Institut für Biochemie und Biologie

Engineering Of The Microviridin Post-translational Modification Enzymes For The Production Of Synthetic Protease Inhibitors

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium

im Fach Mikrobiologie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

von
M. Sc. Emmanuel Reyna González
geboren in Mexiko-Stadt

Referees

Prof. Dr. Elke Dittmann, University of Potsdam: Department of Microbiology.
Prof. Dr. Mohamed A. Marahiel, University of Marburg: Department of Chemistry.
Prof. Dr. Nediljko Budisa, Technical University of Berlin: Department of Biocatalysis.

Thesis defense held on Monday, the 22nd of January 2018.

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-406979
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-406979

Herewith I declare that I prepared the doctoral thesis “Engineering Of The Microviridin Post-translational Modification Enzymes For The Production Of Synthetic Protease Inhibitors” on my own and with no other sources and aids than quoted.

Potsdam, 22nd of January 2018.

Emmanuel Reyna González

“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

Isaac Asimov

Abstract

Natural products and their derivatives have always been a source of drug leads. In particular, bacterial compounds have played an important role in drug development, for example in the field of antibiotics. A decrease in the discovery of novel leads from natural sources and the hope of finding new leads through the generation of large libraries of drug-like compounds by combinatorial chemistry aimed at specific molecular targets drove the pharmaceutical companies away from research on natural products. However, recent technological advances in genetics, bioinformatics and analytical chemistry have revived the interest in natural products. The ribosomally synthesized and post-translationally modified peptides (RiPPs) are a group of natural products generated by the action of post-translationally modifying enzymes on precursor peptides translated from mRNA by ribosomes. The great substrate promiscuity exhibited by many of the enzymes from RiPP biosynthetic pathways have led to the generation of hundreds of novel synthetic and semisynthetic variants, including variants carrying non-canonical amino acids (ncAAs). The microviridins are a family of RiPPs characterized by their atypical tricyclic structure composed of lactone and lactam rings, and their activity as serine protease inhibitors. The generalities of their biosynthetic pathway have already been described, however, the lack of information on details such as the protease responsible for cleaving off the leader peptide from the cyclic core peptide has impeded the fast and cheap production of novel microviridin variants. In the present work, knowledge on leader peptide activation of enzymes from other RiPP families has been extrapolated to the microviridin family, making it possible to bypass the need of a leader peptide. This feature allowed for the exploitation of the microviridin biosynthetic machinery for the production of novel variants through the establishment of an efficient one-pot *in vitro* platform. The relevance of this chemoenzymatic approach has been exemplified by the synthesis of novel potent serine protease inhibitors from both rationally-designed peptide libraries and bioinformatically predicted microviridins. Additionally, new structure-activity relationships (SARs) could be inferred by screening microviridin intermediates. The significance of this technique was further demonstrated by the simple incorporation of ncAAs into the microviridin scaffold.

Zusammenfassung

Naturstoffe und ihre Derivate waren schon immer eine Quelle von Leitstrukturen. Insbesondere haben bakterielle Verbindungen eine wichtige Rolle bei der Arzneimittelentwicklung gespielt, zum Beispiel im Bereich der Antibiotika. Die Abnahme von Entdeckungen neuer Leitstrukturen aus natürlichen Quellen und die Hoffnung, neue Leitstrukturen in großen Bibliotheken medikamentenähnlicher Verbindungen zu finden, welche auf spezifische molekulare Ziele gerichtet sind und mithilfe kombinatorischer Chemie erstellt wurden, trieben die Pharmaunternehmen weg von der Naturstoffforschung. Allerdings haben moderne technologische Fortschritte in der Genetik, der Bioinformatik und der analytischen Chemie das Interesse an Naturstoffen wiederbelebt. Die ribosomal synthetisierten und posttranslational modifizierten Peptide (RiPPs) sind eine Gruppe von Naturstoffen, die durch das Einwirken posttranslational modifizierender Enzymen auf Präkursorpeptide entstehen, welche ihrerseits aus mRNA durch Translation an den Ribosomen hervorgehen. Die durch viele der Enzyme aus RiPP Biosynthesewege gezeigte große Substrat-Promiskuität führte zur Erzeugung hunderter neuartiger synthetischer und halbsynthetischer Varianten, einschließlich Varianten mit nicht-kanonischen Aminosäuren. Die Microviridine sind eine Familie von RiPPs, die durch ihre atypische trizyklische Struktur aus Lacton- und Lactamringen und ihre Aktivität als Serin-Protease-Inhibitoren gekennzeichnet sind. Die Grundlagen ihres Biosyntheseweges sind bereits beschrieben worden, aber wesentliche Fragestellungen, zum Beispiel die für die Spaltung des Leader-Peptids vom zyklischen Core-Peptid verantwortliche Protease betreffend, sind weitgehend ungeklärt und erschweren die schnelle und kostengünstige Herstellung neuer Microviridinvarianten. In der vorliegenden Arbeit wurde das Wissen über die durch Leader-Peptid Aktivierung von Enzymen aus anderen RiPP-Familien auf die Microviridinfamilie extrapoliert, wodurch es möglich wurde, die Notwendigkeit eines Leader-Peptids zu umgehen. Diese Besonderheit erlaubt nunmehr, die Microviridin-Biosynthese-Enzyme für die Herstellung von neuartigen Varianten durch die Etablierung einer effizienten in vitro Synthese-Plattform auszunutzen. Die Relevanz dieses chemoenzymatischen Ansatzes wurde durch die Synthese von neuen potenten Serin-Protease-Inhibitoren aus sowohl rational gestalteten Peptidbibliotheken als auch bioinformatisch vorhergesagten Microviridinen veranschaulicht. Darüber hinaus wurden durch das Screenen von Microviridinzwischenprodukten neue Struktur-Funktionsbeziehungen abgeleitet. Die Bedeutung dieser Technik wurde durch den einfachen Einbau von nicht-kanonischen Aminosäuren in das Microviridin-Gerüst weiter demonstriert.

Contents

List of figures

List of tables

Abbreviations

AcCoA: acetyl coenzyme A

BGC: biosynthetic gene cluster

ESI-HR-MS/MS: Electrospray ionization- high resolution- tandem mass spectrometry

ESI-Q-TOF: Electrospray ionization- quadrupole- time of flight

HCCA: α-Cyano-4-hydroxycinnamic acid

HPLC: high-performance liquid chromatography

Lan: lanthionine

LP-MvdC: Engineered MvdC with its leader peptide attached to the N-terminus

LP-MvdD: Engineered MvdD with its leader peptide attached to the N-terminus

MALDI-TOF: Matrix-assisted laser desorption ionization- time of flight

MeLan: 3-methyllanthionine

Mdn: microviridin (cyclic core peptide)

MS: mass spectrometry

MS/MS: tandem mass spectrometry

ncAA: non-canonical amino acid

NRPS: non-ribosomal peptide synthetase

PBS: phosphate-buffered saline

PKS: polyketide synthase

PRPS: post-ribosomal peptide synthesis

PTM: post-translational modification

RiPP: ribosomally synthesized and post-translationaly modified peptide

RRE: RiPP precursor peptide recognition element

RS: recognition sequence

SAR: structure-activity relationship

SPPS: solid-phase peptide synthesis

wHTH: winged helix-turn-helix

1. Introduction

1.1. Natural products as source of drugs

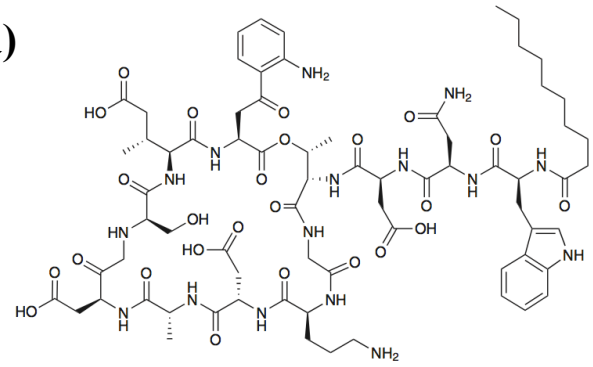
Since the beginning of civilization humans have used natural remedies, mainly plant-derived preparations, to cure diseases and ailments. Cultures such as the Mesopotamian, Egyptian, Chinese, Indian, and Greco-Roman, all documented the use of such preparations [1, 2]. It is now clear that the pharmacological effects of natural remedies observed by ancient cultures are in fact caused by natural products. A natural product can be broadly defined as a biologically active compound produced originally by a plant, animal or microorganism [3-5]. For many authors, however, the definition of natural product is restricted to secondary metabolites.

Natural products have played a big role in drug development, particularly in the area of antibiotics. The accidental discovery of penicillin by Alexander Fleming in 1928 and the isolation of streptomycin in 1943 by Selman Waksman marked the dawn of the golden era of antibiotics [6]. This was followed by the medicinal chemistry era, characterized by the development of improved synthetic versions of the natural antibiotics discovered in the golden era [7]. More than 50% of the antibiotics in clinical use now are either natural products or derived from them [8]. Beside their classical strengths as antibiotics and antitumor agents, natural products have also been marketed as immunosuppressants, antivirals, antiparasitics, and enzyme inhibitors, among others (Fig. 1.1) [9]. Newman and Cragg have published a series of reviews on the sources of drugs approved by the FDA for the treatment of human diseases. In their last publication, covering drugs from 1981 to 2014, they report that about 65% of all small-molecule approved drugs in this period are either natural products or compounds derived from or inspired by them [10]. In a similar review, an analysis of first-in-class drugs (drugs that act on a novel target or pathway) approved by the FDA between 1999 and 2013 revealed that of 113 first-in-class drugs, about 30% are natural products or derivatives thereof [11].

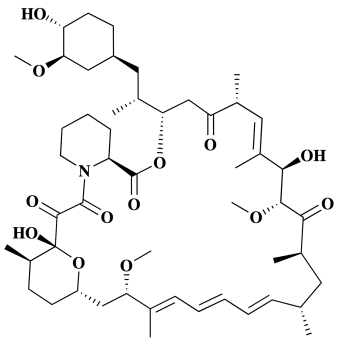
Despite the tremendous impact that natural products have had on drug development, many pharmaceutical companies turned away from them for drug discovery research in the 1990s [5, 12]. This was due to a number of factors: The introduction of high-throughput screening directed at specific molecular targets coupled with the creation of large collections of drug-like compounds by combinatorial chemistry, the difficulty to isolate natural products in suitable quantities from complex mixtures, concerns about the rediscovery of known compounds and possible exhaustion of structural novelty, and concerns about intellectual

property rights arising from the Rio Convention on Biological Diversity [2, 5, 8, 12-15]. The consequence of this was a decrease in the number of new small-molecule natural products with pharmaceutical potential discovered in the 1990s and early 2000s [10, 15].

a)

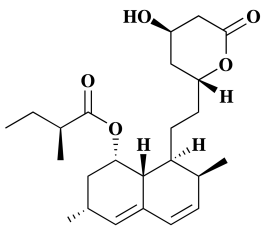


Daptomycin
Antibiotic

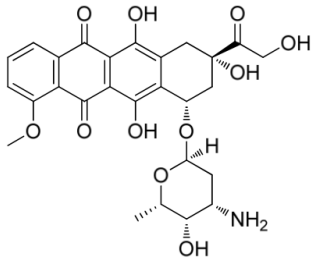


Rapamycin
Immunosuppressant

b)

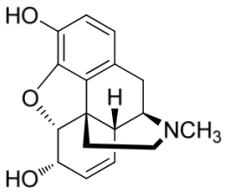


Lovastatin
Anticholesteremic

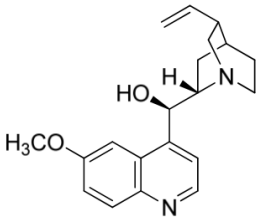


Doxorubicin
Anticancer

c)



Morphine
Analgesic



Quinine
Antimalarial

Figure 1.1. Selected natural products approved for use as therapeutics. a) Bacterial compounds: Daptomycin [16] and rapamycin [1]; **b)** Fungal compounds: Lovastatin [1] and doxorubicin [2]; **c)** Plant-derived compounds: Morphine and quinine [2].

Nevertheless, the recent technological development in the area of genetics, bioinformatics and analytical chemistry is promoting a comeback of natural products in drug discovery [1, 8, 13, 15]. In a recent study, Pye *et al*. address the concern that the structural novelty in modern discoveries of natural products is declining by analyzing the number and source of compounds discovered as well as compound novelty [17]. Some major conclusions

of this analysis are: First, that although a significant number of the new products display structural similarity to known scaffolds, some (23% of all molecules) exhibit novel structures; second, that exploring new taxonomic groups holds a high probability of yielding novel chemical entities, albeit with a limited period of expected novel compound discovery; third, that significant innovation in discovery approaches is required in order to avoid rediscovery of known compounds from continued investigations of the same classes of organisms. These conclusions can be corroborated by the information gathered on whole genomes, generated by next-generation sequencing, and new biosynthetic gene clusters (BGCs), discovered by genome mining with the aide of bioinformatic tools. For example, in the genomes of the actinomycetes there are many more BGCs than those accountable for the known metabolites produced by this order of bacteria [8, 9, 16, 18]. The genomes of microorganisms that have not yet been cultivated in the laboratory (estimated to be 95-99.9% of the microorganisms existing in nature [19]), accessed through metagenomics, and the genomes of species from underexplored environments have also revealed a trove of novel natural products [9, 13, 17, 20, 21]. Developments in the area of analytical chemistry, particularly in NMR (e.g. cryoprobes) and mass spectrometry (e.g. Orbitrap technology), have also rekindled the interest in natural products by simplifying the labour of compound characterization and dereplication [8, 13, 22, 23]. Considering all these facts, it is possible to predict that natural products will continue to be a very prolific source of drug leads.

1.2. Cyanobacteria as source of natural products

Microorganisms have always played a major role in the discovery of natural products, in particular bacteria. Of all known active compounds from microorganisms, about 40% were produced by fungi, 30% by actinomycetes, and the rest by non-filamentous bacteria (mainly *Bacillus* and *Pseudomonas*). In the case of antibiotics, over 50% come from the actinomycetes, between 10% and 15% from non-filamentous bacteria, and approximately 20% from fungi [24]. Genera such as *Streptomyces*, *Bacillus* and *Pseudomonas*, and the group of the myxobacteria have been very fruitful regarding natural products [25]. Nevertheless, the increasing resistance of pathogenic bacteria against the traditional antibiotics has created a need for products with new chemical scaffolds, leading to the search for new sources of natural products. In this regard, microorganisms from underexplored environments, such as the marine and extreme environments, have attracted special attention over the last years. One example of these microorganisms are the cyanobacteria.

Cyanobacteria, formerly known as blue-green algae, are Gram-negative bacteria capable of oxygenic photosynthesis; some strains are also capable of fixing atmospheric nitrogen [26, 27]. They are among the oldest organisms on Earth, with evidence of their existence dating back 3.5 billion years, and are considered to have played a big part in the establishment of the current oxygen-rich atmosphere on Earth [28]. These bacteria constitute a very diverse group, with different morphologies and niche habitats, both in terrestrial and aquatic environments, and even in those with extreme conditions [29]. It is this diversity which makes cyanobacteria special from the natural products' point of view, for the reason that their heterogeneity is translated into a great variety of natural compounds (Fig. 1.2) [27, 30].

Traditionally, research on natural products from cyanobacteria had focused on the toxins (cyanotoxins) they produce due to the danger they represent to humans and livestock, for example microcystins and nodularins (hepatotoxins), anatoxin-a (neurotoxin), and lyngbyatoxin (dermotoxin), among others [26, 29, 31-34]. Nonetheless, cyanobacteria are now regarded as a prolific source of novel compounds with a wide range of biological activities [35, 36]. From a drug development perspective, the most successful cyanobacterial isolates have been the cytotoxic dolastatins. A derivative of dolastatin 10 has been conjugated with an antibody to selectively treat Hodgkin's lymphoma and systemic anaplastic large cell lymphoma, being this the first approved drug substance with cyanobacterial origins [37, 38]. Another interesting group of compounds for antitumor therapeutics are the cryptophycins [28]. Regarding anti-infective activity, several cyanobacterial secondary metabolites with antibacterial or antifungal properties have been isolated (e.g. antibacterial: hapalindoles, comnostins; antifungal: hassallidins, laxaphycins), however their concomitant general cytotoxicity makes them unsuitable candidates for drug development. By contrast, the lectin scytovirin seems to be a promising lead for antiviral therapeutics, while the cyclic peptide aerucyclamide B seems to be propitious for drugs against malaria [37]. Another example of compounds with interesting bioactivity are the anti-quorum sensing molecules malyngolide, malyngamide, and lyngbyoic acid, which have the potential to be exploited as indirect antibacterials and as agents to prevent marine fouling [29, 37]. Finally, protease inhibition is another common bioactivity in many cyanobacterial metabolites such as the aeruginosins, the microginins, and the microviridins. Agents targeting protease activity are very important from the drug development point of view, since proteases play key roles in several pathological

processes, for instance elastases in pulmonary emphysema, the 20S proteasome in cancer, and diverse viral proteases, among others [28, 39].

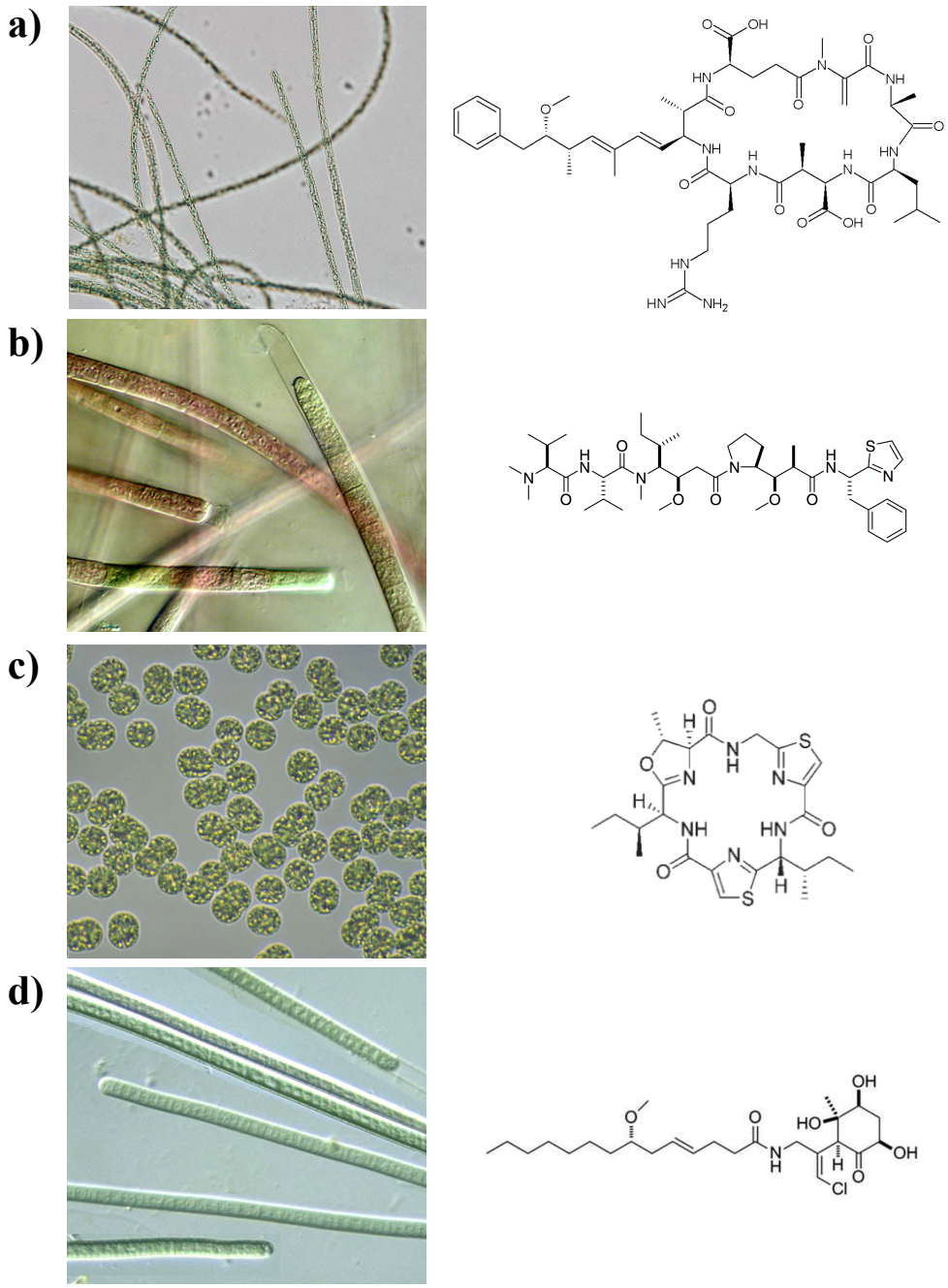


Figure 1.2. Selected bioactive cyanobacterial compounds and their producers. a) The hepatotoxin microcystin LR [40] produced by *Planktothrix agardhii* [41]; **b)** The cytotoxic dolastatin 10 [42] produced by *Symploca hydnoides* [43]; **c)** The antiplasmodial aerucyclamide B [37] produced by *Microcystis aeruginosa* [44]; **d)** The anti-quorum sensing compound malyngamide 2 [35] produced by *Moorea sp.* [45].

1.3. Mechanisms of peptide synthesis in cyanobacteria: The ribosomal pathway

As discussed in the previous section, cyanobacteria are prolific producers of natural products. A great part of these compounds are peptides or show peptidic substructures [28, 46]. Most of them are produced through the activity of nonribosomal peptide synthetases (NRPSs) or the

combined action of NRPSs with polyketide synthases (PKSs) [37, 46, 47]. Nevertheless, the post-ribosomal peptide synthesis (PRPS) also contributes to the production of peptidic metabolites with interesting bioactivities. These compounds are designated as ribosomally synthesized and post-translationally modified peptides (RiPPs).

The PRPS starts with the ribosomal synthesis of a precursor peptide and its associated post-translational modification (PTM) enzymes, whose precursor genes are normally grouped together in clusters in prokaryotes (Fig. 1.3) [48]. The precursor peptide, typically between 20-110 amino acids long, is composed of different segments, with the core region or core peptide being the one that, after PTMs and excision from the other segment(s), will ultimately become the mature natural product [49]. The set of PTMs that a core peptide may undergo is very varied, and may involve only a few amino acids or almost every one in the molecule, allowing for a great amount of structural diversity despite the original constraint of being assembled from the 20 proteinogenic amino acids [50]. Appended to the core peptide, in most of the cases at the N-terminus, is a leader peptide believed to be important for recognition by many of the PTM enzymes. In some eukaryotic peptides, a signal sequence is frequently found N-terminal to the leader peptide and is responsible for intracellular localization. A C-terminal recognition sequence important for excision and cyclization can also be observed in some precursor peptides [51].

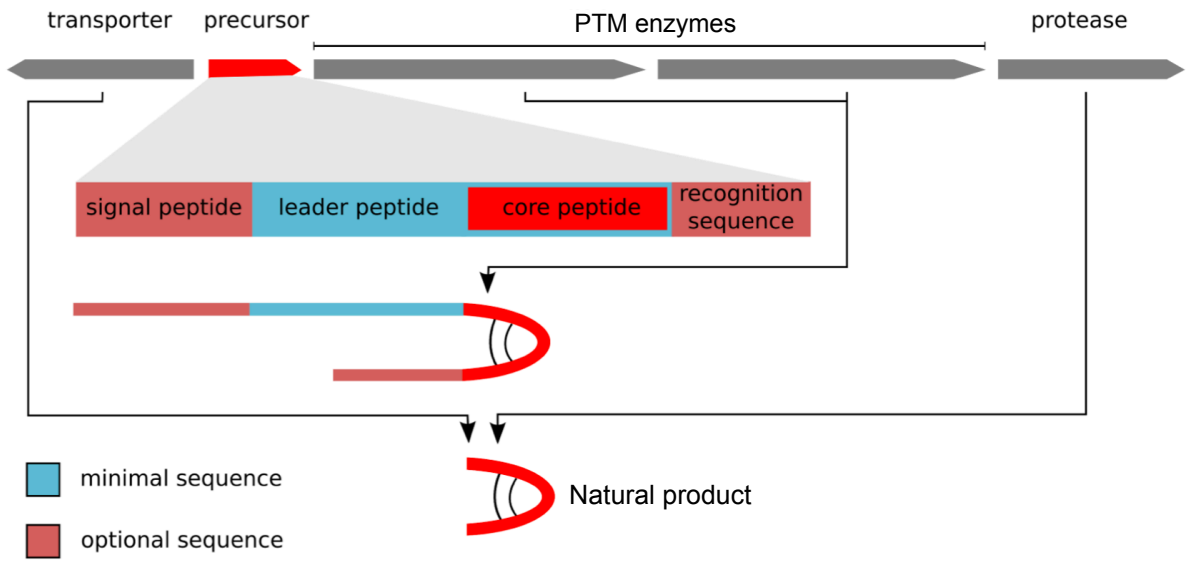


Figure 1.3. Genes in a hypothetical RiPP BGC. The precursor peptide is translated as a multisegmented polypeptide. After the associated PTM enzymes have acted on the core peptide, a protease cleaves off the leader peptide and optional additional regions, and a transporter (which sometimes is fused with the protease) exports the mature natural product out of the cell. Scheme modified from [52].

The exact role of the leader peptide in the PRPS has been extensively discussed. Different functions for the leader peptide have been proposed in the past: A function as a secretion signal, as a chaperone-like sequence that assists in the PTMs of the core peptide, as a protective sequence that maintains the precursor peptide inactive inside the producer cell, and as a recognition and activation sequence for the PTM machinery [53]. More recently, successful crystallization of the class I lanthipeptide dehydratase NisB, the cyanobactin cyclodehydratase LynD and the microviridin ATP-grasp ester ligase MdnC bound to their cognate leader peptides [54-56], and a bioinformatics study by Burkhart *et al*. on leader peptide recognition by RiPP biosynthetic enzymes [57] have gathered experimental support for the last proposed function. The crystallization studies have shown that conserved motifs within the leader peptides of the three mentioned RiPP classes are structurally ordered and interact directly with ordered residues of the mentioned PTM enzymes. The molecular basis for the interaction of the dehydratases NisB and LynD with their cognate leader peptide is similar, and involves a winged helix-turn-helix (wHTH) domain of the PTM enzyme that interacts with the conserved motif in the leader peptide, which is ordered in a β-strand conformation [54, 55, 58, 59]. In the case of the microviridin family, the conserved motif in the leader peptide adopts an α-helix conformation and interacts with an α-helix from the PTM enzyme, differing from the two RiPP classes named before [56]. These observations are corroborated by the aforementioned bioinformatics study. In this work the authors were able to identify a conserved peptide-binding domain, designated as RiPP precursor peptide Recognition Element (RRE), in PTM enzymes from the majority of the prokaryotic RiPP families, including the class I of the lanthipeptides, and the cyanobactins, but not the microviridins [57].

The ability of many PTM enzymes of the natural product ribosomal pathway to recognize a highly conserved leader peptide and accept variations in the core peptide is a feature that allows for easier evolvability compared to the non-ribosomal pathway [49]. It is indeed this ability which makes the PRPS interesting from a bioengineering point of view since the introduction of new chemical features to RiPPs may be possible by simply editing the sequence of the precursor peptides and posterior incubation with PTM enzymes from the huge repertoire already available in nature. Some of the major RiPP classes will be briefly reviewed next with the objective of drawing similarities, differences, and work already realized towards their bioengineering.

1.3.1. Lanthipeptides

The lanthipeptides, short for lanthionine-containing peptides, are the longest known family of RiPPs [49]. Moreover, genome database mining revealed that they are the largest occurring group of RiPPs in the genomes sequenced until now [48]. These natural products show varied bioactivities, including morphogenetic, antinociceptive, antiallodynic, and antibiotic activities (lantibiotics), with nisin being the first discovered lantibiotic (Fig. 1.4a) [60]. The lanthipeptide classification is restricted to those peptides harboring lanthionine (Lan) and 3-methyllanthionine (MeLan) residues in their structure as a result of a PTM process of dehydration of Ser and Thr, respectively, and posterior attack of a Cys residue onto the dehydro amino acids (Fig. 1.4b) [51]. Furthermore, the lanthipeptides are subdivided in 4 groups, depending on the type of PTM enzymes that generate the Lan/MeLan residues.

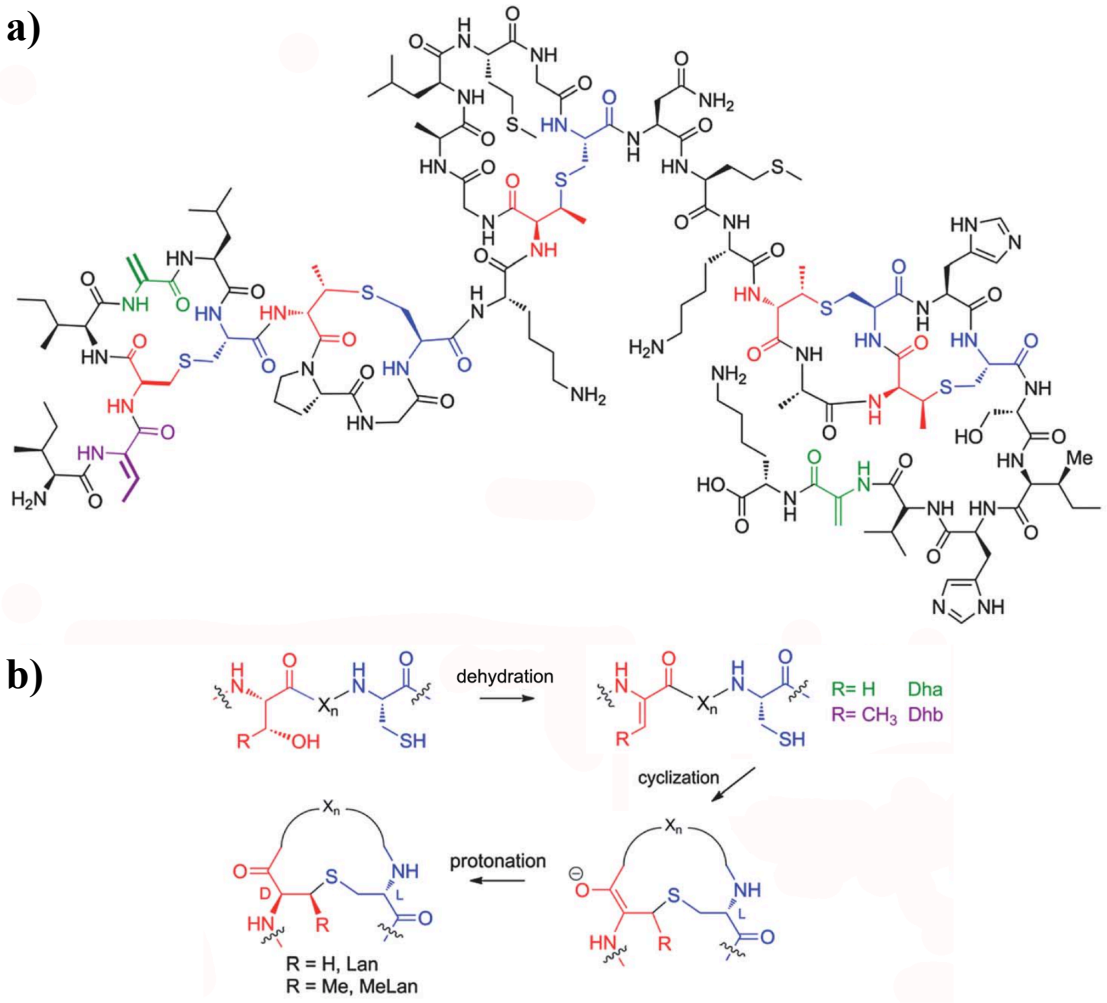


Figure 1.4. a) Structure of nisin, the first lantibiotic discovered; **b)** Steps in the formation of Lan and MeLan in lanthipeptides. Dha: Dehydroalanine; Dhb: Dehydrobutyrine. Image adapted from [51].

The Lan/MeLan residues in class I lanthipeptides are introduced through the combined activity of two independent enzymes, a dehydratase and a cyclase. In the case of class II, III and IV lanthipeptides, the Lan/MeLan motifs are installed by single proteins with different enzymatic domains (Fig. 1.5). Regarding the class II lanthionine synthetases, the N-terminal domain dehydrates the Ser and Thr residues, and the C-terminal domain catalyzes the cyclization with Cys. Concerning the classes III and IV synthetases, the dehydration step is achieved through the successive work of a central kinase domain and an N-terminal lyase domain, and the cyclization step by the action of the C-terminal domain. The cyclase domains of the synthetases of the classes II-IV show sequence homology with the cyclases from class I, however, the cyclization domains of the class III synthetases lack three zinc-binding residues conserved in the other sequences [51].

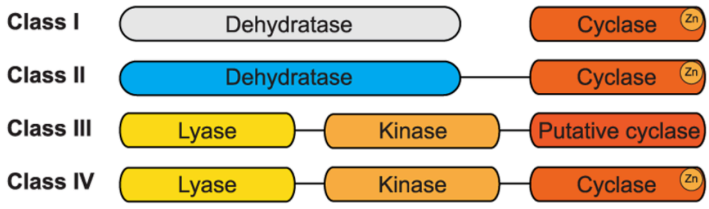


Figure 1.5. Lanthipeptide subclassification. The subgroups are defined by the type of enzymatic machinery that introduces the characteristic thioether crosslinks. The cyclase domains of the class II-IV enzymes show homology with the class I cyclases, with the exception that the class III cyclase domains do not possess the metal-binding residues. Image modified from [60].

The introduction of the characteristic Lan/MeLan residues into the core peptide depends on the presence of the leader peptide. In the case of the class I lanthipeptide nisin, a conserved four-residue motif, FNLD, localized in the N-terminal half of the leader peptide had been recognized as key for the production and secretion of nisin [61]. This was corroborated by the already mentioned co-crystallization of the NisB dehydratase with the nisin leader peptide. The N-terminal half of the leader peptide, ordered in a β-strand conformation, was shown to interact directly with a wHTH domain of the dehydratase, in particular with the three-stranded antiparallel β-sheet “wing” of the domain [55]. In contrast to this, the leader peptides of class II lanthipeptides posses a conserved C-terminal ELxxBx motif, where B is a Val, Leu or Ile [61]. Finally, in the class III leader peptides a semiconserved N-terminal ILELQ motif important for post-translational processing has been identified too [62].

The process of leader peptide removal also differs among the lanthipeptide subgroups. For most of the class I and some class II lanthipeptides, dedicated proteases catalyze the

scission of the leader peptide, in some cases intracellularly and in others after secretion out of the cell. Conversely, bifunctional enzymes carry out the excision of the leader peptide and the export of the modified core peptide of the majority of class II and some class I lanthipeptides [61]. Regarding the classes III and IV, the removal of the leader peptide seems to be done in a stepwise manner by currently unidentified amino peptidases [63].

