BIOCHEMICAL, STRUCTURAL, AND BIOENGINEERING STUDIES OF CYTOCHROME P450 ENZYMES INVOLVED IN BIOSYNTHESIS OF SECONDARY METABOLITES

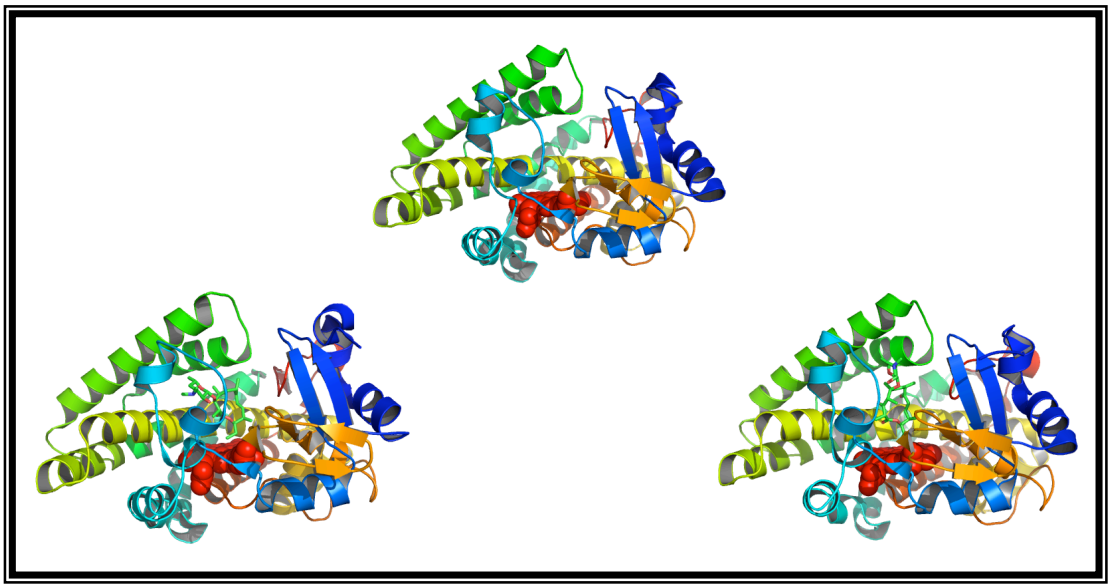
by

Shengying Li

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Medicinal Chemistry)
in The University of Michigan
2009

Doctoral Committee:

Professor David H. Sherman, Chair
Professor Paul F. Hollenberg
Professor Ronald W. Woodard
Assistant Professor Sylvie Garneau-Tsodikova
Assistant Professor Jason E. Gestwicki



To Ling, Anna,
and our families on the other side of the earth

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. David H. Sherman for his guidance and support throughout my graduate education. I greatly appreciate that Dr. Sherman offered me the opportunity and flexibility to work in a variety of interesting projects in collaboration with a group of excellent scientists, providing me broad training and valuable experiences in biochemistry, enzymology, structural biology, and organic chemistry.

I would like to acknowledge my committee members, Dr. Paul Hollenberg, Dr. Ronald Woodard, Dr. Sylvie Garneau-Tsodikova, and Dr. Jason Gestwicki for their continued guidance, suggestions, and encouragement.

In the past five years, it was a series of exciting interdisciplinary collaborations to drive my dissertational research moving forward smoothly and productively. Therefore, I am grateful to my collaborators, Dr. Larissa Podust at University of California, San Francisco, Dr. Yojiro Anzai at Toho University, Japan, and Dr. John Montgomery, Dr. Sabine Grüschow, Mr. Jacob Carlson, Dr. Mani Raj Chaulagain, Ms. Allison Knauff from the University of Michigan, for their outstanding work and inspiring ideas.

I would also like to thank all members in the Sherman laboratory for their kind help. You create a motivating yet comfortable atmosphere to work in and let me learn how to enjoy the excitement as well as the frustration in science? Especially, I owe thanks to Dr. Sabine Grüschow, Dr. Jeffery Kittendorf, Dr. Liangcai Gu, Dr. Fengan Yu, Mr. Yousong Ding, Mr. Rafay Shareef, and Dr. Patricia Cruz Lopez for their helpful instructions, discussions, and assistance during my graduate research and study.

I am grateful for the funding from NIH Grant GM078553 and the travel grants awarded by Rackham Graduate School.

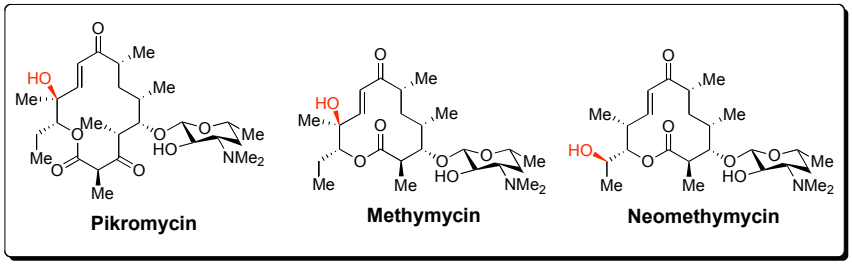
Finally, I would love to thank my wife Ling and my daughter Anna, who are the source of my power, motivation, ideas, and happiness.

Preface

This thesis contains ten chapters covering my dissertational studies on diverse biosynthetic P450 enzymes including PikC, MycCI, MycG, and TamI, an FAD-dependent oxidase TamL, two *O*-methyltransferases MycE anf MycF, and a type III polyketide synthase RppA (see next page for details). Chapter 1 is an overall introduction of the cytochrome P450 enzymes, which represent my major research targets. Chapter 2-5 are focused on the biochemical, structural, and bioengineering investigations on the cytochrome P450 PikC, which are adapted from four papers including *Journal of Biological Chemistry*, **2006**, 281, 26289-26297; **2009**, 284, 5723-5730; *Journal of the American Chemical Society*, **2007**, 129, 12940-12941; *Proceedings of National Academy of Sciences*, **In Press**, respectively. Chapter 6 and 7 demonstrate the results from the collaboration with Dr. Yojiro Anzai in the project of mycinamicin post-PKS biosynthesis that is focused on two P450 enzymes and two *O*-methyltransferases, which are adapted from two papers including *Chemistry & Biology*, **2008**, 15, 950-959 and *ChemBioChem*, **2009**, 10, 1297-1301. Chapter 8 describes the studies of the tirandamycin biosynthetic pathways in collaboration with Mr. Jacob Carlson. Chapter 9 about the type III PKS RppA is adapted from the publication of *Journal of Biological Chemistry*, **2007**, 282, 12765-12772. In chapter 10, I summarize some future directions on the basis of the current dissertational research.

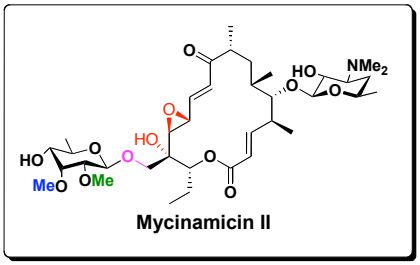
PikC

(Chapter 2-5)



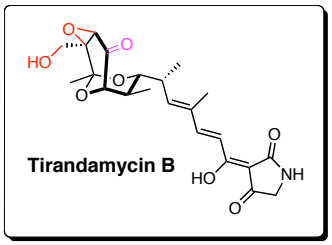
MycG, MycCI, MycE, and MycF

(Chapter 6 & 7)



TamI & TamL

(Chapter 8)



RppA

(Chapter 9)

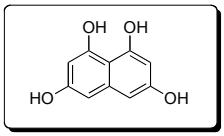


Table of Contents

List of Figures

Figure

List of Tables

Table

ABSTRACT

BIOCHEMICAL, STRUCTURAL, AND BIOENGINEERING STUDIES OF CYTOCHROME P450 ENZYMES INVOLVED IN BIOSYNTHESIS OF SECONDARY METABOLITES

by

Shengying Li

Chair: David H. Sherman

The superfamily of cytochrome P450 monooxygenases is involved in diverse oxidative processes including xenobiotic catabolism, steroid synthesis, and biosynthetic tailoring of diverse natural products. During the past decade, the synthetic potential of biosynthetic P450 enzymes from microorganisms has gained special attention due to their non-membrane bound nature, considerable catalytic efficiency, and high regio- and stereoselectivity. However, current barriers to their application in synthetic chemistry include their instability, inherent dependence on separate redox partners, and narrow substrate spectra. As these hurdles have been gradually overcome, it is likely that these biosynthetic P450s will find expanded use in the production of chemicals, fragrances, pharmaceutical compounds, biofuels, and application in bioremediation.

My dissertation research has focused on the bacterial cytochrome P450 PikC from the pikromycin macrolide antibiotic biosynthetic pathway in *Streptomyces venezuelae*. The inherent substrate flexibility and hydroxylation pattern of PikC suggests its unique oxidative mechanism and synthetic potential. Starting from the crystal structures of PikC, we not only elucidated the structural basis for its substrate flexibility, but also discovered a unique desosamine sugar anchoring functionality of this enzyme. These observations directly inspired a substrate engineering strategy that utilizes the desosamine

anchor to deliver diverse structures into the PikC active site for selective oxidation. Using this approach, the substrate spectrum of PikC has been significantly broadened. Specifically, by using an engineered $PikC_{D50N}$-RhFRED with self-sufficiency and significantly higher catalytic efficiency, a series of carbocyclic rings linked to the desosamine glycoside were effectively hydroxylated in a regioselective manner. Associated analysis of co-crystal structures of PikC with selected unnatural desosaminyl substrates provided significant insights into the mechanism of its oxidative selectivity control. Taken together, these results offer an applicable enzymatic solution of a central challenge in synthetic chemistry - the selective oxidation of an unactivated sp^3 C-H bond.

Moreover, a number of other biosynthetic P450 enzymes, two *O*-methyltransferases, an FAD-dependent oxidase, and a type III polyketide synthase were also studied during the course of my dissertation research. Together, these studies provide new insights into biosynthesis of secondary metabolites and how these enzymes can be adapted for biotechnological use.

Chapter 1
Introduction of Cytochrome P450 Enzymes

1.1 Background

Cytochrome P450 (CYP) hemoenzymes (EC 1.14.x.x) are entitled as "the most versatile biological catalyst in nature" (1) since this superfamily of enzymes catalyze a vast variety of reactions including hydroxylation, epoxidation, dealkylation, phenolic coupling, ring formation and expansion, dehydration, isomerization, etc. and their substrate spectra are extraordinarily broad (2). Not surprisingly, the amazing catalytic power of P450 enzymes has been extensively utilized by almost all types of organisms to serve diverse essential oxidative processes including xenobiotic catabolism, steroid synthesis, and biosynthetic tailoring of diverse secondary metabolites (1, 3), and starts to be intensively exploited regarding their great potential in synthetic application due to their high regio- and stereoselectivity, cost-effectiveness, and environmental-friendliness (4).

The total number of family members of CYP was growing stably since its discovery (5, 6) and began exploding in the modern genome era. For example, the complete human genome contains 57 P450 genes (7); 58 P450 genes have been identified from the genome sequence of *Glycine max* (soybean) so far (8); the model actinomycete *Streptomyces coelicolor* A3(2) that produces actinorhodin and undecylprodigiosin revealed the presence of 18 different P450 genes (9); the CYP complement (CYPome) of wood decaying fungus *Phanerochaete chrysosporium* has demonstrated more than 100 CYP genes (10). To date, there have been totally 11,292 P450 genes reported (11).

Despite the highly divergent evolution of P450 enzymes leading to tremendous diversity and low amino acid sequence identity (< 20%) among different family members

(12), some key features that enable the common dioxygen activation for catalysis, including the obligate heme cofactor with the iron protoporphyrin IX center coordinated to the thiolate provided by an absolutely conserved cystein residue; the similarity of three-dimensional structures arisen from the similar arrangement of common structural elements comprising the well-conserved helices denoted A-L; and a portion of helix I near the heme-iron that is involved in the proton delivery for the O-O bond cleavage to generate the highly reactive species (see below), are well conserved (2, 13).

1.1.1 A Brief History for Study of P450 Enzymes

Cytochrome P450 protein was first reported in 1958 as a pigment from rat liver microsomes (5). In 1964, Sato and Omura revisited this system and named the "p"igment as P450 because it displayed a signature absorption peak at 450 nm upon carbon-monoxide (CO) binding and reduction. The early studies of P450 enzymes were focused on metabolism of carcinogens (14), drugs (15), pesticides (16), vitamins (17), and steroids (18) using tissues or microsomes containing CYPs. In the late 1970s, the purified microsomal P450s started to be used for functional analysis (19, 20). Along with the development of recombinant DNA technology (21) and heterologous protein expression system (22, 23) in 1980s, a much broader range of eukaryotic and prokaryotic cytochome P450s became accessible to the growing number of P450 researchers. Notably, in 1987, Nebert developed the P450 nomenclature system based on the sequence identity that enables individual P450 enzymes to be discussed unambiguously (24). Another milestone in the P450 field, the first P450 crystal structure of $P450_{cam}$ from *Pseudomonas putida* was also achieved in the same year (25), after which a fast growing number of high-resolution crystal structures of were reported, including the first self-sufficient $P450_{BM3}$ structure (26) and the structures of the human P450 2D6 (27) and 3A4 (28). In the past two decades, much of the interest in P450 research was focused on the roles of the major human hepatic P450s such as 3A4 since it is of the significant concern for pharmaceutical industry to predict the bioavailability, drug-drug interaction, and toxicity (29). Currently, discussion on detailed mechanisms of reactive species generation and electron/proton transfer for dioxygen activation (30), structural elucidation of more unique P450 enzymes with biomedical significance and/or application potential (13),

research on P450 inhibitors (31), and functional characterization of novel P450s through genome mining (32) remain the major topics in P450 research. Of particular significance, along with the major limitations for application of cytochrome P450s in synthetic chemistry including instability, inherent dependence on separate redox partners, and narrow substrate spectrum (4) overcome by directed evolution to enhance the thermal and process stability (33), use of native (34) or engineered (35) self-sufficient fusion P450 enzymes, and extension of reaction and substrate spectrum via protein (36, 37) and/or substrate engineering (38), respectively, it is promising to take advantage of the synthetic capacity of these efficient and selective monooxygenases for production of fine chemicals, fragrances, pharmaceutical compounds, biofuels and application in bioremediation (4, 39).

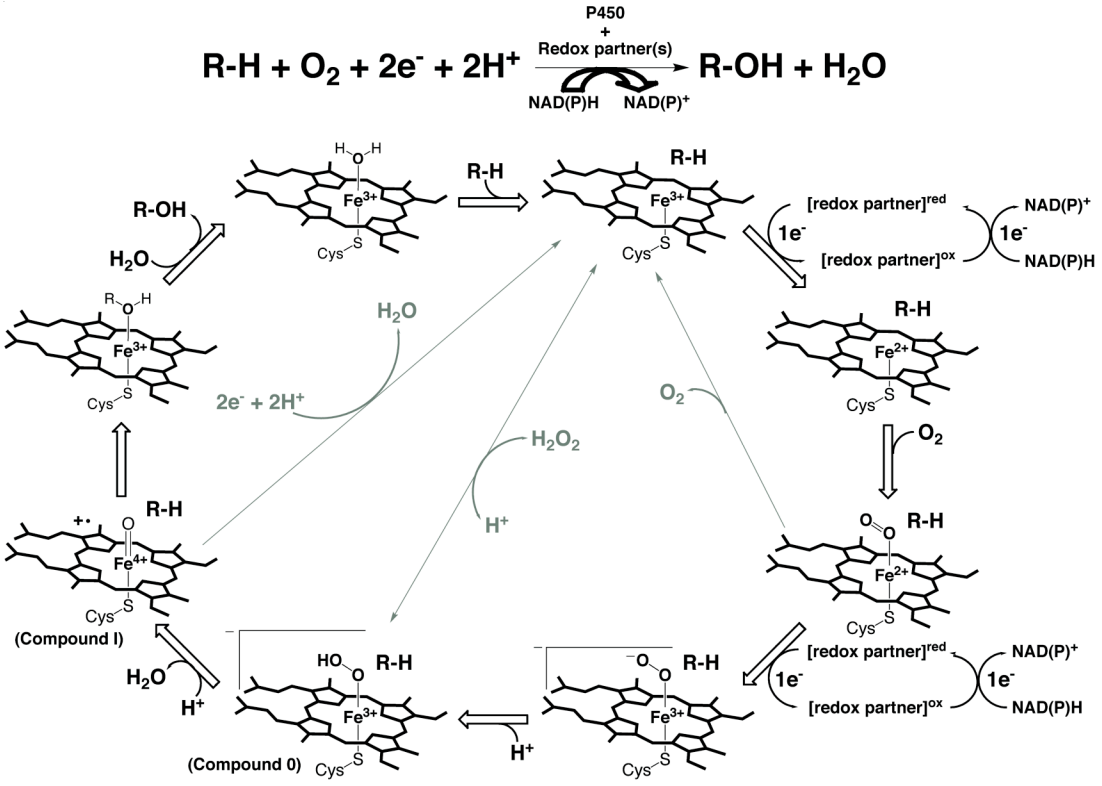


Figure 1-1. Reaction catalyzed by cytochrome P450 monooxygenases and consensus catalytic cycle. The grey arrows indicate the uncoupling reactions when electron transport and/or protonation is not timely, or substrate is absent or improperly positioned for oxidative attack.

1.1.2 Mechanism of Cytochrome P450s

Cytochrome P450 enzymes are highly divergently evolved. In contrast, various P450 enzymes share a common mechanism to activate dioxygen, thereby generating the highly reactive species for catalysis of diverse reactions. Through the observation of intermediates in the catalytic cycle by spectroscopic techniques, the use of diagnostic substrates with mechanistically revealing rearrangements during oxidation, the parallel study of the chemistry of synthetic metalloporphyrins, and the theoretical computational

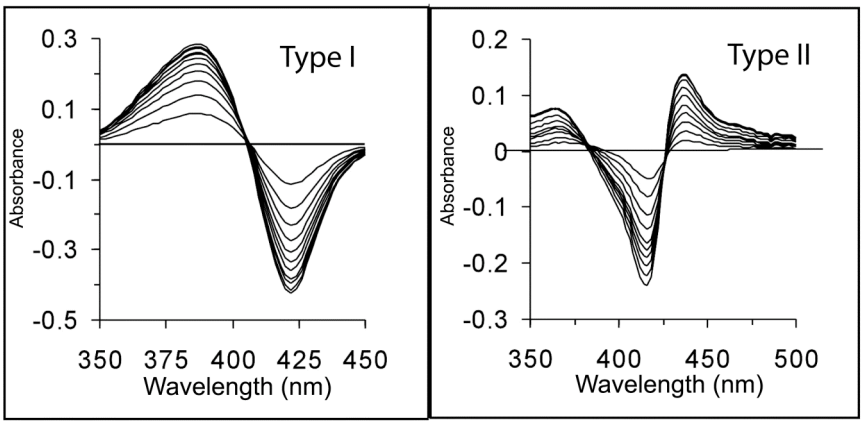


Figure 1-2. Type I and Type II difference spectrta.

investigations, the consensus catalytic mechanism of cytochrome P450s (13) has been established (Fig. 1-1) despite of ongoing debates about the exact nature of the active species responsible for the oxygen insertion (30, 40).

At the resting state, a water molecule acts as the sixth axial ligand to the ferric heme-iron (Fe^{3+}), resulting in a low spin state of the heme-iron with the representative absorbance at 420 nm. Upon substrate binding, the displacement of the water molecule from the active site induces the change of the spin state of the heme iron from the low-spin to the high-spin, giving rise to a spectral change of the enzyme with an increase in absorbance at 390 nm and a decrease at 420 nm. This difference spectrum referred to as "Type I" difference spectrum (41) (Fig. 1-2) can be conveniently measured by spectrometry and used to deduce the substrate binding affinity to P450 enzymes. In contrast, inhibitors and few certain substrates that coordinate directly to the heme iron generate the "Type II" difference spectrum (41) (Fig. 1-2), with a maximal absorbance at 430 nm and a minimum at 390 nm. The loss of the ligand water results in subtle change of the iron position relative to the plane of the porphyrin ring, making the heme a better

electron sink to trigger the first electron transferred from NAD(P)H via redox partners. Upon the one electron reduction, the ferrous heme iron (Fe^{2+}) forms a relatively stable dioxygen adduct. Subsequently, the second electron reduction (often the rate-limiting step) followed by a quick protonation generates the first reactive species – the ferric hydroperoxy (Fe^{3+}–OOH) referred to as the "Compound 0". Following second protonation and heterolytic cleavage of the O–O bond with concurrent production of a water molecule gives rise to the formation of the most applauded oxidative species – the ferryl-oxo species (Fe^{4+}=O) referred to as the "Compound I". This highly reactive cation radical is responsible for insertion of the oxygen atom into the substrate. Based on a number of experimental evidences and theoretical researches (30, 40), both "Compound 0" and "Compound I" are possible to behave as the real catalyst. The choice is believed to be dependent on the substrate and the rate of the proton transfer steps relevant to the nature of certain P450 enzymes. Finally, the oxidized product release completes the catalytic cycle. Since cytochrome P450s always only use a single oxygen atom from dioxygen molecule for oxidation, they represent a class of monooxygenases.

1.1.3 Structure of Cytochrome P450 Enzymes

Structurally, the overall three-dimensional structures of diverse P450s are conservative although their sequence identity could be as low as 20% (12, 13). In general, P450 structures have helices A-L (A, B, B', and C-L) in common and often five additional β sheets (β1-β5) (Fig. 1-3A) (42). The closer to the heme, the peptide sequences and structures are more conserved. Thus, the helices I and L that directly contact the heme are the most conserved regions. Another highly conserved region is the "Cys-Pocket" located at the β-bulge segment immediately prior to the helix L, which houses the absolutely conserved cystein that coordinate to the iron protoporphyrin center. The helix I (except for the substrate recognition site 4, see below) is conserved because it contains several residues directly involved in the dioxygen activation (13). Of particular significance, a threonine residue involved in a local helical distortion participates in a hydrogen bond network together with several peptide carbonyl groups and several water molecules. The arrangement of this H-bond network is believed to be important for the proper delivery of protons to the reactive center for the O–O bond cleavage to generate

the reactive oxidative species. Interestingly, this threonine is not strictly conserved. In $P450_{eryF}$ (43), it substitute an alanine residue for the threonine. To maintain a similar pattern of this essential hydrogen bond network, a water molecule takes the place of the threonine side chain OH and a hydroxy group from the substrate is specially involved.

In contrast, the regions controlling substrate specificity in P450 enzymes are very different from one another. These regions defined as "substrate recognition sites" (SRSs, Fig. 1-3B) include: SRS-1, helix B' and flanking regions (or a part of open B/C loop region); SRS-2, the C-terminal end of helix F; SRS-3, the N-terminal end of helix G; SRS-4, the N-terminal half of helix I; SRS-5, the β strand (in β3 area) immediately following the helix K; and SRS-6, a central region of β5 (44, 45). So far, most of the residues that take part in the substrate binding process in different P450 enzymes have been found to be within these six SRSs. Therefore, SRSs analysis would provide useful information for molecular design to guide P450 engineering, thereby gaining new substrate specificities.

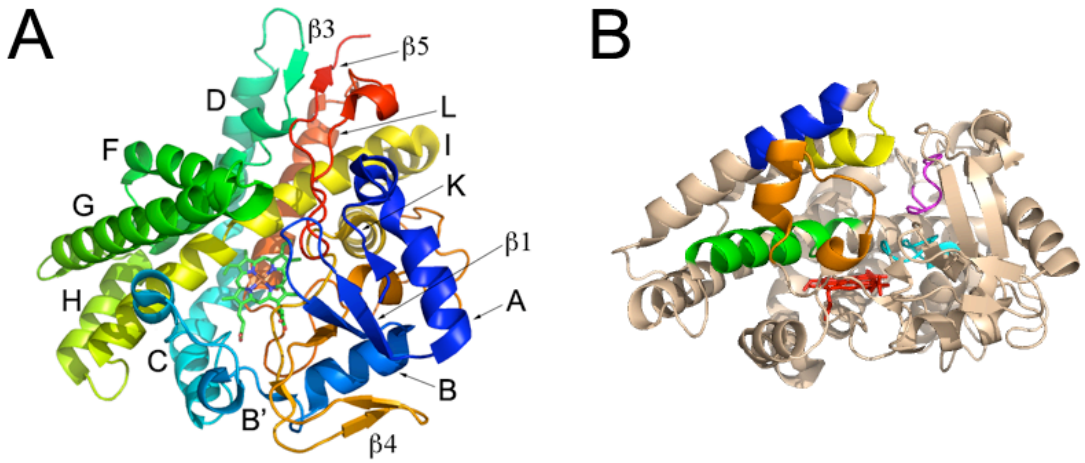


Figure 1-3. The tertiary structure of P450 monooxygenase as exemplified by the crystal structure of ligand-free PikC (PDB ID: 2BVJ) (A) Structure with the conserved secondary structure components labeled. Helices E, J, and β2 sheet are not visible from this viewing angle. (B) Structure with six substrate recognition sites highlighted in orange (SRS1); yellow (SRS2); blue (SRS3); green (SRS4); cyan (SRS5); and magenta (SRS6), respectively.

1.1.4 Classification of Cytochrome P450s

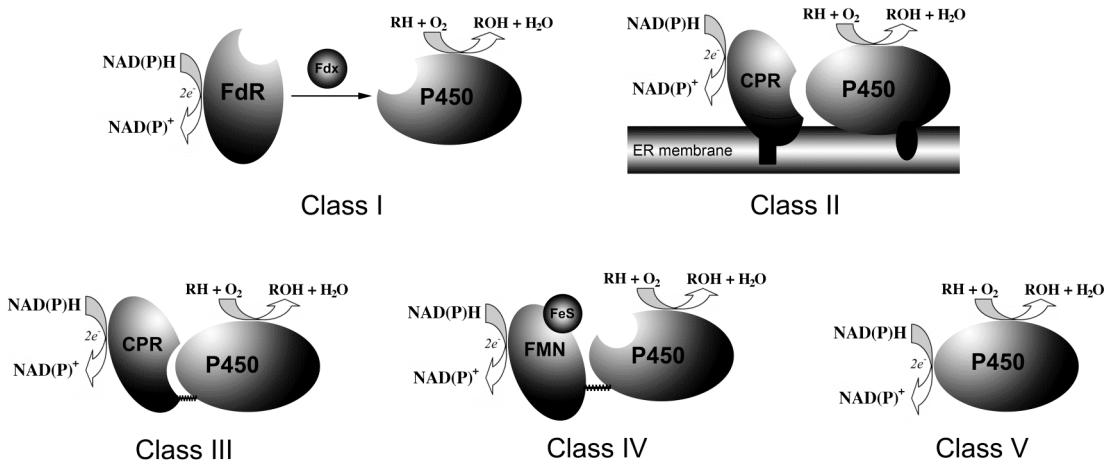


Figure 1-4. Classification of P450 monooxygenases on the basis of redox partners.

There have been a number of classification systems for cytochrome P450 monooxygenases based on different classification criteria. For example, diverse P450 enzymes are classified into hydroxylase, epoxidase, dealkylase, desaturase, dehydrotase, isomerase, and so on (2) in terms of the catalyzed reaction types. They can also be called human, plant, bacterial, or fungal P450 enzymes according to their resources. Alternatively, on the basis of sequence identity, P450s with more than 40% protein sequence identity are grouped into families (1, 2, 3…), subfamilies (A, B, C…, more than 55% identity), and individual P450s (1, 2, 3…) and thus named accordingly, such as CYP1B1, CYP3A4, CYP107L, and etc. Another important classification system of P450 enzymes is based on the redox partners required for the electron transfer to support their catalytic activities. Due to its close relatedness to this dissertational study, herein, this classification system is summarized.

Classically, there are two major redox partner systems (46,47), including an FAD containing reductase with a small iron-sulfur redoxin for most bacterial and mitochondrial P450s (Class I, Fig. 1-4), and a single FAD/FMN containing flavoprotein – cytochrome P450 reductase (CPR) for eukaryotic microsomal P450s (Class II, Fig. 1-4). $P450_{BM3}$ (48), which is naturally fused to a eukaryotic-like CPR, is representative of the Class III P450 monooxygenase (Fig. 1-4). Recently, a new class (Class IV, Fig. 1-4) of self-sufficient cytochrome P450s exemplified by $P450_{RhF}$ from *Rhodococcus* sp. NCIMB 9784 was discovered to be naturally fused to a novel FMN/Fe_2S_2 containing

reductase domain that is similar to the phthalate family of dioxygenase reductases (49, 50). Notably, the discovery of diverse self-sufficient P450 enzymes could greatly benefit the application of these monooxygenases in biotechnology due to their independence on separate redox partner(s) and the significantly higher activity of native or artificial fusion enzymes derived from a more efficient electron transport system upon covalent linkage (51). Finally, not all cytochrome P450s require redox partners. These exceptional enzymes including the soluble $P450_{nor}$ from *Fusarium oxysporum* (52), the membrane bound $P450_{TxA}$ from fall into the small Class V (Fig. 1-4) (53).

1.1.5 Microbial Secondary Metabolites Biosynthetic P450 Enzymes

Cytochrome P450 enzymes from microorganisms are enormously diverse. A significant majority of microbial P450s have been found to be involved in the biosynthesis of secondary metabolites (54). However, the roles of the vast majority of them have yet to be elucidated. Functional analysis of these many soluble enzymes is likely to result in new biochemistry, more insights into biosynthesis of natural products and biodegradation, more candidates for application in biotechnology, and more facile models for understanding CYP structures and catalytic mechanisms.

To date, actinomycetes (especially streptomycetes) represent the richest resource of secondary metabolites, accounting for more than two-thirds of microbially derived compounds. These compounds are important signal or chemical-defense molecules for the life cycles of the host microbes. A lot of them have shown significant biomedical

Table 1-1. Bioactive secondary metabolites produced by *Streptomyces*.

Streptomyces sp.	Secondary metabolite	Bioactivity	Biosynthetic CYP(s) involved	Reference
S. erythraea	Erythromycin	Antibacterial	EryF, EryK	(56, 57)
S. venezuelae	Pikromycin	Antibacterial	PikC	(59)
S. tirandis	Tirandamycin	Antibacterial	TamI	Chapter 8
S. antibioticus	Oleandomycin	Antibacterial	OleP	(61)
S. nodosus	Amphotericin	Antifungal	AmphL, AmphN	(62)
S. noursei	Nystatin	Antifungal	NysL, NysN	(63)
S. avermitilis	Avermectin	Antiparasitic	AveE	(64)
S. sp. MA6548	FK506	Immunosuppressor	FkbD	(65)
S. hygroscopius	Rampamycin	Immunosuppressor	RapJ, RapN	(66)
S. peucetius	Doxorubicin	Anticancer	DoxA	(67)
S. larvendulae	Mitomycin	Anticancer	MmcN	(68)
S. fradiae	Tylosin	Growth promotor	TylHI, TylI	(69)

activities such as antibacterial erythromycin (55-57), pikromycin (58, 59), tirandamycin (60), and oleandomycin (61); antifungal amphotericin (62) and nystatin (63); antiparasitic avermectin (64); immunosuppressors FK506 (65) and rapamycin (66); anticancer doxorubicin (67) and mitomycin (68); and livestock growth promoter tylosin (69). It is worth noting that CYPs participate in biosynthesis of all the mentioned bioactive secondary metabolites (Table 1-1). The known structures of the products and intermediates of certain biosynthetic pathway greatly facilitate the functional analysis of these P450 enzymes. These demonstrate that identification of natural product biosynthetic gene cluster is an important method to discover new microbial P450s.

In current genomic era, the whole genome sequencing represents another highly efficient process to identify new P450 genes. Recent efforts have uncovered an unexpected large number of genes encoding P450 enzymes. For example, the model actinomycete *Streptomyces coelicolor* A3(2) that produces actinorhodin and undecylprodigiosin revealed the presence of 18 different P450 genes (9), whereas *Streptomyces avermitilis* MA-4680, the avermectin producer, contains 33 P450s (64), and *Saccharopolyspora erythraea* NRRL 23338, the erythromycin-producing bacterium, encodes 36 P450s (70). However, it is of a significant difficulty to determine the physiological functions of the large number of these P450 gene products.

1.2 Prior Work

Pikromycin was isolated in 1951 (71), representing the first discovered macrolide antibiotic. Interestingly, the two structurally related 12-membered macrolides methymycin and neomethymycin were also found to be produced by the 14-membered pikromycin producer strain – *Streptomyces venezuelae* (Fig. 1-5) (72, 73). In the past more than a decade, the Sherman laboratory has been focused on elucidation of the pikromycin biosynthetic pathway and understanding the detailed enzymological mechanisms of the unique modular type I polyketide synthase (PKS) and a group of intriguing tailoring enzymes. The knowledge accumulated during these investigations has greatly advanced the understanding of macrolide biosynthesis, making the pikromycin biosynthetic pathway, together with the erythromycin biosynthetic pathway (74), become a model system in the field of natural product biosynthesis (75, 76).

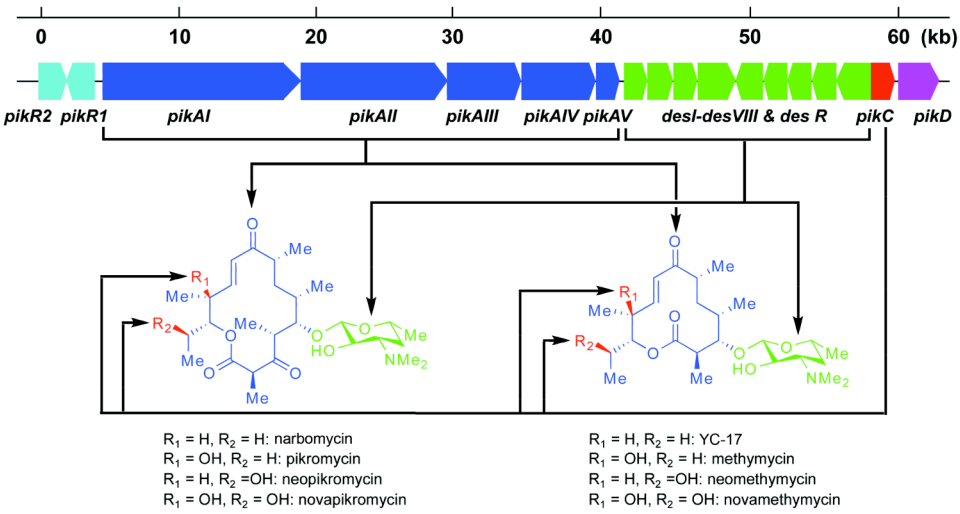


Figure 1-5. The pikromycin biosynthetic pathway. The *pikA* locus (blue) encodes a type I PKS that directs the biosynthesis of the 12- and 14-membered ring macrolides, 10-deoxymethonolide and narbonolide (shown in blue), respectively. The *des* locus (green) encodes all proteins responsible for the generation of the desosamine sugar moity (shown in green) and attachment to macrolactone, giving rise to macrolides narbomycin and YC-17. The *pikC* (red) encodes a cytochrome P450 monooxygenase responsible for installation of all hydroxy groups shown in red. The figure is adapted from Kittendorf J. D. and Sherman, D. H. (74)

The first milestone of pikromycin biosynthesis studies was sequencing and assembly of the pikromycin biosynthetic gene cluster (58), which identified a cluster of genes encoding a five-modular type I polyketide synthase (*pikAI-V*), a group of enzymes responsible for synthesis of 6-deoxysugar desosamine (*desI-VIII* and *desR*), a cytochrome P450 (*pikC*), and a number of regulatory (*pikD*) and self-resistant (*pikR1* and *pikR2*) genes (Fig. 1-5). Shortly afterwards, genetic disruption studies unambiguously demonstrated that the pikromycin PKS is unique from other known PKS systems since it is solely capable of generating two distinct macrolactone products, 10-deoxymethonolide and narbonolide (Fig. 1-5) (77, 78).

The subsequent researches on the desosamine biosynthesis not only functionally characterized all enzymes (DesI-DesVI) required for transformation of the primary metabolite D-glucose-1-phosphate to thymidine diphosphate D-desosamine (TDP-desosamine) (75, 79-82), but also identified a glycosyltransferase DesVII (paired with the auxiliary protein DesVIII) (83), displaying amazing substrate tolerance toward both aglycone and the sugar moiety (84, 85). This unique substrate flexibility has been shown

to be very useful for generating novel unnatural products potentially with improved biological activity (82, 85, 86).

The cytochrome P450 PikC represents the third component (another two are the thioesterase in module V (77) and the glycosyltransferase DesVII) with enzymatic versatility in pikromycin pathway. Physiologically, PikC catalyzes hydroxylation at C10 of the 12-membered ring macrolide YC-17 to yield methymycin, or almost equally at the C12 position to generate neomethymycin. It is also able to further hydroxylate methymycin at C12, giving rise to double hydroxylated product novamethymycin (59,87). With respect to 14-membered ring substrate narbomycin, the PikC catalyzed hydroxylation predominantly occurs at the C12 position, leading to pikromycin (59). Recently, identification of two additional 14-membered ring macrolide antibiotics, neopikromycin and novapikromycin in much smaller amounts compared to pikromycin, from extracts of *S. venezuelae* fermentation broth clearly suggested more functionalities of PikC (88). Further *in vitro* reactions confirmed that the minor activity of PikC to convert narbomycin into neopikromycin and hydroxylate pikromycin to form novapikromycin (88).

This dissertational study started from the collaboration with Dr. Larissa M. Podust, attempting to understand the structural basis for the amazing substrate flexibility and hydroxylation versatility of PikC. The observation of a unique desosamine anchoring feature initiate the following biochemical studies regarding catalytic mechanism and substrate specificity, and the bioengineering work to develop PikC into an efficient and cost-effective synthetic enzyme by harnessing its unique desosamine anchoring functionality. To date, the results have demonstrated that 1) PikC is a good model to help answer some fundamental questions of P450 catalysis; 2) PikC possesses great potential to be developed into an applicable catalyst in synthetic chemistry. Moreover, knowledge and experience gained during the PikC investigation has been applied to the study of other biosynthetic P450s.

1.3 References

1. Coon, M. J. (2005) *Ann. Rev. Pharmacol. Toxicol.* **45**, 1-25
2. Guengerich, P. F. (2001) *Chem. Res. Toxicol.* **14**, 611-650
3. Bernhardt, R. (2006) *J. Biotechnol.* **124**, 128-145
4. Urlacher, V. B., and Eiben, S. (2006) *Trends Biotechnol.* **24**, 324-330
5. Klingenberg, M. (1958) *Arch. Biochem. Biophys.* **75**, 376-386
6. Omura, T., and Sato, R. (1964) *J. Biol. Chem.* **239**, 2379-2385
7. International, Human, Genome, Sequencing, and Consortium. (2001) *Nature* **409**, 860-921
8. *http://www.phytozome.net/soybean*
9. Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) *Nature* **417**, 141-147
10. Martinez, D., Larrondo, L. F., Putnam, N., Gelpke, M. D., Huang, K., Chapman, J., Helfenbein, K. G., Ramaiya, P., Detter, J. C., Larimer, F., Coutinho, P. M., Henrissat, B., Berka, R., Cullen, D., and Rokhsar, D. (2004) *Nat. Biotechnol.* **22**, 695-700
11. Nelson, D. (2009) *http://drnelson.utmem.edu/*
12. Bernhardt, R. (2004) *Chem. Biol.* **11**, 287-288
13. Thomas, P. L., and Johnson, E. F. (2005) *Cytochrome P450: Structure, Mechanism, and Biochemistry (3rd Ed.) Ortiz de Montellano, P. R. Ed. Kluwer Academic / Plenum Publishers, New York*, 87-114
14. Conney, A. H., Miller, E. C., and Miller, J. A. (1956) *Cancer Res.* **16**, 450-459
15. Brodie, B. B., Gillette, J. R., and LaDu, B. N. (1958) *Annu. Rev. Biochem.* **27**, 427-454
16. Kamataki, T., and Neal, R. A. (1976) *Mol. Pharmacol.* **12**, 933-944
17. Capdevila, J. H., Zeldin, D., Makita, K., Karara, A., and Falck, J. R. (1995) *Cytochrome P450: Structure, Mechanism, and Biochemistry (2nd Ed.) Ortiz de Montellano, P. R. Ed. Plenum Press, New York*, 443-471
18. Estabrook, R. W., Cooper, D. Y., and Rosenthal, O. (1963) *Biochem. Z.* **338**, 741-755
19. Lu, A. Y. H., and Coon, M. J. (1968) *J. Biol. Chem.* **243**, 1331-1332
20. Wang, P. P., Beaune, P., and Guengerich, F. P. (1980) *Arch. Biochem. Biophys.* **199**, 206-219
21. Gonzalez, F. J. (1989) *Pharmacol. Rev.* **40**, 243-288
22. Larson, L. R., Coon, M. J., and Porter, T. D. (1991) *J. Biol. Chem.* **266**, 7321-7324
23. Barnes, H. J., Arlotto, M. P., and Waterman, M. R. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5597-5601
24. Nebert, D. W., Adesnik, M., Coon, M. J., Estabrook, R. W., Gonzalez, F. J., and Guengerich, F. P. (1987) *DNA* **6**, 1-11
25. Poulos, T. L., Finzel, B. C., and Howard, A. J. (1987) *J. Mol. Biol.* **195**, 687-700
26. Ravichandran, K. G., Boddupalli, S. S., Hasermann, C. A., Peterson, J. A., and Deisenhofer, J. (1993) *Science* **261**, 731-736
27. Rowland, P., Blaney, F. E., Smyth, M. G., Jones, J. J., Leydon, V. R., Oxbrow, A. K., Lewis, C. J., Tennant, M. G., Modi, S., Eggleston, D. S., Chenery, R. J., and Bridges, A. M. (2006) *J. Biol. Chem.* **281**, 7614-7622
28. Williams, P. A., Cosme, J., Vinkovic, D. M., Ward, A., Angove, H. C., Day, P. J., Vonrhein, C., Tickle, I. J., and Jhoti, H. (2004) *Science* **305**, 683-686

29. Guengerich, P. F. (2005) *Cytochrome P450: Structure, Mechanism, and Biochemistry (3rd Ed.) Ortiz de Montellano, P. R. Ed. Kluwer Academic / Plenum Publishers, New York*, 377-530
30. Meunier, B., de Visser, S. P., and Shaik, S. (2004) *Chem. Rev.* **104**, 3947-3980
31. Correia, M. A., and Ortiz de Montellano, P. R. (2005) *Cytochrome P450: Structure, Mechanism, and Biochemistry (3rd Ed.) Ortiz de Montellano, P. R. Ed. Kluwer Academic / Plenum Publishers, New York*, 247-322
32. Anzai, Y., Li, S., Chaulagain, M. R., Kinoshita, K., Kato, F., Montgomery, J., and Sherman, D. H. (2008) *Chem. Biol.* **15**, 950-959
33. Cherry, J. R. (2000) *Cur. Opin. Biotechnol.* **11**, 250-254
34. Munro, A. W., David, L. G., Kirsty, M. J., Ker, M. R., Tobias, O. W. B., Simon, D., Caroline, M. S., Stephen, C. K., Dominikus, L. A., Christopher, M. C., Christopher, P. C., and Leslie, D. P. (2002) *Trends Biochem. Sci.* **27**, 250-257
35. Li, S., Podust, L. M., and Sherman, D. H. (2007) *J. Am. Chem. Soc.* **129**, 12940-12941
36. Landwehr, M., Carbone, M., Otey, C. R., Li, Y., and Arnold, F. H. (2007) *Chem. Biol.* **14**, 269-278
37. Arnold, F. H. (2006) *Nat. Biotechnol.* **24**, 328-330
38. Griengl, H., and de Raadt, A. (2002) *Curr. Opin. Biotechnol.* **13**, 537-542
39. Ramanavicius, A., and Ramanaviciene, A. (2009) *Fuel Cells* **1**, 25-36
40. Ortiz de Montellano, P. R., and De Voss, J. J. (2002) *Nat. Prod. Rep.* **19**, 477-493
41. Schenkman, J. B., Remmer, H., and Estabrook, R. W. (1967) *Mol. Pharmacol.* **3**, 113-123
42. Hasemann, C. A., Kurumbail, R. G., Boddupalli, S. S., Peterson, J., and Deisenhofer, J. (1995) *Structure* **3**, 41-62
43. Nagano, S., Cupp-Vickery, J. R., and Poulos, T. L. (2005) *J. Biol. Chem.* **280**, 22102-22107
44. Gotoh, O. (1992) *J. Biol. Chem.* **267**, 83-90
45. Podust, L. M., Stojan, J., Poulos, T. L., and Waterman, M. R. (2001) *J. Inorg. Biochem.* **87**, 227-235
46. Lewis, D. F. V., and Hlavica, P., 353-374. (2000) *Biochim. Biophys. Acta.* **1460**, 353-374
47. Munro, A. W., Girvan, H. M., and McLean, K. J. (2007) *Nat. Prod. Rep.* **24**, 585-609
48. Ruettinger, R. T., and Fulco, A. J. (1981) *J. Biol. Chem.* **256**, 5728-5734
49. Roberts, G. A., Grogan, G., Greter, A., Flitsch, S. L., and Turner, N. J. (2002) *J. Bacteriol.* **184**, 3898-3908
50. De Mot, R., and Parret, A. H. A. (2002) *Trends Microbiol.* **10**, 502-508
51. Munro, A. W., Girvan, H. M., and McLean, K. J. (2007) *Biochim. Biophys. Acta* **1770**, 345-359
52. Daiber, A., Shoun, H., and Ullrich, V. (2005) *J. Inorg. Biochem.* **99**, 185-193
53. Wang, L.-H., and Kulmacz, R. J. (2002) *Prostaglandins Other Lipid Mediat.* **68-69**, 409-422
54. Kelly, S. L., Kelly, D. E., Jackson, C. J., Warrilow, A. G. S., and Lamb, D. C. (2005) *Cytochrome P450: Structure, Mechanism, and Biochemistry (3rd Ed.) Ortiz de Montellano, P. R. Ed. Kluwer Academic / Plenum Publishers, New York*, 585-618
55. Weber, J. M., Leung, J. O., Maine, G. T., Potenz, R. H. B., Poulus, T. J., and DeWitt, J. P. (1990) *J. Bacteriol.* **172**, 2372-2383
56. Anderson, J. F., and Hutchinson, C. R. (1992) *J. Bacteriol.* **174**, 725-735
57. Lambalot, R. H., Cane, D. E., Aparicio, J. J., and Katz, L. (1995) *Biochemistry* **34**, 1858-1866
58. Xue, Y., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**(12111-12116)

59. Xue, Y., Wilson, D., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Chem. Biol.* **5**, 661-667
60. Reusser, F. (1976) *Antimicrob. Agents Chemother.* **10**, 618–622
61. Rodriguez, A. M., Olano, C., Méndez, C., Hutchinson, C. R., and Salas, J. A. (1995) *FEMS Microbiol. Lett.* **127**, 117-120
62. Caffrey, P., Lynch, S., Flood, E., Finnan, S., and Oliynyk, M. (2001) *Chem. Biol.* **8**, 713-723
63. Brautaset, T., Sekurova, O. N., Sletta, H., Ellingsen, T. E., Strøm, A. R., Valla, S., and Zotchev, S. B. (2000) *Chem. Biol.* **7**, 395-403
64. Ikeda, H., Nonomiya, T., Usami, M., Ohta, T., and Ōmura, S. (1999) *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9509-9514
65. Motamedi, H., and Shafiee, A. (1998) *Eur. J. Biochem.* **256**, 528-534
66. Schwecke, T., Aparicio, J. F., Molnár, I., König, A., Khaw, L. E., Haydock, S. F., Oliynyk, M., Caffrey, P., Cortés, J., and Lester, J. B. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92**, 7839–7843
67. Lomovskaya, N., Otten, S. L., Doi-Katayama, Y., Fonstein, L., Liu, X. C., Takatsu, T., Inventi-Solari, A., Filippini, S., Torti, F., Colombo, A. L., and Hutchinson, C. R. (1999) *J. Bacteriol.* **181**, 305-318
68. Mao, Y., Varoglu, M., and Sherman, D. H. (1999) *Chem. Biol.* **6**, 251-263
69. Baltz, R. H., and Seno, E. T. (1988) *Ann. Rev. Microbiol.* **42**, 547-574
70. Oliynyk, M., Samborskyy, M., Lester, J. B., Mironenko, T., Scott, N., Dickens, S., F., H. S., and Leadlay, P. F. (2007) *Nat. Biotechnol.* **25**, 447–453
71. Brockmann, H., and Henkel, W. (1951) *Chem. Ber.* **84**, 284-289
72. Donin, M. N., Pagano, J., Dutcher, J. D., and McKee, C. M. (1953) *Antibiot. Annu.* **1**, 179-185
73. Perlman, D., and O'Brien, E. (1954) *Antibiot. Chemother.* **4**, 894-898
74. Stephen, F. J. C., Gareth, A. H., Bevitt, R. D. J., and Leadlay, P. F. (1990) *Nature* **348**, 176-178
75. Xue, Y., and Sherman, D. H. (2001) *Metab. Eng.* **3**, 15-26
76. Kittendorf, J. D., and Sherman, D. H. (2009) *Bioorg. Med. Chem.* **17**, 2137-2146
77. Xue, Y., and Sherman, D. H. (2000) *Nature* **403**, 571-575
78. Xue, Y., Wilson, D., and Sherman, D. H. (2000) *Gene* **245**, 203-211
79. Szu, P. H., He, X., Zhao, L., and Liu, H.-w. (2005) *Angew. Chem. Int. Ed.* **44**, 6742-6746
80. Chen, H., Yamase, H., Murakami, K., Chang, C. W., Zhao, L., Zhao, Z., and Liu, H.-w. (2002) *Biochemistry* **41**, 9165-9183
81. Zhao, L., Borisova, S., Yeung, S., and Liu, H.-w. (2001) *J. Am. Chem. Soc.* **132**, 7909-7910
82. Borisova, S. A., Zhang, C., Takahashi, H., Zhang, H., Wong, A. W., Thorson, J. S., and Liu, H.-w. (2006) *Angew. Chem. Int. Ed.* **45**, 2748-2753
83. Borisova, S. A., Zhao, L., Melançon III, C. E., Kao, C. L., and Liu, H.-w. (2004) *J. Am. Chem. Soc.* **126**, 6534-6535
84. Kao, C., Borisova, S. A., Kim, H. J., and Liu, H.-w. (2006) *J. Am. Chem. Soc.* **128**, 5606-5607
85. Borisova, S. A., Kim, H. J., Pu, X., and Liu, H.-w. (2008) *ChemBioChem* **9**, 1554-1558
86. Thibodeaux, C. J., Melançon, C. E., and Liu, H.-w. (2007) *Nature* **446**, 1008-1016
87. Zhang, Q., and Sherman, D. H. (2001) *J. Nat. Prod.* **64**, 1447-1450
88. Lee, S. K., Park, J. W., Kim, J. W., Jung, W. S., Park, S. R., Choi, C. Y., Kim, E. S., Ahn, J. S., Sherman, D. H., and Yoon, Y. J. (2006) *J. Nat. Prod.* **69**, 847-849

Chapter 2
The Structural Basis for Substrate Anchoring, Active Site Selectivity, and Product Formation by P450 PikC from *Streptomyces venezuelae*

2.1 Introduction

Macrolide antibiotics comprise a large group of medicinal agents characterized by a macrocyclic lactone ring, to which one or more sugar residues are covalently linked. Despite considerable structural variation, macrolides represent a homogeneous group of therapeutic drugs with similar activity spectra and mode of action as anti-infective agents. The success of macrolide antibiotics is attributed to their propensity to bind to the large subunit of prokaryotic ribosomes and inhibit protein synthesis, thereby preventing bacterial growth (1, 2). The first generation macrolide introduced into clinical practice over 50 years ago was erythromycin. Since then, macrolide antibiotics have been further optimized, resulting in improved 14-, 15-, and 16-membered ring macrolides (the second generation), acylides, and ketolides (the third generation) (3).

