2017

INVESTIGATION OF THE MECHANISM OF ACTION FOR MITHRAMYCIN AND THE BIOSYNTHESIS OF L-REDNOSE IN SAQUAYAMYCINS

Stevi Weidenbach
University of Kentucky, sjo248@uky.edu
Digital Object Identifier: https://doi.org/10.13023/ETD.2017.295

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Stevi Weidenbach, Student

Dr. Jürgen Rohr, Major Professor

Dr. David Feola, Director of Graduate Studies

INVESTIGATION OF THE MECHANISM OF ACTION FOR MITHRAMYCIN AND THE BIOSYNTHESIS OF L-REDNOSE IN SAQUAYAMYCINS

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Pharmacy at the University of Kentucky

By

Stevi Weidenbach

Lexington, Kentucky

Director: Dr. Jürgen Rohr

Professor of Pharmaceutical Sciences

Lexington, Kentucky

2017

ABSTRACT OF DISSERTATION

INVESTIGATION OF THE MECHANISM OF ACTION FOR MITHRAMYCIN AND THE BIOSYNTHESIS OF L-REDNOSE IN SAQUAYAMYCINS

Natural products continue to be a major chemical lead matter for drug discovery due to their diverse chemical structures and bioactivities. Clinically significant natural products include anti-cancer and anti-infective compounds and while many more of these compounds show promising bioactivity, their clinical relevance is often limited by toxicity or poor solubility. Combinatorial biosynthesis can be employed to modify existing chemical scaffolds towards reducing these limitations. To fully take advantage of these biochemical tools, it is important to understand the biosynthesis and mechanism of action of the molecules.

Saccharides in glycosylated natural products provide specific interactions with cellular targets and are often crucial for a compound's bioactivity. Genetic engineering of sugar pathways can modify glycosylation patterns leading to the diversification of natural products. Saquayamycins (SQN) H and I are cytotoxic angucycline antibiotics containing five deoxyhexoses including the rare amino sugar rednose. Elucidating the biosynthetic pathway of rednose could add to the arsenal of combinatorial biosynthesis tools for drug development. Our research goal of investigating the rednose biosynthetic pathway was pursued through two specific aims: the identification of the *Streptomyces* sp. KY 40-1 gene cluster involved in the biosynthesis of SQN H and I (*sqn*) (specific aim 1), and the validation of the proposed L-rednose biosynthetic pathway up to the glycosyl transfer through enzymatic synthesis of NDP-3,6-dideoxy-L-idosamine (specific aim 2). The *sqn* gene cluster revealed deoxysugar biosynthetic genes that could be used to alter glycosylation patterns to generate novel compounds while the enzymatic synthesis afforded novel genetic engineering tools to generate novel TDP-deoxysugars that could be used to diversify compounds such as aminoglycosides to circumvent resistance mechanisms. The first step to generate TDP-glucosamine enzymatically was accomplished, however later steps were unsuccessful.

The aureolic acid mithramycin (MTM) was recently tested in clinical trials for Ewing sarcoma following the discovery of MTM as a potent inhibitor of the oncogenic transcription factor EWS-FLI1 present only in Ewing sarcoma cells It is understood that MTM binds the minor groove of G/C rich DNA as an Mg^{2+}-coordinated dimer disrupting

transcription of proto-oncogenes; however, the DNA recognition rules were not completely understood, making further interrogation of MTM's DNA binding preferences necessary. This research goal of further understanding the mechanism of action for MTM was approached through two specific aims: the investigation of the dimerization of MTM (specific aim 3), and the investigation of MTM's DNA binding preferences (specific aim 4). This work established that MTM and its biosynthetic precursor premithramycin B (PreMTM B), and several MTM analogues with modified 3-side chains: mithramycin SDK (MTM SDK), mithramycin SA tryptophan (MTM SA-Trp), and mithramycin SA alanine (MTM SA-Ala) dimerize even in the absence of DNA under physiologically relevant conditions. The study also demonstrated that modification of the 3-side chain modulates DNA binding affinity of MTM analogues, established a minimum MTM binding site on DNA, and revealed MTM DNA recognition is driven by direct (sequence) and not indirect (conformation) readout laying the foundation for subsequent research based on the interaction between MTM, DNA, and the oncogenic transcription factor EWS-FLI1 in the rational design of new MTM analogues for the treatment of Ewing sarcoma.

KEYWORDS: Deoxysugar Biosynthesis, Enzymatic Synthesis, Gene Cluster, Divalent Metal Ion Coordination, DNA Binding

Stevi Weidenbach

07/24/2017

INVESTIGATION OF THE MECHANISM OF ACTION FOR MITHRAMYCIN AND THE BIOSYNTHESIS OF L-REDNOSE IN SAQUAYAMYCINS

By

Stevi Weidenbach

Dr. Jürgen Rohr

(Director of Dissertation)

Dr. David Feola

(Director of Graduate Studies)

07/23/2017

ACKNOWLEDGEMENTS

The following dissertation could not have been possible without the support and guidance of a few exceptional individuals. First, I would like to thank my Dissertation Chair Dr. Jürgen Rohr for his continued encouragement, motivation, and brilliant ideas. He has taught me to work independently and inspired me through his deep understanding and love for science. Second, I would like to thank my Co-mentor Dr. Oleg Tsodikov for providing excellent hands-on training both in the lab and in scientific writing. His enthusiasm and dedication for his work has given me something to strive for. Third, my Dissertation Committee Members Dr. Steven Van Lanen, and Dr. Robert Perry. Their expertise, suggestions, and encouragement was critical to the development of this dissertation.

Additionally, I would like to thank all the past and present Rohr lab members for their ideas and support; Dr. Guojun Wang, Dr. Jhong-Min Chen, Dr. Daiani Savi, Dr. Amit Jha, Dr. Prithiba Mitra Dr. Abhisek Mandal, Redding Gober, and Ryan Wheeler. I would especially like to thank Dr. Theresa Downey for her patience, support, and teaching me from the ground up in molecular cloning and bioinformatics. Special thanks to Dr. Shaimaa Salem for her contributions to the Saquayamycin project and guidance in gene annotation, protein purification and expression, and *Streptomyces* manipulations. I would also like to thank Dr. Caixia Hou for her contributions on the Mithramycin project.

Next, I wish to thank my friends and family who were equally as important in my doctoral process. My husband, Jake, who stood by me and kept me going. Without his love and support, this past 5 years of work would not have been possible. My mother

Deb, for her reassurance during the tough times and shared celebration during the successes. Her strength, love, and guidance was essential.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

LIST OF ABBREVIATIONS

ACP	acyl carrier protein
AT	acyl transferase
BLAST	basic local alignment search tool
bp	base pairs
CoA	coenzyme A
CYC	cyclase
CTP	cytidine triphosphate
DH	dehydratase
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	*Escherichia coli*
ER	enoyl reductase
EWS	Ewing sarcoma
FAD	flavin adenine dinucleotide
$FADH/FADH_2$	flavin adenine dinucleotide, reduced form

FLI1	Friend Leukemia Integration 1
GT	glycosyltransferase
His6	6x polyhistidine tag
IPTG	isopropyl-β-D-thiogalactoside
KS	ketoacyl synthase
KR	keto reductase
LB	lysogeny broth
MTM	mithramycin
MTM SK	mithramycin short side chain, keto
MTM SDK	mithramycin short side chain, diketo
MTM SA	mithramycin short side chain, acid
MTM SA-Trp	mithramycin SA-tryptophan
MTM SA-Ala	mithramycin SA-alanine
MW	molecular weight
NAD^+	nicotinamide adenine dinucleotide, oxidized form
NADH	nicotinamide adenine dinucleotide, reduced form

$NADP^+$	nicotinamide adenine dinucleotide phosphate, oxidized form
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NDP	nucleoside diphosphate
ORF	open reading frame
PCP	peptidyl carrier protein
PCR	polymerase chain reaction
PKS	polyketide synthase
PLP	pyridoxal 5’phosphate
PMP	pyridoxal 5’phosphate
Pre MTM B	premithramycin B
SAM	*S*-adenosyl methionine
TDP	thymidine diphosphate
TMP	thymidine monophosphate
TTP	thymidine triphosphate
TE	thioesterase
UTP	uridine triphosphate

UV	ultraviolet

1. INTRODUCTION TO POLYKETIDES AND DEOXYSUGAR BIOSYNTHESIS

Secondary metabolites are small molecules that are not directly involved in reproduction, development, or growth of the organism under normal conditions, but provide an evolutionary advantage for the producing organism such as antibiotics for defense, attractants like flower scent to attract pollinators, pigments to attract or warn other species, and siderophores for iron scavenging during low-iron conditions. Natural products produced by plants, insects, and microorganisms are derived from secondary metabolic pathways and are of interest to humans for pesticides, flavorings, dyes and most importantly pharmaceuticals. The diverse chemical scaffolds and various bioactivities of natural products have long since rendered them significant sources for drug discovery. Natural products are categorized into 6 groups based on their building blocks and method of biosynthesis (figure 1): specialized carbohydrates (1), terpenoids and steroids (2), alkaloids (3), peptides (4), aromatic amino acids and phenylpropanoids (5), and fatty acids and polyketides (6).

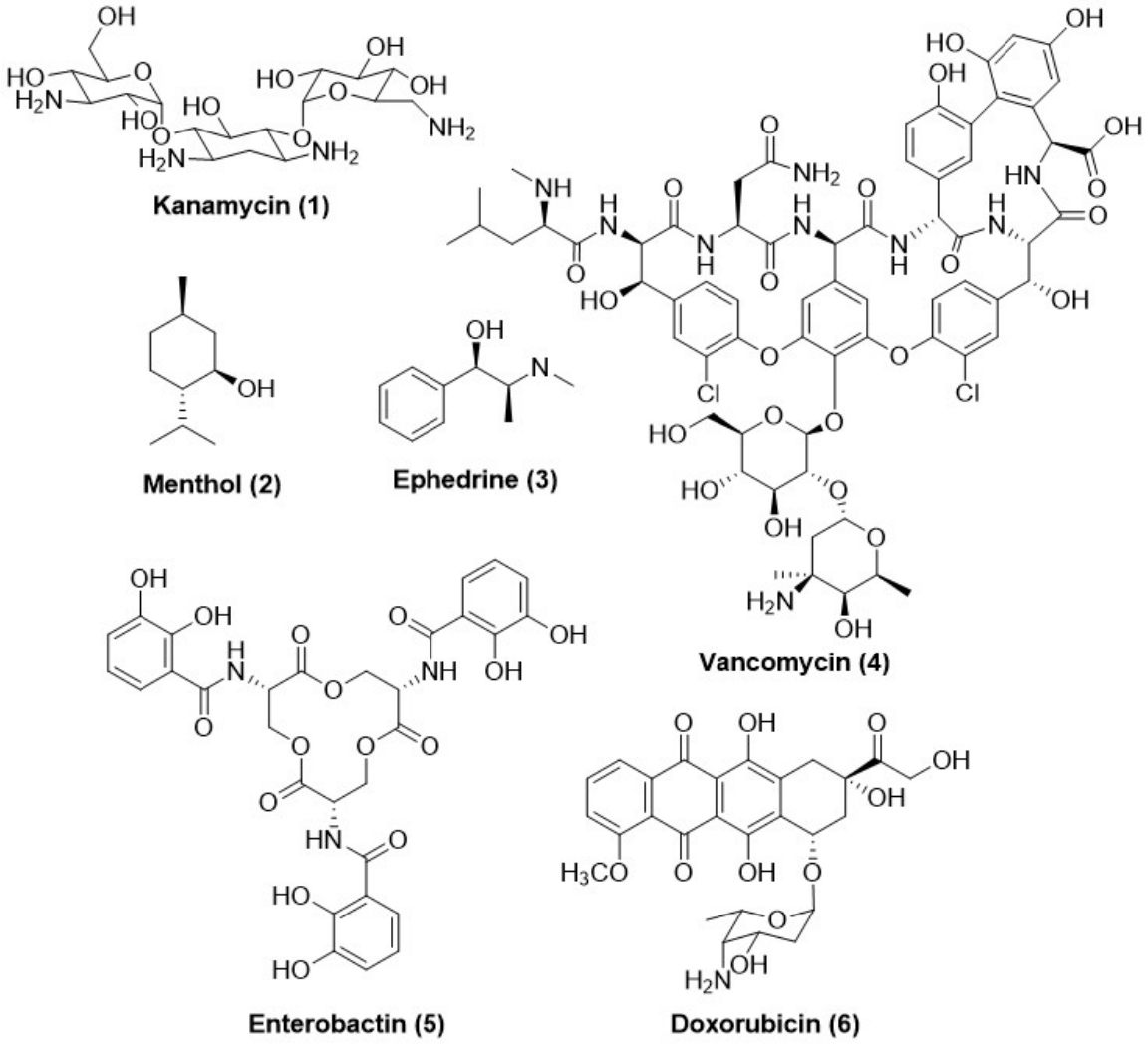


Figure 1. Representative compounds from the classes of natural products

Of particular interest are the polyketide natural products due to their structural diversity and pharmacological impact including anticancer, antibiotic, antifungal, antiparasitic, and immunosuppressive activities. Polyketides are produced primarily by microorganisms through a series of decarboxylative Claisen condensations of acyl CoA esters catalyzed by polyketide synthases (PKS) to build a linear polyketide chain.

Polyketides are classified based on the type of PKS (Type I, Type II, or Type III) responsible for their biosynthesis. Type I PKS are large multifunctional, modular enzymes that function non-iteratively in an assembly line fashion with multiple acyl carrier proteins (ACP) to pass along the growing polyketide chain (figure 2). Typical products for Type I PKS include macrolides, polyethers, and macrolactones. Clinically relevant erythromycin A (7) and amphotericin B are both type I PKS derived molecules (figure 3). Type II PKS are made up of a complex of multiple enzymes with specific functions including two ketosynthase/acyl transferases ($KS_{\alpha}$and KS_{β}) and an ACP that catalyze the extension, and initial cyclization reactions iteratively (figure 2). Type II PKS molecules are polycyclic aromatics such as angucyclines and anthracyclines and include such heavy hitters as doxorubicin (6) and mithramycin (8) (figure 3). Type III PKS act iteratively with only one homodimer forming ketosynthase/acyltransferase (KS) enzyme and no ACP (figure 2). Stilbene-type phytoalexins and flavonoids (9) are typical products of type III PKS (figure 3). Following construction of the polyketide core scaffolds by PKS, post-PKS modifying enzymes diversify the polyketide backbones catalyzing cyclizations, aromatizations, reductions, epimerizations, alkylations, glycosylations, and other tailoring reactions providing much of the structural diversity within the polyketide group.

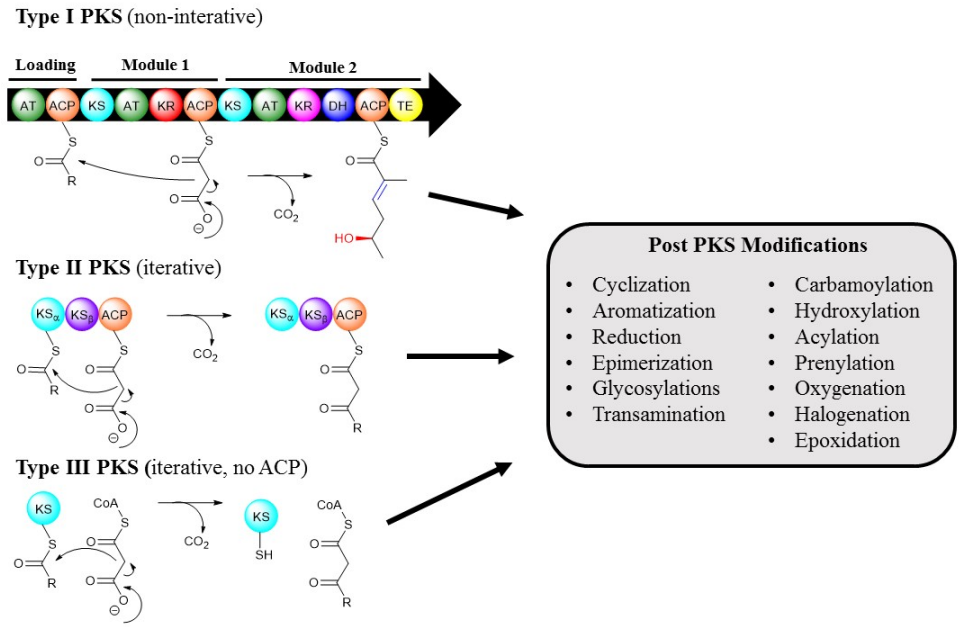


Figure 2. Polyketide synthases (type I, type II, type III) and post-PKS modifications

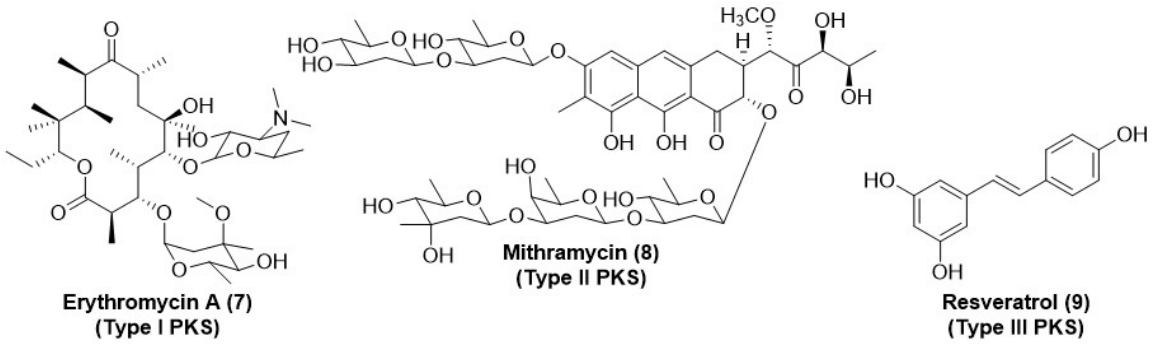


Figure 3. Representative polyketides from PKS (type I, type II, type III)

Natural products, including polyketides, are often glycosylated, and pharmacological properties of these compounds are greatly impacted by their carbohydrate components of which provide specific interactions with cellular targets and are often critical for bioactivity.[1] A recent review of close to 16,000 natural products revealed that of them, 21.5% were glycosylated.[2] Doxorubicin (6), erythromycin (7), amphotericin B, and avermectin all contain sugar residues, removal of which renders the

compounds inactive.[1] Furthermore, sugars are highly soluble and increase the overall solubility of the glycosylated compounds.

Glycosylation requires glycosyl donors, mainly nucleotide diphosphate (NDP)-bound sugars, and an aglycone or other suitable glycosyl acceptor and employs specific enzymes known as glycosyltransferases evolved to carry out the attachment of the activated sugar donor to the acceptor. Glycosidic bonds can be S-, O-, N-, or C-, although linkage through oxygen is the most common. According to Klyne's rule, linkages in glycoside natural products are typically β-configuration for D-sugars and α-configuration for L-sugars.[3] Additionally, specific deoxysugar tailoring enzymes have evolved to modify sugar donors by further deoxygenation, oxidation, reduction, rearrangement, group transfer, isomerization, and epimerization to generate a wide array of unusual sugars. Genes that code for biosynthetic enzymes in the same pathway are generally "clustered" in the chromosomes of bacteria. This allows for horizontal gene transfer in bacteria and inconsequently provides researchers with easy access to this biosynthetic machinery. Deoxysugar biosynthetic pathways can be genetically modified to alter glycosylation patterns to optimize desirable drug properties of existing compounds.

Deoxysugar biosynthetic pathways begin with the glycolytic intermediate D-glucose-6-phosphate (10) which undergoes isomerization carried out by phosphoglucomutase to yield D-glucose-1-phosphate (11). This D-glucose-1-phosphate (10) is considered the precursor of the typical deoxysugar biosynthetic pathway. Secondary metabolism begins with a phosphoanhydride exchange reaction catalyzed by NDP-D-glucose synthase from NTP (CTP, UTP, or TTP (12) in bacteria) and D-glucose-

1-phosphate (12) to yield pyrophosphate and NDP-D-glucose (13). The pyrophosphate is hydrolyzed by a ubiquitous cytosolic enzyme, inorganic pyrophosphatase which provides energy for the otherwise endergonic reaction.[4] The sugar is thus activated for downstream reactions (figure 4).

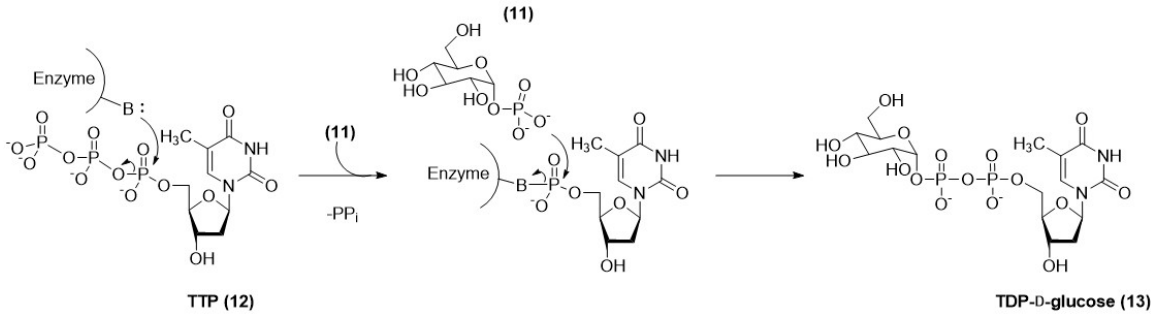


Figure 4. Activation of sugar substrate by glucose synthase

1.1 Deoxygenation

Carbon-oxygen bond cleavage is among the most ubiquitous reactions in biological systems and introduces great structural diversity in compounds in which it is carried out. Deoxysugars are common constituents of both lipopolysaccharides in gram negative bacterial cell walls, and secondary metabolites, and require C-O bond cleavage for their formation. Four types of deoxygenation reactions have been identified: C-6, C-2, C-3, and C-4 deoxygenation.

1.1.1 C-6 Deoxygenation

The majority of unusual deoxysugars are 6-deoxysugars. In this case, an NAD dependent 6-deoxygenation reaction is carried out by an NDP-D-glucose oxidoreductase known as 4,6-dehydratase, whereby NAD^+ oxidizes C-4 OH activating the C-5 proton to

form the enolate (16) which facilitates the elimination of C-6 OH (17). Finally, a hydride transfer from NADH saturates the double bond to form NDP-4-keto-6-deoxy-D-glucose (18) (figure 5).[5-9] This 4-keto-6-deoxy-D-glucose is a universal intermediate that undergoes a series of downstream tailoring reactions to generate a wide array of sugar donors that are used by glycosyltransferases to decorate acceptors. Chemically interesting tailoring reactions that pertain to 6-deoxy hexoses are discussed in the following sections.

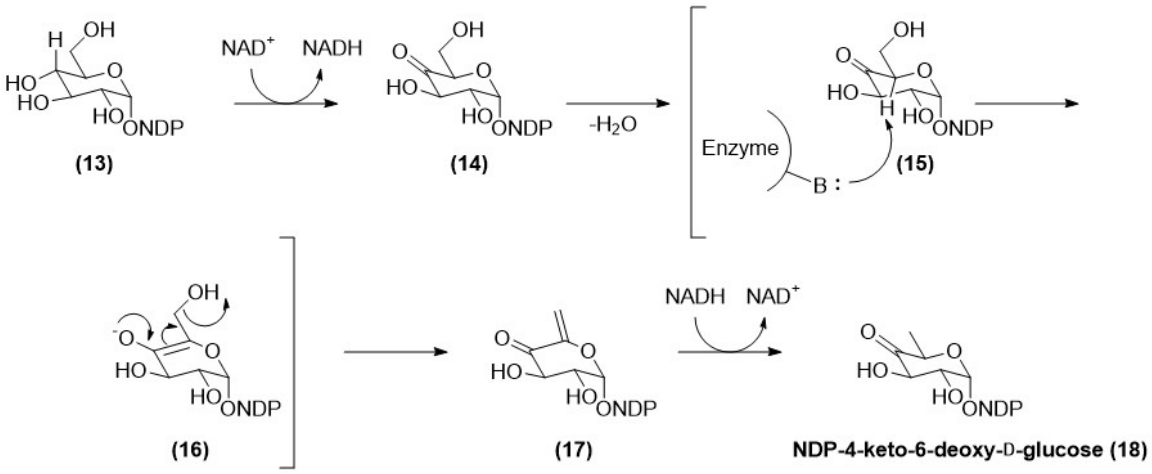


Figure 5. Mechanism of C-6 deoxygenation by 4,6-dehydratase

1.1.2 C-2 Deoxygenation

2-deoxygenation is the 2nd most common hexose deoxygenation event after 6-deoxygenation and the resulting 2-deoxysugars are common structural components of natural products primarily found in aminoglycoside and macrolide antibiotics. C-2 deoxygenation is similar to C-6 deoxygenation employing the 4-ketogroup to activate the C-3 proton for abstraction followed by dehydration across C-2-3 which is then reduced to form NDP-4-keto-2,6-dideoxy-D-glucose. However, unlike C-6 deoxygenation, C-2 requires two enzymes that work in a concerted fashion, a

dehydrase and a reductase. The 2,3-dehydratase carries out the dehydration across C2-C3 to form an unstable 3,4-diketo intermediate (22). A second enzyme, an NAD(P)H-dependent 3-ketoreductase helps stabilize the intermediate and reduces the C-3 ketone to form the final product (23) (figure 6).[10, 11] In the case of tylosin biosynthesis, a 3-methylation,4-ketoreduction and 5-epimerization follow 2-deoxygenation to convert TDP-4-keto-2,6-dideoxy-D-glucose (23) to TDP-L-mycarose.[11]

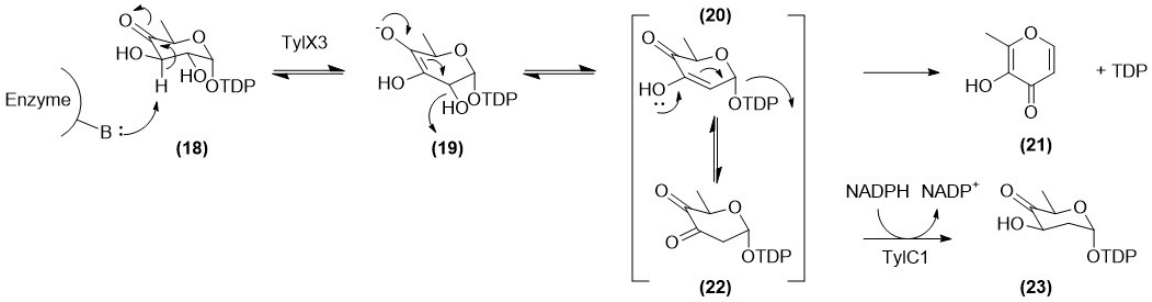


Figure 6. Mechanism of C-2 deoxygenation by 2,3-dehydratase and 3-ketoreductase

1.1.3 C-3 Deoxygenation

3,6-dideoxysugars are common components of gram negative bacterial cell wall lipopolysaccharides and are epitopes responsible for virulence and pathogenicity. In contrast to C-2 or C-6 deoxygenation, C-3 deoxygenation cannot be promoted by an initial α-H abstraction which requires activation by a 4-keto. Instead, C-3 deoxygenation is an elimination of the hydroxyl group from the α-C directly next to the 4-keto and requires two cooperating enzymes, a pyridoxamine 5'-phosphate (PMP)-dependent dehydratase/reductase and a reductase [2S·2Fe] flavoprotein. The dehydratase/reductase (E_1) active site contains one noncovalently-bound equivalent of PMP and one [2S·2Fe] center per subunit.[12] The reductase (E_3) active site contains a ferredoxin [2S·2Fe] center and flavin adenine dinucleotide (FAD).[13, 14] The first half-reaction, a PMP-dependent

reversible dehydration is carried out by E_1 and triggered by the formation of a Schiff base between the amino group of PMP and C-4 keto of the substrate (25).[15] Next, an abstraction of the *pro-S* hydrogen of C-4' of the PMP-ketimine complex by His_{220} in E_1 leads to an expulsion of 3-OH in the substrate to generate a $\Delta^{3,4}$-glucoseen intermediate (27).[16] The second half-reaction is a pyridoxal-based radical reduction in which E_1 and E_3 act as an electron transport chain where a hydride from NADH reduces FAD in E_3, which reduces the E_3 iron sulfur cluster, which reduces the iron sulfur cluster in E_1, which in turn reduces the $\Delta^{3,4}$-glucoseen intermediate (27) to form the NDP-4-keto-3,6-dideoxy-D-glucose (37)[13, 17-19] (figure 9).[11]

E_1 has evolved to have a specific reductase (E_3) as is encoded in the *asc* gene cluster responsible for ascarylose biosynthesis in *Yersinia pseudotuberculosis* Serogroup VA.[20] However, there are examples where E_1 has evolved without a specific reductase to regenerate its [2S·2Fe] center and can function with a general cellular reductase; in this case only E_1 is coded for in the gene cluster as in spinosyn *spn* responsible for forosamine biosynthesis.[21] Forosamine biosynthesis first requires 6-deoxygenation (38), 2-deoxygenation (40), then 3-deoxygenation (41) followed by 4-transamination (42), and 4-dimethylation to yield TDP-D-forosamine (43) (figure 7).[22] Ascarylose biosynthesis requires 6-deoxygenation (45), then 3-deoxygenation (46) followed by epimerization (47) and 4-ketoreduction the yield the final product (48) (figure 8).[20, 23]

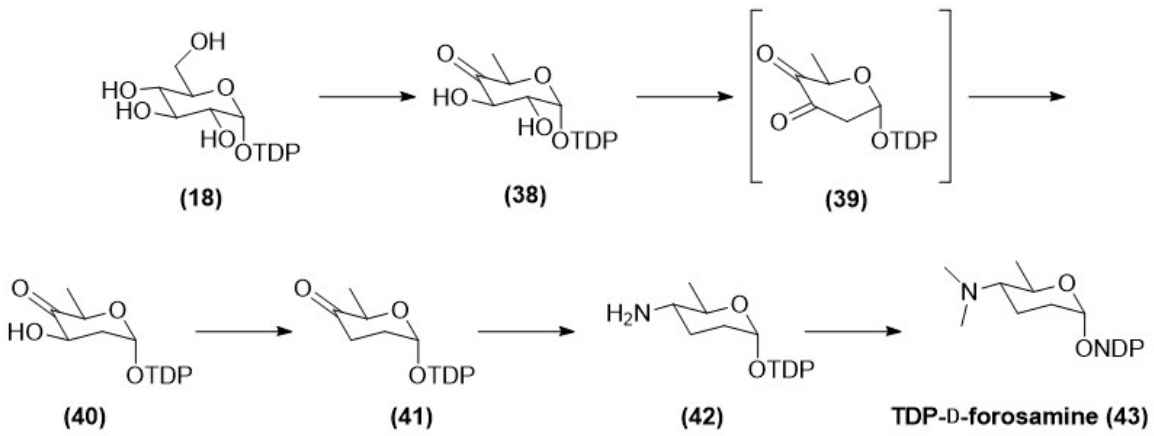


Figure 7. Forosamine biosynthesis

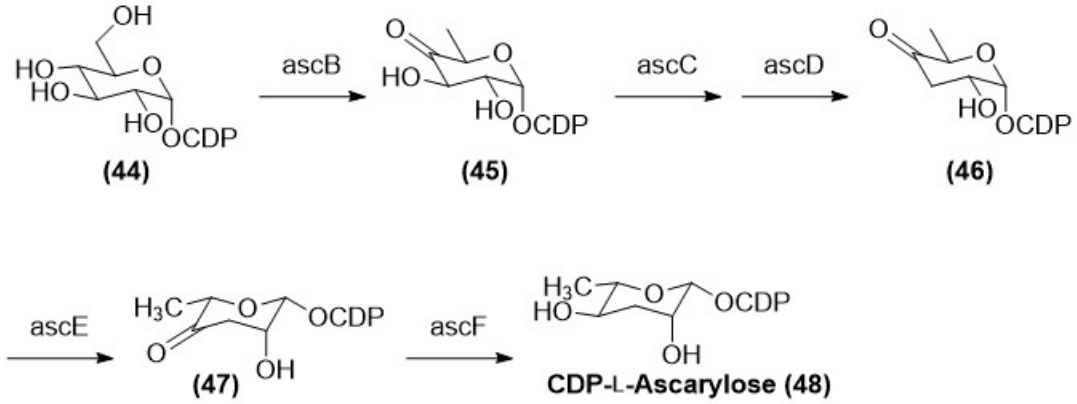


Figure 8. Ascarylose biosynthesis

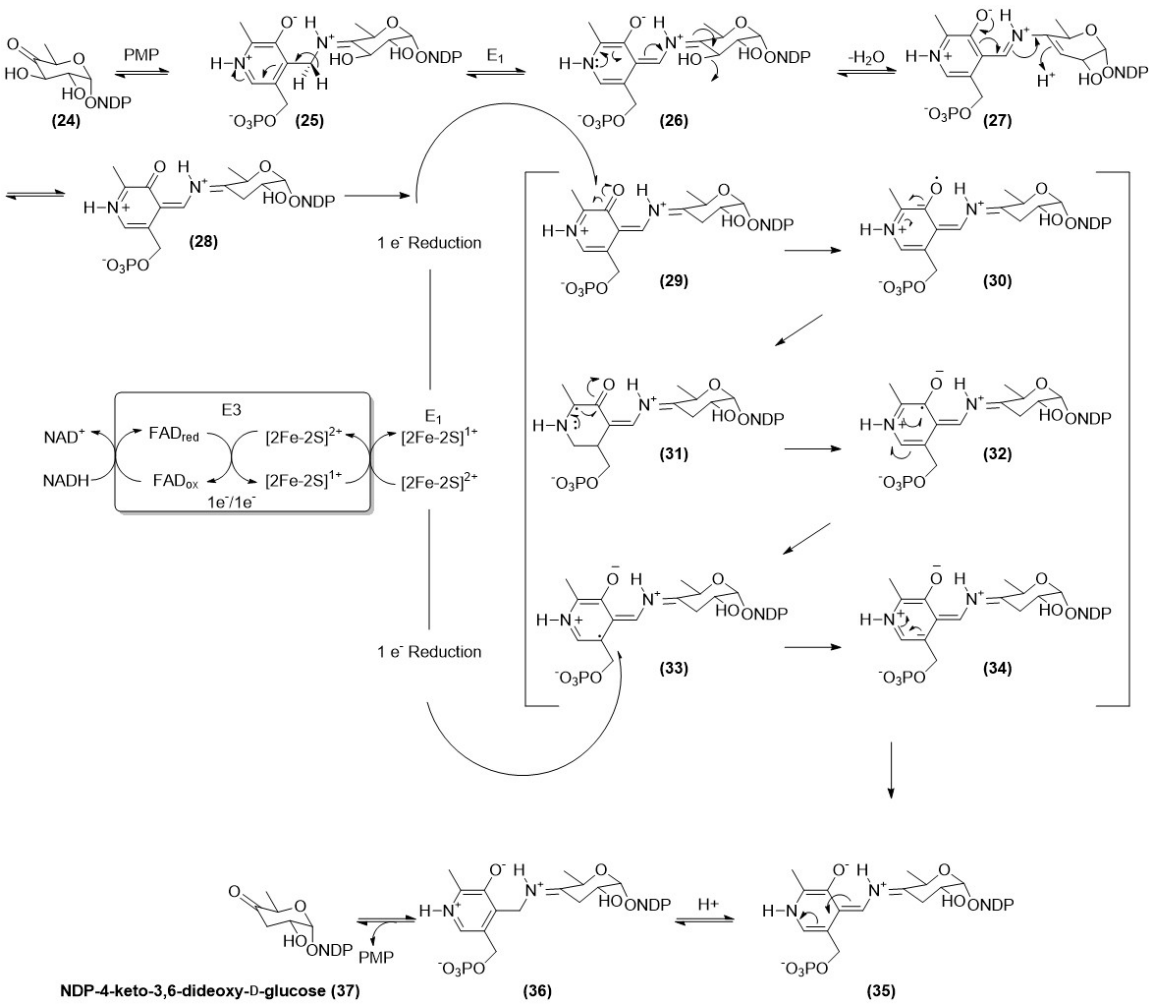


Figure 9. Mechanism of C-3 deoxygenation by 3,4-dehydratase and reductase

1.1.4 C-4 Deoxygenation

C-4 deoxygenation is quite rare, most notably seen in desosamine biosynthesis, a carbohydrate constituent of several macrolides including erythromycin. Two enzymes are required to catalyze C-4 deoxygenation and due to the difficulty in studying these particular enzymes, the mechanistic details are unclear. The first step is carried out by a pyridoxal 5'phosphate (PLP)-dependent aminotransferase to generate a 4-aminosugar intermediate.[24] The details regarding the

mechanism are discussed in the next section, see *Transamination*. The second step is a deamination by a radical *S*-adenosyl-L-methionine (SAM)-dependent enzyme.[25] Radical SAM-dependent enzymes contain a $[4Fe\text{-}4S]^{+}$ center that must be continuously reduced in order to function, making them particularly difficult to study. Catalysis begins with one electron transferred from the $[4Fe\text{-}4S]^{+}$ to SAM triggering homolytic cleavage of the C-5'-*S* bond of SAM yielding methionine and a 5'-deoxyadenosyl radical.[26] Next, the 5'-deoxyadenosyl radical abstracts a hydrogen atom from C-3 generating a C-3 α-hydroxyalkyl radical (50). Mechanistic details of the following steps are unclear but two mechanisms have been proposed. In the first (figure 10, red), a radical-induced 1,2-amino shift forms an aminol radical (58). A hydrogen is reclaimed from 5'deoxyadenosine (59) and an ammonium ion is eliminated to yield the final product (60). In the second (figure 10, blue), the C-3 hydroxyl group is deprotonated to form a ketyl radical anion (52) followed by β-elimination of the 4-amino group driven by the resonance of the ketyl radical anion (53).[26, 27]

