INVESTIGATION OF THE AZINOMYCIN BIOSYNTHETIC PATHWAY

A Dissertation

by

LAUREN ASHLEY WASHBURN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Coran M. H. Watanabe
Committee Members,	David Barondeau
	Tadhg Begley
	Margaret Glasner
Head of Department,	Simon North

August 2020

Major Subject: Chemistry

ABSTRACT

The azinomycins are hybrid nonribosomal peptide and polyketide natural products isolated from *Streptomyces sahachiroi*. They are of interest due to their antitumor and broad antibiotic activity, which comes from the ability to form interstrand crosslinks with dsDNA through electrophilic epoxide and aziridine rings. Fermentation studies using $^{18}O_2$ identified oxygen atoms present in azinomycin installed through enzymatic oxidation. Four oxygens were identified, including the oxygen of the epoxide moiety responsible for azinomycin's activity.

Due to a large portion of azinomycin biosynthesis occurring by nonribosomal peptide synthetase and polyketide synthase enzymes, a probe was developed to directly trap and characterize thioester-bound intermediates commonly found in these systems. The method was developed with the PKS AziB and validated the previous result of 2-methylbenzoic acid production by AziB. The probe was then used to identify intermediates bound to the NRPS AziA3 isolated from the azinomycin producing organism. These studies determined AziA3 bound the final epoxide moiety, with only condensation with the naphthoate and azabicycle to be performed.

The naphthoate moiety plays a key role in DNA-azinomycin adduct formation. To further study AziB's role in naphthoate formation, domain separation was used to increase protein expression, yield, and efficiency of post-translational modification. The separated domains were still active when incubated together. Additionally, AziG's role

was validated *in vivo* and the Δ*aziG* strain was used to probe the ability to produce benzoic-azinomycin derivatives.

The azabicycle moiety is the most complex of the four azinomycin moieties. Previous studies showed glutamic acid is the precursor to this moiety, but it does not contain the two carbons necessary for aziridine ring formation. AziC5 and AziC6 were identified as domains of a thiamin dependent transketolase responsible for installing the final two carbons. This result establishes the final carbon framework of the azabicycle and azinomycin. Additionally, gene disruption and complementation studies were performed with *aziW* to determine azinomycin production dependence on the amino-acyl carrier protein.

These studies will help guide future *in vitro* reconstitutions, gene disruptions, and synthetic biology efforts to produce azinomycin and azinomycin analogues that can be used in further cancer therapy development studies.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Watanabe, and my committee members, Dr. Barondeau, Dr. Begley, and Dr. Glasner, for their guidance and support throughout the course of this research.

I am grateful to my past and present labmates, classmates, and the Texas A&M University Chemistry Department staff. I would like to thank Dr. Keshav Nepal for teaching me all things *Streptomyces* in my first years, Dr. Shogo Mori for teaching me the fundamentals of PKS systems, and Dr. Vishruth Gowda for the many helpful discussions, science or otherwise, through my middle years of graduate work. To my current labmates, Brendan, Brett, Jean and Hieu, I wish you luck on your future endeavors. I would like to acknowledge Dr. Yohannes Rezenom for the continual guidance regarding mass spectrometry and Dr. Sumedh Joshi in the lab of Dr. Begley for help purifying many natural products.

I would like to thank my family. To my parents and grandparents, thank you for your encouragement and support. To Catherine, Lauren, and Theresa, thank you for the group texts and weekend trips to keep me in touch with the outside world. Finally, thank you to my wife, Elizabeth. I couldn't have done this without you, and I can't wait for our next chapter.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Coran M. H. Watanabe (advisor) and Professors David Barondeau and Tadhg Begley of the Department of Chemistry and Professor Margaret Glasner of the Department of Biochemistry and Biophysics.

The MS/MS data for Chapter II was gathered by Dr. Yohannes Rezenom. Undergraduate Researcher Katie Glaeser contributed to Chapter IV through cloning of plasmids. The Native Mass Spectrometry data for Chapter V was gathered and analyzed by Dr. Chris Boone and Dr. Arthur Laganowsky. The following Undergraduate Researchers contributed to Chapter V through cloning and protein expression: Kenneth Williams, Kendall Pryor, Emily Rimes, Alyssa Hippchen, and Sorin Miller. Undergraduate Researcher Alex Kuck contributed to Chapter VII through assisting with fermentation experiments. All other work conducted for the dissertation was completed by the student.

Funding Sources

This work was also made possible in part by the National Science Foundation under Grant Number 1904954. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Science Foundation.

NOMENCLATURE

NRPS	Nonribosomal Peptide Synthetase
NRP	Nonribosomal Peptide
PKS	Polyketide Synthase
PK	Polyketide
NMR	Nuclear Magnetic Resonance
TLC	Thin Layer Chromatography
DTT	Dithiothreitol
IPTG	Isopropyl β-d-1-thiogalactopyranoside
PPTase	Phosphopantetheinyl Transferase
Ppant	Phosphopantetheine
ACP	Acyl Carrier Protein
PCP	Peptide Carrier Protein
CoA	Coenzyme A
NADPH	Dihydronicotinamide-adenine Dinucleotide Phosphate
TCEP	Tris(2-carboxyethyl)phosphine Hydrochloride
ATP	Adenosine Triphosphate
AMP	Adenosine Monophosphate
PPi	Pyrophosphate
LC-MS	Liquid Chromatography – Mass Spectrometry
HPLC	High Performance Liquid Chromatography

EIC	Extracted Ion Chromatogram
NFAS	Non-fatty Acid Synthase Associated
FAS	Fatty Acid Synthase Associated
F6P	Fructose-6-phosphate
X5P	Xylulose-5-phosphate
HP	Hydroxypyruvate

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF SCHEMES

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION: NATURAL PRODUCTS

Introduction

Nature has long been a source for the identification of therapeutic agents and the isolation of specialty compounds, such as dyes. Historically, plants have played a necessary role in human health and medicinal practices.[1, 2] Plant extracts have been used to treat malaria (quinine and artemisinin), cancer (paclitaxel) and pain (morphine).[3-5] These bioactive compounds are produced by living organisms in response to their environment and are referred to as natural products or secondary metabolites. Secondary metabolites are normally not required for the growth and development of the producing organism. The metabolite is typically produced in a response to outside stress, such as a defense response to a predator. The production of secondary metabolites will give the producing organism an advantage over its surroundings.[6-7] Biological processes such as glycolysis, Krebs Cycle and photosynthesis produce the intermediates necessary for natural product biosynthesis.[3, 8] As these metabolic pathways have evolved to produce a bioactive molecule, they are a great source for identification of new drug leads.

Natural Products in Drug Discovery

The "Golden Age" of natural product commercialization was brought upon by advances in cultivation of microorganisms.[9-10] This age was kicked off by the discovery and isolation of penicillin from the fungus *Penicillium chrysogenum* in the 1940s.[3] The

following thirty years (1940s-1970s) led to the identification of thousands of natural products that exhibited antibacterial or antifungal activity, as inhibition of microbial growth was the easiest trait to screen for at that time. The largest source of natural products in this time period were actinobacteria, such as *Streptomyces*. [8-9, 11-12]

Several factors led to a decrease in the research and development of natural products drugs in the 1980s and 1990s.[3] The limitations of technology required for high quality microbial libraries, production of useful amounts of each natural product, and synthesis of natural products have been implicated in the decrease in this time period.[11, 13] Coincidentally, a 20-year low in drug discovery was reached during this time period.[14] The shift to combinatorial libraries and high-throughput screening in the 1990s focused on the development and identification of new chemical entities with novel carbon skeletons. The limitations of this approach are apparent as there is only one FDA approved drug containing a new chemical entity or carbon skeleton (1981-2014).[3, 12] It became clear that the structural and chemical diversity secondary metabolites provide was necessary for the development of drug leads. In the late 1990's there was a shift to diversity-oriented synthesis to develop synthetic drugs and natural product mimics. This approach used synthetic chemistry to produce compounds resembling natural product scaffolds.[3, 12] Since this time period, 51% of approved drugs were derived from natural products or natural product mimics (Figure 1). Natural product leads account for 65% of new cancer drugs and 46% of infectious disease drugs (antibacterial, fungal, parasitic, viral).[15]

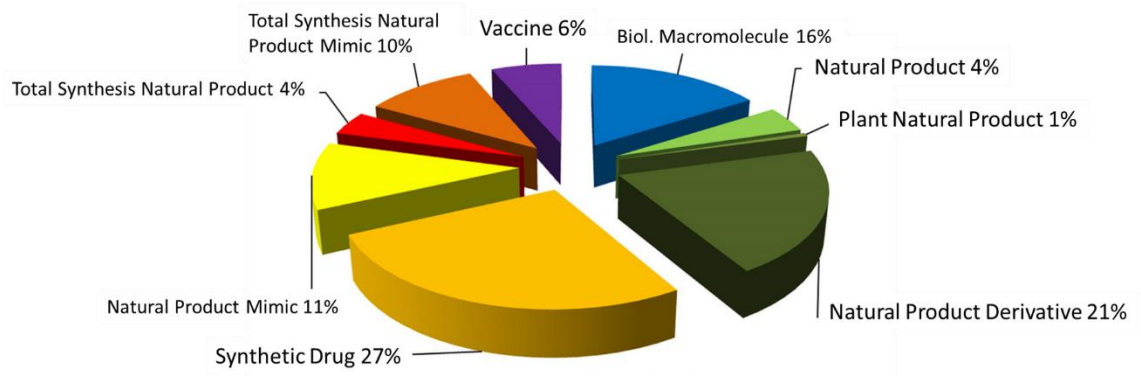


Figure 1 Percentage of Sources for Clinical Drugs[15]
Reprinted from Newman, 2016.

The last ten years has seen a new natural products boom with the advances in genome sequencing, bioinformatics, and high-throughput mass spec technology.[3, 9] These advances have identified cryptic secondary metabolite gene clusters coding for novel secondary metabolites, ushering in a second "Golden Age" of drug development from natural product leads. Genome sequencing has revealed the genomes of actinomycetes, such as *Streptomyces*, contain 20 to 50 secondary metabolite gene clusters each.[16-18] While it is estimated that less than 10% of these gene clusters are naturally expressed in large enough amounts to characterize the natural product from fermentation analysis, recombinant technology for cloning, expression, and activation of *Streptomyces* cryptic gene clusters has rapidly developed.[19-24]

Therapeutic *Streptomyces* Natural Products

Streptomyces are gram-positive soil bacteria known for their high G+C DNA content, the ability to form filamentous mycelia and spores through their life cycle, and their ability to produce secondary metabolites.[25-27] The first antibiotic discovered from a

Streptomyces culture was from *Streptomyces antibioticus* in 1940.[28] Since this time, *Streptomyces* have produced more than 80% of antibiotics from actinomycetes. More recently, a metabolite isolated from *Streptomyces spectabilis* showed activity against methicillin-resistant *Staphylococcus aureus* (MRSA).[29] Additionally, *Streptomyces* produce herbicides, immunosuppressants, and antitumor drugs.[30-32]

Many of the first antitumor drugs were first identified as antibiotic compounds. Actinomycin, a polyketide from *Streptomyces parvulus*, was the first antibiotic to show antitumor activity in 1940.[28, 33] Actinomycin functions by intercalation into DNA, preventing transcription from occurring.[34] A common mechanism of action of antitumor drugs is interaction with DNA, through intercalation, alkylation, or formation of interstrand crosslinks. While intercalation and monoadduct formation can lead to cell death through disrupting or slowing the transcription and replication processes, interstrand crosslinks can completely stop replication.[35-36] The interstrand crosslinking drug mitomycin was first isolated in 1956 from *Streptomyces caespitosis*.[37] Mitomycin forms interstrand crosslinks at very specific DNA sequences, making an interstrand crosslink at the rate of 1 per 20,000 bases on average.[38-40] Its activity stems from the presence of a highly reactive aziridine ring (Figure 2). It was demonstrated that feeding derivatives of precursors found in the mitomycin biosynthetic pathway can produce clinically relevant analogues of mitomycin directly in the cell culture.[41]

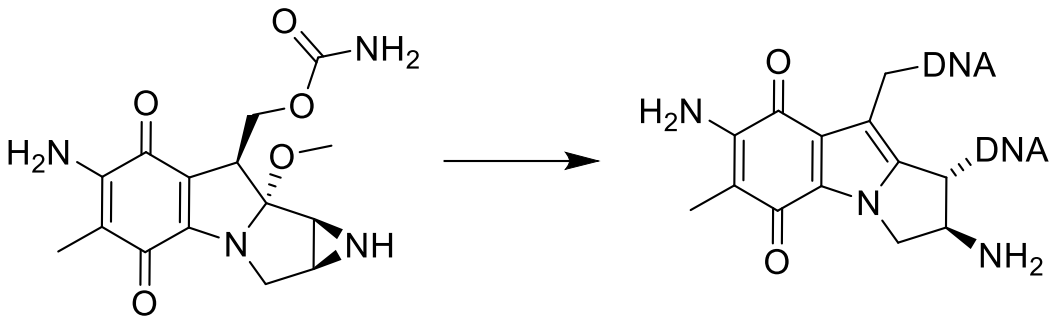


Figure 2 Mitomycin C DNA Interstrand Crosslink

Secondary metabolites can be classified not only by their activity, but also by their biosynthetic origins. There are four main classes of natural products: terpenoids, alkaloids, polyketides, and nonribosomal peptides (Figure 3). Terpenes are derived from

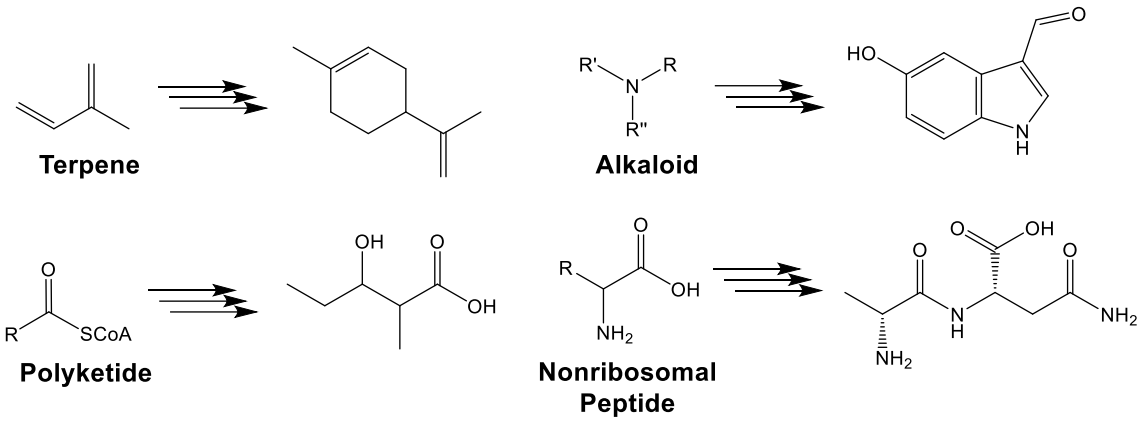


Figure 3 Natural Product Classes

isoprene starter units, alkaloids from nitrogen-based units, polyketides from acetyl and malonyl units, and nonribosomal peptides from amino and aryl acids. *Streptomyces* are capable of producing natural products encompassing the four classes and hybrids of two

or more classes.[42-43] A large portion of secondary metabolite gene clusters found in *Streptomyces* contain genes coding for polyketides (PK), nonribosomal peptides (NRP), and hybrids of the two.[44-45] Figure 4 includes examples of bioactive *Streptomyces* NRP and PK secondary metabolites.

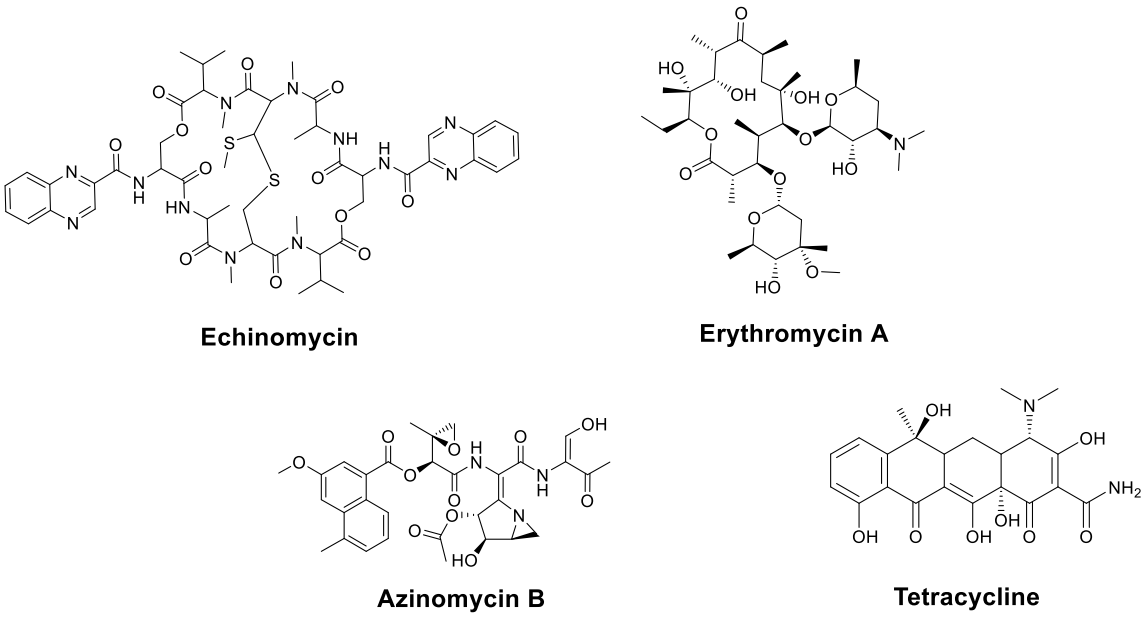


Figure 4 PKS and NRPS Streptomyces Natural Products
NRPS: Echinomycin. PKS: Tetracycline and Erythromycin A. Hybrid: Azinomycin B

Biosynthetic Principles of NRPS and PKS Systems

Nonribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) enzymes are large multidomain proteins that allow for great diversity to be installed into natural product structures. Figure 5 includes a list of possible domains for PKS and NRPS proteins.[46] The great structural diversity of polyketide and nonribosomal peptide natural products comes from the ability of these proteins to form multi-protein assembly

lines. These assembly lines contain proteins that by themselves contain several domains and the ability to functionalize or condense substrates. A common feature is the presence of a thiolation domain (T). In PKS, this domain is referred to as the Acyl Carrier Protein (ACP) and in NRPS it is the Peptide Carrier Protein (PCP). The names also point towards difference in substrates for the two families. NRPS proteins catalyze reactions with amino acid substrates and PKS with acyl-Coenzyme A substrates.[44, 46-47] The thiolation domain is of importance as this domain makes a thioester bond to the

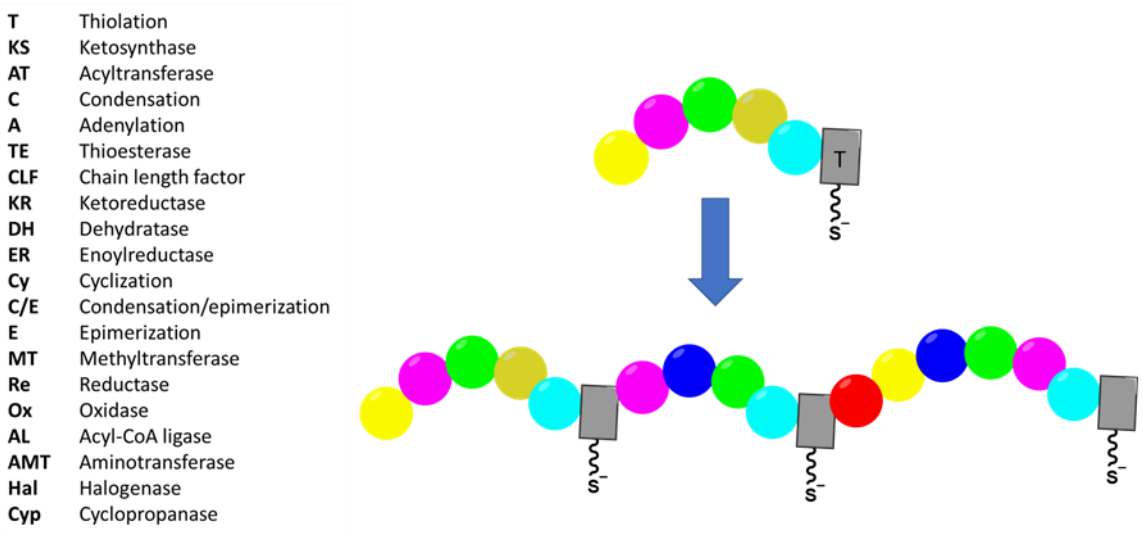


Figure 5 PKS and NRPS Domains and Assembly Line
List of possible NRPS/PKS domains. Representation of how individual proteins form assembly line. Each circle represents a different possible domain.

substrate, allowing the substrate to be carried through the successive catalytic cycle. The formation of the thioester bond requires a post-translational modification to be performed by a phosphopantetheinyl transferase (PPTase).[48] The PPTase catalyzes the addition of the phosphopantetheine arm to a serine on the thiolation domain (Figure 6).

Once the modification has been performed, the enzymatic reaction can begin. While there are several differences in activity between the two protein families, the product release mechanism involving a thioesterase is very similar between PKS and NRPS systems.[49]

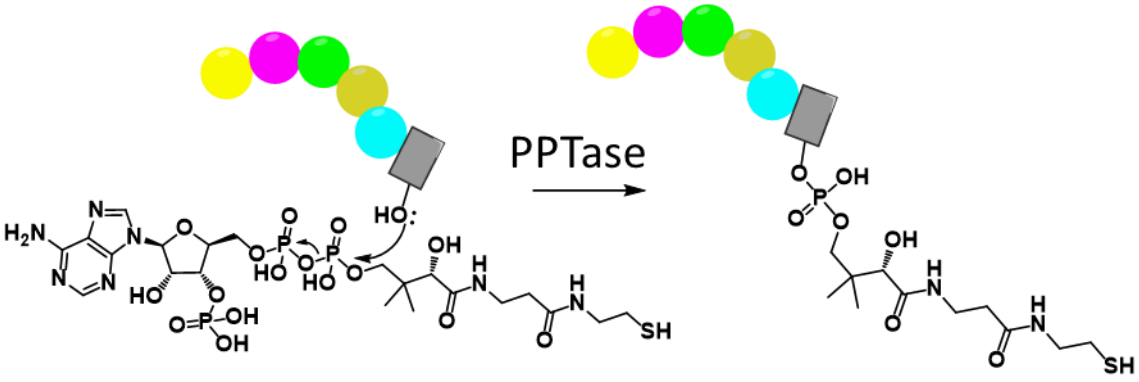


Figure 6 Post-translational Modification of PKS/NRPS

Nonribosomal Peptide Synthetase

Nonribosomal peptide synthetases are typically non-iterative multidomain modular systems, with each module consisting of an adenylation domain, a peptide carrier domain, and catalytic domains such as condensation and ketoreduction.[44, 47, 49] The substrates for NRPS proteins are largely amino acids or amino acid derivatives. NRPS require substrate activation by the adenylation domain. This domain uses ATP to activate the substrate by the addition of AMP (Figure 7). The activated substrate can then be loaded onto the ppant arm of the peptide carrier protein (PCP) domain. This flexible arm can then shuttle the substrate between the catalytic domains for

functionalization or condensation with another amino acid substrate. The condensation domain will form an amide or ester bond between the substrates. In addition to

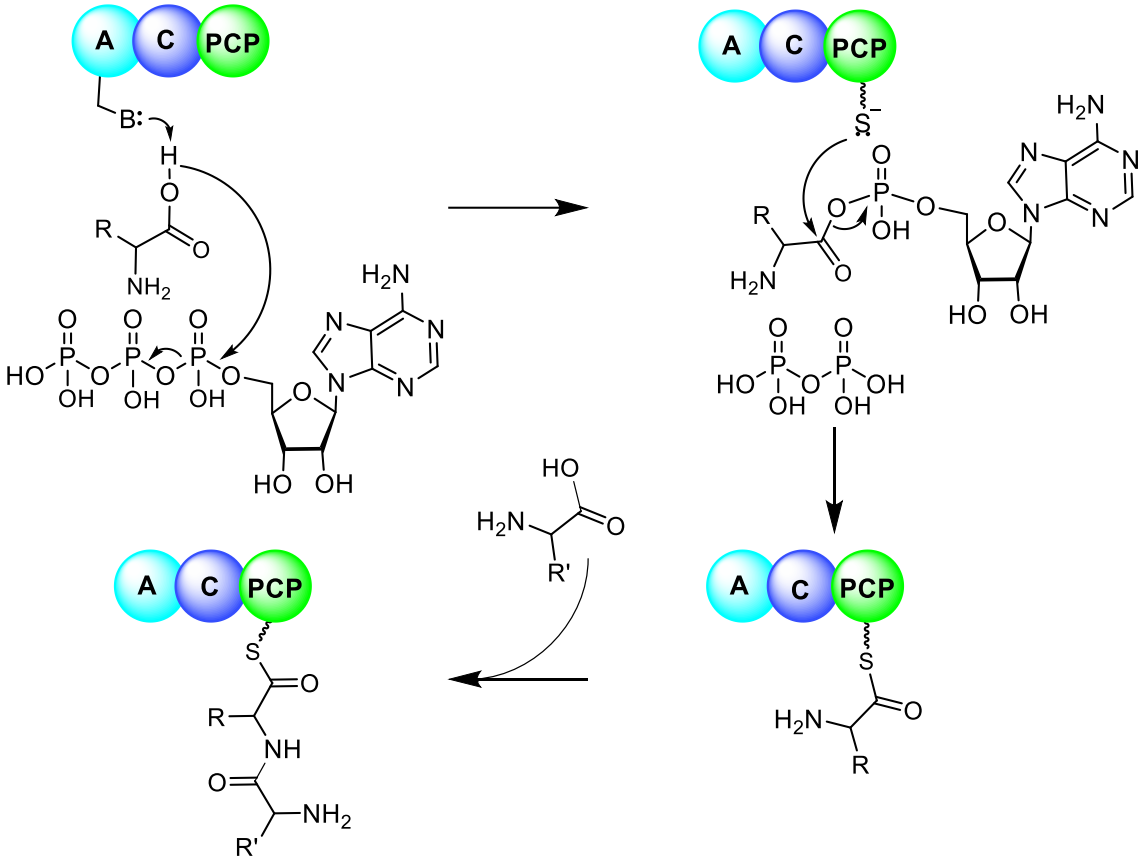


Figure 7 General NRPS Mechanism
A: Adenylation Domain. C: Condensation Domain. PCP: Peptide Carrier Protein Domain

tailoring reactions from domains within the protein (*cis*-acting domains), *trans*-acting tailoring domains can perform reactions on the ppant bound substrate. That is, catalytic proteins that are not on the same peptide chain as the NRPS can interact with the NRPS module and perform reactions on the thioester bound natural product. [50-51]

Polyketide Synthase

Polyketide synthases vary from NRPS in that they can be iterative or non-iterative. Iterative PKS can produce the final polyketide product from one module whereas non-iterative will consist of multiple modules to generate the product. PKS substrates are acyl-Coenzyme A units, common starter units are acetyl-CoA, malonyl-CoA, and methylmalonyl-CoA. The essential PKS domains are the ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP).[45-46, 49, 52-53] The other domains will be tailoring domains and will vary by system. The AT domain selects for the substrate that will be loaded onto the ACP ppant arm (Figure 8). The starter unit is then transferred to an active site cysteine in the KS domain and a second substrate loaded onto the ACP arm. Condensation of the two units will take place resulting in loss of carbon dioxide and formation of a new carbon-carbon bond between the two substrates. This process can continue, or tailoring reactions can be performed to functionalize the product. Figure 8 shows the reactions that can take place to go from ketone to alkane chain catalyzed by KR, DH, and ER domains. Similarly, to NRPS systems, there can be *trans*-acting tailoring domains in PKS pathways.[54-55]

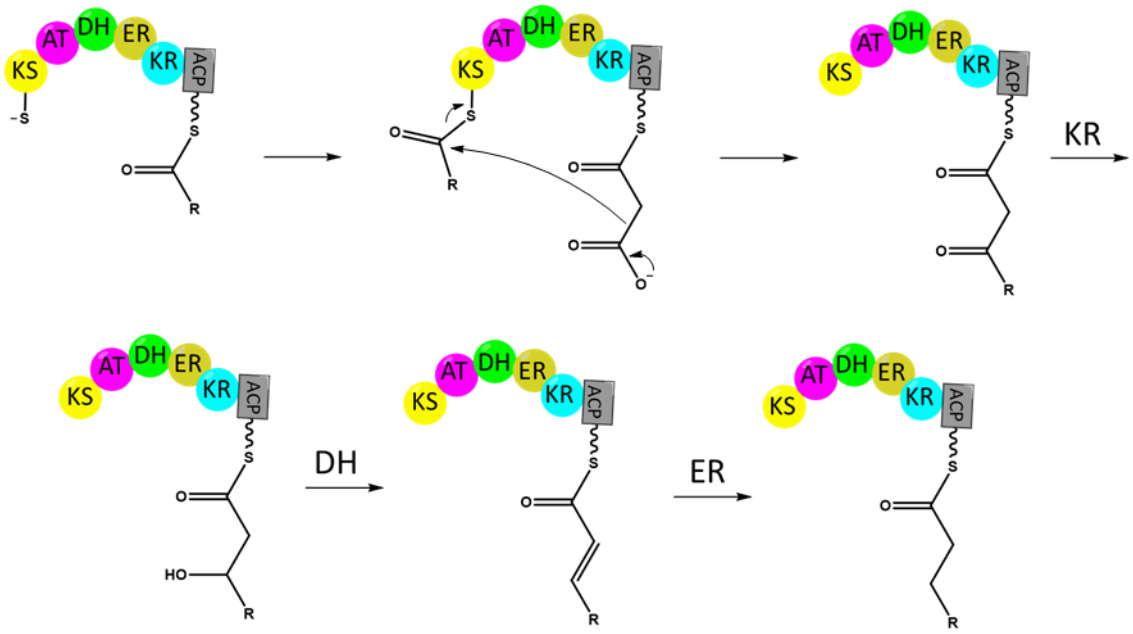


Figure 8 General PKS Mechanism
KS: Ketosynthase. AT: Acyl Transferase. DH: Dehydratase. ER: Enoyl Reductase. KR: Ketoreductase. ACP: Acyl Carrier Protein.

Thioesterase

Once the catalytic cycle of the NRPS or PKS is completed, product release must be initiated to break the thioester bond. This will generate the free natural product and the protein for another enzymatic cycle. A thioesterase (TE) will catalyze product release by either hydrolysis of the thioester bond or cyclization of the product in cases where a cyclic natural product is formed.[49, 56-57] The proposed mechanism of hydrolysis of the thioester bond includes product transfer to an active site serine of the thioesterase (Figure 9). This new ester intermediate can then be hydrolyzed by water to form the free carboxylic acid product.[58] Thioesterases can be *cis* or *trans* acting. Type I Thioesterases are *cis*-acting and will be a part of the peptide chain containing the rest of the

PKS/NRPS protein. Type II Thioesterases are *trans*-acting as depicted in Figure 9.[56] Thioesterases can also work to regenerate the free thiol of the ppant arm if the wrong substrate was loaded.[59]

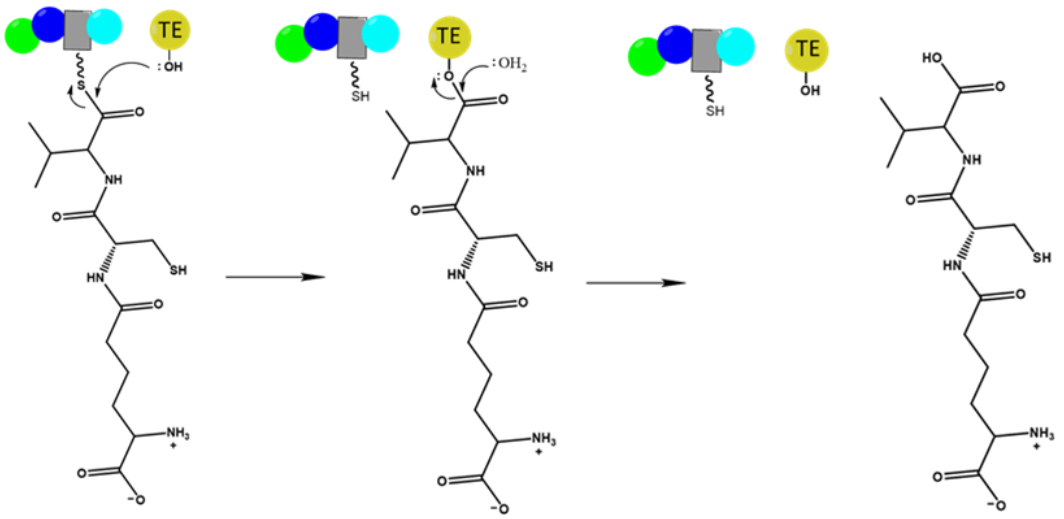


Figure 9 Thioesterase Product Release

Recent Advances in Secondary Metabolite Production

Improvements in sequencing, cloning, and fermentation tools has opened the door for the use of *Streptomyces* and other microbes in the overexpression or heterologous expression of secondary metabolite gene clusters. *Streptomyces* make good hosts for heterologous expression because they already contain adequate substrates and genes from the production of their own natural products. The heterologous expression of the PKS gene cluster for the antitumor agent isomigrastatin in *Streptomyces albus* resulted in a titer of 46 mg/L.[60]

Metabolic engineering and combinatorial biosynthesis are used to improve the yields and diversify the natural products produced. Metabolic engineering targets biosynthetic gene clusters and regulatory genes to improve the yield of a metabolite. For example, the production of the polyketide Nanchangmycin was increased by 8-fold as compared to wild type production by the disruption of the *nsdA* regulator gene.[26] The development of CRISPR tools for use in *Streptomyces* has greatly increased efficiency in this area.[61-63] While metabolic engineering can be used to increase yields, combinatorial biosynthesis is used to produce novel compounds. The multidomain and modular nature of NRPS and PKS systems allows for domains to be swapped between proteins (Figure 10). PKS and NRPS domain swapping has been used to successfully produce analogues

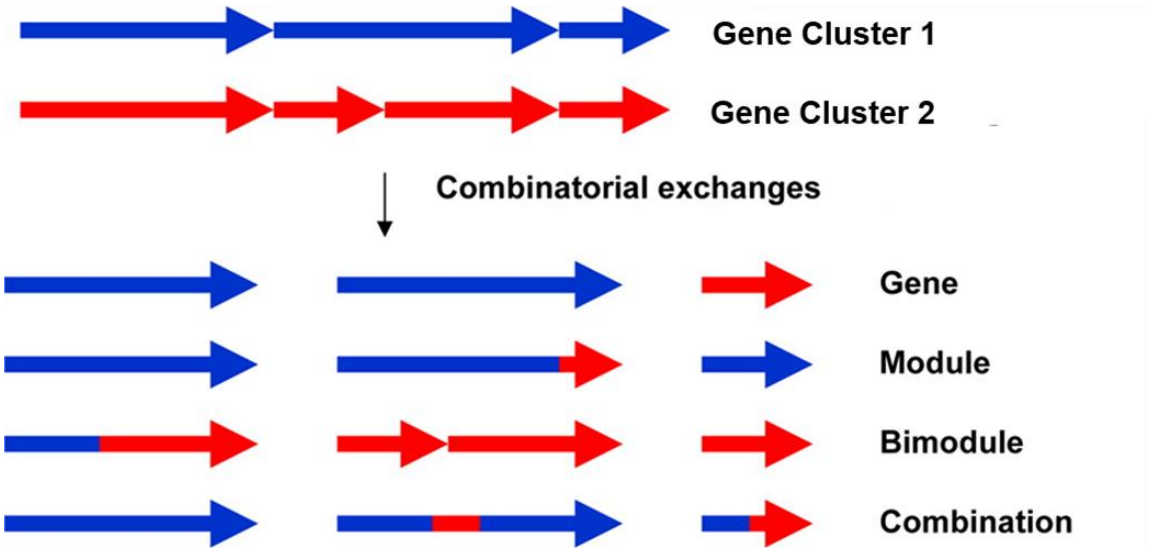


Figure 10 Combinatorial Exchange for Two Gene Clusters[65]
Reprinted from Baltz, 2014.

of erythromycin, novel lipopeptides, and biofuels.[52-53, 64] Merging metabolic engineering and combinatorial biosynthesis strategies leads to the production of large amounts of novel compounds. A key step in the ability to use domain swapping and engineering is understanding the enzymology and substrate acceptance of each protein or domain. The development of more effective therapies will require the use of fundamental biochemistry combined with cutting edge synthetic biology tools.

Azinomycins

Background

The azinomycins (Figure 11) are antitumor natural products produced by the gram-positive soil bacterium *Streptomyces sahachiroi*. Azinomycin B, originally denoted as carzinophilin A, was isolated from *S. sahachiroi* in 1954. It exhibited antibacterial activity against gram-negative bacteria and showed antitumor activity against sarcoma cells.[66-67] In 1986, azinomycin B was re-isolated from *S. griseofuscus*, along with azinomycin A. It was during this time that the structure of the azinomycins was determined (Figure 11).[68-69] The azinomycins are nearly structurally identical, with the only difference being azinomycin B contains an enol substituent. The structure suggests a combination of PKS, NRPS, and alkaloid origins.

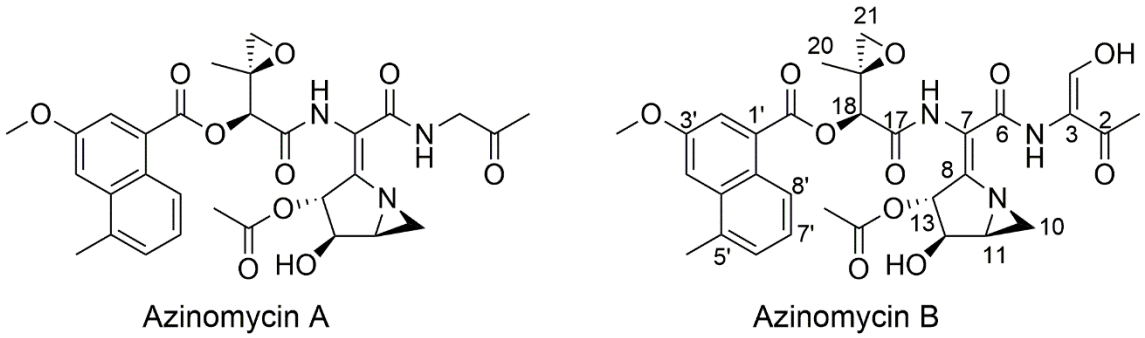


Figure 11 Structure of the Azinomycins

Additional testing after re-isolation in 1986 showed broad range anti-tumor activity against *in vitro* cell lines and *in vivo* murine tumors. The azinomycins were tested *in vitro* against L5178Y cells (mouse lymphoma cell line), azinomycin A exhibited an IC_{50} of 0.07 µg/ml and azinomycin B an IC_{50} of 0.11 µg/ml. Azinomycin B is more effective than azinomycin A *in vivo*, showing activity against 36 cases of malignant neoplasms, as well as an increased life span (ILS) of 193 % at 32 µg/kg against P388 murine transplantable tumors. This dosage is much lower than the commonly used interstrand crosslinking drug mitomycin C which showed 204 % ILS at 1 mg/kg. Additionally, this study showed azinomycin antibiotic activity against gram positive and negative bacteria, but little activity against yeast and fungi.[70]

Azinomycin Mode of Action

The azinomycins induce cell death through the formation of interstrand crosslinks with double stranded DNA. Covalent linkages are formed between DNA bases and the electrophilic epoxide and aziridine rings, and the naphthoate moiety

provides non-covalent interactions with the non-polar regions of DNA (Figure 12).[71] Azinomycin B's sequence selectivity was initially established using ^{32}P labeled synthetic oligonucleotides. It was determined the linkages are made between G and R (A or G) two base pairs downstream on the complementary strand. The covalent bond to guanine occurs at the N7 position.[72] Additional studies using 15-mer dsDNA containing a varied triplet (purine-pyridine-pyridine) sequence showed GCC is the optimal sequence for

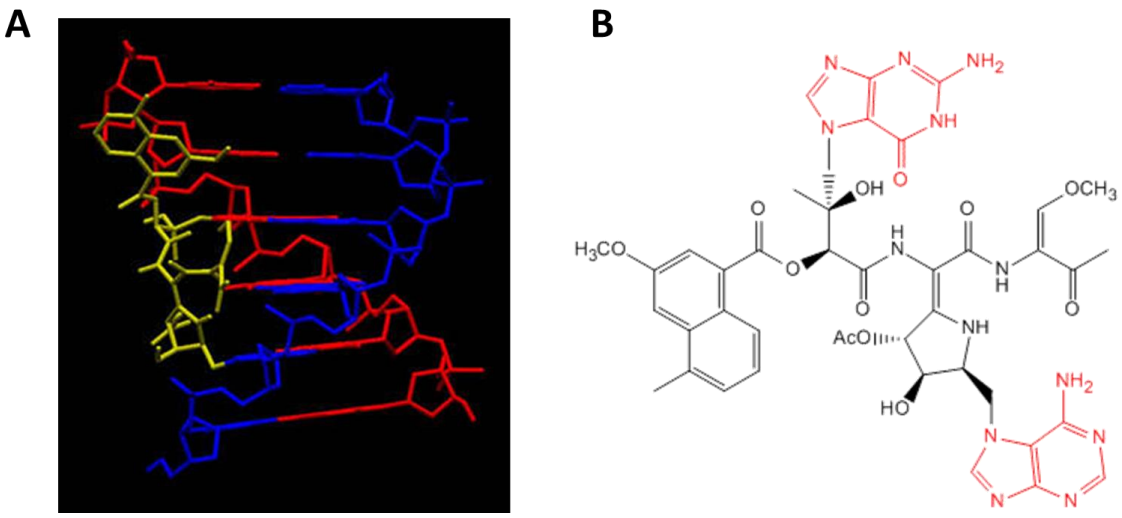


Figure 12 DNA Binding Activity of Azinomycin
A. Azinomycin binding in major groove of DNA. Yellow = azinomycin. Red and Blue = DNA strands. **B**. Covalent binding of Azinomycin B to DNA Bases on separate strands through electrophilic epoxide and aziridine rings.

azinomycin B binding.[71] It was also determined that the aziridine-DNA adduct forms first before the epoxide-DNA adduct.[73] In addition, the naphthoate epoxyamide, or "Left-Half" of azinomycin can form DNA monoadducts resulting in cytotoxic activity. Analogs that contained the epoxide moiety, but not the naphthoate, did not show

activity. This confirms that while the naphthoate does not covalently bind DNA, its non-covalent interactions play a role in adduct formation.[74]

Biosynthesis

The azinomycins are hybrid nonribosomal peptide and polyketide natural products, with alkaloid origins for the aziridine containing moiety. The naphthoate moiety has polyketide origins and the amide and ester backbone are assembled by NRPS modules. To confirm the biosynthetic origins of the individual moieties, isotopic labeling and feeding studies were performed. ^{13}C labeled precursors were fed to *S. sahachiroi* during azinomycin production. Isolated azinomycin was then analyzed by ^{13}C NMR to observe incorporation of the labeled precursor. ^{13}C-acetate was first used to identify PK origins of the naphthoate moiety. Acetate was incorporated in a sequential pattern which is consistent with PKS biosynthesis. In addition to the naphthoate moiety, acetate was incorporated into the keto-enol unit of azinomycin B and the azabicycle (aziridino[1,2-a]pyrrolidine) moiety. It was proposed that the keto-enol incorporation comes from the conversion of acetate to threonine through oxaloacetate and the azabicycle through ketoglutarate derived amino acids.[75]

The incorporation of acetate into the keto-enol and azabicycle fragments guided isotope feeding studies with these two moieties. It was confirmed that the keto-enol originates from threonine.[76] Proposed advanced intermediates were not accepted suggesting the NRPS responsible for condensation of the keto-enol moiety accepts threonine as its substrate and subsequent tailoring reactions occur after loading. The

fourth moiety of azinomycin A was identified as deriving from aminoacetone.[77] Feeding studies identified glutamic acid as the precursor to the azabicycle moiety.

Based on structural similarity, valine and valine analogs were ^{13}C labeled and fed to *S. sahachiroi* to monitor for incorporation into the epoxide moiety. Successive valine analogs did show incorporation into azinomycin suggesting valine is derivatized before being condensed to the naphthoate and azabicycle moieties through ester and amide bonds. The final analog incorporated was the alkene keto-acid (3-methyl-2-oxobut-3-enoic acid).[78]

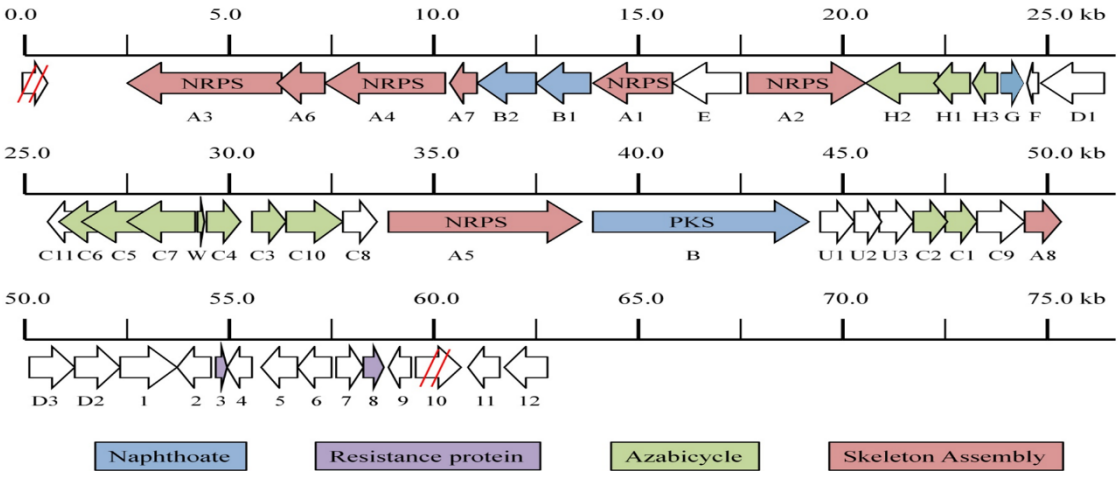


Figure 13 Azinomycin Biosynthetic Gene Cluster
Genes with unknown function are in white.

The identification and sequencing of the azinomycin gene cluster in 2008 allowed for enzymatic studies to be pursued. The gene cluster contains one PKS, five NRPS, and four TE genes (Figure 13). The remaining code for tailoring enzymes, resistance and transport proteins, and a fair amount are genes coding for proteins of

unknown function.[79] Gene disruption and *in vitro* reconstitution have proven the PKS AziB is responsible for naphthoate moiety production. AziB1 and B2 are responsible for methoxy installation on the naphthoate.[79-80] Based on homologous sequences Figure 13 has been color coded with proposed function or moiety involvement. Coupling together the ^{13}C-feeding results, gene disruptions, and *in vitro* reconstitutions, the current biosynthetic pathway to the azinomycins is found in Figure 14.

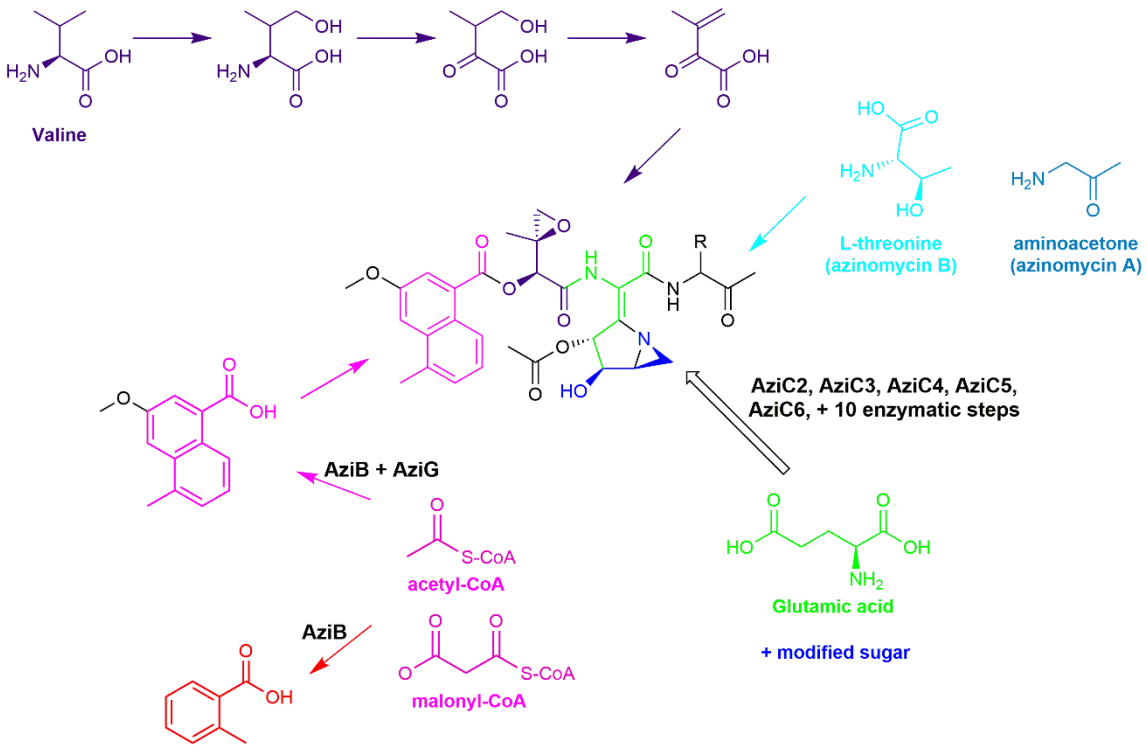


Figure 14 Biosynthetic Origins of the Azinomycins

Statement of Purpose

Secondary metabolites have been of great importance in the development of pharmaceuticals. More recently the advances in sequencing and synthetic biology

methods have led to a boom in combinatorial biosynthesis and the design of new pathways for production of known natural product and the engineering of non-natural metabolites. In order to engineer new pathways, understanding protein function is necessary. In the following chapters, characterization of key azinomycin biosynthetic proteins will be discussed along with the development of a NRPS/PKS thioester intermediate capture strategy. Both will help develop azinomycin analogues and will be useful in engineering new pathways and identifying new natural products.