Most of the natural product macrolide antibiotics are produced by *Streptomyces* sp. and related bacteria, in which assembly of polyketides from simple carboxylic acid precursors is catalyzed by modular polyketide synthases. Over the past 15 years, advances in understanding the modular architecture of polyketide biosynthetic machinery has enabled development of metabolic engineering approaches for production of new antibiotics (4-7). However, significant additional structural variability of polyketide-derived natural products is due to postpolyketide synthase biosynthetic modifications that typically include hydroxylation/epoxidation and/or glycosylation. In most antibiotic biosynthetic pathways, hydroxylation(s) occur(s) at the late stages of assembly after formation of the natural product scaffold and often after glycosylation events. Such

modifications are often necessary to impart or enhance biological activity (8). The 3-(dimethylamino)-3,4,6-trideoxy sugar desosamine (or mycaminose) confers biological activity to a number of macrolide antibiotics such as erythromycin, troleandomycin, mycinamicin, megalomicin (desosamine), tylosin, carbomycin, and spiramycin (mycaminose) and is the only glycoside present on pikromycin, methymycin, and the highly potent semisynthetic ketolide telithromycin (9).

Both macrolides and ketolides were shown crystallographically to bind in a specific pocket in the ribosomal tunnel via interactions with 23 S rRNA and act to block sterically egress of the nascent protein chains (10). The desosamine sugar (or mycaminose) attached to C5 of the 14-, 15-, and 16-membered ring macrolactones extends up the tunnel toward the peptidyl transferase center. Interactions between desosamine and the ribosome play a key role in both macrolide selectivity and macrolide resistance (11).

Previous studies have shown that metabolic engineering approaches can provide structurally variable macrolactones (12, 13) and/or extend a repertoire of deoxysugars coupled to macrolide aglycones (14-18). The Pik biosynthetic gene cluster of *Streptomyces venezuelae* represents an effective system for the synthesis of novel polyketide antibiotics due to (i) its ability to generate two macrolactone ring systems; (ii) the presence of a flexible desosaminyl transferase (DesVII) that is tolerant of changes in macrolactone structure (13, 18, 19), as well as modifications to the sugar substituent (14-

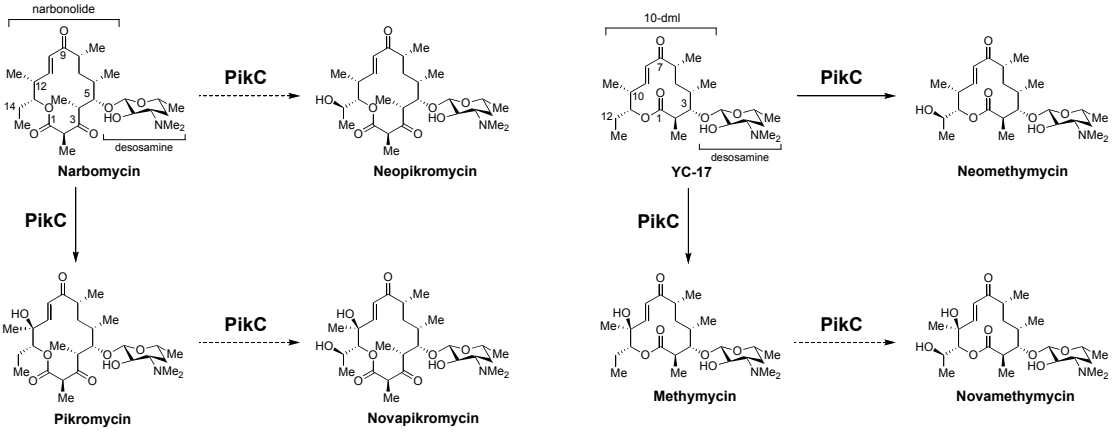


Figure 2-1. Catalytic roles of PikC. Hydroxylated natural products obtained from the Pik pathway are shown. 10-dml, 10-deoxymethynolide. Solid arrows indicate reactions leading to the primary products. Dashed arrows indicate minor reactions.

17); and finally, (iii) the unusual P450 monooxygenase (CYP) PikC (CYP107L1), responsible for the diverse pattern of hydroxylated natural products obtained from the Pik pathway (Fig. 2-1).

A series of particularly intriguing observations were made several years ago regarding the substrate selectivity of PikC that appeared to correlate directly with hydrogen bond functionality at the C3' position of the glycoside, with no tolerance for modifications at C4' (14-17). Such a stringent requirement toward sugar structure contrasts with the apparent flexibility of PikC toward the macrolactone core of a substrate. Thus, PikC is able to catalyze mono-hydroxylation at either C10 or C12 of the 12-membered ring macrolactone of YC-17, giving rise to methymycin and neomethymycin, respectively, as well as at C12 of the 14-membered ring of narbomycin, giving rise to pikromycin (20). Hydroxylation at the C14 position of pikromycin, giving rise to neopikromycin, occurs with a very low yield, which precluded isolation and characterization of this metabolite until recently (21). Dihydroxylation of YC-17 results

Table 2-1. Data collection and refinement statistics

PEG, polyethylene glycol; MME, monomethyl ether.

PDB ID	Ligand-free 2BVJ	Narbomycin-bound 2C7X	YC-17-bound 2C6H
Data collection			
Resolution, Å	2.10	1.70	2.35
Wavelength, Å	1.00000	0.97918	1.72196
Space group	C2	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, *b*, *c*, Å	218.2, 64.2, 73.6	59.0, 64.9, 92.8	60.5, 104.7, 153.6
α, β, γ, °	90.00, 107.15, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Molecules in an asymmetric unit	2	1	2
Solvent content, %	52	34	52
R_{sym} [a,b], %	6.0 (26.0)	5.9 (37.3)	10.0 (49.0)
I/σ	19.5 (3.7)	65.3 (3.6)	23.8 (4.3)
Unique reflections	55910	35331	33864
Completeness, %	98.2 (87.6)	97.3 (80.9)	98.3 (95.8)
Redundancy	3.7 (3.1)	11.2 (6.5)	6.8 (6.2)
Crystallization conditions	10% PEG 8000, 0.1 M HEPES, 7.5	0.8 M $(NH_4)_2SO_4$, 0.1 M MES, 6.0	30% PEG MME 5000, 0.2 M $(NH_4)_2SO_4$, 0.1 M MES, 6.5
Refinement			
Reflections used in refinement	53812	34218	72999
R_{cryst}/R_{free} [c], %	20.2/24.3	20.7/23.2	19.5/23.1
No. atoms			
Protein	6046	3212	6101
Heme	86	43	86
Substrate	0	36	64
Water/ion	317/16	171/0	216/25
Wilson plot *B*-values, Å^2	13.2	28.6	20.5
Mean *B*-factors, Å^2			
Protein	28.3	35.9	23.2
Heme	17.7	23.2	14.6
Substrate	NA	45.1	20.12
Water	30.3	41.9	23.9
r.m.s. deviations			
Bond lengths, Å	0.006	0.01	0.006
Bond angles, °	1.2	1.4	1.2
Ramachandran [d]	A, 89.6/10.4; B, 90.9/9.1	A, 92.3/7.4/0.3	A, 91.1/8.9; B, 92.6/7.4

[a] $R_{sym} = \Sigma | I_i - \langle I \rangle | / \Sigma I_i$, where I_i is the intensity of the *i*th observation, and $\langle I \rangle$ is the mean intensity of reflection.
[b] Numbers in parentheses correspond to the highest resolution shell.
[c] $R_{cryst} = \Sigma \| F_o | - | F_c \| / \Sigma | F_o |$, calculated with the working reflection set. R_{free} is the same as R_{cryst} but calculated with the reserved reflection set.
[d] Program PROCHECK (28).

in novamethymycin (Fig. 2-1), a product that is found in small quantities *in vivo* and appears to be converted from methymycin *in vitro* (22). Dihydroxylation of narbomycin results in a very low but detectable production of novapikromycin (21).

Here we provide new information regarding the diverse hydroxylation pattern of the PikC monooxygenase by determining its crystal structure in a variety of forms, including ligand-free and bound to endogenous substrates YC-17 and narbomycin (see Table 2-1). This series of crystal structures has revealed that in the absence of a substrate, PikC adopts both "open" and "closed" conformations due to repositioning of the BC-loop and the F and G helices. This dynamic process appears to enable substrate access to the active site; however, it does not account for PikC substrate flexibility, which is largely due to the combination of a single macrolactone binding site with two alternative desosamine binding pockets. In both pockets, the desosamine C3' dimethylamino group is sandwiched between two carboxyl-containing (Glu, Asp) amino acids found in the BC-region, forming a salt bridge to the more proximal carboxyl group. Site-directed mutagenesis revealed that formation of this salt bridge is essential for PikC function.

2.2 Results and Discussion

2.2.1 Ligand-free PikC

The crystal structure of the ligand-free form of PikC revealed that two molecules in an asymmetric unit adopt two different conformations, referred to as open and closed depending on the position of the F and G helices and the BC-loop, the latter being partially disordered in the open conformation (Fig. 2-2). These regions are generally flexible in P450s and readily reposition in response to inhibitor or substrate binding in the active site (23-25). To our knowledge, this is the first time that a P450 demonstrates such protein dynamics (linear repositioning of the FG-loop in the ligand-free PikC is within 12 Å) in the absence of a bound ligand, probably as a result of natural motions, although the first structural indications of open/close differences in the FG-region were noticed for a ligand-free $P450_{BM3}$ (26). The most compelling need for such protein dynamics is to allow substrate access to the active site and subsequent product release. A water molecule is bound as the sixth ligand to the hexa-coordinated heme Fe^{3+} atom. Scattered electron density accommodates a few additional water molecules that together form a hydrogen

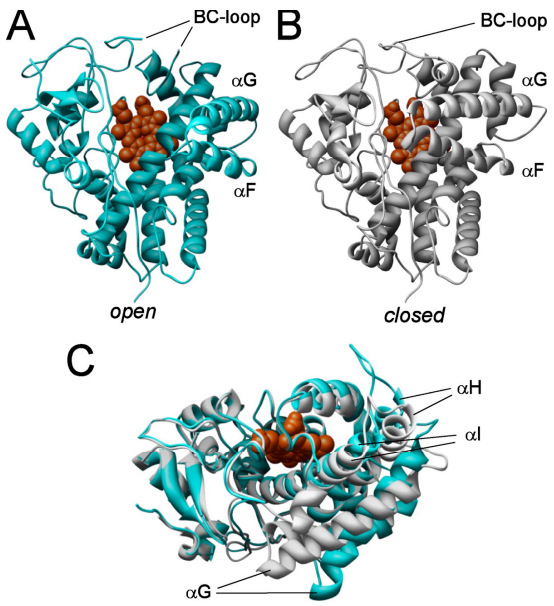


Figure 2-2. Ribbon representation of ligand-free PikC. A, open, and B, closed conformations of ligand-free PikC (2BVJ). C, overlay of both conformations, open (cyan) and closed (gray), demonstrating that in the open form, the F and G helix are bent away from the heme to enable substrate access to the active site. Molecules in C are rotated ~90° toward the viewer along a horizontal axis in the plane of drawing when compared with A and B. The F helix is not seen in this orientation. Closed conformation is related within r.m.s. deviations of 0.58 Å for C_α atoms to catalytically relevant YC-17- and narbomycin-bound forms. The heme co-factor is shown in red.

bond network within the active site. Water molecules are displaced upon substrate binding.

2.2.2 X-ray Structure of Narbomycin-bound PikC

The 14-membered ring macrolide narbomycin is one of the two endogenous substrates for PikC (20). Hydroxylation of narbomycin at C12 (Fig. 2-1) of the macrolactone ring gives rise to the antibiotic pikromycin. In the co-crystal structure (1.75 Å resolution), narbomycin is unambiguously positioned in the active site, and its electron density is well defined (Fig. 2-3A). Of particular significance are interactions between the C3' dimethylamino group of the desosamine sugar and two carboxyl-containing residues localized in the BC-region (Glu-85 and Asp-50) (Fig. 2-4A). Although proximal Glu-85 (3.2 Å distance) provides a salt bridge contact, distal Asp-50 (5.3 Å distance) may compensate for a partial positive charge of the protonated tertiary amine. A calculated pKa of 8.85 for the desosamine nitrogen atom (calculated by using the on-line program

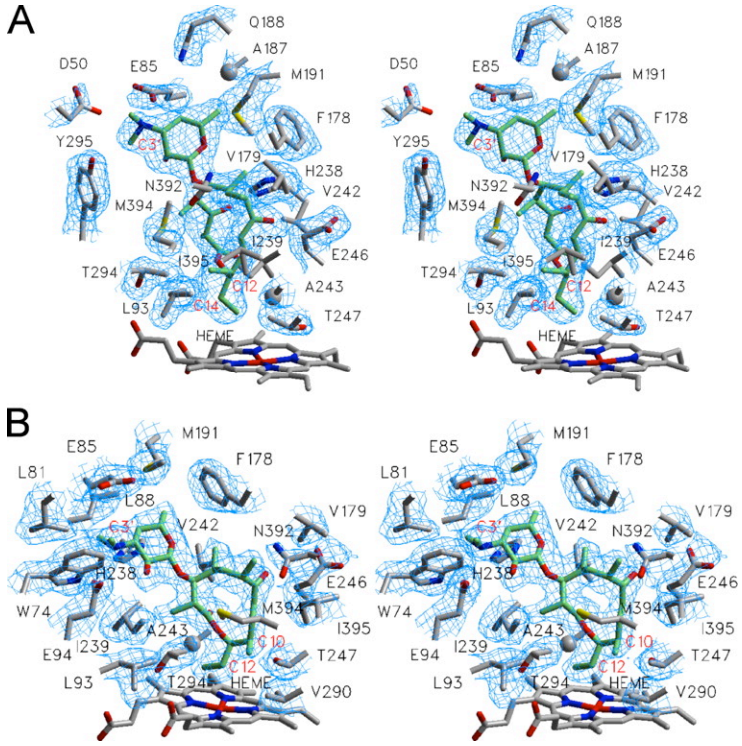


Figure 2-3. Substrate binding in the PikC active site. Stereo views of narbomycin (2C7X) (A) and YC-17 (2C6H) (B) bound in the active site of PikC are shown. Amino acid residues within 4.0 Å (labeled in black) are also shown. Carbon atoms are labeled in red. Fragments of 2Fo - Fc electron density composite omit map contoured at 0.8 σ are in blue. To avoid excessive cluttering of stereo views with the electron density bulks, a number of residues, including heme and those that project on top of the substrate or each other, are excluded from the map calculation. Excluded in addition to heme in panel A are: Val-179, Ile-239, Val-242, Asn-392, and Ile-395; Excluded in addition to heme in panel B are: His-238, Val-242, Ala-243, and Met-394.

SPARC (27)) is in good agreement with the values experimentally obtained for erythromycin (28), suggesting that at neutral pH, desosamine exists primarily in the protonated form. Interestingly, electrostatic interactions involving the C3' dimethylamino group are also involved in macrolide recognition in the ribosome tunnel where a positive charge is neutralized by a negative charge of an rRNA phosphate group invariantly positioned within 5 Å in structurally defined macrolide/ribosome complexes (29-32); however, no salt bridge is formed (Fig. 2-4C). In this regard, macrolide resistance in bacterial pathogens is often attributed to mono- or dimethylation at the A2058 N6 position (*Escherichia coli* numbering) or to mutations that change nucleotide identity, both eliminating interactions with desosamine (11). In addition, PikC amino acid side chains Phe-178, Ala-187, Gln-188, Met-191, and Tyr-295 are within4Åofthe desosamine

moiety (Fig. 2-3A and 2-4A). Although substrate anchoring by a biosynthetic P450 involving desosamine is unique, another type of anchoring mechanism was observed previously for the epothilone P450 EpoK from Sorangium cellulosum. In EpoK, a pendant thiazole moiety linked to the macrolactone ring system is involved in π-π interactions with two aromatic amino acids that determine substrate specificity (33).

Significantly, the macrolactone ring of narbomycin is bound almost entirely via hydrophobic interactions with amino acid side chains Leu-93, Phe-178, Val-179, Ile-239, Val-242, Ala-243, Thr-247, Thr-294, Met-394, and Ile-395 plus three hydrophilic ones, Glu-246, His-238, and Asn-392, but no hydrogen bonding is involved. The Thr-247 hydroxyl is involved in hydrogen bonding with the I helix main chain atoms, causing a local distortion observed in many if not all structurally defined P450s. The Glu-246 side chain is in a conformation that is flipped away from the substrate (Fig. 2-3A). These two I helix residues, Glu-246 and Thr-247, are highly conserved among P450s and play an important part in activation and delivery of protons to the iron-linked dioxygen (34). Surprisingly, the allylic C12 carbon (analog of C10 in YC-17) is positioned 7.1 Å away from the heme iron, whereas the distance between methylene C14 carbon (analog of C12 in YC-17) and the heme iron is only 5.3 Å (Fig. 2-4A). Despite this distance differential that would appear to favor hydroxylation of the more proximal C14 atom, hydroxylation

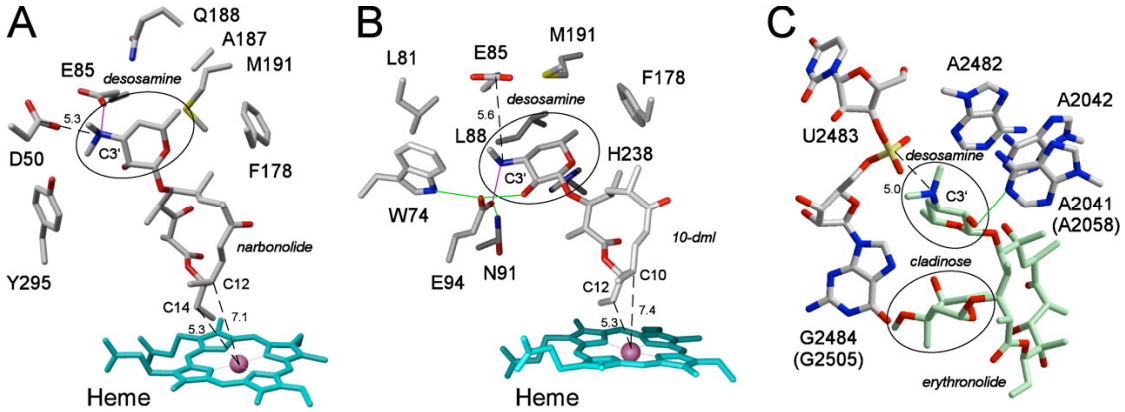


Figure 2-4. Desosamine anchoring. Desosamine interactions in narbomycin (A) and YC-17 (B) binding sites of PikC are shown. Amino acid residues within 4.0 Å from desosamine are also shown. C, erythromycin (in green) is shown bound to the large ribosomal subunit from eubacteria *Deinococcus radiodurans* (1JZY) (35). The phosphate group of G2484 (G2505) interacts with the dimethylamino group of desosamine. Nucleotide numbering according to the E. coli sequence is indicated by parentheses. Hydrogen bonds are highlighted in green, salt bridges are highlighted in magenta, and the distances are in Angstroms.

of narbomycin at this position occurs with very low yield (~40:1 ratio *in vivo*) (21). The structure suggests that unfavorable stereochemistry of the C-H bonds at C14 (pointing away from the oxygen scission site) is the basis for this, which could be due to rotation of the C14-C15 bond toward the heme iron, possibly to avoid steric interference with the side chains of amino acids Leu-93 and Thr-294 (Fig. 2-3A). In contrast, a C-H bond at C12 is positioned favorably toward the oxygen scission site. Whether hydroxylation is controlled by regio- and stereochemistry of the C-H bond, by distance to the heme iron, or by the iron-oxo reactive intermediate (or a combination of these variables) remains to be elucidated. It is also possible that Fe^{+3} reduction followed by oxygen binding might result in repositioning of narbomycin, as is indicated by NMR paramagnetic relaxation studies on the P450BM3-laurate complex, in which 6 Å movement of the substrate into the correct position for hydroxylation has been detected (35).

Alternatively, hydroxylation at the 7 Å remote allylic C12 atom in narbomycin may result from the iron-peroxo intermediate Fe^{+3}-OO- (or its protonated form) as an "active oxygen" as opposed to a widely accepted P450 oxidant oxo-ferryl radical,

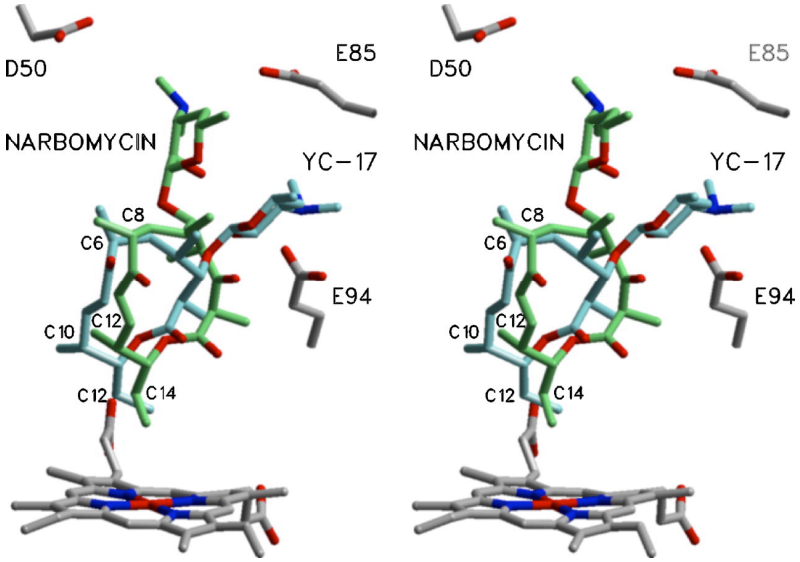


Figure 2-5. Two desosamine anchoring modes in PikC. A stereo view of overlaid YC-17 and narbomycin bound in the PikC active site is shown. Narbomycin is in green, and YC-17 is in cyan. The S-configuration observed for the C6 atom of YC-17 in the crystal structure conflicts with the R-configuration derived based on the methymycin, neomethymycin, narbomycin, and pikromycin structures (42-44). Formation of this unnatural stereoisomer may have been a result of adventitious epimerization at C6 via the enol intermediate during long-term storage of 10-deoxymethynolide.

compound I (36). This frequently discussed alternative pathway requires a homolytic O-O bond cleavage and is generally not thought to be involved in the normal monooxygenation process (37). Classical examples of the role of a nucleophilic iron-peroxo species in P450 catalysis are the oxidative decarbonylation of aldehydes and the final step of the aromatization reaction of androstenedione to estrone catalyzed by human aromatase (38). However, growing evidence supports the active role of peroxy-iron as an electrophilic oxidant in hydroxylation of C-H bonds in the hypersensitive cyclopropane probe (39) and epoxidation of olefins (40, 41). In narbomycin and YC-17, the allylic C-H bond is weaker than an alkanyl C-H by about 16 kcal/mol (36). In addition, the C10=C11 double bond in narbomycin and YC-17 is conjugated to the carbonyl group, which leads to a polarization of the double bond and draws electron density toward the oxygen atom away from the conjugated system. This serves to further weaken the allylic C-H bond, thus decreasing the barrier for its activation.

2.2.3 X-ray Structure of YC-17-bound PikC

The smaller 12-membered ring macrolide YC-17 is the second endogenous substrate of PikC (20). In contrast to narbomycin, it is hydroxylated *in vivo* and *in vitro* at two different positions (~1:1 ratio). Modification at allylic C10 results in methymycin, whereas modification at methylene C12 results in formation of neomethymycin (Fig. 2-1). In the co-crystal structure determined to 2.35 Å resolution, YC-17 is unambiguously positioned in one orientation within the active site, and its electron density is well defined (Fig. 2-3B). As observed with narbomycin, the macrolactone ring of YC-17 is bound almost entirely via hydrophobic interactions projecting the more hydrophilic surface of the lactone toward the I helix and is devoid of hydrogen bonding. The binding site for the 12-membered ring lactone overlaps the entire binding site for the 14-membered ring macrocycle plus an extra amino acid residue, Val-290. As observed for narbomycin, the C3' dimethylamino group of the desosamine moiety in YC-17 is sandwiched between two carboxyl groups (Glu-85 and Glu-94) in a different binding pocket (Fig. 2-5) that is more buried in the protein interior when compared with the narbomycin desosamine binding site (Figs. 2-3B and 2-4B). Although proximal Glu-94 (2.7 Å distance) provides a salt bridge, distal Glu-85 (5.8 Å distance) may potentially compensate for the partial positive

charge at the desosamine tertiary amine. Trp-74, Asn-91, and Glu-94 participate in a network of hydrogen bond contacts with the desosamine moiety (Fig. 2-4B). The BC-loop is more extensively involved in interactions with desosamine in this binding pocket, providing additional hydrophobic contacts via Trp-74, Leu-81, and Leu-88. The remaining desosamine-protein interactions are via Met-191 and His-238. Interestingly, the desosamine interaction pattern with amino acid residues in the PikC active site mimics interactions of desosamine with the bacterial ribosome drug target, although these involve specific RNA nucleotides and the phosphate backbone (Fig. 2-4C).

Both YC-17- and narbomycin-bound PikC differ slightly (~2 Å) in positioning of the FG-region, which is most dynamic in the absence of substrate. Overall r.m.s. deviation between the two complexes is 0.51 Å and is almost the same (0.6 Å) for carboxyl residues anchoring desosamine. This indicates that specific substrate-anchoring modes largely account for the diversity of PikC products rather than an induced fit mechanism.

The two YC-17 hydroxylation sites, allylic C10 and methylene C12 atoms, are within 7.5 and 5.3 Å from the heme iron, respectively (Figs. 2-3B and 2-4B). However, the C12-C13 bond of YC-17 is rotated slightly away from the heme iron, thus presenting a C-H bond at C12 for catalysis. This favorable stereochemical arrangement in combination with the 5.3 Å distance to the heme iron apparently leads to hydroxylation of YC-17 at C12, giving rise to neomethymycin (Fig. 2-1). However, the question remains whether the C10 position in YC-17 is hydroxylated from this same binding orientation and/or via the same mechanism since it is separated by > 7 Å from the heme iron atom.

2.2.4 Site-directed Mutagenesis in Desosamine Binding Pockets

To investigate directly the role of carboxyl functionalities in the desosamine binding pockets, site-directed mutagenesis was performed to replace Asp-50, Glu-85, and Glu-94 with alanine or glutamine/asparagines (Fig. 2-6). Functional activity of the mutant enzymes was assessed *in vitro*. Reaction products were analyzed by high pressure liquid chromatography. Substitution of the proximal carboxyl group forming a salt bridge to the C3' dimethylamino group, Glu-94 for YC-17 and Glu-85 for narbomycin, almost entirely

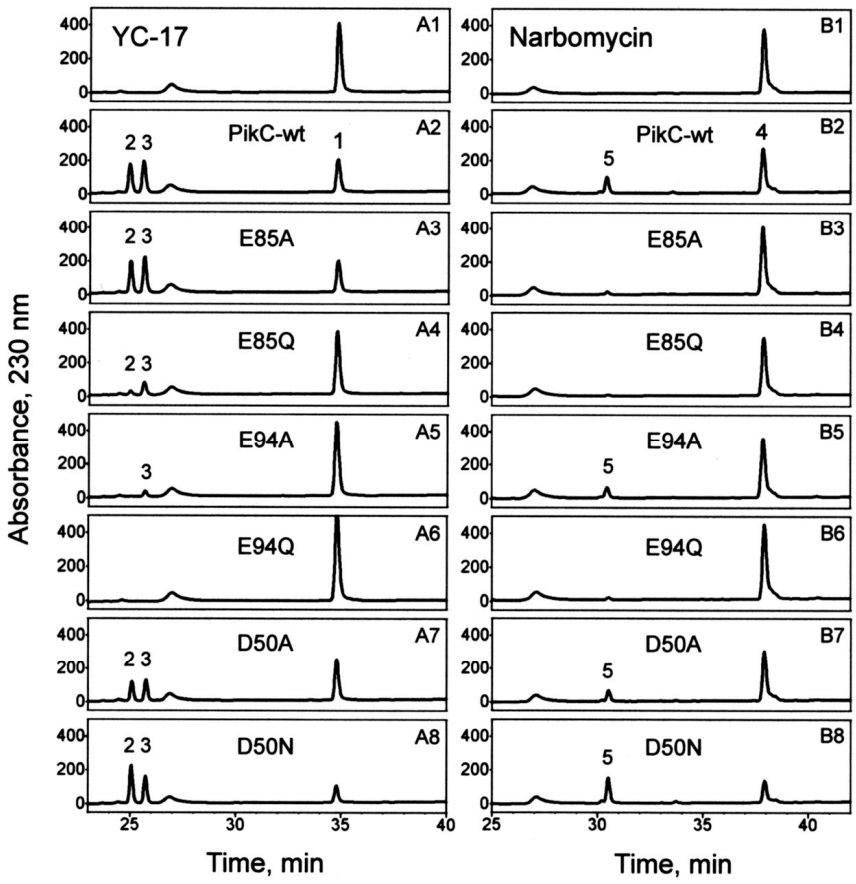


Figure 2-6. Functional activity of PikC mutants. High pressure liquid chromatography analyses of PikC-catalyzed reactions using YC-17 (A series) and narbomycin (B series) as substrate are shown. A1/B1, negative control in the absence of PikC. A2/B2, PikC wild type (PikC-wt). Mutants are used as indicated in the figure. Compound identities are as follows: 1, YC-17; 2, neomethymycin; 3, methymycin; 4, narbomycin; 5, pikromycin. Conversion of narbomycin at low efficiency is probably due to use of exogenous redox partners (e.g. spinach ferredoxin reductase) in P450 reconstitution assays.

abolishes conversion of the corresponding substrate, indicating that both charge and hydrogen-bonding capacities in the proximal position are essential for desosamine anchoring. This result is in agreement with a stringent requirement of hydrogen-bonding functionality in the C3' position in genetically modified sugars (14-17). Substitution of the carboxyl at distal amino acid groups (Glu-85 for YC-17 and Asp-50 for narbomycin) generated PikC mutants that retain partial to complete catalytic activity, indicating that the negative charge of the distal carboxyl is dispensable. Interestingly, the E85Q mutant, which lacks a negative charge but retains a hydrogen-bonding capability, shifts the product ratio of YC-17 hydroxylation toward methymycin (Fig. 2-6). This change in product profile does not rule out the possibility of interchangeable roles of proximal and

distal carboxyl groups in desosamine binding, which may result in realignment of YC-17 in the active site, bringing C10 closer to the heme iron. Substitution of a third residue not participating in binding, Asp-50 for YC-17 and Glu-94 for narbomycin, allows full (YC-17) or significant residual (narbomycin) conversion to occur. It is particularly interesting that the D50N mutant provides more efficient hydroxylation of both macrolide substrates when compared with native PikC. Double mutant E85A/E94A is catalytically incompetent toward both endogenous macrolactone substrates (not shown).

2.3 Summary

In summary, we have determined four x-ray structures for the PikC P450 monooxygenase that catalyzes hydroxylation of two macrolide substrates, leading to four dominant reaction products. Based on these data, a unique anchoring mechanism has been identified that involves a specific salt bridge of glutamate amino acid residues to the C3' dimethylamino group of the deoxysugar substituent. Site-directed mutagenesis of Glu-85, Glu-94, and Asp-50 amino acid residues in the substrate binding pocket were conducted to confirm the role of each in desosamine anchoring and their impact on the product ratio. Formation of a salt bridge between the protein carboxyl group and desosamine dimethylamino group is essential for catalytic conversion of each substrate. The data suggest that substrate tolerance and diverse product distribution occur from two specific anchoring orientations rather than induced fit mechanisms. The 12- and 14-membered ring macrolactone portions of both macrolide substrates bind in the active site, utilizing virtually the same set of protein-substrate interactions. In contrast, the desosamine group binds in two alternative binding pockets, each providing two carboxyl residues, one of which involves a salt bridge to a C3' dimethylamino group of desosamine. Moreover, the presence of two desosamine binding pockets provides a unique opportunity to develop rational design of unnatural glycosylated substrates for PikC. It also enables selection of one of the two desosamine binding sites by site-directed mutagenesis to shift reaction product distribution toward a single desirable bioactive metabolite. It is particularly intriguing that the "acceptor" carbon atoms are located in non-equivalent positions with respect to the heme iron. Based on the x-ray structure, there is a ~5 Å distance between the heme iron atom and the corresponding reactive

methylene C atom that results in hydroxylation of analogous positions, C12 in YC-17, leading to neomethymycin, and C14 in narbomycin, leading to neopikromycin (Fig. 2-1). The latter hydroxylation product is formed in very low amounts, which might be due to the unfavorable stereochemistry of the C14-C15 bond that points toward the heme iron with both C14-H bonds directed away from the oxygen scission site. Two other analogous allylic hydroxylation sites, C10 in YC-17 and C12 in narbomycin, are both within 7.5 Å from the heme iron. This is far enough to raise questions about the ability of these positions to be hydroxylated (i) from this same substrate orientation as opposed to an alternative, as yet uncovered substrate orientation or (ii) via an oxo-ferryl P450 intermediate as opposed to a peroxy-iron species.

2.4 Experimental Procedures

Expression and Purification of PikC — The PikC expression vector (20) was used for site-directed mutagenesis according to the QuikChange (Stratagene) protocol and for protein expression. The desired mutations were confirmed by DNA sequencing. Proteins were expressed in HMS174 (DE3) *E. coli* strain at 20 °C with an induction period of 20 h. Purification on nickel-nitrilotriacetic acid resin (Qiagen) was followed by flow-through chromatography on S-Sepharose (Amersham Biosciences) and binding to Q-Sepharose (Amersham Biosciences), from which PikC was eluted by a NaCl gradient, 0-0.5 M. Protein was concentrated to 1 mM by using a Centriprep concentrating device. Quality and quantity of purified protein were accessed by SDS-PAGE and CO-reduced difference spectra (45).

Crystallization and Diffraction Data Collection — Purified PikC with the His tag at 0.2 mM concentration was either subjected to crystallization alone or mixed with substrate at 2 mM concentration (1:10 ratio) and then subjected to crystallization. In all cases, prior to crystallization, the protein was incubated for 1 h on ice with 1 mM dithiothreitol to prevent dimerization via Cys-375, which is readily observed during prolonged storage of PikC. Co-crystallization with ligands, either inhibitor or substrates, resulted in crystals of different size and morphology grown by vapor diffusion hanging drop method from a variety of crystallization conditions as indicated in Table 2-1.

Ligand-free PikC crystals grew in the presence of 5 mM azole inhibitor 4-phenylimidazole (not found in the final crystal structure). All diffraction data were collected at 100-110 K at the Southeast Regional Collaborative Access Team (SER-CAT) 22ID and Structural Biology Center 19ID beamlines, Advanced Photon Source, Argonne National Laboratory, Argonne, IL. Glycerol mixed with the reservoir solution was used as a cryoprotectant at concentrations ranging from 18 to 25%, which have been determined empirically for each crystal type. The images were integrated and the intensities were merged by using HKL2000 (46).

Crystal Structure Determination — Molecular replacement was used to determine all crystal structures (Table 2-1). The ligand-free form of PikC was initially determined using the program CNS (47) and EryF (46% sequence identity) coordinates (48) as a search model. Crystal structures of YC-17- and narbomycin-bound PikC were similarly determined by molecular replacement using CNS (47), with ligand-free PikC coordinates as a search model. All reflections were used for refinement. The final atomic models (Table 2-1) were obtained after a number of iterations of refinement (CNS (47)) and manual model building with the program O (49). Ramachandran statistics (50) (Table 2-1) indicate no outliers with an exception of one residue, Pro-181, in the narbomycin-bound PikC structure (Protein Data Bank ID 2C7X). From 10 to 13 N-terminal residues, depending on the crystal form, are missing in all protein chains due to insufficient electron density in this region. Also, six BC-loop residues are missing in an open ligand-free form due to the same reason.

PikC Product Profile Analysis — The enzymatic conversions of YC-17 or narbomycin *in vitro* were carried out using a previously developed assay (20). The standard reaction contained 1 μM PikC (native enzyme or mutants), 0.5 mM YC-17 or narbomycin, 3.5 μM spinach ferredoxin, 0.01 units of spinach ferredoxin-$NADP^+$ reductase, and 1 mM NADPH in 200 μl of 50 mM Na-PO4 (pH 7.3), 1 mM EDTA, 0.2 mM dithiothreitol, 10% (v/v) glycerol. The reaction was terminated after 40 min of incubation at 30 °C by the addition of 300 μl of chloroform, and products were subsequently extracted, dried, dissolved in 120 μl of methanol, and analyzed by reverse-phase (C18) high pressure

liquid chromatography (Waters Corp., Milford, MA) by using a linear gradient (0-100%) of water : trifluoroacetic acid : triethylamine (99.8 : 0.1 : 0.1) /acetonitrile : trifluoroacetic acid : triethylamine (99.8 : 0.1 : 0.1) as a mobile phase at the flow rate of 1.0 ml/min.

2.5 References

1. Retsema, J., and Fu, W. (2001) *Int. J. Antimicrob. Agents* **18**, Suppl. 1, S3-S10
2. Furneri, P. M., and Nicoletti, G. (1991) *J. Chemother.* **3**, Suppl. 1, 24-27
3. Goldman, R. C., and Scaglione, F. (2004) *Curr. Drug Targets Infect Disord.* **4**, 241-260
4. Hutchinson, C. R. (1998) *Curr. Opin. Microbiol.* **1**, 319-329
5. Baltz, R. H. (1998) *Trends Microbiol.* **6**, 76-83
6. Menzella, H. G., Reid, R., Carney, J. R., Chandran, S. S., Reisinger, S. J., Patel, K. G., Hopwood, D. A., and Santi, D. V. (2005) *Nat. Biotechnol.* **23**, 1171-1176
7. Reeves, C. D. (2003) *Crit. Rev. Biotechnol.* **23**, 95-147
8. Rix, U., Fischer, C., Remsing, L. L., and Rohr, J. (2002) *Nat. Prod. Rep.* **19**, 542-580
9. Ackermann, G., and Rodloff, A. C. (2003) *J. Antimicrob. Chemother.* **51**, 497-511
10. Yonath, A. (2005) *Mol. Cells* **20**, 1-16
11. Auerbach, T., Bashan, A., and Yonath, A. (2004) *Trends Biotechnol.* **22**, 570-576
12. Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) *Annu. Rev. Biochem.* **68**, 219-253
13. Yoon, Y. J., Beck, B. J., Kim, B. S., Kang, H. Y., Reynolds, K. A., and Sherman, D. H. (2002) *Chem. Biol.* **9**, 203-214
14. Zhao, L., Sherman, D. H., and Liu, H.-w. (1998) *J. Am. Chem. Soc*. **120**, 10256-10257
15. Zhao, L., Que, N. L. S., Xue, Y., Sherman, D. H., and Liu, H.-w. (1998) *J. Am. Chem. Soc.* **120**, 12159-12160
16. Zhao, L., Ahlert, J., Xue, Y., Thorson, J. S., Sherman, D. H., and Liu, H.-w. (1999) *J. Am. Chem. Soc.* **121**, 9881-9882
17. Borisova, S. A., Zhao, L., Sherman, D. H., and Liu, H.-w. (1999) *Org. Lett.* **1**, 133-136
18. Tang, L., and McDaniel, R. (2001) *Chem. Biol.* **8**, 547-555
19. Xue, Y., and Sherman, D. H. (2001) *Metab. Eng.* **3**, 15-26
20. Xue, Y., Wilson, D., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Chem. Biol.* **5**, 661-667
21. Lee, S. K., Park, J. W., Kim, J. W., Jung, W. S., Park, S. R., Choi, C. Y., Kim, E. S., Kim, B. S., Ahn, J. S., Sherman, D. H., and Yoon, Y. J. (2006) *J. Nat. Prod.* **69**, 847-849
22. Zhang, Q., and Sherman, D. H. (2001) *J. Nat. Prod.* **64**, 1447-1450
23. Yano, J. K., Koo, L. S., Schuller, D. J., Li, H., Ortiz de Montellano, P. R., and Poulos, T. L. (2000) *J. Biol. Chem.* **275**, 31086-31092
24. Park, S. Y., Yamane, K., Adachi, S., Shiro, Y., Weiss, K. E., Maves, S. A., and Sligar, S. G. (2002) *J. Inorg. Biochem.* **91**, 491-501
25. Scott, E. E., He, Y. A., Wester, M. R., White, M. A., Chin, C. C., Halpert, J. R., Johnson, E. F., and Stout, C. D. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13196-13201
26. Ravichandran, K. G., Boddupalli, S. S., Hasemann, C. A., Peterson, J. A., and Deisenhofer, J. (1993) *Science* **261**, 731-736
27. Hilal, S., Karickhoff, S. W., and Carreira, L. A. (1995) *Quant. Struc.-Act. Relat.* **14**, 348-354
28. Goldman, R. C., Fesik, S. W., and Doran, C. C. (1990) *Antimicrob. Agents Chemother.* **34**, 426-431
29. Schlunzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001) *Nature* **413**, 814-821
30. Hansen, J. L., Ippolito, J. A., Ban, N., Nissen, P., Moore, P. B., and Steitz, T. A. (2002) *Mol. Cell* **10**, 117-128
31. Schlunzen, F., Harms, J. M., Franceschi, F., Hansen, H. A., Bartels, H., Zarivach, R., and Yonath, A. (2003) *Structure* **11**, 329-338
32. Berisio, R., Schluenzen, F., Harms, J., Bashan, A., Auerbach, T., Baram, D., and Yonath, A. (2003) *Nat. Struct. Biol.* **10**, 366-370

33. Nagano, S., Li, H., Shimizu, H., Nishida, C., Ogura, H., Ortiz de Montellano, P. R., and Poulos, T. L. (2003) *J. Biol. Chem.* **278**, 44886-44893
34. Nagano, S., and Poulos, T. L. (2005) *J. Biol. Chem.* **280**, 31659-31663
35. Modi, S., Sutcliffe, M. J., Primrose, W. U., Lian, L. Y., and Roberts, G. C. (1996) *Nat. Struct. Biol.* **3**, 414-417
36. de Visser, S. P., Ogliaro, F., Sharma, P. K., and Shaik, S. (2002) *J. Am. Chem. Soc.* **124**, 11809-11826
37. Sono, M., Roach, M. P., Coulter, E. D., and Dawson, J. H. (1996) *Chem. Rev.* **96**, 2841-2888
38. Meunier, B., deVisser, S. P., and Shaik, S. (2004) *Chem. Rev.* **104**, 3947-3980
39. Chandrasena, R. E., Vatsis, K. P., Coon, M. J., Hollenberg, P. F., and Newcomb, M. (2004) *J. Am. Chem. Soc.* **126**, 115-126
40. Vaz, A. D. N., McGinnity, D. F., and Coon, M. J. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 3555-3560
41. Jin, S., Makris, T. M., Bryson, T. A., Sligar, S. G., and Dawson, J. H. (2003) *J. Am. Chem. Soc.* **125**, 3406-3407
42. Anliker, R., Dvornik, D., Gubler, K., Heusser, H., and Prelog, V. (1956) *Helv. Chim. Acta.* **39**, 1785-1790
43. Djerassi, C., and Halpern, O. (1957) *J. Am. Chem. Soc.* **79**, 3926-3928
44. Djerassi, C., Halpern, O., Wilkinson, D. I., and Eisenbraun, E. J. (1958) *Tetrahedron* **4**, 369-381
45. Omura, T., and Sato, R. (1964) *J. Biol. Chem.* **239**, 2379-2385
46. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307-326
47. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., and Pannu, N. S. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905-921
48. Cupp-Vickery, J. R., and Poulos, T. L. (1995) *Nat. Struct. Biol.* **2**, 144-153
49. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110-119
50. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283-291

Notes:

This work has been published as "The structural basis for substrate anchoring, active site selectivity, and product formation by P450 PikC from *Streptomyces venezuelae*." Sherman, D. H., **Li, S.**, Yermalitskaya, L. V., Kim, Y., Smith, J. A., Waterman, M. R., and Podust, L. M. ***J. Biol. Chem.*** 2006, 281 (36), 26289-26297.

