Published in final edited form as:
Chemistry. 2013 June 10; 19(24): . doi:10.1002/chem.201300401.

Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products: New Insights Into the Role of Leader and Core Peptides During Biosynthesis

Xiao Yang[b] and **Prof. Wilfred A. van der Donk**[*,a]

[a]W. A. van der Donk, Department of Chemistry and Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, 600 S Mathews Ave, Urbana, Illinois, 61801, USA, Fax: (+1)217-244-8533

[b]X. Yang, Department of Chemistry, University of Illinois at Urbana-Champaign, 600 S Mathews Ave, Urbana, Illinois, 61801, USA

Abstract

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are a major class of natural products with a high degree of structural diversity and a wide variety of bioactivities. Understanding the biosynthetic machinery of these RiPPs will benefit the discovery and development of new molecules with potential pharmaceutical applications. In this review, we discuss the features of the biosynthetic pathways to different RiPP classes, and propose mechanisms regarding recognition of the precursor peptide by the posttranslational modification enzymes. We propose that the leader peptides function as allosteric regulators that bind the active form of the biosynthetic enzymes in a conformational selection process. We also speculate how enzymes that generate polycyclic products of defined topologies may have been selected for during evolution.

Keywords

natural products; peptides; biosynthesis

Introduction

Natural products have played prominent roles in science and medicine over the past century. The complex and diverse chemical scaffolds of natural products have inspired organic chemists to devise new methodologies and new strategies for total synthesis,[1] their biological activities have served as tools in cell biology,[2] and they have been important lead compounds for development of new pharmaceuticals.[3]

Ribosomally synthesized and post-translationally modified peptides (RiPPs) are natural products that have only recently been recognized as a major class of compounds as a result of the genome sequencing efforts of the past decade.[4] Because of their extensive posttranslational modifications (PTMs), RiPPs have greater structural diversity and more rigid structures compared to linear peptides,[5] and the PTMs endow them with expanded chemical functionalities, improved target recognition, and increased metabolic and chemical stability.[4]

. vddonk@illinois.edu.

RiPP biosynthesis is initiated with a ribosomally generated precursor peptide encoded by a structural gene. This precursor peptide usually contains an N-terminal leader peptide (Figure 1) that is important for recognition by PTM enzymes and for export from the cell. This leader sequence is fused to a core peptide that is transformed into the final natural product. In some cases a C-terminal recognition sequence (RS) that is important for excision and cyclization of the core peptide is also present (Figure 1). Upon binding of the precursor peptide to the modifying enzyme, various post-translational modifications are installed in the core peptide. The leader peptide is usually removed by proteolytic cleavage in a late step of the maturation process, which yields the final natural product.

This review will discuss new insights into the roles of the leader and core peptides during post-translational modification of bacterial RiPPs.

Overview of RiPP Structures

Leader peptides have been proposed to play multiple roles during RiPP biosynthesis, including acting as a secretion signal, a recognition motif for the PTM enzymes, a chaperone-like sequence that assists precursor peptide folding and stabilization, or a protective sequence that keeps the precursor peptide inactive inside the host until secretion and proteolysis.[6] Among these potential roles, the recognition and modification of the precursor sequence by the biosynthetic machinery is of particular interest. A better understanding of this process will benefit the engineering of RiPP natural products, which could potentially lead to increased biological activity and/or stability. Although at present the molecular details of the interaction of the peptide substrates with the PTM machinery are still mostly unknown, reports published in recent years have started to provide insights into the mechanism(s) of leader peptide recognition and enzyme activation. In the next sections, we will briefly present the PTMs of the RiPP classes produced by bacteria for which information on leader peptide recognition is available or that have unusual leader peptides. For other RiPP classes, such as the thiopeptides and glycopeptides as well as RiPPs from higher organisms, we refer to a recent comprehensive review.[4] Subsequently, we will address five questions: 1) is the leader peptide always required for the activity of the biosynthetic enzymes, 2) how does leader peptide binding activate these enzymes, 3) how do the biosynthetic enzymes recognize the leader peptide, 4) do the biosynthetic enzymes also recognize the core peptide sequences, and 5) how can one enzyme make rings of different sizes, sequences, and even stereochemistry?

Lanthipeptides

Lanthipeptides are a large group of polycyclic RiPPs classified by their intramolecular thioether crosslinks named lanthionine (Lan) and methyllanthionine (MeLan) (Figure 2A).[7] Lanthipeptides with antimicrobial activities are called lantibiotics.[8] The polycyclic structure of lantibiotics constrains the conformational flexibility of the peptides, thereby conferring improved affinity for their targets, which thus far have been small molecules rather than macromolecules. One well-studied example is nisin produced by certain strains of *Lactococcus lactis*, which possesses high antimicrobial potency against a wide range of gram-positive bacteria (Figure 2B). This natural product contains five thioether rings formed by one Lan and four MeLan residues[9] that are critical for its antimicrobial activity.[10]

The installation of the thioether residues in lanthipeptides is achieved in a two-step process: dehydration of Ser/Thr residues to dehydroalanine (Dha) and dehydrobutyrine (Dhb), and stereoselective intramolecular addition of Cys thiols to the resulting dehydro amino acids (Figure 2). The lanthipeptides are categorized into four classes according to the types of (Me)Lan synthetases.[7f] Class I lanthipeptides, including nisin, utilize a dehydratase (LanB) and a cyclase (LanC) to generate (Me)Lan residues (Figure 2B); Class II lanthipeptides use a

single bifunctional synthetase (LanM).[11] Most of the well-studied lanthipeptides discussed in this review belong to these two classes. The other two more recently discovered classes (III and IV) both feature trifunctional synthetases that contain an N-terminal phospholyase domain, a central kinase domain, and a C-terminal cyclase domain.[12] Interestingly, a subset of the class III synthetases produce unique carbocyclic labionin residues (Figure 2A).[13] After the introduction of the thioether crosslinks, the modified core peptide is typically exported from the producing cell through an ABC transporter (LanT), and the leader peptide is cleaved by a protease LanP (class I lanthipeptides) or the protease domain of LanT (class II lanthipeptides).[14]

Lasso peptides

Lasso peptides are characterized by a unique structure that assembles a threaded lasso (Figure 3).[15] These peptides usually consist of 16-21 residues in which the N-terminal amine and the side chain carboxylate of a Glu/Asp residue at position 8 or 9 form a macrolactam, through which the C-terminal peptide tail is threaded. In class I and class III lasso peptides, this constrained lasso structure is reinforced by one (class III) or two (class I) additional disulfide bonds that connect the lasso ring to the peptide tail, whereas in class II lasso peptides, bulky residues are often located close to the lasso that sterically lock the threaded conformation (Figure 3B).[16] Structural analysis has also revealed that the lasso rings of all tested peptides wrap their tail portion in a right-handed conformation (Figure 3C).[15] This unique structure provides lasso peptides with remarkable stability against chemical and enzymatic degradation, and imparts diverse bioactivities.[15, 17] Disruption of the lasso structure in microcin J25 (MccJ25) showed that it is a prerequisite for its antimicrobial activity.[18]

Thus far, the biosynthetic machinery of lasso peptides is best understood for MccJ25[19] and capistruin.[20] Using MccJ25 as a prototypical example, the *mcjA* gene encodes for a precursor peptide that contains a 37-amino acid leader peptide and a 21-amino acid core sequence. The McjB enzyme is an ATP-dependent cysteine protease[19a] that cleaves off the leader sequence.[19c, 19d] The McjC enzyme is an Asn synthetase homolog[19b] that adenylates the side chain carboxylic acid of Glu8 and catalyzes lactam formation.[19a-c] The production of the correct lasso fold requires the presence of both McjB and McjC, indicating that they are functionally interdependent.[19a, 19d] In the last step, the McjD enzyme, an ABC transporter, is believed to export the final product from the cytoplasm.[19a, 19c]

Linear azol(in)e containing peptides

Linear azol(in)e containing peptides (LAPs) are non-macrocyclized RiPPs featuring multiple thiazole and (methyl)oxazole heterocycles, and sometimes their corresponding 2-electron reduced azolines (e.g. Figure 4), which conformationally constrain the peptide.[4, 21] LAP family members exhibit various bioactivities, such as the DNA gyrase inhibitor microcin B17,[22] the β hemolytic factor streptolysin S,[23] and the antibiotic plantazolicin.[24] The thiazol(in)e and oxazol(in)e heterocycles are critical for the biological functions of the LAPs.[21]

LAPs are generated from a precursor peptide comprised of a leader sequence and a core peptide rich in residues with a β-nucleophile (Ser, Thr and Cys). Typically, a subset of these residues are modified to form azol(in)e rings. The first step is the ATP-dependent cyclodehydration of Ser, Thr, and Cys to produce azoline heterocycles by a cyclodehydratase protein complex (C and D proteins).[22b, 25] Recent work has shown that the D protein, a member of the YcaO/DUF181 protein family, is responsible for the cyclodehydration reaction and uses ATP to phosphorylate the peptide backbone.[26] In the second step, a subset or all of the azoline rings are oxidized to the aromatic azoles by a

flavin-dependent dehydrogenase designated the B-protein.[22b, 27] Subsequently, the leader peptide is proteolytically removed and the mature LAP is exported from the cell.

Cyanobactins

Cyanobacticins are a group of head-to-tail macrocyclized peptides produced by various cyanobacteria, many of which are further modified with azol(in)e heterocycles and prenylated Ser, Thr, or Tyr residues (Figure 5). These cyclic peptides have various biological activities that have recently been reviewed.[28]

The cyanobactin precursor peptides contain a leader sequence and multiple cassettes that can contain different core sequences.[29] In each cassette, the core sequence is sandwiched between two flanking recognition sequences: a N-terminal protease recognition sequence that typically consist of G(L/V)E(A/P)S, and a C-terminal recognition sequence that contains AYDG(E).[30]

The first step is the generation of azol(in)e rings from Ser, Thr and Cys residues catalyzed by the cyclodehydratase D (and the oxidase domain of another protein designated G), which shares similarity with the LAP biosynthetic machinery.[22b, 29b, 31] Subsequently, a serine protease (the A protein) removes the N-terminal recognition sequence, generating a free amine. Next, a second serine protease (the G protein) recognizes the C-terminal recognition sequence and removes it to form an acyl enzyme intermediate, that then is attacked by the N-terminus of the peptide to achieve macrocyclization.[30, 32] Prenylation occurs after the cyclization step, and it has been recently shown that the prenyltransferase only acts on the cyclized peptide in the absence of the leader sequence.[33]

Bottromycins

Bottromycins represent a class of heavily modified RiPPs that contain unusual structures including a macrocyclic amidine, a decarboxylated C-terminal thiazole, and multiple carbon-methylated amino acids (Figure 6). Bottromycins and their derivatives exhibit potent antimicrobial activity against bacterial pathogens such as methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant enterococci (VRE).[34]

In contrast to other RiPPs characterized thus far, bottromycin does not have an N-terminal leader peptide. Instead, a 37-residue sequence is attached to the C-terminus of the core peptide as a follower sequence, which is believed to have a similar function to the leader peptide.[35] Multiple radical S-adenosylmethionine (SAM) methyltransferases catalyze the methylation of Pro, Phe and Val residues, while an O-methyltransferase is responsible for the methylation of an Asp residue. Another two proteins with sequence similarity to the D-protein in LAP biosynthesis are proposed to be involved in thiazoline formation, and perhaps the macrocyclodehydration reaction that generates the amidine structure. Other genes in the cluster encode several proteases and a cytochrome P450 enzyme that are likely involved in removal of the N-terminal Met and the follower peptide, as well as oxidative decarboxylation of the C-terminal residue.[35]

Microviridins

Microviridins are a class of N-acetylated polycyclic peptides mostly produced by cyanobacteria. Microviridins are characterized by ester and amide crosslinks (Figure 7). These lactone or lactam structures are formed between the carboxyl groups of Asp/Glu and the hydroxyl groups of Ser/Thr, or with the amino groups of Lys, respectively.[36] The precursor peptide of microviridins contains an N-terminal leader sequence and a C-terminal core sequence. To date, the biosynthetic process has been partially elucidated.[37] The ester and amide bonds are introduced by two classes of ATP grasp ligases in a strictly ordered

process. The ester bonds are generated prior to amide formation, with the larger lactone forming first and the smaller one forming second.[37c] After leader peptide removal, acetylation of the N-terminus yields the mature microviridin.

