Expression, characterisation, and engineering of cyanobacterial natural product biosynthetic pathways

Author:
Soeriyadi, Angela

Publication Date:
2019

DOI:
https://doi.org/10.26190/unsworks/2275

Expression, characterisation, and engineering of cyanobacterial natural product biosynthetic pathways

A dissertation submitted by

Angela Hertania Soeriyadi

In fulfilment of the requirements for the award of

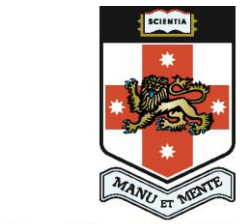
Doctor of Philosophy

at the

School of Biotechnology and Biomolecular Sciences

The University of New South Wales

Sydney, Australia



UNSW SYDNEY

October 2019

Supervisors

Prof. Brett A. Neilan

School of Environmental Life Sciences

The University of Newcastle, and

School of Biotechnology and Biomolecular Sciences

The University of New South Wales

Australia

A/Prof. Wallace Bridge

School of Biotechnology and Biomolecular Sciences

The University of New South Wales

Sydney, Australia

Co-supervisor

Dr. Sarah E. Ongley

School of Environmental Life Sciences

The University of Newcastle, and

School of Biotechnology and Biomolecular Sciences

The University of New South Wales

Australia

Surname/Family Name	:	**Soeriyadi**
Given Name/s	:	**Angela Hertania**
Abbreviation for degree as give in the University calendar	:	**PhD**
Faculty	:	**Science**
School	:	**School of Biotechnology and Biomolecular Sciences**
Thesis Title	:	**Expression, characterisation, and engineering of cyanobacterial natural product biosynthetic pathways**

Abstract 350 words maximum:

Microorganisms are a valuable source of natural products with medically and industrially relevant activities. Cyanobacteria are one of the most chemically diverse microbial phyla, but have been largely underexplored due to their slow growth and intractability to genetic engineering. In recent years, genomic and metagenomic investigations of the aquatic environment have uncovered the untapped diversity of cyanobacterial nonribosomal peptide and polyketide biosynthetic gene clusters. This dissertation describes the application of emerging synthetic biology techniques using *Escherichia coli* as a heterologous host, focussing on translating the bioactive potential of cyanobacteria into industrial applications, while simultaneously characterising and tailoring this biochemical capacity.

E. coli GB05-MtaA was previously shown to be a suitable host for the relatively simple cyanobacterial nonribosomal peptide synthetase (NRPS) pathway lyngbyatoxin (LTX). A synthetic biology approach for characterising and expanding lyngbyatoxin chemical diversity was explored. This thesis reports an *in vitro* investigation of wild-type LtxA NRPS activity that unravelled the multispecificity of the first adenylation domain to L-Val related amino acids, which correlates with the formation of novel lyngbyatoxin analogues *in vivo*. Efficient site-directed mutagenesis of the adenylation domain to modulate substrate specificities was performed through Red/ET-based recombineering, resulting in a library of mutated LTX pathways. Investigation of *in vitro* and *in vivo* pathway activity revealed the complexity of LTX biosynthesis, dictated by tailoring enzymes.

The benefits of using this approach to probe LTX biosynthesis motivated the use of a similar application for the directed production of neosaxitoxin (neoSTX). NeoSTX is a paralytic shellfish toxin (PST) which has medically-relevant activities. The polyketide synthase (PKS)-like enzyme SxtA, which initiates PST biosynthesis, was expressed in *E. coli* to assess the suitability of this host. Efficient production of PST intermediates by heterologously expressed SxtA paved the way for the expression of an engineered neoSTX biosynthetic pathway in *E. coli*. Several variants of *E. coli* strains were successfully constructed to produce neoSTX. This synthetic biology process, with an *E. coli* expression system, is a feasible strategy to facilitate the commercial production of neoSTX.

COPYRIGHT STATEMENT

Signed …….....

Date …….....**.....**

AUTHENTICITY STATEMENT

‘I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.’

Signed …….....

Date …….....**.....**

ORIGINALITY STATEMENT

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.’

Signed …….....

Date …….....**.....**

INCLUSION OF PUBLICATIONS STATEMENT

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The student contributed greater than 50% of the content in the publication and is the "primary author", ie. the student was responsible primarily for the planning, execution and preparation of the work for publication
- The student has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not.

☐ *This thesis contains no publications, either published or submitted for publication*

☐ *Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement*

☒ *This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below*

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Name	Signature	Date (dd/mm/yy)
Angela Hertania Soeriyadi		10/10/2019

Postgraduate Coordinator's Declaration

I declare that:

- the information below is accurate
- where listed publication(s) have been used in lieu of Chapter(s), their use complies with the Thesis Examination Procedure
- the minimum requirements for the format of the thesis have been met.

PGC's Name	PGC's Signature	Date (dd/mm/yy)
Paul Waters		10/10/2019

For each publication incorporated into the thesis in lieu of a Chapter, provide all of the requested details and signatures required

Details of publication #1:
Full title: Chapter 17 Bioprospecting and Insights into the Biosynthesis of Natural Products from Marine Microalgae
Authors: Angela H. Soeriyadi, Sarah E. Ongley, Caitlin S. Romanis, and Brett A. Neilan
Journal or book name: Blue Biotechnology: Production and Use of Marine Molecules (eds S. La Barre and S. S. Bates)
Volume/page numbers: Volume 2/ pp.553-581
Date accepted/ published: December 2018

Status	*Published*		*Accepted and In press*	*In progress (submitted)*	
		x			

The Candidate's Contribution to the Work
The candidate wrote the majority of the book chapter including section 17.2 and 17.3, combining pieces of writing from different authors, as well as responsible for the liaison with the revision and proof of the book chapter. Caitlin S. Romanis wrote the first draft of section 17.1.1 and parts of Table 17.1, and section 17.2.2 was written by Sarah E. Ongley.

Location of the work in the thesis and/or how the work is incorporated in the thesis:
Chapter 1. Introduction: The biosynthesis of cyanobacterial specialised metabolites

Primary Supervisor's Declaration
I declare that:
- the information above is accurate
- this has been discussed with the PGC and it is agreed that this publication can be included in this thesis in lieu of a Chapter
- All of the co-authors of the publication have reviewed the above information and have agreed to its veracity by signing a 'Co-Author Authorisation' form.

Supervisor's name	*Supervisor's signature*	*Date (dd/mm/yy)*
Wallace Bridge		10/10/2019

Abstract

Microorganisms are a valuable source of natural products with medically and industrially relevant activities. Cyanobacteria are one of the most chemically diverse microbial phyla, but have been largely underexplored due to their slow growth and intractability to genetic engineering. In recent years, genomic and metagenomic investigations of the aquatic environment have uncovered the untapped diversity of cyanobacterial nonribosomal peptide and polyketide biosynthetic gene clusters. This dissertation describes the application of emerging synthetic biology techniques using *Escherichia coli* as a heterologous host, focussing on translating the bioactive potential of cyanobacteria into industrial applications, while simultaneously characterising and tailoring this biochemical capacity.

E. coli GB05-MtaA was previously shown to be a suitable host for the relatively simple cyanobacterial nonribosomal peptide synthetase (NRPS) pathway lyngbyatoxin (LTX). A synthetic biology approach for characterising and expanding lyngbyatoxin chemical diversity was explored. This thesis reports an *in vitro* investigation of wild-type LtxA NRPS activity that unravelled the multispecificity of the first adenylation domain to L-Val related amino acids, which correlates with the formation of novel lyngbyatoxin analogues *in vivo*. Efficient site-directed mutagenesis of the adenylation domain to modulate substrate specificities was performed through Red/ET-based recombineering, resulting in a library of mutated LTX pathways. Investigation of *in vitro* and *in vivo* pathway activity revealed the complexity of LTX biosynthesis, dictated by tailoring enzymes.

The benefits of using this approach to probe LTX biosynthesis motivated the use of a similar application for the directed production of neosaxitoxin (neoSTX). NeoSTX is a paralytic shellfish toxin (PST) which has medically-relevant activities. The polyketide synthase (PKS)-like enzyme SxtA, which initiates PST biosynthesis, was expressed in *E. coli* to assess the suitability of this host. Efficient production of PST intermediates by heterologously expressed SxtA paved the way for the expression of an engineered neoSTX biosynthetic pathway in *E. coli*. Several variants of *E. coli* strains were successfully constructed to produce neoSTX. This

synthetic biology process, with an *E. coli* expression system, is a feasible strategy to facilitate the commercial production of neoSTX.

Acknowledgements

So many people have been involved in making this dissertation possible. For that, I am very grateful, and I would like to express my deepest gratitude.

Firstly, I'd like to thank my supervisor, Prof. Brett Neilan for providing me with the opportunity to join the extremely supportive (and very groovy) BGGM group at UNSW and work with challenging but rewarding projects. It has been such a journey of learning, both personally and scientifically. I'd also like to thank my supervisor A/Prof. Wallace Bridge for taking me on as a student and guiding me through the final stages of this remarkable project.

A special thank you goes to my co-supervisor, Dr. Sarah Ongley, whom I will say again, is the best mentor I could ever hope for. You have been my mentor and my buddy since I first joined this lab. I'm so grateful to have you as my co-supervisor. Thank you for introducing me to these marvellous cloning techniques and for all your help, guidance, and hugs especially during my thesis writing process.

Thank you also to Dr. Ralf Kellmann who gave me the opportunity to work with the almighty neosaxitoxin project. It's definitely been an exciting roller coaster project and reminded me why scientific research has a special place in my heart. Thank you for your guidance throughout my PhD. Thank you to Mark Van Asten and Frank Hoover for your helpful engagement in the neosaxitoxin project. A special thank you also goes to Prof. Elke Dittmann who allowed me to visit her lab to perform the experiments for the lyngbyatoxin project. I sincerely appreciate how very welcoming you were, and I enjoyed my time in Germany so much! I also would like to thank Prof. Adrian Lee for establishing the travel scholarship that enabled me to travel to both Norway and Germany to visit these two labs.

Thank you to Dr. Russell Pickford for the assistance with the LC-MS analysis and for teaching me the data processing. Thanks A/Prof. Christopher P. Marquis and Dr. Helene Lebhar for their support in performing the recombinant protein expression and purification work, and also for their help throughout my candidature. Thank you to BABS PGC, Dr. Paul Waters for the guidance and help towards finishing my candidature.

To the people who helped me during my research visit at the University of Bergen, Prof. Nils-Kåre Birkeland, Sudeep, Bagmi, Kareem, Fedor, Antonio, Helge, Birte and the several others whom I haven't named, thank you so much for your assistance and kindness. Thank you for my fellow inhabitants of Allégaten 19 residence for becoming my 'family' in Norway. I cherished all those midnight hikes! Thank you to the members of Prof. Elke Dittmann's lab, especially Dr. Jan-Christoph Kehr and Sabine Meyer who helped me with the assay.

I would also like to thank all the past and present members of Neilan lab of Microbial and Molecular Diversity (BGGM) family, especially Dr. Rabia Mazmouz and Dr. Jason Woodhouse for giving me scientific input and also spiritual support, Dr. Leanne Pearson for your assistance in reading my thesis, Rocky, Alescia, Caitlin, Angie, Jesse, and many more people who supported me throughout my candidature.

Special warm thanks to my fellow 'The lost children', Bakir, Carolin, JP, and Tian. You guys were always there and I'm so privileged to have you as my colleagues and my dear friends. Everything is awesome when you are part of a team! Yay!

Special thanks to my parents, my brother, and my sister for all the years of support and for picking me up from the lab on all those late nights. Thank you Rob for helping me proof reads my drafts. Thank you also to mister Paweł for all your support.

Lastly, thank you to all the cyanobacteria and *E. coli*! And to Andy from Headspace for keeping my head sane-ish.

I don't know what else to say other than thank you to God and the whole universe for everything that is amazing and all the inexplicable things. It has been a super fun and crazy ride and I have learned a lot during this journey… and with that note, Angela's out! ☺

Minima maxima sunt

- The smallest things are most important

Table of Contents

Abbreviations

A-domain	NRPS adenylation domain
ACP	acyl carrier protein
Ala	alanine
Amp	ampicillin
AONS	8-amino-7-oxononanoate synthase
AT	acyltransferase
CoA	coenzyme A
C-toxin	paralytic shellfish toxin analogues sulphated at the N- and O-position
dc-toxin	decarbamoyl paralytic shellfish toxin
demetInt-A’	demethylated intermediate-A’
DH	dehydratase
DNA	deoxyribonucleic acid
ELISA	enzyme-linked immunoadsorbent assay
Genta	gentamicin
Gly	glycine
GTX	Gonyautoxins; mono-sulphated PSTs
HPLC	high-performance liquid chromatography
HPLC-ESI-HRMS	HPLC-coupled electrospray ionisation high-resolution mass spectrometry
Ile	isoleucine
ILI	indolactam-I
ILL	indolactam-L
ILV	indolactam-V
Int-A’	intermediate-A’
Km	kanamycin
KR	ketoreductase
LB	lysogeny broth
LCHR	linear DNA plus circular DNA homologous recombination

Leu	leucine
LTX	lyngbyatoxin
LtxA-A1	first A-domain of the lyngbyatoxin NRPS responsible for incorporating L-valine
LWT	paralytic shellfish toxin analogues produced by *L. wollei*
MATE	multi-drug and toxic compound extrusion protein
Me-LTX (Ile)	methyl-lyngbyatoxin (isoleucine)
Me-LTX (Leu)	methyl-lyngbyatoxin (leucine)
MDA	multiple displacement amplification
MLP	MbtH-like proteins
MNBA	mouse neuroblastoma bioassay
MS/MS	tandem mass spectrometry
MT	methyltransferase
MtaA	broad-range substrate specificity PPTase from the myxobacterium *Stigmatella aurantiaca* DW4/3-1
NADPH	nicotinamide adenine dinucleotide phosphate
NeoSTX	neosaxitoxin
NMIT	N-methyl-L-isoleucyl-L-tryptophanol
NMLT	N-methyl-L-leucyl-L-tryptophanol
NMR	nuclear magnetic resonance
NMVT	N-methyl-L-valyl-L-tryptophanol
NRP	nonribosomal peptide
NRPS	nonribosomal peptide synthetase
Orn	ornithine
PAPS	3’-phosphoadenosine, 5’- phosphosulfate
PCP	peptidyl-carrier-protein
PCR	polymerase chain reaction
Phe	phenylalanine
Phg	phenylglycine

PKC	protein kinase C
PK	polyketide
PKS	polyketide synthase
PPTase	4’-phosphopantetheinyl transferase
PST	paralytic shellfish toxin
Ptet	tetracycline-inducible promoter
RiPPS	ribosomally synthesised and postranslationally modified peptides
RNA	ribonucleic acid
mRNA	messenger RNA
SAM	*S*-adenosylmethionine
SDM	site-directed mutagenesis
SDS-PAGE	sodium dodecylsulphate-polyacrylamide gel electropohoresis
SPE	solid phase extraction
STX	saxitoxin
TAR	transformation-associated recombination
TB	terrific broth
Trp	tryptophan
Tyr	tyrosine
Val	valine
WT	wild-type

List of Publications and Proceedings

Publications

Cullen, A., Pearson, L. A., Mazmouz, R., Liu, T., **Soeriyadi, A. H**., Ongley, S. E., and Neilan, B. A. (2018) Heterologous expression and biochemical characterisation of cyanotoxin biosynthesis pathways, 36, p. 1117-1136, *Natural product reports*. doi:10.1039/C8NP00063H.

Soeriyadi, A. H., Ongley, S. E., Romanis, C. S. and Neilan, B. A. (2018). Bioprospecting and Insights into the Biosynthesis of Natural Products from Marine Microalgae In Blue Biotechnology (eds S. La Barre and S. S. Bates), p.553-581 doi:10.1002/9783527801718.ch17.

Publications in preparation

Soeriyadi, A. H., Mazmouz, Pickford, R., Al-Sinawi, B., Kellmann, R., and Neilan, B. A. (2019). Expression of the PKS-related enzyme SxtA leads to production of saxitoxin intermediates in *E. coli*. (*In preparation*)

Soeriyadi, A. H., Ongley, S. E., Kehr, J.C., Pickford, R., Dittmann, E., and Neilan, B. A. (2019). Tailoring enzyme stringency masks the multispecificity of a lyngbyatoxin nonribosomal peptide synthetase. (*In preparation*)

Presentations

2018 Poster presentation, Gordon Research Conference (GRC) and Gordon Research Seminar (GRS) on Natural Products and Bioactive Compounds, New Hampshire, USA.

2017 Oral and poster presentation, 2017 UNSW-Indonesia Research Roadshow, Jakarta, Indonesia.

2017 Lightning oral presentation and poster presentation, Directing Biosynthesis V Conference, Warwick, UK.

2016 Poster presentation, John Innes/Rudjer Bošković Summer Schools in Applied Molecular Microbiology, Dubrovnik, Croatia.

2015 Oral and poster presentation, UNSW Science Research Competition, Sydney, Australia.

2015 Poster presentation, Joint Academic Microbiology Seminars (JAMS) 4th Annual Symposium and Dinner, Sydney, Australia.

2014 Oral presentation, The 2014 BABS Research Symposium, Sydney, Australia

List of figures

List of tables

Chapter 1. Introduction: The biosynthesis of cyanobacterial specialised metabolites

Preamble

Cyanobacteria (also known as blue-green algae) are photosynthetic prokaryotes ubiquitously found in diverse environmental niches, including marine, terrestrial and freshwater systems. The extreme temperature, pH, and high salinity, as well as the presence of other organisms in cyanobacterial habitat, drive the evolutionary production of various specialised metabolites. Many of these metabolites have bioactivities applicable for the pharmaceutical and agricultural industry. However, cyanobacteria also offer challenges as unculturability and genetic intractability limit commercial application. This thesis focuses on the utilisation of a synthetic biology approach to gain access to the cyanobacterial specialised metabolites via heterologous expression in the model organism, *Escherichia coli*. In addition, biosynthetic pathway modulation through control of gene expression and pathway engineering were explored.

The first two sections of this chapter (section 1.1 and 1.2) provide an overview of selected cyanobacterial biosynthetic pathways, highlighting the distinct features in the enzymology and approaches that have been taken to characterise each pathway. The specific context and aims of this dissertation are then outlined in section 1.3.

The first part of the literature review (section 1.1) presents an overview of specialised metabolite biosynthesis pathways isolated from marine microalgae. The review was published as a chapter in the book entitled "Blue Biotechnology: Production and Use of Marine Molecules". The University of New South Wales requires identical content of a publication in the case of inclusion of a publication in lieu of a thesis chapter. As the book chapter is a collaborative work involving multiple authors, several sections of the chapter are not directly relevant to this thesis. These include the section 1.1.1.1, and parts of Table 1.1, written by Caitlin S. Romanis, which give an overview of natural products found in other microalgae, and section 1.1.2.2, written by Sarah E. Ongley discussing ribosomally synthesised natural products.

The majority of this thesis is dedicated to the expression and characterisation of the saxitoxin biosynthetic pathway from a freshwater cyanobacterium. Section 1.2 provides relevant background information surrounding the bioactivity, discovery, and an overview of the saxitoxin biosynthetic pathway from a genetic and biochemical perspective.

1.1 Bioprospecting and insights into the biosynthesis of natural products from marine microalgae

Authors:

Angela H. Soeriyadi[1], Sarah E. Ongley[1,2], Caitlin S. Romanis[1,2], and Brett A. Neilan[1,2]

[1]The University of New South Wales, School of Biotechnology and Biomolecular Sciences, Kensington, Sydney, NSW 2052, Australia

[2]The University of Newcastle, School of Environmental and Life Sciences, Newcastle, NSW 2308, Australia

This section of the introduction has been published as:

Soeriyadi, A. H., Ongley, S. E., Romanis, C. S. and Neilan, B. A. (2018). Bioprospecting and Insights into the Biosynthesis of Natural Products from Marine Microalgae. In Blue Biotechnology (eds S. La Barre and S. S. Bates), p. 553-581. doi:10.1002/9783527801718.ch17.

Extent to which the publication is your own:

I wrote the majority of the book chapter including section 1.1.2 and 1.1.3 and was responsible for the liaison of the revisions and final proofs. Caitlin S. Romanis wrote the first draft of section 1.1.1.1 and parts of Table 1.1, and section 1.1.2.2 was written by Sarah E. Ongley.

Comments/deviation from the publication:

Remarks of "See chapter XX" which refer to other chapters within the published book were removed. The numbering and the referencing style have been changed to align with the thesis format. The Acknowledgement section has been removed.

Abstract

The production of metabolites eliciting biological activity by marine microalgae is not a recent discovery. *In situ*, these specialized metabolites may play roles in regulating physiological changes during the life cycle of the organism or the maintenance of host–symbiont associations, but the precise role of most metabolites remains unknown. Technological advances over the years have enabled the study of the biosynthesis and regulation of these remarkable specialized metabolites, as well as facilitating new methods for extracting the full metabolite potential from marine microalgae. In this chapter, we describe unique aspects of specialized metabolite biosynthesis in marine cyanobacteria and modern methods for the discovery of novel metabolites.

1.1.1 Introduction

Marine microalgae represent a novel source of natural products that natively possess a range of biological activities. The immense structural diversity of natural products provides chemical scaffolds or inspiration for chemical synthesis of novel pharmaceuticals. The marine environment promotes the evolution of unique adaptative strategies by marine microalgae, adaptations to high hydrostatic pressure and high salinity, temperature and pH translate into novel natural products (Table 1.1) (Jebbar *et al.*, 2015, Imada, 2013).

1.1.1.1 Dinoflagellates

Dinoflagellates are a large group of unicellular marine microorganisms, some of which form symbiotic relationships with sponges and coelenterates. In coral reef ecosystems, dinoflagellates from the genus *Symbiodinium* are the predominant form of symbiotic relationships and occur between the microalgae and diverse group of hosts, for example, cnidarian, molluscan, and poriferan (Stat *et al.*, 2008). These relationships are typically associated with the production of harmful toxins, sometimes as an anti-predation manoeuvre. Seafood poisoning syndromes related to dinoflagellates include neurotoxic shellfish poisoning, paralytic shellfish poisoning, diarrheic shellfish poisoning, and ciguatera fish poisoning (Wang, 2008, Hernandez-Becerril *et al.*, 2007).

Diversity within the phyla is generated by the repeated replacement of their ancestral plastid during a tertiary endosymbiotic event, the process by which a heterotrophic eukaryote incorporated a secondary plastid-containing alga or where a plastid was lost and replaced through sequential secondary endosymbiosis of a primary plastid-containing alga (Yoon *et al.*, 2005). As such, the dinoflagellate group comprises plastids from four heterogenous sources: *Chlorophyta, Crytophyta, Stramenopiles* and *Haptophyta* (Heimann & Huerlimann, 2015, Gabrielsen *et al.*, 2011). Dinoflagellate blooms can form at any given location for numerous reasons (e.g., stoichiometry of nutrients such as N and P, concentration by currents, and water column structure), although it is almost impossible to predict the nature, timing, or species diversity of a given bloom (Anderson & Rengefors, 2006). There is an association between bloom formation and the ratio of nutrients in coastal waters. Nitrogenous runoff from industrial or agricultural sources has been correlated with the sudden formation of blooms and is regarded as causative agents (Camacho *et al.*, 2007).

Dinoflagellates are known to produce a number of bioactive specialised metabolites, for example, several analogues of palytoxin, a lethal marine polyalcohol toxin produced by the dinoflagellate genus *Ostreopsis* (Patocka *et al.*, 2015, Ramos & Vasconcelos, 2010). Dinoflagellates are also a principal producer of polyether ladders, for example, the lipid soluble neurotoxins brevetoxin and maitotoxin. Neurotoxicity by these toxins is mediated by precise interactions with the ion channels that mediate neurotransmission (Wang, 2008). Thirty-four macrolides (designated as amphidinolides) have been isolated from the dinoflagellate *Amphidinium* sp., which is a symbiont of the marine flatworm *Amphiscolops* sp., and they possess potent cytotoxic activities (Kobayashi, 2008). Marine dinoflagellates represent an untapped source of bioactive polyketides and represent a relative gold-mine for natural product prospecting.

While the genetic basis for the biosynthesis of natural products in dinoflagellates remains elusive, it is speculated that many dinoflagellate toxins are predominantly of polyketide origin, being synthesised through sequential Claisen condensations of carboxylic units. Considering the nature of polyketides derived from dinoflagellates, it is assumed this

mechanism of biosynthesis is a type I modular polyketide synthase (PKS) system, which is composed of multiple enzymatic domains located on a singular polypeptide chain (Rein & Snyder, 2006, Kellmann *et al.*, 2010).

1.1.1.2 Cyanobacteria

Cyanobacteria are one of the oldest phyla on the planet, existing in the fossil record for ~3.5 billion years, and are responsible for the earliest known oxygen generation (Christaki, 2014, Manivasagan & Kim, 2015). Marine cyanobacteria are distributed relatively ubiquitously, featured in both tropical climates and Artic conditions of the Polar Regions. Cyanobacteria are exceedingly efficient at nutrient uptake and storage, a feature that is enhanced by rapid vertical migration through the water column using alterations in cell buoyancy possessed by many planktonic species (Manivasagan & Kim, 2015). Many species of cyanobacteria also form endosymbiotic relationships with marine invertebrates, for example, cyanobacterial inhabitation of marine sponges provides carbon and nitrogen, while the sponge provides protection. Furthermore, sponge habitation by cyanobacteria involves peptide production that could help to protect the host sponge, for example, the nonribosomal cyclic peptide leucamide A, which is an antitumor agent isolated from the *Leucetta microraphis* sponge (Thomas *et al.*, 2010). Peptides such as leucamide A are a prime example of the secondary metabolites produced by cyanobacteria. The majority of these products are nitrogen-containing compounds synthesised by nonribosomal peptide synthetases (NRPSs) or hybrid nonribosomal peptide synthetases and polyketide synthesis modules (NRPS/PKS) (Tan, 2007). Derivatives of this class of metabolites (albeit not cyanobacterial products) include the antibiotic vancomycin, the immunosuppressant cyclosporine, and the anticancer agent bleomycin (Tan, 2007, Chang *et al.*, 2002). Specialised metabolites are often subsequently used as scaffolds for the development of more pharmaceutically efficacious synthetic analogues such as curacin A (Wipf *et al.*, 2002b) and dolastatin 10 (Taori *et al.*, 2009, Verdier-Pinard *et al.*, 1998a).

Table 1.1 Examples of natural products from marine microalgae.

Natural Product	Organism	Activity
Cyanobacteria		
Cyanovirin-N (CV-N)	*Nostoc ellipsosporum*	Virucidal agent: inhibits fusion of viral and host cell membranes
Microcolin A	*Lyngbya majuscula*[a)]	Immunosuppressive: inhibits the production of IL-2 and IL-2 receptors
Anatoxin-A	*Anabaena flo-aquae*	Toxin: irreversible inhibitor of acetylcholinesterase
Curacin A	*Lyngbya majuscula* [a)]	Antimitotic agent: inhibits microtubule assembly and prevents the binding of colchicine to tubulin
Jamaicamides	*Lyngbya majuscula* [a)]	Toxic: block sodium channels
Tolytoxin	*Scytonema ocellatum*	Cytostatic: depolymerises actin *in vivo* disrupting cell division in eukaryotic organisms
Dinoflagellates		
Symbiodinolide	*Symbiodinium sp.*	Demonstrated voltage-dependent N-type Ca^{2+} channel opening capabilities
Durinskiol	*Durinskia sp.*	Activity not yet described
Green Microalgae		
Dermochlorella DP	*Chlorella* sp.	Cosmetics: anti-elastase agent
Astaxanthin	*Haematococcus pluvialis*	Varied activities: immunomodulation, anticancer, antioxidant, and anti-inflammatory
Haptophytes		
Nonacosadienes	*Emiliania huxleyi*	Activity not yet described
Red Microalgae		
Sulfated polysaccharides	*Porphyridium* sp.	Antiviral: inhibits viral penetration of host cell

[a)] This species has been reclassified as *Moorea producens*.

1.1.2 Biosynthesis of natural products from cyanobacteria

Cyanobacteria from marine environments have been shown to be a rich source of chemically diverse and biologically significant metabolites, with many being exploited for pharmaceutical interests (Montaser & Luesch, 2011, Martins *et al.*, 2014). Several biosynthetic clusters from marine cyanobacteria have been characterised. Characterisation of biosynthetic clusters from

marine environment is hindered by the difficulty in growing the cells in the laboratory, as well as the abundance of microbial and invertebrate symbionts of cyanobacteria (Simmons *et al.*, 2008). Recently, with advances in DNA sequencing technologies and analytical chemistry techniques, more biosynthetic pathways of both ribosomal and nonribosomal derived natural products have been uncovered. Herein, we describe unique aspects of the biosynthesis of a few select ribosomal and nonribosomal (thio-templated) specialised metabolites from marine cyanobacteria, providing reference to wider literature to direct the reader toward a complete description of biosynthetic pathways.

1.1.2.1 NRPS/PKS thio-templated

The majority of specialised metabolites from marine cyanobacteria are synthesised nonribosomally by the activity of NRPSs or PKSs, and in many cases, a hybrid of the two systems (NRPS/PKS) creating structurally and chemically diverse compounds. Many of these metabolites exhibit pharmacological relevant biological activities, such as polymerisation of tubulin (dolastatin 10 (Verdier-Pinard *et al.*, 1998a, Taori *et al.*, 2009), curacin A (Verdier-Pinard *et al.*, 1998b)), actin (hectochlorin (Marquez *et al.*, 2002), majusculamide), neurotoxins (jamaicamides (Edwards & Gerwick, 2004)), cytotoxins (coibamide A (Medina *et al.*, 2008, Serrill *et al.*, 2015), apratoxin (Luesch *et al.*, 2001b)), protein kinase C (PKC) activators (lyngbyatoxin (Basu *et al.*, 1992), aplysiatoxin (Eliasson *et al.*, 1983) antimalarials (bastimolide A (Shao *et al.*, 2015) and carmabins (McPhail *et al.*, 2007)), and antiviral (aplysiatoxin (Gupta *et al.*, 2014)). In addition, analogues of these have served as therapeutic leads. For example, brentuximab vedotin 63 (analogue of dolastatin 10) was recently FDA approved for the treatment of Hodgkin and systemic anaplastic large cell lymphoma (Martins *et al.*, 2014).

NRPSs are biocatalysts that are responsible for the synthesis of complex bioactive peptides independent of a nucleic acid-derived template. These peptides are initially created from a pool of ~ 500 amino acid residues, including proteinogenic and non-proteinogenic, and L- and D- conformations (Calcott & Ackerley, 2014, Gaudelli *et al.*, 2015). On the other hand,

PKSs are enzymes responsible for the formation of polyketides through Claisen condensations between an activated carboxylic acid starter unit and thioesterified malonate derivative.

Both NRPSs and PKSs are multimodular enzymes consisting of several domains, each with their specific function (Figure 1.1). Generally, they are composed of at least three core domains catalysing peptide/polyketide elongation and associated modifications. The adenylation domain (A-domain) and acyltransferase (AT) domains are responsible for the substrate recognition/selection and activation (Turgay *et al.*, 1992). The carrier protein peptidyl-carrier-protein (PCP) in NRPS or acyl carrier protein (ACP) in PKS domains tether the growing peptide or polyketide chain via covalent attachment to a 4'-phosphopantetheine moiety as a thioester and relocates it to the following module (Stachelhaus *et al.*, 1996). In NRPS, the condensation domain catalyses peptide bond formation of the activated aminoacyl intermediate and the peptidyl-bound intermediate (Stachelhaus *et al.*, 1998), whereas a ketoacyl synthase (KS) domain elongates the polyketide chain through decarboxylative condensation (Donadio & Katz, 1992). Additional tailoring domains, for example, ketoreductase (KR), methyltransferase (MT), and dehydratase (DH), catalyse further structural modifications. The number of modules present in an NRPS and their domain organisation typically determines the structure of the product (Du *et al.*, 2001). Due to the modular nature and complexity of NRPS catalytic ability, the resulting peptides are extremely diverse featuring linear, cyclic, branched or other complex primary structures. Additionally, the incorporation of amide or ester bonds, inclusion of halogens, and integration of glycosyl, and prenyl side groups by synthetase-associated tailoring enzymes further increases the level of complexity and diversity within these peptides (Calcott & Ackerley, 2014, Konz & Marahiel, 1999). PKSs are classified into different groups based on their protein structure and biosynthesis mechanism, that is, type I, type II and type III (reviewed in (Shen, 2003)). Type I modular PKSs are the most common mechanism for polyketide biosynthesis in cyanobacteria.

The NRPSs or PKSs are produced as an inactive apo-form and are posttranslationally modified by a phosphopantetheinyl transferase (PPTase), which attaches a 4'-phosphopantetheinyl arm to a conserved serine residue in the PCP or ACP domain, respectively

(Walsh *et al.*, 1997b, Suo *et al.*, 2001, Lambalot *et al.*, 1996). Some PPTases exhibit broad range activity, modifying many different carrier protein domains from diverse NRPS/PKSs (Copp & Neilan, 2006, Fu *et al.*, 2012, Quadri *et al.*, 1998b), while some have been shown to be specific (Roberts *et al.*, 2009).

A large number of cyanobacterial secondary metabolites are of mixed NRPS/PKS origin. Furthermore, many of the cyanobacterial metabolites discovered from marine environments are halogenated, resulting in vinyl halides, alkyl halides, and acetylenic bromide functional group (mechanisms reviewed in (Neumann *et al.*, 2008)). Studying the biosynthesis of these metabolites will provide insights to harness these gene clusters for the production of novel bioactive compounds through genetic engineering. In the following examples (Figure 1.2) we provide an overview of the origins, bioactivity, and unique or interesting aspects of the biosynthesis of each metabolite.

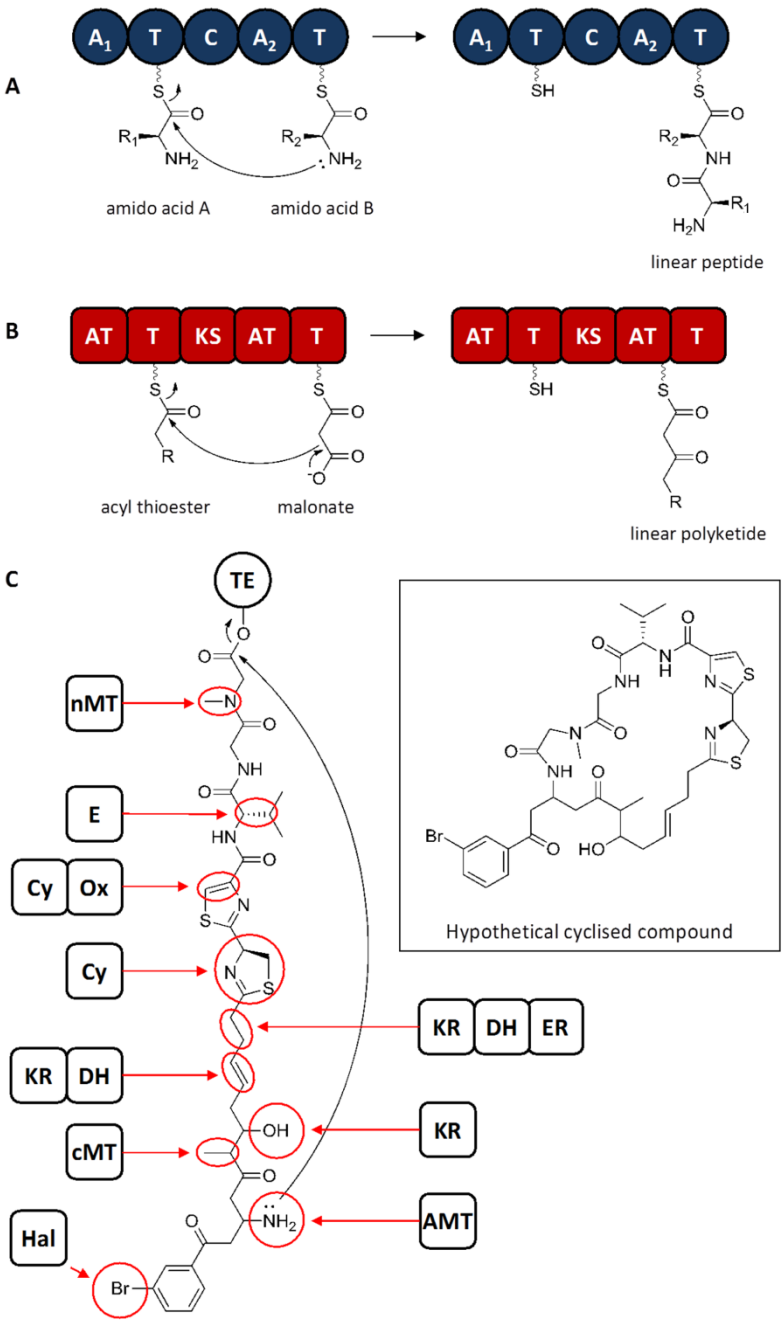


Figure 1.1 Outline of NRPS and PKS biosynthesis and associated tailoring modifications. Mechanism of one iteration of NRPS (A) or PKS (B) chain extension. A hypothetical compound, tethered to a thioesterase (TE), highlighting the structural diversity generated by NRPS and PKS tailoring domains (C). MT: methyltransferase; E: epimerase; Cy: cyclisation-condensation; Ox: oxidase; DH: dehydratase; KR: ketoreductase, ER: enoylreductrase; AMT: aminotransferase; Hal: halogenase.

1.1.2.1.1 Lyngbyatoxin

Lyngbyatoxins A-C (LTX A-C) are secondary metabolites isolated from a Hawaiian cyanobacterium *Moorea producens* (previousky known as *Lyngbya majuscula*) (Engene *et al.*, 2012), and are the common cause of dermatological condition known as "swimmer's itch" (Cardellina *et al.*, 1979, Aimi *et al.*, 1990). LTX A possesses structural similarities to the teleocidin family produced by Actinobacteria and is identical to one of the isomers, teleocidin A-1 (Sakai *et al.*, 1986). Recently, LTX analogues, 2,3-seco-2,3-dioxo-lyngbyatoxin have been identified (Youssef *et al.*, 2015). Members of the lyngbyatoxin family are indole alkaloids that possess tumour promoter activity via the activation of PKC, which regulates cell proliferation and differentiation (Basu *et al.*, 1992, Fujiki *et al.*, 1981). The indolactam V (ILV) ring of LTX, and teleocidin, is responsible for this activity (Basu *et al.*, 1992, Cardellina *et al.*, 1979, Youssef *et al.*, 2015). Based on this characteristic, interest has been drawn to the production of LTX and ILV analogues as PKC antagonists for cancer treatments (Kozikowski *et al.*, 1989).

The 11.3 kbp biosynthetic pathway (*ltxA-D*) was identified from a genomic DNA library constructed from bloom material (Edwards & Gerwick, 2004). LTX is produced by a bimodular NRPS (LtxA) from L-valine and L-typtophan, with the dipeptide release by an unusual NADPH-dependent reductase (Read & Walsh, 2007), cyclisation by a P450-dependent monoxogengenase (LtxB), and geranylation by a reverse prenyltransferase (LtxC) (Edwards & Gerwick, 2004). Interestingly, LtxB and LtxC both have relaxed substrate specificity for valyl group analogues (Huynh *et al.*, 2010). LtxD, an oxidase/reductase, is assumed to be responsible for the conversion of LTX A to LTX B and LTX C. An Alder-ene reaction is proposed for the formation of these two analogues (Videau *et al.*, 2016). Heterologous expression of the LTX biosynthetic pathway was attempted in *Streptomyces coelicolor* A3(2), but this host failed to produce full length mRNA transcripts of the NRPS (Jones *et al.*, 2012). Successful heterologous expression of LTX has been performed in an engineered *Escherichia coli* strain (Ongley *et al.*, 2013b) and the filamentous freshwater cyanobacterium *Anabaena* sp. PCC 7120 (Videau *et al.*, 2016).

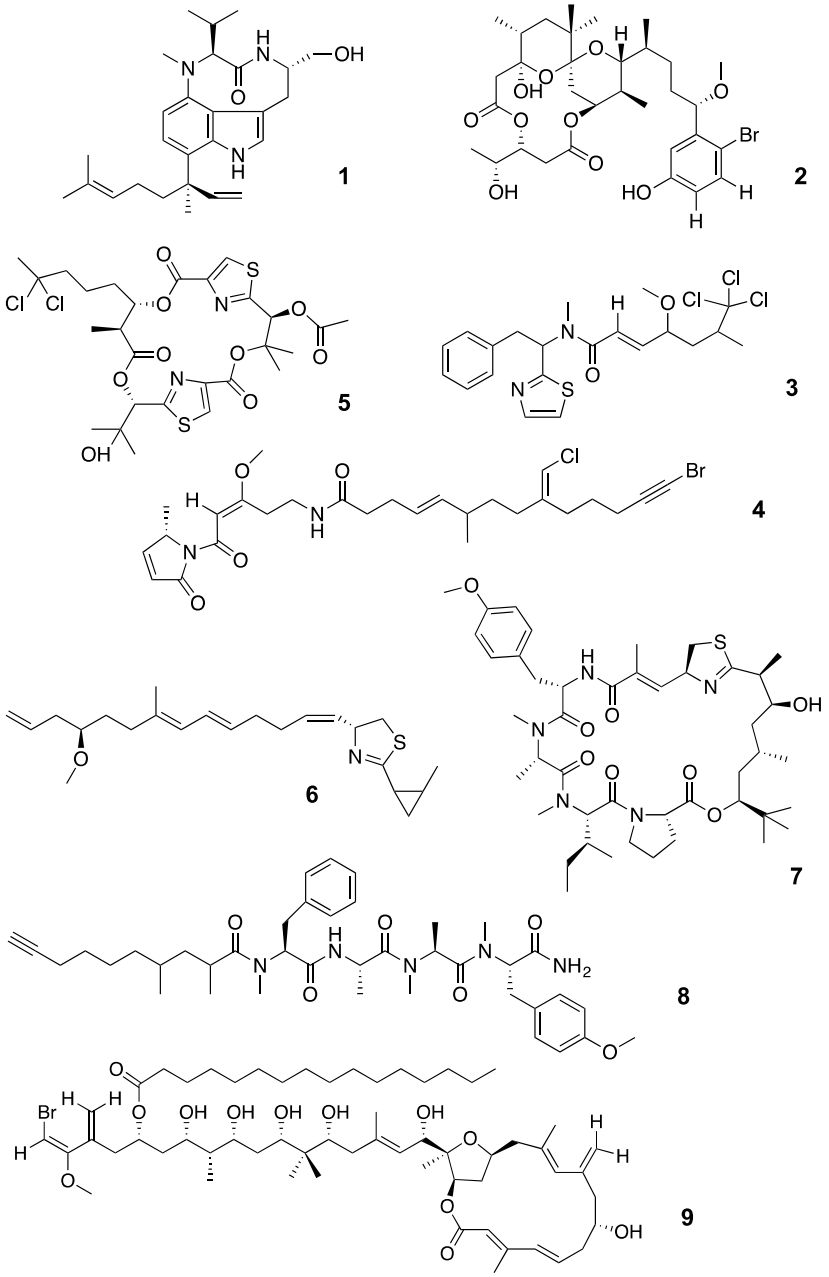


Figure 1.2 Overview of NRPS/PKS structures. Numbering of structures corresponds to the following respective sections (1.1.2.1.x), which provide an overview of each specialised metabolite.

1.1.2.1.2 Aplysiatoxin

Aplysiatoxins are polyketide marine toxins that, akin to LTX A, exhibit tumor promoter activity via the activation of PKC. Recent studies indicate that the modulation of PKC isozymes may present a strategy for the treatment of cancer, Alzheimer's disease, and acquired immunodeficiency syndrome (AIDS) (Irie & Yanagita, 2014). PKC activators, for example,

bryostatin 1 (a PKC activator with minor tumour-promoting activity) are currently one of the leading compounds for cancer treatment. Several studies have shown that the aplysiatoxin and synthetic analogues have potent anti-proliferative properties as potential anticancer agents, with some showing a low tumour promoting effect (Chlipala *et al.*, 2010, Eliasson *et al.*, 1983, Irie *et al.*, 2012, Irie & Yanagita, 2014, Kamachi *et al.*, 2013, Nakagawa *et al.*, 2011, Yanagita *et al.*, 2010). In addition, aplysiatoxin-related metabolites also show antiviral activity against the chikungunya virus (Gupta *et al.*, 2014).

Aplysiatoxins and their analogues have been isolated from several cyanobacterial species including *M. producens* (Chlipala *et al.*, 2010), *Schizothrix calcicola* (Moore *et al.*, 1984), *Oscillatoria nigro-viridis* (Moore *et al.*, 1984), and *Trichodesmium erythraeum* (Gupta *et al.*, 2014). Several natural aplysiatoxin analogues have been identified, including debromoaplysiatoxin, anhydroaplysiatoxin, anhydrodebromoaplysiatoxin, 3-methoxydebromoaplysiatoxin, and 3-methoxyaplysiatoxin (Chlipala *et al.*, 2010, Gupta *et al.*, 2014). Despite these discoveries, the genetic basis for biosynthesis of this metabolite remains unknown. However, the structure suggests the involvement of a PKS. Recently, a study identified two aplysiatoxin-related compounds nhatrangins A and B, which were suggested to be a possible intermediate of aplysiatoxin biosynthesis (Chlipala *et al.*, 2010).

1.1.2.1.3 Barbamide

Barbamide is an unusual trichlorinated lipopeptide with molluscidal activity, isolated from the marine cyanobacterium *M. producens* 19L, as a product of a mixed NRPS/PKS from acetate, L-phenylalanine, L-leucine, and L-cysteine precursors (Orjala & Gerwick, 1996, Chang *et al.*, 2002). The chemical structure of barbamide features a thiazole ring, N-methyldolaphenine moiety, and trichloroisovaleric acid (Sitachitta *et al.*, 1998, Williamson *et al.*, 1999). Dechlorobarbamide was previously isolated from Curaçao collections of the marine cyanobacterium *M. Producens* (Sitachitta *et al.*, 2000). In addition, several polychlorinated peptides, for example, barbaleucamides A-B (Harrigan *et al.*, 2001), dysidin, dysidinin, and a

series of chlorinated diketopiperazines from marine sponge *Dysidea* sp. (Flatt *et al.*, 2005) are believed to be of cyanobacterial origin.

The gene cluster for the biosynthesis of barbamide contains 12 open reading frames (ORFs) (*barA*-*barK*), spanning 26 kbp (Chang *et al.*, 2002). The biosynthesis of barbamide is initiated by the activation of L-leucine by BarD, which is then transferred to BarA (PCP) (Chang *et al.*, 2002, Sitachitta *et al.*, 2000). The tethered amino acids then undergo chlorination catalysed by BarB1/BarB2, which is subsequently released by BarC (Chang *et al.*, 2002). Downstream, a multifunctional NRPS/PKS module encoded by *barE* is responsible for the activation of the trichloroleucine intermediate as well as the loading of malonyl-CoA (supported by precursor study (Sitachitta *et al.*, 2000)), which is then modified by O-methylation and incorporated by BarF (Chang *et al.*, 2002). Further, the two modules within BarG, then catalyse the addition of N-methylated phenylalanine followed by condensation of cysteine forming a thiazoline ring (Chang *et al.*, 2002, Sitachitta *et al.*, 2000, Williamson *et al.*, 1999). The thioesterase (TE) domain at the C-terminus of BarG then release the barbamide intermediate, which then undergoes oxidative decarboxylation catalysed by BarJ, BarH and BarI, forming the thiazole ring in the final product barbamide (Chang *et al.*, 2002).

Barbamide biosynthesis has shown a unique NRPS/PKS system, as modules from one NRPS (BarA and BarD) and one PKS (BarE and BarF) are encoded on separate ORFs. In addition, the halogenation reactions in barbamide biosynthesis are particularly interesting as the trichlorination of leucine occurs in the initial biosynthesis stage and is carried out in tandem by nonheme Fe^{II} halogenases BarB2 followed by BarB1. Specifically, BarB2 catalyses dihalogenation *in trans*, and then the product is further converted to trichloroleucine by BarB1 (Flatt *et al.*, 2006, Galonic *et al.*, 2006). Interestingly, BarB2 is also able to chlorinate monochloroleucyl attached to BarA (Galonic *et al.*, 2006). Heterologous expression of the barbamide gene cluster in *Streptomyces venezuelae* DHS 2001 resulted in the production of 4-O-demethylbarbamide, which exhibits increased molluscicidal activity over barbamide (Kim *et al.*, 2012).

1.1.2.1.4 Jamaicamides

Jamaicamides are highly functional mixed polyketide/peptides isolated from *M. producens* JHB cultured from an environmental sample from Hector's Bay, Jamaica (Edwards *et al.*, 2004). The chemical structure of jamaicamides contains a rare alkynyl bromide (vinyl chloride, but only in jamaicamide A), an unusual β-methoxy enone, and a pyrrolinone ring. Several jamaicamide analogues have been identified, that is, jamaicamide A (brominated), jamaicamide B (debromo analogue) and jamaicamide C (slightly more hydrophobic), with only partial structures currently solved for jamaicamides D-F (Edwards *et al.*, 2004). Activity studies have shown that these compounds exhibit cytotoxicity to human lung and neuroblastoma cell lines (LC_{50}~15 μM), sodium channel blocking activity, and weak ichthyotoxicity to goldfish (Edwards *et al.*, 2004). In addition, jamaicamide B was shown to exhibit weak antimalarial activity (IC_{50} 18.4 μM) and a similar level of cytotoxicity (IC_{50} 16.2 μM) to Vero cells (McPhail *et al.*, 2007). Due to its activity, several attempts was conducted to chemically synthesise jamaicamide (Graf *et al.*, 2009, Tanaka & Usuki, 2011, Tanaka-Yanuma *et al.*, 2015, Watanabe *et al.*, 2013).

The mixed NRPS/PKS biosynthetic pathway of jamaicamides, *jam*, spans 58 kbp and is flanked by several transposases and hypothetical proteins (Edwards *et al.*, 2004). The cluster contains 17 ORFs mostly transcribed unidirectionally with only *jamQ* in the reverse direction (Edwards *et al.*, 2004, Dorrestein *et al.*, 2006). Several environmental factors, for example, salt concentration (Boopathi & Ki, 2014) and light intensity (Kaebernick *et al.*, 2000) have been shown to affect secondary metabolite production in cyanobacteria. Within the *jam* pathway, several strong internal promoters have been identified (Jones *et al.*, 2009), in particular, long intergenic regions between the first gene in the pathway (*jamA*) and the closest upstream ORF (a putative transposase). This region contains two putative transcription factors, protein 5335 and 7968, which are similar to the RcaD transcription factor (regulatory protein in response to light intensity and wavelength) (Jones *et al.*, 2009). In addition, another study proposed that the bromination reactions are affected by environmental parameters such as light intensity (Esquenazi *et al.*, 2011). Interestingly, these proteins have also been identified in *M. producens*

3L, suggesting similar regulatory mechanisms may be employed for the biosynthesis of other pathways.

1.1.2.1.5 Hectochlorin

Hectochlorin was isolated from marine isolates of *M. producens* collected from Hector's Bay (Jamaica) and Boca del Drago Beach (Bocas del Toro, Panama) (Marquez *et al.*, 2002). Based on nuclear magnetic resonance (NMR) data, the compound possesses a *gem*-dichloro group and two units of α,β-dihydroxyisovalerate (DHIV), with two thiazole rings (Marquez *et al.*, 2002). Hectochlorin shows antifungal activity against *Candida albicans*, and cytotoxic activity, by promoting actin polymerisation with activity similar to jasplankinolide, a cyclic peptide isolated from a marine sponge (Marquez *et al.*, 2002, Bubb *et al.*, 1994). Hectochlorin exhibits potent cytotoxicity against several tumour cell lines including colon, ovarian, renal, melanoma (Marquez *et al.*, 2002, Suntornchashwej *et al.*, 2005), and small cell lung (Suntornchashwej *et al.*, 2005). Hectochlorin, and its derivative deacetylhectochlorin, has also been isolated from the mollusc *Bursatella leachii*, which may be of cyanobacterial origin (Suntornchashwej *et al.*, 2005). Based on the bioactivity of hectochlorin, research has focused on the chemical synthesis of this compound (Cetusic *et al.*, 2002) and characterisation of its biosynthetic pathway (Ramaswamy *et al.*, 2007).

The biosynthetic pathway of hectochlorin spans 38 kbp and comprises eight ORFs (Ramaswamy *et al.*, 2007). The biosynthesis is initiated with HctA, an acyl-ACP synthetase similar to *jamA* in jamaicamide pathway (Ramaswamy *et al.*, 2007, Edwards *et al.*, 2004). The N-terminal region of HctB, a nonheme iron and α-ketoglutarate halogenase, is suggested to be responsible for the chlorination, based on its similarity with BarB1/BarB2 (barbamide (Chang *et al.*, 2002)) and SyrB2 (syringomycin (Vaillancourt *et al.*, 2005)), and later supported by its biochemical characterisation (Pratter *et al.*, 2014). Further investigation of HctB has shown that it has a multifunctional metal centre which catalyses oxygenation, dichlorination, and formation of vinyl chloride. Chlorination occurs in two distinct steps in a chloride-dependent fashion with

a relaxed substrate promiscuity. Interestingly, chloride levels also affect the O_2 activation efficiency.

1.1.2.1.6 Curacin

Curacin A is an active antiproliferative and cytotoxic compound initially isolated from strains of *M. producens* from the Curaçao collection (Gerwick *et al.*, 1994). The mechanism of curacin's cytotoxicity is similar to that of dolastatin, that is, interaction with colchicine binding sites of tubulin inhibits microtubule polymerisation, which ultimately blocks the cell cycle and induces cell death (Blokhin *et al.*, 1995, Verdier-Pinard *et al.*, 1998b). Natural analogues of curacin A have been discovered, that is, curacin B and curacin C (isomers exhibiting similar toxicity with curacin A) (Yoo & Gerwick, 1995), an 8-desmethyl analogue curacin D (lower toxicity) (Marquez *et al.*, 1998), and curazole (containing a thiazole moiety instead of thiazoline) (Simmons *et al.*, 2008). Structurally, curacin A is composed of two lipid chains with a unique thiazoline and cyclopropane moiety, which indicates a mixed polyketide-nonribosomal peptide biosynthesis (Gerwick *et al.*, 1994, Chang *et al.*, 2004). Due to the high *in vitro* bioactivity of curacin, optimisation of curacin's biosynthesis has been attempted (Rossi *et al.*, 1997). However, curacin is a highly lipophilic molecule, which limits its clinical applications. Hence research has been conducted to modify curacin A and its analogues through synthetic or semisynthetic approaches to improve its water solubility (Muir *et al.*, 2002, Wipf *et al.*, 2002b, Wipf *et al.*, 2002a).

The biosynthetic cluster for the curacin A biosynthesis was sequenced from *M. producens* 19L, spanning 64 kbp containing 14 ORFs *curA-curN* in one direction (Chang *et al.*, 2004). Nine type I PKS modules and a hybrid PKS/NRPS module were identified in the cluster. The first AT domain of *curA* is proposed to be inactive, and the second domain, AT_{A2}, is proposed to load both the malonyl-CoA starter unit and the first extending malonyl-CoA to their respective domains. The KS domain in CurA then catalyses the formation of a diketide intermediate from these two malonyl units. This is followed by the formation of the unique cyclopropyl ring, which is important for biological activity (Blokhin *et al.*, 1995, Verdier-Pinard

et al., 1998b, Wipf *et al.*, 2002a), by a 3-hydroxy-3-methylglutaryl (HMG)-CoA synthase (HCS)-like enzyme (Chang *et al.*, 2004, Gu *et al.*, 2006) similar to that which forms the vinyl chloride substituent in jamaicamide (Edwards *et al.*, 2004).

1.1.2.1.7 Apratoxin A

Apratoxins are potent cytotoxic cyclodepsipeptides isolated from different species and strains of the genus *Moorea*. The structure of apratoxin A, initially isolated from *M. producens* from Finger's Reef (Apra Harbor, Guam), features both polypeptide fragments of three methylated amino acids (O-methyl-L-tyrosine, N-methyl-L-alanine, N-methyl-L-isoleucine), an L-proline, an unsaturated modified D-cysteine (thiazoline ring) and a polyketide-derived unit, 3,7-dihydroxy-2,5,8,8-tetramethylnonanoic acid (Dtena) (Luesch *et al.*, 2001b). Several natural analogues have been identified, including desmethyl apratoxin B and apratoxin C (*Lyngbya* sp., Guam and Palau) (Luesch *et al.*, 2002a), apratoxin H (with pipecolic acid instead of proline), and apratoxin A sufoxide with (oxidised sulfur atom) (*M. producens*, Red sea) (Thornburg *et al.*, 2013), apratoxin F and G (with N-methyl alanine substituting proline) (*Lyngbya bouilonii*, Palmyra Atoll) (Tidgewell *et al.*, 2010), apratoxin D (with 3,7-dihydroxy-2,5,8,10,10-pentamethylundecanoic acid as the polyketide moiety) (*Lyngbya sordida* and *M. producens*, Papua New Guinea) (Gutierrez *et al.*, 2008), and apratoxin E with a dehydrated polyketide unit (*Lyngbya bouillonii*, Guam) (Matthew *et al.*, 2008). Apratoxins have been shown to be promising cancer drug leads with IC_{50} within the nanomolar range against different cancer cell lines (Huang *et al.*, 2016). Extensive studies have been performed to identify the structure-activity relationship of apratoxins in order to develop the best drug leads (refer to (Thornburg *et al.*, 2013) for details). Recently, it has been proposed that the cytotoxic activity of apratoxin A manifests through interference of cotranslational translocation, leading reversible inhibition of the secretory pathway (Liu *et al.*, 2009), in particular through the binding to Sec61 (Huang *et al.*, 2016). Due to its high cytotoxic activity, numerous studies have attempted to chemically synthesise apratoxin A and its analogues (Gilles *et al.*, 2011, Chen *et al.*, 2011, Chen *et al.*, 2014, Dey *et al.*, 2015).

The biosynthetic cluster of apratoxin was deduced by employing a novel approach combining single-cell genomic sequencing based on multiple displacement amplification (MDA) and metagenomic library screening (Grindberg *et al.*, 2011). The *apr* biosynthetic cluster was isolated from marine cyanobacterium *L. bouillonii*, spanning 58 kbp with 12 ORFs (*aprA-L*), encoding a type I modular mixed PKS/NRPS flanked by putative transposases and hypothetical proteins. It has been suggested that the adenylation domains of the NRPS enzymes might possess relaxed substrate specificity due to the presence of naturally occurring apratoxin analogues. In addition, cryptic genes and/or environmental factors might also play a role in the diversification of apratoxin family (Thornburg *et al.*, 2013).

1.1.2.1.8 Carmabins

Carmabin A and B are linear lipotetrapeptides previously isolated from the marine cyanobacterium *M. producens*. Initial isolation showed that the compounds have potential antiproliferative activity (Hooper *et al.*, 1998). The chemical structure of carmabin A features three methylated amino acids (N-Me,O-MeTyr, N-MeAla, and N-MePhe), Ala, and an unusual fatty acid chain acetylene functionality, 4-dimethyl-9-decynoic acid. In carmabin B, the acetylene unit is replaced with a methyl ketone (Hooper *et al.*, 1998). Related bioactive compounds including dragonamide, dragomabin, dragonamide B (Jimenez & Scheuer, 2001, McPhail *et al.*, 2007), and almiramides A-C (Sanchez *et al.*, 2010) have been previously isolated from a Panamanian strain of *M. producens*. Carmabin A shows antimalarial activity (IC_{50} 4.3 μM) and cytotoxic activity against Vero cells (IC_{50} 9.8 μM) (McPhail *et al.*, 2007).

Genome sequencing of carmabin A-producing *M. producens* 3L led to the identification of genes putatively involved in carmabin A biosynthesis, based on the bioinformatically-predicted substrate specificities of the NRPS (Jones *et al.*, 2011). However, sequencing of the *M. producens* 19L genome facilitated the identification of a 42 kbp mixed PKS/NRPS pathway, containing 10 ORFs, for carmabin A biosynthesis (Ramaswamy, 2005).

1.1.2.1.9 Phormidolide

Phormidolide is a brominated macrolide isolated from a laboratory culture of marine cyanobacterium *Phormidium* sp. (reassigned as *Leptolyngbya* sp.) collected originally from Sulawesi, Indonesia (Williamson *et al.*, 2002). The unique structure of the phormidolide backbone comprises a 16-membered lactone ring with unique vinyl bromine functionality. Phormidolide A was shown to have potent brine shrimp toxicity (LC_{50} of 1.5 μM) (Williamson *et al.*, 2002). Related compounds have been isolated including cytotoxic phormidolides B and C from marine sponges (Lorente *et al.*, 2015), as well as oscillariolide from the marine cyanobacterium *Oscillatoria* sp. (Murakami *et al.*, 1991). Recently, chemical synthesis of the bromomethoxydiene moiety common to oscillariolide and phormidolides A–C has been described (Gil *et al.*, 2016). The biosynthesis of phormidolide was deduced based on genome sequence, followed by bioinformatic analysis in conjunction with results from stable isotope feeding experiments.