Author contributions:
Shengying Li, David H. Sherman, and Larissa M. Podust designed the experiments;
Shengying Li performed the biochemistry experiments and analyzed the crystal structures;
Larissa M. Podust, Liudmila V. Yermalitskaya, Youngchang Kim, Jarrod A. Smith, and Michael R. Waterman solved and analyzed the crystal structures.

Chapter 3
Analysis of Transient and Catalytic Desosamine-binding Pockets in Cytochrome P-450 PikC from *Streptomyces venezuelae*

3.1 Introduction

Macrolides are a large family of secondary metabolites belonging to the polyketide class of natural products generated by diverse genera of actinomycetes bacteria. The large macrolactone ring systems are derived from polymerization of simple carboxylic acid precursors catalyzed by modular polyketide synthases and often require further modification by specific tailoring enzymes (1) to acquire or enhance biological activity. The modular architecture of polyketide synthase gene clusters has led to the development of combinatorial biosynthetic approaches that aim to generate novel secondary metabolites through rational engineering of new combinations of polyketide synthase modules (2–4). Tailoring enzymes, including cytochrome P-450 monooxygenases (P-450) (2) are usually encoded within macrolide biosynthetic pathways (5). P-450 enzymes mainly serve to introduce hydroxyl or epoxide functional groups to nascent macrolactone structures or their glycosylated products (1, 3). To date, only three macrolide P-450 monooxygenases including EryF, EpoK, and PikC have been studied at both enzymatic and structural levels. Therefore, the principles of substrate recognition and regio- and stereochemical selectivity are just beginning to emerge for this intriguing group of biosynthetic enzymes.

Streptomyces venezuelae P-450 PikC displays a relatively broad substrate and regiospecificity compared with EryF (6) and EpoK (7). This characteristic combined with robust catalytic efficiency as a single component engineered biocatalyst (8) has motivated us to further its development as a prototype P-450 monooxygenase directed toward

metabolic engineering and synthetic chemical applications (3). Thus, PikC performs multiple hydroxylations of structurally variant macrolides including the 12-membered ring YC-17 and 14-membered ring narbomycin, leading to methymycin/neomethymycin and the natural ketolide antibiotic pikromycin, respectively (10) (Fig. 3-1). Ketolides are

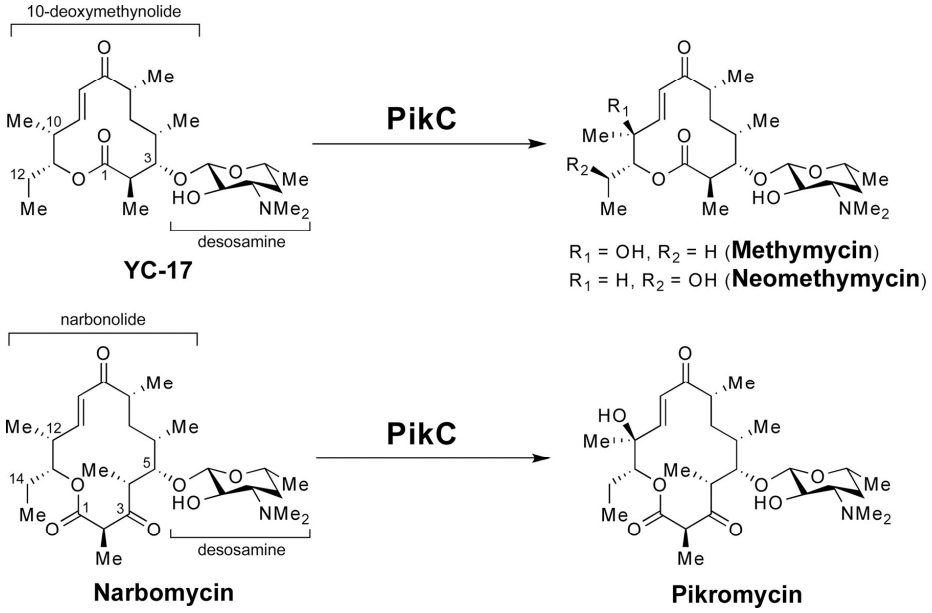


Figure 3-1. Structures of the PikC native substrates and their hydroxylated products.

macrolide derivatives characterized by a C-3 keto group that have received significant attention recently because of their enhanced activity against drug-resistant microbial pathogens (11).

Both endogenous PikC substrates are glycosylated with the 3-(dimethylamino)-3,4,6-trideoxy sugar desosamine that confers antibiotic activity to a number of macrolide antibiotics such as erythromycin, troleandomycin, mycinamicin, megalomicin (desosamine), tylosin, carbomycin, spiramycin (mycaminose, having an additional hydroxyl group at the C-4 position of the sugar ring), and a highly potent semisynthetic ketolide telithromycin (11–13). PikC catalyzes hydroxylation of variant macrolide substrates modified with altered sugar moieties through metabolic engineering (14–18) or with unnatural macrolactone ring systems (19, 20). PikC has also been shown to function effectively when immobilized on a microfluidic biochip (21), and when fused to a heterologous electron donor (8), the reductase domain of a self-sufficient P-450RhF from *Rhodococcus* sp. NCIMB 9784 (22).

Recent analysis of the x-ray crystal structures (23) revealed that YC-17 and narbomycin bind in the PikC active site via overlapping modes sharing the macrolactone-binding site and utilizing distinct desosamine binding regions, including buried and surface-exposed pockets, respectively. In both modes, the protonated dimethylamino group of desosamine binds between two negatively charged carboxyl groups of amino acid residues forming a salt bridge with the proximal (relative to the dimethyamino moiety) carboxyl and an ionic contact with the distal one. The triad of carboxylate residues Asp50, Glu85, and Glu94 located in the BC loop provides this set of interactions. Elimination of the negative charge at Glu85 or Glu94 by site-directed mutagenesis virtually inactivates (Glu94) or substantially reduces (Glu85) conversion of both substrates (23). In contrast, elimination of the surface-exposed negative charge at Asp50 via substitution of this residue with asparagine significantly enhances catalytic activity of PikC. To address the specific role for each desosamine-binding pocket, we analyzed the x-ray structures of the catalytically superior $PikC_{D50N}$ mutant co-crystallized with narbomycin or YC-17. In $PikC_{D50N}$, YC-17 adopts the same binding mode as observed previously in the wild type, with desosamine bound in the buried pocket. In contrast to the previously observed binding mode in wild type PikC, narbomycin was also found predominantly in the buried pocket in the corresponding D50N mutant form, suggesting the possibility of initial substrate recognition in the "surface-exposed site," with subsequent relocation to the catalytic "buried site." We herein report PikC substrate binding, enzyme mutagenesis, and kinetic data to support this hypothesis and provide evidence for kinetic control over substrate dissociation versus relocation to the PikC catalytic pocket.

3.2 Results

3.2.1 Substrate Binding

Binding affinities of PikC wild type and the mutants for the native substrates, YC-17 and narbomycin, and their aglycone precursors, 10-deoxymethynolide and narbonolide, were deduced from a low to high spin iron spectral shift, known as type I binding (24). Both narbonolide and 10-deoxymethynolide bound PikC with the significantly reduced affinities compared with the corresponding glycosylated substrates

Table 3-1. The values of dissociation constants (K_D) for natural PikC substrates and their aglycones

Enzyme	K_D		K_D	
	YC-17	Narbomycin	10-Deoxymethynolide	Narbononolide
	μM		mM	
$PikC_{wild\ type}$	98.9 ± 1.9	234.5 ± 15.0	2.1 ± 0.8	21.1 ± 12.2
$PikC_{D50N}$	27.2 ± 0.6	171.9 ± 14.7	0.9 ± 0.3	2.1 ± 1.5
$PikC_{E85Q}$	340.3 ± 31.5	351.6 ± 54.1	ND[b]	ND
$PikC_{E94Q}$	–[a]	1056.8 ± 212.4	ND	ND
$PikC_{D50N/E85Q}$	289.6 ± 28.8	459.0 ± 64.7	ND	ND
$PikC_{D50N/E94Q}$	–[a]	861.0 ± 567.4	ND	ND

[a] Binding curves (Fig. 1, *panels A4* and *A6*) could not be fitted.
[b] ND, not determined.

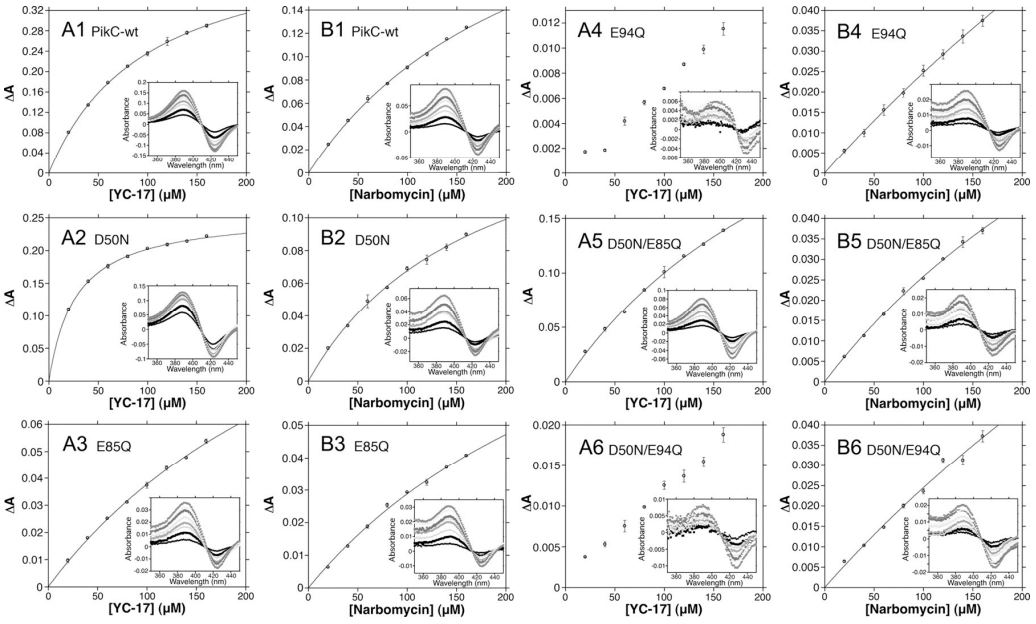


Figure 3-2. Binding of YC-17 and narbomycin to PikC. The concentration dependence of YC-17 (A panels) and narbomycin (B panels) binding deduced from the difference absorption changes (shown in insets) obtained from titration of PikC wild type or mutant enzymes, as indicated in each panel with increasing concentrations of substrates, as shown. Titration experiments were performed in 50 mM sodium phosphate, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol. The absence of the fitting curves in panels A4 and A6 indicates an unsuccessful fit caused by a lack of enzyme saturation. The values for the binding constants obtained from the curve fitting are presented in Table 3-1.

(Table 3-1), which may explain, at least in part, lack of hydroxylation by PikC. Binding affinity for YC-17 (K_D = 98.9 μM) was 2.3 times higher than narbomycin (K_D = 234.5 μM) (Fig. 3-2, panels A1 and B1). Compared with wild type enzyme, binding affinities of the catalytically more active $PikC_{D50N}$ increased about 4-fold for YC-17 (27.2 μM) and 1.3-fold for narbomycin (171.9 μM) (panels A2/B2). $PikC_{E85Q}$ had reduced binding affinities for both substrates compared with wild type enzyme (panels A3/B3). Binding

affinities for $PikC_{E94Q}$ were even further reduced (panels A4/B4), although the K_D values could not be determined for YC-17 because saturation of enzyme could not be approached because of both low binding affinity and limited substrate solubility. Additionally, D50N substitution did not significantly affect binding in the double mutants (Table 3-1 and Fig. 3-2, panels 5 and 6).

3.2.2. Kinetics of Narbomycin Binding

Binding of narbomycin to PikC wild type and the $PikC_{D50N}$ and $PikC_{E85Q}$ mutants was addressed by stopped flow UV-visible spectroscopy. Binding kinetics were biphasic and accurately described by a double exponential function revealing a fast first phase followed by a slow second one (Fig. 3-3). The kinetic rate of the first binding step (accounting for ~50% of the reaction amplitude) was fast and protein-dependent (Table

Table 3-2. Pre-steady state kinetic parameters of narbomycin binding

Enzyme	k_1	k_2
	s^{-1}	s^{-1}
$PikC_{wild\ type}$	703.0 ± 29.1	13.5 ± 0.08
$PikC_{D50N}$	822.0 ± 21.8	13.2 ± 0.07
$PikC_{E85Q}$	421.4 ± 18.4	13.7 ± 0.10

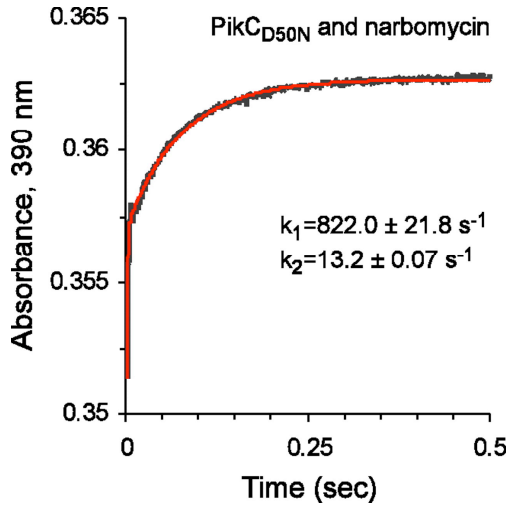


Figure 3-3. Kinetics of narbomycin binding. Typical kinetic transient of narbomycin binding (black) fitted to a double exponential function (red) is shown for $PikC_{D50N}$. Stopped flow experiments were conducted in 50 mM Tris-HCl, pH 7.5, 0.5% Me_2SO, and 0.5 mM EDTA.

3-2). The rate of the second step was slow, 13 s^{-1}, and virtually identical between the wild type and both mutant forms (D50N and E85Q) of PikC. The first binding rate of the catalytically superior $PikC_{D50N}$ was twice as fast as that of the mutant $PikC_{E85Q}$ with low activity: 822.8 ± 21.8 s^{-1} *versus* 421.4 ± 18.4 s^{-1}, respectively, with the wild type determined to be more similar to the D50N mutant at 703.0 ± 29.1 s^{-1}. We assume that the first binding step represents a pseudo first order reaction of the bimolecular encounter between PikC and narbomycin, whereas the second slow step may reflect monomolecular conformational adjustments of substrate position in the active site assisted by PikC dynamics (23). Because of small changes in the absolute spectra used to monitor the reaction and the limited solubility of narbomycin, we were unable to accurately estimate the k_{on} from the concentration dependence of the observed rates. In addition, the pseudo first order rate for $PikC_{D50N}$ may be underestimated because of instrument limitations for monitoring fast reactions.

Table 3-3. Data collection and refinement statistics

Parameter	Value(s)[a] or determination	
	Narbomycin (Protein Data Bank code 2VZM)	YC-17 (Protein Data Bank code 2VZ7)
Wavelength (Å)	1.11587	1.11587
Resolution (Å)	1.85	3.2
Unique reflections	87273	17248
Redundancy	3.9 (3.1)	4.3 (4.4)
Completeness (%)	98.7 (90.0)	100.0 (99.9)
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions: *a*, *b*, *c* (Å)	60.4, 109.5, 153.1	59.9, 109.3, 153.0
Molecules in asymmetric unit	2	2
Solvent content (%)	53	53
R_{sym} (%)[b]	6.5 (33.1)	12.8 (46.3)
I/σ	23.8 (2.3)	9.9 (2.7)
Refinement		
Reflections used in refinement	77423	16262
R_{cryst}/R_{free} (%)[c]	16.5/21.1	17.9/26.0
No. of atoms		
Protein	6233	6116
Prosthetic groups and ligands	158	150
Water	641	176
Wilson plot *B* values (Å^2)	22.4	50.6
Mean *B* factor (Å^2)	24.4	46.7
Protein	23.7	45.1
Heme	12.9	28.9
Substrate	32.7	41.7
Water	31.4	29.6
Root mean square deviation deviations		
Bond length (Å)	0.016	0.008
Bond angles (°)	1.6	1.3
Ramachandran statistics (%)[d]		
A	93.5/5.6/0.3/0.6	84.0/14.8/1.2/0.0
B	92.0/8.0/0.0/0.0	84.8/14.9/0.3/0.0

[a] The numbers in parentheses correspond to the highest resolution shell.
[b] $R_{sym} = \Sigma|I_i - \langle I\rangle|/\Sigma I_i$, where I_i is the intensity of the i^{th} observation and $\langle I\rangle$ is the mean intensity of reflection.
[c] $R_{cryst} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, calculated with the working reflection set. R_{free} is the same as R_{cryst} but calculated with the reserved reflection set.
[d] Program PROCHECK, portions of the protein residues in most favored/additional allowed/generously allowed/disallowed regions.

3.2.3 X-ray Structure of the $PikC_{D50N}$-YC-17 Complex

The crystal structure of the $PikC_{D50N}$-YC-17 complex was determined to a resolution of 3.2 Å (Table 3-3). Although determined at relatively low resolution, the electron density for YC-17 in both protein monomers in an asymmetric unit was well defined, showing no ambiguity for macrolide binding (Fig. 3-4). In both monomers (superimposable with the root mean square deviation of 0.43 Å), YC-17 was bound in the active site with desosamine positioned in the buried pocket anchored via the dimethylamino group to the side chain of Glu94 and Glu85, similar to what has been observed for wild type PikC (23).

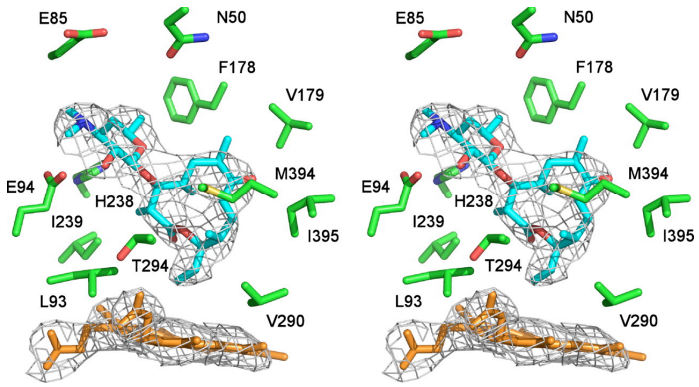


Figure 3-4. YC-17 binding in $PikC_{D50N}$. Stereo view of YC-17 (highlighted in cyan) bound in the active site of $PikC_{D50N}$ (Protein Data Bank code 2VZ7, chain A) surrounded by heme and amino acid side chains (green) within 4 Å plus N50 is shown. Fragments of the 2Fo – Fc electron density composite omit map contoured at 0.8 σ are shown as gray mesh. Heme is orange, nitrogen atoms are blue, and oxygen atoms are red. The image was generated using PYMOL program (30)

3.2.4 X-ray Structure of the $PikC_{D50N}$-Narbomycin Complex

Although they share the same space group and the unit cell dimensions as the $PikC_{D50N}$-YC-17 (Table 3-3), $PikC_{D50N}$-narbomycin crystals diffracted to a higher resolution (1.85 Å), revealing new details of the protein-substrate interactions. Although two protein monomers in the asymmetric unit had virtually identical overall conformations (root mean square deviation of 0.34 Å), the desosamine moiety of narbomycin was found in both alternative pockets. In the chain B form, desosamine was unambiguously bound in the buried pocket with the dimethylamino group positioned between the Glu94 and Glu85 side chains (Fig. 3-5A). This binding mode has been previously characterized for YC-17 but not for narbomycin (23). In the chain A form,

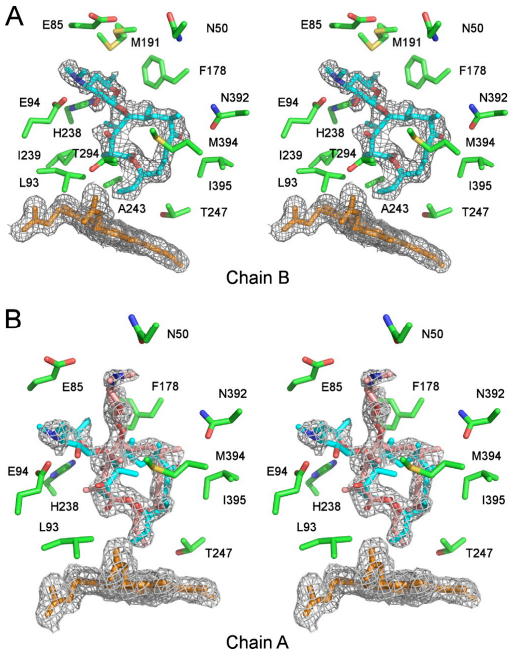


Figure 3-5. Narbomycin binding in PikC$_{D50N}$. Stereo views of narbomycin (cyan or pink) bound in the active site of PikC$_{D50N}$ (Protein Data Bank code 2VZM) (A) in the chain B and (B) in the chain A are shown. Selected amino acid side chains (green) are within 4 Å. In A, alternative conformations for Met191 are shown. Heme is in orange, nitrogen atoms are in blue, oxygen atoms are in red, and sulfur atoms in yellow. Fragments of the 2Fo – Fc electron density composite omit map contoured at 0.8 σ are shown as gray mesh.

electron density for the macrolactone ring of narbomycin was clearly defined, whereas electron density for desosamine was split between both buried and surface-exposed pockets, suggesting two alternative conformations (Fig. 3-5B), each refined with the occupancy of 0.5. The buried conformation (cyan in Fig. 3-5B) superimposed well between the A and B monomer, whereas the surface-exposed conformation in monomer A (pink) superimposed with the position observed for narbomycin in the wild type enzyme (23). Interestingly, in the surface-exposed pocket of PikC$_{D50N}$, the dimethylamino group of desosamine was located between the carboxyl group of Glu85 and the carboxamide group of Asn50 substituting for Asp50. In addition, subtle differences were observed for the macrolactone portion of narbomycin between the two binding modes (Fig. 3-5B).

3.2.5 Catalytic Activity of PikC Mutants

The catalytic activity of the double mutants PikC$_{D50N/E94Q}$ and PikC$_{D50N/E85Q}$ was addressed simultaneously with the panel of single PikC mutants characterized previously

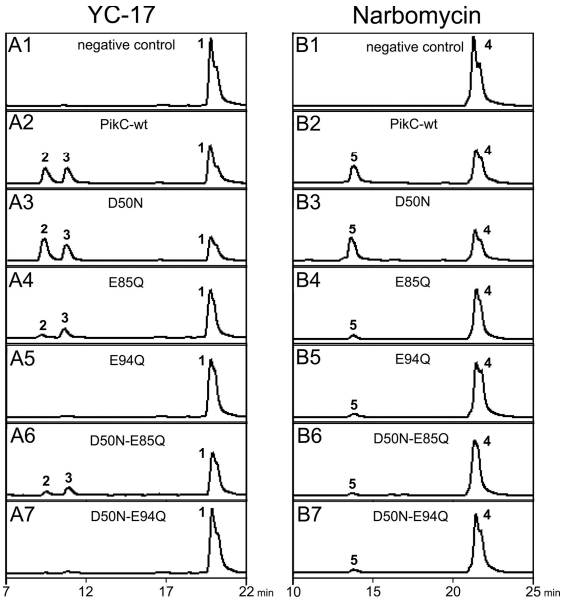


Figure 3-6. Catalytic activity of PikC mutants. High pressure liquid chromatography analysis of PikC-catalyzed reactions using YC-17 (A panels) and narbomycin (B panels) as substrate are shown. A1/B1, negative control in the absence of PikC. A2/B2, PikC wild type (PikC-wt). The mutants are used as indicated in the figure. The compound identities are as follows: 1, YC-17; 2, neomethymycin; 3, methymycin; 4, narbomycin; 5, pikromycin. The peak identity in each HPLC trace was determined by mass spectrometry and compared with authentic compounds with respect to HPLC retention time and UV spectrum. The area under each peak was quantified, and the results for each chromatogram are presented in Table 3-4.

Table 3-4. Conversion of YC-17 & narbomycin by different PikC forms as shown in Fig. 3-6

Enzyme	Panel in Fig. 5	Conversion of YC-17		Conversion of narbomycin (% of total)
		% of total	Product 2:3 ratio	
PikC-wt	*A2/B2*	40.0	0.81	35.0
D50N	*A3/B3*	60.0	1.15	44.4
E85Q	*A4/B4*	18.0	0.26	8.1
E94Q	*A5/B5*	0		5.4
D50N/E85Q	*A6/B6*	18.4	0.38	3.5
D50N/E94Q	*A7/B7*	0		5.3

(23). Enhanced catalytic activity for $PikC_{D50N}$ and attenuated catalytic activity for $PikC_{E94Q}$ were confirmed (Fig. 3-6 and Table 3-4). As before, residual catalytic conversion by $PikC_{E94Q}$ was observed for narbomycin, 5.4% versus 35.0% for the wild type, but not for YC-17 (Fig. 3-6, panels A5 and B5 and Table 3-4). In accord with previous results (23), E85Q substitution notably reduced but did not entirely eliminate both catalytic activities (Fig. 3-6, panels A4/B4, and Table 3-4). Significantly, introduction of the D50N substitution failed to rescue catalytic activity of the double

mutants (Fig. 3-6, compare panels 4 and 5 with panels 6 and 7, respectively), revealing the dominant impact of the E94 interaction for PikC function.

3.3 Discussion

Desosamine plays a fascinating parallel role both in targeting macrolide molecules to the 50 S bacterial ribosomal subunit (25, 26) and in tailoring of these important antibiotics by biosynthetic P-450 enzymes. In each case, the *N,N*-dimethylamino group of the sugar provides a key interaction for anchoring the macrolactone in its specific binding site. Antibiotic action is achieved through blockage of the peptide exit site resulting from precise macrolide interactions (27). Sugar anchoring also provides the basis for hydroxylation in the active site of P-450 monooxygenases PikC (23), EryK (28), MycCI, and MycG (29). As shown previously for PikC, the specific interactions between the protonated dimethylamino group of desosamine and the carboxyl-containing residues in the binding site are required for the catalytic activity. Non-glycosylated substrate precursors narbonolide and 10-deoxymethynolide, are very weakly bound with K_D values in the millimolar range (Table 3-1) and do not serve as substrates for PikC. Thus, given the importance of negative charges associated with Glu85 and Glu94, the enhancement of catalytic activity in $PikC_{D50N}$ was surprising.

In a search for determinants of improved catalytic function of $PikC_{D50N}$, we have determined the x-ray structures for this mutant complexed with the native substrates YC-17 and narbomycin. Although YC-17 showed no binding ambiguity (Fig. 3-4), narbomycin adopted two alternative conformations in the active site of the mutant enzyme (Figs. 3-5 and 3-7). Moreover, for the 14-membered ring macrolide, 25% of protein molecules in the $PikC_{D50N}$-narbomycin crystals had desosamine bound in the surface-exposed pocket, whereas the remaining molecules revealed desosamine associated with the buried pocket. The latter binding mode has not been observed for narbomycin previously. In the surface-exposed pocket of $PikC_{D50N}$, the dimethylamino group of desosamine formed a salt bridge to the proximal Glu85 residue as it does in the wild type. However, the original ionic contact with the distal Asp50 was no longer available because of removal of the negative charge through asparagine substitution. We

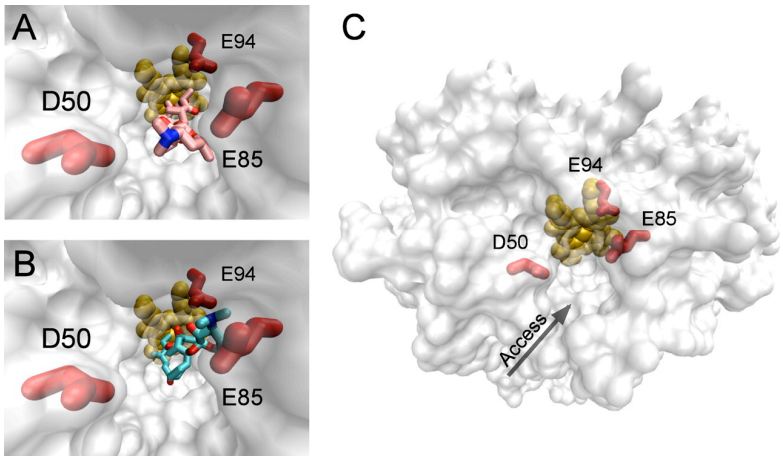


Figure 3-7. Transient and catalytic desosamine-binding pockets. (A) Narbomycin in the surface-exposed pocket is highlighted in pink. (B) Narbomycin in the buried pocket is highlighted in cyan. Zoomed images in (A) and (B) are shown in perspective. (C) Semi-transparent surface of the $PikC_{D50N}$ shows overall disposition of the negatively charged residues. Heme (orange) is shown as van der Waals' spheres. N atoms of desosamine are blue, O atoms red. Carboxylic residues Asp50, Glu85, and Glu94 are shown as red sticks. Images were generated using VMD program (9).

surmised that the loss of an electrostatic component might weaken protein-desosamine interactions in the surface-exposed pocket and facilitate desosamine relocation to the buried position, suggesting a catalytic role for this site. If so, the $PikC_{D50N/E94Q}$ double mutant should remain sensitive to the E94Q substitution. Alternatively, if D50N substitution favors hydroxylation from the surface-exposed pocket (given that the macrolactone is positioned similarly in respect to the iron center in both binding modes (Fig. 3-5B)), $PikC_{D50N/E94Q}$ should be insensitive to the E94Q mutation. To test both hypotheses, the catalytic activities of two double mutants, $PikC_{D50N/E94Q}$ and $PikC_{D50N/E85Q}$, were analyzed. The D50N substitution could not rescue or even affect the functional activity of the double mutants (Fig. 3-6 and Table 3-4), supporting the facilitated relocation to the buried pocket as a factor that favors catalysis.

Both the PikC mutagenesis studies and the crystal structure analysis suggest a two-step substrate binding mechanism whereby desosamine of the macrolide substrate initially binds to the surface-exposed pocket (Fig. 3-7A) and then relocates to the catalytic buried site (Fig. 3-7B). Because narbomycin is comprised of a 14-membered macrolactone ring, it could have a higher energetic barrier for conversion from one site to another and therefore was trapped entirely in the surface-exposed pocket in the crystals of

the wild type enzyme (23) and partially in the crystals of the $PikC_{D50N}$ mutant. The hampered relocation to the buried pocket might explain the reduced binding affinity of narbomycin compared with YC-17: 234.5 ± 15.0 μM *versus* 98.9 ± 1.9 μM (Table 3-1), because the substrate is presumed to be more easily released from the surface-exposed pocket than the buried one. Accordingly, the affinity of $PikC_{D50N}$ is higher toward both substrates because of facilitated access for desosamine to the buried pocket, which reduces dissociation of substrate from the surface-exposed site. Thus, the residual level of the $PikC_{E94Q}$ catalytic activity toward narbomycin (Fig. 3-6, panel B5, and Table 3-4) is readily explained by marginal hydroxylation from the surface-exposed pocket. In contrast, YC-17 is not hydroxylated by this mutant (Fig. 3-6, panel A5) because of rapid relocation to the buried pocket, where catalytic activity depends entirely on the interactions with Glu94.

The residual level of hydroxylation demonstrated by $PikC_{E85Q}$ toward both substrates (Fig. 3-6, panels A4 and B4, and Table 3-4) suggests that relocation of each macrolide is affected in this mutant. Glu85 is centered between the two pockets (Fig. 3-7) and therefore may serve as a pivot in the transition of desosamine, because the latter may remain electrostatically "tethered" to the Glu85 carboxylate group during the relocation. Because introduction of the D50N substitution could not rescue catalysis (Fig. 3-6, panels A6 and B6, and Table 3-4), the residual hydroxylation by $PikC_{E85Q}$ is likely to occur from the surface-exposed pocket as well.

Because of the dynamic constraints imposed upon narbomycin, we were able to assess kinetic binding rates between the PikC mutants by UV-visible spectroscopy combined with stopped flow kinetic studies. The kinetic trends were consistent with the equilibrium binding and functional data, because the catalytically superior $PikC_{D50N}$ bound narbomycin twice as fast as the poorly active $PikC_{E85Q}$ mutant enzyme under the same experimental conditions (Table 3-1). This observation suggests that the first kinetic rate likely reflects formation of the final buried complex rather than a transient one, and hence, transition from the surface-exposed to the buried site is the rate-limiting step of the substrate binding reaction.

The regiospecificity of PikC is a particularly interesting characteristic of this enzyme. In all PikC/substrate co-crystal structures reported to date, the allylic carbon

hydroxylation center in YC-17 (C-10) and narbomycin (C-12) is >7 Å away from the iron. The allylic C-H bond is the predominant hydroxylation site in narbomycin and one of two equally hydroxylated sites in YC-17, the second (methylene at C-12) positioned within 5 Å from the iron. Differences in the ratio of the YC-17 hydroxylation products between $PikC_{D50N}$, $PikC_{E85Q}$, and the wild type (Fig. 3-6 and Table 3-4) may be explained by dynamic variations in substrate binding making one site more favorably positioned for hydroxylation. At the same time, despite its 5 Å proximity to the iron reaction center, we have not observed hydroxylation of the methylene C-14 site in narbomycin by any PikC form *in vitro*, suggesting the possibility of additional steric or electronic factors that shift hydroxylation to the allylic C-12 site exclusively.

Collectively, our data indicate that each of the triad carboxylic acid residues interacting with the positive charge of desosamine plays a particular role in PikC catalysis; surface-exposed Asp50 appears to function as a gate for substrate access to the active site through kinetic control over substrate dissociation in solution from the transient site versus transition to the catalytic buried site. Glu85 centered between two desosamine-binding pockets may serve as a pivot in desosamine relocation. Finally, Glu94 evidently plays a major role in tuning substrate orientation for effective catalysis.

3.4 Experimental Procedures

Preparation of Protein Samples — Expression vectors for PikC wild type and mutants were used to transform *Escherichia coli* strain HMS174 (DE3) to express and subsequently purify proteins according to the protocol described elsewhere (23). The double mutants $PikC_{D50N/E94Q}$ and $PikC_{D50N/E85Q}$ represent new constructs prepared in this study. The quality of the purified proteins was assessed by the SDS-PAGE and UV-visible spectroscopy, and the concentration was determined at 450 nm from the difference spectra between the carbon monoxide-bound ferrous and water-bound ferric forms using the extinction coefficient of 91,000 M^{-1} cm^{-1} (31).

Preparation of Substrates — Substrates YC-17 and narbomycin were obtained from the pikC knock-out strain *S. venezuelae* AX-906 (10). Alternatively, narbomycin was harvested from fermentation culture of *Streptomyces narbonensis* NRRL B-1680 in

soluble complete medium, containing 15 g of soluble starch, 20 g of soytone, 1.5 g of yeast extract, 10.5 g of MOPS, 0.1 g of $CaCl_2$ at pH 7.2, per 1 liter of deionized water.

Equilibrium Binding Assay — Spectroscopic substrate binding assay was carried out at room temperature using a UV-visible spectrophotometer 300 Bio (Cary). Protein dissolved in 50 mM sodium phosphate, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol at concentrations ranging from 1 to 2 μM was titrated with the substrate dissolved in Me_2SO (20 mM) in 1-μl aliquots. The same amounts of Me_2SO alone were added to the protein in the reference cuvette followed by recording of the difference spectra. Absorbance differences ΔA ($A_{peak\ 389\ nm} - A_{trough\ 422\ nm}$) were plotted *versus* substrate concentration, and data from duplicated experiments were fitted to the hyperbolic function $\Delta A = Amax \bullet S/(K_D + S)$, where S is the total ligand concentration, Amax is the maximal absorption shift at saturation, and K_D is the apparent dissociation constant for the enzyme-ligand complex.

Catalytic Activity Assay — Enzymatic conversion of YC-17 or narbomycin *in vitro* were performed using a previously developed assay (10). The standard reaction contained 1 μM PikC (wild type or mutant form), 0.5 mM YC-17 or narbomycin, 3.5 μM spinach ferredoxin, 0.01 units of spinach ferredoxin-$NADP^+$ reductase, and 1 mM NADPH in 100 μl of 50 mM sodium phosphate, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% (v/v) glycerol. The reaction was terminated after 40 min of incubation at 30 °C by the addition of 3 x 300 μl of chloroform. The extractions were subsequently combined, dried, and dissolved in 120 μl of methanol and subjected to the reverse phase HPLC using a X-Bridge C18, 5 μm, 250-mm column (Waters Corporation) in a linear gradient (30–60%) of acetonitrile in 10 mM ammonium acetate, pH 8.1, at the flow rate of 1.0 ml/min. Detection was at 238 nm.

Pre-steady State Binding Assay — Stopped flow kinetic analysis experiments were conducted at 23 °C using a Hi-Tech Scientific instrument equipped with a photodiode array detector and controlled by the KinetAsyst software (Bradford on Avon, UK). Protein and narbomycin solutions were prepared in 50 mM Tris-HCl, pH 7.5, and 0.5

mM EDTA. Narbomycin stock solution was in Me_2SO; therefore, solvent concentration was adjusted in both syringes and maintained constant (0.5%) throughout the experiments. Protein concentration after mixing was 10 μM. To maintain pseudo first order reaction conditions, all of the measurements were carried out in 10-fold excess of narbomycin over the protein. Kinetic transients were recorded for 0.5 s at 390 nm corresponding to the peak of the difference spectra between the high spin substrate-bound and the low spin substrate-free PikC forms. Multiple kinetic traces for each sample were measured and averaged prior to data analysis. The data points from 2 to 500 ms were fitted to a standard double exponential function using SPECFIT software.

Crystallization and Data Collection — $PikC_{D50N}$ from the 1 mM frozen stock stored at –80 °C in 20 mM Tris-HCl, pH 7.5, 200 mM NaCl, and 0.5 mM EDTA was mixed with either YC-17 or narbomycin, both dissolved in Me_2SO at 50 mM stock concentration, to final concentrations of 0.2 mM for protein and 2 mM for the ligand. The mixtures were subjected to the automated screening of crystallization conditions (hanging drop crystallization protocol) using a nanoliter drop setter Mosquito (TTP LabTech) and commercial high throughput screening kits purchased from Hampton Research and Qiagen. Crystallization conditions generating crystals were identified and optimized further in 24-well crystallization plates. The crystals used for x-ray data collection were obtained at 23 °C for the $PikC_{D50N}$-YC-17 complex from 20% polyethylene glycol 4000, 0.1 M Bis-Tris, pH 6.5, and 0.2 M Li_2SO_4, and for the $PikC_{D50N}$-narbomycin complex from 16% polyethylene glycol 8000, 0.1 M sodium cacodilate, pH 6.5, and 0.15 M Li_2SO_4. Prior to data collection, the crystals were cryo-protected by plunging into a drop of reservoir solution supplemented with 20% glycerol and flash frozen in the liquid nitrogen. Diffraction data were collected at 100–110 K at beamline 8.3.1, Advanced Light Source, Lawrence Berkeley National Laboratory. The images were integrated, and the intensities were merged by using the HKL2000 software suite (32).

Structure Determination and Refinement — Structures were determined by molecular replacement using the CCP4 (33) program suit and an A chain of the YC-17-bound PikC (Protein Data Bank code 2C6H) as a search model. Model building was performed with

the programs COOT (34, 35) and O (36) and refined using REFMAC5 (33, 37). Simulated annealing was performed, and electron density composite omit maps were generated using CNS (38). The data collection and refinement statistics are shown in Table 3-3.

3.5 References

1. Rix, U., Fischer, C., Remsing, L. L., and Rohr, J. (2002) *Nat. Prod. Rep.* **19**, 542–580
2. Hutchinson, C. R. (1998) *Curr. Opin. Microbiol.* **1**, 319–329
3. Walsh, C. T. (2002) *Chembiochem.* **3**, 125–134
4. Sherman, D. H. (2005) *Nat. Biotechnol.* **23**, 1083–1084
5. Podust, L. M., Bach, H., Kim, Y., Lamb, D. C., Arase, M., Sherman, D. H., Kelly, S. L., and Waterman, M. R. (2004) *Protein Sci.* 13, 255–268
6. Andersen, J. F., Tatsuta, K., Gunji, H., Ishiyama, T., and Hutchinson, C. R. (1993) *Biochemistry* **32**, 1905–1913
7. Ogura, H., Nishida, C. R., Hoch, U. R., Perera, R., Dawson, J. H., and Ortiz de Montellano, P. R. (2004) *Biochemistry* **43**, 14712–14721
8. Li, S., Podust, L. M., and Sherman, D. H. (2007) *J. Am. Chem. Soc.* **129**, 12940–12941
9. Humphrey, W., Dalke, A., and Schulten, K. (1996) *J. Mol. Graph.* **14**, 33–38
10. Xue, Y., Wilson, D., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Chem. Biol.* **5**, 661–667
11. Ma, Z., and Nemoto, P. A. (2002) *Curr. Med. Chem.* **1**, 15–34
12. Ackermann, G., and Rodloff, A. C. (2003) *J. Antimicrob. Chemother.* **51**, 497–511
13. Blasi, F., Cazzola, M., Tarsia, P., Aliberti, S., Baldessari, C., and Valenti, V. (2006) *Future Microbiol.* **1**, 7–16
14. Zhao, L., Sherman, D. H., and Liu, H.-w. (1998) *J. Amer. Chem. Soc.* **120**, 10256–10257
15. Zhao, L., Que, N. L. S., Xue, Y., Sherman, D. H., and Liu, H.-w. (1998) *J. Am. Chem. Soc.* **120**, 12159–12160
16. Zhao, L., Ahlert, J., Xue, Y., Thorson, J. S., Sherman, D. H., and Liu, H.-w. (1999) *J. Amer. Chem. Soc.* **121**, 9881–9882
17. Borisova, S. A., Zhao, L., Sherman, D. H., and Liu, H. W. (1999) *Org. Lett.* **1**, 133–136
18. Tang, L., and McDaniel, R. (2001) *Chem. Biol.* **8**, 547–555
19. Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) *Annu. Rev. Biochem.* **68**, 219–253
20. Yoon, Y. J., Beck, B. J., Kim, B. S., Kang, H. Y., Reynolds, K. A., and Sherman, D. H. (2002) *Chem. Biol.* **9**, 203–214
21. Srinivasan, A., Bach, H., Sherman, D. H., and Dordick, J. S. (2004) *Biotechnol. Bioeng.* **88**, 528–535
22. Nodate, M., Kubota, M., and Misawa, N. (2006) *Appl. Microbiol. Biotechnol.* **71**, 455–462
23. Sherman, D. H., Li, S., Yermalitskaya, L. V., Kim, Y., Smith, J. A., Waterman, M. R., and Podust, L. M. (2006) *J. Biol. Chem.* **281**, 26289–26297
24. Schenkman, J. B., Remmer, H., and Estabrook, R. W. (1967) *Mol. Pharmacol.* **3**, 113–123
25. Hansen, J. L., Ippolito, J. A., Ban, N., Nissen, P., Moore, P. B., and Steitz, T. A. (2002) *Mol. Cell* **10**, 117–128
26. Auerbach, T., Bashan, A., and Yonath, A. (2004) *Trends Biotechnol.* **22**, 570–576
27. Jenni, S., and Ban, N. (2003) *Curr. Opin. Struct. Biol.* **13**, 212–219
28. Lambalot, R. H., Cane, D. E., Aparicio, J. J., and Katz, L. (1995) *Biochemistry* **34**, 1858–1866
29. Anzai, Y., Li, S., Chaulagain, M. R., Kinoshita, K., Kato, F., Montgomery, J., and Sherman, D. H. (2008) *Chem. Biol.* **15**, 950–959
30. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA
31. Omura, T., and Sato, R. (1964) *J. Biol. Chem.* **239**, 2379–2385
32. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326

33. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
34. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30**, 1022–1025
35. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
36. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110–119
37. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
38. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., and Pannu, N. S. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921

Notes:

This work has been published as "Analysis of transient and catalytic desosamine-binding pockets in cytochrome P-450 PikC from *Streptomyces venezuelae*." **Li, S.**, Ouellet, H. Sherman, D. H., and Podust, L. M. ***J. Biol. Chem.*** 2009, 284 (9), 5723-5730.

Author contributions:
Shengying Li, David H. Sherman, and Larissa M. Podust designed the experiments;
Shengying Li performed the biochemistry experiments and analyzed the crystal structures;
Larissa M. Podust, and Hugues Ouellet solved and analyzed the crystal structures.

