EXPANDING THE GENETIC AND CHEMICAL BIOSYNTHETIC DIVERSITY OF CYANOBACTIN FAMILY OF NATURAL PRODUCTS

Antti Mattila

Department of Microbiology

Faculty of Agriculture and Forestry

University of Helsinki, Finland

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, for public examination in Auditorium 2041 at Biocenter 2, Viikinkaari 5, Helsinki on November 27th 2020, at 12 o'clock noon.

Helsinki 2020

Supervisors:	Principal supervisor Docent David Fewer Department of Microbiology University of Helsinki, Finland
	Professor Kaarina Sivonen Department of Microbiology University of Helsinki, Finland
Pre-examiners:	Professor Marcel Jaspars Marine Biodiscovery Centre Department of Chemistry University of Aberdeen Scotland, UK
	Docent Jarmo Niemi Department of Biochemistry University of Turku, Finland
Thesis committee:	Professor Per Saris Department of Microbiology University of Helsinki, Finland
Opponent:	Professor Nadine Ziemert Interfaculty Institute of Microbiology and Infection Medicine Institute for Bioinformatics and Medical Informatics German Centre for Infection Research (DZIF) Partnersite Tübingen, Germany
Custos:	Professor Kaarina Sivonen Department of Microbiology University of Helsinki, Finland

Dissertationes Schola Doctoralis Scientiae Circumiectalis, Alimentariae, Biologicae

ISSN 2342-5423 (Print)

ISSN 2342-5431 (Online)

ISBN 978-951-51-6766-8 (Paperback)

ISBN 978-951-51-6767-5 (PDF)

http://ethesis.helsinki.fi

Cover image: General workflow of natural product discovery: From genome to molecule (chemical structure of muscoride B).

Painosalama
Turku 2020

Table of Contents

List of original publications 4
Related publications 4
Abbreviations 5
Abstract 6
Tiivistelmä 7
1 Introduction 8
1.1 Natural products 8
1.1.1 Biosynthesis of natural products 9
1.1.2 Ribosomally synthesized and post-translationally modified peptides (RiPPs) 9
1.1.3 Genome mining for new natural product biosynthetic pathways 11
1.2 Cyanobacteria and their specialized metabolites 12
1.3 Cyanobactins 14
1.3.1 Chemical diversity of cyanobactins 14
1.4 Cyanobactin biosynthesis 16
1.4.1 General biosynthetic pathway 16
1.5 Enzymes involved in cyanobactin biosynthesis 18
1.5.1 Proteases 18
1.5.2 Heterocyclase 19
1.5.3 Oxidase 20
1.5.4 Prenyltransferase 20
1.5.5 Proteins and domains of unknown functions in cyanobactin BGCs 23
1.6 Ecological role of cyanobactins and bioactivity 24
1.7 Antimicrobial peptides (AMPs) 25
1.7.1 Disulfide bridge formation 26
1.7.2 α-helical peptides 29
2 Study aims 30
3 Summary of materials and methods 30
4 Summary of results and discussion 31
4.1 Discovery and biosynthesis of peptide alkaloids muscoride A and B (I) 32
4.2 Discovery of truncated cyanobactin BGC and unusual cyanobactins (II) 33
4.3 Biosynthesis of prenylated anacyclamide (III) 38
5 Conclusions 39
6 Acknowledgements 40
7 References 41

List of original publications

This thesis is based on the following publications, which are referred to in the text by their Roman numerals:

I **Antti Mattila**, Rose-Marie Andsten, Mikael Jumppanen, Michele Assante, Jouni Jokela, Matti Wahlsten, Kornelia M. Mikula, Cihad Sigindere, Daniel H. Kwak, Muriel Gugger, Harri Koskela, Kaarina Sivonen, Xinyu Liu, Jari Yli-Kauhaluoma, Hideo Iwaï, David P. Fewer. 2019. Biosynthesis of the bisprenylated alkaloids muscoride A and B. *ACS Chemical Biology.* 14 (12): 2683–2690.

II **Antti Mattila**, Niina Leikoski, Matti Wahlsten, Jouni Jokela, Pavel Hrouzek, Petra Kučerová, Muriel Gugger, Kaarina Sivonen, David P. Fewer. 2020. Discovery of unusual cyanobactins by genome mining. Manuscript.

III Luca Dalponte, Anirudra Parajuli, Ellen Younger, **Antti Mattila**, Jouni Jokela, Matti Wahlsten, Niina Leikoski, Kaarina Sivonen, Scott A. Jarmusch, Wael E. Houssen, David P. Fewer. 2018. N-prenylation of tryptophan by an aromatic prenyltransferase from the cyanobactin biosynthetic pathway. *Biochemistry*. 57 (50): 6860–6867.

Author contribution

I Antti Mattila participated in the design of the study. He executed most of the experiments, interpreted the results and wrote the manuscript.

II Antti Mattila participated in the design of the study. He executed most of the experiments, interpreted the results and wrote the manuscript.

III Antti Mattila participated in the design of the study. He executed some of the experiments, interpreted some of the results and contributed to the writing of the manuscript.

Related publications

IV Jan Mareš, Jan Hájek, Petra Urajová, Andreja Kust, Jouni Jokela, Kumar Saurav, Tomas Galica, Kateřina Čapková, **Antti Mattila**, Esa Haapaniemi, Perttu Permi, Ivar Mysterud, Olav Skulberg, Jan Karlsen, David Fewer, Kaarina Sivonen, Hanne Tonnesen, Pavel Hrouzek. 2019. Alternative biosynthetic starter units enhance the structural diversity of cyanobacterial lipopeptides. *Applied and Environmental Microbiology*. 85 (4): e02675-18.

Abbreviations

AMP	antimicrobial peptide
BGC	biosynthetic gene cluster
BLAST	basic local alignment search tool
bp	base pair
DSMZ	Deutsche Sammlung von Mikro-organismen und Zellkulturen
HAMBI	University of Helsinki Culture Collection
JCM	Japan Collection of Microorganisms
LC-MS	liquid chromatography-mass spectrometry
m/z	mass to charge ratio
NMR	nuclear magnetic resonance
NRPS	non-ribosomal peptide synthetase
ORF	open reading frame
PCC	Pasteur Culture Collection of Cyanobacteria
PCR	polymerase chain reaction
PKS	polyketide synthase
PTM	post-translational modification
Q-TOF	quadrupole time-of-flight
RiPP	ribosomally synthesized and post-translationally modified peptide
TDOR	thiol-disulfide oxidoreductase
UHCC	University of Helsinki Cyanobacterial Culture Collection

Abstract

Natural products are small metabolites produced by living organisms with versatile chemical structures and diverse bioactivities. Bacteria are a rich but yet underexplored source of natural products. Natural product biosynthetic pathways encode a huge array of biosynthetic enzymes which can catalyze stereospecific and regiospecific modification of intermediates to produce complex natural products from simple precursors. The biosynthetic mechanisms underlying the production of natural products vary substantially. Cyanobactins are synthesized through the post-translational modification of precursor peptides and represent one of the most abundant natural product classes produced by cyanobacteria. Cyanobactins have a macrocyclic or linear structure with antimicrobial and cytotoxic bioactivities, although their ecological function is unclear.

In this study, genome mining was used to catalog the genetic diversity of cyanobactin biosynthetic pathways from bacterial and archaeal genomes. Bioinformatic analysis of microbial genomes was used to identify cyanobactin biosynthetic pathways. New cyanobactin biosynthetic pathways were studied using a combination of bioinformatics, molecular biology, microbiology, biochemistry and structural chemistry. Bioinformatic predictions of cyanobactin biosynthetic gene clusters were performed based on DNA sequence data. Subsequently, candidate strains identified based on the predictions were grown in the laboratory and subjected to mass spectrometric analysis and heterologous expression. Novel cyanobactin natural products were discovered and the enzymes underpinning their biosynthesis were characterized.

In the first part of the study, the muscoride biosynthetic gene cluster was described and a new muscoride variant was discovered. Two specific prenyltransferases from the muscoride biosynthetic gene cluster, which were biochemically characterized, catalyzed the regiospesific prenylation of the N- and C-termini of the linear polyoxazole muscorides. In the second part of the study, genome mining revealed a truncated cyanobactin biosynthetic gene cluster across the bacterial and archaeal domains for the production of structurally unusual cyanobactins. In the third part of the study, a regiospecific prenyltransferase was discovered, belonging to the anacyclamide cyanobactin pathway and catalyzing the N-prenylation of the tryptophan residue of the substrate peptide.

The main aim of this work was to uncover novel cyanobactin biosynthetic gene clusters from bacteria and expand knowledge regarding the biosynthetic potential of cyanobactin pathways. The results from this thesis broaden the chemical diversity of the cyanobactin family and expand upon the biosynthetic logic underlying cyanobactin biosynthesis.

Tiivistelmä

Luonnonyhdisteet ovat molekyylejä, joita tuottavat kaikki elävät organismit. Näillä yhdisteillä on hyvin monipuolisia rakenteita ja bioaktiivisuuksia. Bakteerit ovat runsas, mutta silti suurelta osin kartoittamaton, lähde luonnonyhdisteille. Luonnonyhdisteiden biosynteesireitit koodaavat valtavan määrän biosynteettisiä entsyymejä, jotka katalysoivat erilaisia muokkausreaktioita tuottaakseen monimutkaisia rakenteita omaavia lopputuotteita. Syanobaktiinit ovat luonnonyhdistepeptidejä, jotka syntetisoidaan translaation jälkeisillä muokkausreaktioilla ja ne ovat yksi suurimmista syanobakteerien tuottamista luonnonyhdisteluokista. Syanobaktiinit ovat kemiallisilta rakenteiltaan syklisiä tai lineaarisia ja niillä on antimikrobisia sekä sytotoksisia bioaktiivisuuksia, vaikkakin niiden ekologinen rooli on vielä epäselvä.

Tässä väitöskirjatutkimuksessa tutkittiin bioinformatiivisella genomilouhinnalla syanobaktiinien biosynteesireittien esiintyvyyttä bakteerien ja arkeonien genomeissa. Uusia syanobaktiinien biosynteesireittejä tutkittiin käyttämällä hyödyksi bioinformatiikkaa, molekulaarista biologiaa, mikrobiologiaa, biokemiaa sekä rakenteellista kemiaa. Syanobaktiineja tuottavat geenit ovat yleensä toistensa lähellä genomissa muodostaen geeniklusterin ja näiden geenien ennustukset tehtiin bakteerien DNA:n sekvensointitietoon perustuen. Kandidaattibakteerikantoja kasvatettiin laboratoriossa ja kasvatuksen jälkeen niitä tutkittiin massaspektrometrialla sekä heterologisella ekspressiolla, jotta saataisiin selville tuottavatko kyseiset bakteerikannat bioinformatiikalla ennustettuja luonnoyhdisteitä. Tässä työssä löydettiin uusia syanobaktiineja, sekä tunnistettiin näiden biosynteesissä tarvittavia entsyymejä.

Työn ensimmäisessä osassa tunnistettiin kaksi biosynteettistä geeniklusteria luonnonyhdiste muskoridin tuottamiseen, sekä pystyttiin rakenteellisesti kuvaamaan uusi muskoridi variantti. Muskoridin biosynteettisessä geeniklusterissa on kaksi biokemiallisesti karakterisoitua spesifistä prenyylitransferaasientsyymiä, jotka katalysoivat lineaarisen muskoridin amino- ja karboksiterminaalien prenylaatiota. Työn toisessa osassa bioinformatiivisella genomilouhinnalla löydettiin bakteerien ja arkeonien genomeista tunnustettua lyhyempiä syanobaktiinien biosynteettisiä geeniklustereita, joiden ennustetut lopputuotteet ovat rakenteellisesti uudenlaisia syanobaktiineja. Työn kolmannessa osassa anasyklamidin biosynteettisesta geeniklusterista löydettiin spesifinen prenyylitransferaasientsyymi, joka liittää prenyyliryhmän substraattipeptidin tryptofaanin indolin typpiatomiin.

Tämän väitöskirjatutkimuksen päätarkoituksena oli löytää ja karakterisoida uusia syanobaktiinien biosynteettisia geeniklustereita bakteereista sekä laajentaa tietämystä niiden biosynteettisesta potentiaalista. Tämän väitöskirjan tulokset ovat laajentaneet syanobaktiinien kemiallista monimuotoisuutta sekä monipuolistanut tietämystä niiden entsymaattisesta biosynteesistä.

1 Introduction

1.1 Natural products

Natural products are specialized metabolites that are produced by secondary metabolism in a wide variety of living organisms (Demain and Sanchez 2009; Jensen 2016; Katz and Baltz 2016). These secondary metabolites exhibit high structural and chemical diversity (Berdy 2005; Jensen 2016). They have different and usually essential functions, conferring a fitness advantage under specific environmental conditions, but they are not required for growth and reproduction under typical conditions (Jensen 2016; Katz and Baltz 2016; Walsh and Tang 2017). Secondary metabolites from microorganisms can also influence the survival and performance of other organisms (Demain and Sanchez 2009). Bacteria often produce secondary metabolites to fight against competing microorganisms in their natural habitat and for bacterial interactions (Baltz 2008; Li and Rebuffat 2019). Natural products display a tremendous diversity of different bioactivities, and many pharmaceuticals and chemically synthesized drugs used today owe their origin to natural products (Demain and Sanchez 2009; Newman and Cragg 2016). Microorganisms are an abundant source of secondary metabolites, many of these being peptides with exotic structures (Donia et al. 2008; Arnison et al. 2013). Peptides are an important class of natural products which are used as antibiotics as well as anti-tumour and antiviral agents (Gurevich et al. 2018). At present, natural products or their derivatives, in general, represent an important source of novel therapeutic agents and pharmaceuticals (Newman and Cragg 2016; Ziemert et al. 2016).

The major classes of natural products include peptides, polyketides, isoprenoid/terpenoid scaffolds, alkaloids, purines and pyrimidines, phenylpropanoid scaffolds, and flavonoids (Walsh and Tang 2017). Natural products are relevant in nature as toxins (Du et al. 2019), antimicrobial agents (Lewis 2017), siderophores (Kramer et al. 2019), and signaling molecules (Dutta et al. 2019) to name but a few examples. **A principal driver for natural product research is the discovery of new natural products with the potential to act as drug leads**. The increasing prevalence of antibiotic resistance along with a steady decrease in the discovery of novel antibiotics is threatening the prevention and treatment of infections caused by different pathogens worldwide (Mookherjee et al. 2020). To battle this crisis, the pharmaceutical industry is in constant need of new drug candidates. Bacteria are a voluminous source of natural products, many of which are potential drug leads, and microbial genomes are widely believed to encode new antimicrobial compounds with utility for the pharmaceutical industry and biotechnology (Newman and Cragg 2016). Key steps in natural product research include the identification of biosynthetic genes from producer organisms, which are usually clustered in close proximity to each other in the genome (biosynthetic gene clusters, BGSs); structural determination of the end-products from pathways; activity testing of the enzymes involved in the biosynthesis of natural products; and bioactivity testing of the compounds. The process of natural product discovery is highly likely to provide novel enzymes and molecules for use in biotechnology and other industries.

1.1.1 Biosynthesis of natural products

Polyketides and nonribosomal peptides are synthesized by large specific multifunctional enzymes: polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) (Finking and Marahiel 2004; Kopp and Marahiel 2007; Katz and Baltz 2016). Nonribosomal peptides and polyketides are synthesized by protein machinery via a thiotemplate chemical logic assembly line (Walsh and Tang 2017). NRPSs are composed of modules, in turn, composed of domains that incorporate one selected amino acid into the polypeptide chain under synthesis (Katz and Baltz 2016). PKSs incorporate, in a processive manner, specific acyl-CoA units into a growing polyketide chain which is attached to the enzyme as a thioester until set free at the end of synthesis (Katz and Baltz 2016; Walsh and Tang 2017). These type I PKSs also consist of modules which themselves are composed of enzymatic domains catalyzing a single step in the biosynthesis of the polyketide backbone (Katz and Baltz 2016). Microbes frequently use hybrid nonribosomal-polyketide assembly lines for the biosynthesis of specialized metabolites (Walsh and Tang 2017). The other major PKS classes utilize a different assembly line logic (Walsh and Tang 2017). Peptides mediate diverse biological processes in all domains of life (Li and Rebuffat 2019) and are produced non-ribosomally (NRPSs) (Marahiel et al. 1997) or through ribosomal synthesis and post-translational modification(s) (RiPPs) (Arnison et al. 2013).

1.1.2 Ribosomally synthesized and post-translationally modified peptides (RiPPs)

Peptides, structurally diverse natural products, are now united into a single natural product class by their ribosomally synthesized and post-translationally modified peptide pathway (RiPP) (Arnison et al. 2013) apart from polyketides and non-ribosomal peptides. These peptides are found in all three domains of life: Bacteria, Eukarya and Archaea (Arnison et al. 2013). RiPPs encompass a vast genetic, functional and chemical space, and are related by a universal biosynthetic machinery shared by more than twenty known RiPP families (Arnison et al. 2013; Dunbar and Mitchell 2013; Ortega and van der Donk 2016; Hetrick and van der Donk 2017). RiPPs serve as a useful platform for novel natural products, unusual chemistry and new enzymology (Arnison et al. 2013; Ortega and van der Donk 2016; Hetrick and van der Donk 2017), and they exhibit many roles in the physiology and social behavior of microorganisms (Li and Rebuffat 2019). RiPPs have been shown *in vitro* and/or *in vivo* to be involved in communication, competition, and diverse physiological functions such as morphological development and biofilm formation (Li and Rebuffat 2019). Compounds resulting from this group of natural products are chemically highly diverse with a multitude of bioactivities (Arnison et al. 2013; Ortega and van der Donk 2016). The number of natural products belonging to the RiPP class is rapidly increasing, each adorned with unique chemical features (Arnison et al. 2013; Ortega and van der Donk 2016). The genetic coding of RiPP BGCs enables detailed structural predictions of the end-product (Skinnider et al. 2016), especially when the genes encoding for post-translational modification (PTM) enzymes are annotated (Hetrick and van der Donk 2017). In general, RiPP gene clusters are small (Arnison et al. 2013; Ortega and van der Donk 2016; Santos-Aberturas et al. 2019). The genes encoding precursor peptides are often short and frequently unannotated in microbial genomes (Hetrick and van der Donk 2017), which makes them challenging to identify by genome mining (Santos-Aberturas et al. 2019). The discovery of numerous new RiPPs encoded within sequenced bacterial genomes may be possible by genome mining (Ortega and van der Donk 2016; Skinnider et al. 2016), although at present the identification of undiscovered RiPPs, or even possibly complete RiPP families, is still an unsolved problem and challenge (Santos-Aberturas et al. 2019).

RiPP BGCs can be recognized by the presence of specific modifying enzymes together with a precursor protein that incorporates sequence motifs distinctive to a particular RiPP family (Santos-Aberturas et al. 2019). The end-product from a RiPP pathway can be accurately predicted from the genome sequence once the biosynthetic logic of the gene cluster is understood, which grants RiPP pathways potential for bioengineering (Velázquez and van der Donk 2011). RiPP BGCs are easily overlooked because of their small size (Hetrick and van der Donk 2017). However, the methods for RiPP discovery have been revolutionized due to the massive increase in publicly available genome sequence data (Bachmann et al. 2014; Ziemert et al. 2016). In addition to sequence data, advances in understanding biosynthesis and structural chemistry are important for the discovery of new RiPPs.

RiPP BGCs comprise precursor gene(s) flanked by genes encoding post-translational enzymes which modify the precursor peptide (Arnison et al. 2013; Figure 1). The ribosomally translated precursor peptide consists of a leader peptide and a core peptide (Figure 1). Eukaryotic RiPPs often also contain an N-terminal signal sequence before the leader peptide that guides the peptide to specialized compartments of the cell for modification and secretion (Arnison et al. 2013). C-terminal recognition sequences after the core peptide(s) (i.e. follower peptide) are also occasionally present (McIntosh et al. 2009; Li and Rebuffat 2019; Figure 1) for cleavage and macrocyclization, for example, in cyanobactins and cyclotides (Arnison et al. 2013). The leader sequence is usually conserved, and the enzymes which modify the core typically recognize sequences in the leader (Arnison et al. 2013; Sardar et al. 2015). The leader peptides are believed to have numerous roles in post-translational modifications (Arnison et al. 2013), and the core peptide encoded into the precursor peptide is eventually produced to form the mature RiPP through post-translational modifications (Figure 1). The leader peptide might aid in the export of the core peptide outside from the cell and might also have a role in the activation, folding and stabilization of the bound modifying enzyme (Ortega and van der Donk 2016). The importance of the leader peptide has been demonstrated by deletion of the leader peptide abolishing core peptide modifications (Ortega and van der Donk 2016). A RiPP BGC can contain one or multiple precursor proteins, and one precursor protein can include multiple core peptides, which are all modified. In general, RiPP post-translational enzymes show high plasticity and substrate tolerance (Hudson and Mitchell 2018). As a demonstration of RiPP enzyme flexibility, a promiscuous prochlorocin biosynthetic enzyme from *Prochlorococcus* MIT9313 has been shown to have the capacity to transform 29 distinct gene-derived linear RiPPs to polycyclic RiPPs with remarkably varying ring topologies (Li et al. 2010).