Sactipeptides

Sactipeptides are characterized by thioether crosslinks, but unlike lanthipeptides, which have thioether crosslinks between Cys and the β-carbon of formerly Ser and Thr residues, the thioether bridges in sactipeptides are generated between cysteine and the α-carbon of a variety of amino acids (Figure 8).[38] These thioether bridges fold the backbone of sactipeptides into a hairpin-like structure. Structure determination has revealed that several sactipeptides, such as subtilosin A,[38-39] thuricin CD,[40] and thuricin H,[41] possess an amphipathic structure, which is likely responsible for their antimicrobial bioactivities.[41-42] The precursor peptide of sactipeptides contains an N-terminal leader sequence. During the biosynthesis, thioether bond formation is catalyzed by a radical SAM enzyme in a leader peptide-dependent manner.[43]

Proteusins

Proteusins are a very recently classified RiPP family with polytheonamides as their first characterized members.[4],[46a] Polytheonamides are extensively modified peptides containing many unusual residues such as a novel N-acyl moiety, *tert*-leucines and other carbon-methylated residues, as well as multiple D-configured amino acids that alternate with L-amino acids throughout the peptide backbone (Figure 9).[44] Due to their unique structures and strong hydrophobicity, polytheonamides are able to form helical structures that insert into the cell membrane as single molecule channels.[45]

The precursor peptides possess a nitrile hydratase-like leader peptide (NHLP) or a Nif11 nitrogen fixing protein-like leader peptide (N11P).[46] The D-configured amino acids are generated by the epimerase PoyD, and the Asn N-methylation is catalyzed by the SAM-dependent methyltransferase PoyE.[46a] PoyF, which resembles the dehydratase domain of the class II lanthipeptide synthetase, LanM, is responsible for the first step in generation of the acylated N-terminus by dehydrating a Thr residue.[46a]

Role(s) of Leader and Core Peptides

Is the leader peptide always required for the biosynthetic enzymes?

The leader peptides play an important role in many but not all steps of RiPP biosynthesis. With respect to the introduction of the characteristic PTMs of each compound class, the leader peptide has been shown to be very important because deletion of the leader sequence often results in reduced or abolished enzymatic formation of these PTMs.[6a] However, many compound-specific posttranslational modifications have been shown not to require the leader peptide. For instance, C-terminal decarboxylation,[47] hydroxylation,[48] and oxygenation[49] in lanthipeptides, as well as prenylation of cyanobactins[33] takes place in the absence of the leader peptide. Of course, most post-translational modification reactions that require the free N-terminus of the core peptide (and hence take place after removal of the leader peptide) also do not require the leader peptide such as head-to-tail cyclization of cyanobactins[30, 32a, 32b] or N-terminal modifications in lanthipeptides.[50] In the remainder of this review, we will only focus on the formation of the PTMs for which the leader peptide has been shown to play a role.

How does leader peptide binding activate enzyme activity?

In one possible model, the leader peptide could induce some type of conformational change to activate the biosynthetic enzymes for catalysis. Indeed, this model is supported by several

experimental observations. For instance, a number of studies have illustrated that “random” peptides attached to leader peptides are acted upon by the biosynthetic enzymes. In the case of lanthipeptides, various non-natural peptides that were attached to the nisin leader peptide could be dehydrated, cyclized and secreted by the nisin modification and export system in *L. lactis*.[51] Examples included analogues of small therapeutic peptides with engineered thioether rings,[52] and a randomized hexapeptide library with Dha/Dhb introduced at a fixed position.[53] Similarly, the class II lanthionine synthetase LctM was shown to dehydrate and cyclise non-lantibiotic peptides attached to the leader peptide of its substrate LctA.[54] Non-natural cyclic peptides have also been generated by the cyclization machinery of cyanobactins and lasso peptides by insertion of non-natural peptide sequence in pace of native core peptides,[31],[55] and several RiPP biosynthetic enzymes have been shown to accept chimeric peptides with the leader peptide of one compound and core peptide of another (*vide infra*). These investigations show the very high substrate tolerance with respect to the core peptide and seem to require a specific activation mechanism to prevent the biosynthetic enzymes to act on just any peptide in the RiPP-producing cell.

More recently, an alternative model has been put forth in which the leader peptide does not induce a conformational change in the biosynthetic enzymes but instead traps an active conformation that is present in very small amounts in the absence of the leader peptide (Figure 10). This alternative model was first introduced when it was observed that *in vitro*, the lanthipeptide synthetase LctM that generates lacticin 481 (Figure 11) was able to very slowly dehydrate the core peptide of its substrate LctA in the absence of a leader peptide. If leader peptide binding were required to actively induce a productive conformation of the LctM enzyme, no activity should have been observed. However, if a small concentration of an active conformation of LctM were always present, it could explain a low basal level of activity even in the absence of leader peptide. Support for such a model also operating inside the producing organism is the recent observation that co-expression of just the core peptide of the nisin precursor peptide NisA with the dehydratase NisB and the cyclase NisC in *L. lactis* also results in a low level of posttranslational modification of the core peptide.[56]

The model in Figure 10 also explains the observations in several studies in which the leader peptide was provided to the biosynthetic enzymes but was not covalently attached to the core peptide. *In trans* addition of the leader and core peptides of the precursor LctA to LctM resulted in dehydration *in vitro* that was enhanced compared to when the leader peptide was absent.[57] Furthermore, increasing the concentration of the leader peptide resulted in complete processing of the core peptide and the formation of bioactive lacticin 481 (for structure, see Figure 11).[58] The most active system was obtained when the leader peptide was fused to the LctM enzyme with a $(GlySer)_{15}$ linker. This fusion enzyme was able to convert the LctA core peptide into lacticin 481 with much improved efficiency, presumably because of the increased effective concentration of the leader peptide.[58] The *in trans* activation of the leader peptide for core peptide modification was also observed recently in a study in which the leader and core peptides for nisin were co-expressed independently in a *L. lactis* strain containing the NisB dehydratase and NisC cyclase.[56]

In trans activity has now been observed in the case of the class I lanthipeptide nisin,[56] class II lanthipeptide lacticin 481,[57-58] and class III lanthipeptide catenulipeptin,[59] suggesting it may be a general property of lanthipeptide synthetases.

Whether these observations will translate to all RiPP classes remains to be established. Most investigations that have concluded that the leader peptides are absolutely indispensable have been conducted in bacterial cells, and it may be that the amount of processing of core peptides without leader peptides attached is so small that it has not been detected or that the core peptides were not expressed or were degraded. A case in point is the observation that

formation of the lasso peptide MccJ25 in *E. coli* requires the last eight residues of the leader peptide[60] whereas a recent *in vitro* study concluded that a very low level of lasso peptide formation occurs by incubation of just the core peptide with the two biosynthetic enzymes McjB and McjC.[19d] The same study also demonstrated *in vitro* activity when the leader and core peptides were provided *in trans*. In contrast, some *in vitro* reconstituted RiPP biosynthetic systems did not display *in trans* activity such as the class III lanthipeptide labyrinthopeptin A2[61] (Figure 11) and the LAP microcin B17 (Figure 4).[62] It is possible that their biosynthetic enzymes are activated by the leader peptide in a different manner than the model in Figure 10, or that perhaps the *in vitro* studies could not achieve sufficiently high concentrations of the leader and core peptides to generate the required ternary complex. Indeed, for many RiPPs the solubility of the precursor peptides is very poor and limits the concentrations of core and leader peptides that can be attained.

The model in Figure 10 is consistent with many studies on conformational behaviour of enzymes and their relationship to catalysis that have been reported over the past decade.[63] These studies, which have utilized NMR methods to detect even very small concentrations of different enzyme conformations within bulk populations, have provided an alternative to the classical Koshland “induced fit” hypothesis. The latter is not unlike the model of leader peptide binding inducing an active form of the PTM enzymes described above. In the more recent “conformational selection” model,[63a] active and inactive protein conformations have been observed to both exist in an ensemble, and a ligand, whether substrate or an allosteric effector, selects the conformation that it has the highest affinity for. In doing so, substrate or allosteric effectors can shift the distribution of conformations within the ensemble.[63a] The proposed role of the leader peptide in Figure 10 is analogous. The leader peptide seeks out the active form of the enzyme, which in the figure is arbitrarily shown as an open form but it could also be a closed form with separate leader and core binding pockets.

Figure 10 also illustrates an additional possible role of the enzyme, which is that of a pseudosubstrate or intrasteric regulator.[64] A pseudosubstrate is part of an enzyme that resembles the substrate and that binds to the substrate binding site, thus keeping the ensemble mostly composed of the inactive enzyme form. In the case of the RiPP biosynthetic enzymes, the proteins could contain a domain (a “lid” domain) that docks onto the leader peptide binding site in the absence of substrate, thus keeping the enzyme mostly inactive. Because the binding affinity of the lid domain cannot be very tight (otherwise no catalysis would be possible), a small population of the enzyme would have the lid open and allow association of the leader peptide to its binding site. In turn, leader peptide binding would then shift the equilibrium towards the active form of the PTM enzyme. We emphasize that the model could be readily expanded to more complex models.[64b] For instance the leader peptide of one substrate could serve as the allosteric effector for the core peptide of another substrate, or the leader peptide could bind to the lid. We also note that the biosynthetic enzymes other than the proteases do not necessarily strictly distinguish between the parts of the precursor peptide we call the leader peptide and the core peptide, because this distinction is only introduced after proteolysis. Hence, for some enzymes it is possible that the binding site on the precursor peptide that activates a particular PTM enzyme by the model in Figure 10 could span a stretch of amino acids that is on both leader peptide and core peptide.

In addition to the two models presented thus far, in which the leader peptide binds in one specific binding pocket, alternative mechanisms in which the leader peptide moves over or through the enzyme have been proposed.[56] At present, no firm experimental support exists for or against such a mechanism, except that several PTM enzymes do not require any energy in the form of ATP,[59, 65] which appears to argue against any motor-like pulling on the leader peptide. Further studies will be required to determine whether the leader peptide

binds to one or multiple sites on each biosynthetic enzyme, and whether the same stretch of amino acids on the leader peptide is involved (see also next section).

How do the biosynthetic enzymes recognize the leader peptide?