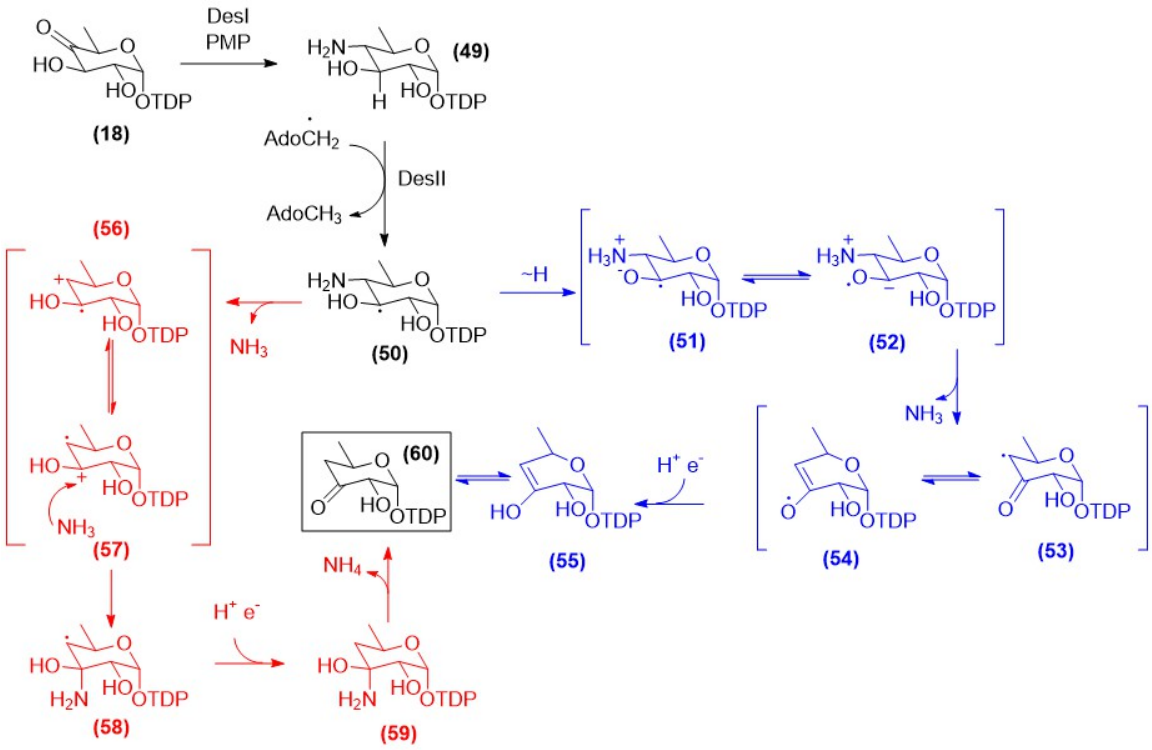


Figure 10. Two proposed mechanisms of C-4 deoxygenation by DesI, a pyridoxal-5′-phosphate (PLP)-dependent aminotransferase, and DesII, a radical S-adenosyl-L-methionine (SAM)-dependent enzyme

1.2 Transamination

Amino-sugars are found in natural products such as aminoglycosides, polyene antifungals, glycopeptides, anthracyclines, and macrolides as well as bacterial cell walls, flagellar glycolipids, and bacterial toxins. Clinically relevant antibiotics such erythromycin (7) and kanamycin (1) as well as the chemotherapeutic agent doxorubicin (6) contain amino-sugars which are critical for their bioactivity.[1] In addition to mediating cell-target interaction, amino-sugars also serve to increase a molecule's solubility. Aminotransferases introduce an amino group from an amino

acid donor, to an NDP-ketosugar acceptor. The enzyme and the location of the keto group, primarily C-3 or C-4 generated in a previous dehydrogenation or isomerization biosynthetic step dictate the substitution position.[28, 29]

Transamination is PLP dependent with a PLP cofactor covalently bound to a lysine residue in the active site of the aminotransferase forming an internal aldimine (61) at the start of the catalytic cycle. The first half reaction (figure 11) proceeds with the exchange of another amino acid donor, usually L-glutamine or L-glutamate for the conserved lysine resulting in the formation of an external aldimine (64). Hydrogen bonding now holds PLP in place as it is no longer covalently bound to the enzyme. Next, an abstraction of the α-proton from the PLP-amino acid adduct leads to a quininoid (66) and is immediately followed by a reprotonation of the 4'-C of PLP to generate a ketamine (67). The ketamine (67) is hydrolyzed to form PMP, and an α-ketoacid is released (70). The second half reaction (figure 12) begins with ketamine formation (74) between the ketosugar substrate and PMP. The 4'-C of PMP is deprotonated to form a quininoid (75), which is followed by reprotonation of α-C to form an external aldimine (77) which is then hydrolyzed to release the amino-sugar and reform the internal aldimine.[30-33]

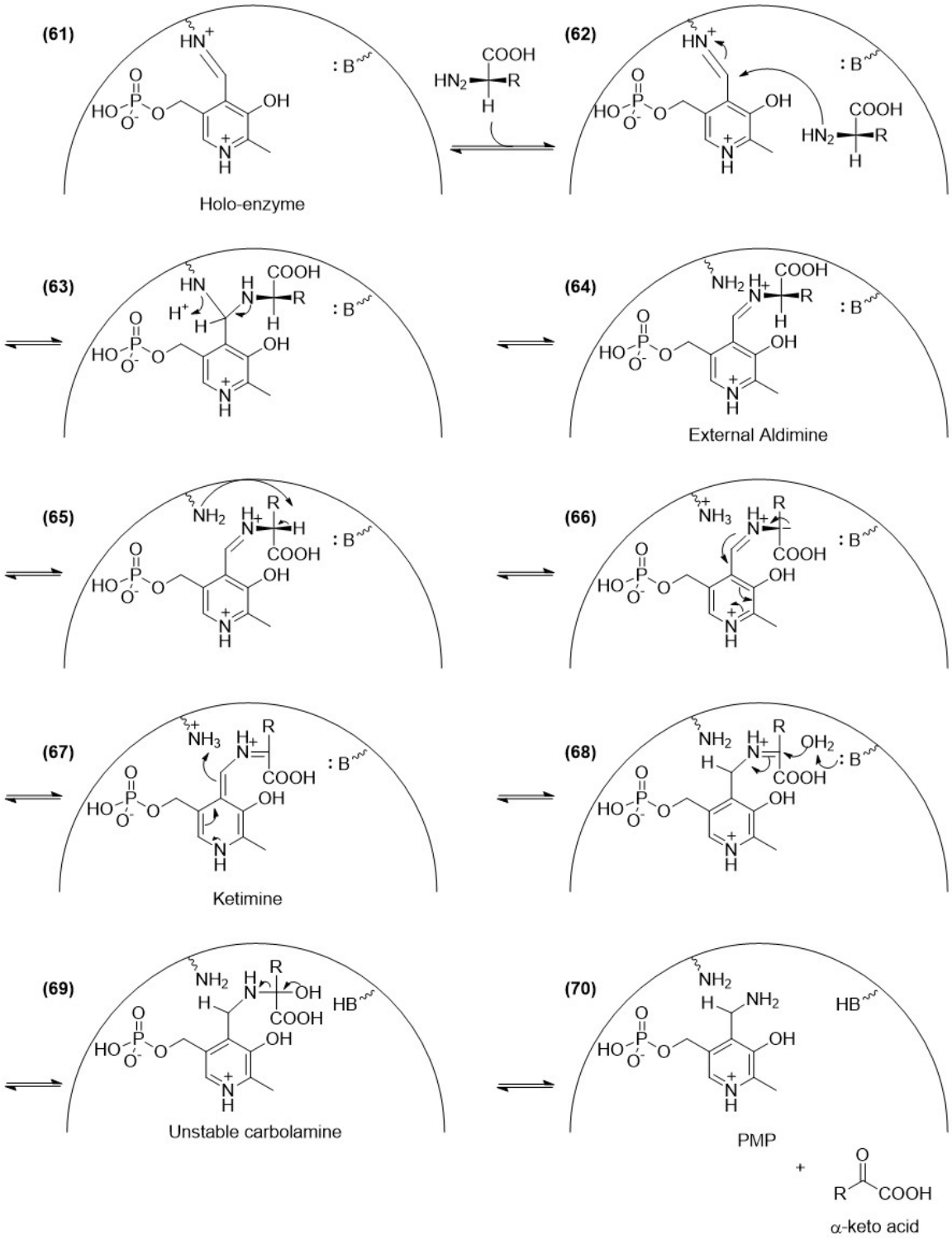


Figure 11. 1st half-reaction of PLP-dependent transamination

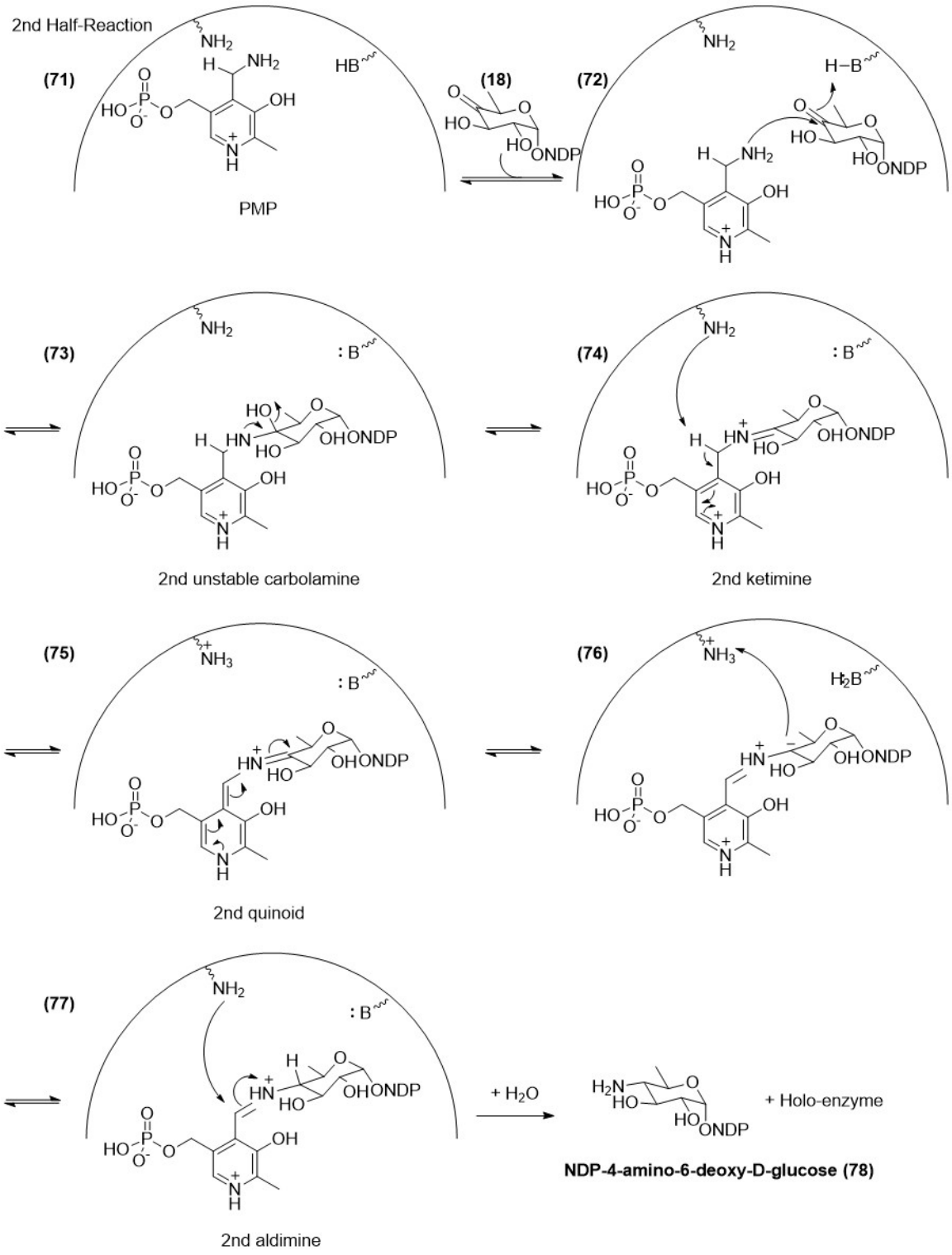


Figure 12. 2nd half-reaction of PLP-dependent transamination

1.3 Epimerization

Epimerization is the inversion of configurational stereochemistry at one or more chiral centers and provides a convenient method to generate a wide array of sugars with structural diversity from just a few common sugar precursors. Carbohydrate epimerization reactions are carried out by epimerases on C-2, C-3, C-4, and C-5 usually through the addition and removal of hydrogen. For C-2 and C-4, the mechanism is similar; an oxidation/reduction cycle requiring NAD^+ to deprotonate the C-2/C-4 hydroxy group to form a carbonyl, followed by reprotonation on the opposite face to invert configuration at C-2 as in CDP-D-tyvelose in *O*-antigen biosynthesis of *Yersinia pseudotuberculosis* IVA (figure 13) [34, 35] or C-4 as in the conversion of galactose to glucose for glycolysis (figure 14). [36, 37] C-3 and C-5 epimerization involves keto-enol tautomerization and requires a C-4 keto to activate the α-H for abstraction followed by reprotonation at C-3 or C-5 on the opposite face. Following C-5 epimerization, steric hindrance is minimized by maintaining the hydroxymethyl group (C-6) in the equatorial position which flips the ring from $^4C_1$ to $^1C_4$ and converts the sugar from D to L configuration.[38, 39] An example of C-5 epimerization can be seen in colitose biosynthesis (figure 15).[38]

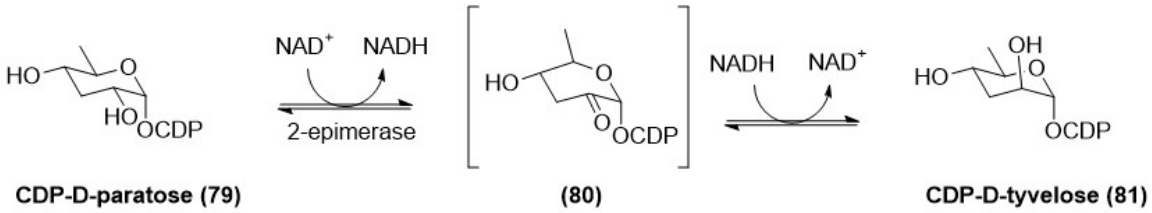


Figure 13. Mechanism of C-2 epimerization

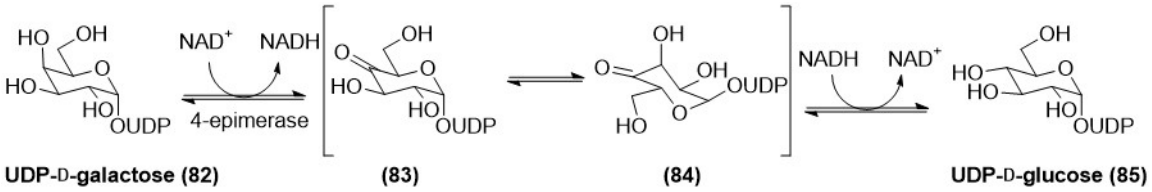


Figure 14. Mechanism of C-4 epimerization

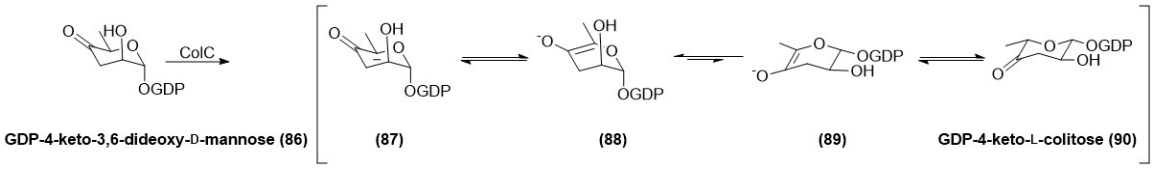


Figure 15. Mechanism of C-5 epimerization

1.4 Branched-Sugar Formation

Branched-sugars can be derived one of two ways. Either through the attachment of a methyl or a two-carbon side chain from an exogenous donor such as SAM or pyruvate to a pyranose ring or through the intramolecular rearrangement of a pyranose precursor to form a furanose with a formyl or hydroxymethyl side chain. Methyl or two-carbon side chain containing sugars are classified as Group I while formyl or hydroxymethyl side chain containing sugars belong to Group II. It is extremely rare for a sugar to bear a branch chain longer than two carbons. Branched sugar biosynthesis has not been well studied. However, two examples, namely the methyl-side chain containing L-mycarose from the antibiotic tylosin produced by *Streptomyces fradiae*, and yersiniose A, a two-carbon side chain carbohydrate

component of the *O*-antigen in the lipopolysaccharide of *Yersinia pseudotuberculosis* have been investigated. L-mycarose biosynthesis requires a SAM-dependent *C*-methyltransferase. Catalysis is initiated by C-3 proton abstraction to form an enediolate intermediate (93) followed by nucleophilic attack on the electrophilic methyl group in SAM to yield TDP-4-keto-D-mycarose (94) (figure 16).[40] A 5-epimerization and 4-ketoreduction complete the biosynthesis to generate L-mycarose.[41] Yersiniose A biosynthesis requires a TPP-dependent coupling enzyme which facilitates a de facto umpolung (charge reversal) reaction. First, the TPP cofactor binds pyruvate followed by decarboxylation to generate a carbanionic adduct (99). Next, the TPP-bound carbanion attacks the 4-keto group and TPP is cleaved to yield the 2-carbon side chain branched-sugar yersiniose A (101) (figure 17).[42]

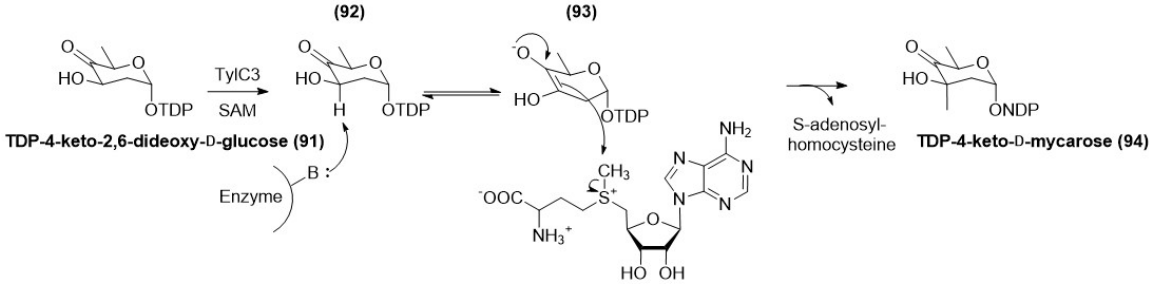


Figure 16. Mechanism of C-methylation in D-mycarose biosynthesis

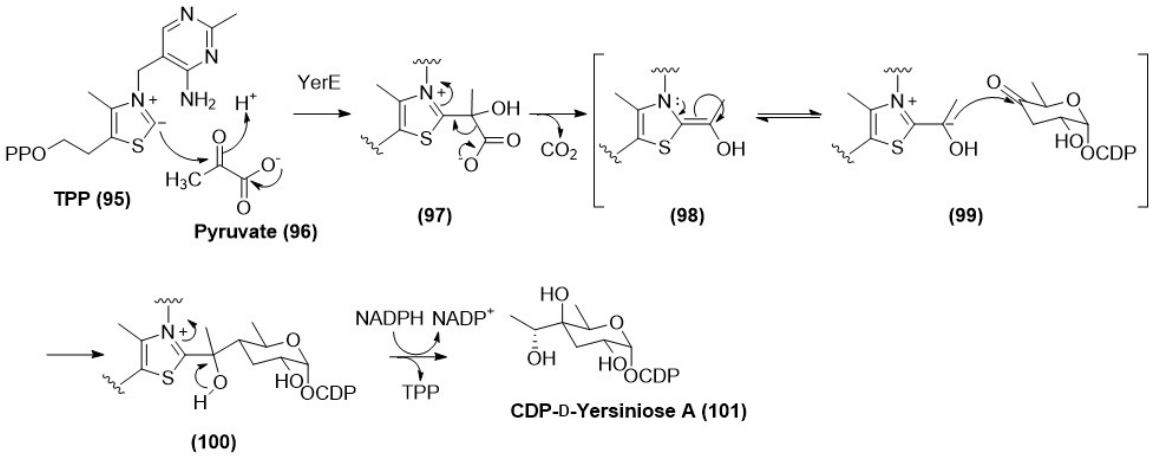


Figure 17. Mechanism of branched-sugar formation, D-yersiniose biosynthesis

1.5 Reduction

The most prevalent reaction in deoxysugar biosynthesis is ketoreduction. Ketoreductases are NAD(P)H-dependent and catalyze hydride reduction of C-3 and C-4-ketosugars to form corresponding secondary alcohols. Both axial and equatorial C-3 and C-4 hydroxyl groups can be generated. Ketoreductases are widespread due to the requirement of ketosugar intermediates which activate α-H for abstraction in many tailoring reactions including epimerization, alkylation (figure 16-17), and both 2 and 3-dehydration and for reactions that occur directly at the ketone as in 4-epimerization (figure 14) or 4-transamination (figure 11-12). C-2 dehydration also requires a 3-ketoreductase to stabilize and reduce the labile NDP-3,4-diketosugar intermediate (figure 6).

1.6 Conclusions

Deoxysugar components are key to the activity of glycosylated natural products including polyketides such as doxorubicin, mithramycin, avermectin, amphotericin B,

and others. The following guidelines for deoxysugar biosynthetic pathways have been derived from bioinformatic, biochemical, and genetic studies. Prior to derivatization, sugars are first activated through nucleotidyl transfer to yield NDP-sugars (figure 4). Second, 4,6-dehydration (figure 5) occurs prior to any other tailoring reactions for all 6-deoxyhexoses. Many reactions rely on the 4-keto group to activate C-3 and C-5 protons including C-3 or C-5 epimerization (figure 15), alkylation (figure 16-17), C-2 and C-3 dehydration (figure 7-8), as well as reactions requiring the 4-keto site directly as in 4-transamination (figure 11-12) or 4-ketoreduction. Next, C-6 deoxygenation precedes C-2, C-3, and C-4 deoxygenation. C-3-ketoreduction or 3-transamination of the unstable NDP-3,4-diketosugar intermediate immediately follows C-2 deoxygenation (figure 6). C-4 deoxygenation proceeds through 4-transamination (figure 10). Epimerization, ketoreduction, and group transfer reactions (except 4-transamination) follow deoxygenation steps.

Bacterial biosynthetic enzymes are typically encoded in gene clusters which allow horizontal gene transfer between organisms and provide genetic engineering tools for researchers interested in altering glycosylation patterns of existing scaffolds to improve their pharmacological properties. Together, understanding the mechanism of action of a compound, along with the comprehension of unusual sugar biosynthetic pathways and access to their biosynthetic machinery expands the drug development toolbox for derivatization of natural products to increase drug-target interaction and solubility.

1.7 Specific Aims

Organisms have evolved to produce natural products that are optimized for interaction with biological targets but that are not necessarily developed for druglikeness. While potent, MTM's clinical use as an anticancer agent has been limited by its toxicity. Combinatorial biosynthesis is one method for addressing these limitations and to increase drug-target interaction. Modification of glycosylation patterns is a popular strategy to diversify compounds since it has been shown that carbohydrate constituents of glycosylated natural products are often crucial for bioactivity. L-rednose is an extremely rare amino sugar and access to its biosynthetic pathway would provide an excellent building block to derivatize aminoglycoside antibiotics to circumvent resistance mechanisms or mithramycin towards more specific activity to decrease side effects. To effectively develop drugs for particular targets, complete understanding of the mechanism of action is key. Identification of the L-rednose biosynthetic pathway and investigation of mechanism of action of MTM was pursued through four specific aims.

1.7.1 Specific Aim 1: Identification of the *Streptomyces* sp. KY 40-1 Gene Cluster Involved in the Biosynthesis of Saquayamycins H and I (*sqn*)

Whole genome sequencing of *Streptomyces* sp. KY 40-1 along with primer walking to fill sequence gaps provided contigs and a comparative genomic approach was taken to identify the contigs containing the *sqn* gene cluster. The gene cluster provides the DNA sequences necessary to express the L-rednose (109 A) deoxysugar biosynthetic enzymes as well as those for L-aculose (103 A), D-olivose (121), L-rhodinose (124), and

L-cinerulose (123) all of which are attractive for future combinatorial biosynthetic manipulations of saccharide chain containing drugs.

1.7.2 Specific aim 2: Validation of the Proposed L-rednose Biosynthetic Pathway Up to the Glycosyl Transfer Through Enzymatic Synthesis of TDP-3,6-dideoxy-L-idosamine

Genes for proposed L-rednose biosynthetic enzymes were cloned and heterologously expressed in *E. coli*. Following expression and purification, enzymes were combined in a one pot reaction. Enzymatic synthesis of TDP-3,6-dideoxy-L-idosamine (132) would provide confirmation of the L-rednose biosynthetic pathway. Additionally, each step would generate TDP-amino sugars that could be used to alter glycosylation patterns in mithramycin and other compounds to increase drug-target interaction or in aminoglycosides to circumvent resistance mechanisms. Enzymatic synthesis of TDP-glucosamine (128), the first step in the proposed pathway was achieved, however later steps were unsuccessful.

1.8.3 Specific Aim 3: The Investigation of the Dimerization of MTM

Previous studies involving dimerization of MTM (138) were not carried out in physiologically relevant salt conditions and MTM analogues had not been investigated. Therefore, it was necessary to further investigate dimerization of MTM (138) and its analogues. MTM (138) and analogues were titrated into buffer containing divalent metals and physiologically relevant salt concentrations and change in fluorescence was monitored to determine dimerization. Although it was understood that DNA binding by

MTM (138) was dependent upon dimer formation, this study investigated if the MTM dimer formed upon binding, or in the absence of DNA and provided a better understanding of the mechanism of action of MTM (138) which is key in development of the drug as a chemotherapeutic.

1.8.4 Specific Aim 4: The Investigation of MTM's DNA Binding Preferences

Interest in MTM (138) has been renewed following a discovery of MTM (138) as a potent inhibitor of the driver of Ewing sarcoma, the oncogenic transcription factor EWS-FLI1. It is unclear how MTM (138) inhibits EWS-FLI1 since MTM (138) binds G/C rich DNA while EWS-FLI1 prefers GGAA microsatellites. Thus, it was necessary to further probe MTM's DNA binding preferences. MTM (138) and MTM analogues with 3-chain modifications were titrated into buffer containing physiologically relevant divalent metal and salt concentrations with DNA oligos of various sequence and confirmation and the change in fluorescence was monitored. This work identified whether modification of the MTM 3-side chain impacts DNA binding affinity, established a minimum MTM binding site on DNA, and determined if MTM (138) recognized DNA based on sequence or confirmation. These findings laid the foundation for future research based on the interaction between MTM (138), DNA, and EWS-FLI1 toward development of MTM analogues for the treatment of Ewing sarcoma.

2. IDENTIFICATION OF THE L-REDNOSE BIOSYNTHETIC PATHWAY IN SAQUAYAMYCINS H AND I

2.1 Introduction

Saquayamycins (SQN) are type II polyketide-derived angucycline antibiotics first discovered in 1985.[43] Fifteen SQN have been reported thus far including SQN A-K (102-112)[43, 44] and Z (113)[45] and three analogues (SQN A1 (114), SQN B1 (115), SQN C1 (116)) generated by partial acid hydrolysis of SQN A-C (102-104) [43] (figure 18). SQN contain the same aquayamycin aglycone and two saccharide chains (except for SQN A1-C1 (114-116) whose second saccharide chain is hydrolyzed) with C-glycosidic D-olivose attached at C-9 and O-glycosidic L-rhodinose attached at C-3 of the chromophore. Glycosylation patterns vary from there containing combinations of L-cinerulose, L-aculose, L-rhodinose, L-rednose, L-olivose, and in the case of SQN Z, L-oliose and 4-keto-L-aculose. SQN Z (113) is the largest with both a pentasaccharide and tetrasaccharide chain. Excluding SQN A1-C1, SQN G (108) is the smallest with one disaccharide and one monosaccharide chain. The remaining SQN all bear two disaccharide chains. In addition to antibiotic activity, SQN are also anti-cancer agents. SQN A-D (102-105) , isolated from *Streptomyces nodosus* MHI90-16F3 exhibited a cytostatic effect against P388 leukemia cells.[43] SQN E (106) and F (107) produced by *Actinomyces* strain MK290-AF1 are farnesyltransferase inhibitors,[46] which impede function of the cell signaling protein Ras whose overactive pathway is typical of cancer.[47] Saquayamycin Z (113) was isolated from *Micromonospora* sp.strain Tü6368 and displayed cytostatic effects against gastric epithelial HM02 and liver hepatocellular carcinoma HepG2 cells.[45] SQN G-K (108-112) produced by *Streptomyces* KY40-1 are cytotoxic, active against human prostate cancer PC3 cells and non-small cell lung cancer H460 cell lines.[44] SQN B (103) is the major product of *S.* KY40-1 and exhibited among the highest activity in both cell lines.[44] This is typical of the major product since

organisms evolve to produce the most potent compounds in the largest scale. The organism's goal is to produce efficacious antibiotics to outcompete other bacteria in the same microbiota. It is not evolutionarily advantageous to spend resources to make large quantities of ineffective metabolites. Interestingly, SQN H (109) showed slightly higher activity than SQN B (103) against H460 cells.[44] SQN H (109) and B (108) differ only in the second sugar of the C-3 disaccharide chain, L-rednose vs. L-aculose respectively, indicating the amino group located at β-position of the L-aculose moiety impacts activity.[44]

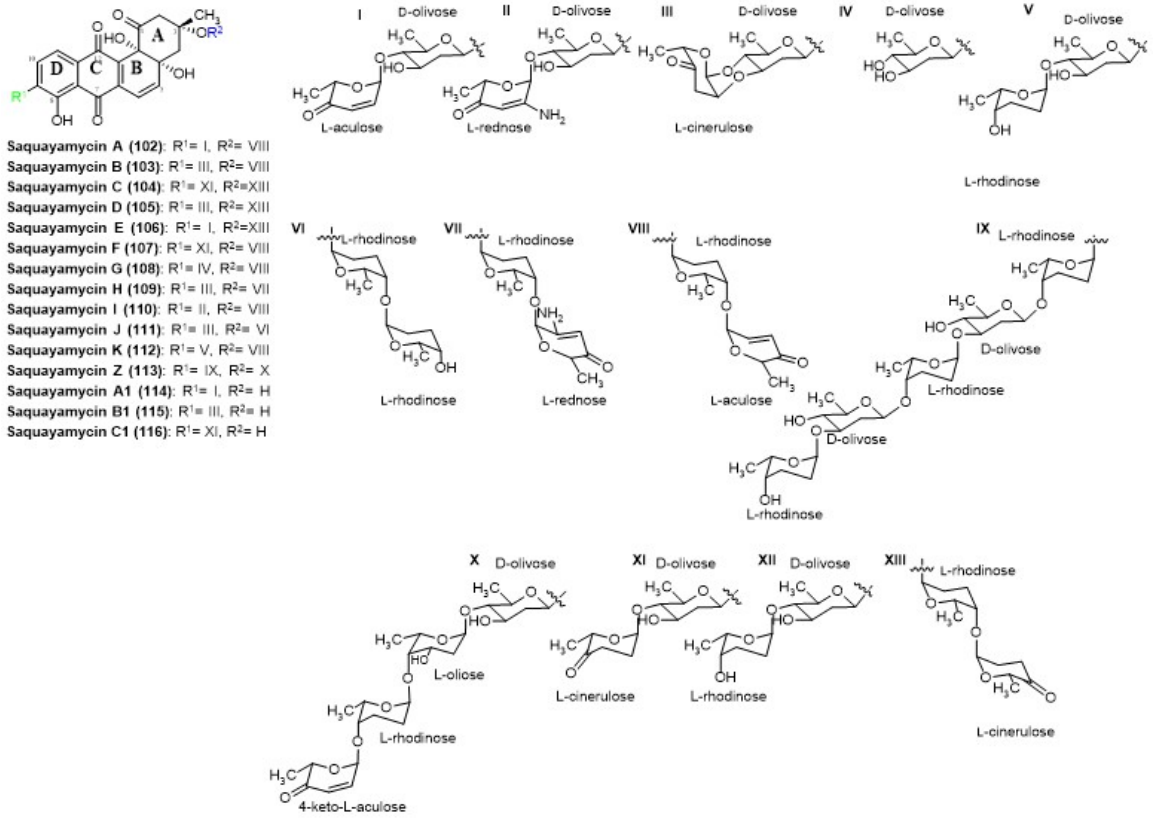


Figure 18. Chemical structures of the saquayamycins

Of particular interest are SQN H (109) and I (110) due to their extremely rare amino sugar moiety rednose, which was previously only reported once, in an

anthracycline rudolfomycin.[43] SQN H (109) and I (110) differ in the second sugar in both disaccharides where saquayamycin H (109) has L-rednose O-glycosidically linked to L-rhodinose and L-cinerulose O-glycosidically linked to D-olivose and saquayamycin I (110) contains O-glycosidic L-aculose attached to L-rhodinose and O-glycosidic L-rednose attached to D-olivose.

