2014

INVESTIGATING STRUCTURE AND PROTEIN-PROTEIN INTERACTIONS OF KEY POST-TYPE II PKS TAILORING ENZYMES

Theresa E. Downey
University of Kentucky, theresa.downey@gmail.com

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.

Theresa E. Downey, Student

Dr. Jürgen Rohr, Major Professor

Dr. Jim Pauly, Director of Graduate Studies

INVESTIGATING STRUCTURE AND PROTEIN-PROTEIN INTERACTIONS OF KEY POST-TYPE II PKS TAILORING ENZYMES

DISSERTATION

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

By
Theresa Downey

Pittsburgh, Pennsylvania

Director: Dr. Jürgen Rohr
Professor of Pharmaceutical Sciences
Lexington, Kentucky

2014

ABSTRACT OF DISSERTATION

INVESTIGATING STRUCTURE AND PROTEIN-PROTEIN INTERACTIONS OF KEY POST-TYPE II PKS TAILORING ENZYMES

Type II polyketide synthase (PKS) produced natural products have proven to be an excellent source of pharmacologically relevant molecules due to their rich biological activities and chemical scaffolds. Type II-PKS manufactured polyketides share similar polycyclic aromatic backbones leaving their diversity to stem from various chemical additions and alterations facilitated by post-PKS tailoring enzymes. Evidence suggests that post-PKS tailoring enzymes form complexes in order to facilitate the highly orchestrated process of biosynthesis. Thus, protein-protein interactions between these enzymes must play crucial roles in their structures and functions. Despite the importance of these interactions little has been done to study them. In the mithramycin (MTM) biosynthetic pathway the Baeyer−Villiger monooxygenase (BVMO) MtmOIV and the ketoreductase MtmW form one such enzyme pair that catalyze the final two steps en route to the final product. MtmOIV oxidatively cleaves the fourth ring of the mithramycin intermediate premithramycin B (PreB) via a Baeyer−Villiger reaction, generating MTM's characteristic tricyclic aglycone core and highly functionalized pentyl side chain at position 3. This Baeyer−Villiger reaction precedes spontaneous lactone ring opening, decarboxylation, and the final step of MTM biosynthesis, a reduction of the 4′-keto group catalyzed by the ketoreductase MtmW.

Another example of co-dependent post-PKS tailoring enzymes from the gilvocarcin biosynthetic pathway is composed of GilM and GilR. These two enzymes form an unusual synergistic tailoring enzyme pair that does not function sequentially. GilM exhibits dual functionality by catalyzing the reduction of a quinone intermediate to a hydroquinone and stabilizes O-methylation and hemiacetal formation. GilM mediates its reductive catalysis through the aid of GilR that provides its covalently bound FADH(2) for the GilM reaction, through which FAD is regenerated for the next catalytic cycle. A few steps later, following glycosylation related events unique to each gilvocarcin derivative, GilR dehydrogenates the hemiacetal moiety created by GilM to establish the formation of a lactone and the final gilvocarcin chromophore. To achieve a better understanding of post-type II PKS tailoring enzymes and their protein-protein

interactions for the benefit of future combinatorial biosynthetic efforts two specific aims were devised.

Specific aim 1 was to investigate the structure of MtmOIV and the role of active site residues in its catalytic mechanism.

Specific aim 2 was to integrate the function of GilM and its protein-protein interactions with GilR that lead to their synergistic activity and sharing of GilR's bicovalently bound FAD moiety.

KEYWORDS: Mithramycin, Baeyer-Villiger monooxygenase, *O*-methyltransferase, Gilvocarcin

Theresa Downey

2014

INVESTIGATING STRUCTURE AND PROTEIN-PROTEIN INTERACTIONS OF KEY POST-TYPE II PKS TAILORING ENZYMES

By

Theresa Downey

Dr. Jürgen Rohr
(Director of Dissertation)

Dr. Jim Pauly
(Director of Graduate Studies)

2014

Dedicated to my Grandmother

ACKNOWLEDGEMENTS

I would like to take a moment to acknowledge the individuals who contributed to the successful completion of my dissertation. Without your support and guidance I would have never made it through this journey. Although I am excited about starting the next stage of my career I will always look back fondly and remember all the good (and not so good) times that have shaped me as a scientist and as a human being.

It gives me great pleasure to thank my mentor Dr. Jürgen Rohr for providing me the chance to work, learn, and grow scientifically in his group. I am grateful to him for his expert guidance and insightful discussions throughout my graduate studies. I am also grateful for his allowing me the chance and space to pursue and develop my own scientific ideas. I would also like to thank my committee members Dr. Steven Van Lanen, Dr. Todd Porter, Dr. Joseph Chappell and Dr. Chang-Gou Zhan for their invaluable advice and assistance over the years. I sincerely thank Dr. Douglas Andres for agreeing to serve as an outside examiner for my dissertation defense.

Special thanks to all the past and present members of Dr. Rohr's lab for their support and ideas; Dr. Guojun Wang, Dr. Nidhi Tibrewal, Dr. Daniel Scott, Jhong-Min Chen, Stevi Weidenbach, Dr. Madan Kharel, Dr. Pallab Pahari, Dr. Khaled Shaaban, and Dr. Micah Shepherd. I would especially like to thank Dr. Mary Bosserman for her expert guidance and excellent enzymology training. Additionally, I would like to thank our collaborator Dr. Nick Noinaj for the contribution of his superb structural biology skills.

I would also like to express my gratitude to all my friends and family for their much appreciated love and support. I would especially like to thank Chelsea Geil and Dr.

Dayna Hayes. Without you both graduate school would have been a very dull and frustrating place.

TABLE OF CONTENTS

List of Tables

List of Figures

LIST OF ABBREVIATIONS

ACP	acyl carrier protein
AT	acyl transferase
BLAST	basic local alignment search tool
BVMO/s	Baeyer-Villiger monooxygenase/s
bp	base pairs
CCP4	Collaborative Computational Project 4
CoA	coenzyme A
COSY	correlation spectroscopy
COOT	Crystallographic Object-Oriented Toolkit
CYC	cyclase
CYP450	cytochrome P-450 monooxygenase
DH	dehydratase
DNA	deoxyribonucleic acid
DMSO	dimethylsulfoxide
E. coli	*Escherichia coli*
ER	enoyl reductase
ESI	electrospray ionization
FAD	flavin adenine dinucleotide
FADH/$FADH_2$	flavin adenine dinucleotide, reduced form

FPLC	fast protein liquid chromatography
GT	glycosyltransferase
GM	Gilvocarcin M
GV	gilvocarcin V
IPTG	isopropyl-β-D-thiogalactoside
KS	ketoacyl synthase
KR	keto reductase
LB	lysogeny broth
MCAT	malonyl CoA acyl carrier protein transacylase
MT	Methyltransferase
MTM	Mithramycin
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
NMR	nuclear magnetic resonance
ORF	open reading frame
PCP	peptidyl carrier protein
PCR	polymerase chain reaction

PHENIX	Python-based Hierarchical Environment for Integrated X-ray crystallography
PKS	polyketide synthase
PMB	premithramycin B
PreB	premithramycin B
RMSD	root-mean square deviation
SAR	structure activity relationship
SAM	*S*-adenosyl methionine
SER-CAT	Southeast Regional Collaborative Access Team
TE	thioesterase
UV	ultraviolet

1 Introduction to Bifunctional and Co-dependent Post-PKS Tailoring Enzymes: The New Frontier of Polyketide Biosynthesis

Natural products have proven an excellent and limitless source for drug discovery. In particular polyketides are well suited for drug discovery applications due to their rich chemical scaffolds and diverse biological activities. This broad group of natural products is characterized by their method of biosynthesis consisting of multiple decarboxlative Claisen condensations of malonyl-CoA derived extender units by a group of enzymes collectively known as polyketide synthases (PKSs). Bacterial polyketides can be further categorized into type I, type II, or type III PKS derived compounds based on the type of PKS responsible for their biosynthesis. Type I PKSs are non-iterative large highly modular proteins with multiple acyl carrier proteins (ACPs) passing along the growing polyketide. Type II PKSs are composed of aggregates of iteratively acting monofunctional proteins in which one ACP carriers the growing polyketide chain. Type III PKSs are much less complex ketosynthase homodimers that preform only a few elongation steps, adding CoA-activated building blocks to non-polyketide starter units (figure 1). Once the polyketide backbone has been synthesized they are diversified through a plethora of possible post-PKS tailoring events most commonly including glycosylations, oxygenations, reductions, cyclizations, aromatizations, methylations, and epimerizations (figure 1). The events carried out by post-PKS tailoring enzymes are responsible for the majority of the diversity in chemical structure seen within the polyketide family. Prior to these tailoring events it is rare that unaltered parent scaffolds produced directly from a PKS have any biological activity. Thus, post-PKS tailoring enzyme modifications can drastically impact a compounds chemical and biological properties making them lucrative targets for combinatorial biosynthetic efforts in the search for new medically relevant molecules. Contrary to the previously recognized belief that post-PKS enzymes are single function free floating cytosolic proteins it has become increasingly evident that many post-PKS tailoring enzymes exhibit remarkable multi-functionality or function as part of a multi-enzyme complex or at least a co-dependent enzyme pair. Thus, protein-protein interactions between these enzymes must

play crucial roles in their structures and functions. Despite the importance of these interactions little has been done to study them. If fact, many researchers do not immediately recognize that they are dealing with a multifunctional or co-dependent enzyme. Although countless reviews have been published on specific polyketides, polyketide tailoring enzymes, and polyketide synthases themselves this review is the first to focus on particularly interesting examples of post-PKS tailoring enzyme interdependence and bifunctionality spanning the literature from 2002-2014. This review's purpose is to function as source of enlightenment for those who unknowingly may be working with such complex enzyme systems and as a source of inspiration for what needs to be done in future research of post-PKS tailoring enzymes to utilize these systems for combinatorial biosynthetic applications.

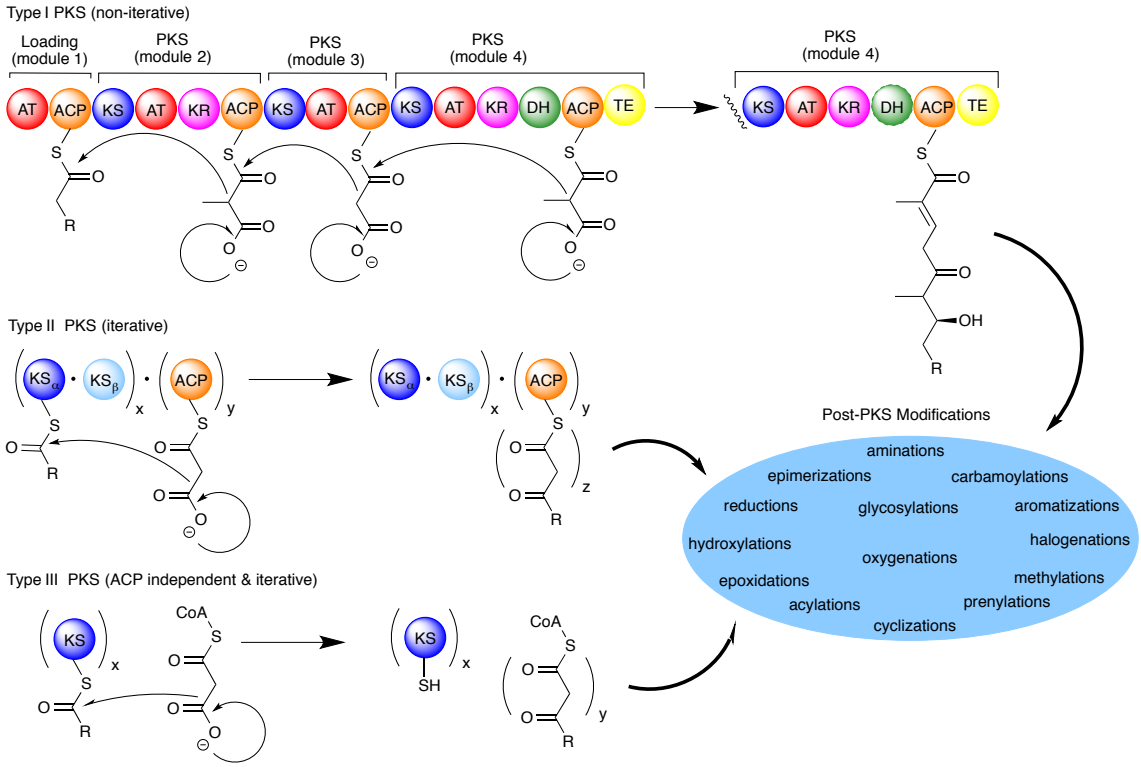


Figure 1. Types and mechanisms of polyketide synthases and common post-PKS modifications.

1.2 Bifunctional Post-PKS Tailoring Enzymes

Bifunctional post-PKS tailoring enzymes have been recognized in the literature as far back as 1992.[1] Generally speaking bifunctional post-PKS tailoring enzymes can exhibit two completely unique functionalities carried out by distinct domains or they can preform similar functionalities on different natural substrates within the same biosynthetic pathway. For the purpose of this discussion these multifunctional enzymes can be grouped into two major categories: Iteratively functioning and non-iteratively functioning.

1.2.1 Iteratively Functioning Post-PKS Bifunctional Tailoring Enzymes

Iteratively functioning post-PKS bifunctional tailoring enzymes can be defined as post-PKS tailoring enzymes that function multiple times consecutively within the same biosynthetic pathway.

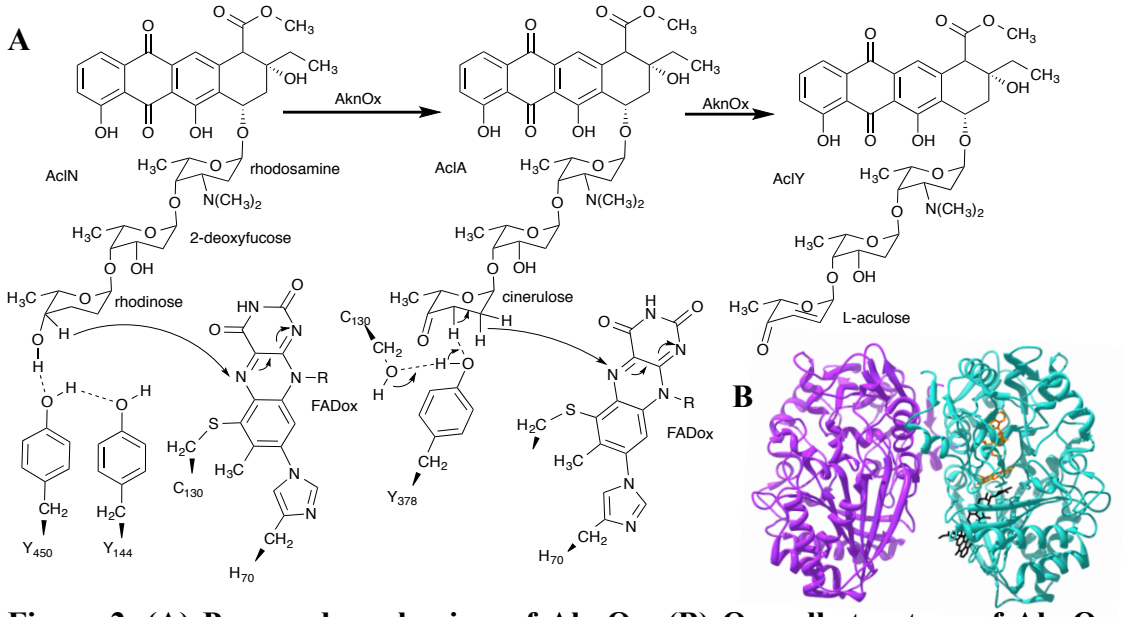


Figure 2. (A) Proposed mechanism of AknOx. (B) Overall structure of AknOx homodimer. FAD is shown in yellow and the product AclY is shown in black in the right monomer.

1.2.1.1 AknOx

Aclacinomycin (Acl) oxidoreductase (AknOx) catalyzes the final two steps of

Acl group polyketide antibiotic biosynthesis in which a terminal sugar moiety is oxidized from rhodinose to L-aculose. AknOx is a secreted flavin dependent enzyme that was first purified from *Streptomyces galilaes* MA144-M1 (ATCC31133).[2] What is interesting about this enzyme is that it utilizes different residues within the same active site to catalyze two consecutive FAD-dependent reactions using the same bicovalently bound cofactor (figure 2A). The X-ray crystal structure (PDB 2IPI) reveals AknOx to have two distinct structural domains, the FAD binding F domain and the substrate binding S domain. The F domain consists of two α + β subdomains while the S domain is composed of a seven-stranded antiparallel β-sheets flanked on one side by four α-helices (figure 2B). During its catalytic mechanism (figure 2A) AknOx first converts the rhodinose sugar of AclN to cinerulouse by oxidizing the hydroxyl group at position C-4 to a keto group yielding AclA. It is theorized based on structural and site directed mutagenic studies that residue Tyr-450 removes a proton from rhodinose's C-4 hydroxyl group. This proton is then transferred to the solvent via a proton transfer system that includes residue Tyr-144. Simultaneously, hydride transfer from the same C-4 position of the sugar moiety to the N-5 position of FAD facilitates formation of the C-4 keto group yielding the cinerulose-containing product AclA. The FAD moiety is reoxidized via molecular oxygen and is then ready to perform its oxidative role in the next reaction. AknOx's second functionality converts the AclA's cinerulose sugar to L-aculose and yields AclY as a final product (figure 2A). To facilitate this catalytic functionality it is hypothesized again based on structural and site directed mutagenic studies that residue Tyr-378 initiates this second reaction by abstracting a proton from position C-3 of cinerulose, which is then transferred to the solvent by a relay system including residues Ser-376 and Glu-374. Next, hydride transfer from position C-2 of the carbohydrate to N-5 of FAD results in the C—C double bond formation that produces L-aculose and thus creates the final product AclY.[3]

1.2.1.2 AurH

The cytochrome P450 monooxygenase AurH from the aureothin biosynthetic

pathway is responsible for furan ring formation via two consecutive C—O bond formations (figure 3A). Based on bioinformatic analysis AurH only contains a single heme binding motif indicating that both C—O bond formations must be catalyzed by

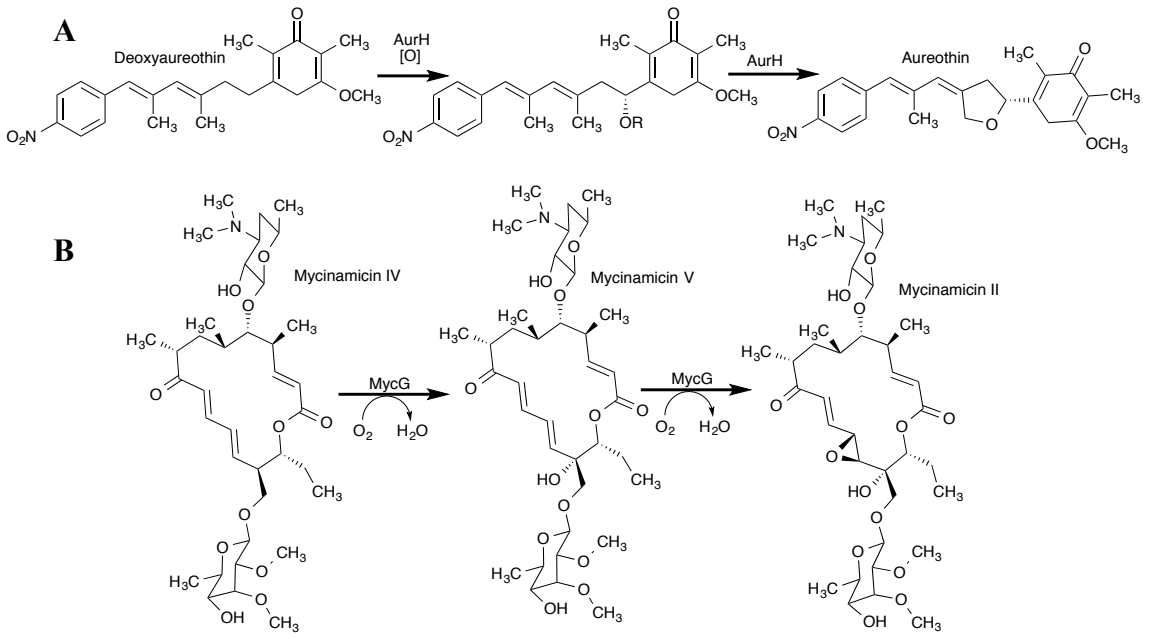


Figure 3. (A) AurH catalyzed reactions. (B) MycG catalyzed reactions.

the same active site. This is conceivable as both of the oxidized carbons are located in such a close proximity that it would only require a slight change in the orientation of the bound substrate to facilitate catalysis. In the first stage of catalysis the methylene (C-7) adjacent to the pyrone ring is attacked setting the stereochemistry of the resulting heterocycle. This is followed by oxidation at C-9a resulting in furan ring closure. AurH is the first example of a cytochrome P450 enzyme catalyzing a two-step cyclization.[4]

1.2.1.3 MycG

MycG is a P450 monooxygenase from biosynthetic pathway of the 16-membered ring macrolide antibiotic mycinamicin II. Characterization of MycG determined that this enzyme catalyzes both the terminal hydroxylation and epoxidation events that lead to the final natural product (figure 3B).[5,6] MycG's first catalytic task is a hydroxylation at C-14 of the macrolide converting the substrate

mycinamicin IV to mycinamicin II. Next, MycG performs an epoxidation converting the double bond between C-12 and C-13 into an epoxide. The enzyme itself has the typical P450 scaffold that is known to accommodate substrates of broad structural variety. Based on crystallographic and paramagnetic NMR relaxation data it was found that a sequential translocation of the substrate molecule from recognition site to the heme catalytic site is most likely responsible for MycG's ability to perform both of its enzymatic functions.[7]

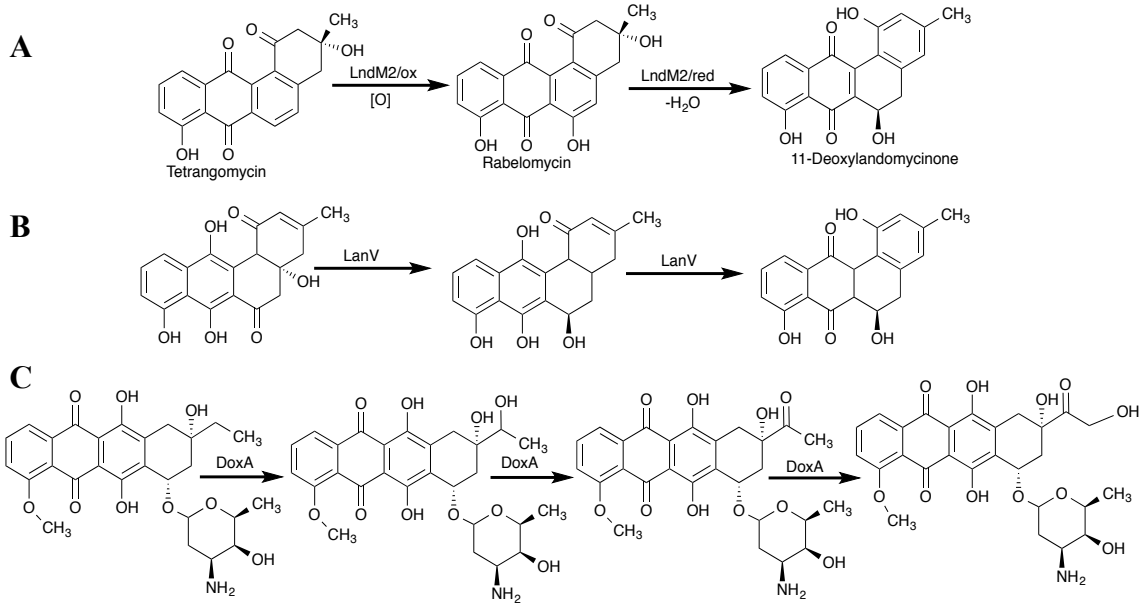


Figure 4. (A) LndM2 catalyzed reactions. (B) LanV catalyzed reactions. (C) DoxA catalyzed reactions.

1.2.1.4 LndM2

LndM2 is a bifunctional post-PKS oxygenase-reductase from the landomycin E biosynthetic pathway. Through a combination of inactivation and feeding experiments it was determined that LndM2 is responsible for both the conversion of tetrangomycin to rabelomycin and rabelomycin to 11-deoxylandomycinone (figure 4A). First LndM2 attaches an oxygen atom as a hydroxyl group at the C-6 position of the tetracyclic polyketide core producing rabelomycin from tetrangomycin. This is immediately followed by a reduction event creating the pathway intermediate 11-deoxylandomycinone that can then continue through the biosynthetic pathway to

landomycin E. These two steps of biosynthesis were found to be crucial for the dearomatization of the B-ring of the aglycone, an essential feature for recognition by the first glycosyltansferase in landomycin E biosynthetic pathway.[8]

1.2.1.5 LanV

LanV is a bifunctional aromatase and ketoreductase from the early portion of the landomycin A biosynthetic pathway. Utilizing gene inactivation experiments it was found that LanV is first responsible for a 6-ketoreduction followed by a 5,6-dehydration that leads to the aromatization of the A ring of the aglycone core (figure 4B). Without this aromatization of the A ring the landomycin A intermediate is not recognized by the first glycosyltansferase (LanGT2) and is therefore extremely important to this polyketides biosynthesis.[9]

1.2.1.6 DoxA

DoxA is a P450 monooxygenase from the doxorubicin biosynthetic pathway and pushes the limits of bifuntionality to its extreme catalyzing three steps in the polyketides biosynthetic pathway.[10,11] DoxA is responsible for the final three steps of doxorubicin's biosynthesis in which the enzyme is responsible for three consecutive oxidative transformations utilizing NADPH and O_2 (figure 4C). DoxA first catalyzes a hydroxylation at C-13 yielding 13-dihydrodaunorubicin from 13-deoxydaunorubicin. Next the enzyme catalyzes a second hydroxylation at either C-13 or C-14 that ultimately results in a spontaneous dehydration and rearrangement to a ketone moiety at C-14 producing daunorubicin. DoxA then hydroxylates daunorubicin a final time at C-14 yielding the final product doxorubicin.[10]

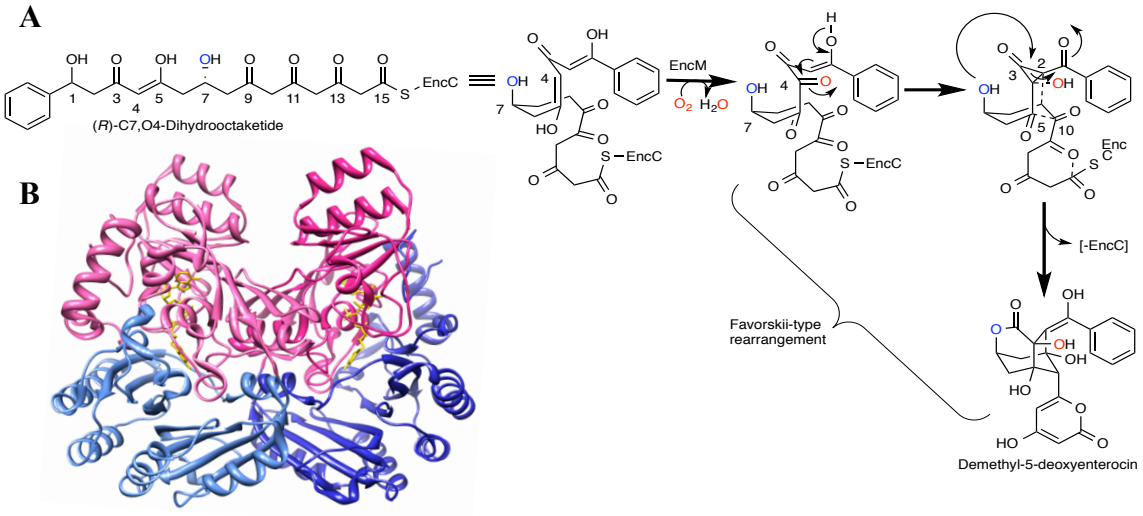


Figure 5. (A) EncM C-4 dual oxidation and rare Favorskii-type rearrangement from the enterocin biosynthetic pathway. (B) Structure of EncM homodimer with chain A shown in a lighter shade than chain B. The FAD binding domain is shown in pink, the substrate binding domain in shown in blue, and the FAD cofactor is shown in yellow.

1.2.1.7 EncM

EncM is an FAD-dependent 'favorskiiase' flavoprotein from the enterocin antibiotic biosynthetic pathway responsible for the atypical polyketide aromatization of its poly(β-carbonyl) chain into a unique tricyclic caged core. EncM accomplishes this task through a rare oxidative Favorskii-type rearrangement.[12,13] *In vivo* characterization of the *encM* gene via mutagenesis and heterologous biosynthesis proved this enzyme to be solely responsible for oxidative C—C rearrangement, two aldol condensations, and two heterocycle forming reactions creating a total of five chiral centers and four rings in the final enterocin chemical structure (figure 5A).[13] EncM was found to exist as a homodimer each with distinct domains for substrate and FAD binding (figure 5B). The FAD binding domain sequesters the flavin at the interface between the two domains positioning the reactive isoalloxazine moiety in optimal proximity for interaction with the polyketide substrate. Additionally, this FAD moiety is bound to EncM by a covalent bond between residue His-78 and the C-

8 methyl group of the flavin's isoalloxazine ring system. Utilizing co-crystallization with the natural substrate, a substrate analogue, and site directed mutagenesis studies it was revealed that the terminal benzene group sits at the hydrophobic end of a long tunnel and forms aromatic–aromatic interactions with Tyr-150 and Trp-152 and van der Waals interactions with Leu-357. Additionally, it was found likely that the C-1 enol engages in hydrogen bonding with O-4 of the flavin and that the C-3 ketone twists away from the flavin and may accept a hydrogen bond from the side chain of Glu-355 and possibly from Tyr-249. The orthogonal sections of the bent EncM ligand-binding pocket separate the C-1—C-6 triketide head from the C-8—C-15 pantothenate-linked tetraketide tail to uncouple the reactivity of the entire C-1—C-16 poly(β-carbonyl) chain preventing kinetically favourable but unwanted cyclization–aromatization reactions and forcing the less favourable Favorskii-type rearrangement (figure 5A). Based on structural data, EncM's proposed mechanism is as follows: EncM first preforms two iterative oxidations at C-4 of substrate (R)-C7,O4-dyhydroctaketide effectively converting a 1,3-diketone to a 1,2,3-triketone and setting up C4 for electrophilic cyclization. This cyclization then triggers the rare Favorskii-like rearrangement. Finally, EncM catalyzes two aldol condensations between C-2 and C-9 and between C-10 and C-5, along with heterocycle formation (order unknown) to release the final product demethyl-5-deoxyenterocin.[14]

1.2.1.8 FkbD

FkbD is a cytochrome P450 hydroxylase from the terminal steps of the FK506 macrocyclic polyketide biosynthetic pathway. Although, the exact sequence of events in this pathway remains elusive it is know that this enzyme functions twice iteratively either before or after 31-*O*-methylation catalyzed by FkbM (figure 6). It has also been suggested that this conversion can happen simultaneous in either order making FkbD quite promiscuous in terms of substrate flexibility. FkbD catalyzes consecutive hydroxylation and further oxidation to a ketone at the C-9 position of the parent compound 9-deoxy-31-*O*-demethyl-FK506. In conjunction with FkbM the final product of FK506 is then produced (figure 6).

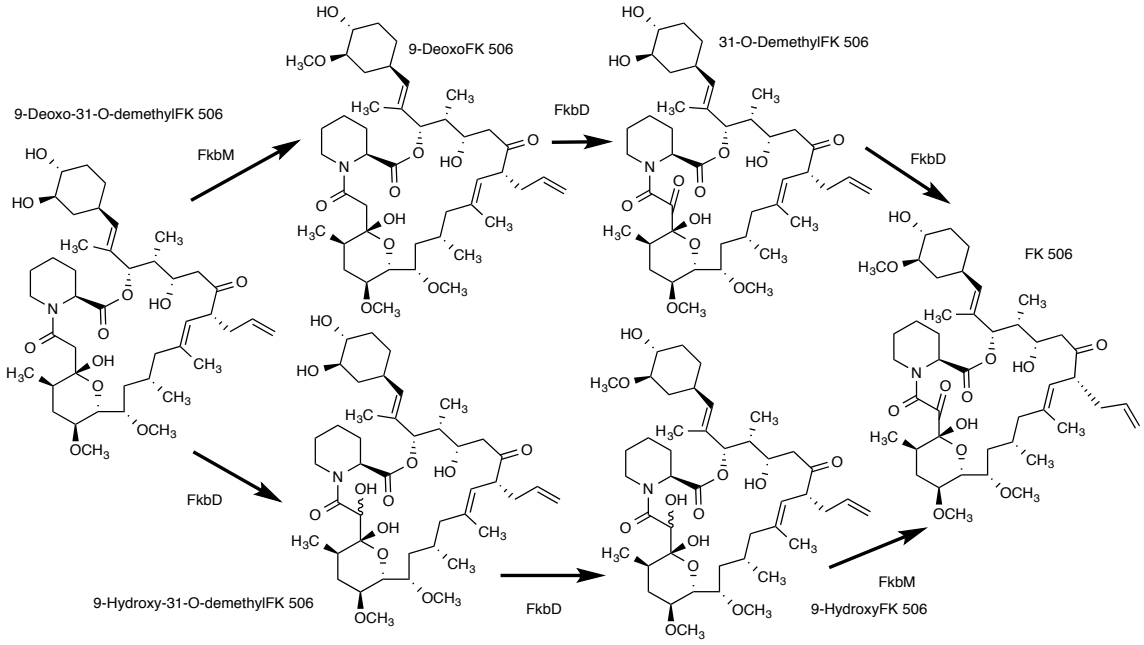


Figure 6. Proposed FkbD catalytic events from the FK506 biosynthetic pathway

1.2.2 Non-iteratively Functioning Post-PKS Bifunctional Tailoring Enzymes

Non-iteratively functioning post-PKS bifunctional tailoring enzymes can be defined as post-PKS tailoring enzymes that function multiple times at different non-consecutive points within the same biosynthetic pathway.