Whereas the model described in the previous section can account for many experimental observations of how leader peptides might activate PTM enzymes of RiPP biosynthesis, very little is understood regarding the exact molecular interactions that are involved in leader peptide binding. At present, no crystal or NMR structures are available of any enzyme in complex with the leader peptide. Perhaps the best understood systems are the transporters that secrete substrates containing leader peptides that end in the double glycine motif, and that also contain a protease domain that cleaves the leader peptide from the modified core peptide. ComA, a transporter involved in cleavage and export of the quorum-sensing signal precursor ComC in *Streptococcus pneumonia* has homology with the transporters involved in RiPP biosynthesis. The leader peptide of ComC appears to attain a helical conformation upon binding to ComA, but not in solution.[66] A crystal structure of ComC showed that its protease active site is located at the end of a narrow cleft explaining the selectivity for the double-glycine cleavage site.[67] The authors also identified a shallow hydrophobic surface that was proposed to interact with the α-helical ComC leader peptide. The residues making up this proposed binding site are conserved in many of the transporters involved in RiPP biosynthesis.[67]

In the absence of any direct structural information, site-directed mutagenesis has been used extensively to investigate the importance of specific residues in RiPP leader peptides. Collectively, these studies have demonstrated that the biosynthetic enzymes generally demonstrate very relaxed substrate specificity with respect to the leader peptides and that the great majority of single mutations are well tolerated.[15, 68] Exceptions are the proteases that remove the leader peptide, which as expected have a much lower tolerance for substitution near the cleavage site.[14b, 68f, 69] It is possible that the generally observed plasticity with respect to leader peptide mutations ensures high evolvability of the core peptide sequence without loss of production as a consequence of leader peptide mutations. These studies have also shown a propensity for helical structures of the leader peptides,[61-62, 68d, 68e, 70] which is also supported by structure prediction programs. Most of these proposed helices are amphiphilic. In many cases, however, a helical conformation is not observed in aqueous buffer and secondary structure-inducing solvents such as trifluoroethanol are required to obtain a helical structure.[61-62, 70a-c] The idea of an unstructured precursor peptide that folds into a helical structure upon binding to their cognate biosynthetic enzymes is not without precedent. Several recent studies have demonstrated that structurally disordered proteins fold into defined structures upon binding their targets in the cell.[71] A general observation is that leader peptides within each RiPP class are quite diverse but have some highly conserved features and that most leader peptides have a preponderance of negatively charged residues.[6]

In the case of lanthipeptides, *in vivo* and *in vitro* binding assays have shown a direct interaction between the nisin leader sequence and the NisB/NisC enzymes.[72] The binding interaction likely involves a conserved FNLD motif.[68f, 72] Leader peptide truncation for the class II lanthipeptide lacticin 481 showed that the C-terminal portion of the leader sequence of LctA is critical for dehydration activity[11b] and that disruption of a proposed helical structure[73] within this C-terminal stretch by proline mutations strongly hampered the dehydration efficiency. The leader peptides of class III lanthipeptides contain an I-L-D/E-L-Q putative helix-forming motif near the N-terminus that is highly conserved. This recognition motif has been shown to be essential for enzymatic production of labyrinthopeptin A2 (Figure 11), while the C-terminal part of the leader has been proposed to act as a spacer between the binding motif and the core sequence.[61] A length requirement

between the proposed enzyme binding site on the leader peptide and residues in the core peptide that can be modified by the PTM enzymes has also been observed for lacticin 481 and nisin.[68f, 74]

In other RiPP systems, truncation of the leader peptide of the lasso peptide MccJ25 showed that only the eight C-terminal residues are important for producing the mature MccJ25 in *E. coli*, whereas the 25 residues on the N-terminal portion are dispensable.[60] These observations resemble the findings with the lanthipeptide lacticin 481 for which the N-terminal 9 amino acids were not required for *in vitro* dehydration and cyclization.[11b] The leader peptide of microcin B17 (Figure 4) is also believed to form a helical structure upon binding to its synthetase.[62] The presence of the helix and several key hydrophobic residues in the leader region are important for the enzyme to carry out the heterocyclization reaction on the native core peptide[22b, 62] and artificial core peptides.[31b, 70e] The first modified amino acid in the core peptide is separated from the end of the leader peptide by an essential[70f] Gly-rich spacer peptide (Figure 4). Like the lanthipeptides discussed above, it has been proposed that this linker is required to allow the core region to reach the active site(s) of the PTM enzymes for modification. The importance of residues in the leader peptide has also been investigated by alanine scanning mutagenesis for streptolysin S, another linear azol(in)e containing peptide.[70e] These studies revealed TQV and FXXXB (B = hydrophobic amino acid) motifs that are important for recognition similar to the findings for microcin B17. A short helix in the leader peptide has also been proposed to be important for recognition by the synthetase for heterocycle formation in cyanobactins,[70a, 70d] and the leader peptides of microviridins contain a strictly conserved PFFARFL recognition sequence that forms a putative helical structure that is important for lactone and lactam formation.[37d]

Biosynthetic enzymes in one pathway can recognize different parts of the substrate peptides

An important point with respect to leader peptide recognition is that not all the PTM enzymes need to recognize the same sequence of the leader peptide. Indeed, in cyanobactin biosynthesis, the enzymes involved in cyclodehydration likely require the leader peptide based on LAP biosynthesis,[62] whereas the enzymes that are responsible for head-to-tail cyclization require the recognition sequences that are immediately flanking the core peptides.[30, 32, 75] Similarly, mutagenesis studies have shown that the double Gly motif in the leader peptide of the class II lanthipeptide lacticin 481 is essential for proteolytic processing,[68c, 69b] but not for installation of lanthionine rings.[68d] Conversely, pull down experiments have shown that both the dehydratase NisB and the cyclase NisC bind the precursor peptide NisA, consistent with *in vitro* activity and binding data,[72a, 76] and mutagenesis experiments suggest both enzymes recognize the FNLD box on their NisA substrates.[68f, 72] Although these experiments show that the leader peptide is important for binding to NisB,[72a] as discussed in the next section, it is likely that NisB also has affinity for the core peptide.

Although both NisB and NisC recognize and bind the leader peptide, not all the enzymes that require the presence of the leader peptide for in vivo activity need to actually bind to the leader peptide themselves if they are part of a multi-enzyme complex. Indeed, many RiPP PTM enzymes are part of such complexes.[15, 22b, 26, 37d, 77] In principle, only one of the proteins in the complex needs to bind the leader peptide if the function of the leader peptide is just to bring the core peptide in proximity to the active sites of the proteins in the complex.

Biosynthetic enzymes in different RiPP pathways can recognize similar leader peptides

Not only do multiple enzymes that carry out very different reactions during the biosynthetic pathway to one particular RiPP recognize the same leader peptide as discussed in the previous section, PTM enzymes that produce very different classes of RiPPs can use leader peptides that are homologous. For instance, lanthipeptide synthetases, azoline cyclodehydratases, and radical SAM enzymes all act on members of the nitrile hydratase-like leader peptide (NHLP) family, as well as the nif11-like leader peptide family.[46, 78] The NHLP and nif11 leader peptides are uncharacteristically long and display clear sequence homology to the alpha subunit of the enzyme nitrile hydratase and the Nif11 nitrogen-fixing proteins from cyanobacteria, respectively.[46b] As such, these leader peptides appear to have been repurposed from enzymatic functions to allosteric regulation, the first clues of the evolutionary origins of leader peptides.[46b] Along similar lines, leader peptides ending in the double glycine motif with sequence homology to RiPP leader peptides are also found in bacteriocins that do not undergo posttranslational modifications. Hence, the leader peptides of these RiPP precursor peptides may have evolved from a role in secretion to include additional roles in activating the enzymes involved in posttranslational modification.

Is the Core Peptide Recognized by the Biosynthetic Enzymes?

As discussed in previous sections, the RiPP biosynthetic enzymes display remarkable plasticity with respect to processing variants of the core peptides.[15-16, 31a, 35a, 52, 55, 70e, 74, 79] Particularly striking examples are the attachment of entirely foreign sequences to lanthipeptide leader peptides that are dehydrated and cyclised by their biosynthetic enzymes,[52, 69a] and natural combinatorial biosynthesis in which the core peptides are hypervariable but the leader peptides or recognition sequences are highly conserved.[31b, 78] These observations raised the question whether the PTM enzymes recognize the core peptides at all. Recent studies in the lanthipeptide area suggest they typically do. First, the observation that the biosynthetic enzymes can dehydrate the core peptides in the absence of the leader peptide[56-58] can only be explained if the enzymes have some affinity for the core peptide. Second, the observed activity when the leader peptide is provided *in trans*, whether *in vitro*,[57] *in vivo*,[72b] or attached to the lanthionine synthetase,[58] requires that the enzyme must have affinity for the core peptide.

It is informative to consider the challenges faced by the various posttranslational modification enzymes with respect to core peptide recognition. Some enzymes such as the Ser/Thr dehydratases in lanthipeptide biosynthesis or the Ser/Thr/Cys cyclodehydratases involved in azol(in)e biosynthesis act on all or nearly all cognate amino acids in the core peptides. Hence, the main challenge is which cognate amino acids are “skipped”.[74-75]

One possible explanation for modification of most but not all Ser/Thr and/or Cys residues is that the enzymes display a strict order in which they catalyze their iterative chemical reactions, such that certain residues can be “protected” from posttranslational modification (e.g. because a ring is formed around or next to it it or because a D-stereocenter is introduced on a flanking residue). Indeed, several leader peptide dependent reactions in RiPP biosynthesis have been shown to display directionality. This was first shown for the heterocylization for microcin B17, which takes place from the N-terminus of the core peptide to the C-terminus.[80] Similarly, the dehydration and cyclization reactions during the biosynthesis of the class I lantibiotic nisin and the class II lantibiotics lacticin 481 and haloduracin move from the N-terminus of the core peptide towards the C-terminus.[81] These studies also suggest that dehydration and cyclizations may be tightly coupled alternating activities. Conversely, for the class III lanthipeptides labyrinthopeptin A2 and catenulipeptin, the synthetases display a C- to N-terminal processing mode.[82] A predominantly C-to-N terminal directionality has also been reported for the LAP

plantazolicin.[79o] For all investigated examples, the enzymatic reactions that are directional are not processive, but rather distributive.[79o, 80, 81b, 82] That is, the enzymes release the intermediates after each enzymatic reaction rather than keeping them bound, at least *in vitro*. These observations are in keeping with the moderate affinities that have been reported for most RiPP enzymes for their substrate peptides,[72a, 76b] which in turn is beneficial to avoid product inhibition. Another example of specific order of biosynthetic reactions is found for the microviridins where the order appears not to be determined as much by directionality as by chemoselectivity, with the ester bonds being formed before the amide bonds. Although many of these enzymatic processes display overall directionality *in vitro*, most are not strictly directional as two residues that are localized close in sequence often are not strongly differentiated. It is possible that *in vivo* the directionality may be tighter. First, other proteins that make up multi-enzyme complexes but that are absent in *in vitro* studies could provide additional control. Secondly, *in vitro* reactions have typically been conducted with the precursor peptide in excess over the biosynthetic enzymes whereas *in vivo* it is likely that this is not the case. Excess substrate can compete with substrates that are being modified, potentially leading to disruption of strict directionality.

How can one enzyme make rings of different sizes, sequences, and even stereochemistry?

Crosslinking enzymes have considerably bigger challenges to overcome, as they typically generate a single ring topology when many different topologies are possible. For instance, consider the cyclization process for nisin (Figure 2B). The dehydrated NisA peptide contains five Cys nucleophiles and eight dehydro amino acid electrophiles (three Dha, five Dhb). Statistically, the number of constitutional isomers that can be generated by a Michael type cyclization process that lacks regio- and chemoselectivity would be 6,720. Moreover, if taking into account all of the possible stereoisomers that can be formed during the Michael addition, this number would be even greater (the exact number depends on whether Lan or MeLan products are formed, but would be at least 8.6×10^5). However, NisC generates a single product out of all these potential structures. Although the control over ring topology is most relevant to the lanthipeptides, similar considerations are also in play for other crosslinking reactions such as thioether formation in sactipeptides or lactone/lactam formation in microviridins.

