Complex peptide natural products: Biosynthetic principles, challenges and opportunities for pathway engineering

Sebastian L. Wenski\textsuperscript{a,b,1}, Sirinthra Thiengmag\textsuperscript{a,b,1}, Eric J.N. Helfrich\textsuperscript{a,b,*}

\textsuperscript{a} Institute for Molecular Bio Science, Goethe University Frankfurt, 60438, Frankfurt am Main, Germany
\textsuperscript{b} LOEWE Center for Translational Biodiversity Genomics (TBG), 60325, Frankfurt am Main, Germany

\begin{abstract}
Complex peptide natural products exhibit diverse biological functions and a wide range of physico-chemical properties. As a result, many peptides have entered the clinics for various applications. Two main routes for the biosynthesis of complex peptides have evolved in nature: ribosomally synthesized and post-translationally modified peptide (RiPP) biosynthetic pathways and non-ribosomal peptide synthetases (NRPSs). Insights into both bioorthogonal peptide biosynthetic strategies led to the establishment of universal principles for each of the two routes. These universal rules can be leveraged for the targeted identification of novel peptide biosynthetic blueprints in genome sequences and used for the rational engineering of biosynthetic pathways to produce non-natural peptides. In this review, we contrast the key principles of both biosynthetic routes and compare the different biochemical strategies to install the most frequently encountered peptide modifications. In addition, the influence of the fundamentally different biosynthetic principles on past, current and future engineering approaches is illustrated. Despite the different biosynthetic principles of both peptide biosynthetic routes, the arsenal of characterized peptide modifications encountered in RiPP and NRPS systems is largely overlapping. The continuous expansion of the biocatalytic toolbox of peptide modifying enzymes for both routes paves the way towards the production of complex tailor-made peptides and opens up the possibility to produce NRPS-derived peptides using the ribosomal route and vice versa.
\end{abstract}

1. Introduction

Bioactive peptides can be found across all kingdoms of life [1]. Textbook knowledge separates the biosynthesis of complex peptide natural products (NPs) into ribosomally synthesized peptides and peptides that are produced independently of the classical ribosomal route [1–3]. The ribosomal route can be further subdivided into the evolutionary ancient class of cationic and amphiphilic antimicrobial peptides (AMPs), mainly involved in primary immune defense of plants and animals [3], and ribosomally synthesized and post-translationally modified peptides (RiPPs), which are distributed throughout the tree of life and show a wide diversity of biological functions [1]. AMPs predominantly target cell membranes and are usually active against a broad range of bacteria [4]. RiPPs, on the other hand, exhibit diverse biological functions ranging from antimicrobials (e.g., lantibiotics) [5], co-factors (e.g., PQQ) [6], to hormones (e.g., tri- and tetraiodothyronine) [7]. RiPP-derived peptides are ribosomally synthesized and then undergo an extensive post-translational modification process, before proteolytic cleavage results in the release of the mature peptide NP [3,8]. Moreover, five different ribosome-independent routes for the production of complex peptide NPs have been described [9]. These types differ in their carrier protein or tRNA dependency, the mode of amide bond formation, and their modular or non-modular protein architecture (for a comprehensive review see Ref. [9]). Among these, the most prominent group of enzymatic machineries for the biosynthesis of complex peptides are modular non-ribosomal peptide synthetases (NRPSs) which act in an assembly line-like fashion [2,9]. Non-ribosomal peptides (NRPs), unlike ribosomally synthesized peptides, are not limited to the 20 proteinogenic amino acids (AAs) but harbor a variety of naturally occurring, non-proteinogenic AAs in their peptide backbones [1,10]. This review focuses on the comparison of multi-modal NRPSs and RiPPs.

Peer review under responsibility of KeAi Communications Co., Ltd.

\textsuperscript{*} Corresponding author. Institute for Molecular Bio Science, Goethe University Frankfurt, 60438, Frankfurt am Main, Germany.

\textsuperscript{1} authors contributed equally to the work.

https://doi.org/10.1016/j.synbio.2022.01.007

Received 16 November 2021; Received in revised form 19 January 2022; Accepted 20 January 2022

2405-805X/© 2022 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Extensive bioactivity-guided screening efforts have resulted in the isolation of numerous highly complex peptide NPs since the beginning of the golden age of antibiotics [11]. These efforts have led to the discovery of most NP-derived drugs that we are using today. Examples of complex peptides of medical relevance include the immunosuppressant cyclosporin (1) (NRP) [12], the antibiotic of last resort vancomycin (2) (NRP) [12], the drug candidate for the treatment of cystic fibrosis duramycin (3) (RiPP) [13] and the antibiotic nisin (4) (RiPP) (Fig. 1) [5,14]. Some of these metabolites were isolated more than half a century before the genetic blueprints that govern their biosynthesis were deciphered. As a result of the presence of many non-proteinogenic AAs in these peptides [12], it was speculated that they originate from a ribosome-independent route. It was not until 1988, or 49 years after the initial isolation, that the first genes responsible for the biosynthesis of the NRP gramicidin were identified [15–17]. It took about a decade to decipher the universal principles that govern NRPS biosynthesis [18–20]. Even though the biosynthesis of epidermin, which is structurally related to nisin (4), was elucidated in 1988 and showed that complex peptide NPs can also be biosynthesized via the ribosomal route [21], the historic misconception that complex peptides that harbor non-proteinogenic AAs are likely NRPS-derived, persists up to this day. Recent examples of RiPP-derived peptides that were initially believed to be of NRPS origin include the polytheonamides which are extremely cytotoxic 48-mer pore formers. Polytheonamides biosynthesis involves a total of 49 post-translational modifications resulting in the formation of 28 non-proteinogenic amino acids [22,23]. Moreover, the hexapeptide tryptorubin A which is characterized by a highly rigid three-dimensional (3-D) shape was initially believed to be a NRPS product [24]. Reevaluation of tryptorubin biosynthesis, however, led to the realization that tryptorubin A is the first member of a new RiPP family [25]. This misconception can, at least in part, be attributed to the better understanding of NRPS biosynthesis. Insights into NRPS biosynthesis have resulted in the development of sophisticated bioinformatic platforms for the identification and annotation of NRPS biosynthetic gene clusters (BGCs) in genome sequences and the structural prediction of the associated peptides. RiPP BGCs that do not belong to the well characterized RiPP families, on the other hand, are significantly more challenging to be identified in genome sequences. As a result, and even though RiPPs have been the fastest growing class of NPs, we are currently not able to chart the full RiPP biosynthetic potential encoded in microbial genome sequences. While only 22 RiPP families had been reported in 2013 [1], we now know more than 40 RiPP families [8]. Since the arsenal of AA modifications encountered in RiPP and NRPS systems is widely overlapping, it becomes increasingly more difficult to assign a peptide to its biosynthetic origin without having access to the genome sequence of the producer. On the other hand, the realization that both RiPPs and NRPSs use different strategies to introduce a largely overlapping arsenal of peptide modifications opens up an entirely new avenue: the production of NRPS-derived peptides using the ribosomal route and vice versa.

Fig. 1. Structures of the medically relevant peptides cyclosporin (1) (NRPs), vancomycin (2) (NRPs), duramycin (3) (RiPP), and nisin (4) (RiPP).
This review will briefly contrast the key biosynthetic features of NRPS and RiPP biosynthesis, describe the biosynthetic origins of the most common peptide modifications and highlight the different genome mining approaches for the targeted identification of novel RiPP and NRPS-derived peptides. We will then showcase the progress, promise and obstacles in the engineering of both peptide biosynthetic routes and venture a look ahead into the future of engineering peptide biosynthetic pathways.

2. Biosynthesis of complex peptide natural products

The biosynthetic principles of NRPSs and RiPPs fundamentally differ with respect to recognition, activation and condensation of the AA building blocks as well as the subsequent modification and maturation of the corresponding peptides (Fig. 2). The following section briefly introduces and contrasts the biosynthetic principles of NRPS and RiPPs pathways.

2.1. NRPS-derived peptide biosynthesis

NRPSs can be subdivided into a nonlinear, iterative, or linear type [26]. This review focuses on linear NRPSs, which constitute the most prominent subclass. Modular, non-iterative NRPSs are large mega enzyme complexes that resemble assembly lines. These systems, much like assembly lines in manufacturing processes, can be subdivided into individual modules. Each module incorporates one building block into the nascent oligopeptide chain. The biosynthesis of NRPs is directional and starts with the loading of the first AAs at the N-terminal module and ends with the release of the peptide at the C-terminal module. The primary sequence of the resulting oligopeptide results from the selection of the incorporated AAs by each module. Facultative enzymatic domains in each module modify the AA incorporated by the respective module. This correlation between NRPS architecture and the structure of the associated peptide is referred to as the colinearity rule.

A module is composed of an adenylation (A) domain, which selects and activates an AA, a peptidyl carrier protein (PCP/P), responsible for the tethering of the activated AA and the nascent peptide intermediate, and a condensation (C) domain that catalyzes peptide bond formation [27] (Fig. 2). Similar to other adenylation enzymes like acyl-CoA synthetases and firefly luciferases, A domains catalyze two reactions: an initial adenylation of a free AA followed by the transfer of the activated AA onto the phosphopantetheinyl arm of a PCP domain [19, 28]. Since a variety of non-proteinogenic AAs can be selectively activated, these domains are the driving force of structural diversity of NRPs [2, 10].