Glycodiversification through the incorporation of amino sugars is one important strategy to modulate bioactivity and solubility of glycosylated natural products. Many natural products rely on amino sugars for their bioactivity. Primarily, the broad-spectrum aminoglycoside antibiotics which inhibit protein synthesis by binding to the bacterial 30S ribosomal subunit causing misreading of t-RNA. Aminoglycosides rely on the biosynthesis of diverse amino sugars and specialized glycosyltransferases for their attachment. An important resistance mechanism for bacteria is the use of modifying enzymes, which acetylate, adenylate, or phosphorylate aminoglycosides preventing ribosomal binding. Incorporation of L-rednose (109 A), which lacks the 3-OH or 3-NH_2 handle typically recognized by modifying enzymes, would potentially circumvent aminoglycoside inactivation. Amino sugars are also key components of macrolides such as erythromycin which rely on D-desosamine to bind the bacterial 50S ribosomal subunit to block tRNA and inhibit protein translation.[48, 49] Widely used anthracycline anticancer agents doxorubicin and daunorubicin both contain the amino sugar L-daunosamine which plays a role in DNA binding.[50, 51] These compounds intercalate DNA inhibiting topoisomerase II function disrupting DNA and RNA synthesis.[52, 53] The incorporation of L-rednose (109 A) into aminoglycosides or polyketides such as mithramycin for example, may modulate bioactivity and increase the interaction between the drug and its target.

Access to the gene cluster encoding biosynthetic enzymes for L-rednose (109 A) and understanding of the biosynthetic pathway will add to the toolbox for glycodiversification. Enzymatic synthesis of key intermediates will not only validate the L-rednose pathway but generate novel TDP-amino sugars to be used in artificial pathways to generate novel compounds.

2.2 Results and Discussion

2.2.1 Identification of the Saquayamycin Biosynthetic Gene Cluster

This section was reproduced in part with permission from ACS Chemical Biology, submitted for publication. Unpublished work copyright 2017 American Chemical Society.

To identify which contig the *sqn* cluster was located on, all reads were initially mapped to the *Streptomyces* sp. SCC 2136 Sch 47554/ Sch 47555 (*sch*) gene cluster.[54] Sch 47554 (119) and Sch 47555 (120) share the same chromophore as SQN produced by *Streptomyces* sp. KY40-1 and BLAST analysis revealed high similarity between *sqn* and *sch* minimal PKS genes (79-90%). However, the carbohydrate residues differ from SQN which have C-glycosidic D-olivose attached to C-9 and O-glycosidic L-rhodinose attached to C-3 while Sch bear O-glycosidic L-aculose attached to C-9 and C-glycosidic D-amicetose attached to C-3 (figure 19). Following the publication of the *Streptomyces lusitanus* SCSIOLR32 grincamycin (*gcn*) gene cluster,[55] *sqn* contigs were mapped to the *gcn* cluster. Grincamycin (117) and SQN share the same aquayamycin aglycone and O-glycosidic L-rhodinose attached to C-3 exhibiting a stronger structural similarity than between SQN and Sch (figure 19). Heterologous expression of the *gcn* gene cluster in

Streptomyces coelicolor M512 yielded P-1894B (118), containing L-aculose in place of L-cinerulose as the last sugar in the trisaccharide chain as well as the disaccharide chain.[55] It was shown that GcnQ, a FAD/FMN containing dehydrogenase in the *gcn* cluster can convert L-rhodinose to L-aculose.[55] Taking both grincamycin (117) and P-1894B (118), all of the sugars in SQN A-B (102-103), and G-K (108-112) are found (D-olivose, L-rhodinose, L-cinerulose, and L-aculose) except for L-rednose (figure 19). It was expected that comparison of the deoxysugar genes in both clusters would reveal the unique genes in *S.* KY40-1 that could be responsible for biosynthesis of rednose. The N-terminal border region of the *sqn* gene cluster was determined by the location of a primary metabolism gene (lactate transporter) upstream of *sqnZ* indicating *sqnZ* as the beginning of the *sqn* cluster. Downstream of *sqnY* lies a 1.5 kb non-coding region of DNA, indicative of the C-terminal border region of the *sqn* cluster, followed by an unknown protein and an ABC transporter. To confirm this as a non-coding region and not a sequencing error, this region was PCR amplified and sequenced again which resulted in a 100% sequence identity to the genomic sequencing data.

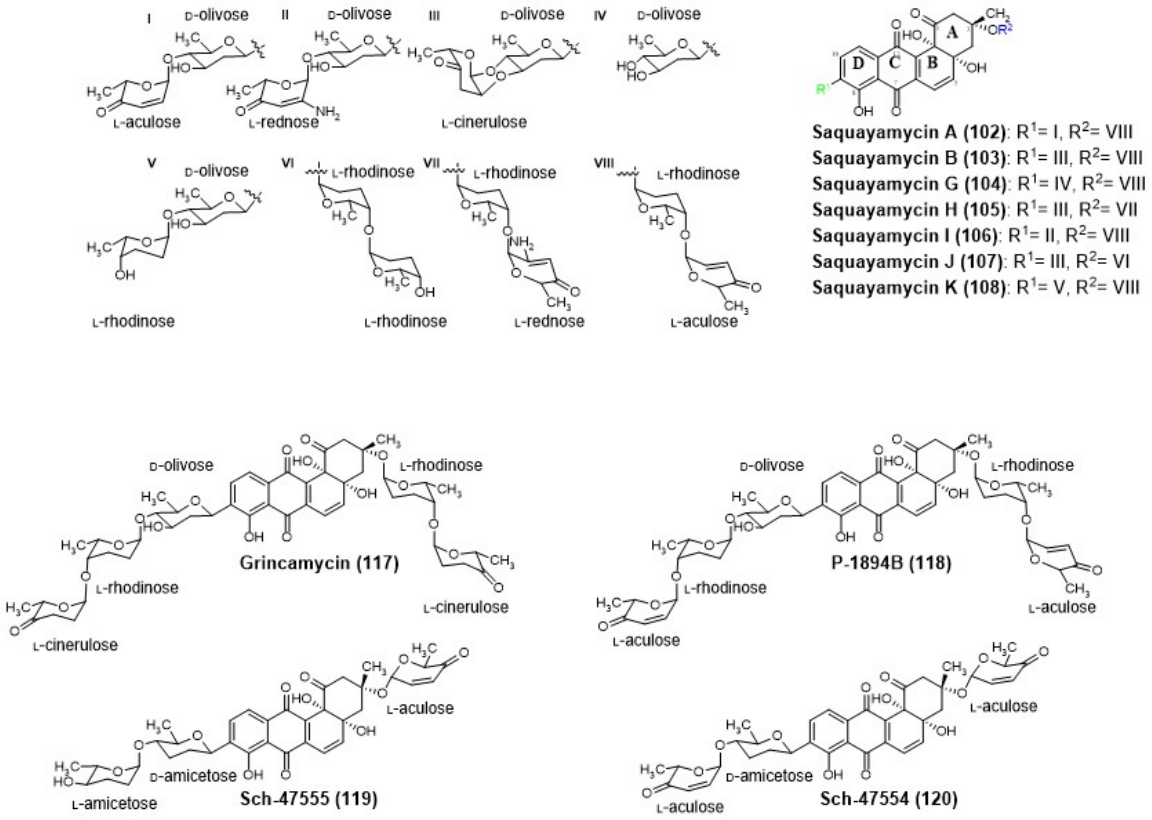


Figure 19. Chemical structures of saquayamycins A, B and G-K, grincamycin, P-1894B, Sch 42554, and Sch 47555.

The *Streptomyces* sp. KY40-1 *sqn* biosynthetic gene cluster spans 44.7 kb and consists of 40 open reading frames (figure 20). The gene cluster encodes the minimal PKS (SqnHIJ), 2 cyclases (SqnBBL), a ketoreductase (SqnK), and a 4'-phosphopantetheinyl transferase (SqnCC) responsible for polyketide biosynthesis. Also found in the cluster were seven oxidoreductases (SqnACFMTUW) and a decarboxylase (SqnP) for post modifications, eight proteins involved in deoxysugar biosynthesis (SqnS1-S8), three glycosyltransferases for sugar attachment (SqnG1-G3), a dehydrogenase (SqnQ) for post-sugar modification, three regulatory proteins (SqnDRV), four transporters (SqnZBNX), and six proteins with unknown functions

(SqnAAEODDEEY). Interestingly, the *sqn* gene cluster lacks an aminotransferase that we would expect to see for the biosynthesis of the amino sugar rednose.

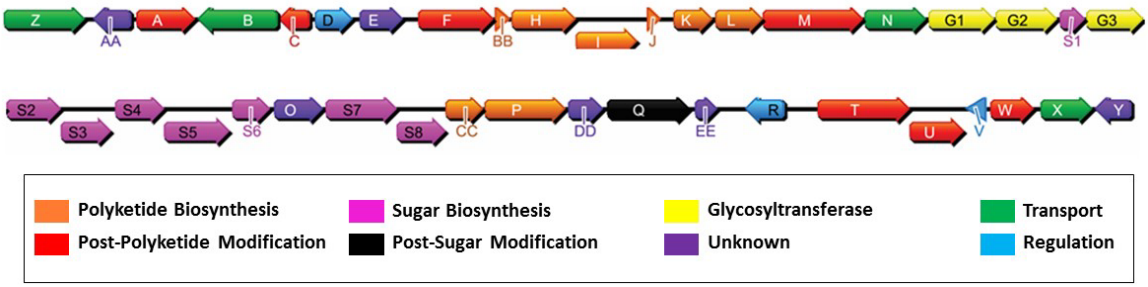


Figure 20. Saquayamycin biosynthetic gene cluster

Table 1. Proposed functions for ORFs in *sqn* gene cluster

ORF	Putative Function	Protein Homologue, Strain	% Identity	Accession #
SqnZ	major facilitator superfamily transporter, efflux	*Streptomyces olivochromogenes*	88	KUN37937.1
SqnAA	hypothetical protein	*Streptomyces sp. LaPpAH-108*	70	WP_018546489.1
SqnA	putative oxidoreductase	GcnA, *Streptomyces lusitanus*	86	AGO50604.1
SqnB	major facilitator superfamily transporter, efflux	GcnB, *Streptomyces lusitanus*	84	AGO50605.1
SqnC	NADPH-dependent FMN reductase	UrdO, *Streptomyces fradiae*	62	WP_059130309.1
SqnD	TetR family transcriptional regulator	GcnD, *Streptomyces lusitanus*	85	AGO50607.1
SqnE	hypothetical protein	SaqP *Micromonospora sp. Tu 6368*	46	ACP19350.1

SqnF	oxidoreductase	UrdE, *Streptomyces fradiae*	90	CAA60567.1
SqnBB	cyclase	UrdF, *Streptomyces fradiae*	86	CAA60568.1
SqnH	ketoacylsynthase alpha subunit	UrdA, *Streptomyces fradiae*	94	CAA60569.1
SqnI	ketoacyl synthase CLF	UrdB, *Streptomyces fradiae*	87	CAA60570.1
SqnJ	acyl carrier protein	UrdC, *Streptomyces fradiae*	92	CAA60571.1
SqnK	ketoreductase	UrdD, *Streptomyces fradiae*	91	CAA60572.1
SqnL	cyclase	UrdL, *Streptomyces fradiae*	88	AAF00205.1
SqnM	oxygenase-reductase	UrdM, *Streptomyces fradiae*	80	AFU51427.1
SqnN	major facilitator superfamily transporter, efflux	UrdJ2, *Streptomyces fradiae*	65	AAF00207.1
SqnG1	glycosyltransferase	SaqGT2, *Micromonospora sp. Tu 6368*	55	ACP19363.1
SqnG2	glycosyltransferase	SaqGT4, *Micromonospora sp. Tu 6368*	57	ACP19365.1
SqnS1	NDP-hexose 3,5-epimerase	UrdZ1, *Streptomyces fradiae*	68	AAF00208.1
SqnG3	glycosyltransferase	UrdGT2, *Streptomyces fradiae*	75	AAF00209.1
SqnS2	glucose-1-phosphate thymidylyltransferase	UrdG, *Streptomyces fradiae*	80	AAF00210.1
SqnS3	NDP-glucose 4, 6-dehydratase	UrdH, *Streptomyces fradiae*	87	AAF00211.1
SqnS4	NDP-hexose 4-ketoreductase	UrdZ3, *Streptomyces fradiae*	67	AAF72549.1
SqnS5	NDP-hexose 3,4-dehydratase	UrdQ, *Streptomyces fradiae*	91	AAF72550.1

SqnS6	NDP-hexose 4-ketoreductase	UrdR, *Streptomyces fradiae*	82	AAF72551.1
SqnO	hypothetical protein	GcnO, *Streptomyces lusitanus*	72	AGO50626.1
SqnS7	NDP-hexose 2,3-dehydratase	UrdS, *Streptomyces fradiae*	81	AAF72552.1
SqnS8	NDP-hexose 3-ketoreductase	SaqT *Micromonospora sp. Tu 6368*	67	ACP19378.1
SqnCC	4'-phosphopantetheinyl transferase	SprT, *Streptomyces sp. TK08046*	79	BAV17018.1
SqnP	methylmalonyl-CoA decarboxylase	*Streptomyces nodosus*	95	WP_043444110.1
SqnDD	hypothetical Protein	*Micromonospora sp. Tu 6368*	65	ACP19381.1
SqnQ	FAD/FMN-containing dehydrogenase	AknOx, *Streptomyces galilaeus ATCC 31615*	56	BAB72054.1
SqnEE	hypothetical Protein	*Streptomyces nodosus*	78	WP_043449810.1
SqnR	OmpR family transcriptional regulator	GcnR, *Streptomyces lusitanus*	88	AGO50631.1
SqnT	2-oxoacid: ferredoxin oxidoreductase subunit alpha	GcnT, *Streptomyces lusitanus*	90	AGO50632.1
SqnU	2-oxoacid: ferredoxin oxidoreductase subunit beta	GcnU, *Streptomyces lusitanus*	84	AGO50633.1
SqnV	HxlR family transcriptional regulator	*Streptomyces sp. TP-A0356*	86	WP_055493122.1
SqnW	NAD(P)-dependent oxidoreductase	*Streptomyces sp. TP-A0356*	99	WP_055493121.1

SqnX	transporter	RarD, *Streptomyces sp. 142MFCol3.1*	88	WP_051371622.1
SqnY	hypothetical protein	*Streptomyces avermitilis*	72	WP_030379785.1

Database comparison of the deduced deoxysugar biosynthetic proteins are shown in table 1. All but one (SqnS8), of the deduced deoxysugar biosynthetic proteins showed highest similarity with homologs from the urdamycin producer *Streptomyces fradiae.*[56] SqnS1, a 164-amino acid (aa) protein, revealed highest similarity to UrdZ1 (accession number AAF00208.1), a NDP-hexose 3,5-epimerase homolog (67% identical aa)[57]. The deduced 355 amino acid sequence of SqnS2 should yield a 37.192 Da MW protein with highest homology (80% identical aa) to UrdG (accession number AAF00210.1), a glucose synthetase homolog. SqnS3, a 327-aa protein with 36.369 kDa MW shows highest similarity to UrdH (accession number AAF00211.1), a NDP-hexose-4,6-dehydratase homolog (87% identical aa). SqnS4, a 317-amino acid (aa) protein (32525.16 kDa MW), has closest similarity to UrdZ3 (accession number: AAF72549.1), a NDP-hexose 4-ketoreductase homolog (67% identical aa). The deduced protein of SqnS5 has 434 aa (molecular weight 48.104 kDa) and shows highest similarity to UrdQ (accession number AAF72550.1), a NDP-hexose 3,4-dehydratase homolog (91% identical aa). SqnS6 encodes a 254-aa protein with 27.296 kDa molecular weight most resembling UrdR (accession number AAF72551.1), a NDP-hexose 4-ketoreductase homolog (82% identical aa). The SqnS7 reading frame should yield a 466-aa protein of 51.995 kDa molecular weight and shows highest homology (81% identical aa) to UrdS, a NDP-hexose 2,3dehydratase homolog. The deduced protein of 318 aa and 33.820 kDa molecular weight, SqnS8, exhibits highest homology to SaqT (accession number

ACP19378.1), a NDP-hexose 3-ketoreductase homolog from the saquayamycin Z producer *Micromonospora sp. Tu 6368* (67% identical aa).[58] Downstream from SqnS8 lies an open reading frame that should yield a 542-aa protein with a molecular weight of 58.745 kDa with highest homology (56% identical aa) to AknOx (accession number BAB72054.1), a FAD/FMN-containing dehydrogenase homolog from the aclacinomycin producer *Streptomyces galilaeus*.[59]

2.2.2 Derivation of Hypothetic L-rednose Biosynthetic Pathway

Comparison of putative proteins encoded by the *sqn* cluster involved in biosynthesis of deoxysugar moieties other than L-rednose (109 A) should reveal any unique enzymes that could be responsible for L-rednose biosynthesis. Based on the sequencing data, the biosynthetic pathways from glucose 1-phosphate (11) to NDP-D-olivose (121), NDP-L-cinerulose (123), NDP-L-rhodinose (124) and NDP-L-aculose (103 A) together should comprise nine steps including NDP-D-glucose synthesis[60] (SqnS2), a 4,6-dehydratation[61] (SqnS3), a 2,3-dehydratation (SqnS7) and a 3-ketoreduction (SqnS8)[10] to generate NDP-4-keto-2,6-dideoxy-D-glucose (23). The remaining steps diverge with a 4-ketoreduction (SqnS6) to yield NDP-D-olivose (121), or a 3-deoxygenation (SqnS5), and 5-epimerization (SqnS1) to generate NDP-L-cinerulose (123). A subsequent 4-ketoreduction (SqnS4) of NDP-L-cinerulose (123) generates NDP-L-rhodinose (124) followed by an oxidation (SqnQ) to prepare L-aculose (103 A). Based on the reaction catalyzed by the SqnQ homolog, AknOx from the aclacinomycin gene cluster where the final sugar oxidation from rhodinose to aculose is carried out after rhodinose is transferred to the acceptor,[59] we propose that NDP-L-rhodinose (124) is

transferred by a glycosyltransferase preceding the C-4 oxidation and subsequent introduction of a double bond between C-2 and C-3 to yield the final L-aculose (103 A) (figure 21). Together, these biosynthetic pathways require all the deoxysugar biosynthetic proteins encoded by the *sqn* gene cluster and do not reveal any enzymes unique to L-rednose biosynthesis. Additionally, the *sqn* gene cluster lacks an aminotransferase that we would expect to see for the biosynthesis of the amino sugar rednose. These findings led to the proposal of a unique biosynthetic pathway beginning with glucosamine 6-phosphate (126) in contrast with the typical glucose 6-phosphate (10) and suggests that the Sqn deoxysugar biosynthetic enzymes are permissive enough to accept both glucose and glucosamine derived substrates.

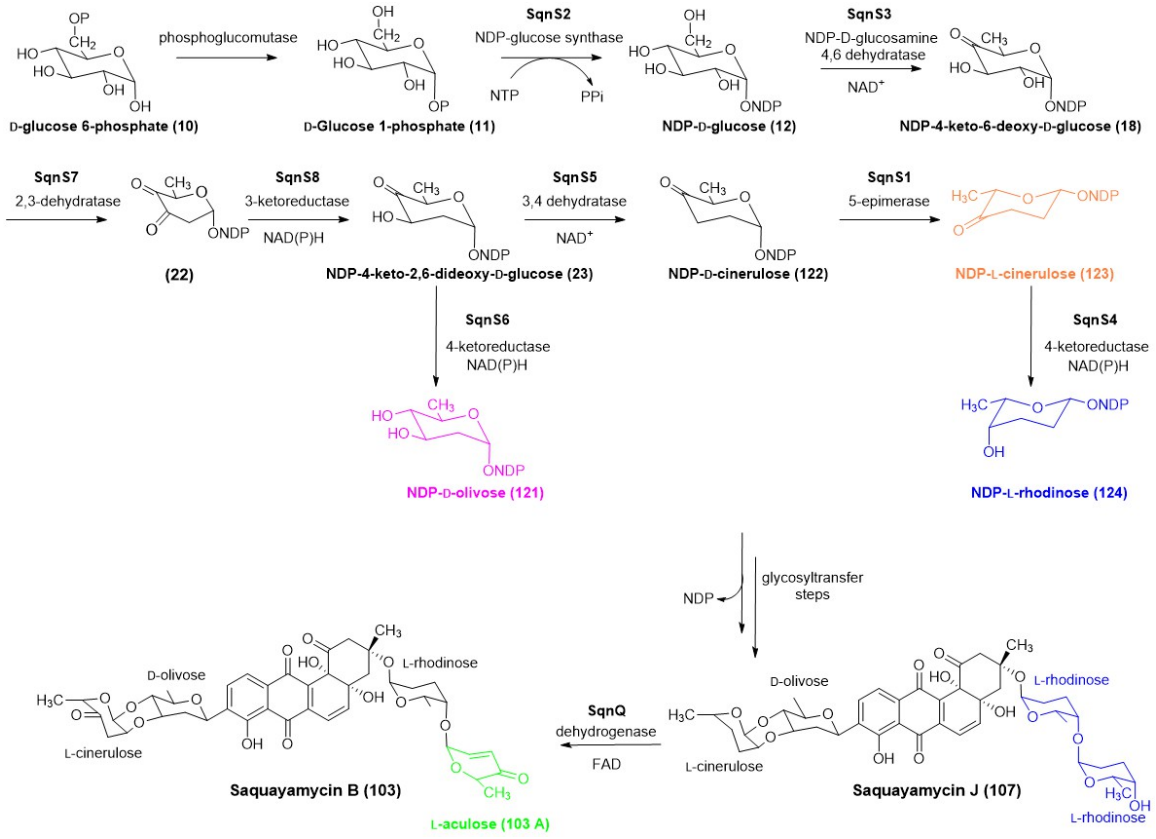


Figure 21. Biosynthetic pathway to NDP-D-olivose, NDP-L-cinerulose, NDP-L-rhodinose, and NDP-L-aculose in *S.* KY 40-1.

The proposed L-rednose biosynthetic pathway comprises seven steps from glucosamine 1-phosphate (127) to L-rednose (109 A) (figure 22). Generation of the precursor is based on the *N*-acetyl-D-glucosamine biosynthetic pathway required for peptidoglycan and lipopolysaccharides in bacterial cell wall biosynthesis where glutamine-fructose-6-phosphate transaminase carries out both an isomerization and transamination to generate D-glucosamine 6-phosphate (126)[62] followed by an isomerization by phosphoglucosamine mutase to yield D-glucosamine 1-phosphate (127).[63] It is also possible that the precursor is derived from *N*-acetyl-D-glucosamine 1-

phosphate but would require a deacetylation step en route to L-rednose (109 A). Since there is not a deacetylase in the *sqn* cluster, it would require a general deacetylase encoded elsewhere in the chromosome. Another possibility is that the amine comes from a transaminase outside of the gene cluster, but this would require a C-2 oxidation to generate the ketone necessary for the transamination. The gene cluster lacks a 2-oxidoreductase as well, so this hypothesis seems less likely.

Based on sequencing data, the first five secondary metabolism steps are analogous to rhodinose biosynthesis in the urdamycin (*urd*)[56] biosynthetic pathway albeit replacing the glucose substrate with its glucosamine counterpart. First, D-glucosamine 1-phosphate (127) is activated by glucosamine/glucose synthase (SqnS2) to become NDP-D-glucosamine (128). Second, an NAD-dependent 4,6-dehydratase (SqnS3) catalyzes C-6 deoxygenation and C-4 oxidation to yield NDP-4-keto-6-deoxy-D-glucosamine (129) analogous to the common intermediate NDP-4-keto-6-deoxy-D-glucose (18) in typical deoxyhexose biosynthesis. Subsequently, a 3-deoxygenation is carried out by 3,6-dehydratase (SqnS5) to yield NDP-4-keto-3,6-dideoxy-D-glucosamine (130). An epimerization by a 5-epimerase (SqnS1) follows, resulting in NDP-4-keto-3,6-dideoxy-L-mannosamine (131). Due to the high homology between SqnS5 and UrdQ (91% identical aa), and between SqnS1 and UrdZ3 (68% identical aa) it is likely that the 3-deoxygenation precedes epimerization as in the urdamycin biosynthetic pathway.[56] The production of urdamycin M (134) containing both D-rhodinose and L-rhodinose moieties by a *urdR* (4-ketoreductase) mutant, clearly demonstrates that the epimerization follows the 3-deoxygenation (Figure 5).[56] If the epimerization occurred prior to 3-deoxygenation,

only L-rhodinose would be generated.[56] Next, a reduction of the C-4 ketone is catalyzed by a ketoreductase (SqnS4) to generate NDP-3,6-dideoxy-L-idosamine (132). Finally, we propose that NDP-3,6-dideoxy-L-idosamine (132) is transferred by a glycosyltransferase prior to the oxidation of C-4 and subsequent introduction of a double bond between C-2 and C-3 by the dehydrogenase SqnQ to yield the final L-rednose (109 A) (figure 22).

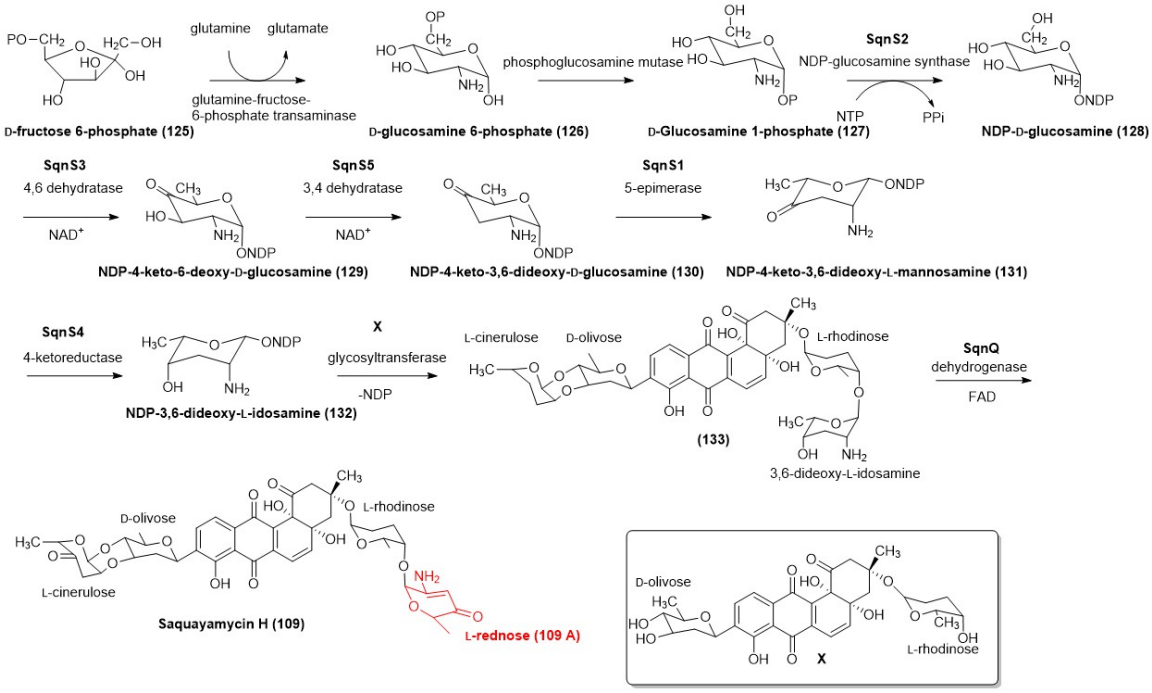


Figure 22. Proposed L-rednose biosynthetic pathway

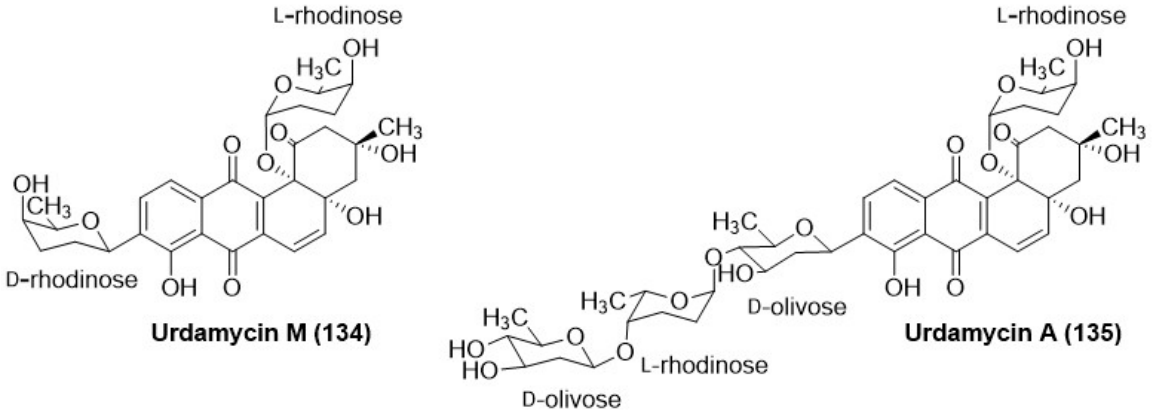
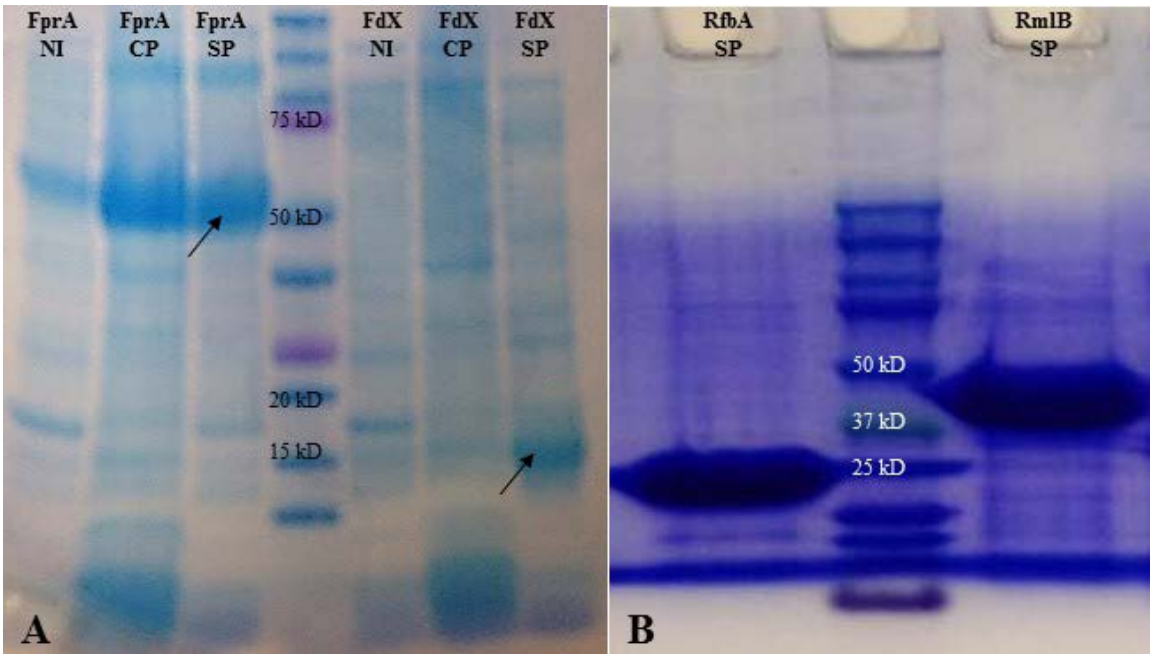


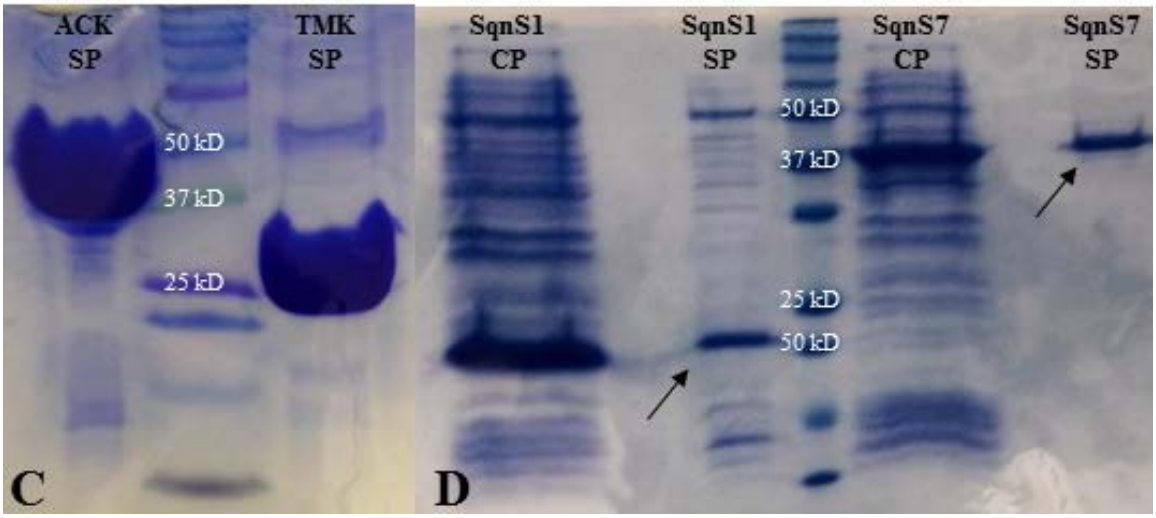
Figure 23. Chemical structures of urdamycins M and A

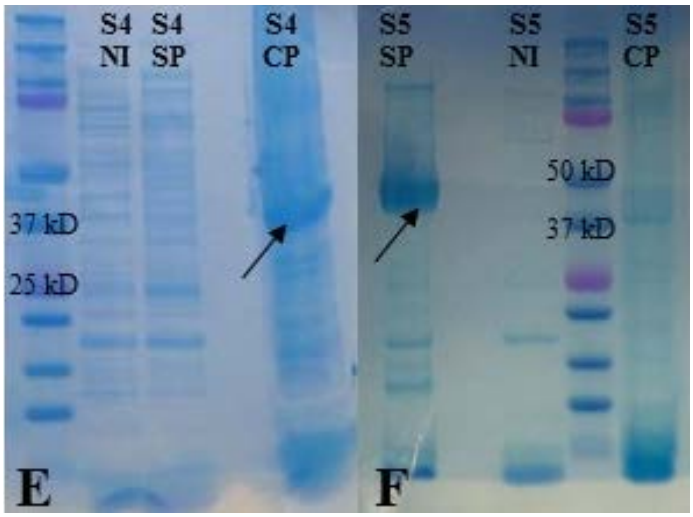
2.2.3 Enzymatic Synthesis of NDP-3,6-dideoxy-L-idosamine

To validate the proposed L-rednose (109 A) biosynthetic pathway up to the glycosyl transfer, the generation of NDP-3,6-dideoxy-L-idosamine (132), a mix-and-match method was employed. Previous constructs included those for *rfbA* encoding a thymidyltransferase SqnS2 homolog (30% identical aa) and *rmlB*, encoding a TDP-D-glucose-4,6 dehydratase SqnS3 homolog (42.4% identical aa) from *Salmonella typhimurium.*[64] Additional deoxysugar enzymes including SqnS1, and SqnS4-8 were heterologously expressed in *E. coli*, and purified through immobilized metal affinity chromatography (IMAC) (figure 24). All except for SqnS4 and SqnS6 resulted in soluble protein after expression in *E. coli* using a *N*-terminal 6x polyhistidine (His6) tag and subsequent purification, and the sizes observed on the sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) were in good agreement with calculated values (figure 24). SqnS4 expression was negligible and SqnS6 was largely insoluble. While expression of all the Sqn deoxysugar enzymes would provide valuable tools for genetic engineering, the priority was the expression of those predicted in L-rednose biosynthesis. So, it was critical to have SqnS4 in the arsenal. To increase expression of

SqnS4, isopropyl-*β*-D-thiogalactoside (IPTG) concentration and induction temperature were varied as well as expression with a *C*-terminal His_6 tag attempted. However, the expression level did not improve. Following the rationale that the RNA polymerase in *E. coli* may have difficulty reading codons from *Streptomyces*, SqnS4 was then codon optimized for expression in *E. coli*. This greatly improved expression of SqnS4, however it was completely insoluble (figure 24). This is likely due to improper protein folding. The next step would be to try heterologous expression in *S. lividans* to increase solubility of SqnS4. As it is expected that the 4-ketoreduction is the final step before glycosyl transfer, it was possible to continue forward with the earlier steps in L-rednose (109 A) biosynthesis.