CHAPTER II

MOLECULAR OXYGEN INCORPORATION INTO AZINOMYCIN*

Introduction

The azinomycins are antitumor natural products produced by *Streptomyces sahachiroi*. As outlined in Chapter I, ^{13}C feeding studies were performed to identify the precursors involved in the biosynthetic pathway of the azinomycins. While the ^{13}C feeding studies were highly effective at establishing the carbon framework of the azinomycins, the origin of the oxygen atoms was still unknown. Identifying the origins of oxygen atoms provides a map of oxidative transformations and places mechanistic constraints on the biosynthesis of the azinomycins. The oxygens could originate from the building blocks identified in the ^{13}C feeding studies, H_2O, or molecular oxygen (O_2) as incorporated through enzymatic oxidation. To identify incorporation of molecular oxygen, $^{18}O_2$ can be used and incorporation tracked by observing an upfield shift in the ^{13}C NMR signal of the neighboring carbon atoms.[81-82] Producing azinomycin in a 50:50 atmosphere of $^{18}O_2$: $^{16}O_2$ allows for a mixture of labeled and unlabeled azinomycin to be present. Having a portion of azinomycin unlabeled will serve as an internal standard, where there will be a standard ^{13}C signal next to the shifted heavy oxygen -^{13}C signal. An internal standard is necessary as the shift resulting from heavy oxygen is very small, ranging from approximately 0.010 to 0.035 p.p.m. for a single bonded oxygen and 0.30

to 0.55 p.p.m. for double bonded oxygens.[81] In addition to NMR studies, incorporation of heavy oxygen can be tracked using mass spectrometry.

Preliminary Research

An oxygen atmosphere recirculation system was used to maintain a 50:50 atmosphere of $^{18}O_2$: $^{16}O_2$. The closed system is required to avoid diffusion exchange with air and to retain expensive $^{18}O_2$. The recirculation design requires 5 M KOH to scrub carbon dioxide and traps to catch over-bubbling of the KOH to protect the culture (Figure 15). $^{18}O_2$ was introduced after 18 h of fermentation to allow for production of azinomycin to begin. Exposure of *S. sahachiroi* to $^{18}O_2$ for 57 h produced 35.9 mg of purified azinomycin B for evaluation (Thom Kelly).

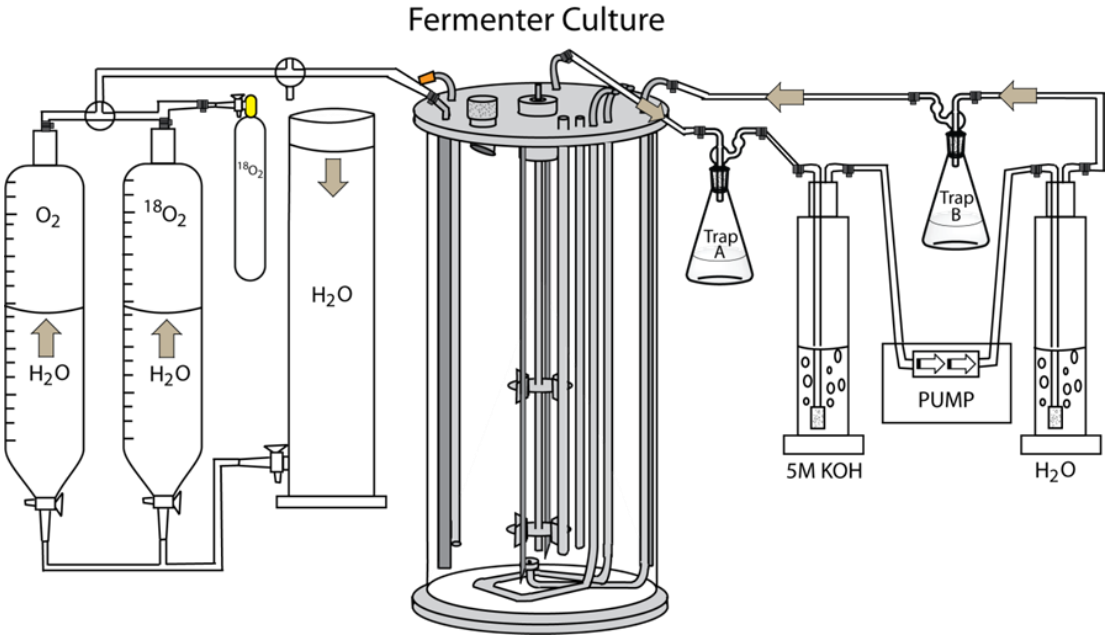


Figure 15 Fermenter Culture Oxygen Recirculation System
Thom Kelly

Results and Discussion

The ^{13}C NMR (Figure 16, Table 1) and Mass Spec (Appendix Figure 1-7) determined the site of four heavy oxygen atoms. The ^{13}C NMR shifts are congruent with the previously reported shifts for single bonded oxygen atoms.

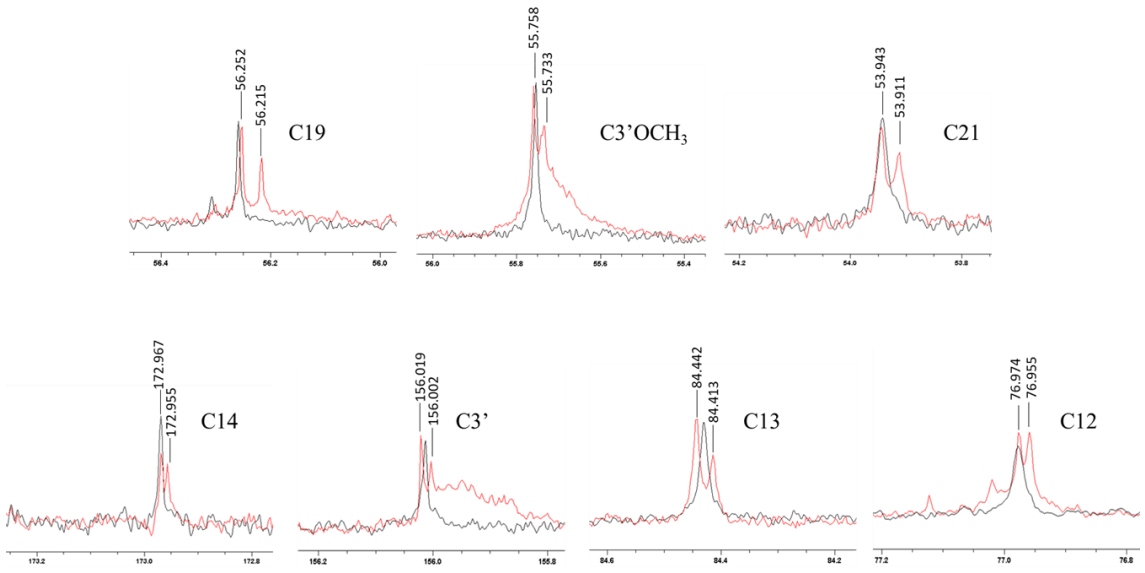


Figure 16 ^{13}C Signals Displaying a Shift in $^{18}O_2$ Incorporation
Shifts were observed in carbons 12, 13, 14, 19, 21, 3' and the 3' methoxy. Black trace = natural. Red trace = $^{18}O_2$. To observe the C12 shift CD_2Cl_2 was used.

Table 1 ^{13}C Shifts from ^{18}O Incorporation in Azinomycin B

Carbon Position	δ/ppm ^{16}O	δ/ppm ^{18}O	δ Difference (ppm)
3'	156.019	156.002	0.017
3'OCH_3	55.758	55.733	0.025
21	53.943	53.911	0.032
19	56.252	56.215	0.037
13	84.442	84.413	0.029
14	172.967	172.955	0.012
12	76.974	76.955	0.019

C-3' Position

It was previously established that the P450 oxygenase AziB1 is responsible for the hydroxylation of the naphthoate moiety at the 3' position (Figure 17).[83] This result supports the involvement of the P450 monooxygenase AziB1 in the formation of 3-methoxy-5-methylnaphthoic acid (ESI-MS/MS Appendix Figure 4, 5, and 7).

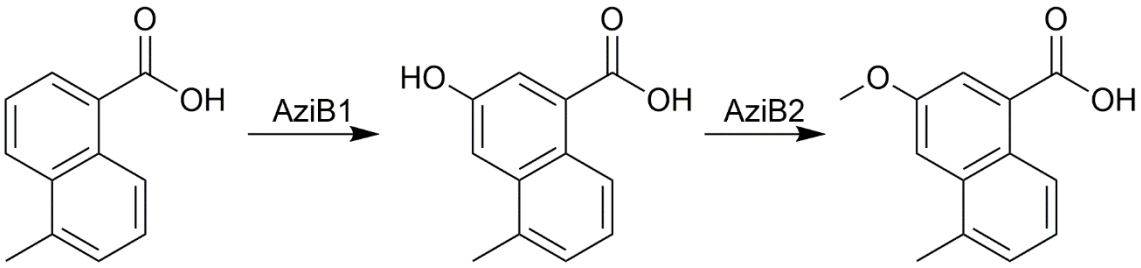


Figure 17 AziB1 Oxidation of 3' Position

C-19 and C-21 Positions

Incorporation of molecular oxygen is observed at the epoxide oxygen as seen by the shifts at C-19 and C-21 (Figure 16, Table 1) and ESI-MS/MS (Appendix Figure 5 and 7). The oxygenase and exact substrate for oxidation is unknown currently (Figure 18). The corresponding epoxy-keto acid was fed to *S. sahachiroi* with ^{13}C label and did not lead to incorporation. It was determined that the epoxy-keto acid was unstable in aqueous medium which limited its application in fermentation feeding studies.[78] Reconstitution of this pathway *in vitro* is underway to establish timing of epoxide formation and identification of the oxygenase responsible.

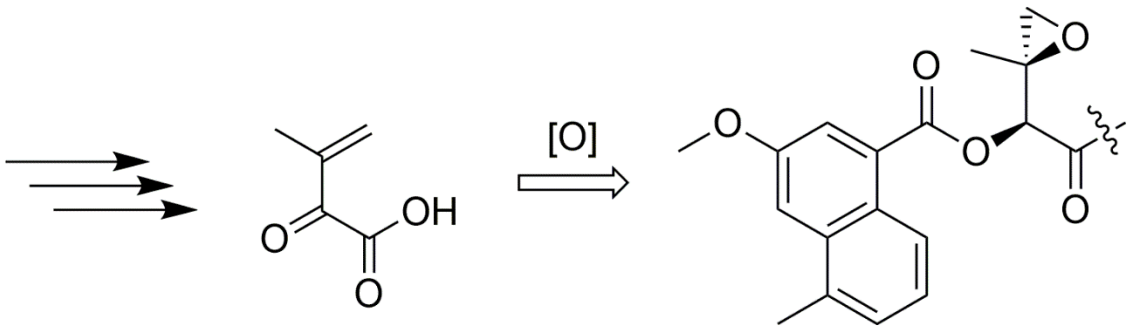


Figure 18 Oxidation at C-19 and C-21 Positions

C-12 and C-13/14 Positions

Heavy molecular oxygen was also observed at both the C-12 and C-13/14 positions of the azabicycle moiety (Figure 16, Table I and Appendix Figure 6-7). This is an interesting observation as there was already an unlabeled oxygen bonded to C-12 originating from the glutamic acid precursor (Figure 19). Feeding studies and *in vitro* reconstitution has established the first four steps of the biosynthetic pathway to the azabicycle. The presence of heavy molecular oxygen at C-12 and C-13/14 places further biosynthetic constraints on the pathway, including a dehydration to produce an alkene or alkane template for introduction of heavy molecular oxygen at both positions. Based on the stereochemical requirements, we have proposed two possible mechanisms for dihydroxylation.

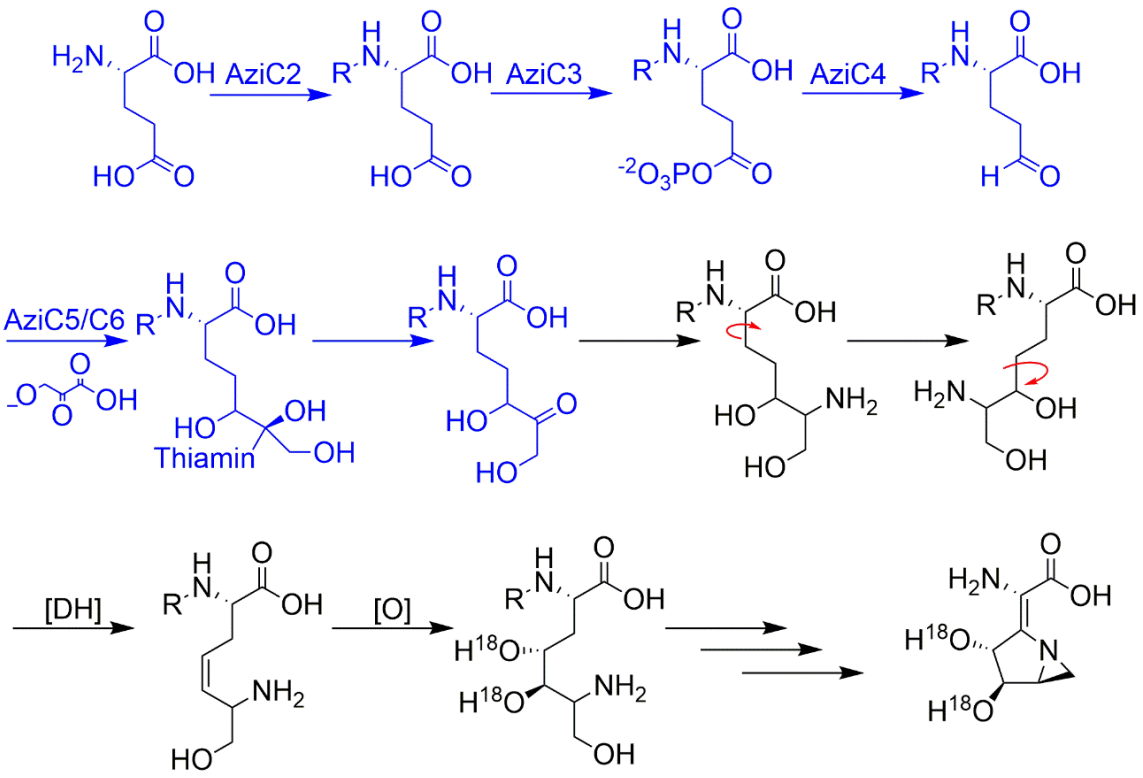


Figure 19 Oxidation in Azabicycle Biosynthesis
Blue steps have been validated by *in vitro* reconstitution. Black steps are proposed. R = $COCH_3$ or AziW

Route A involves the formation of an epoxide and ring opening by water to give the diol (Figure 20). This route only fulfills the stereochemical requirements if there is selective addition of water on both faces. Route B involves a syn dihydroxylation of a trans oriented starting material to result in the necessary stereochemistry. Identification of the oxygenase performing the reaction will be necessary to solve the mechanism.

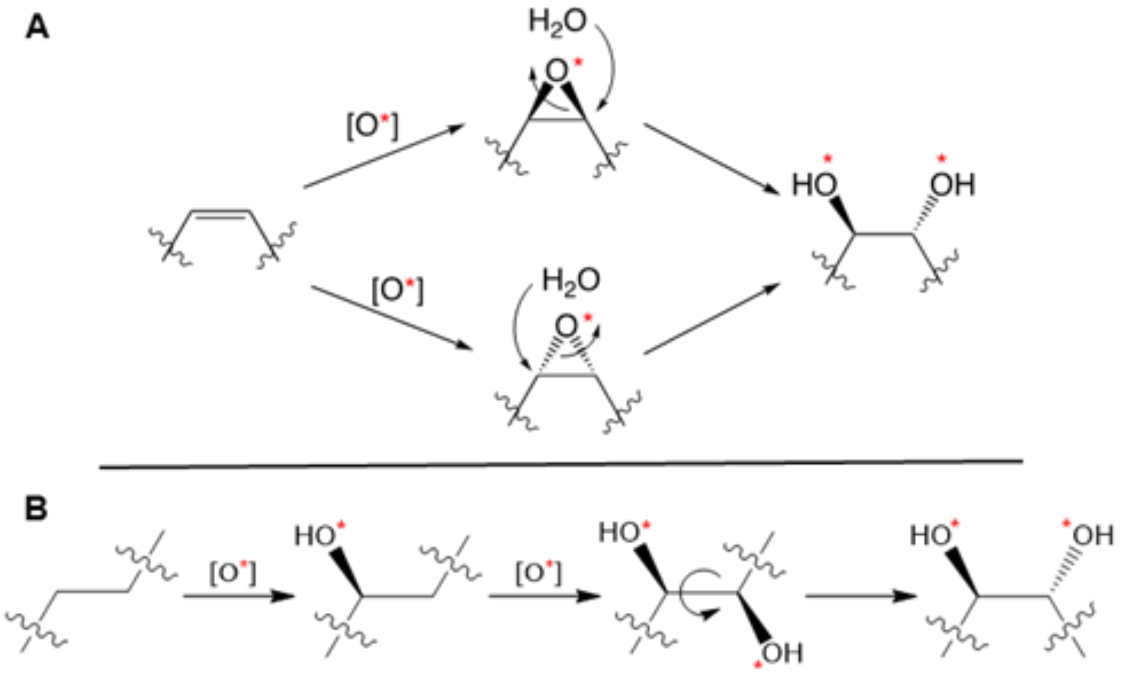


Figure 20 Proposed Route for Azabicycle Oxidation

Identification of the 2-oxoglutarate Fe-dependent Oxygenase AziU1

In the original annotation of the azinomycin biosynthetic gene cluster, there were only two oxygenases identified, AziB1 and AziC9. AziB1 was previously characterized as a cytochrome P450, which performs the oxidation on the naphthoate moiety. AziC9 is another cytochrome P450. AziC9 has not yet been characterized. With the current assignments, it was proposed that one of the P450s would need to carry out multiple oxidations. Based on this information, re-annotation of the gene cluster was performed to identify other possible oxygenases. This revealed AziU1 as a 2-oxoglutarate Fe-

dependent oxygenase. Multiple sequence alignment with AziU1 and characterized 2-oxoglutarate Fe-dependent oxygenase is shown in Figure 21.

Figure 21 AziU1 Multiple Sequence Alignment
* = Conserved Residue for Iron Coordination

There are two key features in 2-oxoglutarate Fe-dependent oxygenase structures. The non-heme iron is ligated by two histidine residues and either a glutamic acid or an aspartic acid residue. The sequence typically follows the pattern HX(D/E)X(n)H where X(n) is 50 to 90 amino acids.[84] In the case of AziU1, there is a conserved aspartic acid which we propose is ligated to the iron. Based on our bioinformatic and structure prediction studies we believe these conserved residues are His 162, Asp 164, and His 216. The other key structural feature of this enzyme family is a double-stranded β-helix fold, which the ligated iron is located within (Figure 22).[84-87]

Figure 22 AziU1 Model
Residues His162, Asp164, and His216 proposed to bind to iron shown in green

The 2-oxoglutarate Fe-dependent oxygenase family of enzymes are very similar to P450s in the types of oxidation reactions they can catalyze, but 2-oxoglutarate Fe-dependent oxygenases contain a non-heme iron and 2-oxoglutarate to perform their reactions. The proposed catalytic mechanism for hydroxylation is shown in Figure 23. The catalytic cycle is initiated by binding of 2-oxoglutarate, which displaces two water ligands from the Fe(II) center. Substrate entering the active site displaces the third water molecule, triggering binding of O_2, and generation of the Fe(III)-superoxo intermediate. The superoxo intermediate forms a new carbon-oxygen bond with the 2-oxoglutarate. This species will undergo decarboxylation and form the reactive Fe(IV)-oxo species.

This Fe(IV)-oxo species will abstract a hydrogen atom from the substrate to generate Fe(III)-OH and substrate radical. The hydroxylated product is then formed, regenerating a Fe(II) species. Product dissociation allows for another catalytic cycle to take place.[86]

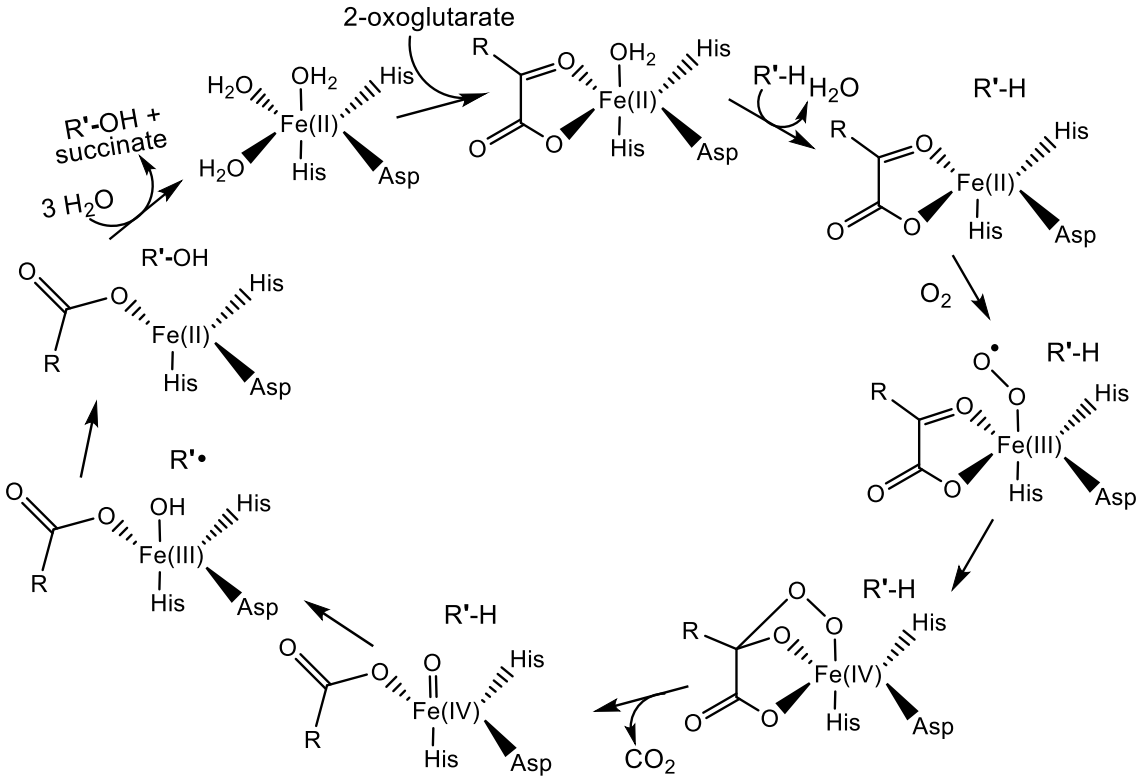


Figure 23 2-oxoglutarate Fe-dependent Oxygenase Mechanism

Significance

The incorporation of molecular oxygen at four sites in azinomycin is of great importance because it places biosynthetic and enzymatic constraints on the proposed biosynthetic routes to azinomycin. While the oxidation on at the C-3' position of the naphthoate has been characterized, the other three positions are still being studied. The unknown positions are of the utmost importance as one is the reactive epoxide ring that

gives azinomycin one half of its activity, and the other two sites adorn the aziridine containing azabicycle moiety responsible for the other half of activity.

Additionally, this experiment led to the reannotation of the azinomycin biosynthetic gene cluster. The original annotation was made based on the DNA sequence. This work highlighted the need for annotation using the protein sequence as well, as preferences in codon usage and differences in nucleotide percentage can hide homologous protein sequences when using gene annotation. The identification of AziU1 opens the door for further oxidative transformations to be explored within the context of producing azinomycin and analogues using synthetic biology efforts. Both AziC9 and AziU1 have been codon optimized for expression in *E. coli* for characterizations of each oxidative transformation necessary for azinomycin production. Determination of their substrates will play a critical role in resolving the final steps of azinomycin biosynthesis and opens the possibility for synthetic biology approaches to producing azinomycin and azinomycin analogs for targeted cancer therapy.

Material and Methods

General Methods and Materials

Mass Spectra (ESI) were obtained at the Laboratory for Biological Mass Spectrometry at the Department of Chemistry, Texas A&M University, with a Thermo Fisher Scientific Q Exactive Focus spectrometer. Gas chromatography and low-resolution mass spectra were recorded on a Thermo Fisher Scientific DSQ II GCMS spectrometer. APCI was recorded on a Thermo Fisher Scientific LC-Q DECA mass spectrometer. ^{1}H and ^{13}C

NMR spectra were recorded on a Varian Inova 500. ^{1}H NMR chemical shifts are reported as δ values in ppm relative to $CDCl_3$ (7.26 ppm). Coupling constants are reported in hertz (Hz). Unless indicated otherwise, $CDCl_3$ served as an internal standard for all $^{13}C\{^{1}H\}$ spectra. Fermentations were run on a Fermentation Design Inc. Model # MS21 (Allentown, PA, USA) with a total capacity of 15 L. Thin layer chromatography (TLC) was used to monitor azinomycin production. Flash column chromatography was performed using 60 Å Silica gel (Silacycle, 230 – 400 mesh) as a stationary phase. The $^{18}O_2$ was obtained from ICON Services Summit, NJ 07901.

Organism

Streptomyces sahachiroi (NRRL 2485) was obtained from the American Type Culture Collection (ATCC).

Culture Conditions and Fermentation

A 1 cm^2 piece of GYM plate containing *S. sahachiroi* spores was used to inoculate a first stage culture of 100 ml of PS5 medium (0.5 g Starch, 0.5 g Pharmamedia per 100 ml water) in a 250 ml flask. The culture was incubated at 30 °C for 24 h at 250 rpm. The second stage culture was prepared by inoculating 2 x 600 ml PS5 with 25 ml of first stage culture. The culture was incubated at 30 °C for 24 h at 250 rpm. The fermenter was prepared by autoclaving while containing 10 L of reduced PS5 (12.5 g Pharmamedia, 12.5 g Starch per 10 L water). Once the media reached room temperature after

sterilization, the fermenter was inoculated with the 2 second stage cultures and the fermenter agitated at 300 rpm for 72 h.[76]

Solid Medium

Solid medium cultures were inoculated by streaking a loop full of *S. sahachiroi* spore stock onto prepared GYM plates (4 g glucose, 4 g yeast extract, 10 g malt extract, 2 g $CaCO_3$, and 1 L water, pH 6.8). The plates were grown at 30 °C for 5 – 7 days before using.

Spore Stocks

S. sahachiroi spores were prepared from dehydrated GYM plates and were plated onto MS plates (20 g mannitol, 20 g Soy Flour per 1 L water). MS plates were incubated at 30 °C for 15 days. The gray spores were removed with sterile water and agitated. The spores were then filtered through sterile cotton, washer three times with sterile water, centrifuged at 3000 rom, resuspended in a minimal amount of 10% glycerol solution, flash frozen, and stored at -80 °C.

General Feeding Conditions

To facilitate the uptake of dioxygen during azinomycin biosynthesis, *S. sahachiroi* was grown in a closed system to allow for replacement of consumed oxygen with a mixture of $^{18}O_2$ and $^{16}O_2$. *S. sahachiroi* was grown under normal conditions until major

azinomycin production began at 18 h. A total of 3.7 L of $^{18}O_2$ was consumed during the experiment.

Extraction of azinomycin occurred 50 h after introduction of the mixed $^{18}O_2$:$^{16}O_2$ environment. Purification of azinomycin B afforded 35.9 mg of purified product. The sample was analyzed by NMR and Mass Spectrometry to identify sites of ^{18}O incorporation.

Isolation and Purification of Azinomycin B

Following fermentation, the cultures were centrifuged at 7,000 rpm at 4 °C. The cell pellets were discarded, and the medium extracted with an equal volume of dichloromethane. The organic layer was collected, dried over anhydrous magnesium sulfate, and concentrated in vacuo. The resulting crude extract was stored under diethyl ether at -80 °C. The solid was dissolved in a minimal amount of dichloromethane and precipitated with the addition of hexane to give a ratio of 1:29 CH_2Cl_2/hexane. The resulting suspension was centrifuged at 1,500 rpm and the supernatant discarded. Diethyl ether (2 mL) was added to the pellet, which was subsequently agitated, centrifuged at 3000 rpm and the supernatant discarded. The resulting residue was dissolved in dichloromethane (600 µL) to which hexanes (2 mL) was added. The heterogeneous mixture was centrifuged at 3,000 rpm and the supernatant retained. To the solution was added hexanes (4 mL) and the suspension centrifuged at 3,000 rpm to give azinomycin B as a solid.

If full purification is not achieved, azinomycin B can be further purified by flash column chromatography (95:5 dichloromethane: methanol). By TLC azinomycin exhibits a Rf of 0.23. A short column should be utilized to minimize overall contact with the silica gel and degradation by hydrolysis. The process can be repeated if necessary. The compound can be safely stored at -80 ºC under anhydrous diethyl ether.

Azinomycin B Characterization

Pale-white amorphous powder (1:9 CH_2Cl_2:hexane); IR (neat) ν_{max} 3338.4 (br), 2957.1, 2925.3, 2872.8, 1725.92 (br), 1619.3, 1601.7, 1511.2, 1417.6 cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ 12.40 (1H, br), 12.32 (1H, s), 8.54 (1H, dd, J = 3.6, 7.0Hz,), 8.20 (1H, br), 7.94 (1H, d, J = 2.9Hz), 7.46 (1H, d, J = 2.9Hz), 7.32 (1H, s), 7.32 (1H, s), 7.32 (1H, s), 5.50 (1H, d, J = 4.0Hz), 5.12 (1H, s), 4.64 (1H, dd, J = 4.0, 4.8Hz), 3.96 (3H, s), 3.96 (1H, br), 3.36 (1H, m), 2.98 (1H, d, J = 4.3Hz), 2.80 (1H, d, J = 4.3Hz), 2.70 (1H, s), 2.66 (3H, s), 2.30 (1H,s), 2.24 (1H, s), 2.18 (1H, s), 1.52 (1H, s); $^{13}C\{^1H\}$ NMR (75MHz, 125MHz, $CDCl_3$) δ 191.5, 173.0, 165.7, 164.0, 162.0, 156.0, 153.0, 150.8, 134.5, 133.3, 128.1, 127.9, 127.0, 125.4, 123.9, 122.3, 119.3, 118.6, 108.5, 84.4, 77.4*, 77.1*, 56.2, 55.7, 53.9, 46.4, 36.7, 24.5, 21.0, 20.3, 17.2. (* obscured by $CDCl_3$ solvent peak). $^{13}C\{^1H\}$ NMR (125 MHz, CD_2Cl_2) δ 191.3, 172.8, 165.8, 164.3, 162.2, 156.2, 154.1, 150.6, 134.6, 133.7, 128.4, 127.9, 126.9, 125.3, 123.8, 122.3, 191.3, 118.6, 108.4, 84.5, 76.9, 77.0, 56.1, 55.8, 53.8, 46.9, 36.9, 24.4, 20.9, 20.1, 17.1. APCI-MS (LRMS) 624.2, found 624.2.

Building Homology Model

Homology modeling was performed using SWISS-MODEL. Model quality was evaluated by having a QMEAN score of less than 4. The model for AziU1 was built upon the crystal structure of a 2-oxoglutarate-Fe^{2+}-dependent dioxygenase from *Methylibium petroleiphilum* (PDB 3pl0).[88] AziU1 and 3pl0 have 30 % sequence identity. The model was visualized using Pymol.

Table 2 Strains used in Chapter II

Strain	**Characteristic**	**Source**
S. sahachiroi	Azinomycin producing organism	ATCC

CHAPTER III

THIOESTER INTERMEDIATE CAPTURE STRATEGY DEVELOPMENT

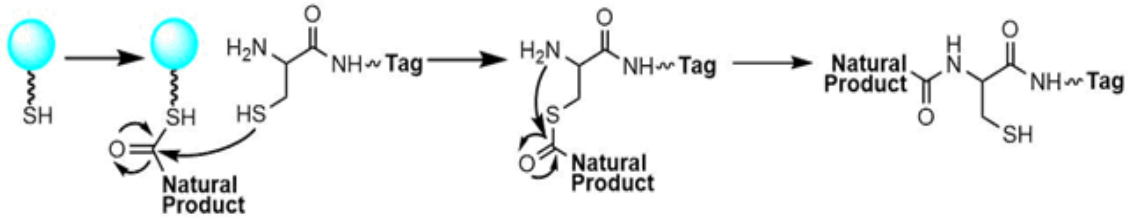
Introduction

As outlined in Chapter I, PKS and NRPS enzymes use carrier proteins to covalently bind substrates through a thioester bond. Once the catalytic cycle of the protein is complete, a thioesterase domain will hydrolyze the thioester bond, releasing the natural product either as the final product or making it available to the next enzyme in the biosynthetic pathway.[56] The azinomycin biosynthetic gene cluster contains five NRPS, one PKS, and four thioesterase genes.[79] The PKS AziB and the Type II thioesterase AziG are responsible for construction of the naphthoate moiety.[80] The five NRPS and three remaining thioesterases are predicted to construct the amide and ester bonded back bone of the azinomycins. Due to the importance of these enzymes in the construction of azinomycin, it is crucial to develop tools to study and analyze these systems.

Several methods are currently in use to characterize the enzymatic activity of these multidomain proteins and their thioester-bound enzyme intermediates. Thioesterase domains can be used to catalyze product release *in vitro,* but this method can be limited by finding a thioesterase specific to the system.[49] A mass spectrometry method, PPant Ejection Assay, is used to produce a fragment containing the phosphopantetheine arm with the thioester bound intermediate.[89-90] This method is limited by detection limits and ability to produce this fragment as many systems have

been unable to do so. Acid/base hydrolysis can be used to generate the free carboxylic acid of the intermediate. Not every system is susceptible to this hydrolysis, and many of the free carboxylic acids can be difficult to detect and ionize by LC-MS.[91] Cleavage with a nucleophile has proven to be more effective than hydrolysis for several systems and allows for direct functionalization for LC-MS analysis.[92-93] This chapter will detail the development of a cysteine containing probe for direct trapping and characterization of thioester-bound enzyme intermediates.

The thioester intermediate capture strategy is based on native chemical ligation in which a cysteine thiol carries out a nucleophilic attack on the carboxy group of a thioester. A rearrangement then occurs to produce an amide bond.[94] The capture agent contains a cysteine attached to a tag, either an affinity tag or a fluorophore. The tag allows for purification or identification from a complex mixture of compounds that might be present in a protein reaction or cell lysate, as well as provides direct derivatization for analysis (Scheme 1).



Scheme 1 Intermediate Capture Strategy

Preliminary Research

It was previously determined that the PKS AziB is responsible for production of 5-methylnaphthoic acid when the thioesterase AziG is present. In the absence of AziG, AziB produces a truncated product of 2-methylbenzoic acid (Figure 24).[80] Since the AziB system was previously validated *in vitro*, it was chosen as a test system to develop the capture strategy. AziB was initially characterized by acid/base hydrolysis *in vitro.* While product was eventually seen using hydrolysis and LC-MS, ionization and detection of the free carboxylic acids was difficult. Additionally, product was unable to be detected by the PPant Ejection Assay.

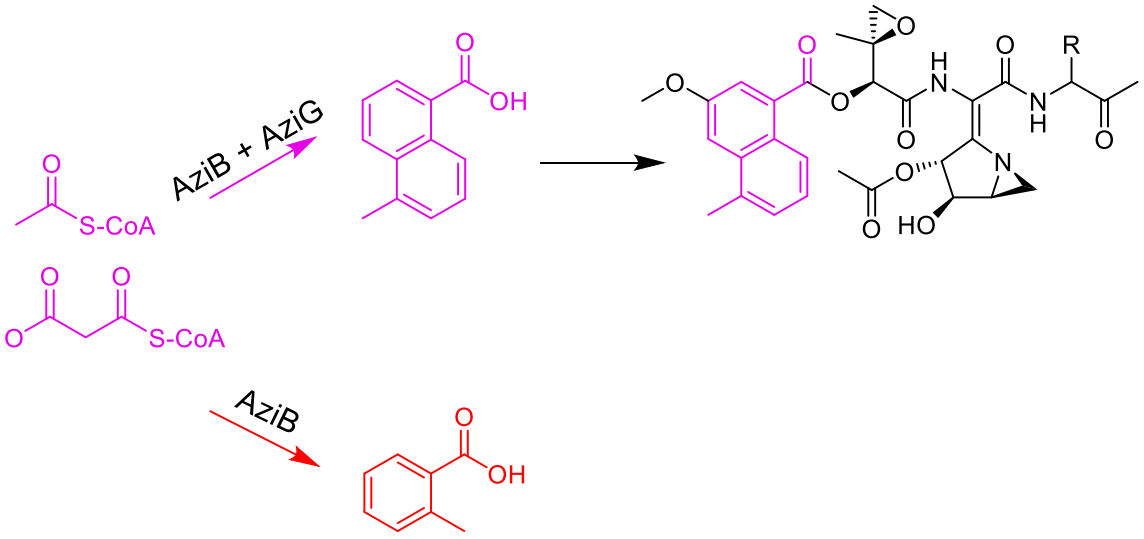


Figure 24 Role of AziB and AziG in Azinomycin Biosynthesis
R = H or CHOH

Results and Discussion

Synthesis and Evaluation of Capture Agents

Four compounds were synthesized to be tested as capture agents (Figure 25, NMR Appendix Figures 9-16). Three contained fluorophore tags, with differences in spectral and solubility properties, and the fourth contained a biotin tag that would allow for pull down or purification by streptavidin. The fluorophore capture agents were

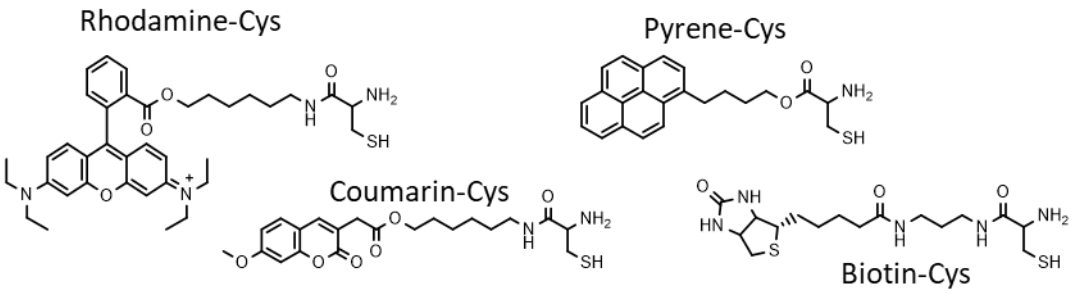


Figure 25 Capture Strategy Agents

unstable in the aqueous reaction buffer (Figure 26). The stability was evaluated by placing the compound in buffer for 12 h in the dark, lyophilizing the solution, and then performing LC-MS. Stability was determined by the presence of species other than starting capture agent in the analysis. Pyrene-Cys (Figure 26B) was the most stable and will be pursued in the future.

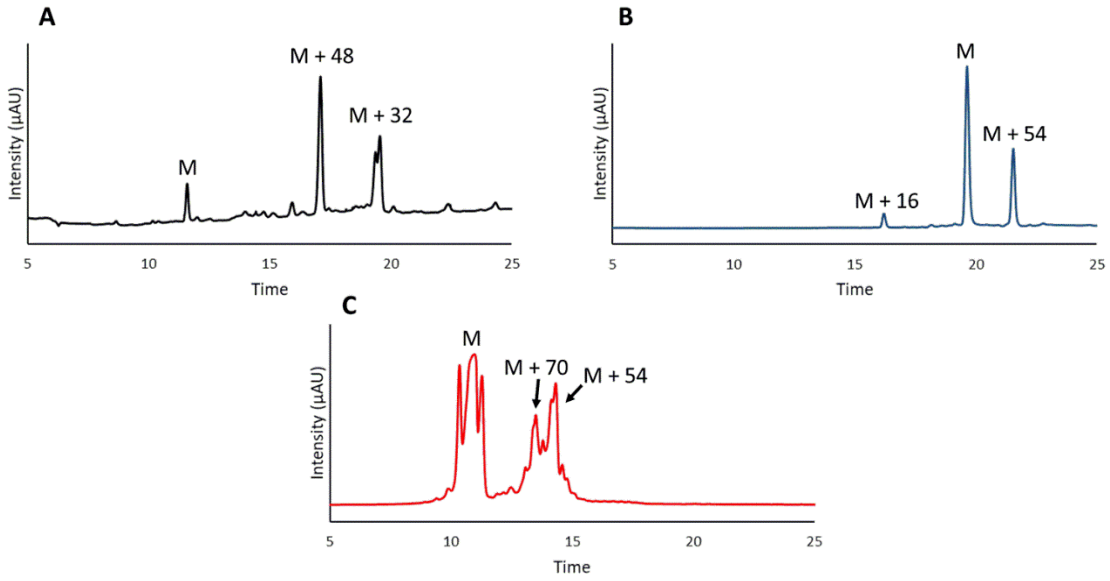
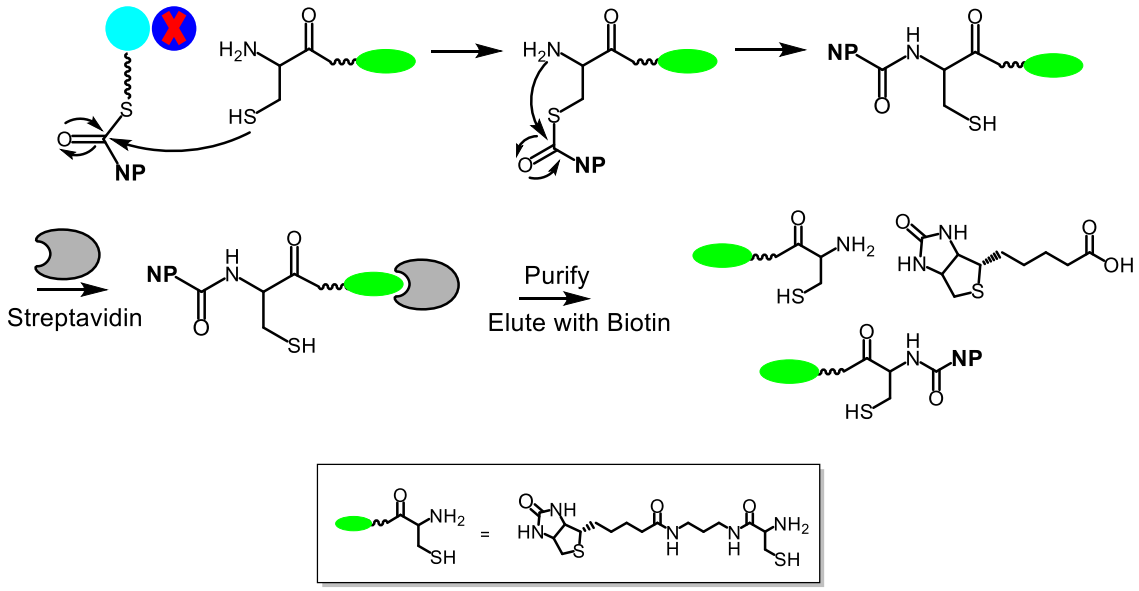


Figure 26 LC-MS Analysis of Fluorophore Capture Agents
The fluorophore capture agents were evaluated for their stability. **A**. Coumarin-Cys, λ = 330 nm. **B**. Pyrene-Cys, λ = 330 nm. **C**. Rhodamine-Cys, λ = 550 nm.

Due to the instability of the fluorophores, Biotin-Cys was pursued as the primary capture agent. For this approach, the next protein in the biosynthetic pathway will be absent to cause a buildup of the intermediate on the protein of interest. This can be performed *in vitro* by leaving out the next protein in the pathway, such as the thioesterase responsible for product release. Utilizing biotin as the tag allows for purification from the crude reaction using streptavidin (Scheme 2). Using a mutated streptavidin, free biotin elutes the capture agents, producing a mixture of unreacted starting capture agent, biotin, and capture agent-trapped intermediate for analysis.



Scheme 2 Biotin Capture Strategy Approach
NP = Natural Product

Validation of Capture Strategy with the PKS AziB

To develop the capture strategy, the PKS AziB was used. AziB, when incubated with malonyl-CoA and acetyl-CoA *in vitro*, produces 2-methylbenzoic acid. Purified AziB was incubated with the substrates to produce 2-methylbenzoic acid bound to AziB as a thioester. The protein was buffer exchanged to remove excess malonyl-CoA and acetyl-CoA, and then Biotin-Cys was added to facilitate trapping of the intermediate (Figure 27A). The reaction passed through a streptavidin matrix for purification and analyzed by LC-MS (Figure 27B). The formation of product was confirmed by LC-MS by mass ($[M+H]^+$ = 522.22) and by comparison to a synthetic standard of the expected compound (NMR Appendix Figure 17-18). Additionally, the MS/MS gives a characteristic fragmentation pattern for the expected compound (Figure 27C).

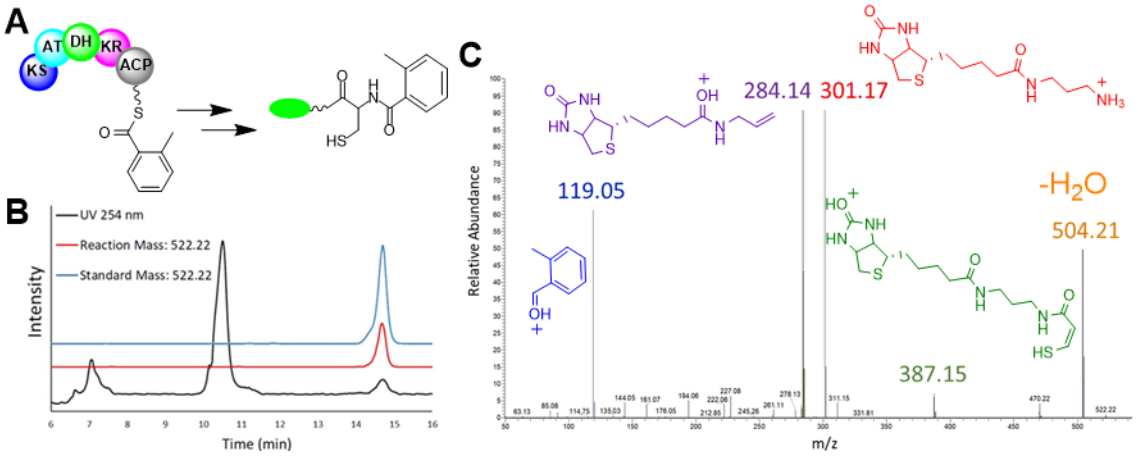


Figure 27 AziB + Biotin-Cys Reaction and Analysis
A. Domain architecture of AziB with 2-methylbenzoic acid bound to the acyl carrier protein domain. Domain abbreviations: Ketosynthase (KS); Acyltransferase (AT); Dehydratase (DH); Ketoreductase (KR); Acyl carrier protein (ACP). **B**. LC-MS traces of Biotin-Cys reaction with AziB and synthetic standard of Biotin-Cys-2-methylbenzoic acid in the reaction trace, unreacted capture agent is seen at 7 min, and biotin from purification on the streptavidin column is seen at 10.5 min. Black trace = UV 254 nm. Red trace = extracted ion chromatogram (EIC) of 522.22. Blue trace = EIC of 522.22 in synthetic standard sample. **C**. MS/MS fragmentation of 522.22 species at 14.7 min. Characteristic fragments are color coded.

Validation of Capture Strategy with the NRPS ClbN

To further validate the strategy, the Biotin-Cys capture agent was tested *in vitro* with the NRPS ClbN found in the colibactin biosynthetic pathway.[95] ClbN is responsible for production of the N-acyl-asparagine prodrug scaffold which is cleaved off at the end of colibactin biosynthesis to give the final active molecule. This N-acyl asparagine prodrug resistance mechanism is prevalent in NRPS containing gene clusters.[96-97] The NRPS ClbN accepts L-asparagine and performs a condensation reaction with the amino acid and an acyl-CoA substrate (Figure 28A). Purified ClbN was incubated with its

substrates to produce octanoyl-asparagine as a thioester, the capture agent Biotin-Cys was added, and the resulting mixture purified by streptavidin and analyzed by LC-MS (Appendix Figure 19-23). Analysis confirmed the trapping of octanoyl-asparagine by the capture agent in the free amide form and the cyclized version (Figure 28B).