Fig. 2. Schematic overview of the key principles of NRP and RiPP biosynthesis. (A) NRPS and BGC-encoded tailoring enzymes are ribosomally translated. The NRPS assembles the peptide, which is modified on-line (red circle) and released via intramolecular cyclization (purple bond). Posttranslational modifications (brown circle) via BGC-encoded enzymes result in the mature peptide. (B) Typical RiPP BGCs harbor genes encoding a precursor peptide, modifying enzymes and a protease. The ribosomally biosynthesized precursor peptide is post-translationally modified. The leader peptide (grey balls) serves as a recognition sequence for tailoring enzymes that modify the core peptide sequence (colored balls). Once the core peptide is fully modified (purple bond, small red and brown circles) it is released from the leader peptide by a protease. A: Adenylation domain; P: Peptidyl carrier protein; C: Condensation domain; T: Thioesterase domain. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
Commonly occurring non-proteinogenic AAs in NRPs are frequently hydroxylated, methylated or halogenated. In addition, β- or homo-AAs are often incorporated [2,10]. These unusual building blocks are derived from primary metabolites that are either modified in a NRPS-dependent manner or by other gene cluster-encoded enzymes (Fig. 2). The NRPS-dependent generation of non-proteinogenic AAs is catalyzed by stand-alone NRPSs with the domain architecture A-PCP or A-PCP-X (X: variable modification domain). AAs are activated, loaded onto these monomodular NRPSs, and presented to trans-acting modifying enzymes [2,10]. The generation of hydroxylated AAs illustrates the diversity of building block biosynthesis: AAs can be hydroxylated either in an NRPS-dependent (e.g., echinomycin or kutzneride) or in a calcium-dependent antibiotic CDA fashion [29–31]. The corresponding hydroxylation is performed either via cytochrome P450-monoxygenases (e.g., echinomycin) or α-ketoglutarate-dependent hydroxylases (e.g., kutzneride or CDA) [29–31]. Similarly, halogenation can also occur on free as well as on NRPS-bound AAs [32,33]. Modified AAs can be hydrolytically released and subsequently re-activated by an A domain of the core NRPS or aminoacyl transferases catalyzing the direct transfer of the modified AAs from the PCP of the stand-alone NRPS to the corresponding PCP of the core NRPS [2,34] (Fig. 2).

A domains harbor a binding pocket with multiple specificity-conferring residues responsible for the selective AA recognition [19]. These conserved residues can be used to predict the A domain’s substrate specificity [19,20]. In some cases, A domains are able to recognize multiple, structurally related AAs as the specificity-conferring residues exhibit nearly identical binding affinities for multiple building blocks [2,35]. Consequently, the corresponding NRPS generates a mixture of NRP analogs which, based on the so-called screening hypothesis, was postulated to function as a means of adaptation to changing environmental challenges [36–38]. Consequently, this biosynthetic promiscuity can be regarded as a feature rather than a bug of the system.

PCP domains belong to the large family of carrier proteins [39]. Phosphopantetheinyl transferases catalyze the attachment of a phosphopantetheinyl arm to carrier proteins which results in the formation of active holo PCP domains [40]. During NRPS biosynthesis, AA building blocks and peptide intermediates are covalently attached to the phosphopantetheinyl arm of PCP domains via a thioester bond. PCPs are responsible for shuttling activated AAs and the growing peptide chain to the catalytic centers of the module-encoded domains and trans-acting modifying enzymes [40]. Since building blocks and peptidyl intermediates are covalently bound to the PCP domains, the NRPS serves as a template for the biosynthesis of complex peptides and the biosynthetic principle is therefore also referred to as thiotemplated biosynthesis.

As soon as the PCPs of two adjacent modules are loaded, the C domain catalyzes the formation of a peptide bond via nucleophilic attack of the C-terminal α-amino residue onto the thioester of the N-terminal intermediate. As a result, the growing oligopeptide is elongated with one building block, and then transferred onto the C-terminal PCP before the elongation process repeats at the next module [41]. In addition to canonical C domains, several homologues have been described which are involved in heterocyclization, epimerization, attachment of carboxylic acids or oligopeptide release ([42]; for a comprehensive review see Ref. [43]).

After the final elongation, the oligopeptide is transferred onto a type 1 thioesterase (TEI) domain via transesterification to form a TE oligopeptide ester. TEI domains catalyze either a water-mediated hydrolyzation to release a linear product or an intramolecular attack from a hydroxyl or amino group which results in macro lactone or macro lactam formation (Fig. 2) [44,45].

Insights into the biosynthesis of NRPS-derived peptides have resulted in the development of several generations of highly sophisticated bioinformatic platforms for the identification and annotation of NRPS BGCs in genome sequences and the prediction of A domain substrate specificities. NRPS BGCs can be identified by hard-coded biosynthetic rules as NRPS core genes are composed of different arrangements of conserved biosynthetic domains [46–48]. This conservation enables domain annotations by profile Hidden Markov Models. The identification of specificity-conferring residues in the substrate binding pocket of A domains, also referred to as the Stachelhaus code, allows the prediction of the incorporated AA building blocks and thus the prediction of the peptide’s primary AA sequence [19,20]. Several algorithms for the prediction of A domain substrate specificities have been developed since the initial description of the Stachelhaus code which has resulted in predictions with an ever-increasing accuracy. This increased accuracy can be attributed to the evolution of genome mining algorithms and the steady increase in characterized NRPs that were used to train novel genome mining algorithms [46–51]. Despite these advancements in A domain substrate specificity predictions, the promiscuity of many A domains often impedes the accurate prediction of the primary AA sequence. Even though the primary peptide sequence of NRPS-derived peptides cannot be predicted with as much confidence as in the case of ribosomally synthesized peptides, NRPSs nevertheless belong to the NP biosynthetic machineries that can be best studied using current genome mining platforms [46,48,50,52]. This advantage over other NP classes such as RiPPs can be attributed to its well-studied assembly line-like character. Conceptually, new NRP scaffolds are the results of novel arrangements of the limited set of NRPS modules with different A domains to form new NRPs architectures. Due to this simple biosynthetic principle, the full biochemical space of canonical NRPSs can be charted by state-of-the-art bioinformatic platforms.

2.2. RiPP-derived peptide biosynthesis

Unlike multi-domain mega synthases that catalyze the biosynthesis of complex peptides in an assembly line-like fashion, RiPPs utilize the classical ribosomal route for the production of peptide precursors [53]. The architecture of a typical RiPP BGC consists of a structural gene encoding a precursor peptide, gene(s) involved in peptide modification and a protease (Fig. 2). A ribosomally biosynthesized precursor peptide generally comprises an N-terminal leader peptide, a C-terminal core peptide, and occasionally a C-terminal follower peptide (e.g. botromycin) [1,8]. In rare cases, N-terminal leader and C-terminal follower peptides can be found within the same precursor peptide (e.g., pantocin) [54,55]. During ribosomal peptide biosynthesis, aminoacyl-tRNA synthetases have a function similar to A domains in NRPSs. Each aminoacyl-tRNA synthetase catalyzes the ATP-dependent activation of its dedicated AA to form an aminoacyl-adenylate [10]. The activated AA is then transferred to its corresponding tRNA in a transsterification reaction, resulting in an aminoacyl-tRNA which is then used in ribosomal peptide biosynthesis [10,56]. However, A domains and aminoacyl-tRNA synthetases share neither sequence nor structural similarities [57]. Following translation by the ribosome, the precursor peptides undergo extensive post-translational modifications (PTMs) catalyzed by an ever-increasing set of modifying enzymes (Fig. 2). The N-terminal leader and C-terminal follower peptides serve as recognition sequences for modifying enzymes. In addition, leader/follower peptides play a role in guiding PTM enzymes to perform modifications in the correct order [1,8]. Once the core peptide is fully modified, it is released from the leader and/or follower peptide by proteolytic cleavage (Fig. 2). Typically, specific proteases are encoded in RiPP BGCs but more and more pathways are reported that utilize ubiquitous cellular proteases for the release of the mature RiPP [1,8,25]. Exceptions from this simple RiPP biosynthetic principle include the cyanobactins, dikaritins and cyclotides BGCs in which multiple core peptides are encoded in one large precursor gene [58]. In these cases, the precursor peptide contains an N-terminal leader peptide followed by multiple core peptides. Highly conserved N-(RSII) and C-terminal (RSIII) recognition sequences flank each core peptide. These recognition sequences are required to guide proteolytic cleavage and macrocyclization. Depending on the RiPP
family, these multiple core peptides can either have the same sequence or differ extensively as shown for the highly variable core peptides of the cyanobactins [59].

Even though RiPPs follow the seemingly simple and universal biosynthetic principle outlined above, they are a very inhomogeneous subclass of peptide NPs. Thus, each RiPP family bears a unique biosynthetic logic [60,61]. This inhomogeneity is reflected in the wide range of genome mining platforms developed to chart the biosynthetic

**Fig. 3.** The most frequently encountered modifications in NRPs and RiPPs. The color code indicates whether a modification has been described for NRPS-derived peptides (blue), RiPPs (green) or was reported for both RiPP and NRPS routes (yellow). Modifications labeled with an asterisk (*) can occur at different positions and atoms. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
space of individual families rather than the entire RiPP biosynthetic diversity [60]. While precursor peptides and modifying enzymes show a high degree of homology within each RiPP family, no core genes or domains are conserved between all RiPP families [62,63]. Some RiPP families are characterized by conserved precursor genes, others employ characteristic tailoring enzymes based on which they can be identified in a bait-based approach. These bait-based approaches allow the identification of novel members of characterized RiPP families with high confidence [48,50,52,64–67]. Based on the biosynthetic insights from other members of the same RiPP family, the precursor sequence and the type and regioselectivity of these modifications are usually not predictable [68]. To chart the RiPP biosynthetic potential beyond RiPP family boundaries, classical homology-based approaches have been complemented with several machine learning-based algorithms for the targeted identification of novel RiPP families [69–72]. These algorithms target putative precursor genes, recognition elements that are conserved amongst several RiPP families and tailoring enzymes with homology to modifying enzymes of characterized RiPPs [69–72]. While these highly sophisticated approaches will likely expand the RiPP biosynthetic space significantly, there is currently no individual tool available which is capable of charting the full biosynthetic space of RiPP-derived NPs [68, 73].

