

**Exploration of the Diverse Functions of Cytochrome P450
Monooxygenases Towards the Development of Biocatalysts**

by

Karoline C. Chiou

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2013**

Doctoral Committee:

Professor David H. Sherman, Co-Chair

Professor Carol A. Fierke, Co-Chair

Professor John Montgomery

Professor Stephen W. Ragsdale

© Karoline C. Chiou
All rights reserved
2013

To my parents, sister, and Kyle

Acknowledgements

I would like to express my gratitude to my advisor, Dr. David H. Sherman, for his guidance and support throughout my graduate education. I appreciate the opportunity he provided to work on multiple projects that challenged me to think creatively and independently.

I would like to acknowledge my current committee members Dr. Carol Fierke, Dr. John Montgomery, and Dr. Stephen Ragsdale and previous members Dr. Jason Gestwicki, and Dr. Kate S. Carroll for their suggestions and encouragement.

I have been very fortunate to be a part of dynamic interdisciplinary collaborations driving my dissertational research. Therefore, I am grateful to Dr. Larissa Podust at University of California, San Francisco, Dr. John Montgomery, and Ms. Solymar Negretti from the University of Michigan, Dr. Erik Sorensen and Dr. Junjia Liu from Princeton University, and Dr. Richard Demaio from Eli Lilly and Company for their motivation, encouragement, and feedback.

I would also like to sincerely thank all current and past Sherman laboratory members. They each challenged me to constantly strive for success and persevere through the scientific frustrations. I especially want to thank Dr. Shengying Li, Dr. Yousong Ding, Dr. Patricia Cruz Lopez, Carl Averang, Dr. Ashootosh Tripathi, Dr. Avi Raveh, Mr. Eli Eisman, Dr. Andrew Lowell, Dr. Christopher Rath, Dr. SungRyeol Park, Dr. Tyler Nusca, Dr. Jamie Scaglione, and Ms. Pam Schultz for their insightful discussions, collegial support, and technical assistance throughout my graduate research. Their willingness to provide support and objective perspectives on my work helped me to develop as an independent scientist able to approach projects with a well-rounded standpoint.

I am grateful for funding from the Pharmaceutical Sciences Training Program, Rackham Graduate Student Research Grant, and the Rackham Travel Grant awarded by the Rackham Graduate School.

I would like to express my deep gratitude and love to my parents, Mrs. Bair and Mr. Der-Ming Chiou. They worked to always nurture my passions and support my goals. Without their encouragement and tireless devotion, it would have been impossible to achieve what I have.

They sacrificed everyday to instill a strong work ethic and fostered a will to persevere in me. These qualities have synergistically sustained me through my most difficult times. I am indebted to my grandmother who cared for me as a young child, raising me to be empathetic and patient. Her gentle presence and spirit have been a calming and constant force. Also, I would like to express my love and appreciation for my sister, Ms. Stefanie C. Chiou. She has always been a role model and confidant as my older, wiser sister. I continue to seek her counsel and advice even today.

Finally, I would like to thank my husband, Mr. Kyle Bolduc, for his unwavering support and love. His friendship and unconditional love inspires and motivates me every day to challenge myself to work harder and be better. He has been the source of my self-confidence and renewed enthusiasm not only for science but for my life.

Preface

This thesis contains five chapters covering my dissertational studies on several bacterial biosynthetic cytochrome P450 monooxygenases, including PikC, TyII, and TyIHI. Chapter 1 encompasses a brief overall introduction to cytochrome P450 enzymes, which comprise my major research targets. Chapter 2 is focused on the collaborative bioengineering efforts conducted on an engineered chimeric P450, PikC_{D50N}-RhFRED, in conjunction with Dr. Larissa Podust (University of California, San Francisco) and Dr. John Montgomery and Ms. Solymar Negretti (University of Michigan). This chapter was adapted from a submitted paper to the *Journal of the American Chemical Society*. Chapter 3 investigates the application of PikC_{D50N}-RhFRED for the production of novel bioactives based on FDA-approved drug scaffolds and chemoenzymatic preparation of *in vivo* metabolites of FDA-approved drugs. Chapter 4 is focused on the biochemical and kinetic characterization of two P450s derived from the tylosin biosynthetic pathway, TyII and TyIHI, in collaboration with Dr. Ashootosh Tripathi (University of Michigan) and Dr. Shengying Li (previously at University of Michigan, currently at Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences). Finally, in Chapter 5, I discuss potential future directions for P450 research based on the current dissertational research.

Table of Contents

Dedication	ii
Acknowledgements.....	iii
Preface	v
List of Figures	viii
List of Tables.....	xi
Abstract.....	xii
Chapter	
1. Introduction to Cytochrome Monooxygenase P450 Enzymes.....	1
1.1 Background	1
1.2 P450 History	2
1.3 Catalytic Mechanism of P450s	4
1.4 P450 Structure.....	5
1.5 Classification of P450s	6
1.6 Biosynthetic P450s in Microbial Secondary Metabolism.....	7
1.7 The Pikromycin System.....	9
1.8 References.....	13
2. Regioselective C-H Bond Oxidation by an Engineered P450 Monooxygenase Employing Simple Removable Directing Groups	17
2.1 Introduction.....	17
2.2 Results and Discussion	19
2.2.1 Regioselective Hydroxylation of Synthetic Cycloalkanes	19
2.2.2 Altered Regioselective of 10-dml C-H Functionalization with Linear Desosamine Mimics	23
2.2.3 Product Analysis.....	25
2.2.4 Structural Basis for Regioselectivity of 12-Membered Ring Aglycones with Simplified Linker Anchors.....	27
2.2.5 Antibacterial Activities of Synthetic Compounds Containing Desosaminyll Mimics ...	31
2.3 Conclusion	32

2.4 Experimental Procedures	34
2.5 Supplementary Information	36
2.6 References.....	42
3. P450-Mediated Oxidation of FDA-Approved Drug Scaffolds	46
3.1 Introduction.....	46
3.2 Background.....	48
3.2.1 Tiamulin and Other Pleuromutilins	48
3.2.2 Tamoxifen and Related Analogs	49
3.3 Results and Discussion	51
3.3.1 PikC-Mediated Oxidation of Putative Semisynthetic Substrates	52
3.3.2 PikC Oxidation of Tamoxifen and Related Compounds	53
3.3.3 Regioselective Oxidation of Pleuromutilin Derivatives.....	54
3.4 Conclusion	56
3.5 Experimental Procedures	58
3.6 Supplementary Information	59
3.7 References.....	65
4. Functional Analysis of Engineered P450s TyII-RhFRED and TyIHI-RhFRED Involved in the Biosynthesis of Tylosin	68
4.1 Introduction.....	68
4.2 Results and Discussion	72
4.2.1 Protein Sequence Analysis of TyII and TyIHI.....	72
4.2.2 Heterologous Expression of TyII- and TyIHI-RhFRED.....	73
4.2.3 Functional Analysis of TyII-RhFRED <i>In Vitro</i>	74
4.2.4 Functional Analysis of TyIHI-RhFRED <i>In Vitro</i>	76
4.2.5 Measurement of Substrate Dissociation Constants	77
4.2.6 Steady-State Kinetic Analysis of TyII- and TyIHI-RhFRED.....	76
4.3 Conclusion	78
4.4 Experimental Procedures	80
4.5 Supplementary Information	84
4.6 References.....	93
5. Future Work	96

List of Figures

Figure

1-1 Proposed general P450 catalytic cycle and the peroxide shunt pathway.....	3
1-2 Oxygen rebound mechanism between Compound I and Compound II.....	5
1-3 Classification of P450s based upon redox partner structure.....	7
1-4 Examples of biologically active natural products that have P450s involved in their biosynthesis	8
1-5 Structures of pikromycin, methymycin, and neomethymycin.....	9
1-6 Biosynthetic cluster that produces pikromycin.....	10
1-7 PikC catalyzed reactions and products	11
2-1 Major physiological reactions catalyzed by PikC.....	19
2-2 Schematic of substrate engineering concept using simplified synthetic linkers.....	20
2-3 Comparison of synthetic diastereomeric C6, C7, and C8 standards (6a-d) to the observed PikC _{D50N} -RhFRED oxidation products of compound 6.....	22
2-4 Stacked plot of ¹ H-NMR of hydrolyzed reaction products from substrates 7-9.....	26
2-5 Cycloalkane derivative in the PikC _{D50N} active site.....	28
2-6 Structural insights into regioselectivity of macrocycle hydroxylation	30
2-7 Effect of the D50N mutation on the overall PikC structure.....	31
S2-1 Representative LC-MS Q-TOF analyses of PikC _{D50N} -RhFRED oxidations	38
S2-2 Structural determination of mono-hydroxylated products of 8 through LC-MS comparison of synthetic authentic standards regarding retention times	39
3-1 Enzymatic scheme of PikC-mediated oxidation of unnatural substrates.....	48

3-2 Structures of tiamulin, retapamulin (SB-275833), and azamulin	49
3-3 Structures of tamoxifen and toremifene.....	50
3-4 Endogenous PikC-catalyzed reactions of YC-17 and narbomycin.....	51
S3-1 Semisynthetic compounds not tolerated by PikC _{D50N} -RhFRED as substrates	59
S3-2 Commercially available FDA-approve compounds not tolerated by PikC _{D50N} -RhFRED as substrates	59
S3-3 LC-MS Q-TOF traces of PikC enzymatic reactions against semisynthetic compounds	60
S3-4 LC-MS Q-TOF traces of PikC enzymatic reactions.....	61
S3-5 LC-MS Q-TOF traces of PikC enzymatic reactions against pleuromutilin derivatives	62
S3-6 ¹ H-NMR of 3-(dimethylamino)propanoic acid	63
S3-7 ¹ H-NMR of 4	64
4-1 Structures of erythromycin A and semisynthetic erythromycin analogs and the oxidation patterns of EryF (red) and EryK (blue)	69
4-2 Structures of pikromycin, methymycin, and neomethymycin with the sites of PikC hydroxylation highlighted in green	70
4-3 The tylosin post-PKS biosynthetic pathway and organization of the tylosin biosynthetic gene cluster	71
4-4 Phylogenetic tree of select macrolide biosynthetic P450s.....	73
4-5 First <i>tylI</i> catalyzed reaction.....	74
4-6 Second <i>tylI</i> catalyzed reaction	75
4-7 Hypothesized reaction catalyzed by <i>tylHI</i>	76
S4-1 Purification of TylI- and TylHI-RhFRED	85
S4-2 CO-bound reduced difference spectra	86
S4-3 Representative LC-MS Q-TOF traces of TylI-RhFRED and TylHI-RhFRED enzymatic reactions	87

S4-4 Biochemical characterization of TyII-RhFRED	88
S4-5 Biochemical characterization of TyIHI-RhFRED	88

List of Tables

Tables

2-1 Panel of initial ether and ester linker compounds screened against PikC _{D50N} -RhFRED.....	22
2-2 10-dml analogs with desosaminyl mimics screened against PikC _{D50N} -RhFRED.....	24
2-3 Linker induced regioselectivity differences and Michaelis-Menten fitted standard-state kinetic parameters	27
S2-1 Binding affinities of PikC _{WT} , PikC _{WT} -RhFRED, and PikC _{D50N} -RhFRED for compounds 7-9	37
S2-2 Regioselectivity and product yield differences exhibited by PikC _{WT} , PikC _{WT} -RhFRED, and PikC _{D50N} -RhFRED against compounds 7-9	40
S2-3 Antibacterial activities of synthetic compounds containing desosaminyl mimics	41
3-1 Regioselectivity and conversion of PikC against semisynthetic compounds	53
3-2 PikC regioselectivity and conversion of tamoxifen derivatives	54
3-3 PikC regioselective oxidation of pleuromutilin derivatives.....	56
4-1 Binding and steady-state kinetic analyses of TylI- and TylHI-RhFRED	78
S4-1 ¹³ C and ¹ H NMR chemical shift tabulation for compound 1.....	89
S4-2 ¹³ C and ¹ H NMR chemical shift tabulation for compound 2.....	90
S4-3 ¹³ C and ¹ H NMR chemical shift tabulation for compound 4.....	91
S4-4 ¹³ C and ¹ H NMR chemical shift tabulation for compound 5.....	92

ABSTRACT

Exploration of the Diverse Functions of Cytochrome P450 Monooxygenases Towards the Development of Biocatalysts

by

Karoline C. Chiou

Chair: David H. Sherman

The superfamily of cytochrome P450 monooxygenases is involved in diverse oxidative processes including xenobiotic detoxification, steroid synthesis, and biosynthetic tailoring of diverse secondary metabolites known as natural products. Members of this superfamily catalyze a vast range of reactions from hydroxylations and epoxidations to isomerizations and ring formations and expansions. Given the ease of cloning, protein overexpression, and purification, bacterial biosynthetic P450s have garnered special interest because of their catalytic efficiency and high regio- and stereoselectivity. Nevertheless, substantial barriers still exist to the application of P450s for biotechnology and synthetic means, which include inherent instability, dependence on redox partners, and a limited substrate scope. Current research efforts are ongoing to overcome those limitations, pushing this superfamily of enzymes towards their successful integration in the production of fine chemicals, pharmaceuticals, biofuels, and bioremediation tools.

My dissertational work has focused on investigating P450s as potential biocatalysts for the production of high value pharmaceuticals through exploration of substrate binding and conversion. First, a substrate-engineering approach elaborated on a preexisting anchoring hypothesis revealing unprecedented flexibility of an engineered chimeric P450, PikC_{D50N}-RhFRED, previously developed from utilizing a reductase domain, RhFRED, from a

Rhodococcus species. Furthermore, we gained valuable insight into the fundamental, yet elusive, factors affecting P450-mediated oxidation, such as substrate binding, orientation, and product formation. The results culminated in the development of an optimized linear linker that efficiently replaced the natively used sugar anchor. These advancements eliminated labor-intensive synthetic steps to acquire and attach the sugar to potential substrates, opening to the door to chemoenzymatic elaboration.

To further expand upon PikC_{D50N}-RhFRED research, the chimeric enzyme was employed to selectively oxidize structurally distinct scaffolds with pharmaceutical applications, including tamoxifen and tiamulin. A proportion of these compounds has already been approved by the Federal Drug Administration (FDA) and are currently used for human patients. These results underscored PikC's flexibility and potential as a biocatalyst, as it can accept very different molecules for oxidation as long as a terminal *N,N*-dimethylamino linker is attached. Furthermore, the regioselective oxidation of these compounds in modest yields highlighted the potential use of this enzyme towards predictive production of *in vivo* metabolites and development of potential drug analogs.

Finally, two P450s from the tylosin biosynthetic pathway, TyII and TyIHI, were functionally characterized to confirm their identities as P450s and determine their biosynthetic order. The enzymes were utilized as chimeras with RhFRED, a reductase domain from a *Rhodococcus* species, to afford single component, self-sufficient P450s. Using substrates isolated from *Streptomyces fradiae* mutant fermentation, TyII-RhFRED and TyIHI-RhFRED were confirmed to be P450 hydroxylases of tylosin intermediates and their sites of oxidation confirmed through rigorous structural elucidation. TyII-RhFRED was especially interesting given its ability to perform sequential oxidations to form an aldehyde crucial to tylosin's bioactivity. However, this study also leaves a number of unanswered questions for future exploration of these unique P450s, mainly the method of substrate recognition employed by these two biosynthetic P450s and the source of the inherent flexibility of a P450 to perform sequential oxidations.

Chapter 1

Introduction to Cytochrome Monooxygenase P450 Enzymes

1.1 Background

Cytochrome P450 monooxygenases (P450s) have been coined “Nature’s most versatile biological catalyst” given their exquisite ability to catalyze an impressive range of regio- and stereoselective reactions that include hydroxylations, epoxidations, dealkylations, phenolic couplings, ring formations and expansions, dehydrations, and isomerizations.^{1,2} Furthermore, the presence of P450s is ubiquitous, having been found across all domains of life, including archaea, bacteria, plants, and mammals. These enzymes are commonly used in essential oxidative processes such as xenobiotic detoxification, steroid synthesis, and the biosynthetic tailoring of chemically diverse secondary metabolites.^{1,3} The P450-installed functionality often provides an important layer of structural variability in the mature natural products that significantly influence biological activity.^{4,5} From a chemical standpoint, P450s catalyze an eclectic variety of challenging, synthetically relevant reactions on physiologically and biotechnologically important molecules with a high degree of regio- and stereoselectivity under mild conditions, emphasizing their synthetic potential as biocatalysts.⁶⁻⁸

In spite of the number of P450s steadily increasing since their original discovery, powerful next-generation genome sequencing and bioinformatics technology has allowed for a unique and rapid boom in the number of identified P450 family members.⁹⁻¹¹ To date, 57 human P450s (CYPs) have been identified through the human genome project; 58 P450s have been found in the genome of *Glycine max* (soybean) and 18 distinct P450s have been annotated in *Streptomyces coelicolor*.¹²⁻¹⁴ Furthermore, the Protein Data Bank (PDB) currently houses nearly 500 P450 structures, including 30 mammalian P450s. To date, there are >18,000 P450 sequences available and the number is quickly growing.¹⁵

Despite the divergent evolution of P450 enzymes that has led to an exceedingly diverse enzyme superfamily characterized by rather low amino acid sequence identity (<20%), there are key features that facilitate a general catalytic mechanism of dioxygen activation, including the obligate heme cofactor with the iron protoporphyrin IX center coordinated to a thiolate ligand provided by the absolutely conserved cysteine residue. Furthermore, P450s share a highly similar three-dimensional structure originating from the equivalent arrangement of common structural elements comprising of well-conserved helices denoted A-L and a portion of helix I proximal to the heme-iron implicated in the proton delivery of the O-O bond cleavage to generate the highly reactive oxidant.^{2,16}

1.2 P450 History

P450s were initially reported in 1958 as a pigment found in rat liver microsomes but did not acquire their characteristic name until later.⁹ Following that report, Omura and Sato rediscovered the system and formally named it P450 where the “P” arose from the word pigment and 450 was derived from the characteristic spectral maximum at 450 nm upon binding of carbon monoxide (CO) to the ferrous heme iron.^{10,11} Early studies focused on P450 involvement in metabolism of carcinogens, drugs, pesticides, vitamins, and steroids using tissue samples or tissue-derived microsomes containing CYPs.¹⁶⁻²⁰ It was not until 1970 that the first purified microsomal P450s were functionally analyzed. With the dawn of recombinant DNA technology and heterologous protein overexpression methods in the 1980s, an increased number of eukaryotic and prokaryotic P450s became widely accessible for scrutiny.²¹⁻²⁵ Then in 1987, the currently used P450 nomenclature system rooted in sequence identity was developed by Nebert to facilitate unambiguous discussions of members of the enzyme family.²⁶ Later the same year, the first crystal structure of a P450 was reported for P450_{cam} from *Pseudomonas putida* after which a number of high-resolution crystal structures were solved for various P450s, including the first self-sufficient P450_{BM3} from *Bacillus megaterium* and 2 separate human P450s, 2D6 and 3A4.²⁷⁻³⁰

In the past two decades, a considerable proportion of P450 research has concentrated on the major roles of human hepatic CYPs, given their substantial role in xenobiotic metabolism. Therefore, the pharmaceutical industry devoted a significant amount of resources to predicting bioavailability, drug-drug interactions, and toxicity.¹⁶ However, there was mounting interest in

more fully understanding P450s to make technological headway in areas such as plant breeding, bioremediation, insect control, and chemical carcinogenesis.^{3,31} Current discussion within the P450 research community targets mechanistic details of reactive species generation during the catalytic cycle, electron/proton transfer for dioxygen activation, structural elucidation of unique P450s with biomedical significance or application, discovery of P450 inhibitors, and functional characterization of novel P450s found through genome mining efforts.^{16,32,33} In particular, there is increasing interest in developing P450s for synthetic applications through overcoming major limitations such as intrinsic instability, reliance on separate redox partners, and narrow substrate specificity through directed evolution, use of native or engineered self-sufficient P450s, and the expansion of the substrate spectrum via protein and/or substrate engineering.^{6,34-40} P450s offer an exceptional and promising pool of enzymes with the potential to be developed as powerful biocatalysts able to produce industrial and fine chemicals, pharmaceutically relevant compounds, and participate in bioremediation.^{6,41}

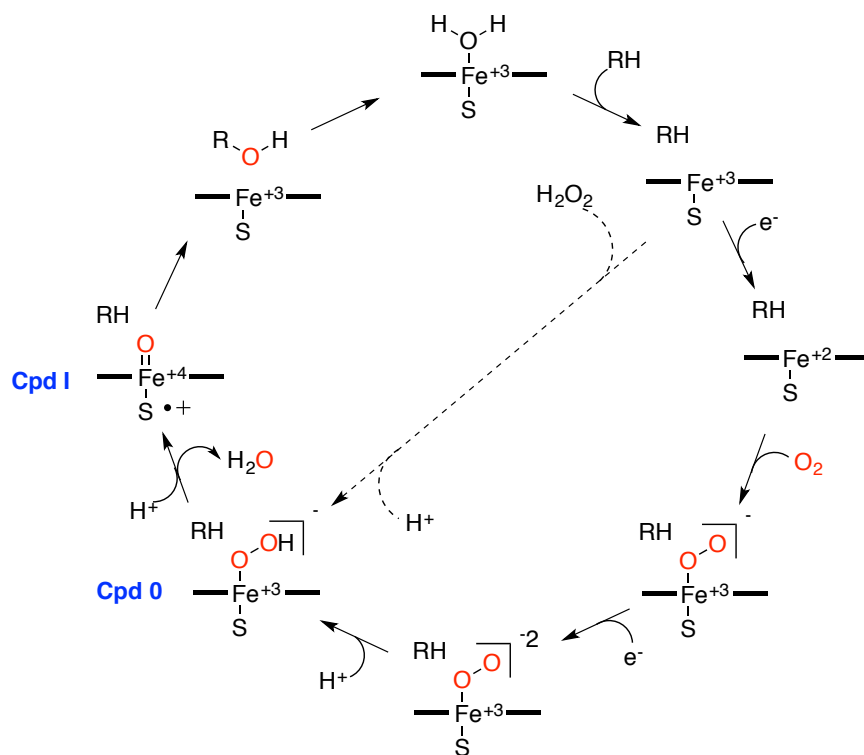


Figure 1-1. Proposed general P450 catalytic cycle and the peroxide shunt pathway. The catalytic cycle of a typical P450 is depicted to highlight the hypothesized oxidative intermediates that include Compound 0 (Cpd 0) and Compound I (Cpd I).

1.3 Catalytic Mechanism of P450s

Although the evolution of P450s is highly divergent, P450s share a common mechanism of reductively activating molecular oxygen through the generation of specific reactive intermediates. Although the exact nature of some of these reactive species is still hotly debated within the literature, observation of some intermediates using various spectroscopic techniques, radical clocks, simultaneous studies of synthetic metalloporphyrins, and theoretical studies have established a general paradigm for the catalytic mechanism of cytochrome P450s.^{16,32,42,43}

In the resting state of the enzyme, a water molecule acts as the sixth ligand to the ferric heme-iron (Fe^{+3}), resulting in a low spin of the heme-iron with a representative absorbance at 420 nm. Upon substrate binding, the water molecule is displaced, inducing a change in the spin state of the heme-iron from low-spin to high-spin. This gives rise to the spectral change of the enzyme with an increase in absorbance at 390 nm and a simultaneous decrease in the absorbance at 420 nm. This difference spectrum is referred to as a “Type I” difference spectrum, which is easily measured by ultraviolet/visible (UV/VIS)-spectrometry and indicates the binding of a substrate.⁴⁴ On the other hand, inhibitors and some other substrates coordinate directly to the heme to generate a “Type II” difference spectrum where the maximal absorbance occurs at 430 nm and the minimum occurs at 390 nm.⁴⁴ In addition to the change of the heme-iron to high-spin, loss of the water ligand also induces subtle changes in the position of the iron relative to the plane of the porphyrin ring, allowing the heme to be a better electron sink and trigger the first electron to be transferred from NAD(P)H via redox partners. Upon this one electron reduction, the ferrous heme-iron (Fe^{+2}) forms a relatively stable dioxygen adduct. The subsequent second electron transfer, often the rate-limiting step of the mechanism, is followed by a rapid protonation step to generate the first reactive species known as the ferric hydroperoxy (Fe^{+3} -OOH) intermediate also referred to as Compound 0. Following a second protonation and heterolytic cleavage of the O-O bond with concurrent production of a H_2O molecule, the oxidative ferryl-oxo species ($\text{Fe}^{+4}=\text{O}$), otherwise known as Compound I, is formed. This is a highly reactive cation radical believed to be directly responsible for oxygen insertion. A number of studies indicate that either Compound 0 or Compound I are able to act as the oxidative catalytic species.^{16,32,43} The choice between the two of them has been suggested to be substrate dependent and the rate of proton transfer steps relevant to the inherent nature of particular P450s.

In the final step, the oxidized product is released from the active site completing the catalytic cycle and restoring the enzyme to its resting state.

Other aspects of the catalytic mechanism include the peroxide shunt pathway and the oxygen rebound mechanism. The peroxide shunt pathway illustrated in **Fig. 1-1** with the dotted lines utilizes hydrogen peroxide (H_2O_2) or alternative oxygen donors to directly generate Compound 0 (Cpd 0) from the ferric substrate-bound form. The oxygen rebound mechanism shown in **Fig. 1-2** highlights hydrogen abstraction from the substrate by compound I to form compound II before rebound of the hydroxyl to the substrate radical to form the oxidized product.

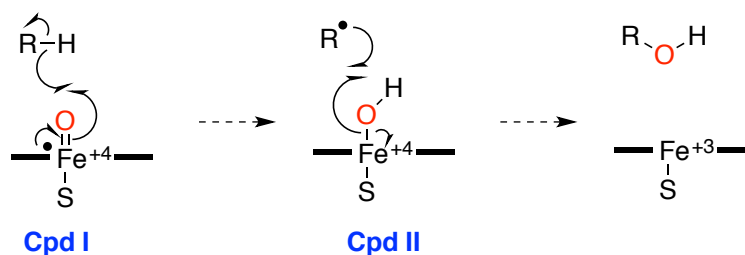


Figure 1-2. Oxygen rebound mechanism between Compound I and Compound II. This expands upon the region of the catalytic cycle between Cpd I and hydroxylated product formation and release to show the hydrogen abstraction from the substrate by Cpd I to form Cpd II, prior to rebound of the hydroxyl to the substrate radical to form the hydroxylated product.

1.4 P450 Structure

The three-dimensional structure of P450s is generally conserved in spite of low sequence identity across this superfamily of enzymes.^{16,45} The P450 general structure exhibits a number of helices denoted A-L and five additional β sheets denoted β 1-5; the sequence and structure of the helices and the sheets are more conserved with increased proximity to the heme-iron. Thus, helices I and L, found closest to the heme, contain the most highly conserved regions. In addition, the “Cys-Pocket” located at the β -bulge segment immediately prior to helix L houses the absolutely conserved cysteine residue that coordinates to the iron protoporphyrin center. Furthermore, many portions of helix I are well-conserved as it contains several residues directly responsible for oxygen activation. Of particular importance is a threonine residue, which produces local helical distortions and participates in a hydrogen bond network with several other peptide carbonyl groups and several water molecules. However, this threonine is not absolutely conserved as alanine is occasionally found in specific P450s.⁴⁶ In those cases, a water molecule

often replaces the threonine side chain hydroxyl group and a hydroxyl group from the substrate also participates.⁴⁶ This hydrogen bond network plays a crucial role in maintaining proper delivery of protons to the reactive center for O-O cleavage and generation of reactive oxidative species.

1.5 Classification of P450s

There have been many systems developed to classify P450s based on a number of criteria. One of those systems relies on the identified enzymatic activity of a P450, which includes hydroxylases, epoxidases, desaturases, and isomerases.² Another system classifies them according to their sources, such as humans, fungus, bacteria, or plants. Alternatively, P450s can be classified based on their sequence similarity, with 40% protein sequence identity defining families and more than 55% identity found among subfamily members. Perhaps the most useful and important classification system for these P450s is based on the redox partners required for electron transfer in the catalytic cycle.

Traditionally, there were two major classes of redox systems (**Fig. 1-3**). Class I was a FAD-containing reductase with a small iron-sulfur redoxin for most bacterial and mitochondrial P450s. Class II was a FAD/FMN-containing flavoprotein-cytochrome P450 reductase (CPR) for eukaryotic microsomal P450s. However, when P450_{BM3} was discovered, it was clear that this P450 signified a new class because it was a naturally fused to a eukaryotic-like CPR.⁴⁷ More recently, another self-sufficient P450 emerged in the form of P450_{RhF} from *Rhodococcus* sp. NCIMB 9784 and found to be naturally fused to a novel FMN/Fe₂S₂-containing reductase domain similar to the phthalate family of dioxygenase reductases.^{48,49} The emergence of these naturally occurring self-sufficient P450s became the third class of P450s. These enzymes were unique given their independence from separate redox partners and exposed their potential as biocatalysts. In addition, these fusion P450s often demonstrated higher activity, most likely derived from a more efficient electron transport system given the covalent linkage between the redox partner and the P450.^{36,50} Finally, there exists a class of P450s that are independent of the activity of redox partners for catalysis (Class V) exemplified by the soluble P450_{nor} from *Fusarium oxysporum*,⁵¹

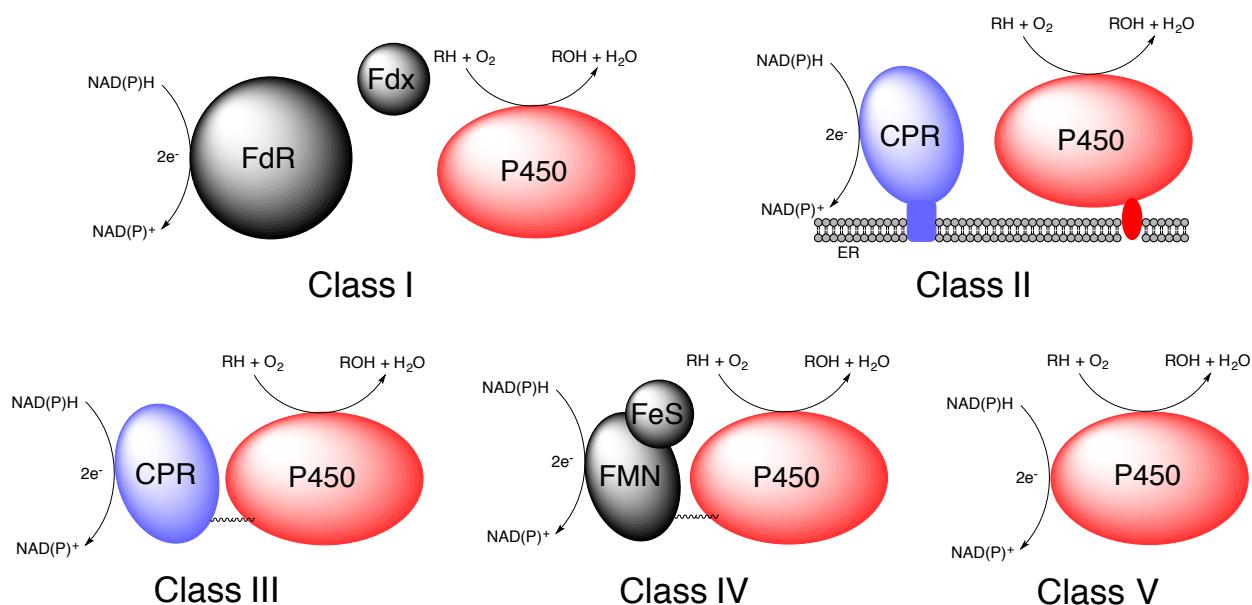


Figure 1-3. Classification of P450s based upon redox partner structure. Fdr: Ferredoxin reductase; Fdx: Ferredoxin; CPR: Cytochrome P450 reductase.

1.6 Biosynthetic P450s in Microbial Secondary Metabolism

The majority of P450s found from microbial sources are involved in the biosynthesis of secondary metabolites.^{16,52} Presently, the native role of many of these enzymes has yet to be elucidated. Therefore, functional analysis of soluble forms of the enzymes will most likely yield new chemistry, give additional insight into the biosynthesis of novel bioactive natural products, offer additional examples to build a consensus model of CYP structure and catalysis, and provide additional candidates for development into biocatalysts.

Actinomycetes, especially *Streptomyces*, are a rich source of secondary metabolites and account for more than two-thirds of microbially derived compounds. These compounds have important roles in signaling or chemical defense for the host organism and many of these compounds have demonstrated biological activity that include antibacterial (erythromycin, pikromycin, tirandamycin), antifungal (amphotericin, nystatin), antiparasitic (ivermectin), immunosuppressive (FK506, rapamycin), growth promotion (tylosin), and anticancer (doxorubicin, mitomycin C) (**Fig. 1-4**).⁵³⁻⁶⁷ All of those compounds mentioned have P450s involved in their biosynthesis, underscoring the importance of these enzymes for producing naturally bioactive molecules. Therefore, elucidation of biosynthetic pathways as well as structural identification of P450 substrates, intermediates, and products, allows for functional

analysis of novel P450s towards the eventual production of crucial biologically active compounds.

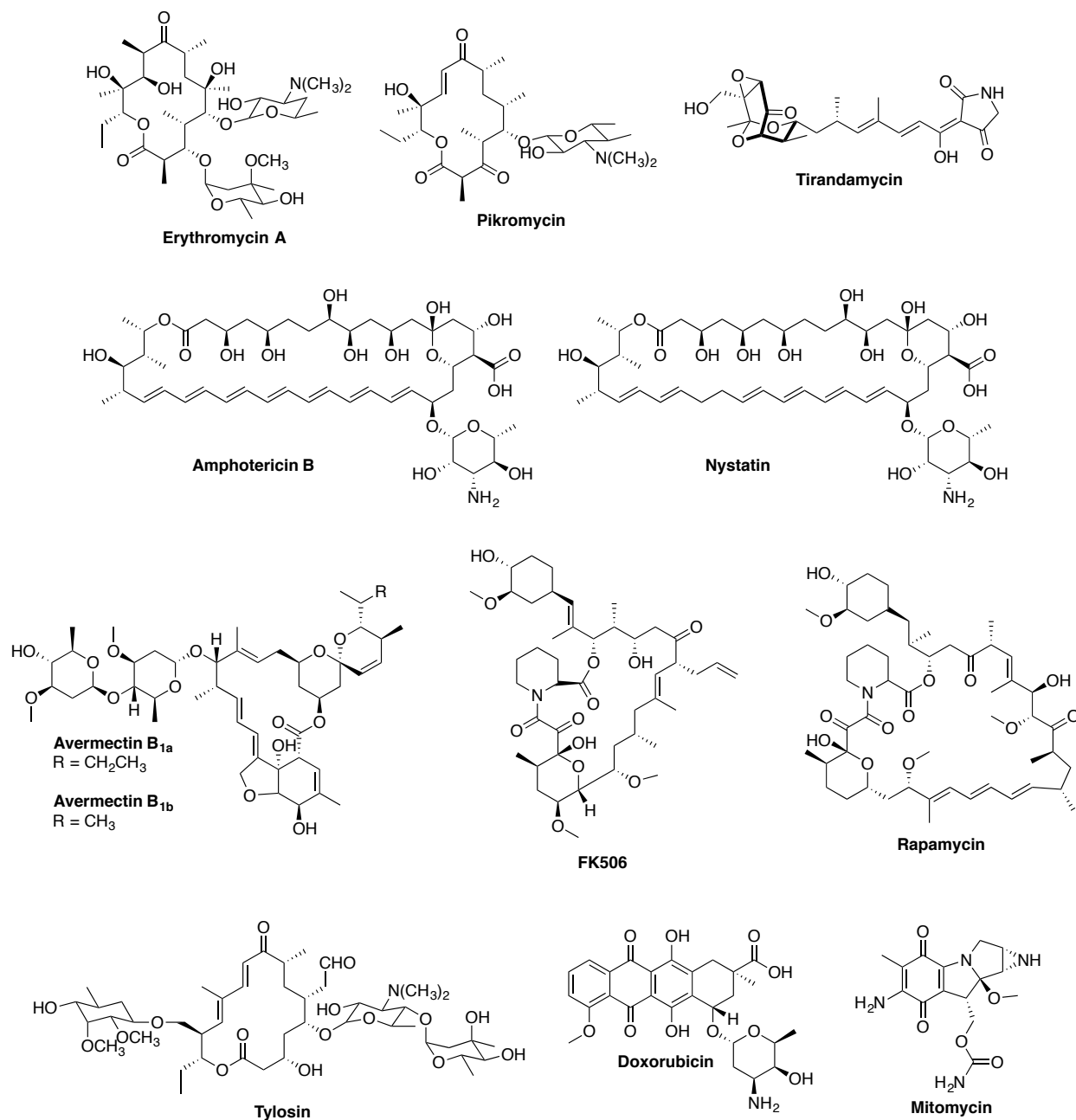


Figure 1-4. Examples of biologically active natural products that have P450s involved in their biosynthesis.