The control over site selectivity is an astounding feat considering that the crosslinks formed by one enzyme often involve very different amino acid sequences and ring sizes (e.g. NisC makes rings that range from four amino acids to seven amino acids with very different sequences, Figure 2B). How one active site of defined size could actively “enforce” formation of these very different rings is difficult to comprehend. Moreover, the shape of the substrate peptide is changing dramatically with each cyclization, especially for RiPPs with overlapping rings (eg lacticin 481, Figure 11), which again is difficult to reconcile with a single, well-defined active site. Another puzzling system is the class II lanthipeptide synthetase ProcM. This enzyme is produced by a *Prochlorococcus* strain that encodes only one lanthipeptide synthetase but no less than 29 different substrate peptides on its genome.[78] The leader peptides of these substrates are highly homologous, but the core peptides are highly diverse. ProcM converts these core peptides into a library of polycyclic peptides with high structural diversity (e.g. Figure 12).[83]

These observations prompt us to speculate that perhaps the core peptides themselves determine the outcome of the cyclization process. Additional observations that point in this direction are the observed phylogenomic divergence of cyclase enzymes that generate very similar ring topologies,[84] and the recent remarkable observation that rings formed from a DhxDhxXxxXxxCys motif (Dhx = Dha or Dhb; Xxx is any amino acid) in the substrate peptides for three different lanthipeptides have different stereochemistry from the canonical (2*S*, 6*R*) and (2*S*, 3*S*, 6*R*) configurations. During the cyclization process of these

lanthipeptides (haloduracin β and both peptides of cytolysin, Figures 11 and 12), a single synthetase catalyzes ring formation with both the canonical stereochemistry, resulting from attack of the Cys onto the *Si* face of the Dhx residue in an overall *anti* addition, as well as alternative stereochemistry, resulting from overall *anti* addition in which the Cys attacks the *Re* face of the Dhx residue. Again, it is difficult to envision how an enzyme active site could actively enforce both types of geometries in a well defined pocket.

Instead, we propose that the enzyme surface may have different patches of hydrophobic and/or charged residues that are complementary to the different rings that the enzyme makes. In this model, a highly promiscuous enzyme like ProcM has not evolved to actively favor the formation of specific rings but instead makes rings that the substrate inherently favors. Unless certain ring topologies are energetically very highly favored, this model suggests ProcM may generate multiple products from one substrate peptide, something that needs to be investigated in more detail. The predecessors of the modern enzymes such as NisC, which are now very specific in the rings they make, initially may also have made rings determined by the sequence of the precursor peptides. Once one of the products conferred an evolutionary advantage upon the producing strain, any enzyme that would favor the formation of this product would be selected for. In our model, this could mean that the surface of the enzyme near the zinc ion that activates the substrate[76a, 85] would start to evolve such that it is most complementary to the residues in the rings found in the product that confers the advantage. The observation that one enzyme can make rings of different sizes and different stereochemistry can be explained by this model by having different complementary patches on the enzyme for the different rings of the bioactive product. The alternative stereochemistry observed for cytolysin and haloduracin would then be the result of the original propensity of a peptide containing two consecutive Dhx residues to attain a different conformation that presents the opposite face of the alkene to the zinc-bound Cys nucleophile. An alternative, or complementary, model to select for the bioactive topology during evolution could involve the development of alternating dehydratase and cyclase activities that could favor the bioactive ring topology by preventing alternative ring patterns.

Much research will be needed to test this mechanistic hypothesis, but some data already is available. Early model studies of the Michael-type additions that generate the thioether crosslinks in lanthipeptides concluded that indeed, the peptides have a propensity to form the same stereochemical products in the absence of any enzymes as is observed for authentic lanthipeptides when individual rings are studied.[86] However, when multiple dehydro amino acids and cysteines are present, a biomimetic approach did not provide the desired ring topology for the one case that has been investigated (the nisin A and B rings).[87] Hence, the NisC cyclase must in some way guide ring formation. It is worth noting that these model studies were carried out before it was known that a Zn^{2+} ion is used in lanthionine cyclases to promote the Michael-type additions,[76a, 85] and that it may be that base-catalysed conjugate addition is not a good mimic for metal-promoted cyclization.

Support for the notion that the precursor peptide may have an inherent propensity to form specific ring structures is provided by various studies in which chimeric peptides were constructed made up of a leader peptide of one lanthipeptide and the core peptide of a different lanthipeptide. When a chimera consisting of the subtilin leader and nisin core peptides was expressed in a subtilin producing *Bacillus* strain, a fully processed product was observed.[88] Similarly, when a chimera containing a subtilin leader and nisin core peptide was expressed in a nisin producing *Lactococcus* strain the core peptide was processed.[89] These early studies support the idea that the core peptides contain the information for correct ring formation, but the subtilin and nisin ring topologies are essentially identical (Figures 2 and 12), and hence it is still possible that the NisC cyclase recognized the subtilin core peptide and vice versa. The same argument can be made for the observed *in vitro* production

of nukacin ISK-1 by the lacticin 481 synthetase LctM from a chimeric substrate consisting of the lacticin 481 leader peptide and the nukacin core peptide,[68d] because lacticin 481 and nukacin ISK-1 have the same ring topology. However, more recent studies have expanded the use of chimeric peptides. For instance, a chimera of the nisin leader peptide and the core peptide of a class II lantibiotic pneumococcin that has no sequence homology to nisin could be modified and secreted as a bioactive product by the class I nisin machinery.[90] Similarly, a NisA-ElxA chimeric peptide containing the core peptide of epilancin 15X fused to the C-terminus of the nisin leader was processed by the nisin dehydratase NisB and cyclase NisC in *E. coli*.[84]

As mentioned above, ProcM has very relaxed substrate specificity and processes 29 very diverse core peptides, suggesting it does not enforce any specific topology. We reasoned that ProcM may therefore be the best available biosynthetic enzyme to test the idea that the core peptide determines the site selectivity of cyclization. The lacticin 481 core sequence is not found in the 29 core peptides that ProcM acts on. Therefore, we probed whether ProcM, an enzyme from a marine cyanobacterium, could make lacticin 481, a compound produced by a lactic acid bacterium; LctM and ProcM have just 22% sequence identity. When a chimeric peptide consisting of the ProcA leader peptide and lacticin 481 core peptide was co-expressed with ProcM, the resulting product after removal of the leader peptide indeed contained a significant amount of lacticin 481.[84] Therefore, for certain scaffolds, the ring topology of the final lanthipeptide is not just determined by the synthetase, but also by the sequence of the substrate. Although any tests of this proposal have been limited thus far to lanthipeptides, it is worth noting that sactipeptides have a very distinct hairpin structure and it is possible that this secondary structure is already present before the thioether crosslinks are formed. Indeed, it is has been proposed that the thioether bridges are installed via a zipper mechanism, in which the first thioether bond formation aids the formation of the second thioether bridge.[43, 91] This mechanism would be facilitated by an inherent propensity of the substrate to already attain a conformation that brings the cysteine residues close to their partners to which they become crosslinked. However, the precursor peptide to the lasso peptide MccJ25 does not appear to attain a conformation that resembles the lasso fold in the absence of the biosynthetic enzymes.[92] Hence, not all RiPP precursor peptides need to have an intrinsic propensity to exist in conformations that are close that those found in the final crosslinked product.

Summary

The leader peptide is important for most but not all biosynthetic enzymes. For those enzymes that are activated by binding of the leader peptide, we suggest based on the available information that leader peptide binding functions as an allosteric effector that changes the distribution of enzyme populations towards the active conformation. At present, the details of the molecular recognition of the leader peptide by the biosynthetic enzymes are still largely unknown. Indirect evidence suggests that the enzymes recognize secondary structure rather than specific residues given the high tolerance that has been demonstrated with respect to leader peptide mutations. It is possible that this is a built-in safety feature to allow high evolvability of the core peptide sequence without placing too much of a restriction on maintaining the sequence of the leader peptide. Recent studies have also demonstrated that the enzymes clearly must recognize the core peptides of their substrates, and possibly that the core peptides may in part determine the ring topology and stereochemistry of the final products.

Acknowledgments

Our work on RiPPs is supported by the National Institutes of Health (GM 058822).

References

[1]. a) Wender PA, Miller BL. Nature. 2009; 460:197–201. [PubMed: 19587760] b) Wilson RM, Danishefsky SJ. J. Org. Chem. 2006; 71:8329–8351. [PubMed: 17064003]

[2]. a) Hung DT, Jamison TF, Schreiber SL. Chem. Biol. 1996; 3:623–639. [PubMed: 8807895] b) Cardenas ME, Sanfridson A, Cutler NS, Heitman J. Trends in biotechnology. 1998; 16:427–433. [PubMed: 9807840]

[3]. Newman DJ, Cragg GM. J. Nat. Prod. 2012; 75:311–335. [PubMed: 22316239]

[4]. Arnison PGB, M. J. Bierbaum G, Bowers AA, Bulaj G, Camarero JA, Campopiano DJ, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian K-D, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Reaney MJT, Rebuffat S, Ross RP, Sahl H-G, Saris P, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Süssmuth RE, Tagg JR, Tang G-L, Vederas JC, Walsh CT, Walton JD, Willey JM, van der Donk WA. Nat. Prod. Rep. 2013; 30:108–160. [PubMed: 23165928]

[5]. a) McIntosh JA, Donia MS, Schmidt EW. Nat. Prod. Rep. 2009; 26:537–559. [PubMed: 19642421] b) Dunbar KL, Mitchell DA. ACS Chem. Biol. 2013

[6]. a) Oman TJ, van der Donk WA. Nat. Chem. Biol. 2010; 6:9–18. [PubMed: 20016494] b) Plat A, Kuipers A, Rink R, Moll GN. Curr. Protein Pept. Sci. 2013 in press.

[7]. a) Chatterjee C, Paul M, Xie L, van der Donk WA. Chem. Rev. 2005; 105:633–684. [PubMed: 15700960] b) Cotter PD, Hill C, Ross RP. Curr. Protein Pept. Sci. 2005; 6:61–75. [PubMed: 15638769] c) Nagao J, Asaduzzaman SM, Aso Y, Okuda K, Nakayama J, Sonomoto K. J. Biosci. Bioeng. 2006; 102:139–149. [PubMed: 17046525] d) Willey JM, van der Donk WA. Annu. Rev. Microbiol. 2007; 61:477–501. [PubMed: 17506681] e) Bierbaum G, Sahl HG. Curr. Pharm. Biotechnol. 2009; 10:2–18. [PubMed: 19149587] f) Knerr PJ, van der Donk WA. Annu. Rev. Biochem. 2012; 81:479–505. [PubMed: 22404629]

[8]. Schnell N, Entian K-D, Schneider U, Götz F, Zahner H, Kellner R, Jung G. Nature. 1988; 333:276–278. [PubMed: 2835685]

[9]. Gross E, Morell JL. J. Am. Chem. Soc. 1971; 93:4634–4635. [PubMed: 5131162]

[10]. a) Chan WC, Leyland M, Clark J, Dodd HM, Lian LY, Gasson MJ, Bycroft BW, Roberts GC. FEBS Lett. 1996; 390:129–132. [PubMed: 8706842] b) Hsu ST, Breukink E, Tischenko E, Lutters MA, De Kruijff B, Kaptein R, Bonvin AM, Van Nuland NA. Nat. Struct. Mol. Biol. 2004; 11:963–967. [PubMed: 15361862] c) Lubelski J, Rink R, Khusainov R, Moll GN, Kuipers OP. Cell. Mol. Life Sci. 2008; 65:455–476. [PubMed: 17965835]

[11]. a) Siezen RJ, Kuipers OP, de Vos WM. Antonie van Leeuwenhoek. 1996; 69:171–184. [PubMed: 8775977] b) Xie L, Miller LM, Chatterjee C, Averin O, Kelleher NL, van der Donk WA. Science. 2004; 303:679–681. [PubMed: 14752162]

[12]. Goto Y, Li B, Claesen J, Shi Y, Bibb MJ, van der Donk WA. PLoS Biol. 2010; 8:e1000339. [PubMed: 20351769]

[13]. Meindl K, Schmiederer T, Schneider K, Reicke A, Butz D, Keller S, Guhring H, Vertesy L, Wink J, Hoffmann H, Bronstrup M, Sheldrick GM, Süssmuth RD. Angew. Chem. Int. Ed. 2010; 49:1151–1154.