Chapter 4
Engineering and Analysis of a Self-Sufficient Biosynthetic Cytochrome P450 PikC Fused to the RhFRED Reductase Domain

4.1 Introduction

Cytochrome P450 enzymes (P450s) are highly attractive biocatalysts due to their ability to catalyze a variety of regio- and stereospecific oxidation reactions of complex organic compounds. These reactions occur under mild conditions by taking advantage of the two-electron activated dioxygen that is often challenging in organic synthesis (1, 2). To activate molecular oxygen, redox partners are required to sequentially transfer two reducing equivalents from NAD(P)H to P450 (3). Classically, there are two major redox partner systems, including an FAD containing reductase with a small iron−sulfur (Fe_2S_2) redoxin for most bacterial and mitochondrial P450s (class I) and a single FAD/FMN containing flavoprotein for eukaryotic microsomal P450s (class II) (4, 5). The inherent requirement of cytochome P450s for separate protein partner(s) significantly limits their application in biotechnology.

The discovery of the first self-sufficient $P450_{BM3}$, which is naturally fused to a eukaryotic-like reductase, represents an effective solution to this limitation (6, 7). The fusion nature of this enzyme dramatically improves electron-transfer efficiency and coupling with the oxidative process, enabling it to be the most efficient P450 enzyme characterized to date (8). Based upon the self-sufficiency of this naturally fused enzyme, a number of engineered proteins of diverse eukaryotic P450s bearing a reductase domain from $P450_{BM3}$ have been generated with *in vitro* activities (9, 10). This provides ready access to the great catalytic versatility of the membrane-bound eukaryotic P450s. In contrast, the biosynthetic P450s (class I) lack such a universal reductase that can be used

to engineer diverse self-sufficient P450s for either functional identification or potential industrial application.

Recently, a new class of self-sufficient cytochrome P450s exemplified by $P450_{RhF}$ from *Rhodococcus* sp. NCIMB 9784 was discovered to be naturally fused to a novel FMN/Fe2S2 containing reductase partner (11, 12). Although the physiological function of $P450_{RhF}$ remains unknown, its reductase domain (RhFRED), which is similar to the phthalate family of dioxygenase reductases, is capable of transferring electrons from NADPH to the heme domain of the monooxygenase, supporting 7-ethoxycoumarin dealkylation activity (13, 14). Moreover, recent reports from Misawa *et al.* demonstrated that this reductase domain could be used to reconstitute the catalytic activities of various class I P450s *in vivo* through expression of corresponding genes fused to RhFRED in *Escherichia coli* cells (15, 16). This suggests that RhFRED might be developed into a generally effective redox partner for biosynthetic bacterial P450s. However, the lack of corresponding *in vitro* data could not unambiguously exclude in trans involvement of additional cellular redox partners.

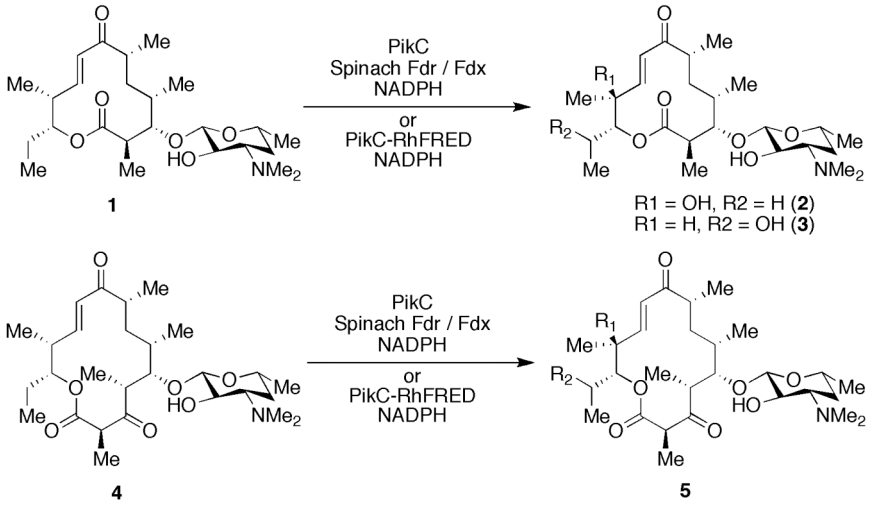


Figure 4-1. Major Physiological Reactions Catalyzed by PikC

4.2 Results and Discussion

Herein, we describe the first *in vitro* characterization of a single component bacterial biosynthetic cytochrome P450 fused to RhFRED that demonstrates high catalytic efficiency. The PikC cytochrome P450 in this study is involved in the

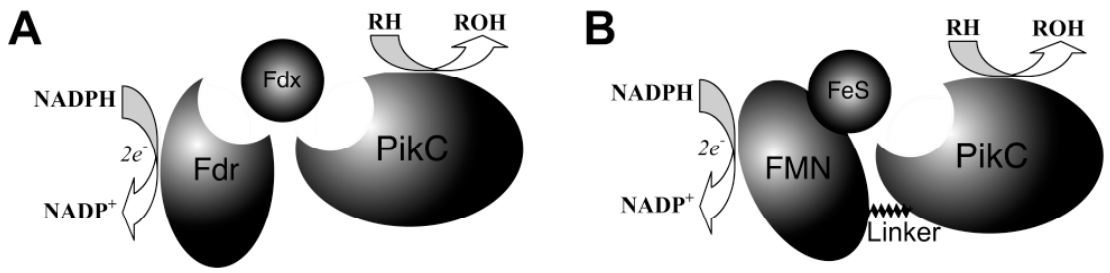


Figure 4-2. Two redox partner systems (electron transfer pathways) used in this study for PikC. A, three-components system; B, one-component RhFRED system.

pikromycin biosynthetic pathway of *Streptomyces venezuelae* (17). PikC catalyzes the final hydroxylation step toward both the 12-membered ring macrolactone YC-17 (**1**) and the 14-membered ring macrolactone narbomycin (**4**) to produce methymycin/neomethymycin (**2**/**3**) and pikromycin (**5**) as major products (Fig. 4-1) (18, 19). Recently, we elucidated the structural basis for the remarkable substrate flexibility by analyzing ligand-free and substrate-bound structures of PikC (20). However, since the native redox partner of PikC remains unknown, its *in vitro* activity has depended on expensive spinach ferredoxin reductase (Fdr) and ferredoxin (Fdx) (Fig. 4-2A), as are many other biosynthetic P450s (21-23). To investigate an alternative electron-transfer pathway mimicking the fusion organization in $P450_{RhF}$ (Fig. 4-2B), *pikC* was linked to the RhFRED gene including the native 16 amino acid linker sequence (Fig. 4-S1). The hybrid sequence was cloned into pET28b(+), and overexpressed in *E. coli* BL21 (DE3) to generate N-terminal His_6-tagged PikC-RhFRED (Fig. 4-S2). After Ni-NTA chromatography, the purified red-colored recombinant P450 displayed (upon reduction) the signature peak at 450 nm in the CO-difference spectrum (Fig. 4-S3). Interestingly, gel filtration chromatography (Fig. 4-S4) indicated that PikC-RhFRED predominantly dimerizes in storage buffer solution containing 0.2 mM dithioerythritol (DTE). In contrast, wild type (wt) PikC was shown to be monomeric under the same conditions. It was thus unclear whether the intermonomer electron transfer could occur in the dimeric PikC-RhFRED as in $P450_{BM3}$ (24).

We next tested the ability of PikC-RhFRED to hydroxylate **1** and **4** *in vitro* when provided electron donor NADPH. We were gratified to observe that this chimeric protein showed significantly improved catalytic activity compared to wt PikC in the presence of

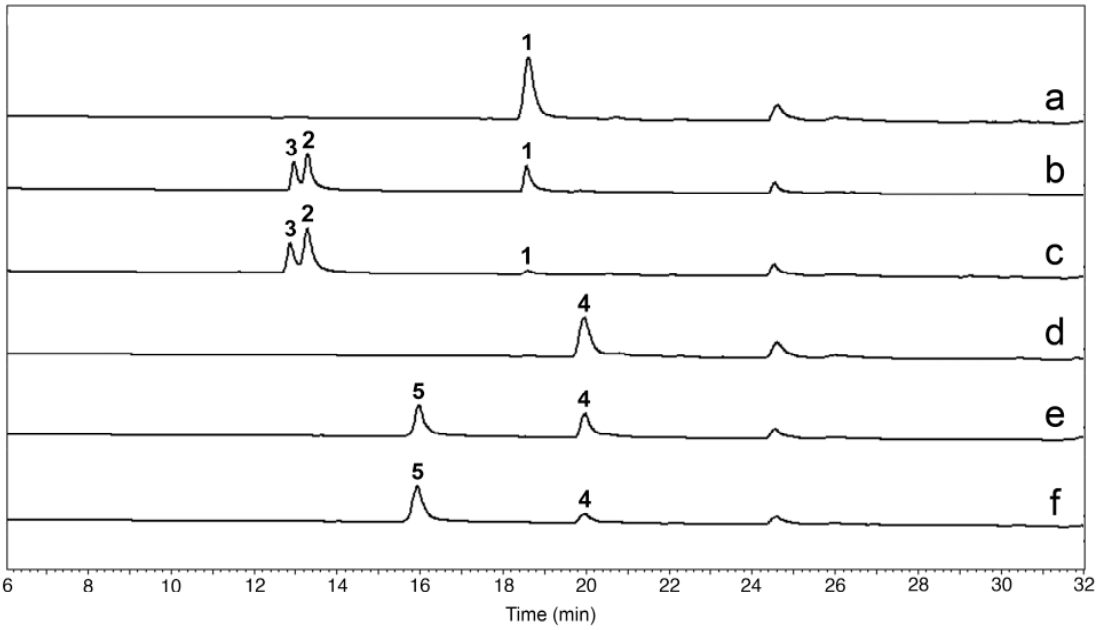


Figure 4-3. HPLC analysis of reactions (1 h) catalyzed by wt PikC and fusion enzyme PikC-RhFRED. a) Negative control of **1** in absence of P450. b) **1** with wt PikC in presence of Fdr, Fdx, and NADPH. c) **1** with PikC-RhFRED in presence of only NADPH. d) Negative control of **4** in absence of P450. e) **4** with wt PikC in presence of Fdr, Fdx, and NADPH. f) **4** with PikC-RhFRED in presence of only NADPH.

exogenous redox partners (spinach Fdr and Fdx), producing higher yields of **2**/**3** and **5** under identical reaction conditions (Fig. 4-3). This result unambiguously confirms that PikC-RhFRED is a self-sufficient P450 enzyme. Interestingly, we also constructed the pET21b(+)-*pikC-RhFRED* and obtained the purified C-terminal His_6-tagged PikC-RhFRED. This protein showed a similar CO-difference spectrum as its N-terminal His_6-tagged counterpart. However, it lacks catalytic activity, which is consistent with a similar C-terminal His_6-tagged form of original P450RhF (13, 14). This provides additional evidence for the importance of the C terminus of RhFRED for electron transfer.

As mentioned above, one benefit of the fusion arrangement is that the covalent linkage presumably stabilizes the interaction between the P450 and redox partner, thus enhancing electron-transfer efficiency. As such, one would expect this to improve the catalytic activity in terms of k_{cat}, whereas the substrate specificity would not be changed significantly (25, 26). To test whether this also applies to PikC-RhFRED, we first determined the substrate binding affinity of **1** and **4** toward both PikC and PikC-RhFRED. As expected, **1** and **4** bind to PikC-RhFRED with K_d values of 92.6 ± 0.5 μM

and 215.0 ± 4.2 μM, respectively, which are similar to 112.9 ± 1.9 μM (**1**) and 288.3 ± 7.1 (**4**) toward wt PikC (Fig. 4-S5). This indicates that attachment of the heterologous reductase domain has no significant impact on substrate binding to PikC. Subsequently, we compared the kinetic parameters of PikC-RhFRED with those of the PikC-Fdr-Fdx three-component system. As previously reported (18, 27), substrate inhibition was observed in all cases when substrate concentration was greater than 250 μM. Moreover, the solubility limitation (less than 500 μM) of macrolides in aqueous solution prevented us from deducing the K_i value. Therefore, we determined the apparent specificity constants (k_{cat}/K_m) by fitting the low-concentration data to the linear region of the Michaelis−Menten curve (Fig. 4-S6). By directly monitoring the substrate consumption by HPLC, the k_{cat}/K_m values of PikC-RhFRED were determined to be 0.96 and 1.20 $\mu M^{-1} \cdot min^{-1}$ for **1** and **4**, respectively. In contrast, the specificity constants of wt PikC partnered by Fdr and Fdx were 0.24 $\mu M^{-1} \cdot min^{-1}$ for **1** and 0.31 $\mu M^{-1} \cdot min^{-1}$ for **4**. It is evident that the fusion enhanced catalytic activity approximately 4-fold for both **1** and **4**. Notably, the kinetic parameters of wt PikC differ significantly from those previously determined indirectly, using a NADPH depletion assay (18, 19, 27), suggesting that the stoichiometric ratio between NADPH and substrate hydroxylation could not be 1:1. The presumed decoupling between electron transfer and hydroxylation might account for this difference.

Finally, when RhFRED was fused to another prototype biosynthetic P450 EryF (21, 22), a more active self-sufficient biocatalyst was obtained once again (Fig. 4-S7). Together with previous *in vivo* work (15, 16), our studies demonstrate that further development of RhFRED as the basis for an efficient cost-effective redox partner for bacterial biosynthetic P450s is warranted. Further efforts to understand this unique reductase, especially the electron-transfer process involving heterologous fusion systems, are now in progress.

4.3 Experimental Procedures

Molecular cloning — The gene encoding the reductase domain of $P450_{RhF}$ (RhFRED) and the linker sequence was amplified by PCR under standard conditions using forward primer: 5'-GG*GAATTC*GTGCTGCACCGCCATCAACCG-3' (the italic bases represent

*Eco*RI restriction site), reverse primer 1: 5'-TTAGAGCTCCAGAGGCGCAGGGCCAGGCG-3' (the *Sac*I cutting site is underlined) for amplifying *RhFRED* gene without the stop codon, and reverse primer 2: 5'-ACATCAAGCTTTCAGAGGCGCAGGGCCAG-3' (the *Hind*III cutting site is underlined) for cloning *RhFRED* gene retaining the stop codon. The cDNA without and with the stop codon were digested by *Nde*I/*Sac*I and *Nde*I/*Hind*III restriction enzyme pairs, respectively, and then ligated into the *Nde*I/*Sac*I digested pET21b(+) (Novagen) and *Nde*I/*Hind*III digested pET28b(+) (Novagen) correspondingly, generating the recombinant plasmid pET21b(+)-*RhFRED* and pET28b(+)-*RhFRED*. On the other hand, using previously prepared plasmid pET28a-*pikC* (18) as template, the *pikC* gene with stop codon removed was amplified by PCR under standard conditions using a pair of primers as follow: forward, 5'-GGAGTTC*CATATG*CGCCGTACCCAGCAG-3', reverse: 5'-GATAGAATTCACCGGTACGGCGGCCCGC-3' (The italic and underlined bases represent the introduced *Nde*I and *Eco*RI restriction sites for following cloning manipulation). The *Nde*I/*Eco*RI double digested *pikC* gene was then ligated into the previously *Nde*I/*Eco*RI-digested pET21b(+)-*RhFRED* and pET28b(+)-*RhFRED* to generate vectors pET21b(+)-*pikC*-RhFRED and pET28b(+)-*pikC*-*RhFRED* (Fig. 4-S1) for overexpressing C-His_6-tagged and N-His_6-tagged fusion protein PikC-RhFRED.

Overexpression and purification of fusion proteins — The *E. coli* BL21 (DE3) transformant carrying certain overexpression vector was grown at 37 ^oC overnight in LB media containing 50 μg/ml of kanamycin (or ampicillin), then 10 ml of culture was used to inoculate 1 liter of LB media containing certain selective antibiotic (50 μg/ml), thiamin (1 mM), 10% glycerol and a rare salt solution. Cells were grown at 37 ^oC for 3~4 h until A_{600} 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 ^oC overnight. The following protein purification was carried out according to previously developed procedure (18).

Spectral substrate binding assay — Spectral substrate binding assay was carried out on UV-visible spectrophotometer 300 Bio (Cary) at room temperature by titrating 30 mM

substrate DMSO solution (blank DMSO for reference group) into 1 ml of 1 μM P450 sample in 1 μl aliquots, leading to substrate concentration ranging from 30 to 360 μM. The series of Type I difference spectra were used to deduce ΔA ($A_{peak(389)}$-$A_{trough(424)}$). Then, the data were fit to Michaelis-Menten equation to obtain the dissociation constant K_d.

HPLC analysis of reactions catalyzed by PikC enzymes — The reactions for PikC enzymes were carried out using previously developed assays with minor modifications (18). The typical assay contains 1 μM PikC-RhFRED (or 1 μM PikC with 3.5 μM spinach ferredoxin and 0.1 U/ml spinach ferredoxin-$NADP^+$ reductase), 200 μM YC-17 (**1**) or narbomycin (**4**), and 0.5 mM NADPH in 100 μl of desalting buffer (50 mM NaH_2PO_4, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, 10% glycerol). The reaction was stopped and extracted after 1 h of incubation at 30 ^oC by addition of 2×200 μl of chloroform. The extracts were dried, dissolved in 120 μl of methanol and analyzed by Xbridge C18 5 μm 250 mm reverse-phase column at 230 nm, using 10-70% solvent B (A: deionized water + 10 mM ammonium acetate, B: acetonitrile) at 1 ml/min over 30 min. The peak identity in each HPLC trace was determined by mass spectrometry and comparison with authentic compounds regarding HPLC retention time and UV spectrum.

Steady-state kinetics — The standard reaction contains 100 nM of PikC-RhFRED (or 150 nM of wt-PikC with 3.5 μM spinach ferredoxin and 0.1 U/ml spinach ferredoxin-$NADP^+$ reductase), 20~250 μM substrate in 200 μl of desalting buffer (50 mM NaH_2PO_4, pH 7.3, 1 mM EDTA, 0.2 mM DTE, 10% glycerol). After pre-incubation at 30 ^oC for 5 min, the reaction was initiated by adding 1 μl of 50 mM NADPH and 50 μl aliquots were taken at 0s, 20s, and 40s (or 0s, 30s, and 60s when substrate concentration greater than 100 μM) to thoroughly mix with 2×100 μl of chloroform. The nitrogen-dried samples were redissolved in 150 μl of methanol for subsequent HPLC analysis. The HPLC conditions were: Xbridge C18 5 μm 250 mm reverse-phase column, 20-80% solvent B (A: deionized water + 0.1% trifluoroacetic acid, B: acetonitrile + 0.1% trifluoroacetic acid) at 0.8 ml/min over 20 min, UV wavelength 230 nm. The initial velocity of substrate consumption was deduced from decreased area under the curve (AUC) of specific

substrate peaks. Finally, the data from duplicated experiments were fit to Michaelis-Menten equation.

4.4 Supplementary Information

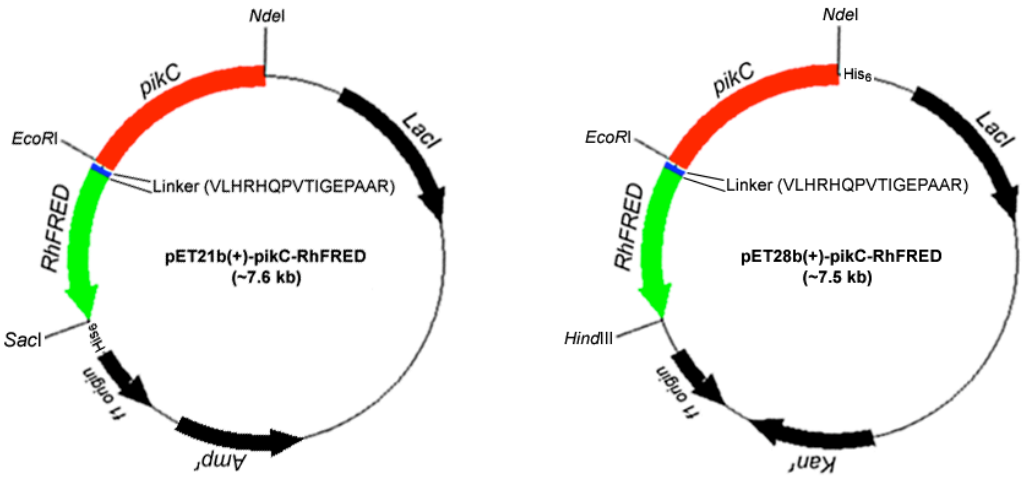


Figure 4-S1. Maps for N-terminal (left) and C-terminal His_6-tagged (right) PikC-RhFRED overexpression vectors.

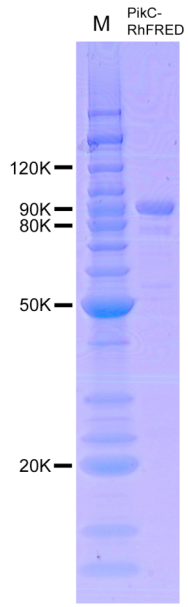


Figure 4-S2. SDS-PAGE analysis of functional N-terminal His_6-tagged PikC-RhFRED. The calculated molecular mass is about 83 kDa.

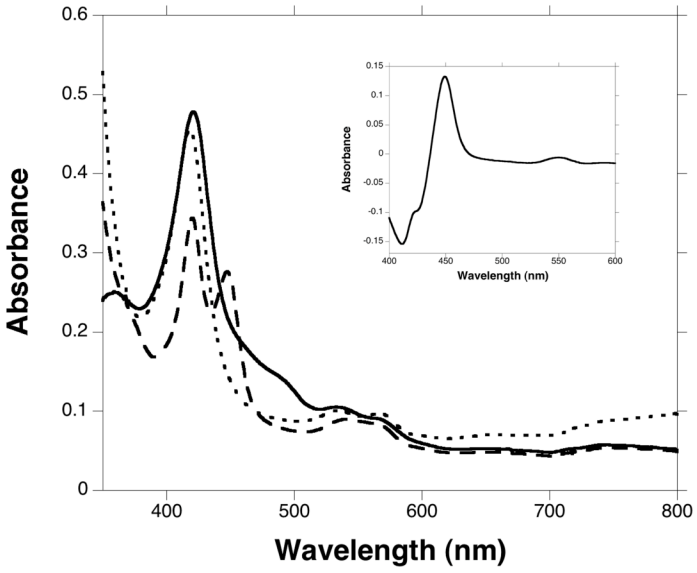


Figure 4-S3. UV-visible absorption spectra for purified N-terminal His_6-tagged PikC-RhFRED (~1.6 μM): oxidized (solid line), sodium dithionite reduced (dotted line), and CO-reduced (dashed line) are shown. The CO-bound reduced difference spectrum is shown in inset. This assay is also employed to determine the concentration of functional P450 enzyme using the extinction coefficient of 91,000 $M^{-1}\cdot cm^{-1}$ (Omura, T., and Sato, R. (1964) *J. Biol. Chem.* **239**, 2379.).

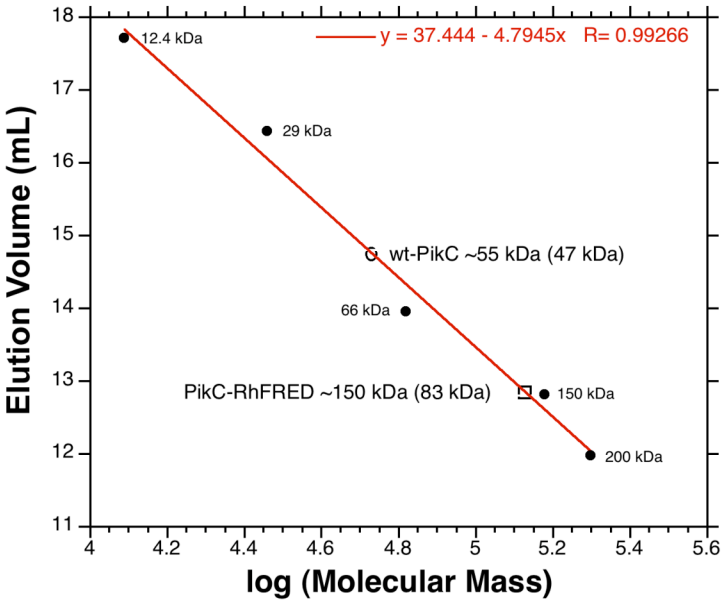


Figure 4-S4. The calculation of the apparent molecular mass of wt-PikC and PikC-RhFRED on the basis of the gel-filtration standard curve. The deduced molecular mass of wt-PikC and PikC-RhFRED is approximately 55 kDa and 150 kDa, respectively, suggesting the monomeric (wt) and dimeric (fusion) form in 50 mM sodium phosphate pH 7.3 buffer containing 300 mM NaCl, 10% glycerol. The theoretical molecular mass of each monomer is shown in parentheses.

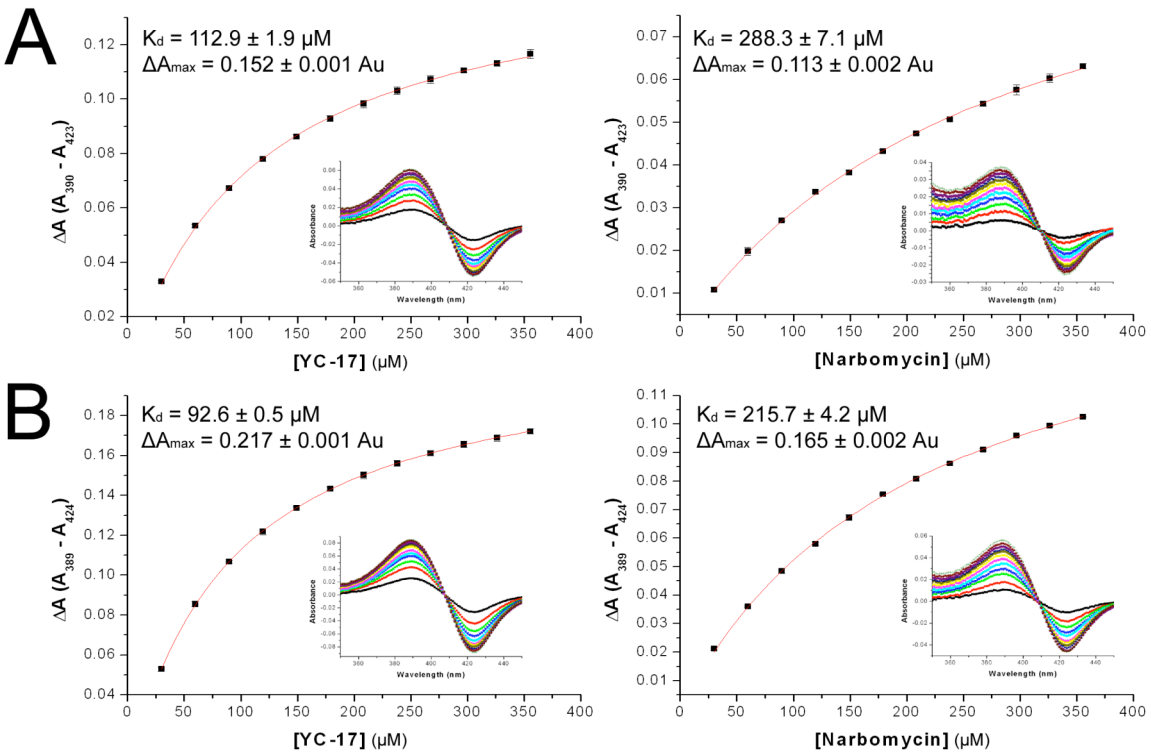


Figure 4-S5. The substrate binding affinity (K_d) of YC-17 (1) and narbomycin (4) toward wt-PikC (A) and PikC-RhFRED (B). The insets show the Type I binding spectra. All experiments were performed in duplicate.

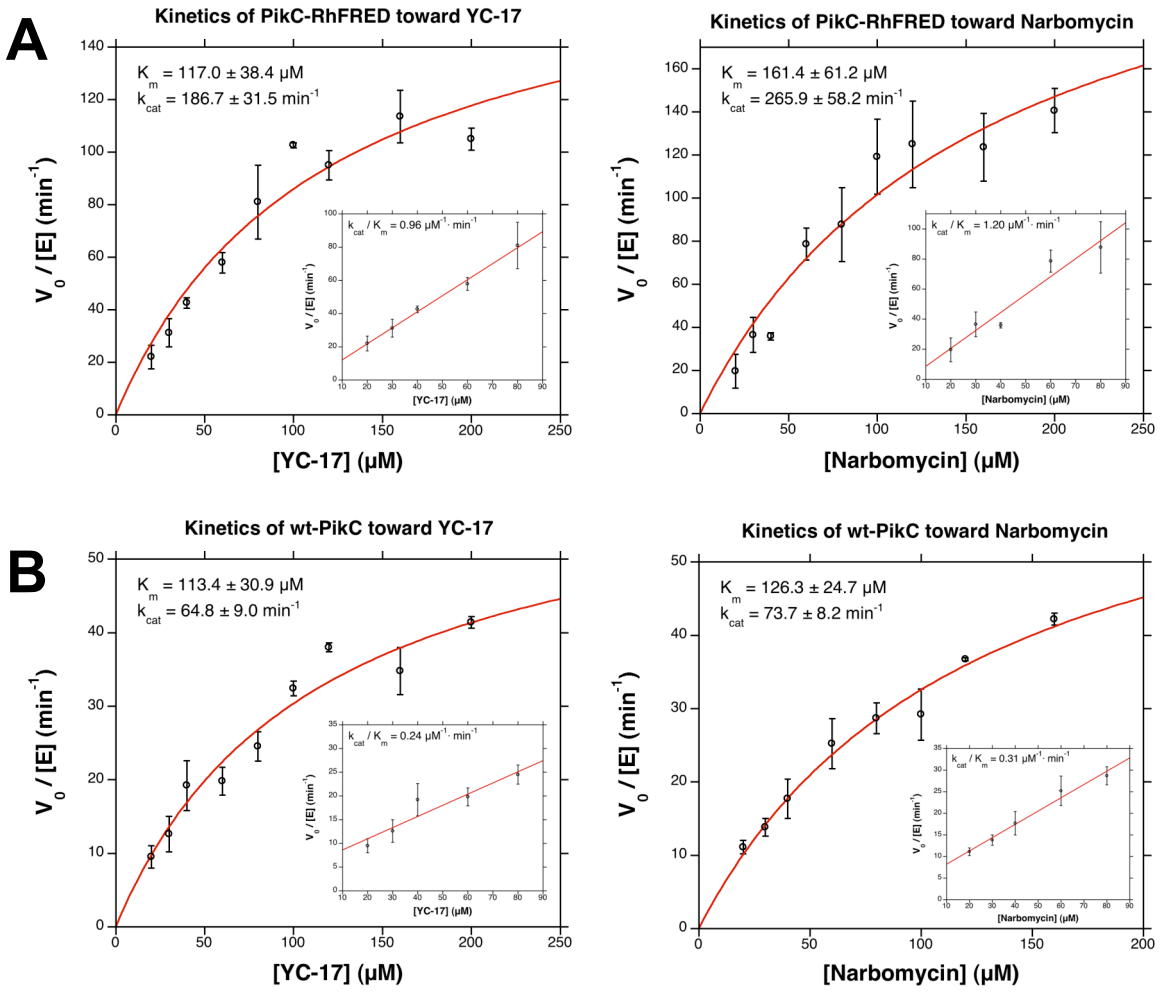


Figure 4-S6. Michaelis-Menten curves of PikC-RhFRED (A) and wt-PikC (B) using YC-17 (1) or narbomycin (4) as substrate. The apparent K_m and k_{cat} values are deduced from the unsaturated Michaelis-Menten curve. The k_{cat}/K_m values deduced from the linear fit of low-concentration data are shown in insets.

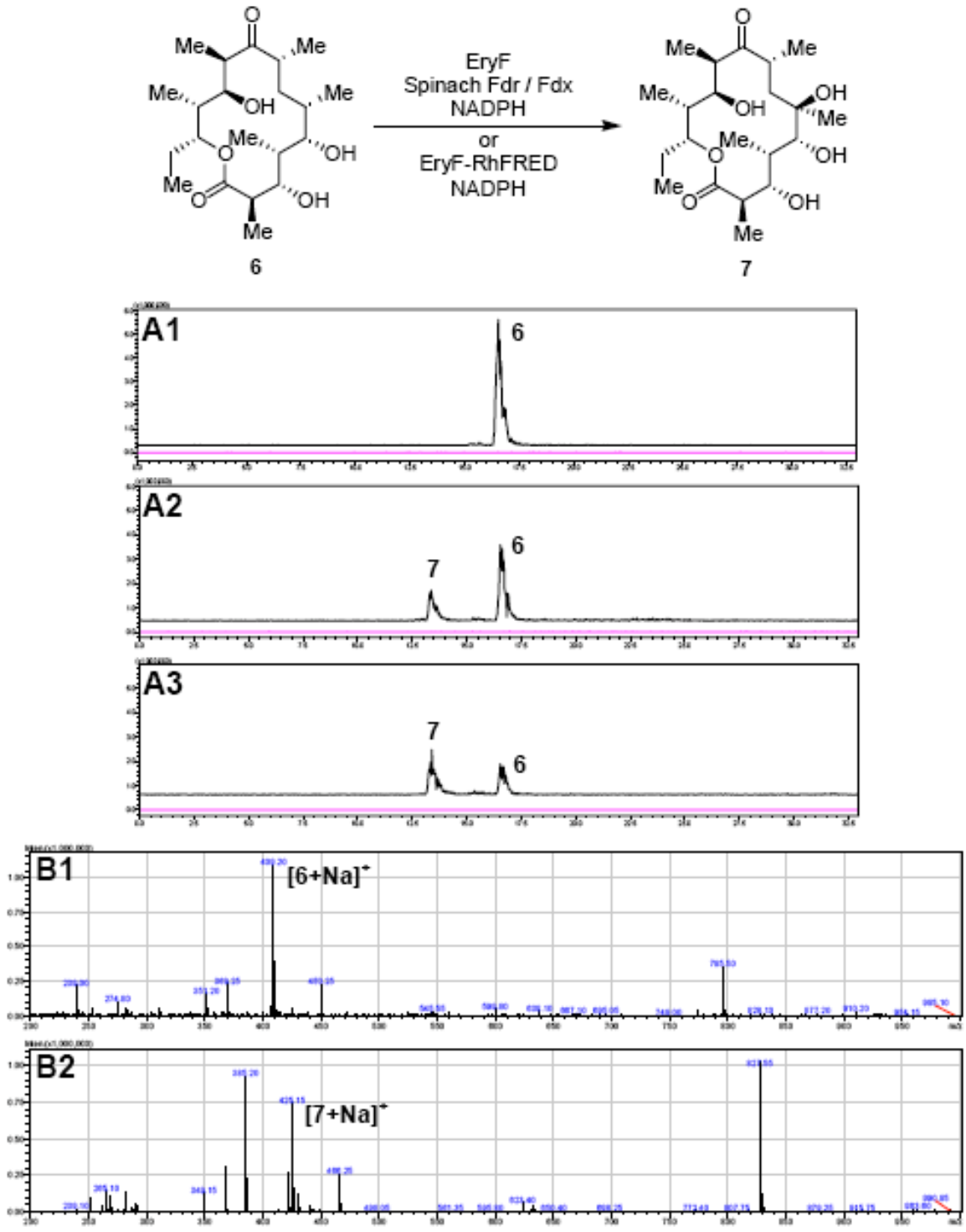


Figure 4-S7. LC-MS analysis of *in vitro* activity of EryF-RhFRED. A1. Negative control of 6-deoxyerythronolide B (**6**) in absence of P450. A2. **6** with wt EryF in presence of Fdr, Fdx, and NADPH. A3. **6** with EryF-RhFRED in presence of only NADPH. B1. Mass spectrum of **6**. B2. Mass spectrum of erythronolide B (**7**). The sample preparation was similar to that for typical PikC assay. The LC conditions were: Xbridge C18 3.5 µm 150 mm reverse-phase column, 20-100% solvent B (A: deionized water + 0.1% formic acid, B: acetonitrile + 0.1% formic acid) at 0.2 ml/min over 20 min.

4.5 References

1. Ortiz de Montellano, P. R. (1995) Cytochrome P450: Structure, Mechanism and Biochemistry, 2nd ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 149-162
2. Guengerich, F. P. (2001) *Chem. Res. Toxicol.* **14**, 611-650
3. Hannemann, F. B. A., Ewen, K. M., and Bernhardt, R. (2007) *Biochim. Biophys. Acta* **1770**, 330-344
4. Lewis, D. F. V., and Hlavica, P. (2000) *Biochim. Biophys. Acta*. **1460**, 353-374
5. Munro, A. W., Girvan, H. M., and McLean, K. J. (2007) *Nat. Prod. Rep.* **24**, 585-609
6. Ruettinger, R. T., and Fulco, A. J. (1981) *J. Biol. Chem.* **256**, 5728-5734
7. Otey, C. R., Bandara, G., Lalonde, J., Takahashi, K., and Arnold, F. H. (2005) *Biotechnol. Bioeng. 93*, 494-499
8. Munro, A. W., Leys, D. G., McLean, K. J., Marshall, K. R., Ost, T. W. B., Daff, S., Miles, C. S., Chapman, S. K., Lysek, D. A., Moser, C. C., Page, C. C., and Dutton, P. L. (2002) *Trends Biochem. Sci.* **27**, 250-257
9. Fairhead, M., Giannini, S., Gillam, E. M. J., and Gilardi, G. (2005) *J. Biol. Inorg. Chem.* **10**, 842-853
10. Dodhia, V. R., Fantuzzi, A., and Gilardi, G. (2006) *J. Biol. Inorg. Chem.* **11**, 903-916
11. De Mot, R., and Parret, A. H. A. (2002) *Trends Microbiol.* **10**, 502-508
12. Roberts, G. A., Grogan, G., Greter, A., Flitsch, S. L., and Turner, N. J. (2002) *J. Bacteriol.* **184**, 3898-3908
13. Roberts, G. A., Çelik, A., Hunter, D. J. B., Ost, T. W. B., White, J. H., Chapman, S. K., Turner, N. J., and Flitsch, S. L. (2003) *J. Biol. Chem.* 48914-48920
14. Hunter, D. J. B., Roberts, G. A., Ost, T. W. B., White, J. H., Müller, S., Turner, N. J., Flitsch, S. L., and Chapman, S. K. (2005) *FEBS Lett.* **579**, 2215-2220
15. Kubota, M., Nodate, M., Yasumoto-Hirose, M., Uchiyama, T., Kagami, O., Shizuri, Y., and Misawa, N. (2005) *Biosci. Biotechnol. Biochem.* **69**, 2421-2430
16. Nodate, M., Kubota, M., and Misawa, N. (2006) *Appl. Microbiol. Biotechnol.* **71**, 455-462
17. Xue, Y., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12111-12116
18. Xue, Y., Wilson, D., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Chem. Biol.* **5**, 661-667
19. Lee, S. K., Park, J. W., Kim, J. W., Jung, W. S., Park, S. R., Choi, C. Y., Kim, E. S., Ahn, J. S., Sherman, D. H., and Yoon, Y. J. (2006) *J. Nat. Prod.* **69**, 847-849
20. Sherman, D. H., Li, S., Yermalitskaya, L. V., Kim, Y., Smith, J. A., Waterman, M. R., and Podust, L. M. (2006) *J. Biol. Chem.* **281**, 26289-26297
21. Lambalot, R. H., and Cane, D. E. (1995) *Biochemistry* **34**, 1858-1866
22. Andersen, J. F., and Hutchinson, R. C. (1992) *J. Bacteriol.* **174**, 725-735
23. Ogura, H., Nishida, C. R., Hoch, U. R., Perera, R., Dawson, J. H., and Ortiz de Montellano, P. R. (2004) *Biochemistry* **43**, 14712-14721
24. Neeli, R., Girvan, H. M., Lawrence, A., Warren, M. J., Leys, D., Scrutton, N. S., and Munro, A. W. (2005) *FEBS Lett.* **579**, 5582-5588
25. Munro, A. W., Girvan, H. M., and McLean, K. J. (2007) *Biochim. Biophys. Acta.* **1770**, 345-359
26. Yabusaki, Y. (1995) *Biochimie* **77**, 594-603
27. Graziani, E. I., Cane, D. E., Betlach, M. C., Kealey, J. T., and McDaniel, R. (1998) *Bioorg. Med. Chem. Lett.* 3117-3120

Notes:

This work has been published as "Engineering and analysis of a self-sufficient biosynthetic cytochrome P450 PikC fused to the RhFRED reductase domain." **Li, S.**, Podust, L. M., and Sherman, D. H. ***J. Am. Chem. Soc.*** 2007, 129 (43), 12940-12941.

Author contributions:
Shengying Li, Larissa M. Podust, and David H. Sherman designed the experiments; Shengying Li performed the experiments;

Chapter 5
Selective Oxidation of Carbolide C-H Bonds by an Engineered Macrolide P450 Monooxygenase ($PikC_{D50N}$-RhFRED)

5.1 Introduction

The superfamily of cytochrome P450 (CYP450) enzymes (monooxygenases) is involved in diverse oxidative processes including xenobiotic catabolism, steroid synthesis, and biosynthetic tailoring of diverse natural products (1-3). Among various reactions catalyzed by P450 enzymes, the regio- and stereoselective oxidation of an unactivated sp^3 C-H bond represents a central challenge in organic chemistry (4-8). Considerable effort has been devoted to identifying biological catalysts or simpler mimics that function by mechanisms typically involving a metal oxo reactive site (9). Alternatively, transition metal complexes have been identified for C-H bond oxidations that proceed through mechanisms completely distinct from biological systems. A key challenge in developing useful C-H oxidation procedures is the control of site-selectivity among similar C-H bonds. Successful approaches have typically involved either relying on the inherent reactivity differences of various C-H bonds based on steric and electronic considerations (10-15), or the incorporation of directing groups that orient the catalyst active site towards a specific C-H bond (16-18). The selective oxidation of a C-H bond that is neither inherently more reactive than alternate sites nor positioned adjacent to a directing group poses the most difficult application in site-selective C-H bond functionalization.

Recent reports have shown that supramolecular organometallic assemblies can provide some success in this challenge for synthetic chemistry (19-24). Alternatively, biological catalysts may provide unique potential to selectively oxidize bonds that are chemically similar, yet remote from directing influences. Thus, we were drawn to

investigate the potential role of biosynthetic CYP450 monooxygenases despite their fundamental dependence on substrate-enzyme complementarity, which might limit their application in synthetic chemistry (25). A number of previous efforts have sought to overcome this limitation by employing protein engineering strategies, including scanning chimeragenesis (26, 27) and directed-evolution (28-31) to generate non-natural cytochrome P450s (e.g., $P450_{BM3}$) with novel substrate specificities and abilities to selectively oxidize target substrates.

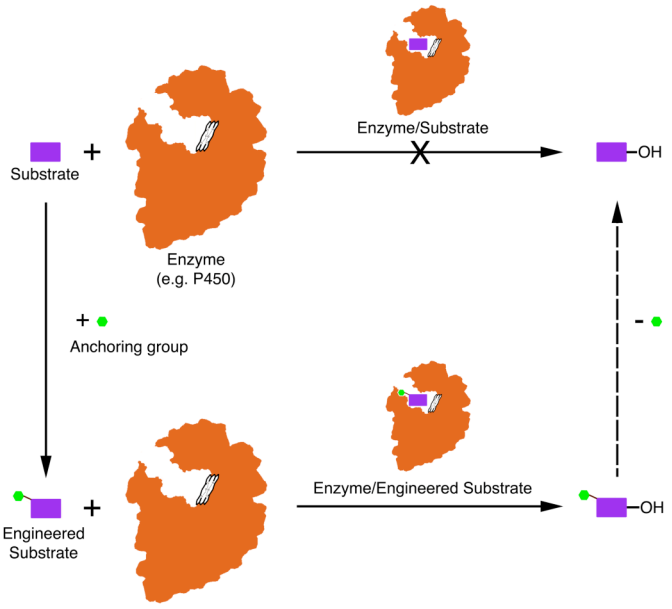


Figure 5-1. Schematic strategy of substrate engineering.

Based on a series of investigations involving CYP450 monooxygenases with remarkable substrate flexibility in macrolide biosynthetic systems (32, 33), we were motivated to explore their potential value in selective C-H bond activation reactions (25). Their unique requirement for a sugar-appended substrate lead us to assess an underexplored strategy involving "substrate engineering" (Fig. 5-1) that offers another way to broaden the substrate landscape of biological catalysis (34, 35). Our anchoring/tethering approach was inspired by an earlier observation that certain functional groups facilitate enzyme-substrate interactions, where a direct functional group linkage (termed anchoring group in this study) to previously inaccessible compounds enables selective oxidation of the modified substrates *in vivo* (36, 37). However, due to the complexity of the cell-based system, the precise mechanism behind the substrate

engineering approach has remained unclear. An initial application (36) hypothesized that the anchoring group might be involved in substrate recognition, productive binding, or control of binding orientation in the active site of certain oxidative enzymes (e.g. monooxygenases).

Herein, we report the first *in vitro* implementation of substrate engineering for selective C-H bond oxidations by using an optimized form of the macrolide CYP450 monooxygenase PikC (32, 38). For this study, a series of carbocyclic rings linked to the desosamine glycoside (referred to as carbolides) were effectively hydroxylated in a regioselective manner. Furthermore, analysis of a series of high-resolution enzyme-substrate co-crystal structures provided significant new insights into the function of the aminosugar-derived anchoring group for control of reaction site selectivity. Finally, unexpected biological activity of a select number of these carbolide systems revealed their potential as a new class of antibiotics.

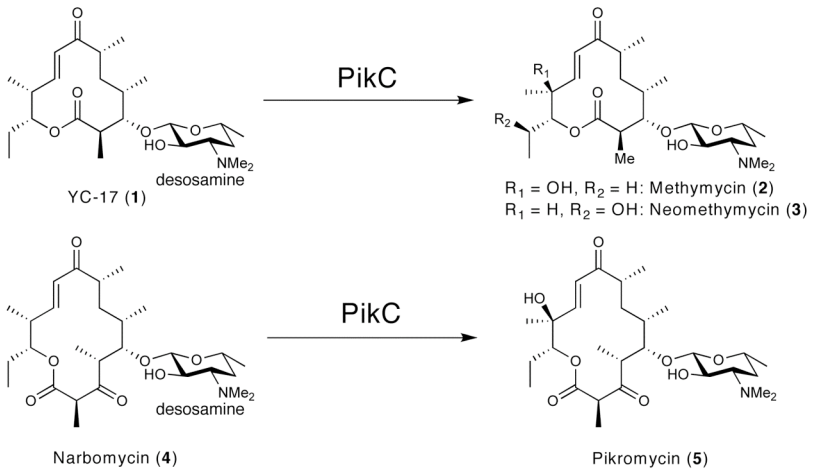


Figure 5-2. Major physiological reactions catalyzed by PikC.

