2015

Characterization of a Biosynthetic Pathway Yielding Anticancer Natural Products from a Marine Bacterium

Elle D. James
University of North Florida, ellejames89@gmail.com

CHARACTERIZATION OF A BIOSYNTHETIC PATHWAY YIELDING ANTICANCER NATURAL PRODUCTS FROM A MARINE BACTERIUM

by

Elle D. James

A thesis submitted to the Department of Biology
in partial fulfillment of the requirements for the degree of

Master of Science in Biology

University of North Florida

College of Arts and Sciences

July 2015

Certificate of Approval

The thesis of Elle D. James is approved: (Date)

_____ _____
Dr. Amy L. Lane, Ph.D.

_____ _____
Dr. Terri Ellis, Ph.D.

_____ _____
Dr. Judith Ochrietor, Ph.D.

Accepted for the Biology Department:

_____ _____
Dr. Cliff Ross, Ph.D.
Chair

Accepted for the College of Arts and Sciences:

_____ _____
Barbara A. Hetrick, Ph.D.
Dean

Accepted for the University:

_____ _____
Dr. John Kantner
Dean of the Graduate School

Acknowledgements

First, I would like to express immense gratitude for my advisor, Dr. Amy Lane, who has so generously shared her contagious passion for research. I am grateful for the example she has set by her leadership, patience, and dedication, all of which have molded me into the scientist I am today. I would like to thank my committee members, Dr. Terri Ellis and Dr. Judith Ochrietor, who have never hesitated to provide insight and guidance throughout this journey. I would like to thank Dr. Bryan Knuckley, whose generosity and willingness to troubleshoot and assist my research has played an instrumental role in my success. I am grateful for the collaborative work and support from Dr. Rajesh Viswanathan, who has worked tirelessly to share our results with the scientific community. I would also like to acknowledge the faculty and staff from both the Biology and Chemistry departments who create a positive, light-hearted and comical environment to work. I am grateful for funding from both the UNF Biology Department and UNF Graduate School that has supported my research endeavors and allowed me to travel to share my research with other scientists.

Special thanks goes to the members of the Lane lab, new and old, many of which have become great friends and outlets in times of “no one understands what I do!” I owe a great deal of my success to our ability to have fun and laugh, even on the most frustrating of days in the pursuit to achieve our goals.

I would like to express my appreciation for my family and friends who have loved, encouraged and cheered me on along the way. Your support is invaluable to where I am today. Last but not least, I’d like to thank Will for always encouraging me to pursue my goals no matter the cost. Thank you for being the light and laughs at the end of the tunnel every day.

Table of Contents

List of Figures and Tables

Abstract

Natural products are bioactive secondary metabolites produced by living organisms and are prevalently utilized as pharmaceutical drugs. Marine adapted organisms are a promising source of new natural products possessing unique chemical structures and biological activities. By studying the biosynthetic pathways employed by living organisms to produce natural products, insights into new strategies to generate molecules to combat disease and overcome drug resistance may be gained. This thesis study aimed to uncover the biosynthetic pathway employed by a marine actinomycete, *Nocardiopsis* sp. CMB-M0232, to catalyze the assembly of the nocardioazines. These molecules are a group of 2,5-diketopiperazine natural products that feature structurally unique functional groups. Nocardioazine A, the hypothesized end product of the nocardioazine biosynthetic pathway, exhibits anticancer activity. Bioinformatics analyses revealed three biosynthetic gene clusters from *Nocardiopsis* encoding proteins with hypothesized roles in nocardioazine A biosynthesis. Two cyclodipeptide synthases (CDPSs), NozA and NcdA, were biochemically characterized *in vivo* and *in vitro* to reveal that both are substrate specific enzymes that utilize tryptophan-charged tRNA substrates to catalyze assembly of *cyclo*(L-Trp-L-Trp), a proposed precursor of nocardioazines. Fidelity is uncommon amongst characterized CDPSs, making NozA and NcdA important CDPS family additions. This study also aimed to characterize NozD and NozE, two cytochrome P450 homologs with predicted roles as diketopiperazine-tailoring enzymes. Heterologous expression of these enzymes in *Streptomyces* strains was not able

to confirm the functions of NozD and NozE but set the stage for future studies to optimize conditions for probing their roles in nocardioazine A biosynthesis. The results gathered from this study, along with future work to better understand the engineering of unique functional groups from *Nocardiopsis* may provide opportunities to produce new bioactive molecules.

Chapter I

Introduction

Marine microorganisms: goldmines of natural products

Modern medicine requires an ongoing search for new, clinically useful drugs to combat disease and overcome drug resistance. It is estimated that less than one third of over 30,000 human diseases have treatment options and those options continue to diminish as antibiotic resistance increases at a faster rate than the development of new antibiotics (Subramani et al. 2012, Netzker et al. 2015). A promising source of new drugs is natural products, bioactive secondary metabolites produced by living organisms. These molecules can be utilized as medicines such as antibacterials, anticancer agents and anti-inflammatory drugs (Newman and Cragg, 2012).

Primary metabolites are molecules directly involved in the growth, development, and reproduction of an organism; these include fats, proteins, carbohydrates and nucleic acids. Natural products are secondary metabolites, which are not necessary for growth, development or reproduction, but are commonly used as agents to increase their competitive fitness. For example, some natural products are utilized as chemical defenses for organisms to adapt to their environment and survive amongst predators or pathogens (Dias et al. 2012). Organisms, including plants, fungi, algae and bacteria, produce natural products by way of biosynthesis, enzyme-catalyzed processes where simple chemical building block substrates are converted into more complex products

(Dias et al. 2012). Secondary metabolites are generally assembled by specialized metabolic pathways that utilize intermediates from primary metabolic processes such as photosynthesis and glycolysis. These biosynthetic pathways presumably evolved due to selective pressure from biotic and abiotic factors such as viruses, environmental shifts and radiation, producing new metabolites that allow organisms to adapt and prolong their existence. Because living organisms adapt to and overcome a diverse variety of environments and predators, the secondary metabolites produced tend to be unique to the organism (Dias et al. 2012).

Historically, natural products have been the leading source of new drugs, offering novel and complex chemical structures that are often too difficult to produce via traditional synthetic chemistry. The 21st century has presented great strides in the field of natural product drug discovery due to advances in molecular biology, bioinformatics and sequencing technologies. These advances provide fast access to genomic information, identification of biosynthetic pathways and the necessary tools to utilize this information and facilitate drug discovery (Lane and Moore, 2011). Interdisciplinary collaborations such as those between biologists, chemists and engineers are also credited with the expansion of this field. Networking among experts has resulted in a better understanding of the biosynthesis of structurally novel, bioactive natural products as well as how that structure links them to their physiological and ecological roles (Dalisay et al. 2013). Use of these modern approaches has greatly facilitated scientists' efforts to study biosynthetic pathways and their products (Giessen and Marahiel, 2012).

For example, recombinant protein expression, used in the current thesis study, exploits expression hosts such as *E. coli* or yeast to study the products of enzymatic synthesis. This biomolecular approach to synthesis provides an environmentally friendly alternative to synthetic chemistry by using far less toxic chemicals, therefore offering a less harmful experimental strategy for scientists and for the environment (Dias et al. 2012).

Of the many organisms that yield natural products, microorganisms are the leading producers. An unrivaled number (>70%) of novel compounds have been derived from actinomycetes, bacteria that are members of the order *Actinomycetales* such as the genera *Actinomadura, Nocardiopsis, Salinispora* and *Stremptomyces* (Manivasagan et al. 2014). Most of these compounds were derived from terrestrial actinomycetes, which were found to be prolific producers of structurally and chemically diverse metabolites. However, following over four decades of exploration, terrestrial organisms have been largely exhausted as sources of new drugs due to repetitive discovery of identical compounds. Research has now shifted significantly toward marine actinomycetes, which are found in abundance throughout the ocean, from plant- and sponge-associates to ocean floor sediments (Dalisay et al. 2013, Manivasagan et al. 2013). Marine adapted organisms hold great potential as sources of new bioactive compounds possessing unique chemical structures and biological activities. It is hypothesized that differences in environmental conditions of marine and terrestrial species may result in different secondary metabolites in marine organisms compared to those found in terrestrial organisms. Seventy percent of the Earth's surface is covered

in water that may contain approximately one million cells per milliliter, setting the stage for nearly infinite sources of new, marine microorganisms and natural products (Subramani et al. 2012, Williams 2009).

Lane and Moore (2011) discussed the concept of the marine natural product “Dogma” to describe the connection between genes, enzymes and natural products when studying natural product biosynthetic pathways. Studying how organisms employ genes and enzymes to produce natural products offers insight into new ways to combat disease and overcome drug resistance by providing opportunities to modify existing natural products and discover new bioactive metabolites (Jensen and Fenical, 2006).

Natural product biosynthetic pathways: the keys to unlocking the potential of organisms to produce natural products

There are several obstacles faced in marine natural product drug discovery that have been overcome by advancements in molecular genetics and analytical tools. For example, despite the known presence of novel bioactive compounds in marine actinomycetes, their use as drugs is hindered due to the minimal quantities often produced by the organisms. Modern mass spectrometry (MS) enables the detection of these miniscule quantities of secondary metabolites, and offers understanding of molecular formulas and structural insights. Even with the knowledge of their presence, the isolation of compounds in reasonable quantities for drug development from the native organisms is highly

challenging. A historical remedy to this issue has been chemical synthesis, but synthetic approaches face roadblocks in the case of many natural products due to their complex chemical structures. A study presented by Long et al. (2005) represents one of the first examples of gene-based strategies to biosynthesize cyclic dipeptides, the patellamides, from the cyanobacterium, *Prochloron didemni.* The researchers cloned the entire patellamide biosynthetic gene cluster from *P. didemni* into a heterologous host, *E. coli*, biosynthesizing the petallamides in larger quantities than available from the native producing organism. They also discovered that upon mutating single genes in the biosynthetic pathway, several variations of the cyclic-peptide natural products were synthesized, exploiting what is referred to as combinatorial biosynthesis.

An alternative to *in vivo* biosynthesis is *in vitro* biosynthesis, an approach that exploits enzymes as biocatalysts to generate complex metabolites outside of a heterologous host. By combining such enzymes with desired precursors *in vitro*, studies have been able to achieve total synthesis of many structurally complex compounds. In addition, *in vitro* biosynthesis presents the opportunity to exploit enzymes to decorate existing natural products with new functional groups, and therefore alter their bioactivity. These types of molecular biology approaches have allowed us to elucidate the origins of many secondary metabolites, exploit biosynthetic machinery to produce greater quantities of secondary metabolites and produce new, bioactive derivatives of known compounds (Gulder and Moore, 2010, Lane and Moore, 2012, Smanski et al. 2012).

Another obstacle that recent studies have begun to investigate is the notable difference between the large presence of bioinformatics-predicted biosynthetic gene clusters from microbial genomes compared to the dramatically smaller number of secondary metabolites observed under standard laboratory conditions (Netzker et al. 2015). For example, genome sequencing of the marine actinomycete *Salinispora tropica* revealed that almost 10% of its 5-Mbp genome was predicted to have involvement in the biosynthesis of at least 19 different types of secondary metabolites. Contrary to this number of compounds capable of being produced, standard laboratory conditions only yielded four different classes of secondary metabolites from *S. tropica* (Gulder and Moore, 2009; Lane and Moore, 2011). Application of molecular biology techniques, such as gene elimination mutations, has led to discovery of additional metabolites encoded by the genome (Lane and Moore, 2011). This observation has been noted among genomes of many bacteria and fungi, suggesting a wealth of new bioactive metabolites that are not accessible by standard fermentation conditions in the laboratory (Challis, 2007, Giessen and Marahiel, 2012).

Several hypotheses exist for the evolutionary maintenance of such energy consuming secondary metabolic pathways, with roles described in habitat defense, predator inhibition and as intraspecific signaling molecules. The majority of secondary metabolic pathways appear to be conditionally expressed, only turned on under specific environmental conditions (Netzker et al. 2015). This fine-tuned pathway expression may offer many benefits to the organism. For example, if secondary metabolite pathways were continually expressed in their

native environment, neighbors exposed would quickly develop resistance and therefore defeat the purpose of the metabolites, which are commonly used for defense.

General fermentation conditions used in laboratories do not typically trigger production of such compounds. By employing methods such as combinatorial biosynthesis and heterologous expression, researchers are beginning to unlock the mechanisms to express these silent gene clusters in organisms and reveal their potential as significant natural product producers (Ozaki et al. 2013). Through the coevolution of species and their secondary metabolites over time, nature may lead us to new, naturally evolved and modified pathways providing us new metabolites and the means to synthesize altered metabolites (Dias et al. 2012).

Nocardioazines: structurally unique natural products from a marine actinomycete

Nocardiopsis sp. CMB-M0232 is a marine actinomycete that was first isolated from sediment samples collected off the coast of Australia in 2009 by Dr. Robert J. Capon and co-workers. Capon and his research team discovered that *Nocardiopsis* sp. CMB-M0232 produced a plethora of interesting, biomedically active secondary metabolites (Raju et al. 2010, Raju et al. 2011, Raju et al. 2013). For example, nocardiopsins A-D were isolated from this strain, and are the only members of the macrolide-pipecolate natural product family discovered

in over ten years (Raju et al. 2011, Raju et al. 2013). The biosynthetic pathway responsible for the production of nocardiopsins A-D from *Nocardiopsis* sp. was recently elucidated by the Lane group (Bis et al. 2015).

Another novel class of compounds isolated from *Nocardiopsis* sp. are the nocardioazines (Figure 1), novel prenylated 2,5-diketopiperazines (DKPs) (Figure 2) (Raju et al. 2011). DKPs are secondary metabolites known for their potential medical application as antibacterials, antitumor agents and immunosuppressives. Special attention has been placed on 2,5-DKP natural products due to the large number isolated from microorganisms. DKPs are the simplest cyclic peptides, formed when two amino acids undergo condensation, resulting in the formation of two peptide bonds that produce a cyclodipeptide (CDP) (Figure 2) (Borthwick 2012, Bonnefond et al. 2011, Gondry et al. 2009).

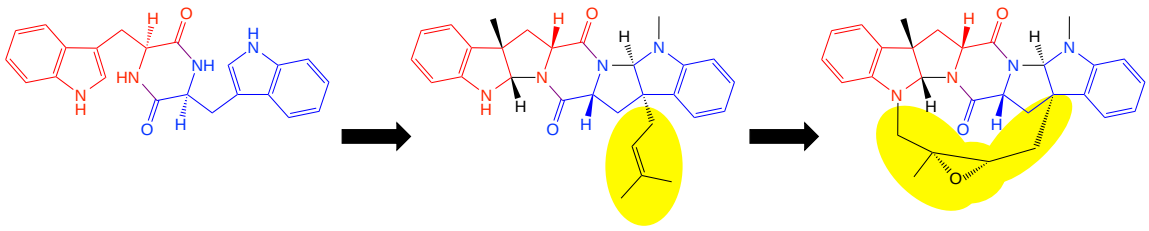


Figure 1. 2,5-DKP-derived natural products from *Nocardiopsis* sp. CMB-M0232 presented in their hypothesized biosynthetic pathway: precursor, intermediate and product (left to right). Prenyl-derived groups are shown in yellow.

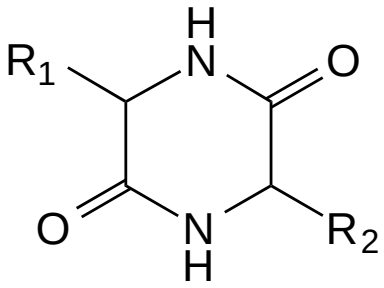


Figure 2. A 2,5-diketopiperazine. These are formed by the condensation of two α-amino acids, where R_1 and R_2 represent amino acid R groups.

Nocardiopsis sp. produces three notable DKP natural products when fermented under high salinity conditions: *cyclo*(L-Trp-L-Trp) and nocardioazines A-B (Figure 1). All three compounds share a common apparently tryptophan-derived 2,5-DKP core. This leads to the hypothesis that *cyclo*(L-Trp-L-Trp) is the precursor, nocardioazine B an intermediate, and nocardioazine A the final product in the nocardioazine biosynthetic pathway (Figure 1). Many DKP natural products have been identified, but nocardioazines A-B are unique as the only known C3-prenylated DKPs from a prokaryote and the only indole-C3-normal prenylated DKP from any source, suggesting the presence of a unique and novel biosynthetic pathway from *Nocardiopsis* sp. Prenyl groups are branched 5-carbon structures that are of considerable interest because of their prevalence amongst bioactive natural products (Raju et al. 2011). The unique prenylation of the nocardioazines strongly suggests that novel enzyme(s) carry out this biosynthetic transformation.

Furthermore, Dr. Capon's lab discovered that nocardioazine A functions as an anticancer agent (Raju et al. 2011). Many cancer cell lines become

resistant to antitumor agents by overexpressing the ABC transporter, P-glycoprotein (P-gp), which allows cells to eject the drug. Capon's group found that nocardioazine A reverses drug resistance in colon cancer cells by acting as a noncytotoxic inhibitor of P-gp, preventing drugs from leaving the cell. Nocardioazine A proved to work as effectively as verapamil, a synthetic P-gp inhibitor (Raju et al. 2011).

The objective of this thesis study was to identify and characterize selected genes and corresponding enzymes involved in the biosynthesis of nocardioazine A from *Nocardiopsis* sp. CMB-M0232. Bioinformatics was used to identify candidate biosynthetic gene clusters with predicted roles in nocardioazine assembly. Selected genes within the clusters were characterized through heterologous expression and profiling of metabolites produced *in vivo,* as well as *in vitro* assays with purified pathway enzymes. These experiments established the enzymatic basis for assembly of *cyclo*(L-Trp-L-Trp) as the precursor of nocardioazine A-B and provided additional candidate biosynthetic enzymes for characterization in future studies.

Chapter II

Characterization of two cyclodipeptide synthases from *Nocardiopsis* sp. CMB-M0232

Introduction

An impressive number of naturally occurring, structurally diverse DKPs have been discovered, but their biosynthetic pathways often remain poorly understood (Giessen and Marahiel, 2014). For years, a class of enzymes referred to as nonribosomal peptide synthetases (NRPSs) were recognized as catalysts for DKP assembly. NRPSs are highly studied, relatively massive (>100-kDa) multimodular enzymes organized so that each module is involved in a single stage of synthesis beginning with a single amino acid modified one step at a time to produce a final peptide product. NRPSs are responsible for biosynthesis of the cyclized amino acid core of DKPs via direct, dedicated DKP-yielding NRPSs or by prematurely releasing a DKP intermediate from the NRPS assembly line for a more complex peptide natural product (Giessen et al. 2013).

In 2009, a new class of enzymes referred to as cyclodipeptide synthases (CDPSs) were first reported to also produce DKPs. CDPSs are unique in their small size (~30-kDa) relative to massive NRPSs that produce analogous molecules. Further, CDPSs use a mechanism that is distinct from what is used by NRPSs to form the peptide bond between two amino acids. This recently discovered family of enzymes does not employ several modules to activate and incorporate free amino acids. Instead, CDPSs ‘hijack’ aminoacyl-tRNAs (aa-

tRNAs) from the ribosomal machinery in primary metabolism to directly catalyze formation of peptide bonds to yield 2,5-DKPs (Figure 3) (Gondry et al. 2009, Belin et al. 2014, Moutiez et al. 2014). By using aa-tRNAs as substrates, CDPSs bypass the amino acid activation step that NRPSs use and instead redirect amino acids away from the ribosome, linking primary and secondary metabolism (Giessen and Marahiel, 2012, Bonnefond et al. 2011, Giessen and Marahiel, 2014).

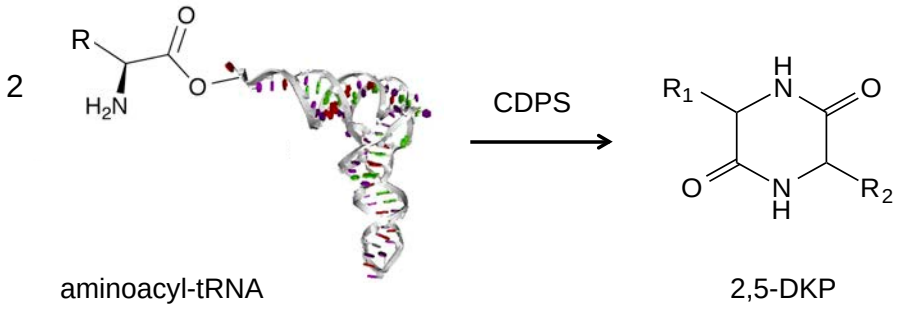


Figure 3. CDPSs are a relatively newly described enzyme family that employs aminoacyl-tRNAs as substrates for 2,5-DKP assembly. R refers to an amino acid R group.

To date, over 150 putative CDPS enzymes are logged in the GenBank database, but less than 12 have been biochemically characterized (Giessen and Marahiel, 2014). Thus, much remains to be known about this apparently common biosynthetic enzyme class. Protein sequence comparisons of biochemically characterized CDPSs reveal varying degrees of identity ranging from 19-70%; only 8 residues are conserved across all functionally characterized CDPSs and no consensus sequences are noted for these enzymes (Figure 5). Despite variation in protein sequences, x-ray structures of the three crystallized CDPSs

superimpose relatively well, providing evidence for conserved amino acid residues appearing in the active site and playing critical roles in the substrate specificity of the enzymes. CDPSs generally exhibit promiscuity with regard to aa-tRNA substrates, and most produce several 2,5-DKP products. Although most CDPSs can produce several DKPs, they tend to preferentially synthesize a single DKP over other products. Single point mutations in *albC* from *Streptomyces noursei* were sufficient to alter the specificity of the enzyme and revealed the one-at-a-time incorporation of aminoacyl charged-tRNAs in the active site (Moutiez et al. 2014). Moutiez et al. (2014) also recently unveiled the process of how CDPSs catalyze reactions with aa-tRNA substrates using as a sequential ping-pong mechanism (Figure 4).

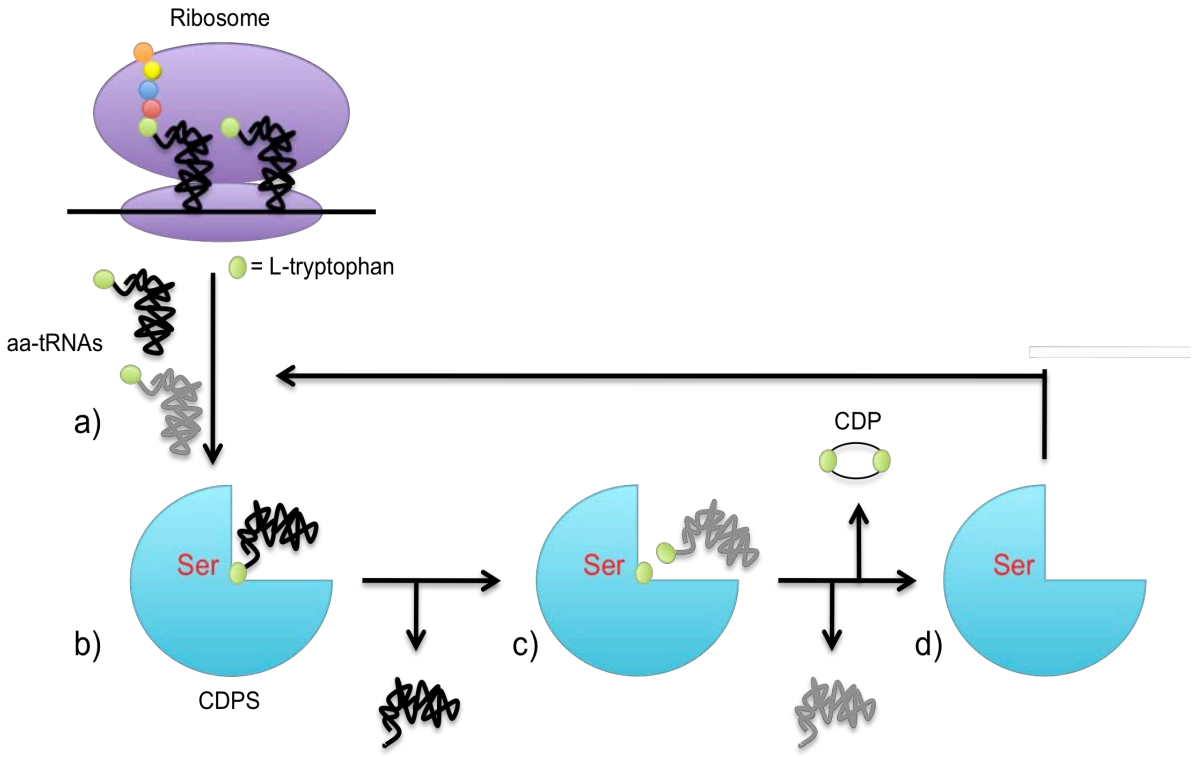


Figure 4. Overview of the mechanism employed by cyclodipeptide synthases to catalyze the production of DKPs. (a) Activated aa-tRNA's are sequestered from primary metabolism. (b) Binding of the first aminoacyl moiety from an aa-tRNA to the CDPS's serine active site. (c) The first aminoacyl group forms a peptide bond with an aminoacyl moiety from a second aa-tRNA substrate. (d) The resulting dipeptidyl intermediate undergoes intramolecular peptide bond formation to close the DKP ring and releases it from the active site.

