UNIVERSITY OF CALIFORNIA SAN DIEGO

Heterologous Expression of Biosynthetic Gene Clusters from Marine Cyanobacteria

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science

in

Biology

by

Brienna Diaz

Committee in charge:

James Golden, Chair
Rachel Dutton
Lena Gerwick
Susan Golden

2019

The Thesis of Brienna Diaz is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

2019

Table of Contents

List of Abbreviations

A – Adenylation

ACP – Acyl carrier protein

AT – Acetyltransferase

Bp – Base pair

BGC – Biosynthetic gene cluster

C – Condensation

Cm – Chloramphenicol

gDNA – Genomic DNA

Kb – Kilobase

Km – Kanamycin

KS – Ketoacyl synthase

LC-MS/MS – Liquid chromatography-Mass spectrometry

MALDI-TOF MS – Matrix Associated Laser Desorption/Ionization-Time of Flight Mass Spectrometry

Nm – Neomycin

NMR – Nuclear Magnetic Resonance

NP – Natural Product

NRPS – Nonribosomal peptide synthase

NS – Neutral site

PCP – Peptidyl carrier protein

PCR – Polymerase chain reaction

PKS – Polyketide synthase

PPTase – phosphopantetheinyl transferase

Rpm – Rotations per minute

SNP – Single nucleotide polymorphism

Sp/Sm – Spectinomycin/streptomycin

Tc - Tetracycline

WT – Wild type

List of Figures

List of Tables

List of Images

Acknowledgements

I would like to thank all members of the Golden Lab for providing a warm and welcoming place to work during my research as both an undergraduate and a graduate student at UCSD. Thank you, Jim and Susan, for welcoming me into your lab and providing an encouraging environment to foster my growth as a scientist. I specifically want to thank Arnaud for teaching me the foundations of cyanobacterial microbiology. Thank you, Arnaud, for always being so patient with me and for being such a great mentor. Thank you also to members of the Gerwick Lab: Andrew Ecker, Raphael Reher, Sebastian Rorher, Lena Gerwick and Bill Gerwick for collaborations on this work.

I would also like to express my gratitude for all of my family's unconditional love and moral support. Lastly, I need to thank my best friend Sean Aceto for simultaneously being my biggest inspiration and strongest supporter. I cannot thank you all enough, you've made such an incredible impact on my life.

Chapter 3 is coauthored with Diaz, Brienna; Taton, Arnaud; Ecker, Andrew; Reher, Raphael; Rohrer, Sebastian; Gerwick, Lena; Gerwick, William H.; and Golden, James. The thesis author was the primary author of this chapter.

ABSTRACT OF THESIS

Heterologous Expression of Biosynthetic Gene Clusters from Marine Cyanobacteria

by

Brienna Diaz

Master of Science in Biology

University of California, San Diego, 2019

Professor James Golden, Chair

Cyanobacteria are prolific producers of natural products with diverse structures and biological activities. Natural products are important for drug discovery and some of those isolated from cyanobacteria were found to have activities against cancer cells and infectious disease, or to relieve minor pain as anti-inflammatory drugs. However, slow growth rates, resistance to genetic engineering, and low yields limit the thorough investigation of many of these secondary metabolites within their native organisms. In this study, we describe and demonstrate the successful heterologous production of two bioactive secondary metabolites from

marine cyanobacteria, cryptomaldamide and columbamide, in *Anabaena* sp. PCC 7120. This work resulted in the development and validation of new genetic tools and methods for the heterologous expression of large biosynthetic gene clusters. These tools include cloning vectors for transformation associated recombination in yeast and a new engineered strain of *Anabaena* along with a CRISPR/cpf1-based system that enables genetic engineering of large biosynthetic gene clusters in both *Synechococcus elongatus* PCC 7942 or *Anabaena*. This study showed that *Anabaena* is a more suitable host than *S. elongatus* for the production of marine cyanobacterial natural products. The heterologous expression of these biosynthetic gene clusters from marine cyanobacteria represents a significant opportunity to produce valuable compounds at a larger scale and for the further study of the mechanisms involved in the synthesis of these compounds.

Chapter One: Introduction

1.1 Cyanobacteria

Cyanobacteria represent a diverse phylum, containing species with different physiologies, morphologies, habitats, and genomes. They are the only prokaryotes on Earth capable of oxygenic photosynthesis (1). Their diversity is represented by over 2,500 species from over 300 different genera, with an estimation of over 3,000 species yet to be discovered (2). The diversity of cyanobacteria is supported by their great genetic diversity, with G/C content ranging from 31%-71% and genome sizes ranging from 1.44 Mb to 9.05 Mb (3, 4). Cyanobacteria are known to exist in a wide range of habitats such as freshwater, ocean, and terrestrial ecosystems, where they play critical roles in the global nitrogen and carbon cycles.

Like plants, cyanobacteria fix carbon dioxide and use energy from light to synthesize sugars while releasing oxygen as a byproduct. The unique ability to produce oxygen has allowed cyanobacteria to fill the critical role of oxygenating the Earth's atmosphere (5, 6). Some cyanobacterial species have also evolved to fix nitrogen from the atmosphere to produce inorganic forms such as nitrate and ammonia -- they are one of only a few groups of organisms able to do so. The nitrogenase enzyme that fixes atmospheric nitrogen is highly sensitive to oxygen, therefore, nitrogen-fixing cyanobacterial species have developed strategies to separate nitrogen fixation from oxygenic photosynthesis either spatially or temporally (7). Some filamentous nitrogen-fixing cyanobacteria such as *Anabaena* sp. PCC 7120 have evolved to produce differentiated cells called heterocysts, which physically isolate the organism's nitrogen fixation machinery from oxygenic photosynthesis (7).

Many cyanobacteria are also known for their production of toxins (8). Some strains are responsible for harmful algal blooms (HABs), which occur both in oceans and freshwater lakes

and can have adverse consequences for humans (8). While the cyanotoxins produced by algal blooms are often feared and viewed by the general public as dangerous and unfavorable, cyanobacteria are also responsible for making hundreds of potentially beneficial natural products (NPs) which have yet to be investigated thoroughly (9-11).

Not only do cyanobacteria play very important roles in their natural environments, but this phylum has also been very significant for scientific research. Along with their ability to fix nitrogen and CO_2 and to produce oxygen, many other features make some species of cyanobacteria of interest to science such as their genetic tractability, circadian rhythms, and production of a variety of secondary metabolites that may be useful to humans. Some cyanobacteria, particularly those among the marine filamentous species, are known to make natural products that include a wide range of molecules with distinct bioactivities (12,13). A number of these secondary metabolites have the potential to be used against cancer, as antibiotics to fight infectious disease, or to relieve minor pain as anti-inflammatory drugs (9,13).

1.2 Natural Products

Cyanobacteria are well known to produce a variety of natural products. In fact, 90 different genera were reported to produce bioactive metabolites (13). Furthermore, all lineages of the cyanobacterial phylum were found to have biosynthetic gene clusters (BGCs) encoding natural products (10, 11, 14).

Natural products from cyanobacteria tend to have unique structures with biologically active properties. Many of these natural products were found to have cytotoxic properties but others harbor antibacterial, antifungal, antioxidant, antiviral, immunosuppressive, or anti-inflammatory activities (9, 13). These secondary metabolites are important for drug discovery --

over 40% of all approved drugs between the years 1981-2014 are natural products themselves or inspired by or derived from natural products in some fashion (15). The large majority of known natural products have yet to be tested for biological activities and genomics data also suggest that many more have yet to be discovered (10,11,14).

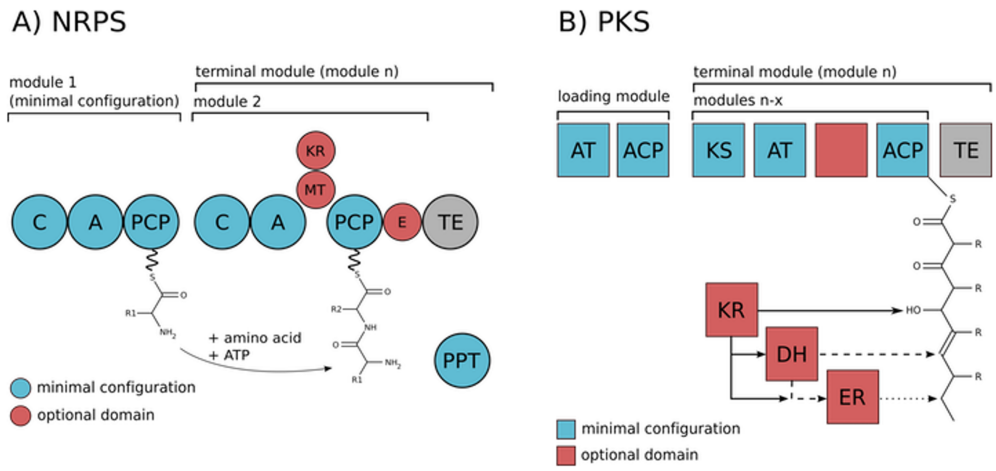


Figure 1.1 Schematic example of NRPS/PKS synthases. Image source: Kehr et al. (16). The domains not specified in the text are: KR, ketoreductase; MT, methyltransferase; E, esterification; DH, dehydrogenase; ER, enoyl reductase.

Most of these biologically active cyanobacterial natural products are synthesized from large biosynthetic enzymes called non-ribosomal peptide synthases (NRPS) and polyketide synthases (PKS), or a hybrid of the two (13) encoded by large biosynthetic gene clusters (BGCs). The non-ribosomal peptides (NRP) and polyketides derived from these megasynthases are built from sequential additions of amino acid or carboxylic acid monomers, respectively (17). The NRPS and PKS systems both represent modular biosynthetic enzymes, with different modules modifying an existing chain of non-ribosomal peptide or polyketide sequentially in an assembly-line fashion (17). Both NRPSs and Type I PKSs are organized as modules which are further

organized as domains. Each module represents an active site responsible for catalyzing a condensation and elongation step of the growing peptide or polyketide chain (**Figure 1.1**).

NRPSs accept both traditional and non-traditional amino acids and output a polypeptide. In the case of NRPSs, a module is responsible for adding a peptide unit to the growing chain. Each module minimally contains three domains: an activation (A), peptidyl carrier protein (PCP), and condensation (C) domain (17-20). The A domain recognizes the incoming amino acid monomer, activates it through a reaction with ATP, and transfers it to the PCP domain. The PCP domain is post-translationally modified to contain a phosphopantetheine group, which is responsible for transferring the growing chain to the next catalytic domain. Lastly, the C domain catalyzes the formation of peptide bonds between the growing peptide chain and the amino acid monomer (17-20).

Type I, or modular PKSs follow a similar scheme with their core domains being: an acetyltransferase (AT), acyl carrier protein (ACP), and ketoacyl synthase (KS) domain. The initial AT domain of the PKS recognizes starter units (Acetyl-CoA), and any subsequent AT domains will recognize extender units (Malonyl-CoA) and transfer them onto the ACP domain (21-24). The ACP domain is post-translationally modified to contain a phosphopantetheine (PPant) moiety, which loads the growing polyketide chain onto the extender unit. Finally, the KS domain catalyzes the decarboxylative condensation between the growing chain and the extender unit (21-24). In both NRPS and PKS systems, this sequential process can continue on for as many modules as exist, and additional optional domains are able to make specific modifications to the peptide or ketide backbone.

The post-translational modification of the PCP and ACP domains within NRPSs and PKSs via a 4'-phosphopanetetheinyl transferase (PPTase) activity is necessary for the activation

of these domains and therefore, for the production of their non-ribosomal peptides and polyketides (25-27). The PPTase works by catalyzing the addition of the 4' phosphopantetheine (PPant) arm onto the PCP or ACP domain to generate its active, *holo* form (25-27).

While cyanobacteria are prolific producers of natural products, not much is known about many of these compounds. Much more research is needed to investigate both the roles of cyanobacterial natural products in nature, as well as how humans may be able to adapt these compounds for medical use.

1.3 Marine Cyanobacterial Strains and their Associated Natural Products

Marine filamentous cyanobacterial strains are known as a rich source of bioactive metabolites (28). The genus *Moorea* is of particular interest because it is reported to be the source of over 40% of all known marine cyanobacterial natural products, yielding over 190 NPs in the past two decades (29, 30). These tropical, marine, metabolic "superproducers" have approximately 18% of their genome dedicated to secondary metabolites, with an average of 38 BGCs per genome, in comparison to 5 BGCs per genome for cyanobacteria in general (11, 29). Species within the genus *Moorea*, previously classified as *Lyngbya*, are characterized by thick filaments (25-65 μm) covered with a polysaccharide sheath and are commonly found in tropical oceans on coral reefs and rocks 0.3-30 m below the surface (31). They are photosynthetic organisms incapable of nitrogen fixation (31). Some of the species have red or purple-like color, due to the phycoerythrin they contain (31). *Moorea producens* JHB 22AUG96-1, *Moorea bouillonii* PNG 19MAY05-8, and *Moorea producens* PAL 15AUG08-1 (hereby referred to as JHB, PNG, and PAL, respectively) are three representatives of this genus, known for their

production of interesting secondary metabolites. These three strains are the native hosts of the three natural products this study is focused on.

One species of *Moorea,* JHB, has 10 biosynthetic pathways encoded in its genome, including those encoding for production of hectochlorin and jamaicamide (32). Originally isolated from Hector's Bay, Jamaica, JHB led to the discovery of a novel natural product after the culture was grown with ^{15}N-sodium nitrate and subsequently analyzed by Matrix-Assisted Laser Desorption/Ionization-Time of Flight (MALDI-TOF) mass spectrometry (33). The novel product, **cryptomaldamide**, was determined to be encoded by a 28.7 kb biosynthetic gene cluster (33). Interestingly, it is structurally similar to a sponge-derived natural product, hemiasterlin, which is known to have antimitotic properties (33).

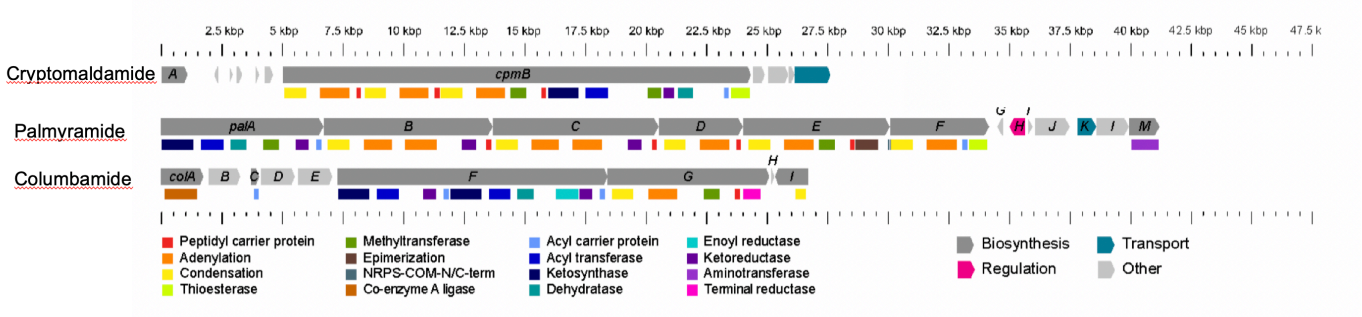


Figure 1.2 Schematic representation of cryptomaldamide, palmyramide, and columbamide biosynthetic gene clusters.

PNG was collected from Pigeon Island, Papua New Guinea (29, 32). Seven PKS/NRPS biosynthetic pathways are found within the PNG genome and PNG was found to produce **columbamides** A-C via Nuclear Magnetic Resonance (NMR) data (32). The columbamides are derived from a 27 kb BGC designated *col* (32). Columbamide A and B were found to be very strong ligands for the cannabinoid receptors CB_1 and CB_2, suggesting potential for use as a pain-relieving agent (32).

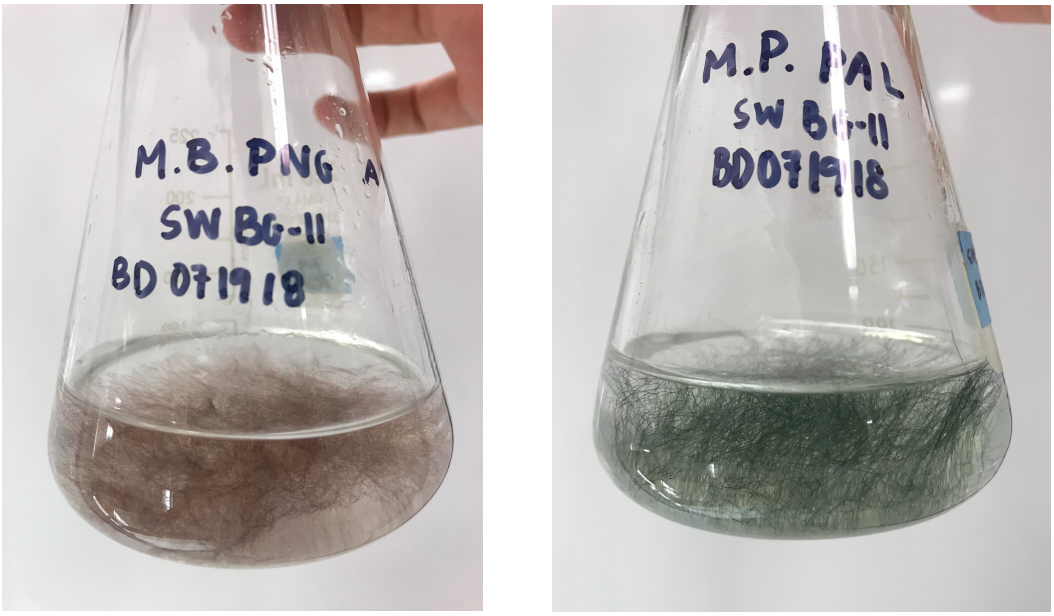
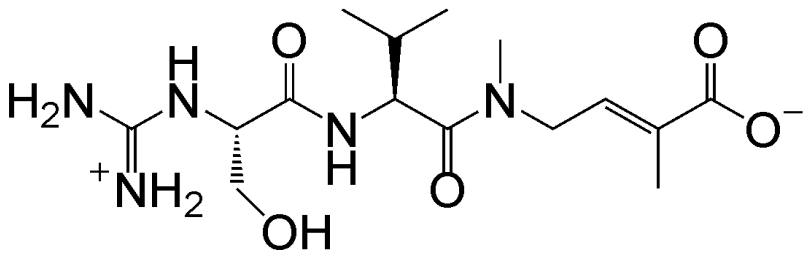


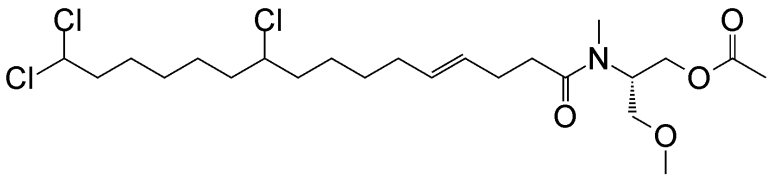
Image 1.1 *Moorea bouillonii* PNG (left) and *Moorea producens* PAL (right) as cultured in the laboratory.