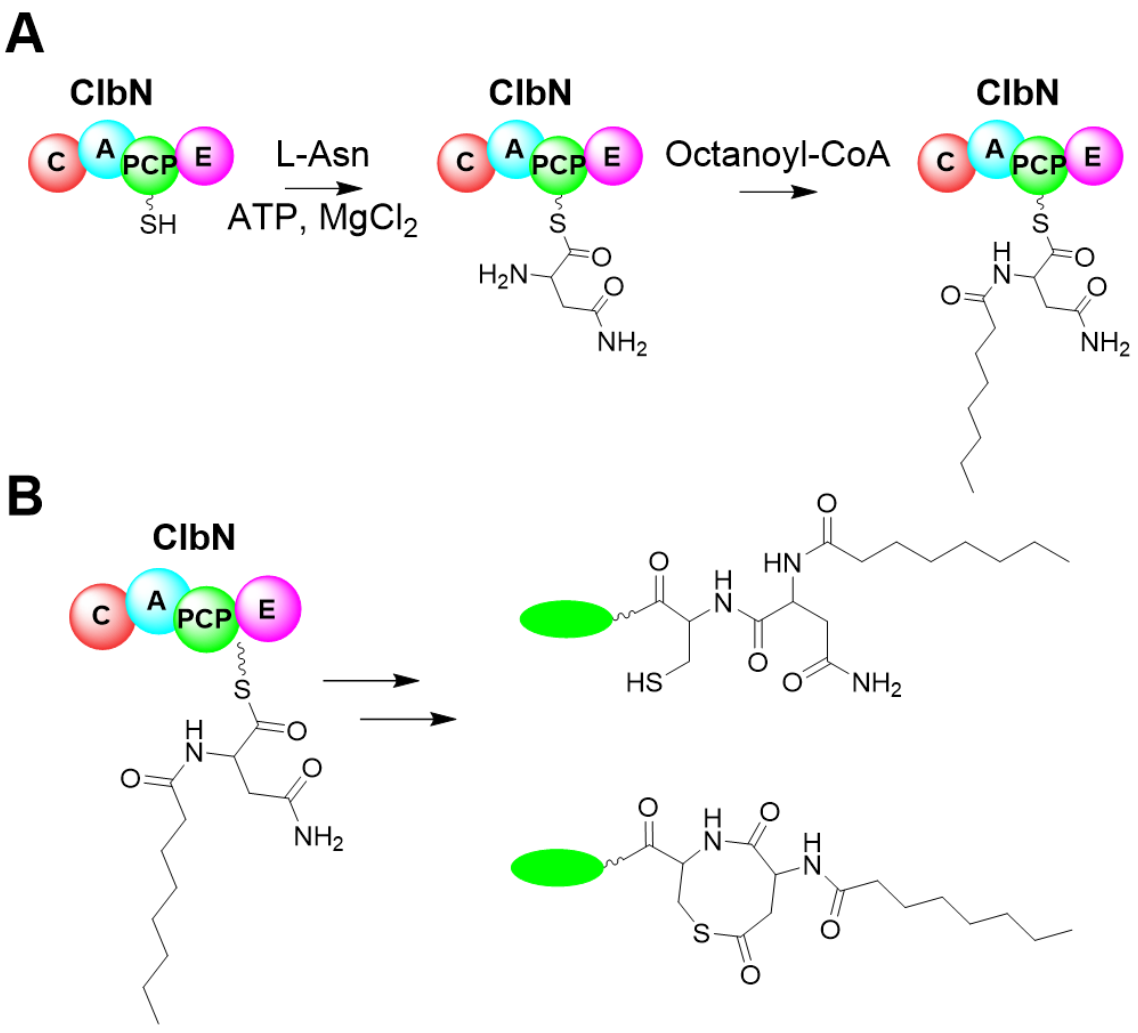


Figure 28 ClbN Reaction
A. ClbN Reaction with L-asparagine and octanoyl-Coenzyme A. Domain abbreviations: Condensation (C); Adenylation (A); Peptide Carrier Protein (PCP); Epimerization (E). **B**. ClbN reaction with Biotin-Cys. Free amide product ($[M+H]^+$ = 644.325) and cyclized product ($[M+H]^+$ = 627.299) were identified and characterized using Mass Spectrometry.

Significance

As azinomycin's backbone and the naphthoate moiety are produced and condensed by enzyme systems using thioester bound intermediates, it was of great importance to develop a reliable tool for the identification and characterization of enzyme products or intermediates in the biosynthetic pathway. This tool further confirms the production of 2-methylbenzoic acid by AziB in the absence of AziG and provides a protocol for use with other PKS and NRPS proteins. The cyclized product observed in the ClbN reaction points out a potential side reaction that can occur between the free cysteine thiol and a reactive group on natural product intermediates and will be considered when using this method in the future.

Furthermore, NRPS and PKS enzymes are extremely common in biosynthetic pathways and are being engineered for production of therapeutic metabolites, biofuels, and specialty chemicals. The development of the thioester intermediate capture strategy opens the door for the characterization of cryptic metabolite pathways for the identification of new drug or specialty chemical leads.

Material and Methods

General Procedures

All chemicals were purchased from Sigma-Aldrich and all bacterial media were purchased from Becton-Dickinson. Reactions were monitored by TLC using silica gel plates. ^{1}H and $^{13}C\{^{1}H\}$ NMR spectra were recorded on either a Bruker Avance 500 equipped with a cryoprobe or a Bruker Ascend 400. Deuterated solvents were purchased

from Cambridge Isotope Laboratories. Yields are reported for chromatographically pure compounds unless otherwise stated. Mass Spectra were obtained at the Laboratory for Biological Mass Spectrometry at the Department of Chemistry, Texas A&M University. A Phenomenex column (Prodigy 5 µm C18 150 Å, 150 x 4.6 mm) was used during HPLC and LC-MS purification and analysis. LC-MS was performed on a Thermo Scientific Ultimate 3000 UHPLC with a Thermo Fisher Scientific Q Exactive Focus mass spectrometer. Mass analysis of synthesized compounds was performed on an ion trap mass spectrometer (LCQ-DECA, Thermo Fisher Scientific). HPLC purifications were performed on an Agilent 1260 HPLC with an automated fraction collector. . pET28a-ClbN-NHis was a gift from Emily Balskus (Addgene plasmid # 51497 ; http://n2t.net/addgene:51497 ; RRID:Addgene_51497).

Bacterial Strains and Media

Escherichia coli BL21 (DE3) was used for *in vitro* protein expression.

Overexpression, Phosphopantetheinylation, and Purification of AziB

AziB was expressed, purified, and post-translationally modified as previously described.[80] *E. coli* BL21 (DE3) containing pET-24-aziB was grown in 5 ml of LB medium with 50 µg/ml kanamycin overnight. The culture was used to inoculate 1 L of LB medium and culture at 37 °C at 250 rpm until an OD_{600} of 0.6 was reached. Induction was performed by the addition of 1 ml of 1M *β*-D-1-thiogalactopyranoside (IPTG). The cultures were incubated at 16 °C for 24 h at 250 rpm. The cells were harvested by

centrifugation at 6500 rpm and resuspended in buffer containing 20 mM potassium phosphate, 500 mM NaCl, 1 mM dithiothreitol (DTT), 5 mM imidazole, and 20% glycerol, pH 7.4. Resuspended cells containing AziB were mixed with cells containing the phosphopantetheinyl transferase Svp. The cells were lysed by sonication, clarified using centrifugation to remove cellular debris. The supernatant containing AziB and Svp was used to post-translationally modify AziB to generate holo-AziB. Coenzyme A, 4 mg, was added to the supernatant and the mixture incubated at 28 °C for 1 h. After the reaction, the mixture was purified using a HisTrap FF 5 ml column (GE Healthcare Life Sciences). The purified protein was concentrated using a 50 kDa centrifugal ultrafiltration unit (Amicon Ultra centrifugal filter unit, Millipore).

Capture Strategy with AziB and Rhodamine-Cys

Purified holo-AziB was buffer exchanged into reaction buffer containing 50 mM potassium phosphate pH 7.5 and 20% glycerol. AziB (20 mg) was incubated with 50 µM acetyl coenzyme A, 250 µM malonyl coenzyme A, 10 µM dihydronicotinamide-adenine dinucleotide phosphate (NADPH), and 1 mM dithiothreitol (DTT) at 28 °C for 12 h. Following incubation, guanidine hydrochloride (8 M) was used to denature the protein. Rhodamine-Cys was added in 10:1 molar excess to AziB alongside 1 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and the mixture was incubated in the dark for 12 h. The reaction was extracted with 10 ml dichloromethane and the solvent removed in vacuo. As a control, Rhodamine-Cys was incubated in the dark for 12 h in reaction buffer with 1 mM TCEP and 8 M guanidine hydrochloride and extracted with

DCM. The extracted compounds were analyzed by LC-MS using a Phenomenex column (Prodigy 5 µm C18 150 Å, 150 x 4.6 mm). The sample was monitored by UV at λ = 550 nm. The column was pre-equilibrated with 80% A (water, 0.1% formic acid) and 20% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-80% B-20%; 1 min, A-80% B-20%; 23 min, A-0% B-100%; 33 min, A-0% B-100%; 35 min, A-80% B-20%; 40 min, A-80% B-20%, flow rate 0.75 µL/min. LC-MS results found in Appendix Figure 8.

Evaluation of Fluorophore Capture Agents

To evaluate stability of the fluorophore containing capture agents (Rhodamine-Cys, Coumarin-Cys, Pyrene-Cys), each agent was incubated in reaction buffer with 1 mM TCEP for 12 h in the dark after which the reaction was lyophilized and analyzed by LC-MS. The sample was monitored by UV (Rhodamine-Cys λ = 550 nm, Coumarin-Cys λ = 330 nm, Pyrene- Cys λ = 330 nm). The column was pre-equilibrated with 80% A (water, 0.1% formic acid) and 20% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-80% B-20%; 1 min, A-80% B-20%; 23 min, A-0% B-100%; 33 min, A-0% B-100%; 35 min, A-80% B-20%; 40 min, A-80% B-20%, flow rate 0.75 µL/min. Stability was determined by presence of species other than starting capture agent in HPLC trace and by mass.

Capture Strategy with AziB and Biotin-Cys

Purified holo-AziB was buffer exchanged into reaction buffer containing 50 mM potassium phosphate pH 7.5 and 20% glycerol. AziB (20 mg) was incubated with 50 µM acetyl coenzyme A, 250 µM malonyl coenzyme A, 10 µM NADPH, and 1 mM DTT at 28 °C for 12 h. Following incubation, AziB was buffer exchanged into fresh reaction buffer to remove unreacted CoA substrates. Biotin-Cys was then added in 10:1 molar excess to AziB alongside 1 mM TCEP and 8 M guanidine hydrochloride. The mixture was incubated for 12 h. Biotinylated compounds were purified from the crude reaction mixture using Roche Streptavidin Mutein Matrix following the manufacturer's instructions. Eluted fractions were analyzed by LC-MS. The sample was monitored by UV at $\lambda = 254$ nm. The column was pre-equilibrated with 80% A (water, 0.1% formic acid) and 20% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-80% B-20%; 1 min, A-80% B-20%; 23 min, A-0% B-100%; 33 min, A-0% B-100%; 35 min, A-80% B-20%; 40 min, A-80% B-20%, flow rate 0.75 µL/min. The reaction was compared to a control lacking AziB as well as a synthetic standard of Biotin-Cys-2-methylbenzoic acid

Overexpression and Purification of ClbN

ClbN was expressed and purified following previously established procedure.[95] A starter culture of pET-28a-ClbN E. coli (60 ml) in LB medium with kanamycin was grown for overnight at 37 °C. The starter culture was used to inoculate 6 x 1 L LB medium with kanamycin. Cultures were incubated at 37 °C, induced with IPTG at $OD_{600} = 0.6$, and

then incubated at 16 °C for 16 h. The cells were harvested by centrifugation and resuspended in lysis buffer (20 mM Tris-HCl pH 7.8, 500 mM NaCl, 10 mM MgCl2) with 5 mM imidazole and 1 mM phenylmethylsulfonyl fluoride (PMSF). The cells were disrupted by sonication and cell debris removed by centrifugation. ClbN was purified from the crude cell lysate on a HisTrap HP 5 mL column (GE Healthcare Life Sciences). A stepwise imidazole gradient was used to purify and elute pure protein, with at least two column volumes of each being used (25 mM, 50 mM, 75 mM, 100 mM, 125 mM, 150 mM, and 250 mM) Pure protein was obtained in elution fractions containing 100 mM to 250 mM imidazole as observed by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS PAGE).

Capture Strategy with ClbN and Biotin-Cys

ClbN was desalted and buffer exchanged into ClbN reaction buffer (40 mM HEPES buffer pH 7.5, 33 mM NaCl, 4 mM MgCl2, 400 µM DTT) To post-translationally modify ClbN 125 µM Coenzyme A and 250 nM Sfp were added along with ClbN and reaction mixture incubated at room temperature for 1 h. To initiate the ClbN reaction, 4 mM L-Asn, 5 mM ATP, 900 µM octanoyl-CoA, and 6 % DMSO were added and allowed to incubate for 3 h. The control reaction to verify product formation was quenched with methanol and products bound hydrolyzed with KOH and HCl. Octanoyl-asparagine was detected by MS in hydrolyzed positive control and not in a negative control lacking ClbN (Appendix Figure 19). Positive control octanoyl-asparagine HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{12}H_{22}N_2O_4$; 259.1652, found 259.1651.

For Biotin-Cys + ClbN reaction, ClbN with octanoyl-asparagine bound was buffer exchanged into reaction buffer (50 mM potassium phosphate pH 7.5, 20% glycerol) with 8 M guanidine hydrochloride and 1 mM TCEP. Biotin-Cys was added in 10:1 molar excess to ClbN and incubated at 4 °C for 16 h. Biotinylated compounds were purified from the crude reaction mixture using Roche Streptavidin Mutein Matrix following the manufacturer's instructions. Eluted fractions were analyzed by LC-MS. The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min. The reaction was compared to a negative control lacking ClbN as well as a synthetic standard reaction of Biotin-Cys + octanoyl-asparagine products (Appendix Figures 20-23).

Synthesis of Biotin-Cys

N-(tert-Butoxycarbonyl)-S-trityl-L-cysteine (222 mg) in 3 ml dimethylformamide (DMF) was stirred at 4 °C. To it was added 4-dimethylaminopyridine (DMAP) (100 mg), N-(3-Dimethylaminopropyl)-N-ethylcarbodiimide (EDC) (200 mg), and trifluoroacetate N-(3-Aminopropyl)biotinamide trifluoroacetate (250 mg). After 20 min, the reaction was warmed to room temperature and stirred for 3 h after which the solvent was removed. The crude reaction dissolved in dichloromethane (DCM) and extracted with water 2 times to give the protected biotin-cys(Boc)Trityl after removal of solvent.

Deprotection was performed by dissolving the protected biotin-cys(Boc)Trityl in 1 ml of water with 5 ml trifluoroacetic acid (TFA) and 100 mg triethylsilane for 4 hours. The reaction mixture was evaporated in vacuo and dissolved in 15 ml of 1:1 water:dichloromethane. The mixture was extracted two times with dichloromethane and the aqueous layer lyophilized to give the crude Biotin-Cys. Biotin-Cys (137 mg, 70.8%) was purified by HPLC to give Biotin-Cys as a clear thick gel. The column was pre-equilibrated with 80% A (water, 0.1% formic acid) and 20% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-80% B-20%; 1 min, A-80% B-20%; 23 min, A-0% B-100%; 33 min, A-0% B-100%; 35 min, A-80% B-20%; 40 min, A-80% B-20%, flow rate 0.75 µL/min. ^{1}H NMR (400 MHz, D_2O) δ 4.61 (dd, 1H, J = 7.9, 4.8), 4.43 (dd, 1H, J = 7.8, 4.8), 4.14 (t, 1H, J = 6.0), 3.38-3.20 (m, 6H), 3.06 (dd, 1H, J = 9.4, 4.4), 3.00 (dd, 1H, J = 12.9, 4.8), 2.78 (d, 1H, J = 13.1), 2.27 (t, 2H, J = 7.4), 1.80–1.53 (m, 6H), 1.48–1.35 (m, 2H); ^{13}C NMR δ 176.8, 167.8, 165.3, 62.0, 60.2, 55.3, 54.5, 39.6, 36.9, 36.4, 35.4, 27.9, 27.8, 27.6, 25.1, 24.8. HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{16}H_{30}N_5O_3S_2$; 404.1790, found 404.1779

Synthesis of Biotin-Cys-2-methylbenzoic acid

o-Toluic acid (40 mg) was dissolved in DMF and stirred at 4 °C. Biotin-Cys (161 mg) and EDC (60 mg) were added and the reaction was brought to room temperature. After 12 h the reaction was lyophilized and purified by HPLC to give product (134 mg, 85.9%). The column was pre-equilibrated with 80% A (water, 0.1% formic acid) and 20% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions

were as follows: time 0 min, A-80% B-20%; 1 min, A-80% B-20%; 23 min, A-0% B-100%; 33 min, A-0% B-100%; 35 min, A-80% B-20%; 40 min, A-80% B-20%, flow rate 0.75 µL/min. For comparison to the Biotin-Cys + AziB reaction, synthetic Biotin-Cys-2-methylbenzoic acid was placed in reaction buffer for 12 h and purified with Roche Streptavidin Mutein Matrix. Eluted fractions were analyzed by LC-MS and compared to the reaction of Biotin-Cys with AziB. HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{24}H_{36}N_5O_4S_2$; 522.2209, found 522.2206.

Synthesis of Rhodamine-Cys

Rhodamine B (120 mg) in DCM was stirred at 4 °C. To the rhodamine solution, DMAP (30 mg), 6-(Boc-amino)-1-hexanol (108 mg), and EDC (50 mg) were added and the reaction brought to room temperature. After 4 h, 20 ml of 5% HCl was added to the reaction and DCM used to extract the product. The Rhodamine-Linker-Boc intermediate was purified using silica column chromatography (ethyl acetate/ethanol 3:1). Deprotection to give the free amino linker was performed by dissolving the protected intermediate in 1:1 DCM/TFA for 12 h with stirring. The reaction was washed with 15 ml saturated NaHCO3 and extracted with ethyl acetate. The solvent was removed in vacuo and the product was used directly in the coupling to cysteine.

N-(tert-Butoxycarbonyl)-S-trityl-L-cysteine (70 mg) in DCM was stirred at 4 °C. To the cysteine solution was added DMAP (15 mg), Rhodamine B-linker (80 mg), and EDC (30 mg). The reaction brought to room temperature and stirred for 6 h. The reaction was washed with 20 ml of 5% HCl and extracted with DCM. The extract was washed with 10

ml saturated $NaHCO_3$ and NaCl. Protected Rhodamine B-Cys intermediate was purified using silica column chromatography (ethyl acetate/ethanol 1:1) and deprotection was performed by dissolving the intermediate in 1.5 ml DCM with 1 ml TFA and 29 mg triethylsilane under nitrogen. The reaction was stirred for 48 h after which the solvent was removed and the product purified with flash column chromatography (DCM/95% ethanol 9:2) to give a purple solid (48 mg, 49%). HRMS (ESI) m/z $[M]^+$ calcd for $C_{37}H_{49}N_4O_4S^+$; 645.3469, found 645.3481.

Synthesis of Coumarin-Cys

7-methoxycoumarin-4-acetic acid (70 mg) was dissolved in chloroform and stirred at 4°C. To the fluorophore was added, 6-(Boc-amino)-1-hexanol (109 mg), DMAP (40 mg), and EDC (60 mg). The reaction was warmed to room temperature and allowed to stir for 4 h. The reaction was washed with 15 ml 1 M HCl, 10 ml saturated $NaHCO_3$, and 10 ml saturated NaCl. The intermediate was purified using silica column chromatography (ethyl acetate/hexanes 3:1) to give a thick yellow gel (97 mg, 74%). The Boc protecting group was removed by dissolving the intermediate in 1:1 DCM/TFA for 12 h after which the solvent was removed and the deprotected coumarin-linker used directly. N-(tert-Butoxycarbonyl)-S-trityl-L-cysteine (111 mg) in 5 ml DCM was stirred at 4 °C. To the solution was added the deprotected coumarin-linker (90 mg), DMAP (15 mg) and EDC (60 mg). The reaction was stirred at room temperature for 12 h and subsequently washed with 20 ml 1 M HCl, 15 ml saturated NaHCO3, and 15 ml saturated NaCl. The intermediate was purified using silica column chromatography

(ethyl acetate/hexanes 3:1) to give a light green compound (103 mg, 59%). Deprotection to give Coumarin-Cys was performed by dissolving the protected intermediate in DCM with 1.5 ml TFA and 40 mg triethylsilane. The reaction proceeded with stirring for 12 h after which the solvent was removed. The residue was rinsed with water, filtered to remove insoluble compounds, and the aqueous solution lyophilized to give Coumarin-Cys (36 mg, 62%) as an off white solid. HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{21}H_{29}N_2O_6S$; 437.1746, found 437.1742

Synthesis of Pyrene-Cys

N-(tert-Butoxycarbonyl)-S-trityl-L-cysteine (185 mg) in DCM was stirred at 4 °C. Pyrene-butanol (200 mg) and EDC (80 mg) were added and the reaction brought to room temperature and stirred for 5 h. The reaction was washed with 10 ml 1 M HCl, followed by 10 ml saturated $NaHCO_3$ and 10 ml saturated NaCl. The solvent was removed in vacuo and the protected intermediate purified using silica column (ethyl acetate/hexanes 1:1). Deprotection to give Pyrene-Cys was performed by dissolving the protected intermediate in DCM with 2.5 ml TFA and 50 mg triethylsilane. The reaction proceeded with stirring for 4 h after which the solvent was removed. The reaction was purified using silica column chromatography (ethyl acetate/hexanes/methanol 15:5:1) to give purified Pyrene-Cys (83 mg, 54.9%) as a white solid. HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{23}H_{24}NO_2S$; 378.1528, found 378.1522.

Synthesis of Biotin-Cys-Octanoyl-Asparagine

Octanoyl-asparagine was synthesized following a previously established protocol.[95] HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{12}H_{22}N_2O_4$; 259.1652, found 259.1650. ^{1}H NMR (400 MHz, CD_3OD) δ 0.92 (t, J = 7.4 Hz, 3H) 1.34 (m, 8H), 1.64 (m, 2H) 2.26 (t, J = 7.1 Hz, 2H) 2.72 (m, 1H) 2.78 (dd, J = 5.3, 8.1 Hz, 1H) 4.75 (q, J = 7.4 Hz, 1H) $^{13}C\{^1H\}$ NMR (400 MHz, CD_3OD) δ 14.1, 22.4, 25.4, 28.6, 28.8, 31.4, 35.2, 36.6, 48.9, 171.2, 172.3, 173.2.

Octanoyl-asparagine (20.3 mg) was dissolved in DMF and stirred at 4 °C. Biotin-Cys (30 mg) and EDC (50 mg) were added to the reaction and the reaction stirred for 12 h at room temperature. The solvent was removed in vacuo and the crude reaction used directly. For comparison to the product of the Biotin-Cys + ClbN reaction, the crude reaction was incubated in Biotin-Cys reaction buffer for 12 h and the reaction was lyophilized, purified with streptavidin, and compared to the reaction of Biotin-Cys with ClbN by LC-MS (Appendix Figures 20-23). The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min. Free amide product HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{28}H_{49}N_7O_6S_2$; 644.3259, found 644.3246. Cyclized product HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{28}H_{46}N_6O_6S_2$; 627.2993, found 627.2983.

Table 3 Strains and Plasmids used in Chapter III

Strain	Characteristic	Source
E. coli BL21 (DE3)	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3)	Invitrogen
E. coli BL21 (DE3) pET-24-aziB	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with AziB expression plasmid	Mori *et. al.*, 2016[80]
E. coli BL21 (DE3) pET-21-svp	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with Svp expression plasmid	Mori *et. al.*, 2016[80]
pET-24-aziB	*E. coli* expression vector for AziB	Mori *et. al.*, 2016[80]
pET-21-svp	*E. coli* expression vector for Svp	Mori *et. al.*, 2016[80]

CHAPTER IV

CHARACTERIZATION OF THE NRPS AZIA3

Introduction

As outlined in Chapter I, a feeding study with ^{13}C labeled substrates resulted in a proposed biosynthetic pathway to the epoxide moiety. A pathway from valine to 3-methyl-2-oxobut-3-enoic acid resulted in incorporation into the azinomycins, but the exact route from 3-methyl-2-oxobut-3-enoic acid to azinomycin remains uncharacterized (Figure 29).[78] From 3-methyl-2-oxobut-3-enoic acid to the final epoxide moiety present in azinomycin, a ketone reduction to an alcohol to facilitate condensation with the naphthoate moiety and an oxidation to form the epoxide are needed. A fermentation study using $^{18}O_2$ suggests the epoxide is formed enzymatically, likely from a P450 or 2-oxoglutarate Fe-dependent oxygenase found in the azinomycin biosynthetic gene cluster, but the order of keto-reduction and epoxide formation is unknown.[98]

The amide and ester bonded backbone of azinomycin is proposed to be constructed by the NRPS proteins found in the pathway. The azinomycin biosynthetic gene cluster contains five NRPS and four thioesterase genes.[79] The exact role of each NRPS in unknown at this time. The NRPS AziA3 consists of three domains, adenylation (A), ketoreductase (KR), and peptide carrier protein (PCP). In this chapter, we will outline the role of AziA3 in the formation of the epoxide moiety of azinomycin using *in vitro* assays and the intermediate capture strategy developed in Chapter III. The capture strategy approach was modified for protein isolated from the producing organism, *S. sahachiroi*.

This allows for the trapping of intermediates from proteins where the native substrate and product are unknown.

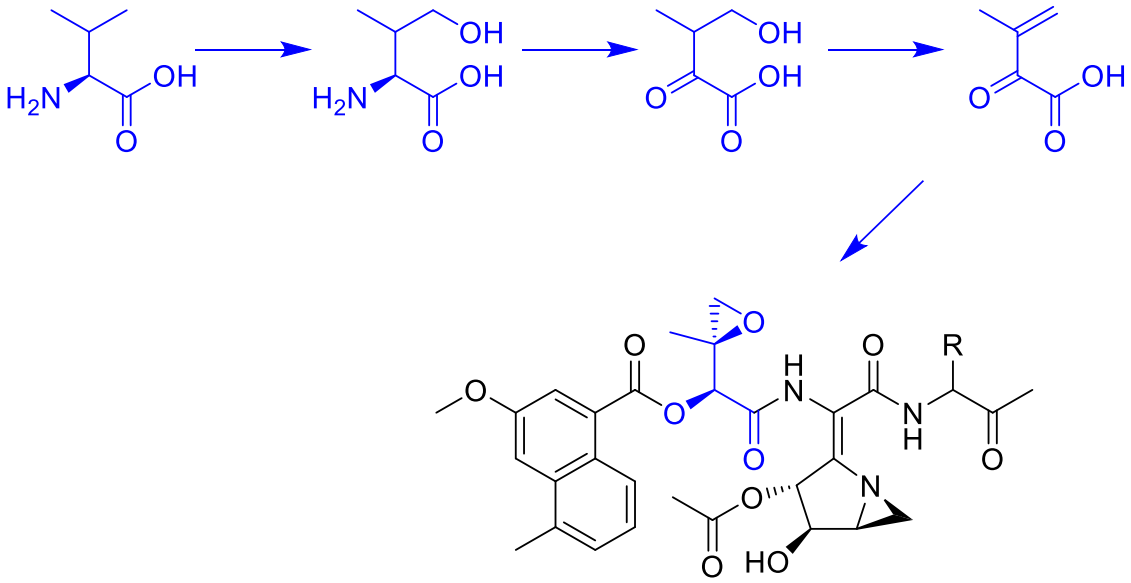


Figure 29 Biosynthesis of Epoxide Moiety
Blue = incorporated into azinomycin

Preliminary Research

AziA3 is a NRPS containing only one tailoring domain, a ketoreductase. Based on similarity to NRPS proteins in cereulide and kutzneride biosynthesis, it is believed AziA3 reduces the ketone of a small keto-acid to a hydroxy acid (Figure 30).[99-100]

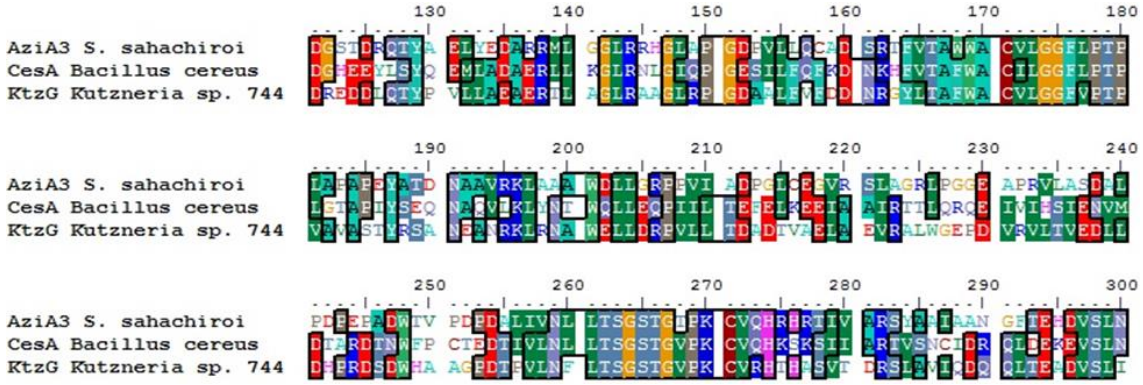


Figure 30 AziA3 Sequence Alignment with Homologous Enzymes
Dr. Shogo Mori

To assess the substrate specificity of the adenylation domain of AziA3, an ATP-[^{32}P]PPi exchange assay was performed.[99] Adenylation domains catalyze the activation of a carboxylic acid substrate by the addition of adenosine monophosphate (AMP) from adenosine triphosphate (ATP), generating an AMP-substrate intermediate and pyrophosphate (PPi). Since the reaction catalyzed by the A domain is reversible, ATP can be converted into [^{32}P]ATP overtime in the presence of [^{32}P]PPi. [^{32}P]ATP (or ATP) is selectively trapped by activated charcoal and detected by a liquid scintillation counter. The following compounds were used as test substrates: an epoxide species (2-((S)-2-methyloxiran-2-yl)-2-oxoacetic acid), valine, pyruvate, and tryptophan (Figure 31). The epoxide analog possessed the best activity followed by a moderate activity of pyruvate. Valine and tryptophan showed no activity, suggesting that the AziA3 A domain has a high selectivity for α-keto carboxylic acids. At the time of the ATP-[^{32}P]PPi exchange assay, the post-translational modification to install the phosphopantetheine arm was inefficient. Due to this, *in vitro* reconstitution was not performed at the time.

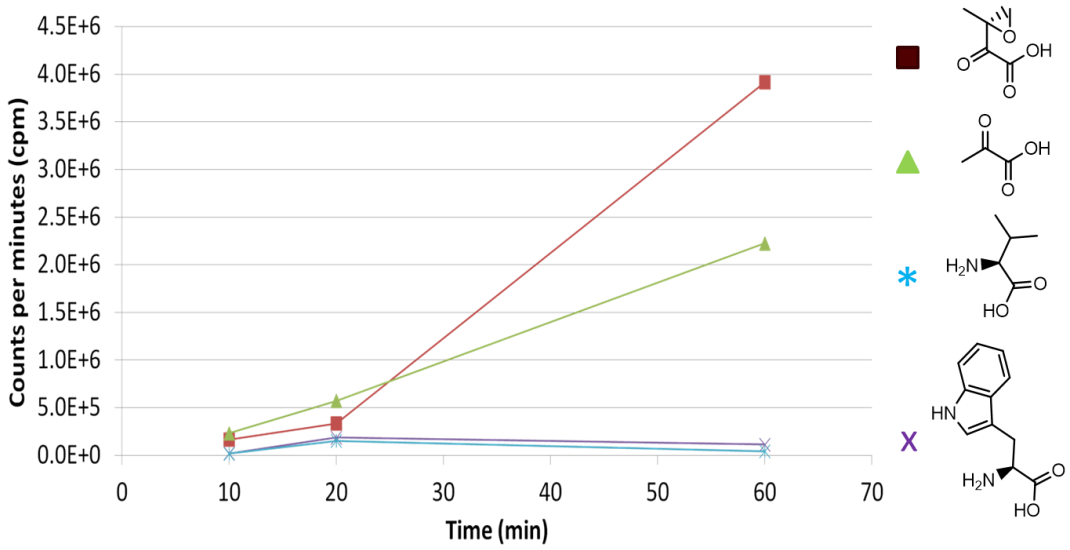


Figure 31 ATP-[^{32}P] Exchange Assay
Dr. Shogo Mori

Based on the ATP-[^{32}P]PPi exchange assay and bioinformatic results, a mechanism was proposed for AziA3. The A domain activates the substrate by transferring AMP from ATP onto the substrate carboxylic acid. The PCP domain's active site serine residue is post-translationally modified with a flexible phosphopantetheinyl arm from Coenzyme A, catalyzed by a phosphopantetheinyl transferase (PPTase). This flexible arm forms a thioester bond with the activated substrate, and the KR domain subsequently catalyzes the reduction of the α-ketone to an alcohol functional group (Figure 32).

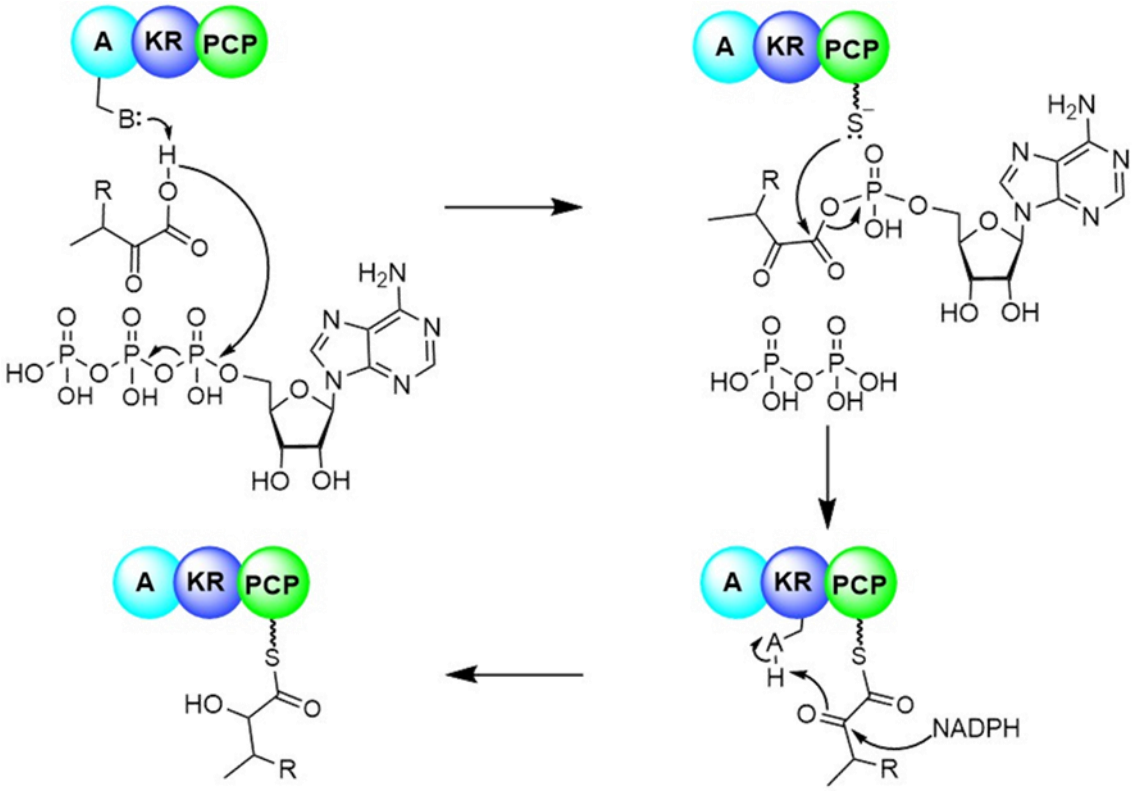


Figure 32 Proposed AziA3 Mechanism

Due to the possible promiscuity of the adenylation domain and the low level of post-translational modification with commercial PPTases, the native substrate of AziA3 could not be determined *in vitro*. Here we present the use of the intermediate capture strategy with protein isolated from the producing organism, *S. sahachiroi.* The AziA3 post-translational modification will be performed by PPTases present in *S. sahachiroi*, and the substrate for AziA3 will be provided by *S. sahachiroi* and does not require the identification or feeding of substrates.

Results and Discussion

To adapt the intermediate capture strategy for use with protein isolated from the native organism, the gene coding for the protein following AziA3 in the biosynthetic pathway must be disrupted to facilitate buildup of product on AziA3. As AziA3 does not contain a thioesterase or condensation domain, it was hypothesized that product release from AziA3 would need to be facilitated by a thioesterase. A Type II thioesterase, AziA6, is adjacent to AziA3 in the gene cluster and is under the same promoter as AziA3. To probe for an interaction between AziA3 and AziA6, a size exclusion assay was performed in which AziA3 and AziA6 were separately expressed and purified and then incubated together to detect complex formation. A new species was detected in the chromatogram. The fraction collected from this new species showed AziA3 and AziA6 (Figure 33). The assay was performed with the other thioesterases in the gene cluster, and AziA6 was the only one that showed an interaction with AziA3.

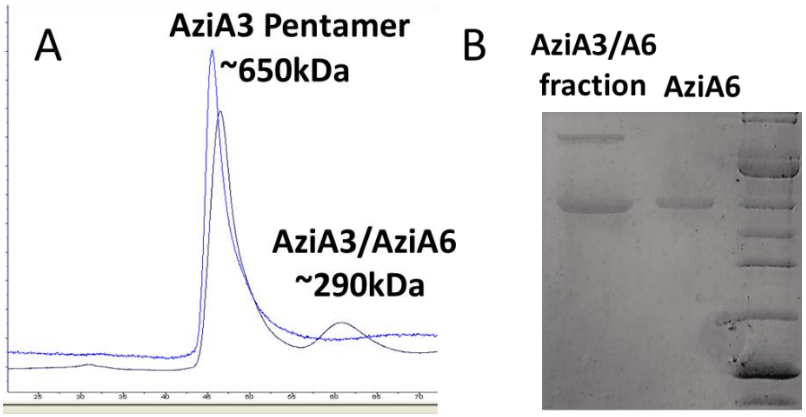


Figure 33 Size Exclusion Chromatography
A. UV 280 nm trace. **B**. SDS-PAGE of new ~290 kDa species vs purified AziA6

To facilitate intermediate build up on AziA3, *aziA6* was disrupted in *S. sahachiroi* by homologous recombination. Regions upstream and downstream of *aziA6* were cloned into the disruption plasmid pKC1139 to generate the plasmid pKCAziA6 (Figure 34). The disruption plasmid was transformed into *E. coli* S17 for conjugal transformation into *S. sahachiroi*. Incorporation of the plasmid was screened for by using the apramycin resistance gene (*apr*) located on the pKC1139 plasmid back bone,

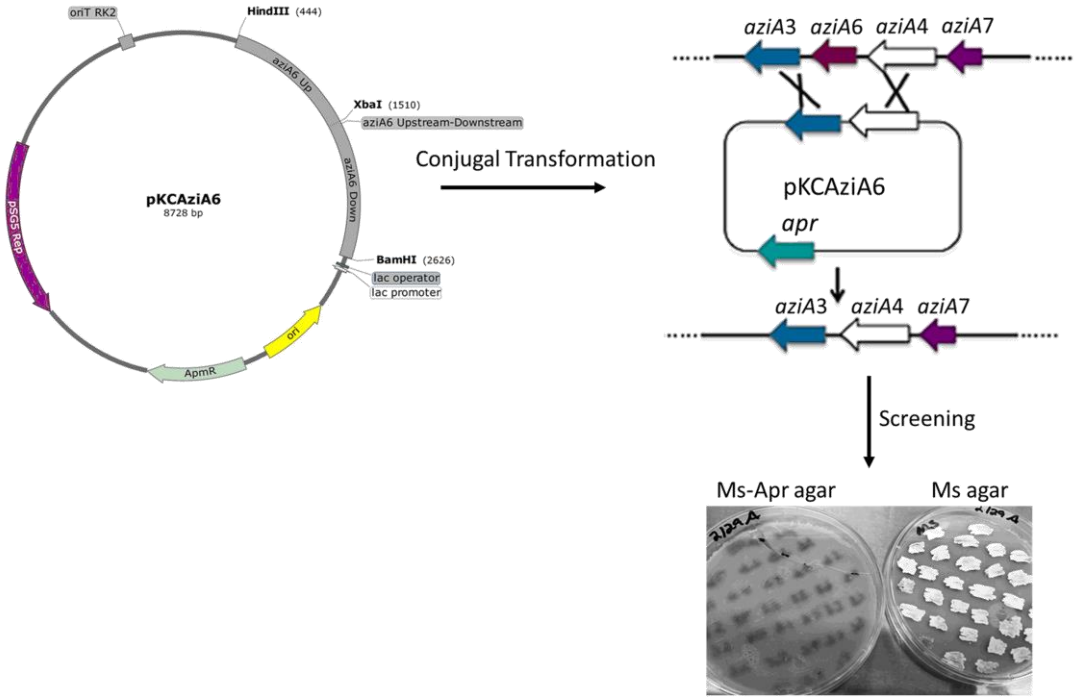


Figure 34 Generation of Δ*aziA6* strain

colonies with successful plasmid incorporation are referred to here as ex-conjugants. Subsequent screening was performed to identify double crossovers which have lost *apr* and *aziA6*. To confirm loss of *apr* and *aziA6*, PCR and Southern Blot was used.

Additionally, RT-PCR was employed to evaluate gene expression of the genes surrounding *aziA6* (Figure 35).

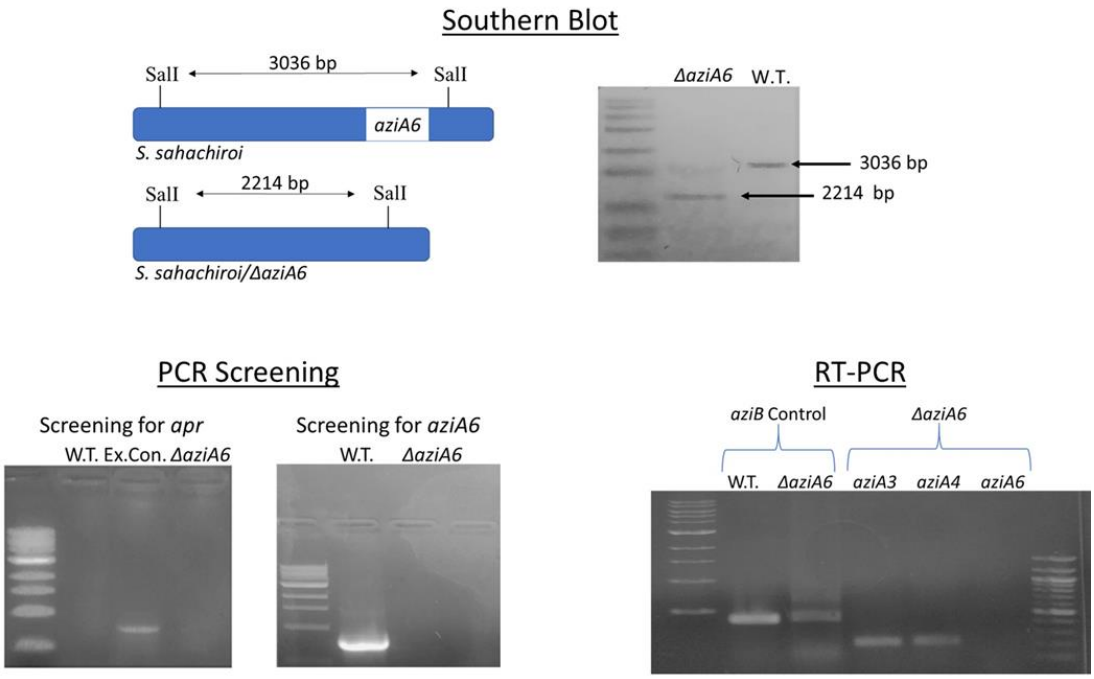


Figure 35 Confirmation of Δ*aziA6* Strain

Once the disruption of *aziA6* was confirmed, fermentation of the strain was performed to confirm the loss of azinomycin production and to identify the buildup of intermediates or byproducts resulting from the loss of azinomycin production (Appendix Figure 24). In addition to the azinomycins, excess naphthoate moiety is typically seen in wild type crude extracts. In the crude extract of Δ*aziA6* strain, there was both loss of the azinomycins and the naphthoate moiety. After loss of azinomycin production was confirmed, an AziA3 expression construct was introduced to the Δ*aziA6* strain.

pSETAziA3 is an integrative vector with the constitutive *ermE** promoter for continual expression of AziA3-6XHis. AziA3-6XHis was expressed and purified using His-Trap purification (Figure 36).

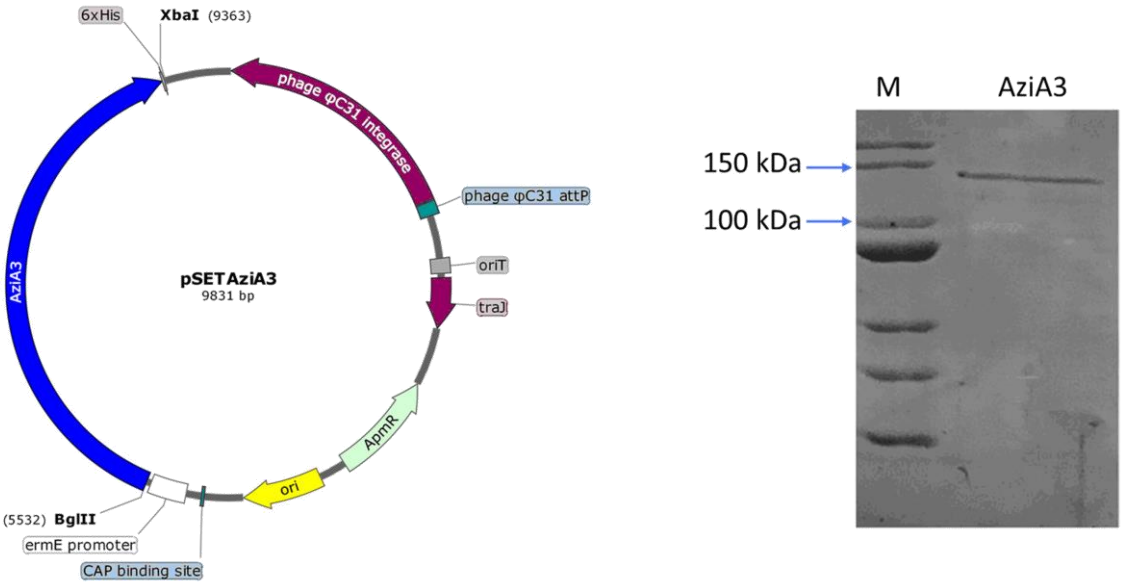


Figure 36 Overexpression of AziA3 in *S. sahachiroi*

Immediately after purification, Biotin-Cys probe was incubated with AziA3-6XHis, the biotin containing compounds purified by Streptavidin Mutein Matrix, and the resulting mixture analyzed by LC-MS. Potential products were proposed based on the feeding and ATP exchange assay results (Figure 37).

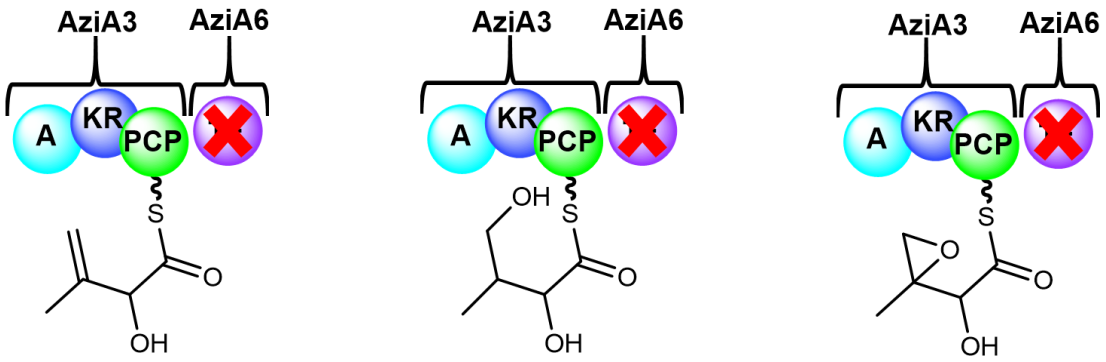


Figure 37 Proposed AziA3 Products

Incubation of AziA3 with Biotin-Cys led to the trapping of the proposed epoxide compound as characterized by the resulting cyclized product (Figure 38A). Presence of the cysteine thiol led to ring opening of the epoxide as confirmed by LC-MS ($[M+H]^+$ = 518.21) (Figure 38B), MS/MS (Figure 38C), and NMR (Appendix Figure 26-27). This result suggests the epoxide is fully formed before the moiety is assembled into the azinomycin backbone. After ketone reduction, condensation with the naphthoate and azabicycle moieties is needed. As AziA3 does not contain a condensation domain, the epoxide moiety would be transferred to another NRPS in the gene cluster for condensation. In addition to pursuing the NRPS responsible for condensation, a mutant of AziA3 with the ketoreductase domain inactive was next overexpressed and purified from *S. sahachiroi* to trap the starting material of AziA3.