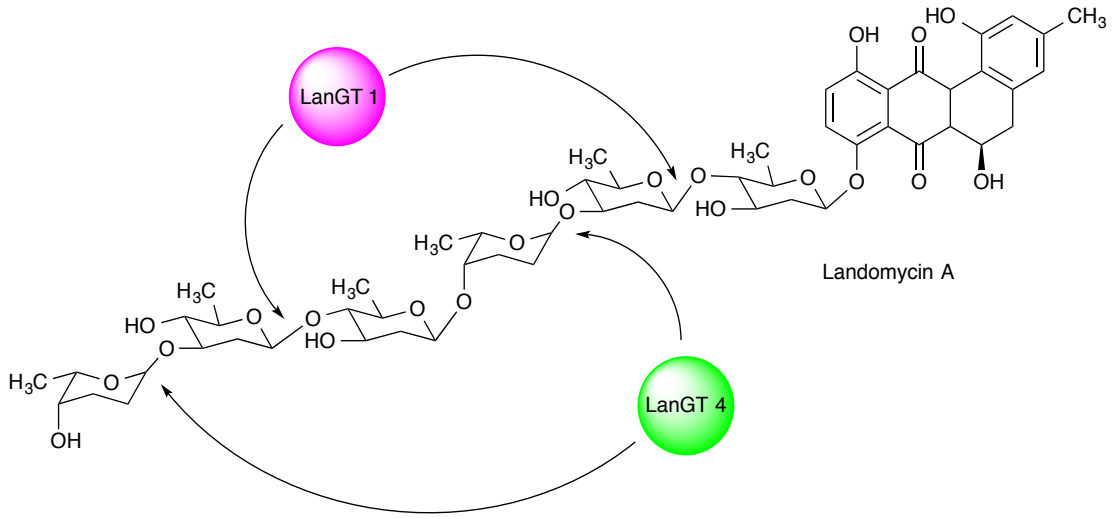


Figure 7. Landomycin A with glycosylation events catalyzed by LanGT1 (pink) and LanGT4 (green) highlighted.

1.2.2.1 LanGT1

LanGT1 is a glycosyltansferase from the landomycin biosynthetic pathway. More specifically, LanGT1 is a D-olivosyltransferase responsible for linking a second D-olivose to the 4-OH group of another D-olivose.[17] LanGT1's first function in landomycin biosynthesis is to add the second sugar to the growing saccharide side chain. It then functions again after several other glycosylations catalyzed by different glycosyltransferases to add the fifth sugar in side chain (figure 7B). Landomycin A has a hexasaccharide side chain that is important to its biological activity and any alterations in this side chain noticeably change the compound's action, thus the function of LanGT1 is crucial to the compounds biosynthesis. It is not usual to see glycosyltransferases act multiple times non-consecutively suggesting that this glycosyltansferase may have relaxed substrate specificity and could be of use in the combinatorial biosynthesis of new biologically active compounds.[18]

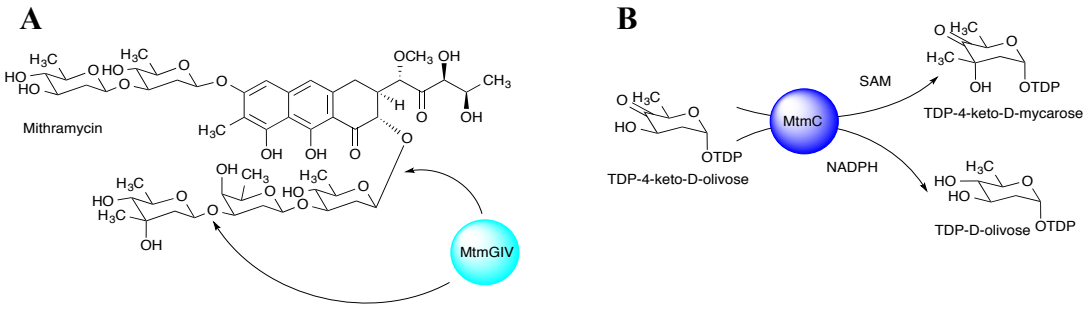


Figure 8. (A) MtmGIV catalyzed glycosylation from the mithramycin biosynthetic pathway. (B) Dual reactions catalyzed by MtmC from the mithramycin biosynthetic pathway.

1.2.2.2 LanGT4

LanGT4, like the previously discussed LanGT1, is a glycosyltansferase from the landomycin biosynthetic pathway. LanGT4 differs in that it is a L-rhodinosyltransferase and is responsible for attaching L-rhodinose to the 3-OH group of a D-olivose.[17] First, LanGT4 attaches the third L-rhodinose sugar moiety to the growing saccharide side chain. Then, after two more glycosylations by LanGT3 and LanGT1 respectively, LanGT4 adds the final L-rhodinose to the side chain yielding the final product landomycin A (figure 7B). This enzyme is of interest for the same reasons previously mentioned for LanGT1.[18]

1.2.2.3 MtmGIV

MtmGIV is a glycosyltansferase from the mithramycin biosynthetic pathway. This enzyme's first role in mithramycin's biosynthesis is to create the linkage of the first sugar moiety, D-olivose, of what will become the trisaccharide side chain to the 12a-position of premithramycinone creating premithramycin A1 (figure 8A).[19,20] MtmGIV's second action in the mithramycin biosynthetic pathway is to attach the third sugar, D-mycarose, to complete the compounds trisaccharide side chain (figure 8A). MtmGIV also was found to be capable of accepting TDP-D-quinovose, TDP-D-

digitoxose, and TDP-4-keto-D-digitoxose as substrates for its first functionality creating the new mithramycins D-quinovosylpremithramycinone, D-digitoxosylpremithramycinone, and 4-keto-D-digitoxosylpremithramycinone respectively. Interestingly, this range of substrate promiscuity was not found to be the same for MtmGIV's second functionality. It is important to note that MtmGIV is the first glycosyltansferase found to have two unique donor and acceptor substrates. This distinctive feature makes MtmGIV an enticing target for future enzyme engineering towards the creation of new mithramycin analogues.[19]

1.2.2.4 MtmC

MtmC is a reductase/methyltransferase from the deoxysugar portion of the mithramycin biosynthetic pathway. Remarkably, this enzyme can utilize one substrate to catalyze reactions creating two different TDP activated sugar moieties. During MtmC's reductase functionality the enzyme reduces its substrate TDP-4-keto-D-olivose to TDP-D-olivose by reducing the 4-keto group to a hydroxyl group with the aid of the cofactor NADPH. During MtmC's methyltransferase functionality the enzyme methylates the same substrate, TDP-4-keto-D-olivose, at the C-3 position using SAM as a methyl donor to yield TDP-4-keto-D-mycarose. How this enzyme dictates which catalytic event a substrate will be subjected to somehow involves its interaction with MtmGIV. This interaction will be discussed in more detail in section 1.3.2.2 of this review.[19]

1.2.2.5 Anm21

Anm21 is a dual functioning carbamoyltransferases from extended ansamitocin biosynthesis responsible for the creation of 4"-carbamoyl ansamitocinoside P-3. Through *in vivo* and *in vitro* validation this enzyme was found to be not only responsible for a carbamoylation at the C-7 position of the polyketide macrolactam backbone, but also was found to be responsible for the carbamoylation

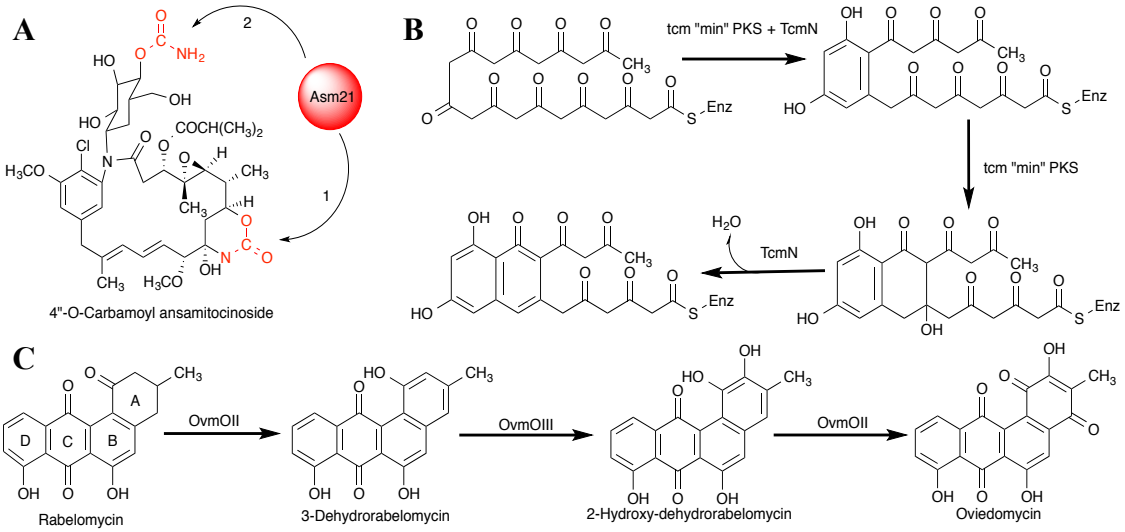


Figure 9. (A) Carbamoylations catalyzed by Asm21 from extended ansamitocin biosynthesis. Numbering denotes the order in which the carbamoylations occur. (B) TcmN functions in the early post-PKS steps of the tetracenomycin biosynthetic pathway. (C) Enzymatic activity of OvmII from the oviedomycin biosynthetic pathway.

of the C-4 hydroxyl group of the *N*-glycosyl moiety in ansamitosinoside P-3 (figure 9A). Interestingly, Asm21 was found to contain about 100 additional amino acids at it C-terminus when compared to other carbamoyltransferases from various antibiotic biosynthetic pathways. Additionally, the enzyme was able to accept a variety of ansamitocinoside structures as substrates demonstrating very high substrate flexibility. Due to this this broad substrate range Asm21 could theoretically be used to develop a variety of ansamitocin drug candidates.[21]

1.2.2.6 TcmN

TcmN is a dual functioning enzyme from the early post-PKS modification portion of the tetracenomycin biosynthetic pathway. TcmN's first function in tetracnomycin's biosynthesis is to dictate the regiospecificity of the first ring cyclization. In this role TcmN assures that an intermolecular aldol condensation occurs between the C-9 carbonyl and the C-14 methylene. If TcmN is not present this reaction can occur one of two ways. The first of which is between the C-9 carbonyl and the C-14 methylene as previously mentioned and the second is between the

between the C-7 carbonyl and the C-12 methylene group (figure 9B). This implicates that TcmN functions to determine regiospecificity but does not preform the aldol condensation itself. TcmN's second functionality is that of an aromatase aromatizing the second ring of the polyketide intermediate shown in figure 9B.[22] Additionally, the C-terminal portion of TcmN has been implicated in the post-PKS *O*-methylation of a tetracenomycin C precursor although this function has never been unequivocally verified.[1]

1.2.2.7 OvmOII

OvmOII is a dehydratase/oxygenase form the oviedomycin biosynthetic pathway (figure 9C). This enzymes activity has thus far been deduced via *in vivo* methods. OvmOII's first catalytic function is a 2,3-dehydration event that aromatizes the A ring of the polyketide creating 3-dehydrorabelomycin from rabelomycin. Next, OvmOIII preforms a 2-hydroxylation yielding 2-hydroxy-dehydrorabelomycin in preparation for OvmII's second functionality. OvmII then converts 2-hydroxy-dehydrorabelomycin to the final product oviedomycin via a 4-oxidation.[23]

1.3 Co-dependent Post-PKS Tailoring Enzymes

Until recently, it was commonly accepted the post-PKS tailoring enzymes were independently functioning. Recently, this has increasingly proven not to be the case. Co-dependent enzymes rely on their partners in one or more ways such as by providing structural support, providing cofactors, and through sequestration of highly reactive intermediates from solvent. For the purpose of this discussion, these co-dependent enzymes can be grouped into two major categories: consecutively functioning and non-consecutively functioning.

1.3.1 Consecutively Functioning Co-dependent Post-PKS Tailoring Enzymes

Consecutively functioning co-dependent post-PKS enzymes can be defined as post-PKS enzymes that require one or more enzyme partners that each function consecutively in a polyketide natural product's biosynthetic pathway.

1.3.1.1 LanZ4-LanZ5

LanZ4 and LanZ5 are a co-dependent enzyme pair from the landomycin biosynthetic pathway. LanZ4 is a FMN-NADPH reductase while LanZ5 is an oxygenase. LanZ4 supports LanZ5 by providing the reduced $FMNH_2$ cofactor required for catalysis. The major collaborative activity of this enzyme pair is the conversion of landomycin H to landomycin I via a hydroxylation at C-11. Landomycin I can then be subjected to various glycosylation events to create the major products landomycins A, B, D, and E (figure 10A). In order for LanZ4-LanZ5 to function appropriately, LanGT2 must glycosylate the aglycone prior to their activity. If glycosylation does not occur LanZ4-LanZ5 creates the minor products of anhydrolandomycinone and tetrangulol from 11-deoxylandomycinone that can go on to be glycosylated to create the minor products landomycin M, O, P, R, T, U, and W (Figure 10B).[24]

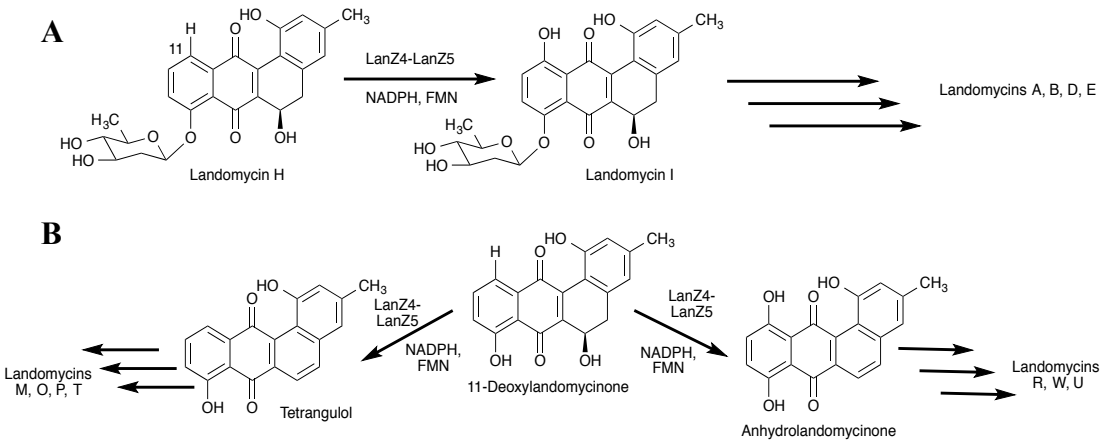


Figure 10. (A) Major pathway of LanZ4-LanZ5 from landomycin biosynthesis. (B) Minor pathways of LanZ4-LanZ5 from landomycin biosynthesis.

1.3.1.2 TamI-TamL

TamI and TamL are an oxidative co-dependent enzyme pair from the tirandamycin biosynthetic pathway. TamI is a P450 enzyme that catalyzes at least two hydroxylations and an epoxidation at three distinct sites during the conversion of tirandamycin C (TirC) to tirandamycin B (TirB) (figure 11A). The TamI/TamL

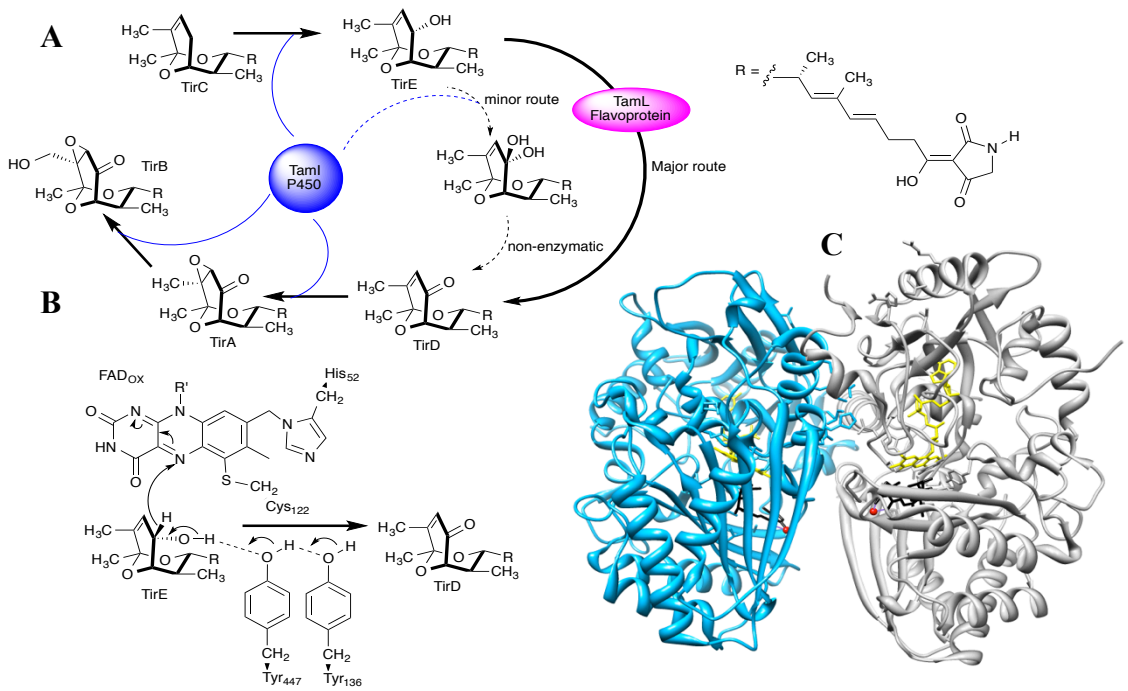


Figure 11. (A) Complete oxidative cascade of TamI and TamL catalyzed reaction from the tirandamycin biosynthesis. (B) Mechanism of dehydrogenation at C-10 of TirE by TamL. (C) Ribbon representation of TamL homodimer with one monomer shown in blue and the other shown in grey (PDB ID 2Y3S). FAD cofactor is shown in yellow, Mg^{2+} is shown in red, and TirE is shown in black.

catalyzed cascade occurs in a precisely defined order without any detectable promiscuity. The cascade starts when TirC is oxidized at position C-10 to tirandamycin E (TirE) by TamI. TamL then oxidizes this hydroxyl group at C-10 to a ketone to produce tirandamycin D (TirD). Next, TamI epoxidates TirD at positions C-11/C-12 yielding tirandamycin A (TirA). To complete the cascade, TirI hydroxylates position C-18 to produce TirB. This unique mechanism was further investigated through the crystallization and X-ray structural determination of TamL alone (PDB ID 2Y08) and the TamL-tirandomycin complex (PDB IDs 2Y3R, 2Y4G, and 2Y3S). Together, these structures reveal that TamL functions as a homodimer in which the tirandomycins are bound in the active site with the bicyclic ketal group facing the FAD cofactor and the tetramic acid moiety extended toward the mouth of the substrate binding cleft (figure 11C). Additionally, the FAD cofactor was shown to be bicovalently bound by His-52 and Cys-122. Based on residue positioning in the active

site, it is hypothesized that TamL oxidation is likely to proceed via proton extraction from the C-10 hydroxyl group of the substrate TirE by Tyr-447 and simultaneous hydride transfer from the C-10 carbon to the N-5 nitrogen of FAD. Tyr-136 is also thought to precipitate in this proton transfer network by initiating the proton abstraction from Tyr-447 (figure 11B).[25] It is known that the oxidative modifications of the tirandamycin biosynthetic pathway are key to antibiotic potency,[26] which may have created evolutionary pressure toward the development of the TamI/TamL system.

1.3.1.3 PgaE-PgaM

PgaE and PgaM are a co-dependent sequentially functioning pair of enzymes from the early steps of angucycline biosynthesis. PgaE functions as a hydroxylase while PgaM functions as a reductase and as an oxygenase in this pathway. PgaE first catalyzes a hydroxylation of UMW6 at position C-12 followed by quinone formation. Next, PgaE passes the substrate to PgaM where the reductase domain preforms a 2,3-dehydration reaction. PgaM's oxygenase domain then preforms a C-12b hydroxylation utilizing water as an oxygen donor. This cascade is completed by a stereospecific ketoreduction at position C-6 also catalyzed by PgaM. Mechanistically, the oxidative reactions of PgaM are thought to proceed through a highly polarized quinone methide intermediate (o-QM) (figure 12A).[27]

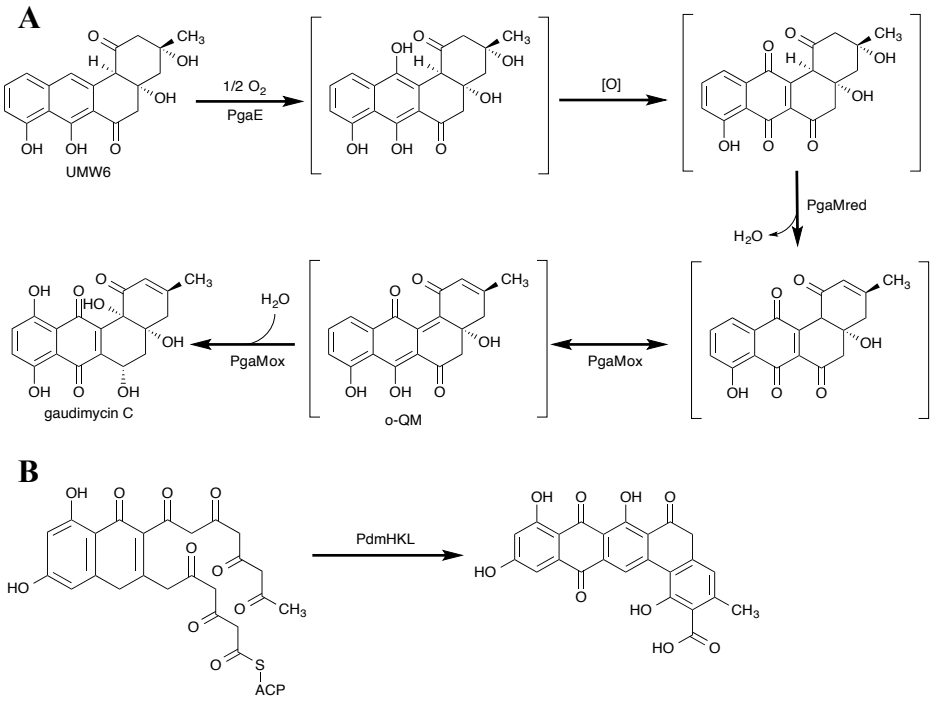


Figure 12. (A) PgaE and PgaM catalyzed reactions from gaurdimycin C biosynthesis. (B) Reactions catalyzed by the PdmH-PdmK-PdmL trio of enzymes from pradimicin A biosynthesis.

1.3.1.4 PdmH-PdmK-PdmL

PdmH, PdmK, and PdmL are a synergistic trio of co-dependent enzymes from the pradimicin A biosynthetic pathway. PdmH is a monooxygenase while PdmK and PdmL are cyclases. It was originally predicted that each of these enzymes could work independently, but following *in vivo* studies it became clear that upon removal of any one of these genes from the cluster the pradimicin aglycone was no longer produced. These surprising results lead to the conclusion that PdmH, PdmK, and PdmL form a complex that oxidizes the cyclized-B ring and cyclizes portions of the polyketide substrate as shown in figure 12B. PdmH is theorized to be positioned in close proximity to the B-ring to hydroxylate C-8, while PdmK and PdmL are likely positioned near the two ends of the polyketide chain to preform the aldol condensations between C-5 and C-17 and between C-4 and C-21 yielding the product

shown in figure 12B.[28]

1.3.1.5 MtmOIV-MtmW

MtmOIV and MtmW are a pair of co-dependent enzymes from the end of the mithramycin biosynthetic pathway. MtmOIV is a Baeyer-Villeger monooxygenase responsible for creating the lactone intermediate and subsequent ring cleavage to mithramycin DK from premithramycin B shown in figure 13A.[29,30] MtmW is a ketoreductase that reduces the 4' ketone of mithramycin DK's pentyl side chain yielding the final product of mithramycin.[31] Based on *in vitro* and *in vivo* studies the intermediate mithramycin DK rapidly degrades to mithramycin SK, SDK, and SA in solution (figure 13A). Because of this rapid degradation it can only be concluded that MtmOIV and MtmW must be in interacting closely to enable the passing of mithramycin DK to MtmW. MtmOIV has been studied much more extensively then MtmW because despite multiple attempts MtmW has yet to be produced heterologously in soluble form. MtmOIV's structure has been solved with and without substrate (premithramycin B) as shown in figure 13B.[29,32,33] This structure is one of only a few Baeyer-Villeger monooxygenases and is the most complex in terms of the enzyme and substrate structures know to date.[32] MtmOIV functions as a homodimer and is FAD and NADPH dependent. Prior to this enzymes activity, none of the tetracyclic premithramycin intermediates are biologically active making MtmOIV key to this natural product's biosynthesis.[29,30,34]

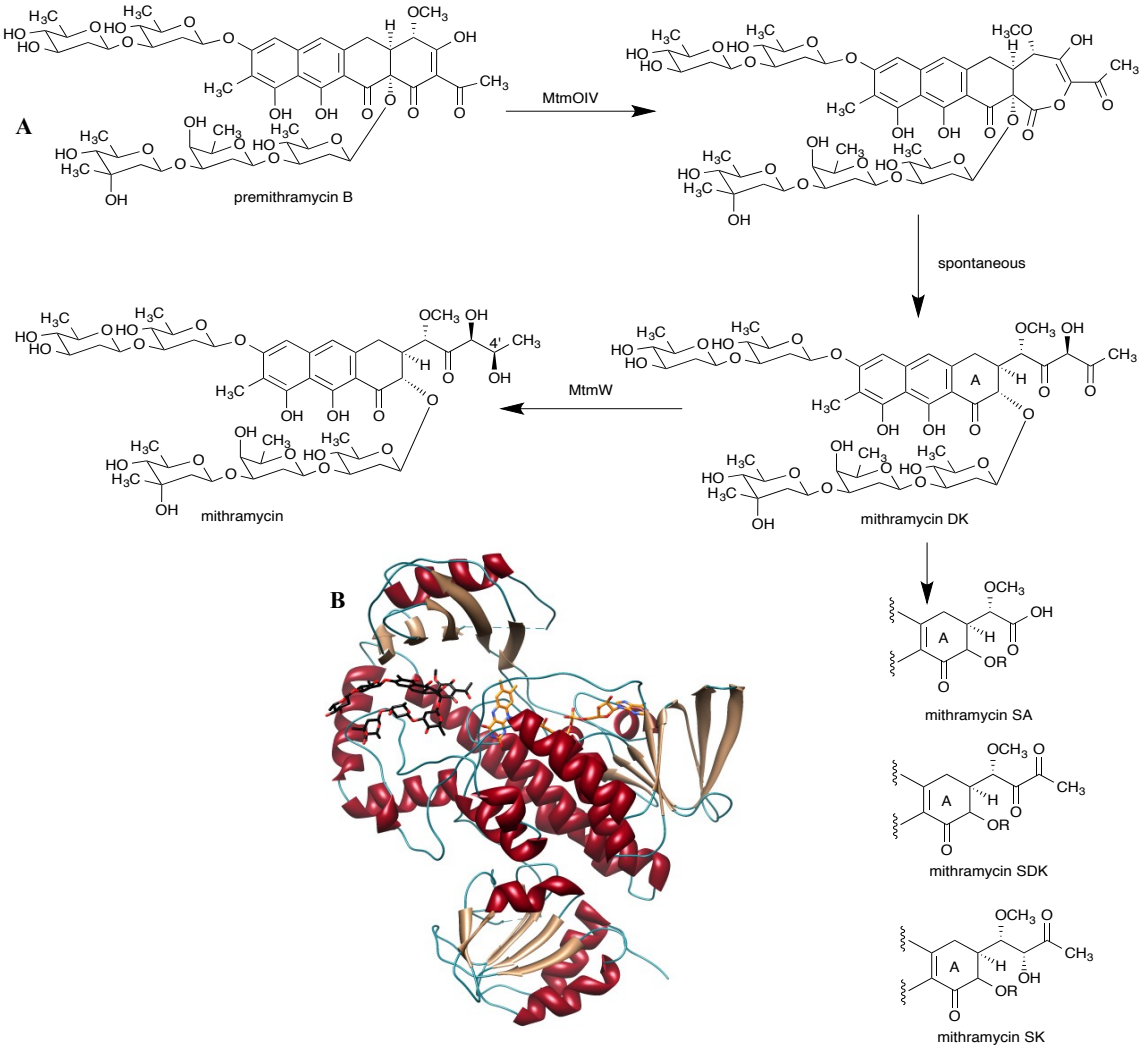


Figure 13. (A) MtmOIV and MtmW catalyzed reactions from the mithramycin biosynthetic pathway. Overall structure of MtmOIV monomer in complex with premithramycin B (black) and FAD (orange). Alpha helices are shown in red while beta sheets are shown in beige.

1.3.2 Non-consecutively Functioning Co-dependent Post-PKS Tailoring Enzymes

Non-consecutively functioning co-dependent post-PKS enzymes can be defined as post-PKS enzymes that require one or more enzyme partners that each function at different non-consecutive points in a polyketide natural product's biosynthetic pathway.

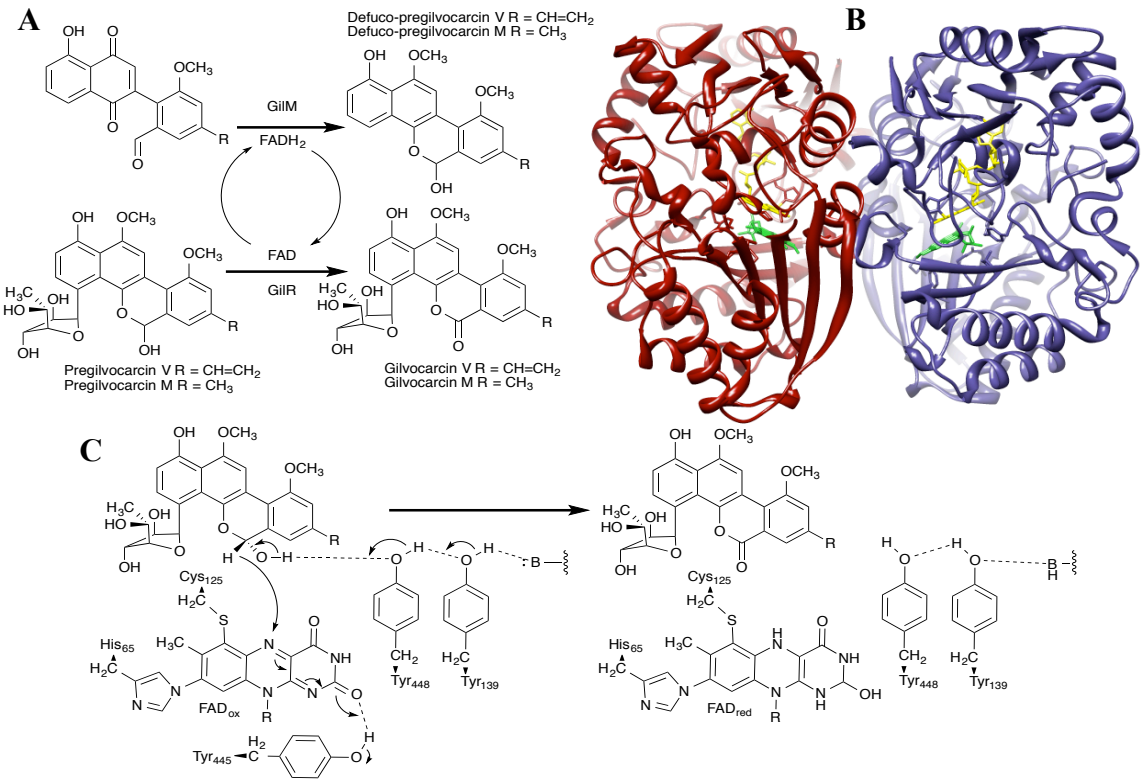


Figure 14. (A) GilM and GilR co-dependent reactions from the gilvocarcin biosynthetic pathway emphasizing the shared FAD cofactor. (B) GilR homodimer crystal structure with one dimer shown in red and the other in blue. Pregilvocarcin V substrate is shown in green and the FAD cofactor is shown in yellow. (C) Proposed GilR catalytic mechanism.