Lastly, PAL was isolated from Palmyra Atoll in the Northern Pacific Ocean. This species is predicted to harbor seven orphan gene clusters, meaning they are unique to only its genome (29). PAL was found to produce **palmyramide** A, encoded by a BGC just over 40 kb in length (29).

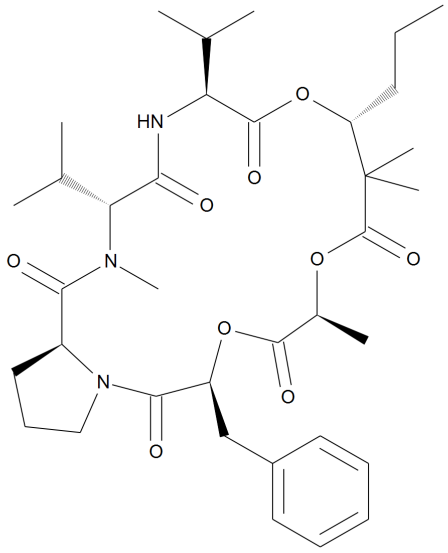
While these strains can be cultured in the lab (**Image 1.1**), they are very slow-growing, and their robust filaments complicate even relatively straightforward molecular biology protocols such as genomic DNA extraction. Furthermore, no synthetic biology tools currently exist in order to genetically manipulate these strains to study the BGCs or produce larger amounts of the compounds. As a result, further study of the BGCs native to these strains and their corresponding natural products requires heterologous expression in a more genetically compliant strain.



Cryptomaldamide



Columbamide B



Palmyramide A

Figure 1.3 Natural products from marine cyanobacteria of the genus *Moorea.*

1.4 Cyanobacterial Biotechnology

Genetic tools and synthetic biology techniques for cyanobacteria are most advanced in only a few select species, such as *Synechococcus elongatus* PCC 7942, *Anabaena sp.* PCC 7120, and *Synechocystis sp.* PCC 6803 (35, 36). Exogenous DNA can be transformed into cyanobacterial strains via natural transformation, conjugation, and electroporation with either an autonomous-replicating vector, or a vector designed for genome integration (35). Genomic integration and knock-in mutations require the identification of "neutral sites" (NS) within the chromosome of the target organism. Neutral sites are sites within the genome where exogenous DNA can be inserted without a known effect on phenotype, and several have been developed for the more commonly engineered cyanobacterial strains listed previously (35). Many genetic tools are available for these model strains. Such tools include promoters, transcription factors, and terminators for transcriptional control, as well as various systems involved in translational control, including characterized ribosome binding sites and riboswitches (34, 36). There is also some progress being made towards the standardization of these biological parts and tools, aiming for a streamlined system like those existing for more commonly used organisms for synthetic biology, such as *Escherichia coli, Saccharomyces cerevisiae* (yeast), and *Bacillus subtilis* (35). For example, there are now a number of online tools for the design of cyanobacterial vectors that help to streamline the design process for cyanobacterial vectors, as well as in silico predictive tools for regulatory sequences within the genome of cyanobacterial model organisms (35, 38). Recently, there have also been many efforts to take advantage of cyanobacterial photosynthesis to efficiently express high-value products (39).

Previous work done by Roulet, et al. helped to lay the foundation for developing *S. elongatus* as a heterologous expression platform strain for PKS-derived compounds (40). This

work resulted in the development of a module containing a tightly regulated and theophylline-inducible synthetic riboswitch controlling the translation of the promiscuous Sfp-PPTase from *B. subtilis* (26, 27, 36, 40). This system was successful in activating the heterologous PKS for producing multimethyl-branched esters (40).

Further expanding on the development of new tools for the use of cyanobacteria as heterologous hosts, the yeast Transformation-Associated Recombination (TAR) cloning strategy from Kouprina et al. (41) has been adapted for cyanobacteria in our laboratory (Brooke Anderson and Arnaud Taton) with the help of Bradley Moore's laboratory at the Scripps Institute of Oceanography. The TAR cloning strategy utilizes the natural homologous recombination machinery of yeast to capture large DNA fragments, like BGCs, into plasmid vectors and enable the transfer of these BGCs to other organisms.

New TAR destination vectors (pAM5570 and pAM5571), shown in **Figure 1.4,** were designed to capture large gene clusters from strains that are not genetically tractable and enable the recombination of these gene clusters into *S. elongatus.* Our TAR cloning vectors contain several parts that enable their replication in yeast and *E. coli,* and their recombination into the *S. elongatus* chromosome. Yeast elements include the yeast origin of replication and centromere (CEN6_ARS4), which is necessary to allow the plasmid to replicate in yeast, and a tryptophan synthase (*trpI*), to enable plasmid selection in tryptophan drop-out medium (containing all standard amino acids excluding tryptophan). The *ura3* gene is also present; it is used as a counterselection mechanism for identifying TAR capture vectors with inserts greater than 130 bp because *ura3* generates the toxic compound 5-fluorouracil in the presence of 5-fluoroorotic acid (5-FOA), causing cell death. The plasmid also contains an *E. coli* origin of replication (ori-pBR322), an *E. coli* origin of transfer (oriT), and an antibiotic selection marker that works in *E.*

coli and cyanobacteria (*aphI* or *aadA*, which provide resistance to kanamycin or spectinomycin and streptomycin, respectively). Lastly, these plasmids contain left and right "arms", which are homologous to a neutral recombination site in the *S. elongatus* chromosome (NS2).

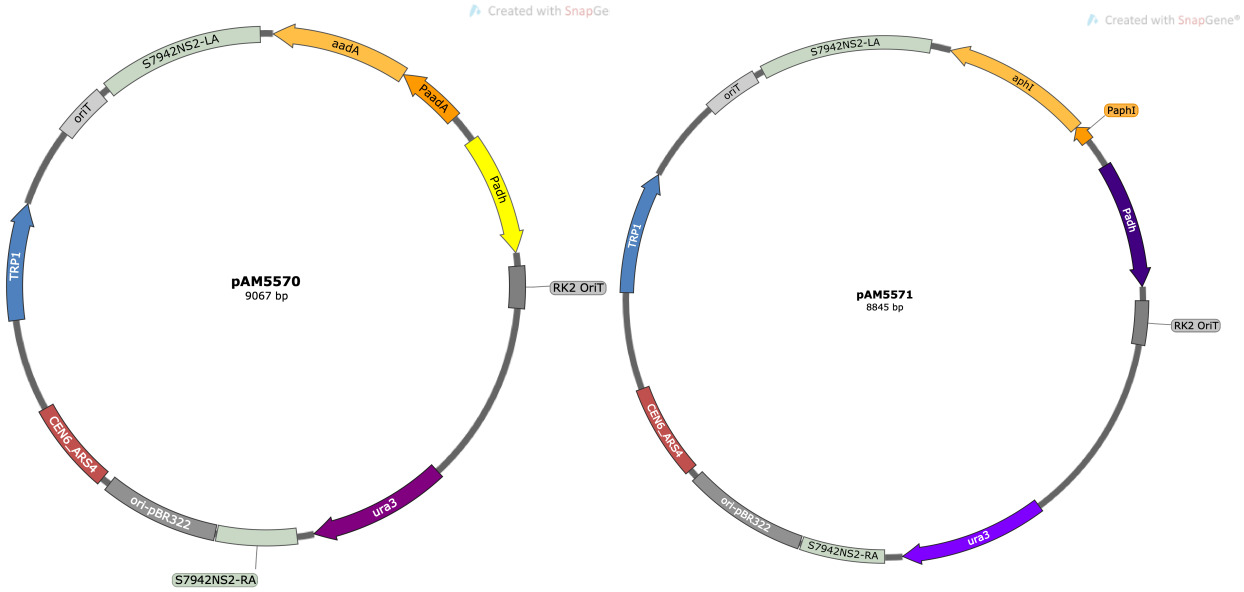


Figure 1.4 Plasmid maps of TAR capture vectors designed to integrate into the *S. elongatus* chromosome.

A recent study suggested that *Anabaena* is a suitable host for the expression of natural product pathways from marine cyanobacteria. It demonstrated the production of Lyngbyatoxin A from a small NRPS gene cluster from *Moorea* and confirmed the activity of several *Moorea* promoters in *Anabaena* (42). With this in mind, our lab developed another recent advancement to the heterologous expression strategy – the creation of an *Anabaena* strain engineered to accept plasmid constructs designed for *S. elongatus* in order to avoid having to re-clone BGC pathways in a separate destination vector designed specifically for recombination into *Anabaena*'s chromosome. This was done by integrating the *S. elongatus* neutral site 2, abbreviated as NS2

(43-44), into an *Anabaena* neutral site (45-48) via conjugation of pAM5569 into the WT strain of *Anabaena*. This plasmid backbone (like many that are used for engineering the *Anabaena* chromosome) contains *sacB*, which is used for selection of double recombinants in the presence of 5% sucrose (48). This counterselection technique is adopted in genetic engineering with *Anabaena* because in comparison to *S. elongatus*, chromosomal segregation is more difficult to achieve. Any cells carrying a chromosome that has undergone a single recombination event with the plasmid will be counter-selected when plated on BG-11 plates supplemented with sucrose. A nourseothricin (NT) cassette was placed within the *S. elongatus* NS2, which enables the differentiation of single recombinants and double recombinants after integration of the target BGC. This engineered *Anabaena* strain is now part of the Golden Lab Collection (AMC2556) and is available for this study. AMC2556 was constructed in collaboration with Diego B. Genuário and Christopher Peterson. The creation of this strain is a significant development because it eliminates the need to create a capture vector specified for *Anabaena* and thus allows the engineering of BGCs into two different expression hosts while only having to capture the BGC once, in a single vector.

In recent years, CRISPR-Cas9 technology has revolutionized synthetic biology by enabling precise, markerless mutations of DNA sequences in many species. Initially, the Cas9 nuclease was found to be toxic to cyanobacteria, rendering this technology ineffective for cyanobacterial genetics, but a CRISPR system using the Cpf1 nuclease was shown to work very efficiently with no toxicity in several strains of cyanobacteria (50, 51). All of these synthetic biology tools can be integrated to allow the cloning and manipulation of large BGCs in cyanobacteria like *S. elongats* and *Anabaena*.

1.5 Objectives

Our understanding of cyanobacterial BGCs and natural products remains limited because most known bioactive compound producers are not genetically tractable and tend to grow slowly compared to the model organisms used for genetics experiments. In addition, BGCs from cyanobacteria are often quite large, which complicates their study in heterologous hosts.

The overall goal of this project is to enable the biosynthesis of selected natural products from marine filamentous cyanobacteria in a suitable cyanobacterium host through the heterologous expression of their respective BGCs. This would represent a significant opportunity to produce these valuable compounds at larger scale and allow for their further investigation. Using cyanobacteria to express pathways isolated from other cyanobacterial strains is more likely to succeed because of their genetic relatedness and this was found to be the case for the expression of polybrominated diphenyl ether biosynthesis genes from cyanobacterial metagenomic sequences (52). In addition, using cyanobacteria as a platform for the production of biochemicals and natural products is attractive because only carbon dioxide, water, and light are needed to power their metabolism and biosynthesis.

Chapter Two: Materials and Methods

2.1 Strains and Culture Conditions

All strains used in this study are listed in **Table 2.1**. Marine filamentous cyanobacterial strains were grown in SW-BG-11 (BG-11 with 33 g/L Instant Ocean Aquarium Salt) as 100-mL cultures at 25-30°C at a light intensity of 30-40 μmol photons m^{-2} s^{-1}. Both wild-type (WT) *S. elongatus* and engineered strains were grown in BG-11 liquid medium as 100-mL cultures or on agarose plates (40 mL BG-11 media solidified with 1.5% agar), at 30°C under continuous light with intensity ranging from 15-110 μmol photons m^{-2} s^{-1}. WT and engineered strains of *Anabaena* were grown in similar conditions but were kept at approximately 25-70 μmol photons m^{-2} s^{-1} on a shaker. Appropriate antibiotics were added to the media when necessary for the engineered cyanobacterial strains.

S. cerevisiae VL6-48N was either grown on YPD agar plates and incubated at 30°C in the dark for 2-3 days, or isolated colonies were grown in 5 mL YPD liquid media and incubated overnight at 30°C. Liquid cultures were kept on a roller drum or on a shaker during overnight incubations.

E. coli strains were grown on LB agar plates or in LB liquid medium (5 mL cultures) with appropriate antibiotics and incubated at 37°C. Liquid cultures in culture tubes were grown on a roller drum.

Table 2.1 List of strains used in this research. Plasmids are described in section 2.4.

Strain	Description	Antibiotic Resistance	Source/Creator
VL6-48N	*S. cerevisiae*, cloning host		Yamanaka, et al. (53), from Bradley Moore Lab
DH5α	*E. coli*, chemically competent cloning host		Golden Lab Collection
DH10β	*E. coli*, electrocompetent cloning host		Golden Lab Collection
EPI400	*E. coli*, electrocompetent cloning host; optimized for cloning unstable DNA		Golden Lab Collection
NEB Stable	*E. coli*, electrocompetent cloning host; optimized for cloning repeat elements and unstable inserts		Golden Lab Collection
DB3.1	*E. coli,* electrocompetent cloning host		Golden Lab Collection
AM1356	*E. coli,* DH10β possessing pAM4413, pRL623, and pRL443 for conjugation		Golden Lab Collection
AM5413	*E. coli*, carrying pCRCT plasmid	Ap	Bao, et al. (54)
AM5570	*E. coli,* carrying pAM5570	SpSm	Golden Lab Collection/B. Anderson & A. Taton
AM5571	*E. coli,* carrying pAM5571	Km/Nm	Golden Lab Collection/B. Anderson & A. Taton
AM5569	*E. coli*, carrying pAM5569		Golden Lab Collection/D.B. Genuário, A. Taton & C. Peterson
AM5572	*E. coli,* DH5α carrying pAM5572	SpSm	Golden Lab Collection/T. Gilderman & A. Taton
AM5558	*E. coli,* DH5α carrying pAM5558	Km/Nm	This study
AM5559	*E. coli,* EPI400 carrying pAM5559	SpSm	This study
AM5560	*E. coli,* DH5α carrying pAM5560	Km	This study
AM5561	*E. coli,* AM1359 carrying pAM556	Km/Nm, Ap, Tc, Cm	This study
AM5562	*E. coli,* D3.1 carrying pAM5562	Km/Nm	This study
AM5563	*E. coli,* AM1359 carrying pAM5564	Km/Nm, Ap, Tc, Cm	This study
AM5564	*E. coli,* DB3.1 carrying pAM5564	Km/Nm	This study
AM5565	*E. coli*, DH5α carrying pAM5565	SpSm	This study
AM5566	*E. coli,* AM1359 carrying pAM5565	SpSm, Ap, Cm, Tc	This study
AM5567	*E. coli,* DH5α carrying pAM5567	Ap	This study
AM5568	*E. coli*, DH5α carrying pAM5568	Ap	This study
AMC2302	S. elongatus PCC 7942 WT (AMC06) cured of its small plasmid pANS		Chen et al. 2016 (55)
AMC2566	AMC2302 carrying an *sfp* PPTase from *B. subtilis* under riboswitch control in 7942 NS3	Cm	Golden Laboratory collection/ A. Taton
AMC2556	*Anabaena sp.* 7120, engineered to contain *S. elongatus* neutral site within native neutral site 2	Nm	Golden Laboratory collection/ A. Taton, D. Genuário, C. Peterson
M. bouillonii JHB	Marine cyanobacteria, native host of cryptomaldamide		Gerwick Laboratory collection
M. bouillonii PNG	Marine cyanobacteria, native host of columbamide		Gerwick Laboratory collection
M. producens PAL	Marine cyanobacteria, native host of palmyramide		Gerwick Laboratory collection

Table 2.1 List of strains used in this research. Plasmids are described in section 2.4. (continued)

Strain	Description	Antibiotic Resistance	Source/Creator
AMC2557	AMC2556 carrying columbamide pathway - clone 9	Nm	This study
AMC2558	AMC2556 carrying columbamide pathway - clone 10	Nm	This study
AMC2559	AMC2556 carrying columbamide pathway - clone 11	Nm	This study
AMC2560	AMC2556 carrying cryptomaldamide pathway – clone 1	Nm	This study
AMC2561	AMC2556 carrying cryptomaldamide pathway – clone 2	Nm	This study
AMC2562	AMC2556 carrying cryptomaldamide pathway – clone 3	Nm	This study
AMC2563	AMC2560 conjugated with pAM5565 to force segregation– clone 1	Nm, SpSm	This study
AMC2564	AMC2560 conjugated with pAM5565 to force segregation– clone 2	Nm, SpSm	This study
AMC2565	AMC2560 conjugated with pAM5565 to force segregation– clone 3	Nm, SpSm	This study

2.2 Genomic and Plasmid DNA Extraction Protocols

Genomic DNA was isolated from *M. producens* PAL, *M. bouillonii* PNG, and *M. bouillonii* JHB using the Qiagen Genomic-tip 20/G kit with modifications for the sample preparation and lysis protocol for the filamentous marine cyanobacteria. A 1.5 cm^3 clump of the filamentous cells was isolated with sterile forceps, washed with sterile water, dried on a filter, flash frozen with liquid nitrogen, and subsequently ground with a pre-chilled mortar and pestle. The resulting powder was then used as starting extraction material for the Qiagen Genomic-tip 20/G kit following the manufacturer's instructions.

