem.* **1998**, *62* (3), 469.

(58) Caruso, A.; Bushin, L. B.; Clark, K. A.; Martinie, R. J.; Seyedsayamdost, M. R. Radical Approach to Enzymatic β-Thioether Bond Formation. *J. Am. Chem. Soc.* **2019**, *141* (2), 990.

(59) Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; van der Donk, W. A.; Mitchell, D. A. Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. *ACS Cent. Sci.* **2017**, *3* (6), 629.

(60) Zhang, Q.; Ortega, M.; Shi, Y.; Wang, H.; Melby, J. O.; Tang, W.; Mitchell, D. A.; van der Donk, W. A. Structural investigation of ribosomally synthesized natural products by hypothetical structure enumeration and evaluation using tandem MS. *Proc. Natl. Acad. Sci. USA* **2014**, *111* (33), 12031.

(61) Wang, J.; Zhang, L.; Teng, K.; Sun, S.; Sun, Z.; Zhong, J. Cerecidins, novel lantibiotics from Bacillus cereus with potent antimicrobial activity. *Appl. Environ. Microbiol.* **2014**, *80* (8), 2633.

(62) Voller, G. H.; Krawczyk, J. M.; Pesic, A.; Krawczyk, B.; Nachtigall, J.; Süssmuth, R. D. Characterization of new class III lantibiotics--erythreapeptin, avermipeptin and griseopeptin from

Saccharopolyspora erythraea, Streptomyces avermitilis and Streptomyces griseus demonstrates stepwise N-terminal leader processing. *ChemBioChem* **2012**, *13* (8), 1174.

(63) Lohans, C. T.; Vederas, J. C. Structural characterization of thioether-bridged bacteriocins. *J. Antibiot.* **2014**, *67* (1), 23.

(64) Shenkarev, Z. O.; Finkina, E. I.; Nurmukhamedova, E. K.; Balandin, S. V.; Mineev, K. S.; Nadezhdin, K. D.; Yakimenko, Z. A.; Tagaev, A. A.; Temirov, Y. V.; Arseniev, A. S.; Ovchinnikova, T. V. Isolation, structure elucidation, and synergistic antibacterial activity of a novel two-component lantibiotic lichenicidin from Bacillus licheniformis VK21. *Biochemistry* **2010**, *49* (30), 6462.

(65) Monneau, Y. R.; Ishida, Y.; Rossi, P.; Saio, T.; Tzeng, S.-R.; Inouye, M.; Kalodimos, C. G. Exploiting E. coli auxotrophs for leucine, valine, and threonine specific methyl labeling of large proteins for NMR applications. *J. Biomol. NMR* **2016**, *65* (2), 99.

(66) Freeman, M. F.; Gurgui, C.; Helf, M. J.; Morinaka, B. I.; Uria, A. R.; Oldham, N. J.; Sahl, H. G.; Matsunaga, S.; Piel, J. Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* **2012**, *338* (6105), 387.

(67) Benjdia, A.; Guillot, A.; Ruffié, P.; Leprince, J.; Berteau, O. Post-translational modification of ribosomally synthesized peptides by a radical SAM epimerase in Bacillus subtilis. *Nat. Chem. Biol.* **2017**, *9* (7), 698.

(68) Schramma, K. R.; Bushin, L. B.; Seyedsayamdost, M. R. Structure and biosynthesis of a macrocyclic peptide containing an unprecedented lysine-to-tryptophan crosslink. *Nat. Chem.* **2015**, *7* (5), 431.

(69) Bushin, L.B.; Clark, K.A.; Pelczer, I.; Seyedsayamdost, M.R. Charting an unexplored streptococcal biosynthetic landscape reveals a unique peptide cyclization motif. *J. Am. Chem. Soc.* **2018**, *140* (50), 17674.

(70) Khaliullin, B.; Ayikpoe, R.; Tuttle, M.; Latham, J. A. Mechanistic elucidation of the mycofactocin-biosynthetic radical S-adenosylmethionine protein, MftC. *J. Biol. Chem.* **2017**, *292* (31), 13022.

(71) Khaliullin, B.; Aggarwal, P.; Bubas, M.; Eaton, G. R.; Eaton, S. S.; Latham, J. A. Mycofactocin biosynthesis: modification of the peptide MftA by the radical S-adenosylmethionine protein MftC. *FEBS Lett.* **2016**, *590* (16), 2538.

Chapter 5: Reactivity-Based Screening for Citrulline-Containing Natural Products Reveals a Family of Bacterial Peptidyl Arginine Deiminases[3]

Lonnie A. Harris, Patricia M. B. Saint-Vincent, Xiaorui Guo, Graham A. Hudson, Adam J. DiCaprio, Lingyang Zhu, and Douglas A. Mitchell[1]

5.1 Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) comprise a natural product (NP) family defined by a common biosynthetic logic, wherein a genetically encoded precursor peptide is modified by associated biosynthetic enzymes to form the mature product.[1, 2] RiPP precursor peptides are typically bipartite, composed of an N-terminal leader region responsible for enzyme recognition and a C-terminal core region which is enzymatically modified. Often, a locally encoded protease will remove the leader region prior to yielding the mature RiPP. Owing to substrate recognition being guided by motifs within the leader region, RiPP biosynthetic enzymes are often highly permissive of sequence variation in the core region, which facilitates access to analogs.[3] One RiPP class gaining increased attention is the lasso peptides, which are defined by a unique lariat topology that grants most members high levels of thermal stability and protease resistance.[4,5] These collective properties are stimulating interest in the design of lasso peptide-based therapeutics.[6, 7, 8, 9]

After ribosomal synthesis of the lasso precursor peptide, the next biosynthetic step involves RiPP recognition element (RRE) engagement of the precursor peptide, which occurs through binding to a recognition sequence within the N-terminal leader region. The RRE domain mediates recruitment of the subsequent enzymes[10, 11, 12, 13, 14] and is found as either a discretely encoded protein or fused to the leader peptidase. Upon RRE binding, the leader peptidase (an enzyme homologous to transglutaminases) removes the leader region from the precursor peptide.[15, 16, 17] The characteristic

[3] A.J.D. contributions: All NMR data acquisition and assisted in manuscript preparation. Assisted Lonnie Harris in data processing and interpretation.

lariat topology is then formed by an ATP-dependent lasso cyclase (homologous to asparagine synthetases), which installs a macrolactam linkage between the newly formed N-terminus of the core with the side chain carboxylate of a downstream Asp or Glu acceptor residue, trapping the C-terminal tail inside the ring.

As with other RiPP classes, post-translational modifications (PTMs) beyond the class-defining, threaded macrolactam are known, among them are disulfide formation (e.g. BI-32169, etc.),[18] C-terminal *O*-methylation (lassomycin),[19] Asp β-hydroxylation (canucin A),[20] Lys ε-acylation (albusnodin),[21] Trp 7-hydroxylation (RES-701-2),[22, 23] epimerization to D-Trp (MS-271),[24] Ser *O*-phosphorylation (paeninodin),[25] and Arg deimination (citrulassin A).[26] Citrulassin A was discovered along with several other new lasso peptides using the Rapid ORF Description and Evaluation Online (RODEO) automated genome-mining tool.[26] Briefly, RODEO uses profile hidden Markov models (pHMMs), heuristic scoring, and supervised machine learning to identify putative RiPP biosynthetic gene clusters (BGCs) and score potential precursors based on characterized family members. Since its initial application to lasso peptides, additional RODEO modules have been developed, which have helped catalog and classify other RiPP classes, such as the thiopeptides,[27] lanthipeptides,[28] sactipeptides, and ranthipeptides.[29]

During our previous RODEO-guided lasso peptide discovery effort, an uncharacterized family of 55 lasso peptides was bioinformatically identified (**Figure 1**).[26] The first member of this family, citrulassin A, was isolated and characterized from *Streptomyces albulus* NRRL B-3066. Citrulassin A was so named owing to an unprecedented citrulline at position nine of the core region (Cit9) rather than the genetically encoded Arg. Because Cit is a non-proteinogenic amino acid, it was hypothesized that an additional tailoring enzyme would be required for Arg conversion to Cit. One possible candidate would be a peptidyl arginine deiminase (PAD), as these enzymes carry out

post-translational deimination of Arg for a variety of biochemical functions. However, genome sequencing of *S. albulus* NRRL B-3066 revealed no enzymes with predicted PAD activity near the citrulassin A BGC (lasso cyclase NCBI accession identifier: WP_079136914.1). Moreover, heterologous expression of the citrulassin A BGC with ~20 kb 5' and 3' flanking regions in *Streptomyces lividans* resulted in *des*-citrulassin A production, which contains Arg rather than Cit (**Figure 1**). Given this result, we concluded that the gene(s) responsible for Cit formation was distally encoded.[26]

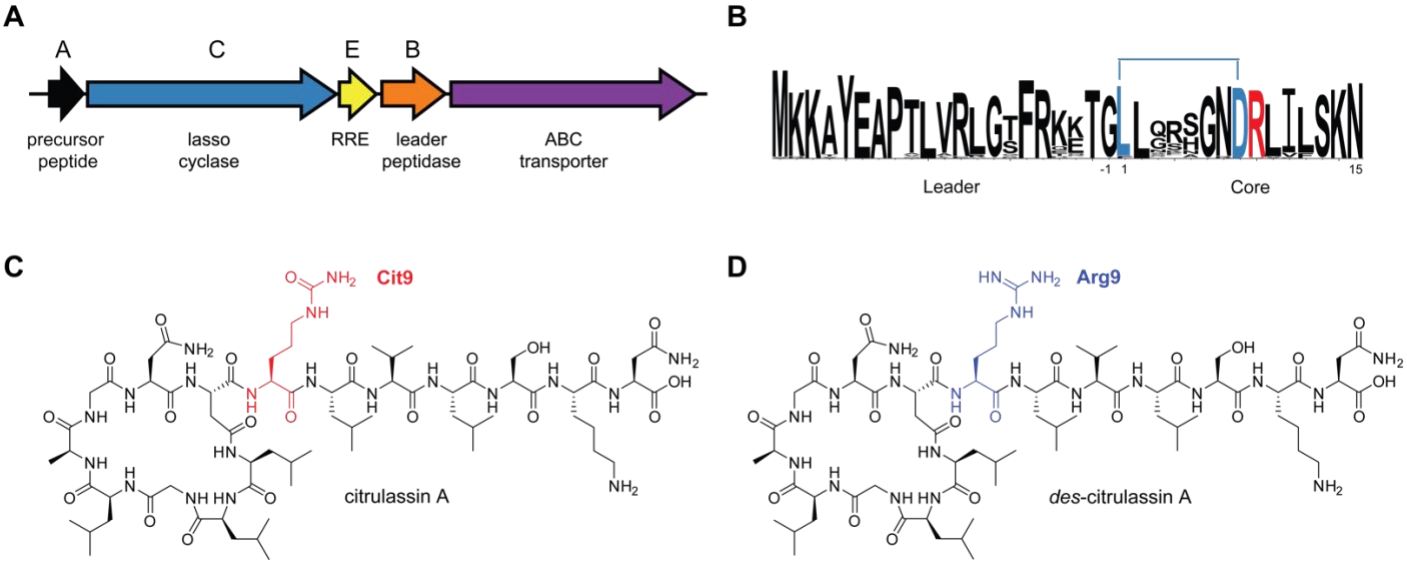


Figure 5.1. Overview of citrulassin lasso peptides. (**A**) Citrulassin biosynthetic gene cluster (BGC) architecture. (**B**) Sequence logo (http://weblogo.berkeley.edu/logo.cgi)[30] of the 181 citrulassin precursor peptides bioinformatically identified in this study. Isopeptide bond formation between the Asp side chain and N-terminal Leu of the core region is indicated with a blue bar. The Arg9 residue modified to Cit in citrulassin A is red. (**C**) Structure and molecular mass of citrulassin A from *S. albulus* with Cit9 highlighted in red. (**D**) Structure and molecular mass of of *des*-citrulassin

Figure 5.1. (cont.) A heterologously obtained from *S. lividans* with Arg9 highlighted in blue. Non-threaded, two-dimensional versions of the structures are shown for clarity.

To the best of our knowledge, only one bacterial PAD has been characterized. This protein from *Porphyromonas gingivalis* functions as a virulence factor conserved among periodontal pathogens.[31] The *P. gingivalis* PAD is known to deiminate human fibrinogen, leading to gingivitis and an increased risk of rheumatoid arthritis.[32] This PAD belongs to Protein Family[33] PF08527 (PAD_porph) and thus is distinct from characterized human PADs (PF03068). PAD4 is the best studied human isoform and known to deiminate Arg residues within histones, which affects transcription as well as cellular differentiation. Dysfunctional PAD4 activity has been implicated in rheumatoid arthritis and in certain cancers.[34]

To identify the gene(s) responsible for converting Arg to Cit during citrulassin biosynthesis, a comparative genomics approach was employed on Arg- versus Cit-containing citrulassin producers. Arg deimination is easily overlooked by routine mass spectrometry (MS) given that the PTM results in a small mass deviation (replacement of NH with O, +0.98 Da). Moreover, the commonly observed modification of Asn/Gln to Asp/Glu is isobaric with Arg deimination, further complicating matters.[35] We previously described strategies for the rapid discovery of NPs bearing specific organic functional groups in the extracts of cultured bacteria, termed reactivity-based screening (RBS, **Figure S1**).[36] RBS involves the chemoselective labeling of an organic functional group present within a NP to enable rapid detection via comparative MS of reacted and unreacted samples. RBS has facilitated the discovery and characterization of RiPPs, non-ribosomal peptides,[37] polyketides,[38,39] and hybrid compounds.[40] RBS is therefore a valuable tool for bridging the theoretical, bioinformatics-guided identification of NPs with experimental detection of the

bona fide compound(s). RBS-based discovery strategies are agnostic to the biological activity of the NP, which allow the NP hunter to focus on chemical novelty by through the rapid dereplication of already-characterized NPs[36] and detection of low-abundance species using sensitive MS-based techniques such as matrix-assisted laser/desorption ionization time-of-flight MS (MALDI-TOF-MS). Further, RBS is compatible with variable probe chemistries that enhance detection due to unique isotopic patterns of labeled compounds or selective enrichment through affinity purification.[37, 41, 42]

Herein, we describe the expansion of RBS as a tool to aid in identifying the PAD responsible for the Arg to Cit conversion required to form a Cit-containing lasso peptide. Repurposing work from Thompson,[43, 35] a brominated phenylglyoxal probe was used to specifically label primary ureido groups over primary guanidino groups. This probe was used to survey for the presence of Cit in a subset of lasso peptides and enable partial phylogenetic profiling, thus revealing a PAD gene putatively responsible for each citrulassin producer. Using functional expression of the PAD from *Streptomyces glaucescens*, we demonstrated the deimination of a citrulassin precursor peptide in vivo, providing a gene-to-function link for RiPP citrulline formation. A broader bioinformatic survey of bacterial PADs was also performed, which revealed a wide taxonomic breadth and diversity of genomic contexts for these understudied enzymes.

5.2 Results and Discussion

5.2.1 Validation of 3-bromophenylglyoxal as a selective probe for primary ureido groups

Previous work has demonstrated selective labeling of Cit in eukaryotic cellular extracts at low pH with phenylglyoxal probes (**Figure 2**).[35,43] Under mildly acidic conditions, guanidino groups are protonated and thus unreactive towards the glyoxal moiety; however, ureido groups remain nucleophilic and thus react rapidly.[43] The reaction of Cit with phenylglyoxal has been further

elaborated to readily identify labeled species in complex samples by employing brominated phenylglyoxal derivatives. Natural abundance bromine provides a distinct isotopic pattern (~1:1 $^{79}Br/^{81}Br$) while the UV laser-absorbing arene increases signal intensity in MALDI-TOF-MS.[44] Thus, we imagined that 3-bromophenylglyoxal (**1**) could selectively target primary ureido group-containing NPs within bacterial extracts. As predicted from the known reactivity of glyoxal, the free amino acid L-Cit was robustly modified by **1** while L-Arg was unmodified under identical conditions (**Figure S2**). To evaluate probe suitability within a bacterial extract, *S. albulus* NRRL B-3066 (the native producer of citrulassin A) and *Streptomyces lividans* 3H4 (a heterologous producer of *des*-citrulassin A)[37] were grown for 7 d prior to extraction of metabolites with methanol. When treated with **1** under identical reaction conditions, citrulassin A was nearly fully labeled while no product was observed in the *des*-citrulassin A sample (**Figure 2**). The non-ribosomal peptide *deimino*-antipain, another Cit-containing NP from *S. albulus* NRRL B-3066, was also labeled under these conditions (**Figure S3**).[37] Extract from *S. lividans* heterologously producing antipain confirmed selectivity for Cit over Arg, with antipain displaying negligible labeling by **1**. Taken together, these experiments validate **1** as useful for the rapid identification of primary ureido group-containing NPs.

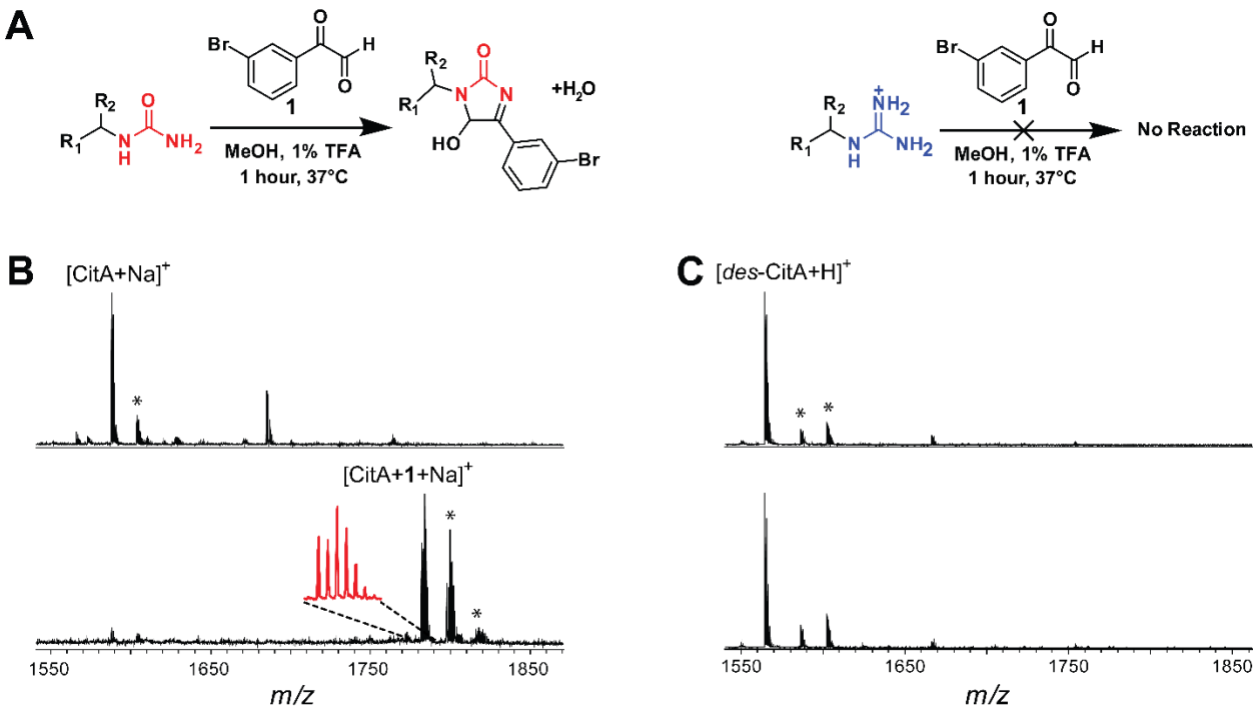


Figure 5.2. Validation of 3-bromophenylglyoxal (**1**) selectivity. (**A**) Reaction conditions that selectively label citrulassin A (*left*) over *des*-citrulassin A (*right*). TFA = trifluoroacetic acid. (**B**) MALDI-TOF mass spectra of a methanolic extract of *Streptomyces albulus* (citrulassin A producer) prior to addition of **1** (*top*) and after reaction with **1** (*bottom*). Magnified inset shows the ^{79}Br:^{81}Br isotopic pattern. Peaks indicated with an asterisk are (from *left* to *right*): $[CitA+K]^+$, $[CitA+\mathbf{1}+K]^+$, and $[CitA+\mathbf{1}+K+H_2O]^+$. (**C**) MALDI-TOF mass spectra of a methanolic extract of *des*-citrulassin A heterologously expressed in *Streptomyces lividans* prior to addition of **1** (*top*) and after reaction with **1** (*bottom*).[26] Peaks indicated with an asterisk are $[des\text{-}CitA+Na]^+$ and $[des\text{-}CitA+K]^+$.