1.3.2.1 GilM-GilR

GilM and GilR are a non-consecutively functioning pair of co-dependent enzymes from the gilvocarin biosynthetic pathway. GilM is an $FADH_2$ dependent reductive *O*-methyltransferase responsible for quinone reduction, hemiacetal

formation, and *O*-methylation. Structurally, it is known the GilM is a homodimer in solution and co-purifies with the SAM cofactor. Its interaction with GilR was proven via *in vitro* experimentation during which GilM produced an increased amount of defucogilvocarcin when GilR was present. GilM functions first during biosynthesis prior to glycosylation and catalyzes quinone reduction, hemiacetal formation, and *O*-methylation of the substrate shown in the reaction in figure 14A with the aid of SAM to yield defuco-pregilvocarcin. Defuco-pregilvocarcin is then glycosylated by GilGT before being passed to GilR as pregilvocarcin.[35] GilR is an FAD dependent oxidoreductase that catalyzes the terminal step of gilvocarcin V biosynthesis and establishes the lactone core of the polyketide-derived gilvocarcin chromophore (figure 14A).[36,37] GilR catalyzes the final step in the pathway by the mechanism shown in figure 13C during which Tyr-448 aided by Tyr-139 abstracts the proton of the 6-OH group of pregilvocarcin, with concomitant hydride transfer of 6-H to N-5 of the FAD cofactor. The latter process is facilitated by Tyr-445, which protonates the 2-C=O group of FAD, thereby enhancing the electrophilicity of its N-5 position. Structurally, GilR functions as a homodimer as it is found as a homodimer in solution and crystalized. The structure of GilR also reveals the FAD cofactor is bicovalently bound by His-65 and Cys-125 (figure 14B).[37] The combined knowledge of GilR's FAD cofactor being bicovalently bound and GilM needing GilR to function properly lead to the hypothesis that this pair share the FAD cofactor that is bound to GilR effectively creating a shared FAD regeneration system.[35,37]

1.3.2.2 MtmC-MtmGIV

MtmC and MtmGIV are a non-consecutively functioning pair of co-dependent enzymes from the mithramycin biosynthetic pathway. These enzymes were both discussed previously as non-iteratively bifunctional enzymes in sections 1.2.2.5 and 1.2.2.4 respectively. As stated previously, MtmC is a reductase/methyltransferase from the deoxysugar portion of the mithramycin biosynthetic pathway (figure 8B) and MtmGIV is a glycosyltansferase responsible for the transfer of the first and third sugar moieties to the trisaccharide side chain (figure 8A). In addition to their

remarkable bifunctionality MtmC and MtmGIV have been shown to work cooperatively via *in vitro* enzyme reactions. During *in vitro* characterization of MtmC the methylated product of MtmC, TDP-4-keto-D-mycarose, was unable to be observed due to instability even though the SAM cofactor utilized by this reaction was being depleted. This indicated that the product was being formed but was unstable in solution. To overcome these instability issues MtmC's reactions were coupled with MtmGIV's reactions producing the full trisaccharide side chain intermediated premithramycin A3 from the substrate premithramycin A2' necessitating that these reactions are linked and that the enzymes are co-dependent.[19]

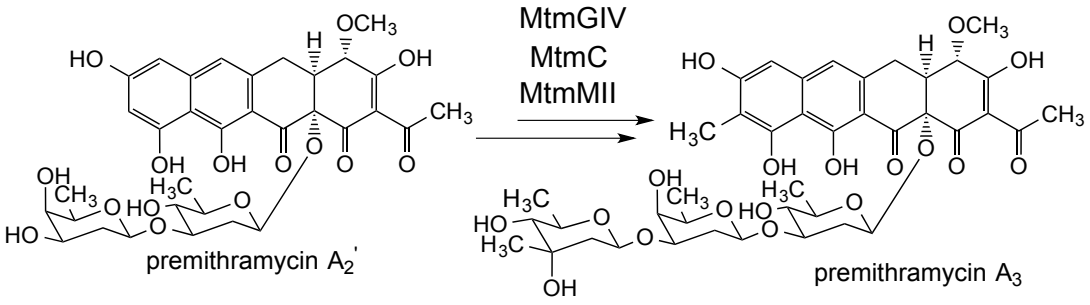


Figure 15. Conversion of premithramycin A_2' to premithramycin A_3.

1.4 Conclusions

If post-PKS tailoring enzyme bifunctionality and co-dependence were considered a common possibility many currently perplexing experimental outcomes could easily be explained or more thoroughly investigated for understanding. Bifunctional post-PKS tailoring enzymes appear to be more common then co-dependent enzymes but that may be due to the difficulty of detecting and confirming enzyme co-dependence. With regard to co-dependent enzymes there is a need to study the mechanism of their protein-protein interactions, as there are currently no in depth studies of these interfaces. For both bifunctional and co-dependent post-PKS enzymes there is a need for more structural information. Of the few structures that have been solved and studied very unique and interesting features have been found such as covalently and bicovalently bound FAD cofactors and complex binding sites. Post-PKS enzymes offer a real opportunity to create new medically relevant compounds

via combinatorial biosynthesis and enzyme engineering. If this opportunity is to be fully exploited more information about these complex systems is needed making bifunctional and co-dependent tailoring enzymes the new frontier of combinatorial biosynthesis.

1.5 Specific Aims

1.5.1 Specific aim 1 was to investigate the structure of MtmOIV and the role of active site residues in its catalytic mechanism.

The drastically improved protein structure of MtmOIV along with the high-resolution structure of MtmOIV in complex with its natural substrate premithramycin B reveals previously undetected key residues that are important for substrate recognition and catalysis. Kinetic analyses of selected mutants allowed probing of the substrate binding pocket of MtmOIV and confirmation of the putative NADPH binding site. This is the first substrate-bound structure of MtmOIV and the most complex BVMO-substrate complex structure known to date providing new insights into substrate recognition and catalysis. This investigation paves the way for the future design of a tailored enzyme for the chemo-enzymatic preparation of novel mithramycin analogues.

1.5.2 Specific aim 2 was to integrate the function of GilM and its protein-protein interactions with GilR that lead to their synergistic activity and sharing of GilR's bicovalently bound FAD moiety.

GilM was determined to be a key enzyme of the gilvocarcin pathway exhibiting dual functionality as a reductase and *O*-methyltransferase responsible for the reduction of a quinone intermediate to a hydroquinone, associated stabilizing *O*-methylation, and hemiacetal formation. Additionally, GilM was found to mediate its reductive catalysis through the assistance of GilR. GilR provides its bicovalently bound $FADH_2$ cofactor for the GilM reaction. In turn GilM regenerates FAD for GilR's next catalytic cycle establishing a unique co-dependent relationship. This unusual synergy eventually completes the biosynthesis of the polyketide-derived defuco-gilvocarcin chromophore. Additionally, the nessessary enzyme-enzyme interaction between GilM and GilR was

probed revealing the absence of a long-lived interface between the two and alluding to a transient relationship during catalysis.

2 Molecular Insight into Substrate Recognition and Catalysis of Baeyer-Villiger Monooxygenase MtmOIV, the Key Frame-modifying Enzyme in the Biosynthesis of Anticancer Agent Mithramycin.

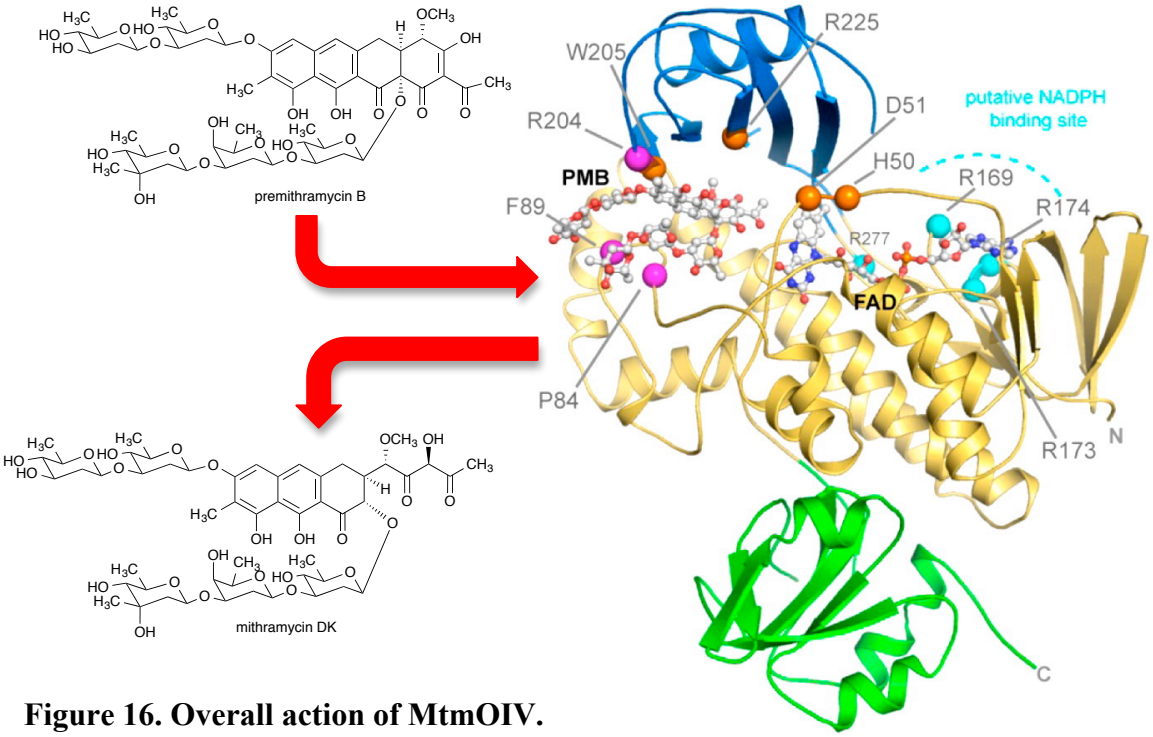


Figure 16. Overall action of MtmOIV.

2.1 Abstract

Baeyer-Villiger monooxygenases (BVMOs) have been shown to play key roles for the biosynthesis of important natural products. MtmOIV, a homodimeric FAD- and NADPH-dependent BVMO, catalyzes the key frame-modifying steps of the mithramycin biosynthetic pathway, including an oxidative C-C bond cleavage, by converting its natural substrate premithramycin B into mithramycin DK, the immediate precursor of mithramycin. The drastically improved protein structure of MtmOIV along with the high-resolution structure of MtmOIV in complex with its natural substrate premithramycin B are reported here, revealing previously undetected key residues that are important for substrate recognition and catalysis. Kinetic analyses of selected mutants allowed us to probe the substrate binding pocket of MtmOIV and also to discover the putative NADPH binding site. This is the first substrate-bound structure of MtmOIV providing new

insights into substrate recognition and catalysis, which paves the way for the future design of a tailored enzyme for the chemo-enzymatic preparation of novel mithramycin analogues.

2.2 Introduction

Baeyer–Villiger monooxygenases (BVMOs), an important subclass of flavoprotein monooxygenases, not only have proven to be potent biocatalysts for synthetic organic chemistry applications[38] but also have been found to hold critical roles in the biosynthesis of many natural products.[39] Additionally, BVMOs are known to be responsible for pro-drug activation[40] as well as biodegradation reactions.[41–43]

MtmOIV is a homodimeric BVMO from the mithramycin biosynthetic pathway that catalyzes the key frame-modifying step responsible for producing bioactivity. This FAD- and NADPH-dependent enzyme reacts with its natural substrate premithramycin B to form mithramycin DK, the substrate of the final enzymatic step of the mithramycin biosynthetic pathway.[29,44] Mithramycin (MTM, also known as aureolic acid, mithracin, LA-7017, PA-144, and plicamycin) is an aureolic acid-type polyketide anticancer antibiotic produced by the soil bacterium *Streptomyces argillaceus* (ATCC 12956) and various other species of streptomycetes. The small distinct group of aureolic acid-type anticancer antibiotics includes MTM, chromomycin, olivomycin, UCH9, and durhamycin, which all contain the same polyketide-derived tricyclic aromatic core with a highly functionalized pentyl side chain attached at C-3 but vary with respect to their saccharide patterns and side chains in the 7-position.[45–47]

The structure of MTM consists of a tricyclic aglycone core, a highly functionalized pentyl side chain at the 3-position, and five deoxysugars linked as a trisaccharide at the 2-position and a disaccharide at the 6-position, respectively. MTM biosynthesis proceeds through the polyketide-derived tetracyclic premithramycinone, to which the five sugar moieties and one C-methyl group are added leading to premithramycin B.[19,45–49] Then, MtmOIV oxidatively cleaves the fourth ring via a Baeyer–Villiger reaction, generating MTM's characteristic tricyclic aglycone core and

the highly functionalized pentyl side chain at position 3. The Baeyer–Villiger reaction precedes lactone opening, decarboxylation, and the final step of MTM biosynthesis, the reduction of the 4′-keto group catalyzed by ketoreductase MtmW (Figure 16).[31]

MTM exhibits anticancer activity by inhibiting replication and transcription by cross-linking of DNA strands predominantly in GC-rich regions used by Sp promoters.[50–52] MTM has been used clinically to treat certain cancers, such as testicular carcinoma,[53,54] and bone diseases, such as Paget's disease,[55–58] and particularly to address cancer-related hypercalcemia.[59–62] Recent discoveries include MTM's potential (i) as therapeutic to treat Huntington's disease,[63–65] (ii) as a lead drug out of a screen of 50,000 compounds affecting Ewing sarcoma promotion factors that previously were believed to be "undruggable",[66] and (iii) as a repressor of ABCG2-related stem cell signaling in lung and esophageal cancers.[67] Yet, the clinical use of MTM is still hindered because of its hardly manageable toxicity, which could potentially be reduced through the generation of new, less toxic analogues. In this context, the function of the Baeyer–Villiger monooxygenase (BMVO) MtmOIV needs to be fully understood to pave the way for further chemo-enzymatic and/or combinatorial biosynthetic approaches toward new MTM analogues.

Besides our low resolution structure of MtmOIV[44] only a few crystal structures of Baeyer–Villiger reaction-performing enzymes are available to date,[38] including phenylacetone monooxygenase from the thermophilic bacterium *Thermobifida fusca*,[68] the oxygenating component of 3,6-diketocamphane monooxygenase,[69–71] a dimeric BVMO catalyzing the lactonization of 2-oxo-Δ(3)-4,5,5-trimethylcyclopentenylacetyl-CoA, a key intermediate in the metabolism of camphor by *Pseudomonas putida*,[42] and cyclohexanone monooxygenase (CHMO) from *Rhodococcus* sp. HI-31,[72] which could also be solved in complex with cyclohexanone as well as NADP and FAD.[73] Most recently, the toxoflavin lyase structure was solved, and the enzyme was found to catalyze a degradation reaction via a Baeyer–Villiger oxidation reaction.[43] Various BVMOs have been described and studied, and most of these appear to be involved in oxidative degradation pathways, for which substrates were identified ranging from simple ketones,

such as phenylacetone or cyclohexanone and their derivatives, to steroids, such as progesterone.[38,43,74] In contrast to these mostly monofunctional molecules, the MtmOIV substrate premithramycin B (PMB) is a multifunctional and richly decorated molecule and thus remains the most complex substrate and one of very few unambiguously proven *natural* BVMO substrates.[75]

We have previously reported the native structure of MtmOIV that provided initial insights into the intriguing enzymatic BV mechanism.[44] However, crystals for this structure suffered from a high solvent content that led to many disordered side chains. Here, we report the significantly improved crystal structure of native MtmOIV to 2.0 Å resolution, as well as the crystal structure of MtmOIV in complex with its natural substrate premithramycin B to 1.85 Å resolution, allowing us to identify key residues involved in substrate recognition and catalysis. This provides the structural information required to re-engineer MtmOIV to have altered substrate specificity for the generation of new bioactive analogues of MTM. Our structure of MtmOIV in complex with premithramycin B provides insight into substrate recognition and catalysis in MtmOIV and other important enzymes from this family. Furthermore, mutagenesis, kinetic assays, and docking studies based on a low resolution NADPH-bound co-crystal structure allowed us to probe the substrate binding pocket of MtmOIV and to discover the putative NADPH binding site, which sits opposite the substrate binding site, yet still in close proximity to the FAD cofactor.

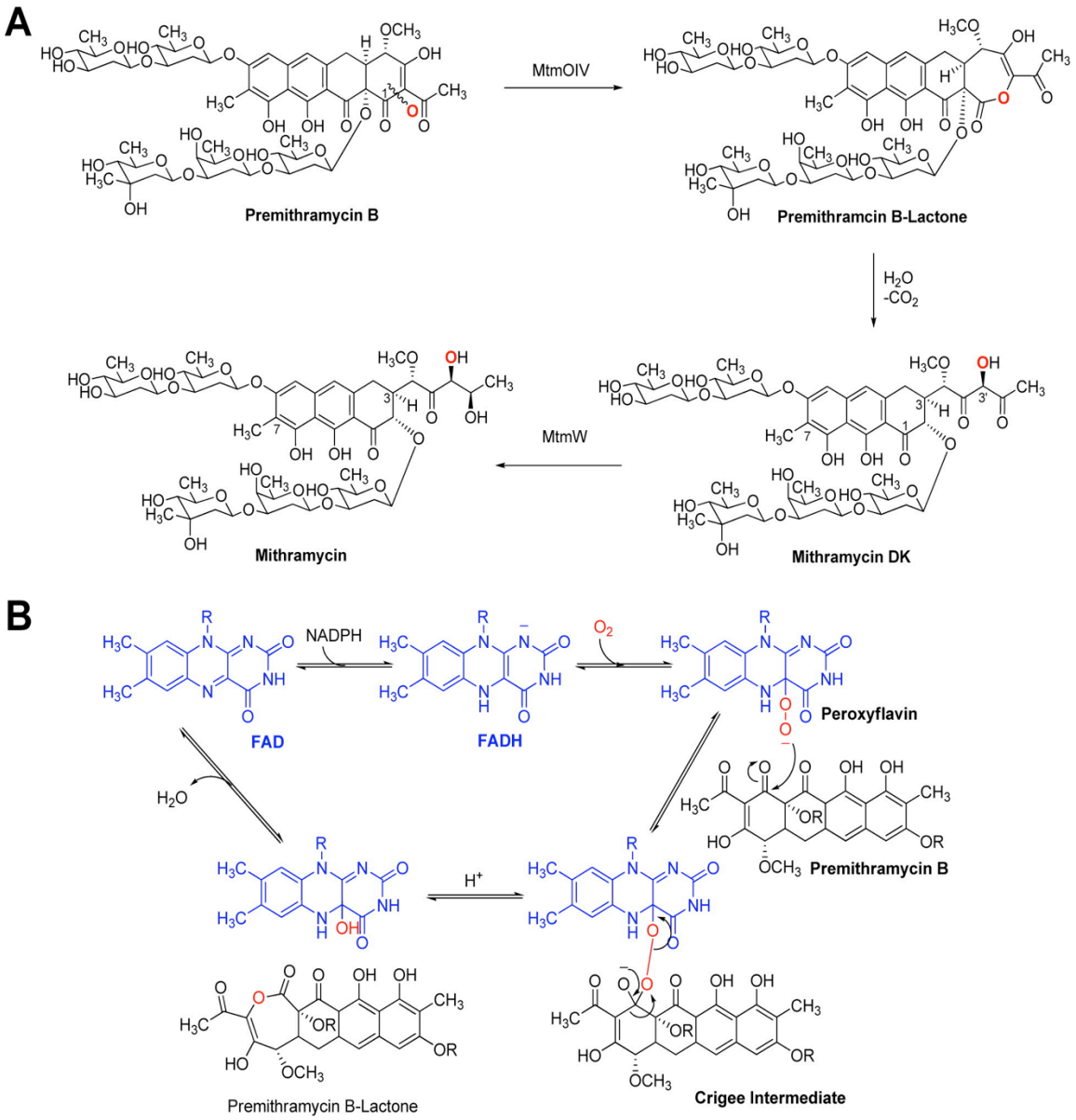


Figure 17. Mithramycin biosynthetic pathway. (A) The reaction catalyzed by the Baeyer–Villiger monooxygenase (BVMO) MtmOIV yields premithramycin B-lactone, which is further converted to mithramycin DK. Ketoreductase MtmW catalyzes the final step of mithramycin biosynthesis. (B) Suggested Baeyer–Villiger oxidation of premithramycin B (sugar residues not shown) to premithramycin B-lactone involving cofactors FADH and NADPH (for FADH regeneration).

2.3 Results and Discussion

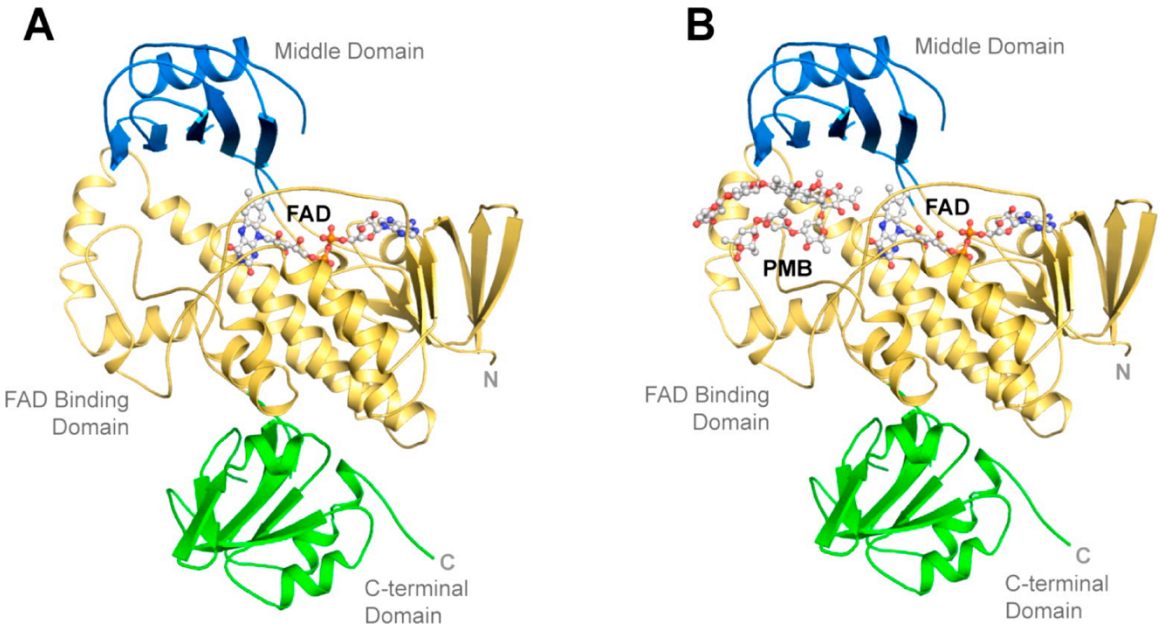


Figure 18. Overall structures of MtmOIV and the MtmOIV-premithramycin B complex. (A) The improved native MtmOIV crystal structure to 2.0 Å resolution. FAD is shown in ball and stick representation. (B) The MtmOIV-premithramycin B complex crystal structure showing FAD and premithramycin B in ball and stick representation. For both panels, the FAD domain is shown in gold, the middle domain in blue, and the C-terminal domain in green.

2.3.1 Overall Structure of MtmOIV

The crystal structure of native MtmOIV has previously been described.[44] However, the N-terminal fusion tag was not removed, and the structure suffered from high solvent content (~70%) that led to a large number of disordered side chains. Here, we have removed the N-terminal fusion tag, which led to crystal conditions with overall improved quality and significantly improved electron density quality and side chain order with ~50% solvent content. The new structure was solved to 2.0 Å resolution in space group *P*2 with one molecule per asymmetric unit. The final structure contains 499 of 533 total residues (Ala10–Ala509) for MtmOIV and a single noncovalently bound FAD cofactor (Figure 17A). Final data collection and refinement statistics are summarized in Table 1. Figure 17A shows the three domains of MtmOIV: the FAD binding domain

(residues 10–183 and 276–394), the middle domain (residues 184–275), and the C-terminal domain (residues 395–510). A loop containing residues 218–224, which lies at the perimeter of the substrate binding pocket, was found ordered in the new structure. However, the loop containing residues 233–239, which has been suggested to be involved in NADPH binding,[44] was again found disordered along with a small loop within the C-terminal domain containing residues 402–403.

2.3.2 Structure of MtmOIV-PMB Complex

For the structure of MtmOIV in complex with its substrate premithramycin B (PMB), PMB was added to MtmOIV prior to concentrating the protein for crystallization trials. Co-crystals were grown in final conditions consisting of 200mM ammonium acetate and 30% poly(ethylene glycol) (PEG) 1000, and the structure was solved to 1.85 Å resolution in space group $P222_1$, with two molecules per asymmetric unit. The final structure contains 500 of 533 total residues (Ala10–Asn510) for MtmOIV and a single noncovalently bound FAD cofactor per subunit (Figure 17B) with the same domain architecture as the native structure (root-mean-square deviation (RMSD) of 0.266 Å for 2933 aligned atoms). A difference density map revealed the presence of unmodeled density along the putative substrate binding site in which PMB was modeled and refined (Figures 17B and 18). Final data collection and refinement statistics are summarized in Table 1.

As shown in Figure 16B, PMB binds MtmOIV within the large binding pocket formed between the FAD binding domain and the middle domain, and the sugar chains of PMB interact with the inner walls of the middle and the FAD binding domains within the pocket. Ignoring any ordered water molecules found in the structure, only a few residues from MtmOIV appear to make direct hydrogen bonding interactions with PMB including R225 and W205 (Figure 19A), indicating that van der Waals and hydrophobic interactions may play a major role in substrate binding and selectivity. When we include the ordered water molecules found within the crystal structure, it becomes clear that the ordered solvent may also play a major role in substrate binding by forming a hydrogen

bonding network that bridges PMB binding within the binding pocket of MtmOIV (Figure 19B). A Ligplot analysis of the interactions between PMB and MtmOIV depicts best the vast number of interactions involved in substrate binding involving not only protein side chains but also many water molecules that are crucial for binding (Figure 19C).

The improved native (substrate-free) MtmOIV crystal structure described above allowed us to compare the substrate-free and substrate-bound crystal structures. The overall conformation of MtmOIV in the complex structure was the same as in the native structure; however, a minimal domain shift was observed within the middle domain relative to the FAD binding domain (more open), which likely serves to accommodate the bound substrate. While only minimal rearrangements were observed for most of the side chains of residues lining the substrate binding pocket, the largest shifts appear in W205 (1.6 Å along atom CH_2) and F89 (0.9 Å along atom C_ζ), which sit on opposite sides at the entrance to the substrate binding pocket. Aside from these slight shifts, no major conformational differences were observed between the substrate-free and substrate-bound states of MtmOIV.

Table 1. Data Collection and Refinement Summary

	Native	PMB
Data Collection		
λ (Å)	1.0	1.0
space group	*P*2	$P222_1$
mol/ASU	2	1
a (Å)	83.3	55.72
b (Å)	56.05	79.75
c (Å)	124.9	154.26
α (deg)	90	90
β (deg)	95	90
γ (deg)	90	90
resolution (Å)	50–2.0 (2.07–2.0)	50–1.85 (1.92–1.85)
completeness (%)a	99.4 (97.7)	96.1 (91.5)
redundancya	3.8 (3.0)	5.6 (4.4)
R_{sym}[b,a]	0.10 (0.66)	0.09 (0.71)
I/σ (I)a	12.2 (1.7)	15.1 (1.8)

	Native	PMB
Refinement		
resolution (Å)	20–2.0	20–1.85
$R^c/R_{free}{}^d$	0.21/0.25	0.18/0.22
bonds (Å)	0.01	0.007
angles (deg)	1.138	1.139
protein atoms	7238	3734
ligand atoms	106	130
water molecules	358	508
B-Factors		
protein	48.7	20.7
ligand	34.6	22.8
water molecules	42.7	31.8
Ramachandran Analysis		
core (%)	90.9	93.1
allowed (%)	8.9	6.6
generously allowed (%)	0.1	0.3
disallowed (%)	0	0
PDB ID	4K5R	4K5S

[a] Indicates statistics for last resolution shell shown in parentheses.

[b] $R_{sym} = \Sigma_{hkl,j} (|I_{hkl} - I_{hkl}|)/\Sigma_{hkl,j} I_{hkl}$, where I_{hkl} is the average intensity for a set of *j* symmetry-related reflections and Ihkl is the value of the intensity for a single reflection within a set of symmetry-related reflections.

[c] *R* factor = $\Sigma_{hkl} (F_o| - |F_c)/\Sigma_{hkl} |F_o|$ where F_o is the observed structure factor amplitude and F_c is the calculated structure factor amplitude.

[d] $R_{free} = \Sigma_{hkl,} T (F_o| - |F_c)/\Sigma_{hkl,} T |F_o|$, where a test set, *T* (5% of the data), is omitted from the refinement.

[e] Performed using Procheck.

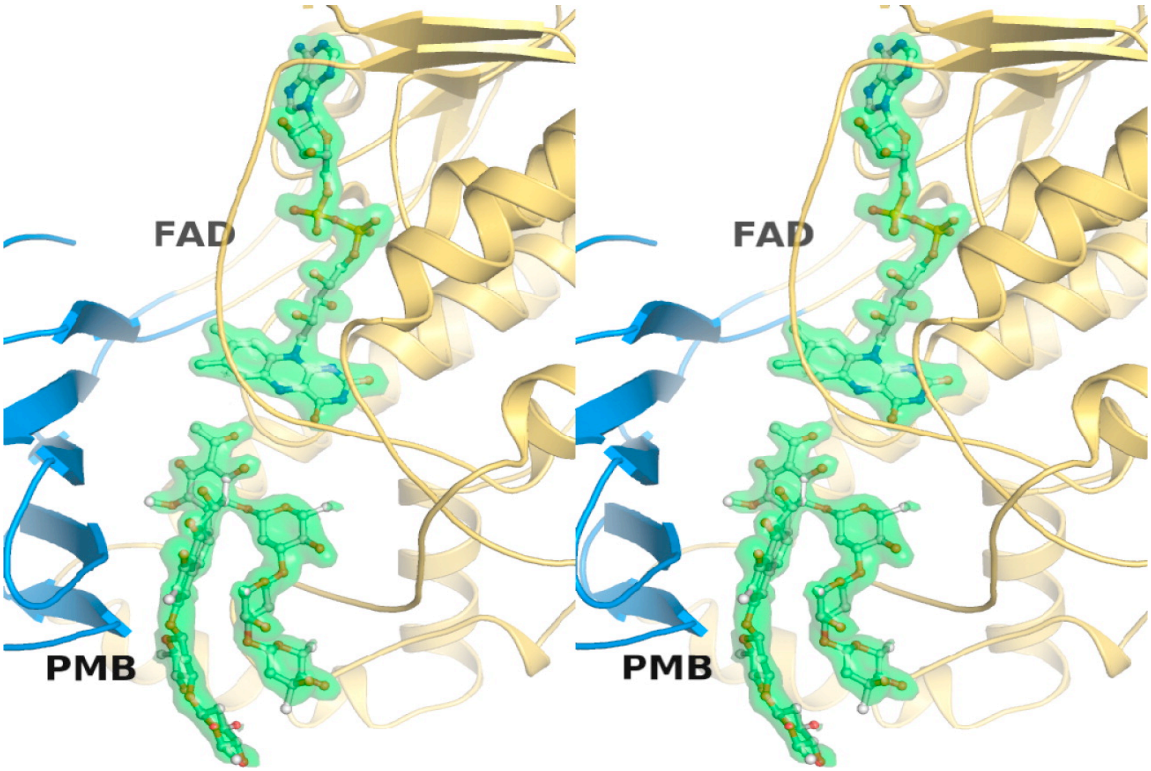


Figure 19. Stereoview showing the electron density for FAD and premithramycin B. The middle domain in shown in blue, the FAD domain is in gold, both FAD and premithramycin B (PMB) are shown in ball and stick representation, and the electron density (SA-omit $F_o - F_c$ map contoured to 2.5 σ) is shown as a green transparent isosurface.