The variety of the lanthipeptide group is in great part due to the many tailoring modifications that adorn their structures. These tailoring modifications, i.e. PTMs that normally enhance the activity or the stability of the molecule, are generally class restricted, with a higher prevalence in the class I and class II, and do not depend on the presence of the leader peptide. More than 10 tailoring modifications have been reported, including halogenation, glycosylation, and incorporation of D-amino acids, and some of the enzymes responsible for these modifications are very permissive regarding substrate selectivity, thus making them attractive for bioengineering [64, 65].

Many different bioengineering attempts have been carried out on this group of RiPPs, concentrating on those members whose biosynthetic pathway has already been detailed, for example nisin and lacticin 481. Three main approaches have been followed to generate novel lanthipeptide analogs with improved properties: *in vivo* engineering in native producers, *in vivo* engineering in heterologous hosts, and *in vitro* engineering [64]. The *in vivo* methodologies remodel the wild type precursor by site-directed mutagenesis or provide a new precursor peptide gene in a plasmid with simultaneous knockout of the native gene in the case of working with native producers, or achieve heterologous production by co-expressing the precursor peptide and its PTM enzymes in more amenable hosts, e.g. *E. coli*. This has proven particularly useful for the characterization of compounds from silent clusters. The *in vitro* approach utilizes the heterologous overexpression of PTM enzymes, and the production of semisynthetic precursor substrates via heterologous expression or solid-phase peptide synthesis (SPPS). Both methods have rendered analogs with improved bioactivities, and analogs carrying non-proteinogenic amino acids [49]. Nevertheless, the excision of the leader peptide in both ways remains problematic. Strategies involving the introduction of artificial cleavage sites have been developed, e.g. for trypsin, factor Xa protease or alkaline hydrolysis [66, 67], with the drawback of requiring an additional step for the removal of the leader peptide.

In a very innovative bioengineering work by Oman *et al*. [68], the authors took advantage of the biosynthetic pathway of lacticin 481, and described the creation of an

engineered synthetase capable of modifying substrates without a leader peptide. The trick consisted of attaching the cognate leader peptide to the N-terminus of the enzyme by means of a linker so that the enzyme was in a constitutively active state. By using this fusion enzyme, improved analogs of lacticin 481 harboring ncAAs were easily produced, bypassing the dilemma of the leader peptide. This example paved the way for other RiPP groups in which the need for a leader peptide hinders the production of novel variants.

1.3.2. Cyanobactins

The cyanobactins were originally defined as N-to-C macrocyclic peptides produced via PRPS which contain heterocyclized Ser, Thr and/or Cys (or their oxidized derivatives), and/or prenylated amino acids [51, 69]. However, with the recent discovery of new related compounds this definition has had to be widened, for example, to include linear variants [70]. These compounds have been isolated from free-living and symbiont cyanobacteria, and from marine animals such as ascidians and sponges. Nevertheless it is currently believed that cyanobacteria are the true and only producers of cyanobactins, with about 30% of all cyanobacterial species being able to synthesize them [71]. The members of the cyanobactin family are varied, including the patellamides, trunkamides, aeruginosamides and microcyclamides, among others, and exhibit bioactivities such as cytotoxic, antibacterial and antimalarial activity [49].

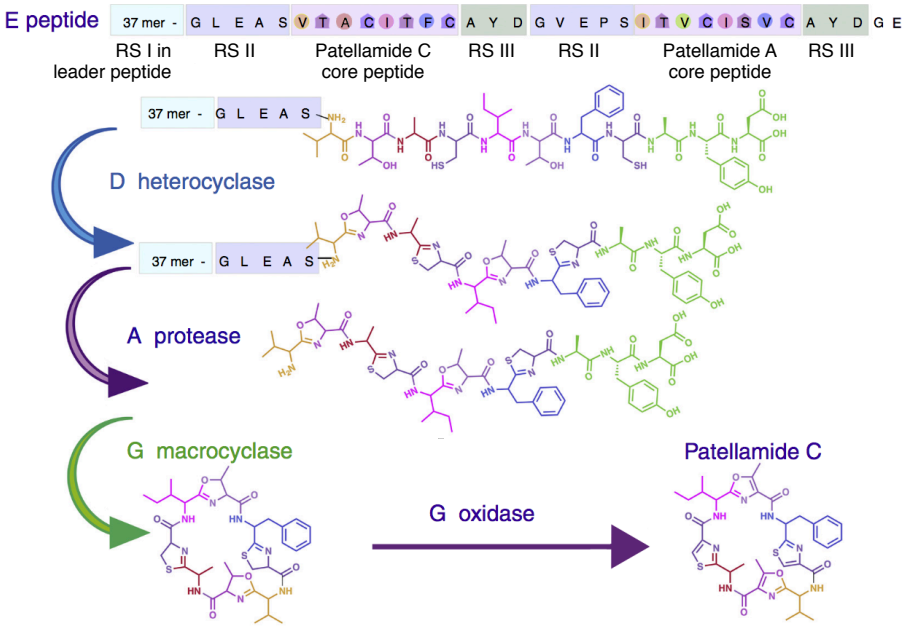


Figure 1.6. Patellamide biosynthesis. The precursor peptide E contains two core peptides. First, the heterocyclase D catalyzes the formation of heterocycles. Then, the N-terminal protease A cleaves off the leader peptide and RS II. Finally, the macrocyclase G removes the RS III, and catalyzes macrocyclization and heterocycle oxidation. Scheme modified from [72].

The minimal cyanobactin BGC is normally composed of an N-terminal protease gene, denoted by the letter *A*, two genes, *B* and *C*, coding for hypothetical proteins of unknown function and non-essential for the biosynthetic pathway, a precursor peptide gene *E*, and a C-terminal protease gene *G*, also responsible for the macrocyclization and in some cases, of heterocycle oxidation (Fig. 1.6). Additional tailoring enzyme genes can be found in the clusters, e.g. a cyclodehydratase *D*, a prenyltransferase *F*, a methyltransferase, a dedicated oxidase, and/or other genes coding for proteins of unknown function [71]. The PTM of the cyanobactin precursor peptides starts with the introduction of heterocycles (azolines) by an ATP-dependent cyclodehydratase, when present. The cyclodehydratases are activated by a conserved first recognition sequence (RS I), LAELSEEAL, located in the leader peptide, and upon activation they catalyze the formation of the heterocycles between the hydroxyl or sulfhydryl group from a Ser, Thr or Cys, and the amide bond N-terminal to it. After heterocylization, the N-terminal protease removes the leader peptide by acting on a semiconserved second RS (RS II). It is worth mentioning that the cyanobactin precursor peptides can contain up to 4 core peptides, each of them flanked by a RS II at their N-terminus and a RS III at their C-terminus. The C-terminal protease cleaves off the RS III and catalyzes the N-C macrocyclization (except on the linear cyanobactins). Finally, tailoring reactions such as oxidation of the heterocycles, prenylation or methylation take place [71, 72].

The cyanobactin family has been subject of many bioengineering approaches. The easiness to manipulate the biosynthetic pathway of this group resides in the fact that the PTM machinery is guided by defined transportable RSs, which allows for promiscuity regarding core peptide processing [73]. Some biotechnological endeavors achieved for this RiPP family *in vivo* in *E. coli* are the heterologous expression of patellamides [74, 75], the synthesis of eptidemnamide, a non-native cyanobactin analogous to a clinical anticoagulant [76], and the creation of a trunkamide pathway library, with which over 300 novel variants were produced [77]. Achievements done *in vitro* include the establishment of a system that couples cell-free translation with the cyclodehydratase PatD to generate azoline-containing peptides from synthetic DNA templates [78], the fluorescent labelling of cyanobactin analogs containing orthogonal reactive groups [79], the production of natural and unnatural cyanobactin derivatives by using mixtures of PTM enzymes from different pathways [80], and the creation of hybrid macrocycles containing amino acids and other functional groups such as aryl rings and polyethers [81].

As is common for PTM enzymes of PRPS pathways, the cyanobactin heterocyclases depend on the leader peptide, particularly the RS I, to be activated. The molecular basis for this has been illustrated in a study where the co-crystallization of LynD and its cognate leader peptide was reported [54]. As pointed out in section 1.3, the molecular interaction between this enzyme-substrate pair is very similar to that of the lanthipeptide dehydratase NisB and its leader peptide, involving a β-strand conformation of the conserved sequence in the leader peptide and a wHTH domain of the enzyme. In the same study the authors described the creation of a LynD cyclodehydratase fused to its leader peptide via the N-terminus, analogous to the lacticin 481 synthetase fusion enzyme mentioned in section 1.3.1. As in the case of the lanthipeptide synthetase, the engineered cyanobactin heterocyclase was constitutively active and could process precursor peptides without leader sequence in them [54].

1.3.3. Lasso peptides

Just as their name suggests, the structure of these RiPPs ressembles a threaded cowboy's lariat (Fig. 1.7). They are normally composed of 13-22 amino acids, with an N-terminal lactam ring formed between the N-terminal amino group of the core peptide and the ω-carboxyl group of an Asp or Glu at position 7, 8 or 9 [82]. The C-terminal sequence of the core peptide, the "tail", is trapped inside the lactam ring before cyclization, and the structure is stabilized by steric interactions between the ring and specific residues in the tail. Further structural stabilization can be provided by disulfide bridges between the ring and the tail, and the presence or absence of them is a feature used to subclassify lasso peptides into three groups: Variants of the class I contain two disulfide bonds, class II lasso peptides, which are the great majority, exhibit no disulfide bonds, and the only member of class III possesses one disulfide bond [51]. A fourth class has been recently proposed for a variant identified by genome mining which exhibits a C-terminal disulfide bridge that does not link the tail to the ring [83].

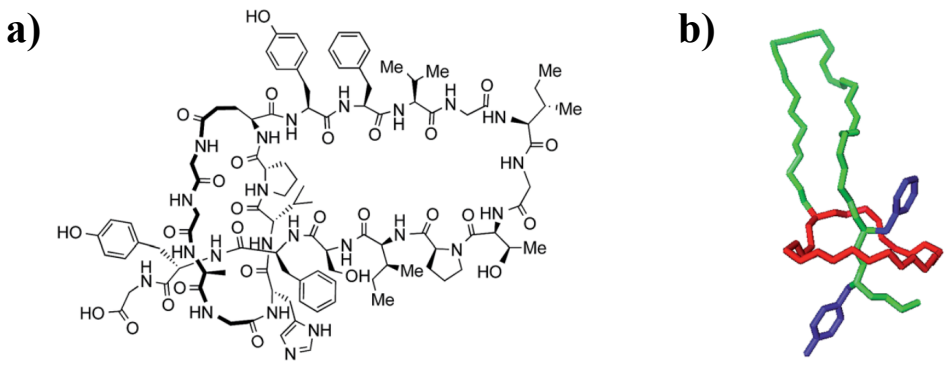


Figure 1.7. The lasso peptide microcin J25. a) 2D structure; **b)** 3D model. Image taken from [84].

In general, the minimal biosynthetic machinery required for lasso peptide production is comprised of a cysteine protease (designated with the letter B), in charge of cleaving off the leader peptide from the precursor (designated A), and an ATP-dependent lactam synthetase responsible for macrocyclization (designated C) (Fig. 1.8). Supplementary proteins can be encoded in the BGCs but they are non-essential for the production of most of the lasso peptides known to date. In some cases the cysteine protease is encoded as two discrete proteins, one homologous to the N-terminal domain and one to the C-terminal (proteolytic) domain of the full length protease [85]. The RRE identified by Burkhart *et al*. was also found in the N-terminal domain of many cysteine proteases from lasso peptide BGCs, both in full length proteases and in split proteases [57]. By aligning lasso peptide precursor sequences and RREs, the semiconserved motif YxxPxLx$_3$Gx$_5$Tx was identified in the C-terminal region of the leader peptide [83]. Recently, the independent studies of Cheung *et al.* [86] and Zhu *et al.* [87] showed that the interaction between the N-terminal domain of the B proteins and the lasso leader peptide is also facilitated through a wHTH motif and a β-strand conformation, respectively, as is the case for the class I lanthipeptides and the cyanobactins.

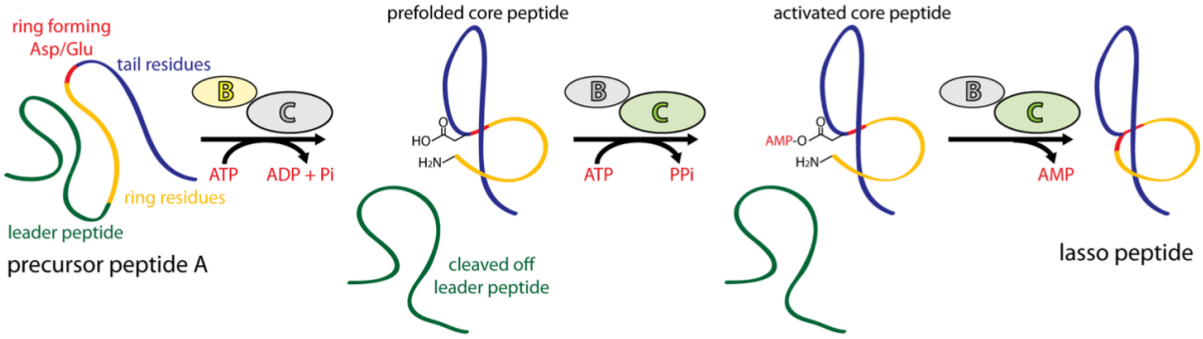


Figure 1.8. Proposed biosynthetic process of the lasso peptides. The cysteine protease B removes the leader peptide and possibly assists in the prefolding of the core peptide, and the lactam synthetase C catalyzes the N-terminal macrocyclization. Image taken from [82].

Members of this family of RiPPs have displayed resistance against thermal and proteolytic denaturation, and bioactivities such as antibiotic, receptor antagonist, and enzyme inhibiting activity [88]. Because of this, lasso peptides are interesting from a drug development point of view. The efforts toward the discovery of novel variants and their bioengineering have been plenty, and some of the most recent include a mapping of the lasso peptide chemical space, with over 1300 compounds identified [83], the chemical synthesis of lasso peptide analogs [84], the incorporation of ncAAs into capistruin [89], the

overexpression of chaxapeptin under co-culture conditions [90], and the grafting of an integrin antagonist onto the microcin J25 scaffold [91, 92], just to name a few.

1.4. Microviridins

The microviridins are another family of RiPPs. Since the present work was aimed toward engineering this type of natural products, they will be reviewed in more depth, presenting with more detail all the information gathered until now about this exciting group of RiPPs.

1.4.1. Brief history of microviridins

Microviridins owe their name to the cyanobacterium *Microcystis viridis*, species from which the first member of this group, microviridin A (Mdn A), was isolated in 1990 [93]. Twelve further congeners, namely Mdn B-J, Mdn SD1634, SD1652 and SD1684, had been isolated from laboratory and field cyanobacteria by 2006 as a result of screening programs for protease inhibitors [39, 94-98]. Along with IC_{50} values for different serine proteases such as elastase, trypsin and chymotrypsin, the structure of these metabolites was described too. They were reported to be N-acetylated trideca- and tetradecapeptides that bear one or two lactone rings and one lactam ring. As noted by Hemscheidt [99], it is now clear that those variants described as having only one lactone ring and a methyl ester in lieu of the second lactone ring are indeed isolation artifacts, product of the methanolysis of the tricyclic compounds. Apart from the fact that they were produced by some genera of cyanobacteria, nothing was known about their natural role or their biosynthesis. A study by Rohrlack *et al*. showed that Mdn J causes a lethal molting disruption in microcrustaceans of the genus *Daphnia*, which feed on planktonic cyanobacteria, pointing to a function as grazer deterrents [100]. Nevertheless, this function could only apply to those microviridin variants with an inhibitory profile similar to that of Mdn J. Since other variants show different inhibitory activity or no activity at all, the role of microviridin as grazer deterrent can not be generalized, and up to date the natural function of microviridin remains unclear. Regarding their biosynthesis, it was not until 2008 that the studies of Ziemert *et al.* [101] and Philmus *et al.* [102], by means of different approaches, finally shed light on the ribosomal origin of these compounds. The former group, using a fosmid carrying the biosynthetic microviridin genes from a strain of *M. aeruginosa*, achieved heterologous expression of Mdn B in *E. coli*, while the latter, using the purified microviridin PTM enzymes from a strain of *P. agardhii*, achieved *in vitro* cyclization and acetylation of the microviridin precursor peptide MvdE. Further research on microviridin has

been focused on characterizing with more detail its biosynthesis, on better comprehending its protease inhibitory activity to improve it, and on describing its natural diversity and occurrence. These aspects will be described separately in the following sections.

*Note on microviridin nomenclature: The group of Ziemert *et al*. named *mdnA-E* the microviridin biosynthetic genes from *M. aeruginosa*, while the group of Philmus *et al*. named *mvdA-F* those from *P. agardhii*. Since in the present work the microviridin biosynthetic machinery from *P. agardhii* was employed the latter nomenclature will be used, and the abbreviation Mdn will be utilized to refer to post-translationally modified microviridin core peptides, i.e. cyclic microviridins.

1.4.2. BGC and biosynthesis

The BGC of microviridin in *P. agardhii* is comprised of 6 genes (Fig. 1.9). *mvdA* codes for an ABC transporter which is thought to have an scaffolding role for a biosynthetic enzyme complex *in vivo* [103]. *mvdB* encodes a GNAT-type N-acetyltransferase, a tailoring enzyme responsible for the acetylation of the cyclized core peptide of microviridin. The genes *mvdC* and *mvdD* code for two ATP-grasp ligases, an ω-amide ligase and an ω-ester ligase, respectively. It is these PTM enzymes the ones that cyclize the microviridin core peptide. Finally, *mvdE* and *mvdF* encode two microviridin precursor peptides. Nevertheless, only the microviridin product of *mvdE* has been isolated from *P. agardhii* [102]. This microviridin variant, designated Mdn K, possesses the amino acid sequence YGNTMKYPSDWEEY, and will be used as reference in this work.

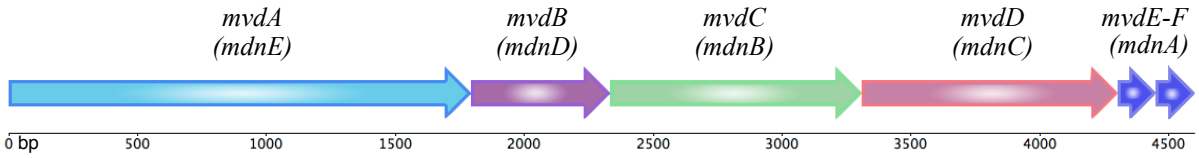


Figure 1.9. Microviridin BGC in *P. agardhii* NIVA-CYA 126. The cluster is ~5 kbp long and contains six genes: Two precursor peptides (*E* and *F*), two PTM enzymes (*D* and *C*), one tailoring enzyme (*B*) and one ABC transporter (*A*). The gene homologs from *M. aeruginosa* are indicated in parenthesis; in *M. aeruginosa* there is only one precursor peptide gene, *mdnA*.

The order in which these enzymes act has been elucidated too (Fig. 1.10). First, the ATP-grasp ligase MvdD introduces the large lactone ring between the β-hydroxy group of the threonine and the ω-carboxyl group of the aspartate, followed by the introduction of the small lactone ring between the β-hydroxy group of the serine and the ω-carboxyl group of the first

glutamate [104]. Then, the ATP-grasp ligase MvdC introduces the lactam ring between the ε-amino group of the lysine and the ω-carboxyl group of the second glutamate. It has been shown that the cyclization of microviridin follows this strict order, and that a precursor with only one lactone ring is not further processed by the ω-amide ligase. It was shown too that MvdD is a distributive enzyme, i.e. the intermediate product of the first enzymatic reaction is released, since a monocyclic intermediate could be detected in *in vitro* assays [102]. After cyclization by the ATP-grasp ligases, the leader peptide is excised. Until now, the protease responsible for this reaction has not been described. Finally, the N-acetyltransferase product of *mvdB* catalyzes the transfer of the acetyl group from acetyl coenzyme A (AcCoA) to the N-terminus of the cyclic peptide [102]. As mentioned before, it is believed that the ABC transporter MvdA plays a role in the assembly of a biosynthetic enzyme complex *in vivo* [103], and possibly in the export of microviridins out of the cell, though this has not been demonstrated.

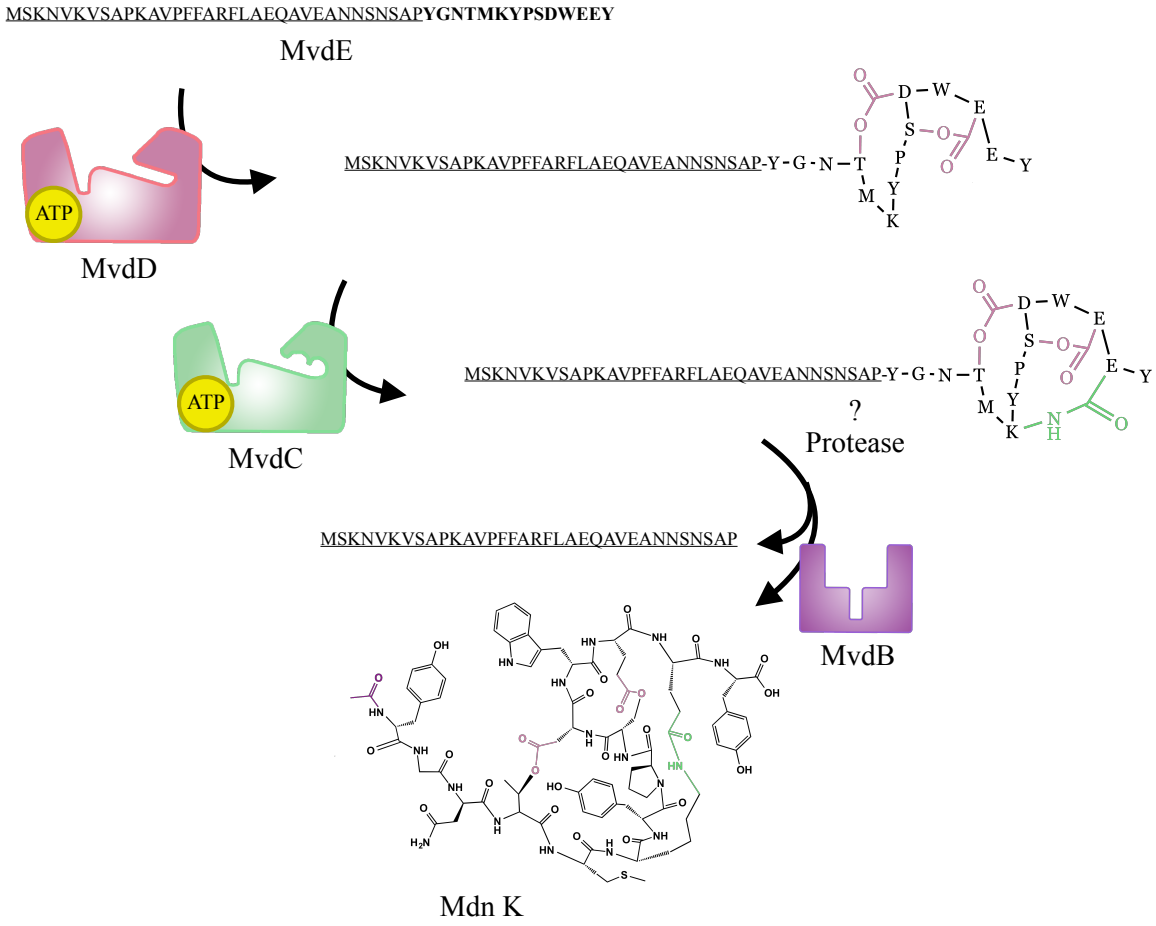


Fig. 1.10. PTM of microviridin. First, the ω-ester ligase installs the two lactone rings in the core peptide (in bold). Then, the ω-amide ligase introduces the lactam ring. After this, an unknown protease removes the leader peptide (underlined), and finally the N-acetyltransferase catalyzes the acetylation of the tricyclic peptide.

1.4.3. Important structural features and first bioengineering efforts

As pointed out before, the leader peptide is important for activation of some PTM enzymes in the ribosomal pathway. In the case of microviridin, the leader peptide is necessary for the activation of the ATP-grasp ligases. A conserved motif in the leader peptide, PFFARFL, had been identified as crucial for the activity of the ligases [103]. As described in section 1.3, a recent crystallization study of MdnC, an MvdD homolog from *M. aeruginosa*, has shown that the PFFARFL motif adopts an α-helix conformation and interacts with an α-helix from the ATP-grasp ligase, thus demonstrating the direct interaction between this enzyme and the leader peptide [56]. In the case of the core peptide, the conserved motif TxKxPSDWE(D/E) has been identified too [103, 105].

The first attempts to engineer microviridins were aimed at modifying the size of the rings. These trials were carried out *in vitro* using the purified PTM enzymes from *P. agardhii* and mutated MvdE precursor peptides expressed using the pET28b vector. The results of this study showed that the ring size is fixed, and modification or displacement of the amino acids directly involved in the cyclization abolish ring formation [104]. Further attempts to engineer these metabolites were aimed at improving their inhibitory activity and installing a leader peptide cleavage site. It had been hypothesized that the 5th amino acid in the microviridin sequence played a role in the inhibitory selectivity of the variant, based on the observation that known variants with different residues at this position showed different protease selectivity. This hypothesis was experimentally corroborated using a minimal expression platform based on the pDrive vector harboring the microviridin BGC of *M. aeruginosa* NIES 843 and heterologous expression in *E. coli*. The results of this work showed that at position 5 Leu conferred selectivity towards elastase, Met towards subtilisin, and Arg and Lys towards trypsin. Moreover, in the same study the protease trypsin was co-crystallized with Mdn J, revealing the interaction between some residues of the inhibitor, particularly of the Arg at position 5, and residues of the enzyme, thus demonstrating the key role of the amino acid at this position for the inhibitory activity of microviridin (Fig. 1.11) [106]. Since the protease responsible for cleaving off the leader peptide of microviridin in cyanobacteria has not been described, artificial cleavage sites for trypsin, factor Xa protease or hydroxylamine were introduced between the leader and the core peptide using the same minimal platform to improve the yield of the expected 14-membered microviridins. This approach not only failed to achieve the latter, but also abolished completely the synthesis of microviridins in the case

of the introduction of the cleavage site for the factor Xa protease and for hydroxylamine [107].

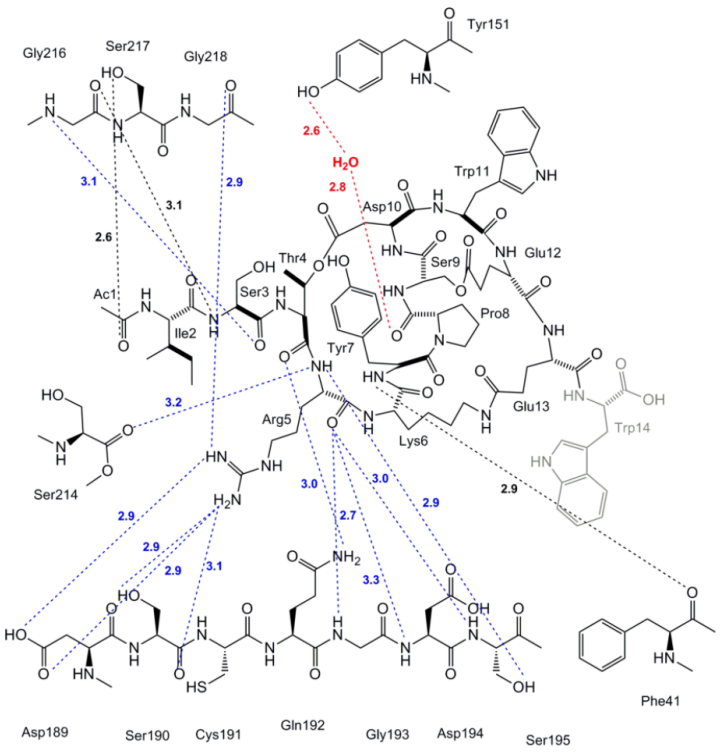


Figure 1.11. Interaction of trypsin and Mdn J. It can be appreciated that the N-terminus of Mdn J interacts with residues 216-217 of trypsin, and that the Arg at position 5 interacts with the residues 189-190, and 218 of trypsin. The scheme was taken from the supp. inf. of [106].

1.4.4. Natural diversity of microviridins

All the reported members of this class of RiPPs have been isolated from the same source, the cyanobacteria. Furthermore, studies with laboratory strains and metagenomic samples of *Microcystis* revealed the ability of this kind of bacteria to produce diverse microviridins [105, 108]. Nonetheless, early bioinformatic analyses had shown related BGC present in genera of bacteria not belonging to the cyanobacterial phylum, for example *Sorangium*, *Microscilla*, and *Alteromonas* [102, 103]. Indeed, recent work on genome mining for BGC of RiPPs identified microviridin biosynthetic clusters in several genera of the phyla Bacteroidetes and Proteobacteria, apart from several novel clusters of the cyanobacterial phylum [48]. Altogether, these findings confirm the great microviridin diversity waiting to be exploited.

1.5. Aims of this study

As described in this section, recent technological advances have made the collection of new valuable information about the RiPPs possible, including insights into their biosynthetic pathways, their distribution in nature, and their potential applications. In some cases, the extrapolation of knowledge from one RiPP family to another has been possible too, contributing to the fast development of the field of natural products. The elucidation of the molecular basis of the requirement of a leader peptide for PTM of the core peptide not only has allowed for the expansion of the diversity of precursor peptides but has even circumvented the need of a leader peptide in some cases, further promoting the production of novel natural products. The goal of the present work was to extrapolate this knowledge to the family of the microviridins to establish an efficient *in vitro* platform for the easy production and characterization of synthetic analogs. The creation of synthetic microviridins would pursue two objectives: the production of synthetic and cryptic variants with improved bioactivity, and the synthesis of variants harboring ncAAs as molecular tools for further research.

2. Materials and methods

2.1. Materials

2.1.1. Organisms

For this work the following strains were used:

1. *Planktothrix agardhii* NIVA-CYA 126, obtained from the Culture Collection of Algae at the Norwegian Institute for Water Research, Norway.
2. *Algicola sagamiensis* DSM 14643, DNA obtained from the Leibniz Institute German Collection of Microorganisms and Cell Cultures (DSMZ), Germany.
3. *Cyanothece* sp. PCC 7822, obtained from the Pasteur Culture Collection, France.
4. *Escherichia coli* TOP10 chemically competent cells for cloning procedures. (Invitrogen, USA).
5. *Escherichia coli* BL21 (DE3) chemically competent cells for protein expression. (Merck Chemicals Ltd., UK).

2.1.2. Chemicals and antibiotics

2-Mercaptoethanol	Ferak-Berlin GmbH (Germany)
α-Cyano-4-hydroxycinnamic acid (HCCA)	Sigma-Aldrich (Germany)
Acetic acid	Carl Roth (Germany)
Acetonitrile HPLC grade	VWR Chemicals (USA)
Acetyl coenzyme A sodium salt	Sigma-Aldrich (Germany)
Ampicillin sodium salt	Carl Roth (Germany)
APS	Carl Roth (Germany)
ATP disodium salt hydrate	Carl Roth (Germany)
BAPNA	Sigma-Aldrich (Germany)
Bromophenol blue	Carl Roth (Germany)
Coomassie Brilliant Blue-G250	Thermo Scientific (USA)
DMSO	Carl Roth (Germany)
Disodium phosphate dihydrate	Carl Roth (Germany)
dNTP	Thermo Scientific (USA)
EDTA disodium salt dihydrate	Carl Roth (Germany)
Ethanol denatured 96%	VWR Chemicals (USA)
Ethidium bromide	Sigma-Aldrich (Germany)
Glycerol Rotipuran	Carl Roth (Germany)

Hydrochloric acid 37%	VWR Chemicals (USA)
Imidazole	Carl Roth (Germany)
IPTG	VWR Chemicals (USA)
Kanamycin sulfate	Carl Roth (Germany)
LE agarose	Biozym (Germany)
Magnesium chloride hexahydrate	Carl Roth (Germany)
Methanol HPLC grade	VWR Chemicals (USA)
Monosodium phosphate monohydrate	Carl Roth (Germany)
Ni-NTA agarose	Cube Biotech (Germany)
o-Phosphoric acid 85%	Carl Roth (Germany)
Potassium chloride	Carl Roth (Germany)
Powdered milk blotting grade	Carl Roth (Germany)
Rotiphorese Gel 30	Carl Roth (Germany)
SDS ultra-pure	Carl Roth (Germany)
Sodium chloride	Carl Roth (Germany)
Sodium hydroxide	Merck Chemicals Ltd. (UK)
Suc-AAA-*p*NA	Sigma-Aldrich (Germany)
Suc-AAPF-*p*NA	Sigma-Aldrich (Germany)
TEMED	Carl Roth (Germany)
Trifluoroacetic acid	Thermo Scientific (USA)
Tris	Carl Roth (Germany)
Tween 20	Carl Roth (Germany)
Urea	M.P. Biomedicals Inc. (Germany)
X-Gal	Carl Roth (Germany)
Xylene cyanol	Carl Roth (Germany)

2.1.3. Media and solutions

The following preparations are referred to as 1 L of ddH_2O unless otherwise specified.