5.2.2 3-bromophenylglyoxal (1) screening of citrulassin producers

With **1** validated, we next sought to identify which members of the citrulassin family of lasso peptides contain Cit. Previous screening of bacterial extracts suggested that not all citrulassins bear

a Cit residue (the Arg-containing counterpart is referred to as *des*-citrulassin),[26] but owing to a small mass difference between guanidino- versus ureido-containing products, we tested all putative citrulassin producers through reactivity with **1**. To identify probable citrulassin producers, a BLASTP search[45] was performed using the citrulassin A lasso cyclase as the query. Proteins with an expectation value below 10^{-200} were then subjected to RODEO analysis. The identified citrulassin precursor peptides were well conserved, with only a few positions in the ring and loop regions showing sequence variation (**Figure 1** and **Supplemental Dataset 1**). In addition, co-occurrence analysis indicated that no putative PADs were encoded in the neighborhood of the citrulassin BGCs, in accord with the earlier heterologous expression study (**Table S2**).[26] With this list in hand, a selection of 109 strains predicted to contain citrulassin BGCs were obtained from the Agriculture Research Service (ARS) Culture Collection and grown for 10 d on a variety of media at 30 ˚C. The cells were then harvested, extracted with methanol, and analyzed by MALDI-TOF-MS. Samples possessing a mass within the expected molecular weight range for a citrulassin-type lasso peptide (n = 14) were screened for Arg deimination using reactivity towards **1** (**Table 1**). Extracts from eight of these strains underwent labeling and were subjected to high-resolution and tandem mass spectrometry (HR-MS/MS), which confirmed the molecular formula and location of Cit (**Figure S4**). Several citrulassins identified contain Arg within the ring, typically at position four, in addition to the conserved Arg9 encoded within the loop region. Cit formation was confined to core position nine for all citrulassins with one exception: citrulassin K from *Streptomyces* sp. NRRL S-920 contained a second site of deimination within the ring (at Arg3 or Arg4). The extent of deimination was accurately reflected by reactivity towards **1**, with all singly

Table 5.1. Characterized citrulassins

Bacterial strain (NRRL)	Core Sequence	Citrulassin	Deimination	PAD
Streptomyces albulus B-3066	LLGLAGND**R**LVLSKN	A	Y	Y
Streptomyces aurantiacus B-2806	LLQRHGND**R**LIFSKN	B	Y	Y
Streptomyces tricolor B-16925	LLQRSGNDRLILSKN	C	N	N
Streptomyces katrae B-16271	LLGRSGNDRLVLSKN	D	N	N
Streptomyces glaucescens B-11408	LLQRHGND**R**LILSKN	E	Y	Y
Streptomyces avermitilis B-16169	LLGRSGNDRLILSKN	F	N	N
Streptomyces torulosus S-189	LLGRSGND**R**LILSKN	F	Y	Y
Streptomyces auratus 8097	LLNSSGNDRLILSKN	G	N	N
Streptomyces sp. S-118	LLAFHGND**R**LILSKN	H	Y	Y
Streptomyces sp. F-5140	LLGRHGND**R**LILSKN	I	Y	Y
Streptomyces natalensis B-5314	LLEFRGNDRLILSKN	J	N	N
Streptomyces sp. S-920	LL<u>RR</u>SGND**R**LILSKN*	K	Y	Y
Streptomyces sp. S-481	LLGDAGNDRLILSKN	L	N	Y
Streptomyces pharetrae B-24333	LLQRNGND**R**LILSKN	M	Y	Y

Confirmed deiminated Arg (i.e. positions containing Cit) are shown in red.

*Citrulassin K contained a second Cit at Arg3 or Arg4 (underlined).

deiminated citrulassins showing exactly one labeling event, while the doubly deiminated citrulassin K was labeled twice (**Figure S4**). Thus, RBS is useful for determining the equivalents of targetable organic functional groups within a compound of interest.

5.2.3 Characterization of citrulassin and des-citrulassin F

During our screen for citrulassin producers, it was noted that two sequenced bacteria encoded the same citrulassin core sequence; however, one strain (*Streptomyces avermitilis* NRRL B-16169), exclusively produced *des*-citrulassin F, while the other (*Streptomyces torulosus* NRRL S-189) exclusively produced citrulassin F. No solution structure for any citrulassin family member has

been reported, thus we set out to determine the solution structure of *des*-citrulassin F as this compound was produced at a relatively high level in *S. avermitilis*. HR-MS/MS of the two HPLC-purified lasso peptides confirmed primary structures consistent with the RODEO-predicted core regions (**Figure S4**). MS/MS fragmentation analysis also confirmed that only Arg9 was converted to Cit in citrulassin F. Two-dimensional, homonuclear NMR data (^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY) were collected for *des*-citrulassin F and used to assign chemical shifts of individual residues (**Table S3**, **Figures S5-S6**). Strong NOE correlations between the Leu1-H_N and the Asp8 β protons confirmed the isopeptide linkage while NOE correlations between Gly3-Leu12 and Asp8-Leu12 suggested Leu12 as the most probable steric-locking residue. (**Figure S6**). The ^{1}H-NMR spectrum of *des*-citrulassin F taken in D_2O revealed that several backbone amides were resistant to deuterium exchange, specifically Leu1, Asp8, Leu10, Ile11, and Leu12 (**Figure S7**). Recalcitrance of amide NH groups to deuterium exchange is consistent with a lasso topology and further supports the assignment of Leu12 as the steric-locking residue.[46] Distance constraints obtained from the NOESY data were used to calculate an ensemble structure (**Figure 3** and **S8**)[47]. The resulting ensemble formed the expected right-handed lasso topology, with the C-terminal tail passing through the macrocycle between Ile11 and Leu12. Analysis of the solvent-accessible surface of the lowest energy structure revealed that the topological surface area of side chains of Ile11 and Leu12 are larger than the inner diameter of the ring and thus are competent to serve as steric-locking residues. After determining the solution structure of *des*-citrulassin F, the antibacterial activity of *des*-citrulassin F and citrulassin F were assessed against a small panel of Gram-positive and Gram-negative strains. Neither lasso peptide showed activity up to 500 μM, consistent with published assays performed using citrulassin A.[26]

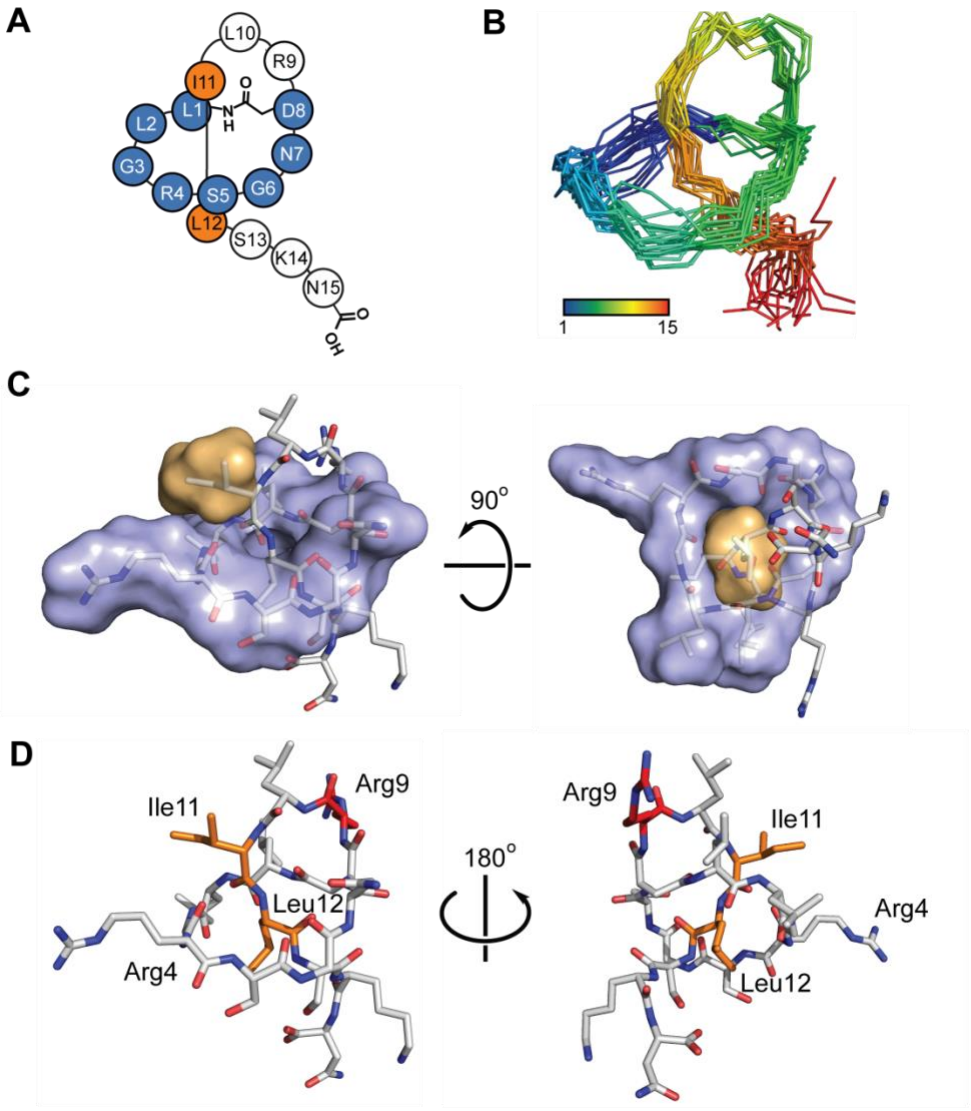


Figure 5.3. Structure of *des*-citrulassin F. (**A**) Topological diagram of *des*-citrulassin F. Blue, ring region; orange, steric-locking residues. (**B**) NOE-based NMR ensemble of the 20 lowest energy structures of *des*-citrulassin F (amide backbone is shown for clarity). (**C**) Surface-filling mode of *des*-citrulassin F, with the ring and steric-locking residues colored blue and orange, respectively. (**D**) Lowest energy structure of *des*-citrulassin F. Arg4 and Arg9 are labeled, with the latter being modified to Cit in citrulassin F (red). Ile11 and Leu12 are highlighted in orange.

5.2.4 Partial phylogenetic profiling identifies a PAD for citrulline formation

The fact that two distinct bacterial strains (*S. avermitilis* and *S. torulosus*) produced lasso peptides differing only in the state of Arg9 deimination was unique among the citrulassin producers screened. This scenario presented an opportunity to identify genomic differences between the organisms to identify the gene(s) responsible for Cit9 formation. Using the profiling algorithm PhyloProfile,[48, 49] a search was conducted for genes present in the *S. albulus* genome (7,734 predicted genes) and in four other citrulassin producers (**Table 1**, NRRL identifiers: S-189, S-118, B-11408, and B-24333) but absent from four *des*-citrulassin producers (B-16925, B-16271, B-16169, and 8097). This analysis returned 32 *S. albulus* genes that correlated with the Cit phenotype, one of which encoded a hypothetical protein containing a PAD domain (WP_064069847.1, **Supplemental Dataset 1**). This gene encodes a member of protein family PF03068 and thus is homologous to human PADs. No significant sequence homology is shared between the putative *S. albulus* PAD and the only other characterized bacterial PAD from *P. gingivalis* (WP_005873463.1, **Table S4**).[33,31] Position-specific iterative BLAST (PSI-BLAST)[50] of citrulassin producers confirmed the correlation between the genome of a bacterial strain encoding a PAD and the presence of Cit in the mature citrulassin, with only one strain producing *des*-citrulassin while encoding a PAD in the genome (*Streptomyces* sp. S-481, *des*-citrulassin L, **Table S1** and **Figure S4**). Analogously, all organisms that lacked a PAD were associated with *des*-citrulassin productions. However, the small sample size prevented the establishment of a statistically significant correlation between genotype and phenotype.[51] Nonetheless, having identified a candidate PAD, we sought to confirm the role of this PAD in citrulassin biosynthesis.

5.2.5 PAD complementation in a des-citrulassin producer results in citrulassin production

The co-occurrence of PAD with the production of citrulassin provided circumstantial support for the proposed role in Arg deimination. To further evaluate this putative biochemical role, a functional expression experiment was designed to introduce the PAD-encoding gene from a verified citrulassin producer to the genome of a verified *des*-citrulassin producer that lacked a PAD-encoding gene. Using the *Streptomyces* phage φC31 integrase, plasmids containing an *attP* locus can be inserted into the chromosome of a recipient organism at the corresponding *attB* locus. Nearly all identified *des*-citrulassin producer genomes contain a φC31 *attB*, enabling *Escherichia coli* to *Streptomyces* conjugative transfer.[52,53] The integrative plasmid placed the PAD under strong constitutive promotion (*ermE*p*) to facilitate a higher level of expression within the heterologous host.[53,54] Owing to a consistently higher yield of the citrulassin from *Streptomyces glaucescens*, we chose to integrate the I. *glaucescens* PAD into *Streptomyces katrae* NRRL B-16271, a native producer of *des*-citrulassin D (**Figure 4**). Exconjugants were isolated by iterative cultivation on a selective medium and chromosomal insertion of the PAD gene was confirmed via DNA sequencing (**Figure S9**). Successfully integrated strains were then grown on a medium that elicited *des*-citrulassin D production in wild-type *S. katrae*. Metabolites were extracted with MeOH and subsequently analyzed via MALDI-TOF-MS. PAD integration in *S. katrae* resulted in the production of an ion consistent with Cit-containing citrulassin D (**Figure 4**). Reaction with **1** resulted in complete labeling, while wild-type *des*-citrulassin D was not labeled, suggesting that PAD integration had resulted in the conversion of one Arg within *des*-citrulassin D to Cit. HR-MS/MS confirmed that the produced lasso peptide was 0.98 Da heavier than expected for *des*-citrulassin D and elicited a fragmentation pattern consistent with the sequence of citrulassin D, featuring Arg9 converted to Cit. Notably, the site-specific deimination of Arg9 of *des*-citrulassin

D over Arg4 confirms our previously assigned site of deimination on citrulassin and demonstrates that expression of PAD is necessary and sufficient for conversion of Arg to Cit in the context of citrulassin biosynthesis. Future work will be necessary to uncover the origins of the selectivity and the timing of deimination during citrulassin biosynthesis.

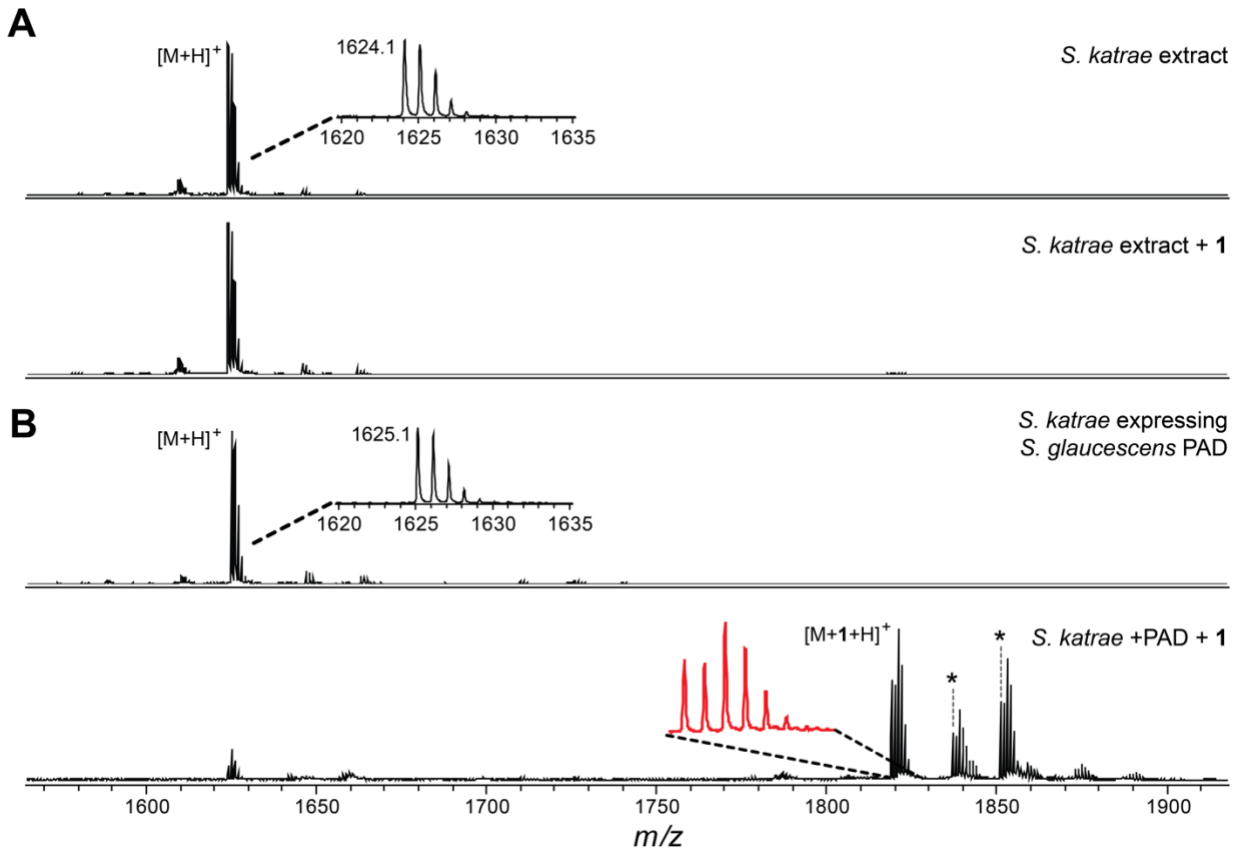


Figure 5.4. Citrulassin production in a *des*-citrulassin producer by PAD functional expression (**A**) MALDI-TOF mass spectrum of *Streptomyces katrae* extract containing *des*-citrulassin D **Figure 5.4 (cont.)** unreacted (*top*) and reacted (*bottom*) with **1**. (**B**) MALDI-TOF mass spectrum of *S. katrae* extract expressing the PAD from *Streptomyces glaucescens*. Shown are samples unreacted (*top*) and reacted (*bottom*) with **1**. Insets show zoomed regions to visualize the isotopic

Figure 5.4 (cont.) ratios. Asterisks denote solvated adducts corresponding to $[M+ \mathbf{1}+H_2O+H]^+$ (*left*) and $[M+\mathbf{1}+MeOH+H]^+$ (*right*).