FprA NI
FprA CP
FprA SP
75 kD
50 kD
20 kD
15 kD
FdX NI
FdX CP
FdX SP
A
RfbA SP
RmlB SP
50 kD
37 kD
25 kD
B


ACK SP
50 kD
37 kD
25 kD
TMK SP
C
SqnS1 CP
SqnS1 SP
50 kD
37 kD
25 kD
50 kD
SqnS7 CP
SqnS7 SP
D



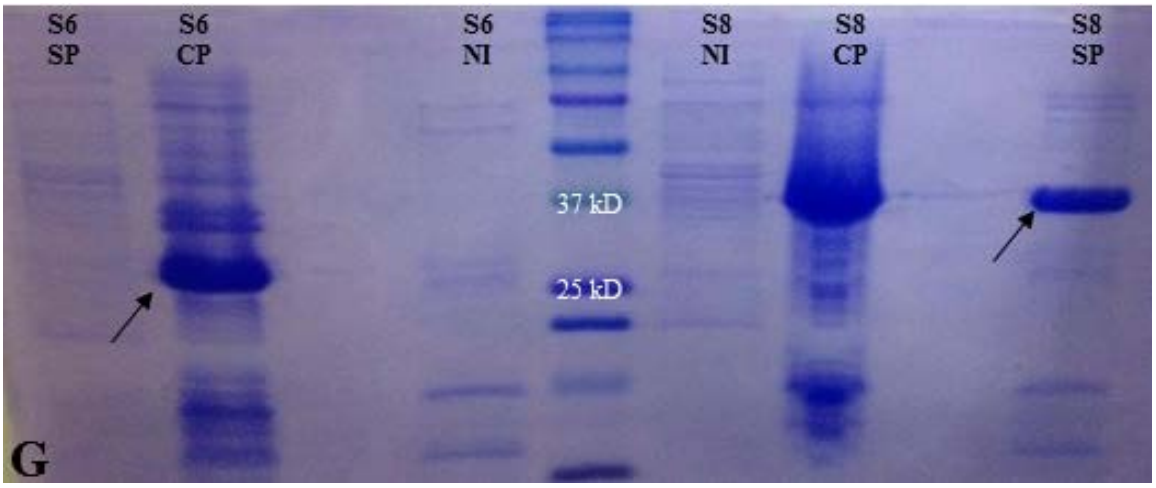


Figure 24. SDS PAGE analyses of the purified proteins. A: Lanes 1-3, MT FprA (49.3 kDa); Lanes 5-6 MT FdX (11.8), B: Lane 1 RfbA (32.7 kDa), Lane 3 RmlB (40.7 kDa); C: Lane 1 ACK (43.3 kDa), Lane 3 TMK (23.8 kDa), D: Lanes 1,3 SqnS1 (18.6 kDa), Lanes 5,6 SqnS7 (53 kDa), E: SqnS4 (33.5 kDa), F: SqnS5 (49.1 kDa), G: Lanes 1-4 SqnS6 (28.3 kDa), Lanes 6-9 SqnS8 (34.8 kDa). Abbreviations: Non-induced (NI), Soluble portion (SP), Cell Pellet (CP).

All but one (SqnS4) of the biosynthetic enzymes were in hand including SqnS2 homolog RfbA and SqnS3 homolog RmlB. These homologs represent well studied enzymes thymidyltransferases and 4,6-dehydratases and it was expected that together they should generate TDP-4-keto-6-deoxy-D-glucosamine (18). TDP (137) is the most common NDP used to activate deoxysugars, and high homology of Sqn enzymes to Urd, deoxysugar enzymes that recognize TDP (137), is an indication that Sqn enzymes likely use TDP (137). In order to be cost effective, thymidine-triphosphate (TTP) (138) was generated in situ using a system previously devised where thymidine monophosphate kinase (TMK) converts ATP to ADP to phosphorylate TMP (136) and generate TDP (137). Acetate kinase (ACK) further phosphorylates TDP (137) to TTP (138) and regenerates ATP using a phosphate from the conversion of acetyl phosphate to acetic acid (Figure 25). [64] RmlB, Rfba, TMK, and ACK were expressed in *E. coli* and purified as previously reported[64] (Figure 24) .Considering that TDP (137) can serve as an inhibitor for the downstream pathway enzymes, overall TDP-3,6-dideoxy-L-idosamine biosynthesis (132) was divided into two stages in the same pot. In the first stage, TMP (136), glucosamine 1-phosphate (127), and ATP were incubated with ACK, TMK, RfbA and RmlB. For the second stage, ACK, TMK, RfbA, and RmlB were removed and SqnS5 and SqnS1 along with NADPH and PMP were added. SqnS5 is a PMP dependent 3,4-dehydratase with a [2S·2Fe] center that must be continually regenerated in order to reduce the $\Delta^{3,4}$-glucoseen intermediate (27) to generate TDP-4-keto-3,6-dideoxy-D-glucosamine (130). This requires a second enzyme, either a reductase that has evolved to be specific for the dehydratase or a general cellular reductase. There is not a designated reductase for 3-deoxygenation in the *sqn* gene cluster akin to urdamycin[56], landomycin[65],

and spinosyn.[21] A pair of enzymes from *Mycobacterium tuberculosis* flavodoxin (Mt Fdx) and flavodoxin reductase (Mt FprA) were added to generate an electron transport chain. Mt Fdx oxidizes NADPH to reduce FAD, to reduce the [2S·2Fe] center in Mt FprA, which in turn reduces the [2S·2Fe] center in the dehydratase SqnS5. MtFdx and Mt FprA were heterologously expressed in *E. coli* and purified using TALON-IMAC (figure 24). The SqnS5 reaction was also carried out in anaerobic conditions in a glovebox using reducing agents in place of Mt Fdx/Mt FprA. After the one-pot enzymatic synthesis, MS data verified the production of only TDP-glucosamine (128) (figure 27). Further conversion of TDP-glucosamine (128) was unsuccessful. To rule out *N*-acetylglucosamine as the precursor, the first two steps of the reaction were also carried out using *N*-acetyl-D-glucosamine 1-phosphate in place of glucosamine 1-phosphate (127). MS data verified the production of only TDP-*N*-acetyl-D-glucosamine (figure 28). The dehydration to TDP-4-keto-6-deoxy-*N*-acetyl-D-glucosamine was unsuccessful (figure 26). Since the biosynthesis did not proceed past the activation of either TDP-glucosamine (128) or TDP-*N*-acetyl-D-glucosamine, it is likely that RmlB, the SqnS3 homolog, is not flexible enough to accept the glucosamine derived substrate. It may be very substrate-specific for TDP-glucose (12). To rule this out, SqnS3 needs to be overexpressed, purified, and used in place of RmlB, however due to funding and time constraints, this was not possible within the scope of this dissertation.

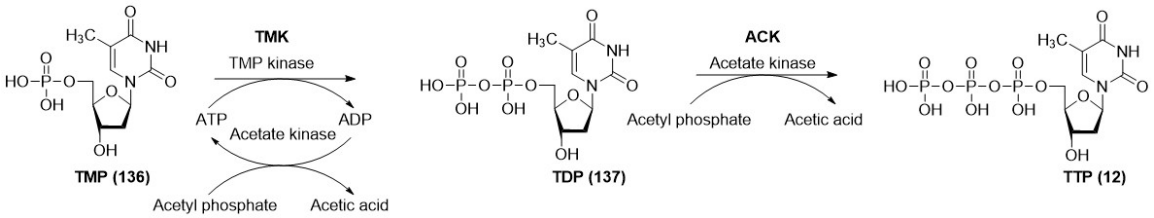


Figure 25. Enzymatic synthesis of TTP

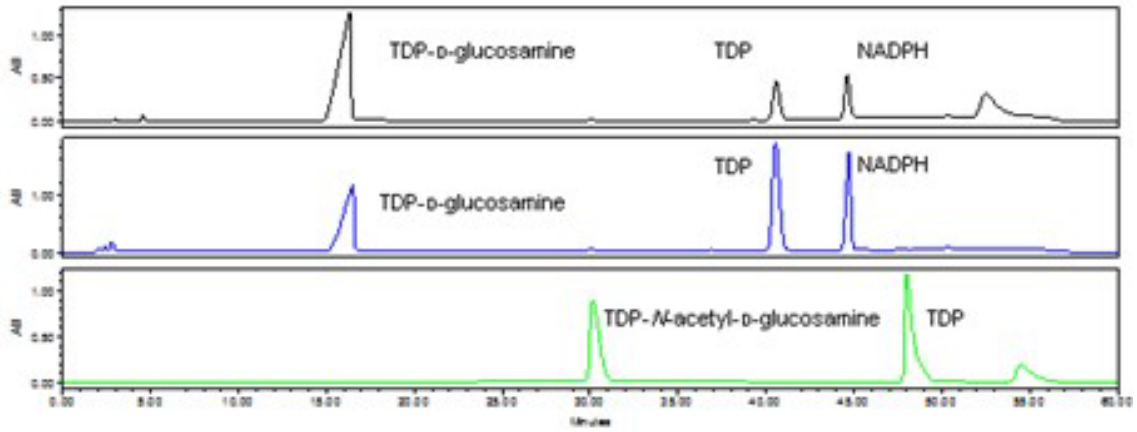


Figure 26. Representative HPLC traces. Panel A. glucosamine-1-phosphate (127), TMP (136), PMP, ATP, NAD^+ and NADPH incubated with ACK, TMK, RfbA, RmlB, SqnS5, Mt Fdx, Mt FprA and SqnS1 under aerobic conditions. Panel B. Reaction mixture from panel A with reducing agents replacing Mt Fdx/Mt FprA for SqnS5 catalysis in anaerobic conditions. Panel C. *N*-acetylglucosamine-1-phosphate, TMP (136), ATP, and NAD^+ incubated with ACK, TMK, RfbA, RmlB

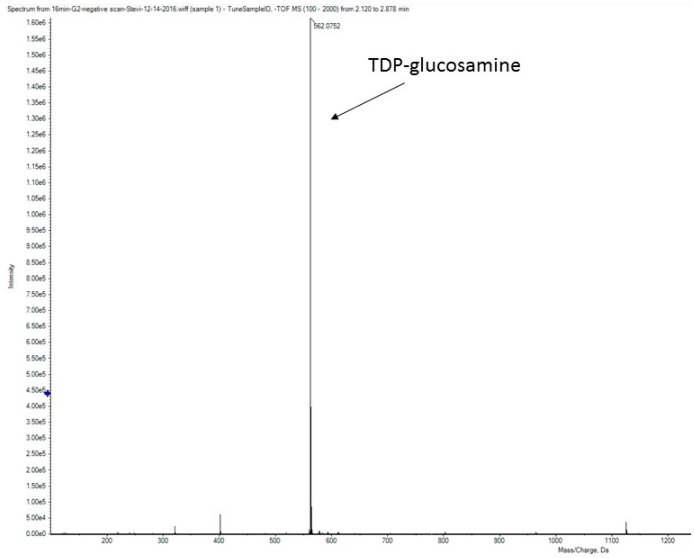


Figure 27. Representative HR-MS spectrum (negative mode). TDP-D-glucosamine (128), calculated molecular weight 563.09.

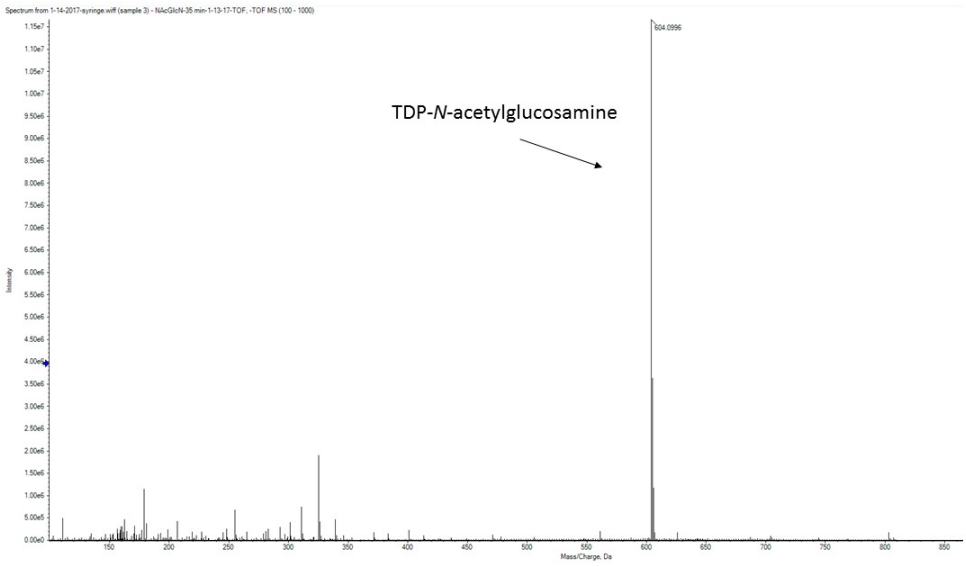


Figure 28. Representative HR-MS spectrum (negative mode). TDP-*N*-acetyl-D-glucosamine, calculated molecular weight 605.10.

2.3 Methods

2.3.1 DNA Sequencing and Analysis:

A sequence was obtained from a pure culture containing only *Streptomyces* sp. KY40-1. This organism was originally isolated from a soil sample collected near Cave Run Lake, Kentucky, USA at the foothills of the Appalachian Mountains[58]. From a glycerol stock, *Streptomyces* sp. KY40-1was cultivated on M2-agar plates at 28 °C for 7 days. From the plate, a piece of agar was excised and added to a 250 mL Erlenmeyer flask containing 100 mL of SG medium and was grown at 28 °C (250 rpm) for 48 hours. The mycelium was pelleted and *Streptomyces* sp. KY40-1 genomic DNA was isolated using MOBIO Ultraclean microbial DNA isolation kit (bioMérieux, France). The genome

was sequenced by Advanced Genetic Technologies Center (Lexington, KY, USA; Illumina Miseq platform) and the processed reads were assembled using Newbler (GS *De Novo* Assembler v 2.8, Roche 454 Life Sciences). Sequence gaps were filled using Universal GenomeWalker 2.0 kit (Clontech/Takara Bio, Mountain View, CA) and subsequent sequencing by ACGT, Inc (Wheeling, IL, USA). Geneious v.6.1.5 (Biomatters)[66] was used for pairwise alignment, assembly of contigs, and gene annotation. Assembly was achieved through a series of de novo assembly of contigs until a single, final contig resulted. Open reading frames were assigned with the aid of protein prediction features in Geneious and the location of ribosomal binding sites. The open reading frames were translated into amino acid sequences, and using the bioinformatic tool BLASTp, putative gene functions were assigned. The sequence of the saquayamycin (*sqn*) cluster has been deposited in GenBank (accession: KX356541).

Preparation of *Streptomyces fradiae* Genomic DNA

From a glycerol stock, *Streptomyces fradiae* was cultivated on M2-agar plates at 28 °C for 7 days. From the plate, a piece of agar was excised and added to a 250 mL Erlenmeyer flask containing 100 mL of SG medium and was grown at 28 °C (250 rpm) for 48 hours. The mycelium was pelleted and *Streptomyces fradiae* genomic DNA was isolated using MOBIO Ultraclean microbial DNA isolation kit (bioMérieux, France).

2.3.2 Cloning of SqnS1, SqnS4-8, and UrdR

sqnS1, *sqnS4-S8* were PCR amplified from *Streptomyces* sp. KY40-1. genomic DNA and *urdZ3* from *Streptomyces fradiae* genomic DNA using primers listed in table 2 and Q5 Hot Start High-Fidelity DNA Polymerase (New England BioLabs). Thermal cycle conditions consisted of 1 denaturation cycle at 98 °C for 30 s; 30 amplification cycles of incubation at 98 °C for 10 s and 72 °C for 30 s/500 bp and then 1 extension cycle at 72 °C for 2 min and finally a hold at 4 °C after the cycles completed. PCR products were purified using GeneJET Gel Extraction and DNA Cleanup Micro Kit (Thermo Scientific). Each gene and vector were individually dually digested with restriction enzymes for 1 hour at 37 °C and dephosphorylated by Shrimp Alkaline Phosphatase (New England BioLabs) for 30 minutes at 37 °C. All but SqnS1, S4, and UrdZ3 were cloned into and pET28a(+) vector with *NdeI* and *HindIII*. SqnS1 was cloned into and pET28a(+) vector using *Nde*I and *Xho*I. SqnS4 was cloned into pET28a(+) vector with *EcoR*I and *Hind* III and into pET22b(+) with *Eco*RI and *Xho*I. UrdZ3 was cloned into pET28a(+) vector with *Eco*RI and *Hind* III and subcloned into pET22b(+) in the same sites. Each gene fragment and vector was purified using Thermo Scientific GeneJET Gel Extraction and DNA Cleanup Micro Kit. Concentrations of each gene insert and linearized vector were determined using a Thermo Scientific NanoDrop 2000. Gene inserts and corresponding linearized vector were ligated in an insert:vector ratio of 5:1 using T4 DNA ligase (New England BioLabs) and incubated overnight at 16 °C. Each ligation reaction was then transformed into NEB 5-α chemically competent cells by heat shocking the cells combined with 1 μL from the ligation reaction for 30 seconds at 42 °C. The cells were incubated at 37 °C for 1 hour in 200 μL LB (Lysogeny Broth) to

allow the cells to recover from heat shock, and then plated on LB agar plates supplemented with appropriate antibiotic for resistance selection. For pET28a(+) constructs, 50 µg/mL kanamycin was used, and for pET28a(+) constructs, 100 µg/mL ampicillin was used. Each plate was incubated overnight at 37 °C and several of the resulting colonies were selected and grown overnight in a 10 mL culture tube with 3 mL LB liquid media supplemented with appropriate antibiotic. Plasmid was recovered from each of the 3 mL overnight cultures using Thermo Scientific GeneJET Plasmid Miniprep Kit. Each gene was PCR amplified from the plasmid to confirm the insert. The plasmids were sequenced (to ensure fidelity) by ACGT, Inc (Wheeling, IL, USA) using their stock T7 promoter and T7 terminator sequencing primers and high GC content DNA sequencing protocol.

Table 2. Primers used to PCR amplify *sqnS1, S4-S8, urdZ3*

	Primer Sequence (5'- 3')	Restriction Enzyme
sqnS1	F: GGGAAA**CATATG**CGCAGTGACGCGCTGGA R: GGGAAA**CTCGAG**TCACGAGGTGCCTGCCTTGTCG	*Nde*I *Xho*I
sqnS4	F: TTTGGG**GAATTC**GTGCACCTGATCGCCCGCACTG R: GGGTTT**AAGCTT**TCATGGTCGTACCCCTTCCTCGGC	*Eco*RI *Hind*III
sqnS4[1]	F: AAAGGG**GAATTC**GTGCACCTGATCGCCCGCA R: GGGAAA**CTCGAG**TGGTCGTACCCCTTCCTCGGC	*Eco*RI *Xho*I

sqnS5	F: GGGAAA**CATATG**AGCGACCGCAAGGACCGGATCCTG R: AAAGGG**AAGCTT**TCAGCCGCGCGCGGCCAC	*Nde*I *Hind*III
sqnS6	F: GGGAAA**CATATG**GACATTGTGGGAAACGGTTTCCTGGC R: AAAGGG**AAGCTT**TCAGGCGTGCGCGGACGTG	*Nde*I *Hind*III
sqnS7	F: GGGAAA**CATATG**CCGTCTTCGCTCCTTCAAGC R: AAAGGG**AAGCTT**TCATCGCGCGCCCGC	*Nde*I *Hind*III
sqnS8	F: GGGAAA**CATATG**ATCACGCCGGTAC R: AAAGGG**AAGCTT**TCATGCGGACAGACGG	*Nde*I *Hind*III
urdZ3	F: GGGAAA**GAATTC**ATGGTCGTCCTGGGAAGCGG R: TTTCCC**AAGCTT**TCACTTCCTTTCGTCAACCGGCGC	*Eco*RI *Hind*III
urdZ3[1]	F: GGGAAA**GAATTC**ATGGTCGTCCTGGGAAGCGG R: TTTCCC**AAGCTT**CTTCCTTTCGTCAACCGGCGC	*Eco*RI *Hind*III

[1] for cloning into pET22b(+), restriction sites in bold

2.3.3 Expression and Purification of Enzymes

sqnS1, S4-S8, urdZ3, rfba, rmlB, mt fdx, mt fdR genes were expressed using pET expression constructs in *E. coli* BL21 (DE3) cells with a His6 tag. Constructs for Mt Fdx

and Mt FdR were gifted from Dr. Steven Van Lanen, University of Kentucky. From LB plates with appropriate antibiotic (50 µg/mL kanamycin for pET28a(+)/100 µg/mL ampicillin for pET22b(+) single colonies were transferred to a 10 mL culture tube with 3 mL LB liquid supplemented with appropriate antibiotic and grown at 37 °C and 250 rpm for 8 h. Subsequently, 2 L borosilicate glass baffled flasks with 500 mL terrific broth (TB) liquid supplemented with 50 µg/mL kanamycin/100 µg/mL ampicillin was inoculated with 1 mL of the culture and was grown at 18 °C until OD_{600} reached to 0.5. Gene expression was induced with isopropyl-*β*-D-1-thiogalactopyranoside (IPTG, 0.2 mM final concentration) and the culture was incubated at 18 °C for an additional 16 h. The cell pellets were collected by centrifugation (4000 × g, 15 min) and resuspended in lysis buffer (50 mM Tris buffer pH 8.0 with 25 mg lysozyme/ 50 mL lysis buffer, 10 % glycerol, and 300 mM NaCl) in a 1 g/ 5 mL pellet weight:lysis buffer volume ratio. The cells were sonicated and the crude soluble enzyme fractions were collected through centrifugation (16000 × g, 1h). The supernatant was then passed through a 0.45 µm syringe driven multigrade glass fiber filter (Millipore) and incubated with Talon metal affinity resin (Clontech) in 50 mL falcon tubes on a platform shaker at 4 °C for 30 minutes. The protein bound resin was then centrifuged (4000 × g, 5 min) and washed with equilibration buffer (50 mM Tris buffer pH 8.0, 10 % glycerol, and 300 mM NaCl). The resin was then transferred to a 2 mL gravity-flow column and washed with wash buffer (50 mM Tris buffer pH 8.0, 10 % glycerol, and 300 mM NaCl, 30 mM imidazole). The enzymes were then eluted with elution buffer (50 mM Tris buffer pH 8.0, 10 % glycerol, and 300 mM NaCl, 250 mM imidazole). The purified proteins were concentrated using an Amicon Ultra centrifugal filter (Millipore Corp.) and stored as

20% glycerol stocks at -20 °C. The concentration of protein was determined by using a Thermo Scientific NanoDrop 2000.

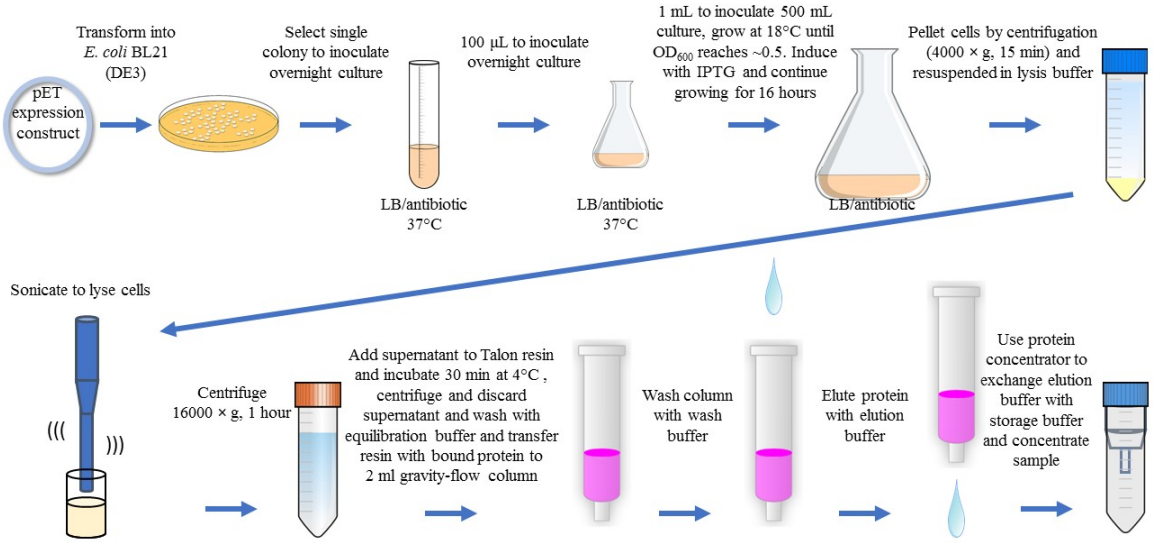


Figure 29. Diagram of SqnS1, S4-S8, and UrdZ3 protein expression and purification.

2.3.4 Enzymatic Reactions

2.3.4.1 Two-Stage One-Pot Enzymatic Synthesis of TDP-4-keto-3,6-dideoxy-L-idosamine in Aerobic Conditions

Overall enzymatic synthesis of TDP-4-keto-3,6-dideoxy-L-idosamine (132) was divided into two stages. For the first stage, a typical assay mixture containing glucosamine 1-phosphate (127) (5.1 mM), TMP (136) (5.0 mM), ATP (5 µM), acetyl phosphate (20 mM), NAD^+ (5 µM) (ACK (4 µM), TMK (6 µM), RfbA (5 µM), RmlB (10

μM), phosphate buffer (50 mM, pH 7.5) and MgCl2 (20 mM) was incubated at 37◦C for up to 8 hours while monitoring the reaction hourly by HPLC at 254 nm λ using a Waters HPLC system equipped with a photodiode array detector. For each sample injection (50 μL), the enzymes were removed by heating the solution to 100 ◦C for 5 min, followed by centrifugation (13,000 x G, 5 min), and filtration (Amicon, MW cut-off 3000Da). The enzyme-free solution was analyzed using a Dionex CarboPack PA1 (4 x 250 mm) column. A linear gradient of 5-20% 500 mM ammonium acetate, pH 7.0 (solvent B) in water (solvent A) over 15 min, 20 to 60% B in A over 20 min, 60-100% B in A over 2 min, 100 % B for 3 min, 100-5% B in A over 5 min, and 5% B in A for 15 minutes was employed to separate the assay products. In the second stage, Rfba, RmlB, ACK, and TMK were removed by heating, centrifugation, and filtration. PMP (250 μM), NADPH (4 Mm), SqnS5 (30 μM), SqnS1 (30 μM), Mt Fdx (30 μM) and Mt FprA (30 μM) were then added to the assay mixture and incubation at 28 °C was continued monitoring the reaction hourly up to 5 hours, and then overnight by HPLC. The peaks were collected and analyzed by HR-MS.

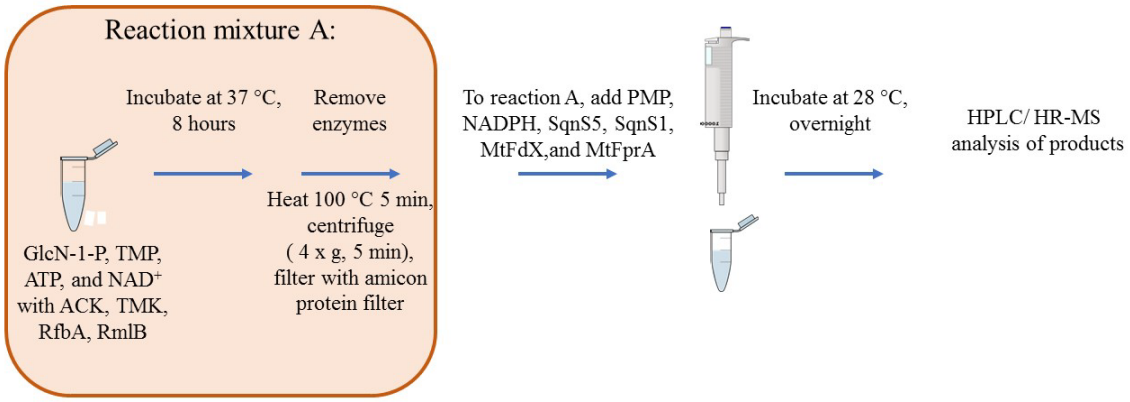


Figure 30. Diagram of two-stage one-pot enzymatic synthesis of TDP-4-keto-3,6-dideoxy-L-idosamine, aerobic conditions

2.3.4.2 Three-Stage One-Pot Enzymatic Synthesis of TDP-4-keto-3,6-dideoxy-L-idosamine in Partial Anaerobic Conditions

The reaction from section 2.3.4.1 was also carried out in partial anaerobic conditions at room temperature. All reagents, buffers, and enzymes were degassed immediately prior to transferring to the glovebox. SqnS5 was reconstituted in the glovebox using dithiothreitol (DTT) to a final concentration of 10 mM and incubation for 20 min. Next, 5 molar equivalents (SqnS5 is 1 equivalent) of PMP (250 μM), $Fe(NH_4)_2(SO_4)_2$, and Na_2S were added and allowed to incubate for 15 min. Addition and 15 min incubation of $Fe(NH_4)_2(SO_4)_2$, and Na_2S were repeated 3 times for a total of 4 cycles. Excess reagents were removed using an Econ-Pac 10 DG column (Bio-Rad) according to the manufacturer's directions; the sample was diluted to 3 mL using protein storage solution (50 mM Tris buffer pH 8.0, 20 % glycerol, and 300 mM NaCl) and was added the column preequilibrated with protein storage solution. The column was washed

with 4 mL of protein storage solution, collecting the first 3 mL that eluted. The protein was then concentrated to about 10 µL using an Amicon protein filter (MW cutoff of 3 Da) and microcentrifuge in the glovebox. Finally, 20 equivalents of 0.6 mM sodium dithionite were added and incubated for 40 min. To begin the 3-deoxygenation (SqnS5) reaction, reconstituted SqnS5 was added to 100 µL of the assay mixture from stage 1 (Reaction A) with enzymes (Rfba, RmlB, ACK, TMK) removed and incubated in the glovebox for 3 hours. The SqnS5 reaction mixture was removed from the glovebox and SqnS1 and NADPH (4 mM) were added and incubated at 37 ◦C for three hours followed by HPLC/HR-MS analysis.