Generally, prokaryotic CDPSs produce DKPs that incorporate phenylalanine, leucine, tyrosine and/or methionine. Only one CDPS from eukaryotic origin has been identified and that CDPS, Nvec_CDPS2, is able to incorporate the same four amino acids as well as tryptophan. DKPs that incorporate tryptophan are recognized as promising bioactive metabolites and are relatively common molecules despite the scarce number of identified CDPSs that incorporate tryptophan. Thus, this has left open the question of how living things yield the wealth of observed Trp-containing DKPs. Recently, Giessen et al.

(2013) characterized Amir_4627, a CDPS from *Actinosynnema mirum,* and found this enzyme as the first to catalyze the production of *cyclo*(L-Trp-L-Trp). This discovery also provided the first substrate specific CDPS and the first CDPS identified from prokaryotic origin to catalyze formation of a Trp-containing DKP.

In this chapter we aim to identify biosynthetic genes and corresponding enzymes from the draft genome of *Nocardiopsis* sp. CMB-M0232 with hypothesized roles in the first step of nocardioazine A biosynthesis, the catalysis of *cyclo*(L-Trp-L-Trp) assembly. It is understood that genes involved in the production of a specific secondary metabolite are generally clustered in the genomes of bacteria; this facilitates the identification of biosynthetic pathways (Giessen and Marahiel, 2012). By identifying the enzyme(s) responsible for the production of *cyclo*(L-Trp-L-Trp), we may locate other genes involved in nocardioazine A biosynthesis nearby and thus gain access to the entire biosynthetic pathway.

Methods

Genome sequencing and bioinformatics analyses of Nocardiopsis sp. draft genome to locate candidate enzymes from the nocardioazine A biosynthetic pathway

A clone library of *Nocardiopsis* sp. CMB-M0232 genomic DNA was created according to procedures described in Bis et al. (2015). Sequencing of *Nocardiopsis* sp. CMB-M0232 gDNA was performed by Cofactor Genomics (St.

Louis, MO) using Illumina with paired-end 80 bp reads and 454 with single end 400 bp reads. BLASTP was used to analyze ORFs for homology to nonribosomal peptide synthases and cyclodipeptide synthases, putative enzymes in *cyclo*(L-Trp-L-Trp) biosynthesis. Geneious Bioinformatics Software (Biomatters) and BLASTP were used to compare protein sequence homology between CDPSs (ClustalW). The phylogenetic tree was created using Geneious Treebuilder with no outgroup.

Generation of constructs for recombinant expression of Nocardiopsis sp. CDPS genes, nozA and ncdA, in E. coli

The genes encoding NozA and NcdA from *Nocardiopsis* sp. draft genome sequence contig 557 and contig 96, respectively, were codon optimized for expression in *E. coli* and were cloned into the pMA-T vector by GeneArt (Life Technologies) (Appendix A-B). The pMA-T/*nozA* and pMA-T/*ncdA* constructs were introduced into *E. coli* JM109 by heat shock (Green and Sambrook, 2012). Isolated colonies were used to inoculate LB media (10 g/L peptone, 5 g/L yeast extract, 10 g/L sodium chloride) supplemented with ampicillin (100 µg/mL) and were cultured at 37 °C and 200 rpm overnight. Plasmid DNA was purified from the cultures using a Qiagen Miniprep Kit and isolated DNA was then digested in a reaction mixture containing 1.0 µL (~1 µg) of the purified plasmid DNA (pMA-T/*nozA* or pMA-T/*ncdA*), 4.0 µL of 10X Thermo FastDigest Green Buffer, 2.0 µL of restriction enzymes (Thermo Scientific) and 31.0 µL of molecular biology grade water to prepare the CDPS genes for ligation into expression vectors. For

purposes of *in vivo* characterization of enzyme function, both *nozA* and *ncdA* were introduced into pQE30 expression vectors and *E. coli* M15[pREP4] via protocols that follow. The expression of NcdA was too low in pQE30 for the *in vitro* experiments, so the following methods were used to instead introduce *ncdA* into pET30a(+) expression vector and *E. coli* BL21 (DE3).

The pMA-T/*nozA* was digested with SacI and HindIII and pMA-T/*ncdA* was digested with NdeI and HindIII. The reactions were incubated at 37 °C for 1 hour followed by heat inactivation at 70 °C for 10 minutes. Expression vectors, pQE30 (Qiagen) and pET30a(+) (Novagen), were digested in reactions containing 22 µL of molecular biology grade water, 4 µL of 10X FastDigest Green Buffer, 10 µL (~1 ug) of purified plasmid (pQE30 or pET30a(+)) and 2 of µL of each restriction enzyme. The reactions incubated at 37 °C for 45 minutes followed by heat inactivation at 70 °C for 10 minutes. The digests were purified using the Qiagen QIAquick PCR Cleanup Kit and then subjected to dephosphorylation in reactions containing 20 µL of molecular biology grade water, 4 µL of Alkaline Phosphatase 10X Buffer (Promega), 15 µL (~15 µg) of restriction digested vector and 1 µL of Calf Intestinal Alkaline Phosphatase (Promega). The reactions were incubated at 37 °C for 1 hour followed by heat inactivation at 70 °C for 15 minutes. Digested CDPSs and expression vectors were verified by agarose gel electrophoresis and the appropriate sized DNA fragments were cut from the gels and purified using the Qiagen QIAquick Gel Extraction Kit.

A ligation reaction to anneal the synthetic *nozA* gene with pQE30 contained 50 ng of digested pQE30, 37.5 ng of NozA insert DNA, 0.5 µL of 10X T4 DNA Ligase Buffer (Thermo Scientific), 0.5 µL of T4 DNA Ligase (Thermo Scientific) and sufficient molecular biology grade water to yield a final volume of 5 µL. The ligations were incubated overnight at 4 °C. The same reaction was set up to ligate the synthetic *ncdA* gene with the pET30a(+). The pQE30/*nozA* construct was propagated in *E. coli* JM109 and transformed into *E. coli* M15[pREP4] cells using standard heat shock methods (Green and Sambrook, 2012). The pET30a(+)/*ncdA* construct was propagated in *E. coli* JM109 and then transformed into *E. coli* BL21 (DE3) via heat shock. The constructs were confirmed by DNA sequencing (Operon). As no biosynthetic gene controls, *E. coli* M15[pREP4] and *E. coli* BL21 (DE3) were transformed with pQE30 and pET30a(+) plasmid vector, respectively.

In vivo characterization of functions of NozA and NcdA

The pQE30/*nozA* construct in *E. coli* M15[pREP4] was cultured in 10 mL LB broth supplemented with ampicillin (100 µg/mL) and kanamycin (25 µg/mL). The pET30a(+)/*ncdA* construct in *E. coli* BL21 (DE3) was cultured in 10 mL LB broth supplemented with kanamycin (25 µg/mL). Control *E. coli* M15[pREP4] pQE30 and *E. coli* BL21 (DE3) pET30a(+) containing no biosynthetic genes were cultured equivalently to treatments. All treatment and control cultures were incubated at 37 °C and 200 rpm until an optical density (OD_{600}) of 0.4 was

reached. Aliquots of each culture (500 µL) were taken to analyze protein expression prior to induction. The remaining ~9.5 mL cultures were induced with a final concentration of 1 mM IPTG and were incubated at 19 °C and 240 rpm for an additional 19 hours. Following incubation, aliquots of each culture (500 µL) were taken to analyze protein expression following induction. The remaining ~9.0 mL cultures were centrifuged at 6,000 rpm for 10 minute and the supernatants were set aside for metabolite profiling by HPLC (described below).

Cell pellets (corresponding to 500 µL culture volume) taken before and after induction with IPTG were subjected to -80 °C for 1 hour to facilitate cell lysis and then resuspended in 200 µL of B-PER (Bacterial Protein Extraction Reagent, Thermo Scientific) with lysozyme (1 mg/mL). The homogenates were incubated in a 30 °C water bath for 30 minutes and then centrifuged at 13,000 rpm for 6 minutes. Supernatant (20 µL) from each sample was combined with 4X LDS Non-Reducing Sample Buffer (Thermo Scientific) and subjected to SDS-PAGE to evaluate expression of recombinant NozA and NcdA.

For analysis of the DKP metabolite profiles from treatment and control cultures, supernatants (~9 mL) were filter sterilized (0.2 uM cellulose acetate membrane) into a scintillation vial and dried using a speedvac concentrator (Savant ISS1100, Thermo Scientific) overnight. The dried supernatants were resuspended in 500 µL MilliQ water and were transferred to HPLC vials. HPLC was conducted using an Aligent 1100 HPLC system with a ZORBAX SB-C18 column (4.6 x 150 mm, particle size: 5 µm, pore size of 110 Å). A flow rate of 0.6

mL/min was used, with solvent A consisting of MilliQ water, solvent C consisting of acetonitrile and a linear gradient of 0 to 95% solvent C over 50 minutes following a 3 minute hold at 100 % solvent A. Synthetic *cyclo*(L-Trp-L-Trp) (0.2 mg/mL) was run as a positive control and provided by Alqahtani et al. (2015). Metabolite profiles were compared using ChemStation (Agilent).

Site directed mutagenesis of NozA at residues conserved among NozA, NcdA and Amir_4627

NozA was mutated at several residues unique to L-tryptophan-specific CDPSs to determine whether the residues played a role in catalysis of *cyclo*(L-Trp-L-Trp) formation. Six primer sets were designed with the goal of yielding single and double amino acid point mutations (S36A, L131R, S198L, N114R, N114A, V111R+N114A) by PCR mutagenesis. The PCRs contained 17.9 of µL molecular biology grade water, 0.1 µL of each forward and reverse primer (Appendix C, SDM_Mutant_F/R) (0.5 µM final concentration), 5 µL of 5X Phusion GC Buffer (Thermo Scientific), 0.5 µL of High Fidelity dNTPs (Roche, 10 mM final concentration), 0.75 µL of DMSO, 0.5 µL of template gDNA (0.4 µg) (pQE30/*nozA*) and 0.125 µL of Phusion DNA Polymerase (Thermo Scientific). The PCRs ran for an initial denaturation of 98 °C for 30 s, 25 cycles of 98 °C for 10 s and 72 °C for 60 s, and a final extension of 72 °C for 5 min with a hold at 4 °C. The PCR products were verified for size (4.1 Kb) by agarose gel electrophoresis and the appropriate sized DNA fragments were cut from the gels

and purified using the Qiagen QIAquick Gel Extraction Kit. The mutated constructs were ligated on themselves in a reaction containing 2.5 µL (~50 ng) of the gel purified constructs, 3 µL of 2X DNA ligase buffer and 0.5 µL of T4 DNA ligase (Thermo Scientific) and were incubated at 25 °C for 8 min. The cyclized constructs were introduced into *E. coli* JM109 for propagation, purified by miniprep and then transformed into *E. coli* M15[pREP4] by heat shock (Green and Sambrook, 2012) and incubated on LB agar with ampicillin (100 µg/mL) and kanamycin (25 µg/mL) at 37 °C overnight. The pQE30 vector containing no biosynthetic genes served as the negative control and pQE30/*nozA* served as the positive control.

Control constructs were transformed in *E. coli* M15[pREP4] and were treated identically to the mutated constructs in *E. coli* M15[pREP4]. In order to run duplicate experiments, two colonies from each agar plate (*E. coli* M15[pREP4] pQE30/*nozA*-S36A, -L131R, -S198L, -N114R, -N114A, -V111R+N114A and controls) were inoculated in 10 mL LB broth with ampicillin (100 µg/mL) and kanamycin (25 µg/mL) and were incubated at 37 °C and 200 rpm until an optical density (OD_{600}) of 0.4 was reached. The cultures were induced with a final concentration of 1 mM IPTG and were incubated for an additional 20 hours at 19 °C and 240 rpm. Following incubation, the optical density (OD_{600}) was measured to ensure the site directed mutations did not interfere with *E. coli* growth. The cultures were centrifuged at 6,000 rpm for 10 minutes and the supernatants were filtered through a 0.2 µM cellulose acetate membrane. For LC/MS analysis, 450 µL of the supernatants were combined with

50 µL of a synthetic *cyclo*-3-methyl-L-Trp-N1'-methyl-L-Trp DKP internal standard synthesized by Alqahtani et al. (2015). The remaining supernatants were dried using a speedvac concentrator (Savant ISS1100, Thermo Scientific) overnight, resolubilized in 500 µL MilliQ water and subjected to analytical HPLC as described above.

Expression and purification of NozA

From an overnight culture of *E. coli* M15[pREP4] transformed with pQE30/*nozA,* 1 mL was used to inoculate 500 mL LB medium supplemented with ampicillin (100 µg/mL) and kanamycin (25 µg/mL). The culture was incubated for 4 hours at 37 °C and 200 rpm until an optical density (OD_{600}) of ~0.4 was reached. The culture was induced with a final concentration of 1 mM IPTG and incubated for an additional 16 hours at 16 °C and 240 rpm. Cells were harvested by centrifugation (6,000 rpm, 4°C), resuspended in 16 mL lysis buffer (100 mM Tris HCl, 150 mM NaCl, 5% glycerol, 2 mM DTT, pH 8) and lysozyme (20 mg) and were incubated at -20 °C and 200 rpm for 30 minutes. Following incubation, the suspension was sonicated (Branson Digital Sonifier) at 60% amplitude for 2 minutes (1 s pulse:10 s rest). Cell debris was removed by centrifugation (12,000 rpm, 4 °C) and the soluble lysate was filtered through a 0.45 µM cellulose acetate membrane (Corning). The lysate, containing soluble proteins, was combined with 3 mL of HisPur Ni-NTA Resin (Thermo Scientific), spiked with a final concentration of 5 mM imidazole and incubated at 4 °C and 200 rpm for 30

minutes. Lysis buffer with increasing concentrations of imidazole (10 mM, 50 mM, 200 mM and 500 mM) was run through the column and collected. A sample (20 µL) of each Ni-column fraction was combined with 3 µL 4X LDS Non-Reducing Sample Buffer (Thermo Scientific), boiled briefly and loaded into a 12% Precise Protein Gel (Thermo Scientific) along with PageRuler Plus Prestained Protein Ladder (Life Technologies). The gel was stained with Imperial Protein Stain (Thermo Scientific) for 30 minutes and destained in Milli-Q water. The Ni-column fractions containing NozA (~28-kDa) were concentrated using an Amicon Ultra Centrifugal Filter Device with a 10,000 MWCO (Millipore) to a final volume of 500 µL and resuspended in 5 mL Buffer A (100 mM Tris HCl, 50 mM NaCl, 5% glycerol, 2 mM DTT, pH 8) for further purification via FPLC. The sample was loaded onto a HiTrap Q FF. (5x5 mL) anion exchange column (GE Healthcare) that was pre-equilibriated with buffer A and the protein was eluted with a gradient of 0 to 100% buffer B (100 mM Tris HCl, 1 M NaCl, 5% glycerol, 2 mM DTT, pH 8) for 50 mL at a rate of 1 mL/min. Protein elution was monitored at 280 nm and all fractions containing detectable protein were analyzed via SDS-PAGE as described previously. The fractions containing NozA (~28-kDa) were concentrated and the buffer was exchanged to assay buffer (50 mM Tris HCl, 300 mM NaCl, pH 8) using an Amicon Ultra Centrifugal Filter Device with a 10,000 MWCO. The purified protein was stored on ice while *in vitro* assays were prepared.

Expression and purification of NcdA

A -80 °C freezer stock of *E. coli* BL21 (DE3) transformed with pET30a(+)/*ncdA* was streaked onto LB agar supplemented with kanamycin (25 µg/mL) and incubated overnight at 37 °C. LB media (5 mL) supplemented with kanamycin (25 µg/mL) was inoculated with an isolated colony and grown at 37 °C and 200 rpm for ~4 hours. Following incubation, 1 L of LB broth with kanamycin was inoculated with 1 mL of this starter culture and grown at 37 °C and 200 rpm until an optical density (OD_{600}) of ~0.4 was reached. The culture was induced with a final concentration of 1 mM IPTG and incubated for an additional 16 hours at 19 °C and 240 rpm. The cells were harvested by centrifugation (6,000 rpm, 4°C) and were resuspended in 30 mL lysis buffer with lysozyme (37.5 mg). The 30 mL suspension was split between two 50 mL falcon tubes and each aliquot (15 mL) was sonicated at 60% amplitude for 2 minutes (1 s pulse:10 s rest) on ice. Cell debris was removed by centrifugation (12,000 rpm, 4 °C) and the soluble lysate was filtered through a 0.45 µM cellulose acetate membrane (Corning). The lysate was combined with 3mL HisPur Ni-NTA Resin (Thermo Scientific), spiked with a final concentration of 5 mM imidazole and incubated at 4 °C and 200 rpm for 30 minutes. Lysis buffer with increasing concentrations of imidazole (10 mM, 50 mM, 100 mM x 2, 200 mM x 2 and 500 mM) was run through the column and collected. An aliquot of each Ni-column fraction was analyzed by SDS-PAGE as described previously. The Ni-column fractions containing NcdA (~29-kDa) were concentrated using an Amicon Ultra Centrifugal Filter Device with a 10,000 MWCO (Millipore) to a final volume of 500 µL and resuspended in

5 mL Buffer A (100 mM Tris HCl, 50 mM NaCl, 5% glycerol, 2 mM DTT, pH 7.5) for further purification via fast protein liquid chromatography (FPLC). The sample was loaded onto a HiTrap Q FF. (5x5 mL) anion exchange column (GE Healthcare) pre-equilibriated with buffer A and the protein was eluted with a gradient of 0 to 30% buffer B (100 mM Tris HCl, 1 M NaCl, 5% glycerol, 2 mM DTT, pH 7.5) for 50 mL at a rate of 1 mL/min. Protein elution was monitored at 280 nm and all fractions containing detectable protein were analyzed via SDS-PAGE gel as described previously. The fractions containing NcdA (~29 kDa) were concentrated and the buffer was exchanged to assay buffer (50 mM Tris HCl, 300 mM NaCl, pH 8) using an Amicon Ultra Centrifugal Filter Device with a 10,000 MWCO. The purified protein was stored in 1 mM TCEP and 5% glycerol at -80 °C until *in vitro* enzyme assays.

In vitro characterization of NozA and NcdA functions and substrate specificity

To evaluate the function and substrate specificity of the *Nocardiopsis* CDPSs *in vitro*, 100 µL assays containing purified recombinant NozA and NcdA were incubated with tRNA from *E. coli* (50 µM, Sigma-Aldrich), *E. coli* aminoacyl-tRNA synthetase mix from *E. coli* (250 units, Sigma, Aldrich), L-Tryptophan (1 mM), ATP (5 mM), $MgCl_2$ (10 mM), KCl (30 mM) and DTT (2 mM) in Tris buffer (50 mM Tris HCl, 300 mM NaCl, pH 8). Control assays that (1) lacked enzyme, (2) lacked amino acids, (3) lacked tRNA and/or (4) lacked aminoacyl-tRNA

synthetase mix were run alongside treatments. All assays were set up at least in duplicate. All assay components other than the CDPS enzyme were pre-incubated at 37 °C for 20 minutes to allow charging of the tRNA with tryptophan. NozA assays were set up the same day as purification, so the enzyme was added to the pre-incubated assay directly following purification. Purified NcdA was thawed from -80 °C by incubation in 30 °C water prior to addition to assays.

To evaluate substrate specificity of NozA and NcdA, *in vitro* assays identical to those described above were designed with alternative aromatic amino acids (L-phenylalanine and L-tyrosine, 1 mM). Treatment and control assays were incubated overnight at 37 °C.

The reactions were stopped by the addition of 5 µL of trichloroacetic acid (100% w/v) and centrifugation at 13,000 rpm for 30 minutes to precipitate all proteins. The supernatants (90 µL) were transferred to HPLC vials and combined with 10 µL of methyl-*cyclo*(L-Trp-L-Trp) (1 µM), which served as an internal standard for quantification via HPLC and LC/MS as described previously.

Results and Discussion

Nocardioazine biosynthetic gene clusters hypothesized via bioinformatics analyses of Nocardiopsis sp. draft genome

Sequencing of *Nocardiopsis* sp. CMB-M0232 genomic DNA yielded a ~6.4 Mbp draft of the genome, from which over ~5500 open reading frames (ORFs)

were predicted using GeneMark (Besemer et al. 2001). These included 11 putative biosynthetic gene clusters, as predicted by AntiSmash (Weber et al. 2015). Analysis of the putative biosynthetic gene clusters by Basic Local Alignment Search Tool for proteins (BLASTP) revealed two candidate ORFs with similarity to nonribosomal peptide synthetases (NRPSs) and two with similarity to cyclodipeptide synthases (CDPSs). Both classes of enzymes have been previously found to be capable of catalyzing the formation of the DKP backbones, suggesting these enzymes might play roles in assembly of the predicted *cyclo*(L-Trp-L-Trp) precursor leading to nocardioazines A-B (Figure 2). Further bioinformatics analyses of the adenylation domains from the putative NRPSs revealed that none of the NRPSs were likely to incorporate two tryptophan substrates. Additionally, genes identified upstream and downstream from the putative NRPS are similar to those found in a hybrid polyketide synthase-nonribosomal peptide synthetase biosynthetic pathway known to produce natural products other than prenylated diketopiperazines. Outside the scope of this thesis, one of these NRPSs was found to play a role in nocardiopsin assembly (Bis et al. 2015).

NozA and NcdA are putative cyclodipeptide synthases with homology to Amir_4627, revealed via bioinformatics

The putative CDPSs NozA and NcdA, identified by bioinformatics analyses are the most likely candidates for production of the *cyclo*(L-Trp-L-Trp) DKP precursor in nocardioazine biosynthesis. Comparison of NozA and NcdA to

other known CDPSs revealed both enzymes share all 8 residues conserved among biochemically characterized CDPSs (Figure 5) (Moutiez et al. 2014). Of the conserved residues, serine-36 is conserved among the active sites and is necessary for recognition and covalent linkage of CDPSs to the aminoacyl group from amioacyl-charged tRNA substrates (Moutiez et al. 2014).