The biosynthetic cluster of phormidolide (*phm*), spanning 94 kbp, contains a *trans*-AT PKS pathway (Bertin *et al.*, 2016) in which freestanding AT domains load substrates to the modular ACP (refer to (Shen, 2003) for a review of PKS biosynthesis). An interesting feature of phomidolide biosynthesis is that the starter unit is not acetate derived; rather, it is a three-carbon starter unit proposed to be 1,3-bisphosphoglycerate. The proposed biosynthesis of phormidolide is initiated with the loading and modification (dephosphorylation, dehydration and O-methylation) of the starter unit on PhmE. PKS modules within PhmE then further extend the molecule to form β-ketone, which is further modified by HCS cassettes to form a branching vinyl group. Next, the final PKS module in PhmE adds another acetate residue (ketone group is reduced) followed by five more additions by PhmF, each with modification by KR domains and S-adenosylmethionine (SAM) methylation. The penultimate ketone undergoes another modification by HCS cassette. The biosynthesis continues with elongation by PhmH, coupled with the formation of an olefin intermediate. The four PKS modules within PhmI, together with the HCS cassette enzymes, synthesise the remainder of the phormidolide backbone, ending with

cyclisation and subsequent release by the final domain in PhmI. Bromination of the backbone is proposed to be catalysed by PhmJ, while the transfer of palmitic acid is presumed to be catalysed by PhmL and PhmM (Bertin *et al.*, 2016). Apart from phormidolide, only a few *trans*-AT PKS systems have been discovered in cyanobacteria, including tolytoxin, swinholide, and luminaolide, some of which are produced by marine symbionts. It could be anticipated that, genome and metagenome analysis may reveal that more of these unique biosynthesis systems of cyanobacterial origin. In addition, this finding highlights the use of genome mining as a tool to characterise complex marine cyanobacterial natural product biosynthetic pathways.

1.1.2.1.10 Columbamides

The columbamides are dichlorinated (columbamide A) and trichlorinated (columbamide B) acyl amides, recently discovered through the integration of mass spectrometry-based metagenomic approach and genomic analysis of *Moorea bouillonii* PNG-198 (Kleigrewe *et al.*, 2015). Columbamide C, found in small amounts, is presumed to be the precursor of columbamide A, with the absence of acetoxy group. Columbamide A and B show a high affinity toward cannabinoid receptors CB_1 and CB_2 (Kleigrewe *et al.*, 2015). Cannabinoid receptor ligands have been extensively studied for various biological applications especially for the treatment of nervous system-related disorders (further reading (Fijal & Filip, 2016, Nikan *et al.*, 2016)).

A ~ 60 kbp gene cluster, *col*, putatively responsible for columbamide biosynthesis, has been identified (Kleigrewe *et al.*, 2015). The biosynthesis of columbamide is initiated with the loading of dodecanoic acid by an acyl-CoA synthase ColA onto the ACP encoded by *colC*. Halogenations of the terminal end (one for columbamide A and two for columbamide B), and at the *ω*-7 position, are catalyzed by ColD and ColE, new type of halogenases distinct from those found in the barbamide pathway. Next, a bimodular PKS ColF is postulated to catalyze two rounds of acetate incorporation, followed by dehydration. The adjacent NRPS module, ColG, then incorporates serine and releases the product by a reductase similar to NADPH-dependent reductive cleavage in LTX biosynthesis (Edwards & Gerwick, 2004). An AT ColI then acetylates the product to form columbamides A and B (Kleigrewe *et al.*, 2015).

The discovery of the columbamide biosynthetic pathway is another example of the use of genomic tools in combination with mass spectrometry. In particular, the use of previously described regulatory domains from *M. producens* 3L and *M. producens* JHB further highlights the need to investigate the closely related regulatory mechanism employed in NRPS/PKS biosynthesis.

1.1.2.2 Ribosomally synthesised

As previously described, marine organisms, in particular cyanobacteria, hold a high potential as a source of bioactive metabolites. Apart from NRPs and PKs, marine cyanobacteria also produce several bioactive peptides that are ribosomally produced, termed RiPPs; these are ribosomally synthesised and postranslationally modified peptides (Figure 1.3). Due to the high variability in precursor peptide sequences, the natural promiscuity of modifying enzymes, and the modularity of pathways, RiPPs present a promising platform for combinatorial engineering of new peptides (for a review see (Sardar & Schmidt, 2016)). Herein, two classes of RiPPs produced by marine cyanobacteria are briefly described.

1.1.2.2.1 Cyanobactins

Cyanobactins are members of class of RiPP synthesised as a precursor peptide flanked with a recognition sequence that is matured through proteolytic cleavage and cyclisation and may undergo a variety of postranslational modifications. Several cyanobactins have shown to possess bioactivities, for example, cytotoxicity, antifungal, antibacterial, antimalarial, antiviral, and immunomodulation properties (reviewed in (Sivonen *et al.*, 2010)). The most described cyanobactins from marine cyanobacteria are patellamides, trunkamides, and trichamides.

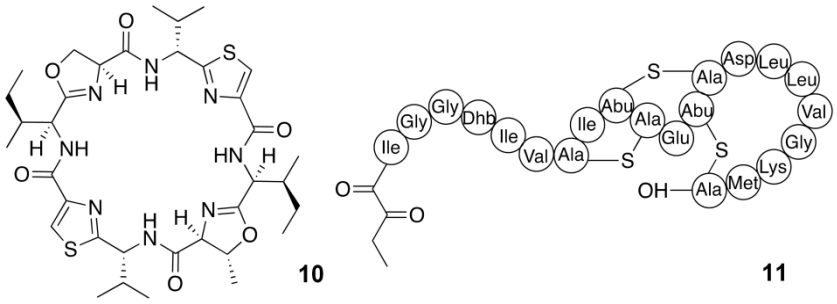


Figure 1.3 Marine cyanobacterial ribosomally-synthesised (RiPP) natural products. 10, patellamide A; 11, prochlorosin. The lanthipeptide (**11**) undergoes posttranslational modification by the ProcM tailoring system which cyclises the peptide through the formation of thioether crosslinks (C-S-C).

Patellamides are cytotoxic cyclic peptides originally isolated from the marine ascidian *Lissoclinum patella* (Ireland *et al.*, 1982). Later, it was shown that these compounds are produced ribosomally by its cyanobacterial symbiont, *Prochloron didemni* (Schmidt *et al.*, 2005). *Prochloron* spp. are symbiotic cyanobacteria that are associated with marine animals and stromatolites (Withers *et al.*, 1978, Burns *et al.*, 2004). Patellamide A is an octapeptide containing two oxazoline and two thizole rings. The biosynthesis of patellamide A was proposed through genome sequencing of *P. didemni* and confirmed by heterologous expression in *E. coli* (Schmidt *et al.*, 2005). The *pat* cluster, spanning 11 kbp, comprises *patA*-*G*. The precursor peptide for patellamides A and C (PatE) contains the leader peptide, N-terminal recognition elements, and a core peptide with two eight-residue peptides that are converted to patellamide A and C. Firstly, the core peptide within PatE undergoes heterocylisation of cysteine, serine, and threonine residues by PatD (and potentially PatF), forming thiazoline and oxazoline rings (Schmidt *et al.*, 2005, Sudek *et al.*, 2006). The subtilisin-like protease PatA then removes the leader peptide and N-terminal recognition sequence (Lee *et al.*, 2009, Schmidt *et al.*, 2005). PatG then catalyzes cleavage of the C-terminal recognition sequence in tandem with macrocyclization and potentially is involved in the oxidation of thiazolines to form thiazoles (Schmidt *et al.*, 2005, Lee *et al.*, 2009). Interestingly, an *in vitro* study of PatG shows that this enzyme has a very relaxed substrate specificity, as long as the C-terminal recognition sequence is present, which makes it a suitable catalyst for general peptide cyclisation (Lee *et al.*, 2009).

For further structural details and enzymology, refer to (Houssen *et al.*, 2012, Koehnke *et al.*, 2013, Koehnke *et al.*, 2014, Koehnke *et al.*, 2012).

Trunkamides, and their related cyclic peptide patellins, were initially isolated from the marine ascidian from *Lissoclinum* sp. (Carroll *et al.*, 1996). They were subsequently shown to have a cyanobacterial origin from *Prochloron* sp. (an oligate symbiotic unicellular cyanobacteria of the ascidian *L. patella*, found in tropical reefs) (Donia *et al.*, 2008). The patellin biosynthetic pathway (~ 11 kb), *tru*, is highly similar to the *pat* pathway (Carroll *et al.*, 1996). In contrast to patellamide, these metabolites possess prenylated serine and threonine residues and a reduced form of thiazoline (heterocylisation of cysteine). In addition, TruF1-TruF2 (homologues of PatF) together with TruD (homologues of PatD), are proposed to perform prenylation of patellins, instead of heterocyclisation observed in patallamides (Donia *et al.*, 2008).

Trichamide, isolated from the filamentous marine cyanobacterium *T. erythraeum* (Sudek *et al.*, 2006) is biosynthesised by the 12.5 kbp *tri* pathway containing 11 ORFs (*triA-K*). In this pathway, the precursor peptide TriG is modified by TriA through heterocyclization of cysteine, forming a thiazoline group, which is then oxidized by TriD. TriA and TriK cleave the propeptide, leading to the cyclization and maturation by TriH and TriK (Sudek *et al.*, 2006).

Due to the mechanism of biosynthesis, cyanobactins possess a high level of structural diversity, creating a wide range of potentially bioactive compounds. Variation in the PatE precursor peptide sequence resulted in the natural diversification of the patellamides and the related peptides ulithiacyclamides and lissoclinamides (Donia *et al.*, 2006). Similar observations are seen within the *tru* pathway with different variants of product coding regions, *truE2* and *truE3*, responsible for the formation of octapeptide patellin 6 and heptapeptide trunkamide, respectively (Donia *et al.*, 2008). These pathways have exhibited relaxed sequence selectivity *in vitro* and *in vivo* (Donia *et al.*, 2008, Ruffner *et al.*, 2015), highlighting the potential of exploiting this pathway mechanism to produce new drug leads through *in vivo* and *in vitro* means (Tianero *et al.*, 2012, Houssen *et al.*, 2014, Goto *et al.*, 2014, Ruffner *et al.*, 2015). In fact, an attempt to exploit this versatile mechanism, by creating a mutant library of *tru* pathway

in *E. coli*, yielded > 300 new compounds despite some substrate restrictions (Ruffner *et al.*, 2015).

1.1.2.2.2 Lanthipeptides

Lanthionine-containing peptides, or lanthipeptides, are a family of RiPPs that are characterised by the presence of intramolecular thioether cross-linkages formed between a dehydrated Ser or Thr residue and a Cys residue. Many of these RiPPs possess antimicrobial activity and are thus termed lanthibiotics (Piper *et al.*, 2009). There are four different classes of lanthipeptides based on their biosynthetic routes (reviewed in (Yu *et al.*, 2013)). To date, only class II lanthipeptides have been described from cyanobacteria. In class II lanthipeptides, the precursors peptides (generic term LanAs) have C-terminal core peptides, which are modified by a bifunctional synthase (generic name LanM) with dehydration and cyclisation.

Prochlorosins are class II lanthipeptides produced by the unicellular marine cyanobacteria *Prochlorococcus* and *Synechococcus* (Li *et al.*, 2010). The expression of prochlorosins is dependent upon environmental conditions, and thus it may serve a physiological purpose. Interestingly, the genome of *Prochlorococcus* MIT9313 contains 29 genes encoding for LanA homologues (referred as *proc*As), as the substrate for a single LanM homologue (*procM*). In addition, *proc*As have highly conserved leader peptides (directing the modification), but extremely diverse core peptides varying in length and composed of serines, threonines and cysteines at every position (Li *et al.*, 2010). ProcM catalyses the postranslational modification of ProcA (Piper *et al.*, 2009, Tang & van der Donk, 2012). The presence of the full-length leader peptides is not necessary for the posttranslational modification of some of the precursor peptides (Voigt et al., 2014). ProcM contains conserved zinc-binding motifs, which are known to be important for cyclisation. However, the ligands for the zinc ion in ProcM comprise three cysteine residues, whereas other known LanM homologues contain only two. This unique characteristic of ProcM was hypothesised to increase substrate selectivity of ProcM, enabling the enzyme to catalyse a diverse range of precursor peptides (procAs) (Zhang et al., 2014). A closer look at the kinetic properties of ProcM has shown that, due to its relaxed

substrate specificity, this enzyme possesses a lower catalytic efficiency than HalM2. HalM2 is a LacM homologue for haloduracin biosynthesis in *Bacillus halodurans*, which has an activity specific to one substrate (Thibodeaux *et al.*, 2014, McClerren *et al.*, 2006). As such, the promiscuous nature of ProcM allows for a great level of diversification, possibly allowing the generation of new cyclic compounds.

1.1.3 Tools for the discovery and characterisation of marine bioactive natural products

1.1.3.1 Genome mining

The advancement of genomic technology, combined with analytical chemistry and dereplication strategies, has allowed marine cyanobacteria to be a treasure trove for the rapid discovery of new bioactive compounds. Some are from a new, or known, genus; others have yet unknown bioactive metabolite capacities (for recent examples refer to (Gunasekera *et al.*, 2016, Shao *et al.*, 2015, Ogawa *et al.*, 2016)). In addition, techniques such as tandem mass spectroscopy (MS/MS) can be used as a dereplication (screening) strategy to avoid isolating known compounds (Yang *et al.*, 2013). Genomics and metagenomics approaches have also been useful to discover new bioactive compounds (reviewed in (Micallef *et al.*, 2015)). These tools are especially noteworthy for natural products derived from NRPSs and PKSs. Due to the modular structures of NRPS and PKS, it is possible to deduce the biosynthetic pathway based on chemical structure and vice versa. For example, the elucidation of bartoloside was guided by the investigation of its biosynthesis derived from genomic data (Leao *et al.*, 2015). In addition, homologous enzymes (jamaicamide and cyanobactins), regulatory genes (columbamide), and metagenomic analysis (columbamide) have also been subjects of targeted investigations.

Recently, MDA has been used for biosynthetic cluster discovery. Previous approaches of genome mining, for example, polymerase chain reaction (PCR) screening of genomic libraries and genomic bioinformatic analysis, rely on the ability to isolate high-quality DNA, which can be challenging to obtain for marine cyanobacteria. In the case of filamentous bacteria, the polysaccharide sheath surrounding the filaments form a habitat for a rich

population of heterotrophic bacteria and creates difficulties for establishing axenic cultures as well as the extraction of adequate yields of high-quality pure cyanobacterial genomic DNA. For example, mild antibiotic application resulted in bleaching and death of *L. bouillonii* (Grindberg *et al.*, 2011). However, by using high-throughput single-cell genome amplification, the biosynthetic pathway responsible for the production of apratoxin was located (Grindberg *et al.*, 2011). This approach can be further adapted for pathway characterisation for organisms that are intractable.

1.1.3.2 Heterologous expression

Despite the discovery of new biosynthetic pathways from cyanobacteria, the industrial-scale production of compounds, as well as functional characterisation of the genes within the cluster, is difficult to achieve, because of low compound yields, slow growth, and a lack of genetic modification techniques. For example, most species of cyanobacteria do not have established DNA transformation methods, rendering targeting knockouts or mutagenesis nearly impossible. In some cases, comparative genomics between different organisms producing the same or similar spectrum of compounds can serve as "virtual knockout or modification" to assist in the assigning gene function. However, this is only effective and true if all the genes involved in biosynthesis are contained within a known cluster. Heterologous expression can bypass these problems to facilitate full characterisation of the biosynthetic cluster and thus to produce sustainably the compound for further use. In particular, when it is not possible to insert an exogenous promoter to forcibly drive expression of the pathway, heterologous expression can be essential for elucidating products from orphan or cryptic natural product pathways. Examples for heterologous expression of marine cyanobacterial natural products include patellamide (Schmidt *et al.*, 2005), 4-O-demethylbarbamide (Kim *et al.*, 2012), and LTX (Ongley *et al.*, 2013b). As more heterologous host systems are developed and fine-tuned, it is foreseeable that heterologous expression will play an integral part in furthering natural product discovery and our understanding of cyanobacterial natural product biosynthesis.

1.1.4 Conclusion

Marine dinoflagellates and cyanobacteria produce a wide array of specialized metabolites. Biosynthesis of these metabolites is typically directed by a thio-templated mechanism, namely, NRPSs or PKSs. These enzymes incorporate substrates from a greater pool than ribosomal synthesis allows, providing the observed structural diversity and a wide range of bioactivities. In addition, the nature of the marine environment has led to the evolution of specialized metabolites where halogen (i.e., Cl, Br) decorations are common, moieties that generally increase the bioactivity of compounds and are typically less frequent in specialized metabolites from terrestrial sources. The modularity of enzymes for both thio-templated and ribosomally synthesized metabolites lends themselves to the creation of "designer biosynthetic pathways" or applications in chemoenzymatic biosynthesis (where enzymes are utilized to catalyze difficult reactions in chemical synthesis). It is anticipated that advances in technology will enable the full natural product capacity of marine microalgae to be realized and harnessed for the sustainable production of these compounds in the biotechnological and pharmacological industries.

1.2 Saxitoxin

The paralytic shellfish toxins (PSTs) are neurotoxic alkaloids globally associated with paralytic shellfish poisoning (PSP). Interestingly, PSTs native producers span two domains of life, the prokaryotic freshwater cyanobacteria (Llewellyn *et al.*, 2001, Lagos *et al.*, 1999, Mahmood & Carmichael, 1986, Foss *et al.*, 2012, Carmichael *et al.*, 1997, Stucken *et al.*, 2010, Pomati *et al.*, 2000, Borges *et al.*, 2015, Smith *et al.*, 2011) and the marine dinoflagellates (Lefebvre *et al.*, 2008, Oshima *et al.*, 1993, Usup *et al.*, 1994) (for more details about dinoflagellates see section 1.1.1.1), raising questions into their biosynthetic evolutionary and ecological role (Stuken *et al.*, 2011, Murray *et al.*, 2015, Orr *et al.*, 2013). PSTs cause human and animal intoxication through exposure to contaminated freshwater and ingestion of shellfish, which have accumulated the toxins (Turner *et al.*, 2018, Etheridge, 2010, Landsberg, 2002). This illness is attributed to the reversible binding of PSTs onto site 1 of voltage-gated Na^+ channels, causing neuronal

transmission blockage (Noda *et al.*, 1989, Terlau *et al.*, 1991). Intoxication manifests in nausea, numbness, paralysis, and in some severe cases respiratory failure and death. The parent molecule saxitoxin is the most lethal PST with an i.p. LD_{50} of 3–10 mg/kg body weight in mice (Cusick & Sayler, 2013). PSTs also bind to voltage-gated K^+ (Wang *et al.*, 2003) and Ca^{2+} channels (Su *et al.*, 2004), and can cross the blood-brain barrier (Andrinolo *et al.*, 1999). Recently, several PSTs have been studied for pharmaceutical applications such as anesthetic agents (Rodriguez-Navarro *et al.*, 2009, Lobo *et al.*, 2015, Epstein-Barash *et al.*, 2009, Kohane *et al.*, 2000) and treatments for anal fissures (Garrido *et al.*, 2007).

Saxitoxin and its analogues feature a unique heterocyclic guanidine with different functional groups (Figure 1.4). More than 57 analogues of saxitoxin have been reported, including the non-sulfated forms, saxitoxin (STX) and neosaxitoxin (neoSTX); the monosulfated gonyautoxins (GTXs); the disulfated C-toxins; the acetylated variants of *Lyngbya wollei* toxins (LWT1-6); the decarbamoylated toxins (dc-toxins); and the deoxy-decarbamoylated toxins (doSTX and doGTX 1-3) (Wiese *et al.*, 2010). The complex structures of PSTs along with their high polarity contributed by the guanidium moieties have made chemical synthesis of these molecules difficult to achieve.

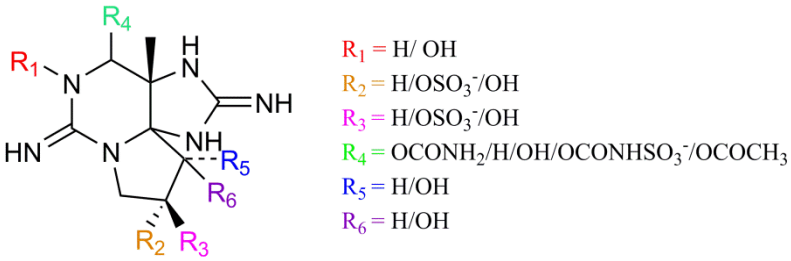


Figure 1.4 Structures of paralytic shellfish toxins (PSTs). Functional group substitutions (R_1-R_6) of the analogues are listed on the right.

The saxitoxin biosynthetic cluster was first discovered in the freshwater cyanobacterium *Raphidiopsis raciborskii* T3 (Kellmann *et al.*, 2008). Their biosynthesis is initiated by the formation of 4-amino-3oxo-guanidinoheptane (Intermediate A' or Int-A') by the polyketide synthase-like enzyme, SxtA (Figure 1.5). The AT domain of the enzyme, SxtA2, first loads malonyl-CoA to the ACP domain, SxtA3. An S-adenosylmethionine (SAM)-

dependent methyltransferase, SxtA1, then methylates the acyl-ACP to form propionyl-ACP. A Claisen condensation with arginine catalysed by SxtA4, a class II aminotransferase, produces the first saxitoxin precursor, Int-A'. Next, SxtG, an amidinotransferase catalyses the addition to Int-A'of a second amidino group from arginine. A series of reactions including cyclisation, desaturation, and epoxidation are proposed to subsequently catalysed by modifying enzymes encoded in the cluster, to form the saxitoxin precursor of decarbomyl STX (dcSTX). The addition of the carbamoyl group is presumably performed by O-carbamoyltransferase SxtI, converting dcSTX into STX (Kellmann *et al.*, 2008). The biosynthesis of saxitoxin is unique as it involves only one round of the PKS-type reaction followed by multiple downstream modifying reactions (Figure 1.5). The complexity of this pathway makes the characterisation and expression of saxitoxin biosynthesis challenging.

Several studies have described the gene clusters for the biosynthesis of PSTs in other species of cyanobacteria including *Anabaena circinalis* AWQC131C (Mihali *et al.*, 2009), *Aphanizomenon sp.* NH-5(Mihali *et al.*, 2009), *Raphidiopsis brookii D9* (Stucken *et al.*, 2010), *Lyngbya wollei* (Mihali *et al.*, 2011), and *Scytonema crispum* (Cullen *et al.*, 2018a). The clusters vary in size, gene composition and organisation. The presence of different tailoring enzymes is thought to be involved in the diversification of STX, and to have resulted in the different toxin profiles observed between producer strains (D'Agostino *et al.*, 2019). Bioinformatic analysis, in some cases supported by biochemical characterisation, has been used to assign the functions of the saxitoxin tailoring enzymes. The tailoring gene *sxtX*, which encodes for a homolog of cephalosporine hydrolase enzyme, is proposed to convert STX to neoSTX through N1 hydroxylation. The enzyme is present in all strains, except the non-N1 hydroxylated PST producers (*D. circinale* AWQC131C and *R. Brooki* D9) (Mihali *et al.*, 2009). Recently, SxtDIOX, a Rieske oxygenase was experimentally shown to catalyse the formation of C11 hydroxylated STX analogues (Lukowski *et al.*, 2018). These analogues are the substrates for StxSUL (a putative sulfotransferase) which catalyses the production of the O-sulfated congeners of gonyautoxin (GTX 1-4). The biosynthesis of sulfated analogues GTX5, but not GTX6, is catalysed by the N-sulfotransferase, SxtN (Cullen *et al.*, 2018a, Soto-Liebe *et al.*, 2010). The

production of C-toxins, the doubly sulfated analogues, is likely to involve the combined catalytic activity of SxtDIOX, SxtSUL and SxtN (Soto-Liebe *et al.*, 2010). The sulfur required for sulfonation is considered to be supplied by 3'-phosphoadenosine,5'- phosphosulfate (PAPS) produced by SxtO (adenylyl sulfate kinase) (Mihali *et al.*, 2009, Cullen *et al.*, 2018a). A class of C-13 acetylated STX variants termed *Lyngbya wollei* toxins (LWTX1-6) are exclusively found in *M. wollei*. These analogues are predicted to be produced by the activity of SxtSUL, SxtDIOX, and a novel enzyme, SxtACT (a putative acetyltransferase), which solely found in *Lyngbya wollei* cluster (Mihali *et al.*, 2011).

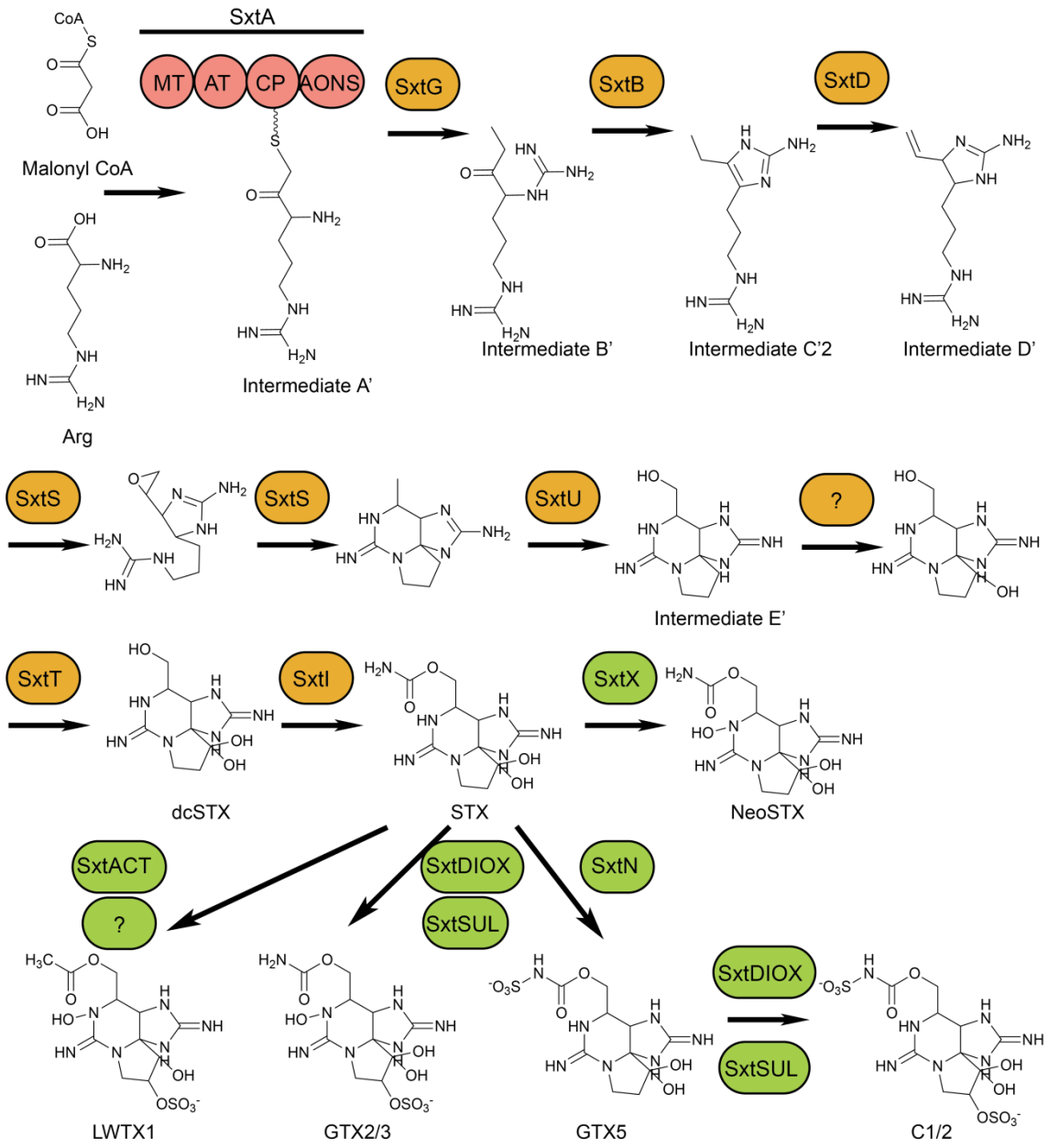


Figure 1.5 Saxitoxin biosynthesis. The biosynthesis of saxitoxin is initiated the formation of Intermediate-A' by a PKS-like enzyme SxtA (red). The Intermediate-A' undergo a series of additional reactions carried out by modifying enzymes (orange), forming saxitoxin (STX). The diversification of saxitoxin then catalysed by STX tailoring enzymes (green), which resulted in the formation of various STX analogues. MT, methyltransferase; AT, acyltransferase, CP, carrier protein; AONS, 8-amino-7-oxononanoate synthase. Figure reproduced from Cullen *et* al. (Cullen *et al.*, 2018b)

1.3 Context of the thesis and objectives

Cyanobacteria are a known rich source of bioactive specialised metabolites, many of which are produced by the megasynthetase enzymes NRPSs and/or PKSs. Unfortunately, low compound yield, unculturability, and lack of genetic tools to engineer the producers limit access to these compounds. Recently, heterologous expression of specific enzymes or the entire biosynthetic pathways has gained interest as an alternative approach for biochemical characterisation and compound production. Additionally, the modularity structure of NRPS/PKS enzymes, along with the step-wise enzyme process allows opportunities for pathway engineering. The use of synthetic biology to manipulate the native biosynthesis has been applied in numerous pathways (Kries *et al.*, 2014, Evans *et al.*, 2011, Crüsemann *et al.*, 2013, Taori *et al.*, 2009, Liu *et al.*, 2011, Liu *et al.*, 2017, Awan *et al.*, 2017, Yamanaka *et al.*, 2014), yet such strategy is underexplored in the cyanobacterial pathway. This dissertation aimed to utilise an *E. coli* system to express, characterise, and modify cyanobacterial natural products pathways.

The research described in Chapter 2 focused on a simple lyngbyatoxin (*ltx*) pathway originating from the filamentous cyanobacterium *Moorea producens*, which was previously expressed in *E. coli*. Recently, studies have demonstrated the inter-modular and inter-domain dynamic of NRPS enzymes. The pathway was characterised in detail through the expression and *in vitro* substrate specificity assessment of the bimodular NRPS, LtxA. The *in vitro* NRPS activity was then correlated with metabolite production, providing a representation of the pathway activity. Further, modification of the NRPS binding pocket was performed in an attempt to alter the production of lyngbyatoxin variants. It was anticipated that the study would provide insights into the development of LTX analogue production.

Heterologous expression in *E. coli* has proven to be a useful tool for the biosynthetic characterisation and production of cyanobacterial metabolites, such as lyngbyatoxin and microcystin. This approach could therefore be employed for a more complicated pathway, such as saxitoxin (STX). Chapters 3 and 4 describe the development of STX expression in *E. coli*. In Chapter 3, the expression of the initial biosynthetic enzyme SxtA was reported. An inducible

Pm promoter was inserted in front of the *sxtA* gene and co-expressed with the broad substrate PPTase Sfp. The successful expression of SxtA, was extended to the complete *sxt* pathway expression described in Chapter 4. The use of pathway engineering to direct production of neoSTX was performed. However, unlike in Chapter 2, where the initial enzyme and culture condition was modified, the presence of Sxt tailoring enzymes was altered to direct production to one particular analogue.

Overall, this dissertation aimed to develop a production platform for cyanobacterial specialised metabolites for academic and industrial applications. The implementation of synthetic biology to modulate substrate selectivity and metabolite production was explored. This allowed the characterisation of NRPS/PKS enzymes and pathways, providing insights into strategies for the production of specialised metabolites and their analogues.

Chapter 2. Tailoring enzyme stringency masks the multispecificity of a lyngbyatoxin nonribosomal peptide synthetase

This chapter is being prepared for publication.

Authors: **Angela H. Soeriyadi**, Sarah E. Ongley, Jan-Christoph Kehr, Russel Pickford, Elke Dittmann, Brett A. Neilan

Current Status: In preparation

The Extent to which the publication is your own:

I have developed the expression and purification of full-length LtxA. Together with Jan-Christoph Kehr (University of Potsdam, Germany), I investigated the *in vitro* substrate activation assay of wild-type and mutated LtxA. I developed the method and analysed the production of lyngbyatoxin analogues and their intermediate in *E. coli*. The site-directed mutagenesis of LtxA A-domain was performed during my BSc (Hons) thesis. However, the expression of the recombinant protein, and characterisation of the *in vitro* and *in vivo* pathway activity were performed for the completion of this chapter as a part of my PhD dissertation. I prepared the first draft of the manuscript, and subsequent revisions based on advice from other authors.

Abstract

Biochemical characterisation of nonribosomal peptide synthetase pathways is aimed at understanding the essentiality of enzymatic mechanisms to inform any potential industrial and pharmaceutical applications. Here, the marine cyanobacterial NRPS pathway for lyngbyatoxin biosynthesis, which was previously expressed in *E. coli*, was investigated. It was discovered that *in vitro*, full-length LtxA has multispecificity towards L-Val related amino acids, specifically L-Ile and L-Leu. This multispecificity was demonstrated to enable the production of previously undescribed methylated LTX and its intermediates in *E. coli*, which was confirmed by labelled precursor feeding with L-Leu-5,5,5-d_3 and L-Ile-d_{10}. The *in vitro* substrate activation of full-length LtxA was shown to be relatively congruent with its *in vivo* activity. Efforts to alter analogue production, performed via site-directed mutagenesis of the LtxA A-domain binding pocket, resulted in a slight alteration or abolishment of enzyme activity. Semi-quantitative assessment of the LTX pathway activity revealed the substrate specificity of the tailoring enzymes LtxB and LtxC, which surprisingly dictated the metabolite flux in LTX biosynthesis. It is recommended that future modification of LTX biosynthesis should include engineering of the tailoring enzyme to extend its substrate specificity.

2.1 Introduction

Cyanobacteria from marine environments are a rich source of chemically diverse and biologically significant metabolites. Several of these metabolites have bioactivities relevant to pharmaceutical applications such as anticancer (dolastatin, apratoxin, symplostatin, largazole) (Luesch *et al.*, 2001a, Luesch *et al.*, 2002b, Mooberry *et al.*, 2003, Taori *et al.*, 2009, Luesch *et al.*, 2001b), antifungal (hassallidin) (Neuhof *et al.*, 2005, Vestola *et al.*, 2014), antimalarial (gallinamide A) (Linington *et al.*, 2009), immunosuppressant (microcolin A) (Zhang *et al.*, 1997), and antiviral (cyanovirin-N, scytovirin) (Boyd *et al.*, 1997, Bokesch *et al.*, 2003) agents. Many of these compounds are synthesised by nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs). NRPSs are multimodular enzymes consisting of three core

domains catalysing peptide elongation and associated modifications. Due to the modular nature of NRPSs, studies have employed a synthetic biology approach to accelerate the discovery and to tailor the production of both known and novel natural product derived medicine (Alexander *et al.*, 2010, Doekel *et al.*, 2008, Nguyen *et al.*, 2010, Butz *et al.*, 2008, Mootz *et al.*, 2002). This approach has aided the investigation of NRPS enzymology, structural biology, and the expression of chimeric pathways to understand the biosynthetic mechanism, providing insights into drug design strategies (Drake *et al.*, 2016, Reimer *et al.*, 2016, Marahiel, 2016, Stachelhaus *et al.*, 1999, Meyer *et al.*, 2016, Bloudoff *et al.*, 2016, Weissman, 2015, Gulick, 2016, Miller *et al.*, 2016, Miller & Gulick, 2016, Belshaw *et al.*, 1999).

One NRPS protein engineering approach targets the adenylation domain (A-domain) substrate specificity (Kries *et al.*, 2014, Schneider *et al.*, 1998, Eppelmann *et al.*, 2002, Villiers & Hollfelder, 2011, Zhang *et al.*, 2013) which would modify the peptide backbone. The A-domain is responsible for the selection and activation of the substrate amino acid, and is mainly dictated by ten amino acid residues within the A-domain binding pocket referred to as the 'specificity-conferring code' (Stachelhaus *et al.*, 1999). The specificity code was later refined by Rausch *et al.* to include other active sites as well (Rausch *et al.*, 2005). Alteration of this amino acid sequence can potentially result in different substrate specificity. Examples of successful implementation of this A-domain engineering include an increase of andrimid NRPS promiscuity (Evans *et al.*, 2011), the conversion of relaxed specificity into a stringent specificity in fusaricidin (Han *et al.*, 2012), luminmide (Bian *et al.*, 2015), and anabaenopeptin (Kaljunen *et al.*, 2015) biosynthesis, as well as the production of novel daptomycin analogues (Thirlway *et al.*, 2012). Other studies have also used natural evolution to guide A-domain re-engineering (Crüsemann *et al.*, 2013, Vobruba *et al.*, 2017, Meyer *et al.*, 2016). The recent advent of synthetic biology methods allows for the feasible manipulation of NRPSs enzyme activity (Zhang *et al.*, 2013, Steiniger *et al.*, 2017, Gao *et al.*, 2018, Yan *et al.*, 2018b, Lundy *et al.*, 2018, Cleto & Lu, 2017). Yet, pathway engineering of cyanobacterial peptides remains under-explored despite their biological potential.

Lyngbyatoxins A-C are dermatoxic cyclic peptides (Cardellina *et al.*, 1979, Aimi *et al.*, 1990) isolated from a Hawaiian cyanobacterium *Moorea producens*. Lyngbyatoxins possess structural similarities to the teleocidin family produced by Actinobacteria. In particular, lyngbyatoxin A (LTX) is identical to one of the isomers, teleocidin A-1 (Sakai *et al.*, 1986). In addition to lyngbyatoxin A-C, several natural analogues have been previously identified, including 12-epi-lyngbyatoxin A (Jiang *et al.*, 2014b), 2-oxo-3(*R*)-hydroxy-lyngbyatoxin A (Jiang *et al.*, 2014a), 2-oxo-3(*R*)-hydroxy-13-N-desmethyl-lyngbyatoxin A (Jiang *et al.*, 2014a), and 2,3-seco-2,3-dioxo-lyngbyatoxin A (Youssef *et al.*, 2015). The biochemical pathways of these analogues have not been described. Nevertheless, this illustrates the chemical diversity of microbial nonribosomal peptides.

Members of the lyngbyatoxin family possess tumour-promoter activity via activation of protein kinase C (PKC), which regulates cell proliferation and differentiation (Basu *et al.*, 1992, Fujiki *et al.*, 1981). The indolactam-V (ILV) ring of LTX and teleocidin is essential for this activity (Basu *et al.*, 1992, Cardellina *et al.*, 1979, Youssef *et al.*, 2015). PKC modulators have been exploited as therapeutic targets for several diseases including various cancers and neurodegenerative disorders (Mochly-Rosen *et al.*, 2012, Nelson & Alkon, 2009, Nakagawa, 2012). In the case of indolactam-derived molecules, they have also been employed as molecular tools to study disease progression (Meseguer *et al.*, 2000, Masuda *et al.*, 2002, Yanagita *et al.*, 2008). Due to the biological implications, various studies have performed the total synthesis of ILV (Xu *et al.*, 2011, Fine Nathel *et al.*, 2014, Noji *et al.*, 2015), ILV analogues (Meseguer *et al.*, 2000, Ma *et al.*, 2002, Masuda *et al.*, 2002, Yanagita *et al.*, 2008, Fine Nathel *et al.*, 2014, Nakagawa *et al.*, 2001) and LTX (Fine Nathel *et al.*, 2014), as well as other structurally related alkaloids, including pendolmycin (Fine Nathel *et al.*, 2014) and teleocidin A-2 (Fine Nathel *et al.*, 2014). Synthetic chemistry approaches, however, are often limited by the complexity of chemical reactions and low overall yields. Genetic engineering of the biosynthetic pathway potentially overcomes these problems. Therefore, the present study explores the potential of a synthetic biology approach to produce different variants of indolactam alkaloids.

The lyngbyatoxin biosynthetic pathway, encoded by the *ltx* gene cluster was identified from a *Moorea producens* genomic DNA library (Edwards & Gerwick, 2004). This cluster spans 11.3 kbp and consists of four unidirectionally transcribed open reading frames (*ltx*A-D, see Figure 2.1). Lyngbyatoxin biosynthesis is initiated by the formation of dipeptide NMVT from L-valine and L-tryptophan catalysed by a bimodular NRPS LtxA, which is released by an unusual NADPH-dependent reductase (Read & Walsh, 2007). The dipeptide then undergoes cyclisation performed by a P450-dependent monooxygenase (LtxB), and further geranylation by a reverse prenyltransferase (LtxC) forming lyngbyatoxin A (Edwards & Gerwick, 2004). The tailoring enzymes LtxB and LtxC have shown relaxed substrate specificity *in vitro* for valyl group analogues (Huynh *et al.*, 2010). The conversion of lyngbyatoxin A to lyngbyatoxin B and C has not yet been biochemically characterised. However, LtxD, an oxidase/reductase, is predicted to catalyse an Alder-Ene reaction for the formation of these two analogues (Videau *et al.*, 2016).

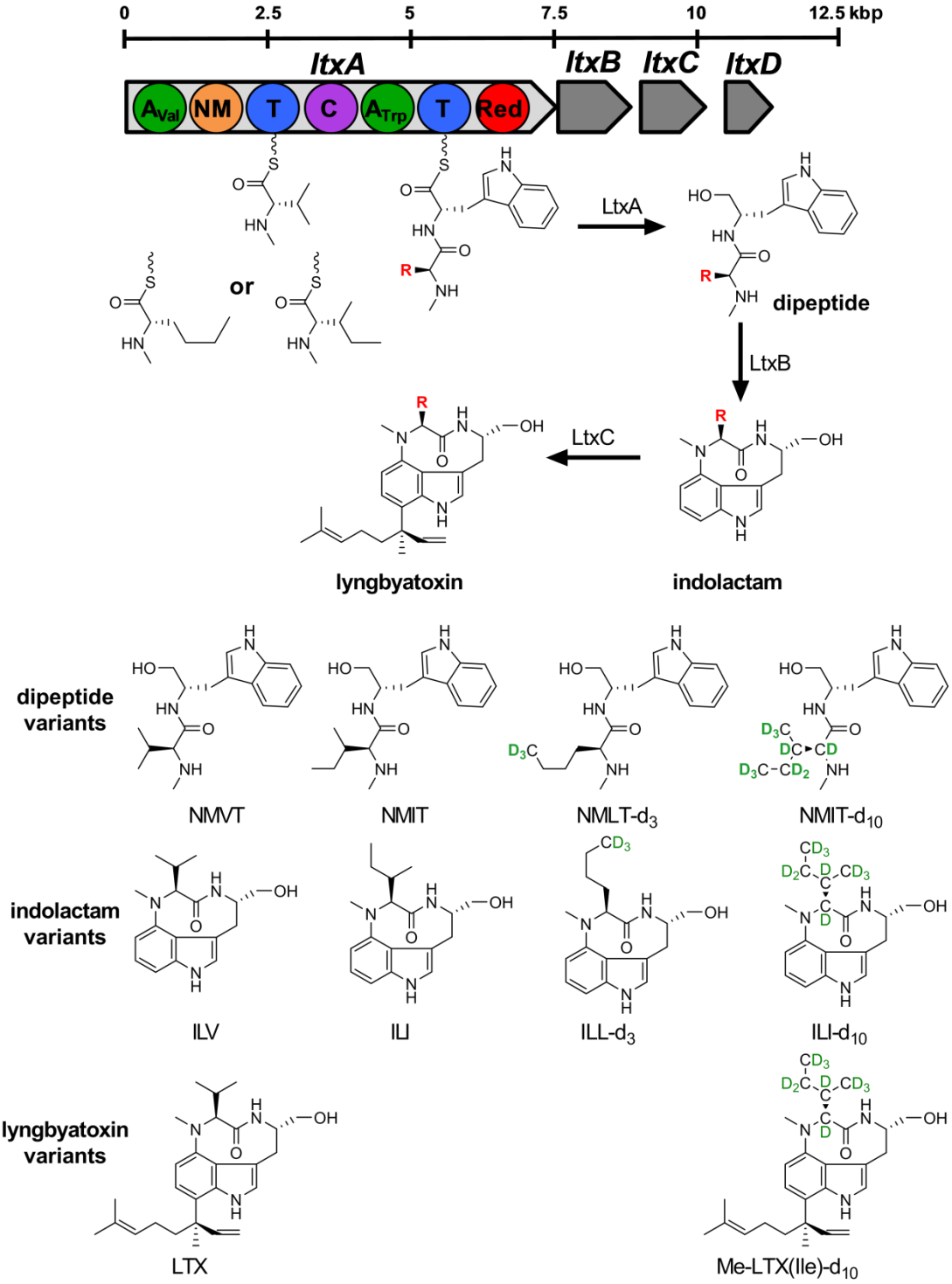


Figure 2.1 The biosynthesis of lyngbyatoxin A and its analogues observed in this study. The genetic arrangement of the lyngbyatoxin biosynthetic pathway (11.3 kbp). The deuterated variants are observed in the precursor feeding studies. NMVT, N-methyl-L-valyl-L-tryptophanol; NMIT, N-methyl-L-isoleucyl-L-tryptophanol; NMLT, N-methyl-L-leucyl-L-tryptophanol; ILV, Indolactam V; ILI, Indolactam I; ILL, Indolactam L; LTX, Lyngbyatoxin A; Me-LTX(Ile), methyl-lyngbyatoxin A (Ile).

The first attempt to heterologously express the LTX biosynthetic pathway was unsuccessful, partly due to the differences in GC content and codon bias of *Streptomyces* compared to *M. producens* (Jones *et al.*, 2012). Further attempts utilising *E. coli* (Ongley *et al.*, 2013b) and *Anabaena* sp. (Videau *et al.*, 2016) as heterologous hosts led to successful LTX production. In *E. coli*, a high titre of LTX (25.6 mg/L) and ILV (150 mg/L) was achieved. This high level of production and the relatively simple pathway prompted us to further characterise the LTX biosynthesis and apply a synthetic biology approach to tailor analogue production in *E. coli*. The present study aimed to investigate the activity of the full-length bimodular LtxA and the effect of mutations within the first A-domain on its *in vitro* activity. In addition, product formation was analysed to understand the pathway activity and how the alteration of NRPS activity translated *in vivo*.

2.2 Methods

2.2.1 Bacterial strains and culture conditions

E. coli strains were maintained on lysogeny broth (LB) medium at 37°C. Either LB, Terrific Broth (TB) medium or M9 minimal medium were used for LTX pathway expression. Expression in M9 minimal medium was supplemented with 100 mg/L L-leucine to satisfy auxotrophic requirements, except when cultures were supplemented with L-leucine-5,5,5-d_3. Where appropriate, media were supplemented with antibiotics (chloramphenicol 20 μg/mL, ampicillin 50 μg/mL, kanamycin 50 μg/mL) for plasmid selection and maintenance. The preparation of recombineering-proficient *E. coli* was performed as previously described (Chai *et al.*, 2012) with electroporation at 1400 V (Bio-Rad Gene Pulser). Standard plasmid electroporation followed a similar procedure, but without the addition of arabinose. *E. coli* strains and plasmids used in this study are listed in Table 2.1.

Table 2.1 Strains and plasmids used in this study.

Strains and plasmids	Genotype or relevant features	Ref
***E. coli* strains**		
GB05	(HS996, Δ*recET*, Δ*ybcC*). Endogenous *recET* locus and the DLP12 prophage *ybcC*, which encodes a putative exonuclease similar to Redα were deleted.	(Fu *et al.*, 2010)
GB05-red	(GB05, *ybcC*::*araC*-BAD-αβγA) Lambda *red* operon and *recA* under P_{BAD} promoter were inserted at the *ybcC* locus.	(Fu *et al.*, 2012)
GB05-red-gyrA462	(GB05-red, *gyrA*462), GyrA mutation of R462C.	(Wang *et al.*, 2014)
DB3.1	(*gyrA*462) Used for propagating plasmids containing *ccdB*.	Invitrogen
GB05-MtaA	(GB05, *mtaA-genta*). A derivative of GB05 with MtaA PPTase from *Stigmatella aurantiaca* DW4/3-1 integrated via transposition.	(Fu *et al.*, 2012)
EPI300-T1^R	Contains mutant *trfA* gene under the control of L-arabinose inducible promoter, pBAD, for induction of pCC1FOS to high copy number. Used for plasmid propagation.	Epicentre
Plasmids		
pCFOS-Ptet-*ltx*	pCC1FOS (Epicentre) containing the lyngbyatoxin gene cluster with *genta^R-tetR-Ptet$_O$* inserted upstream of *ltxA*.	(Ongley *et al.*, 2013b)
pET28b::*ltx*A	pET28b vector with *ltxA* gene under T7 promoter.	This study
pET28b::genta-Ptet-*ltx*A	pET28b vector containing gentamycin resistance gene and *ltxA* gene under Ptet promoter.	This study
pET28::amp-Ptet-*ltx*A	pET28b vector containing ampicillin resistance gene and *ltxA* gene under Ptet promoter.	This study
pCFOS-Ptet-His*ltxAltxBCD*	pCFOS-Ptet-*ltx* with 6xHis was inserted at the N-terminus of LtxA, with *genta^R-tetR-Ptet$_O$* replaced with *amp^R-tetR-Ptet$_O$*.	This study
pCFOS-Ptet-ltx239-ccdBamp	pCFOS-Ptet-*ltx* with *ccdB-amp^R* inserted at position 239 of LtxA-A1 binding pocket (LtxA amino acid 170)	This study
pCFOS-Ptet-ltx299-ccdBamp	pCFOS-Ptet-*ltx* with *ccdB-amp^R* inserted at position 299 of LtxA-A1 binding pocket (LtxA amino acid 232)	This study
pCFOS -Ptet-*ltx*Y239M	pCFOS-Ptet-*ltx* with Y239M mutation in LtxA-A1 (LtxA amino acid 170).	This study

pCFOS -Ptet-*ltx*Y239F	pCFOS-Ptet-*ltx* with Y239F mutation in LtxA-A1 (LtxA amino acid 170).	This study
pCFOS -Ptet-*ltx*W299L	pCFOS-Ptet-*ltx* with W299L mutation in LtxA-A1 (LtxA amino acid 232).	This study
pCFOS -Ptet-*ltx*W299C	pCFOS-Ptet-*ltx* with W299C mutation in LtxA-A1 (LtxA amino acid 232).	This study
pCFOS -Ptet-*ltx*Y239M W299L	pCFOS-Ptet-*ltx* with Y239M and W299L mutation in LtxA-A1 (LtxA amino acid 170 and 232).	This study
pCFOS -Ptet-*ltx*Y239F W299L	pCFOS-Ptet-*ltx* with Y239F and W299L mutation in LtxA-A1 (LtxA amino acid 170 and 232).	This study
pCFOS -Ptet-*ltx*Y239M W299C	pCFOS-Ptet-*ltx* with Y239M and W299C mutation in LtxA-A1 (LtxA amino acid 170 and 232).	This study
pCFOS -Ptet-*ltx*Y239F W299C	pCFOS-Ptet-*ltx* with Y239F and W299C mutation in LtxA-A1 (LtxA amino acid 170 and 232).	This study

2.2.2 *In silico* analysis of A-domain amino acid binding pockets

The 'specificity-conferring code' of the LtxA adenylation domain was analysed using NRPSpredictor2 (http://nrps.informatik.uni-tuebingen.de) (Röttig *et al.*, 2011). An analysis to predict the amino acid residue alterations to the substrate preference was also performed. The tool predicts A-domain substrates using Transductive Support Vector Machines (TSVMs). The output prediction classified A-domains on four hierarchical levels ranging from gross physiochemical properties of the substrate (hydrophobic-aromatic, hydrophobic-aliphatic and hydrophilic) to single amino acid substrates.

2.2.3 Expression plasmid constructions

2.2.3.1 Construction of expression plasmid pET28b::*ltxA*

Construction of the *ltxA* expression plasmid containing N-terminal polyhistidine tags (6xHis-tag) was performed using linear-circular homologous recombineering (LCHR) (Fu *et al.*, 2012). The linear fragment of pET28b with homology arms to *ltxA* was amplified by using Velocity DNA polymerase (NEB) and primers pETHisTagLtxA_F and pETHisTagLtxA_R. The manufacturer's protocol was followed with an annealing temperature of 59°C and an extension

time of 4 min. Fragments were subjected to DpnI restriction enzyme digestion (NEB) before purification by agarose gel electrophoresis (Gel DNA Recovery Kit, ZymoResearch). Purified fragments were electroporated into recombineering-proficient *E. coli* GB05-red harbouring pCFOS-Ptet-*ltx*. Cells were selected using kanamycin selection to ensure the construction of pET28::*ltxA*. The colonies were screened by colony PCR using universal primers T7_prom and T7_term as well as *ltxA* specific primer LTX299seqR and sequenced using the BigDye Terminator v3.1 cycle sequencing kit and an ABI 3730 DNA analyser sequencer (Applied Biosystems) at the Ramaciotti Centre for Genomics. Overall plasmid integrity was confirmed by diagnostic restriction enzyme digestion with EcoRI. Primer oligonucleotide sequences are listed in Appendix Table A 1.1.

2.2.3.2 Promoter exchange of pET28b::*ltxA*

Overexpression of LtxA using T7 promoter in pET28b::*ltxA* did not yield soluble protein (see Appendix Table A 1.2 and Figure A 1.1 for examples of protein expression attempts). To address this, the T7 promoter was exchanged into a Ptet promoter using Gibson assembly (NEB). Briefly, the *genta*^{R}-Ptet region was amplified from pCFOS::Ptet-*ltx* using primers genta-Ptet5_pET28HA and genta-Ptet3_pET28HA. The expression plasmid pET28b::*ltxA* constructed previously was digested with BglII and NcoI (NEB). Both fragments were purified via agarose gel electrophoresis (DNA Recovery Kit, ZymoResearch) before being assembled using NEBuilder HiFi DNA Assembly (NEB), forming pET28b::genta-Ptet-*ltxA*. This was then transformed into chemically competent GB05 cells and grown under kanamycin and gentamycin selection. Clones were screened by pET_T7promUp and genta-Ptet3_pET28HA then verified by DNA sequencing.

2.2.3.3 Construction of LtxA-A1 binding-pocket mutants

Site-directed mutagenesis of residues 239 and 299 within the first LtxA A-domain binding pocket was performed using two-step selection-counterselection recombineering as previously performed by Wang *et al* (Wang *et al.*, 2014), with minor modifications (Appendix Figure A 1.3, panel A). The first recombineering involved the insertion of the *ccdB-amp*^{R} cassette from

p15A-ccdB-amp kindly provided by Prof. Rolf Müller (Saarland University, Germany). The cassette was amplified by PCR using Velocity DNA polymerase and primers 239ccdBamp5 and 239ccdBamp3, or 299ccdBamp5 and 299ccdBamp3 (Appendix Table A 1.1). The linear fragments were gel purified (Gel DNA recovery kit, ZymoResearch) then electroporated into recombineering-proficient GB05-red-gyrA462 harbouring pCFOS-Ptet-ltx to replace the 239 and 299 codons with the *ccdB-amp*^{R} cassette. The *E. coli* strain GB05-red-gyrA462 carries a mutation in the *gyrA* gene, conferring its resistance to CcdB toxicity. Cells were grown under chloramphenicol and ampicillin selection. Positive recombinants were confirmed by colony PCR using LTX239seqF and LTX299seqR as well as diagnostic restriction enzyme digestion with EcoRV. The functionality of CcdB counterselection was confirmed by transformation into GB05, with few or no resulting colonies.

In the second recombineering step, *ccdB-amp*^{R} cassettes were replaced with the desired amino acid change encoded within the single-stranded DNA repair oligonucleotides (Appendix Table A 1.1). Briefly, 50 ng of the appropriate repair oligonucleotide (Appendix Table A 1.1) were transformed by electroporation into the recombineering-proficient GB05-red-gyrA462 harbouring pCFOS-Ptet-ltx239-ccdBamp or pCFOS-Ptet-ltx299-ccdBamp. Cells were incubated for 1.5 h at 37°C. An addition of 1 mL of LB with chloramphenicol was given and the cells were incubated further for 16 h. The plasmid was extracted from transformants grown overnight and subsequently electroporated into GB05, and then grown under chloramphenicol selection. Colonies were screened by PCR using primers LTX239seqF and LTX299seqR. To confirm the mutation introduced, plasmids were sequenced using the BigDye Terminator v3.1 cycle sequencing kit and an ABI 3730 DNA analyser sequencer (Applied Biosystems). Overall plasmid integrity was confirmed by diagnostic restriction enzyme digestion with EcoRV.

2.2.3.4 Insertion of polyhistidine tags to pCFOS::Ptet-*ltx* wild-type and mutant LtxA

Replacement of the *genta*^{R} cassette with the *amp*^{R} cassette was performed in pET28b::genta-Ptet-*ltxA* to mediate N-terminal polyhistidine tag (6xHis-tag) insertion into the *ltx* expression

fosmid. Briefly, the *amp*^{R} (*bla*) cassette was amplified from pGEM-T Easy using Velocity DNA Polymerase (NEB) using primers amp5_sgrAI and amp3_bglII. The PCR product was gel purified (Gel DNA recovery kit, ZymoResearch) prior to digestion by BglII and SgrAI restriction enzymes (NEB). The vector plasmid, pET28b::genta-Ptet-*ltxA* was cut with the same enzymes. The fragments were gel purified (Gel DNA recovery kit, ZymoResearch) and ligated with T4 DNA ligase (NEB). The ligated plasmid (pET28::amp-Ptet-*ltxA*) was transformed into chemically competent *E. coli* GB05 and was grown under kanamycin and ampicillin selection. Successful ligations were screened by colony PCR using pET_T7promUp and LtxAR primers and verified by DNA sequencing.

Insertion of N-terminal polyhistidine tags to both wild-type and mutant LtxA in the expression vector pCFOS-Ptet-*ltx* was performed using linear-circular homologous recombineering. Firstly, linear fragments comprising *amp*^{R}-Ptet-6xHis-*ltxA* with appropriate homology arms were amplified from pET28::*amp*-Ptet-*ltxA* using primers amp5-HApfos and LtxAR. The manufacturer's protocol was followed with an annealing temperature of 62°C and 1 min extension time. Purified fragments were electroporated into recombineering-proficient *E. coli* GB05-red harbouring either wild-type pCFOS-Ptet-*ltx* or its mutated variants. Cells were selected under chloramphenicol and ampicillin selection to ensure the construction of pCFOS-Ptet-His*ltxAltxBCD*. The colonies were screened by colony PCR using primers amp5-HApfos and *ltxA* specific primer LTX299seqR, and sequenced to verify 6xHis-tag insertion. Overall plasmid integrity was confirmed by restriction enzyme digestion with EcoRI.

2.2.4 Characterisation of bimodular NRPS LtxA activity *in vitro*

2.2.4.1 Functionality of recombinant LTX pathways via HPLC analysis

To investigate the functionality of the recombinant protein (N-HisLtxA), heterologous expression of the lyngbyatoxin pathway was performed. Plasmid pCFOS-Ptet-His*ltxAltxBCD* was electroporated into *E. coli* GB05-MtaA and grown overnight in LB at 30°C with shaking at 950 rpm (Eppendorf Thermomixer). Cultures were inoculated 1:100 into 50 mL Terrific Broth in 250 mL Erlenmeyer flasks stoppered with a cotton wool bung. Cultures were incubated at

30°C with shaking at 200 rpm to an OD_{600} of ~0.4, and then the temperature was reduced to 18°C. At an OD_{600} of ~0.5, expression was induced by the addition of tetracycline at 0.5 mg/L. Cells were incubated at 18°C with shaking at 200 rpm for three days before the addition of 2% (v/v) Amberlite XAD-7 polymeric resin, then incubated for a further 16-18 h. Cells and resin were harvested by centrifugation, and metabolites were extracted with 50 mL methanol shaking at 18°C for 1.5 h. The extract was clarified by centrifugation, passed through a Whatman No. 1 filter, evaporated to dryness and redissolved in 1 mL methanol.

Extracts were analysed using a Dionex Ultimate HPLC (ThermoFisher Scientific) for the production of lyngbyatoxin pathway metabolites. Extracts (50 µL) from induced cultures were compared to an uninduced negative control. Separation was achieved using a Luna C18(2) (250 mm x 10 mm, 5 µm) (Phenomenex) column with a linear gradient of water + 0.1% (v/v) trifluoroacetic acid (solvent A) to acetonitrile (solvent B) at a flow rate of 4.5 mL/min. The gradient was initiated with 5% B and increased to 95% B within 30 min, followed by 95% B for 10 min and decreased to 5% B within 5 min and held there for a further 15 min. UV absorption was monitored at λ= 280 nm. The lyngbyatoxin pathway metabolites eluted with retention times of 11.45 min, 18.10 min, and 30.20 min for NMVT, ILV, and LTX, respectively (see Figure A 1.4).