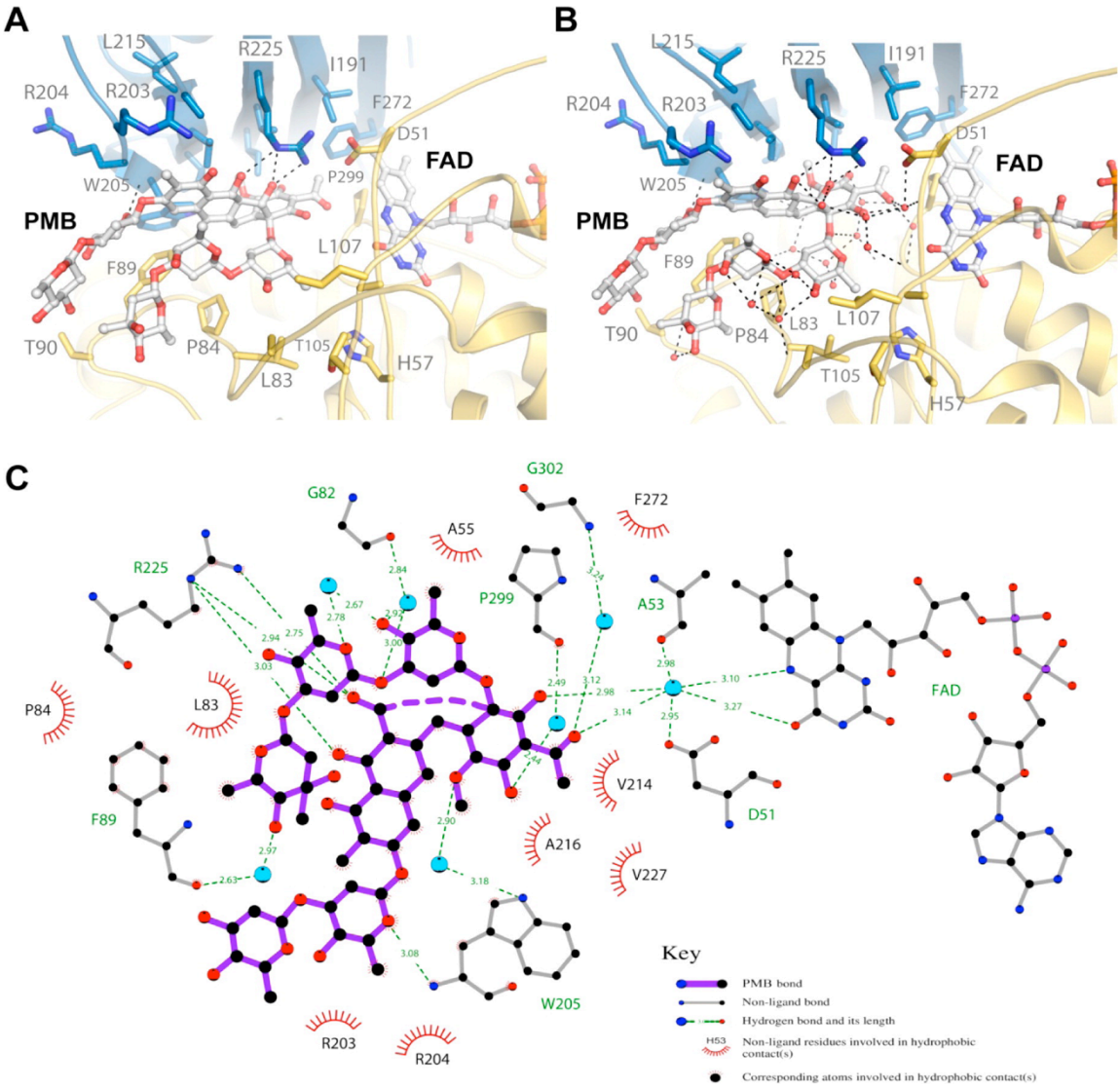


Figure 20. Premithramycin B binding site in MtmOIV. (A) Shown are the interactions (dashed lines) of MtmOIV with premithramycin B (PMB) bound along the active site with water molecules removed from structure. Without water molecules, there appear to be two primary interactions with W205 and R225, with van der Waals and hydrophobic interactions contributing a large part to the binding energy. (B) Shown are the interactions of MtmOIV with PMB, including the water molcules that were observed in the crystal structure (1.85 Å resolution). Once solvation is included, a vast number of interactions (dashed lines, distance cutoff of 3.3 Å) are observed that are bridged by ordered water molecules, indicating that hydrogen bonding networks play a major role in substrate binding. (C) Ligplot of PMB bound to MtmOIV indicating important interactions (dashed lines, distance cutoff of 3.3 Å). For clarity, atom names for PMB have been removed, and only water molecules (cyan spheres) having two or more hydrogen bonding interactions are shown.

2.3.3 Mutagenesis and Steady State Kinetics Assays

Using the structure of MtmOIV in complex with PMB, we were able to identify residues important for substrate recognition (P84, F89, R204) and catalysis (H50, D51, W205, R225) (Figure 20A). We mutated these residues to alanine and then used a kinetic assay to determine the effect each mutation had on substrate binding and catalysis (Figure 20B–D and Table 2). While the substrate binding of mutants was relatively tolerant to most of the mutations, turnover rates for several of the mutants were severely affected. Interestingly, none of the substrate recognition site mutants significantly affect substrate binding except the double mutant P84A/F89A with a ~50-fold decrease in K_M compared to the wild-type enzyme, leading to a 3-fold increased catalytic efficiency. The single mutants F89A, P84A, and R204A decrease K_M only by 3–10-fold; however, their catalytic efficiencies are even greater than that of the double mutant due to a less severe decrease in k_{cat}. The active site mutant W205A is interesting since its catalytic efficiency is drastically decreased (~30-fold), although the substrate binding is comparable to that of the wild-type enzyme, likely because of its role as gatekeeper for the entry into the binding pocket (see below). Arginine-225 appears crucial; its mutation led to a noncatalytic enzyme. Figure 19 illustrates that R225 stabilizes the polyketide core of PMB. Interestingly, the D51A mutant, like the previously reported F89A mutant, showed a ~2-fold increase in k_{cat} and a drastically increased catalytic efficiency, which proves the crucial role of this amino acid for MtmOIV's catalysis of the BV reaction. It is possible that changing the carboxylate residue of aspartate into the simple methyl group of alanine may decrease both sterically and electronically disadvantageous interactions between the substrate PMB and the peroxyflavin (Table 2, Figures 16 and 19).

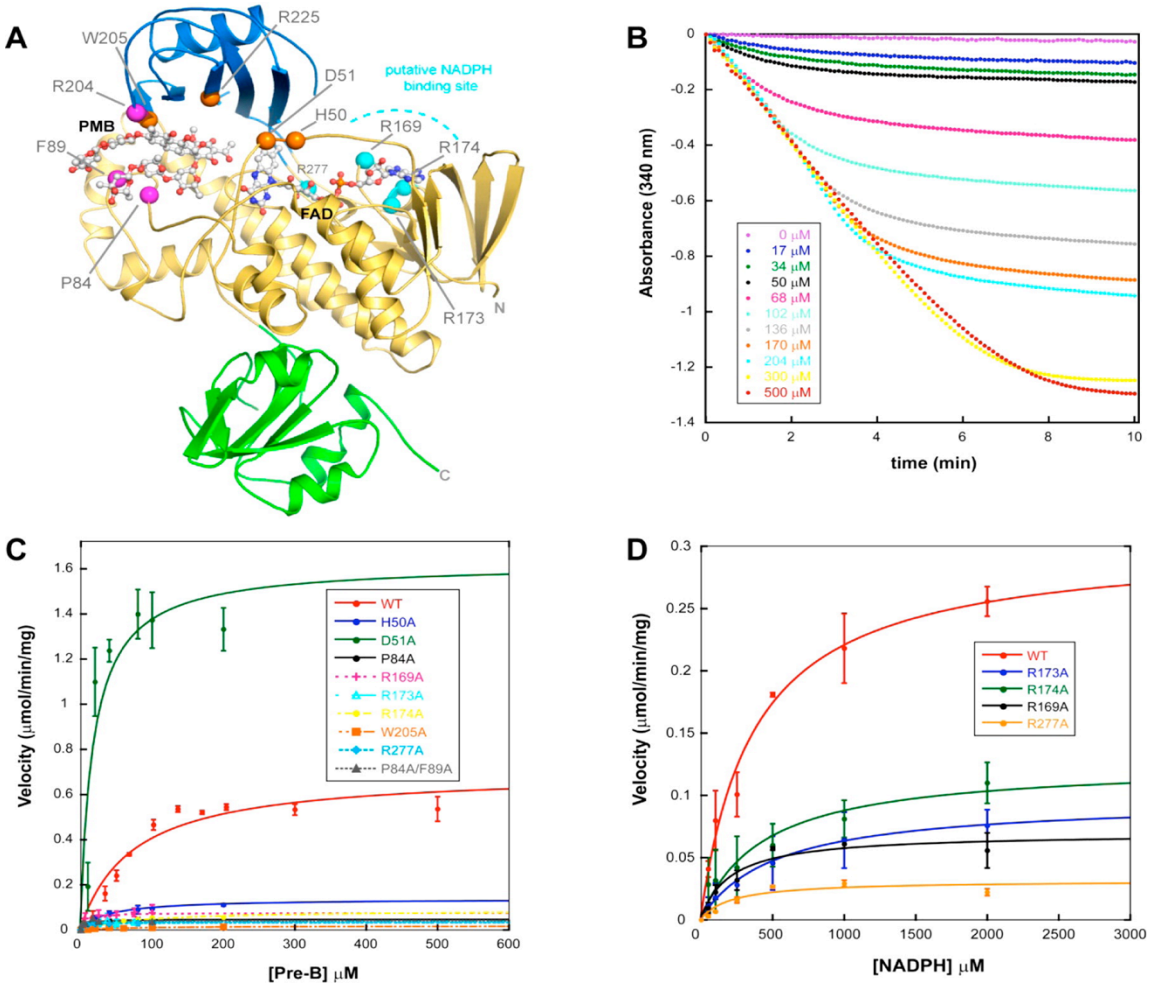


Figure 21. Probing substrate recognition and catalysis in MtmOIV. On the basis of structural analysis, three regions were initially targeted for this study to determine their contribution to substrate binding and catalysis (see Table 2). (A) Shown is the MtmOIV-premithramycin B (PMB) crystal structure with the location of the mutations that were made shown as spheres and color coded by their predicted role in catalysis: (i) substrate binding (magenta), (ii) active site (orange), and (iii) putative NADPH binding (cyan). The FAD domain is shown in gold, the middle domain in blue, the C-terminal domain in green, and both FAD and PMB are shown in ball and stick representation. (B) Sample of raw kinetic data collected at λ = 340 nm. (C) Michaelis–Menten curve fitting of data for WT MtmOIV and mutants. (D) Curve fitting of data for WT MtmOIV and the mutants R169A, R173A, R174A, and R277A.

Table 2. Summary Table of MtmOIV Mutagenesis and Kinetics Assays

mutation(s)	target region	K_M (μM)	V_{max} (μmol min^{-1} mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} mM^{-1})
			NADPH		
wild-type		372 ± 68	0.3 ± 0.02	0.29 ± 0.02	0.8
R169A	NADPH site	211 ± 71	0.07 ± 0.007	0.07 ± 0.006	0.3
R173A	NADPH site	528 ± 63	0.1 ± 0.004	0.09 ± 0.0042	0.2
R174A	NADPH site	437 ± 159	0.1 ± 0.017	0.12 ± 0.02	0.3
R277A	NADPH site	221 ± 118	0.03 ± 0.004	0.03 ± 0.0005	0.1
			Pre-B		
wild-type		73 ± 20	0.7 ± 0.07	0.7 ± 0.06	9
P84A	substrate binding	7.8 ± 2.4	0.04 ± 0.002	0.07 ± 0.007	9
F89A[8]	substrate binding	27.1 ± 8.9	1.3 ± 0.1	1.27 ± 0.11	47
P84A/F89A	substrate binding	1.5 ± 0.9	0.04 ± 0.005	0.04 ± 0.004	26
R204A[8]	substrate binding	18.0 ± 2.6	0.77 ± 0.02	0.73 ± 0.02	41
H50A	active site	40.4 ± 15.3	0.14 ± 0.02	0.13 ± 0.02	3
D51A	active site	18.1 ± 9.9	1.62 ± 0. 2	1.55 ± 0.2	86
W205A	active site	62.1 ± 29.0	0.02 ± 0.003	0.02 ± 0.003	0.3
R225A	active site	no activity	no activity	no activity	no activity
R169A	NADPH site	80.2 ± 42.3	0.08 ± 0.02	0.08 ± 0.02	1
R173A	NADPH site	5.3 ± 4.0	0.04 ± 0.005	0.02 ± 0.003	4
R174A	NADPH site	5.4 ± 3.0	0.09 ± 0.02	0.07 ± 0.007	13
R277A	NADPH site	4.8 ± 1.4	0.03 ± 0.001	0.03 ± 0.004	6

2.3.4 Identification of the Putative NADPH Binding Site of MtmOIV

We recently grew crystals of MtmOIV in the presence of 5 mM NADPH that diffracted to ~3.5 Å (data not shown). We were able to solve the structure and observed density for the putative NADPH binding loop (residues 233–239), which had previously always been found disordered, as well as for a possible NADPH molecule found at a novel binding site along the second identified Rossmann-type fold (residues 163–180 correspond to a single β-α-β super secondary structural motif commonly found in Rossmann folds), which is distinct from either the FAD or the substrate binding sites (Figure 22A and B). However, due to the low resolution of the data and partial disorder within the ligand, we were not able to model the entire NADPH molecule, and our efforts to improve the resolution of the crystals have so far been unsuccessful. Yet, on the basis of the location of the density, we were able to identify four residues (R169, R173, R174, R277) that seem well positioned to interact with NADPH at this site (Figure 21A and B). Mutating these residues to alanine significantly altered the catalysis of MtmOIV, although the K_M values of the mutants still indicate similar NADPH binding as found for the wild-type enzyme. We assume that inactivating single arginine residues leads to a distorted NADPH binding that precludes efficient reduction of FAD and therefore efficient enzyme turnover. Unexpectedly, PMB binding seems to be enhanced upon changing the latter three of these four arginine residues (Table 2),[76] possibly because the access of substrate PMB is facilitated if NADPH is bound less efficiently to the enzyme. All four arginine residues are fully conserved in other FAD- and NADPH-dependent monooxygenases, e.g., PgaE and CabE,[77] two enzymes with a similar fold as MtmOIV (Figure 21C and D). These R to A mutants seem not to affect FAD binding, since they keep the same bright yellow color as the wild-type enzyme.

Using this information along with the recently reported protein structure of CHMO in complex with NADP[73] and the suggested role of the disordered loop containing residues 233–239 for NADPH binding, we performed steered docking studies using AutoDock 4 to form an *in silico* model for the NADPH-bound state of MtmOIV

(Figure 22). The putative NADPH binding site is formed by the outer edge of the FAD binding domain and the middle domain and is located on the opposite side of the substrate binding pocket, but in close proximity to the FAD cofactor. Together, our structural and mutagenesis studies have allowed us not only to identify both the exact FAD and substrate binding sites but also to identify the putative NADPH binding site.

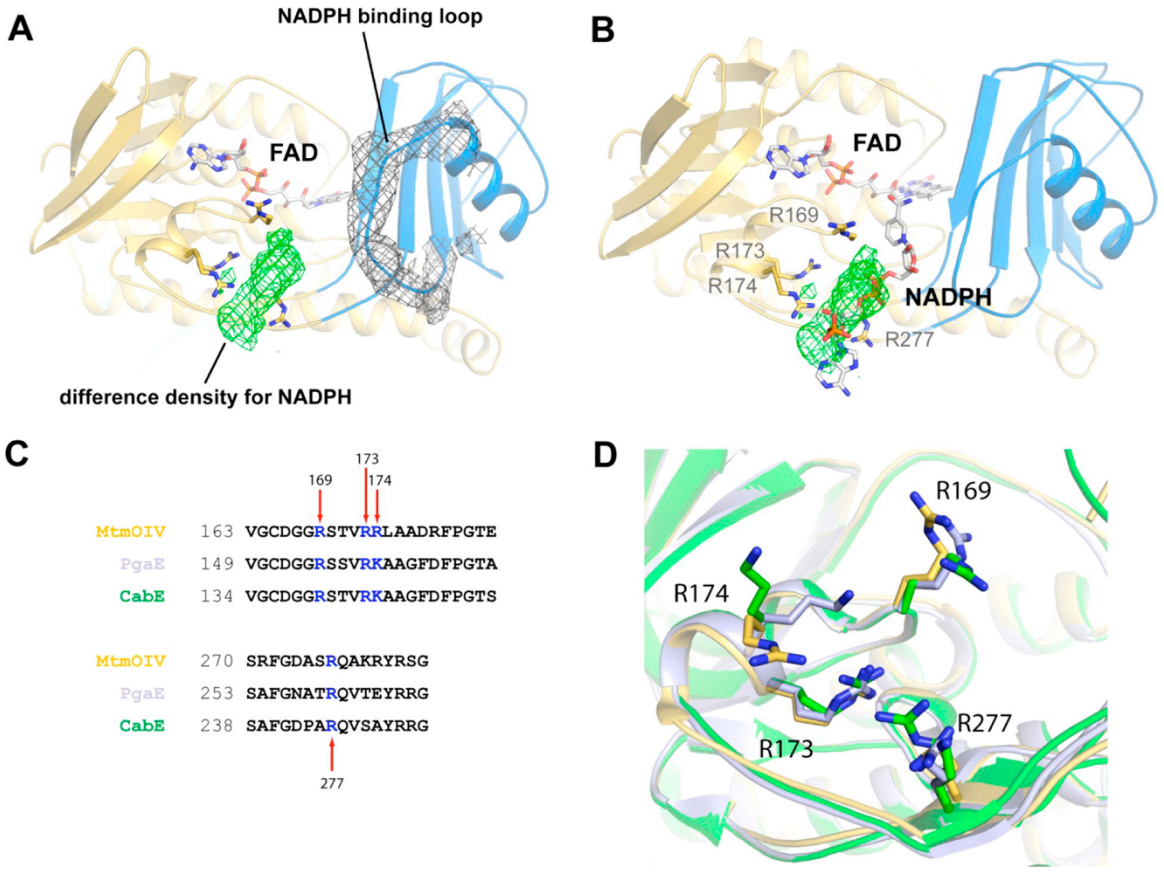


Figure 22. Low resolution crystal structure of NADPH bound to MtmOIV. (A) The ordered NADPH binding loop of MtmOIV showing $2F_o - F_c$ electron density (gray) contoured at 0.8 σ. Difference density (green) contoured at 2.5 σ for NADPH that was observed only in this crystal structure and not within any of the previous MtmOIV structures. The space group for the NADPH-bound co-crystal structure was *P1* with 6 molecules in the asymmetric unit. The resolution was 3.5 Å, and final R/R_{free} values are 0.22/0.26. (B) Rigid body placement of NADPH along the difference density within the MtmOIV structure, which was used as a starting point for subsequent docking studies. (C) Sequence alignment showing the conservation of basic residues at the putative NADPH binding site. (D) Structural alignment of MtmOIV (gold), PgaE (light purple), and CabE (green) depicting residues proposed to interact with NADPH shown in stick representation.

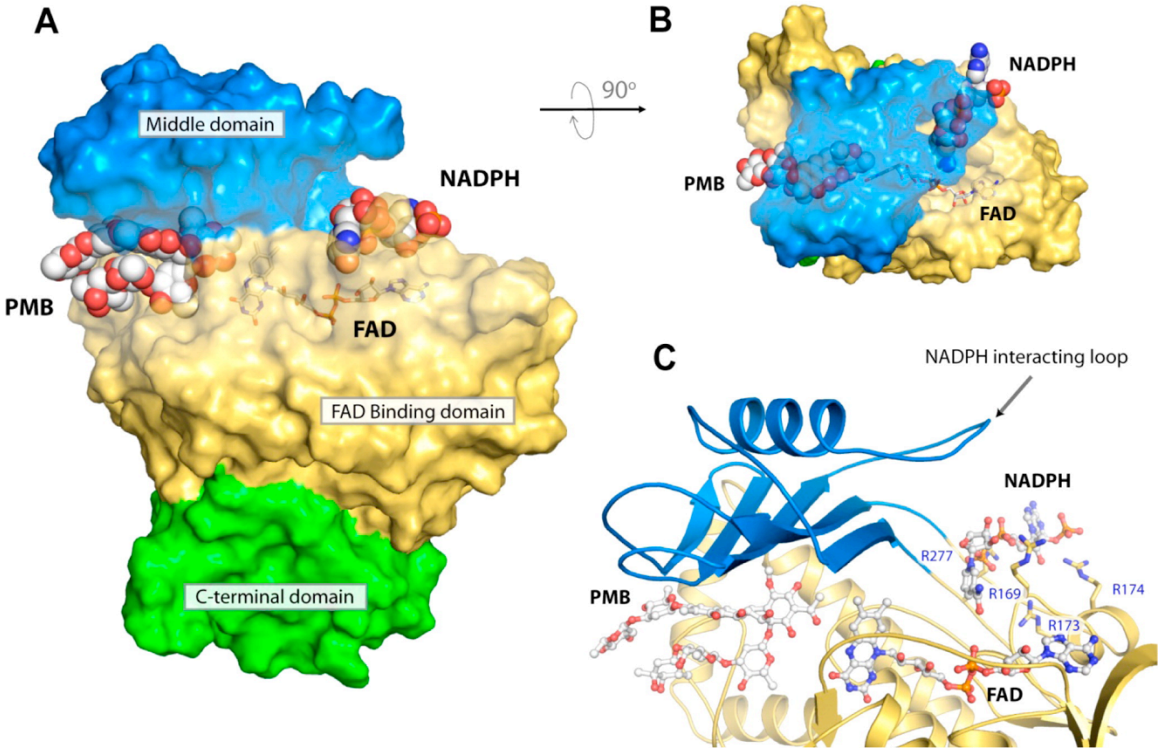


Figure 23. NADPH docked into the putative NADPH binding pocket in MtmOIV. (A) The MtmOIV structure showing the bound premithramycin B in the substrate binding pocket and the lowest energy docked model (−2.8 kcal/mol) of NADPH positioned within the putative NADPH binding pocket. (B) Top-down view of the MtmOIV structure shown in panel A through the middle domain (blue). The binding pockets in panels A and B are shown as transparent surfaces, and for clarity, FAD is shown in stick. (C) Close-up view of the putative NADPH binding site shown in proximity to FAD and PMB (ball and stick), which was co-crystallized within the substrate binding pocket. Middle domain loop consisting of residues 233–239 has been postulated to participate in binding NADPH (ball and stick) and was found disordered in all known structures of MtmOIV except our low resolution NADPH co-crystal structure. Those residues that were identified from our low resolution crystal structure to interact with NADPH (R169, R173, R174, R277) are shown in stick representation.

2.4 Conclusions

Our new refined structure is the first substrate-bound crystal structure of MtmOIV. The MtmOIV structure with its bound substrate PMB revealed significant differences to the previous computer model, in which PMB was docked into the low resolution structure of MtmOIV.[44] Most important, two glutamine residues (Q78 and Q91) and a proline residue (P84) were recognized as critical to hold the trisaccharide chain in place. The trisaccharide chain is essential for the DNA interaction of the anticancer drug mithramycin (MTM) with DNA,[78,79] and BVMO MtmOIV, which catalyzes the second to last step in MTM biosynthesis, seems to be optimized to exclusively process PMB, i.e., a molecule already fully glycosylated with the 5 sugars found in the final mithramycin (MTM) molecule. The three above-mentioned trisaccharide stabilizing residues were not identified in the previous PMB–MtmOIV binding site model but are likely critical for our future attempts to further modify MTM's trisaccharide side chain. Preliminary SAR studies showed that modifications of this trisaccharide chain can significantly change the biological activity of mithramycins,[80,81] and it was shown that changing the terminal D-mycarose of the trisaccharide chain into D-digitoxose led to a significant reduction of toxicity.[82] A re-engineered MtmOIV will be essential for future approaches to bioengineer MTM variants with alternate trisaccharide residues. Proline residue (P84) appears to be pivotal to accommodate the trisaccharide chain of the substrate PMB. Interestingly, a P84A mutation increased the catalytic activity of MtmOIV almost 1.5-fold, but seems less effective than the previously generated F89A mutant, which led to a 5-fold activity increase. Combining these successful mutations in a P84A/F89A double mutant failed, since the double mutant led to only a 3-fold increased catalytic activity. In contrast to the longer trisaccharide chain that interacts well inside the PMB binding pocket, the shorter disaccharide chain of PMB seems to establish no particular interaction with that pocket (Figure 19) and may serve as a mechanical block along with the gatekeeper residues W205 and F89 (see below) to allow only a certain degree of penetration of the PMB molecule into its binding pocket.

The new MtmOIV protein structures discussed here also revealed that a tryptophan (W205), arginine (R204), and phenylalanine (F89) residue that play key roles in guarding the entry loop during the substrate binding process. Comparing the substrate-free and the substrate-bound structures on MtmOIV showed the largest shifts for W205 (1.6 Å) and F89 (0.9 Å), which sit on opposite sides at the entrance to the substrate binding pocket. To confirm our hypothesis regarding the W205 residue as key gatekeeper, we prepared the W205A mutant, which turned out to be the only mutant that negatively affects substrate binding, while the previously published F89A and R204A mutants had only marginal effects on the substrate binding and slightly increased catalytic activity. In contrast, the W205A mutation decreased the catalytic activity about 30-fold compared to that of the wild-type enzyme (see Table 2), corroborating W205's role as key gatekeeper that needs to stay intact to correctly place the PMB molecule. The new structures also revealed that a histidine (H50) and an aspartate (D51) residue seem to be important to fine-tune the positioning of the FAD cofactor. Modifying these could affect the catalytic activity of the enzyme. Indeed, in the H50A and D51A mutants an increased catalytic activity of 1.5- and 3-fold, respectively, was observed.

On the basis of a low-resolution complex structure of MtmOIV in complex with NADPH, we identified four arginine residues (R169, R173, R174, R277) potentially important for binding NADPH. These residues are quite distant from the substrate binding site, and mutants were made to confirm the putative NADPH binding site. The mutants showed a clearly diminished catalytic activity of 10–50% relative to the wild-type enzyme in their ability to reduce NADPH, supportive of the hypothesis that these residues contribute to the NADPH binding. Future ligand binding studies to completely delineate this enzyme's complex mechanism may shed greater light on the role each NADPH binding arginine plays, especially if product inhibition plays a role and becomes the rate-limiting step of catalysis.

In summary, the high resolution structures of MtmOIV revealed new details of this important bottleneck enzyme of the MTM biosynthesis pathway, and the subsequent

structural mutations pointed out possible strategies for the broadening of its substrate specificity and enhancement of its catalytic activity.

2.5 Methods

2.5.1 MtmOIV Cloning

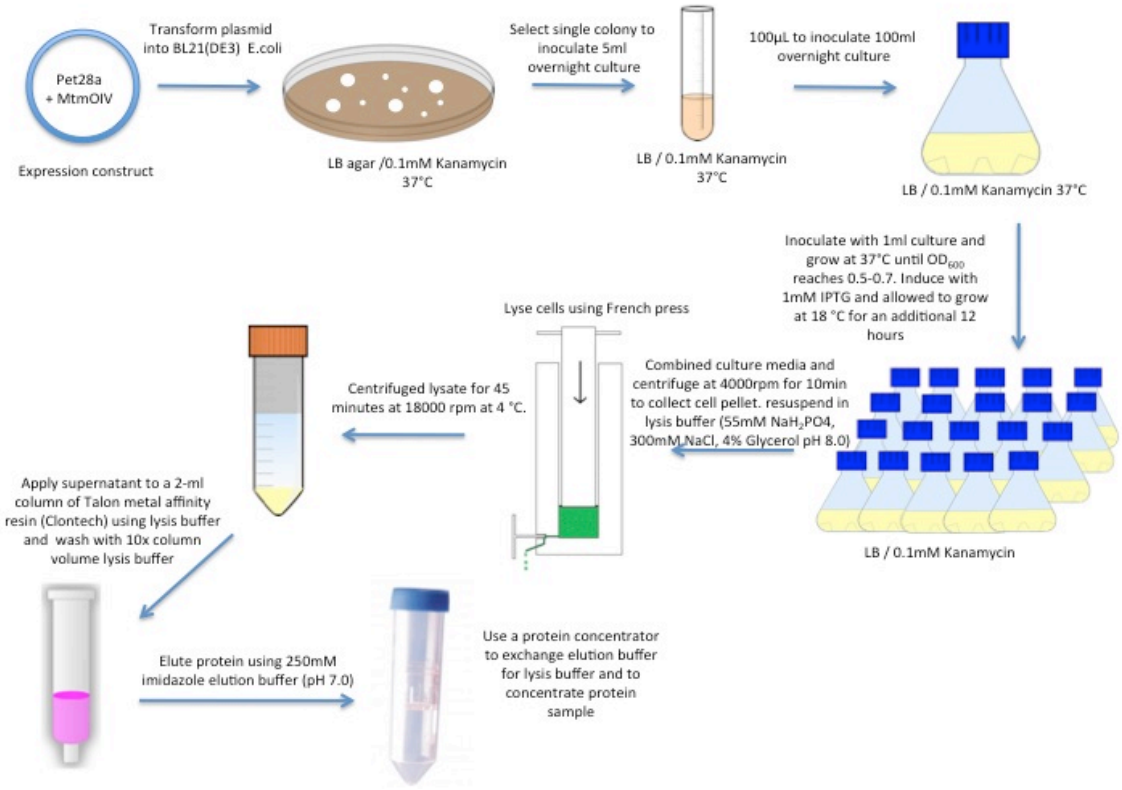


Figure 24. Diagram of MtmOIV protein expression and purification.

MtmOIV was cloned from genomic DNA using Pfu polymerase. The gene and Pet28a(+) vector were individually dually digested with *NdeI* and *EcoRI* restriction enzymes for 1 hour at 37 °C and purified using 8% agarose gel electrophoresis and Qiagen miniprep columns. 1 µl of digested vector and 9 µl digested insert gene were combined in a 20 µl ligation reaction with T4 ligase and allowed to incubate overnight at room temperature. Each ligation reaction was then transformed into XL1-blue chemically competent sells by heat shocking the cells combined with the 20 µl ligation reaction for

30 seconds at 42 °C. The cells were allowed to recover for 1 hour in 200 µl LB (Lysogeny Broth) at 37 °C and then plated on LB agar plates supplemented with 50µg/mL kanamycin for resistance selection. Each plate was incubated overnight at 37 °C and several of the resulting colonies were selected and grown overnight in 5 ml LB media supplemented with 50 µg/mL kanamycin. Plasmid was recovered from each of these 5 ml overnight cultures using Qiagen miniprep spin columns. 5 µl of each plasmid was digested with *NdeI* and *EcoRI* for 1 hour at 37 °C and checked for the correct size gene insert using 8% agarose gel electrophoresis. Digested plasmid with the correct size insert and vector backbone were sent to Genewiz Inc. for DNA sequencing using their stock T7 and T7 terminator sequencing primers and high gc content DNA sequencing protocol.

2.5.2 Protein Expression, Purification, and Crystallization

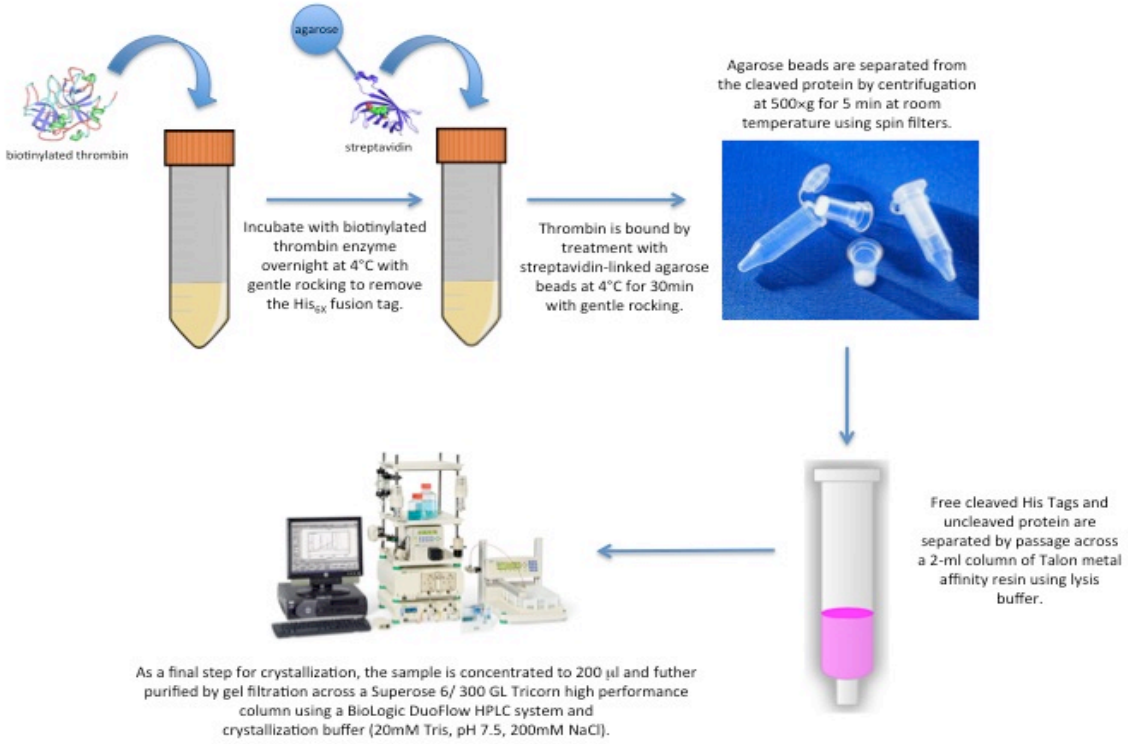


Figure 25. Diagram of histag cleavage and further purification for X-ray crystallography.

The cloning and overexpression of MtmOIV was performed basically as previously reported[44] except that we chose here to remove the N-terminal His_6-tag leading to better crystallization conditions. Briefly, the *mtmOIV* coding sequence (1.5 kb) was amplified and subcloned into pET28a(+) and expressed with an N-terminal His_6 fusion tag for purification and removed using a Thrombin Cleavage Capture Kit (Novagen) prior to crystallization. This construct was then transformed into BL21(DE3) cells. A 1 L culture using LB medium supplemented with 0.1 mM kanamycin was grown at 37 °C until the OD_{600} reached 0.5–0.7 and then was induced with 1 mM IPTG (isopropyl-β-d-1-thiogalactopyranoside) and allowed to grow at 18 °C for an additional 12 hours (figure 23).

For purification, cells were collected by centrifugation, resuspended in lysis buffer (50 mM KH_2PO_4, 300 mM KCl, 10 mM imidazole, pH 8.0), lysed by two passes through a French press, and then centrifuged for 45 min at 18000 rpm at 4 °C. The supernatant was then passed through a 0.45 μm syringe driven multigrade glass fiber filter (Millipore), applied to a Ni-NTA (nitrilotriacetic acid) column using a Profinia protein purifier (Bio-Rad), and eluted with 250 mM imidazole. Fractions containing MtmOIV were pooled and incubated with biotinylated thrombin (Thrombin Cleavage Capture Kit, Novagen) overnight at 4 °C with gentle rocking to remove the His_6 fusion tag. Thrombin was then removed by treatment with streptavidin-linked agarose beads at 4 °C with gentle rocking for 30 min. The beads were separated from the cleaved protein by centrifugation at 500 × *g* for 5 min at room temperature using spin filters (Novagen). Free cleaved His-tags and uncleaved protein were then separated by passage across a 2 mL column of Talon metal affinity resin (Clontech) using lysis buffer. As a final step for crystallization, the sample was concentrated to ~200 μL using a Ultracel 30K filter (Millipore) and further purified by gel filtration across a Superose 6/300 GL Tricorn high performance column (Amersham Biosciences) on a BioLogic DuoFlow HPLC system (Bio-Rad) and crystallization buffer (20 mM Tris, pH 7.5, 200 mM NaCl). Fractions were collected and concentrated to a final concentration of ~10 mg mL^{-1}. For the MtmOIV–premithramycin B (PMB) complex structure, PMB (5 mM) was added to the protein

solution (~200 μL) 1 h prior to concentration to 10 mg mL^{-1} for crystallization. Protein concentrations were determined by Bradford assay (Sigma) for crystallization and kinetic studies.