In addition to traditional methods of identifying new P450s, modern whole genome sequencing technology offers a decidedly more efficient process for accessing new gene products. Recent efforts have demonstrated that a large number of P450 genes can be identified. For example, the model actinomycete *Streptomyces coelicolor* A3(2) was found to have 18

distinct P450s, and *Streptomyces avermitilis* MA-4680 contained 33 P450s.^{14,65} Nevertheless, there still remains difficulty in confirming the physiological function of each of these P450s.

1.7 The Pikromycin System

Pikromycin was originally isolated in 1951, representing the first discovered macrolide antibiotic.⁶⁸ Two structurally related macrolides, neomethymycin and methymycin, were later discovered to be produced by *Streptomyces venezuelae* (**Fig. 1-5**).^{69,70} For more than a decade, the Sherman laboratory has investigated many aspects of pikromycin biosynthesis that spans pathway elucidation and comprehension of enzymatic mechanisms of a unique modular type I polyketide synthase (PKS) and tailoring enzymes. The knowledge acquired greatly advanced understanding of general macrolide biosynthesis, joining the erythromycin pathway as a model pathway in the field of natural product biosynthesis.^{71,72}

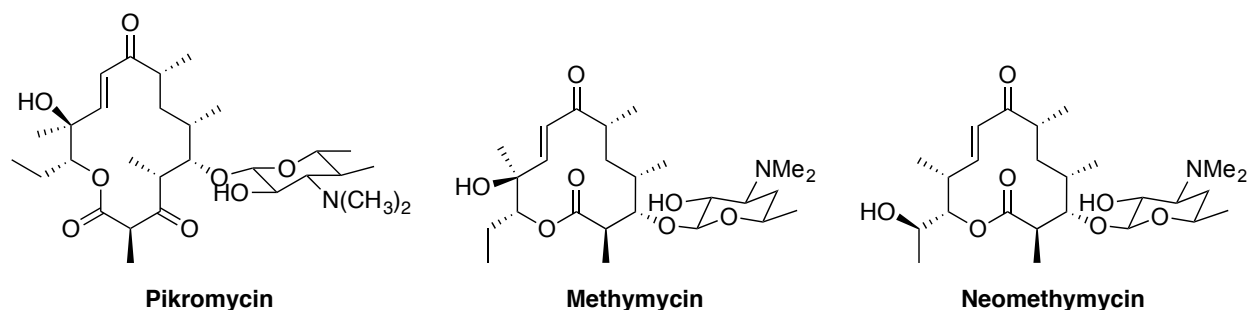


Figure 1-5. Structures of pikromycin, methymycin, and neomethymycin.

The first hurdle for the inquiries was sequencing and assembling the pikromycin pathway, which revealed a gene cluster encoding a five modular type I polyketide synthase (*pikAI-V*), a group of genes responsible for the synthesis of the 6-deoxysugar desosamine (*desI-VIII* and *desR*), a P450 (*pikC*), and a number of regulatory (*pikD*) and self-resistant (*pikR1* and *pikR2*) genes. Subsequent gene disruption studies unambiguously demonstrated that the PKS was unique from other established PKS systems given its ability to generate two distinct macrolactone scaffolds, 10-deoxymethynolide (10-dml) and narbonolide (**Fig. 1-6**).^{73,74} Other research groups have studied desosamine biosynthesis and functionally characterized all enzymes required for transformation of D-glucose-1-phosphate to thymidine diphosphate D-desosamine (TDP-desosamine).^{71,75-78} Specifically, DesVII in conjunction with DesVIII demonstrated unprecedented substrate flexibility towards the aglycone and the sugar moiety, which was later exploited to generate novel unnatural substrates with improved biological

activity.⁷⁹⁻⁸² In addition to the sugar genes, the thioesterase of module V and the glycosyltransferase DesVII have also demonstrated unusual enzymatic flexibility.⁷³

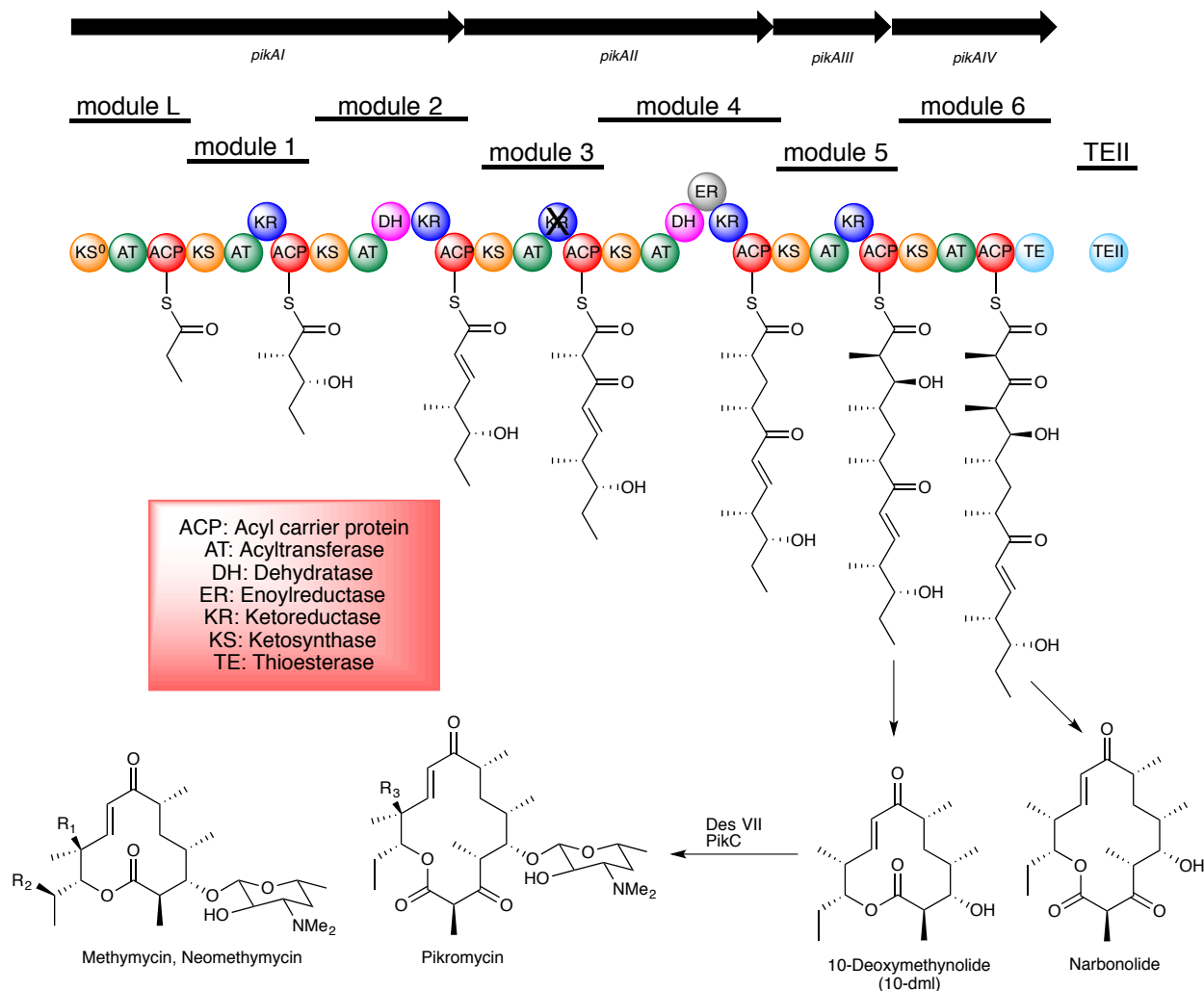


Figure 1-6. Biosynthetic cluster that produces pikromycin. The pikromycin polyketide synthase is comprised of a one loading module and six elongation modules that span four polypeptide chains (PikAI-PikAIV). Chain elongation through PikAIII followed by thioesterase catalyzed termination results in the 12-membered ring macrolactone product, 10-deoxymethynolide, while continued elongation of the polyketide chain produces the 14-membered ring macrolactone, narbonolide. Both products undergo further processing via a glycosyltransferase and P450 hydroxylase to yield methymycin and pikromycin. PikAV is a type II thioesterase that functions to remove pre-maturely decarboxylated extended units from ACP domains.

Of additional interest is the cytochrome P450 involved in the biosynthetic pathway of pikromycin, *pikC*. *PikC* catalyzes the hydroxylation at the C10 position of the 12-membered ring macrolide YC-17 to yield methymycin and at the C12 position to generate neomethymycin (**Fig. 6**). It is able to further oxidize methymycin at C12 to give rise to the doubly hydroxylated product novamethymycin.^{60,83} Furthermore, *PikC* is able to hydroxylate the 14-membered ring

substrate, narbomycin, at the C12 position leading to pikromycin and the C14 position leading to neopikromycin (**Fig. 1-7a**). Similarly, PikC can further hydroxylate pikromycin to yield novapikromycin (**Fig. 1-7b**).⁸⁴

A.

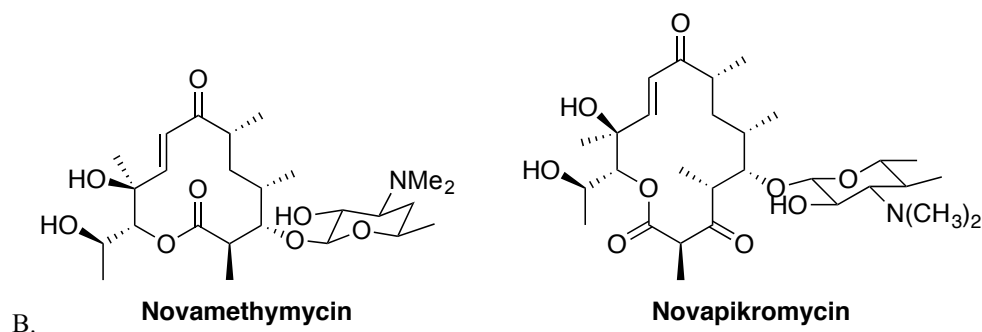
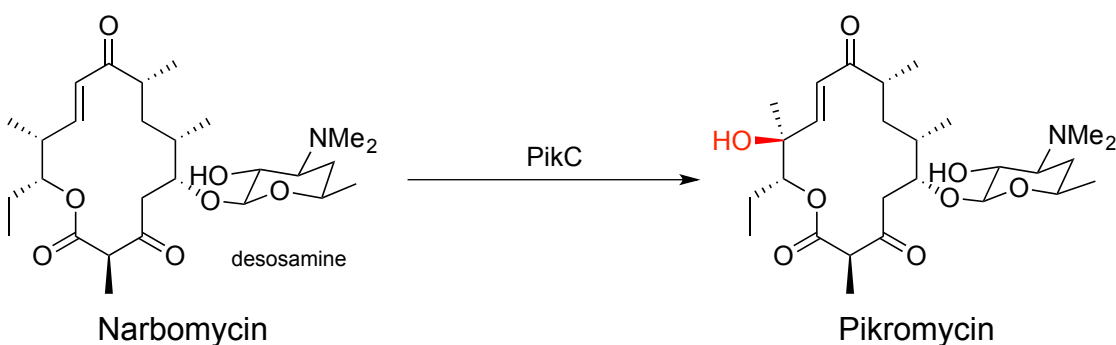
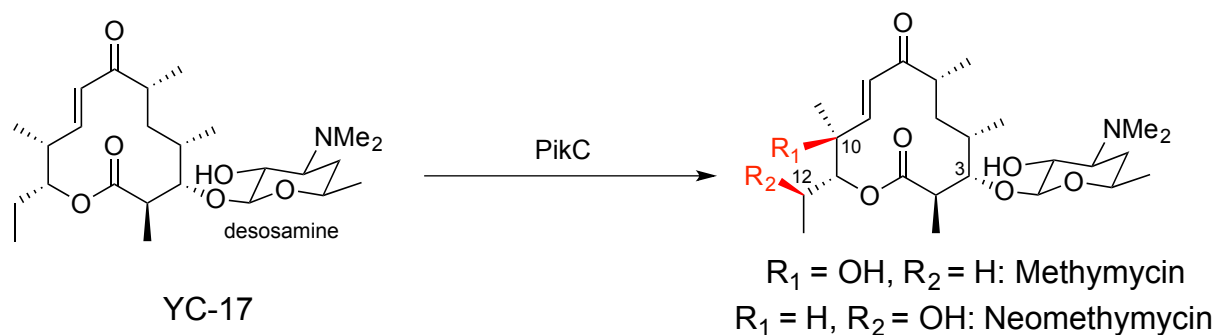


Figure 1-7. PikC catalyzed reactions and products. (A) Endogenous reactions catalyzed by PikC to convert YC-17 to either methymycin or neomethymycin and narbomycin to pikromycin (B) Structures of novamethymycin and novapikromycin.

This dissertational study grew from the work of Dr. Shengying Li and his efforts to understand the structural basis of the substrate flexibility and oxidation variability of PikC in collaboration with Dr. Larissa M. Podust (University of California San Francisco) and Dr. John Montgomery (University of Michigan). Building upon Dr. Li's original desosamine anchoring hypothesis, enzymatic studies exposed even more impressive substrate flexibility and tolerance

than originally observed, propelling PikC further towards being an efficient and cost-effective biocatalyst. To date, the results have demonstrated that firstly, PikC may serve as a model for studying fundamental properties of P450s and their catalytic abilities and secondly, PikC has enormous potential to be developed into an applicable catalyst for synthetic chemistry. During the investigations of PikC with synthetic substrates, PikC was also found to act on chemically distinct scaffolds found in several U.S. Food and Drug Administration (FDA)-approved drugs. Moreover, knowledge and experience gained during the PikC investigation was expanded to include two unique biosynthetic P450s from the tylosin biosynthetic pathway, one of which was hypothesized to be responsible for installation of the rarely observed aldehyde moiety of the mature natural product.

1.8 References:

- 1 Coon, M. J. Cytochrome P450: Nature's most versatile biological catalyst. *Ann Rev Pharmacol Toxicol* **45**, 1-25 (2005).
- 2 Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem Res Toxicol* **14**, 611-650 (2001).
- 3 Bernhardt, R. Cytochrome P450s as versatile biocatalysts. *J Biotechnol* **124**, 128-145 (2006).
- 4 Betlach, M. C., Kealey, J. T., Betlach, M. C., Ashley, G. W. & McDaniel, R. Characterization of the macrolide P-450 hydroxylase from *Streptomyces venezuelae* which converts narbomycin to picromycin. *Biochemistry* **37**, 14937-14942 (1998).
- 5 Mendes, M. V. *et al.* Engineered biosynthesis of novel polyenes: a pimaricin derivative produced by targeted gene disruption in *Streptomyces natalensis* *Chem Biol* **8**, 635-644 (2001).
- 6 Urlacher, V. B. & Eiben, S. Cytochrome P450 monooxygenases: perspectives for synthetic application. *Trends in Biotechnology* **24**, 324-330 (2006).
- 7 Sono, M., Roach, M. P., Coutler, E. D. & Dawson, J. H. Heme-containing oxygenases. *Chemical Reviews* **96**, 2841-2887 (1996).
- 8 Guengerich, F. P. Mechanisms of cytochrome P450 substrate oxidation. *J Biochem Mol Toxicol* **21**, 163-168 (2007).
- 9 Klingenberg, M. Pigments of rat liver microsomes. *Arch Biochem Biophys* **75**, 376-386 (1958).
- 10 Omura, T. & Sato, R. The carbon monoxide-binding pigment of liver microsomes I. Evidence for its hemoprotein nature. *Journal of Biological Chemistry* **239**, 2370-2378 (1964).
- 11 Omura, T. & Sato, R. The carbon monoxide-binding pigment of liver microsomes II. Solubilization, purification, and properties. *Journal of Biological Chemistry* **239**, 2379-2385 (1964).
- 12 <http://www.phytozome.net/soybean>.
- 13 International Human Genome Sequencing and Consortium. *Nature* **409**, 860-921 (2001).
- 14 Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147 (2002).
- 15 Guengerich, F. P. & Cheng, Q. Orphans in the human cytochrome P450 family: approaches to discovering function and relevance to pharmacology. *Pharmacol Rev* **63**, 684-699 (2011).
- 16 Montellano, P. R. O. d. *Cytochrome P450: Structure, Mechanism, and Biochemistry*. 3 edn, (Kluwer Academic/Plenum Publishers, 2005).
- 17 Estabrook, R. W., Cooper, D. Y. & Rosenthal, O. The high reversible carbon monoxide inhibition of the steroid C21-hydroxylase system of the adrenal cortex. *Biochem Z* **338**, 741-755 (1963).
- 18 Brodie, B. B., Gillette, J. R. & Du, B. N. L. Enzymatic metabolism of drugs and other foreign compounds. *Annu Rev Biochem* **27**, 427-454 (1958).
- 19 Conney, A., Miller, E. & Miller, J. The metabolism of methylated aminoazo dyes: V. Evidence for induction of enzyme synthesis in the rat by 3-methylcholanthrene. *Cancer Res* **16**, 450-459 (1956).
- 20 Kamataki, T. & Neal, R. A. Metabolism of diethyl *p*-nitrophenyl phosphorothionate (parathion) by a reconstituted mixed-function oxidase enzyme system: studies of the covalent binding of the sulfur atom. *Mol Pharmacol* **12**, 933-944 (1976).
- 21 Barnes, H. J., Arlotto, M. P. & Waterman, M. R. Expression and enzymatic activity of recombinant cytochrome P450 17 α -hydroxylase in *Escherichia coli*. *Proc Natl Acad Sci USA* **88**, 5597-5601 (1991).
- 22 Larson, J. R., Coon, M. J. & Porter, T. D. Alcohol-inducible cytochrome P-450IIE1 lacking the hydrophobic NH₂-terminal segment retains catalytic activity and is membrane-bound when expressed in *Escherichia coli*. *J Biol Chem* **266**, 7321-7324 (1991).
- 23 Lu, A. Y. H. & Coon, M. J. Role of hemoprotein P-450 in fatty acid ω -hydroxylation in a soluble enzyme system from liver microsomes. *J Biol Chem* **243**, 1331-1332 (1968).
- 24 Wang, P., Mason, P. S. & Guengerich, F. P. Purification of human liver cytochrome P-450 and comparison to the enzyme isolated from rat liver. *Arch Biochem Biophys* **199**, 206-219 (1980).
- 25 Gonzalez, F. J. The molecular biology of cytochrome P450s. *Pharmacol Rev* **40**, 243-288 (1988).
- 26 Nebert, D. W. *et al.* The P450 gene superfamily: recommended nomenclature. *DNA* **6**, 1-11 (1987).
- 27 Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *J Mol Biol* **195**, 687-700 (1987).
- 28 Ravichandran, K. G., Boddupalli, S. S., Hasermann, C., Peterson, J. A. & Deisenhofer, J. Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450s. *Science* **261**, 731-736 (1993).
- 29 Rowland, P. *et al.* Crystal structure of human cytochrome P450 2D6. *J Biol Chem* **281**, 7614-7622 (2006).

- 30 Williams, P. A. *et al.* Crystal structures of human cytochrome P450 3A4 bound to metyrapone and
progesterone. *Science* **305**, 683-686 (2004).
- 31 O'Reilly, E., Köhler, V., Flitsch, S. L. & Turner, N. J. Cytochromes P450 as useful biocatalysts: addressing
the limitations. *ChemComm* **47**, 2490-2501 (2011).
- 32 Meunier, B., Visser, S. P. d. & Shaik, S. Mechanism of oxidation reactions catalyzed by cytochrome P450
enzymes. *Chem Rev* **104**, 3947-3980 (2004).
- 33 Anzai, Y. *et al.* Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in
biosynthesis of mycinamicin macrolide antibiotics. *Chem. Biol.* **15**, 950-959 (2008).
- 34 Cherry, J. R. Directed evolution of microbial oxidative enzymes. *Curr Opin Biotechnol* **11**, 250-254
(2000).
- 35 Munro, A. W. *et al.* P450 BM3: the very model of a modern flavocytochrome. *Trends Biochem Sci* **27**,
250-257 (2002).
- 36 Li, S., Podust, L. M. & Sherman, D. H. Engineering and analysis of self-sufficient biosynthetic cytochrome
P450 PikC fused to the RhFRED reductase domain. *Journal of the American Chemical Society* **129**, 12940-
12941 (2007).
- 37 Li, S. *et al.* Selective oxidation of carbolide C-H bonds by an engineered macolide P450 mono-oxygenase.
Proc. Natl. Acad. Sci. USA **104**, 18463-18468 (2009).
- 38 Landwehr, M., Carbone, M., Otey, C. R., Li, Y. & Arnold, F. H. Diversification of catalytic function in a
synthetic family of chimeric cytochrome P450s. *Chem Biol* **14**, 269-278 (2007).
- 39 Arnold, F. H. Fancy footwork in the sequence space shuffle. *Nat Biotechnol* **24**, 328-330 (2006).
- 40 Raadt, A. d. & Griengl, H. The use of substrate engineering in biohydroxylation. *Current Opinion in
Biotechnology* **13**, 537-542 (2002).
- 41 Ramanavicius, A. & Ramanaviciene, A. Hemoproteins in design of biofuel cells. *Fuel Cells* **1**, 25-36
(2009).
- 42 Montellano, P. R. O. d. & Voss, J. J. D. Oxidizing species in the mechanism of cytochrome P450. *Nat Prod
Rep* **19**, 477-493 (2002).
- 43 Rittle, J. & Green, M. T. Cytochrome P450 Compound I: capture, characterization, and C-H bond
activation kinetics. *Science* **330**, 933-937 (2010).
- 44 Schenkman, J. B., Remmer, H. & Estabrook, R. W. Spectral studies of drug interaction with hepatic
microsomal cytochrome. *Mol Pharmacol* **3**, 113-123 (1967).
- 45 Bernhardt, R. Optimized chimeragenesis: creating diverse P450 functions. *Chem Biol* **11**, 287-288 (2004).
- 46 Cupp-Vickery, J. R., Han, O., Hutchinson, C. R. & Poulos, T. L. Substrate-assisted catalysis in cytochrome
P450eryF. *Nature* **3**, 632-637 (1996).
- 47 Ruettinger, R. T. & Fulco, A. J. Epoxidation of unsaturated fatty acids by a soluble cytochrome P-450-
dependent system from *Bacillus megaterium*. *J Biol Chem* **256**, 5728-5734 (1981).
- 48 Roberts, G. A., Grogan, G., Greter, A., Flitsch, S. L. & Turner, N. J. Identification of a new class of
cytochrome P450 from a *Rhodococcus* sp. *J Bacteriol* **184**, 3898-3908 (202).
- 49 Mot, R. D. & Parret, A. H. A. A novel class of self-sufficient cytochrome P450 monooxygenases in
prokaryotes. *Trends Microbiol.* **10**, 502-508 (2002).
- 50 Munro, A. W., Girvan, H. M. & McLean, K. K. Cytochrome P450-redox partner fusion enzymes.
Biochimica et Biophysica Acta **1770**, 345-359 (2007).
- 51 Daiber, A., Shoun, H. & Ullrich, V. Nitric oxide reductase (P450_{nor}) from *Fusarium oxysporum* *J Inorg
Biochem* **99**, 185-193 (2005).
- 52 Kelly, S. L., Lamb, D. C., Jackson, C. J., Warrilow, A. G. & Kelly, D. E. The biodiversity of microbial
cytochrome P450s. *Adv Microb Physiol* **47**, 131-186 (2003).
- 53 Baltz, R. H. & Seno, E. T. Genetics of *Streptomyces Fradiae* and Tylosin Biosynthesis. *Ann Rev Microbiol*
42, 547-574 (1988).
- 54 Lomovskaya, N. *et al.* Doxorubicin overproduction in *Streptomyces peucetius*: cloning and characterization
of the *dnrU* ketoreductase and *dnrV* genes and the *doxA* cytochrome P-450 hydroxylase gene. *J Bacteriol*
181, 305-318 (1999).
- 55 Mao, Y., Varoglu, M. & Sherman, D. H. Molecular characterization and analysis of the biosynthetic gene
cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* NRRL 2564. *Chem Biol* **6**,
251-263 (1999).
- 56 Weber, J. M. *et al.* Organization of a cluster of erythromycin genes in *Saccharopolyspora erythraea*. *J
Bacteriol* **172**, 2372-2383 (1990).

- 57 Andersen, J. F. & Hutchinson, C. R. Characterization of *Saccharopolyspora erythraea* cytochrome P-450
genes and enzymes, including 6-deoxyerythronolide B hydroxylase. *J Bacteriol* **174**, 725-735 (1992).
- 58 Lambalot, R. H., Cane, D. E., Aparicio, J. J. & Katz, L. Overproduction and characterization of the
erythromycin C-12 hydroxylase, EryK. *Biochemistry* **34**, 1858-1866 (1995).
- 59 Xue, Y., Zhao, L., Liu, H.-W. & Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in
Streptomyces venezuelae: Architecture of metabolic diversity. *Proc. Natl. Acad. Sci. USA* **95**, 12111-12116
(1998).
- 60 Xue, Y., Wilson, D., Zhao, L., Liu, H.-W. & Sherman, D. H. Hydroxylation of macrolactones YC-17 and
narbomycin is mediated by the pikC-encoded cytochrome P450 in *Streptomyces venezuelae*. *Chem. Biol.* **5**,
661-667 (1998).
- 61 Reusser, F. Effect of lincomycin and clindamycin on peptide chain initiation. *Antimicrob Agents
Chemother* **10**, 618-622 (1975).
- 62 Rodriguez, A. M., Olano, C., Mendez, C., Hutchinson, C. R. & Salas, J. A. A cytochrome P450-like gene
possibly involved in oleandomycin biosynthesis by *Streptomyces antibioticus*. *FEBS Microbiol Lett* **127**,
117-120 (1995).
- 63 Caffrey, P., Lynch, S., Flood, E., Finnan, S. & Oliynyk, M. Amphotericin biosynthesis in *Streptomyces
nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem Biol* **8**, 713-723 (2001).
- 64 Brautaset, T. *et al.* Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei*
ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway. *Chem Biol* **7**, 395-
403 (2000).
- 65 Ikeda, H., Nonomiya, T., Usami, M., Ohta, T. & Omura, S. Organization of the biosynthetic gene cluster
for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proc Natl Acad Sci USA*
96, 9509-9514 (1999).
- 66 Motamedi, H. & Shafiee, A. The biosynthetic gene cluster for the macrolactone ring of the
immunosuppressant FK506. *Eur J Biochem* **256**, 528-534 (1998).
- 67 Schwecke, T. *et al.* The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc
Natl Acad Sci USA* **92**, 7839-7843 (1995).
- 68 Brockmann, H. & Henkel, W. Pikromycin, ein bitter schmeckendes antibiotikum aus actinomyceten
(antibiotica aus actinomyceten, VI. mitteil. *Chem Ber* **84**, 284-288 (1951).
- 69 Donin, M. N., Pagano, J., Dutcher, I. D. & McKee, C. M. Methymycin, a new crystalline antibiotic.
Antibiot Annu **1**, 179-185 (1954).
- 70 Perlman, D. & O'Brien, E. Microbiological production of methymycin and related antibiotics. *Antibiot
Chemther* **4**, 894-898 (1954).
- 71 Xue, Y. & Sherman, D. H. Biosynthesis and combinatorial biosynthesis of pikromycin-related macrolides
in *Streptomyces venezuelae* *Metab Eng* **3**, 15-26 (2001).
- 72 Kittendorf, J. D. & Sherman, D. H. The methymycin/pikromycin pathway: a model for metabolic diversity
in natural product biosynthesis. *Bioorganic & Medicinal Chemistry* **17**, 2137-2146 (2009).
- 73 Xue, Y. & Sherman, D. H. Alternative modular polyketide synthase expression controls macrolactone
structure. *Nature* **403**, 571-575 (2000).
- 74 Xue, Y., Wilson, D. & Sherm, D. H. Genetic architecture of the polyketide synthases fir methymycin and
pikromycin series macrolides. *Gene* **245**, 203-211 (2000).
- 75 Szu, P.-h., He, X., Zhao, L. & Liu, H.-w. Biosynthesis of TDP-D-desosamine: identification of a strategy
for C4 deoxygenation. *Angew Chem Int Ed Engl* **44**, 6742-6746 (2005).
- 76 Chen, H. *et al.* Expression, purification, and characterization of two *N,N*-dimethyltransferases, TylM1 and
DesVI, involved in the biosynthesis of mycaminose and desosamine. *Biochemistry* **41**, 9165-9183 (2002).
- 77 Zhao, L., Borisova, S., Yeung, S.-M. & Liu, H.-w. Study of C-4 deoxygenation in the biosynthesis of
desosamine: evidence implicating a novel mechanism. *J Am Chem Soc* **123**, 7909-7910 (2001).
- 78 Borisova, S. A. *et al.* Substrate specificity of the macrolide-glycosylating enzyme pair DesVII/DesVIII:
opportunities, limitations, and mechanistic hypotheses. *Angew Chem Int Ed Engl* **45**, 2748-2753 (1006).
- 79 Thibodeaux, C. J., III, C. E. M. & Liu, H.-w. Natural-product sugar biosynthesis and enzymatic
glycodiversification. *Angew Chem Int Ed Engl* **47**, 9814-9859 (2008).
- 80 Borisova, S. A., Zhao, L., III, C. E. M., Kao, C.-L. & Liu, H.-w. Characterization of the glycosyltransferase
activity of DesVII: analysis of the implications for the biosynthesis of macrolide antibiotics. *J Am Chem
Soc* **126**, 6534-6535 (2004).

- 81 Kao, C.-L., Borisova, S. A., Kim, H. J. & Liu, H.-w. Linear aglycones are the substrates for glycosyltransferase DesVII in methymycin biosynthesis: analysis and implications *J Am Chem Soc* **128**, 5606-5607 (2006).
- 82 Borisova, S. A., Kim, H. J., Pu, X. & Liu, H.-w. Glycosylation of acyclic and cyclic aglycone substrates by macrolide glycosyltransferase DesVII/DesVIII: analysis and implications *ChemBioChem* **9**, 1554-1558 (2008).
- 83 Zhang, Q. & Sherman, D. H. Isolation and structure determination of novamethymycin, a new bioactive metabolite of the methymycin biosynthetic pathway in *Streptomyces venezuelae*. *J Nat Prod* **64**, 1447-1450 (2001).
- 84 Lee, S. K. *et al.* Neopikroycin and novapikromycin from the pikromycin biosynthetic pathway of *Streptomyces venezuelae*. *J Nat Prod* **69**, 847-849 (2006).

Chapter 2

Regioselective C-H Bond Oxidation by an Engineered P450 Monooxygenase Employing Simple Removable Directing Groups

2.1 Introduction

The regio- and stereoselective oxidation of unactivated C-H bonds under mild conditions remains a substantial barrier to the efficient synthesis of complex bioactive molecules.¹⁻⁶ Considerable efforts have focused on both directed and non-directed reactions. In the absence of a covalently installed directing functionality, the intrinsic influences of the substrate, including steric interactions, proximity to electron donating and withdrawing groups, and stereochemical orientation, play the major role in governing regioselectivity of oxidations. In contrast, directed processes are highly effective at overcoming the inherent reactivity of specific C-H bonds by relying on entropic considerations.^{5,7,8} The regioselectivity of these processes is typically governed by selection of a C-H bond through a five- or six-membered transition state involving delivery of a reactive catalyst to a region of the molecule that is proximal to the directing group. The functionalization of regions of the molecular structure that are remote to directing groups, while overriding the inherent reactivity of specific bonds, is not possible by these well-studied approaches. Previous strategies that have begun to address this challenge included the utilization of supramolecular assemblies or specialized directing groups with extended linear functionality to reach remote parts of the substrate.⁹⁻¹⁶ This challenge has long been recognized beginning with the pioneering work by Breslow using specially-designed protecting groups for photochemically mediated oxidations and free radical reactions.^{12,13,16} Since then, enzymes have inspired identification of biological catalysts or simpler mimics that utilize a metal oxo-reactive site. Nevertheless, subsequent development of enzymes as alternative tools as synthetic reagents has

only seen a few examples of success for site-selective in organic chemistry.¹⁷

Biological catalysts provide an attractive opportunity to selectively oxidize chemically similar bonds that are remote from directing influences. In particular, the heme-containing superfamily of cytochrome P450 monooxygenases (P450s) involved in diverse oxidative processes that include xenobiotic catabolism, steroid synthesis, and late-stage biosynthetic tailoring of natural products offer a promising source of powerful biocatalysts given their capacity to perform oxidations in a regio- and stereoselective manner.¹⁹⁻²³ Therefore, we were motivated to explore the potential of bacterial biosynthetic P450 monooxygenases as oxidizing agents by increasing substrate flexibility, and exploiting their relatively ready access through gene cloning, overexpression, and purification.²⁴ Previous efforts have improved limitations through protein engineering, such as scanning chimeragenesis and directed evolution to produce mutant cytochrome P450s with desirable target oxidation activity.²⁵⁻²⁹

Due to its innate flexibility and distinct substrate anchoring mode, PikC promised to be an excellent candidate for biocatalysis.^{30,31} The apparent requirement for the presence of a sugar moiety revealed by co-crystal structures with natural substrates YC-17 and narbomycin inspired a “substrate engineering” strategy to expand the substrate scope of catalysis (**Fig. 2-2**).³²⁻³⁵ Our initial hypothesis grew from observations that particular functional groups promoted exact enzyme-substrate interactions with this “anchor” enabling selective oxidation of modified substrates *in vivo*.^{32,33} The substrate tolerance of PikC was extended to unnatural substrates and employed as a biocatalyst for the moderately regioselective hydroxylation of a number of carbolides (desosamine-linked carbocyclic rings) and linear derivatives, confirming the crucial interaction between the aminosugar and enzyme active site residues.³⁶

Herein we report *in vitro* implementation of substrate engineering for selective C-H bond oxidation using an optimized chimera of the macrolide biosynthetic P450 monooxygenase PikC fused to the exogenous electron donor, PikC_{D50N}-RhFRED.^{31,37} While desosamine proved to be an effective anchoring group for unnatural substrates in the PikC active site, the synthetic challenges in the attachment and removal of glycosidic bonds compromised the generality and utility of this biocatalytic hydroxylation methodology, limiting the substrate scope to molecules containing this sugar moiety.³⁰ For this study, a series of carbocyclic rings were synthetically attached to several simplified linear linkers with a terminal *N,N*-dimethylamino moiety to achieve effective regioselective hydroxylation. This study culminated in a series of compounds

containing the natural macrolactone, 10-deoxymethynolide (10-dml), attached through a linear ester being oxidized in greater than 95% yield with high regioselectivity. Furthermore, substantial modulation of the observed regioselectivity was achieved through manipulation of the anchoring group linker length. Analysis of high-resolution enzyme-substrate co-crystal structures provided valuable insight into the function of the *N,N*-dimethylamino moiety of the linear and aminosugar-derived anchoring groups to control reaction site selectivity through formation of specific interactions between active site residues and the substrate. Finally, the ease of chemical installation and removal of an ester-containing linear linker highlighted the potential for anchoring group incorporation into the synthesis of elaborated macrocycles to create structurally diverse bioactive molecules.