2.2.4.2 Protein expression

Wild-type and single mutants of LtxA protein were expressed in *E. coli* GB05-MtaA in LB medium. Expression was induced at OD ~0.5 and cultures were incubated for 18 h at 23°C with shaking at 220 rpm. Cells were lysed by sonication in lysis buffer (100 mM Tris-HCl [pH 7.5], 150 mM NaCl, 20 mM imidazole, 10% [v/v] glycerol) and LtxA proteins were purified by bench purification using Ni-NTA resin (Qiagen) with buffer A (100 mM Tris-HCl [pH7.5], 150 mM NaCl, 20 mM imidazole) and buffer B (100 mM Tris-HCl [pH7.5], 150 mM NaCl, 250 mM imidazole). For ATP-$^{32}PP_i$ Exchange assay, the protein sample was desalted using a PD-10 column into assay reaction buffer of 100 mM Tris-HCl (pH 7.5), 150 mM NaCl, 10% (v/v) glycerol. For the *in vitro* product formation assay, the protein was desalted and

concentrated into 50 mM HEPES (pH 7.5), 150 mM NaCl using Amicon Ultra-15 centrifugal filter unit 100 kDa MWCO (Millipore) and purified further through size exclusion chromatography using AKTA purifier system fitted with Superdex 200 10/300 GL column with a running buffer of 50 mM HEPES (pH 7.5), 150 mM NaCl, 10% (v/v) glycerol. The eluates were concentrated using Amicon Ultra-15 centrifugal filter unit 100 kDa MWCO (Millipore) with the same buffer. Protein concentration was determined by using protein assay kit (Bio-Rad) and stored at -80°C.

2.2.4.3 ATP-$^{32}PP_i$ exchange Assay

The ATP-$^{32}PP_i$ exchange Assay was used to demonstrate the substrate specificity of LtxA enzymes. The assays were performed in collaboration with Prof. Roderich Sussmuth (Technical University Berlin, Germany). A total of ten amino acids were tested for activity: Val, Ile, Tyr, Leu, Orn, Trp, Phe, Ala, Gly, phenylglycine (Phg). In brief, the reaction was performed in 2 mL tubes containing 100 µL of ATP-$^{32}PP_i$ mix (0.1 mM Na_4PPi, 5 mM ATP, 10 mM $MgCl_2$, $^{32}PP_i$ equivalent to 200,000 counts/min; PerkinElmer), 20 µL of 100 mM amino acid stock solution (dissolved in 1 M Tris-HCl [pH 8.0] or water). The reaction was started with the addition of 100 µL enzyme solution and was incubated for 10 min at 25°C. The reaction was quenched through the addition of 1 mL of stop solution (1.4% Norit A, 5.4% [v/v] 70% perchloric acid, 168.5 mM Na_4PP_i). Samples were then washed with MilliQ under vacuum, and the charcoal-bound radioactivity was measured in a Hidex 300 SL scintillation counter (Hidex) according to Cherenkov using Mikrowin 2000 software (Sislab).

2.2.5 Characterisation of lyngbyatoxin pathway activity *in vivo*

2.2.5.1 Heterologous expression of the lyngbyatoxin pathway in *E. coli*

Heterologous expression of the lyngbyatoxin pathway product was performed in *E. coli* GB05-MtaA, containing the myxobacterial PPTase from *Stigmatella aurantiaca* (Fu *et al.*, 2012). *E. coli* GB05-MtaA harbouring either the empty pCFOS or the pCFOS::Ptet-*ltx* variants were grown in LB supplemented with 10 µg/mL chloramphenicol, incubated overnight at 30°C under agitation (200 rpm). The cultures were then subsequently diluted 1:50 into M9 minimal medium

supplemented with 100 mg/L L-leucine and were grown overnight under the same conditions. Cultures were inoculated at 1:50 dilution into fresh 10 mL M9 minimal medium in 50 mL Erlenmeyer flasks sealed with a cotton bung. For isotopic labelling feeding experiments, cultures were supplemented with L-leucine-5,5,5-d_3 (200 mg/L) and L-isoleucine-d_{10} (200 mg/L) (Cambridge Isotope Laboratories). The cultures were grown at 30°C, while shaking 200 rpm to an OD_{600} of ~0.4 then transferred into 18°C, shaking at 200 rpm for 15-30 min to reach an OD_{600} of ~0.5. Tetracycline was added to a final concentration of 0.5 µg/mL to induce the expression of lyngbyatoxin NRPS. Cultures were grown at 18°C, shaking 200 rpm for three days. Before harvesting, 2% (v/v) of Amberlite XAD-7 polymeric resin was added into the culture, followed by a further 18 h incubation allowing the adsorption of the extracellular metabolites. Cells and resin were harvested by centrifugation for 30 min at 3,250 *g* and stored at -20°C.

Intracellular and extracellular LTX pathway metabolites were extracted with 10 mL methanol by vortex mixing and shaking (200 rpm) at 18°C for 2 h. Caffeine (Interlink Scientific Services) served as an internal standard and was added to the culture prior to extraction to a final concentration of 500 ng/mL. The extracts were clarified by centrifugation for 5 min at 3,250 *g* and filtered through Whatman No. 1 membrane before being evaporated to dryness by using a Syncore Dryer (Buchi). Extracts were solubilised in 1 mL methanol and diluted 1:3 (v/v) for HPLC-HRMS/MS analysis. Dipeptide Gly-Phe (Sigma) was used as an internal standard for the LCMS analysis and was added to the LC-MS extracts at a concentration of 500 ng/mL.

2.2.5.2 HPLC-ESI-HRMS/MS analysis of lyngbyatoxin biosynthetic pathway products

Pathway products were identified via HPLC-HR-MS/MS analysis. The analysis was performed on 20 µL of the sample using a Dionex U3000 UHPLC interfaced to a Q-Exactive Plus (ThermoFisher Scientific) via a heated electrospray interface. The separation was achieved by Waters BEH C18 column (50 mm x 2.1 mm, 1.9 µm) (Waters) using a linear gradient of 0.1%

(v/v) formic acid in water (A2) and acetonitrile (B2) with a flow rate of 0.4 mL/min. The gradient was initiated with 100% A for 3 min, followed by a linear gradient of 0-100% B over 7 min and held at 100% for 2 min before re-equilibration to 0% B. The lyngbyatoxin pathway metabolites eluted with retention times of 5.50 min, 7.05 min, and 9.51 min for NMVT, ILV, and LTX, respectively (see Figure A 1.7 for the chromatograms for each compound).

HPLC was coupled to a Q-Exactive Plus (ThermoFisher Scientific) where samples were ionised using positive ionisation mode. Mass spectra were acquired in centroid mode with a range of 100-700 *m/z* at a resolution of R=70,000. Ion trap mass spectrometry (ITMS) was performed on the inclusions list; masses correspond to the predicted metabolites with their $[M+H]^+$ as the precursor ions with width 1.4 *m/z*. Predicted ions included were 290.19, 304.20, 314.26, 307.22, 318.22, 328.28, 321.24, 302.19, 316.20, 326.26, 319.22, 438.31, 452.33, 462.39, 455.35. The normalised collision energy was set to 35%. All data were processed using Xcalibur software (Thermo Scientific), and statistical analysis was performed with Prism 7 (GraphPad).

2.3 Results and discussion

Efficient heterologous expression of LTX pathway in *E. coli* has been previously reported to produce high titres of lyngbyatoxin and its indolactam intermediates (Ongley *et al.*, 2013b). In the present study, both *in vitro* substrate preference and *in vivo* product formation were analysed to provide insight into the enzymology of LtxA, as well as the stringency of the lyngbyatoxin pathway. To achieve this, site-directed mutagenesis of the LtxA A-domain specificity conferring code was performed, followed by *in vitro* analysis of precursor activation, which was then correlated with the capacity of the modified enzyme to produce LTX analogues via *in vivo* product formation.

2.3.1 LtxA is multispecific *in vitro*

LtxA (275 kDa) is a bimodular NRPS with A1-T1-C-A2-T2-R domain organisation. The activity of NRPS enzymes has been shown to be dependent on inter-domain interactions (Reimer *et al.*, 2016, Drake *et al.*, 2016). In the current study, the full-length holo-LtxA was

expressed and characterised in order to preserve the protein structure and its dynamic interactions. Attempts to express the stand-alone LtxA containing N-terminal 6xHis-tag (pET28::*ltxA*) in *E. coli* yielded no or only insoluble protein (see Appendix Table A 1.2 and Figure A 1.1 for example of protein expression attempts). Several parameters such as induction conditions and buffer composition were tested, however these attempts failed to yield any soluble protein. Promoter exchange of T7 promoter with Ptet promoter did not result in any significant improvement (see appendix Table A 1.2 and Figure A 1.2). In the present study, successful expression of soluble and active LtxA was only observed with the co-expression of the tailoring enzymes LtxB, LtxC and LtxD.

To assist protein purification, a 6xHis-tag was inserted using recombineering into the N-terminus of LtxA expressed in the fosmid pCFOS-Ptet-*ltx*. The inserted tag was shown to have no detrimental effect on LtxA enzymatic activity, as consistent production of pathway products *in vivo* was still achieved (see Appendix Figure A 1.4 for HPLC analysis results). Protein purification through affinity chromatography yielded LtxA that was co-purified with other proteins, including LtxB (52 kDa) and the chaperone GroES/EL. Both proteins were co-eluted after separation using size exclusion chromatography (see Appendix Figure A 1.5 for SDS PAGE gel of purified LtxA). The importance of LtxB co-expression, along with its co-purification with LtxA on gel filtration, suggested that LtxA and LtxB form a protein complex, which could affect the NRPS activity.

Various studies have employed protein purification of NRPS A-domains with subsequent ATP-$^{32}PP_i$ exchange assays to assess their catalytic activity (Meyer *et al.*, 2016, Zhang *et al.*, 2018, Otten *et al.*, 2007). The A-domain is responsible for the substrate activation through an ATP-dependent reaction, forming amino acyl-adenylate and releasing PP_i (Turgay *et al.*, 1992). The presence of $^{32}PP_i$ in the assay drives the reverse reaction forming ^{32}P-ATP, which correlates to the relative activation of each amino acid (Linne & Marahiel, 2004). However, recently it was shown that the condensation domain (C-domain) could influence the substrate activation (Meyer *et al.*, 2016, Degen *et al.*, 2019). Based on this knowledge, the substrate activation activity of the full-length LtxA was assesed.

The substrate specificity of wild-type and the altered LtxA adenylation domains predicted by NRPSpredictor2 are tabulated in Table 2.2. For ATP-$^{32}PP_i$ exchange assays. A total of ten amino acids were assayed, including the cognate substrate and several other structurally related amino acids. The assay was performed with three technical replicates. Bioinformatic analysis predicted L-Val to be the substrate of LtxA-A1 with L-Phe being the cognate substrate for LtxA-A2 (Table 2.2), however previous *in vivo* analysis has indicated that L-Trp is the prefered substrate of LtxA-A2 (Edwards & Gerwick, 2004, Ongley *et al.*, 2013b). The wild-type enzyme LtxA was shown to possess activity above the background level with L-Trp, L-Phe, L-Try, L-Val, L-Ile, and L-Leu, with the highest for L-Trp (Figure 2.2). It is important to note that the activities observed in the assay were substrate activation by both A-domains within LtxA. LtxA was shown to activate aromatic amino acids L-Phe and L-Tyr, but with much less activity (10-20%) in comparison to L-Trp. Thus, L-Trp was indicated to be the preferred substrate of the second A-domain of LtxA (LtxA-A2), which is consistent with *in vivo* production of NMVT. As such, the level of enzymatic activity on L-Trp was used as a normalisation factor to calculate the relative activity for other amino acids. The *in vitro* analysis of LtxA-A2 substrate was in agreement with the previous *in vivo* findings (Edwards & Gerwick, 2004, Ongley *et al.*, 2013b), though it was interesting to note that this substrate is different to the bioinformatic prediction by NRPSpredictor2 (Röttig *et al.*, 2011, Rausch *et al.*, 2005) (Table 2.2), although the amino acid class remained the same. A similar observation was reported with specific yet unreliable small clusters substrate prediction of TycA using NRPSpredictor2 (Villiers & Hollfelder, 2011).

Table 2.1 Substrate predictions of LtxA binding pocket studied.

A-domain	A-domain binding pocket		Predicted A-domain specificities	
	239	299	Small clusters	Nearest neighbour
LtxA-A1				
Wild-type	Y	W	Val, Leu, Ile, Abu, Iva	Val
Y239M	M	W	Val, Leu, Ile, Abu, Iva	Ile
Y239F	F	W	Val, Leu, Ile, Abu, Iva	Phg
W299L	Y	L	-	Ile
W299C	Y	C	Gly, Ala	Val
Y239M W299L	M	L	Val, Leu, Ile, Abu, Iva	Ile
Y239F W299L	F	L	-	Ile
Y239M W299C	M	C	Gly, Ala	Ile
Y239F W299C	F	C	Gly, Ala	Ala
LtxA-A2				
Wild-type			Phe, Trp	Phe

-, indicates that no small clusters prediction were available for the binding pocket. Phg, phenylglycine; Abu, 2-amino-butyric acid; Iva, isovaline.

The native dipeptide formation, as well as the bioinformatic analysis, indicated L-Val to be the substrate of LtxA-A1. The assay result for aliphatic amino acids was consistent with this, as wild-type LtxA strongly activates L-Val. Interestingly, activation above the background was also observed for L-Leu, and to a lesser extent, L-Ile (Figure 2.2). Although the substrate specificity of the two LtxA A-domains are not distinguishable, the activity for the aliphatic amino acids was suspected to be mainly performed by LtxA-A1.

The first A-domain of LtxA was predicted to activate aliphatic amino acids. From the data presented at Figure 2.2, the detected activities towards L-Val are 3 times lower than towards L-Trp. As the assay was performed on the full-length LtxA protein, the activity of the second A-domain is also represented. The lower L-Val activation relative to L-Trp might be due to the substrate multispecificity of LtxA-A1 in comparison to the stringency of LtxA-A2. Nevertheless, the results indicated that WT LtxA-A1 possesses multispecific activity for L-Val related amino acids *in vitro*.

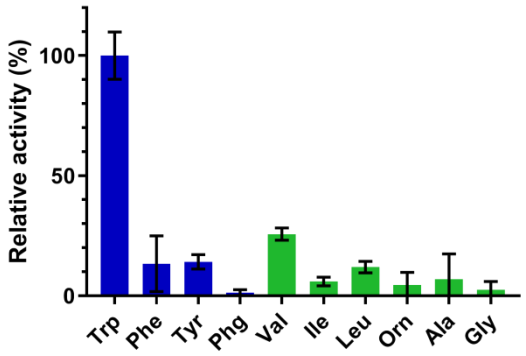


Figure 2.2 *In vitro* activity of wild-type LtxA. Amino acid activation profile of LtxA determined using ATP-$^{32}PP_i$ exchange assay. The activity towards L-tryptophan was normalised to 100%, as this is the substrate of the unaltered second A-domain of LtxA (LtxA-A2). In blue are the level of activation of aromatic amino acids putatively due to the activity of LtxA-A2 and in green are the level of activation for aliphatic amino acid putatively due to the activity of LtxA-A1. Error bars indicate SD from the normalised mean of three technical replicates (n = 3). Negative values were interpreted as no activation (=0). Phg, phenylglycine

2.3.2 Multispecificity of Ltx enzymes results in novel methylated metabolites

The heterologous expression of the lyngbyatoxin pathway in *E. coli* yielded the pathway intermediates NMVT, ILV and LTX (Ongley *et al.*, 2013b). The multispecificity of LtxA *in vitro* prompted us to determine whether the substrate activation for various amino acids resulted in the formation of pathway analogues. Fermentation of *E. coli* strains harbouring pCFOS::Ptet-*ltx* in Terrific Broth showed the formation of previously undescribed LTX intermediates NM(I/L)T and IL(I/L) (data not shown). These intermediates have an additional methyl group, putatively due to the incorporation of either L-Ile or L-Leu instead of L-Val by the first LtxA module. However, the level of production was far lower than that of LTX intermediates.

Precursor-directed biosynthesis was used to exploit the relaxed specificity of MicC to produce novel micacocidin derivatives (Kreutzer *et al.*, 2014). In the present study, *E. coli* cultures were subjected to amino acid supplementation, in an attempt to increase LTX analogue production. Fermentations of recombinant *E. coli* strains in M9 minimal media supplemented

with L-Ile and L-Leu significantly increased the production of methylated LTX intermediates (Appendix Figure A 1.6). The production level was shown to be dose-dependent, indicating that the amino acid availability influences the production of LTX intermediates. Unfortunately, the yields of the new methylated metabolites were insufficient for structural confirmation by NMR. Instead, a method using labelled precursor feeding was developed to investigate metabolite structure.

The strain used for LTX pathway expression was *E. coli* GB05-MtaA, which contains a myxobacterial PPTase necessary for the post-translational modification of apo-LtxA. *E.coli* GB05-MtaA is a derivative of DH10β (Fu *et al.*, 2008, Fu *et al.*, 2012), which is a leucine auxotroph. Hence, no endogenous L-Leu would be present under normal growth conditions and any unlabelled Me-LTX and intermediates can be identified as L-Ile derivatives. Based on this characteristic, a labelled precursor feeding method using L-Ile-d_{10} and L-Leu-5,5,5-d_3 was developed. An amino acid concentration of 200 mg/L was used, as this concentration was shown to maximise the yield of methylated products (Appendix Figure A 1.6).

Methyl lyngbyatoxin biosynthesis is hypothesised to commence with the formation of N-methyl-L-isoleucyl-L-tryptophanol (NMIT) and N-methyl-L-leucyl-L-tryptophanol (NMLT) (for structure of the novel analogues and the position of methyl group, refer to Figure 2.1). As L-Ile and L-Leu are isomers, different deuterated precursors were used in this study, namely L-Ile-d_{10} and L-Leu-5,5,5-d_3. The products formed were analysed using HPLC-ESI-HRMS/MS. Incorporation of endogenous and deuterated L-Ile and L-Leu precursors resulted in dipeptides with different molecular ions (NMIT/NMLT *m/z* = 318, NMIT-d_{10} *m/z* = 328, and NMLT-d_3 *m/z* = 321) with different deuterium locations (Figure 2.1, dipeptide variants row).The biosynthesis of indolactams and subsequent lyngbyatoxin variants (Figure 2.1, indolactam and lyngbyatoxin variants rows) from the labelled dipeptides also contained a different number of deuteriums which could be identified and verified by MS/MS analysis. This method allows the elucidation of the methyl group within Me-LTX, which presumably originates from the two amino acid isomers, isoleucine and leucine.

Labelled precursor feeding successfully confirmed the structural identity of Me-LTX and its intermediates. LC-MS analysis revealed the presence of new molecular ions that were present in label-fed cultures and absent in non-labelled control cultures. The new ions eluted at the same retention time as Me-LTX and intermediates observed in the non-labelled control sample (Appendix Figure A 1.7). Further HPLC-ESI-HRMS/MS analysis confirmed the chemical structure of the metabolites. Comparison of their structural features to the Val-derived product and non-labelled Ile/Leu-derived product are shown in Figure 2.3, Figure 2.4, and Figure 2.5 for dipeptides, indolactams, and lyngbyatoxins respectively. Both L-Ile-d_{10} and L-Leu-5,5,5-d_3 were shown to be utilised by LtxA, forming the dipeptide N-methyl-L-isoleucyl-L-tryptophanol-d_{10} (NMIT-d_{10}) and N-methyl-L-leucyl-L-tryptophanol-d_3 (NMLT-d_3), respectively (Figure 2.3 C and D). Also, the formation of indolactam I-d_{10} (ILI-d_{10}) and indolactam L-d_3 (ILL-d_3) was observed (Figure 2.4 C and D). Interestingly, only methyl-lyngbyatoxin (isoleucine)-d_{10} (Me-LTX(Ile)-$d_{10,}$ Figure 2.5 C) was detected but not methyl-lyngbyatoxin (leucine)-d_3 (Me-LTX(Leu)-d_3). This phenomenon is likely due to the low level of the L-Leu-derived LTX precursor, ILL-d_3.

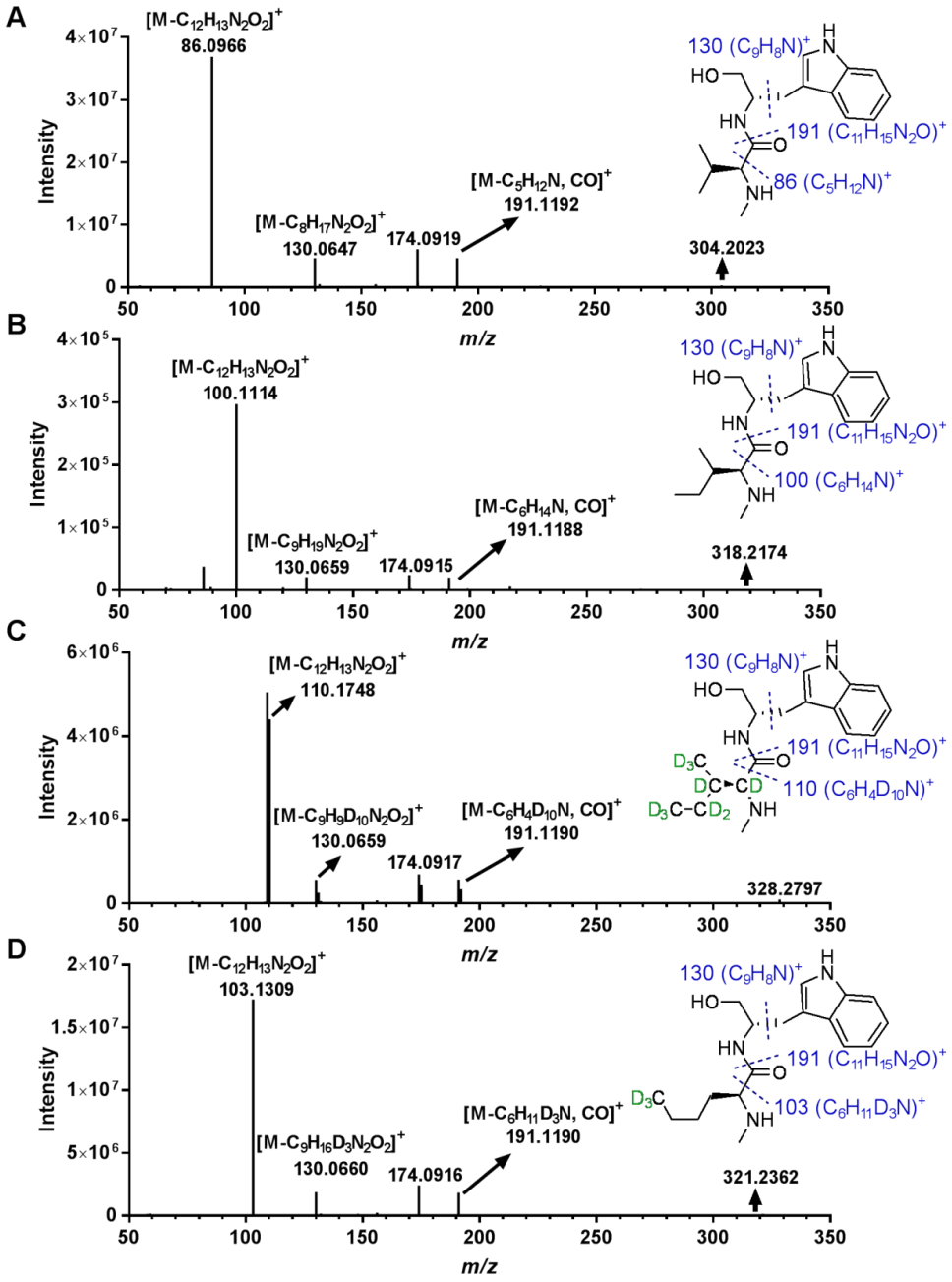


Figure 2.3 HPLC-ESI-HRMS/MS analysis of dipeptide analogues from induced *E. coli* GB05-MtaA pCFOS -Ptet-*ltx* cultured in M9 minimal medium. Cultures were supplemented with L-leucine-5,5,5-d_3 and L-isoleucine-d_{10}, each to a concentration of 200 mg/L. Fragmentation profile of **C**) N-methyl-L-isoleucyl-L-tryptophanol-d_{10} (NMIT-d_{10}) and **D**) N-methyl-L-leucyl-L-tryptophanol-d_3 (NMLT-d_3) confirmed the incorporation of L-Ile and L-Leu in comparison to L-Val in **A**) N-methyl-L-valyl-L-tryptophanol (NMVT). Shown in **B**) is the fragmentation pattern of N-methyl-L-isoleucyl-L-tryptophanol (NMIT). The tandem mass spectrometry was performed on $[M+H]^+$ as precursor ions for NMVT (*m/z* = 304.20), NMIT (*m/z* = 318.22), NMIT-d_{10} (*m/z* = 328.28), and NMLT-d_3 (*m/z* = 321.24), with width of 1.4 *m/z*. The precursor ions are annotated in each spectra.

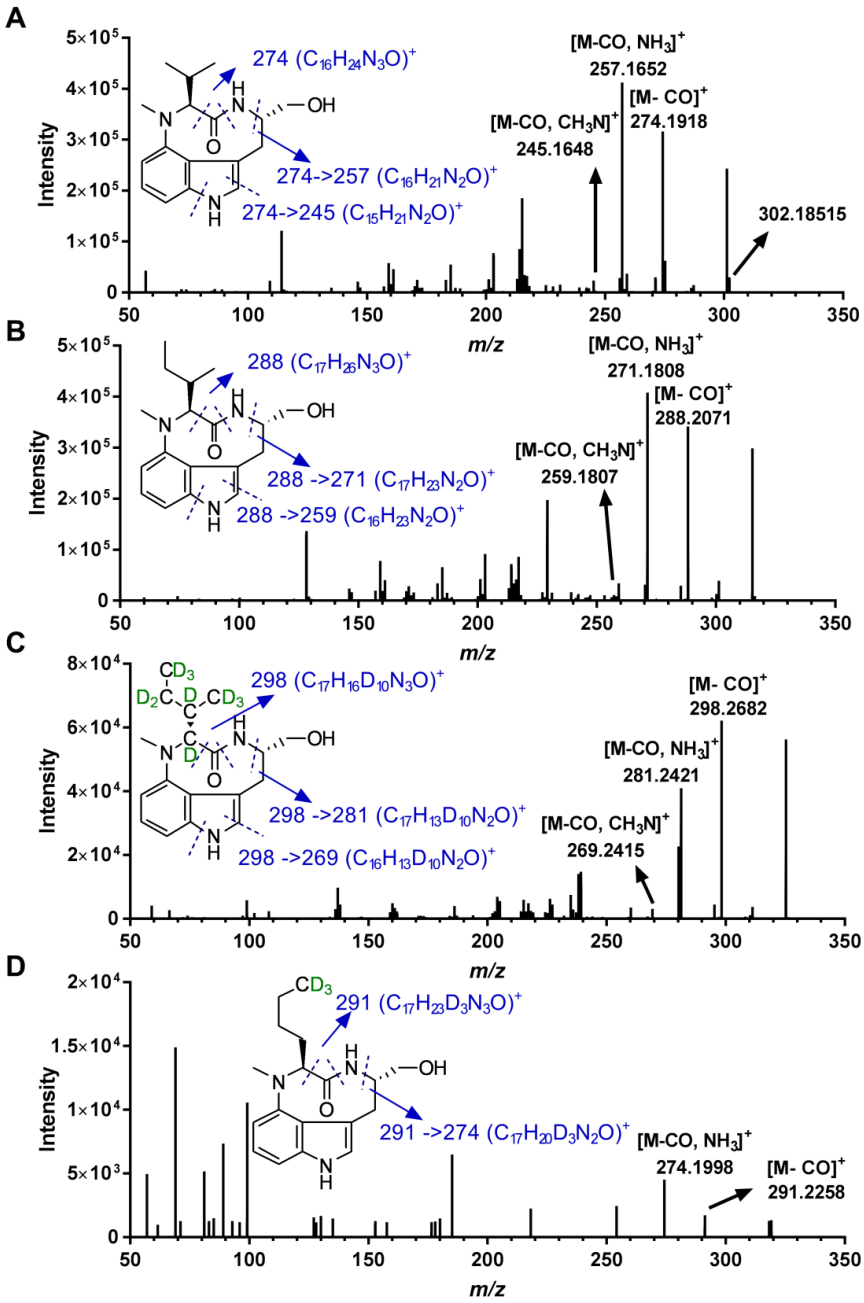


Figure 2.4 HPLC-ESI-HRMS/MS analysis of indolactam analogues from induced *E. coli* GB05-MtaA pCFOS -Ptet-*ltx* cultured in M9 minimal medium. Cultures were supplemented with L-leucine-5,5,5-d_3 and L-isoleucine-d_{10}, each to a concentration of 200 mg/L. Fragmentation profile of **C**) indolactam I -d_{10} (ILI-d_{10}) and **D**) indolactam L-d_3 (ILL-d_3) confirmed the incorporation of L-Ile and L-Leu in comparison to L-Val in **A**) indolactam V (ILV). Shown in **B**) is the fragmentation pattern of indolactam I (ILI). The tandem mass spectrometry was performed on $[M+H]^+$ as precursor ions for ILV (*m/z* = 302.19), ILI ((*m/z* = 316.20), ILI-d_{10} (*m/z* =326.26), and ILL-d_3 (*m/z* = 319.22), with width of 1.4 *m/z*.

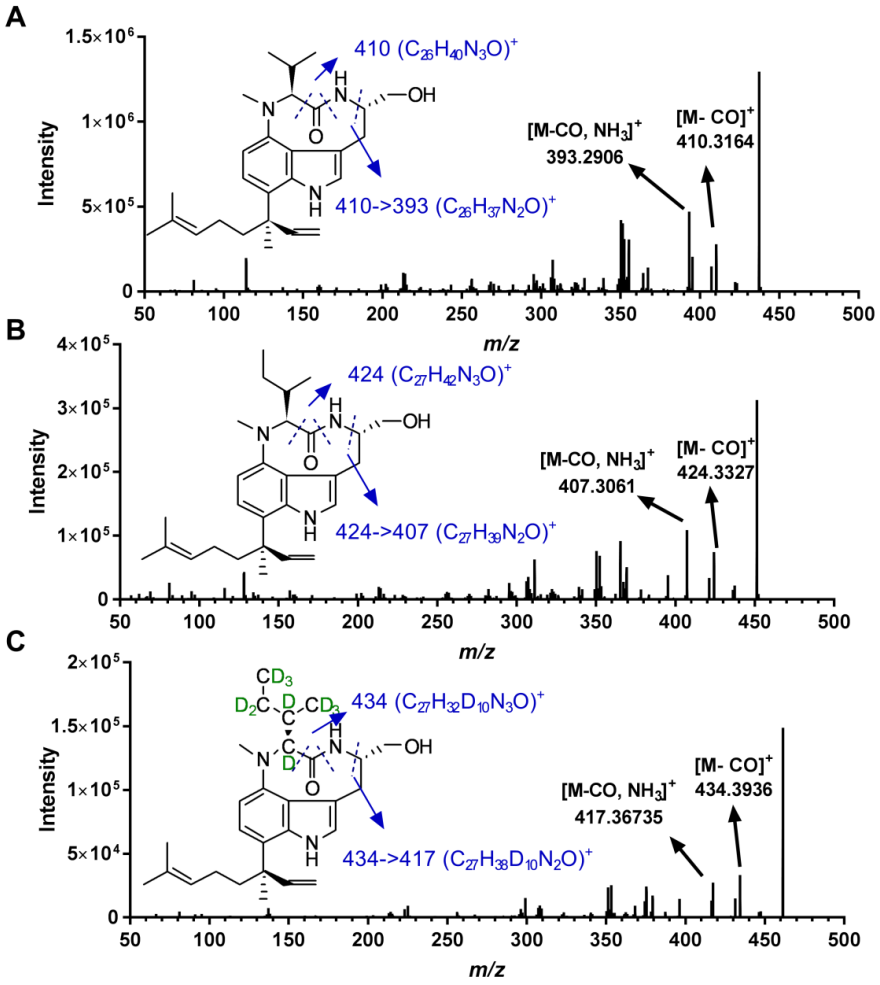


Figure 2.5 HPLC-ESI-HRMS/MS analysis of lyngbyatoxin analogues from induced *E. coli* GB05-MtaA pCFOS-Ptet-*ltx* cultured in M9 minimal medium. Cultures were supplemented with L-leucine-5,5,5-d_3 and L-isoleucine-d_{10} supplemented to a concentration of 200 mg/L. Fragmentation profile of **C**) methyl-lyngbyatoxin (isoleucine)-d_{10} (Me-LTX(Ile)-d_{10}) confirmed the incorporation of L-Ile in comparison to L-Val in **A**) lyngbyatoxin A (LTX). Shown in **B** is the fragmentation pattern of methyl-lyngbyatoxin (isoleucine) (Me-LTX(Ile)). The tandem mass spectrometry was performed on $[M+H]^+$ as precursor ions for LTX (*m/z* = 438.31), Me-LTX(Ile) (*m/z* = 452.33), and (*m/z* = 462.39), with width of 1.4 *m/z*.

2.3.3 Site-directed mutagenesis of LtxA-A1 marginally affects substrate preferences *in vitro*

The native lyngbyatoxin biosynthetic pathway heterologously expressed in *E. coli* GB05-MtaA produces different analogues due to the promiscuity of the LtxA NRPS. Mutagenesis of the first A-domain, LtxA-A1 was performed in an attempt to reprogram the biosynthetic flux. *In silico* prediction using NRPSpredictor2 (Röttig *et al.*, 2011) was used to predict the effect of amino acid substitution on the 'specificity-conferring code' (Stachelhaus *et al.*, 1999) (Appendix Table A 1.3). The software was unable to provide a substrate prediction for more than 40% of single substitutions. When it did make a prediction, minimal influence of a single substitution onto LtxA-A1 activity was predicted, with most of the mutations not altering their preference for L-Val. Substitutions at positions 239 and 299 were predicted to have the most significant influence on substrate specificity. Thus, seamless site-directed mutagenesis of amino acids in LtxA-A1 was targeted at position 239 and 299 (Table 2.2). The mutation was introduced using selection/counterselection λ-Red/ET recombineering. Four substitutions (Y → F, M, N, P) were chosen for position 239, and two (W → C, L) for 299. Preliminary fermentations of *E. coli* strains harbouring the mutated pCFOS-Ptet-*ltx* showed a deleterious effect of Y239N and Y239P substitution. As such, these mutated constructs were not investigated further.

The effect of the introduced point mutations towards the activity of LtxA A-domain was assessed by ATP-$^{32}PP_i$ exchange assay performed in three technical replicates. The mutated proteins were expressed, purified and assayed as previously described with the wild-type protein. Based on the bioinformatics tool NRPSpredictor2 (Rausch *et al.*, 2005, Röttig *et al.*, 2011), the mutations were anticipated to alter the enzyme specificity towards Ile (Y239M and W299L) and phenylglycine (Phg) (Y239F), while W299C was predicted to retain the specificity for Val (Table 2.2).

In the *in vitro* assay, the enzyme activation of L-Trp, the substrate of LtxA-A2, was assumed to remain stable between LtxA mutants. Thus, the degree of substrate activation of

other amino acids was normalised to the activity for L-Trp. It was shown that point mutations Y239F and W299C did not significantly alter the substrate specificities in comparison to the wild-type LtxA (Figure 2.6 B and D). However, the point mutation Y239M narrowed the substrate specificity towards Val and Leu. In particular, the enzyme showed a significant increase in preference towards Leu (one-way ANOVA, multiple comparison $p < 0.05$, Figure 2.5 A) from 0.5-fold compared to Val in wild-type, to up to 1.5-fold compared to Val in Y239M. The point mutation W299L increases the overall preferences towards aliphatic amino acids, especially for Val (one-way ANOVA, multiple comparison $p < 0.05$, Figure 2.5 C). There was only a slight increase preference towards Val in comparison to Ile and Leu.

Overall, it was shown that the single amino acid substitution did not significantly modify LtxA substrate preferences. The accuracy of the bioinformatic tool was limited in predicting the amino acid class. As valine, isoleucine, and leucine are closely related, it is understandable that the mutated LtxA activity did not precisely reflect the bioinformatic prediction. In the present study, the adenylation activity was purposely assessed using the full-length NRPS to accommodate inter-domain and inter-module interactions, not accounting for the C-domain gatekeeping mechanism.

The present investigation suggests that LtxA NRPS has a degree of substrate stringency. NRPS gene clusters have been shown to contain MbtH-like proteins (MLP), a class of small ~70 amino acid proteins homologs of MbtH contained in the mycobactin gene cluster in *Mycobacterium tuberculosis* (Quadri *et al.*, 1998a, Baltz, 2011). Several studies have shown that MbtH-like genes are associated with the NRPS A-domain (Baltz, 2011), and in some cases they affect their protein expression and solubility (Zolova & Garneau-Tsodikova, 2012, Felnagle *et al.*, 2010) as well as their activity and substrate specificity (Zolova & Garneau-Tsodikova, 2012, Zhang *et al.*, 2010, Felnagle *et al.*, 2010, Davidsen *et al.*, 2013, Boll *et al.*, 2011). Structural studies have shown the MLP does not directly affect substrate activation (Herbst *et al.*, 2013), rather it changes the dynamics within the enzyme which subsequently influences its activity (Tarry *et al.*, 2017). Nevertheless, the actual role of MLP in substrate activation remains unknown. The lyngbyatoxin biosynthetic cluster contains a domain

homologous to MLP encoded at the N-terminal region of *ltxB* (Edwards & Gerwick, 2004). Although it is unusual for MLP to be fused with a tailoring enzyme, the presence and effect of LtxB MLP-like domain on substrate specificity have not yet been determined. In the present study, LtxB was co-expressed and co-purified with LtxA. Thus, the possible influence of a MLP-like domain within LtxB which was included in the assay cannot be discounted. Future studies could investigate the effect of LtxB MLP-like domain gene inactivation followed by *in vivo* activity assays. In addition, co-expression and purification of either MLP-like domain or LtxB with LtxA followed by *in vitro* activity assay and structural studies might also highlight its influence.

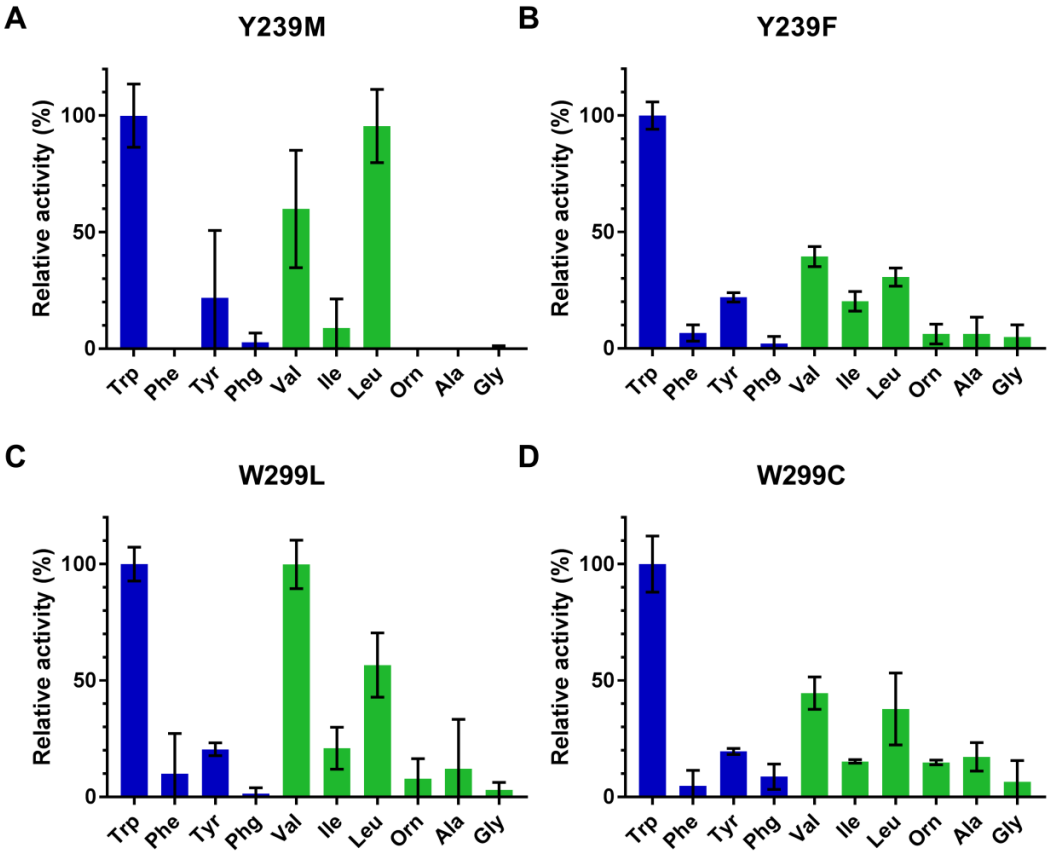


Figure 2.6 Amino acid activation profile of LtxA variants determined using ATP-^{32}PP$_i$ exchange assay. The activity towards L-tryptophan was normalised to 100%, as this is the substrate of the unaltered second A-domain of LtxA. In blue are the levels of activation of aromatic amino acids putatively due to the activity of LtxA-A2 and in green are the levels of activation towards aliphatic amino acid putatively due to the activity of LtxA-A1. Error bars indicate SD from the normalised mean of three technical replicates (n = 3). Negative values were interpreted as no activation (=0).

2.3.4 LtxB activity determines metabolite flux *in vivo*

A precursor feeding experiment using L-Ile-d_{10} and L-Leu-5,5,5-d_3 was performed to semi-quantitatively assess LTX pathway activity *in vivo* under excess availability of L-Ile and L-Leu. The production of different LTX analogues and intermediates by the wild-type and recombinant pathways were analysed by HPLC-ESI-HRMS/MS. Ion abundance values were normalised against internal standards of caffeine and Gly-Phe, as well as OD_{600} of the *E. coli* culture at the time the XAD resin was added. The supplementation of 200 mg/L isotope labelled precursor was shown to be sufficient to minimise the bias effect of endogenous L-Ile. This was evident by the significantly higher amount of L-Ile-d_{10} derived compounds (~30 times) relative to the endogenous L-Ile derived ones (Appendix Figure A 1.8). Cell cultures that were grown in media containing non-labelled L-Ile and L-Leu (200 mg/L), as well as cell cultures harbouring empty pCC1FOS (Epicentre) were used as negative controls.

The *in vivo* pathway activity of the wild-type LtxA is relatively consistent with the *in vitro* result. Both L-Ile and L-Leu were utilised by LtxA efficiently to form a high level of dipeptide NMIT-d_{10} and NMLT-d_3, respectively (Appendix Figure A 1.8). This is congruent with the *in vitro* activation of both amino acid precursors. The level of NMVT detected was highest as expected. This was followed by a 2.5 times lower level of NMIT-d_{10} and 4 times lower level of NMLT-d_3 (Appendix Figure A 1.8). However, NMIT-d_{10} and NMLT-d_3 were not efficiently processed by LtxB. This was shown by the significant lower levels of both ILI-d_{10} and ILL-d_3 (Student's t-test, unpaired, $p < 0.05$, Appendix Figure 1.8), 5- and 90-fold respectively, in comparison to ILV (Appendix Figure A 1.8). Notably, the amount of ILI-d_3 was 17 times higher than ILL (Student's t-test, unpaired, $p < 0.05$, Appendix Figure A 1.8) despite the similar level of dipeptide formation (NMLT was 1.5 times more than NMIT-d_{10}). In addition, a 25-fold lower level of Me-LTX (Ile)-d_{10} compared to LTX was observed, with no detectable formation of Me-LTX (Leu).

The activity of the wild-type pathway and the mutated pathway were compared. In addition to the single mutant pathways, eight double mutants were constructed combining the

substitution in the two different sites (239 and 299) (Table 2.2). A comparison of the total LTX pathway products of single mutants and double mutants with respect to wild-type levels is shown in Table 2.3. Except for mutant W299C, all mutants have reduced levels of dipeptide and indolactam production (Student's t-test, unpaired, $p < 0.05$, Table 2.3, Appendix Figure A 1.8 and Figure A 1.9). This observation is similar to other attempts to engineer A-domains which have led to poor yields of new compound derivatives (Williams, 2013, Bian *et al.*, 2015). The most deleterious effect was observed for the double mutants, especially with mutant Y239M/W299C that showed less than 1% of wild-type product. The results suggested that position 239 is critical for the catalytic activity of LtxA, with any change to the amino acid drastically altering activity. A previous study has also shown that the amino acid at position 239 is highly variable and linked to substrate specificity (Stachelhaus *et al.*, 1999). A more significant impact from tyrosine (Y) into methionine (M) substitution in Y239M rather than into phenylalanine (F) in Y239F (Table 2.3) was observed, possibly due to the more conserved amino acid substitution in the latter.

Table 2.3 Total production of LTX pathway from M9 minimal medium fermentations.

LtxA	Total Pathway products		
	Dipeptide	Indolactam	Lyngbyatoxins
WT	1.00	1.00	1.00
Y239M	0.14	0.09	0.20
Y239F	0.54	0.62	0.65*
W299L	0.50	0.68	0.82*
W299C	0.64*	0.64*	0.78*
Y239M/W299L	0.04	0.02	0.06
Y239F/W299L	0.38	0.40	0.58
Y239M/W299C	0.00†	0.00†	0.02
Y239F/W299C	0.19	0.11	0.16

Total products refer to the combined amounts of valine, isoleucine and leucine derived compounds. Dipeptide refers to NMVT, NMIT and NMLT; Indolactam refers to ILV ILI and ILL; Lyngbyatoxins refer to LTX, Me-LTX (Ile) and Me-LTX (Leu). Proportions were calculated as fractions compared to wild-type (WT) pathway; * denotes not significantly different from WT levels (Student's t-test, unpaired, $p > 0.05$); † Compounds were detected albeit significantly lower abundances compared to the WT levels; Experiments were performed in 3 biological replicates.

The *in vivo* activity can be analysed by an evaluation of the analogue yield proportion for all three pathway intermediates. Interestingly, although the amount of the compounds detected varied amongst biological replicates, the proportion of analogues remained consistent (see Appendix, Table A 1.4). This trend did not occur when the production level was low.

Similar to the wild-type pathway, the production pattern of the different analogues aligned with *in vitro* adenylation activity. For example, mutants Y239F and W299C showed similar production profiles compared to the wild-type, whereas W299L showed slight stringency towards Val (Figure 2.7). However, Y239M have reduced both Ile and Leu containing products in comparison to wild-type LtxA with a slight increase in preference towards Val (Figure 2.7). This observation was contradictory to the *in vitro* data where Y239M had increased substrate activation towards Leu (Figure 2.6A). This observation could be caused by limitations of the ATP- PP_i exchange assays, as other factors besides amino acid substrate activation affect the formation of dipeptide and subsequent downstream processing. This includes the efficiency of PCP domain thioester bond formation with the activated amino acids or a C-domain gatekeeping role (Meyer *et al.*, 2016, Belshaw *et al.*, 1999), both in donor and acceptor sites (Belshaw *et al.*, 1999).

Similar to mutants W299L and Y239M, pathway metabolite production by the LtxA double mutants showed decreased Ile incorporation. This coincided with an increase in production of Val-derived pathway product, without changes in the level of Leu incorporation, except for mutant Y239F/W299C. Interestingly, mutant Y239F/W299C produced an increased level of NMLT (Leu-containing dipeptide), and a decreased NMVT, but downstream pathway production did not follow the same pattern. The level of ILL and ILV observed remained similar in comparison to the wild-type with no detectable level of Leu-derived Me-LTX. This might be due to the substantially higher level of LtxB's preferred substrate of NMVT, which prevented the formation of Leu-derived indolactam (ILL). The relatively well conserved ratio of isoleucine and leucine indolactam and lyngbyatoxin products shows that the alteration of

dipeptide formation was not enough to overcome the control posed by tailoring domains, such as LtxB.

The findings illustrate that the tailoring enzymes LtxB, and to an extent LtxC, have a substrate specificity which dictates the metabolite flux in LTX biosynthesis. The trend observed is consistent with a previous study which showed that the tailoring enzymes LtxB and LtxC possess relaxed substrate specificity *in vitro*, with relative substrate preference ranked as L-valyl > L-isoleucyl > L-leucyl intermediates (Huynh *et al.*, 2010). The natural biosynthesis, therefore, has LtxB-dependent activity despite precursor similarity. This was particularly evident in the LtxA mutants when comparing *in vitro* substrate activation with *in vivo* product formation.

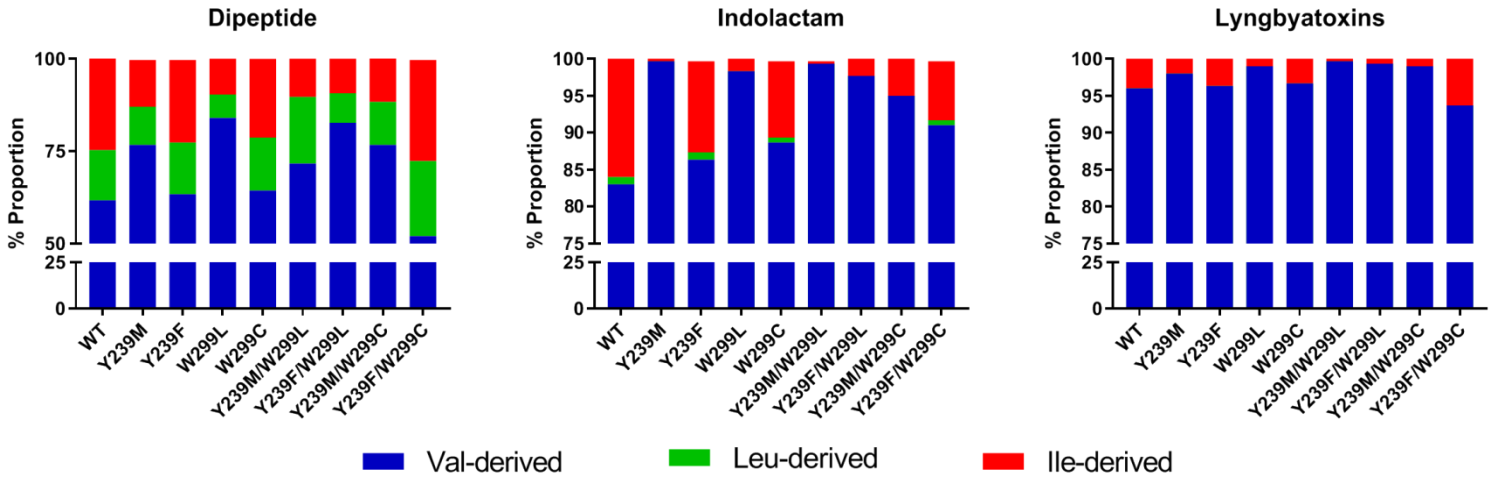


Figure 2.7 Proportion (%) of LTX analogues to each total intermediate pathway products. Total pathway products are the combined amounts of Val, Ile and Leu containing compounds in each different LtxA variants from minimal medium fermentation determined by HPLC-ESI-HRMS. Value of Ile-derived and Ile_{d10}-derived were combined as Ile-derived compounds. Values were calculated based on metabolite levels in 3 biological replicates.

2.3.5 Conclusion

This work reports an *in vitro i*nvestigation of the wild-type and mutated first A-domain of lyngbyatoxin NRPS, and further correlates the findings with the *in vivo* pathway activity when heterologously expressed in *E. coli*. Wild-type LtxA possesses multispecificity to valine related amino acids, which resulted in the production of novel LTX and intermediate analogues. The low level production of the analogues and the isomeric characteristic of isoleucine and leucine hindered direct compound identifications by LC-MS/MS analysis. To address this, an effective alternative analytical method was developed that involved precursor feeding of L-Ile-d_{10} and L-Leu-5,5,5-d_3 using a leucine auxotroph *E. coli* strain GB05-MtaA.

It can be concluded that with regards to NRPS enzyme activity, *in vitro* adenylation assays of full-length LtxA can be used to predict the corresponding *in vivo* activity. However, as the pathway involves reactions by the tailoring enzyme, *in vivo* pathway expression should also be investigated to provide a clearer picture of the natural metabolite flux inside the cells. In particular, the tailoring enzymes LtxB and LtxC were shown to have a strong preference for Val-derived dipeptides. Therefore, further downstream pathway modifications would be required for directing the *in vivo* synthesis to produce novel LTX-derived metabolites.

LtxA NRPS has some degree of substrate stringency. Attempts to reprogram the LTX pathway by site-directed mutagenesis of LtxA-A1 resulted in only minor differences for both *in vitro* and *in vivo* activities in comparison to the wild-type biosynthesis. Future studies could also investigate the influence of the MLP-like domain within LtxB which may affect LtxA enzyme activity.

Chapter 3. Expression of the PKS-related enzyme SxtA leads to production of saxitoxin intermediates in *E. coli*

This chapter has been prepared for publication.

Authors: **Angela H. Soeriyadi**, Rabia Mazmouz, Russell Pickford, Bakir Al-Sinawi, Ralf Kellmann, and Brett A. Neilan

Current Status: In preparation

The Extent to which the publication is your own:

I have constructed the SxtA expression plasmid and performed experiments to assess the protein expression and metabolite production. I developed the method and analysed the production of Int-A' and demetInt-A in *E. coli* including confirmation using precursor feeding study. I processed the LC-MS data analysis and interpretation. I prepared the first draft of the manuscript, and subsequent revisions based on advice from other authors.

Abstract

Saxitoxin is a neurotoxic alkaloid produced by freshwater cyanobacteria and marine dinoflagellates that is implicated in paralytic shellfish poisoning. The biosynthesis of saxitoxin is initiated with the formation of Int-A' (4-amino-3-oxo-guanidinoheptane) by a polyketide synthase-related protein, SxtA. Here, we heterologously expressed SxtA in *E. coli* and analysed its activity *in vivo*. Int-A' and a truncated analogue demetInt-A' were detected by HPLC-ESI-HRMS analysis. The chemical structures of these products were further verified by tandem mass spectrometry and labelled-precursor feeding with [guanidino-$^{15}N_2$] arginine and [1,2-$^{13}C_2$] acetate. This study has shown that Int-A' is not an exclusive enzyme product formed *in vivo*. The successful production of Int-A' opens up the possibility to characterise the pathway and produce paralytic shellfish toxins (PSTs) in *E. coli* through a heterologous expression approach.

3.1 Introduction

Saxitoxin (STX) and its chemical analogues, collectively known as paralytic shellfish toxins (PSTs), are neurotoxins associated with paralytic shellfish poisoning produced by freshwater cyanobacteria and marine dinoflagellates (Pearson *et al.*, 2016). PSTs feature a unique tricyclic backbone with two guanidium moieties. So far, more than 57 natural analogues of saxitoxin have been reported, containing different combinations of functional groups (Wiese *et al.*, 2010). The neurotoxicity of PST is attributed to their reversible binding to the toxin binding site 1 of the voltage-gated sodium channels, causing muscle paralysis (Henderson *et al.*, 1973, Terlau *et al.*, 1991). PSTs are known to cause human intoxication through the ingestion of contaminated seafood (Etheridge, 2010). In contrast, several PSTs have shown pharmaceutical potential as anaesthetic agents and as muscle relaxing agents (Epstein-Barash *et al.*, 2009, Kohane *et al.*, 2000, Garrido *et al.*, 2004, Garrido *et al.*, 2005, Lattes *et al.*, 2009).

The *sxt* biosynthetic gene cluster, responsible for saxitoxin formation, was first identified in the cyanobacterium *Raphidiopsis raciborskii* T3 (formerly *Cylindrospermopsis raciborskii* T3) (Aguilera *et al.*, 2018))(Kellmann *et al.*, 2008). Since then, *sxt* gene clusters have been identified in several other cyanobacterial species including *Dolichospermum*

circinale AWQC131C (formerly *Anabaena circinalis* AWQC131C) (Wacklin *et al.*, 2009) (Mihali *et al.*, 2009), *Aphanizomenon sp.* NH-5 (Mihali *et al.*, 2009), *Raphidiopsis brookii* D9 (Stucken *et al.*, 2010), *Microseira wollei* (formerly *Lyngbya wollei*) (McGregor & Sendall, 2015) (Mihali *et al.*, 2011), and *Scytonema crispum* (Cullen *et al.*, 2018a). The *sxt* gene clusters vary in gene content and organization (Murray *et al.*, 2011). Yet, the presence of the *sxtA* gene is conserved across all STX-producing organisms. SxtA is a polyketide synthase (PKS) related enzyme, which initiates the biosynthesis of saxitoxin and its analogues. It is, therefore, imperative to understand its activity.

SxtA was predicted by Kellmann and co-workers to catalyse the first step in the biosynthesis of saxitoxin, by condensing arginine and one methylated acetate unit to produce the 4-amino-3-oxo-guanidinoheptane, here referred as intermediate A' (Int-A', **3**, Figure 3.1). Int-A' was detected via mass spectrometry in *R. raciborskii* cellular extracts, and proposed to be modified by downstream enzymes to form saxitoxin (**4**) (Kellmann *et al.*, 2008). Tsuchiya *et al.* successfully chemically synthesized Int-A' and confirmed its presence in the extracts of several PSTs producers (Tsuchiya *et al.*, 2014).

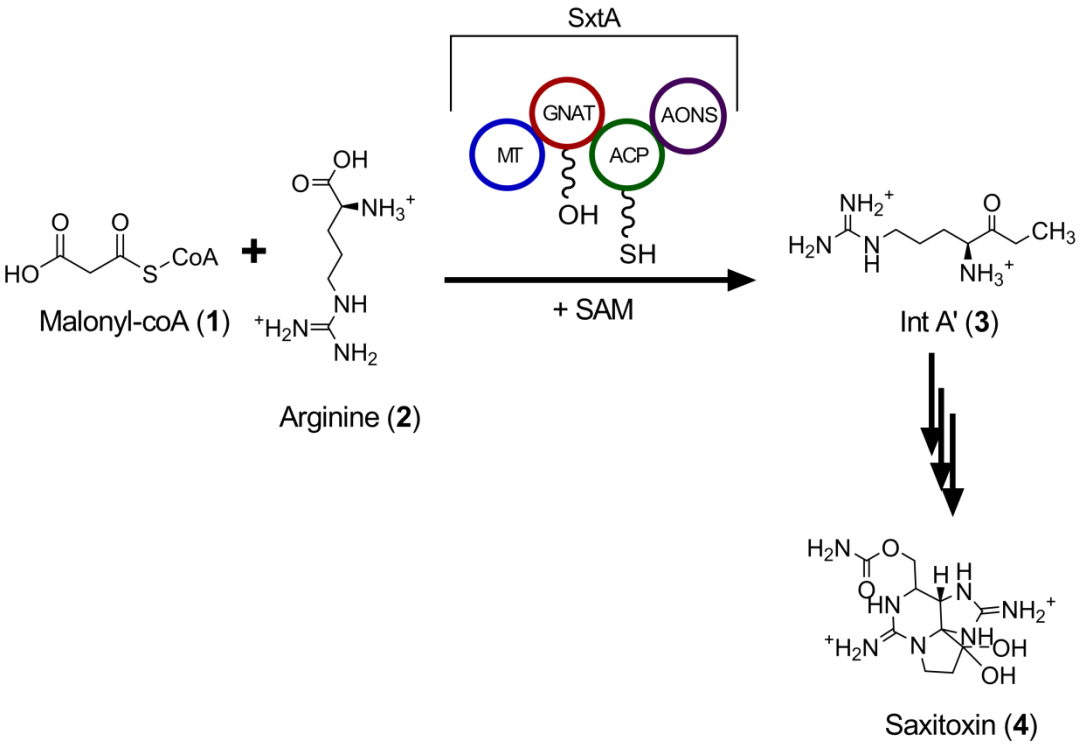


Figure 3.1 The initial step of saxitoxin biosynthesis. The formation of Int-A'(**3**) catalysed by a PKS-related enzyme SxtA, through Claisen condensation of malonyl-CoA (**1**) and arginine (2). The enzymatic reaction is catalysed in a stepwise manner by the four domains within SxtA; acyltransferase domain (ACP), a S-adenosylmethionine (SAM) dependent methyltransferase (MT) domain, a GCN5-related N-acetyltransferase (GNAT) domain, and an 8-amino-7-oxononanoate synthase (AONS) domain. Int-A' is then modified by several enzymes encoded within the *sxt* cluster forming saxitoxin (**4**).

Recently, Chun *et al*. performed an enzymatic characterization of SxtA (Chun *et al.*, 2018). Interestingly, *in vitro*, SxtA from *R. raciborskii* was able to use different carboxylic acid-CoA substrates, which resulted in two different enzyme products, ethyl ketone and methyl ketone here referred as Int-A' (**B**) and demethylated Int-A' (demetInt-A', **5**), respectively. Formation of Int-A' is exclusively observed when either propionyl-CoA or methylmalonyl-CoA are present. Whereas demetInt-A' is produced in the presence of acetyl-CoA. Both Int-A' and demetInt-A' are produced at a similar ratio when malonyl-CoA is used as the substrate. However, methylation of malonyl-CoA by the methyltransferase domain in the presence of SAM is essential to produce Int-A' (Chun *et al.*, 2018). Despite the current understanding, there

is no knowledge on how this substrate usage will affect the biosynthesis of Int-A' and saxitoxin *in vivo*.

Several synthetic chemistry studies performed by Tsuchiya *et al.* gave insights into the reactions of saxitoxin biosynthesis (Tsuchiya *et al.*, 2015, Tsuchiya *et al.*, 2016, Tsuchiya *et al.*, 2017). However, these studies did not investigate the involvement of the genes within the pathway. Heterologous expression of cyanobacterial nonribosomal peptide synthetase (NRPS) and PKS pathways in *E. coli* have been shown to be a useful approach for their biochemical characterization (Ongley *et al.*, 2013b, Liu *et al.*, 2017, Cullen *et al.*). Therefore, this study aimed to heterologously express SxtA to assess its activity *in vivo*, and pave the way for characterization of this novel and complex biosynthetic pathway.