5.2 Results and Discussion

5.2.1 The Engineered PikC$_{D50N}$-RhFRED Is Capable of Hydroxylating Carbolides Effectively and Regioselectively.

PikC is the cytochrome P450 involved in pikromycin biosynthesis from *Streptomyces venezuelae* (32, 39). The physiological function of this monooxygenase is to hydroxylate both the 12-membered ring macrolide YC-17 **1** and the 14-membered ring

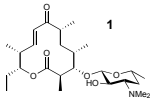
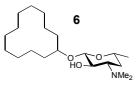
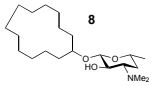
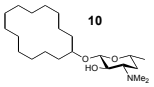
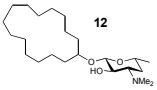
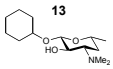
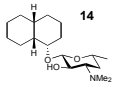
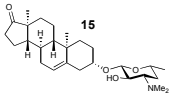
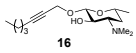
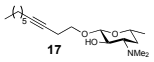
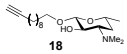
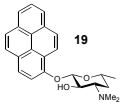
macrolide narbomycin **4**, giving rise to methymycin/neomethymycin **2**/**3** and pikromycin **5**, respectively, as major products (Fig. 5-2). Recent analysis of x-ray co-crystal structures of PikC (40, 41) involving endogenous substrates revealed that the macrolactone ring contacts the active site residues entirely via nonspecific hydrophobic interactions, likely accounting for the tolerance of PikC toward the variant macrolactone ring size and functionalization. In contrast, the desosamine sugar employs two distinct binding pockets and anchors the substrate through a number of hydrogen bonds and ionic interactions, in particular, a unique salt bridge between the protonated dimethylamino group of desosamine and a glutamate residue, either Glu94 or Glu85 in the B/C loop region. Based on these previously recognized molecular interactions that specify substrate binding affinity and orientation in the binding pocket, we reasoned that desosamine could be an effective anchoring group to direct positioning of various unnatural molecules in the active site of PikC for selective C-H bond hydroxylations.

To test this hypothesis, we synthesized the unnatural cyclic carbolide substrate desosaminyl cyclododecane **6** to mimic the structure of the natural substrate **1** (YC-17) using a recently developed synthetic strategy (33), which was subsequently employed as a general approach to derivatize diverse alcohols with desosamine (Fig. 5-S1). An enzyme-substrate analysis showed that **6** binds to wild type PikC ($PikC_{wt}$) with a dissociation constant (K_d) of 1,379 μM, about 12 times higher than the K_d value (116 μM) of **1**. The decreased binding affinity of **6** could result from 1) an entropic penalty upon binding due to high conformational freedom of the saturated ring system, 2) lack of hydrophobic interactions between the functional groups on the macrolactone ring of **1** and PikC active site residues, and/or 3) loss of some specific interactions with desosamine as observed in the co-crystal structure with PikC (see below). When using the more active $PikC_{D50N}$ mutant (40, 41), the binding affinities of both carbolide **6** and macrolide **1** were shown to be approximately four times higher with K_d values of 390 and 32 μM, respectively. Moreover, we recently engineered a self-sufficient fusion enzyme PikC-RhFRED (38) that displayed approximately 4-fold enhanced catalytic activity (k_{cat}) compared to $PikC_{wt}$. Combining these two beneficial properties, the resulting engineered form of the P450 enzyme $PikC_{D50N}$-RhFRED (k_{cat}/K_m = 7.44 $\mu M^{-1} \bullet min^{-1}$ for **1**, Fig. 5-S2) is approximately 13 times more active than $PikC_{wt}$ (38). Interestingly, **1** and **6** bound to

this mutant enzyme with slightly improved K_d values of 19 and 309 μM (Table 5-1), respectively. Due to enhanced substrate conversion and ease of use in the absence of expensive exogenous redox partners, we elected to employ $PikC_{D50N}$-RhFRED to hydroxylate carbolide **6** and all other substrates for this work.

LC-MS analysis (Fig. 5-3A) of the extract obtained following reaction with $PikC_{D50N}$-RhFRED showed that 47% of carbolide **6** was converted into seven different monohydroxylated products **7a-g** (no multi-hydroxylated products were observed) with expected *m/z* = 358.19 for $[\mathbf{6} + OH + H]^+$ using 5 μM of $PikC_{D50N}$-RhFRED in 3 h (the

Table 5-1. The activity of $PikC_{D50N}$-RhFRED toward various substrates.

Substrates[a]	K_d (μM)	Yield[b]	Number of products[c]
 **1**	19	>99%	2
 **6**	309	47%	7
 **8**	218	65%	6
 **10**	289	63%	6
 **12**	243	35%	9
 **13**	NB[d]	0	0
 **14**	NB	0	0
 **15**	NB	< 1%	0
 **16**	NB	0	0
 **17**	2,900	8%	5
 **18**	2,300	14%	7
 **19**	>5,000	4%	1

[a]Racemic cis-decahydro-1-naphthol was used in preparation of **14**, leading to the structure shown and its diastereomer. The mixture of diastereomers was subjected to the PikC oxidation reaction.

[b]The yield shown is based on product formation calculated with $AUC_{total\ products}$ / ($AUC_{total\ products}$ + $AUC_{unreacted\ substrate}$) by assuming ionization efficiency of substrate and various hydroxylated products are same since the ionization site of this series of desosaminyl derivatives should be the dimethylamino group. The calculated yield is fairly close (difference < ±3%) to the substrate conversion ratio calculated by 1 – $AUC_{unreacted\ substrate}$ / $AUC_{total\ substrate}$ ($AUC_{total\ substrate}$ is derived from the control reaction using boiled enzyme under identical conditions), indicating the assumption is correct. (AUC: area under curve of the ion count chromatograms as shown in Fig. 5-3).

[c]The number of products refers to the number of distinct peaks seen in the LCMS. Comparison to synthetic standards illustrated a few cases where regioisomers or diastereomers were not distinguished by LCMS, so the number provided is a lower limit. See supporting information for details.

conversion can be driven further by increasing enzyme concentration and/or reaction time). All product ions displayed the same MS/MS spectra (Fig. 5-S3) at *m/z* = 158.02, corresponding to [desosamine – OH]$^+$. The unmodified desosamine moiety indicates that all hydroxylations occur on the cyclododecane ring. In contrast, cyclododecanol lacking an appended desosamine was unable to serve as a substrate for PikC P450 under identical conditions. Therefore, it is evident that desosamine is indispensable for this biochemical transformation. Notably, PikC$_{wt}$, PikC$_{D50N}$, and PikC$_{wt}$-RhFRED generated similar product profiles compared to PikC$_{D50N}$-RhFRED, albeit with lower efficiency. These results indicate that neither the point mutation nor the C-terminal RhFRED-fusion with PikC has a significant impact on the binding mode of **6**.

To assess the regio- and stereoselectivity of PikC$_{D50N}$-RhFRED toward the selected substrate, we sought to obtain detailed structural information on the reaction products. Challenging preparative-scale separations as well as the high similarity of the methylene groups on the cyclododecane ring complicated structure determination of the individual compounds by NMR analysis. Due to the large number of potential isomers (23 total) that could result from oxidation of the 12-membered ring of **6**, it was impractical to synthesize all possible products. Therefore, guided by the 2.0 Å co-crystal structure of mutant PikC$_{D50N}$ enzyme with **6** (Fig. 5-4), we synthesized a series of authentic standards with a hydroxyl group installed at the C7 and C6/C8 position (the two diastereotopic carbons were numbered differently for clarity) that were likely to be the major reaction products. During the chemical synthesis, but prior to desosamine installation, racemic mixtures of two diastereomers of a mono-protected diol leading to C6/C8-oxidized authentic materials were prepared. Glycosylation with the glycosyl fluoride of acetate-protected desosamine followed by deprotection provided the four diastereomers that correspond to C6/C8 oxidized products (see Supplementary Information). Similarly, the two diastereomers of C7-oxidized authentic material were obtained.

Through comparison of LC-MS retention times and confirmation by co-injections (Fig. 5-S4), six of the seven PikC-derived hydroxylated products of **6** were assigned as the C7 and C6/C8 oxidized materials. Since authentic samples of oxidized products were prepared as diastereomeric pairs, hydroxyl positioning was determined, but the precise

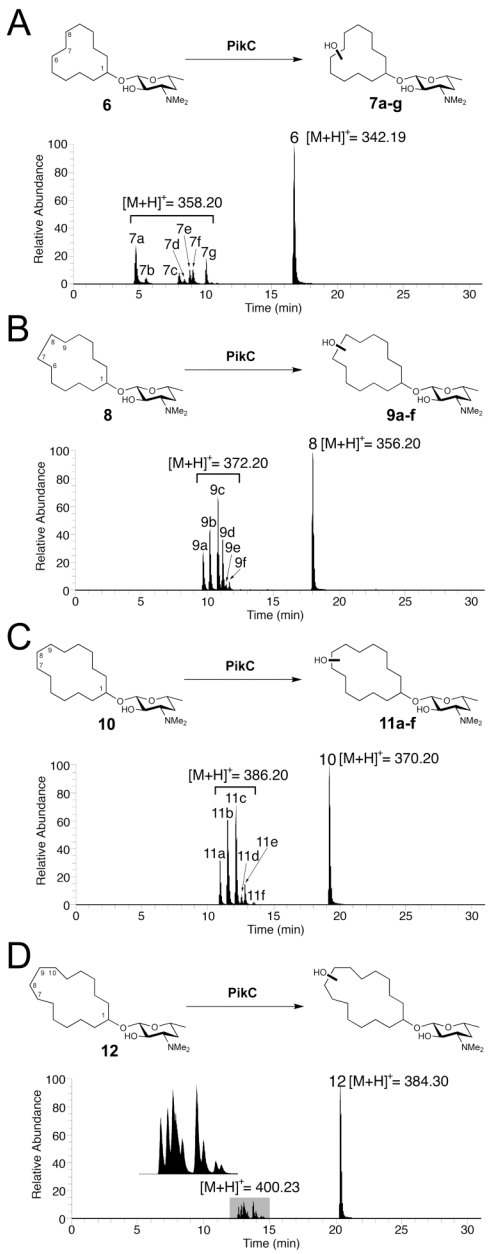


Figure 5-3. LC-MS analysis of PikC$_{D50N}$-RhFRED catalyzed reactions using different cyclized carbolides as substrates. (Ion count chromatograms are shown) (A) Desosaminyl cyclododecane **6** reaction, **7b** and **7f** correspond to two diastereomers generated by C7 hydroxylation, **7a**, **7c**, **7e**, and **7g** correspond to four diastereomers generated by C6/C8 hydroxylation; (B) Desosaminyl cyclotridecane **8** reaction, **9a**, **9d** and **9f** correspond to diastereomers arising from C6/C9 hydroxylation, **9b** and **9c** correspond to diastereomers originated from C7/C8 or C6/C9 hydroxylation. The number of products that peak **9c** contains is undetermined due to product overlap; (C) Desosaminyl cyclotetradecane **10** reaction; (D) Desosaminyl cyclopentadecane **12** reaction. The details of product assignment for **6** and **8** based on correlation with synthesized authentic standards regarding retention time and co-injection confirmation are shown in Supplementary Information.

stereostructure was not. The ratio of C7 oxidized products (**7b** and **7f**) to C6/C8 oxidized

products (**7a**, **7c**, **7e**, and **7g**) was approximately 1:4 (see Supplementary Information for product assignment). Thus, it is evident that PikC catalyzed hydroxylation occurs primarily at sites most remote from the desosamine-anchoring group, as predicted by the crystal structure (Fig. 5-4). The C7 and C6/C8 oxidized compounds account for 95% of the mass of mono-hydroxylated material, and the only unidentified minor product **7d** (5%) might be one of the C5 hydroxylated products. Considering the abundance of secondary C-H bonds on the 12-membered ring with almost equal reactivities, this regioselectivity is considerable, but not as strict as that observed toward the native macrolide substrates **1** (YC-17) and **4** (narbomycin).

We next sought to determine if C-H hydroxylation still occurs at the sites remote from desosamine in the cases of larger hydrocarbon rings, thus the 13-membered ring carbolide **8** was synthesized and treated with $PikC_{D50N}$-RhFRED in a similar way to **6**. Consistently, the major peaks **9a-d** (accounting for 94% of all products, Fig. 5-3B) were determined to be C7/C8 or C6/C9 mono-hydroxylated products (Fig. 5-S6). However, due to high structural similarity between these molecules, some synthetic hydroxylated authentic standards containing different diastereomeric pairs were not distinguishable from one another due to identical retention times (e.g. Fig. 5-S6, traces D and E**)**, which prevented determination of the exact number of diastereomers generated in the reaction.

Furthermore, two larger cyclic carbolides, including the 14-membered ring **10,** and 15-membered ring **12** were investigated. The reactivity of desosaminyl cyclotetradecane **10** was similar to that of 13-membered ring carbolide **8**, with 65% substrate converted to hydroxylation products **11a-f** (Fig. 5-3C). In contrast, the desosaminyl cyclopentadecane **12**, although having comparable binding affinity with **8** and **10** (Table 5-1), showed lower conversion (35%) and decreased selectivity reflected by the increased number of products (Fig. 5-3D), suggesting that this large cyclic substrate might not be located in a suitable position within the PikC active site.

5.2.2 Analysis of $PikC_{D50N}$-RhFRED Reactivity toward Other Types of Desosaminyl Derivatives.

Due to the surprising activity and regioselectivity that $PikC_{D50N}$-RhFRED showed toward 12-, 13-, and 14-membered ring carbolide derivatives, we decided to investigate

the ability of desosamine to function as an anchor for other classes of compounds. Thus, the activity of $PikC_{D50N}$-RhFRED toward additional cyclic and linear synthetic substrates (Table 5-1) was assessed. The results revealed that PikC only marginally hydroxylated cyclic derivatives **13**-**15**, suggesting that these small ring systems might not be accommodated within the active site, or cannot reach the iron-oxo center for catalysis.

Surprisingly, the linear derivatives **17** and **18** were oxidized to form multiple mono-hydroxylated products with 8% and 14% yields, respectively (Table 5-1, Fig. 5-S9 and 5-S10). The predominant hydroxylation site for **18** was identified as the C9 propargylic position (Fig. 5-S11). When aromatic derivative **19** was used as substrate (Fig. 5-S12), we observed a single oxidized product determined to be the 8-hydroxy desosaminyl pyrene (Table 5-S1), albeit with a conversion of 4%. This observation suggests that rigidity of the substrate might be a key factor to gain high selectivity when using PikC as a biocatalyst for oxidation. The low level of conversion might result from its poor binding and/or higher C-H bond dissociation energy in this aromatic ring system. Taken together, these data reveal that PikC-catalyzed oxidation occurs primarily at the site most remote from the desosamine anchoring group when linear or aromatic substrates are used, suggesting a similar mechanism for control of regioselectivity compared to the carbolides.

5.2.3 Structural Basis for Regioselectivity of PikC toward 12- and 13-membered Ring Carbolides.

To understand the structural basis for the regioselectivity of PikC and predict the hydroxylation sites of the 12- and 13-membered ring carbolides to direct the synthesis of authentic standards, we solved the crystal structures (Table 5-S2, Fig. 5-4 and 5-5) of $PikC_{D50N}$ in complex with unnatural substrates **6** and **8**.

In the co-crystal structure of $PikC_{D50N}$ and **6**, the electron density for **6** is well defined in one monomer of the asymmetric unit (Fig. 5-4A). In contrast, the other asymmetric unit showed an unambiguously positioned cyclododecane ring, while the dispersed electron density for desosamine indicated at least two alternative conformations (Fig. 5-4B). A satisfactory fit was achieved when **6** was docked in two flipped orientations. Carbolide **6** binds in the active site in the L-shaped conformation bringing

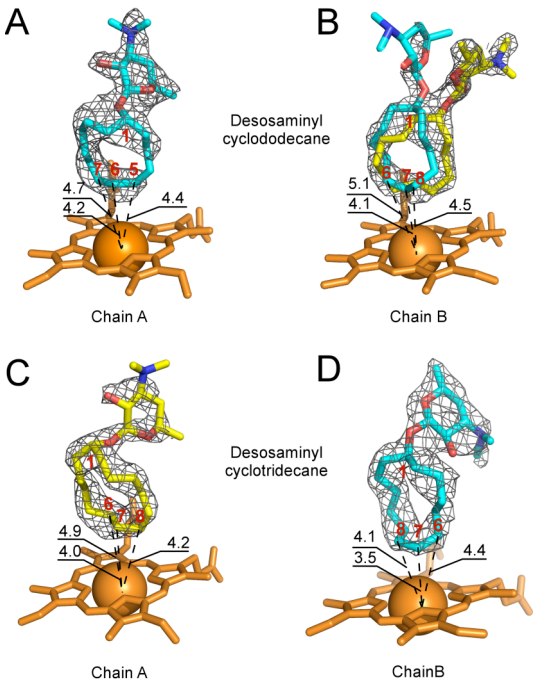


Figure 5-4. Multiple binding modes of desosaminyl cycloalkanes. Orientations of **6** in the active site of $PikC_{D50N}$ (*A*) in chain A and (*B*) in chain B, and orientations of **8** in the active site of $PikC_{D50N}$ (*C*) in chain A and (*D*) in chain B, as defined by the fragments of the electron density map (*gray mesh*) contoured at 0.8 σ are shown. In (*B*), **6** is docked in the flipped-over orientations allowing hydroxylation on the both sides of the ring. In (C) and (D), **8** is in flipped over orientations. Heme is shown in *orange*. Oxygen atoms are in *red*, nitrogen in *blue*, iron in *orange*. Atoms of the cycloalkane ring are labeled in *red*. Distances are in Angstroms. Images are generated using PYMOL.

four cyclododecane carbons most remote from the desosamine anchoring group within 5 Å of the Fe reaction center (Fig. 5-4A and 5-4B), revealing that C5, C6, C7, and C8 are the likely hydroxylation sites. The observed pattern of regioselectivity could arise from sporadic contacts of the desosamine moiety of **6** with a number of amino acid residues (Fig. 5-5A). Indeed, the specific salt-bridge involving E94 is found in only one conformer (*pink* in Fig. 5-5A) of **6**, and a number of the chain A active site residues (*green*), including F178, I239, V242, and M394, adopt alternative conformations indicative of dynamic interactions with the substrate. In chain B, although the salt-bridge to E94 is lost, E246 (*ice blue* in Fig. 5-5A) is located within electrostatic distance from the dimethyamino group of the proximal conformer (*cyan*). However, the E246 side chain is missing from the electron density map of chain A. In addition, the sub-optimal regioselectivity could be due to the inherent flexibility of the large cycloalkane ring. Since all six diastereomers arising from the C7 and C6/C8 hydroxylated regioisomers

were observed, there is clearly a lack of stereoselectivity from this enzymatic transformation. The co-crystal structure suggests that compromised stereoselectivity is likely due to flipping (or rotating) of the carbolide substrate in the PikC active site resulting in oxidation on both faces of the ring (Fig. 5-4A and 5-4B).

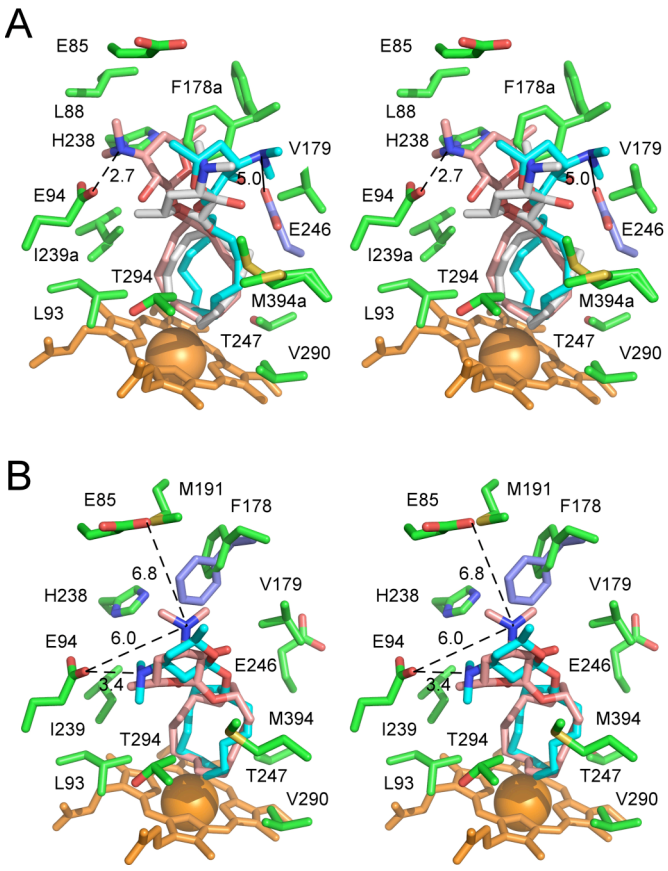


Figure 5-5. Desosaminyl cycloalkane binding sites. (A) Stereo view of the $PikC_{D50N}$ binding site with the three superimposed **6** conformers highlighted in *gray* (chain A), *pink* and *cyan* (chain B) surrounded by the chain A amino acid side chains within 5 Å plus E85 (*green*) is shown. E246 of chain B is highlighted in *ice blue*. To increase clarity, V242 was omitted from the drawing. (B) Stereo view of the $PikC_{D50N}$ binding site with two superimposed **8** conformers highlighted in *pink* (chain A) and *cyan* (chain B) surrounded by the chain A amino acid side chains within 5 Å plus E85 is shown. F178 of chain B is highlighted in *ice blue*. Heme is shown in *orange*. Oxygen atoms are in *red*, nitrogen in *blue*, iron in *orange*. Atoms of the cycloalkane ring are labeled in *red*. Distances between tertiary amine and carboxylic groups are in Angstroms. Lower-case *a* in the residue label indicates that alternative conformations are shown.

Further inspection of the co-crystal structures revealed that whereas three orientations are distinguishable for **6**, only two orientations (one in each protein monomer in the asymmetric unit) are observed for **8** in the 2.2 Å x-ray co-crystal structure (Fig. 5-4

and 5-5), suggesting improved complementarity and/or more limited conformational freedom of the larger ring in the active site. This could explain the increased binding affinity (K_d = 218 μM) and hence reactivity (65% yield) of **8** compared to **6** (K_d = 309 μM, 47% yield). In one monomer (*cyan* in Fig. 5-5B), the desosamine moiety is in salt-bridge contact with E94, while in the other monomer (*pink*), it is within electrostatic distance from both E94 and E85. Amino acid side chains in the active site are stabilized in a single conformation, with the exception of F178, which is represented by different conformers in chain A (*green*) and chain B (*ice blue*) (Fig. 5-5B). Similar to **6**, the co-crystal structure suggests that flipping-over (or rotation) of **8** in the active site could enable hydroxylation (likely at C6/C9 or C7/C8) on both faces of the ring. However, a more limited resolution of the product profile in the LCMS analysis (Fig. 5-S6) (unlike oxidations of compound **6**, where oxidation products were better distinguished by LCMS, Fig. 5-S4) prevents verification of this prediction, or further evaluation of the reaction stereoselectivity.

Despite previously demonstrated dynamics of PikC (40), no fit-induced conformational changes were observed in response to binding of the various carbolides, which could potentially prevent the substrate from wobbling in the active site. This observation contradicts the concept of "conformational plasticity" applied for mammalian P450 enzymes (42), and limits the role of PikC conformational dynamics to substrate access to and product release from the active site.

5.2.4 Antibacterial Activities of Synthetic Desosaminyl Derivatives.

In macrolide antibiotics, desosamine (or a related deoxysugar moiety) has been found to play a crucial role in the interaction of these important compounds with the 23S ribosomal RNA (the major drug target for macrolides) through a number of specific contacts with ribonucleotides in the peptidyl-transferase center (43). Accordingly, we surmised that our series of synthetic desosaminyl derivatives might possess antibacterial activity. As shown in Table 5-S3, the cyclic carbolides including **6**, **8**, **10**, and **12** displayed some inhibitory activities against selected microbial targets, while their corresponding aglycones were inactive, confirming the significance of desosamine for bioactivity. Remarkably, the aromatic derivative **19** showed similar or even higher

bioactivity compared to the natural macrolide antibiotics **2** and **5**. We also compared the bioactivity of **8** with its synthetic hydroxylated products. Unexpectedly, upon installation of a hydroxyl group at either the C7/C8 or C6/C9 positions (the four authentic standards, each of which contains a pair of diastereomers), the products lost activity. Although hydroxylation of **8** had a detrimental impact on biological activity against a limited number of microbial targets, installation of this functional group might enable a useful chemical handle for further functionalization (44) and subsequent generation of more potent therapeutics. Further analysis of these carbolides, and direct investigation of their presumed binding to ribosomal targets will provide important insights into the impact of PikC-mediated regioselective hydroxylation for development of new small molecules for treatment of microbial pathogens and other human diseases.

5.3 Summary

The work described in this report is based on the innate substrate flexibility of the biosynthetic CYP450 PikC enzyme that could be harnessed using a substrate engineering strategy *in vitro* due to its desosamine anchoring functionality. In contrast to previous chemical approaches that have exploited proximal directing influences or inherent differences in C-H bond strengths, our approach utilizes the interactions of an engineered bacterial P450 enzyme with aminosugar (desosamine)-linked substrates that otherwise lack all of the functionality of the native biosynthetic intermediates. This approach provides the unique opportunity to obtain high resolution crystallographic insights into the nature of the substrate-catalyst interactions, which has generally not been available with biomimetic and organometallic systems that achieve catalytic oxidation of unactivated C-H bonds. Attachment of desosamine to a number of unnatural aglycones (esp. cyclized aglycones) via an acetal linkage leads to productive enzyme binding and considerable regioselectivity in the hydroxylation reactions. However, the level of regio- and stereoselectivity observed for natural macrolide substrate YC-17 **1** and narbomycin **4** was not achieved, which infers synergistic contributions of both desosamine and macrolactone portions of these endogenous substrates. Insights gained from structural analysis in the current study suggest that selectivity could be improved by (i) rigidifying the substrate structure, (ii) limiting conformational freedom by optimizing the

complementarity between substrate size/shape and the volume of the active site, and (iii) increasing structural complexity of compounds to improve the specificity of their binding.

Currently, further investigations, including the applicability of this approach in selective oxidation of complex compounds of medicinal significance, the development of alternative and more chemically accessible glycosides as well as alternative molecular anchors, and the development of diverse linkage groups, are in progress. Additionally, the observed selectivity could be potentially exploited for building libraries of compounds with hydroxyl group functional handles for further synthetic transformations. Moreover, the ability to obtain enzyme/substrate co-crystal structures offers unique mechanistic insights as well as a means to engineer CYP450 enzymes that display greater regio- and stereoselectivity. This type of reagent engineering is particularly amenable to biological catalysts, and provides effective approaches to tailor the catalytic outcome toward a desired synthetic goal. Finally, these findings suggest that in addition to PikC, diverse natural product P450s can be further harnessed for a substrate engineering approach to selectively oxidize C-H bonds, and to facilitate further chemical diversification of both synthetic and natural product molecules of biological interest.

5.4 Experimental Procedures

Preparation of $PikC_{D50N}$-RhFRED — Using previously prepared pET28b-PikC-RhFRED as a template (38), site-directed mutagenesis was performed by following the QuikChange (Stratagene) protocol. The primers for mutagenesis were: forward, 5'-CACCCCCGAGGGGAATGAGGTGTGGCTGG-3'; reverse, 5'-CCAGCCACACCTCATTCCCCTCGGGGGTG-3'. Protein expression and purification of $PikC_{D50N}$-RhFRED followed the procedure developed previously (38).

$PikC_{D50N}$-RhFRED assay — The standard assay contains 5 μM $PikC_{D50N}$-RhFRED, 0.5 mM substrate, 2.5 mM NADPH, 0.25 Unit of glucose-6-phosphate dehydrogenase, and 5 mM glucose-6-phosphate for NADPH regeneration in 100 μl of reaction buffer (50 mM NaH_2PO_4, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol). The reaction was carried out at 30 ^oC for 3 h and terminated by extraction using 3 × 200 μl of

$CHCl_3$. The resulting organic extract was dried by N_2 and redissolved in 120 μl of methanol. The subsequent LC-MS analysis was performed on a ThermoFinnigan LTQ linear ion-trap instrument (Department of Pharmacology, University of Michigan) equipped with electrospray source and Surveyor HPLC system by using an XBridge™ C18 3.5 μm 150 mm reverse-phase HPLC column under following conditions: mobile phase (A = deionized water + 0.1% formic acid, B = acetonitrile + 0.1 % formic acid), 20% B for 3 min, 20-100% B over 25 min, 100% B for 5 min, 100-20% B over 1 min, 20% B for 15 min; flow rate, 0.21 ml/min. The substrate binding assays were performed following the previous report (38).

Crystallization, data collection and structure refinement — Crystallization conditions for $PikC_{D50N}$ complexes with **6** and **8** were identified by using commercial high throughput screening kits available in deep-well format (Hampton Research), a nanoliter drop-setting Mosquito robot (TTP Lab Tech) operating with 96-well plates, and a hanging drop crystallization protocol. Optimization of conditions was carried out manually in 24-well plates. The protein from the 1 mM stock was diluted to 0.2 mM by mixing with **6** or **8** dissolved at 2 mM in 10 mM Tris-HCl, pH 7.5. Crystals of $PikC_{D50N}$-**6** complex were obtained from 15% PEG 4000, 0.1 M Tris-HCl, pH 7.5; 200 mM $MgCl_2$. Crystals of the $PikC_{D50N}$-**8** complex were obtained from 12% PEG 8000, 0.1 M sodium cacodilate, pH 6.5, and 200 mM Li_2SO_4. Prior to data collection the crystals were cryo-protected by plunging them into a drop of reservoir solution supplemented with 20% glycerol. 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 HKL2000 software suite. Crystal structures were determined by molecular replacement using the atomic coordinates of the 2C6H (PDB ID) as a search model (Table 5-S2).

Antibacterial assay — Two-fold serial dilutions of 40 mM (10 mM for positive control erythromycin) DMSO solutions of test compounds with DMSO generated a series of stock solutions with concentrations ranging from 0.31 ~ 40 mM. Then, a ten-time dilution of each was performed by using dd H_2O to make a 31 ~ 4000 μM series. Cultures of

target strains were grown in appropriate media at 37 °C (30 °C for *Deinococcus radiodurans*) with shaking (180 rpm). An overnight seed culture was diluted to an OD_{600} of 0.05, grown to an OD_{600} of 0.4 ~ 0.6, back diluted to an OD_{600} of 0.004, and 45 µL of this diluted culture was added to each well of a 384-well microtiter plate that contained 5 µL of a given dilution of compound in dd H_2O. Plate cultures were grown for 16 h (60 h for *D. radiodurans*) and OD_{600} measurements were taken. All measurements were performed in duplicate.

Steady-state kinetics of PikC$_{D50N}$-RhFRED — The reaction contains 40 nM of PikC-RhFRED-D50N with 3.5 µM spinach ferredoxin and 0.1 U/ml spinach ferredoxin-$NADP^+$ reductase, 10~160 µM substrate in 400 µl of desalting buffer (50 mM NaH_2PO_4, pH 7.3, 1 mM EDTA, 0.2 mM DTE, 10% glycerol). After pre-incubation at 30 ^oC for 5 min, the reaction was initiated by adding 4 µl of 50 mM NADPH and 100 µl aliquots were taken at 0s, 20s, and 40s (or 0s, 30s, and 60s when substrate concentrations are greater than 100 µM) to thoroughly mix with 100 µl of methanol for reaction termination. After centrifugation at 13,000 g for 15 min to pellet protein, the supernatants were analyzed by HPLC. The HPLC conditions were: Xbridge C18 5 µm 250 mm reverse-phase column, 20-80% solvent B (A: deionized water + 0.1% trifluoroacetic acid, B: acetonitrile + 0.1% trifluoroacetic acid) at 1.0 ml/min over 18 min, UV wavelength 226 nm. The initial velocity of substrate consumption was deduced from decreased area under the curve (AUC) of specific substrate peaks. Finally, the data from duplicated experiments were fit to Michaelis-Menten equation.

Preparation of the oxidized product of desosaminyl pyrene 19 — An enzymatic reaction containing 0.8 µM PikC$_{D50N}$-RhFRED, 750 µM glucose-6-phosphate, 0.15 Unit/ml glucose-6-phosphate dehydrogenase, 500 µM NADPH, and 15 mg **19** in 200 ml reaction buffer (50 mM NaH_2PO_4, pH7.3, 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol) was carried out at 30 ^oC for 20 h. The hydroxylated product **20** (yield ~8% based on HPLC analysis) was purified by reverse phase C18 preparative HPLC. Upon chloroform extraction from the aqueous solution of **20**, 0.2 mg of **20** was recovered. The

oxidation site was determined based on the 600M ^{1}H NMR and COSY performed by using this small amount of **20** in $CDCl_3$.

5.5 Supplementary Information

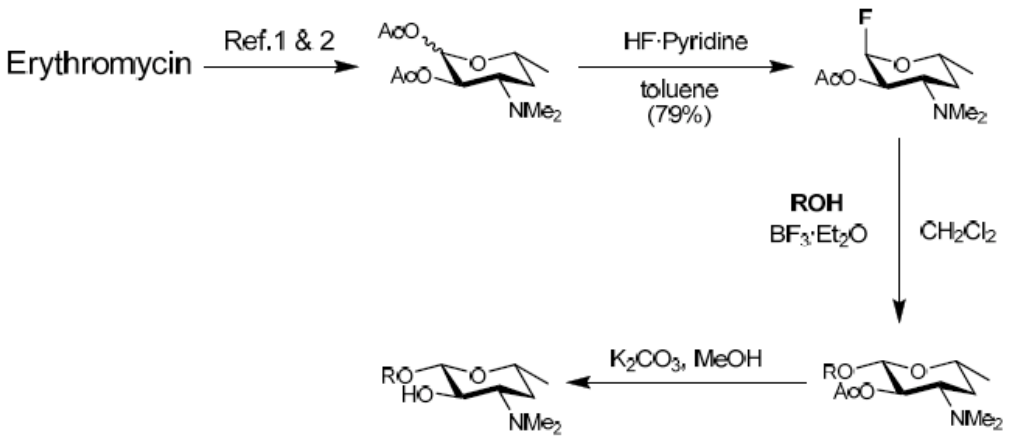


Figure 5-S1. General synthetic strategy for glycosylation of diverse alcohols with desosamine. [Ref.: 1. Chen, H., Yamase, H., Murakami, K., Chang, C.-w., Zhao, L., Zhao, Z. and Liu, H.-w. (2002) *Biochemistry* **41**, 9165–9183; 2. Anzai, Y., Li, S., Chaulagain, M. R., Kinoshita, K., Kato, F., Montgomery, J. and Sherman, D. H. (2008) *Chem. Biol.* **15,** 950-959]

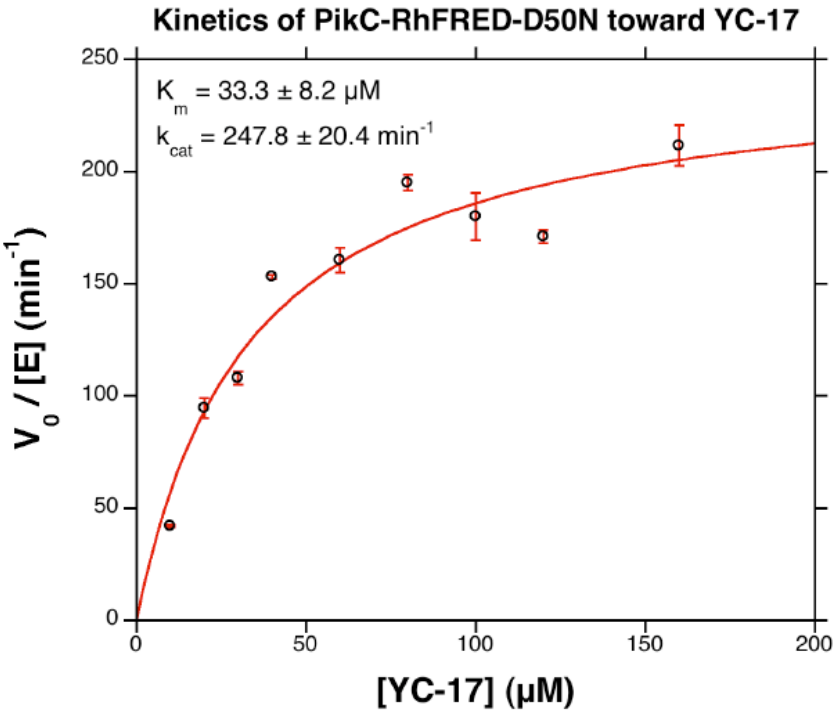


Figure 5-S2. Michaelis-Menten curve of PikC-RhFRED-D50N using YC-17 1 as substrate.

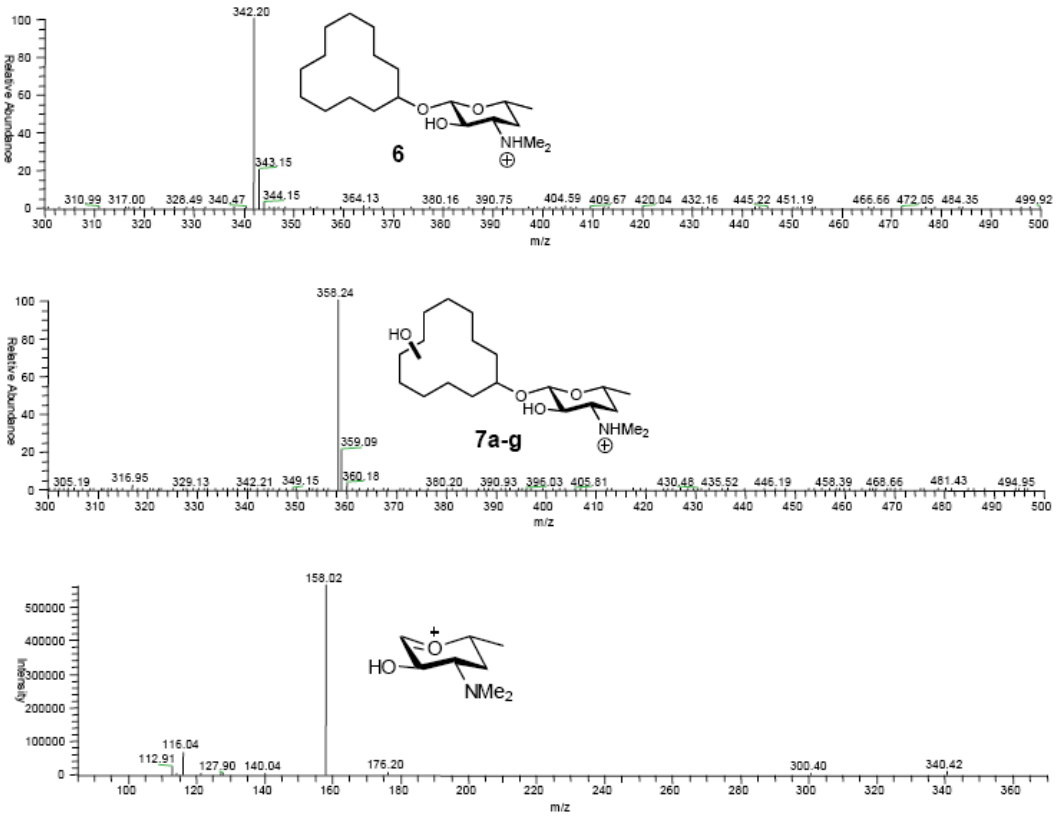


Figure 5-S3. Mass spectra of 6 (top panel) and its hydroxylated products 7a-g (middle panel). The bottom panel shows the MS-MS pattern of one of products. Notably, all desosaminyl derivatives in this study show the same MS-MS pattern.

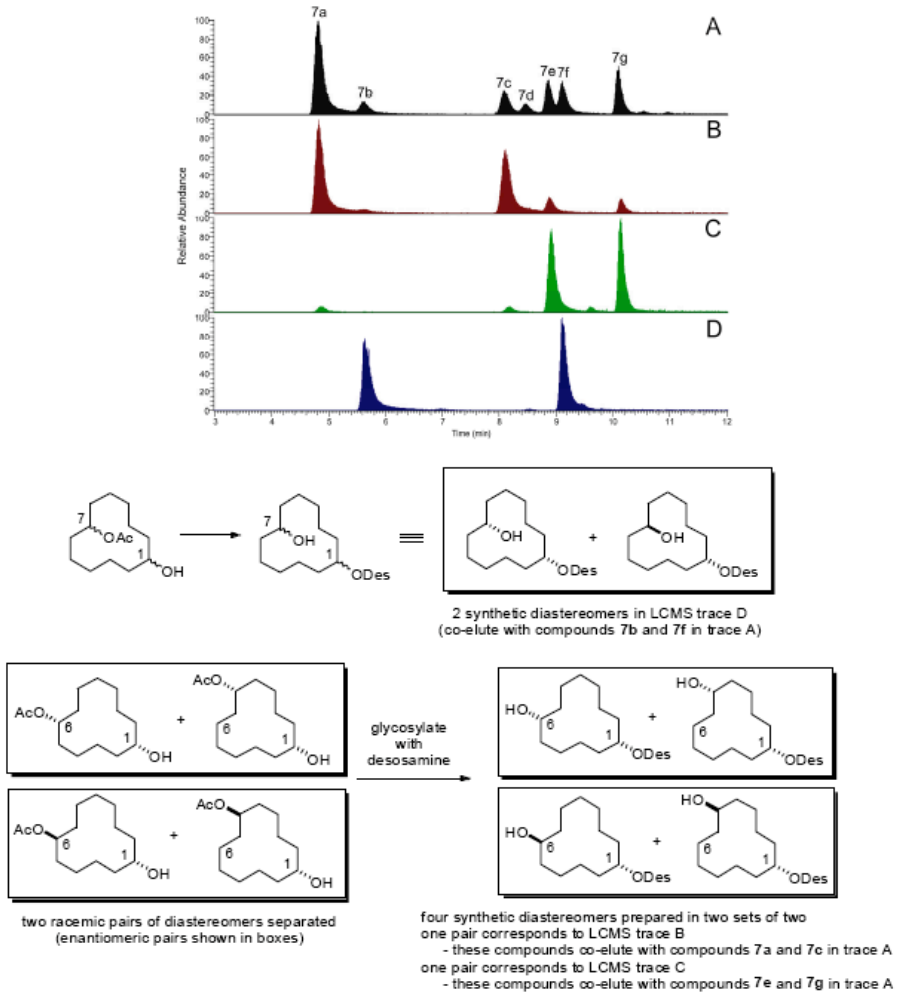


Figure 5-S4. Structural determination of mono-hydroxylated products of 6 through LC-MS comparison of synthetic authentic standards to 7a-g regarding retention times. (A) Product profile of $PikC_{D50N}$-RhFRED reaction using **6** as substrate; (B) Authentic standard containing a pair of C6/C8 hydroxylated diastereomers; (C) Authentic standard containing the other pair of C6/C8 hydroxylated diastereomers; (D) Authentic standard containing the pair of C7 hydroxylated diastereomers. The products assignment was further confirmed by co-injections.

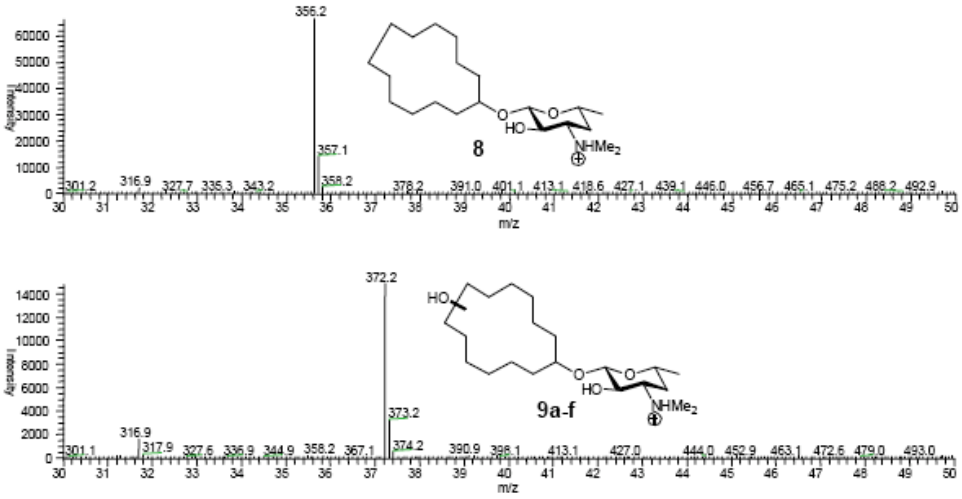


Figure 5-S5. Mass spectra of 8 (upper panel) and its hydroxylated products 9a-f (lower panel).