3. Peptide modifications

In comparison to RiPPs, NRPS-derived peptides are modified at different biosynthetic stages. The chemical diversity of NRPs is mainly based on adenylation domain substrate specificities, module order, composition and the presence of trans-acting enzymes. In rare cases, NRPs are modified after the oligopeptidase has been released. In contrast, only proteinogenic AAs are incorporated into RiPP-derived peptides. After translation, the precursor peptide is modified, leading to the conversion of proteinogenic AAs into a wide range of non-proteinogenic AAs. In addition, some RiPP families are characterized by complex 3-D structures. In the following paragraphs differences and similarities of modifications in NRPS and RiPP derived peptides are discussed (Fig. 3).

3.1. Epimerization

The incorporation of D-configured AAs influences the overall structure and bioactivity of peptides and prevents degradation by proteases [74]. NRPSs introduce D-AAs via epimerization (E), dual condensation/epimerization (CE) and A domains (Fig. 3) [42,75–77]. The epimerization via E and CE domains takes place after the incorporation of an L-AA. The position of the module-encoded E or CE domains in the NRPS assembly line determines which AA is epimerized. The catalytic mechanism of E domains is proposed to comprise α-proton abstraction via the catalytic glutamate of the E domain followed by racemization [42,78]. Interestingly, in CE domains the catalytic glutamate is missing, suggesting an alternative mode of epimerization [2,42,78]. Epimerization domains are followed by a distinct C domain subtype, the $^\text{6}C_l$ type [42]. The stereoselectivity of the $^\text{6}C_l$ domain enables the exclusive incorporation of the $\alpha$-isomer from the racemic mixture generated by the E domain [79]. In rare cases, A domains were shown to directly activate D-AAs that are generated via cytoplasmic racemases [75–77].

In contrast to NRPs, direct incorporation of D-AAs into RiPP scaffolds does not occur in nature since ribosomal peptides are restricted to the 19 proteinogenic L-configured AAs and glycine. Therefore, incorporation of D-AAs requires post-translational epimerization by PTM enzymes. In RiPPs, epimerization can occur via a direct or indirect mechanism [5]. Direct epimerization of L-AAs to D-AAs has been shown to follow a radical mechanism catalyzed by radical S-adenosylmethionine (rSAM) enzymes. PoyD is the first characterized rSAM epimerase that performs a total of 18 epimerizations in an alternating fashion during polytheonamide biosynthesis in a C- to N-directional manner [22,23,80]. AA epimerization by PoyD-like enzymes is initiated by reductive cleavage of SAM to generate a 5′-deoxyadenosyl radical ($^\text{5}′$-dA•) which abstracts the C6 H-atom from L-AAs to form a carbon-centered radical. The thiolate proton of cysteine in the epimerase is then transferred to the radical intermediate leading to the formation of D-AAs in a radical rebound mechanism [81]. In contrast to this radical mechanism, the indirect epimerization mechanism requires two enzymatic steps that have been reported for some lanthipeptides. Epimerization is initiated by dehydration of $\gamma$-serine to form dehydroalanine (Dha) by the dehydratase LanB. Once Dha is formed, the dehydrogenase LanJ catalyzes the diastereoselective hydrogenation to yield D-alanine [82,83]. Moreover, Both-H-like enzymes, belonging to the subfamily of $\alpha$-/β-hydrolase (ABH) fold proteins, have been shown to be involved in yet another epimerization route that converts $\gamma$-aspartate to D-aspartate during bottromycin biosynthesis. The proposed mechanism involves the self-abstraction of an $\alpha$-proton by the carboxylic acid group of the AA side chain followed by proton transfer from a water molecule to yield the epimerized AA [84].

3.2. Heterocyclization

Thiazolines and (methyl) oxazolines are characteristic for multiple NRP and RiPP families (Fig. 3). During NRPs biosynthesis, cyclization (C) domains, homologues of C domains, catalyze heterocycle formation in a bifunctional manner: First, the NRP intermediate is elongated by the C domain either with cysteine, threonine or serine. Subsequently, C domains catalyze the nucleophilic attack of the thiol or hydroxy group of the AA side chain onto the carbonyl carbon of the amide bond. Subsequent dehydrogenation results in (thi/o)xazoline formation [85]. Thiazoline- or oxazoline heterocycles can either be reduced via trans-acting reductases to form thiazolines/oxazolidines or oxidized via cis or trans-acting oxygenases to form thiazols/oxazols [86–90]. Cysteine-, threonine- or serine-derived heterocycles have also been reported for many RiPPs including linear azol(in)e-containing peptides (e.g., microcin B17) [91], bottromycin [92], and cyanobactins (e.g., patellamide A) [93]. The mechanism underlying thiazole and oxazole formation requires an enzyme complex that is comprised of one or two cyclodehydratase(s) (C/D protein) and a dehydrogenase (B protein). First, cysteine, threonine or serine undergo ATP-dependent cyclodehydration catalyzed by the cyclodehydratase to form an azoline heterocycle. Then, a flavin-dependent dehydrogenase oxidizes the azoline rings to yield azole heterocycles. In some cases, including cyanobactins (e.g., turamidine) and half of known linear azol(in)e-containing peptides (LAPs), C- and D-proteins are fused into one single enzyme [8,94, 95].

3.3. Methylation

Peptide methylation is a simple way to influence polarity, improve proteolytic resistance, facilitate cellular uptake and increase the half-life of a peptide, all of which are important parameters for bioavailability [96]. In NRPSs, the majority of methylation (M) domains act in cis and are directly integrated into a flexible loop of the A domain. M domains catalyze mainly N- but also O-, C- or S-methylations in a SAM-dependent manner [2,97–99] (Fig. 3). Remarkably, during cyclosporin biosynthesis, methylation is critical for the overall peptide assembly and cyclization [100]. In rare cases, methylation takes place after peptide release via tailoring enzymes like the stand-alone methyltransferase MtFa in glycopeptide biosynthesis [101].

Polytheonamide, one of the most densely modified RiPPs, has been shown to carry eight N-methylations and seventeen C-methylations. A rSAM-dependent methyltransferase, PoyE, is responsible for all N-methylations. While the two vitamin B$_2$-dependent C-methyltransferases, PoyB/C catalyzes all C-methylations [22,23]. PoyC has been shown to catalyze the homolytic cleavage of SAM to generate $^\text{5}′$-dA• that abstracts the C6 H-atom of $\gamma$-valine residues. The
carbon-centered valine radical likely reacts with the methyl radical to yield Cβ methyl-valine [80]. Computational studies of polytheonamide B revealed that N-methylation stabilizes the unusual β-helix conformation [102]. Beside N- and C-methylation, a rare S-methylation of cysteine residues mediated by a SAM-dependent methyltransferase takes place during the biosynthesis of a cryptic protein in the sponge symbiont "Candidatus Entotheonella factor" (Fig. 3) [103]. Moreover, a rSAM-dependent methyltransferase that methylates carboxyl groups has been reported during the final step of botromycin biosynthesis resulting in methyl ester formation at the aspartate residue [92].

3.4. Macrocyclization

Macrocyclization is a common feature found in peptide NPs, derived from both NRPS and RiPP pathways [104]. Macrocyclization plays a crucial role for the biological activity of many compounds and enhances peptide stability by protecting from proteolytic digestion [53, 105]. Crucial role for the biological activity of many compounds and enhances a carbonyl carbon resulting in cyclization by carbon-carbon bond formation (e.g., glycocin), the formation of lactone/lactam rings via Cα Cβ bond formation (e.g., glycocin), the formation of lactone/lactam rings between AA side chains (e.g., microviridins) as well as macrocyclization via C–C and C–S bond formation [112] (Fig. 3). Head-to-tail macrocyclization is a common PTM found in RiPPs [105]. Several enzymes involved in macrocyclization have been experimentally characterized including the PatG protease that is involved in patellamide biosynthesis [8]. PatG contains a subtilisin-like serine protease domain which recognizes and cleaves a signal sequence on the precursor peptide. This cleavage results in the formation of an acyl-enzyme intermediate, where a serine residue of the catalytic triad of the protease is bound to the core peptide. In the next step, the acyl-enzyme intermediate is attacked by the N-terminal amino group of the peptide to form a macro lactam. This mechanism is similar to macrocyclizations catalyzed by TE domains in NRPs that contain the same catalytic triad as the serine protease PatG and that likewise mediate macrocyclization via an acyl-enzyme intermediate [104]. Macrocyclization via Michael-type addition is an alternative mechanism to construct cyclic structures in RiPPs as shown for the formation of lanthionine bridges during lanthipeptide biosynthesis [112]. A lanthionine (Lan) bridge is a thioether crosslink between the β-carbons of serine/threonine and cysteine [83]. In nisin biosynthesis, for instance, it has been shown that Lan and (methyl)Lan formation is a native mechanism to construct cyclic structures in RiPPs as shown for subtilosin biosynthesis. The rSAM enzyme AlbA catalyzes the formation of the uniquely defined sactionine thioether cross-link between the thiol residue of cysteine and the α-carbon of phenylalanine or threonine [115]. The carbon-centered radical then attacks the sulfur atom to form the thioether bond [112].

3.5. β-amino acids

The incorporation of β-AAs leads to structural diversity and increases proteolytic resistance of peptides [116]. Many NRPs are composed of one or multiple β-AAs like β-tyrosine in chondramide [117], β-lysine and 2,3-diaminopropionate in vionycin [118] or β-alanine in cryptophycins [119], for a comprehensive review see Ref. [116]). In NRPS systems, A domains recognize and subsequently activate free β-AAs as unusual building blocks, derived from a variety of different pathways [116]. In contrast, during RiPP biosynthesis β-AAs cannot be directly incorporated but proteinogenic AAs can be converted into β-AAs [120]. Piel, Morinaka and co-workers have recently identified a rSAM-dependent mode to convert proteinogenic AAs into α-keto-β-AAs. The reaction involves the unusual radical excision of thymamine and the rejoicing of the remaining peptide fragments to create an α-keto-β-AA [120]. Moreover, the installation of a β-AA has been described for the lanthipeptide OlvA (BCSβ). In the case of OlvA(BCSβ), a SAM-dependent O-methyltransferase catalyzes the conversion of aspartate to the β-AA i-aspartate through methylation, followed by imide formation between the carboxyl group of the AA side chain and the amide bond of the peptide backbone. Subsequent imide hydrolysis can result in the transfer of the peptide backbone to the former aspartate side chain, resulting in β-AA formation [121].