5.2.6 Bioinformatic survey of bacterial members of PF03068

Having identified the PAD responsible for citrulassin deimination, we next surveyed bacterial strains encoding members of protein family PF03068. An hmmscan-based comparison of the *S. albulus* PAD against the pHMMs that define PF03068 and PF04371 confirmed that the *S. albulus* PAD is a member of PF03068 (**Table S4**). Moreover, an alignment of the *S. albulus* PAD with human PAD4, the best-characterized isoform,[55] shows conservation of functionally important residues, such as the catalytic Cys645, stabilizing triad of Asp350-His471-Asp473, as well as numerous Ca^{2+}-binding residues (**Figure S10**).[55] In contrast, the *P. gingivalis* (PAD_porph) is calcium-independent and cannot be aligned with the *S. albulus* PAD outside of short motifs near the active site.[56] Indeed, few residues are conserved between members of PF03068 and PF04371 beyond those common to all members of the guanidino-modifying enzyme superfamily.[57,58]

The *S. albulus* PAD was used to generate a dataset of 837 bacterial PF03068 PADs using a PSI-BLAST against all bacterial genomes in the NCBI non-redundant database (**Supplemental Dataset 1**).[50] Protein accession identifiers were compiled, analyzed by RODEO,[26] and then used to generate a sequence similarity network (SSN) and a maximum-likelihood phylogenetic tree to visualize the distribution of bacterial PADs belonging to PF03068 (**Figures 4** and **S11**).[59] Homologs of the *S. albulus* PAD are present in diverse bacterial phyla, but are predominantly in Actinobacteria, Cyanobacteria, and Proteobacteria, with very few homologs in Bacteroidetes and Firmicutes. Genomes that encode a member of PF03068 within these genera is fairly limited, as illustrated by only 1.6%, 0.1%, and 3.9% of sequenced Actinobacteria, Proteobacteria, and

Cyanobacteria, respectively, featuring an annotated homolog. The highest occurrence of annotated PADs is found within the Riflebacteria, which despite having only 11 genomes in the NCBI database currently, includes 15 annotated PADs. With a few notable examples, there is little evidence of horizontal gene transfer with PF03068, as proteins tend to cluster with other PADs from species within the same phylum (**Figures 5** and **S11**). Additionally, the %GC content of the genes encoding the PAD correlate strongly with the overall %GC content of the genome (**Figure S12**).[60] These data indicate bacterial members of PF03068 are not disseminated across bacterial genomes due to recent horizontal gene transfer, although we cannot rule out the possibility that the genes were acquired by organisms having similar GC content. Notably, several PAD proteins encoded by Proteobacteria are more closely related to homologs from Actinobacteria, which includes those implicated in citrulassin biosynthesis. Intriguingly, most of these PAD-containing bacteria are uncultured strains detected from aquatic and sediment metagenomic studies, raising questions concerning the genomic origin and ecology of PADs.

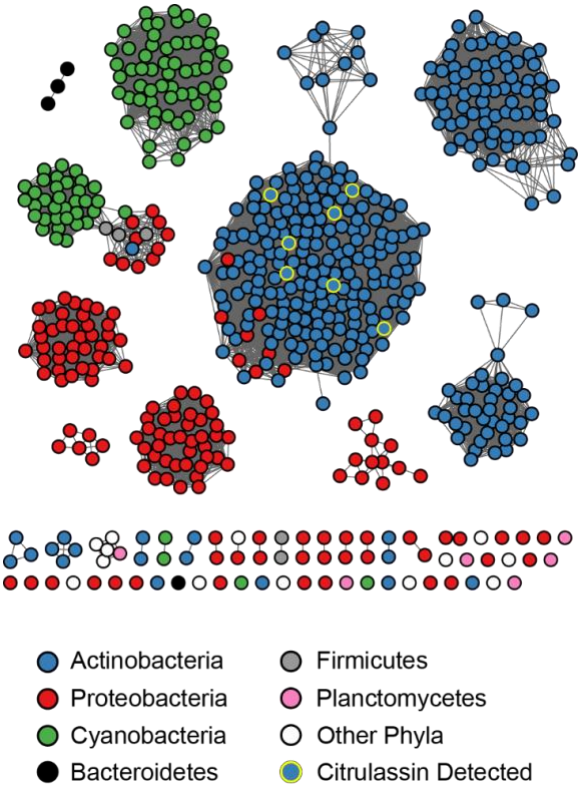


Figure 5.5. Protein sequence similarity network (SSN) of bacterial PADs. The SSN was generating using EFI-EST[61], Sequences sharing 100% identity are conflated as a single node, and connected nodes indicate an alignment score of 125, see methods) and visualized by Cytoscape.[62] Nodes are colored by phylum and detection of citrulassin (deiminated product) within the same strain.

Co-occurrence analysis of citrulassin BGCs (± 8 protein-coding sequences) shows little correlation beyond genes encoding lasso peptide biosynthetic proteins and furthermore demonstrates that PADs do not appear within lasso peptide biosynthetic operons (**Table S7**). Even PADs, which exhibit a higher degree of similarity, appear in diverse genome neighborhoods that do not bear resemblance to RiPP BGCs (**Figure S13**). Thus, these PADs are likely to exert their activity on distally encoded substrate(s).

5.3 Conclusion

Advances in genome sequencing and bioinformatics have prompted a renaissance in NP discovery. The versatility of genome-guided NP discovery is strongly complemented by reactivity-based screening (RBS) which exploits chemoselective targeting of organic functional groups to facilitate metabolite identification and isolation. In this work, we expanded the RBS toolkit to include primary ureido groups (i.e. citrulline) through their reactivity towards glyoxal-based probes, which, under mildly acidic conditions, are largely inert towards guanidino groups (i.e. Arg). Harnessing this reactivity led to the discovery of 13 new citrulassins and provided a means to perform comparative genomic analysis to discover the enzyme responsible for the previously observed Arg deimination in citrulassin biosynthesis. Our data implicated bacterial members of a protein family previously only described in eukaryotes for this activity. Members of this PAD family are unrelated to the PAD described from *P. gingivalis*. We confirmed the Arg deimination activity of this PAD by converting a *des*-citrulassin producer into a citrulassin producer through heterologous PAD expression. This work highlights the synergy between genome- and reactivity-guided discovery approaches and provides a platform for identifying additional bacteria that produce NPs featuring the rare, non-proteinogenic primary ureido functional group. More broadly, we expect this workflow to be amenable towards uncovering both the prevalence and enzymatic origins of other reactive moieties present in RiPPs and other NPs.

5.4 References

(1) G. Arnison, P.; J. Bibb, M.; Bierbaum, G.; A. Bowers, A.; S. Bugni, T.; Bulaj, G.; A. Camarero, J.; J. Campopiano, D.; L. Challis, G.; Clardy, J.; D. Cotter, P.; J. Craik, D.; Dawson, M.; Dittmann, E.; Donadio, S.; C. Dorrestein, P.; Entian, K.-D.; A. Fischbach, M.; S. Garavelli, J.; Göransson, U.; W. Gruber, C.; H. Haft, D.; K. Hemscheidt, T.; Hertweck, C.; Hill, C.; R. Horswill,

A.; Jaspars, M.; L. Kelly, W.; P. Klinman, J.; P. Kuipers, O.; James Link, A.; Liu, W.; A. Marahiel, M.; A. Mitchell, D.; N. Moll, G.; S. Moore, B.; Müller, R.; K. Nair, S.; F. Nes, I.; E. Norris, G.; M. Olivera, B.; Onaka, H.; L. Patchett, M.; Piel, J.; T. Reaney, M. J.; Rebuffat, S.; Paul Ross, R.; Sahl, H.-G.; W. Schmidt, E.; E. Selsted, M.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; D. Süssmuth, R.; R. Tagg, J.; Tang, G.-L.; W. Truman, A.; C. Vederas, J.; T. Walsh, C.; D. Walton, J.; C. Wenzel, S.; M. Willey, J.; Donk, W. A. van der. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(2) Ortega, M. A.; van der Donk, W. A. New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products. *Cell Chem. Biol.* **2016**, *23* (1), 31–44. https://doi.org/10.1016/j.chembiol.2015.11.012.

(3) Hudson, G. A.; Mitchell, D. A. RiPP Antibiotics: Biosynthesis and Engineering Potential. *Curr. Opin. Microbiol.* **2018**, *45*, 61–69. https://doi.org/10.1016/j.mib.2018.02.010.

(4) Hegemann, J. D.; Zimmermann, M.; Xie, X.; Marahiel, M. A. Lasso Peptides: An Intriguing Class of Bacterial Natural Products. *Acc. Chem. Res.* **2015**, *48* (7), 1909–1919. https://doi.org/10.1021/acs.accounts.5b00156.

(5) Hegemann, J. D. Factors Governing the Thermal Stability of Lasso Peptides. *ChemBioChem* **2020**, *21* (1–2), 7–18. https://doi.org/10.1002/cbic.201900364.

(6) Pavlova, O.; Mukhopadhyay, J.; Sineva, E.; Ebright, R. H.; Severinov, K. Systematic Structure-Activity Analysis of Microcin J25. *J. Biol. Chem.* **2008**, *283* (37), 25589–25595. https://doi.org/10.1074/jbc.M803995200.

(7) Knappe, T. A.; Manzenrieder, F.; Mas-Moruno, C.; Linne, U.; Sasse, F.; Kessler, H.; Xie, X.; Marahiel, M. A. Introducing Lasso Peptides as Molecular Scaffolds for Drug Design: Engineering of an Integrin Antagonist. *Angew. Chem. Int. Ed.* **2011**, *50* (37), 8714–8717. https://doi.org/10.1002/anie.201102190.

(8) Al Toma, R. S.; Kuthning, A.; Exner, M. P.; Denisiuk, A.; Ziegler, J.; Budisa, N.; Süssmuth, R. D. Site-Directed and Global Incorporation of Orthogonal and Isostructural Noncanonical Amino Acids into the Ribosomal Lasso Peptide Capistruin. *ChemBioChem* **2015**, *16* (3), 503–509. https://doi.org/10.1002/cbic.201402558.

(9) J. Piscotta, F.; M. Tharp, J.; R. Liu, W.; James Link, A. Expanding the Chemical Diversity of Lasso Peptide MccJ25 with Genetically Encoded Noncanonical Amino Acids. *Chem. Commun.* **2015**, *51* (2), 409–412. https://doi.org/10.1039/C4CC07778D.

(10) Burkhart, B. J.; Hudson, G. A.; Dunbar, K. L.; Mitchell, D. A. A Prevalent Peptide-Binding Domain Guides Ribosomal Natural Product Biosynthesis. *Nat. Chem. Biol.* **2015**, *11* (8), 564–570. https://doi.org/10.1038/nchembio.1856.

(11) Zhu, S.; Fage, C. D.; Hegemann, J. D.; Mielcarek, A.; Yan, D.; Linne, U.; Marahiel, M. A. The B1 Protein Guides the Biosynthesis of a Lasso Peptide. *Sci. Rep.* **2016**, *6* (1), 1–12. https://doi.org/10.1038/srep35604.

(12) Chekan, J. R.; Ongpipattanakul, C.; Nair, S. K. Steric Complementarity Directs Sequence Promiscuous Leader Binding in RiPP Biosynthesis. *Proc. Natl. Acad. Sci.* **2019**, *116* (48), 24049–24055. https://doi.org/10.1073/pnas.1908364116.

(13) Hegemann, J. D.; Schwalen, C. J.; Mitchell, D. A.; Donk, W. A. van der. Elucidation of the Roles of Conserved Residues in the Biosynthesis of the Lasso Peptide Paeninodin. *Chem. Commun.* **2018**, *54* (65), 9007–9010. https://doi.org/10.1039/C8CC04411B.

(14) Sumida, T.; Dubiley, S.; Wilcox, B.; Severinov, K.; Tagami, S. Structural Basis of Leader Peptide Recognition in Lasso Peptide Biosynthesis Pathway. *ACS Chem. Biol.* **2019**, *14* (7), 1619–1627. https://doi.org/10.1021/acschembio.9b00348.

(15) Yan, K.-P.; Li, Y.; Zirah, S.; Goulard, C.; Knappe, T. A.; Marahiel, M. A.; Rebuffat, S. Dissecting the Maturation Steps of the Lasso Peptide Microcin J25 in Vitro. *ChemBioChem* **2012**, *13* (7), 1046–1052. https://doi.org/10.1002/cbic.201200016.

(16) DiCaprio, A. J.; Firouzbakht, A.; Hudson, G. A.; Mitchell, D. A. Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin. *J. Am. Chem. Soc.* **2019**, *141* (1), 290–297. https://doi.org/10.1021/jacs.8b09928.

(17) Koos, J. D.; Link, A. J. Heterologous and in Vitro Reconstitution of Fuscanodin, a Lasso Peptide from Thermobifida Fusca. *J. Am. Chem. Soc.* **2019**, *141* (2), 928–935. https://doi.org/10.1021/jacs.8b10724.

(18) Knappe, T. A.; Linne, U.; Xie, X.; Marahiel, M. A. The Glucagon Receptor Antagonist BI-32169 Constitutes a New Class of Lasso Peptides. *FEBS Lett.* **2010**, *584* (4), 785–789. https://doi.org/10.1016/j.febslet.2009.12.046.

(19) Gavrish, E.; Sit, C. S.; Cao, S.; Kandror, O.; Spoering, A.; Peoples, A.; Ling, L.; Fetterman, A.; Hughes, D.; Bissell, A.; Torrey, H.; Akopian, T.; Mueller, A.; Epstein, S.; Goldberg, A.; Clardy, J.; Lewis, K. Lassomycin, a Ribosomally Synthesized Cyclic Peptide, Kills Mycobacterium Tuberculosis by Targeting the ATP-Dependent Protease ClpC1P1P2. *Chem. Biol.* **2014**, *21* (4), 509–518. https://doi.org/10.1016/j.chembiol.2014.01.014.

(20) Xu, F.; Wu, Y.; Zhang, C.; Davis, K. M.; Moon, K.; Bushin, L. B.; Seyedsayamdost, M. R. A Genetics-Free Method for High-Throughput Discovery of Cryptic Microbial Metabolites. *Nat. Chem. Biol.* **2019**, *15* (2), 161–168. https://doi.org/10.1038/s41589-018-0193-2.

(21) Zong, C.; Ling Cheung-Lee, W.; E. Elashal, H.; Raj, M.; James Link, A. Albusnodin: An Acetylated Lasso Peptide from Streptomyces Albus. *Chem. Commun.* **2018**, *54* (11), 1339–1342. https://doi.org/10.1039/C7CC08620B.

(22) Ogawa, T.; Ochiai, K.; Tanaka, T.; Tsukuda, E.; Chiba, S.; Yano, K.; Yamasaki, M.; Yoshida, M.; Matsuda, Y. RES-701-2, -3 and -4, Novel and Selective Endothelin Type B Receptor Antagonists Produced by Streptomyces Sp. I. Taxonomy of Producing Strains, Fermentation, Isolation, and Biochemical Properties. *J. Antibiot. (Tokyo)* **1995**, *48* (11), 1213–1220. https://doi.org/10.7164/antibiotics.48.1213.

(23) Oves-Costales, D.; Sánchez-Hidalgo, M.; Martín, J.; Genilloud, O. Identification, Cloning and Heterologous Expression of the Gene Cluster Directing RES-701-3, -4 Lasso Peptides Biosynthesis from a Marine Streptomyces Strain. *Mar. Drugs* **2020**, *18* (5), 238. https://doi.org/10.3390/md18050238.

(24) Feng, Z.; Ogasawara, Y.; Nomura, S.; Dairi, T. Biosynthetic Gene Cluster of a D-Tryptophan-Containing Lasso Peptide, MS-271. *ChemBioChem* **2018**, *19* (19), 2045–2048. https://doi.org/10.1002/cbic.201800315.

(25) Zhu, S.; Hegemann, J. D.; Fage, C. D.; Zimmermann, M.; Xie, X.; Linne, U.; Marahiel, M. A. Insights into the Unique Phosphorylation of the Lasso Peptide Paeninodin. *J. Biol. Chem.* **2016**, *291* (26), 13662–13678. https://doi.org/10.1074/jbc.M116.722108.

(26) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A New Genome-Mining Tool Redefines the Lasso Peptide Biosynthetic Landscape. *Nat. Chem. Biol.* **2017**, *13* (5), 470–478. https://doi.org/10.1038/nchembio.2319.

(27) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. Bioinformatic Expansion and Discovery of Thiopeptide Antibiotics. *J. Am. Chem. Soc.* **2018**, *140* (30), 9494–9501. https://doi.org/10.1021/jacs.8b03896.

(28) Walker, M. C.; Mitchell, D. A.; van der Donk, W. A. Precursor Peptide-Targeted Mining of More than One Hundred Thousand Genomes Expands the Lanthipeptide Natural Product Family. *bioRxiv* **2020**, 2020.03.13.990614. https://doi.org/10.1101/2020.03.13.990614.

(29) Hudson, G. A.; Burkhart, B. J.; DiCaprio, A. J.; Schwalen, C. J.; Kille, B.; Pogorelov, T. V.; Mitchell, D. A. Bioinformatic Mapping of Radical S-Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. *J. Am. Chem. Soc.* **2019**, *141* (20), 8228–8238. https://doi.org/10.1021/jacs.9b01519.

(30) Crooks, G. E.; Hon, G.; Chandonia, J.-M.; Brenner, S. E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14* (6), 1188–1190. https://doi.org/10.1101/gr.849004.

(31) Gabarrini, G.; Smit, M. de; Westra, J.; Brouwer, E.; Vissink, A.; Zhou, K.; Rossen, J. W. A.; Stobernack, T.; Dijl, J. M. van; Winkelhoff, A. J. van. The Peptidylarginine Deiminase Gene Is a Conserved Feature of Porphyromonas Gingivalis. *Sci. Rep.* **2015**, *5* (1), 1–8. https://doi.org/10.1038/srep13936.

(32) Nielen, M. M. J.; Schaardenburg, D. van; Reesink, H. W.; Stadt, R. J. van de; Horst-Bruinsma, I. E. van der; Koning, M. H. M. T. de; Habibuw, M. R.; Vandenbroucke, J. P.; Dijkmans, B. A. C. Specific Autoantibodies Precede the Symptoms of Rheumatoid Arthritis: A Study of Serial Measurements in Blood Donors. *Arthritis Rheum.* **2004**, *50* (2), 380–386. https://doi.org/10.1002/art.20018.

(33) El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S. R.; Luciani, A.; Potter, S. C.; Qureshi, M.; Richardson, L. J.; Salazar, G. A.; Smart, A.; Sonnhammer, E. L. L.; Hirsh, L.; Paladin, L.; Piovesan, D.; Tosatto, S. C. E.; Finn, R. D. The Pfam Protein Families Database in 2019. *Nucleic Acids Res.* **2019**, *47* (D1), D427–D432. https://doi.org/10.1093/nar/gky995.

(34) Fuhrmann, J.; Clancy, K. W.; Thompson, P. R. Chemical Biology of Protein Arginine Modifications in Epigenetic Regulation. *Chem. Rev.* **2015**, *115* (11), 5413–5461. https://doi.org/10.1021/acs.chemrev.5b00003.

(35) Clancy, K. W.; Weerapana, E.; Thompson, P. R. Detection and Identification of Protein Citrullination in Complex Biological Systems. *Curr. Opin. Chem. Biol.* **2016**, *30*, 1–6. https://doi.org/10.1016/j.cbpa.2015.10.014.

(36) Cox, C. L.; Tietz, J. I.; Sokolowski, K.; Melby, J. O.; Doroghazi, J. R.; Mitchell, D. A. Nucleophilic 1,4-Additions for Natural Product Discovery. *ACS Chem. Biol.* **2014**, *9* (9), 2014–2022. https://doi.org/10.1021/cb500324n.