For crystallization, broad matrix screening was performed using a TTP LabTech Mosquito crystallization robot, and lead conditions were optimized accordingly. Final conditions for the native MtmOIV crystals consisted of 100 mM Bis-Tris 6.5 and 28% PEG MME (poly(ethylene glycol) monomethyl ether) 2000, while the final conditions for the MtmOIV–PMB co-crystals consisted of 200 mM ammonium acetate and 30% PEG 1000. The best crystals typically grew within 3–5 days at 21 °C, were harvested directly from the crystallization drop, and then were plunged into liquid nitrogen for storage until data collection.

2.5.3 Data Collection and Structure Determination

Lead conditions were initially screened using our in-house X-ray diffractometer (Rigaku MicroMax-007 HF microfocus X-ray generator, Raxis IV++ detector). Final data sets were collected at the SER-CAT (Southeast Regional Collaborative Access Team) beamline at the Advanced Photon Source at Argonne National Laboratory. All data were processed using HKL2000[83] and statistics are summarized in Table 1. Both the native MtmOIV and the MtmOIV–PMB complex structures were solved by molecular replacement using PHASER[84] within CCP4 (Collaborative Computational Project 4).[85] An initial search model was prepared using the previously deposited MtmOIV coordinates (PDB code 3FMW). A difference map clearly indicated the locations of the FAD molecules within the native MtmOIV structure and the locations of FAD and PMB molecules within the MtmOIV–PMB complex structure.

All model building was performed using COOT (Crystallographic Object-Oriented Toolkit),[86] and refinement (including automated water picking followed by visual inspection) was performed using PHENIX (Python-based Hierarchical Environment for Integrated X-ray crystallography)[87] and CCP4.[85] All figures were prepared using PyMOL

(Schrödinger), and final editing was done in Adobe Illustrator. RMSD analysis was performed within PyMOL (Schrödinger).

2.5.4 Modeling the Putative NADPH Binding Site

We were able to grow crystals of MtmOIV in the presence of NADPH. These crystals diffracted to ~3.5 Å, and while we were able to observe extra density indicating the presence of NADPH, the density was mostly disordered and incomplete, preventing us from modeling the ligand accurately and with confidence. Nevertheless, on the basis

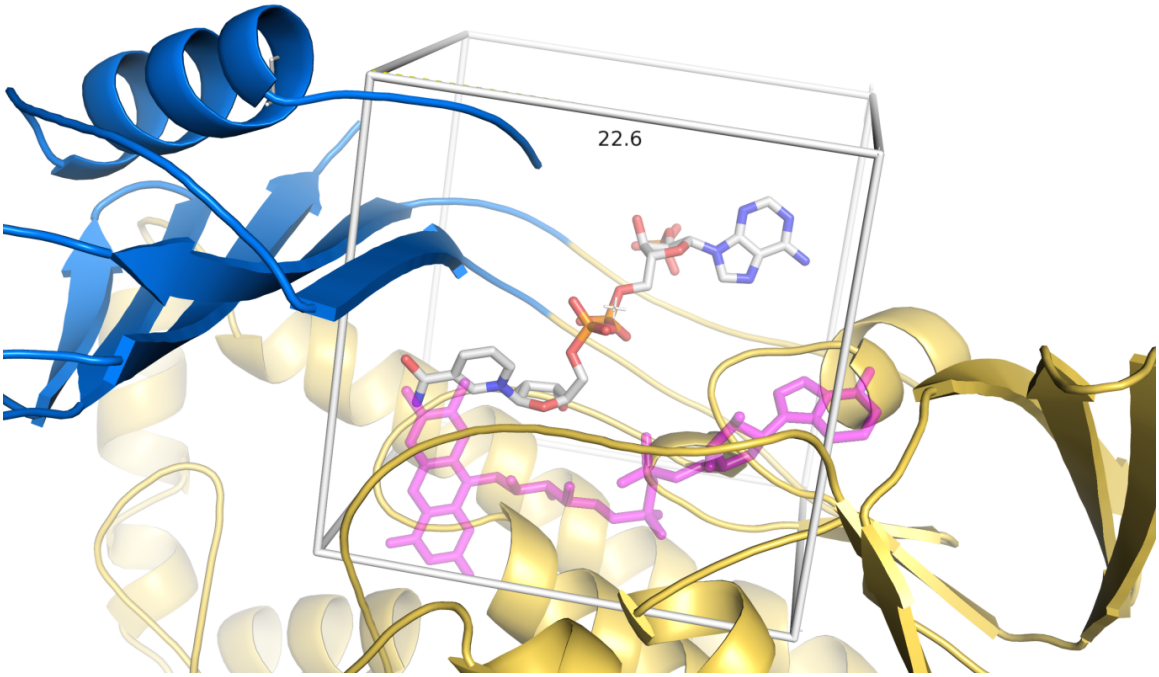


Figure 26. NADPH docking box used in the Autodock procedure to model the putative NADPH binding site of MtmOIV.

of the location of the difference density, we were able to identify four residues at this site that interacted with the putative NADPH molecule, namely, R169, R173, R174, and R277. Mutagenesis of three of these residues proved to have a major impact on catalysis. Further, a disordered loop containing residues 233–239 is located along this putative NADPH binding site and has been suggested to play a role in binding NADPH. This information has allowed us to propose the putative NADPH binding site, which in comparison with the substrate binding site sits on the opposite side of the MtmOIV

structure, yet still in proximity to the FAD cofactor. To model this, we used the NADP(+)-bound crystal structure of cyclohexanone monooxygenase (PDB code 3GWF) and aligned to MtmOIV along the FAD cofactor, which positioned the NADP(+) molecule in the proximity of the putative NADPH binding pocket in a 23Å x 23Å x 23Å box with 0.386Å grid spacing and a gridcenter at -23.082, 46.586, -114.336. We then added NADPH to our structure at this site and manually optimized the fit to agree with the observable density in our low-resolution NADPH-bound structure. This was used as a starting point for docking studies to optimize the NADPH position within the binding pocket. Here, we used AutoDock 4.2[88–90] and the AutoDockTools[89,91–93] to perform the docking analysis using the FAD-bound MtmOIV crystal structure as the receptor molecule and NADPH as the ligand molecule. Based on the low-resolution co-crystal structure with NADPH, the putative NADPH binding loop was added to the receptor molecule prior to the analysis. A grid was calculated about the placed ligand, and an autodocking procedure was performed using the Local Search algorithm with default parameters that preform 300 iterations of lowest energy calculations. The docking analysis produced coordinates for the top 50 solutions, which were further sorted on the basis of lowest energy, with the top solution, having the lowest energy of −2.8 kcal/mol, being used for subsequent analysis.

2.5.5 Mutagenesis and Kinetics Assays

MtmOIV mutants were prepared using an Agilent Quikchange site-directed mutagenesis kit. Primers used for mutagenesis are shown in Table 4. All kinetic assays were repeated in triplicate with no less than two separate purifications of the enzyme. The kinetic assays utilized here with respect to the premithramycin-B substrate have been previously reported.[44] Briefly, we monitored the conversion (100 mM Tris-HCl pH 8.25, 30 °C, 10 min) of premithramycin B by MtmOIV in a continuous assay measuring the oxidation of NADPH at λ = 340 nm (ε 340 = 6220 M^{-1} cm^{-1}) in the presence of FAD and O_2 (0.25 mM NADPH/NADH; 0.1 mM FAD added; open cuvette/96 well plate). The reactions were initiated by addition of 0.25, 1, 5, or 10 μM wild-type or mutant MtmOIV. Kinetic parameters were determined by fitting with Kaleidagraph 4.0 (Synergy).

The kinetic assays utilized with respect to the substrate NADPH were monitored via a fluorescent assay due to the inability to monitor the reaction at high concentrations of NADPH by absorbance because of absorbance values exceeding the detection limits of available spectrophotometers. Fluorescence spectroscopy was used to monitor the oxidation of NADPH in the presence of FAD and O_2 (2 mM Pre-B; 0.1 mM FAD added; open 96-well plate). This conversion was continuously monitored at an excitation $\lambda = 355$ nm and an emission $\lambda = 460$ nm (100 mM Tris-HCl pH 8.25, 30 °C, 10 min). The excitation $\lambda = 355$ nm was used instead of the typical $\lambda = 340$ nm due to the available equipment filter selection. Assays were preformed utilizing the FLOUstar Omega plate reader (BMG LABTECH) and nontreated white optical flat bottom nonsterile 96-well plates (Nunc). The reactions were initiated by addition of 0.25, 1, 5, or 10 μM wild-type or mutant MtmOIV. To determine the value of NADPH oxidation a standard curve was created utilizing known concentrations of NADPH. Kinetic parameters were determined by fitting with Kaleidagraph 4.0 (Synergy)

Table 3. Mutagenesis primrs

Mutagenesis Primers	
R169A forward	5' -GCG ACG GCG GGG CCA GCA CGG TGC- 3'
R169A reverse	5' -GCA CCG TGC TGG CCC CGC CGT CGC- 3'
R173A forward	5' -GCAGCA CGG TGG CGC CGC GGC TGG CCG- 3'
R173A reverse	5' -CGG CCA GCC GCG CCA CCG TGC TGC- 3'
P84A forward	5' -CCA AGG GGC TGG CCT TCG CCG GG-3'
P84A reverse	5' -CCC GGC GAA GGC CAG CCC CTT GG- 3'
F89A forward	5' -CTT CGC CGG GAT CGC CAC CCA GGG CTT G- 3'
F89A reverse	5' -CAG GCC CTG GGT GCC GAT CCC GGG GAA G-3'
R204A forward	5' -GAG GTG CCG CGC GCC TGG GAG CGC AC- 3'
R204A reverse	5' -GTG CGC TCC CAG GCG CGC GGC ACC TC-5'
H50A forward	5' -ACG CCG TCG GCG CCG ACC GGG CGG-3'
H50A reverse	5' -CCG CCC GGT CGG CGC CGA CGG GCT-3'
D51A forward	5' -GTC GGC CAC GCC CGG GCG GGG- 3'
D51A reverse	5' -CCC CGC CCG GGC GTG GCC GAC- 3'
W205A forward	5' -TGC CGC GCC GCG CGG AGC GCA CCC- 3'
W205A reverse	5' -GGG TGC GCT CCG CGC GGC GCG GCA- 3'
R225A forward	5' -GCG GCC TTG GGG CCG GGT GGT CG-3'
R225A reverse	5' -CGA CCA CCC GGC CCC AAG GCC GC- 3'

2.6 Reflections and Future Directions

If the experiments preformed in this chapter were to be repeated several changes could be made to improve the validity of the findings presented. One major change would be to perform all the kinetic analysis using the fluorescence based method used when studying MtmOIV's NADPH binding dicussed in section 2.5.5 This would allow the experimenter to use a much higher concentration of NADPH when studying PreB binding and produce more universally accepted K_M values so that MtmOIV's catalytic efficiency could be compared to other enzymes and not just between wildtype and the generated mutants. Another point of improvement would be to monitor the uncoupling of PreB turnover from NADPH oxidation. This could be accomplished my monitoring the

creation of hydrogen peroxide generated when uncoupling in present using one of the many commercially available horseradish peroxidase assays. A third point of improvement entails verifying that the mutations introduced into MtmOIV's amino acid sequence are not disrupting then enzymes overall structure. One crude way to perform this analysis would be to utilize circular dichroism to compare the specta of mutant enzymes to the wildtype enzyme. This would be a fast and inexpensive way to verify the mutant enzymes were folded correctly and that the overall tertiary structure was undisrupted. A second more thorough approach to this problem would be to crystallize and solve the X-ray crystal structure of each generated mutant with substrate. In addition to ensuring the overall structure of the enzyme was intact, an intricate view of what interactions each mutation may be disrupting would be gained. However, this method would be relatively expensive and very time consuming.

The next major step in this project would be to obtain soluble purified MtmW enzyme to confirm the existence of the theorized codependence between MtmOIV and MtmW. Though production of soluble MtmW has been attempted numerous times it has yet to be successfully obtained. If one were to continue this project, expression of MtmW in Streptomyces using vector pXY200 is a promising approach to accomplishing this task. Table 4 lists unsuccessful attempts at producing soluble MtmW. Additionally, the plasmid maps associated with these attempts are shown in the appendix of this dissertation.

Table 4. MtmW Expression Constructs

	Expression Construct	**Organism**	**Inclusion Bodies**
1	MtmW-Pet28a- N His6x	E. coli	Yes
2	MtmW-Pet28a- C His6x	E. coli	Yes
3	MtmW codon optimized- Pet28a- N His6x	E. coli	Yes
4	MtmW codon optimized- Pet28a- C His6x	E. coli	Yes
5	MtmW-PetSUMO- N His6x	E. coli	Yes
6	MtmW-PetNusA- N His6x	E. coli	Yes
*Varying IPTG concentrations and induction temperature were attempted for each construct. Refolding of inclusion bodies was also attempted in L-arginine rich buffer.			

This chapter has been reproduced with permission from the following publication: copyright ACS publications

Mary A. Bosserman*[+], **Theresa Downey*[&]**, Nicholas Noinaj*[#], Susan K. Buchanan, and Jürgen Rohr. "Molecular Insight into Substrate Recognition and Catalysis of Baeyer–Villiger Monooxygenase MtmOIV, the Key Frame-Modifying Enzyme in the Biosynthesis of Anticancer Agent Mithramycin" *ACS Chemical Biology* 2013 *8* (11), 2466-2477.

*These authors contributed equally to this this work

[+]Production of original MtmOIV-Pet28a cloning vector and site-directed mutagenesis of active site mutations.

[&]Site-directed mutagenisis of NADPH cofactor binding site mutations and P84A/F89A double mutant. All kinetic assays and data analysis. MtmW expression attempts.

[#]Crystalization, X-ray crystallography, and Autodocking experiments.

3 Role of the Synergistic Reductive *O*-Methyltransferase GilM of the Gilvocarcin Biosynthetic Pathway and Investigation of Enzyme-Enzyme Interactions with its Co-dependent Partner GilR.

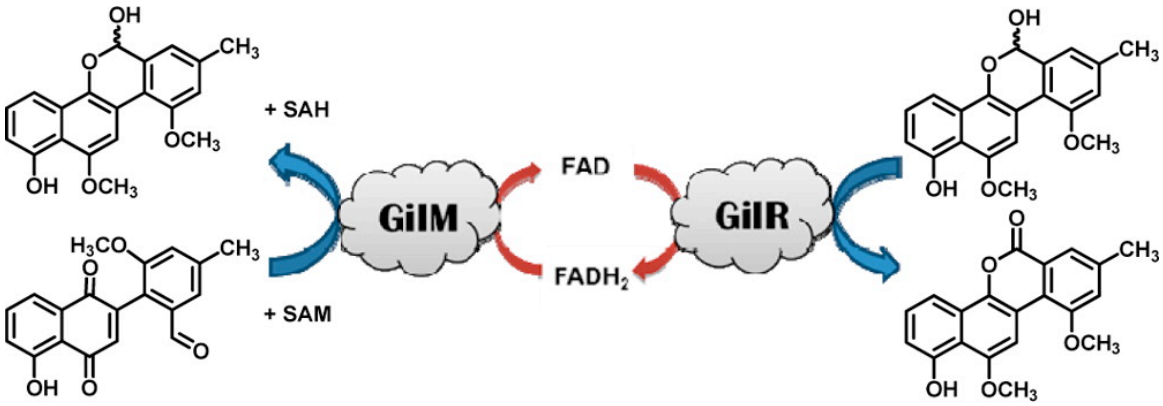


Figure 27. Overall representation of the co-dependent GilR-GilM catalytic cascade.

3.1 Abstract

GilM an enzyme from the gilvocarcin biosynthetic pathway with an unknown function was investigated *in vitro* utilizing purified recombinant enzymes and a synthetically prepared pathway intermediate. This study reveals GilM as a key enzyme of the gilvocarcin pathway exhibiting dual functionality as a reductase and *O*-methyltransferase responsible for both the reduction of a quinone intermediate to a hydroquinone and the associated stabilizing *O*-methylation and hemiacetal formation. Additionally, GilM was found to mediate its reductive catalysis through the assistance of GilR. GilR provides its bicovalently bound $FADH_2$ cofactor for the GilM reaction. In turn GilM regenerates FAD for GilR's next catalytic cycle establishing a unique co-dependent relationship. This unusual synergy eventually completes the biosynthesis of the polyketide-derived defuco-gilvocarcin chromophore. Additionally, the necessary enzyme-enzyme interaction between GilM and GilR was probed revealing the absence of a long-lived interface between the two, alluding to a transient relationship during catalysis.

3.2 Introduction

The gilvocarcins belong to the angucycline group of natural products. Angucylines are the largest group of polyketide-derived natural products and are rich in biological activities and interesting chemical scaffolds. The unique oxidative scaffold rearrangements and other post-PKS tailoring events were what first attracted researchers to the group.[94] The gilvocarcin family of angucyclines were originally discovered in the mid-1950s.[95] However, their structures were not solved until 1981 when X-ray structure[96,97] was able to determine relative conformation of gilvocarcin M, leading to the structural elucidation of other group members.[97–100] The key characteristic of the gilvocarcin family is the polyketide-derived benzo[**d**]naphtho[1,2-**b**]pyran-6-one chromophore and variable *C*-glycosidically linked sugar moiety.[101–105] Gilvocarcins are known for their strong antitumor activities, low toxicities, and unique mode of action.[99,103,106] The benzo[**d**]naphtho[1,2-**b**]pyran-6-one core's 8-vinyl side chain undergoes photoactivated [2+2]-cycloaddition with DNA's thymine bases under irradiating conditions with low energy UV or visible light.[107–109] Additionally, the sugar moiety is essential for observed gilvocarcin-mediated cross-linking of histone H3 or heat shock protein GRP78 with DNA resulting in the disruption of DNA replication and transcription.[110] In addition to anticancer activity gilvocarcins exhibit strong antibacterial[95] and antiviral properties.[111] In spite of these many positive properties, poor solubility remains a major impediment toward development into useful therapeutics.[112,113]

The aromatic core of the gilvocarcins is produced by a type II polyketide synthase (PKS),[114,115] initially yielding an angucycline biosynthetic intermediate (e.g., **3** in figure 26). Oxidative post-PKS modifications convert this intermediate into the unique defuco-gilvocarcin chromophore.[116] The role of many of the gilvocarcin post-PKS tailoring gene products were assigned based on gene inactivation, cross-feeding and labeling studies.[36,114–119] In fact most of the genes of the three known biosynthetic clusters[120] of this group have been assigned to certain enzymatic activities with *gilM* (encoding an enzyme of unknown function) being one of a few a notable exceptions as this gene did not bare high sequence similarity to other know sequences. Thus, the function of its gene

product remained obscure despite appearing to be crucial for the biosynthesis of gilvocarcin, as mutant strains created by deletion of this individual gene did not accumulate any isolable metabolites. Here we report the assignment of the GilM enzyme as a dual function reductase and *O*-methyltransferase responsible for both the reduction of a quinone intermediate to a hydroquinone and the associated stabilizing *O*-methylation and hemiacetal formation. It was also discovered that GilM requires the assistance of oxidoreductase GilR to mediate its reductive functionality. GilR provides its bicovalently bound $FADH_2$ cofactor for the GilM reaction. In turn, GilM regenerates FAD for GilR's next catalytic cycle establishing a unique co-dependent relationship. This unusual synergy eventually completes the biosynthesis of the polyketide-derived defuco-gilvocarcin chromophore. Additionally, the proposed enzyme-enzyme interaction between GilM and GilR was investigated and did not reveal a prolonged interaction between the two, thus indicating their co-dependent relationship may be transient in nature.

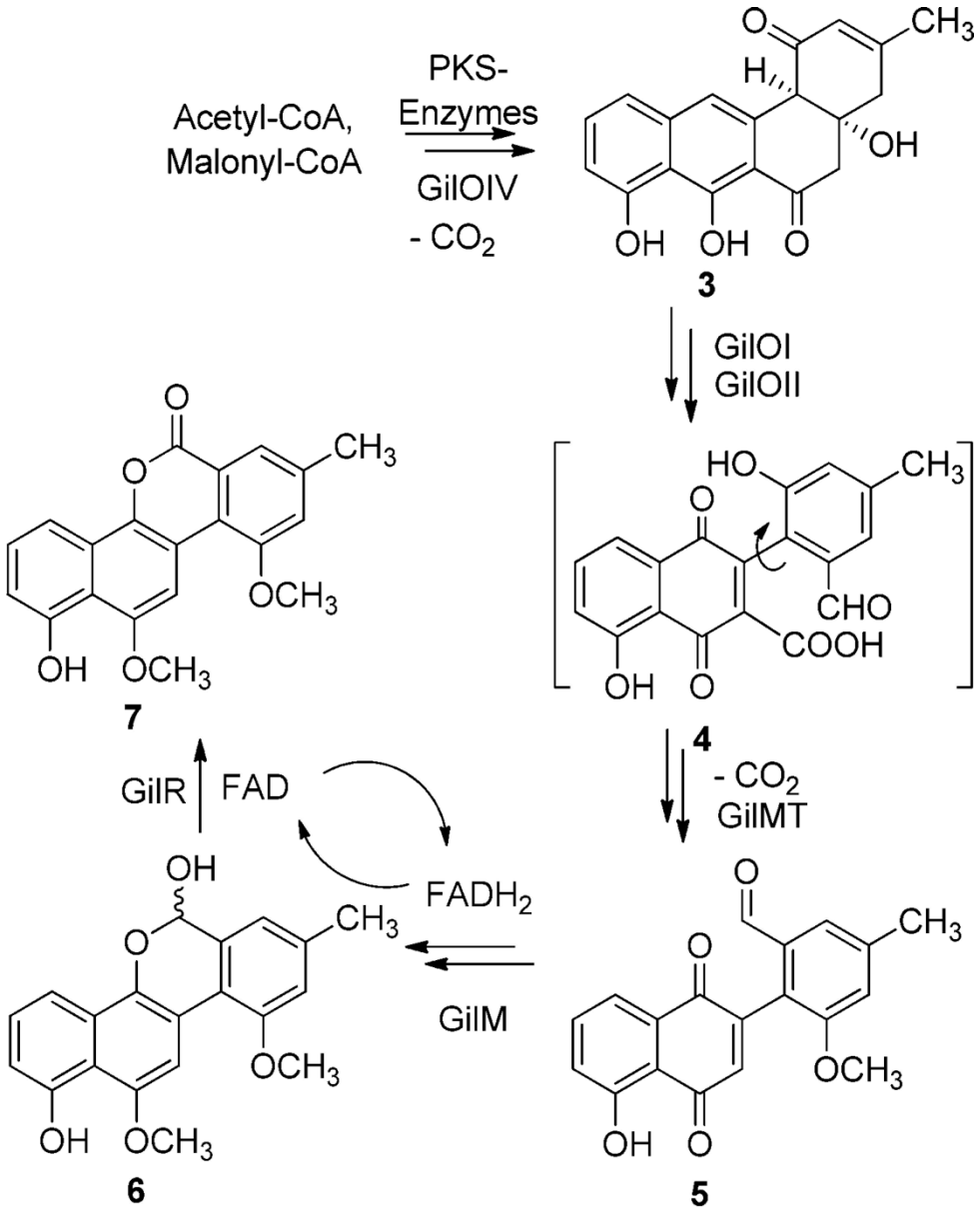


Figure 28. Key oxidative rearrangement and follow-up sequence of events involving GilM and GilR reactions en route to defuco-gilvocarcin M (7).

3.3 Results and Discussion

3.3.1 Determing GilM's Catalytic Activity

BLAST analysis[121,122] revealed GilM to have low similarity (34-37% sequence identity) to nucleotidyl-*S*-transferases, such as thiopurine-*S*-methyltransferases from *Rhodococcus equi* or *Mycobacterium avium,* and benzoquinone methyltransferases from *Mycobacterium tuberculosis* (33% sequence identity). The conceptual translated amino acid sequence of *GilM* contains the amino acid sequence VLDLGCGLG at position 49-57, which is predicted to be a SAM binding motif following the general pattern

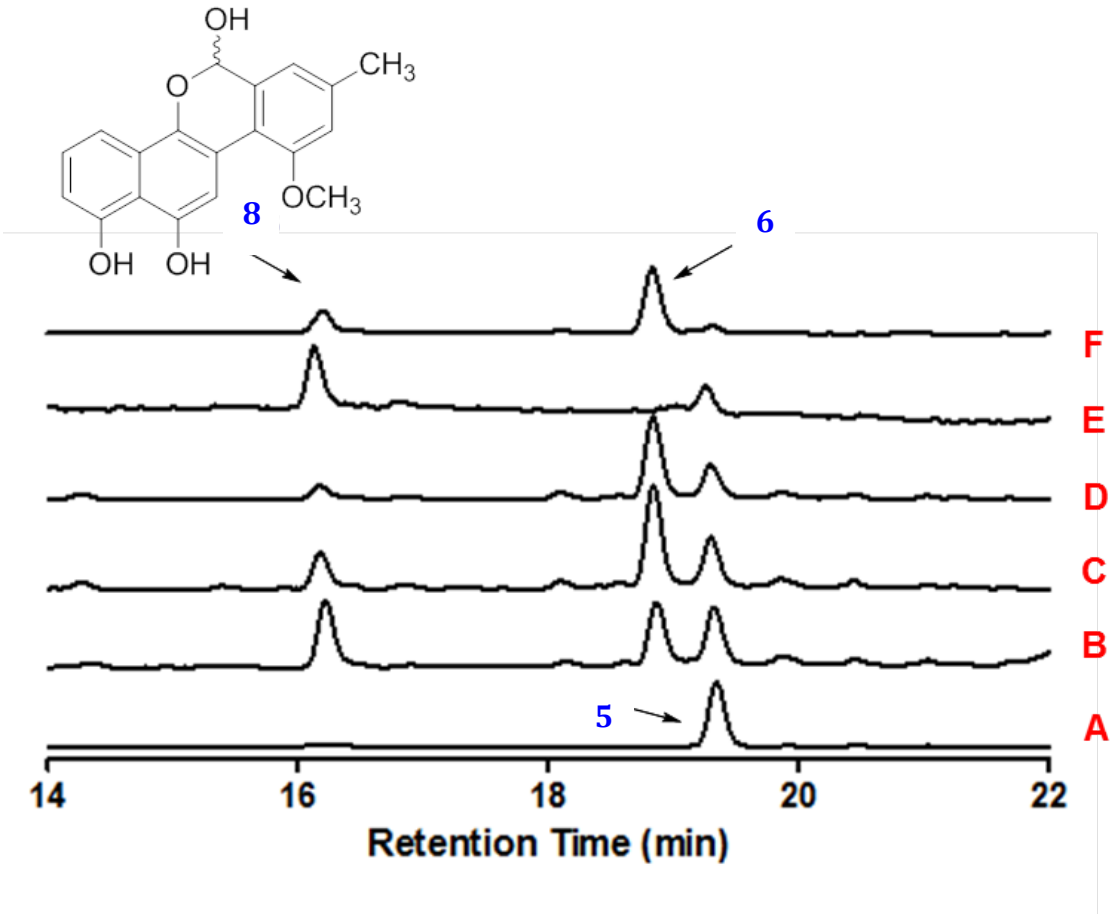


Figure 29. HPLC traces of the enzymatic reactions with 5. (A) Standard 5; (B) 5 + GilM + SAM + 5 min; (C) 5 + GilM + SAM + 10 min; (D) 5 + GilM + SAM + 15 min; (E) intermediate 8 isolated from HPLC (mixture of closed and open form); (F) 8 + GilM.

hh(D/E)hGXGXG (where h represents a hydrophobic residue). To determine the function of GilM compound **5** was chemoenzymatically produced by co-worker Nidhi Tibrewal and used as substrate. This reaction accumulated compound **6** proving GilM was responsible for quinone reduction, hemiacetal formation, and *O*-methylation (figure 27, trace D). Surprisingly, the GilM reaction seemed to be able convert **5** to **6** even in absence of SAM, which prompted further investigation of GilM for any bound SAM cofactor.

To further analyze the reaction sequence catalyzed by GilM, compound **5** was incubated with GilM and the reaction was studied at different time points (Figure 27). With stoichiometric enzyme quantities, the substrate was 80% converted into the product after 15 min (Figure 27, trace D). When the reaction was analyzed by reversed phase HPLC-MS in 3–5 min intervals, a new peak (**16**) appeared with shorter retention time than the overall product **6** (Figure 27). Although NMR analysis of **8** was impossible due to its instability, LC-MS analysis suggested it to be demethyl-defuco-pregilvocarcin M (*m*/*z* 323 $[M - H]^-$). To prove that **8** was an intermediate en route to **6** and not a shunt product, **8** was incubated with GilM, and, as anticipated, was rapidly converted to **6** (Figure 27, trace F). Overall, the results show that GilM catalyzes a sequence of reactions: (i) quinone reduction, (ii) hemiacetal formation and (iii) *O*-methylation, to construct the tetracyclic core of the gilvocarcins.

3.3.2 Identification of GilM's SAM cofactor

GilM catalyzes reductive O-methylation without the addition of any external cofactor. However, the BLAST analysis showed a very vague similarity to thiopurine-S-methyltransferases as previously discussed. To analyze any bound cofactor, the enzyme was boiled for 5 minutes and centrifuged (12000 ×g, 5 min). The supernatant was then subjected to LC-MS analysis. A linear gradient of acetonitrile and 0.1% formic acid-water (solvent A = 0.1% formic acid-H_2O; solvent B = acetonitrile; 0-15 min 25% B to 100% B; 16-24 min 100% B; 25-26 min 100% to 25% B; 27-29 min 25% B) with flow rate of 0.5 mL/min was used to separate the compounds in a Waters Symmetry C_{18} (4.6 × 250 mm, 5 μm) column. The supernatant showed UV-absorption at λ=260 nm, typical of

an adenosine spectrum, indicating co-purified SAM. To further verify presence of loosely bound *S*-adenosyl methionine, a standard solution of *S*-adenosyl methionine was prepared in 50 mM phosphate buffer and was used in parallel for comparison. GilM was found to be co-purifying with *S*-adenosyl methionine (Figure 28).

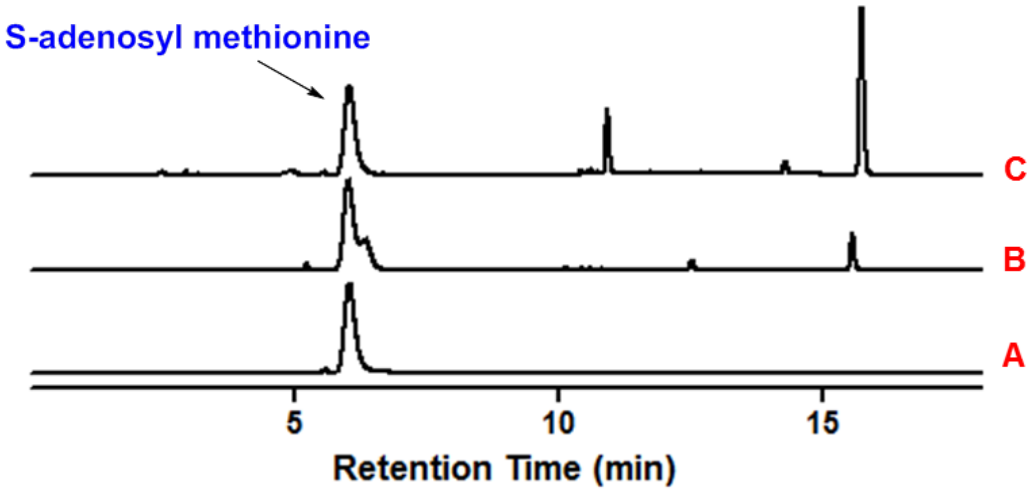


Figure 30. HPLC traces of the released cofactor: (A) standard *S*-adenosyl methionine; (B) Cofactor released from GilM; (C) *S*-adenosyl methionine boiled for 5 min.