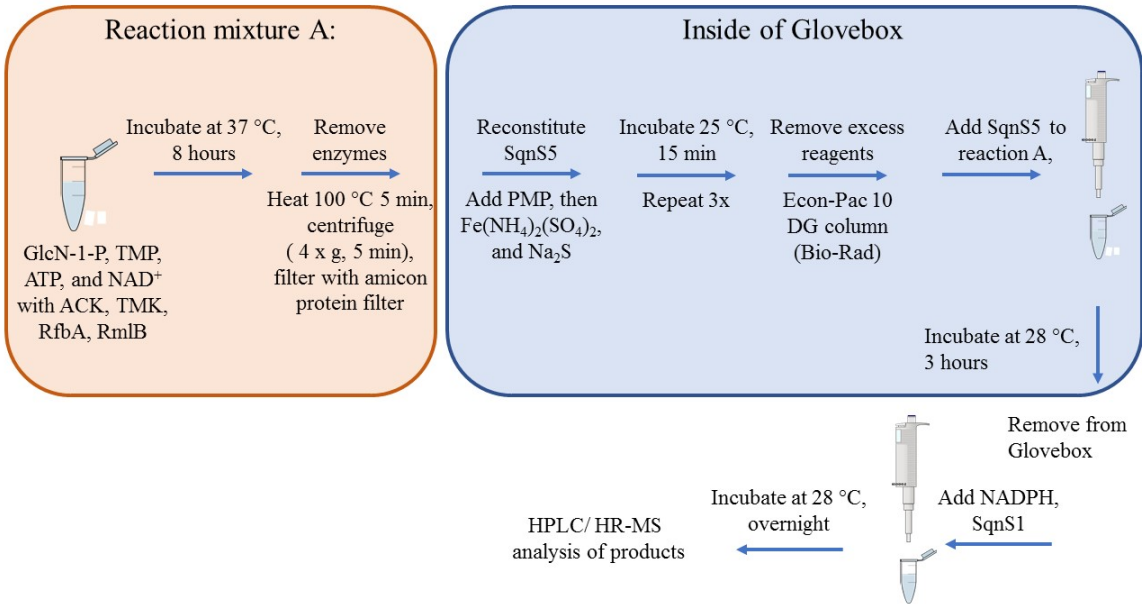


Figure 31. Diagram of three-stage one-pot enzymatic synthesis of TDP-4-keto-3,6-dideoxy-L-idosamine, partial anaerobic conditions

2.4 Conclusions

DNA analysis of the *Streptomyces* sp. KY40-1 genome revealed a 44.7 kb gene cluster with 40 open reading frames involved in production of SQN A, B, G-K. Six deoxysugar genes were putatively assigned to L-rednose (109 A) biosynthesis including a nucleotidyltransferase (SqnS2), a 4,6-dehydratase (SqnS3), a 3,4-dehydratase (SqnS5), a 5-epimerase (SqnS1), a 4-ketoreductase (SqnS4), and a dehydrogenase (SqnQ). A biosynthetic pathway was proposed with glucosamine 6-phosphate (128) as the starter unit. The hypothetical pathway proceeds with activation of the sugar by thymidyltransfer followed by C-6 deoxygenation, C-3 deoxygenation, 5-epimerization to convert the sugar from D to L conformation, and a 4-ketoreduction to yield TDP-3,6-dideoxy-L-idosamine (132). It is proposed that TDP-3,6-dideoxy-L-idosamine (132) is transferred to the acceptor and a final oxidation generates L-rednose (109 A). Four of the SQN deoxysugar biosynthetic enzymes (SqnS1, S5, S7-S8) were successfully cloned from *S.* sp. KY40-1 genomic DNA and were heterologously expressed in *E. coli* and purified resulting in soluble protein. SqnS4 and SqnS6 remained insoluble. Enzymatic synthesis of TDP-4-keto-3,6-dideoxy-L-idosamine was attempted using a mix-and-match method with SqnS2 homolog RfbA and SqnS3 homolog RmlB, SqnS5, Mt FdX, Mt FprA and SqnS1 in a two stage one-pot reaction aerobically and a three-stage one-pot reaction with reducing agents in place of Mt FdX/Mt FprA for SqnS5 catalysis in anaerobic conditions. Only the first step, thymidyltransfer was successful and generated both TDP-glucosamine (128) and TDP-*N*-acetylglucosamine. The future directions should include cloning and expression of SqnS3 in *E. coli* and since codon optimization for expression in *E. coli* was unsuccessful, subcloning SqnS4 into PXY200 and expression in *S. lividans TK24*. With both SqnS3 and S4 in hand, enzymatic synthesis of TDP-3,6-dideoxy-L-idosamine (132)

can be attempted. Together this study revealed valuable deoxysugar biosynthetic tools which can be used to modify artificial pathways, and generated TDP-glucosamine (128) and TDP-*N*-acetylglucosamine building blocks to engineer new glycosylated natural products.

3. INVESTIGATION OF THE MECHANISM OF ACTION FOR MITHRAMYCIN

Weidenbach, S., Hou, C., Chen, J.M., Tsodikov, O.V., Rohr, J. Dimerization and DNA recognition rules of mithramycin and its analogues. J. Inorg. Biochem. 2016, 156, 40-47.

3.1 Introduction

After its discovery in 1960, an aureolic acid natural product mithramycin (MTM (138); figure 33) was initially used in the treatment of Paget's disease[67-69] and disseminated testicular carcinoma.[69] Due to toxicity of MTM (138), its more recent clinical use has been limited to treatment of hypercalcemia associated with cancers. Interest in MTM (138) was renewed after it was identified by high-throughput screening of 50,000 compounds and then validated as a potent antagonist of the oncogenic transcription factor EWS-FLI1 (Ewing Sarcoma- Friend Leukemia Integration 1).[70] EWS-FLI1 is a principal driver of Ewing sarcoma,[71] a rare devastating bone cancer affecting primarily children and young adults. MTM (138) displayed potent cytotoxic and anti-tumor activities in xenograft models of Ewing sarcoma in mice,[70] owing to its action

as an EWS-FLI1 inhibitor. On the basis of these findings, MTM (138) entered clinical trials as a novel Ewing sarcoma chemotherapeutic (National Cancer Institute, clinical trial NCT01610570) however the study was terminated due to off-target toxicity.

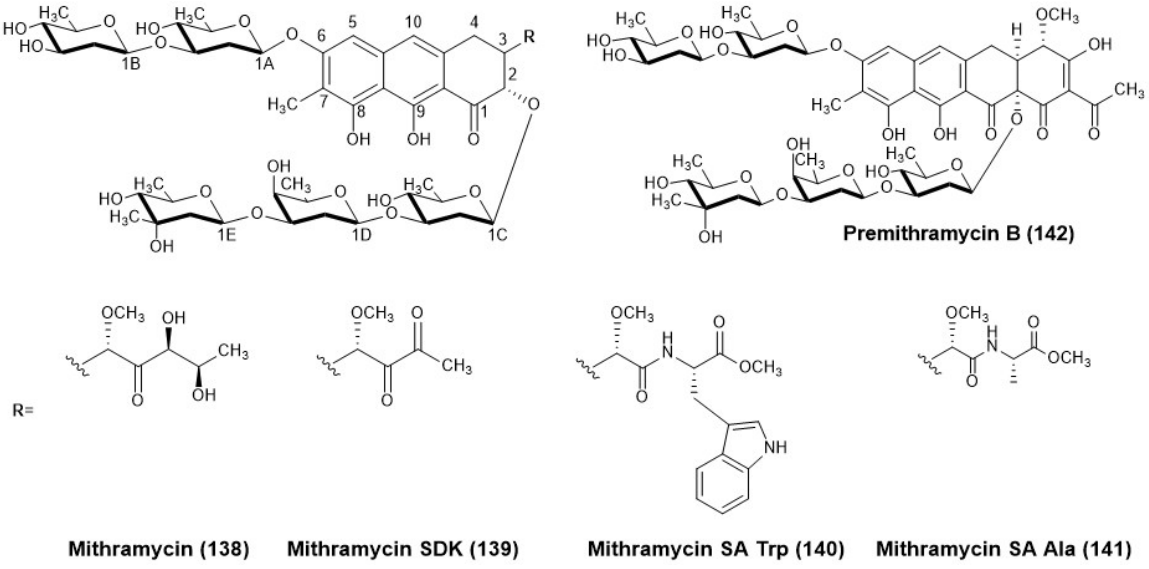


Figure 32. Chemical structures of MTM, its analogues MTM SDK, MTM SA-Trp, MTM SA-Ala and PreMTM B.

Aureolic acid natural products bind in the minor groove of GC-rich DNA, thereby disrupting transcription of protooncogenes.[72-80] MTM (138) binding was shown to lead to structural changes in DNA that were consistent with the widening of the minor groove.[81, 82] MTM (138) and other known aureolic acid compounds, except chromocylomycin, consist of a tricyclic polyketide-derived aglycone with two oligosaccharide chains at C-2 and C-6 positions of the aglycone and two aliphatic side chains, branching off at C-3 and C-7 (figure 33). Combinatorial biosynthetic and semisynthetic approaches yielded several

MTM analogues including MTM SDK (139) (for Short side chain, DiKeto)[83], MTM SK (for Short side chain, Keto), MTM SA (for Short side chain, Acid)[84], and MTM SA-amino acid residue derivatives (figure 33).[85] Genetic engineering of the MTM-producing bacterium *Streptomyces argillaceus* to inactivate the gene *mtmW* encoding the enzyme catalyzing the final step of the MTM biosynthetic pathway, the ketoreductase MtmW, yielded derivatives MTM SA, MTM SK[84], and MTM SDK (139).[83] These MTM analogues all contained shortened 3-side chains, which had been shown by previous NMR studies of MTM (138)[86, 87] to interact with the sugar-phosphate backbone of DNA. Both MTM SK[84] and MTM SDK (139)[83] displayed higher potency against several human cancer cell lines and lower toxicity than those of MTM (138), demonstrating that 3-side chain modifications can modulate activity of MTM analogues. MTM SA had a decreased cytotoxicity[84] but the carboxylic moiety provided a chemical handle that could be readily and selectively modified to further diversify this analogue towards derivatives with increased activity. The amino acid residue derivatives of MTM SA contained glycine, valine, cysteine, and alanine residues, of which MTM SA-Ala (141) was the most promising, although still not as potent as MTM SK or MTM SDK (139) against a human non-small cell lung cancer cell line A549.[85] Consequently, additional MTM SA amino acid derivatives have been generated including the most promising MTM SA-Trp (140).

MTM (138) binds DNA as a dimer, in which the O-1 carbonyl and the O-9 enolate anion of the aglycone of each monomer are coordinated via a divalent metal ion (Me^{2+}).[88, 89] DNase I footprinting studies of MTM (138) bound to DNA showed that MTM (138) protects three base pair sites, with a preference to sites containing GG dinucleotide steps.[90] Pioneering solution NMR studies revealed that the MTM_2-Me^{2+}

binds a palindromic DNA oligomer containing the sequence GGCC, where the 2-fold rotational symmetry of the MTM dimer coincides with the symmetry of the palindromic DNA.[86, 87] These structures indicated that the C-8 hydroxyl groups on the chromophores of each MTM monomer form hydrogen bonds with the NH_2 group of the central guanine bases on the complementary strands, while the flexible saccharide chains of the two MTM monomers are extended along the minor groove of the DNA in a symmetric antiparallel fashion. A crystal structure of chromomycin A_3 (CHR), a related natural product containing the same aglycone core as MTM (138), in complex with a GGCC-containing DNA was consistent with the solution structure of MTM-DNA complex, showing additionally the octahedral coordination of the divalent metal ion and the widened minor groove of the DNA.[91] These structural studies, in agreement with DNA binding studies in solution, demonstrated that the DNA binding sites for MTM (138) and CHR were four base pairs long.

The rules for DNA sequence recognition of MTM are still not fully understood. For example, whether MTM (138) recognizes a specific DNA sequence (direct readout), DNA conformation (indirect readout) or both, is not clear. [92] MTM (138) preferentially binds G/C-rich sequences.[90, 93-95] Binding to a GC sequence is strengthened by flanking TA steps, and a GC site is preferred over a CG site.[90, 94, 96] DNase I footprinting using DNA fragments containing all 64 symmetrical hexanucleotide sequences demonstrated that MTM (138) binds best to AGCGCT and GGGCCC.[95] However, this study also showed that MTM (138) also bound well at non-GC sites, such as ACCGGT. While these studies systematically investigated sequences at which MTM (138) is footprintable, other sequences that form thermodynamically stable, but relatively short-lived complexes with

MTM (138) may have eluded detection, leaving open the question of the minimum DNA recognition sequence of MTM (138).

In this study, we performed equilibrium titrations at a physiologically relevant salt concentration, to investigate the propensity of MTM (138) and its novel analogues to dimerize and recognize DNA of different sequences and conformations.

3.2 Results and Discussion

3.2.1 Dimerization of MTM and its Analogues

Whether MTM (138) is a monomer or a dimer when free in solution at physiological salt concentrations is an open question, since previous dimerization studies of MTM (138) dimerization, although rigorous, have been performed at non-physiologically relevant salt concentrations.[97-99] We investigated the effect of Mg^{2+} concentration on the dimer stability of MTM (138) and its analogues at 100 mM NaCl, i.e. at a physiologically relevant concentration of monovalent cations, by monitoring the change of intrinsic fluorescence of MTM (138) (figure 34) and its analogues (figure 35), including a novel analogue MTM SA-Trp (140), upon their dimerization. The fluorescence of MTM (138) and all analogues, except PreMTM B (142) (figure 35), was quenched upon dimerization with increasing concentration of Mg^{2+}. The chromophore of PreMTM B (142) is tetracyclic; therefore, the observed increase of its fluorescence may arise from a different orientation of the two chromophores in a PreMTM B dimer from that in MTM (138). These data were in good agreement with the cooperative 2MTM:Me^{2+} binding model, which yielded best-fit values of the equilibrium dimerization constants K_d (table 3). MTM (138) and MTM SA-Trp (140) exhibited a

similar propensity to dimerize, stronger than that of MTM SDK (139) and PreMTM B (142). PreMTM B (142) was by far the most defective in dimer formation, with a 40-fold higher K_d than that of MTM (138). For MTM (138) and MTM SDK (139) at 5 mM Mg^{2+} these K_d values correspond to the values of the apparent dimerization equilibrium constant (for the process $2M \rightleftharpoons M_2$) of 0.7 mM and 2.8 mM, respectively. These values indicate that at 5 mM or higher concentrations, MTM (138) and MTM SDK (139) are dimeric at these conditions at 5 mM Mg^{2+} and 100 mM NaCl, i.e. in the physiological range of salt concentrations. In contrast, PreMTM B (142) is mostly dimeric only at ~30 mM and higher concentrations. Indeed, the fluorescence intensity reaches a plateau by 5 mM Mg^{2+} for MTM (138) and MTM SDK (139), but not for PreMTM B (142) (figure 35). In addition, we also tested the effect of Fe^{2+} on MTM dimerization. In agreement with a previous report,[98] Fe^{2+} is a more potent MTM dimerizer, with K_d at least 100-fold smaller than in Mg^{2+} (table 3). MTM SA-Ala (141) was prone to aggregation in the absence of divalent metal ions yielding an unstable signal. Nevertheless, MTM SA-Ala (141) formed dimers as well, as described in section 3.2.

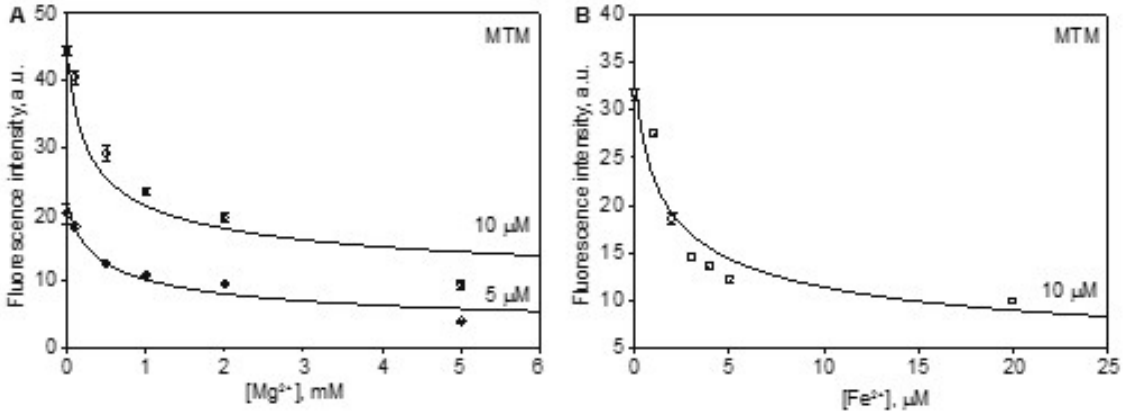


Figure 33. Divalent metal ion-dependent dimerization of MTM observed through quenching of intrinsic MTM fluorescence. A. Mg^{2+}-dependent dimerization of MTM. B. Fe^{2+}-dependent dimerization of MTM. The curves correspond to the cooperative dimerization model given by eq. (1) with the best-fit parameter values in Table 1. The concentrations of MTM in each titration are indicated.

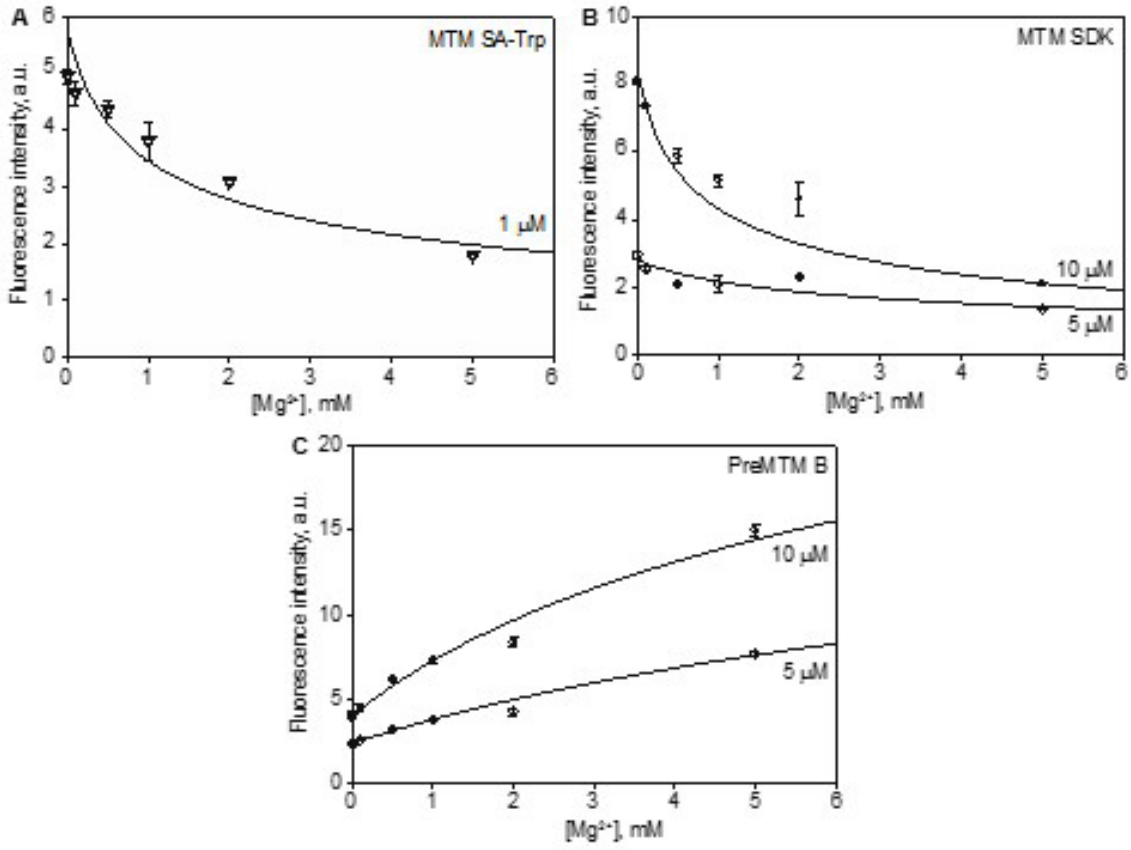


Figure 34. Mg^{2+}-dependent dimerization of MTM analogues: MTM SA-Trp (panel A), MTM SDK (panel B) and Pre MTM B (panel C) at specified concentrations.

Table 3 Equilibrium dimerization constants for MTM and its analogues.

MTM analogue	Metal ion	K_d, $(\mu M)^2$
MTM	Mg^{2+}	$(3.6 \pm 1.7)\times10^3$
MTM	Fe^{2+}	< 40[a]
MTM SA-Trp	Mg^{2+}	$(1.8 \pm 0.9)\times10^3$
MTM SDK	Mg^{2+}	$(13.9 \pm 4.0)\times10^3$
PreMTM B	Mg^{2+}	$(120 \pm 70)\times10^3$

[a]Only an upper bound could be determined from the data.

3.2.2 DNA Binding by MTM and its Analogues

In order to define DNA sequence and conformational specificity determinants for binding MTM (138), we chose 13 palindromic 8-mer DNA sequences (table 4). Crystal structures of respective double-stranded DNA oligonucleotides covering the entire sequence or a relevant part of it were previously determined. These sequences comprised B- and A-form DNA and included a sequence containing a well-established MTM binding motif GGCC that was used in the solution NMR studies of MTM-DNA complex.[83,84] Among B-form DNA, we included G/C-rich oligomers as well as oligomers with a widened minor groove. The length of 8 base pairs was chosen to minimize the potential of binding of multiple MTM dimers to a single DNA oligomer; shorter DNA would have a compromised duplex stability. Binding of MTM (138) and its analogues to these DNA oligonucleotides was studied at salt concentrations close to physiological, where the dominant species are dimers, as established by the studies described in the previous section. Furthermore, at these salt concentrations and assay temperature (21 °C), the DNA oligomers are predominantly in the double-stranded form. The calculated DNA melting temperatures at these conditions at 6 mM of each oligomer, or 3 mM of double-stranded oligomer, which is approximately the lowest DNA concentration used in the titrations, range from 28 °C for CTCTAGAG to 59 °C for CGGCGCCG. [100] Even though the melting temperatures calculated by this method were shown to be accurate for 8-mer and larger oligomers,[100] we carried out a DNA melting experiment with oligomer GGTTAACC (at 3 mM in terms of duplexes) with one of the lowest calculated melting temperatures of 32.2 °C. In this experiment, we monitored the hyperchromic effect by measuring absorbance at 260 nm as a function of gradually increasing temperature. The

observed melting curve (figure 36) yielded the melting temperature of ~32 °C, in excellent agreement with the calculation. These data also demonstrate that at the assay temperature of 21 °C this DNA oligomer is in the double-stranded state.

Table 4. Structural and MTM analogue binding characteristics of the DNA oligomers.

DNA	Analogue	DNA form	PDB ID[a]	Average Zp[b] (Å)	Minor groove width (Å)	Major groove width (Å)	K_d (μM)
CGGCGCCG	MTM	B	1CGC	-0.1	12.8	16.5	< 1[c]
CCCCGGGG	MTM	A	187D	2.5	15.4	15.1	< 1
GGGATCCC	MTM	A	3ANA	2.4	16.1	16.0	< 1
GGGTACCC	MTM	A	1VT9	2.5	16.4	15.3	< 1
AGGTACCT	MTM	B	1SS7	0.03	12.8	16.1	< 1
GTAGCTAC	MTM	B	141D	0.0	13.2	21.2	< 1
GATCGATC	MTM	B	1D23	-0.1	12.2	17.3	< 1
GAGGCCTC	MTM	B	1BD1	-0.6	11.7	17.3	2.0 ± 0.5
GTACGTAC	MTM	A	28DN	2.4	15.2	16.4	7.8 ± 0.3
GTGTACAC	MTM	A	1D78	1.9	17.0	13.4	13 ± 1
GGAATTCC	MTM	B	1SGS	-0.7	12.2	18.0	59 ± 2
GTCTAGAC	MTM	A	1D82	2.1	15.5	16.2	25 ± 1
CTCTAGAG	MTM	A	1D93	2.1	15.4	15.6	82 ± 9
CGGCGCCG	MTM SA-Trp	see above					1.2 ± 0.2
GAGGCCTC	MTM SA-Trp						2.1 ± 0.3
CCCCGGGG	MTM SA-Trp						2.1 ± 0.6
GGGTACCC	MTM SA-Trp						3.7 ± 1.5
GGGATCCC	MTM SA-Trp						4.3 ± 1.5
GAGGCCTC	MTM SDK						< 1
GAGGCCTC	MTM SA-Ala						19.6 ± 3.8
GAGGCCTC	PreMTM B						> 300

[a] The crystal structures were previously published as follows: 1CGC,[101] 187D,[102] 3ANA,[103] 1VT9,[104] 1SS7,[105] 141D,[106] 1D23,[107] 1BD1,[108] 28DN,[109] 1D78,[109] 1SGS,[110] 1D82,[111] 1D93.[112]

[b] Z_p, a difference between coordinates of projections of the two phosphorus atoms of a base pair onto the DNA helical axis, was averaged ether among all base pairs or for those that constituted a binding site (a stretch of two or more (G/C) base pairs), where such stretch was present.

[c] The upper bound on K_d of ~1 mM was defined by the detection limit of the assay.

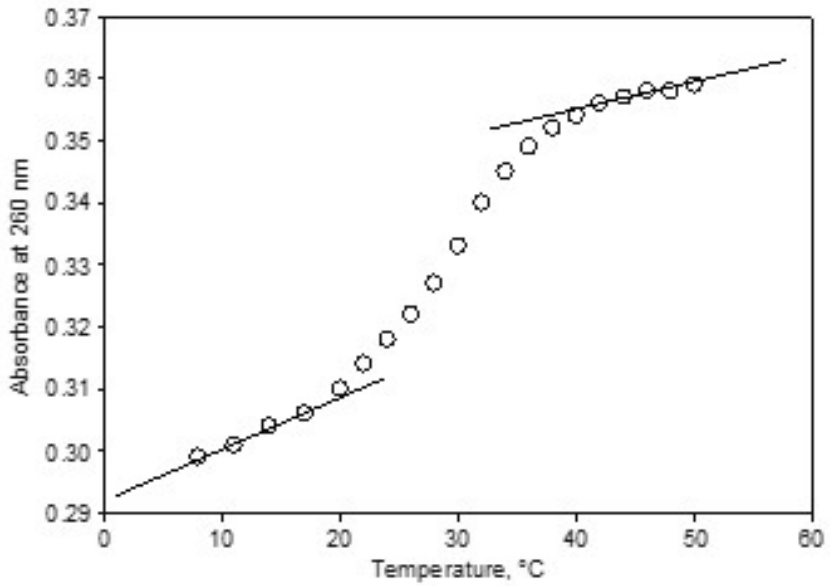


Figure 35. The DNA melting curve for oligomer GGTTAACC (at 6 mM in single strands) in the DNA binding buffer. The lines at low and high temperatures show the approximate baselines corresponding to the double- and the single-stranded states, respectively.

Strong MTM-DNA binding resulted in an ~8-fold increase of fluorescence intensity at the fluorescence maximum (figure 37), apparently due to unquenching of the chromophore rings. Such unquenching may occur upon reorientation of the MTM molecules in the dimer upon DNA binding. Fluorescence enhancement upon MTM binding to DNA was observed previously. [113] These and the other DNA titration data fit well to a 1:1 binding isotherm. The best-fit K_d values for all titrations and the DNA structural parameters are given in Table 4. Affinities stronger than 1 mM could not be measured and were indicated as bounds in Table 4, because of the low fluorescence signal at concentrations of MTM (138) and analogues required for such measurements.

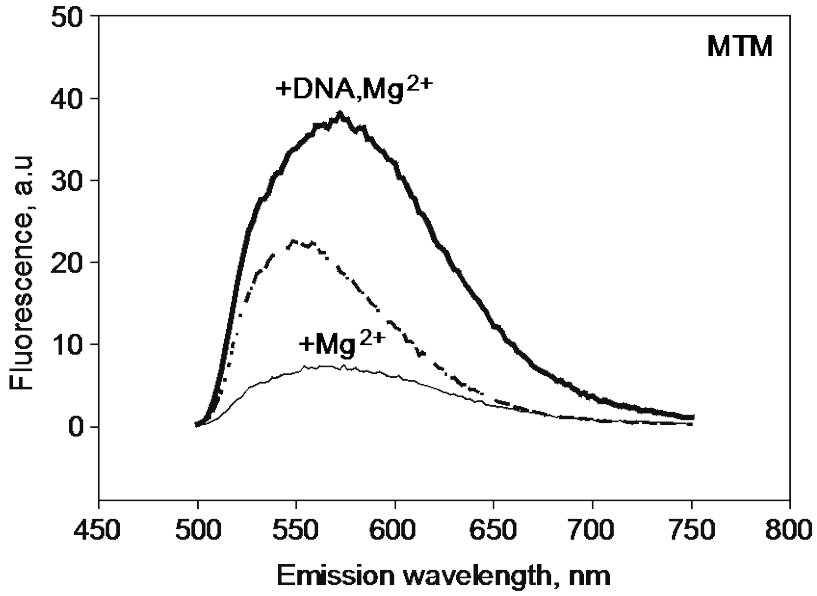


Figure 36. Fluorescence emission spectra of MTM (10 μM) in the absence of divalent metal and DNA (the dashed line), with 5 mM $MgCl_2$ and no DNA (the thin solid line), and with 5 mM $MgCl_2$ and 10 μM GAGGCCTC DNA (the thick solid line). The excitation wavelength was 470 nm.

Binding of MTM (138) to DNA was tested most extensively, with all 13 DNA oligomers; binding of the novel analogue, MTM SA-Trp (140) was tested with 5 oligomers and binding of MTM SDK (139) and PreMTM B (142) was tested with one oligomer, GAGGCCTC, which was tested with all the compounds. Representative titration data for MTM (138) with high-affinity and low-affinity oligomers are shown in figures 38A and 38B, respectively. The strongest MTM-DNA binding affinity was observed for 8 DNA oligomers, with K_d values of 2 mM or smaller. All of these sequences contain at least one GG, GC, CG or CC site (note that a CC sequence is automatically present if a GG sequence is present due to complementarity). Notably, sequence GGAATTCC containing a terminal GG site did not bind MTM (138) well (K_d = 59 mM). On the other hand, sequence AGGTACCT was a high-affinity sequence. These results indicate that a binding site needs to contain at least one flanking base pair on each side. MTM (138) displayed low affinity for all sequences that did not contain an internal GG, GC or CG site. Therefore, 2-base pair sequences GG (CC), GC and CG that are flanked by at least one base pair on each side are preferred sequences for binding MTM (138), whereas the other 2-base pair sequences are non-preferred. In other words, the minimum DNA recognition sequence for MTM (138)) is a 4-base pair sequence X(G/C)(G/C)X, where either G or C in any order, and X is any base.

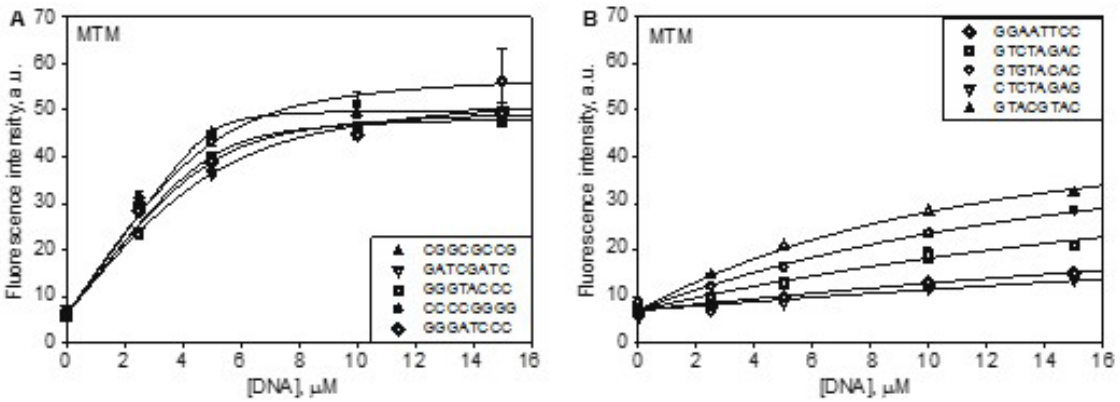


Figure 37. Binding of MTM (10 μM) to double-stranded DNA oligomers of different sequences. A. Representative titrations of MTM with high-affinity DNA oligomers. A. Representative titrations of MTM with low-affinity DNA oligomers. The curves are the best fit to the 1:1 binding model with the values given in Table 4. The DNA sequences are specified in the inset.

Binding affinity to the high-affinity sequences was modulated by the base pairs flanking the GG/GC/CG site. Sequence GATCGATC bound MTM (138) with at least a 7-fold higher affinity than did GTACGTAC. This effect of flanking sequences is consistent with previous observations.[114] Sequence GTAGCTAC, containing a GC in place of the CG in the same context, was a high affinity sequence, reversing the effect of the flanking DNA, in agreement with a previous study showing a preference of MTM for an isolated GC over an isolated CG site.[114] These results suggest that the conformation or position of an MTM dimer bound to a GC site differ somewhat from that bound to a CG site, resulting in different interactions with flanking DNA. Structural investigation of this adaptability of MTM binding different sequences is in progress.

Next, we examined whether DNA affinity of MTM (138) was additionally defined by or correlated with DNA structural parameters (table 4). Because aureolic acid natural products were shown to bind in the minor groove, we focused on the geometric parameters of the minor and major grooves of the DNA oligomers of interest, as calculated by 3DNA software. [115] Binding of MTM (138) did not appear to be correlated with any of the structural parameters: the conformation of the DNA backbone (B- *vs*. A form), the Z_p values, or the widths of the minor and the major grooves. These observations are consistent with a previous study demonstrating strong binding to two G/C rich sequences, one B- and the other A-form.[116] Nevertheless, because the best MTM binding sequences were somewhat enriched in G/C stretches, which are known to have a widened minor groove,[108, 117] this effect of intrinsic sequence-dependent minor groove widening on MTM binding cannot be further uncoupled from the (G/C)(G/C) sequence recognition, at least with naturally occurring nucleotides.

We next tested whether modifications at position 3 had an effect on DNA binding. DNA binding of the novel analogue, MTM SA-Trp (140), was measured with 5 oligomers that displayed high affinity towards MTM (138) (table 4, figure 39). MTM SA-Trp (140) exhibited K_d values in the range 1-4 μM. Thus, the DNA affinities of MTM SA-Trp (140) were either similar or somewhat weaker than those of MTM. Sequence GAGGCCTC bound with a similar affinity (K_d = 2 μM) to MTM (138) and MTM SA-Trp (140) (table 4, figure 40A). Additionally, binding of the other analogues, MTM SDK (139), MTM SA-Ala ((141) and PreMTM B (142), to a strong binding sequence GAGGCCTC was measured (table 4, figures 40A and 40B). MTM SDK (139) bound the DNA at least 2-fold more tightly than did MTM (138), possibly due to reduced steric

hindrance by the shorter side chain of MTM SDK (139). To check whether MTM SDK (139) binding to DNA was still Mg^{2+}-mediated and was not due to a new binding mode that became possible because of its specific side chain, we performed the same titration in the absence of Mg^{2+} with this analogue (figure 40B). No binding was observed, indicating that the binding mode was unchanged by the side chain modification. Interestingly MTM SA-Ala (141) bound ~10-fold weaker than MTM (138) did, apparently due to the unfavorable environment of the Ala methyl group. PreMTM B (142) did not display any observable binding, indicating that this tetracyclic scaffold cannot adapt the same way to the features of the minor groove.