(37) Maxson, T.; Tietz, J. I.; Hudson, G. A.; Guo, X. R.; Tai, H.-C.; Mitchell, D. A. Targeting Reactive Carbonyls for Identifying Natural Products and Their Biosynthetic Origins. *J. Am. Chem. Soc.* **2016**, *138* (46), 15157–15166. https://doi.org/10.1021/jacs.6b06848.

(38) Molloy, E. M.; Tietz, J. I.; Blair, P. M.; Mitchell, D. A. Biological Characterization of the Hygrobafilomycin Antibiotic JBIR-100 and Bioinformatic Insights into the Hygrolide Family of Natural Products. *Bioorg. Med. Chem.* **2016**, *24* (24), 6276–6290. https://doi.org/10.1016/j.bmc.2016.05.021.

(39) Castro-Falcón, G.; Millán-Aguiñaga, N.; Roullier, C.; Jensen, P. R.; Hughes, C. C. Nitrosopyridine Probe To Detect Polyketide Natural Products with Conjugated Alkenes:

Discovery of Novodaryamide and Nocarditriene. *ACS Chem. Biol.* **2018**, *13* (11), 3097–3106. https://doi.org/10.1021/acschembio.8b00598.

(40) Castro-Falcón, G.; Hahn, D.; Reimer, D.; Hughes, C. C. Thiol Probes To Detect Electrophilic Natural Products Based on Their Mechanism of Action. *ACS Chem. Biol.* **2016**, *11* (8), 2328–2336. https://doi.org/10.1021/acschembio.5b00924.

(41) Palaniappan, K. K.; Pitcher, A. A.; Smart, B. P.; Spiciarich, D. R.; Iavarone, A. T.; Bertozzi, C. R. Isotopic Signature Transfer and Mass Pattern Prediction (IsoStamp): An Enabling Technique for Chemically-Directed Proteomics. *ACS Chem. Biol.* **2011**, *6* (8), 829–836. https://doi.org/10.1021/cb100338x.

(42) Capehart, S. L.; Carlson, E. E. Mass Spectrometry-Based Assay for the Rapid Detection of Thiol-Containing Natural Products. *Chem. Commun.* **2016**, *52* (90), 13229–13232. https://doi.org/10.1039/C6CC07111B.

(43) Bicker, K. L.; Subramanian, V.; Chumanevich, A. A.; Hofseth, L. J.; Thompson, P. R. Seeing Citrulline: Development of a Phenylglyoxal-Based Probe To Visualize Protein Citrullination. *J. Am. Chem. Soc.* **2012**, *134* (41), 17015–17018. https://doi.org/10.1021/ja308871v.

(44) Choi, M.; Song, J.-S.; Kim, H.-J.; Cha, S.; Lee, E. Y. Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry Identification of Peptide Citrullination Site Using Br Signature. *Anal. Biochem.* **2013**, *437* (1), 62–67. https://doi.org/10.1016/j.ab.2013.03.003.

(45) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215* (3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

(46) Xie, X.; Marahiel, M. A. NMR as an Effective Tool for the Structure Determination of Lasso Peptides. *ChemBioChem* **2012**, *13* (5), 621–625. https://doi.org/10.1002/cbic.201100754.

(47) Schwieters, C. D.; Bermejo, G. A.; Clore, G. M. Xplor-NIH for Molecular Structure Determination from NMR and Other Data Sources. *Protein Sci.* **2018**, *27* (1), 26–40. https://doi.org/10.1002/pro.3248.

(48) Vallenet, D.; Calteau, A.; Dubois, M.; Amours, P.; Bazin, A.; Beuvin, M.; Burlot, L.; Bussell, X.; Fouteau, S.; Gautreau, G.; Lajus, A.; Langlois, J.; Planel, R.; Roche, D.; Rollin, J.; Rouy, Z.; Sabatet, V.; Médigue, C. MicroScope: An Integrated Platform for the Annotation and Exploration of Microbial Gene Functions through Genomic, Pangenomic and Metabolic Comparative Analysis. *Nucleic Acids Res.* **2020**, *48* (D1), D579–D589. https://doi.org/10.1093/nar/gkz926.

(49) Tran, N.-V.; Greshake Tzovaras, B.; Ebersberger, I. PhyloProfile: Dynamic Visualization and Exploration of Multi-Layered Phylogenetic Profiles. *Bioinformatics* **2018**, *34* (17), 3041–3043. https://doi.org/10.1093/bioinformatics/bty225.

(50) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25* (17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

(51) McDonald, J. H. *Handbook of Biological Statistics*, 3rd ed.; Sparky House Publishing: Baltimore, MD, 2014.

(52) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Nagaraja Rao, R.; Schoner, B. E. Plasmid Cloning Vectors for the Conjugal Transfer of DNA from Escherichia Coli to Streptomyces Spp. *Gene* **1992**, *116* (1), 43–49. https://doi.org/10.1016/0378-1119(92)90627-2.

(53) Herrmann, S.; Siegl, T.; Luzhetska, M.; Petzke, L.; Jilg, C.; Welle, E.; Erb, A.; Leadlay, P. F.; Bechthold, A.; Luzhetskyy, A. Site-Specific Recombination Strategies for Engineering Actinomycete Genomes. *Appl. Environ. Microbiol.* **2012**, *78* (6), 1804–1812. https://doi.org/10.1128/AEM.06054-11.

(54) Bibb, M. J.; Janssen, G. R.; Ward, J. M. Cloning and Analysis of the Promoter Region of the Erythromycin Resistance Gene (ErmE) of Streptomyces Erythraeus. *Gene* **1985**, *38* (1), 215–226. https://doi.org/10.1016/0378-1119(85)90220-3.

(55) Bicker, K. L.; Thompson, P. R. The Protein Arginine Deiminases: Structure, Function, Inhibition, and Disease. *Biopolymers* **2013**, *99* (2), 155–163. https://doi.org/10.1002/bip.22127.

(56) Montgomery, A. B.; Kopec, J.; Shrestha, L.; Thezenas, M.-L.; Burgess-Brown, N. A.; Fischer, R.; Yue, W. W.; Venables, P. J. Crystal Structure of Porphyromonas Gingivalis Peptidylarginine Deiminase: Implications for Autoimmunity in Rheumatoid Arthritis. *Ann. Rheum. Dis.* **2016**, *75* (6), 1255–1261. https://doi.org/10.1136/annrheumdis-2015-207656.

(57) Shirai, H.; Mokrab, Y.; Mizuguchi, K. The Guanidino-Group Modifying Enzymes: Structural Basis for Their Diversity and Commonality. *Proteins Struct. Funct. Bioinforma.* **2006**, *64* (4), 1010–1023. https://doi.org/10.1002/prot.20863.

(58) Linsky, T.; Fast, W. Mechanistic Similarity and Diversity among the Guanidine-Modifying Members of the Pentein Superfamily. *Biochim. Biophys. Acta BBA - Proteins Proteomics* **2010**, *1804* (10), 1943–1953. https://doi.org/10.1016/j.bbapap.2010.07.016.

(59) Zallot, R.; Oberg, N.; Gerlt, J. A. The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. *Biochemistry* **2019**, *58* (41), 4169–4182. https://doi.org/10.1021/acs.biochem.9b00735.

(60) Lawrence, J. G.; Hartl, D. L. Inference of Horizontal Genetic Transfer from Molecular Data: An Approach Using the Bootstrap. *Genetics* **1992**, *131* (3), 753–760.

(61) The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. - Abstract - Europe PMC https://europepmc.org/article/med/31553576 (accessed Apr 10, 2020).

(62) Kohl, M.; Wiese, S.; Warscheid, B. Cytoscape: Software for Visualization and Analysis of Biological Networks. In *Data Mining in Proteomics: From Standards to Applications*; Hamacher, M., Eisenacher, M., Stephan, C., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2011; pp 291–303. https://doi.org/10.1007/978-1-60761-987-1_18.

Chapter 6: Structure Prediction and Synthesis of Pyridine-Based Macrocyclic Natural Products[4]

Graham A. Hudson, Annie R. Hooper, Adam J. DiCaprio, David Sarlah, Douglas A. Mitchell

6.1 Introduction

Advances in genome sequencing and analysis have rapidly expanded our knowledge of natural product biosynthetic space.[1,2] This is especially true for the ribosomally synthesized and post-translationally modified peptides (RiPPs), where characterization of new compounds is outpaced by an increasingly inundating number of biosynthetic gene clusters (BGCs).[3–6] Natural product discovery efforts often involves time-consuming and frequently unsuccessful techniques, such as traditional "grind-and-find" or heterologous BGC expression.[7] Given sufficient knowledge of the biosynthetic enzymes, the structure of the final product can be predicted, permitting access to the compound by chemical synthesis.[8–11] RiPP structures are often more readily predicted from genomic information compared to other natural product classes given their direct ribosomal origin.[12,13] Structural predictions can enable directed synthetic efforts, which offer several advantages including an unparalleled capacity for analog generation (often with scalability) and obviating the need for tedious isolation from a producing organism.

Recently, we reported a comprehensive genomic landscape survey of the thiopeptides, an extensively modified RiPP class.[14] All thiopeptides feature azole/azoline heterocycles (from Cys, Ser, Thr), dehydroamino acids (from Ser/Thr), and (dehydro)piperidine/pyridine post-translational modifications (PTMs).[12,14,15] While certain thiopeptides contain a significant number of ancillary modifications, the class-defining PTM is the six-membered, *N*-heterocycle, which arises from an

[4] A.J.D. contributions: All NMR data acquisition, processing and interpretation of chemoenzymatically prepared samples. Generation of all NMR figures and assistance with manuscript preparation.

enzyme-catalyzed [4+2]-cycloaddition of two dehydroalanine (Dha) residues and the amide backbone.

6.2 Results and Discussion

We sought to determine if orthologs of the thiomuracin pyridine synthase, TbtD,[15,16] occurred in contexts beyond canonical thiopeptide BGCs. A bioinformatic search yielded several RiPP-like BGCs that lacked an azoline-forming cyclodehydratase while still possessing a split LanB, putatively responsible for Ser glutamylation and elimination, to yield Dha,[17,18] as well as a pyridine synthase (Figures 1A, S1, and Supplementary Dataset 1).[19,20] No other local genes were conserved, suggesting pyridine formation as the sole PTM for these RiPP products. The neighboring precursor peptides also generally lacked Cys, which would be required for thiazole/thiazoline formation, a PTM ubiquitous to all known thiopeptides. Given the lack of sulfur (and thus, thiazol(in)e rings), we termed these RiPPs as "pyritides".

We next tested our genome-guided structural prediction using the pyritide BGC from *Micromonospora rosaria* owing to strain availability. The C-terminal core regions of MroA1 and MroA2 (precursor peptides) contained two Ser residues (Figure 1), both presumably converted to Dha. A subsequent [4+2]-cycloaddition would result in 14- and 17-atom macrocyclic products, hereafter pyritide A1 (**1**) and pyritide A2 (**2**). To assess the validity of these predictions, we attempted to isolate **1** and **2** from the native producer; however, we were unsuccessful in detecting either pyritide from *M. rosaria*. Instead, an orthogonal, 11-step total synthesis (Figure S2) and chemoenzymatic route were first devised to obtain **1** (Scheme 1 and Supplementary Methods).

Pursuant to the chemoenzymatic approach, we recombinantly expressed and purified the MroA1 precursor peptide and MroD pyridine synthase as fusions to maltose-binding protein (MBP, Table S1 and Figure S3). To sidestep the tRNA-dependency of MroB/C, we carried out a

dehydrothiolation using dibromohexanediamide (**3**, DBHDA)[20–23] of an MroA1 variant with Ser1 and Ser6 replaced with Cys (**4**). The resulting di-dehydrated intermediate (**5**) was detected by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS, Figure 2). As the predicted substrate for the pyridine synthase, **5** was then reacted with MroD, yielding masses congruent with the predicted structure of **1** and the leader peptide (LP) bearing a C-terminal carboxamide, **6**.

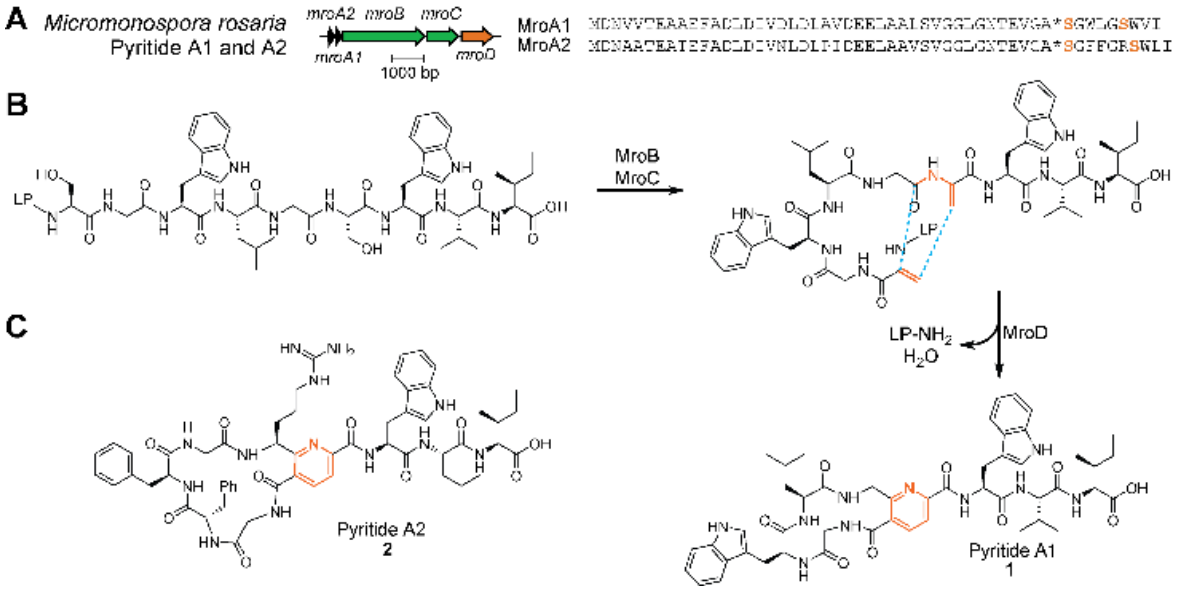
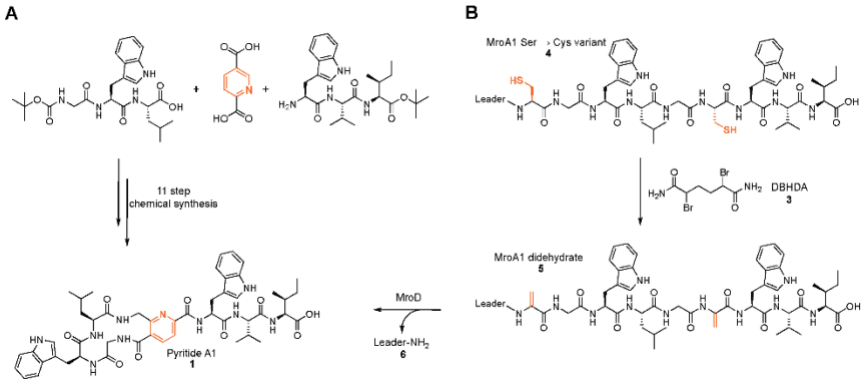


Figure 6.1. Biosynthesis of pyritides. (A) BGC from *Micromonospora rosaria*, encoding two precursor peptides, split LanB dehydratase, and pyridine synthase. (B) Proposed biosynthetic route to pyritide A1 (**1**). (C) Structure of pyritide A2 (**2**). LP, leader peptide.

This chemoenzymatic route was then carried out on a larger scale to produce a quantity of **1** sufficient for spectroscopic comparison to the total synthesis product. Approximately 30 L of MBP-MroA1 and 8 L of MBP-MroD culture was used to produce 2.5 mg (post-HPLC yield) of **1**. Several analytical techniques were used to characterize the synthetic and chemoenzymatic samples

of **1**, including high-resolution and tandem mass spectrometry (HR-MS/MS), C_{18} reverse-phase HPLC retention time, and ^{1}H NMR (Figures S4-S9, Tables S2-S3). Data from all three methods confirmed that the structure of **1** was correctly predicted and the synthetic and chemoenzymatically prepared samples were indistinguishable.

Scheme 6.1. Orthogonal routes to obtain **1**. (A) Chemical synthesis beginning from pyridine-2,5-dicarboxylic acid and protected tripeptides (see Figure S2). (B) Chemoenzymatic synthesis using MroA1 double cysteine variant to allow DBHDA-mediated (**3**) dehydrothiolation and subsequent enzymatic [4+2]-cycloaddition.



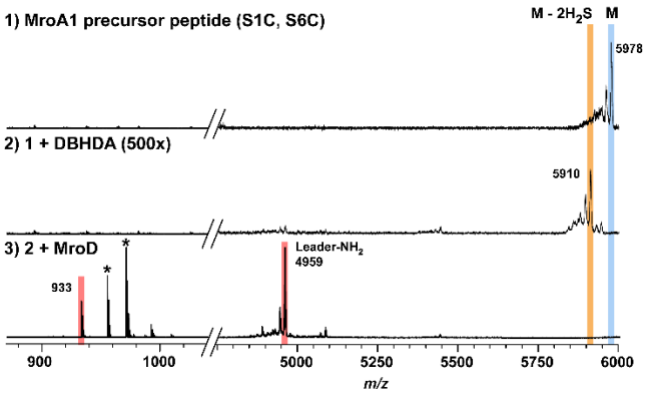


Figure 6.2. MALDI-TOF mass spectra detailing the chemoenzymatic synthesis of **1**. Compound **4** (*m/z* 5978 Da) was didehydrothiolated to **5** (*m/z* 5910 Da) followed by reaction with MroD. The resulting masses correspond to the $[M+H]^+$ ions for **1** (*m/z* 933 Da) and the eliminated leader peptide (**6**, *m/z* 4959 Da). Asterisks, Na^+/K^+ ions.

Using the chemoenzymatic route, we obtained an equivalent quantity of pyritide A2 (**2**). The principle difference between **1** and **2** is the presence of an Arg in **2** adjacent to the pyridine. This position is conserved in several other bioinformatically predicted pyritides (Figure S1). Analogous to MroA1, the two Ser within the core region of MroA2 were substituted with Cys prior to chemical dehydrothiolation by **3**. This resulted in the production of a mass consistent with the predicted structure of **2**, which was verified by ^{1}H NMR and HR-MS/MS (Figures S10-S13, Tables S2 and S3). The LP byproduct of MroA2 was detected as a C-terminal carboxamide (Figure S14). Purified **1** and **2** were then subjected to a brief panel of microbial growth-suppression assays; however, unlike the well-known activity of thiopeptides towards Gram-positive bacteria,[24] we did not observe any antimicrobial activity for **1** or **2** individually or in combination (Table S4). Further studies are warranted to investigate other potential bioactivities of pyritides.

Our data show that MroD catalyzes the formation of pyridine-based macrocycles of 14 or 17 atoms in size, mirroring the smallest size reported using enzymes and engineered substrates from the

lactazole biosynthetic pathway.[25] MroD is unique amongst characterized RiPP-related pyridine synthases in that it does not require preinstallation of azole heterocycles prior to yielding a macrocyclic product.[16,26] It should be noted that while a chemoenzymatic route offers convenient access to pyritide analogs, a drawback is that installing Dhb from Thr-containing ribosomal peptides cannot employ a dehydrothiolation route. However, alternate strategies using unnatural amino acids to incorporate Dhb have been developed.[27,28] Further, total synthesis provides a complementary avenue to produce Dhb-containing pyritide analogs.