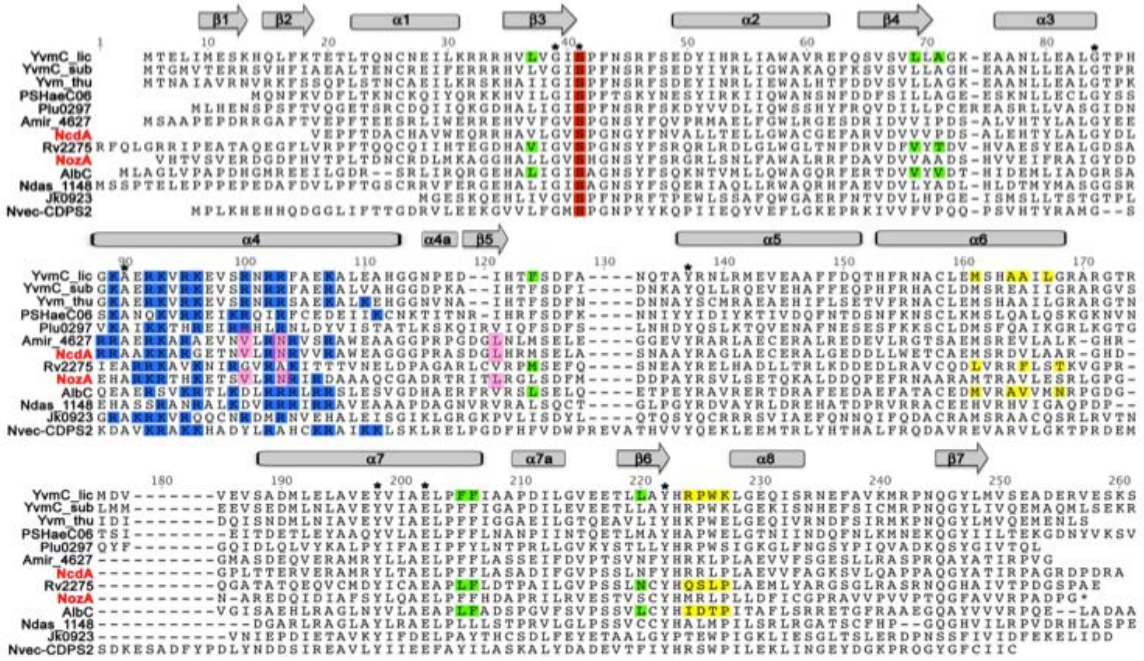


Figure 5. Alignment of CDPSs NozA and NcdA encoded by *Nocardiopsis* sp. CMB-M0232 with all biochemically characterized CDPSs. Secondary structures (grey) are based upon features conserved for the structures of crystallized CDPSs AlbC, Rv2275, and YvmC_lic. Residues conserved among all CDPSs are indicated with a *; these include the serine active site (red). For crystallized CDPSs, basic residues from the α4 helix (blue) play roles in recognition of the first tRNA group. Residues in green and yellow comprise the binding pocket of the first and second aminoacyl groups, respectively, for indicated enzymes. Residues in pink refer to those specific to NozA, NcdA and Amir_4627, the only substrate specific CDPSs the only ones known to catalyze assembly of *cyclo*(L-Trp-L-Trp).

Phylogenetic analysis of 11 previously characterized CDPSs based on protein sequence homology organized the enzymes into three monophyletic groups: one group incorporating aromatic amino acids, a second group representative of the only CDPS from eukaryotic origin, and a third group incorporating non-aromatic amino acids (Figure 6). NozA and NcdA were both most closely related to CDPSs that incorporate aromatic substrates. This comparison also revealed that NozA and NcdA share 35 and 67% identity, respectively, to Amir_4627 from *Actinosynnema mirum*, the only characterized CDPS that incorporates two Trp residues (Figure 6). Amir_4627 is also the only functionally characterized CDPS with specificity for one substrate. This relatedness is suggestive that NcdA (67% ID to Amir_4627) may exhibit similar activity to Amir_4627. The sequence of NozA, more distantly related to Amir_4627 (35% ID), is not as suggestive of biosynthesis of *cyclo*(L-Trp-L-Trp), but is likely to incorporate aromatic amino acids based on its phylogenetic relationship to other CDPSs catalyzing formation of aromatic DKPs (Figure 6). Based on these bioinformatics predictions, characterization of both NozA and NcdA was then undertaken to reveal the substrate specificity of both enzymes and whether they may be involved in nocardioazine A biosynthesis by catalyzing the production of the *cyclo*(L-Trp-L-Trp) precursor.

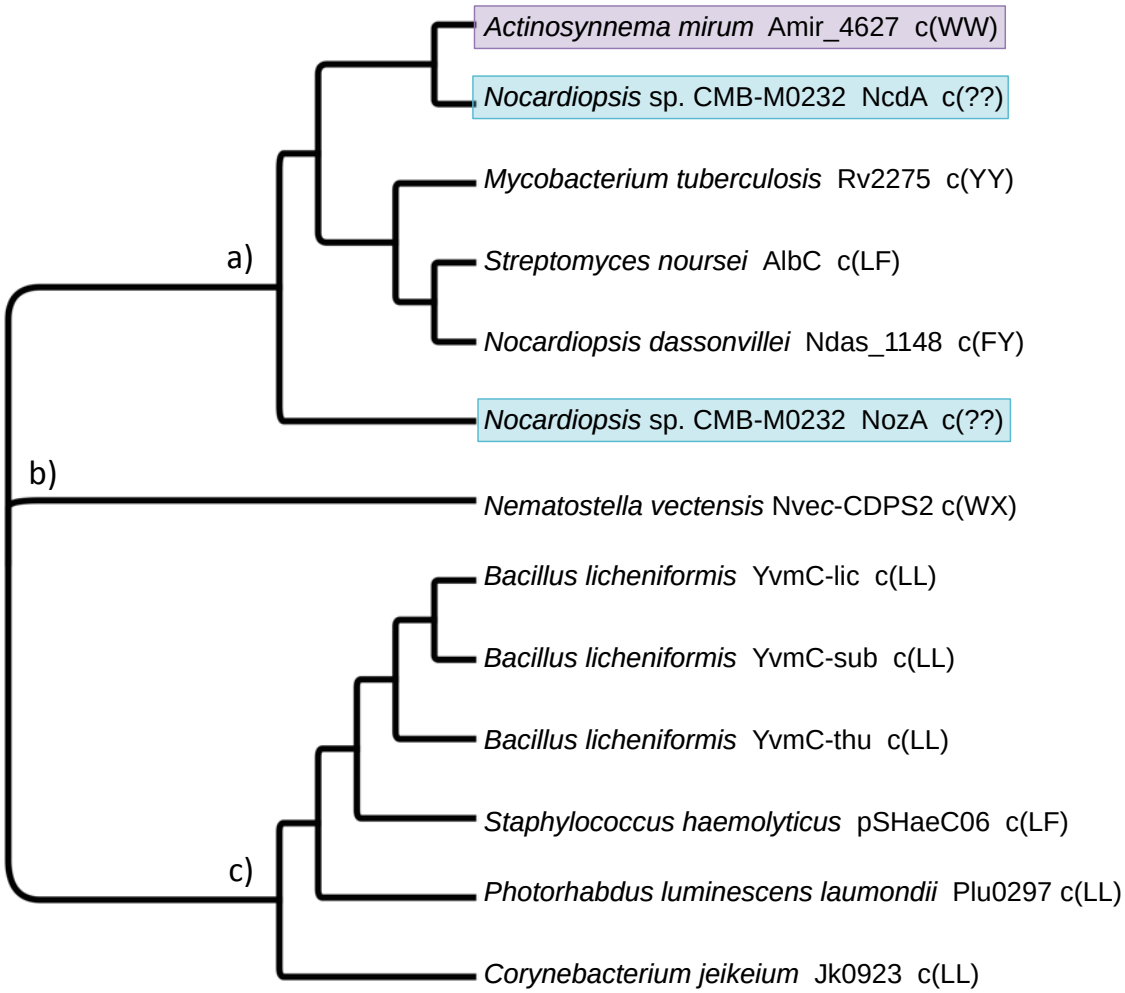


Figure 6. Phylogenetic tree of NcdA and NozA relative to all previously characterized CDPSs. The CDPSs organized into three monophyletic groups: (a) CDPSs that incorporate aromatic amino acids, (b) the single characterized CDPS from eukaryotic origin, and (c) CDPSs that incorporate non-aromatic amino acids. The only previously characterized CDPS that yielded *cyclo*(L-Trp-L-Trp), Amir_4627, is shown in purple. Amir_4627 and NcdA (top blue box) share 67% identity. Amir_4627 and NozA (bottom blue box) share 35% identity. Enzymes characterized in this current study are highlighted in blue. The only previously characterized CDPS that yielded *cyclo*(L-Trp-L-Trp) is shown in purple. The DKP product is shown with one-letter abbreviations for L-amino acids, e.g. *cyclo*(L-Trp-L-Trp) is listed as cWW. Most CDPSs are promiscuous; only the product formed in greatest yield is indicated. The Nvec-CDPS2 product is shown as cWX, since multiple amino acids were non-preferentially incorporated as the second residue of the DKP. Protein accession numbers are listed in the appendix (Appendix D).

Recombinant expression of NozA and NcdA in E. coli

In order to produce NozA and NcdA in greater quantities than what is provided by the native organism (*Nocardiopsis* sp.) and to facilitate enzyme purification through tagging with an affinity label (e.g. hexahistidine), the enzymes were recombinantly expressed in *E. coli.* Codon optimized *nozA* was introduced into the pQE30 plasmid vector. Codon optimized *ncdA* was introduced into plasmid vector pQE30 for *in vivo* characterization and pET30a(+) for *in vitro* characterization. The constructs were then introduced into *E. coli*. Agarose gel electrophoresis confirmed *E. coli* M15[pREP4] pQE30/*nozA, E. coli* M15[pREP4] pQE30/*ncdA*, and *E. coli* BL21 (DE3) pET30a(+)/*ncdA* were successfully constructed, showing both the vector and insert were present following purification of plasmid DNA and restriction digest (Figure 7). This finding was further confirmed by DNA sequencing, which revealed a perfect match between the expected and experimental DNA sequence of gene inserts.

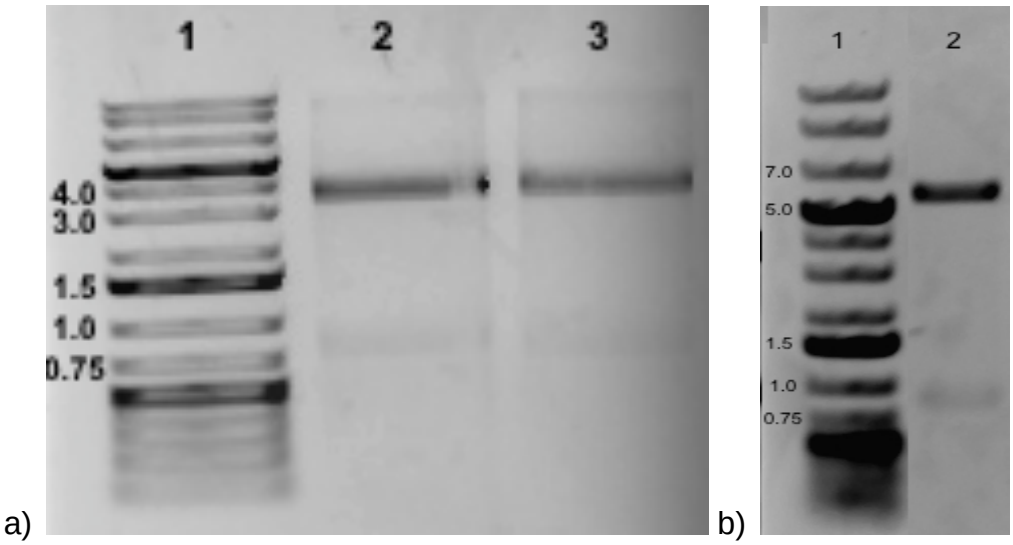


Figure 7. Verification of assembly of constructs for *E. coli* expression of *nozA* and *ncdA.* (a) Lane 1 contains GeneRuler 1 kb Plus DNA Ladder, lane 2 contains *nozA* (711 bp) digested from pQE30 plasmid vector (3.4-Kb) and lane 3 contains *ncdA* (753 bp) digested from pQE30 plasmid vector (3.4-Kb). (b) Lane 1 contains GeneRuler 1 kb Plus Ladder and lane 2 contains *ncdA* (753 bp) digested from pET30a(+) (5.4-Kb).

E*. coli* pQE30/*nozA* and pQE30/*ncdA* were induced with IPTG to upregulate expression of the genes encoding NozA and NcdA. The optical density (OD_{600}) measured before (~0.4) and after induction (~1.0) confirmed uniform growth among all treatment (carrying *nozA* or *ncdA*) and control cultures (carrying vector without biosynthetic genes). SDS-PAGE analysis of induced control and both uninduced and induced treatment soluble lysates suggested successful overexpression of both CDPSs as soluble proteins upon induction with IPTG (Figure 8). With this confirmation of the expression of both NozA and NcdA, the catalytic function of these enzymes was next probed.

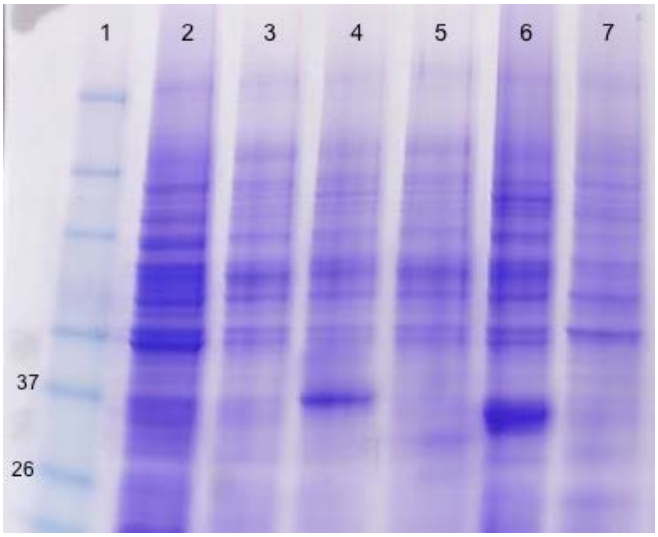


Figure 8. SDS-PAGE verifying the overexpression of NozA and NcdA in the soluble lysates of recombinant *E. coli* upon induction with IPTG. Lane 1 contains BenchMark Pre-Stained Protein Ladder (Life Technologies) and lane 2 contains soluble proteins from the induced *E. coli* M15[pREP4] pQE30 vector control, containing no biosynthetic genes. Lane 3 contains soluble proteins from the uninduced *E. coli* M15[pREP4] pQE30 vector control, containing no biosynthetic genes. Lanes 4-5 contain soluble proteins from induced and uninduced M15[pREP4] pQE30/*ncdA*, respectively. Lanes 6-7 contain soluble proteins from induced and uninduced M15[pREP4] pQE30/*nozA*, respectively. Overexpression of proteins in observed in treatment lanes lanes 4 and 6 at the correct NcdA and NozA respective molecular weights of 29- and 28-kDa, respectively.

Characterization of the functions of recombinantly expressed NozA and NcdA in vivo

Metabolite profiles of supernatants from induced treatment and control cultures were compared using HPLC to determine the presence of *cyclo*(L-Trp-L-Trp) upon expression of NozA and NcdA *in vivo*. When compared to the empty vector control, recombinant *E. coli* expressing NozA or NcdA produced a unique HPLC UV signal (~23.7 min), matching the retention time of the signal created by a synthetic *cyclo*(L-Trp-L-Trp) standard (Figure 9). High resolution LC/MS analyses of the metabolites revealed almost identical molecular weights between

the metabolites and synthetic standard, further supporting that both NozA and NcdA catalyze the production of *cyclo*(L-Trp-L-Trp), the putative precursor in nocardioazine biosythesis. The observed *in vivo* $[M+Na]^+$ *m/z* for synthetic *cyclo*(L-Trp-L-Trp) was 395.1472, for NozA *cyclo*(L-Trp-L-Trp) was 395.1470, and for NcdA *cyclo*(L-Trp-L-Trp) was 395.1474.

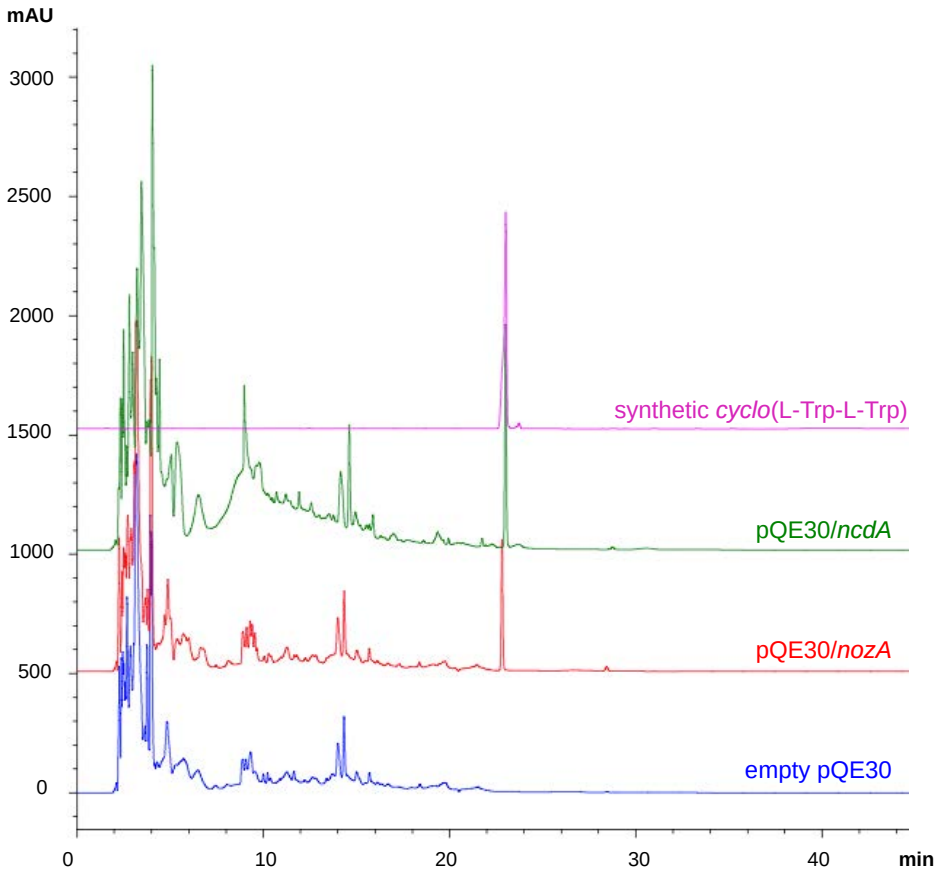


Figure 9. HPLC metabolite profiles for supernatants from IPTG-induced *E. coli* cultures pQE30/*nozA* or pQE30/*ncdA* relative to control cultures featuring empty pQE30 vector. A distinct signal at 23.7 minutes was observed in supernatants from treatment *E. coli* but was absent from control *E. coli* carrying empty pQE30 vector. The retention time of the metabolite matched that of synthetic *cyclo*(L-Trp-L-Trp) standard, supporting that both NozA and NcdA are able to catalyze *cyclo*(L-Trp-L-Trp) assembly. Chromatograms show absorbance at 280 nm.

Probing the roles of conserved amino acid residues in catalysis by NozA

NozA, NcdA, and Amir_4627 are unique as the only CDPSs known to catalyze formation of *cyclo*(L-Trp-L-Trp). Several residues conserved between the protein sequences of NozA, NcdA and Amir_4627 were identified and mutated in NozA using PCR-targeted site directed mutagenesis in order to determine whether or not the conserved residues played a role in the enzyme's preference for tryptophan. Furthermore, a residue identified in AlbC proven to play a critical role in the substrate preference of the enzyme was mutated in NozA via PCR-targeted site directed mutagenesis to determine if changing that residue would alter the substrate preference of NozA. The serine-36 residue conserved among the functionally characterized CDPSs was also mutated in NozA to determine whether that residue played a role in the active site of the enzyme.

The optical density (OD_{600}) measured before (~0.4) and after induction with IPTG (~1.0) confirmed uniform growth among control (*E. coli* M15[pREP4] carrying NozA with no site directed mutations) and treatment cultures (*E. coli* M15[pREP4] carrying NozA with mutated residues). SDS-PAGE analysis of treatment and control culture lysates verified uniform protein expression among all cultures suggesting that *nozA* site directed mutations did not interfere with *E. coli* growth (Figure 10).

Metabolite profiles of NozA site directed mutagens alongside controls analyzed by HPLC revealed production of *cyclo*(L-Trp-L-Trp) was abolished upon

mutation of the serine residue to an alanine in NozA's active site (S36A). These data imply that serine-36 plays an essential role in NozA catalysis and suggest that NozA employs a catalytic mechanism analogous to that recently unveiled for AlbC, (Moutiez et al. 2014) a CDPS 33% homologous to NozA. During the first step of AlbC catalysis, the aminoacyl moiety from aa-tRNA is transferred onto the serine active site. This aminoacyl group forms a peptide bond with an aminoacyl moiety donated from a second aa-tRNA substrate, and the resulting dipeptidyl intermediate then undergoes intramolecular peptide bond formation to close the DKP ring and concomitantly release from the active site.

The other mutations did not alter the enzyme's ability to catalyze the production of *cyclo*(L-Trp-L-Trp), suggesting those conserved residues are not critical in the enzyme's substrate specificity (Figure 11). Additionally, production of alternative cyclodipeptide metabolites due to the mutations made to NozA were not observed.

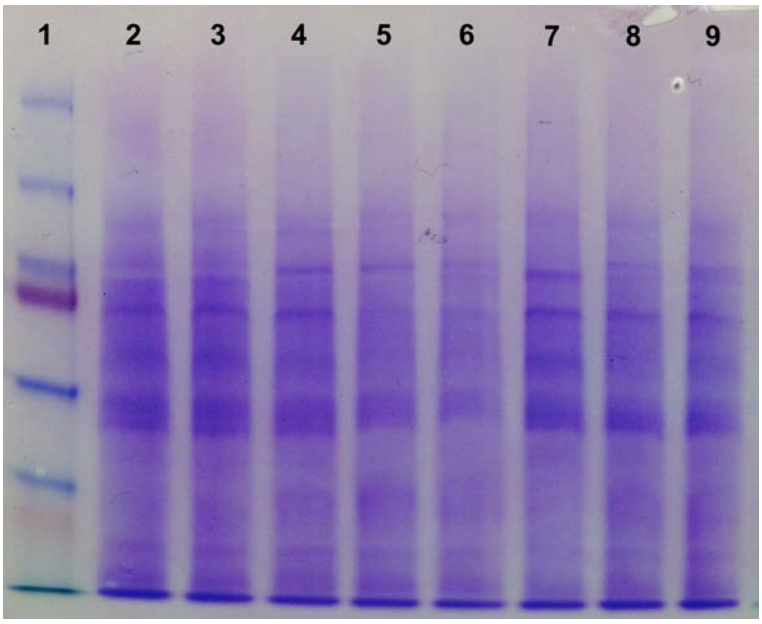


Figure 10. SDS-PAGE verifying uniform protein expression among *E. coli* M15[pREP4] pQE30 carrying *nozA* with and without site directed mutations. Lane 1 contains PageRuler Plus Prestained Protein Ladder, lane 2 contains pQE30 carrying no biosynthetic genes, lane 3 contains pQE30/*nozA* (no mutations), lane 4 contains pQE30/*nozA*-S36A, lane 5 contains pQE30/*nozA*-L131R, lane 6 contains pQE30/*nozA*-S198L, lane 7 contains pQE30/*nozA*-N114R, lane 8 contains pQE30/*nozA*-N114A and lane 9 contains pQE30/*nozA*-V111R+N114A. The optical density (OD_{600}) following fermentation with IPTG was ~1.0 for each culture, suggesting that the *nozA* site directed mutations did not interfere with *E. coli* growth.