Post-translational modifications include, among others, miscellaneous cyclizations, dehydrations, rearrangements, terminal capping, hydroxylations, oxidations, phosphorylations, glycosylations, reductions, targeting moieties, as well as crosslinks between residues via oxazole, thiazole, lanthionine and disulfide bridges (Arnison et al 2013; Wang 2014). The end-products of RiPP pathways can achieve a high degree of chemical variety via extensive PTMs and structural variation encoded in the core (Arnison et al. 2013; Ortega and van der Donk 2016). All these factors endow the mature peptide with a rigidified structure and bioactivity (Dunbar and Mitchell 2013).

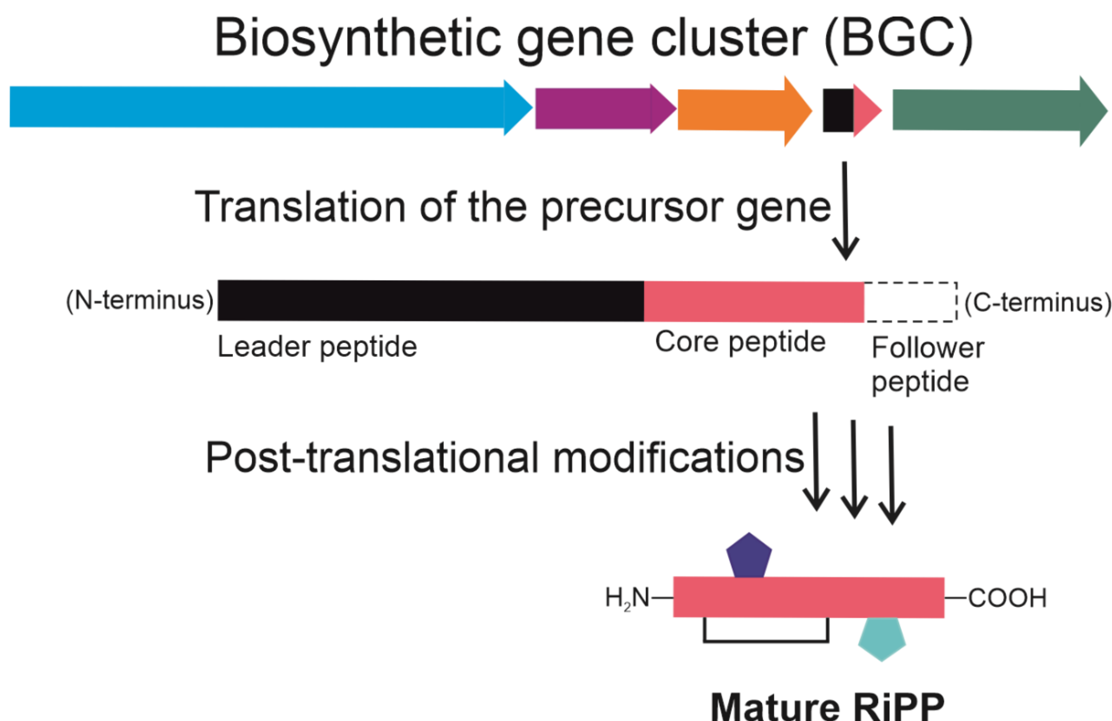


Figure 1. General biosynthetic pathway for RiPPs. The gene encoding for the precursor peptide is flanked by genes encoding post-translational enzymes. The precursor peptide contains a core sequence that is transformed to the mature natural product via post-translational modifications. The figure has been adapted from Arnison et al. (2013) and Tan et al. (2019).

1.1.3 Genome mining for new natural product biosynthetic pathways

Natural product discovery, the search for new compounds from nature, can be divided into two distinct categories, known as structure-guided (top-down) and genome-guided (bottom-up) strategies. The structure-guided approach denotes natural product discovery without genomic data and no prior knowledge regarding the biosynthesis of the secondary metabolite in question (Luo et al. 2014). Structure-guided methods include comparative metabolomics profiling, fraction bioactivity testing, and structural characterization (Luo et al. 2014). The classic structure-guided approach for natural product discovery usually includes screening of crude extracts from natural sources, such as bacteria (Winter et al. 2011). However, this method is often problematic in the search for novel compounds because of the high probability to rediscover known natural products (Winter et al. 2011). In the genome-guided approach, genomics and bioinformatics are utilized in the search for new biosynthetic pathways of natural products in a process called genome mining (Luo et al. 2014; Ziemert et al. 2016). The rapid accumulation of genome sequence data in public databases combined with advancements in the understanding of natural product biosynthesis has enabled unprecedented opportunities for the discovery of novel natural products from a broad diversity of organisms (Challis 2008; Velázquez and van der Donk 2011; Bachmann et al. 2014; Jensen 2016; Ziemert et al. 2016; Kalkreuter et al. 2020). The huge quantity of accessible genome data serves as the base for the examination and discovery of new specialized metabolites (Zerikly and Challis 2009) through comparative genomic approaches (Micallef et al. 2015). Although the diversity of secondary metabolites between and within different families is vast, the biosynthetic machineries are often conserved (Ziemert et al.

2016). Biosynthetic gene clusters within microorganisms encode enzymatic machineries that catalyze the production of chemically complex natural products, and many of these have been repurposed as pharmaceutical, agricultural, and manufacturing agents (Epstein et al. 2018).

Genome mining is a powerful tool in the search for new secondary metabolites, as it predicts the biosynthesis of new natural products directly from genome sequences (Zerikly and Challis 2009) and has the theoretical potential to eliminate the rediscovery problem (discovery of already known compound) (Bachmann et al. 2014). Genome mining includes the examination of the sequenced genome of a microorganism to infer the presence of gene clusters involved in the production of novel natural products (Scheffler et al. 2013). Genome mining can be further combined with modern spectroscopy techniques, such as liquid chromatography-mass spectrometry (LC-MS) and quadrupole time-of-flight (Q-TOF), to identify the products of novel gene clusters (Scheffler et al. 2013) which can eventually be used for heterologous expression and bioactivity testing (Micallef et al. 2014). In general, genome mining can be performed manually or using bioinformatic tools, such as antiSMASH (Weber et al. 2015), BAGEL3 (van Heel et al. 2013), and PRISM (Skinnider et al. 2016). Genome mining specifically for RiPPs can also be performed either manually or using bioinformatic tools, such as RiPPquest (Mohimani et al. 2014), RODEO (Tietz et al. 2017), RiPPER (Santos-Aberturas et al. 2019) or the recently developed tool DeepRiPP (Merwin et al. 2020). RiPPquest combines bioinformatics and tandem mass spectrometry in the search for new RiPPs (Mohimani et al. 2014). DeepRiPP employs multiomics and machine learning for the automated discovery for novel RiPPs (Merwin et al. 2020). RODEO uses machine learning, heuristic scoring and hidden Markov model based analysis to predict and identify RiPP BGCs and precursor peptides (Tietz et al. 2017) Using RODEO, researchers have been able to expand the RiPP family of lassopeptides by an order of magnitude, as well as characterizing six novel lasso peptides, demonstrating it as a potent tool (Tietz et al. 2017). Advances in genome sequencing and bioinformatic tools have facilitated the discovery of thousands of BGCs encoding novel natural products. The field of natural product discovery is exhibiting considerable development in the genomic era (Epstein et al. 2018; Kalkreuter et al. 2020).

1.2 Cyanobacteria and their specialized metabolites

Cyanobacteria are one of the oldest and most successful life forms on earth, and they can be found in terrestrial, aerial, glacier, brackish, marine, and fresh water environments. They represent a specialized group of microorganisms with a high degree of versatility adapted to various ecological niches (Dixit and Suseela 2013). Cyanobacteria are able to produce an array of bioactive compounds, which have potential applications in many industries, including the pharmaceutical, cosmetic, and nutritional industries (Tan 2007; Demay et al. 2019). Cyanobacteria represent a source of different natural product classes, including peptides, alkaloids, polyketides, terpenes, fatty acids and UV-absorbing compounds (Micallef et al. 2014; Demay et al. 2019). These specialized metabolites can be, among others, bioactive peptides or toxins (Burja et al. 2001; Demay et al. 2019).

Fourteen major bioactivities have been listed for cyanobacterial specialized metabolites, including neurotoxicity, lethality, dermal toxicity, cytotoxicity, hepatotoxicity, antioxidant, anti-inflammatory, antiviral, antibacterial, anti-microalgal, antifungal, and antiprotozoal activities; as well as enzyme and protease inhibition activities (Demay et al. 2019). Cyanobacterial natural products are promising for drug discovery, and some of them have entered clinical trials, but thus far only one is in use as an actual therapeutic agent. In 2011 the United States Food and Drug Administration (FDA) approved brentuximab vedotin, or SGN-35 (Adcetris; Seattle Genetics), for the treatment of Hodgkin lymphoma (Deng et al. 2013). It is an antibody-drug conjugate derived from the cyanobacterial

compound dolastatin 10 (Senter and Sievers 2012; Deng et al. 2013; Figure 2). Examples of antimicrobial peptides from cyanobacteria include antifungal hassallidins (Vestola et al. 2014), antibacterial brunsvicamides (Müller et al. 2006), and the anti-cyanobacterial compound spiroidesin (Kaya et al. 2002) (Figure 2). The BGCs of antimicrobial bacteriocins have been demonstrated to be widespread in cyanobacteria by genome mining (Wang et al. 2011). There has been a considerable expansion (71 %) of new natural products reported from cyanobacteria since 2016 (Carroll et al. 2019). The 48 novel compounds discovered from nine genera of cyanobacteria for the year 2017 represents the highest number reported for this phylum since the initiation of the annual "Marine natural products" review article (Carroll et al. 2019). The currently established broad spectrum of cyanobacterial natural products includes over 1100 specialized metabolites with divergent chemical structures described from 39 genera of cyanobacteria (Dittmann et al. 2015).

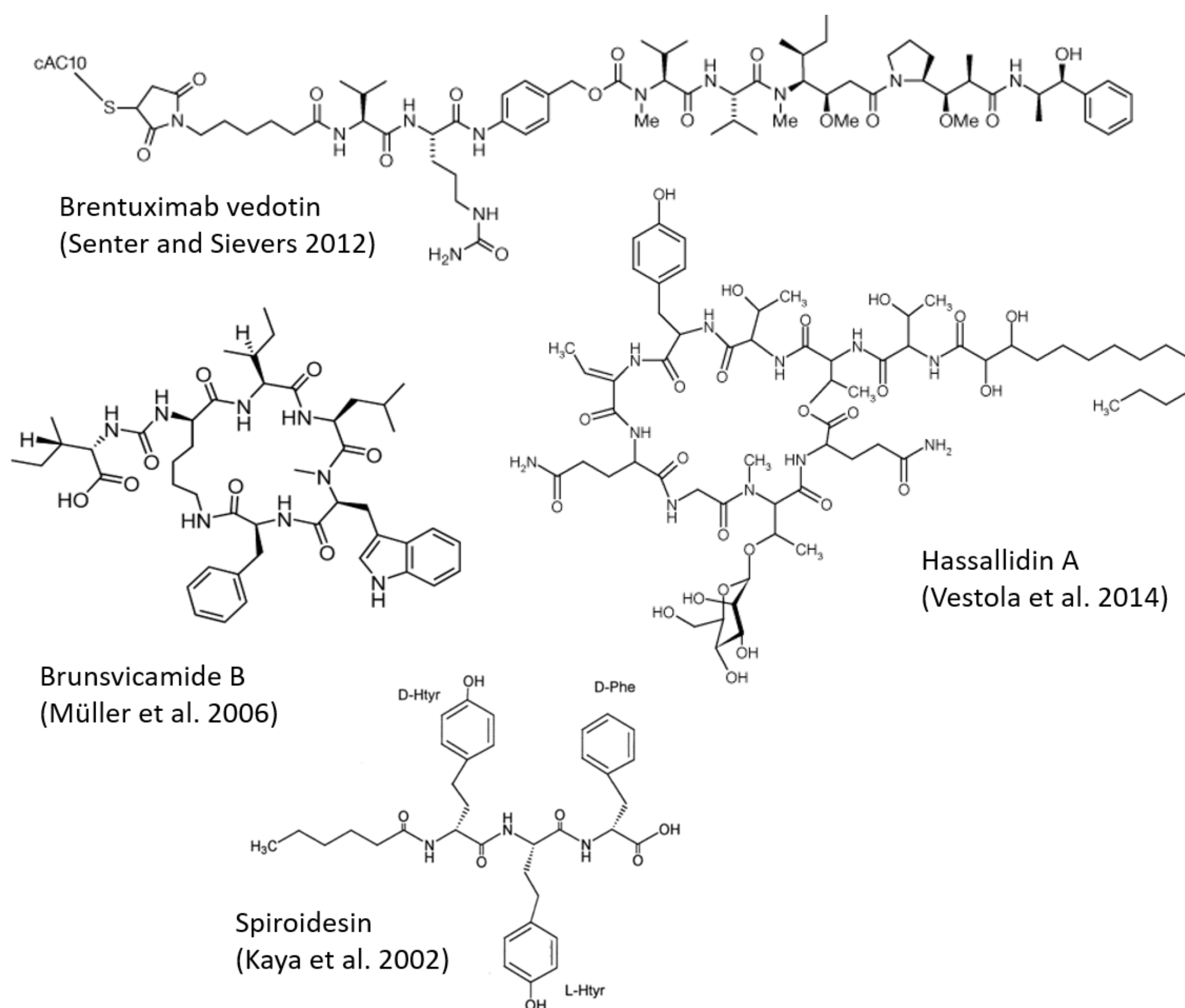


Figure 2. Selection of bioactive natural products from cyanobacteria.

1.3 Cyanobactins

Cyanobactins produced by cyanobacteria are a growing family of N-to-C macrocyclic and linear natural product peptides, which are synthesized through the post-translational modification of precursor proteins (Sivonen et al. 2010; Arnison et al. 2013; Leikoski et al. 2013). Ulicyclamide and ulithiacyclamide were the first cyanobactins reported four decades ago from the tunicate *Lissoclinum patella* (Ireland and Scheuer 1980). *L. patella* has an obligate cyanobacterial symbiont, and it was hypothesized that the symbiont, *Prochloron*, is actually responsible for the production of cyanobactins (Sings and Rinehart 1996). Two studies demonstrated that the cyanobacterium *Prochloron* is the biosynthetic source of patellamides isolated from the *L. patella* ascidian (Long et al. 2005; Schmidt et al. 2005). Long et al. (2005) used shotgun cloning of *Prochloron* genomic DNA introduced to *E. coli* and verified the heterologous expression of patellamides by LC-MS (Long et al. 2005). Schmidt et al. (2005) used sequence data to identify the patellamide BGC and demonstrated that patellamides are synthesized using ribosomal machinery (Schmidt et al. 2005). The term cyanobactin was originally introduced in 2008 by Donia et al. to describe natural products that are macrocyclic, contain azoles and are produced through the cyanobactin biosynthetic pathway. At present, a growing number of compounds are emerging from the cyanobactin family that do not strictly adhere to the original definition of cyanobactins. Cyanobactin BGCs are widespread within cyanobacterial species (10–30 % of all cyanobacteria) (Donia et al. 2008; Leikoski et al. 2009; Leikoski et al. 2013), representing one of the largest natural product family produced by cyanobacteria (Schmidt and Donia 2009; Donia and Schmidt 2010).

1.3.1 Chemical diversity of cyanobactins

Cyanobactins were originally defined as macrocyclic peptides that contain azole (heterocyclic) residues (Donia et al. 2008; Sivonen et al. 2010). Anacyclamides, kawaguchipeptins (A and B), prenylagaramides, piricyclamides, scytodecamide and trikoramide A are to date the only cyanobactins that have no azol(in)es in their structures and consequently do not encode a heterocyclase within their BGCs (Ishida et al. 1996; Ishida et al. 1997; Leikoski et al. 2010; Donia and Schmidt 2011; Leikoski et al. 2012; Martins and Vasconcelos 2015; Parajuli et al. 2016; Crnkovic et al. 2019; Phyo et al. 2019). Cyanobactins frequently contain heterocycles or prenyl group(s), and occasionally contain both (Schmidt and Donia 2009; Gu et al. 2018). Most cyanobactins consist of L-amino acids, however, D-amino acids infrequently occur adjacent to a heterocyclized amino acid (Ishida et al. 1996; Milne et al. 2006; McIntosh et al. 2009; Gu et al. 2018).

Short (3–5 amino acids) heavily modified linear cyanobactins have been reported from the cyanobacterial strains *Microcystis aeruginosa* PCC 9432 and *Oscillatoria nigroviridis* PCC 7112 (Leikoski et al. 2013). Ulithiacyclamide (Ireland and Scheuer 1980; Figure 3) and piricyclamides (Leikoski et al. 2012) are to date the only cyanobactins containing intramolecular disulfide bridges. The cyanobactin BGC does not encode an enzyme for disulfide bridge formation, and it is suggested to be a spontaneous modification based on chemical proximity (McIntosh et al. 2009: Burkhart et al. 2017) of thiol groups of cysteines in disulfide bridge containing cyanobactins. The extreme chemical diversity of cyanobactins originates from the characteristic genetic features of this class of natural products, including high variation in core peptide sequences, differing combinations of PTMs in biosynthesis, and diverging lengths of the mature cyanobactins from the pathways (Martins and Vasconcelos 2015; Burkhart et al. 2017). Examples of known cyanobactins and their chemical structures are presented in Figure 3.

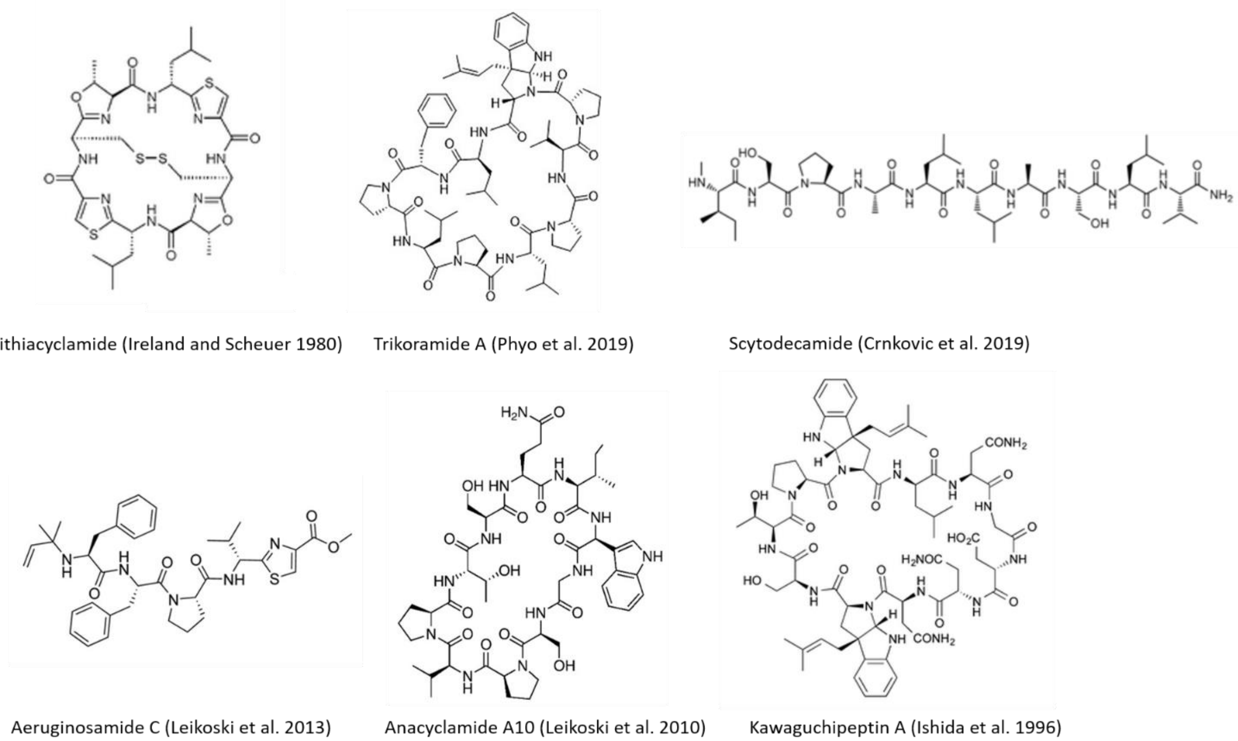


Figure 3. Chemical structures of selected cyanobactins.