3.2 Methods

3.2.1 Construction of an expression plasmid for *sxtA*.

The *sxtA* gene from *R. raciborskii* T3 was cloned onto a broad host range vector with the RK2 minimal replicon, pVB expression plasmid (Vectron Biosolutions). This vector contains a positively regulated XylS/*Pm* promoter system from *Pseudomonas putida* that is inducible by benzoic acid derivatives such as *m*-toluic acid (Gallegos *et al.*, 1996, Worsey & Williams, 1975, Winther-Larsen *et al.*, 2000). Briefly, the expression plasmid pVB was digested with EcoRI and NdeI (New England Biolabs). Native *sxtA* from *R. raciborskii* T3 was amplified by PCR using Velocity DNA polymerase (Bioline). Fragments were purified (Zymoclean Gel DNA Recovery Kit, ZymoResearch) and assembled using the NEBuilder HiFi DNA Assembly (New England Biolabs). The *E. coli* strain BL21(DE3) was electroporated with the assembled fragments and grown under ampicillin selection. Successful constructs of pVB::*sxtA* were screened by colony PCR with primers flanking the insertion site (pVB_upstreamF and pVB_R). The purified plasmid (PureLink Miniprep, Invitrogen) was then sequenced using a BigDye Terminator v3.1 sequencing kit and detected by an ABI 3730 analyzer sequencer (Applied Biosystems) at the Ramaciotti Centre for Genomics (UNSW Sydney, Australia). Plasmid integrity was confirmed

by restriction enzyme digestion using BamHI (NEB). The construction of a plasmid encoding the N- and C-terminal His-tagged SxtA was performed using a similar method. All primers used for these cloning experiments are listed in Table 3.1.

3.2.2 Construction of an *sfp* expression plasmid

The phosphopanteinyl transferase (PPTase) expression plasmid (pET28b::*sfp*) was constructed using restriction enzyme cloning. Briefly, *sfp* from *Bacillus subtilis* was amplified from genomic DNA using Velocity DNA polymerase (Bioline). The PCR amplicon was purified (DNA Clean and Concentrator 5X Kit, ZymoResearch) and subjected to a restriction enzyme digestion using NcoI and XhoI (NEB). Digested products were purified and ligated into NcoI and XhoI digested pET28b (Novagen), containing a T7 promoter system with N and C-terminal histidine tags. After transformation into electrocompetent *E. coli* GB2005 cells (Fu *et al.*, 2008), clones were screened by performing PCR on the purified plasmid (PureLink Miniprep, Invitrogen) with the universal T7 promoter and T7 terminator primers. The inserts were sequenced as described for pVB::*sxtA*.

3.2.3 Expression and purification of SxtA and Sfp

For the expression of holo-SxtA, starter cultures of *E. coli* BL21(DE3) pVB::His-*sxtA* pET28b::*sfp* were grown overnight in lysogeny broth (LB) supplemented with 50 µg/mL kanamycin and 50 µg/mL ampicillin at 30°C shaking at 200 rpm. The overnight cultures were inoculated (1:100) into 200 mL Terrific Broth (TB) in 1 L flasks with cotton stopper as the lid, supplemented with 50 µg/mL kanamycin and 50 µg/mL ampicillin. The cultures were incubated at 30°C with shaking at 200 rpm to an OD_{600} of ~0.4 before being transferred to 18°C to reach an OD_{600} of ~0.5. Expression of *sfp* and *sxtA* was induced by the addition of 0.05 mM isopropyl-β-D-thiogalactoside (IPTG) and 0.5 mM *m*-toluic acid. Cells were incubated for 24 h at 18°C before being harvested by centrifugation (4,000 rpm, at 4°C for 20 min). The cell pellet was washed with chilled 20 mM Na_2HPO_4 twice before being stored at -20°C until further purification.

For protein purification, the cell pellet was resuspended in 5 mL of lysis buffer (20 mM sodium phosphate buffer [pH 7.5], 500 mM NaCl, 20 mM imidazole, 10% glycerol, protease inhibitor [Roche]). Cells were lysed by sonication (Branson Digital Sonifier M450, 3 mm probe, 25% of amplitude, 3 min at 4°C with cycles of 10 s power on and 59 s power off). The resulting suspensions were centrifuged at 20,000 rpm, at 4°C for 60 min. SxtA and Sfp proteins were purified by bench purification using Ni-NTA HisBind Resin (Merck) with buffer A (20 mM sodium phosphate buffer, pH 7.5, 500 mM NaCl, 20 mM imidazole), buffer B1 (20 mM sodium phosphate buffer, pH 7.5, 500 mM NaCl, 68 mM imidazole), and buffer B (20 mM sodium phosphate buffer, pH 7.5, 500 mM NaCl, 500 mM imidazole). The collected fractions were analyzed using 10% polyacrylamide gel electrophoresis under denaturing conditions (SDS-PAGE). Protein concentration was determined by using a protein assay kit (BioRad). Purified protein was stored at -80°C. The identities of the purified protein were determined following trypsin digestion and peptide mass fingerprinting analysis at the Bioanalytical Mass Spectrometry Facility (UNSW Sydney, Australia).

Table 3.1 Primers used for the construction of pVB::*sxtA* and pET28::*sfp* expression plasmids.

Primer Name	Nucleotide sequence 5'-3'
Fragment amplification	
***sxtA*-nat-rbs8-F**	**CAATAATAATGGAGTCATGAAC**ATGTTACAAAAGATTAA TCGTTATACTCACGGCT
H*sxtA*-pVB-R	**AAAGCTGACTCTAGCTAGAGGATCC**TCAATGTTTCAGG ATATTTTCTTC
his*sxtA*F_pVBHA	**CAATAATAATGGAGTCATGAAC**ATGGGCAGCAGCCATC ATC
his*sxtA*R_pVBHA	**AAAGCTGACTCTAGCTAGAGGATCC**TCAGTGGTGGTGG TGGTGG
***sfp* forward**	GGCATCCATGGGCAAGATTTACGGAA
***sfp* reverse**	TATCTCGAGAGTGTTGATTTCGTTGGCTG
Screening and sequencing primers	
pVB_upstreamF	ATCTGTTGTTTGTCGGTGAAC
***sxtA*_ 716R**	TCAATGTGCCTTTCGTGACC
***sxtA*_ 589F**	GAGCGAGCCTTGAATACAGC
***sxtA*_ 1983R**	TCGTTGCTCATTGTGTAGCT
***sxtA*_ 1903F**	GGTGTAACTCGCTGTCGA
***sxtA*_ 3362R**	AATCCGGCTAAGGCACTT
pVB_R	AGTAATCGCAACATCCGC

Bolded bases are the overlapping sequences for Gibson assembly. Underlined bases are the RBS spacer region insertions.

3.2.4 Expression and extraction of Int-A' and demetInt-A'

For the production of the intermediate compounds by SxtA, *E. coli* BL21(DE3) pET28b::*sfp* and pVB::*sxtA* were grown overnight in LB supplemented with 50 µg/mL kanamycin and 50 µg/mL ampicillin at 30°C shaking at 200 rpm. Cultures were inoculated at 1:100 dilution into 50 mL TB pH 6.8 supplemented with 50 µg/mL kanamycin and 50 µg/mL ampicillin in 250 mL Erlenmeyer flasks with cotton stopper as the lid. A previous study had shown that pH 6.8 is suitable for the activity of *Pm* promoter (Winther-Larsen *et al.*, 2000). Cultures were incubated at 30°C with shaking at 200 rpm to an OD_{600} of ~0.4 before being transferred to 18°C to reach an OD_{600} of ~0.5. Expression of *sfp* and *sxtA* were induced through the addition of 0.05 mM isopropyl β-D-thiogalactoside (IPTG) and 0.5 mM *m*-toluic acid. The cultures were incubated

for 24 h and harvested by centrifugation. The cell pellet was washed with chilled MilliQ water twice before being stored at -20°C until further processing.

For LC-MS analysis, metabolites were extracted in 0.1% formic acid with 1:3 v/w (volume:wet weight) ratio. Cells were lysed by boiling at 98°C for 5 min. The cell lysates were then clarified by centrifugation and the supernatants were evaporated to dryness using a freeze dryer (Martin Christ). Extracts were resuspended in 200 µL of 0.1% formic acid and diluted five times to a final solvent of 80% acetonitrile 0.1% formic acid. Samples were clarified by centrifugation and transferred into HPLC vials for LC-MS analysis.

3.2.5 Stable isotope precursor-feeding study

Cell cultures in the precursor feeding experiments were grown in 1 L M9 minimal medium with the supplementation of an isotope labelled precursor of either [guanidino-$^{15}N_2$] arginine or [1,2-^{13}C] acetate, at a final concentration of 100 mg/L. In order to decrease non-specific incorporation, feeding was performed in two steps, one third of the labelled substrates was added at the beginning of the culture and the remainder was added during induction (OD_{600} ~0.5). Culture conditions were as described in the production of the intermediates except for a longer incubation of 48 h.

3.2.6 Analysis of Int-A' and demetInt-A' production

LC-MS analysis of the intermediates was performed on 20 µL of the extracts using a U3000 UHPLC connected to a Q-Exactive Plus (ThermoFisher Scientific) via a heated electrospray interface. For the analysis of metabolite production in the precursor feeding study, 100 µL of the extracts were analysed. Liquid chromatography was performed with an AQUITY UPLC BEH amide column (130 Å, 1.7 µm, 2.1 mm x 100 mm [Waters]). Mobile phase A was 5 mM ammonium formate and 5 mM formic acid. Mobile phase B was 5 mM ammonium formate and 5 mM formic acid in 95% acetonitrile. The gradient program was as follows: 0-5 min 85% B, 5-20 min 85-70% B, 20-22 min 70% B, 22-23 min 70-85% B, 23-29 min 85% B. The flow rate was 0.3 mL/min. Ionization was performed in positive mode under default source conditions, as suggested by the manufacturer's tune software. Mass spectra were acquired with full scan mode

with a range of 160-400 *m/z* and a resolution of R = 70,000. An inclusion list was used to direct tandem mass spectrometry (MS/MS) analysis of the precursor masses corresponding to the predicted metabolites with their $[M+H]^+$ (Int-A' *m/z* = 187.15530, demetInt-A' *m/z* = 173.13969, $^{15}N_2$ Int-A' *m/z* = 189.14941, $^{15}N_2$ demetInt-A' *m/z* = 175.13376, $^{13}C_2$ Int-A' *m/z* = 189.16205, $^{13}C_2$ demetInt-A' *m/z* = 175.14640). All data were processed with Xcalibur V2 software (Thermo Scientific). Compounds were identified by mass with < 5 ppm error and confirmed by MS/MS spectra. Statistical analysis was performed using Progenesis QI (Nonlinear Dynamics) and GraphPad Prism 7. Prediction of the fragmentation pattern based on the chemical structure was performed using Mass Frontier software 7 (Thermo Scientific).

3.3 Results and discussion

SxtA is a PKS-related enzyme, thus requires heterologous expression with a PPTase for conversion to its active form. In the present study, the PPTase enzyme Sfp from *Bacillus subtilis* was co-expressed with SxtA due to its efficiency in converting SxtA-ACP to its active form (Liu *et al.*, 2018). The *sxtA* gene from *R. raciborskii* T3 was cloned into a pVB expression vector containing a XylS/*Pm* promoter system using Gibson cloning (Gibson *et al.*, 2009). SxtA was expressed concurrently with the *sfp* gene controlled under T7 promoter. For protein analysis, two histidine tags, one at each terminus, were cloned to assist the purification of SxtA and Sfp. The tagged constructs pVB::His-*sxtA* and pET28::*sfp* were co-expressed in *E. coli* BL21(DE3). The level of Sfp and SxtA protein were different as expected, as they were expressed under different promoters. Protein purification yielded fractions containing SxtA and Sfp, as confirmed by peptide mass fingerprinting (Figure 3.2, Appendix Figure A 2.1 and 2). This confirmed that the SxtA protein had been successfully expressed at a high-level.

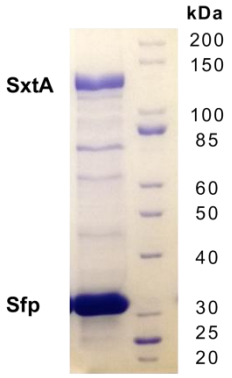


Figure 3.2. SDS-PAGE of the purified protein after Immobilized Metal Ion Affinity Chromatography (IMAC). Lane 1: protein fraction containing SxtA (147 kDa) and Sfp (27 kDa). Lane 2: unstained protein standard broad range, 10–200 kDa (New England Biolabs). The identities of proteins were confirmed by trypsin digestion and MS/MS (Appendix Figure A 2.1 and 2).

To determine the biosynthetic products of SxtA, intracellular metabolites of *E. coli* cells harbouring pET28b::*sfp* and pVB::*sxtA* cultured in Terrific Broth (TB) medium were extracted and analysed by liquid chromatography mass spectrometry (LC-MS). Extracts from induced cells harbouring pET28b::*sfp* and pVB plasmids were used as negative controls. HPLC-ESI-HRMS analysis revealed the presence of an ion with a retention time of 4.7 min and a mass-to-charge ratio (*m/z*) of 187.1553 $[M+H]^+$, corresponding to the protonated mass of Int-A' ($C_8H_{18}N_4O$) (Figure 3.3, panel A and B). This ion was detected in *E. coli* pVB::*sxtA* in both uninduced and induced samples. Collision-induced fragmentation produced tandem MS spectra that were in accordance with those reported previously for Int-A' by Tsuchiya et al. (Tsuchiya *et al.*, 2014), including the diagnostic fragments with *m/z* of 170, 128, 110, and 60 (Appendix Figure A 2.3, panel A). A trace amount of an interfering compound with a similar *m/z* was detected in the vector control, but its MS2 spectrum was highly dissimilar to Int-A'. The experiment was performed with six biological replicates and the data were analysed by Progenesis QI (Nonlinear Dynamics) to determine any difference in expression between samples. Int-A' was shown to be produced at significantly higher level in the induced cells compared to the uninduced cells (Figure 3.3, panel C) (one-way ANOVA p-value < 0.0001, Appendix Table A 2.1). The low level production of Int-A' in the uninduced cells suggests

leakiness of the XylS/*Pm* promoter system, consistent with a previous report (Balzer *et al.*, 2013). However, the extent of promoter activity in the uninduced condition was not established due to the lack of statistical difference in ion abundances between the uninduced cells and the negative control (one-way ANOVA *p*-value > 0.01, Appendix Table A 2.1).

Production of demetInt-A' (**5**), an analogue of Int-A' lacking the methyl group was also observed. An ion with an *m/z* ratio corresponding to protonated demetInt-A' ($C_7H_{16}N_4O$ *m/z* 173.1397) was detected in the pVB::*sxtA* uninduced and induced samples, but not in the vector control (Figure 3.3, panel D and E). The demetInt-A' that was detected eluted at 5.6 min, slightly after Int-A'. Tandem mass spectrometry of this molecular ion showed a fragmentation pattern indicating structural similarity to Int-A' with the lack of a methyl group (Appendix Figure A 2.3, panel B). A previous study of SxtA enzymology has shown that SxtA possesses broad substrate specificity *in vitro* forming both Int-A' and demetInt-A' (Chun *et al.*, 2018). Our results have shown that in *E. coli,* both compounds were produced *in vivo* albeit at an unknown quantity and production ratio. In the present study, we did not investigate the substrates used by SxtA and their relative contribution towards Int-A' and demetInt-A' formation. We do not discount the possibility that the formation of Int-A' *in vivo* could also arise from the usage of various acyl-CoA precursors.

It was observed that *E. coli* pET28::*sfp* pVB::*sxtA* was able to produce demetInt-A' in the absence of an expression inducer (Figure 3.3, panel D and F, Appendix Table A 2.2). The leakiness of the pVB promoter does not explain the high level of demetInt-A' with a low level of Int-A' detected in the uninduced cells. The results suggest that the demetInt-A' and Int-A' could arise from different substrates, which would be consistent with the *in vitro* activity of SxtA described by Chun et al. (Chun *et al.*, 2018). The study also showed that the disruption of the methyltransferase domain causes exclusive formation of demetInt-A' from malonyl-CoA (Chun *et al.*, 2018). Therefore, we postulate that a truncated SxtA with an inactive methyltransferase domain might be present. An analysis of the SxtA protein coding sequence using Virtual Footprint (Munch *et al.*, 2005) revealed the presence of a putative promoter element with high homology to σ^{70} elements (Appendix Table A 2.3). The -10 region was predicted to be 21 bp

downstream of the *sxtA* start codon. Several ORFs were predicted by Geneious R9.0.4 (Biomatters) located downstream of the putative promoter, including 93, 105, and 138 bp downstream of the predicted -10 sequence with a potential ribosomal binding site within the region (Appendix Figure A 2.4). As the activity of internal promoters is not in the scope of this study, experimental verification of the putative promoters was not performed (Chun *et al.*, 2018). It is hypothesized that the putative promoter and/or potentially other internal promoter may lead to the expression of a truncated SxtA, leading to the production of demetInt-A' in the uninduced cells.

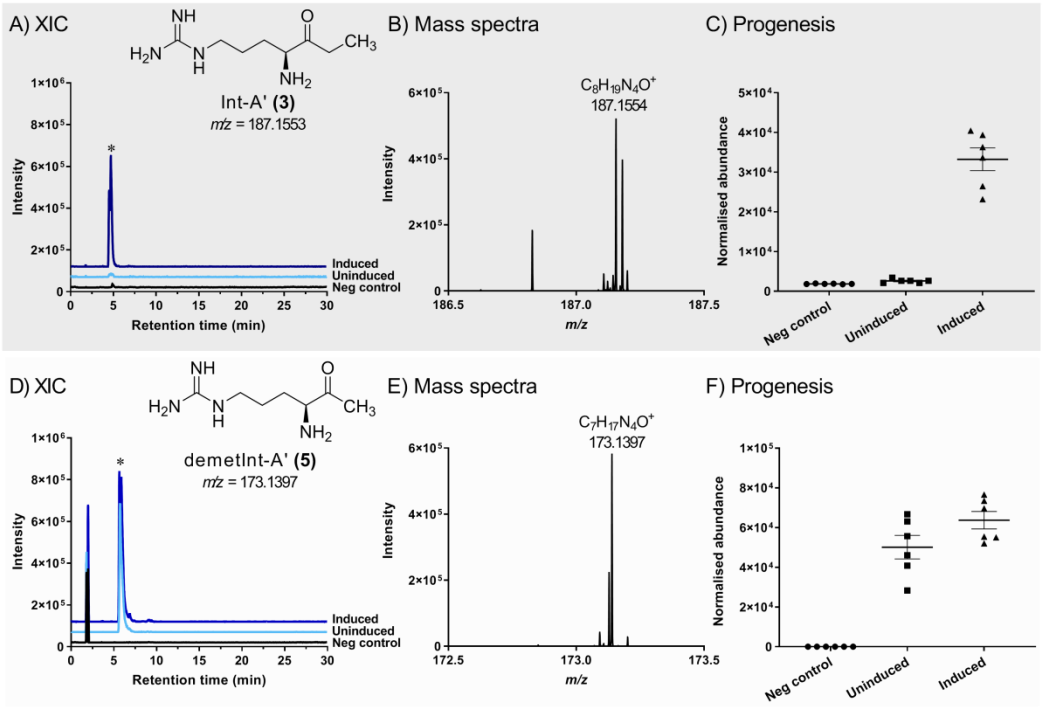


Figure 3.3. HPLC-ESI-HRMS analysis of intermediates production. Analysis of Int-A' (top panel) and demetInt-A' (bottom panel) production in *E. coli* pET::*sfp* and pVB::*sxtA* and controls. Extracted ion chromatograms for *m/z* = 187.1553 (**A**) and 173.1397 (**D**) (± 5 ppm) from *E. coli* culture extracts with pET::*sfp* and pVB vector (black), *E. coli* with pET::*sfp* and pVB::*sxtA* uninduced (red) and induced (blue). Mass spectrum of induced extract for Int-A' (**B**, RT = 4.7 min) and demetInt-A' (**E**, RT = 5.6 min). Progenesis QI analysis comparing the expression of Int-A' (**C**) and demetInt-A' (**F**) between extracts of negative control (*E. coli* pVB pET::*sfp*, circles) and extracts of *E. coli* pVB::*sxtA* pET::*sfp* culture uninduced (square) and induced (triangles). Data obtained from 6 biological replicates (n = 6) with error bars represent the standard error of the mean (SEM). *, denotes peak of interests.

In order to confirm the identity of the compounds produced, a feeding study with stable isotope labelled precursors was performed. *E. coli* BL21(DE3) harbouring both pVB::*sxtA* and pET28b::*sfp* was cultured in M9 minimal medium supplemented with either [guanidino-$^{15}N_2$] arginine or [1,2-$^{13}C_2$] acetate. LC-MS analysis showed the presence of new molecular ions in label-fed cultures that are absent in non-labelled control cultures (Appendix Table A 2.4 and 5). The retention times of these ions were the same as Int-A' and demetInt-A' (Appendix Figure A 2.5). For cultures fed with [guanidino-$^{15}N_2$] arginine, molecular ions with an *m/z* of 189.1494 ($C_8H_{19}{}^{15}N_2N_2O^+$) and 175.1338 ($C_7H_{17}{}^{15}N_2N_2O^+$) were found (Figure 3.4, Appendix Table A 2.4 and A 2.5). This +2 Da difference indicates the incorporation of two ^{15}N atoms of arginine for Int-A' and demetInt-A'. In the cultures fed with [1,2-$^{13}C_2$] acetate, molecular ions with an *m/z* of 189.1620 ($^{13}C_2C_6H_{19}N_4O^+$) and 175.1464 ($^{13}C_2C_5H_{17}N_4O^+$) were found, indicating the incorporation of two ^{13}C atoms from the acetate. The spectra obtained through tandem mass spectrometry showed the presence of the expected fragment ions containing 2+ Da carbon and nitrogen shifts (Figure 3.4, for detail mass lists see Appendix Table A 2.4 and 5). This supports the conclusion that structures of the identified compounds are as expected and that both arginine and acetate are incorporated during the biosynthesis of Int-A' and demetInt-A'.

The precursor feeding study showed efficient incorporation of 50-60% of [guanidino-$^{15}N_2$] arginine into both intermediate products (Figure 3.4, Appendix Figure A 2.5). However, incorporation of [1,2-$^{13}C_2$] acetate into the intermediates was relatively low, with only 8% and 2% of Int-A' and demetInt-A' containing labelled acetate derivatives, respectively (Appendix Figure A 2.5). *E. coli* has been shown to take up excreted acetate from outside of cell and convert it back to acetyl-CoA through the activity of acetyl-CoA synthetase (Brown *et al.*, 1977, Kumari *et al.*, 2000). However, this conversion only occurs in a sugar-depleted environment, in particular at the beginning of a stationary phase (Wolfe, 2005). Hence, it is not clear whether the high level of labelled acetate was converted efficiently into acetyl-CoA and subsequently into other CoA derivatives, which are the precursors of the intermediates. In addition to being less abundant than the unlabelled acetyl-CoA, [1,2-^{13}C]acetyl CoA is likely to be consumed by other

competing pathways, as it is one of the central metabolites in *E. coli* metabolism having a rapid turnover time (Wolfe, 2005).

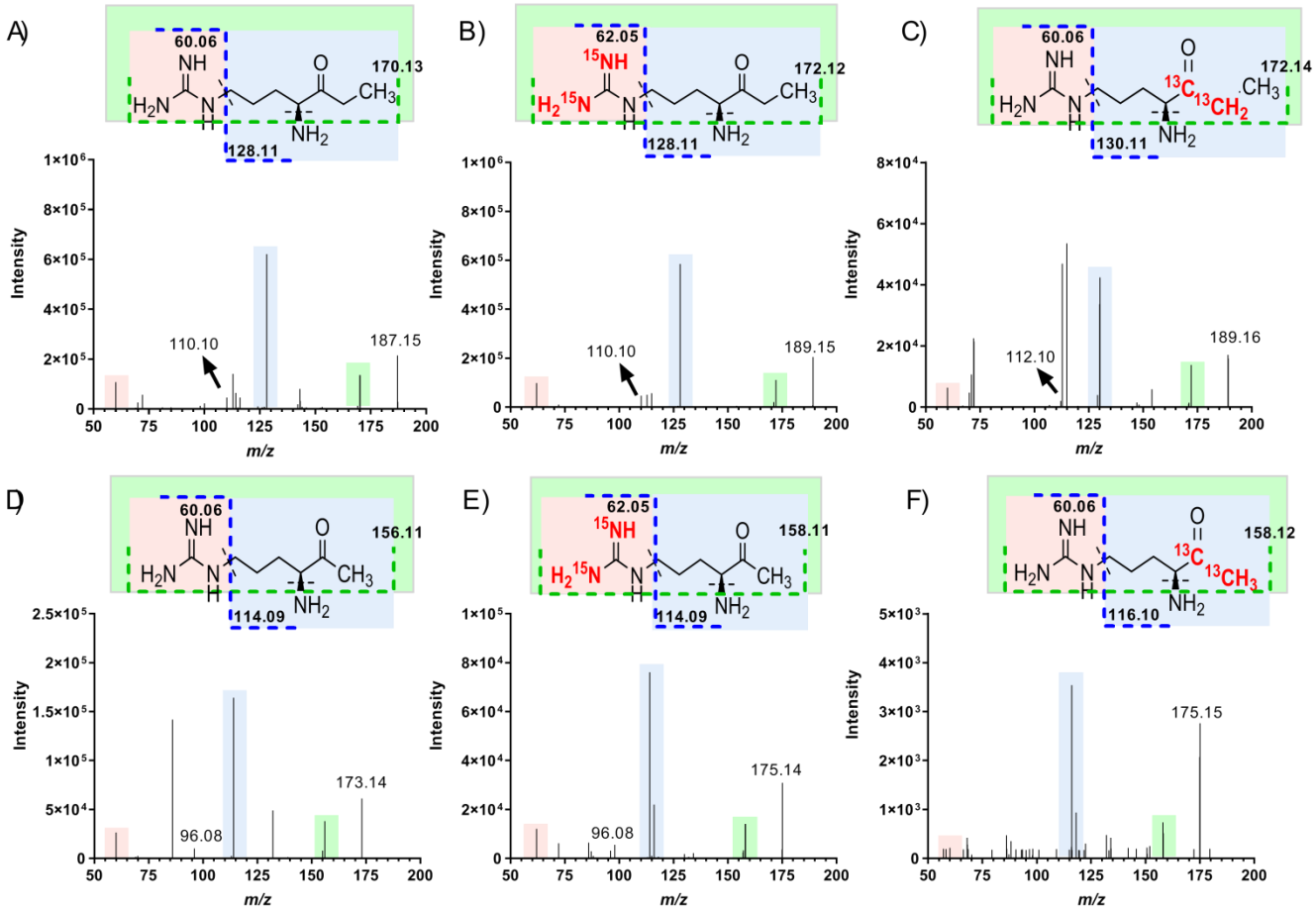


Figure 3.4. HPLC-ESI-HRMS/MS analysis Int-A' and demetInt-A' from isotope labelled precursor feeding experiment. Fragmentation patterns of Int-A' (**A**), $^{15}N_2$-Int-A' (**B**), and $^{13}C_2$-Int-A' (**C**) as well as demetInt-A' (**D**), $^{15}N_2$-demetInt-A' (**E**) and $^{13}C_2$-Int-A' (**F**), detected from the extracts of *E. coli* pVB::*sxtA* pET28::*sfp* cultured in M9 minimal media. Mass shifts were observed from both compounds when the cultures were supplemented with [guanidino-$^{15}N_2$] arginine (**B** and **E**) and [1,2-^{13}C]acetate (**C** and **F**) in comparison to non-supplemented media (**A** and **D**).

We also observed that the detected signal ratio between the Int-A' and demetInt-A' varied in cells cultured in different media, being high in M9 minimal media and low in TB (Figure 3.3, Appendix Figure A 2.5). This was not the focus of the present study, and thus the production of intermediates was not quantified under different conditions. Nevertheless, this shows that culture conditions might influence the flux of intermediate production. A further comparative investigation of substrate usage by SxtA *in vivo* would benefit the optimization of Int-A' production, and ultimately saxitoxin production.

In conclusion, we have demonstrated a heterologous expression of the initial enzyme in saxitoxin biosynthesis, SxtA in *E. coli*. Production of both Int-A' and demetInt-A' *in vivo* were observed. The successful production of Int-A' opens up the possibility to characterize the pathway *in vivo* and produce PSTs in *E. coli* using a heterologous expression approach.

Chapter 4. Heterologous expression of the neosaxitoxin biosynthetic pathway in *E. coli*

Abstract

Paralytic shellfish toxins (PSTs) are a group of toxic alkaloids produced by freshwater cyanobacteria and marine dinoflagellates. Neosaxitoxin, a member of this toxin family, has potential pharmaceutical application as an anaesthetic agent. The biosynthesis of PSTs have been observed in several cyanobacterial strains, but the pathways have not been well-characterised due to the slow growth and genetic intractability of the native producers. Herein, the emerging technology of synthetic biology was utilised to develop directed heterologous production of neoSTX in *E. coli* and further facilitate the pathway characterisation. The saxitoxin biosynthetic cluster from *R. raciborskii* T3 was re-arranged, optimised for *E. coli* heterologous expression system, and integrated into for *E. coli* BL21(DE3). The present study reports the extension of previous work which had integrated a partial *sxt* biosynthetic cluster. *E. coli* strains producing neoSTX were successfully constructed, and as a result several genes within the cluster were determined to be unnecessary for biosynthesis. NeoSTX production was observed using three independent assessments; HPLC-ESI-HRMS analysis, enzyme-linked immunoadsorbent assay (ELISA) and mouse neuroblastoma bioassay (MNBA). This study demonstrated a proof of concept process for neosaxitoxin production in a heterologous host using synthetic biology techniques.

4.1 Introduction

Saxitoxin (STX) and its analogues, collectively known as paralytic shellfish toxins (PSTs), are neurotoxic alkaloids with voltage-gated sodium channel blocking activity. Known producers of PSTs include freshwater cyanobacteria of the genera *Aphanizomenon* (Mahmood & Carmichael, 1986), *Cylindrospermum* (Borges *et al.*, 2015), *Dolichospermum* (previously Anabaena) (Llewellyn *et al.*, 2001), *Phormidium* (Borges *et al.*, 2015), *Planktotrix* (Pomati *et al.*, 2000), *Raphidiopsis* (including strains previously under the genera *Cylindrospermosis*) (Stucken *et al.*, 2010, Lagos *et al.*, 1999), *Microseira* (previously *Lyngbya*) (Foss *et al.*, 2012, Carmichael *et al.*, 1997), and *Scytonema* (Smith *et al.*, 2011), as well as marine dinoflagellates (eukaryotes) of the

genera *Alexandrium* (Lefebvre *et al.*, 2008), *Gymnodinium* (Oshima *et al.*, 1993), and *Pyrodinium* (Usup *et al.*, 1994). One member of the PST family, neosaxitoxin (neoSTX), is a N-1 hydroxylated toxin currently being investigated in clinical trials as a prolonged-action local anaesthetic agent (Lattes *et al.*, 2009, Rodriguez-Navarro *et al.*, 2009, Wylie *et al.*, 2012). This compound could potentially overcome the limitations of current pain therapies, which are mainly based on opioids.

The current state-of-the-art method for producing neoSTX for research and environmental monitoring applications is via chemical extraction from the native cyanobacterial producers, or from contaminated shellfish (Laycock *et al.*, 1994, González *et al.*, 2015, Rubio *et al.*, 2015). However, extraction from native producers and environmental samples is costly, tedious, low yield, inefficient, and prone to contamination by impurities (Ongley *et al.*, 2013a, Rey *et al.*, 2015, D'Agostino *et al.*, 2019). As such, utilising native neoSTX producers is not a feasible industrial production route. They are intractable to genetic modification, fastidious with slow growth rates and low metabolite titres, and do not grow under axenic conditions which provides obstacles for seeking regulatory approval for this pharmaceutical production route (Vu *et al.*, 2018, Burja *et al.*, 2001, Lurling *et al.*, 2013, Šulčius *et al.*, 2017). In addition, neoSTX is commonly produced together with other types of PSTs which are challenging to purify individually (D'Agostino *et al.*, 2019, Lawrence *et al.*, 2005). Heterologous expression of the neoSTX biosynthetic pathway is an alternative approach, which could provide a sustainable, scalable, economically feasible production process for this molecule. *E. coli* strains are among the most commonly used Food and Drug Administration (FDA)-approved organisms for biopharmaceutical production (Ferrer-Miralles *et al.*, 2009). In this study, *E. coli* was chosen as the heterologous host based on its fast growth, simple process set-up, and well-defined capacity for genetic manipulation useful for directing single compound manufacture (Chi *et al.*, 2019, Ferrer-Miralles *et al.*, 2009, Ongley *et al.*, 2013b, Ongley *et al.*, 2013a). The extensive knowledge of *E. coli* metabolism is useful for further metabolic engineering strategies to increase product yield. Additionally, by using biosynthetic pathway engineering, the production of other PSTs could also be possible.

The (neo)saxitoxin (*sxt*) biosynthetic gene cluster was first described in the freshwater cyanobacterium *Raphidiopsis raciborskii* T3 (previously known as *C. raciborskii* T3) (Aguilera *et al.*, 2018, Kellmann *et al.*, 2008) (Figure 4.1). Since then, studies have identified various *sxt* gene clusters in several other cyanobacteria (Mihali *et al.*, 2009, Mihali *et al.*, 2011, Stucken *et al.*, 2010, Ballot *et al.*, 2016). The gene clusters were found to vary in size and were not as well-conserved as other cyanotoxin biosynthetic pathways, particularly in regard to gene composition (Figure 4.1). These genomic differences putatively lead to altered toxin profiles in different PST producers (D'Agostino *et al.*, 2019). Despite this, several genes are consistently present in all *sxt* clusters, including *sxtABCDEGHIJKPQRSTU*, some of which were proposed to be directly involved in the biosynthesis of saxitoxin analogues (Cullen *et al.*, 2018b).

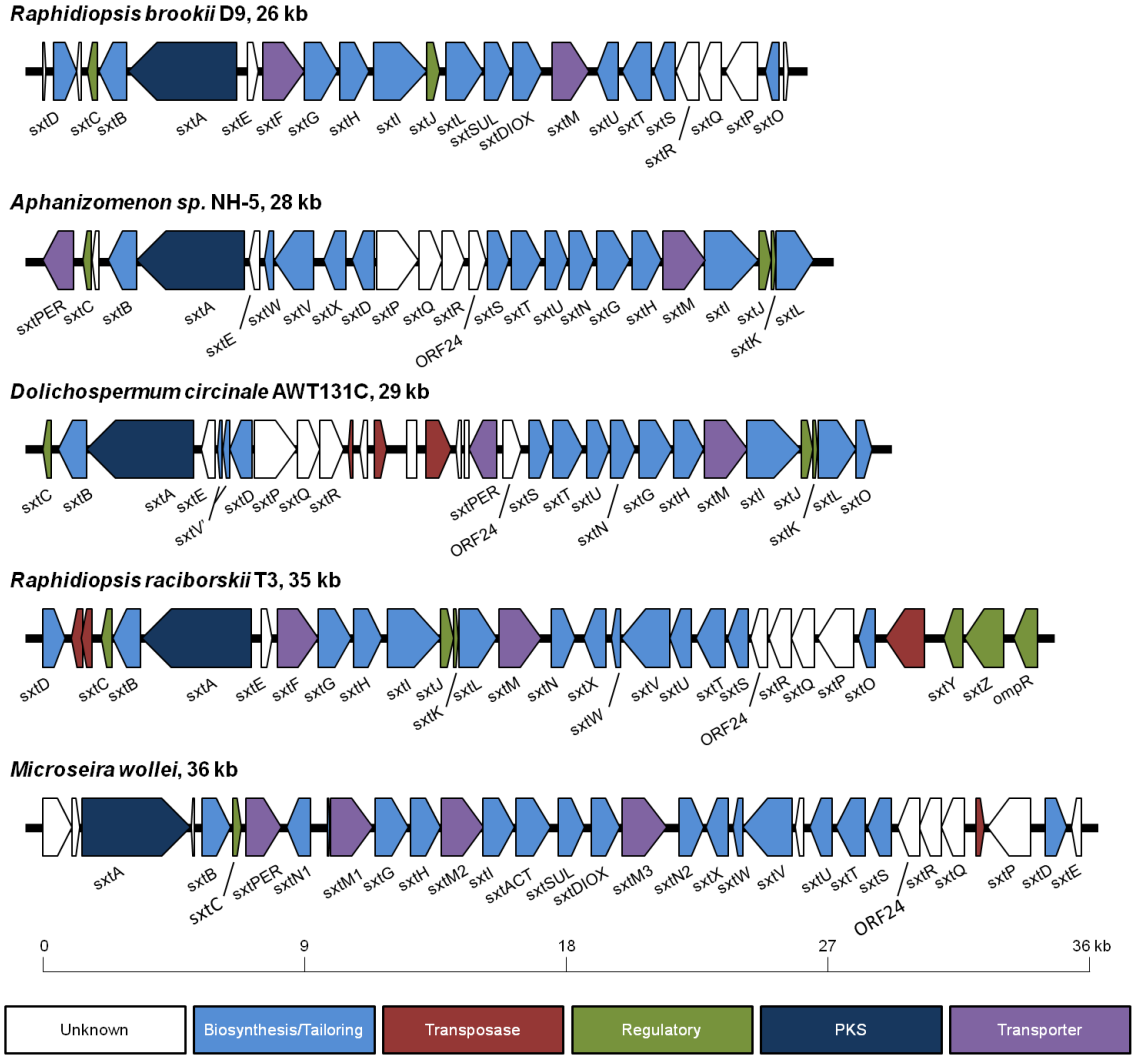


Figure 4.1 Organisation of PST biosynthetic gene clusters from various cyanobacteria species. The *sxt* biosynthetic cluster was previously deduced from the nucleotide sequence analysis, and the size of the ORFs was predicted based sequence homology. The lengths of the clusters are given as kilo base pairs (kb). ORFs homologous to PKS enzymes are indicated in dark blue. Putative modifying enzyme ORFs are indicated in light blue. ORFs which were predicted to have regulatory or transport functions are indicated in green and purple, respectively. Regions homologous to transposases are indicated in red, while ORFs with unknown function are shown in white.

The current knowledge of the saxitoxin biosynthetic pathway is based on synthetic chemistry studies and bioinformatic analysis of *sxt* genes, in combination with their biochemical characterisation (see Chapter 1 Figure 1.5 for the proposed STX pathway). The biosynthesis of saxitoxin is initiated with the formation of Int-A' by SxtA (PKS-like enzyme) from arginine, malonyl-CoA, and S-adenosylmethionine (Shimizu *et al.*, 1984, Tsuchiya *et al.*, 2014, Chun *et al.*, 2018, Kellmann *et al.*, 2008). The formation of the heterocyclic intermediate Int-C'2 is

proposed to occur by the addition of the second amidino group by SxtG (amidinotransferase), and a retroaldol-like condensation by SxtB (cytidine deaminase-like enzyme) (Kellmann *et al.*, 2008, Tsuchiya *et al.*, 2015, Tsuchiya *et al.*, 2014).

The other heterocycles are putatively formed through the catalytic activity of a sterol-like desaturase SxtD and a multifunctional enzyme SxtS (nonheme iron 2-oxoglutarate-dependent dioxygenase). A series of reactions, including epoxidation, hydroxylation, desaturation, and cyclisation, is expected to be carried out by SxtS (Kellmann *et al.*, 2008). An alcohol dehydrogenase SxtU is proposed to convert the terminal aldehyde into alcohol, forming the key intermediate Int-E'. The presence of Int-E' has been confirmed in the PST producers *Alexandrium tamarense* (Axat-2) and *Dolichospermum circinale* (TA04) (Kellmann *et al.*, 2008, Tsuchiya *et al.*, 2017). Int-E' was hydroxylated at the C12 α position by SxtT (Lukowski *et al.*, 2018) and β with an unsolved enzyme, forming dcSTX, which is then converted into STX through the activity of the O-carbamoyltransferase SxtI (Kellmann *et al.*, 2008). It is important to note, however, that the hydroxylation of Int-E' at C12β position is unclear, and does not involve SxtH as initially predicted (Lukowski *et al.*, 2018).

Related to the diversification of STX, several genes have been biochemically characterised, including SxtN (Cullen *et al.*, 2018a), SxtDIOX (Lukowski *et al.*, 2018), and SxtSUL (Lukowski *et al.*, 2019). Although the studies have provided valuable insights into the biosynthetic pathway of saxitoxin, the majority of proposed gene functions in the core biosynthesis have not been supported by biochemical characterisation, while in some cases the reaction orders remain elusive. The functions of several genes within the cluster also remain unknown. These include *sxtC*, *sxtE*, *sxtQ*, *sxtR*, *sxtJ*, *sxtK*, and *orf24*.

This project aimed to express the *sxt* biosynthetic gene cluster in *E. coli* to direct the production of neoSTX for industrial applications, and to provide insight into its biosynthetic pathway. Efficient expression of SxtA, a PKS-like enzyme that initiates saxitoxin biosynthesis has been previously achieved in *E. coli* (Chapter 3). Building on this, we utilised *E. coli* and the XylS/*Pm* promoter system for expression of the *sxt* gene cluster from *R. raciborskii* T3. Pathway engineering was designed to achieve the most efficient expression of neoSTX. This

included the re-organisation of the cluster into three transcriptional fragments fitted with a XylS/*Pm* promoter, insertion of the *E. coli* RBS, codon-optimisation for *E. coli*, and deletion of unnecessary genes or genetic regions. The genes included in the recombinant pathway were: *sxtABDGHISTUX*, which were hypothesised to be necessary for neoSTX biosynthesis; *sxtVW*, which are involved in redox reactions; and genes with unknown function *sxtCJKELQR* and *orf24*. The pathway was then integrated into the *E. coli* BL21(DE3) genome through a series of recombineering steps. The first and second transcriptional fragments encoding *sxtABC* and *sxtDEGHIJKL*, along with *T3PPTase* (a phosphopantetheinyl transferase (PPTase) from *R. raciborskii* T3) coding gene were previously integrated into the lactose, maltose, and mannitol operons, respectively (unpublished data). The present work reports the integration of the third fragment (*sxtQRSTUVWX* and *orf24*) into the recombinant *E. coli* strain and subsequent evaluation of this heterologous system for the production of PSTs.

4.2 Materials and methods

4.2.1 Bacterial strains and culture conditions

E. coli strains and plasmids used in this study are listed in Table 4.1. *E. coli* strains were maintained in lysogeny broth (LB) at either 37°C or 30°C when transformed with the thermo-sensitive plasmid. Terrific Broth (TB) was used for metabolite expression. Where appropriate, media were supplemented with antibiotics (3 µg/mL tetracycline, 15 µg/mL kanamycin or 50 µg/mL ampicillin).

Table 4.1 *E. coli* strains and plasmids used in this study.

Strains and plasmids	Genotype or relevant features	Reference
***E. coli* strains**		
BL21(DE3) *T3PPTase* sxt1 sxt2	(BL21(DE3), Δ*malEFG::T3PPTase* Δ*lacAYZ::xylSsxtABC* Δ*mtlADR::sxtDEGHIJKL*)	Unpublished
	A derivative of BL21(DE3) expression strain, integrated with *T3PPTase* (T7 promoter) and *sxt* fragment 1 and 2 (*Pm* promoter) genes from *R. raciborskii* T3 that were codon-optimised.	
BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v1	(BL21(DE3) *T3PPTase* sxt1 sxt2, Δ*xylABF*::*sxtQRorf24sxtSTUVWX*)	This study
	Variant of *sxt* fragment 3 encoding SxtQRSTUVWX and ORF24 (under the control of *Pm* promoter) inserted into the xylose operon.	
BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v2	(BL21(DE3) *T3PPTase* sxt1 sxt2 Δ*xylABF::sxtQRSTUVWX*)	This study
	Variant of *sxt* fragment 3 encoding SxtQRSTUVWX (under the control of *Pm* promoter) inserted into the xylose operon.	
BL21(DE3) *T3PPTase* sxt1 sxt2 sxt v3	(BL21(DE3) *T3PPTase* sxt1 sxt2 Δ*xylABF::sxtSTUVWX*)	This study
	Variant of *sxt* fragment 3 encoding SxtSTUVWX (under the control of *Pm* promoter) inserted into the xylose operon.	
BL21(DE3) *sfp* NSXv1-3	(BL21(DE3) sxt1 sxt2 sxt3, Δ*malEFG::sfp*)	This study
	Derivatives of BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3 v1, v2, and v3, with *T3PPTase* gene was replaced with *sfp* gene from *E. coli* BAP1 strain.	
T-SACK	(W3110 *araD*::*tetA-sacB*-amp *fliC*::*cat argG*::Tn5)	(Li *et al.*, 2013)
	Template for *tetA-sacB* cassettes amplification.	

Plasmids		
pRedET	Low copy number plasmid, with thermo-sensitive pSC101 replicon ($repA^{ts}$), expressing Red/ET recombinases genes (under the control of p_{araBAD} promoter) and tetracycline resistance gene (tet^R).	GeneBridges
p707- FLPe	FLPe expression plasmid encoding FLPe recombinase and tetracycline resistance gene (tet^R).	GeneBridges
pMK-RQ-3- 1	GeneArt cloning vector containing a homology region to *xylF* in xylose operon, *Pm* promoter, and codon-optimised genes from *R. raciborskii* T3 (*sxtQ* and *sxtR)*. The plasmid was used as a PCR template for *sxt* 3 fragment construction. Synthesised by LifeTechnology.	This study
pMK-RQ-3-2	GeneArt cloning vector containing codon-optimised genes from *R. raciborskii* (*sxtR*, *ORF24*, *sxtS*, *sxtT*, and fragment of *sxtU*). The plasmid was used as a PCR template for *sxt* 3 fragment construction. Synthesised by LifeTechnology.	This study
pMK-RQ-3-3	GeneArt cloning vector containing codon-optimised genes from *R. raciborskii* (fragment of *sxtU*, *sxtV*, *sxtW*, *sxtX*), kan^R, FRT sites and homology region to *xylB* in the xylose operon. The plasmid was used as a PCR template for *sxt* 3 fragment construction. Synthesised by LifeTechnology.	This study
pSIM5	Low copy number plasmid, with thermo-sensitive pSC101 replicon, expressing recombinase genes and chloramphenicol resistance gene (cm^R). The recombinase genes were expressed from the p_L operon under the control of temperature sensitive (ts) repressor CI857.	(Datta *et al.*, 2006)

4.2.2 Assembly of saxitoxin (*sxt*) fragment 3 and integration into the *E. coli* BL21(DE3) T3PPTase sxt1 sxt2 genome

Previous work has successfully integrated amp^R and *T3PPTase* (under the control of T7 promoter) into the mannitol operon, *sxt* fragment 1 (*xylS* and *sxtABC*, under the control of XylS/*Pm* promoter) into the lactose operon, and *sxt* fragment 2 (*sxtDEGHIJKL*, under the control of XylS/*Pm* promoter) into the maltose operon using the Quick & Easy *E. coli* Gene Deletion Kit (Cat. K006, Gene Bridges), generating the recombinant strain *E. coli* BL21(DE3) T3PPTase sxt1 sxt2. The present study reports the integration of the codon-optimised

(LifeTechnology, using GeneArt software) *sxt* fragment 3 (*sxtQRorf24sxtSTUVWX,* under the control of XylS/*Pm* promoter) into the xylose operon of *E. coli* BL21(DE3) T3PPTase sxt1 sxt2. The cloning strategies used are illustrated in Figure 4.2.

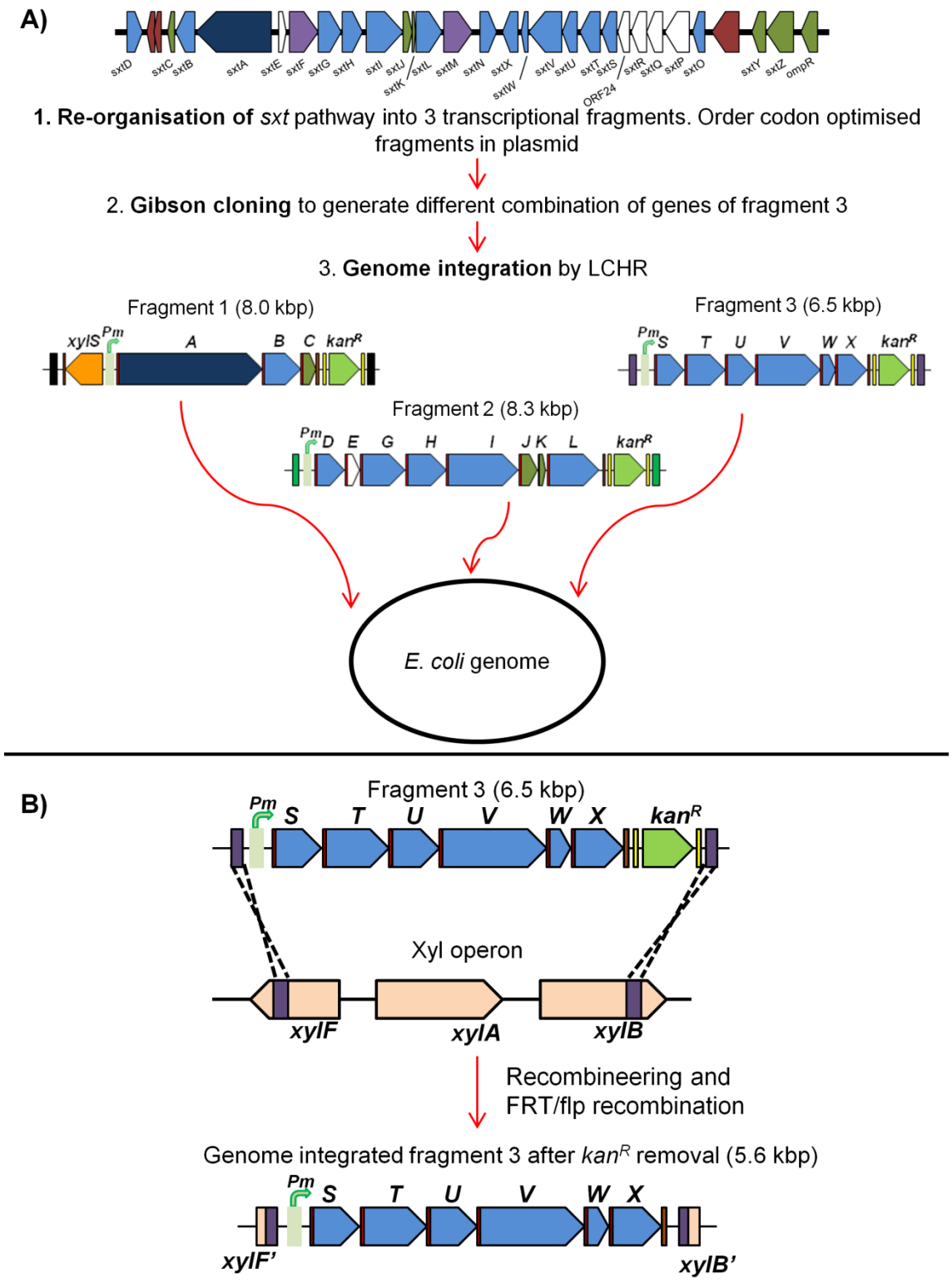


Figure 4.2 Cloning approach utilised for the integration of saxitoxin biosynthetic (*sxt*) gene cluster in *E. coli*. (**A**) Cloning steps to integrate *sxt* fragment 1-3 into the *E. coli* genome. (**B**) The linear *sxt* fragment 3 with 50 bp homology regions to *xylF* and *xylB* genes in the xylose operon of *E. coli* was integrated following recombineering. The kan^R cassette was removed following FRT/FLP recombination.

The *sxt* fragment 3 was synthesised by LifeTechnology in three separate plasmids (Figure 4.3A). pMK-RQ-3- 1 (containing a 50 bp homology region to *xylF*, *Pm* promoter, *sxtQ*, and a fragment of *sxtR*), pMK-RQ-3- 2 (containing a fragment of *sxtR*, *orf24*, *sxtS*, *sxtT*, and a fragment of *sxtU*), and pMK-RQ-3- 3 (a fragment of *sxtU*, *sxtV*, *sxtW*, *sxtX*, kan^R, FRT sites, and a 50 bp homology region to *xylB*). The assembly of *sxt* fragment 3 was performed using Gibson Assembly (New England Biolabs). Three variants of linear constructs were generated, encoding *sxtQRorf24sxtSTUVWX* (v1), *sxtQRSTUVWX* (v2), and *sxtSTUVWX* (v3). All constructs contain the kanamycin resistance gene (kan^R) as a selective marker, flanked by the FRT sites. Briefly, linear fragments from each sub-fragment were amplified using Phusion DNA polymerase (New England Biolabs) and primers designed for Gibson assembly. The primers used in this study are listed in Appendix Table A 3.1. Gibson assembly was performed as per the manufacturer's protocol. The assembled linear fragments then served as PCR templates to generate inserts for *E. coli* genome integration.

The linear fragment constructs contained a 50 bp homology region to the 5' and 3' ends of the xylose operon (*xylF* and *xylB*) in the *E.coli* genome. The inserts were amplified using Phusion DNA polymerase (New England Biolabs) with primers xylFH1F and xylBH2R (Appendix Table A 3.1). Integration of *sxt* fragment 3 was performed using the pRed/ET(tet) recombinase plasmid based on the Quick & Easy *E. coli* Gene Deletion kit (Gene Bridges). The manufacturer's instructions were followed with some modifications. The Red/ET expression plasmid pRedET, containing λ phage redγβα operon, was transformed into the electrocompetent *E. coli T3PPTase* sxt1 sxt2 strain. A recombineering-proficient strain was then electroporated with 200 ng of the purified linear DNA fragment and incubated at 37°C for 3 h with shaking at 200 rpm, before plating under the selection of kanamycin. As pRed/ET has a thermo-sensitive pSC101 replicon, the plasmid is lost upon incubation at 37°C. The recombinants were screened by culturing on simplified differential MacConkey agar plates (Appendix Supplementary material A 3.1). Successful gene insertion was indicated by white colony morphology due to the lack of acid formation from xylose fermentation as the cells have the disrupted xylose operon.

Positive clones were screened by colony PCR using primers flanking the insertion site (xylF4232_4249F and xylB_7987-7970R2) and within *sxtQ* (sxtQ_R) and *sxtT* (sxt3_5539-5557R). The inserted genes were amplified using OneTaq (New England Biolabs) and sequenced using the BigDye Terminator v3.1 cycle sequencing kit and an ABI 3730XL DNA analyser sequencer (Applied Biosystems) at the Molecular Biology Institute (University of Bergen, Norway). The resulting recombinant *E. coli* variants from this cloning procedure were *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v1 (s*xtQRorf24sxtSTUVWX*), *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v2 (s*xtQRSTUVWX*), and *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v3 (*sxtSTUVWX*) (Figure 4.3B).

The inserted selective marker (kan^R) was flanked by FRT-sites to assist its removal by FLP-mediated site-directed recombination using plasmid p707- FLPe (Gene Bridges). Briefly, temperature sensitive p707- FLPe plasmid was transformed into the electrocompetent recombinant strains. Transformants were initially grown in 1 mL of LB medium without antibiotics at 30°C for 3 h. The temperature was then increased to 37°C and cells were incubated overnight. During this period, the FLP/FLPe protein mediated the FRT recombination removing the kan^R cassette, and the expression plasmid lost due to the inability to replicate at 37°C. The culture was then plated onto a fresh LB agar plate, and successful antibiotic marker removal was confirmed by a kanamycin sensitivity test and colony PCR.

4.2.3 Replacement of the *T3PPTase* with *sfp*

The *T3PPTase* gene in BL21(DE3) *T3PPTase* sxt1 sxt2 sxt v1-3 was replaced with the *sfp* gene originated from *B. subtilis*. To prepare the linear fragment of *sfp* under the control of T7 promoter for genome integration, the expression plasmid pET30b::*sfp* was constructed using restriction enzyme cloning.

The *sfp* gene was amplified from the genomic DNA of *E. coli* BAP1 (Pfeifer *et al.*, 2001) using Phusion DNA polymerase (New England Biolabs). The PCR amplicon was purified (DNA Clean and Concentrator 5X Kit, ZymoResearch) and treated with BamHI and NdeI restriction enzymes (Thermo Fischer Scientific). The expression vector pET30b was digested

with the same restriction enzymes. Treated PCR products and the digested vector backbone were purified by gel electrophoresis (Gel DNA Recovery Kit, ZymoResearch), and ligated with T4 DNA ligase (New England Biolabs). *E. coli* DH5α was then chemically transformed with the ligated plasmids, and the resulting colonies screened by colony PCR using the universal T7 promoter and T7 terminator primers. The inserts were sequenced as described previously.

The *T3PPTase* genes that were previously integrated into the mannitol operon of *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3 v1, v2, and v3 (constructed in section 4.2.2) were replaced with the *sfp* gene. Seamless homologous recombination was performed following a two-step recombineering approach utilising the *tetA-sacB* counter-selection cassette as previously described by Li et al.(Li *et al.*, 2013)(see Appendix Figure A 3.1).

In the first recombineering step, the *tetA-sacB* cassette was amplified from *E. coli* T-SACK, and kindly provided by Donald L. Court (Gene Regulation and Chromosome Biology Laboratory, National Cancer Institute at Frederick, Frederick, MD) using primers tetA_mtlA_ F and sacB_mtlD_R. The targeted *E. coli* strain was transformed with the pSIM5 plasmid (provided by Donald L. Court) containing the Red recombinase protein system used for the dsDNA recombination. The pSIM5 plasmid replicates under the control of the temperature sensitive (ts) repressor CI857 and is lost under high temperatures (37°C). The purified DNA fragment was electroporated into the recombineering proficient cells. Colonies were screened by colony PCR, and successful recombination was confirmed by testing sensitivity to sucrose.

In the second recombineering round, the *tetA-sacB* cassette was replaced with the *sfp* gene from *B. subtilis*. Briefly, the *sfp* gene was PCR-amplified from pET30b::*sfp* constructed above using primers T7 sfp _ mtlA_F/T7term_ mtlD_R. The purified DNA fragment was electroporated into the recombineering proficient cells, prepared as described above, and was subsequently grown on TetA-SacB counterselection medium (Li *et al.*, 2013) for 2-3 days at 42°C. The medium contained fusaric acid and 10% sucrose to screen for the insertion of *sfp* gene, which was further confirmed by colony PCR. The insert was sequenced using the BigDye Terminator v3.1 cycle sequencing kit and an ABI 3730XL DNA analyser sequencer (Applied Biosystems) at the Molecular Biology Institute (University of Bergen). The substitution of

T3PPTase with *sfp* resulted in the generation of *E. coli* BL21(DE3) *sfp* sxt1 sxt2 sxt3v1 (s*xtQRorf24sxtSTUVWX*), *E. coli* BL21(DE3) *sfp* sxt1 sxt2 sxt3v2 (*sxtQRSTUVWX*), and *E. coli* BL21(DE3) *sfp* sxt1 sxt2 sxt3v3 (*sxtSTUVWX*), herein named *E. coli* BL21(DE3) *sfp* NSXv1, *E. coli* BL21(DE3) *sfp* NSXv2, and *E. coli* BL21(DE3) *sfp* NSXv3, respectively. Genome integration of *sfp* into the mannitol operon of *E. coli* BL21(DE3) was also performed to construct *E. coli* BL21(DE3) *sfp*, which was used in subsequent fermentation experiments as a negative control strain for neoSTX production.

4.2.4 Heterologous expression of the neosaxitoxin pathway in *E. coli*

The heterologous production of PSTs in *E. coli* was investigated by LC-MS analysis, and by bioassay using enzyme-linked immunoadsorbent assay (ELISA) and mouse neuroblastoma bioassay (MNBA).

For LC-MS analysis, crude extracts of recombinant *E. coli* strains harbouring *sxt* biosynthetic cluster were analysed. Starter cultures of *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3 v1, v2, and v3 were grown overnight in LB medium at 30°C shaking at 200 rpm. The culture was inoculated 1:100 into 25 mL of TB (pH 6.8) medium in 100 mL Erlenmeyer flasks stoppered with cotton bungs. Cultures were incubated at 30°C with shaking (200 rpm) to an OD_{600} of ~0.7, before transferring to 19°C until an OD_{600} of ~0.8 was reached. The expression of *T3PPTase* and *sxt* genes was induced through the addition of 0.2 mM IPTG and 1 mM toluic acid, and the cultures were incubated at 19°C with shaking (200 rpm) for a further 48 h. Cells were harvested by centrifugation and pellets were washed with chilled MilliQ water twice, before being stored at -20°C until further processing. Cultures of *E. coli* BL21(DE3) *T3PPTase* were used as negative control. Intracellular metabolites were extracted from the cell pellets with 0.8 ml of 0.1 M acetic acid in methanol. The cell lysate was clarified by centrifugation (16,000 x g for 10 min), and the supernatant was diluted 1:5 to give a final solvent composition of 8% methanol, 72% acetonitrile, and 0.1% formic acid. Samples were clarified by centrifugation and transferred into HPLC vials for LC-MS analysis.