Genomic DNA was extracted from engineered *S. elongatus* and *Anabaena* using a phenol-chloroform extraction with cetyltrimethylammonium bromide (CTAB) (56). The protocol is as follows: 10 mL of culture was centrifuged and resuspended in 1000 µl of BG-11. The cells were pelleted with another 3-minute centrifugation at 16,000 g and resuspended in 340 µl STET buffer (8% sucrose, 5% Triton X-100, 50 mM EDTA [pH 8.0]). Lysozyme and RNase A were added to a final concentration of 2 mg/mL and 0.1 mg/mL respectively, and the cell suspension was incubated at 37°C for 45 minutes. After the incubation, 2 µl of a 10 mg/mL proteinase K

stock solution and 20 µl of 20% sarkosyl was added, the mixture was vortexed for 20 seconds, and then incubated at 55°C for 30 minutes. Afterwards, 57 µl of 5 M NaCl and 45 µl of 10% CTAB in 0.7 M NaCl was added, the mixture was vortexed again, and incubated at 65°C for 10 minutes. The sample was then extracted twice with 500 µl 24:1 chloroform:isoamyl alcohol. After vortexing the solution mixture for 10 seconds and centrifugation at 16,000 g for 10 minutes, the upper aqueous layer was removed and extracted once more with the same procedure. The resulting aqueous phase was removed and precipitated with 2 volumes of 100% ethanol. The DNA was pelleted with a centrifugation (16,000 g for 10 minutes) and washed with 1 mL of 70% ethanol before evaporation and resuspension in 50 µl sterile water.

Plasmid DNA from *E. coli* cells was extracted from 5-mL overnight cultures using either an ethanol precipitation or the Qiagen QIAprep Spin Miniprep Kit, according to the manufacturer's instructions.

A phenol-chloroform extraction method was used to isolate plasmid DNA from *S. cerevisiae* (53). After growing overnight on a shaker, a 5-mL culture was pelleted via centrifugation at 16,000 g and resuspended in 200 µl of a resuspension solution (10% sucrose, 50 mM Tris-HCl (pH 8.0), 10 mM EDTA) containing 0.5 mg/mL zymolyase-20T (MP Biomedicals) and 0.2% BME. The suspension was incubated on a roller drum for 2 hours at 37°C. RNase was added to a final concentration of 100 µg/mL and the suspension was mixed with 400 µl 0.2 M NaOH and 1% SDS (w/v) by inverting the tube several times. 300 µl of a buffer solution containing 3 M KAc (pH 8.0) was added and the solution was centrifuged for 10 minutes at 16,000 g. The resulting supernatant was extracted with 1 volume of phenol-chloroform, vortexed, and centrifuged for 2 minutes at 4°C at 16,000 g. The DNA was precipitated from the upper aqueous layer by the addition of 1 volume isopropanol and

centrifugation at 16,000 g for 5 minutes. The DNA was washed twice with 70% ethanol and then dissolved in 50 µl water after drying overnight. 1-3 µl of this extraction was used in PCR reactions and 5 µl was used in *E. coli* transformations.

2.3 General PCR Protocols

PCR amplifications were performed with Q5 High-fidelity polymerase from New England BioLabs (NEB) according to the manufacturer's instructions. A typical PCR reaction used the following cycles: an initial denaturation of 30 seconds at 98°C; 30 cycles of 10 seconds at 98°C, 30 seconds at the recommended annealing temperature (65-70°C), and 30 seconds/kb at 72°C; and at the end, a final extension period of 2 minutes at 72°C. 5-kb fragments of each BGC were amplified using Q5 polymerase. Some PCR screens were done using Taq Polymerase from NEB; thermocycler cycles for Taq Polymerase were as follows: 30 second initial denaturation at 95°C; 30 cycles of 30 seconds at 95°C, 60 seconds at optimal annealing temperature (45-68°C), 1 minute/kb extension at 68°C; and a final extension period of 5 minutes at 68°C. PCR products were cleaned and concentrated using the Zymo Research DNA Clean and Concentrator Kit for downstream use. Primers used in this research are listed in **Table 2.2**.

Table 2.2 List of primer sequences used in this study

Name	Sequence (5' to 3')	Purpose
colA_U739F	TGCATCAAGACATTGAACATTATCGTC	amplify columbamide pathway
colD_1389R	CGCATGGTATAGTATTGGCAAGAAAAG	amplify columbamide pathway
colC_326F	TTTAGGCAATGAACTCAAAGGAGAGG	amplify columbamide pathway
colF_2644R	GGAAGGGATATGTCGGTAAGGCTAC	amplify columbamide pathway
colF_1786F	CATCTATTTTAGACCAAACCGCCTATAC	amplify columbamide pathway
colF_7685R	ACTTGTTTACCATTTTTTGAAACCCC	amplify columbamide pathway
colF_6723F	GTAGTCATTTCAGGGGACTCTATTGCC	amplify columbamide pathway
colG_1518R	AGGAAGTTGTCCTACTCTCTGCTCTGG	amplify columbamide pathway
colG_601F	TGATTATCAAAGAACTGCCACAACTC	amplify columbamide pathway
colG_6519R	TGAAAAAAGCGACAAAAGAGGATAC	amplify columbamide pathway
colG_5792F	TCCACTATTGGGTCTTGACGAGTC	amplify columbamide pathway
colI_U893R	CTATCCTGTCCATGGCAGTAGTTCC	amplify columbamide pathway
colA_U790F_H	GCCTCCCATGGTATAAATAGTGGCtgcatcaagacattgaacattatcgtc	create hooks for columbamide pathway
colA_U74R_H	gagatagtttaaactatacttcggtgtgcttatttttgagc	create hooks for columbamide pathway
colI_U926R_H	GCAGCACGTTCCTTATATGTAGCTTTCGACATActatcctgtccatggcagtagttcc	create hooks for columbamide pathway
colI_48F_H	agtatagtttaaactatctcctctttgacctctgttccttc	create hooks for columbamide pathway
palA_U975F	AAACGTAAAATCAGTGTTTCTTGTGC	amplify palmyramide pathway
palA_6705R	AATACTTCCTGGTTATTTCCCTCTACTC	amplify palmyramide pathway
palA_4098F	CCTCAAGGCAAGCCAGAACTTACC	amplify palmyramide pathway
palB_3287R	GTTAACGTTCTGGGAGCATTTTCTGG	amplify palmyramide pathway
palB_2929F	CTTATGTTCCCCTGGCTCGCTTAC	amplify palmyramide pathway
palC_1452R	TGTTGCTGTTATTAGGGGGAGTTTTG	amplify palmyramide pathway
palC_874F	TTATGACTCTGTTGAGTGCATTTTTTG	amplify palmyramide pathway
palC_6537R	ATTCCTGGGCTTGACTCTTTCTTG	amplify palmyramide pathway
palC_5931F	GCTGAGGTGTTACGTCCTAAACTGATAG	amplify palmyramide pathway
palE_1331R	CTCCTTCGTACCAACCCACTTCTAAG	amplify palmyramide pathway
palE_692F	TTACACTGCTACACTGGAAGAAAAACC	amplify palmyramide pathway
palF_103R	CTTGCATAGGAGAAAGGGGATAGAG	amplify palmyramide pathway
palF_U321F	ACCTAGGTCAGTTTAAGAGCAGTCAATC	amplify palmyramide pathway
palH_325R	AATGGCAACGTAAAGCTCAAGAATAC	amplify palmyramide pathway
palG_186F	GATCTGCTGATGGGTCAACTCTTC	amplify palmyramide pathway
palL_461R	GTCTCCGTCAATCCCACATACTTTAG	amplify palmyramide pathway
palK_996F	CATGCGGTGTTTAATGGTAAGGTTTC	amplify palmyramide pathway
palM_D1385R	CAAGAGGACAAGAGTTAATAGAGCAGCG	amplify palmyramide pathway
palA_U975F_ALH	TGGCGAACGAGCGCAAGGTTTCGGTCTCCACGCAaaacgtaaaatcagtgtttcttgtgc	create hooks for palmyramide pathway
palA_U147R_ALH	ggccaaatgaatcaccccatccagttgtttaaacaattataaaattcatctgctcctggg	create hooks for palmyramide pathway
palC_5857F_ARH	acacccaggagcagatgaattttataattgtttaaacaactggatggggtgattcatttg	create hooks for palmyramide pathway
palC_6537R_ARH	GAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTattcctgggcttgactctttcttg	create hooks for palmyramide pathway
palC_5931F_BLH	TGACGCCTCCCATGGTATAAATAGTGGCgctgaggtgttacgtcctaaactgatag	create hooks for palmyramide pathway
palC_6708R_BLH	ctccaccacatgtaggtttaaacttccatctcaaaagcctgtcgtaag	create hooks for palmyramide pathway
palM_414F_BRH	cgacaggcttttgagatggaagtttaaacctacatgtggtggagttaaaaaatcagc	create hooks for palmyramide pathway
palM_D1385R_BRH	CAGCACGTTCCTTATATGTAGCTTTCGACATAaagaggacaagagttaatagagcagcg	create hooks for palmyramide pathway
palC_6630F	aatttctttgagttgggcggtaattctc	sequence palmyramide destination vector
ColA_D390F	CTGGTGCCGCGCGGCAGCCAatgaccgagttaaccagtcgctttgc	left arm hook for col under pT7 plasmid
ColA_D1131R	cgacacttccccgtttaaacaaccttacggacacctgtaacctcgc	left arm hook for col under pT7 plasmid
ColI_D818F	gtccgtaaggttgtttaaacggggaagtgtcggtaagttaggc	right arm hook for col under pT7 plasmid
ColI_D1377R	TGCGGCCGCAAGCTTGTCGAttgcaagaacccagaattgaggaag	right arm hook for col under pT7 plasmid
pT7col_repair_F	CCCATGGATTTTGAGACACAACG	PCR homologous repair template from pT7_pcvd015
pT7col_repair_R	GCGATTTTAGCAGGCTGAGTTAGG	PCR homologous repair template from pT7_pcvd015
pT7col_pcvd015_448F	GAAGTGAAAAATGGCGCACATTGTG	PCR gRNA from pT7col_pcvd015
pT7col_pcvd015_618R	GGTCCACGTTGTGTCTCAAAATCC	PCR gRNA from pT7col_pcvd016
pCRCT_colgRNA_1F	CAAAACttgaaaaacgctatctaccaG	screen for correct ligation of col_gRNA into pCRCT
pCRCT_colgRNA_D529R	cgctatttgtgccgatatctaagcc	screen for correct ligation of col_gRNA into pCRCT
NT_U248F	agatGACAGCTTATCATCGAATTA	Oligos for CRISPR gRNA to target NT gene in 7120_NS2
NT_U228R	agacTAATTCGATGATAAGCTGTC	Oligos for CRISPR gRNA to target NT gene in 7120_NS3
pconII-RSwB-T7pol_pAM5470_F-pcvd015	ggccaataacccagggatttTAGAAAGCTTCAAAAAGGCCATCCG	amplify pconII, RSwB, and T7 pol from pAM5470
pconII-RSwB-T7pol_pAM5470_R	cgaaaccggtaGAATTGACAATTAATCATCGGCGCG	amplify pconII, RSwB, and T7 pol from pAM5470
pT7-yfp-pAM5467_F	ttgtcaattcTACCGGTTTCGAATTGAGATTGACG	amplify pT7 and yfp from pAM5467
pT7-yfp-pAM5467_R-pcvd015	gccggggagctccttcatttTAGAAAGCTTCAAAAAGGCCATCCG	amplify pT7 and yfp from pAM5467

2.4 Plasmid Construction Methods

Plasmids and strains used and created in this study are listed in **Table 2.3**. Plasmids were constructed *in silico* using the CYANO-VECTOR system and assembled as published previously (38). Insert DNA fragments were obtained by PCR and subsequently assembled in a linearized vector backbone.

Restriction enzyme digests were performed for 3-5 hours, using 5U of enzyme per μg of DNA in a volume at least 50 times greater than that of the volume of enzyme used. PCR and digest samples were cleaned and concentrated using the Zymo Clean and Concentrator TM-5 spin column kit according to the manufacturer's instructions and DNA concentrations were measured with a UV-Vis spectrophotometer Nanodrop 2000c before being used to construct plasmids. Most plasmids were constructed using the GeneArt Seamless Cloning and Assembly Kit from Life Technologies, according to the manufacturer's instructions. Golden Gate cloning was used with BsaI-HF and T4 Ligase from New England BioLabs to clone guide RNA template in CRISPR plasmids (pAM5572 and pCRCT). The protocol followed that of the manufacturer's instructions. Two methods were used to simultaneously assemble plasmids and transform *S. cerevisiae*: Transformation Associated Recombination (TAR) cloning, and a LiAc yeast transformation protocol. See section 2.5 for full details on these protocols.

Table 2.3 List of plasmids used in this study.

Plasmid	Description	Source/Creator
pRL623	Helper plasmid containing Mob(ColK) and *Anabaena* methylases M.AvaI, MEco47II, and M.EcoT22I; Cm resistance	Elhai, et al. (57)
pRL443	Conjugal plasmid; Ap and Tc resistance	Elhai, et al. (57)
pCRCT	Plasmid encoding Cas9, tracrRNA and crRNAs; used for CRISPR in yeast; URA3 selection in yeast	Bao, et al. (54)
pAM5570	Yeast capture vector for recombination in *S. elongatus;* SpSm resistance S7942NS2RK2B-Y_AADA URA3-CAPT-S	Golden Lab Collection/B. Anderson & A. Taton
pAM5571	Yeast capture vector for recombination in *S. elongatus;* Km resistance S7942NS2RK2B-Y_APHI URA3-CAPT-S	Golden Lab Collection/B. Anderson & A. Taton
pAM5558	Yeast capture vector with columbamide hooks; Km resistance	This study
pAM5559	Yeast capture vector with palmyramide part A hooks; SpSm resistance	This study
pAM5560	Yeast capture vector with palmyramide part B hooks; Km resistance	This study
pAM5562	Columbamide (#45) BGC captured in TAR vector (pAM5558); sequence confirmed; Km resistance	This study
pAM5564	Cryptomaldamide (#19) BGC captured in TAR vector; Km resistance	This study
pAM5565	CRISPR/Cpf1 plasmid targeting NT cassette within AMC2556 chromosome; SpSm resistance	This study
pAM5567	gBlock of T7 system cloned into pcvd015 for preservation; Ap resistance	This study
pAM5568	pCRCT with gRNA targeting *colA* for promoter refactoring; URA3 for selection in yeast	This study
pAM5569	*Anabaena* 7120 per015 with S7942NS2 carrying Nt cassette; Nt resistance	Golden Lab Collection/D.B. Genuário, C. Peterson & A. Taton
pAM5572	CRISPR/Cpf1 system in RSF1010 backbone; SpSm resistance	Golden Lab Collection/T. Gilderman & A. Taton

2.5 Transformation in *S. cerevisiae*

Two transformation methods were used for *S. cerevisiae.* Transformation-Associated Recombination (TAR) cloning was used to assemble the 5-kb PCR fragments of each BGC into its respective destination vector following the protocol published by Kouprina and Larionov with some slight modifications from Bradley Moore's laboratory (Scripps Institute of Oceanography) and Brooke Anderson (41, 53).