Another pitfall of the chemoenzymatic approach is that it assumes one knows the extent of dehydration to a precursor peptide. In the cases of **1** and **2**, the predicted structures were unambiguous as only two Ser were present and both were required for pyridine formation. For other pyritide core sequences that contain additional Ser, Thr, or Cys, several distinct structures could result, which requires further exploration.

Advances in genomics has led to a renaissance in natural product discovery. Bioinformatic analyses offer a means to prioritize BGCs predicted to give rise to novel compounds, mitigating the rediscovery of known natural products. This paradigm shift is further enhanced by the unification of biosynthetic enzymology and total synthesis as it provides a dual platform to access new molecules. Here, we applied this methodology by identifying an unusual genomic context for RiPP-related pyridine synthases, leading to the structural prediction, chemical synthesis, and enzymatic verification of the pyritides. Given that the only PTM present in pyritide A1 and A2 is a trisubstituted pyridine (formed via an *aza*-[4+2]-cycloaddition), and the same PTM is class-defining for the thiopeptides, we propose that thiopeptides be reclassified as thiazole/thiazoline-containing pyritides.

We anticipate future investigations into pyritides will yield greater insight into the requisite enzymology and biological activity, which will undoubtedly be augmented by complementary chemical synthesis. Moreover, we predict that the interplay between natural product enzymology and chemical synthesis will become an increasingly useful means of accessing natural products and analogs thereof.

6.3 References

(1) Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland-Brown, L. C.; Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; Birren, B. W.; Takano, E.; Sali, A.; Linington, R. G.; Fischbach, M. A. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158* (2), 412–421. https://doi.org/10.1016/j.cell.2014.06.034.

(2) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A New Genome-Mining Tool Redefines the Lasso Peptide Biosynthetic Landscape. *Nat Chem Biol* **2017**, *13*, 470. https://doi.org/10.1038/nchembio.2319 https://www.nature.com/articles/nchembio.2319#supplementary-information.

(3) de los Santos, E. L. C. NeuRiPP: Neural Network Identification of RiPP Precursor Peptides. *Sci. Rep.* **2019**, *9* (1), 13406. https://doi.org/10.1038/s41598-019-49764-z.

(4) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(5) van Heel, A. J.; de Jong, A.; Song, C.; Viel, J. H.; Kok, J.; Kuipers, O. P. BAGEL4: A User-Friendly Web Server to Thoroughly Mine RiPPs and Bacteriocins. *Nucleic Acids Res.* **2018**, *46* (W1), W278–W281. https://doi.org/10.1093/nar/gky383.

(6) Skinnider, M. A.; Johnston, C. W.; Edgar, R. E.; Dejong, C. A.; Merwin, N. J.; Rees, P. N.; Magarvey, N. A. Genomic Charting of Ribosomally Synthesized Natural Product Chemical Space Facilitates Targeted Mining. *Proc. Natl. Acad. Sci.* **2016**, *113* (42), E6343. https://doi.org/10.1073/pnas.1609014113.

(7) Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H. Microbial Genome Mining for Accelerated Natural Products Discovery: Is a Renaissance in the Making? *J. Ind. Microbiol. Biotechnol.* **2014**, *41* (2), 175–184. https://doi.org/10.1007/s10295-013-1389-9.

(8) Vila-Farres, X.; Chu, J.; Ternei, M. A.; Lemetre, C.; Park, S.; Perlin, D. S.; Brady, S. F. An Optimized Synthetic-Bioinformatic Natural Product Antibiotic Sterilizes Multidrug-Resistant Acinetobacter Baumannii-Infected Wounds. *mSphere* **2018**, *3* (1), e00528-17. https://doi.org/10.1128/mSphere.00528-17.

(9) Chu, J.; Vila-Farres, X.; Inoyama, D.; Ternei, M.; Cohen, L. J.; Gordon, E. A.; Reddy, B. V. B.; Charlop-Powers, Z.; Zebroski, H. A.; Gallardo-Macias, R.; Jaskowski, M.; Satish, S.; Park, S.; Perlin, D. S.; Freundlich, J. S.; Brady, S. F. Discovery of MRSA Active Antibiotics Using Primary Sequence from the Human Microbiome. *Nat. Chem. Biol.* **2016**, *12* (12), 1004–1006. https://doi.org/10.1038/nchembio.2207.

(10) Chu, J.; Vila-Farres, X.; Inoyama, D.; Gallardo-Macias, R.; Jaskowski, M.; Satish, S.; Freundlich, J. S.; Brady, S. F. Human Microbiome Inspired Antibiotics with Improved β-Lactam Synergy against MDR Staphylococcus Aureus. *ACS Infect. Dis.* **2018**, *4* (1), 33–38. https://doi.org/10.1021/acsinfecdis.7b00056.

(11) Chu, J.; Vila-Farres, X.; Brady, S. F. Bioactive Synthetic-Bioinformatic Natural Product Cyclic Peptides Inspired by Nonribosomal Peptide Synthetase Gene Clusters from the Human

Microbiome. *J. Am. Chem. Soc.* **2019**, *141* (40), 15737–15741. https://doi.org/10.1021/jacs.9b07317.

(12) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Goransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Muller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Sussmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat Prod Rep* **2013**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(13) Funk, M. A.; van der Donk, W. A. Ribosomal Natural Products, Tailored To Fit. *Acc. Chem. Res.* **2017**, *50* (7), 1577–1586. https://doi.org/10.1021/acs.accounts.7b00175.

(14) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. Bioinformatic Expansion and Discovery of Thiopeptide Antibiotics. *J. Am. Chem. Soc.* **2018**, *140* (30), 9494–9501. https://doi.org/10.1021/jacs.8b03896.

(15) Hudson, G. A.; Zhang, Z.; Tietz, J. I.; Mitchell, D. A.; van der Donk, W. A. In Vitro Biosynthesis of the Core Scaffold of the Thiopeptide Thiomuracin. *J Am Chem Soc* **2015**, *137* (51), 16012–16015. https://doi.org/10.1021/jacs.5b10194.

(16) Zhang, Z.; Hudson, G. A.; Mahanta, N.; Tietz, J. I.; van der Donk, W. A.; Mitchell, D. A. Biosynthetic Timing and Substrate Specificity for the Thiopeptide Thiomuracin. *J Am Chem Soc* **2016**, *138* (48), 15511–15514. https://doi.org/10.1021/jacs.6b08987.

(17) Bothwell, I. R.; Cogan, D. P.; Kim, T.; Reinhardt, C. J.; van der Donk, W. A.; Nair, S. K. Characterization of Glutamyl-TRNA–Dependent Dehydratases Using Nonreactive Substrate Mimics. *Proc. Natl. Acad. Sci.* **2019**, *116* (35), 17245. https://doi.org/10.1073/pnas.1905240116.

(18) Ortega, M. A.; Hao, Y.; Zhang, Q.; Walker, M. C.; van der Donk, W. A.; Nair, S. K. Structure and Mechanism of the TRNA-Dependent Lantibiotic Dehydratase NisB. *Nature* **2015**, *517* (7535), 509–512. https://doi.org/10.1038/nature13888.

(19) Cogan, D. P.; Hudson, G. A.; Zhang, Z.; Pogorelov, T. V.; van der Donk, W. A.; Mitchell, D. A.; Nair, S. K. Structural Insights into Enzymatic [4+2] Aza-Cycloaddition in Thiopeptide Antibiotic Biosynthesis. *Proc Natl Acad Sci USA* **2017**.

(20) Wever, W. J.; Bogart, J. W.; Baccile, J. A.; Chan, A. N.; Schroeder, F. C.; Bowers, A. A. Chemoenzymatic Synthesis of Thiazolyl Peptide Natural Products Featuring an Enzyme-Catalyzed Formal [4 + 2] Cycloaddition. *J. Am. Chem. Soc.* **2015**, *137* (10), 3494–3497. https://doi.org/10.1021/jacs.5b00940.

(21) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernández-González, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. Methods for Converting Cysteine to Dehydroalanine on Peptides and Proteins. *Chem. Sci.* **2011**, *2* (9), 1666–1676. https://doi.org/10.1039/C1SC00185J.

(22) Wright, T. H.; Bower, B. J.; Chalker, J. M.; Bernardes, G. J. L.; Wiewiora, R.; Ng, W.-L.; Raj, R.; Faulkner, S.; Vallée, M. R. J.; Phanumartwiwath, A.; Coleman, O. D.; Thézénas, M.-L.; Khan, M.; Galan, S. R. G.; Lercher, L.; Schombs, M. W.; Gerstberger, S.; Palm-Espling, M. E.;

Baldwin, A. J.; Kessler, B. M.; Claridge, T. D. W.; Mohammed, S.; Davis, B. G. Posttranslational Mutagenesis: A Chemical Strategy for Exploring Protein Side-Chain Diversity. *Science* **2016**, aag1465. https://doi.org/10.1126/science.aag1465.

(23) Galan, S. R. G.; Wickens, J. R.; Dadova, J.; Ng, W.-L.; Zhang, X.; Simion, R. A.; Quinlan, R.; Pires, E.; Paton, R. S.; Caddick, S.; Chudasama, V.; Davis, B. G. Post-Translational Site-Selective Protein Backbone α-Deuteration. *Nat. Chem. Biol.* **2018**, *14* (10), 955–963. https://doi.org/10.1038/s41589-018-0128-y.

(24) Vinogradov, A. A.; Suga, H. Introduction to Thiopeptides: Biological Activity, Biosynthesis, and Strategies for Functional Reprogramming. *Cell Chem. Biol.* **2020**. https://doi.org/10.1016/j.chembiol.2020.07.003.

(25) Vinogradov, A. A.; Shimomura, M.; Goto, Y.; Ozaki, T.; Asamizu, S.; Sugai, Y.; Suga, H.; Onaka, H. Minimal Lactazole Scaffold for in Vitro Thiopeptide Bioengineering. *Nat. Commun.* **2020**, *11* (1), 2272. https://doi.org/10.1038/s41467-020-16145-4.

(26) Bogart, J. W.; Kramer, N. J.; Turlik, A.; Bleich, R. M.; Catlin, D. S.; Schroeder, F. C.; Nair, S. K.; Williamson, R. T.; Houk, K. N.; Bowers, A. A. Interception of the Bycroft–Gowland Intermediate in the Enzymatic Macrocyclization of Thiopeptides. *J. Am. Chem. Soc.* **2020**, *142* (30), 13170–13179. https://doi.org/10.1021/jacs.0c05639.

(27) Goto, Y.; Iwasaki, K.; Torikai, K.; Murakami, H.; Suga, H. Ribosomal Synthesis of Dehydrobutyrine- and Methyllanthionine-Containing Peptides. *Chem. Commun.* **2009**, No. 23, 3419–3421. https://doi.org/10.1039/B904314D.

(28) Yang, B.; Wang, N.; Schnier, P. D.; Zheng, F.; Zhu, H.; Polizzi, N. F.; Ittuveetil, A.; Saikam, V.; DeGrado, W. F.; Wang, Q.; Wang, P. G.; Wang, L. Genetically Introducing

Biochemically Reactive Amino Acids Dehydroalanine and Dehydrobutyrine in Proteins. *J. Am. Chem. Soc.* **2019**, *141* (19), 7698–7703. https://doi.org/10.1021/jacs.9b02611.

Chapter 7: Bioinformatics-Guided Expansion and Discovery of Graspetides[5]

Sangeetha Ramesh, Xiaorui Guo, Adam J. DiCaprio, Ashley Delio, Bryce L. Kille[|], Taras V. Pogorelov, Douglas A. Mitchell[†]

7.1 Introduction

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a diverse family of natural products that display a range of biological activities.[1] Although structurally and functionally diverse, RiPPs are united by a common biosynthetic logic. RiPP biosynthesis begins with the ribosomal synthesis of the precursor peptide, which typically consists of an N-terminal leader region and a C-terminal core region. The leader region includes a recognition sequence bound by biosynthetic proteins to allow enzymatic modification of the core region. The leader region is subsequently removed enzymatically, occasionally by a peptidase encoded in the biosynthetic gene cluster (BGC). Further tailoring of the core region can ensue by the action of leader-independent modification enzymes, followed by cellular export of the mature RiPP. A compelling feature of RiPP biosynthesis is that a direct gene product (ribosomal peptide) can be converted into a structurally complex natural product through only a few enzymatic steps.[1] Furthermore, genome sequencing efforts and improvements in RiPP genome-mining tools in the past decade have facilitated the emergence of RiPPs as a major family of natural products with over 40 classes described to date.[2–4]

Microviridins **(Figure 1)** are RiPPs exhibiting a tricyclic structure resulting from ATP-grasp ligase-dependent installation of macrolactones/macrolactams.[5,6] The ATP-grasp ligases (InterPro: IPR011761)[7] include well-studied enzymes such as D-Ala-D-Ala ligase, glutathione synthetase,

[5] A.J.D. contributions: All NMR data acquisition, processing, and interpretation. Generation of all NMR figures and assistance with manuscript preparation. Generation of NOE constraints for structure generation.

and biotin carboxylase, which use ATP to phosphorylate a carboxylate moiety within a small molecule or peptidic substrate.[8] The ligase then directs a nucleophile towards the electrophilic acyl phosphate intermediate to yield a new carbonyl-containing functional group, often an amide or ester. MdnC/MvdD and MdnB/MvdC homologs (InterPro families: IPR026439 and IPR026446)[7] that install the macrolactone and macrolactam linkages, respectively, in microviridins are homologous to RimK (InterPro: IPR023533), which appends the amino group of L-Glu to the C-terminus of the ribosomal protein S6.[5,6] Nineteen microviridin congeners have been characterized thus far, and all are inhibitors of serine proteases such as trypsin, chymotrypsin, subtilase, and elastase.[9–12] Aside from canonical microviridins, a structurally related peptide, marinostatin, comprised of two macrolactone rings, has also been characterized from *Alteromonas* sp. B-10-31.[13] Several genome-mining studies have been conducted on microviridin-like RiPPs. One early bioinformatic study revealed the presence of BGCs across diverse bacterial phyla, encoding an assortment of peptides enriched in Asp/Glu and Ser/Thr/Lys residues, which serve as the acceptor and donor residues, respectively, for macrolactone/macrolactam formation.[14] A more recent genome-mining study identified 174 microviridin and marinostatin BGCs present within Proteobacteria, Cyanobacteria, and Bacteroidetes, revealing a greater diversity of microviridin precursor peptides, the majority of which are uncharacterized.[11]

Non-microviridin RiPPs containing ATP-grasp ligase-mediated macrolactones, namely plesiocin and thuringinin (**Figure 1**), have been expressed heterologously in *E. coli*.[15,16] Plesiocin and thuringinin both portray a hairpin (ring-within-a-ring) topology, formed through two macrolactone linkages on a conserved "TTx$_4$EE" (where x is any amino acid) and "Tx$_2$Tx$_3$Ex$_2$DxD" motif, respectively. Unified by an ATP-grasp ligase-dependent biosynthetic pathway, macrolactone-containing microviridins, plesiocins, and thuringinins were termed the omega ester-containing

peptides (OEPs).[16] The most recent genome-mining study significantly expanded the OEP family, identifying 1,724 BGCs across 12 groups, each exhibiting distinct core peptide consensus sequences.[17] The OEPs were recently renamed as the graspetides, a term that describes the enzyme that installs both the class-defining modifications.[4] Currently, only graspetides from groups 1-6 and 11 have at least one characterized member[11,15–18], underscoring that most structural and functional diversity is yet to be uncovered (**Figure 1**).

While the previous bioinformatic studies have illustrated the prevalence of diverse graspetides, the current repertoire of RiPP mining tools does not readily facilitate the genomic detection of graspetides. Tools such as antiSMASH[19], RiPP-PRISM[20], and RiPPER[21] allow genome-mining for microviridins but do not confidently predict plesiocins, thuringinins, or other members of the graspetide family. To enable rapid identification of putative graspetide precursor peptides and BGCs, we enhanced the capabilities of Rapid ORF Description and Evaluation Online (RODEO)[22] by including a graspetide scoring module. RODEO uses profile hidden Markov model (pHMM)-based analysis to predict the function of genes flanking a query protein. RODEO utilizes a combination of heuristic scoring, motif analysis, and support vector machine classification to provide the most accurate, RiPP class-specific predictions for precursor peptides currently available. The current version identifies precursor peptides for lanthipeptides, lasso peptides, linaridins, ranthipeptides, sactipeptides, and thiopeptides.[22–26]

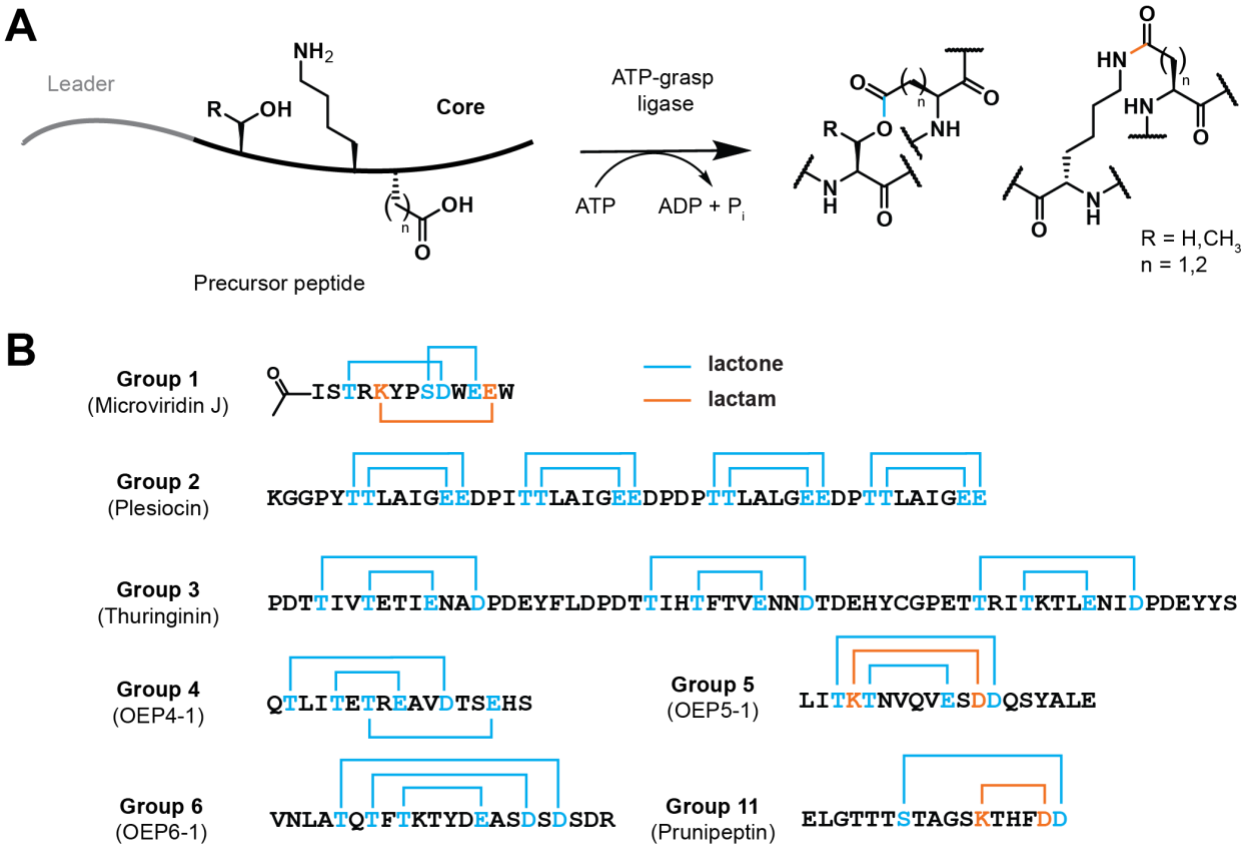


Figure 7.1. Graspetide biosynthesis and ring topologies. A. Formation of a macrolactone (blue) and macrolactam (orange) linkage during graspetide biosynthesis. B. Ring topology of graspetide representatives from each characterized group. In group 1 graspetides, the macrolactone/macrolactam rings are installed by two distinct ATP-grasp ligases, whereas in group 5 and 11 graspetides, a single ATP-grasp ligase installs both linkage types.