Genome mining aims to identify new cyanobactins through comparison of biosynthetic gene clusters (Schmidt and Donia 2009). Once the cyanobactin BGCs have been mapped, predictions for precursor peptide protease cleavage site(s) and core peptide sequences can be performed (Velázquez and van der Donk 2011). In cyanobactin BGCs, the most highly conserved genes encode the N- and C-terminal proteases, a feature which can be utilized in the search for new cyanobactin pathways using BLAST (Donia and Schmidt 2011). However, certain cyanobactin pathways may remain unrecognized if the C-terminal protease is used for BLAST searching. The C-terminal protease might be absent from cyanobactin gene clusters, as has been shown in the case of scytodecamide (Crnkovic et al. 2019).

The presence of cyanobactins outside the cyanobacterial phyla remains elusive. However, recent genome mining studies have reported three cyanobactin precursor peptides in the genera *Streptomyces* and *Nocardiopsis* (Poorinmohammad et al. 2019), seven cyanobactin BGCs in the genus *Streptomyces* (Belknap et al. 2020), and one cyanobactin BGC in *Amycolatopsis alkalitolerans* sp. nov. (Rao et al. 2020). However, in these studies, the BGCs were not identified and the precursor peptides lacked the obvious recognition sequences characteristic of cyanobactin precursor proteins.

1.4 Cyanobactin biosynthesis

1.4.1 General biosynthetic pathway

Cyanobactins are ribosomally synthesized and post-translationally modified (RiPP) natural products. Cyanobactins are 3–20 amino acids in length, and the hypervariable core peptide sequences are completely variable, not just altered by point mutation (Leikoski et al. 2013; Martins and Vasconcelos 2015; Sardar et al. 2015). Cyanobactins are defined as a natural product class by a shared biosynthetic logic (Arnison et al. 2013). They are produced by the post-translational modification of a short genetically encoded precursor peptide (Arnison et al. 2013). The precursor peptide is divided into leader, core and recognition sequences (Sardar et al. 2015). The core sequence is modified and excised to produce the mature natural product (Czekster et al. 2016; Gu et al. 2018). In many cases, the precursor peptide encodes multiple core peptides which all are transformed to mature natural products (Ortega and van der Donk 2016). Certain cyanobacterial strains encode cyanobactin pathways with multiple precursor peptides with varying cores, such as the pathway for cyanothecamides (Donia and Schmidt 2011; Houssen et al. 2012). The existence of multiple non-identical cores in the precursor is possible because of substrate promiscuity of cyanobactin tailoring enzymes and contributes to the enormous chemical diversity of cyanobactins (Arnison et al. 2013; Donia and Schmidt 2010; Sivonen et al. 2010). The core peptide is flanked by leader (usually around 40 residues) or follower sequences that are proteolytically cleaved and thereby discarded from the final product (Arnison et al. 2013).

The canonical cyanobactin biosynthetic pathway includes the ribosomal synthesis of the precursor peptide (E-gene), followed by heterocyclization (D-gene) if the mature product contains heterocycles. This is followed by proteolytic cleavage of the leader sequence (A-gene), and cleavage/macrocyclization (G-gene) (Figure 4; Czekster et al. 2016; Gu et al. 2018). The size of a cyanobactin BGC is approximately 8–19 kb in length (Martins and Vasconcelos 2015), and the gene organization is not strictly conserved (Sivonen et al. 2010). Additional tailoring enzymes are often present in cyanobactin biosynthetic pathways, including prenyltransferases, oxidases, and methyltransferases (Hao et al. 2016; Sardar et al. 2017; Gao et al. 2018; Crnkovic et al. 2019). Recently, a novel linear cyanobactin, scytodecamide, was demonstrated to be amidated in the C-terminus by the amidotransferase ScdTA (Crnkovic et al. 2019; Figure 3). These additional enzymes usually, but not always, catalyze late-stage reactions in cyanobactin biosynthesis (Martins and Vasconcelos 2015; Czekster et al. 2016; Sardar et al. 2017; Gu et al. 2018).

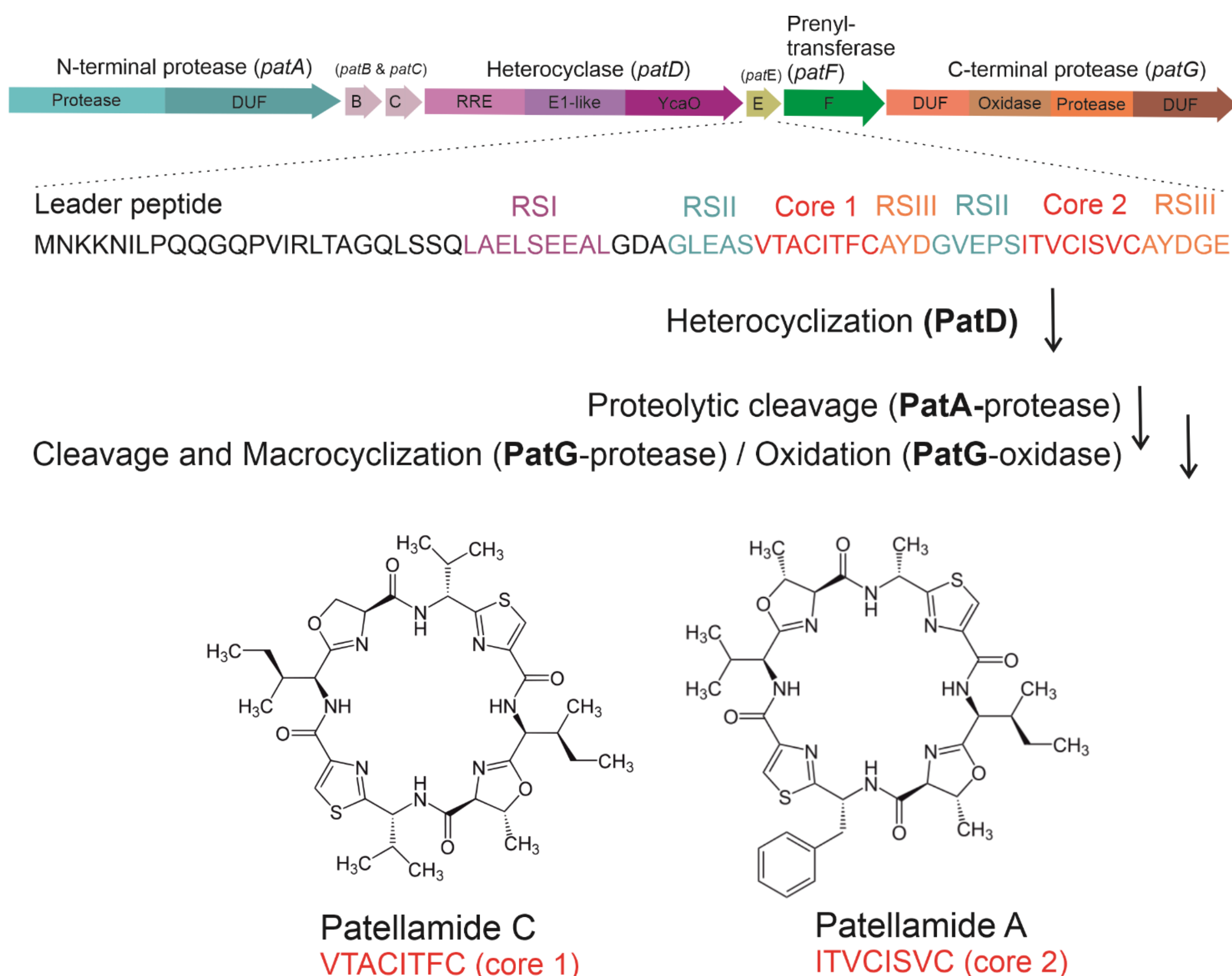


Figure 4. Schematic depiction of canonical cyanobactin pathway represented by the patellamide (*pat*) pathway. The figure has been adapted from Schmidt et al. (2005) and Gu et al. (2018).

Cyanobactin enzymes have a natural ability for plasticity enabling them to process several substrates, but the end-product(s) of the pathways is usually specific (Martins and Vasconcelos 2015; Czekster et al. 2016; Gu et al. 2018). Identical enzymes thus modify hypervariable core peptides within cyanobactin pathways (Sardar et al. 2015). The relaxed substrate selectivity of cyanobactin pathways was shown by Tianero et al. (2012) who used the *tru* pathway to generate cyanobactin mutants. In addition to mutations to the core peptide, non-proteogenic amino acids were introduced to the core and the products produced (Tianero et al. 2012). The PTM enzymes were tolerant to changes made to the core (Donia et al. 2006; Tianero et al. 2012), suggesting that cyanobactin pathways are ideal for bioengineering using synthetic biology (Martins and Vasconcelos 2015; Sardar et al. 2015).

Recognition sequences (RSs) are short sequences encoded into the precursor protein that guide enzymes that modify the core (Oman and van der Donk 2010) but are not part of the final natural product (Gu et al. 2018). Enzymes bind to RSs and therefore focus their biosynthetic efforts on a single gene product (Gu et al. 2018). Cyanobactin core peptides are hypervariable, but the leader sequences including recognition sequences are to some extent conserved. The heterocyclase (D-protein) recognizes RSI, while RSII and RSIII guide the activities of the protease (A-protein) and macrocyclase (G-protein), respectively (Lee et al. 2009; Figure 4). If the precursor peptide encodes more than one core, each one is sandwiched by N-terminal (RSII) and C-terminal (RSIII) RSs which

can comprise four or five amino acids (Martins and Vasconcelos 2015; Figure 4). It is unknown which recognition elements are required for the activity of the oxidase and prenyltransferase (Koehnke et al. 2015). The prenyltransferase is thought to act on the released product, and therefore no RSs would be needed for the catalytic activity of a cyanobactin prenyltransferase.

The canonical cyanobactin pathway encompasses two proteases: PatA and PatG. Leikoski et al. (2013) reported linear cyanobactins, aeruginosamide A & B and viridisamide A, possessing both proteases (A and G) within their BGCs. The C-terminal protease (G-protein) in the linear cyanobactin pathway (*age* and *vir*) linearizes and hydrolyzes instead of macrocyclization (Leikoski et al. 2013; Sardar et al. 2017). A better understanding of the enzyme mechanisms involved in the biosynthesis of the evermore expanding cyanobactin class of natural products would provide new interesting (bio)chemistry and insights to the concept of cyanobactin BGC. With the progress of genome mining studies, unusual cyanobactins are highly likely to be discovered in the near future expanding the scope of the biosynthetic potential of cyanobactin BGCs.

The minimal set of genes needed for cyanobactin biosynthesis and present in cyanobactin gene clusters includes genes encoding the precursor peptide, two short conserved hypothetical proteins, and two proteases (Donia and Schmidt 2010). The prenyltransferase is present in most cyanobactin BGCs, including those whose end-product is not prenylated (Schmidt and Donia 2009; Sivonen et al. 2010). Scytodecamide from *Scytonema* sp. UIC 10036 is a novel linear cyanobactin featuring N-terminal N-methylation and C-terminal amidation (Crnkovic et al. 2019; Figure 3), which are both unprecedented PTMs in cyanobactin biosynthesis. Scytodecamide is a unique linear cyanobactin decorated only with the N-terminal protease (A-protein) and missing the G-protein completely from its BGC, suggesting that the concept of minimal cyanobactin BGC needs to be redefined. What is minimal to one cyanobactin pathway might not be minimal for another. Genome mining leading to the discovery of structurally novel cyanobactins expands the enzymatic machinery of cyanobactin biosynthesis.

1.5 Enzymes involved in cyanobactin biosynthesis

1.5.1 Proteases

The canonical cyanobactin gene cluster encompasses two proteases which are responsible for cleavage of the precursor peptide and cyclization of the mature product, PatA and G -homologs (Lee et al. 2009; Donia and Schmidt 2010; Sivonen et al. 2010). They are paralogous enzymes and have relatively high sequence similarity (~ 40 %) (Lee et al. 2009; Agarwal et al. 2012; Czekster et al. 2016). These proteases were first characterized when the *pat* gene cluster was discovered (Schmidt et al. 2005). The PatA homolog is responsible for the proteolytic cleavage of the N-terminal, while the PatG homolog catalyzes cleavage of the C-terminal recognition sequence together with simultaneous cyclization (Lee et al. 2009). The PatA homolog recognizes the G(L/V)E(A/P)S recognition sequence at the N-terminus, and the PatG homolog recognizes the AYDG(E) recognition sequence at the C-terminus (Lee et al. 2009; Figure 4). The N-terminal G(L/V)E(A/P)S recognition sequence is not strictly conserved and varies between cyanobactin pathways, for instance, in the case of cyanothecamides with the non-canonical flanking sequences GFEAP and SNCIG (Houssen et al. 2012). The PatA and PatG -homologs are both subtilisin-like serine proteases (Oman and van der Donk 2010), and they both have a domain of unknown function (DUF) (Lee et al. 2009; Agarwal et al. 2012). Both proteases recognize the RSII and RSIII (Ortega and van der Donk 2016; Figure 4), and eventually the 50–150 amino acid long precursor protein is excised and modified to form the

mature product (Donia and Schmidt 2010; Sivonen et al. 2010). The catalytic triad of the PatA protease domain is composed of Asp23-His58-Ser218 (Agarwal et al. 2012). Two disulfide bridges are present in the PatA protease domain, between Cys156 and Cys158, and between Cys258 and Cys269, respectively (Agarwal et al. 2012; Houssen et al. 2012).

Most cyanobactins are macrocyclic via ligation by peptide bond of their C- and N-termini, and not via amino acid side chains like many other macrocyclic secondary metabolites (Agarwal et al. 2012; Czekster et al. 2016; Gu et al. 2018). PatG includes a protease domain, an oxidase domain, and N- and C-terminal DUFs (Gu et al. 2018; Figure 4). The catalytic triad of the PatG protease domain is composed of Asp548-His618-Ser783 (Agarwal et al. 2012). PatG functions via a transamidation mechanism where the reactive-site Ser of PatG forms an enzyme-substrate tetrahedral intermediate with the evanescence of the C-terminal recognition sequence (Lee et al. 2009). An exception to this rule lies in the *age* pathway where the PatG homolog (AgeG) linearizes and hydrolyzes instead of macrocyclization resulting in a linear product (Leikoski et al. 2013; Sardar et al. 2017). PatG has value as a biotechnological tool as a macrocyclase because it has high substrate tolerance, but the C-terminal recognition sequence must be present (Lee et al. 2009). Additionally, the substrate is required to have at least one proline or thiazoline and is between 6–11 amino acid residues in size in order for PatG to catalyze macrocyclization efficiently (Koehnke et al. 2012). However, the proline/thiazoline requirement has been bypassed using double cysteine and triazole, creating two synthetic routes for the PatG macrocyclase (Oueis et al. 2017). The promiscuity of the PatG macrocyclase has been demonstrated with unnatural substrates producing patellamide-like macrocyclic compounds (Oueis et al. 2015). The enzyme substrate specificity of the OscG cyanobactin macrocyclase from the oscillacyclamide pathway from *Oscillatoria* sp. PCC 6506 was tested for 19 linear peptide substrates (Alexandru-Crivac et al. 2017). Compared to PatG, OscG was shown to possess less rigorous recognition determinants, wider substrate tolerance, and the ability to process longer substrates including D-amino acids (Alexandru-Crivac et al. 2017).

1.5.2 Heterocyclase

Many natural product RiPPs contain Thr-, Ser- and Cys-derived heterocycles (McIntosh et al. 2009; Arnison et al. 2013), and the heterocyclization (also known as cyclodehydration) of these amino acid residues is a common modification in cyanobactin pathways (Koehnke et al. 2015; Martins and Vasconcelos 2015). Peptide heterocycles add chemical diversity, reduce polarity and increase conformational stability in peptides (Czekster et al. 2016). Heterocyclic structures are formed when the side-chain hydroxyl of Thr, Ser or Cys or thiol attacks the carbonyl carbon of the adjoining amino acid yielding methyloxazoline, oxazoline, or thiazoline, respectively (McIntosh et al. 2009). The oxidized derivatives of these structures are named methyloxazole, oxazole and thiazole, respectively (McIntosh et al. 2009). Cyanobactins usually contain oxazoline and thiazoline structures, or their oxidized derivatives oxazoles and thiazoles (Sivonen et al. 2010; McIntosh and Schmidt 2010; Velázquez and van der Donk 2011; Gu et al. 2018). Heterocyclases recognize the RSI binding site and catalyze the alterations of one or more serine, threonine or cysteine, into the equivalent azoline (McIntosh and Schmidt 2010). When the RSI is present, substrate peptides are effectively heterocyclized, while in the absence of RSI, heterocyclization is very inefficient or totally abolished (Sardar et al. 2015). Cyanobactin pathways which lack the heterocyclase, RSI is absent from the precursor peptide (Gu et al. 2018). Adding RSI to cyanobactin precursor peptides, which originally lack RSI, allows access to the element and enables heterocyclization (Sardar et al. 2015). LynD anchors the substrate peptide by binding a conserved region within the leader sequence of PatE (Koehnke et al. 2015). The leader peptide of the substrate also forms contacts with the LynD catalytic domain stabilizing a loop within the active site (Koehnke et al. 2015). Heterocycles are not contiguous

in the cyanobactins reported to date but are instead interspersed in the macrocycle occurring singly (McIntosh et al. 2009; Sivonen et al. 2010; Martins and Vasconcelos 2015).

Many cyanobactins include a heterocycle at the C-terminus of the core peptide (either proline or azol(in)e derived from cysteine, serine, or threonine) (Donia and Schmidt 2010; Donia and Schmidt 2011; Leikoski et al. 2013). Heterocyclases require ATP and Mg^{2+} for their activity, and they share the YcaO domain as a catalytic unit (Dunbar et al. 2012). The cyanobactin heterocyclase has two domains in addition to YcaO: an E1-like domain and a substrate recognition domain (RiPP recognition element, RRE) (Burkhart et al. 2015; Gu et al. 2018; Figure 4). YcaO family heterocyclases are relatively widespread in cyanobactin gene clusters, suggesting that heterocyclization offers a favorable change in the chemical construction of cyanobactins (Burkhart et al. 2017). The TruD cyanobactin heterocyclase was shown to be an adenylase and catalyze cysteine heterocyclization in strict order from the C- to N-terminus (Koehnke et al. 2013). The leader peptide can be discarded and TruD will act on a single specific cysteine (Koehnke et al. 2013). Heterocyclization most likely occurs as the first step of cyanobactin biosynthesis for the unmodified precursor. MicD, a cyanobactin heterocyclase from *Microcystis aeruginosa*, can heterocyclize both cysteines and threonines (Ge et al. 2019). Cysteines are heterocyclized before threonines from the C- to N-terminus by MicD, if they are present in the same peptide substrate (Ge et al. 2019).

1.5.3 Oxidase

The oxidase enzyme catalyzes the formation of (methyl)oxazoles and thiazoles from methyloxazolines, thiazolines, and oxazolines which are made by the heterocyclase (Czekster et al. 2016). The cyanobactin oxidase does not seem to compose a complex with the heterocyclase, but is either found as a freestanding protein or fused together as an oxidase domain with the macrocyclase enzyme and a DUF (G-protein) (Donia and Schmidt 2011; Gao et al. 2018). The ThcOx oxidase from *Cyanothece* sp. PCC 7425 catalyzes thiazoline rings oxidation to thiazoles with FMN as a cofactor, but this enzyme is unable to use linear peptides as a substrate (Houssen et al. 2014). An enzyme from *A. platensis* homologous enzyme to ThcOx, ApOx, was able to oxidize thiazoline into thiazole on linear and macrocyclic cyanobactin precursors (Houssen et al. 2014). ThcOx consists of a C-terminal FMN binding the nitroreductase catalytic domain and two RRE domains at the N-terminus (Bent et al. 2016). Recently, ThcOx was shown to convert thiazolines to thiazoles regardless of the presence of the leader peptide, and the reaction was shown to be faster and more efficient on the leaderless substrate (Gao et al. 2018). Oxidation by ThcOx proceeds in a defined order from the N- to C-terminus (Gao et al. 2018). Since the order does not depend on the leader, it is a feature of the enzyme and core sequence (Gao et al. 2018). Cyanobactin oxidases can catalyze oxidation on linear and macrocyclic substrates, and they could therefore be used as flexible enzymes in biotechnological applications.

1.5.4 Prenyltransferase

Prenylation is a widespread modification in nature. Prenylation enhances the lipophilicity of natural products and increases their affinity for biological membranes (Botta et al. 2005; Wong et al. 2018). According to current understanding, F-type cyanobactin prenyltransferases act at the final step of biosynthesis (McIntosh et al. 2011). Cyanobactin prenylation is catalyzed by ABBA prenyltransferases (Hao et al. 2016; Gu et al. 2018). Two cyanobactin prenyltransferases have been crystallized, the inactive PatF (Bent et al. 2013) and active PagF (Hao et al. 2016; Figure 5 A),

demonstrating the ABBA fold. The structure of the ABBA prenyltransferase is composed of a central barrel of 10 antiparallel β-strands surrounded by α-helices (Kuzuyama et al. 2005; Metzger et al. 2009; Bonitz et al. 2011). The central barrel forms a catalytic cavity where substrate binding and catalysis occurs (Kuzuyama et al. 2005; Metzger et al. 2009; Hao et al. 2016). The sidechains of catalytic amino acid residues stick into the barrel which control the donor substrate DMSPP (or DMAPP) (Figure 5 A; Hao et al. 2016), which enables substrate prenylation (Figure 5 B). The cyanobactin prenyltransferases have a broad substrate selectivity and they accept a wide range of macrocyclic and linear peptide substrates (Hao et al. 2016; Gu et al. 2018). The plasticity of cyanobactin prenyltransferases makes them valuable catalysts in synthetic biology creating structural diversity to a wide array of compounds.