The *S. cerevisiae* VL6-48N strain was grown on YPD plates containing 1% yeast extract, 2% peptone, 2% agar, and supplemented with 100 mg/L adenine. To begin the TAR cloning procedure, a single yeast colony was grown in 5 mL YPD liquid medium and incubated overnight at 30°C on a roller drum. The following day, the 5 mL overnight culture was transferred to a 125-mL flask containing 35 mL YPD media and grown at 30°C in a shaker at 200 rpm to an OD_{600} of 0.7-1.0. The flask was placed on ice for 10 minutes and the cell suspension centrifuged for 3 minutes at 1,800 g at 4°C to pellet the cells. The cells were washed twice with 25 mL ice cold sterile water, centrifuging as before. After decanting the water, cells were then resuspended in 25 mL of ice cold 1 M sorbitol by vortexing. Cells were left covered, on ice at 4°C overnight for osmotic stabilization. The next day, cells were inverted to resuspend and centrifuged again for 3 minutes at 1,800 g at 4°C. The supernatant was removed and cells were resuspended in 20 mL SPE solution (1 M sorbitol, 5 µM HEPES pH 7.5, 1 µM EDTA pH 8.0) by vortexing. 20 µl β-mercaptoethanol and 40 µl zymolyase 20T (10 mg/mL zymolyase in 50 mM Tris buffer pH 7.5, 50% glycerol) for every 25 mL of total initial yeast culture was added to the mix and incubated for 40-50 minutes at 30°C with gentle shaking. The zymolyase enzyme should be titrated prior to the spheroplast process to determine the appropriate concentration to use; in the TAR experiments performed in this study, a concentration of 50 mg/mL was used to achieve the desired level of spheroplasts within 50 minutes of incubation. The optical density (OD_{600}) was monitored throughout the incubation to assess the spheroplasting progress. For this, at each timepoint, the cells were diluted in 1 M sorbitol or 2% SDS and the OD_{600} were measured. Spheroplasts are considered ready when there is a 5-20 fold difference between these measurements. After cells have become spheroplasts, they are extremely fragile so the remainder of the protocol following this point was carried out on ice. Chilled (4°C) 1 M sorbitol was added

to the spheroplast mixture up to 25 mL, the solution was mixed by inversion, and then centrifuged at 4°C for 10 minutes at 600 g. The supernatant was poured off, the cells were gently resuspended in 10 mL 1 M sorbitol that had been chilled at 4°C, and the final volume was raised up to 25 mL using the same sorbitol solution. The cell suspension was centrifuged for 10 minutes at 600 g and the supernatant was removed. The cells were resuspended in 1 mL room temperature STC solution (10% $CaCl_2$, 1% Tris buffer, 1 M sorbitol) using a serological pipette. The spheroplasts were incubated at room temperature for 10 minutes before adding 200 µl to tubes containing 100-500 ng of the transforming DNA. The cell and DNA mixture was then incubated at room temperature for 10 minutes before adding 800 µl 20% PEG solution, pH 7.5 (10% $CaCl_2$, 1% Tris buffer, 20% PEG800), inverting the tube 10 times to mix. The mixture was incubated for 20 minutes at room temperature and the cell suspension was centrifuged at 4°C for 10 minutes at 700 g. The supernatant was removed with a pipette and the cells were resuspended in 800 µl of SOS solution (6.5% $CaCl_2$, 0.25% yeast extract, 1% peptone, 1 M sorbitol) by pipetting with a wide pipette tip. The mixture was incubated at 30°C for 30-40 minutes, mixed with melted top-selective agar (1.1 M sorbitol, 3% agar, 2.2% glucose, 100 mg/L adenine hemisulfate, 1X N-base solution), and then poured onto selective bottom agar plates (1.1 M sorbitol, 2% agar, 2.2% glucose, 100 mg/L adenine hemisulfate, 1X N-base solution). 10X N-base solution contains 1.7 g yeast nitrogen base without amino acids, 1.9 g yeast synthetic drop-out medium supplements without the amino acid being used for prototrophic selection (tryptophan if using TRP1 or uracil if using URA3 as the selection marker) and 4.5 g ammonium sulfate in 100 mL ddH_2O. The plates were then incubated in the dark for 2-3 days at 30°C to allow colony growth.

Alternatively, a lithium-acetate transformation protocol with a single-stranded carrier DNA published by Gietz and Schiestl was used when transforming smaller plasmids (<15 kb) into yeast (58).

2.6 Transfer of recombinant DNA in cyanobacteria

Recombinant DNA was transferred in *S. elongatus* by natural transformation (43,59,60). 5 mL of culture was centrifuged at 4500 g and the cell pellet was resuspended in 1.5 mL 10 mM NaCl. The cells were centrifuged, the solution was decanted, and the pellet was resuspended in 200 µl BG-11. 50 ng-1 µg transforming DNA (1-2 µg for plasmids greater than 15 kb) was added to the cells and the mixture was left to incubate 4-18 hours at 30°C on a shaker. The cells were plated on selective BG-11 agar plates and the plates were incubated at 30°C in continuous light at 150-300 µmol photons $m^{-2}s^{-1}$ for 4-6 days to allow colony growth.

Recombinant DNA was transferred in *Anabaena* by biparental conjugation (57,61,62). Conjugal donor *E. coli* strain AM1359 was transformed with the shuttle vector of interest using electroporation. The transformed cells were grown in 2 x 5-mL LB with antibiotics at 37°C on a roller drum overnight. About 50 mL of a one-month *Anabaena* culture (the cyanobacterial cells can be in any growth phase from an early actively growing culture to an older stationary phase culture) was fragmented by sonication (50 seconds at 20% amplitude, 5 sec on/off cycles), the fragmented cells were sedimented by centrifugation for 3 minutes at 4000 g and resuspended in fresh BG-11 to recover overnight. The *E. coli* cultures were then split into 5 x 2-mL aliquots (each 2-mL culture being used for 1 conjugation). *E. coli* cells were centrifuged for 3 minutes at 4000 g, washed twice with LB to remove the antibiotics, and then concentrated in 200 µl of LB, being careful not to vortex the cells to ensure the conjugal pili remain intact. 50 mL of

fragmented *Anabaena* was centrifuged, washed with BG-11, and resuspended in 10 mL of BG-11. The *E. coli* and *Anabaena* strains were mixed at a ratio of 1 mL concentrated *Anabaena* to 200 µl concentrated *E. coli* cells. The mixture of cells was centrifuged for 5 minutes at 4000 g to collect the cells, which were then resuspended in 200 µl of BG-11 and plated on BG-11 plates containing 5% LB without antibiotics. The plates were incubated at 30°C with 10-20 µmol photons $m^{-2}s^{-1}$ of light. After 2-3 days of growth, cells were scraped from the plates, resuspended in 1 mL of BG-11, and then plated on several BG-11 agar plates with selective antibiotics, transferred to higher light (50-100 µmol photons $m^{-2}s^{-1}$), and incubated at 30°C with high humidity for up to two weeks. Individual exconjugant colonies were picked and streaked on BG-11 agar containing appropriate antibiotics to promote segregation.

Chapter Three: Results

3.1 Capture of Biosynthetic Gene Clusters

The general strategy to transfer each NP BGC pathway to a genetically tractable cyanobacterial strain followed a similar scheme. Initially, the pathway was captured and assembled into a plasmid construct in *S. cerevisiae* utilizing the cell's native homologous recombination machinery. Because the yeast cells possess only one copy of this plasmid, the recombinant plasmid was extracted from a yeast culture and then transformed into *E. coli* cells where it replicates at high copy number. Plasmids are then extracted from *E. coli* for screening for the correct sequence by restriction digests and sequencing. Finally, the confirmed plasmids are transferred into a genetically tractable cyanobacterial strain for production of the NP.

3.1.1 Cryptomaldamide

In previous work done by Arnaud Taton and Nathan Moss (Gerwick Lab, SIO), the cryptomaldamide gene cluster was captured using three PCR products of approximately 10 kb in length that overlap with one another. The first and last PCR products shared 30 bp of homology with the capture plasmid backbone. Yeast plasmids were extracted and screened via PCR for the correct insert. Positive plasmids were pooled, transformed into *E. coli,* and the cells were grown in LB supplemented with Km (50 μg/mL). After extraction from *E. coli,* plasmids were screened by restriction analysis with NcoI-HF and 90 positive clones were archived at -80°C.

3.1.2 Columbamide

One risk with capturing a BGC from PCR products is the high chance of getting point mutations in large DNA fragments (~ 30 kb). Therefore, in addition to using PCR to amplify the

entire BGC as ~ 5 kb overlapping fragments (so far, the most dependable method of capture in our lab), we decided to capture the BGC directly from gDNA. Because attempts to capture a BGC from gDNA using 50 bp hooks (as recommended in the original TAR cloning protocol) had previously failed in our lab, we constructed a TAR destination vector with 1 kb of homology flanking the BGC (**Figure 3.1**). We hypothesized that these longer hooks might facilitate BGC capture directly from gDNA. Upon gDNA extraction, it is possible that the ~40 kb columbamide BGC is fragmented over several DNA molecules. In addition, a ~40 kb BGC represents only a small fraction of the PNG genome (about 0.5%). Therefore, in case the capture of the entire pathway from gDNA might still fail, we decided to facilitate the recombination of sheared gDNA fragments carrying the BGC with selected PCR products that would be present in higher amounts.

To construct the TAR destination cloning vector for the columbamide, pAM5558, plasmid pAM5571 was digested with XhoI and NdeI and the ~1 kb columbamide hooks were cloned into the plasmid backbone using GeneArt Seamless Assembly.

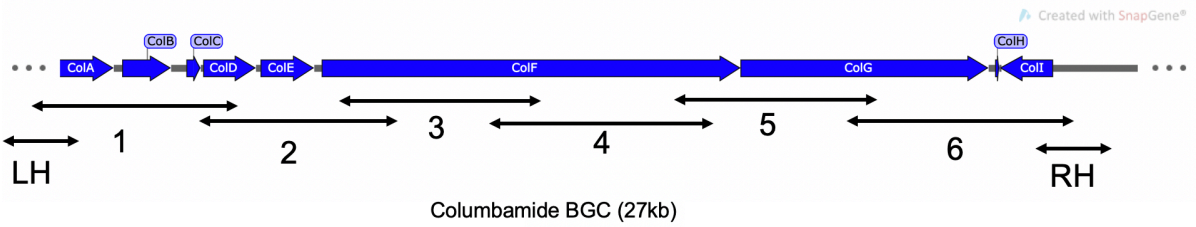


Figure 3.1 Schematic of PCR products used to amplify the columbamide gene cluster. LH and RH represent left and right "hooks" respectively, and these two PCR products were used to generate the TAR destination cloning vector for cloning the columbamide BGC.

The spheroplasting step within the TAR cloning strategy requires a zymolyase to break down the tough yeast cell wall and create competent yeast spheroplasts that can take up large DNA fragments (41). The activity of the zymolyase enzyme is known to fluctuate from batch to

batch (upon preparation) and over time, as reduction of the enzyme activity can occur during storage at 4°C (41). Therefore, before carrying out the TAR cloning experiment, we titrated the zymolyase enzyme to determine the amount of enzyme needed to achieve 90% spheroplasting within a 30-minute incubation. To quantify spheroplasting, we used the ratio of OD_{660} measurements of the zymolyase-treated cell suspension diluted in 1 M sorbitol (representing intact spheroplasts) and of the same suspension diluted in 2% SDS (representing lysed spheroplasts). A 5-20 fold difference between sorbitol and SDS OD_{660} readings represents the point when spheroplasts are ready and indicates 80-95% spheroplast conversion (41, 53). For this assay, 5 different concentrations of zymolyase were tested and OD_{660} readings were measured every 10 to 20 minutes over a 50-minute period. It is recommended to not exceed 50-minute incubations with zymolyase. We concluded from this assay that a concentration of 50 mg/mL zymolyase was the most effective and practical because this concentration reached the desired spheroplasting level within the recommended time while also not requiring the wasteful use of the enzyme (**Figure 3.2**).

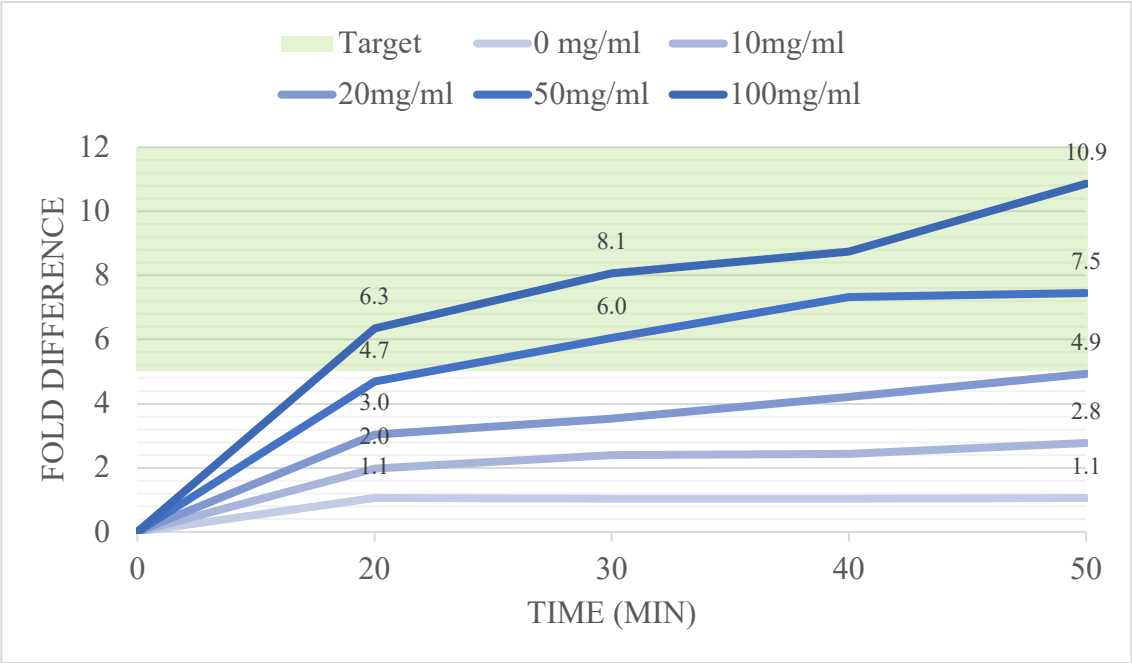


Figure 3.2 Graph of spheroplasting progress over time using various concentrations of zymolyase. The green background highlights an increase of at least 5-fold in the production of spheroplasts.

The columbamide BGC and flanking sequence was produced as 6 fragments by PCR (**Figure 3.3a**). Then TAR cloning assembly reactions were prepared using the columbamide BGC destination vector and either gDNA only, PCR products only, or gDNA plus PCR products #2 and #4 **(Figure 3.3b).** The transformations were plated on synthetic dropout medium lacking tryptophan to select for prototrophic cells carrying the plasmid. As shown in **Table 3.1**, the TAR cloning reaction with gDNA alone produced only 8 colonies, compared to over 100 colonies produced for the reactions containing PCR products only. Positive and negative controls were as expected, indicating successful transformation and near complete linearization of the capture vector.

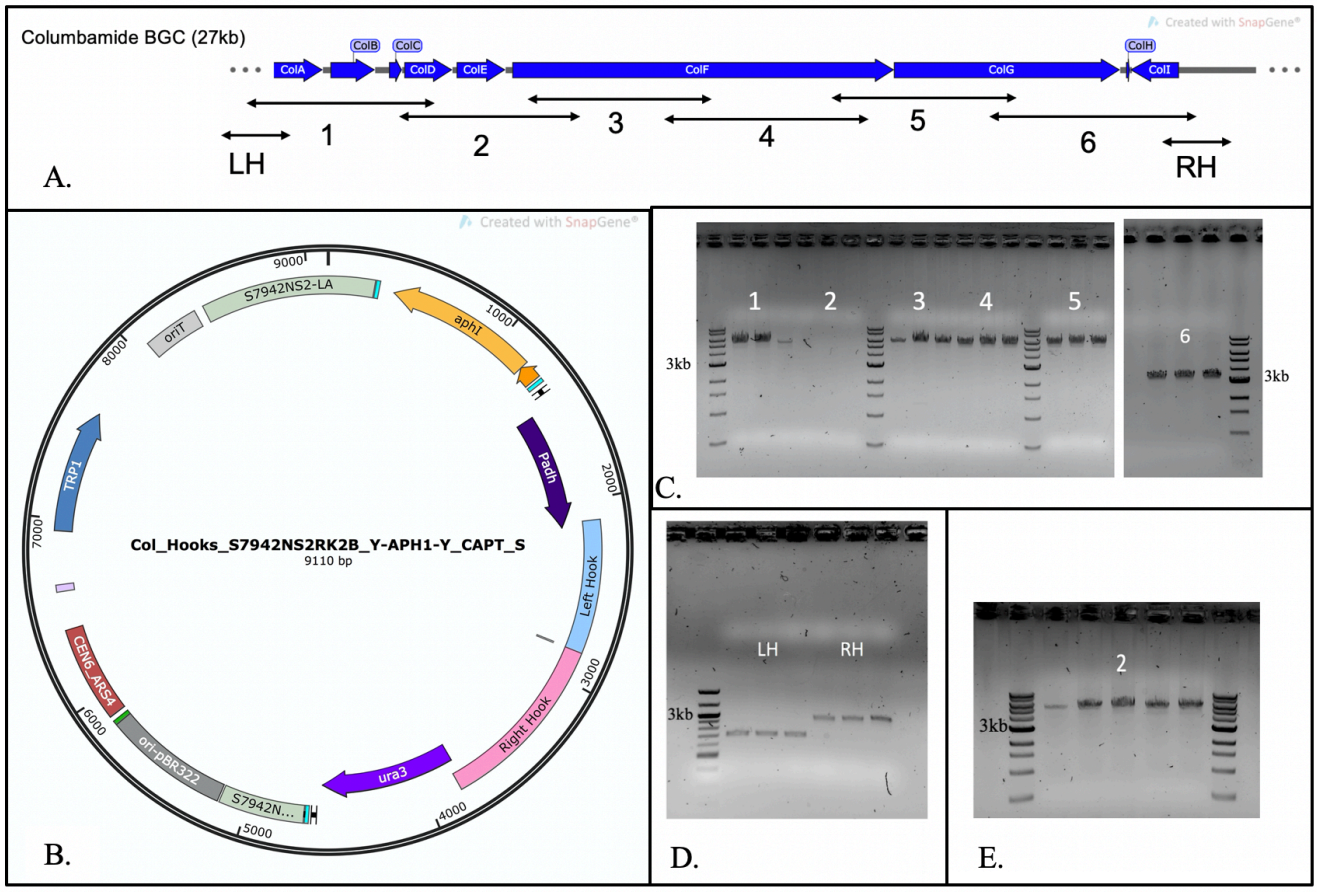


Figure 3.3 TAR cloning of the columbamide BGC. a) Schematic of columbamide BGC pathway. Black arrows with numbers represent 3- to 6-kb PCR products used to amplify the entire gene cluster. b) Plasmid map of columbamide capture vector. c) Gel showing 3- to 6-kb PCR products, numbers correspond to those shown on pathway schematic above. d) Gel showing PCR amplified left and right "hooks". e) Repeat PCR amplification of PCR product #2

Table 3.1 Number of colonies for each TAR cloning reaction capturing the columbamide BGC.