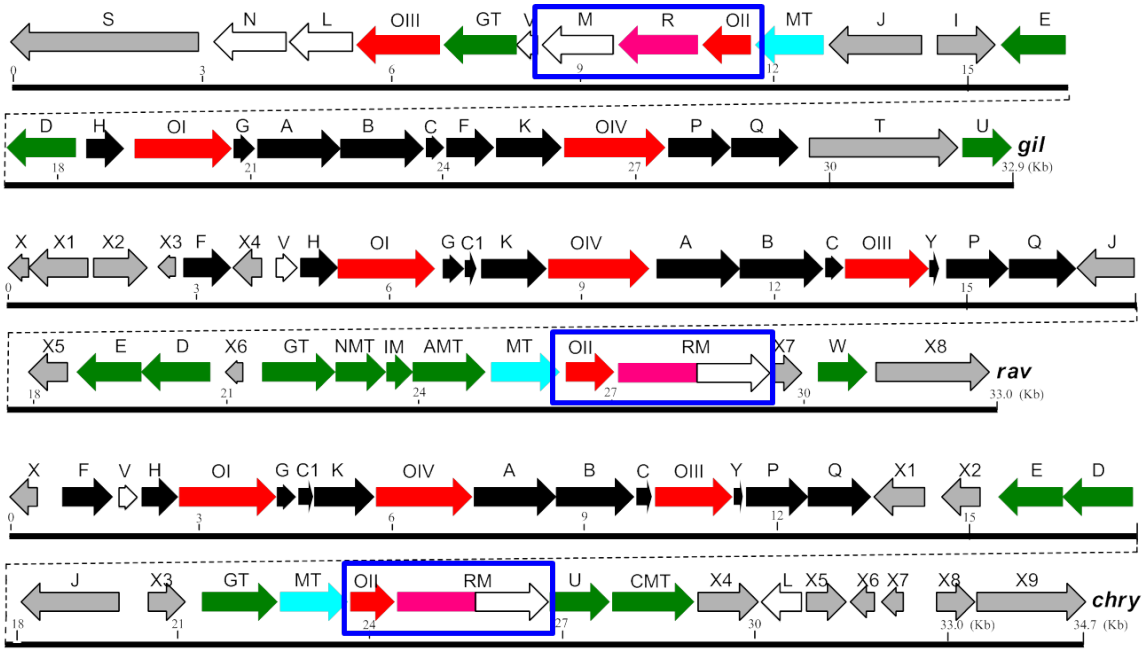


Figure 31. R and M genes in gilvocarcin (gil), ravidomycin (rav) and chrysomycin (chry) gene clusters.

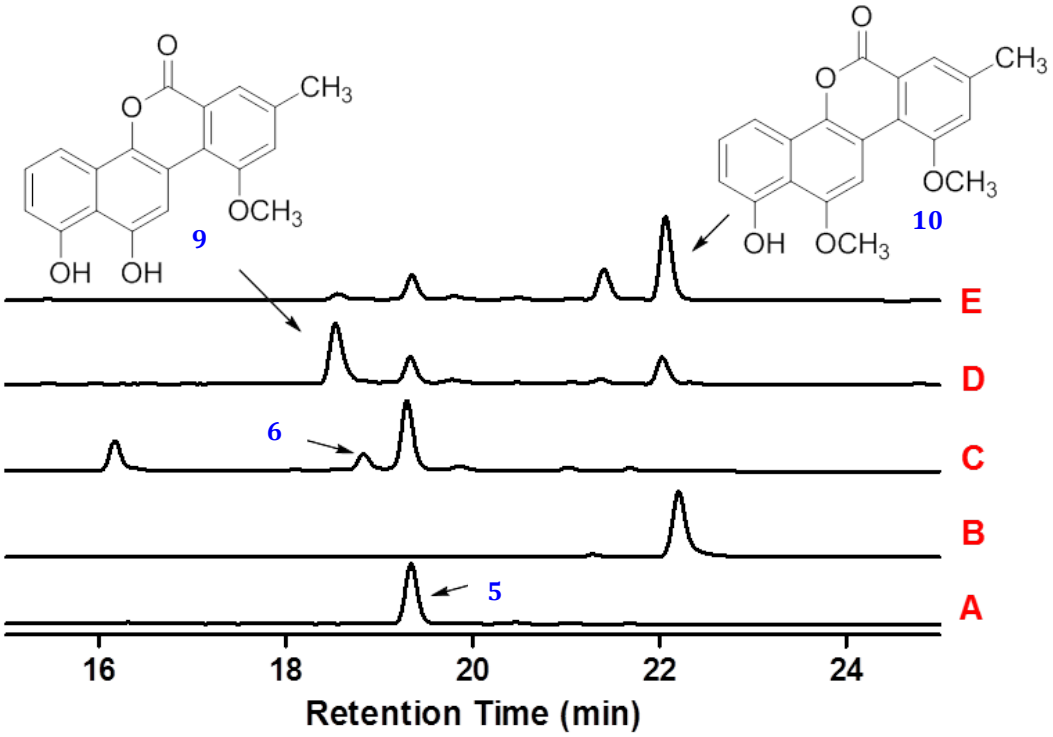


Figure 32. HPLC traces of the enzymatic reactions with 5: (A) standard 5; (B) standard defuco-gilvocarcin M (45); (C) 5 + GilM; (D) 5 + GilM + GilR; (E) 5 + GilM + GilR + SAM.

3.3.3 Co-dependence between GilM and GilR

One remaining question was how the regeneration of GilM occurred after the reduction of the quinone. Interestingly, the homologous GilM-activity from the biosynthetic pathways of other structurally similar compounds, chrysomycin A (chry) and ravidomycin V (rav), is encoded on the same gene as the oxidoreductase GilR-activity,[120] while in the gilvocarcin pathway, *gilM* and *gilR* are encoded by two separate genes (Figure 29).[114] This prompted the proposal that GilM could be working in conjunction with GilR, an FAD dependent oxidoreductase that catalyzes the very last step in gilvocarcin biosynthesis by converting pregilvocarcin to gilvocarcin. Additionally, GilR was shown to be able to convert sugar free defuco-pregilvocarcins.[37,119] We hypothesized that GilM utilizes the reduced flavin generated in the GilR reaction to reduce the quinone thereby regenerating oxidized flavin for the next catalytic cycle. To validate this hypothesis, compound **5** was incubated with GilM and GilR (Figure 30). The assistance of GilR for the reducing capability of GilM was established by measuring the amount of product formed during the reactions. Catalytic amounts of GilM alone mainly accumulated starting material with a minor production of **6** (Figure 30, trace C). In the absence of SAM, GilM and GilR accumulated a new peak (Figure 30, trace D) that corresponds to demethyl-defuco-gilvocarcin M[123] (**9**, *m*/*z* 321 $[M-H]^-$). Adding SAM to the GilM-GilR reaction mixture led to the accumulation of **10** (Figure 2.5, trace E). A 10-fold increase in the formation of **10** in the GilM-GilR reaction versus using GilM alone confirms that GilM works synergistically with GilR.

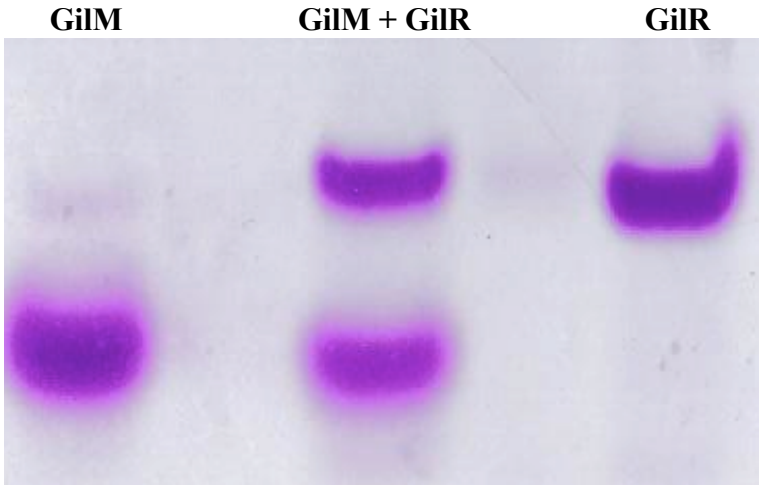


Figure 33. Native PAGE probing for GilM-GilR enzyme-enzyme interaction.

3.3.4 Investigating GilM-GilR Enzyme-Enzyme Interactions

GilM's synergistic activity with GilR necessitates the presence of an intimate enzyme-enzyme interface as the shared FAD cofactor is known to be bicovalently bound to GilR.[37] To investigate this enzyme-enzyme interaction two experiments were preformed to test it stability. First, equal amounts of heterologously expressed and purified GilM and GilR were mixed, incubated at room temperature for 10 minutes, and then subjected to native polyacrylamide gel electrophoreses (PAGE). When compared to controls of GilR and GilM alone, the enzyme mixture did not appear to form a stable complex as no new bands indicating complex formation where see in the native PAGE results (figure 31).

To confirm this finding was not due to the limitations of the native PAGE technique a second study was conducted utilizing size exclusion fast protein liquid chromatography (FPLC). When equal amounts of GilM and GilR where mixed in the presence of SAM and subjected to FPLC purification and the FPLC chromatogram of the GilM-GilR mixture was compared to the FPLC chromatograms of either enzyme alone no evidence of a GilM-GilR complex was seen (figure 32). The lack of evidence of complex formation in both native PAGE and FPLC enzyme interaction studies combined with the knowledge that GilM and GilR must cooperate suggests that the GilM-GilR

interaction is transient in nature or requires different condition to initiate and/or stabilize complex formation.

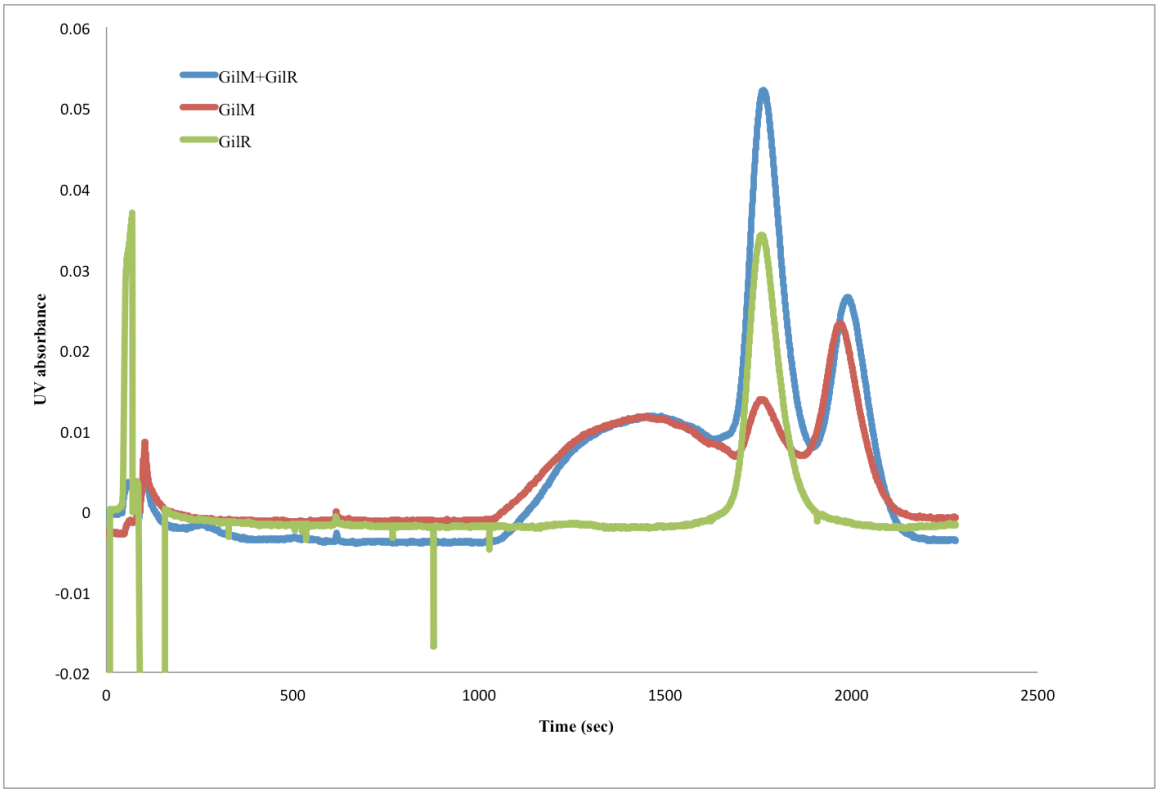


Figure 34. FPLC chromatograms of GilM+GilR (blue), GilM (red), and GilM (green).

3.4 Conclusions

The gene product of *GilM* from the gilvocarcin biosynthetic gene cluster was identified via *in vitro* characterization as a SAM-dependent multifunctional enzyme responsible for the reduction of a quinone intermediate to a hydroquinone and the associated stabilizing *O*-methylation and hemiacetal formation. This enzyme is one of only a limited number of bifunctional post-PKS tailoring enzymes identified to date as discussed in Chapter 1. In addition to being identified as a bifunctional enzyme, GilM was also identified as a co-dependent enzyme requiring the assistance of the oxidoreductase GilR to function. GilM is hypothesized to utilize GilR's bicovalently bound reduced flavin generated in the GilR reaction for quinone reduction thereby regenerating oxidized flavin for the next catalytic cycle. To further interrogate the GilM-GilR relationship two different enzyme-enzyme interaction studies were conducted. These studies led to the conclusion that the GilM-GilR interaction is transient as no enzyme complex was detected in either case. The identification of GilM's catalytic functionality and its co-dependent relationship with GilR elucidates the end of the gilvocarcin biosynthetic pathway and establishes the groundwork for future combinatorial biosynthetic work that may lead to the production of new gilvocarcin analogues with improved druggability. Further investigation of the GilM-GilR interaction is needed to fully understand the complexity of their relationship. The X-ray crystal structure of GilR has previously been solved and the crystallization and structure elucidation of GilM has been commenced. Once GilM's structure has been solved docking studies or even co-crystallization of both enzymes may reveal more details about this curious transient relationship.

3.5 Methods

3.5.1 Cloning of GilM and GilR

For GilR pfu-polymerase (Stratagene) was used to amplify the ~1.5 kb nucleotide sequence of *gilR* from the cosmid G9B3 using the respective set of primers GilR_for: 5'-CGCCATATGACCGCTTCCGTACCGCCGTTCACGGTG-3', and GilR_rev: 5'-CCAGAATTCTCAGAGTCCTATGGACATGCTGTG-3'). The generated PCR product

was cloned into Zero Blunt® TOPO® vector (Invitrogen), and the positive clones were sequenced to make sure that no mutation had been incorporated during amplification. *GilR* was recovered as an *Nde*I-*EcoR*I fragment from the TOPO-GilR construct and ligated at the identical sites of pET28a(+) (Novagen) to generate a GilR expression construct.

For GilM pfu-polymerase (Stratagene) was used to amplify the ~800 bp nucleotide sequence of *gilM* from the cosmid G9B3 using the respective set of primers GilM_for: 5'-CGCCATATGCCAACGGGCAGCACG -3'; and GilM_rev: 5'-CCAGAAATCGGAGAGCCCGGCATGA -3'). The generated PCR product was cloned into Zero Blunt® TOPO® vector (Invitrogen), and the positive clones were sequenced to make sure that no mutation had been incorporated during amplification. *GilM* was recovered as an *Nde*I-*EcoR*I fragment from the TOPO-GilM construct and ligated at the identical sites of pET28a(+) (Novagen) to generate a GilM expression construct.

3.5.2 Expression and Purification of Enzymes

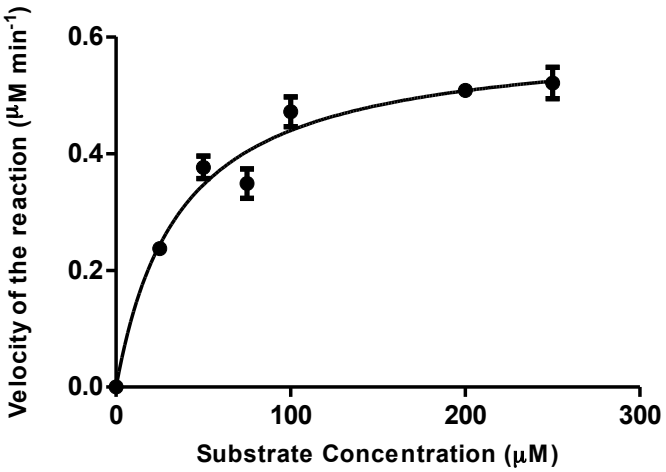
gilM and *gilR* genes were expressed using pET28a(+) expression constructs in *E. coli* BL21 (DE3) with a 6x *N*-terminal polyhistidine tag. A single colony was transferred to 10 mL LB supplemented with 50 µg/mL kanamycin and grown at 37 °C and 250 rpm for 12 h. Subsequently, 500 mL LB supplemented with 50 µg/mL kanamycin was inoculated with 5 mL of the culture and was grown at 37 °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 allowed to grow at 18 °C for 16 h. The cell pellets were collected by centrifugation (4000 × g, 15 min) and were washed twice with 20 mL of lysis buffer (50 mM KH_2PO_4, 300 mM KCl, 10 mM imidazole, pH 8.0). The cells were lysed using a French Press and the crude soluble enzyme fractions were collected through centrifugation (16000 × g, 1h). The crude enzymes were loaded onto a Talon metal affinity resin (BD Biosciences) column and were washed three times with lysis buffer. The enzymes were then eluted with elution buffer (50 mM KH_2PO4, 300 mM KCl, 250 mM imidazole, pH 8.0). The purified proteins were concentrated using an Amicon Ultra centrifugal filter (Millipore Corp.) and stored as 25% glycerol stocks at -20

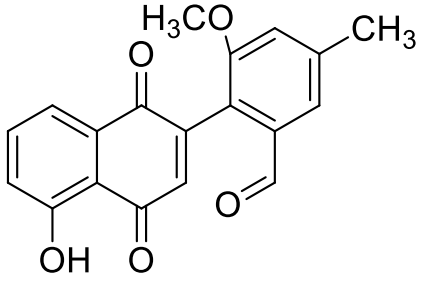
°C. The concentration of protein was determined by the Bradford method using a calibration curve from known concentrations of BSA.

3.5.3 Cofactor Analysis of GilM

GilM catalyzes reductive methylation without any external cofactor. However, the BLAST analysis showed a very vague similarity to thiopurine-S-Methyltransferase. To analyze any co-purified cofactor, the enzyme was boiled for 5 minutes and centrifuged (12000 ×g, 5 min). The supernatant was then subjected to LC-MS analysis. A linear gradient of acetonitrile and 0.1% formic acid-water (solvent A = 0.1% formic acid-H_2O; solvent B = acetonitrile; 0-15 min 25% B to 100% B; 16-24 min 100% B; 25-26 min 100% to 25% B; 27-29 min 25% B) with flow rate of 0.5 mL/min was used to separate the compounds in a Waters Symmetry C_{18} (4.6 × 250 mm, 5μm) column. The supernatant showed UV-absorption at λ=260nm, typical of adenosine spectrum. To further verify presence of loosely bound S-adenosyl methionine a standard solution of S-adenosyl methionine was prepared in 50 mM phosphate buffer and was used in parallel for comparison. GilM was found to be co-purifying with S-adenosyl methionine.

3.5.4 GilM Kinetic Profile





5

Figure 35. GilM kinetic profile with substrate 5.

A typical reaction mixture (50 μM) composed of substrate (**92**), 50 mM phosphate buffer, 20 μM enzyme (final concentration) was incubated at 25 °C. After 5 min reaction

was extracted twice with EtOAc (300 µl) and combined organic layers were dried under vacuum. The residue was then dissolved in 50 µl acetonitrile and 20 µl was injected onto HPLC following the protocol described in Section 3.5.3. The amount of product formed was estimated by plotting the peak area in a standard calibration curve. The data resulting from incubating seven different substrate concentrations with enzyme were fit to the Michael-Menten equation with nonlinear regression. k_{cat} and K_M values were calculated using GraphPad Prism 5.0. The analysis was done in triplicate and the average was taken. The analysis yielded a k_{cat} of 0.02930 min^{-1} and a K_M of 35.38 µM.

3.5.5 GilM-GilR Intraction Native PAGE Analysis

A 10% native PAGE gel was prepared by combining 3.4 mL acrylamide/bis-acrylamide (30%/0.8% w/v), 6.49 mL 0.375 M tris-HCl (pH 8.8), 100 µl 10% ammonium persulfate (w/v), and 10 µl TEMED in a 15 mL falcon tube and mixing by inversion. This solution was then quickly pipetted into the glass gel casting apparatus until 3/4 full. The separating gel was the allowed to gelate for 30 minutes. Next, the stacking gel solution was prepared by combining 4.275 ml 0.375 M tris-HCl (pH 8.8), 0.67 mL acrylamide/bis-acrylamide (30%/0.8% w/v), 0.05 mL 10% ammonium persulfate (w/v), and 5 µl TEMED in a 15 mL falcon tube and mixing by inversion. The stacking gel solution was pipetted on top of the separating gel, a comb was inserted and the solution was allowed to gelate for 30 minutes.

10 µl of each protein sample was then mixed with 10 ul sample buffer composed of 62.5 mM Tris-HCl (pH 6.8), 25% glycerol, and 1% bromophenol blue. The samples were loaded onto the native gel and electrophoresed for 1.5 hours at 150 volts. The gel was carefully removed from between the glass plates of the gel cast and stained in coomassie blue staining solution (0.3% coomassie brilliant blue R-250 w/v, 45% methanol v/v, 10% glacial acetic acid v/v, 45% distilled water v/v) for 45 minutes. Finally, the gel was rinsed with distilled water and destained in destaining solution (20% methanol v/v, 10% glacial acetic acid v/v, 70% distilled water v/v) overnight.

3.5.6 GilM-GilR Interaction FPLC Analysis

GilR and GilM were first purified using the method discussed in section 3.5.2 with the exception of using different lysis, wash, and elution buffers. The cell lysis buffer was composed of 40 mM Tris pH 8.0, 400 mM NaCl, 10% glycerol (w/w), and 2 mM β-mercaptoethanol. The wash buffer was composed of 40mM Tris pH 8.0, 400mM NaCl, 10% glycerol (w/w), 2mM β-mercaptoethanol, and 20mM imidazole. The elution buffer was composed of 40mM Tris pH 8.0, 400mM NaCl, 10% glycerol (w/w), 2mM β-mercaptoethanol, and 200mM imidazole. Each protein stock solution was diluted to a concentration of 1mg/ml using 40 mM Tris-HCl pH 8.0, 100 mM NaCl, and 2mM β-mercaptoethanol. 25µl of 1mg/ml GilM was mixed with 72µl of 40mM Tris-HCl pH 8.0, 100 mM NaCl, and 2mM β-mercaptoethanol. 3µl of 50mM SAM prepared in DMSO was then added to the solution followed by the addition of 25µl of GilR 1mg/ml stock. The enzyme mixture was passed through a size-exclusion S-200 column equilibrated in 40 mM Tris-HCl pH 8.0, 100mM NaCl, and 2mM β-mercaptoethanol using the Bio-Rad BioLogic DuoFlow system. Controls of GilM and GilR alone were then run for comparison of the chromatograms.

3.6 Reflections and Future Directions

If the experiments preformed in this chapter were to be repeated several changes could be made to improve the validity of the findings presented. One major change would be to use C-terminally histagged GilM during all experimentation. During initial crystallization screening it was found that a C-terminally histagged *GilM* expression construct (GilM-Pet22b(+)-C His6x shown in the appendix) produced greater yields of soluble GilM protein. Additionally, it was found that C-terminally histagged GilM was more stable then N-terminally histagged GilM. It should also be noted that upon removal of either the N or C terminal histag GilM rapidly precipitates. Another major improvement would be to change the method of GilM reaction monitoring with regard to kinetic analysis. Rather then monitoring the less stable GilM products it would be better to include a large quantity of GilR and monitor the formation of the final stable product defuco-gilvocarcin M.

The next major steps in this project include verification of transient protein-protein interactions between GilM and GilR and structural elucidation of GilR. Either an *E. coli* based two-hybrid assay or an alpha screen could be used to positively identify the presence of a transient protein-protein interaction. Another method of verification could be obtained by solving the X-ray crystal structures of GilM and a fusion homologue, such as RavRM or ChryRM. The structure of GilM could then be computationally docked to the previously solved structure of GilR using the structure of the fusion homologue as a guide.

Portions of this chapter have been reproduced with permission from the following publication: copyright ACS publications

Tibrewal, N.*****, **Downey, T. E.**[+], Van Lanen, S. G., Ul Sharif, E., O'Doherty, G. A., Rohr, J. "Roles of the Synergistic Reductive O-Methyltransferase GilM and of O-Methyltransferase GilMT in the Gilvocarcin Biosynthetic Pathway." *J. Am. Chem. Soc.*, **2012**, *134*, 12402–12407.

*Chemoenzymatic synthesis of substrate 5, GilM reaction analysis, and collection of kinetic data.

[+] Kinetic assay development, analysis of kinetic GilM kinetic data, construction of GilM-Pet28a vector for enzyme-enzyme interaction studies, native PAGE analysis, and FPLC method development and analysis.

4 Summary

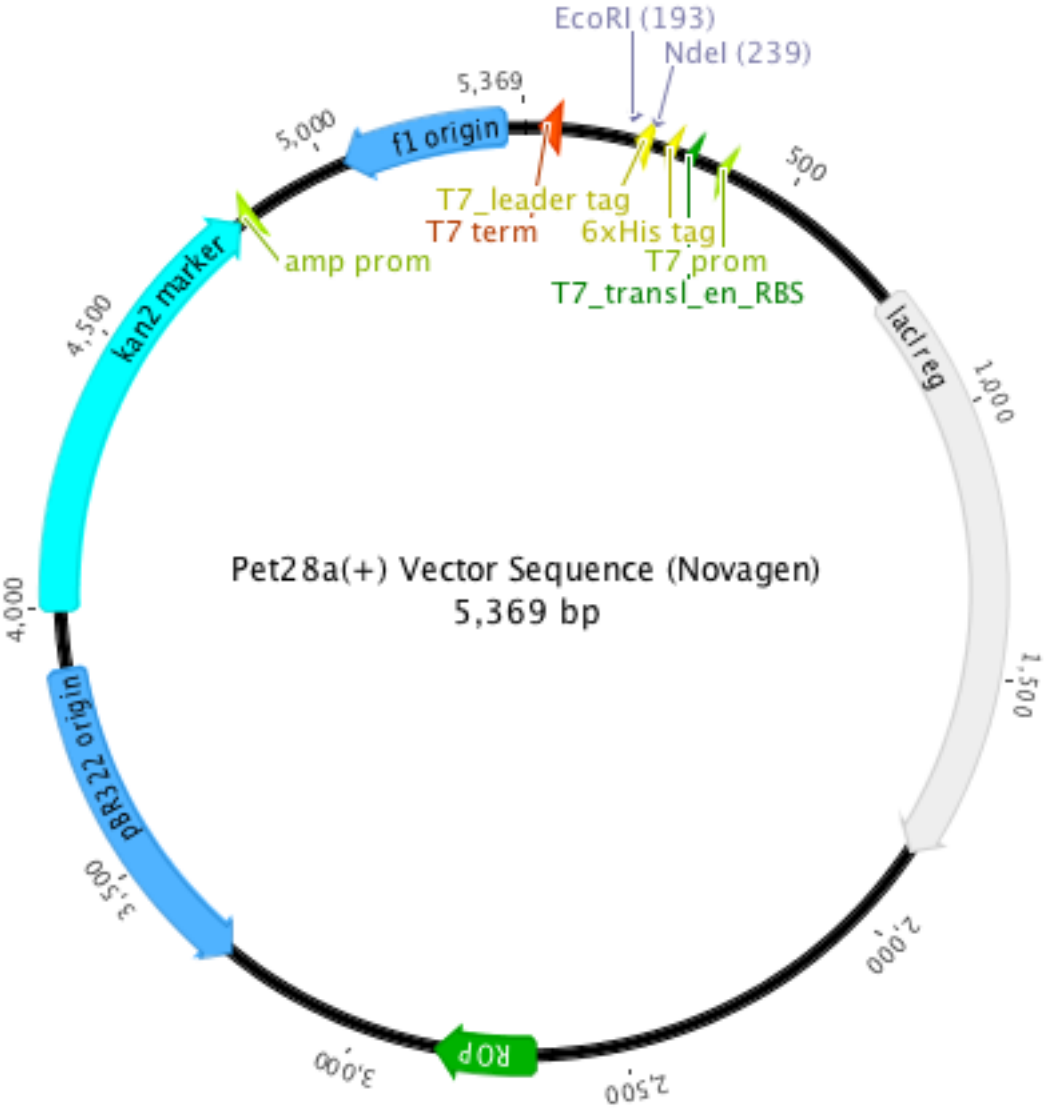
Post-PKS tailoring enzymes offer a unique opportunity to biosynthetically engineer new molecules for medicinal and industrial applications. If scientists are to fully exploit their potential the structures and functions of these enzymes must be studied in detail not only as lone entities, but also synergistically as co-dependent partners. For both bifunctional and co-dependent post-PKS enzymes there is a need for more functional and structural information. Through the work presented in this dissertation a better understand of the enzyme MtmOIV, of the MtmOIV-MtmW co-dependent enzyme pair from the mithramycin biosynthetic pathway, and the enzyme GilM, a bifunctional partner of the GilM-GilR co-dependent enzyme pair from the gilvocarcin biosynthetic pathway, has been gained.

Our new refined MtmOIV structure is the first substrate-bound crystal structure of MtmOIV. The MtmOIV structure with its bound substrate PMB revealed significant differences to the previous computer model, in which PMB was docked into the low resolution structure of MtmOIV. The new MtmOIV protein structures revealed that a tryptophan (W205), arginine (R204) residue and a phenylalanine (F89) residue play key roles in guarding the entry loop during the substrate binding process. These new structures also show that a histidine (H50) and an aspartate (D51) residue seem to be important to fine-tune the positioning of the FAD cofactor. Additionally, two glutamine residues (Q78 and Q91) and a proline residue (P84) were recognized as critical to holding the substrate's trisaccharide chain, essential for anticancer activity,[78,79] in place. These three trisaccharide stabilizing residues were not identified in the previous PMB–MtmOIV binding site model but are likely critical for our future attempts to further modify MTM's trisaccharide side chain through enzyme engineering. Furthermore, using a low-resolution complex structure of MtmOIV in complex with NADPH combined with site directed mutagenesis studies, four arginine residues (R169, R173, R174, R277) important for binding NADPH were identified and confirmed.[44] With this new structural information it may be possible to design a more promiscuous MtmOIV enzyme capable of creating new and improved mithramycin analogues.

GilM from the gilvocarcin biosynthetic gene cluster was identified via *in vitro* characterization as a SAM-dependent multifunctional enzyme responsible for both the reduction of a quinone intermediate to a hydroquinone and the associated stabilizing *O*-methylation and hemiacetal formation. This enzyme is one of only a limited number of bifunctional post-PKS tailoring enzymes identified to date. In addition to being identified as a bifunctional enzyme GilM, was identified as a co-dependent enzyme requiring the assistance of the oxidoreductase GilR to function. GilM is hypothesized to utilize GilR's bicovalently bound reduced flavin generated in the GilR reaction for quinone reduction thereby regenerating oxidized flavin for the next catalytic cycle. To further interrogate the GilM-GilR relationship, enzyme-enzyme interaction studies were conducted. These studies led to the conclusion that the GilM-GilR interaction is transient as no enzyme complex was detected in either case. The identification of GilM's catalytic functionality and co-dependent relationship with GilR elucidates the end of the gilvocarcin biosynthetic pathway and establishes the groundwork for future combinatorial biosynthetic work that may lead to the production of new gilvocarcin analogues with improved druggability.