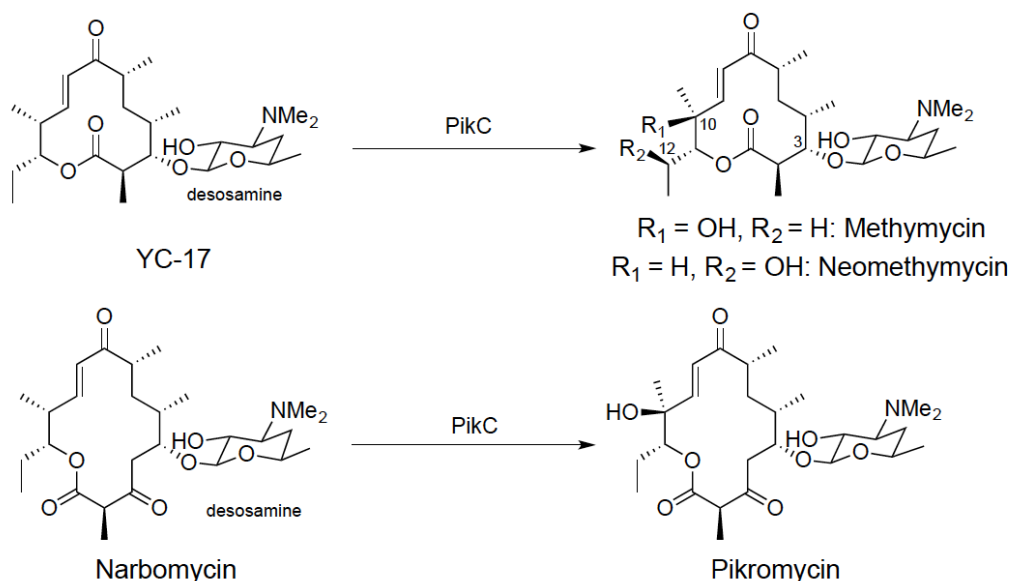


Figure 2-1. Major physiological reactions catalyzed by PikC. Endogenous reactions catalyzed by PikC include the oxidation of YC-17 and narbomycin to yield methymycin, neomethymycin, and pikromycin, respectively.

2.2 Results and Discussion

2.2.1 Regioselective Hydroxylation of Synthetic Cycloalkanes.

PikC is the versatile cytochrome P450 monooxygenase involved in the final tailoring steps for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*.^{31,38} The confirmed endogenous function of this enzyme involves hydroxylation of the 12-membered macrolide YC-17 and the 14-membered macrolide narbomycin to produce methymycin/neomethymycin and pikromycin, respectively (**Fig. 2-1**). Previous analysis of X-ray co-crystal structures of PikC with

its natural substrates revealed largely non-specific hydrophobic interactions between the substrate macrolactone rings and active site residues. However, the aminosugar desosamine, present in both natural macrolide substrates, was located in two separate binding pockets utilizing several hydrogen bonds and ionic interactions for substrate positioning and stabilization in the active site.^{39,40} Specifically, the protonated dimethylamino moiety of desosamine and a glutamate residue in the B/C loop region of the protein participated in an electrostatic salt bridge to achieve specific substrate binding.⁴⁰ In addition, desosamine has already been demonstrated to direct unfunctionalized carbocycles to the active site of PikC for regioselective oxidation, albeit at lower levels of selectivity compared to YC-17 and narbomycin.³⁶ The necessity of harvesting and synthetic installation of desosamine motivated the exploration of alternative, simplified anchoring groups with an ability to mediate similar electrostatic and H-bonding interactions for effective PikC-catalyzed C-H bond activation.^{30,41}

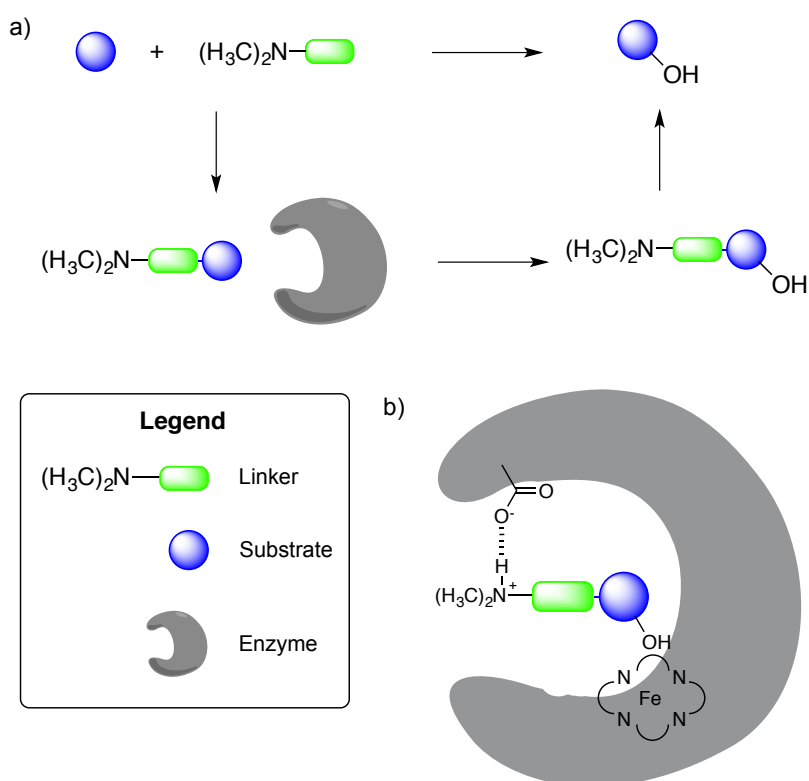


Figure 2-2. Schematic of substrate engineering concept using simplified synthetic linkers. a) Targeted substrates for PikC oxidation were synthetically appended to linkers bearing a terminal *N,N*-dimethylamino to assess directed C-H bond oxidation. Products were then submitted for linker removal to recover oxidized materials. b) The terminal *N,N*-dimethylamino group made crucial contacts with enzyme active site residues bearing a carboxylic acid for substrate oxidation.

To gain insight into the factors involved in productive substrate binding, the activity of PikC_{D50N}-RhFRED was screened against a panel of cycloalkane derivatives containing linear linkers bearing a terminal *N,N*-dimethylamino group (**Table 2-1**). A range of substrate conversions and varying degrees of regioselectivity were observed using ether- and ester-based linkers. Compared to the desosamine-fused cycloalkane (compound **1**), we observed similar product yields that were affected through linker length (compounds **2** vs. **4**), cycloalkane size (compounds **3** vs. **4**) and the presence of a linker ester group, which notably increased yield while simultaneously decreased the number of products (compounds **3** vs. **5** and **4** vs. **6**). Whereas seven products were observed in oxidations of **1**, comparisons must consider that diastereotopic protons in **1** (due to desosamine chirality) become enantiotopic when achiral linkers are employed. Decreased conversion and regioselectivity of oxidations of the simple cycloalkane-derived substrates compared to the macrolactone system was likely attributed to suboptimal substrate binding affinity in the PikC active site due to a lack of functionality of the carbocyclic rings. The decreased chemical complexity potentially caused (i) an entropic penalty associated with the flexibility of the saturated ring systems; (ii) a lack of hydrophobic interactions between functional groups in the substrate macrolactone ring and PikC active site residues; and (iii) a loss of interactions between PikC active site and functional groups of macrolactone ring or desosamine (while retaining the salt-bridge between E94 and the dimethylamino group).

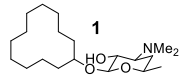
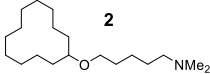
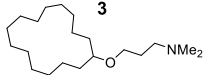
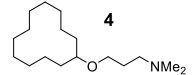
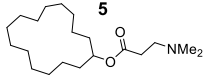
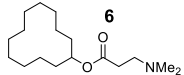
Compound	No. of products	Product Ratio	Total yield (%)
	7	N.D.	47
	5	6:10:25:9:1	50
	6	4:3:5:6:3:1	30
	5	1:1.3:3*	30
	5	3:1:6:2:1	35
	4	4:6:5:1	34

Table 2-1. Panel of initial ether and ester linker compounds screened against $\text{PikC}_{\text{D50N}}$ -RhFRED. The number of diastereomeric products, product ratios, and total conversion were determined by LC-MS Q-TOF. (*Note: Not all peaks in **Fig. S2-2** corresponding to products for substrate **4** were completely separable on the LC-MS column; therefore, product ratio offered only an approximation of selectivity.)

Binding orientation of cycloalkane derivative **6** within the PikC active site was assessed using authentic standards bearing a hydroxyl group at C6, C7, and C8 (**Fig. 2-3**). These compounds were synthesized as diastereomeric pairs with the cis/trans configurations of the ester and hydroxyl substituents. Identities for three of the four PikC -hydroxylated products were correlated using the synthetic standards, while the fourth product could not be determined because one of the C6/C8-hydroxylated diastereomers co-eluted with one of the C7-hydroxylated diastereomers (**Fig. S2-2**).

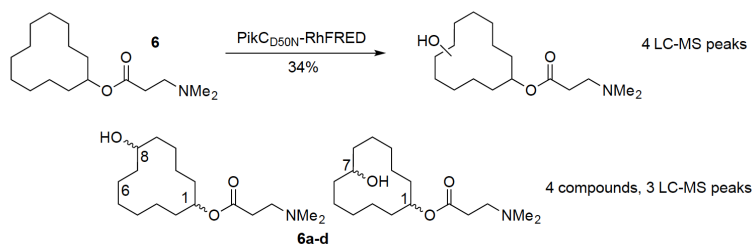
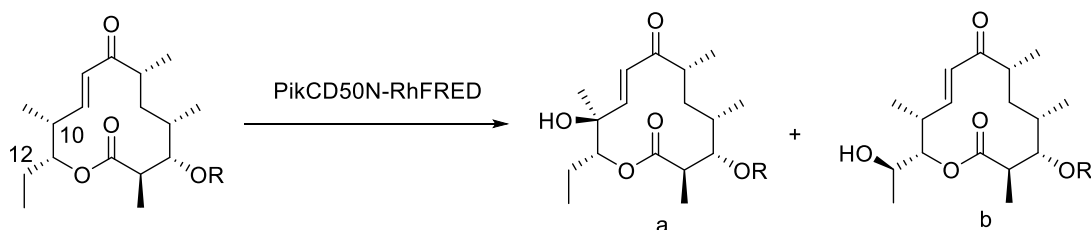


Figure 2-3. Comparison of synthetic diastereomeric C6, C7, and C8 standards (6a-d) to the observed $\text{PikC}_{\text{D50N}}$ -RhFRED oxidation products of compound 6. Initial analysis for product distribution and yield were determined by LC-MS Q-TOF. Product identities were subsequently confirmed by use of authentic standards.

2.2.2 Altered Regioselectivity of 10-dml C-H Functionalization with Linear Desosamine Mimics

Our studies on the carbocyclic substrates above demonstrated that the ester functionality offered an attractive alternative to labor-intensive desosamine attachment by providing a chemically labile anchoring group. Thus, we pursued ester-containing linkers to probe the effect of an additional hydrogen-bond acceptor on the regioselectivity of targeted C-H bond oxidation in the macrolactone system. To evaluate the impact of this functionality, 10-dml, the macrolactone core of YC-17 with a C-3 OH, was esterified by a series of linear anchors (**Table 2-2**). In these reactions, 5 μM of enzyme was incubated overnight with NADPH and the glucose-6-phosphate/glucose-6-phosphate dehydrogenase NADPH regeneration system.^{42,43} The terminal *N,N*-dimethylamino group was essential for catalytic conversion by $\text{PikC}_{\text{D50N}}$ -RhFRED as derivatives lacking this moiety were catalytically incompetent. Thus, 10-dml bearing the C-3 OH failed to generate products under identical reaction conditions (data not shown). Furthermore, 10-dml attached to a linker bearing a terminal isopropyl group in place of dimethylamino (compounds **10** and **11**) was not accepted as a substrate for $\text{PikC}_{\text{D50N}}$ -RhFRED, confirming the indispensability of the salt-bridge interaction in the enzyme active site for efficient regioselective hydroxylation (**Table 2-2**).³⁹



Compound	R	No. of Prods.	Product ratio	Total yield	Products
YC-17		2	1:1.2	>99	Methymycin, Neomethymycin
7		2	1:1.3	95	7a, 7b
8		2	1:2	98	8a, 8b
9		2	2:1	68	9a, 9b
10		0	-	0	-
11		0	-	0	-
12		0	-	0	-
13		0	-	0	-

Table 2-2. 10-dml analogs with desosaminyl mimics screened against PikC_{D50N}-RhFRED.

Activity of 3-(dimethylamino)propanoate (DMAp) derivative, compound **8**, was comparable to YC-17 by total yield with altered C-10/C-12 hydroxylation product distribution profile: 1:2 *versus* 1:1.2 for YC-17 (**Table 2-2**).⁴⁴ The products were determined to be methynolide and neomethynolide analogs **8a** and **8b**, with the hydroxylation occurring at the tertiary allylic position (C-10), and favored at the secondary exocyclic methylene (C-12) of the ethyl side-chain, respectively. Shortening the linker by one methylene unit in the *N,N*-dimethylamino ethanoate (DMAE) derivative, compound **7**, resulted in a 95% yield of the combined C-10/C-12 hydroxylated products **7a** and **7b** in a ratio of 1:1.3, slightly favoring the C-12 hydroxylated analog. Extending the linker by one methylene unit in *N,N*-dimethylamino butanoate (DMAb) derivative, compound **9**, resulted in 68% overall conversion by the engineered PikC biocatalyst, producing the C-10/C-12 analogs with a product ratio of 2:1,

favoring the allylic tertiary carbon atom (**9a**). These results demonstrated that linker length of the linear anchoring group mediated regioselective oxidation at distal C-H bonds through direct substrate engineering.

In a survey of the various forms of the PikC biocatalyst, we observed that PikC_{WT}, PikC_{WT}-RhFRED, and PikC_{D50N}-RhFRED generated similar product profiles with slightly different ratios, while PikC_{D50N}-RhFRED provided a substantially higher overall efficiency (**Fig. S2-3**). Furthermore, substrate-binding assays of compounds **7** – **9** with the different forms of the biocatalyst revealed comparable dissociation constants (**Table S2-1**). These results confirmed that neither the D50N mutant form of PikC nor the RhFRED fusion had a significant impact on substrate orientation for compounds **7** – **9**.

2.2.3 Product Analysis

Analysis of the PikC reactions was based on the ability to correlate hydrolysis products of **7a,b**, **8a,b** and **9a,b** to methynolide and neomethynolide. These two macrolactones were employed in this study as authentic standards to determine sites of hydroxylation, and have been previously described as products of fermentation by mutant strains of *S. venezuelae*, through chemical degradation of methymycin and neomethymycin, and by total synthesis.⁴⁵⁻⁵¹ Therefore, a sequential scheme involving substrate synthesis (**7-13**), enzymatic oxidation, ester hydrolysis, and LCMS analysis was followed for initial assessment of reaction products (**Fig. S2-1**), followed by enzymatic scale-up and isolation for rigorous structure elucidation and correlation (**Table 2, Fig. 2-4**). Preparative-scale PikC-mediated reactions afforded a sufficient quantity of the oxidation products for NMR analysis after hydrolysis, and the spectra were directly compared to the authentic standards of methynolide and neomethynolide. Thus, hydrolysis of the ester linker in **8** demonstrated that the major product was neomethynolide and the minor compound was methynolide (**Fig. 2-4**). The ¹H-NMR spectra stacked plot (**Fig. 4**, trace E) showed signals corresponding to alkene protons (C-8/C-9), while the characteristic doublet of doublets from C-9 disappeared following PikC catalyzed oxidation at the C-10 tertiary allylic position observed in the methynolide analog. In the neomethynolide analog, the C-9 signal remained as a doublet of doublets. The C-8 signal corresponded to the C-H adjacent to the C-7 carbonyl group and was split into a doublet in both methynolide and neomethynolide. Analogous hydrolysis experiments were performed for PikC oxidations of compounds **7** and **9** as

well. These signals were present both in the $^1\text{H-NMR}$ spectra of hydrolyzed PikC reaction products from substrates **7**, **8** and **9**. Therefore, the $^1\text{H-NMR}$ confirmed the sites of oxidation to be at C-10 and C-12 and that variations in regioselectivity observed in **7a,b**, **8a,b** and **9a,b** result from changes in substrate orientation within the active site of PikC conveyed through linker length.

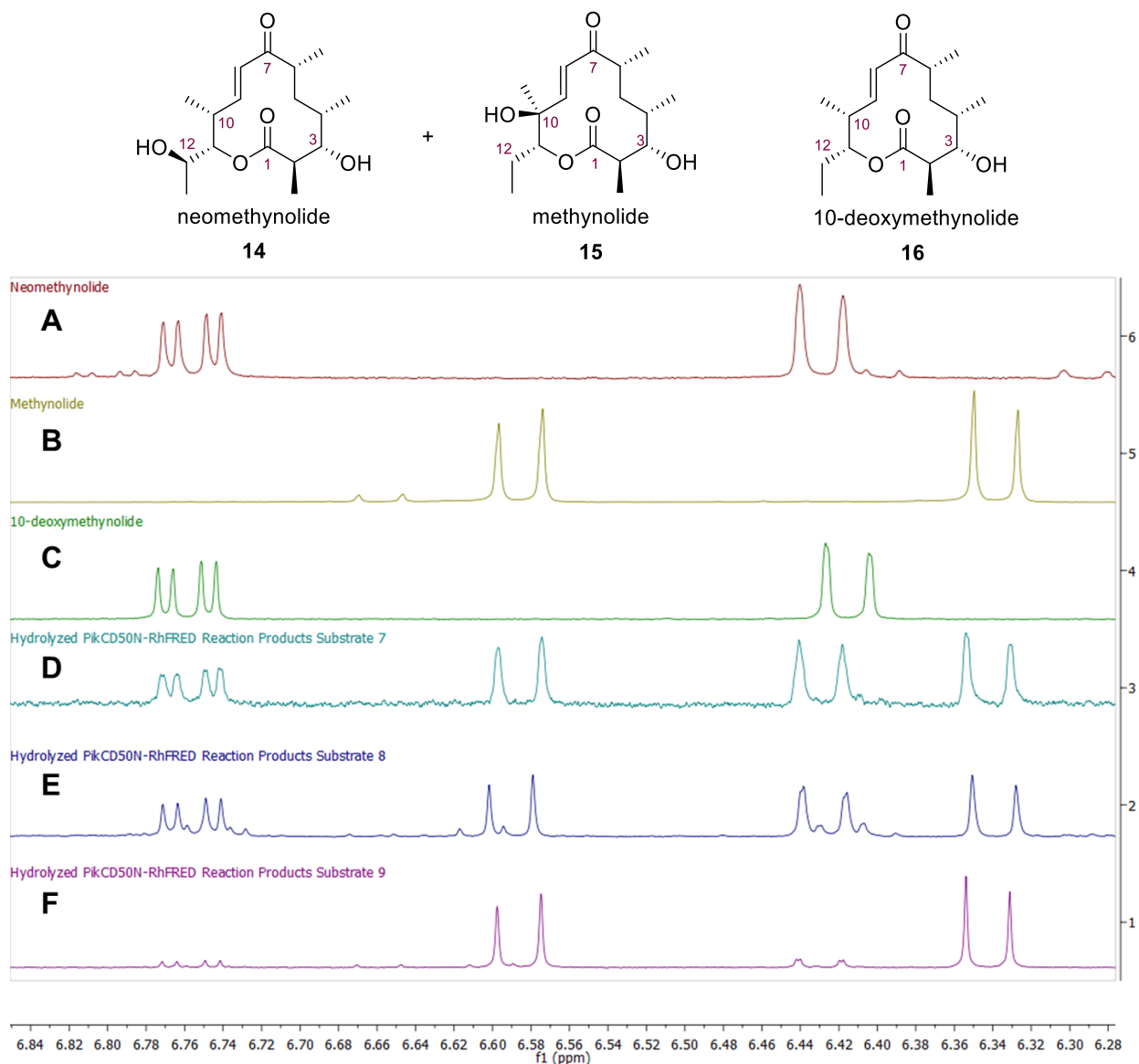


Figure 2-4. Stacked plot of $^1\text{H-NMR}$ of hydrolyzed reaction products from substrates 7-9. From top to bottom (A-F): (A) Neomethynolide, (B) Methynolide, (C) 10-deoxymethynolide, (D) Hydrolyzed Pik_{D50N}-RhFRED reaction products from substrate **7**, (E) Hydrolyzed Pik_{D50N}-RhFRED reaction products from substrate **8**, (F) Hydrolyzed Pik_{D50N}-RhFRED reaction products from substrate **9**. From left to right: C₉-H signal, C₈-H signal.

To gain further understanding into binding affinity and conversion of new 10-dml analogs bearing the DMAE, DMAp and DMAb synthetic linkers, we established the steady-state kinetic parameters of the oxidation of compounds **7** – **9** using PikC_{D50N}-RhFRED (**Table 2-3**). Although the single-component enzyme system demonstrated good to excellent overall product yields and high regioselectivity for compounds **7** – **9**, the kinetic data indicated that PikC_{D50N}-RhFRED catalyzed hydroxylation of the semisynthetic compounds less efficiently compared to native macrolide substrate YC-17. Consistent with the reduced levels of oxidation of **9** by PikC_{D50N}-RhFRED, both Michaelis constant (K_m) and enzyme efficiency (k_{cat}) were significantly attenuated toward this substrate.

Compound	Product Ratio 15:14	Kinetic Parameters		
		K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\mu\text{M}^{-1}\text{min}^{-1}$)
YC-17	1:1.2	33.3 ± 8.20	247.8 ± 20.4	7.5 ± 2.5
7	1:1.3	138.7 ± 50.5	17.9 ± 3.2	0.129 ± 0.063
8	1:2	245.2 ± 50.5	25.3 ± 3.3	0.103 ± 0.065
9	2:1	330.9 ± 124.8	27.6 ± 6.9	0.0834 ± 0.0554

Table 2-3. Linker induced regioselectivity differences and Michaelis-Menten fitted steady-state kinetic parameters.

2.2.4 Structural Basis for Regioselectivity of 12-Membered Ring Aglycones with Simplified Linker Anchors.

To establish the substrate binding orientation within the active site, we determined the x-ray structures of catalytically competent compounds **2**, **7** and **8**, bound within the active site of PikC_{WT} and/or PikC_{D50N} (**Table S2-4**). The crystallization propensity of the enzyme-substrate complexes often directly correlated with their catalytic potential as less active complexes failed to generate suitable quality crystals. The only structurally characterized carbocyclic derivative, compound **2**, was represented by a 2.7 Å structure (PDB ID 4BF4) containing 16 molecules in the asymmetric unit with two monomers possessing interpretable electron density in the active site to allow fitting of the substrate (**Fig. 2-5**). Notably, a fragment of electron density was

present adjacent to E94 virtually in each monomer suggesting that the salt-bridge contact between the *N,N*-dimethylamino group of the linker and E94 was preserved in the complex. A fragment of the cyclododecane ring proximal to the heme Fe was also visible, whereas the remainder of the substrate molecule was considerably disordered. The orientation of **2** in the active site was consistent with hydroxylation of C-6, C-7 and C-8, and in agreement with the limited regioselectivity of this compound due to the inherent flexibility of the non-functionalized ring system.

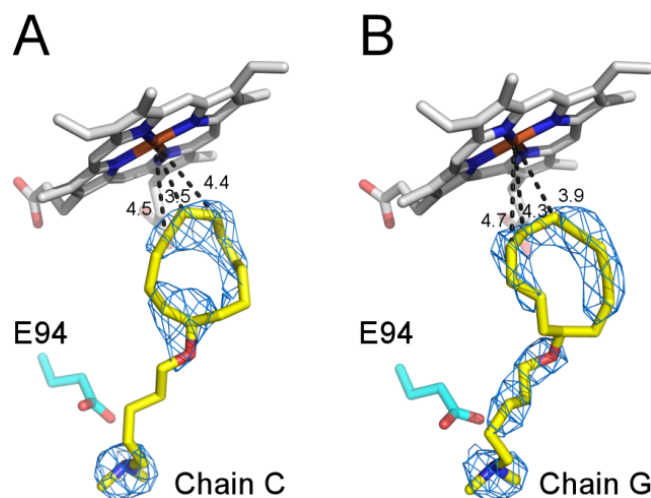


Figure 2-5. Cycloalkane derivative in the *PikC*_{D50N} active site. Fragments of the 2F_o-F_c electron density map (blue mesh) countoured at 1.0 σ delineate position of the cyclododecane substrate mimic (yellow) in the *PikC*_{D50N} (PDB ID 4BF4). Fragment of electron density next to E94 was present in the majority of the 16 monomers constituting the asymmetric unit, providing strong evidence for a dominant role of the salt-bridge interaction between E94 and the terminal dimethylamino group of the linear anchor. The heme cofactor is shown in gray sticks. Distances to the heme Fe are in Angstroms.

Unlike the carbocyclic compound **2**, the macrocyclic derivatives **7** and **8** contained well-defined electron density in the corresponding high-resolution x-ray structures (**Fig. 2-6A-C**). In addition to the preserved salt-bridge, a novel hydrogen-bonding interaction between H238 and the ester carbonyl group of the linker emerged in the active site. This engineered H-bond may account for relocating the macrolactone ring towards the heme by ~ 1.4 Å as compared to YC-17 (**Fig. 2-6D**). In the *PikC*_{WT}, both the C-10 and C-12 sites were significantly closer in **8** than in the native substrate YC-17.³⁹ The distance between C-12 and the heme iron varied between 4.0 and 5.3 Å for synthetic mimic **8** and YC-17, respectively, while the C-10 site varied between 5.1 and 7.4 Å. This differences were consistent with the observed preference for oxidation of C-12 over the allylic carbon (C-10). Furthermore, the allylic site was captured in a closer proximity to

the heme Fe than has been observed previously, increasing its accessibility for hydroxylation by the oxo-ferryl intermediate, Compound I.²¹

In the PikC_{D50N} mutant, the macrolactone core of **8** was located further away from the heme Fe (**Fig. 2-6E**), at distances comparable with YC-17 in the wild type enzyme. However, the salt-bridge contact between E94 and the tertiary *N,N*-dimethylamino group was maintained through adjusting the positioning of the linker moiety. In addition, this ionic interaction persisted after shortening the linker by one methylene unit in **7** (**Fig. 2-6F**), confirming the critical role of this interaction in the PikC active site. The H-bond between H238 and the linker carbonyl group of **7** was also preserved at the expense of straining H-bonding geometry and pulling the macrolactone ring away from the heme cofactor by 1.3 Å. These differences in substrate binding within the active sites of both PikC_{WT} and PikC_{D50N} were consistent with the yield and regioselectivity observed in the *in vitro* enzymatic reactions for both **7** and **8**. Based on kinetic (**Table 3**) and structural data (**Fig. 2-6**), perturbation of the substrate structure resulting in repositioning of reactive sites by ~1.5 Å with respect to the heme Fe led to catalytically competent and regioselective complexes, but attenuated both K_m and k_{cat} values by up to an order of magnitude compared to the natural macrolide substrates.

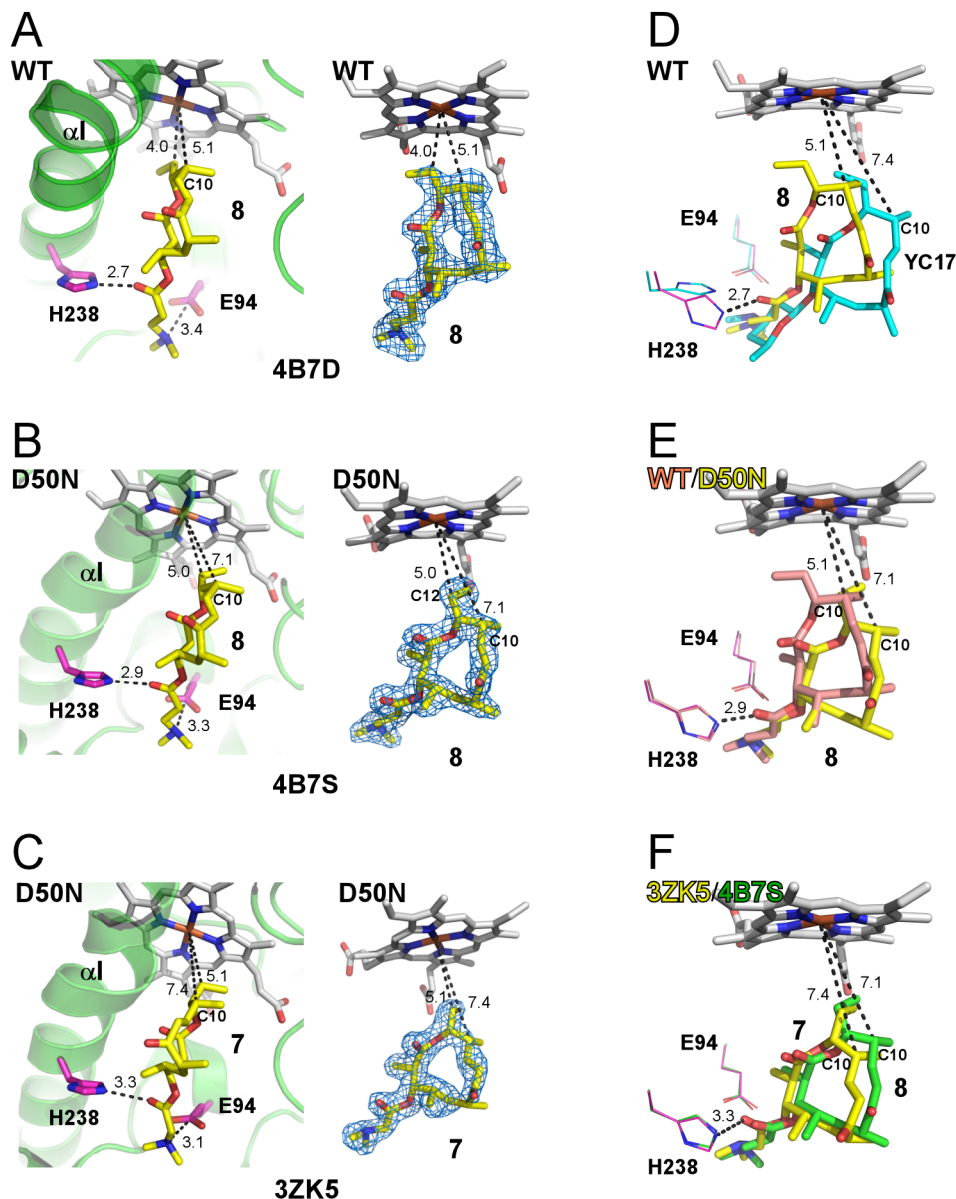


Figure 2-6. Structural insights into regioselectivity of macrocycle hydroxylation. A-C. Binding of compounds **7** and **8** (yellow sticks) in the PikC active site. $2F_o-F_c$ electron density map (blue mesh) is contoured a 1.0σ . Amino acid residues are in pink sticks, fragments of protein structure are shown as green ribbons, distances are in Angstroms. D-F. Superimposition of substrates in the active sites that emphasize differences between (D) YC-17 (cyan) and compound **8** (yellow), (E) compound **8** in the wild type (pink) and $\text{PikC}_{\text{D50N}}$ mutant (yellow), and (F) compounds **8** (yellow) and **7** (green) in $\text{PikC}_{\text{D50N}}$.

The D50N substitution was empirically identified previously to enhance PikC catalysis perhaps by positively affecting the substrate binding kinetics.^{36,37} The difference in the Y295 interactions observed in the wild-type (Fig. 2-7A) and the $\text{PikC}_{\text{D50N}}$ mutant, both co-crystallized with compound **8**, suggested a destabilizing role for the D50N mutation (Fig. 2-7B). Conformational variability of the $\beta 1$ -strand between the wild type and mutant may be directly

affected by this mutation as the interactions of D50 with Y295 were lost in the mutant. Perturbation in the β 1-interacting FG-loop was likely a secondary effect associated with the mutation. These perturbations may have destabilized the substrate access channel and facilitated substrate progression to a catalytically productive binding mode.⁴⁰

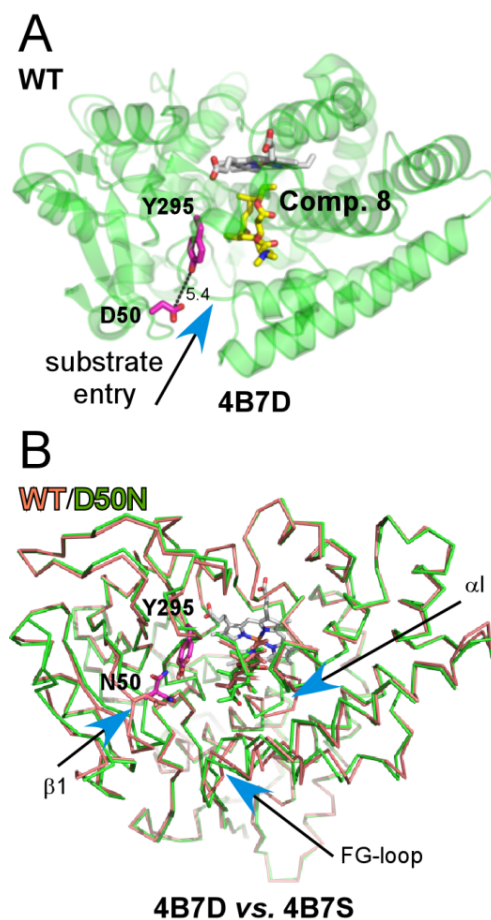


Figure 2-7. Effect of the D50N mutation on the overall PikC structure. **A.** Salt-bridge contact between D50 and Y295 stabilized the substrate entry channel in PikC (green ribbon). Compound **8** is in yellow sticks. **B.** Disruption of the D50-Y295 interaction by site-directed mutagenesis perturbed the channel structure, favorably affecting substrate progression deeper into the active site.³⁸ Wild type C α carbons are traced in pink, PikC_{D50N} C α carbons are traced in green. Conformational variability in the I-helix was due to variations of the G244 torsion angles causing violations in the α -helix H-bonding pattern.

2.2.5 Antibacterial Activities of Synthetic Compounds Containing Desosaminyl Mimics.

Macrolide antibiotics rely heavily on the presence of a deoxysugar to make several crucial ribonucleotide interactions in the peptidyl-transferase center of the 23S ribosomal RNA subunit (the major drug target of macrolides).⁵² Therefore, we sought to determine whether any of substrates with a linear desosamine anchor replacement would maintain measurable

bioactivity (**Table S2-3**). Of particular interest was the observed bioactivity of compounds 3 and 5 against a variety of relevant human pathogenic bacterial strains compared to the activity of YC-17 and erythromycin (**Table S2-3**). We surmised that the bioactivity was derived from the accumulation of the cyclic hydrocarbons within the cellular membranes affecting structural and functional properties that may have included inhibition of primary ion pumps and increasing proton permeability.⁵³ However, minimal inhibitory concentration (MIC) assays of the other synthetic compounds revealed that most lacked activity, which was likely due to the absence of a functional group such as desosamine, capable of mediating ribosomal interactions resulting in antibiotic activity.⁵² It was also conceivable that the general lack of biological activity results from the abundant presence of microbial esterases that cleave the linker.