3.6. Carbon-carbon bond formation

The introduction of carbon-carbon bonds strongly influences the structure and bioactivity of NPs. The bioactivity of vancomycin-like NRPs, for instance, depends on the formation of the rigid aglycon structure that is crucial for target binding [122–124]. During vancomycin biosynthesis, the peptide scaffold is cyclized by the cytochrome P450 oxygenases OxyA, OxyB, and OxyC through the introduction of two diaryl ether bridges and a biaryl bond (Fig. 3) [122]. Remarkably, the trans-acting oxygenases are recruited by a non-catalytic C domain-like domain, called X domain, which is embedded in the NRPS assembly line [125]. However, inactivation of the oxygenases has no influence on the biosynthesis of the peptide backbone [126].

Carbon-carbon bond formation in RiPPs is frequently catalyzed by rSAM enzymes (Fig. 3) [127]. The rSAM catalyzed carbon-carbon bond formation between non activated carbons of tryptophan and the β-carbon of lysine, for instance, has been characterized for darobactin biosynthesis [128]. Recently, a second mode of carbon-carbon bond formation has been proposed. Tryptorubin A, a hexapeptide that is characterized by an unusual complex 3-D shape features a carbon-carbon and two carbon-nitrogen bonds between two AA side chains and between one AA side chain and the peptide backbone, respectively, that are putatively installed by a single cytochrome P450 in an atropospecific fashion [25]. Shortly after the putative tryptorubin BGC was identified, similar BGC architectures were reported. The bicyclic tetrapeptide, cittilin contains biaryl and aryl-oxgen-aryl ether bonds that are installed by the cytochrome P450 enzyme, GtB [129]. Similarly, the cytochrome P450 monooxygenase, BytO is responsible for the installation of an unusual biaryl bridge between tyrosine and histidine in a short pentapeptide precursor during the biosynthesis of the biaryltyide tripeptide [130].
3.7. Lipidation/prenylation

Peptide lipidation increases stability, decreases polarity and can enable membrane interactions [131]. During NRPS biosynthesis, N-terminal C domains, so-called \( \epsilon \)-thiuram domains, catalyze the condensation of carboxylic acids, in the form of fatty acyl-coenzyme A esters, with the \( \epsilon \)-amino group of the first AA (Fig. 3). This N-aclylation of NRPs is a common feature, connecting NRPSs with fatty acid metabolism [42, 132]. In addition, diketopiperazines (e.g., cycloamidizine, echinulim, notoamide) but also linear peptides (e.g., cyclomarins) are frequently lipidated via prenylation with isoprene moieties [133–135] (Fig. 3). In contrast to N-aclylation, prenyltransferases install prenyl groups on carboxylic acids or oxygen atoms after peptide release [135,136].

Lipidation in RiPPs is rare. Piel and co-workers have recently reported a new family of RiPP-derived lipopeptides, the selidamides [137]. Kamaipotamide, phaseonamide and nostolysamides are the first members of the selidamide family of RiPP-derived lipopeptides that is characterized by fatty acyl moieties attached to the side chain of (hydr)xylose, threonine or lysine, respectively. Heterologous expression studies suggest that every peptide of the selidamide family is selectively modified by a fatty acid with fixed chain length by members of the GNAT–like family of peptide matruses [137]. In the case of the lipolanthine family of RiPPs (e.g., microvin and goadvion), fatty acids are attached to the N-terminal amino group of the peptide backbone by members of the GNAT family [138–140]. Since lipolanthines are biosynthesized by hybrid polyketide (PK)-RiPP pathways, they will be discussed in the crosstalk section below.

Similarly, N-terminal acylation of RiPPs has been reported for lasso peptides (e.g., albusnodin), the LAP (e.g., goadsporin) and microviridins, which is likewise catalyzed by members of the GNAT superfamily [140–142]. Prenylation is a common modification in cyanobactin biosynthesis (Fig. 3). The first characterized prenyltransferase, LynF, has been shown to catalyze the head or tail installation of isoprene units onto serine, threonine and tyrosine [143]. However, after enzymatic O-prenylation via the isoene’s C3 carbon (head), a non-enzymatic Claisen rearrangement results in the transfer of the prenyl group to the ortho position while simultaneously inverting the prenyl linkage from the C3 (head) to the C1 (tail) carbon of the isoprene unit [143,144]. Apart from cyanobactins, the quorum-sensing compound ComX is also prenylated. The prenyltransferase ComQ catalyzes the prenylation at the indolic C3 position of tryptophan in the core peptide with geranyl prenylation. The prenyltransferase ComQ catalyzes the prenylation at the N-terminus [162–164]. In contrast to the fatty acid attachment mediated by \( \epsilon \)-thiuram domains, the formylation reaction of the first AA is catalyzed via formyl transferases under consumption of formyltetrahydrofolate [42,162–164]. To the best of our knowledge, no formylated RiPPs have been reported to date.

In peptide NPs, peptide bonds can be converted into thioamides. The YeaO homolog TvaH has been proposed to catalyze thioamidation during the biosynthesis of thiocyclamidamide, a member of the thiamidamide RiPP family [165].

In addition to the frequently encountered peptide modifications, different modes for the formation of non-proteinogenic AAs have been described. In NRPS systems, non-proteinogenic AAs are directly incorporated, whereas proteinogenic AAs have to be transformed into non-proteinogenic AAs after incorporation into the precursor peptide in RiPPs. The conversion of arginine into the non-proteinogenic AA ornithine by an arginase during landornamide A biosynthesis displays an example for this biosynthetic strategy [166].

4. Crosstalk to other NP pathways

The diversity of NRPS- and RiPP-derived peptides is not restricted to the incorporation of different AA building blocks or the ever-increasing number of characterized tailoring reactions and enzymatic domains. In addition, NRPSs as well as RiPPs interact with other pathways to generate hybrid NPs. This hybridization takes place in a precursor-directed manner via the integration of intermediates or products from other pathways. Furthermore, the biosynthetic core enzymes directly interact either in a covalent or non-covalent manner. This chapter briefly describes different types of hybrids with involvement of NRPSs or RiPPs.

4.1. NRPS

NRPSs are predestined for hybridization with other pathways as A domains are able to recognize a wide variety of substrates beyond the 20 proteinogenic AAs. Moreover, NRPSs are biosynthesized in a thiotemplated, assembly line-like fashion, a common mode of biosynthesis that NRPSs share with polyketide synthases (PKSs) and fatty acid synthases (FASs).

Since the boundary between the incorporation of unusual building blocks and the interaction with further metabolic pathways is fluid, we consider any exchange between metabolic pathways as crosstalk (Fig. 4). The non-proteinogenic AA dihydroxyphenylglycine, for instance, is generated by the type III PKS DpgA and is subsequently incorporated into the growing balhimycin NRP, a vancomycin-type antibiotic [167,168]. Similarly, during cyclosporin biosynthesis, a PKS is involved in the formation of the non-proteinogenic AA 4-(2-butenyl)-4-methyl-threonine [169]. In addition, unusual building blocks, derived from primary metabolic pathways like the citrate cycle, are often incorporated after further modifications (170, for a comprehensive review see [2]). During the biosynthesis of CDA, for instance, \( \alpha \)-ketoglutarate is methylated and subsequently transaminated to generate 3-methyl glutamate [170].

Furthermore, PCP domains, which belong to the carrier protein family, play a central role during hybridization of NRPSs with other...
4.2. RiPPs

In contrast to NRPSs, RiPP hybrids are rare. This infrequent occurrence might be due to the unique RiPP biosynthetic principles that make it inherently more difficult to form hybrids. Alternatively, it might be a result from a characterization bias meaning that hybrid BGCs might have been bioinformatically identified but not experimentally validated. Microvionin is the first member of the lipolanthine RiPP family of lanthipeptide-fatty acid hybrids. It is composed of a triamino-dicarboxylic acid moiety, the so-called avionin residues (Fig. 3), and a bismethylated guanidino fatty acid (MGFA) at the N-terminus. The microvionin BGC harbors a total of ten genes in addition to those that encode enzymes responsible for lantipeptide biosynthesis. As the corresponding proteins are homologues of FAS- or type II PKS-derived enzymes, a FAS/PKS-RiPP hybrid biosynthesis was postulated [138]. In addition, genome mining revealed more than 80 lipolanthine BGCs from actinobacterial genomes, leading to the classification of four lipolanthine subtypes. These BGC subtypes are distinguished by the presence of type I or type II PKS-related genes and lanD, which encodes a cysteine decarboxylase. Interestingly, the subtype I as well as an additional unique cluster harbor genes coding for NRPS-derived domains, indicating a putative PKS-NRPS-RiPP hybrid (Fan [139]).

5. Engineering

In addition to identifying novel BGCs and characterizing the corresponding products, engineering of NRPS and RiPP pathways is an important means to increase structural diversity. Furthermore, dissection and engineering of individual pathway components opens up the possibility to gain deeper biosynthetic insights. Engineering NRPS and RiPP pathways require different strategies and techniques due to their fundamentally different biosynthesis principles. While NRPS engineering mainly focuses on manipulating the AA sequence of NRPs, RiPP engineering targets the variety of modification processes and their applicability to different core peptides.