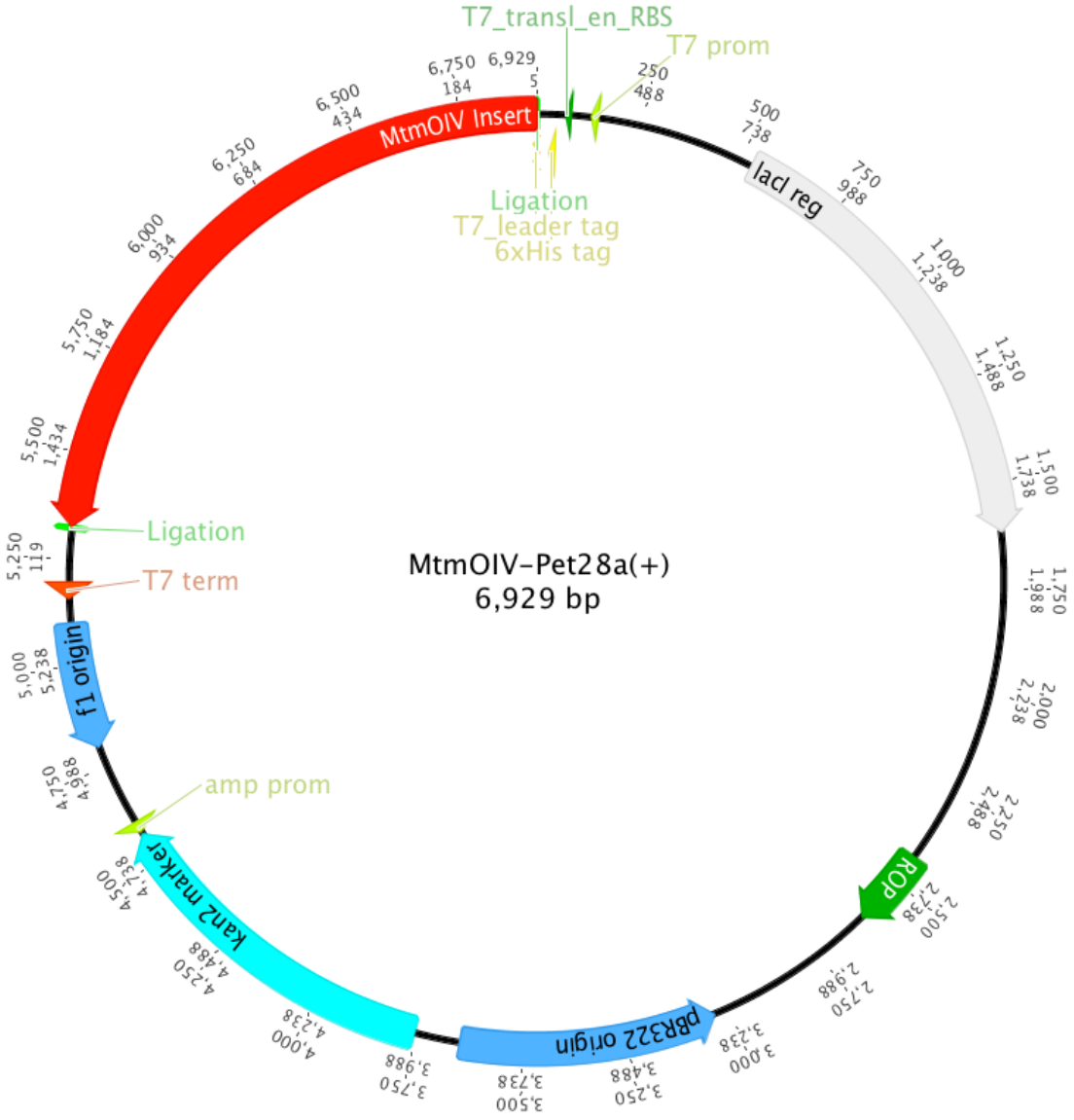
5 Appendix

5.1 Pet28a(+) Plasmid Map



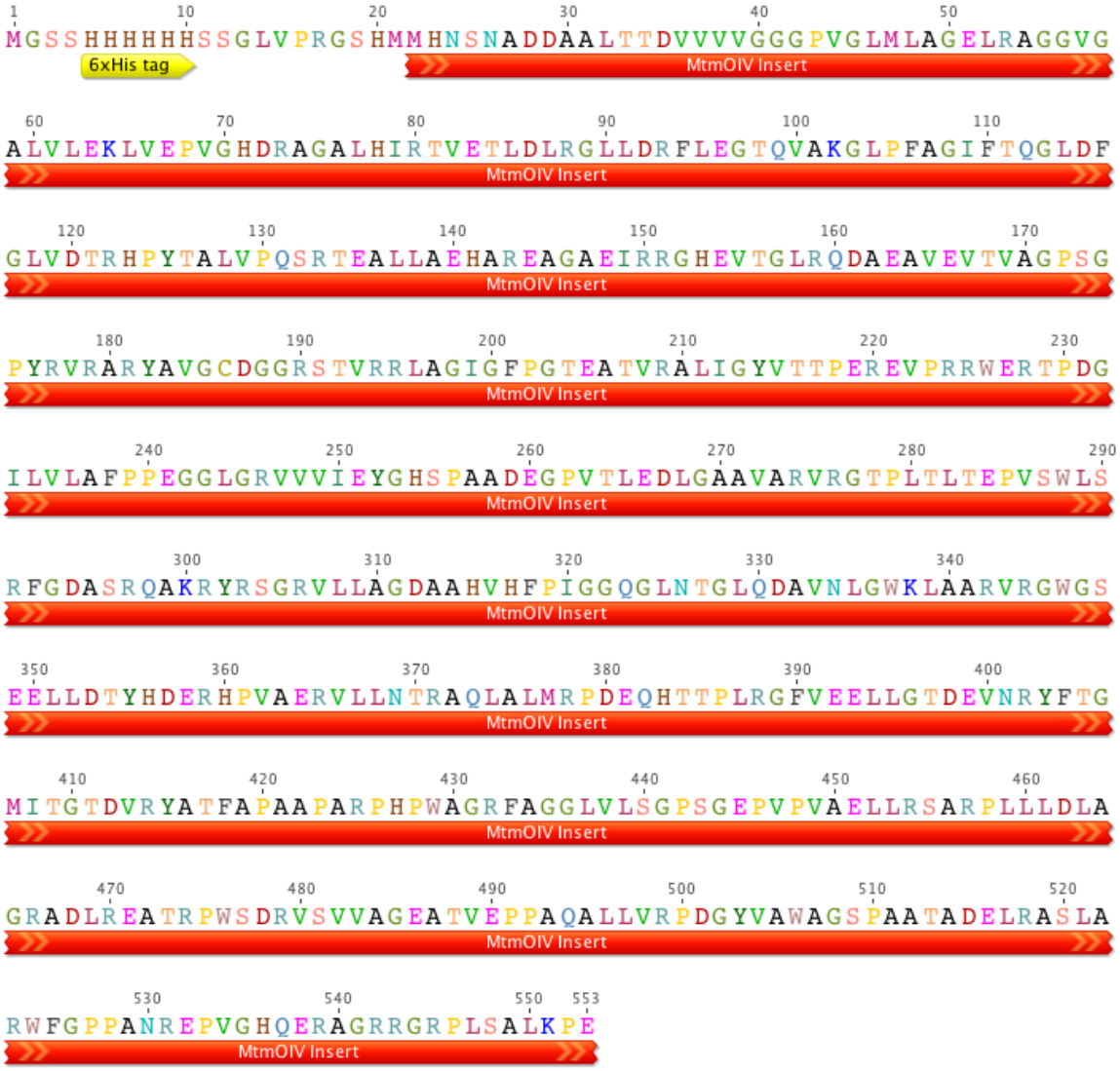
5.2 MtmOIV

5.2.1 MtmOIV-Pet28a(+) Plasmid Map



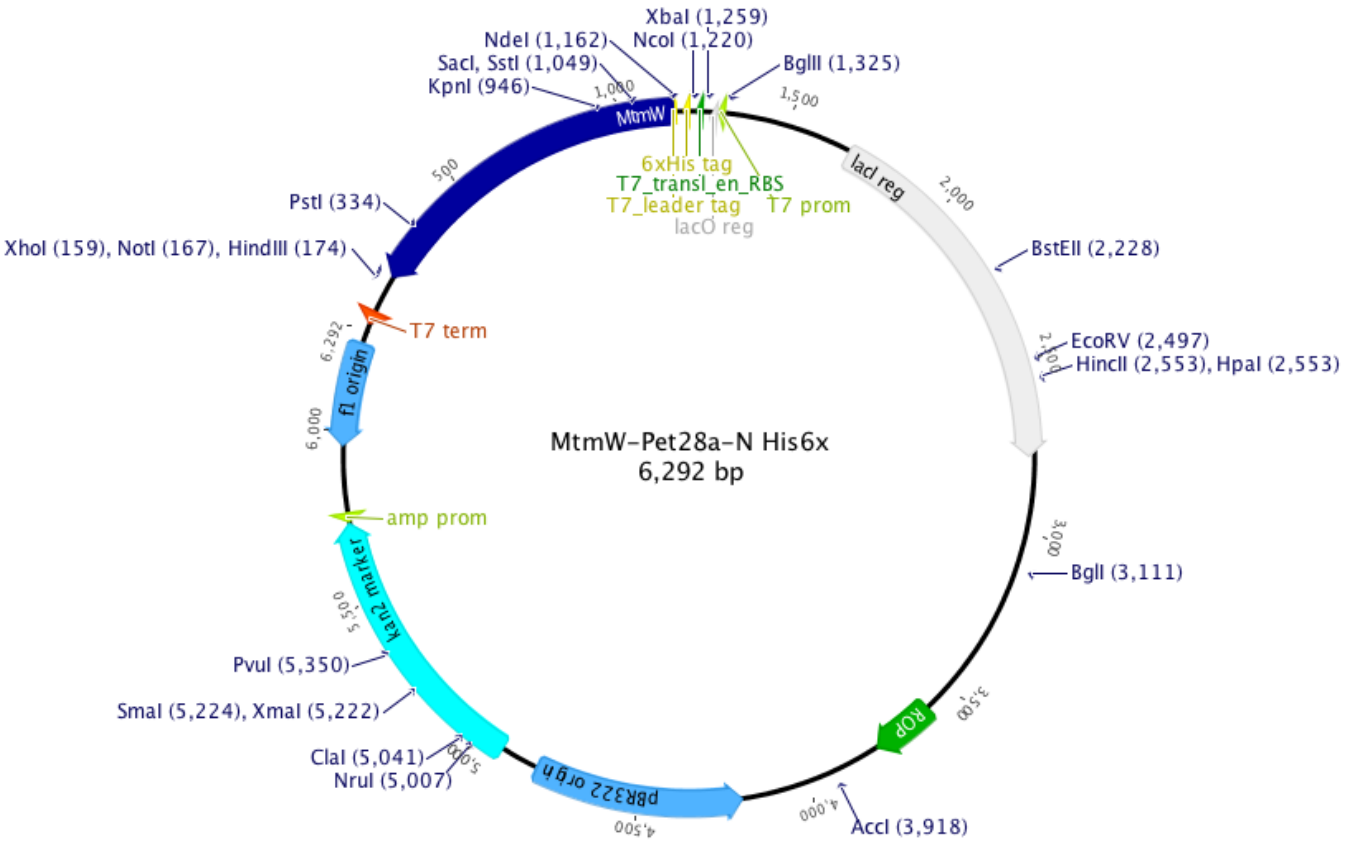
5.2.2 *MtmOIV* Gene Sequence

5.2.3 Pet28a(+) Expressed MtmOIV Enzyme Seqeunce

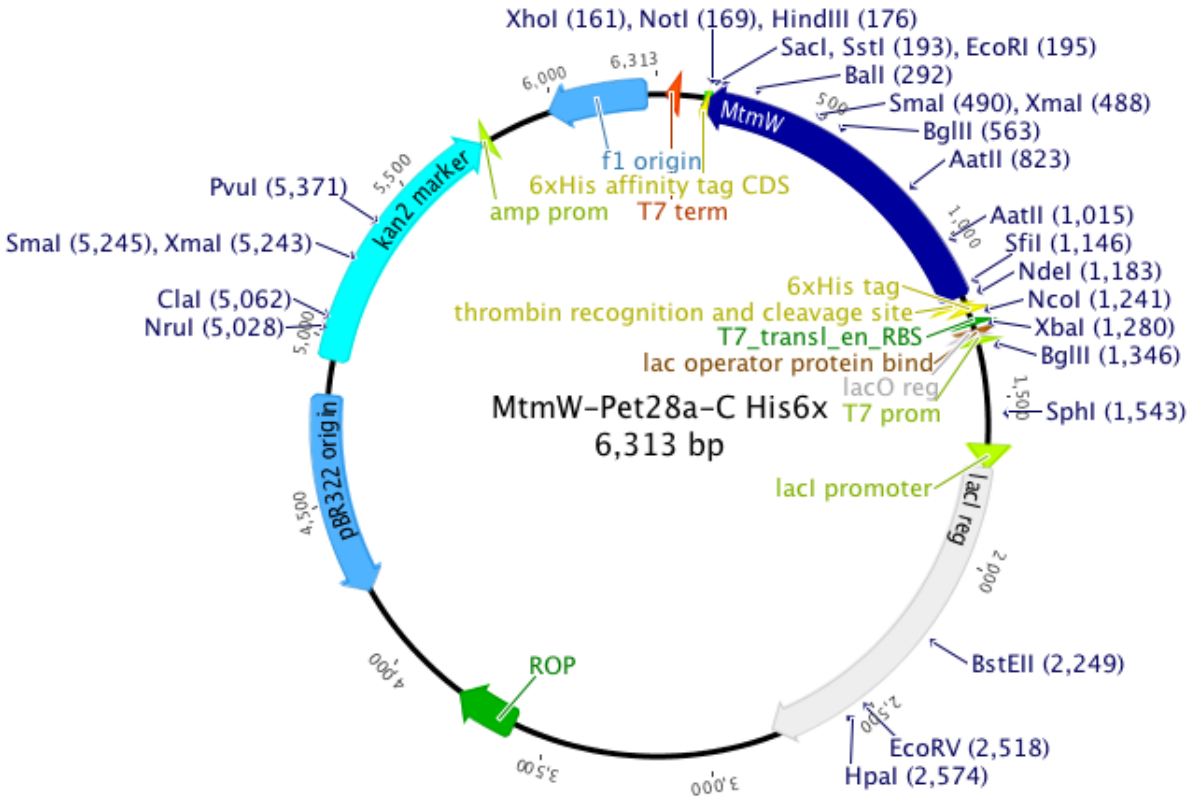


5.3 MtmW

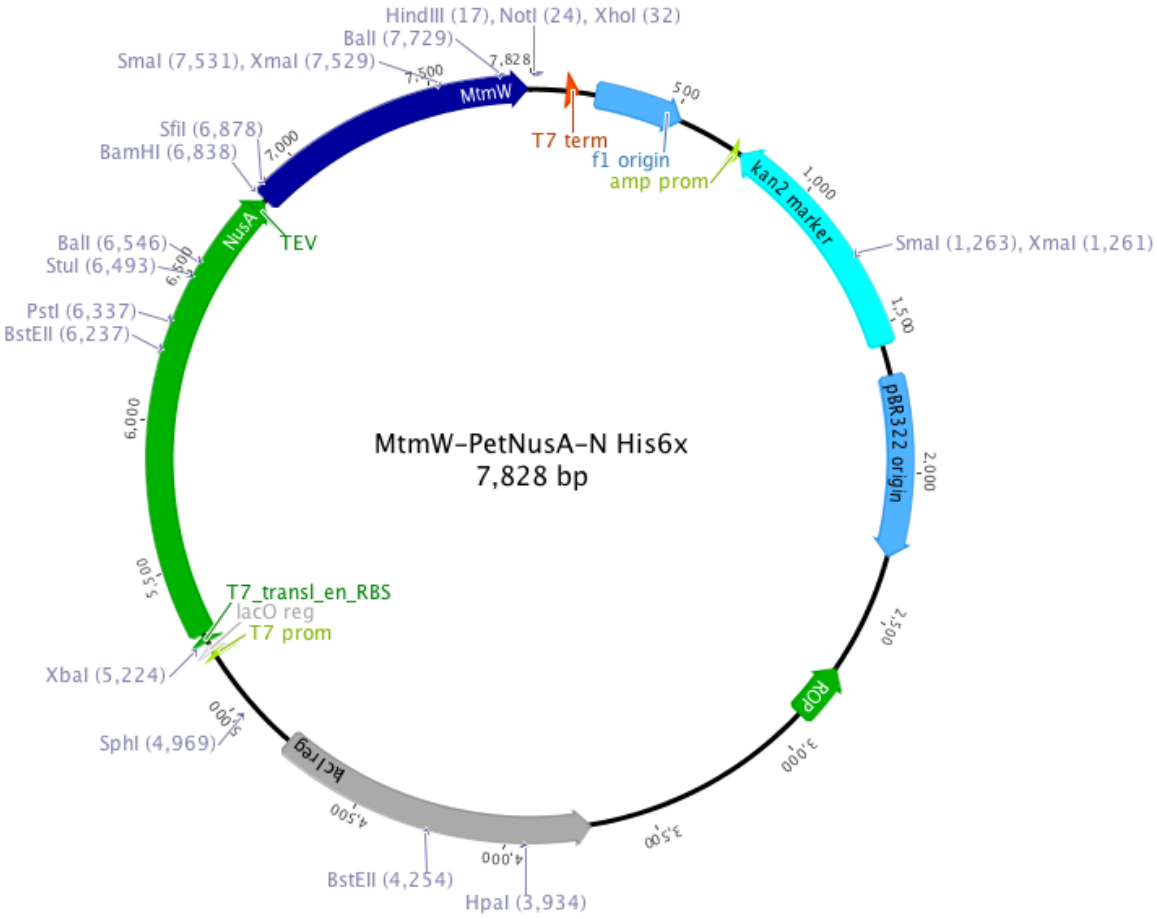
5.3.1 MtmW-Pet28a-N His6x



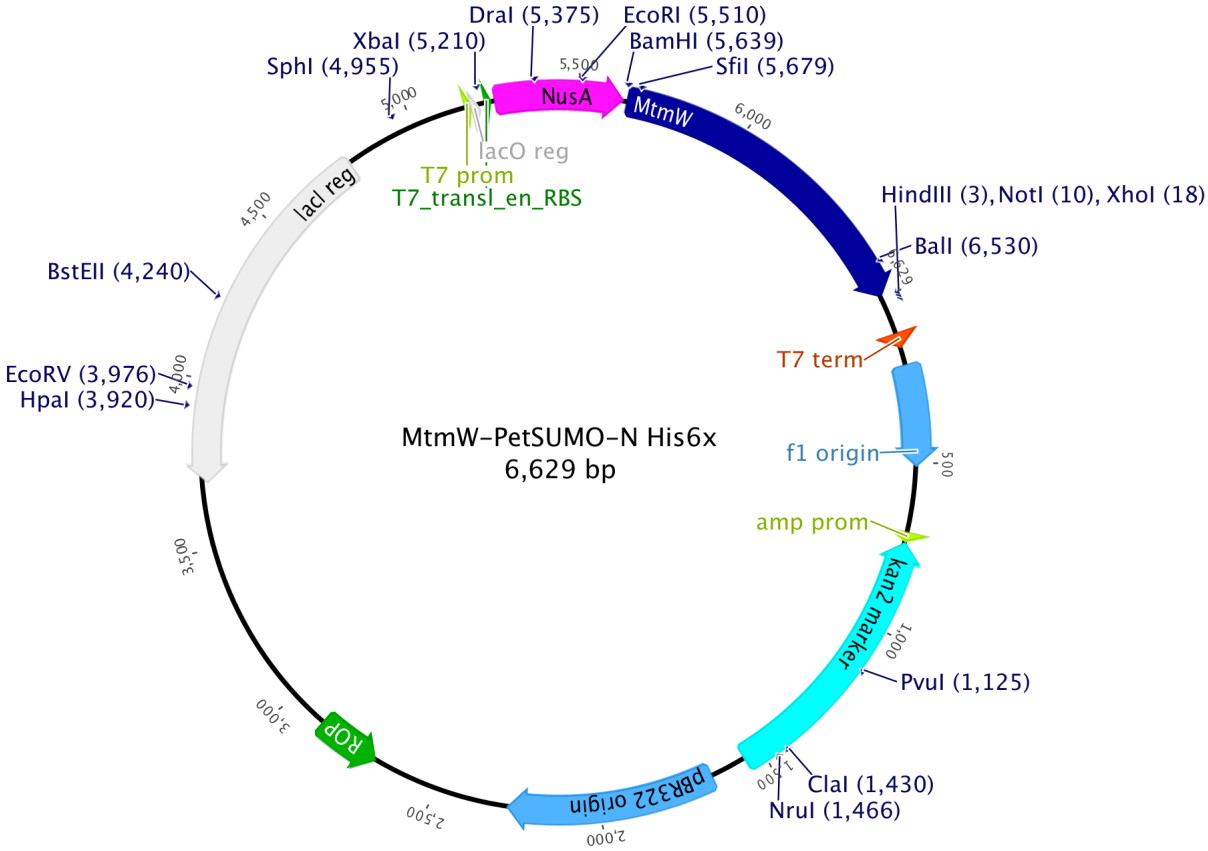
5.3.2 MtmW-Pet28a-C His6x



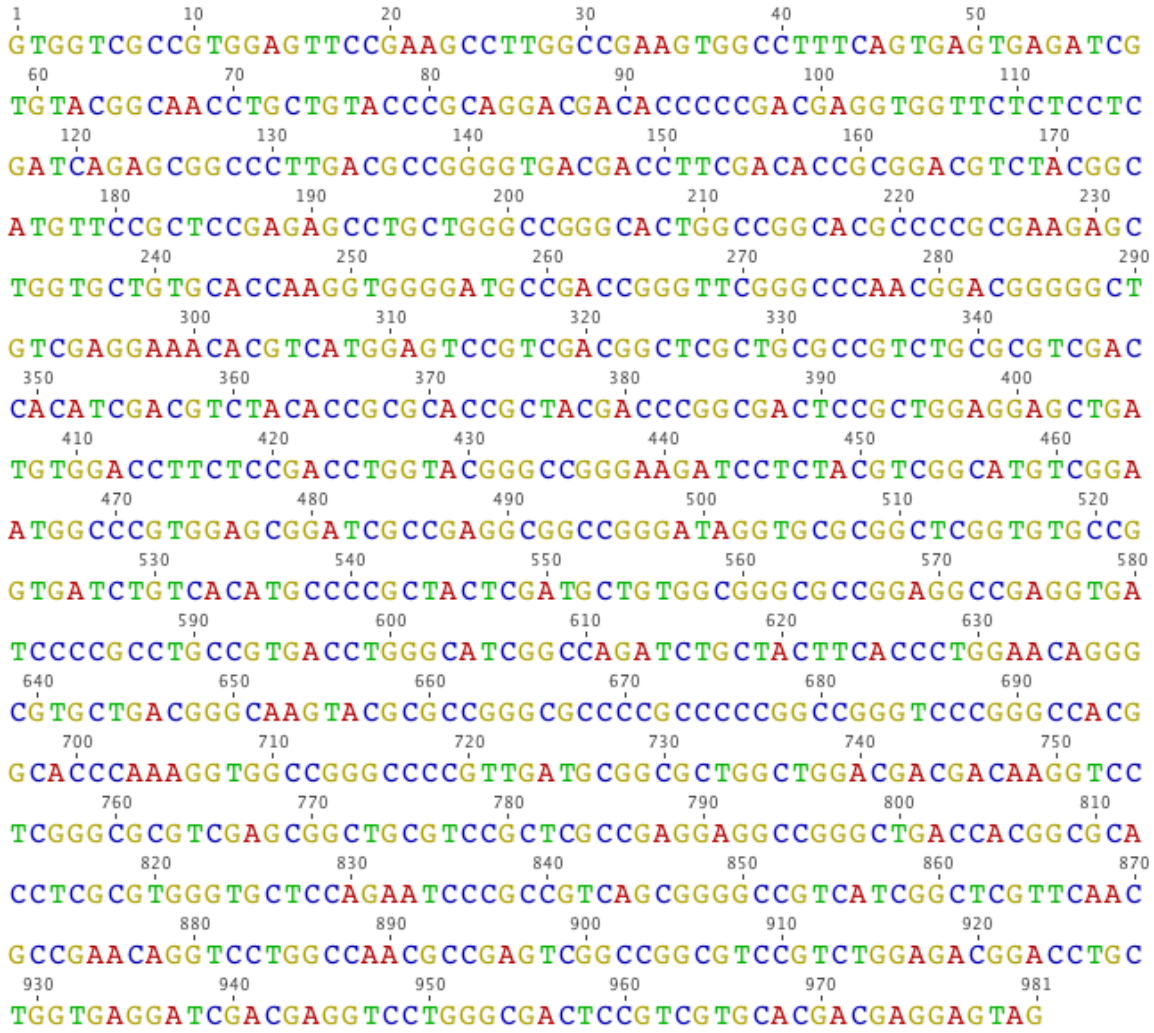
5.3.3 MtmW-PetNusA-N His6x



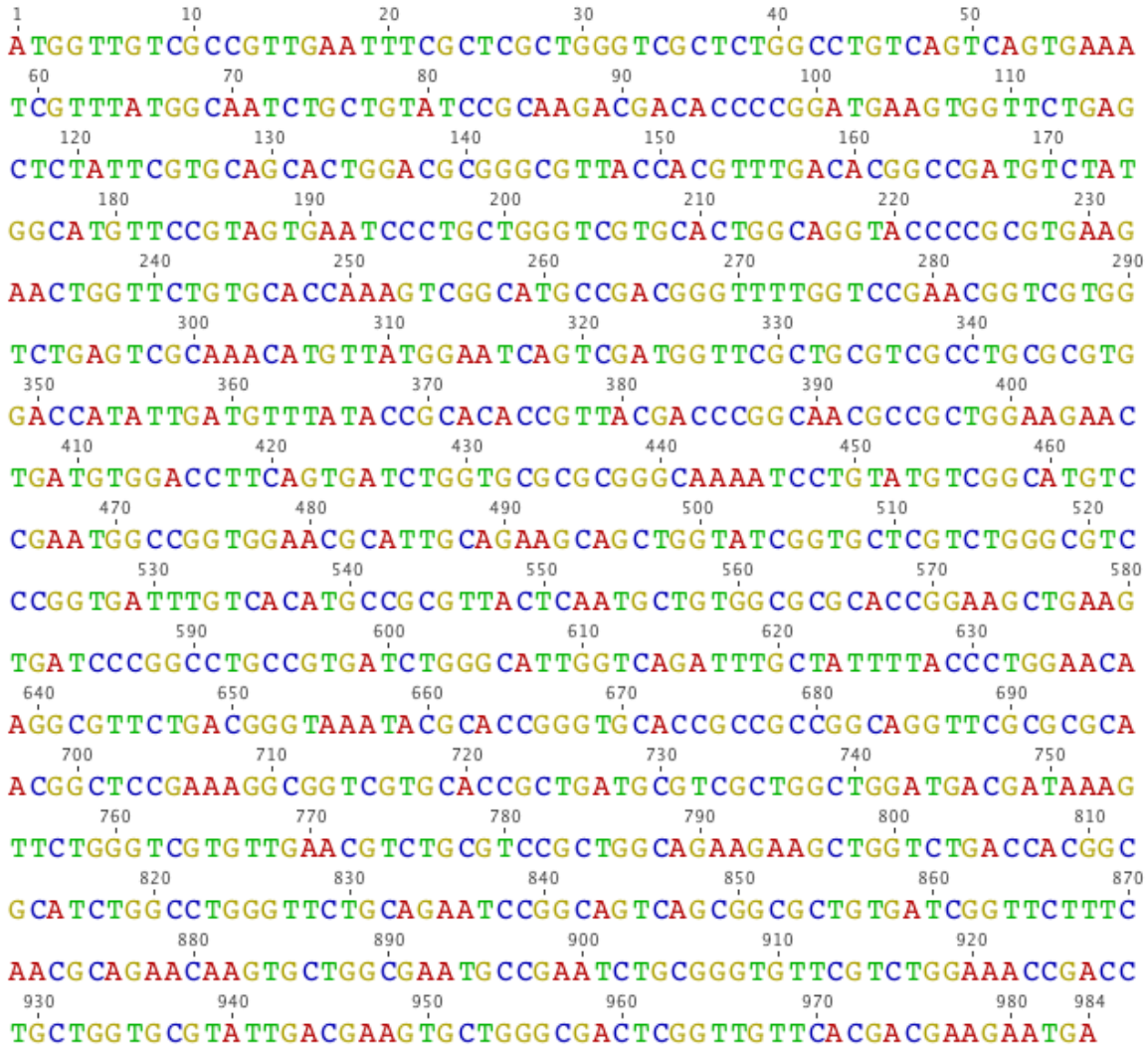
5.3.4 MtmW-PetSUMO-N His6x



5.3.5 *MtmW* Gene Sequence

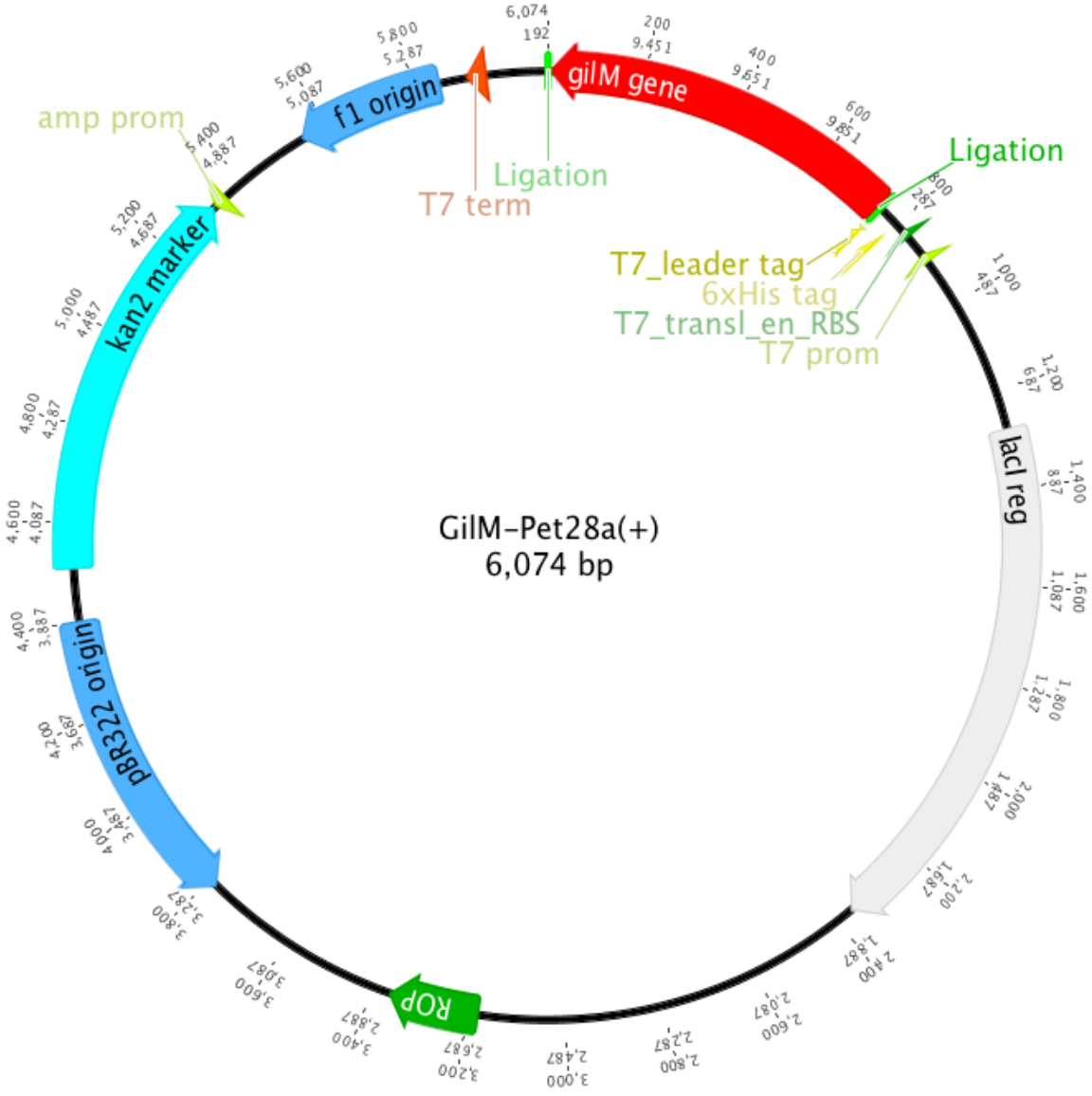


5.3.6 Codon optimized *MtmW* Gene Sequence



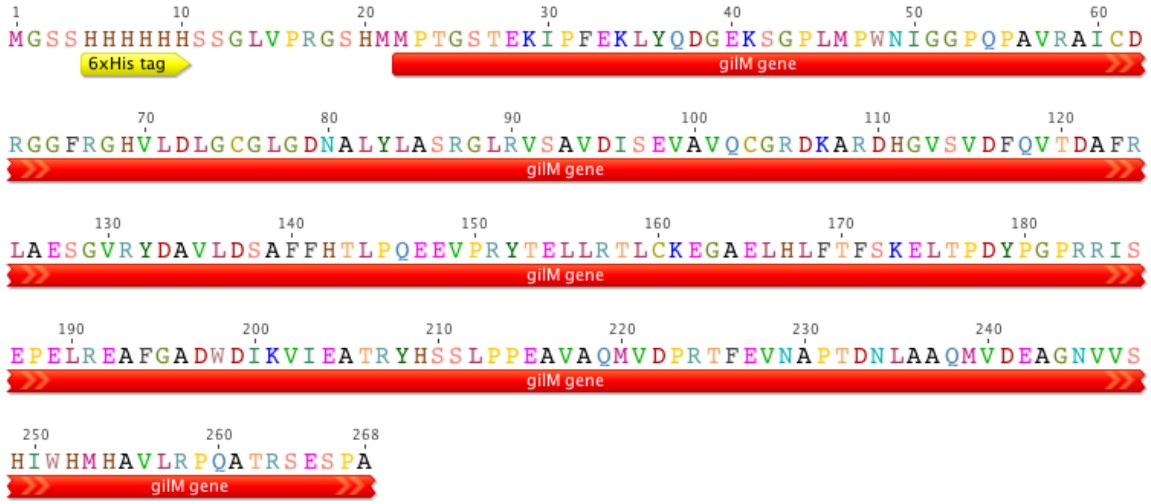
5.4 GilM

5.4.1 GilM-Pet28a(+) Plasmid Map

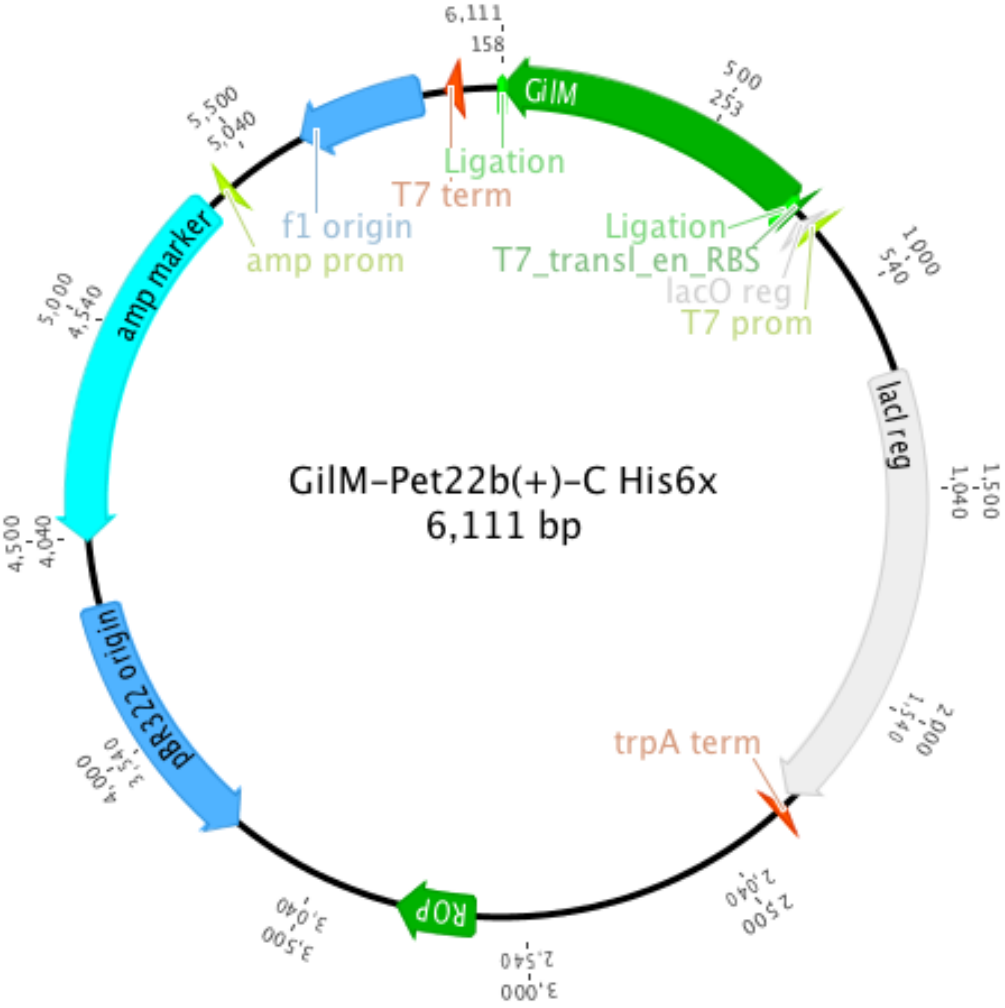


5.4.2 *GilM* Gene Sequence

5.4.3 Pet28a(+) Expressed GilM Enzyme Sequence

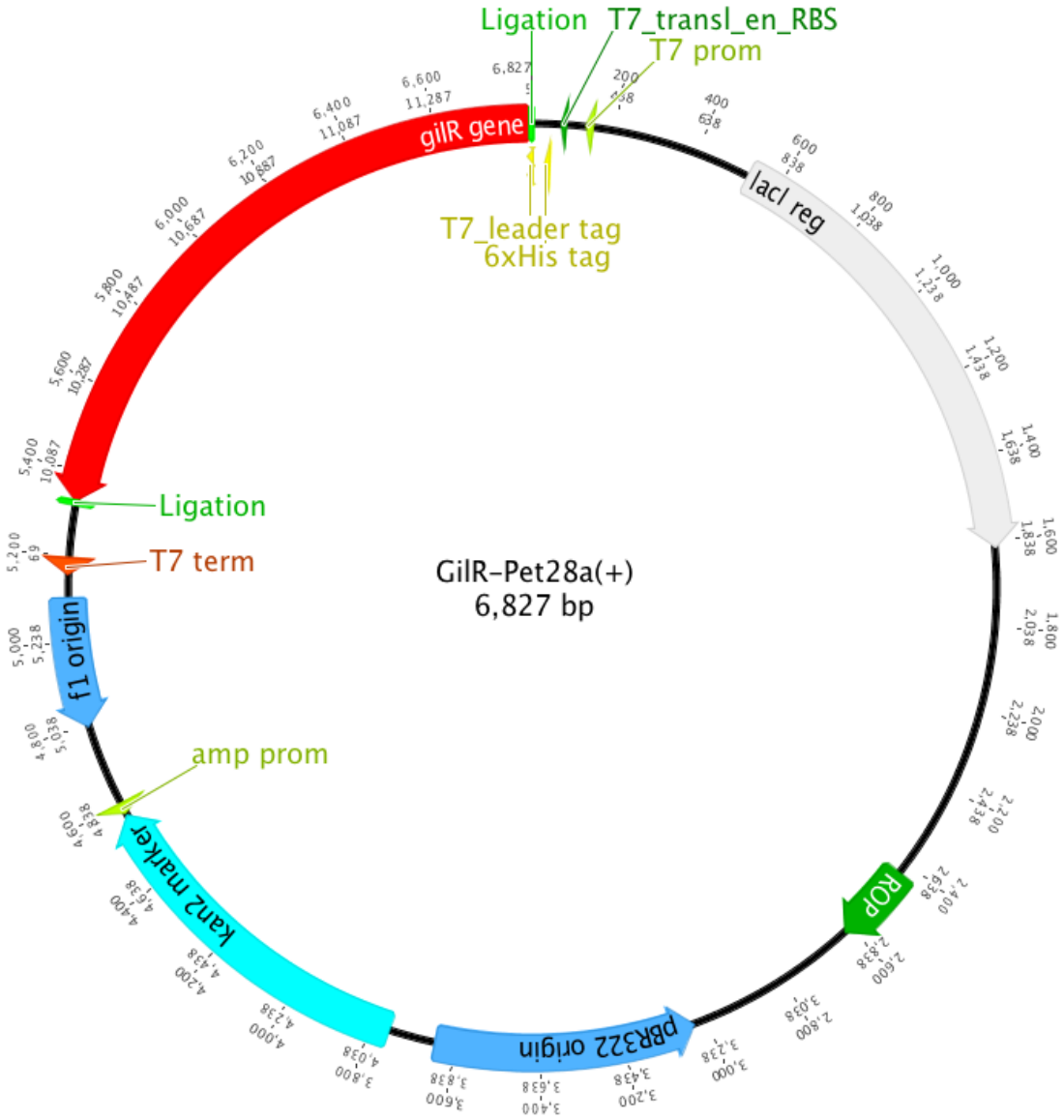


5.4.4 GilM-Pet22b(+) Plasmid Map



5.5 GilR

5.5.1 GilR-Pet28a(+) Plasmid Map



5.5.2 *GilR* Gene Sequence

5.5.3 Pet28a(+) Expressed GilR Enzyme Sequence

6. References

(1) Summers, R. G.; Wendt-Pienkowski, E.; Motamedi, H.; Hutchinson, C. R. *J. Bacteriol.* **1992**, *174*, 1810–1820.

(2) Yoshimoto, A.; Ogasawara, T.; Kitamura, I.; Oki, T.; Inui, T.; Takeuchi, T.; Umezawa, H. *J. Antibiot. (Tokyo)* **1979**, *32*, 472–481.

(3) Alexeev, I.; Sultana, A.; Mäntsälä, P.; Niemi, J.; Schneider, G. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 6170–6175.

(4) He, J.; Müller, M.; Hertweck, C. *J. Am. Chem. Soc.* **2004**, *126*, 16742–16743.

(5) Inouye, M.; Takada, Y.; Muto, N.; Beppu, T.; Horinouchi, S. *Mol. Gen. Genet. MGG* **1994**, *245*, 456–464.

(6) Anzai, Y.; Li, S.; Chaulagain, M. R.; Kinoshita, K.; Kato, F.; Montgomery, J.; Sherman, D. H. *Chem. Biol.* **2008**, *15*, 950–959.

(7) Li, S.; Tietz, D. R.; Rutaganira, F. U.; Kells, P. M.; Anzai, Y.; Kato, F.; Pochapsky, T. C.; Sherman, D. H.; Podust, L. M. *J. Biol. Chem.* **2012**, *287*, 37880–37890.

(8) Zhu, L.; Ostash, B.; Rix, U.; Nur-E-Alam, M.; Mayers, A.; Luzhetskyy, A.; Mendez, C.; Salas, J. A.; Bechthold, A.; Fedorenko, V.; Rohr, J. *J. Org. Chem.* **2005**, *70*, 631–638.

(9) Mayer, A.; Taguchi, T.; Linnenbrink, A.; Hofmann, C.; Luzhetskyy, A.; Bechthold, A. *Chembiochem Eur. J. Chem. Biol.* **2005**, *6*, 2312–2315.

(10) Dickens, M. L.; Priestley, N. D.; Strohl, W. R. *J. Bacteriol.* **1997**, *179*, 2641–2650.

(11) Walczak, R. J.; Hines, J. V.; Strohl, W. R.; Priestley, N. D. *Org. Lett.* **2001**, *3*, 2277–2279.

(12) Cheng, Q.; Xiang, L.; Izumikawa, M.; Meluzzi, D.; Moore, B. S. *Nat. Chem. Biol.* **2007**, *3*, 557–558.

(13) Xiang, L.; Kalaitzis, J. A.; Moore, B. S. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15609–15614.

(14) Teufel, R.; Miyanaga, A.; Michaudel, Q.; Stull, F.; Louie, G.; Noel, J. P.; Baran, P. S.; Palfey, B.; Moore, B. S. *Nature* **2013**, *503*, 552–556.

(15) Trefzer, A.; Fischer, C.; Stockert, S.; Westrich, L.; Künzel, E.; Girreser, U.; Rohr, J.; Bechthold, A. *Chem. Biol.* **2001**, *8*, 1239–1252.

(16) Luzhetskyy, A.; Fedoryshyn, M.; Dürr, C.; Taguchi, T.; Novikov, V.; Bechthold, A. *Chem. Biol.* **2005**, *12*, 725–729.

(17) Wang, G.; Pahari, P.; Kharel, M. K.; Chen, J.; Zhu, H.; Van Lanen, S. G.; Rohr, J. *Angew. Chem. Int. Ed.* **2012**, *51*, 10638–10642.

(18) Trefzer, A.; Blanco, G.; Remsing, L.; Künzel, E.; Rix, U.; Lipata, F.; Braña, A. F.; Méndez, C.; Rohr, J.; Bechthold, A.; Salas, J. A. *J. Am. Chem. Soc.* **2002**, *124*, 6056–6062.

(19) Li, Y.; Zhao, P.; Kang, Q.; Ma, J.; Bai, L.; Deng, Z. *Chem. Biol.* **2011**, *18*, 1571–1580.

(20) McDaniel, R.; Hutchinson, C. R.; Khosla, C. *J. Am. Chem. Soc.* **1995**, *117*, 6805–6810.

(21) Lombó, F.; Abdelfattah, M. S.; Braña, A. F.; Salas, J. A.; Rohr, J.; Méndez, C. *Chembiochem Eur. J. Chem. Biol.* **2009**, *10*, 296–303.

(22) Kharel, M. K.; Pahari, P.; Shaaban, K. A.; Wang, G.; Morris, C.; Rohr, J. *Org. Biomol. Chem.* **2012**, *10*, 4256–4265.

(23) Carlson, J. C.; Li, S.; Gunatilleke, S. S.; Anzai, Y.; Burr, D. A.; Podust, L. M.; Sherman, D. H. *Nat. Chem.* **2011**, *3*, 628–633.

(24) Carlson, J. C.; Li, S.; Burr, D. A.; Sherman, D. H. *J. Nat. Prod.* **2009**, *72*, 2076–2079.

(25) Kallio, P.; Liu, Z.; Mäntsälä, P.; Niemi, J.; Metsä-Ketelä, M. *Chem. Biol.* **2008**, *15*, 157–166.

(26) Zhan, J.; Watanabe, K.; Tang, Y. *ChemBioChem* **2008**, *9*, 1710–1715.

(27) Gibson, M.; Nur-e-alam, M.; Lipata, F.; Oliveira, M. A.; Rohr, J. *J. Am. Chem. Soc.* **2005**, *127*, 17594–17595.

(28) Rodríguez, D.; Quirós, L. M.; Braña, A. F.; Salas, J. A. *J. Bacteriol.* **2003**, *185*, 3962–3965.

(29) Remsing, L. L.; González, A. M.; Nur-e-Alam, M.; Fernández-Lozano, M. J.; Braña, A. F.; Rix, U.; Oliveira, M. A.; Méndez, C.; Salas, J. A.; Rohr, J. *J. Am. Chem. Soc.* **2003**, *125*, 5745–5753.

(30) Bosserman, M. A.; Downey, T.; Noinaj, N.; Buchanan, S. K.; Rohr, J. *ACS Chem. Biol.* **2013**, *8*, 2466–2477.

(31) Wang, C.; Gibson, M.; Rohr, J.; Oliveira, M. A. *Acta Crystallograph. Sect. F Struct. Biol. Cryst. Commun.* **2005**, *61*, 1023–1026.

(32) Prado, L.; Fernández, E.; Weissbach, U.; Blanco, G.; Quirós, L. M.; Braña, A. F.; Méndez, C.; Rohr, J.; Salas, J. A. *Chem. Biol.* **1999**, *6*, 19–30.

(33) Tibrewal, N.; Downey, T. E.; Van Lanen, S. G.; Ul Sharif, E.; O'Doherty, G. A.; Rohr, J. *J. Am. Chem. Soc.* **2012**, *134*, 12402–12405.

(34) Kharel, M. K.; Pahari, P.; Lian, H.; Rohr, J. *Chembiochem Eur. J. Chem. Biol.* **2009**, *10*, 1305–1308.

(35) Noinaj, N.; Bosserman, M. A.; Schickli, M. A.; Piszczek, G.; Kharel, M. K.; Pahari, P.; Buchanan, S. K.; Rohr, J. *J. Biol. Chem.* **2011**, *286*, 23533–23543.

(36) Leisch, H.; Morley, K.; Lau, P. C. K. *Chem. Rev.* **2011**, *111*, 4165–4222.

(37) Walsh, C. T.; Wencewicz, T. A. *Nat. Prod. Rep.* **2013**, *30*, 175–200.

(38) Dover, L. G.; Corsino, P. E.; Daniels, I. R.; Cocklin, S. L.; Tatituri, V.; Besra, G. S.; Fütterer, K. *J. Mol. Biol.* **2004**, *340*, 1095–1105.

(39) Koma, D.; Sakashita, Y.; Kubota, K.; Fujii, Y.; Hasumi, F.; Chung, S. Y.; Kubo, M. *Appl. Microbiol. Biotechnol.* **2004**, *66*, 92–99.

(40) Leisch, H.; Shi, R.; Grosse, S.; Morley, K.; Bergeron, H.; Cygler, M.; Iwaki, H.; Hasegawa, Y.; Lau, P. C. K. *Appl. Environ. Microbiol.* **2012**, *78*, 2200–2212.

(41) Philmus, B.; Abdelwahed, S.; Williams, H. J.; Fenwick, M. K.; Ealick, S. E.; Begley, T. P. *J. Am. Chem. Soc.* **2012**, *134*, 5326–5330.

(42) Beam, M. P.; Bosserman, M. A.; Noinaj, N.; Wehenkel, M.; Rohr, J. *Biochemistry (Mosc.)* **2009**, *48*, 4476–4487.

(43) Rohr, J.; Méndez, C.; Salas, J. A. *Bioorganic Chem.* **1999**, *27*, 41–54.

(44) Lombó, F.; Menéndez, N.; Salas, J. A.; Méndez, C. *Appl. Microbiol. Biotechnol.* **2006**, *73*, 1–14.

(45) Rawlings, B. J. *Nat. Prod. Rep.* **1999**, *16*, 425–484.

(46) Wang, G.; Kharel, M. K.; Pahari, P.; Rohr, J. *Chembiochem Eur. J. Chem. Biol.* **2011**, *12*, 2568–2571.

(47) O'Connor, S. *Chem. Biol.* **2004**, *11*, 8–10.

(48) Nehls, M. C.; Brenner, D. A.; Gruss, H. J.; Dierbach, H.; Mertelsmann, R.; Herrmann, F. *J. Clin. Invest.* **1993**, *92*, 2916–2921.

(49) Seznec, J.; Silkenstedt, B.; Naumann, U. *J. Neurooncol.* **2011**, *101*, 365–377.

(50) Sachrajda, I.; Ratajewski, M. *Mol. Genet. Genomics MGG* **2011**, *285*, 57–65.

(51) Kennedy, B. J. *Cancer* **1970**, *26*, 755–766.

(52) Kennedy, B. J.; Torkelson, J. L. *Med. Pediatr. Oncol.* **1995**, *24*, 327–328.

(53) Elias, E. G.; EVANS, J. T. *J. Bone Jt. Surg.* **1972**, *54*, 1730–1736.

(54) 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.; Catapano, C. V. *Nucleic Acids Res.* **2006**, *34*, 1721–1734.

(55) Lee, T.-J.; Jung, E. M.; Lee, J. T.; Kim, S.; Park, J.-W.; Choi, K. S.; Kwon, T. K. *Mol. Cancer Ther.* **2006**, *5*, 2737–2746.

(56) Hadjipavlou, A. G.; Gaitanis, L. N.; Katonis, P. G.; Lander, P. *Eur. Spine J. Off. Publ. Eur. Spine Soc. Eur. Spinal Deform. Soc. Eur. Sect. Cerv. Spine Res. Soc.* **2001**, *10*, 370–384.

(57) Kiang, D. T.; Loken, M. K.; Kennedy, B. J. *J. Clin. Endocrinol. Metab.* **1979**, *48*, 341–344.

(58) Lumachi, F.; Brunello, A.; Roma, A.; Basso, U. *Anticancer Res.* **2009**, *29*, 1551–1555.

(59) Perlia, C. P.; Gubisch, N. J.; Wolter, J.; Edelberg, D.; Dederick, M. M.; Taylor, S. G., 3rd. *Cancer* **1970**, *25*, 389–394.

(60) Zojer, N.; Keck, A. V.; Pecherstorfer, M. *Drug Saf. Int. J. Med. Toxicol. Drug Exp.* **1999**, *21*, 389–406.

(61) Ferrante, R. J.; Ryu, H.; Kubilus, J. K.; D'Mello, S.; Sugars, K. L.; Lee, J.; Lu, P.; Smith, K.; Browne, S.; Beal, M. F.; Kristal, B. S.; Stavrovskaya, I. G.; Hewett, S.;

Rubinsztein, D. C.; Langley, B.; Ratan, R. R. *J. Neurosci. Off. J. Soc. Neurosci.* **2004**, *24*, 10335–10342.

(62) Qiu, Z.; Norflus, F.; Singh, B.; Swindell, M. K.; Buzescu, R.; Bejarano, M.; Chopra, R.; Zucker, B.; Benn, C. L.; DiRocco, D. P.; Cha, J.-H. J.; Ferrante, R. J.; Hersch, S. M. *J. Biol. Chem.* **2006**, *281*, 16672–16680.

(63) Sleiman, S. F.; Langley, B. C.; Basso, M.; Berlin, J.; Xia, L.; Payappilly, J. B.; Kharel, M. K.; Guo, H.; Marsh, J. L.; Thompson, L. M.; Mahishi, L.; Ahuja, P.; MacLellan, W. R.; Geschwind, D. H.; Coppola, G.; Rohr, J.; Ratan, R. R. *J. Neurosci. Off. J. Soc. Neurosci.* **2011**, *31*, 6858–6870.

(64) 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.; Helman, L. J. *J. Natl. Cancer Inst.* **2011**, *103*, 962–978.

(65) Zhang, M.; Mathur, A.; Zhang, Y.; Xi, S.; Atay, S.; Hong, J. A.; Datrice, N.; Upham, T.; Kemp, C. D.; Ripley, R. T.; Wiegand, G.; Avital, I.; Fetsch, P.; Mani, H.; Zlott, D.; Robey, R.; Bates, S. E.; Li, X.; Rao, M.; Schrump, D. S. *Cancer Res.* **2012**, *72*, 4178–4192.

(66) Torres Pazmiño, D. E.; Baas, B.-J.; Janssen, D. B.; Fraaije, M. W. *Biochemistry (Mosc.)* **2008**, *47*, 4082–4093.

(67) Beecher, J.; Grogan, G.; Roberts, S.; Willetts, A. *Biotechnol. Lett.* **1996**, *18*, 571–576.

(68) McGhie, E. J.; Isupov, M. N.; Schröder, E.; Littlechild, J. A. *Acta Crystallogr. D Biol. Crystallogr.* **1998**, *54*, 1035–1038.

(69) Pasta, P.; Carrea, G.; Gaggero, N.; Grogan, G.; Willetts, A. *Biotechnol. Lett.* **1996**, *18*, 1123–1128.

(70) Mirza, I. A.; Yachnin, B. J.; Wang, S.; Grosse, S.; Bergeron, H.; Imura, A.; Iwaki, H.; Hasegawa, Y.; Lau, P. C. K.; Berghuis, A. M. *J. Am. Chem. Soc.* **2009**, *131*, 8848–8854.