For ELISA and MNBA bioassays, expression of neoSTX pathway in *E. coli* BL21(DE3) *sfp* NSXv3 was performed as above, except with an upscaled culture volume of 50 mL and inducer concentration of 0.05 mM IPTG and 0.5 mM toluic acid added at an OD_{600} of ~0.5. A culture of *E. coli* BL21(DE3) *sfp* was used as a negative control. Intracellular metabolites were extracted and purified using solid phase extraction (SPE) as described in section 4.2.7 and 4.2.8.

4.2.5 Transcript analysis of *E. coli* BL21(DE3) *sfp* NSXv3

Total RNA was isolated from *E. coli* BL21(DE3) *sfp* NSXv3 cell pellets, previously lysed with TRI reagent (Sigma), using the Direct-zol RNA Miniprep kit (Zymo Research) according to the manufacturer's instructions. Contaminating genomic DNA was digested using the Turbo DNA-Free kit (Ambion). The absence of contaminating DNA was confirmed via PCR with primers targeting the 16S rRNA gene. cDNA synthesis was then performed using the ProtoScript II First Strand cDNA Synthesis Kit (New England Biolabs) following the manufacture protocol with 200 ng of total RNA. PCR was performed using BIO*Taq* DNA polymerase with primers targeting 16S rRNA, *sfp* and several *sxt* genes listed in Table A 3.1.

4.2.6 HPLC-ESI-HRMS analysis of the saxitoxin pathway products

The products of the STX biosynthetic pathway were analysed by LC-MS using a Dionex 3000 Ultimate RSLC coupled to a Q-Exactive MS (Thermo Fischer Scientific) via a heated electrospray interface. Liquid chromatography was performed on a 40 µL sample using a Synchronis HILIC 2.1 x 50 mm column with a 1.7 µm particle size (Thermo Fisher Scientific). The separation was achieved using a linear gradient of acetonitrile (mobile phase D) with 10 mM ammonium formate pH = 3.9 (C) against 10 mM ammonium formate pH = 3.9 at a flow rate of 500 µL/min. The column temperature was held at 40°C. The gradient program was as follows: 0-3 min 5 % C, 3-10 min 5%-70% C, 10-12 min 70% C, 12-15 min 5% C. Ionisation was performed in positive mode (3.5 kV), and mass spectra were acquired in SIM mode with a range of 299-317 *m/z* and a resolution of 140,000. All data were processed using the Xcalibur software (Thermo Scientific).

4.2.7 Analysis of heterologously produced saxitoxins by enzyme-linked immunosorbent assay (ELISA)

The diagnostic test of toxin production was performed together with Dr Ralf Kellmann at the University of Bergen using a Saxitoxin (PSP) ELISA kit (Abraxis). Several saxitoxin (PSP) ELISAs have been developed as rapid and sensitive tools to detect PSTs from environmental samples (Huang *et al.*, 1996, Dubois *et al.*, 2010, Metcalf & Codd, 2003). The assay used in this study is based on the competitive binding between saxitoxin in the sample with the saxitoxin-horseradish peroxidise (STX-HRP), which is a conjugate of a rabbit anti-saxitoxin antibody immobilised on the microtiter plate. After the addition of the substrate, the formation of blue colour is inversely proportional to the concentration of saxitoxin present.

Intracellular metabolites from *E. coli* BL21(DE3) *sfp* NSXv3 were extracted and purified by solid phase extraction prior to the ELISA test. Metabolites from *E. coli* BL21(DE3) *sfp* NSXv3 and BL21(DE3) *sfp* cultured in 50 mL TB (~1.2g) (Section 4.2.4) were extracted in 0.1% formic acid with 1:4 v/w (volume: wet weight) ratio and lysed by boiling at 95°C for 5 min. The cell lysate was then clarified by centrifugation (16,000 x g for 10 min) and was purified by solid phase extraction on SampliQ silica columns (Agilent, 1 ml, 100 mg) using the method of Jansson *et al* (Jansson & Astot, 2015). Purified extracts were evaporated to dryness by vacuum centrifugation, and reconstituted in 100 µL of Saxitoxin (PSP) ELISA sample buffer included in the kit. The assay was calibrated using standard solutions of Std0 (blank), Std1 (0.0668 nM STX), and Std5 (1.3365 nm STX) included in the kit. A reference sample containing 100 nM neoSTX was used as a positive control.

4.2.8 Analysis of heterologously produced saxitoxins by mouse neuroblastoma bioassay (MNBA)

A mouse neuroblastoma bioassay (MNBA) was performed on the Neuro-2a cell line to assess the production of PSTs. As PST native producers are freshwater cyanobacteria and dinoflagellates, the assays developed previously were optimised for the relevant samples (cyanobacterial extracts, freshwater samples and of shellfish samples) (Jellett *et al.*, 1992,

Okumura *et al.*, 2005, Hayashi *et al.*, 2006, Humpage *et al.*, 2007). Purification of PSTs by SampliQ silica columns described by Jansson *et al.*(Jansson & Astot, 2015) was not compatible with the MNBA assay, as the extracts induced cell death likely due to the high levels of sodium retained in the extracts. In the present study, optimisation of the assay for *E. coli* extracts as the biological matrix was conducted.

The optimised method used for MNBA bioassay was as follows. Intracellular metabolites of BL21(DE3) *sfp* NSXv3 were purified on a silica-based hydrophilic weak cation exchange column (Sep-Pak Accell Plus CM Cartridge, 360 mg, Waters). Metabolites from *E. coli* BL21(DE3) *sfp* NSXv3 and BL21(DE3) *sfp* cultured in 50 mL TB (~1.2 g) (Section 4.2.4) were extracted in 0.1% formic acid with 1:3 v/w (volume: wet weight) ratio and lysed by boiling at 95°C for 5 min. The cell lysate was clarified by centrifugation, and the supernatant was adjusted to a final concentration of 0.15% formic acid. The solid phase extraction column matrix was equilibrated with 3.6 mL methanol, followed by 3.6 mL 0.15% formic acid. A total volume of 3.6 mL of sample was loaded into the column and then washed with 1.8 mL MilliQ and 1.8 mL acetonitrile. Metabolites were eluted in 7.2 mL MeOH/5% formic acid and evaporated to dryness. Purified extracts were reconstituted in 100 µL 0.01% formic acid. Extracts from BL21(DE3) *sfp* were used as a negative control, whereas the extract of BL21(DE3) *sfp* spiked with 10 pmol neoSTX (final concentration in the extract was 100 nm) before purification was used as a positive control. A standard curve was constructed using an extract from BL21(DE3) *sfp* strain spiked with neoSTX standard to a concentration of 0-800 nM.

The mouse neuroblastoma bioassay was performed on Neuro-2a cells (American Tissue Culture Collection) using the method developed by Humpage et al. (Humpage *et al.*, 2007). Neuro-2a cells were cultured in RPMI 1640 Medium with GlutaMAX™ Supplement (Gibco, Life Technologies) supplemented with 10% Fetal Bovine Serum (FBS) (Sigma Aldrich), 1% sodium pyruvate (Sigma-Aldrich), and 1% penicillin-streptomycin (Sigma-Aldrich). Cell cultures were incubated at 37°C with 5% CO_2 in a humidified atmosphere.

For the assay, Neuro-2a cells were inoculated into a 96-well plate to a cell density of 10,000 cells per/well and were incubated for 24 h. The culture medium was removed, and the cells were exposed to the toxin mixture containing final concentrations of 50 µM veratridine and 500 µM ouabain. Purified *E. coli* intracellular extract was added (5 µL per well) for treated cells. The cells were given fresh medium before being incubated for a further 24 h. Cell viability was measured by the MTT (thiazolyl blue tetrazolium bromide) colourimetric assay. The assay measures mitochondrial reductase activity (Mosmann, 1983) and was therefore used for assessing cell viability and proliferation.

4.3 Results

4.3.1 Integration of the neosaxitoxin biosynthetic (*sxt*) gene cluster in *E. coli*

Previous work by Alessandra Estaquio, Antonio Garcia Moyano and Sudeep Karki (University of Bergen, Norway) successfully integrated parts of the *sxt* biosynthetic gene cluster into the *E. coli* genome. The resulting recombinant *E. coli* strain, *E. coli* BL21(DE3) T3PPTase sxt1 sxt2, did not produce neosaxitoxin. The present study describes the integration of *sxt* fragment 3 which encodes the remaining genes in the pathway of *sxtQR*ORF24*sxtSTUVWX*.

The *sxt* fragment 3 encoded SxtS (nonheme iron 2-oxoglutarate-dependent dioxygenase) and SxtU (alcohol dehydrogenase), which are enzymes involved the formation of heterocycle rings; SxtT (phenylpropionate dioxygenase) which is required for C12-hydroxylation; and SxtX (cephalosporin hydroxylase), which is required for the conversion of STX to neoSTX (see Chapter 1 Figure 1.5 for the proposed saxitoxin biosynthesis pathway). Other enzymes, namely SxtV (succinate dehydrogenase) and SxtW (ferredoxin), were also included as they are proposed to mediate biosynthesis by performing redox reactions. Several genes within the *sxt* biosynthetic cluster in *R. raciborskii* T3 encode the proteins SxtQ, SxtR, and ORF24, which have unknown function. These genes are also detected in several other *sxt* clusters, and yet their involvement within the pathway remains elusive. No homology or predicted functions were assigned for both ORF24 and SxtQ, while SxtR is a homolog of

acyltransferase and is possibly involved in regulation or analogue formation (Kellmann *et al.*, 2008, Mihali *et al.*, 2009). As such, genes encoding SxtQ, SxtR, and ORF24 were included in the *sxt* fragment 3; however, variants of recombinant *E. coli* strains were constructed in order to investigate the necessity of these genes in neoSTX biosynthesis (Table 4.3).

Table 4.2 List of *E. coli* strains for neoSTX pathway engineering. The strain BL21(DE3) *sfp* NSXv3 was used as the parent strain for gene deletion and insertion. The right column lists the alteration compared to Sfp NSXv3 strain.

***E. coli* BL21(DE3)**	**Genotype**	**Variations to *sfp* NSXv3**
Sfp NSXv1	Δ*lacAYZ::xylSsxtABC*	Insertion of *sxtQRorf24*
(20 genes)	Δ*mtlADR::sxtDEGHIJKL*	
	Δ*xylABF::sxtQR*ORF24*sxtSTUVWX*	
Sfp NSXv2	Δ*lacAYZ::xylSsxtABC*	Insertion of *sxtQR*
(19 genes)	Δ*mtlADR::sxtDEGHIJKL*	
	Δ*xylABF::sxtQRSTUVWX*	
Sfp NSXv3	Δ*lacAYZ::xylSsxtABC*	
(17 genes)	Δ*mtlADR::sxtDEGHIJKL*	
	Δ*xylABF::sxtSTUVWX*	

The *sxt* fragment 3 containing *sxtQR*ORF24*sxtSTUVWX* was synthesised by LifeTechnology in three separate plasmids (Figure 4.3A) pMK-RQ-3- 1 (containing a 50 bp homology region to *xylF*, *Pm* promoter, *sxtQ*, and a fragment of *sxtR*), pMK-RQ-3- 2 (containing a fragment of *sxtR*, *orf24*, *sxtS*, *sxtT*, and a fragment of *sxtU*), and pMK-RQ-3- 3 (a fragment of *sxtU*, *sxtV*, *sxtW*, *sxtX*, kan^R, FRT sites, and a 50 bp homology region to *xylB*). As the *sxt* fragment 3 encodes proteins of unknown function (SxtQ, SxtR, and ORF24), the one-pot Gibson assembly was performed to generate linear variants of *sxt* 3 fragments namely v1 (*sxtQR*ORF24*sxtSTUVWX*), v2 (*sxtQRSTUVWX),* and v3 (*sxtSTUVWX*) (Figure 4.2B). This allowed an investigation into the functions of the unknown genes.

Each linear fragment variant was then integrated into *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 through recombineering with the dispensable xylose sugar operon as integration sites {Albermann, 2010 #67}. Successful gene insertion was indicated by white colony morphology due to the lack of acid formation from xylose fermentation as the cells have the disrupted xylose operon. Upon integration of the linear fragments of sxt3, three *E. coli* strains were constructed, namely *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v1 (s*xtQRorf24sxtSTUVWX*), *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v2 (*sxtQRSTUVWX*), and *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3v3 (*sxtSTUVWX*). Successful integration was confirmed by PCR amplification

using one primer targeting a region of the *E. coli* genome (xylF4232_4249F) and the second targeting the inserted gene (sxt3_5539-5557R), allowing confirmation of each construct variant (Figure 4.4, for primer binding sites and the expected amplicon size refer to Figure 4.3B).

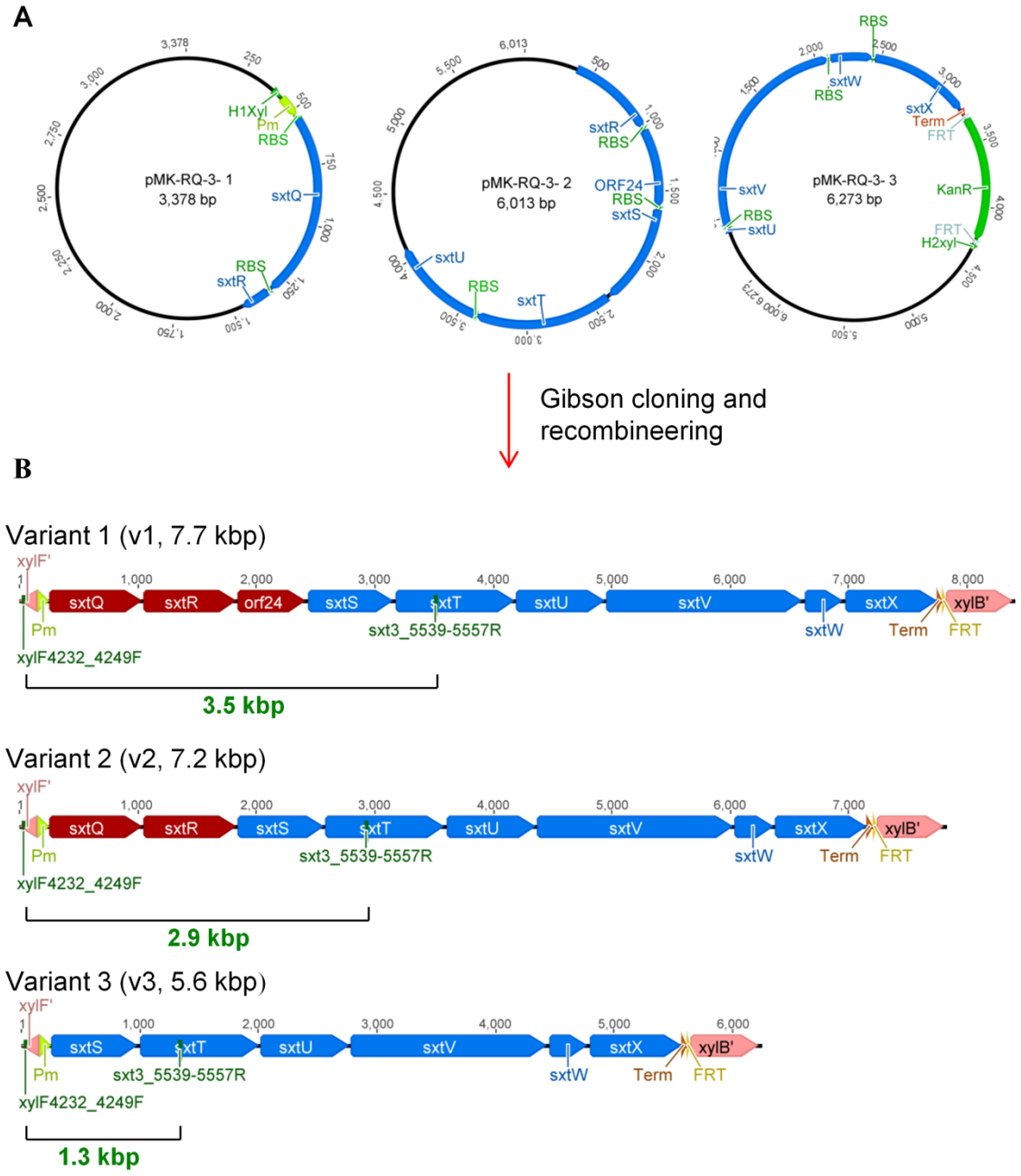


Figure 4.3 Re-assembly and integration of *sxt* fragment 3. (**A**) Plasmid map of GeneArt vector plasmids pMK-RQ encoding *sxt* fragment 3 sub-fragments. (**B**) Three variants of re-assembled *sxt* fragment 3 integrated into the xylose operon of *E. coli*, the gene variations are highlighted in red. Diagnostic primers of xylF4232_4249F and sxt3_5539-5557R were used to test the successful construction of each variant. The binding location for each primer along with the size of the PCR products are denoted in green in panel B. Construct map was generated using Geneious 9.0.4 (Biomatters).

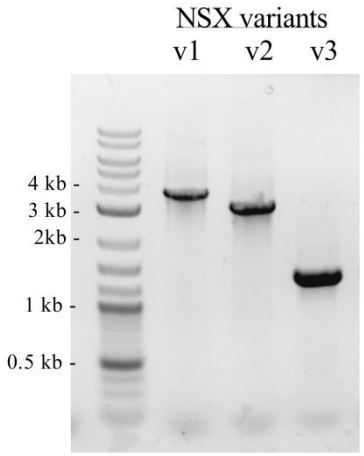


Figure 4.4 Integration of *sxt* fragment 3 creation of *E. coli* BL21(DE3) T3PPTase sxt1 sxt2 sxt3 variants. PCR amplification of strain variants showing the correct insertion and variations of gene composition. Lane 1, 2-log DNA ladder. Lane 2-4, PCR amplification of *sxt* fragments (variants 1-3) integrated into the *E. coli* genome and flanking regions. PCR amplification was performed using xylF4232_4249F and sxt3_5539-5557R primers. The binding location for each primer is denoted in green in panel B in Figure 4.3. Using these primers, the expected sizes of amplicons are 3.5 kbp for *E. coli* BL21(DE3) T3PPTase sxt1 sxt2 sxt3 variant 1, 2.9 kbp for *E. coli* BL21(DE3) T3PPTase sxt1 sxt2 sxt3 variant 2, and 1.3 kbp for *E. coli* BL21(DE3) T3PPTase sxt1 sxt2 sxt3 variant 3.

PKS enzymes require a post-translational modification by the phosphopantetheinyl transferase (PPTase) enzyme family. PPTases catalyse the transfer of a phosphopantetheinyl (Ppant) moiety onto a conserved serine residue in the acyl carrier protein domain, which acts as a tethering arm during polyketide assembly (Lambalot *et al.*, 1996, Walsh *et al.*, 1997a). Previous studies have demonstrated the optimisation of natural products yields through PPTase modulation (Jiang *et al.*, 2013, Zhang *et al.*, 2017). A broad range PPTase from *Bacillus subtilis*, Sfp, has been shown to be efficient for the activation of various PKS and NRPS enzymes (Trauger & Walsh, 2000, Pfeifer *et al.*, 2001, Liu *et al.*, 2018), including those of cyanobacterial origin (Liu *et al.*, 2018, Liu *et al.*, 2017). In an attempt to increase neoSTX production, the *T3PPTase* gene in BL21(DE3) *T3PPTase* sxt1 sxt2 sxt v1-3 was replaced with the *sfp* gene which was known to be active towards SxtA resulting in the efficient production of Int-A' (Chapter 3). The

substitution was performed using a two-step recombineering, utilising *tetA-sacB* as the selective/counterselective cassette (see Appendix Figure A 3.1). The TetA protein could act as both a positive and negative selection marker depending on the media composition, i.e. the TetA protein increases cell sensitivity to lipophilic chelators, such as fusaric acid, and also confers tetracycline resistance (Stavropoulos & Strathdee, 2001, Bochner *et al.*, 1980, Maloy & Nunn, 1981). On the other hand, *sacB* encodes a levansucrase, which confers sucrose sensitivity in *E. coli* (Gay *et al.*, 1983, Steinmetz *et al.*, 1983). In the first recombineering step inserting the *tetA-sacB* cassette, colonies resistant to tetracycline were confirmed by colony PCR, and successful recombination was confirmed by testing sensitivity to sucrose. While in the second recombineering step, replacement of *tetA-sacB* cassette with the *sfp* gene was confirmed by growing the colonies on fusaric acid and sucrose containing media, followed by colony PCR. The resulting recombinant strains were BL21(DE3) *sfp* sxt1 sxt2 sxt v1-3, which from here on are referred to as BL21(DE3) *sfp* NSXv1-3.

4.3.2 Confirmation of neosaxitoxin production by recombinant *E. coli* variants using HPLC-ESI-HRMS

The production of neosaxitoxin by recombinant *E. coli* variants was investigated by HPLC-ESI-HRMS. The presence of a molecular ion with an *m/z* of 316.13639 at 4.2-4.5 min was observed from the intracellular crude extracts of all variants of *E. coli* T3PPTase sxt1 sxt2 sxt3 (Figure 4.5B, C, and D). The levels of abundance were in a similar range. The elution profile of ions with an *m/z* of 316.13639 was consistent with that of the authentic neosaxitoxin standard (Figure 4.5E), while the corresponding peak was absent in chromatograms generated for extracts of the negative control strain of BL21(DE3) *T3PPTase* (Figure 4.5A). This data strongly suggests that neosaxitoxin is successfully produced by the recombinant *E. coli* strains, albeit at low levels. This finding was further supported by saxitoxin-ELISA and a mouse neuroblastoma bioassay (MNBA) (see section 4.3.4).

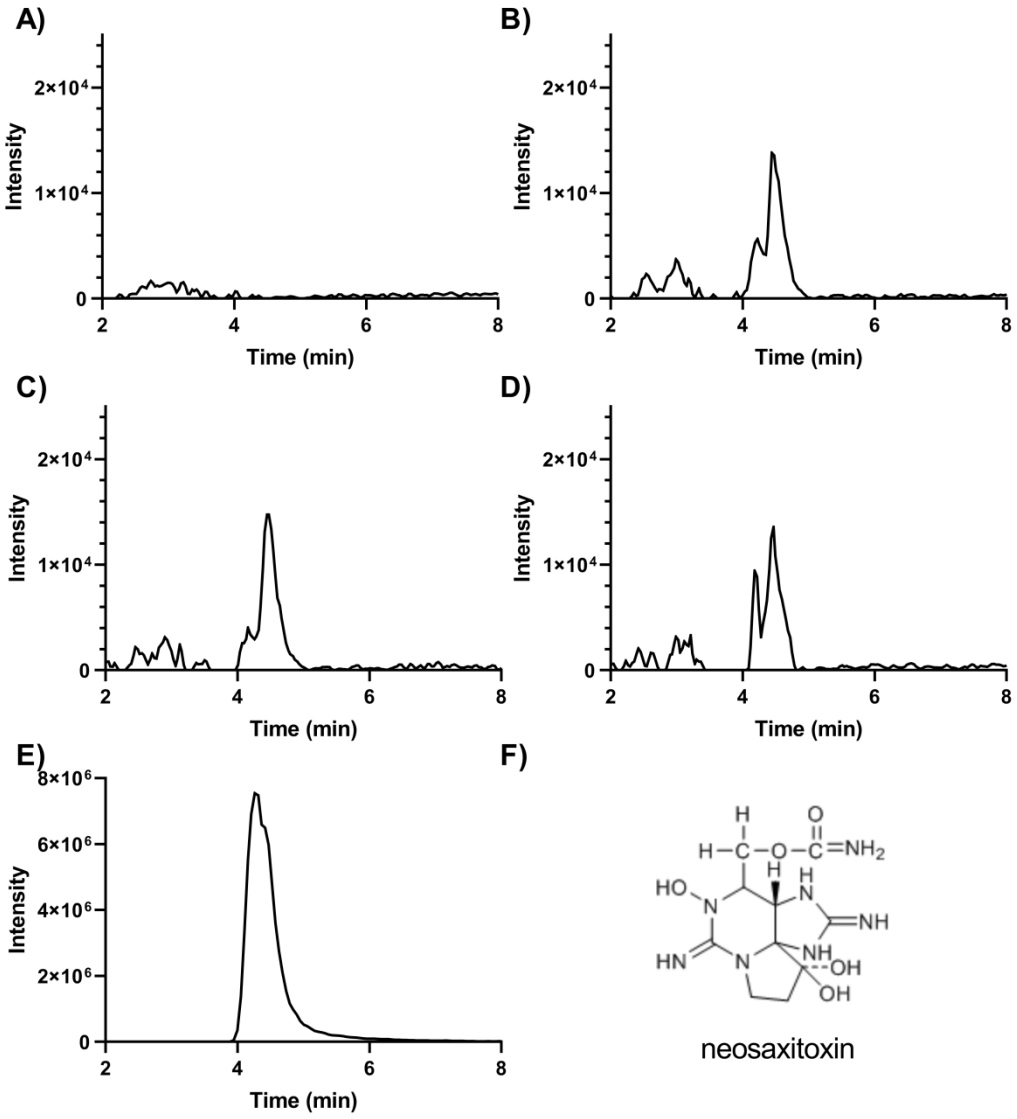


Figure 4.5 HPLC-ESI-HRMS analysis of neosaxitoxin production by *E. coli* variants. Extracted ion chromatograms of *m/z* = 316.13639 (± 5 ppm) with a retention time of 2-8 min, from the crude extracts of **A**) control strain BL21(DE3) *T3PPTase*, **B**) BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3 v1, **C**) BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3 v2, and **D**) BL21(DE3) *T3PPTase* sxt1 sxt2 sxt3 v3. **E**) Extracted ion chromatogram of *m/z* = 316.13639 (± 5 ppm) with a retention time of 2-8 min from the neosaxitoxin standard. **F**) Chemical structure of neosaxitoxin. The *E. coli* culture was grown in the presence of IPTG and toluic acid.

4.3.3 Transcript analysis of *E. coli* BL21(DE3) *sfp* NSXv3

To determine transcription of the complete integrated saxitoxin biosynthetic cluster in *E. coli* BL21(DE3) *sfp* NSXv3, PCR amplification from bacterial cDNA was performed targeting each mRNA fragment. PCR amplification of the *sfp* gene was also performed to test the functionality of the T7 promoter. The genes amplified were *sxtA* and *sxtC* (*sxt* fragment 1); *sxtD, sxtI*, and

sxtL (*sxt* fragment 2); as well as *sxtS*, *sxtU*, and *sxtX* (*sxt* fragment 3). PCR amplifications of these genes using genomic DNA of *E. coli* BL21(DE3) *sfp* NSXv3 (positive control) resulted in amplicons that corresponded to the correct theoretical sizes. On the other hand, no PCR amplification occurred using cDNA of the negative control strain, *E. coli* BL21(DE3) *sfp*. The negative control confirmed the specificity of the primers and PCR conditions used (Figure 4.6).

Transcripts of *sfp* and *sxt* genes were detected in induced cells of *E. coli* BL21(DE3) *sfp* NSXv3 (Figure 4.6). All amplicons from the cDNA corresponded to the sizes determined by PCR amplification of the same genes from genomic DNA of *E. coli* BL21(DE3) *sfp* NSXv3 with the same primer sets (Appendix Figure A 3.2). Amplification of all targeted *sxt* genes was also detected from the cDNA of uninduced *E. coli* BL21(DE3) *sfp* NSXv3 cDNA. This is most likely due to the leakiness of *Pm* promoter in *E. coli*, which was also observed in Chapter 3. Amplification of *sxtS* from the uninduced sample was minimal and only visible with higher loading volumes (see Appendix Figure A 3.2 for visualisation of low concentration amplicons). The level of amplification detected from the cDNA of both induced and uninduced *E. coli* BL21(DE3) *sfp* NSXv3 varied between target *sxt* genes (Figure 4.6). It is possible that the codon-optimised constructs contain internal promoters which might lead to the transcription of downstream genes.

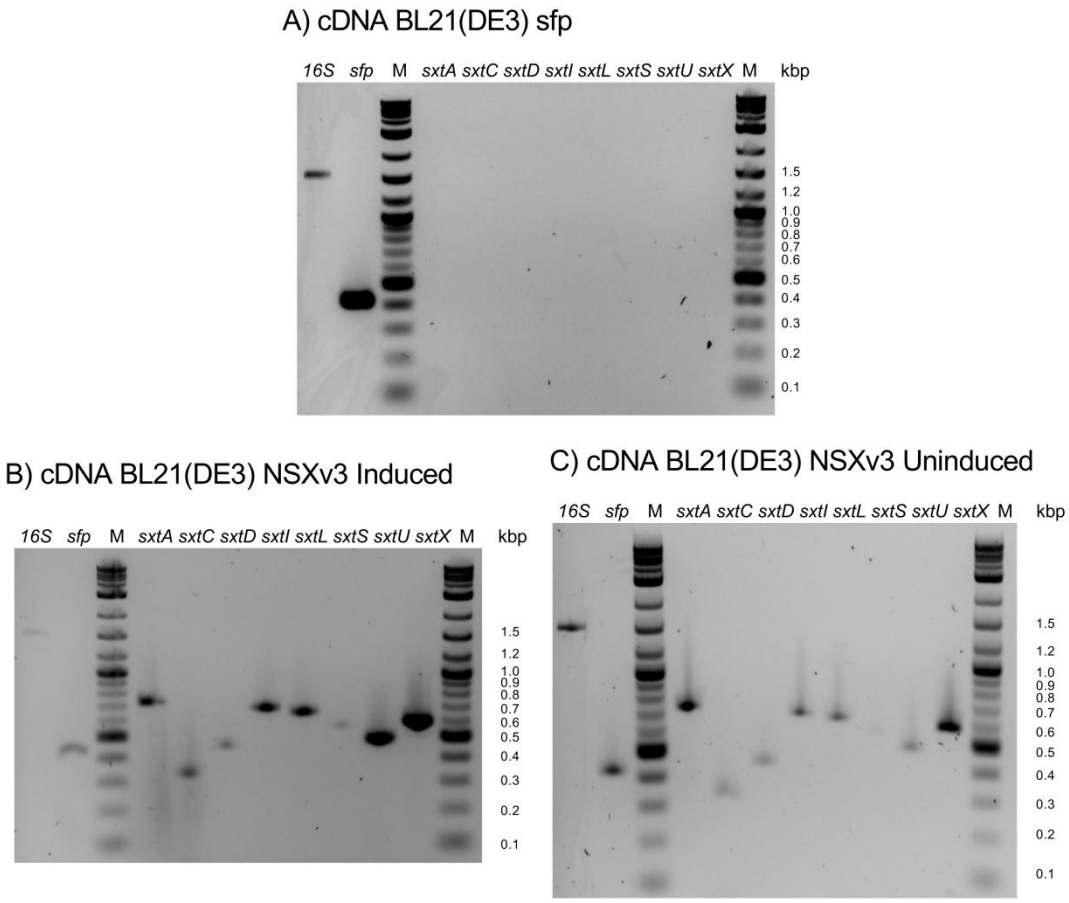


Figure 4.6 Transcriptional analysis of *sxt* gene expression. PCR amplification of 16S rRNA, *sfp*, and *sxt* genes from cDNA reverse transcribed from mRNA of *E. coli* BL21(DE3) *sfp* induced (**A**), *E. coli* BL21(DE3) *sfp* NSXv3 induced (**B**), and *E. coli* BL21(DE3) *sfp* NSXv3 uninduced (**C**). Induction of gene expression was performed with 0.05 mM IPTG and 0.5 mM m-toluic acid over a 24 h incubation at 19°C. M denotes the marker of 2-log DNA ladder (New England Biolabs). Equal volume of PCR reactions were loaded onto a 2% agarose gel. For a clearer image of low concentration amplicons see Appendix Figure A 3.2.

4.3.4 Confirmation of PST production through activity assay

To analyse the production of PSTs from recombinant *E. coli* BL21(DE3) *sfp* NSXv3, a diagnostic test was performed using Saxitoxins ELISA kit (Abraxis). The results of the ELISA test are provided in Table 4.2. As shown, the normalised absorbance is inversely correlated with the concentration of saxitoxin. The Abraxis kit employs polyclonal saxitoxin antibodies, which have 1.3% cross-reactivity to neoSTX (Abraxis). The accuracy of the STX ELISA for neoSTX measurement, the assay was performed on the neat neoSTX reference standard and results were compared with the data from STX calibration standards. The results indicated that the cross-reactivity to neoSTX was slightly lower than expected. A concentration of 100 nM neoSTX in

cell extracts resulted in a lower normalised absorbance compared to a neat sample of 1.34 nM STX. To estimate the recovery of neoSTX by solid phase extraction, crude extract of *E. coli* BL21(DE3) *sfp* was spiked with 100 nM neoSTX, purified, and then subjected to ELISA test. The SPE method applied had a 92.2 % recovery rate. The purification and the assay therefore, could be used as a diagnostic tool for neoSTX. The purified extract of recombinant *E. coli* BL21(DE3) *sfp* NSXv3 yielded a 0.22 normalised absorbance. This demonstrated that the *E. coli* BL21(DE3) *sfp* NSXv3 produces metabolites that bind the anti-saxitoxin antibodies.

Table 4.3 Normalised absorbance obtained from ELISA with saxitoxin standards and intracellular *E. coli* extracts.

Sample	Theoretical concentration (nM)	Normalised absorbance$_{\lambda 450}$	STX equivalent concentration (nM
STX Std 0	0.00	1.00	0
STX Std 1	0.07	0.73	0.07
STX Std 2	1.34	0.14	1.34
neoSTX standard	100.00	0.04	119.9
Purified extracts containing neoSTX standard [1]		0.09	110.5
Purified *E. coli* BL21(DE3) Sfp NSXv3 extract		0.22	

[1]*E. coli* BL21(DE3) Sfp crude extract spiked with 100 nM neoSTX and then purified by solid phase extraction

A mouse neuroblastoma bioassay (MNBA) was performed on the Neuro-2a cell line to assess the production of PSTs. In the present study, optimisation of the assay for *E. coli* extracts as the biological matrix was conducted. Purification of *E. coli* cell extracts using a weak cation exchange Sep-Pak column (Waters) was performed to minimise the *E. coli* extract matrix effect in the MNBA. A standard curve was constructed by spiking the biological matrix (purified of *E. coli* BL21(DE3) *sfp* extract) with the authentic neoSTX standard. The antagonistic effect of neoSTX on veratridine and ouabain toxicity was shown to be dose-dependent (Appendix Figure A 3.3). However, the response was shown to have a saturation point at higher concentrations of

neoSTX (100-200 nM). Nevertheless, this observation confirmed that the MNBA assay could be used to evaluate the production of PSTs from *E. coli* cells.

Thepurified extracts of *E. coli* BL21(DE3) *sfp* NSXv3 showed significantly higher cell survival in comparison to cells subjected to ouabain and veratridine treatment alone (negative control) ($P < 0.001$ by Student's *t*-test analysis) (Figure 4.7), and in comparison to the SPE purified extracts of the negative control strain BL21(DE3) *sfp* ($P < 0.001$ by Student's *t*-test analysis). The solid phase extraction purification performed was required to allow differentiation between the matrix effect, shown by the purified extracts of BL21(DE3) *sfp* result and the sodium channel blocking activity. The assay was performed with three independent extracts from of *E. coli* BL21(DE3) *sfp* NSXv3, demonstrating the stable measurement of sodium channel blocking activity.

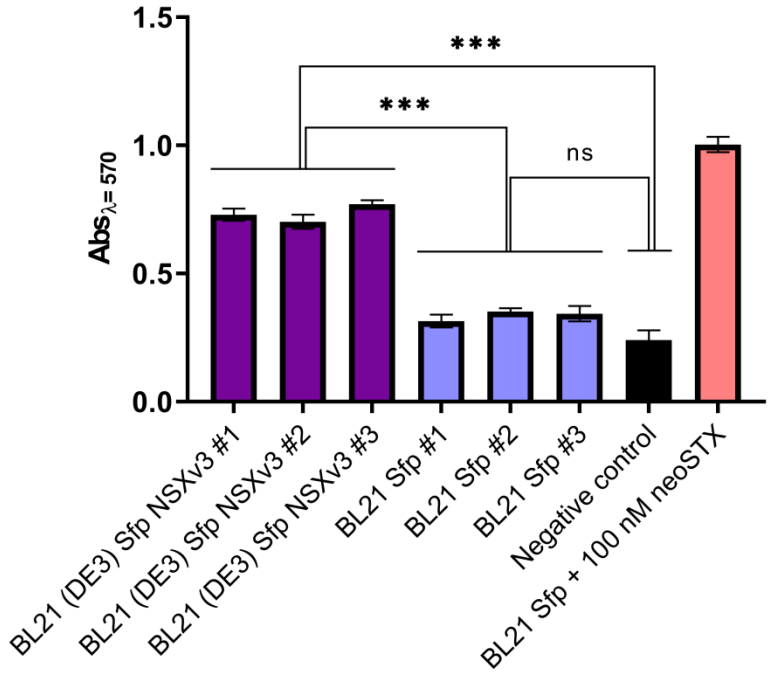


Figure 4.7 Antagonism of veratridine and ouabain toxicity by *E. coli* cell extracts in Neuro-2a cells. Cells were exposed to 50 µM veratridine (V) and 500 µM ouabain (O) and purified *E. coli* cell extracts for 24 h. Cells exposed to ouabain and veratridine treatment alone served as the negative control while cells exposed to ouabain and veratridine and purified *E. coli* extracts spiked with neoSTX served as the positive control. The assay was performed with four

biological replicates (n = 3). Ns, not significant; *** represents $P < 0.001$ by Student's *t*-test analysis, GraphPad Prism 7.

4.4 Discussion

Paralytic shellfish toxins (PSTs) are a family of neurotoxic alkaloids of which some have medically-relevant bioactivities, particularly as pain killers and muscle relaxants (Kohane *et al.*, 2000, Garrido *et al.*, 2004, Garrido *et al.*, 2005, Rodriguez-Navarro *et al.*, 2007, Lattes *et al.*, 2009, Rodriguez-Navarro *et al.*, 2009). To date, at least 57 STX analogues have been identified, including the non-sulfated forms, saxitoxin (STX) and neosaxitoxin (neoSTX); the monosulfated gonyautoxins (GTXs); the disulfated C-toxins; the acetylated variants of *Lyngbya wollei* toxins (LWT1-6); the decarbamoylated toxins (dc-toxins); and the deoxy-decarbamoylated toxins (doSTX and doGTX 1-3). Neosaxitoxin is N-1 hydroxylated saxitoxin which currently is one of the most promising pain therapy agents being investigated as alternatives to opioid-derived drugs, which are known to have many undesirable side effects (Kohane *et al.*, 2000, Rodriguez-Navarro *et al.*, 2007, Rodriguez-Navarro *et al.*, 2009). Unfortunately, the traditional method of extracting neoSTX from the native producers is inefficient as they are difficult to cultivate at large scale and produce an array of PST analogues. It is therefore imperative to develop a platform for neoSTX targeted production that would be suitable for industrial applications. Recently, the emerging technology of synthetic biology was shown to facilitate the expression and manipulation of large biosynthetic cluster in different hosts (Ongley *et al.*, 2013b, Yan *et al.*, 2018a, Cook & Pfleger, 2019, Liu *et al.*, 2017, Luo *et al.*, 2016, D'Agostino & Gulder, 2018). The aim of this project was, therefore, to apply a synthetic biology approach to generate a neoSTX producing strain of *E. coli*, that would be suitable for the industrial production of neoSTX for medical applications. This platform is also of interest from an academic point of view for the characterisation of saxitoxin biosynthesis pathway and the functions of its biosynthesis genes.

The *sxt* biosynthetic cluster from *R. raciborskii* T3 contains more than 20 genes putatively involved in neoSTX biosynthesis. The pathway was re-designed into three transcriptional fragments containing genes which are predicted to be necessary for neoSTX

production while preserving gene arrangements where possible. The three fragments were fragment 1 containing *sxtABC,* fragment 2 containing *sxtDEGHIJKL,* and fragment 3 *sxtQRorf24sxtSTUVWX*. Each operon was fitted with a positively regulated XylS/*Pm* promoter system (originating from *Pseudomonas putida* TOL (toluene degradation) plasmid pWWO) to drive transcription of the *sxt* biosynthetic genes (Gallegos *et al.*, 1996, Worsey & Williams, 1975). Each gene was codon-optimised for *E. coli* and a host specific RBS was inserted upstream to the start codon. All transcriptional units were integrated into the sugar operons in the *E. coli* genome. Previously, Alessandra Estaquio, Antonio Garcia Moyano and Sudeep Karki (University of Bergen, Norway) have integrated several *sxt* genes encoded in fragment 1 and 2 (*sxtABC and sxtDEGHIJKL* under the control of the XylS/*Pm* promoter system) into the genome of *E. coli* BL21(DE3). The resulting strain, *E. coli* BL21(DE3) *T3PPTase* sxt1 sxt2 did not produce any PSTs. This result was expected as the strain lacked several genes which were hypothesised to be necessary for neosaxitoxin biosynthesis. The present study reported the extension of previous work to integrate the remaining genes in the pathway encoded in *sxt* fragment 3 and successfully generated *E. coli* strains that produce neoSTX.

The cloning described in Section 4.3.1 generated three *E. coli* variants of v1 (*sxtQRorf24sxtSTUVWX*), v2 (*sxtQRSTUVWX),* and v3 (*sxtSTUVWX*). LC-MS analysis showed the presence of molecular ions corresponding to the *m/z* of neosaxitoxin in crude extracts of all three variants. This demonstrated that the genes encoded in *sxt* fragment 3 were successfully expressed and that these genes are crucial for the production of neoSTX.

The LC-MS analysis confirmed the production of neoSTX. The level detected, however, was relatively low, causing difficulties in product purification for further LC-MS/MS analysis and NMR, due to a high level of interference from the *E. coli* matrix. As a result, PST-specific assays such as the Saxitoxin ELISA and mouse neuroblastoma bioassay (MNBA) were performed to verify the production of neoSTX.

ELISA immunoassay is traditionally used as one of the monitoring tools for detecting the presence of PSTs in environmental samples, and several kits are commercially available. The assay used in the present study was a direct competitive ELISA based on the recognition of

saxitoxin by specific anti-saxitoxin antibody immobilised on the microtiter plate. The assay has cross-reactivity with other PSTs including neoSTX, and was therefore used in this study. The challenge, however, was to adapt this assay for the detection of neoSTX from *E. coli* BL21(DE3) *sfp* NSXv3 extracts as the method was designed for samples derived from cyanobacteria, dinoflagellates, and contaminated shellfish. Crude extracts of *E. coli* control strains yielded a false-positive result, which was presumably caused by unspecific binding of intracellular metabolites with the antibodies. Method optimisation through metabolite purification by SampliQ SPE column minimalised the matrix effect permitting positive detection of PST in *E. coli* BL21(DE3) *sfp* NSXv3.

Another method that is commonly used as a screening assay for determination of PSTs presence in environmental samples is the *in vitro* mouse neuroblastoma bioassay (MNBA). In this assay, the effect of intracellular cell extracts on cell viability was assessed in the presence of ouabain and veratridine, which act as enhancers on Na^+ influx into the cells (Barnes & Hille, 1988, Ulbricht, 1998, Blaustein *et al.*, 1998). The sodium channel blockers activity of PSTs counteracts the toxic effect of veratridine, enhancing cell survival. Crude extracts and SampliQ SPE purified extracts were shown to cause Neuro-2a cell death, presumably due to high salt content. In the present study, the extraction method was optimised and the weak cation exchange purified intracellular extracts were shown to counteract the effects of ouabain and veratridine toxicity and rescue cells from cell death due to sodium influx. This demonstrates sodium channel blocking activity by the metabolites present in the *E. coli* extracts.

R. raciborskii T3 produces STX, neoSTX, GTXs, dcSTX, and deoxydcSTXs (D'Agostino *et al.*, 2019). As the present study aimed to target the production of neoSTX, the genes involved in the biosynthesis of sulphated PSTs (*sxtN* and *sxtO*) (Cullen *et al.*, 2018a, Soto-Liebe *et al.*, 2010, Kellmann *et al.*, 2008) were purposely not included in the engineered pathway. This was considered to be successful as none of the other PSTs were detected in the LC-MS assays. The LC-MS findings were supported by the ELISA and MNBA assays, which confirmed the production of PSTs by *E. coli* BL21(DE3) *sfp* NSXv3. However, due to the limitation of these bioassays, we cannot discount the possibility of interference on ELISA and

MNBA by other unknown STX intermediates and/or other PSTs that may have been expressed at levels below the detection limits of the LC-MS system. Future studies could look to improve expression levels and the analytical methods for identifying and quantifying the amounts of PSTs and their intermediates.

Successful heterologous expression of a biosynthetic pathway starts with the transcription of the corresponding genes. Often, the insertion of a strong inducible promoter that is known to be compatible with the host is necessary to ensure optimum levels of gene expression (Ongley *et al.*, 2013b, Liu *et al.*, 2017). In the present study, a T7 promoter was used for the expression of the PPTase and a XylS/*Pm* promoter system was used to drive the transcription of the engineered *sxt* cluster. The positive regulator, XylS, forms a dimer complex upon binding to inducers, such as m-toluic acid, and then induces gene transcription from the *Pm* promoter (Ramos *et al.*, 1987, Gallegos *et al.*, 1996, Gawin *et al.*, 2017). The broad-host-range XylS/*Pm* promoter has been previously used for industrial and medical processes (Sletta *et al.*, 2004, Sletta *et al.*, 2007) based on its favourable properties of strong induction by low-cost inducers, and that are rarely present in the environment (Balzer *et al.*, 2013). Moreover, the promoter activity is tunable by varying media composition and inducer concentration (Winther-Larsen *et al.*, 2000, Saida *et al.*, 2006, Lale *et al.*, 2011), and can be further optimised by mutagenesis strategies (Berg *et al.*, 2009, Bakke *et al.*, 2009, Lale *et al.*, 2011, Zwick *et al.*, 2012, Berg *et al.*, 2012). Analysis of transcripts in induced cultures of *E. coli* BL21(DE3) *sfp* NSXv3 showed that *sfp* and the *sxt* cluster were fully transcribed. Transcription of *sxt* genes was also detected in the uninduced cells showing the leakiness of *Pm* promoter in *E. coli* in the absence of m-toluic acid, as was expected (Balzer *et al.*, 2013) (see Chapter 3). Overall, the transcriptional analysis showed that the XylS/*Pm* promoter is suitable for the expression of the *sxt* cluster in *E. coli*.

The *sxt* biosynthetic cluster from *R. raciborskii* T3 contains several genes with unconfirmed functions, including those putatively involved in tailoring reactions. Section 4.3.1 described the generation of *E. coli T3PPTase* sxt1 sxt2 sxt3 variants, differing in the presence of the genes *sxtQ*, *sxtR*, and *orf24* (Figure 4.3B). Based on the LC-MS analysis, these three genes

are not essential for PST biosynthesis, as the BL21(DE3) *sfp* NSXv3 strain was able to produce PSTs. Future studies could use this strain as the base for further pathway engineering to build the minimal gene cluster for neoSTX biosynthesis, while simultaneously shedding light on the function of specific genes. The Red/ET recombineering technique was shown to be a useful technique for gene insertion (Section 4.2.2, 4.2.3, and 4.3.1). Similar techniques could be used for the deletion of several genes of unconfirmed function such as *sxtCEJKL*. The *sxt* biosynthetic gene cluster encodes genes putatively involved in PST transport including the MATE transporter homologues SxtF and SxtM (Kellmann *et al.*, 2008, Soto-Liebe *et al.*, 2013, Mihali *et al.*, 2009). Export of metabolites is beneficial for industrial production as it simplifies fermentation and downstream processing, as well as reducing product-inhibition of the biosynthetic enzymes. Insertion of *sxtFPM* in the engineered cluster can be expected to yield higher extracellular and overall PST production.

4.5 Conclusion and future directions

The present study has established a new platform to study and produce PSTs by using modern molecular technology. The saxitoxin biosynthetic cluster was efficiently integrated into *E. coli* BL21(DE3) together with the Sfp PPTase from *B. subtilis*. NeoSTX was detected from the recombinant *E. coli* strain by three assays with independent principles, namely, qualitative LC-MS analysis, ELISA, and MNBA, providing evidence for successful neoSTX production. Method development was performed to adapt the standardised practice of PSTs detection commonly used in research and industry. This work can be further developed to enable the future production and analysis of PSTs in an *E. coli* host.

The chemical configuration of the metabolites produced has not been confirmed due to the limitations of the LC-MS/MS and NMR analysis, and the matrix effect of *E. coli* metabolites. The development of instrumental methods for PSTs detection is known to be challenging due to their lack of a useful chromophore, and the non-volatility and high polarity of these compounds (Schantz *et al.*, 1975). Adaptation of a recently developed LC-MS method

for PSTs detection in cyanobacteria (D'Agostino *et al.*, 2019) coupled with optimisation of extraction methods would be worthwhile for future qualitative and quantitative analysis of PSTs and intermediates in *E. coli*. Moreover, isotope labelled precursor feeding together with HRMS/MS analysis performed in Chapters 2 and 3 could also aid the elucidation of the structure of the PSTs and their intermediates.

In Chapter 3, we hypothesised that SxtA might have an internal promoter where RNA polymerase can bind and initiate transcription. Herein, refactorisation of the *sxt* pathway was achieved, which involved the rearrangement of the *sxt* cluster, the inclusion of host-specific inducible promoter, RBS and codon optimisation to maximise gene expression, and the deletion of intergenic nucleotides. However, the presence of internal promoters in the codon optimised genes was not the scope of the study. Although the sequence of *sxtA* is dissimilar and has a different predicted internal promoter, we do not discount the possibility of the presence of internal promoters in *sxtA* and the other *sxt* genes cloned in this study. This might lead to the formation of truncated mRNA transcripts, which in turn might affect downstream metabolite production. Future studies could use qPCR analysis to quantitatively determine the level of transcription of all genes in the recombinant pathway, and if relevant, investigate the presence of internal promoters. At the protein level, a previous study has suggested that there is an optimum level of PPTase to maximise PKS activity (Liu *et al.*, 2018). In the case of microcystin, reduction of the product yield occurred when the PPTase was oversupplied, which may be due to occlusion of the ACP domain (Liu *et al.*, 2018). It is anticipated that the modulation of culture conditions, including inducer concentration, could optimise the balance of PPTase and SxtA expression levels, which in turn might increase metabolite production.

In summary, this study has demonstrated the application of synthetic biology to achieve a heterologous expression of neoSTX in *E.coli*. We have engineered the *sxt* gene cluster to direct the production of neosaxitoxin, while uncovering the redundancy of several genes. This study serves as a proof of concept that synthetic biology is a feasible strategy for the directed synthesis of neosaxitoxin and provides a platform for further STX pathway characterisation and the production of other PSTs.

Chapter 5. Concluding remarks and future directions

5.1 Research motivation and objectives

Microorganisms have historically been recognised as valuable sources of natural products that elicit medically and industrially relevant activities. Cyanobacteria are one of the most chemically diverse microbial phyla, but have been largely overlooked by the biotechnology sector due to the technical difficulties associated with culturing and genetically modifying most species. In recent years, genome sequencing advances have uncovered various nonribosomal peptide (NRP) and polyketide (PK) biosynthetic gene clusters from marine and freshwater cyanobacteria. However, the majority of these pathways remain uncharacterised and their bioactive potential has not been realised. Synthetic biology is an emerging, but powerful technology with the potential to advance the discovery, heterologous expression, and biochemical characterisation of natural products from a wide range of microorganisms, including fastidious or non-culturable cyanobacteria.

Genetic engineering of NRP and PK biosynthesis pathways is challenging due to the large size (10-100 kbp) and complexity of the associated gene clusters, which comprise multiple enzymes. Heterologous expression of these pathways often requires extensive optimisation of genetic constructs and culture conditions. Recently, several synthetic biology tools such as RecET recombineering, yeast transformation associated recombination (TAR) cloning, Gibson assembly, and Golden Gate cloning have been developed, increasing the feasibility of cloning and manipulating large biosynthetic clusters. These techniques can be used to engineer expression strains for the high-yield production of natural products and novel derivatives thereof.

Previous studies focussing on the microcystin pathway demonstrated that the heterologous expression of large complex cyanobacterial biosynthetic clusters in *E. coli* is feasible (Liu *et al.*, 2017). The objective of this dissertation was to utilise similar synthetic biology techniques to characterise the biosynthetic pathways of lyngbyatoxin (LTX) and saxitoxin (STX), and provide alternative access to these biomedically relevant natural products. In Chapter 2, the modularity of NRPS scaffolds was exploited via site-directed mutagenesis to

alter the relatively-simple lyngbyatoxin biosynthesis pathway and direct the heterologous production of LTX variants. Chapters 3 and 4 described the cloning and heterologous expression of the saxitoxin biosynthetic pathway (*sxt*), resulting in the production of STX. In addition to cloning and expressing these pathways in *E. coli*, new strategies were also developed for the chemical diversification of LTX and STX, and for improving product titres.

5.2 Key findings

5.2.1 Expression and manipulation of LTX biosynthetic cluster leads to product discoveries and pathway characterisation

In the past few decades, many cyanobacterial bioactive compounds have been discovered. With the recent developments of synthetic biology methods, several examples of successful production of cyanobacterial natural products in heterologous hosts have been described. Thus far, several species have been assessed as heterologous hosts; including *E. coli* (production of patellamide (Schmidt *et al.*, 2005), lyngbyatoxin (Ongley *et al.*, 2013b), and microcystin (Liu *et al.*, 2017)), *Streptomyces* (production of 4-O-demethylbarbamide (Kim *et al.*, 2012)), and more recently cyanobacteria species of *Anabaena* sp. strain PCC 7120 (production of lyngbyatoxin (Videau *et al.*, 2016)) and *Synechococcus* sp. UTEX 2973 (production of hapalindole (Knoot *et al.*, 2019)).

A previous study described the heterologous production of lyngbyatoxin, a dermatoxin produced by the marine cyanobacterium *Moorea producens*, in *E. coli* (Ongley *et al.*, 2013b). The use of heterologous expression for pathway characterisation and further enzyme engineering was exploited in Chapter 2. It was shown that LtxA, the NRPS of the *ltx* pathway, possesses multispecificity towards L-Val related amino acids, specifically L-Ile and L-Leu, resulting in the production of novel methylated analogues of lyngbyatoxin in *E. coli*. However, it was shown that LtxA also possesses some degree of substrate stringency as site-directed mutagenesis did not significantly alter *in vitro* substrate activation nor the *in vivo* dipeptide production. Nevertheless, the results illustrate that heterologous expression is a useful approach

for characterising cyanobacterial biosynthetic pathways and directing the key metabolite flux using synthetic biology methods.

5.2.2 Towards the production of paralytic shellfish toxins (PSTs)

Paralytic shellfish toxins (PSTs) are a family of neurotoxic alkaloids produced by freshwater cyanobacteria and marine dinoflagellates. Neosaxitoxin (neoSTX), a member of the PSTs is currently being evaluated in clinical trials for pain therapy (Kohane *et al.*, 2000, Epstein-Barash *et al.*, 2009). Efficient chemical synthesis of PSTs is challenging due to their complex structures. Chapters 3 and 4 described the heterologous expression and refactorisation of the *sxt* biosynthetic cluster, providing an alternative source of neoSTX.

In achieving heterologous expression of NRPS/PKS biosynthetic pathways, refactorising of the native cluster was often performed (Menzella *et al.*, 2006, Liu *et al.*, 2017, Ongley *et al.*, 2013b, Kang *et al.*, 2016). Successful heterologous expression often required the adjustment of multiple factors to promote compound production; including genetic components (such as promoters, RBSs, terminators, codon-usage, insertion/deletion of regulatory elements), metabolic engineering of the host organisms, and co-expression of a suitable PPTase. In Chapter 3, the efficient production of saxitoxin Intermediate-A' (Int-A') in *E. coli* was achieved through the co-expression of the first enzyme in saxitoxin biosynthesis, SxtA (a PKS-like enzyme) and Sfp (a PPTase). The *Pm* promoter was shown to be suitable for driving the expression of *sxtA* while the T7 promoter was used for the expression of *sfp*.

Building on the results of Chapter 3, the *sxt* cluster from *R. raciborskii* T3 was redesigned and integrated into the *E. coli* genome (Chapter 4). The *sxt* biosynthetic cluster from *R. raciborskii* T3 consists of more than 20 genes, most of which have not been functionally characterised. The generation of different *E. coli* derivatives in Chapter 4 uncovered the redundancy of *sxtQ*, *sxtR*, and *orf24* for the production of neosaxitoxin. In addition, pathway engineering was performed to direct PST production towards neosaxitoxin by inclusion of the tailoring gene *sxtX* (catalysing N-1 hydroxylation) and exclusion of *sxtN* and *sxtO* (genes involved in sulfonation of saxitoxin). Results from LC-MS analysis, supported by ELISA and

mouse neuroblastoma activity assays, demonstrated the production of neoSTX by the engineered *E. coli*. This study serves as a proof-of-concept that synthetic biology is a feasible strategy for the sustainable production of neosaxitoxin and for the directed synthesis of PST-derivatives.

5.2.3 In vitro and *in vivo* substrate preference revealed the complexity of NRPSs PKS enzymology

Since their discovery, the modularity of NRPS and PKS enzymes has been exploited to manipulate discrete steps in natural product biosynthetic pathway via combinatorial biosynthesis, resulting in the structural diversification of pathway products. For NRPSs, the new enzyme activity is commonly evaluated through the *in vitro* substrate activation of the recombinant A-domains or NRPS modules. However, recent studies on NRPSs and PKSs showed their structural dynamics, which are affected by inter-domain and inter-module interactions revealing a more complex chemoenzymatic activity than previously thought (Meyer *et al.*, 2016, Reimer *et al.*, 2016, Drake *et al.*, 2016, Gulick, 2016). It is, therefore, necessary to biochemically characterise the activity of intact NRPSs and PKSs to understand their native role in natural product biosynthesis pathways and to exploit their activity for the rational design of drugs.

In Chapter 2, both *in vitro* substrate activation and *in vivo* product formation in lyngbyatoxin biosynthesis was performed to gain insight into the enzymology of lyngbyatoxin biosynthesis. Manipulation of the pathway through site-directed mutagenesis of the first domain of LtxA (NRPS) was performed. It was shown that the *in vitro* adenylation assays of full-length LtxA could be used to predict the corresponding *in vivo* activity. However, it was also shown that environmental condition (i.e. amino acid availability) influences metabolite production, and therefore precursor feeding was performed to increase the production of methylated analogues. In the present study, substrate stringency of LtxA was observed as site-directed mutagenesis of the A-domain only marginally altered LtxA activity. The tailoring enzyme LtxB contains an MbtH-like protein (MLP) domain, which in some cases is associated with the NRPS A-domain

and affects enzyme activity. Interestingly, the LtxB protein co-expressed and co-purified with LtxA demonstrating a strong interaction between the two proteins. This suggests that the MLP domain within LtxB could influence LtxA substrate selection and activation.

In Chapter 3, the *in vivo* activity of the PKS-like enzyme in saxitoxin biosynthesis, SxtA was characterised. It was shown that the saxitoxin Int-A' was produced along with a truncated analogue of demethylated Int-A' (demetInt-A'). SxtA therefore has broad substrate specificity *in vivo*. The formation of either Int-A' and demetInt-A is likely dependent on the acyl-CoA substrate utilised. As was observed for the LTX pathway, the production ratio of the two pathway products was dependent on the growth media. As the truncated intermediate might be a shunt product of the pathway, this current study highlights the need for metabolic engineering and optimisation of culture conditions to improve the production of Int-A' and ultimately STX or other PSTs.

5.2.4 Tailoring enzyme characterisation provides insights for pathway engineering

The use of synthetic biology for natural product discovery and structural diversification has been emerging for other phyla such as actinomycetes and fungi. Yet, the manipulation of the cyanobacterial biosynthetic pathways has been limited. Engineering of the adenylation domain of LtxA, the NRPS in lyngbyatoxin biosynthesis, was described in Chapter 2. Upon investigation of the full wild-type and mutant *ltx* pathways *in vivo*, it was found that the tailoring enzyme LtxB possess substrate specificity towards L-valyl dipeptide (NMVT), followed by L- isoleucyl (NMIT) and L-leucyl (NMLT) dipeptides. This preference dictated the metabolic flux in LTX biosynthesis and was particularly evident in the LtxA mutants, which had a relatively unaltered ratio of isoleucine and leucine indolactam and lyngbyatoxin products, compared to the strain expressing the native full cluster. This result showed that the alteration of dipeptide formation was not enough to overcome the control exerted by tailoring domains, such as LtxB. Different strategies such as engineering of LtxB, to broaden its specificity, should be employed for reprogramming analogue production in *E. coli*. This study also highlights the need

for tailoring enzyme characterisation through *in vivo* metabolite analysis to provide a clearer picture of their native activity inside the cells.

5.3 Future directions

5.3.1 Confirmation of metabolites produced during STX biosynthesis and 'core' enzyme characterisation

In comparison to other cyanotoxin biosynthetic pathways, saxitoxin biosynthesis is regarded as unique as it involves only one PKS module with a series of modifying enzymes. The complexity of the STX pathway has confounded researchers, resulting in multiple revisions of biosynthetic steps and intermediates (Shimizu *et al.*, 1984, Kellmann *et al.*, 2008, Tsuchiya *et al.*, 2014, Tsuchiya *et al.*, 2016, Tsuchiya *et al.*, 2017, Lukowski *et al.*, 2018, Cullen *et al.*, 2018b). The cloning and expression of the *sxt* cluster from *R. raciborskii* described in Chapter 4 has provided a new platform for characterising this unique pathway.