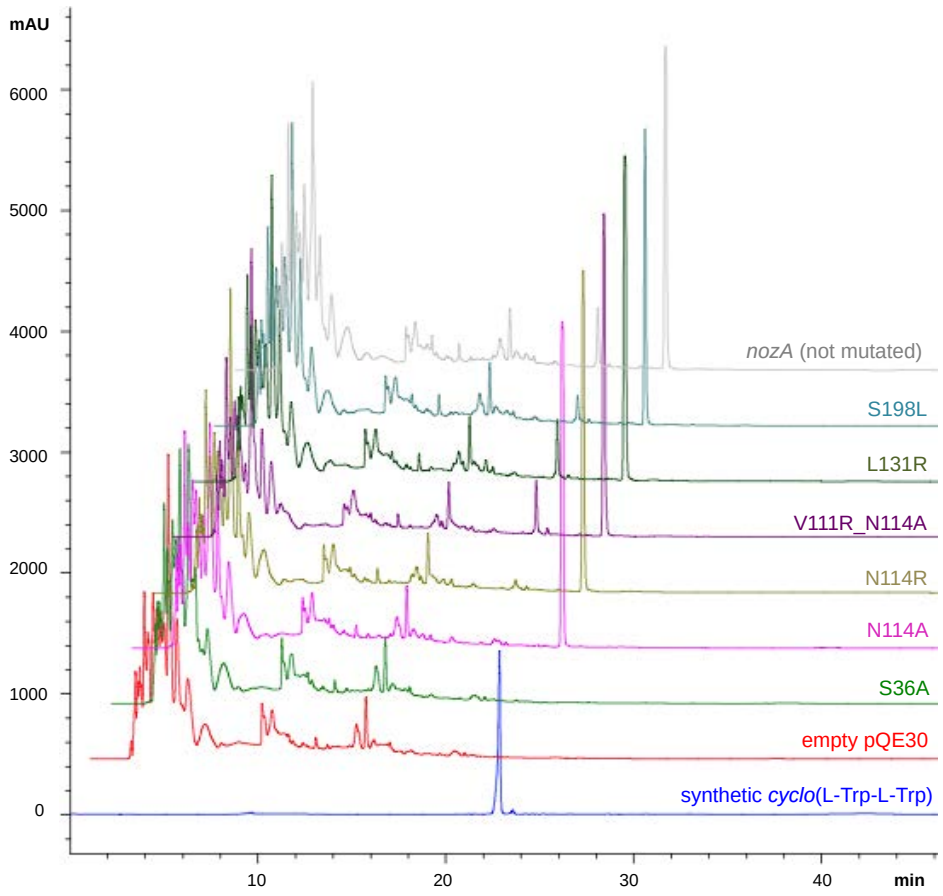


Figure 11. HPLC metabolite profiles for *E. coli* M15[pREP4] producing NozA site directed mutagens. A distinct signal at 23.7 minutes, indicating *cyclo*(L-Trp-L-Trp) production, was observed in the supernatants from all treatments and controls except the empty vector control carrying no biosynthetic genes and the S36A mutant, where the serine active site residue was replaced with alanine. This confirmed the critical role of the serine-36 residue for the catalytic function of NozA. Chromatograms show absorbance at 280 nm.

Characterization of the function and substrate preferences of purified recombinant NozA and NcdA in vitro

By exploiting the hexahistidine tags incorporated into NozA and NcdA by the expression plasmids, the purification of both enzymes began with standard Ni-affinity chromatography (Figure 12, a-b). Subsequent purification by anion exchange FPLC resulted in >90% pure NozA and NcdA for *in vitro* analysis of enzyme functions and substrate preferences (Figure 12, c-d).

Previously characterized CDPSs were found to use aminoacyl-tRNAs as substrates for the catalysis of 2,5-DKP biosynthesis (Moutiez et al. 2014). LC/MS analysis following incubation of NozA and NcdA with tryptophanyl-tRNA *in vitro,* revealed production of *cyclo*(L-Trp-L-Trp). Eliminating any one component of the assay mixture (e.g. enzyme (NozA or NcdA), tryptophan, tRNA or aminoacyl-tRNA synthetase mix) resulted in no production of *cyclo*(L-Trp-L-Trp) (Figure 13 a). These results confirmed that NozA and NcdA, like other characterized CDPSs, use aminoacyl-tRNA's as their substrates to catalyze the production of *cyclo*(L-Trp-L-Trp). Both NozA and NcdA are most closely related to CDPSs that incorporate aromatic aminoacyl-tRNAs (Figure 6). In order to elucidate the substrate specificity of NozA and NcdA, both enzymes were incubated with tRNAs charged with alternative aromatic amino acids, L-tyrosine and L-phenylalanine. LC/MS analysis revealed that NozA and NcdA were not able to catalyze the biosynthesis of alternative 2,5-DKPs using these aromatic aminoacyl-tRNA substrates, revealing the substrate specificity of both enzymes for L-tryptophan charged tRNA (Figure 13 a).

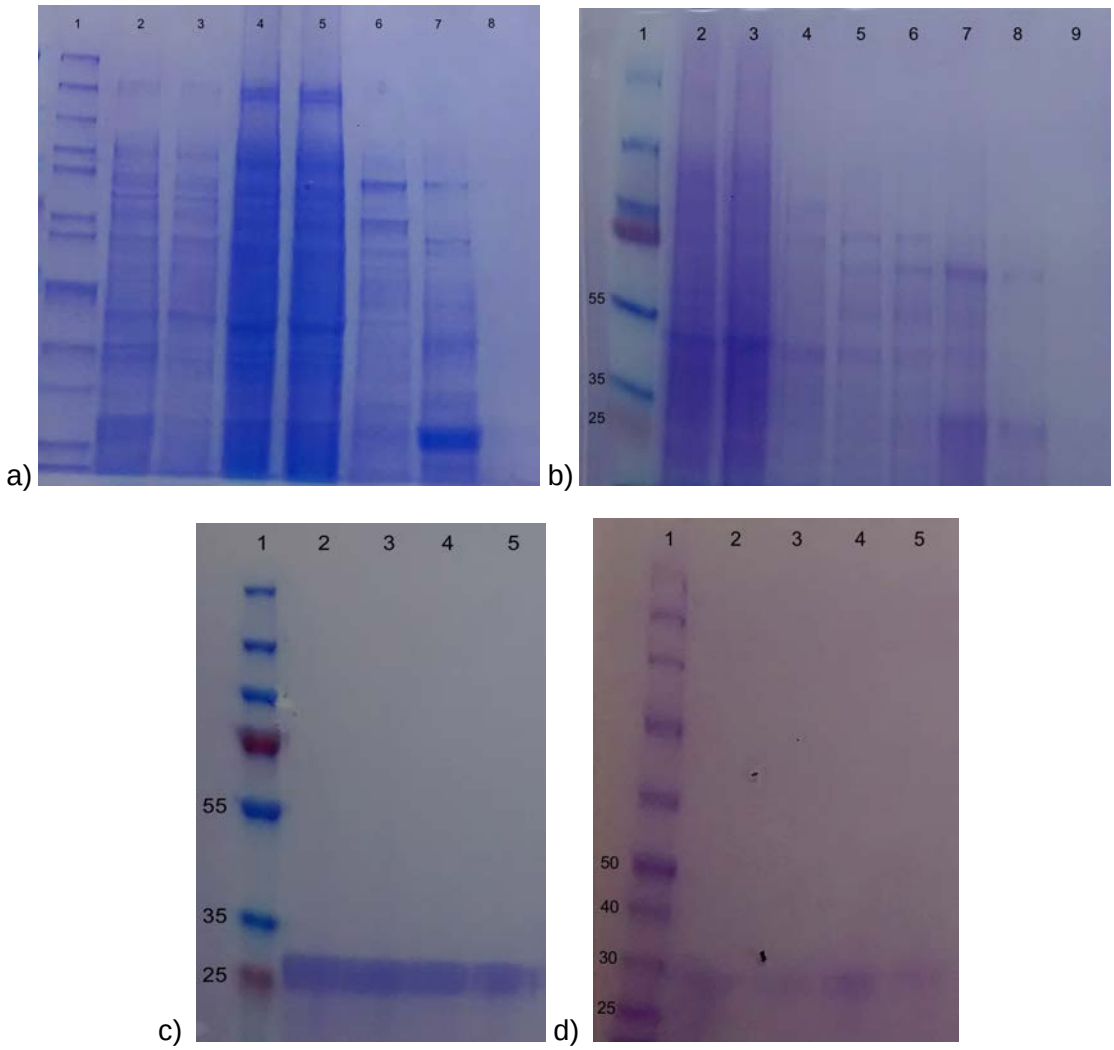


Figure 12. SDS-PAGE analysis of NozA and NcdA purification via Ni-column and FPLC. (a) NozA Ni-column eluents. Lane 1 contains PageRuler Unstained Protein Ladder (Thermo Scientific), lane 2 contains total proteins, lane 3 contains total soluble proteins, lane 4 contains the flow through fraction (containing 5 mM imidazole), lane 5 contains 10 mM imidazole eluent, lane 6 contains 50 mM imidazole eluent, lane 7 contains 200 mM imidazole eluent and lane 8 contains 500 mM imidazole eluent. NozA (~28 kDa) was eluted in the 200 mM imidazole fraction. (b) NcdA Ni-column eluents. Lane 1 contains PageRuler Prestained Protein Ladder (Thermo Scientific), lane 2 contains the flow through fraction (containing 5 mM imidazole), lane 3 contains 10 mM imidazole eluent, lane 4 contains 50 mM imidazole eluent, lanes 5-6 contain sequential 100 mM imidazole eluent, lanes 7-8 contain sequential 200 mM imidazole eluent and lane 9 contains 500 mM imidazole eluent. NcdA (~29 kDa) was eluted in both 200 mM imidazole fractions. (c-d) FPLC fractions containing >90% pure NozA and NcdA. (c) Lane 1 contains PageRuler Prestained Protein Ladder (Thermo Scientific) and lanes 2-5 contain purified NozA (~28 kDa). (d) Lane 1 contains PageRuler Unstained Protein Ladder (Thermo Scientific) and lanes 2-5 contain purified NcdA (~29 kDa). These >90% pure fractions were pooled for *in vitro* analyses of enzyme functions.

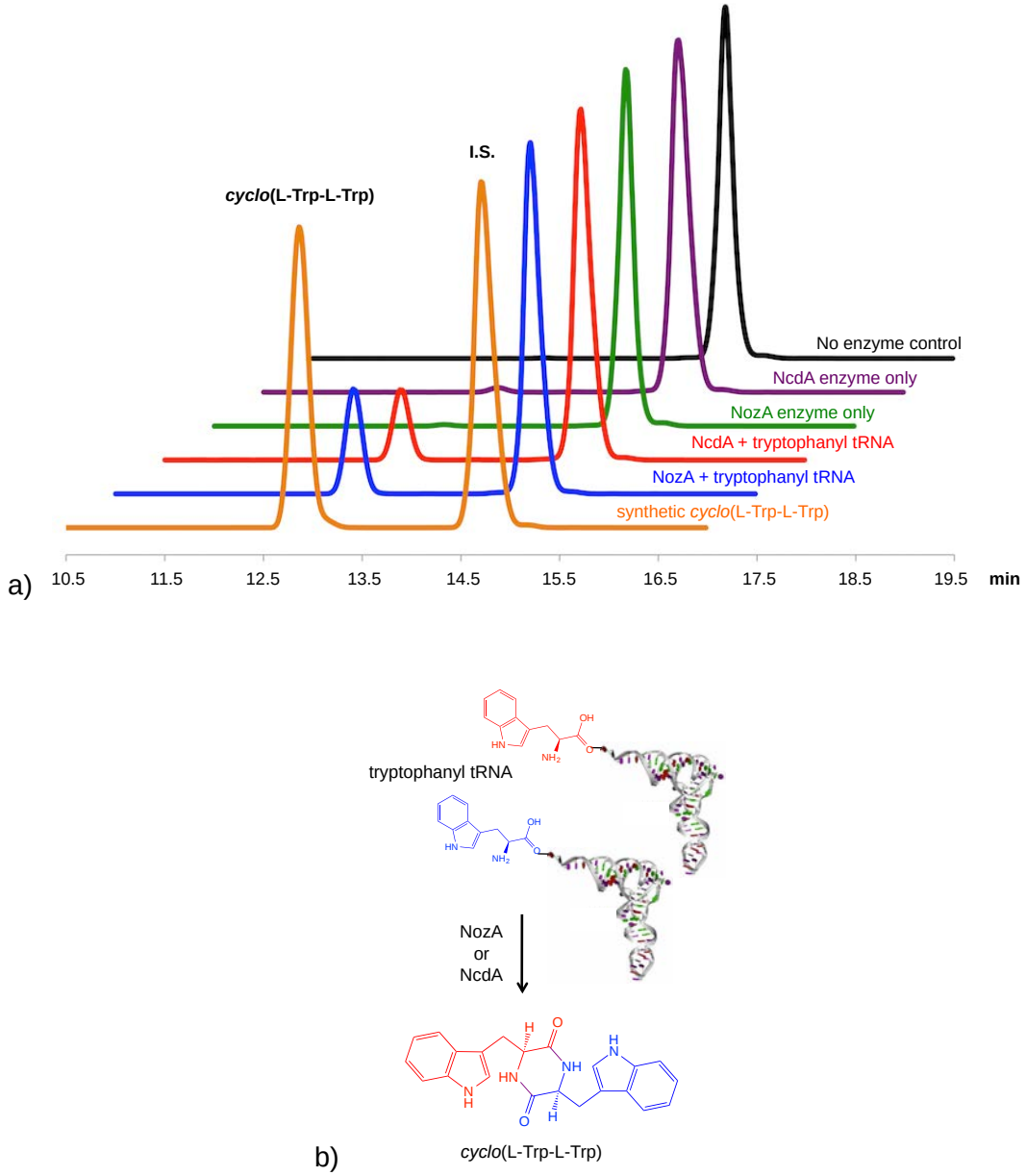


Figure 13. *In vitro* characterization of NozA and NcdA encoded by the *Nocardiopsis* sp. CMB-M0232 genome. (a) LC/MS revealed that incubation of purified recombinant NozA and NcdA with tryptophanyl-tRNA substrates yielded *cyclo*(L-Trp-L-Trp) *in vitro.* Incubation with alternative aromatic amino acids yielded no alternative 2,5-DKPs (data not shown), revealing substrate specificity of both enzymes for tryptophan-charged tRNA. Plot shows LC/MS selected ion recording for *cyclo*(L-Trp-L-Trp) and for *cyclo*-3-methyl-L-Trp-N1'-methyl-L-Trp DKP internal standard (I.S.). (b) These data support the overall reaction catalyzed by both NozA and NcdA is to employ tryptophanyl-tRNA substrates to yield *cyclo*(L-Trp-L-Trp) product.

These experiments confirmed that NozA and NcdA from *Nocardiopsis* sp. CMB-M0232 catalyze the production of *cyclo*(L-Trp-L-Trp) both *in vivo* and *in vitro*. These results suggest both enzymes are involved in the biosynthesis of nocardioazine A by producing the 2,5-DKP precursor in the biosynthetic pathway, *cyclo*(L-Trp-L-Trp). These findings make NozA and NcdA two of only three characterized CDPSs from prokaryotic origin to use tryptophanyl-tRNA as a substrate and two of only three characterized CDPSs from any origin to exhibit specificity for a single substrate. Fidelity is uncommon amongst characterized CDPSs, making NozA and NcdA important CDPS family additions.

Characteristic to CDPS-containing biosynthetic pathways are DKP-tailoring enzymes capable of modifying the CDPS-assembled substrates to produce alternative natural products (Gondry et al. 2009, Sauguet et al. 2011, Seguin et al 2011, Belin et al. 2012, Giessen and Marahiel, 2012, Giessen et al. 2013). Chapter three aims to identify putative *cyclo*(L-Trp-L-Trp) modifying enzymes in the noz gene cluster to determine their roles in the modification of *cyclo*(L-Trp-L-Trp) to produce nocardioazines A-B.

Chapter III

Evaluation of tailoring enzymes with hypothesized roles in nocardioazine A biosynthesis

Introduction

In prokaryotic organisms, genes involved in secondary metabolic pathways are typically organized in operon-like structures within the genome, spanning up to 100 kilobases or more (Lane and Moore, 2012). This organization in prokaryotes greatly facilitates location of entire biosynthetic pathways compared to eukaryotic organisms, where genes involved in a biosynthetic pathway are often dispersed among the entire genome. Once a prokaryotic biosynthetic gene cluster is located, modern molecular techniques allow probing of individual players in the pathway to elucidate their enzymatic functions and engineer synthesis of alternative metabolites.

In the case of peptide derived natural products, two biological strategies exist for generating structural diversity: (1) changing the peptide backbone itself or (2) modifying the scaffold with various tailoring enzymes (Giessen and Marahiel, 2012). The majority of functionally characterized CDPSs exhibit promiscuity in their aa-tRNA substrate specificity and are able to incorporate multiple amino acids to generate several different 2,5-DKP backbones. NozA and NcdA are two of only three CDPSs that exhibit substrate specificity, solely catalyzing the production of *cyclo*(L-Trp-L-Trp) assembly. We focused on gaining insight into the function of cytochrome P450 enzymes encoded within the same operon as *nozA.* These enzymes were hypothesized to play roles in the late

stages of biosynthesis of the nocardioazines (Figure 14), which share the *cyclo*(Trp-Trp) backbone provided by the *Nocardiopsis* sp. CDPSs (Chapter 2).

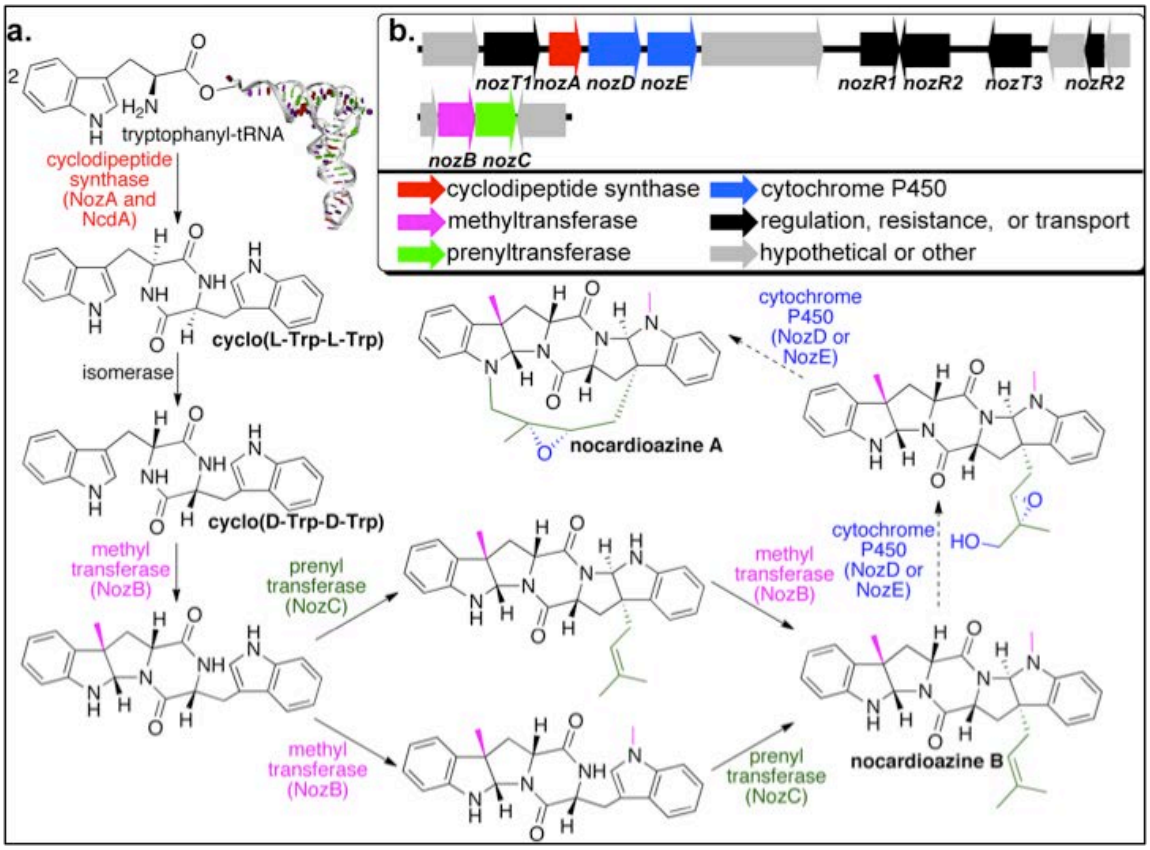


Figure 14. Nocardioazine biosynthesis overview. (a) Schematic of proposed intermediates in nocardioazine assembly, as established by comparison of synthetic standards to metabolites from *Nocardiopsis* sp. (Alqahtani et al. 2015). In this thesis, gene elimination mutations and feeding studies were employed in an effort to evaluate the functions of NozD and NozE in the late stages of nocardioazine assembly. (b) Overview of genes with predicted roles in nocardioazine assembly.

All CDPS-encoding biosynthetic pathways contain at least a single gene encoding a DKP-modifying enzyme positioned upstream or downstream from the CDPS gene, indicating that the DKPs generated are ultimately modified further to

yield diverse chemical structures (Belin et al. 2012, Giessen and Marahiel, 2014). This organization generates a mechanism for biosynthesis of standard cyclic peptide scaffolds that can be modified by a variety of enzymes, incorporating functional groups that result in unique biologically activities. Exploiting this molecular method provides endless opportunities to synthesize new natural products and therefore, new candidates for pharmaceutical drugs.

There are several enzymes characteristic to CDPS-containing pathways that can be organized into three groups: modifying enzymes, regulators and transporters (Giessen and Marahiel, 2014). Common classes of modifying enzymes are hydrolases, methyltransferases, and acyl-CoA transferases. Among all, the most common class of tailoring enzymes are the oxidases, with a relatively high majority homologous to cytochrome P450s, enzymes that catalyze a wide variety of oxidation reactions, including oxygenation reactions on a variety of substrates by adding hydroxy and epoxide groups (Giessen and Marahiel, 2014). Common regulatory enzymes associated with CDPS-containing gene clusters are transcription factors. Transcription factors trigger the expression of genes in the biosynthetic pathways in response to environmental cues such as oxidative stress, toxins and antibiotics. Membrane transporters are also common among CDPS-containing gene clusters, generally responsible for exporting the products of biosynthesis from the cell. Based on these three enzyme groups, it is suggested that environmental cues trigger synthesis of CDPs, which are then transported to the periplasm or extracellularly (Belin et al. 2012 & Giessen and Marahiel, 2014).

NozD and NozE, two cytochrome P450 homologs located directly downstream from *nozA* (Figure 14 & 16; Table 1), were predicted via bioinformatics analyses to modify the unique C3'-prenyl group of nocardioazine B to form the epoxide ring characteristic to nocardioazine A and is critical to its anticancer activity. In this chapter we aim to elucidate the roles of NozD and NozE in nocardioazine A biosynthesis by expressing the genes in heterologous *Streptomyces* hosts optimized for the expression of biosynthetic pathways from actinomycetes. Constructs engineered to express the gene cluster containing *nozA, nozD* and *nozE* (Gust et al. 2004, Smanski et al. 2012) as well as three additional constructs featuring gene elimination mutations in the gene cluster were engineered for expression in *Streptomyces*. To probe the functions of the three genes from contig 557 (*nozA, nozD* and *nozE*) the heterologous *Streptomyces* were exposed to nocardioazine B via *in vivo* and *in vitro* methods and metabolite profiles were screened via HPLC to detect accumulation of nocardiazine A.

Methods

Isolation of nocardioazines A-B from wild type Nocardiopsis sp. CMB-M0232

M1 liquid media (20 mL) (10 g/L potato starch, 4 g/L yeast, 2 g/L peptone, 1 g/L $CaCO_3$, 40 mg/L $Fe_2(SO_4)_5H_2O$, 100 mg/L KBr, pH 7) was inoculated with wild type *Nocardiopsis* sp. CMB-M0232 from a -80 °C freezer stock and

incubated for 6 days and 200 rpm at 30 °C. Following incubation, 5 x 1 L of M1 liquid media in Fernbach flasks were inoculated with 1 mL each of the 6-day starter culture and incubated at 30 °C and 200 rpm for seven days. Following incubation, the 5 L culture was filtered over Celite 545 (Fisher) and extracted with an equal volume of ethyl acetate. The ethyl acetate was removed from the aqueous extracts using a vacuum controlled rotavapor (BUCHI) and the remaining traces of liquid were dried using a SpeedVac concentrator (Savant ISS1100, Thermo Scientific) overnight.

Nocardiopsis sp. crude extracts (~600 mg from 5 L culture) were subjected to sequential trituration with water, *n*-hexane, dichloromethane (CH_2Cl_2), and methanol (MeOH) to begin to separate the nocardioazines from other compounds. Specifically, the crude extracts were solubilized in 5 mL of Milli-Q water and centrifuged (5,000 rpm for 10 minutes). The supernatant, consisting of water with soluble polar compounds, was discarded. The remaining pellet composed of insoluble, nonpolar compounds was dried using a SpeedVac concentrator (Savant ISS1100) and resolublized in *n*-hexane (50 mL). The suspension was centrifuged and the supernatant was transferred to a scintillation vial, dried using a vacuum controlled rotavapor (BUCHI), and set aside for LC/MS analysis. The remaining pellet from the hexane suspension was then resolublized in CH_2Cl_2 (50 mL) and handled identically to the hexane suspension. This was again repeated using MeOH (50 mL). Dried *n*-hexane, CH_2Cl_2 and MeOH extracts were resolubilized in MeOH to a final concentration of 10 mg/mL

and subjected to LC/MS analysis to determine which fraction contained the majority of nocardioazine A-B.