[14]. a) van den Berg van Saparoea HB, Bakkes PJ, Moll GN, Driessen AJ. Appl. Environ. Microbiol. 2008; 74:5541–5548. [PubMed: 18621866] b) Nishie M, Sasaki M, Nagao J, Zendo T, Nakayama J, Sonomoto K. J. Biol. Chem. 2011; 286:11163–11169. [PubMed: 21303905]

[15]. Maksimov MO, Pan SJ, James Link A. Nat. Prod. Rep. 2012; 29:996–1006. [PubMed: 22833149]

[16]. Knappe TA, Linne U, Robbel L, Marahiel MA. Chem. Biol. 2009; 16:1290–1298. [PubMed: 20064439]

[17]. Maksimov MO, Pelczer I, Link AJ. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:15223–15228. [PubMed: 22949633]

[18]. Ducasse R, Yan KP, Goulard C, Blond A, Li Y, Lescop E, Guittet E, Rebuffat S, Zirah S. ChemBioChem. 2012; 13:371–380. [PubMed: 22287061]

[19]. a) Duquesne S, Destoumieux-Garzon D, Zirah S, Goulard C, Peduzzi J, Rebuffat S. Chem. Biol. 2007; 14:793–803. [PubMed: 17656316] b) Clarke DJ, Campopiano DJ. Org. Biomol. Chem. 2007; 5:2564–2566. [PubMed: 18019529] c) Pan SJ, Rajniak J, Cheung WL, Link AJ. ChemBioChem. 2012; 13:367–370. [PubMed: 22213148] d) Yan KP, Li Y, Zirah S, Goulard C, Knappe TA, Marahiel MA, Rebuffat S. ChemBioChem. 2012; 13:1046–1052. [PubMed: 22488892]

[20]. Knappe TA, Linne U, Zirah S, Rebuffat S, Xie X, Marahiel MA. J. Am. Chem. Soc. 2008; 130:11446–11454. [PubMed: 18671394]

[21]. a) Melby JO, Nard NJ, Mitchell DA. Curr. Opin. Chem. Biol. 2011; 15:369–378. [PubMed: 21429787] b) Walsh CT, Malcolmson SJ, Young TS. ACS Chem. Biol. 2012; 7:429–442. [PubMed: 22206579]

[22]. a) Vizan JL, Hernandez-Chico C, del Castillo I, Moreno F. Embo J. 1991; 10:467–476. [PubMed: 1846808] b) Li YM, Milne JC, Madison LL, Kolter R, Walsh CT. Science. 1996; 274:1188–1193. [PubMed: 8895467]

[23]. Lee SW, Mitchell DA, Markley AL, Hensler ME, Gonzalez D, Wohlrab A, Dorrestein PC, Nizet V, Dixon JE. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:5879–5884. [PubMed: 18375757]

[24]. Scholz R, Molohon KJ, Nachtigall J, Vater J, Markley AL, Süssmuth RD, Mitchell DA, Borriss R. J. Bacteriol. 2011; 193:215–224. [PubMed: 20971906]

[25]. Fischbach MA, Walsh CT. Chem. Rev. 2006; 106:3468–3496. [PubMed: 16895337]

[26]. Dunbar KL, Melby JO, Mitchell DA. Nat. Chem. Biol. 2012; 8:569–575. [PubMed: 22522320]

[27]. Milne JC, Roy RS, Eliot AC, Kelleher NL, Wokhlu A, Nickels B, Walsh CT. Biochemistry. 1999; 38:4768–4781. [PubMed: 10200165]

[28]. a) Houssen WE, Jaspars M. ChembioChem. 2010; 11:1803–1815. [PubMed: 20632432] b) Donia, MS.; Schmidt, EW. Comprehensive Natural Products II Chemistry and Biology. Moore, BS.; Crews, P., editors. Vol. 2. Elsevier; Oxford: 2010. p. 539-558.

[29]. a) Long PF, Dunlap WC, Battershill CN, Jaspars M. ChemBioChem. 2005; 6:1760–1765. [PubMed: 15988766] b) Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J. Proc. Natl. Acad. Sci. U. S. A. 2005; 102:7315–7320. [PubMed: 15883371] c) Sivonen K, Leikoski N, Fewer DP, Jokela J. Appl. Microbiol. Biotechnol. 2010; 86:1213–1225. [PubMed: 20195859]

[30]. Lee J, McIntosh J, Hathaway BJ, Schmidt EW. J. Am. Chem. Soc. 2009; 131:2122–2124. [PubMed: 19166292]

[31]. a) Donia MS, Ravel J, Schmidt EW. Nat. Chem. Biol. 2008; 4:341–343. [PubMed: 18425112] b) Donia MS, Hathaway BJ, Sudek S, Haygood MG, Rosovitz MJ, Ravel J, Schmidt EW. Nat. Chem. Biol. 2006; 2:729–735. [PubMed: 17086177]

[32]. a) McIntosh JA, Robertson CR, Agarwal V, Nair SK, Bulaj GW, Schmidt EW. J. Am. Chem. Soc. 2010; 132:15499–15501. [PubMed: 20961047] b) Koehnke J, Bent A, Houssen WE, Zollman D, Morawitz F, Shirran S, Vendome J, Nneoyiegbe AF, Trembleau L, Botting CH, Smith MC, Jaspars M, Naismith JH. Nat. Struct. Mol. Biol. 2012; 19:767–772. [PubMed: 22796963] c) Agarwal V, Pierce E, McIntosh J, Schmidt EW, Nair SK. Chem. Biol. 2012; 19:1411–1422. [PubMed: 23177196]

[33]. McIntosh JA, Donia MS, Nair SK, Schmidt EW. J. Am. Chem. Soc. 2011; 133:13698–13705. [PubMed: 21766822]

[34]. a) Shimamura H, Gouda H, Nagai K, Hirose T, Ichioka M, Furuya Y, Kobayashi Y, Hirono S, Sunazuka T, Omura S. Angew. Chem. Int. Ed. Engl. 2009; 48:914–917. [PubMed: 19115340] b) Kobayashi Y, Ichioka M, Hirose T, Nagai K, Matsumoto A, Matsui H, Hanaki H, Masuma R, Takahashi Y, Omura S, Sunazuka T. Bioorg. Med. Chem. Lett. 2010; 20:6116–6120. [PubMed: 20833545]

[35]. a) Huo L, Rachid S, Stadler M, Wenzel SC, Müller R. Chem. Biol. 2012; 19:1278–1287. [PubMed: 23021914] b) Crone WJK, Leeper FJ, Truman AW. Chem. Sci. 2012; 3:3516–3521.c) Gomez-Escribano JP, Song L, Bibb MJ, Challis GL. Chem. Sci. 2012; 3:3522–3525.d) Hou Y,

Tianero MD, Kwan JC, Wyche TP, Michel CR, Ellis GA, Vazquez-Rivera E, Braun DR, Rose WE, Schmidt EW, Bugni TS. Org. Lett. 2012; 14:5050–5053. [PubMed: 22984777]

[36]. Ishitsuka MO, Kusumi T, Kakisawa H, Kaya K, Watanabe MM. J. Am. Chem. Soc. 1990; 112:8180–8182.

[37]. a) Philmus B, Christiansen G, Yoshida WY, Hemscheidt TK. ChemBioChem. 2008; 9:3066–3073. [PubMed: 19035375] b) Ziemert N, Ishida K, Liaimer A, Hertweck C, Dittmann E. Angew. Chem., Int. Ed. Engl. 2008; 47:7756–7759. [PubMed: 18683268] c) Philmus B, Guerrette JP, Hemscheidt TK. ACS Chem. Biol. 2009; 4:429–434. [PubMed: 19445532] d) Weiz AR, Ishida K, Makower K, Ziemert N, Hertweck C, Dittmann E. Chem. Biol. 2011; 18:1413–1421. [PubMed: 22118675]

[38]. Kawulka K, Sprules T, McKay RT, Mercier P, Diaper CM, Zuber P, Vederas JC. J. Am. Chem. Soc. 2003; 125:4726–4727. [PubMed: 12696888]

[39]. Kawulka KE, Sprules T, Diaper CM, Whittal RM, McKay RT, Mercier P, Zuber P, Vederas JC. Biochemistry. 2004; 43:3385–3395. [PubMed: 15035610]

[40]. Sit CS, McKay RT, Hill C, Ross RP, Vederas JC. J. Am. Chem. Soc. 2011; 133:7680–7683. [PubMed: 21526839]

[41]. Sit CS, van Belkum MJ, McKay RT, Worobo RW, Vederas JC. Angew. Chem. Int. Ed. 2011; 50:8718–8721.

[42]. Thennarasu S, Lee DK, Poon A, Kawulka KE, Vederas JC, Ramamoorthy A. Chem. Phys. Lipids. 2005; 137:38–51. [PubMed: 16095584]

[43]. Flühe L, Knappe TA, Gattner MJ, Schäfer A, Burghaus O, Linne U, Marahiel MA. Nat. Chem. Biol. 2012; 8:350–357. [PubMed: 22366720]

[44]. Hamada T, Matsunaga S, Yano G, Fusetani N. J .Am. Chem. Soc. 2005; 127:110–118. [PubMed: 15631460]

[45]. Hamada T, Matsunaga S, Fujiwara M, Fujita K, Hirota H, Schmucki R, Guntert P, Fusetani N. J. Am. Chem. Soc. 2010; 132:12941–12945. [PubMed: 20795624]

[46]. a) Freeman MF, Gurgui C, Helf MJ, Morinaka BI, Uria AR, Oldham NJ, Sahl HG, Matsunaga S, Piel J. Science. 2012; 338:387–390. [PubMed: 22983711] b) Haft DH, Basu MK, Mitchell DA. BMC Biol. 2010; 8:70. [PubMed: 20500830]

[47]. Kupke T, Kempter C, Jung G, Götz F. J. Biol. Chem. 1995; 270:11282–11289. [PubMed: 7744764]

[48]. Ökesli A, Cooper LE, Fogle EJ, van der Donk WA. J. Am. Chem. Soc. 2011; 133:13753–13760. [PubMed: 21770392]

[49]. Shi Y, Bueno A, van der Donk WA. Chem. Commun. 2012; 48:10966–10968.