5.1. NRPS

The collinearity between the NRPS architecture and its product is predestined for engineering. The first successful attempts to manipulate the primary AA sequence of NRPs were described in 1995 [183]. At the time, the Marahiel lab was able to swap A-P didomains in the surfactin BGC to obtain non-natural NRPs [183]. In the following years, many comparable approaches that relied on multiple domain exchanges were published (for a comprehensive review see Ref. [2]) (Fig. 5). In many
cases, however, domain exchanges resulted in a decrease in peptide titers, while the diversity of the resulting NRPs remained limited. Since the identification of the A domain binding pocket and the subsequent deciphering of the Stachelhaus code, the A domain moved into the focus of engineering efforts [19,20]. Single or multiple AAs in the A domain’s binding pocket were exchanged to alter its substrate specificity without drastic decreases in peptide titers (Fig. 5). However, using such minimal invasive approaches, the A domain selectivity only marginally changed regarding structure or polarity of the AAs [184]. Following studies reported the incorporation of AA analogs like 3-methyl glutamine or O-propargyl-l-tyrosine using site-directed mutagenesis [185,186]. Furthermore, using a directed evolution approach, the specificity of an A domain was successfully changed from l-phenylalanine to l-alanine. During this approach, the non-conserved, specificity-conferring binding pocket residues of a promiscuous A domain were engineered via successive saturation mutagenesis [187]. Although many examples of successful NRPS engineering have been reported, universal principles were still lacking. Until 2018, a traditional NRPS module was defined as C-A-P (E) (Fig. 5) [107,163,188]. With the introduction of the so-called exchange units (XUs) module boundaries were redefined [163]. Bode and co-workers used a conserved motif in the C-A linker as recombination site for their engineering efforts, thereby defining A-P-C(E) as a XU (Fig. 5) [163]. A shift by one domain to redefine module boundaries was also postulated for modular PKSs [189-192]. Using XUs from Gram-positive and Gram-negative bacteria, functional NRPSs were generated that produced both known and artificial NRPs with only moderately reduced titers [163]. Since the XU concept was proposed to be limited by the downstream C domain specificity, the Bode lab subsequently refined the engineering strategy and developed the exchange unit condensation domain (XUC) concept [163,193]. Here, the fusion point is located in a flexible loop within the C domain. Thus, the C domain is divided into its N- and C-terminal subdomain, leading to the module architecture C_{sub}A-P-C_{sub} for XUCs (Fig. 5). The advantage of this division is that the C domain maintains the original interface with the adjacent A domain [193]. Disadvantages are that two XUCs, one containing a C and the other containing a CE domain, cannot be combined. Furthermore, XUCs from different genera are not compatible [193]. Both limitations were attributed to structural differences of the C (E) domains, that prevent functional fusions of the N- and C-terminal subdomains [193]. However, the XU and the XUC concept can be combined to overcome their respective limitations. Functional NRPSs were generated using the different recombination sites of each concept [163,193]. Remarkably, both engineering strategies as well as multiple others strictly respect C domain specificities [163,193-195]. However, recent studies showed that C domains do not harbor binding pocket-like structures. Therefore, it was speculated that C domains do not exhibit a specificity [196-198]. Furthermore, based on a series of in vitro and in vivo experiments, the C-A interface was identified as a specificity-conferring structure with an “extended gatekeeping” function, which influences the A domain specificity [198]. Evolutionary analysis also revealed that NRPS diversity is based on recombination events restricted to the A_{core} domain independent of the C domain [196,199,200]. Consequently, A_{sub} domain exchanges were reevaluated for NRPS engineering. Using this A_{sub} domain exchange strategy, multiple NRPSs were successfully engineered, yet no universal rules have been established to date [196]. Prior to the new focus on A domain engineering, comparable results had been obtained in previous studies using A_{sub} domain swaps (Fig. 5) [35,201,202]. Taken together, these results indicate that domain interfaces, in particular of the C-A didomain, are more crucial for a functional module than previously expected. For a better understanding, more structural data is required to compare interfaces of NRPSs with low homology and to identify crucial structural motifs.

Cloning strategies for the engineering of NRPSs are complicated which can be attributed to the size of the corresponding genes and the homology of domains in every module. With the definition of engineering concepts, the investigation of high-throughput cloning and recombination strategies was the logical next step to increase NRP chemical space. Naturally, the interaction between multiple NRPSs is mediated via N- and C-terminal docking or communication-mediating domains [203]. These domains can also be introduced into artificially split NRPSs without losing catalytic activity [203,204]. Recently, a study combined the XU concept to artificially split native single protein NRPSs with the introduction of synthetic zippers instead of docking domains to simplify the NRPS engineering workflow [205]. The cloning of the artificially split NRPS subunits on different plasmids allows the efficient recombination of different subunits to generate large peptide libraries in vivo [205,206]. In a comparable approach, zinc finger proteins, which bind specifically to 9 base pair DNA motifs, were used to modify the gramicidin S NRPS. The NRPS was split into stand-alone modules and different zinc fingers were added to each module. In combination with a DNA scaffold, containing the binding sites of the zinc fingers, a functional DNA-templated assembly line was obtained [207].

Insights from these combined efforts have increased our understanding of how NRPSs can be engineered. Despite these advancements,
truly universal recombination concepts have yet to be established that allow the recombination of any conceivable module architecture beyond phylum boundaries. However, the combination of structural insights, especially into domain interfaces, evolutionary analysis and time-saving high-throughput cloning methods could pave the way towards more efficient and universal engineering strategies.

5.2. RiPPs

Leader peptide, core peptide, PTM enzymes and proteases are the four components of RiPP biosynthesis. The focus of RiPP engineering efforts lies on the modification of the core peptide sequence and the recombination of PTM enzymes through leader peptide engineering.

Instead of exchanging modules or domains as it is the case for NRPSs, the primary sequence of RiPPs can be altered by changing the gene sequence associated with the core peptide [8, 208]. The challenge here is that the modifying enzymes must still be able to modify the engineered core peptide. Many studies targeted the core peptide of various RiPP families to generate artificial libraries using site-directed and random mutagenesis (Fig. 6A) [209–212]. ProcM, a promiscuous lanthipeptide synthetase, has been used for the in vivo production of a precursor gene-encoded library of $10^6$ lanthipeptides. Screening the obtained library resulted in the identification of a HIV p6 protein-human TSG101 protein interaction inhibitor [213]. This approach is, however, limited as not all biosynthetic enzymes exhibit natural substrate promiscuity. Therefore, studies focused on the engineering of biosynthetic enzymes to expand their substrate promiscuity. Enzyme libraries were generated using random mutagenesis and screened for tailoring enzymes with high substrate tolerance. For example, a dehydratase mutant library of NisB with $10^5$ variants, was generated via error-prone PCR. Subsequent high-throughput screening based on cell surface display of the peptide products, revealed a NisB variant that showed substrate flexibility against non-natural substrates [214]. Moreover, covalent fusion of PTM enzymes with their cognate leader peptides has been shown to achieve substrate promiscuity and tolerance towards non-natural peptides. Microviridin variants were successfully generated using leader peptides linked with the two ATP-grasp ligases MvdD and MvdC that were subsequently used to modify core peptide libraries [215]. Although engineering of the core peptides as well as modifying enzymes has been shown to increase structural diversity and improve bioactivity of RiPPs, the structural variety of the produced peptides is still limited.

As an alternative to core peptide and PTM enzyme engineering, combinatorial approaches for RiPP biosynthesis have been developed. Here, PTM enzymes from different pathways are used to modify a single core peptide (Fig. 6B). Using this approach, new-to-nature RiPPs were generated, combining characteristic features of two or more RiPP pathways in one product. Examples include thiazoline-containing, prenylated and macrocyclized derivatives of cyanobactins that have been produced in vitro using a set of PTM enzymes from different cyanobactin pathways [216]. To further expand the toolset of modifying enzymes that can modify a core peptide beyond RiPP family boundaries, leader peptide engineering moved into the focus of subsequent studies [1, 8]. As RiPP systems follow a leader peptide-guided biosynthetic logic, the exchange of leader peptides enables the modification of the core peptide

Fig. 6. Overview of RiPP engineering strategies. (A) Core peptide engineering by site-directed/random mutagenesis to create core peptide libraries. (B) Combinatorial biosynthesis to create new-to-nature RiPP products. (1) In vitro production of hybrid RiPPs using enzymes from different RiPP pathways. (2) Leader peptide engineering. The combination of two recognition sequences (RS) of PTM enzymes from different RiPP families allows in vivo production of hybrid RiPP products. (3) Sortase A (StrA)-mediated leader peptide exchange (LPX) strategy. In vitro production of RiPP hybrids by swapping leader peptides to allow core peptide modifications by different tailoring enzymes from different RiPP families.
Based on the observation that many PTM enzymes recognize their serine and/or threonine residues as well as (methyl)lanthionine bridges developed. In this approach, the leader peptide is enzymatically swapped after each round of modification using a sortase (StrA) [220]. StrA is a transpeptidase that catalyzes the cleavage of the peptide bond between threonine and glycine in the LPXTG motif. Subsequently, the C-terminal threonine residue is linked to the amino group of glycine of the new leader peptide (Fig. 6) [221]. The so-called StrA-based LPX technique has been used to generate novel RiPP products using PTM enzymes from cyanobactin and microviridin pathways. To do so, the leader peptide that is recognized by the cyanobactin heterocyclase (LynD) was fused with the StrA recognition motif and the microviridin J core peptide (MdnA). Following heterocyclization by LynD, the leader peptide was exchanged with the leader peptide that is recognized by the ATP-grasp ligase MdnC. Modifications by both enzymes resulted in a RiPP that harbored two thiazoline and two ω-ester cross-links [220].