2.3 Conclusion

In the current study, the inherent substrate flexibility of PikC was exploited to expand substrate scope using desosamine mimics as anchoring groups. In contrast to typical synthetic C-H bond activation strategies that leverage proximal groups or differing C-H bond strengths, our substrate-engineering approach involved simplifying the chemical functionality of the natural substrate to produce a removable linear anchor with a terminal *N,N*-dimethylamino group to achieve regioselective oxidation of remote sites on the target molecule. By this approach, regions of substrates were oxidized distal to the *N,N*-dimethylamino functionality, rendering the regioselectivity patterns orthogonal to that which can be achieved by either directed or non-directed small molecule oxidation catalysts.^{1,4-6,8-11} Replacement of desosamine in semi-synthetic substrates with a simple and removable group eliminates a number of labor-intensive steps for harvesting and synthetic attachment of desosamine, and enables facile recovery of the oxidized product(s) for additional chemical or enzymatic elaboration.

In this work, an ester linker attached to the 12-membered ring macrolactone 10-dml achieved high regioselectivity and substrate conversion by an engineered chimeric mutant form of PikC. Linker length variation produced meaningful changes in regioselectivity of C-H bond oxidation, substantiating the substrate-engineering strategy as an effective means of expanding the substrate scope of this P450. The alteration of regioselectivity and product yield achieved with semi-synthetic substrates contrasted significantly with a drop (up to 10-fold) in k_{cat} and increase in K_m compared to the natural substrate YC-17. Based on the x-ray structure analysis,

substrate orientation within the PikC active site was largely controlled by the electrostatic interaction of the *N,N*-dimethylamino moiety with the carboxylate group of E94. A novel H-bond between the ester carbonyl group of the linker and H238 compensated only partially for the loss of desosamine interactions in semi-synthetic substrates. Although we observed only poor to modest yields with some of the initial semisynthetic unnatural substrates using PikC, compounds like **6** offered significant insight into the rules governing substrate binding through its oxidation pattern and compound **2** through x-ray structural studies.

There is a long-standing demand for enzymes with improved activity and substrate tolerance that can be applied toward chemoenzymatic synthesis, and the development of chemical probes, reagents and novel therapeutics. Bacterial biosynthetic P450s offer a unique source of potential biocatalysts given their propensity for expeditious cloning, over-expression in *E. coli*, and rapid purification. The substrate-engineering approach offers a foothold towards gaining valuable insight into the rules that govern productive substrate binding and conversion of potential P450 substrates, and also towards developing a general method for regio- and stereoselective oxidation of highly complex molecules, including secondary metabolites. This is possible because the products of the enzymatic reactions are readily recoverable and further chemically or enzymatically elaborated towards the development of novel compounds. Finally, these results underscore the potential of bacterial biosynthetic P450s, such as PikC, to be developed as biocatalysts for the selective oxidation of C-H bonds for diversification of natural products and creation of a broad range of bioactive synthetic molecules. The catalytic conversion rate, reaction specificity, and yield could be further maximized by directed evolution and rational or random protein mutagenesis methods.

2.4 Experimental Procedures

Expression and purification of PikC_{D50N}-RhFRED

Protein expression and purification followed a previously described procedure.³⁷

PikC_{D50N}-RhFRED Enzymatic Assay

The standard assay contained 5 μ M PikC_{D50N}-RhFRED, 0.5 mM substrate, 2.5 mM NADPH, 0.25 units of glucose-6-phosphate dehydrogenase, and 5 mM glucose-6-phosphate for NADPH regeneration in 100 μ L of reaction buffer (50 mM NaH₂PO₄, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol). The reaction was carried out at 30°C overnight and quenched by extraction using 3 \times 200 μ L of CHCl₃. The organic extracts were combined, dried under N₂, and redissolved in 100 μ L of acetonitrile. The subsequent liquid chromatography mass spectrometry (LC-MS) analysis was performed on an Agilent Q-TOF HPLC-MS (Department of Chemistry, University of Michigan) equipped with an high resolution electrospray mass spectrometry (ESI-MS) source and a Beckmann Coulter reverse-phase HPLC system using an Waters XBridge C18 3.5 μ m, 2.1 \times 150 mm under the following conditions: mobile phase (A = deionized water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid), 20% B for 3 min, 20 to 100% B over 25 min, 100% B for 5 min, 100 to 20% B over 1 min, 20% B for 10 min; flow rate, 0.21 mL/min. The substrate-binding assays were performed as previously described.³⁷

Steady-state kinetic assays

The standard assays contained 0.1 μ M PikC_{D50N}-RhFRED and various concentrations of substrates (20-250 μ M) in a total volume of 195 μ L of P450 desalting buffer (50 mM NaH₂PO₄, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, 10% glycerol). After pre-incubation at 30°C for 5 min, the reactions were initiated by adding 5 μ L of 50mM NADPH. Reaction termination and analysis was carried out as follows: three aliquots (50 μ L) were taken at three time points (0, 20, 40s for concentrations below 100 μ M or 0, 30, 60s for concentrations above 100 μ M) within the linear range to thoroughly mix with 100 μ L methanol. The protein was pelleted by centrifugation at 13,000g at 4°C for 15 min. The supernatant was subjected to HPLC analysis to monitor substrate consumption within the linear range to measure the initial velocity of the reactions. The HPLC conditions were: XBridge C18 3.5 μ m, 2.1 \times 150 mm under the following conditions: mobile phase (A = deionized water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid),

10% B for 2 min, 10 to 100% B over 10 min, 100% B for 1 min, 100 to 10% B over 2 min; flow rate, 0.21 mL/min. The initial velocity of substrate consumption was deduced from decreased area under the curve (AUC) of specific substrate peaks. All measurements were performed in duplicate, and velocities determined under different substrate concentrations were fit to the Michaelis-Menten equation to calculate kinetic parameters

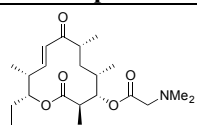
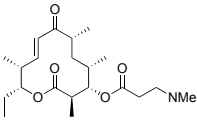
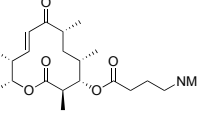
2.5 Supplementary Information

Hydroxylation product analysis and identification

General Hydrolysis Procedure: The crude PikC_{D50N}-RhFRED reaction mixture was dissolved in a 4:1 mixture of MeOH:CHCl₃ (0.015 M) and 2.5 equiv of potassium carbonate were added. The reaction mixture was allowed to stir at rt for 72 hours or until all starting material was consumed. The reaction mixture was concentrated under reduced pressure, quenched by addition of saturated ammonium chloride and extracted thrice with portions of EtOAc. The combined organic layers were dried over Na₂SO₄, and filtered through a plug of silica gel and concentrated under reduced pressure. The products were analyzed by ¹H-NMR and determined to be a mixture of methynolide and neomethynolide (comparison to previously reported spectra), the ratio of the two components depended on the length of the ester anchoring group in the substrate.

Crystallization, data collection and structure determination

Prior to crystallization, the PikC and PikC_{D50N} protein stocks solutions stored at -80°C were diluted to 0.2 mM in 10 mM Tris-HCl, pH 7.5 buffer supplemented with 1 mM compound of interest. Crystallization conditions in each case were determined using commercial high-throughput screening kits available in deep-well format (Hampton Research), a nanoliter drop-setting Mosquito robot (TTP LabTech) operating with 96-well plates, and a hanging drop crystallization protocol. Crystals were further optimized in 96-well or 24-well plates for diffraction data collection. Prior to data collection, all crystals were cryo-protected by plunging them into a drop of reservoir solution supplemented with 20-25% ethylene glycol or glycerol, then flash frozen in liquid nitrogen. Diffraction data were collected at 100-110 K at beamline 8.3.1, Advanced Light Source, Lawrence Berkeley National Laboratory, USA. Data indexing, integration, and scaling were conducted using MOSFLM and the programs implemented in the ELVES software suite.⁵⁴ The crystal structures were initially determined by molecular replacement using the structures of PikC (PDB ID 2BVJ) as search models. The final structures were built using COOT and refined using REFMAC5 (Collaborative Computational Project, 1994; software.⁵⁵⁻⁵⁷

Compound	Enzyme	Dissociation constant (μM)
 7	PikC _{WT}	130
	PikC _{WT} -RhFRED	- ^a
	PikC _{D50N} -RhFRED	135
 8	PikC _{WT}	179
	PikC _{WT} -RhFRED	322
	PikC _{D50N} -RhFRED	273
 9	PikC _{WT}	475
	PikC _{WT} -RhFRED	342
	PikC _{D50N} -RhFRED	87

^a Binding curve could not be fitted.

Table S2-1. Binding affinities of PikC_{WT}, PikC_{WT}-RhFRED, and PikC_{D50N}-RhFRED for compounds 7 – 9 were deduced from a low to high spin iron spectral shift, known as type I binding.⁵⁸

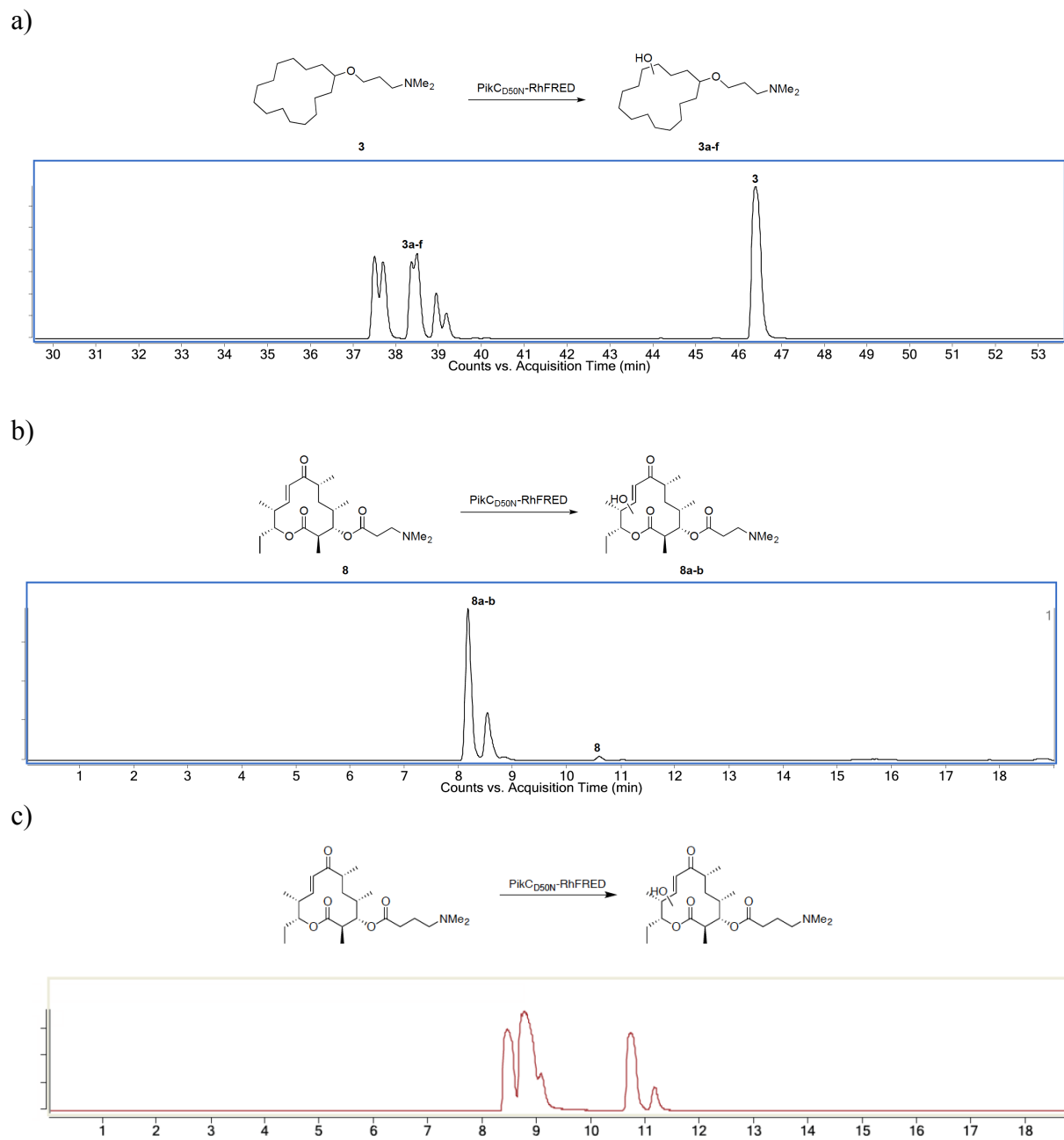


Figure S2-1. Representative LC-MS Q-TOF analyses of PikC_{D50N}-RhFRED oxidations of (a) compound 3 (b) compound 8 (c) compound 9.

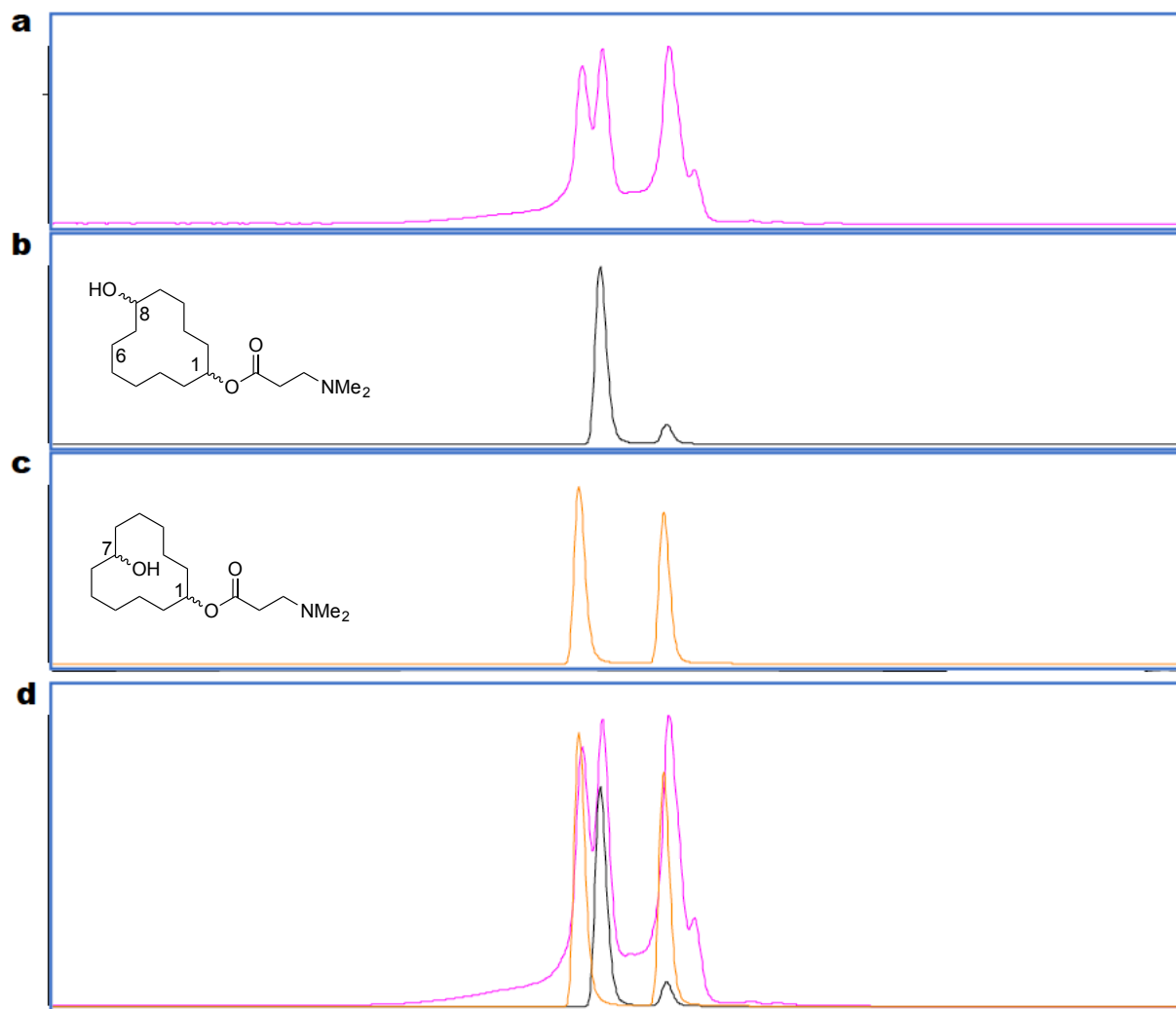


Figure S2-2. Structural determination of mono-hydroxylated products of 8 through LC-MS comparison of synthetic authentic standards regarding retention times. (A) Product profile of $\text{PikC}_{\text{D50N}}$ -RhFRED reaction using **8** as substrate; (B) Authentic standard containing a pair of C6/C8 hydroxylated diastereomers; (C) Authentic standard containing the pair of C7 hydroxylated diastereomers; (D) Overlay of product profile from enzymatic reaction and authentic synthetic standards.

Compound	Enzyme	Product Ratio	% Yield
7	PikC _{WT}	1:1.4	89
	PikC _{WT} -RhFRED	1:1.3	92
	PikC _{D50N} -RhFRED	1:1.3	95
8	PikC _{WT}	1:1.7	59
	PikC _{WT} -RhFRED	1:1.7	75
	PikC _{D50N} -RhFRED	1:2.0	98
9	PikC _{WT}	1.8:1	58
	PikC _{WT} -RhFRED	1.9:1	64
	PikC _{D50N} -RhFRED	2.0:1	68

Table S2-2. Regioselectivity and product yield differences exhibited by PikC_{WT}, PikC_{WT}-RhFRED, and PikC_{D50N}-RhFRED against compounds 7 – 9. Product ratios were determined by LC-MS.

	<i>Deinococcus radiodurans</i> ^a	<i>Bacillus subtilis</i> DHS5333	<i>Staphylococcus aureus</i> ATCC6538P	<i>Kocuria rhizophila</i> ATCC6538P ^b	<i>E. coli</i> TolC ^c	<i>Staphylococcus aureus</i> NorA ^d	<i>Acinetobacter baumannii</i>	Multi-drug resistance <i>Staphylococcus aureus</i>
Media	Special media ^e	LB	LB	Nutrient broth	Mueller Hinten	Mueller Hinten	Mueller Hinten	Mueller Hinten
Compound								
DMSO	-	-	-	-	-	-	-	-
Erythromycin	6.2	<0.8	<0.8	<0.8	<0.8	<0.8	25	>100
YC-17	400	100	100	100	100	100	>400	>400
2	200	400	400	>400	200	100	>400	>400
3	100	50	25	>400	25	12.5	400	50
4	200	400	400	400	400	400	>400	>400
5	12.5	25	>400	25	25	400	>400	400
6	100	400	400	200	100	>400	>400	>400
7	>400	>400	>400	>400	>400	>400	>400	>400
8	400	>400	>400	>400	400	>400	>400	>400
9	400	>400	>400	>400	400	>400	>400	>400
10	>400	>400	>400	>400	>400	>400	>400	>400
11	>400	>400	>400	>400	>400	>400	>400	>400
12	>400	>400	>400	>400	>400	>400	>400	>400

^a A gift from Professor Ada E. Yonath (Structural Biology Department, Weizmann Institute of Science, Rehovot, Israel)

^b Previously known as *Micrococcus luteus* ATCC 9341, which is sensitive to macrolide antibiotics.

^c *E. coli* W3110 TolC disruption mutant is more sensitive to antibiotics.

^d *S. aureus* 8325 NorA disruption mutant is more sensitive to antibiotics.

^e Recipe: 10 g caseine peptide (tryptic digest), 5 g yeast extract, 5 g NaCl, and 5 g glucose (add following sterilization) in 1 L water at pH 7.2.

* All MIC values are in μ M.

Table S2-3. Antibacterial activities of synthetic compounds containing desosaminyl mimics.

2.6 References:

- 1 Chen, M. S. & White, M. C. A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis. *Science* **318**, 783-787 (2007).
- 2 Gutekunst, W. & Baran, P. S. C-H functionalization logic in total synthesis. *Chem. Soc. Rev.* **40**, 1976-1991 (2011).
- 3 Newhouse, T. & Baran, P. S. If C-H bonds could talk: selective C-H bond oxidation. *Angew. Chem. Int. Ed. Engl.* **50**, 3362-3374 (2011).
- 4 Chen, M. S. & White, M. C. Combined effects on selectivity in Fe-catalyzed methylene oxidation. *Science* **327**, 566-571 (2010).
- 5 Lyons, T. W. & Sanford, M. S. Palladium-catalyzed ligand-directed C-H functionalization reactions. *Chem. Rev.* **110**, 1147-1169 (2010).
- 6 Wang, D. H., Engle, K. M., Shi, B. F. & Yu, J. Q. Ligand-enabled reactivity and selectivity in a synthetically versatile aryl C-H olefination. *Science* **327**, 315-319 (2010).
- 7 Dick, A. R. & Sanford, M. S. Transition metal catalyzed oxidative functionalization of C-H bonds. *Tetrahedron* **62**, 2439-2463 (2006).
- 8 Simmons, E. M. & Hartwig, J. F. Catalytic functionalization of unactivated primary C-H bonds directed by an alcohol. *Nature* **481**, 70-73 (2012).
- 9 Leow, D., Li, G., Mei, T.-S. & Yu, J.-Q. Activation of remote meta-C-H bonds assisted by an end-on template. *Nature* **486**, 518-522 (2012).
- 10 Das, S., Incarvito, C. D., Crabtree, R. H. & Brudvig, G. W. Molecular recognition in the selective oxygenation of saturated C-H bonds by a dimanganese catalyst. *Science* **312**, 1941-1943 (2006).
- 11 Das, S., Brudvig, G. W. & Crabtree, R. H. High turnover remote catalytic oxygenation of alkyl groups: How steric exclusion of unbound substrate contributes to high molecular recognition selectivity. *J. Am. Chem. Soc.* **130**, 1628-1637 (2008).
- 12 Breslow, R. *et al.* Remote oxidation of steroids by photolysis of attached benzophenone groups. *J. Am. Chem. Soc.* **95**, 3251-3262 (1973).
- 13 Yang, J., Gabriele, B., Belvedere, S., Huang, Y. & Breslow, R. Catalytic oxidations of steroid substrates by artificial cytochrome P-450 enzymes. *J. Org. Chem.* **67**, 5058-5067 (2002).
- 14 Grieco, P. A. & Stuk, T. L. Remote oxidation of unactivated carbon-hydrogen bonds in steroids via oxometalloporphinates. *J. Am. Chem. Soc.* **112**, 7799-7801 (1990).
- 15 Groves, J. T. & Neumann, R. Regioselective oxidation catalysis in synthetic phospholipid vesicles. Membrane-spanning steroidal metalloporphyrins. *J. Am. Chem. Soc.* **111**, 2900-2909 (1989).
- 16 Breslow, R. Biomimetic control of chemical selectivity. *Acc. Chem. Res.* **13**, 170-177 (1980).
- 17 Jr, L. Q. & Tolman, W. B. Biologically inspired oxidation catalysis. *Nature* **455**, 333-340 (2008).
- 18 Coelho, P. S., Brustad, E. M., Kannan, A. & Arnold, F. H. Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. *Science* **339**, 307-310 (2013).
- 19 Coon, M. J. Cytochrome P450: Nature's most versatile biological catalyst. *Annual Review of Pharmacology and Toxicology* **45**, 1-25 (2005).
- 20 Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem. Res. Toxicol.* **14**, 611-650 (2001).
- 21 Montellano, P. R. O. d. *Cytochrome P450, Structure, Mechanism and Biochemistry*. Vol. 3 (Kluwer Academic/Plenum Publishers, 2005).
- 22 Guengerich, F. P. Cytochrome P450s and other enzymes in drug metabolism and toxicity. *AAPS J.* **8**, E101-E111 (2006).
- 23 Larsen, A. T., May, E. M. & Auclair, K. Predictable stereoselective and chemoselective hydroxylations and epoxidations with P450 3A4. *J. Am. Chem. Soc.* **133**, 7853-7858 (2011).
- 24 Urlacher, V. B. & Eiben, S. Cytochrome P450 monooxygenases: Perspectives for synthetic application. *Trends Biotechnol.* **24**, 324-330 (2006).
- 25 Jung, S. T., Lauchil, R. & Arnold, F. H. Cytochrome P450: Taming a wild type enzyme. *Curr. Opin. Biotech.* **22**, 809-817 (2011).
- 26 Chen, C. *et al.* Scanning chimeragenesis: The approach used to change the substrate selectivity of fatty acid monooxygenase CYP102A1 to that of terpene w-hydroxylase CYP4C7. *J. Biol. Inorg. Chem.* **15**, 159-174 (2010).

- 27 Lewis, J. C. *et al.* Combinatorial alanine substitution enables rapid optimization of cytochrome P450BM3 for selective hydroxylation of large substrates. *ChemBioChem* **11**, 2502-2505 (2010).
- 28 Kille, S., Zilly, F. E., Acevedo, J. P. & Reetz, M. T. Regio- and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nat. Chem. Biol.* **3**, 738-743 (2011).
- 29 Sawayama, A. *et al.* A panel of cytochrome P450 BM3 variants to produce drug metabolites and diversify lead compounds. *Chem. Eur. J.* **15**, 11723-11729 (2009).
- 30 Anzai, Y. *et al.* Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in biosynthesis of mycinamicin macrolide antibiotics. *Chem. Biol.* **15**, 950-959 (2008).
- 31 Xue, Y., Wilson, D., Zhao, L., Liu, H.-W. & Sherman, D. H. Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the pikC-encoded cytochrome P450 in *Streptomyces venezuelae*. *Chem. Biol.* **5**, 661-667 (1998).
- 32 Raadt, A. d., Griengl, H. & Weber, H. The concept of docking and protecting groups in biohydroxylation. *Chem. Eur. J.* **7**, 27-31 (2001).
- 33 Braunegg, G. *et al.* The concept of docking/protecting groups in biohydroxylation. *Angew. Chem. Int. Ed. Engl.* **38**, 2763-2766 (1999).
- 34 Lairson, L. L., Watts, A. G., Wakarchuk, W. W. & Withers, S. G. Using substrate engineering to harness enzymatic promiscuity and expand biological catalysis. *Nat. Chem. Biol.* **2**, 724-728 (2006).
- 35 Raadt, A. d. & Griengl, H. The use of substrate engineering in biohydroxylation. *Curr. Opin. Biotech.* **13**, 537-542 (2002).
- 36 Li, S. *et al.* Selective oxidation of carbolide C-H bonds by an engineered macolide P450 mono-oxygenase. *Proc. Natl. Acad. Sci. USA* **104**, 18463-18468 (2009).
- 37 Li, S., Podust, L. M. & Sherman, D. H. Engineering and analysis of self-sufficient biosynthetic cytochrome P450 PikC fused to the RhFRED reductase domain. *Journal of the American Chemical Society* **129**, 12940-12941 (2007).
- 38 Xue, Y., Zhao, L., Liu, H.-W. & Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity. *Proc. Natl. Acad. Sci. USA* **95**, 12111-12116 (1998).
- 39 Sherman, D. H. *et al.* The structural basis for substrate anchoring, active site selectivity, and product formation by P450 PikC from *Streptomyces venezuelae*. *Journal of Biological Chemistry* **281**, 26289-26397 (2006).
- 40 Li, S., Ouellet, H., Sherman, D. H. & Podust, L. M. Analysis of transient and catalytic desosamine-binding pockets in cytochrome P450 PikC from *Streptomyces venezuelae*. *J. Biol. Chem.* **284**, 5723-5730 (2009).
- 41 Chen, H. *et al.* Expression, purification, and characterization of two *N,N*-dimethyltransferases, TylM1 and DesVI, involved in the biosynthesis of mycaminose and desosamine. *Biochemistry* **41**, 9165-9183 (2002).
- 42 Chenault, H. K. & Whitesides, G. M. Regeneration of nicotinamide cofactors for use in organic synthesis. *Applied Biochemistry and Biotechnology* **14**, 147-197 (1986).
- 43 Wong, C.-H., Gordon, J., Cooney, C. L. & Whitesides, G. M. Regeneration of NAD(P)H using glucose-6-sulfate and glucose-6-phosphate dehydrogenase. *Journal of Organic Chemistry* **46**, 4676-4679 (1981).
- 44 Borisova, S. A., Zhao, L., Sherman, D. H. & Liu, H.-w. Biosynthesis of desosamine: Construction of a new macrolide carrying a genetically designed sugar moiety. *Organic Letters* **1**, 133-136 (1999).
- 45 Oh, H.-S. & Kang, H.-Y. Total synthesis of neomethymycin and novamethymycin. *Tetrahedron* **66**, 4307-4317 (2010).
- 46 Oh, H.-S., Xuan, R. & Kang, H.-Y. Total synthesis of methymycin. *Org. Biomol. Chem.* **7**, 4458-4463 (2009).
- 47 Inanaga, J., Kawanami, Y. & Yamaguchi, M. Total synthesis of neomethynolide. *Bull. Chem. Soc. Jpn.* **59**, 1521-1528 (1986).
- 48 Ditrich, K. Total synthesis of methynolide. *Liebigs Ann. Chem.* **1990**, 789-793 (1990).
- 49 Oikawa, Y., Tanaka, T. & Yonemitsu, O. Highly stereoselective synthesis of methynolide, the aglycone of the 12-membered ring macrolide methymycin, from D-glucose. *Tet. Lett.* **27**, 3647-3650 (1986).
- 50 Vedejs, E., Buchanan, R. A. & Wantabe, Y. Total synthesis of d,1-methynolide. Sulfur removal and remote stereocontrol. *J. Am. Chem. Soc.* **111**, 8430-8438 (1989).
- 51 Kittendorf, J. D. & Sherman, D. H. The methymycin/pikromycin pathway: a model for metabolic diversity in natural product biosynthesis. *Bioorganic & Medicinal Chemistry* **17**, 2137-2146 (2009).
- 52 Schlünzen, F. *et al.* Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **413**, 814-821 (2001).

- 53 Sikkema, J., Bont, J. A. M. d. & Poolman, B. Interactions of cyclic hydrocarbons with biological membranes. *J. Biol. Chem.* **269**, 8022-8028 (1994).
- 54 Holton, J. & Alber, T. Automated protein crystal structure determination using ELVES. *Proc. Natl. Acad. Sci. USA* **101**, 1537-1542 (2004).
- 55 Emsley, P. & Cowtan, K. Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132 (2004).
- 56 Murshudov, G., Vagin, A. & Dodson, E. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240-255 (1997).
- 57 Collaborative Computational Project, Number 4. *Acta Crystallogr* **50**, 760-763 (1994).
- 58 Schenkman, J., Remmer, H. & Estabrook, R. *Mol Pharmacol* **3**, 113-123 (1967).

Notes:

During preparation of this chapter, this work has been submitted to the *Journal of American Chemical Society* for review.

The atomic coordinates and structure factors (PDB ID codes 4BF4, 3ZK5, 4B7D and 4B7S) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

Author Contributions

Karoline C. Chiou and Solymar Negretti contributed equally to this study. Karoline C. Chiou, Solymar Negretti, Petrea M. Kella, Alison R.H. Narayan, Larissa M. Podust, John Montgomery, and David H. Sherman conducted the experimental design; Karoline C. Chiou, Solymar Negretti, Petrea M. Kella, and Alison R.H. Narayan performed the experimental work. Solymar Negretti and Alison R.H. Narayan synthesized substrates and authentic hydroxylated standards, and performed product characterization. Karoline C. Chiou performed the enzymatic reactions, kinetic experiments, and anti-bacterial assays. Larissa M. Podust obtained crystallographic data and solved the structures. Karoline C. Chiou, Solymar Negretti, Petrea M. Kella, Alison R.H. Narayan, Larissa M. Podust, John Montgomery, and David H. Sherman evaluated the data.

Chapter 3

P450-Mediated Oxidation of FDA-Approved Drug Scaffolds

3.1 Introduction

Cytochrome P450 monooxygenases (P450s or CYPs) have been implicated in human metabolism and have therefore been a major research target for many decades. In humans, 57 distinct P450s have been identified and we continue to learn more about each of those enzymes and their roles in human diseases.¹ Most of the medically relevant P450s act on important endogenous substrates to control physiological levels, as well as exogenous compounds such as medications we ingest. Therefore, alteration in activity levels of these enzymes in the human body can result in serious diseases. Additionally, human CYPs play an integral role in the metabolism of xenobiotics and pharmaceuticals that include over 80% of the drugs routinely prescribed.² During the body's initial clearance, most drugs are metabolized in the liver by resident P450s.^{3,4} Some of the subsequent metabolites are biologically active themselves; therefore, understanding the bioactivity of *in vivo* metabolites is crucial in determining efficacy, toxicity, and pharmacokinetics.⁵ With respect to those goals, terfenadine, an antihistamine which was marketed by Hoechst Marion Roussel (now Sanofi-Aventis), exemplified the importance of determining the active metabolites for identifying well-tolerated and safer therapies.⁶

Synthesis of sufficient quantities of pure putative metabolites is labor-intensive and difficult, so exploitation of P450s to generate metabolites is quite attractive. Hepatic microsomes are a potential source of human CYPs, but their availability is limited and they often exhibit variable expression levels, complicating preparative-scale metabolite synthesis. Although some human enzymes have been successfully expressed in recombinant hosts such as *Escherichia coli*, most human membrane-bound, multi-protein systems commonly suffer from

misfolding and aggregation or do not express in an active form. Previous groups have demonstrated measurable metabolite preparation in *E. coli* and insect cells, but these systems were costly and demonstrate low productivities given limited stability and slow enzymatic reaction rates.⁷⁻¹⁰

An alternative approach is employing engineered bacterial P450s in place of microsomal CYPs. For example, Otey et al. demonstrated that an engineered version of cytochrome P450 BM3, a fatty acid hydroxylase from *Bacillus megaterium*, could be used in conjunction with hydrogen peroxide to prepare authentic human metabolites of propranolol, a multi-function beta-adrenergic blocker used to treat hypertension, arrhythmia, angina, migraines, overactive thyroids, and anxiety.^{11,12} Of note is the use of BM3 for the study because this P450 is known to be self-sufficient, opening the door to employing other bacterial P450s for the synthesis of human cytochrome P450 monooxygenase (CYP) metabolites of pharmaceuticals.