BG-11 Medium	As described by Rippka *et al.* 1979 [109].
LB agar	Carl Roth (Germany)
LB broth	Carl Roth (Germany)
SOC medium	3.6 g glucose, 0.18 g KCl, 2.0 g $MgCl_2$•$6H_2O$, 0.5 g NaCl, 20 g tryptone, 5.0 g yeast extract,

	pH 7.0
Coomassie blue solution	50 g $Al_2(SO_4)_3$•$16H_2O$, 100 mL ethanol 96%, 0.20 g Coomassie blue, 94 mL *o*-phosphoric acid 85%
Destaining solution (Coomassie blue)	100 mL ethanol 96%, 20 mL *o*-phosphoric acid 85%
Gel filtration buffer	8.77 g NaCl, 2.42 g Tris-HCl, pH 7.5
His-tagged protein purification:	
Base buffer	17.53 g NaCl, 5.99 g NaH_2PO_4, pH 8.0
Lysis buffer (10 mM imidazole)	Base buffer + 0.68 g imidazole
Wash buffer 1 (20 mM imidazole)	Base buffer + 1.36 g imidazole
Wash buffer 2 (50 mM imidazole)	Base buffer + 3.40 g imidazole
Elution buffer (250 mM imidazole)	Base buffer + 17.02 g imidazole
Buffer B	12.00 g NaH_2PO_4, 10 ml 1 M Tris-HCl pH 8.0, 480.48 g urea, pH 8.0
HCCA matrix stock solution	10 mg HCCA in 1 mL 84% acetonitrile/ 13% ethanol/ 3% TFA 0.1%
HCCA matrix working solution	35 µL stock solution + 65 µL 84% acetonitrile/ 13% ethanol/ 3% TFA 0.1%
6x DNA loading dye	90 mg bromophenol blue, 6.7 g EDTA, 180 mL glycerine, 3 mL 1 M Tris-HCl pH 7.6, 90 mg xylene cyanol in 300 mL ddH_2O
5x SDS loading dye	1.0 g bromophenol blue, 100 mL glycerine, 50 mL 2-mercaptoethanol, 62.5 mL 1 M Tris-HCl pH 6.8, 100 mL 20% SDS
Phosphate-buffered saline (PBS)	0.2 g KCl, 0.24 g KH_2PO_4, 8 g NaCl, 1.44 g Na_2HPO_4, pH 7.4
Protease inhibition assays:	
Elastase working solution	Dilution of the commercial elastase solution to 1.5 U mL^{-1} in buffer 50 mM Tris-HCl, 20 mM $CaCl_2$, pH 7.5
Trypsin stock solution	1 mg trypsin in 1 mL buffer 50 mM Tris-HCl, 20 mM $CaCl_2$, pH 7.5

Trypsin working solution	1:10 dilution (in same buffer) of trypsin stock solution
Subtilisin stock solution	1 mg subtilisin in 1 mL PBS
Subtilisin working solution	1:10 dilution (in PBS) of subtilisin stock solution
BAPNA stock solution	43.3 mg in 1 mL DMSO
BAPNA working solution	1:20 dilution (in trypsin's assay buffer) of BAPNA stock solution
Suc-AAA-*p*NA working solution	1 mg Suc-AAA-*p*NA in 1 mL elastase's assay buffer
Suc-AAPF-*p*NA stock solution	2.5 mg Suc-AAPF-*p*NA in 500 µL DMSO
Suc-AAPF-*p*NA working solution	1:100 dilution (in PBS) of Suc-AAPF-*p*NA stock solution
SDS-PAGE running buffer	14.42 g glycine, 1 g SDS, 3.03 g tris
50x TAE	57.1 mL acetic acid, 100 mL 0.5 mM EDTA pH 8.0, 242 g tris
TBS-T	8.77 g NaCl, 6.06 g tris, 1 mL tween 20, pH 7.4
Western Blot:	
Transfer buffer	14.42 g glycine, 3.03 g tris, 10 mL ethanol 96%
Blocking solution	1 g powdered milk in 20 mL TBS-T
Primary antibody solution	1:10 000 dilution of monoclonal anti-poly histidine antibody produced in mouse in TBS-T
Secondary antibody solution	1:10 000 dilution of anti-mouse IgG (whole molecule)–peroxidase antibody produced in rabbit in TBS-T

2.1.4. Commercial Kits

Bio-Rad Protein Assay	Bio-Rad (Germany)
GeneJET Gel Extraction	Thermo Scientific (USA)
GeneJET Plasmid Miniprep Kit	Thermo Scientific (USA)
innuPREP PCR pure kit	Analytik Jena (Germany)
Metagenomic DNA Isolation Kit for Water	Epicentre Biotechnologies (USA)
Serva Light Polaris CL HRP WB Substrate kit	Serva (Germany)

2.1.5. Cloning and expression vectors

pDrive: routine TA cloning (Qiagen PCR cloning kit, Germany).

pQE80L: protein expression with N-terminal 6x His-tag (Qiagen cis-Repressed pQE Vector Set, Germany).

2.1.6. Commercial enzymes and antibodies

The restriction enzymes used in this work as well as the DreamTaq and Phusion High-Fidelity DNA polymerases, and the T4 DNA ligase were all purchased from Thermo Scientific (USA). The proteases elastase (from porcine pancreas, E1250), subtilisin (protease, P8038), and trypsin (from porcine pancreas, T7409) as well as the monoclonal anti-poly histidine antibody produced in mouse and the anti-mouse IgG (whole molecule)–peroxidase antibody produced in rabbit were all purchased from Sigma-Aldrich (Germany).

2.1.7. Peptides

The peptides employed in this work are all synthetic and were either obtained from a collaboration or purchased from private companies. In both cases they were synthesized via solid-phase synthesis, and in the case of the peptides bought from private companies the requested purity was always > 75%.

Table 2.1. Precursor peptides used in this work

Sequence	Monoisotopic mass (Da)	Description of the peptide	Code	Manufacturer
MSKNVKVSAPKAVPF FARFLAEQAVEANNS NSAPYGNTMKYPSD WEEY	5382.564	*P. agardhii* NIVA-CYA 126 first microviridin precursor (MvdE)	MvdE	Genscript (USA)
MSKNVKVSAPKAVPF FARFLAEQAVEANNS NSAP	3618.851	*P. agardhii* NIVA-CYA 126 microviridin leader peptide	MvdE leader	Synpeptide (China)
YGNTMKYPSDWEEY	1781.724	*P. agardhii* NIVA-CYA 126 microviridin core peptide (MdnK)	MvdE core/ M5M	
Mdn K peptide library: Substitution of the 5th position in MdnK with the 20 natural amino acids				
YGNTMKYPSDWEEY	1781.724	*P. agardhii* NIVA-CYA 126 microviridin core peptide (MdnK)	MvdE core/ M5M	
YGNTAKYPSDWEEY	1721.721	Substitution with alanine	M5A	

YGNTRKYPSDWEEY	1806.785	Substitution with arginine	M5R	
YGNTNKYPSDWEEY	1764.727	Substitution with asparagine	M5N	
YGNTDKYPSDWEEY	1765.711	Substitution with aspartate	M5D	
YGNTCKYPSDWEEY	1753.693	Substitution with cysteine	M5C	
YGNTEKYPSDWEEY	1779.726	Substitution with glutamate	M5E	
YGNTQKYPSDWEEY	1778.742	Substitution with glutamine	M5Q	
YGNTGKYPSDWEEY	1707.705	Substitution with glycine	M5G	
YGNTHKYPSDWEEY	1787.743	Substitution with histidine	M5H	Bianca Schmid, AG Süssmuth, TU Berlin
YGNTIKYPSDWEEY	1763.768	Substitution with isoleucine	M5I	
YGNTLKYPSDWEEY	1763.768	Substitution with leucine	M5L	
YGNTKKYPSDWEEY	1778.779	Substitution with lysine	M5K	
YGNTFKYPSDWEEY	1797.752	Substitution with phenylalanine	M5F	
YGNTPKYPSDWEEY	1747.737	Substitution with proline	M5P	
YGNTSKYPSDWEEY	1737.716	Substitution with serine	M5S	
YGNTTKYPSDWEEY	1751.731	Substitution with threonine	M5T	
YGNTWKYPSDWEEY	1836.763	Substitution with tryptophan	M5W	
YGNTYKYPSDWEEY	1813.747	Substitution with tyrosine	M5Y	
YGNTVKYPSDWEEY	1749.752	Substitution with valine	M5V	

Mdn B peptide library: Substitution of the 5th position in MdnB with the 20 natural amino acids

FGTTLKYPSDWEEY	1734.778	*M. aeruginosa* NIES 298 microviridin core peptide (MdnB)	L5L	
FGTTAKYPSDWEEY	1692.731	Substitution with alanine	L5A	

FGTTRKYPSDWEEY	1777.795	Substitution with arginine	L5R	Bianca Schmid, AG Süssmuth, TU Berlin
FGTTNKYPSDWEEY	1735.737	Substitution with asparagine	L5N	
FGTTDKYPSDWEEY	1736.721	Substitution with aspartate	L5D	
FGTTCKYPSDWEEY	1724.703	Substitution with cysteine	L5C	
FGTTEKYPSDWEEY	1750.736	Substitution with glutamate	L5E	
FGTTQKYPSDWEEY	1749.752	Substitution with glutamine	L5Q	
FGTTGKYPSDWEEY	1678.715	Substitution with glycine	L5G	
FGTTHKYPSDWEEY	1758.753	Substitution with histidine	L5H	
FGTTIKYPSDWEEY	1734.778	Substitution with isoleucine	L5I	
FGTTMKYPSDWEEY	1752.734	Substitution with methionine	L5M	
FGTTKKYPSDWEEY	1749.789	Substitution with lysine	L5K	
FGTTFKYPSDWEEY	1768.762	Substitution with phenylalanine	L5F	
FGTTPKYPSDWEEY	1718.746	Substitution with proline	L5P	
FGTTSKYPSDWEEY	1708.726	Substitution with serine	L5S	
FGTTTKYPSDWEEY	1722.741	Substitution with threonine	L5T	
FGTTWKYPSDWEEY	1807.773	Substitution with tryptophan	L5W	
FGTTYKYPSDWEEY	1784.757	Substitution with tyrosine	L5Y	
FGTTVKYPSDWEEY	1720.762	Substitution with valine	L5V	
Genome-mined microviridins library: Selection of bioinformatically predicted microviridins encoded in the genome of *Cyanothece* sp. PCC 7822, *P. agardhii*, and *A. sagamiensis*				
YQNTLKYPSDWEEY	1834.794	14 amino acids long with Leu at position 5	Cth 1	
-TVTRKYPSDWEDY	1658.747	13 amino acids long with Arg at position 5	Cth 2	
-YVTKKYPSDWEEY	1706.772	13 amino acids long with Lys at position 5	Cth 3	

--ATLKYPSDWEEY	1500.666	12 amino acids long with Leu at position 5	Cth 4	Bianca Schmid, AG Süssmuth, TU Berlin
--FTLKYPSDWEDE	1528.661	12 amino acids long with Leu at position 5 and Glu at the C-terminus	Cth 5	
TIWTFKWPSDWEDS	1796.794	*P. agardhii* NIVA-CYA 126 second microviridin core peptide	MvdF core	
FATMRYPSDSDE	1417.571	Marinostatin core peptide	Mst	Biomatik (USA)
Mdn J with non-canonical amino acids peptide library: Introduction of non-canonical amino acids in the sequence of Mdn J				
-ISTRK*YPSDWEEW	1921.8665	*Microcystis* UOWOCC MRC microviridin core peptide (MdnJ) with biotinylated Lys at position 6	J-MidBiotin	Synpeptide (China)
K*ISTRKYPSDWEEW	2049.9618	Introduction of a biotinylated Lys at the N-terminus	J-Nbiotin	Bianca Schmid, AG Süssmuth, TU Berlin
-ISTRKYPSDWEEWK*	2049.9618	Introduction of a biotinylated Lys at the C-terminus	J-Cbiotin	

2.1.8. General Equipment

Falcon centrifuge	Heraeus multifuge 1S-R (Thermo Scientific, USA)
High speed centrifuge	6-16K with rotors 12169H and 12500H (Sigma Laborzentrifugen GmbH, Germany)
Microcentrifuge	PerfectSpin 24R refrigerated centrifuge. (Peqlab Biotechnologie GmbH, Germany)
Gel documentation system	ChemiDoc XRS+ System (Bio-Rad, Germany)
Spectrophotometer	UV-1800 UV Spectrophotometer (Shimadzu, Germany)
SDS-PAGE and western blot equipment	Bio-Rad, Germany

2.2. Methods

2.2.1. Cultivation

Cyanobacteria strains were cultivated in liquid BG-11 medium under constant light conditions of approximately 20 µmol photons $m^{-2}s^{-1}$ at 23ºC.

E. coli overnight cultures were incubated in LB broth at 37ºC under agitation of 220 rpm for 16 hours. Expression cultures (400 mL) were incubated under agitation of 220 rpm at different temperatures for either 16 hours or 3 days. LB agar plates were incubated at 37ºC for 16 hours. When required the media were supplemented with antibiotics.

2.2.2. Genomic DNA extraction from cyanobacteria

For the extraction of cyanobacterial genomic DNA the "Metagenomic DNA isolation kit for water" was used. Between 5 and 10 mL of a medium- to high-density culture was centrifuged at 10000 g for 10 min. The supernatant was discarded and the manufacturer's protocol was followed from step 7.

2.2.3. *mvdB, mvdC,* and *mvdD* cloning

Note: All the primers used in this work are listed in the appendix in Table A1.

The genes *mvdB*, *mvdC*, and *mvdD* were amplified from DNA of *P. agardhii* NIVA-CYA 126 using the Phusion polymerase and primers designed to introduce restriction sites (BamHI and PstI) flanking the genes. The PCRs were performed on a C1000 Thermal Cycler (Bio-rad, Germany) with the conditions described in table 2.2.

Table 2.2. Phusion PCR settings.

PCR mix	PCR conditions			
1 µL DNA template [50-100 ng/µL]		Temperature (ºC)	Time (s)	
1 µL Fwd primer [10 µM]	Initial denaturation	98	180	
1 µL Rv primer [10 µM]	Denaturation	98	60	
4 µL 5x Green HF buffer	Annealing	-----	30	30 cycles
0.4 µL dNTPs [10 mM]	Extension	72	30/Kb	
0.2 µL Phusion Pol.	Final extension	72	600	
12.4 µL H_2O				
20 µL total				

After amplification, 2 µL of the PCRs was analyzed via gel electrophoresis with 1% agarose gels containing 50 µg/mL ethidium bromide under 90-110 V. Lambda DNA (Thermo Scientific, USA) digested with PstI restriction enzyme and mixed with bromophenol blue was used as DNA ladder. The gels were visualized in a gel documentation system. Successful PCRs were then purified with the "innuPREP PCR pure kit". 35 µL of the purified products was mixed with 4 µL 10x DreamTaq buffer, 1 µL dNTPs [10 mM], and 1 µL DreamTaq and incubated at 72ºC for 10 min for polyadenylation. The polyadenylated products were then cloned into pDrive using the "Qiagen PCR cloning kit" for maintenance and posterior subcloning.

5 µL of the ligation reaction was added to a 50 µL aliquot of *E. coli* TOP10 chemically competent cells for transformation. The cells were left on ice for 30 minutes and then heat shocked at 42ºC for 45 seconds. Immediately after, 500 µL of SOC medium was added and the samples were placed on ice again for 2 minutes. The cells were incubated afterwards at 37ºC for 1 hour under agitation of 220 rpm, and plated on selective LB agar with X-gal and IPTG for blue/white screening. 4 white colonies were normally selected per plate and used to prepare overnight cultures. The plasmids were purified from these cultures using the "GeneJET Plasmid Miniprep Kit", and they were later analyzed via restriction reaction and gel electrophoresis. The plasmids containing an insert of the expected size were then sent to GATC Biotech (Germany) to be sequenced. DNA sequences were analyzed using BLAST and SnapGene Viewer 3.1.4.

The enzymes BamHI and PstI were used to cleave the inserts with correct sequence from pDrive. The digestion reactions were set up according to the manufacturer's protocol. After incubation the digestion was analyzed via agarose gel electrophoresis, and the bands with the expected size were excised from the gel and purified with the "GeneJET Gel Extraction" kit. Finally, the inserts' concentration was measured with a Nanodrop ND 2000 (Thermo Scientific, USA). The vector pQE80L, which is cis-repressed, provides Amp resistance and encodes an N-terminal His-tag, was chosen for expression. This vector was also digested with BamHI and PstI, and after purification the open vector and inserts were ligated with the T4 DNA ligase for 16 hours at 4ºC. 5 µL of the ligation was used to transform *E. coli* BL21 (DE3) as previously mentioned. After purification the plasmids were sent to sequencing, and the constructs with correct sequence were used for expression.

2.2.4. Optimization of the expression of MvdC and MvdD and affinity chromatography

A first expression test was carried out by inoculating 50 mL of LB medium (supplemented with ampicillin) with 2 mL of an overnight culture of transformed *E. coli* BL21 (DE3) cells (1:25 dilution). The cultures were grown at 37ºC under agitation of 220 rpm to an OD_{600} of 0.5 and then IPTG was added to a final concentration of 1 mM to induce protein expression. The cultures were further incubated at 37ºC for 5 hours, and 8 mL of each culture were harvested by centrifugation. The cell pellets were washed once with 1 M Tris-HCl pH 7.5 and split into two pellets per culture. One pellet was resuspended in 1 mL lysis buffer and the other pellet was resuspended in 500 µL buffer B. The pellets were then submitted to ultrasonication (HD 3100, Bandelin, Germany) for 10 min (3 s on/ 3 s off) with an amplitude of 70% on ice. The samples were centrifuged afterwards at 4ºC, maximum speed for 5 min, and the supernatants were analyzed via discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and western blot. For this, 40 µL of the supernatants were mixed with 10 µL of 5x SDS loading dye and incubated at 95ºC for 10 min. The samples were then loaded onto the stacking gel (4% acrylamide) together with PageRuler Prestained Protein Ladder (Thermo Scientific, USA) and run through a 12.5% acrylamide resolving gel in a Tris-glycine buffer. The SDS-PAGE was done in duplicates to use one gel for Coomassie blue staining and one for western blot. After electrophoresis, the gels were washed thrice with ddH_2O. In the case of Coomassie staining, the gel was incubated in 20 mL of Coomassie blue solution under light agitation overnight. The gel was then washed with ddH_2O and submerged in 20 mL of destaining solution for one hour before being photographed in a gel documentation system. For western blot a 0.45 µm nitrocellulose membrane (GE Healthcare, Germany) was used. After blotting, the membrane was washed four times with TBS-T and incubated in blocking solution for one hour at RT. The membrane was then washed once with TBS-T and incubated with the primary antibody for one hour at RT. Afterwards, the membrane was washed four times with TBS-T and incubated with the secondary antibody for one hour at RT. Finally, the secondary antibody solution was washed away with TBS-T, and the membrane was developed with the "Serva Light Polaris CL HRP WB Substrate" kit and photographed in a gel documentation system.

Different expression conditions were tested: 30ºC for 5 hours and 16ºC for 16 hours. The cultures were harvested and lysed in 10 mL lysis buffer as aforementioned. After centrifugation at maximum speed, a sample of the supernatants was taken (Lysate, L) and they were left in incubation with 1 mL Ni-NTA agarose at 4ºC for 1 hour under constant rotation.

The samples were then centrifuged and a sample of the supernatants was taken (Flow through, FT) before being discarded. 5 mL of wash buffer 1 was added to the agarose, and the sample was vortexed and centrifuged again. A sample was taken (Wash 1, W1) and the wash step was repeated once more (Wash 2, W2) before proceeding to elute the proteins in two steps (Elution 1 and Elution 2) with 1.25 mL of elution buffer in each step. The samples were analyzed via SDS-PAGE and western blot.

The expression tests were then escalated to 400 mL cultures maintaining the 1:25 dilution of overnight cultures, and two different expression conditions were tried: 16ºC for 16 hours and room temperature for 3 days. In the case of the culture incubated at room temperature for 3 days the second wash step was done with wash buffer 2. For the rest of the project, the expression of the ATP-grasp ligases of *P. agardhii* was carried out at room temperature for 3 days, with the wash buffer 2 being used for the second wash step.

2.2.5. Gel filtration chromatography and enzyme concentration for crystallography

One of the objectives of the project was to solve the crystal structure of the ATP-grasp ligases from *P. agardhii*. For this reason another step of protein purification was necessary to prepare high purity solutions of the enzymes for crystallization. After purification with Ni-NTA agarose the enzyme eluates were further purified by gel filtration chromatography (HiLoad 16/600 SuperDex 200 column, Äkta Prime Plus, GE Healthcare, Germany) with a buffer 150 mM NaCl, 20 mM Tris-HCl, pH 7.5. Fractions of 2.5 mL were collected and the ones containing the enzyme were pooled together and concentrated by centrifugation (Amicon Ultra-4 centrifugal filter units, Millipore, Germany), with its concentration being determined through the use of a “Bio-Rad protein assay” kit. The concentrated protein solutions were added with glycerol to a final concentration of 10% before being sent to the group of Prof. Michael Groll at the Technical University of Munich for crystallization studies.

2.2.6. Engineering of the microviridin ATP-grasp ligases

The enzymes were engineered by overlap extension PCR as described in [68] (Fig. 2.1). Briefly, two short DNA sequences with complementary endings were synthesized by PCR in the first step. The first sequence contained the *Sac*I restriction site, the partial gene coding for the leader peptide of MvdE, the *Spe*I restriction site, a spacer, the *Kpn*I restriction site, and the first nucleotides of the ligase gene. The second sequence was composed of the last nucleotides of the partial gene coding for the leader peptide of MvdE, the *Spe*I restriction site,

a spacer, the *Kpn*I restriction site, the ligase gene, and the *Pst*I restriction site. These two sequences were combined in an overlap extension PCR, and the final product, starting with the *Sac*I restriction site and ending with the *Pst*I restriction site, was cloned into pQE80L. The linker coding for the 15x GlySer repetition was prepared from self-complementary oligonucleotides containing overhangs identical to those generated by the enzymes *Spe*I and *Kpn*I. The construct was then digested with *Spe*I and *Kpn*I, and the linker was ligated to it.

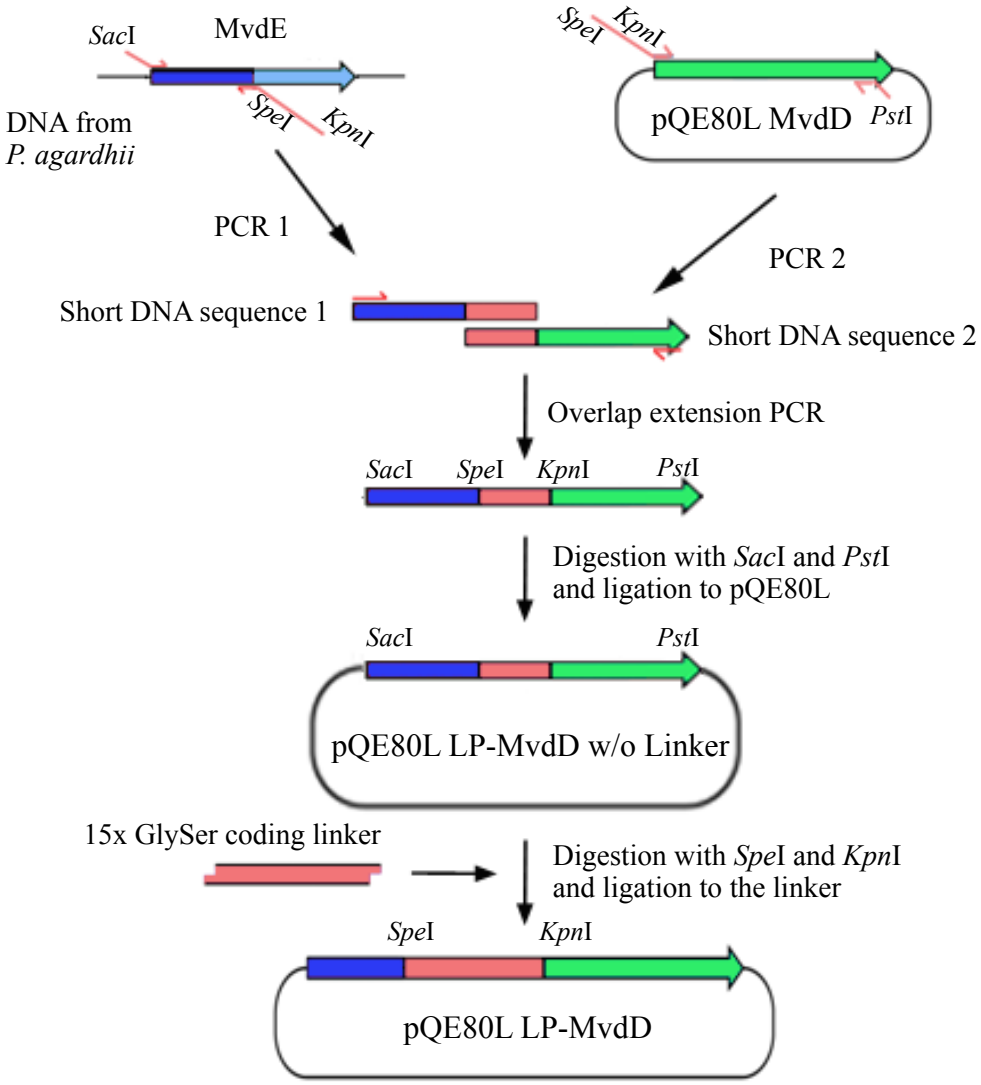


Figure 2.1. Scheme of the engineering of the ATP-grasp ligases. A PCR is conducted to amplify the leader peptide of MvdE with sites for restriction enzymes and the first nucleotides of the ligase gene. Another PCR is realized to amplify the gene of the ligase with sites for restriction enzymes and the last nucleotides of the leader peptide partial gene. The two products are joint in an overlap extension PCR and the final product is cloned into pQE80L. Finally the construct is digested with *Spe*I and *Kpn*I and ligated to the linker. Figure modified from the Supp. Inf. of [68].

2.2.7. *Algicola sagamiensis* ATP-grasp ligase engineering

The gene in the *A. sagamiensis* genome predicted to be a protein belonging to the MvdD family was also engineered by overlap extension PCR as described above. Two engineered enzymes with different leader peptide length were created: one with the full length leader

peptide (long leader peptide) and one with a leader peptide that ends at a putative double glycine motif (short leader peptide) (Fig. 2.2). These constructs were also cloned into pQE80L for expression.

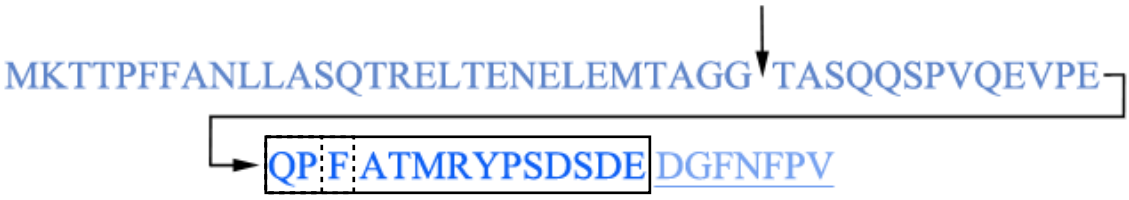


Figure 2.2. Marinostatin precursor peptide. The leader peptide is shown in the upper line and the core peptide and follower peptide in the lower line. A vertical arrow marks where the "short leader peptide" ends. The rectangles in the core peptide show the different amino acid compositions of marinostatin depending on the proteolysis it undergoes.

2.2.8. *Cyanothece* ATP-grasp ligase gene assembly

The approach used to clone and engineer the MvdD-like ligase from *Cyanothece* was PCR assembly. The online tool Primerize (https://primerize.stanford.edu, Stanford University, USA) was used to design the primers for the assembly. The sequence of the enzyme containing the N-terminal leader peptide was divided in two and these sequences were used as input for the tool. A list of the primers and their recommended annealing temperatures were obtained for each of the sequences. The last primer of the first sequence and the first primer of the second sequence were edited manually to join both sequences. In total 36 primers were utilized and the suggested protocol with primer sub-pools from Primerize was followed. In sum, instead of assembling all the primers in one PCR, the primers were divided in sub-pools and assembled separately. In some cases a nested PCR was required to generate enough DNA from a sub-pool. Three primer sub-pools were assembled: Pool 1.- P1-P8, Pool 2.- P7-P24, and Pool 3.- P23-36. Pool 2 and pool 3 were then assembled together, and the product from this PCR was finally assembled with pool 1. The final construct was cloned first into pDrive for conservation, and then into pQE80L for expression.

2.2.9. Expression of the engineered ATP-grasp ligases and MvdB

The strain *E.* coli BL21 (DE3) was also used to express these proteins, and the same optimized expression conditions and purification process (including gel filtration) for MvdC and MvdD were used to express and purify the engineered ATP-grasp ligases and the N-acetyl transferase from *P. agardhii*, and the engineered ATP-grasp ligases from *A. sagamiensis.* In the case of the ATP-grasp ligase from *Cyanothece sp*. PCC 7822 the expression conditions were 16ºC for 3 days.

2.2.10.Cyclization assays and one-pot reactions

The reactions were set up following a previously described protocol [110]: 100 mM Tris-HCl pH 8.0, 2.5 mM ATP, 5 mM $MgCl_2$, 25 mM KCl, 25 µg MvdD/LP-MvdD (in the case of tricyclization also 25 µg MvdC/LP-MvdC), and 80 µg of precursor peptide (in the case of *trans* activation, also 80 µg of the leader peptide). After 16 hours of incubation at 37ºC the reactions were quenched with 2 µL 0.5 M EDTA pH 8.0 and 200 µL methanol 100%, and evaporated in vacuum. In the first attempts to establish a one-pot reaction 10 µg of MvdB and AcCoA (final concentration 1 mM) were also added to the reaction mix. The one-pot reactions were then optimized and carried out in two steps: after the initial incubation with the ATP-grasp ligases for 16 hours MvdB and AcCoA were added to the mix and further incubated at 37ºC for 6 hours, followed by quenching with EDTA and methanol, and evaporation.

2.2.11.HPLC analysis of the reactions

The samples were redissolved in methanol 80% and filtered through a 0.45 µm membrane (Acrodisc 4 mm syringe filter, Pall, Germany) prior to injection into the HPLC (Prominence UFLC XR HPLC, Shimadzu, Germany). The column used was a Symmetry Shield RP18, particle size 3.5 µM, 4.6 x 100 mm (Waters, Germany), and the following linear gradient was used: aq acetonitrile (20%)+TFA (0.05%) to aq acetonitrile (38%)+TFA (0.05%) over 20 min at 1.0 mL min^{-1} flow rate. Fractions containing candidate peaks were collected and either evaporated in vacuum for posterior mass spectrometry analysis or lyophilized for quantification.

2.2.12.ESI-Q-TOF analysis of the cyclization of the MvdE precursor

The samples of the first cyclization trial, i.e. cyclization of the full-length MvdE precursor with MvdD and MvdC, were analyzed via ESI-Q-TOF (Bruker maXis II mass spectrometer) by Dr. Ines Starke from the department of Analytical Chemistry of the University of Potsdam.

2.2.13.MALDI-TOF analysis

Dry fractionated samples were resuspended in 10 µL TFA 0.1%. 0.3 µL of the sample solution was spotted with 0.3 µL HCCA matrix working solution and analyzed with a Bruker microflex LRF equipped with a nitrogen laser (λ= 337 nm). The analysis was done in positive-ion reflectron mode (accelerating voltage of 19 kV) and in general approx. 1000 laser

shots were averaged to generate a spectrum. The spectra were analyzed with the open source software mMass [111].

2.2.14.LC-MS analysis of the pilot one-pot reaction

To verify the PTM of the core region of MvdE by the established *in vitro* platform, samples of the reactions were analyzed via LC-MS in the department of Biochemistry of the Technical University of Berlin.

2.2.15.ESI-HR-MS/MS and MALDI-TOF/TOF studies

To confirm the position of the rings in the synthesized microviridins, samples of bicyclic Mdn K, and bicyclic and tricyclic Mdn B, as well as bicyclic Mdn Cth2 and tricyclic Mdn Cth3 were analyzed via ESI-HR-MS/MS. A methanolic extract of *Cyanothece sp*. PCC 7822 was analyzed via MALDI-TOF/TOF to determine whether this species produces microviridins *in vivo* and if those microviridins are identical to the microviridins produced *in vitro*. These measurements were conducted by Dr. Daniel Petras and Vincent Wiebach from the group of Prof. Roderich Süssmuth at the Technical University of Berlin.

2.2.16.Protease inhibition assays

2000 µM sample solutions were prepared for the inhibition assays by weighing the lyophilized samples separated by HPLC (see section 2.2.11) in an analytical balance (MX5 microbalance, Mettler Toledo) and redissolving them in the appropriate volume of ddH_2O. The assays were based on a previously described protocol [106] and were optimized for this work as follows: In a 96-well microtiter plate 30 µL of the enzyme working solution and 60 µL of the buffer (see table 2.3) were pipetted into a series of 6 wells (Fig. 2.3). 10 µL of a 2000 µM microviridin solution was then added to well 1 and serially diluted until well 5; well 6 was used as a control. After completing the serial dilution of the microviridin solution, 10 µL was subtracted from well 5 to maintain an equal volume in all wells. The plates were then incubated for 5 min at room temperature, and subsequently 90 µL of the substrate working solution (see table 2.3) was added to each well (180 µL final volume). The enzymatic activity was measured spectrophotometrically (Varioskan Flash, Thermo Scientific, USA) at 410 nm after incubation at 37ºC for 15 min (elastase assay) or 30 min (trypsin and subtilisin assays). The assays were done twice for each sample, and the obtained values were plotted in Origin to determine the IC_{50} of each microviridin variant. In the case of the Mdn K and Mdn B

libraries a pre-screen against trypsin and subtilisin was done with the bicyclic form of the peptides. When an active variant was found, the tricyclic form and the acetylated tricyclic form were synthesized and used in the protease inhibition assays to compare their IC_{50}. In the case of the *Cyanothece* variants the amino acids at position 5 were either Lys, Arg or Leu, thus inhibitory activity against trypsin or leucine was expected.

Table 2.3. Working solutions of the reagents used for the protease inhibition assays

Protease working solution	Substrate working solution	Buffer
Elastase [1.5 U mL^{-1}]	N-Succinyl-Ala-Ala-Ala-*p*-nitroanilide (Suc-AAA-*p*NA) [1 mg mL^{-1}]	50 mM Tris-HCl, 20 mM $CaCl_2$, pH 7.5
Trypsin [100 µg mL^{-1}] (approx. [150 U mL^{-1}])	*N*α-Benzoyl-DL-arginine 4-nitroanilide hydrochloride (BAPNA) [2.17 mg mL^{-1}]	50 mM Tris-HCl, 20 mM $CaCl_2$, pH 7.5
Subtilisin [100 µg mL^{-1}] (approx. [1.05 U mL^{-1}])	N-Succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-*p*NA) [0.05 mg mL^{-1}]	PBS

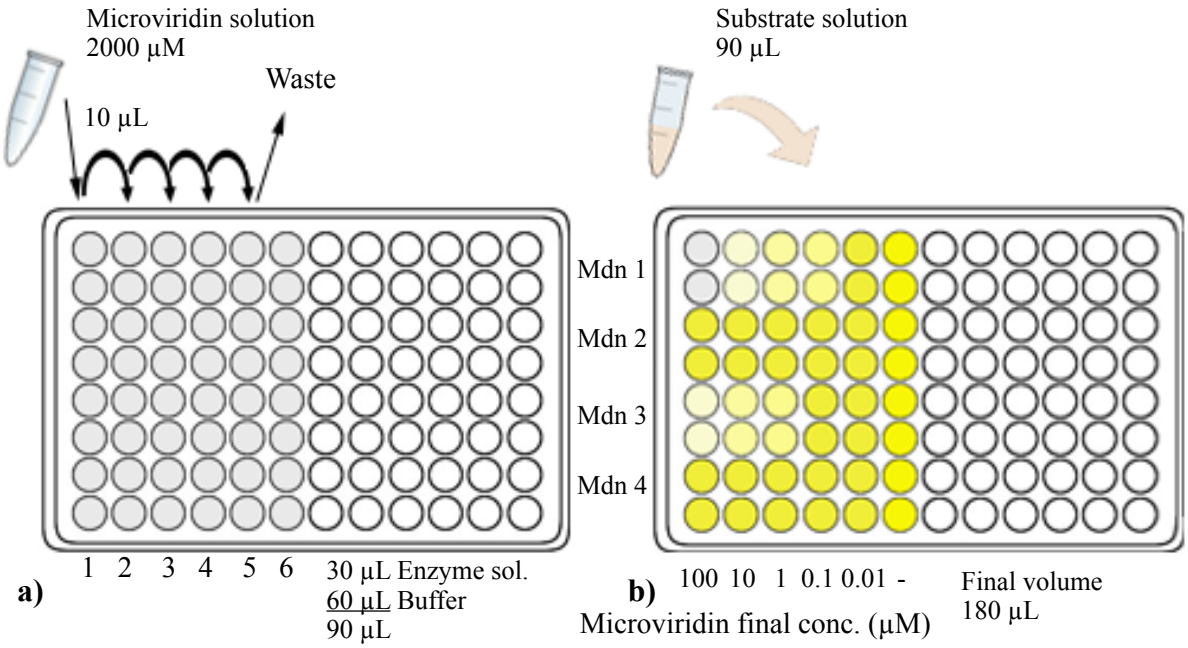


Figure 2.3. Scheme of a protease inhibition assay. **a)** 30 µL of the enzyme solution and 60 µL of the correspondent buffer were pipetted into each well. A serial dilution of the microviridin solution is carried out from well 1 to 5 and the plate is incubated for 5 min at room temperature. Well 6 is the control. **b)** 90 µL of the substrate solution is added to each well and after incubation at 37ºC the plate is read. The assay is done twice per microviridin variant.