The realization that both NRPSs and RiPP pathways produce complex peptide NPs that can harbor largely overlapping AA modifications resulted in the hypothesis that NRPS-derived peptides can be biosynthesized using synthetic RiPP pathways and vice versa. Particularly, the production of NRPs using RiPP pathways has attracted a lot of interest from the scientific community as RiPP primary sequences can be altered easily and hence large libraries of structural analogs generated for subsequent structure-activity relationship studies or to improve physico-chemical properties of peptides. The first steps towards this goal were recently taken when a lanthipeptide BGC was designed to produce RiPPs that mimic the structure of the NRPS-derived antimicrobial peptide brevicidine. Brevicidine is a cyclic depsipeptide containing N-acetylated, positively charged non-canonical AA residues in the peptide backbone and a lactone ring at the C-terminus. To mimic brevicidine using the tailor-made RiPP pathway, 6-AAs were replaced with 7-AAs, positively charged ornithins replaced with lysine and a methyl-lanthionine ring was installed to mimic the lactone ring of brevicidine. For the formation of the methylanthionine bridge, the C-terminal AA was replaced with cysteine to facilitate ring formation. The designed core peptides were fused with nisin leader peptides and co-expressed with the genes involved in peptide modification and cleavage. Even though the exact brevicidine was not produced using this strategy, the structurally similar lanthipeptides showed bioactivity against Gram-negative bacteria [222].

6. Conclusion

NRPSs are large multi-enzyme complexes that biosynthesize NRPs in a modular assembly line-like fashion. The chemical diversity encountered in NRPs is based on the incorporation of a large variety of (unusual) amino and carboxylic acids, the presence of module-encoded facultative enzymatic domains, on and off-line tailoring reactions catalyzed by trans-acting enzymes and the fusion with other biosynthetic pathways. Moreover, novel scaffolds have evolved through recombination of serial module arrangements from different biosynthetic pathways. RiPP BGCs, on the other hand, are small and encode monofunctional enzymes. Since RiPP precursor peptides are restricted to the set of 20 proteinogenic AAs, the conversion into heavily modified
non-proteinogenic AAs is achieved through an ever-increasing toolbox of RPP modifying enzymes. Despite these fundamentally different biosynthetic strategies the arsenal of AA modifications between both biosynthetic routes is overlapping. The biochemical strategies that result in the same set of modifications, however, can vary extensively. As a result of the different biosynthetic strategies, genome-based identification and engineering strategies vary.

Deciphering of universal NRPS biosynthetic principles resulted in the development of several generations of highly sophisticated genome mining tools that can be applied to all families of linear, non-iterative NRPS systems. State-of-the-art NRPS genome mining tools predominantly rely on hard-coded biosynthetic rules. The universal biosynthetic principles are applied to annotate the canonical set of module-encoded domains. Structures of the associated peptides can be predicted based on the A domain’s substrate specificity, and the colinearity principle. Since some A domains are promiscuous and can incorporate multiple AAs, the accurate peptide sequence can, in many cases, not be predicted with high confidence. As a consequence, NRPSs usually produce small libraries of related peptides rather than a single NRP. The presence of module-encoded facultative enzymatic domains, on the other hand, allows the regiospecific prediction of modification reactions. Even though universal biosynthetic principles have also been established for RPP biosynthesis, these universal biosynthetic rules cannot be used to chart the full RPP biosynthetic space. RPP precursor genes and genes encoding characteristic modifying enzymes for each RPP family are conserved within a RPP family. Since this conservation is usually restricted to a single family, tools have been developed for the annotation of individual RPP families. The majority of these tools relies on the hard-coded, homology-based detection BGCs of one RPP family. Machine learning-based approaches, on the other side, have shown a lot of promise for the identification of putative RPP BGCs beyond family boundaries. In the case of characterized RPP families, the core peptide sequence can be predicted with high confidence, the regioslectivity of the modifying enzymes and the number of modifications introduced by a modifying enzyme, however, is currently not predictable.

The fundamentally different biosynthetic principles between both bio-orthogonal routes are also reflected in their respective engineering strategies. NRPS engineering focuses on the manipulation of specificity-conferring components of the assembly line to alter the primary sequence of the NRP. The variability of these concepts ranges from minimal invasive methods to change the substrate specificity of single A domains to the exchange and recombination of module series. Especially the latter approach can be transformed into the high-throughput engineering of large numbers of NRPS BGCs. Changing the core peptide sequence in RPPs can be achieved by simple alternations of three bases in the precursor gene sequence. While these simple changes can be used to generate large peptide libraries with varying core peptide sequences, modifying enzymes might no longer be able to cope with the changed peptide sequence. As a result, the leader peptide has been the focus of many engineering approaches. Leader peptide swapping, the generation of chimeric leader peptides or the fusion of leader peptides are just a few examples of how modifying enzymes from other RPP families were successfully recruited to modify non-natural core peptide sequences. The ever-increasing number of characterized RPP families that coincides with the constant expansion of the toolset of RPP modifying enzymes, is paving the way towards the design of chimeric RPP BGCs for the production of tailor-made peptides, including those that were initially reported to be of NRPS origin.

Declaration of competing interest

The authors declare no competing interest.

Acknowledgements

The authors gratefully acknowledge funding by the LOEWE Center for Translational Biodiversity Genomics (LOEWE-TBG). All figures except Figs. 1 and 3 were created in BioRender.

References

[1] Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, et al. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. Nat Prod Rep 2013;30:108–66. https://doi.org/10.1039/C2NP20855F.
[2] Süssmuth RD, Mainz A. Nonribosomal peptide synthesis-principles and prospects. Angew Chem Int Ed Engl 2017;56:3770–821. https://doi.org/10.1002/anie.201609079.
[3] Zasloff M. Antimicrobial peptides of multicellular organisms. Nature 2002;415:389–95. https://doi.org/10.1038/415389a.
[4] Fjell CD, Hiso JA, Hancock REW, Schneider G. Designing antimicrobial peptides: form follows function. Nat Rev Drug Discov 2011;11:37–51. https://doi.org/10.1038/nrd3591.
[5] Dang T, Süssmuth RD. Bioactive peptide natural products as lead structures for medicinal use. Acc Chem Res 2017;50:1566–76. https://doi.org/10.1021/acs.accounts.7b00159.
[6] Puhringer S, Melittizky M, Schwarzenbacher R. The pyrroloquinoline quinone biosynthesis pathway revisited: a structure approach. BMC Biochem 2008;9:8. https://doi.org/10.1186/1471-2091-9-8.
[7] Tokita SE, Adler JM, McTaminey PM, Watson Jr JA. Efficient use and recycling of the micronutrient iodide in mammals. Biochimie 2010;92:1227–35. https://doi.org/10.1016/j.biochi.2010.02.013.
[8] Montalban-Lopez M, Scott TA, Ramesh S, Rahman IR, Van Heel AJ, Veli JL, et al. New developments in RPP discovery, enzymology and engineering. Nat Prod Rep 2021;38:1230–239. https://doi.org/10.1039/d0np00276c.
[9] Dell M, Dunbar KL, Hertweck C. Ribosome-independent peptide biosynthesis: the challenge of a unifying nomenclature. Nat Prod Rep 2021. https://doi.org/10.1039/d1np00194c.
[10] Walsh CF, O’Brien RV, Khoda C. Nonproteinogenic amino acid building blocks for nonribosomal peptide and hybrid polypeptide scaffolds. Angew Chem Int Ed Engl 2013;52:7908–124. https://doi.org/10.1002/anie.201208344.
[11] Baltz RH. Natural product drug discovery in the genomic era: realities, conjectures, misconceptions, and opportunities. J Ind Microbiol Biotechnol 2019;46:281–99. https://doi.org/10.1007/s10295-018-2115-4.
[12] Felnagle EA, Jackson EE, Chan YA, Podevels AM, Berti AD, McMahon MD, et al. Nonribosomal peptide synthetases involved in the production of medically relevant natural products. Mol Pharm 2008;5:191–211. https://doi.org/10.1021/mp700137g.
[13] Hsu L, Okels A, Zhao M, van der Donk WA. Insights into the biosynthesis of duramycin. Appl Environ Microbiol 2017;83. https://doi.org/10.1128/AEM.02968-16.
[14] Shin JH, Gwak JW, Kamaraean P, Fenko JC, Rickard AH, Kapila YL. Biomedical applications of nisin. J Appl Microbiol 2016;120:1449–65. https://doi.org/10.1111/jam.13033.
[15] Dubos RJ, Cattaneo C. Studies on a bactericidal agent extracted from a soil Bacillus: III. preparation and activity of a protein-free fraction. J Exp Med 1939;70:249–56. https://doi.org/10.1084/jem.70.3.249.
[16] Honckin RD, Dubos RJ. Fractionation of the bactericidal agent from cultures of a soil Bacillus. J Biol Chem 1940;132:791–2.
[17] Krase M, Marahiel MA. Organization of the biosynthesis genes for the peptide antibiotic granamicidin S. J Bacteriol 1988;170:4669–74. https://doi.org/10.1128/JB.170.7.4669-4674.1988.
[18] Marahiel MA, Stachelhaus T, Mootz HD. Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem Rev 1997;97:2651–74. https://doi.org/10.1021/cr960029c.
[19] Stachelhaus T, Mootz HD, Marahiel MA. The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. Chem Biol 1999;6:493–505. https://doi.org/10.1016/S1074-5521(99)80082-9.
[20] Chaliss IL, Ravel J, Townsend CA. Predictive, structure-based model of amino acid recognition by nonribosomal peptide synthetase adenylation domains. Chem Engl 2013;52:769–80. https://doi.org/10.1002/chem.201203499.
[21] Schnell N, Entian KD, Schneider U, Gmünder F, Zott F, Zährer H, Kellner R, et al. Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide rings. Nature 1988;333:276–8. https://doi.org/10.1038/333276a0.
[22] Freeman MF, Gurgui C, Helf MJ, Morinaka BI, Uria AR, Oldham NJ, et al. Metagene mining reveals polytheonamides as posttranslationally modified ribosomal peptides. Science 2012;338:387–90. https://doi.org/10.1126/science.1226121.
[23] Freeman MF, Helf MJ, Bhusan A, Morinaka BI, Pier J. Seven enzymes create extraordinary molecular complexity in an un cultivated bacterium. Nat Chem 2017;9:387–95. https://doi.org/10.1038/nchem.2666.