Perhaps equally important to being able to predict *in vivo* metabolites of potential pharmaceuticals, growing antibiotic resistance among pathogenic bacterial strains represents a significant global health threat.¹³ A promising strategy to combat resistance is to introduce antibiotics possessing a unique mechanism of action. In particular, the ribosome is a common target of many antibacterials given the distinct differences between pathogenic bacterial and eukaryotic ribosomes. Specifically, position 2058 within the peptidyl transferase center (PTC) in the macrolide binding pocket (adenine in eubacteria, guanine in eukaryotes) offers an efficient target for antibiotic selectivity given the use of a different base in eubacteria compared to eukaryotes. The PTC is almost completely conserved across kingdoms of life, yet it also serves as a target for several families of antibiotics. One family is the pleuromutilins, which target the bacterial ribosome and still show with a slow development of resistance.¹⁴

In the current study, we used a previously engineered self-sufficient P450, PikC_{D50N}-RhFRED, originally derived from the pikromycin pathway to produce oxidation products consistent with potential hydroxylated human and animal metabolites of tamoxifen, toremifene, and tiamulin in yields ranging from 16-35% (**Fig. 3-1**).¹⁵ In addition, we tested several synthetic undecorated hydrocarbons and macrocycles with terminal *N,N*-dimethylamino moieties to explore the ability of PikC to oxidize these unnatural substrates. We hypothesized that the terminal functionality of the linkers facilitated regioselective oxidation by PikC through a salt

bridge comparable to that observed with the endogenous PikC substrates, YC-17 and narbomycin, and other synthetic unnatural substrates previously tested (see Chapter 2).¹⁶

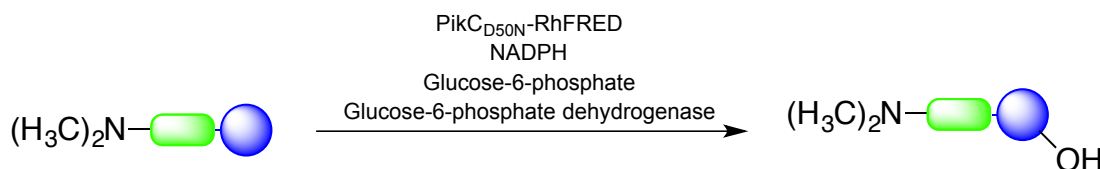


Figure 3-1. Enzymatic scheme of PikC-mediated oxidation of unnatural substrates. Potential PikC substrate with a linker ending in a terminal *N,N*-dimethylamino to facilitate substrate binding and oxidation with NADPH and the NADPH regeneration system of glucose-6-phosphate and glucose-6-phosphate dehydrogenase.

3.2 Background

3.2.1 Tiamulin and Other Pleuromutilins

Tiamulin (**Fig. 3-2**) is a semisynthetic derivative of the natural product pleuromutilin discovered to be produced by *Pleurotus mutilus* (now known as *Clitopilus scyphoides*).¹⁷ The compound contains a tricyclic mutilin core of a cyclo-pentanone, cyclo-hexyl and cyclo-octane, a C21 keto group critical for bioactivity, and a (((2-diethylamino)ethyl)thio)-acetic acid side-chain on C14 of the octane ring.^{17,18} During the early 1980s, significant strides were made towards developing formulation methods for azamulin (**Fig. 3-2**), a related compound to tiamulin, to accelerate its entry into clinical use because it demonstrated activity against many clinical isolates of resistant bacteria, including erythromycin- and tetracycline-resistant strains. However, the compound strongly inhibited CYP activity, severely limiting its clinical use.¹⁹ Nevertheless, continued efforts to develop pleuromutilins yielded a number of semisynthetic derivatives, including retapamulin (SB-275833; SmithKline Beecham) (**Fig. 3-2**) which demonstrated potent activity against Gram-positive pathogens and a low propensity to develop resistance in all strains of *Staphylococcus aureus* and *Streptococcus pyogenes* with MICs $\leq 0.5 \mu\text{g/ml}$.²⁰⁻²⁹ Tiamulin and the other pleuromutilins exhibit broad-spectrum activity against a variety of clinically relevant bacterial strains, but unlike retapamulin, many suffer from limited oral bioavailability, restricting the use of this class of compounds in humans.^{26,30,31} Nevertheless, with the increasing number of dangerous human pathogens resistant to currently used medications, there is increasing interest in developing new antibiotic scaffolds such as pleuromutilin for human therapy.³²⁻³⁶

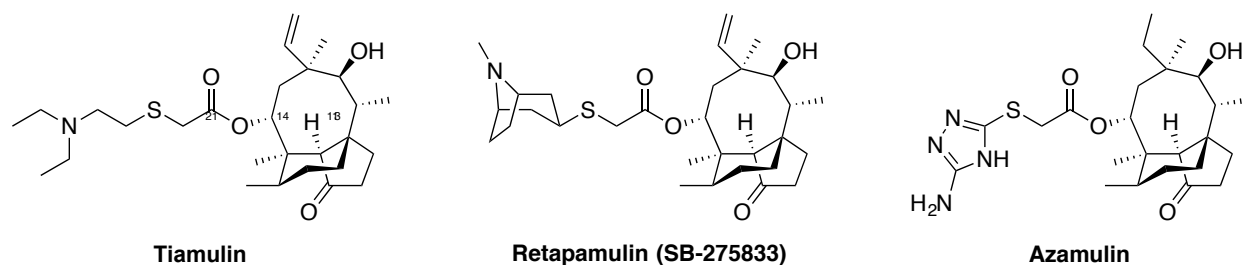


Figure 3-2. Structures of tiamulin, retapamulin (SB-275833), and azamulin.

Specific interactions between pleuromutilin compounds and the ribosome, the predicted biological target, have been found using the 50S ribosomal subunit from *Deinococcus radiodurans* (D50S) in complex with tiamulin.³⁷ The crystal structure unambiguously localized tiamulin in the large subunit through an extensive network of hydrophobic interactions with the 23S rRNA.³⁷ In addition, a number of related compounds were also crystallized in the *D. radiodurans*' ribosome and all demonstrated that C14 extensions were localized to the PTC and held in place by a hydrogen-bond network between G2061 and the essential C21 keto of the bound compound.¹⁴ Furthermore, the C14 extensions were located between the acetylated and peptidylated tRNA CCA ends.¹⁴ However, this area of the ribosome contained only a few potential candidates for interaction so various chemical moieties of the inhibitor should be tolerated, explaining why all C14 extensions did not have extensive interactions with the PTC. Therefore, longer extensions may be tolerated and allow for substantial flexibility in drug design.

3.2.2 Tamoxifen and Related Analogs

Breast cancer affects over 1 million women and kills approximately 400,000 every year in the Western world.³⁸ Breast cancer is often classified according to its estrogen (ER) or progesterone (PR) receptor status to divide the patients into appropriate treatment groups. As more than two thirds of breast cancers are positive in immunohistochemical staining to the ER, personalized hormonal therapies have become a large proportion of treatment options for these breast cancer patients. Additionally, differentiating breast cancer into these categories allows for substantial increases in treatment efficacy as it was found that ER-positive (ER+) tumors are highly dependent on estrogen signaling for growth and replication. These ER+ tumors can then be treated with aromatase inhibitors (AI), while PR-positive tumors are not estrogen dependent and tend to respond better to alternative treatments.³⁹

Tamoxifen (**Fig. 3-3**), a selective estrogen receptor modulator (SERM) introduced to clinical use in 1977, has subsequently become the gold standard for treating ER+ early, metastatic, and adjuvant cancer situations to become the first form of molecular targeted therapy, allowing ER+ patients to derive the greatest benefit.³⁹⁻⁴² However, tamoxifen's indications have expanded to include advanced breast cancer in premenopausal women and men, as well as the adjuvant therapy for node-positive and -negative disease.³⁹ Tamoxifen has prolonged disease-free survival, significantly reduced mortality rates, and a 39% reduction in the risk for developing contralateral breast cancer.⁴¹ The therapeutic mechanism of action of tamoxifen is antagonistically binding to the ER α normally present in the mammary gland to block the ER from binding its endogenous ligands.^{40,43} Tamoxifen also acts agonistically in the bone in addition to its anticancer activity to increase bone density and prevent osteoporosis.⁴⁰ The biological activity of tamoxifen and other selective ER modulators (SERMs) depends on the balance between co-activators and co-repressor proteins, differently represented in healthy tissue and breast cancer.^{44,45}

Toremifene, a tamoxifen analog (**Fig. 3-3**), received Food and Drug Administration approval in 1997 under the name Fareston for use in metastatic breast cancer in postmenopausal women.⁴⁶ Toremifene's pharmacologic profile is comparable to that of tamoxifen in terms of its ER binding, antitumor activity, and estrogenic activity, but it also demonstrates lower estrogenic activity than tamoxifen at low and moderate doses.^{47,48} The major difference between toremifene and tamoxifen is the hepatocarcinogenicity reported in animal studies, as high-dose tamoxifen was hepatocarcinogenic in rats while the same dose of toremifene was not.⁴⁹ Furthermore, the 2 major active metabolites of tamoxifen, 4-hydroxy-*N*-desmethyl-tamoxifen (endoxifen) and *Z*-4-hydroxytamoxifen (4HT), are 100 times more potent than the parent drug.⁵⁰⁻⁵²

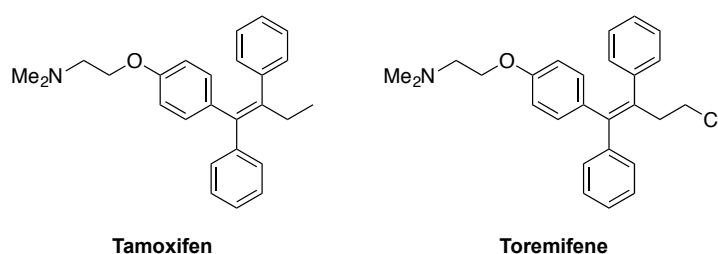


Figure 3-3. Structures of tamoxifen and toremifene.

3.3 Results and Discussion

PikC is a biosynthetic cytochrome P450 monooxygenase involved in the post-polyketide synthase tailoring steps of macrolide antibiotic biosynthesis in *Streptomyces venezuelae*.^{53,54} The endogenous activity of this enzyme is the hydroxylation of the 12-membered macrolide, YC-17, and the 14-membered macrolide, narbomycin, to produce methymycin/neomethymycin and pikromycin, respectively (**Fig. 3-4**). Previous analysis of X-ray co-crystal structures of PikC with its natural substrates highlighted largely non-specific hydrophobic interactions between the substrate macrolactone rings and active site residues. However, the aminosugar desosamine, present in both natural macrolide substrates, was notably localized in two distinct binding pockets participating in hydrogen bonding and ionic interactions for appropriate substrate positioning and stabilization in the active site.^{16,55} Specifically, the protonated dimethylamino moiety of desosamine and a glutamate residue in the B/C loop region of the protein participated in an electrostatic salt bridge to achieve specific substrate binding.⁵⁵ In addition, desosamine has already been shown to direct unfunctionalized carbocycles to the active site of PikC for regioselective oxidation, albeit at lower levels of selectivity than YC-17 and narbomycin.¹⁵ The difficulty of harvesting and synthetically installing of desosamine motivated the exploration of simplified anchoring groups with an ability to mediate similar electrostatic and H-bonding interactions for effective PikC-catalyzed C-H bond activation.^{56,57}

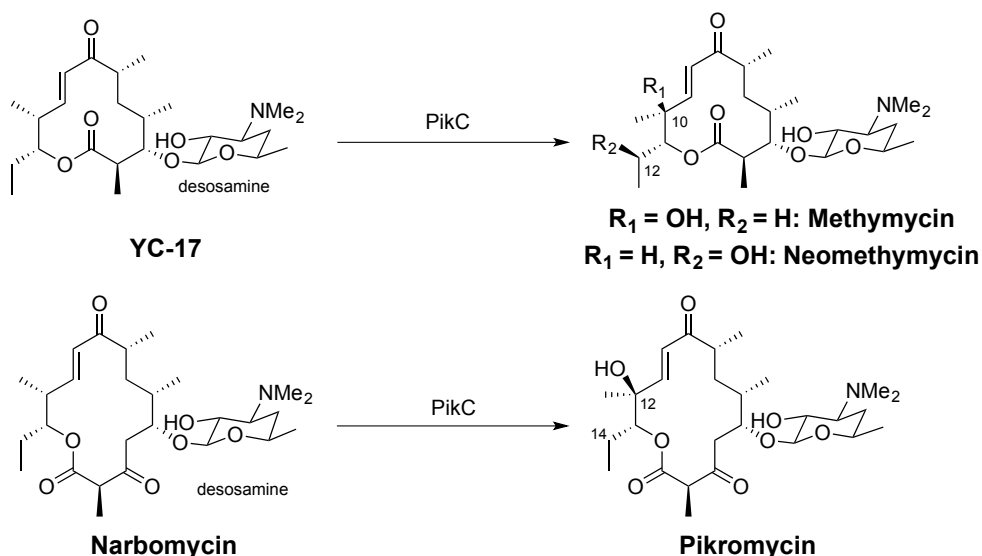


Figure 3-4. Endogenous PikC-catalyzed reactions of YC-17 and narbomycin. Endogenous reactions catalyzed by PikC include the oxidation of YC-17 and narbomycin to yield methymycin, neomethymycin, and pikromycin, respectively.

3.3.1 PikC-Mediated Oxidation of Putative Semisynthetic Substrates

To build upon earlier investigations into PikC's ability to oxidize synthetic hydrocarbons and macrocycles with mimics of desosamine (see Chapter 2), we began our studies in collaboration with Dr. John Montgomery (University of Michigan) with small semisynthetic molecules bearing a series of various linkers all terminating in a *N,N*-dimethylamino group. We focused on linkers that we had previously confirmed to target other substrates into the PikC active site for efficient regioselective hydroxylation. We subsequently tested a number of compounds of varying size and degree of functionality and observed a range of conversions up to 21% (**Table 3-1**) using the engineered chimeric enzyme, PikC_{D50N}-RhFRED (to be referred to as PikC-RhFRED).

From these results, we had several observations concerning substrate targeting for PikC oxidation. First, from the P450 oxidation of **1** and **2**, we observed that the acyclic ester linker ending in a *N,N*-dimethylamino moiety successfully facilitated PikC oxidation albeit at lower levels than **1**. These results were surprising given the degree of functionality present in these molecules as compared to previous work with an unfunctionalized hydrocarbon with the same linker. In that case, a 34% conversion was observed across 4 products, an almost 3-fold increase in conversion. Nevertheless, the data suggested a threshold of functionality required for productive substrate binding and conversion. Second, the oxidation pattern and conversion of **3** reconfirmed previously observed data that suggested that the terminal *N,N*-dimethylamino was crucial to directing the substrate to the enzyme active site. Compounds **5** and **6** revealed that the length of the linker was crucial to reaching the catalytic heme and achieve regioselective oxidation. Third, enzymatic conversion of **4** stood in contrast to previously obtained data of the observed oxidation of a similar compound that contained an ester in place of the amide of **4** (see Chapter 2). The decreased conversion with **4** suggested the critical role of the oxygen of the ester, perhaps as an acceptor in a hydrogen bond network. The presence of the nitrogen in the amide-containing compound would then clearly prohibit this interaction. Alternatively, the ester might allow for a more optimal compound configuration that was disallowed by the amide. Finally, increased functionality of the macrolactone facilitated oxidation by PikC as demonstrated by compound **7**, a narbonolide derivative. This phenomenon was likely due to favorable enzyme-substrate interactions that may have mirrored the interactions observed with 10-deoxymethynolide (10-dml) derivatives (previous data).

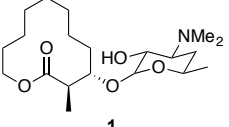
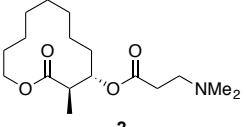
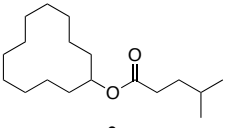
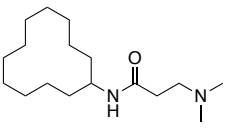
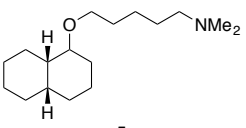
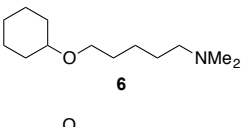
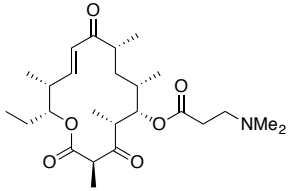
<u>Compound</u>	<u>No. products</u>	<u>Conversion</u>
 1	4	12
 2	4	6
 3	3	Trace
 4	4	16
 5	3	14
 6	0	0
 7	2	21

Table 3-1. Regioselectivity and conversion of PikC against semisynthetic compounds. The enzymatic reactions of semisynthetic compounds were analyzed by LC-MS Q-TOF after overnight incubation.

3.3.2 PikC Oxidation of Tamoxifen and Related Compounds

While the enzymatic C-H bond activation test compounds were interesting, the predicted lack of biological activity of both the substrates and products motivated expansion of our biocatalyst advancements towards useful synthetic applications. Therefore, investigations into the potential of PikC to oxidize structurally distinct scaffolds with medical relevance were begun by conducting a high-throughput screen of currently United States Food and Drug Administration (FDA)-approved drugs to determine if any of those scaffolds would be accepted

as substrates by PikC-RhFRED (SI). Approximately 100 compounds were tested from the library; of those compounds, approximately 20 compounds were regioselectively oxidized with conversions varying from less than 10% to 30% (data not shown). One of the compounds included in the screen was toremifene, which was confirmed to be oxidized by the biocatalyst in a modest yield (**Table 3-2**). Commercially available analogs to toremifene, tamoxifen and endoxifen, (**Table 3-2**) were tested against PikC-RhFRED for regioselectivity and total conversion. Tamoxifen was oxidized at a 34% yield across 4 products while toremifene was verified to be oxidized in a 16% yield across 2 products (**Table 3-2**). Interestingly, oxidation of endoxifen was not observed, further substantiating that the *N,N*-dimethylamino was crucial to achieving productive substrate binding.

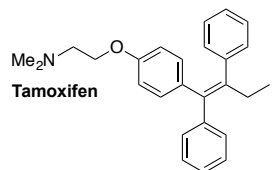
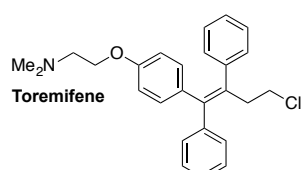
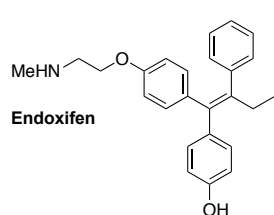
<u>Compound</u>	<u>No. products</u>	<u>Conversion</u>
 <p>Tamoxifen</p>	4	34
 <p>Toremifene</p>	2	16
 <p>Endoxifen</p>	0	0

Table 3-2. PikC regioselectivity and conversion of tamoxifen derivatives. The enzymatic reactions of commercially available tamoxifen derivatives were analyzed by LC-MS Q-TOF after overnight reaction incubation.

3.3.3 Regioselective Oxidation of Pleuromutilin Derivatives

To further expand our studies with biologically relevant chemical scaffolds, we examined pleuromutilins as potential PikC substrates. In collaboration with Dr. Erik Sorensen's laboratory (Princeton University), we were able to test a number of pleuromutilin derivatives similar to tiamulin, the commercially available veterinary medication. Against tiamulin, PikC-RhFRED

yielded a 31% yield across 2 products while pleuromutilin, lacking the thioester with the terminal *N,N*-dimethylamino and instead ending in a carboxylic acid, resulted in only an 8% yield with 1 product. Furthermore, analogs **1** and **2** also demonstrated poor yields of 3% across 2 products, while analog **3** showed a 35% yield across 6 products (**Table 3-3**). From these results, it was clear that the length of the thioester linker was crucial, leading to the hypothesis that the terminal moiety was participating in the same salt bridge previously observed. It appeared that the endocyclic double bond in conjunction with the absence of the hydroxyl group at C14 of the pleuromutilin derivatives facilitated hydroxylation while the exocyclic double bond and the presence of the hydroxyl group inhibited PikC-mediated oxidation.

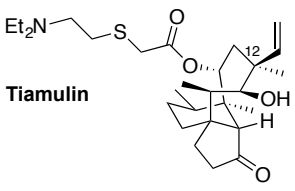
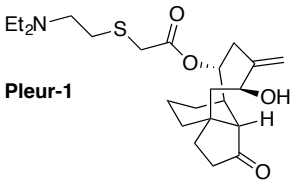
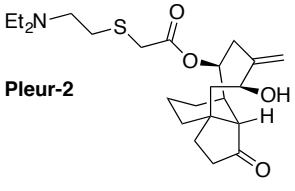
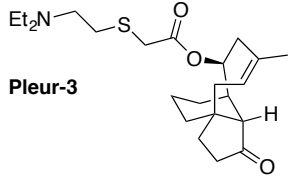
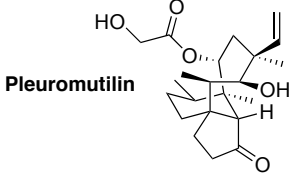
<u>Compound</u>	<u>No. products</u>	<u>Conversion</u>
 Tiamulin	2	34
 Pleur-1	2	3
 Pleur-2	2	3
 Pleur-3	6	38
 Pleuromutilin	1	8

Table 3-3. PikC regioselective oxidation of pleuromutilin derivatives. The enzymatic reactions of synthetic pleuromutilin derivatives were analyzed by LC-MS Q-TOF following an overnight reaction for the overall conversion.

3.4 Conclusion

In the current study, we demonstrated that PikC regioselectively oxidized several structurally distinct scaffolds with biological relevance. Although the semisynthetic compounds (**Table 3-1**) were successfully oxidized, the compounds lacked predictable bioactivity given the lack of chemical functionality. Therefore, a screen of FDA-approved drugs (**Table 3-2**, **Table 3-3**) was explored to expand the practical use of PikC. In particular, tamoxifen and pleuromutilin

and their derivatives are FDA-approved for breast cancer and as a veterinary antibiotic, respectively. The driving hypothesis for the enzymatic conversion of drug scaffolds came from a substrate-engineering approach previously used that exploited PikC's inherent substrate flexibility using linear desosamine replacements. Although additional work remains to characterize the location of oxidation through NMR, it is reasonable to conclude that the regions of the substrates that were oxidized were distal to the *N,N*-dimethylamino functionality based on previous results, rendering the oxidation pattern orthogonal to what has been achieved using directed or non-directed small molecule oxidation catalysts.⁵⁸⁻⁶⁵ In spite of the surprisingly diverse substrate scope demonstrated in this work, directed evolution or rational site-directed mutagenesis methods could be employed to further optimize both conversion and selectivity of the mutant chimera.

Although there has been long-standing interest in developing P450s as potential industrial catalysts, there have only been a select number of enzymes that have been integrated into commercial use. This scarcity is attributed to the limited chemical repertoire of natural enzymes. Thus, enzymes exhibiting improved activity and increased substrate tolerance offer improved different tools for chemoenzymatic synthesis, chemical probes, reagents, metabolite production, and novel therapeutics. Specifically, bacterial biosynthetic P450s are a unique pool of candidates given their ease of acquisition through traditional molecular biology and protein purification techniques. Therefore, our substrate-engineering approach not only offers an opportunity to gain insight into P450 substrate binding and conversion, but to also produce analogs of potent bioactives already approved through the FDA and confirm *in vivo* metabolites through *in vitro* methods. For example, the pleuromutilin class of antibiotics is under extensive investigation as a valuable pool of antibiotics with a novel mechanism of action that could be used to combat the growing problem of antibiotic resistance. As a practical chemoenzymatic tool, our engineering method can thus be integrated as a general method of regio- and stereoselective oxidation of highly complex molecules, including secondary metabolites. Finally, the results underscore the potential of bacterial biosynthetic P450s to be developed as biocatalysts for selective C-H bond oxidation towards chemical diversification.

3.5 Experimental Procedures

Materials

Unless otherwise specified, all chemical reagents were ordered from Sigma-Aldrich. Protein purification used QIAGEN (Valencia, CA) Ni-NTA resin, HiTrap Chelating columns from GE Healthcare (Piscataway, NJ), Millipore (Billerica, MA) Amicon Ultra centrifugal filter, and PD-10 desalting columns from GE Healthcare (Piscataway, NJ). Luria broth and Terrific broth components were purchased from EMD Sciences (Gibbstown, NJ). Preparative-scale enzymatic reactions utilized SPE reservoirs, frits, and caps from Sigma-Aldrich.

Expression and purification of PikC_{D50N}-RhFRED

Protein expression and purification followed a previously described procedure.⁶⁶

PikC_{D50N}-RhFRED Enzymatic Assay

The standard assay contained 5 μ M PikC_{D50N}-RhFRED, 0.5 mM substrate, 2.5 mM NADPH, 0.25 units of glucose-6-phosphate dehydrogenase, and 5 mM glucose-6-phosphate for NADPH regeneration in 100 μ L of reaction buffer (50 mM NaH₂PO₄, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol). The reaction was carried out at 30°C overnight and quenched by extraction using 3 \times 200 μ L of CHCl₃. The organic extracts were combined, dried under N₂, and redissolved in 100 μ L of acetonitrile. The subsequent liquid chromatography mass spectrometry (LC-MS) analysis was performed on an Agilent Q-TOF HPLC-MS (Department of Chemistry, University of Michigan) equipped with a high resolution electrospray mass spectrometry (ESI-MS) source and a Beckmann Coulter reverse-phase HPLC system using an Waters XBridge C18 3.5 μ m, 2.1x150 mm under the following conditions: mobile phase (A = deionized water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid), 20% B for 3 min, 20 to 100% B over 25 min, 100% B for 5 min, 100 to 20% B over 1 min, 20% B for 10 min; flow rate, 0.21 mL/min.

3.6 Supplementary Information

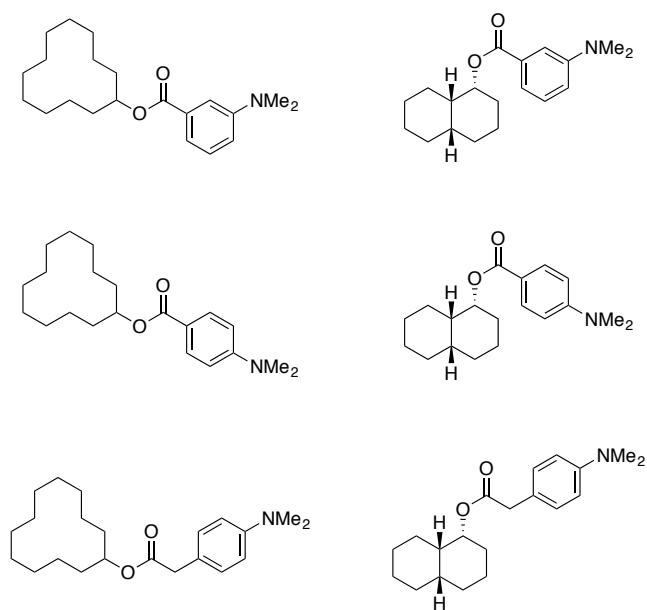


Figure S3-1. Semisynthetic compounds not tolerated by $\text{PikC}_{\text{D50N}}\text{-RhFRED}$ as substrates.

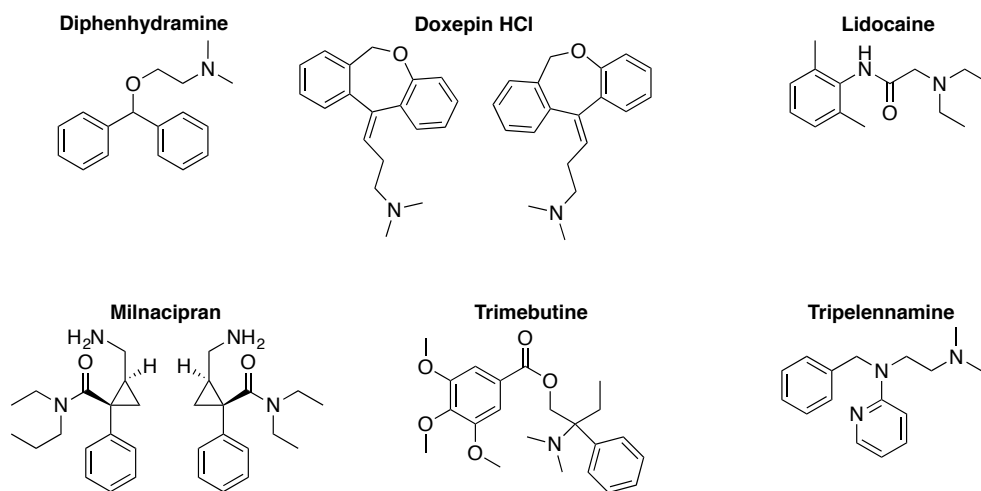


Figure S3-2. Commercially available FDA-approved compounds not tolerated by $\text{PikC}_{\text{D50N}}\text{-RhFRED}$ as substrates.

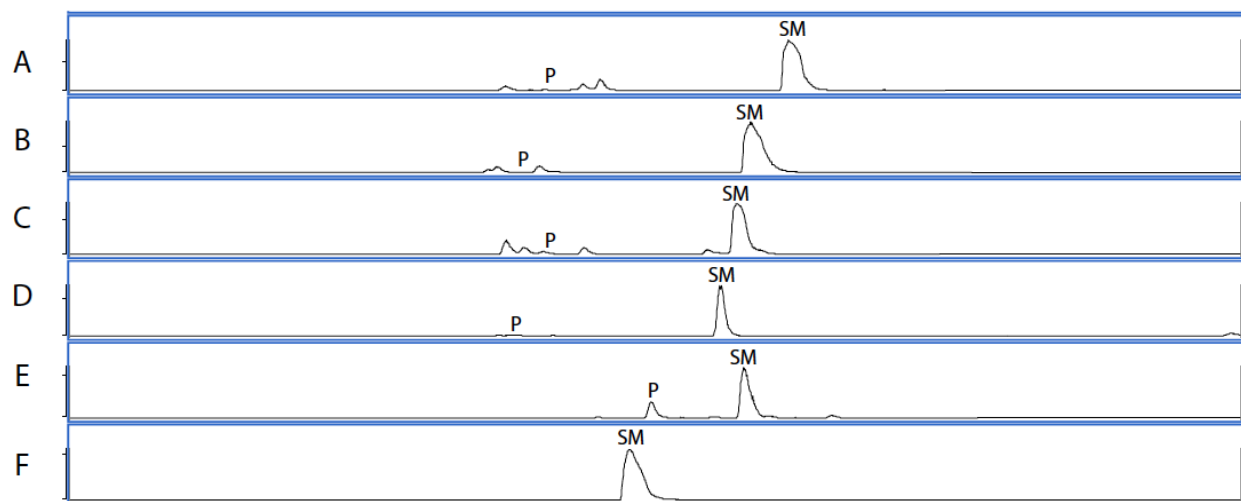


Figure S3-3. LC-MS Q-TOF traces of PikC enzymatic reactions against semisynthetic compounds. (A) Compound 5, (B) 4, (C) 1, (D) 2, (E) 7, and (F) 6. Trace for compound 3 not shown given the minimal conversion.

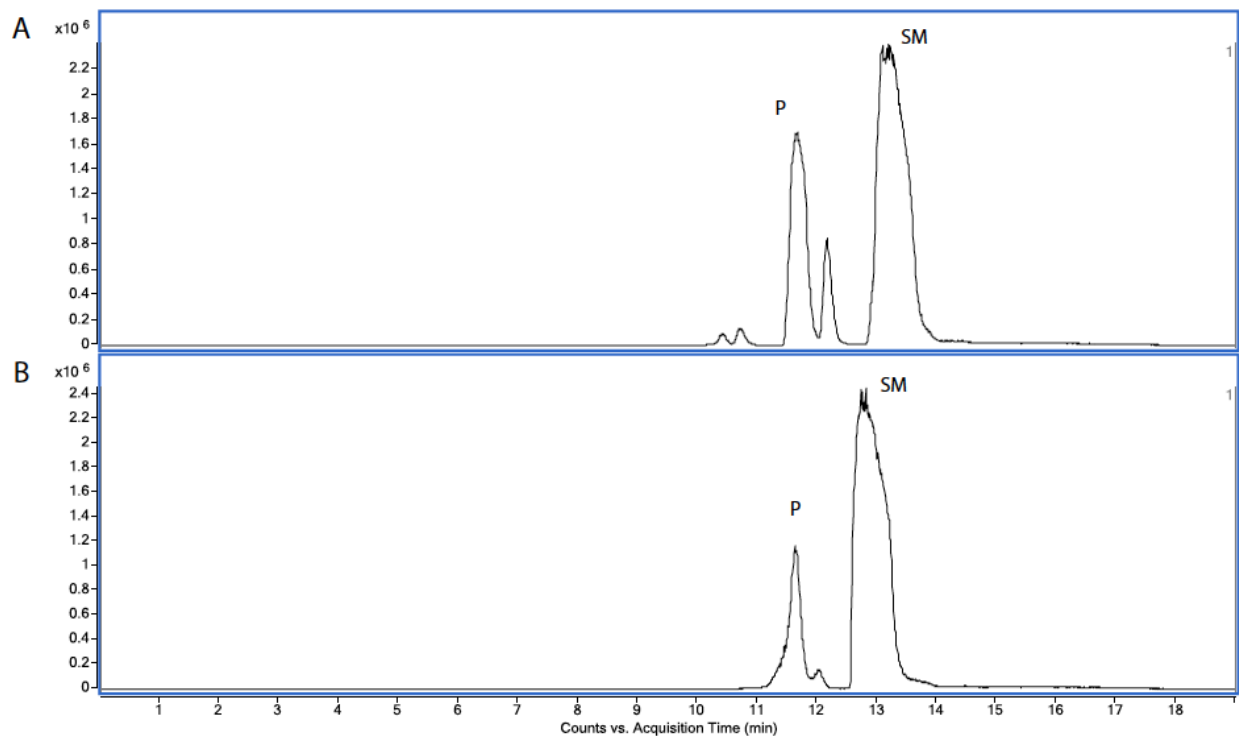


Figure S3-4. LC-MS Q-TOF traces of PikC enzymatic reactions with (A) tamoxifen and (B) toremifene. Abbreviations: starting material (SM), product(s) (P).

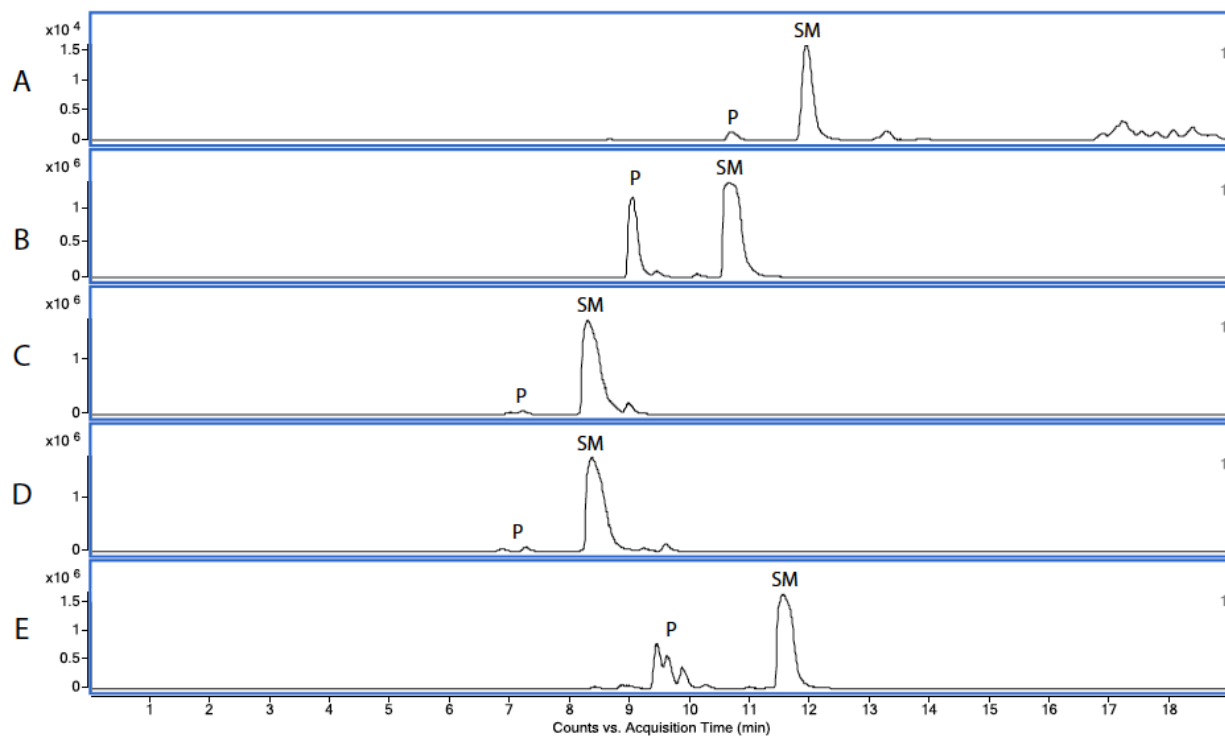
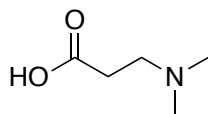


Figure S3-5. LC-MS Q-TOF traces of PikC enzymatic reactions against pleuromutilin derivatives. (A) Pleuromutilin, (B) Tiamulin, (C) Pleur-1, (D) Pleur-2, and (E) Pleur-3. Abbreviations: starting material (SM), product(s) (P).