Here, we developed an improved version of RODEO and compiled a dataset of all observable graspetides in the NCBI non-redundant database, revealing 4,356 (3,455 unique) high-confidence graspetide precursor peptides distributed amongst 3,923 BGCs. Significant portions of the RODEO code have been rewritten, which renders the future incorporation of scoring modules less cumbersome. This report features the most comprehensive analysis of the graspetides to date, consisting of 24 graspetide groups. Of these, 12 groups are new and contain 1,769 unique

Figure 7.1. (cont.) graspetide BGCs. Thatisin, from a newly identified group (group 16), was structurally characterized to feature two macrolactones with an interlocking ring topology. Interestingly, thatisin could be isolated as two atropisomers at room temperature. NMR spectroscopy and computational analyses attribute *cis/trans* Pro isomerization to the observed atropisomerism in thatisin.

7.2 Results and Discussion

7.2.1 Development of a graspetide-specific RODEO scoring module

The class-defining modification for the graspetides are macrolactone/macrolactam linkages installed by an ATP-grasp ligase.[4] Thus, the minimum requirement for a graspetide BGC are two genes encoding for an ATP-grasp ligase and a precursor peptide, which requires at least one Asp/Glu and one Ser/Thr/Lys residue in the core region. Such minimal attributes for a graspetide precursor peptide render discriminating bona fide substrates from hypothetical, non-coding sequences a challenging task, for which machine learning seemed ideally suited to solve. Another obstacle to constructing a database the encompasses all graspetides is that the ATP-grasp ligases are one of the largest enzyme superfamilies comprising over 444,000 members (InterPro: IPR011761). However, the vast majority are involved in primary metabolic pathways, including glutathione, purine, cell wall, and ribosomal biosynthesis.[8] To restrict our dataset to include only ATP-grasp ligases potentially involved in graspetide biosynthesis, a Position-Specific Iterative BLAST[27] search was performed against the NCBI non-redundant database using ATP-grasp ligases from known graspetide BGCs (one from each of the seven characterized groups) as queries.[6,17,18] The search was restricted to 3 iterations with an expectation value threshold of e^{-30} to exclude unlikely candidates. The resulting non-redundant set of 6,625 ATP-grasp ligases was

analyzed using the newly developed RODEO graspetide module (described below) to annotate the local genomic region and identify putative graspetide precursor peptides.

The graspetide precursor peptide scoring module utilizes a combination of heuristic scoring, MEME-based motif analysis, and support vector machine learning similar to previously reported RODEO modules.[22–26] The relevant parameters (**Tables S1-S2**) were chosen based on trends observed among graspetide groups with characterized members (groups 1-6, 11).[11,17] Unlike precursor peptides from most other RiPP classes, which tend to be short and are missed by automated gene finders, graspetide precursor peptides are often longer. Indeed, the average length of group 1-6 graspetide precursor peptides is nearly 100 amino acids. Some representatives feature multiple repeats and can be over several hundred amino acids in length. For example, a putative plesiocin (group 2) precursor peptide from *Acidobacteria bacterium* isolate gp7 AA5 (ATP-grasp ligase accession: PYQ56175.1) is 225 amino acids long and contains 13 "TTx$_4$EE" repeats.[17] Such precursors are not typically scored by RODEO, given the default upper limit for precursor peptide length is 110 amino acids. However, increasing this parameter results in a more significant false-positive rate and an increase in computation time. To solve this problem, we incorporated RADAR (Rapid Automatic Detection and Alignment of Repeats)[28] into the graspetide scoring module. The number of polypeptide sequences subject to graspetide precursor peptide scoring is reduced by removing large sequences (< 400 amino acids) that lack RADAR-identified repeats enriched in Asp, Glu, Ser, Thr, or Lys residues from the analyses. The detailed graspetide module workflow is available in **Figure S1**. Seven custom pHMMs were constructed (one each for groups 2-6 and two for group 1) and included in the graspetide module to readily distinguish graspetide-modifying ATP-grasp ligases from other members of the ATP-grasp ligase family (**Supplemental Dataset 1**). The graspetide module was trained using a randomly selected set of characterized and high-

confidence, predicted graspetide precursors from groups 1-6 published in a previous study.[17] Upon testing an optimized module (**Figure S2**) with the remaining set of predicted, high-confidence graspetide precursors from groups 1-6[17], we obtained precision (true positives/observed positives) and recall (true positives/actual positives) values of 0.98 and 0.98 at a cut-off score of 12.

Hypothetical peptides were considered a candidate graspetide precursor peptide if one of the following criteria were met: *(i)* the translated open-reading frame (ORF) received a RODEO score greater than 11, given the optimized balance between precision and recall, *(ii)* the translated ORF scored between 8-11 and also formed a group with similar members on a sequence similarity network (SSN). The precursor peptides identified by these stringent criteria were analyzed further using an SSN to remove less likely candidates resulting in 4,356 precursor peptides (3,455 unique) associated with 4,329 ATP-grasp ligases distributed amongst 3,923 unique BGCs (**Supplemental Dataset 2**). Many ORFs scoring below the validity threshold recapitulate some features of higher confidence graspetide precursors and may represent false negatives; however, we have excluded these from further analysis owing to the lower score (**Figure S2**). As new graspetide members are characterized, the scoring module can be retrained for confident identification of divergent members of this family, as recently demonstrated for the lasso peptides.[29]

7.2.2 Bioinformatic analysis of the graspetide family

To examine the sequence diversity of the graspetides and subsequently classify them into different groups, SSNs for the precursor peptides and the ATP-grasp ligases (**Figures S3** and **2, respectively**) were generated. Among the 3,923 RODEO-identified graspetide BGCs, 2,154 BGCs belong to groups 1-12, which were also identified in a previous genome-mining study.[17] In this study, graspetide (OEPs, previously) precursor peptides were identified by searching for sequence patterns enriched in macrolactone/macrolactam donor and acceptor residues within protein

sequences less than 200 amino acids long and encoded within 3 kb of an ATP-grasp ligase gene. This tedious process is now automated by RODEO. A detailed bioinformatic analysis of group 1-12 graspetides can be found in the aforementioned study. The remaining 1,769 graspetide BGCs are classified into 12 groups based on the cooccurrence of genes and conservation of motifs in the leader and core regions of the precursor peptide.

The largest of the newly cataloged groups (group 13) is the Grasp-with-Methyltransferase (GwM) group, comprising 1,326 BGCs. The GwM group is almost exclusively found in Actinobacteria (99%), although a few examples derive from Chloroflexi and Acidobacteria (1%, **Figures 2** and **S3; Table S3**). The methyltransferases found in GwM BGCs (**Figure S4**) are annotated as *S*-adenosylmethionine (SAM) dependent *O*-methyltransferases (InterPro: IPR026448). Members of this family are related to InterPro families IPR000682, which convert L-isoaspartyl and D-aspartyl residues to L-Asp, and IPR027573 that frequently cooccur with lanthipeptide biosynthetic enzymes.[30] Recently, a member of the latter family, $OlvS_A$, was implicated in lanthipeptide biosynthesis.[31] $OlvS_A$ catalyzes a three-step reaction (methylation, cyclization, and hydrolysis) of a highly conserved L-Asp to the corresponding L-isoaspartyl residue in the lanthipeptide OlvA (BCS_A). The prevalence of conserved Asp residues in the C-terminus of the GwM group precursor peptides (**Figure S5**) suggests that these residues may be similar targets for the methyltransferase. The GwM group can be further classified into five subgroups based on conserved motifs in the N-terminal region of the precursor peptide, potentially involved in leader peptide recognition. GwM subgroup 1 features a microviridin-like leader region motif, “PFAFRFL”. The “PFFARFL” motif in the microviridin leader peptide mediates recognition of the precursor peptide by the cognate ATP-grasp ligases MdnB and MdnC for macrocyclizations. The GwM subgroup 1 core regions, however, are unlike any characterized graspetide. GwM subgroups 3 and 4 are classified further

based on conserved motifs in the core regions enriched in Asp, Glu, Ser, Thr, and Lys residues. Several precursor peptides from subgroups 3 and 4 contain a "DPxTQb" motif (where x is any amino acid and b is a branched-chain amino acid, **Figure S5**), suggesting a functional role.

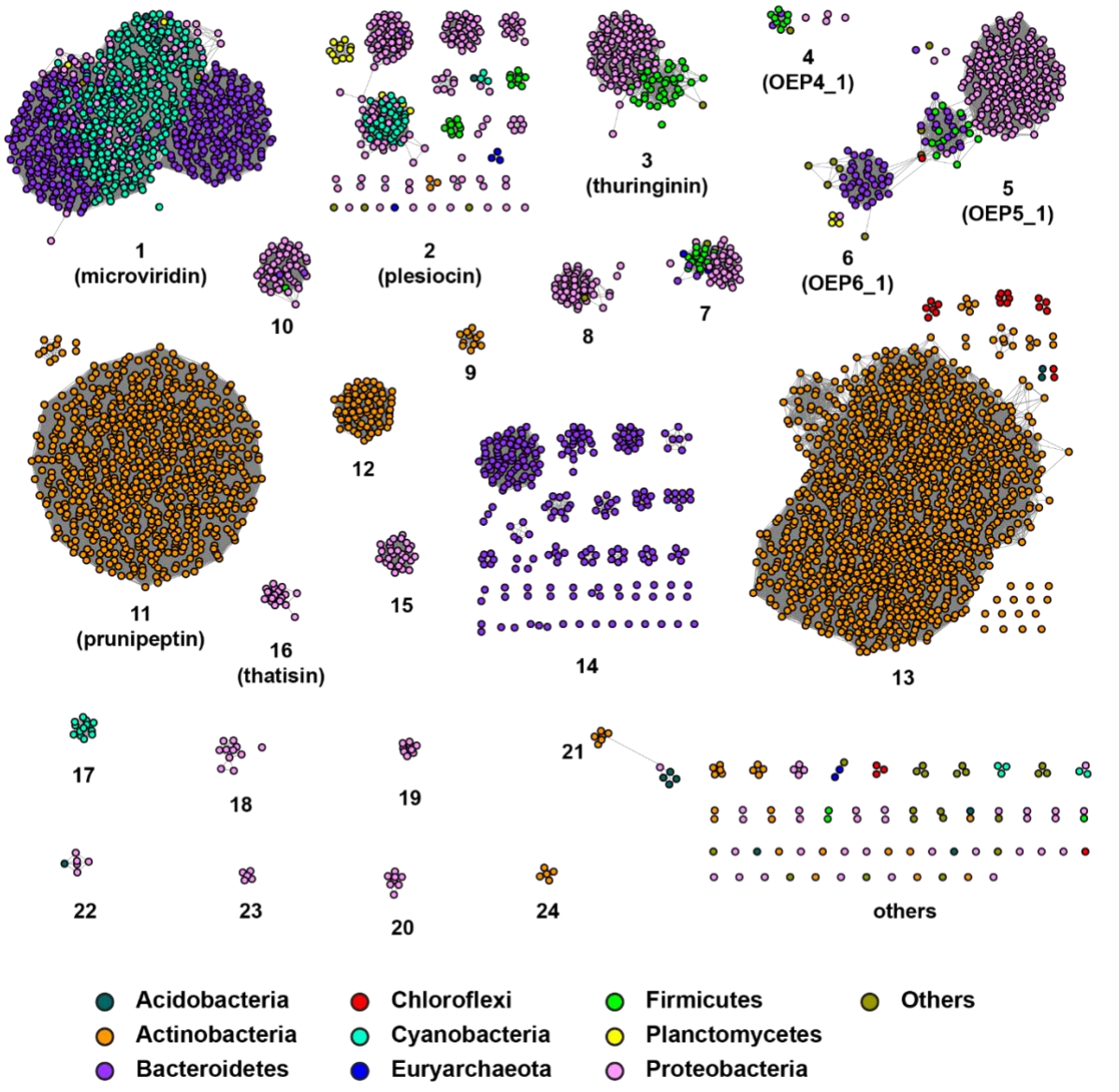


Figure 7.2. Sequence similarity network of graspetide-associated ATP-grasp ligases. Non-redundant ATP-grasp ligases with a RODEO-identified precursor peptide were used as input for this SSN (n = 4,329). The SSN was generated using EFI-EST[32] and visualized using Cytoscape[33]

Figure 7.2. (cont.) at an alignment score of 70. Nodes are numbered and colored according to graspetide group and phylum, respectively. An SSN for the corresponding precursor peptides and sequence analysis for **Figure 7.2 (cont.)** the new graspetide groups identified in this study are available in the supplementary information document.

The second-largest newly identified group (group 14) includes 245 BGCs, predominantly from Bacteroidetes, and is the Grasp-with-SPASM (GwS) group (**Figure 2**). All GwS BGCs (**Figure S4**) include an ATP-grasp ligase and a "SPASM" domain-containing enzyme (InterPro: IPR026497). "SPASM" is an acronym derived from the founding members of the family: subtilosin, pyrroloquinoline quinolone, anaerobic sulfatase-maturing enzyme, and mycofactosin, all radical SAM (rSAM) enzymes.[4] rSAM enzymes featuring an auxiliary SPASM domain are involved in various RiPP biosynthetic reactions that include C-C and C-heteroatom macrocyclizations, site-specific epimerizations, and other chemically demanding transformations.[4,34] SPASM domain-containing enzymes from GwS clusters, however, lack the 4Fe-4S cluster in the N-terminal domain required for binding rSAM. The GwS group is further divided into 14 subgroups based on conserved motifs observed in the leader and core regions of the precursor peptide. The extent of precursor peptide diversity within the GwS subgroups (**Figure S6**), combined with the likely processing by the neighboring, yet-to-be characterized SPASM-domain-containing enzyme, increases the appeal of the GwS group for uncovering new graspetide structures. The majority (>90%) of GwS precursor peptides harbor a Gly-Gly motif, a common RiPP leader peptide cleavage site that is often a substrate for peptidase-containing ATP-binding transporters (PCATs, Interpro: IPR005897).[35] The conserved leader peptide motifs of the GwS group are enriched in Lys and Leu residues with variations among the different subgroups, likely

involved in binding interactions with the biosynthetic proteins (**Figure S7**). The recent RODEO-mediated expansion of the lanthipeptide family also noted a similar trend among class I lanthipeptide precursors from Bacteroidetes, which feature an "LxLxKx_5L" motif and a double Gly motif at the leader/core junction and a clan C39 peptidase in the BGC.[25]

Among the remaining graspetide BGCs, 104 can be classified into 10 new groups (groups 15-24 **Figure S4**) based on distinct conserved motifs in the leader and core regions of the precursor peptides (**Figure S8**). Notably, group 15 precursors are confined to the taxonomic order Myxococcales and are about 200 amino acids long and feature multiple core repeats. The BGCs typically encode an α-ketoglutarate/Fe(II)-dependent oxygenase (Interpro: IPR005123) (**Figure S9**). A small population of graspetide BGCs (94 total) is either singletons or forms groups of less than five unique BGCs. Among these, some natural hybrid RiPP BGCs were observed (**Figure S10**). For instance, three graspetide precursors from *Nitrospira* sp. display a "YxxP" motif in the leader region, a known recognition sequence (i.e., the motif directly engaged by the RiPP Recognition Element, RRE) from lasso peptide biosynthesis.[4] Besides harboring two ATP-grasp ligases, the *Nitrospira* sp. BGCs also encode a lasso peptide-like leader peptidase, lasso peptide-like RRE[36], and several other modification enzymes. These BGCs, however, lack the cyclase responsible for lasso peptide formation (**Figure S10**). Five BGCs encode a TfuA and YcaO pair known to install thioamide linkages in polypeptides, in addition to an ATP-grasp ligase.[4,23,37]

Actinobacteria, Bacteroidetes, and Proteobacteria encode the vast majority of the graspetide BGCs (87%). (**Figure 2, Table S3**). Most graspetide groups are confined to a single phylum, with a few exceptions. Although comprising only ~20% of the graspetide dataset (n = 753 BGCs), Proteobacteria represent the most diverse taxonomic group with members in 19 of the 24 defined graspetide groups. Several RiPP scoring modules in RODEO support the prediction of leader

peptide cleavage sites, allowing quantitative analysis of the amino acid composition of the core peptide.[22,23,26] However, graspetides present a unique challenge in that they cover a wide range of lengths (**Figure S11**), and little to no information is known about leader peptide removal. Thus, a similar analysis of the core region could not be performed for the graspetide precursors identified in this study.

To understand the relationship between graspetide-modifying ATP-grasp ligases and other members of the ATP-grasp ligase superfamily, we generated an SSN of representative members of the InterPro family IPR011761. The graspetide-modifying ATP-grasp ligases were artificially introduced into this dataset since most of them were not part of the Uniprot database. This SSN suggests that graspetide-modifying ATP-grasp ligases are more closely related to RimK and LysX than to other members of the ATP-grasp ligase superfamily (**Figure S12**). While RimK appends L-Glu residues to the C-terminus of the ribosomal protein S6, LysX is involved in the biosynthesis of Lys via the α-amino adipate pathway.[8] The SSN further teases the possibility that the graspetide-modifying ATP-grasp ligases emerged once from an ancestral ATP-grasp ligase and diversified later on to introduce different ring topologies. To probe this further, we generated a maximum-likelihood tree using representative graspetide-modifying ATP-grasp ligases and rooted the tree using RimK. In addition to supporting the hypothesis, the tree also indicates that ATP-grasp ligases modifying groups 2 and 15 graspetides diverged early from the rest of the graspetide-modifying ATP-grasp ligases and are more closely related to the last common ancestral ATP-grasp ligase (**Figure S13, Supplemental Dataset 3**). This observation seems reasonable from an evolutionary standpoint since the substrates for groups 2 and 15 more closely resemble a protein and comprise some of the largest known graspetide precursor peptides (**Figure S11**). Parametric %GC content comparison between the gene encoding the ATP-grasp ligase and the genome of the organism

confirms that most graspetide BGCs cataloged in this study were vertically acquired (**Figure S14**). However, we can't exclude the possibility of horizontal gene transfer between organisms with similar %GC content.

7.2.3 Thatisin, a new graspetide, exhibits atropisomerism

We next sought to isolate and characterize a graspetide from *Lysobacter antibioticus* ATCC 29479 that belongs to a newly cataloged group, group 16. Based on a conserved $Tx_2Tx_4DDx_2D$ motif in the core region, we expected that mature graspetides from group 16 would be structurally unique (**Supplemental Dataset 2**). The BGC is comprised of genes encoding a precursor peptide (ThtA, WP_057916646.1), ATP-grasp ligase (ThtB, WP_057916645.1), PCAT (ThtC), and HlyD-like transporter (ThtD, **Figure 3**). Although *thtA* was annotated as a protein-coding gene, ten of the 16 precursor peptides from group 16 are not annotated as genes in the NCBI database. ThtA precursor peptide was co-expressed as a tobacco etch virus (TEV) protease-cleavable fusion to the C-terminus of maltose-binding protein (MBP) with the ATP-grasp ligase, ThtB. After an induction period, amylose affinity chromatography was used to purify modified and unmodified MBP-ThtA peptides from the cell pellet before TEV protease treatment and analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS). Upon co-expression of ThtA with ThtB, we observed a mass 36 Da lighter than a control lacking ThtB, suggestive of two dehydrations (**Figure S15**). Analysis of group 16 (ThtA-like) precursor peptides revealed strictly conserved Thr5, Thr8, Asp13, Asp14, and Asp17 (**Figure 3**). Given the lack of a conserved Lys, the 36 Da mass loss was hypothesized to arise from two Thr-Asp (macrolactone) linkages.