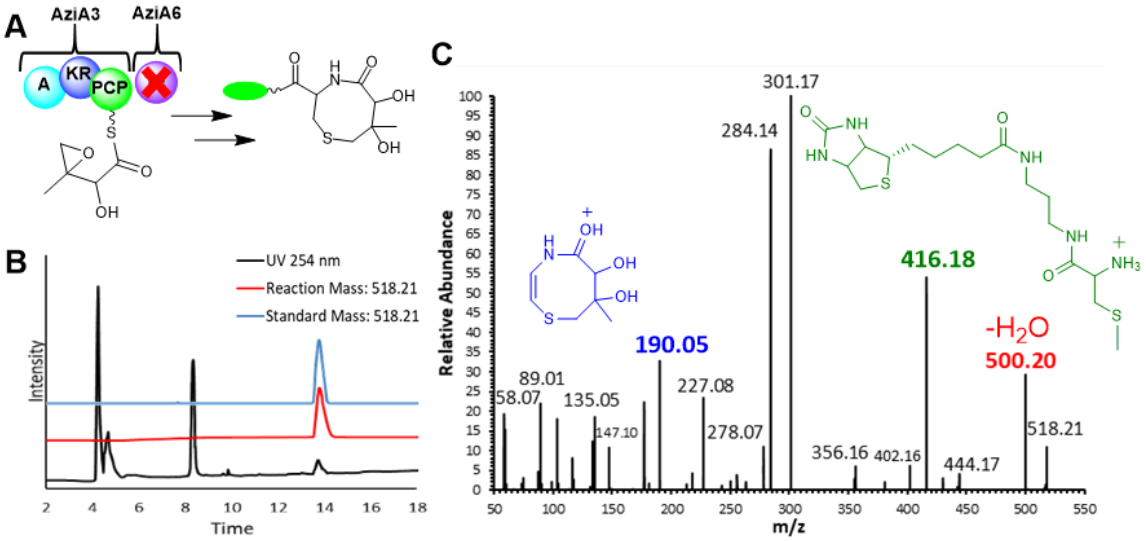


Figure 38 AziA3 + Biotin-Cys Reaction and Analysis
A. Domain architecture of AziA3 with compound 15 bound to the peptide carrier protein domain. Domain abbreviations: Adenylation (A); Ketoreductase (KR); Peptide carrier protein (PCP). **B**. LC-MS traces of Biotin-Cys reaction with AziA3 and synthetic standard of cyclized product. In the reaction trace, unreacted capture agent is seen at 4.2 min, disulfide dimer of capture agent at 4.7 min, and biotin from purification on the streptavidin column is seen at 8.3 min. Black trace = UV 254 nm. Red trace = extracted ion chromatogram (EIC) of 518.21. Blue trace = EIC of 518.21 in synthetic standard sample. **C**. MS/MS fragmentation of 518.21 species (cyclized product) at 13.8 min. Characteristic fragments are color coded.

To first validate the necessity of the ketoreductase domain, a strain of *S. sahachiroi* with the KR domain of AziA3 disrupted was generated using conjugal transformation and homologous transformation. A pKCAziA3ΔKR plasmid was generated using regions upstream and downstream of the KR domain of AziA3 (Figure 39A). PCR was used to screen for the presence of the *apr* gene (Figure 39B) and the KR domain (Figure 39C). RT-PCR was used to confirm the expression of the other AziA3 domains as well as a gene downstream of AziA3 (Figure 39D). Fermentation of the *aziA3*Δ*KR* strain was performed to confirm the necessity of the KR domain in the formation of azinomycin (Appendix Figure 25). LC-MS analysis confirmed loss of

azinomycin production. Unlike the case of the disruption of *aziA6*, excess naphthoate moiety was still seen in the crude extract. This is an interesting result and may point to a larger system of protein-protein interactions necessary for azinomycin production.

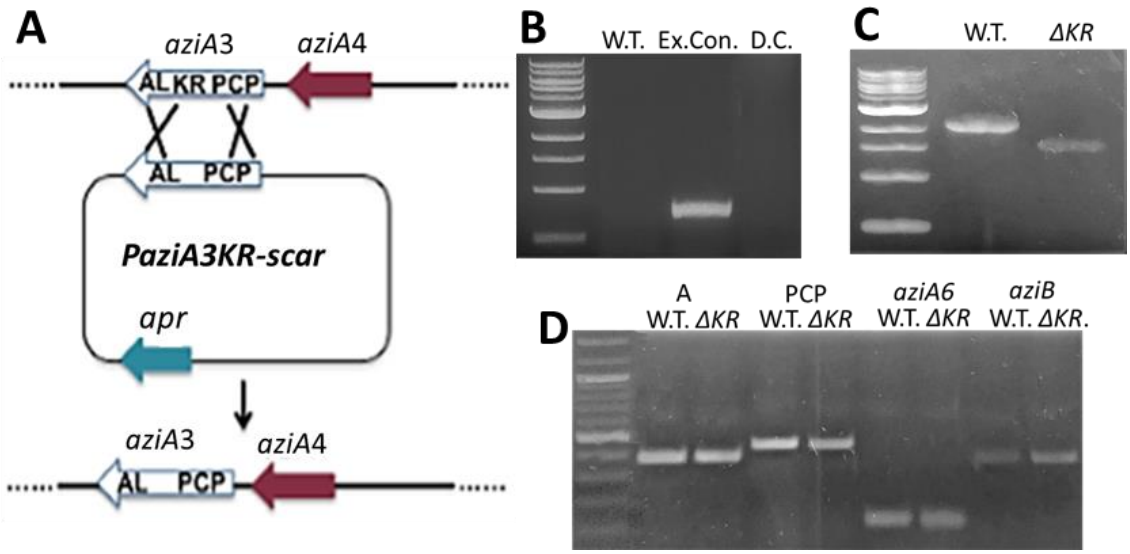


Figure 39 Generation and Confirmation of *aziA3*Δ*KR* Disruption Strain
A. Scarless gene disruption diagram. **B**. PCR screening for *apr* gene. **C**. PCR confirmation of absence of KR domain in disruption strain. **D**. RT-PCR confirming expression of surrounding domains and downstream genes in. W.T. = Wild Type. D.C. = Double Crossover.

Once the KR domain proved to be vital for azinomycin production, AziA3ΔKR-6xHis was overexpressed in the Δ*aziA6* strain to facilitate trapping of the AziA3 starting material using the Biotin-Cys probe. The pSETAziA3ΔKR plasmid was generated and incorporated into the Δ*aziA6* strain. Based on previous feeding studies and the result of the intermediate trapping strategy with full length AziA3, three possible intermediates were proposed (Figure 40). The absence of the ketoreductase domain and *aziA6* should

stall progression of the intermediate, trapping it in a keto-acid form. While the product trapped in the AziA3 reaction was the final epoxide, *trans*-acting tailoring domains could install the epoxide while the moiety is bound to the NRPS AziA3. In this case, the substrate could be the alkene or hydroxy compound. Incubation of AziA3ΔKR-6XHis with Biotin-Cys has been performed and confirmation of the results by LC-MS and synthesis of standards is underway.

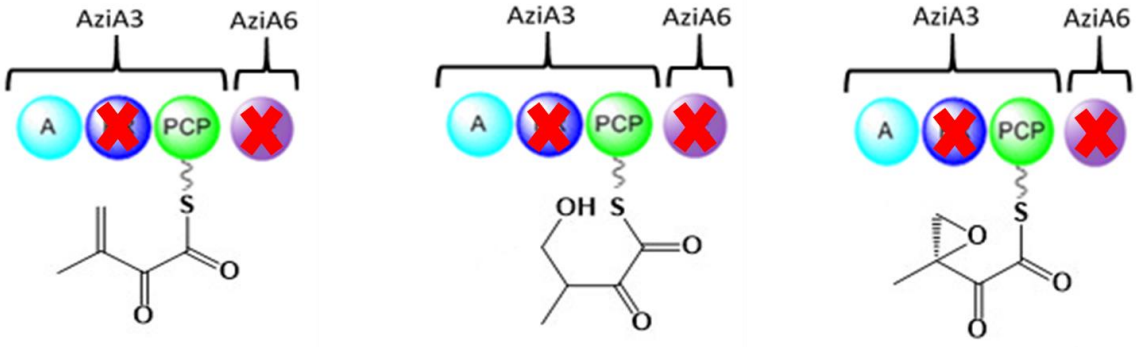


Figure 40 Proposed AziA3ΔKR Intermediates

Significance

Using the intermediate capture strategy, AziA3 was determined to function in the final steps of epoxide moiety construction. Many factors hindered the *in vitro* identification of the AziA3 substrate and product, and the intermediate capture strategy provides an alternative method for characterization. It was previously proposed that the epoxide was installed after condensation with the other moieties. This result disproves

that hypothesis and determined that the epoxide moiety is fully formed before condensation occurs.

In addition to the AziA3 characterization, this finding also confirms the ability of the intermediate capture strategy to be used with protein isolated directly from the metabolite producing organism. The ability to use this method without prior substrate identification opens the door for the probe to be used in characterization of cryptic metabolite pathways.

Materials and Methods

General Procedures

All chemicals were purchased from Sigma-Aldrich and all bacterial media were purchased from Becton-Dickinson. Molecular biology reagents were obtained from New England Biolabs. RT-PCR reagents were purchased from Thermo-Fisher. Reactions were monitored by thin layer chromatography (TLC) using silica gel plates. ^{1}H and $^{13}C\{^{1}H\}$ NMR spectra were recorded on either a Bruker Avance 500 equipped with a cryoprobe or a Bruker Ascend 400. Deuterated solvents were purchased from Cambridge Isotope Laboratories. Yields are reported for chromatographically pure compounds unless otherwise stated. Mass Spectra were obtained at the Laboratory for Biological Mass Spectrometry at the Department of Chemistry, Texas A&M University. A Phenomenex column (Prodigy 5 µm C18 150 Å, 150 x 4.6 mm) was used during HPLC and LC-MS purification and analysis. LC-MS was performed on a Thermo Scientific Ultimate 3000 UHPLC with a Thermo Fisher Scientific Q Exactive Focus mass

spectrometer. Mass analysis of synthesized compounds was performed on an ion trap mass spectrometer (LCQ-DECA, Thermo Fisher Scientific). HPLC purifications were performed on an Agilent 1260 HPLC with an automated fraction collector.

Bacterial Strains and Media

Streptomyces sahachiroi (NRRL 2485) was obtained from the American Type Culture Collection (ATCC). *Escherichia coli* S17 was used as the donor host for intergenic conjugal transformation. *E. coli* DH10B was used in the cloning of plasmids. 2xYT (16 g tryptone, 10 g yeast extract, 5 g NaCl in 1 L of water) was used for general liquid culturing of *S. sahachiroi.* GYM agar (4 g glucose, 4 g yeast extract, 2 g CaCO3, 10 g malt extract, 8 g agar in 1 L water adjusted to pH 6.8 with NaOH) was used for plating *S. sahachiroi* spores. MS agar (20 g agar, 20 g soya flour, 20 g mannitol, 2.03 g MgCl2 in 1 L water) was used in plating of the intergenic conjugal transformation. YEME (3 g yeast extract, 5 g peptone, 3 g malt extract, 10 g glucose, 170 g sucrose in 1 L water) was used in protein expression in *S. sahachiroi*.

Construction of Disruption Plasmid pKCAziA6

A 1078 bp upstream fragment and 1125 bp downstream fragment to *aziA6* was amplified from the genomic DNA of *S. sahachiroi*. The upstream fragment was amplified using Taq 2x Master mix and primers AziA6UF/AziA6UR (Table 5). The downstream fragment was amplified using Phusion High Fidelity Polymerase and primers AziA6DF/AziA6DR (Table 5). pGEM-Teasy cloning was utilized and the upstream and

downstream products were digested with XbaI/HindIII and XbaI/BamHI, respectively, to clone into the corresponding site within the pKC1139 plasmid to generate pKCAziA6.

Generation of ΔaziA6 Strain

Intergenic conjugal transformation was used to introduce the pKCAziA6 plasmid *to S. sahachiroi* to facilitate gene deletion by homologous recombination. The pKCAziA6 plasmid was transformed into *E. coli* S17 cells. An overnight culture of pKCAziA6/S17 cells was grown and washed twice with LB medium and resuspended in 600 µL of 2xYT medium. *S. sahachiroi* spores from 10 MS agar plates were collected using 2xYT medium. The collected spores were subjected to heat shock at 65 °C for 10 min and incubated at 37 °C for 3 h. Following incubation, the spores were collected and washed with 20 ml 2xYT two times and finally resuspended in 4 ml 2xYT. The recipient cells were mixed with the 600 µl of pKCAziA6/S17 donor cells and 800 µl of the mixture was plated on MS agar plates. The plates were incubated at 28 °C and overlaid with 1 ml of water containing nalidixic acid (500 µg) and apramycin (70 µg) to select for *S. sahachiroi* exconjugants after 24 h. Incubation continued for 10 days to allow for appearance of apramycin resistant exconjugants. Exconjugants were screened for several generations until the double crossover allelic exchange was detected by the presence of apramycin sensitivity (Figure 34).

The Δ*aziA6* strain was confirmed by PCR, Southern Blot, and RT-PCR (Figure 35). PCR was used to screen for the loss of the apramycin resistance gene (AprF/AprR primers) as well as the *aziA6* gene (AziA6F/AziA6R primers). Southern Blot was performed using

Roche DIG-High Prime DNA Labeling and Detection Starter Kit I. Digestion of genomic DNA was performed with SalI to generate characteristic DNA fragments. Taq 2x Master mix was used to generate the probe using primers AziA6-SBF and AziA6-SBR. Southern blot was performed according to manufacturer's instructions with Roche DNA Molecular Weight Marker VII, DIG Labeled. RT-PCR was performed using Superscript IV First Strand Synthesis Kit. *S. sahachiroi* RNA was isolated following previously published protocol.[101] Oligo$(dT)_{20}$ primer was used to generate cDNA. To evaluate expression of *aziA6, aziA3* (upstream), *aziA4* (downstream), and *aziB* (control), cDNA was used as the template for PCR and primer pairs AziA6RTF, AziA6RTR AziA3F, AziA3R, AziA4F, AziA4R, AziBF, AziBR were used respectively (Table 5). Primers were designed such that PCR products were ~200-400 bp. Thermocycling conditions were optimized to give a PCR band for the *aziB* control when analyzed by agarose gel electrophoresis. Optimized conditions utilized Phusion Master Mix with the following conditions: 30 s at 98 °C, 30 cycles of 10 s at 98 °C, 30 s at 60 °C, and 30 s at 72 °C, and a final extension of 10 min at 72 °C.

*Fermentation and Analysis of Δ*aziA6 strain*

A 1 cm^2 piece of GYM plate containing *S. sahachiroi* spores was used to inoculate a first stage culture of 100 ml of PS5 medium (0.5 g Starch, 0.5 g Pharmamedia per 100 ml water) in a 250 ml flask. The culture was incubated at 30 °C for 24 h at 250 rpm. The second stage culture was prepared by inoculating 2 x 600 ml PS5 with 25 ml of first stage culture. The culture was incubated at 30 °C for 24 h at 250 rpm. The fermenter was

prepared by autoclaving while containing 10 L of reduced PS5 (12.5 g Pharmamedia, 12.5 g Starch per 10 L water). Once the media reached room temperature after sterilization, the fermenter was inoculated with the 2 second stage cultures and the fermenter agitated at 300 rpm for 72 h.

Following fermentation, the cultures were centrifuged at 7,000 rpm at 4 °C. The cell pellets were discarded, and the medium extracted with an equal volume of dichloromethane. The organic layer was collected, dried over anhydrous magnesium sulfate, and concentrated in vacuo. The resulting crude extract was stored under diethyl ether at -80 °C. Crude extracts were analyzed by LC-MS. The sample was monitored by UV at $\lambda = 254$ nm. The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min.

Construction of Overexpression Plasmid pSETAziA3

Oligonucleotide primers pSETAziA3F and pSETAziA3R were used to amplify *aziA3* from *S. sahachiroi* genomic DNA. The primers incorporated restriction sites (XbaI and BglII) for cloning as well as a polyhistidine tag. PCR was performed with Phusion Master Mix polymerase and resulted in a PCR product of 3792 bp. The PCR product and pSET152 plasmid were digested using XbaI and BglII and ligated together using T4 DNA ligase.

*Generation of pSETAziA3Δ*aziA6* Strain

Intergenic conjugal transformation was used to introduce the pSETAziA3 plasmid to the Δ*aziA6* strain. The pSETAziA3 plasmid was transformed into *E. coli* S17 cells. An overnight culture of pSETAziA3/S17 cells was grown and washed twice with LB medium and resuspended in 600 µL of 2xYT medium. *S. sahachiroi* spores from 10 GYM agar plates were collected using 2xYT medium. The collected spores were subjected to heat shock at 65 °C for 10 min and incubated at 37 °C for 3 h. Following incubation, the spores were collected and washed with 20 ml 2xYT two times and finally resuspended in 4 ml 2xYT. The recipient cells were mixed with the 600 µl of pKCAziA6/S17 donor cells and 800 µl of the mixture was plated on ISP4 agar plates. The plates were incubated at 28 °C and overlaid with 1 ml of water containing nalidixic acid (500 µg) and apramycin (70 µg) to select for *S. sahachiroi* exconjugants after 24 h. Incubation continued for 8 days to allow for appearance of apramycin resistant exconjugants. The pSET152 backbone of the pSETAziA3 plasmid facilitated integration into the genome of *S. sahachiroi* generating a stable apramycin resistant strain.

Overexpression and Purification of AziA3

A starter culture of *S. sahAziA3*Δ*aziA6* (100 ml) in 2xYT medium was grown for 24 h at 28 °C. Protein expression was performed in YEME medium supplemented with 1% starter culture and 50 µg/ml apramycin. Cultures were grown at 28 °C for 120 h. The cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM β-mercaptoethanol (BME), 10% glycerol) with 5 mM

imidazole and 1 mM phenylmethylsulfonyl fluoride (PMSF). The cells were disrupted by sonication and cell debris removed by centrifugation. AziA3 was purified from the crude cell lysate on a HisTrap HP 5 mL column (GE Healthcare Life Sciences). Pure protein was obtained in elution fractions containing 100 mM imidazole and 250 mM imidazole as observed by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS PAGE).

Capture Strategy with AziA3

AziA3 was desalted and buffer exchanged into reaction buffer (50 mM potassium phosphate pH 7.5, 20% glycerol) with 8 M guanidine hydrochloride and 1 mM TCEP. Biotin-Cys was added in 10:1 molar excess to AziA3 and incubated at 4 °C for 16 h. Biotinylated compounds were purified from the crude reaction mixture using Roche Streptavidin Mutein Matrix following the manufacturer's instructions. Eluted fractions were analyzed by LC-MS. The sample was monitored by UV at $\lambda = 254$ nm. The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min. The reaction was compared to a control lacking AziA3 as well as a synthetic standard of cyclized product Biotin-Cys-epoxy acid.

Synthesis of Biotin-Cys-epoxy acid

Epoxy acid precursor (2-(2-methyloxiran-2-yl)-2-oxoacetic acid) was synthesized following a previously established protocol.[78] HRMS (ESI) m/z $[M-H]^-$ calcd for $C_5H_7O_4$; 131.0344, found 131.0336. ^1H NMR (400 MHz, D_2O) δ 1.24 (s, 3H), 2.72 (d, J = 4.12 Hz, 1H), 2.84 (d, J = 4.12, 1H), 4.23 (s, 1H); $^{13}C\{^1H\}$ NMR (400 MHz, D_2O) δ 17.9, 52.3, 57.9, 75.2, 170.4.

Epoxy acid precursor (2-(2-methyloxiran-2-yl)-2-oxoacetic acid) (19.8 mg) was dissolved in DMF and stirred at 4 °C. Biotin-Cys (80 mg) and EDC (40 mg) were added to the reaction and the reaction stirred for 12 h at room temperature. The solvent was removed in vacuo and cyclized product Biotin-Cys-epoxy acid purified by HPLC (29 mg, 39.2%). The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min. For comparison to the product of the Biotin-Cys + AziA3 reaction, cyclized product Biotin-Cys-epoxy acid purified was incubated in reaction buffer for 12 h and the reaction was lyophilized, purified with streptavidin, and compared to the reaction of Biotin-Cys with AziA3. HRMS (ESI) m/z $[M+H]^+$ calcd for $C_{21}H_{36}N_5O_6S_2$; 518.2107, found 518.2102. NMR Appendix Figure 26-27.

Construction of Disruption Plasmid pKCAziA3ΔKR

A 1.1 Kb upstream fragment and 1.2 Kb downstream fragment to the KR domain of AziA3 was amplified from the genomic DNA of *S. sahachiroi*. The fragments were amplified using Phusion Master mix and primers AziA3-KRUF/AziA3-KRUR and AziA3-KRDF/AziA3-KRDR, respectively (Table 5). pGEM-Teasy cloning was utilized and the upstream and downstream products were digested with EcoRI/XbaI and XbaI/HindIII, respectively, to clone into the corresponding site within the pKC1139 plasmid to generate pKCAziA3ΔKR.

Generation of aziA3ΔKR strain

Intergenic conjugal transformation was used to introduce the pKCAziA3ΔKR plasmid to *S. sahachiroi* to facilitate gene deletion by homologous recombination. The pKCAziA3ΔKR plasmid was transformed into *E. coli* S17 cells. An overnight culture of pKCAziA3ΔKR/S17 cells was grown and washed twice with LB medium and resuspended in 600 µL of 2xYT medium. *S. sahachiroi* spores from 10 GYM agar plates were collected using 2xYT medium. The collected spores were subjected to heat shock at 65 °C for 10 min and incubated at 37 °C for 3 h. Following incubation, the spores were collected and washed with 20 ml 2xYT two times and finally resuspended in 4 ml 2xYT. The recipient cells were mixed with the 600 µl of pKCAziA3ΔKR/S17 donor cells and 800 µl of the mixture was plated on ISP4 agar plates. The plates were incubated at 28 °C and overlaid with 1 ml of water containing nalidixic acid (500 µg) and apramycin (70 µg) to select for *S. sahachiroi* exconjugants after 24 h. Incubation

continued for 10 days to allow for appearance of apramycin resistant exconjugants. Exconjugants were screened for several generations until the double crossover allelic exchange was detected by the presence of apramycin sensitivity (Figure 39).

The AziA3ΔKR strain was confirmed by PCR and RT-PCR (Figure 39). PCR was used to screen for the loss of the apramycin resistance gene (AprF/AprR primers) as well as the KR domain of AziA3 (AziA3-2/AziA3R primers). RT-PCR was performed using Superscript IV First Strand Synthesis Kit. *S. sahachiroi* RNA was isolated following previously published protocol.[101] Oligo$(dT)_{20}$ primer was used to generate cDNA. To evaluate expression of *aziA6*, *aziA3-PCP domain* (upstream), *aziA3-A domain* (downstream), and *aziB* (control), cDNA was used as the template for PCR and primer pairs AziA6RTF, AziA6RTR AziA3PCPF, AziA3R, AziA3AF, AziA3AR, AziBF, AziBR were used respectively (Table 5). Primers were designed such that PCR products were ~200-400 bp. Thermocycling conditions were optimized to give a PCR band for the *aziB* control when analyzed by agarose gel electrophoresis. Optimized conditions utilized Taq Master Mix with the following conditions: 30 s at 98 °C, 30 cycles of 10 s at 98 °C, 30 s at 55 °C, and 30 s at 72 °C, and a final extension of 10 min at 72 °C.

Fermentation and Analysis of aziA3ΔKR strain

A 1 cm^2 piece of GYM plate containing *S. sahachiroi* spores was used to inoculate a first stage culture of 100 ml of PS5 medium (0.5 g Starch, 0.5 g Pharmamedia per 100 ml water) in a 250 ml flask. The culture was incubated at 30 °C for 24 h at 250 rpm. The second stage culture was prepared by inoculating 2 x 600 ml PS5 with 25 ml of first

stage culture. The culture was incubated at 30 °C for 24 h at 250 rpm. The fermenter was prepared by autoclaving while containing 10 L of reduced PS5 (12.5 g Pharmamedia, 12.5 g Starch per 10 L water). Once the media reached room temperature after sterilization, the fermenter was inoculated with the 2 second stage cultures and the fermenter agitated at 300 rpm for 72 h.

Following fermentation, the cultures were centrifuged at 7,000 rpm at 4 °C. The cell pellets were discarded, and the medium extracted with an equal volume of dichloromethane. The organic layer was collected, dried over anhydrous magnesium sulfate, and concentrated in vacuo. The resulting crude extract was stored under diethyl ether at -80 °C. Crude extracts were analyzed by LC-MS. The sample was monitored by UV at $\lambda = 254$ nm. The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min.

Construction of Overexpression Plasmid pSETAziA3ΔKR

Oligonucleotide primers pSETAziA3-AF and pSETAziA3-AR were used to amplify the adenylation domain of *aziA3* from *S. sahachiroi* genomic DNA. The primers incorporated restriction sites (BglII and HindIII) for cloning as well as a ribosome binding site. PCR was performed with Phusion Master Mix polymerase and resulted in a PCR product of 2638 bp. Oligonucleotide primers pSETAziA3-PCPF and pSETAziA3-

PCPR were used to amplify the peptide carrier protein domain of *aziA3* from *S. sahachiroi* genomic DNA. The primers incorporated restriction sites (HindIII and XbaI) for cloning as well as a polyhistidine tag. PCR was performed with Phusion Master Mix polymerase and resulted in a PCR product of 648 bp. The PCR products were digested with their respective restriction enzymes and pSET152 plasmid was digested using XbaI and BglII. All three pieces were then ligated together using T4 DNA ligase.

Generation of pSETAziA3ΔKRΔaziA6

Intergenic conjugal transformation was used to introduce the pSETAziA3ΔKR plasmid to the Δ*aziA6* strain. The pSETAziA3ΔKR plasmid was transformed into *E. coli* S17 cells. An overnight culture of pSETAziA3ΔKR/S17 cells was grown and washed twice with LB medium and resuspended in 600 µL of 2xYT medium. *S. sahachiroi* spores from 10 GYM agar plates were collected using 2xYT medium. The collected spores were subjected to heat shock at 65 °C for 10 min and incubated at 37 °C for 3 h. Following incubation, the spores were collected and washed with 20 ml 2xYT two times and finally resuspended in 4 ml 2xYT. The recipient cells were mixed with the 600 µl of pKCAziA6/S17 donor cells and 800 µl of the mixture was plated on ISP4 agar plates. The plates were incubated at 28 °C and overlaid with 1 ml of water containing nalidixic acid (500 µg) and apramycin (70 µg) to select for *S. sahachiroi* exconjugants after 24 h. Incubation continued for 8 days to allow for appearance of apramycin resistant exconjugants. The pSET152 backbone of the pSETAziA3ΔKR plasmid facilitated

integration into the genome of *S. sahachiroi* generating a stable apramycin resistant strain.

Overexpression and Purification of AziA3ΔKR

A starter culture of *S. sahAziA3ΔKR* (100 ml) in 2xYT medium was grown for 24 h at 28 °C. Protein expression was performed in YEME medium supplemented with 1% starter culture and 50 µg/ml apramycin. Cultures were grown at 28 °C for 120 h. The cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM β-mercaptoethanol (BME), 10% glycerol) with 5 mM imidazole and 1 mM phenylmethylsulfonyl fluoride (PMSF). The cells were disrupted by sonication and cell debris removed by centrifugation. AziA3 was purified from the crude cell lysate on a HisTrap HP 5 mL column (GE Healthcare Life Sciences). Pure protein was obtained in elution fractions containing 100 mM imidazole and 250 mM imidazole as observed by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS PAGE).

Capture Strategy with AziA3ΔKR

AziA3ΔKR was desalted and buffer exchanged into reaction buffer (50 mM potassium phosphate pH 7.5, 20% glycerol) with 8 M guanidine hydrochloride and 1 mM TCEP. Biotin-Cys was added in 10:1 molar excess to AziA3ΔKR and incubated at 4 °C for 16 h. Biotinylated compounds were purified from the crude reaction mixture using Roche Streptavidin Mutein Matrix following the manufacturer's instructions. Eluted fractions

were analyzed by LC-MS. The sample was monitored by UV at λ = 254 nm. The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min.

Table 4 Strains and Plasmids used in Chapter IV

Strain or plasmid	**Characteristic**	**Source**
S. sahachiroi	Azinomycin producing organism	ATCC
S.sah Δ*aziA6*	*S. sahachiroi* lacking *aziA6*	This study
*S. sahAziA3*Δ*aziA6*	*S.sah*Δ*aziA6* with overexpression of *aziA3*	This study
*S.sah AziA3*Δ*KR*	*S. sahachiroi* lacking KR domain of AziA3	This study
*S. sahAziA3*Δ*KR*Δ*aziA6*	*S.sah*Δ*aziA6* with overexpression of *aziA3*Δ*KR*	This study
pKCAziA6	Plasmid for disruption of *aziA6*	This study
pSETAziA3	Plasmid for overexpression of His-AziA3 in *S. sah*	This study
pKCAziA3ΔKR	Plasmid for disruption of KR domain of *aziA3*	This study
pSETAziA3ΔKR	Plasmid for overexpression of His-AziA3ΔKR in *S. sahachiroi*	This study

Table 5 Primers used in Chapter IV

Primer	Sequence (5'-3')
AziA6UF	TCAAGCTTCACACGAGCGAGAACGCG
AziA6UR	GTCTAGAGGAAGGACCTCTTCGTGA
AziA6DF	ATCTAGACGTGCTCAACGCGTCCCC
AziA6DR	CAGGATCCCGAGAACTATCCCGACCT
AziA6SBF	ATCGTCAACCTGCTCACCTC
AziA6SBR	GCCAGGAACACGTCCCGG
AziA6RTF	TTCCGATCACCACGGTTCG
AziA6RTR	TGAGTCCGACGACGAAGTAG
pSETAziA3F	TCAGATCTAGGAGGCTCTTCGTGACCACCGCCACCAGC
pSETAziA3R	TCTAGAAATTCAGTGGTGGTGGTGGTGGCCGGCGATCCAGGCCGCCAG
AprF	CGAATGGCGAAAAGCCGAGC
AprR	GCATCGCATTCTTCGCATCC
AziA6F	GTCAAGCTTATGGACTCGCCCTGGGT
AziA6R	GAGCTCGAGCGAAGAGGTCCTTCCG
AziA3F	TGGTGACCTCACGGATAGAG
AziA3R	CCCAGGCTGAAGAACGAGT
AziA4F	CCGACGAGGAGAACAGCAC
AziA4R	GTGGGAGTGGCGAAGAAGTC
AziBF	CGCGGGGGCGCAGAGCGCCG

Table 5 Continued

Primer	Sequence (5'-3')
AziBR	CCCGGCCGCGGCCGCGACGG
AziA3-KRUF	TGAATTCGCGGACCGCCACGGATA
AziA3-KRUR	ATCTAGACCGGGGATGAACCAGGGC
AziA3-KRDF	ATCTAGAGCCCCGGGGCAGCACATC
AziA3-KRDR	TCAAGCTTTGGACCAGGAAGGAGAGG
AziA3-2	ATGGTGATCGTCAACGGGG
AziA6RTF	TTCCGATCACCACGGTTCG
AziA6RTR	TGAGTCCGACGACGAAGTAG
AziA3PCPF	GGATGAACCAGGGCAGT
AziA3AF	TTCGTCACCGCCTGGTG
AziA3AR	CATCCAGTTGAGCGAGACGT
pSETAziA3-AR	TCAAGCTTGCCCCGGGGCAGCACATCAGC
pSETAziA3-PCPF	CCAAGCTTCCGGGGATGAACCAGGGCAGT
pSETAziA3-PCPR	ATCTAGAAATTCAGTGGTGGTGGTGGTGGCCGGCGATCCTGGCCGCCAG

CHAPTER V

PROBING THE LIMITS OF THE NAPHTHOATE PATHWAY

Introduction

As discussed in Chapter I and III, it was previously determined *in vitro* that the PKS AziB is responsible for 5-methylnaphthoic acid production when the thioesterase AziG is present. In the absence of AziG, AziB produces a truncated product of 2-methylbenzoic acid (Figure 41).[80]

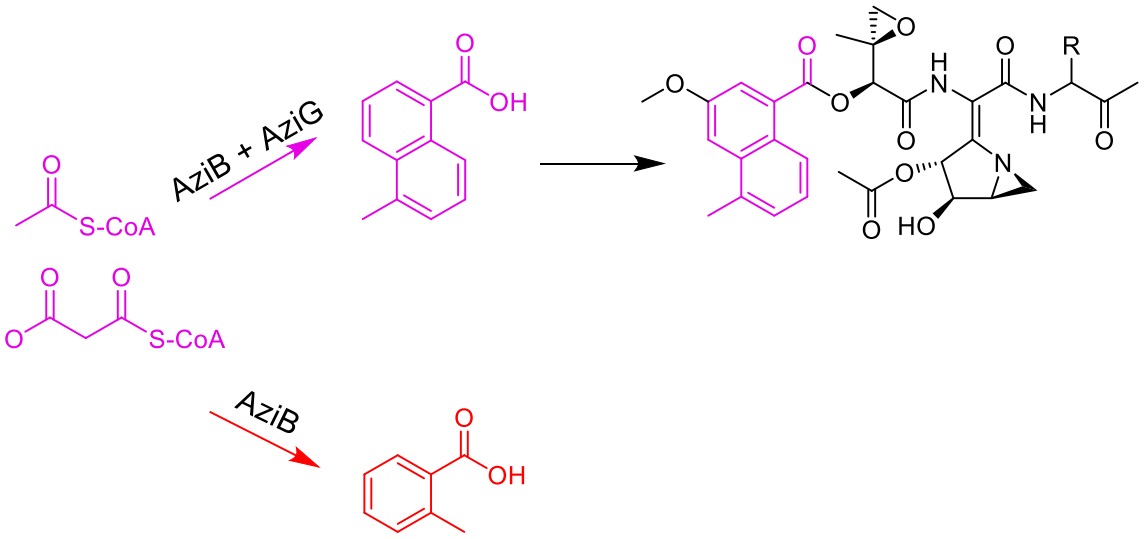


Figure 41 Role of AziB and AziG in Naphthoate Production

The naphthoate moiety plays an important role in the activity of azinomycin. Studies have shown the hydrophobicity of the naphthoate moiety allows for interaction with the DNA bases, enhancing the activity of azinomycin.[74] Additionally, activity was detected in derivatives containing only the naphthoate and epoxide moiety, commonly

referred to as the Left-Half of azinomycin.[73] Improving production and our understanding of the naphthoate moiety allows for cancer therapies to be developed using synthetic biology techniques for development of analogues such as antibody-drug conjugants. In this chapter, the *in vitro* expression and post-translational modification efficiency was improved by the separation of AziB domains. In addition, the AziB/AziG system was investigated *in vivo*.

Preliminary Research

AziB was previously heterologously expressed in *E. coli* and post-translationally modified by a promiscuous PPTase, Svp, from *Bacillus subtilis*. Both expression and PPTase modification are very low, resulting in low yield for 2-methylbenzoic acid and 5-methylnaphthoic acid products (Figure 42).[80] Two PPTases, Non-Fatty Acid Synthase

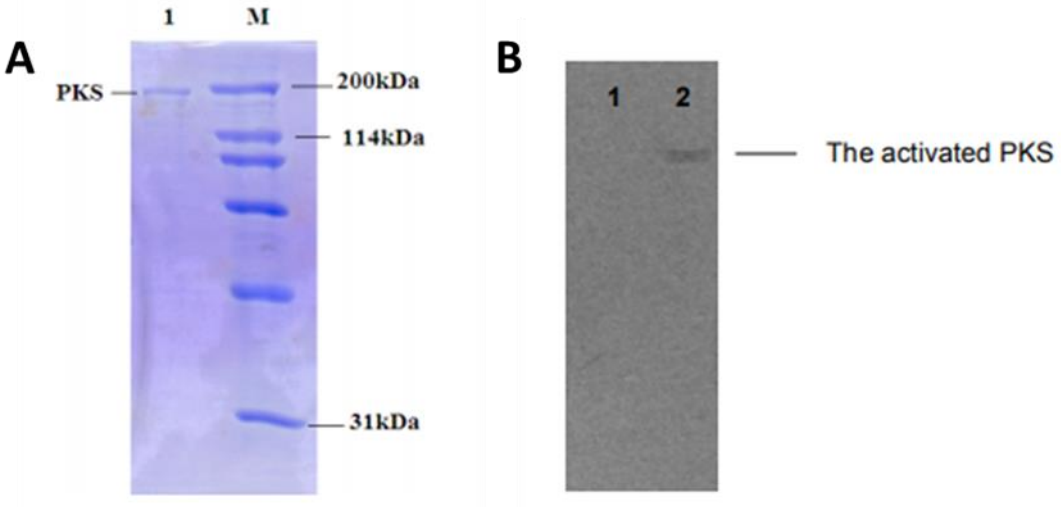


Figure 42 Expression and Phosphopantetheinylation of AziB[80]
A. SDS-PAGE with purified AziB. **B**. Phosphopantetheinylation of AziB with ^{3}H-pantetheine-CoA as detected by autoradiography: Lane 1: Negative control (no PPTase added), Lane 2: AziB (PKS) + Svp.

Associated (NFAS) PPTase and Fatty Acid Synthase Associated (FAS) PPTase, were identified in the *S. sahachiroi* genome and were cloned for expression in *E. coli*, but their expression was too low for regular use (Dr. Shogo Mori and Dr. Huitu Zhang).

Results and Discussion

Domain Separation and PPTase Assays

To increase yield, stability, and post-translational modification efficiency, the ACP domain was separated from the catalytic domains (KS-AT-DH-KR) of AziB. In order to retain the activity of AziB, the Udwary-Merski algorithm (UMA) was used to identify flexible linker sites for separation (Figure 43). A low UMA score indicates the amino acid site is in a short chain with low conservation in structure and sequence and should contain hydrophilic side-chains.[102] These factors indicate a site is a good candidate for separation as evolutionarily, sites required for function should be highly conserved.

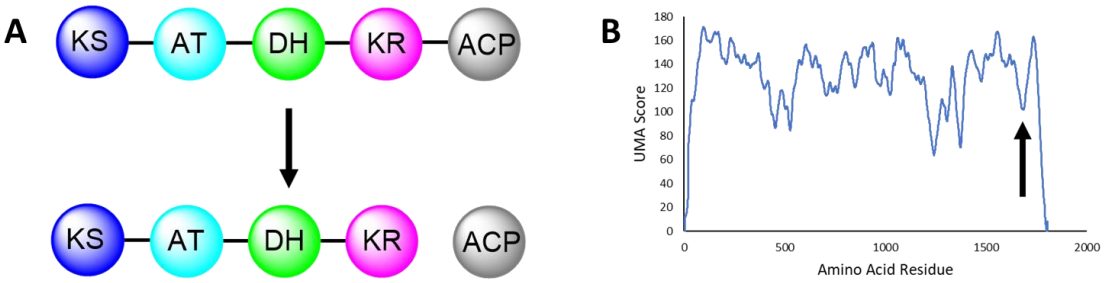


Figure 43 AziB Domain Separation and UMA
A. Domain Separation. **B**. UMA Graph for AziB. Arrow indicates cut site.

The separated domains were cloned into pET overexpression vectors for expression in *E. coli*. Separation of the ACP domain from KS-AT-DH-KR domains resulted in a ten-fold expression increase as compared to full length AziB. Stability and oligomeric state of AziB and KS-AT-DH-KR was determined using Native Mass Spectrometry (Figure 44). Full length AziB was unstable and heterogeneous and oligomeric state could not be determined (Figure 44A). KS-AT-DH-KR domains of AziB exhibited greater stability and was determined to be a dimer (Figure 44B), which is congruent with homologous PKS proteins.[103]

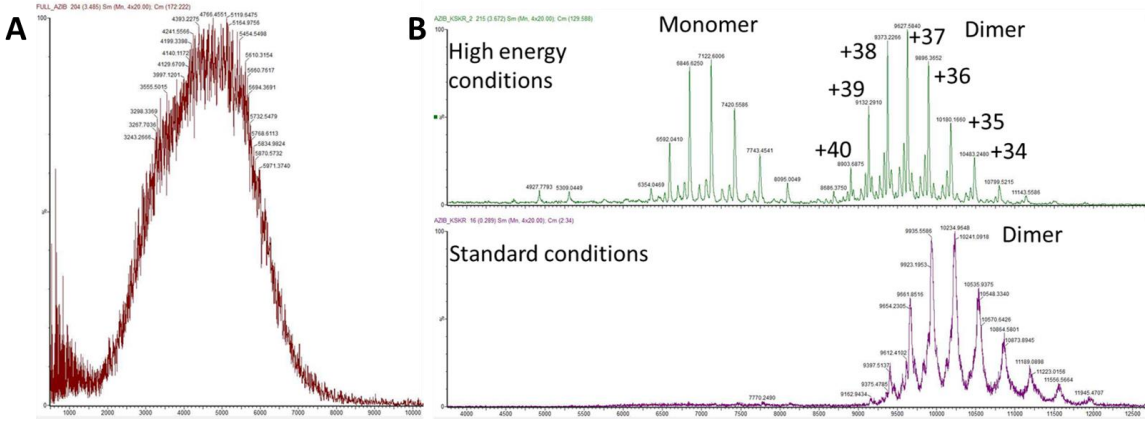


Figure 44 Native Mass Spec of Full AziB and KS-AT-DH-KR
A. Native Mass Spectrum of Full length AziB. **B**. Native Mass Spectrum of KS-AT-DH-KR domains of AziB. Native MS shows a dimer for KS-AT-DH-KR. Charge states denoted for dimer.

To address the low levels of post-translational modification of the ACP domain, the two PPTases found in *S. sahachiroi* were codon optimized by Genscript for expression in *E. coli* and cloned into pET expression vectors. Once overexpression and purification was achieved, a modified ELISA was used to determine post-translational

modification levels with full length AziB, the separated ACP domain of AziB, and four PPTases (Svp, Sfp, NFAS PPTase, and FAS PPTase).[104] To determine modification efficiency of the four PPTases toward full length AziB and the separated ACP domain of AziB, a Biotin-Coenzyme A derivative was synthesized and used to label the carrier protein domain (Appendix Figure 28). A streptavidin coated plate binds to the biotin labeled carrier protein. Unlabeled protein is removed with washes. Both full length AziB and ACP contain a 6X-His-tag for purification. To quantify the amount of biotin labeled protein in each reaction, a 6X-His Monoclonal Antibody – Horse Radish Peroxidase Conjugant was used with 3,3′,5,5′-Tetramethylbenzidine (TMB) to produce a color

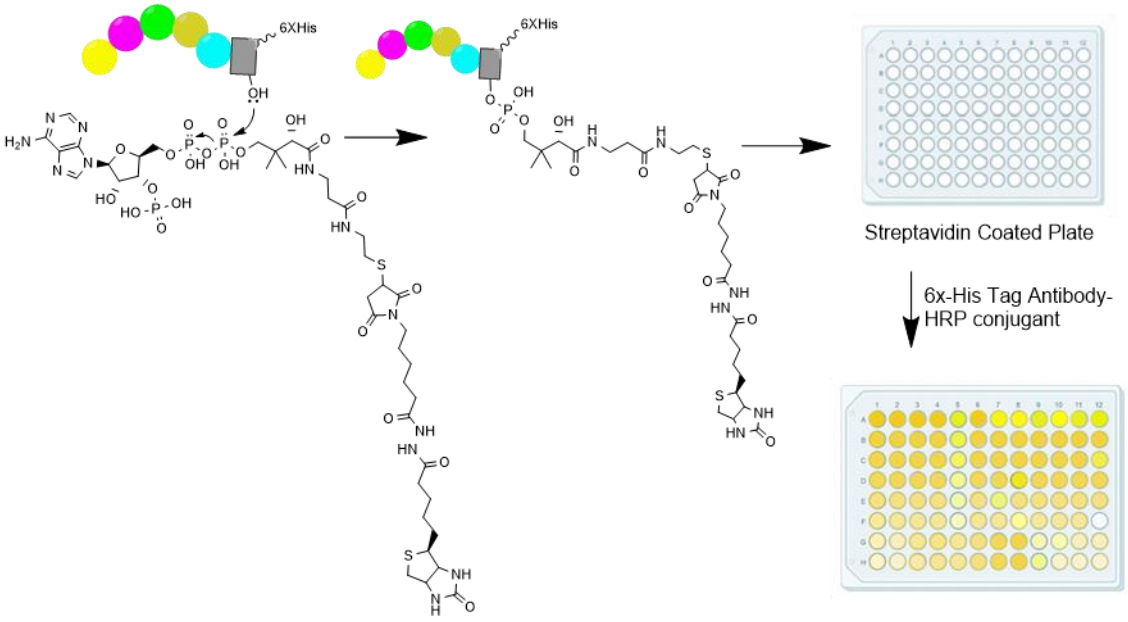


Figure 45 Biotin-CoA PPTase Activity Assay

change, which was quenched with acid and quantified using UV. The color change will only occur if the protein was labeled with Biotin-Coenzyme A and has a 6X-His-tag present to bind the antibody-HRP conjugant. The assay showed that separated ACP was labeled more efficiently than full length AziB, and that the promiscuous PPTase Svp was still the most efficient at labeling ACP and AziB (Appendix Figure 29). AziB and the ACP domain were able to be modified by the FAS PPTase and another promiscuous PPTase, Sfp. This assay was later repeated and optimized by Brendan Foley.

Based on the initial post-translational modification assay results, the production of 2-methylbenzoic acid by full length AziB and the separated domains was tested. Each ACP containing protein (full length AziB and separated ACP domain) was modified by either Svp or FAS PPTase and then incubated with substrates. Production of 2-methylbenzoic acid was detected by LC-MS (Figure 46). The results showed the separated domains, when incubated together, can still produce the desired 2-methylbenzoic acid product (Figure 46A and 46D). ACP and AziB labeled by Svp produced more 2-methylbenzoic acid then those labeled by FAS PPTase, which is consistent with the ELISA assay results which show Svp post-translationally modifies AziB and the ACP domain more efficiently than FAS PPTase. Both AziB and the ACP domain labeled by FAS PPTase produced 2-methylbenzoic acid in similar amounts (Figure 46C and 46D). This again agrees with the PPTase assay results as both full length AziB and the ACP domain were labeled at similar amounts by FAS PPTase.

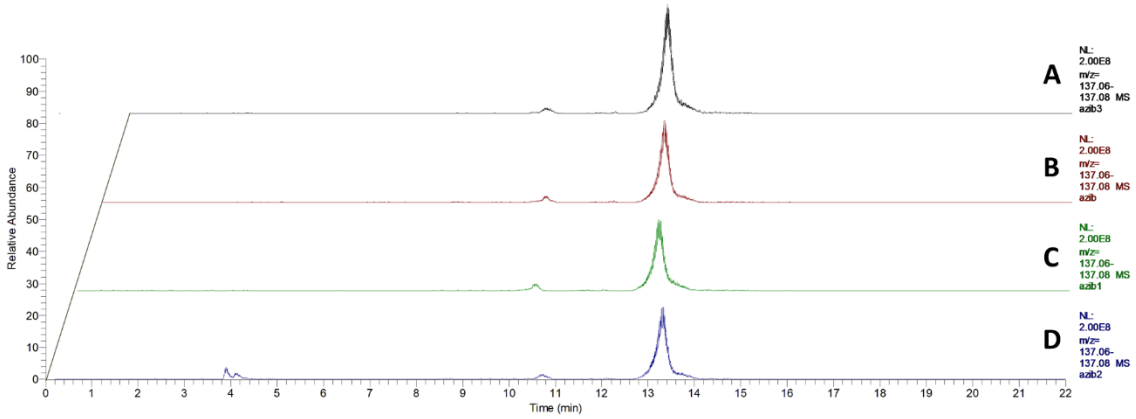


Figure 46 2-methylbenzoic Acid Production by AziB Domains and PPTases
EIC for 2-methylbenzoic acid $[M+H]^+$ = 137.07. **A**: ACP + KS-AT-DH-KR + Svp. **B**: AziB + Svp. **C**: AziB + FAS PPTase. **D**: ACP + KS-AT-DH-KR + FAS PPTase

Investigation of AziG in vivo

To validate the AziB/AziG *in vitro* reconstitution results, *aziG* was disrupted in *S. sahachiroi* by homologous recombination. Regions upstream and downstream of *aziG* were cloned into the disruption plasmid pKC1139 to generate the plasmid pKCAziG. The disruption plasmid was transformed into *E. coli* S17 for conjugal transformation into *S. sahachiroi*. Incorporation of the plasmid was screened for by using the apramycin resistance gene (*apr*) located on the pKC1139 plasmid back bone. Subsequent screening was performed to identify double crossovers which have lost *apr* and *aziG.* To confirm loss of *apr* and *aziG*, PCR and sequencing were used. Additionally, RT-PCR was employed to evaluate gene expression of the genes surrounding *aziG* (Figure 47).