Synthesis of 3-(dimethylamino)propanoic acid



A 0.25 M solution of methyl 3-(dimethylamino)propionate (100 mg, 0.000762 mol) was added to a 3:1 mixture of tetrahydrofuran and water solution. Two hundred fourteen milligrams of potassium hydroxide (0.00381 mol) was then added and the reaction was stirred vigorously for 1 h at room temperature. The reaction was subsequently quenched with 2N hydrochloric acid until the reaction was acidified. Then, the reaction was extracted twice with ethyl acetate. The combined organic layers were dried over sodium sulfate and evaporated to dryness *in vacuo* to afford 83 mg (83% yield) of 3-(dimethylamino)propanoic acid, a white flaky solid.

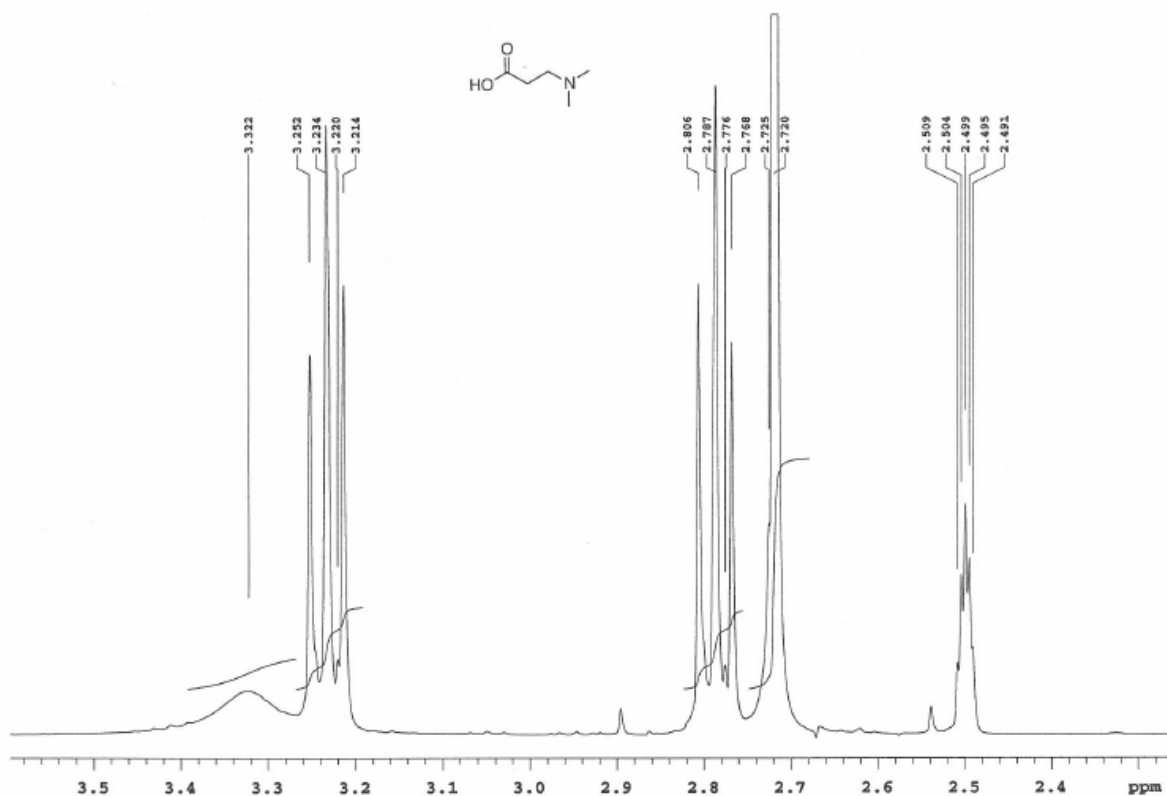
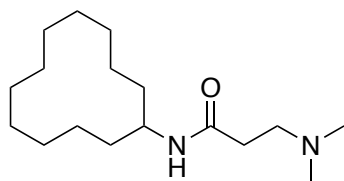


Figure S3-6. ¹H NMR of 3-(dimethylamino)propanoic acid. ¹H NMR spectra taken at 500 MHz in CDCl₃.

Synthesis of Compound 4



Seventy five milligrams of 3-(dimethylamino)propanoic acid (0.000638 mmol) was resuspended in dichloromethane and treated with 245 mg of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.00128 mmol) and 8 mg (0.0000638 mmol) 4-dimethylaminopyridine. Fresh 0.00191 mmol triethylamine and 0.000958 mmol cyclododecylamine were added and the mixture was stirred at room temperature for 18 h. The reaction was quenched with aqueous saturated ammonium chloride and diluted with 5 mL of dichloromethane and 5 mL of water. The aqueous layer was extracted with additional dichloromethane and the organic layers were combined and washed with brine. The organic layer was dried over sodium sulfate, concentrated, and purified through flash chromatography (silica gel, 10% methanol in dichloromethane with ammonium hydroxide) to afford 88 mg (50% yield) of **4**, a white crystalline powder.

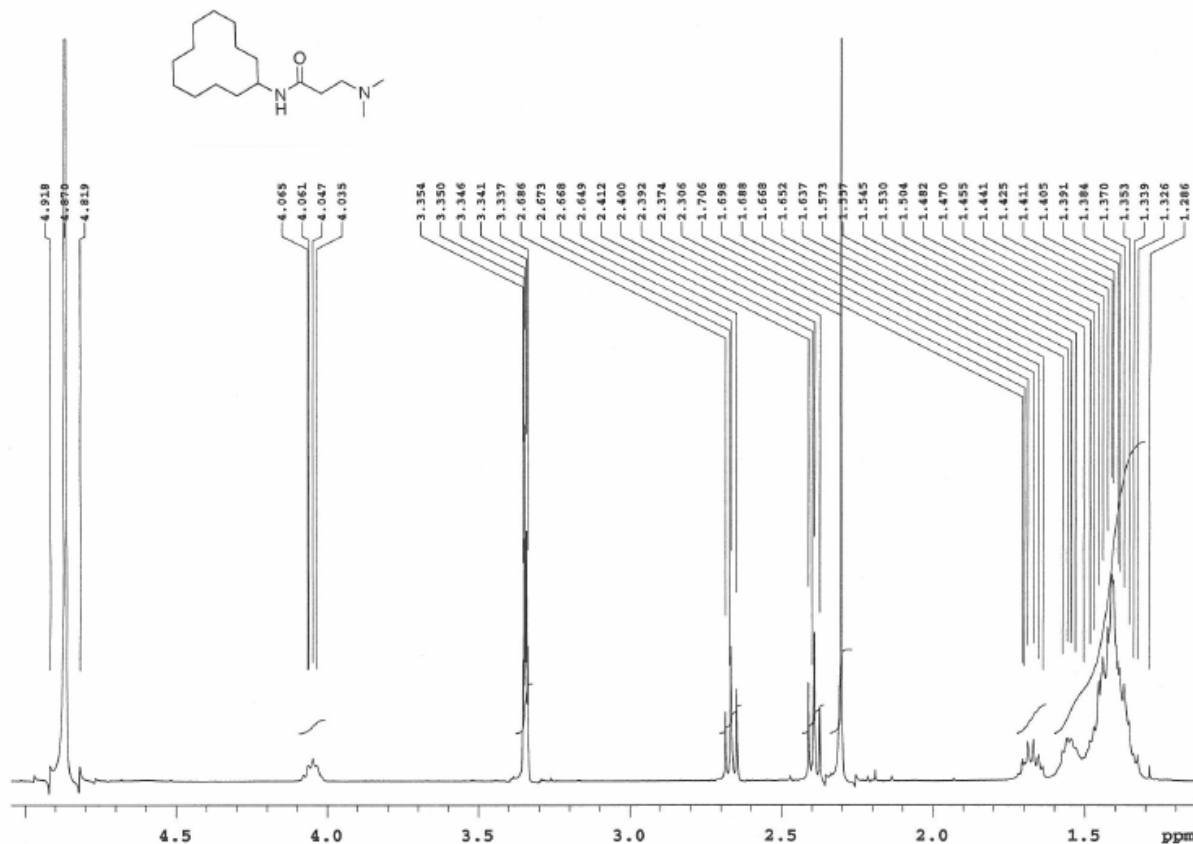


Figure S3-7. ^1H NMR of **4**. ^1H NMR spectra taken at 500 MHz in CDCl_3 .

3.7 References:

- 1 Montellano, P. R. O. d. *Cytochrome P450: Structure, Mechanism, and Biochemistry*. 3 edn, (Kluwer Academic/Plenum Publishers, 2005).
- 2 Guengerich, F. P. Cytochrome P450, drugs, and diseases. *Mol Interv* **3**, 194-204 (2003).
- 3 Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem Res Toxicol* **14**, 611-650 (2001).
- 4 Lewis, D. F. 57 varieties: The human cytochrome P450. *Pharmacogenomics* **5**, 305-318 (2004).
- 5 Johnson, M. D. *et al.* Pharmacological characterization of 4-hydroxy-*N*-desmethyl tamoxifen, a novel metabolite of tamoxifen. *Breast Cancer Res Treat* **207**, 1-9 (2004).
- 6 Markham, A. & Wagstaff, A. J. Fexofenadine. *Drugs* **55**, 269-274 (1998).
- 7 Guengerich, F. P., Gillam, E. M. & Shimada, T. New applications of bacterial systems to problems in toxicology. *Crit Rev Toxicol* **26**, 551-583 (1996).
- 8 Parikh, A., Gillam, E. M. & Guengerich, F. P. Drug metabolism by *Escherichia coli* expression human cytochromes P450. *Nat Biotechnol* **15**, 784-788 (1997).
- 9 Rushmore, T. H., Reider, P. J., Slaughter, D., Assang, C. & Shou, M. Bioreactor systems in drug metabolism: Synthesis of cytochrome P450-generated metabolites. *Metab Eng* **2**, 115-125 (2000).
- 10 Vail, R. B., Homann, M. J., Hanna, I. & Zaks, A. Preparative synthesis of drug metabolites using human cytochrome P450s 3A4, 2C9 and 1A2 with NADPH-P450 reductase in *Escherichia coli*. *J Ind Microbiol Biotechnol* **32**, 67-74 (2005).
- 11 Otey, C. R., Bandara, G., Lalonde, J., Takahashi, K. & Arnold, F. H. Preparation of human metabolites of propranolol using laboratory-evolved bacterial cytochromes P450. *Biotechnol Bioeng* **93**, 494-499 (2005).
- 12 Shand, D. G. Drug therapy: Propranolol. *N Engl J Med* **293**, 280-285 (1975).
- 13 Diekema, D. J. *et al.* Antimicrobial resistance trends and outbreak frequency in United States hospitals. *Clin Infect Dis* **38**, 78-85 (2004).
- 14 Davidovich, C. *et al.* Induced-fit tightens pleuromutilins binding to ribosomes and remote interactions enable their selectivity. *Proc Natl Acad Sci USA* **104**, 4291-4296 (2007).
- 15 Li, S. *et al.* Selective oxidation of carbolide C-H bonds by an engineered macolide P450 mono-oxygenase. *Proc. Natl. Acad. Sci. USA* **104**, 18463-18468 (2009).
- 16 Sherman, D. H. *et al.* The structural basis for substrate anchoring, active site selectivity, and product formation by P450 PikC from *Streptomyces venezuelae*. *Journal of Biological Chemistry* **281**, 26289-26397 (2006).
- 17 Kavanagh, F., Hervey, A. & Robbins, W. J. Antibiotic substances from *Basidiomycetes VIII Pleurotus mutilus* (fr.) Sacc. and *Pleurotus Passeckerianus* Pilat. *Proc Natl Acad Sci USA* **37**, 570-574 (1951).
- 18 Egger, H. & Reinshagen, H. New pleuromutilin derivatives with enhanced antimicrobial activity II. Structure-activity correlations. *J Antibiot (Tokyo)* **29**, 923-927 (1976).
- 19 Berner, H., Turnowsky, F., Laber, G. & Hildebrandt, J. New pleuromutilin derivatives, their production and pharmaceutical compositions containing them. (1980).
- 20 Free, A. *et al.* Retapamulin ointment twice daily for 5 days vs oral cephalexin twice daily for 10 days for empiric treatment of secondarily infected traumatic lesions of the skin *Skinmed* **5**, 224-232 (2006).
- 21 Parish, L. C. *et al.* Topical retapamulin ointment (1%, wt/wt) twice daily for 5 days versus oral cephalexin twice daily for 10 days in the treatment of secondarily infected dermatitis: results of a randomized controlled trial. *J Am Acad Dermatol* **55**, 1003-1013 (2006).
- 22 Kosowska-Shick, K. *et al.* Single- and multistep resistance selection studies on the activity of retapamulin compared to other agents against *Staphylococcus aureus* and *Streptococcus pyogenes*. *Antimicrob Agents Chemother* **50**, 765-769 (1006).
- 23 Pankuch, G. A. *et al.* Activity of retapamulin against *Streptococcus pyogenes* and *Staphylococcus aureus* evaluated by agar dilution, microdilution, E-test, and disk diffusion methodologies. *Antimicrob Agents Chemother* **50**, 1727-1730 (2006).
- 24 Jones, R. N., Fritsche, T. R., Sader, H. S. & Ross, J. E. Activity of retapamulin (SB-275833), a novel pleuromutilin, against selected resistant gram-positive cocci *Antimicrob Agents Chemother* **50**, 2583-2586 (2006).
- 25 Goldstein, E. J. C. *et al.* Compactive in vivo activities of retapamulin (SB-275833) against 141 clinical isolates of *Propionibacterium* spp., including 117 *P. acnes* isolates. *Antimicrob Agents Chemother* **50**, 379-381 (2006).

- 26 Boyd, B. & Castaner, J. Retapamulin. Pleuromutilin antibiotic. *Drugs Future* **31**, 107-113 (2006).
- 27 Ross, J. E. & Jones, R. N. Quality control guidelines for susceptibility testing of retapamulin (SB-275833) by reference and standardized methods. *J Clin Microbiol* **43**, 6212-6213 (2005).
- 28 Rittenhouse, S. *et al.* Selection of retapamulin, a novel pleuromutilin for topical use. *Antimicrob Agents Chemother* **50**, 3882-3885 (2006).
- 29 Rittenhouse, S., Singley, C., Hoover, J., Page, R. & Payne, D. Use of the surgical wound infection model to determine the efficacious dosing regimen of retapamulin, a novel topical antibiotic. *Antimicrob Agents Chemother* **50**, 3886-3888 (2006).
- 30 Hunt, E. Pleuromutilin antibiotics. *Drugs Future* **25**, 1163-1168 (2000).
- 31 Brooks, G. *et al.* Pleuromutilins. Part I: The identification of novel mutilin 14-carbamates. *Bioorg Med Chem* **9**, 1221-1231 (2000).
- 32 Springer, D. M. *et al.* Synthesis and activity of a C-8 ketone pleuromutilin derivative. *Bioorg Med Chem Lett* **13**, 1751-1753 (2003).
- 33 Bacque, E., Pautrat, F. & Zard, S. Z. A concise synthesis of the tricyclic skeleton of pleuromutilin and a new approach to cycloheptenes. *Org Lett* **5**, 325-328 (2003).
- 34 Brooks, G. *et al.* Pleuromutilins. Part 1. The identification of novel mutilin 14-carbamates. *Bioorg Med Chem* **9**, 1221-1231 (2001).
- 35 Bacque, E., Pautrat, F. & Zard, S. Z. A flexible strategy for the divergent modification of pleuromutilin. *Chem Comm (Camb)* **20**, 2312-2313 (2002).
- 36 Pearson, N. D. *et al.* Design and synthesis of conformationally restricted eight-membered ring diketones as potential serine protease inhibitors. *Bioorg Med Chem Lett* **12**, 2359-2362 (2002).
- 37 Schlünzen, F., Pyetan, E., Fucini, P., Yonath, A. & Harms, J. M. Inhibition of peptide bond formation by pleuromutilins: the structure of the 50S ribosomal subunit from *Deinococcus radiodurans* in complex with tiamulin. *Mol Microbiol* **54**, 1287-1294 (2004).
- 38 IARC. (World Health Organization, Geneva, 2008).
- 39 Re, M. D., Michelucci, A., Simi, P. & Danesi, R. Pharmacogenetics of anti-estrogen treatment of breast cancer. *Cancer Treat Rev* **38**, 442-450 (2012).
- 40 Saladores, P. H., Precht, J. C., Schroth, W., Brauch, H. & Schwab, M. Impact of metabolizing enzymes on drug response of endocrine therapy in breast cancer. *Expert Rev Mol Diagn* **13**, 349-365 (2013).
- 41 Early Breast Cancer Trialists' Collaborative Group: Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. *Lancet* **339**, 1-15, 71-85 (1992).
- 42 Buzdar, A. U. & Hortobagyi, G. N. Tamoxifen and toremifene in breast cancer: Comparison of safety and efficacy *J Clin Oncol* **16**, 348-353 (1998).
- 43 Bourassa, P. *et al.* Locating the binding sites of anticancer tamoxifen and its metabolites 4-hydroxytamoxifen and endoxifen on bovine serum albumin. *Eur J Med Chem* **46**, 4334-4353 (2011).
- 44 Cato, A., Nestl, A. & Mink, S. Rapid actions of steroids receptors in cellular signalling pathways *Sci STKE* **138**, Re9 (2002).
- 45 Silva, E., Kabil, A. & Kortenkamp, A. Cross-talk between non-genomic and genomic signalling pathways-distinct effect profiles of environmental estrogens *Toxicol Appl Pharmacol* **245**, 160-170 (2010).
- 46 Vogel, C. L. Phase II and III clinical trials of toremifene for metastatic breast cancer. *Oncology (Williston Park)* **12**, 9-13 (1998).
- 47 Kangas, L. Introduction to toremifene. *Breast Cancer Res Treat* **16**, S3-S7 (1990).
- 48 Howell, A., Downey, S. & Anderson, E. New endocrine therapies for breast cancer. *Eur J Cancer* **32A**, 576-588 (1996).
- 49 Hard, G. C., Iatropoulos, M. J., Jordan, K. & al, e. Major difference in the hepatocarcinogenicity and DNA adduct forming ability between toremifene and tamoxifen in female Crl:CD(BR) rats. *Cancer Res* **53**, 4534-4541 (1993).
- 50 Kisanga, E. R. *et al.* Tamoxifen and metabolite concentrations in serum and breast cancer tissue during three dose regimens in a randomized preoperative trial. *Clin Cancer Res* **10**, 2336-2343 (2004).
- 51 Furr, B. J. & Jordan, V. C. The pharmacology and clinical uses of tamoxifen. *Pharmacol Ther* **25**, 127-205 (1984).
- 52 Jin, Y. *et al.* CYP2D6 genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment. *J Natl Cancer Inst* **97**, 30-39 (2005).

- 53 Xue, Y., Wilson, D., Zhao, L., Liu, H.-W. & Sherman, D. H. Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the pikC-encoded cytochrome P450 in *Streptomyces venezuelae*. *Chem. Biol.* **5**, 661-667 (1998).
- 54 Xue, Y., Zhao, L., Liu, H.-W. & Sherman, D. H. A gene cluster for macrolide antibiotic biosynthesis in *Streptomyces venezuelae*: Architecture of metabolic diversity. *Proc. Natl. Acad. Sci. USA* **95**, 12111-12116 (1998).
- 55 Li, S., Ouellet, H., Sherman, D. H. & Podust, L. M. Analysis of transient and catalytic desosamine-binding pockets in cytochrome P450 PikC from *Streptomyces venezuelae*. *J. Biol. Chem.* **284**, 5723-5730 (2009).
- 56 Anzai, Y. *et al.* Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in biosynthesis of mycinamicin macrolide antibiotics. *Chem. Biol.* **15**, 950-959 (2008).
- 57 Chen, H. *et al.* Expression, purification, and characterization of two *N,N*-dimethyltransferases, TylM1 and DesVI, involved in the biosynthesis of mycaminoside and desosamine. *Biochemistry* **41**, 9165-9183 (2002).
- 58 Chen, M. S. & White, M. C. A predictably selective aliphatic C-H oxidation reaction for complex molecule synthesis. *Science* **318**, 783-787 (2007).
- 59 Chen, M. S. & White, M. C. Combined effects on selectivity in Fe-catalyzed methylene oxidation. *Science* **327**, 566-571 (2010).
- 60 Lyons, T. W. & Sanford, M. S. Palladium-catalyzed ligand-directed C-H functionalization reactions. *Chem. Rev.* **110**, 1147-1169 (2010).
- 61 Wang, D. H., Engle, K. M., Shi, B. F. & Yu, J. Q. Ligand-enabled reactivity and selectivity in a synthetically versatile aryl C-H olefination. *Science* **327**, 315-319 (2010).
- 62 Simmons, E. M. & Hartwig, J. F. Catalytic functionalization of unactivated primary C-H bonds directed by an alcohol. *Nature* **481**, 70-73 (2012).
- 63 Leow, D., Li, G., Mei, T.-S. & Yu, J.-Q. Activation of remote meta-C-H bonds assisted by an end-on template. *Nature* **486**, 518-522 (2012).
- 64 Das, S., Incarvito, C. D., Crabtree, R. H. & Brudvig, G. W. Molecular recognition in the selective oxygenation of saturated C-H bonds by a dimanganese catalyst. *Science* **312**, 1941-1943 (2006).
- 65 Das, S., Brudvig, G. W. & Crabtree, R. H. High turnover remote catalytic oxygenation of alkyl groups: How steric exclusion of unbound substrate contributes to high molecular recognition selectivity. *J. Am. Chem. Soc.* **130**, 1628-1637 (2008).
- 66 Li, S., Podust, L. M. & Sherman, D. H. Engineering and analysis of self-sufficient biosynthetic cytochrome P450 PikC fused to the RhFRED reductase domain. *Journal of the American Chemical Society* **129**, 12940-12941 (2007).

Chapter 4

Functional Analysis of Engineered P450s TyII-RhFRED and TyIHI-RhFRED Involved in the Biosynthesis of Tylosin

4.1 Introduction

Cytochrome P450 monooxygenases (P450s or CYPs) are a family of heme proteins that catalyze a wide range of oxidative transformations in all domains of life and are well-known for their roles in xenobiotic metabolism, human steroid metabolism, and the biosynthesis of antibiotics and signaling molecules.^{1,2} These remarkable enzymes are able to catalyze difficult synthetically relevant reactions on chemically distinct substrates with a high degree of regio- and stereoselectivity. Among the myriad reactions catalyzed by P450s are aliphatic and aromatic hydrocarbon hydroxylation; alkene and arene epoxidations; *O*-, *N*-, and *S*-dealkylation; oxidative amination, dehalogenation, and decarboxylation; dehydrogenation and dehydration; NO, *N*-oxide, and epoxide reduction; reductive dehalogenation and denitration; and isomerization and ring formation, contraction, and expansion.^{3,4} Furthermore, recent bacterial genome sequencing efforts have revealed an unprecedented number of genes encoding for P450s. For example, the actinomycete *Streptomyces coelicolor* A3(2) that produces actinorhodin and undecylprodigiosin was found to have 18 distinct P450s genes, while *Streptomyces avermitilis* MA-4680, the avermectin producer, contains 33 P450s and *Saccharopolyspora erythraea* NRRL 23338, the erythromycin producer, encodes 36 P450s.⁵⁻⁷

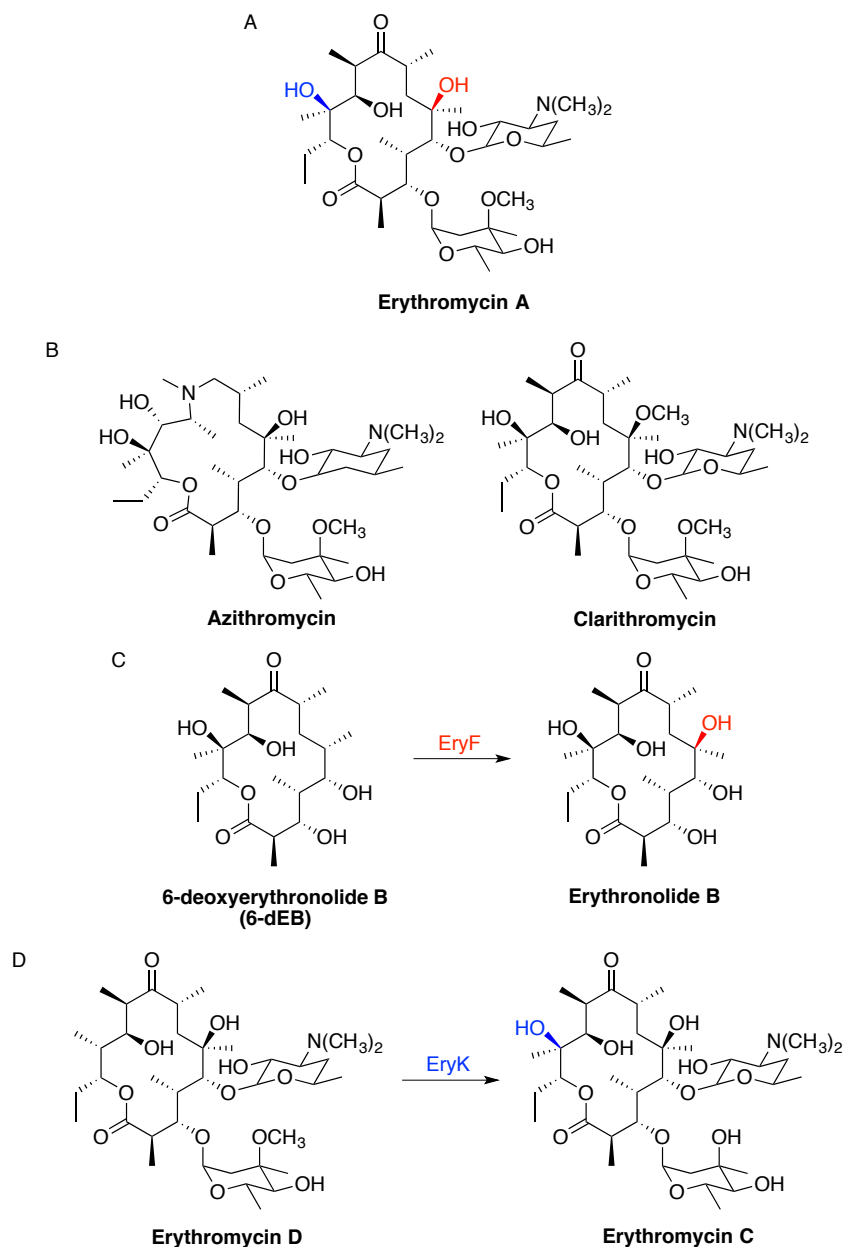


Figure 4-1. Structures of erythromycin A and semisynthetic erythromycin analogs and the oxidation patterns of EryF (red) and EryK (blue). (A) Erythromycin A (B) Azithromycin and clarithromycin (C) Enzymatic oxidation catalyzed by EryF on 6-deoxyerythronolide B (D) Enzymatic oxidation catalyzed by EryK on erythromycin D.

Many P450s have been implicated in secondary metabolic pathways to produce compounds such as macrolides, but only a relatively small number of them have been studied. These compounds are attractive secondary metabolites given their antibacterial, antifungal, anti-inflammatory, antiparasitic, antitumor, and/or immunomodulatory activities. P450s catalyze late-stage tailoring, often regio- and stereoselective oxidation that contributes to increased chemical

diversity and improved bioactivity. Erythromycin A, a macrolide antibiotic whose biosynthesis requires 2 P450s, EryK and EryF, is a commercially available bacteriostatic drug that selectively inhibits bacterial protein synthesis via reversible binding to the large ribosomal subunit (**Fig. 4-1A**).⁸⁻¹⁰ Specifically, EryF hydroxylates the macrolactone precursor 6-deoxyerythronolide B whereas EryK is a macrolide hydroxylase resulting in the formation of erythromycin D (**Fig. 4-1C** and **Fig. 4-1D**). As prototypic P450 hydroxylases, EryF and EryK exhibit strict substrate specificity, tolerating only minor changes in substrate structure.¹¹ In contrast, the pikromycin P450, PikC, which is involved in the biosynthesis of methymycin/neomethymycin and pikromycin in *Streptomyces venezuelae* demonstrates broad substrate promiscuity (**Fig. 4-2**).^{12,13} Related compounds to erythromycin are clarithromycin and azithromycin (**Fig. 4-1B**), which were developed as semisynthetic strategies to combat antibiotic resistance given their broad spectrum of efficacy and improved pharmacokinetic profiles. Nevertheless, given the difficulty of accessing similar macrolides, bacterial biosynthetic P450s represent another tool for the production of novel macrolides.

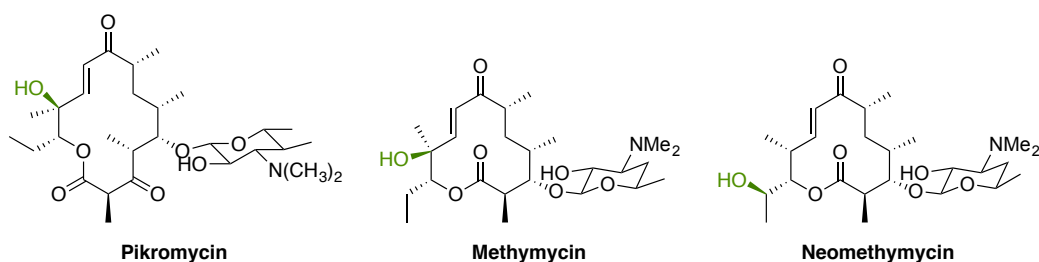


Figure 4-2. Structures of pikromycin, methymycin, and neomethymycin with the sites of PikC hydroxylation highlighted in green.

Tylosin is commercially produced by *Streptomyces fradiae*, but is also produced by *Streptomyces rimosus* and *Streptomyces hygroscopicus*.¹⁴⁻¹⁷ Tylosin is composed of a 16-member branched lactone (tylonolide) and three sugars (mycarose, mycaminose, and mycinose) (**Fig. 4-3**) and that is related to a series of other natural products of juvenimycin and rosamicin.¹⁸⁻²³ Tylosin is currently used as a veterinary medication and as a feed supplement for farm animals due to its growth promoting qualities.²⁴ The seminal studies that first located the tylosin-biosynthetic gene cluster within the *S. fradiae* genome were conducted at Lilly Research Laboratories, Indianapolis and involved the purification and partial sequence analysis of the terminal enzyme of tylosin biosynthesis macrocin *O*-methyltransferase (MOMT) (data not published). Identification allowed for the synthesis of deoxyoligonucleotide probes designed to recognize the MOMT-encoding gene. The probes were used against a genomic *S. fradiae* DNA

library to identify the structural genes required for tylosin biosynthesis.²⁵⁻²⁷ Portions of the genome were subsequently sequenced by several research groups and gene functions were assigned through database comparisons, *in vitro* enzymatic analysis, complementation of mutants, and targeted gene disruptions.²⁸⁻³¹

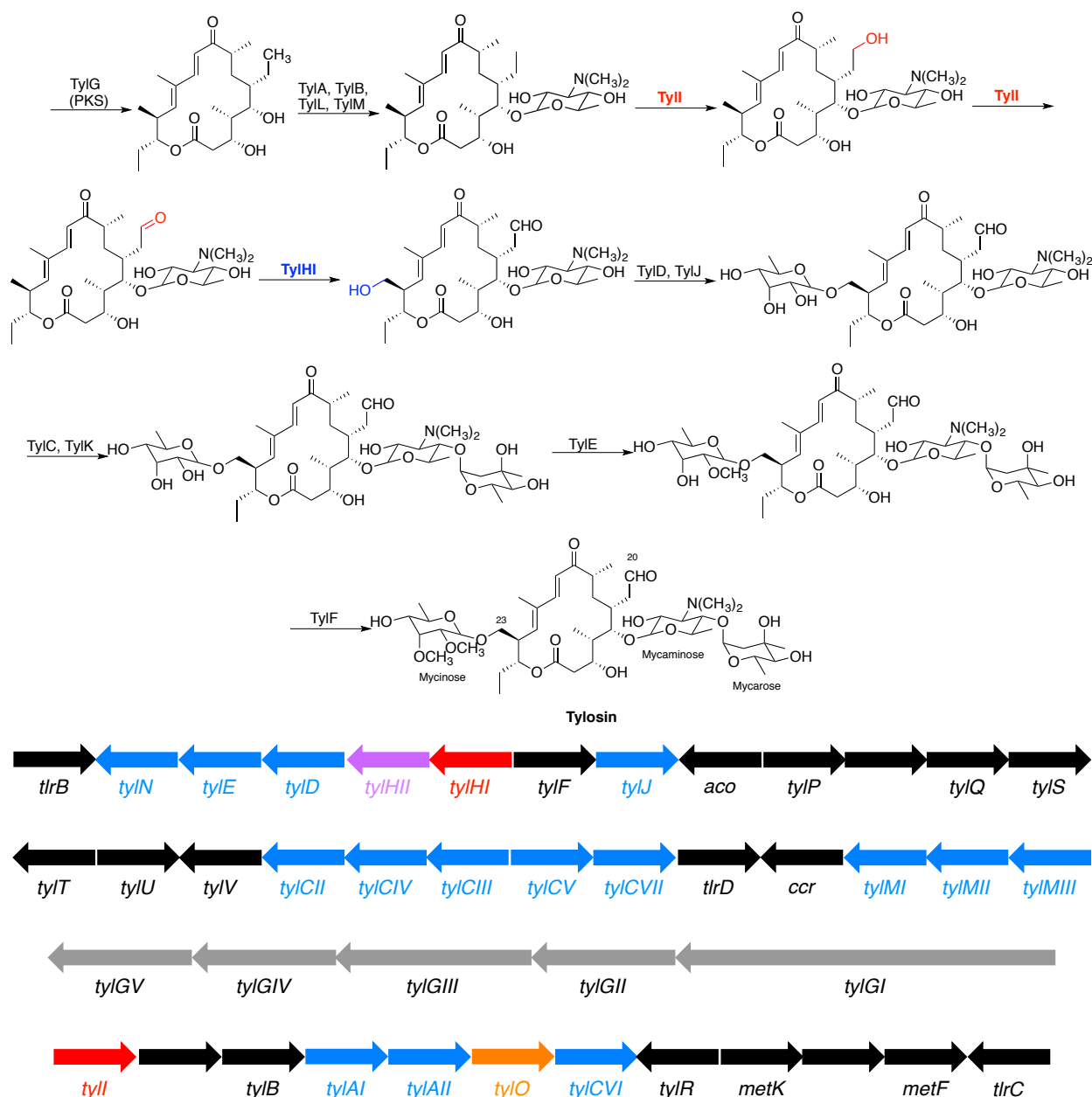


Figure 4-3. The tylosin post-PKS biosynthetic pathway and organization of the tylosin biosynthetic gene cluster. Oxidation steps catalyzed by cytochrome P450 enzymes in the biosynthetic pathway are highlighted in red and blue. Color codes in the *tyl* gene cluster are as follows: red, cytochrome P450 genes; blue, deoxysugar biosynthesis and incorporation genes; purple, ferredoxin gene; blue, deoxysugar biosynthesis and incorporation genes; gray, polyketide synthetase genes; orange, polyketide editing gene; black, mixture of proposed reductases, oxidases, efflux, ancillary, transmethylation, regulatory, and unknown proteins.