3. Results

3.1. Enzyme preparation optimization and establishment of the *in vitro* platform

3.1.1. Optimization of the expression of MvdC and MvdD

Work on the expression of the ATP-grasp ligases from *P. agardhii* had been previously carried out in the group of Prof. Dittmann, however no active MvdC could be obtained [110]. Thus, the first task of this study was to achieve overexpression of functional ligases, not only because this was a prerequisite for the efficient expression of engineered variants but also to prepare high concentration solutions for crystallography. To achieve this goal, the vector pQE80L was chosen as expression vector and conditions with different temperatures and incubation times were tested. In a first attempt, protein expression was done at 37ºC for 5 h. It is worth noting that the bulk of the protein produced with these conditions was found in the insoluble fraction. For this reason conditions with lower temperatures had to be tested: 30ºC for 5 hours, 16ºC for 16 hours, and room temperature for 3 days. A clear improvement in the expression of the enzymes was observed when incubating the cultures at room temperature for 3 days (Fig. 3.1). A typical enzyme yield would be of approx. 12 µg enzyme/mL culture for MvdD, and 9 µg enzyme/mL culture for MvdC. Although these results were still not perfect, they were good enough to set these expression conditions as standard for the rest of the work.

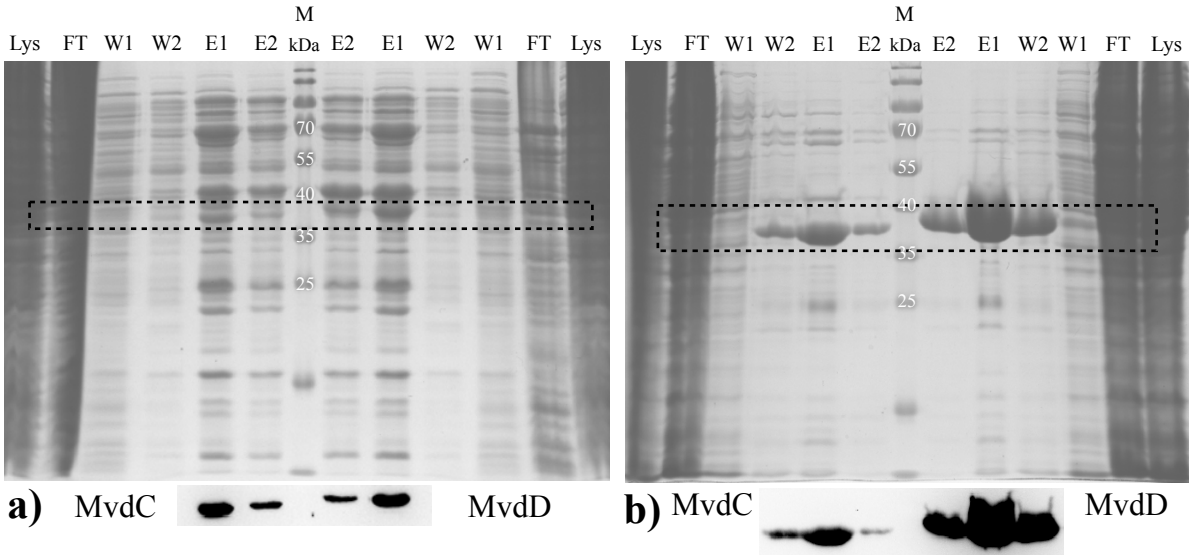


Figure 3.1. Comparison of the expression of the ATP-grasp ligases from *P. agardhii* with two different conditions. SDS-PAGE gel and western blot (anti-His tag antibody) of the proteins expressed at **a)** 16ºC for 16 hours after induction with IPTG. **b)** Room temperature for 3 days after induction with IPTG. The samples corresponding to MvdC (MW: 36.8 kDa) can be seen on the left part of the gels and the samples corresponding to MvdD (MW: 37.5 kDa) on the right part. M: Protein ladder; Lys: Lysate; FT: Flow-through; W1-W2: Wash fractions; E1-E2: Elution fractions.

3.1.2. Activity tests of MvdC and MvdD

To test whether the expression conditions used in combination with the pQE80L vector produced active ligases, *in vitro* cyclization assays with the full-length MvdE precursor were carried out. After analyzing the result of the reactions by HPLC and observing a shift in the retention time of the precursor peak (Fig. 3.2), the fractionated samples were analyzed in the department of Analytical Chemistry of the University of Potsdam by Dr. Starke via ESI-Q-TOF. The results of these measurements can be seen in figures 3.3-3.5. By using the m/z value of the lightest isotope of each spectrum to calculate the monoisotopic mass of the peptide the following masses were obtained: 5382.619 Da for MvdE, 5346.6024 Da for the product of the reaction between MvdE and MvdD, and 5328.5875 Da for the product of the reaction between MvdE and the two ATP-grasp ligases. By comparing these masses it was observed that the product of the reaction with MvdD had a mass 36 Da lighter than the linear core peptide MvdE, and the product of the reaction with both ligases had a mass 54 Da lighter than the linear core peptide MvdE. This corresponded to the expected 18 Da mass loss for each lactone or lactam ring introduced in the molecule, thus it was assumed that the expressed ATP-grasp ligases were active.

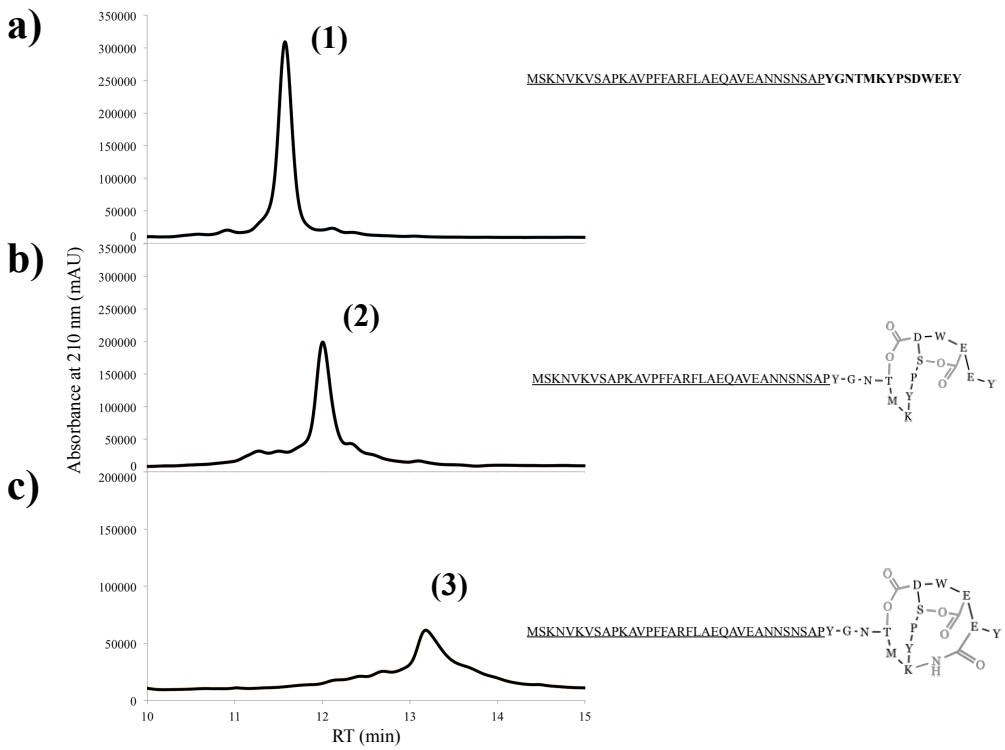


Figure 3.2. PTM of the precursor peptide MvdE. HPLC chromatograms of **a)** The precursor peptide MvdE (1); **b)** Bicyclic product (2) result of the incubation of MvdE with MvdD; **c)** Tricyclic product (3) result of the incubation of MvdE with MvdD and MvdC.

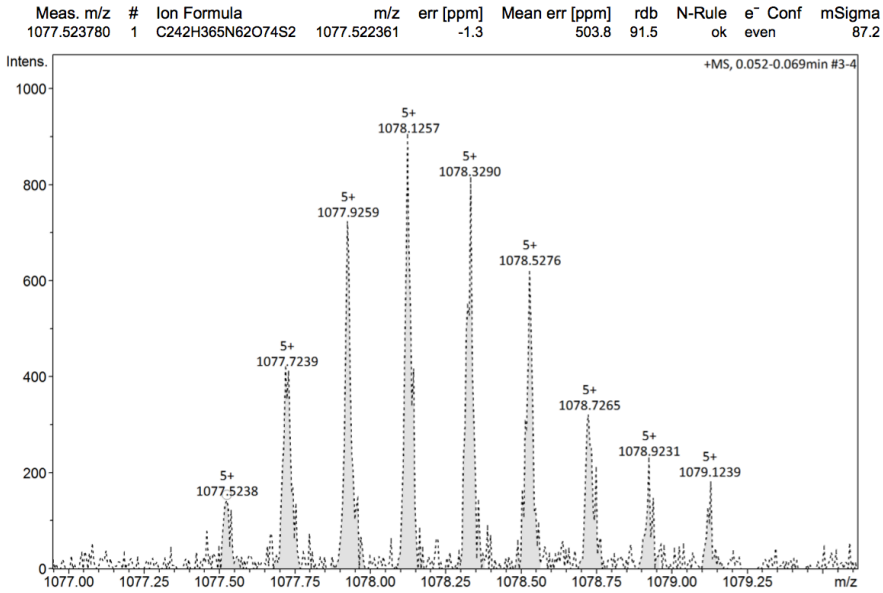
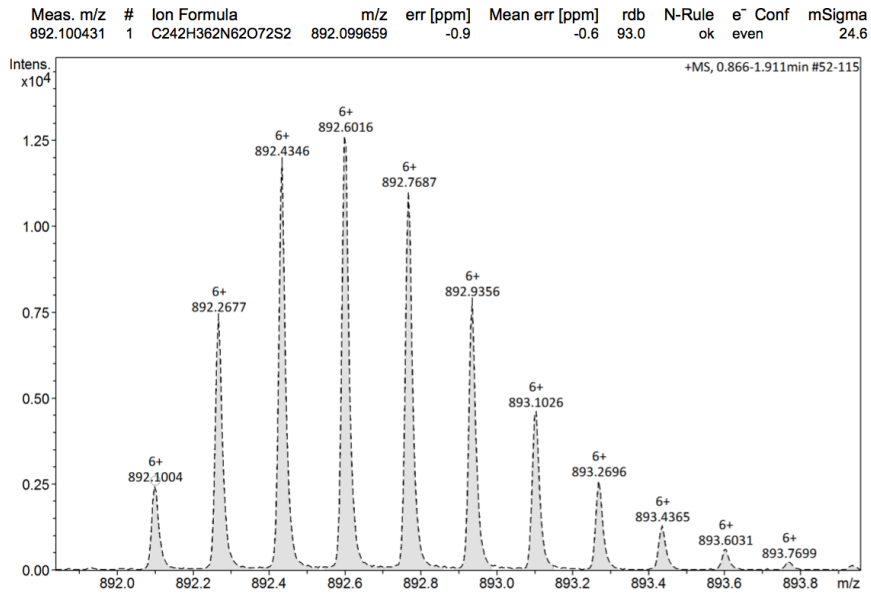


Figure 3.3. ESI-Q-TOF analysis of (1). The determined monoisotopic mass of this sample is 5382.619 Da. (See Fig. 3.2a.



F**igure 3.4. ESI-Q-TOF analysis of (2).** The determined monoisotopic mass of this sample is 5346.6024 Da. (See Fig. 3.2b.

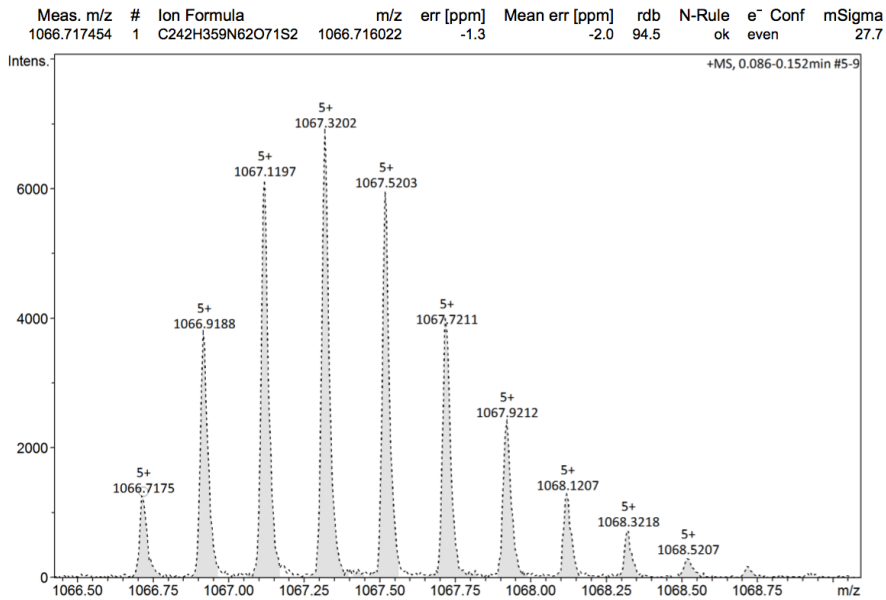


Figure 3.5. ESI-Q-TOF analysis of (3). The determined monoisotopic mass of this sample is 5328.5875 Da. (See Fig. 3.2c.

After corroborating that the enzymes were active, the decision was made to test if they were also capable of processing the MvdE core region when the leader region was added in *trans*, i.e. as a separate peptide. These reactions were analyzed first by HPLC, and the fractionated samples were then analyzed by MALDI-TOF. The results of these analyses showed that both ligases are able to cyclize the MvdE core peptide when the leader peptide is added in *trans*, as depicted in figure 3.6.

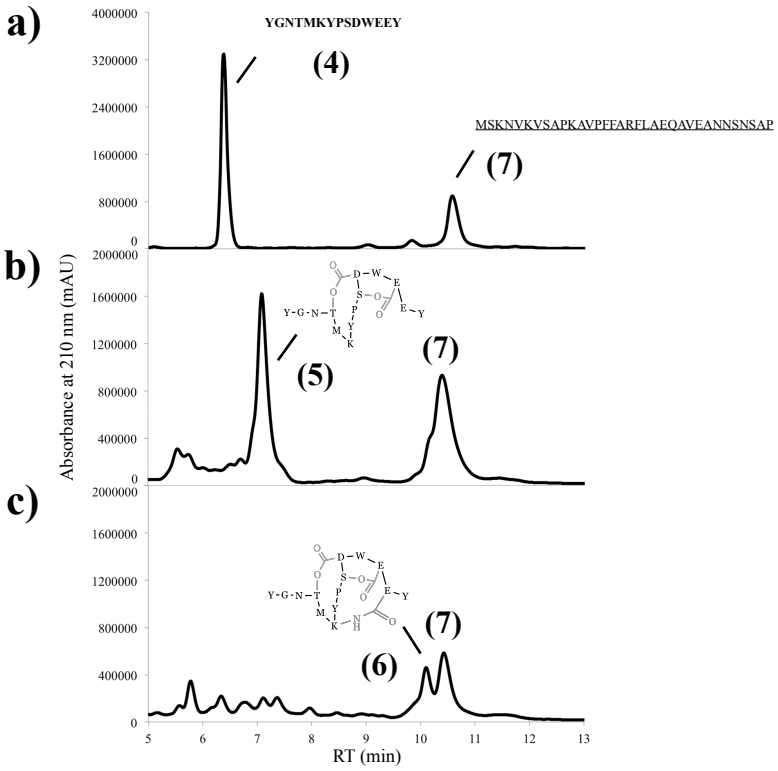


Fig. 3.6. Activation in *trans* of the microviridin PTM enzymes. HPLC chromatograms of **a)** The leader (7) and core (4) regions of MvdE; **b)** Bicyclic product (5) result of the incubation of the core peptide with MvdD in presence of the leader peptide; **c)** Tricyclic product (6) result of the incubation of the core peptide with MvdD and MvdC in presence of the leader peptide.

3.1.3. Expression of the engineered ATP-grasp ligases and MvdB

As mentioned in the introduction of this study, the modification of PTM enzymes from other RiPP families through the covalent attachment of their cognate leader peptide to their N-terminus successfully rendered constitutively active enzymes capable of processing cognate precursor peptides without leader region. Therefore, one of the main goals of the present work was to emulate this achievement with the microviridin biosynthetic machinery (Fig. 3.7). After creating the envisioned constructs, the engineered enzymes were expressed using the same optimized conditions as for the natural ATP-grasp ligases from *P. agardhii*. The microviridin tailoring enzyme MvdB was expressed with these conditions too (Fig. 3.8).

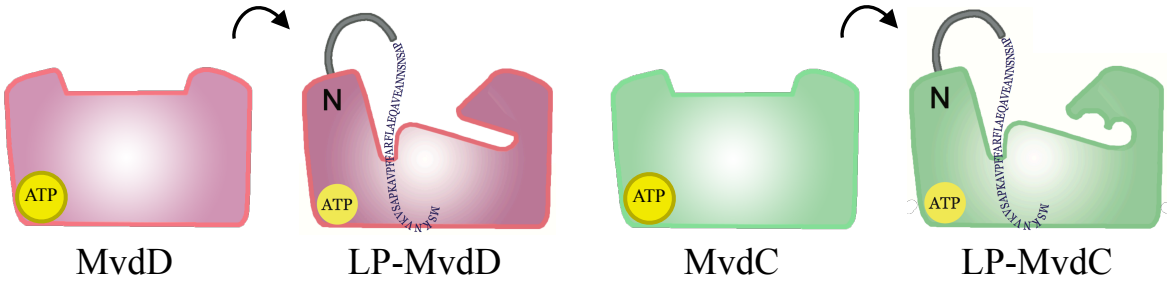


Figure 3.7. Schematic representation of the *cis* activation of the microviridin PTM enzymes. It was hypothesized that the attachment of the leader peptide via a linker to the N-terminus of the enzymes would activate them constitutively. LP: Leader peptide.

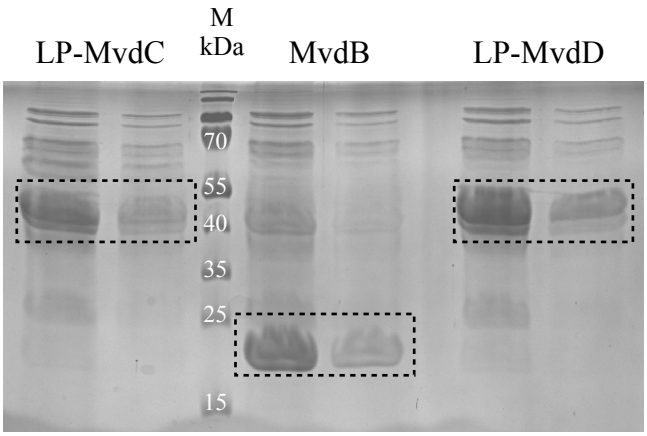


Figure 3.8. Expression of the engineered ATP-grasp ligases and N-acetyltransferase from *P. agardhii*. SDS-PAGE gel of the eluate fraction of the engineered ATP-grasp ligases (LP-MvdC MW: 43.2 kDa; LP-MvdD MW: 44 kDa) and N-acetyltransferase (MvdB MW: 19.6 kDa) from *P. agardhii*. M: Marker.

3.1.4. Gel filtration chromatography and crystallization studies

Gel filtration chromatography was an essential step in the purification of the engineered ligases for successful peptide cyclization, as demonstrated by assays carried out with enzymes that had only been purified through affinity chromatography where no peptide cyclization could be observed. Gel filtration was also necessary to prepare high purity solutions of the natural ATP-grasp ligases from *P. agardhii* for crystallization. Typical gel filtration chromatograms of the natural and engineered ligases from *P. agardhii* can be seen in figures 3.9-3.10. By comparing the chromatograms of the ligases to those of molecular weight markers, an apparent molecular weight of approx. 65 kDa and over 70 kDa can be assigned to the natural and engineered ligases, respectively. This is in agreement with the expected homodimeric structure of an ATP-grasp ligase [56]. Since the enzymes had been previously purified via affinity chromatography, few contaminants can be observed in the gel filtration chromatograms.

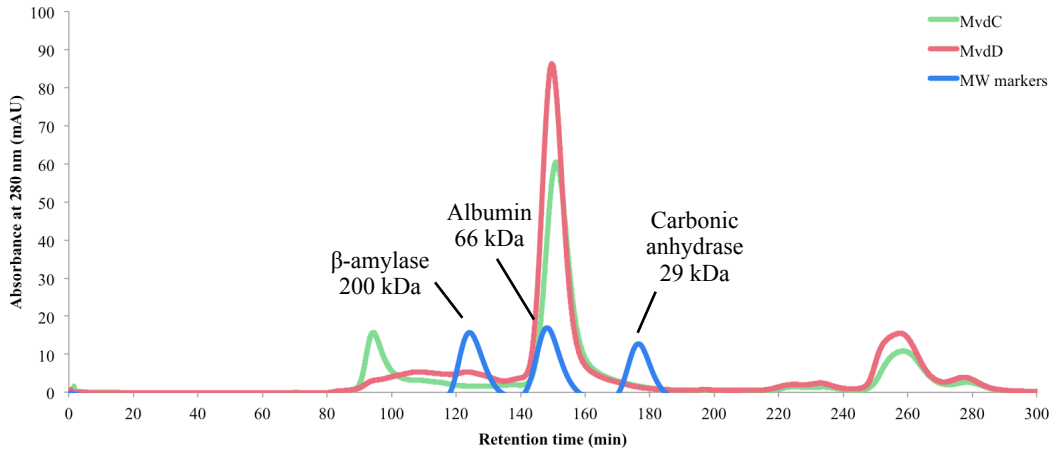


Figure 3.9. Gel filtration chromatography of MvdC and MvdD. In green the chromatogram correspondent to MvdC. In red the chromatogram correspondent to MvdD. In blue the chromatogram of MW markers. Both ATP-grasp ligases eluted around minute 150, similar to albumin. The buffer used was 150 mM NaCl, 20 mM Tris-HCl, pH 7.5.

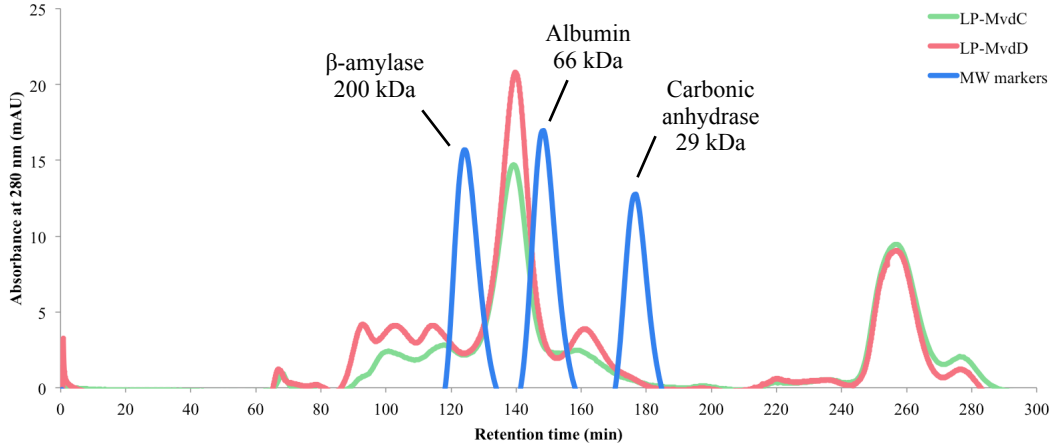


Figure 3.10. Gel filtration chromatography of LP-MvdC and LP-MvdD. In green the chromatogram correspondent to LP-MvdC. In red the chromatogram correspondent to LP-MvdD. In blue the chromatogram of MW markers. Both engineered ATP-grasp ligases eluted around minute 140. The buffer used was 150 mM NaCl, 20 mM Tris-HCl, pH 7.5.

For the crystallization studies highly concentrated and pure solutions of MvdC and MvdD were prepared (Fig. 3.11). Despite numerous crystallization attempts carried out by Prof. Groll's group at the Technical University of Munich and including the addition of $MgCl_2$, ATP, ATP analogues, the MvdE precursor, the leader and core peptides, and the expression of the enzymes using the SUMO vector for solubility enhancement, no crystals were obtained. The crystallization attempts were stopped after Li *et al.* published the solved structures of MdnB and MdnC [56].

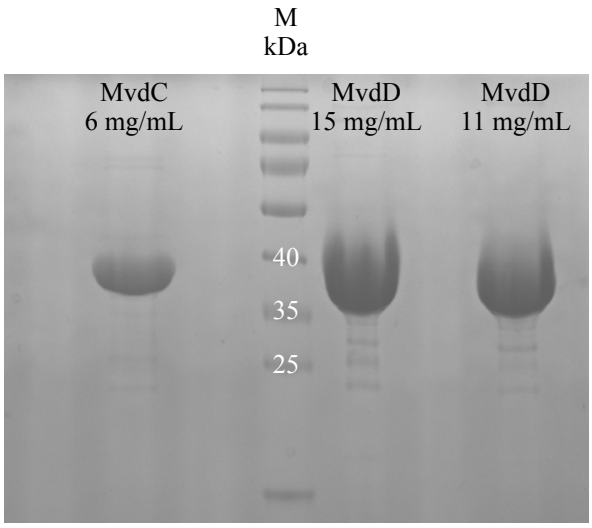


Figure 3.11. MvdC and MvdD solutions. Highly concentrated and pure solutions of the ligases from *P. agardhii* were used for crystallization studies.

3.1.5. Activity tests and development of the one-pot reaction

To determine if the engineered enzymes were functional, a first test with their cognate substrate, the core region of MvdE, was carried out. Three different assays with the core peptide of MvdE were prepared: One containing only LP-MvdD, one containing LP-MvdD and LP-MvdC, and one containing LP-MvdD, LP-MvdC and MvdB (one-step one-pot reaction). The reactions were analyzed via HPLC after being quenched with methanol and EDTA. Since the cyclized peptides are more hydrophobic than the linear precursors, a higher retention time was expected for each peptide with the introduction of each ring. Acetylation would in turn also increase the retention time of the peptide. For this reason when a peptide peak was observed to be shifted to a later retention time after incubation with the enzymes, it was considered that the reaction had been successful and the collected fractions were then analyzed via MALDI-TOF for verification. When the MALDI-TOF results showed a mass loss of 18 Da, 36 Da or 54 Da compared to the monoisotopic mass of the linear precursor, it was assumed that one, two or three cyclizations had occurred, respectively. When a mass increment of 42 Da was observed, it was assumed that the peptide had been acetylated. The MALDI-TOF spectra can be seen in the figure A1 in the appendix.

In the HPLC chromatogram of the one-step one-pot reaction three peaks with a higher retention time were witnessed (Fig. 3.12). MALDI-TOF analyses revealed that these peaks corresponded to an acetylated monocyclic peptide, an acetylated bicyclic peptide and the acetylated tricyclic peptide (mature microviridin). To increase the yield of the latter, the decision was made to split the one-pot reaction in two steps: An initial incubation period with only the ligases followed by the addition of AcCoA and the N-acetyltransferase.

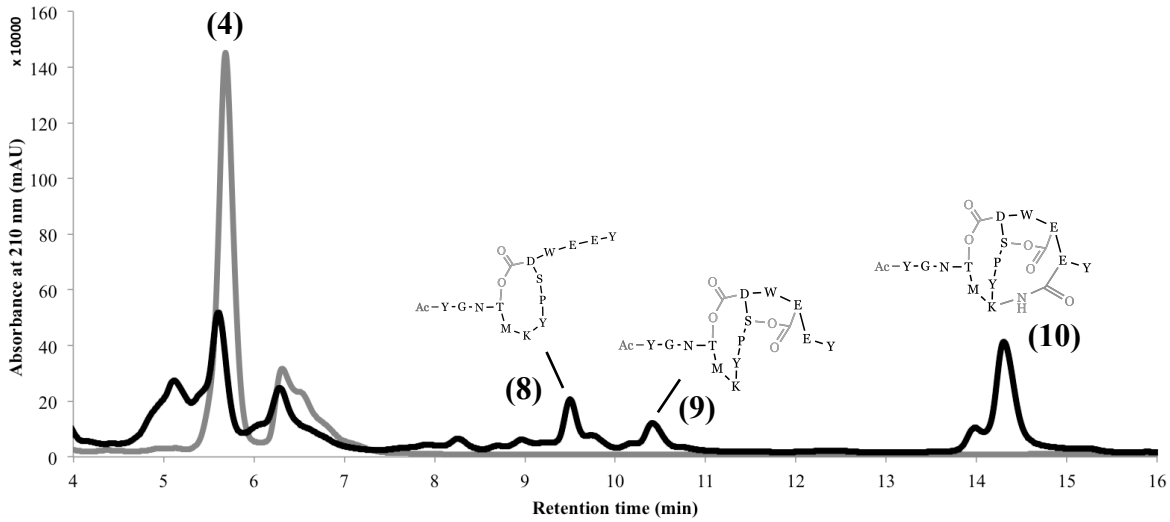


Figure 3.12. HPLC analysis of the one-step one-pot reaction of the synthesis of microviridin K. In the gray chromatogram the peak of the precursor core peptide MvdE (4) can be seen before minute 6. In the black chromatogram three peaks with a higher retention time can be observed, corresponding to two acetylated intermediates (8-9) and the mature Mdn K (10).

The cyclization assays were repeated including a two-step one-pot reaction. These reactions were then analyzed via LC-MS in the department of Biochemistry of the Technical University of Berlin. The results of these analyses can be observed in figure 3.13. For the rest of the work one-pot reactions were carried out in two steps.

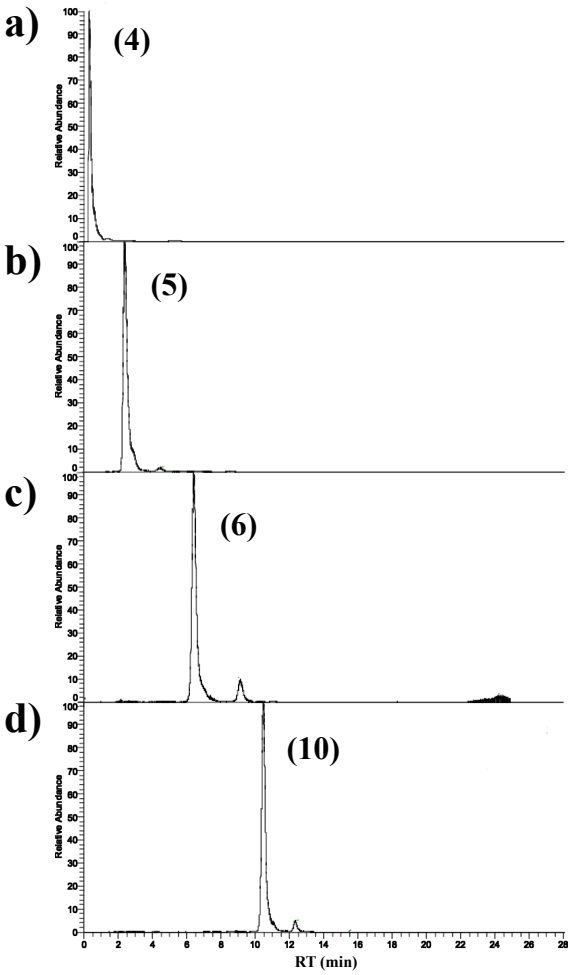


Figure 3.13. ***In vitro*** **maturation of the MvdE core peptide.** LC-MS chromatograms of **a)** The core region of MvdE (4); **b)** Bicyclic product (5) result of the incubation of the core peptide of MvdE with LP-MvdD; **c)** Tricyclic product (6) result of the incubation of the core peptide of MvdE with LP-MvdD and LP-MvdC; **d)** Mature Mdn K (10) result of the incubation of the core peptide of MvdE with LP-MvdD and LP-MvdC in a first step, followed by addition of MvdB and AcCoA (two-steps one-pot reaction).

3.2. Implementation of the *in vitro* platform for the synthesis of microviridins

3.2.1. Production of synthetic microviridin variants from peptide libraries

After demonstrating the functionality of the *in vitro* platform, its first application was to produce a microviridin library to screen for novel variants which showed enhanced protease inhibition. The Mdn K and Mdn B amino acid sequences, YGNTMKYPSDWEEY and FGTTLKYPSDWEEY respectively, were selected as backbones for the libraries since Mdn K is the variant produced by *P. agardhii,* and Mdn B, similar in sequence to Mdn K, had previously shown potent protease inhibition [39]. The focus of these libraries was the 5th amino acid position due to its pivotal role in protease inhibition. Thus the peptide libraries, synthesized via SPPS, were composed of the backbone of Mdn K or Mdn B with a substitution at the 5th position with all 20 proteinogenic amino acids. The workflow for the screening was planned as follows: First, bicyclic variants would be generated by incubation with LP-MvdD and their mass would be determined via MALDI-TOF. Then, the bicyclic variants would be screened for inhibitory activity against trypsin and subtilisin. Finally, for those active variants found, the tricyclic and acetylated tricyclic forms would be synthesized and their inhibitory activity would also be measured. An overview of this workflow can be seen in figure 3.14.