Reaction	**Reaction components**	**Number of Colonies**
Positive control	Undigested capture vector	58
Negative control	Digested capture vector	1
Experimental	PNG gDNA	8
Experimental	PNG gDNA + PCR products	27
Experimental	PCR products only	>150

Recombinant plasmids were then extracted from 5 mL liquid cultures of the yeast clones and screened by PCR for the presence of the BGC. PCR-positive plasmids were transformed individually into *E. coli*, and further screened by restriction digestion with NcoI-HF.

A large proportion of plasmids derived from gDNA and PCR reactions gave positive results after the PCR screen of DNA from yeast **(Table 3.2)**. However, after transforming these plasmids into *E. coli* and using a NcoI-HF digest to confirm the presence of the BGC, the pattern from the digests showed that these clones contained only empty capture vectors. The same result was observed for the single gDNA-only plasmid that had been confirmed in yeast. After using PCR to screen dozens of plasmids isolated from yeast clones, it was found that only those that were derived from reactions with solely PCR products gave positive results from the restriction digest screens of *E. coli* clones (**Table 3.2** and **Figure 3.4**). We decided to sequence-verify six of these plasmids.

Table 3.2 Number of columbamide plasmids screened at each stage.

Plasmid Type	**Total # Screened by PCR from yeast**	**# Confirmed by PCR from yeast**	**# Screened by digest from *E. coli***	**# Confirmed by digest from *E. coli***
gDNA only	4	1	1	0
gDNA + PCR	19	19	19	0
PCR only	10	10	9	9

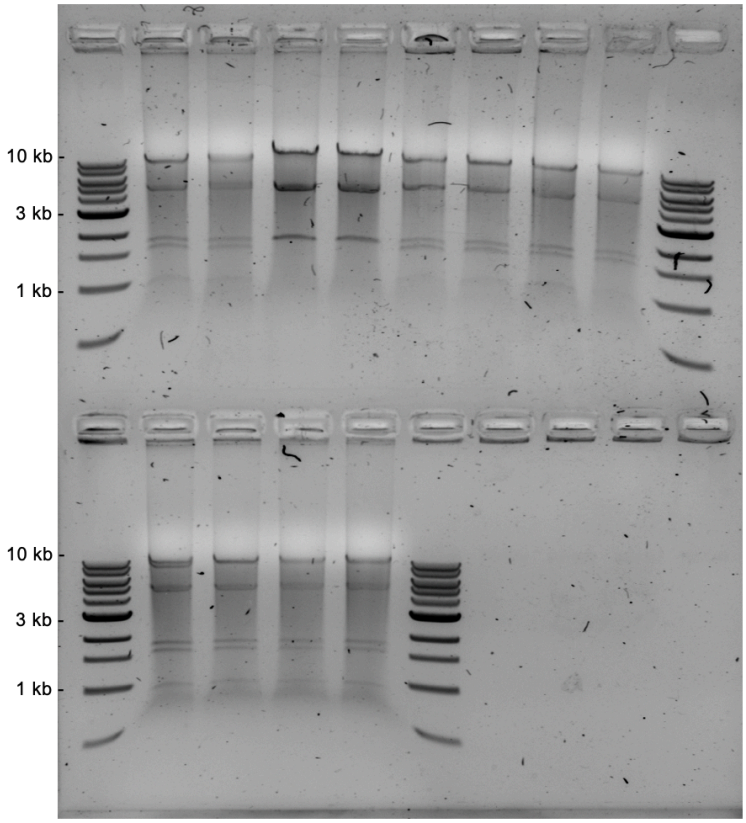


Figure 3.4 NcoI-HF digest screen of columbamide capture vector carrying columbamide BGC. Expected band sizes for columbamide within destination vector in bp are: 13502, 12124, 5440, 1804, 1579, and 904.

3.1.3 Palmyramide

The strategy to capture the palmyramide BGC closely followed the strategy used to capture the columbamide pathway. However, because palmyramide is slightly larger at 41 kb, it was decided to split the palmyramide BGC into two parts (part A and part B) which would be cloned in two separate vectors and then transformed sequentially into *S. elongatus* to reconstitute the entire BGC in the chromosome. However, the two TAR cloning backbone plasmids (pAM5570 and pAM5571) contain four different regions of homology (*ura3, trp1,* the alcohol dehydrogenase promoter pADH, and the *S. elongatus* right arm) which create an opportunity for

the second integrative plasmid to recombine improperly with the first plasmid and "loop out" the first part of the BGC. To avoid these unwanted recombinations, the palmyramide part A (PalA) capture vector (pAM5559) was assembled via the LiAc yeast transformation method described previously in section 2.5 to remove the homologous regions of the backbone plasmid. The yeast cells were able to successfully recombine the linearized pAM5559 (cut with NdeI) and two PCR products representing left and right hooks. The palmyramide part B (PalB) capture vector (pAM5560) was constructed by digesting pAM5571 with XhoI followed by a GeneArt Seamless Assembly to incorporate the 1 kb hooks into the backbone. Screening methods followed the same strategy as the columbamide pathway, and a table summarizing the results from the TAR cloning reactions is shown below (**Table 3.3**).

Table 3.3 Number of colonies for each TAR cloning reaction capturing the palmyramide BGC.

BGC Part	**Reaction**	**Reaction components**	**Number of Colonies**
PalA	Positive control	Undigested capture vector	>200
	Negative control	Digested PalA capture vector	68
	Experimental	PAL gDNA	15
	Experimental	PAL gDNA + PCR products	31
	Experimental	PCR products only	83
PalB	Positive control	Undigested capture vector	~200
	Negative control	Digested PalB capture vector	26
	Experimental	PAL gDNA	54
	Experimental	PAL gDNA + PCR products	76
	Experimental	PCR products only	42

For the palmyramide part A, as in previous TAR cloning experiments, more clones were obtained in reactions where the BGC was carried on PCR products only compared to reactions utilizing gDNA. Note that our negative control suggests that the backbone was not completely

digested. The results from the PalB transformation show a minimal difference between the three experimental conditions. It would be useful to repeat these two experiments with a more thoroughly digested capture vector and more replicates in an attempt to reliably quantify the recombination efficiency of these different experimental methods.

Table 3.5 Number of palmyramide part A plasmids screened at each stage.

Plasmid Type	**Total # Screened by PCR from yeast**	**# Confirmed by PCR from yeast**	**# Screened by digest from *E. coli***	**# Confirmed by digest from *E. coli***
gDNA only	15	0	0	0
gDNA + PCR P.	17	0	0	0
PCR P. only	20	20	12	8

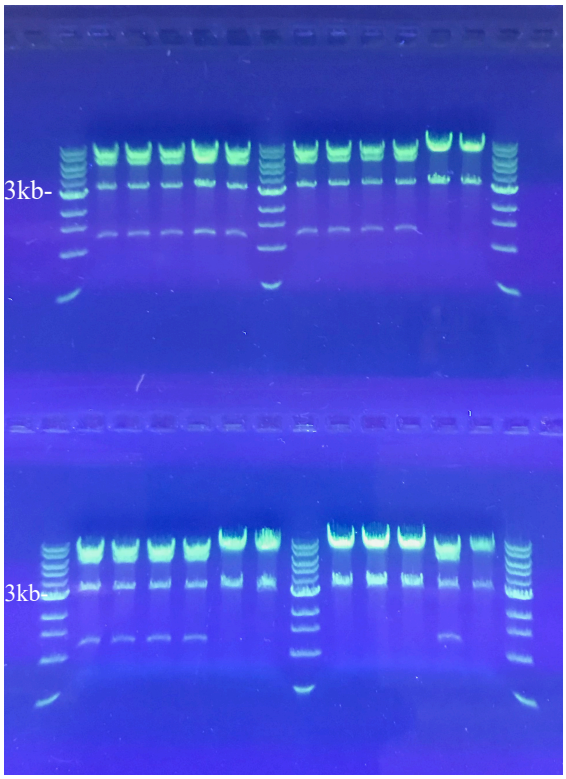


Figure 3.5 EcoRV digest of PalA plasmids; expected band sizes in kb are: 8.6, 7.3, 6, 3.4, and 1.2.

Eight PalA plasmids were confirmed by restriction digest analysis using EcoRV (**Table 3.5** and **Figure 3.5).** After screening 46 palmyramide part A plasmids derived from reactions with gDNA or gDNA plus PCR products but not finding any promising results, we focused on clones obtained from PCR products only. Therefore, these were the only clones screened for palmyramide part B constructs. 17 plasmids from PCR only reactions were screened via PCR and all gave positive results. 12 of these plasmids were then electroporated into NEB Stable cells and later digested with EcoRV to identify the correct clones; 7 out of these 12 plasmids gave the expected banding pattern (**Table 3.6** and **Figure 3.6**). Eight different plasmids for palmyramide part A and seven different plasmids for palmyramide part B were confirmed by restriction analysis using EcoRV and were archived until further verification by sequencing. This is the extent to which this report will discuss the palmyramide pathway.

Table 3.6 Number of palmyramide part B plasmids screened at each stage.

Plasmid Type	**Total # Screened by PCR from yeast**	**# Confirmed by PCR from yeast**	**# Screened by digest from *E. coli***	**# Confirmed by digest from *E. coli***
gDNA only	0	0	0	0
gDNA + PCR P.	0	0	0	0
PCR P. only	17	17	12	7

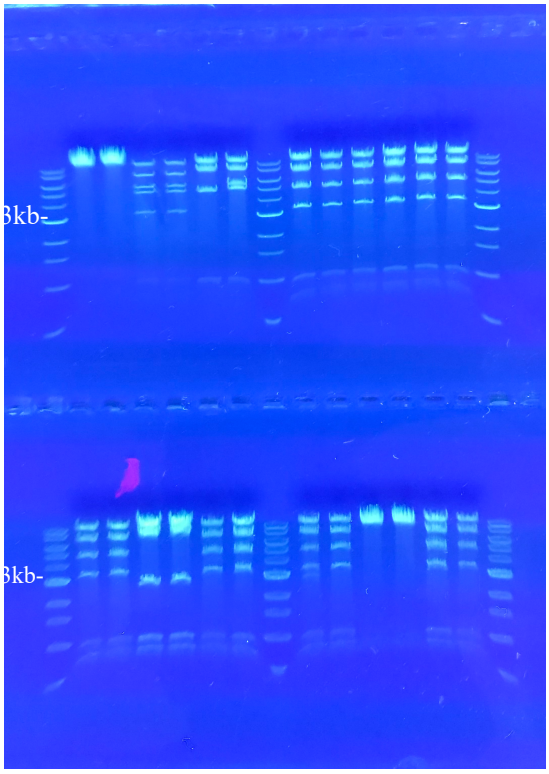


Figure 3.6 EcoRV digest of PalB plasmids; expected band sizes in kb are: 12.6, 7.8, 5.2, 3.4, 1, and 700.

3.2 Production of Natural Products in *S. elongatus*

3.2.1 Cryptomaldamide

Previous experiments, carried out by Arnaud Taton and Nathan Moss on a small number of clones (4), suggested that the cryptomaldamide gene cluster was transcribed in *S. elongatus* but the compound was not produced. Therefore, to overcome the risk associated with critical point mutations in the BGC acquired during PCR, a high throughput screen for the production of cryptomaldamide was developed. Dozens (~ 90) of plasmids were verified by restriction digests, positive clones were pooled together in equal amounts and then the pool was transformed in AMC2566. AMC2566 is an engineered *S. elongatus* strain, cured of its small plasmid pANS and carrying an *sfp* PPTase from *B. subtilis* under riboswitch control within neutral site 3. 68 *S.*

elongatus clones were obtained, grown in BG-11 96-well plates supplemented with Km (5 μg/mL) and theophylline and screened via MALDI MS. No product was detected from these preliminary screens. As an alternative to this high-throughput screen, six plasmids carrying the cryptomaldamide pathway were sequenced using Illumina MiSeq with a very high coverage and the results were analyzed using the Bowtie Sequence Alignment software (version 1.2.3) (63). As shown in **Table 3.7**, five plasmid clones carried 3 to 8 single nucleotide polymorphisms (SNPs) and only one out of the six plasmids had only one mutation. This mutation was located in the intergenic region between cpmA and cpmB and was conserved in all 6 clones leading to the assumption that there was actually a mistake in the original sequencing of the genome and this SNP found within the intergenic region of clone 92 could be disregarded.

Table 3.7 Summary of sequencing results for 6 plasmids carrying the cryptomaldamide gene cluster. The sixth SNP within clone #3 was in a transposase which is not part of the biosynthetic pathway.

Clone #	Intergenic SNPs		Gene SNPs		Total SNPs
	Type	Number	Mutations (*cpmB*)	Number	
3	SUB	1	R1199R, E2086G, A2683S, L4494F, S4610L	5 (6*)	7
15	SUB	2	A1522A, G5049C	2	4
31	SUB	1	P220P, C1250C, L2050I	3	4
33	SUB	1	Q88K, P3544T	2	3
84	SUB	3	S191L, E176G, F204Y, P3850P, R4767H, L6130I	6	9
92	SUB	1	-	0	1

Clone #92 was used to transform *S. elongatus* strain AMC2566. Three independent segregated clones were cultured in 100 mL BG-11 with the appropriate antibiotics in continuous light (100 μmol photons m^{-2} s^{-1}). One day before the cultures were to be screened by LC-MS/MS, 0.5 mM theophylline per 0.1 OD_{750} (or a maximum of 5 mM) was added to the cultures to induce expression of the *sfp*-PPTase, but cryptomaldamide could not be detected. In an attempt to boost the metabolism of the recombinant strains, the cryptomaldamide AMC2566 cultures were also grown with air bubbling at 70-150 μmol photon m^{-2} s^{-1} (light intensity was increased as the culture became denser) but no product could be detected.

3.2.2 Columbamide

Six plasmids were sequenced similarly to the cryptomaldamide plasmids. Sequencing results showed that while most plasmids contained 2 to 5 SNPs, one out of the six plasmids, #45, was found to have no mutations within the BGC at all (**Table 3.8**).

Table 3.8 Summary of sequencing results for 6 plasmids carrying the columbamide gene cluster

Clone #	**Intergenic SNPs**		**Gene SNPs**		**Total SNPs**
	Type	Number	Mutations (*colF*)	Number	
29B	SUB	1	L1122L	1	2
29D	SUB	3	P339H, L1122L,	2	5
42	-	0	G1745V, V2329I	2	2
44	SUB	2	Q967K, P2227T	2	4
45	-	0	-	0	0
46	SUB	3	S2161Y	1	4

The columbamide plasmid confirmed via sequencing (clone #45) was used to transform AMC2566 via natural transformation and twelve single clones carrying the pathway were grown as patches on BG-11 agar supplemented with Km (5 μg/mL) and Cm (7.5 μg/mL). Problems quickly arose when it became evident that the culture of *S. elongatus* carrying the columbamide BGC bleached when exposed to light greater than 16 μmol photons m^{-2} s^{-1}. The clones struggled to grow both on plates and in BG-11 liquid medium and would only survive at very low light exposure (**Image 3.1)**.

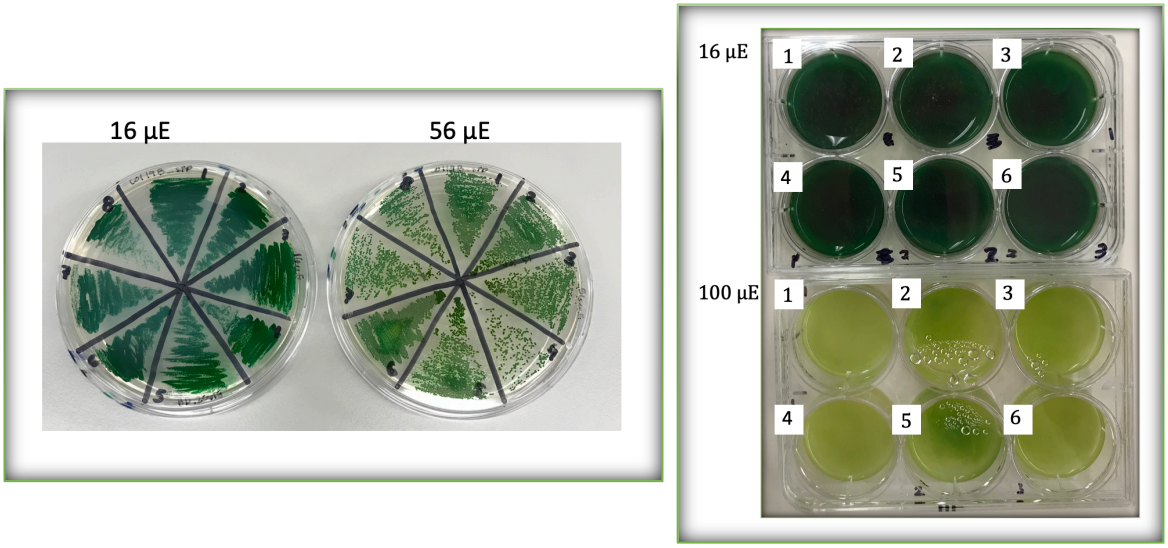


Image 3.1 Plate and liquid growth assay showing the light-sensitivity of *S. elongatus* strains carrying columbamide both on BG-11 agar and in BG-11 liquid media. BG-11 agar plates (left) were incubated at the indicated light levels for two weeks. Liquid cultures normalized to an OD_{700} of 0.2 and split amongst two 6-well plates. Both 6-well plates (right) were placed in 16 µmol photons m^{-2} s^{-1} for five days, and one plate was then transferred to high light (103 µmol photons m^{-2} s^{-1}) for two more days of growth.