Of particular medicinal interest is the C20 aldehyde moiety present in the mature natural product, tylosin. In spite of its rarity in natural products, the aldehyde functionality was demonstrated to be crucial to tylosin's antibiotic activity. Crystallization studies revealed the presence of a definitive, but reversible covalent bond between the ethylaldehyde substituent at the C6 position of the macrolide and the target bacterial ribosome to form a carbinolamine.⁹ Loss of the aldehyde functionality resulted in 100-fold increased minimal inhibitory concentration (MIC) values over unmodified tylosin for a variety of Gram-positive bacteria.³²⁻³⁶ A related 16-membered macrolide discovered in 1972 rosamicin also was discovered to contain an aldehyde, which is crucial to its bioactivity.³⁷ Furthermore, the aldehyde functionality was predicted through fermentation of random *Streptomyces fradiae* mutants to be produced through the sequential action of a single bacterial P450.^{38,39}

Given the looming antibiotic crisis, there is an overwhelming need to identify new pathways towards developing novel macrolide antibiotics. With a thorough comprehension of the biosynthetic machinery that constructs antibiotics such as tylosin, we will be able to develop these enzymes as potential powerful biocatalysts for the production of novel macrolides. *TylII* and *TylHI* exemplify this idea as they are predicted to be oxidative enzymes in the tylosin biosynthetic pathway where *tylI* is predicted to act sequentially in the pathway to produce the final aldehyde functionality. *TylII* has a 71% sequence identity to a previously confirmed P450, *RosC*, from the rosamicin biosynthetic pathway that performed sequential oxidations to produce the aldehyde in the final natural product.^{29,37} Furthermore, *TylHI* has a 48% sequence identity to a P450 from the oleandomycin pathway, *OleP*, and a 55% identity with a P450 from the biosynthesis mycinamicin pathway, *MycCI*.^{40,41} In the current study, *tylII* and *tylHI* genes were overexpressed in *Escherichia coli*, and the functions of purified oxidative *TylII* and *TylHI* enzymes were determined through *in vitro* efforts. Putative substrates obtained from engineered strains of *S. fradiae* that accumulate key tylosin intermediates were employed to characterize the enzymes through biochemical and kinetic studies.

4.2 Results and Discussion

4.2.1 Protein Sequence Analysis of *TylII* and *TylHI*

Comparison of the deduced amino acid sequences of *TylII* and *TylHI* showed relatively low sequence identity (34%), suggesting separate evolutionary sources of the 2 P450s.

Construction of a phylogenetic tree containing select bacterial macrolide P450s, TylII and TylHI were found in distinct branches, confirming the hypothesis that these 2 enzymes may have arisen from discrete ancestral genes as opposed to through divergent evolution from a single parental gene (**Fig. 4-4**). TylII clustered with EryK, responsible for hydroxylation of macrolactone erythromycin D, an intermediate of the erythromycin biosynthetic pathway, while TylHI clustered with MycCI and ChmHI, both reported to oxidize a methyl group on 16-membered macrolides. While the phylogenetic tree accurately predicted the functionality of the P450s (epoxidases vs. hydroxylases) and also correlated the correct substrate size with the acting P450, there was some inconsistency between the predicted site of action of TylII and the known site of EryK (methyl vs. macrolactone site). Nevertheless, the phylogenetic tree suggested that evolutionary selection of secondary metabolic P450s was largely based on product structure dictated by the upstream PKS biosynthetic machinery.

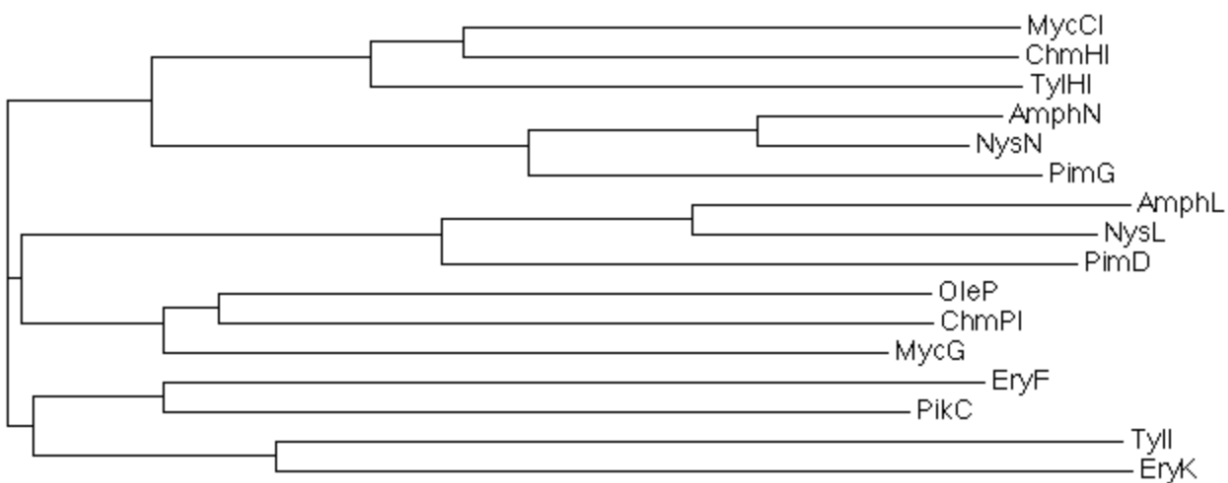


Figure 4-4. Phylogenetic tree of select macrolide biosynthetic P450s. The particular P450s include MycCI and MycG (mycinamicin biosynthesis), ChmHI and ChmPI (chalconic acid biosynthesis), TylII and TylHI (tylosin biosynthesis), AmphN and AmphL (amphotericin biosynthesis), NysN and NysL (nystatin biosynthesis), PimG and PimD (pimaricin biosynthesis), OleP (oleandamycin biosynthesis), EryF and EryK (erythromycin biosynthesis), and PikC (pikromycin biosynthesis).

4.2.2 Heterologous Expression of TylII- and TylHI-RhFRED

The *tylII-RhFRED* and *tylHI-RhFRED* genes were overexpressed in *E. coli* BL21(DE3) and BL21(DE3) pRARE, respectively, and the resulting proteins were purified via affinity chromatography. After purification using Fast Protein Liquid Chromatography (FPLC) in conjunction with a nickel affinity chromatography column, the individual polypeptides showed

molecular weights of approximately 80 kDa each, corresponding to the estimated masses of N-terminal 6× His tagged TyII- and TyIHI-RhFRED. Expression of both genes was placed under the control of a T7 promoter including a C-terminal His₆ affinity tag. Both proteins were purified to homogeneity and CO-bound reduced difference spectra were obtained to confirm that both enzymes were P450s (**Fig. S4-1**).

4.2.3 Functional Analysis of TyII-RhFRED *In Vitro*

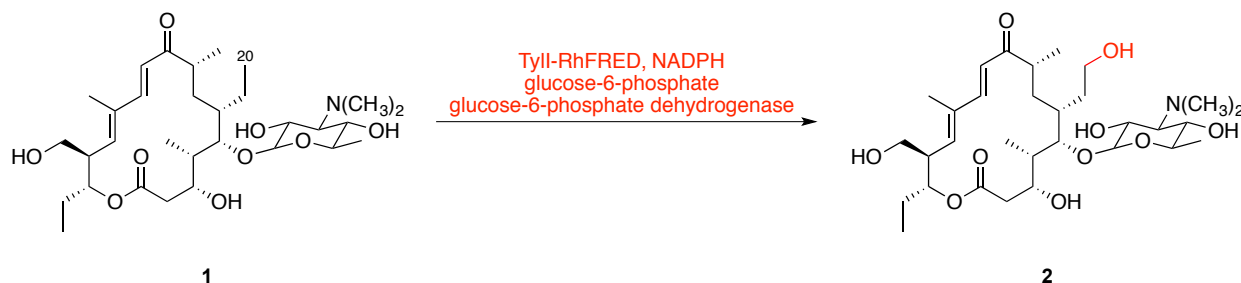


Figure 4-5. First *tyII* catalyzed reaction. In addition to TyII-RhFRED and the appropriate substrate, NADPH and the NADPH regeneration system of glucose-6-phosphate and glucose-6-phosphate dehydrogenase was used to oxidize **1** to its hydroxylated counterpart, **2**.

To establish the role of each tylosin P450, we exploited a previously used covalently attached redox partner derived from *Rhodococcus* sp. NCIMB 9784 found to be naturally fused to a novel FMN/Fe₂S₂ containing reductase partner (**Fig. 4-5**) attached to the P450, TyII.⁴²⁻⁴⁴ The N-terminal His₆-tagged recombinant TyII-RhFRED was heterologously overexpressed and purified from *E. coli* to yield an orange-red enzyme solution that was characterized by ultraviolet-visible (UV-vis) spectroscopy using standard techniques.⁴⁵ The enzyme solution displayed an absorbance peak at 420 nm, with a peak at 450 nm emerging after reduction using sodium dithionite and bubbling of gaseous CO through the solution.

We first tested whether TyII was capable of hydroxylation the predicted substrate **1** at the C20 position. The recombinant enzyme, TyII-RhFRED, was co-incubated with the putative substrate **1** in reaction buffer, the NADPH regeneration system of glucose-6-phosphate and glucose-6-phosphate dehydrogenase, and NADPH overnight. Liquid chromatography-mass spectrometry quadrupole time-of-flight (LC-MS Q-TOF) analysis of the reaction supernatant demonstrated the successful conversion of **1** to a single monohydroxylated product in a greater than 90% yield. Through preparative-scale enzymatic reactions, **2** was enzymatically prepared and was subsequently purified via reverse-phase high-pressure liquid chromatography (RP-

HPLC). Rigorous 1- and 2-dimensional nuclear magnetic resonance (NMR) spectroscopy confirmed the identity of **2** as the C20 hydroxylated counterpart of **1**, confirming the hypothesized first reaction catalyzed by TyII (Fig. 4-5).

Notably, TyII used in conjunction with the commercially available heterologous spinach ferredoxin and ferredoxin reductase system also demonstrated a comparable product profile as the chimeric TyII-RhFRED (data not shown). Nevertheless, the overall conversion was compromised when utilizing the spinach system presumably due to a lower effective concentration of P450 in relation to redox partner than the covalently attached P450-RhFRED system. Additionally, TyII-RhFRED was tested against the putative substrate of TyIHI-RhFRED (**4**) to simultaneously elucidate the proposed biosynthetic order of TyII and TyIHI and to simultaneously probe the enzyme's inherent substrate promiscuity. However, analysis of the reaction mixture following extraction of an overnight reaction between the chimera and **4** revealed unsuccessful oxidation of this compound and starting material was quantitatively recovered.

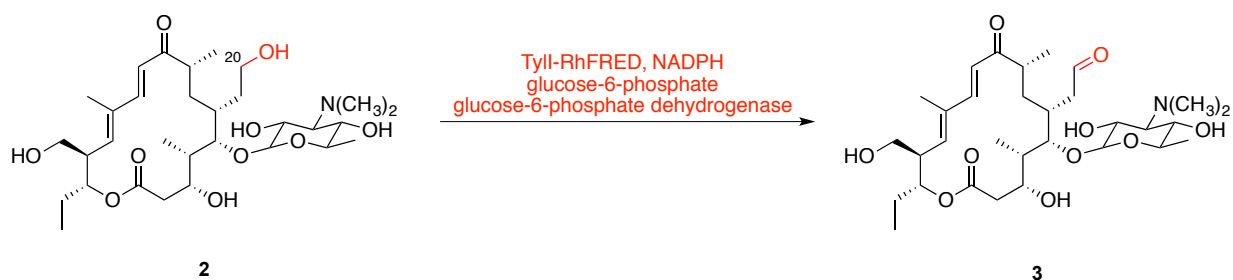


Figure 4-6. Second *tyII* catalyzed reaction. In addition to TyII-RhFRED and the appropriate substrate, NADPH and the NADPH regeneration system of glucose-6-phosphate and glucose-6-phosphate dehydrogenase was used to oxidize **2** to its hydroxylated counterpart **3**.

As a preliminary experiment to determine whether the aldehyde found in **3** could be produced by TyII, a small-scale enzymatic reaction was prepared using **1** as the substrate. Following an overnight incubation using NADPH and the NADPH regeneration system, the reaction mixture was extracted to remove contaminating protein and salt, and the organic layer was dried under nitrogen to obtain a mixture of compounds **1** and **2**. The mixture was then co-incubated with fresh TyII-RhFRED enzyme to test if TyII could accept **2** to install the C20 aldehyde moiety found in the fully matured natural product. After an overnight incubation, there was the emergence of a small additional peak according to the LC-MS Q-TOF data consistent

with production of the aldehyde-containing product **3** (data not shown) in a 2% yield. Given the poor yield, structural elucidation of **3** and biochemical characterization could not be performed and substrate dissociation constants could not be determined.

The poor yield of this second TyII reaction was attributed to the nonnative nature of compounds **1** and **2** as substrates for TyII as neither **1** nor **2** were proposed to be the endogenous substrates for TyII in the hypothesized biosynthetic pathway. Comparison of the structure of **1** to the proper initial biosynthetic intermediate that served to be TyII's first substrate demonstrated that the endogenous substrate lacked the additional hydroxyl group at C23. The presence of this additional functionality could have exacerbated TyII's preference against **2** through suboptimal substrate binding and orientation. Alternatively, the poor conversion could also have been attributed the presence of an undiscovered P450 that is responsible for the aldehyde installation.

4.2.4 Functional Analysis of TyIHI-RhFRED *In Vitro*

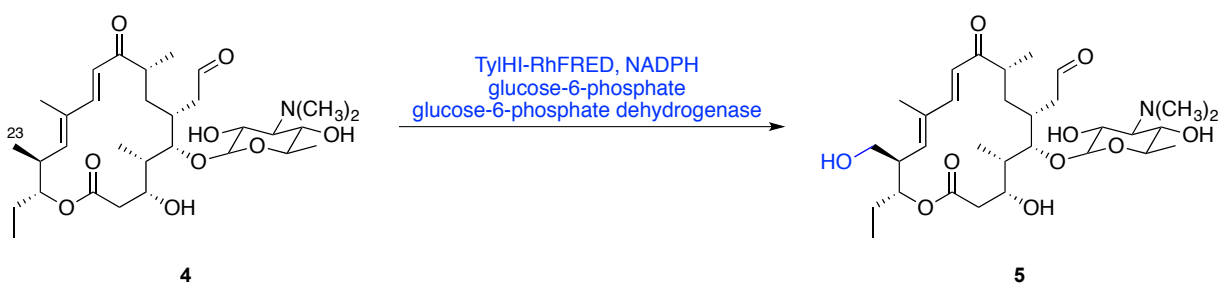


Figure 4-7. Hypothesized reaction catalyzed by *tyIHI*. In addition to TyIHI-RhFRED and the appropriate substrate, NADPH and the NADPH regeneration system of glucose-6-phosphate and glucose-6-phosphate dehydrogenase were used to oxidize **4** to its hydroxylated counterpart **5**.

As in the case of TyII, we utilized a previously used covalently attached redox partner derived from *Rhodococcus* sp. NCIMB 9784 (**Fig. 4-7**) attached to TyIHI for *in vitro* reactions.⁴²⁻⁴⁴ The N-terminal His₆-tagged recombinant TyIHI-RhFRED was heterologously overexpressed and purified from *E. coli* to yield an orange-red enzyme solution that was characterized by UV-vis spectroscopy using standard techniques.⁴⁵ The enzyme solution displayed an absorbance peak at 420 nm, with a peak at 450 nm emerging after reduction using sodium dithionite and bubbling gaseous CO through the solution.

We first tested whether TyIHI was capable of hydroxylating the predicted substrate **4**. To confirm the oxidative capabilities of TyIHI, the chimera TyIHI-RhFRED was incubated with **4** in

the presence of NADPH in protein storage buffer and the NADPH regeneration system. After an overnight reaction at 30°C, a greater than 80% yield of a single monohydroxylated product was observed with a single dose of enzyme using LC-MS Q-TOF.

Notably, TyIHI-RhFRED was unsuccessful at oxidizing **1** and **2**, allowing for the compounds to be recovered in quantitative yield. These results highlighted the ability of the P450 to differentiate between potential substrates, which is extremely important given the potential substrate pool available *in vivo*.

4.2.5 Measurement of Substrate Dissociation Constants

To understand how the tylosin P450 enzymes differentiated between potential substrates with subtle structural differences, we performed spectrophotometric substrate binding studies to determine dissociation constants (K_d) of the putative P450 substrates. As shown in Table __, **1** bound to TyII-RhFRED with a K_d of $482 \pm 88 \mu\text{M}$, suggesting a modest binding ability of the substrate to the active site. This was unsurprising given the proposed biosynthetic pathway that hypothesized that the native TyII substrate was unoxidized at C23. Nevertheless, the binding and conversion observed underscored TyII's impressive capabilities and suggested an underlying substrate promiscuity not often seen in similar bacterial biosynthetic P450s.

Compound **4** bound to TyIHI-RhFRED with a K_d of $122 \pm 15 \mu\text{M}$, suggesting moderately tight binding of the substrate in the enzymatic chimera. This was in agreement with the relatively high conversion observed in an overnight incubation of the enzyme with the substrate. Simultaneously, substrate binding data for **4** with TyII-RhFRED could not be fitted to the Michaelis-Menten equation, suggesting negligible binding of the compound to the TyII-RhFRED active site. This was in agreement with no conversion being observed upon TyII-RhFRED incubation with the hypothesized substrate for TyIHI.

4.2.6 Steady-State Kinetic Analysis of TyII- and TyIHI-RhFRED

Assuming a 1:1 stoichiometric relationship between NADPH consumption and substrate oxidation, the steady-state kinetic parameters for both TyII- and TyIHI-RhFRED were determined using purified **1** and **4**, respectively. In conjunction with the exogenous redox partner RhFRED, TyII demonstrated a K_m of $37.8 \pm 12.3 \mu\text{M}$ and k_{cat} of $20.9 \pm 1.7 \text{ min}^{-1}$ for C20 methyl group hydroxylation.

Following the complete kinetic analyses of TyII-RhFRED with both compounds **1** and **2**, the steady-state kinetic parameters were determined for TyIHI-RhFRED using purified compound **4**. Under the same stoichiometric relationship between NADPH consumption and substrate oxidation as previously used, the K_m was determined to be $152 \pm 87 \mu\text{M}$ and the k_{cat} was found to be $5.55 \pm 1.54 \text{ min}^{-1}$ for the C23 hydroxylation of the methyl group.

Enzyme	Substrate	$K_d (\mu\text{M})$	$K_m (\mu\text{M})$	$k_{\text{cat}} (\text{min}^{-1})$	$k_{\text{cat}}/K_m (\mu\text{M}^{-1} \text{min}^{-1})$
TyII-RhFRED	1	482 ± 88	37.8 ± 12.3	20.9 ± 1.7	0.553 ± 0.135
TyIHI-RhFRED	4	122 ± 15	152 ± 87	5.55 ± 1.54	0.0365 ± 0.0177

Table 4-1. Binding and steady-state kinetic analyses of TyII- and TyIHI-RhFRED.

4.3 Conclusion

Tylosin is a veterinary 16-membered macrolide antibiotic isolated from fermentation of *Streptomyces fradiae* and sold by Eli Lilly and Company. The majority of structural diversity is derived from post-PKS tailoring modifications achieved through a series of glycosylation, oxidation, and methylation reactions. Therefore, this biosynthetic pathway offers an invaluable system to explore the mechanism and significance of secondary metabolite diversification.

In the current study, we provide a detailed study of 2 P450s derived from the tylosin pathway, TyII and TyIHI, and their indispensable participation in late-stage tailoring modifications. Through gene cloning, protein expression, purification, and *in vitro* enzymatic reactions with a foreign covalently attached redox partner, we unambiguously determined their physiological roles as hydroxylases of the C20 and C23 positions of tylosin biosynthetic intermediates. The use of RhFRED, an exogenous reductase derived from a *Rhodococcus* species, served as a highly efficient redox partner for the electron transfer from NADPH required for enzymatic activity. Furthermore, the hypothesized biosynthetic order of two reactions catalyzed by TyII followed by a hydroxylation catalyzed by TyIHI was confirmed through *in vitro* enzymatic reactions against putative substrates. Previous crystallographic data obtained from PikC and its natural substrates (YC-17 and narbomycin) highlighted the role of a sugar, desosamine, to act as a substrate anchor, responsible for facilitating productive binding and proper positioning of substrate within the enzyme active site. Therefore, we surmise that mycaminose plays a similar role to mediate hydroxylation by TyIHI and TyII, respectively.

The current study demonstrated that TyII is potentially a more flexible enzyme than TyIHI, providing an example of a unique biosynthetic P450 able to catalyze sequential oxidations to ultimately install an aldehyde found in the mature natural product through a hydroxyl intermediate. This work also demonstrated that TyII activity depended on the presence of mycaminose given the lack of binding and subsequent oxidation of the macrolactone, tylactone. Although we cannot determine the precise interactions that take place between the TyII substrate and the active site residues without structural information from either through X-ray crystallography or NMR, the high binding affinity of **1** with TyII indicated that there is a specific active site-binding pocket in the polypeptide to accommodate mycaminose to improve substrate affinity. Evidently, the second oxidation step seems to impart a pronounced effect upon the TyIHI oxidation of **3**, enabling it to be accepted by the second P450.

TyIHI was confirmed to be a C23 hydroxylase that acted prior to attachment of mycaminose through the action of *tylD* and *tylJ*. From genomic sequencing work, there was a potential P450 redox partner identified as TyIHI, which was found to have a highly similar sequence to an identified redox partner from the mycinamicin pathway, MycCII. Therefore, the activity observed using RhFRED could be substantially improved upon if the endogenous redox partner was utilized. However, while cloning of TyIHII was successful, protein overexpression was subsequently unsuccessful and may require additional exploratory efforts.

Bacterial biosynthetic P450s such as TyII and TyIHI offer a unique source of potential biocatalysts given the propensity for expeditious cloning, over-expression in *E. coli*, and rapid purification. With a thorough comprehension of the binding and catalytic properties of P450s, these enzymes have been pushed another step towards practical application. Given the unusual sequential reactions catalyzed by TyII, this enzyme is uniquely suited for further development through rational or random mutagenesis and directed evolution efforts to expand its substrate scope. Through understanding the rules that govern productive substrate binding and conversion of biosynthetic intermediates, we gain a foothold towards understanding P450-mediated selective C-H bond activation for the diversification of natural products that can be applied towards biocatalyst development for the production of pharmaceuticals and fine chemicals.

4.4 Experimental Procedures

Materials

Unless otherwise specified, all chemical reagents were ordered from Sigma-Aldrich. Protein purification used QIAGEN (Valencia, CA) Ni-NTA resin, HiTrap Chelating columns from GE Healthcare (Piscataway, NJ), Millipore (Billerica, MA) Amicon Ultra centrifugal filter, and PD-10 desalting columns from GE Healthcare (Piscataway, NJ). LB broth and TB broth components were purchased from EMD Sciences (Gibbstown, NJ). Preparative-scale enzymatic reactions utilized SPE reservoirs, frits, and caps from Sigma-Aldrich. All tylosin-like compounds were obtained from the fermentation broth of *S. fradiae* GS-77 and GS-76 from Eli Lilly and Company (Indianapolis, IN).

Protein Overexpression and Purification

TyII- and TyIHI-RhFRED overexpression and purification followed previous procedures with minor modifications.^{12,42} For TyII-RhFRED, *E. coli* BL21(DE3) transformants carrying pET28b-*tyII-RhFRED* were grown at 37°C in 1L of TB broth containing thiamine (1 mM), 20% glycerol, 50 µg/ml kanamycin, and a rare salt solution (6750 µg/L FeCl₃, 500 µg/L ZnCl₂, CoCl₂, Na₂MoO₄, 250 µg/L CaCl₂, 465 µg/L CuSO₄, and 125 µg/L H₃BO₃) until OD₆₀₀ reached 0.6-1.0. Then isopropyl β-D-thiogalactoside (IPTG, 0.1 mM) and δ-aminolevulinic acid (1 mM) were added, and the cells were cultured at 18°C overnight. For TyIHI-RhFRED, *E. coli* BL21(DE3) pRARE transformants carrying pET28b-*tyIHI-RhFRED* was grown at 37°C in 1L of TB broth containing thiamine (1 mM), 20% glycerol, 50 µg/ml kanamycin, and a rare salt solution (6750 µg/L FeCl₃, 500 µg/L ZnCl₂, CoCl₂, Na₂MoO₄, 250 µg/L CaCl₂, 465 µg/L CuSO₄, and 125 µg/L H₃BO₃) until OD₆₀₀ reached 0.6-1.0. Then isopropyl β-D-thiogalactoside (IPTG, 0.1 mM) and δ-aminolevulinic acid (1 mM) were added, and the cells were cultured at 15°C overnight. After harvesting the cells by centrifugation, 35 mL of lysis buffer (50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl, 10% glycerol, and 10 mM imidazole) was used to resuspend the cell pellet. Lysis was accomplished on a Model 500 Sonic Dismembrator (ThermoFischer Scientific). The insoluble cell debris was separated by centrifugation (35,000 × g, 30 min at 4°C). The soluble fraction was collected and flowed over a HiTrap Chelating HP column using a ÄKTAFPLC instrument from GE Healthcare Life Sciences. The protein was eluted over a linear gradient of 10 – 500 mM imidazole (50 mM NaH₂PO₄ pH 8.0, 300 mM NaCl, and 10% glycerol). The eluted protein

fractions were concentrated with an Amicon Ultra 4, Ultracel – 30K. Subsequent desalting was attained by buffer exchange using a PD-10 column into storage buffer (50 mM NaH₂PO₄ pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol).

CO-Bound Reduced Difference Spectra

The confirmation of TyII-RhFRED and TyIHI-RhFRED as active P450 enzymes was performed through obtaining the CO-bound reduced difference spectra using a UV-visible spectrophotometer (SpectraMax M5, Molecular Devices). CO was bubbled through a solution of enzyme in storage buffer for approximately 30 sec. The spectra was recorded from 350 – 550 nm and compared to the spectra of the enzyme again in storage buffer reduced by adding several milligrams of sodium dithionite (Na₂S₂O₄). The identical assay was used for functional P450 concentration using the extinction coefficient of 91,000 M⁻¹ cm⁻¹.⁴⁵

Functional Analysis of *In Vitro* Activities of TyII- and TyIHI-RHFRED

The standard conversion was done by adding 5 μM of desalted protein, whose functional concentration was determined by the UV-visible absorption spectrum method, 0.5 mM putative substrate, 0.5 mM NADPH in a total volume of 100 μL of desalting buffer. The reaction lacking protein was used as a negative control. The reaction proceeded for 3 h at 30°C and was terminated by extraction using 3 × 200 μL of CHCl₃. The organic layer was dried under nitrogen and redissolved in 100 μL methanol. The LC-MS analysis of reaction extract was performed on an Agilent Q-TOF HPLC-MS using an XBridge™ C18 3.5 μm 150 mm reverse-phase HPLC column under the following conditions: mobile phase, 10-100% solvent B over 15 min (A = deionized water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid); flow rate: 0.2 mL/min; UV wavelength: 254 and 280 nm.

Spectral Substrate Binding Assay

Spectra substrate binding assays were performed on UV-visible spectrophotometer (SpectraMax M5, Molecular Devices) at room temperature by titrating a DMSO solution of the substrate into a total volume of 1 mL of 1 μM P450 solution in 0.5 μL aliquots. A DMSO only sample was used as a reference. The series of Type I difference spectra was used to deduce ΔA (ΔA_{peak} - ΔA_{trough}).

Data from duplicated experiments were fit to the Michaelis-Menten or Hill equation to obtain the dissociation constant, K_d .

Substrate Purification from Fermentation of *S. fradiae* GS-77 and GS-76

Given the difficulty of synthesizing putative substrates for both TyII- and TyIHI-RhFRED, proposed substrates were isolated from the fermentation broths of two *S. fradiae* mutants generously provided by Eli Lilly and Company.³⁸ Glycerol stocks of GS-77 and GS-76 were streaked out on tryptic soy agar (TSA) to revive the organism and colonies were inoculated in tryptic soy broth (TSB) in preparation for fermentation. After five days of growth, 1L of fermentation media (2% dextrose, 1.5% corn meal, 0.9% fish meal, 0.9% corn gluten, 0.1% NaCl, 0.04% $(\text{NH}_4)_2\text{HPO}_4$, 0.2% CaCO_3 , 3% crude soybean oil) was inoculated with the seed culture and allowed to proceed for an additional five days. The broth was harvested through centrifugation and extracted using ethyl acetate. The organic layers were combined and dried *in vacuo*. After resuspension in methanol, the crude extract was purified using reverse-phase high-performance liquid chromatography (RP-HPLC) to obtain the proposed substrates. Consecutive XBridge™ Prep C18 5 μm OBD™ 19×150 mm and XBridge™ Prep C18 5 μm 10×250 mm columns were used under the following conditions: mobile phase, 10-100% solvent B over 15 min (A = deionized water + 0.1% trifluoroacetic acid, B = acetonitrile + 0.1% trifluoroacetic acid); flow rate: 3.5 mL/min; UV wavelength: 254 and 280 nm.

Steady-State Kinetics of TyII- and TyIHI-RhFRED

The standard reaction conditions contained 0.2 μM of TyII- or TyIHI-RhFRED, 10-500 μM substrate, and 250 μM NADPH in storage buffer. After preincubation at 30°C for 5 min, the reactions with varying substrate concentrations were initiated by adding 5 μL of 50 mM NADPH and removing aliquots at 0, 1, 2, and 3 min for substrate concentrations below 100 μM ; 0, 1, 2, and 5 min between 100-120 μM ; 0, 1, 5, and 10 min for 200 μM ; 0, 5, 10, and 20 min for 300 μM ; and 0, 5, 10, and 30 min for 500 μM . The rate of product formation was monitored via Agilent Q-TOF HPLC-MS. The initial velocities of product formation were deduced from the absorbance curves within the linear range. The results from duplicated experiments were fit to Michaelis-Menten equation to obtain steady-state kinetic parameters.

Preparative-scale Enzymatic Reactions

Preparative-scale enzymatic reactions were conducted in SPE reservoirs with fitted frits and caps in a total volume of 5 mL of protein storage buffer with 5 μM enzyme, 250 μM NADPH, 500 μM substrate, 5 mM glucose-6-phosphate, and 12.4 U glucose-6-phosphate dehydrogenase. After enzymatic oxidation, the reactions were quenched by adding approximately twice the volume of LC-MS grade ethyl acetate and extracted thrice. The organic layer was dried *in vacuo* and the oxidation product was purified by RP-HPLC using a XBridgeTM Prep C18 5 μm 10 \times 250 mm column under the following conditions: mobile phase, 10-100% solvent B over 15 min (A = deionized water + 0.1% trifluoroacetic acid, B = acetonitrile + 0.1% trifluoroacetic acid); flow rate: 3.5 mL/min; UV wavelength: 254 and 280 nm.

4.5 Supplementary Information

Culture maintenance and fermentation. *Streptomyces fradiae* GS-76 and GS-77 generously provided by Eli Lilly and Company were maintained on tryptic soy agar (TSA) plates and as glycerol stocks at -80°C. All culture incubations were at 30°C, and with shaking at 160 rpm for liquid cultures. Seed cultures were 50 mL of tryptic soy broth (TSB) inoculated with a loopful of vegetative cells from plate cultures. Fermentation cultures were shake flasks of 2% dextrose, 1.5% corn meal, 0.9% fish meal, 0.15 sodium chloride, 0.04% ammonium phosphate, 0.2% calcium carbonate, and 3% crude soybean oil and seeded with a 1% inoculum of TSB culture and grown for 5 – 7 days.

Cloning and preparation of recombinant enzymes. The *tylI* and *tylHI* genes were PCR amplified from *Streptomyces fradiae* C373.10 genomic DNA using primers that append *NdeI* and *EcoRI* sites at the 5'- and 3'-termini, respectively, with the stop codon removed. These PCR products were digested with *NdeI/EcoRI* and ligated into pET28b-*pikC-RhFRED* to replace the *pikC* gene previously reported.⁴² The identities of all constructs were verified by DNA Sanger sequencing (University of Michigan DNA Sequencing Core). The confirmed constructs were transformed into *E. coli* BL21(DE3) for pET28b-*tylI-RhFRED* and BL21(DE3) pRARE for pET28b-*tylHI-RhFRED* overexpression hosts and the resulting proteins were purified using nickel affinity chromatography as described below.

Overexpression and purification of TylI-RhFRED and TylHI-RhFRED recombinant protein. The construct pET28b-*tylI-RhFRED* and pET28b-*tylHI-RhFRED* were used to transform *E. coli* BL21(DE3) and BL21(DE3) pRARE, respectively, for protein overexpression. The resulting transformant for pET28b-*tylI-RhFRED* was grown at 37°C in 1 L of TB broth containing thiamine (1 mM), 4% glycerol, and 50 µg/mL kanamycin until OD₆₀₀ reached 0.6-1.0. Then isopropyl β-D-thiogalactoside (IPTG, 0.1 mM) and δ-aminolevulinic acid (1 mM) were added, and the cells were cultured at 18°C overnight. For pET28b-*tylHI-RhFRED*, the resulting transformant was grown at 37°C in 1 L of TB broth containing thiamine (1 mM), 4% glycerol, 50 µg/mL spectinomycin, and 50 µg/mL kanamycin until OD₆₀₀ reached 0.6-1.0. Then isopropyl β-D-thiogalactoside (IPTG, 0.1 mM) and δ-aminolevulinic acid (1 mM) were added, and the cells

were cultured at 15°C overnight. After harvesting the cells by centrifugation (5000 x g, 30 min), the cell pellets were thoroughly resuspended in 35 mL of lysis buffer (50 mM NaH₂PO₄, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10% glycerol). Lysis was achieved on a Model 500 Sonic Dismembrator (ThermoFisher Scientific). The cell lysate was clarified by centrifugation (37,000 x g, 30 min at 4°C). The soluble fraction was collected and passed over a HiTrap Chelating 5 mL column pre-charged with 0.1 M NiSO₄ (GE Healthcare). The column was washed with 50 mL of lysis buffer and the proteins were eluted over a linear imidazole gradient from 10 mM to 500 mM over 50 mL. The eluted protein fractions were combined and concentrated with a Amicon Ultra 4, Ultracel – 50K. Desalting was achieved by buffer exchange into storage buffer (50 mM NaH₂PO₄, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, 10% glycerol) with a PD-10 column (GE Healthcare).