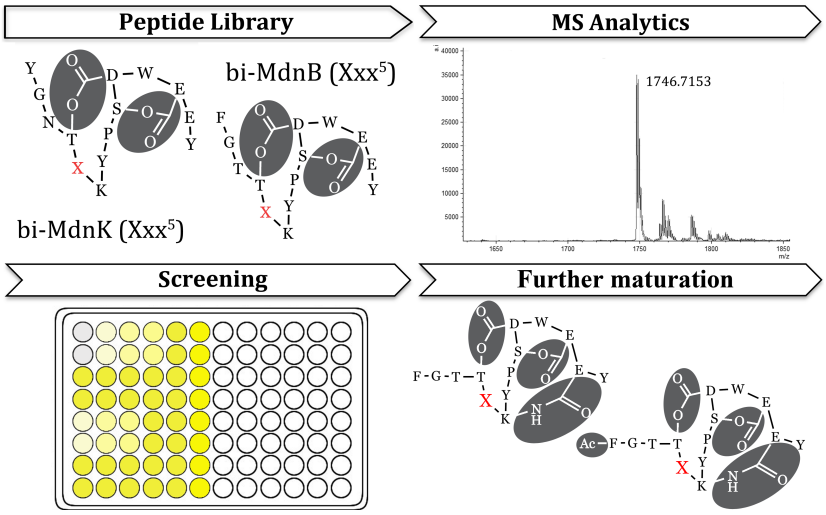


Figure 3.14. Overview of the production of synthetic microviridins with improved bioactivity. The Mdn K and Mdn B libraries were bicyclized and analyzed via MALDI-TOF. Then, the bicyclic variants were screened for bioactivity. Finally, the maturated forms of the active variants were synthesized and their bioactivity was measured too.

The great majority of the variants were successfully bicyclized, normally with yields >80%, however there were two noteworthy cases in both libraries where the cyclization was hindered: when the amino acid at position 5 was a Cys or a Pro. In the case of the Mdn K library, the variant with Thr could not be cyclized either (Fig. 3.15 and Table A2).

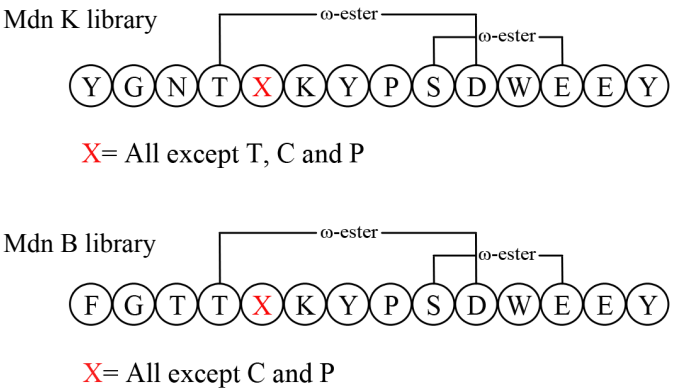
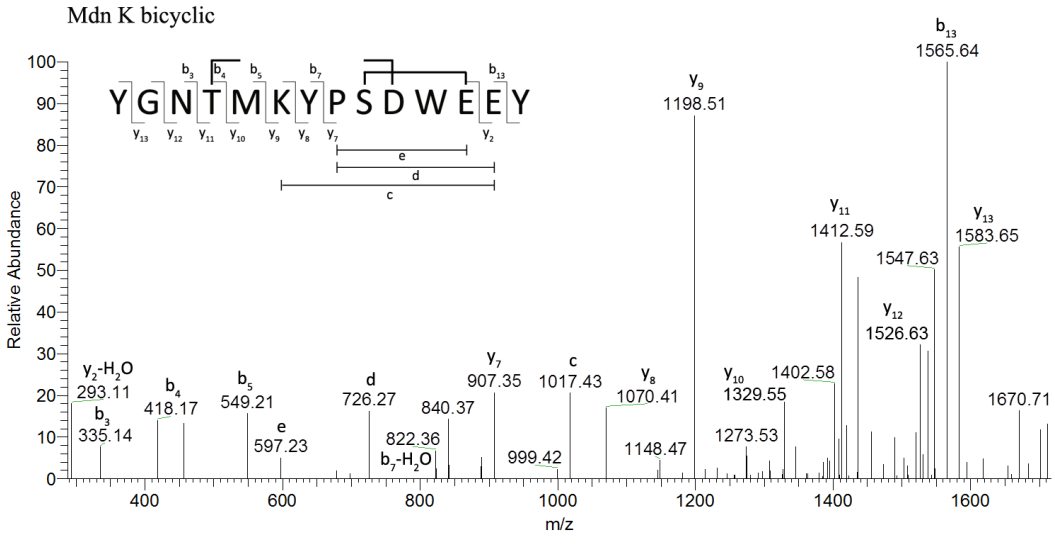


Figure 3.15. Summary of the bicyclization assays with the Mdn K and Mdn B libraries. The results of the bicyclization were similar for both libraries, except that for the Mdn K library, bicyclization of the variant carrying Thr at position 5 was not possible.

As mentioned before, the modification of the variants was verified by MALDI-TOF analysis, and the spectra of the active variants can be found in the appendix. Nevertheless, MALDI-TOF analysis only renders the mass of the modified peptide. Although the 18 Da mass difference corresponds to the loss of a water molecule, this does not necessarily demonstrate the formation of a ring, nor does it show its position in the peptide. For this reason further analysis was required to prove the formation of the lactone and lactam rings, and their positions in the peptide. Tandem mass spectrometry was chosen for this purpose over NMR because of the small sample amounts needed for the measurements. Samples of Mdn K and Mdn B were analyzed via ESI-HR-MS/MS since it was rationalized that these variants, as lead compounds of the libraries, would serve as proof-of-principle for the correct cyclization of the libraries. A summary of the results of these measurements can be seen in figure 3.16, demonstrating the expected lactonization and lactamization of the peptides.



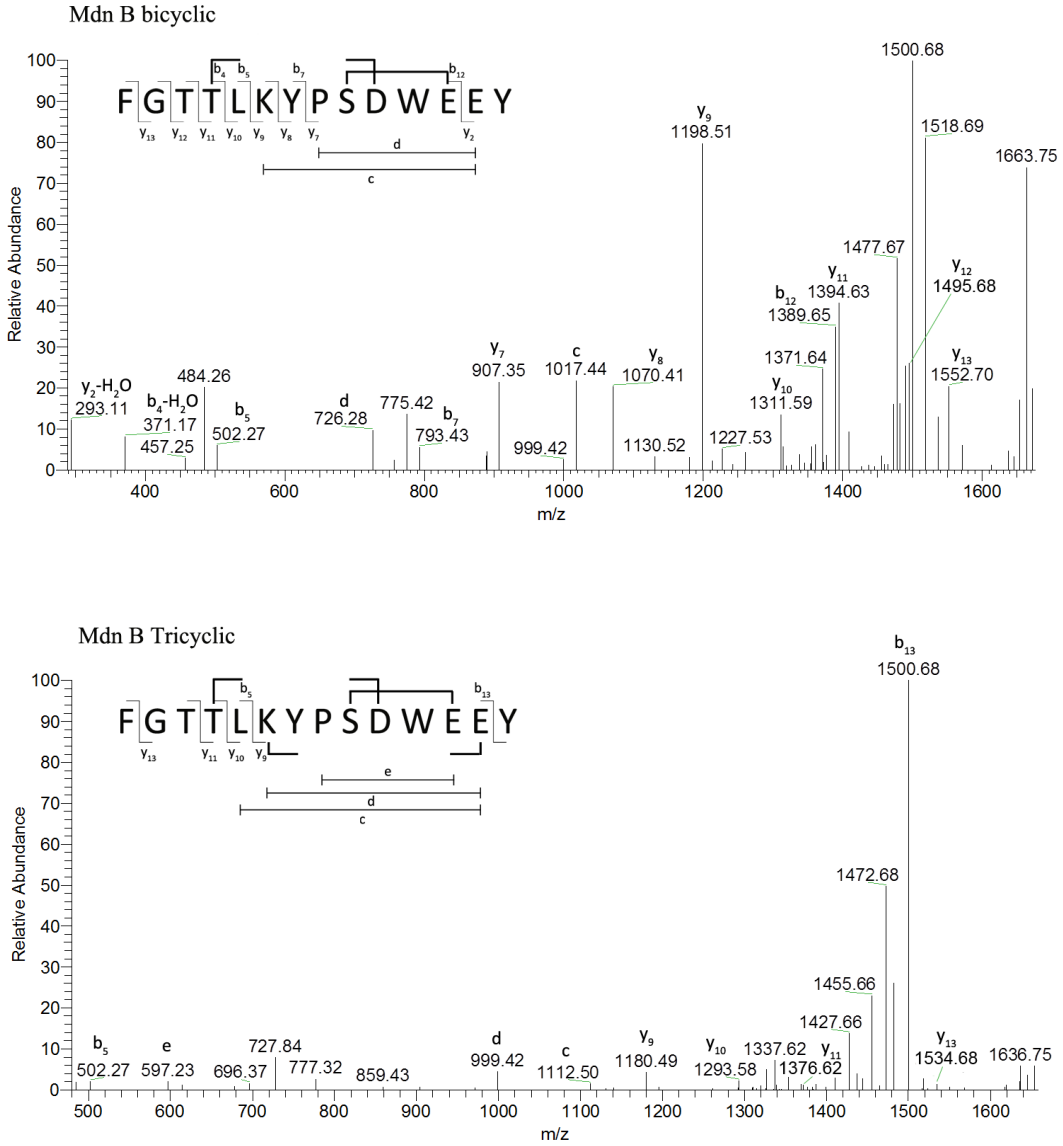


Figure 3.16. Tandem mass spectrometry analysis of Mdn K and Mdn B. A scheme of each variant depicts the ions found on each deconvoluted MS/MS spectrum.

The active variants found in the screening against the representative serine proteases trypsin and subtilisin can be found in table 3.1. In the case of M5F and L5D the amount of bicyclic peptide generated was not enough to use in the assays, therefore these variants were not tested. In the case of L5Y, only the bicyclic variant could be tested due to an scarce supply of the synthetic peptide. All the active variants could be tricyclized and acetylated with satisfactory yields.

Table 3.1. Results of the protease inhibition assays for the Mdn K and Mdn B libraries. ND: Not determined.

Microviridin	Trypsin IC_{50} µM	Subtilisin IC_{50} µM	Microviridin	Trypsin IC_{50} µM	Subtilisin IC_{50} µM
Mdn K library: YGNTMKYPSDWEEY			**Mdn B library:** FGTTLKYPSDWEEY		
M5M Bicyclic	Inactive (>100)	**8.3 ± 0.06**	L5R Bicyclic	**0.76 ± 0.02**	Inactive (>100)
M5M Tricyclic	ND	**3.9 ± 0.04**	L5R Tricyclic	**0.09 ± 0.002**	ND
M5M Ac-tricyclic	ND	**5.2 ± 0.32**	L5R Ac-tricyclic	**0.05 ± 0.01**	ND
M5R Bicyclic	**1.78 ± 0.23**	Inactive (>100)	L5Q Bicyclic	Inactive (>100)	**26.7 ± 2.6**
M5R Tricyclic	**0.41 ± 0.05**	ND	L5Q Tricyclic	ND	**10.4 ± 1.1**
M5R Ac-tricyclic	**0.20 ± 0.02**	ND	L5Q Ac-tricyclic	ND	**4.8 ± 0.39**
M5K Bicyclic	**3.13 ±0.10**	Inactive (>100)	L5M Bicyclic	Inactive (>100)	**9.2 ± 1.4**
M5K Tricyclic	**0.48 ± 0.11**	ND	L5M Tricyclic	ND	**4.3 ± 0.14**
M5K Ac-tricyclic	**0.17 ± 0.01**	ND	L5M Ac-tricyclic	ND	**4.8 ± 0.14**
			L5K Bicyclic	**0.22 ± 0.05**	Inactive (>100)
			L5K Tricyclic	**0.36 ± 0.01**	ND
			L5K Ac-tricyclic	**0.35 ± 0.03**	ND
			L5Y Bicyclic	Inactive (>100)	**38.2 ± 0.20**

3.2.2. Synthesis of genome-mined microviridins

A second application for the designed platform was to synthesize bioinformatically predicted microviridins. Microviridin BGCs were mined from the database at NCBI by the group of Prof. David Fewer at the University of Helsinki. The BGCs found belonged to the cyanobacteria, bacteroidetes, and proteobacteria phyla, with the majority encoding the ω-ester and ω-amide ATP-grasp ligases. The microviridin precursor peptides could be divided in three main classes according to their structure (Fig. 3.17). Precursor peptides belonging to the first class contain a leader region per core region. The marinostatin precursor peptide, which was also assigned to the first class, possesses a double glycine motif and short sequences N- and C-terminal to the microviridin motif which are not present in the final product (Fig. 2.2 and

3.17). Class II precursor peptides exhibit a single leader region for up to five core regions. The core regions can be consecutive or separated by double glycine motifs. Finally, the precursor peptides of the third class show a single leader region separated by a double glycine motif from a long region which contains the conserved microviridin motif at the C-terminus.

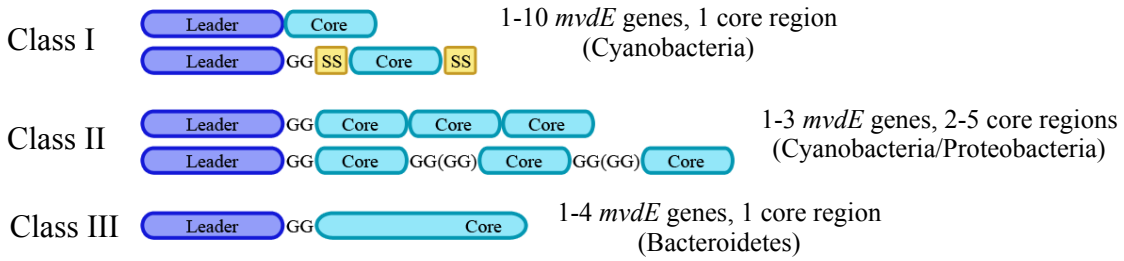


Figure 3.17. Classification of the microviridin precursor peptides. Three major classes based on structural features were assigned to the mined microviridin precursor peptides. Noteworthy, class I precursors are characteristic of cyanobacteria, precursors from class II can be found in cyanobacteria and proteobacteria, and class III precursors are largely restricted to the bacteroidetes. GG: double glycine motif; SS: short sequence stretch.

The microviridin BGC of the species *Cyanothece* sp. PCC 7822, which belongs to the class I, is an exceptional example since it codifies for ten precursor peptides (Fig. 3.18). The *Cyanothece* sp. core peptides varied between 12-14 amino acids and carried either Leu, Arg or Lys at the key position 5, prompting the hypothesis that this species produces a cocktail of microviridins with protease inhibitory activity (Table 3.2). For this reason, five representative core peptides were selected out of the ten encoded in the genome of *Cyanothece* sp. PCC 7822 for a third peptide library (Table 3.2). Another member included in this library was the core peptide of the microviridin-related compound marinostatin. The BGC of marinostatin, which is found in the genome of the species *Algicola sagamiensis* (previously known as *Alteromonas sagamiensis*), does not encode an ω-amide ligase, hence the mature product is a twelve amino acids-long peptide that carries only two lactone rings [112]. The last member of this library was the second precursor peptide encoded in the BGC of *P. agardhii* NIVA-CYA 126, MvdF.

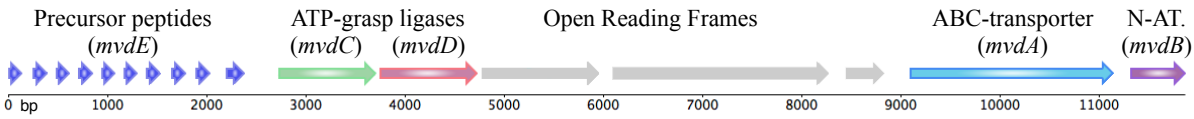


Figure 3.18. Microviridin BGC of *Cyanothece* sp. PCC 7822. Ten microviridin precursor peptides are encoded in the genome of this species. Three additional open reading frames of unknown function are found in the BGC. The homologous genes of *P. agardhii* are indicated in parenthesis. N-AT.: N-acetyltransferase.

Table 3.2. Bioinformatically predicted microviridin core peptides in the genome of *Cyanothece* sp. PCC 7822. The variants selected for the library (black rectangle) exhibit different length and amino acids at position 5 (underlined).

Putative microviridins		Putative microviridins	
Cth 1	YQNTLKYPSDWEEY	Cth 6	-YVTKKYPSDWEDY
Cth 2	-TVTRKYPSDWEDY	Cth 7	YQNTLKYPSDWEDY
Cth 3	-YVTKKYPSDWEEY	Cth 8	YQNTLKYPSDWEDY
Cth 4	--ATLKYPSDWEEY	Cth 9	-YVTKKYPSDWEEY
Cth 5	--FTLKYPSDWEDE	Cth 10	-IFTLKFPSDWEDS

A schematic representation of the results of the PTM assays with this library can be seen in figure 3.19. and in more detail in table A3. Of particular note is that Cth 1, being similar in sequence to the core region of MvdE, could only be bicyclized with satisfactory yields. It is also surprising that the core peptide of MvdF could not be processed by its cognate biosynthetic machinery. Finally, the core peptides that are twelve amino acids long were poorly modified or not modified at all.

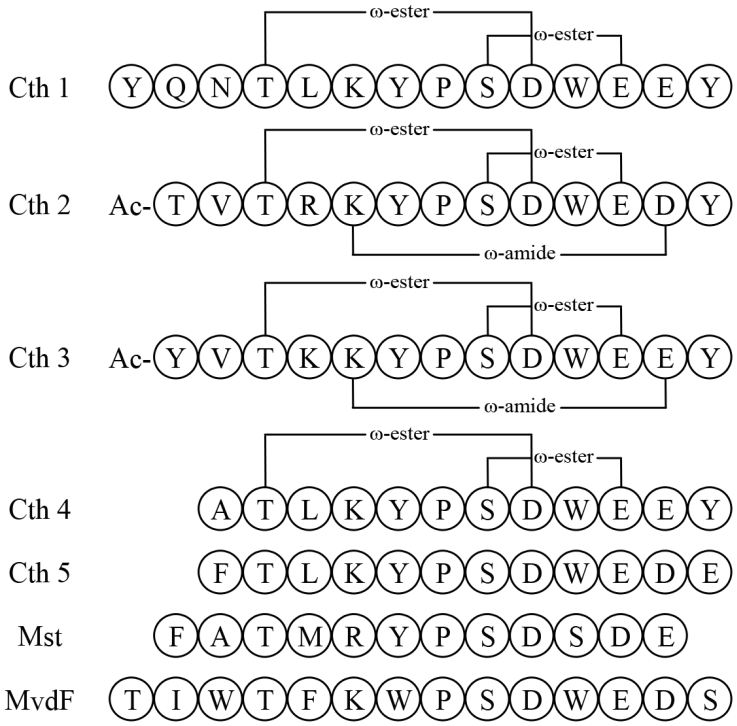


Figure 3.19. Results of the PTM assays with the genome-mined microviridins. While Cth 2 and Cth 3 could be fully maturated with satisfactory yields, Cth 1 and Cth 4 could only be bicyclized. Cth 5, Mst, and MvdF could not be processed at all.

A methanolic extract of a culture of *Cyanothece* sp. PCC 7822 was analyzed via MALDI-TOF/TOF to determine whether this species produces microviridins *in vivo* and if those microviridins are identical to the microviridins synthesized *in vitro*. Two *Cyanothece* sp. microviridin variants were then analyzed via ESI-HR-MS/MS to demonstrate that although the ATP-grasp ligases from *P. agardhii* are not their cognate enzymes, the precursor peptides from *Cyanothece* sp. were correctly processed by them. The summarized results of these analyses can be seen in Figure 3.20-3.21.

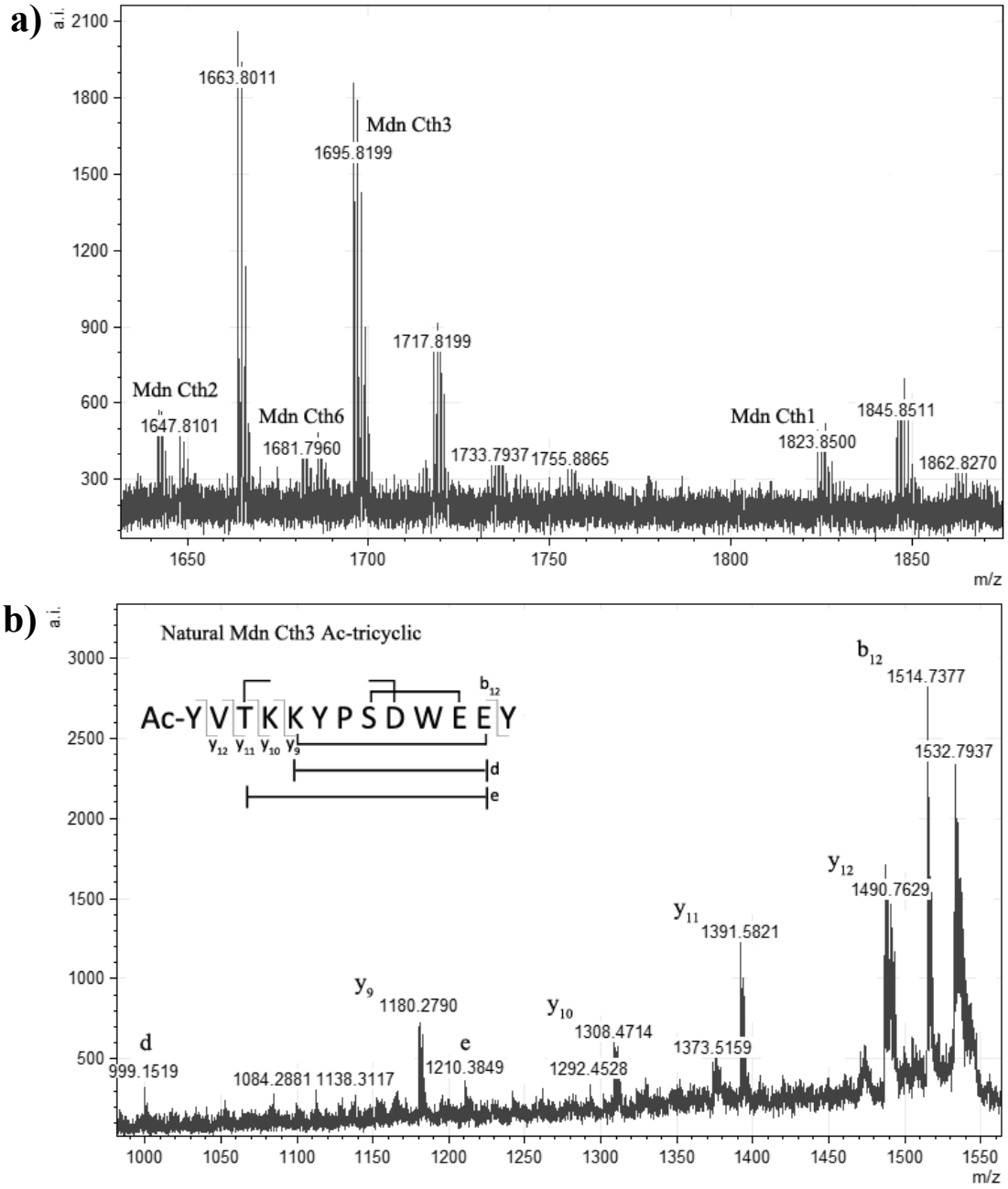


Figure 3.20. MALDI-TOF and MALDI-TOF/TOF analyses of the methanolic extract of *Cyanothece* sp. PCC 7822. a) MALDI-TOF scan of the methanolic extract. 4 masses depicted in the scan correspond to acetylated tricyclic microviridins encoded in the microviridin BGC of *Cyanothece* sp. PCC 7822. **b)** MALDI-TOF/TOF of the extract. The ion with a m/z of 1695.8199 was fragmented to corroborate that it represents microviridin Mdn Cth3. A scheme of this variant depicts the ions found on the deconvoluted MS/MS spectrum.

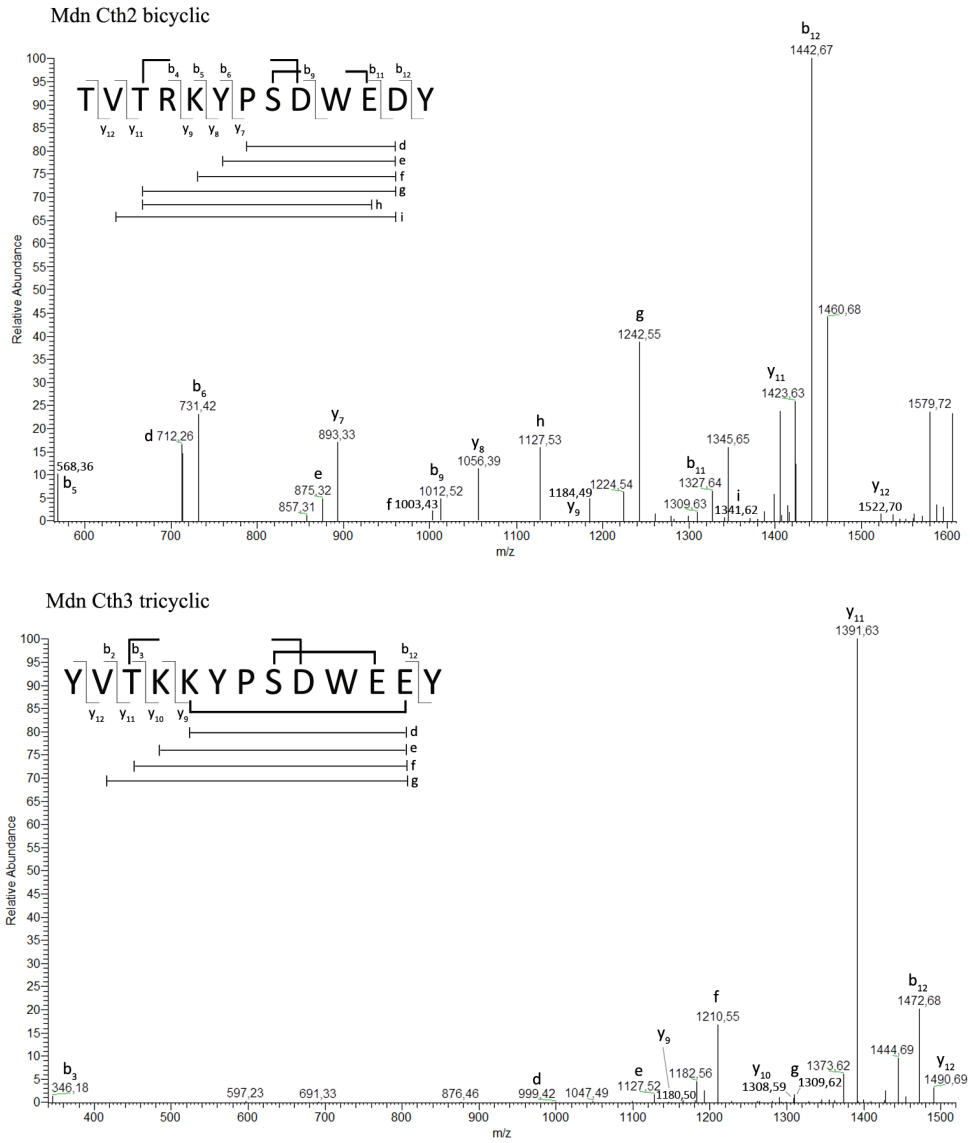


Figure 3.21. Tandem mass spectrometry analysis of the *Cyanothece* sp. microviridin variants Mdn Cth2 and Mdn Cth3. A scheme of each variant depicts the ions found on each deconvoluted MS/MS spectrum.

The bicyclic Cth 1 and Cth 4, and the acetylated tricyclic Cth 2 and Cth 3 were used for the inhibition assays of the microviridins from *Cyanothece* sp. (Fig. 3.22). It was expected that they were active against trypsin or elastase due to the Arg and Lys, or Leu at their 5th position, respectively. The bicyclic variants had a potency in the low micromolar range while the fully maturated variants were active in the low nanomolar range (Table 3.3).

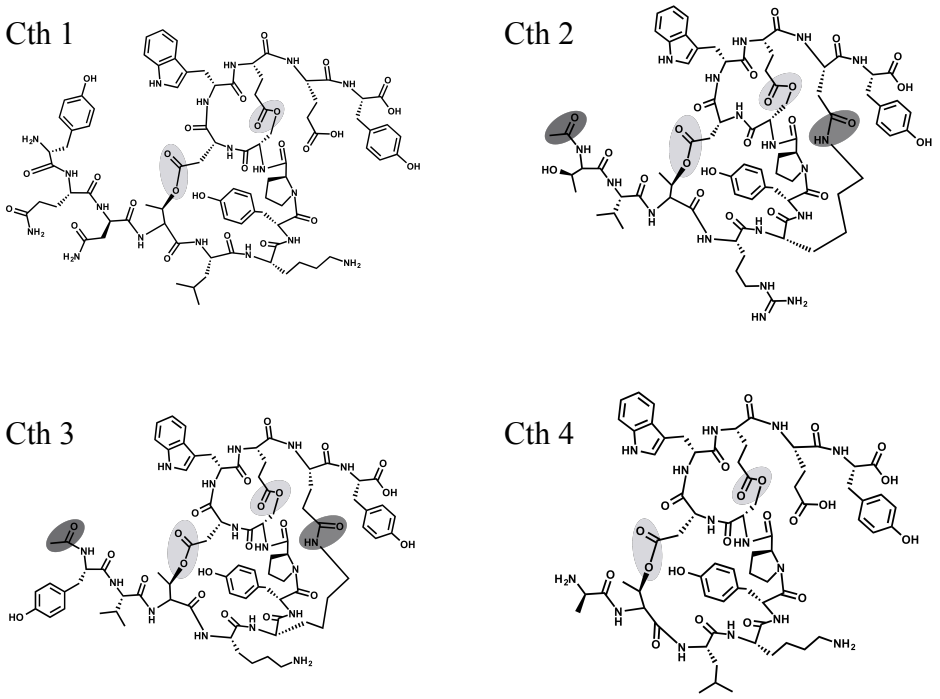


Figure 3.22. Structures of the *Cyanothece* sp. microviridins used for inhibition assays. The ω-ester and ω-amide bonds, as well as the acetyl group are marked in gray ellipses.

Table 3.3. Results of the protease inhibition assays with the *Cyanothece* sp. microviridins. ND: Not determined.

Microviridin	Trypsin IC_{50} nM	Elastase IC_{50} µM
Cth 1 Bicyclic	ND	**1.05 ± 0.20**
Cth 2 Ac-tricyclic	**28.4 ± 4.0**	ND
Cth 3 Ac-tricyclic	**21.5 ± 0.6**	ND
Cth 4 Bicyclic	ND	**21.85 ± 3.35**

After becoming aware of the potent inhibition that these variants had shown, it was agreed that the ω-ester ligases from *Cyanothece* sp. and *A. sagamiensis* would be cloned and engineered in an attempt to cyclize the shorter core peptides Cth 5 and Mst. The method used to engineer the ligase from *A. sagamiensis* was overlap extension PCR with two different leader peptide lengths, however for the ligase from *Cyanothece* sp. a gene assembly methodology was followed (see 2.2.7-.8). Upon sequencing the final *Cyanothece* sp. ligase construct, it was noted that the linker region was shorter than designed: Instead of coding for a 15x GlySer repetition the sequence coded for a 8x GlySer repetition. Nevertheless it was decided to use this construct based on the fact that Oman *et al.* showed that an engineered enzyme with a 10x GlySer repetition linker was constitutively activated too but to a lesser degree than that with a 15x GlySer repetition [68]. In the case of the ligase from

A. sagamiensis the expression conditions used were the same as for the ligases from *P. agardhii*, but for the ligase from *Cyanothece* sp. the expression temperature had to be lowered to 16ºC (Fig. 3.23).

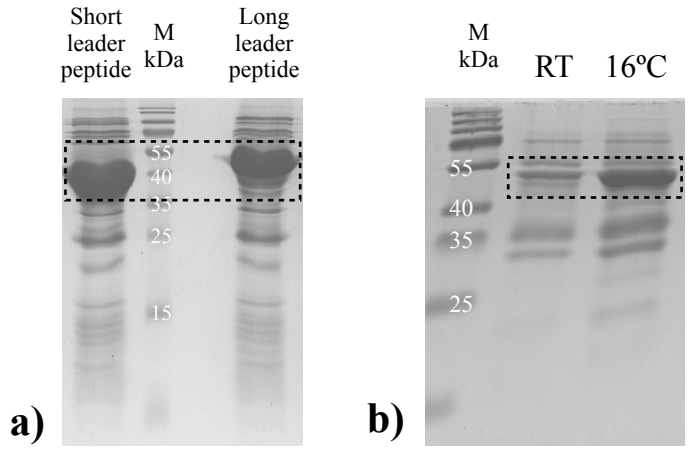


Figure 3.23. Expression of the engineered ATP-grasp ligases from *Cyanothece* sp. PCC 7822 and *A. sagamiensis*. SDS-PAGE gels of the eluate fraction of the engineered ATP-grasp ligases from **a)** *A. sagamiensis,* engineered with two different-length leader peptides (short leader peptide ligase MW: 43.5 kDa; long leader peptide ligase MW: 44.9 kDa); and **b)** *Cyanothece* sp. PCC 7822 (MW: 42 kDa) expressed at two different temperatures. M: Marker; RT: Room temperature.

The expressed engineered ligases from these two species were purified following the same protocol as for the ligases from *P. agardhii,* and they were utilized in cyclization assays with controls and the short core peptides Cth 5 and Mst. Unfortunately no modification could be observed for any of the core peptides used. Finally, it was decided against trying to optimize the assays due to the amount of time required.

3.2.3. Synthesis of microviridins harboring non-canonical amino acids

A further aim of this work was to determine whether the microviridin structure would support the introduction of ncAAs. Apart from the possibility of enhancing the inhibitory activity, microviridin scaffolds harboring ncAAs could be employed as a molecular tool for research, for example to identify binding partners of microviridin (pull-down assays) or to determine the intracellular localization of microviridin (fluorescent labeling). Biotinylated lysine was chosen for the first experiments with ncAAs because of its potential to be used in pull-down assays with avidin. For this library the backbone of Mdn J, ISTRKYPSDWEEW, was used since it had been demonstrated that this particular variant causes lethal molting disruption in microcrustaceans [100], making it interesting to use in a pull-down assay with extracts from these particular microcrustaceans. A biotinylated lysine was introduced either at the N- or

C-terminus of the peptide, or replacing the proteinogenic lysine at position 6. It was deemed rational that the production of bicyclic variants would suffice for pull-down assays based on the fact that bicyclic microviridins also exhibit bioactivity. The inhibitory activity of these microviridins was to be tested nevertheless to verify their functionality. The peptide with the Lys(biotin) in the middle of the molecule could not be cyclized at all, whereas the peptides with the Lys(biotin) at the N- or the C-terminus could be bicyclized or monocyclized, respectively (Fig. 3.24). The modifications were corroborated by MALDI-TOF analyses and the correspondent spectra can be found in figure A.3 in the appendix. Finally, both the mono- and the bicyclic variants showed inhibitory activity against trypsin in the low micromolar range, accounting for an specific interaction (Table 3.4).