LC/MS determined that the *n*-hexane fraction contained the highest concentrations of nocardioazines A-B. An aliquot (500 µL) of the *n*-hexane extract (10 mg/mL) was subjected to HPLC using an Aligent 1100 HPLC system with a ZORBAX SB-C18 column (9.4 x 250 mm, particle size: 5 µm, pore size of 110 Å). A flow rate of 3.5 mL/min was used, with solvent A consisting of MilliQ water, solvent C consisting of acetonitrile and a linear gradient of 10 to 100% solvent C over 25 minutes following an 5 minute hold at 10% solvent C. Fractions were collected coming off the column in two-minute increments, concentrated and analyzed by LC/MS to identify which peaks in the spectra (280 nm) correspond to nocardioazine A-B. The respective peaks were collected from the remainder of the hexane extracts, dried using a SpeedVac concentrator (Savant ISS1100) and solubilized in 100% chloroform-d (Alfa Aesar) to verify their chemical structures by nuclear magnetic resonance (NMR).

Generation of heterologous constructs for expression of contig 557 in Streptomyces coelicolor M1152 and Streptomyces lividans TK24

The SuperCos 1 cosmid clone carrying contig 557, featuring the noz gene cluster (Figure 14 & 16, Table 1), was modified for introduction and expression in *Streptomyces* hosts following protocols described by Smanski et al. (2012). Specifically, SuperCos-1 (Agilent) carrying contig 557 from a cosmid clone library

generated as described by Bis et al. (2015) was modified by λ-RED *E. coli* recombination to introduce *oriT*, an origin of conjugal transfer, *Streptomyces θ*C31 integrase to facilitate integration into the *attB* site of *Streptomyces'* chromosome, and *aac(3)IV,* an apramycin resistance gene for selection.

A 270 bp fragment of the 3' end of the *bla* gene from SuperCos 1 was amplified using methods and primers described by Smanski et al. and cloned into the XbaI/BamHI sites in the *Streptomyces*-integrating vector, pSET152 (Bierman et al. 1992) which encodes *aac(3)IV, oriT* and *θ*C31 integrase. The pSET152/3'*bla* construct was then linearized at BamHI and EcoRI restriction sites and introduced into λ-RED recombination proficient *E. coli* BW25113/pIJ790 carrying SuperCos/contig 557 to undergo homologous recombination between the homologous pUC and 3' -*bla* sites of the SuperCos backbone. PCR and DNA sequencing were used to confirm the presence of *aac(3)IV, oriT* and *θ*C31 integrase in SuperCos/contig 557A, the product of this recombination event (Appendix C, SuperpSET_F/R). *E. coli* ET12567/pUZ8002 was transformed with SuperCos/contig 557A to facilitate conjugation into *Streptomyces coelicolor* M1152 and *Streptomyces lividans* TK24 using standard methods described by Gust et al. (2004). PCR confirmed successful integration of contig 557 genes into the *S. coelicolor* M1152 and *S. lividans* TK24 chromosome (Appendix C, NozACtg557_F/R).

In vivo and in vitro characterization of NozD and NozE heterologously expressed in Streptomyces

NozD and NozE, cytochrome P450 homologs (Table 1) were predicted via bioinformatics analyses as the most likely candidates for the modifications made to nocardioazine B to produce nocardioazine A, which requires cyclizing of the C3'-prenyl group on nocardioazine B to form the epoxide ring on nocardioazine A (Figure 14). To determine the role of NozD and NozE in this catalytic reaction, *in vivo* and *in vitro* experiments were designed using heterologous *Streptomyces* carrying the entire contig 557 including NozA, NozD and NozE.

For the *in vivo* approach, *S. coelicolor* M1152 and *S. lividans* TK24 heterologously expressing the contig 557 gene cluster (including NozA, NozD and NozE) were cultured in media supplemented with nocardioazine B, extracted and analyzed for the presence of nocardioazine A. Specifically, 100 µL of 3-day starter cultures of heterologous *S. coelicolor* M1152 and *S. lividans* TK24 carry contig 557 were inoculated into 25 mL of M1 media supplemented with 50 µg/mL apramycin in 250 mL Erlenmeyer flasks. Wild type *S. coelicolor* M1152 and *S. lividans* TK24 carrying no biosynthetic genes were cultured alongside as controls, lacking apramycin. Two 25 mL cultures of each treatment and control strain were prepared and supplemented as follows prior to fermentation. One culture of each wild type strain and heterologous strain was supplemented with 250 µg nocardioazine B that was solubilized in dimethyl sulfoxide (DMSO). The second set of cultures was supplemented with equal volumes of DMSO to serve

as controls. The treatment and control cultures were incubated at 30 °C and 200 rpm for 6 days.

Following incubation, the cultures were extracted with 25 mL of EtOAc and the chemical extracts were concentrated using a vacuum controlled rotavapor (BUCHI) followed by a SpeedVac concentrator (Savant ISS1100, Thermo Scientific) overnight. The extracts were resolubilized in 250 uL MeOH for analysis by HPLC to determine whether NozD and NozE were able to convert nocardioazine B to nocardioazine A. The HPLC method mentioned previously was run with known concentrations of *cyclo*(L-Trp-L-Trp) (0.2 mg/mL), nocardioazine B (25 µM) and nocardioazine A (25 µM) to serve as positive controls.

The *in vitro* approach offered an alternative that exposed nocardioazine B to cell lysates containing NozD and NozE rather than requiring nocardioazine B enter the cell as in the *in vivo* methods. Specifically, 200 µL of 3-day starter cultures of heterologous *S. lividans* TK24 carrying contig 557 (NozA, NozD and NozE) were inoculated into 50 mL of M1 media supplemented with 50 µg/mL apramycin in 250 mL Erlenmeyer flasks. Wild type *S. lividans* TK24 was cultured alongside as a control, lacking apramycin. Following 3 days and 6 days of incubation at 30 °C and 200 rpm, 25 mL of each culture (wild type control and treatment carrying contig 557) were transferred to 50 mL falcon tubes. The cells were harvested by centrifugation (6,000 rpm, 4°C) and were resuspended in 3 mL lysis buffer (100 mM Tris HCl, 150 mM NaCl, 5% glycerol, 2 mM DTT, pH 8) with lysozyme (3.75 mg). The cell suspensions were sonicated at 60% amplitude

for 3 minutes (1 s pulse:10 s rest). Cell debris was removed by centrifugation (12,000 rpm, 4 °C) and the soluble lysates were filtered through a 0.45 μM cellulose acetate membrane (Corning). In 1.5 mL microtubes, 100 μL of cell lysates were combined with 20 μL of DMSO-solubilized nocardioazine (100 μg). To serve as controls, duplicate assays were set up that combined 100 μL of cell lysates with 20 μL of DMSO (no addition of nocardioazine B). The assays were incubated ~16 hours at 30 °C. Following incubation, the assays were extracted with 1 mL EtOAc and the chemical extracts were dried using a vacuum controlled rotavapor (BUCHI) followed by a SpeedVac concentrator (Savant ISS1100, Thermo Scientific) overnight. The extracts were resolubilized in 100 uL MeOH for analysis by HPLC, described previously, to determine whether NozD and NozE converted nocardioazine B to nocardioazine A.

Generation of PCR-targeted gene replacements in the contig 557 gene cluster for heterologous expression in Streptomyces coelicolor M1152 and Streptomyces lividans TK24

Additional constructs were created with PCR-targeted gene deletions to further probe the enzymatic functions of NozA and the two cytochrome P450s encoded downstream from *nozA*, referred to a NozD and NozE. Three constructs were designed, each with one of the three abovementioned genes deleted in a manner that would not interfere with expression of the other genes following

procedures described by Smanski et al. (2012) and Gust et al. (2004) (Figure 15).

Specifically, gene disruption was carried out via PCR to replace each targeted biosynthetic gene with an apramycin resistance cassette flanked with FLP recognition targets (FRTs) (Figure 15 a). The FRTs were engineered into the cassettes for FLP recombinase-induced excision of the apramycin resistance gene to yield an in-frame scar in place of the gene. First, PCR was conducted to amplify the FLP-flanked apramycin resistance gene by using primers with regions homologous to the region immediately up- and downstream from the biosynthetic gene targeted for replacement (Figure 15 a). The PCRs contained 36 µL of molecular biology grade water, 0.5 µL of each forward and reverse primer (NozX::scarKO_F/R, Appendix C) (0.5 µM final concentration), 5 µL of 10X Expand High Fidelity Buffer with $MgCl_2$ (Roche), 0.5 µL of High Fidelity dNTPs (10 mM final concentration of each dNTP), 2.5 µL of DMSO, 0.5 µL (~50 ng) of template DNA (Apr^R cassette from pIJ773 vector) and 1 µL of Expand High Fidelity Enzyme Mix (Roche). The PCR reactions ran for an initial denaturation of 94 °C for 2 min, 10 cycles of 94 °C for 45 s, 50 °C for 45 s, 72 °C for 90 s, 15 cycles of 94 °C for 45 s, 55 °C for 45 s, 72 °C for 90 s and a final extension of 72 °C for 5 min. The PCR products were verified for size by agarose gel electrophoresis and the remaining PCR product was purified using a QIAquick PCR Purification Kit.

Following amplification of the apramycin resistance cassette, SuperCos/contig 557 was introduced into *E. coli* BW25113/pIJ790 via

electroporation (Figure 15 b). Next, the PCR-amplified apramycin resistance cassette was introduced into *E. coli* BW25113/pIJ790/SuperCos/contig 557 (Figure 15 b). In three separate constructs, one of each of the desired genes (*nozA, nozD* or *nozE)* was replaced with the FRT-flanked apramycin resistance cassette, *aac(3)IV*, via λ-RED recombination to generate SuperCos/*nozA*::Apr^R, SuperCos/*nozD*::Apr^R and SuperCos/*nozE*::Apr^R constructs (Figure 15 c). PCR confirmed recombination between the SuperCos/contig 557 cosmids and the Apr^R cassette.

Prior to conjugation with *Streptomyces,* the SuperCos/contig 557 constructs carrying apramycin disruption cassettes in place of *nozA, nozD,* or *nozE* required further modification. This modification required exploiting FLP recombinase to replace the disruption cassette with an 81 bp scar to create non-polar, in-frame deletions (*nozA*::scar, *nozD*::scar, *nozE*::scar). SuperCos carrying the apramycin disruption cassettes were introduced into *E. coli* DH5α/BT340 (*E. coli* Genetic Stock Center: Strain# 7629) via standard electroporation (Figure 15 d). Incubation at 42 °C induced expression of the FLP-recombination plasmid, BT340, to express FLP-recombinase, which recognizes the FRT regions, resulting in excision of the apramycin resistance cassette. This leaves behind an 81 bp “scar” in the correct reading frame (Figure 15 e). PCR confirmed the presence of the 81 bp “scars” in each construct (Appendix C, NozX_KOCon_F/R).

Following protocols mentioned in the previous section, these three SuperCos cosmids now carrying the contig 557 with an 81 bp “scar” in place of

NozA, NozD or NozE were again introduced into λ-RED recombination proficient *E. coli* BW25113/pIJ790 carrying the previously constructed pSET152/3'*bla* linearized plasmid DNA (Smanski et al. 2012). Homologous recombination between the pUC origin of replication and 3' -*bla* sites introduced *oriT*, *Streptomyces θ*C31 integrase, and *aac(3)IV* into SuperCos to facilitate conjugation (Figure 15 f). PCR and DNA sequencing were used to confirm the presence of *aac(3)IV, oriT* and *θ*C31 integrase in the products of the recombination event (Appendix C, SuperpSET_F/R).

Each of the SuperCos/*noz*::scar constructs (*nozA*::scar, *nozD*::scar, *nozE*::scar) were introduced into *E. coli* ET12567/pUZ8002 to facilitate conjugation into *S. coelicolor* M1152 and *S. lividans* TK24 using standard methods described by Gust et al. (2004) (Figure 16 g-h). PCR amplification confirmed successful integration of the *nozA*::scar, *nozD*::scar and *nozE*::scar constructs into *Streptomyces* chromosome (Appendix C, NozX_KOCon_F/R).

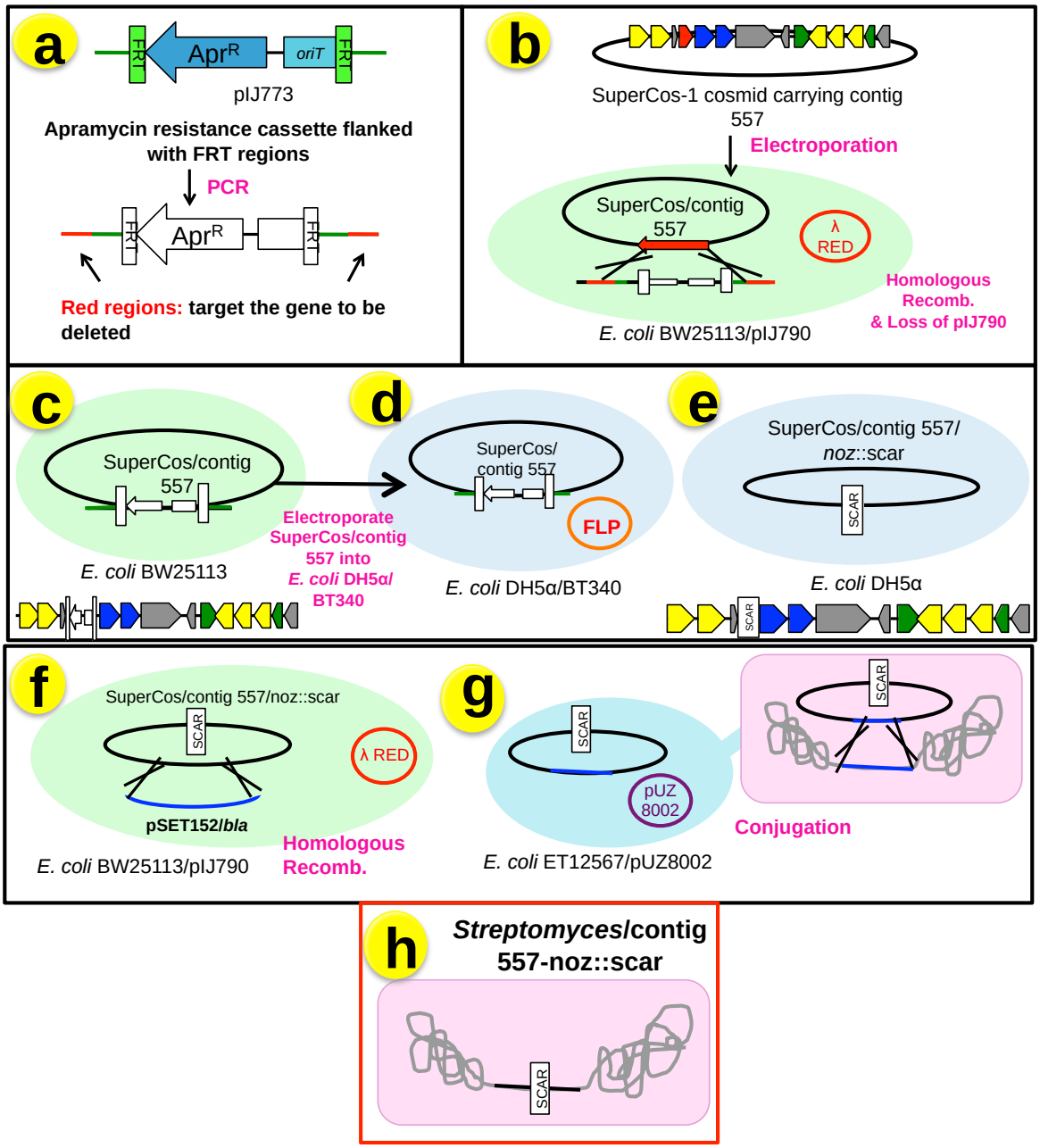


Figure 15. Methods for generating PCR-targeted gene deletions in the noz gene cluster and introducing the constructs into heterologous *Streptomyces* via conjugation. (a) An apramycin resistance cassette from pIJ773 was PCR amplified using primers that introduced homologous regions of the gene to be deleted both upstream and downstream from the apramycin resistance cassette. (b) SuperCos1-contig 557 from the cosmid clone library and the PCR amplified apramycin resistance cassette were both introduced into *E. coli* BW25113/pIJ790 to undergo homologous recombination. (c) Homologous recombination replaced the targeted gene in contig 557 with the apramycin resistance cassette. (d) SuperCos/contig 557/noz::Apr^R was introduced into *E. coli* DH5α/BT340. (e) FLP-recombinase was expressed by the BT340 plasmid, which recognized FRT sequences flanking the apramycin resistance cassette and excised the cassette, leaving behind an 81-bp, in frame 'scar'. (f) SuperCos/contig 557/noz::scar was

introduced into *E. coli* BW25113/pIJ790 and underwent homologous recombination with pSET152/*bla* to introduce elements into the plasmid DNA necessary for conjugation with *Streptomyces* chromosome. (g) SuperCos/contig 557/noz::scar underwent conjugation with *Streptomyces* to integrate the SuperCos plasmid into *Streptomyces* via homologous recombination. (h) SuperCos/contig 557/noz::scar was introduced into *Streptomyces* chromosome.

Comparison of metabolite profiles between heterologous S. coelicolor M1152 and S. lividans TK24 carrying contig 557 or contig 557/nozA::scar, nozD::scar or nozE::scar

Metabolite profiles were compared between *Streptomyces* featuring heterologous *noz* cluster expression constructs (contig 557, *nozA*::scar, *nozD*::scar and *nozE*::scar) to screen for the production of alternative metabolites as a result of the PCR-targeted gene replacements. Specifically, one hundred microliters of 3-day starter cultures of heterologous *S. coelicolor* M1152 and *S. lividans* TK24 constructs (contig 557, *nozA*::scar, *nozD*::scar, *nozE*::scar) were inoculated into 10 mL of M1 media (lacking Instant Ocean) supplemented with 50 µg/mL apramycin in 50 mL falcon tubes. Wild type *S. coelicolor* M1152 and *S. lividans* TK24 were cultured alongside as controls lacking the addition of apramycin. The treatment and control cultures were run in duplicates and incubated at 30 °C and 200 rpm for 5 days. Following incubation, the cultures were extracted with 10 mL EtOAc. The chemical extracts were dried using a vacuum controlled rotavapor (BUCHI) and the remaining extracts were dried using a SpeedVac concentrator (Savant ISS1100, Thermo Scientific) overnight. The extracts were resolubilized in 150 µL MeOH for analysis by HPLC, described

previously. Chemical profiles produced were compared using ChemStation (Agilent).

Results and Discussion

Bioinformatics analyses predicted biosynthetic pathways from Nocardiopsis sp.

FGENESB (Softberry) was used to predict putative genes and their corresponding protein sequences upstream and downstream from *nozA* and *ncdA,* CDPSs responsible for catalyzing *cyclo*(L-Trp-L-Trp) assembly. The CDPS enzymes were found to reside in two different regions of *Nocardiopsis* sp. genome, resulting in two CDPS-containing biosynthetic gene clusters, referred to as contig 557 and contig 96 (Table 1, Figure 16 a-b). The contig 557 gene cluster spans approximately 15 Kb and contains a CDPS homolog (*nozA*) along with several other putative DKP tailoring enzymes. Two tailoring enzymes of interest within the contig 557 gene cluster were identified as cytochrome P450 homologs and named NozD and NozE. The genes encoding NozD and NozE are located directly downstream from *nozA* and play hypothesized roles in the cyclization of the C3'-normal prenyl group of nocardioazine B to yield the group observed in nocardioazine A (Table 1, Figure 16 a). Additional bioinformatics analyses of *Nocardiopsis* sp. CMB-M0232 genome revealed another gene cluster, contig 119, with hypothesized roles in nocardioazine biosynthesis (Table 1, Figure 16 c). A putative prenyltransferase, NozC, is hypothesized to catalyze the C3'-

normal prenylation of the DKP core first seen in nocardioazine B. In this same gene cluster and directly upstream from NozC is NozB, a putative methyltransferase with hypothesized roles in C- and N-methylations of the DKP scaffold also first seen in nocardioazine B (Figure 14).

Table 1. Predicted functions and homologs of putative proteins encoded by the hypothesized nocardioazine biosynthetic gene clusters: contig 557, contig 96 and contig 119. # aa = number of amino acid residues; ID = % identity; Sim = % Similarity.

contig 557 genes (bp)	**# aa**	**BLASTP annotation**	**NCBI accession number of homolog**	**ID / Sim (%)**
1512-2777	421	pyridoxas-5'-phosphate-dependent protein subunit beta	WP_037674793	56/69
2842-4050	402	major facilitator superfamily transporter	AEF16056	58/72
4093-4155	20	beta-N-acetylglucosaminidase	KFA31228	73/90
4235-4939	234	tRNA-dependent cyclodipeptide synthase	WP_033352817	42/53
5023-6219	398	cytochrome P450	WP_045298286	40/54
6312-7373	353	cytochrome P450	WP_045298286	40/54
7451-9994	847	adenylocsuccinate synthetase	ELQ79671	58/69
10221-10550	109	4-carboxymuconolactone decarboxylase	WP_020270970	75/87
10810-11733	307	XRE family transcriptional regulator	WP_044572329	80/89
11721-12809	362	ABC transporter ATP-binding protein	WP_046471802	73/82
12806-13675	289	molybdate ABC transporter	WP_017566834	76/88
13947-14780	277	molybdate-binding protein	WP_046471807	71/78
14773-15183	136	MerR family transcriptional regulator	WP_017544945	76/84
15254-15904	216	hypothetical protein	ERT00338	36/52
contig 96 genes (bp)	**# aa**	**BLASTP annotation**	**NCBI accession number of homolog**	**ID / Sim (%)**
1-379	126	cell division protein	WP_033301593	72/83
431-1063	210	hypothetical protein	WP_033301600	36/49
1527-3068	513	AMP-dependent synthetase	WP_042418882	74/82
3083-3811	242	CDPS	WP_035283888	70/80
3897-4760	287	cell wall hydrolase	WP_026126175	63/71
4924-5283	119	hypothetical protein	WP_017620737	69/84
5502-6452	316	erythromycin esterase	WP_046725140	38/51
6846-7481	211	transcriptional regulator	WP_011211565	58/74
7689-9269	526	peptide ABC transporter substrate-binding protein	WP_041588039	40/57
9435-9824	129	extradiol dioxygenase	WP_033381084	78/85
9947-10945	332	nickel ABC transporter permease	WP_015051569	41/55
10970-11806	278	serine protease	WP_037956543	55/72
11949-12185	78	multidrug transporter	WP_031172074	57/70
12310-12423	37	polyketide synthase PKS1	AGL28386	52/58
12967-13338	123	beta-galactosidase	WP_019609329	57/62
contig 119 genes (bp)	**# aa**	**BLASTP annotation**	**NCBI accession number of homolog**	**ID / Sim (%)**
105-1118	337	hypothetical protein	WP_017620974	73/82
1223-2152	309	3-ketoacyl-ACP reductase	WP_026126235	77/84
2501-3805	434	glycosyl transferase family 1	WP_017620972	85/92
3916-4554	212	methyltransferase	WP_046468151	86/92
4551-5627	358	prenyltransferase	WP_033301783	79/84
5668-8013	781	hypothetical protein	WP_040270987	69/81

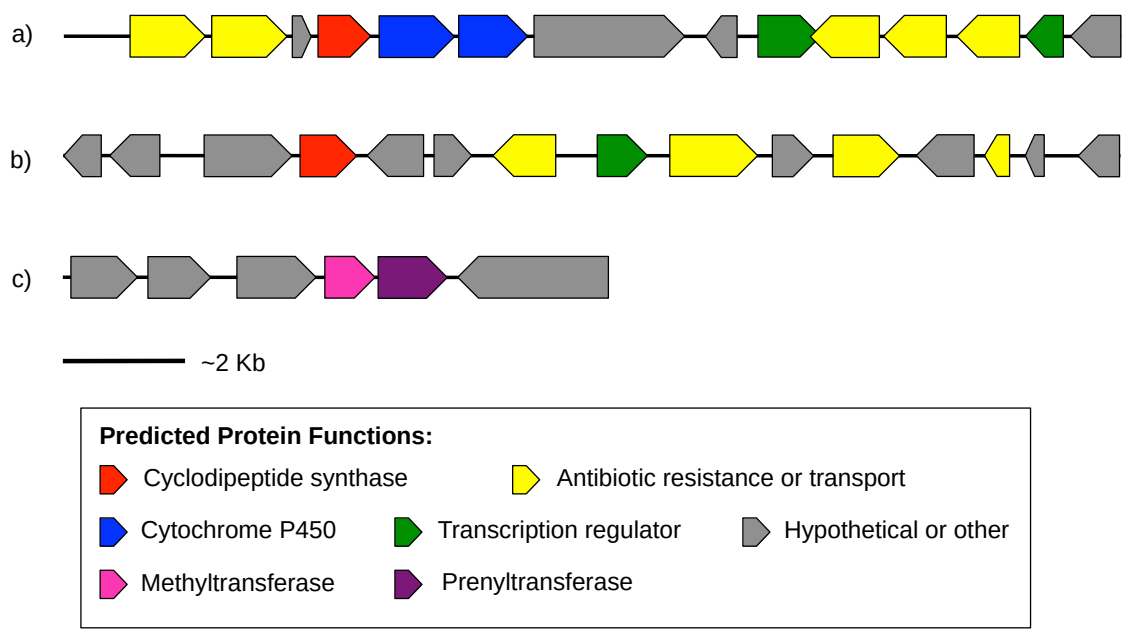


Figure 16. Bioinformatics-predicted *Nocardiopsis* sp. gene clusters with hypothesized roles in *cyclo*(L-Trp-L-Trp) and nocardioazine A-B biosynthesis. Cluster a) represents contig 557 which spans ~15 Kb, b) represents contig 96 which spans ~15 Kb and c) represents contig 119, spanning ~8 Kb. Arrows indicate the predicted direction of gene transcription. Chapter 2 established the function of the cyclodipeptide synthases (red) in production of *cyclo*(L-Trp-L-Trp).