CRediT authorship contribution statement

Sebastian L. Wenski: Conceptualization, Visualization, Writing – original draft, Writing – review & editing. Srinithra Thienmsg: Conceptualization, Visualization, Writing – original draft, Writing – review & editing. Eric J.N. Helfrich: Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition, Project administration.
S.L. Wenski et al.  

Synthetic and Systems Biotechnology 7 (2022) 631–647

24. Wyche TP, Ruzicci AG, Schwab L, Carriere CR, Clardy J, Tryptostrubin A. A polycyclic peptide from a fungus-derived streptomyces. J Am Chem Soc 2017; 139:12899–904. https://doi.org/10.1021/jacs.7b05376.

25. Reisberg SH, Gao Y, Walker AS, Helfrich EJN, Clardy J, Baran PS. Total synthesis of the anticancer antibiotic, a daptomycin-type lipopeptide. ACS Chem Biol 2007;2:187–90. https://doi.org/10.1021/cb070002y.

26. Medeira HS, Dusser O, Marahiel MA. yeast models for predicting adenylation domains in nonribosomal peptide synthetases (NRPS). J Am Chem Soc 2007;129:13590–93. https://doi.org/10.1021/ja0764704.

27. Zhang C, Kong L, Liu Q, Lei X, Zhu T, Yin J, et al. In vitro characterization of a comprehensive resource for deciphering chemical structures of RiPPs based on prediction of natural product chemical structures from microbial genomes. Nucleic Acids Res 2017;45:W36–40. https://doi.org/10.1093/nar/gkw437.

28. Blin K, Kopp F, Kohlbacher O, Wohlleben W, Huson DH. Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res 2005;33:5799–808. https://doi.org/10.1093/nar/gki481.

29. Wolf T, Tvetegost MR, Walsh CT, Marahiel MA, et al. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. Nucleic Acids Res 2017;45:W36–41. https://doi.org/10.1093/nar/gkx319.

30. Chevreton MG, Alcher F, Kohlbacher O, Currie CR, Medema MH. SANDPUMA: ensemble predictions of nonribosomal peptide chemistry reveal biosynthetic diversity across Actinomycetes. Nucleic Acids Res 2015;33:2002–10. https://doi.org/10.1093/bioinformatics/btv400.

31. Blin K, Shaw S, Klosterman AM, Charlop-Powers Z, van Welzen GP, Medema MH, et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 2021;49:e29–35. https://doi.org/10.1093/nar/gkab355.

32. Funk MA, van der Donk WA. Ribosomal natural products, tailored to fix. Acc Chem Res 2017;50:1577–86. https://doi.org/10.1021/acs.accounts.7b00175.

33. Liu M, Liu T, Wright SA, Raney K, Chvapil M, Walsh CT. Structural and functional analysis of pantocin A: an antibiotic from Pantoea agglomerans discovered by heterologous expression of cloned genes. Angew Chem Int Ed Engl 2003;42:2898–901. https://doi.org/10.1015/00345105.

34. Bignell SV, Biemann KA, Bassett SJ, Redinho MR, Bowers AA. Post-translational Clnais Condensation and Decarboxylation en Route to the Bicyclic Core of Pantocin A. J Am Chem Soc 2016;138:5847–90. https://doi.org/10.1021/jacs.6b06529.

35. Bha M, Soll D. Aminoacyl-tRNA synthetase. Ann Rev Biochem 2006;75:619–70. https://doi.org/10.1146/annurev.biochem.69.1.617.

36. Sieber SA, Marahiel MA. Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem Rev 2005;101:75–138. https://doi.org/10.1021/cr0480919.

37. Rubin GM, Ding Y. Recent advances in the biosynthesis of RiPPs from multicore-containing precursor peptides. J Biol Chem 2017;332:6579–794. https://doi.org/10.1073/pnas.0504411102.

38. Martinis J, Vasoncellos V. Cyanobacterial gene clusters. current genetic and chemical state of knowledge. Mar Drugs 2015;13:6910–46. https://doi.org/10.3390/md13116910.

39. Zhang H, He B, Li Y, Li X. Challenges and advances in genome mining of ribosomally synthesized and post-translationally modified peptides. Nat Rev Drug Discov 2020;19:198–203. https://doi.org/10.1038/s41573-019-00457-9.

40. Russell AH, Truman AW. Genome mining strategies for ribosomally synthesized and post-translationally modified peptides. Nat Rev Drug Discov 2019;18:638–51. https://doi.org/10.1038/s41573-019-00458-8.

41. Klosterman AM, Stollenwerk KE, van Welzen GP, Medema MH, Mitchell DA, RRfinder: a genome-mining tool for class-independent RiPP discovery. mSystems 2020;5:2002–10. https://doi.org/10.1128/mSystems.00267-20.

42. Blin K, Medema MH, Kazempour D, Fischbach MA, Breitling R, Takano E, et al. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res 2013;41:W204–12. https://doi.org/10.1093/nar/gkq449.

43. Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucolesi R, et al. antiSMASH 3.0—a comprehensive resource for the genome mining of gene clusters. Nucleic Acids Res 2015;43:W237–43. https://doi.org/10.1093/nar/gkv437.

44. Blin K, Shaw S, Steinke S, Villebro R, Ziemert N, Lee SY, et al. antiSMASH 4.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res 2019;47:W817–7. https://doi.org/10.1093/nar/gkz310.

45. Skinnider MA, Merwin NJ, Johnston CW, Magarey NA. Prim 3: expanded prediction of natural product chemical structures from microbial genomes. Nucleic Acids Res 2017;45:W49–54. https://doi.org/10.1093/nar/gkw320.

46. Biermann F, Helfrich EJN. Hidden treasures: microbial natural product biosynthesis off the beaten path. mSystems 2021;6:e00462-21. https://doi.org/10.1128/mSystems.00462-21.

47. Agrawal P, Khater S, Gupta M, Sain N, Mohanty D. RiPPMiner: a bioinformatics resource for deciphering chemical structures of RiPPs based on prediction of cleavage and cross-links. Nucleic Acids Res 2017;45:W808–10. https://doi.org/10.1093/nar/gkx408.

48. de los Santos ELC. NeuLiPP: neural network identification of RIPP precursor peptides. Sci Rep 2019;9:13406. https://doi.org/10.1371/journal.pone.0056772.

49. Russell AH, Truman AW. Genome mining strategies for ribosomally synthesized and post-translationally modified peptides. Nat Rev Drug Discov 2019;18:638–51. https://doi.org/10.1038/s41573-019-00458-8.

50. Martinis J, Vasoncellos V. Cyanobacterial gene clusters. current genetic and chemical state of knowledge. Mar Drugs 2015;13:6910–46. https://doi.org/10.3390/md13116910.

51. Huisman P, Maritz JM, Klempir O, Dzamba D, Woelk CH, Hazuda DJ, et al. Bacterial bioactive peptide library design. Nat Chem 2019;11:88–96. https://doi.org/10.1038/s41557-018-00412-3.

52. Blin K, Medema MH, Kazempour D, Fischbach MA, Bregtling K, Takano E, et al. antiSMASH 2.0—a versatile platform for genome mining of secondary metabolite producers. Nucleic Acids Res 2013;41:W204–12. https://doi.org/10.1093/nar/gkq449.

53. Raiszadeh P, Khan M, Doherty P, Amin S, Wu Y, Chiu C, et al. RiPPMiner: a bioinformatics resource for deciphering chemical structures of RiPPs based on prediction of cleavage and cross-links. Nucleic Acids Res 2017;45:W808–10. https://doi.org/10.1093/nar/gkx408.
S.L. Wenski et al.  

"Synthetic and Systems Biotechnology 7 (2022) 631–647"

**References**

[76] Li J, Jensen SE. Nonribosomal biosynthesis of furacinicidin by Paenibacillus polymyxa PKB1 involves direct activation of a D-amino acid. Chem Bio 2008;15:118-27. https://doi.org/10.1016/j.chembiol.2008.01.014.

[77] Hoffmann K, Schneider-Scherzer E, Kleinka H, Zocher R. Purification and characterization of eukaryotic alanine racemase acting as key enzyme in cyclopiazonic acid biosynthesis. Biocatal Biotransform 1999;17:1270-4. https://doi.org/10.1562/0270-124X(1999)017[1270:PARCET]2.0.CO;2.

[78] Samel SA, Choquet A, Parent A, Charrier G, Balty C, Lefranc B, et al. The epimerization domain of tyrocidine synthetase Acta Crystallogr D Biol Crystallogr 2014;70:1442-52. https://doi.org/10.1107/S1399004714043906.

[79] Clugston SL, Sieber SA, Marahiel MA, Walsh CT. Chirality of peptide bond-forming condensation domains in nonribosomal peptide synthetases: the CS domain of tyrocidine synthetase is a DC(D)L catalytic. Biochemistry 2003;42:12095-104. https://doi.org/10.1021/bi0350090.

[80] Parent A, Guillet A, Parent A, Charrier G, Balty C, Lefranc B, et al. The B12-radical SAM enzyme PoyC catalyses valine ω-methylation during polytheonamide biosynthesis. J Am Chem Soc 2016;138:15515-8. https://doi.org/10.1021/jacs.6b06697.

[81] Parent A, Ameda A, Guillet A, Kubiak X, Balty C, Lefranc B, et al. Mechanistic investigations of PoyD, a radical-S-Adenosyl-l-methionine enzyme catalyzing iterative and directional epimerizations in polytheonamide A biosynthesis. J Am Chem Soc 2018;140:24697-77. https://doi.org/10.1021/jacs.8b07767.