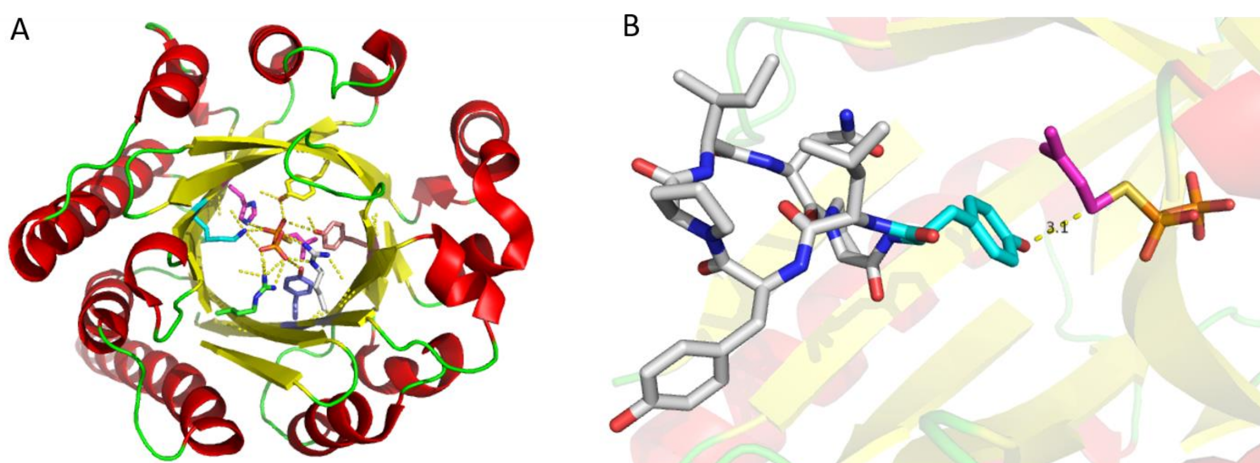


Figure 5. A) PagF cyanobactin prenyltransferase (PDB: 5tu6) in complex with DMAPP demonstrating the ABBA-fold (Hao et al. 2016). B) Cyclic peptide substrate (INPYLYP) in complex with DMSPP (Hao et al. 2016). The Tyr-residue (colored in cyan) is the prenyl (colored in magenta) acceptor. Figures were created with Pymol based on PDB structures.

The catalytic activity of a cyanobactin prenyltransferase does not depend on the recognition sequence in the precursor protein, and therefore it has relatively relaxed substrate selectivity (Hao et al. 2016). Cyanobactin prenyltransferases are known for their promiscuous substrate selectivity. Nevertheless, they are highly specific regarding the donor substrate and the regio/stereo positioning of the prenyl relative to the peptide substrate (Gu et al. 2018). Owing to these characteristics, they could be useful for the prenylation of particular residues of divergent peptide substrates (Gu et al. 2018). Prenyltransferases from cyanobactin pathways can catalyze forward and reverse prenylation on cyclic and linear substrates.

The trunkamide cyanobactin biosynthetic gene cluster encodes the TruF1 and TruF2 enzymes, and trunkamide is bisprenylated on threonine and serine (Donia et al. 2008; McIntosh et al. 2011). Heterologous expression studies have been carried out by Tianero et al. (2016) to demonstrate the activities of the potential prenyltransferases TruF1 and TruF2. The genes encoding TruF1 and TruF2 were knocked out individually or together, and it was shown that TruF1 is individually responsible for the prenylation in the *tru* pathway while TruF2 is a nonprenylating homolog. The deletion of TruF1 led to the production of macrocyclic peptides without prenyl moieties (Tianero et al. 2016). The TruF1 prenyltransferase *O*-prenylates Thr, Tyr and Ser on cyclic or linear peptide substrates (Hao et al. 2016).

Prenylagaramides are forward *O*-prenylated on Tyr by PagF (Donia and Schmidt 2011). Aestuaramides are reverse *O*-prenylated on Tyr by LynF (McIntosh et al. 2011; McIntosh et al. 2013). Reverse prenylated Tyr intermediates can undergo a spontaneous Claisen rearrangement, yielding a forward C-prenylated moiety, even without an enzyme (McIntosh et al. 2013). PagF *O*-prenylates Tyr in forward orientation (Hao et al. 2016). The recently discovered sphaerocyclamide is forward *O*-prenylated on Tyr by the SphF prenyltransferase (Martins et al. 2018; Figure 7).

Croissamide is reverse *N*-prenylated on Trp (Iwasaki et al. 2018). The biosynthetic gene cluster has not been described, but the structure suggests that croissamide is a cyanobactin. Recently, trikoramide from *Symploca hydnoides* was reported with a C-prenylated cyclotryptophan in forward orientation (Phyo et al. 2019; Figure 3). Kawaguchipeptin A produced by *Microcystis aeruginosa* NIES-88 is a cyclic peptide and C-bisprenylated on two Trp residues in forward orientation (Ishida et al. 1996; Figure 3). The kawaguchipeptin gene cluster has been identified and, heterologously expressed, and it has been shown that the bisprenylation of Trp is catalyzed by the prenyltransferase KgpF encoded in the gene cluster (Parajuli et al. 2016). The prenylation by KgpF was shown *in vitro* on several linear and cyclic substrates using DMAPP as a donor substrate (Parajuli et al. 2016). The stereospecificity of forward C-prenylation by KgpF has been confirmed using Fmoc-protected Trp (Okada et al. 2016).

A smaller proportion of cyanobactins are linear by structure, and are reverse prenylated on the N-terminus and methylated on the C-terminus (Leikoski et al. 2013; Figure 3). This linear peptide termini protection is carried out by the bifunctional enzyme AgeMTPT (Sardar et al. 2017). Leikoski et al. (2013) have hypothesized that the N- and C-termini prenylated natural product peptide muscoride (Nagatsu et al. 1995) might be the product of a cyanobactin BGC (Leikoski et al. 2013).

Piricyclamide cyanobactins from *Microcystis aeruginosa* PCC 7005 have been suggested to be geranylated instead of prenylation by prenyltransferase PirF (Leikoski et al. 2012). Later this enzyme PirF has been shown to be a geranyltransferase catalyzing forward *O*-geranylation of Tyr, an unprecedented post-translational modification in cyanobactin biosynthesis (Morita et al. 2018). Mutagenesis of only one amino acid (Phe222 to Ala222) residue in the binding pocket of the PirF ABBA prenyltransferase alters the donor specificity from prenyl to geranyl (Estrada et al. 2018). This might be due to amino acid sidechains sticking into the catalytic chamber of the enzyme. When there is Phe at the catalytic site, it possibly prevents geranyl substrate binding because of space restriction (Figure 6). These high plasticity characteristics promote cyanobactin prenyltransferases as valuable tools in biotechnology.

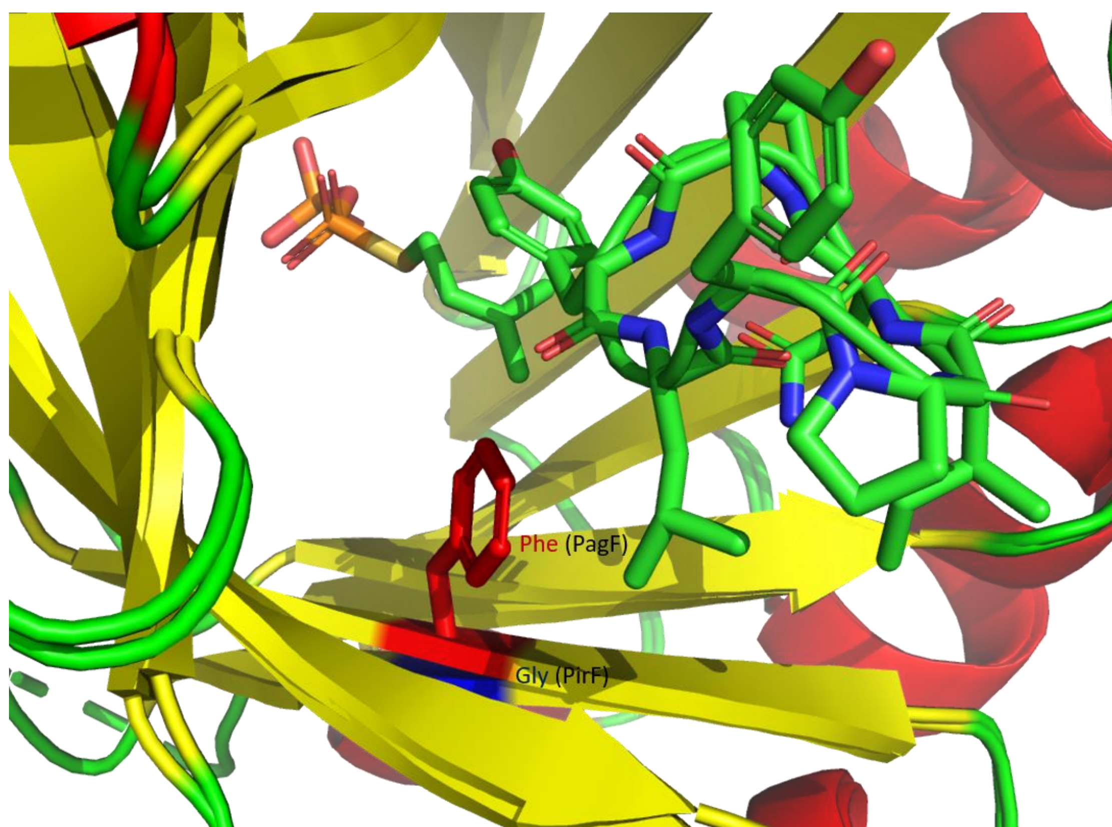


Figure 6. Structural alignment of PirF geranyltransferase (PDB: 6pgm) and PagF prenyltransferase (PDB: 5tu6). The peptide substrate (INPYLYP) is coloured in green. The diphosphate group of DMAPP is shown in brown and the prenyl group of DMAPP in green.

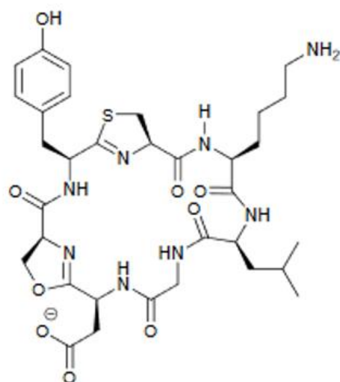
1.5.5 Proteins and domains of unknown functions in cyanobactin BGCs

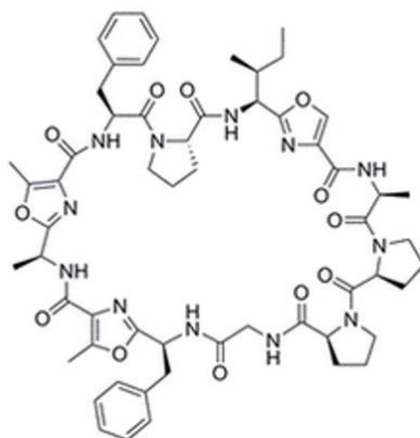
Cyanobactin pathways encode proteins with unknown functions as well as domains of unknown functions (DUFs). The genes encoding for B- and C-proteins occur in cyanobactin gene clusters, as do homologous C-terminal DUFs of A- and G-proteins (Lee et al. 2009; Agarwal et al. 2012; Figure 4), but their functions are unknown. The B- and C-proteins are short proteins without recognized functions and no resemblance to proteins associated with known functions (Gu et al. 2018). Since DUFs are conserved throughout cyanobactin pathways, they are likely to have an operational role in biosynthesis (Sikandar and Koehnke 2019). Because DUFs from PatA and PatG are homologous, it has been suggested that these domains could form heterodimers of PatA and PatG through Zn^{2+} -mediated dimerization, which might be essential for the *in vivo* -activity of these proteins (Sikandar and Koehnke 2019). Such a complex could possibly reduce precursor peptide degradation by substrate channeling (Sikandar and Koehnke 2019). Epimerization has been suggested as a possible role for DUF, but the epimerization to the D-form has been noted to be most likely spontaneous and non-enzymatic (Wipf et al. 1998; Milne et al. 2006: McIntosh et al. 2009; Koehnke et al. 2014). It is possible that epimerization enables the macrocycles to accept the thermodynamically preferred form (Gu et al. 2018). The structure of the DUF of PatG was determined, but no binding between precursor peptide and this domain was shown (Mann et al. 2014). The fold of the DUF was new but yielded no insights regarding its potential function (Mann et al. 2014). The PatG -DUF has one disulfide bridge between Cys1136 and Cys1142 and a single Zn^{2+} ion in the part of the domain that was characterized (Mann et al. 2014). The heterologous expression of patellamides demonstrated that PatB and PatC

are nonessential in patellamide biosynthesis (Donia et al. 2006). Nevertheless, these proteins and domains are universally found in cyanobactin BGCs, and they therefore presumably have important roles in cyanobactin biosynthesis.

1.6 Ecological role of cyanobactins and bioactivity

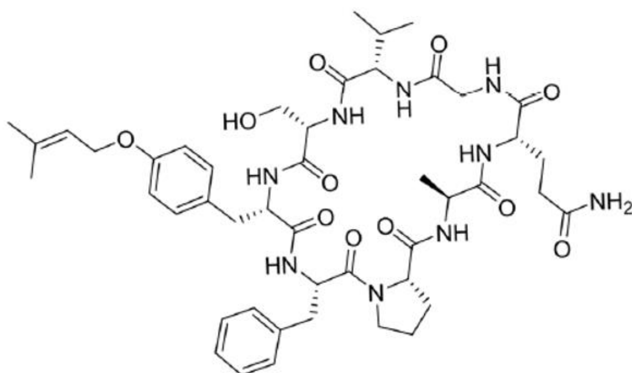
The ecological role of cyanobactins is not fully understood. Macrocyclization is a common modification in cyanobactins, supposedly because of the advantages that macrocycles provide in activity, membrane permeability and stability properties compared to their linear counterparts (Agarwal et al. 2012; Koehnke et al. 2012). Cyanobactins have selective metal chelating properties, and they might therefore have a role in metal transport, storage or detoxification (Bertram and Pattenden 2007; Houssen and Jaspars 2010; Bauermeister et al. 2019). Cyanothecamides were detectable after *Cyanothece* PCC 7425 was subjected to heat shock, suggesting that the expression of these cyanobactins might be due to stress responses (Houssen et al. 2012; Figure 7). Trikoramide was shown to be cytotoxic against AML2 and MOLT-4 cancer cell lines (Phyo et al. 2019; Figure 3). Wewakazole B was shown to be cytotoxic against human H460 lung cancer cells and human MCF7 breast cancer cells (Lopez et al. 2016; Figure 7). Ulithiacyclamide showed potent cytotoxic activity against L1210 murine leukemia cells and it was suggested that the efficacy of this cyanobactin was associated with the intramolecular disulfide bridge (Shioiri et al. 1987; Figure 3). The cytotoxic effect of ulithiacyclamide was decreased when the compound was treated with dithiothreitol (DTT) to reduce the disulfide bridge to prevent the correct conformation of the molecule (Shioiri et al. 1987). Further development of ulithiacyclamide was, however, stopped because of its high toxicity (Houssen and Jaspars 2010). Bisprenylated patellin 2 might inhibit potassium channels by blocking (Tianero et al. 2016). Antibacterial cyanobactins include the prenylated sphaerocyclamide against *Halomonas aquamarina* (CECT 5000) (Martins et al. 2018; Figure 7), muscoride A against *Bacillus subtilis* (Nagatsu et al. 1995; Figure 13), and kawaguchipeptins A (Figure 3) & B against *Staphylococcus aureus* at minimal inhibition concentrations (MIC) of 1 µg/mL (Ishida et al. 1997). Whether these bioactivities are associated with the ecological roles of cyanobactins remains an enigma (Gu et al. 2018). Despite limited understanding regarding the bioactivity and ecological role of cyanobactins, they still show promise in synthetic biology applications. The PTM enzymes from cyanobactin pathways could have applications in the production of modified peptides.





Cyanothecamide C
(Houssen et al. 2012)

Wewakazole B
(Lopez et al. 2016)



Sphaerocyclamide (Martins et al. 2018)

Figure 7. Selection of bioactive cyanobactins.

1.7 Antimicrobial peptides (AMPs)

Antibiotic resistance is a global health threat, especially owing to the lack of new antimicrobial drugs (Tossi et al. 2000; Hassan et al. 2012: Mookherjee et al. 2020). As a consequence, new therapeutic agents are urgently needed which would be less likely to cause antibiotic resistance. Antimicrobial peptides (AMPs) represent attractive candidates for the development of new antimicrobial agents for multidrug resistant bacteria (Ageitos et al. 2017). An antimicrobial peptide (AMP) is an agent that kills microorganisms or inhibits their growth. One indisputable function of RiPPs is related to AMPs (Hudson and Mitchell 2018), which operate as chemical weapons in competition and defense (Li and Rebuffat 2019). AMPs are important factors in the innate immunity in multicellular eukaryotes (Zasloff 2002), and in microorganisms, AMPs are generally considered to be chemical weapons allowing microbes to eradicate rivals from other species fighting for the same ecological niche (Hassan et al. 2012; Li and Rebuffat 2019). Regardless of their biological origin, all AMPs include a linear or cyclic structure and are small in size (7–100 amino acids) (Ageitos et al. 2017). Most naturally occurring AMPs have an amphipathic structure and are cationic (Lai and Gallo 2009). Usually AMPs are co-expressed as groups that act in synergy, although this depends on the cell (Lai and Gallo 2009).

AMPs can be broadly divided into four groups: (1) linear α-helix structure peptides with amphipathic properties; (2) β-sheet peptides often stabilized by disulfide bridges; (3) extended peptides with a relatively high amount of specific amino acids (most often Arg, Gly, Pro, Trp, and/or His); and (4) peptides possessing one or more D-amino acids or modified amino acid(s) in their structure (Lai and Gallo 2009; Ageitos et al. 2017; Mookherjee et al. 2020; Figure 8).

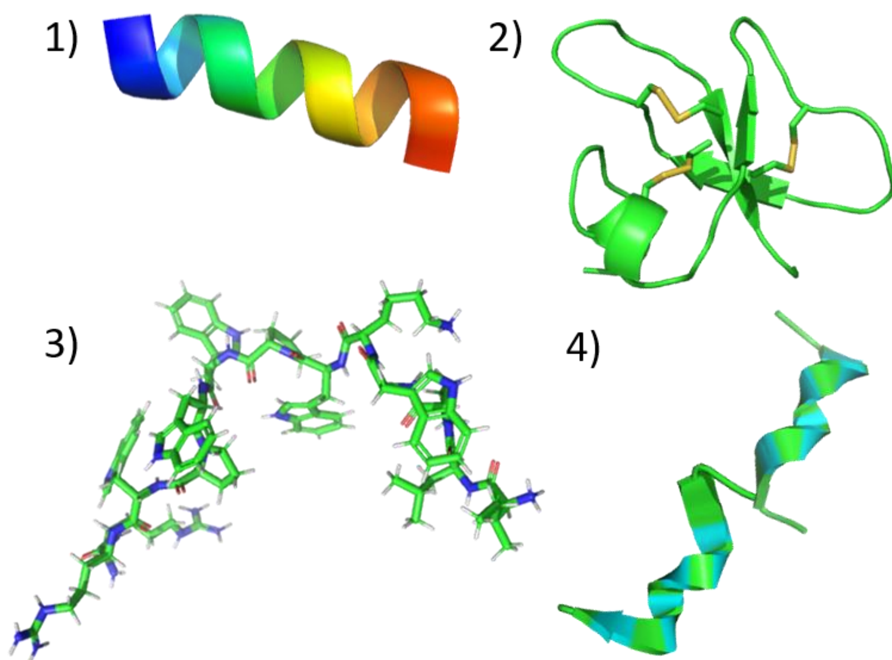


Figure 8. Structures of selected AMP´s. 1) α-helical LL-37 (PDB: 2FBS); 2) Human beta-defensin-2, where disulfide bridges are highlighted as light brown (PDB: 1FQQ); 3) Indolicidin, which is unusually rich in tryptophan and proline (PDB: 1G89); 4) Gramicidin B, where D-amino acids are highlighted in cyan (PDB: 1JO3). All structures were drawn using PyMol based on PDB structures.