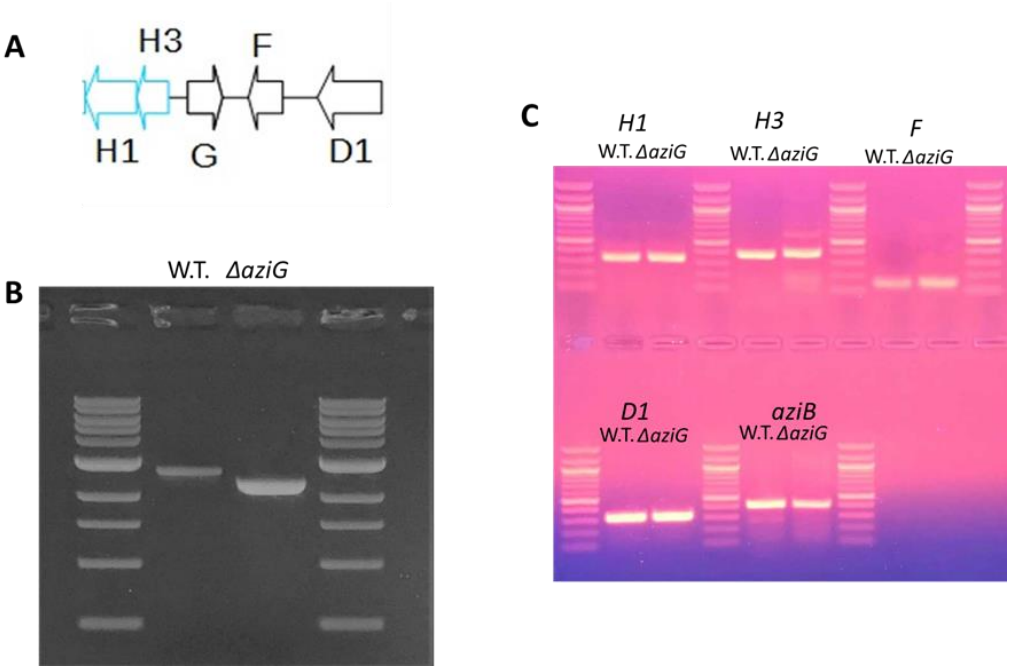


Figure 47 Confirmation of Δ*aziG* Strain
A. Azinomycin Gene Cluster near *aziG*. **B**. PCR confirming *aziG* disruption. **C**. RT-PCR for expression of surrounding genes

Once the disruption of *aziG* was confirmed, fermentation of the strain was performed to confirm the loss of naphthoate and azinomycin production and to identify the buildup of intermediates or byproducts resulting from the loss of azinomycin production. Of interest is the potential production of the 2-methylbenzoic acid truncated product. Comparison of Δ*aziG* and wild type extracts showed loss of naphthoate and azinomycin production in the Δ*aziG* strain, confirming AziG's role in naphthoate production (Figure 48). Truncated products based on the 2-methylbenzoic acid moiety were not detected. Due to the absence of AziG, the 2-methylbenzoic acid may be bound to the AziB carrier protein domain as a thioester. Further work is underway to explore this possibility.

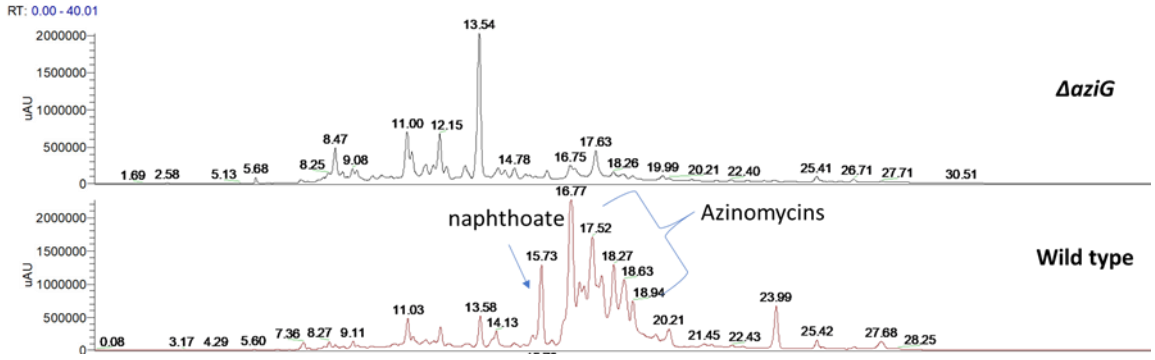


Figure 48 Δ*aziG* vs Wild Type LC-MS
UV λ = 254 nm trace

To explore the production of an azinomycin analogue that contains a benzoic moiety instead of a naphthoic one, 2-methylbenzoic acid derivatives were fed to the Δ*aziG* strain (Figure 49). Feeding of 2-methylbenzoic acid will also probe the substrate acceptance of AziB1 and AziB2 by looking for conversion to a methoxy-benzoic derivative. Feeding was performed during the fermentation process and extraction and analysis followed the same procedures for wild type *S. sahachiroi*. LC-MS analysis and

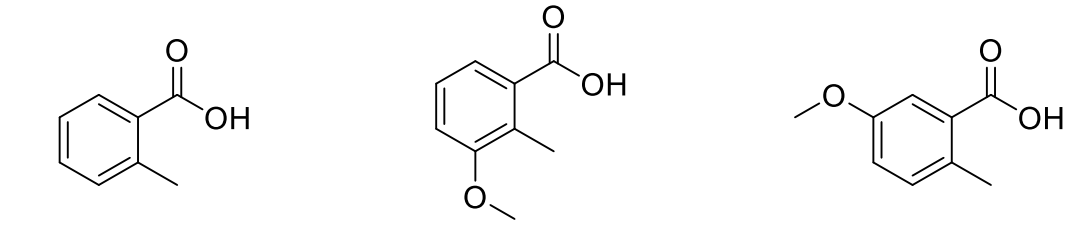


Figure 49 Derivatives for Feeding Study

EICs for the expected benzoic-azinomycin derivatives revealed the incorporation of 3-methoxy-2-methylbenzoic acid and 5-methoxy-2-methylbenzoic acid, but not 2-

methylbenzoic acid (Figure 50, Appendix Figures 30-31). This result is consistent with the previous finding that the NRPS AziA1, which is proposed to activate the naphthoate moiety for future condensation with the epoxide moiety, accepts 3-methoxy-5-methylnaphthoic acid. Activity of AziA1 dramatically decreased if the methoxy was not present on the naphthoate substrate.[105] Additionally, methoxy-2-methylbenzoic acid was not detected in the 2-methylbenzoic acid feeding sample, indicating it was not accepted by AziB1 and AziB2 *in vivo*.

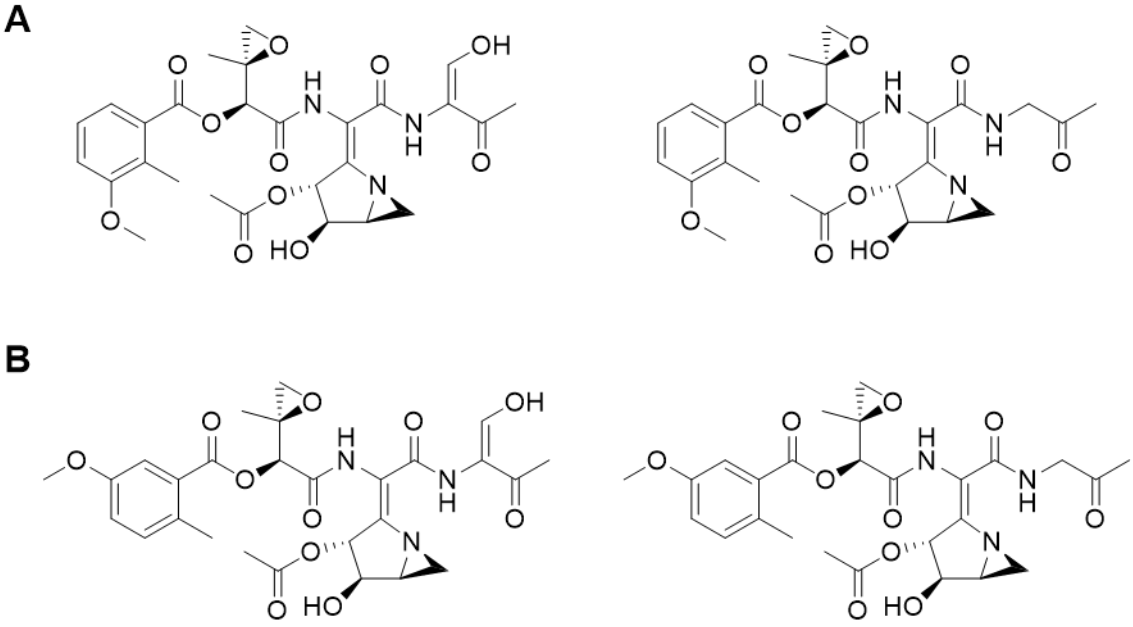


Figure 50 Benzoic-Azinomycin Derivatives Detected by LC-MS
A. 3-methoxy-2-methylbenzoic acid derivatives. **B**. 5-methoxy-2-methylbenzoic acid derivatives.

In addition to the benzoic-azinomycin derivatives, new compounds were detected in the 16 to 18 minute region for all feeding samples (Figure 51). While the UV 254 nm

shows potential new compounds, their identity could not be determined by the Mass Spectrum alone. Purification of the new 16 min to 18 min species will be performed to identify the compounds by NMR and using other mass spec ionization methods. Additionally, the benzoic acid rings showed poor solubility in the fermentation media. To overcome this, a liposomal encapsulation method is now being employed and feeding of the liposomally encapsulated derivatives will be repeated to increase yield of the benzoic-azinomycin derivatives. Another potential reason for lack of incorporation or low levels of incorporation is that the next protein in the biosynthetic pathway may prefer thioester substrates over the free carboxylic acid compounds. N-acetyl-cysteamine derivatives will be synthesized and fed to the strain to test this hypothesis.

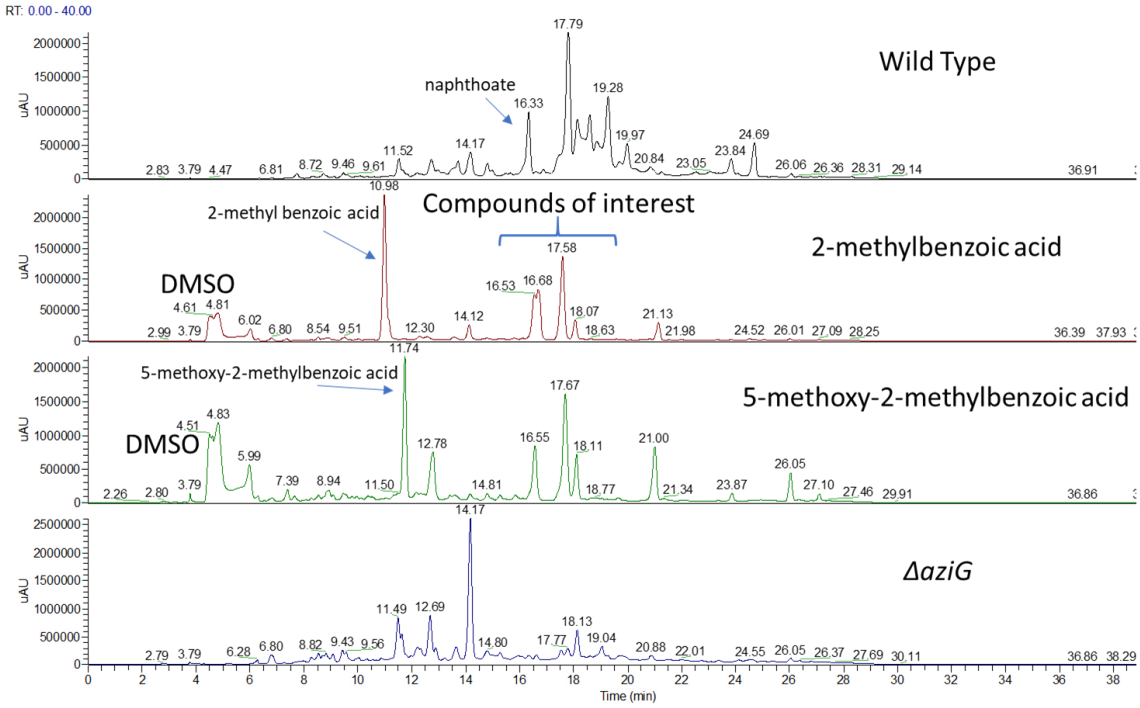


Figure 51 Δ*aziG* Feeding Study LC-MS Results

Significance

Separating the domains of a large multidomain protein can result in increased expression, stability, and opens the door for domain swapping between proteins. The separation of the catalytic domains of AziB from the acyl carrier protein domain increased expression in *E. coli*. This system also produced more homogenous and stable protein, which allows for crystal structure studies to be pursued in the future. The domain separation still produced active protein, resulting in the production of 2-methylbenzoic acid when incubated together with the other domain(s) and substrates. Additionally, optimization of the PPTase reaction conditions improved product yield and showed the ACP domain is more efficiently modified when separated from the rest of the large protein.

Understanding the biosynthesis of the naphthoate moiety *in vivo* allows for production of azinomycin analogues. The role of AziG in naphthoate production was confirmed by loss of naphthoate production in the Δ*aziG* strain. The Δ*aziG* strain can produce methoxy-benzoic-azinomycin derivatives when methoxy-benzoic compounds are supplemented in the fermentation process. Further work is needed to probe the breadth of the acceptance of benzoic or naphthoic based analogues.

Material and Methods

General Procedures

All chemicals from Sigma-Aldrich, all bacterial media were purchased from Becton-Dickinson, and ELISA and PCR reagents were purchased from ThermoFisher Scientific.

Restriction enzymes and molecular biology reagents were purchased from New England Biolabs. RT-PCR reagents were purchased from Thermo-Fisher. Mass Spectra were obtained at the Laboratory for Biological Mass Spectrometry or at the Department of Chemistry, Texas A&M University. A Phenomenex column (Prodigy 5 µm C18 150 Å, 150 x 4.6 mm) was used during HPLC and LC-MS purification and analysis. LC-MS was performed on a Thermo Scientific Ultimate 3000 UHPLC with a Thermo Fisher Scientific Q Exactive Focus mass spectrometer. Protein mass spectrometry was recorded on an Agilent 6560 ion mobility Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA). Reactions were monitored by thin layer chromatography (TLC) using silica gel plates. ^{1}H and $^{13}C\{^{1}H\}$ NMR spectra were recorded on either a Bruker Avance 500 equipped with a cryoprobe or a Bruker Ascend 400. Deuterated solvents were purchased from Cambridge Isotope Laboratories. Yields are reported for chromatographically pure compounds unless otherwise stated. HPLC purifications were performed on an Agilent 1260 HPLC with an automated fraction collector.

Bacterial Strains and Media

Escherichia coli BL21 (DE3) was used for *in vitro* protein expression. *E. coli* DH10B was used in the cloning of plasmids. *Streptomyces sahachiroi* (NRRL 2485) was obtained from the American Type Culture Collection (ATCC). *E. coli* S17 was used as the donor host for intergenic conjugal transformation. *E. coli* DH10B was used in the cloning of plasmids. 2xYT (16 g tryptone, 10 g yeast extract, 5 g NaCl in 1 L of water) was used for general liquid culturing of *S. sahachiroi.* GYM agar (4 g glucose, 4 g yeast

extract, 2 g CaCO3, 10 g malt extract, 8 g agar in 1 L water adjusted to pH 6.8 with NaOH) was used for plating *S. sahachiroi* spores. MS agar (20 g agar, 20 g soya flour, 20 g mannitol, 2.03 g MgCl2 in 1 L water) was used in plating of the intergenic conjugal transformation.

Cloning of Phosphopantetheinyl Transferases

The two PPTases located in the genomic DNA of *S. sahachiroi*, FAS PPTase and NFAS PPTase, were codon optimized for expression in *E. coli* by Genscript. They were cloned into a pET21 expression plasmid using NdeI and XhoI restriction enzymes. T4 DNA Ligase was used to ligate the final constructs.

Overexpression and Purification of PPTases

E. coli BL21 (DE3) containing a PPTase expression vector was grown in 5 ml of LB medium with 50 µg/ml ampicillin overnight. The culture was used to inoculate 1 L of LB medium and culture at 37 °C at 250 rpm until an OD_{600} of 0.6 was reached. Induction was performed by the addition of 1 ml of 1 M IPTG. The cultures were incubated at 16 °C for 24 h at 250 rpm. The cells were harvested by centrifugation at 6500 rpm and resuspended in buffer containing 20 mM potassium phosphate, 500 mM NaCl, 1 mM DTT, 5 mM imidazole, and 20% glycerol, pH 7.4. The cells were lysed by sonication, clarified using centrifugation to remove cellular debris. The supernatant containing the PPTase was purified using a HisTrap FF 5 ml column (GE Healthcare Life Sciences).

The purified protein was concentrated using a 10 kDa centrifugal ultrafiltration unit (Amicon Ultra centrifugal filter unit, Millipore).

Separation and Cloning of AziB Domains

The UMA program was used to identify cut sites to separate the domains of AziB without loss of function. Homologous sequences to AziB were located on BLASTp and a multiple sequence alignment was produced using CLUSTAL. The PredictProtein Server generated secondary structure predictions for the protein sequences in the multiple sequence alignment. UMA uses these files to generate a score for each amino acid in AziB. Amino acid regions with a low UMA score were chosen as cut sites.

The ACP domain of *aziB* was amplified by PCR using the pET24b-aziB vector as template DNA, One Taq 2X Master Mix with GC Buffer, and ACPF/ACPR primers. The thermocycles condition were as follows; initial denaturing (94°C for 30 sec), 30 cycles of denaturing (94°C for 30 sec), annealing (60°C for 30 sec), and extension (68°C for 30 sec), and a final extension (68°C for 5 min). The gel extracted PCR product was ligated into the intermediate pGEM-Teasy Vector. NdeI and XhoI were used to digest pET24b and ACP. The two digested fragments were ligated together using T4 DNA Ligase.

The ketosynthase-acyl transferase-dehydratase-ketoreductase (KS-AT-DH-KR) domains of *aziB* were amplified by PCR using pET24b-aziB vector as the template DNA, Platinum SuperFi Green PCR Master Mix, and KS-KRF/KS-KRR primers. The thermocycle conditions were as follows; initial denaturing (98°C for 30 sec), 30 cycles

of denaturing (98°C for 10 sec), annealing (62°C for 10 sec), and extension (72°C for 2 min), and a final extension (72°C for 5 min). The PCR product was isolated using PureLink gel purification. NdeI and XhoI were used to digest pET21a and the fragments ligated together using Gibson Assembly (NEB HiFi DNA Assembly Cloning Kit).

Overexpression of AziB Domains

E. coli BL21 (DE3) containing either the pET-24-acp or pET-21-ks-kr expression vector was grown in 5 ml of LB medium with 50 µg/ml antibiotic overnight. Ampicillin was used for pET-21-ks-kr and Kanamycin for pET-24-acp. The culture was used to inoculate 1 L of LB medium and culture at 37 °C at 250 rpm until an OD_{600} of 0.6 was reached. Induction was performed by the addition of 1 ml of 1 M IPTG. The cultures were incubated at 16 °C for 24 h at 250 rpm. The cells were harvested by centrifugation at 6500 rpm and resuspended in buffer containing 20 mM potassium phosphate, 500 mM NaCl, 1 mM DTT, 5 mM imidazole, and 20% glycerol, pH 7.4. The cells were lysed by sonication, clarified using centrifugation to remove cellular debris. The supernatant containing the PPTase was purified using a HisTrap FF 5 ml column (GE Healthcare Life Sciences). The purified protein was concentrated using a 10 kDa centrifugal ultrafiltration unit for ACP and a 50 kDa unit for KS-AT-DH-KR (Amicon Ultra centrifugal filter unit, Millipore).

Protein Mass Spectrometry

Protein Mass Spectrometry was recorded on an Agilent 6560 ion mobility Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA) in the Laganowsky Lab to determine stability and oligomeric state of the protein.

Characterization of PPTase Activity using Modified ELISA

To synthesize Biotin-Coenzyme A, 25 mg of Biotin-Maleimide was dissolved in 500 µl DMSO and added to a solution containing 45 mg Coenzyme A in 2 ml of Phosphate buffer pH 6.0. The reaction was stirred at room temperature overnight. The reaction was lyophilized and analyzed by LC-MS to confirm product formation (Appendix Figure 28).

To perform phosphopantetheinylation reactions, the four PPTase (Sfp, Svp, NFAS PPTase, and FAS PPTase) were purified along with AziB and the ACP domain of AziB. 100 µM Biotin-Coenzyme A, 1 µM PPTase, and 50 µM ACP containing protein were combined in 100 µl of 50 mM HEPES Buffer pH 7.5 with 10 mM $MgCl_2$. The reaction was incubated at 37 °C for 1 hr. A streptavidin coated plate was washed three times with 25 mM Tris, 150 mM NaCl pH 7.2 with 0.1% BSA and 0.05% Tween 20. Serial dilutions of the 100 µl PPTase reaction were made and 100 µl of each was added to one well and incubated for 2 hours at room temperature. The wells were washed three times. 100 µl of 6x-His Tag Monoclonal Antibody-Horse Radish Peroxidase conjugant was added to each well and incubated for 30 minutes. The wells were washed three times to remove excess antibody. 100 µl of 3,3',5,5'-tetramethylbenzidine (TMB) substrate was

added to each well and incubated for 30 minutes to develop color. 100 µl of 2 M sulfuric acid was added to stop the reaction and the UV was recorded at 450 nm.

Production of 2-methylbenzoic Acid by AziB and Separated Domains

Following purification, the ACP containing protein (Full length AziB or separated ACP domain) was incubated with a PPTase (Svp or FAS PPTase) and 4 mg of Coenzyme A for 1 h at 28 °C. After the incubation period, the post-translationally modified protein was incubated with 50 µM acetyl-Coenzyme A, 250 µM malonyl-Coenzyme A, 10 µM dihydronicotinamide-adenine dinucleotide phosphate (NADPH), and 1 mM dithiothreitol (DTT). For the reactions containing separated ACP domain, the KS-AT-DH-KR protein was added in a 1:1 ratio to ACP protein alongside the substrates to initiate the reaction. Reactions were incubated at 28 °C for 12 h. After 12 h, the reaction was quenched, and products hydrolyzed by the addition of 2 M NaOH to bring the pH to 12 for 15 min. Concentrated HCl was then added dropwise to bring mixture to pH 2. The reaction was extracted with ethyl acetate, the solvent was removed, and the products analyzed by LC-MS. The sample was monitored by UV at $\lambda = 254$ nm. The column was pre-equilibrated with 80% A (water, 0.1% formic acid) and 20% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-80% B-20%; 1 min, A-80% B-20%; 23 min, A-0% B-100%; 33 min, A-0% B-100%; 35 min, A-80% B-20%; 40 min, A-80% B-20%, flow rate 0.75 µL/min.

Generation of ΔaziG Strain

The plasmid pKCAziG was designed and constructed by Genscript using a 1026 bp region upstream and a 1215 bp region downstream of *aziG*. The homology arms were introduced using the EcoRI and HindIII sites of pKC1139. Intergenic conjugal transformation was used to introduce the pKCAziG plasmid to *S. sahachiroi* to facilitate gene deletion by homologous recombination. The pKCAziG plasmid was transformed into *E. coli* S17 cells. An overnight culture of pKCAziG/S17 cells was grown and washed twice with LB medium and resuspended in 600 µL of 2xYT medium. *S. sahachiroi* spores from 10 GYM agar plates were collected using 2xYT medium. The collected spores were subjected to heat shock at 65 °C for 10 min and incubated at 37 °C for 3 h. Following incubation, the spores were collected and washed with 20 ml 2xYT two times and finally resuspended in 4 ml 2xYT. The recipient cells were mixed with the 600 µl of pKCAziG/S17 donor cells and 800 µl of the mixture was plated on ISP4 plates. The plates were incubated at 28 °C and overlaid with 1 ml of water containing nalidixic acid (500 µg) and apramycin (70 µg) to select for *S. sahachiroi* exconjugants after 24 h. Incubation continued for 10 days to allow for appearance of apramycin resistant exconjugants. Exconjugants were screened for several generations until the double crossover allelic exchange was detected by the presence of apramycin sensitivity The Δ*aziG* strain was confirmed by PCR and RT-PCR (Figure 47). PCR was used to screen for the loss of the apramycin resistance gene (AprF/AprR primers) as well as the *aziG* gene (AziGUpF/AziGDownR primers). Amplification of the region surrounding *aziG* was achieved using Phusion Master Mix with the following conditions: 30 s at 98

°C, 30 cycles of 10 s at 98 °C, 30 s at 60 °C, and 60 s at 72 °C, and a final extension of 10 min at 72 °C. RT-PCR was performed using Superscript IV First Strand Synthesis Kit. *S. sahachiroi* RNA was isolated following previously published protocol.[101] Oligo(dT)$_{20}$ primer was used to generate cDNA. To evaluate expression of *aziH1, aziH3*, *aziF*, *aziD1*, and *aziB* (control), cDNA was used as the template for PCR and primer pairs AziAH1F, AziAH1R AziH3F, AziH3R, AziFF, AziFR, AziD1F, AziD1R, AziBF, and AziBR were used respectively (Table 9). Primers were designed such that PCR products were ~200-400 bp. Thermocycling conditions were optimized to give a PCR band for the *aziB* control when analyzed by agarose gel electrophoresis. Optimized conditions utilized Phusion Master Mix with the following conditions: 30 s at 98 °C, 30 cycles of 10 s at 98 °C, 30 s at 60 °C, and 30 s at 72 °C, and a final extension of 10 min at 72 °C.

Culture Conditions and Fermentation

A 1 cm^2 piece of GYM plate containing *S. sahachiroi* spores was used to inoculate a first stage culture of 100 ml of PS5 medium (0.5 g Starch, 0.5 g Pharmamedia per 100 ml water) in a 250 ml flask. The culture was incubated at 30 °C for 24 h at 250 rpm. The second stage culture was prepared by inoculating 2 x 600 ml PS5 with 25 ml of first stage culture. The culture was incubated at 30 °C for 24 h at 250 rpm. The fermenter was prepared by autoclaving while containing 10 L of reduced PS5 (12.5 g Pharmamedia, 12.5 g Starch per 10 L water). Once the media reached room temperature after

sterilization, the fermenter was inoculated with the 2 second stage cultures and the fermenter agitated at 300 rpm for 72 h.

Following fermentation, the cultures were centrifuged at 7,000 rpm at 4 °C. The cell pellets were discarded, and the medium extracted with an equal volume of dichloromethane. The organic layer was collected, dried over anhydrous magnesium sulfate, and concentrated in vacuo. The resulting crude extract was stored under diethyl ether at -80 °C. Crude extracts were analyzed by LC-MS. The sample was monitored by UV at $\lambda = 254$ nm. The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min.

Feeding Conditions

Fermentation was carried out using above conditions. After 24 hrs, 500 mg of compound was added and fermentation allowed to continue. A second 500 mg was added after an additional 24 hrs. The fermentation culture was processed and analyzed the same as above.

Table 6 Strains and Plasmids used in Chapter V

Strain	Characteristic	Source
E. coli BL21 (DE3)	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3)	Invitrogen
E. coli BL21 (DE3) pET-24-aziB	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with AziB expression plasmid	Mori *et. al.*, 2016[80]
E. coli BL21 (DE3) pET-24-acp	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with ACP domain expression plasmid	This study
E. coli BL21 (DE3) pET-21-ks-kr	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with KS-AT-DH-KR domain expression plasmid	This study
E. coli BL21 (DE3) pET-21-svp	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with Svp expression plasmid	Mori *et. al.*, 2016[80]
E. coli BL21 (DE3) pET-21-sfp	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with Sfp expression plasmid	Mori *et. al.*, 2016[80]
E. coli BL21 (DE3) pET-21-FASPPTase	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with FAS PPTase expression plasmid	This study
E. coli BL21 (DE3) pET-21-NFASPPTASE	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with NFAS PPTase expression plasmid	This study
pET-24-aziB	*E. coli* expression vector for AziB	Mori *et. al.*, 2016[80]
pET-24-acp	*E. coli* expression vector for ACP domain	Mori *et. al.*, 2016[80]
pET-21-ks-kr	*E. coli* expression vector for KS-AT-DH-KR domains	Mori *et. al.*, 2016[80]
pET-21-svp	*E. coli* expression vector for Svp	Mori *et. al.*, 2016[80]
pET-24-sfp	*E. coli* expression vector for Sfp	Mori *et. al.*, 2016[80]
pET-21-FASPPTase	*E. coli* expression vector for FAS PPTase	This study
pET-21-NFASPPTase	*E. coli* expression vector for FAS PPTase	This study
S. sahachiroi	Azinomycin producing organism	ATCC
S.sah Δ*aziG*	*S. sahachiroi* lacking AziG	This study
pKCAziG	Plasmid for the disruption of *aziG*	Genscript

Table 7 Primers used in Chapter V

Primer	**Sequence (5'-3')**
KS-KRF	CTTTAAGAAGGAGATATACATATGGCTGAGAATGTGCAGAAC
KS-KRR	AGTGGTGGTGGTGGTGGTGCTCGAGCTGCTCGTGGGTCACCTTC
ACPF	CATATGCGGTTCGCGCTGCCGTA
ACPR	CTCGAGGACGGCGTCGAGCTGC
AprF	CGAATGGCGAAAAGCCGAGC
AprR	GCATCGCATTCTTCGCATCC
AziGUpF	TTGAATTCCTCGCTGTGGATCTTCTC
AziGDownR	TCAAGCTTGCTCCAGTACAACCTGTG
AziH1F	CAGAAGTACCCCTCACGCTT
AziH1R	TGGACGTGTGGCAGTACAT
AziH3F	GATGACGAGGTCGGGCTTC
AziH3R	GACACCGAGTGGAGGTCCTG
AziFF	CGGAACTGCGGAATGCG
AziFR	GAGGAACTCGCGGCTCTCA
AziD1F	GAGGTAGAGGGCGAAGGAG
AziD1R	GAACTGGCTCTCCATGTGGT
AziBF	CGCGGGGGCGCAGAGCGCCG
AziBR	CCCGGCCGCGGCCGCGACGG

CHAPTER VI

CHARACTERIZATION OF THE TRANSKETOLASE AZIC5/C6*

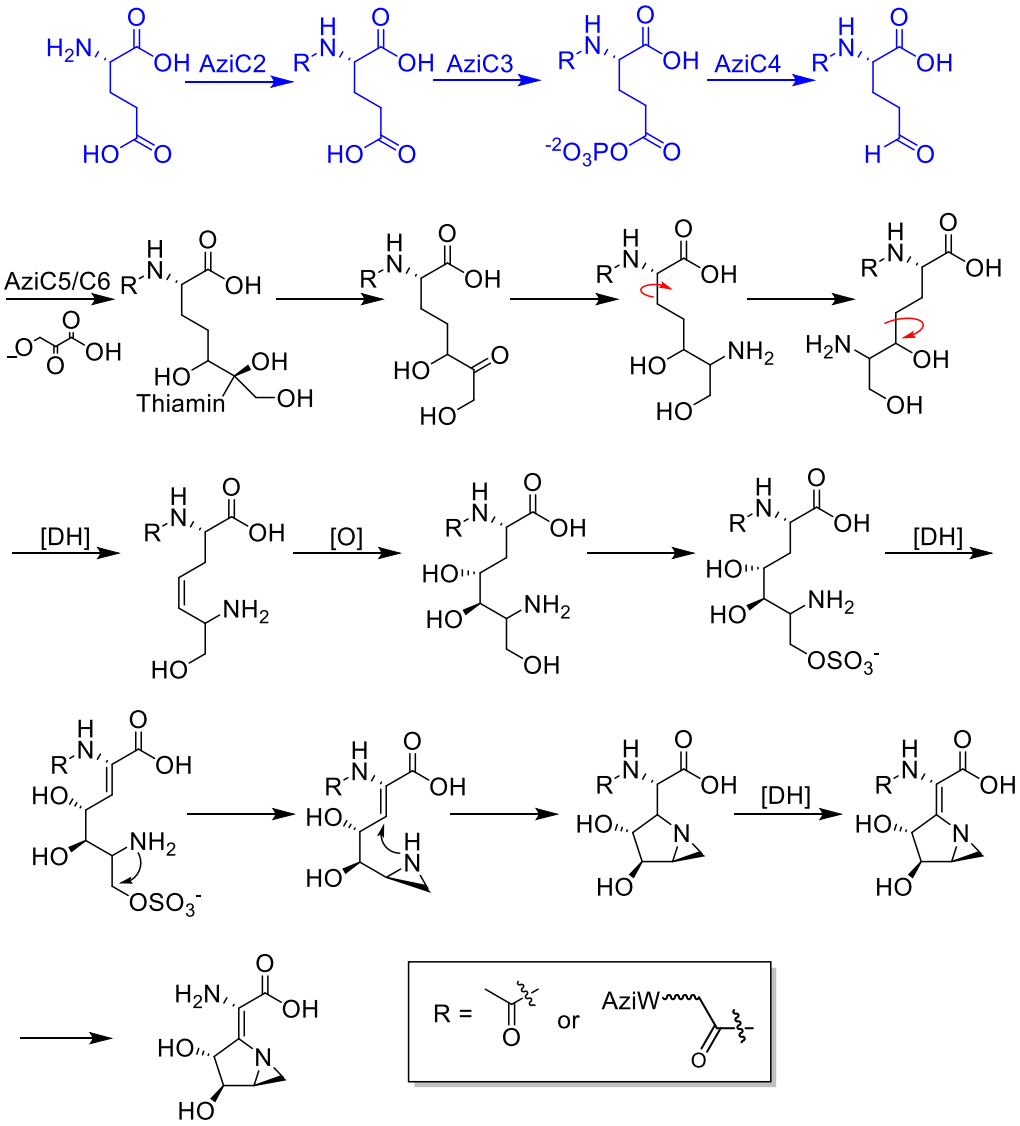
Introduction

The biosynthesis of the azabicycle (aziridino[1,2-a]pyrrolidine) moiety of the azinomycins has proven to be particularly complex. Scheme 3 shows a proposed biosynthetic route for formation of the azabicycle moiety, with previously validated steps shown in blue. Whole cell feeding studies revealed glutamic acid as the starting precursor.[101] Additionally, studies have shown AziC2 carries out a protection step on the amino-group giving N-acetyl glutamic acid. Bioinformatics suggest this protection might also be facilitated with an amino group carrier protein, AziW.[106] Protection is followed by AziC3 and AziC4 catalyzing the formation of N-acetyl-glutamate 5-semialdehyde via an acyl phosphate intermediate.[107]

Transketolase enzymes catalyze the transfer of a dihydroxylethyl group from a ketose donor molecule to an aldose acceptor resulting in an α,α-dihydroxyketone molecule.[108] Transketolases are prevalent in the pentose phosphate pathway and Calvin cycle, as well as numerous biosynthetic pathways. Thiamin and a divalent cation M^{2+} are necessary cofactors for transketolase activity.[109] DNA annotation suggests that *aziC5* and *aziC6* encode C- and N-terminal subunits of a transketolase enzyme. Based on their

*Reproduced with permission from Washburn, L.; Foley, B.; Martinez, F.; Lee, R.; Pryor, K.; Rimes, E.; Watanabe, C. "Transketolase Activity in the Formation of the Azinomycin Azabicycle Moiety" *Biochemistry* **2019**, *58* (52), 5255-5258. COPYRIGHT 2020 American Chemical Society.

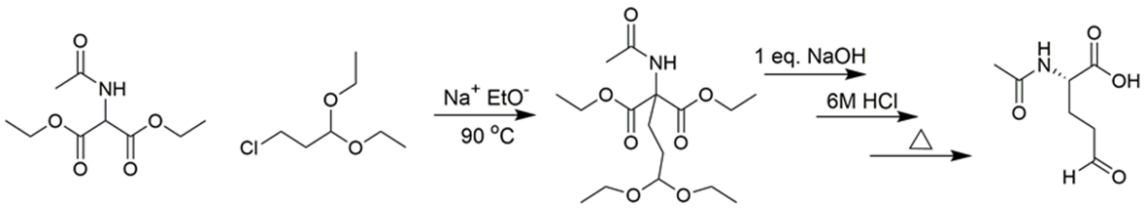
54% similarity to the transketolase in vazabitide biosynthesis, it was proposed that AziC5 and AziC6 use thiamin pyrophosphate (TPP) and magnesium (II) as cofactors to catalyze the transfer of a two-carbon unit onto the N-acetyl-glutamate 5-semialdehyde product of the AziC4 reaction (Scheme 3).[106]



Scheme 3 Biosynthesis of the Azabicycle Moiety
Characterized steps shown in blue.

Preliminary Research

aziC5 and *aziC6* were successfully cloned, overexpressed, and purified in *E. coli* BL21 (DE3) to evaluate this step of the pathway (Rachel Lee). Synthesis of the aldehyde starting material was achieved following Scheme 4 (Flor Martinez).



Scheme 4 Synthesis of AziC5/C6 Aldehyde Substrate
Flor Martinez

Transketolase activity was initially probed using a tetrazolium red colorimetric assay in which the α-hydroxy ketone product produced by the transketolase reaction reduces a colorless tetrazolium to a red formazan precipitate.[110] Hydroxypyruvate is often used as a general donor molecule. This is because the transketolase reaction is reversible *in vivo*, and use of hydroxypyruvate results in loss of CO_2 and an irreversible reaction.[108, 110-111] Red precipitate only formed when the transketolase was incubated with N-acetyl-glutamate 5-semialdehyde, thiamin pyrophosphate and the hydroxypyruvate donor molecule (Figure 52). This suggests product only forms in the presence of the AziC5/C6 enzyme and cofactors.

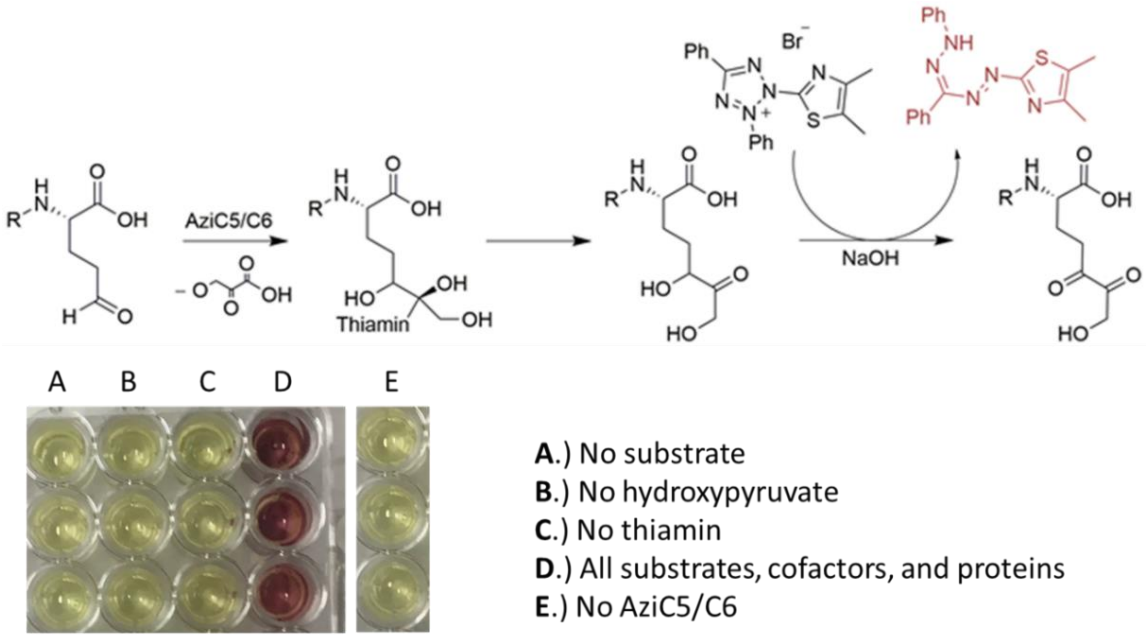


Figure 52 AziC5/C6 Colorimetric Assay
Brendan Foley

Results and Discussion

To further demonstrate the thiamin dependency and binding capabilities of AziC5/C6 a thiochrome assay was performed. Oxidation of thiamin produces a highly fluorescent thiochrome which can be characterized by HPLC.[112] Exposure of purified AziC5/C6 to the oxidant resulted in production of thiochrome as characterized by HPLC. The thiochrome produced from AziC5/C6 was compared to a synthetic thiochrome standard (Figure 53).

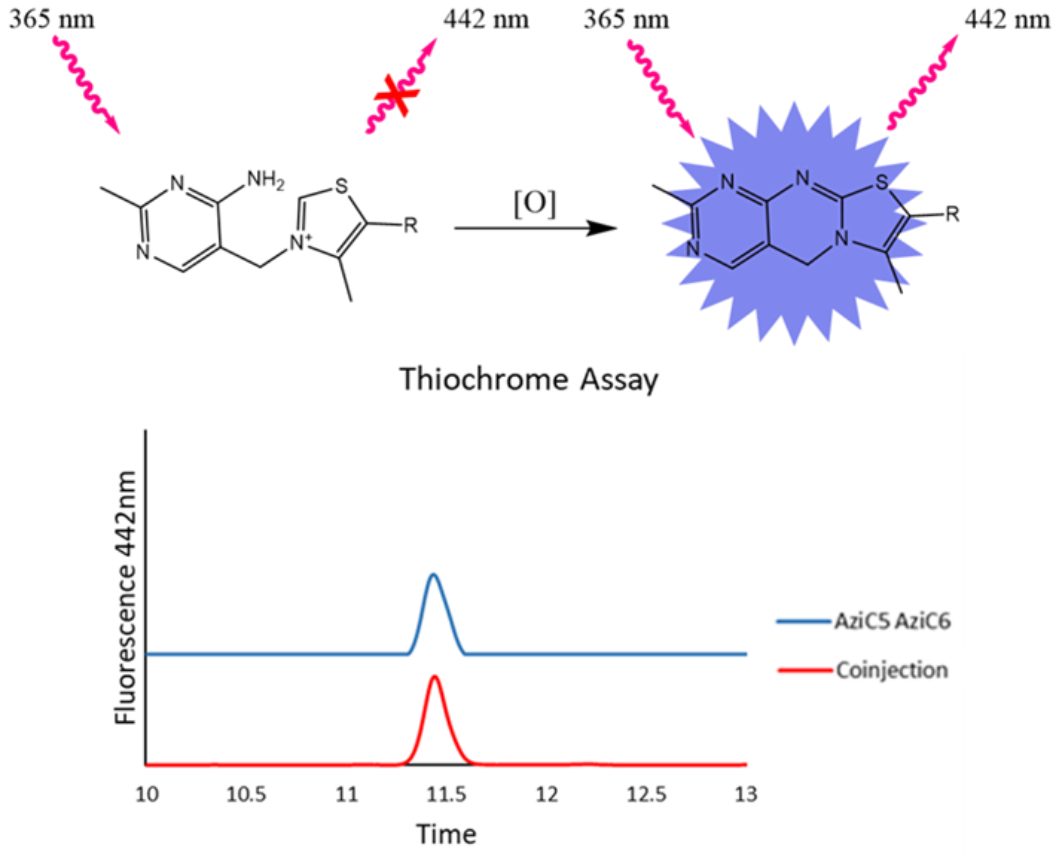


Figure 53 AziC5/C6 Thiochrome Assay

A homology model based on the structure of human transketolase was generated and the binding pocket identified with highly conserved histidine residues, including the Lys272-His286 catalytic dyad (Figure 54).[113] The predicted structure is a tetramer with a homodimer of AziC5 and a homodimer of AziC6. The first step in the proposed transketolase catalytic cycle of AziC5/C6 involves the formation of the thiamin ylide, which attacks the carbonyl carbon of the keto donor substrate. Deprotonation by His286 of AziC5 leads to the dihydroxylethyl-thiamin enamine intermediate, which reacts with

the aldose acceptor forming a new carbon-carbon bond. Subsequent deprotonation by His129 of AziC6 generates the final α,α-dihydroxyketone product as well as regenerated thiamin ylide for another catalytic cycle.[108, 114] His75 and His139 of AziC5 are also highly conserved and are proposed to function in H-bonding of the substrate during the catalytic cycle.

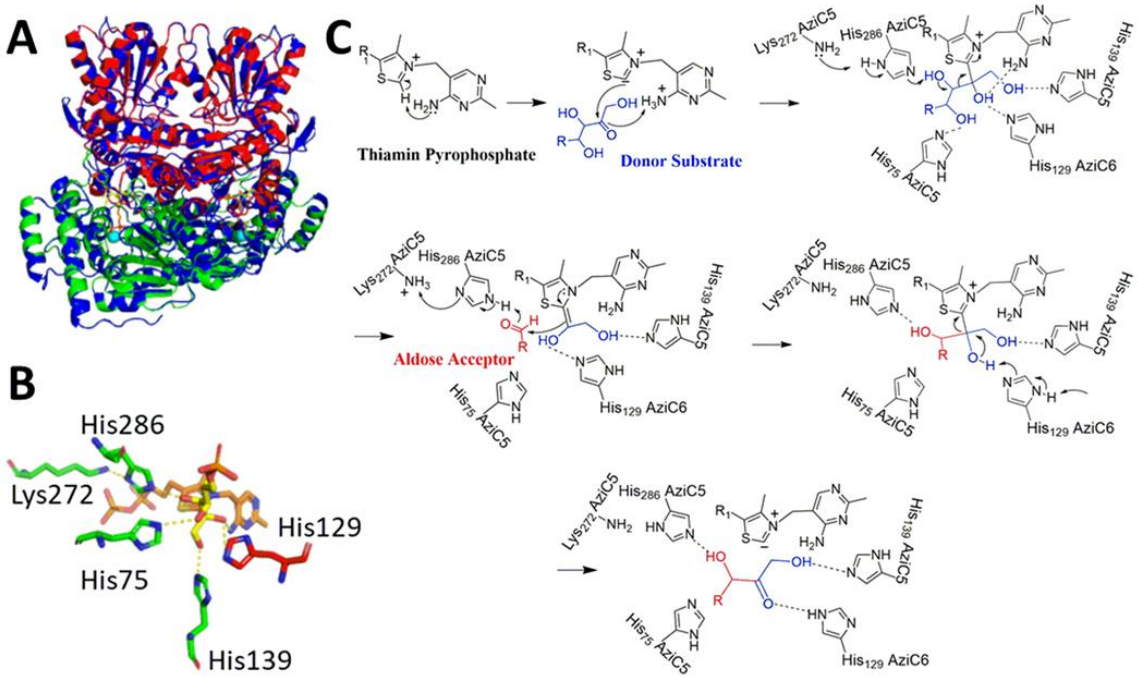


Figure 54 AziC5/C6 Model and Proposed Mechanism
A) Overlay of AziC5/C6 model with Human Transketolase (4KXU). Blue = 4KXU, Green = AziC5, Red = AziC6, Orange = thiamin pyrophosphate, Yellow = D-fructose-6-phosphate, Cyan = Mg2+. **B**) Zoom in on one catalytic site with conserved residues involved in substrate binding and enzymatic activity shown. **C**) Predicted mechanism based on the homology model and conserved residues.

The activity of AziC5/C6 toward dihydroxylethyl donor compounds was measured spectrophotometrically using potassium ferricyanide absorbance at 420 nm.

The α-carbanion intermediate is oxidized in the presence of ferricyanide, resulting in a decrease in absorbance at 420 nm as the ferricyanide is reduced.[109, 111, 115-116] Fructose-6-Phosphate (F6P), xylulose-5-phosphate (X5P), and hydroxypyruvate (HP) were used as the dihydroxylethyl group donor molecules (Figure 55). F6P showed the highest activity. followed by X5P and then HP. The Km values and Michaelis-Menten Plot can be found in Appendix Figure 32 and Appendix Table 2 and 3. The cellular concentration of these types of metabolites and their Km values may suggest that enzyme activity will fluctuate with substrate availability.[117]

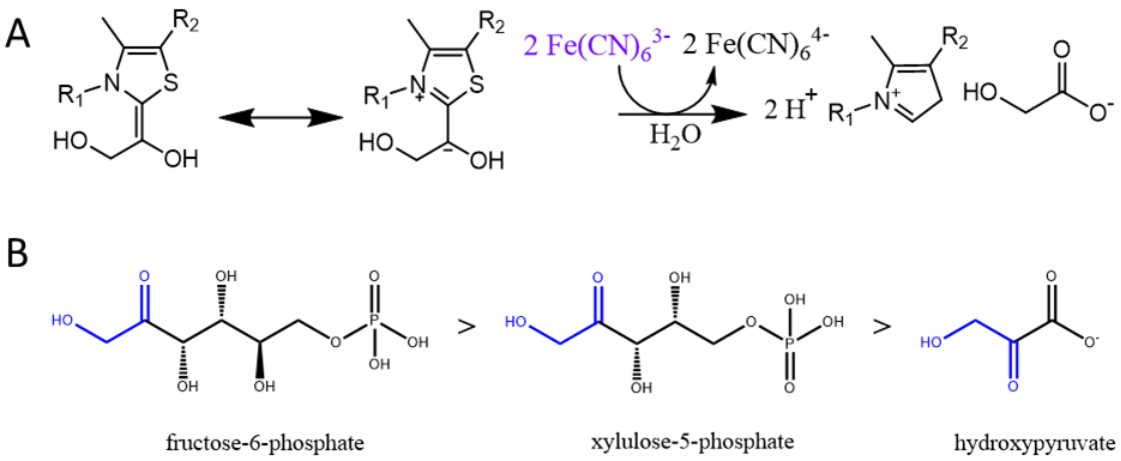


Figure 55 AziC5/C6 Donor Molecule Screening

Significance

Previous ^{13}C feeding studies identified the origins of the azinomycin carbon framework, except the two carbons needed for aziridine ring formation. This study demonstrated the transketolase mediated two carbon extension necessary for the

generation of the carbon framework of the aziridino[1,2-a]pyrrolidine moiety and identifies potential sources of the carbons previously unidentified through ^{13}C feeding experiments. N-acetyl-glutamate 5-semialdehyde is converted to 2-acetamido-5,7-dihydroxy-6-oxoheptanoic acid by AziC5/C6, a thiamin dependent transketolase. AziC5 and AziC6 encode the N- and C-terminal subunits of the enzyme, catalyzing the 2-carbon extension with a suitable donor, e.g. hydroxypyruvate. Scheme 3 depicts the overall proposed route to the formation of 2-acetamido-5,7-dihydroxy-6-oxoheptanoic acid and its role in the formation of the aziridino[1,2-a]pyrrolidine based upon previous work and bioinformatic analysis of the azinomycin gene cluster.[79] In the long term, understanding the mechanism of formation of the aziridino[1,2-a]pyrrolidine moiety will enable genetic engineering of novel azabicycle containing agents.