Therefore, cultures for chemical analysis were grown at 16 µmol photons m^{-2} s^{-1} with air bubbling for 5-10 days and light intensity was increased to 50-100 µmol photons m^{-2} s^{-1} for the last 2-3 days of growth. 0.5 mM Theophylline per 0.1 OD750 (or a maximum of 5 mM) was added to the cultures in order to induce expression of the *sfp*-PPTase 48 hours before the cultures were harvested and analyzed by LC-MS/MS for production of columbamide. No product was detected from initial screens and it was hypothesized this could be from the lack of Cl ions in the BG-11 medium. The columbamide compounds contain 2-3 chlorine groups (**Figure 3.7**) and the native producer *M. producens* PNG is a marine strain grown in seawater BG-11 (swBG-11) which has a higher NaCl concentration than standard BG-11 media because it contains 33 g/L Instant Ocean Sea Salt.

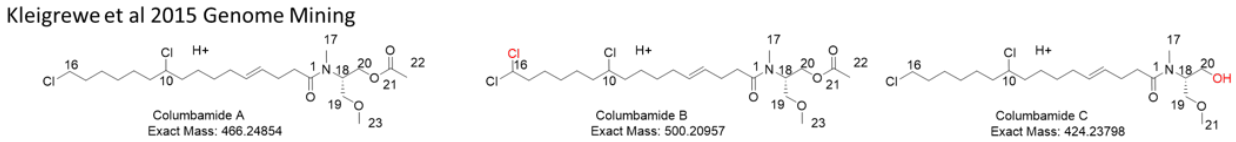


Figure 3.7 Molecular structures of columbamides A-C

Assuming that the availability of Cl ions could limit the organism's ability to produce columbamide, the previous experiment was repeated with 0.1 M NaCl added to the cultures, but the strains still failed to produce cryptomaldamide at detectable level. PNG, the native producer, was isolated from the ocean, which has an NaCl concentration of approximately 0.6M; this is far greater than the additional concentration of 0.1 M NaCl added to the media of our engineered *S. elongatus* strains. However, it has been shown that the addition of 0.2 M NaCl significantly decreases the growth rate of *S. elongatus* (64).

3.3 Use of *Anabaena* sp. PCC 7120 for natural product production

3.3.1 Cryptomaldamide

After the difficulties with producing these two natural products in *S. elongatus*, a new genetically tractable host was used. We thought that *Anabaena* PCC 7120 might be a more suitable heterologous host for the cryptomaldamide and the columbamide pathways based on previous work done at Oregon State University by the Philmus group. This work demonstrated the production of Lyngbyatoxin A by expressing the BGC (NRPS of ~ 10 kb) in *Anabaena* PCC 7120 and the activity of several promoters from *M. producens* in *Anabaena*, (42). *Anabaena* carries a native sfp-like PPTase, which might be better for cyanobacterial carrier proteins. It also has a larger genome than *S. elongatus* and might, for example, encode for chaperones essential to

properly fold PKS and NRPS megasynthases. Furthermore, it also carries several of its own NRPS, PKS, or hybrid secondary metabolites pathways (11, 14).

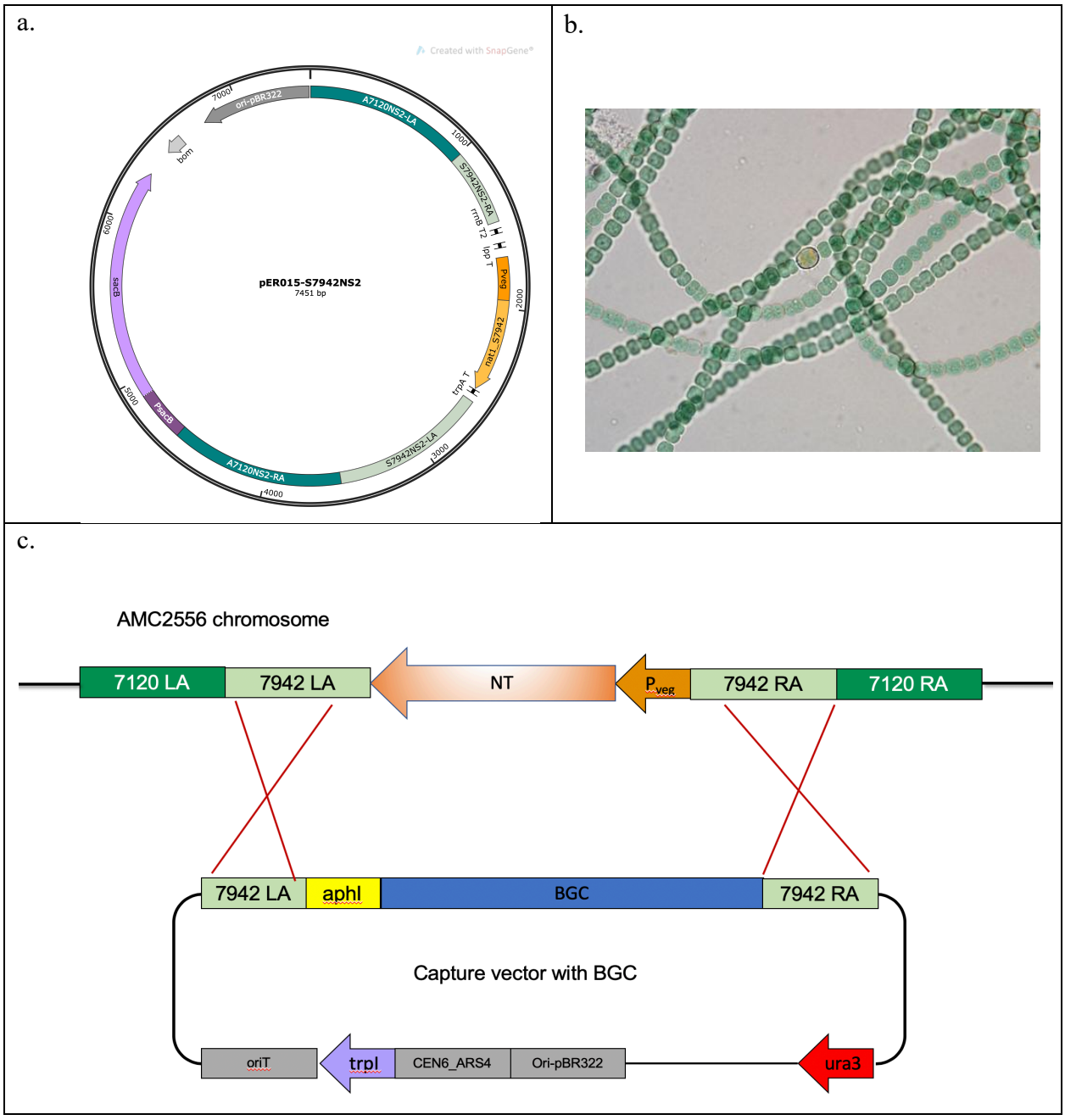


Figure 3.8 Schematic of AMC2556 design. a) per015-S7942NS2 plasmid map. *S. elongatus* neutral sites are shown in light green, flanking the NT cassette. *sacB* is used to select against single recombinants. b) *Anabaena* PCC 7120 filaments. c) Schematic showing recombination of BGC capture vector with engineered *Anabaena* strain AMC2556.

The plasmid carrying cryptomaldamide (pAM5564) was conjugated into *Anabaena* AMC2556. Isolated colonies of AMC2556 exconjugants were streaked several times on BG-11 agar plates containing Nm (25 μg/mL), so that the antibiotic selection over several generations would help segregate the chromosomes. Three clones were grown in 100-mL flasks and screened for production. The strains inoculated at an OD_{750} of approximately 0.02 were grown in 150 mL BG-11 medium bubbled with air to an OD_{750} of approximately 0.8. These three strains were then extracted and screened by Andrew Ecker (Gerwick Lab, SIO) for production of cryptomaldamide via LC-MS/MS, and relatively high amounts of cryptomaldamide were found both in media and cell extracts (**Table 3.9**).

Table 3.9 Screen for production of cryptomaldamide in *Anabaena* for 3 clones carrying the cryptomaldamide BGC and one control strain. Cultures were grown in duplicates. +, cryptomaldamide detected; -, cryptomaldamide not detectable; nd, not determined.

Culture	Biomass (EtOH)	Media (BuOH)	Media (aq)
Cl. 1A	+	nd	nd
1B	+	nd	nd
2A	-	+	+
2B	-	+	+
3A	-	+	+
3B	+	+	+
Control	-	nd	nd
Control	-	-	-

Although the production of cryptomaldamide was confirmed in *Anabaena,* initial analyses were performed on cultures of *Anabaena* that were not segregated double recombinants. Further studies, such as quantifying the amount of compound produced in different culture conditions, engineering the BGC to produce more compound, or interrogating the pathways by genetic manipulation in *Anabaena*, will require segregated, clonal double recombinant strains.

Therefore, we envisioned a CRISPR/Cpf1-based strategy to make a double strand break at the recombination site within AMC2556's chromosomes that had not integrated the BGC or had integrated the entire BGC-containing plasmid by a single recombination event. This strategy will cut unwanted chromosomes, thereby forcing the segregation of double recombinant chromosomes.

We used an RSF1010-based plasmid that autonomously replicates in *Anabaena*, and carries the gene encoding for the Cpf1 nuclease and a CRISPR array to express engineered guide RNAs (pAM5565). For the guide RNA, designed to target the NT resistance gene located between the recombination sites, a short DNA fragment produced by annealing complementary oligonucleotides was cloned into pAM5572 at the AarI site to produce pAM5565, which is shown in **Figure 3.9**.

This new construct, pAM5565, was then conjugated via bi-parental mating into AMC2560, which carries the cryptomaldamide BGC and was previously shown to produce cryptomaldamide. The mating reaction was plated on BG-11 agar plates containing 5% LB and Nm and Sp/Sm for selection of the BGC and CRISPR/Cpf1 plasmid, respectively.

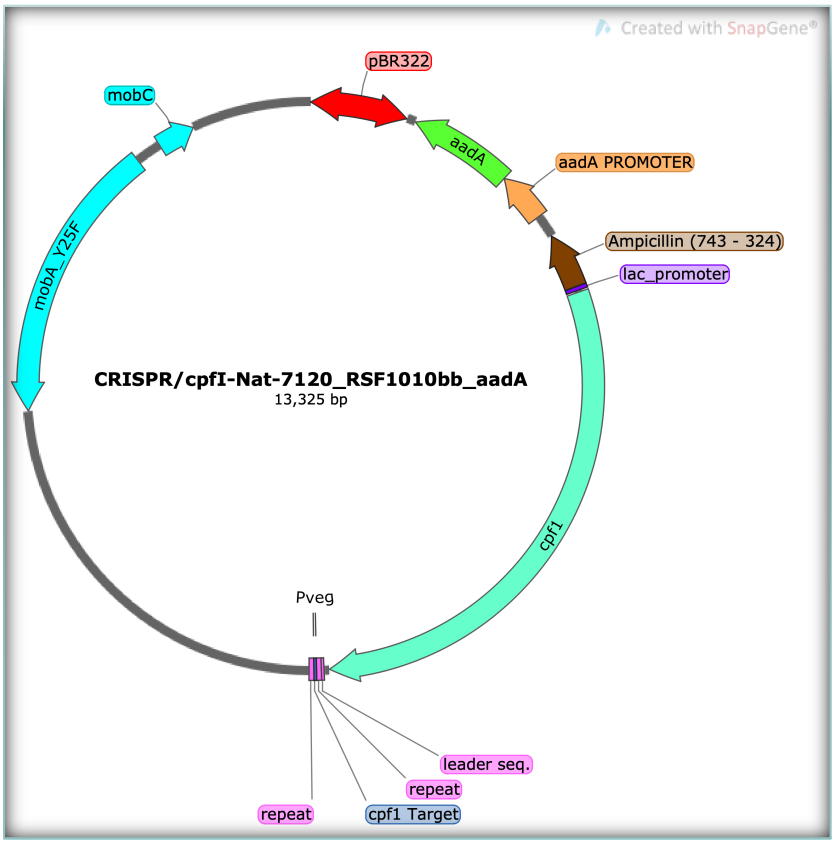


Figure 3.9 Plasmid map of pAM5565 containing the CRISPR/Cpf1 system used to target the Nt cassette in AMC2566 chromosomes that have not undergone a double recombination event.

Single colonies from these plates were then streaked on selective BG-11 plates and isolated colonies were patched on plates containing Nm and Sp/Sm or Nt to identify clones that lost resistance against Nt, which indicated that segregation was complete. Isolated colonies were streaked and patched 4 times on selective BG-11 plates, until we found clones that no longer could grow on Nt. At this point, gDNA was extracted from three clones and PCR analysis was used to screen for segregation. Expected fragment sizes and PCR gel analysis results are shown in **Table 3.10** and **Figure 3.10,** respectively. Clones 2 and 3 of the engineered *Anabaena* strains carrying the cryptomaldamide pathway (strains AMC2564 and AMC2565) are completely segregated, as they produce no PCR product for primers that bind to the left and right arms of the *S. elongatus* NS2. This is only possible if a double recombination event has occurred and the ~30

kb BGC disrupts the ability to produce the PCR product. Crypto_AMC2556 clone 1 (AMC2563) shows a very faint band for this PCR reaction, suggesting not all of its chromosomes are fully segregated. Segregated clones where confirmed to produce cryptomaldamide and are currently being grown as large cultures for a complete NMR characterization of the compound.

Table 3.10 Summary of the expected PCR products for single or double recombination of the cryptomaldamide BGC in AMC2556.

Region	Primer Sets	WT	Single Recombination	Double Recombination
Middle	Crypto_2015F and crypto_2788R	No product	774 bp	774 bp
Right	Crypt_27823F and S7942NS2LA_274R	No product	1449 bp	1449 bp
Left	aphI_528R and crypt_215R	No product	1703 bp	1703 bp
Out	S7942NS2-LA-F and S7942-RA-R	1517 bp	1517 bp	No product

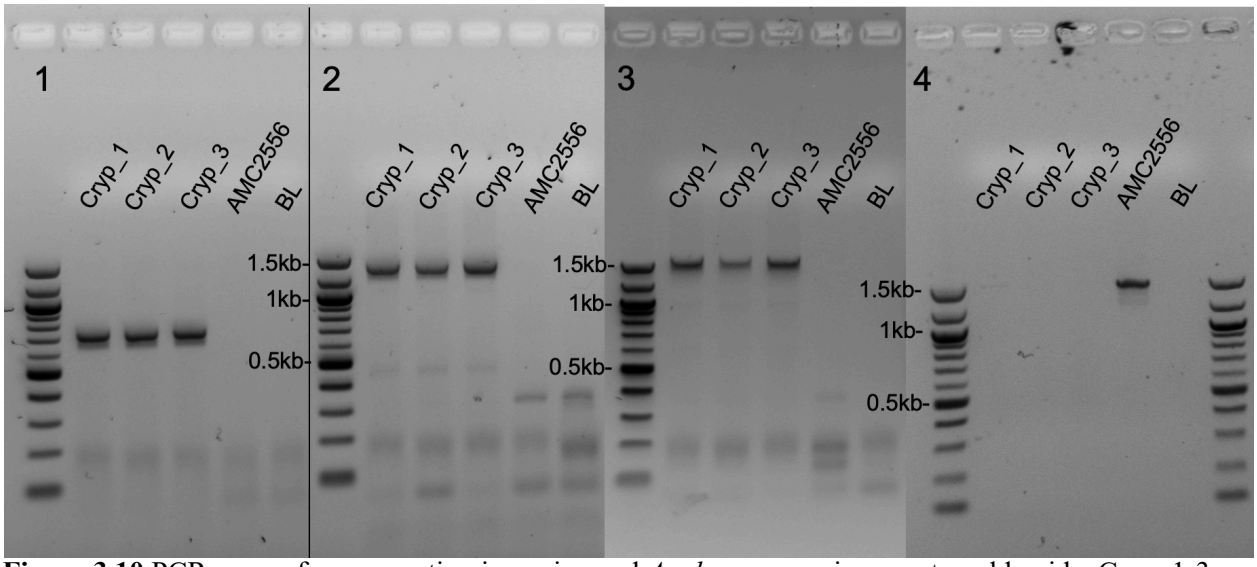


Figure 3.10 PCR screen for segregation in engineered *Anabaena* carrying cryptomaldamide. Cryp_1-3 represent the 3 different clones screened, 7120_NS2 is WT AMC2556, and BL is a negative control. Panel 1 contains reactions for the middle region, panel 2 for the right, panel 3 for the left, and panel 4 shows a screen for the 1.5 kb region spanning the *S. elongatus* left and right arms, with no recombinant DNA inserted between the arms.