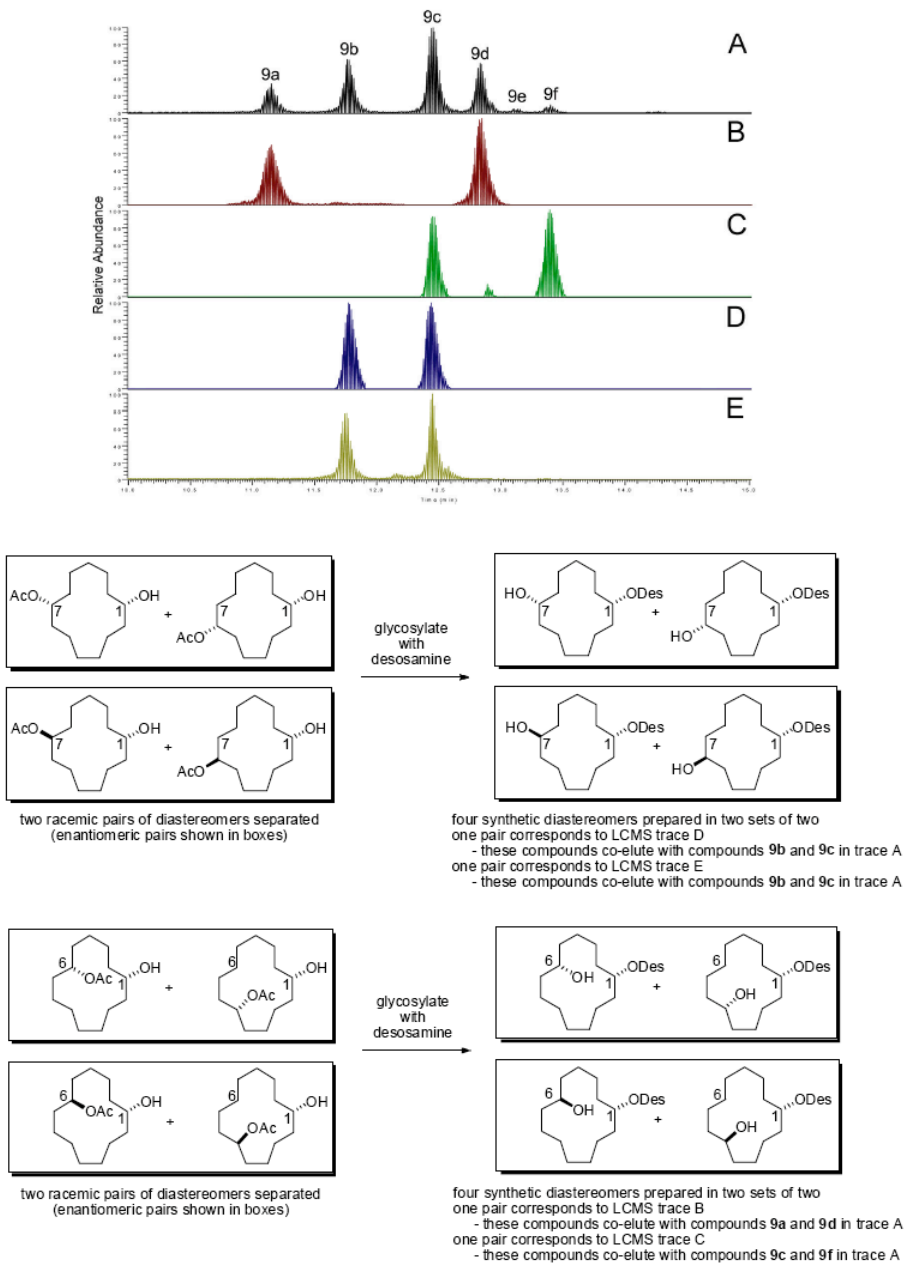


Figure 5-S6. Structural determination of mono-hydroxylated products of 8 through LC-MS comparison of synthetic authentic standards to 9a-f regarding retention times. (A) Product profile of $PikC_{D50N}$-RhFRED reaction using **8** as substrate; (B) Authentic standard containing a pair of C6/C9 hydroxylated diastereomers; (C) Authentic standard containing the other pair of C6/C9 hydroxylated diastereomers; (D) Authentic standard containing a pair of C7/C8 hydroxylated diastereomers; (E) Authentic standard containing the other pair of C7/C8 hydroxylated diastereomers. The products assignment was further confirmed by co-injections. Some diastereomers are not distinguishable due to identical retention times (e.g. traces D and E) Notably, to get a better separation of similar diastereomers, an optimized LC condition other than the one described in the main text was employed: mobile phase, 10% B for 3 min, 10-80% B over 25 min, 80-100% B over 1 min, 100% B for 5 min, 100-10% solvent B over 2 min, 10% solvent B for 15 min. All other conditions remained the same.

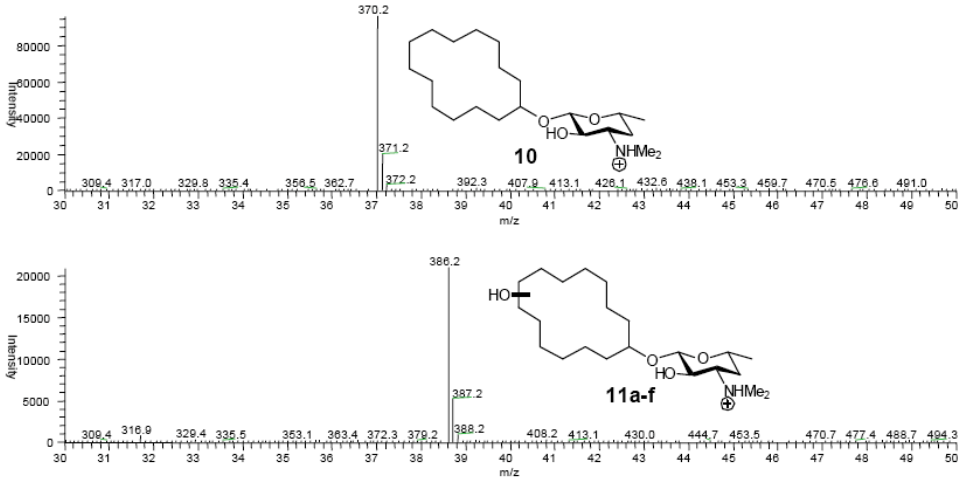


Figure 5-S7. Mass spectra of 10 (upper panel) and its hydroxylated products 11a-f (lower panel).

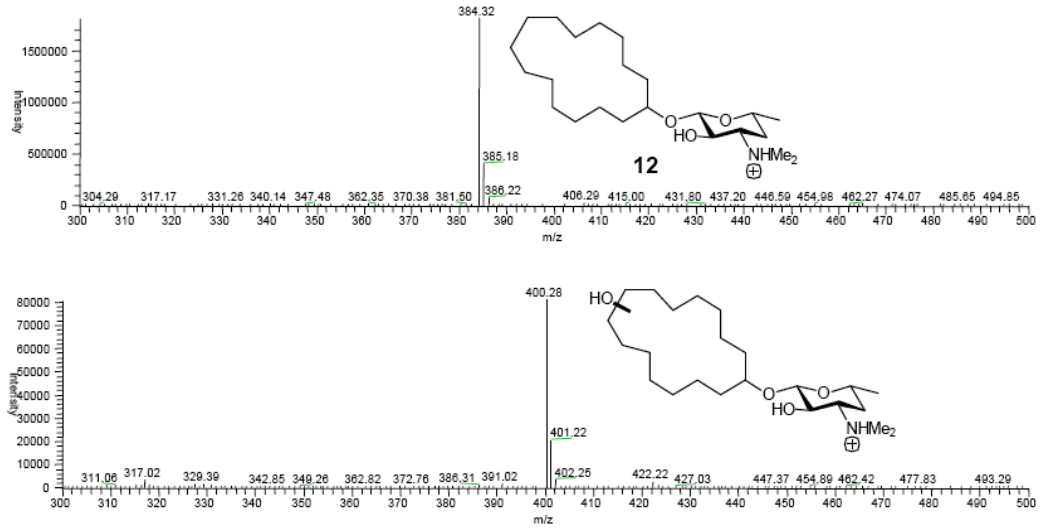


Figure 5-S8. Mass spectra of 12 (upper panel) and its hydroxylated products (lower panel).

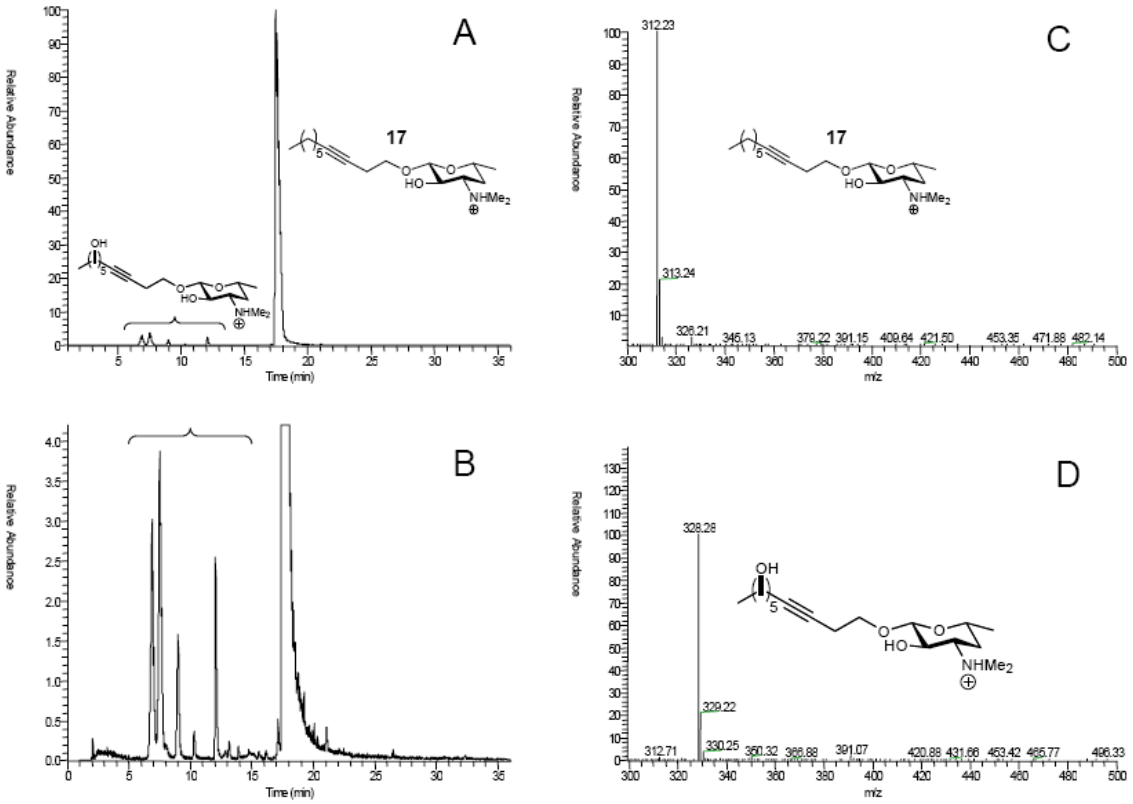


Figure 5-S9. (A) Product profile of $PikC_{D50N}$-RhFRED reaction using linear desosaminyl derivative **17** as substrate. (B) Amplified product profile to visualize the hydroxylated products in small amounts. Mass spectra of **17** (C) and its hydroxylated products (D).

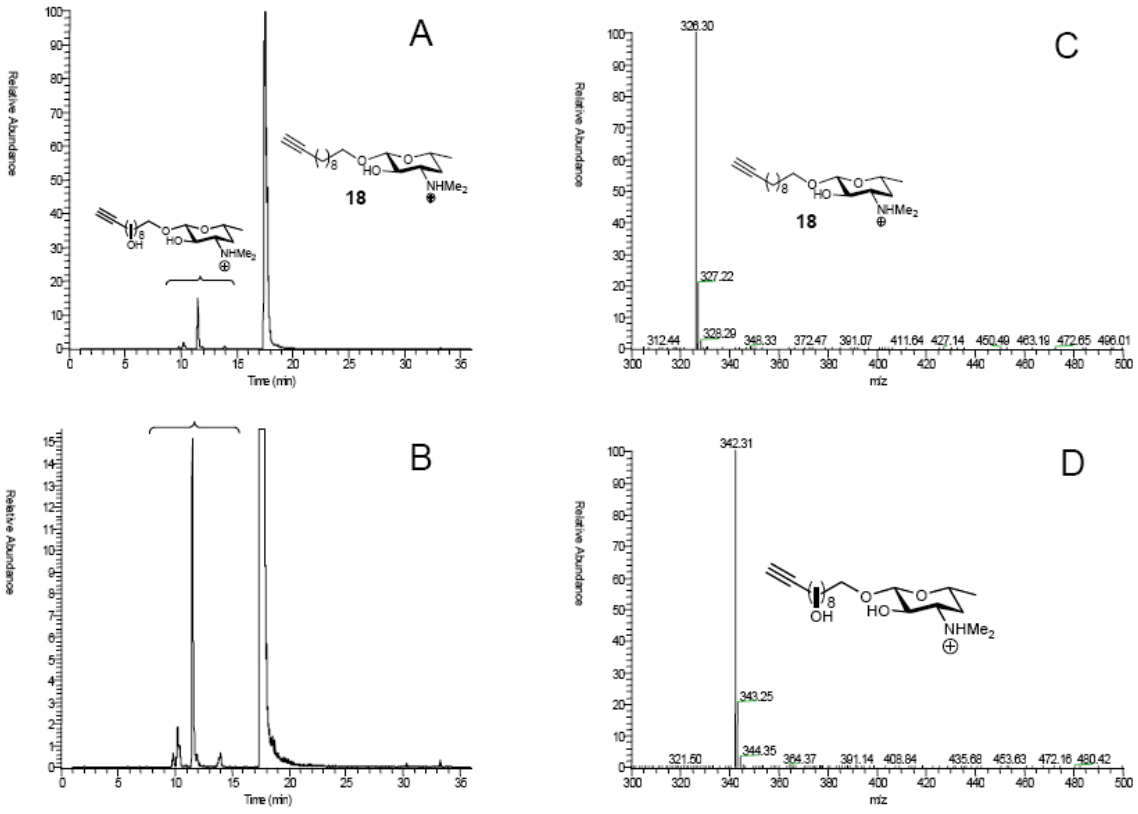


Figure 5-S10. (A) Product profile of $PikC_{D50N}$-RhFRED reaction using linear desosaminyl derivative **18** as substrate. (B) Amplified product profile to visualize the hydroxylated products in small amounts. Mass spectra of **18** (C) and its hydroxylated products (D).

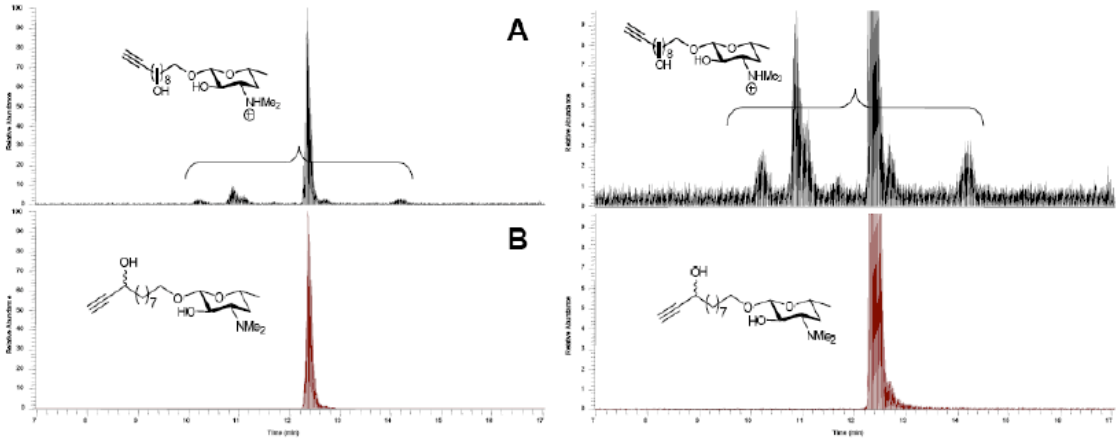


Figure 5-S11. Product identification of desosaminyl derivative 18. (A) Product profile of $PikC_{D50N}$-RhFRED reaction using linear desosaminyl derivative **18** as substrate. (B) Authentic standard containing a pair of C9 hydroxylated diastereomers. The products assignment was further confirmed by co-injection. The right panel shows the amplified chromatograms to visualize the hydroxylated products in small amounts.

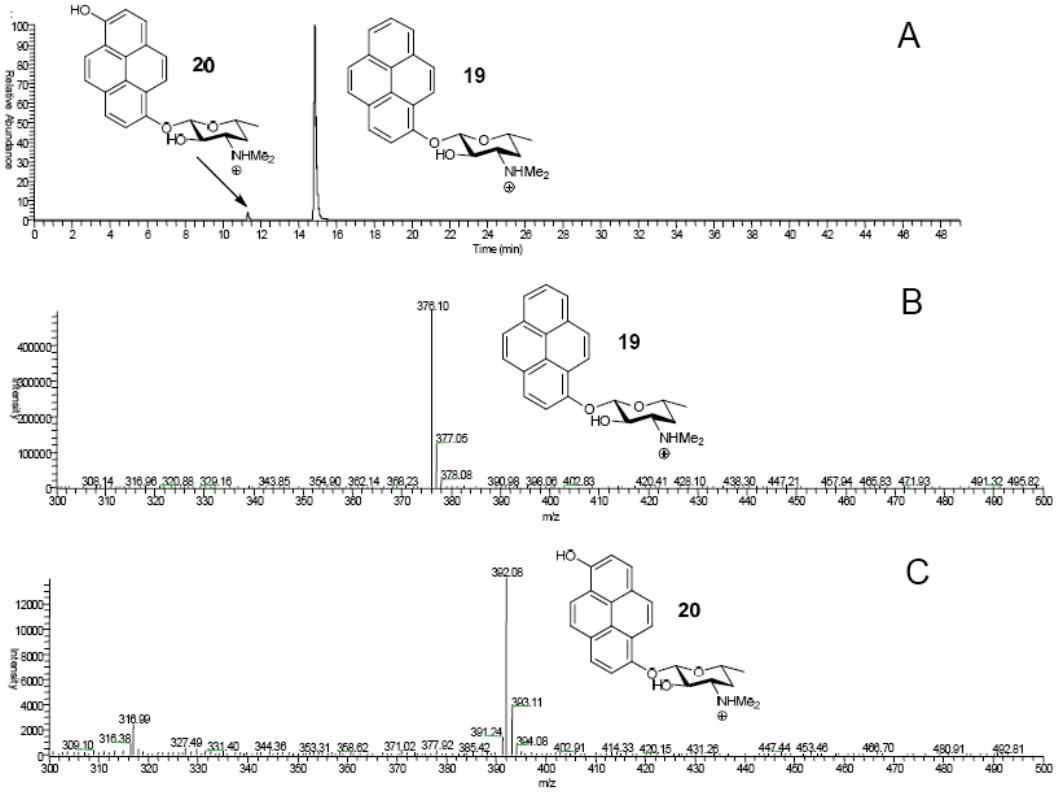


Figure 5-S12. (A) Product profile of $PikC_{D50N}$-RhFRED reaction using aromatic desosaminyl pyrene **19** as substrate. Mass spectra of **19** (B) and its C8 hydroxylated product **20** (C).

Structural determination of 20

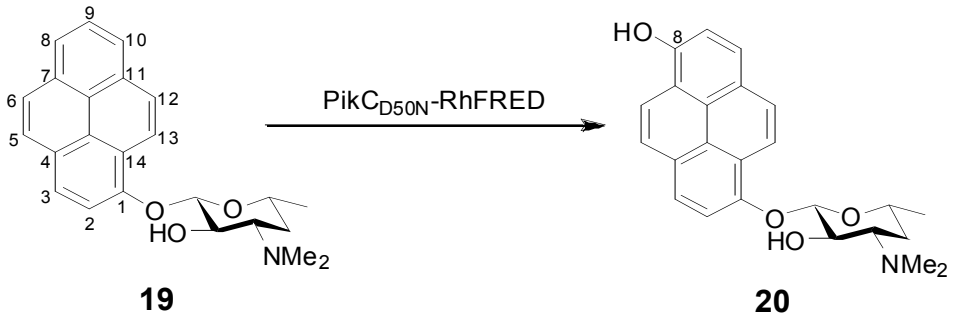


Table 5-S1. ^{1}H NMR data of 19 and 20

Proton Position	Proton assignment (ppm) in $CDCl_3$	
	19	**20**
H-2	7.81 (d, J = 8.3 Hz, 1H)	7.78 (d, J = 8.3 Hz, 1H)
H-3	8.12 (d, J = 8.3 Hz, 1H)	8.02 (d, J = 8.3 Hz, 1H)
H-5	7.94 (d, J = 9.1 Hz, 1H)	7.95 (d, J = 9.3 Hz, 1H)
H-6	8.00 (d, J = 9.1 Hz, 1H)	8.17 (d, J = 9.1 Hz, 1H)
H-8	8.13 (d, J = 7.7 Hz, 1H)	-
H-9	7.96 (t, J = 7.7, 7.7 Hz, 1H)	7.44 (d, J = 8.1 Hz, 1H)
H-10	8.14 (d, J = 7.7 Hz, 1H)	7.96 (d, J = 8.2 Hz, 1H)
H-12	8.06 (d, J = 9.1 Hz, 1H)	7.92 (d, J = 9.2 Hz, 1H)
H-13	8.63 (d, J = 9.1 Hz, 1H)	8.41 (d, J = 9.1 Hz, 1H)

The assignment of the aromatic protons of desosaminyl pyrene **19** was conducted based on a full set of 1D and 2D NMR spectra and consistency with the published ^{1}H NMR data for 1-hydroxypyrene (Lambert, M., Kremer, S., Sterner, P., and Anke, H. (1994) *App. Environ. Microbiol,* **60**, 3597-3601). The structure of **20** was determined based on the 600M ^{1}H NMR and COSY spectra, using **19** as reference.

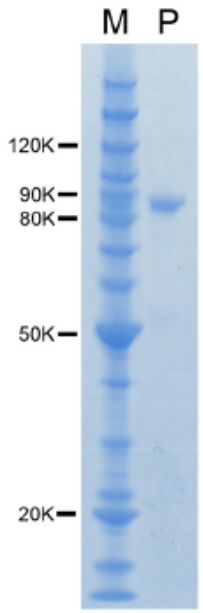


Figure 5-S13. SDS-PAGE analysis of purified $PikC_{D50N}$-RhFRED (M, Marker; P, Purified $PikC_{D50N}$-RhFRED)

Table 5-S2. Crystallographic data and statistics.

Protein complex **PDB ID**	**PikC-cyclododecane** **PDB ID 2WI9**	**PikC-cyclotridecane** **PDB ID 2WHW**
Data collection		
Wavelength, Å	1.11587	1.11587
Resolution, Å	2.0	2.2
Unique reflections	68456	52386
Redundancy	3.9 (3.4)[a]	4.2 (4.3)
Completeness, %	98.2 (87.7)	100.0 (99.8)
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions (a,b,c), Å (α, β, γ), °	65.3, 109.1, 153.2 90.0, 90.0, 90.0	60.2, 109.5, 153.1 90.0, 90.0, 90.0
Molecules in asymmetric unit	2	2
Solvent content, %	55.1	55.2
R_{sym}[b], %	7.2 (39.0)	9.8 (49.7)
I/σ	21.2 (2.7)	20.9 (3.6)
Refinement		
Reflections used in refinement		49541
$R_{cryst}(R_{free})$[c], %	18.9/23.9	17.6/25.6
No. of atoms	6978	7098
Protein	6229	
Heme	86	86
Ligand	72	
Water/ions	576/15	/10
Wilson plot B-values, $Å^2$	30.0	29.8
Mean B-factor, $Å^2$	31.2	29.46
Protein	30.8	27.0
Heme	16.4	12.6
Ligand	**A:** 62.9, **B:** 35.3	**A:** 65.9, **B:** 57.0
Water/ions	34.8/85.3	35.3/84.8
r.m.s. deviations		
Bond length, Å	0.022	0.024
Bond angles, °	1.9	1.9
Ramachandran[d] (%)	**A:** 89.9/9.2/0.6/0.3 **B:** 92.3/7.7/0.0/0.0	**A:** 92.0/7.4/0.3/0.3 **B:** 92.3/7.7/0.0/0.0

[a]Numbers in parentheses correspond to the highest resolution shell.

[b]$R_{sym}=\Sigma|I_i - \langle I\rangle|/\Sigma I_i$, where I_i is the intensity of the i^{th} observation, and $\langle I\rangle$ is the mean intensity of reflection.

[c]$R_{cryst}=\Sigma\ ||Fo|-|Fc||/\Sigma|Fo|$, calculated with the working reflection set. R_{free} is the same as R_{cryst} but calculated with the reserved reflection set.

[d]Program PROCHECK (Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton J. M. 1993, PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283-291), portions of the protein residues in most favored/additional allowed/generously allowed/disallowed regions.

Table 5-S3. Antibacterial activities of desosaminyl derivatives against selected strains.

	Minimal inhibitory concentration (μM)							
	Kocuria rhizophila ATCC9341[a]	*Staphylococcus aureus* ATCC6538P	*Bacillus subtillis* DHS5333	*Deinococcus radiodurans*[b]	*Escherichia. coli* TolC[c]	*S. aureus* NorA[d]	*Acinetobacter baumannii*	Multidrug resistant *S. aureus* (MRSA)
Media	Nutrient broth	LB	LB	Special media[f]	Mueller Hinten	Mueller Hinten	Mueller Hinten	Mueller Hinten
Compound								
Erythromycin	< 0.8	< 0.8	< 0.8	6.2	< 0.8	< 0.8	25	> 100
DMSO	-	-	-	-	-	-	-	-
1	100	100	100	400	100	100	> 400	> 400
2	50	100	100	200	50	100	> 400	> 400
4	< 3.1	6.2	< 3.1	100	25	3.1	> 400	> 400
5	6.2	25	12.5	100	100	12.5	> 400	> 400
6	50	> 400	200	100	50	200	> 400	400
8	50	400	200	100	50	200	> 400	200
10	50	> 400	200	50	50	400	> 400	400
12	50	200	100	50	50	100	> 400	100
13	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400
14	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400
15	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400
8-C6OH-a[e]	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400
8-C6OH-b[e]	> 400	> 400	> 400	400	> 400	> 400	> 400	> 400
8-C7OH-a[e]	> 400	> 400	> 400	400	> 400	> 400	> 400	> 400
8-C7OH-b[e]	> 400	> 400	> 400	100	400	> 400	> 400	> 400
17	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400
18	> 400	> 400	> 400	> 400	400	> 400	> 400	> 400
19	25	25	6.2	25	25	25	100	25
Cyclododecanol	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400
Cyclotridecanol	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400
Cyclopentadecanol	> 400	> 400	> 400	> 400	> 400	> 400	> 400	> 400

[a] Previously known as *Micrococcus luteus* ATCC 9341, which is sensitive to macrolide antibiotics.

[b] A gift from Prof. Ada E. Yonath (Structural Biology Department, Weizmann Institute of Science, Rehovot, Israel)

[c] *E. coli* W3110 TolC disruption mutant that is more sensitive to antibiotics

[d] *S. aureus* 8325 NorA disruption mutant that is more sensitive to antibiotics

[e] **8**-C6OH-a, b and **8**-C7OH-a, b correspond to synthetic C6/C9 and C7/C8 hydroxylated standards of 13-membered carbolide, each of which contains a pair of diastereomers as shown in trace B, C and D, E, respectively in Fig. S6

[f] Media recipe: 10 g caseine peptone (tryptic digest), 5 g yeast extract, 5 g NaCl and 5 g glucose (add after sterilization) in 1 liter water, pH 7.2.

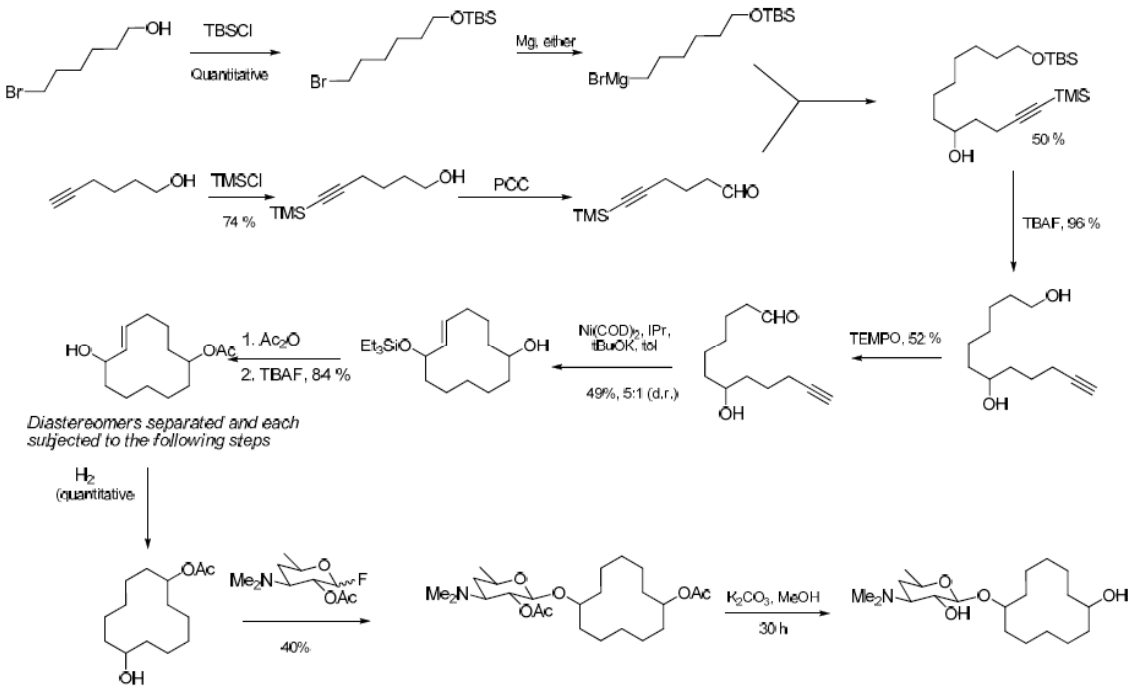


Figure 5-S14. Synthesis of C-7 hydroxylated authentic compounds in 12-membered ring series.

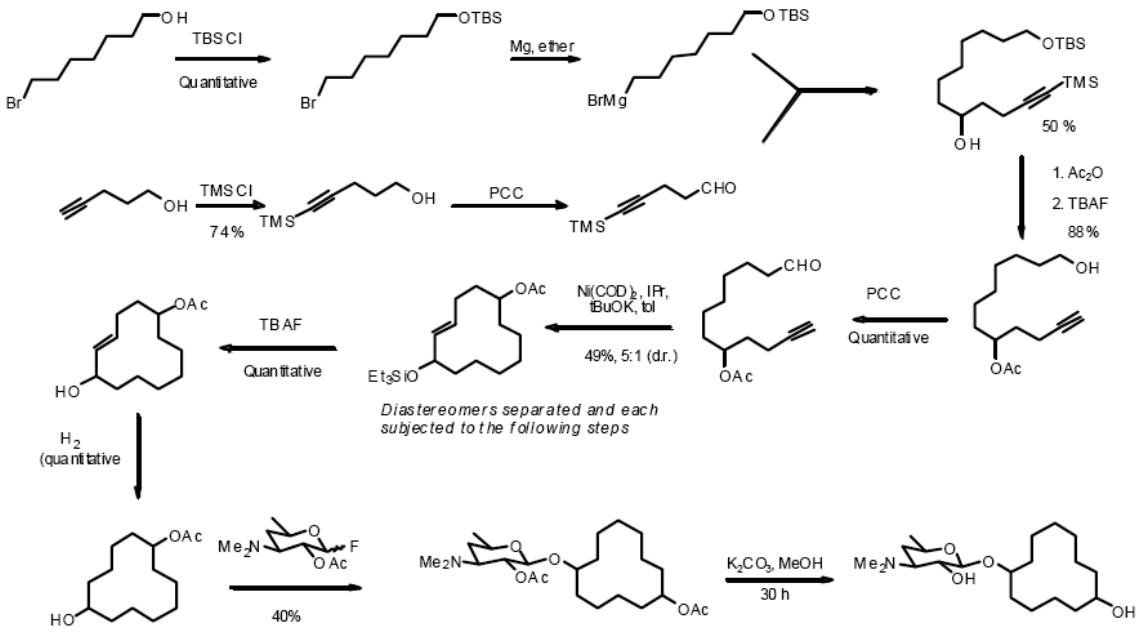


Figure 5-S15. Synthesis of C-6/C-8 hydroxylated authentic compounds in 12-membered ring series.

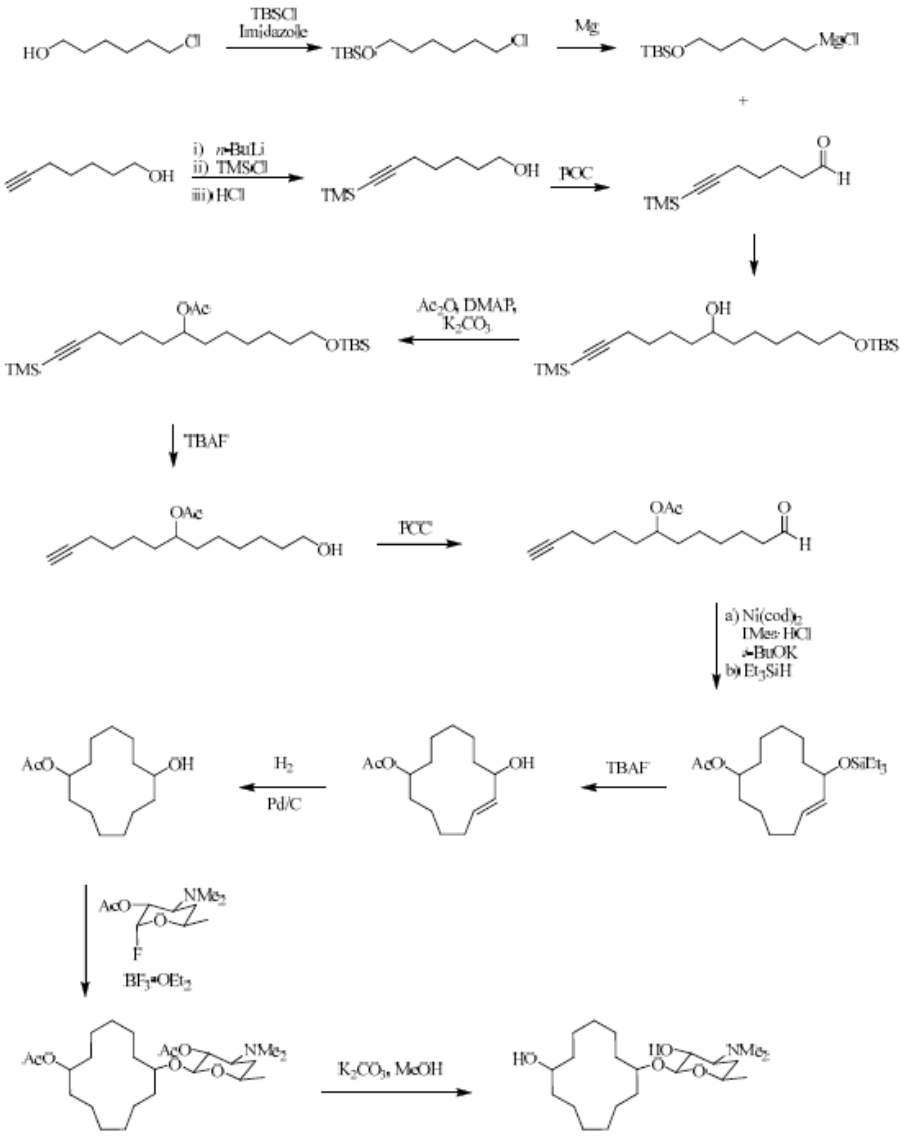


Figure 5-S16. Synthesis of C-7/C-8 hydroxylated authentic products in 13-membered ring series.

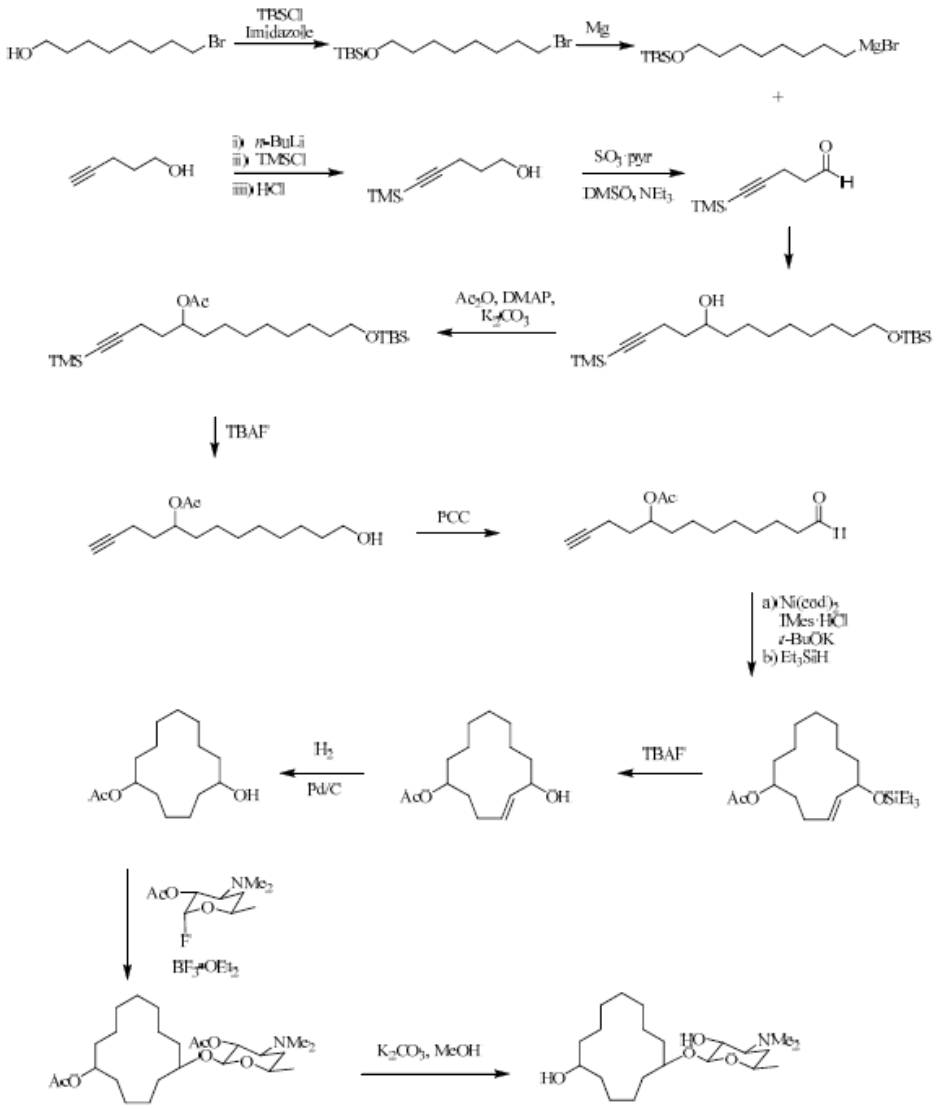


Figure 5-S17. Synthesis of C-6/C-9 hydroxylated authentic products in 13-membered ring series.

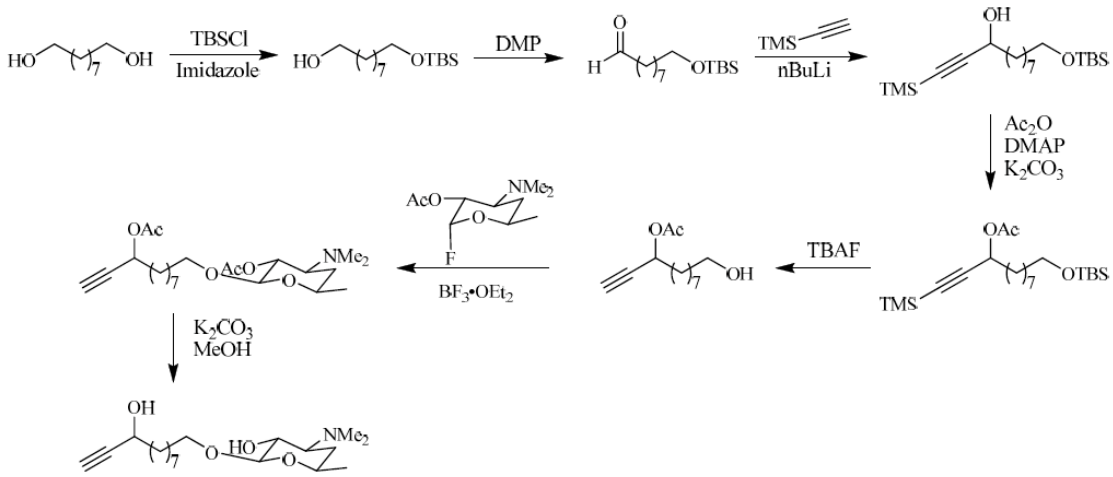


Figure 5-S18. Synthesis of authentic propargyl alcohol.

5.6 References

1. Ortiz de Montellano, P. R. (1995) Cytochrome P450: Structure, mechanism, and biochemistry. 2nd Ed. *Plenum Press, New York*.
2. Guengerich, F. P. (2001) *Chem. Res. Toxicol.* **14**, 611-650
3. Coon, M. J. (2005) *Ann. Rev. Pharmacol. Toxicol.* **45**,1-25
4. Dick, A. R., and Sanford, M. S. (2006) *Tetrahedron* **62**, 2439-2463
5. Crabtree, R. H. (2001) *J. Chem. Soc., Dalton Trans.* 2437-2450
6. Bergman, R. G. (2007) *Nature* **446**, 391-393
7. Labinger, J. A. and Bercaw, J. E. (2002) *Nature* **417**, 507-514
8. Godula, K., and Sames, D. (2006) *Science* **312**, 67-72
9. Jr Que, L. and Tolman, W. (2008) *Nature* **445**, 333-340
10. Chen, M. S., and White, C. M. (2007) *Science* **318**,783-787
11. Brodsky, B. H. and Du Bois, J. (2005) *J. Am. Chem. Soc.* **127**, 15391-15393
12. Wender, P. A., Hilinski, M. K., and Mayweg, A. V. (2005) *Org. Lett.* **7**, 79-82
13. Lee, S. and Fuchs, P. L. (2002) *J. Am. Chem. Soc.* **124**,13978-13979
14. Chen, H., Schlecht, S., Semple, T. C., and Hartwig, J. F. (2000) *Science* **287**, 1995-1997
15. Cook, B. R., Reinert, T. J., and Suslick, K. S. (1986) *J. Am. Chem. Soc.* **108**, 7281-7286
16. Desai, L. V., Hull, K. L., and Sanford, M. S. (2004) *J. Am. Chem. Soc.* **126**, 9542-9543
17. Dick, A. R., Hull, K. L., and Sanford, M. S. (2004) *J. Am. Chem. Soc.* **126**, 2300-2301.
18. Dangel, B. D., Johnson, J. A., and Sames, D. (2001) *J. Am. Chem. Soc.* **123**, 8149-8150
19. Das, S., Incarvito, C. D., Crabtree, R. H., and Brudvig, G. W. (2006) *Science* **312**, 1941-1943
20. Das, S., Brudvig, G. W., and Crabtree, R. H. (2008) *J. Am. Chem. Soc.* **130**, 1628-1637
21. Breslow, R. *et al.* (1973) *J. Am. Chem. Soc.* **95**, 3251-3262
22. Yang, J., Gabriele, B., Belvedere, S., Huang, Y., and Breslow, R. (2002) *J. Org. Chem.* **67**, 5057-5067
23. Grieco, P. A. and Stuk, T. L. (1990) *J. Am. Chem. Soc.* **112**, 7799-7801
24. Groves, J. T. and Neumann, R. (1989) *J. Am. Chem. Soc.* **111**, 2900-2909
25. Urlacher, V. B. and Eiben, S. (2006) *Trends Biotechnol.* **24**, 324-330
26. Landwehr, M., Carbone, M., Otey, C. R., Li, Y., and Arnold, F. H. (2007) *Chem. Biol.* **14**, 269-278
27. Sieber, V., Martinez, C. A., and Arnold, F. H. (2001) *Nat. Biotechnol.* **19**, 456-460
28. Arnold, F. H. (2006) *Nat. Biotechnol.* **24**, 328-330
29. Bloom, J. D. *et al.* (2005) *Curr. Opin. Struc. Biol.* **15**, 447-452
30. Kumar, S., Chen, C. S., Waxman, D. J., and Halpert, J. R. (2005) *J. Biol. Chem.* **280**, 19569-19575
31. Bloom, J. D. and Arnold, F. H. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9995-10000
32. Xue, Y., Wilson, D., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Chem. Biol.* **5**, 661-667
33. Anzai, Y*., et al.* (2008) *Chem. Biol.* **15**, 950-959
34. Griengl, H. & de Raadt, A. (2002) *Curr. Opin. Biotechnol.* **13**, 537-542
35. Lairson, L. L., Watts, A. G., Wakarchuk, W. W., and Withers, S. G. (2006) *Nat. Chem. Biol.* **2**, 724-728
36. Braunegg, G. *et al.* (1999) *Angew. Chem. Int. Ed.* **38**, 2763-2766
37. de Raadt, A., Griengl, H., and Weber, H. J. (2001) *Chem. Eur. J.* **7**, 27-31
38. Li, S., Podust, L. M., and Sherman, D. H. (2007) *J. Am. Chem. Soc.* **129**, 12940-12941
39. Xue, Y., Zhao, L., Liu, H.-w., and Sherman, D. H. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 12111-12116
40. Sherman D.H. *et al*. (2006) *J. Biol. Chem.* **281**, 26289-26297

41. Li, S., Ouellet, H., Sherman, D. H., and Podust, L. M. (2009) *J. Biol. Chem.* **284**, 5723-5730
42. Muralidhara, B. K. and Halpert, J. R. (2007) *Drug. Metab. Rev.* **39**, 539-556
43. Schlünzen, F. *et al.* (2001) *Nature* **413**, 814-821
44. Rentmeister, A., Arnold, F. H., and Fasan, R. (2009) *Nat. Chem. Biol.* **5**, 26-28

Notes:

During preparation of this chapter, this work has been accepted for publication on ***Proc. Natl. Acad. Sci. U.S.A.*** as "Selective oxidation of carbolide C-H bonds by an engineered macrolide P450 monooxygenase." **Li, S.**, Chaulagain, M. R., Knauff, A. R., Podust, L. M., Montgomery, J., and Sherman, D. H. (***In Press***)

Author contributions:
Shengying Li, John Montgomery, Larissa M. Podust, and David H. Sherman designed the experiments;
Shengying Li performed the biochemistry experiments and analyzed the crystal structures;
Mani Raj Chaulagain, and Allison R. Knauff performed the organic synthesis;
Larissa M. Podust solved and analyzed the crystal structures.

Chapter 6
Functional Analysis of MycCI and MycG, Cytochrome P450 Enzymes Involved in Biosynthesis of Mycinamicin Macrolide Antibiotics

6.1 Introduction

The cytochrome P450 enzymes (P450s) form a very large family of oxidative hemeproteins that are responsible for diverse oxidative transformations across most life forms (1, 2). These reactions typically involve modification of physiologic and xenobiotic compounds and include the biosynthesis of various bioactive compounds (e.g., steroids, antibiotics, and signaling molecules). Recent bacterial genome sequencing efforts have uncovered an unexpected large number of genes encoding P450 enzymes. For example, the model actinomycete *Streptomyces coelicolor* A3(2) that produces actinorhodin and undecylprodigiosin revealed the presence of 18 different P450 genes (3), whereas *Streptomyces avermitilis* MA-4680, an avermectin producer, contains 33 P450s (4), and *Saccharopolyspora erythraea* NRRL 23338, the erythromycin-producing bacterium, encodes 36 P450s (5). In secondary metabolic pathways, it is typical that P450 genes are integrated within the biosynthetic cluster, where their products catalyze regio- and stereospecific oxidation of precursors leading to structural diversity as well as improved bioactivities of these molecules (6, 7). Thus, cytochrome P450 enzymes EryF (8) and EryK (9) that are encoded within the erythromycin biosynthetic gene cluster are involved in the biosynthesis of erythromycin A. Specifically, EryF hydroxylates the macrolactone precursor 6-deoxyerthronolide B, whereas EryK is a macrolide hydroxylase resulting in formation of erythromycin D. As prototypic P450 hydroxylases involved in secondary metabolism, EryF and EryK exhibit strict substrate specificity. In contrast,

PikC cytochrome P450 involved in the methymycin/neomethymycin and pikromycin biosynthetic pathway of *Streptomyces venezuelae* has broader substrate tolerance (10). PikC catalyzes the final hydroxylation step toward the 12-membered ring macrolide YC-17 and the 14-membered ring macrolide narbomycin to produce methymycin/neomethymycin and pikromycin as major products.