The preliminary evidence shows that the integrated *sxt* gene cluster is functionally expressed, resulting in the production of neosaxitoxin. However, it was shown that the production of neosaxitoxin, as detected by LC-MS analysis, was at a low level. The results were supported by ELISA and the mouse neuroblastoma assay (MNBA). However, it is possible that interference by other unknown STX intermediates and/or other PSTs yet to be identified by LC-MS analysis, could have resulted in false positives. The high polarity of PSTs renders chemical extractions and analysis difficult. Future studies should optimise metabolite extraction and purification, followed by metabolomic analysis of PSTs and intermediates produced by the recombinant *E. coli*. This could also be assisted by stable isotopic feeding and metabolic products tracing. Recently D'Agostino *et al.* (D'Agostino *et al.*, 2019) developed a method for PSTs extraction and detection from cyanobacteria. This could be used as a reference for optimisation of PST analysis for various saxitoxin analogues.

The applicability of synthetic biology methods for pathway characterisation was explored in Chapter 4, through cluster refactorisation and stepwise integrations of *sxt* operons.

Future efforts could expand a similar strategy by creating strains with 'core' gene knockouts (i.e. of *sxtG, sxtS, sxtU*) and/or deletion of gene with unknown functions followed by metabolic analysis to build the minimal 'core' gene clusters, assign gene functions, identify pathway intermediates and catalytic steps. It is also anticipated that this will provide insights into pathway metabolite flux and identify bottlenecks within the pathway to inform future optimisation strategies.

5.3.2 Yield optimisation and pathway engineering

Chapter 4 described the proof of concept for neoSTX production using heterologous expression in *E. coli*. The titre achieved, however, is not enough for structural studies by NMR nor other research and commercial use. The benefit of using the well-characterised *E. coli* as a heterologous host, together with the insights obtained by the pathway characterisation described above uncover a vast potential to further increase end product levels.

In the expression of SxtA described in Chapter 3, it was shown that a truncated analogue of Int-A', demetInt-A', was produced. This truncated intermediate may be a shunt product or a biosynthetic competitor consuming the substrates of arginine and acyl-CoA. A future study could investigate the fate of this metabolite in the STX pathway to inform further optimisation strategies. Both for LtxA and SxtA (Chapters 2 and 3), it was shown that environmental conditions affect product yield. The product titre could therefore be increased by determining of SxtA substrate specificity *in vivo* and mutating SxtA enzyme, coupled with host metabolic engineering and optimisation of culture conditions.

A recent investigation of toxin profiles from cyanobacterial strains revealed a very high overall PST production by *Aphanizomenon* sp. NH-5 in comparison to *R. raciborskii* T3 and several other producers (D'Agostino *et al.*, 2019). The cause of this is not yet understood nor is the reason for different toxin profiles among PST producers. The application of –omics tools and techniques could be used to study the effect of gene expression and regulation, post-translational modification, enzyme catalytic efficiency, and host metabolism, on toxin production. In addition, future studies could also explore the synthetic biology approach by

creating strains either with deleted or over-expressed putative regulatory genes, followed by metabolic analysis of neoSTX and related PSTs. It could be anticipated that this investigation would provide insight for pathway optimisation.

Synthetic biology also allows gene expression tuning at the transcriptional level. Chapter 3 uncovered the presence of a putative internal promoter, which may affect metabolite production. Quantitative qPCR analysis to determine the level of transcription of *sxt* genes in the recombinant pathway could be beneficial to investigate the presence of internal promoters within the pathway. The transcription level of the recombinant pathway can also be compared with the native producers to provide insight into optimisation strategies.

5.3.3 Characterisation of *sxt* tailoring enzymes and production of saxitoxin variants

To date, 57 analogues of saxitoxin have been reported. Several studies have attempted to characterise the function of the various tailoring enzymes within the *sxt* cluster with a view to understanding and directing PST biosynthesis. The biochemically characterised enzymes include SxtDIOX and SxtN (Cullen *et al.*, 2018a, Lukowski *et al.*, 2018). In the present study, the tailoring gene *sxtX*, which was predicted to catalyse N1 hydroxylation of STX was included in the cluster to direct the production of neosaxitoxin. The function of SxtX could be conclusively established through gene deletion using a similar recombineering approach to that described in Chapter 4. This approach could also be implemented for the characterisation of other tailoring enzymes, while simultaneously directing the production of other PSTs.

5.3.4 Characterisation of the LtxB MLP-domain and lyngbyatoxin pathway engineering

LtxB contains an MbtH-like protein (MLP) domain, which in some cases have been reported to be vital for A-domain protein solubilisation and activation (Zolova & Garneau-Tsodikova, 2012, Felnagle *et al.*, 2010), albeit via an unclear mechanism. Our observation indicates that an MLP domain might be essential for LtxA expression. The specific role of MLP on NRPS substrate activation remains unknown. Several studies have shown that various MLPs could

flexibly interact with A-domains from different NRPS (Lautru *et al.*, 2007, Zhang *et al.*, 2010). Recently, Mori et al. (Mori *et al.*, 2018) exploited this flexible nature, showing the alteration of the TioK A-domain *in vitro* substrate promiscuity using 12 different MLPs. Future efforts to alter LtxA activity and other NRPS enzymes could benefit from an investigation of the LtxB MLP-like domain and co-expression of non-cognate MLPs.

Teleocidin B, an indole alkaloid structurally related to lyngbyatoxin A, is a potent protein kinase C activator produced by *Streptomyces* (Fujiki *et al.*, 1979, Fujiki *et al.*, 1981). The biosynthesis pathway of teleocidin B involved lyngbyatoxin A as a precursor, which was further modified by a SAM-dependent methyl transferase TleD (Awakawa *et al.*, 2014, Zhang *et al.*, 2016). TleD was reported to catalyse C-methylation of the prenyl chain to further trigger terpene cyclisation (Awakawa *et al.*, 2014). In Chapter 2, the production of previously undescribed methylated lyngbyatoxin analogues was reported. Future studies could co-express *tleD* in the *E. coli* strain harbouring the *ltx* cluster for further structural diversification. We acknowledge that currently, only a small fraction of indolactam was converted into lyngbyatoxin A. The metabolic engineering of *E. coli* hosts to increase cellular availability of geranyl pyrophosphate required for methyl-lyngbyatoxin A production also needs to be performed simultaneously.

5.4 Conclusion

This dissertation described the heterologous production of cyanobacterial specialised metabolites for research and industrial applications. The implementation of synthetic biology to modulate substrate selectivity and metabolite production was explored. The *in vitro* and *in vivo* investigation of lyngbyatoxin biosynthesis revealed the pathway stringency that renders the mutation of the LtxA A-domain inefficient for generating lyngbyatoxin analogues *in vivo*. The results, however, provided insight into strategies for the production of specialised metabolites and their analogues. The heterologous expression of a PST biosynthetic pathway was achieved following the refactorisation and integration of the *sxt* cluster in *E. coli*. Several genes within

the cluster were found to be non-essential, allowing pathway characterisation of this complex metabolite. Ultimately, these platforms pave the way for the development of a more scalable and sustainable alternative source of cyanobacterial bioactive metabolites, while simultaneously allowing characterisation of their respective complex biosynthesis pathways.

Appendix 1. Chapter 2 Supplementary information

Table A 1.1 Primers and repair oligonucleotides used in this study.

Primer name	Sequence 5'-3'
Primers for 6xHis-tag LtxA construction	
pETHisTagLtxA_F	TTCATCAACTCTAAATTCTTCCCTCAAATAACTTATACAGGAAAGCTCGAGCACCACCACCAC
pETHisTagLtxA_R	TACTATCGCCCGCCTACTTCCACTCCAAGGTTGATTCATAATCATGCTGCTGTGATGATGATGAT
amp5_sgrAI	CTGAATCACCGGTGTATTTCACACCGCATCAGGT
amp3_bglII	ATGCTAGAGATCTGTCTTGATCCGGCAAACAAAC
amp5-HApfos	TATTCCTTATACCATCAAGATCCTTTGATTTTGAGGGCTATATTTCACACCGCATCAGGTG
genta-Ptet3_pET28HA	GTCGATCCTCTTCTCTATCACTG
genta-Ptet5_pET28HA	GGCGTAGAGGATCGAGAAGGCACGAACCCAGTTG
Screening and sequencing primers	
T7 prom	TAATACGACTCACTATAGGG
T7 term	GCTAGTTATTGCTCAGCGGT
pET_T7promUp	ATCTTCCCCATCGGTGAT
LtxAR	TACTATCGCCCGCCTACTTCC
LTX239seqF	TTATGTACACCTCAGGCTCG
LTX299seqR	TCGTCAAGTACATAGACCTGG
Primers for LtxA-A1 binding pocket mutants and repair oligonucleotides	
239ccdBamp5	GAGTCCTATTTCACAGTTCTCTAGGCTTTGATATCTCAAATTCCGCTAGCGCTTTGTTTAT
239ccdBamp3	ATAACTAACTGTCCAGATCGTAAAAGAGGCACCCACAGTTCAGCCCCATACGATATAAGTTG
299ccdBamp5	CTGAGGAAAATCCCAGATGCCTTGCAGGTGTTCAACAAGTATCCGCTAGCGCTTTGTTTAT
299ccdBamp3	ATCCGCTGAATAGCAGCTACTGAAGCTTGTTCCCCTCCAATAGCCCCATACGATATAAGTTG
Y239Foligo	CGAGTCCTATTTCACAGTTCTCTAGGCTTTGATATCTCAAAT**TTT**GAACTGTGGGTGCCTCTTTTACGATCTGGACAGTTAGTTATG
Y239Noligo	CGAGTCCTATTTCACAGTTCTCTAGGCTTTGATATCTCAAAT**AAC**GAACTGTGGGTGCCTCTTTTACGATCTGGACAGTTAGTTATG
Y239Moligo	CGAGTCCTATTTCACAGTTCTCTAGGCTTTGATATCTCAAAT**ATG**GAACTGTGGGTGCCTCTTTTACGATCTGGACAGTTAGTTATG
Y239Poligo	CGAGTCCTATTTCACAGTTCTCTAGGCTTTGATATCTCAAAT**CCG**GAACTGTGGGTGCCTCTTTTACGATCTGGACAGTTAGTTATG

W299Coligo	ACTGAGGAAAATCCCAGATGCCTTGCAGGTGTTCAACAAG TA**TGC**ATTGGAGGGGAACAAGCTTCAGTAGCTGCTATTCA GCGGATG
W299Loligo	ACTGAGGAAAATCCCAGATGCCTTGCAGGTGTTCAACAAG TA**CTG**ATTGGAGGGGAACAAGCTTCAGTAGCTGCTATTCA GCGGATG

*Underlined sequences are homology arms used to mediate for linear-circular homologous recombination (LCHR) recombineering or Gibson cloning, or restriction sites sequences. The codon mutations for position 239 and 299 are bolded in the repair oligonucleotides.

Table A 1.2 Several examples of LtxA protein expression attempts.

***E. coli* strain**	**Plasmid**	**Induction condition**			**Lysis buffer[1]**	**LtxA expression**		
		Inducer	**OD_{600}**	**Length**		**Soluble**	**Insoluble**	**Ni-NTA[3]**
BL21(DE3)	pET28b::*ltxA*,pRARE	100 µM IPTG	0.6	16-18 h	20 mM NaH_2PO_4 pH 8.0	-	+	-
BL21(DE3)	pET28b::*ltxA*,pRARE	200 µM IPTG	0.6	16-18 h	20 mM NaH_2PO_4 pH 8.0	-	-	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	400 µM IPTG	0.6	16-18 h	20 mM NaH_2PO_4 pH 8.0	-	-	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	500 µM IPTG	0.8	16-18 h	20 mM NaH_2PO_4 pH 8.0	-	+	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	800 µM IPTG	0.8	16-18 h	20 mM NaH_2PO_4 pH 8.0	-	+	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	500 µM IPTG	1.0	16-18 h	20 mM NaH_2PO_4 pH 8.0	-	-	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	800 µM IPTG	1.0	16-18 h	20 mM NaH_2PO_4 pH 8.0	-	-	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	50 µM IPTG	0.6	16-18 h	20 mM NaH_2PO_4 pH 8.0 + additives[2]	-	+	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	50 µM IPTG	0.8	16-18 h	20 mM NaH_2PO_4 pH 8.0 + additives	-	+++	-
BL21(DE3)	pET28b::*ltxA*,pRARE	100 µM IPTG	0.8	16-18 h	20 mM NaH_2PO_4 pH 8.0 + additives	-	+++	-
BL21(DE3)	pET28b::*ltxA*,pRARE	1 mM IPTG	0.8	16-18 h	20 mM NaH_2PO_4 pH 8.0 + additives	-	+	N/A
BL21(DE3)	pET28b::*ltxA*,pRARE	50 µM IPTG	0.8	16-18 h	50 mM Hepes pH 7.5 + additives	-	+	-
BL21(DE3)	pET28b::*ltxA*,pRARE	10 µM IPTG	0.8	16-18 h	50 mM Hepes pH 7.5 + additives	-	+	-
BL21(DE3)	pET28b::*ltxA*,pRARE	25 µM IPTG	0.8	16-18 h	50 mM Hepes pH 7.5 + additives	-	+	N/A
BL21(DE3)	pET28b::*ltxA*, pRARE	50 µM IPTG	0.8	16-18 h	50 mM Hepes pH 7.5 + additives	-	+	-
BL21(DE3)	pET28b::genta-Ptet-*ltxA*	0.5 µg/mL	0.8	16-18 h	50 mM Hepes pH 7.5 + additives	-	+	-
BL21(DE3)	pET28b::genta-Ptet-*ltxA*, pRARE	0.5 µg/mL	0.8	16-18 h	50 mM Hepes pH 7.5 + additives	-	+	-
BL21(DE3)	pCFOS-Ptet-His*ltxAltxBCD*	0.5 µg/mL	0.5	48 h	50 mM Hepes pH 7.5 + additives	-	-	-
GB05-MtaA	pCFOS-Ptet-His*ltxAltxBCD*	0.5 µg/mL	0.5	72 h	50 mM Hepes pH 7.5 + additives	+	-	+++

[1] All lysis buffers contains 20 mM imidazole, 500 mM NaCl, and 10% glycerol.

[2] The additives for the lysis buffer are 0.5 mM TCEP, 5u/ml benzonase, and 2ul/ml protease inhibitor.

[3] LtxA protein were purified by bench purification using Ni-NTA resin (Milipore).

Table A 1.2 *In silico* predicted changes of substrate specificity with single mutations.

X	Binding pocket substitutions and predicted A-domain specificity				
	Y239X	**F278X**	**W299X**	**T330X**	**V331X**
Cys (C)	-	Val	-	-	Val
Glu (E)	-	Val	Tyr	-	Val
Phe (F)	-	Val	-	-	Val
Gly (G)	-	Val	-	-	Orn, Val
Ile (I)	-	Val	-	Orn	Val
Leu (L)	-	Val	Ile	-	Val
Met (M)	Ile	Val	-	-	Val
Asn (N)	Orn	Val	-	-	Val
Pro (P)	Tyr	Val	-	-	Val
Tyr (Y)	-	Orn, Val	-	-	Val

Substrate specificity predictions were performed for all amino acid substitutions (X) at all positions of the eight-residue specificity conferring code using NRPSpredictor2. Substitutions that altered the specificity are shown. A dash (-) indicates that no *in silico* prediction was available; Amino acids not shown, either had no prediction or no change, similarly to other substitutions at that site.

Table A 1.3 Ratio of Ile and Leu- derived pathway products in comparison to Val-derived product from M9 minimal medium fermentations.

	NMIT		NMLT		ILI		ILL		Me-LTX (Ile)	
	Ratio to NMVT	SD	Ratio to NMVT	SD	Ratio to ILV	SD	Ratio to ILV	SD	Ratio to LTX	SD
WT	0.41	0.01	0.22	0.04	0.19	0.01	0.01	0.00	0.04	0.01
Y239M	0.17	0.01	0.14	0.00	0.00	0.01	0.00	0.00	0.02	0.01
Y239F	0.35	0.03	0.22	0.00	0.14	0.05	0.01	0.00	0.04	0.01
W299L	0.12	0.01	0.07	0.01	0.02	0.01	0.00	0.00	0.01	0.00
W299C	0.33	0.04	0.22	0.04	0.12	0.03	0.01	0.01	0.03	0.01
Y239M/W299L	0.14	0.02	0.26	0.09	0.00	0.01	0.00	0.00	0.01	0.01
Y239F/W299L	0.11	0.01	0.10	0.03	0.02	0.01	0.00	0.00	0.01	0.00
Y239M/W299C	0.16	0.02	0.15	0.03	0.06	0.10	0.00	0.00	0.01	0.01
Y239F/W299C	0.52	0.06	0.40	0.04	0.09	0.08	0.01	0.01	0.07	0.03

NMVT: N-methyl-L-isoleucyl-L-tryptophanol; NMIT: N-methyl-L-isoleucyl-L-tryptophanol; NMLT: N-methyl-L-leucyl-L-tryptophanol; ILV: indolactam V; ILI: indolactam I; ILL: indolactam L: LTX: lyngbyatoxin A: Me-LTX (Ile); methyl-lyngbyatoxin A (Ile). Production level of Ile derived metabolites refers to combined amounts of Ile and Ile-d_{10} derived. Experiments were performed in 3 biological replicates.

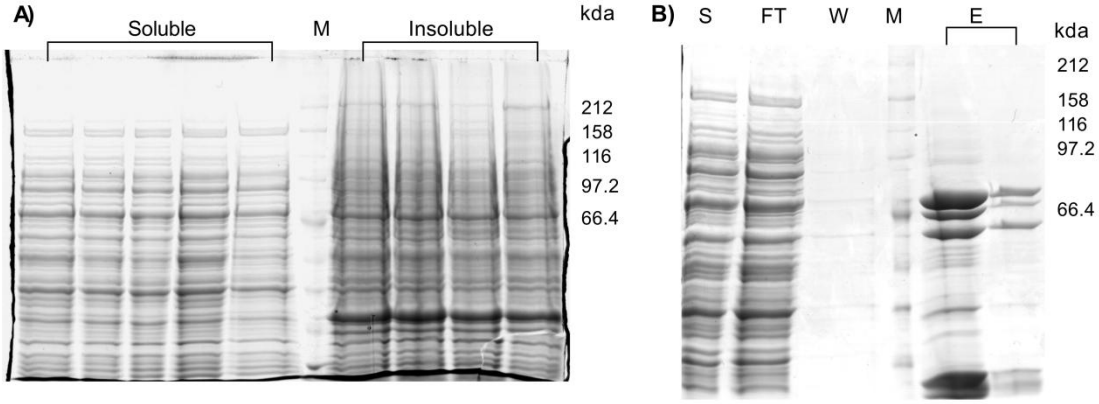


Figure A 1.3 SDS-PAGE gel of LtxA expression from *E. coli* BL21(DE3) pET28b::*ltxA*,pRARE. A) Soluble and insoluble fractions of total protein extracted from E. coli Bl21(DE3) pET28b::*ltxA*,pRARE cell pellet. Cells were cultured with different induction condition, OD600 ~ 0.6 with 50 (Lane 1 and 7) and 1000 μM IPTG (Lane 2 and 8), OD600 ~ 0.8 with 50 (Lane 3 and 9) and 1000 μM IPTG (Lane 10). **B**) Fractions from Ni-NTA protein purification of total soluble protein from Bl21(DE3) pET28b::*ltxA*,pRARE culture induced at OD_{600} ~ 0.6 with 50 μM IPTG.M, unstained protein marker, broad range, 2-212 kDa (NEB) S, soluble; FT, flowthrough fraction; W, wash fraction; E, elution fraction. LtxA only detected in insoluble fractions. Purification from soluble fractions did not yield any LtxA protein.

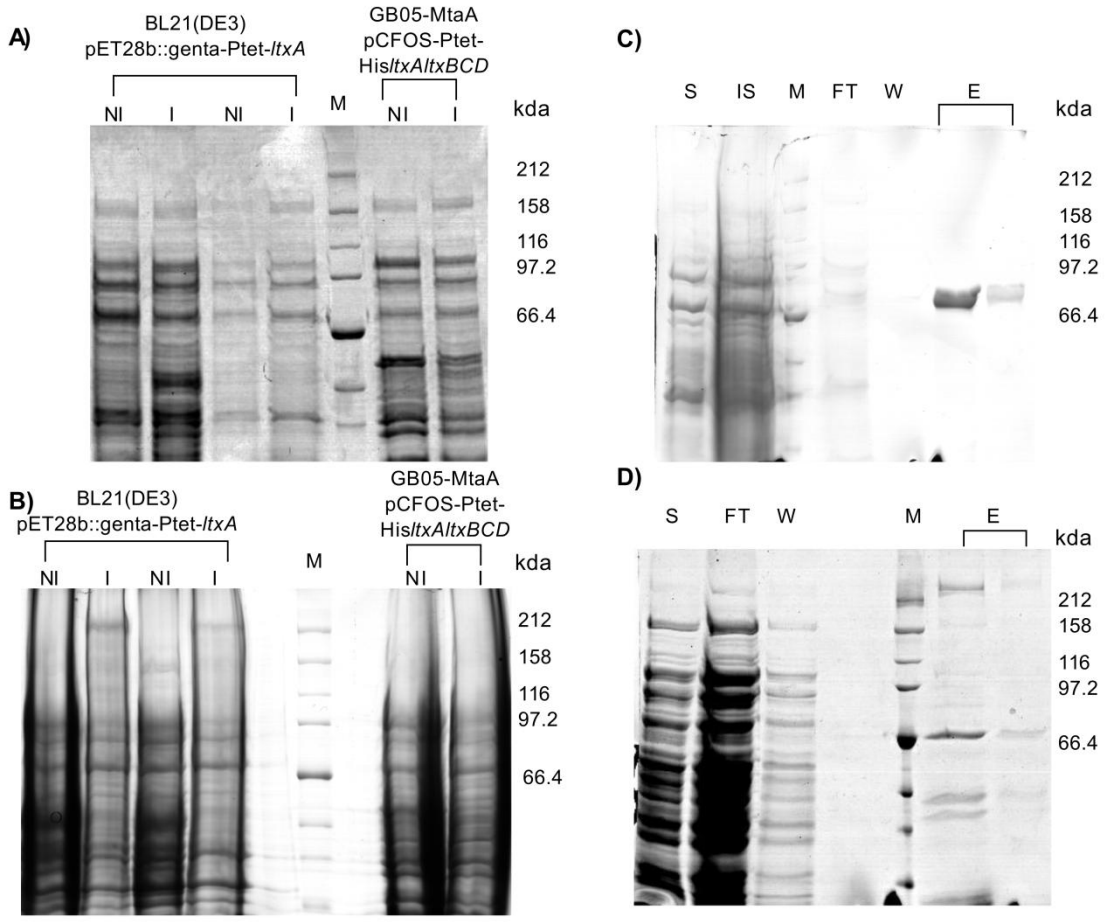


Figure A 1.4 SDS-PAGE gel of LtxA expression from *E. coli* strains harbouring various LtxA expression plasmids. A) Total cellular soluble protein extracted from BL21(DE3) pET28b::genta-Ptet-*ltxA* (Lane 1 and 2), BL21(DE3) pET28b::genta-Ptet-*ltxA*, pRARE (Lane 2 and 3) and GB05-MtaA pCFOS-Ptet-His*ltxAltxBCD* (Lane 6 and 7). **B**) Total cellular insoluble protein extracted from BL21(DE3) pET28b::genta-Ptet-*ltxA* (Lane 1 and 2), BL21(DE3) pET28b::genta-Ptet-*ltxA*, pRARE (Lane 2 and 3) and GB05-MtaA pCFOS-Ptet-His*ltxAltxBCD* (Lane 8 and 9). **C**) Fractions from Ni-NTA protein purification of total soluble protein from BL21(DE3) pCFOS-Ptet-His*ltxAltxBCD*. **D**) Fractions from Ni-NTA protein purification of total soluble protein from GB05-MtaA pCFOS-Ptet-His*ltxAltxBCD*. M, unstained protein marker, broad range, 2-212 kDa (NEB), IS, insoluble S, soluble; FT, flowthrough fraction; W, wash fraction; E, elution fraction.NI, not induced culture, I, induced culture with 0.5 μg/mL tetracycline. LtxA successfully expressed and purified in soluble fraction only from expression in GB05-MtaA pCFOS-Ptet-His*ltxAltxBCD*.

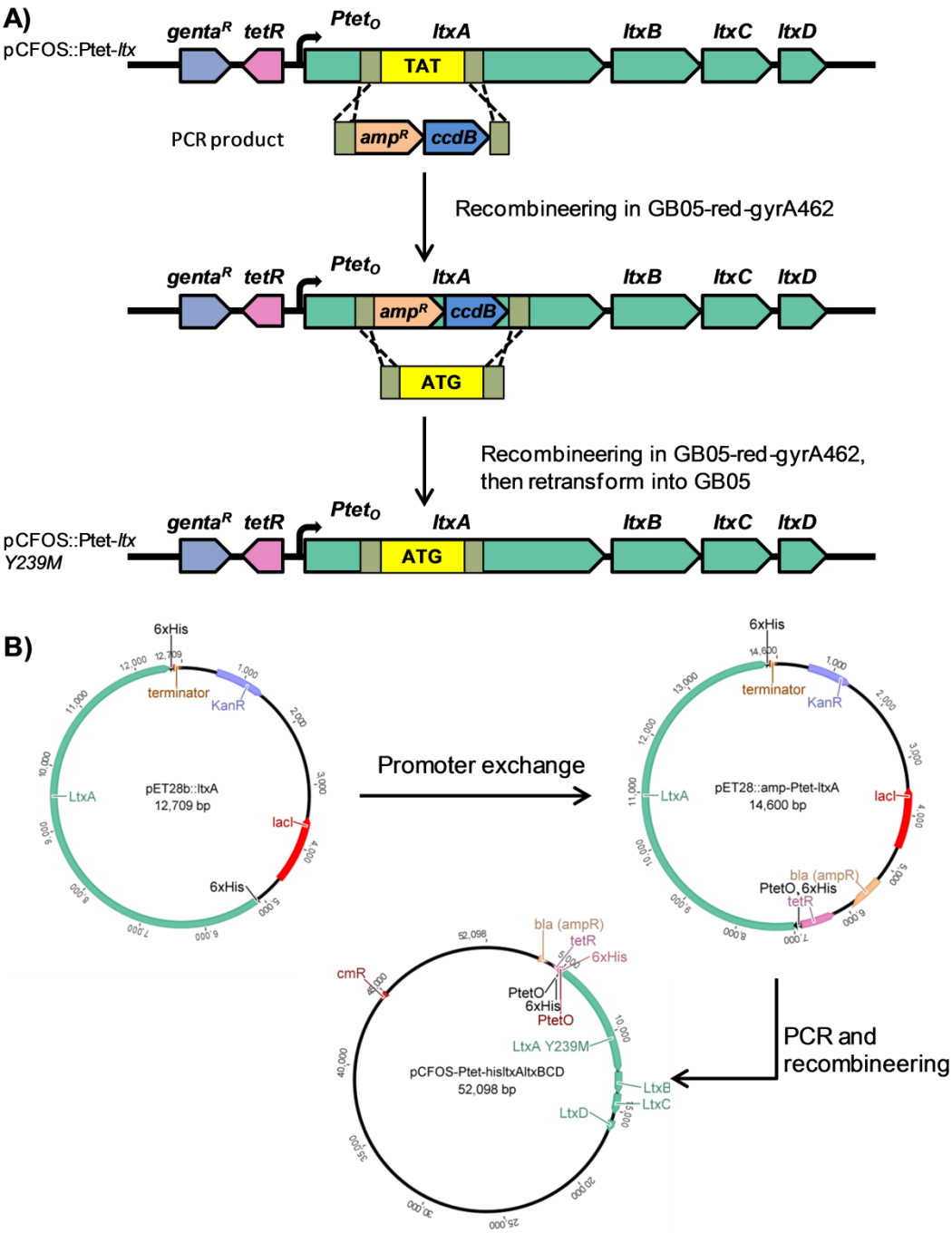


Figure A 1.3 Expression plasmid construction of wild-type and engineered Ltx biosynthetic pathway. (**A**) Site-directed mutagenesis using selection-counterselection recombineering. Firstly, the *amp*^{R}*-ccdB* cassette was inserted into the *ltx* fosmid replacing the codon to be altered. Next, a single-stranded oligonucleotide containing the specific mutation was inserted to replace the *amp*^{R}*-ccdB* cassette. Successful recombination was selected by growth in *ccdB* sensitive *E. coli* strain. The pathway is under the control of a tetracycline-inducible promoter. In the absence of tetracycline, the repressor protein binds the operator region of the promoter (Pteto) to inhibit transcription. **(B)** Plasmid map of His-tagged LtxA vectors constructed in this study.

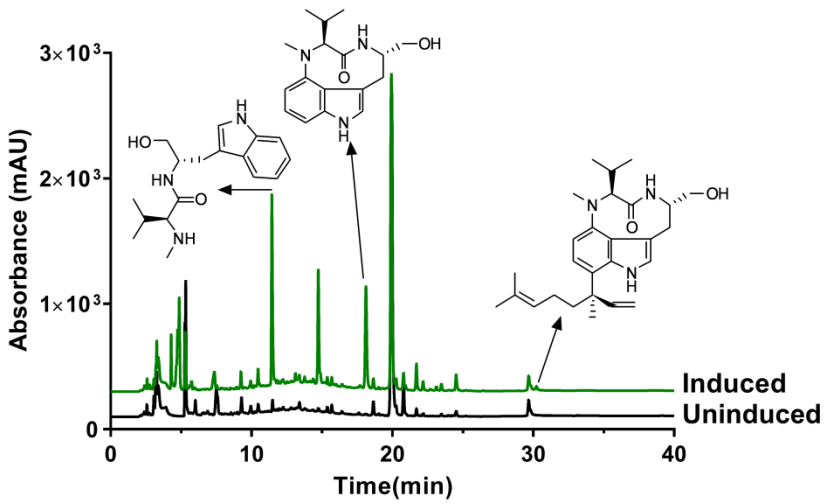


Figure A 1.4 Production of LTX pathway metabolites in *E. coli*-MtaA harbouring pCFOS-Ptet-HisltxAltxBCD. HPLC chromatogram of the tetracycline-induced and the uninduced cultures. Lyngbyatoxin biosynthetic pathway were eluted with retention times of 5.50 min (N-methyl-L-isoleucyl-L-tryptophanol), 7.05 min (indolactam-V), and 9.51 min (lyngbyatoxin A)

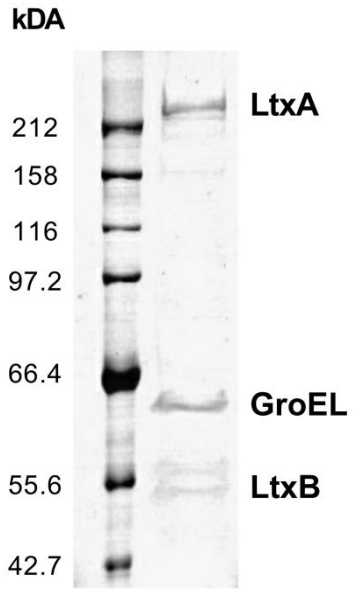


Figure A 1.5 SDS PAGE of purified N-HisLtxA. Full-length bimodular N-HisLtxA was purified using affinity chromatography and gel filtration. Lane 1: unstained protein marker, broad range, 2-212 kDa (NEB). Lane 2: protein fraction containing N-HisLtxA (277 kDa); chaperone GroEL, (57 kDa) and LtxB (52 kDa). The identities of proteins were confirmed by trypsin digestion and MS/MS. The yield of LtxA protein cultured in TB media was 0.085 mg/L.

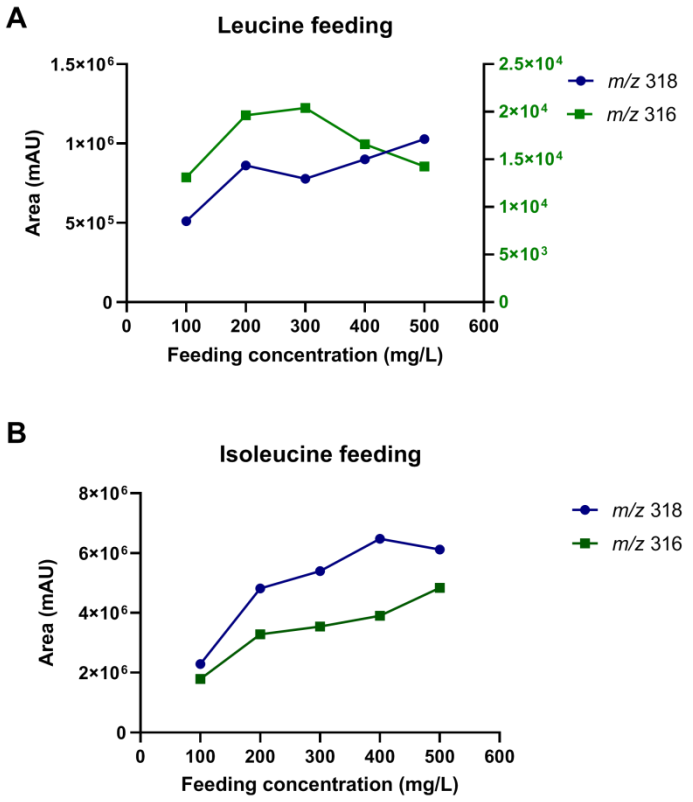


Figure A 1.6 The dose-dependent effect of amino acid supplementation into lyngbyatoxin intermediate production. LC-MS analysis of extracts of induced *E. coli* GB05-MtaA pCFOS - Ptet-ltx cultured in M9 minimal medium supplemented various concentration of with L-leucine (**A**) and L-isoleucine (**B**) of methylated lyngbyatoxin intermediates. The detected ions with an m/z of 318 are putatively either NMIT (N-methyl-L-isoleucyl-L-tryptophanol) or NMLT (N-methyl-L-leucyl-L-tryptophanol). The detected ions with an m/z of 316 are putatively either ILI (Indolactam-I) or ILL (Indolactam-L). As *E. coli* GB05-MtaA is a leucine auxotroph, M9 minimal under the supplementation of 100 mg/L L-Leucine was used as the baseline (=0 feeding concentration).

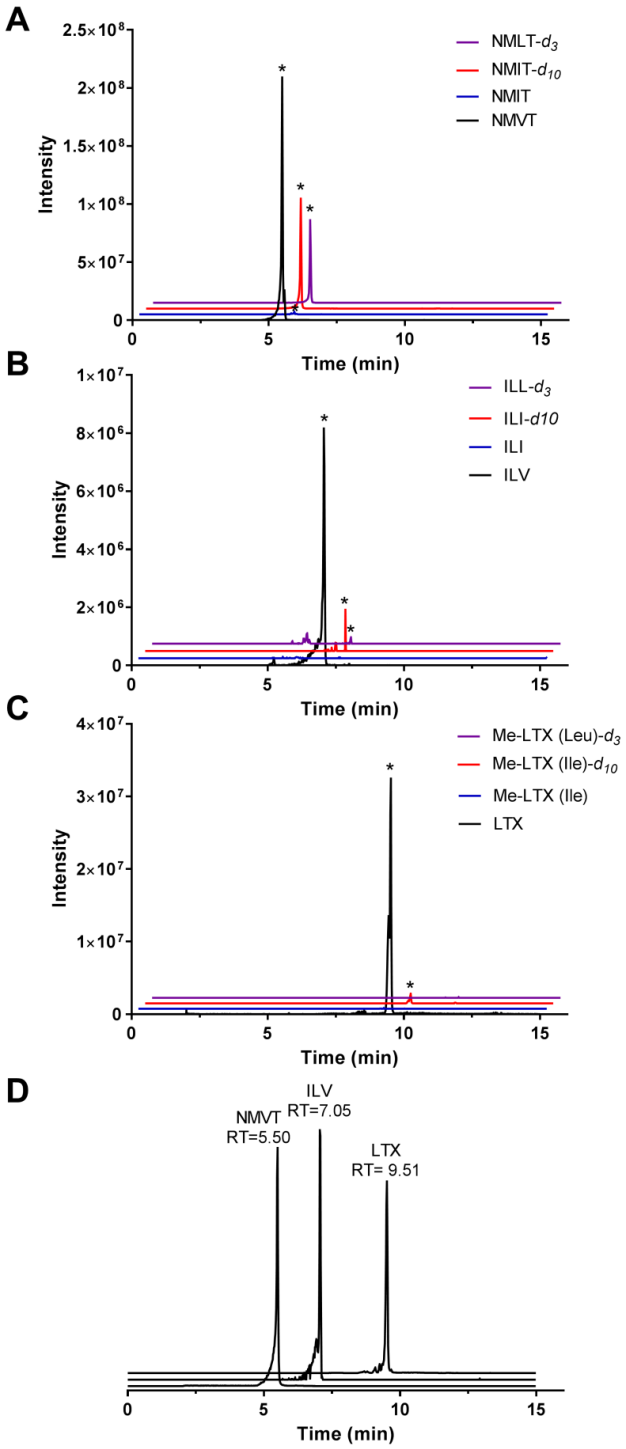


Figure A 1.7 Production lyngbyatoxin pathway products from induced *E. coli* GB05-MtaA pCFOS -Ptet-ltx cultured in M9 minimal medium supplemented with L-leucine-5,5,5-d_3 and L-isoleucine-d_{10}. Chromatogram of HPLC-ESI-HRMS analysis of (A) dipeptide (B) Indolactam (C) Lynbyatoxins (D) authentic standards of Val-containing products. The corresponding HPLC-ESI-HRMS/MS data were presented in Figure 2.3, 2.4, and 2.5 for dipeptide, indolactam, and lyngbyatoxin, respectively.

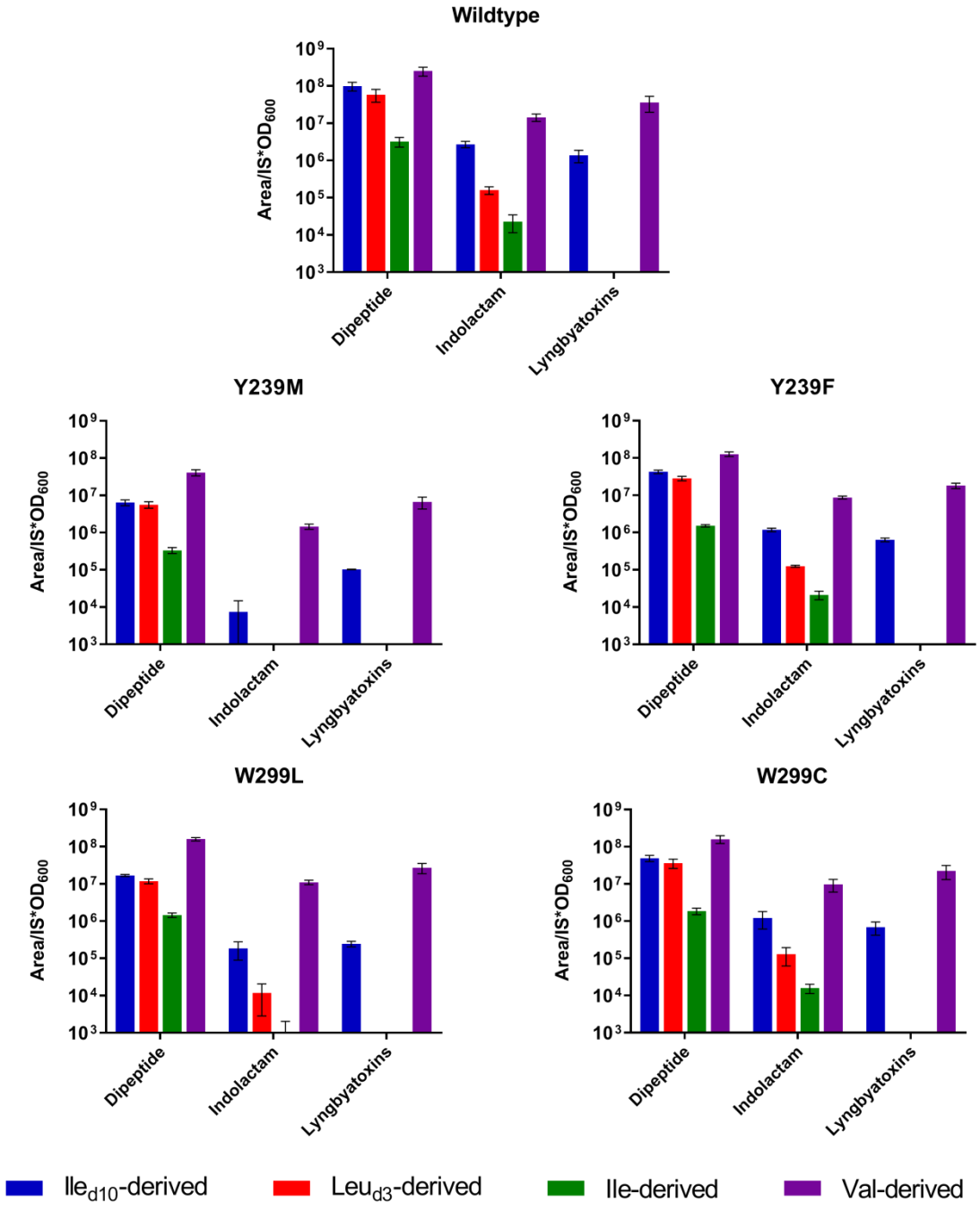


Figure A 1.8 Yields of LTX analogues and intermediates from minimal medium fermentation of LtxA-A1 wild-type single mutants determined by HPLC-ESI-HRMS. Cultures were supplemented with L-leucine-5,5,5-d_3 and L-isoleucine-d_{10} supplemented to a concentration of 200 mg/L. Measurement of peak area were normalised against caffeine and Gly-Phe as internal standards as well as culture OD_{600} at XAD-7 addition. Error bars represent SEM from 3 biological replicates.

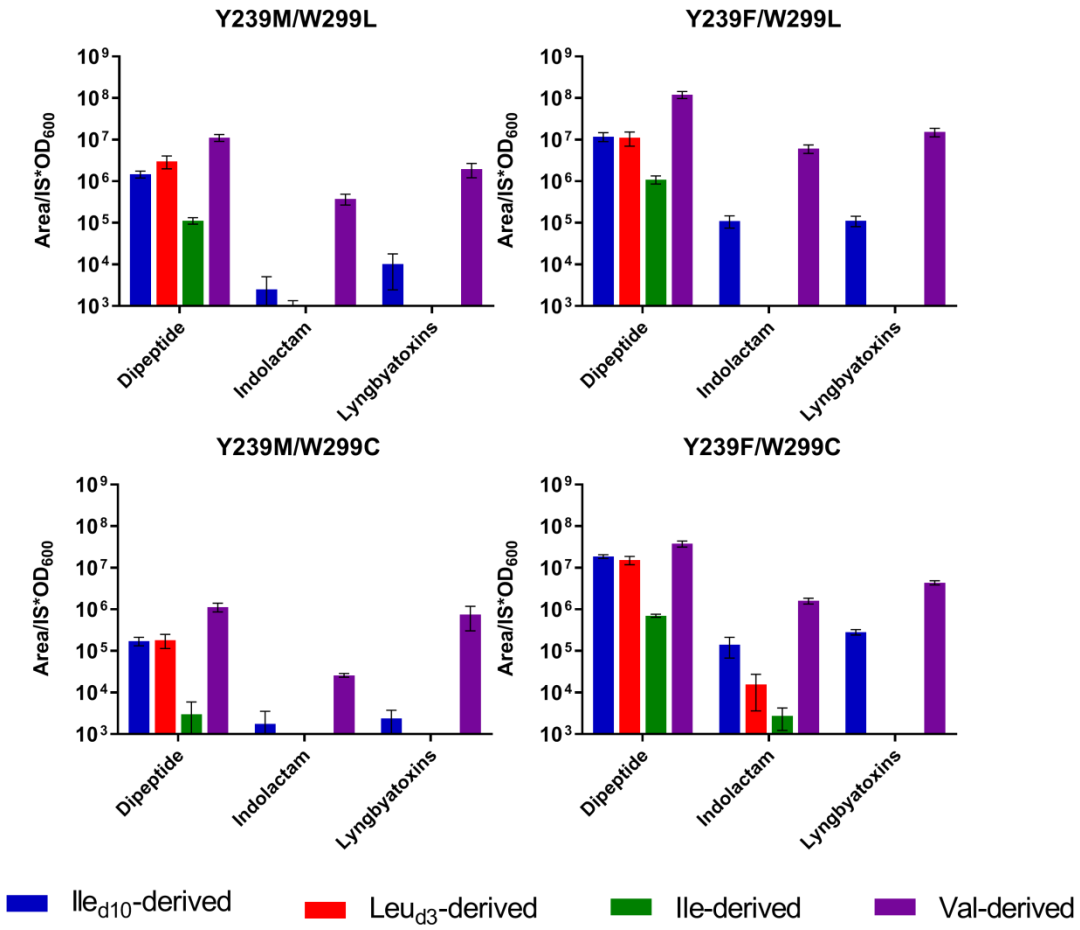


Figure A 1.9 Yields of LTX analogues and intermediates from minimal medium fermentation of LtxA-A1 double mutants determined by HPLC-ESI-HRMS. Cultures were supplemented with L-leucine-5,5,5-d_3 and L-isoleucine-d_{10} supplemented to a concentration of 200 mg/L. Measurement of peak area were normalised against caffeine and Gly-Phe as internal standards as well as culture OD_{600} at XAD-7 addition. Error bars represent SEM from 3 biological replicates.

Appendix 2. Chapter 3 Supplementary information

Table A 2.1 One-Way ANOVA analysis of LC-MS data on the metabolite with an *m/z* of 187.

	Negative Control	**Uninduced**	**Induced**
Mean ± SEM	1 843 ± 31	2 573 ± 194	33 238 ± 2869
One-Way ANOVA			
P value	<0.0001		
P value summary	****		
R square	0.9395		
Tukey's multiple comparisons test			
	Negative control vs. Uninduced	**Negative control vs. Induced**	**Uninduced vs. Induced**
Mean Diff.	-729.5	-31 395	-30 666
P Value	0.9484	<0.0001	<0.0001
Significant	NS	***	***

Values were obtained from analysis using Progenesis software and Graphpad Prism based on *E. coli* fermentations performed with six biological replicates (n = 6) per group. NS, not significant; SEM, standard error of the mean; * denotes the degree of significance.

Table A 2.2 One-Way ANOVA analysis of LC-MS data on the metabolite with an *m/z* of 173.

	Negative Control	**Uninduced**	**Induced**
Mean ± SEM	6 ± 2	50 100 ± 5924	63 694 ± 4385
One-Way ANOVA			
P value	<0.0001		
P value summary	***		
R square	0.8923		
Tukey's multiple comparisons test			
	Negative control vs. Uninduced	**Negative control vs. Induced**	**Uninduced vs. Induced**
Mean Diff.	-50 093	-63 687	-13 594
P Value	<0.0001	<0.0001	0.0933
Significant	****	****	NS

Values were obtained from analysis using Progenesis software and Graphpad Prism based on *E. coli* fermentations performed with six biological replicates (n = 6) per group. NS, not significant; SEM, standard error of the mean; * denotes the degree of significance.

Table A 2.3 Putative internal promoter within sxtA corresponds to σ^{70} promoter elements.

	(–35)-N(x)-(–10)
***E. coli* consensus sequence**	TTGACA-----(N_{17})------TATAAT
Predicted internal *sxtA* promoter	G**TTACA**------(N_{13})------**TATA**C**T**

The bolded text indicates homology to the *E. coli* consensus sequence

Table A 2.4 Predicted and observed primary and secondary ions in MS2 spectra from the stable isotope precursor feeding experiment correspond to the Int-A' masses.

		Chemical formula	Predicted *m/z*	Observed *m/z*		
				M9 media	M9 + [guanidino-$^{15}N_2$]arginine	M9 + [1,2^{13}C]acetate
Int-A'						
Parent ion	$[M+H]^+$	$C_8H_{19}N_4O^+$	187.1553	187.1560	187.1566	187.1560
Fragment ions	$[M-NH_3]^+$	$C_8H_{16}N_3O^+$	170.1288	170.1295	170.1297	170.1293
	$[M-CH_5N_3]^+$	$C_7H_{14}NO^+$	128.1070	128.1075	128.1074	128.1073
	$[M-CH_7N_3O]^+$	$C_7H_{12}N^+$	110.0964	110.0968	110.0968	110.0968
	$[M-C_7H_{13}NO]^+$	$CH_6N_3^+$	60.0556	60.0563*	60.0563*	60.0562*
$^{15}N_2$-Int-A'						
Parent ion	$[M+H]^+$	$C_8H_{19}{}^{15}N_2N_2O^+$	189.1494	ND	189.1503	ND
Fragment ions	[M-NH_3]+	$C_8H_{16}{}^{15}N_2NO^+$	172.1229	-	172.1234	-
	[M-$CH_5{}^{15}N_2N$]+	$C_7H_{14}NO^+$	128.1070	-	128.1073	-
	[M-$CH_7{}^{15}N_2N_3$]+	$C_7H_{12}N^+$	110.0964	-	110.0968	-
	[M-$C_7H_{13}NO$]+	$CH_6{}^{15}N_2N^+$	62.0497	-	62.0504*	-
$^{13}C_2$-Int-A'						
Parent ion	$[M+H]^+$	$^{13}C_2C_6H_{19}N_4O^+$	189.1620	ND	ND	189.1627
Fragment ions	[M-NH_3]+	$^{13}C_2C_6H_{16}N_3O^+$	172.1355	-	-	172.1362
	[M-CH_5N_3]+	$^{13}C_2C5H_{14}NO^+$	130.1137	-	-	130.1143
	[M-CH_7N_3O]+	$^{13}C_2C_5H_{12}N^+$	112.1031	-	-	112.1036
	[M-$^{13}C_2C_5H_{13}NO$]+	$CH_6N_3^+$	60.0556	-	-	60.0564*

Listed fragmented ions are based on *m/z* reported by Tsuchiya et al.(Tsuchiya *et al.*, 2014) and fragment predictions by Mass Frontier. The underlined values indicate mass shifts from the incorporation of stable isotope labelled precursors. Mass accuracy is < 10 ppm. ND, not detected; *, indicate mass detected with error < 15 ppm.

Table A 2.5 Predicted and observed primary and secondary ions in MS2 spectra from the stable isotope precursor feeding experiment correspond to the demetInt-A' masses.

		Chemical formula	Predicted *m/z*	Observed *m/z*		
				M9 media	M9 + [guanidino-$^{15}N_2$]arginine	M9 +[1,2^{13}C]acetate
demetInt-A'						
Parent ion	$[M+H]^+$	$C_7H_{17}N_4O^+$	173.1397	173.1408	173.1411	173.1409
Fragment ions	$[M-NH_3]^+$	$C_7H_{14}N_3O^+$	156.1131	156.1137	156.1138	156.1137
	$[M-CH_5N_3]^+$	$C_6H_{12}NO^+$	114.0913	114.0917	114.0917	114.0917
	$[M-CH_7N_3O]^+$	$C_6H_{10}N^+$	96.0808	96.0811	96.0811	96.0811
	$[M-C_6H_{11}NO]^+$	$CH_6N_3^+$	60.0556	60.0563*	60.0563*	60.0563*
$^{15}N_2$-demetInt-A'						
Parent ion	$[M+H]^+$	$C_7H_{17}{}^{15}N_2N_2O^+$	175.1338	ND	175.1342	ND
Fragment ions	$[M-NH_3]^+$	$C_7H_{14}{}^{15}N_2NO^+$	158.1072	-	158.1080	-
	$[M-CH_5{}^{15}N_2N]^+$	$C_6H_{12}NO^+$	114.0913	-	114.0917	-
	$[M-CH_7{}^{15}N_2N_3]^+$	$C_6H_{10}N^+$	96.0808	-	96.0811	-
	$[M-C_6H_{11}NO]^+$	$CH_6{}^{15}N_2N^+$	62.0497	-	62.0503	-
$^{13}C_2$-demetInt-A'						
Parent ion	$[M+H]^+$	$^{13}C_2C5H_{17}N_4O^+$	175.1464	ND	ND	175.1465
Fragment ions	$[M-NH_3]^+$	$^{13}C_2C5H_{14}N_3O^+$	158.1198	-	-	158.1192
	$[M-CH_5N_3]^+$	$^{13}C_2C4H_{12}NO^+$	116.0981	-	-	116.0984
	$[M-CH_7N_3O]^+$	$^{13}C_2C_4H_{10}N^+$	98.0875	-	-	ND†
	$[M-^{13}C_2C_4H_{13}NO]^+$	$CH_6N_3^+$	60.0556	-	-	60.0560

Listed fragmented ions are based on fragment predictions by Mass Frontier. The underlined values indicate mass shifts from the incorporation of stable isotope labelled precursors. Mass accuracy is < 10 ppm. ND, not detected; *, indicate mass detected with error < 15 ppm; †, not detected possibly due to low parent ion abundance.

MATRIX SCIENCE **MASCOT Search Results**

Protein View: gi|114462352

polyketide synthase-related protein [Cylindrospermopsis raciborskii T3]

Database:	NCBInr
Score:	7866
Monoisotopic mass (M_r):	140019
Calculated pI:	5.87
Taxonomy:	Cylindrospermopsis raciborskii T3

Sequence similarity is available as an NCBI BLAST search of gi|114462352 against nr.

Search parameters

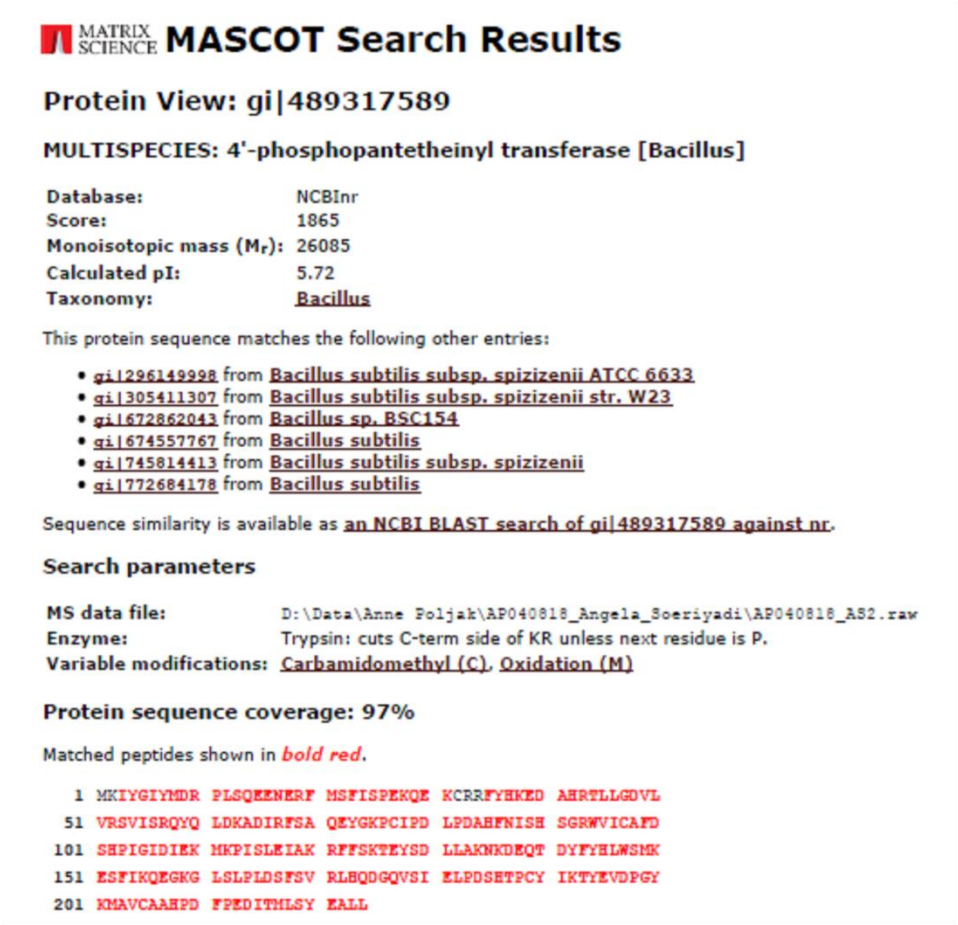
MS data file:	D:\Data\Anne Poljak\AP040818_Angela_Soeriyadi\AP040818_AS1.raw
Enzyme:	Trypsin: cuts C-term side of KR unless next residue is P.
Variable modifications:	Carbamidomethyl (C), Oxidation (M)

Protein sequence coverage: 88%

Matched peptides shown in ***bold red***.

```
   1 MLQKINRYTH GFVAVPVILA CREKGVFELL ADESPLSLNQ MVEHLGANSG
  51 HFQVALRMLE SLHWLSRNKE LKYSLTAEAA IHNKISEDIL QLYNLPIQSY
 101 LEGKQGNLLG RWIERSCQLW NLDNPLMADF LDGLLVIPLL LALHKHNLLA
 151 DSEDKPLLSS LSSTVQEELG KLFLHLGWAD LTAGRLTITE LGRFMGERAL
 201 NTAIVASYTP MLSRIHDVLF GNCLSVFQRD ASGHERHIDR TLNVIGSGFQ
 251 HQKYFADLEE SILSVFNQLP LEEQPKYITD MGCGDGTLLK RVWETIQFKS
 301 ARGKALEQYP LRLIGVDYNE ASLKATTRTL ASLPHLVLQG DIGNPEQMVR
 351 SLEAHGIHDP ENILHIRSFL DHDRLFIPPQ KRNELKERAH LPYQSVCVDD
 401 QGELIPPHVM VQSLVEHLER WSQVVNKHGL MILEVHCLEP RVVYQFLDKS
 451 ENLHFDAPQG FSQQYLVEAE VFLMSAAQVG LPPKLELSKR YPKTFPFTRI
 501 TLNYFEKRPY KISHAYLSDL PALVDLEVKC WPENLRASTH EIRRRLELNP
 551 QGNLVLIIED QIIGAIYSQT ITSTEALENV KYAQVPTLHT PQGSVIQLLA
 601 LNILPEFQAR GLGNELRDFM LYYCTLKGGI ESVVGVTRCR NYVNYSQMPM
 651 MEYLKLHNEQ RQLLDPIVGF HVSGGAEIRG IIANYRPEDT DNLGMGILIE
 701 YNLRDSALHS PGDRKGPYIN SAIGSLVPKA TSATKENKTV ADLVKECILK
 751 VMGSQRQAAY APQQKLLDMG LDSLDLLELQ TLLEERLGIN LSGTFFLQKN
 801 TPTAIITYFQ NQVVQEKQSD LAPPVDSANE INTLENVVNQ QKIPQVTRVV
 851 TEQQGRKVLI DGHWVIDFAS CNYLGLDLHP KVKEAIPPAL DKWGTHPSWT
 901 RLVASPAIYE ELEEELSKLL GVPDVLVFPA VTLLQIGILP LLTGNNGVIF
 951 GDIAAHRCIY EACCLAQHKG AQFIQYRHND LNDLAEKLAK YPPEQVKIIV
1001 IDGVYSMSAD FPDLPAYVHL AKEYNALIYM DDAHGFGILG ENPSSDMPYG
1051 YKGNGMVNYF DLRFAEDNII YVAGLSKAYS SYAAFLTCGD RRIKTNFRNA
1101 WTAIFSGPSP VASLASALAG LQVNRQEGEQ LRKQIYHLTH KLVTQARAIG
1151 FEVDNYGYVP IVGVLVGDAQ HMIDVCQLLW EYGILITPAI FPIVPLNKSA
1201 LRFSITAANT EEEIDQAIKS LKAVWDLLQK RKALPCKQEE NILKH
```

Figure A 2.1 Identification of purified SxtA protein by trypinolysis and peptide mass fingerprinting. The band on the SDS-PAGE (Figure 3.2) was cut and digested by trypsin prior to LC-MS/MS analysis. The results were blasted against the Mascot server at the Bioanalytical Mass Spectrometry Facility (UNSW Sydney, Australia). The band immediately below 150 kDa was identified as SxtA.



MATRIX SCIENCE **MASCOT Search Results**

Protein View: gi|489317589

MULTISPECIES: 4'-phosphopantetheinyl transferase [Bacillus]

Database:	NCBInr
Score:	1865
Monoisotopic mass (M_r):	26085
Calculated pI:	5.72
Taxonomy:	Bacillus

This protein sequence matches the following other entries:

- gi|296149998 from Bacillus subtilis subsp. spizizenii ATCC 6633
- gi|305411307 from Bacillus subtilis subsp. spizizenii str. W23
- gi|672862043 from Bacillus sp. BSC154
- gi|674557767 from Bacillus subtilis
- gi|745814413 from Bacillus subtilis subsp. spizizenii
- gi|772684178 from Bacillus subtilis

Sequence similarity is available as an NCBI BLAST search of gi|489317589 against nr.