3.3.2 Columbamide

The plasmid carrying the sequenced columbamide pathway (pAM5562) was also conjugated into *Anabaena* AMC2556 and plated on BG-11 agar plates containing Nm (25 μg/mL). The columbamide-carrying clones were again very difficult to grow. Initially, when single colonies from the conjugation were picked and patched on new plates, the cells would not grow well; the single clones had a very sticky phenotype and would produce bubbles. Very little material was able to be grown from these patches and we were never able to grow a liquid culture of these single clones. Every colony from the original conjugation/mating plates was collected, pooled, and grown in 100 mL BG-11 with Nm (25μg/mL). This pooled, non-clonal culture could be maintained with shaking at relatively low light (35 μmol photons m^{-2} s^{-1}). Aliquots of this culture were then inoculated in two tissue culture flasks containing 250 mL BG-11 plus 0.1 M NaCl and incubated bubbled with air at a light intensity of 40-50 μmol photons m^{-2} s^{-1}. After about 4 days of growing in these conditions, the cultures began to bleach. However, after 5 more days of growth, the cultures revived and became very dense (**Figure 3.11**). These non-clonal cultures were transferred to the Gerwick Lab and screened by Raphael Reher and Sebastian Rohrer for the production of columbamide by LC/MS-MS and further characterized by NMR. The production of columbamides A-C was confirmed, and in addition, four potentially new analogs were identified (**Figure 3.12** and **Figure 3.13**).

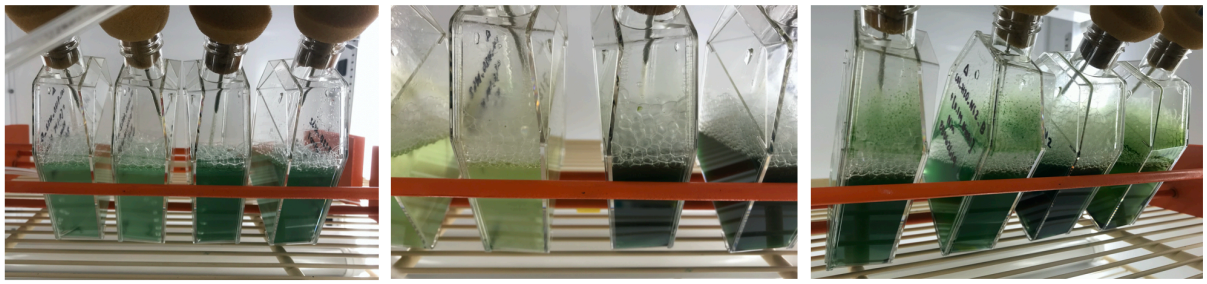


Figure 3.11 Growth progression over 12 days of non-clonal *Anabaena* carrying the columbamide BGC.

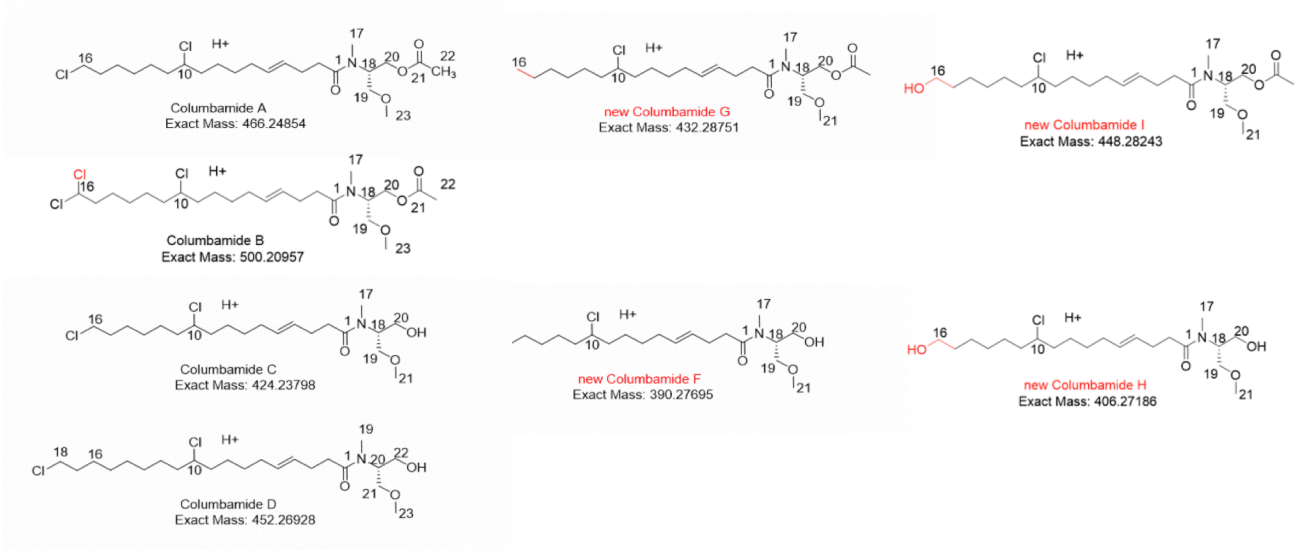


Figure 3.12 Molecular structures of known and potentially new columbamides; created by Raphael Reher of the Gerwick Lab

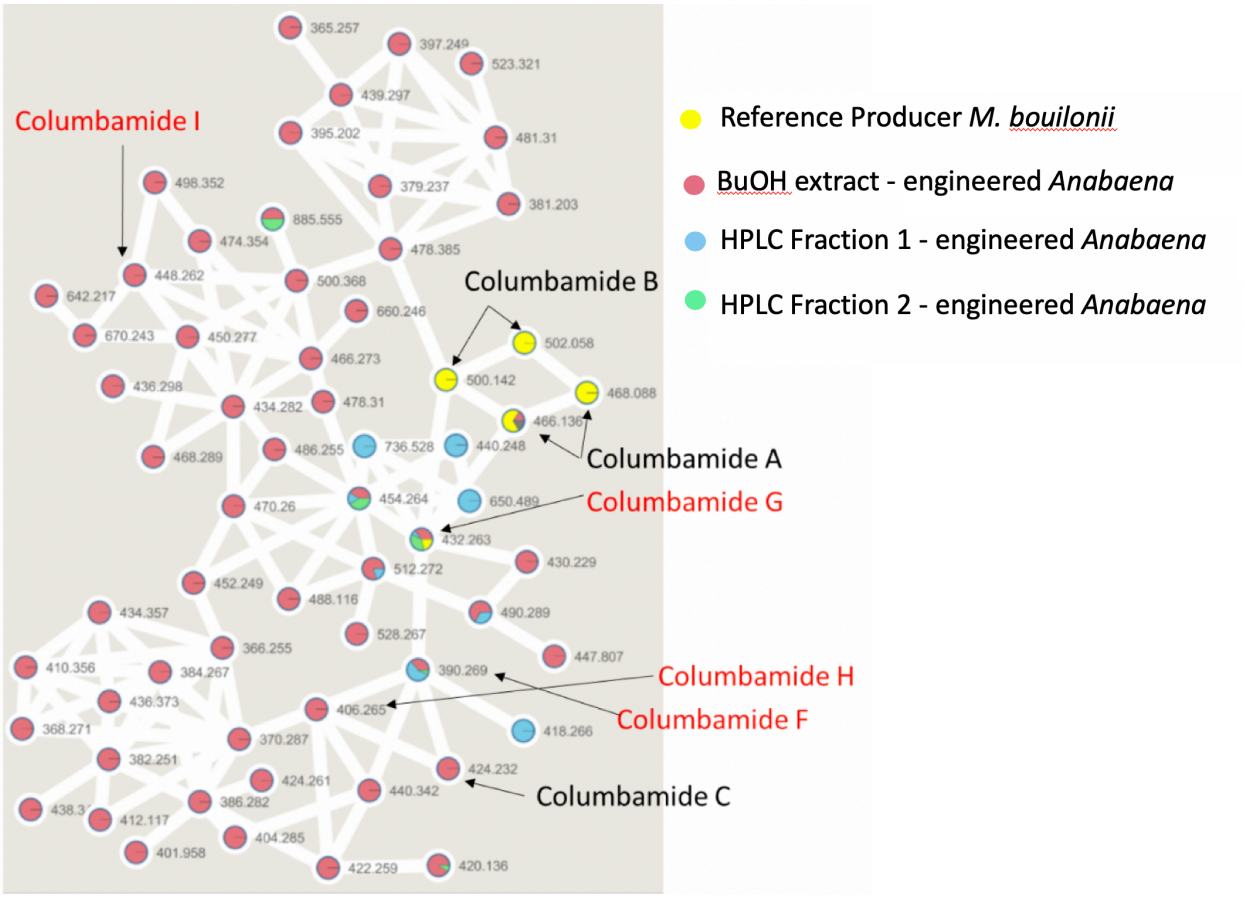


Figure 3.13 Molecular Networking Analysis of HPLCMS-MS showing relationship between known and putatively new columbamides. Network created by Raphael Reher of the Gerwick Lab.

After confirmation of the production of columbamide from the non-clonal culture, the next immediate goal was to acquire a segregated exconjugant clone from this culture. Serial dilutions of the non-clonal culture were plated on BG-11 agar plates supplemented with Nm (25 µg/mL) and after repeatedly picking and streaking single colonies, cells were isolated that grew as dense patches. Three isolated clones were transferred to 100-mL flasks of BG-11 with half the concentration of Nm typically used (12.5 µg/mL) to alleviate some stress for these already fragile strains. These three clones were grown in 250 mL BG-11 media with 0.1M NaCl and bubbled with air to be screened similarly to the non-clonal culture. The production of

columbamide from these clonal cultures were confirmed and the next step is to acquire segregated double recombinants using the CRISPR/Cpf1 technique described previously.

3.4 Pathway Editing

Using an approach that involves both CRISPR and TAR cloning in yeast (mCRISTAR) (65), we aimed to replace the native promoter of the biosynthetic pathway with an inducible T7 promoter system. The mCRISTAR strategy utilizes: (1) a yeast CRISPR/Cas9 plasmid, pCRCT, encoding a guide RNA that targets the Cas9 nuclease on the native promoter, (2) the shuttle plasmid carrying the BGC, and (3) a repair template produced by PCR and that includes the T7 promoter flanked by homology sequences to the linearized plasmid. The pCRCT plasmid is first introduced into yeast cells, followed by the shuttle plasmid and repair template. The shuttle plasmid is cleaved at the native promoter site by Cas9, and the T7 promoter fragment is used as repair template and places the native promoter using the yeast recombination machinery.

We have begun this experiment and constructed pAM5568, which is the pCRCT plasmid carrying a guide RNA targeting the promoter region of *colA* (20 bp upstream of *colA*). A gBlock DNA fragment (synthesized by IDT) that contains guide RNA and direct repeats flanked by two BsaI restriction sites was cloned into pCRCT by the Golden Gate method (65). The guide RNA was designed to target a 20-bp sequence just upstream of the *colA* gene. The T7 promoter, *lac* operator, and an RBS was PCR amplified from that same gBlock fragment.

pAM5568 was transformed into yeast cells by LiAc, and colonies from this transformation are ready to be screened. Next, we will co-transform the plasmid carrying the columbamide pathway (pAM5562) along with the homologous repair template. Previous studies have found that this system only works if the pCRCT plasmid is transformed into yeast cells

before the target plasmid, likely because this gives pCRCT enough time to express its CRISPR system (65).

The new construct with the engineered promoter will then be transferred into our platform strain of *Anabaena*, together with another plasmid carrying T7 RNA polymerase placed under control of a theophylline-inducible riboswitch and the *lacI* gene to further repress the expression of the pathway in the absence of IPTG.

Chapter 3 is coauthored with Diaz, Brienna; Taton, Arnaud; Ecker, Andrew; Reher, Raphael; Rohrer, Sebastian; Gerwick, Lena; Gerwick, William H.; and Golden, James. The thesis author was the primary author of this chapter.

Chapter Four: Conclusion

4.1 Discussion

Marine filamentous cyanobacteria are known to produce diverse bioactive secondary metabolites and their biosynthesis requires unique enzymatic functions. Many obstacles prevent a thorough investigation of these compounds, specifically, their functions and biosynthetic origins within their native hosts. The slow-growing nature of these marine cyanobacterial strains makes it difficult to explore these natural products and their BGCs in a timely manner. Furthermore, these natural products are often produced in low yields in their native host, making them difficult to acquire in high concentrations. Thus, the ability to perform chemical analyses to investigate the bioactivities and pharmaceutical potential of the compounds is inherently limited. Lastly, these marine cyanobacterial strains are not genetically tractable, creating another barrier for the investigation of these NRPS/PKS pathways. In this study we worked towards establishing an approach to address these limitations.

Heterologous expression of entire BGCs from cyanobacteria is challenging and has been achieved in only a few instances. Aside from smaller ribosomally synthesized and post-translationally modified peptides (RiPPs) pathways (66-69), only a few entire BGCs from cyanobacteria have been expressed in a heterologous host. The lyngbyatoxin A and microcystin toxins, from *M. producens* and *Microcystis aeruginosa* PCC 7806 respectively, have been expressed in *E. coli* (70,71). The native promoters and regulatory regions were not sufficiently functional to allow heterologous expression in *E. coli* and the biosynthesis of these natural products at a reasonable yield required refactoring of the regulatory regions of the marine BGCs (70,71). There has also been some work done using *Streptomyces* as a heterologous expression platform for cyanobacterial BGCs. This strategy has been applied to an ~11-kb BGC from *M.*

producens encoding for enzymes involved in the biosynthesis of lyngbyatoxin A, using *Streptomyces coelicolor* A3(2) as an expression host (72). However, although the two nonmodular proteins (LtxB and LtxC) were successfully expressed, the product of the *ltxA* gene, a large NRPS, was never produced due to premature transcriptional termination, and the secondary metabolite of interest lyngbyatoxin A was never produced in this heterologous host (72). The biosynthesis of another natural product, barbamide from *M. producens*, was attempted in *Streptomyces venezuelae* and resulted in a closely related derivative, 4-O-demethylbarbamide, with considerably low yields (73). These heterologous expression attempts suggest that differences in %GC content, codon usage, and protein folding and stability might have been barriers to successful and reliable production of these BGCs from cyanobacterial strains in actinobacteria. To overcome these concerns, *Moorea*'s lyngbyatoxin A BGC was expressed in *Anabaena* PCC 7120 and it led to the production of lyngbyatoxin A with yields comparable to those of *Moorea* (42).

In this study, we report the successful heterologous expression of two biosynthetic gene clusters from *Moorea* strains for the production of cryptomaldamide (29 kb BGC) and columbamides (27 kb BGC) in *Anabaena* PCC 7120. Previous attempts in *S. elongatus* had failed, possibly because *S. elongatus* has no endogenous NP BGCs, and a smaller genome and a higher %GC than do marine filamentous cyanobacteria and *Anabaena*.

The availability of precursors, substrates, and cofactors, as well as %GC content and codon usage are all known obstacles for heterologous expression, and these challenges can be greatly reduced when the host and expression strains are more closely related genetically and evolutionarily (74,75). Our results further reinforce this idea and further showed that *Anabaena*

is a suitable host for the production of secondary metabolites from marine filamentous cyanobacteria, even with the native promoters and RBSs from two different pathways.

On more technical grounds, we found that the capture of BGCs through TAR cloning in yeast directly from gDNA (excluding the Pal part B constructs because these were not screened) was difficult and failed in our hands. However, we showed that reconstructing large functional BGCs in yeast from PCR products is possible. The assembly of overlapping large PCR products (~ 5 kb) in yeast is highly efficient, however, requires an additional sequencing step to identify plasmid clones that do not contain any critical sequence mismatches introduced by PCR or cloning.

We found that the culture conditions in which our strains were grown were important for the production of the compounds. Neither the cryptomaldamide nor columbamide analogs were produced at detectable levels when the cultures were simply grown on a shaker at low light or non-shaking on a shelf. For the strains of *Anabaena* carrying the cryptomaldamide BGC, bubbling fresh cultures with air boosted the growth and likely the overall metabolism of the strains which lead to production of detectable amounts of cryptomaldamide. For the production of the columbamides, which are halogenated, we found that in addition to bubbling the cultures with air, additional chloride (as NaCl) had to be provided, as cultures grown in the same conditions lacking additional chloride did not produce the compounds. In conclusion, it is highly important to consider both the composition of the compound of interest as well as the habitat of its native organism when attempting to obtain high yield of natural products by heterologous expression.

This study's successful heterologous expression of two natural products further validate *Anabaena* PCC 7120 as a heterologous expression host for marine cyanobacterial natural

products. The heterologous expression of marine cyanobacterial natural products including cryptomaldamide and columbamide will pave the road for drug discovery and the investigation of new biosynthetic gene clusters.

4.2 Future Directions

Ultimately, the production of these natural products in significant amounts will permit further investigation of columbamides and cryptomaldamide as therapeutic drug leads, enable detailed studies of these large gene clusters and the biosynthetic enzymes, and may provide an opportunity to study their ecological roles for the native producers.

The next steps forward for this project will be: 1, to obtain fully segregated strains carrying the columbamide pathway. This is required to obtain stable strains which will allow us to reliably quantify the amount of NP produced and to set up experiments using different growth conditions to determine the best conditions for optimal yield. 2, Determine the best growth conditions for the strains carrying the cryptomaldamide pathway, which were already confirmed to be fully segregated.

Furthermore, now that these two BGCs are functioning in a genetically tractable cyanobacterial strain, this provides a means to study the mechanisms of the biosynthetic pathways. For example, as suggested by one of collaborators (Raphael Rohrer, Gerwick Lab), it has been shown that two SNPs at the active site of another L-arginine:glycine amidinotransferase, which is similar to *cpmA* and involved in the biosynthesis of cylindrospermopsin, decreased substrate specificity (76). It would be interesting to mutate these positions in *cpmA* (cryptomaldamide) to determine if this results in the production of novel compounds. This could shed light on the function of these enzymes within BGCs and could

contribute to the synthetic production of various compounds. Another idea brought up for this project was to study the halogenation process in the biosynthesis of columbamide. The columbamide BGC carries two halogenases; both genes could be inactivated or modified to investigate the role of these genes in the production of the compound.

Lastly, the robust production of these and other NP compounds will allow more extensive investigation of their properties. The bioactivities of these NP compounds can be explored further with assays testing for antibiotic, anti-virulence, anticancer, anti-inflammatory, and other properties.