Material and Methods

General Procedures

All chemicals were purchased from Sigma-Aldrich and all bacterial media were purchased from Becton-Dickinson. UV spectra were measured on a Thermo Fisher GENESYS 10 Bio spectrophotometer.

Expression and Purification of AziC5/C6

The construct was transformed in *E. coli* BL21 (DE3) for protein overexpression. A 10 mL overnight culture in LB medium containing 50 µg/mL kanamycin was added to 1 L of medium and grown until an OD_{600} of 0.8 was reached. The culture was induced with 1

mM IPTG and incubated at 16 °C for 22 h prior to harvest. The cells were centrifuged at 6,000 rpm for 20 min, washed twice with cold 50 mM phosphate buffer (300 mM NaCl and 10% glycerol at pH 7) and re-suspended in the same buffer containing 1 mM dithiothreitol (DTT) and 1 mM phenyl methylsulfunyl fluoride (PMSF). The cell pellets were sonicated on ice water bath using Branson Sonifier 450 (Branson Ultrasonics) fitted with a micro tip, output setting 6, duty cycle 50%, for 8 cycles for 30 sec each and centrifuged for 1 h at 12,000 rpm. The soluble protein was purified by Nickel affinity chromatography, His Trap FF 5 mL column (GE Healthcare Life Sciences), according to the instructions provided by the supplier. The purified fractions were desalted and concentrated using Amicon 10K Ultra centrifugal filter (EMD Millipore).

Thiochrome Assay

AziC5/C6 were expressed in the presence of 100 mM thiamin pyrophosphate. The harvested protein was purified on a His Trap Column and the purified fractions desalted and concentrated with an Amicon 10K Ultra centrifugal filter (EMD Millipore). The isolated protein was quenched with an equal volume of TCA and centrifuged. To the resulting supernatant (100 μL) was added potassium acetate (50 μL of 4 M), followed by 7 M NaOH, resulting in the formation of the desired thiochrome phosphate. The oxidation reaction was neutralized with 6 M HCl and analyzed by HPLC.

Building Homology Model

Homology modeling was performed using SWISS-MODEL. Model quality was evaluated by having a QMEAN score of less than 4. The model for AziC5/C6 was built upon the crystal structure of a human transketolase with d-fructose-6-phosphate bound and thiamin bound (PDB 4KXU).5 AziC5/C6 and 4KXU have 33 % sequence identity. The model was visualized using Pymol to identify residues that may play a role in transketolase activity of AziC5/C6.

Screening of Donor Molecules

The following method was used to monitor the rate of formation of the keto donor molecule carbanion intermediate. The reaction contained 50 mM glycylglycine pH 7.6, 2 mM $MgCl_2$, 0.2 mM TPP, 1 mM $K_3[Fe(CN)_6]$, 0.4 mg/mL AziC5/C6, and donor molecule substrates (F6P, X5P, and HP, utilizing a range of concentrations from 0.5 mM to 40 mM) in a final volume of 1 mL. The reaction was initiated by the addition of AziC5/C6 and was performed in triplicate. Ferricyanide reduction was monitored at 420 nm.

Table 8 Strains and Plasmids used in Chapter VI

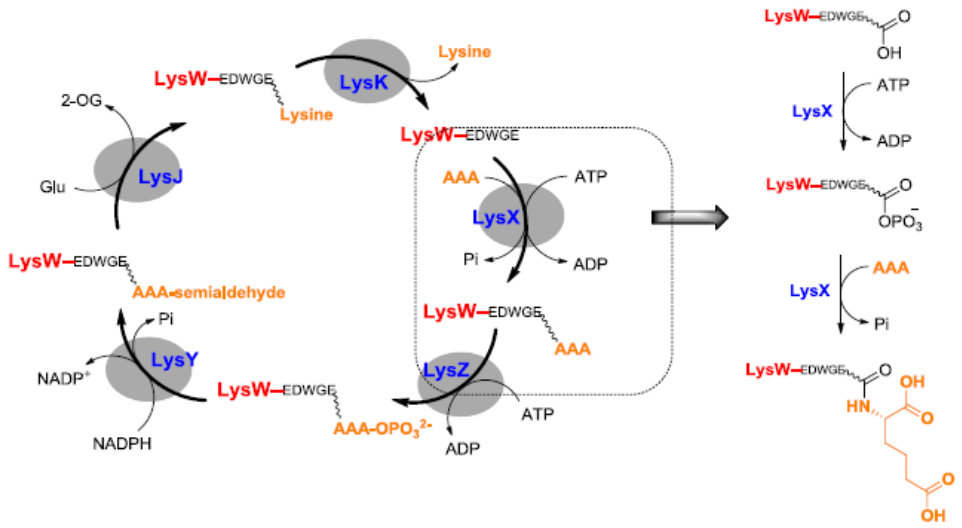
Strain or Plasmid	Characteristic	Source
E. coli BL21 (DE3)	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3)	Invitrogen
E. coli BL21 (DE3) pET-24-aziC5	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with AziC5 expression plasmid	This Study
E. coli BL21 (DE3) pET-24-aziC6	B; F–ompT hsdSB(rb- mB-) gal dcm (DE3) with AziC6 expression plasmid	This Study
pET-24-aziC5	*E. coli* expression vector for AziC5	This Study
pET-24-aziC6	*E. coli* expression vector for AziC6	This Study

CHAPTER VII

INVESTIGATION OF AZIW *IN VIVO*

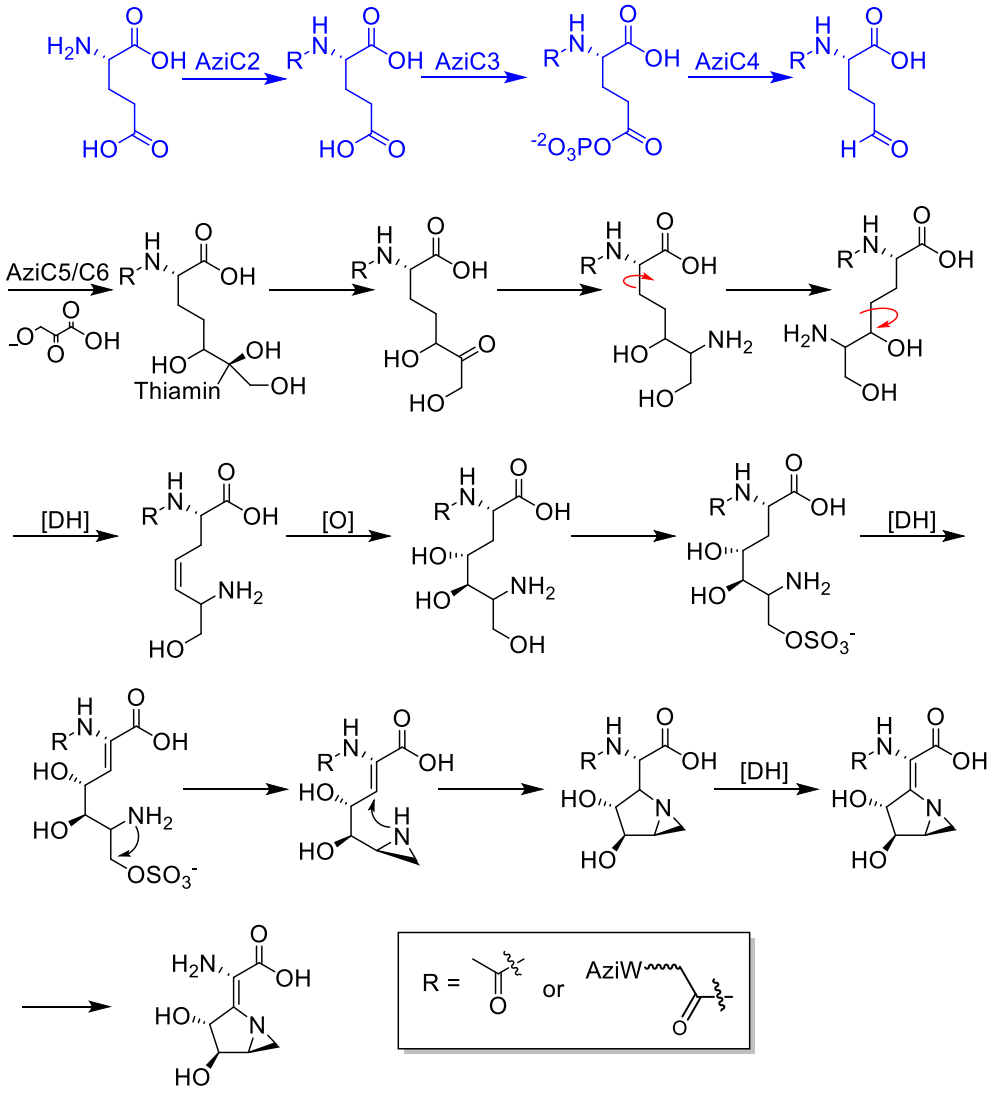
Introduction

Amino-acyl carrier proteins have been implicated in the biosynthesis of primary metabolites (lysine and arginine) as well as secondary metabolites (vazabitide).[106, 118-122] The first characterized protein of this family, LysW, was involved in Lysine biosynthesis (Scheme 5). This family of carrier proteins functions by forming an amide bond between the substrate and a C-terminus glutamic acid of the carrier protein. This step requires a transferase (LysX) and ATP. The protein substrate complex functions as a protecting group and carrier protein, leading the substrate through several successive reactions. Finally, the amide bond is hydrolyzed by a release protein (LysK) yielding the free amine of the product and the carboxylic acid of the amino group carrier protein.



Scheme 5 LysW Function in Lysine Biosynthesis
AAA = α-aminoadipic acid

Bioinformatic studies suggest that *aziW*, a gene in the azinomycin biosynthetic gene cluster, may code for an amino-acyl carrier protein. Based on homology to the lysine and vazabitide gene cluster, we propose AziW functions in the biosynthetic pathway to the azabicycle moiety, forming a covalent amide bond with the glutamic acid precursor (Scheme 6).



Scheme 6 Azabicycle Biosynthetic Pathway

Preliminary Research

Bioinformatic analysis of AziW revealed it to have 50 % sequence identity to LysW. The proteins in the amino-acyl carrier protein family have a conserved C-terminal amino acid sequence of EDWGE (Figure 56). The C-terminal glutamic acid of the protein forms the bond with the substrate. The conjugation of the substrate to LysW required the protein LysX and ATP. In the azinomycin gene cluster, AziC2 shows 44 % sequence identity to LysX.

Figure 56 Sequence Alignment of AziW and LysW Homologs
Rachel Lee

As mentioned in Chapter I, feeding studies revealed glutamic acid to be the precursor to the azabicycle moiety. The *aziW* gene was originally not annotated as part of the azinomycin gene cluster, because of this, it was proposed that AziC2 functioned solely as an acetyl-transferase. Studies with AziC2 showed it accepts glutamic acid and acetyl-Coenzyme A as substrates, producing N-acetyl glutamic acid. Disruption of *aziC2* and subsequent feeding of N-acetyl glutamic acid to the Δ*aziC2* strain resulted in azinomycin A production.[101] Based on the identification of AziW, a reaction scheme for the AziW and AziC2 was proposed (Figure 57)

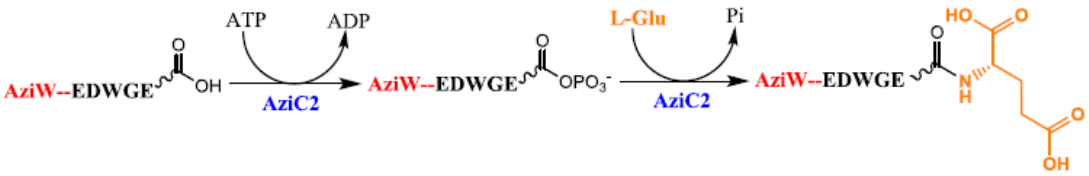


Figure 57 Proposed AziW and AziC2 Reaction

To probe this reaction, a malachite green assay was performed to measure the free phosphate released in the enzymatic reaction. Inorganic phosphate and malachite green molybdate form a complex under acidic conditions which can be monitored at 620 nm. Incubation of AziC2 and AziW with glutamic acid showed phosphate release as compared to the control reactions with different or no substrate (Figure 58).

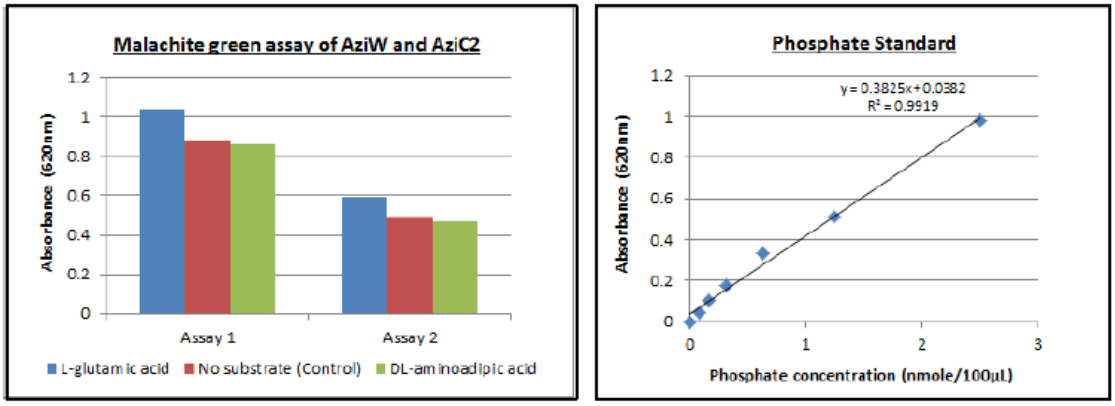
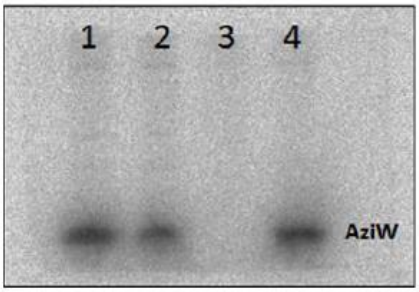


Figure 58 Malachite Green Assay with AziC2 and AziW
Rachel Lee

To further probe this AziW and AziC2 reaction, [γ-^{32}P]-ATP used, the reaction run on SDS-PAGE, and the gel visualized using autoradiography (Figure 59).

AziW was labeled in the presence and absence of glutamic acid (Lane 1 and 2). The most interesting result was the control with no AziC2 (Lane 4). Based on the proposed mechanisms for LysW-like proteins, we would not expect AziW to be labeled in the absence of AziC2. This result suggests AziW is self-phosphorylating. To further probe the reaction, *in vitro* reconstitution with AziW and AziC2 was attempted. To date, no AziW-glutamic acid product has been detected. To probe AziW's role *in vivo*, AziW was previously disrupted and the Δ*aziW* strain does not produce azinomycin (Dr. Keshav Nepal).



Lane	2mg/mL AziC2	2mg/mL AziW	0.1mCi/mL [γ-^{32}P]-ATP	5mM L-Glu	1mM $MgSO_4$
1	+	+	+	-	+
2	+	+	+	+	+
3	+	-	+	+	+
4	-	+	+	+	+

Figure 59 [γ-^{32}P]-ATP AziW and AziC2 Reaction
Rachel Lee

Results and Discussion

We aimed to validate the Δ*aziW* and Δ*aziC2* fermentation results through gene complementation and feeding of N-acetyl glutamic acid to the Δ*aziW* strain. To complement *aziW* back into the Δ*aziW* strain, the pSETaziW plasmid was constructed by Genscript. Conjugal transformation was used to introduce pSETaziW into the Δ*aziW strain*. PCR confirmed successful integration of the plasmid into the genome of *S.*

sahachiroi Δ*aziW*. Fermentation was performed of wild type *S. sahachiroi, AziW:*Δ*aziW* complementation strain, Δ*aziW,* and Δ*aziW* + N-acetyl glutamic acid feeding. All four samples were analyzed by LC-MS for direct comparison. A representative Extracted Ion Chromatogram (EIC) for Azinomycin B is shown in Figure 60. Mass Spectra for

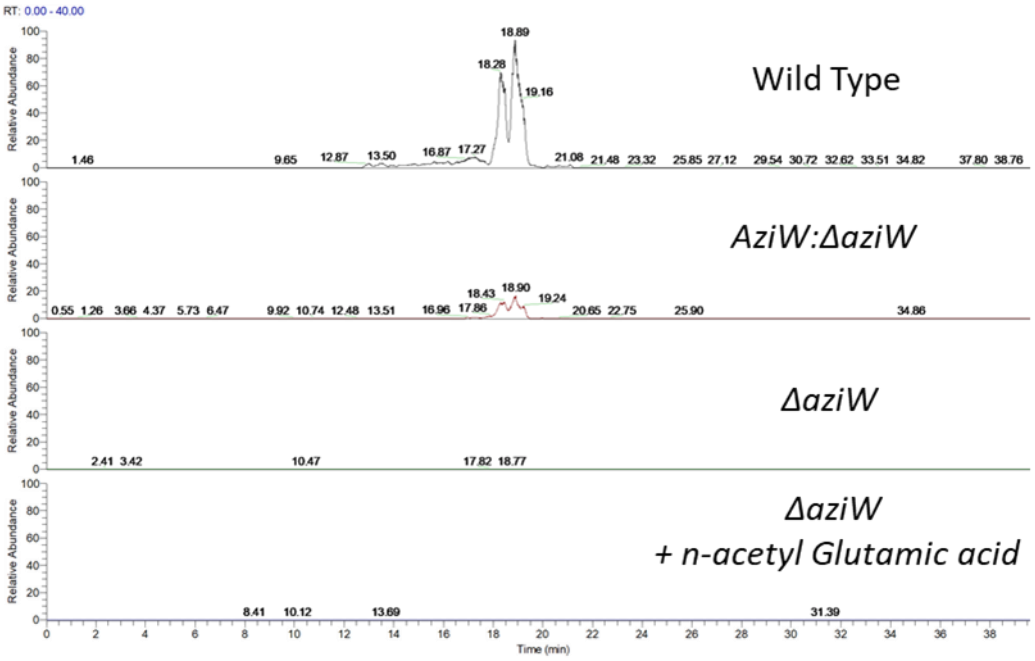


Figure 60 Extracted Ion Chromatogram for Azinomycin B in Δ*aziW* Study

Azinomycin A and B and EIC for Azinomycin A can be found in Appendix Figures 33-35. The results showed azinomycin production in wild type and the *AziW:*Δ*aziW* complementation samples, but not the Δ*aziW* and Δ*aziW* + N-acetyl glutamic acid feeding samples. This result validates the necessity of AziW in azinomycin production but raises questions about the AziW/AziC2 interaction previously thought to be vital for AziW activity and azinomycin production. Further validation with *in vitro* reconstitution

and additional feeding studies will shed light on the role of AziC2 and AziW in production of the azinomycins.

Significance

AziW has proven to be vital for production of the azinomycins as is evident by gene deletion and complementation studies. N-acetyl glutamic acid does not seem to bypass the role of AziW in the Δ*aziW* strain. This is an interesting result as gene deletion studies with *aziC2* showed N-acetyl glutamic acid can lead to azinomycin A production.[101] Combined with the preliminary phosphate assays, this result suggests there are differences in the azinomycin biosynthetic system as compared to the amino-acyl carrier protein system found in Lysine biosynthesis. The lack of azinomycin production when N-acetyl glutamic acid was fed to the Δ*aziW* strain suggests there are different roles for AziW and AziC2 *in vivo*. Lack of AziW-glutamic acid production *in vitro* when incubated with AziC2 and substrate also suggests there are unidentified roles or missing substrates found in the AziW system *in vivo.* If AziW is self-phosphorylating, AziC2 may not be necessary for activation. A feeding study where [^{13}C]N-acetyl glutamic acid is fed to the Δ*aziC2* strain would shed light on the origins of the glutamic acid in the produced azinomycin A. Additionally, *in vitro* reconstitution with the enzymes in the biosynthetic pathway to the azabicycle has shown the first four steps to accept an N-acetyl substrate. These results together show the azinomycin pathway may allow for promiscuous reactions to occur.

Material and Methods

General Procedures

All chemicals were purchased from Sigma-Aldrich and all bacterial media were purchased from Becton-Dickinson. Molecular biology reagents were obtained from New England Biolabs. Mass Spectra were obtained at the Laboratory for Biological Mass Spectrometry at the Department of Chemistry, Texas A&M University. A Phenomenex column (Prodigy 5 µm C18 150 Å, 150 x 4.6 mm) was used during LC-MS analysis. LC-MS was performed on a Thermo Scientific Ultimate 3000 UHPLC with a Thermo Fisher Scientific Q Exactive Focus mass spectrometer.

Bacterial Strains and Media

Streptomyces sahachiroi (NRRL 2485) was obtained from the American Type Culture Collection (ATCC). *E. coli* S17 was used as the donor host for intergenic conjugal transformation. 2xYT (16 g tryptone, 10 g yeast extract, 5 g NaCl in 1 L of water) was used for general liquid culturing of *S. sahachiroi.* GYM agar (4 g glucose, 4 g yeast extract, 2 g CaCO3, 10 g malt extract, 8 g agar in 1 L water adjusted to pH 6.8 with NaOH) was used for plating *S. sahachiroi* spores. ISP4 agar was used in plating of the intergenic conjugal transformation.

Generation of pSETAziW:ΔaziW Complementation Strain

The plasmid pSETaziW was designed and constructed by Genscript. Intergenic conjugal transformation was used to introduce the pSETaziW plasmid to *S. sahachiroi* to facilitate

gene complementation by integration. The pSETaziW plasmid was transformed into *E. coli* S17 cells. An overnight culture of pSETaziW/S17 cells was grown and washed twice with LB medium and resuspended in 600 µL of 2xYT medium. *S. sahachiroi* Δ*aziW* spores from 10 GYM agar plates were collected using 2xYT medium. The collected spores were subjected to heat shock at 65 °C for 10 min and incubated at 37 °C for 3 h. Following incubation, the spores were collected and washed with 20 ml 2xYT two times and finally resuspended in 4 ml 2xYT. The recipient cells were mixed with the 600 µl of pSETaziW/S17 donor cells and 800 µl of the mixture was plated on ISP4 agar plates. The plates were incubated at 28 °C and overlaid with 1 ml of water containing nalidixic acid (500 µg) and apramycin (70 µg) to select for *S. sahachiroi* exconjugants after 24 h. Incubation continued for 10 days to allow for appearance of apramycin resistant exconjugants. The pSETaziW:Δ*aziW* strain was confirmed by PCR. PCR was used to screen for the apramycin resistance gene (AprF/AprR primers)

Culture Conditions and Fermentation

A 1 cm^2 piece of GYM plate containing *S. sahachiroi* spores was used to inoculate a first stage culture of 100 ml of PS5 medium (0.5 g Starch, 0.5 g Pharmamedia per 100 ml water) in a 250 ml flask. The culture was incubated at 30 °C for 24 h at 250 rpm. The second stage culture was prepared by inoculating 2 x 600 ml PS5 with 25 ml of first stage culture. The culture was incubated at 30 °C for 24 h at 250 rpm. The fermenter was prepared by autoclaving while containing 10 L of reduced PS5 (12.5 g Pharmamedia, 12.5 g Starch per 10 L water). Once the media reached room temperature after

sterilization, the fermenter was inoculated with the 2 second stage cultures and the fermenter agitated at 300 rpm for 72 h.

Following fermentation, the cultures were centrifuged at 7,000 rpm at 4 °C. The cell pellets were discarded, and the medium extracted with an equal volume of dichloromethane. The organic layer was collected, dried over anhydrous magnesium sulfate, and concentrated in vacuo. The resulting crude extract was stored under diethyl ether at -80 °C. Crude extracts were analyzed by LC-MS. The sample was monitored by UV at $\lambda = 254$ nm. The column was pre-equilibrated with 90% A (water, 0.1% formic acid) and 10% B (75% methanol, 24.9% isopropanol, and 0.1% formic acid). Solvent conditions were as follows: time 0 min, A-90% B-10%; 1 min, A-90% B-10%; 5 min, A-65% B-35%; 23 min, A-5% B-95%; 28 min, A-35% B-65%; 31 min, A-90% B-10%; 40 min, A-90 B-10%, flow rate 0.75 µL/min.

Feeding Conditions

Fermentation was carried out using above conditions. After 24 hrs, 500 mg of N-acetylglutamic acid was added and fermentation allowed to continue. A second 500 mg was added after an additional 24 hrs. The fermentation culture was processed and analyzed the same as above.

Table 9 Strains and Plasmids used in Chapter VII

Strain or plasmid	**Characteristic**	**Source**
S. sahachiroi	Azinomycin producing organism	ATCC
S.sah ΔaziW	*S. sahachiroi* lacking AziW	Previous unpublished work
S. sah AziW:ΔaziW	*S.sahΔaziA6* with complementation of *aziW*	This study
pSETAziW	Plasmid for complementation of AziW	Genscript

Table 10 Primers used in Chapter VII

Primer	**Sequence (5'-3')**
AprF	CGAATGGCGAAAAGCCGAGC
AprR	GCATCGCATTCTTCGCATCC

CHAPTER VIII

CONCLUSION

Natural products have played a vital role in human health since the beginning of our history. Technological advances have pushed us into a second "Golden Age" of discovery and development of natural products as drug leads. Many natural products were first discovered in the mid 1900's but could not be extensively studied due to limitations in biochemical techniques. The advances in sequencing, cloning, and high-throughput experimentation has allowed for new natural product discovery and more thorough characterization and development of previously discovered natural products, as is the case with the azinomycins.

Produced by the soil microbe *Streptomyces sahachiroi*, the azinomycins are an interesting target as they are highly potent as an antitumor agent, but they are difficult to synthesize, and their biosynthesis is not fully understood. Unlocking the key to their biosynthesis opens the door for increased production and the design of azinomycin analogues. The azinomycin backbone is assembled by PKS and NRPS enzymes, which carry their substrates as a thioester bound intermediate. To trap and characterize these intermediates a probe was developed based on the native chemical ligation reaction. The probe was first tested on the PKS AziB, which plays a role in formation of the naphthoate moiety. The utilization of this system *in vitro* identified Biotin-Cys as the best probe and pointed out flaws with the fluorophore-based probes. This system also

validated the production of 2-methylbenzoic acid by AziB when the thioesterase AziG is not present.

As the “Left-Half” (naphthoate epoxyamide) of azinomycin has also shown antitumor activity, understanding the biosynthesis of the epoxide moiety will allow for analogs containing this “Left-Half” to be pursued. The NRPS AziA3 was identified as playing a role in the final steps of epoxide moiety formation based on its homology to NRPS enzymes that reduce keto acids to hydroxy acids. This reduction to an alcohol allows for the ester bond to be made with the naphthoate moiety. Due to low PPTase post-translational modification efficiency with AziA3, *in vitro* reconstitution could not be pursued, additionally, the native substrate of AziA3 was unknown. To overcome these issues, AziA3 was overexpressed in the azinomycin producing organism. This allowed for *S. sahachiroi* native PPTases to modify AziA3 and the organism supplied the native substrate as part of the azinomycin biosynthetic pathway. To ensure buildup of this intermediate on AziA3, the thioesterase responsible for product release was disrupted. Reaction of AziA3 with the Biotin-Cys probe revealed the final epoxide moiety as the AziA3 product. While moiety condensation and installation of the epoxide need to be characterized, this study validated the use of the thioester intermediate capture strategy with protein isolated from the producing organism.

To further study NRPS and PKS proteins, protein expression and post-translational modification efficiency need to be optimized *in vitro.* The ACP domain of AziB was separated from the four other domains (KS-AT-DH-KR). This increased expression and stability of the domains as determined by Native Mass Spec.

Additionally, two PPTases found in *S. sahachiroi* were cloned for *in vitro* expression. While the promiscuous PPTase Svp modified AziB and the separated ACP domain most efficiently, the FAS PPTase from *S. sahachiroi* could also modify both. When incubated together, the separated domains produced 2-methylbenzoic acid, demonstrating that domain separation is a viable way to increase expression and product yield *in vitro*. Additionally, *aziG*'s role in naphthoate moiety production was validated in *S. sahachiroi* through gene disruption. The Δ*aziG* strain was utilized to produce benzoic-azinomycin derivatives.

The biosynthetic pathway of the azabicycle moiety has remained difficult to characterize. Feeding studies identified glutamic acid as its precursor, but this precursor is missing the two carbons required to form the reactive aziridine ring. A transketolase AziC5/C6 was identified and proposed to perform the two-carbon extension. A series of *in vitro* experiments confirmed this proposal, identified AziC5/C6's thiamin dependence, and screened potential two-carbon donor molecules. The two-carbon extension establishes the final carbon framework for the aziridine ring and azabicycle moiety. In addition to establishing the carbon framework, the presence of an amino-acyl carrier protein, *aziW,* in the gene cluster suggests the azabicycle is formed as a protein-bound intermediate. To validate AziW's role, gene disruption, gene complementation, and feeding studies were performed. These studies showed that AziW is necessary for azinomycin production, and that N-acetyl-glutamic acid cannot bypass the need for AziW. This differs from the results with AziC2 gene disruption and feeding

experiments, highlighting a potential difference in the role of AziC2 and AziW as compared to homologous pathways.

The insights into azinomycin biosynthesis opens the door for increased production of azinomycin and analogues *in vitro, in situ,* and *in vivo*. Further studies of the condensation reactions between moieties and the installation of the epoxide and aziridine rings will directly influence what analogues are capable of being produced. Additionally, improvements to and utilization of the intermediate capture strategy will allow for the pursuit of cryptic metabolite gene clusters to be identified and key NRPS and PKS steps characterized without extensive *in vitro* studies beforehand. Insights into biosynthetic pathways will advance our ability to develop new natural product based therapies. The research discussed in this dissertation bring us closer to this goal.

REFERENCES

1. Cragg, G. M.; Newman, D. J., Natural products: A continuing source of novel drug leads. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013,** *1830* (6), 3670-3695.

2. Borchardt, J. K., The Beginnings of Drug Therapy: Ancient Mesopotamian Medicine. *Drugs News Perscpet.* **2002,** *15*, 187-192.

3. Dias, D. A.; Urban, S.; Roessner, U., A historical overview of natural products in drug discovery. *Metabolites* **2012,** *2* (2), 303-336.

4. Wongsrichanalai, C.; Pickard, A. L.; Wernsdorfer, W. H.; Meshnick, S. R., Epidemiology of drug-resistant malaria. *The Lancet Infectious Diseases* **2002,** *2* (4), 209-218.

5. Kingston, D. G. I., Taxol and its Analogs. *Anticancer Agents from Natural Products* **2005,** *1*, 89-122.

6. Williams, D. H.; Stone, M. J.; Hauck, P. R.; Rahman, S. K., Why Are Secondary Metabolites (Natural Products) Biosynthesized? *Journal of Natural Products* **1989,** *52* (6), 1189-1208.

7. Maplestone, R. A.; Stone, M. J.; Williams, D. H., The evolutionary role of secondary metabolites — a review. *Gene* **1992,** *115* (1), 151-157.

8. Mushtaq, S.; Abbasi, B. H.; Uzair, B.; Abbasi, R., Natural products as reservoirs of novel therapeutic agents. *EXCLI journal* **2018,** *17*, 420-451.

9. Katz, L.; Baltz, R. H., Natural product discovery: past, present, and future. *Journal of Industrial Microbiology & Biotechnology* **2016,** *43* (2), 155-176.

10. Scriabine, A., Discovery and Development of Major Drugs Currently in Use. *Pharmaceutical Innovation* **1999**, 148-270.

11. Butler, M. S., The Role of Natural Product Chemistry in Drug Discovery. *Journal of Natural Products* **2004,** *67* (12), 2141-2153.

12. Newman, D. J., Natural Products as Leads to Potential Drugs: An Old Process or the New Hope for Drug Discovery? *Journal of Medicinal Chemistry* **2008,** *51* (9), 2589-2599.

13. Lam, K. S., New aspects of natural products in drug discovery. *Trends in Microbiology* **2007,** *15* (6), 279-289.

14. Butler, M. S., Natural products to drugs: natural product derived compounds in clinical trials. *Natural Product Reports* **2005,** *22* (2), 162-195.

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

16. Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Ōmura, S., Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nature Biotechnology* **2003,** *21* (5), 526-531.

17. Challis, G. L.; Hopwood, D. A., Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species. *Proceedings of the National Academy of Sciences* **2003,** *100* (suppl 2), 14555-14561.

18. Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S., Genome Sequence of the Streptomycin-Producing Microorganism Streptomyces griseus IFO 13350. *Journal of Bacteriology* **2008,** *190* (11), 4050-4060.

19. Baltz, R. H., Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters. *Journal of Industrial Microbiology & Biotechnology* **2010,** *37* (8), 759-772.

20. Rutledge, P. J.; Challis, G. L., Discovery of microbial natural products by activation of silent biosynthetic gene clusters. *Nature Reviews Microbiology* **2015,** *13* (8), 509-523.

21. Baltz, R. H., Strain improvement in actinomycetes in the postgenomic era. *Journal of Industrial Microbiology & Biotechnology* **2011,** *38* (6), 657-666.

22. Yoon, V.; Nodwell, J. R., Activating secondary metabolism with stress and chemicals. *Journal of Industrial Microbiology & Biotechnology* **2014,** *41* (2), 415-424.

23. Baltz, R. H., Genetic manipulation of secondary metabolite biosynthesis for improved production in Streptomyces and other actinomycetes. *Journal of Industrial Microbiology & Biotechnology* **2016,** *43* (2), 343-370.

24. Ochi, K.; Tanaka, Y.; Tojo, S., Activating the expression of bacterial cryptic genes by rpoB mutations in RNA polymerase or by rare earth elements. *Journal of Industrial Microbiology & Biotechnology* **2014,** *41* (2), 403-414.

25. Barka, E. A.; Vatsa, P.; Sanchez, L.; Gaveau-Vaillant, N.; Jacquard, C.; Klenk, H.-P.; Clément, C.; Ouhdouch, Y.; van Wezel, G. P., Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews* **2016,** *80* (1), 1-43.

26. Liu, R.; Deng, Z.; Liu, T., Streptomyces species: Ideal chassis for natural product discovery and overproduction. *Metabolic Engineering* **2018,** *50*, 74-84.

27. Hopwood, D., Streptomyces in Nature and Medicine: the Antibiotic Makers. *Oxford University Press* **2007**.

28. Waksman, S. A.; Woodruff, H. B., Bacteriostatic and Bactericidal Substances Produced by a Soil Actinomyces. *Proceedings of the Society for Experimental Biology and Medicine* **1940,** *45* (2), 609-614.

29. Liu, Y.; Chen, X.; Li, Z.; Xu, W.; Tao, W.; Wu, J.; Yang, J.; Deng, Z.; Sun, Y., Functional Analysis of Cytochrome P450s Involved in Streptovaricin Biosynthesis and Generation of Anti-MRSA Analogues. *ACS Chemical Biology* **2017,** *12* (10), 2589-2597.

30. Ōmura, S.; Crump, A., Ivermectin: panacea for resource-poor communities? *Trends in Parasitology* **2014,** *30* (9), 445-455.

31. Jizba, J.; Sedmera, P.; Zima, J.; Beran, M.; Blumauerová, M.; Kandybin, N. V.; Samoukina, G. V., Macrotetrolide antibiotics produced byStreptomyces globisporus. *Folia Microbiologica* **1991,** *36* (5), 437-443.

32. Minotti, G.; Menna, P.; Salvatorelli, E.; Cairo, G.; Gianni, L., Anthracyclines: Molecular Advances and Pharmacologic Developments in Antitumor Activity and Cardiotoxicity. *Pharmacological Reviews* **2004,** *56* (2), 185-229.

33. Hollstein, U., Actinomycin. Chemistry and mechanism of action. *Chemical Reviews* **1974,** *74* (6), 625-652.

34. Sobell, H. M., Actinomycin and DNA transcription. *Proceedings of the National Academy of Sciences of the United States of America* **1985,** *82* (16), 5328-5331.

35. Rajski, S. R.; Williams, R. M., DNA Cross-Linking Agents as Antitumor Drugs. *Chemical Reviews* **1998,** *98* (8), 2723-2796.

36. Lawley, P. D.; Lethbridge, J. H.; Edwards, P. A.; Shooter, K. V., Inactivation of bacteriophage T7 by mono- and difunctional sulphur mustards in relation to cross-linking and depurination of bacteriophage DNA. *Journal of Molecular Biology* **1969,** *39* (1), 181-198.

37. Hata, T. S., Y.; Sugawara, R.; Matsumae, A.; Kanomori, K.; Shima, T.; Hoshi, T., Mitomycin, a New Antibiotic from Streptomyces. I. *J. Antibiot. Ser. A* **1956,** *141*.

38. Prakash, A. S.; Beall, H.; Ross, D.; Gibson, N. W., Sequence-selective alkylation and cross-linking induced by mitomycin C upon activation by DT-diaphorase. *Biochemistry* **1993,** *32* (21), 5518-5525.

39. Li, V. S.; Kohn, H., Studies on the bonding specificity for mitomycin C-DNA monoalkylation processes. *Journal of the American Chemical Society* **1991,** *113* (1), 275-283.

40. Kumar, S.; Lipman, R.; Tomasz, M., Recognition of specific DNA sequences by mitomycin C for alkylation. *Biochemistry* **1992,** *31* (5), 1399-1407.

41. Bush, J. A. C., C. A.; Doyle, T. W.; Nettleton, D. E.; Moseley, J. E.; Kimball, D.; Kammer, M. F.; Veitch, J., New Mitomycin Analogues Produced by Directed Biosynthesis. *J. Antibiot.* **1986,** *39*, 437.

42. Nett, M.; Ikeda, H.; Moore, B. S., Genomic basis for natural product biosynthetic diversity in the actinomycetes. *Natural Product Reports* **2009,** *26* (11), 1362-1384.

43. Hwang, K.-S.; Kim, H. U.; Charusanti, P.; Palsson, B. Ø.; Lee, S. Y., Systems biology and biotechnology of Streptomyces species for the production of secondary metabolites. *Biotechnology Advances* **2014,** *32* (2), 255-268.

44. Felnagle, E. A.; Jackson, E. E.; Chan, Y. A.; Podevels, A. M.; Berti, A. D.; McMahon, M. D.; Thomas, M. G., Nonribosomal peptide synthetases involved in the production of medically relevant natural products. *Molecular pharmaceutics* **2008,** *5* (2), 191-211.

45. Mullis, M. M.; Rambo, I. M.; Baker, B. J.; Reese, B. K., Diversity, Ecology, and Prevalence of Antimicrobials in Nature. *Frontiers in microbiology* **2019,** *10*, 2518-2518.

46. Fischbach, M. A.; Walsh, C. T., Assembly-Line Enzymology for Polyketide and Nonribosomal Peptide Antibiotics: Logic, Machinery, and Mechanisms. *Chemical Reviews* **2006,** *106* (8), 3468-3496.

47. Süssmuth, R. D.; Mainz, A., Nonribosomal Peptide Synthesis—Principles and Prospects. *Angewandte Chemie International Edition* **2017,** *56* (14), 3770-3821.

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

49. Du, L.; Lou, L., PKS and NRPS release mechanisms. *Natural Product Reports* **2010,** *27* (2), 255-278.

50. Hur, G. H.; Vickery, C. R.; Burkart, M. D., Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Natural Product Reports* **2012,** *29* (10), 1074-1098.

51. Bloudoff, K.; Schmeing, T. M., Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2017,** *1865* (11, Part B), 1587-1604.

52. Hans, M.; Hornung, A.; Dziarnowski, A.; Cane, D. E.; Khosla, C., Mechanistic Analysis of Acyl Transferase Domain Exchange in Polyketide Synthase Modules. *Journal of the American Chemical Society* **2003,** *125* (18), 5366-5374.

53. Cai, W.; Zhang, W., Engineering modular polyketide synthases for production of biofuels and industrial chemicals. *Current opinion in biotechnology* **2018,** *50*, 32-38.

54. Miyanaga, A.; Ouchi, R.; Ishikawa, F.; Goto, E.; Tanabe, G.; Kudo, F.; Eguchi, T., Structural Basis of Protein–Protein Interactions between a trans-Acting Acyltransferase and Acyl Carrier Protein in Polyketide Disorazole Biosynthesis. *Journal of the American Chemical Society* **2018,** *140* (25), 7970-7978.

55. Zhang, C.; Ke, D.; Duan, Y.; Lu, W., The Combinatorial Biosynthesis of "Unnatural" Products with Polyketides. *Transactions of Tianjin University* **2018,** *24* (6), 501-512.

56. Kotowska, M.; Pawlik, K., Roles of type II thioesterases and their application for secondary metabolite yield improvement. *Applied Microbiology and Biotechnology* **2014,** *98* (18), 7735-7746.

57. Cantu, D. C.; Chen, Y.; Reilly, P. J., Thioesterases: a new perspective based on their primary and tertiary structures. *Protein science : a publication of the Protein Society* **2010,** *19* (7), 1281-1295.

58. Horsman, M. E.; Hari, T. P. A.; Boddy, C. N., Polyketide synthase and non-ribosomal peptide synthetase thioesterase selectivity: logic gate or a victim of fate? *Natural Product Reports* **2016,** *33* (2), 183-202.

59. Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A., Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases. *Proceedings of the National Academy of Sciences* **2002,** *99* (22), 14083-14088.

60. Feng, Z.; Wang, L.; Rajski, S. R.; Xu, Z.; Coeffet-LeGal, M. F.; Shen, B., Engineered production of iso-migrastatin in heterologous Streptomyces hosts. *Bioorganic & Medicinal Chemistry* **2009,** *17* (6), 2147-2153.

61. Cobb, R. E.; Wang, Y.; Zhao, H., High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System. *ACS Synthetic Biology* **2015,** *4* (6), 723-728.

62. Wang, Y.; Cobb, R. E.; Zhao, H., Chapter Twelve - High-Efficiency Genome Editing of Streptomyces Species by an Engineered CRISPR/Cas System. In *Methods in Enzymology*, O'Connor, S. E., Ed. Academic Press: 2016; Vol. 575, pp 271-284.

63. Zeng, H.; Wen, S.; Xu, W.; He, Z.; Zhai, G.; Liu, Y.; Deng, Z.; Sun, Y., Highly efficient editing of the actinorhodin polyketide chain length factor gene in Streptomyces coelicolor M145 using CRISPR/Cas9-CodA(sm) combined system. *Applied Microbiology and Biotechnology* **2015,** *99* (24), 10575-10585.

64. Nguyen, K. T.; Ritz, D.; Gu, J.-Q.; Alexander, D.; Chu, M.; Miao, V.; Brian, P.; Baltz, R. H., Combinatorial biosynthesis of novel antibiotics related to daptomycin. *Proceedings of the National Academy of Sciences* **2006,** *103* (46), 17462-17467.

65. Baltz, R. H., Combinatorial Biosynthesis of Cyclic Lipopeptide Antibiotics: A Model for Synthetic Biology To Accelerate the Evolution of Secondary Metabolite Biosynthetic Pathways. *ACS Synthetic Biology* **2014,** *3* (10), 748-758.

66. Hata, T.; Koga, F.; Sano, Y.; Kanamori, K.; Matsumae, A.; Sugawara, R.; Hoshi, T.; Shima, T.; Ito, S.; Tomizawa, S., Carzinophilin, a new tumor inhibitory substance produced by streptomyces. I. *The Journal of antibiotics* **1954,** *7* (4), 107-112.

67. Shimada, N.; Uekusa, M.; Denda, T.; Ishii, Y.; Iizuka, T.; Sato, Y.; Hatori, T.; Fukui, M.; Sudo, M., Clinical studies of carzinophilin, an antitumor substance. *The Journal of antibiotics* **1955,** *8* (3), 67-76.

68. Nagaoka, K.; Matsumoto, M.; Oono, J.; Yokoi, K.; Ishizeki, S.; Nakashima, T., Azinomycins A and B, new antitumor antibiotics. I. Producing organism, fermentation, isolation, and characterization. *The Journal of antibiotics* **1986,** *39* (11), 1527-32.

69. Yokoi, K.; Nagaoka, K.; Nakashima, T., Azinomycins A and B, new antitumor antibiotics. II. Chemical structures. *Chem Pharm Bull (Tokyo)* **1986,** *34* (11), 4554-61.

70. Ishizeki, S.; Ohtsuka, M.; Irinoda, K.; Kukita, K.; Nagaoka, K.; Nakashima, T., Azinomycins A and B, new antitumor antibiotics. III. Antitumor activity. *The Journal of antibiotics* **1987,** *40* (1), 60-5.

71. Coleman, R. S.; Perez, R. J.; Burk, C. H.; Navarro, A., Studies on the Mechanism of Action of Azinomycin B: Definition of Regioselectivity and Sequence Selectivity of DNA Cross-Link Formation and Clarification of the Role of the Naphthoate. *Journal of the American Chemical Society* **2002,** *124* (44), 13008-13017.

72. Armstrong, R. W.; Salvati, M. E.; Nguyen, M., Novel interstrand cross-links induced by the antitumor antibiotic carzinophilin/azinomycin B. *Journal of the American Chemical Society* **1992,** *114* (8), 3144-3145.

73. Fujiwara, T.; Saito, I.; Sugiyama, H., Highly efficient DNA interstrand crosslinking induced by an antitumor antibiotic, carzinophilin. *Tetrahedron Letters* **1999,** *40* (2), 315-318.

74. Coleman, R. S.; Burk, C. H.; Navarro, A.; Brueggemeier, R. W.; Diaz-Cruz, E. S., Role of the Azinomycin Naphthoate and Central Amide in Sequence-Dependent DNA Alkylation and Cytotoxicity of Epoxide-Bearing Substructures. *Organic Letters* **2002,** *4* (20), 3545-3548.

75. Corre, C.; Lowden, P. A. S., The first biosynthetic studies of the azinomycins: acetate incorporation into azinomycin B. *Chemical Communications* **2004,** (8), 990-991.

76. Kelly, G. T.; Sharma, V.; Watanabe, C. M. H., An improved method for culturing Streptomyces sahachiroi: Biosynthetic origin of the enol fragment of azinomycin B. *Bioorganic Chemistry* **2008,** *36* (1), 4-15.

77. Sharma, V.; Kelly, G. T.; Foulke-Abel, J.; Watanabe, C. M. H., Aminoacetone as the Penultimate Precursor to the Antitumor Agent Azinomycin A. *Organic Letters* **2009,** *11* (17), 4006-4009.

78. Sharma, V.; Kelly, G. T.; Watanabe, C. M. H., Exploration of the Molecular Origin of the Azinomycin Epoxide: Timing of the Biosynthesis Revealed. *Organic Letters* **2008,** *10* (21), 4815-4818.

79. Zhao, Q.; He, Q.; Ding, W.; Tang, M.; Kang, Q.; Yu, Y.; Deng, W.; Zhang, Q.; Fang, J.; Tang, G.; Liu, W., Characterization of the Azinomycin B Biosynthetic Gene Cluster Revealing a Different Iterative Type I Polyketide Synthase for Naphthoate Biosynthesis. *Chemistry & Biology* **2008,** *15* (7), 693-705.

80. Mori, S.; Simkhada, D.; Zhang, H.; Erb, M. S.; Zhang, Y.; Williams, H.; Fedoseyenko, D.; Russell, W. K.; Kim, D.; Fleer, N.; Ealick, S. E.; Watanabe, C. M. H.,

Polyketide Ring Expansion Mediated by a Thioesterase, Chain Elongation and Cyclization Domain, in Azinomycin Biosynthesis: Characterization of AziB and AziG. *Biochemistry* **2016,** *55* (4), 704-714.

81. Vederas, J. C., The use of stable isotopes in biosynthetic studies. *Nat Prod Rep* **1987,** *4* (3), 277-337.

82. Watanabe, C. M. H.; Townsend, C. A., Incorporation of molecular oxygen in aflatoxin B-1 biosynthesis. *J. Org. Chem.* **1996,** *61* (6), 1990-1993.