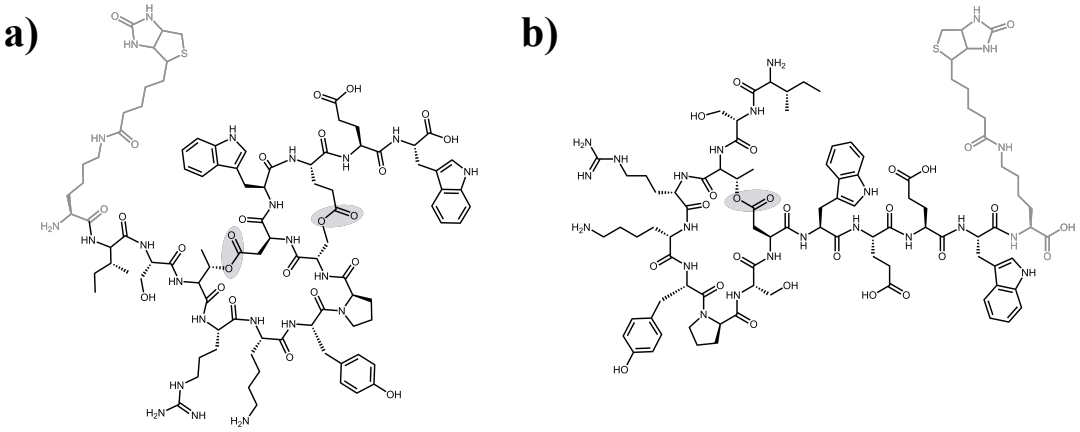


Figure 3.24. Biotinylated microviridins. a) Bicyclic structure of a Mdn J variant carrying an N-terminal biotinylated Lys (in gray). **b)** Monocyclic structure of a Mdn J variant carrying a C-terminal biotinylated Lys (in gray). The ω-ester bonds are marked in gray ellipses.

Table 3.4. Results of the trypsin inhibition assays with biotinylated microviridins.

Peptide	Trypsin IC_{50} µM
Bicyclic N-biotinylated Mdn J	**0.42 ± 0.09**
Monocyclic C-biotinylated Mdn J	**4.84 ± 0.24**

4. Discussion

Recent technological advances have enabled a better understanding of the PRPS. Knowledge on leader peptide-dependent activation of PTM enzymes has, in turn, permitted the creation of hundreds of synthetic RiPP analogs, some of them with improved bioactivities. The present study exploited this information to establish an efficient *in vitro* platform for the production of microviridin analogs. The utilization of this platform has provided more information regarding the substrate specificity of the microviridin PTM enzymes. Moreover, the modular nature of the microviridin biosynthetic pathway allowed for the production and isolation of intermediates that could be used to analyze SARs between microviridins with increasing level of PTMs and their target proteases. On the whole, the insights into RiPP synthesis gained through this work contribute to the common goal of synthesis of novel RiPPs through the modularization of the PRPS pathways. These results will be discussed in the present section.

4.1. Heterologous overexpression of MvdC, MvdD, and MvdB

Although heterologous expression of the *P. agardhii* microviridin ATP-grasp ligases had been achieved before, the yield of soluble enzyme was low and in some cases the enzymes produced were inactive [102, 110]. For this reason a new attempt to overexpress these proteins was carried out. It was observed that expression at room temperature resulted in higher amounts of soluble protein compared to expression at 37ºC. A possible explanation for this is that at low temperatures the cellular processes slow down, enhancing correct protein folding and avoiding protein aggregation [113]. The optimized methodology yielded active enzymes in amounts greater than for any methodology previously used in Prof. Dittmann's group, nevertheless the yield of MvdC, the amide ligase, was always lower than that of MvdD, in agreement with previous results [110]. The difference in the yield is notable since the ligases share a sequence similarity of 57%. It had been previously proposed that interaction with other microviridin biosynthetic proteins is required for stabilization of the amide ligase. Weiz *et al*. observed that expression in *E. coli* of the microviridin cassette from *M. aeruginosa* NIES 843 lacking the ABC-transporter resulted in decreased abundance of the amide ligase, as compared with the expression of a cassette containing all the genes [103]. This could offer an explanation for the difference in the yield of the ligases. A second explication might be the presence of motifs that negatively regulate gene expression in the sequence of MvdC. By using the GenScript rare codon analysis tool (https://www.genscript.com/tools/rare-codon-analysis, GenScript, USA), it was determined that the

DNA sequences of both enzymes are suitable for heterologous expression in *E. coli*. However, the sequence of MvdC contains four negative *cis* elements compared to none in the sequence of MvdD. Finally, the yield of the N-acetyltransferase MvdB was outstanding, in line with the results of Philmus *et al*. [102].

4.2. Substrate recognition and activation of the microviridin ATP-grasp ligases

A great deal has been achieved in recent years regarding the understanding of the mechanism by which the PTM enzymes recognize their substrates. It had long been suspected that the leader region of the precursor peptide had a major role in substrate recognition and activation of the PRPS biosynthetic machineries based on different pieces of evidence, for example that removal of the leader peptide would, in some cases, completely abolish RiPP production. The identification of the leader peptide recognition motif RRE (see section 1.3) helped to comprehend the molecular basis underlying the interaction between the PTM enzymes and the leader peptides of the majority of the prokaryotic RiPP families known to date (Fig. 4.1). Nonetheless, the RRE could not be identified in the ligases of the microviridin pathway [57].

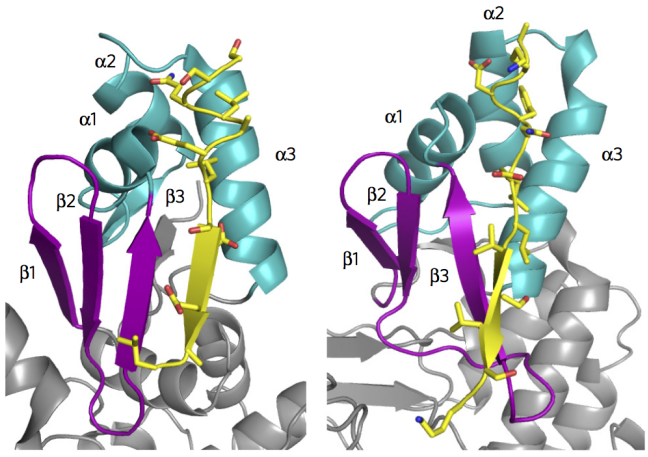


Figure 4.1. Comparison of the RRE in the cyclodehydratase LynD and the dehydratase NisB. Part of the leader peptide (yellow) adopts a β-stand conformation and interacts with the three-stranded antiparallel β-sheet of the RRE domain (purple and cyan), resulting in a four-stranded β-sheet. Image taken from [57].

Although the microviridin PTM enzymes do not recognize their leader peptide through the mentioned RRE, interaction with the former does promote enzyme activation. Modification of the microviridin core peptide by the ATP-grasp ligases could not be observed when the leader peptide was missing, however, when the latter was added in *trans*, cyclic core peptides could be detected. This is congruent with findings from analogous experiments

carried out with other RiPP biosynthetic machineries [54, 68, 78, 114], and supports the hypothesis that the leader peptide induces a conformational change in the PTM enzymes that allows for successive modification of the core peptide [58].

To shed light on the mechanism behind the interaction of the microviridin PTM enzymes and their leader peptide, crystallization attempts were carried out in a cooperation with the group of Prof. Groll from the TU Munich. Unfortunately despite numerous efforts no crystals of MvdD or MvdC could be obtained. Parallel to the present study, the group of Li *et al*. published the solved structures of MdnC and MdnB, enzymes homologous to MvdD and MvdC, respectively [56]. The ester ligase was crystallized in complex with its cognate leader peptide, providing an insight into the mechanism underlying substrate recognition (Fig. 4.2). The conserved PFFARFL motif in the microviridin leader peptide adopts an α-helix conformation and interacts with the helix α7 of the ligase. This differs from the mode of interaction between leader peptide and PTM enzymes that possess the RRE. Of particular importance are the Arg in the PFFARFL motif, and the Glu191, Asp192 and Asn195 of MdnC α7 due to the electrostatic interactions established between them [56].

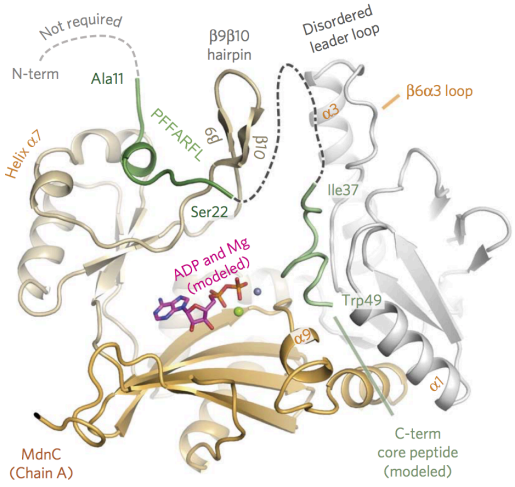


Figure 4.2. Model of the interaction between the microviridin precursor peptide and the ester ligase MdnC. The structure of MdnC and the leader peptide could be resolved. The nucleotide, Mg^{2+} and core peptide were modeled. It can be appreciated that the conserved motif PFFARFL forms an α-helix and is in close contact with the helix α7 of the enzyme. The image was modified from [56].

Homology models of the ligases from *P. agardhii* could be constructed based on the structures of MdnC (co-crystallized with its leader peptide) and MdnB (sequence identity of 79% and 70% with MvdD and MvdC respectively. See Fig. A.4 in the appendix). Since the ATP-grasp ligases of *P. agardhii* share a sequence similarity of 57%, an alignment of the homology models could illustrate conformational changes triggered by the leader peptide. As

noted in the aforementioned article, it seems that the binding of the leader peptide to the ligases provokes a rearrangement in its tertiary structure, namely a shift of the β9β10 hairpin (Fig. 4.3). The authors suggest that the latter gives the core region of the precursor peptide access to the catalytic center of the ligase, allowing the cyclizations to occur (Fig. 4.2 and 4.3). The shift of a loop in the C-terminal domain can also be observed, possibly allowing the interaction with ATP.

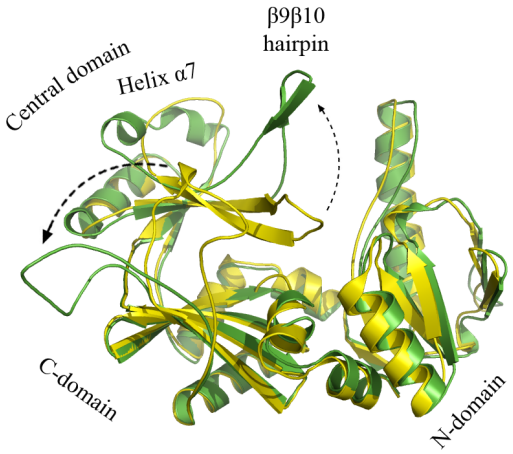


Figure 4.3. Superimposition of the homology models of MvdD and MvdC. In this alignment of the monomers of MvdD (green) and MvdC (yellow), a difference in the position of the β9β10 hairpin can be observed between the two structures. In MvdC, the β9β10 hairpin occludes the access to the catalytic core of the enzyme. In MvdD, a loop in the C-domain is shifted toward the outside of the enzyme. The homology models were created with SWISS model [115] based on the structures of MdnC (co-crystallized with its leader peptide; PDB file 5IG9) and MdnB (PDB file 5IG8), and subsequently aligned with PyMol [116].

The affinity of the microviridin precursor peptide for the ligases was measured in the aforementioned study by means of isothermal titration calorimetry-based binding assays. It was evidenced by these results that the affinity of MdnC for its cognate substrate (K_d linear precursor: 112 nM) is two orders of magnitude greater than that of MdnB for its own (K_d bicyclic precursor: 6.4 μM). The authors propose that a faster processing of the linear precursor peptide is required because of its lower stability [56]. An alternative explanation for the observed difference in substrate affinity could be that in the case of the amide ligase, substrate recognition is assisted by the ester ligase. This can be compared to the lasso peptide biosynthetic pathways where the B protein is split: The B1 protein (RRE domain) recognizes the leader peptide and presents the precursor peptide to the B2 protein for cleavage [87]. Moreover, a recent phylogenomic analysis of the microviridin biosynthetic pathway revealed that the ligase genes possess a strictly conserved order in all BGCs and are translationally coupled in many strains [117], suggesting a close interaction of these enzymes. On top of that, an scaffolding role has been previously proposed for MvdA, the ABC-transporter in the

microviridin biosynthetic machinery [103]. The stability of the amide ligase might be supported by one or more proteins in an enzyme complex as discussed before, therefore it is not unreasonable to think that substrate recognition might be assisted too. Indeed, there are known systems where the formation of a protein complex is needed for biosynthesis, for example microcin B17 and the lantibiotic nisin [118, 119]. It is important to note that not every microviridin BGC has been found to harbor an homolog of MvdA, however, it is possible that the enzyme is encoded elsewhere in the genome [103]. An interaction between the ATP-grasp ligases could be demonstrated with experiments such as yeast two-hybrid screening or pull-down assays. Another approach could be to compare the rate of conversion of bicyclic precursor to tricyclic precursor in presence and absence of the ester ligase.

Similar to the cases of the lanthipeptide synthetase LctM and the cyanobactin cyclodehydratase LynD [54, 68], activation of the microviridin ligases was possible when the leader peptide was covalently attached to their N-terminus via a peptidic linker. It has been demonstrated that the activation of the PTM enzymes depends mainly on the interaction with the leader peptide [56], therefore it is irrelevant if it is supplied in *trans* as one piece with the core peptide or if its part of the PTM enzyme. In this regard, it is noteworthy that the cyclization assays with the engineered ester ligases from *Cyanothece* sp. and *A. sagamiensis* did not work. It should be mentioned that the *in vitro* system with the enzymes from *P. agardhii* had been previously established and the conditions for the reaction had been optimized [102, 110]. Despite the high sequence similarity between the ester ligases (90% and 60% for the ligases from *Cyanothece* sp. and *A. sagamiensis* respectively, compared to the one from *P. agardhii.* Fig. 4.4), it is now clear that the conditions for the cyclization reactions need to be optimized individually to assure enzymatic function. Concerning the biosynthetic pathway of the microviridin-related marinostatin, further characterization is needed in order to successfully reconstitute it *in vitro*.

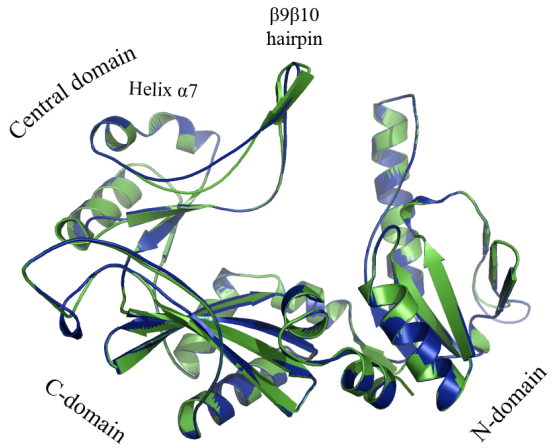


Figure 4.4. Superimposition of the homology models of MvdD and the ester ligase of *A. sagamiensis*. Alignment of the monomers of MvdD (green) and the ω-ester ligase of *A. sagamiensis* (blue). The enzymes seem similar overall, however, slight differences can be seen in the β9β10 hairpin. The homology models were created with SWISS model [115] based on the structure of MdnC (co-crystallized with its leader peptide; PDB file 5IG9) and subsequently aligned with PyMol [116].

4.3. Bioengineering potential of microviridin

The microviridin one-pot *in vitro* platform developed in this work can be defined as a milestone in the field of RiPP synthesis. The reconstitution of multistep RiPP pathways in one-pot reactions has only recently become possible due to the technological development in the area of genetics and bioinformatics, and only a few examples haven been reported to date [80, 120, 121]. The synthesis of microviridin analogs had been largely hampered by the absence of an adequate peptidase to remove the leader peptide from the cyclic core peptide. The previous *in vivo* approach by Weiz *et al*. was an important accomplishment in the direction of microviridin engineering, however the large-scale fermentation needed for the production of single variants made it a non-trivial project. Moreover, the yield and processing efficiency varied for the different microviridins produced [106]. The new platform not only allows the simple and cost-effective production of microviridin analogs, but also the isolation and characterization of microviridin intermediates.

4.3.1. Mdn K and Mdn B libraries

The new chemoenzymatic technique was tested with the core peptide of Mdn B, which differs in four N-terminal amino acids from the Mdn K core peptide. After successful cyclization and acetylation, these two core peptides were used as leads for the generation of peptide libraries focused on the important fifth position. The efficiency of bicyclization was not the same for all the variants. While some variants were completely bicyclized after 3 hours of incubation, e.g. the lead core peptides, others were poorly processed even after 16 hours of incubation. Of the latter, the most notable cases are the cores bearing a Pro or a Cys at position 5. An

explanation for the unsuccessful processing of the cores with a Pro substitution may involve the inherent conformational rigidity of this amino acid, which could prevent a possible structural rearrangement needed for the installation of the large lactone ring. Regarding the Cys substitution, it is very probable that interpeptidic disulfide bonds prevented the cyclization. The resuspension of these core peptides in a reducing buffer previous to cyclization may improve the yield of cyclized product [122].

As described before, PTM of the peptides was evaluated via HPLC and MALDI-TOF. Nonetheless, a more conclusive technique was required for the confirmation of the insertion of the rings and their position in the molecule. The fragmentation pattern of bicyclic Mdn K, and bicyclic and tricyclic Mdn B generated by means of ESI-HR-MS/MS revealed ions correlating with the installment of the expected lactone and lactam rings, particularly the presence of dehydrated Thr and Ser (Fig. 3.15) [123].

The bicyclic intermediates produced by incubation with LP-MvdD were used against the serine proteases trypsin and subtilisin in assays to screen for inhibitory activity. As expected, variants carrying Arg or Lys showed activity against trypsin. The reason for trypsin inhibition by microviridins with positively-charged residues at position 5 resides in the substrate-like interaction between the protease and the depsipeptide [106]. Notably, variants with a His substitution did not show inhibitory activity. A possible explanation for this is the length of the side chain of this amino acid, which may be too short to interact with the protease. Variants harboring Met at the fifth position as well as the variants L5Q and L5Y inhibited subtilisin (Table 3.1). If a substrate-like interaction is also considered for subtilisin, then large uncharged residues at microviridin position 5 should be expected to inhibit the enzyme [124]. This is true for L5Y and to some extent for the variants with Met, but not for L5Q. Furthermore, the variants with Tyr or Gln substitutions from the Mdn K library showed no inhibition. These results seem to suggest that in the case of substilisin, the residues at the N-terminus of microviridin are important for inhibition too. In any case, a structure of the subtilisin-microviridin complex would help to understand these results.

The bicyclic microviridins that exhibited inhibition were selected for further maturation, rendering tricyclic and acetylated tricyclic variants in satisfactory yields. Concerning trypsin inhibition, a trend pointing toward significant microviridin bioactivity improvement as a result of the insertion of the lactam ring and acetylation, respectively, was revealed (Table 3.1). This can also be explained with the crystallized trypsin-microviridin complex where the direct interaction between the acetyl group and the enzyme can be

appreciated (Fig. 1.11). Opposed to these findings, the tricyclization and acetylation of microviridin seemed to be less relevant for activity against subtilisin.

4.3.2. Genome-mined microviridin libraries

As mentioned in section 1.1, a great diversity of natural products awaits exploitation. Nonetheless, many microorganisms predicted to be producers of natural products have proven to be impossible to culture in the laboratory, and even when culturable, many BGCs are not expressed under laboratory conditions. Synthesis approaches that bypass the need of microbial cultures are thus a very attractive alternative for the production of natural products, as exemplified by the case of the humimycins [21]. In line with this, the platform established in the present study offers a culture-free route for the synthesis of bioinformatically predicted microviridins. To prove the latter, a third peptide library containing putative microviridin core peptides from *Cyanothece* sp., the MvdF core peptide from *P. agardhii*, and the core peptide of marinostatin from *A. sagamiensis* was assembled. A notable finding was revealed after incubation of the library with LP-MvdD, namely that bicyclization was apparently inhibited by substitution of the Tyr14 by a polar residue (Fig. 3.19). This was unexpected since several mined microviridin core peptides possess Ser or Glu at the last position. Moreover, the presence of Glu at the last position had prompted the hypothesis that an alternative lactone or lactam ring in the microviridin scaffold was possible. The lack of enzymatic activity of LP-MvdD on the MvdF core peptide explains why the expected microviridin deriving from this precursor peptide has never been found in extracts of *P. agardhii* [102]. Concerning the marinostatin core peptide, it is highly possible that not only the polar substitution at position 14 hampered the PTM, but also the substitutions of the strictly conserved Lys6 and Trp11. The four variants from *Cyanothece* sp. that were bicyclized (Cth 1-4) were then incubated with LP-MvdC and MvdB. Of these, only Cth 2 and Cth 3 could be fully maturated with satisfactory yields. In the case of Cth 1 the yields were surprisingly low, while for Cth 4 acetylation could not be detected. The shorter sequence of Cth 4 is probably the reason for a poor tricyclization and absence of activity from the N-acetyltransferase. The unusually poor maturation of Cth 1 does not correlate with the results obtained from the first two libraries. Consequently, the most plausible explication for the latter is that this particular peptide was very sensitive and it deteriorated extremely fast in solution. A repetition of the assays with fresh Cth 1 core peptide would most likely render higher yields of mature product.

The fully maturated Cth 2 and Cth 3, and the bicyclic Cth 1 and Cth 4 were used in inhibition assays against trypsin and elastase, respectively. The microviridins Cth 2 and Cth 3 inhibited trypsin with IC_{50} values in the low nanomolar range, representing the most potent microviridin inhibition toward trypsin reported to date. The bicyclic Cth 1 and Cth 4 inhibited elastase with IC_{50} values in the low micromolar range. Again, this can be explained from a substrate-like interaction perspective since elastase shows selectivity for cleavage of peptidic bonds involving amino acids with small hydrophobic residues [125, 126]. Extrapolating the results from the other libraries, it would be expected that lactamization and N-acetylation of the latter variants enhance their inhibitory activity against elastase. The potent inhibition exhibited by the microviridins from *Cyanothece* sp. supports the hypothesis that this species produces a cocktail of serine protease inhibitors.

Analogous to the analyses carried out with the lead microviridins from the first two libraries, bicyclic Cth 2 and tricyclic Cth 3 were used in tandem mass spectrometry measurements to confirm the insertion and position of the rings. Furthermore, a methanolic extract of *Cyanothece* sp. PCC 7822 was also tested for the presence of microviridins produced *in vivo*. Four of the ten predicted depsipeptides could be detected in the extract, albeit in very little amounts, and notably only 13- and 14-membered variants. The fragmentation pattern of the ion representing the Cth 3 microviridin resembled the one of the synthetic tricyclic Cth 3, corroborating its correct maturation *in vitro* (Fig. 3.20-21).

4.3.3. Microviridins harboring non-canonical amino acids

The presence of ncAAs in the backbones of RiPPs is not as common as in non-ribosomal peptides. The incorporation of these particular building blocks in the scaffold of RiPPs is a common objective in the field of natural product research. The interest in the latter resides in the fact that ncAAs confer unique physicochemical properties to the peptides, from improved stability and bioactivities, to the possibility of further PTM by chemical or photoactivation. An approach to this goal is to introduce the ncAAs *in vivo* via selective pressure incorporation or stop codon suppression. Significant achievements have been attained through this approach, particularly for antimicrobial peptides [127]. The chemoenzymatic platform developed in this study was thus employed as an *in vitro* alternative to test the possibility of inclusion of ncAAs into the microviridin scaffold.

The interest in the generation of microviridin variants carrying ncAAs lay beyond the improvement of the inhibitory activity and was aimed at the creation of molecular tools for

further microviridin research. Of particular interest were tools for the identification of microviridin binding partners and for the intracellular localization of microviridin. A system involving a modified microviridin that could be targeted in pull-down assays was then envisioned for the identification of *in vivo* binding partners. Biotinylated lysine was therefore selected as a viable approach to develop such a system. Three variants of the Mdn J backbone were created, harboring a Lys(biotin) either at the N- or C-terminus, or replacing the natural Lys6. Since the sequence requirements for the installment of the lactam ring are more stringent than for the lactone rings, only the bicyclization was expected to be possible. Nonetheless, bicyclic variants were believed to be functional in light of the results observed with the first two microviridin libraries. After incubation with LP-MvdD, a bicyclic and a monocyclic variant derived from the peptides with Lys(biotin) at their N- or C-terminus were obtained, respectively (Fig. 3.24). The reaction products were analyzed by HPLC and MALDI-TOF and the correspondent mass loss was confirmed. Since biotin does not possess any reactive group capable of forming lactones, it was assumed that the position of the rings in the molecules was the same as for the rest of the microviridins created before. The remarkable loss of production with the L6L(biotin) substitution and the single cyclization of the variant with the biotin tag at the C-terminus are probably due to steric hindrance by the non-native biotin group. In contrast with the latter, it seems that the engineered ester ligase is able to accommodate the biotin group if it is located at the N-terminus of the core peptide. The Mdn J backbone has an Arg5, therefore trypsin inhibition was expected. The monocyclic and bicyclic products showed activity against trypsin in protease inhibition assays with IC_{50} in the low micromolar range. The inhibition of trypsin by the bicyclic depsipeptide was tenfold more potent than that of the monocyclic one, in line with previous results. These findings demonstrate the functionality of the non-natural microviridins despite the insertion of the biotin group. It can thus be expected that the interaction with their *in vivo* partners will not be disrupted either, making them eligible as bait peptides for pull-down experiments (Fig. 4.5).

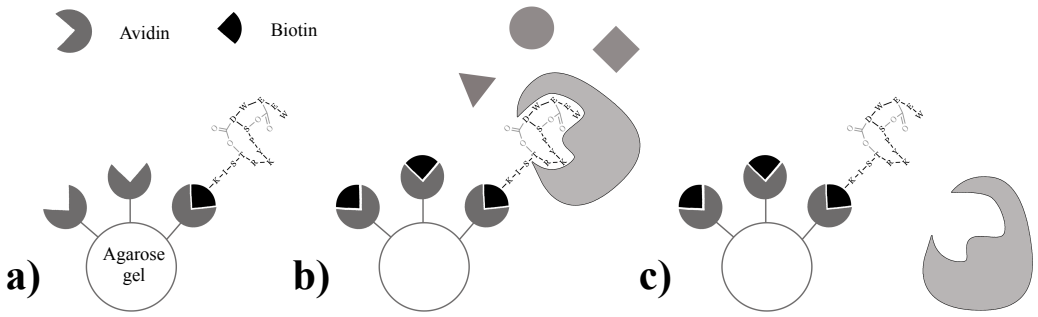


Figure 4.5. Scheme of a pull-down experiment with biotinylated microviridin. a) Biotinylated microviridin binds avidin; **b)** After blocking avidin with free biotin, a lysate containing the binding partner of microviridin is added to the gel; **c)** After washing the non-bound proteins away, the binding partner of microviridin is eluted.

4.3.4. Limitations of the platform

The advantages of the developed chemoenzymatic technique have been demonstrated by the results obtained so far. Nevertheless, there is an intrinsic limitation of the platform, namely the conserved microviridin core sequence TxKxPSDWE(D/E). The characteristic intramolecular ω-ester and ω-amide crosslinks are interesting features providing structural stability, hence harnessing dedicated enzymes capable of introducing these PTM onto ribosomal peptides is a powerful tool for the synthesis of designer natural products. As mentioned in the introduction, early microviridin bioengineering efforts focused on altering the size of the rings, with the outcome that this was fixed and only the substitution of some surrounding amino acids was possible in the case of the large lactone ring [104]. The study by Li *et al*. demonstrated by comparing the K_d of precursor peptide, leader peptide, and the PFFARFL region to MdnC and MdnB that substrate binding is mainly directed by the leader peptide, in particular the PFFARFL region [56]. Nonetheless, the results of the present and previous studies point toward a structural semirigidity in the microviridin scaffold. In order to gain a better understanding of the latter, the precise catalytic mechanism behind the activity of the ligases and structural characterization of the cyclic intermediates by spectroscopic methods will be needed. A novel natural product resembling microviridin has been reported in a very recent publication. The new 45-membered RiPP, plesiocin, harbors eight lactone rings introduced by an ATP-grasp ligase related to MvdD, and has been shown to be a potent inhibitor of elastase and chymotrypsin (Fig. 4.6) [128]. The sequence required for installment of the lactone rings in plesiocin seems to be TTxxxxEE, varying from the semiconserved TxKxPSDWE(D/E) of microviridin.

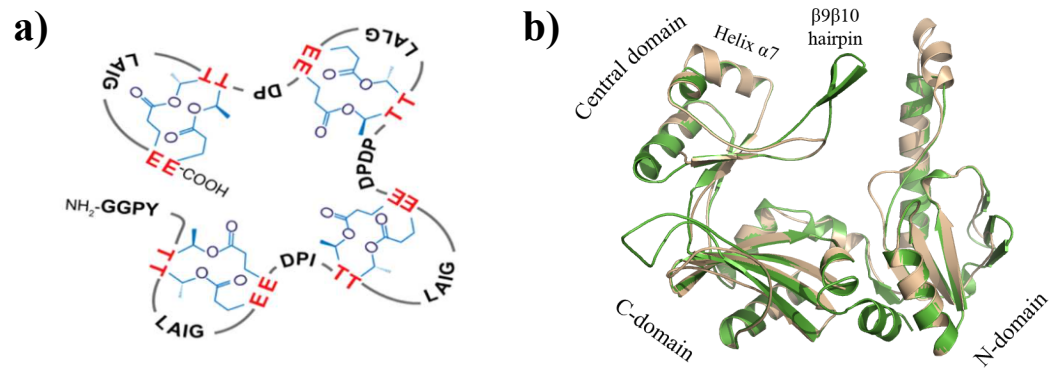


Figure 4.6. The recently discovered plesiocin. a) Structure of plesiocin featuring eight lactone rings, resulting in a complex structure with four hairpins. Image taken from [128]; **b)** Alignment of the homology models of the monomers of MvdD (green) and PsnB (beige), the ω-ester ligase from the plesiocin pathway. The β9β10 hairpin and the loop in the C-domain are notably shorter in the plesiocin ligase, possibly to allow the accommodation of the larger core peptide of plesiocin. The homology models were created with SWISS model [115] based on the structure of MdnC (co-crystallized with its leader peptide; PDB file 5IG9) and subsequently aligned with PyMol [116].

The discovery of plesiocin offers many opportunities for the development of RiPP research. The new ester ligase, PsnB, is a potential tool for the production of designer RiPPs ready to be exploited. Further characterization of this ligase as well of congeners from the microviridin and the marinostin pathway could help identify the mechanism underlying lactone cyclization, assisting in a rational design of ester ligases. Another potential application could be the generation of multi-targeted serine protease inhibitors possessing a microviridin-plesiocin hybrid structure (Fig. 4.7), similar to the hybrid RiPPs produced by combining a cyclodehydratase from the microcin family with enzymes from the sactipeptide or the lanthipeptide pathways [129].

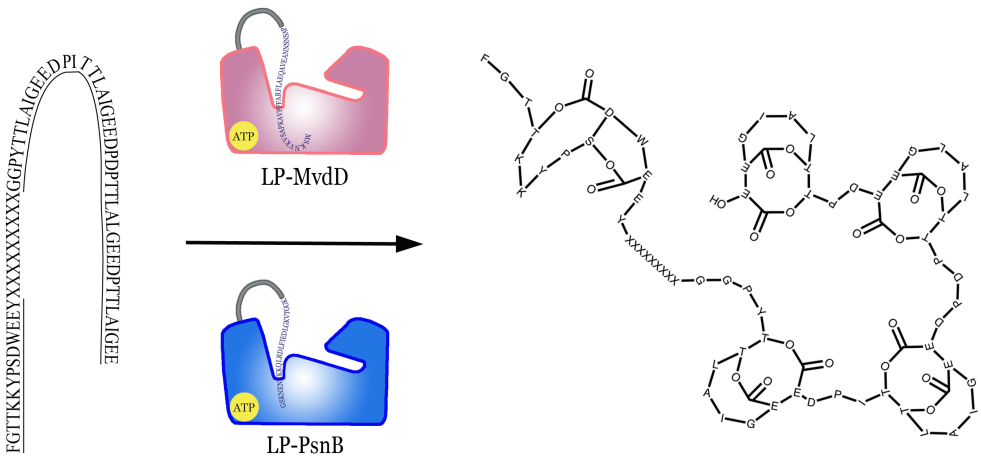


Figure 4.7. Hypothetical generation of a multi-targeted serine protease inhibitor. A precursor peptide comprised of the microviridin core peptide, a spacer and the plesiocin core peptide is incubated with the constitutively activated ω-ester ligases LP-MvdD and LP-PsnB. LP-MvdD introduces two lactone rings into the microviridin core region while LP-PsnB installs eight lactones in the plesiocin core region. It is hypothesized that such a microviridin-plesiocin hybrid would inhibit trypsin and elastase at the same time.

5. Conclusions and perspectives

The present work is an example of the bioengineering potential of RiPPs. The biosynthetic enzymes of the microviridin pathway, MvdC and MvdD, were engineered to catalyze the introduction of lactone and lactam rings into synthetic short core peptides fed *in trans*. The successive N-acetylation of the cyclic peptides was possible with the additional activity of the microviridin tailoring enzyme MvdB, successfully reconstituting the microviridin pathway in a time and cost-effective one-pot reaction. Four core peptide libraries were then designed with different objectives: The first two peptide libraries, based on the backbones of Mdn B and Mdn K, were screened for variants with improved bioactivity. Furthermore, the modular nature of the platform permitted the production of intermediates that were used to establish a SAR between microviridin and trypsin; the third library was aimed at the production of cryptic microviridins and marinostatin, a microviridin-related serine protease inhibitor; finally, the objective of the last library was to assess the possibility of incorporating ncAAs into the microviridin scaffold. The results obtained can be summarized as follows: First, that it is possible to engineer microviridin core peptides to enhance their bioactivity as long as the conserved motif is preserved; second, the enzymes from *P. agardhii* can be used to synthesize bioinformatically predicted microviridins from other species containing the conserved motif; third, the production of microviridins harboring ncAAs is possible. In the particular case of biotinylated lysine, its position has to be either N- or C-terminal. Everything considered, the established platform is a powerful technique for the generation of rationally designed serine protease inhibitors and molecular tools that can be used to further deepen the existing knowledge on microviridin. To overcome the current limitations of the platform, additional research on the mode of action of the ligases will be needed. To further enhance the inhibitory activity of microviridin against subtilisin, more insights into the mechanism of inhibition are required. The microviridin ATP-grasp ligases and the tailoring N-acetyltransferase constitute modular tools for the production of novel RiPPs. Through the rational design of hybrid precursors and combination of the microviridin PTM enzymes with enzymes from other RiPP pathways, the production of tailored new-to-nature natural products will be broadened.