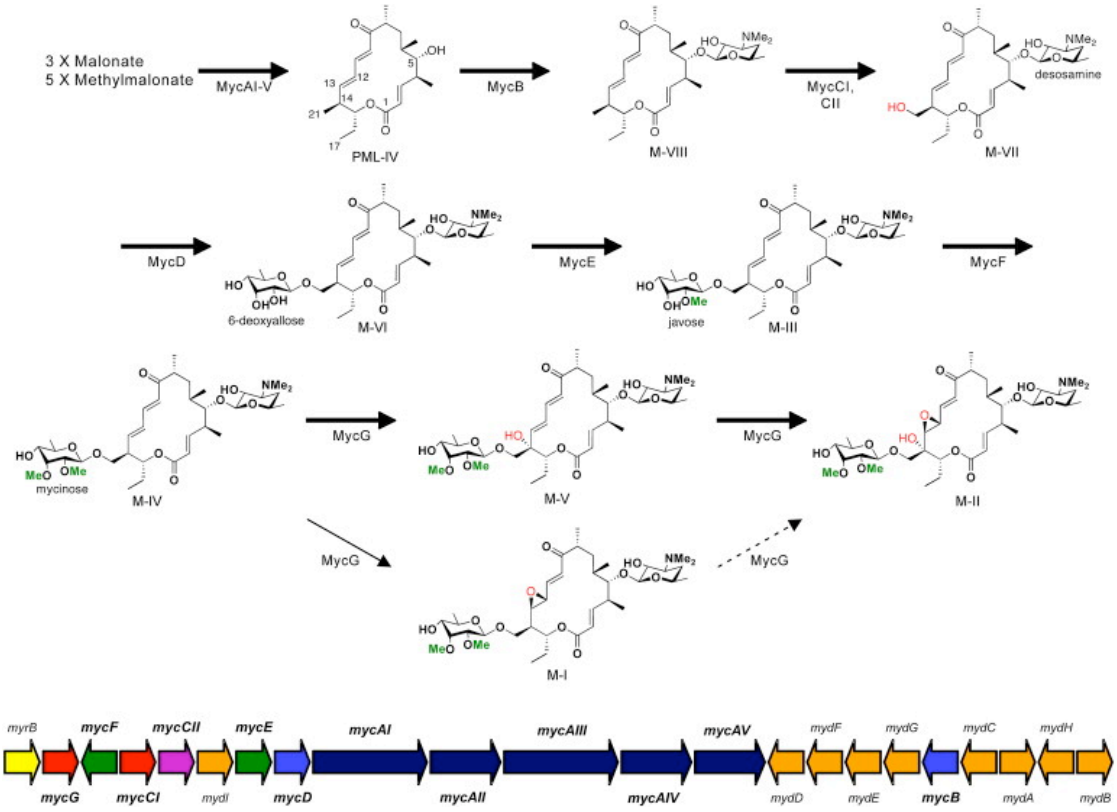


Figure 6-1. Mycinamicin post-PKS biosynthetic pathway and organization of the mycinamicin biosynthetic gene cluster. Methyl groups introduced by *O*-methyltransferases and oxidation steps catalyzed by cytochrome P450 enzymes are highlighted in green and red, respectively. Bold arrows represent the main pathway, whereas thin arrows indicate a shunt pathway. The dashed arrow indicates low-level conversion, as reported elsewhere (17). Color codes in the myc gene cluster are as follows: red, cytochrome P450 genes; purple, ferredoxin gene; indigo blue, polyketide synthetase genes; blue, glycosyltransferase genes; green, O-methyltransferase genes; orange, deoxysugar biosynthetic genes; and yellow, self-resistance gene (rRNA methyltransferase gene).

Mycinamicins, a series of macrolide antibiotics produced by the rare actinomycete *Micromonospora griseorubida*, have shown impressive activities against a spectrum of Gram-positive strains, especially some antibiotic-resistant human pathogens (11-13). Stucturally, the major mycinamicin products of wild-type strain *M. griseorubida*

A11725, including mycinamicin I (M-I), II (M-II), IV (M-IV), and V (M-V) (Fig. 6-1), consist of a 16-membered ring polyketide macrolactone substituted with 6-deoxyhexose sugars desosamine and mycinose. Partial characterization of the biosynthetic pathway for mycinamicins has been obtained through analysis of blocked mutants and corresponding bioconversion studies (13, 14). More recently, the nucleotide sequence of the complete mycinamicin biosynthetic gene cluster has been reported (15), wherein two putative P450 genes mycCI and mycG were identified (Fig. 6-1).

Analysis of the 5′ region of the myc gene cluster upstream from the PKS locus revealed that *mycCI* is located adjacent to *mycCII*, which encodes a putative ferredoxin (15). Since the deduced amino acid sequences of *mycCI* and *mycCII* show high sequence similarities to TylHI and TylHII (Fig. 6-2), respectively, that are likely responsible for hydroxylation at the C23 methyl group of tylactone (16), the function of MycCI and MycCII was accordingly proposed to mediate hydroxylation at the analogous C21 methyl group of protomycinolide IV (PML-IV) (15) (Fig. 6-1). On the basis of genetic complementation analysis of a targeted mutant strain of *M. griseorubida*, mycG was presumed to encode a P450 enzyme that catalyzes both hydroxylation and epoxidation at C14 and C12/13 on the macrolactone ring of mycinamicin (13, 17). In the current study, *mycCI* and *mycG* genes were overexpressed in *Escherichia coli*, and the functions of purified MycCI and MycG proteins were determined *in vitro* using natural substrates derived from semi-synthesis or isolated from wild-type or engineered strains of *M. griseorubida* that accumulate key mycinamycin intermediates. Moreover, specific roles have been proposed for both the desosamine and mycinose sugar residues in the oxidative cascade leading to M-II, the final product in the pathway.

6.2 Results

6.2.1 Protein Sequence Analysis of MycCI and MycG

Comparison of the deduced amino acid sequences of MycCI and MycG showed relatively low sequence identity (33%). In the phylogenetic tree (Fig. 6-2) of selected bacterial macrolide biosynthetic P450 enzymes, they were clustered in distinct branches, suggesting these two unrelated P450 genes may have been integrated into the mycinamicin biosynthetic gene cluster sequentially from different ancestors, as opposed

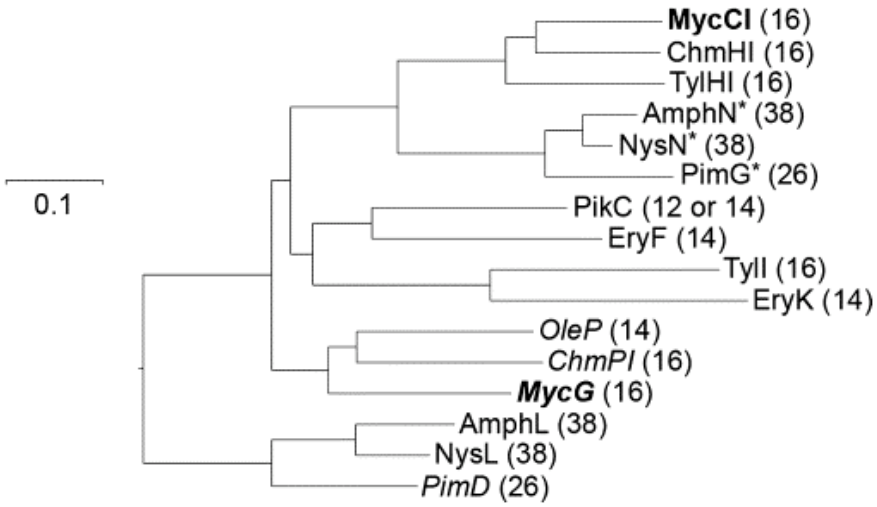


Figure 6-2. Phylogenetic tree of macrolide biosynthetic P450 monooxygenases. The selected cytochrome P450s include OleP (accession number AAA92553) [oleandomycin pathway], ChmHI and ChmPI (accession numbers AAS79447 and AAS79453) [chalcomycin pathway], MycCI and MycG [mycinamicin pathway] (accession numbers BAC57023 and BAA03672), AmphL and AmphN (accession numbers AAK73504 and AAK73509) [amphotericin pathway], NysL and NysN (accession numbers AAF71769 and AAF71771) [nystatin pathway], PimD and PimG (accession numbers CAC20932 and CAC20928) [pimaricin pathway], TylHI and TylI (accession numbers AAD41818 and AAA21341) [tylosin pathway], EryF and EryK (accession numbers AAA26496 and YP-001102980) [erythromycin pathway], and PikC (accession numbers AAC68886) [pikromycin pathway]. The numbers in parentheses indicate the macrolactone ring size of the corresponding P450 substrate. Unless otherwise specified, all selected P450 enzymes are hydroxylases. The epoxidases are noted in italics. The enzymes marked with an asterisk are presumed to mediate carboxylic acid formation. MycCI and MycG are highlighted in bold.

to being derived from divergent evolution following duplication of a parental gene. Specifically, MycG, as the first biosynthetic monooxygenase characterized with an ability to catalyze both hydroxylation and epoxidation steps, is clustered with two other epoxidases — OleP (18) and ChmPI (19), whereas MycCI is closely related to ChmHI (19) and TylHI (16), both of which are responsible for methyl group hydroxylations of 16-membered ring macrolides. Interestingly, we revealed that the clustering of macrolide biosynthetic P450s correlated with not only the functionalities (i.e., hydroxylases, epoxidases, or ones responsible for carboxylic acid formation) but also the substrate macrolactone ring size, suggesting that selection based on product structure derived from the upstream PKS biosynthetic system might be among the significant factors directing evolution of secondary metabolic P450s.

6.2.2 Heterologous Expression of MycCI, MycCII, and MycG

The *mycCI*, *mycCII*, and *mycG* genes were overexpressed in *E. coli* BL21(DE3), and the resulting proteins were purified (see Supplementary Information). MycCI was overexpressed as either a C- or N-terminal 6× His tagged protein. After purification using Ni-NTA agarose chromatography, the individual polypeptides showed molecular weights of approximately 44 and 45 kDa, corresponding to the estimated masses of C- and N-terminal 6× His tagged MycCI. The wild-type MycCII (MycCII-wt, m.w. ~ 8 kDa) and N-terminal 6× His-tagged ferredoxin MycCII (MycCII-NH, m.w. not, ~ 10 kDa) were purified to homogeneity (see Experimental Procedures). MycG expression was placed under the control of a T7 promoter including a 6× His-tag introduced at its amino terminus. Similarly, one-step purification was performed using a Ni-NTA agarose column to obtain homogeneous protein with expected molecular weight of 46 kDa. Subsequently, the CO-bound reduced difference spectra confirmed the identity of both MycCI and MycG as cytochrome P450 enzymes (see Supplementary Information).

4.2.3 Synthesis of Mycinamicin VIII (M-VIII) from Protomycinolide IV (PML-IV)

Because of the limited quantity (< 0.1 mg) of totally available M-VIII isolated from *M. griseorubida* fermentation culture, the presumed substrate of MycCI, we developed an effective chemical glycosylation strategy to obtain this important intermediate by coupling desosamine as sugar donor to the readily available PML-IV aglycone precursor (20). Glycosylation with desosamine has previously employed either thioglycosides or glycosyl fluorides, with a methyl carbonate as the C2 protecting group

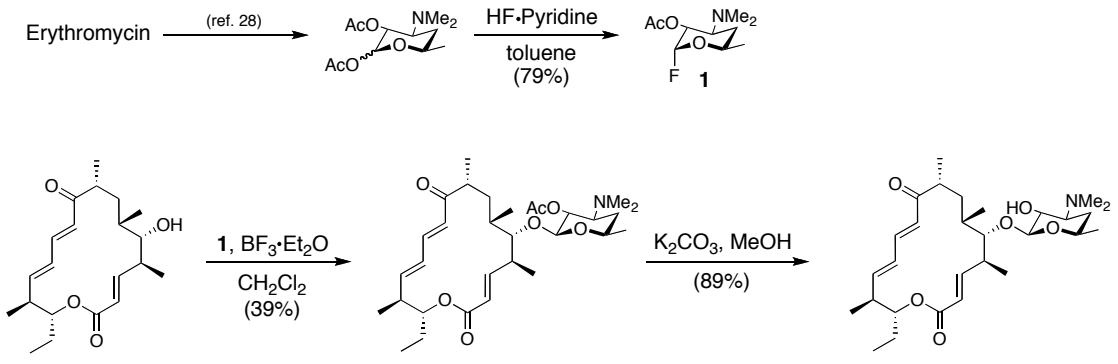


Figure 6-3. Synthetic scheme for M-VIII.

(21-27). After examining several combinations of protecting groups and anomeric leaving groups, we found that the C2 acetoxy-protected glycosyl fluoride of desosamine was conveniently obtained, easily purified, stable to storage, and effective in the transfer of desosamine to a range of aglycones. In this procedure (Fig. 6-3), erythromycin hydrate was hydrolyzed under acidic conditions to obtain a crude sample of desosamine hydrochloride that was converted to the bis(acetate) of desosamine in 88% isolated yield, as reported elsewhere (28). Further treatment of this material with HF·pyridine afforded the C2 acetoxy glycosyl fluoride of desosamine 1, which serves as a stable and convenient precursor for chemical glycosylations. Direct treatment of PML-IV (Fig. 6-1) with the glycosyl fluoride and $BF_3{\cdot}Et_2O$ in CH_2Cl_2 afforded C2 acetoxy mycinamycin VIII as a 2:1 mixture of β and α anomers in 39% (not optimized) isolated yield. Higher yielding desosamine glycosylations with other mycinamicins have been reported (22); however, these alternative procedures with different glycosyl donors were not explored for comparison. Purification by preparative HPLC (SiO_2) afforded the pure β-anomer (Fig. 6-3), which was deprotected with K_2CO_3 in methanol to afford semisynthetic M-VIII in 89% isolated yield. The ^{1}H NMR spectrum of this synthetic material was coincident with previously reported data (29).

6.2.4 Functional Analysis of MycCI *In Vitro*

To establish the role of each mycinamicin P450, we first tested whether MycCI is capable of hydroxylating the predicted substrate M-VIII using a typical heterologous spinach ferredoxin and ferredoxin reductase system (10). As expected, M-VIII was converted to the corresponding C21 hydroxymethyl product M-VII (Fig. 6-4), albeit inefficiently. Notably, the C- and N-terminal His-tagged MycCI showed similar activities, indicating that the location of His-tag has no significant impact on MycCI activity. We selected the C-terminal His-tagged MycCI to perform biochemical studies and refer to this form as MycCI for convenience. The low conversion ratio of M-VIII (11.0 ± 0.8%) might be due to the higher activation energy required to functionalize the primary C-H bond. Alternatively, the fact that a specific ferredoxin gene *mycCII* is clustered with *mycCI* suggests that MycCI might possess higher ferredoxin specificity than MycG. Thus, the capability of MycCII (together with spinach ferredoxin reductase,

since the native ferredoxin reductase for MycCI remains unknown) to reconstitute the in vitro M-VIII hydroxylation activity of MycCI was assessed and compared to spinach ferredoxin. Thus, when partnered by MycCII-wt and MycCII-NH, MycCI converted 28.6 ± 2.0% and 37.4 ± 1.4% of M-VIII to M-VII, respectively, suggesting MycCII mediates the electron transfer from spinach ferredoxin reductase to MycCI more efficiently than does spinach ferredoxin (Fig. 6-4). However, it is currently unclear why the activity of the His-tagged MycCII is higher than that of MycCII-wt. Considering that spinach ferredoxin and ferredoxin reductase are natural partners, we reason that the higher activity of MycCII could derive from a more favorable interaction with the MycCI P450, compared with the spinach ferredoxin reductase. Therefore, the binding affinities of different ferredoxins toward MycCI were compared with one another. As expected, the most active form (MycCII-NH) exhibited the lowest dissociation constant (K_d) of 7.0 ± 0.1 μM. In contrast, the binding affinity of spinach ferredoxin to MycCI (K_d = 148.9 ±

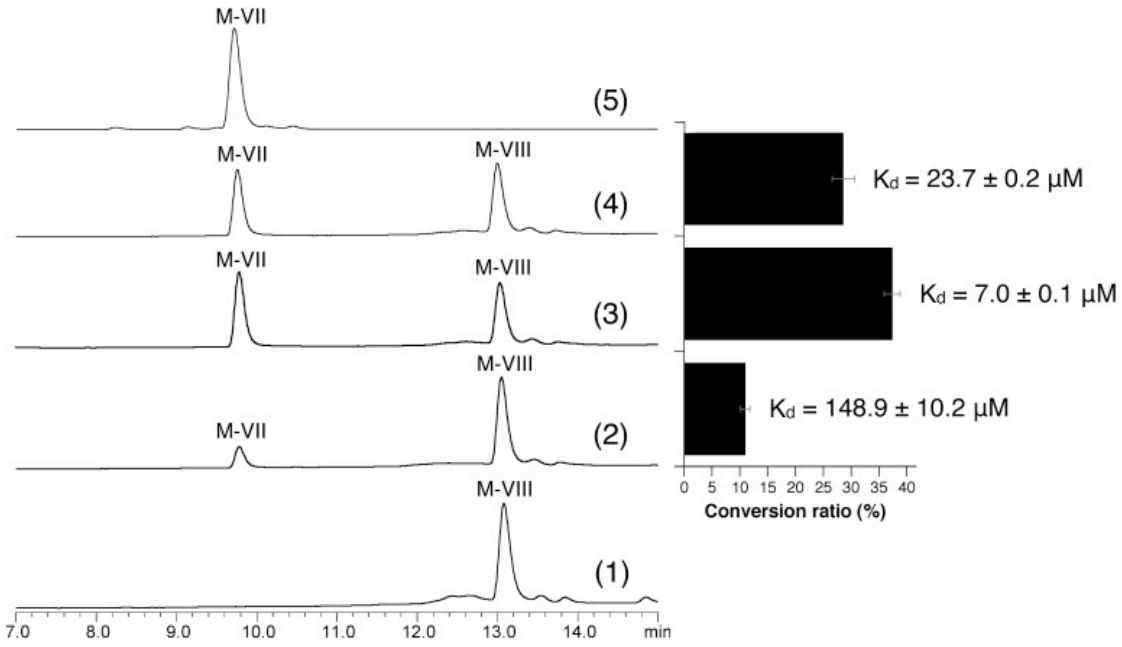


Figure 6-4. ***In vitro*** **M-VIII conversions catalyzed by MycCI.** All LC traces were analyzed at 280 nm: (1) negative control, M-VIII + boiled MycCI; (2) M-VIII + MycCI + spinach ferredoxin; (3) M-VIII + MycCI + MycCII-NH; (4) M-VIII + MycCI + MycCII-wt; and (5) M-VII standard. The right diagram shows M-VIII conversion ratios calculated by using Equation 1-$AUC_{unreacted\ M\text{-}VIII}/AUC_{total\ M\text{-}VIII}$ (AUC: area under curve) based on the corresponding LC-traces in parallel. Error bars are shown since the conversion ratios (mean ± s.d.) are calculated from two independent experiments. In addition, dissociation constants (K_d) with standard errors of the spinach and MycCII ferredoxins toward MycCI are shown.

10.2 μM) is much lower than that of either MycCII-wt or MycCII-NH. Interestingly, we found MycCII is unable to support *in vitro* activity of MycG (see Supplementary Information), the other P450 enzyme involved in mycinamicin biosynthesis. These results strongly suggest MycCII has evolved to selectively serve MycCI catalysis.

In addition, similar MycCI reactions were assessed with the PML-IV, M-VI, M-III, M-IV, and M-V (Fig. 6-1), but no products were observed, and starting materials were quantitatively recovered. These results establish that MycCI catalyzes the first oxidation step, converting the C21 methyl group to the corresponding hydroxymethyl that is subsequently functionalized with 6-deoxyallose, followed by two methylation steps to mycinose. Moreover, the substrate requires desosamine modification at the C5 hydroxyl group in order to be accepted by MycCI. This is analogous to PikC, shown through cocrystal structure analysis with natural 12- and 14-membered ring macrolides to involve desosamine-mediated substrate anchoring (30).

6.2.5 Functional Analysis of MycG *In Vitro*

Previous genetic studies on the mycinamicin biosynthetic pathway identified a gene fragment that complemented a M-II nonproducing mutant of *M. griseorubida* (17). DNA sequence analysis revealed an open reading frame whose translated product showed high-level amino acid sequence similarity to cytochrome P450 enzymes. The mycinamicin biosynthetic intermediate isolated from the M-II nonproducing *M. griseorubida* strain lacked both C14 hydroxyl and C12-C13 epoxide functionalities, and the subcloned fragment bearing the putative P450 gene was able to restore production of M-II (17). These data provided strong but indirect evidence that the mycG gene product is capable of catalyzing both oxidation steps at adjacent positions on the macrolactone ring (Fig. 6-1).

To confirm the dual function of MycG, our *in vitro* analysis was initiated by testing directly the ability of MycG to convert the putative substrate M-IV. As shown in Fig. 6-5A, the LC trace of the reaction extract [2] at 280 nm showed two peaks **a** and **b**. The corresponding mass spectra indicate M-V ($[M+H]^+$ = 712.35) and M-II ($[M+H]^+$ = 728.35) co-eluted as peak **a**, and peak **b** includes M-IV ($[M+H]^+$ = 696.35) and M-I ($[M+H]^+$ = 712.35). This LC-MC analysis clearly demonstrates the dual function of

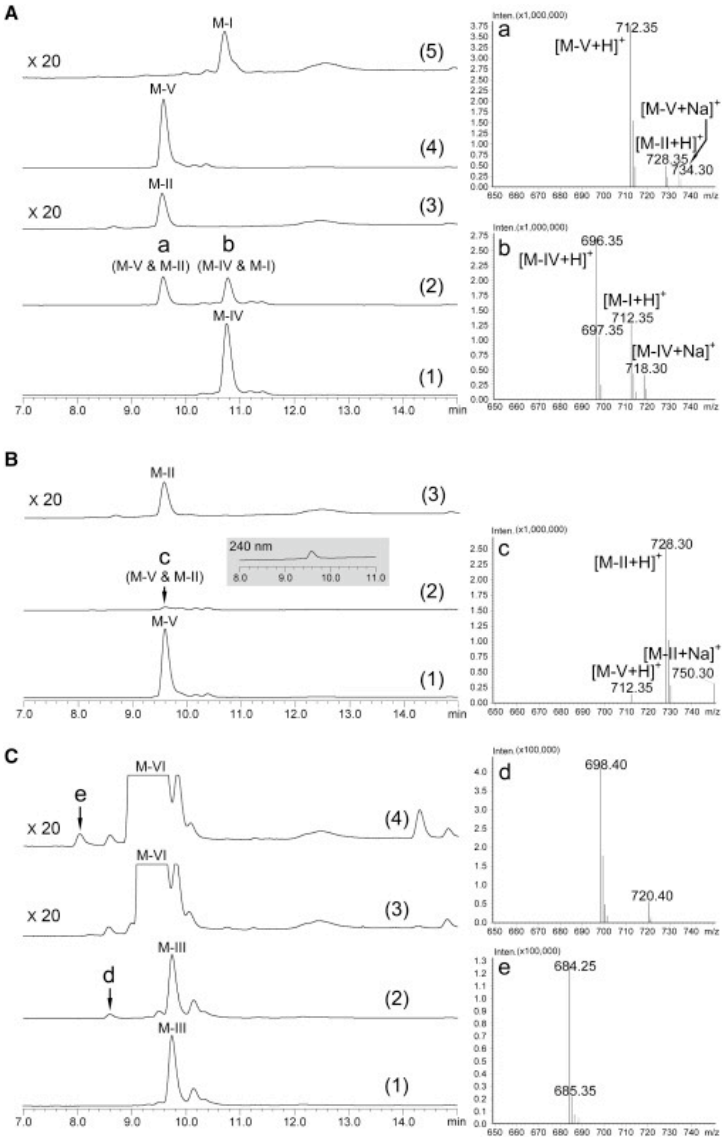


Figure 6-5. LC-MS analysis of *in vitro* conversions catalyzed by MycG. All LC traces were analyzed at 280 nm unless otherwise specified. (A) MycG reaction using M-IV as substrate: (1) negative control, M-IV + boiled MycG; (2) M-IV + MycG. The right panel shows the mass spectra of peak **a** and **b**; (3) M-II standard amplified 20× because of its poor absorbance at 280 nm; (4) M-V standard; (5) M-I standard amplified 20×. (B) MycG reaction using M-V as substrate: (1) negative control, M-V + boiled MycG; (2) M-V + MycG. The inset shows the 240 nm LC trace, under which M-II has better absorbance. The right panel shows the mass spectrum of peak c; (3) M-II standard amplified 20×. (C) MycG reactions using M-III and M-VI as alternative substrates: (1) negative control, M-III + boiled MycG; (2) M-III + MycG; (3) negative control, M-VI + boiled MycG; (4) M-VI + MycG. LC traces (3) and (4) amplified 20 × compared to (1) and (2) to visualize the peak **e** in trace amount. The right panel shows the mass spectra of peak **d** and **e**.

MycG as hydroxylase and epoxidase. However, despite efforts to optimize reverse-phase HPLC conditions, we were unable to separate M-IV and M-V from their epoxidized

products.

To determine the precursor of M-II, M-V and M-I were incubated with MycG in separate reactions. When M-V was used as substrate, it was almost completely converted into M-II. Since M-II (and M-I) lacks strong absorbance at 280 nm (11), only a minor peak c was observed in the LC trace, although substantial M-II was formed on the basis of mass spectral analysis (Fig. 6-5B; notably, the M-II product was visible at 240 nm). However, M-I was not converted into M-II *in vitro* (see Supplementary Information), indicating that M-II can only be generated from M-V as substrate. Epoxidation of M-IV prior to hydroxylation evidently results in termination of this post-PKS tailoring pathway (Fig. 6-1).

To assess further enzyme selectivity and ability to catalyze alternative oxidative reactions against the mycinamicin aglycone and several early glycosylated intermediates (PML-4, M-VIII, M-VII, M-VI, and M-III), MycG-mediated reactions were performed. No products were observed when using PML-4, M-VIII, and M-VII (see Supplementary Information), but two new compounds appeared at low levels in LC-MS analysis of the M-VI and M-III reactions (Fig. 6-5C). When M-III was used as a substrate, a minor new peak **d** was detected with m/z = 698.40, which is 16 Da higher than the molecular weight of M-III. Moreover, in the MycG reaction with M-VI, a new peak **e** (m/z = 684.25), presumed to be oxidized M-VI, was detected at an even lower conversion level. Both peak **d** and **e** were visible under 280 nm, and their retention times were faster than corresponding starting materials, strongly suggesting they are hydroxylated instead of epoxidized products. Moreover, the MS-MS analysis of d and e indicated that both hydroxylations occurred on the macrolactone ring, although we could not further determine the site of modification because of limited amounts of each compound. Interestingly, the C14 hydroxylated forms of M-III and M-VI have already been reported as minor components of the mycinamicin pathway designated as mycinamicin IX (M-IX) and mycinamicin XV (M-XV), respectively (31). Accordingly, it is likely that **d** and **e** correspond to M-IX and M-XV, providing direct evidence for the in vivo origin of these two minor compounds. More importantly, we established that efficient catalysis by MycG requires mycinose, derived from dimethylation of the second sugar residue as C21 OH linked 6-deoxyallose.

6.2.6 Measurement of Substrate Dissociation Constants

To understand why mycinamicin P450 enzymes (especially MycG) behave differently toward compounds with subtle structural differences, we performed spectrophotometric substrate binding assays to determine dissociation constants (K_d). As shown in Table 6-1, M-VIII binds to MycCI with a K_d value of 28.1 ± 3.2 μM, which is significantly higher than 0.7 ± 0.1 μM, the K_d value of M-IV toward MycG. Evidently, the substrate binding to MycG is much tighter than that to MycCI, which is likely attributed to additional interactions of mycinose within the P450 substrate-binding pocket. Interestingly, the C14 hydroxylated M-V showed approximately 14 times lower binding affinity toward MycG than M-IV, suggesting that M-V is a less suitable substrate for MycG. However, this finding is inconsistent with the results from in vitro assays (described in Fig. 6-4C above) in which M-V was converted to oxidized product more effectively than M-IV (kinetic analysis [see below] provided insights into this apparent paradox). In contrast to M-V, the epoxidized product M-I was not able to function as a substrate since we did not observe a Type-I difference spectra even at high substrate concentrations. This result is consistent with a lack of MycG activity toward M-I *in vitro*. It was particularly surprising that the final product M-II bound to MycG with a dissociation constant of 71.5 ± 8.4 μM, indicating the compromising effect of the epoxide moiety toward substrate binding could be overridden by an adjacent C14 hydroxyl group. The reason for this unusual finding will await substrate cocrystallographic analysis of MycG, which is currently in progress.

Table 6-1. Binding and steady-state kinetic analysis of MycCI and MycG

Enzyme	Substrate	K_d (μM)	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($\mu M^{-1}\bullet min^{-1}$)
MycCI	M-VIII	28.1 ± 3.2	34.5 ± 5.5[a]	71.7 ± 3.2[a]	2.1[a]
			5.8 ± 0.7[b]	104.1 ± 1.8[b]	17.9[b]
MycG	M-IV	0.7 ± 0.1	ND[c]	ND[c]	ND[c]
	M-V	10.4 ± 0.2	16.2 ± 3.1[a]	415.7 ± 22.9[a]	25.7[a]
	M-I	NB[d]	-	-	-
	M-II	71.5 ± 8.4	-	-	-

-. No reaction.
[a]Kinetic parameters were determined when spinach ferredoxin was used.
[b]Kinetic constants were determined when MycCII0NH ferredoxin was used.
[c]ND: The kinetic parameters were not determined because multiple reactions occur at the same time.
[d]NB: No binding.

6.2.7 Steady-State Kinetic Analysis of MycCI and MycG

Assuming a 1:1 stoichiometric relationship between NADPH consumption and substrate oxidation, the steady-state kinetic parameters of MycCI were determined using purified synthetic M-VIII as substrate. When partnered with spinach ferredoxin, MycCI demonstrated a K_m of 34.5 ± 5.5 μM and a k_{cat} of 71.7 ± 3.2 min^{-1} for the C21 methyl group hydroxylation step. In contrast, with MycCII-NH coupled with MycCI, catalysis became more efficient with a decreased K_m of 5.8 ± 0.7 μM and an improved k_{cat} of 104.1 ± 1.8 min^{-1}. This finding suggests that MycCII not only induces better substrate binding of MycCI but also improves electron transfer efficiency as reflected by increased turnover number (k_{cat}).

Although MycG is capable of catalyzing three different reactions (M-IV to M-V; M-IV to M-I; and M-V to M-II), we limited our study to kinetic constants for M-V to M-II conversion (K_m = 16.2 ± 3.1 μM; k_{cat} = 415.7 ± 22.9 min^{-1}). This was because of the complexity of the analysis when M-IV was used as substrate, since multiple reactions (M-IV to M-V or M-I and M-V to M-II) did not allow fitting of data to the Michaelis-Menten equation. However, it is noteworthy that we observed the overall NADPH consumption rate at high M-IV concentration (e.g., 100 μM) to be considerably reduced, compared with low M-IV concentration (e.g., 10 μM), suggesting strong substrate inhibition. This could explain the above-mentioned contradiction between the conversion ratio and binding affinity. According to the *in vitro* analysis of MycG activity (Fig. 6-5A and 6-5B), M-V appears to be a better substrate than M-IV on the basis of the conversion ratios. However, this result might be misleading since M-IV has better binding affinity toward MycG (Table 6-1). Thus, on the basis of in vitro reactions using substrate concentrations at 500 μM, we believe that the lower conversion of M-IV is a reflection of substrate inhibition.

6.3 Discussion

Mycinamicins represent a large family of macrolide antibiotics with more than 20 members (31, 32). The structural diversity is derived primarily from post-PKS tailoring modifications, including glycosylation, oxidation, and methylation steps. Therefore, this

biosynthetic pathway represents an important system to explore the mechanism and significance of secondary metabolite diversification.

Here, we provide a detailed analysis of two cytochrome P450 enzymes, MycCI and MycG, and their role in late-stage chemical modifications. By gene cloning, protein expression, purification, and reconstitution assays with heterologous redox partners, we unambiguously confirmed their physiological role in vitro. To the best of our knowledge, the functional and kinetic analysis of MycCI represents the first enzymatic characterization of a biosynthetic monooxygenase responsible for methyl group hydroxylation. In addition, we established that MycCII is a specific ferredoxin, whose corresponding gene (*mycCII*) is immediately adjacent to *mycCI* and capable of effectively supporting MycCI activity. Compared with spinach ferredoxin, MycCII binds to MycCI more tightly, presumably leading to more efficient electron transfer. On the other hand, we established that C21 hydroxylation by MycCI depends on C5-linked desosamine, which is analogous to PikC (10), another cytochrome P450 involved in pikromycin biosynthesis. Recent cocrystallographic analysis of PikC (30) and its natural substrates (YC-17 and narbomycin) revealed that desosamine acts as an indispensable anchor responsible for productive binding and proper positioning of substrate in the active site. Accordingly, we surmise that desosamine plays a similar role in hydroxylation of precursor macrolide M-VIII.

The current study has revealed that MycG is a more versatile enzyme than MycCI, providing the first example of a biosynthetic P450 involved physiologically in catalyzing both hydroxylation and epoxidation reactions. In addition, MycG represents the primary basis for structural diversification in the mycinamicin pathway since it is solely responsible for generating multiple products, including M-V, M-I, and M-II, as well as some minor components, including M-IX and M-XV. This work has also demonstrated that MycG function depends on the presence of both desosamine and the second sugar mycinose in the substrate. It is also noteworthy that both deoxysugars are essential for mycinamicin bioactivity (29). Although the precise role of the second sugar for MycG substrate recognition remains obscure because of the lack of X-ray structural information, on the basis of the high binding affinity of M-IV toward MycG (K_d = 0.7 ± 0.1 μM), we predict that there might be a specific active site-binding pocket in the

polypeptide to accommodate mycinose, thereby significantly improving substrate affinity. Moreover, of particular interest, the activity of MycG against diglycosylated substrates (M-VI, M-III, and M-IV) is related to the extent of methylation mediated by sequential *O*-methyltransfer by MycE and MycF (15). In the absence of methylether groups installed on the second sugar molecule, M-VI bearing 6-deoxyallose appears to be a very poor substrate for MycG. This situation can be improved by the first methylation to form javose in M-III, although its conversion is low (Fig. 6-5C). Subsequently, the second methylation catalyzed by MycF imparts a pronounced effect upon MycG-mediated binding of M-IV. It appears likely that these stepwise methylations decrease the polarity of 6-deoxyallose in M-VI, thus enabling it to be accepted by a putative hydrophobic MycG-binding pocket.

Substrate inhibition occurs in about 20% of all known enzymes (33). In biosynthetic P450 enzymes, this phenomenon has been observed in previous kinetic analysis of PikC (34), EryK (35), and PimD (36). In this study, M-IV inhibition of MycG was also detected indirectly even though the kinetic parameters were not obtained because of the complexity of the multiple-reaction system. Thus, the mechanism of this substrate inhibition cannot be further elucidated without generating detailed kinetic curves. However, we realize that this behavior of MycG might be physiologically significant for maintaining the chemical diversity within the mycinamicin pathway. In the event that M-IV inhibition does not occur, MycG would convert all M-IV into final products M-I and M-II, instead of a mixture of M-I, II, IV, and V as shown in Fig. 6-5A and observed in the fermentation culture of *M. griseorubida* strain A11725 (11). This would ultimately limit the spectrum of its metabolic output and perhaps compromise the ability of the microorganism to adapt to a variable and competitive environment.

6.4 Experimental Procedures

Materials — Unless otherwise specified, all chemical reagents were ordered from Sigma-Aldrich. DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). Molecular cloning used New England Biolabs (Ipswich, MA) restriction enzymes, Stratagene (La Jolla, CA) Pfu Turbo DNA polymerase, Novagen (Madison, WI) pET vectors, Invitrogen (Carlsbad, CA) T4 DNA ligase, and Z-

Competent Escherichia coli Transformation Buffer set from Zymo Research (Orange, CA). Protein purification used QIAGEN (Valencia, CA) Ni-NTA resin, Millipore (Billerica, MA) Amicon Ultra centrifugal filter, PD-10 desalting columns from GE Healthcare (Piscataway, NJ), and Thrombin restriction grade Kit from Novagen (Madison, WI). LB Broth was from EM Sciences (Gibbstown, NJ). All mycinamicins except M-VIII were obtained from the fermentation broth of M. griseorubida A11725. M-VIII was synthesized from PML-IV using the method described above.

DNA Manipulation, Cloning, and PCR — Using cosmid pMR01 (15) as template, *mycCI*, *mycCII*, and *mycG* genes were amplified by PCR under standard conditions with primers as follows: forward, 5′-CAGCATATGGTGGTCTGGCCCATGGACCGCACCTG-3′ for *mycCI*, 5′-GTGCCATATGCGGATAGTCCTGGACGCCGAAC-3′ for *mycCII*, and 5′-CGGTCATATGACTTCAGCTGAACCTAGGGCGTATCC-3′ for *mycG* (the underlined bases represent the introduced NdeI site for further cloning); reverse, 5′-TCGT*AAGCTT*CCACTCGACCAGCAGCTCGTCGATG-3′ for cloning *mycCI* gene without a stop codon, 5′-TCCA*AAGCTT*TCCGCATACCGCACCCCCATTCGTC-3′ for amplifying the *mycCI* gene retaining a stop codon, 5′-TGAC*AAGCTT*ACTCCTGTTGGCCCACCTGTCCCGTG-3′ for *mycCII*, and 5′-GGCA*AAGCTT*CTCCGACGAGATCGTCGAGATCGAC-3′ for *mycG* (the italic letters indicate a *Hind*III restriction site for later cloning). The gel purified cDNAs were rescued by double digestion of *Nde*I and *Hind*III. Then, the fragment containing the *mycCI* gene with stop codon removed was ligated into *Nde*I/*Hind*III treated pET21b to generate recombinant plasmid pET21b-*mycCI* for expression of C-terminal His-tagged MycCI. Genes retaining stop codons, including *mycCI*, *mycCII*, and *mycG*, were ligated into previously *Nde*I/*Hind*III-digested pET28b to generate recombinant plasmids for expression of N-terminal His_6-tagged proteins. After confirming the identity of inserted genes by DNA sequencing, the constructs were used to transform *E. coli* BL21 (DE3) for protein overexpression.

Protein Overexpression and Purification — MycCI and MycG (P450s) overexpression and purification followed previously developed procedures with minor modifications (10, 34). The *E. coli* BL21 (DE3) transformants carrying certain plasmids were grown at 37°C in 1 liter of LB broth containing thiamine (1 mM), 5% glycerol, 50 μg/ml selective antibiotics (ampicillin for pET21b and kanamycin for pET28b), and a rare salt solution (6750 μg/l $FeCl_3$, 500 μg/l $ZnCl_2$, $CoCl_2$, Na_2MoO_4, 250 μg/l $CaCl_2$, 465 μg/l $CuSO_4$, and 125 μg/l H_3BO_3) until OD_{600} reached 0.6–0.8. Then isopropyl β-D-thiogalactoside (IPTG, 0.1mM) and δ-aminolevulinic acid (1mM) were added, and the cells were cultured at 18°C overnight. After harvesting cells by centrifugation, 40 ml of lysis buffer (50 mM NaH_2PO_4 [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 (ThermoFisher Scientific). The insoluble material was separated by centrifugation (35,000 × g, 30 min at 4°C). The soluble fraction was collected and incubated for 1 h at 4°C after addition of 1ml Ni-NTA resin. The slurry was loaded onto an empty column, and the column was then washed with 40 to 80 ml of wash buffer (50 mM NaH_2PO_4 [pH8.0], 300 mM NaCl, 10% glycerol, and 20–30 mM imidazole). The elution buffer (50 mM NaH_2PO_4 [pH8.0], 300 mM NaCl, 10% glycerol, and 250 mM imidazole) was added onto the column, and eluted protein fraction was concentrated with Amicon Ultra 4, Ultracel −30K. Subsequent desalting was attained by buffer exchange into desalting buffer (50 mM NaH_2PO_4 [pH7.3], 1 mM EDTA, 0.2 mM dithioerythritol, and 10% glycerol) with a PD-10 column.

Overexpression and purification of N-terminal 6 × His-tagged ferredoxin MycCII-NH are similar to above procedures with minor modifications as follows: 1) δ-aminolevulinic acid, the precursor for heme biosynthesis, was omitted in the culture broth; and 2) the protein obtained from elution buffer was concentrated using Amicon Ultra 4, Ultracel −5K because of the low molecular weight of MycCII (~ 10 kDa). To obtain MycCII-wt lacking a His-tag, 2.4 mg of MycCII-NH in desalting buffer was digested by 2 units of thrombin at 4°C overnight. The cleaved His-tag and residual thrombin were removed by use of Ni-NTA resin and a 30K size exclusion filter sequentially. The concentration of purified ferredoxins was determined by Coomassie protein assay using spinach ferredoxin as standard.

CO-Bound Reduced Difference Spectra — The identification of MycCI and MycG as active P450 enzymes was performed through getting the CO-bound reduced difference spectra using a UV-visible spectrophotometer 300 Bio (Cary). First, the P450 enzyme in desalting buffer was reduced by adding several milligrams of sodium dithionite ($Na_2S_2O_4$), and a spectrum was recorded from 350 to 600 nm. Then, after CO bubbling of the solution for 30–60 sec, the spectrum of CO-bound reduced P450 species was recorded using previous reduced spectrum as reference. This assay was also used to determine the functional P450 concentration using the extinction coefficient of 91,000 $M^{-1} \cdot cm^{-1}$ (36).

Functional Analysis of *In Vitro* Activities of MycCI and MycG — The standard conversion was accomplished by combining 1 μM of desalted MycCI or MycG, whose functional concentrations were determined using UV-visible absorption spectrum method (35), 0.5 mM mycinamicin biosynthetic intermediate, 3.5 μM spinach ferredoxin or MycCII ferredoxin, 0.1 U/ml spinach ferredoxin-$NADP^+$ reductase, and 0.5 mM NADPH in 100 μl of desalting buffer. The reaction with boiled P450 enzyme was used as a negative control. The reaction was performed for 2 h at 30°C and was terminated by extraction, using 3 × 200 μl of $CHCl_3$. The resulting organic extraction was dried and redissolved in 150 μl of methanol. The LC-MS analysis of reaction extract was performed on LCMS-2010 EV (Shimadzu) by using an XBridge™ C18 3.5 μm 150 mm reverse-phase HPLC column under following conditions: mobile phase, 20–100% solvent B over 18 min (A = deionized water + 0.1% formic acid, B = acetonitrile + 0.1% formic acid); flow rate: 0.2 ml/min; UV wavelength: 240 and 280 nm.

Spectral Substrate Binding Assay — Spectral substrate binding assay was performed on UV-visible spectrophotometer 300 Bio (Cary) at room temperature by titrating substrate DMSO solution (blank DMSO for reference group) into 1 ml of 0.5–1 μM P450 solution in 1 μl aliquots (34). The series of Type I difference spectra were used to deduce ΔA ($A_{peak} - A_{trough}$). Then, the data from duplicated experiments were fit to Michaelis-Menten equation to obtain the dissociation constant K_d.

Spectral Ferredoxin Binding Assay — Spectral ferredoxin titrations were performed as described elsewhere (37) using 0.7 μM MycCI and 10 μM M-VIII in desalting buffer at room temperature. Ferredoxin solutions in appropriate concentrations were used to titrate the P450 solution. Binding of ferredoxins to MycCI induced increased M-VIII binding, thus leading to a larger absorbance difference ΔA ($A_{peak} - A_{trough}$). The data from duplicated experiments were linearized by using a Hanes-Wolf plot to deduce dissociation constant K_d.

Steady-State Kinetics of MycCI and MycG — The standard reactions contain 0.6 μM of MycCI (when partnered with spinach ferredoxin), 0.1 μM of MycCI (when partnered with MycCII-NH), or 0.1 μM of MycG, 1.9 μM spinach or MycCII ferredoxins, 0.02 U/ml spinach ferredoxin-$NADP^+$ reductase, 5–200 μM M-VIII for MycCI and 2–160 μM M-V for MycG in 90 μl of desalting buffer. After preincubation in 96-well plate at room temperature for 5 min, the reactions with different substrate concentrations were initiated spontaneously by adding 10 μl of 2 mM NADPH with a multichannel pipette. The rate of NADPH consumption was monitored continuously over 2 min under 340 nm by SpectraMax M5 spectrophotometer (Molecular Devices). The initial velocities of NADPH consumption were deduced from the absorbance curves within the linear range (0–20 s). Then, assuming a 1:1 stoichiometric relationship between NADPH consumption and substrate oxidation, the initial velocities of hydroxylation reactions were calculated by using the millimolar absorption coefficient 6.22 $mM^{-1} \cdot cm^{-1}$ of NADPH at 340 nm (38). Finally, the results from duplicated experiments were fit to Michaelis-Menten equation to obtain steady-state kinetic parameters.

6.5 Supplementary Information

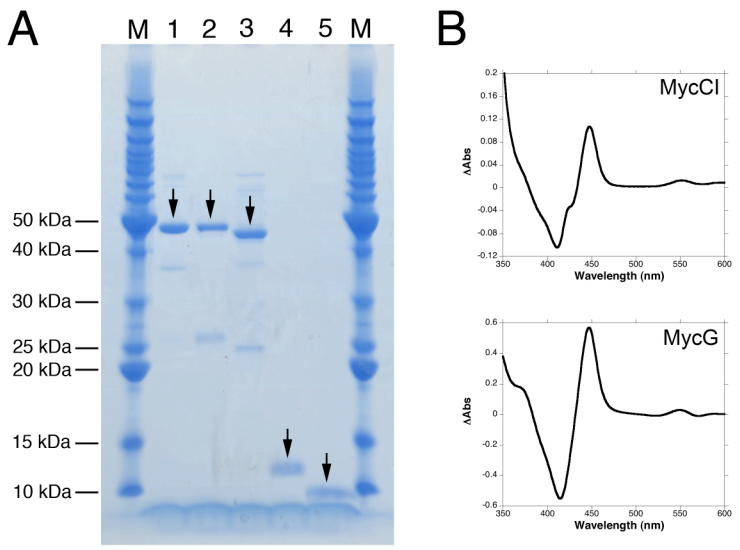


Figure 6-S1. SDS-PAGE analsis and CO-bound reduced difference spectra of MycCI and MycG. (A) Coomassie blue stained SDS-PAGE gel results of purified MycCI, MycG and MycCII. Lanes: M, protein marker; 1, N-terminal 6 × His tagged MycG; 2, N-terminal 6 × His tagged MycCI; 3, C-terminal 6 × His tagged MycCI; 4, N-terminal 6 × His tagged MycCII; 5, wild type MycCII with His-tag removed. (B) CO-bound reduced difference spectra of MycCI and MycG.