83. Ding, W.; Deng, W.; Tang, M.; Zhang, Q.; Tang, G.; Bi, Y.; Liu, W., Biosynthesis of 3-Methoxy-5-Methyl Naphthoic Acid and its Incorporation into the Anti-tumor Antibiotic Azinomycin B. *Mol. BioSys.* **2010,** *6*, 1071-1081.

84. Schofield, C. J.; Zhang, Z., Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Current Opinion in Structural Biology* **1999,** *9* (6), 722-731.

85. Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J., Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *Journal of Inorganic Biochemistry* **2006,** *100* (4), 644-669.

86. Martinez, S.; Hausinger, R. P., Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-dependent Oxygenases. *Journal of Biological Chemistry* **2015,** *290* (34), 20702-20711.

87. Herr, C. Q.; Hausinger, R. P., Amazing Diversity in Biochemical Roles of Fe(II)/2-Oxoglutarate Oxygenases. *Trends in Biochemical Sciences* **2018,** *43* (7), 517-532.

88. Xu, Q.; Grant, J.; Chiu, H.-J.; Farr, C. L.; Jaroszewski, L.; Knuth, M. W.; Miller, M. D.; Lesley, S. A.; Godzik, A.; Elsliger, M.-A.; Deacon, A. M.; Wilson, I. A., Crystal structure of a member of a novel family of dioxygenases (PF10014) reveals a conserved cupin fold and active site. *Proteins: Structure, Function, and Bioinformatics* **2014,** *82* (1), 164-170.

89. Meluzzi, D.; Zheng, W. H.; Hensler, M.; Nizet, V.; Dorrestein, P. C., Top-down mass spectrometry on low-resolution instruments: characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways. *Bioorganic & medicinal chemistry letters* **2008,** *18* (10), 3107-3111.

90. Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L., Facile Detection of Acyl- and Peptidyl- intermediates on Thiotemplate Carrier Domains via Phosphopantetheinyl Elimination Reactions During Tandem Mass Spectrometry. *Biochemistry* **2006,** *45* (42), 12756-12766.

91. Johnson, D. W., Contemporary clinical usage of LC/MS: analysis of biologically important carboxylic acids. *Clinical biochemistry* **2005,** *38* (4), 351-361.

92. Belecki, K.; Townsend, C. A., Biochemical Determination of Enzyme-Bound Metabolites: Preferential Accumulation of a Programmed Octaketide on the Enediyne Polyketide Synthase CalE8. *Journal of the American Chemical Society* **2013,** *135* (38), 14339-14348.

93. Patteson, J. B.; Dunn, Z. D.; Li, B., In Vitro Biosynthesis of the Nonproteinogenic Amino Acid Methoxyvinylglycine. *Angewandte Chemie International Edition* **2018,** *57* (23), 6780-6785.

94. Camarero, J. A.; Muir, T. W., Native Chemical Ligation of Polypeptides. In *Current Protocols in Protein Science*, John Wiley & Sons, Inc.: 2001.

95. Brotherton, C. A.; Balskus, E. P., A Prodrug Resistance Mechanism Is Involved in Colibactin Biosynthesis and Cytotoxicity. *Journal of the American Chemical Society* **2013,** *135* (9), 3359-3362.

96. Reimer, D.; Pos, K. M.; Thines, M.; Grün, P.; Bode, H. B., A Natural Prodrug Activation Mechanism in Nonribosomal Peptide Synthesis. *Nature Chemical Biology* **2011,** *7* (12), 888-890.

97. Dubois, D.; Baron, O.; Cougnoux, A.; Delmas, J.; Pradel, N.; Boury, M.; Bouchon, B.; Bringer, M.-A.; Nougayrède, J.-P.; Oswald, E.; Bonnet, R., ClbP Is a Prototype of a Peptidase Subgroup Involved in Biosynthesis of Nonribosomal Peptides. *Journal of Biological Chemistry* **2011,** *286* (41), 35562-35570.

98. Kelly, G. T.; Washburn, L. A.; Watanabe, C. M. H., The Fate of Molecular Oxygen in Azinomycin Biosynthesis. *The Journal of Organic Chemistry* **2019,** *84* (5), 2991-2996.

99. Magarvey, N. A.; Ehling-Schulz, M.; Walsh, C. T., Characterization of the Cereulide NRPS α-Hydroxy Acid Specifying Modules: Activation of α-Keto Acids and Chiral Reduction on the Assembly Line. *Journal of the American Chemical Society* **2006,** *128* (33), 10698-10699.

100. Fujimori, D. G.; Hrvatin, S.; Neumann, C. S.; Strieker, M.; Marahiel, M. A.; Walsh, C. T., Cloning and characterization of the biosynthetic gene cluster for kutznerides. *Proceedings of the National Academy of Sciences of the United States of America* **2007,** *104* (42), 16498-16503.

101. Mori, S.; Nepal, K. K.; Kelly, G. T.; Sharma, V.; Simkhada, D.; Gowda, V.; Delgado, D.; Watanabe, C. M. H., Priming of Azabicycle Biosynthesis in the Azinomycin Class of Antitumor Agents. *Biochemistry* **2017,** *56* (6), 805-808.

102. Udwary, D. W.; Merski, M.; Townsend, C. A., A Method for Prediction of the Locations of Linker Regions within Large Multifunctional Proteins, and Application to a Type I Polyketide Synthase. *Journal of Molecular Biology* **2002,** *323* (3), 585-598.

103. Herbst, D. A.; Jakob, R. P.; Zähringer, F.; Maier, T., Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **2016,** *531*, 533.

104. Yin, J.; Liu, F.; Li, X.; Walsh, C. T., Labeling Proteins with Small Molecules by Site-Specific Posttranslational Modification. *Journal of the American Chemical Society* **2004,** *126* (25), 7754-7755.

105. Ding, W.; Deng, W.; Tang, M.; Zhang, Q.; Tang, G.; Bi, Y.; Liu, W., Biosynthesis of 3-methoxy-5-methyl naphthoic acid and its incorporation into the antitumor antibiotic azinomycin B. *Molecular BioSystems* **2010,** *6* (6), 1071-1081.

106. Hasebe, F.; Matsuda, K.; Shiraishi, T.; Futamura, Y.; Nakano, T.; Tomita, T.; Ishigami, K.; Taka, H.; Mineki, R.; Fujimura, T.; Osada, H.; Kuzuyama, T.; Nishiyama, M., Amino-group carrier-protein-mediated secondary metabolite biosynthesis in Streptomyces. *Nature Chemical Biology* **2016,** *12* (11), 967-972.

107. Nepal, K. K.; Lee, R. P.; Rezenom, Y. H.; Watanabe, C. M. H., Probing the Role of N-Acetyl-glutamyl 5-Phosphate, an Acyl Phosphate, in the Construction of the Azabicycle Moiety of the Azinomycins. *Biochemistry* **2015,** *54* (29), 4415-4418.

108. Schenk, G.; Duggleby, R. G.; Nixon, P. F., Properties and functions of the thiamin diphosphate dependent enzyme transketolase. *The International Journal of Biochemistry & Cell Biology* **1998,** *30* (12), 1297-1318.

109. Markert, B.; Stolzenberger, J.; Brautaset, T.; Wendisch, V. F., Characterization of two transketolases encoded on the chromosome and the plasmid pBM19 of the facultative ribulose monophosphate cycle methylotroph Bacillus methanolicus. *BMC microbiology* **2014,** *14*, 7-7.

110. Smith, M. E. B.; Kaulmann, U.; Ward, J. M.; Hailes, H. C., A colorimetric assay for screening transketolase activity. *Bioorganic & Medicinal Chemistry* **2006,** *14* (20), 7062-7065.

111. Solovjeva, O. N.; Kochetov, G. A., Inhibition of transketolase by p-hydroxyphenylpyruvate. *FEBS Letters* **1999,** *462* (3), 246-248.

112. Hazra, A. B.; Han, Y.; Chatterjee, A.; Zhang, Y.; Lai, R.-Y.; Ealick, S. E.; Begley, T. P., A Missing Enzyme in Thiamin Thiazole Biosynthesis: Identification of TenI as a Thiazole Tautomerase. *Journal of the American Chemical Society* **2011,** *133* (24), 9311-9319.

113. Lüdtke, S.; Neumann, P.; Erixon, K. M.; Leeper, F.; Kluger, R.; Ficner, R.; Tittmann, K., Sub-ångström-resolution crystallography reveals physical distortions that enhance reactivity of a covalent enzymatic intermediate. *Nature Chemistry* **2013,** *5* (9), 762-767.

114. Mitschke, L.; Parthier, C.; Schröder-Tittmann, K.; Coy, J.; Lüdtke, S.; Tittmann, K., The Crystal Structure of Human Transketolase and New Insights into Its Mode of Action. *Journal of Biological Chemistry* **2010,** *285* (41), 31559-31570.

115. Joshi, S.; Singh, A. R.; Kumar, A.; Misra, P. C.; Siddiqi, M. I.; Saxena, J. K., Molecular cloning and characterization of Plasmodium falciparum transketolase. *Molecular and Biochemical Parasitology* **2008,** *160* (1), 32-41.

116. Aymard, C. M. G.; Halma, M.; Comte, A.; Mousty, C.; Prévot, V.; Hecquet, L.; Charmantray, F.; Blum, L. J.; Doumèche, B., Innovative Electrochemical Screening Allows Transketolase Inhibitors to Be Identified. *Analytical Chemistry* **2018,** *90* (15), 9241-9248.

117. Bennett, B. D.; Kimball, E. H.; Gao, M.; Osterhout, R.; Van Dien, S. J.; Rabinowitz, J. D., Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature Chemical Biology* **2009,** *5* (8), 593-599.

118. Horie, A.; Tomita, T.; Saiki, A.; Kono, H.; Taka, H.; Mineki, R.; Fujimura, T.; Nishiyama, C.; Kuzuyama, T.; Nishiyama, M., Discovery of proteinaceous N-modification in lysine biosynthesis of Thermus thermophilus. *Nature Chemical Biology* **2009,** *5*, 673.

119. Ouchi, T.; Tomita, T.; Horie, A.; Yoshida, A.; Takahashi, K.; Nishida, H.; Lassak, K.; Taka, H.; Mineki, R.; Fujimura, T.; Kosono, S.; Nishiyama, C.; Masui, R.; Kuramitsu, S.; Albers, S.-V.; Kuzuyama, T.; Nishiyama, M., Lysine and arginine biosyntheses mediated by a common carrier protein in Sulfolobus. *Nature Chemical Biology* **2013,** *9*, 277.

120. Yoshida, A.; Tomita, T.; Atomi, H.; Kuzuyama, T.; Nishiyama, M., Lysine Biosynthesis of Thermococcus kodakarensis with the Capacity to Function as an Ornithine Biosynthetic System. *Journal of Biological Chemistry* **2016,** *291* (41), 21630-21643.

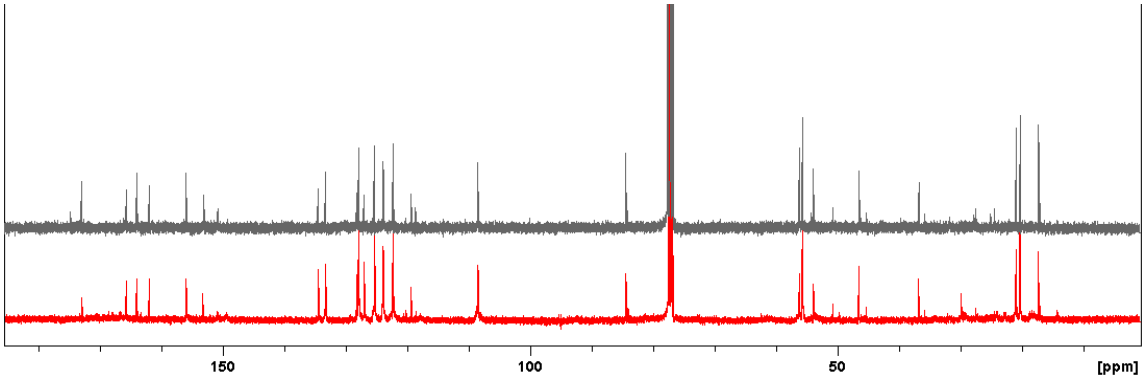
121. Shimizu, T.; Tomita, T.; Kuzuyama, T.; Nishiyama, M., Crystal Structure of the LysY·LysW Complex from Thermus thermophilus. *Journal of Biological Chemistry* **2016,** *291* (19), 9948-9959.

122. Yoshida, A.; Tomita, T.; Fujimura, T.; Nishiyama, C.; Kuzuyama, T.; Nishiyama, M., Structural Insight into Amino Group-carrier Protein-mediated Lysine Biosynthesis: Crystal Structure of the LysZ·LyzW Complex From Thermus Thermophilus *Journal of Biological Chemistry* **2015,** *290* (1), 435-447.

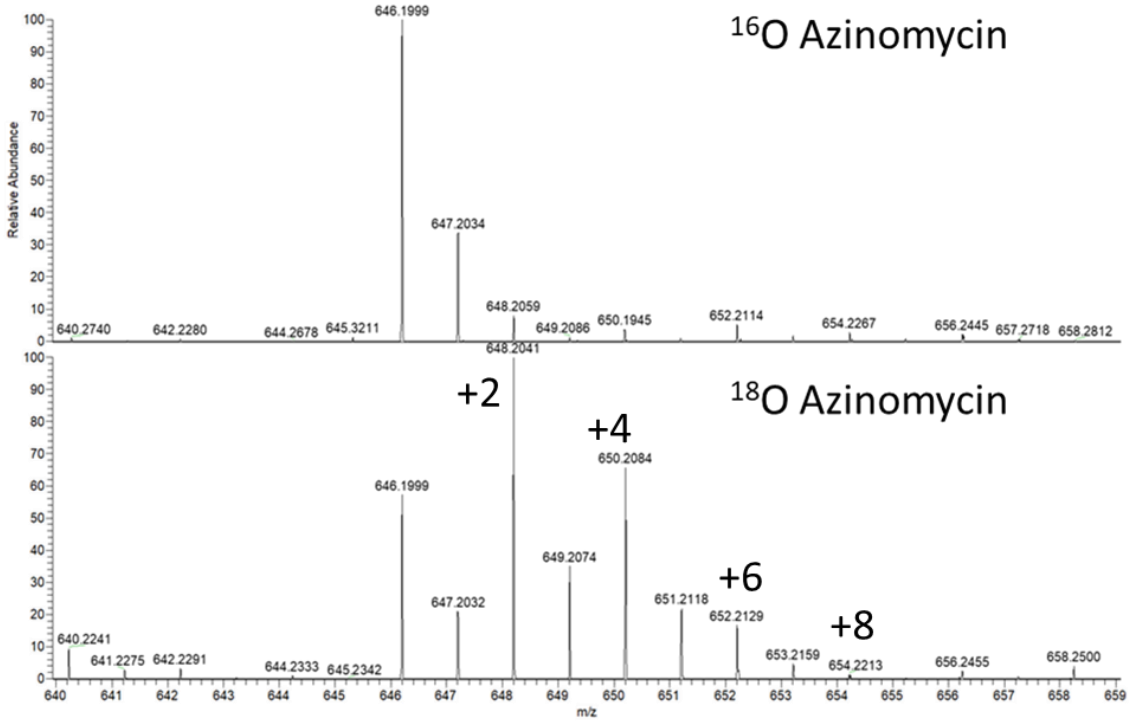
123. Washburn, L. A.; Foley, B.; Martinez, F.; Lee, R. P.; Pryor, K.; Rimes, E.; Watanabe, C. M. H., Transketolase Activity in the Formation of the Azinomycin Azabicycle Moiety. *Biochemistry* **2019,** *58* (52), 5255-5258.

APPENDIX A

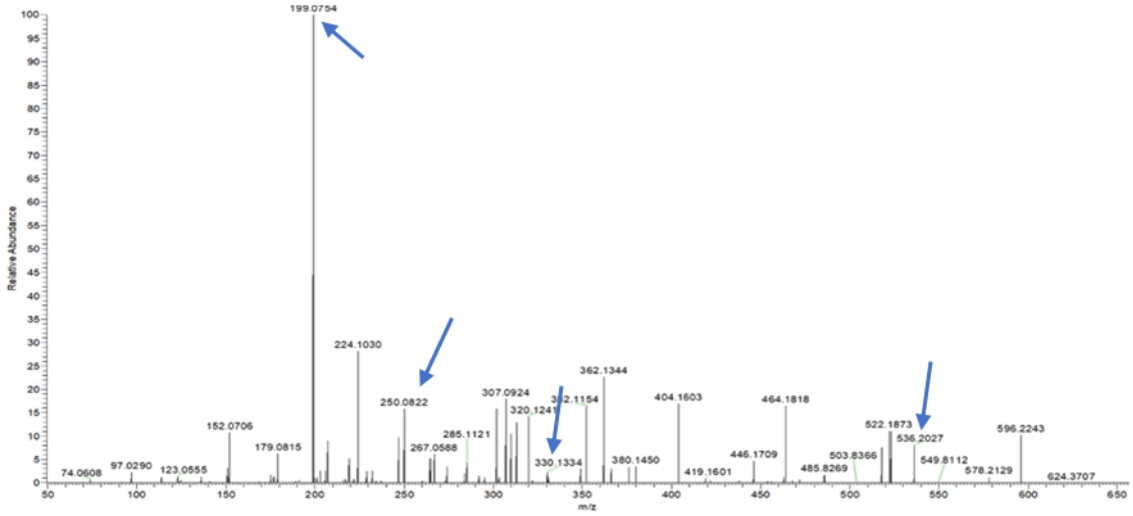
FIGURES AND TABLES



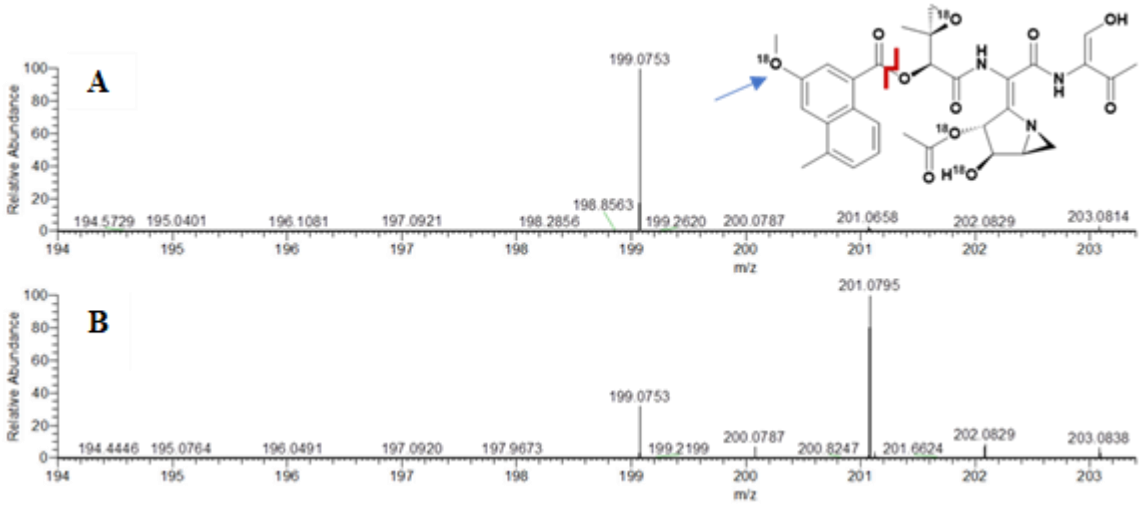
Appendix Figure 1 ^{13}C NMR Comparing Natural and $^{18}O_2$ Incorporated Azinomycin B[98]
Top = natural. Bottom= $^{18}O_2$



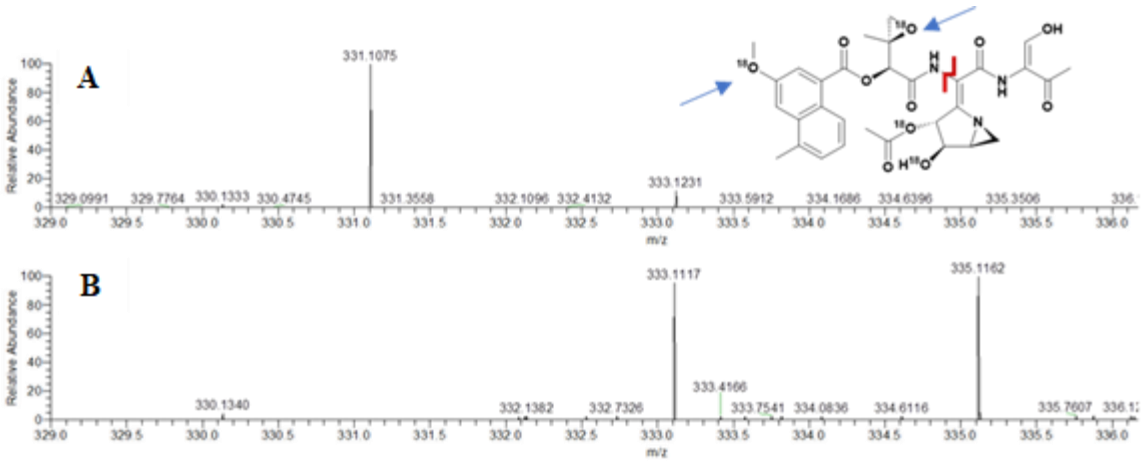
Appendix Figure 2 ESI-MS ^{16}O-Azinomycin B vs ^{18}O-Azinomycin B[98]
Azinomycin B [M+Na] = 646.20



Appendix Figure 3 ESI-MS/MS of Azinomycin from Crude Extract[98]
Azinomycin A (M+H = 596.22). Arrows indicated fragments used to confirm ^{18}O incorporation locations.

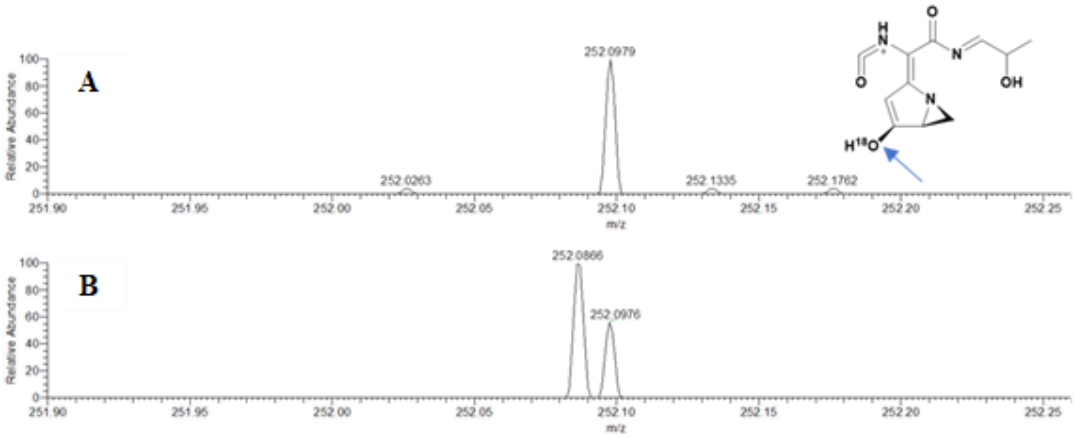


Appendix Figure 4 $^{18}O_2$ Incorporation into the Azinomycin Naphthoate Detected by ESI-MS/MS[98]
A. MS/MS of natural azinomycin (Parent ion of azinomycin B + H = 624.22) showing naphthoate fragment (M+H = 199.07) **B.** MS/MS of $^{18}O_2$ azinomycin (Parent ion of azinomycin + H + 3(^{18}O) = 630.23) with presence of ^{18}O on naphthoate fragment (M+H+^{18}O = 201.08). Red line indicating fragmentation.



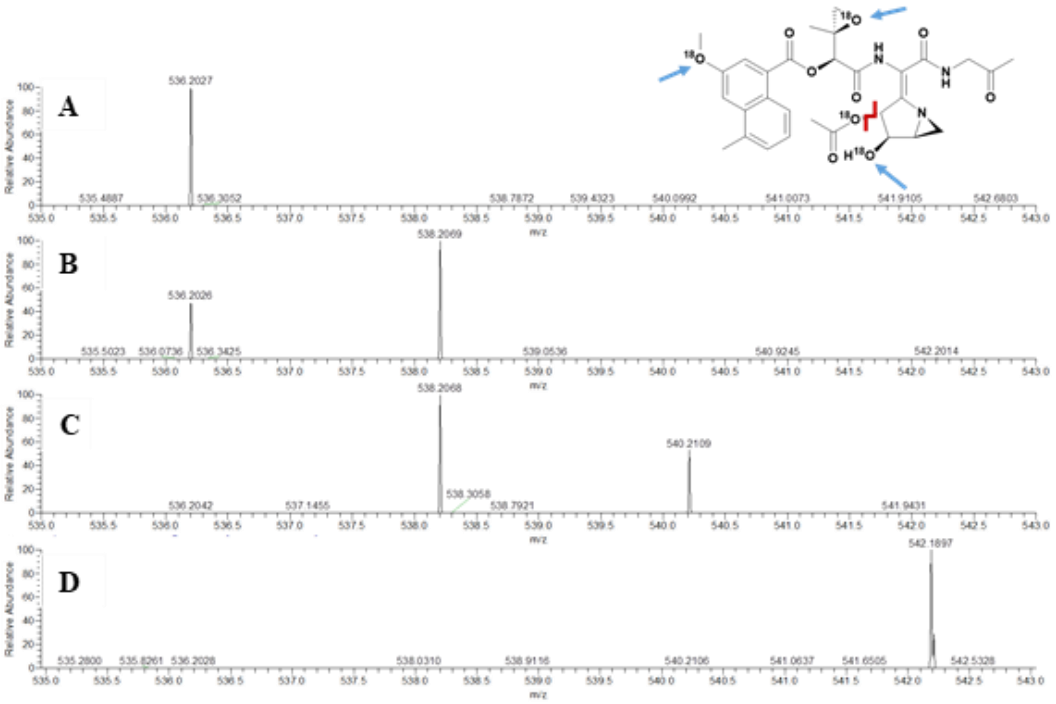
Appendix Figure 5 $^{18}O_2$ Incorporation into the Azinomycin Epoxyamide Detected by ESI-MS/MS[98]

A. Epoxyamide of azinomycin (Parent ion of azinomycin B + H = 624.22) detected by ESI-MS/MS (M+H = 331.10) in natural sample **B**. Epoxyamide of azinomycin in $^{18}O_2$ sample (Parent ion of azinomycin + H + 3(^{18}O) = 630.23) detected by ESI-MS/MS (M+H+^{18}O = 333.11, M+H+2(^{18}O) = 335.12). Red line indicates fragmentation.



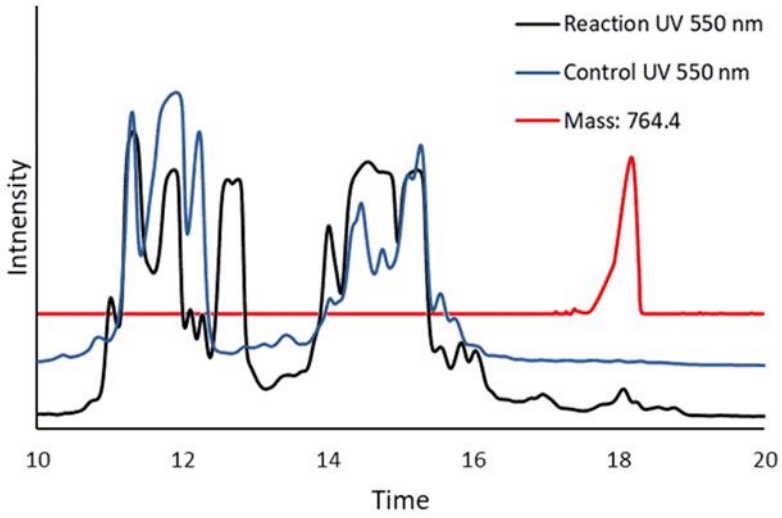
Appendix Figure 6 $^{18}O_2$ Incorporation on C-12 Hydroxyl Detected by ESI-MS/MS[98]

A. Fragment detected by ESI-MS/MS (M+H = 250.08) in natural sample. **B**. Fragment detected in $^{18}O_2$ sample detected by ESI-MS/MS (M+H+^{18}O = 252.08)

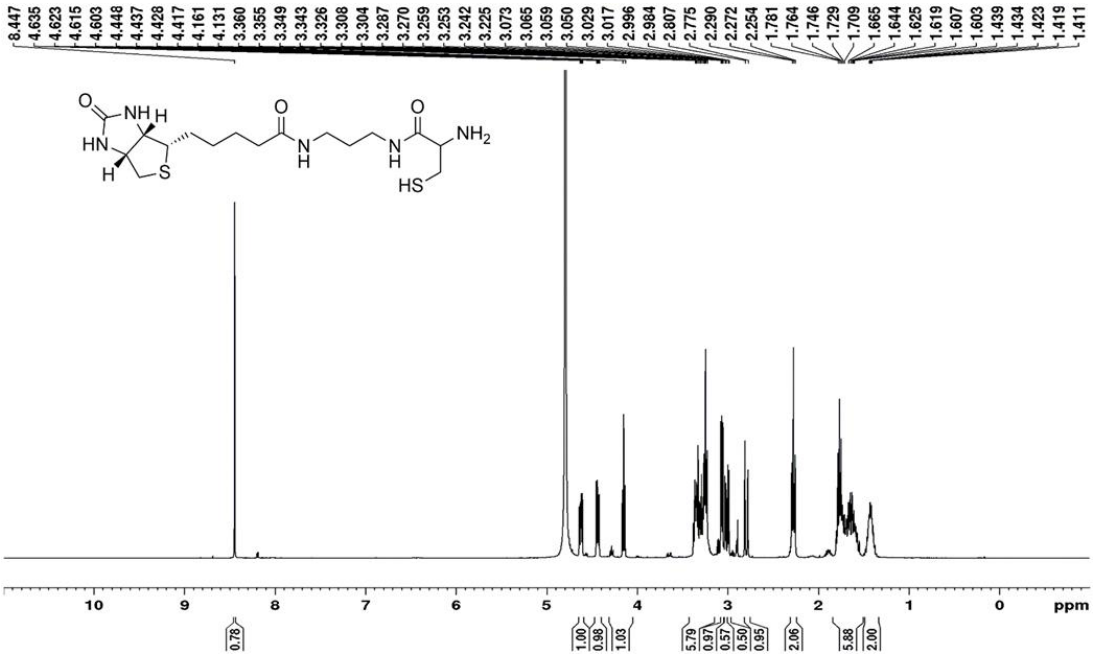


Appendix Figure 7 $^{18}O_2$ Incorporation into Azinomycin O-Acetyl Fragment Detected by ESI-MS/MS[98]

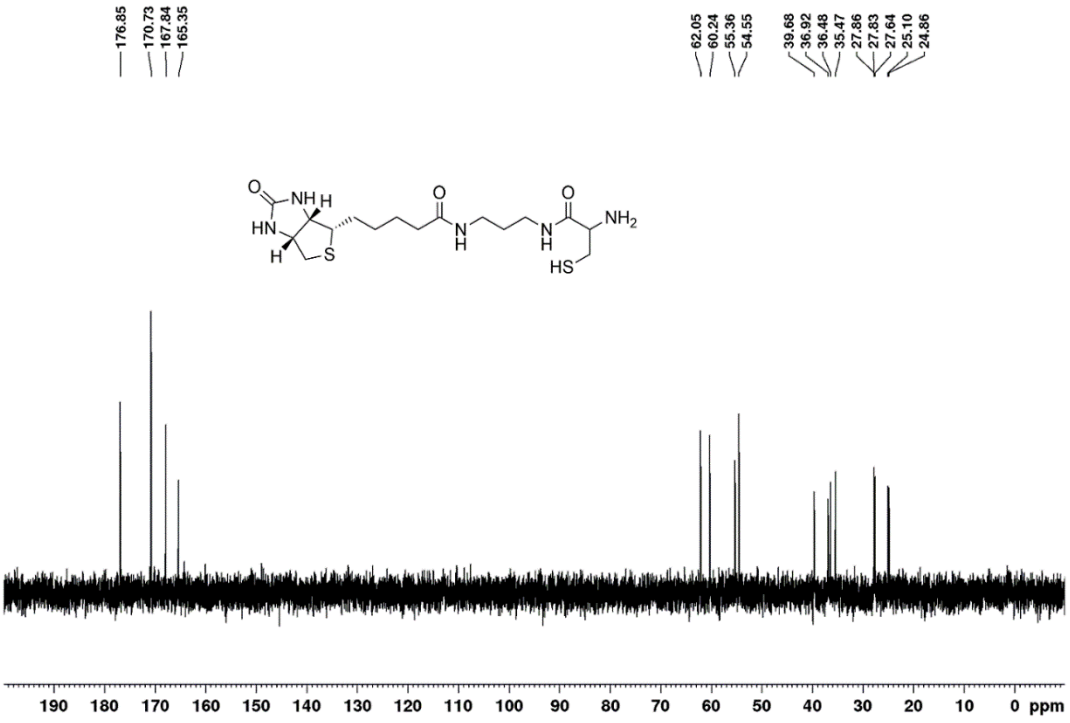
A. MS/MS of natural azinomycin A 596.22 species with O-Acetyl fragmentation in ESI-MS/MS (M+H = 536.20). **B**. MS/MS of azinomycin A 598.22 species (Azinomycin A + 1(^{18}O)) with ^{18}O-Acetyl fragmentation in ESI-MS/MS (M+H+^{18}O = 538.20) in $^{18}O_2$ sample. **C.** MS/MS of azinomycin A 600.22 species (Azinomycin A + 2(^{18}O)) with ^{18}O-Acetyl fragmentation in ESI-MS/MS (M+H+2(^{18}O) = 540.21) in $^{18}O_2$ sample. **D.** MS/MS of azinomycin A 602.22 species (Azinomycin A + 3(^{18}O)) with ^{18}O-Acetyl fragmentation in ESI-MS/MS (M+H+3(^{18}O) = 542.22) in $^{18}O_2$ sample. M+H+4(^{18}O) is not seen as the fourth position is fragmented. Red line indicates fragmentation.



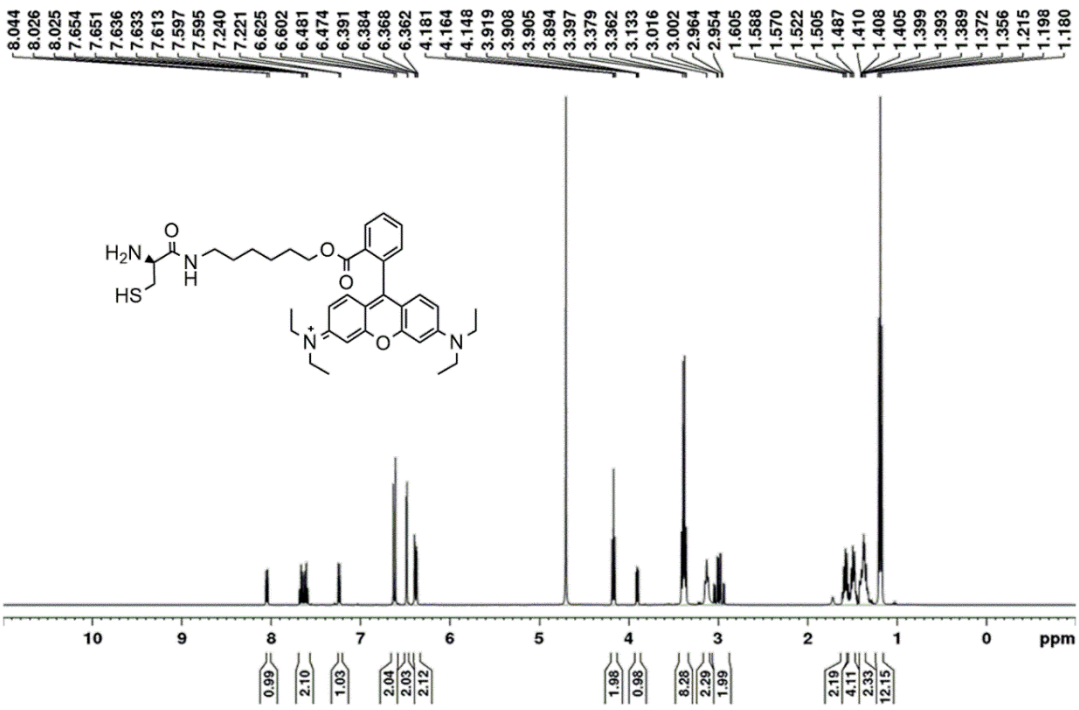
Appendix Figure 8 AziB + Rhodamine-Cys LC-MS Analysis



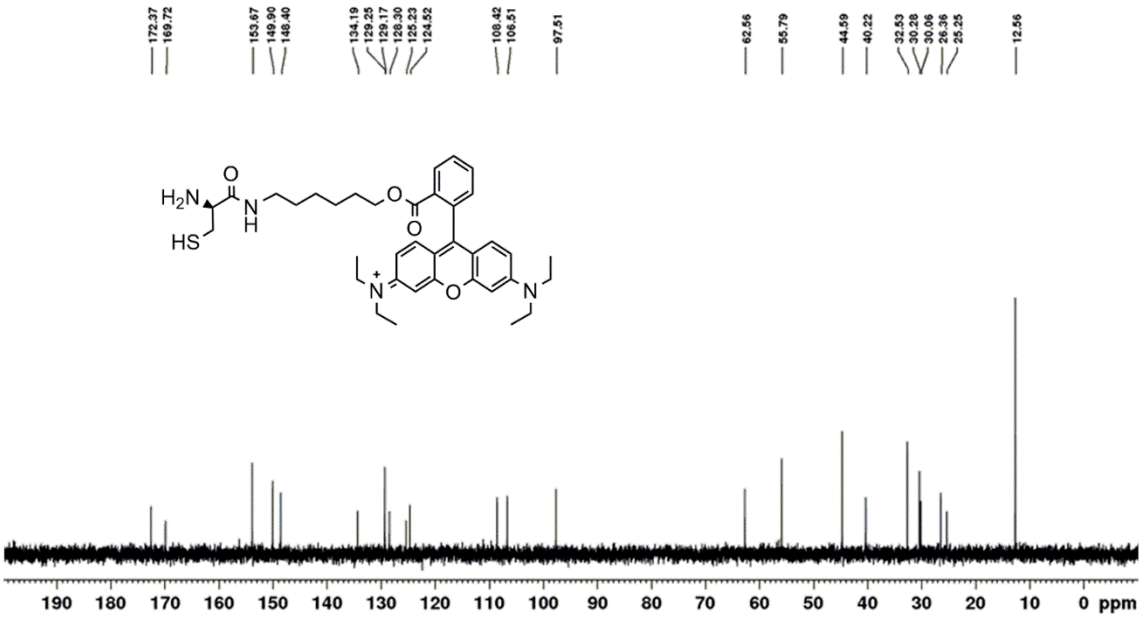
Appendix Figure 9 ^{1}H NMR of Biotin-Cys in D_2O.



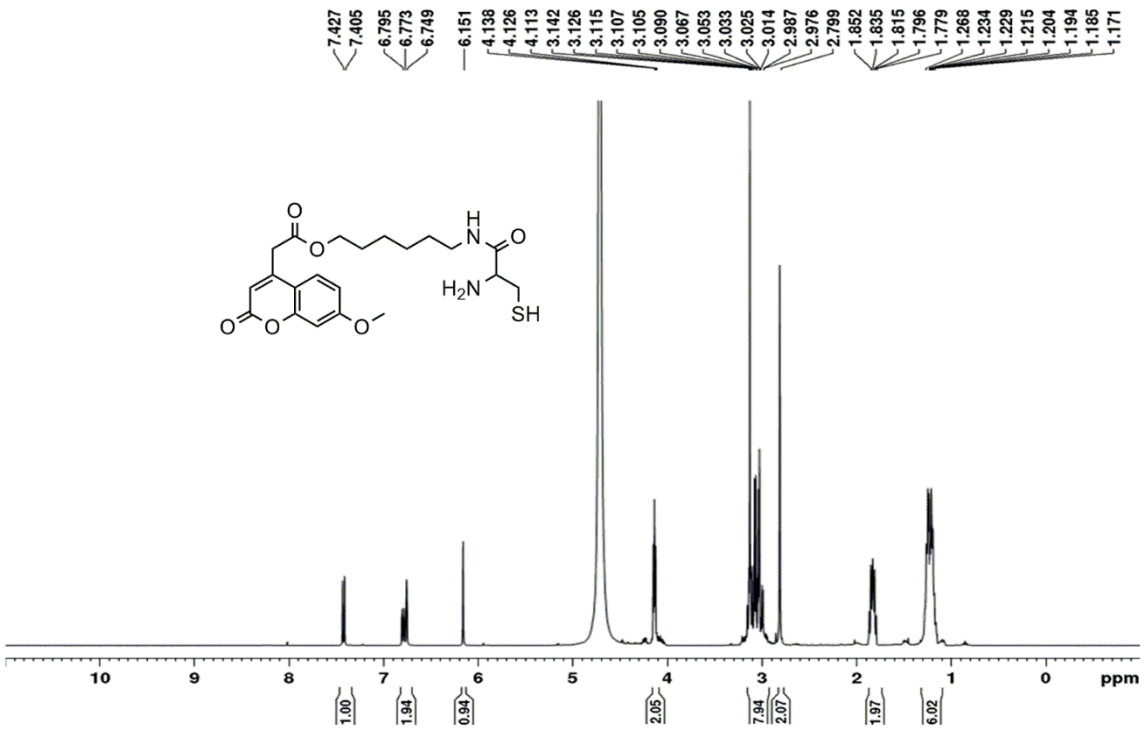
Appendix Figure 10 ^{13}C{^{1}H} NMR of Biotin-Cys in D_2O



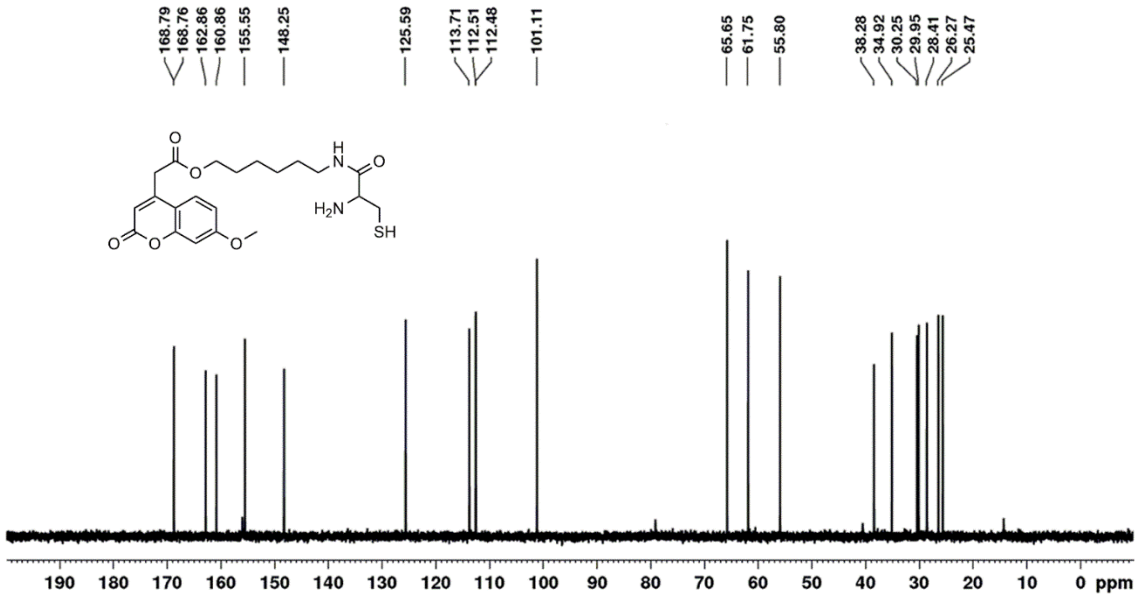
Appendix Figure 11 ^{1}H NMR of Rhodamine-Cys in D_2O



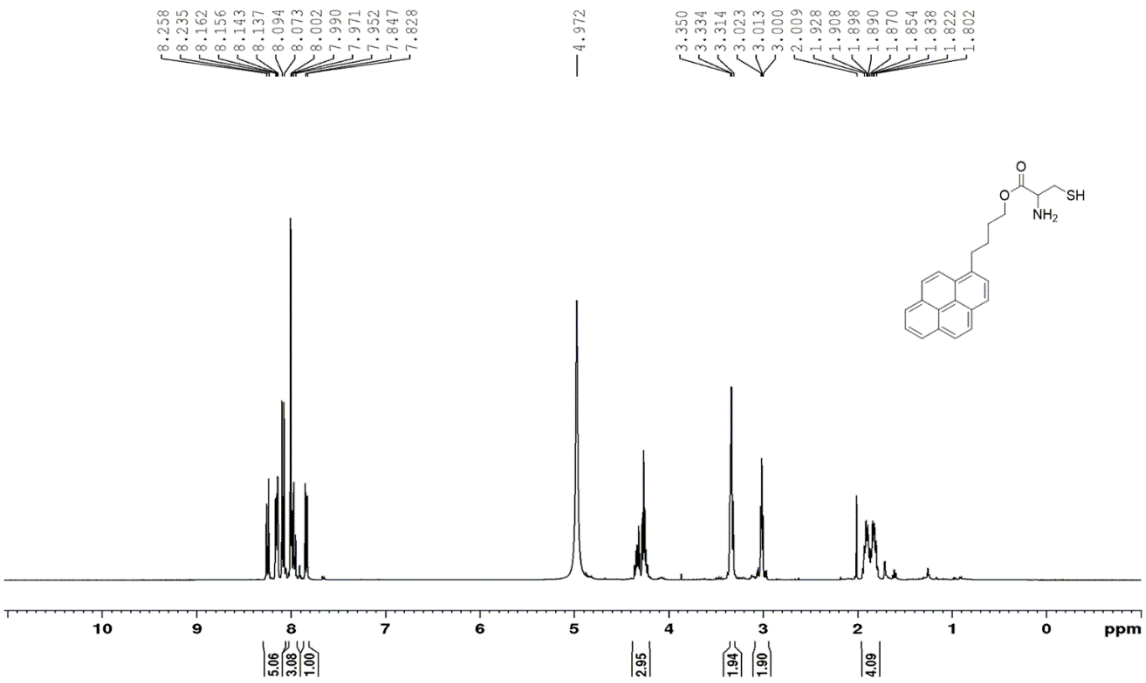
Appendix Figure 12 ^{13}C{^{1}H} NMR of Rhodamine-Cys in D_2O



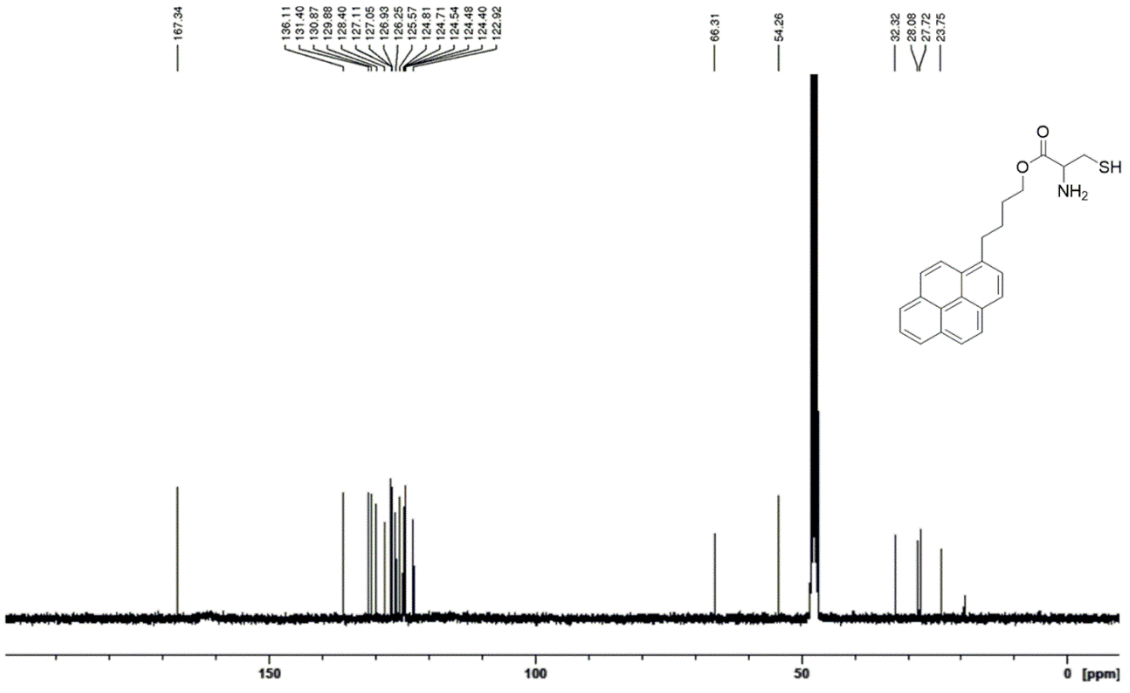
Appendix Figure 13 ^{1}H NMR of Coumarin-Cys in D_2O