[50]. Velásquez JE, Zhang X, van der Donk WA. Chem. Biol. 2011; 18:857–867. [PubMed: 21802007]

[51]. Moll GN, Kuipers A, Rink R. Antonie van Leeuwenhoek. 2010; 97:319–333. [PubMed: 20140513]

[52]. Kluskens LD, Kuipers A, Rink R, de Boef E, Fekken S, Driessen AJ, Kuipers OP, Moll GN. Biochemistry. 2005; 44:12827–12834. [PubMed: 16171398]

[53]. Rink R, Wierenga J, Kuipers A, Kluskens LD, Driessen AJM, Kuipers OP, Moll GN. Appl. Environ. Microbiol. 2007; 73:1792–1796. [PubMed: 17261515]

[54]. Levengood MR, van der Donk WA. Bioorg. Med. Chem. Lett. 2008; 18:3025–3028. [PubMed: 18294843]

[55]. a) Tianero MD, Donia MS, Young TS, Schultz PG, Schmidt EW. J. Am. Chem. Soc. 2012; 134:418–425. [PubMed: 22107593] b) Knappe TA, Manzenrieder F, Mas-Moruno C, Linne U, Sasse F, Kessler H, Xie X, Marahiel MA. Angew. Chem. Int. Ed. 2011; 50:8714–8717.

[56]. Khusainov R, Kuipers OP. ChembioChem. 2012; 13:2433–2438. [PubMed: 23070977]

[57]. Levengood MR, Patton GC, van der Donk WA. J. Am. Chem. Soc. 2007; 129:10314–10315. [PubMed: 17676843]

[58]. Oman TJ, Knerr PJ, Bindman NA, Velasquez JE, van der Donk WA. J. Am. Chem. Soc. 2012; 134:6952–6955. [PubMed: 22480178]

[59]. Wang H, van der Donk WA. ACS Chem. Biol. 2012; 7:1529–1535. [PubMed: 22725258]

[60]. Cheung WL, Pan SJ, Link AJ. J. Am. Chem. Soc. 2010; 132:2514–2515. [PubMed: 20143810]

[61]. Müller WM, Ensle P, Krawczyk B, Süssmuth RD. Biochemistry. 2011; 50:8362–8373. [PubMed: 21905643]

[62]. Roy RS, Kim S, Baleja JD, Walsh CT. Chem. Biol. 1998; 5:217–228. [PubMed: 9545435]

[63]. a) Boehr DD, Nussinov R, Wright PE. Nat. Chem. Biol. 2009; 5:789–796. [PubMed: 19841628] b) Boehr DD, McElheny D, Dyson HJ, Wright PE. Science. 2006; 313:1638–1642. [PubMed: 16973882] c) Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, Pozharski E, Wilson MA, Petsko GA, Karplus M, Hubner CG, Kern D. Nature. 2007; 450:838–844. [PubMed: 18026086]

[64]. a) Newton AC. Chem. Rev. 2001; 101:2353–2364. [PubMed: 11749377] b) Kobe B, Kemp BE. Nature. 1999; 402:373–376. [PubMed: 10586874]

[65]. Chatterjee C, Miller LM, Leung YL, Xie L, Yi M, Kelleher NL, van der Donk WA. J. Am. Chem. Soc. 2005; 127:15332–15333. [PubMed: 16262372]

[66]. Kotake Y, Ishii S, Yano T, Katsuoka Y, Hayashi H. Biochemistry. 2008; 47:2531–2538. [PubMed: 18232718]

[67]. Ishii S, Yano T, Ebihara A, Okamoto A, Manzoku M, Hayashi H. J. Biol. Chem. 2010; 285:10777–10785. [PubMed: 20100826]

[68]. a) van der Meer JR, Rollema HS, Siezen RJ, Beerthuyzen MM, Kuipers OP, de Vos WM. J. Biol. Chem. 1994; 269:3555–3562. [PubMed: 8106398] b) Neis S, Bierbaum G, Josten M, Pag U, Kempter C, Jung G, Sahl HG. FEMS Microbiol. Lett. 1997; 149:249–255. [PubMed: 9141666] c) Chen P, Qi FX, Novak J, Krull RE, Caufield PW. FEMS Microbiol. Lett. 2001; 195:139–144. [PubMed: 11179642] d) Patton GC, Paul M, Cooper LE, Chatterjee C, van der Donk WA. Biochemistry. 2008; 47:7342–7351. [PubMed: 18570437] e) Nagao J, Morinaga Y, Islam MR, Asaduzzaman SM, Aso Y, Nakayama J, Sonomoto K. Peptides. 2009; 30:1412–1420. [PubMed: 19481127] f) Plat A, Kluskens LD, Kuipers A, Rink R, Moll GN. Appl. Environ. Microbiol. 2011; 77:604–611. [PubMed: 21097596]

[69]. a) Kuipers A, De Boef E, Rink R, Fekken S, Kluskens LD, Driessen AJ, Leenhouts K, Kuipers OP, Moll GN. J. Biol. Chem. 2004; 279:22176–22182. [PubMed: 15044440] b) Furgerson Ihnken LA, Chatterjee C, van der Donk WA. Biochemistry. 2008; 47:7352–7363. [PubMed: 18570436]

[70]. a) Houssen WE, Wright SH, Kalverda AP, Thompson GS, Kelly SM, Jaspars M. ChembioChem. 2010; 11:1867–1873. [PubMed: 20715266] b) Sprules T, Kawulka KE, Gibbs AC, Wishart DS, Vederas JC. Eur. J. Biochem. 2004; 271:1748–1756. [PubMed: 15096213] c) Beck-Sickinger, AG.; Jung, G. Nisin and Novel Lantibiotics. Jung, G.; Sahl, H-G., editors. ESCOM; Leiden: 1991. p. 218-230.d) Donia MS, Schmidt EW. Chem. Biol. 2011; 18:508–519. [PubMed: 21513887] e) Mitchell DA, Lee SW, Pence MA, Markley AL, Limm JD, Nizet V, Dixon JE. J. Biol. Chem. 2009; 284:13004–13012. [PubMed: 19286651] f) Sinha Roy R, Belshaw PJ, Walsh CT. Biochemistry. 1998; 37:4125–4136. [PubMed: 9521734]

[71]. Sugase K, Dyson HJ, Wright PE. Nature. 2007; 447:1021–1025. [PubMed: 17522630]

[72]. a) Mavaro A, Abts A, Bakkes PJ, Moll GN, Driessen AJ, Smits SH, Schmitt L. J. Biol. Chem. 2011; 286:30552–30560. [PubMed: 21757717] b) Khusainov R, Heils R, Lubelski J, Moll GN, Kuipers OP. Mol. Microbiol. 2011; 82:706–718. [PubMed: 22011325]

[73]. Beck-Sickinger AG, Jung G. Liebigs Ann. Chem. 1993; 10:1125–1131.

[74]. Chatterjee C, Patton GC, Cooper L, Paul M, van der Donk WA. Chem. Biol. 2006; 13:1109–1117. [PubMed: 17052615]

[75]. McIntosh JA, Donia MS, Schmidt EW. J. Am. Chem. Soc. 2010; 132:4089–4091. [PubMed: 20210311]

[76]. a) Li B, Yu JP, Brunzelle JS, Moll GN, van der Donk WA, Nair SK. Science. 2006; 311:1464–1467. [PubMed: 16527981] b) Li B, van der Donk WA. J. Biol. Chem. 2007; 282:21169–21175. [PubMed: 17513866]

[77]. a) Kiesau P, Eikmanns U, Gutowski-Eckel Z, Weber S, Hammelmann M, Entian K-D. J. Bacteriol. 1997; 179:1475–1481. [PubMed: 9045802] b) Siegers K, Heinzmann S, Entian K-D. J. Biol. Chem. 1996; 271:12294–12301. [PubMed: 8647829] c) Nagao J, Aso Y, Sashihara T,

Shioya K, Adachi A, Nakayama J, Sonomoto K. Biosci. Biotechnol. Biochem. 2005; 69:1341–1347. [PubMed: 16041140]

[78]. Li B, Sher D, Kelly L, Shi Y, Huang K, Knerr PJ, Joewono I, Rusch D, Chisholm SW, van der Donk WA. Proc. Natl. Acad. Sci. U.S.A. 2010; 107:10430–10435. [PubMed: 20479271]

[79]. a) Kuipers OP, Bierbaum G, Ottenwälder B, Dodd HM, Horn N, Metzger J, Kupke T, Gnau V, Bongers R, van den Bogaard P, Kosters H, Rollema HS, de Vos WM, Siezen RJ, Jung G, Götz F, Sahl HG, Gasson MJ. Antonie van Leeuwenhoek. 1996; 69:161–169. [PubMed: 8775976] b) Cotter PD, Deegan LH, Lawton EM, Draper LA, O'Connor PM, Hill C, Ross RP. Mol. Microbiol. 2006; 62:735–747. [PubMed: 17076667] c) Rink R, Wierenga J, Kuipers A, Kluskens LD, Driessen AJM, Kuipers OP, Moll GN. Appl. Environ. Microbiol. 2007; 73:5809–5816. [PubMed: 17660303] d) Pavlova O, Mukhopadhyay J, Sineva E, Ebright RH, Severinov K. J. Biol. Chem. 2008; 283:25589–25595. [PubMed: 18632663] e) Appleyard AN, Choi S, Read DM, Lightfoot A, Boakes S, Hoffmann A, Chopra I, Bierbaum G, Rudd BA, Dawson MJ, Cortés J. Chem. Biol. 2009; 16:490–498. [PubMed: 19477413] f) Levengood MR, Knerr PJ, Oman TJ, van der Donk WA. J. Am. Chem. Soc. 2009; 131:12024–12025. [PubMed: 19655738] g) Islam MR, Shioya K, Nagao J, Nishie M, Jikuya H, Zendo T, Nakayama J, Sonomoto K. Mol. Microbiol. 2009; 72:1438–1447. [PubMed: 19432794] h) Acker MG, Bowers AA, Walsh CT. J. Am. Chem. Soc. 2009; 131:17563–17565. [PubMed: 19911780] i) Bowers AA, Acker MG, Koglin A, Walsh CT. J. Am. Chem. Soc. 2010; 132:7519–7527. [PubMed: 20455532] j) Pan SJ, Link AJ. J. Am. Chem. Soc. 2011; 133:5016–5023. [PubMed: 21391585] k) Li C, Zhang F, Kelly WL. Chem. Commun. 2012; 48:558–560.l) Li C, Zhang F, Kelly WL. Mol. Biosyst. 2011; 7:82–90. [PubMed: 21107477] m) Onaka H, Nakaho M, Hayashi K, Igarashi Y, Furumai T. Microbiology. 2005; 151:3923–3933. [PubMed: 16339937] n) Shioya K, Harada Y, Nagao J, Nakayama J, Sonomoto K. Appl. Microbiol. Biotechnol. 2010; 86:891–899. [PubMed: 19915831] o) Melby JO, Dunbar KL, Trinh NQ, Mitchell DA. J. Am. Chem. Soc. 2012; 134:5309–5316. [PubMed: 22401305] p) Boakes S, Ayala T, Herman M, Appleyard AN, Dawson MJ, Cortes J. Appl. Microbiol. Biotechnol. 2012; 95:1509–1517. [PubMed: 22526797] q) Field D, Connor PM, Cotter PD, Hill C, Ross RP. Mol. Microbiol. 2008; 69:218–230. [PubMed: 18485077] r) Young TS, Dorrestein PC, Walsh CT. Chem. Biol. 2012; 19:1600–1610. [PubMed: 23261603]