References

1. Hamilton, T. L., Bryant, D. A., & Macalady, J. L. (2016). The role of biology in planetary evolution: cyanobacterial primary production in low-oxygen Proterozoic oceans. *Environmental Microbiology*, *18*(2), 325–340.

2. Nabout, J., Rocha, B., Carneiro, F., & C.L, S. (2013). How many species of Cyanobacteria are there? Using a discovery curve to predict the species number. *Biodiversity and Conservation*, *22*.

3. Herdman, M., Janvier, M., Waterbury, J. B., Rippka, R., Stanier, R. Y., & Mandel, M. (1979). Deoxyribonucleic Acid Base Composition of Cyanobacteria. *Journal of General Microbiology*, *111*(1), 63–71.

4. Hess, W. R. (2011). Cyanobacterial genomics for ecology and biotechnology. *Current Opinion in Microbiology*, *14*(5), 608–614.

5. Rippka, R., Waterbury, J., & Cohen-Bazire, G. (1974). A cyanobacterium which lacks thylakoids. *Archives of Microbiology*, *100*(1), 419–436.

6. Schirrmeister, B. E., Gugger, M., & Donoghue, P. C. J. (2015). Cyanobacteria and the Great Oxidation Event: evidence from genes and fossils. *Palaeontology*, *58*(5), 769–785.

7. Stal, L. J. (2015, December 23). Nitrogen Fixation in Cyanobacteria. *ELS*, pp. 1–9.

8. Wood, R. (2016). Acute animal and human poisonings from cyanotoxin exposure — A review of the literature. *Environment International*, *91*, 276–282.

9. Shah, S. A. A., Akhter, N., Auckloo, B. N., Khan, I., Lu, Y., Wang, K., Wu, B., Guo, Y.-W. (2017). Structural Diversity, Biological Properties and Applications of Natural Products from Cyanobacteria. A Review. *Marine Drugs*, *15*(11), 354.

10. Dittmann, E., Gugger, M., Sivonen, K., & Fewer, D. P. (2015). Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. *Trends in Microbiology*, *23*(10), 642–652.

11. Calteau, A., Fewer, D. P., Latifi, A., Coursin, T., Laurent, T., Jokela, J., Kerfeld, C. A., Sivonen, K., Piel, J., & Gugger, M. (2014). Phylum-wide comparative genomics unravel the diversity of secondary metabolism in Cyanobacteria. *BMC Genomics*, *15*(1), 977.

12. Tan, L. T. (2010). Filamentous tropical marine cyanobacteria: a rich source of natural products for anticancer drug discovery. *Journal of Applied Phycology*, *22*(5), 659–676. https://doi.org/10.1007/s10811-010-9506-x

13. Demay, J., Bernard, C., Reinhardt, A., & Marie, B. (2019). Natural Products from Cyanobacteria: Focus on Beneficial Activities. *Marine Drugs*, *17*(6), 320.

14. Shih, P. M., Wu, D., Latifi, A., Axen, S. D., Fewer, D. P., Talla, E., Calteau, A., Cai, F., Tandeau de Marsac, N., Rippka, R., Herdman, M., Sivonen, K., Coursin, T., Laurent, T., Goodwin, L., Nolan, M., Davenport, K. W., Han, C. S., Rubin, E. M., Eisen, J. A., Woyke, T., Gugger, M., & Kerfeld, C. A. (2013). Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(3), 1053–1058.

15. Newman, D. J., & Cragg, G. M. (2016). Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products*, *79*(3), 629–661.

16. Kehr, J.-C., Gatte Picchi, D., & Dittmann, E. (2011). Natural product biosyntheses in cyanobacteria: A treasure trove of unique enzymes. *Beilstein Journal of Organic Chemistry*, *7*, 1622–1635.

17. Cane, D. E., Walsh, C. T., & Khosla, C. (1998). Harnessing the Biosynthetic Code: Combinations, Permutations, and Mutations. *Science*, *282*(5386), 63 LP – 68.

18. Cane, D. E., & Walsh, C. T. (1999). The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. *Chemistry & Biology*, *6*(12), R319–R325.

19. Marahiel, M. A., Stachelhaus, T., & Mootz, H. D. (1997). Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. *Chemical Reviews*, *97*(7), 2651–2674.

20. Mootz, H. D., Schwarzer, D., & Marahiel, M. A. (2002). Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases. *ChemBioChem*, *3*(6), 490–504.

21. Hopwood, D. A. (1997). Genetic Contributions to Understanding Polyketide Synthases. *Chemical Reviews*, *97*(7), 2465–2498.

22. Staunton, J., & Weissman, K. J. (2001). Polyketide biosynthesis: a millennium review. *Natural Product Reports*, *18*(4), 380–416.

23. Khosla, C., Gokhale, R. S., Jacobsen, J. R., & Cane, D. E. (1999). Tolerance and Specificity of Polyketide Synthases. *Annual Review of Biochemistry*, *68*(1), 219–253.

24. Du, L., & Shen, B. (2001). Biosynthesis of hybrid peptide-polyketide natural products. *Current Opinion in Drug Discovery & Development*, *4*(2), 215–228.

25. Beld, J., Sonnenschein, E. C., Vickery, C. R., Noel, J. P., & Burkart, M. D. (2014). The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Natural Product Reports*, *31*(1), 61–108.

26. Yang, G., Zhang, Y., Lee, N. K., Cozad, M. A., Kearney, S. E., Luesch, H., & Ding, Y. (2017). Cyanobacterial Sfp-type phosphopantetheinyl transferases functionalize carrier proteins of diverse biosynthetic pathways. *Scientific Reports*, *7*(1), 11888.

27. Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., & Walsh, C. T. (1998). Characterization of Sfp, a Bacillus subtilis Phosphopantetheinyl Transferase for Peptidyl Carrier Protein Domains in Peptide Synthetases. *Biochemistry*, *37*(6), 1585–1595.

28. 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., Reinhardt, R., Kube, M., Burkart, M. D., Allen, E. E., Dorrestein, P. C., Gerwick, W. H., & Gerwick, L. (2011). Genomic insights into the physiology and ecology of the marine filamentous cyanobacterium *Lyngbya majuscula*. *Proceedings of the National Academy of Sciences*, *108*(21), 8815.

29. Leao, T., Castelão, G., Korobeynikov, A., Monroe, E. A., Podell, S., Glukhov, E., Allen, E. E., Gerwick, W. H., & Gerwick, L. (2017). Comparative genomics uncovers the prolific and distinctive metabolic potential of the cyanobacterial genus Moorea. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(12), 3198–3203.

30. Kleigrewe, K., Gerwick, L., Sherman, D. H., & Gerwick, W. H. (2016). Unique marine derived cyanobacterial biosynthetic genes for chemical diversity. *Natural Product Reports*, *33*(2), 348–364.

31. Engene, N., Rottacker, E. C., Kaštovský, J., Byrum, T., Choi, H., Ellisman, M. H., Komárek, J., & 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*(Pt 5), 1171–1178.

32. 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., & 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*(7), 1671–1682.

33. Kinnel, R. B., Esquenazi, E., Leao, T., Moss, N., Mevers, E., Pereira, A. R., Monroe, E. A., Korobeynikov, A., Murray, T. F., Sherman, D., Gerwick, L., Dorrestein, P. C., & Gerwick, W. H. (2017). A Maldiisotopic Approach to Discover Natural Products: Cryptomaldamide, a Hybrid Tripeptide from the Marine Cyanobacterium Moorea producens. *Journal of Natural Products*, *80*(5), 1514–1521.

34. Heidorn, T., Camsund, D., Huang, H.-H., Lindberg, P., Oliveira, P., Stensjö, K., & Lindblad, P. (2011). Synthetic Biology in Cyanobacteria: Engineering and Analyzing Novel Functions. *Methods in Enzymology*, *497*, 539–579.

35. Sengupta, A., Pakrasi, H. B., & Wangikar, P. P. (2018, July 1). Recent advances in synthetic biology of cyanobacteria. *Applied Microbiology and Biotechnology*, Vol. 102, pp. 5457–5471.

36. Ma, A. T., Schmidt, C. M., & Golden, J. W. (2014). Regulation of Gene Expression in Diverse Cyanobacterial Species by Using Theophylline-Responsive Riboswitches. *Applied and Environmental Microbiology*, *80*(21), 6704 LP – 6713.

37. Wang, B., Wang, J., Zhang, W., & Meldrum, D. (2012). Application of synthetic biology in cyanobacteria and algae. *Frontiers in Microbiology*, Vol. 3, p. 344.

38. Taton, A., Unglaub, F., Wright, N. E., Zeng, W. Y., Paz-Yepes, J., Brahamsha, B., Palenik, B., Peterson, T. C., Haerizadeh, F., Golden, S. S., & Golden, J. W. (2014). Broad-host-range vector system for synthetic biology and biotechnology in cyanobacteria. *Nucleic Acids Research*, *42*(17), e136–e136.

39. Ducat, D. C., Way, J. C., & Silver, P. A. (2011). Engineering cyanobacteria to generate high-value products. *Trends in Biotechnology*, *29*(2), 95–103.

40. Roulet, J., Taton, A., Golden, J., Arabolaza, A., Burkart, M., & Gramajo, H. (2018). Development of a cyanobacterial heterologous polyketide production platform. *Metabolic Engineering*, *49*.

41. Kouprina, N., & Larionov, V. (2008). Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast Saccharomyces cerevisiae. *Nature Protocols*, *3*, 371.

42. Videau, P., Wells, K. N., Singh, A. J., Gerwick, W. H., & 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*, *5*(9), 978–988.

43. Clerico, E. M., Ditty, J. L., & Golden, S. S. (2007). *Specialized Techniques for Site-Directed Mutagenesis in Cyanobacteria BT - Circadian Rhythms: Methods and Protocols* (E. Rosato, ed.).

44. Niederholtmeyer, H., Wolfstädter, B. T., Savage, D. F., Silver, P. A., & Way, J. C. (2010). Engineering cyanobacteria to synthesize and export hydrophilic products. *Applied and Environmental Microbiology*, *76*(11), 3462–3466.

45. Olmedo-Verd, E., Muro-Pastor, A. M., Flores, E., & Herrero, A. (2006). Localized induction of the ntcA regulatory gene in developing heterocysts of Anabaena sp. strain PCC 7120. *Journal of Bacteriology*, *188*(18), 6694–6699.

46. Chaurasia, A. K., Parasnis, A., & Apte, S. K. (2008). An integrative expression vector for strain improvement and environmental applications of the nitrogen fixing cyanobacterium, Anabaena sp. strain PCC7120. *Journal of Microbiological Methods*, *73*(2), 133–141.

47. Puerta-Fernández, E., & Vioque, A. (2011). Hfq is required for optimal nitrate assimilation in the Cyanobacterium Anabaena sp. strain PCC 7120. *Journal of Bacteriology*, *193*(14), 3546–3555.

48. Paz-Yepes, J., Flores, E., & Herrero, A. (2009). Expression and mutational analysis of the glnB genomic region in the heterocyst-forming Cyanobacterium Anabaena sp. strain PCC 7120. *Journal of Bacteriology*, *191*(7), 2353–2361.

49. Cai, Y. P., & Wolk, C. P. (1990). Use of a conditionally lethal gene in Anabaena sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. *Journal of Bacteriology*, *172*(6), 3138–3145.

50. Wendt, K. E., Ungerer, J., Cobb, R. E., Zhao, H., & Pakrasi, H. B. (2016). CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium Synechococcus elongatus UTEX 2973. *Microbial Cell Factories*, *15*(1), 115.

51. Ungerer, J., & Pakrasi, H. B. (2016). Cpf1 Is A Versatile Tool for CRISPR Genome Editing Across Diverse Species of Cyanobacteria. *Scientific Reports*, *6*, 39681.

52. Agarwal, V., Blanton, J. M., Podell, S., Taton, A., Schorn, M. A., Busch, J., Lin, Z., Schmidt, E. W., Jensen, P. R., Paul, V. J., Biggs, J. S., Golden, J. W., Allen, E. E., & Moore, B. S. (2017). Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. *Nature Chemical Biology*, *13*(5), 537–543.

53. Yamanaka, K., Reynolds, K. A., Kersten, R. D., Ryan, K. S., Gonzalez, D. J., Nizet, V., Dorrestein, P. C, 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 of the United States of America*, *111*(5), 1957–1962.

54. Bao, Z., Xiao, H., Liang, J., Zhang, L., Xiong, X., Sun, N., Si, T., & Zhao, H. (2015). Homology-Integrated CRISPR–Cas (HI-CRISPR) System for One-Step Multigene Disruption in Saccharomyces cerevisiae. *ACS Synthetic Biology*, *4*(5), 585–594.

55. Chen, Y., Taton, A., Go, M., London, R. E., Pieper, L. M., Golden, S. S., & Golden, J. W. (2016). Self-replicating shuttle vectors based on pANS, a small endogenous plasmid of the unicellular cyanobacterium Synechococcus elongatus PCC 7942. *Microbiology*, *162*(12), 2029–2041.

56. Singh, S. P., Rastogi, R. P., Häder, D.-P., & Sinha, R. P. (2011). An improved method for genomic DNA extraction from cyanobacteria. *World Journal of Microbiology and Biotechnology*, *27*(5), 1225–1230.

57. Elhai, J., Vepritskiy, A., Muro-Pastor, A. M., Flores, E., & Wolk, C. P. (1997). Reduction of conjugal transfer efficiency by three restriction activities of Anabaena sp. strain PCC 7120. *Journal of Bacteriology*, *179*(6), 1998–2005.

58. Gietz, R. D., & Schiestl, R. H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, *2*(1), 31–34.

59. Golden, S. S., & Sherman, L. A. (1984). Optimal conditions for genetic transformation of the cyanobacterium Anacystis nidulans R2. *Journal of Bacteriology*, *158*(1), 36–42.

60. Golden, S. S., Brusslan, J., & Haselkorn, R. B. T.-M. in E. (1987). [12] Genetic engineering of the cyanobacterial chromosome. In *Recombinant DNA Part D* (Vol. 153, pp. 215–231).

61. Taton, A., Lis, E., Adin, D. M., Dong, G., Cookson, S., Kay, S. A., Golden, S. S., & Golden, J. W. (2012). Gene Transfer in Leptolyngbya sp. Strain BL0902, a Cyanobacterium Suitable for Production of Biomass and Bioproducts. *PLOS ONE*, *7*(1), e30901.

62. Elhai, J., & Wolk, C. P. B. T.-M. in E. (1988). [83] Conjugal transfer of DNA to cyanobacteria. In *Cyanobacteria* (Vol. 167, pp. 747–754).

63. Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, *10*(3), R25.

64. Kaku, N., Hibino, T., Tanaka, Y., Ishikawa, H., Araki, E., Takabe, T., & Takabe, T. (2000). Effects of overexpression of Escherichia coli katE and bet genes on the tolerance for salt stress in a freshwater cyanobacterium Synechococcus sp. PCC 7942. *Plant Science*.

65. Kang, H.-S., Charlop-Powers, Z., & Brady, S. F. (2016). Multiplexed CRISPR/Cas9- and TAR-Mediated Promoter Engineering of Natural Product Biosynthetic Gene Clusters in Yeast. *ACS Synthetic Biology*, *5*(9), 1002–1010.

66. Schmidt, E. W., Nelson, J. T., Rasko, D. A., Sudek, S., Eisen, J. A., Haygood, M. G., & 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 of the United States of America*, *102*(20), 7315–7320.

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

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

69. Ziemert, N., Ishida, K., Liaimer, A., Hertweck, C., & Dittmann, E. (2008). Ribosomal Synthesis of Tricyclic Depsipeptides in Bloom-Forming Cyanobacteria. *Angewandte Chemie International Edition*, *47*(40), 7756–7759.

70. Ongley, S. E., Bian, X., Zhang, Y., Chau, R., Gerwick, W. H., Müller, R., & Neilan, B. A. (2013). High-Titer Heterologous Production in E. coli of Lyngbyatoxin, a Protein Kinase C Activator from an Uncultured Marine Cyanobacterium. *ACS Chemical Biology*, *8*(9), 1888–1893.

71. Liu, T., Mazmouz, R., Ongley, S. E., Chau, R., Pickford, R., Woodhouse, J. N., & Neilan, B. A. (2017). Directing the Heterologous Production of Specific Cyanobacterial Toxin Variants. *ACS Chemical Biology*, *12*(8), 2021–2029.

72. Jones, A. C., Ottilie, S., Eustáquio, A. S., Edwards, D. J., Gerwick, L., Moore, B. S., & Gerwick, W. H. (2012). Evaluation of Streptomyces coelicolor A3(2) as a heterologous expression host for the cyanobacterial protein kinase C activator lyngbyatoxin A. *The FEBS Journal*, *279*(7), 1243–1251.

73. 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., & Yoon, Y. J. (2012). Heterologous Production of 4-O-Demethylbarbamide, a Marine Cyanobacterial Natural Product. *Organic Letters*, *14*(23), 5824–5827.

74. Wang, B., Wang, J., Zhang, W., & Meldrum, D. (2012). Application of synthetic biology in cyanobacteria and algae. *Frontiers in Microbiology*, Vol. 3, p. 344. Retrieved from

75. Wilkinson, B., & Micklefield, J. (2007). Mining and engineering natural-product biosynthetic pathways. *Nature Chemical Biology*, *3*(7), 379–386.

76. Muenchhoff, J., Siddiqui, K. S., & Neilan, B. A. (2012). Identification of two residues essential for the stringent substrate specificity and active site stability of the prokaryotic l-arginine:glycine amidinotransferase CyrA. *The FEBS Journal*, *279*(5), 805–815.