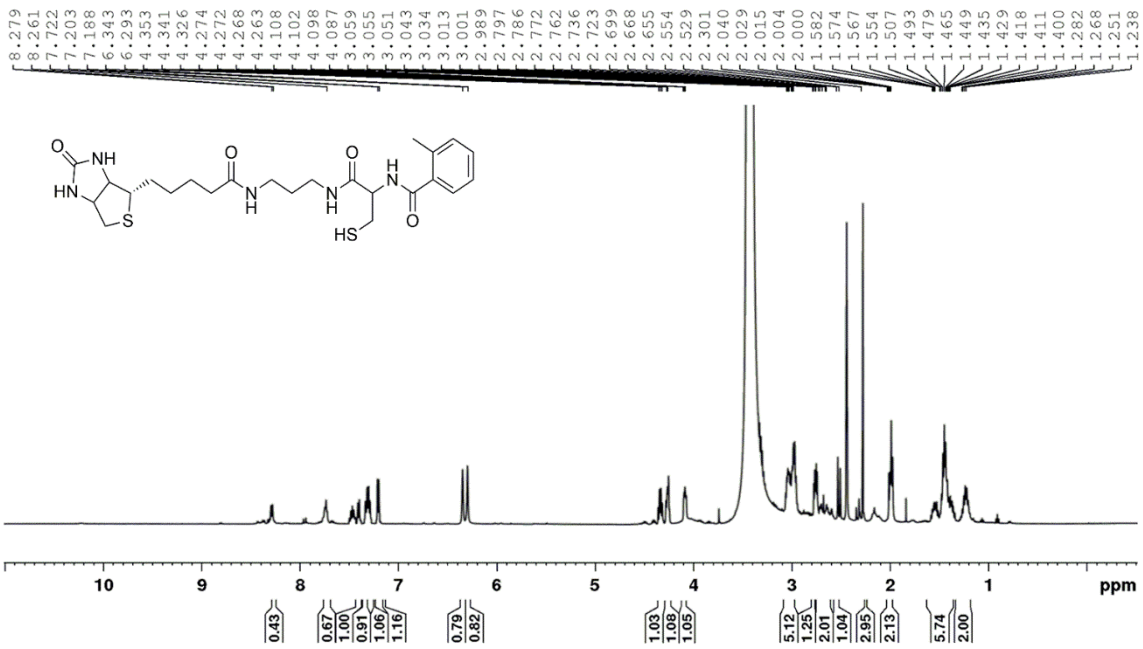
Appendix Figure 14 ^{13}C{^{1}H} NMR of Coumarin-Cys in D_2O



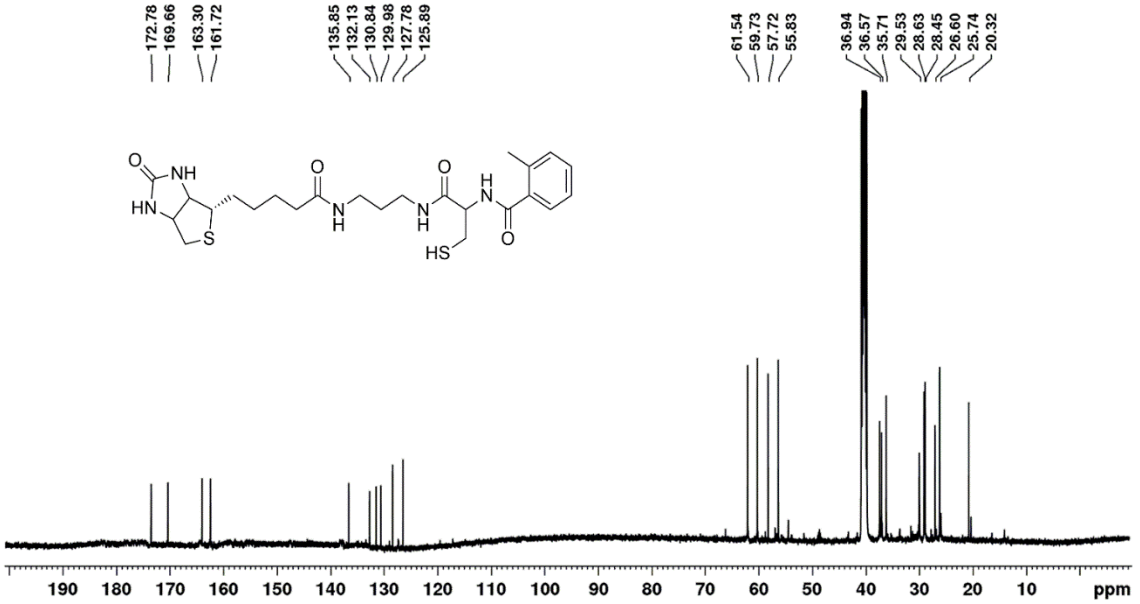
Appendix Figure 15 ^{1}H NMR of Pyrene-Cys in CD_3OD



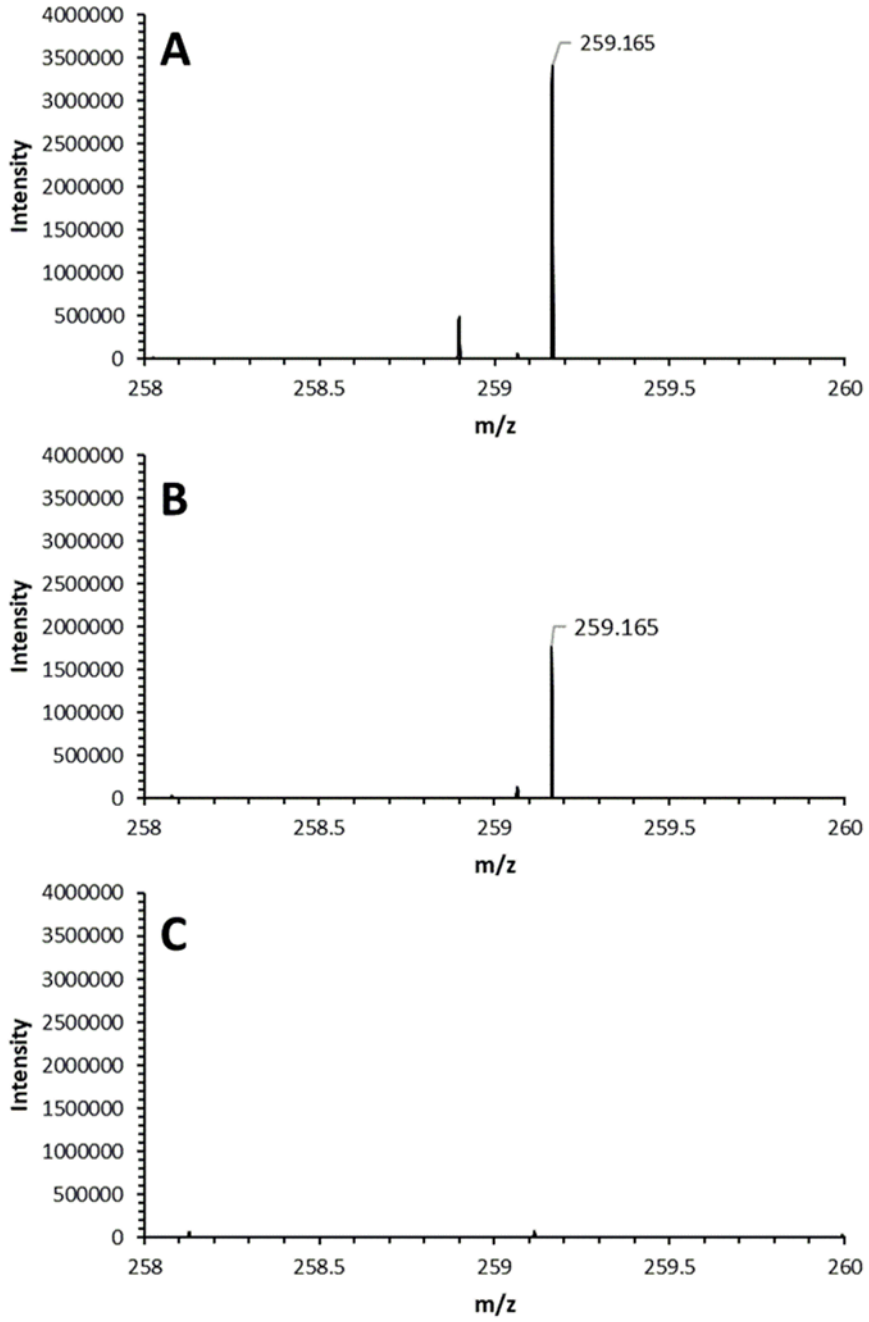
Appendix Figure 16 ^{13}C{^{1}H} NMR of Pyrene-Cys in CD_3OD



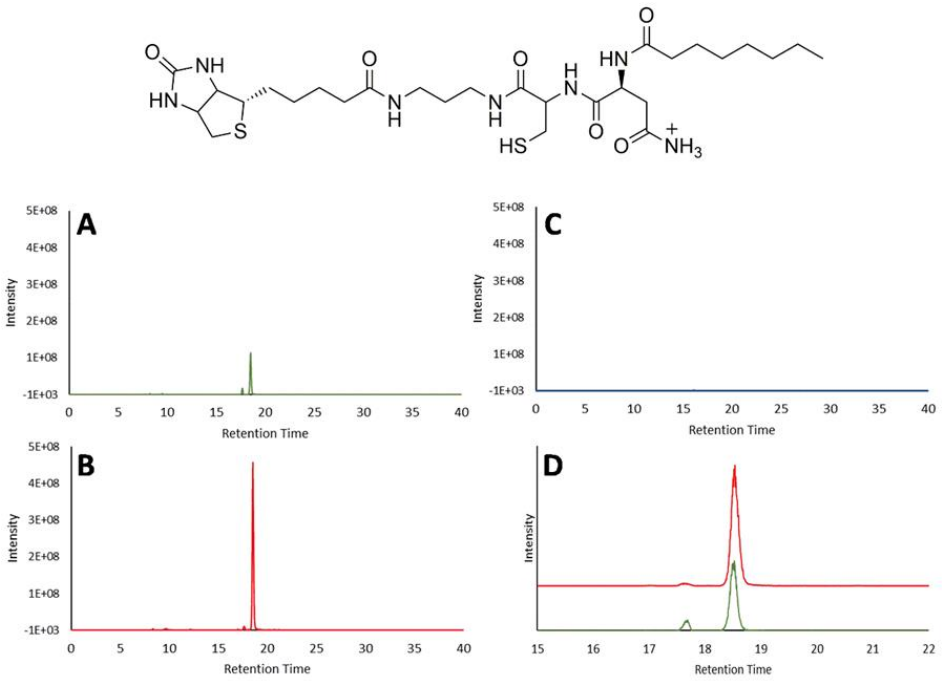
Appendix Figure 17 ^{1}H NMR of Biotin-Cys-2-methylbenzoic acid in DMSO-D$_6$



Appendix Figure 18 ^{13}C{^{1}H} NMR of Biotin-Cys-2-methylbenzoic acid in DMSO-D$_6$

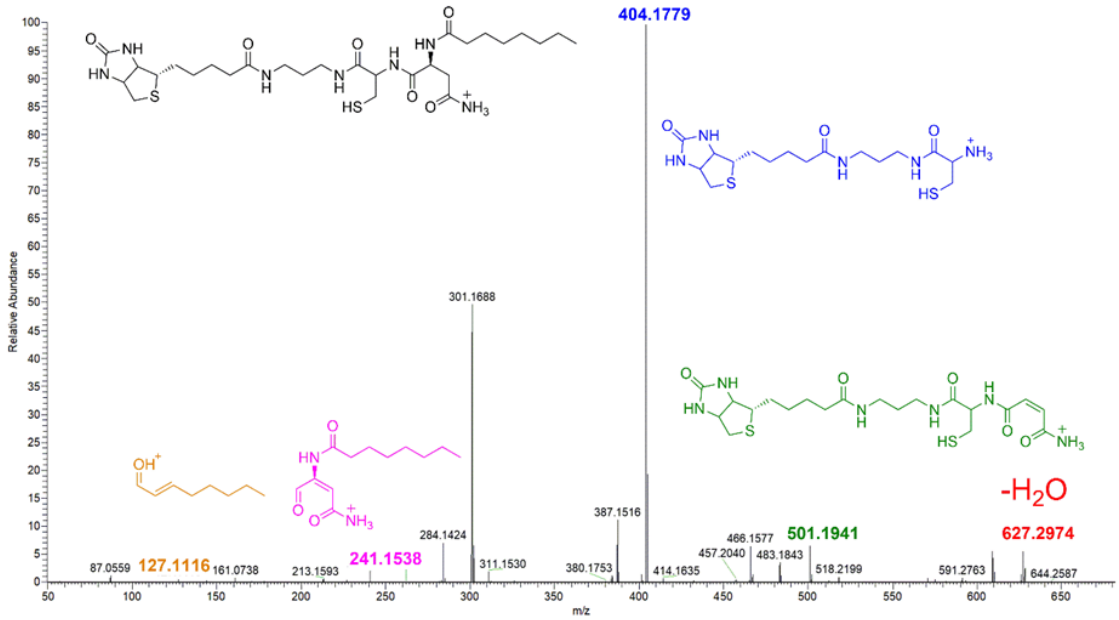


Appendix Figure 19 ClbN Hydrolysis Control Reaction MS
A. ClbN Reaction. **B**. Synthetic Standard of octanoyl-asparagine. **C**. Negative Control lacking ClbN protein. Expected product mass of octanoyl-asparagine m/z $[M+H]^+$ calcd for $C_{12}H_{22}N_2O_4$; 259.165.



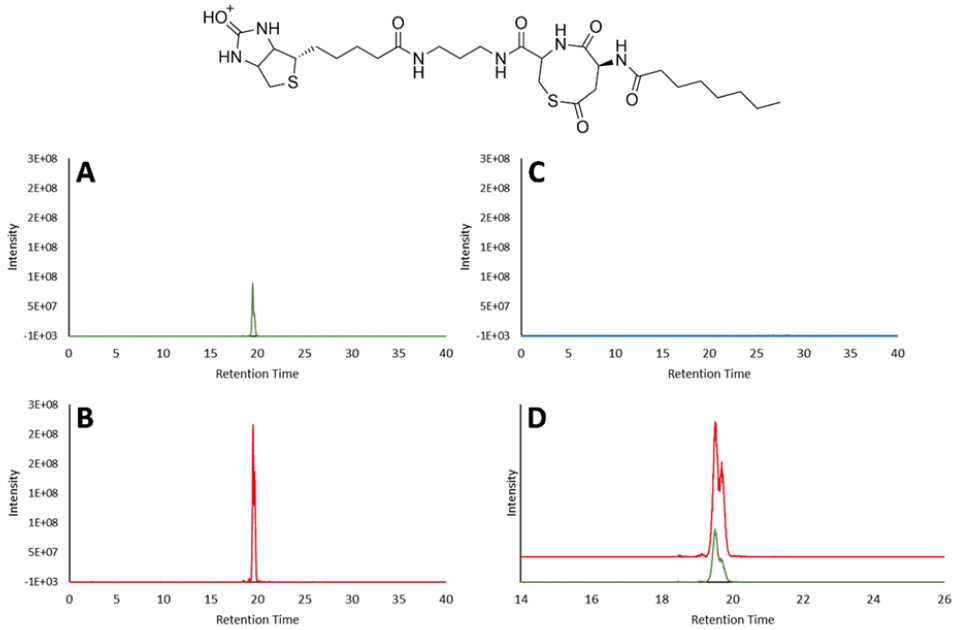
Appendix Figure 20 ClbN LC-MS Traces Free Amide Product

Extracted Ion Chromatograms (EIC) for 644.325. Expected product mass *m/z* $[M+H]^+$ calcd for $C_{28}H_{49}N_7O_6S_2$; 644.325. **A**. Biotin-Cys + ClbN Reaction. **B**. Synthetic Standard of free amide product **C**. Negative Control lacking ClbN protein. **D**. Overlay of EIC for Biotin-Cys + ClbN Reaction and Synthetic Standard.

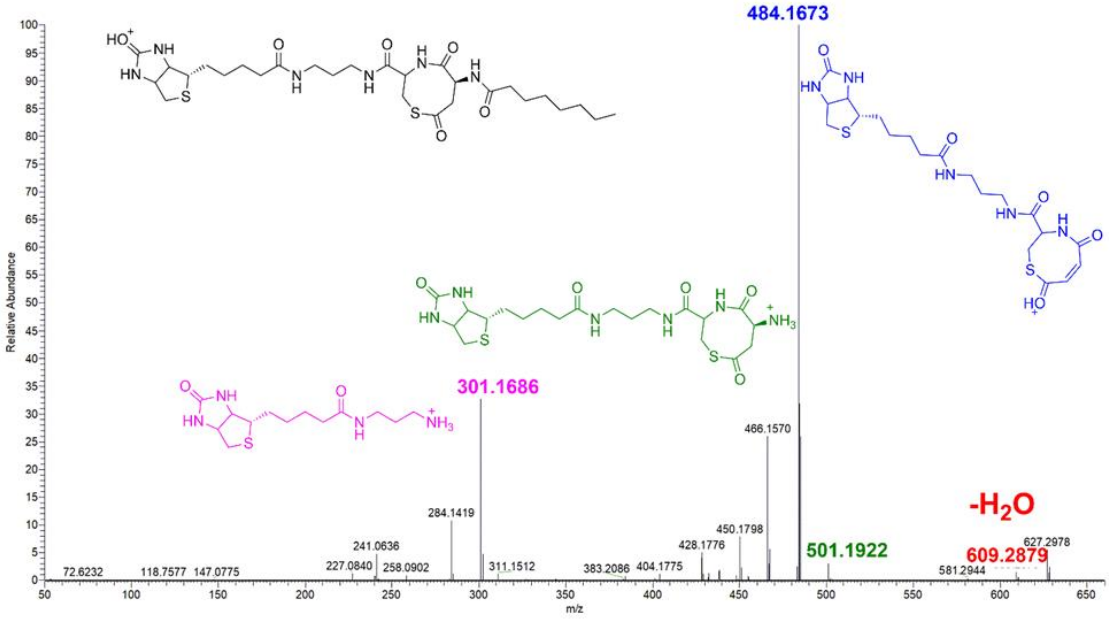


Appendix Figure 21 MS/MS of ClbN Free Amide Product

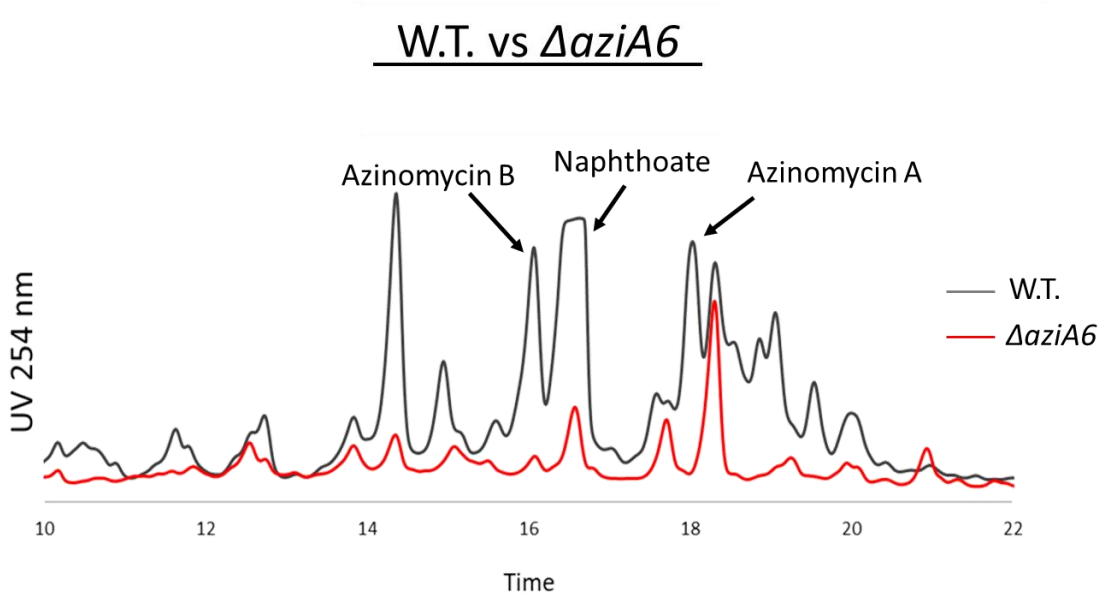
MS/MS Fragmentation for 644.324 Species. Expected product m/z $[M+H]^+$ calcd for $C_{28}H_{49}N_7O_6S_2$; 644.325. Corresponding fragments are color coded.



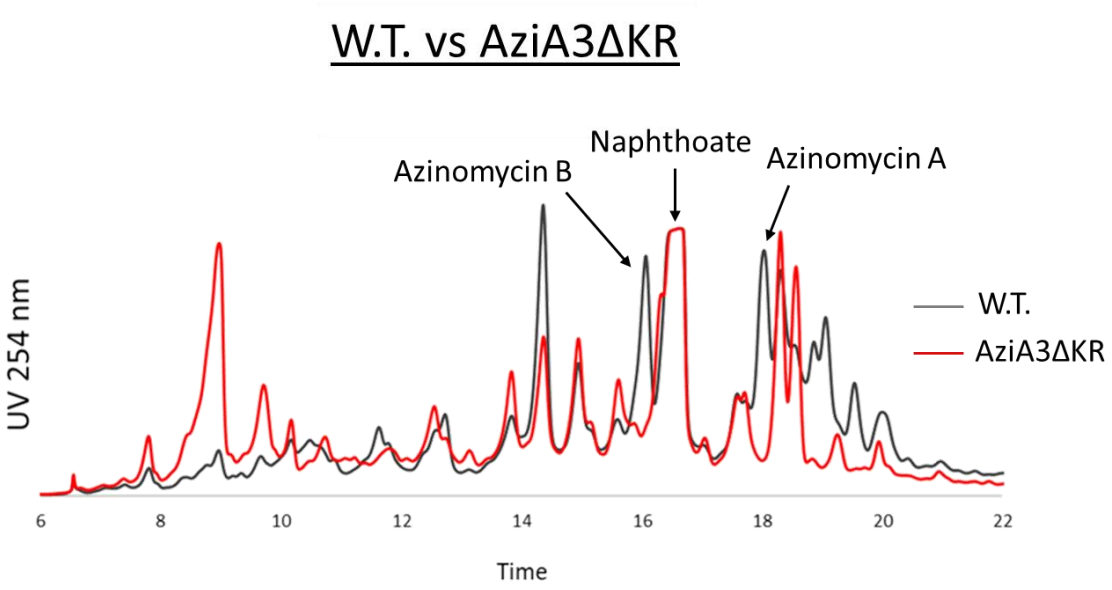
Appendix Figure 22 ClbN LC-MS Traces for Cyclized Product
Extracted Ion Chromatograms (EIC) for 627.299. Expected product m/z $[M+H]^+$ calcd for $C_{28}H_{46}N_6O_6S_2$; 627.299. **A**. Biotin-Cys + ClbN Reaction. **B**. Synthetic Standard of cyclized product. **C**. Negative Control lacking ClbN protein. **D**. Overlay of EIC for Biotin-Cys + ClbN Reaction and Synthetic Standard.



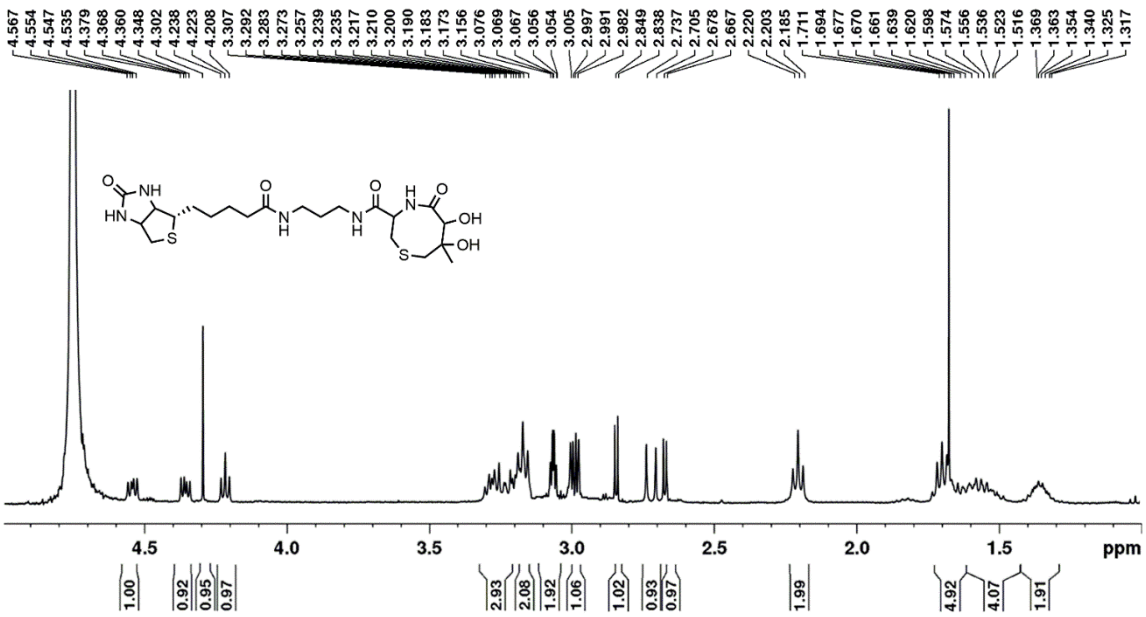
Appendix Figure 23 MS/MS of ClbN Cyclized Product
MS/MS Fragmentation for 627.298 Species. Expected product m/z $[M+H]^+$ calcd for $C_{28}H_{46}N_6O_6S_2$; 627.299. Corresponding fragments are color coded.



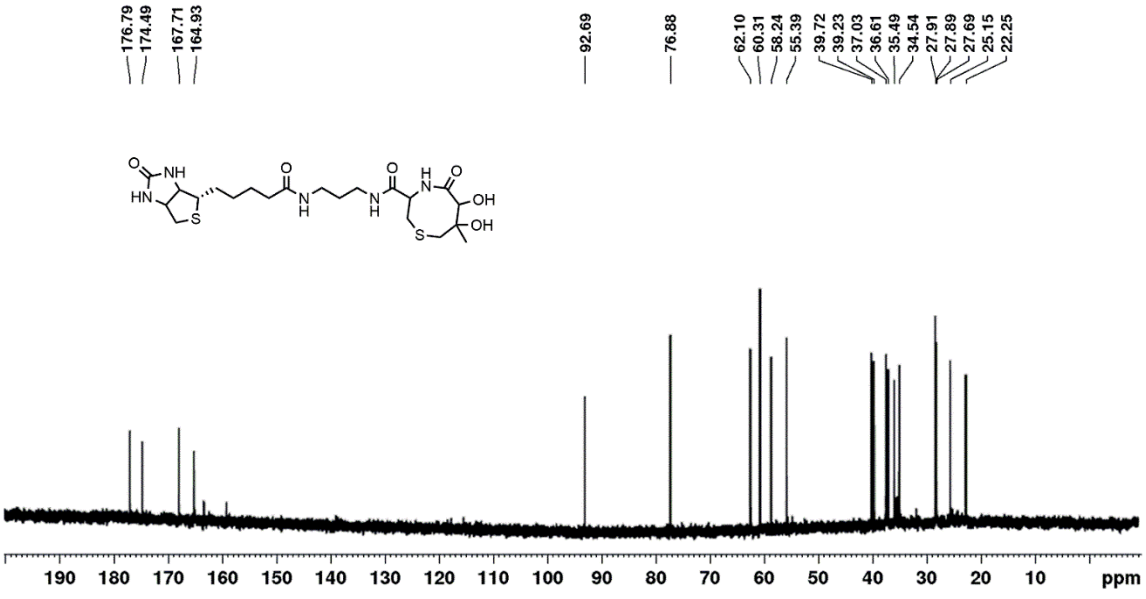
Appendix Figure 24 Δ*aziA6* Crude Extract vs Wild Type Crude Extract LC-MS



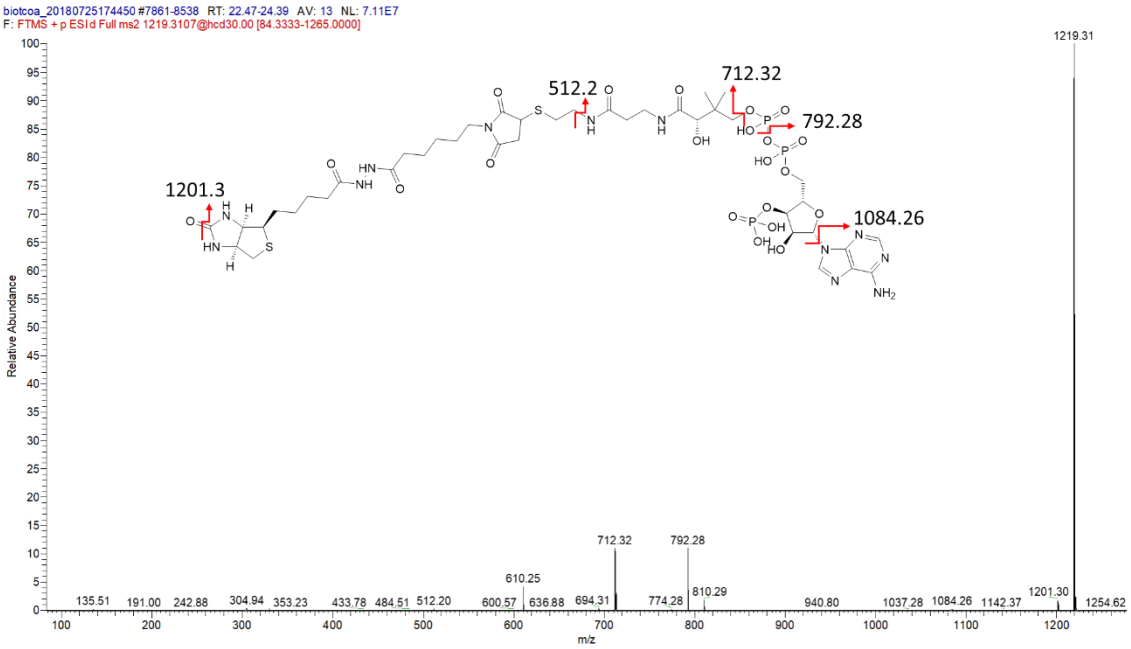
Appendix Figure 25 *aziA3ΔKR* Crude Extract vs Wild Type Crude Extract LC-MS



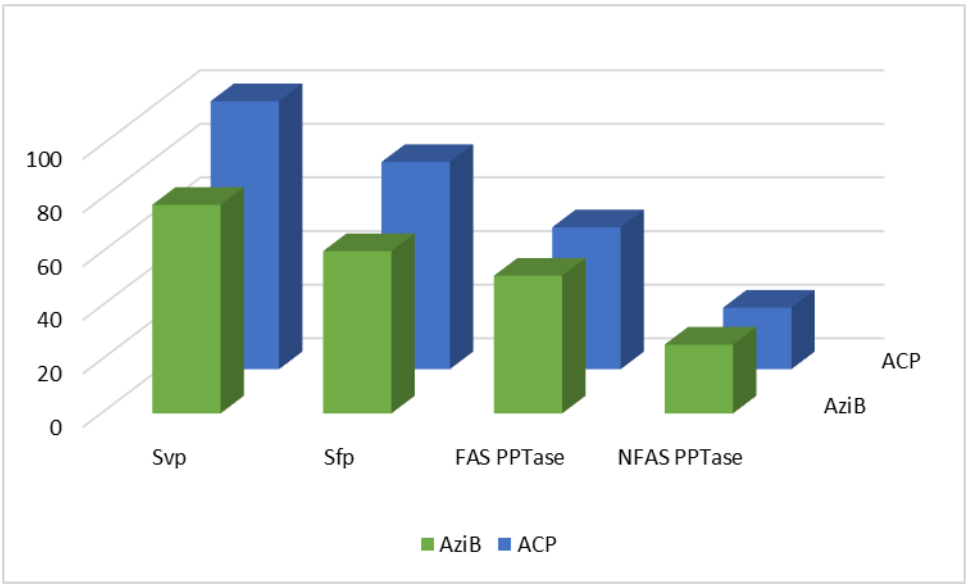
Appendix Figure 26 ^{1}H NMR of Biotin-Cys-epoxyacid in D_2O



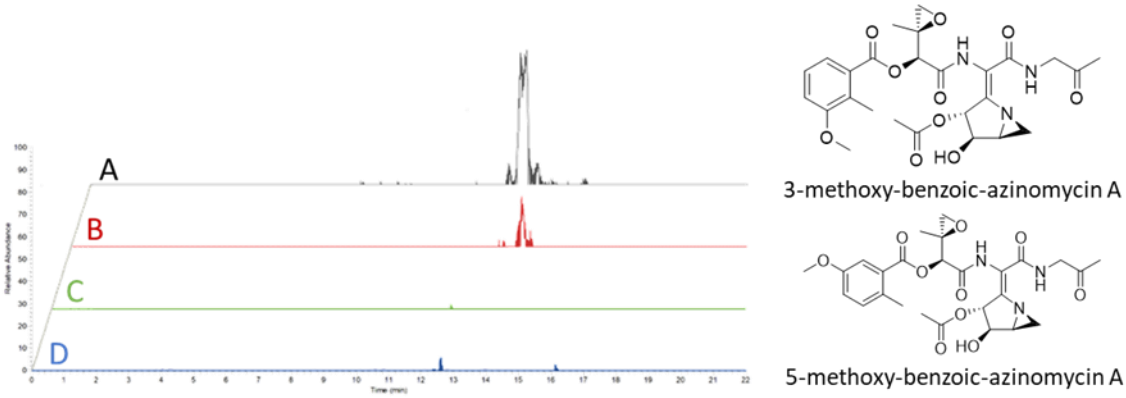
Appendix Figure 27 ^{13}C{^{1}H} NMR of Biotin-Cys-epoxyacid in D_2O



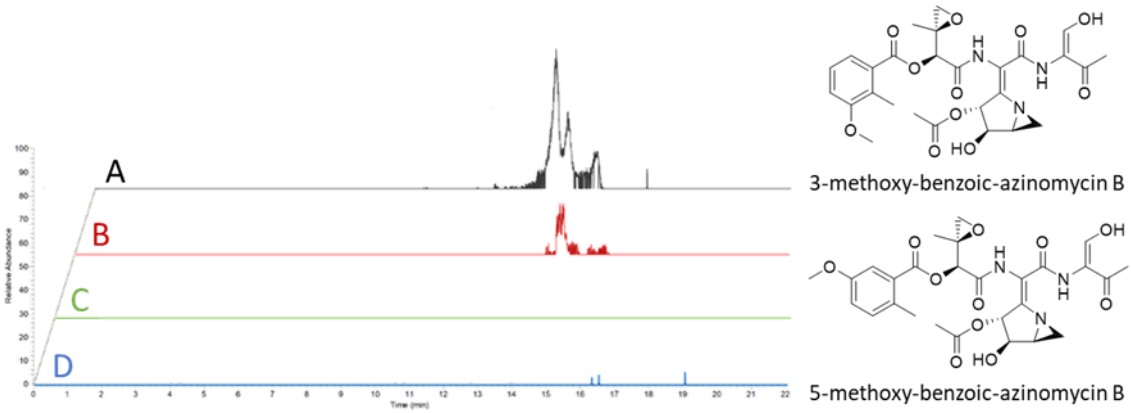
Appendix Figure 28 Biotin-Coenzyme A MS/MS
Expected product m/z $[M+H]^+$ calcd for $C_{41}H_{66}N_{12}O_{21}P_3S_2$; 1219.311. Found $[M+H]^+$ 1219.311.



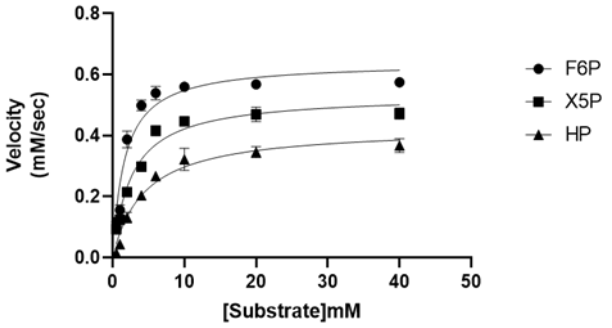
Appendix Figure 29 Relative PPTase Modification Efficiency of AziB and ACP Domain



Appendix Figure 30 EIC Methoxy-benzoic-azinomycin A Derivatives
Extracted Ion Chromatograms (EIC) for 546.208. Expected product *m/z* $[M+H]^+$ calcd for $C_{26}H_{32}N_3O_{10}$; 546.208. **A**. 3-methoxy-2-methylbenzoic acid feeding sample. **B**. 5-methoxy-2-methylbenzoic acid feeding sample. **C**. 2-methylbenzoic acid feeding sample. **D**. Wild type extract.



Appendix Figure 31 EIC Methoxy-benzoic-azinomycin B Derivatives
Extracted Ion Chromatograms (EIC) for 574.203. Expected product *m/z* $[M+H]^+$ calcd for $C_{27}H_{32}N_3O_{11}$; 574.203. **A**. 3-methoxy-2-methylbenzoic acid feeding sample. **B**. 5-methoxy-2-methylbenzoic acid feeding sample. **C**. 2-methylbenzoic acid feeding sample. **D**. Wild type extract.



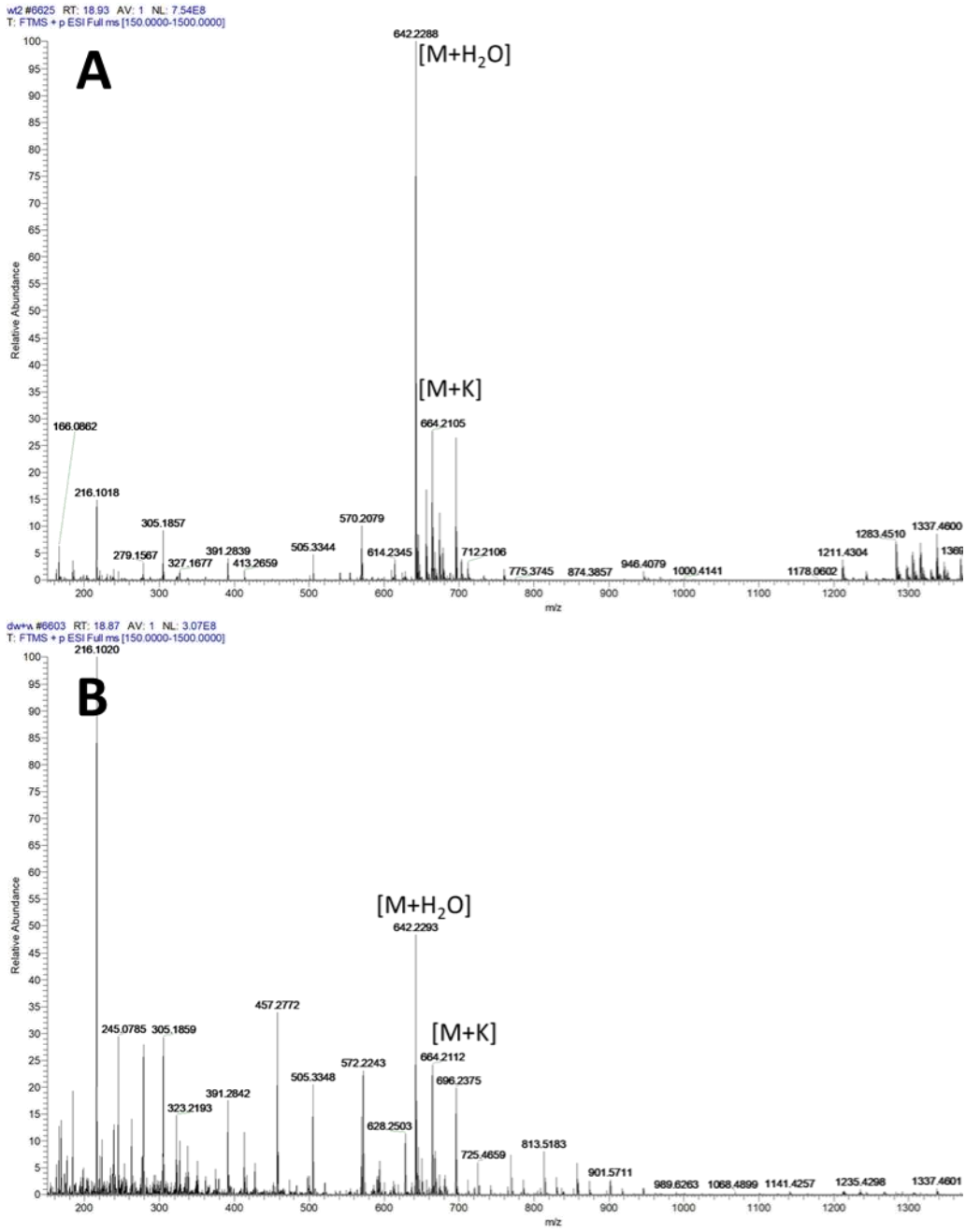
Appendix Figure 32 Michaelis-Menten Kinetic Analysis of AziC5/C6 Donor Compounds[123]

Appendix Table 1 Kinetic Values for AziC5/C6 Donor Molecules[123]

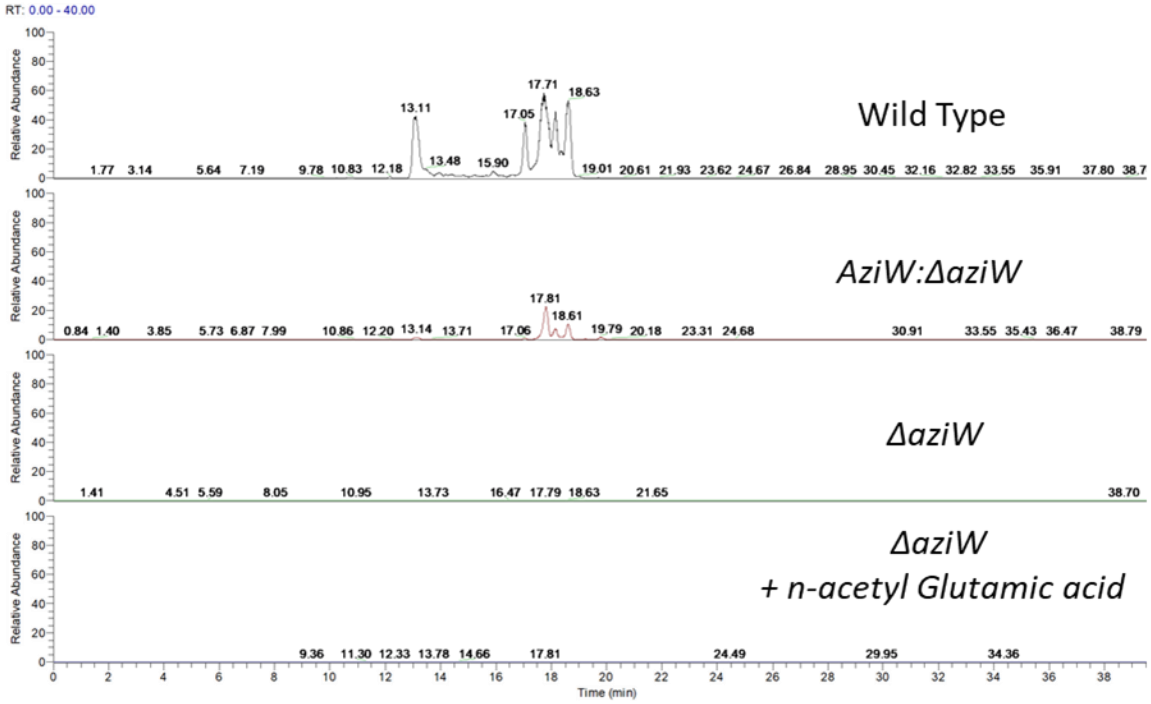
[Substrate] mM	F6P (mM/sec)	X5P (mM/sec)	HP (mM/sec)
0.5	0.117 ± 0.009	0.0938 ± 0.014	0.015 ± 0.003
1	0.155 ± 0.014	0.125 ± 0.007	0.044 ± 0.003
2	0.387 ± 0.022	0.214 ± 0.010	0.131 ± 0.014
4	0.498 ± 0.015	0.297 ± 0.012	0.204 ± 0.008
6	0.539 ± 0.018	0.416 ± 0.013	0.267 ± 0.012
10	0.559 ± 0.003	0.446 ± 0.004	0.322 ± 0.029
20	0.567 ± 0.010	0.469 ± 0.019	0.346 ± 0.015
40	0.575 ± 0.007	0.471 ± 0.015	0.367 ± 0.019

Appendix Table 2 Michaelis-Menten Values for AziC5/C6 Donor Molecules[123]

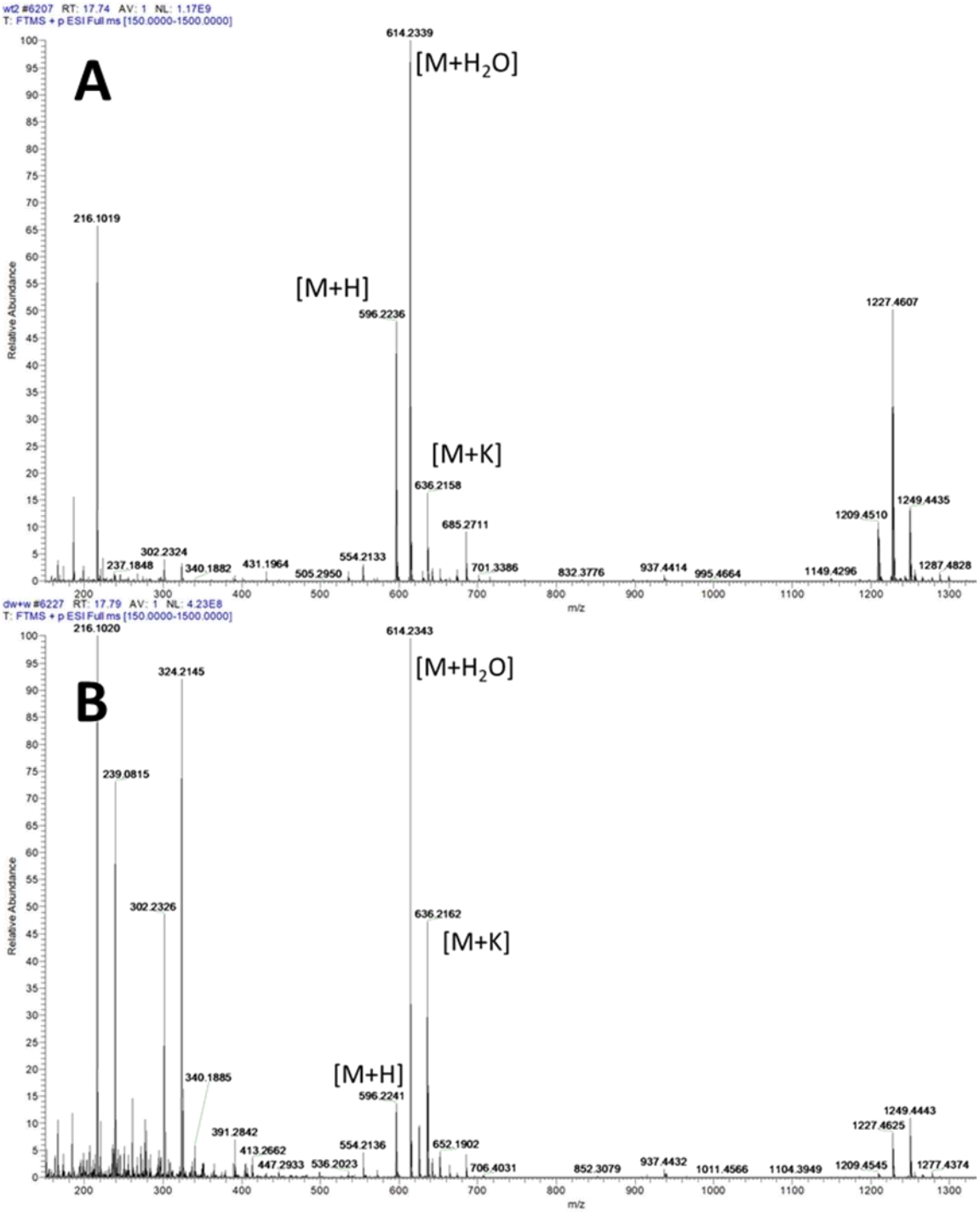
Parameter	F6P	X5P	HP
K_m, mM	2.0 ± 0.4	3.0 ± 0.5	4.0 ± 0.9
k_{cat}, s^{-1}	91 ± 5.4	76 ± 3.6	61 ± 3.8
k_{cat}/K_m, $mM^{-1}\ s^{-1}$	54 ± 13	29 ± 6.3	14 ± 3.6



Appendix Figure 33 AziW Complementation Azinomycin B Masses
A. Azinomycin B in Wild Type. **B**. Azinomycin B in AziW:Δ*aziW* complementation strain. Retention time 18.9 min.



Appendix Figure 34 EIC for Azinomycin A in AziW Study



Appendix Figure 35 AziW Complementation Azinomycin A Masses
A. Azinomycin A in Wild Type. **B**. Azinomycin A in AziW:Δ*aziW* complementation strain. Retention time 17.7 min.