1.7.1 Disulfide bridge formation

The amino acid cysteine is able to form covalent intra -and inter molecular disulfide bridges through its reactive side chain thiol group (SH) (Fass and Thorpe 2018). Disulfide bond formation is a widespread post-translational covalent modification among proteins and peptides, and one of the most common modification in the process of oxidative folding (Walsh et al. 2005). The formation of disulfide bridges can be distinguished into two steps; the oxidation of cysteines to form disulfide bridges, and the isomerization of erroneous disulfide bridges to the *bona fide* pattern (Gleiter and Bardwell 2008). For proteins and peptides with more than two cysteine residues, the possibility for erroneous disulfide bonding increases rapidly with cysteine count (Gleiter and Bardwell 2008). The larger the number of disulfide bridges, the greater the stability of the mature product (in theory) (Góngora-Benítez et al. 2013). In bacteria, proteins with disulfide bridges play vital roles in important cellular processes, including motility, cell division, virulence factors of pathogenic bacteria, reaction to environmental threats, transport of virulence proteins, and outer membrane assembly of Gram-negative bacteria (Landeta et al. 2018; Figure 9). Disulfide bridges can be oxidized (formed), reduced

(broken) or isomerized (shuffled) in the context of protein and peptide folding (Berkmen 2012). In proteins and peptides, disulfide bridges can be of catalytic, signaling or structural type (Berkmen 2012).

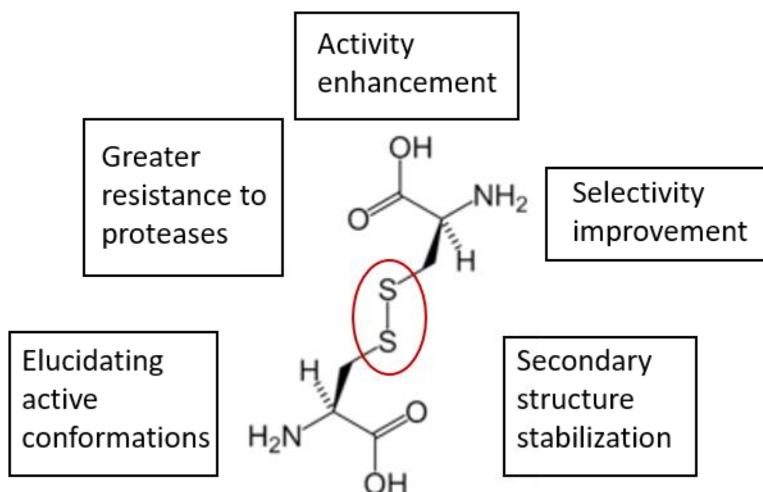


Figure 9. Advantages of disulfide bridge. Adapted from Góngora-Benítez et al. (2013).

Disulfide bridges in proteins and peptides are generated through an oxidative process where a covalent bond is formed between the sulfur atoms of two cysteines (Heras et al. 2009; Landeta et al. 2018; Fass and Thorpe 2018; Figure 10). Disulfide bridges can form spontaneously in atmospheric conditions but the reaction is slow (Bocian-Ostrzycka et al. 2017). Disulfide bridge formation is an enzymatic process and an electron transfer event where a temporary disulfide bridged complex is formed between the enzyme and the substrate (Berkmen 2012). Proteins that catalyze disulfide bridge formation belong to the large group of thiol-disulfide oxidoreductases (TDORs) found in all living cells (Kadokura et al. 2003; Bocian-Ostrzycka et al. 2017). All such proteins have a thioredoxin fold and the catalytic site motif CxxC, which is involved in redox reactions (Bardwell et al. 1991; Berkmen 2012; Landeta et al. 2018). Disulfide bridge formation through oxidative folding by folding catalysts is fundamental for the activities of many proteins and peptides, bacterial virulence, proper growth and division (Heras et al. 2009; Reardon-Robinson and Ton-That 2016).

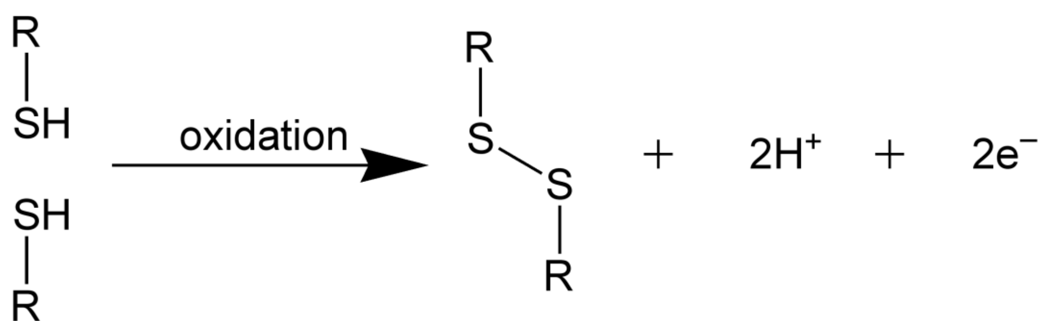


Figure 10. Oxidation of thiol groups of two cysteines resulting in disulfide bridge.

The Dsb family of enzymatic proteins represents a prime example of bacterial oxidative protein folding. The first Dsb protein discovered was DsbA from *E. coli* (Bardwell et al. 1991). DsbA forms disulfide bridges to its substrate by a thiol-disulfide exchange reaction via bestowing its disulfide bridges to the cysteines of the substrate. DsbB interacts with DsbA, keeping it oxidized and thereby active. DsbC and DsbD work as isomerases by repairing incorrect disulfide bridges (Kadokura et al. 2003; Kadokura and Beckwith 2010; Davey et al. 2016; Landeta et al. 2018; Fass and Thorpe 2018). Certain bacterial taxa possess a DsbA homolog while lacking a DsbB homolog. These species, including cyanobacteria and actinobacteria, employ a protein belonging to the VKOR family to reoxidize the DsbA homolog (Dutton et al. 2008). The active site (CXXC) of the *Staphylococcus aureus* DsbA (SaDsbA) has been shown to be evenly stable in its oxidized and reduced shapes, indicating that it does not depend on the DsbB-like protein (Heras et al. 2008). SaDsbA requires solely the presence of oxidizing factors to reach the catalytically active form (Heras et al. 2008). Another individually operating DsbA-like protein called SdbA has been identified in *Streptococcus gordonii* and shown to be crucial in substrate disulfide bridge formation and a novel type of TDOR (Davey et al. 2013). Disulfide bridge formation is a fundamental process in microbial life (Landeta et al. 2018) and the burgeoning count of sequenced bacterial genomes has shown that oxidative folding mechanisms differ extensively (Heras et al. 2009).

1.7.1.1 Bioactive natural products with disulfide bridge(s)

The formation of disulfide bridges is an important post-translational modification increasing the conformational stability of many natural products. Disulfide bridges conserve the macromolecular conformation and are vital for the biological activity and stability of peptides, and are therefore frequently found in peptide therapeutics (Góngora-Benítez et al. 2013). Examples of disulfide-containing therapeutic peptides include the anticancer agent romidepsin from *Chromobacterium violaceum*, antibiotic holomycin from *Streptomyces clavuligerus*, virulence factor gliotoxin from *Aspergillus fumigatus*, peginesatide for the treatment of anemia and linaclotide for the treatment of chronic constipation (Góngora-Benítez et al. 2013; Scharf et al. 2014; Fass and Thorpe 2018). Disulfide bridges can lock the structure of the peptide into a bioactive conformation. These characteristics make cysteine residues and disulfide bridges good tools for improving the pharmacological properties of peptides (Góngora-Benítez et al. 2013; Figure 9).

Defensins are antimicrobial peptides, expressed by animals, fungi and plants for host defence (Ganz 2003; Shafee et al. 2017). They are small (<10 kDa), cysteine-rich (forming three to six disulfide bridges), usually cationic (net charge in the range of +1 to +5) and have low amino acid identity (Shafee et al. 2017). Defensins are active against Gram-positive and Gram-negative bacteria, viruses and fungi (Ganz 2003; Brogden 2005). Other examples of eukaryotic disulfide bridge peptides are cyclotides which are plant defense peptides with three disulfide bridges and exceptional stability (De Veer et al. 2019), as well as conotoxins from cone snails with 2–5 disulfide bridges representing attractive leads for drug development and neurological applications (Jin et al. 2019).

Laterosporulins are defensin-like bacteriocins produced by *Brevibacillus* spp. containing six cysteines with conserved positions forming three intramolecular disulfide bridges (Singh et al. 2012; Baindara et al. 2016). Other defensin-like bacteriocins include cationic bactofensin (O’Shea et al. 2013; O’Connor et al. 2018), pediocin-like bacteriocins from which more than 20 peptides have been identified (Oppegård et al. 2015) and the recently discovered actifensins within the *Actinomyces* genus (Sugrue et al. 2020). Bioactive peptides with disulfide bridges occur more commonly in eukaryotes and infrequently in bacteria.

1.7.2 α-helical peptides

Peptides with an α-helical structure are cationic, and are therefore electrostatically drawn to the negatively charged microbial cell membrane (Tossi et al. 2000), which is the main target for the majority of cationic antimicrobial peptides (Mokherjee et al. 2020). The structure of a peptide of this class indicates that it might function as a pore former in contact with bacterial membranes (Tossi et al. 2000; Brogden 2005; Mokherjee et al. 2020; Figure 11). α-helical peptides assume amphipathic conformations in three dimensions usually only when in contact with membranes, and otherwise these compounds are highly flexible in solution (Hancock and Sahl 2006). Factors that can impact the activity and potency of α-helical peptides include the helical content, sequence, charge of the peptide, amphipathicity, hydrophobicity, and the widths of the hydrophilic and hydrophobic sites (Tossi et al. 2000). Modifying one factor can end up in remarkable changes to one or more of the other factors (Tossi et al. 2000).

The antimicrobial magainin peptides from the skin of the African clawed frog *Xenopus laevis* are structurally amphiphilic α-helices (Zasloff 1987). These peptides have been shown to be active against fungi, protozoa, and Gram-positive and Gram-negative bacteria (Zasloff 1987). A synthetic analog of magainin, Pexiganan, proceeded to phase III clinical trials, targeting the treatment of diabetic foot ulcers, but ultimately failed the trials (Fox 2013).

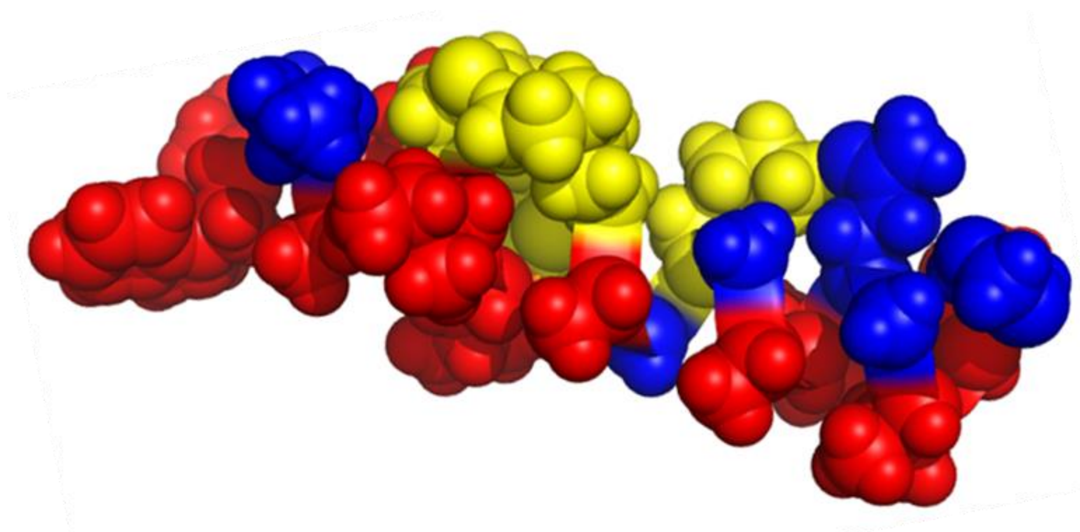


Figure 11. Structure of α-helical, amphipathic and charged peptide. Red color indicates hydrophobic, blue hydrophilic, and yellow charged amino acid residues. The model was created using a peptide sequence (FWGGVFDVLKKAGKGALQGVMS) discovered in this study. The PDB model was created with the iTasser-tool and drawn with PyMol.

2 Study aims

Bacteria are prolific sources of structurally divergent natural products, produced employing an array of promiscuous natural product enzymes (Fewer and Metsä-Ketelä 2019). The biosynthetic pathways of natural products in microorganisms harbor enormous potential for the drug industry and biotechnology. The aim of this study was to widen the understanding regarding the biochemical diversity of the cyanobactin family of natural products through the following means: i. discovering and exploring novel cyanobactin pathways from cyanobacteria and other bacteria; ii. testing the biosynthetic boundaries of the cyanobactin class; and iii. characterizing the end-products and underlying enzymatic machinery through genome mining. General scheme of this study is presented in Figure 12.

3 Summary of materials and methods

Table 1. Methods used in this study.

Method	Article
Genome sequencing	I
Genome mining	I-III
Cloning and heterologous expression	I-III
Bioinformatic analysis of gene clusters	I-III
Mass spectrometry (LC-MS & Q-TOF)	I-III
Compound purification	I, III
NMR	I, III
Cultivation of bacteria	I-III
DNA extraction	I-III
Stable isotope labeling	II
PCR and primer design	I-III
Antimicrobial bioassays and MIC	II
Protein overexpression	I, III
Enzymatic *in vitro* -assays	I, III
Site-directed mutagenesis	II
Organic synthesis	I

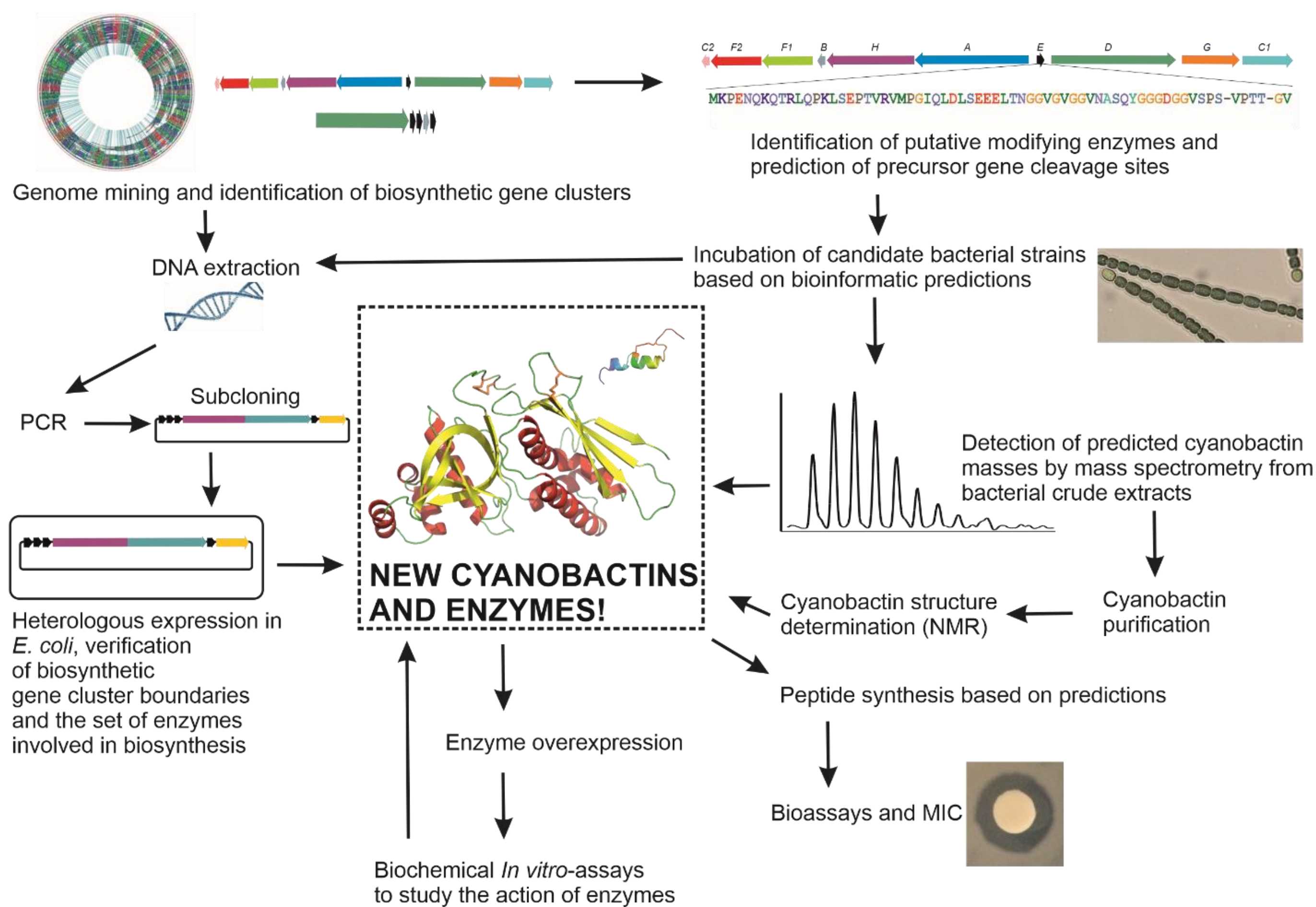


Figure 12. Overview of workflow in this study leading to discovery of new cyanobactin pathways.

4 Summary of results and discussion

A large mass of genome sequence data has shown that RiPP biosynthetic pathways are much more widespread than anticipated (Oman and van der Donk 2010), and an increasing number of structurally diverse peptides are emerging from this biosynthetic pathway type (Arnison et al. 2013). The cyanobactin pathway has now been studied for approximately 15 years, and with the boom of genome sequencing and mining, new enzymes and peptides belonging to the cyanobactin family of natural products continue to be discovered. Genome mining is a powerful approach which, if applied systematically across the bacterial domain, could greatly expand the known chemical and genetic diversity of cyanobactins.

4.1 Discovery and biosynthesis of peptide alkaloids muscoride A and B (I)

Muscoride A was identified from the cyanobacterium *Nostoc muscorum* IAM M-14 (Nagatsu et al. 1995), but the BGC for its biosynthesis was not described. The structure of muscoride A (Nagatsu et al. 1995) suggests that it might be produced through the cyanobactin biosynthetic pathway. We cultivated *Nostoc* sp. PCC 7906, which is an identical strain to *Nostoc muscorum* IAM M-14, to examine whether it produces muscoride A by LC-MS. We detected muscoride A (m/z 513.3), and based on this discovery, we sequenced the genome of *Nostoc* sp. PCC 7906 to mine and annotate the genes for muscoride A BGC (*mus*). The *mus* BGC encodes 10 conserved cyanobactin genes and it differs from the canonical cyanobactin biosynthetic pathway in key features. The *mus* BGC encodes only one protease and two prenyltransferases. Variants of PatF homologs are encoded in nearly all cyanobactin BGCs, but not all cyanobactins are prenylated (Schmidt and Donia 2009; Sivonen et al. 2010; Gu et al. 2018).

Bioinformatic analysis of unpublished *Nostoc* genomes and public databases led to the identification of 12 complete *mus* biosynthetic gene clusters. The precursor proteins encode predicted peptides with VPTT, VPTS, IPTS, IPSTS, or IPTSS sequences. The putative muscoride variant (IPSTS) from *Nostoc* sp. UHCC 0398 was subjected to mass spectrometric analysis, and the compound was detected (*m/z* 580.31) from the bacterial crude extract. A sufficient amount of the compound was purified from *Nostoc* sp. UHCC 0398 and subjected to NMR analysis. The 1D and 2D NMR analysis revealed three contiguous oxazoles and two prenyl moieties in forward orientation in the amino- and carboxy-termini of a linear compound, which was named muscoride B (Figure 13).

Inspection of *mus* BGCs and the structures of muscoride A & B suggest that MusF1 and MusF2 could potentially be regiospecific prenyltransferases. In order to test this hypothesis, we overexpressed the MusF1 and MusF2 proteins from *Nostoc* spp. PCC 7906 and UHCC 0398 and organically synthesized the core of the muscoride A backbone without prenyl moieties. Biochemical *in vitro* assays with overexpressed prenyltransferases together with the muscoride A core as a native substrate demonstrated the prenyltransferase enzymatic activity of the MusF1 and MusF2 proteins. The MusF1 enzymes from *Nostoc* spp. PCC 7906 and UHCC 0398 both catalyze the forward prenylation of the carboxy-terminus. The two MusF2 enzymes differ significantly as the MusF2 from *Nostoc* sp. PCC 7906 catalyze reverse prenylation whereas MusF2 from *Nostoc* sp. UHCC 0398 catalyze forward prenylation (Figure 13). Both muscoride BGCs were heterologously expressed, which provided further support concerning the boundaries of these gene clusters. This study expanded the genetic and chemical diversity of the cyanobactin family to include polyoxazole bisprenylated linear peptides, and the regiospecific activity of cyanobactin prenyltransferases.