Most CDPS-containing natural product biosynthetic pathways carry genes encoding oxidases and almost all contain a gene encoding a cytochrome P450. But, contig 557 presents the only identified natural product biosynthetic pathway thus far that encodes two cytochrome P450s directly downstream from a CDPS gene, suggestive of unique tailoring capabilities from the contig 557 gene cluster.

Isolation of nocardioazines A-B from wild type Nocardiopsis sp. CMB-M0232

Nocardioazines A-B were isolated from wild type *Nocardiopsis* sp. CMB-M0232 in order to 'feed' heterologous *Streptomyces* the putative precursors in nocardioazine A biosynthesis and to provide a standard of the pathway end product, nocardioazine A. Following EtAOc extraction of wild type *Nocardiopsis* sp. CMB-M0232, the crude extracts were subjected to sequential trituration to concentrate the nocardioazines in a single fraction. LC/MS identified the largest quantities of nocardioazine A-B present in the *n*-hexane extracts, a result of their relatively non-polar chemical structures. To determine which peaks from the *n*-hexane fraction HPLC spectra were representative of nocardioazines A-B, eluents were collected from the HPLC column in two-minute increments. LC/MS identified that nocardioazine A eluted from the column between minutes 26 and 28 (Figure 17 a) and nocardioazine B eluted from the column between minutes 28 and 30 (Figure 17 b). LC/MS then confirmed the individual peaks within those time frames that were representative of nocardioazine A-B. From 5 L of *Nocardiopsis* sp. CMB-M0232 culture, ~3.0 mg of nocardioazine B (pale white powder) was purified and ~1.5 mg of nocardioazine A (pale white powder) was purified. NMR validated the chemical structure of the purified compounds, confirming isolation of pure, nocardioazines B and nocardioazine A (Appendix E-F).

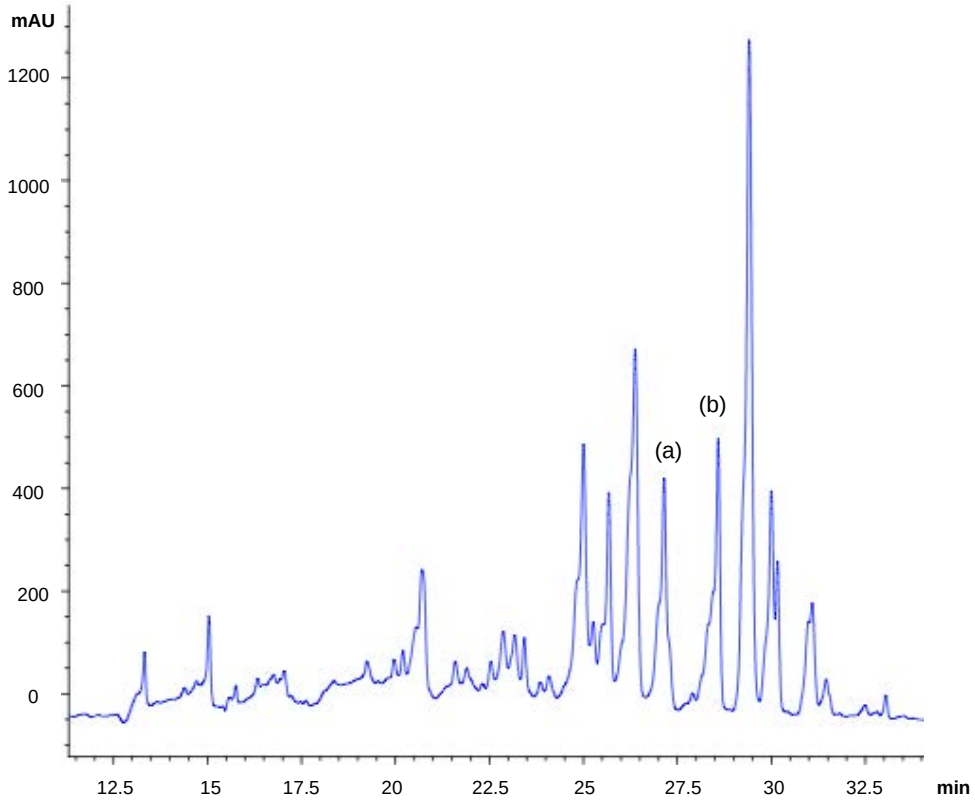


Figure 17. HPLC spectra of *n*-hexane extracts that contain (a) nocardioazine A and (b) nocardioazine B. Peaks (a) and (b) were collected and their contents verified via NMR. Chromatogram shows absorbance at 280 nm.

Generation of noz* cluster *constructs for heterologous expression in Streptomyces

To facilitate conjugation with *Streptomyces,* the SuperCos cosmid from the gDNA clone library carrying contig 557 (called SuperCos/557 and featuring the *noz* cluster of genes including *nozA, nozD* and *nozE*) as well as the SuperCos cosmids carrying contig 557 featuring PCR-targeted mutants (*nozA*::scar, *nozD*::scar and *nozE*:: scar) underwent homologous recombination

with a fragment of the pSET152 plasmid vector to introduce the following elements into SuperCos: *aac(3)IV, oriT* and *θ*C31 integrase. These components are essential for the introduction and expression of genes carried by SuperCos (i.e. *noz* genes) in *Streptomyces* hosts. PCR confirmed homologous recombination between SuperCos constructs carrying mutated (*noz*::scars) and unmutated contig 557 and the linearized pSET152 plasmid vector using two sets of primers that were designed to amplify both ends of the region where this DNA was introduced into SuperCos. The constructs confirmed to carry all three elements were introduced into *S. coelicolor* M1152 and *S. lividans* TK24 (Figure 18) by conjugation.

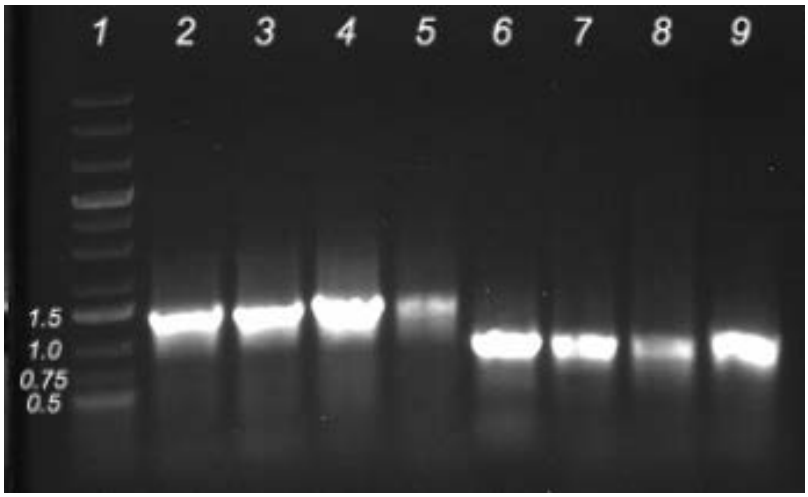


Figure 18. Agarose gel verifying homologous recombination between SuperCos1 featuring *Streptomyces* gene expression elements from pSET152 plasmid vector. Lanes 2-5 contain the amplified region of the constructs spanning the pUC-Apr^R gene and were expected to amplify a 1301 bp product. Lanes 6-9 contain the amplified region of the constructs linking the *bla* from SuperCos with pSET152 and were expected to amplify a 700 bp product. Lane 1 contains GeneRuler 1 kb Plus DNA Ladder, lanes 2 and 6 contain SuperCos carrying contig 557, lanes 3 and 7 contain SuperCos carrying the *nozA*::scar, lanes 4 and 8 contain SuperCos carrying the *nozD*::scar, and lanes 5 and 9 contain SuperCos carrying the *nozE*::scar.

Following conjugation, candidate *Streptomyces* transformants were screened via PCR to confirm the incorporation of contig 557 or the contig 557 featuring PCR-targeted mutants (*nozA*::scar, *nozD*::scar and *nozE*::scar) into *Streptomyces* chromosome. The ~3 Kb region of the contig 557 gene cluster carrying *nozA, nozD* and *nozE* was confirmed present in both *S. coelicolor* M1152 and *S. lividans* TK24 transformants using PCR (Figure 19). The mutants were also confirmed present in both strains of *Streptomyces* transformants (Figure 20). The unmutated contig 557 gDNA served as a positive control to identify successful gene replacement and allow comparison of metabolites between treatments and controls. Colony PCRs of cultures featuring contig 557 revealed amplification of the gene of interest. Replacement of the gene with an 81 bp-scar resulted in the expected reduction of the PCR product size.

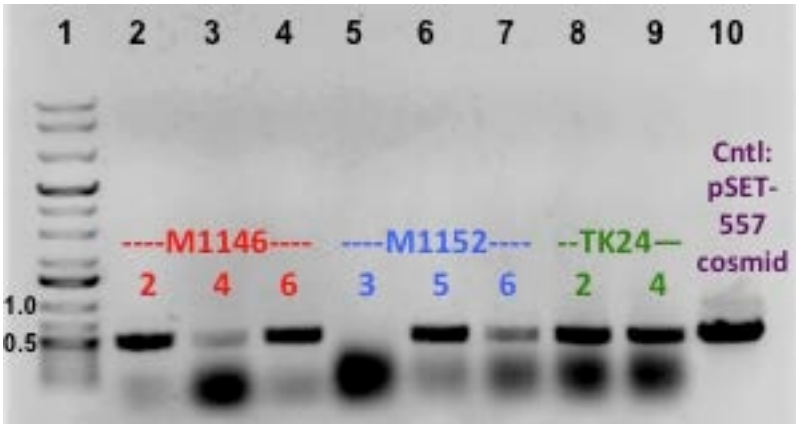


Figure 19. Agarose gel verifying integration of contig 557, containing *nozA, nozD* and *nozE,* into *S. coelicolor* M1152 and *S. lividans* TK24. Lane 1 contains GeneRuler 1kb Plus DNA Ladder and lanes 2-4 contain contig 557 integrated into *S. coelicolor* M1146 which was not used in this current thesis study due to failed conjugation of mutant contig 557 constructs with *S. coelicolor* M1146. Lanes 5-7 contain contig 557 integrated into *S. coelicolor* M1152, lanes 8-9 contain contig 557 integrated into *S. lividans* TK24, and lane 10 contains the pSET152/contig 557 cosmid to serve as a PCR control. PCR amplified an ~500 bp region of *nozA* to confirm the presence of the contig 557 gene cluster in *Streptomyces.*

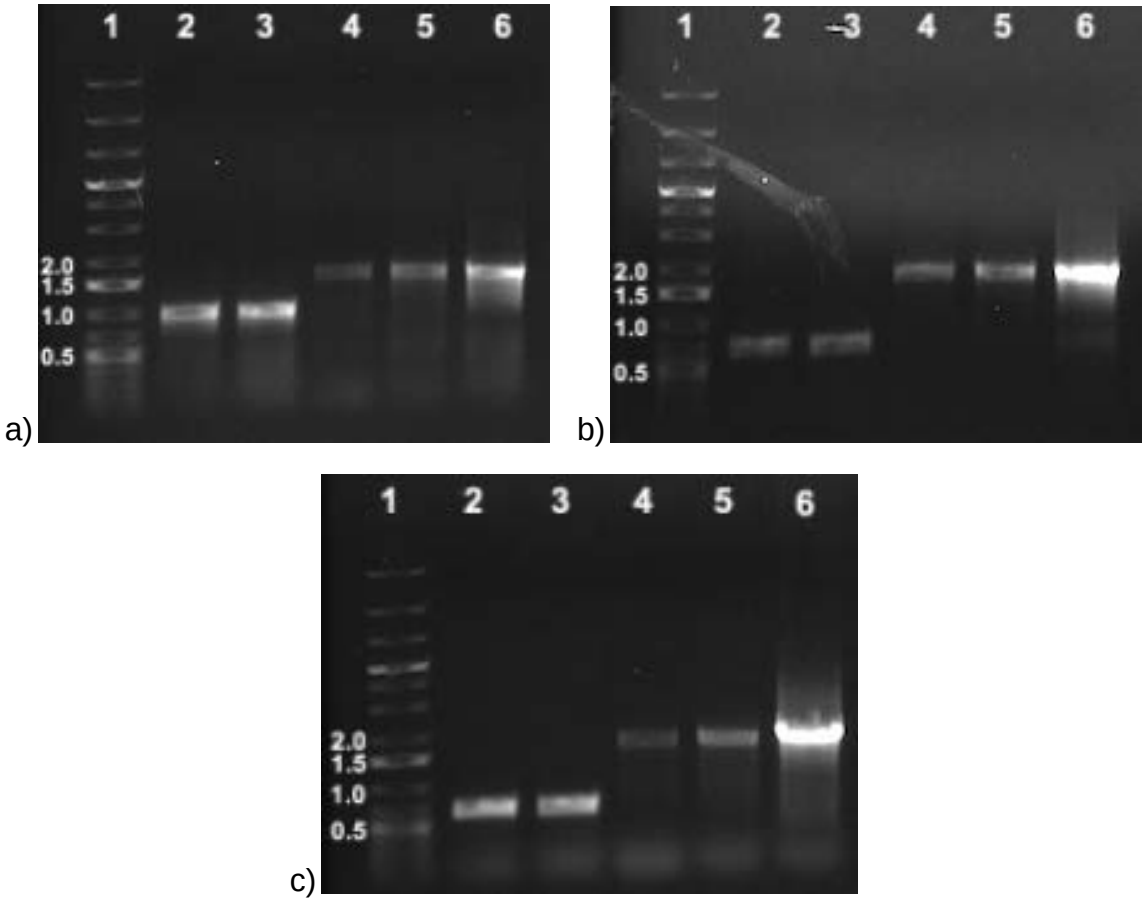


Figure 20. Agarose gels verifying integration of contig 557 gene elimination mutations via PCR-targeted gene deletions into *S. coelicolor* M1152 and *S. lividans* TK24. (a-c) Lane 1 contains GeneRuler 1kb Plus DNA Ladder, lanes 2-3 contain *S. lividans* TK24 and *S. coelicolor* M1152 carrying the contig 557/*noz*::scars, respectively, lanes 4-5 contain *S. lividans* TK24 and *S. coelicolor* M1152 carrying contig 557 (unmutated), respectively, and lane 6 contains SuperCos-1 carrying contig 557 to serve as a PCR control. Gel (a) contains *Streptomyces* carrying the contig557/*nozA*::scar constructs, gel (b) contains *Streptomyces* carrying the contig557/*nozD*::scar constructs, and gel (c) contains *Streptomyces* carrying the contig557/*nozE*::scar constructs. The reduction in size observed in the gene deletion constructs indicates deletion of that gene compared to the constructs carrying the entire gene cluster.

In vivo characterization of functions of NozD and NozE exposed to nocardioazine B to evaluate their roles in biosynthesis of nocardioazine A

Nocardioazine B, purified from wild type *Nocardiopsis* sp. CMB-M0232, was fed to *Streptomyces* expressing contig 557 genes, including *nozA, nozD* and *nozE.* HPLC was used to evaluate accumulation of nocardioazine A from these cultures, since NozD and/or NozE were hypothesized to catalyze the transformation of nocardioazine B to nocardioazine A. Metabolite profiles were compared between control, wild type *S. lividans* TK24 and *S. coelicolor* M1152, with treatments of these strains carrying the contig 557 gene cluster (including *nozA, nozD,* and *nozE*) following fermentation supplemented with nocardioazine B (Figure 21-22). HPLC spectra were compared to synthetic *cyclo*(L-Trp-L-Trp) and pure nocardioazine A-B standards to determine the presence of these metabolites. *Cyclo*(L-Trp-L-Trp) was present in all metabolite profiles from *Streptomyces* carrying the contig 557 gene cluster, supporting that the gene encoding NozA was translated in the correct reading frame and therefore, suggesting NozD and NozE from the same operon were present in the cells as well. Nocardioazine B was detected by HPLC in samples in which it was supplemented, but nocardioazine A was not detected in the chemical profiles of *Streptomyces* expressing NozD and NozE and provided nocardioazine B (Figure 21-22). It is unclear whether or not nocardioazine B was able to enter the cells to interact with NozD and NozE, offering one possibility as to why there was no detectable accumulation of nocardioazine A. Thus, an *in vitro* approach was

undertaken to circumvent the issue of nocardioazine B entering the cell and instead allow for direct contact between nocardioazine B, NozD and NozE.

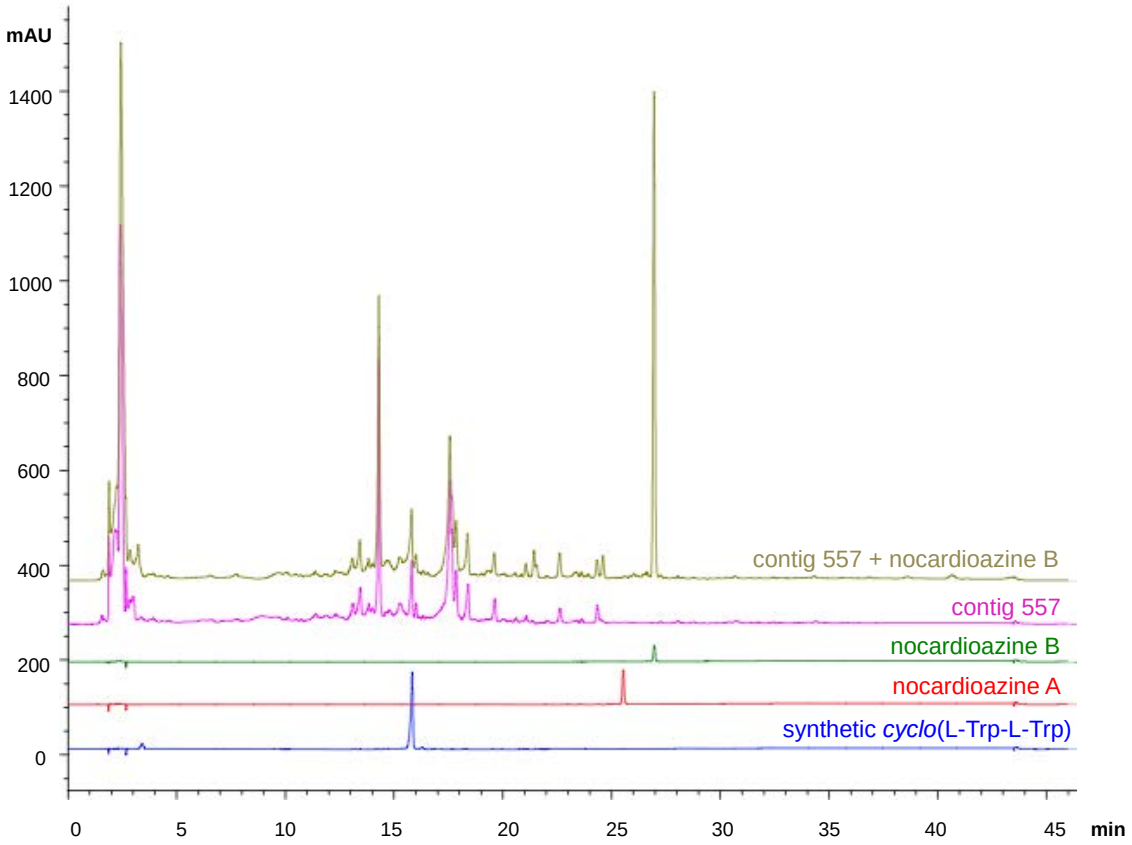


Figure 21. HPLC spectra comparing metabolite profiles from heterologous *Streptomyces lividans* TK24 carrying contig 557 exposed to nocardioazine B and not exposed to nocardioazine B, *in vivo*. Synthetic *cyclo*(L-Trp-L-Trp) and purified nocardioazines A-B serve as controls. Following fermentation of *S. lividans* TK24 carrying contig 557 in media supplemented with nocardioazine B, nocardioazine A was not detected. Chromatograms show absorbance at 254 nm.

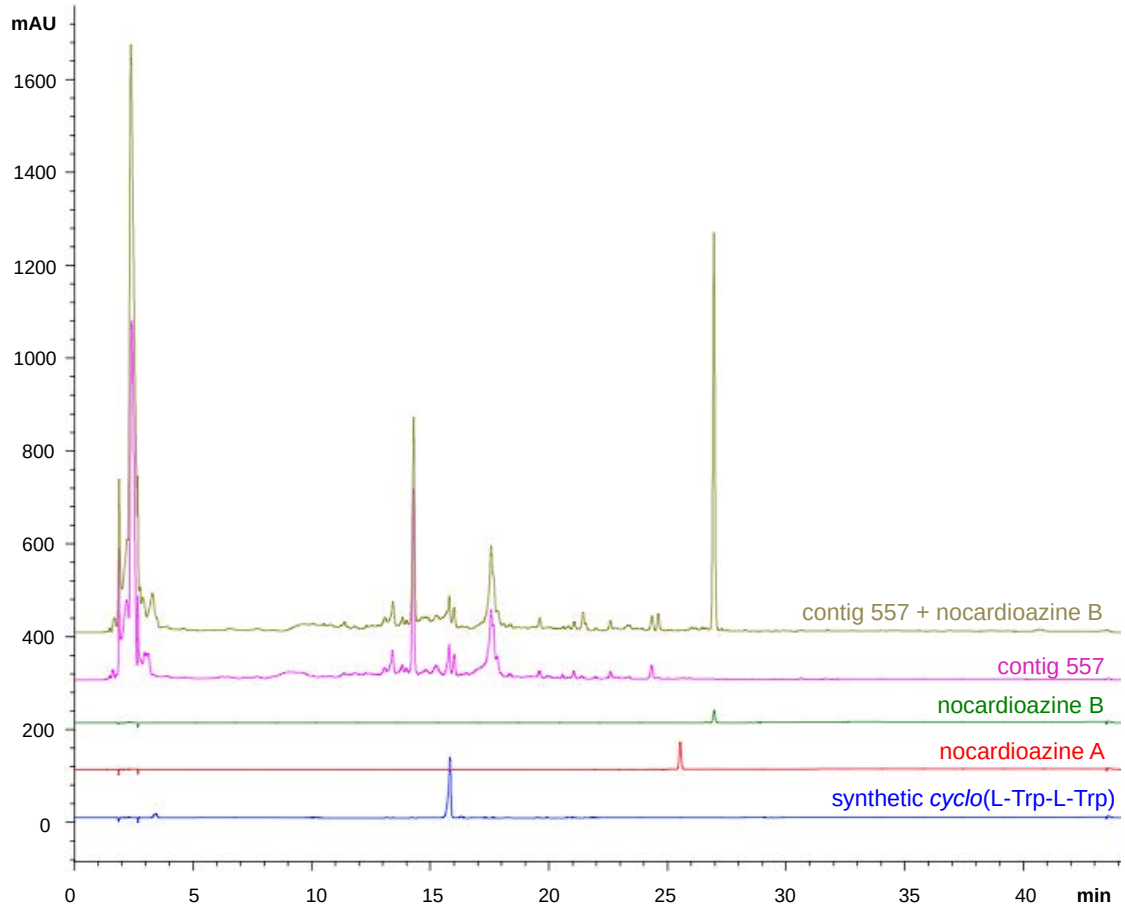


Figure 22. HPLC spectra comparing metabolite profiles from heterologous *Streptomyces coelicolor* M152 carrying contig 557 exposed to nocardioazine B and not exposed to nocardioazine B, *in vivo*. Synthetic *cyclo*(L-Trp-L-Trp) and purified nocardioazines A-B serve as controls. Following fermentation of *S. coelicolor* M1152 carrying contig 557 in media supplemented with nocardioazine B, nocardioazine A was not detected. Chromatograms show absorbance at 254 nm.