To aid the MS-based analysis, we next removed the leader region from modified ThtA using the PCAT LahT (WP_051646490.1, residues 1-150). LahT150 is the soluble, substrate-tolerant Cys protease domain of the PCAT from *Lachnospiraceae* sp. that shares 54% amino acid similarity

(39% identity) to the homologous domain of ThtC (**Figure S16**).[35] LahT150 recognizes and cleaves after Lx_4Lx_4Gy (where x = any amino acid; y = Ala, Gly, or Ser) motif in the leader region.[35] Although ThtA has a slightly altered recognition motif (Lx_4Ix_4Gy, **Figure 3**), it was processed by LahT150. After confirming leader peptidolysis by MALDI-TOF-MS (**Figure S15**), the modified core region of ThtA was purified via HPLC. Analysis of the chromatographic fractions by MALDI-TOF-MS revealed two isobaric products corresponding to doubly dehydrated ThtA core peptide (**Figure S17**). We named the major, earlier-eluting product, and the minor, later-eluting isobaric product as thatisin and *iso*-thatisin, respectively.

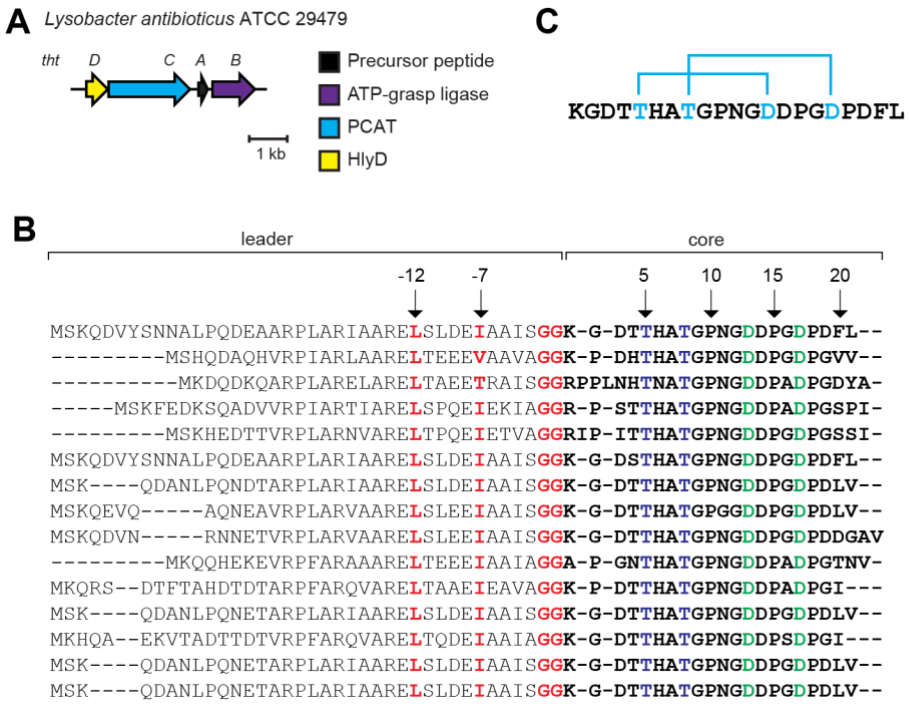


Figure 7.3. Overview of thatisin group graspetides. A. BGC diagram. PCAT, peptidase-containing ATP-binding transporter. B. Alignment of thatisin (group 16) precursor peptides. Top sequence is ThtA. The core region is bolded, and the acceptor and donor residues for macrolactones are in green and blue, respectively. Numbers indicate residue position in ThtA. PCAT recognition sequence is in red.[35] C. Ring topology of thatisin. Cyan line indicates macrolactone.

To confirm or refute if thatisin contains two macrolactone rings, we subjected purified thatisin to methanolysis under basic conditions with subsequent analysis using high-resolution and tandem mass spectrometry (HRMS/MS). Given the selective susceptibility of lactones to methanolysis, this procedure readily discriminates between macrolactone and macrolactam rings.[17] The number of macrolactones is also discernible, given the expected 14 Da increase to each methanolized acceptor residue (Asp). Methanol-treated thatisin yielded a product ($[M + 3H]^{3+}$ = 718.9904) in good agreement with the expected mass of a peptide containing two methyl esters ($[M + 3H]^{3+}$ = 718.9907; error = 0.3 ppm, **Figure S18**). This ion was then subjected to collision-induced dissociation (CID) with the resulting b- and y-ion series consistent with the +14 Da masses localizing to Asp13 and Asp17. The identity of these acceptor residues was further corroborated by heterologous co-expression experiments and MALDI-TOF-MS analysis of ThtA Ala-substituted variants. Two dehydrations were observed for the D3A, D14A, and D19A variants, while only 18 Da mass losses were observed for the D13A and D17A variants (**Figure S19**). Given the invariance of two macrolactone donor residues, Thr5 and Thr8, among ThtA-like precursor peptides (**Figure 3**), and the lack of fragmentation between Thr5 and Asp17 in thatisin (**Figure S18**), we rationalized that Thr5 and Thr8 were the most probable macrolactone donor residues for thatisin. Co-expression of Thr to Ala ThtA variants with ThtB and subsequent analysis of the products by MALDI-TOF-MS were consistent with this hypothesis. While substitution of Thr8 completely abolished thatisin production, a detectable amount of a doubly modified product was observed for the T5A variant (**Figure S19**).

Given the presence of two macrolactone rings, two ring topologies are possible for thatisin: a ring-within-a-ring topology with Thr5-Asp17 and Thr8-Asp13 macrolactones or an interlocked set of macrolactones at Thr5-Asp13 and Thr8-Asp17. To distinguish between these two ring topologies,

we subjected thatisin to partial methanolysis and analyzed the products by HRMS/MS. HRMS/MS spectra of partially methanolized thatisin revealed a series of b- and y-ions consistent with Asp17 being the acceptor residue of the most methanolytic-susceptible macrolactone (**Figure S18**). No fragmentation was observed between Thr5 and Asp13 on the singly methanolized ion, suggesting that the unreacted macrolactone involved Thr5 and Asp13. This fragmentation pattern was consistent with thatisin displaying an interlocked ring topology with macrolactones at Thr5-Asp13 and Thr8-Asp17. Methanolysis and HRMS/MS analysis of thatisin-related products (with one or two modifications) from co-expression of wild-type ThtA and Ala-substituted ThtA-variants (i.e., T5A, T8A, D13A, and D17A) with ThtB (**Table S4**) provide additional support for the proposed ring-topology of thatisin: *(i)* biosynthetically immature Thr5-Asp13 linkage products were observed for wild-type thatisin (**Table S4, Figure S20**); *(ii)* ThtA variants T8A and D17A both formed Thr5-Asp13 linkage products (**Table S4, Figures S21-S22**); and *(iii)* ThtA variants T5A and D13A yielded an alternative, non-wild-type Thr8-Asp17 linked product (**Table S4, Figures S23-S24**).

To rigorously establish the ring connectivity of thatisin, we analyzed all thatisin-related peptides (with one or two modifications) produced upon co-expression of Asp to Ala and Thr to Ala ThtA variants by methanolysis and HRMS/MS. Positions not engaged in macrolactone formation (i.e., D3A, T4A, D14A, D19A) reacted with methanol and fragmented similarly to wild-type thatisin (**Table S4**, **Figures S25-S28**). However, Ala-substituted variants involving the acceptor and donor positions behaved differently, depending on the position. All four variants, T5A, T8A, D13A, and D17A, produced singly modified thatisin intermediates with linkages formed from the unperturbed set of available donor and acceptor residues (**Table S4, Figures S21-S24**). Two separate singly modified species corresponding to a T5-D13 linkage were observed via HPLC for both T8A and

D17A (**Figures S21-S22**) variants, whereas only a single T8-D17 species was observed for T5A and D13A (**Figures S23-S24**) variants. Variants T5A and D13A also utilized a neighboring residue for macrolactone installation in the absence of the native donor or acceptor. Variant T5A produced off-pathway single-modification products with Thr4-Asp13 and Thr4-Asp14 linkages; the Thr4-Asp13 was further processed into a thatisin congener with Thr4-Asp13 and Thr8-Asp17 linkages (**Figure S23**), although in minuscule amounts. Similarly, the D13A variant could utilize Asp14 as an acceptor to form a product with Thr5-Asp14 linkage (**Figures S24**); however, we could not detect further processing of this product to a thatisin congener.

Next, we sought to determine the modification type and ring topology of *iso*-thatisin. Methanolysis and HRMS/MS fragmentation pattern of *iso*-thatisin was identical to that of thatisin, suggesting the presence of two macrolactones and implicating Asp13 and Asp17 as the acceptor residues (**Figure S29**). Furthermore, the CID spectrum of *iso*-thatisin was also indistinguishable from that of thatisin (**Figure S29**), suggesting that the *iso*-thatisin also utilized Thr5 and Thr8 as the macrolactone donor residues. Given the utilization of the same set of acceptor and donor residues as thatisin, we considered the possibility that *iso*-thatisin adopts a ring-within-ring topology. To test this idea, we subjected *iso*-thatisin to an identical partial methanolysis treatment and analyzed the products by HRMS/MS. Partially methanolized *iso*-thatisin upon HRMS/MS, produced b- and y-ion series consistent with Asp13 as the acceptor residue of the most methanolysis-susceptible macrolactone (**Figure S29**). No fragmentation was observed between Thr8 and Asp17 on the singly methanolized ion suggesting that these residues were involved in the intact macrolactone. Although different from that of thatisin, this fragmentation pattern is also consistent with an interlocked ring topology formed via Thr5-Asp13 and Thr8-Asp17 linkages. While both thatisin and *iso*-thatisin have identical ring topology, they show differential susceptibility to partial

methanolysis, suggesting a difference in the conformation of the two molecules. Far UV circular dichroism (CD) spectroscopy of purified thatisin and *iso*-thatisin further supports a conformational difference. While thatisin has an ellipticity minimum at 207.2 nm, *iso*-thatisin has ellipticity minima at 204.3 and 205.7 nm upon analysis by CD spectroscopy (**Figure S30**). Altogether, reverse-phase HPLC, partial methanolysis, HRMS/MS, and CD spectroscopy data suggest that thatisin exhibits atropisomerism.

7.2.4 Atropisomerism in thatisin results from cis/trans Pro isomerization

We considered two potential possibilities that could lead to atropisomerism in thatisin, *cis/trans* Pro isomerization, and position of the $_{9}$GPNG$_{12}$-bridge. Unlike peptide bonds involving other amino acids, which predominantly adopt the *trans* configuration, Xaa-Pro (where Xaa is any amino acid) peptide bonds can adopt both *cis* and *trans* configurations.[38] This is attributed to the lower energy difference and energy barrier between the *trans* and *cis* forms of Xaa-Pro peptide bond (~-0.5 kcal/mol and ~13 kcal/mol, respectively) compared to Xaa-nonPro peptide bonds (~-2.5 kcal/mol and ~20 kcal/mol, respectively). Very few instances of *cis/trans* Pro isomerization have been observed among RiPPs. Microcin J25, an antimicrobial lasso peptide, can adopt both *cis* and *trans* configuration at Pro7 and Pro16.[39] Given the presence of three Pro residues (Pro10, Pro15, and Pro18) in thatisin, we hypothesized that one or more might undergo *cis*/*trans* isomerization. A second possibility is that the $_{9}$GPNG$_{12}$-bridge (part of both the macrocycles) can be positioned either above or below the plane of the macrocycles. Such "non-canonical" atropisomerism, where interconversion of the atropisomers would be theoretically possible but physically impossible through multiple bond rotations, was reported for tryptorubin A during total chemical synthesis (the natural product, however, is produced in an atropselective manner).[40] Lasso peptides and their branched cyclic peptide counterparts are also considered non-canonical atropisomers of each other.

To investigate these two possibilities, we attempted to characterize both thatisin and *iso*-thatisin by NMR spectroscopy. However, this was challenging since each purified compound could interconvert to the other (**Figure S31**). Upon equilibration at room temperature for 18 h, thatisin underwent conversion to ~ 8% *iso*-thatisin. Similarly, *iso*-thatisin also converted to ~8 % thatisin upon equilibration at room temperature. While *iso*-thatisin undergoes slow conversion to thatisin at 4° C, no interconversion was seen for thatisin at this temperature (**Figure S31**). Attempts to characterize the atropisomer mixture by NMR spectroscopy at 25 °C were unsuccessful; therefore, temperature control was utilized to characterize thatisin. After confirming that no conformational exchange occurred at 12 °C, two-dimensional ^{1}H-^{1}H TOCSY and ^{1}H-^{1}H NOESY experiments were performed on thatisin. Residue identities were assigned based on characteristic TOCSY coupling patterns, and sequence positions were assigned based on interresidue NOEs (**Figures S32-S34 and Table S5**). NOE intensity and identity between Asp14 and Pro15 residues confirmed that Pro15 occupied a *cis*-configuration at 12 °C. The assigned NOEs of Pro15 demonstrated strong Pro15$_{H\alpha}$-Asp14$_{H\alpha}$ and Pro15$_{H\delta}$-Asp14$_{H\beta1/2}$ contacts, characteristic of a *cis*-isomer for Asp14-Pro15 (**Figures 4 and S35**). Furthermore, the absence of a Pro15$_{H\delta}$- Asp14$_{H\alpha}$ NOE contact supports this configuration. Proximity of Gly9 and Asp17 to the water signal likely led to collateral suppression

of these signals, preventing observation of inter-residue NOEs to assign the configuration of Gly9-Pro10 and Asp17-Pro18 peptide bonds.

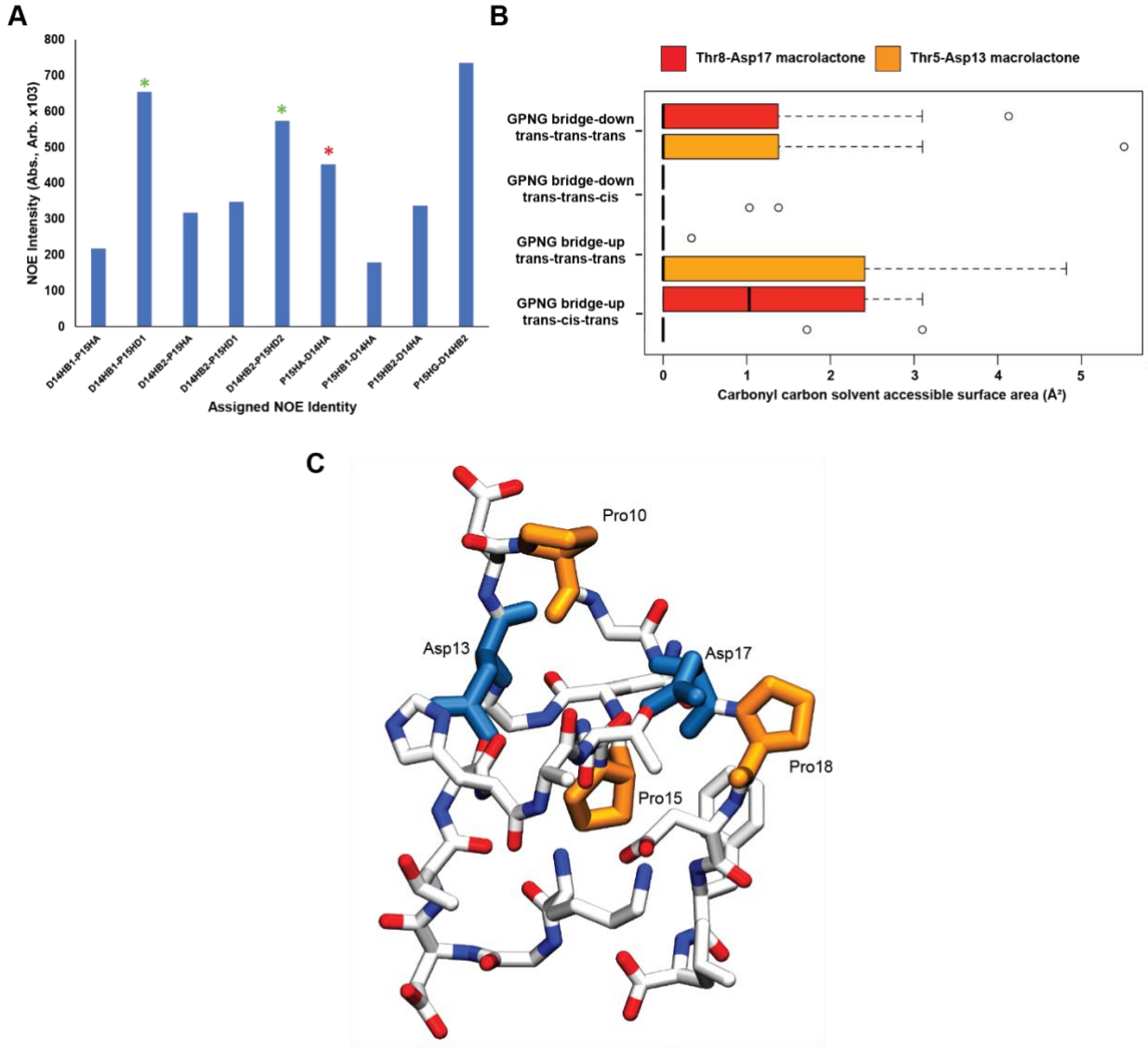


Figure 7.4. Structure of thatisin. A. Absolute NOE intensities between Asp14 and Pro15 indicate *cis* Pro15. Strong inter-residue H_{α}-H_{α} correlations suggest *cis*-isomerization of Pro15 (red asterisk), while the presence of $Pro15_{H\delta}$-$Asp14_{H\beta1/2}$ NOEs (green asterisks) and absence of $Pro15_{H\delta}$- $Asp14_{H\alpha}$ further support a *cis*-isomerization of Pro15. B. SASA box plot for thatisin conformations. Box plot of SASA for all converged structures with $_{9}GPNG_{12down}$ $P10_{trans}$-$P15_{trans}$-$P18_{trans}$, $_{9}GPNG_{12down}$ $P10_{trans}$-$P15_{trans}$-$P18_{trans}$, $_{9}GPNG_{12up}$ $P10_{trans}$-$P15_{trans}$-$P18_{trans}$, and

Figure 7.4 (cont.) $_{9}$GPNG$_{12up}$ P10*$_{trans}$*-*P15$_{cis}$*-*P18$_{trans}$* conformations. Raw data can be found in **Tables S7-S10**. Bold lines indicate the median SASA value and circles indicate outliers. The only conformation with a median SASA greater than 0 is GPNG-bridge-up *trans-cis-trans*. C) Computed structure for thatisin. D) Topology of thatisin.