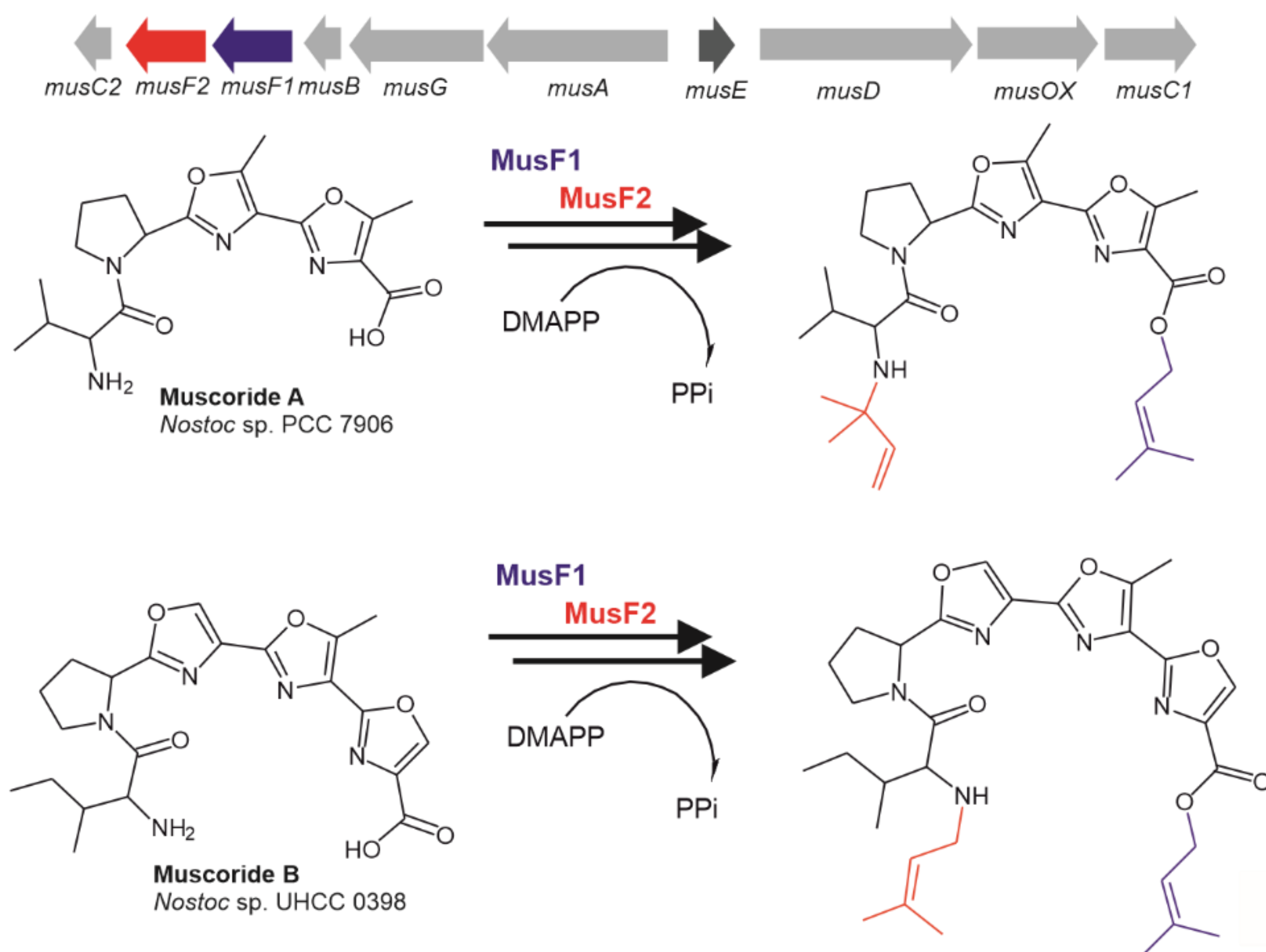


Figure 13. *mus* BGC illustrating regiospecific prenyltransferase activity of MusF1 and MusF2 enzymes.

4.2 Discovery of truncated cyanobactin BGC and unusual cyanobactins (II)

Since the discovery of cyanobactins, genome mining studies have gradually increased the genetic and chemical diversity of the family to include linear peptides which are highly modified (Leikoski et al. 2013). Genome mining was used to identify an unusual truncated cyanobactin BGC distributed across the bacterial domain encoding a cyanobactin protease together with 1–8 precursor proteins. The protease is encoded as a bimodular protein possessing a peptidase domain and a DUF (Figure 14 A). The truncated cyanobactin BGC protease shares about 40 % similarity to known cyanobactin proteases (Lee et al. 2009; Agarwal et al. 2012). An inspection of the precursor proteins revealed that the predicted cores encode an even number of cysteines, which suggests that they would form disulfide bridges (Figure 14 A & D). Disulfide bridge peptides are well known from eukaryotes (Ganz 2003; Shafee et al. 2017; De Veer et al. 2019; Jin et al. 2019), but rarely reported from bacteria.

We studied 12 candidate strains (nine cyanobacterial and three other bacterial strains) for the expression of bioinformatically predicted disulfide bridge stabilized peptides by LC-MS and Q-TOF. Three peptides were discovered by mass spectrometry from *Pseudanabaena* sp. PCC 6802, and were named pseudabactins A–C (data only shown for pseudabactin A) (Figure 14 A). All of these structurally novel cyanobactins had four intramolecular disulfide bridges. The cyanobacterium

Pseudanabaena sp. PCC 6802 was cultured in medium containing the ^{34}S isotope in order to authenticate the number of cysteines encoded in the peptides. Pseudabactin A with a larger mass was detected from bacterial crude extract cultured in media with the ^{34}S isotope (Figure 14 B). Heterologous expression was performed to demonstrate the boundaries of the gene cluster and that the gene cluster is active (Figure 14 C). The truncated cyanobactin BGC was amplified from *Pseudanabaena* sp. PCC 6802 with specific primers, restricted with restriction enzymes, ligated with the Puc57 vector to create a recombinant plasmid (p6802NGC), subcloned to *E. coli* Top10 -cells, and further heterologously expressed in *E. coli* expression strains. The expression level of pseudabactin A was higher compared to the native producer *Pseudanabaena* sp. PCC 6802 (Figure 14 B). A synthetic construct containing the codon optimized PsbA protease and the PsbE2 precursor protein (pseudabactin A) was also constructed to create a recombinant plasmid (p6802OPT). These results yielded further support for the existence of disulfide bridge stabilized novel cyanobactins (Figure 14 D).

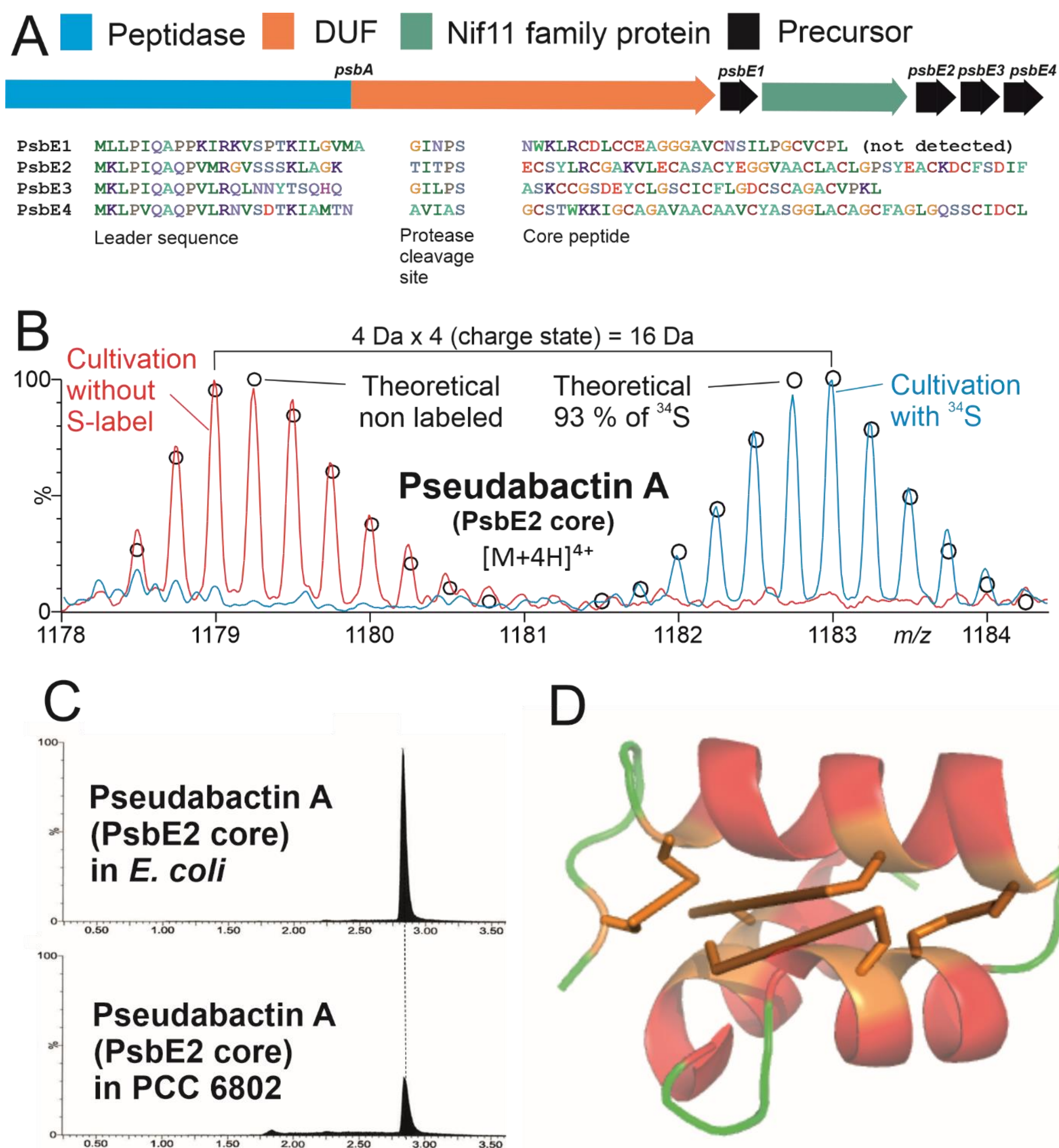


Figure 14. A) Truncated cyanobactin BGC from *Pseudanabaena* sp. PCC 6802 for producing disulfide bridge stabilized peptides from which three out of four were detected. B) In the presence of ^{34}S, a mass unit addition of 16 Da to pseudabactin A is detected, and when divided by charge (+4), a mass unit difference of 4 Da is visible. Pseudabactins A with 4 Da larger mass can be seen in the presence of ^{34}S but not in the presence of ^{32}S. C) Pseudabactin A detected (*m/z* 1179) from native producer cyanobacterium (PCC 6802) and heterologously expressed in *E. coli*. D) Model of pseudabactin A with four intramolecular disulfide bridges between cysteines representing one hypothetical disulfide connectivity pattern out of 105 possible ones. A PDB model of pseudabactin A was created using the iTasser-tool and drawn with Pymol.

Once the production of pseudabactins A–C had been confirmed, more detailed predictions of protease cleavage sites could be performed, followed by further genome mining. A subset of peptides was identified from the same family without cysteines encoded in their cores (Figure 15). The bioinformatic analysis suggested that these peptides would form α-helical structures based on their

core sequences (Figure 15). This is a known structure of antimicrobial peptides primarily produced by eukaryotes (Tossi et al. 2000: Mokherjee et al. 2020). Some of these peptides were synthesized and tested against a panel of bacteria in disc diffusion assays and for their minimal inhibitory concentration (MIC). The peptides were shown to be antimicrobial in the concentration range 0.6–22.5 µM (Table 2). The putative α-helical peptides were not detected from crude extracts from their native producer bacterium. The gene clusters for producing these peptides are most likely silent and become active when in contact with competing bacterial membranes (Hancock and Sahl 2006).

	Leader peptide	Protease cleavage site	Core peptide
PCC 6802 PseE2	MKLPIQAQPVMRGVSSSKLAGK	TITPS	ECSYLRCGAKVLECASACYEGGVAACLACLGPSYEACKDCFSDIF
PCC 6802 PseE3	MKLPIQAQPVLRQLNNYTSQHQ	GILPS	ASKCCGSDEYCLGSCICFLGDCSCAGACVPKL
PCC 6802 PseE4	MKLPVQAQPVLRNVSDTKIAMTN	AVIAS	GCSTWKKIGCAGAVAACAAVCYASGGLACAGCFAGLGQSSCIDCL
Dickeya zeae Ech1591	MKLPTQAQNVNRANRVAEAKAS	GVNPA	FWGEALKLLKKAGQGALSGVLS
Roseibium sp. TrichSKD4 B	MKLPYQAPSINREDRINSAIDS	SVAPA	FWGALASAAAPIAIDLGKKALRGLLR
Vibrio nigripulchritudo SOn1 A	MKLPTQAPGVDRSNRLEAAQAA	GVNPA	FWGTAFDILKKAGKGALQGVMS
Vibrio nigripulchritudo SOn1 B	MKLPTQAPGVDRSNRLEAAQAA	GVNPA	FWGGVFDVLKKAGKGALQGVMS
Yersinia frederiksenii ATCC 33641	MKLPIQAQSVNRSNRIAEAKMA	GVNPA	FWDSVLSVLKKAGQGALAGVLG

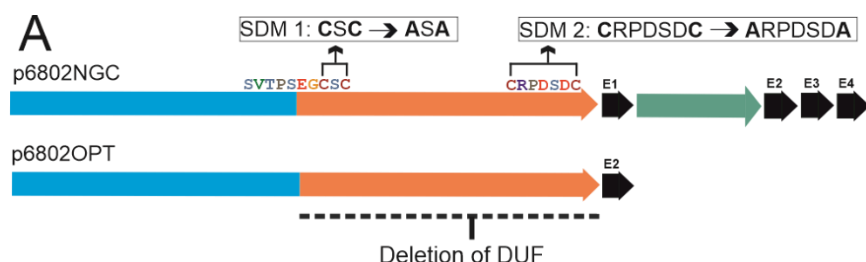
Figure 15. Precursor peptide alignment between detected cysteine-rich pseudabactins A–C (highlighted by black rectangle) and predicted cysteine-poor core peptides with supposed α-helical structures from different bacterial strains. These precursor peptides have conserved residues in their leader sequences (MKLPxQAxxxxR), while the core sequences are more variable. The synthetic analogs of the core peptides with supposed α-helical structures were shown to display antimicrobial activity.

Table 2. Minimal inhibitory concentration (MIC) values of five peptides synthesized based on predictions.

Producer strain (Antimicrobial peptide)	Peptide sequence	MIC (µM)			
		S. aureus	*B. subtilis*	*E. coli*	*M. luteus*
Dickeya zeae Ech1591	FWGEALKLLKKAGQGALSGVLS	22	11	11	2.8
Roseibium sp. TrichSKD4 B	FWGALASAAAPIAIDLGKKALRGLLR	4.7	2.3	18.6	0.6
Vibrio nigripulchritudo SOn1 A	FWGTAFDILKKAGKGALQGVMS	ND	10.8	21.5	2.7
Vibrio nigripulchritudo SOn1 B	FWGGVFDVLKKAGKGALQGVMS	21.8	10.9	21.8	2.7
Yersinia frederiksenii ATCC 33641	FWDSVLSVLKKAGQGALAGVLG	ND	22.5	ND	11.3

ND = not determined

The bioinformatic analysis of the DUF suggested that it might function as a disulfide bridge generating enzyme. We tested for disulfide bridge formation by replacing the Cys residues with Ala residues in the supposed catalytic sites in the DUF (N-terminal CSC site and C-terminal CRPDSDC site) and observed the outcome on pseudabactin A expression by Q-TOF (Figure 16). All thiol-disulfide oxidoreductases have the redox-active catalytic site CxxC, which is involved in disulfide bridge redox reactions (Bardwell et al. 1991; Berkmen 2012; Landeta et al. 2018), and we therefore hypothesized that one of these sites encoded in the DUF might be redox-active in oxidative folding of the substrate peptide pseudabactin A. There was no pseudabactin A production when the N-terminal Cys-residues were replaced with Ala residues (CSC→ASA), or when the DUF was enzymatically deleted (Figure 16). Replacing the Cys residues with Ala residues in the C-terminus (CRPDSDC→ARPDSDA) did not abolish pseudabactin A expression but hindered it (Figure 16).



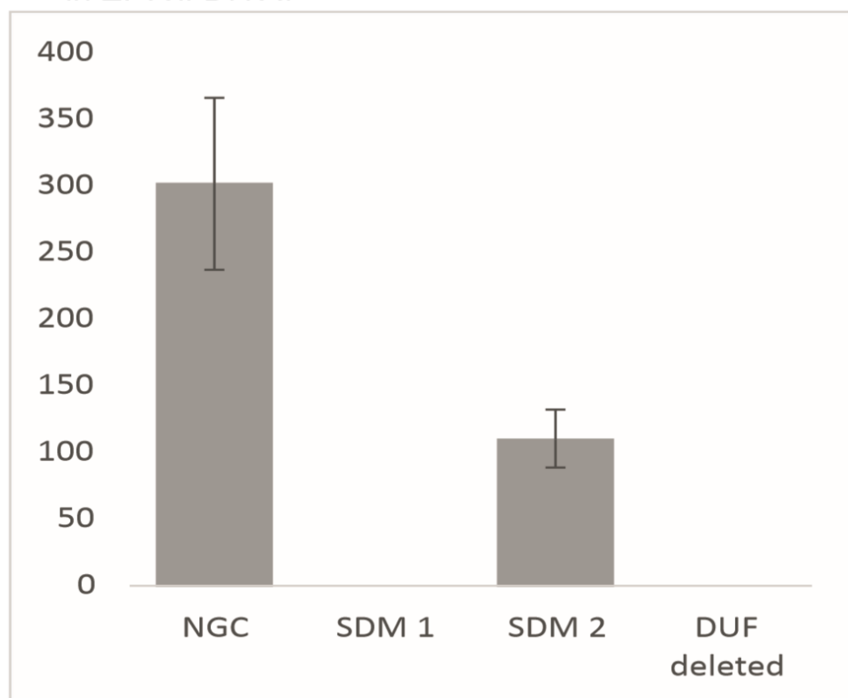


Figure 16. A) Site-directed mutagenesis experiments performed on native gene cluster (p6802NGC). DUF was enzymatically deleted from the optimized gene cluster (p6802OPT). B) Pseudabactin A production was not detected (i. e. disulfide bridge formation did not occur) when supposed catalytic site cysteines were mutated (SDM 1) or when the DUF was deleted enzymatically. The mutation of the CX_5C-site (SDM 2) hindered peptide production but did not abolish it.

These results suggest that the DUF is important in disulfide bridge formation in pseudabactin biosynthesis. It is tempting to speculate that the DUF of the protease is responsible for disulfide bridge formation in the macrocyclic cyanobactins ulithiacyclamide (Ireland and Scheuer 1980) and piricyclamides (Leikoski et al. 2012). In conclusion, this study expanded the genetic and chemical diversity of the cyanobactin family to include disulfide bridge and α-helical peptides, suggesting that the DUF is a potential disulfide bridge generating oxidase, and suggesting re-defining the minimal cyanobactin BGC to include only the genes encoding for the protease and precursor(s).

4.3 Biosynthesis of prenylated anacyclamide (III)

Anacyclamides are highly diverse macrocyclic cyanobactins (Leikoski et al. 2010; Leikoski et al. 2013; Martins et al. 2018). Cyclic anacyclamides consist of 7–20 proteinogenic amino acids and lack heterocyclized residues in their structure (Leikoski et al. 2010). The anacyclamide BGC from *Anabaena* sp. UHCC 0232 encompasses AcyA, AcyB, AcyC, AcyE, AcyF, AcyG, and an additional two open reading frames (ORFs) with unclear functions in anacyclamide biosynthesis (Figure 17). The PTMs for the AcyE core (HQPWHAAP) include proteolytic cleavage, macrocyclization and prenylation (Leikoski et al. 2010). The cyanobacterium *Anabaena* sp. UHCC 0232 was cultivated and its methanol extract was analyzed by LC-MS. It was shown that *Anabaena* sp. UHCC 0232 produces the prenylated anacyclamide A8P (*m*/*z* 993.50), as reported previously (Leikoski et al. 2010). The successful heterologous expression of the *acy* gene cluster confirmed the boundaries of the gene cluster. However, detection of anacyclamide by mass or heterologous expression could not unequivocally designate the orientation or the position of the prenyl moiety.

KgpF is a Trp C-prenyltransferase from the biosynthetic pathway of the cyanobactin kawaguchipeptin (Parajuli et al. 2016; Okada et al. 2016). We hypothesized that AcyF is a Trp N-prenyltransferase from the biosynthetic pathway of the cyanobactin anacyclamide. A two-plasmid expression system was constructed to compare the functions of AcyF and KgpF. A minimal *acy* gene cluster encoding AcyA–AcyG was constructed to produce the pABCFG plasmid. The plasmid pSYNE was constructed by replacing the native anacyclamide HQPWHAAP core with the kawaguchipeptin WLNGDNNWSTP core sequence. Simultaneous expression of these two plasmids in *E. coli* BL21 resulted in the detection of bisprenylated peptides with the same mass as kawaguchipeptin A with the enzymatic machinery from the anacyclamide (*acy*) biosynthetic pathway. Kawaguchipeptins produced by the *kgp* pathway (Parajuli et al. 2016) had different retention times compared to the bisprenylated kawaguchipeptin produced by the *acy* pathway. This result suggests that there are likely to be regioselective prenylation differences between the *acy* and *kgp* biosynthetic pathways.