[82] Tran MW, Kohli RM, Moor RD, Marahiel MA, Walsh CT. Peptide cyclisation catalysed by the thioesterase domain of tyrocidine synthetase. Nature 2000;407:215-9. https://doi.org/10.1038/35025116.

[83] Patellamide A and C biosynthesis by a microcin-like pathway in Prochloron dichotomus. Proc Natl Acad Sci U S A 2005;102:7315-20. https://doi.org/10.1073/pnas.0505980102.

[84] Antimicrobial biosynthesis of novel β-lactam antibiotics related to daptomycin. J Am Chem Soc 2012;134:2008-15. https://doi.org/10.1021/ja2096236.

[85] Bose S, Shokhi S, Kato S, Kaneda M, Kato S, Watanabe H, et al. Synthetic and mode of action studies on the antibiotic nikkomycins. J Med Chem 2021;64:8256-70. https://doi.org/10.1021/acsmedchemeng.0c00863.

[86] Williams DH, Kalman J. Structural and mode of action studies on the antibiotic nikkomycins. J Med Chem 2021;64:8256-70. https://doi.org/10.1021/acsmedchemeng.0c00863.
Baltz RH. Combinatorial biosynthesis of cyclic lipopeptide antibiotics: a model for
Calcott MJ, Ackerley DF. Genetic manipulation of non-ribosomal peptide
Calcott MJ, Owen JG, Ackerley DF. Efficient rational modification of non-
Bozhüyük KAJ, Linck A, Tietze A, Kranz J, Wesche F, Nowak S, et al. Modification
Beck C, Garz Villiers B, Hollfelder F. Directed evolution of a gatekeeper domain in
Kries H, Nakeeb M, Styles M, Struck A-W, et al. Evolution of combinatorial diversity in trans-acyltransferase polyketide assembly lines redefined with a central acyl carrier protein. Proteins 2018;86:2017-5068.

Mimicry of a non-ribosomally produced antimicrobial, gain-of-function in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions. J Am Chem Soc 2011;133:5016-23.

Ruijte F, Kuipers OP. Combinatorial biosynthesis for the generation of new-to-nature peptide antimicrobials. Biochem Soc Trans 2021;49:203-15. https://doi.org/10.1042/bst20200077.

Wu L, Hansen JN. Enhancement of the chemical and antimicrobial properties of subulin by site-directed mutagenesis. J Biol Chem 1992;267:25078-85. https://doi.org/10.1074/jbc.267.42.25078.

Otter PD, Doegan LH, Lawson EM, Deaper LA, O’Connor PM, Hill C, et al. Complete alanine scanning of the two-component lantibiotic lactacin 3147: generating a blueprint for rational drug design. Mol Microbiol 2006;52:735-47. https://doi.org/10.1111/j.1365-2958.2006.05358.x.

Li B, Sher D, Kelly L, Shi Y, Huang K, Knerr PJ, et al. Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. Nat Proc Natl Acad Sci U S A 2010;107:10430-5. https://doi.org/10.1073/pnas.1003236107.

Pan SJ, Link AJ. Sequence diversity in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions. J Am Chem Soc 2011;133:5016-23.

Yang X, Lennard KR, He C, Walker MC, Ball AT, Doigneaux C, et al. A lanthipeptide library used to identify a protein-protein interaction inhibitor. ACS Synth Biol 2020;9:1468-76. https://doi.org/10.1021/acssynbio.0c00327.

Kries H, Wachtel R, Pabst A, Wanner B, Niquille D, Hilvert D. Reprogramming nonribosomal peptide synthetases for ‘clickable’ clinical isolates*. Angew Chem Int Ed Engl 2019;58:9090-3. https://doi.org/10.1002/anie.201811170.

Pan SJ, Link AJ, Steer DL, Kaczmarski JA, Gavriilidou A, Chow KH, Candace Ho YT, Kaczmarski JA, Gavriilidou A, Chow KH, Steer DL, et al. A combination of polyunsaturated fatty acid, nonribosomal peptide and NRPS assembly lines. Biotechnol Bioproc Eng 2020;25:886-900. https://doi.org/10.1002/ebc.26047.

Booth TJ, Bouzhiük AJ, Liston JD, Lacey E, Wilkinson B. Bifurcation drives the evolution of assembly-line biosynthesis. Nat Biotechnol 2020;12:26-30. https://doi.org/10.1038/s41589-018-0008-5.

Booth TJ, Bouzhiük AJ, Liston JD, Lacey E, Wilkinson B. Bifurcation drives the evolution of assembly-line biosynthesis. Nat Biotechnol 2020;12:26-30. https://doi.org/10.1038/s41589-018-0008-5.

Kranz J, Wenski SL, Ditche JF, Bode HB. Influence of condensation domains on activity and specificity of adenylation domains. bioRxiv 2021.

Baunach M, Chowdhury S, Stallforth P, Dittmann E. The landscape of recombination events that create nonribosomal peptide diversity. Mol Biol Evol 2021;38:2116-30. https://doi.org/10.1093/molbev/msab015.

Cristmann M, Kohlhaas C, Pie J. Evolution-guided engineering of nonribosomal peptide synthetase adenylation domains. Chem Sci 2013;4:1041-5. https://doi.org/10.1039/C2SC2172H.

Kries H, Niqueille DL, Goetzert A. A subreddit swap strategy for reengineering nonribosomal peptides. Biochem J 2015;464:641-8. https://doi.org/10.1042/BJ20150415.

Hahn M, Stachhaltaus T. Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediated domains. Proc Natl Acad Sci U S A 2004;101:15585-90. https://doi.org/10.1073/pnas.0404932101.

Kegler C, Bode HB. Artificial splitting of a non-ribosomal peptide synthetase by inserting natural docking domains. Angew Chem Int Ed Engl 2020;59:13565-9. https://doi.org/10.1002/anie.201919589.

Bozhueyuk AJ, Watzel AJ, Abbood N, Bode HB. Synthetic zippers as an enabling tool for engineering of non-ribosomal peptide synthetases. Angew Chem Int Ed 2020;61:17531-8. https://doi.org/10.1002/anie.202002899.

Abbood N, Vo TD, Watzel AJ, Bozhueyuk AJ, Bode HB. Type S Non Ribosomal Peptide Synthetases for the rapid generation of tailor-made peptide libraries. https://doi.org/10.1021/acs.chembio.0c00299. 2021.

Huang H-M, Stephan P, Kries H. Engineering DNA-templated nonribosomal peptide synthesis. Cell Chem Biol 2021;28:221-7. https://doi.org/10.1016/j.chembiol.2020.11.004. e7.

Mimicry of a non-ribosomally produced antimicrobial, gain-of-function in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions. J Am Chem Soc 2011;133:5016-23.

Li B, Sher D, Kelly L, Shi Y, Huang K, Knerr PJ, et al. Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. Nat Proc Natl Acad Sci U S A 2010;107:10430-5. https://doi.org/10.1073/pnas.1003236107.

Pan SJ, Link AJ. Sequence diversity in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions. J Am Chem Soc 2011;133:5016-23.

Yang X, Lennard KR, He C, Walker MC, Ball AT, Doigneaux C, et al. A lanthipeptide library used to identify a protein-protein interaction inhibitor. Nat Chem Biol 2018;14:375-80. https://doi.org/10.1038/nchembio.2869.

Zhao X, Cebrián R, Fu Y, Rink R, Bosma T, Moll GN, et al. High-throughput screening for substrate specificity-adapted mutants of the nisin dehydratase NisB. ACS Synth Biol 2020;9:1468-78. https://doi.org/10.1021/acssynbio.0b00150.

Reyna-Gonzalez E, Schmid B, Petras D, Süssmuth RD, Dittmann E. Leader peptide-free in vivo reconstitution of microcinivir biosynthesis enables design of syntheses: protease-targeted libraries. Angew Chem Int Ed 2016;128:944-7. https://doi.org/10.1002/ange.201604345.

Sardar D, Lin Z, Schmidt EW. Mutability of RIP enzymes enables design of decorated peptides. Chem Biol 2015;22:907-16. https://doi.org/10.1016/j.chembiol.2015.06.014.

Majchrickowska JA, Labelski J, Moll GN, Kuipers A, Bijlsma JJE, Kuipers OP. Production of a class II two-component lantibiotic of Streptococcus pneumoniae using the class I nisin synthetic machinery and leader sequence. Antimicrob Agents Chemother 2016;60:1498-505. https://doi.org/10.1128/AAC.00883-09.

Oman TJ, van der Donk WA. Follow the leader: the use of leader peptides to guide natural product biosynthesis. Nat Chem Biol 2010;6:9-18. https://doi.org/10.1038/nchembio.201005.4.

Burkart BJ, Kakkar N, Hudson GA, van der Donk WA, Mitchell DA. Chimeric leader peptides for the generation of non-natural hybrid RIP products. ACS Cent Sci 2017;3:629-38. https://doi.org/10.1021/acscentsci.7b00411.

Franz L, Koehnke J. Leader peptide exchange to produce hybrid, new-to-nature ribosomal products. Chem Biol 2021;28:2116-30. https://doi.org/10.1016/j.chembiol.2021.01.008.

Selvaraj C, Priya RB, Lee JK, Singh SK. Mechanistic insights of Srm-A/LPTXG blockers targeting the therapeutically relevant mechanism in Streptococcus mutants. ACS Adv 2015;5:10498-510. https://doi.org/10.1021/acs.casr.5b01762.

Zhao X, Li Z, Kuipers OP. Mimicry of a non-ribosomally produced antimicrobial, breviscacin, by ribosomally synthesized and post-translational modified cell. Cell Biotechnol 2020;27:262-71. https://doi.org/10.37362/biotech.2020.07.005.e4.