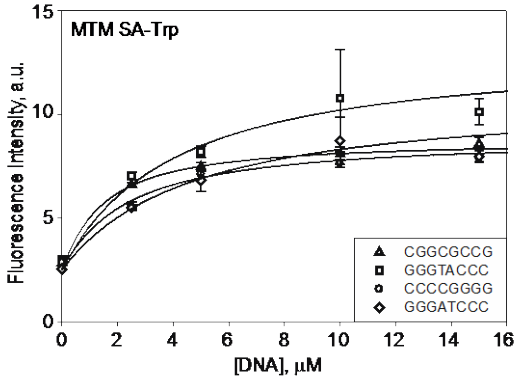


Figure 38. Binding of MTM SA-Trp (1 M) to double-stranded DNA oligomers of different high-affinity sequences, as indicated in the inset.

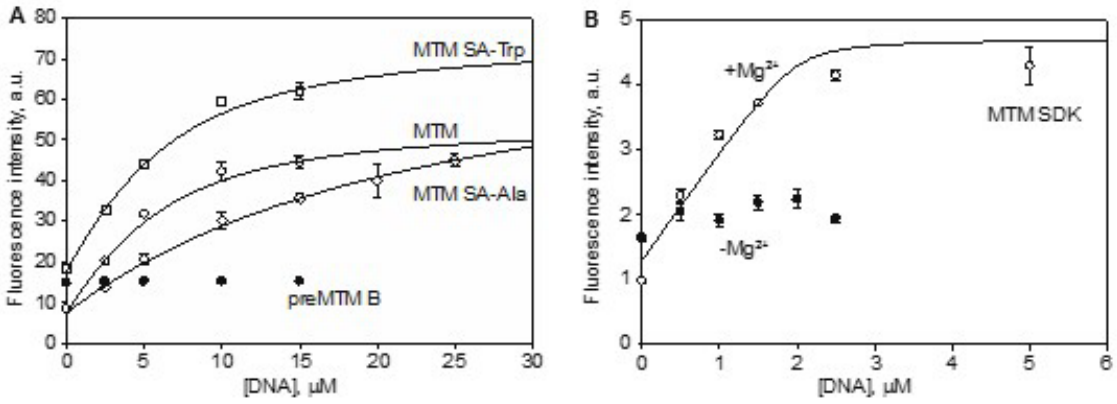


Figure 39. Binding of MTM and its analogues to GAGGCCTC DNA. A. Binding of MTM, MTM SA-Trp, MTM SA-Ala and PreMTM B, each at 10 μM. B. Binding of MTM SDK (4 μM) in the presence (open circles) and in the absence (filled circles) of Mg2+.

3.3 Methods

3.3.1 Dimerization Assays

Concentrations of MTM (138), its analogues and Pre MTM B (142) were determined by measurements based on the dry material mass. The concentrations of MTM (138) and its analogues were consistent with the measurements of absorbance at 400 nm using a Thermo Scientific Biomate 3S spectrophotometer assuming the molar extinction coefficient of 10,000 $M^{-1}cm^{-1}$. The dimerization assays were carried out in 40 mM Hepes pH 7.0, 100 mM NaCl, by titrating $MgCl_2$ into a solution of MTM (138) or analogue at constant concentration at 21 °C. The concentrations of MTM (138), MTM SDK (139), Pre MTM B (142), MTM SA-Trp (140), and MTM SA-Ala (141) were 1 μM, 5 μM or 10 μM, as specified in the text. The dimers were more stable in Fe^{2+} than in

Mg^{2+}; therefore, lower concentrations of the divalent metal ion were used, as specified. Fluorescence of MTM (138) and its analogues, generated for the excitation wavelength of 470 nm and emission wavelength of 540 nm, was measured on a SpectraMax M5 microplate reader. All experiments were carried out in triplicate.

3.3.2 Analysis of the Dimerization Assay Data

MTM dimerization is mediated by one divalent metal ion cooperatively.[99]

$$2M + Me^{2+} \rightleftharpoons M_2 \bullet Me^{2+}, \quad (1)$$

where M designates a monomer of MTM, and Me^{2+} is the divalent metal ion. The equilibrium constant for this process in the direction of dissociation is

$$K_d = [M]^2[Me^{2+}]/[M_2 \bullet Me^{2+}] \quad (2)$$

Conservation of material for mithramycin is given by the following expression:

$$[M]_{tot} = [M] + 2[M_2 \bullet Me^{2+}], \quad (3)$$

where $[M]_{tot}$ is the total concentration of mithramycin in the mixture.

Solving the system of equations (2) and (3) for [M] and $[M_2 \bullet Me^{2+}]$ yields:

$$[M] = \left(\frac{K_d}{4[Me^{2+}]}\right)\left(\sqrt{1+8[M]_{tot}[Me^{2+}]/K_d} - 1\right) \quad (4)$$

$$[M_2 \bullet Me^{2+}] = \frac{1}{2}\left[\left([M]_{tot} - \frac{K_d}{4[Me^{2+}]}\left(\sqrt{1+8[M]_{tot}[Me^{2+}]/K_d} - 1\right)\right)\right] \quad (5)$$

Finally, the observed fluorescence is a sum of contributions from M and $M_2 \bullet Me^{2+}$ species:

$$I = C_1[M] + C_2[M_2 \bullet Me^{2+}] + C_3, \quad (6)$$

where C_1 and C_2 are molar fluorescence yield of M and $M_2 \bullet Me^{2+}$, respectively, and C_3 is the signal in the absence of MTM. Nonlinear regression using SigmaPlot (SysStat) was employed to obtain best-fit values of K_d by fitting eq. (6) to the data for MTM and its analogues.

3.3.3 DNA Binding Assays

DNA oligomers of the palindromic sequences reported in the text were dissolved in 10 mM Na cacodylate buffer pH 6.0 to a final concentration of 2 mM. DNA was self-annealed in the annealing buffer (10 mM Na cacodylate, pH 6.0, 50 mM NaCl) by heating to 95 ºC for 5 min and cooling for 30 min to 4 ºC. The DNA concentrations were then calculated based on the respective extinction coefficients for double-stranded DNA by the IDT online concentration calculation tool.[118] The same tool was used to calculate the DNA melting temperatures.[100]

The DNA binding assays were carried out in 40 mM Hepes pH 7.0, 100 mM NaCl and 5 mM $MgCl_2$ at 21 °C, where the majority of MTM or its analogue is in the dimeric form. DNA was titrated into MTM (138) or its analogues, where all compounds were used at the concentrations specified in the text. Fluorescence was measured as it was in the dimerization assays. All experiments were carried out in triplicate.

The DNA melting experiment for oligomer GGTTAACC (at 6 mM in single strands) was carried out by measuring absorbance at 260 nm at temperatures increasing from 8 °C to 50 °C over 1 hour in the binding buffer. These data were collected on a Shimadzu spectrophotometer TCC-240A UV-1800 equipped with a temperature-controlled cell holder.

3.3.4 Analysis of the DNA Binding Assay Data

Binding of an MTM dimer to a short DNA oligomer is described by the following 1:1 binding equilibrium:

$$M_2{\bullet}Me^2 + D \rightleftharpoons M_2{\bullet}Me^2{\bullet}D, \quad (7)$$

where D and $M_2{\bullet}Me^2{\bullet}D$ denote free DNA and MTM-DNA complex, respectively. The equilibrium constant K_{DNA} in the direction of dissociation is

$$K_{DNA} = [M_2{\bullet}Me^2][D]/[M_2{\bullet}Me^2{\bullet}D] \quad (8)$$

The equations describing conservation of D and M are:

$$[M]_{tot} = 2[M_2{\bullet}Me^{2+}] + 2[M_2{\bullet}Me^2{\bullet}D], \quad (9)$$

$$[D]_{tot} = [D] + [M_2{\bullet}Me^2{\bullet}D], \quad (10)$$

Eqs. (8)-(10) yield $[M_2{\bullet}Me^2]$ and $[M_2{\bullet}Me^2{\bullet}D]$:

$$[M_2 \bullet Me^{2+}] = \frac{1}{4}\left(b + \sqrt{b^2 + 8K_{DNA}[M]_{tot}}\right), \quad (11)$$

where $b = [M]_{tot} - 2[D]_{tot} - 2K_{DNA}$ (12)

and

$$[M_2{\bullet}Me^2{\bullet}D] = 1/2\ ([M]_{tot} - [M_2{\bullet}Me^2]) \quad (13)$$

As in the dimerization assay, the fluorescence signal is a sum of contributions from $M_2{\bullet}Me^{2+}$ and $M_2{\bullet}Me^2{\bullet}D$:

$$I = C_2\,[M_2{\bullet}Me^{2+}] + C_4\,[M_2{\bullet}Me^2{\bullet}D] + C_3, \quad (14)$$

where C_2 and C_3 were defined previously and C_4 is the molar fluorescence of MTM-DNA complex. Nonlinear regression curve fitting to the titration data with SigmaPlot was used to obtain the values of K_d.

The structural parameters of DNA were calculated from the crystal structure coordinates by using 3DNA software.[115]

3.4 Conclusions

In summary, this study exploited the intrinsic MTM fluorescence to interrogate the dimerization and DNA binding properties of MTM (138) and its analogues. We obtained the minimum DNA binding requirements for MTM (138) and established that it is the direct (sequence) and not indirect (conformation) read-out that guides DNA recognition by MTM (138). The 3-side chain derivatization led to modulation of DNA binding affinity of the analogues. Future studies will characterize structural details of binding of MTM (138) and its analogues to different DNA sequences and use the DNA recognition rules established here to map MTM binding sites in oncogenic promoters and other regulatory sites in the genome.

4 SUMMARY

Deoxysugar biosynthetic enzymes provide valuable glycodiversification tools that can be used to engineer molecules for pharmaceutical use. To exploit the full potential of these tools for drug development, the role of these enzymes in the biosynthetic pathway must be determined, and the mechanism of action of the compound to be derivatized should be completely understood. The work in this dissertation provides a better

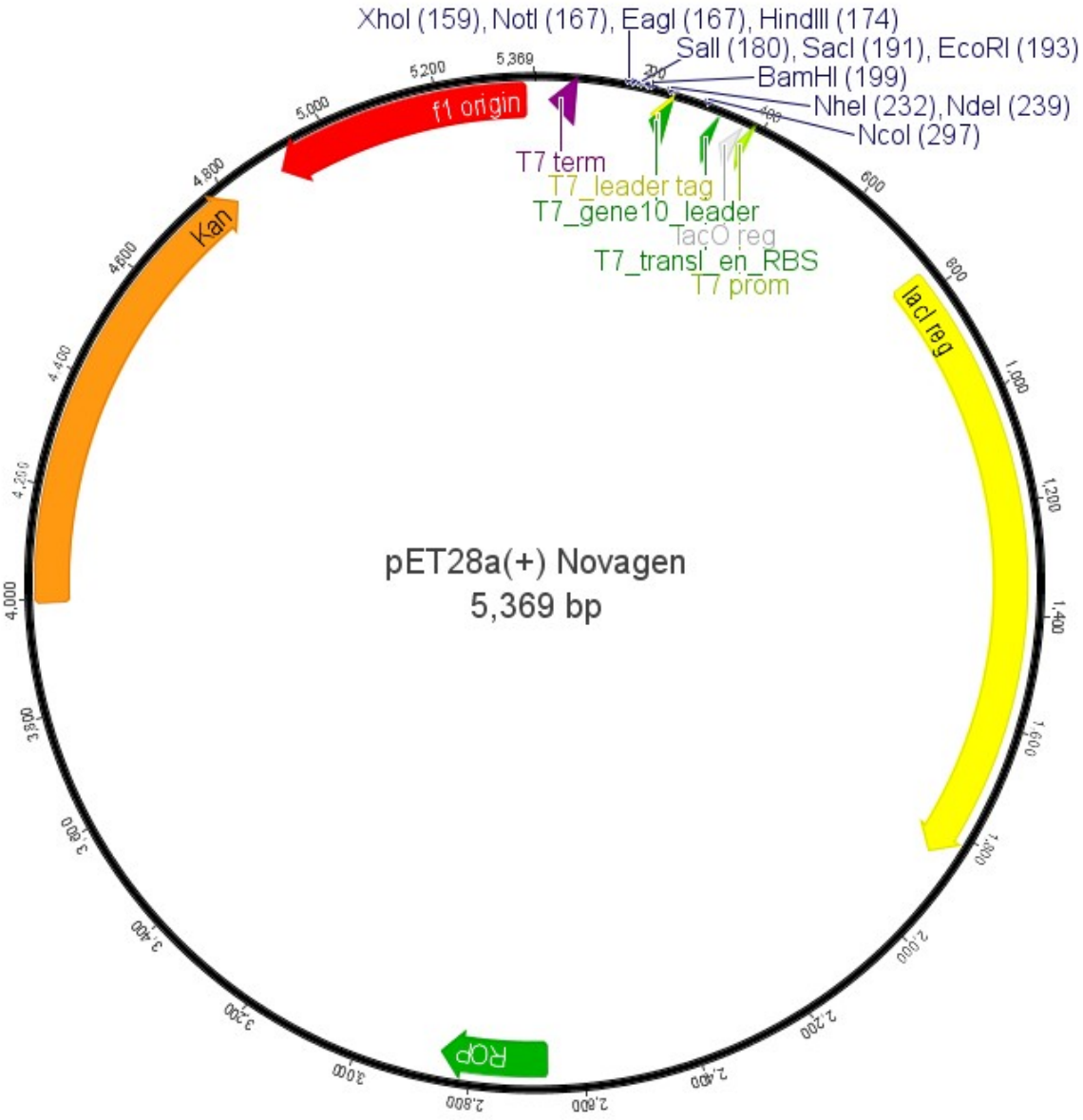
understanding of the mechanism of action of MTM (138), new biosynthetic tools, and TDP-sugar building blocks that can be used to further develop MTM (138) or other compounds for improved druggability.

The identification of the *sqn* gene cluster, responsible for production of SQN A (102), B (103), G-K (104-108) in *Streptomyces* sp. KY40-1 provides eight deoxysugar biosynthetic genes including six that are putatively involved in L-rednose biosynthesis. The proposed pathway is interesting for several reasons. It proposes a glucosamine derived starter unit in place of glucose and suggests that the biosynthetic enzymes are flexible enough to accept both substrates. Four of the cloned SQN deoxysugar biosynthetic enzymes (SqnS1, S5, S7-S8) provided soluble protein. Enzymatic synthesis of TDP-4-keto-3,6-dideoxy-L-idosamine (132) was attempted and while largely unsuccessful, generated TDP-glucosamine (128) and TDP-*N*-acetylglucosamine. Successful enzymatic synthesis of L-rednose (109 A) would validate the proposed biosynthetic pathway, and provide additional novel TDP-amino sugar building blocks to derivatize amino glycosides, MTM (138) or other polyketides.

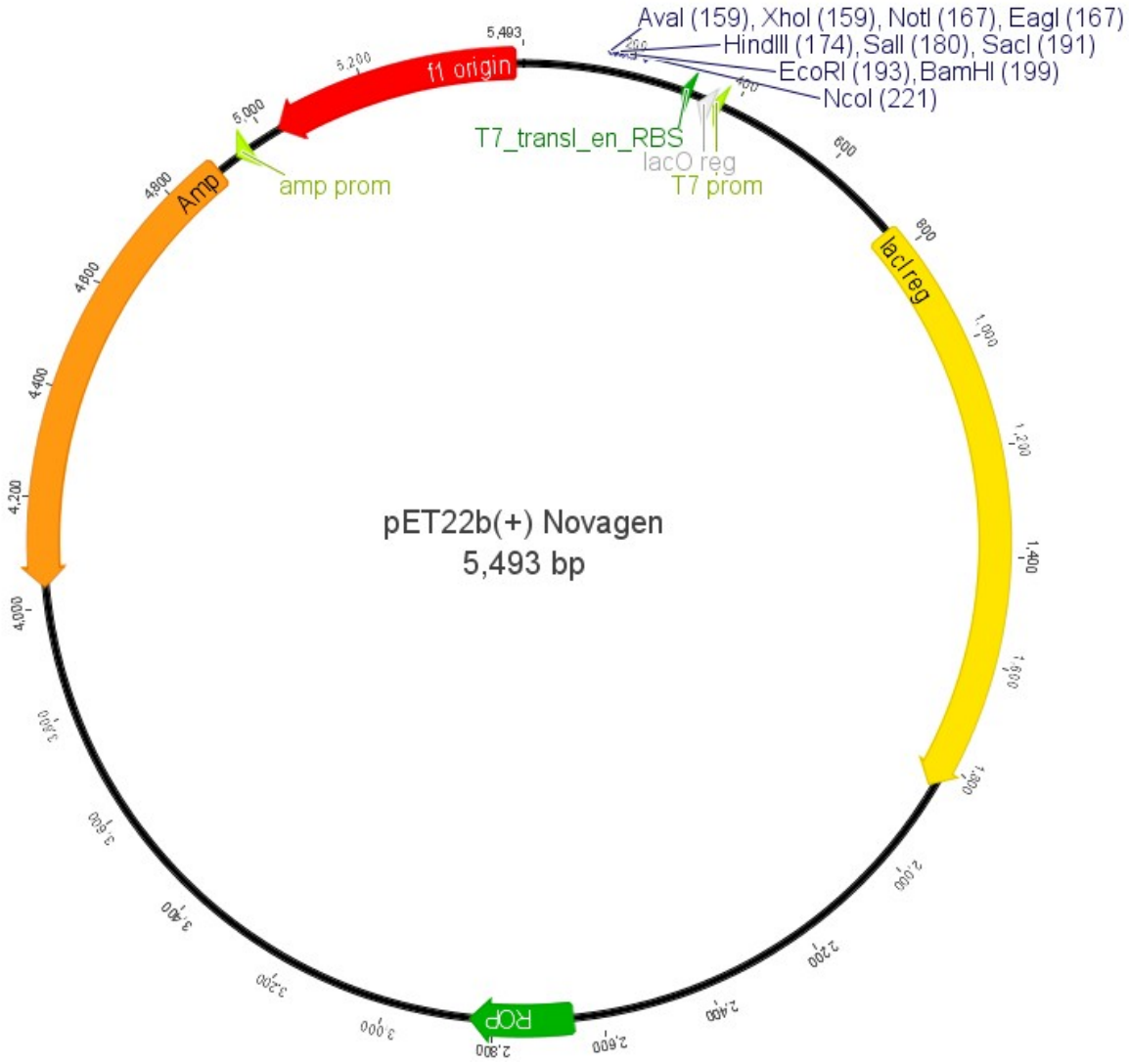
Dimerization and DNA binding studies of MTM (138) and analogues provided a better understanding of the mechanism of action of MTM (138). It was determined that dimerization of MTM (138) and analogues prior to binding DNA, direct (sequence) read-out guides MTM (138) and a minimum DNA binding requirement for MTM (138) was established. It was also revealed that 3-side chain derivatization of MTM (138) effects DNA binding affinity. This information lays initial groundwork necessary to develop MTM (138) as a clinically relevant chemotherapeutic.