Bibliography

1. Cragg, G.M. and Newman, D.J., *Natural products: A continuing source of novel drug leads.* Biochimica et Biophysica Acta (BBA) - General Subjects, 2013. **1830**(6): p. 3670-3695.
2. Dias, D.A., Urban, S., and Roessner, U., *A Historical Overview of Natural Products in Drug Discovery.* Metabolites, 2012. **2**(2).
3. Kinghorn, A.D., *Drug discovery from natural products*, in *Foye's principles of medicinal chemistry*, Lemke, T.L. and Williams, D.A., Editors. 2013, Lippincott Williams & Wilkins: Philadelphia. p. 13-28.
4. *All natural.* Nat Chem Biol, 2007. **3**: p. 351-351.
5. Baker, D.D., Chu, M., Oza, U., and Rajgarhia, V., *The value of natural products to future pharmaceutical discovery.* Natural Product Reports, 2007. **24**(6): p. 1225-1244.
6. Scriabine, A., *Discovery and development of major drugs currently in use*, in *Pharmaceutical innovation: Revolutionizing human health*, Landau, R., Achilladelis, B., and Scriabine, A., Editors. 1999, Chemical Heritage Press: Philadelphia. p. 148-270.
7. Brown, E.D. and Wright, G.D., *Antibacterial drug discovery in the resistance era.* Nature, 2016. **529**(7586): p. 336-343.
8. Wright, G.D., *Something old, something new: revisiting natural products in antibiotic drug discovery.* Canadian Journal of Microbiology, 2014. **60**(3): p. 147-154.
9. Smanski, M.J., Zhou, H., Claesen, J., Shen, B., Fischbach, M.A., and Voigt, C.A., *Synthetic biology to access and expand nature's chemical diversity.* Nat Rev Micro, 2016. **14**(3): p. 135-149.
10. Newman, D.J. and Cragg, G.M., *Natural Products as Sources of New Drugs from 1981 to 2014.* Journal of Natural Products, 2016. **79**(3): p. 629-661.
11. Eder, J., Sedrani, R., and Wiesmann, C., *The discovery of first-in-class drugs: origins and evolution.* Nat Rev Drug Discov, 2014. **13**(8): p. 577-587.
12. Laraia, L. and Waldmann, H., *Natural product inspired compound collections: evolutionary principle, chemical synthesis, phenotypic screening, and target identification.* Drug Discovery Today: Technologies, 2017. **23**: p. 75-82.
13. Harvey, A.L., Edrada-Ebel, R., and Quinn, R.J., *The re-emergence of natural products for drug discovery in the genomics era.* Nat Rev Drug Discov, 2015. **14**(2): p. 111-29.
14. Harvey, A.L., *Natural products in drug discovery.* Drug Discov Today, 2008. **13**(19-20): p. 894-901.
15. Koehn, F.E. and Carter, G.T., *The evolving role of natural products in drug discovery.* Nat Rev Drug Discov, 2005. **4**(3): p. 206-220.

16. Clardy, J., Fischbach, M.A., and Walsh, C.T., *New antibiotics from bacterial natural products.* Nat Biotech, 2006. **24**(12): p. 1541-1550.
17. Pye, C.R., Bertin, M.J., Lokey, R.S., Gerwick, W.H., and Linington, R.G., *Retrospective analysis of natural products provides insights for future discovery trends.* Proceedings of the National Academy of Sciences, 2017. **114**(22): p. 5601-5606.
18. Doroghazi, J.R., Albright, J.C., Goering, A.W., Ju, K.S., Haines, R.R., Tchalukov, K.A., Labeda, D.P., Kelleher, N.L., and Metcalf, W.W., *A roadmap for natural product discovery based on large-scale genomics and metabolomics.* Nat Chem Biol, 2014. **10**(11): p. 963-8.
19. Demain, A.L., *Antibiotics: natural products essential to human health.* Med Res Rev, 2009. **29**(6): p. 821-42.
20. Challinor, V.L. and Bode, H.B., *Bioactive natural products from novel microbial sources.* Annals of the New York Academy of Sciences, 2015. **1354**(1): p. 82-97.
21. Chu, J., Vila-Farres, X., Inoyama, D., Ternei, M., Cohen, L.J., Gordon, E.A., Reddy, B.V., Charlop-Powers, Z., Zebroski, H.A., Gallardo-Macias, R., Jaskowski, M., Satish, S., Park, S., Perlin, D.S., Freundlich, J.S., and Brady, S.F., *Discovery of MRSA active antibiotics using primary sequence from the human microbiome.* Nat Chem Biol, 2016. **12**(12): p. 1004-1006.
22. Breton, R.C. and Reynolds, W.F., *Using NMR to identify and characterize natural products.* Natural Product Reports, 2013. **30**(4): p. 501-524.
23. Moss, N.A., Bertin, M.J., Kleigrewe, K., Leao, T.F., Gerwick, L., and Gerwick, W.H., *Integrating mass spectrometry and genomics for cyanobacterial metabolite discovery.* J Ind Microbiol Biotechnol, 2016. **43**(2-3): p. 313-24.
24. Demain, A.L., *Importance of microbial natural products and the need to revitalize their discovery.* J Ind Microbiol Biotechnol, 2014. **41**(2): p. 185-201.
25. Berdy, J., *Bioactive microbial metabolites.* J Antibiot (Tokyo), 2005. **58**(1): p. 1-26.
26. Silva-Stenico, M.E., Kaneno, R., Zambuzi, F.A., Vaz, M.G., Alvarenga, D.O., and Fiore, M.F., *Natural products from cyanobacteria with antimicrobial and antitumor activity.* Curr Pharm Biotechnol, 2013. **14**(9): p. 820-8.
27. Abed, R.M., Dobretsov, S., and Sudesh, K., *Applications of cyanobacteria in biotechnology.* J Appl Microbiol, 2009. **106**(1): p. 1-12.
28. Chlipala, G.E., Mo, S., and Orjala, J., *Chemodiversity in freshwater and terrestrial cyanobacteria - a source for drug discovery.* Curr Drug Targets, 2011. **12**(11): p. 1654-73.
29. Mazard, S., Penesyan, A., Ostrowski, M., Paulsen, I.T., and Egan, S., *Tiny Microbes with a Big Impact: The Role of Cyanobacteria and Their Metabolites in Shaping Our Future.* Mar Drugs, 2016. **14**(5).

30. Barrios-Llerena, M.E., Burja, A.M., and Wright, P.C., *Genetic analysis of polyketide synthase and peptide synthetase genes in cyanobacteria as a mining tool for secondary metabolites.* Journal of Industrial Microbiology & Biotechnology, 2007. **34**(6): p. 443-456.
31. Burja, A.M., Banaigs, B., Abou-Mansour, E., Grant Burgess, J., and Wright, P.C., *Marine cyanobacteria—a prolific source of natural products.* Tetrahedron, 2001. **57**(46): p. 9347-9377.
32. Blaha, L., Babica, P., and Marsalek, B., *Toxins produced in cyanobacterial water blooms - toxicity and risks.* Interdiscip Toxicol, 2009. **2**(2): p. 36-41.
33. Dittmann, E., Neilan, B.A., and Borner, T., *Molecular biology of peptide and polyketide biosynthesis in cyanobacteria.* Appl Microbiol Biotechnol, 2001. **57**(4): p. 467-73.
34. Dittmann, E., Fewer, D.P., and Neilan, B.A., *Cyanobacterial toxins: biosynthetic routes and evolutionary roots.* FEMS Microbiol Rev, 2013. **37**(1): p. 23-43.
35. Nunnery, J.K., Mevers, E., and Gerwick, W.H., *Biologically active secondary metabolites from marine cyanobacteria.* Curr Opin Biotechnol, 2010. **21**(6): p. 787-93.
36. Dittmann, E., Gugger, M., Sivonen, K., and Fewer, D.P., *Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria.* Trends Microbiol, 2015. **23**(10): p. 642-52.
37. Niedermeyer, T.H., *Anti-infective Natural Products from Cyanobacteria.* Planta Med, 2015. **81**(15): p. 1309-25.
38. Videau, P., Wells, K.N., Singh, A.J., Gerwick, W.H., and Philmus, B., *Assessment of Anabaena sp. Strain PCC 7120 as a Heterologous Expression Host for Cyanobacterial Natural Products: Production of Lyngbyatoxin A.* ACS Synth Biol, 2016. **5**(9): p. 978-88.
39. Okino, T., Matsuda, H., Murakami, M., and Yamaguchi, K., *New microviridins, elastase inhibitors from the blue-green alga Microcystis aeruginosa.* Tetrahedron, 1995. **51**(39): p. 10679-10686.
40. *Microcystin*. [Electronic] 2012 [cited 2017 August]; Available from: http://naturalproductforum.org/tag/microcystin/.
41. *Planktothrix agardhii (Gomont) Anagnostidis et Komárek*. [Electronic] 2013 [cited 2017 July]; Available from: http://ccala.butbn.cas.cz/en/planktothrix-agardhii-gomont-anagnostidis-et-komarek.
42. Tan, L.T., *Bioactive natural products from marine cyanobacteria for drug discovery.* Phytochemistry, 2007. **68**(7): p. 954-79.
43. Lobban, C.S. *Symploca hydnoides filaments showing sheath surrounding trichomes*. [Electronic] [cited 2017 July]; Available from: http://cfb.unh.edu/phycokey/Choices/

Cyanobacteria/cyano_filaments/cyano_unbranched_fil/untapered_filaments/no_heterocysts/vis_sheath/SYMPLOCA/Symploca_Image_page.htm.

44. *Microcystis aeruginosa*. [Electronic] [cited 2017 July]; Available from: http://www.recetox.muni.cz/res/image/pristroj-vybaveni/microcystis_Aeruginosa.jpg.
45. *Lyngbya sp.* The Japanese Freshwater Algae [Electronic] 1977 [cited 2017 July]; Available from: http://protist.i.hosei.ac.jp/PDB/Images/Prokaryotes/Oscillatoriaceae/Lyngbya/Lyngbya_3.html.
46. Welker, M. and von Dohren, H., *Cyanobacterial peptides - nature's own combinatorial biosynthesis.* FEMS Microbiol Rev, 2006. **30**(4): p. 530-63.
47. Cassier-Chauvat, C., Dive, V., and Chauvat, F., *Cyanobacteria: photosynthetic factories combining biodiversity, radiation resistance, and genetics to facilitate drug discovery.* Appl Microbiol Biotechnol, 2017. **101**(4): p. 1359-1364.
48. Skinnider, M.A., Johnston, C.W., Edgar, R.E., Dejong, C.A., Merwin, N.J., Rees, P.N., and Magarvey, N.A., *Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining.* Proc Natl Acad Sci U S A, 2016. **113**(42): p. E6343-E6351.
49. Bindman, N.A. and Der Donk, W.A.V., *RiPPs: Ribosomally Synthesized and Posttranslationally Modified Peptides*, in *Natural Products*, Osbourn, A., Goss, R.J., and Carter, G.T., Editors. 2014, John Wiley & Sons, Inc.: New Jersey. p. 195-217.
50. Truman, A.W., *Cyclisation mechanisms in the biosynthesis of ribosomally synthesised and post-translationally modified peptides.* Beilstein J Org Chem, 2016. **12**: p. 1250-68.
51. Arnison, P.G., Bibb, M.J., Bierbaum, G., Bowers, A.A., Bugni, T.S., Bulaj, G., Camarero, J.A., Campopiano, D.J., Challis, G.L., Clardy, J., Cotter, P.D., Craik, D.J., Dawson, M., Dittmann, E., Donadio, S., Dorrestein, P.C., Entian, K.D., Fischbach, M.A., Garavelli, J.S., Goransson, U., Gruber, C.W., Haft, D.H., Hemscheidt, T.K., Hertweck, C., Hill, C., Horswill, A.R., Jaspars, M., Kelly, W.L., Klinman, J.P., Kuipers, O.P., Link, A.J., Liu, W., Marahiel, M.A., Mitchell, D.A., Moll, G.N., Moore, B.S., Muller, R., Nair, S.K., Nes, I.F., Norris, G.E., Olivera, B.M., Onaka, H., Patchett, M.L., Piel, J., Reaney, M.J., Rebuffat, S., Ross, R.P., Sahl, H.G., Schmidt, E.W., Selsted, M.E., Severinov, K., Shen, B., Sivonen, K., Smith, L., Stein, T., Sussmuth, R.D., Tagg, J.R., Tang, G.L., Truman, A.W., Vederas, J.C., Walsh, C.T., Walton, J.D., Wenzel, S.C., Willey, J.M., and van der Donk, W.A., *Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature.* Nat Prod Rep, 2013. **30**(1): p. 108-60.
52. Kehr, J.C., Gatte Picchi, D., and Dittmann, E., *Natural product biosyntheses in cyanobacteria: A treasure trove of unique enzymes.* Beilstein J Org Chem, 2011. **7**: p. 1622-35.

53. Oman, T.J. and van der Donk, W.A., *Follow the leader: the use of leader peptides to guide natural product biosynthesis.* Nat Chem Biol, 2010. **6**(1): p. 9-18.
54. Koehnke, J., Mann, G., Bent, A.F., Ludewig, H., Shirran, S., Botting, C., Lebl, T., Houssen, W., Jaspars, M., and Naismith, J.H., *Structural analysis of leader peptide binding enables leader-free cyanobactin processing.* Nat Chem Biol, 2015. **11**(8): p. 558-563.
55. Ortega, M.A., Hao, Y., Zhang, Q., Walker, M.C., van der Donk, W.A., and Nair, S.K., *Structure and mechanism of the tRNA-dependent lantibiotic dehydratase NisB.* Nature, 2015. **517**(7535): p. 509-12.
56. Li, K., Condurso, H.L., Li, G., Ding, Y., and Bruner, S.D., *Structural basis for precursor protein-directed ribosomal peptide macrocyclization.* Nat Chem Biol, 2016. **12**(11): p. 973-979.
57. Burkhart, B.J., Hudson, G.A., Dunbar, K.L., and Mitchell, D.A., *A prevalent peptide-binding domain guides ribosomal natural product biosynthesis.* Nat Chem Biol, 2015. **11**(8): p. 564-70.
58. Ortega, M.A. and van der Donk, W.A., *New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-translationally Modified Peptide Natural Products.* Cell Chem Biol, 2016. **23**(1): p. 31-44.
59. Link, A.J., *Biosynthesis: Leading the way to RiPPs.* Nat Chem Biol, 2015. **11**(8): p. 551-2.
60. Dang, T. and Sussmuth, R.D., *Bioactive Peptide Natural Products as Lead Structures for Medicinal Use.* Acc Chem Res, 2017. **50**(7): p. 1566-1576.
61. Repka, L.M., Chekan, J.R., Nair, S.K., and van der Donk, W.A., *Mechanistic Understanding of Lanthipeptide Biosynthetic Enzymes.* Chem Rev, 2017. **117**(8): p. 5457-5520.
62. Muller, W.M., Ensle, P., Krawczyk, B., and Sussmuth, R.D., *Leader peptide-directed processing of labyrinthopeptin A2 precursor peptide by the modifying enzyme LabKC.* Biochemistry, 2011. **50**(39): p. 8362-73.
63. Voller, G.H., Krawczyk, J.M., Pesic, A., Krawczyk, B., Nachtigall, J., and Sussmuth, R.D., *Characterization of new class III lantibiotics--erythreapeptin, avermipeptin and griseopeptin from Saccharopolyspora erythraea, Streptomyces avermitilis and Streptomyces griseus demonstrates stepwise N-terminal leader processing.* Chembiochem, 2012. **13**(8): p. 1174-83.
64. Knerr, P.J. and van der Donk, W.A., *Discovery, Biosynthesis, and Engineering of Lantipeptides.* Annual Review of Biochemistry, 2012. **81**(1): p. 479-505.
65. Ongey, E.L. and Neubauer, P., *Lanthipeptides: chemical synthesis versus in vivo biosynthesis as tools for pharmaceutical production.* Microb Cell Fact, 2016. **15**: p. 97.

66. Shi, Y., Yang, X., Garg, N., and van der Donk, W.A., *Production of Lantipeptides in Escherichia coli.* Journal of the American Chemical Society, 2011. **133**(8): p. 2338-2341.

67. Bindman, N.A., Bobeica, S.C., Liu, W.R., and van der Donk, W.A., *Facile Removal of Leader Peptides from Lanthipeptides by Incorporation of a Hydroxy Acid.* Journal of the American Chemical Society, 2015. **137**(22): p. 6975-6978.

68. Oman, T.J., Knerr, P.J., Bindman, N.A., Velasquez, J.E., and van der Donk, W.A., *An engineered lantibiotic synthetase that does not require a leader peptide on its substrate.* J Am Chem Soc, 2012. **134**(16): p. 6952-5.

69. Sivonen, K., Leikoski, N., Fewer, D.P., and Jokela, J., *Cyanobactins-ribosomal cyclic peptides produced by cyanobacteria.* Appl Microbiol Biotechnol, 2010. **86**(5): p. 1213-25.

70. Leikoski, N., Liu, L., Jokela, J., Wahlsten, M., Gugger, M., Calteau, A., Permi, P., Kerfeld, C.A., Sivonen, K., and Fewer, D.P., *Genome mining expands the chemical diversity of the cyanobactin family to include highly modified linear peptides.* Chem Biol, 2013. **20**(8): p. 1033-43.

71. Martins, J. and Vasconcelos, V., *Cyanobactins from Cyanobacteria: Current Genetic and Chemical State of Knowledge.* Mar Drugs, 2015. **13**(11): p. 6910-46.

72. Czekster, C.M., Ge, Y., and Naismith, J.H., *Mechanisms of cyanobactin biosynthesis.* Curr Opin Chem Biol, 2016. **35**: p. 80-88.

73. Sardar, D., Tianero, M.D., and Schmidt, E.W., *Directing Biosynthesis: Practical Supply of Natural and Unnatural Cyanobactins.* Methods Enzymol, 2016. **575**: p. 1-20.

74. Long, P.F., Dunlap, W.C., Battershill, C.N., and Jaspars, M., *Shotgun Cloning and Heterologous Expression of the Patellamide Gene Cluster as a Strategy to Achieving Sustained Metabolite Production.* ChemBioChem, 2005. **6**(10): p. 1760-1765.

75. Schmidt, E.W., Nelson, J.T., Rasko, D.A., Sudek, S., Eisen, J.A., Haygood, M.G., and Ravel, J., *Patellamide A and C biosynthesis by a microcin-like pathway in Prochloron didemni, the cyanobacterial symbiont of Lissoclinum patella.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(20): p. 7315-7320.

76. Donia, M.S., Hathaway, B.J., Sudek, S., Haygood, M.G., Rosovitz, M.J., Ravel, J., and Schmidt, E.W., *Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians.* Nat Chem Biol, 2006. **2**(12): p. 729-35.

77. Ruffner, D.E., Schmidt, E.W., and Heemstra, J.R., *Assessing the combinatorial potential of the RiPP cyanobactin tru pathway.* ACS Synth Biol, 2015. **4**(4): p. 482-92.

78. Goto, Y., Ito, Y., Kato, Y., Tsunoda, S., and Suga, H., *One-pot synthesis of azoline-containing peptides in a cell-free translation system integrated with a posttranslational cyclodehydratase.* Chem Biol, 2014. **21**(6): p. 766-74.

79. Oueis, E., Adamson, C., Mann, G., Ludewig, H., Redpath, P., Migaud, M., Westwood, N.J., and Naismith, J.H., *Derivatisable Cyanobactin Analogues: A Semisynthetic Approach.* Chembiochem, 2015. **16**(18): p. 2646-50.

80. Sardar, D., Lin, Z., and Schmidt, E.W., *Modularity of RiPP Enzymes Enables Designed Synthesis of Decorated Peptides.* Chem Biol, 2015. **22**(7): p. 907-16.

81. Oueis, E., Nardone, B., Jaspars, M., Westwood, N.J., and Naismith, J.H., *Synthesis of Hybrid Cyclopeptides through Enzymatic Macrocyclization.* ChemistryOpen, 2017. **6**(1): p. 11-14.

82. Hegemann, J.D., Zimmermann, M., Xie, X., and Marahiel, M.A., *Lasso peptides: an intriguing class of bacterial natural products.* Acc Chem Res, 2015. **48**(7): p. 1909-19.

83. Tietz, J.I., Schwalen, C.J., Patel, P.S., Maxson, T., Blair, P.M., Tai, H.C., Zakai, U.I., and Mitchell, D.A., *A new genome-mining tool redefines the lasso peptide biosynthetic landscape.* Nat Chem Biol, 2017. **13**(5): p. 470-478.

84. Saito, F. and Bode, J.W., *Synthesis and stabilities of peptide-based [1]rotaxanes: molecular grafting onto lasso peptide scaffolds.* Chem Sci, 2017. **8**(4): p. 2878-2884.

85. Hegemann, J.D., Zimmermann, M., Zhu, S., Klug, D., and Marahiel, M.A., *Lasso peptides from proteobacteria: Genome mining employing heterologous expression and mass spectrometry.* Biopolymers, 2013. **100**(5): p. 527-42.

86. Cheung, W.L., Chen, M.Y., Maksimov, M.O., and Link, A.J., *Lasso Peptide Biosynthetic Protein LarB1 Binds Both Leader and Core Peptide Regions of the Precursor Protein LarA*. ACS Cent Sci, 2016. **2**(10): p. 702-709.

87. Zhu, S., Fage, C.D., Hegemann, J.D., Mielcarek, A., Yan, D., Linne, U., and Marahiel, M.A., *The B1 Protein Guides the Biosynthesis of a Lasso Peptide.* Sci Rep, 2016. **6**: p. 35604.

88. Fouque, K.J., Lavanant, H., Zirah, S., Hegemann, J.D., Zimmermann, M., Marahiel, M.A., Rebuffat, S., and Afonso, C., *Signatures of Mechanically Interlocked Topology of Lasso Peptides by Ion Mobility-Mass Spectrometry: Lessons from a Collection of Representatives.* J Am Soc Mass Spectrom, 2017. **28**(2): p. 315-322.

89. Al Toma, R.S., Kuthning, A., Exner, M.P., Denisiuk, A., Ziegler, J., Budisa, N., and Sussmuth, R.D., *Site-directed and global incorporation of orthogonal and isostructural noncanonical amino acids into the ribosomal lasso peptide capistruin.* Chembiochem, 2015. **16**(3): p. 503-9.

90. Wakefield, J., Hassan, H.M., Jaspars, M., Ebel, R., and Rateb, M.E., *Dual Induction of New Microbial Secondary Metabolites by Fungal Bacterial Co-cultivation.* Front Microbiol, 2017. **8**: p. 1284.

91. Knappe, T.A., Manzenrieder, F., Mas-Moruno, C., Linne, U., Sasse, F., Kessler, H., Xie, X., and Marahiel, M.A., *Introducing lasso peptides as molecular scaffolds for drug design: engineering of an integrin antagonist.* Angew Chem Int Ed Engl, 2011. **50**(37): p. 8714-7.

92. Hegemann, J.D., De Simone, M., Zimmermann, M., Knappe, T.A., Xie, X., Di Leva, F.S., Marinelli, L., Novellino, E., Zahler, S., Kessler, H., and Marahiel, M.A., *Rational improvement of the affinity and selectivity of integrin binding of grafted lasso peptides.* J Med Chem, 2014. **57**(13): p. 5829-34.

93. Ishitsuka, M.O., Kusumi, T., Kakisawa, H., Kaya, K., and Watanabe, M.M., *Microviridin: A novel tricyclic depsipeptide from the toxic cyanobacterium Microcystis viridis.* J Am Chem Soc, 1990. **112**(22): p. 8180-8182.

94. Shin, H.J., Murakami, M., Matsuda, H., and Yamaguchi, K., *Microviridins D-F, serine protease inhibitors from the cyanobacterium Oscillatoria agardhii (NIES-204).* Tetrahedron, 1996. **52**(24): p. 8159-8168.

95. Murakami, M., Sun, Q., Ishida, K., Matsuda, H., Okino, T., and Yamaguchi, K., *Microviridins, elastase inhibitors from the cyanobacterium Nostoc minutum (NIES-26).* Phytochemistry, 1997. **45**(6): p. 1197-1202.

96. Fujii, K., Sivonen, K., Naganawa, E., and Harada, K.-i., *Non-Toxic Peptides from Toxic Cyanobacteria, Oscillatoria agardhii.* Tetrahedron, 2000. **56**(5): p. 725-733.

97. Rohrlack, T., Christoffersen, K., Hansen, P.E., Zhang, W., Czarnecki, O., Henning, M., Fastner, J., Erhard, M., Neilan, B.A., and Kaebernick, M., *Isolation, characterization, and quantitative analysis of Microviridin J, a new Microcystis metabolite toxic to Daphnia.* J Chem Ecol, 2003. **29**(8): p. 1757-70.

98. Reshef, V. and Carmeli, S., *New microviridins from a water bloom of the cyanobacterium Microcystis aeruginosa.* Tetrahedron, 2006. **62**(31): p. 7361-7369.

99. Hemscheidt, T.K., *Microviridin biosynthesis.* Methods Enzymol, 2012. **516**: p. 25-35.

100. Rohrlack, T., Christoffersen, K., Kaebernick, M., and Neilan, B.A., *Cyanobacterial protease inhibitor microviridin J causes a lethal molting disruption in Daphnia pulicaria.* Appl Environ Microbiol, 2004. **70**(8): p. 5047-50.

101. Ziemert, N., Ishida, K., Liaimer, A., Hertweck, C., and Dittmann, E., *Ribosomal synthesis of tricyclic depsipeptides in bloom-forming cyanobacteria.* Angew Chem Int Ed Engl, 2008. **47**(40): p. 7756-9.

102. Philmus, B., Christiansen, G., Yoshida, W.Y., and Hemscheidt, T.K., *Post-translational modification in microviridin biosynthesis.* Chembiochem, 2008. **9**(18): p. 3066-73.

103. Weiz, A.R., Ishida, K., Makower, K., Ziemert, N., Hertweck, C., and Dittmann, E., *Leader peptide and a membrane protein scaffold guide the biosynthesis of the tricyclic peptide microviridin.* Chem Biol, 2011. **18**(11): p. 1413-21.

104. Philmus, B., Guerrette, J.P., and Hemscheidt, T.K., *Substrate specificity and scope of MvdD, a GRASP-like ligase from the microviridin biosynthetic gene cluster.* ACS Chem Biol, 2009. **4**(6): p. 429-34.

105. Gatte-Picchi, D., Weiz, A., Ishida, K., Hertweck, C., and Dittmann, E., *Functional analysis of environmental DNA-derived microviridins provides new insights into the diversity of the tricyclic peptide family.* Appl Environ Microbiol, 2014. **80**(4): p. 1380-7.

106. Weiz, A.R., Ishida, K., Quitterer, F., Meyer, S., Kehr, J.C., Muller, K.M., Groll, M., Hertweck, C., and Dittmann, E., *Harnessing the evolvability of tricyclic microviridins to dissect protease-inhibitor interactions.* Angew Chem Int Ed Engl, 2014. **53**(14): p. 3735-8.

107. Weiz, A.R., *Characterization and manipulation of the biosynthetic pathway of cyanobacterial tricyclic microviridins in E. coli*. 2012, Humboldt University: Berlin.

108. Ziemert, N., Ishida, K., Weiz, A., Hertweck, C., and Dittmann, E., *Exploiting the natural diversity of microviridin gene clusters for discovery of novel tricyclic depsipeptides.* Appl Environ Microbiol, 2010. **76**(11): p. 3568-74.

109. Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y., *Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria.* Microbiology, 1979. **111**(1): p. 1-61.

110. Picchi, D.G., *On the characterisation of biosynthetic pathways and functional metagenomic approaches of cyanobacterial peptides*. 2013, University of Potsdam: Potsdam.

111. Niedermeyer, T.H. and Strohalm, M., *mMass as a software tool for the annotation of cyclic peptide tandem mass spectra.* PLoS One, 2012. **7**(9): p. e44913.

112. Miyamoto, K., Tsujibo, H., Hikita, Y., Tanaka, K., Miyamoto, S., Hishimoto, M., Imada, C., Kamei, K., Hara, S., and Inamori, Y., *Cloning and Nucleotide Sequence of the Gene Encoding a Serine Proteinase Inhibitor Named Marinostatin from a Marine Bacterium, Alteromonas sp. Strain B-10-31.* Bioscience, Biotechnology, and Biochemistry, 1998. **62**(12): p. 2446-2449.

113. Rosano, G.L. and Ceccarelli, E.A., *Recombinant protein expression in Escherichia coli: advances and challenges.* Front Microbiol, 2014. **5**: p. 172.

114. Khusainov, R. and Kuipers, O.P., *When the leader gets loose: in vivo biosynthesis of a leaderless prenisin is stimulated by a trans-acting leader peptide.* Chembiochem, 2012. **13**(16): p. 2433-8.

115. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.

116. Schrodinger and LLC, *The PyMOL molecular graphics system, version 1.8*. 2015.

117. Ahmed, M.N., Reyna-González, E., Schmid, B., Wiebach, V., Süssmuth, R.D., Dittmann, E., and Fewer, D.P., *Phylogenomic Analysis of the Microviridin Biosynthetic Pathway Coupled with Targeted Chemo-Enzymatic Synthesis Yields Potent Protease Inhibitors.* ACS Chemical Biology, 2017. **12**(6): p. 1538-1546.

118. Milne, J.C., Roy, R.S., Eliot, A.C., Kelleher, N.L., Wokhlu, A., Nickels, B., and Walsh, C.T., *Cofactor requirements and reconstitution of microcin B17 synthetase: a multienzyme complex that catalyzes the formation of oxazoles and thiazoles in the antibiotic microcin B17.* Biochemistry, 1999. **38**(15): p. 4768-81.

119. van den Berg van Saparoea, H.B., Bakkes, P.J., Moll, G.N., and Driessen, A.J., *Distinct contributions of the nisin biosynthesis enzymes NisB and NisC and transporter NisT to prenisin production by Lactococcus lactis.* Appl Environ Microbiol, 2008. **74**(17): p. 5541-8.

120. Wang, J., Ge, X., Zhang, L., Teng, K., and Zhong, J., *One-pot synthesis of class II lanthipeptide bovicin HJ50 via an engineered lanthipeptide synthetase.* Sci Rep, 2016. **6**: p. 38630.

121. Ozaki, T., Yamashita, K., Goto, Y., Shimomura, M., Hayashi, S., Asamizu, S., Sugai, Y., Ikeda, H., Suga, H., and Onaka, H., *Dissection of goadsporin biosynthesis by in vitro reconstitution leading to designer analogues expressed in vivo.* Nat Commun, 2017. **8**: p. 14207.

122. Andreu, D., Albericio, F., Solé, N.A., Munson, M.C., Ferrer, M., and Barany, G., *Formation of Disulfide Bonds in Synthetic Peptides and Proteins*, in *Peptide Synthesis Protocols*, Pennington, Michael W. and Dunn, Ben M., Editors. 1995, Humana Press: Totowa, NJ. p. 91-169.

123. Mayumi, T., Kato, H., Kawasaki, Y., and Harada, K., *Formation of diagnostic product ions from cyanobacterial cyclic peptides by the two-bond fission mechanism using ion trap liquid chromatography/multi-stage mass spectrometry.* Rapid Commun Mass Spectrom, 2007. **21**(6): p. 1025-33.

124. Perona, J.J. and Craik, C.S., *Structural basis of substrate specificity in the serine proteases.* Protein Sci, 1995. **4**(3): p. 337-60.

125. Del Mar, E.G., Largman, C., Brodrick, J.W., Fassett, M., and Geokas, M.C., *Substrate specificity of human pancreatic elastase 2.* Biochemistry, 1980. **19**(3): p. 468-472.

126. Hedstrom, L., *Serine protease mechanism and specificity.* Chem Rev, 2002. **102**(12): p. 4501-24.

127. Baumann, T., Nickling, J.H., Bartholomae, M., Buivydas, A., Kuipers, O.P., and Budisa, N., *Prospects of In vivo Incorporation of Non-canonical Amino Acids for the Chemical Diversification of Antimicrobial Peptides.* Frontiers in Microbiology, 2017. **8**(124).

128. Lee, H., Park, Y., and Kim, S., *Enzymatic Cross-Linking of Side Chains Generates a Modified Peptide with Four Hairpin-like Bicyclic Repeats.* Biochemistry, 2017.

129. Burkhart, B.J., Kakkar, N., Hudson, G.A., van der Donk, W.A., and Mitchell, D.A., *Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products.* ACS Cent Sci, 2017. **3**(6): p. 629-638.

Appendix

Table A.1. Oligonucleotides used in this work.