AcyF was overexpressed in order to further examine the catalytic activity of this putative Trp N-prenyltransferase in biochemical assays. The anacyclamide A10 core (cyclo [TSQIWGSPVP]) (Leikoski et al. 2010), already as synthetic in our lab, was expected to be an applicable substrate for the AcyF enzyme. An *in vitro* assay was performed where the AcyF was mixed together with the substrate and DMAPP as a donor substrate, and analyzed by LC-MS. The mass of the substrate increased by 68.06 Da suggesting the incorporation of a prenyl moiety to the compound. The whole reaction was scaled up, and the reaction product was purified using HPLC and subjected to NMR. NMR analysis was performed to acquire firm evidence of the position in the indole ring of Trp and the orientation of the prenyl moiety. The prenyltransferase AcyF was able to prenylate protected Trp derivatives, z-Trp-OH and Boc-Trp-OH, as well. The analysis demonstrated that the prenyl group is attached to the Trp indole ring nitrogen in forward orientation. In this study, the AcyF regiospecific prenyltransferase from the cyanobactin biosynthetic pathway of *Anabaena* sp. UHCC 0232 was characterized (Figure 17).

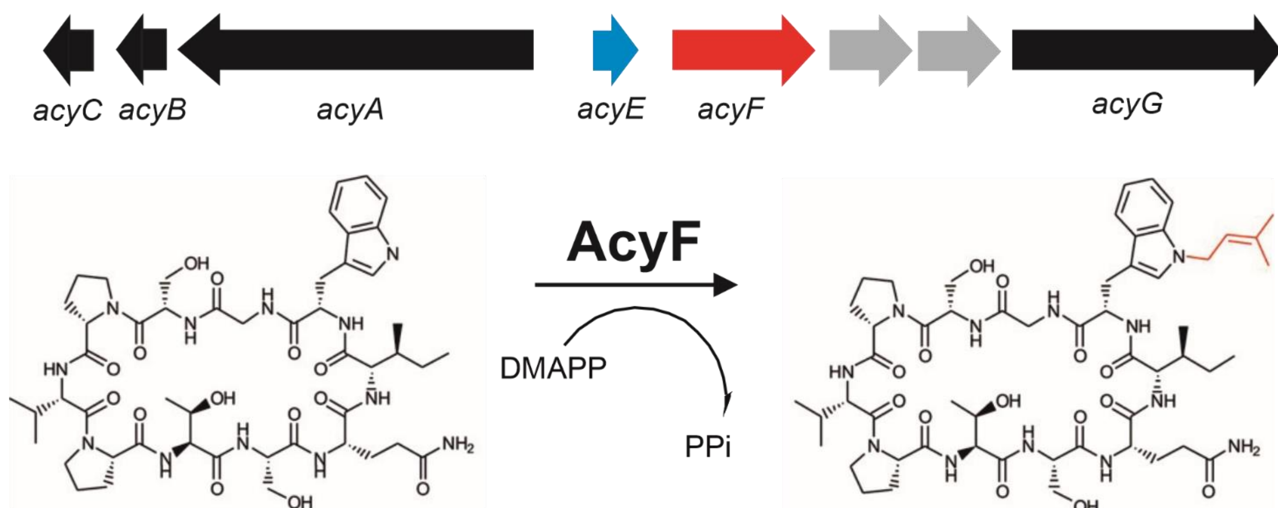


Figure 17. Biosynthetic *acy* gene cluster illustrating regiospecific prenyltransferase activity of AcyF.

5 Conclusions

This thesis demonstrates that the cyanobactin family of natural product peptides is much more widely distributed and structurally diverse than currently appreciated. Genome mining combined with mass spectrometric methods, molecular biology, biochemical assays and structural chemistry has provided us with insights on cyanobactin biosynthesis.

Cyanobactin biosynthesis involves multiple post-translational modifications leading to an astonishing array of compounds. This work has increased knowledge regarding the enzymology of cyanobactin biosynthetic pathways, as well as discovering novel post-translational modification enzymes and structurally novel cyanobactins. We have identified regiospecific prenyltransferases and a hypothetical disulfide bridge generating enzyme. The study has broadened the chemical diversity of cyanobactins to include bis-prenylated polyoxazole linear peptides, disulfide bridge stabilized peptides, and α-helical peptides as products from cyanobactin pathways. Together the results of this PhD thesis expand the cyanobactin family considerably and will affect how cyanobactins are determined in the future. Cyanobactin pathways are an abundant source of modification enzymes and novel natural product peptides, most likely including numerous compounds still awaiting to be discovered.

6 Acknowledgements

This work was carried out at the Department of Microbiology, Faculty of Agriculture and Forestry, University of Helsinki, Finland. Doctoral candidate funding was provided as a salaried position by the Doctoral Programme in Microbiology and Biotechnology, University of Helsinki.

I am deeply grateful to my brilliant supervisor Docent David Fewer for endless discussions, mentoring and motivational teaching throughout all these years. I want to express my sincerest gratitude to my second supervisor Professor Kaarina Sivonen for guidance and advice. Kaarina is the reason I entered this great research group in the first place. I thank Matti Wahlsten for skilled technical work, constructive ideas and discussions. I thank Docent Jouni Jokela for valuable help regarding chemical methods and discussions. Luydmila Saari is acknowledged for her valuable help with cyanobacterial cultures. I thank the members of my steering group: Professor Jari Yli-Kauhaluoma and Associate Professor Mikko Metsä-Ketelä. I want to thank Professor Marcel Jaspars and Docent Jarmo Niemi for pre-examination of this thesis, and Professor Nadine Ziemert for agreeing to act as my opponent. I am grateful for all my colleagues in the Department of Microbiology and my co-authors.

Special thanks go out to my fellow PhD candidates: Lassi, Muhammad, Rafael, Sila, Julia and Maria. Thank you for sharing thoughts and moments in the lab. I thank Johannes Cairns for language revision. A very special thanks go to my colleagues Jonna and Sebastián for endless support in science and for friendship. Also, a big thanks for boosting me to finish this thesis!

I want to express my deepest gratitude to all my friends. Thank you for your friendship and for being there.

Finally, I would like to thank my family. A very special thanks to my wife Silva for everything you have done for me and for our family. I thank Mellu for lending her room for my writing work and boosting my PhD work. I thank my brother Mikko for being there. The biggest thanks go to my beloved son Joel for everything you are.

7 References

Agarwal V, Pierce E, McIntosh J, Schmidt EW and Nair SK (2012). Structures of cyanobactin maturation enzymes define a family of transamidating proteases. Chem. Biol. 19: 1411–1422.

Ageitos JM, Sánchez-Pérez A, Calo-Mata P and Villa TG (2017). Antimicrobial peptides (AMPs): ancient compounds that represent novel weapons in the fight against bacteria. Biochem. Pharmacol. 133: 117–138.

Alexandru-Crivac CN, Umeobika C, Leikoski N, Jokela J, Rickaby KA, Grilo AM, Sjö P, Plowright AT, Idress M, Siebs E, Nneoyi-Egbe A, Wahlsten M, Sivonen K, Jaspars M, Trembleau L, Fewer DP and Houssen WE (2017). Cyclic peptide production using a macrocyclase with enhanced substrate promiscuity and relaxed recognition determinants. Chem. Commun. 53: 10656–10659.

Arnison PG, Bibb MJ, Bierbaum G, Bowers AA, Bugni TS, Bulaj G, Camarero JA, Campopiano DJ, Challis GL, Clardy J, Cotter PD, Craik DJ, Dawson M, Dittmann E, Donadio S, Dorrestein PC, Entian KD, Fischbach MA, Garavelli JS, Göransson U, Gruber CW, Haft DH, Hemscheidt TK, Hertweck C, Hill C, Horswill AR, Jaspars M, Kelly WL, Klinman JP, Kuipers OP, Link AJ, Liu W, Marahiel MA, Mitchell DA, Moll GN, Moore BS, Müller R, Nair SK, Nes IF, Norris GE, Olivera BM, Onaka H, Patchett ML, Piel J, Reaney MJ, Rebuffat S, Ross RP, Sahl HG, Schmidt EW, Selsted ME, Severinov K, Shen B, Sivonen K, Smith L, Stein T, Süssmuth RD, Tagg JR, Tang GL, Truman AW, Vederas JC, Walsh CT, Walton JD, Wenzel SC, Willey JM and van der Donk WA (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. Nat. Prod. Rep. 30: 108–160.

Bachmann BO, Van Lanen SG and Baltz RH (2014). Microbial genome mining for accelerated natural products discovery: is a renaissance in the making? J. Ind. Microbiol. Biotechnol. 41: 175–184.

Baindara P, Singh N, Ranjan M, Nallabelli N, Chaudhry V, Pathania GL, Sharma N, Kumar A, Patil PB and Korpole S (2016). Laterosporulin10: a novel defensin like Class IId bacteriocin from *Brevibacillus* sp. strain SKDU10 with inhibitory activity against microbial pathogens. Microbiology. 162: 1286–1299.

Baltz RH (2008). Renaissance in antibacterial discovery from actinomycetes. Curr. Opin. Pharmacol. 8: 557–563.

Bardwell JC, McGovern K and Beckwith J (1991). Identification of a protein required for disulfide bond formation in vivo. Cell. 67: 581–589.

Bauermeister A, Branco PC, Furtado LC, Jimenez PC, Costa-Lotufo LV, da Cruz Lotufo TM (2019). Tunicates: A model organism to investigate the effects of associated-microbiota on the production of pharmaceuticals. Drug Discov. Today Dis. Models. ISSN 1740-6757.

Belknap KC, Park CJ, Barth BM and Andam CP (2020). Genome mining of biosynthetic and chemotherapeutic gene clusters in Streptomyces bacteria. Sci. Rep. 10: 2003.

Bent AF, Koehnke J, Houssen WE, Smith MCM, Jaspars M and Naismith JH (2013). Structure of PatF from *Prochloron didemni*. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 69: 618–623.

Bent AF, Mann G, Houssen WE, Mykhaylyk V, Duman R, Thomas L, Jaspars M, Wagner A, Naismith JH (2016). Structure of the cyanobactin oxidase ThcOx from *Cyanothece* sp. PCC 7425, the first structure to be solved at diamond light source beamline I23 by means of S-SAD. Acta Crystallogr. D Struct. Biol. 72: 1174–1180.

Berdy J (2005). Bioactive microbial metabolites. A personal view. J. Antibiot. 58: 1–26.

Berkmen M (2012). Production of disulfide-bonded proteins in *Escherichia coli*. Protein Expr. Purif. 82: 240–251.

Bertram A and Pattenden G (2007). Marine metabolites: metal binding and metal complexes of azole-based cyclic peptides of marine origin. Nat. Prod. Rep. 24: 18–30.

Bocian-Ostrzycka KM, Grzeszczuk MJ, Banas AM and Jagusztyn-Krynicka EK (2017). Bacterial thiol oxidoreductases - from basic research to new antibacterial strategies. Appl. Microbiol. Biotechnol. 101: 3977–3989.

Bonitz T, Alva V, Saleh O, Lupas AN and Heide L (2011). Evolutionary relationships of microbial aromatic prenyltransferases. PLoS ONE, 6, e27336.

Botta B, Vitali A, Menendez P, Misiti D and Monache GD (2005). Prenylated flavonoids: Pharmacology and biotechnology. Curr. Med. Chem. 12: 713–739.

Brogden KA (2005). Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? Nat. Rev. Microbiol. 3: 238–250.

Burja AM, Banaigs B, Abou-Mansour E, Burguess JG and Wright PC (2001). Marine cyanobacteria—a profilic source of natural products. Tetrahedron 57: 9347–9377.

Burkhart BJ, Hudson GA, Dunbar KL and Mitchell DA (2015). A prevalent peptide-binding domain guides ribosomal natural product biosynthesis. Nat. Chem. Biol. 11: 564–570.

Burkhart BJ, Schwalen CJ, Mann G, Naismith JH and Mitchell DA (2017). YcaO-dependent posttranslational amide activation: Biosynthesis, structure, and function. Chem. Rev. 117: 5389–5456.

Carroll AR, Copp BR, Davis RA, Keyzers RA and Prinsep MR (2019). Marine natural products. Nat. Prod. Rep. 36: 122–173.

Challis GL (2008). Genome mining for novel natural product discovery. J. Med. Chem. 51: 2618–2628.

Crnkovic CM, Braesel J, Krunic A, Eustáquio AS and Orjala J (2019). Scytodecamide from the cultured *Scytonema* sp. UIC 10036 expands the chemical and genetic diversity of cyanobactins. ChemBioChem. 21: 845–852.

Czekster CM, Ge Y and Naismith JH (2016). Mechanisms of cyanobactin biosynthesis. Curr. Opin. Chem. Biol. 35: 80–88.

Davey L, Ng CKW, Halperin SA, and Lee SF (2013). Functional analysis of paralogous thiol disulfide oxidoreductases in *Streptococcus gordonii*. J. Biol. Chem. 288: 16416–16429.

Davey L, Halperin SA and Lee SF (2016). Thiol-disulfide exchange in gram-positive *Firmicutes*. Trends Microbiol. 24: 902–915.

De Veer SJ, Kan MW and Craik DJ (2019). Cyclotides: from structure to function. Chem. Rev. 119: 12375–12421.

Demain AL and Sanchez S (2009). Microbial drug discovery: 80 years of progress. J. Antibiot. (Tokyo) 62: 5–16.

Demay J, Bernard C, Reinhardt A and Marie B (2019). Natural products from cyanobacteria: Focus on beneficial activities. Mar. Drugs. 17: 320.

Deng C, Pan B and O'Connor OA (2013). Brentuximab vedotin. Clin. Cancer Res. 19: 22–27.

Dittmann E, Gugger M, Sivonen K and Fewer DP (2015). Natural product biosynthetic diversity and comparative genomics of the cyanobacteria. Trends Microbiol. 23: 642–652.

Dixit RB and Suseela MR (2013). Cyanobacteria: potential candidates for drug discovery, Antonie van Leeuwenhoek 103: 947–961.

Donia MS, Hathaway BJ, Sudek S, Haygood MG, Rosovitz MJ, Ravel J and Schmidt EW (2006). Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. Nat. Chem. Biol. 2: 729–735.

Donia MS, Ravel J and Schmidt EW (2008). A global assembly line for cyanobactins, Nat. Chem. Biol. 4: 341–343.

Donia MS and Schmidt EW (2010). Cyanobactins – ubiquitous cyanobacterial ribosomal peptide metabolites. In: Comprehensive natural products II chemistry and biology. Elsevier, Oxford. 539–558.

Donia MS and Schmidt EW (2011). Linking chemistry and genetics in the growing cyanobactin natural products family. Chem. Biol. 18: 508–519.

Du X, Liu H, Yuan L, Wang Y, Ma Y, Wang R, Chen X, Losiewicz MD, Guo H and Zhang H (2019). The diversity of cyanobacterial toxins on structural characterization, distribution and identification: A systematic review. Toxins. 11: 530.

Dunbar KL, Melby JO and Mitchell DA (2012). YcaO Domains Use ATP to activate amide backbones during peptide cyclodehydrations. Nat. Chem. Biol. 8: 569−575.

Dunbar KL and Mitchell DA (2013). Revealing nature´s synthetic potential through the study of ribosomal natural product biosynthesis. ACS Chem. Biol. 8: 473–487.

Dutta S, Mahalanobish S, Saha S, Ghosh S and Sil PC (2019). Natural products: An upcoming therapeutic approach to cancer. Food Chem. Toxicol. 128: 240–255.

Epstein SC, Charkoudian LK and Medema MH (2018). A standardized workflow for submitting data to the Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository: prospects for research-based educational experiences. Stand. Genom. Sci. 13: 16.

Estrada P, Morita M, Hao Y, Schmidt EW and Nair SK (2018). A single amino acid switch alters the isoprene donor specificity in ribosomally synthesized and post-translationally modified peptide prenyltransferases. J. Am. Chem. Soc. 140: 8124–8127.

Fass D and Thorpe C (2018). Chemistry and enzymology of disulfide cross-linking in proteins. Chem. Rev. 118: 1169–1198.

Fewer DP & Metsa-Ketelä M (2019). A pharmaceutical model for the molecular evolution of microbial natural products. FEBS J. 287: 1429–1449.

Finking R and Marahiel MA (2004). Biosynthesis of nonribosomal peptides. Annu. Rev. Microbiol. 58: 453–488.

Fox JL (2013). Antimicrobial peptides stage a comeback. Nat. Biotechnol. 31: 379–382.

Ganz T (2003). Defensins: antimicrobial peptides of innate immunity. Nat. Rev. Immunol. 3: 710–720.

Gao S, Ge Y, Bent AF, Schwarz-Linek U and Naismith JH (2018). Oxidation of cyanobactin precursor peptide is independent of leader peptide and operates in a defined order. Biochemistry 57: 5996–6002.

Ge Y, Czekster CM, Miller OK, Botting CH, Schwarz-Linek U and Naismith JH (2019). Insights into the mechanism of the cyanobactin heterocyclase enzyme. Biochemistry 58: 2125–2132.

Gleiter S and Bardwell JC (2008). Disulfide bond isomerization in prokaryotes. Biochim. Biophys. Acta. 4: 530–534.

Gu W, Dong SH, Sarkar S, Nair SK and Schmidt EW (2018). The biochemistry and structural biology of cyanobactin pathways: Enabling combinatorial biosynthesis. Methods Enzymol. 604: 113–163.

Gurevich A, Mikheenko A, Shlemov A, Korobeynikov A, Mohimani H and Pevzner PA (2018). Increased diversity of peptidic natural products revealed by modification-tolerant database search of mass spectra. Nat. Microbiol. 3: 319–327.

Góngora-Benítez M, Tulla-Puche J and Albericio F (2013). Multifaceted roles of disulfide bonds. Peptides as therapeutics. Chem. Rev. 114: 901–926.

Hancock RE and Sahl HG (2006). Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. Nat. Biotechnol. 24: 1551–1557.

Hassan M, Kjos M, Nes IF, Diep DB and Lotfipour F (2012). Natural antimicrobial peptides from bacteria: Characteristics and potential applications to fight against antibiotic resistance. J. Appl. Microbiol, 113: 723–736.

Hao Y, Pierce E, Roe D, Morita M, McIntosh JA, Agarwal V, Cheatham TE III, Schmidt EW and Nair SK (2016). Molecular basis for the broad substrate selectivity of a peptide prenyltransferase. Proc. Natl. Acad. Sci. USA. 113: 14037–14042.

van Heel AJ, de Jong A, Montalban-Lopez M, Kok J and Kuipers OP (2013). BAGEL3: automated identification of genes encoding bacteriocins and (non-) bactericidal posttranslationally modified peptides. Nucleic Acids Res. 41: 448–453.

Heras B, Kurz M, Jarrott R, Shouldice SR, Frei P, Robin G, Čemažar M, Thöny-Meyer L, Glockshuber R, and Martin JL (2008). *Staphylococcus aureus* DsbA does not have a destabilizing disulfide: a new paradigm for bacterial oxidative folding. J. Biol. Chem. 283: 4261–4271.

Heras B, Shouldice SR, Totsika M, Scanlon MJ, Schembri MA and Martin JL (2009). DSB proteins and bacterial pathogenicity. Nat. Rev. Microbiol. 7: 215–225.

Hetrick KJ and van der Donk WA (2017). Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. Curr. Opin. Chem. Biol. 38: 36–44.

Houssen WE and Jaspars M (2010). Azole-based cyclic peptides from the sea squirt *Lissoclinum patella*: Old scaffolds, new avenues. ChemBioChem. 11: 1803–1815.

Houssen WE, Koehnke J, Zollman D, Vendome J, Raab A, Smith MCM, Naismith JH and Jaspars M (2012). The discovery of new cyanobactins from *Cyanothece* PCC 7425 defines a new signature for processing of patellamides. ChemBioChem 13: 2683–2689.

Houssen WE, Bent AF, McEwan AR, Pieiller N, Tabudravu J, Koehnke J, Mann G, Adaba RI, Thomas L, Hawas UW, Liu H, Schwarz-Linek U, Smith MC, Naismith JH and Jaspars M (2014). An efficient method for the in vitro production of azol(in)e-based cyclic peptides. Angew. Chem. Int. Ed. 53: 14171–14174.

Hudson GA and Mitchell DA (2018). RiPP antibiotics: biosynthesis and engineering potential. Curr. Opin. Microbiol. 45: 61–69.