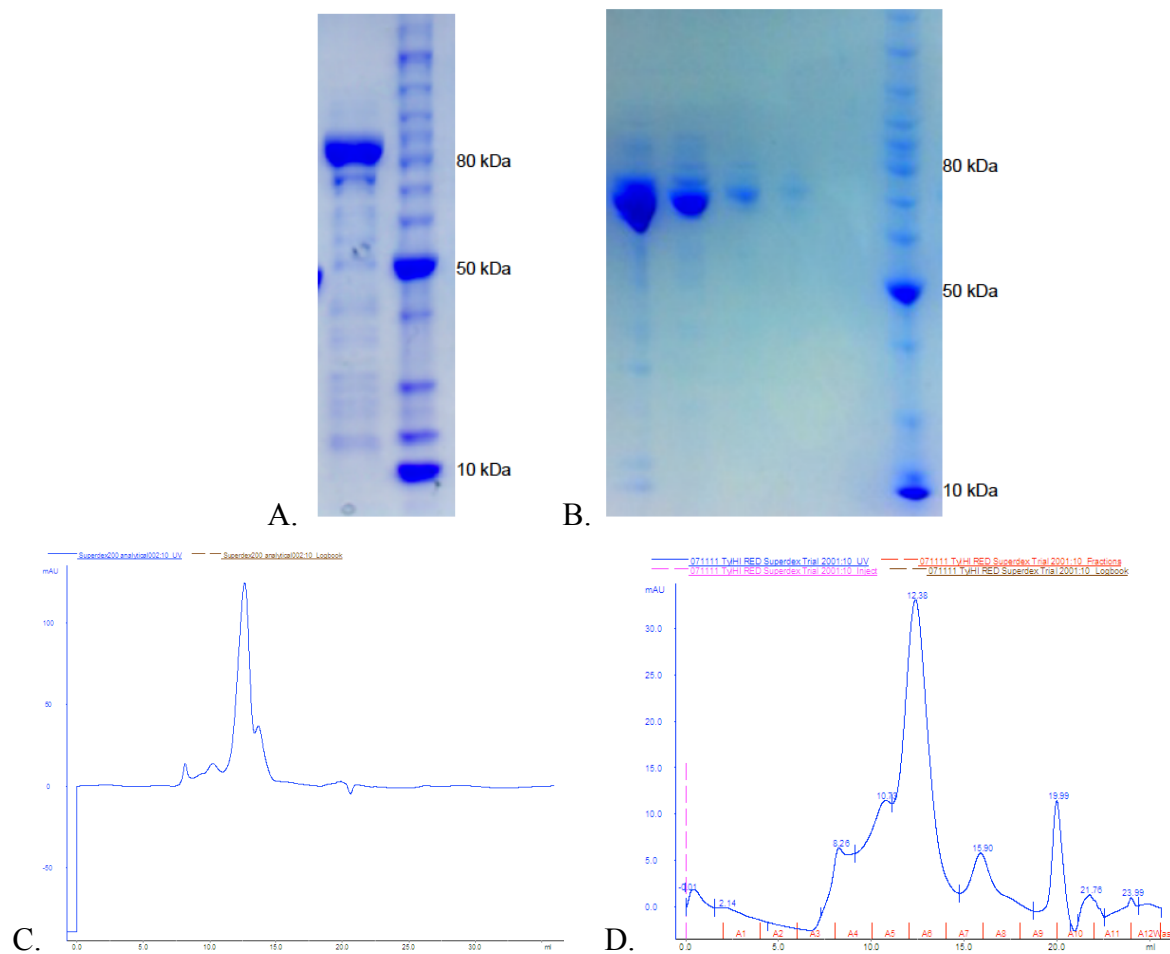


Figure S4-1. Purification of TyII- and TyIHI-RhFRED. (A) SDS-PAGE gel of purified TyII-RhFRED; (B) SDS-PAGE gel purified TyIHI-RhFRED; (C) FPLC gel filtration trace confirming dimerization of TyII-RhFRED in solution; (D) FPLC gel filtration trace confirming dimerization of TyIHI-RhFRED in solution.

CO-bound reduced difference spectra of TyII-RhFRED and TyIHI-RhFRED. Both TyII-RhFRED and TyIHI-RhFRED were identified as functional P450s by obtaining the CO-bound reduced difference spectra using a UV-visible spectrophotometer. Both P450s were reduced by adding several milligrams of sodium dithionite ($\text{Na}_2\text{S}_2\text{O}_4$) and recording the subsequent spectra from 350 to 550 nm. After CO bubbling of the solution for 30 s, the spectrum of CO-bound reduced P450 species was recorded using the previous reduced spectrum as reference. This assay was also utilized to determine the functional P450 concentration using the extinction coefficient of $91,000 \text{ M}^{-1} \text{ cm}^{-1}$.⁴⁵

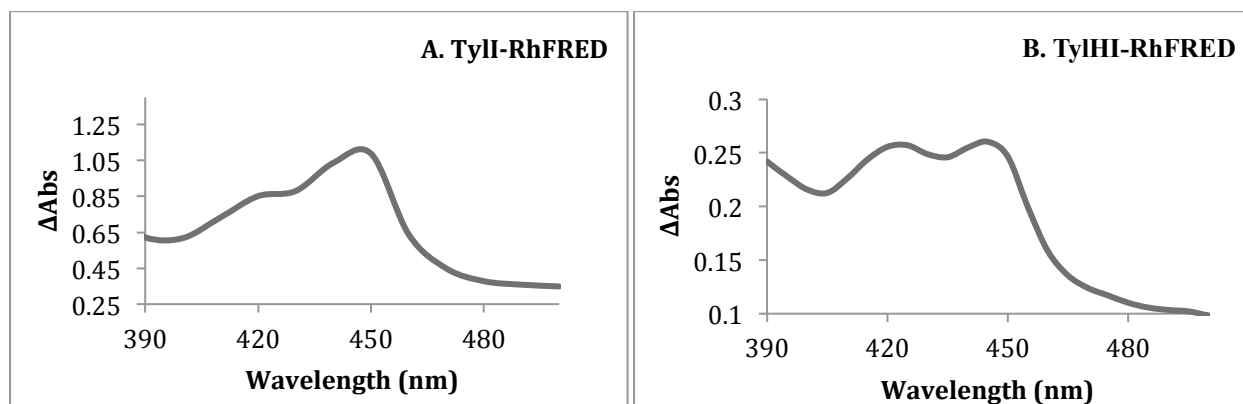


Figure S4-2. CO-bound reduced difference spectra. (A) TyII-RhFRED (B) TyIHI-RhFRED.

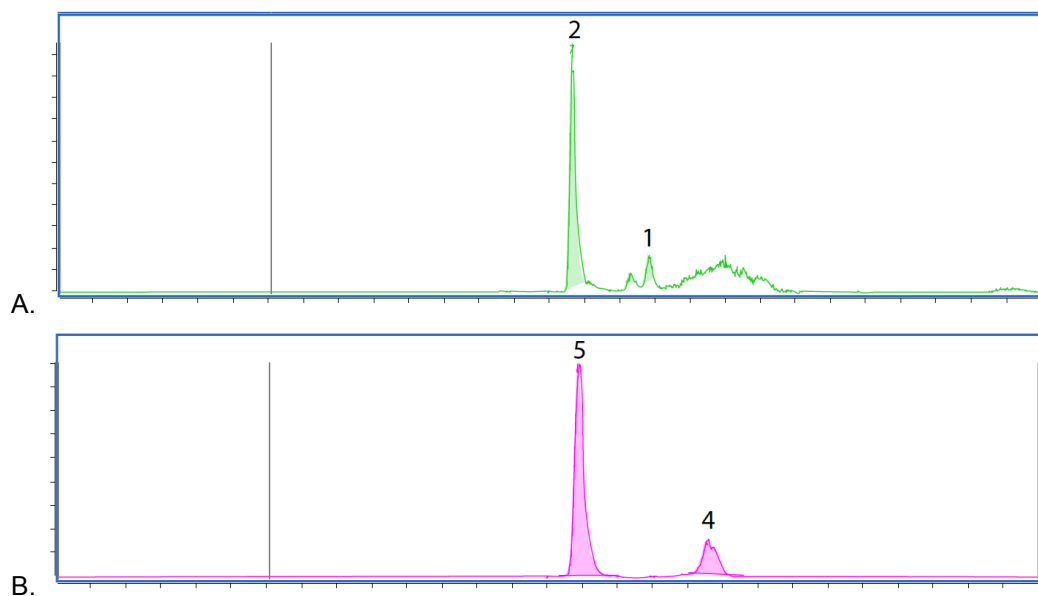


Figure S4-3. Representative LC-MS Q-TOF traces of TylII-RhFRED and TylHI-RhFRED enzymatic reactions. (A) TylII-RhFRED reaction with **1**. LC-MS QTOF trace of an overnight TylII-RhFRED reaction with **1** in the presence of NADPH and the NADPH regeneration system of glucose-6-phosphate and glucose-6-phosphate dehydrogenase. (B) TylHI-RhFRED reaction with **4**. LC-MS QTOF trace of an overnight TylHI-RhFRED reaction with **4** in the presence of NADPH and the NADPH regeneration system of glucose-6-phosphate and glucose-6-phosphate dehydrogenase.

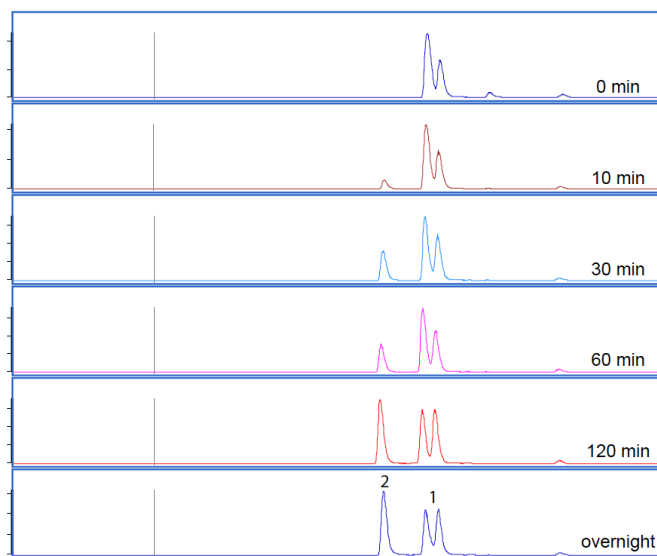


Figure S4-4. Biochemical characterization of TyII-RhFRED. 5 μ M TyII-RhFRED was incubated with 500 μ M **1** in the presence of NADPH with aliquots removed for analysis by LC-MS QTOF at predetermined time points.

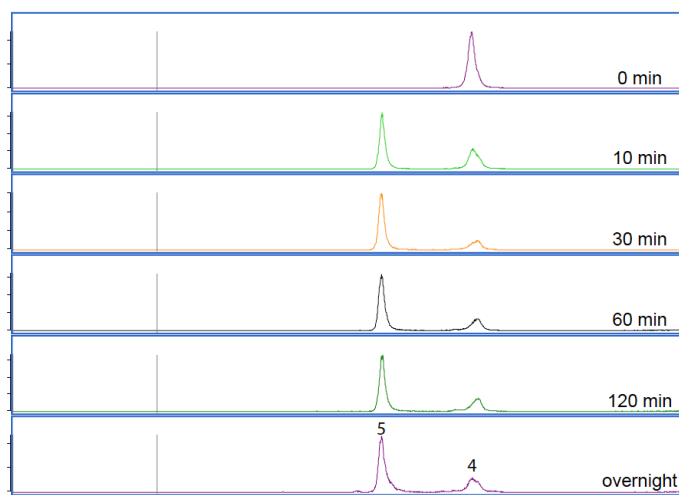
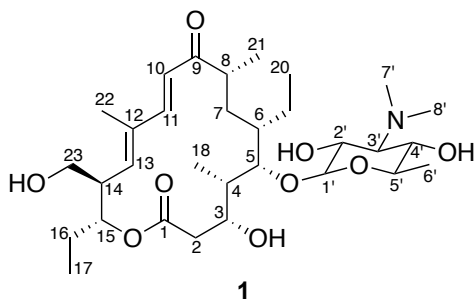
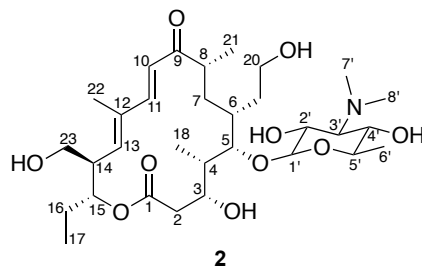


Figure S4-5. Biochemical characterization of TyIHI-RhFRED. 5 μ M TyIHI-RhFRED was incubated with 500 μ M **4** in the presence of NADPH with aliquots removed for analysis by LC-MS QTOF at predetermined time points.



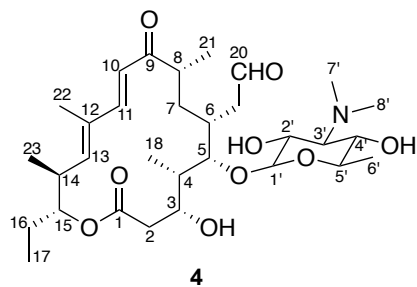
Position	Predicted		Observed	
	¹³ C	¹ H	¹³ C	¹ H
1	172.8	-	173.5	-
2	39.7	2.25, 2.50 (dd)	39.8	2.52
3	70.1	4.03 (ddd)	70.5	4.26 (ddd)
4	43.6	1.60 (m)	44.3	1.59
5	85.6	2.90 (dd)	76.8	2.84
6	41.4	1.40 (m)	46.6	1.43
7	33.1	1.39, 1.14	31.7	1.40
8	40.7	2.36 (m)	33.3	2.45
9	201.8	-	205.6	-
10	119.5	6.33 (d)	120.3	6.44 (dd)
11	147.8	7.40 (d)	148.2	7.26 (dd)
12	134.9	-	135.2	-
13	142.5	5.44 (dd/m)	142.8	5.91 (dd)
14	47.0	2.67 (m)	47.0	2.62
15	75.2	4.50 (m)	74.9	4.94 (m)
16	25.2	1.55 (m)	24.8	1.55
17	10.4	0.89 (t)	11.1	0.89
18	8.9	1.24 (d)	8.63	1.24
19	25.5	1.55 (m)	29.3	1.57
20	12.2	0.99 (m)	11.8	1.00
21	15.7	1.09 (t)	15.9	1.09
22	12.9	1.71 (s)	13.0	1.72
23	62.3	3.38, 3.63 (dd)	61.1	3.39, 3.36 (dd)
1'	110.6	5.40 (d)	103.6	5.11
2'	71.1	3.77 (m)	71.0	3.75
3'	79.5	2.73 (dd)	79.0	2.59
4'	68.2	4.09 (dd)	70.1	3.66
5'	75.8	3.77 (m)	72.7	3.73
6'	16.7	1.26 (d)	16.6	1.27
7'	42.6	2.35 (s)	40.8	3.14 (s)
8'	42.6	2.35 (s)	40.8	3.16 (s)

Table S4-1. ¹³C and ¹H NMR chemical shift tabulation for compound 1.



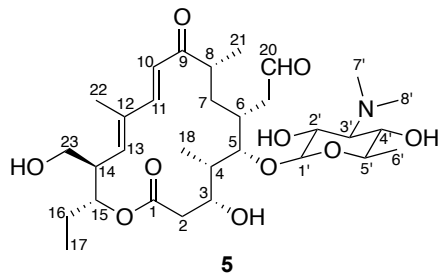
Position	Predicted		Observed	
	¹³ C	¹ H	¹³ C	¹ H
1	173.01	-	162.7	-
2	41.55	2.25, 2.50 (dd)		
3	66.17	4.03 (ddd)	64.33	
4	41.72	1.60 (m)		1.60 (m)
5	84.8	2.90 (dd)	74.3	4.09 (dd)
6	40.38	1.40 (m)		1.18
7	30.38	1.39, 1.14	30.4	1.22
8	40.06	2.36 (m)	33.0	2.66
9	204.06	-	203.2	-
10	124.57	6.33 (d)	119.1	
11	140.46	7.40 (d)	149.6	
12	137.62	-	128.5	-
13	138.72	5.44 (dd/m)	120.7	
14	49.4	2.67 (m)		2.87 (m)
15	78.25	4.50 (m)	73.8	3.85 (m)
16	24.32	1.55 (m)	23.7	1.33 (m)
17	10.17	0.99 (t)	9.58	0.90 (t)
18	10.24	1.24 (d)	10.05	1.24 (d)
19	36.38	1.40 (m)	33.01	1.74 (m)
20	60.94	3.73 (m)	64.41	3.67 (m)
21	15.70	1.09 (t)		
22	14.68	1.71 (s)	14.4	1.87 (s)
23	60.82	3.38, 3.63 (dd)	64.39	3.47, 3.61 (dd)
1'	101.98	5.40 (d)	115.7	5.10
2'	68.69	3.77 (m)	67.01	3.74 (m)
3'	78.98	2.73 (dd)	74.5	2.56 (dd)
4'	71.15	4.09 (dd)	67.1	3.65 (m)
5'	71.15	3.77 (m)	67.1	3.55 (m)
6'	17.16	1.26 (d)	17.6	1.22 (d)
7'	41.91	2.35 (s)	41.8	3.01 (s)
8'	41.91	2.35 (s)	41.8	3.01 (s)

Table S4-2. ¹³C and ¹H NMR chemical shift tabulation for compound 2.



Position	Predicted		Observed	
	¹³ C	¹ H	¹³ C	¹ H
1	172.8	-	175.2	-
2	39.7	2.25, 2.50 (dd)	37.8	2.26, 2.51 (dd)
3	70.1	4.03 (ddd)	70.3	4.06
4	43.3	1.60 (m)	44.9	1.59 (m)
5	85.7	2.90 (dd)	77.0	2.83
6	32.8	1.72 (m)	29.7	1.41
7	32.9	1.39, 1.14	31.9	1.34
8	40.4	2.36 (m)	39.4	2.61
9	201.8	-	203.2	-
10	119.5	6.33 (d)	120.3	6.60 (dd)
11	147.8	7.40 (d)	148.2	7.62 (dd)
12	134.9	-	133.4	-
13	142.5	5.44 (dd/m)	146.0	5.99 (dd)
14	42.3	2.77 (m)	43.8	
15	81.4	4.50 (m)	79.0	4.61
16	24.9	1.55 (m)	24.8	1.56
17	10.4	0.89 (t)	9.65	0.41
18	8.90	0.88 (d)	8.91	1.27 (d)
19	43.1	2.44, 2.19 (m)	38.8	1.48
20	202.2	9.72 (m)	202.9	3.98
21	15.7	1.09 (t)	14.1	
22	12.9	2.12 (s)	12.9	1.82 (s)
23	16.2	0.86	16.2	
1'	110.6	5.40 (d)	100.4	5.11
2'	71.1	3.77 (m)	70.8	3.74 (m)
3'	79.5	2.73 (dd)	78.9	2.51 (dd)
4'	68.2	4.09 (dd)	70.5	3.64 (m)
5'	76.8	3.77 (m)	77.0	3.38
6'	16.7	1.26 (d)	17.5	1.22 (d)
7'	42.6	2.35 (s)	41.7	2.97 (s)
8'	42.6	2.35 (s)	41.7	2.97 (s)

Table S4-3. ¹³C and ¹H NMR chemical shift tabulation for compound 4.



Position	Predicted		Observed	
	¹³ C	¹ H	¹³ C	¹ H
1	172.8	-	172.8	-
2	39.7	2.25, 2.50 (dd)	34.8	2.17, 2.73
3	70.1	4.03 (ddd)	64.0	3.54
4	43.3	1.60 (m)	38.8	1.72
5	85.7	2.90 (dd)	78.1	2.70
6	32.8	1.72 (m)	30.4	1.65
7	32.9	1.39, 1.14	28.1	1.21, 1.45
8	40.4	2.36 (m)	34.6	2.15
9	201.8	-	205.9	-
10	119.5	6.33 (d)	128.3	6.08
11	147.8	7.40 (d)	143.6	7.54
12	134.9	-	130.8	-
13	142.5	5.44 (dd/m)	139.1	5.37
14	47.0	2.67 (m)	42.2	2.35
15	75.2	4.50 (m)	73.5	
16	25.2	1.55 (m)	26.4	1.50
17	10.4	0.89 (t)	14.4	0.95
18	8.90	0.88 (d)	17.6	1.03
19	43.1	2.44, 2.19 (m)	42.1	1.78, 2.11
20	202.2	9.72 (m)	201.5	9.49
21	15.7	1.09 (t)	19.1	1.52
22	12.9	2.12 (s)	18.1	1.95
23	62.3	2.36, 3.38	64.4	3.43, 3.81
1'	110.6	5.40 (d)	101.3	5.21
2'	71.1	3.85 (m)	66.7	3.56
3'	79.5	2.73 (dd)	73.8	2.75
4'	68.2	3.55 (dd)	71.1	4.05
5'	76.8	3.70 (m)	69.1	3.78
6'	16.7	1.11 (d)	23.2	1.26
7'	42.6	2.26 (s)	42.1	3.11
8'	42.6	2.26 (s)	42.1	3.11

Table S4-4. ¹³C and ¹H NMR chemical shift tabulation for compound 5.

4.6 References:

- 1 Coon, M. J. Cytochrome P450: Nature's most versatile biological catalyst. *Ann Rev Pharmacol Toxicol* **45**, 1-25 (2005).
- 2 Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem Res Toxicol* **14**, 611-650 (2001).
- 3 Bernhardt, R. Cytochrome P450s as versatile biocatalysts. *J Biotechnol* **124**, 128-145 (2006).
- 4 Isin, E. M. & Guengerich, F. P. Complex reactions catalyzed by cytochrome P450 enzymes. *Biochimica et Biophysica Acta* **1770**, 314-329 (2007).
- 5 Bentley, S. D. *et al.* Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141-147 (2002).
- 6 Ikeda, H., Nonomiya, T., Usami, M., Ohta, T. & Omura, S. Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. *Proc Natl Acad Sci USA* **96**, 9509-9514 (1999).
- 7 Caffrey, P., Lynch, S., Flood, E., Finnan, S. & Oliynyk, M. Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem Biol* **8**, 713-723 (2001).
- 8 Andersen, J. F. & Hutchinson, C. R. Characterization of *Saccharopolyspora erythraea* cytochrome P-450 genes and enzymes, including 6-deoxyerythronolide B hydroxylase. *J Bacteriol* **174**, 725-735 (1992).
- 9 Hansen, J. L. *et al.* The structures of four macrolide antibiotics bound to the large ribosomal subunit. *Molecular Cell* **10**, 117-128 (2002).
- 10 Stassi, D., Donadio, S., Staver, M. J. & Katz, L. Identification of a *Saccharopolyspora erythraea* Gene Required for the Final Hydroxylation Step in Erythromycin Biosynthesis. *Journal of Bacteriology* **175**, 182-189 (1992).
- 11 Montemiglio, L. C. *et al.* Redirecting P450 EryK specificity by rational site-directed mutagenesis. *Biochemistry* **52**, 3678-3687 (2013).
- 12 Xue, Y., Wilson, D., Zhao, L., Liu, H.-W. & Sherman, D. H. Hydroxylation of macrolactones YC-17 and narbomycin is mediated by the pikC-encoded cytochrome P450 in *Streptomyces venezuelae*. *Chem. Biol.* **5**, 661-667 (1998).
- 13 Li, S. *et al.* Selective oxidation of carbonyl C-H bonds by an engineered macrolide P450 mono-oxygenase. *Proc. Natl. Acad. Sci. USA* **104**, 18463-18468 (2009).
- 14 McGuire, J. M. *et al.* Tylosin, a new antibiotic. I. Microbiological studies. *Antibiot Chemother* **11**, 320-327 (1961).
- 15 Seno, E. T., Pieper, R. L. & Huber, F. M. Terminal stages in the biosynthesis of tylosin. *Antimicrob Agents Chemother* **11**, 455-461 (1977).
- 16 Pape, H. & Brillinger, G. U. Metabolic products of microorganisms. 113. Biosynthesis of thymidine diphospho mycarose in a cell-free system from *Streptomyces rimosus*. *Arch Microbiol* **88**, 25-35 (1973).
- 17 Jensen, A. L., Darken, M. A., Schultz, J. S. & Shay, A. J. Relomycin: flask and tank fermentation studies. *Antimicrob Agents Chemother*, 49-53 (1964).
- 18 Carlson, S. *et al.* Potential chemopreventive activity of a new macrolide antibiotic from a marine-derived *Micromonospora* sp. *Mar Drugs* **11**, 1152-1161 (2013).
- 19 Waitz, J. A., Drube, C. G., Eugene L Moss, J. & Weinstein, M. J. Biological studies with rosamicin, a new *Micromonospora*-produced macrolide antibiotic. *J Antibiot* **25**, 647-652 (1972).
- 20 Kishi, T., Harada, S., Yamana, H. & Miyake, A. Studies on juvenimicin, a new antibiotic. II Isolation, chemical characterization and structures. *J Antibiot* **29**, 1171-1181 (1976).
- 21 Morin, R. B. & Gorman, M. The partial structure of tylosin, a macrolide antibiotic. *Tet. Lett.* **34**, 2339-2345 (1964).
- 22 Morin, R. B., Gorman, M., Hamill, R. L. & Demarco, P. V. The structure of tylosin. *Tet. Lett.* **54**, 4737-4740 (1970).
- 23 Omura, S. & Nakagawa, A. Chemical and biological studies on 16-membered macrolide antibiotics. *J Antibiot* **28**, 401-433 (1975).
- 24 Podojil, M., Blumauerova, M., Culik, K. & Vanek, Z. *The tetracyclines: properties, biosynthesis and fermentation.* (Marcel Dekker, 1984).
- 25 Baltz, R. H. & Seno, E. T. Genetics of *Streptomyces fradiae* and tylosin biosynthesis. *Annu Rev Microbiol* **42**, 547-574 (1988).
- 26 Fishman, S. E. *et al.* Cloning genes for the biosynthesis of a macrolide antibiotic. *Proc Natl Acad Sci USA* **84**, 8248-8252 (1987).

- 27 Birmingham, V. A. *et al.* Cloning and expression of a tylosin resistance gene from a tylosin-producing
strain of *Streptomyces fradiae*. *Mol Gen Genet* **204**, 532-539 (1986).
- 28 Gandecha, A. R., Large, S. L. & Cundliffe, E. Analysis of four tylosin biosynthetic genes from the *tylLM*
region of the *Streptomyces fradiae* genome *Gene* **184**, 197-203 (1997).
- 29 Merson-Davies, L. A. & Cundliffe, E. Analysis of five tylosin biosynthetic genes from the *tylIBA* region of
the *Streptomyces fradiae* genome. *Mol Microbiol* **13**, 349-355 (1994).
- 30 Jr, P. R. R., Reynolds, P. A. & Hershberger, C. L. Homology between proteins controlling *Streptomyces*
fradiae tylosin resistance and ATP-binding transport. *Gene* **102**, 27-32 (1991).
- 31 Gandecha, A. R. & Cundliffe, E. Molecular analysis of *thrD*, an MLS resistance determinant from the
tylosin producer, *Streptomyces fradiae* *Gene* **184**, 197-203 (1996).
- 32 Omura, S., Miyano, K., Matsubara, H. & Nakagawa, A. Novel dimeric derivatives of leucomycins and
tylosin, sixteen membered macrolides. *J Med Chem* **25**, 271-275 (1982).
- 33 Omura, S., Tishler, M., Nakagawa, Y., Hironaka, Y. & Hata, T. Relationship of structures and
microbiological activities of the 16-membered macrolides. *J Med Chem* **15**, 1011-1015 (1972).
- 34 Narandja, A., Suskovic, B., Kelneric, Z. & Djokic, S. Structure-activity relationships among polyhydro
derivatives of tylosin. *J Antibiot (Tokyo)* **47**, 581-587 (1993).
- 35 Kirst, H. A. *et al.* Synthesis and evaluation of tylosin-related macrolides modified at the aldehyde function:
a new series of orally effective antibiotics. *J Med Chem* **31**, 1631-1641 (1988).
- 36 Omura, S. *et al.* Structure-biological activities relationships among leucomycins and their derivatives. *J*
Antibiot (Tokyo) **21**, 532-538 (1968).
- 37 Iizaka, Y. *et al.* Function of cytochrome P450 enzymes RosC and RosD in the biosynthesis of rosamicin
macrolide antibiotic produced by *Micromonospora rosaria*. *Antimicrob Agents Chemother* **57**, 1529-1531
(2013).
- 38 Baltz, R. H. & Seno, E. T. Genetics of *Streptomyces Fradiae* and Tylosin Biosynthesis. *Ann Rev Microbiol*
42, 547-574 (1988).
- 39 Baltz, R. H. & Seno, E. T. Properties of *Streptomyces fradiae* Mutants Blocked in Biosynthesis of
Macrolide Antibiotic Tylosin. *Antimicrobial Agents and Chemotherapy* **20**, 214-225 (1981).
- 40 Anzai, Y. *et al.* Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in
biosynthesis of mycinamicin macrolide antibiotics. *Chem. Biol.* **15**, 950-959 (2008).
- 41 Rodriguez, A. M., Olano, C., Mendez, C., Hutchinson, C. R. & Salas, J. A. A cytochrome P450-like gene
possibly involved in oleandomycin biosynthesis by *Streptomyces antibioticus*. *FEMS Microbiol Lett* **127**,
117-120 (1995).
- 42 Li, S., Podust, L. M. & Sherman, D. H. Engineering and analysis of self-sufficient biosynthetic cytochrome
P450 PikC fused to the RhFRED reductase domain. *Journal of the American Chemical Society* **129**, 12940-
12941 (2007).
- 43 Mot, R. D. & Parret, A. H. A novel class of self-sufficient cytochrome P450 monooxygenases in
prokaryotes. *Trends Microbiol* **10**, 502-508 (2002).
- 44 Roberts, G. A., Grogan, G., Greter, A., Flitsch, S. L. & Turner, N. J. Identification of a new class of
cytochrome P450 from a *Rhodococcus* sp. *J Bacteriol* **184**, 3898-3908 (2002).
- 45 Omura, T. & Sato, R. The carbon monoxide-binding pigment of liver microsomes II. Solubilization,
purification, and properties. *Journal of Biological Chemistry* **239**, 2379-2385 (1964).

Acknowledgements

The authors thank Patricia Cruz Lopez, Carl Averang, and George Chlipala with assistance with obtaining initial NMR spectra on several tylosin-like compounds.

Author Contributions

Karoline C. Chiou, Ashootosh Tripathi, Shengying Li, and David H. Sherman conducted the experimental design.

Karoline C. Chiou, Ashootosh Tripathi, and Shengying Li performed the experimental work.

Karoline C. Chiou performed the *Streptomyces* fermentation, substrate purification, enzyme reactions, kinetic experiments, and preparative-scale enzymatic reactions while Ashootosh Tripathi performed structural elucidation.

Karoline C. Chiou, Ashootosh Tripathi, Shengying Li, and David H. Sherman evaluated the data.

Chapter 5

Future Work

In this Ph.D. dissertational study, I have focused on the investigation of cytochrome P450 monooxygenases as potential biocatalysts for the production of high value pharmaceuticals. First in collaboration with Dr. Larissa M. Podust (University of California, San Francisco) and Dr. John Montgomery (University of Michigan), a substrate-engineering approach elaborated on a previously developed desosamine anchoring hypothesis with PikC, revealing unprecedented flexibility. This has driven PikC towards serving as a model for studying the fundamental factors that influence P450-mediated oxidation, such as substrate binding, orientation, and product formation using an engineered chimeric P450, PikC_{D50N}-RhFRED. The results provided significant insight into the parameters that control P450 oxidation culminating in the development of an optimized linear acyclic linker that efficiently replaced the natively used sugar anchor. These advancements eliminated labor-intensive synthetic steps to acquire and attach the sugar to potential substrates, opening to the door to chemoenzymatic elaboration.

To further expand PikC_{D50N}-RhFRED research, the chimeric enzyme was employed to selectively oxidize several structurally distinct scaffolds with pharmaceutical implications. These compounds included the veterinary drug tiamulin and the anti-cancer agents tamoxifen and toremifene all known to undergo extensive P450-mediated metabolism *in vivo*. These results underscored PikC's flexibility and potential as a biocatalyst, as it can accept very different molecules for oxidation as long as a terminal *N,N*-dimethylamino linker is attached. More importantly, however, was the tangible possibility of using bacterial biosynthetic enzymes such as PikC to increase the number of analogs of the scaffold.

Finally, two P450s from the tylosin biosynthetic pathway were functionally characterized to confirm their identities and determine their biosynthetic order. TyII-RhFRED was especially interesting given its ability to perform sequential oxidations to form an aldehyde crucial to

tylosin's bioactivity. However, this study also leaves a number of unanswered questions for future exploration of these unique P450s.

First, the presence of desosamine or desosamine analogs in diverse secondary metabolite substrates of different biosynthetic P450s suggests that a general sugar anchoring mechanism may be extensively used in nature. MycCI, TylI, TylHI, and EryK are all P450s implicated in the tailoring of macrolides (i.e. mycinamicin, tylosin, and erythromycin) and all four enzymes hydroxylate a sugar-containing substrate (mycinamicin VIII, *O*-mycaminosyltylactone, 23-deoxy-*O*-mycaminosyltylonolide, and erythromycin D, respectively) at positions distal from the sugar substituent. These observations indicate that the oxidative abilities of other bacterial P450s could be channeled through a similar approach as to the substrate-engineering idea pursued against PikC. This methodology could widen the significantly limited substrate scope of biosynthetic P450s and retool chemoenzymatic syntheses toolboxes towards chemical diversification of synthetic, semisynthetic, and natural product molecules of the utmost biological and chemical interest.

Second, although the engineered chimera utilized in the PikC experiments demonstrated regioselective hydroxylation in moderate to high yield of the unnatural substrates containing an optimized linear linker, significant work remains to increase regio- and stereoselectivity and achieve complete conversion. To achieve these goals, the additional enzyme-substrate interaction observed between the linker carbonyl and an active site Histidine reveal that other active site residues are underutilized for increasing substrate binding affinities and promoting selectivity. Rational protein engineering and directed evolution efforts will not only address those concerns, but also provide opportunities to further widen the substrate scope of this impressive enzyme.

Third, we still lack full comprehension of the regioselectivity that PikC exhibits with the natural substrates, YC-17 and narbomycin. Although previous work has solved co-crystal structures with both substrates, the basis of preference for C10 hydroxylation of YC-17 and C12 hydroxylation of narbomycin remains elusive. Although substrate repositioning in the active state and the confirmation of the identity of the oxidative species in the catalytic cycle could offer possible answers, the static crystal structures do not offer satisfactory answers. Therefore, a more dynamic and full picture of protein movement and breathing is required, which may be provided through protein NMR. Through understanding protein movement between open and closed states that often accompany P450 substrate binding, oxidation, and product release, we

may be able to more directly understand P450-mediated oxidation and regio- and stereoselectivity. Other complementary approaches would include testing synthetic metalloporphyrins against endogenous or exogenous substrates and employing selenium-containing versions of PikC to determine the identity of the oxidative species.

Fourth, building on information gained through protein engineering and NMR efforts, additional exploration of alternative linkers capable of targeting substrates to P450 active sites is crucial to further expand potential of P450s as biocatalysts. As studies with our optimized ester-containing linker demonstrated, there are other active site residues available to increase substrate affinity not previously involved in substrate binding of exogenous substrates. Therefore, in conjunction with examination of protein structure and dynamics, functionally diverse linkers should be tested to explore the available space between portions of the substrates and active site residues.

Finally, the knowledge accumulated in the studies of PikC and those of the tylosin P450s are directly applicable to the functional analysis of numerous unexplored and currently unidentified bacterial biosynthetic P450s to understand new biochemistry, gain insight into the biosynthesis of natural products, accumulate new candidates for biotechnology applications, and build more facile models for understanding CYP structure and mechanism. Given the recent advancements in sequencing technology, an unprecedented number of gene products are currently available for study in conjunction with cost-effective and expeditious methods of cloning, protein expression, and protein purification. Therefore, the model P450s studied in this dissertation provide the foundation to exploring the potential of other P450s to participate in not only the production of biologically active pharmaceuticals, but also fine and industrial chemicals.