We next undertook a computational approach to explore thatisin atropisomerism. For this, 16 different thatisin conformations were considered. These conformations differ in the position of the $_{9}$GPNG$_{12}$-bridge (up/down) and the configuration of the Xaa-Pro peptide bonds (*cis*/*trans*). Molecular Dynamic (MD) simulations were performed on all 16 energy-minimized thatisin conformations to determine whether 1) Pro could switch geometries 2) some conformations were preferred over the others. This analysis revealed that five conformations: $_{9}$GPNG$_{12up}$ P10*$_{trans}$*-*P15$_{trans}$*-*P18$_{trans}$*, $_{9}$GPNG$_{12up}$ P10*$_{trans}$*-*P15$_{cis}$*-*P18$_{trans}$*, $_{9}$GPNG$_{12up}$ P10*$_{trans}$*-*P15$_{trans}$*-*P18$_{cis}$*, $_{9}$GPNG$_{12down}$ P10*$_{trans}$*-*P15$_{trans}$*-*P18$_{trans}$*, $_{9}$GPNG$_{12down}$ P10*$_{trans}$*-*P15$_{trans}$*-*P18$_{cis}$* produced more converged structures, *i.e.,* they were more favored over the others (**Table S6**). Only one ($_{9}$GPNG$_{12up}$ P10*$_{trans}$*-*P15$_{cis}$*-*P18$_{trans}$*) of these 5 conformations features a *cis* Pro15, suggesting this as the most likely conformation of thatisin. To probe this further, we initiated solvent accessible surface area (SASA) calculations on the macrolactone carbonyl carbons of the four most favored conformations **(Tables S7-S10**). SASA analysis revealed that $_{9}$GPNG$_{12up}$ P10*$_{trans}$*-*P15$_{cis}$*-*P18$_{trans}$* is the only favored thatisin conformation that shows preferential methanol accessibility to Thr8-Asp17 carbonyl carbon (**Figure 4),** which is in line with the experimental observations for thatisin (**Figure S18**). Similarly, $_{9}$GPNG$_{12up}$ P10*$_{trans}$*-*P15$_{trans}$*-*P18$_{trans}$* shows methanol accessibility preferentially to Thr5-Asp13 carbonyl carbon (**Figure 4**), which was experimentally observed for *iso*-thatisin (**Figure S29**). Taken together, experimental (partial methanolysis, HRMS/MS, NMR

spectroscopy) and computational (MD simulations and SASA calculations) data suggest that atropisomerism observed in thatisin results from *cis/trans* Pro isomerization.

Next, we tested for growth suppression activity of thatisin against a brief panel of Gram-positive and Gram-negative bacteria; however, no such activity was detected under the conditions assayed.

7.3 Conclusion

Several cyanobacterial tricyclic RiPPs termed "microviridins", exhibiting potent serine protease inhibitory activity, were characterized over the past two decades. Until recently, microviridins were the only known RiPPs that featured the ATP-grasp ligase-mediated macrolactone/macrolactam modifications. The discovery of non-microviridin, ATP-grasp ligase-modified RiPPs, plesiocin, and thuringinin, in recent years, prompted the reclassification of the microviridin family of RiPPs to "OEPs" and now, "graspetides". While several genome mining studies show an impressive range of topological novelty among graspetides, the true breadth of possibilities within this family remained hidden due to a lack of bioinformatic tools to identify and analyze such BGCs readily. We expanded the predictive utility of RODEO to enable the confident prediction of graspetide BGCs and precursor peptides. The graspetide-specific scoring module combines the power of RODEO and RADAR to identify both short and long repeat-containing precursor peptides and uses machine learning and heuristics to score potential graspetide precursor peptides. Using this module, we surveyed the NCBI non-redundant database and uncovered 3,923 graspetide BGCs widespread among prokaryotes. Furthermore, bioinformatic analysis of the precursor peptides and the BGCs identified in this study reveals unforeseen diversity in sequence and potential ancillary modifications among graspetides. Utilizing this dataset, we characterized a novel graspetide from *Lysobacter antibioticus* called thatisin, which displays two interlocked macrolactone linkages. Spectroscopic and chromatographic techniques revealed that thatisin

exhibits atropisomerism, a property rarely observed among characterized RiPP natural products. Furthermore, NMR spectroscopy and computational analyses provide evidence that atropisomerism in thatisin stems from Pro *cis/trans* isomerization. Overall, this work captures the breadth of the graspetide RiPP family and we expect that it will assist in the prioritized characterization of topologically novel members in the future.

7.4 References

(1) Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; Cotter, P. D.; Craik, D. J.; Dawson, M.; Dittmann, E.; Donadio, S.; Dorrestein, P. C.; Entian, K.-D.; Fischbach, M. A.; Garavelli, J. S.; Göransson, U.; Gruber, C. W.; Haft, D. H.; Hemscheidt, T. K.; Hertweck, C.; Hill, C.; Horswill, A. R.; Jaspars, M.; Kelly, W. L.; Klinman, J. P.; Kuipers, O. P.; Link, A. J.; Liu, W.; Marahiel, M. A.; Mitchell, D. A.; Moll, G. N.; Moore, B. S.; Müller, R.; Nair, S. K.; Nes, I. F.; Norris, G. E.; Olivera, B. M.; Onaka, H.; Patchett, M. L.; Piel, J.; Reaney, M. J. T.; Rebuffat, S.; Ross, R. P.; Sahl, H.-G.; Schmidt, E. W.; Selsted, M. E.; Severinov, K.; Shen, B.; Sivonen, K.; Smith, L.; Stein, T.; Süssmuth, R. D.; Tagg, J. R.; Tang, G.-L.; Truman, A. W.; Vederas, J. C.; Walsh, C. T.; Walton, J. D.; Wenzel, S. C.; Willey, J. M.; van der Donk, W. A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108–160. https://doi.org/10.1039/C2NP20085F.

(2) Russell, A. H.; Truman, A. W. Genome Mining Strategies for Ribosomally Synthesised and Post-Translationally Modified Peptides. *Computational and Structural Biotechnology Journal* **2020**, *18*, 1838–1851. https://doi.org/10.1016/j.csbj.2020.06.032.

(3) Zhong, Z.; He, B.; Li, J.; Li, Y.-X. Challenges and Advances in Genome Mining of Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs). *Synthetic and Systems Biotechnology* **2020**, *5* (3), 155–172. https://doi.org/10.1016/j.synbio.2020.06.002.

(4) Montalbán-López, M.; Scott, T. A.; Ramesh, S.; Rahman, I. R.; van Heel, A. J.; Viel, J. H.; Bandarian, V.; Dittmann, E.; Genilloud, O.; Goto, Y.; Grande Burgos, M. J.; Hill, C.; Kim, S.; Koehnke, J.; Latham, J. A.; Link, A. J.; Martínez, B.; Nair, S. K.; Nicolet, Y.; Rebuffat, S.; Sahl, H.-G.; Sareen, D.; Schmidt, E. W.; Schmitt, L.; Severinov, K.; Süssmuth, R. D.; Truman, A. W.; Wang, H.; Weng, J.-K.; van Wezel, G. P.; Zhang, Q.; Zhong, J.; Piel, J.; Mitchell, D. A.; Kuipers, O. P.; van der Donk, W. A. New Developments in RiPP Discovery, Enzymology and Engineering. *Nat. Prod. Rep.* **2021**, 10.1039.D0NP00027B. https://doi.org/10.1039/D0NP00027B.

(5) Ziemert, N.; Ishida, K.; Liaimer, A.; Hertweck, C.; Dittmann, E. Ribosomal Synthesis of Tricyclic Depsipeptides in Bloom-Forming Cyanobacteria. *Angewandte Chemie International Edition* **2008**, *47* (40), 7756–7759. https://doi.org/10.1002/anie.200802730.

(6) Philmus, B.; Christiansen, G.; Yoshida, W. Y.; Hemscheidt, T. K. Post-Translational Modification in Microviridin Biosynthesis. *ChemBioChem* **2008**, *9* (18), 3066–3073. https://doi.org/10.1002/cbic.200800560.

(7) Mitchell, A. L.; Attwood, T. K.; Babbitt, P. C.; Blum, M.; Bork, P.; Bridge, A.; Brown, S. D.; Chang, H.-Y.; El-Gebali, S.; Fraser, M. I.; Gough, J.; Haft, D. R.; Huang, H.; Letunic, I.; Lopez, R.; Luciani, A.; Madeira, F.; Marchler-Bauer, A.; Mi, H.; Natale, D. A.; Necci, M.; Nuka, G.; Orengo, C.; Pandurangan, A. P.; Paysan-Lafosse, T.; Pesseat, S.; Potter, S. C.; Qureshi, M. A.; Rawlings, N. D.; Redaschi, N.; Richardson, L. J.; Rivoire, C.; Salazar, G. A.; Sangrador-Vegas, A.; Sigrist, C. J. A.; Sillitoe, I.; Sutton, G. G.; Thanki, N.; Thomas, P. D.; Tosatto, S. C.

E.; Yong, S.-Y.; Finn, R. D. InterPro in 2019: Improving Coverage, Classification and Access to Protein Sequence Annotations. *Nucleic Acids Research* **2019**, *47* (D1), D351–D360. https://doi.org/10.1093/nar/gky1100.

(8) Fawaz, M. V.; Topper, M. E.; Firestine, S. M. The ATP-Grasp Enzymes. *Bioorg. Chem.* **2011**, *39* (5–6), 185–191. https://doi.org/10.1016/j.bioorg.2011.08.004.

(9) Sieber, S.; Grendelmeier, S. M.; Harris, L. A.; Mitchell, D. A.; Gademann, K. Microviridin 1777: A Toxic Chymotrypsin Inhibitor Discovered by a Metabologenomic Approach. *J. Nat. Prod.* **2020**, *83* (2), 438–446. https://doi.org/10.1021/acs.jnatprod.9b00986.

(10) Unno, K.; Kaweewan, I.; Nakagawa, H.; Kodani, S. Heterologous Expression of a Cryptic Gene Cluster from Grimontia Marina Affords a Novel Tricyclic Peptide Grimoviridin. *Appl Microbiol Biotechnol* **2020**. https://doi.org/10.1007/s00253-020-10605-z.

(11) Ahmed, M. N.; Reyna-González, E.; Schmid, B.; Wiebach, V.; Süssmuth, R. D.; Dittmann, E.; Fewer, D. P. Phylogenomic Analysis of the Microviridin Biosynthetic Pathway Coupled with Targeted Chemo-Enzymatic Synthesis Yields Potent Protease Inhibitors. *ACS Chemical Biology* **2017**, *12* (6), 1538–1546. https://doi.org/10.1021/acschembio.7b00124.

(12) Zhang, Y.; Li, K.; Yang, G.; McBride, J. L.; Bruner, S. D.; Ding, Y. A Distributive Peptide Cyclase Processes Multiple Microviridin Core Peptides within a Single Polypeptide Substrate. *Nature Communications* **2018**, *9* (1), 1780. https://doi.org/10.1038/s41467-018-04154-3.

(13) Kanaori, K.; Kamei, K.; Taniguchi, M.; Koyama, T.; Yasui, T.; Takano, R.; Imada, C.; Tajima, K.; Hara, S. Solution Structure of Marinostatin, a Natural Ester-Linked Protein Protease Inhibitor ‡. *Biochemistry* **2005**, *44* (7), 2462–2468. https://doi.org/10.1021/bi048034x.

(14) Iyer, L. M.; Abhiman, S.; Maxwell Burroughs, A.; Aravind, L. Amidoligases with ATP-Grasp, Glutamine Synthetase-like and Acetyltransferase-like Domains: Synthesis of Novel Metabolites and Peptide Modifications of Proteins. *Molecular BioSystems* **2009**, *5* (12), 1636. https://doi.org/10.1039/b917682a.

(15) Lee, H.; Park, Y.; Kim, S. Enzymatic Cross-Linking of Side Chains Generates a Modified Peptide with Four Hairpin-like Bicyclic Repeats. *Biochemistry* **2017**, *56* (37), 4927–4930. https://doi.org/10.1021/acs.biochem.7b00808.

(16) Roh, H.; Han, Y.; Lee, H.; Kim, S. A Topologically Distinct Modified Peptide with Multiple Bicyclic Core Motifs Expands the Diversity of Microviridin-Like Peptides. *ChemBioChem* **2019**, *20* (8), 1051–1059. https://doi.org/10.1002/cbic.201800678.

(17) Lee, H.; Choi, M.; Park, J.-U.; Roh, H.; Kim, S. Genome Mining Reveals High Topological Diversity of ω-Ester-Containing Peptides and Divergent Evolution of ATP-Grasp Macrocyclases. *J. Am. Chem. Soc.* **2020**, jacs.9b12076. https://doi.org/10.1021/jacs.9b12076.

(18) Unno, K.; Kodani, S. Heterologous Expression of Cryptic Biosynthetic Gene Cluster from Streptomyces Prunicolor Yields Novel Bicyclic Peptide Prunipeptin. *Microbiological Research* **2021**, *244*, 126669. https://doi.org/10.1016/j.micres.2020.126669.

(19) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Research* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(20) Skinnider, M. A.; Merwin, N. J.; Johnston, C. W.; Magarvey, N. A. PRISM 3: Expanded Prediction of Natural Product Chemical Structures from Microbial Genomes. *Nucleic Acids Research* **2017**, *45* (W1), W49–W54. https://doi.org/10.1093/nar/gkx320.

(21) Santos-Aberturas, J.; Chandra, G.; Frattaruolo, L.; Lacret, R.; Pham, T. H.; Vior, N. M.; Eyles, T. H.; Truman, A. W. Uncovering the Unexplored Diversity of Thioamidated Ribosomal Peptides in Actinobacteria Using the RiPPER Genome Mining Tool. *Nucleic Acids Research* **2019**. https://doi.org/10.1093/nar/gkz192.

(22) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. A New Genome-Mining Tool Redefines the Lasso Peptide Biosynthetic Landscape. *Nature Chemical Biology* **2017**, *13* (5), 470–478. https://doi.org/10.1038/nchembio.2319.

(23) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. Bioinformatic Expansion and Discovery of Thiopeptide Antibiotics. *Journal of the American Chemical Society* **2018**, *140* (30), 9494–9501. https://doi.org/10.1021/jacs.8b03896.

(24) Hudson, G. A.; Burkhart, B. J.; DiCaprio, A. J.; Schwalen, C. J.; Kille, B.; Pogorelov, T. V.; Mitchell, D. A. Bioinformatic Mapping of Radical *S* -Adenosylmethionine-Dependent Ribosomally Synthesized and Post-Translationally Modified Peptides Identifies New Cα, Cβ, and Cγ-Linked Thioether-Containing Peptides. *J. Am. Chem. Soc.* **2019**, jacs.9b01519. https://doi.org/10.1021/jacs.9b01519.

(25) Walker, M. C.; Eslami, S. M.; Hetrick, K. J.; Ackenhusen, S. E.; Mitchell, D. A.; van der Donk, W. A. Precursor Peptide-Targeted Mining of More than One Hundred Thousand Genomes Expands the Lanthipeptide Natural Product Family. *BMC Genomics* **2020**, *21* (1), 387. https://doi.org/10.1186/s12864-020-06785-7.

(26) Georgiou, M. A.; Dommaraju, S. R.; Guo, X.; Mitchell, D. A. *Bioinformatic and Reactivity-Based Discovery of Linaridins*; preprint; Biochemistry, 2020. https://doi.org/10.1101/2020.07.09.196543.

(27) Kloosterman, A. M.; Shelton, K. E.; van Wezel, G. P.; Medema, M. H.; Mitchell, D. A. RRE-Finder: A Genome-Mining Tool for Class-Independent RiPP Discovery. *mSystems* **2020**, *5* (5), e00267-20, /msystems/5/5/msys.00267-20.atom. https://doi.org/10.1128/mSystems.00267-20.

(28) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25* (17), 3389–3402.

(29) Heger, A.; Holm, L. Rapid Automatic Detection and Alignment of Repeats in Protein Sequences. *Proteins* **2000**, *41* (2), 224–237. https://doi.org/10.1002/1097-0134(20001101)41:2<224::aid-prot70>3.0.co;2-z.

(30) DiCaprio, A. J.; Firouzbakht, A.; Hudson, G. A.; Mitchell, D. A. Enzymatic Reconstitution and Biosynthetic Investigation of the Lasso Peptide Fusilassin. *J. Am. Chem. Soc.* **2019**, *141* (1), 290–297. https://doi.org/10.1021/jacs.8b09928.

(31) Gerlt, J. A.; Bouvier, J. T.; Davidson, D. B.; Imker, H. J.; Sadkhin, B.; Slater, D. R.; Whalen, K. L. Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A Web Tool for Generating Protein Sequence Similarity Networks. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2015**, *1854* (8), 1019–1037. https://doi.org/10.1016/j.bbapap.2015.04.015.

(32) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13* (11), 2498–2504. https://doi.org/10.1101/gr.1239303.

(33) Griffith, S. C.; Sawaya, M. R.; Boutz, D. R.; Thapar, N.; Katz, J. E.; Clarke, S.; Yeates, T. O. Crystal Structure of a Protein Repair Methyltransferase from Pyrococcus Furiosus with Its l -Isoaspartyl Peptide Substrate 1 1Edited by I. A. Wilson. *Journal of Molecular Biology* **2001**, *313* (5), 1103–1116. https://doi.org/10.1006/jmbi.2001.5095.

(34) Acedo, J. Z.; Bothwell, I. R.; An, L.; Trouth, A.; Frazier, C.; van der Donk, W. A. *O* -Methyltransferase-Mediated Incorporation of a β-Amino Acid in Lanthipeptides. *J. Am. Chem. Soc.* **2019**, *141* (42), 16790–16801. https://doi.org/10.1021/jacs.9b07396.

(35) Mahanta, N.; Hudson, G. A.; Mitchell, D. A. Radical *S* -Adenosylmethionine Enzymes Involved in RiPP Biosynthesis. *Biochemistry* **2017**, *56* (40), 5229–5244. https://doi.org/10.1021/acs.biochem.7b00771.

(36) Bobeica, S. C.; Dong, S.-H.; Huo, L.; Mazo, N.; McLaughlin, M. I.; Jiménez-Osés, G.; Nair, S. K.; van der Donk, W. A. Insights into AMS/PCAT Transporters from Biochemical and Structural Characterization of a Double Glycine Motif Protease. *eLife* **2019**, *8*. https://doi.org/10.7554/eLife.42305.

(37) Mahanta, N.; Liu, A.; Dong, S.; Nair, S. K.; Mitchell, D. A. Enzymatic Reconstitution of Ribosomal Peptide Backbone Thioamidation. *Proc Natl Acad Sci USA* **2018**, *115* (12), 3030–3035. https://doi.org/10.1073/pnas.1722324115.

(38) Dong, S.-H.; Liu, A.; Mahanta, N.; Mitchell, D. A.; Nair, S. K. Mechanistic Basis for Ribosomal Peptide Backbone Modifications. *ACS Cent. Sci.* **2019**, acscentsci.9b00124. https://doi.org/10.1021/acscentsci.9b00124.

(39) Pierson, N. A.; Chen, L.; Russell, D. H.; Clemmer, D. E. *Cis – Trans* Isomerizations of Proline Residues Are Key to Bradykinin Conformations. *J. Am. Chem. Soc.* **2013**, *135* (8), 3186–3192. https://doi.org/10.1021/ja3114505.

(40) Jeanne Dit Fouque, K.; Hegemann, J. D.; Zirah, S.; Rebuffat, S.; Lescop, E.; Fernandez-Lima, F. Evidence of Cis/Trans-Isomerization at Pro7/Pro16 in the Lasso Peptide Microcin J25. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (6), 1038–1045. https://doi.org/10.1007/s13361-019-02134-5.

(41) Reisberg, S. H.; Gao, Y.; Walker, A. S.; Helfrich, E. J. N.; Clardy, J.; Baran, P. S. Total Synthesis Reveals Atypical Atropisomerism in a Small-Molecule Natural Product, Tryptorubin A. *Science* **2020**, *367* (6476), 458–463. https://doi.org/10.1126/science.aay9981.