Application	Primer name	Sequence 5' ▶ 3'
MvdC cloning	P1-MvdC-BamHI_FW	GGATCCATGTCTTTGTCTCGTGATGTTGT
	P2-MvdC-PstI_RV	CTGCAGGTTAGGGAATTAGGGTATCGGC
MvdD cloning	P3-MvdD-BamHI_FW	GGATCCATGACGGTTTTAATTTTCACT
	P4-MvdD-PstI_RV	CTGCAGGCTAAAATGGTTTCTGGTTTGT
LP-MvdD cloning	P40 SacI_MvdE Fwd	GTACGCGAGCTCATGTCCAAAAATGTTAAAGTCTCTGCG
	P46 KpnI_SpeI_MvdE Rv	TAAAACCGTCATGGTACCTCCTGATCCACTAGTAGGAGCAGAATTAGAATTGTTAGCTTCC
	P47 SpeI_KpnI_MvdD Fwd	TTCTGCTCCTACTAGTGGATCAGGAGGTACCATGACGGTTTTAATTTTCACT
	P43 cPstI_MvdD Rv	TCGTAACTGCAGGCTAAAATGGTTTCTGGTTTGT
Linker for the engineered ATP-grasp ligases	P50 Link SpeI_GS15_KpnI Fwd	CTAGTGGTTCCGGCTCTGGTAGCGGCTCTGGTTCCGGCTCTGGTAGCGG CTCTGGTTCCGGCTCTGGTAGCGGCTCTGGTTCCGGCTCTGGTAGCGGTAC
	P51 Link KpnI_GS15_SpeI Rv	CGCTACCAGAGCCGGAACCAGAGCCGCTACCAGAGCCGGAACCAGAGCC GCTACCAGAGCCGGAACCAGAGCCGCTACCAGAGCCGGAACCA
LP-MvdC cloning	P40 SacI_MvdE Fwd	GTACGCGAGCTCATGTCCAAAAATGTTAAAGTCTCTGCG
	P52 KpnI_SpeI_MvdE Rv	AGACAAAGACATGGTACCTCCTGATCCACTAGTAGGAGCAGAATTAGAATTGTTAGCTTCC
	P53 SpeI_KpnI_MvdC Fwd	TTCTGCTCCTACTAGTGGATCAGGAGGTACCATGTCTTTGTCTCGTGATGTT
	P54 cPstI_MvdC Rv	TCGTAACTGCAGGTTAGGGAATTAGGGTATCGGC
MvdB cloning	P55 SacI MvdB Fwd	GTACGCGAGCTCATGGGTGATCTTGTTACCAAA
	P56 cPstI MvdB Rv	TCGTAACTGCAGGCTAAGAAGAACGCGATAATTC
Engineering ligase of *A. sagamiensis* with short LP	P57 SacI Leader (start)	GTACGCGAGCTCATGAAAACTACACCATTCTTTGC
	P58 Ligase (start) KpnI SpeI Leader Short (end)	TGCTGCCATGGTACCTCCTGATCCACTAGTACCACCAGCAGTCATTTCTAATT
	P60 SpeI KpnI Ligase (start)	ACTAGTGGATCAGGAGGTACCATGGCAGCAAATACAGTACTCATC
	P61 PstI Ligase (end)	TCGTAACTGCAGCTTAAGCCACGCCTTCATATAA
Engineering ligase of *A. sagamiensis* with long LP	P59 Ligase (start) KpnI SpeI Leader Long (end)	TGCTGCCATGGTACCTCCTGATCCACTAGTCTCTGGTACTTCTTGAACTGGAG
	P66 1F	TTCTAATACGACTCACTATAGAGCTCATGT

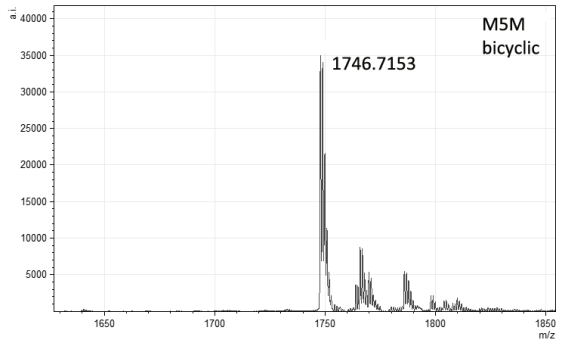
Application	Primer name	Sequence 5’ ▶ 3’
Cloning the engineered ligase of *Cyanothece* sp.	P67 2R	CCCTGGCTTTTCCGGTGTTTTTAGACATGAGCTCTATAGTGAGTCGT
	P68 3F	CCGGAAAAGCCAGGGAAATAAGAGCCGTAGTTCCTTTCTTT
	P69 4R	GCTGGATGTGCCGCTTGTTCTTCTAGAAAACGAGCAAAGAAAGGAACTACGGCTCTTAT
	P70 5F	GCGGCACATCCAGCTAATCCATTACCCACTAGTGGTTCCGGCTCT
	P71 6R	GCCGGAACCAGAGCCGCTACCAGAGCCGGAACCAGAGCCGCTACCAGAGCCGGAACCACT
	P72 7F	GCTCTGGTTCCGGCTCTGGTAGCGGCTCTGGTTCCGGCTCTGGTAGCGGTACCATGACCG
	P73 8R	AGGAATACTTTCGTTATCATTGCTAAAAGTCAAGATTAAGGCGGTCATGGTACCGC
	P74 9F	AGCAATGATAACGAAAGTATTCCTCTAGTGATGGAAGCCATTGAAGCT
	P75 10R	GATATCTGTCGGTATCAAAACGAAAGGCTTTTTCTCCTCTAGCTTCAATGGCTTCCATCA
	P76 11F	TCGTTTTGATACCGACAGATATCCCACCGAAGTGCTACTTGATATCTACTCCGGTGCTT
	P77 12R	GCCGCTTGTCGCCATCGGTAATAAATCCTCCCTCTGAAGCACCGGAGTAGATATCAAGT
	P78 13F	GGCGACAAGCGGCTAGATTTGAGTGAAGTGACCTCGGTTTGGTATCGCCGCATG
	P79 14R	TCCATACTGTCGGGCAGTTTGTGTCCGTATCTCATGCGGCGATACCAA
	P80 15F	TGCCCGACAGTATGGACAAGCAATATCGAGAAGCGGCAATTCAAGAAAGTCGCGCC
	P81 16R	TCCCGGCAGACTGGCAATTAATCCCCTAACAGAGGCGCGACCTTTCTTGAATT
	P82 17F	CCAGTCTGCCGGGATTTCATTTAGATAAAATCTCAAATGTGGACCGGGCTAATAATAAAC
	P83 18R	CCGAGTTCTCGGGCTATTTGTAACTGTAATTGTTTATTATTAGCCCGGTCCACATT
	P84 19F	AGCCCGAGAACTCGGTCTTGTCACTCCGCGCACCCTAACC
	P85 20R	CTGACACTCGGCGGCAAATTGCTTGACTGCCTCTGGATTGTTTGAGGTTAGGGTGCGCGG
	P86 21F	CCGCCGAGTGTCAGGAGCAAGGTATTATTACAAAAATGCTTTCTTCCTTTGCTATCTATG
	P87 22R	TTCTTCTCCCTGGTCCCCATAGATAGCAAAGGAAGAAAGCATT
	P88 23F	GGGACCAGGGAGAAGAAATGGTAGTTTTTACCACTCCCATTACAGAGGATGA
	P89 24R	TTTTCTAGGTCATCCTCTGTAATGGGAGTGGTAAA
	P90 25F	ATTACAGAGGATGACCTAGAAAATATGGAAGGACTGTGTTTTTGTCCGATGACCTTT
	P91 26R	ACTCTAGCGCCTTGGGGATGTTTTCCTGAAAGGTCATCGGACAAAAACACA
	P92 27F	CCCAAGGCGCTAGAGTTACGCACTACTATTGTCGGACATCAAGTATTTACCGCCGCCGTC
	P93 28R	GCCCTCTTTGCGCCAATCATAAGTCGATCCTTCCAATTTCTGGGAGTTGACGGCGGCGGT
	P94 29F	GGCGCAAAGAGGGCAAAACTTTAGTTAAAGAGTGGGAACCCTACAAATTACCGGAA
	P95 30R	AGCCATCAGTTGGAGCAATTTTTTTTCAATATCTTCCGGTAATTTGTAGGGTTC

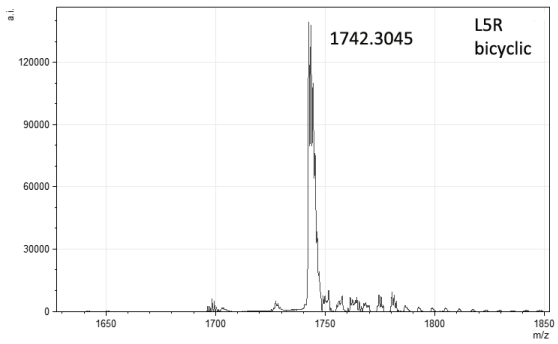
Application	Primer name	Sequence 5' ▶ 3'
	P96 31F	TGCTCCAACTGATGGCTTCTTTGGGATTAAACTATGGGGCCATTGATA
	P97 32R	GACCATCGGGGGTGACAATAATATCAATGGCCCCATAGTTTAATCC
	P98 33F	TCACCCCCGATGGTCAGCATTTATTCCTTGAGATTAACCCGGTTGGTGAGTTTTTCTGGA
	P99 34R	TGGCTTGAGAAATGGGGTATTTTGGTGAATAGATTTCCATCCAGAAAAACTCACCAACC
	P100 35F	ACCCCATTTCTCAAGCCATTGCTGAGATTCTGCTCACGGGGAAAAATTAGCCT
	P101 36R	CTGCAGGCTAATTTTTCCCCGTG
	P103 Cyanothece t701a Fwd	TGGGGACCAGGGAGAAGAAATGGTAGTTTTTACCA
	P104 Cyanothece t701a Rv	TGGTAAAAACTACCATTTCTTCTCCCTGGTCCCCA
	P105 Cyanothece t898g Fwd	GACTTATGATTGGCGCAAAGAGGGCAAAACTTTAGTTAAAGAGTG
	P106 Cyanothece t898g Rv	CACTCTTTAACTAAAGTTTTGCCCTCTTTGCGCCAATCATAAGTC

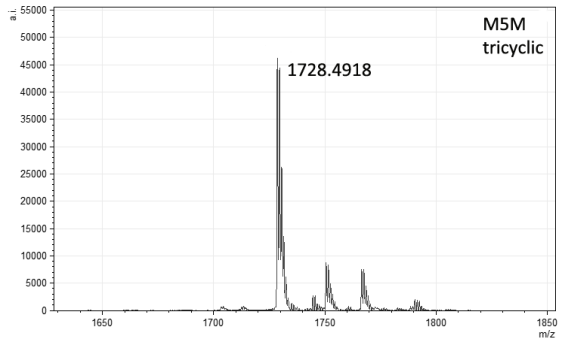
Table A.2. Results of the bicyclization assays with the Mdn K and Mdn B peptide libraries. ✓: >80% of the peptide was modified. **✗**: The modification could not be detected. The variants with Cys and Pro are marked in rectangles.

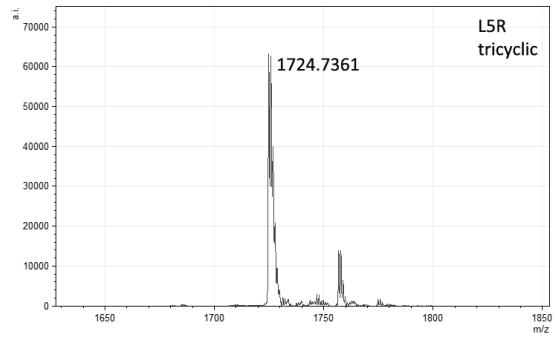
Peptide	Monocyclic intermediate detected	Bicyclization	Peptide	Monocyclic intermediate detected	Bicyclization
Mdn K library: YGNTMKYPSDWEEY			**Mdn B library:** FGTTLKYPSDWEEY		
MvdE core/ M5M	No	✓	L5L	No	✓
M5A	No	✓	L5A	Yes	✓
M5R	Yes	✓	L5R	Yes	✓
M5N	No	✓	L5N	No	✓
M5D	Yes	✓	L5D	No	< 60%
M5C	No	✗	L5C	No	< 30%
M5E	Yes	< 30%	L5E	Yes	✓
M5Q	No	✓	L5Q	No	✓
M5G	Yes	✓	L5G	Yes	✓
M5H	No	✓	L5H	No	✓
M5I	No	✓	L5I	No	✓

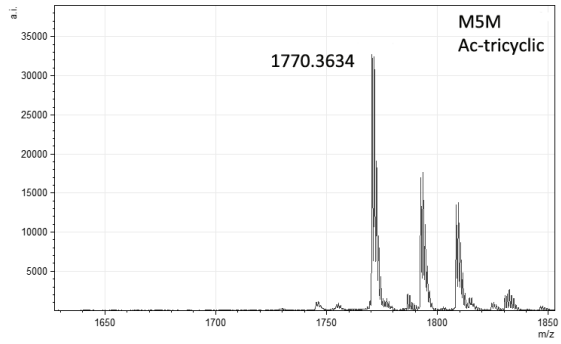
Peptide	Monocyclic intermediate detected	Bicyclization	Peptide	Monocyclic intermediate detected	Bicyclization
M5L	No	✓	L5M	Yes	✓
M5K	Yes	✓	L5K	No	✓
M5F	No	< 60%	L5F	No	✓
M5P	No	✗	L5P	No	✗
M5S	No	✓	L5S	Yes	✓
M5T	No	✗	L5T	No	✓
M5W	No	✓	L5W	No	✓
M5Y	No	✓	L5Y	Yes	✓
M5V	No	✓	L5V	Yes	✓

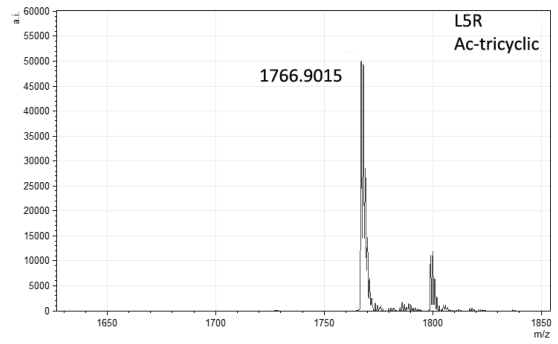


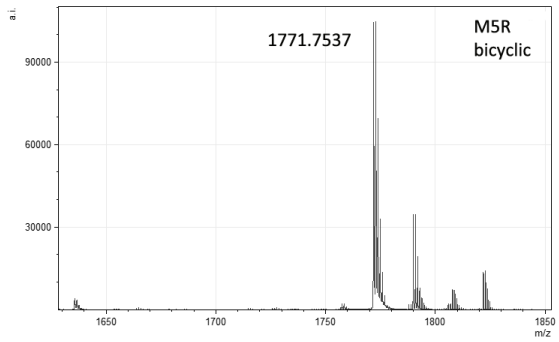


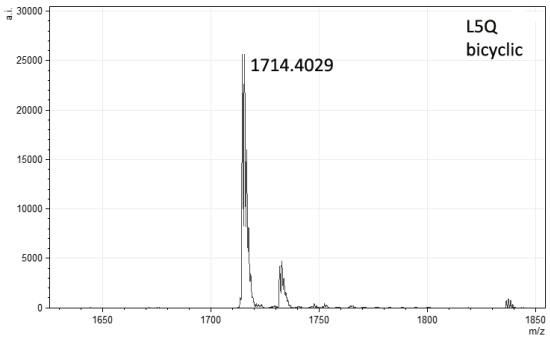


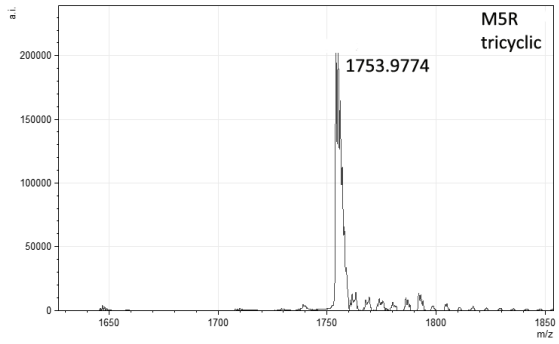


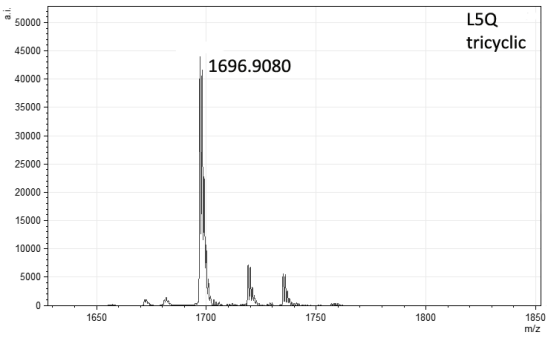


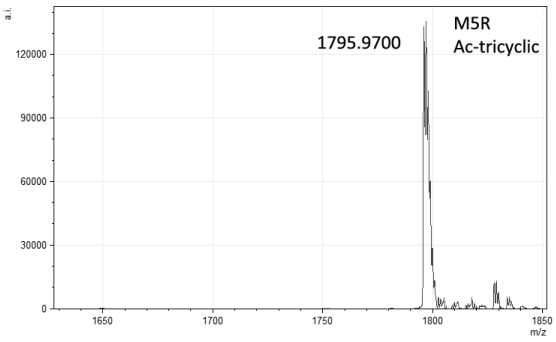


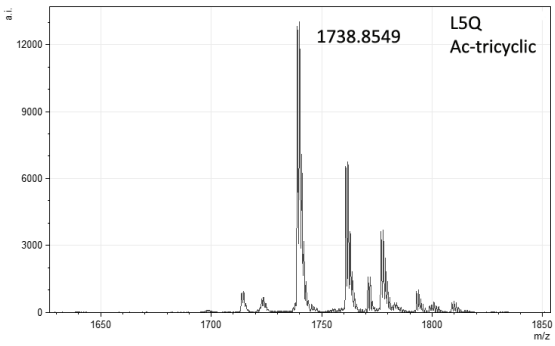

1771.7537
M5R
bicyclic
a.i.
m/z

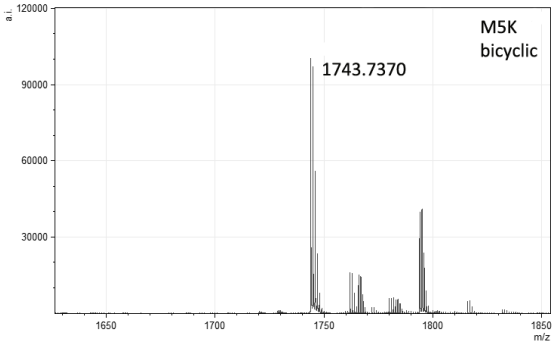

1714.4029
L5Q
bicyclic
a.i.
m/z

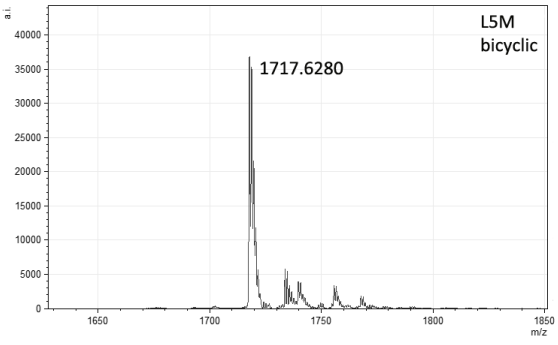

1753.9774
M5R
tricyclic
a.i.
m/z

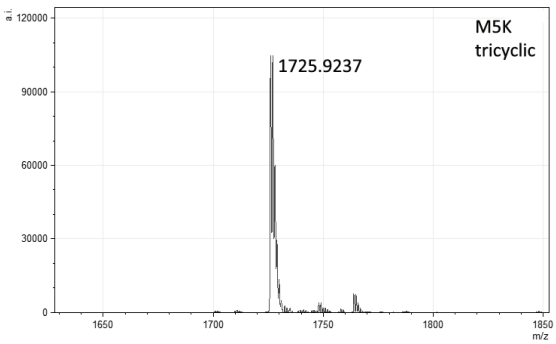

1696.9080
L5Q
tricyclic
a.i.
m/z

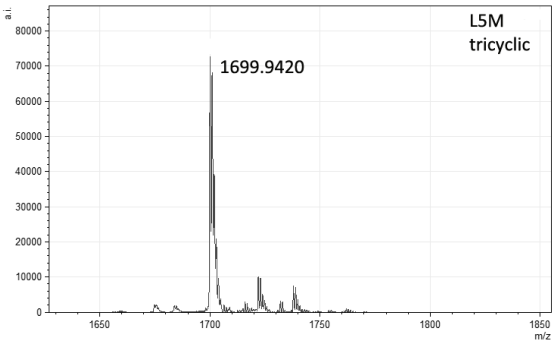

1795.9700
M5R
Ac-tricyclic
a.i.
m/z


1738.8549
L5Q
Ac-tricyclic
a.i.
m/z


1743.7370
M5K
bicyclic
a.i.
m/z


1717.6280
L5M
bicyclic
a.i.
m/z


1725.9237
M5K
tricyclic
a.i.
m/z


1699.9420
L5M
tricyclic
a.i.
m/z

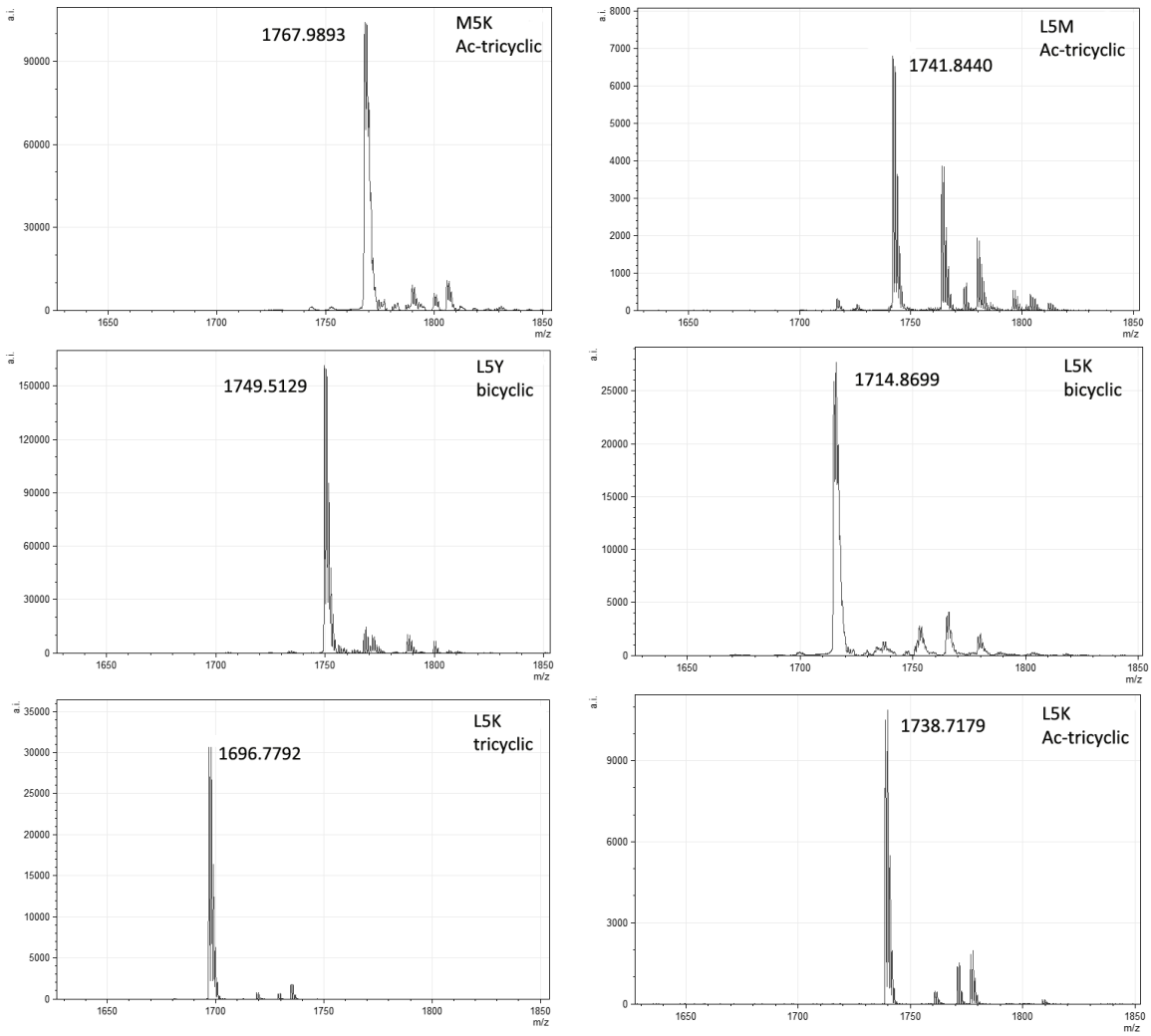


Figure A.1. MALDI-TOF MS spectra of the active microviridin variants from the Mdn K and Mdn B libraries. The mass shown is the correspondent to $[M+H]^+$.

Table A.3. Results of the PTM assays with the genome-mined microviridins. ✓: >80% of the peptide was modified. ✗: The modification could not be detected. ND: Not determined.

Peptide	Bicyclization	Tricyclization	Acetylation
Cth 1: YQNTLKYPSDWEEY	✓	<10%	<10%
Cth 2: -TVTRKYPSDWEDY	✓	✓	✓
Cth 3: -YVTKKYPSDWEEY	✓	✓	✓
Cth 4: --ATLKYPSDWEEY	✓	<10%	✗
Cth 5: --FTLKYPSDWEDE	✗	ND	ND
Mst: FATMRYPSDSDE	✗	ND	ND
MvdF core: TIWTFKWPSDWEDS	✗	ND	ND

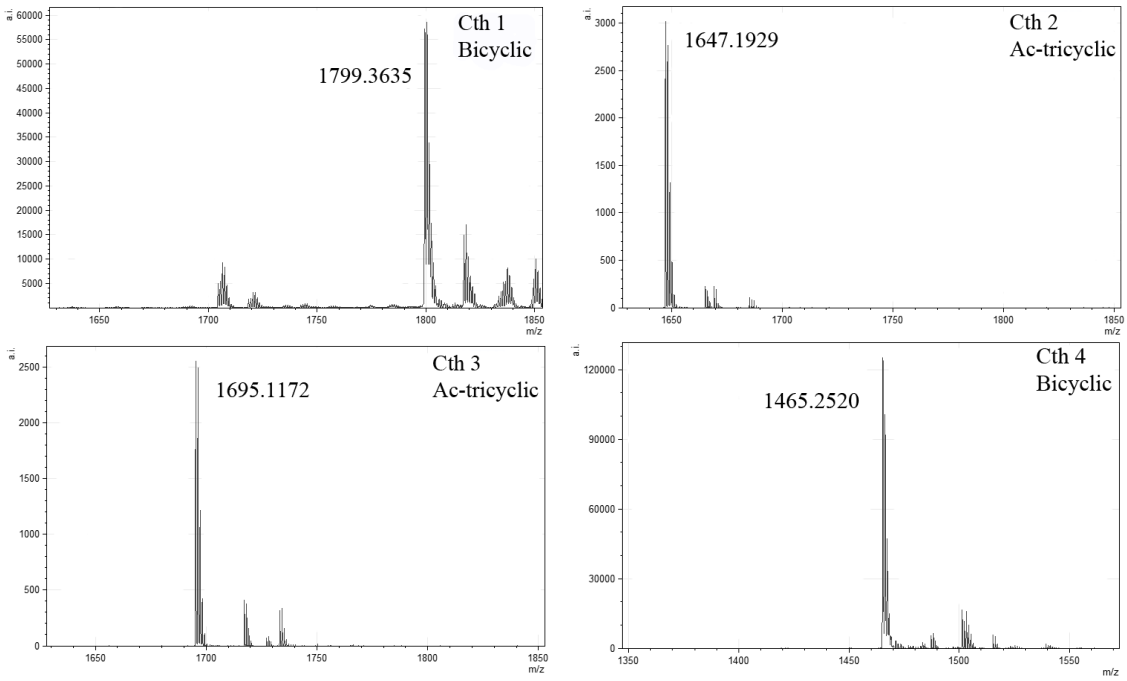


Figure A.2. MALDI-TOF MS spectra of the active variants from the library of genome-mined microviridins. The mass shown is the correspondent to $[M+H]^+$.

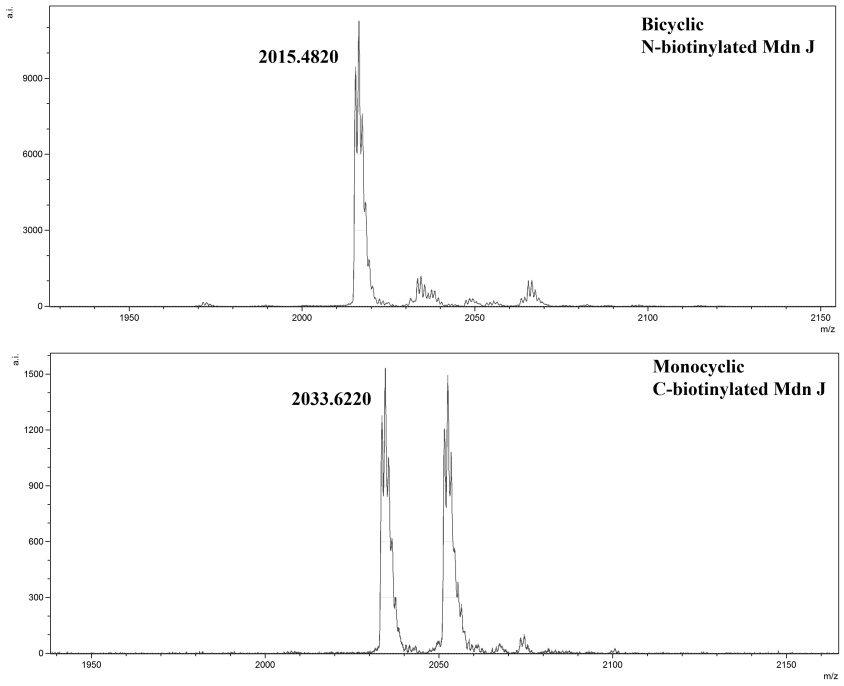


Figure A.3. MALDI-TOF MS spectra of the active microviridin variants carrying a biotinylated Lys. The mass shown is the correspondent to $[M+H]^+$.

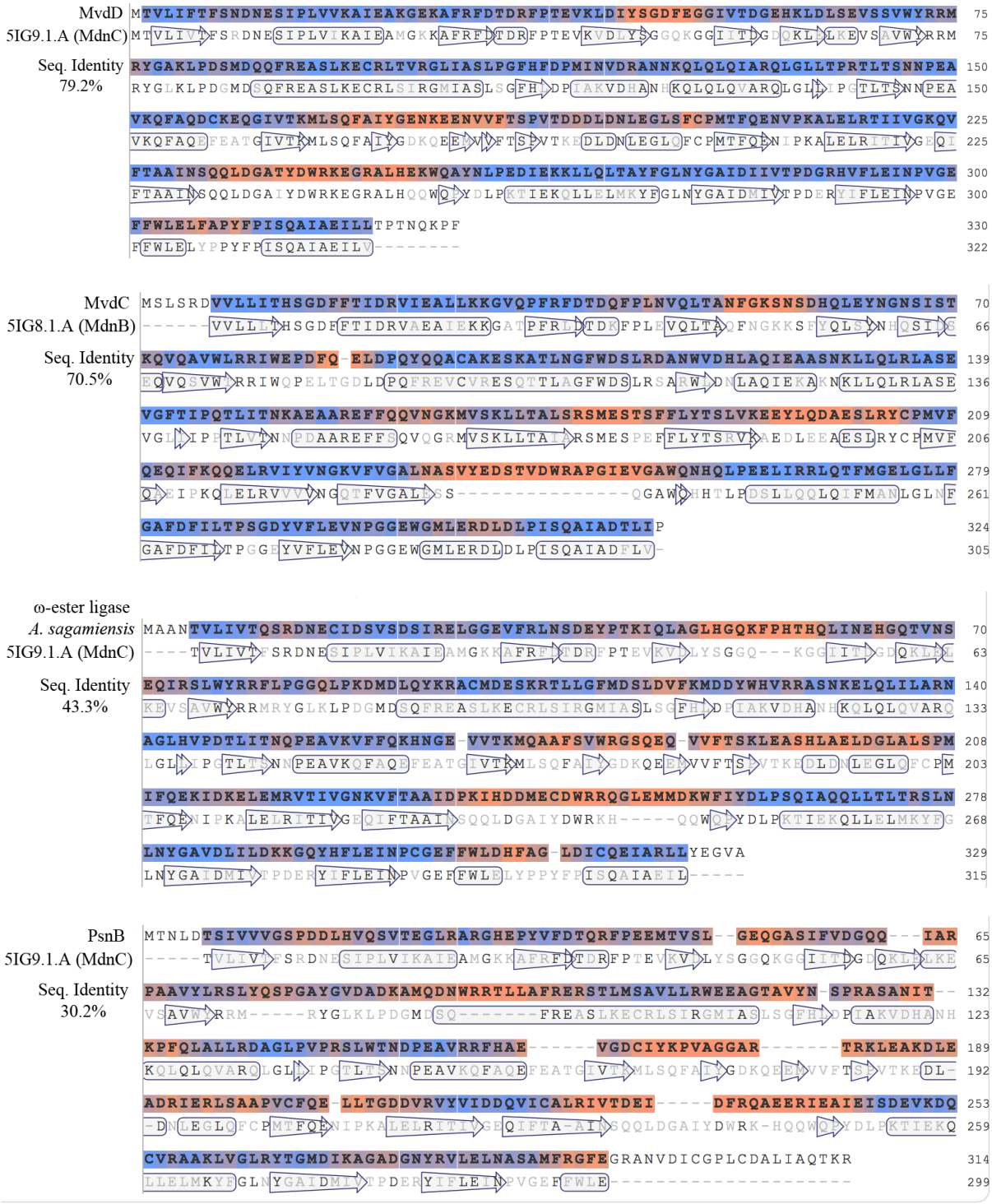


Figure A.4. Sequence alignment of different ATP-grasp ligases with MdnC and MdnB. The alignments were carried out by Swiss Model to create homology models of ATP-grasp ligases from *P. agardhii*, *A. sagamiensis*, and *P. pacifica* from the crystal structures of the ω-ester ligase MdnC (PDB file 5IG9) and the ω-amide ligase MdnB (PDB file 5IG8) from *M. aeruginosa*. The sequence identity is indicated for each alignment. Predicted α-helices and β-strands are indicated in rectangles and arrows, respectively. Motifs in blue exhibit higher homology and motifs in orange exhibit lower homology.

Publications

Parts of this work have already been published:

- "Phylogenomic analysis of the microviridin biosynthetic pathway coupled with targeted chemo-enzymatic synthesis yields potent protease inhibitors." MN. Ahmed*, E. Reyna-González*, B. Schmid, V. Wiebach, RD. Süssmuth, E. Dittmann, DP. Fewer, *ACS. Chem. Biol.* **2017**, *12* (6), 1538-1546. *First authors.
- "Leader peptide-free in vitro reconstitution of microviridin biosynthesis enables design of synthetic target- ed libraries." E. Reyna-González, B. Schmid, D. Petras, RD. Süssmuth, E. Dittmann, *Angew. Chem. Int. Ed.* **2016**, *55*, 9398-9401.

Acknowledgments

I would like to express my most sincere thanks to the following people because, in one way or another, they helped me through this endeavor:

Professor Elke Dittmann, for accepting me in her group in the first place, for enthusiastically supervising the project and in general for her friendly support.

The Cyanos, because we collaborated and helped each other out as a team. It would not be possible to successfully complete such a project without organized team work.

Professor Roderich Süßmuth, Bianca Schmid, Dr. Daniel Petras, and Vincent Wiebach for the sample preparation, sample analysis and contribution of ideas to the project.

Professor David Fewer and Muhammad Ahmed for a successful collaboration.

Dr. Arthur Guljamow for proof-reading the dissertation and for the alemañol tandem.

Tim Schwier, for contributing with his Cambridge English to the proof-reading of the dissertation.

Professor Silke Leimkühler, for allowing me to use the lab facilities of her group when needed.

Dr. Jörg Fettke, for allowing me to use the MALDI equipment of his group.

Dr. Ines Starke, for sample analysis.

Yasemin Mai-Linde, for weighing the microviridin variants.

Luciano Tepper, for providing me with Endnote.

My friends in Germany, who always supported me in all kind of situations, in particular Bárbara Ocampo, Emma Paredes, Diana Alatriste, Fabián Barrera, Duban Canal, and the Spaniards.

My friends in Mexico, who despite the distance still offer me their support, in particular Richard Alvarez for helping me with Photoshop.

Finally my family, which has always supported me unconditionally.