[80]. Kelleher NL, Hendrickson CL, Walsh CT. Biochemistry. 1999; 38:15623–15630. [PubMed: 10569947]

[81]. a) Lubelski J, Khusainov R, Kuipers OP. J. Biol. Chem. 2009; 284:25962–25972. [PubMed: 19620240] b) Lee MV, Ihnken LA, You YO, McClerren AL, van der Donk WA, Kelleher NL. J. Am. Chem. Soc. 2009; 131:12258–12264. [PubMed: 19663480]

[82]. Krawczyk B, Ensle P, Müller WM, Süssmuth RD. J. Am. Chem. Soc. 2012; 134:9922. [PubMed: 22687055]

[83]. Tang W, van der Donk WA. Biochemistry. 2012; 51:4271–4279. [PubMed: 22574919]

[84]. Zhang Q, Yu Y, Velasquez JE, van der Donk WA. Proc. Natl. Acad. Sci. U. S. A. 2012; 109:18361–18366. [PubMed: 23071302]

[85]. Okeley NM, Paul M, Stasser JP, Blackburn N, van der Donk WA. Biochemistry. 2003; 42:13613–13624. [PubMed: 14622008]

[86]. a) Toogood PL. Tetrahedron Lett. 1993; 34:7833–7836.b) Polinsky A, Cooney MG, Toy-Palmer A, Ösapay G, Goodman M. J. Med. Chem. 1992; 35:4185–4191. [PubMed: 1433221] c) Burrage S, Raynham T, Williams G, Essex JW, Allen C, Cardno M, Swali V, Bradley M. Chem.-Eur. J. 2000; 6:1455–1466. [PubMed: 10840968] d) Okeley NM, Zhu Y, van der Donk WA. Org. Lett. 2000; 2:3603–3606. [PubMed: 11073655] e) Zhou H, van der Donk WA. Org. Lett. 2002; 4:1335–1338. [PubMed: 11950356]

[87]. Zhu Y, Gieselman M, Zhou H, Averin O, van der Donk WA. Org. Biomol. Chem. 2003; 1:3304–3315. [PubMed: 14584794]

[88]. Chakicherla A, Hansen JN. J. Biol. Chem. 1995; 270:23533–23539. [PubMed: 7559517]

[89]. Kuipers OP, Rollema HS, de Vos WM, Siezen RJ. FEBS Lett. 1993; 330:23–27. [PubMed: 8370453]

[90]. Majchrzykiewicz JA, Lubelski J, Moll GN, Kuipers A, Bijlsma JJ, Kuipers OP, Rink R. Antimicrob. Agents Chemother. 2010; 54:1498–1505. [PubMed: 20100873]

[91]. Zheng G, Yan LZ, Vederas JC, Zuber P. J. Bacteriol. 1999; 181:7346–7355. [PubMed: 10572140]

[92]. Ferguson AL, Zhang S, Dikiy I, Panagiotopoulos AZ, Debenedetti PG, Link AJ. Biophys. J. 2010; 99:3056–3065. [PubMed: 21044604]

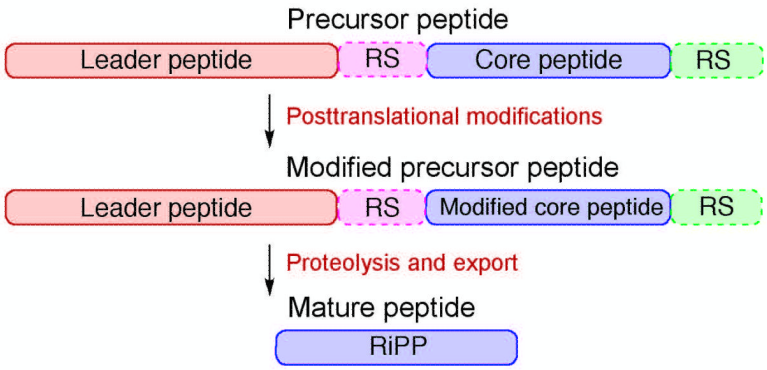

Figure 1.
General biosynthetic pathway of RiPPs. RS = recognition sequence.

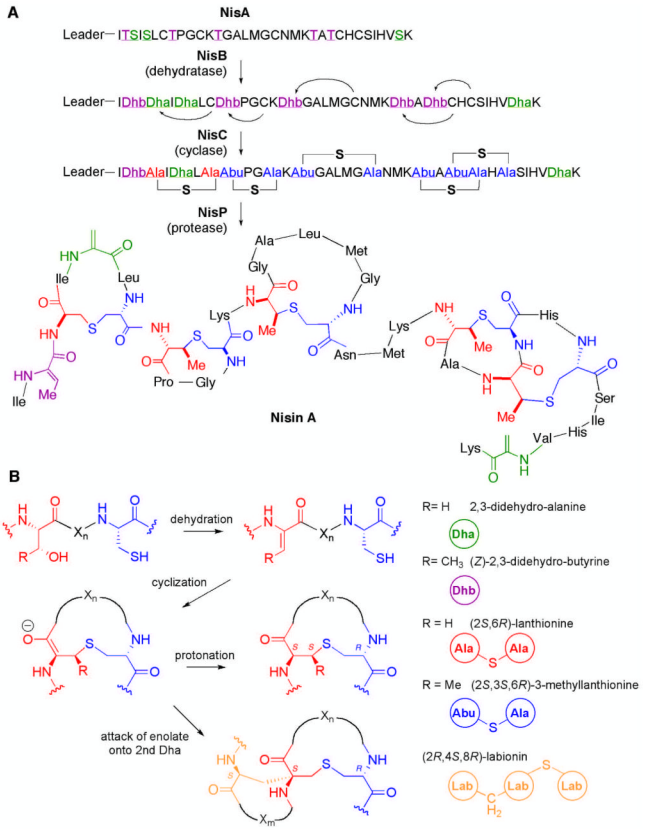

Figure 2.
(A) Biosynthesis of nisin A. (B) Generation of (Me)Lan and labionin motifs.

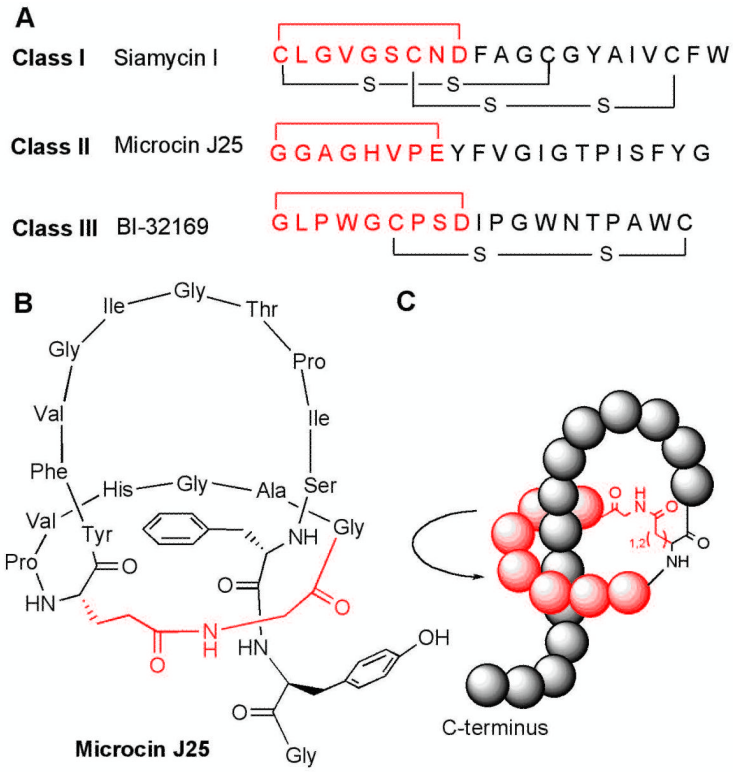

Figure 3.

(A) Three classes of lasso peptides. Residues involved in the macrolactam are shown in red. (B) Structure of microcin J25 as a representative lasso peptide. (C) Righthanded conformation of lasso peptides.

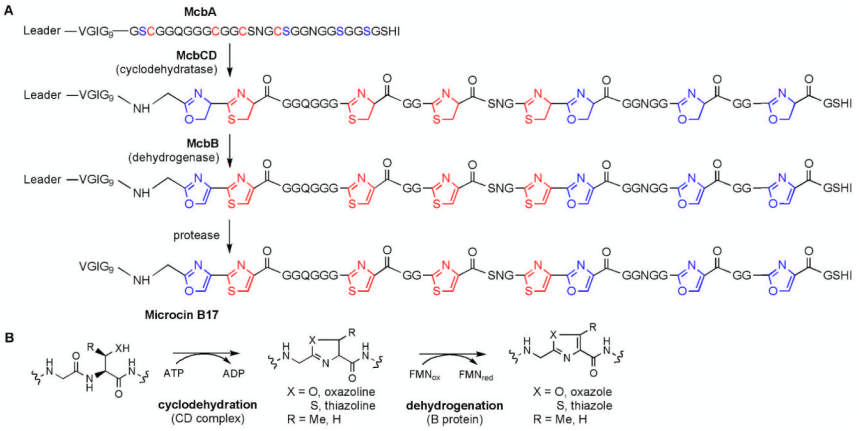

Figure 4.

(A) Biosynthesis of microcin B17 as a representative LAP. (B) Generation of oxazol(in)e and thiazol(in)e motifs.

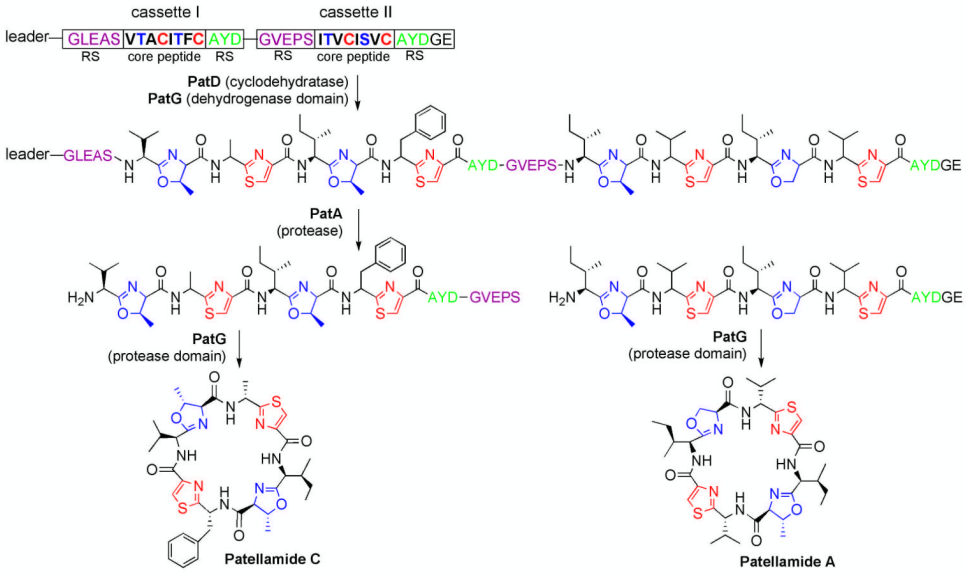

Figure 5

. Proposed biosynthesis of patellamides A and C. The precursor peptide contains an N-terminal leader peptide and two core peptide cassettes. In each cassette, the core peptide sequence is sandwiched between two recognition sequences (purple/green).