In vitro characterization of functions of NozD and NozE when exposed to nocardioazine B to evaluate their roles in biosynthesis of nocardioazine A

Following the *in vivo* results failing to show NozD- or NozE-catalyzed conversion of nocardioazine B to A, an alternative *in vitro* approach was taken to probe the enzymatic activity of NozD and NozE. *In vitro* assays combined cell

lysates from control, wild type *S. lividans* TK24 and treatments of this strain carrying the contig 557 genes (*nozA, nozD and nozE)* with pure nocardioazine B. The metabolite profiles from the previous mentioned assays were compared to the chemical profiles from control and treatment *S. lividans* TK24 lysates that were incubation without the addition of nocardioazine B (Figure 23). This approach did not require that nocardioazine B enters the cell to interact with NozD and NozE, but instead provided direct contact between nocardioazine B and the cell lysates containing the cytochrome P450s, NozD and NozE.

HPLC metabolite profiles were compared to the *cyclo*(L-Trp-L-Trp) and nocardioazine A-B standards to detect the presence of these metabolites. As expected, *cyclo*(L-Trp-L-Trp) was present in all metabolite profiles from *Streptomyces* carrying the contig 557 gene cluster. Nocardioazine B was detected in the samples in which it was supplemented but Nocardioazine A was not detected in the chemical profiles where *Streptomyces* expressing NozD and NozE were provided nocardioazine B *in vitro.* These data suggested that either NozD and/or NozE were not adequately expressed, lacked essential cofactors for activity when expressed in *Streptomyces* hosts, or catalyzed reaction(s) other than the modifications of nocardioazine B to A. Thus, the possibility that these enzymes carry out other modifications of *cyclo*(L-Trp-L-Trp), from NozA, was next explored.

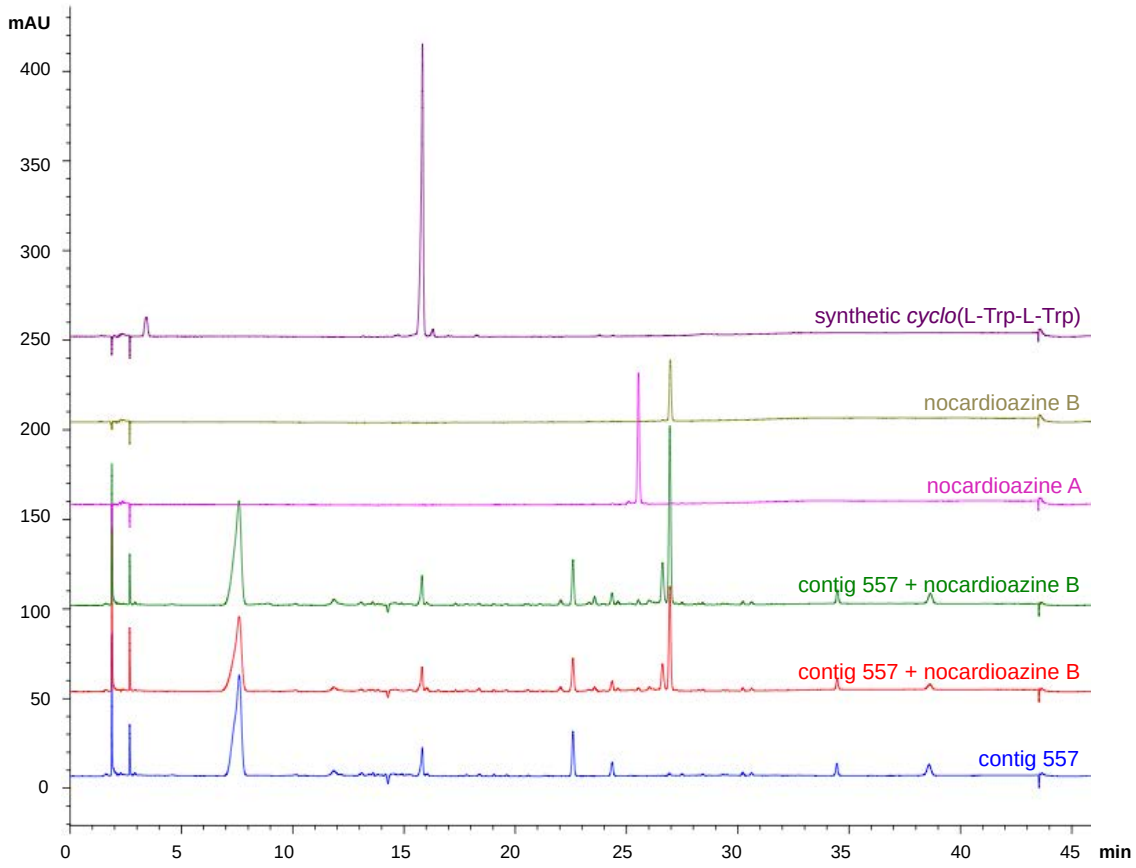


Figure 23. HPLC spectra comparing metabolite profiles from *in vitro* treatment assays that combined cell lysates from *Streptomyces lividans* TK24 carrying contig 557 with nocardioazine B and control assays where lysates incubated without the addition of nocardioazine B (blue). Treatments included cell lysates from 3-day (red) and 6-day cultures of *Streptomyces* treatments incubated with nocardioazine B (green). Following incubation of the lysates with nocardioazine B, nocardioazine A was not detected. Furthermore, no alternative metabolites were detected in assays incubated with nocardioazine B compared to the control that was incubated without the addition of nocardioazine B. Chromatograms show absorbance at 210 nm.

Comparing chemical profiles of heterologous Streptomyces noz pathway cultures carrying PCR-targeted gene mutants to evaluate the accumulation of alternative DKP-derived natural products

Since the previous experiments failed to establish enzymatic functions of NozD and NozE in conversion of nocardioazine B to A, we next explored the

possibility that these enzymes directly modify *cyclo*(L-Trp-L-Trp). Metabolite profiles from *Streptomyces* carrying individual PCR-targeted mutations from the contig 557 cluster (*nozA*::scar, *nozD*::scar or *nozE*::scar) were compared to *Streptomyces* carrying the entire contig 557 gene cluster to detect accumulation of alternative metabolites accumulated upon deletion of any of the three genes (Figure 24-25). *Cyclo*(L-Trp-L-Trp) was present in each metabolite profile from heterologous *Streptomyces* except the *nozA*::scar mutant, where *nozA* was deleted. This further linked *nozA* to the production of *cyclo*(L-Trp-L-Trp) and corroborated the finding from chapter 2. The presence of *cyclo*(L-Trp-L-Trp) in the other constructs suggested that the genes were being translated in the correct reading frame and that polar effects did not result from gene deletions.

No metabolites were present in *Streptomyces* carrying contig 557 but missing from those carrying the *nozD*::scar and *nozE*::scar gene mutants by HPLC (Figure 24-25). Thus, the enzymatic functions of the cytochrome P450s, NozD and NozE, are still yet to be confirmed. Similarities of the chromosomal region surrounding *nozA* to other CDP-modifying biosynthetic pathways, still suggest that contig 557 contains likely candidates for the tailoring of nocardioazine B to form nocardioazine A. Further studies utilizing purified recombinant enzymes may allow establishment of the functions of NozD and NozE by permitting careful control over enzyme concentrations, cofactors, and other variables that strongly impact enzyme activity. This thesis study sets the stage for such future studies.

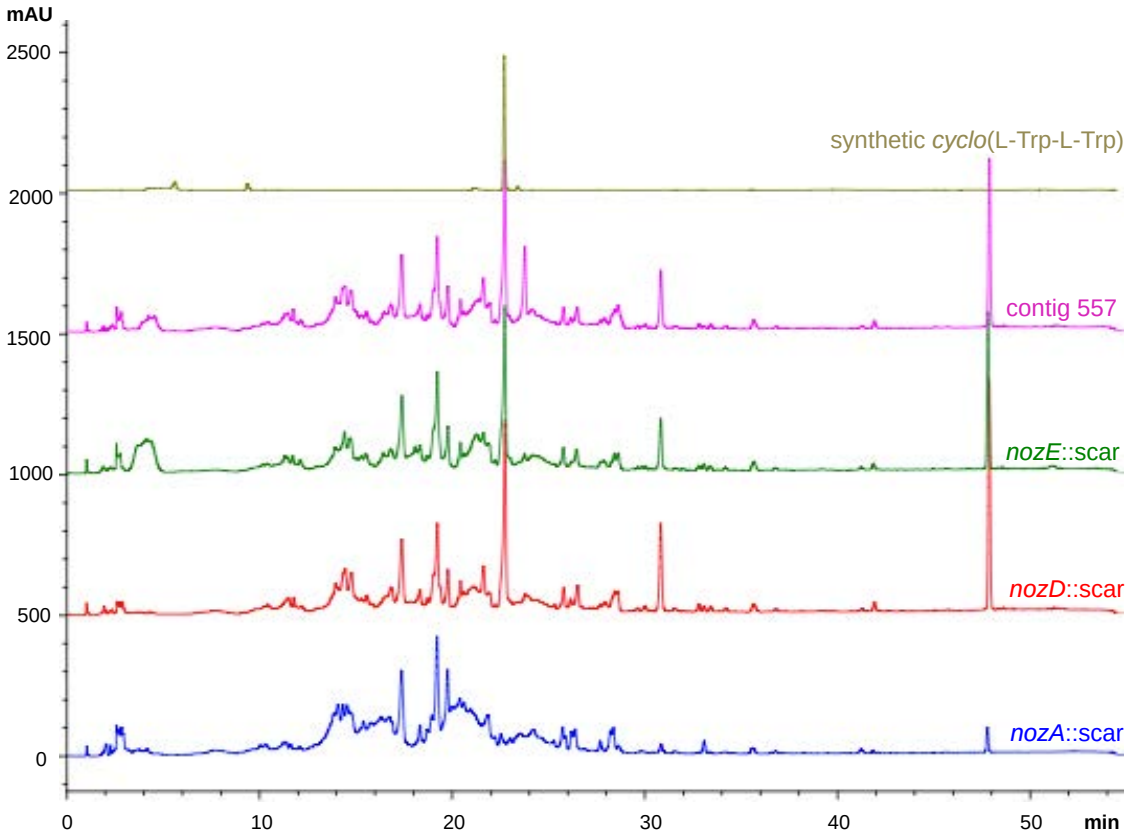


Figure 24. HPLC spectra comparing metabolite profiles between heterologous *Streptomyces lividans* TK24 carrying the full 557 gene cluster (*nozA, nozD* and *nozE)* and those carrying the PCR-targeted gene mutants (*nozA*::scar, *nozD*::scar and *nozE*::scar). Synthetic *cyclo*(L-Trp-L-Trp) served as a control. Depletion of *cyclo*(L-Trp-L-Trp) production in the *nozA*::scar mutant further confirmed the function of *nozA* in *cyclo*(L-Trp-L-Trp) assembly. No metabolites differed between metabolite profiles from *S. lividans* TK24 carrying contig 557 and those carrying contig 557 with PCR-targeted gene deletions. Chromatograms show absorbance at 210 nm.

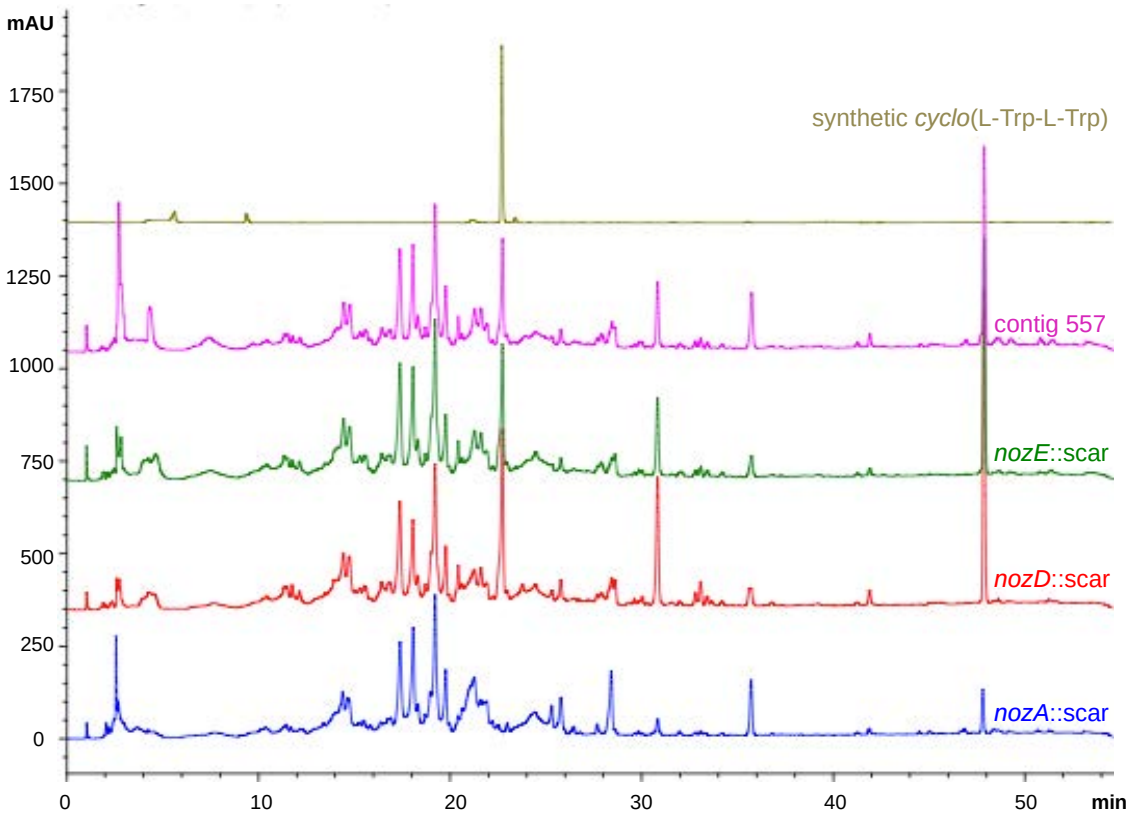


Figure 25. HPLC spectra comparing metabolite profiles between heterologous *Streptomyces coelicolor* M1152 carrying the full 557 gene cluster (*nozA, nozD* and *nozE)* and those carrying the PCR-targeted gene mutants (*nozA*::scar, *nozD*::scar and *nozE*::scar). Synthetic *cyclo*(L-Trp-L-Trp) served as a control. Depletion of *cyclo*(L-Trp-L-Trp) production in the *nozA*::scar mutant further confirmed the function of *nozA* in *cyclo*(L-Trp-L-Trp) assembly. No metabolites differed between metabolite profiles from *S. coelicolor* M1152 carrying contig 557 and those carrying contig 557 with PCR-targeted gene deletions. Chromatograms show absorbance at 210 nm.

Conclusions and Future Directions

This thesis study provides insight into the biosynthetic pathway from *Nocardiopsis* sp. CMB-M0232 responsible for the catalytic assembly of the nocardioazines, novel 2,5-DKP natural products particularly interesting due to their unique chemical structures and the promise of nocardioazine A as an anticancer agent. Two CDPS homologs identified in the draft genome sequence of *Nocardiopsis* were biochemically characterized via *in vivo* and *in vitro* methods to reveal that each catalyzes the production of *cyclo*(L-Trp-L-Trp), the hypothesized precursor in the nocardioazine biosynthetic pathway. Evaluation of the substrate specificity of NozA and NcdA revealed that both CDPSs are specific for tryptophanyl-tRNA substrates, catalyzing the production of *cyclo*(L-Trp-L-Trp) and no other 2,5-DKPs with alternative aromatic aminoacyl-tRNAs. This finding renders NozA and NcdA two of only three characterized CDPSs that are substrate specific, incorporating only one aminoacyl-tRNA substrate. Furthermore, *Nocardiopsis* provides a rare example of a single bacterium encoding two distinct enzymes that catalyze biosynthesis of the same secondary metabolite. The physiological relevance of DKP-derived natural products is still largely speculative, as little is known about the actual use of these metabolites in the native organism. A more thorough understanding of their biological importance may provide clues as to why *Nocardiopsis* has evolved to use two phylogenetically identical enzymes in the biosynthesis of *cyclo*(L-Trp-L-Trp), a precursor of the nocardioazines.

The chromosomal region surrounding *nozA* contains putative CDP-tailoring enzymes with roles predicted in the modifications made to nocardioazine B to produce nocardioazine A. NozD and NozE, two cytochrome P450 homologs downstream from *nozA* were heterologously expressed with NozA in *Streptomyces* to determine their involvement in assembly of nocardioazine A. The approaches employed by this study were not able to confirm the enzymatic activity of these enzymes, but lead the way toward future studies to optimize the conditions for probing NozD and NozE activity. In addition, future work will aim to elucidate the biological properties of the methyltransferase and prenyltransferase homologs, NozB and NozC, respectively, from contig 119. These enzymes are also hypothesized to play roles in nocardioazine assembly (Figure 14). Uncovering their roles in the biosynthesis may link *cyclo*(L-Trp-L-Trp) produced by NozA and NcdA with the modified scaffold characteristic of nocardioazine B.

The results gathered from this thesis study have offered novel cyclodipeptide-synthesizing enzymes, adding to the knowledge of a relatively understudied class of enzymes that offer novel natural products. The characterization of NozA and NcdA combined with future work to better understand the unique engineering of the C3'-prenyl functional group and other nocardioazine groups by *Nocardiopsis* may provide new opportunities for pathway engineering to produce new bioactive drugs.

Appendix

Appendix A. Sequence of the *E. coli* optimized *nozA* gene, optimized and synthesized by GeneArt (Life Technologies).

nozA
GAATTGGCGGAAGGCCGTCAAGGCCACGTGTCTTGTCCAGAGCTCCATATGCA
TACCGTTAGCGTTGAACGTGATGGTGATTTTCATGTTACACCGCTGACCGATAAT
TGTCGTGATCTGATGAAAGCCGGTGGTCATGCACTGCTGGGTGTTAGCCATGG
TAATAGTTATTTTAGCCGTGGTCGTCTGAGCAACCTGTTTGCATGGGCACTGCG
TCGTTTTGATGCAGTTGATGTTGTTGCAGCAGATAGCCATGTTGTGGAAATTTTT
CGTGCCATTGGCTATGATGATGAACATGCACGTAAACGCACCCATAAAGAAACC
AGCGTTCTGCGTAATCGTATTCGTGATGCAGCAGCACAGTGTGGTGCAGATCG
TACCCGTATTACCCTGCGTGGTCTGAGCGATTTTATGGATGATCCGGCATATCGT
AGCGTTCTGAGCGAAACCCAGAAAGCACTGGATGATCAGCCGGAATTTCGTAA
TGCAGCCCGTGCAATGACCCGTGCAGTTCTGGAAAGCCGTCTGGGTCCGGGT
AATGCACGTGAAGATCAGATTGATATTGCCTTTAGCTATCTGCAGGCAGAACTG
CCGTTTTTTCATGATGCACCGCGTATTCTGCGTGTTGAAAGCACCGTTAGCTGT
TATCACATGCGTCTGCCGCTGCTGGATTTTATCTGTGGTCCGCGTGCAGTTGTT
CCGGTTGTGCCGACCCAGGGTTTTGCCGTTGTTCGTCCGGCAGATCCGGGTA
AGCTTGGTACCTGGAGCACAAGACTGGCCTCATGGGCCTTCCGCTCACTG

Appendix B. Sequence of the *E. coli* optimized *ncdA* gene, optimized and synthesized by GeneArt (Life Technologies).

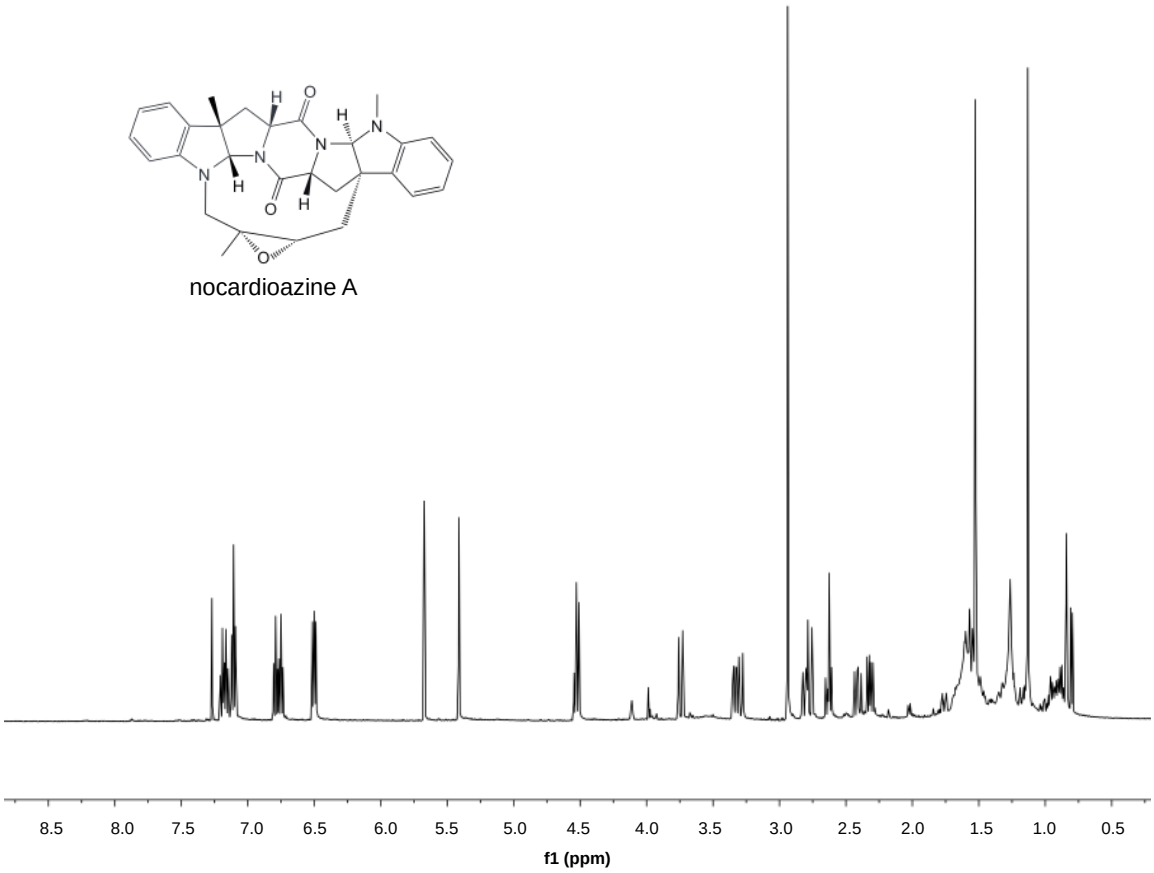
ncdA
ATTGGCGGAAGGCCGTCAAGGCCACGTGTCTTGTCCAGAGCTCCATATGAGC
GGTCCGCATGCAGTTCCGGATGGTGGTAGCGCACCGCCTCCGAGCCGTCGTG
CCAGCTTTACCGTTGAACCGTTTACCGATGCATGTCATGCAGTTTGGGAACAGC
GTCGTCATGCCGTTCTGGGTGTTAGTCCGGGTAATGGCTATTTTAACGTTGCAC
TGCTGACCGAACTGCTGGGTTGGGCATGTGGTGAATTTGCACGTGTTGATGTT
GTTGTTCCGGATAGCGCACTGGAACACACCTATCTGGCACTGGGTTATGATCTG
CGTCGTGCAGCAAAAAAAGCACGTGGTGAAACCAATGTTCTGCGTAATCGTGT
TGTTCGTGCATGGGAAGCCGGTGGTGGTCCGCGTGCAAGTGATGGTCTGCAT
CGTATGAGCGAACTGGCAAGCAATGCAGCATATCGTGCAGGTCTGGCAGAATG
TGAACGTGCACTGGGTGAAGATGATCTGCTGTGGGAAACCTGTGCAGAAATGA
GCCGTGATGTTCTGGCAGCCCGTGGTCATGATGGTCCGCTGACCACCGAACG
TGTTGAACGTGCCATGCGTTATCTGACCGCAGAACTGCCGTTTTTTCTGGCAA
GCGCAGATATTTTTGGTGTTCCGAGCAGCCTGAACTTTTATCATCGTCGTCTGC
CGCTGGCAGAAGTTGTTTTTGCAGGTAAAAGCGTTCTGCAGGCACCGCCTGCA
CAGGGTTATGCAACCATTCGTCCGGCAGGTCGTGATCCTGATCGTGCAAAGCT
TGGTACCTGGAGCACAAGACTGGCCTCATGGGCCTTCCGCTCAC

Appendix C. List of primer sequences used in this study. Regions in bold are FLP recognition targets.