(71) Yachnin, B. J.; Sprules, T.; McEvoy, M. B.; Lau, P. C. K.; Berghuis, A. M. *J. Am. Chem. Soc.* **2012**, *134*, 7788–7795.

(72) Bučko, M.; Schenkmayerová, A.; Gemeiner, P.; Vikartovská, A.; Mihovilovič, M. D.; Lacík, I. *Enzyme Microb. Technol.* **2011**, *49*, 284–288.

(73) Tibrewal, N.; Pahari, P.; Wang, G.; Kharel, M. K.; Morris, C.; Downey, T.; Hou, Y.; Bugni, T. S.; Rohr, J. *J. Am. Chem. Soc.* **2012**, *134*, 18181–18184.

(74) Leadbeater, C.; McIver, L.; Campopiano, D. J.; Webster, S. P.; Baxter, R. L.; Kelly, S. M.; Price, N. C.; Lysek, D. A.; Noble, M. A.; Chapman, S. K.; Munro, A. W. *Biochem. J.* **2000**, *352 Pt 2*, 257–266.

(75) Koskiniemi, H.; Metsä-Ketelä, M.; Dobritzsch, D.; Kallio, P.; Korhonen, H.; Mäntsälä, P.; Schneider, G.; Niemi, J. *J. Mol. Biol.* **2007**, *372*, 633–648.

(76) Sastry, M.; Patel, D. J. *Biochemistry (Mosc.)* **1993**, *32*, 6588–6604.

(77) Sastry, M.; Fiala, R.; Patel, D. J. *J. Mol. Biol.* **1995**, *251*, 674–689.

(78) Baig, I.; Perez, M.; Braña, A. F.; Gomathinayagam, R.; Damodaran, C.; Salas, J. A.; Méndez, C.; Rohr, J. *J. Nat. Prod.* **2008**, *71*, 199–207.

(79) Pérez, M.; Baig, I.; Braña, A. F.; Salas, J. A.; Rohr, J.; Méndez, C. *Chembiochem Eur. J. Chem. Biol.* **2008**, *9*, 2295–2304.

(80) Núñez, L. E.; Nybo, S. E.; González-Sabín, J.; Pérez, M.; Menéndez, N.; Braña, A. F.; Shaaban, K. A.; He, M.; Morís, F.; Salas, J. A.; Rohr, J.; Méndez, C. *J. Med. Chem.* **2012**, *55*, 5813–5825.

(81) Otwinowski, Z.; Minor, W. In *Methods in Enzymology*; Charles W. Carter, J., Ed.; Macromolecular Crystallography Part A; Academic Press, 1997; Vol. Volume 276, pp. 307–326.

(82) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(83) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242.

(84) Emsley, P.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132.

(85) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 1948–1954.

(86) Goodsell, D. S.; Morris, G. M.; Olson, A. J. *J. Mol. Recognit. JMR* **1996**, *9*, 1–5.

(87) Goodsell, D. S. *Cold Spring Harb. Protoc.* **2009**, *2009*, pdb.prot5200.

(88) Norgan, A. P.; Coffman, P. K.; Kocher, J.-P. A.; Katzmann, D. J.; Sosa, C. P. *J. Cheminformatics* **2011**, *3*, 12.

(89) Baba, N.; Akaho, E. *Bioinformation* **2011**, *6*, 387–388.

(90) Cosconati, S.; Forli, S.; Perryman, A. L.; Harris, R.; Goodsell, D. S.; Olson, A. J. *Expert Opin. Drug Discov.* **2010**, *5*, 597–607.

(91) Vistoli, G.; Pedretti, A.; Mazzolari, A.; Testa, B. *J. Comput. Aided Mol. Des.* **2010**, *24*, 771–787.

(92) Kharel, M. K.; Pahari, P.; Shepherd, M. D.; Tibrewal, N.; Nybo, S. E.; Shaaban, K. A.; Rohr, J. *Nat. Prod. Rep.* **2012**, *29*, 264–325.

(93) STRELITZ, F.; FLON, H.; ASHESHOV, I. N. *J. Bacteriol.* **1955**, *69*, 280–283.

(94) Hirayama, N.; Takahashi, K.; Shirahata, K.; Ohashi, Y.; Sasada, Y. *Bull. Chem. Soc. Jpn.* **1981**, *54*, 1338–1342.

(95) Kharel, M. K.; Rohr, J. *Curr. Opin. Chem. Biol.* **2012**, *16*, 150–161.

(96) Balitz, D. M.; O'Herron, F. A.; Bush, J.; Vyas, D. M.; Nettleton, D. E.; Grulich, R. E.; Bradner, W. T.; Doyle, T. W.; Arnold, E.; Clardy, J. *J. Antibiot. (Tokyo)* **1981**, *34*, 1544–1555.

(97) Morimoto, M.; Okubo, S.; Tomita, F.; Marumo, H. *J. Antibiot. (Tokyo)* **1981**, *34*, 701–707.

(98) Findlay, J. A.; Liu, J.-S.; Radics, L.; Rakhit, S. *Can. J. Chem.* **1981**, *59*, 3018–3020.

(99) Sehgal, S. N.; Czerkawski, H.; Kudelski, A.; Pandev, K.; Saucier, R.; Vézina, C. *J. Antibiot. (Tokyo)* **1983**, *36*, 355–361.

(100) Weiss, U.; Yoshihira, K.; Highet, R. J.; White, R. J.; Wei, T. T. *J. Antibiot. (Tokyo)* **1982**, *35*, 1194–1201.

(101) Li, Y.; Huang, X.; Ishida, K.; Maier, A.; Kelter, G.; Jiang, Y.; Peschel, G.; Menzel, K.-D.; Li, M.; Wen, M.; Xu, L.; Grabley, S.; Fiebig, H.-H.; Jiang, C.; Hertweck, C.; Sattler, I. *Org. Biomol. Chem.* **2008**, *6*, 3601–3605.

(102) Nakajima, S.; Kojiri, K.; Suda, H.; Okanishi, M. *J. Antibiot. (Tokyo)* **1991**, *44*, 1061–1064.

(103) Yamashita, N.; Shin-ya, K.; Furihata, K.; Hayakawa, Y.; Seto, H. *J. Antibiot. (Tokyo)* **1998**, *51*, 1105–1108.

(104) Matsumoto, A.; Fujiwara, Y.; Elespuru, R. K.; Hanawalt, P. C. *Photochem. Photobiol.* **1994**, *60*, 225–230.

(105) McGee, L. R.; Misra, R. *J. Am. Chem. Soc.* **1990**, *112*, 2386–2389.

(106) Elespuru, R. K.; Gonda, S. K. *Science* **1984**, *223*, 69–71.

(107) Tse-Dinh, Y. C.; McGee, L. R. *Biochem. Biophys. Res. Commun.* **1987**, *143*, 808–812.

(108) Matsumoto, A.; Hanawalt, P. C. *Cancer Res.* **2000**, *60*, 3921–3926.

(109) Lytle, C. D.; Wagner, S. J.; Prodouz, K. N. *Photochem. Photobiol.* **1993**, *58*, 818–821.

(110) Hosoya, T.; Takashiro, E.; Matsumoto, T.; Suzuki, K. *J. Am. Chem. Soc.* **1994**, *116*, 1004–1015.

(111) Futagami, S.; Ohashi, Y.; Imura, K.; Hosoya, T.; Ohmori, K.; Matsumoto, T.; Suzuki, K. *Tetrahedron Lett.* **2000**, *41*, 1063–1067.

(112) Fischer, C.; Lipata, F.; Rohr, J. *J. Am. Chem. Soc.* **2003**, *125*, 7818–7819.

(113) Shepherd, M. D.; Kharel, M. K.; Zhu, L. L.; van Lanen, S. G.; Rohr, J. *Org. Biomol. Chem.* **2010**, *8*, 3851–3856.

(114) Pahari, P.; Kharel, M. K.; Shepherd, M. D.; van Lanen, S. G.; Rohr, J. *Angew. Chem. Int. Ed Engl.* **2012**, *51*, 1216–1220.

(115) Kharel, M. K.; Zhu, L.; Liu, T.; Rohr, J. *J. Am. Chem. Soc.* **2007**, *129*, 3780–3781.

(116) Liu, T.; Fischer, C.; Beninga, C.; Rohr, J. *J. Am. Chem. Soc.* **2004**, *126*, 12262–12263.

(117) Liu, T.; Kharel, M. K.; Fischer, C.; McCormick, A.; Rohr, J. *Chembiochem Eur. J. Chem. Biol.* **2006**, *7*, 1070–1077.

(118) Kharel, M. K.; Nybo, S. E.; Shepherd, M. D.; Rohr, J. *Chembiochem Eur. J. Chem. Biol.* **2010**, *11*, 523–532.

(119) Altschul, S. F.; Lipman, D. J. *Proc. Natl. Acad. Sci.* **1990**, *87*, 5509–5513.

(120) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.

(121) Frutos, Ó. de; Atienza, C.; Echavarren, A. M. *Eur. J. Org. Chem.* **2001**, *2001*, 163–171.

7. Vita

Theresa Downey

Experience:

PhD Candidate Graduate Research Assistant **June 2010 to Present**

University of Kentucky College of Pharmacy - Lexington, KY

Research Assistant **August 2007 to May 2010**

Saint Vincent College - Latrobe, PA

Research Internship **May 2009 to May August 2009**

University of Pittsburgh NIH-NSF Bioengineering and Bioinformatics Summer Institute

Internship/Volunteer **September 2005 to ay 2006**

Veterans Affairs Hospital - Pittsburgh, PA

Education:

BS in Biochemistry (GPA: 3.9/4.0)
2006 to 2010
Saint Vincent College - Latrobe, PA

Awards:

American Foundation for Pharmaceutical Education Pre-doctoral Fellowship 2013
To encourage outstanding pre-doctoral students to continue their studies and earn a PhD in the pharmaceutical sciences at a U.S. school or college of pharmacy.

Graduated Suma Cum Laude - Saint Vincent College 2010
Achievement of a final GPA above 3.85/4.00.

Society for Analytical Chemists of Pittsburgh College Chemistry Award 2009
Award for outstanding scholastic achievement at the undergraduate level.

Edith Q. Monack Scholarship 2009
Award for scholastic achievement of a female student in the sciences.

Saint Vincent Academic Scholarship 2006
Award based on scholastic achievement.

Veterans Affairs Youth Volunteer Scholarship 2006
Award given for significant volunteer contribution.

Publications:

Molecular insight into substrate recognition and catalysis in the Baeyer-Villiger monooxygenase MtmOIV, the key frame modifying enzyme in the biosynthesis of anticancer agent mithramycin.

Downey T, Bosserman MA, Noinaj N, Buchanan SK, Rohr J.

ACS Chem Biol. 2013 Sep 13. [Epub ahead of print]

Baeyer-Villiger monooxygenases (BVMOs) have been shown to play key roles for the biosynthesis of important natural products. MtmOIV, a homodimeric FAD- and NADPH-dependent BVMO, catalyzes the key frame-modifying steps of the mithramycin biosynthetic pathway, including an oxidative C-C bond cleavage, by converting its natural substrate premithramycin B into mithramycin DK, the immediate precursor of mithramycin. The drastically improved protein structure of MtmOIV along with the high-resolution structure of MtmOIV in complex with its natural substrate premithramycin B are reported here, revealing previously undetected key residues that are important for substrate recognition and catalysis. Kinetic analyses of selected mutants allowed us to probe the substrate binding pocket of MtmOIV, and also to discover the putative NADPH binding site. This is the first substrate-bound structure of MtmOIV providing new insights into substrate recognition and catalysis, which paves the way for the future design of a tailored enzyme for the chemo- enzymatic preparation of novel mithramycin analogues.

Roles of the synergistic reductive O-methyltransferase GilM and of O-methyltransferase GilMT in the gilvocarcin biosynthetic pathway.

Tibrewal N, Downey TE, Van Lanen SG, Ul Sharif E, O'Doherty GA, Rohr J.

J Am Chem Soc. 2012 Aug 1;134(30):12402-5. doi: 10.1021/ja305113d.

Two enzymes of the gilvocarcin biosynthetic pathway, GilMT and GilM, with unclear functions were investigated by in vitro studies using purified, recombinant enzymes along with synthetically prepared intermediates. The studies revealed GilMT as a typical S-adenosylmethionine (SAM) dependent O-methyltransferase, but GilM was identified as a pivotal enzyme in the pathway that exhibits dual functionality in that it catalyzes a reduction of a quinone intermediate to a hydroquinone, which goes hand-in-hand with a stabilizing O-methylation and a hemiacetal formation. GilM mediates its reductive catalysis through the aid of GilR that provides FADH(2) for the GilM reaction, through which FAD is regenerated for the next catalytic cycle. This unusual synergy eventually completes the biosynthesis of the polyketide-derived defuco-gilvocarcin chromphore.

Baeyer-villiger C-C bond cleavage reaction in gilvocarcin and jadomycin biosynthesis.

Tibrewal N, Pahari P, Wang G, Kharel MK, Morris C, Downey T, Hou Y, Bugni TS, Rohr J.

J Am Chem Soc. 2012 Nov 7;134(44):18181-4. doi: 10.1021/ja3081154.

GilOII has been unambiguously identified as the key enzyme performing the crucial C-C bond cleavage reaction responsible for the unique rearrangement of a benz[a]anthracene skeleton to the benzo[d]naphthopyranone backbone typical of the gilvocarcin-type natural anticancer antibiotics. Further investigations of this enzyme led to the isolation of a hydroxyoxepinone intermediate, leading to important conclusions regarding the cleavage mechanism.