Search parameters

MS data file:	D:\Data\Anne Poljak\AP040818_Angela_Soeriyadi\AP040818_AS2.raw
Enzyme:	Trypsin: cuts C-term side of KR unless next residue is P.
Variable modifications:	Carbamidomethyl (C), Oxidation (M)

Protein sequence coverage: 97%

Matched peptides shown in *bold red*.

```
  1 MKIYGIYMDR PLSQEENERF MSFISPEKQE KCRRFYHKED AHRTLLGDVL
 51 VRSVISRQYQ LDKADIRFSA QEYGKPCIPD LPDAHFNISH SGRWVICAFD
101 SHPIGIDIEK MKPISLEIAK RFFSKTEYSD LLAKNKDEQT DYFYHLWSMK
151 ESFIKQEGKG LSLPLDSFSV RLHQDGQVSI ELPDSETPCY IKTYEVDPGY
201 KMAVCAAHPD FPEDITMLSY EALL
```

Figure A 2.2 Identification of purified Sfp protein by trypinolysis and peptide mass fingerprinting. The band on the SDS-PAGE (Figure 3.2) was cut and digested by trypsin prior to LC-MS/MS analysis. The results were blasted against the Mascot server at the Bioanalytical Mass Spectrometry Facility (UNSW Sydney, Australia). The band immediately above 30 kDa was identified as Sfp.

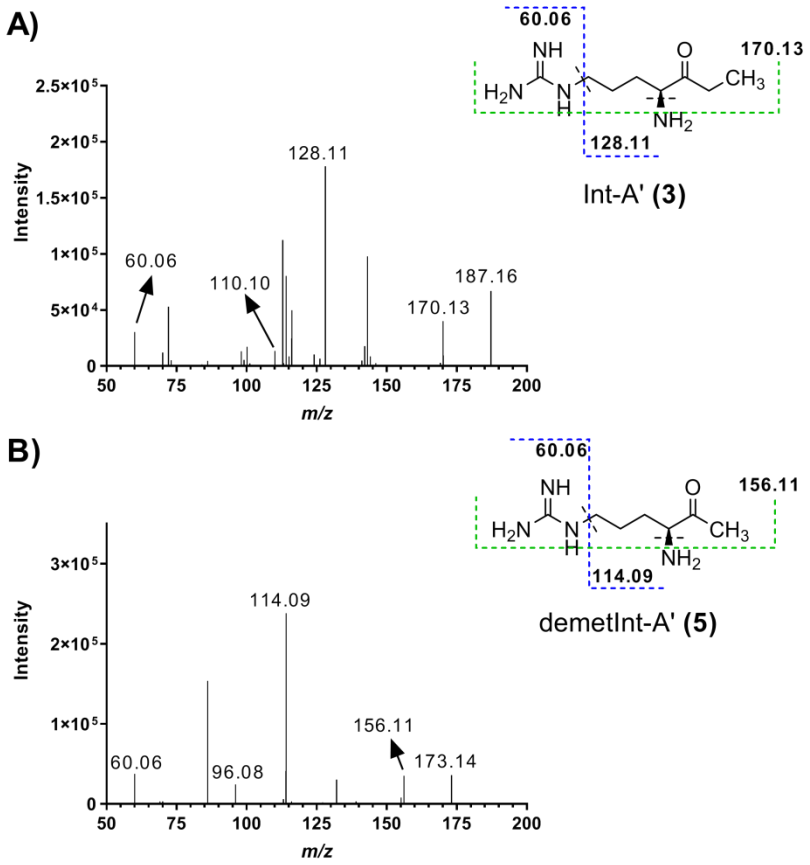


Figure A 2.3 HPLC-ESI-HRMS/MS analysis of Int-A' and demetInt-A' from *E. coli* pVB::*sxtA* pET::*sfp* induced extracts cultured in Terrific Broth. Fragmentation of Int-A' (**A**) and demetInt-A' (**B**), indicating the lack of a methyl group in demetInt-A' in comparison to Int-A'.

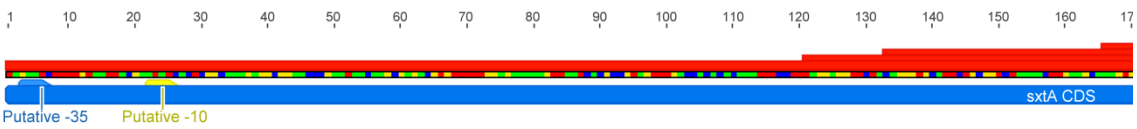


Figure A 2.4 Promoter and ORF analysis performed with Geneious 9.0.4. The SxtA CDS is highlight in blue and the putative ORF predicted by Geneious 9.0.4 (Biomatters) is highlighted in orange.

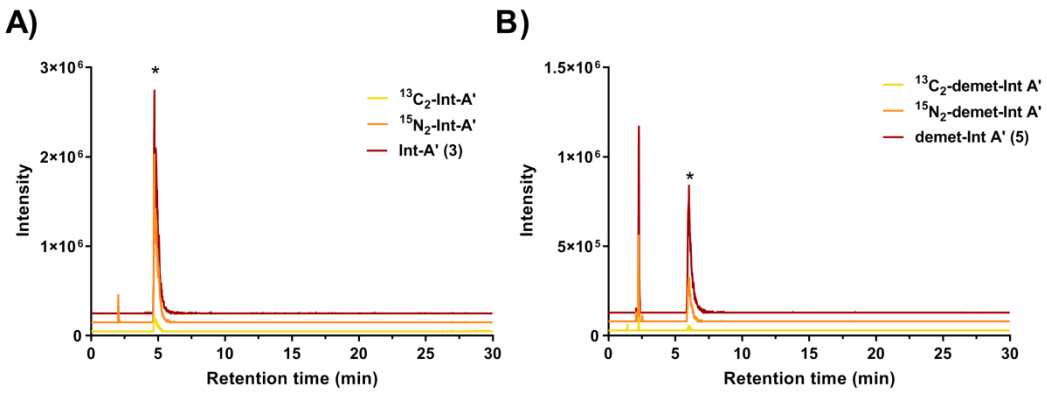


Figure A 2.5 HPLC-ESI-HRMS analysis of Int-A' and demetInt-A' production in M9 minimal medium. Cultures were grown with and without supplementation of [guanidino-$^{15}N_2$] arginine and [1,2^{13}C] acetate. (**A**) Extracted ion chromatograms for *m/z* = 187.1553 (Int-A'), *m/z* = 189.14941 ($^{15}N_2$ Int-A') and *m/z* = 189.16205 ($^{13}C_2$ Int-A'). (**B**) Extracted ion chromatograms for *m/z* = 173.13969 (demetInt-A'), *m/z* = 175.13376 ($^{15}N_2$ demetInt-A'), and *m/z* = 175.14640 ($^{13}C_2$ demetInt-A').

Appendix 3. Chapter 4 Supplementary information

Supplementary material A 3.1 MacConkey agar plates recipe

Peptone (Difco) or Gelysate (BBL)	17.0 g
Proteose peptone (Difco) or Polypeptone (BBL)	3.0 g
Sugars (mannitol/lactose/maltose/xylose)	10.0 g
NaCl	5.0 g
Neutral Red	0.03 g
Agar	13.5 g
Distilled Water	Add to make 1 L

Adjust to pH 7.1 +/-0.2 and sterilise at 121°C for 15 min.

Table A 3.1 Primers used in this study.

Primer name	Sequence 5'-3'
Integration of sxt fragment 3	
Gibson assembly for sxt fragment 3 variants	
3sub1_HA1_F	TCAGAACAGTACCATGACTGATTCGGTGTCAGGAGGCG
3sub1_HA2_R	CTGCCCAAACATCATGGATCAGAAATTTCGGATATG
3sub1PmRBS_HA2_R	GTCAGCATATGTTCATGACTCCATTATTATTGTTTCTG
3sub2sxtR_HA2_F	GATCCATGATGTTTGGGCAGATCCGGCA
3sub2sxtU_HA3_R	TTTCATTAACGTTAACATGATCCGGCTGGGTAACTGCAT
3sub2sxtR_HA2b_R	TATCTCCTTCTCAGTCCAGATCATAATACAGCCAGA
3sub2sxtS_HA2bF	TCTGGACTGAGAAGGAGATATACATATGCTGACC
3sub2sxtU_HA3_R2	ATTTCATTAACGTTAACATGATCCGGCTGGGTAACTG
3sub2sxtS_HA3_F	CATGAACATATGCTGACCGCAGAACAG
3sub3_HA2_F	CCAGCCGGATCATGTTAACGTTAATGAAATTCTGATTC
3sub3_pVB_R	CTGACTCTAGCTAGAGGATCATTAAAGCTGGGACATTGCTC
Fragment amplification	
xylFH1F	GGTGTCAGGAGGCGGGA
xylBH2R	ATTAAAGCTGGGACATTGCTC
Screening and sequencing primers	
xylF4232_4249F	AGCTCGCTCTCTTTGTGG
sxtQ_R	ATCAGCAGATTCGGATGAG
sxt3_5539-5557R	GACAAACCCAAACCAGACC

sxt3_5305-5322F	ACTGGTTCTGTGGCGTAGC
sxt3_6741-6724R	CACTTCTTGACGCAGTGC
sxt3_6514-6533F	GCAATCGATACCGGTAATCC
sxt3_7734-7717R	ATAAAACGGAGCTGCACG
sxt3_8815-8797R	GCATCTGCAGGACAATATGC
sxt3_VF	TTACTGCTGGCACAGGATG
sxt3_8516-8533F	ATTACCCGGAAACGGATC
sxt3_9749-9731R	CGTTTCAGAATGCTGTTCC
sxt3_9528-9546F	CCTGATCGTGGAAGATACC
xylB_7987-7970R2	CTCGCCGGAAAGATAAGG

PPTase replacement (tetA-sacB approach)

Insert amplification

tetA_ mtlA_ F	TATTAAGATCAAAGTGCAAAGCTTTGGTCGTTTCCTCAGCAA CATGGTGATCCTAATTTTTGTTGACACTCTATC
sacB_mtlD_R	TTGCATTGCTTTATAAGCGGTTACCGCCTCGGATACAACCTC GCTGTTGGATCAAAGGGAAAACTGTCCATATGC
T7 sfp _ mtlA_F	TATTAAGATCAAAGTGCAAAGCTTTGGTCGTTTCCTCAGCAA CATGGTGATAATACGACTCACTATAGGGGAATTGTGAG
T7term_ mtlD_R	TTGCATTGCTTTATAAGCGGTTACCGCCTCGGATACAACCTC GCTGTTGGTTCGCCAATCCGGATATAGTTC

Screening and sequencing primers

TP1_mtlA_F	GCTTTGGTCGTTTCCTCAGC
TP2_ mtlD_R	GGTTACCGCCTCGGATACAAC

Integration of *sxtPFM*

Fragment amplification

H1melF	AAGCCTGCCGTCAGGG
H2melR	AGCGCAACGATGGCTTTAAG

Screening and sequencing primers

melR_F	AGTTTGCGGTATTGTTGC
Sxt4_932-949R	GAACATTCAGTGCGCTGG
Sxt4_732-750F	AGTTAGCCGTATTGCAAGC
Sxt4_1978-1959R	TAATATTGGCCAGTGCAACC
Sxt4_1710-1728F	GGCAATTATTCTGGGTTGG
Sxt4_2963-2946R	GCAATACCAAATGATCCG
Sxt4_2744-2763F	TAGCAGTTGCAAATACCAGC
Sxt4_3942-3925R	AACAACGCTTTCCACACC

Sxt4_3731-3748F	TACCGGAATTAGGCCTGG
melB_R	GAACAAATCCGCATCACC

Gene knockout of *sxtC, sxE, sxtL, sxtJKL* (tetA-sacB approach)

Insert amplification

tetA_HstB_F	AAGTGAAACTGCGCTATGCACCGGCAACCCCGAGCAGTAAC AAACTGTGATCCTAATTTTTGTTGACACTCTATC
sacB_sxt1term_R	AAAAAAGGCCCCCCTTTCGGGAGGCCTCTTTTCTGGAATTTG GTACCGAGATCAAAGGGAAAACTGTCCATATGC
sxtB_F	ATGACCCATGTTGCCCTG
Sxt1term_ HsxtB_R	AAAAAAGGCCCCCCTTTCGGGAGGCCTCTTTTCTGGAATTTG GTACCGAGTCACAGTTTGTTACTGCTCGGGG
tetA_HstK_F	TTTTTGCACAGAGCAGCGCAATTGCACCGTTTATCTATACCC TGTTTTGATCCTAATTTTTGTTGACACTCTATC
sacB_ sxt2term_R	AATAGGAACTTCGACAAAAAACCCCTAAGACGGCCGATGG GGTTAGAGTAATCAAAGGGAAAACTGTCCATATGC
Sxt2_9415-9435F	CGCGTTATTATGAACTGATCC
Sxt2term _HsxtK_R	AATAGGAACTTCGACAAAAAACCCCTAAGACGGCCGATGG GGTTAGAGTATCAAAACAGGGTATAGATAAACGG
tetA_HsxtI_F	CTCGTGGTAATAGTGATGAAAGCTGGCAGAAAGAATTTGAG CTGGATTGATCCTAATTTTTGTTGACACTCTATC
Sxt2term_HsxtI_R	AATAGGAACTTCGACAAAAAACCCCTAAGACGGCCGATGG GGTTAGAGTATCAATCCAGCTCAAATTCTTTCTGC
tetA_HsxtG_F	ACCGCCTGGATGATCAGAGCGTTACCGTTAAAACCATTCTG GGTTTCTGATCCTAATTTTTGTTGACACTCTATC
sacB_ HsxtG_R	TCATTTTCCAGCTCTTGGTTGTTCTGATTGGTCATATGTATAT CTCCTTCATCAAAGGGAAAACTGTCCATATGC
sxtD_F	ATGATTGATACCATTAGCGTTCTGCTGC
sxtD_ HsxtG_R	TCATTTTCCAGCTCTTGGTTGTTCTGATTGGTCATATGTATAT CTCCTTCTCAGAAACCCAGAATGGTTTTAACGG

Screening and sequencing primers

Sxt1_7384-7401F	TTGATAGCAGCGTTGTGG
lacA_6017-5999_r	TGTTACCGTGTAGTCATCG
Sxt2_9415-9435F	CGCGTTATTATGAACTGATCC
malG8358-8342R	AGCGGCATAACATTGGC
malK4237-4256F	TCAGGAGATGGCTTAAATCC
sxtG_R	GTTCTTCAAACGGAAAGGTC

Primers for transcriptional analysis

sxtA_F	CAGGTCGTGCAAGAGAAACA
sxtA_R	CATACGGCATATCGCTGCT
sxtC_F	TTATAGCGTCGTTGGCCTG
sxtC_R	CATTTTCATTTCCAGGCGT
sxtD_F	CCGGTGTTATGTATGTTCCGC
sxtD_R	CGGAAACACATAGGTGCCAA
sxtI_F	CAGGCGATCGTATGAAAGGT
sxtI_R	AGGTAATCCATCTCGGTACGC
sxtL_F	CGTGCACTGTCCCAGAAA
sxtL_R	CCAATGACCAACACAACCCT
sxtS_F	CGCCATGAAGTGGAACTGAT
sxtS_R	GGTATTTGCTGCTGCGGTAT
sxtU_F	CACTGCCGATTGTTACCGAT
sxtU_R	GCGGAGTAATGGTTGACAGC
sxtX_F	GCAGCCTATAAAGTCGCACC
sxtX_R	GTTCATCCAGGTCATCACGG
sfp_F	TGCTCGTTCGCTCAGTCATA
sfp_R	ATAGCATGGGGAATGGCTGT
27F	AGAGTTTGATCCTGGCTCAG
1492R	GGTTACCTTGTTACGACTT

*Underlined sequences are homology arms used to mediate linear-circular homologous recombination (LCHR) or Gibson cloning.

Table A 3.2 List of genes and their putative function in the *sxt* biosynthetic gene cluster of *R. raciborskii* T3 included in this study. Description of the protein family and putative functions are adapted from Kellmann et al. (Kellmann *et al.*, 2008)

Gene	Protein family	Putative function
sxtA	PKS-like (Methyltransferase, GNAT, ACP, AONS)	Formation of the first intermediate, Int-A
sxtB	Cysteine deaminase	Cyclisation
sxtC	Unknown	Unknown (possible regulatory*)
sxtD	Sterole desaturase-like protein	Desaturation
sxtE	Unknown	Unknown
sxtG	Amidinotransferase	Amidinotransfer
sxtH	Phenyl propionate dioxygenase	C-12 hydroxylation
sxtI	Carbamoyl transferase	Carbamoylation
sxtJ	Unknown	Unknown, might be coupled with sxtI
sxtK	Unknown	Unknown, might be coupled with sxtI
sxtL	GDSL-lipase	decarbomylation
sxtQ	Unknown	Unknown
sxtR	Acyltransferase	Unknown
ORF24	Unknown	Unknown
sxtS	Phytanoyl-CoA dioxygenase	Ring formation
sxtT	Phenylpropionate dioxygenase	C-12 hydroxylation
sxtU	Alcohol dehydrogenase	Reduction of C-1
sxtV	Succinate dehydrogenase	Dioxygenase reductase
sxtW	Ferredoxin	Electron carrier
sxtX	Cephalosporin hydroxylase	N-1 hydroxylation
sxtF	MATE transporter	Export of PSTs
sxtP	RTX toxin	Binding of PSTs
sxtM	MATE transporter	Export of PSTs

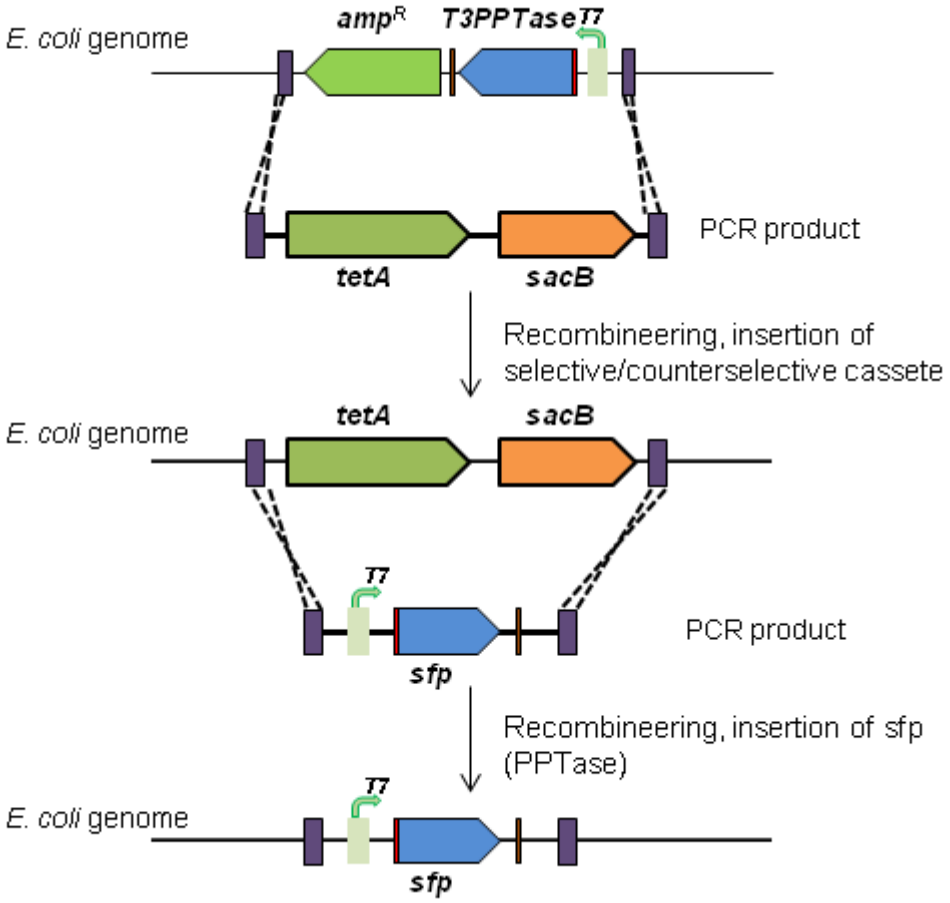


Figure A 3.1 Substitution of *T3PPTase* gene into *sfp* gene encoding a PPTase from *B. subtilis*. The replacement of *T3PPTase* gene with *sfp* gene in *E.coli* BL21(DE3) *T3PPTase* sxt1 sxt2 was performed following a two-step recombineering approach using a *tetA-sacB* dual selective/counterselective cassette. In the first recombineering step, the *T3PPTase* gene integrated in the mannitol operon was replaced with *tetA-sacB* cassette. The second recombineering step then removed the cassette while inserting the *sfp* gene and having the T7 promoter drive the gene expression. Successful homologous recombinations were screened based on tetracycline resistance in the first step and fusaric acid and sucrose sensitivity in the second step.

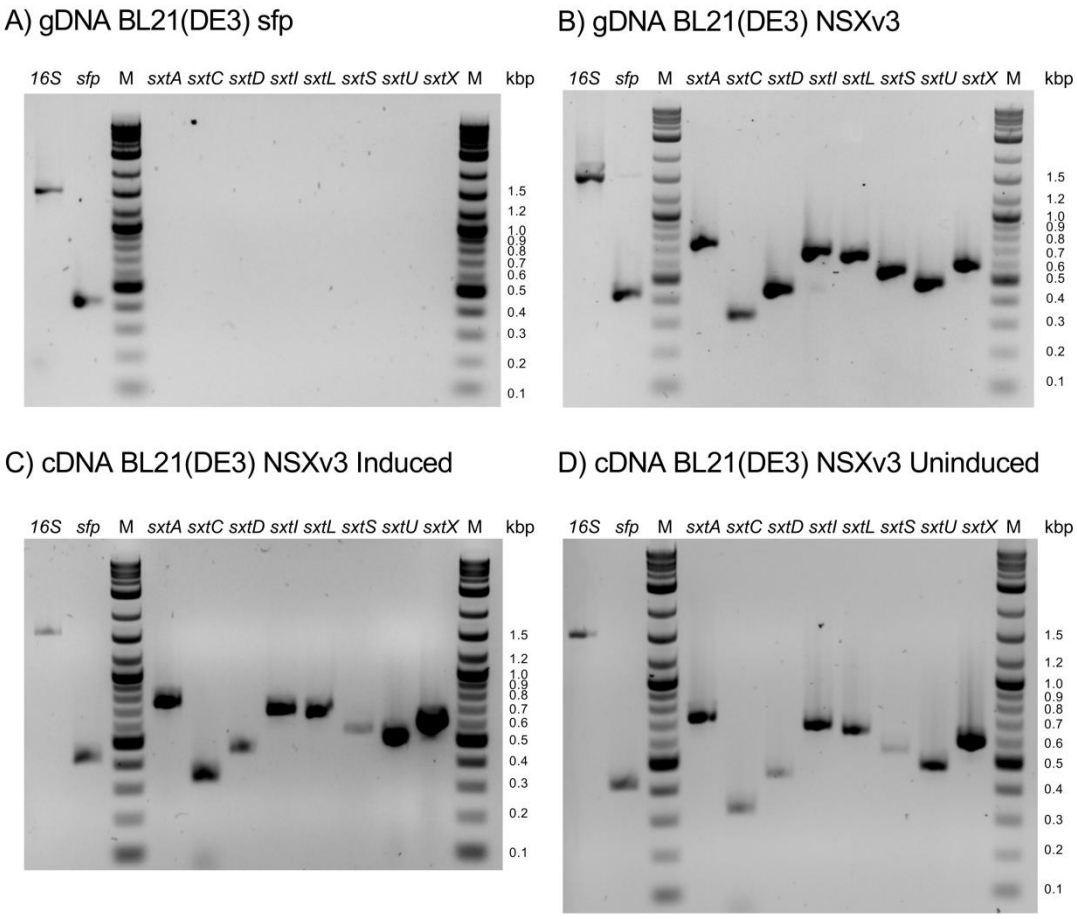


Figure A 3.2 Transcriptional analysis of *sxt* gene expression. PCR amplification of 16S rRNA, *sfp*, and *sxt* genes from gDNA of *E. coli* BL21(DE3) sfp (**A**) and *E. coli* BL21(DE3) *sfp* NSXv3 (**B**), cDNA reverse transcribed from mRNA of *E. coli* BL21(DE3) *sfp* NSXv3 induced (**C**) and *E. coli* BL21(DE3) *sfp* NSXv3 uninduced (**D**). The volumes of PCR reactions loaded for the amplification from gDNA were equal. Whereas for cDNA PCR reactions, the loaded volume was adjusted to ensure clear visualisation of the low-level amplicons.

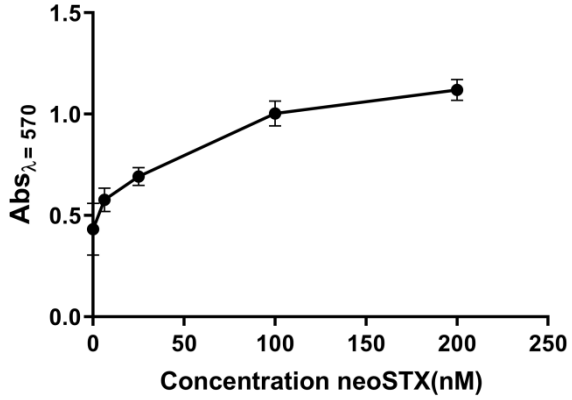


Figure A 3.3 Standard curve of dose-response to neoSTX by neuroblastoma cells (Neuro-2a). Standard curve of neoSTX dose-response by Neuro-2a in the *E. coli* biological matrix (SPE purified BL21(DE3) *sfp* extract spiked with authentic neoSTX standard).

References

Aguilera, A., Gómez, E.B., Kaštovský, J., Echenique, R.O., and Salerno, G.L. (2018) The polyphasic analysis of two native *Raphidiopsis* isolates supports the unification of the genera *Raphidiopsis* and *Cylindrospermopsis* (Nostocales, Cyanobacteria). *Phycologia* **57**: 130-146.

Aimi, N., Odaka, H., Sakai, S.I., Fujiki, H., Suganuma, M., Moore, R.E., and Patterson, G.M.L. (1990) Lyngbyatoxins B and C, two new irritants from *Lyngbya majuscula*. *Journal of Natural Products* **53**: 1593-1596.

Alexander, D.C., Rock, J., Gu, J.-Q., Mascio, C., Chu, M., Brian, P., and Baltz, R.H. (2010) Production of novel lipopeptide antibiotics related to A54145 by *Streptomyces fradiae* mutants blocked in biosynthesis of modified amino acids and assignment of *lptJ*, *lptK* and *lptL* gene functions. *The Journal Of Antibiotics* **64**: 79-87.

Anderson, D.M., and Rengefors, K. (2006) Community assembly and seasonal succession of marine dinoflagellates in a temperate estuary: the importance of life cycle events. *Limnology and Oceanography* **51**: 860-873.

Andrinolo, D., Michea, L.F., and Lagos, N. (1999) Toxic effects, pharmacokinetics and clearance of saxitoxin, a component of paralytic shellfish poison (PSP), in cats. *Toxicon* **37**: 447-464.

Awakawa, T., Zhang, L., Wakimoto, T., Hoshino, S., Mori, T., Ito, T., Ishikawa, J., Tanner, M.E., and Abe, I. (2014) A methyltransferase initiates terpene cyclization in teleocidin B biosynthesis. *Journal of the American Chemical Society* **136**: 9910-9913.

Awan, A.R., Blount, B.A., Bell, D.J., Shaw, W.M., Ho, J.C.H., McKiernan, R.M., and Ellis, T. (2017) Biosynthesis of the antibiotic nonribosomal peptide penicillin in baker's yeast. *Nature Communications* **8**: 15202.

Bakke, I., Berg, L., Aune, T.E., Brautaset, T., Sletta, H., Tondervik, A., and Valla, S. (2009) Random mutagenesis of the *Pm* promoter as a powerful strategy for improvement of recombinant-gene expression. *Applied and Environmental Microbiology* **75**: 2002-2011.

Ballot, A., Cerasino, L., Hostyeva, V., and Cires, S. (2016) Variability in the *sxt* gene clusters of PSP toxin producing *Aphanizomenon gracile* strains from Norway, Spain, Germany and North America. *PLoS ONE* **11**: e0167552.

Baltz, R.H. (2011) Function of MbtH homologs in nonribosomal peptide biosynthesis and applications in secondary metabolite discovery. *Journal of Industrial Microbiology & Biotechnology* **38**: 1747–1760.

Balzer, S., Kucharova, V., Megerle, J., Lale, R., Brautaset, T., and Valla, S. (2013) A comparative analysis of the properties of regulated promoter systems commonly used for recombinant gene expression in *Escherichia coli*. *Microbial Cell Factories* **12**: 26.

Barnes, S., and Hille, B. (1988) Veratridine modifies open sodium channels. *Journal of General Physiology* **91**: 421-443.

Basu, A., Kozikowski, A.P., and Lazo, J.S. (1992) Structural requirements of lyngbyatoxin A for activation and downregulation of protein kinase C. *Biochemistry* **31**: 3824-3830.

Belshaw, P.J., Walsh, C.T., and Stachelhaus, T. (1999) Aminoacyl-CoAs as probes of condensation domain selectivity in nonribosomal peptide synthesis. *Science* **284**: 486-489.

Berg, L., Kucharova, V., Bakke, I., Valla, S., and Brautaset, T. (2012) Exploring the 5'-UTR DNA region as a target for optimizing recombinant gene expression from the strong and inducible *Pm* promoter in *Escherichia coli*. *Journal of Biotechnology.* **158**: 224-230.

Berg, L., Lale, R., Bakke, I., Burroughs, N., and Valla, S. (2009) The expression of recombinant genes in *Escherichia coli* can be strongly stimulated at the transcript production level by mutating the DNA-region corresponding to the 5'-untranslated part of mRNA. *Microbial Biotechnology* **2**: 379-389.

Bertin, M.J., Vulpanovici, A., Monroe, E.A., Korobeynikov, A., Sherman, D.H., Gerwick, L., and Gerwick, W.H. (2016) The phormidolide biosynthetic gene cluster: a trans-AT PKS pathway encoding a toxic macrocyclic polyketide. *ChemBioChem* **17**: 164-173.

Bian, X., Plaza, A., Yan, F., Zhang, Y., and Müller, R. (2015) Rational and efficient site-directed mutagenesis of adenylation domain alters relative yields of luminmide derivatives *in vivo*. *Biotechnology and Bioengineering* **112**: 1343-1353.

Blaustein, M.P., Juhaszova, M., and Golovina, V.A. (1998) The cellular mechanism of action of cardiotonic steroids: a new hypothesis. *Clinical and Experimental Hypertension* **20**: 691-703.

Blokhin, A.V., Yoo, H.D., Geralds, R.S., Nagle, D.G., Gerwick, W.H., and Hamel, E. (1995) Characterization of the interaction of the marine cyanobacterial natural product curacin A with the colchicine site of tubulin and initial structure-activity studies with analogues. *Molecular Pharmacology* **48**: 523-531.

Bloudoff, K., Alonzo, Diego A., and Schmeing, T.M. (2016) Chemical probes allow structural insight into the condensation reaction of nonribosomal peptide synthetases. *Cell Chemical Biology* **23**: 331-339.

Bochner, B.R., Huang, H.C., Schieven, G.L., and Ames, B.N. (1980) Positive selection for loss of tetracycline resistance. *Journal of Bacteriology* **143**: 926-933.

Bokesch, H.R., O'Keefe, B.R., McKee, T.C., Pannell, L.K., Patterson, G.M., Gardella, R.S., Sowder, R.C., 2nd, Turpin, J., Watson, K., Buckheit, R.W., Jr., and Boyd, M.R. (2003) A potent novel anti-HIV protein from the cultured cyanobacterium *Scytonema varium*. *Biochemistry* **42**: 2578-2584.

Boll, B., Taubitz, T., and Heide, L. (2011) Role of MbtH-like proteins in the adenylation of tyrosine during aminocoumarin and vancomycin biosynthesis. *Journal of Biological Chemistry* **286**: 36281-36290.

Boopathi, T., and Ki, J.-S. (2014) Impact of environmental factors on the regulation of cyanotoxin production. *Toxins* **6**: 1951-1978.

Borges, H.L.F., Branco, L.H.Z., Martins, M.D., Lima, C.S., Barbosa, P.T., Lira, G.A.S.T., Bittencourt-Oliveira, M.C., and Molica, R.J.R. (2015) Cyanotoxin production and phylogeny of benthic cyanobacterial strains isolated from the northeast of Brazil. *Harmful Algae* **43**: 46-57.

Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Mori, T., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M., Currens, M.J., Cardellina, J.H., 2nd, Buckheit, R.W., Jr., Nara, P.L., Pannell, L.K., Sowder, R.C., 2nd, and Henderson, L.E. (1997) Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. *Antimicrobial Agents and Chemotherapy* **41**: 1521-1530.

Brown, T.D., Jones-Mortimer, M.C., and Kornberg, H.L. (1977) The enzymic interconversion of acetate and acetyl-coenzyme A in *Escherichia coli*. *Journal of General Microbiology* **102**: 327-336.

Bubb, M.R., Senderowicz, A.M., Sausville, E.A., Duncan, K.L., and Korn, E.D. (1994) Jasplakinolide, a cytotoxic natural product, induces actin polymerization and competitively inhibits the binding of phalloidin to F-actin. *Journal of Biological Chemistry* **269**: 14869-14871.

Burja, A.M., Banaigs, B., Abou-Mansour, E., Burgess, J.G., and Wright, P.C. (2001) Marine cyanobacteria - a prolific source of natural products. *Tetrahedron* **57**: 9347-9377.

Burns, B.P., Goh, F., Allen, M., and Neilan, B.A. (2004) Microbial diversity of extant stromatolites in the hypersaline marine environment of Shark Bay, Australia. *Environmental Microbiology* **6**: 1096-1101.

Butz, D., Schmiederer, T., Hadatsch, B., Wohlleben, W., Weber, T., and Süssmuth, R.D. (2008) Module extension of a non-ribosomal peptide synthetase of the glycopeptide antibiotic balhimycin produced by *Amycolatopsis balhimycina*. *ChemBioChem* **9**: 1195-1200.

Calcott, M.J., and Ackerley, D.F. (2014) Genetic manipulation of non-ribosomal peptide synthetases to generate novel bioactive peptide products. *Biotechnology Letters* **36**: 2407-2416.

Camacho, F.G., Rodriguez, J.G., Miron, A.S., Garcia, M.C., Belarbi, E.H., Chisti, Y., and Grima, E.M. (2007) Biotechnological significance of toxic marine dinoflagellates. *Biotechnology Advances* **25**: 176-194.

Cardellina, J.H., 2nd, Marner, F.J., and Moore, R.E. (1979) Seaweed dermatitis: structure of lyngbyatoxin A. *Science* **204**: 193-195.

Carmichael, W.W., Evans, W.R., Yin, Q.Q., Bell, P., and Moczydlowski, E. (1997) Evidence for paralytic shellfish poisons in the freshwater cyanobacterium *Lyngbya wollei* (Farlow ex Gomont) comb. nov. *Applied and Environmental Microbiology* **63**: 3104-3110.

Carroll, A.R., Coll, J.C., Bourne, D.J., MacLeod, J.K., Zabriskie, T.M., Ireland, C.M., and Bowden, B.F. (1996) Patellins 1-6 and trunkamide A: novel cyclic hexa-, hepta- and octa-peptides from colonial ascidians, *Lissoclinum* sp. *Australian Journal of Chemistry* **49**: 659-667.

Cetusic, J.R., Green, F.R., 3rd, Graupner, P.R., and Oliver, M.P. (2002) Total synthesis of hectochlorin. *Organic Letters* **4**: 1307-1310.

Chai, Y., Shan, S., Weissman, Kira J., Hu, S., Zhang, Y., and Müller, R. (2012) Heterologous expression and genetic engineering of the tubulysin biosynthetic gene cluster using Red/ET recombineering and inactivation mutagenesis. *Chemistry & Biology* **19**: 361-371.

Chang, Z., Flatt, P., Gerwick, W.H., Nguyen, V.-A., Willis, C.L., and Sherman, D.H. (2002) The barbamide biosynthetic gene cluster: a novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit. *Gene* **296**: 235-247.

Chang, Z., Sitachitta, N., Rossi, J.V., Roberts, M.A., Flatt, P.M., Jia, J., Sherman, D.H., and Gerwick, W.H. (2004) Biosynthetic pathway and gene cluster analysis of curacin A, an antitubulin natural product from the tropical marine cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products* **67**: 1356-1367.

Chen, Q.Y., Liu, Y., Cai, W., and Luesch, H. (2014) Improved total synthesis and biological evaluation of potent apratoxin S4 based anticancer agents with differential stability and further enhanced activity. *Journal of Medicinal Chemistry* **57**: 3011-3029.

Chen, Q.Y., Liu, Y., and Luesch, H. (2011) Systematic chemical mutagenesis identifies a potent novel apratoxin A/E hybrid with improved *in vivo* antitumor activity. *Chemistry & Biology* **2**: 861-865.

Chi, H., Wang, X., Shao, Y., Qin, Y., Deng, Z., Wang, L., and Chen, S. (2019) Engineering and modification of microbial chassis for systems and synthetic biology. *Synthetic and Systems Biotechnology* **4**: 25-33.

Chlipala, G.E., Tri, P.H., Hung, N.V., Krunic, A., Shim, S.H., Soejarto, D.D., and Orjala, J. (2010) Nhatrangins A and B, aplysiatoxin-related metabolites from the marine cyanobacterium *Lyngbya majuscula* from Vietnam. *Journal of Natural Products* **73**: 784-787.

Christaki, E. (2014) Mircoalgae as a potential new generation of material for various innovative products. *Journal of Oceanography and Marine Research* **2**: e106.

Chun, S.W., Hinze, M.E., Skiba, M.A., and Narayan, A.R.H. (2018) Chemistry of a unique polyketide-like synthase. *Journal of the American Chemical Society* **140**: 2430-2433.

Cleto, S., and Lu, T.K. (2017) An engineered synthetic pathway for discovering nonnatural nonribosomal peptides in *Escherichia coli*. *mBio* **8**: e01474-01417.

Cook, T.B., and Pfleger, B.F. (2019) Leveraging synthetic biology for producing bioactive polyketides and non-ribosomal peptides in bacterial heterologous hosts. *MedChemComm* **10**: 668-681.

Copp, J.N., and Neilan, B.A. (2006) The phosphopantetheinyl transferase superfamily: phylogenetic analysis and functional implications in cyanobacteria. *Applied and Environmental Microbiology* **72**: 2298-2305.

Crüsemann, M., Kohlhaas, C., and Piel, J. (2013) Evolution-guided engineering of nonribosomal peptide synthetase adenylation domains. *Chemical Science* **4**: 1041-1045.

Cullen, A., D'Agostino, P.M., Mazmouz, R., Pickford, R., Wood, S., and Neilan, B.A. (2018a) Insertions within the saxitoxin biosynthetic gene cluster result in differential toxin profiles. *ACS Chemical Biology* **13**: 3107-3114.

Cullen, A., Pearson, L.A., Mazmouz, R., Liu, T., Soeriyadi, A.H., Ongley, S.E., and Neilan, B.A. (2018b) Heterologous expression and biochemical characterisation of cyanotoxin biosynthesis pathways. *Natural Product Reports* **36**: 1117-1136.

Cusick, K.D., and Sayler, G.S. (2013) An overview on the marine neurotoxin, saxitoxin: genetics, molecular targets, methods of detection and ecological functions. *Marine Drugs* **11**: 991-1018.

D'Agostino, P.M., Boundy, M.J., Harwood, T.D., Carmichael, W.W., Neilan, B.A., and Wood, S.A. (2019) Re-evaluation of paralytic shellfish toxin profiles in cyanobacteria using hydrophilic interaction liquid chromatography-tandem mass spectrometry. *Toxicon* **158**: 1-7.

D'Agostino, P.M., and Gulder, T.A.M. (2018) Direct pathway cloning combined with sequence- and ligation-independent cloning for fast biosynthetic gene cluster refactoring and heterologous expression. *ACS Synthetic Biology* **7**: 1702-1708.

Datta, S., Costantino, N., and Court, D.L. (2006) A set of recombineering plasmids for gram-negative bacteria. *Gene* **379**: 109-115.

Davidsen, J.M., Bartley, D.M., and Townsend, C.A. (2013) Non-ribosomal propeptide precursor in nocardicin A biosynthesis predicted from adenylation domain specificity dependent on the MbtH family protein NocI. *Journal of the American Chemical Society* **135**: 1749-1759.

Degen, A., Mayerthaler, F., Mootz, H.D., and Di Ventura, B. (2019) Context-dependent activity of A domains in the tyrocidine synthetase. *Scientific Reports* **9**: 5119.

Dey, S., Wengryniuk, S.E., Tarsis, E.M., Robertson, B.D., Zhou, G.Q., and Coltart, D.M. (2015) A formal asymmetric synthesis of apratoxin D via advanced-stage asymmetric ACC α,α-bisalkylation of a chiral nonracemic ketone. *Tetrahedron Letters* **56**: 2927-2929.

Doekel, S., Coeffet-Le Gal, M.F., Gu, J.Q., Chu, M., Baltz, R.H., and Brian, P. (2008) Non-ribosomal peptide synthetase module fusions to produce derivatives of daptomycin in *Streptomyces roseosporus*. *Microbiology* **154**: 2872-2880.

Donadio, S., and Katz, L. (1992) Organization of the enzymatic domains in the multifunctional polyketide synthase involved in erythromycin formation in *Saccharopolyspora erythraea*. *Gene* **111**: 51-60.

Donia, M.S., Hathaway, B.J., Sudek, S., Haygood, M.G., Rosovitz, M.J., Ravel, J., and Schmidt, E.W. (2006) Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. *Nature Chemical Biology* **2**: 729-735.

Donia, M.S., Ravel, J., and Schmidt, E.W. (2008) A global assembly line for cyanobactins. *Nature Chemical Biology* **4**: 341-343.

Dorrestein, P.C., Blackhall, J., Straight, P.D., Fischbach, M.A., Garneau-Tsodikova, S., Edwards, D.J., McLaughlin, S., Lin, M., Gerwick, W.H., Kolter, R., Walsh, C.T., and Kelleher, N.L. (2006) Activity screening of carrier domains within nonribosomal peptide synthetases using complex substrate mixtures and large molecule mass spectrometry. *Biochemistry* **45**: 1537-1546.

Drake, E.J., Miller, B.R., Shi, C., Tarrasch, J.T., Sundlov, J.A., Allen, C.L., Skiniotis, G., Aldrich, C.C., and Gulick, A.M. (2016) Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**: 235-238.

Du, L., Sanchez, C., and Shen, B. (2001) Hybrid peptide-polyketide natural products: biosynthesis and prospects toward engineering novel molecules. *Metabolic Engineering* **3**: 78-95.

Dubois, M., Demoulin, L., Charlier, C., Singh, G., Godefroy, S.B., Campbell, K., Elliott, C.T., and Delahaut, P. (2010) Development of ELISAs for detecting domoic acid, okadaic acid, and saxitoxin and their applicability for the detection of marine toxins in samples collected in Belgium. *Food Additives & Contaminants. Part A, Chemistry, Analysis, Control, Exposure & Risk Assessment* **27**: 859-868.

Edwards, D.J., and Gerwick, W.H. (2004) Lyngbyatoxin biosynthesis: Sequence of biosynthetic gene cluster and identification of a novel aromatic prenyltransferase. *Journal of the American Chemical Society* **126**: 11432-11433.

Edwards, D.J., Marquez, B.L., Nogle, L.M., McPhail, K., Goeger, D.E., Roberts, M.A., and Gerwick, W.H. (2004) Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. *Chemistry & Biology* **11**: 817-833.

Eliasson, L., Kallin, B., Patarroyo, M., Klein, G., Fujiki, H., and Sugimura, T. (1983) Activation of the EBV-cycle and aggregation of human blood lymphocytes by the tumor promoters teleocidin, lyngbyatoxin a, aplysiatoxin and debromoaplysiatoxin. *International Journal of Cancer* **31**: 7-11.

Engene, N., Rottacker, E.C., Kaštovský, J., Byrum, T., Choi, H., Ellisman, M.H., Komárek, J., and Gerwick, W.H. (2012) *Moorea producens* gen. nov., sp. nov. and *Moorea bouillonii* comb. nov., tropical marine cyanobacteria rich in bioactive secondary metabolites. *International Journal of Systematic and Evolutionary Microbiology* **62**: 1171-1178.

Eppelmann, K., Stachelhaus, T., and Marahiel, M.A. (2002) Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics. *Biochemistry* **41**: 9718-9726.

Epstein-Barash, H., Shichor, I., Kwon, A.H., Hall, S., Lawlor, M.W., Langer, R., and Kohane, D.S. (2009) Prolonged duration local anesthesia with minimal toxicity. *Proceedings of the National Academy of Sciences USA* **106**: 7125-7130.

Esquenazi, E., Jones, A.C., Byrum, T., Dorrestein, P.C., and Gerwick, W.H. (2011) Temporal dynamics of natural product biosynthesis in marine cyanobacteria. *Proceedings of the National Academy of Sciences USA* **108**: 5226-5231.

Etheridge, S.M. (2010) Paralytic shellfish poisoning: seafood safety and human health perspectives. *Toxicon* **56**: 108-122.

Evans, B.S., Chen, Y., Metcalf, W.W., Zhao, H., and Kelleher, N.L. (2011) Directed evolution of the nonribosomal peptide synthetase AdmK generates new andrimid derivatives *in vivo*. *Chemistry & Biology* **18**: 601-607.

Felnagle, E.A., Barkei, J.J., Park, H., Podevels, A.M., McMahon, M.D., Drott, D.W., and Thomas, M.G. (2010) MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. *Biochemistry* **49**: 8815-8817.

Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J.L., Vázquez, E., and Villaverde, A. (2009) Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories* **8**: 17.

Fijal, K., and Filip, M. (2016) Clinical/therapeutic approaches for cannabinoid ligands in central and peripheral nervous system diseases: mini review. *Clinical Neuropharmacology* **39**: 94-101.

Fine Nathel, N.F., Shah, T.K., Bronner, S.M., and Garg, N.K. (2014) Total syntheses of indolactam alkaloids (−)-indolactam V, (−)-pendolmycin, (−)-lyngbyatoxin A, and (−)-teleocidin A-2. *Chemical Science* **5**: 2184-2190.

Flatt, P.M., Gautschi, J.T., Thacker, R.W., Musafija-Girt, M., Crews, P., and Gerwick, W.H. (2005) Identification of the cellular site of polychlorinated peptide biosynthesis in the marine sponge *Dysidea* (Lamellodysidea) herbacea and symbiotic cyanobacterium *Oscillatoria spongeliae* by CARD-FISH analysis. *Marine Biology* **147**: 761-774.

Flatt, P.M., O'Connell, S.J., McPhail, K.L., Zeller, G., Willis, C.L., Sherman, D.H., and Gerwick, W.H. (2006) Characterization of the initial enzymatic steps of barbamide biosynthesis. *Journal of Natural Products* **69**: 938-944.

Foss, A.J., Phlips, E.J., Yilmaz, M., and Chapman, A. (2012) Characterization of paralytic shellfish toxins from *Lyngbya wollei* dominated mats collected from two Florida springs. *Harmful Algae* **16**: 98-107.

Fu, J., Bian, X., Hu, S., Wang, H., Huang, F., Seibert, P.M., Plaza, A., Xia, L., Muller, R., Stewart, A.F., and Zhang, Y. (2012) Full-length RecE enhances linear-linear homologous recombination and facilitates direct cloning for bioprospecting. *Nature Biotechnology* **30**: 440-446.

Fu, J., Teucher, M., Anastassiadis, K., Skarnes, W., and Stewart, A.F. (2010) A recombineering pipeline to make conditional targeting constructs. *Methods in Enzymology* **477**: 125-144.

Fu, J., Wenzel, S.C., Perlova, O., Wang, J., Gross, F., Tang, Z., Yin, Y., Stewart, A.F., Müller, R., and Zhang, Y. (2008) Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Research* **36**: e113.

Fujiki, H., Mori, M., Nakayasu, M., Terada, M., and Sugimura, T. (1979) A possible naturally occurring tumor promoter, teleocidin B from *Streptomyces*. *Biochemical and Biophysical Research Communications* **90**: 976-983.

Fujiki, H., Mori, M., Nakayasu, M., Terada, M., Sugimura, T., and Moore, R.E. (1981) Indole alkaloids: dihydroteleocidin B, teleocidin, and lyngbyatoxin A as members of a new class of tumor promoters. *Proceedings of the National Academy of Sciences USA* **78**: 3872-3876.

Gabrielsen, T.M., Minge, M.A., Espelund, M., Tooming-Klunderud, A., Patil, V., Nederbragt, A.J., Otis, C., Turmel, M., Shalchian-Tabrizi, K., Lemieux, C., and Jakobsen, K.S. (2011) Genome evolution of a tertiary dinoflagellate plastid. *PLoS ONE* **6**: e19132.

Gallegos, M.T., Marques, S., and Ramos, J.L. (1996) Expression of the TOL plasmid *xylS* gene in *Pseudomonas putida* occurs from a alpha 70-dependent promoter or from alpha 70- and alpha 54-dependent tandem promoters according to the compound used for growth. *Journal of Bacteriology* **178**: 2356-2361.

Galonic, D.P., Vaillancourt, F.H., and Walsh, C.T. (2006) Halogenation of unactivated carbon centers in natural product biosynthesis: trichlorination of leucine during barbamide biosynthesis. *Journal of the American Chemical Society* **128**: 3900-3901.

Gao, L., Guo, J., Fan, Y., Ma, Z., Lu, Z., Zhang, C., Zhao, H., and Bie, X. (2018) Module and individual domain deletions of NRPS to produce plipastatin derivatives in *Bacillus subtilis*. *Microbial Cell Factories* **17**: 84.

Garrido, R., Lagos, N., Lagos, M., Rodriguez-Navarro, A.J., Garcia, C., Truan, D., and Henriquez, A. (2007) Treatment of chronic anal fissure by gonyautoxin. *Colorectal Disease* **9**: 619-624.

Garrido, R., Lagos, N., Lattes, K., Abedrapo, M., Bocic, G., Cuneo, A., Chiong, H., Jensen, C., Azolas, R., Henriquez, A., and Garcia, C. (2005) Gonyautoxin: new treatment for healing acute and chronic anal fissures. *Diseases of the Colon and Rectum* **48**: 335-343.

Garrido, R., Lagos, N., Lattes, K., Azolas, C.G., Bocic, G., Cuneo, A., Chiong, H., Jensen, C., Henriquez, A.I., and Fernandez, C. (2004) The gonyautoxin 2/3 epimers reduces anal tone when injected in the anal sphincter of healthy adults. *Biological Research* **37**: 395-403.

Gaudelli, N.M., Long, D.H., and Townsend, C.A. (2015) β-Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis. *Nature* **520**: 383-387.

Gawin, A., Valla, S., and Brautaset, T. (2017) The XylS/*Pm* regulator/promoter system and its use in fundamental studies of bacterial gene expression, recombinant protein production and metabolic engineering. *Microbial Biotechnology* **10**: 702-718.

Gay, P., Le Coq, D., Steinmetz, M., Ferrari, E., and Hoch, J.A. (1983) Cloning structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. *Journal of Bacteriology* **153**: 1424-1431.

Gerwick, W.H., Proteau, P.J., Nagle, D.G., Hamel, E., Blokhin, A., and Slate, D.L. (1994) Structure of curacin A, a novel antimitotic, antiproliferative and brine shrimp toxic natural product from the marine cyanobacterium *Lyngbya majuscula*. *Journal of Organic Chemistry* **59**: 1243-1245.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison Iii, C.A., and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**: 343-345.

Gil, A., Lamariano-Merketegi, J., Lorente, A., Albericio, F., and Alvarez, M. (2016) Synthesis of (E)-4-bromo-3-methoxybut-3-en-2-one, the key fragment in the polyhydroxylated chain common to oscillariolide and phormidolides A-C. *Chemistry* **22**: 7033-7035.

Gilles, A., Martinez, J., and Cavelier, F. (2011) A flexible synthesis of C33-C39 polyketide region of apratoxin: synthesis of natural and unnatural analogues. *Comptes Rendus Chimie* **14**: 437-440.

González, L., Santiago, M.S., and Santiago, P.S.A., (2015) Method for the industrial purification of biologically active phycotoxins. International patent no. WO 2010/109386 In. European patent office.

Goto, Y., Ito, Y., Kato, Y., Tsunoda, S., and Suga, H. (2014) One-pot synthesis of azoline-containing peptides in a cell-free translation system integrated with a posttranslational cyclodehydratase. *Chemistry & Biology* **21**: 766-774.

Graf, K.M., Tabor, M.G., Brown, M.L., and Paige, M. (2009) Synthesis of (S)-jamaicamide C carboxylic acid. *Organic Letters* **11**: 5382-5385.

Grindberg, R.V., Ishoey, T., Brinza, D., Esquenazi, E., Coates, R., Liu, W.t., Gerwick, L., Dorrestein, P.C., Pevzner, P., Lasken, R., and Gerwick, W.H. (2011) Single cell genome amplification accelerates identification of the apratoxin biosynthetic pathway from a complex microbial assemblage. *PLoS ONE* **6**: e18565.

Gu, L.C., Jia, J.Y., Liu, H.C., Hakansson, K., Gerwick, W.H., and Sherman, D.H. (2006) Metabolic coupling of dehydration and decarboxylation in the curacin A pathway: functional identification of a mechanistically diverse enzyme pair. *Journal of the American Chemical Society* **128**: 9014-9015.

Gulick, A.M. (2016) Structural insight into the necessary conformational changes of modular nonribosomal peptide synthetases. *Current Opinion in Chemical Biology* **35**: 89-96.

Gunasekera, S.P., Imperial, L., Garst, C., Ratnayake, R., Dang, L.H., Paul, V.J., and Luesch, H. (2016) Caldoramide, a modified pentapeptide from the marine cyanobacterium *Caldora penicillata*. *Journal of Natural Products* **79**: 1867-1871.

Gupta, D.K., Kaur, P., Leong, S.T., Tan, L.T., Prinsep, M.R., and Chu, J.J.H. (2014) Anti-chikungunya viral activities of aplysiatoxin-related compounds from the marine cyanobacterium *Trichodesmium erythraeum*. *Marine Drugs* **12**: 115-127.

Gutierrez, M., Suyama, T.L., Engene, N., Wingerd, J.S., Matainaho, T., and Gerwick, W.H. (2008) Apratoxin D, a potent cytotoxic cyclodepsipeptide from Papua New Guinea collections of the marine cyanobacteria *Lyngbya majuscul*a and *Lyngbya sordida*. *Journal of Natural Products* **71**: 1099-1103.

Han, J.W., Kim, E.Y., Lee, J.M., Kim, Y.S., Bang, E., and Kim, B.S. (2012) Site-directed modification of the adenylation domain of the fusaricidin nonribosomal peptide synthetase for enhanced production of fusaricidin analogs. *Biotechnology Letters* **34**: 1327-1334.

Harrigan, G.G., Goetz, G.H., Luesch, H., Yang, S.T., and Likos, J. (2001) Dysideaprolines A-F and barbaleucamides A-B, novel polychlorinated compounds from a *Dysidea* species. *Journal of Natural Products* **64**: 1133-1138.

Hayashi, R., Saito, H., Okumura, M., and Kondo, F. (2006) Cell bioassay for paralytic shellfish poisoning (PSP): comparison with postcolumn derivatization liquid chromatographic analysis and application to the monitoring of PSP in shellfish. *Journal of Agricultural and Food Chemistry* **54**: 269-273.

Heimann, K., and Huerlimann, R., (2015) Chapter 3 - Microalgal classification: major classes and genera of commercial microalgal species In: Handbook of Marine Microalgae. Boston: Academic Press, pp. 25-41.

Henderson, R., Ritchie, J.M., and Strichartz, G.R. (1973) The binding of labelled saxitoxin to the sodium channels in nerve membranes. *The Journal of Physiology* **235**: 783-804.

Herbst, D.A., Boll, B., Zocher, G., Stehle, T., and Heide, L. (2013) Structural basis of the interaction of MbtH-like proteins, putative regulators of nonribosomal peptide biosynthesis, with adenylating enzymes. *Journal of Biological Chemistry* **288**: 1991-2003.

Hernandez-Becerril, D.U., Alonso-Rodriguez, R., Alvarez-Gongora, C., Baron-Campis, S.A., Ceballos-Corona, G., Herrera-Silveira, J., Meave Del Castillo, M.E., Juarez-Ruiz, N., Merino-Virgilio, F., Morales-Blake, A., Ochoa, J.L., Orellana-Cepeda, E., Ramirez-Camarena, C., and Rodriguez-Salvador, R. (2007) Toxic and harmful marine

phytoplankton and microalgae (HABs) in Mexican Coasts. *Journal of Environmental Science and Health. Part A, Toxic/Hazardous Substances & Environmental Engineering* **42**: 1349-1363.

Hooper, G.J., Orjala, J., Schatzman, R.C., and Gerwick, W.H. (1998) Carmabins A and B, new lipopeptides from the Caribbean cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products* **61**: 529-533.

Houssen, W.E., Bent, A.F., McEwan, A.R., Pieiller, N., Tabudravu, J., Koehnke, J., Mann, G., Adaba, R.I., Thomas, L., Hawas, U.W., Liu, H., Schwarz-Linek, U., Smith, M.C., Naismith, J.H., and Jaspars, M. (2014) An efficient method for the *in vitro* production of azol(in)e-based cyclic peptides. *Angewandte Chemie International Edition* **53**: 14171-14174.

Houssen, W.E., Koehnke, J., Zollman, D., Vendome, J., Raab, A., Smith, M.C.M., Naismith, J.H., 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.

Huang, K.C., Chen, Z., Jiang, Y., Akare, S., Kolber-Simonds, D., Condon, K., Agoulnik, S., Tendyke, K., Shen, Y., Wu, K.M., Mathieu, S., Choi, H.W., Zhu, X., Shimizu, H., Kotake, Y., Gerwick, W.H., Uenaka, T., Woodall-Jappe, M., and Nomoto, K. (2016) Apratoxin A shows novel pancreas-targeting activity through the binding of Sec 61. *Molecular Cancer Therapeutics* **15**: 1208-1216.

Huang, X., Hsu, K.-H., and Chu, F.S. (1996) Direct competitive enzyme-linked immunosorbent assay for saxitoxin and neosaxitoxin. *Journal of Agricultural and Food Chemistry* **44**: 1029-1035.

Humpage, A.R., Ledreux, A., Fanok, S., Bernard, C., Briand, J.F., Eaglesham, G., Papageorgiou, J., Nicholson, B., and Steffensen, D. (2007) Application of the neuroblastoma assay for paralytic shellfish poisons to neurotoxic freshwater cyanobacteria: interlaboratory calibration and comparison with other methods of analysis. *Environmental Toxicology and Chemistry* **26**: 1512-1519.

Huynh, M.U., Elston, M.C., Hernandez, N.M., Ball, D.B., Kajiyama, S.-i., Irie, K., Gerwick, W.H., and Edwards, D.J. (2010) Enzymatic production of (-)-indolactam V by LtxB, a cytochrome P450 monooxygenase. *Journal of Natural Products* **73**: 71-74.

Imada, C., (2013) Treasure hunting for useful microorganisms in the marine environment. In: Marine Microbiology. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 21-31.

Ireland, C.M., Durso, A.R., Newman, R.A., and Hacker, M.P. (1982) Anti-neoplastic cyclic-peptides from the marine runicate *Lissoclinum patella*. *Journal of Organic Chemistry* **47**: 1807-1811.

Irie, K., Kikumori, M., Kamachi, H., Tanaka, K., Murakami, A., Yanagita, R.C., Tokuda, H., Suzuki, N., Nagai, H., Suenaga, K., and Nakagawa, Y. (2012) Synthesis and structure-activity studies of simplified analogues of aplysiatoxin with antiproliferative activity like bryostatin-1. *Pure and Applied Chemistry* **84**: 1341-1351.

Irie, K., and Yanagita, R.C. (2014) Synthesis and biological activities of simplified analogs of the natural PKC ligands, bryostatin-1 and aplysiatoxin. *Chemical Record* **14**: 251-267.

Jansson, D., and Astot, C. (2015) Analysis of paralytic shellfish toxins, potential chemical threat agents, in food using hydrophilic interaction liquid chromatography-mass spectrometry. *Journal of Chromatography. A* **1417**: 41-48.

Jebbar, M., Franzetti, B., Girard, E., and Oger, P. (2015) Microbial diversity and adaptation to high hydrostatic pressure in deep-sea hydrothermal vents prokaryotes. *Extremophiles* **19**: 721-740.

Jellett, J.F., Marks, L.J., Stewart, J.E., Dorey, M.L., Watson-Wright, W., and Lawrence, J.F. (1992) Paralytic shellfish poison (saxitoxin family) bioassays: automated endpoint determination and standardization of the *in vitro* tissue culture bioassay, and comparison with the standard mouse bioassay. *Toxicon* **30**: 1143-1156.

Jiang, H., Wang, Y.Y., Ran, X.X., Fan, W.M., Jiang, X.H., Guan, W.J., and Li, Y.Q. (2013) Improvement of natamycin production by engineering of phosphopantetheinyl

transferases in *Streptomyces chattanoogensis* L10. *Applied and Environmental Microbiology* **79**: 3346-3354.

Jiang, W., Tan, S., Hanaki, Y., Irie, K., Uchida, H., Watanabe, R., Suzuki, T., Sakamoto, B., Kamio, M., and Nagai, H. (2014a) Two new lyngbyatoxin derivatives from the cyanobacterium, *Moorea producens*. *Marine Drugs* **12**: 5788–5800.