5 APPENDIX

5.1 pET28a(+) Plasmid Map

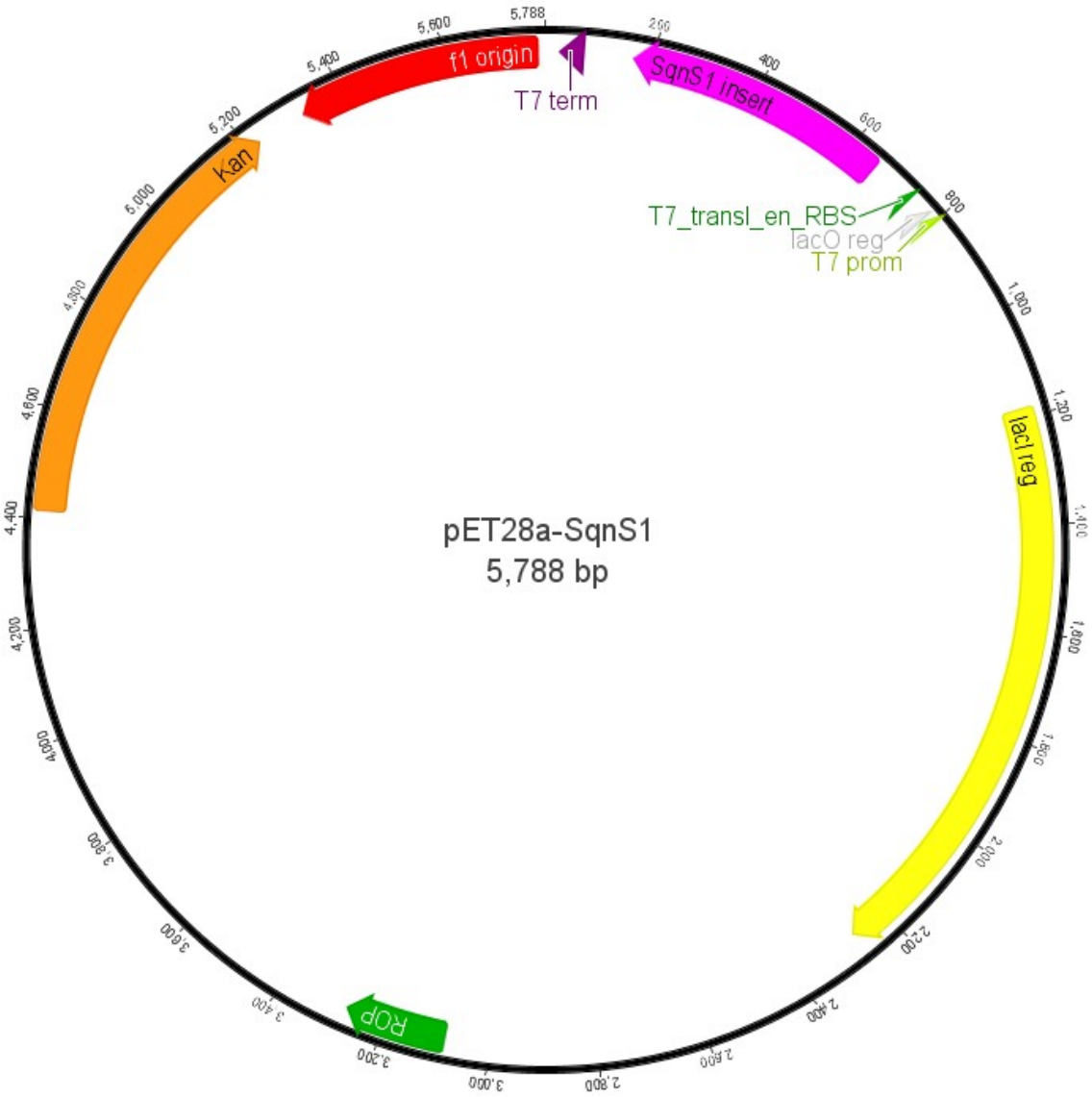


5.2 pET22b(+) Plasmid Map



5.3 SqnS1

5.3.1 Plasmid Map of pET28a- SqnS1_N His6x



5.3.2 *sqnS1* DNA Sequence

ATGCGCAGTGACGCGCTGGAGGACGCCACCGGCGTCTCCTTCGGTCCGCGGC

AGATCAACTACTCCGTGTCGAAGCGCCATACGCTGCGCGGGATCCACAGCGT

GCGCATTCCGCCGGGGCAGGCAAAGTACGTCACCTGTGTACGCGGCTCGCTG

CGCGACATCGTCGTGGACCTGAGGATCGGCTCCCCGACCTTCGGCGAGCACC

GGGTCAATGTGCTGGACGCGGACTCCGGAAGGTCCGTCTACGTGCCCGAGGG

CGTCGGCCACGGCTTCCTCGCACTCAGCGACGACGCCTGCATCTGCTACGTCG

TCTCCACCACGTACGTGCCAGGCACGCAGATCGACGTCAATCCTCTGGACCC

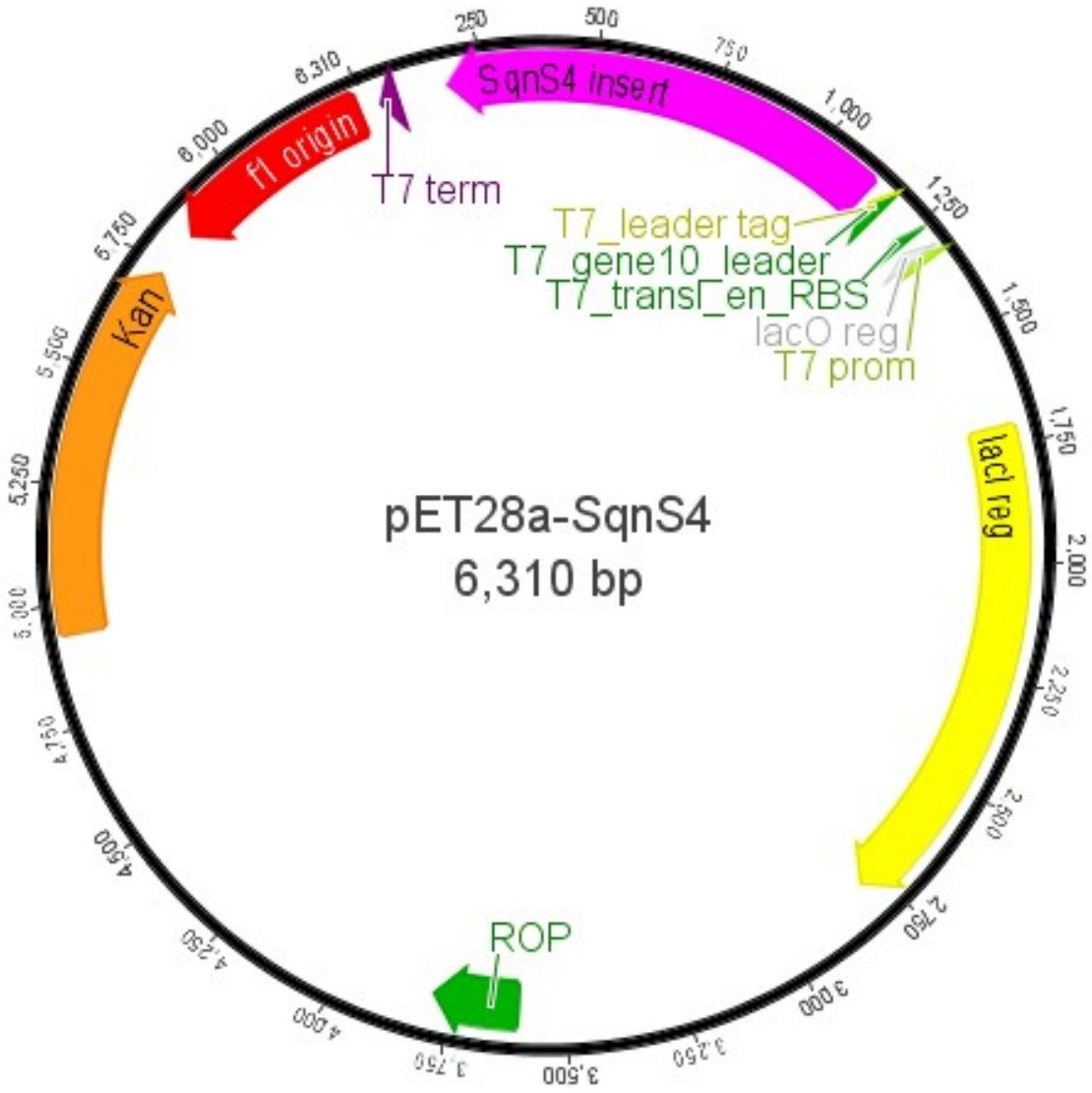
GGATCTCGATCTGCCCTGGGACTGCCCGCAGCCCCCTCTCATCTCGGACAAG

GACGCCAAGGCGCCCACCGTGGCCGAGGCCGTACGGGCAGGCATGCTGCCCC

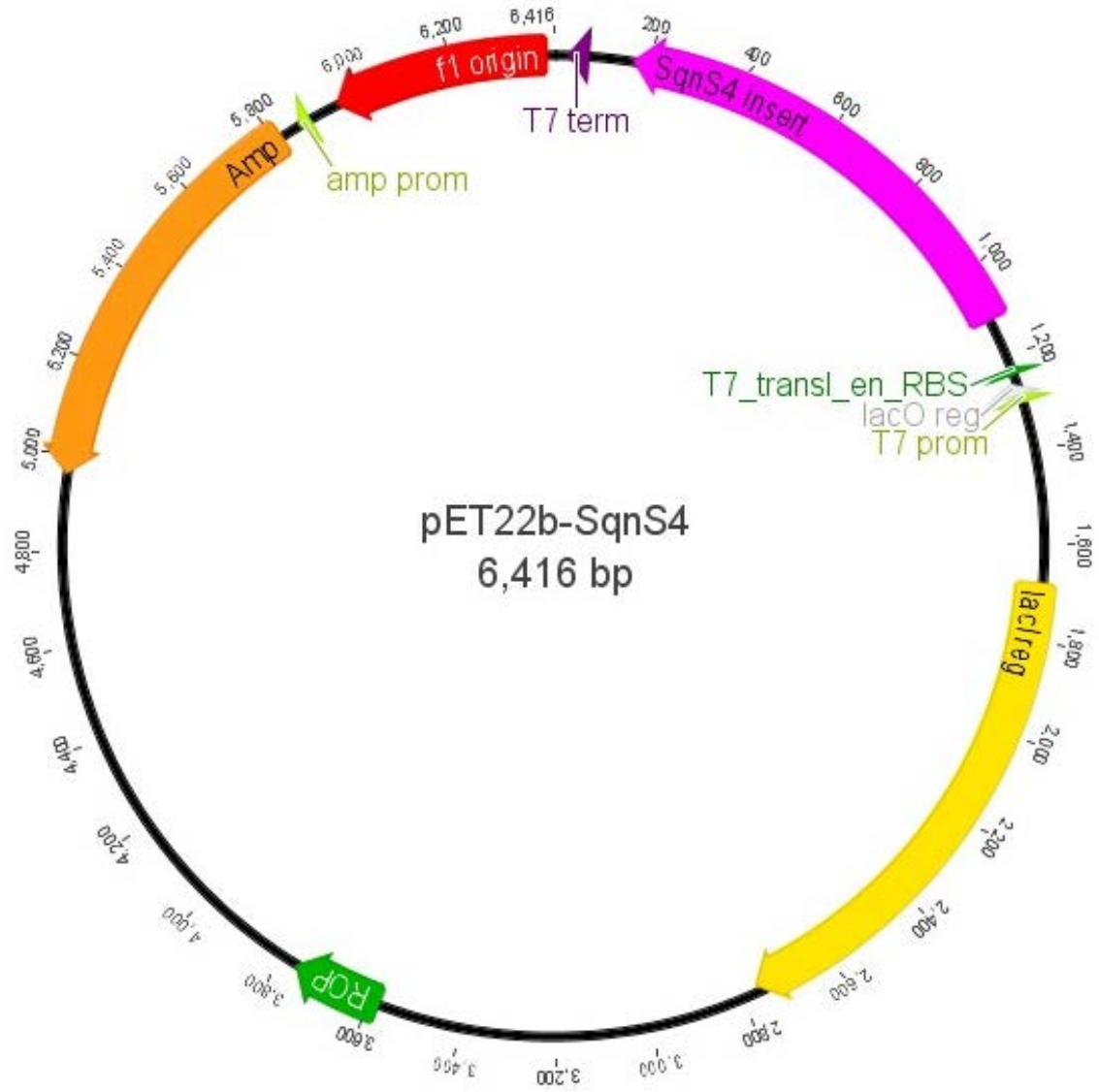
GATTCGACAAGGCAGGCACCTCGTGA

5.4 SqnS4

5.4.1 Plasmid Map of pET28a-SqnS4- N His6x



5.4.2 Plasmid Map of pET22b- SqnS4-C His6x



5.4.3 *sqnS4* DNA Sequence

GTGCACCTGATCGCCCGCACTGCGGCAACCGAGTCGGTGCCGCCCAGCACCT

CCGGAGTGACGACGACGCGGCTGGACCTGCTCACGACGGCACCCGAGGCACT

CACGGAGGTCCTCGCCACAATCGGCGCCGACGTGGTGGTCAACGCGGCGGGC

CGGGCCTGGCGGGCCGACGGGACGGAGATGGCAGCGGGCAACGTCGAACTG

GTCGAACGGGTCGTGACGGCGCTCGCCGCCGCTCCCGGCCCTCCAGTACGGC

TGGTTCAGCTCGGCAGCGTCCACGAGTACGGGGCCGGCGCCCCGGACGGCGC

CACCGGCGAGGACCATGCGCCCACCCCGGTGACGCCGTACGGCCGTACCAAA

CTCCGGGGCACGCAAGCCGTCCTGCGCGGCACACGGGAGCGGGGCGTCGAC

GGGGTGGTGCTCCGGCTCGCCAACGTGATCGGTGCCGGGGTACCGGAGGGCA

GCCTCTTCGGCCGGGTCGCCGGCCACCTCGGCACCTCTGCCCGGGCCGACGC

GCGGGGTGAGAAGCCCGCCGCACTGCATCTGCCGCCCCTGCGTGCGGCCCGC

GACCTGGTGGACACCGGTGACGTCGCCGCGGCGGTCCTGGCCGCGGCCACGG

CACCGGCACGGATCGTCACCGGCCAGGTGATCAATGTGGGCCGCGGCGAGGC

GGTCCCCATGCGCGGGCTGGTCGACCGGATGATCACGCTCAGCGGCCTGGAG

ATCCCGGTGGTCGAGGCGACCGAGGCACCCGCCTCGCGCACCGATGTCGCCC

GGCAGTGCCTGGACGTCTCCCGTGCCCGGCGCCTGCTCGGCTGGCGCCCGCT

GCGCAGCGTCGACGACTCCCTGCGCGACCTTCTCGCATCCGTACTGCCGCCG

GAGCGACCGCCGCTCGGCACACCACTGGGCATCACGGCCGGCGCACCGGCCG

AGGAAGGGGTACGACCATGA

5.4.4 *sqnS4* DNA Sequence Codon Optimized for E. coli

GTGCATTTGATCGCGCGCACAGCGGCAACAGAATCGGTGCCACCTAGTACCT

CAGGAGTAACCACAACCAGACTGGACCTGCTTACGACAGCGCCAGAGGCGCT

GACAGAAGTTTTGGCCACTATCGGGGCAGACGTGGTCGTAAACGCAGCCGGG

AGAGCGTGGCGGGCTGACGGTACCGAAATGGCGGCTGGGAACGTTGAGCTTG

TTGAGCGTGTGGTTACCGCGCTTGCTGCCGCGCCCGGTCCACCAGTAAGATTG

GTTCAACTTGGCAGCGTACACGAATACGGGGCTGGCGCTCCGGACGGCGCTA

CTGGCGAAGACCATGCACCGACGCCAGTCACCCCGTATGGGCGTACGAAGCT

GAGAGGGACGCAAGCAGTCCTGCGTGGAACTCGGGAGCGGGGGGTAGACGG

TGTGGTGCTGAGATTGGCGAATGTCATTGGAGCTGGGGTACCCGAAGGTAGC

CTGTTTGGGCGTGTTGCTGGACATTTAGGAACATCGGCTCGCGCCGATGCTCG

GGGAGAGAAACCCGCCGCGTTGCACTTGCCGCCGCTTCGGGCAGCGCGCGAC

TTAGTTGACACAGGCGACGTGGCGGCAGCTGTCTTAGCCGCCGCTACCGCAC

CAGCCAGAATTGTTACAGGTCAGGTTATCAATGTCGGGCGTGGTGAAGCCGT

ACCGATGCGCGGCTTGGTCGATCGCATGATTACTTTATCCGGTCTTGAAATCC

CGGTCGTAGAAGCGACTGAGGCACCGGCCTCGCGTACGGATGTAGCTCGCCA

ATGTTTGGATGTGTCTCGGGCGAGAAGATTGTTGGGGTGGCGTCCTCTTAGAT

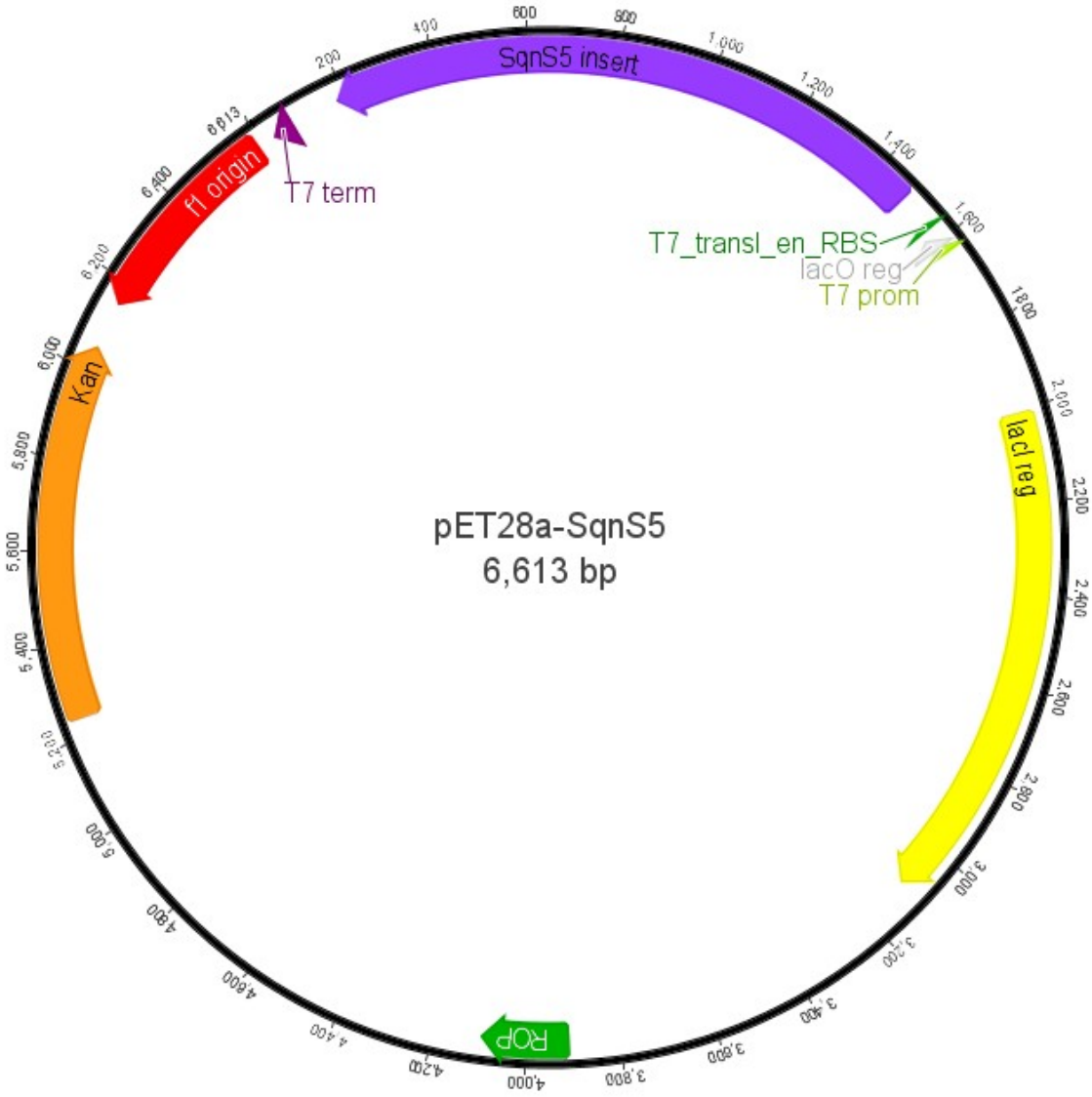
CGGTCGACGATTCTTTACGTGACTTATTAGCATCCGTACTTCCTCCGGAACGC

CCACCTTTGGGCACACCACTTGGGATTACAGCCGGCGCTCCCGCAGAAGAGG

GCGTACGCCCGTGA

5.5 SqnS5

5.5.1 Plasmid Map of pET28a-SqnS5 -N His6x



5.5.2 *sqnS5* DNA Sequence

ATGAGCGACCGCAAGGACCGGATCCTGGAAGAGGTCCGCACATACCACCAG

GAGGTCTCCCCGGAGCGTGAGTTCGTGCCGGGCACGACGGAGATCTGGCCGT

CCGGCGCGGTCCTGGAGGAGGCCGACCGGATGGCCCTGGTGGAGGCGGCGC

TGGATATGCGCATCGCCGCCGGGACCAGCTCCCGCAAGTTCGAGTCGGCCTT

CGCCCGCCGCCTGAAGCGCCGCAAGGCACATCTGACGAACTCCGGTTCGTCG

GCCAATCTGCTGGCGGTGTCGGCACTGACGTCGCACACGCTGGAGGACCGGC

GGCTGCGGCCGGGCGACGAGGTCATCACGGTGGCGGCGGGTTTCCCGACCAC

CGTGAACCCCATCCTGCAGAACGGCCTGATCCCGGTCTTCGTGGACGTGGAC

CTCACCACCTACAACGCGACCGCCGACCGGGTGGCCGCGGCGATCGGTCCCA

AGACCCGAGCCATCATCGTCGCGCACGCCCTGGGCAACCCCTTCGAGGTCGC

GGAGATAGCCCAACTCGCCCAGGAGCACGATCTGTTTTTGATCGAGGACAAC

TGCGACGCGGTCGGCTCGCTCTACGACGGCAAACTCACCGGCACCTTCGGCG

ACATGACCACCGTCAGCTTCTACCCGGCGCACCACCTCACCATGGGCGAGGG

CGGGTGTGTGCTGACCTCCAACCTGGCGCTGGCCAGGATCGTGGAGTCGCTG

CGGGACTGGGGGCGGGACTGCTGGTGCGAGCCGGGCGAGAACGACCGTTGC

ATGAAGCGGTTCAAGTACCAGATGGGCACGCTGCCCGCCGGATACGACCACA

AGTACATCTTCTCGCACGTCGGGTACAACCTGAAGGCCACGGACATCCAGGC

CGCACTCGGTCTGGCGCAGCTGGCCAAGCTGGACTCCTTCATCGAGGCCCGG

CAGCGCAACTGGCGGAGGCTGCGGGAGGGCCTGGACGGGGTTCCCGGGCTG

CTGCTGCCCGAGGCCACGCCGCGTTCCGAGCCGAGCTGGTTCGGGTTCGTGA

TCACCGTCGACCCCGAGGCGCCCTTCGGCCGCGCCGAGTTGGTCGACTTCCTG

GAGTCCCGAAGGATCGGCACCCGCCGCCTGTTCGCCGGCAACCTCACCCGGC

ACCCGGCCTACATCGACCAGCCGCACCGGATCGTGGGCGAACTGACCAACAG

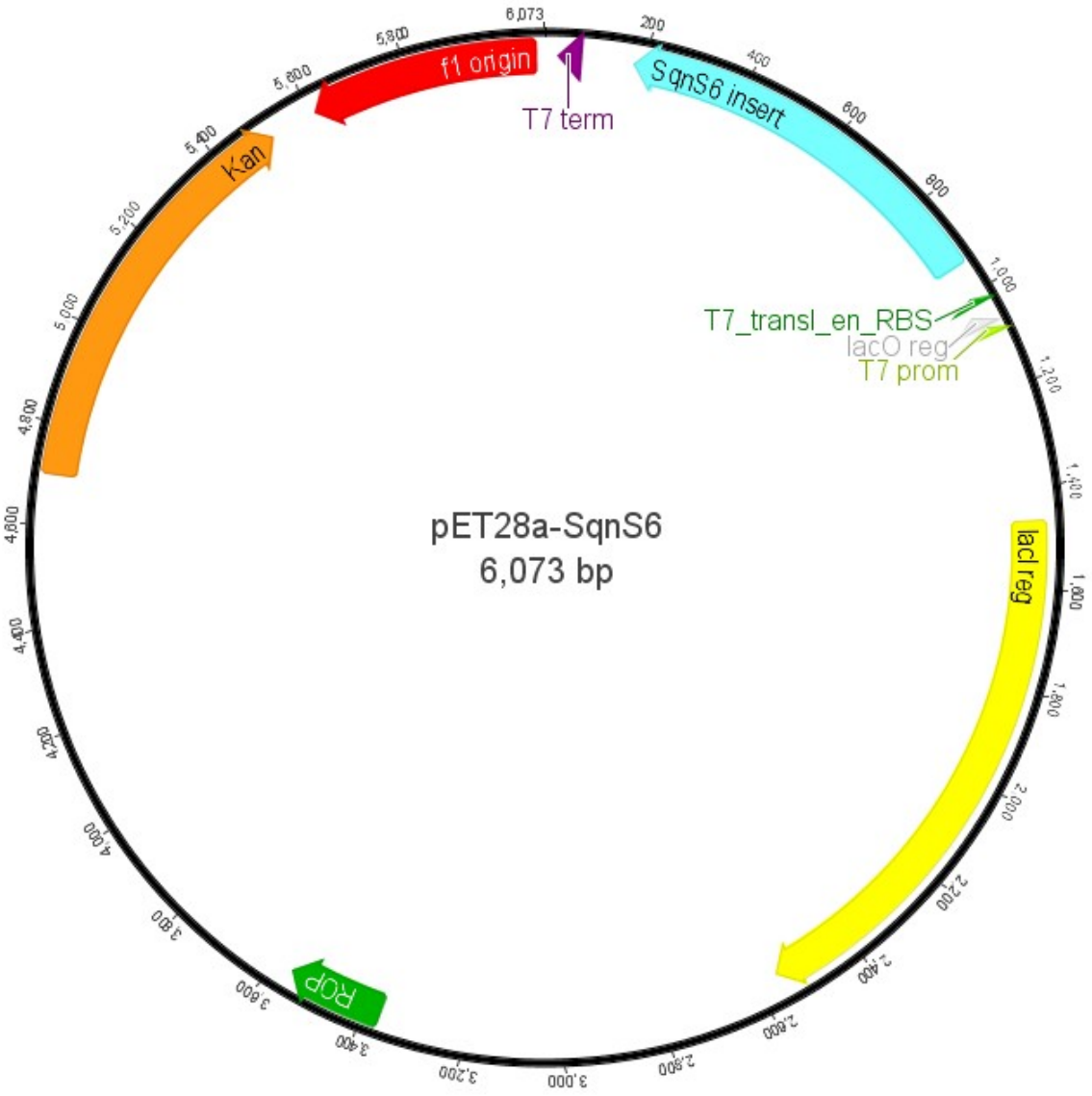
CGACGTCATCACCGAGCACACCTTCTGGATCGGGGTCTACCCGGCGCTCACC

GACGAGATGCTGGACTACGTCACCGCCTCGATCAAGGAGTTCGTGGCCGCGC

GCGGCTGA

5.6 SqnS6

5.6.1 Plasmid Map of pET28a-SqnS6-N His6x



5.6.2 *sqnS6* DNA Sequence

ATGGACATTGTGGGAAACGGTTTCCTGGCGCGAAACCTTCGTTCGCTGGCCG

GCCGGCATCCGGACACCGTGGCCCTGGCAGCGGGTGTGTCCTGGGCGAGCGG

CACCTCCGACGAGGACTTCGCCCGTGAGGCGGCGCTGCTGCACGAGGTCGTC

GAGCGGTGCCAGGACACCGGTCGGCGACTGCTCTTCTTCTCCACGGCGGCGA

CCGGCATGTACGGACTCGCCGAGGGGCCCGGCCGGGAGGACACGCCGGTGA

CGCCCTGCACCCCCTACGGCGCGCACAAACTCGCCCTCGAGCAGCTGCTGCG

CGACTCCGGCGCCGCCCACGTGATCCTTCGGCTCGGCCATCTGGTCGGACCC

GATCAGCCCGAGCACCAGTTACTGCCCACTCTGGTGCGCCAGTTGCGCGAAG

GAGTGGTCCGCGTGCACCGGGAAGCGGCCCGCGACCTGATCGACGTCGGCGA

CGTCATCACGGTCATCGACCGTCTGCTCGCCATGGATCTGCGGGCCGAGACG

GTCAATGTCGCCTCCGGTTTCGCCGTCCCGGTGGGGGACGTCGTCGACCACCT

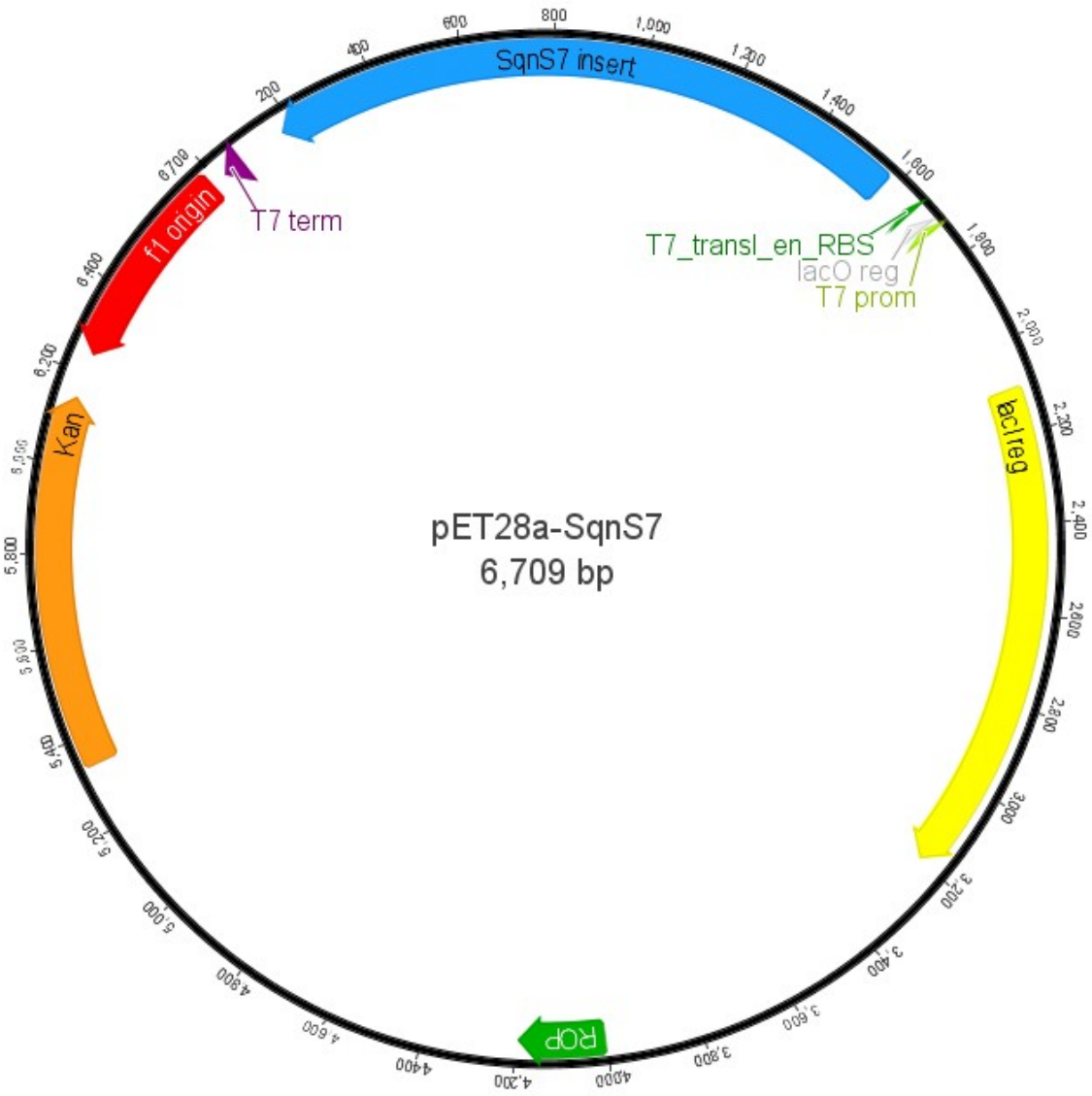
CGCGCGGGCCCTGGGTTTGGAGGCGCGCCGTGAGTTCGTCGACGCGGGCGTC

CGTCAGCACGTCATCTCGACCGAGAAGCTGCGCGCCCTGGTGCCGCAGGTCG

CGGAGATGGGCTTCGGCCCCGGCTACTACCGGCGGATCCTCGGCGACTTCAC

TGCCTCGGTGTCGTCCGGCACGTCCGCGCACGCCTGA

5.7 SqnS7



5.7.1 Plasmid Map of pET28a-SqnS7- N His6x

5.7.2 *sqnS7* DNA Sequence

ATGCCGTCTTCGCTCCTTCAAGCACGCGACAGTGCACGGCAGTTCGCCGACC

GGATCGCCCGGTCGACCGCCACCGCCCGGGGCGTGCACCTGCCGACCGCCGA

CTTCCCGGAATGGCTCGCCGAACGGGGCAGCGTCAACGCCTTCCGCGTGGAC

CGGATCCCCTTCGCGCAACTGGACGGCTGGTCGTTCCAGGAGACCACCGGGA

ACCTGGTGCACCGCAGTGGCCGGTTCTTCACGATCGAGGGCCTGCACGTCAC

CGAGCGGGACGGCCCCTATGGCGACGGCCCGTACGCGGACTGGCACCAACCC

GTCATCAAGCAGCCCGAGGTCGGCATCCTGGGCATCCTGGTCAAGGAGTTCG

ACGGGGTCCTGCACTTCCTGATGCAGGCGAAGATGGAGCCCGGCAACCCGAA

TCTGCTCCAGCTCTCCCCCACCGTGCAGGCGACCCGCAGCAACTACACCAAG

GCCCACCGGGGCGCGGACGTGAAGTACATCGAGCACTTCGTGGGGCCGGGG

CGCGGGCGGATCGTCGCCGACGTCCTGCAGTCCGAACACGGTTCGTGGTTCT

ACCGCAAGTCCAACCGCAACATGATCGTGGAAGCGGTCGGCGACGTGCCCCT

CCTCGACGACTTCTGCTGGCTCACCCTGGGGCAGATCGCCGAGCTCCTGCACC

GCGACAATGTCGTCAACATGGACTCGCGAACGGTGCTGTCATGCCTGCCCGT

CGCACCGAAGGCGAGCAACGCCCTGCACTGCGACGCCGACCTGCTGTCCTGG

ATCACCGGTGAGCGCGCCCGCCACGATGTGCACGCCGAGCGCGTTCCGCTGG

CCGGGCTGCCCGGCTGGCGGCGCCACGACATGGCCATCGAGCACGAGGACG

GACGCTACTTCACGGTGGTGGCGGTGGCCGTCCAGGCCGGCAACCGCGAGGT

CACCAGCTGGACCCAGCCGTTGTTCGAGCCGGTGGGCCTCGGCGTCACCGCC

TTCCTCACCCGGACGTTCGACGGCATACCCCATGTGCTGGTGCACGCCCGGGT

CGAAGGCGGGTTCCTGGACACCGTCGAACTGGGCCCCACGGTCCAGTACACG

CCCGGCAACTACGCCCATCTCCCGGACAAGGAGCGCCCGTTGTTCCTCGACA

CCGTGCTCGAGGTGCCGCCGGACCGGATCCGGTACGAGGCAGTGCACTCGGA

GGAGGGCGGCCGGTTCCTCAACGCGGAGAGCCGGTATCTCGTCGTCGAGGCC

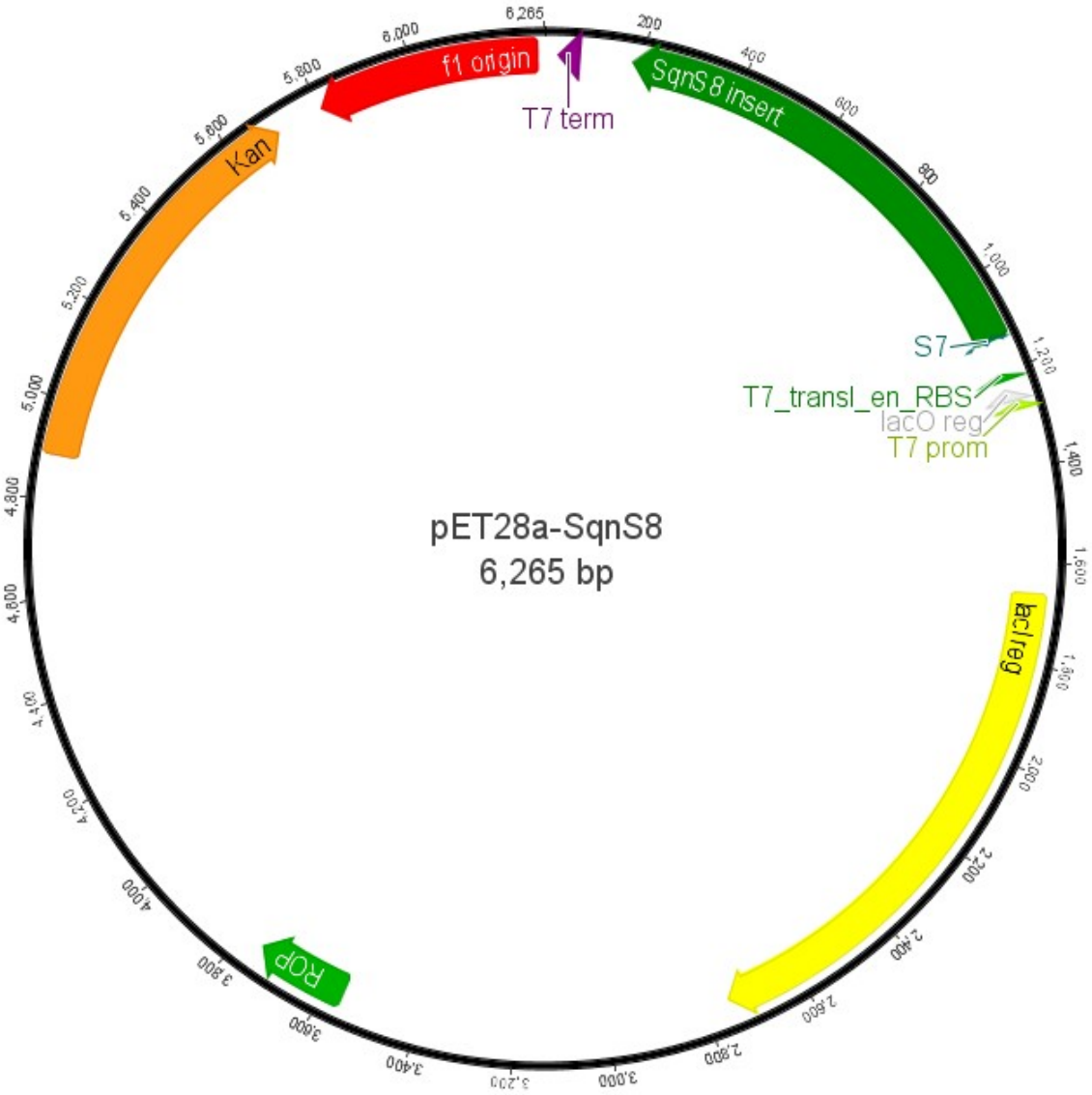
GACGAGCGGGAGGCCCCACTCGATCCGCCGCCCGGCTACGCCTGGGTCACCC

CCGACCAGCTCACCTGGCTGGTCCGCCACGGTCACTACCTCAACGTCCAGGC

CCGCACACTGCTGGCCTGCCTCAACGCCACCTTGGCGGGCGCGCGATGA

5.8 SqnS8

5.8.1 Plasmid Map of pET28a-SqnS8- N His6x



5.8.2 *sqnS8* DNA Sequence

ATGATCACGCCGGTACGCATCGGGGTGCTCGGCTGCGCGGACATCGCGGTAC

GGCGGATGCTGCCGGCGTTCGCCGCATCCCGCGACGTGGAGGTCGCCGCCGT

CGCCAGCCGCGATCCGGCCAGGGCCGAGGAGGTGGCCGCACGCTTCGGCTGC

CGTCCGGTGCGGGGATACGCCGAACTGCTCGACGCCGACGACGTGCAGGCGG

TGTACGTCCCCCTGCCGGCCGCACTGCACGCCCCCTGGGTGGAGGCCGCTCT

GAACGCGGGCAAACACGTGCTGGCCGAGAAGCCCCTGACGACGGACCCGGA

GAGCACCGAGCGTCTCCTCGACCTCGCCGCGAAGCAGGGTGCGGCCCTGATG

GAGAACGTGATGTTCGTCCACCATCCGCGGCACGAGGCGGTGCGGCGCCTCG

TCTCCGCGGGACGGATCGGTGAACTCCGCTCCCTCCACGCGGCGTTCACCATC

CCCCCGCTCCCCGACGGTGACATCCGCTACGCGCCCGAGCTCGGCGGTGGCG

CCCTGTCGGACGTCGGCCTCTACCCGCTGCGCGCCGCCCTGCACTTCCTGGGT

CCCGCTCTCGAGGTCGTCGGCGCCAGGCTCACCCGAGGCGGCGGGCGACAGG

TGGAGACGTCCGGGGCCGCGCTGCTGGGCACACCCGAAGGCGTCACCGCACA

CCTCACGTTCGGGATGGAGCACGCCTACCTCTCGCGGTACGAACTTTGGGGC

AGCGAGGGCCGGATCACGGTGGACCGGGCGTTCACGCCGCCGGCCGACTTCG

TGCCCGTGATCGCCGTGCACGGGCACACCACGGAGGAGATCCGGCTGGAGCC

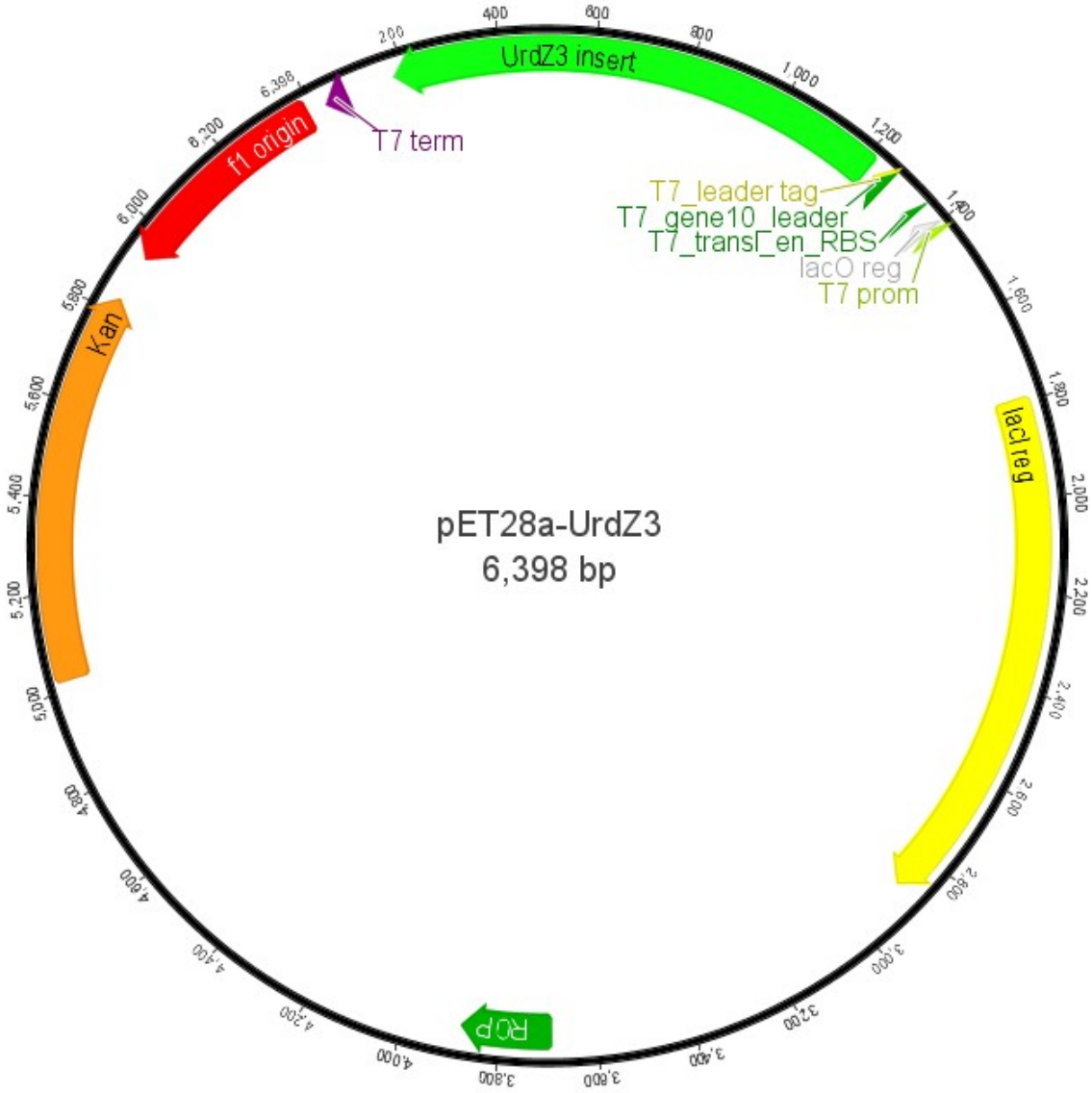
CGCCGACCAGGTCGCGGCCACCGTCGCCGCCTTCGTCACCGCGGTCCGCGCC

GGGTCCGCGGTGCGTGGGGACACCCTGCGGCAGGCCGAGCTGCTGGACGAC

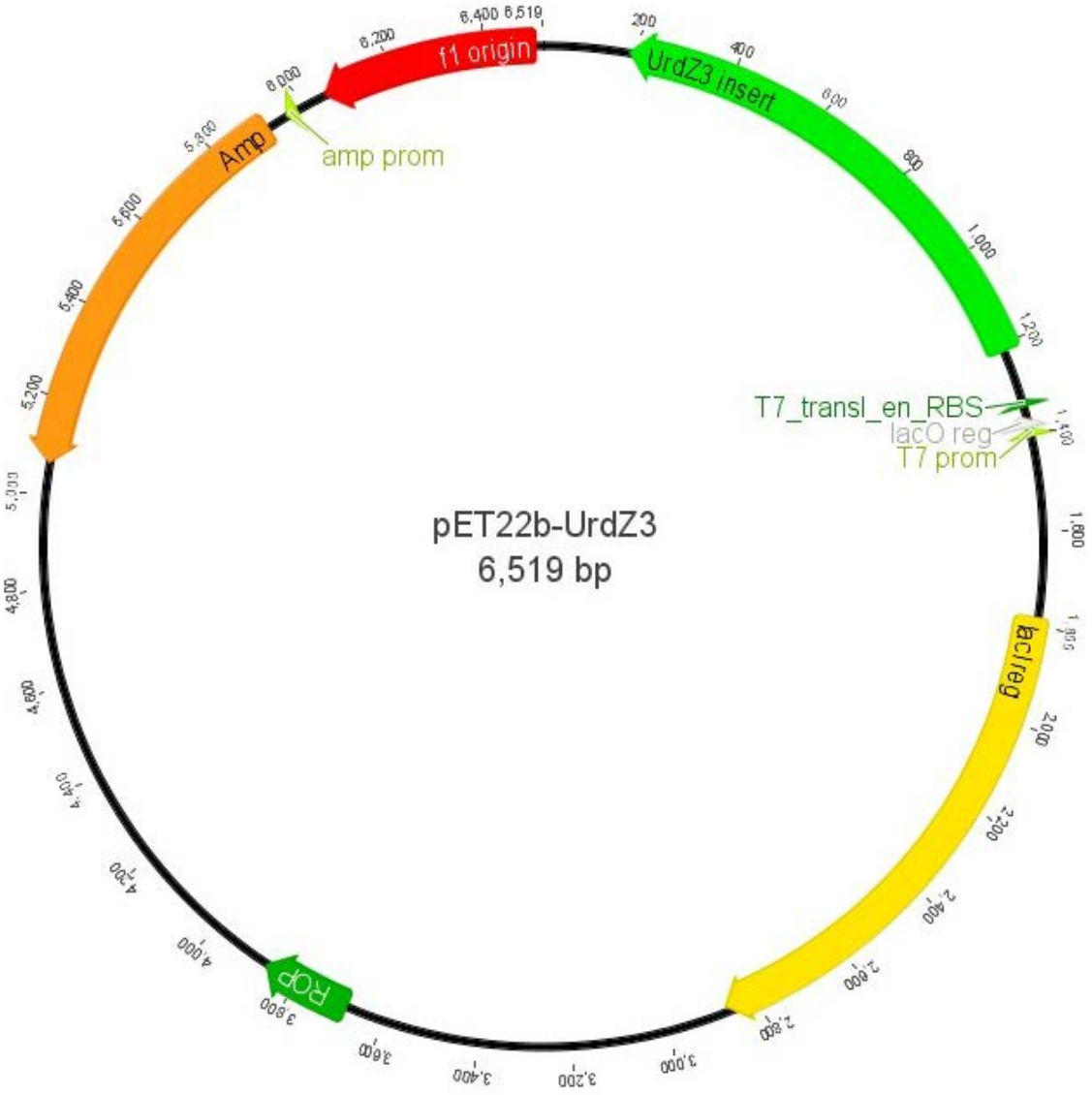
GTACGCCGTCTGTCCGCATGA

5.9 UrdZ3

5.9.1 Plasmid Map of pET28a-UrdZ3- N His6x



5.9.2 Plasmid Map of pET22b-UrdZ3- N His6x



5.9.3 *urdZ3* DNA Sequence

ATGGTCGTCCTGGGAAGCGGCGGTTTTCTGGGGCGCCACATCGGCGACGCCT

TCGCCGCGCAGGGAGCGCGGGTGCACGCCGTGACCCGGAACAGCTCGGACC

GTACTCCCCACGGCGCCGCGACCGGGTCCCGCACGACACCCATCGACCTGCT

GACGACCCCACCGCAGGAGGTCGCCGAGTTCCTCGGCTCGGTCGGCGCCGAT

GTCGTCGTCAACGCCGCCGGCCGGGCCTGGCGGGCCGACGAGGCCCAGATGA

CCGCGGGCAACGCCGAACTGGTCCTGCGCCTCGTCACCGCCCTCGAGCAAGT

GCCCGGCCCGCCGATCCGGCTGGTGCAGCTCGGCAGCGTCCACGAGTACGGG

GCGGGCGCGCCGAACACCGGCACCGGTGAGGAGCACGAGCCCACCCCGGTG

ACCGCCTACGGACGCACCAAGCTCCTCGGGACGCAGACCGTCCTGCGCGCCG

CGCAGGACCGCACTGTCGAGGGCGTCGTGCTCCGGCTCGCCAACGTCATCGG

TGCGGGCGTTCCGGAGGGCAGCCTGTTCGGCAGGGTCGCCGCGCACCTCGCG

GCGGCCGCGCGCGCCGAGTCCCGGGGGGAGAAGGCCGCCGAGCTGCGCCTT

CCGCCACTGGCCGCGGCGCGCGACGTCGTCGACGCGCGTGACGCCGCCGACG

CGGTGGTGGCCGCCGCCACCACGGTCGGTGTCGCCGGACGGGTGATCAACGT

GGGGCGCGGCGAGGCGGTCGGGACACGCGAGCTGATCGACCGGATGGTCCG

GCTCAGCGGCCTGGCCGTCCCGGTCGTGGAAGAGACCGCCGCGGCGATGTCA

CGCACGGACGTGGACTGGCAGCGGCTGGATGTGTCCCGCGCGCGTCGGCTGC

TGGGCTGGCGGCCGCGCCGCAGCCTCGATGCCTCGCTGCGCGACCTTCTCGC

GGCGGCCCTGCCGTCATCCTCGACACGCCTCGGCGGCGTCCCGGACGTCGCA

TCCCGCCGCGCACCACTCGACGCTCCGACCACCGCGCCGGTTGACGAAAGGA

AGATAG

6. REFERENCES

[1] Weymouth-Wilson, A. C. (1997) The role of carbohydrates in biologically active natural products, *Nat Prod Rep 14*, 99-110.