Ishida K, Matsuda H, Murakami M and Yamaguchi K (1996). Kawaguchipeptin A, a novel cyclic undecapeptide from cyanobacterium *Microcystis aeruginosa* (NIES-88). Tetrahedron 52: 9025–9030.

Ishida K, Matsuda H, Murakami M, Yamaguchi K (1997). Kawaguchipeptin B, an antibacterial cyclic undecapeptide from the cyanobacterium *Microcystis aeruginosa*. J. Nat. Prod. 60: 724–726.

Ireland C and Scheuer PJ (1980). Ulicyclamide and ulithiacyclamide, two new small peptides from a marine tunicate. J. Am. Chem. Soc. 102: 5688–5691.

Iwasaki K, Iwasaki A, Sumimoto S, Sano T, Hitomi Y, Ohno O and Suenaga K (2018). Croissamide, a proline-rich cyclic peptide with an *N*-prenylated tryptophan from a marine cyanobacterium *Symploca* sp. Tetrahedron Lett. 42: 3806–3809.

Jensen PR (2016). Natural products and the gene cluster revolution. Trends Microbiol. 24: 968–977.

Jin AH, Muttenthaler M, Dutertre S, Himaya SWA, Kaas Q, Craik DJ, Lewis RJ and Alewood PF (2019). Conotoxins: Chemistry and biology. Chem. Rev. 119: 11510–11549.

Kadokura H, Katzen F and Beckwith J (2003). Protein disulfide bond formation in prokaryotes. Annu. Rev. Biochem. 72: 111–135.

Kadokura H and Beckwith J (2010). Mechanisms of oxidative protein folding in the bacterial cell envelope. Antioxid. Redox Signal. 13: 1231–1246.

Kalkreuter E, Pan G, Cepeda AJ and Shen B (2020). Targeting bacterial genomes for natural product discovery. Trends Pharmacol. Sci. 4: 13–26.

Katz L, Baltz RH (2016). Natural product discovery: past, present, and future. J. Ind. Microbiol. Biotechnol. 43: 155–176.

Kaya K, Mahakhant A, Keovara L, Sano T, Kubo T and Takagi H (2002). Spiroidesin, a novel lipopeptide from the cyanobacterium *Anabaena spiroides* that inhibits cell growth of the cyanobacterium *Microcystis aeruginosa*. J. Nat. Prod. 65: 920–921.

Koehnke J, Bent A, Houssen WE, Zollman D, Morawitz F, Shirran S, Vendome J, Nneoyiegbe AF, Trembleau L, Botting CH, Smith MC, Jaspars M and Naismith JH (2012). The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. Nat. Struct. Mol. Biol. 8: 767–772.

Koehnke J, Bent AF, Zollman D, Smith K, Houssen WE, Zhu X, Mann G, Lebl T, Scharff R, Shirran S, Botting CH, Jaspars M, Schwarz-Linek U and Naismith JH (2013). The cyanobactin heterocyclase enzyme: a processive adenylase that operates with a defined order of reaction. Angew. Chem. Int. Ed. 52: 13991–13996.

Koehnke J, Bent AF, Houssen WE, Mann G, Jaspars M and Naismith JH (2014). The structural biology of patellamide biosynthesis. Curr. Opin. Struct. Biol. 29: 112–121.

Koehnke J, Mann G, Bent AF, Ludewig H, Shirran S, Botting C, Lebl T, Houssen WE, Jaspars M, Naismith JH (2015). Structural analysis of leader peptide binding enables leader-free cyanobactin processing. Nat. Chem. Biol. 11: 558–563.

Kopp F and Marahiel MA (2007). Where chemistry meets biology: the chemoenzymatic synthesis of nonribosomal peptides and polyketides. Curr. Opin. Biotechnol. 18: 513–520.

Kramer J, Özkaya Ö and Kümmerli R (2019). Bacterial siderophores in community and host interactions. Nat. Rev. Microbiol. 18: 152–163.

Kuzuyama T, Noel JP and Richard SB (2005). Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. Nature. 435: 983–987.

Lai Y and Gallo RL (2009). AMPed up immunity: how antimicrobial peptides have multiple roles in immune defense. Trends Immunol. 30: 131–141.

Landeta C, Boyd D and Beckwith J (2018). Disulfide bond formation in prokaryotes. Nat. Microbiol. 3: 270–280.

Lee J, McIntosh J, Hathaway BJ and Schmidt EW (2009). Using marine natural products to discover a protease that catalyses peptide macrocyclization of diverse substrates. J. Am. Chem. Soc. 131: 2122–2124.

Leikoski N, Fewer DP and Sivonen K (2009). Widespread occurrence and lateral transfer of the cyanobactin biosynthesis gene cluster in cyanobacteria. Appl. Environ. Microbiol. 75: 853–857.

Leikoski N, Fewer DP, Jokela J, Wahlsten M, Rouhiainen L and Sivonen K (2010). Highly diverse cyanobactins in strains of the genus *Anabaena*. Appl. Environ. Microbiol. 76: 701–709.

Leikoski N, Fewer DP, Jokela J, Alakoski P, Wahlsten M and Sivonen K (2012). Analysis of an inactive cyanobactin biosynthetic gene cluster leads to discovery of new natural products from strains of the genus *Microcystis*. PLoS ONE 7: e43002.

Leikoski N, Liu L, Jokela J, Wahlsten M, Gugger M, Calteau A, Permi P, Kerfeld CA, Sivonen K and Fewer DP (2013). Genome mining expands the chemical diversity of the cyanobactin family to include highly modified linear peptides. Chem. Biol. 20: 1–11.

Lewis K (2017). New approaches to antimicrobial discovery. Biochem. Pharmacol. 134: 87–98.

Li B, Sher D, Kelly L, Shi Y, Huang K, Knerr PJ, Joewono I, Rusch D, Chisholm SW, van der Donk WA (2010). Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. Proc. Natl. Acad. Sci. USA, 107: 10430–10435.

Li Y and Rebuffat S (2019). The manifold roles of microbial ribosomal peptide–based natural products in physiology and ecology. J. Biol. Chem. 295: 34–54.

Long PF, Dunlap WC, Battershill CN and Jaspars M (2005). Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production. ChemBioChem. 6: 1760–1765.

Lopez JAV, Al-Lihaibi SS, Alarif WM, Abdel-Lateff A, Nogata Y, Washio K, Morikawa M and Okino T (2016). Wewakazole B, a cytotoxic cyanobactin from the cyanobacterium *Moorea producens* collected in the Red Sea. J. Nat. Prod. 79: 1213–1218.

Luo Y, Cobb RE and Zhao H (2014). Recent advances in natural product discovery. Curr. Opin. Biotechnol. 30: 230–237.

Mann G, Koehnke J, Bent AF, Graham R, Houssen W, Jaspars M, Schwarz-Linek U, Naismith JH (2014). The structure of the cyanobactin domain of unknown function from PatG in the patellamide gene cluster. Acta Crystallogr. 70: 1597–1603.

Martins J and Vasconcelos V (2015). Cyanobactins from cyanobacteria: Current genetic and chemical state of knowledge. Mar. Drugs. 13: 6910–6946.

Martins J, Leikoski N, Wahlsten M, Azevedo J, Antunes A, Jokela J, Sivonen K, Vasconcelos V, Fewer DP and Leão PN (2018). Sphaerocyclamide, a prenylated cyanobactin from the cyanobacterium *Sphaerospermopsis* sp. LEGE 00249. Sci. Rep. 8: 14537.

Marahiel MA, Stachelhaus T and Mootz HD (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem. Rev. 97: 2651–2674.

McIntosh JA, Donia MS and Schmidt EW (2009). Ribosomal peptide natural products: Bridging the ribosomal and nonribosomal worlds. Nat. Prod. Rep. 26: 537–559.

McIntosh JA and Schmidt EW (2010). Marine molecular machines: Heterocyclization in cyanobactin biosynthesis. ChemBioChem, 11: 1413–1421.

McIntosh JA, Donia MS, Nair SK and Schmidt EW (2011). Enzymatic basis of ribosomal peptide prenylation in cyanobacteria. J. Am. Chem. Soc. 133: 13698–13705.

McIntosh JA, Lin Z, Tianero MD and Schmidt EW (2013). Aestuaramides, a natural library of cyanobactin cyclic peptides resulting from isoprene-derived Claisen rearrangements. ACS Chem. Biol. 8: 877–883.

Merwin NJ, Mousa WK, Dejong CA, Skinnider MA, Cannon MJ, Li H, Dial K, Gunabalasingam M, Johnston C and Magarvey NA (2020). DeepRiPP integrates multiomics data to automate discovery of novel ribosomally synthesized natural products. Proc. Natl. Acad. Sci. USA. 117: 371–380.

Metzger U, Schall C, Zocher G, Unsöld I, Stec E, Li S, Heide L and Stehle T (2009). The structure of dimethylallyl tryptophan synthase reveals a common architecture of aromatic prenyltransferases in fungi and bacteria. Proc. Natl. Acad. Sci. USA. 106: 14309–14314.

Micallef ML, D'Agostino PM, Al-Sinawi B, Neilan BA and Moffitt MC (2015). Exploring cyanobacterial genomes for natural product biosynthesis pathways. Mar. Genomics. 21: 1–12.

Milne BF, Long PF, Starcevic A, Hranueli D and Jaspars M (2006). Spontaneity in the patellamide biosynthetic pathway. Org. Biomol. Chem. 4: 631–638.

Mohimani H, Kersten RD, Liu WT, Wang MX, Purvine SO, Wu S, Brewer HM, Pasa-Tolic L, Bandeira N, Moore BS, Pevzner PA and Dorrestein PC (2014). Automated genome mining of ribosomal peptide natural products ACS Chem. Biol. 9: 1545–1551.

Mookherjee N, Anderson MA, Haagsman HP and Davidson DJ (2020). Antimicrobial host defence peptides: Functions and clinical potential. Nat. Rev. Drug Discov. 19: 311–332.

Morita M, Hao Y, Jokela JK, Sardar D, Lin Z, Sivonen K, Nair SK and Schmidt EW (2018). Post-translational tyrosine geranylation in cyanobactin biosynthesis. J. Am. Chem. Soc. 140: 6044–6048.

Müller D, Krick A, Kehraus S, Mehner C, Hart M, Küpper FC, Saxena K, Prinz H, Schwalbe H, Janning P, Waldmann H and König GM (2006). Brunsvicamides A–C: sponge-related cyanobacterial peptides with *Mycobacterium tuberculosis* protein tyrosine phosphatase inhibitory activity. J. Med. Chem. 49: 4871–4878.

Newman DJ and Cragg GM (2016). Natural products as sources of new drugs from 1981 to 2014. J. Nat. Prod. 79: 629–661.

O'Connor PM, O'Shea EF, Cotter PD, Hill C and Ross RP (2018). The potency of the broad spectrum bacteriocin, bactofencin A, against staphylococci is highly dependent on primary structure, N-terminal charge and disulphide formation. Sci. Rep. 8: 11833.

Okada M, Sugita T, Akita K, Nakashima Y, Tian T, Li C, Mori T and Abe I (2016). Stereospecific prenylation of tryptophan by a cyanobacterial post-translational modification enzyme. Org. Biomol. Chem. 14: 9639–9644.

Oman TJ and van der Donk WA (2010). Follow the leader: the use of leader peptides to guide natural product biosynthesis. Nat. Chem. Biol. 6: 9–18.

Oppegård C, Fimland G, Anonsen JH and Nissen-Meyer J (2015). The pediocin PA-1 accessory protein ensures correct disulfide bond formation in the antimicrobial peptide pediocin PA-1. Biochemistry. 54: 2967–2974.

Ortega MA and van der Donk WA (2016). New insights into the biosynthetic logic of ribosomally synthesized and post-translationally modified peptide natural products. Cell Chem. Biol. 23: 31–44.

Oueis E, Adamson C, Mann G, Ludewig H, Redpath P, Migaud M, Westwood NJ and Naismith JH (2015). Derivatisable cyanobactin analogues: A semisynthetic approach. ChemBioChem 16: 2646–2650.

Oueis E, Jaspars M, Westwood NJ and Naismith JH (2016). Enzymatic macrocyclization of 1,2,3-triazole peptide mimetics. Angew. Chem. Int. Ed. Engl. 55: 5842–5845.

Oueis E, Stevenson H, Jaspars M, Westwood NJ and Naismith JH (2017). Bypassing the proline/thiazoline requirement of the macrocyclase PatG. Chem. Commun. (Cambridge, U. K.) 53: 12274–12277.

Parajuli A, Kwak DH, Dalponte L, Leikoski N, Galica T, Umeobika U, Trembleau L, Bent A, Sivonen K, Wahlsten M, Wang H, Rizzi E, De Bellis G, Naismith J, Jaspars M, Liu X, Houssen W and Fewer DP (2016). A unique tryptophan C-prenyltransferase from the kawaguchipeptin biosynthetic pathway. Angew. Chem. Int. Ed. 55: 3596–3599.

Phyo MY, Ding CYG, Goh HC, Goh JX, Ong JFM, Chan SH, Yung PYM, Candra H and Tan LT (2019). Trikoramide A, a prenylated cyanobactin from the marine cyanobacterium *Symploca hydnoides*. J. Nat. Prod. 82: 3482–3488.

Poorinmohammad N, Bagheban-Shemirani R and Hamedi J (2019). Genome mining for ribosomally synthesised and post-translationally modified peptides (RiPPs) reveals undiscovered bioactive potentials of actinobacteria. Antonie van Leeuwenhoek. 112: 1–23.

Rao MPN, Li Y-Q, Zhang H, Dong Z-Y, Dhulappa A, Xiao M and Li W-J (2020). *Amycolatopsis alkalitolerans* sp. nov., isolated from *Gastrodia elata* Blume. J Antibiot. 73: 35–39.

Reardon-Robinson ME and Ton-That H (2016). Disulfide bond-forming pathways in Gram-positive bacteria. J. Bacteriol, 198: 746-754.

Santos-Aberturas J, Chandra G, Frattaruolo L, Lacret R, Pham TH, Vior NM, Eyles TH and Truman, AW (2019). Uncovering the unexplored diversity of thioamidated ribosomal peptides in actinobacteria using the RiPPER genome mining tool. Nucleic Acids Res. 47: 4624–4637.

Sardar D, Pierce E, McIntosh JA and Schmidt EW (2015). Recognition sequences and substrate evolution in cyanobactin biosynthesis. ACS Synth. Biol. 4: 167–176.

Sardar D, Hao Y, Lin Z, Morita M, Nair SK and Schmidt EW (2017). Enzymatic N- and C-Protection in cyanobactin RiPP natural products. J. Am. Chem. Soc. 139: 2884–2887.

Scharf DH, Groll M, Habel A, Heinekamp T, Hertweck C, Brakhage AA and Huber EM (2014). Flavoenzyme-catalyzed formation of disulfide bonds in natural products. Angew. Chem. Int. Ed. Engl. 53: 2221– 2224.

Scheffler RJ, Colmer S, Tynan H, Demain AL and Gullo VP (2013). Antimicrobials, drug discovery, and genome mining. Appl. Microbiol. Biotechnol. 97: 969–978.

Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG and Ravel J (2005). Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. Proc. Natl. Acad. Sci. USA. 102: 7315–7320.

Schmidt EW and Donia MS (2009). Cyanobactin ribosomally synthesized peptides — a case of deep metagenome mining Methods. Enzymol, 458: 575–596.

Senter PD and Sievers EL (2012). The discovery and development of brentuximab vedotin for use in relapsed Hodgkin lymphoma and systemic anaplastic large cell lymphoma. Nat. Biotechnol. 30: 631–637.

Shafee TMA, Lay FT, Phan TK, Anderson MA and Hulett MD (2017). Convergent evolution of defensin sequence, structure and function. Cell. Mol. Life. Sci. 74: 663–682.

Singh S, Bhushan KN and Banerjee UC (2005). Bioactive compounds from cyanobacteria and microalgae: An overview. Crit. Rev. Biotechnol. 25: 73–95.

Singh PK, Chittpurna A, Sharma V, Patil PB and Korpole S (2012). Identification, purification and characterization of laterosporulin, a novel bacteriocin produced by *Brevibacillus* sp. strain GI-9. PLoS One 7: e31498.

Sings HL and Rinehart KL (1996). Compounds produced from potential tunicate-blue-green algal symbiosis: A review. J. Ind. Microbiol. Biotechnol. 17: 385–396.

Sikandar A and Koehnke J (2019). The role of protein–protein interactions in the biosynthesis of ribosomally synthesized and post-translationally modified peptides. Nat. Prod. Rep. 36: 1576–1588.

Shioiri T, Hamada Y, Kato S, Shibata M and Kondo Y (1987). Cytotoxic activity of cyclic peptides of marine origin and their derivates: Importance of oxazoline functions. Biochem. Pharmacol. 3: 4181–4185.

Sivonen K, Leikoski N, Fewer DP and Jokela J (2010). Cyanobactins–ribosomal cyclic peptides produced by cyanobacteria. Appl. Microbiol. Biotechnol. 86: 1213–1225.

Skinnider MA, Johnston CW, Edgar RE, Dejong CA, Merwin NJ, Rees PN and Magarvey NA (2016). Genomic charting of ribosomally synthesized natural product chemical space facilitates targeted mining Proc. Natl. Acad. Sci. USA. 113: 6343–6351.

Sugrue I, O'Connor PM, Hill C, Stanton C and Ross RP (2020). *Actinomyces* produces defensin-like bacteriocins (actifensins) with a highly degenerate structure and broad antimicrobial activity. J. Bacteriol. 202: e00529–19.

Tan LT (2007). Bioactive natural products from marine cyanobacteria for drug discovery. Phytochemistry. 68: 954–979.

Tan S, Moore G and Nodwell, J (2019). Put a bow on it: Knotted antibiotics take center stage. Antibiotics. 8: 117–134.

Tianero MD, Donia MS, Young TS, Schultz PG and Schmidt EW (2012). Ribosomal route to small-molecule diversity. J. Am. Chem. Soc. 134: 418–425.

Tianero MD, Pierce E, Sardar D, Raghuraman S, McIntosh JA, Heemstra JR, Schonrock Z, Covington BC, Cox JE, Bachmann BO, Olivera BM, Ruffner DE and Schmidt EW (2016). Metabolic model for diversity generating metabolism. Proc. Natl. Acad. Sci. USA. 113: 1772–1777.

Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai H-C, Zakai UI and Mitchell DA (2017). A new genome-mining tool redefines the lasso peptide biosynthetic landscape. Nat. Chem. Biol. 13: 470.

Tossi A, Sandri L and Giangaspero A (2000). Amphipathic, alpha-helical antimicrobial peptides. Biopolymers. 55: 4–30.

Velásquez JE and van der Donk WA (2011). Genome mining for ribosomally synthesized natural products. Curr. Opin. Chem. Biol. 15: 11–21.

Vestola J, Shishido TK, Jokela J, Fewer DP, Aitio O, Perm iP, Wahlsten M, Wang H, Rouhiainen L and Sivonen K (2014). Hassallidins, antifungal glycolipopeptides, are widespread among cyanobacteria and are the end-product of a nonribosomal pathway. Proc. Natl. Acad. Sci. USA. 111: 1909–1917.

Walsh CT, Garneau-Tsodikova S and Gatto GJ Jr (2005). Protein posttranslational modifications: the chemistry of proteome diversifications. Angew. Chem. Int. ed. 44: 7342–7372.

Walsh CT and Tang Y (2017). Natural product biosynthesis: chemical logic and enzymatic machinery. Croydon: Royal Society of Chemistry (RSC) Publishing, UK.

Wang G (2014). Post-translational modifications of natural antimicrobial peptides and strategies for peptide engineering. Curr. Biotechnol. 1: 72–79.

Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach M, Müller R, Wohlleben W, Breitling R, Takano E and Medema MH (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. 43: 237–243.

Winter JM, Behnken S and Hertweck C (2011). Genomics-inspired discovery of natural products. Curr. Opin. Chem. Biol. 15: 22–31.

Wipf P, Fritch PC, Geib SJ and Sefler AM (1998). Conformational studies and structure-activity analysis of lissoclinamide 7 and related cyclopeptide alkaloids. J. Am. Chem. Soc. 120: 4105–4112.

Wong CP, Awakawa T, Nakashima Y, Mori T, Zhu Q, Liu X and Abe I (2018). Two distinct substrate binding modes for the normal and reverse prenylation of hapalindoles by the prenyltransferase AmbP3. Angew. Chem. Int. Ed. 57: 560–563.

Zasloff M (1987). Magainins, a class of antimicrobial peptides from Xenopus skin: Isolation, characterization of two active forms, and partial cDNA sequence of a precursor. Proc. Natl. Acad. Sci. USA. 84: 5449–5453.

Zasloff M (2002). Antimicrobial peptides of multicellular organisms. Nature 415: 389–395.

Zerikly M and Challis GL (2009). Strategies for the discovery of new natural products by genome mining. ChemBioChem 10: 625−633.

Ziemert N, Alanjary M and Weber T (2016). The evolution of genome mining in microbes - a review. Nat. Prod. Rep. 33: 988–1005.