Jiang, W., Zhou, W., Uchida, H., Kikumori, M., Irie, K., Watanabe, R., Suzuki, T., Sakamoto, B., Kamio, M., and Nagai, H. (2014b) A new lyngbyatoxin from the Hawaiian cyanobacterium *Moorea producens*. *Marine Drugs* **12**: 2748-2759.

Jimenez, J.I., and Scheuer, P.J. (2001) New lipopeptides from the caribbean cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products* **64**: 200-203.

Jones, A.C., Gerwick, L., Gonzalez, D., Dorrestein, P.C., and Gerwick, W.H. (2009) Transcriptional analysis of the jamaicamide gene cluster from the marine cyanobacterium *Lyngbya majuscula* and identification of possible regulatory proteins. *BMC Microbiology* **9**: 247-247.

Jones, A.C., Monroe, E.A., Podell, S., Hess, W.R., Klages, S., Esquenazi, E., Niessen, S., Hoover, H., Rothmann, M., Lasken, R.S., Yates, J.R., 3rd, Reinhardt, R., Kube, M., Burkart, M.D., Allen, E.E., Dorrestein, P.C., Gerwick, W.H., and Gerwick, L. (2011) Genomic insights into the physiology and ecology of the marine filamentous cyanobacterium *Lyngbya majuscula*. *Proceedings of the National Academy of Sciences USA* **108**: 8815-8820.

Jones, A.C., Ottilie, S., Eustaquio, A.S., Edwards, D.J., Gerwick, L., Moore, B.S., and Gerwick, W.H. (2012) Evaluation of *Streptomyces coelicolor* A3(2) as a heterologous expression host for the cyanobacterial protein kinase C activator lyngbyatoxin A. *FEBS Journal* **279**: 1243-1251.

Kaebernick, M., Neilan, B.A., Borner, T., and Dittmann, E. (2000) Light and the transcriptional response of the microcystin biosynthesis gene cluster. *Applied and Environmental Microbiology* **66**: 3387-3392.

Kaljunen, H., Schiefelbein, S.H., Stummer, D., Kozak, S., Meijers, R., Christiansen, G., and Rentmeister, A. (2015) Structural elucidation of the bispecificity of A domains as a basis for activating non-natural amino acids. *Angewandte Chemie International Edition* **54**: 8833-8836.

Kamachi, H., Tanaka, K., Yanagita, R.C., Murakami, A., Murakami, K., Tokuda, H., Suzuki, N., Nakagawa, Y., and Irie, K. (2013) Structure-activity studies on the side chain of a simplified analog of aplysiatoxin (aplog-1) with anti-proliferative activity. *Bioorganic & Medicinal Chemistry* **21**: 2695-2702.

Kang, H.S., Charlop-Powers, Z., and Brady, S.F. (2016) Multiplexed CRISPR/Cas9- and TAR-mediated promoter engineering of natural product biosynthetic gene clusters in yeast. *ACS Synthetic Biology* **5**: 1002-1010.

Kellmann, R., Mihali, T.K., Young, J.J., Pickford, R., Pomati, F., and Neilan, B.A. (2008) Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. *Applied and Environmental Microbiology* **74**: 4044-4053.

Kellmann, R., Stüken, A., Orr, R.J.S., Svendsen, H.M., and Jakobsen, K.S. (2010) Biosynthesis and molecular genetics of polyketides in marine dinoflagellates. *Marine Drugs* **8**: 1011-1048.

Kim, E.J., Lee, J.H., Choi, H., Pereira, A.R., Ban, Y.H., Yoo, Y.J., Kim, E., Park, J.W., Sherman, D.H., Gerwick, W.H., and Yoon, Y.J. (2012) Heterologous production of 4-O-demethylbarbamide, a marine cyanobacterial natural product. *Organic Letters* **14**: 5824-5827.

Kleigrewe, K., Almaliti, J., Tian, I.Y., Kinnel, R.B., Korobeynikov, A., Monroe, E.A., Duggan, B.M., Di Marzo, V., Sherman, D.H., Dorrestein, P.C., Gerwick, L., and Gerwick, W.H. (2015) Combining mass spectrometric metabolic profiling with genomic analysis: a powerful approach for discovering natural products from cyanobacteria. *Journal of Natural Products* **78**: 1671-1682.

Knoot, C.J., Khatri, Y., Hohlman, R.M., Sherman, D.H., and Pakrasi, H.B. (2019) Engineered production of hapalindole alkaloids in the cyanobacterium *Synechococcus* sp. UTEX 2973. *ACS Synthetic Biology* **8**: 1941-1951.

Kobayashi, J. (2008) Amphidinolides and its related macrolides from marine dinoflagellates. *The Journal of Antibiotics* **61**: 271-284.

Koehnke, J., Bent, A., Houssen, W.E., Zollman, D., Morawitz, F., Shirran, S., Vendome, J., Nneoyiegbe, A.F., Trembleau, L., Botting, C.H., Smith, M.C.M., Jaspars, M., and Naismith, J.H. (2012) The mechanism of patellamide macrocyclization revealed by the characterization of the PatG macrocyclase domain. *Nature Structural & Molecular Biology* **19**: 767-772.

Koehnke, J., Bent, A.F., Houssen, W.E., Mann, G., Jaspars, M., and Naismith, J.H. (2014) The structural biology of patellamide biosynthesis. *Current Opinion in Structural Biology* **29**: 112-121.

Koehnke, J., Bent, A.F., Zollman, D., Smith, K., Houssen, W.E., Zhu, X.F., Mann, G., Lebl, T., Scharff, R., Shirran, S., Botting, C.H., Jaspars, M., Schwarz-Linek, U., and Naismith, J.H. (2013) The cyanobactin heterocyclase enzyme: A processive adenylase that operates with a defined order of reaction. *Angewandte Chemie International Edition* **52**: 13991-13996.

Kohane, D.S., Lu, N.T., Gökgöl-Kline, A.C., Shubina, M., Kuang, Y., Hall, S., Strichartz, G.R., and Berde, C.B. (2000) The local anesthetic properties and toxicity of saxitonin homologues for rat sciatic nerve block *in vivo*. *Regional Anesthesia and Pain Medicine* **25**: 52-59.

Konz, D., and Marahiel, M.A. (1999) How do peptide synthetases generate structural diversity? *Chemistry & Biology* **6**: R39-R48.

Kozikowski, A.P., Sato, K., Basu, A., and Lazo, J.S. (1989) Synthesis and biological studies of simplified analogs of lyngbyatoxin A; use of an isoxazoline-based indole synthesis. Quest for protein kinase C modulators. *Journal of the American Chemical Society* **111**: 6228-6234.

Kreutzer, M.F., Kage, H., Herrmann, J., Pauly, J., Hermenau, R., Müller, R., Hoffmeister, D., and Nett, M. (2014) Precursor-directed biosynthesis of micacocidin derivatives with activity against *Mycoplasma pneumoniae*. *Organic & Biomolecular Chemistry* **12**: 113-118.

Kries, H., Wachtel, R., Pabst, A., Wanner, B., Niquille, D., and Hilvert, D. (2014) Reprogramming nonribosomal peptide synthetases for "clickable" amino acids. *Angewandte Chemie International Edition* **53**: 10105-10108.

Kumari, S., Beatty, C.M., Browning, D.F., Busby, S.J., Simel, E.J., Hovel-Miner, G., and Wolfe, A.J. (2000) Regulation of acetyl coenzyme A synthetase in *Escherichia coli*. *Journal of Bacteriology* **182**: 4173-4179.

Lagos, N., Onodera, H., Zagatto, P.A., Andrinolo, D., Azevedo, S.M., and Oshima, Y. (1999) The first evidence of paralytic shellfish toxins in the fresh water cyanobacterium *Cylindrospermopsis raciborskii*, isolated from Brazil. *Toxicon* **37**: 1359-1373.

Lale, R., Berg, L., Stuttgen, F., Netzer, R., Stafsnes, M., Brautaset, T., Vee Aune, T.E., and Valla, S. (2011) Continuous control of the flow in biochemical pathways through 5' untranslated region sequence modifications in mRNA expressed from the broad-host-range promoter *Pm*. *Applied and Environmental Microbiology* **77**: 2648-2655.

Lambalot, R.H., Gehring, A.M., Flugel, R.S., Zuber, P., LaCelle, M., Marahiel, M.A., Reid, R., Khosla, C., and Walsh, C.T. (1996) A new enzyme superfamily - the phosphopantetheinyl transferases. *Chemistry & Biology* **3**: 923-936.

Landsberg, J.H. (2002) The effects of harmful algal blooms on aquatic organisms. *Reviews in Fisheries Science* **10**: 113-390.

Lattes, K., Venegas, P., Lagos, N., Lagos, M., Pedraza, L., Rodriguez-Navarro, A.J., and Garcia, C. (2009) Local infiltration of gonyautoxin is safe and effective in treatment of chronic tension-type headache. *Neurological Research* **31**: 228-233.

Lautru, S., Oves-Costales, D., Pernodet, J.L., and Challis, G.L. (2007) MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore biosynthetic pathways in *Streptomyces coelicolor* M145. *Microbiology* **153**: 1405-1412.

Lawrence, J.F., Niedzwiadek, B., and Menard, C. (2005) Quantitative determination of paralytic shellfish poisoning toxins in shellfish using prechromatographic oxidation and liquid chromatography with fluorescence detection: collaborative study. *Journal of AOAC International* **88**: 1714-1732.

Laycock, M.V., Thibault, P., Ayer, S.W., and Walter, J.A. (1994) Isolation and purification procedures for the preparation of paralytic shellfish poisoning toxin standards. *Natural Toxins* **2**: 175-183.

Leao, P.N., Nakamura, H., Costa, M., Pereira, A.R., Martins, R., Vasconcelos, V., Gerwick, W.H., and Balskus, E.P. (2015) Biosynthesis-assisted structural elucidation of the bartolosides, chlorinated aromatic glycolipids from cyanobacteria. *Angewandte Chemie International Edition* **54**: 11063-11067.

Lee, J., McIntosh, J., Hathaway, B.J., and Schmidt, E.W. (2009) Using marine natural products to discover a protease that catalyzes peptide macrocyclization of diverse substrates. *Journal of the American Chemical Society* **131**: 2122-2124.

Lefebvre, K.A., Bill, B.D., Erickson, A., Baugh, K.A., O'Rourke, L., Costa, P.R., Nance, S., and Trainer, V.L. (2008) Characterization of intracellular and extracellular saxitoxin levels in both field and cultured *Alexandrium* spp. samples from Sequim Bay, Washington. *Marine Drugs* **6**: 103-116.

Li, B., Sher, D., Kelly, L., Shi, Y., Huang, K., Knerr, P.J., Joewono, I., Rusch, D., Chisholm, S.W., and van der Donk, W.A. (2010) Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine cyanobacteria. *Proceedings of the National Academy of Sciences USA* **107**: 10430-10435.

Li, X.T., Thomason, L.C., Sawitzke, J.A., Costantino, N., and Court, D.L. (2013) Positive and negative selection using the *tetA-sacB* cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Research* **41**: e204.

Linington, R.G., Clark, B.R., Trimble, E.E., Almanza, A., Urena, L.D., Kyle, D.E., and Gerwick, W.H. (2009) Antimalarial peptides from marine cyanobacteria: isolation and structural elucidation of gallinamide A. *Journal of Natural Products* **72**: 14-17.

Linne, U., and Marahiel, M.A., (2004) Reactions catalyzed by mature and recombinant nonribosomal peptide synthetases. In: Methods in Enzymology. Academic Press, pp. 293-315.

Liu, T., Chiang, Y.M., Somoza, A.D., Oakley, B.R., and Wang, C.C. (2011) Engineering of an "unnatural" natural product by swapping polyketide synthase domains in *Aspergillus nidulans*. *Journal of the American Chemical Society* **133**: 13314-13316.

Liu, T., Mazmouz, R., and Neilan, B.A. (2018) An *in vitro* and *in vivo* study of broad-range phosphopantetheinyl transferases for heterologous expression of cyanobacterial natural products. *ACS Synthetic Biology* **7**: 1143-1151.

Liu, T., Mazmouz, R., Ongley, S.E., Chau, R., Pickford, R., Woodhouse, J.N., and Neilan, B.A. (2017) Directing the heterologous production of specific cyanobacterial toxin variants. *ACS Chemical Biology***12**: 2021-2029.

Liu, Y., Law, B.K., and Luesch, H. (2009) Apratoxin a reversibly inhibits the secretory pathway by preventing cotranslational translocation. *Molecular Pharmacology* **76**: 91-104.

Llewellyn, L.E., Negri, A.P., Doyle, J., Baker, P.D., Beltran, E.C., and Neilan, B.A. (2001) Radioreceptor assays for sensitive detection and quantitation of saxitoxin and its analogues from strains of the freshwater cyanobacterium, *Anabaena circinalis*. *Environmental science & technology* **35**: 1445-1451.

Lobo, K., Donado, C., Cornelissen, L., Kim, J., Ortiz, R., Peake, R.W., Kellogg, M., Alexander, M.E., Zurakowski, D., Kurgansky, K.E., Peyton, J., Bilge, A., Boretsky, K., McCann, M.E., Berde, C.B., and Cravero, J. (2015) A phase 1, dose-escalation, double-blind, block-randomized, controlled trial of safety and efficacy of neosaxitoxin alone and in combination with 0.2% bupivacaine, with and without epinephrine, for cutaneous anesthesia. *Anesthesiology* **123**: 873-885.

Lorente, A., Gil, A., Fernandez, R., Cuevas, C., Albericio, F., and Alvarez, M. (2015) Phormidolides B and C, cytotoxic agents from the sea: enantioselective synthesis of the macrocyclic core. *Chemistry* **21**: 150-156.

Luesch, H., Moore, R.E., Paul, V.J., Mooberry, S.L., and Corbett, T.H. (2001a) Isolation of dolastatin 10 from the marine cyanobacterium *Symploca* species VP642 and total stereochemistry and biological evaluation of its analogue symplostatin 1. *Journal of Natural Products* **64**: 907-910.

Luesch, H., Yoshida, W.Y., Moore, R.E., and Paul, V.J. (2002a) New apratoxins of marine cyanobacterial origin from Guam and Palau. *Bioorganic & Medicinal Chemistry* **10**: 1973-1978.

Luesch, H., Yoshida, W.Y., Moore, R.E., Paul, V.J., and Corbett, T.H. (2001b) Total structure determination of apratoxin A, a potent novel cytotoxin from the marine cyanobacterium *Lyngbya majuscula*. *Journal of the American Chemical Society* **123**: 5418-5423.

Luesch, H., Yoshida, W.Y., Moore, R.E., Paul, V.J., Mooberry, S.L., and Corbett, T.H. (2002b) Symplostatin 3, a new dolastatin 10 analogue from the marine cyanobacterium *Symploca* sp. VP452. *Journal of Natural Products* **65**: 16-20.

Lukowski, A.L., Denomme, N., Hinze, M.E., Hall, S., Isom, L.L., and Narayan, A.R.H. (2019) Biocatalytic detoxification of paralytic shellfish toxins. *ACS Chemical Biology*: 941-948.

Lukowski, A.L., Ellinwood, D.C., Hinze, M.E., DeLuca, R.J., Du Bois, J., Hall, S., and Narayan, A.R.H. (2018) C–H hydroxylation in paralytic shellfish toxin biosynthesis. *Journal of the American Chemical Society* **140**: 11863-11869.

Lundy, T.A., Mori, S., and Garneau-Tsodikova, S. (2018) Engineering bifunctional enzymes capable of adenylating and selectively methylating the side chain or core of amino acids. *ACS Synthetic Biology* **7**: 399-404.

Luo, Y., Enghiad, B., and Zhao, H. (2016) New tools for reconstruction and heterologous expression of natural product biosynthetic gene clusters. *Natural Product Reports* **33**: 174-182.

Lurling, M., Eshetu, F., Faassen, E.J., Kosten, S., and Huszar, V.L.M. (2013) Comparison of cyanobacterial and green algal growth rates at different temperatures. *Freshwater Biology* **58**: 552-559.

Ma, D., Tang, G., and Kozikowski, A.P. (2002) Synthesis of 7-substituted benzolactam-V8s and their selectivity for protein kinase C isozymes. *Organic Letters* **4**: 2377-2380.

Mahmood, N.A., and Carmichael, W.W. (1986) Paralytic shellfish poisons produced by the freshwater cyanobacterium *Aphanizomenon flos-aquae* NH-5. *Toxicon* **24**: 175-186.

Maloy, S.R., and Nunn, W.D. (1981) Selection for loss of tetracycline resistance by *Escherichia coli*. *Journal of Bacteriology* **145**: 1110-1111.

Manivasagan, P., and Kim, S.-K., (2015) Chapter 34 - An overview of harmful algal blooms on marine organisms. In: Handbook of Marine Microalgae. Boston: Academic Press, pp. 517-526.

Marahiel, M.A. (2016) A structural model for multimodular NRPS assembly lines. *Natural Product Reports* **33**: 136-140.

Marquez, B., Verdier-Pinard, P., Hamel, E., and Gerwick, W.H. (1998) Curacin D, an antimitotic agent from the marine cyanobacterium *Lyngbya majuscula*. *Phytochemistry* **49**: 2387-2389.

Marquez, B.L., Watts, K.S., Yokochi, A., Roberts, M.A., Verdier-Pinard, P., Jimenez, J.I., Hamel, E., Scheuer, P.J., and Gerwick, W.H. (2002) Structure and absolute stereochemistry of hectochlorin, a potent stimulator of actin assembly. *Journal of Natural Products* **65**: 866-871.

Martins, A., Vieira, H., Gaspar, H., and Santos, S. (2014) Marketed marine natural products in the pharmaceutical and cosmeceutical industries: tips for success. *Marine Drugs* **12**: 1066-1101.

Masuda, A., Irie, K., Nakagawa, Y., and Ohigashi, H. (2002) Binding selectivity of conformationally restricted analogues of (−)-indolactam-V to the C1 domains of protein kinase C isozymes. *Bioscience, Biotechnology, and Biochemistry* **66**: 1615-1617.

Matthew, S., Schupp, P.J., and Luesch, H. (2008) Apratoxin E, a cytotoxic peptolide from a guamanian collection of the marine cyanobacterium *Lyngbya bouillonii*. *Journal of Natural Products* **71**: 1113-1116.

McClerren, A.L., Cooper, L.E., Quan, C., Thomas, P.M., Kelleher, N.L., and van der Donk, W.A. (2006) Discovery and *in vitro* biosynthesis of haloduracin, a two-component lantibiotic. *Proceedings of the National Academy of Sciences USA* **103**: 17243-17248.

McGregor, G.B., and Sendall, B.C. (2015) Phylogeny and toxicology of *Lyngbya wollei* (Cyanobacteria, Oscillatoriales) from north-eastern Australia, with a description of *Microseira* gen. nov. *Journal of Phycology* **51**: 109-119.

McPhail, K.L., Correa, J., Linington, R.G., Gonzalez, J., Ortega-Barria, E., Capson, T.L., and Gerwick, W.H. (2007) Antimalarial linear lipopeptides from a Panamanian strain of the marine cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products* **70**: 984-988.

Medina, R.A., Goeger, D.E., Hills, P., Mooberry, S.L., Huang, N., Romero, L.I., Ortega-Barria, E., Gerwick, W.H., and McPhail, K.L. (2008) Coibamide A, a potent antiproliferative cyclic depsipeptide from the panamanian marine cyanobacterium *Leptolyngbya* sp. *Journal of the American Chemical Society* **130**: 6324–6325.

Menzella, H.G., Reisinger, S.J., Welch, M., Kealey, J.T., Kennedy, J., Reid, R., Tran, C.Q., and Santi, D.V. (2006) Redesign, synthesis and functional expression of the 6-deoxyerythronolide B polyketide synthase gene cluster. *Journal of Industrial Microbiology & Biotechnology* **33**: 22-28.

Meseguer, B., Alonso-Diaz, D., Griebenow, N., Herget, T., and Waldmann, H. (2000) Solid-phase synthesis and biological evaluation of a teleocidin library-discovery of a selective PKC down regulator. *Chemistry* **6**: 3943-3957.

Metcalf, J.S., and Codd, G.A. (2003) Analysis of cyanobacterial toxins by immunological methods. *Chemical Research in Toxicology* **16**: 103-112.

Meyer, S., Kehr, J.C., Mainz, A., Dehm, D., Petras, D., Sussmuth, R.D., and Dittmann, E. (2016) Biochemical dissection of the natural diversification of microcystin provides lessons for synthetic biology of NRPS. *Cell Chemical Biology* **23**: 462-471.

Micallef, M.L., D'Agostino, P.M., Al-Sinawi, B., Neilan, B.A., and Moffitt, M.C. (2015) Exploring cyanobacterial genomes for natural product biosynthesis pathways. *Marine Genomics* **21**: 1-12.

Mihali, T.K., Carmichael, W.W., and Neilan, B.A. (2011) A putative gene cluster from a *Lyngbya wollei* bloom that encodes paralytic shellfish toxin biosynthesis. *PLoS ONE* **6**: e14657.

Mihali, T.K., Kellmann, R., and Neilan, B.A. (2009) Characterisation of the paralytic shellfish toxin biosynthesis gene clusters in *Anabaena circinalis* AWQC131C and *Aphanizomenon* sp. NH-5. *BMC Biochemistry* **10**: doi: 10.1186/1471-2091-1110-1188.

Miller, B.R., Drake, E.J., Shi, C., Aldrich, C.C., and Gulick, A.M. (2016) Structures of a nonribosomal peptide synthetase module bound to MbtH-like proteins support a highly dynamic domain architecture. *Journal of Biological Chemistry* **291**: 22559-22571.

Miller, B.R., and Gulick, A.M. (2016) Structural biology of nonribosomal peptide synthetases. *Methods in Molecular Biology* **1401**: 3-29.

Mochly-Rosen, D., Das, K., and Grimes, K.V. (2012) Protein kinase C, an elusive therapeutic target? *Nature Reviews Drug Discovery* **11**: 937– 957.

Montaser, R., and Luesch, H. (2011) Marine natural products: a new wave of drugs? *Future Medicinal Chemistry* **3**: 1475-1489.

Mooberry, S.L., Leal, R.M., Tinley, T.L., Luesch, H., Moore, R.E., and Corbett, T.H. (2003) The molecular pharmacology of symplostatin 1: a new antimitotic dolastatin 10 analog. *International Journal of Cancer* **104**: 512-521.

Moore, R.E., Blackman, A.J., Cheuk, C.E., Mynderse, J.S., Matsumoto, G.K., Clardy, J., Woodard, R.W., and Craig, J.C. (1984) Absolute stereochemistries of the aplysiatoxins and oscillatoxin-A. *Journal of Organic Chemistry* **49**: 2484-2489.

Mootz, H.D., Kessler, N., Linne, U., Eppelmann, K., Schwarzer, D., and Marahiel, M.A. (2002) Decreasing the ring size of a cyclic nonribosomal peptide antibiotic by in-frame module

deletion in the biosynthetic genes. *Journal of the American Chemical Society* **124**: 10980-10981.

Mori, S., Green, K.D., Choi, R., Buchko, G.W., Fried, M.G., and Garneau-Tsodikova, S. (2018) Using MbtH-Like proteins to alter the substrate profile of a nonribosomal peptide adenylation anzyme. *ChemBioChem* **19**: 2186-2194.

Mosmann, T. (1983) Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *Journal of Immunological Methods* **65**: 55-63.

Muir, J.C., Pattenden, G., and Ye, T. (2002) Total synthesis of (+)-curacin A, a novel antimitotic metabolite from a cyanobacterium. *Journal of the Chemical Society-Perkin Transactions 1*: 2243-2250.

Munch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., and Jahn, D. (2005) Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* **21**: 4187-4189.

Murakami, M., Matsuda, H., Makabe, K., and Yamaguchi, K. (1991) Oscillariolide, a novel macrolide from a blue-green alga *Oscillatoria* Sp. *Tetrahedron Letters* **32**: 2391-2394.

Murray, S.A., Diwan, R., Orr, R.J.S., Kohli, G.S., and John, U. (2015) Gene duplication, loss and selection in the evolution of saxitoxin biosynthesis in alveolates. *Molecular Phylogenetics and Evolution* **92**: 165-180.

Murray, S.A., Mihali, T.K., and Neilan, B.A. (2011) Extraordinary conservation, gene loss, and positive selection in the evolution of an ancient neurotoxin. *Molecular Biology and Evolution* **28**: 1173-1182.

Nakagawa, Y. (2012) Artificial analogs of naturally occurring tumor promoters as biochemical tools and therapeutic leads. *Bioscience, Biotechnology, and Biochemistry* **76**: 1262-1274.

Nakagawa, Y., Irie, K., Nakamura, Y., and Ohigashi, H. (2001) The amide hydrogen of (-)-indolactam-V and benzolactam-V8's plays a critical role in protein kinase C binding and tumor-promoting activities. *Bioorganic & Medicinal Chemistry Letters* **11**: 723-728.

Nakagawa, Y., Kikumori, M., Yanagita, R.C., Murakami, A., Tokuda, H., Nagai, H., and Irie, K. (2011) Synthesis and biological evaluation of the 12,12-dimethyl derivative of aplog-1, an anti-proliferative analog of tumor-promoting aplysiatoxin. *Bioscience Biotechnology and Biochemistry* **75**: 1167-1173.

Nelson, T.J., and Alkon, D.L. (2009) Neuroprotective versus tumorigenic protein kinase C activators. *Trends in Biochemical Sciences* **34**: 136-145.

Neuhof, T., Schmieder, P., Preussel, K., Dieckmann, R., Pham, H., Bartl, F., and von Dohren, H. (2005) Hassallidin A, a glycosylated lipopeptide with antifungal activity from the cyanobacterium *Hassallia* sp. *Journal of Natural Products* **68**: 695-700.

Neumann, C.S., Fujimori, D.G., and Walsh, C.T. (2008) Halogenation strategies in natural product biosynthesis. *Chemistry & Biology* **15**: 99-109.

Nguyen, K.T., He, X., Alexander, D.C., Li, C., Gu, J.Q., Mascio, C., Van Praagh, A., Mortin, L., Chu, M., Silverman, J.A., Brian, P., and Baltz, R.H. (2010) Genetically engineered lipopeptide antibiotics related to A54145 and daptomycin with improved properties. *Antimicrobial Agents and Chemotherapy* **54**: 1404-1413.

Nikan, M., Nabavi, S.M., and Manayi, A. (2016) Ligands for cannabinoid receptors, promising anticancer agents. *Life Sciences* **146**: 124-130.

Noda, M., Suzuki, H., Numa, S., and Stuhmer, W. (1989) A single point mutation confers tetrodotoxin and saxitoxin insensitivity on the sodium channel II. *FEBS Letters* **259**: 213-216.

Noji, T., Okano, K., and Tokuyama, H. (2015) A concise total synthesis of (−)-indolactam V. *Tetrahedron* **71**: 3833-3837.

Ogawa, H., Iwasaki, A., Sumimoto, S., Kanamori, Y., Ohno, O., Iwatsuki, M., Ishiyama, A., Hokari, R., Otoguro, K., Omura, S., and Suenaga, K. (2016) Janadolide, a cyclic polyketide-peptide hybrid possessing a tert-butyl group from an *Okeania* sp. marine cyanobacterium. *Journal of Natural Products* **79**: 1862-1866.

Okumura, M., Tsuzuki, H., and Tomita, B. (2005) A rapid detection method for paralytic shellfish poisoning toxins by cell bioassay. *Toxicon* **46**: 93-98.

Ongley, S.E., Bian, X., Neilan, B.A., and Muller, R. (2013a) Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Natural Product Reports* **30**: 1121-1138.

Ongley, S.E., Bian, X., Zhang, Y., Chau, R., Gerwick, W.H., Muller, R., and Neilan, B.A. (2013b) High-titer heterologous production in *E. coli* of lyngbyatoxin, a protein kinase C activator from an uncultured marine cyanobacterium. *ACS Chemical Biology* **8**: 1888-1893.

Orjala, J., and Gerwick, W.H. (1996) Barbamide, a chlorinated metabolite with molluscicidal activity from the Caribbean cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products* **59**: 427-430.

Orr, R.J.S., Stüken, A., Murray, S.A., and Jakobsen, K.S. (2013) Evolutionary acquisition and loss of saxitoxin biosynthesis in dinoflagellates: The second "core" gene, *sxtG*. *Applied and Environmental Microbiology* **79**: 2128-2136.

Oshima, Y., Blackburn, S.I., and Hallegraeff, G.M. (1993) Comparative study on paralytic shellfish toxin profiles of the dinoflagellate *Gymnodinium catenatum* from three different countries. *Marine Biology* **116**: 471-476.

Otten, L.G., Schaffer, M.L., Villiers, B.R., Stachelhaus, T., and Hollfelder, F. (2007) An optimized ATP/PP_i exchange assay in 96-well format for screening of adenylation domains for applications in combinatorial biosynthesis. *Biotechnology Journal* **2**: 232-240.

Patocka, J., Gupta, R.C., Wu, Q.H., and Kuca, K. (2015) Toxic potential of palytoxin. *Journal of Huazhong University of Science and Technology* **35**: 773-780.

Pearson, L.A., Dittmann, E., Mazmouz, R., Ongley, S.E., D'Agostino, P.M., and Neilan, B.A. (2016) The genetics, biosynthesis and regulation of toxic specialized metabolites of cyanobacteria. *Harmful Algae* **54**: 98-111.

Pfeifer, B.A., Admiraal, S.J., Gramajo, H., Cane, D.E., and Khosla, C. (2001) Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291**: 1790-1792.

Piper, C., Cotter, P.D., Ross, R.P., and Hill, C. (2009) Discovery of medically significant lantibiotics. *Current Drug Discovery Technologies* **6**: 1-18.

Pomati, F., Sacchi, S., Rossetti, C., Giovannardi, S., Onodera, H., Oshima, Y., and Neilan, B.A. (2000) The freshwater cyanobacterium *Planktothrix* sp. Fp1: molecular identification and detection of paralytic shellfish poisoning toxins. *Journal of Phycology* **36**: 553-562.

Pratter, S.M., Ivkovic, J., Birner-Gruenberger, R., Breinbauer, R., Zangger, K., and Straganz, G.D. (2014) More than just a halogenase: modification of fatty acyl moieties by a trifunctional metal enzyme. *ChemBioChem* **15**: 567-574.

Quadri, L.E., Sello, J., Keating, T.A., Weinreb, P.H., and Walsh, C.T. (1998a) Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chemistry & Biology* **5**: 631-645.

Quadri, L.E.N., Weinreb, P.H., Lei, M., Nakano, M.M., Zuber, P., and Walsh, C.T. (1998b) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* **37**: 1585-1595.

Ramaswamy, A.V., Sorrels, C.M., and Gerwick, W.H. (2007) Cloning and biochemical characterization of the hectochlorin biosynthetic gene cluster from the marine cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products* **70**: 1977-1986.Ramaswamy, A., V., (2005) Cloning and biochemical characterization of the hectochiorin biosynthetic gene cluster from the marine cyanobacterium *Lyngbya majuscula*. In. Oregon, USA: Oregon State University, pp. 1-179.

Ramos, J.L., Mermod, N., and Timmis, K.N. (1987) Regulatory circuits controlling transcription of TOL plasmid operon encoding meta-cleavage pathway for degradation of alkylbenzoates by *Pseudomonas*. *Molecular Microbiology* **1**: 293-300.

Ramos, V., and Vasconcelos, V. (2010) Palytoxin and analogs: biological and ecological effects. *Marine Drugs* **8**: 2021-2037.

Rausch, C., Weber, T., Kohlbacher, O., Wohlleben, W., and Huson, D.H. (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). *Nucleic Acids Research* **33**: 5799-5808.

Read, J.A., and Walsh, C.T. (2007) The lyngbyatoxin biosynthetic assembly line: chain release by four-electron reduction of a dipeptidyl thioester to the corresponding alcohol. *Journal of the American Chemical Society* **129**: 15762–15763.

Reimer, J.M., Aloise, M.N., Harrison, P.M., and Schmeing, T.M. (2016) Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase. *Nature* **529**: 239-242.

Rein, K.S., and Snyder, R.V. (2006) The biosynthesis of polyketide metabolites by dinoflagellates. *Advances in applied microbiology* **59**: 93-125.

Rey, V., Alfonso, A., Botana, M.L., and Botana, M.A. (2015) Influence of different shellfish matrices on the separation of PSP toxins using a postcolumn oxidation liquid chromatography method. *Toxins* **7**: 1324-1340.

Roberts, A.A., Copp, J.N., Marahiel, M.A., and Neilan, B.A. (2009) The *Synechocystis* sp. PCC6803 Sfp-type phosphopantetheinyl transferase does not possess characteristic broad-range activity. *ChemBioChem* **10**: 1869-1877.

Rodriguez-Navarro, A.J., Lagos, M., Figueroa, C., Garcia, C., Recabal, P., Silva, P., Iglesias, V., and Lagos, N. (2009) Potentiation of local anesthetic activity of neosaxitoxin with bupivacaine or epinephrine: development of a long-acting pain blocker. *Neurotoxicity Research* **16**: 408-415.

Rodriguez-Navarro, A.J., Lagos, N., Lagos, M., Braghetto, I., Csendes, A., Hamilton, J., Figueroa, C., Truan, D., Garcia, C., Rojas, A., Iglesias, V., Brunet, L., and Alvarez, F. (2007) Neosaxitoxin as a local anesthetic: preliminary observations from a first human trial. *Anesthesiology* **106**: 339-345.

Rossi, J.V., Roberts, M.A., Yoo, H.D., and Gerwick, W.H. (1997) Pilot scale culture of the marine cyanobacterium *Lyngbya majuscula* for its pharmaceutically-useful natural metabolite curacin A. *Journal of Applied Phycology* **9**: 195-204.

Röttig, M., Medema, M.H., Blin, K., Weber, T., Rausch, C., and Kohlbacher, O. (2011) NRPSpredictor2—a web server for predicting NRPS adenylation domain specificity. *Nucleic Acids Research* **39**: W362-W367.

Rubio, D.P., Roa, L.G., Soto, D.A., Velasquez, F.J., Gregorcic, N.A., Soto, J.A., Martinez, M.C., Kalergis, A.M., and Vasquez, A.E. (2015) Purification and characterization of saxitoxin from *Mytilus chilensis* of southern Chile. *Toxicon* **108**: 147-153.

Ruffner, D.E., Schmidt, E.W., and Heemstra, J.R. (2015) Assessing the combinatorial potential of the RiPP cyanobactin *tru* pathway. *ACS Synthetic Biology* **4**: 482-492.

Saida, F., Uzan, M., Odaert, B., and Bontems, F. (2006) Expression of highly toxic genes in *E. coli*: special strategies and genetic tools. *Current Protein & Peptide Science* **7**: 47-56.

Sakai, S.I., Hitotsuyanagi, Y., Aimi, N., Fujiki, H., Suganuma, M., Sugimura, T., Endo, Y., and Shudo, K. (1986) Absolute configuration of lyngbyatoxin A (teleocidin A-1) and teleocidin A-2. *Tetrahedron Letters* **27**: 5219-5220.

Sanchez, L.M., Lopez, D., Vesely, B.A., Della Togna, G., Gerwick, W.H., Kyle, D.E., and Linington, R.G. (2010) Almiramides A-C: discovery and development of a new class of leishmaniasis lead compounds. *Journal of Medicinal Chemistry* **53**: 4187-4197.

Sardar, D., and Schmidt, E.W. (2016) Combinatorial biosynthesis of RiPPs: docking with marine life. *Current Opinion in Chemical Biology* **31**: 15-21.

Schantz, E.J., Ghazarossian, V.E., Schnoes, H.K., Strong, F.M., Springer, J.P., Pezzanite, J.O., and Clardy, J. (1975) The structure of saxitoxin. *Journal of the American Chemical Society* **97**: 1238-1239.

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

Schneider, A., Stachelhaus, T., and Marahiel, M.A. (1998) Targeted alteration of the substrate specificity of peptide synthetases by rational module swapping. *Molecular Genetics and Genomics* **257**: 308-318.

Serrill, J.D., Wan, X., Hau, A.M., Jang, H.S., Coleman, D.J., Indra, A.K., Alani, A.W.G., McPhail, K.L., and Ishmael, J.E. (2015) Coibamide A, a natural lariat depsipeptide, inhibits VEGFA/VEGFR2 expression and suppresses tumor growth in glioblastoma xenografts. *Investigational new drugs*: 24-40.

Shao, C.-L., Linington, R.G., Balunas, M.J., Centeno, A., Boudreau, P., Zhang, C., Engene, N., Spadafora, C., Mutka, T.S., Kyle, D.E., Gerwick, L., Wang, C.-Y., and Gerwick, W.H. (2015) Bastimolide A, a potent antimalarial polyhydroxy macrolide from the marine cyanobacterium *Okeania hirsuta*. *Journal of Organic Chemistry* **80**: 7849-7855.

Shen, B. (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* **7**: 285-295.

Shimizu, Y., Norte, M., Hori, A., Genenah, A., and Kobayashi, M. (1984) Biosynthesis of saxitoxin analogues: the unexpected pathway. *Journal of the American Chemical Society* **106**: 6433-6434.

Simmons, T.L., Coates, R.C., Clark, B.R., Engene, N., Gonzalez, D., Esquenazi, E., Dorrestein, P.C., and Gerwick, W.H. (2008) Biosynthetic origin of natural products isolated from marine microorganism-invertebrate assemblages. *Proceedings of the National Academy of Sciences USA* **105**: 4587-4594.

Sitachitta, N., Marquez, B.L., Williamson, R.T., Rossi, J., Roberts, M.A., Gerwick, W.H., Nguyen, V.A., and Willis, C.L. (2000) Biosynthetic pathway and origin of the chlorinated methyl group in barbamide and dechlorobarbamide, metabolites from the marine cyanobacterium *Lyngbya majuscula*. *Tetrahedron* **56**: 9103-9113.

Sitachitta, N., Rossi, J., Roberts, M.A., Gerwick, W.H., Fletcher, M.D., and Willis, C.L. (1998) Biosynthesis of the marine cyanobacterial metabolite barbamide. 1. Origin of the trichloromethyl group. *Journal of the American Chemical Society* **120**: 7131-7132.

Sivonen, K., Leikoski, N., Fewer, D.P., and Jokela, J. (2010) Cyanobactins-ribosomal cyclic peptides produced by cyanobacteria. *Applied microbiology and biotechnology* **86**: 1213-1225.

Sletta, H., Nedal, A., Aune, T.E., Hellebust, H., Hakvag, S., Aune, R., Ellingsen, T.E., Valla, S., and Brautaset, T. (2004) Broad-host-range plasmid pJB658 can be used for industrial-level production of a secreted host-toxic single-chain antibody fragment in *Escherichia coli*. *Applied and Environmental Microbiology* **70**: 7033-7039.

Sletta, H., Tondervik, A., Hakvag, S., Aune, T.E., Nedal, A., Aune, R., Evensen, G., Valla, S., Ellingsen, T.E., and Brautaset, T. (2007) The presence of N-terminal secretion signal sequences leads to strong stimulation of the total expression levels of three tested medically important proteins during high-cell-density cultivations of *Escherichia coli*. *Applied and Environmental Microbiology* **73**: 906-912.

Smith, F.M., Wood, S.A., van Ginkel, R., Broady, P.A., and Gaw, S. (2011) First report of saxitoxin production by a species of the freshwater benthic cyanobacterium, *Scytonema Agardh*. *Toxicon* **57**: 566-573.

Soto-Liebe, K., Lopez-Cortes, X.A., Fuentes-Valdes, J.J., Stucken, K., Gonzalez-Nilo, F., and Vasquez, M. (2013) *In silico* analysis of putative paralytic shellfish poisoning toxins export proteins in cyanobacteria. *PLoS ONE* **8**: e55664.

Soto-Liebe, K., Murillo, A.A., Krock, B., Stucken, K., Fuentes-Valdes, J.J., Trefault, N., Cembella, A., and Vasquez, M. (2010) Reassessment of the toxin profile of *Cylindrospermopsis raciborskii* T3 and function of putative sulfotransferases in synthesis of sulfated and sulfonated PSP toxins. *Toxicon* **56**: 1350-1361.

Stachelhaus, T., Huser, A., and Marahiel, M.A. (1996) Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chemistry & Biology* **3**: 913-921.

Stachelhaus, T., Mootz, H.D., Bergendahl, V., and Marahiel, M.A. (1998) Peptide bond formation in nonribosomal peptide biosynthesis: catalytic role of the condensation domain. *Journal of Biological Chemistry* **273**: 22773-22781.

Stachelhaus, T., Mootz, H.D., and Marahiel, M.A. (1999) The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chemistry & Biology* **6**: 493-505.

Stat, M., Morris, E., and Gates, R.D. (2008) Functional diversity in coral: dinoflagellate symbiosis. *Proceedings of the National Academy of Sciences USA* **105**: 9256-9261.

Stavropoulos, T.A., and Strathdee, C.A. (2001) Synergy between *tetA* and *rpsL* provides high-stringency positive and negative selection in bacterial artificial chromosome vectors. *Genomics* **72**: 99-104.

Steiniger, C., Hoffmann, S., Mainz, A., Kaiser, M., Voigt, K., Meyer, V., and Süssmuth, R.D. (2017) Harnessing fungal nonribosomal cyclodepsipeptide synthetases for mechanistic insights and tailored engineering. *Chemical Science* **8**: 7834-7843.

Steinmetz, M., Le Coq, D., Djemia, H.B., and Gay, P. (1983) Genetic analysis of *sacB*, the structural gene of a secreted enzyme, levansucrase of *Bacillus subtilis* Marburg. *Molecular & General Genetics* **191**: 138-144.

Stucken, K., John, U., Cembella, A., Murillo, A.A., Soto-Liebe, K., Fuentes-Valdes, J.J., Friedel, M., Plominsky, A.M., Vasquez, M., and Glockner, G. (2010) The smallest known genomes of multicellular and toxic cyanobacteria: comparison, minimal gene sets for linked traits and the evolutionary implications. *PLoS ONE* **5**: e9235.

Stuken, A., Orr, R.J., Kellmann, R., Murray, S.A., Neilan, B.A., and Jakobsen, K.S. (2011) Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. *PLoS ONE* **6**: e20096.

Su, Z., Sheets, M., Ishida, H., Li, F., and Barry, W.H. (2004) Saxitoxin blocks L-Type I_{Ca}. *Journal of Pharmacology and Experimental Therapeutics* **308**: 324–329.

Sudek, S., Haygood, M.G., Youssef, D.T., and Schmidt, E.W. (2006) Structure of trichamide, a cyclic peptide from the bloom-forming cyanobacterium *Trichodesmium erythraeum*, predicted from the genome sequence. *Applied and Environmental Microbiology* **72**: 4382-4387.

Šulčius, S., Slavuckytė, K., Januškaitė, M., and Paškauskas, R. (2017) Establishment of axenic cultures from cyanobacterium *Aphanizomenon flos-aquae* akinetes by micromanipulation and chemical treatment. *Algal Research* **23**: 43-50.

Suntornchashwej, S., Chaichit, N., Isobe, M., and Suwanborirux, K. (2005) Hectochlorin and morpholine derivatives from the Thai sea hare, *Bursatella leachii*. *Journal of Natural Products* **68**: 951-955.

Suo, Z., Tseng, C.C., and Walsh, C.T. (2001) Purification, priming, and catalytic acylation of carrier protein domains in the polyketide synthase and nonribosomal peptidyl synthetase modules of the HMWP1 subunit of yersiniabactin synthetase. *Proceedings of the National Academy of Sciences USA* **98**: 99-104.

Tan, L.T. (2007) Bioactive natural products from marine cyanobacteria for drug discovery. *Phytochemistry* **68**: 954-979.

Tanaka-Yanuma, A., Watanabe, S., Ogawa, K., Watanabe, S., Aoki, N., Ogura, T., and Usuki, T. (2015) Synthesis of the polyketide moiety of the jamaicamides. *Tetrahedron Letters* **56**: 6777-6781.

Tanaka, A., and Usuki, T. (2011) Synthesis of the peptide moiety of the jamaicamides. *Tetrahedron Letters* **52**: 5036-5038.

Tang, W., and van der Donk, W.A. (2012) Structural characterization of four prochlorosins: a novel class of lantipeptides produced by planktonic marine cyanobacteria. *Biochemistry* **51**: 4271-4279.

Taori, K., Liu, Y., Paul, V.J., and Luesch, H. (2009) Combinatorial strategies by marine cyanobacteria: symplostatin 4, an antimitotic natural dolastatin 10/15 hybrid that synergizes with the coproduced HDAC inhibitor largazole. *ChemBioChem* **10**: 1634-1639.

Tarry, M.J., Haque, A.S., Bui, K.H., and Schmeing, T.M. (2017) X-Ray crystallography and electron microscopy of cross- and multi-module nonribosomal peptide synthetase proteins reveal a flexible architecture. *Structure* **25**: 783-793.

Terlau, H., Heinemann, S.H., Stuhmer, W., Pusch, M., Conti, F., Imoto, K., and Numa, S. (1991) Mapping the site of block by tetrodotoxin and saxitoxin of sodium channel II. *FEBS Letters* **293**: 93-96.

Thibodeaux, C.J., Ha, T., and van der Donk, W.A. (2014) A price to pay for relaxed substrate specificity: a comparative kinetic analysis of the class II lanthipeptide synthetases ProcM and HalM2. *Journal of the American Chemical Society* **136**: 17513-17529.

Thirlway, J., Lewis, R., Nunns, L., Al Nakeeb, M., Styles, M., Struck, A.W., Smith, C.P., and Micklefield, J. (2012) Introduction of a non-natural amino acid into a nonribosomal peptide antibiotic by modification of adenylation domain specificity. *Angewandte Chemie International Edition* **51**: 7181-7184.

Thomas, T.R.A., Kavlekar, D.P., and LokaBharathi, P.A. (2010) Marine drugs from sponge-microbe association—a review. *Marine Drugs* **8**: 1417-1468.

Thornburg, C.C., Cowley, E.S., Sikorska, J., Shaala, L.A., Ishmael, J.E., Youssef, D.T.A., and McPhail, K.L. (2013) Apratoxin H and apratoxin A sulfoxide from the red sea cyanobacterium *Moorea producens*. *Journal of Natural Products* **76**: 1781-1788.

Tianero, M.D., Donia, M.S., Young, T.S., Schultz, P.G., and Schmidt, E.W. (2012) Ribosomal route to small-molecule diversity. *Journal of the American Chemical Society* **134**: 418-425.

Tidgewell, K., Engene, N., Byrum, T., Media, J., Doi, T., Valeriote, F.A., and Gerwick, W.H. (2010) Evolved diversification of a modular natural product pathway: apratoxins F and G, two cytotoxic cyclic depsipeptides from a Palmyra collection of *Lyngbya bouillonii*. *ChemBioChem* **11**: 1458-1466.

Trauger, J.W., and Walsh, C.T. (2000) Heterologous expression in *Escherichia coli* of the first module of the nonribosomal peptide synthetase for chloroeremomycin, a vancomycin-type glycopeptide antibiotic. *Proceedings of the National Academy of Sciences USA* **97**: 3112-3117.

Tsuchiya, S., Cho, Y., Konoki, K., Nagasawa, K., Oshima, Y., and Yotsu-Yamashita, M. (2014) Synthesis and identification of proposed biosynthetic intermediates of saxitoxin in the cyanobacterium *Anabaena circinalis* (TA04) and the dinoflagellate *Alexandrium tamarense* (Axat-2). *Organic & Biomolecular Chemistry* **12**: 3016-3020.

Tsuchiya, S., Cho, Y., Konoki, K., Nagasawa, K., Oshima, Y., and Yotsu-Yamashita, M. (2015) Synthesis of a tricyclic bisguanidine compound structurally related to saxitoxin and its identification in paralytic shellfish toxin-producing microorganisms. *Chemistry* **21**: 7835-7840.

Tsuchiya, S., Cho, Y., Konoki, K., Nagasawa, K., Oshima, Y., and Yotsu-Yamashita, M. (2016) Biosynthetic route towards saxitoxin and shunt pathway. *Scientific Reports* **6**: 20340.

Tsuchiya, S., Cho, Y., Yoshioka, R., Konoki, K., Nagasawa, K., Oshima, Y., and Yotsu-Yamashita, M. (2017) Synthesis and identification of key biosynthetic intermediates for the formation of the tricyclic skeleton of saxitoxin. *Angewandte Chemie International Edition* **56**: 5327-5331.

Turgay, K., Krause, M., and Marahiel, M.A. (1992) Four homologous domains in the primary structure of GrsB are related to domains in a superfamily of adenylate-forming enzymes. *Molecular Microbiology* **6**: 2743-2744.

Turner, A., Dhanji-Rapkova, M., Dean, K., Milligan, S., Hamilton, M., Thomas, J., Poole, C., Haycock, J., Spelman-Marriott, J., Watson, A., Hughes, K., Marr, B., Dixon, A., and Coates, L. (2018) Fatal canine intoxications linked to the presence of saxitoxins in stranded marine organisms following winter storm activity. *Toxins* **10**: E94.

Ulbricht, W. (1998) Effects of veratridine on sodium currents and fluxes. *Reviews of Physiology, Biochemistry and Pharmacology* **133**: 1-54.

Usup, G., Kulis, D.M., and Anderson, D.M. (1994) Growth and toxin production of the toxic dinoflagellate *Pyrodinium bahamense* var. compressum in laboratory cultures. *Natural Toxins* **2**: 254-262.

Vaillancourt, F.H., Yin, J., and Walsh, C.T. (2005) SyrB2 in syringomycin E biosynthesis is a nonheme Fe^{II} alpha-ketoglutarate- and O_2-dependent halogenase. *Proceedings of the National Academy of Sciences USA* **102**: 10111-10116.

Verdier-Pinard, P., Bai, R., and Hamel, E. (1998a) Retention of dolastatin 10 by cellular tubulin: basis for the higher potency of peptide-like antimitotic agents. *Molecular Biology of the Cell* **9**: 253-265.

Verdier-Pinard, P., Lai, J.Y., Yoo, H.D., Yu, J., Marquez, B., Nagle, D.G., Nambu, M., White, J.D., Falck, J.R., Gerwick, W.H., Day, B.W., and Hamel, E. (1998b) Structure-activity analysis of the interaction of curacin A, the potent colchicine site antimitotic agent, with tubulin and effects of analogs on the growth of MCF-7 breast cancer cells. *Molecular Pharmacology* **53**: 62-76.

Vestola, J., Shishido, T.K., Jokela, J., Fewer, D.P., Aitio, O., Permi, P., 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. *Proceedings of the National Academy of Sciences USA* **111**: E1909-E1917.

Videau, P., Wells, K.N., Singh, A.J., Gerwick, W.H., and Philmus, B. (2016) Assessment of *Anabaena* sp. strain PCC 7120 as a heterologous expression host for cyanobacterial natural products: production of lyngbyatoxin A. *ACS Synthetic Biology*: 978-988.

Villiers, B., and Hollfelder, F. (2011) Directed evolution of a gatekeeper domain in nonribosomal peptide synthesis. *Chemistry & Biology* **18**: 1290-1299.

Vobruba, S., Kadlcik, S., Gazak, R., and Janata, J. (2017) Evolution-guided adaptation of an adenylation domain substrate specificity to an unusual amino acid. *PLoS ONE* **12**: e0189684.

Voigt, K., Sharma, C.M., Mitschke, J., Lambrecht, S.J., Voss, B., Hess, W.R., and Steglich, C. (2014) Comparative transcriptomics of two environmentally relevant cyanobacteria reveals unexpected transcriptome diversity. *The ISME Journal* **8**: 2056-2068.

Vu, C.H.T., Lee, H.G., Chang, Y.K., and Oh, H.M. (2018) Axenic cultures for microalgal biotechnology: establishment, assessment, maintenance, and applications. *Biotechnology Advances* **36**: 380-396.

Wacklin, P., Hoffmann, L., and Komárek, J. (2009) Nomenclatural validation of the genetically revised cyanobacterial genus *Dolichospermum* (RALFS ex BORNET et FLAHAULT) comb. nova. *Fottea* **9**: 59-64.

Walsh, C.T., Gehring, A.M., Weinreb, P.H., Quadri, L.E., and Flugel, R.S. (1997a) Post-translational modification of polyketide and nonribosomal peptide synthases. *Current Opinion in Chemical Biology* **1**: 309-315.

Walsh, C.T., Gehring, A.M., Weinreb, P.H., Quadri, L.E.N., and Flugel, R.S. (1997b) Post-translational modification of polyketide and nonribosomal peptide synthases. *Current Opinion in Chemical Biology* **1**: 309-315.

Wang, D.-Z. (2008) Neurotoxins from marine dinoflagellates: a brief review. *Marine Drugs* **6**: 349-371.

Wang, H., Bian, X., Xia, L., Ding, X., Muller, R., Zhang, Y., Fu, J., and Stewart, A.F. (2014) Improved seamless mutagenesis by recombineering using ccdB for counterselection. *Nucleic Acids Research* **42**: e37.

Wang, J., Salata, J.J., and Bennett, P.B. (2003) Saxitoxin is a gating modifier of HERG K^+ channels. *Journal of General Physiology* **121**: 583–598.

Watanabe, S., Watanabe, S., Aoki, N., and Usuki, T. (2013) Synthesis of the polyketide (E)-olefin of the jamaicamides. *Synthetic Communications* **43**: 1397-1403.

Weissman, K.J. (2015) The structural biology of biosynthetic megaenzymes. *Nature Chemical Biology* **11**: 660-670.

Wiese, M., D'Agostino, P.M., Mihali, T.K., Moffitt, M.C., and Neilan, B.A. (2010) Neurotoxic alkaloids: saxitoxin and its analogs. *Marine Drugs* **8**: 2185-2211.

Williams, G.J. (2013) Engineering polyketide synthases and nonribosomal peptide synthetases. *Current Opinion in Structural Biology* **23**: 603-612.

Williamson, R.T., Boulanger, A., Vulpanovici, A., Roberts, M.A., and Gerwick, W.H. (2002) Structure and absolute stereochemistry of phormidolide, a new toxic metabolite from the marine cyanobacterium *Phormidium* sp. *Journal of Organic Chemistry* **67**: 7927-7936.

Williamson, R.T., Sitachitta, N., and Gerwick, W.H. (1999) Biosynthesis of the marine cyanobacterial metabolite barbamide. 2: Elucidation of the origin of the thiazole ring by application of a new GHNMBC experiment. *Tetrahedron Letters* **40**: 5175-5178.

Winther-Larsen, H.C., Josefsen, K.D., Brautaset, T., and Valla, S. (2000) Parameters affecting gene expression from the *Pm* promoter in gram-negative bacteria. *Metabolic Engineering* **2**: 79-91.

Wipf, P., Reeves, J.R., and Day, B. (2002a) Synthesis and biological evaluation of a targeted library of curacin A analogs. *Abstracts of Papers of the American Chemical Society* **223**: B176-B177.

Wipf, P., Reeves, J.T., Balachandran, R., and Day, B.W. (2002b) Synthesis and biological evaluation of structurally highly modified analogues of the antimitotic natural product curacin A. *Journal of Medicinal Chemistry* **45**: 1901-1917.

Withers, N., Vidaver, W., and Lewin, R.A. (1978) Pigment composition, photosynthesis and fine structure of a non-blue-green prokaryotic algal symbiont (*Prochloron* sp) in a didemnid ascidian from Hawaiian waters. *Phycologia* **17**: 167-171.

Wolfe, A.J. (2005) The acetate switch. *Microbiology and Molecular Biology Reviews* **69**: 12-50.

Worsey, M.J., and Williams, P.A. (1975) Metabolism of toluene and xylenes by *Pseudomonas (putida (arvilla)* mt-2: evidence for a new function of the TOL plasmid. *Journal of Bacteriology* **124**: 7-13.

Wylie, M.C., Johnson, V.M., Carpino, E., Mullen, K., Hauser, K., Nedder, A., Kheir, J.N., Rodriguez-Navarro, A.J., Zurakowski, D., and Berde, C.B. (2012) Respiratory, neuromuscular, and cardiovascular effects of neosaxitoxin in isoflurane-anesthetized sheep. *Regional Anesthesia and Pain Medicine* **37**: 152-158.

Xu, Z., Zhang, F., Zhang, L., and Jia, Y. (2011) Total synthesis of (−)-indolactam V. *Organic & Biomolecular Chemistry* **9**: 2512-2517.

Yamanaka, K., Reynolds, K.A., Kersten, R.D., Ryan, K.S., Gonzalez, D.J., Nizet, V., Dorrestein, P.C., and Moore, B.S. (2014) Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proceedings of the National Academy of Sciences USA* **111**: 1957-1962.

Yan, F., Burgard, C., Popoff, A., Zaburannyi, N., Zipf, G., Maier, J., Bernauer, H.S., Wenzel, S.C., and Muller, R. (2018a) Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production. *Chemical Science* **9**: 7510-7519.

Yan, F., Burgard, C., Popoff, A., Zaburannyi, N., Zipf, G., Maier, J., Bernauer, H.S., Wenzel, S.C., and Müller, R. (2018b) Synthetic biology approaches and combinatorial biosynthesis towards heterologous lipopeptide production. *Chemical Science* **9**: 7510-7519.

Yanagita, R.C., Kamachi, H., Tanaka, K., Murakami, A., Nakagawa, Y., Tokuda, H., Nagai, H., and Irie, K. (2010) Role of the phenolic hydroxyl group in the biological activities of simplified analogue of aplysiatoxin with antiproliferative activity. *Bioorganic & Medicinal Chemistry Letters* **20**: 6064-6066.

Yanagita, R.C., Nakagawa, Y., Yamanaka, N., Kashiwagi, K., Saito, N., and Irie, K. (2008) Synthesis, conformational analysis, and biological evaluation of 1-hexylindolactam-V10 as a selective activator for novel protein kinase C isozymes. *Journal of Medicinal Chemistry* **51**: 46-56.

Yang, J.Y., Sanchez, L.M., Rath, C.M., Liu, X., Boudreau, P.D., Bruns, N., Glukhov, E., Wodtke, A., de Felicio, R., Fenner, A., Wong, W.R., Linington, R.G., Zhang, L., Debonsi, H.M., Gerwick, W.H., and Dorrestein, P.C. (2013) Molecular networking as a dereplication strategy. *Journal of Natural Products* **76**: 1686-1699.

Yoo, H.D., and Gerwick, W.H. (1995) Curacins B and C, new antimitotic natural products from the marine cyanobacterium *Lyngbya majuscula*. *Journal of Natural Products* **58**: 1961-1965.

Yoon, H.S., Hackett, J.D., Van Dolah, F.M., Nosenko, T., Lidie, K.L., and Bhattacharya, D. (2005) Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. *Molecular Biology and Evolution* **22**: 1299-1308.

Youssef, D.T.a., Shaala, L.a., Mohamed, G.a., Ibrahim, S.R.M., Banjar, Z.M., Badr, J.M., McPhail, K.L., Risinger, A.L., and Mooberry, S.L. (2015) 2,3-Seco-2,3-dioxo-lyngbyatoxin A from a Red Sea strain of the marine cyanobacterium *Moorea producens*. *Natural Product Research* **29**: 703-709.

Yu, Y., Zhang, Q., and van der Donk, W.A. (2013) Insights into the evolution of lanthipeptide biosynthesis. *Protein Science* **22**: 1478-1489.

Zhang, B., Tian, W., Wang, S., Yan, X., Jia, X., Pierens, G.K., Chen, W., Ma, H., Deng, Z., and Qu, X. (2017) Activation of natural products biosynthetic pathways via a protein modification level regulation. *ACS Chemical Biology* **12**: 1732-1736.

Zhang, F., Wang, Y., Jiang, Q., Chen, Q., Karthik, L., Zhao, Y.L., and Li, Z. (2018) Substrate selection of adenylation domains for nonribosomal peptide synthetase (NRPS) in bacillamide C biosynthesis by marine *Bacillus atrophaeus* C89. *Journal of Industrial Microbiology & Biotechnology* **45**: 335-344.

Zhang, K., Nelson, K.M., Bhuripanyo, K., Grimes, K.D., Zhao, B., Aldrich, C.C., and Yin, J. (2013) Engineering the substrate specificity of the DhbE adenylation domain by yeast cell surface display. *Chemistry & Biology* **20**: 92-101.

Zhang, L., Hoshino, S., Awakawa, T., Wakimoto, T., and Abe, I. (2016) Structural diversification of lyngbyatoxin A by host-dependent heterologous expression of the *tleABC* biosynthetic gene cluster. *ChemBioChem* **17**: 1407-1411.

Zhang, L.H., Longley, R.E., and Koehn, F.E. (1997) Antiproliferative and immunosuppressive properties of microcolin A, a marine-derived lipopeptide. *Life Sciences* **60**: 751-762.

Zhang, Q., Yang, X., Wang, H., and van der Donk, W.A. (2014) High Divergence of the Precursor Peptides in Combinatorial Lanthipeptide Biosynthesis. *ACS Chemical Biology* **9**: 2686-2694.

Zhang, W., Heemstra, J.R., Jr., Walsh, C.T., and Imker, H.J. (2010) Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins. *Biochemistry* **49**: 9946-9947.

Zolova, O.E., and Garneau-Tsodikova, S. (2012) Importance of the MbtH-like protein TioT for production and activation of the thiocoraline adenylation domain of TioK. *MedChemComm* **3**: 950-955.

Zwick, F., Lale, R., and Valla, S. (2012) Strong stimulation of recombinant protein production in *Escherichia coli* by combining stimulatory control elements in an expression cassette. *Microbial Cell Factories* **11**: 133.