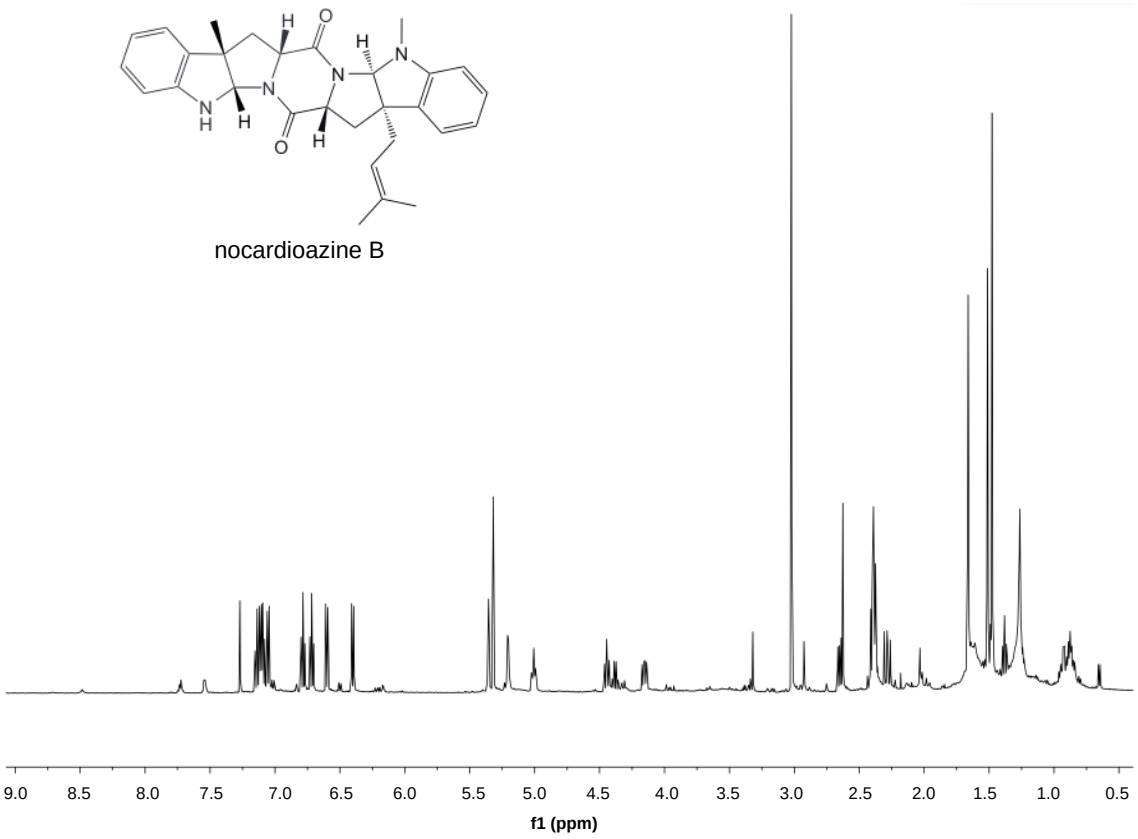
Primer	Sequence (5' --> 3')
1NozA_SDM_F	TGATGCAGCAGCACAGTGTGGTGCAG
2NozA_N114A_R	CGAATACGCGCACGCAGAACGCTGGT
2NozA_N114R_R	CGAATACGACGACGCAGAACGCTGGT
2NozA_V111RN114A_R	CGAATACGCGCACGCAGACGGCTGGT
NozA_L131R_F	GGCATATCGTAGCGTTCGTAGCGAA
NozA_L131R_R	GGATCATCCATAAAATCGCTCAGAC
NozA_S36A_F	CACTGCTGGGTGTTGCCCATGGTAA
NozA_S36A_R	CATGACCACCGGCTTTCATCAGATC
NozA_S198L_F	GTTGAAAGCACCGTTCTGTGTTATCAC
NozA_S198L_R	ACGCAGAATACGCGGTGCATCATGAA
	GCCCCTCGACGCCGATGATGGGAAGGACAGCCCGTGGTG**AT**
NozA::scarKO_F	**TCCGGGGATCCGTCGACC**
	GGGTGCCGGCCCGCAGGCGGGCCGGCGGCCCCTGGATCA**T**
NozA::scarKO_R	**GTAGGCTGGAGCTGCTTC**
	GGCCGAAACGCCACGGCCCCCACCGGAGGTTCCCGTTTG**AT**
NozD::scarKO_F	**TCCGGGGATCCGTCGACC**
	TGCCGGGCCTCGCTCGGGTCCGGGCGCGGCGGCGGGTCA**T**
NozD::scarKO_R	**GTAGGCTGGAGCTGCTTC**
	CCCCGGCGCACAGAGCCGATCCCGACGGGAGACCCCTTG**AT**
NozE::scarKO_F	**TCCGGGGATCCGTCGACC**
	TGGGCGCGGGGATCGGGGACCGCCGGGCGCCGGGGTCA**T**
NozE::scarKO_R	**GTAGGCTGGAGCTGCTTC**
SuperpSET_F1	CATGATCGTGCTCCTGTCGT
SuperpSET_R1	CTACGGAAGGAGCTGTGGAC
SuperpSET_F2	TAAATAGCTGCGCCGATGGT
SuperpSET_R2	GTATCCTGCGTGATGAGCCA
NozACnt557_F	ACTGCCGTGACCTCATGAAG
NozACnt557_R	TCTCCTACCTCCAAGCCGAA
NozA_KOCon_F	GCCGTTAACACCAGCTTCAC
NozA_KOCon_R	GGACCTCTTCCGCCACTATG
NozD_KOCon_F	ACCAGCCCGAATTCCGAAAT
NozD_KOCon_R	GTGGCTGATCACCTCCCAC
NozE_KOCon_F	GGTTCCCCACGATGTACCTG
NozE_KOCon_R	TGCTGAAGAACCACCTCGTC

Appendix D. Origin and accession number (NCBI) of biochemically characterized CDPSs including NozA and NcdA from *Nocardiopsis* and characterized in this thesis.

CDPS	Origin	NCBI accession number	Preferential products
Amir_4627	*Actinosynnema mirum*	ACU38460	c(WW)
YvmC_lic	*Bacillus licheniformis*	AAU25020	c(LL)
YvmC_sub	*Bacillus subtilis*	CAB15512	c(LL)
YvmC_thu	*Bacillus thuringiensis*	EAO57133	c(LL)
Jk0923	*Corynebacterium jeikeium*	CAI37087	c(LL)
Rv2275	*Mycobacterium tuberculosis*	P9WPF9	c(YY)
Nvec_CDPS2	*Nematostella vectensis*	EDO44063	c(WX)
Ndas_1148	*Nocardiopsis dassonvillei*	ADH66589	c(FY)
NcdA	*Nocardiopsis* sp. CMB-M0232	KT184401	c(WW)
NozA	*Nocardiopsis sp.* CMB-M0232	KT184400	c(WW)
Plu0297	*Photorhabdus luminescens laumondii*	CAE12592	c(LL)
pSHaeC06	*Staphylococcus haemolyticus*	BAE05998	c(LL)
AlbC	*Streptomyces noursei*	AAN07909	c(LF)



Appendix E. ^{1}H NMR (500 MHz, $CDCl_3$) spectra confirming the chemical structure and purity of nocardioazine A.



Appendix F. ^{1}H NMR (500 MHz, $CDCl_3$) spectra confirming the chemical structure and purity of nocardioazine B.

References

Alqahtani, N., Porwal, S., James, E., Bis, D. M., Karty, J. A., Lane, A. L. & Viswanathan, R. (2015). Synergism between Genome Sequencing, Tandem Mass Spectrometry and Bio-Inspired Synthesis Reveals Insights into Nocardioazine B Biogenesis. *Org. Biomol. Chem.* In press.

Belin, P., Moutiez, M., Lautru, S., Seguin, J., Pernodet, J.-L., & Gondry, M. (2012). The nonribosomal synthesis of diketopiperazines in tRNA-dependent cyclodipeptide synthase pathways. *Natural Product Reports*, *29*(9), 961–79.

Besemer, J., Lomsadze, A., Borodovsky, M. (2001) GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions, *Nucleic Acids Res. 29*, 2607-2618.

Bierman M., Logan R., O'Brien K., Seno E., Rao R. and Schoner B. (1992). Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene. 116, 43 49*

Bis, D. M., Ban, Y. H., James, E. D., Alqahtani, N., Viswanathan, R., Lane, A. L. (2015). Characterization of the nocardiopsin biosynthetic gene cluster reveals similarities to and differences from the rapamycin and FK-506 pathways. *ChemBioChem.* 16: 990-997.

Bonnefond, L., Arai, T., Sakaguchi, Y., Suzuki, T., Ishitani, R., & Nureki, O. (2011). Structural basis for nonribosomal peptide synthesis by an aminoacyl-tRNA synthetase paralog. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(10), 3912–7.

Borthwick, A. D. (2012). 2,5-Diketopiperazines: synthesis, reactions, medicinal chemistry, and bioactive natural products. *Chem. Rev. 112*, 3641-3716.

Bull, A. T., & Stach, J. E. M. (2007). Marine actinobacteria: new opportunities for natural product search and discovery. *Trends in Microbiology*, *15*(11), 491–9.

Challis, G. L. (2008). Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology*, *154*(6), 1555–1569.

Dalisay, D. S., Williams, D. E., Wang, X. L., Centko, R., Chen, J., & Andersen, R. J. (2013). Marine Sediment-Derived Streptomyces Bacteria from British Columbia, Canada Are a Promising Microbiota Resource for the Discovery of Antimicrobial Natural Products. *PLoS ONE*, *8*(10), 1–14.

Dias, D. a., Urban, S., & Roessner, U. (2012). A Historical Overview of Natural Products in Drug Discovery. *Metabolites*, *2*(4), 303–336.

Eisner, T. (2003). Chemical ecology: can it survive without natural products chemistry? *Proceedings of the National Academy of Sciences of the United States of America*, *100 Suppl* (90002), 14517–8.

Fenical, W., & Jensen, P. R. (2006). Developing a new resource for drug discovery: marine actinomycete bacteria. *Nature Chemical Biology*, *2*(12), 666–673.

Giessen, T. W., Von Tesmar, A. M., & Marahiel, M. A. (2013). Insights into the generation of structural diversity in a tRNA-dependent pathway for highly modified bioactive cyclic dipeptides. *Chemistry & Biology*, *20*(6), 828–38.

Giessen, T. W., Von Tesmar, A. M., & Marahiel, M. A. (2013). A tRNA-dependent two-enzyme pathway for the generation of singly and doubly methylated ditryptophan 2,5-diketopiperazines. *Biochemistry*, *52*(24), 4274–83.

Giessen, T. W., & Marahiel, M. A. (2012). Ribosome-independent biosynthesis of biologically active peptides: Application of synthetic biology to generate structural diversity. *FEBS Letters*, *586*(15), 2065–2075.

Giessen, T., & Marahiel, M. (2014). The tRNA-Dependent Biosynthesis of Modified Cyclic Dipeptides. *International Journal of Molecular Sciences*, *15*(8), 14610–14631.

Gondry, M., Sauguet, L., Belin, P., Thai, R., Amouroux, R., Tellier, C., Tuphile, K., Jacquet, M., Braud, S., Courcon, M., Masson, C., Dubois, S., Lautru, S., Lecoq, A., Hashimoto, S., Genet, R. & Pernodet, J.-L. (2009). Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. Nature Chemical Biology, 5(6), 414–20.

Gulder, T. A M., & Moore, B. S. (2010). Chasing the Treasure of the Sea-Bacterial Marine Natural Products. *Marine Biotechnology*, *12*(3), 252–260.

Gust, B., Challis, G. L., Fowler, K., Kieser, T., & Chater, K. F. (2003). PCR-targeted Streptomyces gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(4), 1541–6.

Gust, B., Rourke, S. O., Bird, N., Kieser, T., & Chater, K. (2004). Recombineering in Streptomyces coelicolor, 1–22.

Green, M.R. & Sambrook, J. (2012) Molecular Cloning: A Laboratory Manual (Fourth Edition). *Cold Spring Harbor Laboratory Press.*

Komatsu, M., Komatsu, K., Koiwai, H., Yamada, Y., Kozone, I., Izumikawa, M., Hashimoto, J., Takagi, M., Omura, S., Shin-ya, K., Cane, D. E. & Ikeda, H. (2013). Engineered Streptomyces avermitilis host for heterologous expression of biosynthetic gene cluster for secondary metabolites. *ACS Synthetic Biology*, *2*(7), 384–96.

Lane, A. L., & Moore, B. S. (2012). A sea of bisynthesis: marine natural products meet the molecular age. *Natural Product Reports*, *28*(2), 411–428.

Lautru, S., Gondry, M., Genet, R., & Pernodet, J. L. (2002). The albonoursin gene Cluster of S noursei biosynthesis of diketopiperazine metabolites independent of nonribosomal peptide synthetases. *Chemistry & Biology*, *9*(12), 1355–64.

Manivasagan, P., Alam, M. S., Kang, K.-H., Kwak, M., & Kim, S.-K. (2015). Extracellular synthesis of gold bionanoparticles by Nocardiopsis sp. and evaluation of its antimicrobial, antioxidant and cytotoxic activities. *Bioprocess and Biosystems Engineering*.

Manivasagan, P., Kang, K.-H., Sivakumar, K., Li-Chan, E. C. Y., Oh, H.-M., & Kim, S.-K. (2014). Marine actinobacteria: an important source of bioactive natural products. *Environmental Toxicology and Pharmacology*, *38*(1), 172–88.

Manivasagan, P., Venkatesan, J., Sivakumar, K., & Kim, S.-K. (2013). Marine actinobacterial metabolites: current status and future perspectives. *Microbiological Research*, *168*(6), 311–32.

Manivasagan, P., Venkatesan, J., Sivakumar, K., & Kim, S.-K. (2013). Pharmaceutically active secondary metabolites of marine actinobacteria. *Microbiological Research*.

Moutiez, M., Seguin, J., Fonvielle, M., Belin, P., Jacques, I. B., Favry, E., Arthur, M. & Gondry, M. (2014). Specificity determinants for the two tRNA substrates of the cyclodipeptide synthase AlbC from Streptomyces noursei. *Nucleic Acids Research*, 1–12.

Moutiez, M., Schmitt, E., Seguin, J., Thai, R., Favry, E., Belin, P., Mechulam, Y. & Gondry, M. (2014). Unravelling the mechanism of non-ribosomal peptide synthesis by cyclodipeptide synthases. *Nature Communications*, *5*(May), 5141.

Netzker, T., Fischer, J., Weber, J., Mattern, D. J., König, C. C., Valiante, V., Schroeckh, V. & Brakhage, A. A. (2015). Microbial communication leading to the activation of silent fungal secondary metabolite gene clusters. *Frontiers in Microbiology*, *6*(April), 1–13.

Newman, D. J. and Cragg, G. M. (2012). Natural products as sources of new drugs over 30 years from 1981 to 2010. *J. Nat. Prod.* 75 (3), 311-335.

Ongley, S. E., Bian, X., Neilan, B. A, & Müller, R. (2013). Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Natural Product Reports*, *30*(8), 1121–38.

Ozaki, T., Nishiyama, M., & Kuzuyama, T. (2013). Novel tryptophan metabolism by a potential gene cluster that is widely distributed among actinomycetes. *The Journal of Biological Chemistry*, *288*(14), 9946–56.

Raju, R., Piggott, A.M., Conte, M., Tnimov, Z., Alexandrov, K. & Capon, R. J. (2010). Nocardiopsins: New FKBP12-Binding macrolide polyketides from an Australian marine-derived actinomycete, *Nocardiopsis* sp. *Chem. Eur. J.* 16, 3194-3200.

Raju, R., Piggott, A. M., Huang, X.-C. & Capon, R. J. (2011). Nocardioazines: a novel bridged diketopiperazine scaffold from a marine-derived bacterium inhibits P-glycoprotein. *Organic Letters*, *13*(10), 2770–3.

Raju, R., Piggott, A. M., Quezada, M. & Capon, R. J. (2013). Nocardiopsins C and D and nocardiopyrone A: new polyketides from an Australian marine-derived *Nocardiopsis* sp. *Tetrahedron*. 69(2), 692-698.

Sauguet, L., Moutiez, M., Li, Y., Belin, P., Seguin, J., Le Du, M.-H., Thai, R., Masson, C., Fonvielle, M., Pernodet, J.-L., Charbonnier, J.-B. & Gondry, M. (2011). Cyclodipeptide synthases, a family of class-I aminoacyl-tRNA synthetase-like enzymes involved in non-ribosomal peptide synthesis. *Nucleic Acids Research*, *39*(10), 4475–89.

Smanski M. J., Casper J., Peterson R. M., Yu Z., Rajski S. R. & Shen B. (2012) Expression of the platencin biosynthetic gene cluster in heterologous hosts yielding new platencin congeners. *J. Nat. Prod. 75*, 2158-2167.

Subramani, R., & Aalbersberg, W. (2012). Marine actinomycetes: an ongoing source of novel bioactive metabolites. *Microbiological Research*, *167*(10), 571–80.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., Lee, S. Y., Fischbach, M. A., Muller, R., Wohlleben, W., Breitling R., Takano, E. & Medema, M. H. (2015). antiSMASH 3.0--a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research*, 1–7.

Williams, P. G. (2009). Panning for chemical gold: marine bacteria as a source of new therapeutics. *Trends in Biotechnology*, *27*(1), 45–52.

Elle James

Education

University of North Florida, Jacksonville, Florida
Bachelors of Science Degree in Biology, April 2012 (3.4 GPA)

Masters of Science Degree in Biology, August 2013-Present (3.6 GPA)
Research: Biochemical characterization of two natural product producing biosynthetic pathways from a marine actinomycete. Focus on *in vivo* and *in vitro* characterization of biosynthetic enzymes.
Advisor: Dr. Amy Lane

Relative Work Experience

University of North Florida, Jacksonville, Florida
Graduate Teaching Assistant, Biology Department, 2014-present
General Biology 1 & General Biology 2 Lab

University of North Florida, Jacksonville, Florida
Chemistry Research Assistant, Summer 2011-2013
Research: Construction of a genomic DNA library of a marine actinomycete in order to target gene sequences responsible for the production of a novel anticancer natural product.
Mentor: Dr. Amy Lane

Skills

Laboratory: Proficient in molecular biology techniques. Sterile technique, PCR, protein and DNA electrophoresis, DNA & RNA extraction, DNA mutagenesis, cloning. Recombinant DNA methodologies, protein expression, isolation and purification and functional protein assay. Chemical extraction and isolation.

Experimental Equipment: UV Spectrophotometry, HPLC, FPLC.

Computer: Microsoft Office, Geneious. Bioinformatics: DNA and protein sequence analysis, primer design and recombinant construct design.

Personal: Strong multi-tasking, problem solving and conflict resolution skills. Works well independently or as part of a team. Good technical and organizational skills with attention to detail. Punctual, reliable, and driven.

Publications

James ED*,* Knuckley BA, Alqahtani N, Porwal S, Karty JA, Viswanathan R, Lane AL. Two distinct cyclodipeptide synthases from a marine actinomycete catalyze biosynthesis of the same diketopiperazine natural product. *ACS Synthetic Biology.* In preparation.

Alqahtani N, Porwal SK, **James ED**, Bis D, Karty JA, Lane AL, Viswanathan R. 2015. Synergism between Genome Sequencing, Tandem Mass Spectrometry and Bio-Inspired Synthesis Reveals Insights into Nocardioazine B Biogenesis. *Organic & Biomolecular Chemistry.* In press. **DOI:** 10.1039/C5OB00537J

Bis DM, Ban YH, **James ED**, Alqahtani N, Viswanathan R, Lane AL. 2015. Characterization of the Nocardiopsin Biosynthetic Gene Cluster Reveals Similarities to and Differences from the Rapamycin and FK-506 Pathways. ChemBioChem. 16: 990-997.

Scientific Presentations

James E & Lane AL. (5/2015). Invited Oral Presentation. Characterization of Natural Product Producing Cyclodipeptide Synthases from a Marine Actinomycete. General Meeting of the American Society for Microbiology. New Orleans, LA.

James E, Alqahtani N, Viswanathan R, Lane AL. (4/2015). Oral Presentation. Characterization of Natural Product Producing Cyclodipeptide Synthases from a Marine Actinomycete. Showcase of Osprey Advancements in Research and Scholarship. University of North Florida, Jacksonville, FL.

James E, Alqahtani N, Viswanathan R, Lane AL. (9/2014). Poster & Invited Oral Presentation. Characterization of Natural Product Producing Cyclodipeptide Synthases from a Marine Actinomycete. The Southeastern Microbiology Summit. Ponte Vedra Beach, FL.

*Awarded 2nd place for graduate student oral presentations.

James E, Alqahtani N, Viswanathan R, Lane AL. (8/2014). Poster. Cyclodipeptide Synthase-Derived Diketopiperazine Natural Products from a *Nocardiopsis* sp. American Society of Pharmacognosy Annual Meeting. Oxford, MS.

James E, Lane AL. (10/2013). Poster. The Characterization of Cyclodipeptide Synthases in a Marine Actinomycete, *Nocardiopsis* sp. Natural Sciences Poster Session. University of North Florida, Jacksonville, FL.

James E, Bis DM, Lane AL. (10/2011). Poster. Exploring the genetic basis for production of anticancer compounds from a marine bacterium, *Nocardiopsis* sp. Natural Sciences Poster Sesssion. University of North Florida, Jacksonville, FL.

Professional Affiliations

American Society for Microbiology (2015-)
Florida Branch of the American Society for Microbiology (2014-)
American Society of Pharmacognosy (2014-)

Awards and Funding

- 2015 Department of Biology Graduate Summer Research Award ($2000).
- 2015 UNF Graduate Student Travel Award for presenting at ASM National Conference ($500).
- 2014 UNF Biology travel grant to present at American Society of Pharmacognosy conference ($500).