[2] Elshahawi, S. I., Shaaban, K. A., Kharel, M. K., and Thorson, J. S. (2015) A comprehensive review of glycosylated bacterial natural products, *Chem Soc Rev 44*, 7591-7697.

[3] Klyne, W. (1950) The configuration of the anomeric carbon atoms in some cardiac glycosides, *Biochem J 47*, xli-ii.

[4] Melo, A., and Glaser, L. (1965) The Nucleotide Specificity and Feedback Control of Thymidine Diphosphate D-Glucose Pyrophosphorylase, *J Biol Chem 240*, 398-405.

[5] Glaser, L., and Zarkowsky, R. (1971) In *The Enzymes* (Boyer, P. D., Ed.) 3d ed., pp 465-480, Academic Press, New York,.

[6] Gabriel, O. (1973) Carbohydrates in solution; a symposium sponsored by the Division of Carbohydrate Chemistry at the 162nd meeting of the American Chemical Society, Washington, D.C., Sept. 14-15, 1971, In *Advances in chemistry series*, (Gould, R., Ed.), pp 387-410, American Chemical Society, Washington,.

[7] Gabriel, O., and van Lenten, L. (1978) In *Biochemistry of carbohydrates II* (Manners, D. J., Ed.), pp 1-36, University Park Press, Baltimore.

[8] Frey, P. A. (1987) Pyridine Nucleotide Coenzymes, (Dolphin, D., Avramovic, O., and Poulson, R., Eds.), pp 461-511, Wiley-Interscience, New York.

[9] Grisebach, H. (1978) Biosynthesis of sugar components of antibiotic substances, *Advances in carbohydrate chemistry and biochemistry 35*, 81-126.

[10] Draeger, G., Park, S.-H., and Floss, H. G. (1999) Mechanism of the 2-Deoxygenation Step in the Biosynthesis of the Deoxyhexose Moieties of the Antibiotics Granaticin and Oleandomycin, *J Am Chem Soc 121*, 2611–2612.

[11] He, X., and Liu, H. W. (2002) Mechanisms of enzymatic CbondO bond cleavages in deoxyhexose biosynthesis, *Curr Opin Chem Biol 6*, 590-597.

[12] Thorson, J. S., and Liu, H. W. (1993) Characterization of the first PMP-dependent iron–sulfur-containing enzyme which is essential for the biosynthesis of 3,6-dideoxyhexoses, *J Am Chem Soc 115*, 7539–7540.

[13] Gassner, G. T., Johnson, D. A., Liu, H. W., and Ballou, D. P. (1996) Kinetics of the reductive half-reaction of the iron-sulfur flavoenzyme CDP-6-deoxy-L-threo-D-glycero-4-hexulose-3-dehydrase reductase, *Biochemistry 35*, 7752-7761.

[14] Miller, V. P., Thorson, J. S., Ploux, O., Lo, S. F., and Liu, H. W. (1993) Cofactor characterization and mechanistic studies of CDP-6-deoxy-Δ 3,4-glucoseen reductase: exploration into a novel enzymatic C-O bond cleavage event, *Biochemistry 32*, 11934-11942.

[15] Gonzalez-Porqué, P. (1987) Pyridine Nucleotide Coenzymes, In *Pyridine Nucleotide Coenzymes* (Dolphin, D., Avramovic, O., and Poulson, R., Eds.), pp 391–419, Wiley-Interscience, New York.

[16] Weigel, T. M., Miller, V. P., and Liu, H. W. (1992) Mechanistic and stereochemical studies of a unique dehydration catalyzed by CDP-4-keto-6-deoxy-D-glucose-3-dehydrase: a pyridoxamine 5'-phosphate dependent enzyme isolated from Yersinia pseudotuberculosis, *Biochemistry 31*, 2140-2147.

[17] Burns, K. D., Pieper, P. A., Liu, H. W., and Stankovich, M. T. (1996) Studies of the redox properties of CDP-6-deoxy-L-threo-D-glycero-4-hexulose-3-dehydrase (E1) and CDP-6-deoxy-L-threo-D-glycero-4-hexulose-3-dehydrase reductase (E3): two important enzymes involved in the biosynthesis of ascarylose, *Biochemistry 35*, 7879-7889.

[18] Thorson, J. S., and Liu, H. W. (1993) Coenzyme B6 as a redox cofactor: a new role for an old coenzyme?, *J Am Chem Soc 115*, 12177–12178.

[19] Johnson, D. A., Gassner, G. T., Bandarian, V., Ruzicka, F. J., Ballou, D. P., Reed, G. H., and Liu, H. W. (1996) Kinetic characterization of an organic radical in the ascarylose biosynthetic pathway, *Biochemistry 35*, 15846-15856.

[20] Thorson, J. S., Lo, S. F., Ploux, O., He, X., and Liu, H. W. (1994) Studies of the biosynthesis of 3,6-dideoxyhexoses: molecular cloning and characterization of the asc (ascarylose) region from Yersinia pseudotuberculosis serogroup VA, *J Bacteriol 176*, 5483-5493.

[21] Hong, L., Zhao, Z., and Liu, H. W. (2006) Characterization of SpnQ from the spinosyn biosynthetic pathway of Saccharopolyspora spinosa: mechanistic and evolutionary implications for C-3 deoxygenation in deoxysugar biosynthesis, *J Am Chem Soc 128*, 14262-14263.

[22] Hong, L., Zhao, Z., Melançon, C. E., 3rd, Zhang, H., and Liu, H. W. (2008) In vitro characterization of the enzymes involved in TDP-D-forosamine biosynthesis in the spinosyn pathway of Saccharopolyspora spinosa, *J Am Chem Soc 130*, 4954-4967.

[23] Thorson, J. S., Lo, S. F., and Liu, H. W. (1993) Molecular basis of 3,6-dideoxyhexose biosynthesis: elucidation of CDP-ascarylose biosynthetic genes and their relationship to other 3,6-dideoxyhexose pathways, *J Am Chem Soc 115*, 5827-5828.

[24] Zhao, L., Borisova, S., Yeung, S. M., and Liu, H. (2001) Study of C-4 deoxygenation in the biosynthesis of desosamine: evidence implicating a novel mechanism, *J Am Chem Soc 123*, 7909-7910.

[25] Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F., and Miller, N. E. (2001) Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods, *Nucleic Acids Res 29*, 1097-1106.

[26] Szu, P. H., Ruszczycky, M. W., Choi, S. H., Yan, F., and Liu, H. W. (2009) Characterization and mechanistic studies of DesII: a radical S-adenosyl-L-methionine enzyme involved in the biosynthesis of TDP-D-desosamine, *J Am Chem Soc 131*, 14030-14042.

[27] Lin, G. M., Choi, S. H., Ruszczycky, M. W., and Liu, H. W. (2015) Mechanistic Investigation of the Radical S-Adenosyl-L-methionine Enzyme DesII Using Fluorinated Analogues, *J Am Chem Soc 137*, 4964-4967.

[28] Thibodeaux, C. J., Melançon, C. E., and Liu, H. W. (2007) Unusual sugar biosynthesis and natural product glycodiversification, *Nature 446*, 1008-1016.

[29] Thibodeaux, C. J., Melançon, C. E., 3rd, and Liu, H. W. (2008) Natural-product sugar biosynthesis and enzymatic glycodiversification, *Angew Chem Int Ed 47*, 9814-9859.
[30] Soda, K., Yoshimura, T., and Esaki, N. (2001) Stereospecificity for the hydrogen transfer of pyridoxal enzyme reactions, *Chem Rec 1*, 373-384.
[31] Jansonius, J. N. (1998) Structure, evolution and action of vitamin B6-dependent enzymes, *Curr Opin Struct Biol 8*, 759-769.
[32] Eliot, A. C., and Kirsch, J. F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations, *Annu Rev Biochem 73*, 383-415.
[33] Percudani, R., and Peracchi, A. (2003) A genomic overview of pyridoxal-phosphate-dependent enzymes, *EMBO Rep 4*, 850-854.
[34] Koropatkin, N. M., Liu, H. W., and Holden, H. M. (2003) High resolution x-ray structure of tyvelose epimerase from Salmonella typhi, *J Biol Chem 278*, 20874-20881.
[35] Hallis, T. M., Zhao, Z., and Liu, H.-w. (2000) New Insights into the Mechanism of CDP-d-Tyvelose 2-Epimerase: An Enzyme-Catalyzing Epimerization at an Unactivated Stereocenter, *J Am Chem Soc 122*, 10493–10503.
[36] Frey, P. A. (1996) The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose, *FASEB J 10*, 461-470.
[37] Thoden, J. B., Wohlers, T. M., Fridovich-Keil, J. L., and Holden, H. M. (2000) Crystallographic evidence for Tyr 157 functioning as the active site base in human UDP-galactose 4-epimerase, *Biochemistry 39*, 5691-5701.
[38] Alam, J., Beyer, N., and Liu, H. W. (2004) Biosynthesis of colitose: expression, purification, and mechanistic characterization of GDP-4-keto-6-deoxy-D-mannose-3-dehydrase (ColD) and GDP-L-colitose synthase (ColC), *Biochemistry 43*, 16450-16460.
[39] Dong, C., Major, L. L., Allen, A., Blankenfeldt, W., Maskell, D., and Naismith, J. H. (2003) High-resolution structures of RmlC from Streptococcus suis in complex with substrate analogs locate the active site of this class of enzyme, *Structure 11*, 715-723.
[40] Chen, H., Zhao, Z., Hallis, T. M., Guo, Z., and Liu Hw, H. (2001) Insights into the Branched-Chain Formation of Mycarose: Methylation Catalyzed by an (S)-Adenosylmethionine-Dependent Methyltransferase, *Angew Chem Int Ed 40*, 607-610.
[41] Takahashi, H., Liu, Y. N., Chen, H., and Liu, H. W. (2005) Biosynthesis of TDP-L-mycarose: the specificity of a single enzyme governs the outcome of the pathway, *J Am Chem Soc 127*, 9340-9341.
[42] Chen, H., Guo, Z., and Liu, H.-w. (1998) Biosynthesis of Yersiniose: Attachment of the Two-Carbon Branched-Chain Is Catalyzed by a Thiamine Pyrophosphate-Dependent Flavoprotein, *J Am Chem Soc 120*, 11796–11797.
[43] Uchida, T., Imoto, M., Watanabe, Y., Miura, K., Dobashi, T., Matsuda, N., Sawa, T., Naganawa, H., Hamada, M., and Takeuchi, T. (1985) Saquayamycins, new aquayamycin-group antibiotics, *J Antibiot 38*, 1171-1181.

[44] Shaaban, K. A., Ahmed, T. A., Leggas, M., and Rohr, J. (2012) Saquayamycins G-K, cytotoxic angucyclines from Streptomyces sp. Including two analogues bearing the aminosugar rednose, *J Nat Prod 75*, 1383-1392.
[45] Antal, N., Fiedler, H. P., Stackebrandt, E., Beil, W., Stroch, K., and Zeeck, A. (2005) Retymicin, galtamycin B, saquayamycin Z and ribofuranosyllumichrome, novel secondary metabolites from Micromonospora sp. Tu 6368. I. Taxonomy, fermentation, isolation and biological activities, *J Antibiot 58*, 95-102.
[46] Sekizawa, R., Iinuma, H., Naganawa, H., Hamada, M., Takeuchi, T., Yamaizumi, J., and Umezawa, K. (1996) Isolation of novel saquayamycins as inhibitors of farnesyl-protein transferase, *J Antibiot 49*, 487-490.
[47] Goodsell, D. S. (1999) The molecular perspective: the ras oncogene, *Stem Cells 17*, 235-236.
[48] Wolfe, A. D., and Hahn, F. E. (1964) Erythromycin: Mode of Action, *Science 143*, 1445-1446.
[49] Schlunzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001) Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria, *Nature 413*, 814-821.
[50] Frederick, C. A., Williams, L. D., Ughetto, G., van der Marel, G. A., van Boom, J. H., Rich, A., and Wang, A. H. (1990) Structural comparison of anticancer drug-DNA complexes: adriamycin and daunomycin, *Biochemistry 29*, 2538-2549.
[51] Pigram, W. J., Fuller, W., and Hamilton, L. D. (1972) Stereochemistry of intercalation: interaction of daunomycin with DNA, *Nat New Biol 235*, 17-19.
[52] Seeber, S., Schilcher, R. B., Brucksch, K. P., Kading, J., and Schmidt, C. G. (1977) Cystostatic efficacy of DNA-complexes of adriamycin, daunomycin, and actinomycin D. II. Comparative in vivo studies in an Ehrlich-ascites tumor, *Z Krebsforsch Klin Onkol Cancer Res Clin Oncol 90*, 307-312.
[53] Tewey, K. M., Rowe, T. C., Yang, L., Halligan, B. D., and Liu, L. F. (1984) Adriamycin-induced DNA damage mediated by mammalian DNA topoisomerase II, *Science 226*, 466-468.
[54] Basnet, D. B., Oh, T. J., Vu, T. T., Sthapit, B., Liou, K., Lee, H. C., Yoo, J. C., and Sohng, J. K. (2006) Angucyclines Sch 47554 and Sch 47555 from Streptomyces sp. SCC-2136: cloning, sequencing, and characterization, *Mol Cells 22*, 154-162.
[55] Zhang, Y., Huang, H., Chen, Q., Luo, M., Sun, A., Song, Y., Ma, J., and Ju, J. (2013) Identification of the grincamycin gene cluster unveils divergent roles for GcnQ in different hosts, tailoring the L-rhodinose moiety, *Org Lett 15*, 3254-3257.
[56] Hoffmeister, D., Ichinose, K., Domann, S., Faust, B., Trefzer, A., Drager, G., Kirschning, A., Fischer, C., Kunzel, E., Bearden, D., Rohr, J., and Bechthold, A. (2000) The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster, *Chem Biol 7*, 821-831.
[57] Erb, A., Luzhetskyy, A., Hardter, U., and Bechthold, A. (2009) Cloning and sequencing of the biosynthetic gene cluster for saquayamycin Z and galtamycin B

and the elucidation of the assembly of their saccharide chains, *Chembiochem 10*, 1392-1401.
[58] Abdelfattah, M. S., Kharel, M. K., Hitron, J. A., Baig, I., and Rohr, J. (2008) Moromycins A and B, isolation and structure elucidation of C-glycosylangucycline-type antibiotics from Streptomyces sp. KY002, *J Nat Prod 71*, 1569-1573.
[59] Alexeev, I., Sultana, A., Mantsala, P., Niemi, J., and Schneider, G. (2007) Aclacinomycin oxidoreductase (AknOx) from the biosynthetic pathway of the antibiotic aclacinomycin is an unusual flavoenzyme with a dual active site, *Proc Natl Acad Sci U S A 104*, 6170-6175.
[60] Parsons, T. F., and Preiss, J. (1978) Biosynthesis of bacterial glycogen. Isolation and characterization of the pyridoxal-P allosteric activator site and the ADP-glucose-protected pyridoxal-P binding site of *Escherichia coli* B ADP-glucose synthase, *J Biol Chem 253*, 7638-7645.
[61] Stockmann, M., and Piepersberg, W. (1992) Gene probes for the detection of 6-deoxyhexose metabolism in secondary metabolite-producing streptomycetes, *FEMS Microbiol Lett 69*, 185-189.
[62] Mengin-Lecreulx, D., and van Heijenoort, J. (1996) Characterization of the essential gene glmM encoding phosphoglucosamine mutase in *Escherichia coli*, *The J Biol Chem 271*, 32-39.
[63] Dutka-Malen, S., Mazodier, P., and Badet, B. (1988) Molecular cloning and overexpression of the glucosamine synthetase gene from *Escherichia coli*, *Biochimie 70*, 287-290.
[64] Kharel, M. K., Lian, H., and Rohr, J. (2011) Characterization of the TDP-D-ravidosamine biosynthetic pathway: one-pot enzymatic synthesis of TDP-D-ravidosamine from thymidine-5-phosphate and glucose-1-phosphate, *Org Biomol Chem 9*, 1799-1808.
[65] Westrich, L., Domann, S., Faust, B., Bedford, D., Hopwood, D. A., and Bechthold, A. (1999) Cloning and characterization of a gene cluster from Streptomyces cyanogenus S136 probably involved in landomycin biosynthesis, *FEMS Microbiol Lett 170*, 381-387.
[66] Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data, *Bioinformatics 28*, 1647-1649.
[67] Ryan, W. G. (1970) Mithramycin for Paget's disease of bone, *N Engl J Med 283*, 1171.
[68] Ryan, W. G., Schwartz, T. B., and Northrop, G. (1970) Experiences in the treatment of Paget's disease of bone with mithramycin, *Jama 213*, 1153-1157.
[69] Brown, J. H., and Kennedy, B. J. (1965) Mithramycin in the Treatment of Disseminated Testicular Neoplasms, *N Engl J Med 272*, 111-118.
[70] Grohar, P. J., Woldemichael, G. M., Griffin, L. B., Mendoza, A., Chen, Q. R., Yeung, C., Currier, D. G., Davis, S., Khanna, C., Khan, J., McMahon, J. B., and Helman, L. J. (2011) Identification of an inhibitor of the EWS-FLI1 oncogenic

transcription factor by high-throughput screening, *J Natl Cancer Inst 103*, 962-978.
[71] Delattre, O., Zucman, J., Plougastel, B., Desmaze, C., Melot, T., Peter, M., Kovar, H., Joubert, I., De Jong, P., Rouleau, G., Aurias, A., and Thomas, G. (1992) Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours, *Nature 359*, 162-165.
[72] Yuan, P., Wang, L., Wei, D., Zhang, J., Jia, Z., Li, Q., Le, X., Wang, H., Yao, J., and Xie, K. (2007) Therapeutic inhibition of Sp1 expression in growing tumors by mithramycin a correlates directly with potent antiangiogenic effects on human pancreatic cancer, *Cancer 110*, 2682-2690.
[73] Campbell, V. W., Davin, D., Thomas, S., Jones, D., Roesel, J., Tran-Patterson, R., Mayfield, C. A., Rodu, B., Miller, D. M., and Hiramoto, R. A. (1994) The G-C specific DNA binding drug, mithramycin, selectively inhibits transcription of the C-MYC and C-HA-RAS genes in regenerating liver, *Am J Med Sci 307*, 167-172.
[74] Lee, T. J., Jung, E. M., Lee, J. T., Kim, S., Park, J. W., Choi, K. S., and Kwon, T. K. (2006) Mithramycin A sensitizes cancer cells to TRAIL-mediated apoptosis by down-regulation of XIAP gene promoter through Sp1 sites, *Mol Cancer Ther 5*, 2737-2746.
[75] Koutsodontis, G., and Kardassis, D. (2004) Inhibition of p53-mediated transcriptional responses by mithramycin A, *Oncogene 23*, 9190-9200.
[76] Tagashira, M., Kitagawa, T., Isonishi, S., Okamoto, A., Ochiai, K., and Ohtake, Y. (2000) Mithramycin represses MDR1 gene expression in vitro, modulating multidrug resistance, *Biol Pharm Bull 23*, 926-929.
[77] Saha, S., Mukherjee, S., Mazumdar, M., Manna, A., Khan, P., Adhikary, A., Kajal, K., Jana, D., Sa, G., Sarkar, D. K., and Das, T. (2015) Mithramycin A sensitizes therapy-resistant breast cancer stem cells toward genotoxic drug doxorubicin, *Transl Res 165*, 558-577.
[78] Kulkarni, K. K., Bankar, K. G., Shukla, R. N., Das, C., Banerjee, A., Dasgupta, D., and Vasudevan, M. (2015) Global gene expression profiling data analysis reveals key gene families and biological processes inhibited by Mithramycin in sarcoma cell lines, *Genom Data 3*, 8-14.
[79] Finotti, A., Bianchi, N., Fabbri, E., Borgatti, M., Breveglieri, G., Gasparello, J., and Gambari, R. (2015) Erythroid induction of K562 cells treated with mithramycin is associated with inhibition of raptor gene transcription and mammalian target of rapamycin complex 1 (mTORC1) functions, *Pharmacol Res 91*, 57-68.
[80] Fernandez-Guizan, A., Mansilla, S., Barcelo, F., Vizcaino, C., Nunez, L. E., Moris, F., Gonzalez, S., and Portugal, J. (2014) The activity of a novel mithramycin analog is related to its binding to DNA, cellular accumulation, and inhibition of Sp1-driven gene transcription, *Chem Biol Interact 219*, 123-132.
[81] Cons, B. M., and Fox, K. R. (1990) The GC-selective ligand mithramycin alters the structure of (AT)n sequences flanking its binding sites, *FEBS Lett 264*, 100-104.
[82] Cons, B. M., and Fox, K. R. (1991) Effects of the antitumor antibiotic mithramycin on the structure of repetitive DNA regions adjacent to its GC-rich binding site, *Biochemistry 30*, 6314-6321.

[83] Albertini, V., Jain, A., Vignati, S., Napoli, S., Rinaldi, A., Kwee, I., Nur-e-Alam, M., Bergant, J., Bertoni, F., Carbone, G. M., Rohr, J., and Catapano, C. V. (2006) Novel GC-rich DNA-binding compound produced by a genetically engineered mutant of the mithramycin producer Streptomyces argillaceus exhibits improved transcriptional repressor activity: implications for cancer therapy, *Nucleic Acids Res 34*, 1721-1734.

[84] Remsing, L. L., Gonzalez, A. M., Nur-e-Alam, M., Fernandez-Lozano, M. J., Brana, A. F., Rix, U., Oliveira, M. A., Mendez, C., Salas, J. A., and Rohr, J. (2003) Mithramycin SK, a novel antitumor drug with improved therapeutic index, mithramycin SA, and demycarosyl-mithramycin SK: three new products generated in the mithramycin producer Streptomyces argillaceus through combinatorial biosynthesis, *J Am Chem Soc 125*, 5745-5753.

[85] Scott, D., Chen, J. M., Bae, Y., and Rohr, J. (2013) Semi-synthetic mithramycin SA derivatives with improved anticancer activity, *Chem Biol Drug Des 81*, 615-624.

[86] Sastry, M., Fiala, R., and Patel, D. J. (1995) Solution structure of mithramycin dimers bound to partially overlapping sites on DNA, *J Mol Biol 251*, 674-689.

[87] Sastry, M., and Patel, D. J. (1993) Solution structure of the mithramycin dimer-DNA complex, *Biochemistry 32*, 6588-6604.

[88] Gao, X. L., and Patel, D. J. (1990) Chromomycin dimer-DNA oligomer complexes. Sequence selectivity and divalent cation specificity, *Biochemistry 29*, 10940-10956.

[89] Demicheli, C., Albertini, J. P., and Garnier-Suillerot, A. (1991) Interaction of mithramycin with DNA. Evidence that mithramycin binds to DNA as a dimer in a right-handed screw conformation, *Eur J Biochem 198*, 333-338.

[90] Cons, B. M., and Fox, K. R. (1989) High resolution hydroxyl radical footprinting of the binding of mithramycin and related antibiotics to DNA, *Nucleic Acids Res 17*, 5447-5459.

[91] Hou, M. H., Robinson, H., Gao, Y. G., and Wang, A. H. (2004) Crystal structure of the [Mg2+-(chromomycin A3)2]-d(TTGGCCAA)2 complex reveals GGCC binding specificity of the drug dimer chelated by a metal ion, *Nucleic Acids Res 32*, 2214-2222.

[92] Arauzo-Bravo, M. J., and Sarai, A. (2008) Indirect readout in drug-DNA recognition: role of sequence-dependent DNA conformation, *Nucleic Acids Res 36*, 376-386.

[93] Cons, B. M., and Fox, K. R. (1989) Interaction of mithramycin with metal ions and DNA, *Biochem Biophys Res Commun 160*, 517-524.

[94] Carpenter, M. L., Marks, J. N., and Fox, K. R. (1993) DNA-sequence binding preference of the GC-selective ligand mithramycin. Deoxyribonuclease-I/deoxyribonuclease-II and hydroxy-radical footprinting at CCCG, CCGC, CGGC, GCCC and GGGG flanked by (AT)n and An.Tn, *Eur J Biochem 215*, 561-566.

[95] Hampshire, A. J., and Fox, K. R. (2008) The effects of local DNA sequence on the interaction of ligands with their preferred binding sites, *Biochimie 90*, 988-998.

[96] Fox, K. R., and Howarth, N. R. (1985) Investigations into the sequence-selective binding of mithramycin and related ligands to DNA, *Nucleic Acids Res 13*, 8695-8714.
[97] Aich, P., and Dasgupta, D. (1995) Role of magnesium ion in mithramycin-DNA interaction: binding of mithramycin-Mg2+ complexes with DNA, *Biochemistry 34*, 1376-1385.
[98] Hou, M. H., and Wang, A. H. (2005) Mithramycin forms a stable dimeric complex by chelating with Fe(II): DNA-interacting characteristics, cellular permeation and cytotoxicity, *Nucleic Acids Res 33*, 1352-1361.
[99] Hsu, C. W., Kuo, C. F., Chuang, S. M., and Hou, M. H. (2013) Elucidation of the DNA-interacting properties and anticancer activity of a Ni(II)-coordinated mithramycin dimer complex, *Biometals 26*, 1-12.
[100] Owczarzy, R., You, Y., Moreira, B. G., Manthey, J. A., Huang, L., Behlke, M. A., and Walder, J. A. (2004) Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures, *Biochemistry 43*, 3537-3554.
[101] Heinemann, U., Alings, C., and Bansal, M. (1992) Double helix conformation, groove dimensions and ligand binding potential of a G/C stretch in B-DNA, *EMBO J 11*, 1931-1939.
[102] Eisenstein, M., and Shakked, Z. (1995) Hydration patterns and intermolecular interactions in A-DNA crystal structures. Implications for DNA recognition, *J Mol Biol 248*, 662-678.
[103] Lauble, H., Frank, R., Blocker, H., and Heinemann, U. (1988) Three-dimensional structure of d(GGGATCCC) in the crystalline state, *Nucleic Acids Res 16*, 7799-7816.
[104] Eisenstein, M., Frolow, F., Shakked, Z., and Rabinovich, D. (1990) The structure and hydration of the A-DNA fragment d(GGGTACCC) at room temperature and low temperature, *Nucleic Acids Res 18*, 3185-3194.
[105] McAteer, K., Aceves-Gaona, A., Michalczyk, R., Buchko, G. W., Isern, N. G., Silks, L. A., Miller, J. H., and Kennedy, M. A. (2004) Compensating bends in a 16-base-pair DNA oligomer containing a T(3)A(3) segment: A NMR study of global DNA curvature, *Biopolymers 75*, 497-511.
[106] Mujeeb, A., Kerwin, S. M., Kenyon, G. L., and James, T. L. (1993) Solution structure of a conserved DNA sequence from the HIV-1 genome: restrained molecular dynamics simulation with distance and torsion angle restraints derived from two-dimensional NMR spectra, *Biochemistry 32*, 13419-13431.
[107] Grzeskowiak, K., Yanagi, K., Prive, G. G., and Dickerson, R. E. (1991) The structure of B-helical C-G-A-T-C-G-A-T-C-G and comparison with C-C-A-A-C-G-T-T-G-G. The effect of base pair reversals, *J Biol Chem 266*, 8861-8883.
[108] Heinemann, U., and Alings, C. (1989) Crystallographic study of one turn of G/C-rich B-DNA, *J Mol Biol 210*, 369-381.
[109] Courseille, C., Dautant, A., Hospital, M., Langlois D'Estaintot, B., Precigoux, G., Molko, D., and Teoule, R. (1990) Crystal structure analysis of an A(DNA) octamer d(GTACGTAC) *Acta Crystallogr A46*, FC9-FC12

[110] Huang, D. B., Phelps, C. B., Fusco, A. J., and Ghosh, G. (2005) Crystal structure of a free kappaB DNA: insights into DNA recognition by transcription factor NF-kappaB, *J Mol Biol 346*, 147-160.
[111] Cervi, A. R., Destaintot, B. L., and Hunter, W. N. (1992) Crystal and Molecular-Structure of D(Gtctagac), *Acta Crystallogr B 48*, 714-719.
[112] Hunter, W. N., D'Estaintot, B. L., and Kennard, O. (1989) Structural variation in d(CTCTAGAG). Implications for protein-DNA interactions, *Biochemistry 28*, 2444-2451.
[113] Majee, S., Sen, R., Guha, S., Bhattacharyya, D., and Dasgupta, D. (1997) Differential interactions of the Mg2+ complexes of chromomycin A3 and mithramycin with poly(dG-dC) x poly(dC-dG) and poly(dG) x poly(dC), *Biochemistry 36*, 2291-2299.
[114] Carpenter, M. L., Cassidy, S. A., and Fox, K. R. (1994) Interaction of mithramycin with isolated GC and CG sites, *J Mol Recognit 7*, 189-197.
[115] Lu, X. J., and Olson, W. K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures, *Nucleic Acids Res 31*, 5108-5121.
[116] Chakrabarti, S., Bhattacharyya, D., and Dasgupta, D. (2000) Structural basis of DNA recognition by anticancer antibiotics, chromomycin A(3), and mithramycin: roles of minor groove width and ligand flexibility, *Biopolymers 56*, 85-95.
[117] Goodsell, D. S., Kopka, M. L., Cascio, D., and Dickerson, R. E. (1993) Crystal structure of CATGGCCATG and its implications for A-tract bending models, *Proc Natl Acad Sci U S A 90*, 2930-2934.
[118] Tataurov, A. V., You, Y., and Owczarzy, R. (2008) Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids, *Biophys Chem 133*, 66-70.

7. VITA

Stevi Weidenbach

EDUCATION

Bachelor of Science in Biology, GPA: 3.39/4.0 May 2012

Cellular/Molecular/Genetics Track

Minor in Chemistry

Northern Kentucky University, Highland Heights KY

EXPERIENCE

Graduate Research Assistant August 2012-Present

Primary Investigator Dr. Jürgen Rohr

University of Kentucky, Lexington KY

Undergraduate Research Assistant June 2010 - May 2012

Northern Kentucky University, Highland Heights KY

Research Advisor Dr. Isabelle Lagadic

PUBLICATIONS

Salem, S., **Weidenbach, S.**, Rohr, J. (2017). A putative glycosyltransferase complex is responsible for the sugar diversity of saquayamycins isolated from *Streptomyces* sp. KY 40-1. *Submitted.*

Hou, C., **Weidenbach, S.**, Cano, K. E., Wang, Z., Mitra, P., Ivanov, D. N., Rohr, J. & Tsodikov, O. V. (2016). Structures of mithramycin analogues bound to DNA and implications for targeting transcription factor FLI1. *Nucl Acids Res 44*(18), 8990-9004.

Weidenbach, S., Hou, C., Chen, J. M., Tsodikov, O. V., & Rohr, J. (2015). Dimerization and DNA recognition rules of mithramycin and its analogues. *J Inorg Biochem*, *156*, 40-47.

Weidenbach, S., & Rohr, J. (2014). Identification and characterization of the biosynthetic gene cluster of saquayamycins H and I. *Planta Med, 80*(10), PR5.

Weidenbach, S., & Rohr, J. (2013). Identification of the saquayamycin gene cluster in *Streptomyces* sp. KY40 – 1. *Planta Med, 79*(10), PL20.