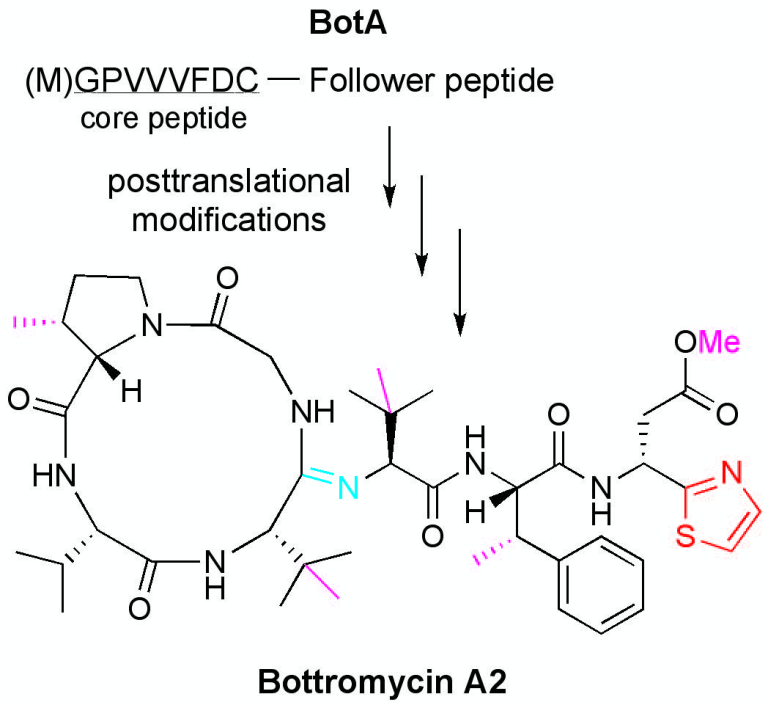

Figure 6.
Structure of bottromycin A2.

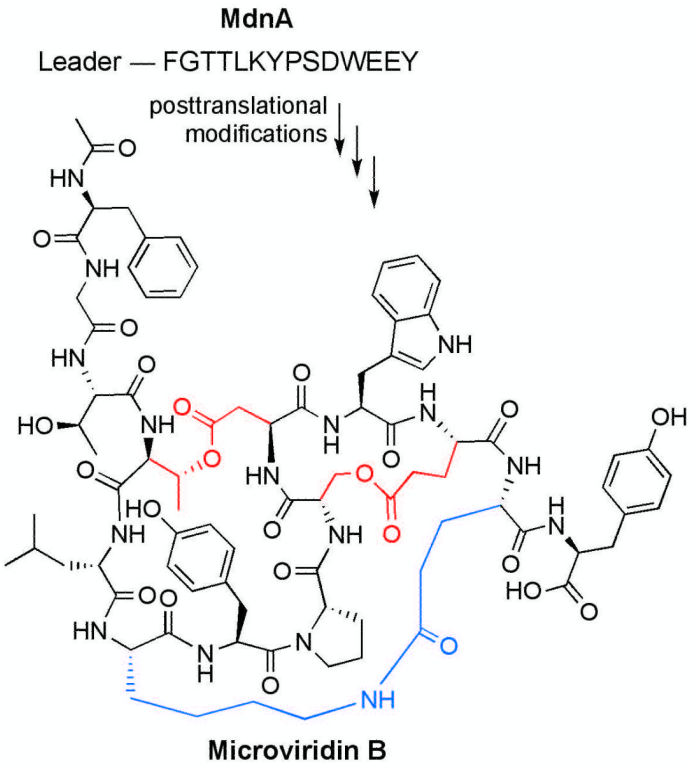

Figure 7.
Structure of microviridin B.

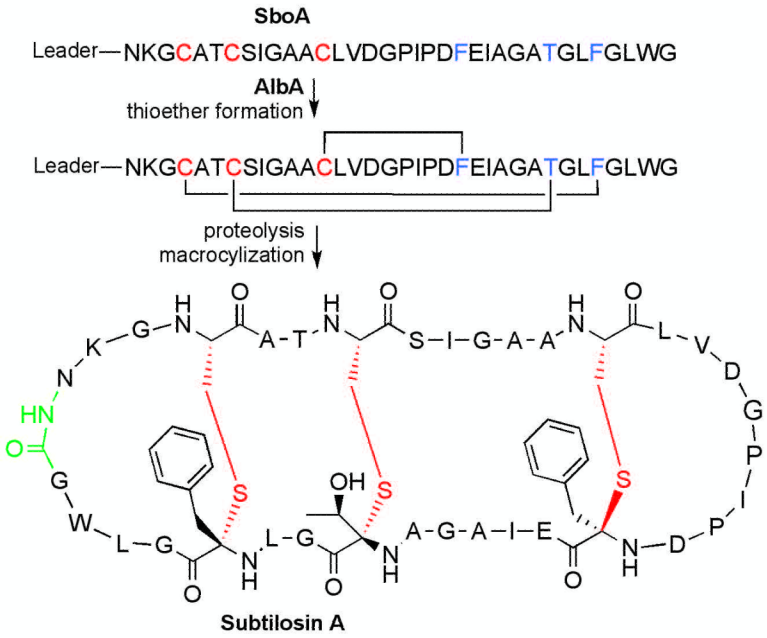

Figure 8.
Proposed biosynthesis of subtilosin A.

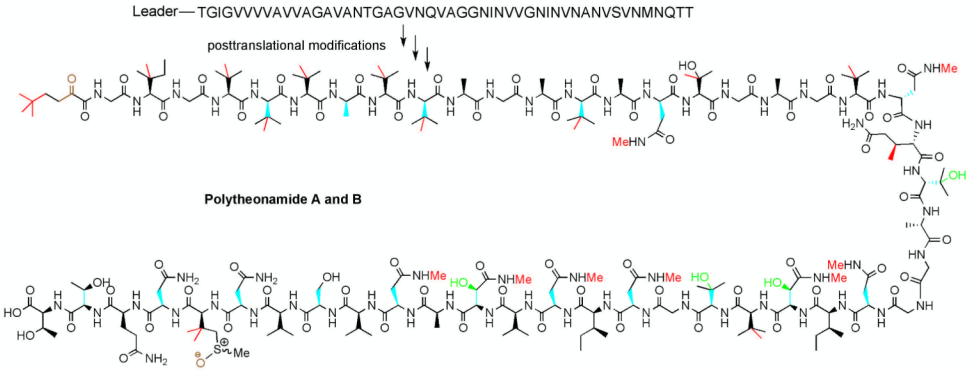

Figure 9.

Structures of polytheonamide A and B. The two peptides differ by the configuration of the sulfoxide moiety. Epimerizations are shown in blue, methylation is shown in red, and hydroxylation is shown in purple.

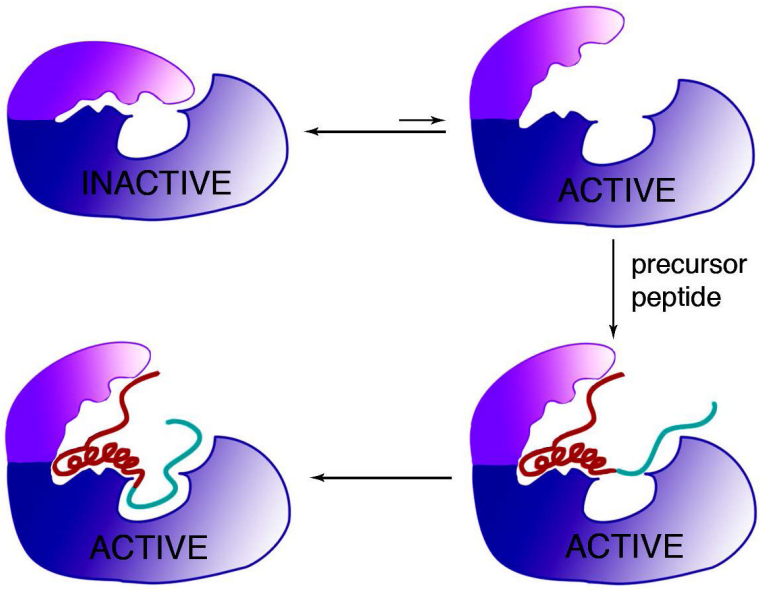

Figure 10.
Proposed role of the leader peptide in activating their biosynthetic enzymes. The leader peptide is shown in red. The core peptide is shown in blue. The potential pseudosubstrate domain of the synthetase is shown in purple.

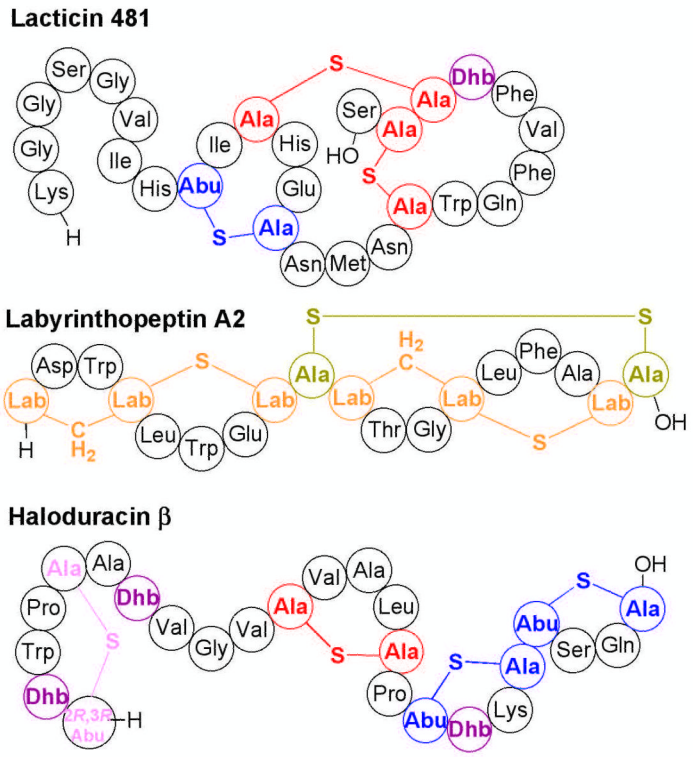

Figure 11.
Some representative lanthipeptides. The same shorthand notation is used as in Fig. 2.

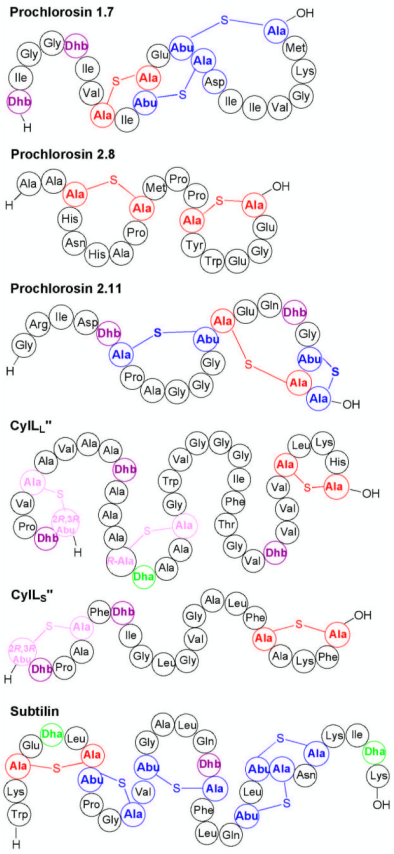

Figure 12.
Lanthipeptide structures that have been used for studies investigating the role of the core peptide. The same shorthand notation is used as in Figure 2. (Me)Lan residues with unusual stereochemistry are shown in pink.