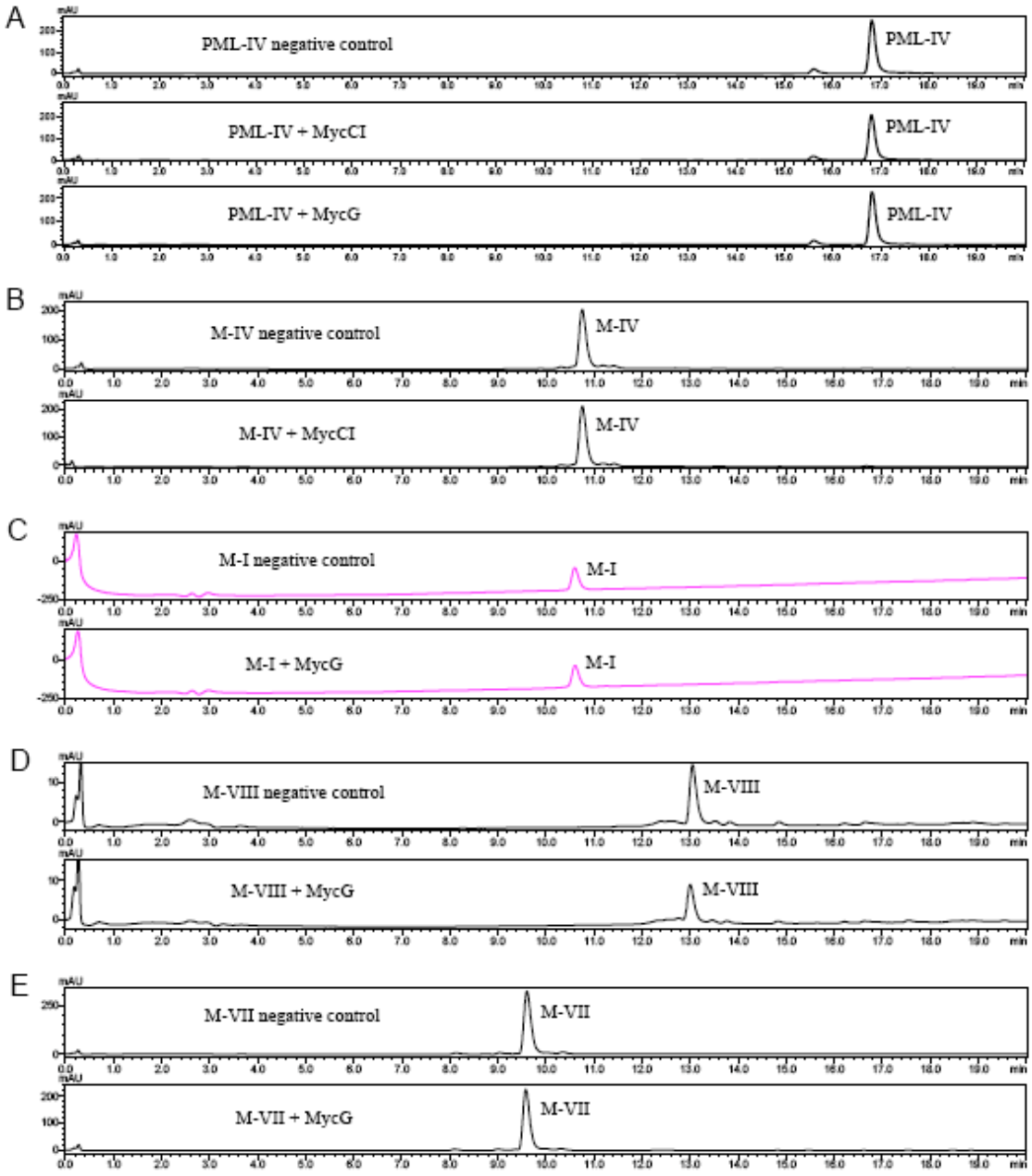


Figure 6-S2. The activities of MycCI and MycG toward various mycinamicin intermediates (all LC traces were recorded under 280 nm except for M-I group at 240 nm). (A) PML-IV reactions; (B) M-IV reactions; (C) M-I reactions; (D) M-VIII reactions; (E) M-VII reactions.

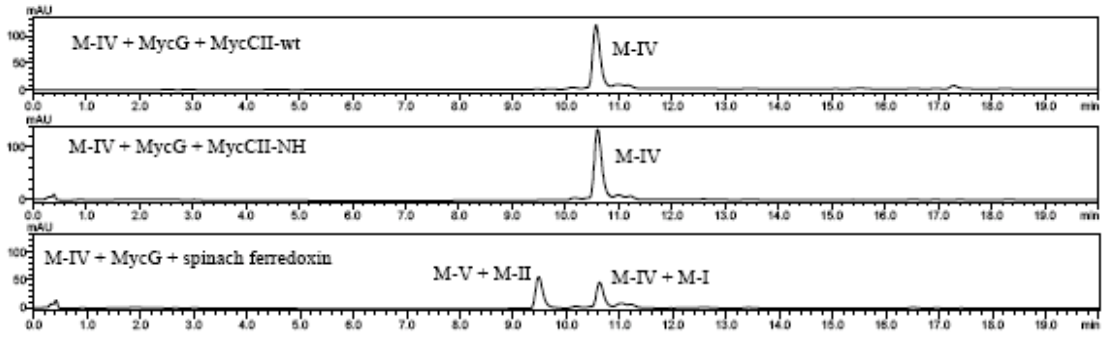


Figure 6-S3. The activities of MycG toward M-IV when partnered by different ferredoxins.

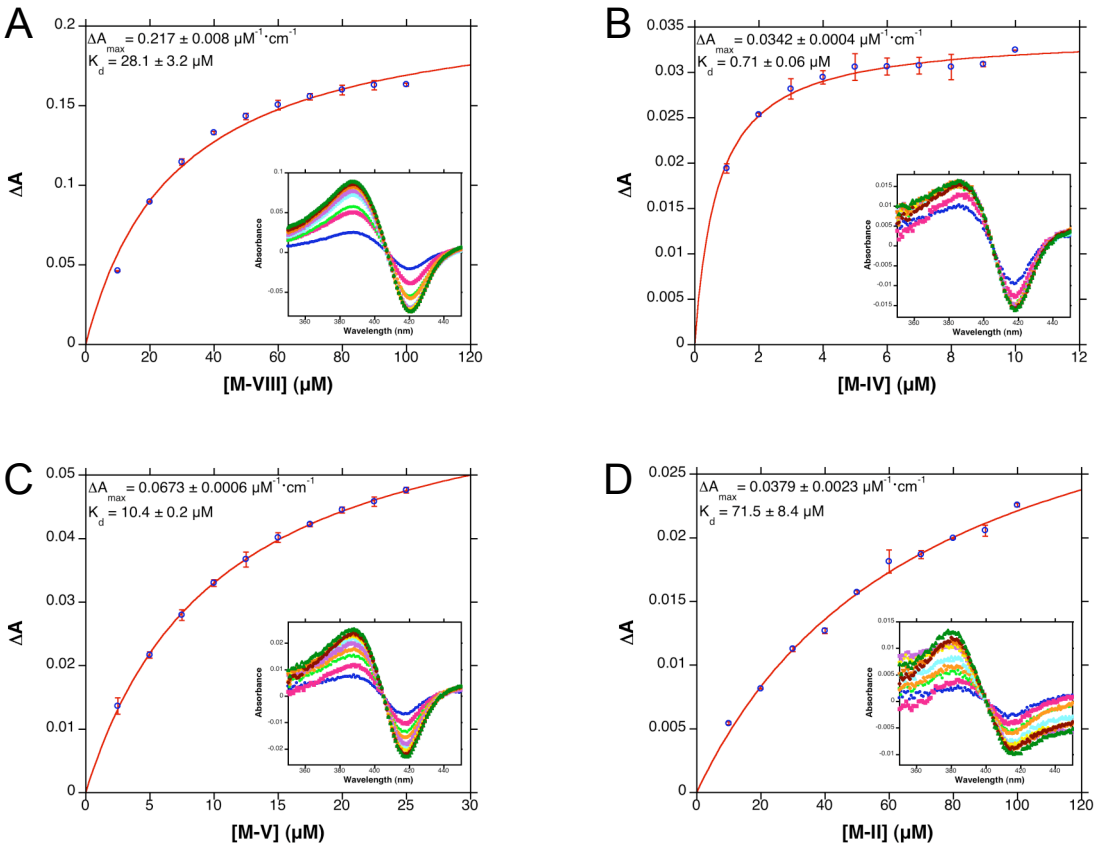


Figure 6-S4. Binding analysis of MycG and MycCI. The substrate dissociation constants (K_d) measurement of (A) M-VIII toward MycCI and (B) M-IV, (C) M-V, and (D) M-II toward MycG. The insets show the Type I difference spectra. All experiments were performed in duplicate.

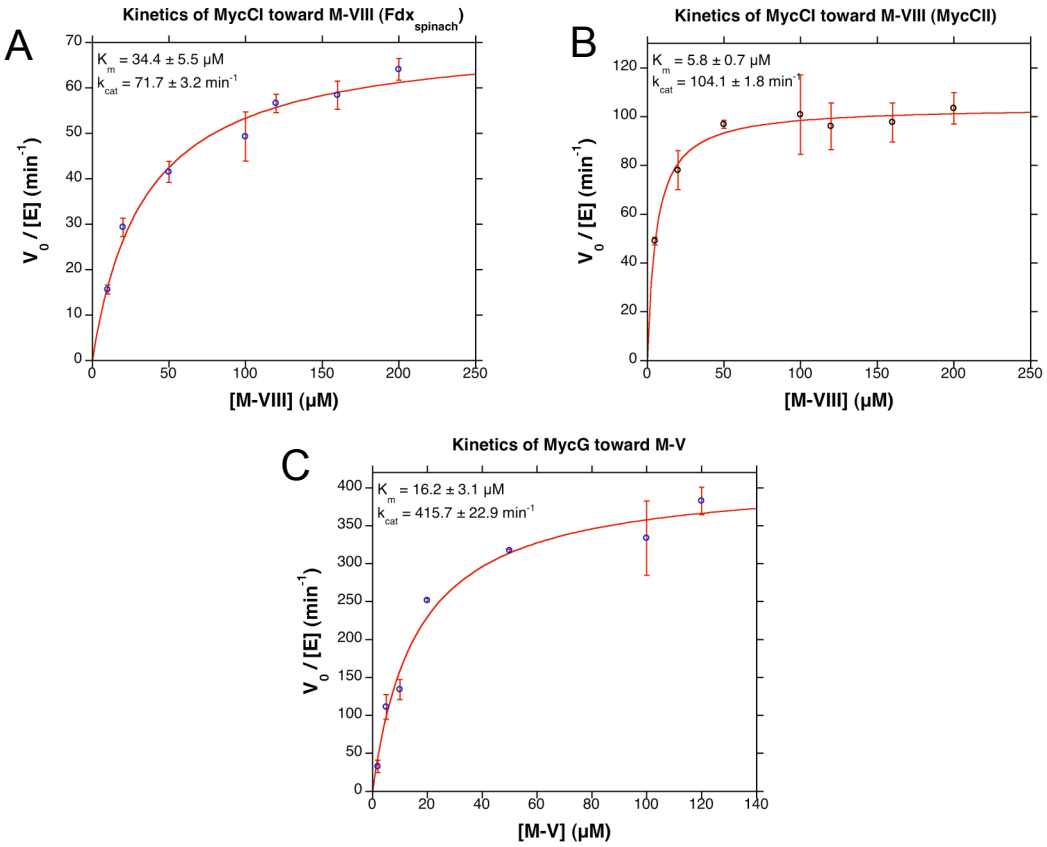


Figure 6-S5. Kinetic analysis of MycG and MycCI. Michaelis-Menten curves of (A) MycCI using M-VIII as substrate (partnered by spinach ferredoxin and ferredoxin reductase), (B) MycCI using M-VIII as substrate (partnered by MycCII-NH and spinach ferredoxin reductase) and (C) MycG using M-V as substrate (partnered by spinach ferredoxin and ferredoxin reductase). All experiments were performed in duplicate.

6.6 References

1. Coon, M. J. (2005) *Annu. Rev. Pharmacol. Toxicol.* **45**, 1–25
2. Guengerich, F. P. (2001) *Chem. Res. Toxicol.* **14**, 611–650
3. Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., and Harper D. *et al.* (2002) *Nature* **417**, 141–147
4. Ikeda, H., Ishikawa, J., Hanamoto, A., Shinose, M., Kikuchi, H., Shiba, T., Sakaki, Y., Hattori, M., and Omura, S. (2003) *Nat. Biotechnol.* **21**, 526–531
5. Oliynyk, M., Samborskyy, M., Lester, J. B., Mironenko, T., Scott, N., Dickens, S., Haydock S. F., and Leadlay, P. F. (2007) *Nat. Biotechnol.* **25**, 447–453
6. Lamb, D. C., Ikeda, H., Nelson, D. R., Ishikawa, J., Skaug, T., Jackson, C., Omura, S., Waterman, M. R., and Kelly, S. L. (2003) *Biochem. Biophys. Res. Commun.* **307**, 610–619
7. Rix, U., Fischer, C., Remsing, L. L., and Rohr, J. (2002) *Nat. Prod. Rep.* **19** 542–580
8. Andersen, J. F., and Hutchinson, C.R. (1992) *J. Bacteriol.* **174**, 725–735
9. Stassi, D., Donadio, S., Staver, M.J., and Katz, L. (1993) *J. Bacteriol.* **175**, 182–189
10. Xue, Y., Wilson, D., Zhao, L., Liu, H.-w., and Sherman, D.H. (1998) *Chem. Biol.* **5**, 661–667
11. Satoi, S., Muto, N., Hayashi, M., Fujii, T., and Otani, M. (1980) *J. Antibiot.* **33**, 364–376
12. Kinoshita, K., Takenaka S., and Hayashi, M. (1988) *J. Chem. Soc. Chem. Commun.* **1988**, 943–945
13. Suzuki, H., Takenaka, S., Kinoshita K., and Morohoshi, T.(1990) *J. Antibiot.* **43**, 1508–1511
14. Kinoshita, K., Takenaka, S., and Hayashi, M. (1991) *J. Chem. Soc., Perkin Trans. 1* **1**, 2547–2553
15. Anzai, Y., Saito, N., Tanaka, M., Kinoshita, K., Koyama, Y., and Kato, F. (2003) *FEMS Microbiol. Lett.* **218**, 135–141
16. Baltz R. H., and Seno, E.T. (1981) *Antimicrob. Agents Chemother.* **20**, 214–225
17. Inouye, M., Takada, Y., Muto, N., Beppu, T., and Horinouchi, S. (1994) *Mol. Gen. Genet.* **245**, 456–464
18. Rodriguez, A. M., Olano, C., Mendez, C., Hutchinson C. R., and Salas, J. A. (1995) *FEMS Microbiol. Lett.* **127**, 117–120
19. Ward, S. L., Hu, Z., Schirmer, A., Reid, R., Revill, W. P., Reeves, C. D., Petrakovsky, O. V., Dong S. D., and Katz, L. (2004) *Antimicrob. Agents Chemother.* **48**, 4703–4712
20. Hayashi, M., Ohara, H., Ohno, M., Sakakibara H., and Satoi, S. (1981) *J. Antibiot.* **34**, 1075–1077
21. Martin, S. F., Hida, T., Kym, P. R., Loft, M., and Hodgson, A. (1997) *J. Am. Chem. Soc.* **119**, 3193–3194
22. Matsumoto, T., Maeta, H., Suzuki, K., and Tsuchihashi, G. (1988) *Tetrahedron Lett.* **29**, 3575–3578
23. Suzuki, K., Maeta, H., Matsumoto, T., and Tsuchihashi, G. (1988) *Tetrahedron Lett.* **29**, 3571–3574
24. Toshima, K., Nozaki, Y., Mukaiyama, S., Tamai, T., Nakata, M., Tatsuta, K., and Kinoshita, M. (1995) *J. Am. Chem. Soc.* **117**, 3717–3727
25. Woodward, R. B., Logusch, E., Nambiar, K. P., Sakan, K., Ward, D. E., Au-Yeung, B. W., Balaram, P., Browne, L. J., Card, P. J., and Chen, C. H. *et al*. (1981) *J. Am. Chem. Soc.* **103**, 3210–3213
26. Woodward, R. B., Logusch, E., Nambiar, K. P., Sakan, K., Ward, D. E., Au-Yeung, B. W., Balaram, P., Browne, L. J., Card, P. J., and Chen, C. H. *et al*. (1981) *J. Am. Chem. Soc.* **103**, 3213–3215

27. Woodward, R. B., Logusch, E., Nambiar, K. P., Sakan, K., Ward, D. E., Au-Yeung, B. W., Balaram, P., Browne, L. J., Card, P. J., and Chen, C. H. *et al*. (1981) *J. Am. Chem. Soc.* **103**, 3215–3217
28. Chen, H., Yamase, H., Murakami, K., Chang, C.-w., Zhao, L., Zhao, Z., and Liu, H.-w. (2002) *Biochemistry* **41**, 9165–9183
29. Kinoshita, K., Imura, Y., Takenaka, S., and Hayashi, M. (1989) *J. Antibiot.* **44**, 1869–1872
30. Sherman, D. H., Li, S., Yermalitskaya, L. V., Kim, Y., Smith, J. A., Waterman, M. R., and Podust, L. M. (2006) *J. Biol. Chem.* **281**, 26289–26297
31. Kinoshita, K., Takenaka, S., Suzuki, H., Morohoshi, T., and Hayashi, M. (1992) *J. Antibiot.* **45**, 1–9
32. Kinoshita, K., Takenaka, S., and Hayashi, M. (1991) *J. Antibiot.* **44**, 1270–1273
33. Copeland, R. A. (2000) Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis (Second Edition), Wiley, New York

34. Li, S., Podust, L. M., and Sherman, D. H. (2007) *J. Am. Chem. Soc.* **129**, 12940–12941
35. Lambalot, R. H., and Cane, D. E.(1995) *Biochemistry* **34**, 1858–1866
36. Omura T., and Sato, R. (1995) *J. Biol. Chem.* **239**, 2379–2385
37. Coghlan, V. M., and Vickery, L. E. (1991) *J. Biol. Chem.* **266**, 18606–18612
38. Greenbaum, E., Austin, R. H., Frauenfelder, H., and Gunsalus, I. C. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1273–1276

Notes:

This work has been published as "Functional analysis of MycCI and MycG, cytochrome P450 enzymes involved in biosynthesis of mycinamicin macrolide antibiotics." Anzai, Y., **Li, S.** (Co-first author), Chaulagain, M. R., Kinoshita, K., Kato, F., Montgomery, J., and Sherman, D. H. ***Chem. Biol.*** 2008, 15 (9), 950-959.

Author contributions:
Shengying Li, Yojiro Anzai, John Montgomery and David H. Sherman designed the experiments;
Shengying Li and Yojiro Anzai performed the biochemistry;
Mani Raj Chaulagain performed the organic synthesis;
Kenji Kinoshita and Fumio Kato provided mycinamicin materials.

Chapter 7
Functional Analysis of MycE and MycF, Two *O*-Methyltransferases Involved in the Biosynthesis of Mycinamicin Macrolide Antibiotics

7.1 Introduction

Deoxysugars are prevalent structural components of many antibiotics and often contribute substantially to their biological properties (1-3). Methylation of hydroxy group(s) on the deoxysugar ring is relatively common (Fig. 7-1), as *O*-methylation not only protects the reactive hydroxy group from undesired modifications, such as oxidation or dehydration, but also alters the solubility and pharmacokinetic properties of the resulting molecule (4). Biosynthetically, these *O*-methylation reactions are mainly catalyzed by a variety of *S*-adenosyl-$_L$-methionine (SAM or AdoMet)-dependent methyltransferases in a site-specific manner. For example, the two SAM-dependent *O*-methyltransferases TylE and TylF in the tylosin biosynthetic pathway of *Streptomyces fradiae* sequentially methylate individual hydroxy groups (C2-OH and C3-OH) in the 6-deoxyallose moiety of demethylmacrocin to generate the macrolide antibiotic tylosin (5-8). ElmMI, ElmMII, and ElmMIII are responsible for the consecutive methylation of three hydroxy groups of $_L$-rhamnose in the antitumor polyketide antibiotic elloramycin (9). Moreover, a growing number of *O*-methyltransferases involved in various deoxysugar methylation reactions such as EryG (10), OleY (11), SpinH, SpinI, and SpinK(12) have been reported in diverse antibiotic biosynthetic systems.

Mycinamicins represent a family of macrolide antibiotics with more than 20 members produced by the rare actinomycete *Micromonospora griseorubida* (13-15). The antibacterial activities of some mycinamicin products against *Staphylococcus aureus* are higher than those of the clinical macrolide antibiotics erythromycin and leucomycin.

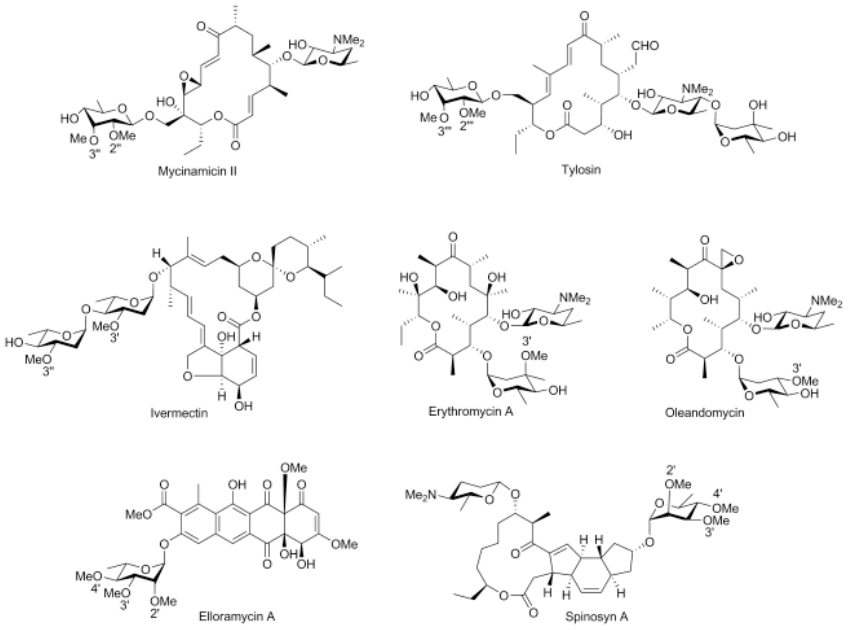


Figure 7-1. Antibiotics containing various *O*-methylated deoxysugars. The methyl groups installed by *O*-methyltransferases are numbered.

More importantly, mycinamicins have shown strong activity against a number of antibiotic-resistant human pathogens (13). Structurally, the major mycinamicin products of wild-type strain *M. griseorubida* A11725, including mycinamicins I, II, IV, and V, are composed of a 16-membered ring macrolactone core, an *N*,*N*-dimethylated deoxysugar desosamine, and a di-*O*-methylated deoxyhexose mycinose (Fig. 7-1). During the past two decades, the biosynthesis of mycinamicin has been elucidated through strain mutagenesis, bioconversion studies (16, 17), and sequence analysis of the complete mycinamicin gene cluster (18), wherein two putative *O*-methyltransferase genes *mycE* and *mycF* were tentatively assigned.

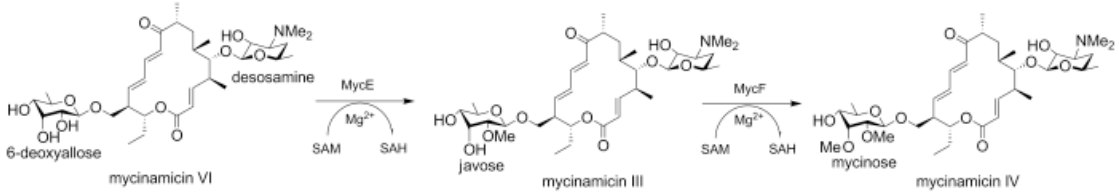


Figure 7-2. Physiological reactions catalyzed by MycE and MycF.

Initial bioinformatics analysis of the corresponding genes revealed that MycE and MycF show high amino acid sequence similarities to TylE (demethylmacrocin *O*-

methyltransferase) and TylF (macrocin *O*-methyltransferase), respectively, in the tylosin biosynthetic pathway (18). Accordingly, the function of MycE was proposed to methylate the C2-OH group of 6-deoxyallose in mycinamicin VI, leading to mycinamicin III, whereas MycF was presumed to transfer a methyl group to the C3-OH group of javose (that is, C2-methylated 6-deoxyallose) in mycinamicin III to generate mycinamicin IV (Fig. 7-2). The proposed functions of MycE and MycF were supported by *in vivo* precursor feeding studies (16, 19). Herein, we report the expression of *mycE* and *mycF* in *Escherichia coli*, and purification of MycE and MycF to establish their biochemical function for regiospecific deoxysugar *O*-methylation in mycinamicin macrolide antibiotics.

7.2 Results and Discussion

7.2.1 Protein Sequence Analysis of MycE and MycF

Comparative analysis revealed that MycE (399 amino acids) is significantly larger than MycF (222 amino acids). Alignment of these two *O*-methyltransferases exhibits low sequence identity (11.3 %), suggesting they might have evolved from distinct ancestors. A protein BLAST search revealed a number of *O*-methyltransferases with either MycE or MycF as the query protein. Interestingly, all candidates with high sequence similarities are *O*-methyltransferases involved in deoxysugar biosynthesis. In the phylogenetic tree (Fig. 7-3A) of selected *O*-methyltransferases with high similarities to MycE and MycF, it is evident that they are located in distinct branches; this indicates their potentially disparate evolutionary origins. In the sequence alignment of MycE with corresponding close relatives (Fig. 7-3B and 7-3C), three conserved motifs (motifs I-III) (9, 20) were identified that are predicted to contribute to SAM binding. In contrast, there are only two conserved SAM binding motifs (motifs I and III) found in MycF.

7.2.2 *In Vitro* Characterization of MycE and MycF

To confirm the proposed function of MycE and MycF, we cloned *mycE* and *mycF* genes into pET28b for overexpression in *E. coli* BL21 (DE3). The recombinant N-terminal His_6-tagged proteins were purified to homogeneity by one-step Ni-NTA agarose chromatography (Fig. 7-4). With the purified enzymes in hand, we initially tested the

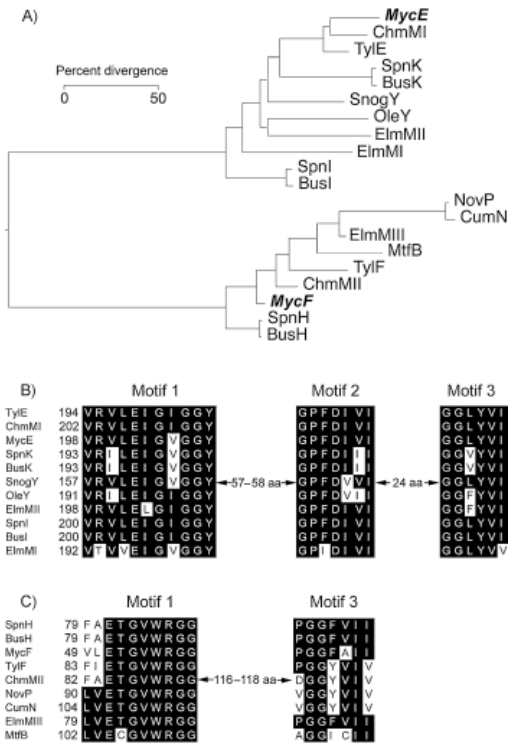


Figure 7-3. Amino acid sequence analysis of MycE and MycF. A) Phylogenetic tree of selected O-methyltransferases generated by MegAlign (DNASTAR) with the Clustal W method. MycE and MycF are highlighted in bold and italic. The selected O-methyltransferases include MycE and MycF (mycinamicin pathway), TylE and TylF (tylosin pathway), ChmMI and ChmMII (chalcomycin pathway), SpnH, SpnI, and SpnK (spinosad pathway), BusH, BusI, and BusK (butenyl spinosyn pathway), SnogY (nogalamycin pathway), OleY (oleandomycin pathway), ElmMI, ElmMII, and ElmMIII (elloramycin pathway), NovP (novobiocin pathway), CumN (coumermycin pathway), and MtfB (mycobacterial serovar-specific glycopeptidolipid pathway). B) SAM binding motifs of MycE and its homologues. C) SAM binding motifs of MycF and its homologues.

activities of MycE and MycF in sodium phosphate buffer (50 mM, pH 7.4) at 30 ^{o}C using SAM as methyl donor and mycinamicin VI and mycinamicin III as substrates (Fig. 7-2). MycE was unreactive toward both substrates, whereas MycF was able to moderately methylate mycinamicin III (but not VI), forming mycinamicin IV (data not shown). Because the dependence of this class of methyltransferases on a metal cofactor is not uncommon, we next investigated the effect of various divalent metal ions on the activities of MycE and MycF. Both enzymes achieved optimal activity in the presence of 10 mM $MgCl_2$ (Fig. 7-5A). A number of alternative divalent ions including Co^{2+}, Fe^{2+}, Mn^{2+}, and Zn^{2+} are capable of supporting suboptimal activities. Interestingly, the Mg^{2+} dependence of MycE appears to be more pronounced than MycF, as MycF remained moderately

active in the absence of Mg^{2+}. The metal dependence of MycF was further assessed by the addition of 2 mM of EDTA; however, this treatment failed to abrogate activity, suggesting that a metal ion might be dispensable in the O-methyltransferase reaction catalyzed by MycF. In contrast, EDTA significantly lowered the activity of MycE even in the presence of 10 mM Mg^{2+}. The optimal pH and temperature range for MycE and MycF assays were determined through comparison of enzymatic activities under various reaction conditions. The optimal reaction conditions of the two *O*-methyltransferases is pH 9.0 (Fig. 7-5B), significantly higher than the corresponding homologues TylE (optimal pH 7.5-8.5) and TylF (pH 7.5-8.0) (7). At pH 9.0, the maximal activities of MycE and MycF were observed at 50 and 37 ^{o}C (Fig. 7-5C), respectively, higher than those of TylE (42 ^{o}C) and TylF (31 ^{o}C) (7).

Under optimal conditions, the in vitro activities of MycE and MycF were analyzed to reveal that MycE methylated the C2-OH of 6-deoxyallose, converting a majority of mycinamicin VI to mycinamicin III (Fig. 7-6B). MycE was incapable of double methylation to generate mycinamicin IV. The second C3-OH methylation of javose was catalyzed by MycF, with a higher conversion than MycE toward mycinamicin VI (Fig. 7-6F). Co-incubation of MycE and MycF with starting substrate mycinamicin VI resulted in the accumulation of both mycinamicins III and IV (Fig. 7-6D). Notably, mycinamicin IV cannot be further methylated by these two methyltransferases, despite the remaining hydroxy group at the C4 position in mycinose. Collectively, it is evident

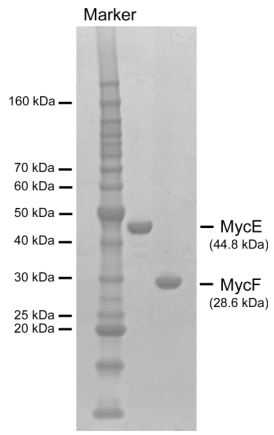


Figure 7-4. SDS-PAGE analysis of purified MycE and MycF.

that both MycE and MycF possess high substrate specificity.

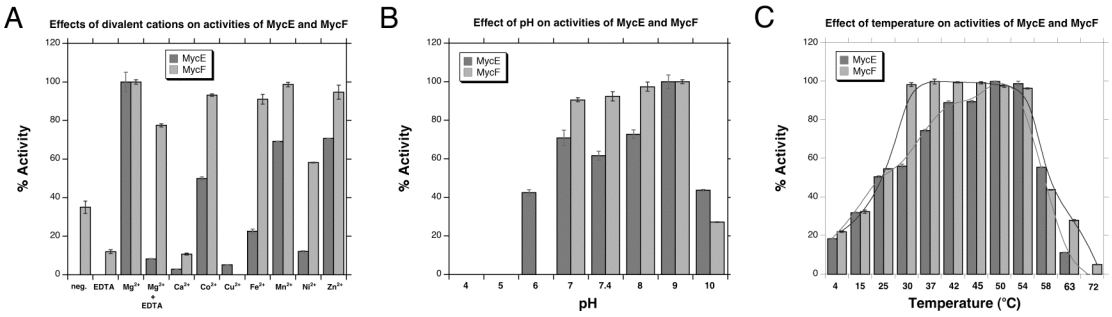


Figure 7-5. Optimization of methylation reactions catalyzed by MycE and MycF. Effects of A) divalent metal ion, B) pH, and C) temperature.

7.2.3 Kinetic Analysis of MycE and MycF

Finally, we determined the steady-state kinetic parameters (Table 7-1) for MycE and MycF based on substrate consumption monitored by HPLC. MycE converted mycinamicin VI to mycinamicin III with a K_M value of 26.4 ± 7.0 μM and a k_{cat} value of 5.0 ± 0.5 min^{-1}. In contrast, MycF methylated mycinamicin III approximately two-fold more efficiently (with respect to k_{cat}/K_M value) than MycE toward mycinamicin VI. Notably, these kinetic data for MycE and MycF are similar to those of TylE and TylF in the tylosin pathway, as previously reported (7).

Table 7-1. Steady-state kinetic parameters of MycE and MycF.

	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\mu M^{-1} \bullet min^{-1}$)
MycE	26.4 ± 7.0	5.0 ± 0.5	0.19
MycF	30.7 ± 6.9	13.5 ± 1.1	0.44

7.3 Summary

The mycinamicin post-PKS (polyketide synthase) tailoring pathway includes two glycosylation steps mediated by two glycosyltransferases (MycB and MycD), four oxidation steps mediated by two cytochrome P450 monooxygenases (MycCI and MycG), and two methylation steps catalyzed by MycE and MycF (16). These post-PKS modifications not only lead to structural diversification, but also confer biologically active properties on the resulting metabolites. We recently confirmed all oxidative

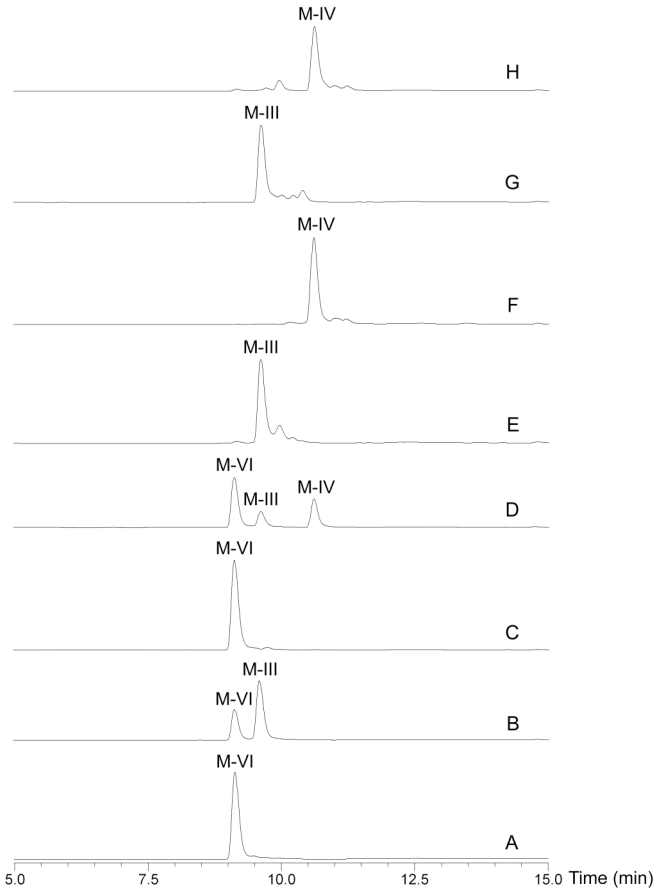


Figure 7-6. LC-MS analysis (UV, 280 nm) of *in vitro* conversions catalyzed by MycE and MycF. A) Mycinamicin VI (M-VI) standard; B) M-VI + MycE; C) M-VI + MycF; D) M-VI + MycE + MycF; E) mycinamicin III (M-III) standard; F) M-III + MycF; G) M-III + MycE; H) mycinamicin IV (M-IV) standard. Compound identity was confirmed by mass spectrometry and comparison with standard compound regarding retention time and co-injection.

tailoring steps *in vitro* through functional analysis of two P450 enzymes (21). This work revealed the importance of both methylation steps in 6-deoxyallose for substrate recognition by the MycG monooxygenase. In this study, we have advanced the knowledge about this pathway by analyzing MycE and MycF *O*-methyltransferases *in vitro*. Taking advantage of the reconstituted optimal *in vitro* assay, the substrate specificity of MycE and MycF and hence the order of sugar modification (mycinamicin VI→III→IV) in this pathway was unambiguously determined. This new information will help facilitate future efforts to manipulate deoxysugar biosynthesis for generation of novel macrolide antibiotics.

7.4 Experimental Procedures

MycE and MycF gene cloning — Using cosmid pMR01 (18) as template, mycE and mycF genes were amplified by PCR under standard conditions with primers as follows: forward, 5'-GGA GTT C*CA TAT G*AC CGC ACA GAC CGA A-3' for *mycE* and 5'-GGA GTT C*CA TAT G*AG CCC GTC GAC CGG A-3' for *mycF* (bases in italics represent the *Nde*I cutting site); reverse, 5'-ACA TCA AGC TTT CAT GTC GCG CCT CCG GA-3' for *mycE* and 5'-ACA TCA AGC TTT CAG GCC GAG CGA CGC CA-3' for *mycF* (underlined bases indicate the *Hind*III restriction site). The gel-purified cDNAs were double digested by *Nde*I and *Hind*III (New England Biolabs), followed by the ligation of fragments containing *mycE* and *mycF* genes into *Nde*I/*Hind*III-treated pET28b (Novagen) to generate recombinant plasmids pET28b-*mycE* and pET28b-*mycF* for the expression of N-terminal His_6-tagged MycE and MycF, respectively. The identity of the inserted gene was confirmed by nucleotide sequencing.

Protein expression and purification — Recombinant plasmids pET28b-*mycE* and pET28b-*mycF* were used to transform *E. coli* BL21 (DE3) cells with a Z-Competent™ Kit (Zymo Research). The resulting transformants were grown at 37 °C in 1 L of LB broth containing kanamycin (50 g/mL) for 2-3 h until OD_{600} reached 0.6-0.8. Isopropyl--$_D$-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.1 mM to induce gene expression, and the cells were cultured at 18 °C overnight. The culture was centrifuged at 5000 g for 10 min to collect cells. The freeze-thaw cell pellet was resuspended in 30 mL of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 10 % glycerol, pH 8.0) and applied to sonication. Cell debris was removed by centrifugation at 35,000 g for 30 min, and the supernatant was mixed with 1 mL of Ni-NTA agarose (Qiagen) for 1 h at 4 °C. The slurry was loaded onto an empty column, and the column was washed stepwise with 10 mL of lysis buffer and 40-60 mL of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 10 % glycerol, pH 8.0). The bound His_6-tagged proteins were eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, 10 % glycerol, pH 8.0). The MycE (45 kDa) and MycF (30 kDa) proteins were further purified and concentrated with 30 kDa and 10 kDa size-exclusion filters (Amicon), respectively. The final desalting step was attained by buffer exchange

into storage buffer (50 mM NaH_2PO_4, 10 % glycerol, pH 7.4) with a PD-10 column (GE Healthcare).

Enzyme assays — The optimized enzyme assay was carried out in 100 L of 50 mM Tris-buffer (pH 9.0) containing 2 M MycE or MycF, 250 M substrate (mycinamicin VI for MycE or mycinamicin III for MycF), 10 mM $MgCl_2$, and 500 M SAM at 50 ^oC (for MycE) or 37 ^oC (for MycE) for 1 h. The reactions were quenched by extraction with CHCl3 (2 × 200 μL). The resulting organic extracts were dried and redissolved in 120 μL methanol. LC-MS analysis of the reaction extract was performed with an LC-MS-2010 EV instrument (Shimadzu) by using an XBridge^{TM} reversed-phase HPLC column (C18, 3.5 μm, 150 mm; Waters) under the following conditions: 20-100 % solvent B over 18 min (solvent A=deionized H_2O+0.1 % formic acid; solvent B=CH_3CN+0.1 % formic acid), flow rate: 0.2 mL/min, UV wavelength: 280 nm.

Steady-state kinetics — The standard reaction buffered with 50 mM Tris-HCl (pH 9.0) contained 0.3 μM MycE or MycF, 500 μM $MgCl_2$, 2-100 M mycinamicin VI for MycE or 3-100 μM mycinamicin III in a total volume of 396 μL. After pre-incubation at optimal temperature for 5 min, the reactions with various substrate concentrations were initiated by adding 4 μL of SAM (50 mM), and three aliquots (100 μL) were taken at three time points (0, 1, 2 min and 0, 2, 4 min for reactions with substrate concentrations < 40 μM and > 60 μM, respectively) within the linear range to thoroughly mix with 100 μL of methanol for reaction termination. The proteins were removed by centrifugation at 16,000 g for 15 min. The supernatant was subject to HPLC analysis to monitor substrate consumption within the linear range, thereby deducing the initial velocity of the *O*-methylation reaction. The HPLC conditions were: XBridge^{TM} reversed-phase HPLC column (C18, 5 μm, 250 mm; Waters) 20-100 % solvent B over 20 min (solvent A=deionized H_2O +0.1 % trifluoroacetic acid; solvent B=CH_3CN +0.1 % trifluoroacetic acid), flow rate: 1.0 mL/min, UV wavelength: 280 nm. All measurements were performed in duplicate, and velocities determined under different substrate concentrations were fit into the Michaelis-Menten equation to calculate the kinetic parameters.

7.5 References

1. Thibodeaux, C. J., Melancon, C. E., and Liu, H.-w. (2007) *Nature* **446**, 1008-1016
2. Trefzer, A., Salas, J. A., and Bechthold, A. (1999) *Nat. Prod. Rep.* **16**, 283-299
3. Weymouth-Wilson, A. C. (1997) *Nat. Prod. Rep.* **14**, 99-110
4. Zubieta, C., He, X., Dixon, R. A., and Noel, J. P. (2001) *Nat. Struct. Biol.* **8**, 271-279
5. Fouces, R., Mellado, E., Diez, B., and Barredo, J. L. (1999) *Microbiology* **145**, 855-868
6. Bauer, N. J., Kreuzman, A. J., Dotzlaf, J. E., and Yeh, W. K. (1988) *J. Biol. Chem.* **263**, 15619-15625
7. Kreuzman, A. J., Turner, J. R., and Yeh, W. K. (1988) *J. Biol. Chem.* **263**, 15626-15633
8. Bate, N., and Cundliffe, E. (1999) *J. Ind. Microbiol. Biotechnol.* **23**, 118-122
9. Patallo, E. P., Blanco, G., Fischer, C., Brana, A. F., Rohr, J., Mendez, C., and Salas, J. A. (2001) *J. Biol. Chem.* **276**, 18765-18774
10. Paulus, T. J., Tuan, J. S., Luebke, V. E., Maine, G. T., DeWitt, J. P., and Katz, L. (1990) *J. Bacteriol.* **172**, 2541-2546
11. Aguirrezabalaga, I., Olano, C., Allende, N., Rodriguez, L., Brana, A. F., Mendez, C., and Salas, J. A. (2000) *Antimicrob. Agents Chemother.* **44**, 1266-1275
12. Waldron, C., Matsushiba, P., Rosteck, P. R., Broughton, M. C., Turner, J., Madduri, K., Crawford, K. P., Merlo, D. J., and Baltz, R. H. (2001) *Chem. Biol.* **8**, 487-499
13. Satoi, I., Muto, N., Hayashi, M., Fujii, T., and Otani, M. (1980) *J. Antibiot.* **33**, 364-376
14. Kinoshita, K., Takenaka, S., Suzuki, H., Morohoshi, T., and Hayashi, M. (1992) *J. Antibiot.* **45**, 1-9
15. Kinoshita, K., Takenaka, S., and Hayashi, M. (1991) *J. Antibiot.* **44**, 1270-1273
16. Suzuki, H., Takenaka, S., Kinoshita, K., and Morohoshi, T. (1990) *J. Antibiot.* **43**, 1508-1511
17. Anzai, Y., Ishii, Y., Yoda, Y., Kinoshita, K., and Kato, F. (2004) *FEMS Microbiol. Lett.* **238**, 315-320
18. Anzai, Y., Saito, N., Tanaka, M., Kinoshita, K., Koyama, Y., and Kato, F. (2003) *FEMS Microbiol. Lett.* **218**, 135-141
19. Inouye, M., Suzuki, H., Takada, Y., Muto, N., Horinouchi, S., and Beppu, T. (1994) *Gene* **141**, 121-124
20. Kagan, R. M., and Clarke, S. (1994) *Arch. Biochem. Biophys.* **310**, 417-427
21. Anzai, Y., Li, S., Chaulagain, M. R., Kinoshita, K., Kato, F., Montgomery, J., and Sherman, D. H. (2008) *Chem. Biol.* **15**, 950-959

Notes:

This work has been published as "Functional analysis of MycE and MycF, two *O*-methyltransferases involved in the biosynthesis of mycinamicin macrolide antibiotics." **Li, S.**, Anzai, Y., Kinoshita, K., Kato, F., and Sherman, D. H. ***ChemBioChem*** 2009, 10 (8), 1297-1301.

Author contributions:
Shengying Li, and David H. Sherman designed the experiments;
Shengying Li performed the experiments;
Anzai Yojiro, Kenji Kinoshita, and Fumio Kato provided mycinamicin materials.

Chapter 8
Structural Diversification of Tirandamycin by Versatile and Codependent Oxidative Tailoring Enzymes

8.1 Introduction

The antibiotic tirandamycin is of the tetramic acid natural product class defined by a 2,4-pyrrolidinedione ring system that typically originates biosynthetically from the condensation of an amino acid to a polyketide-derived acyl chain to yield a 3-acyl tetramate moiety that confers metal chelating activity (1). Beyond this common feature, this class of compounds enjoys a substantial structural and bioactive diversity represented by the HIV-1 integrase inhibitor equisetin (2), the antimycotic dihydromaltophilin (3), and the first discovered tetramate antibiotic streptolydigin (4) (Str, **1**), for which the specific biological target was found to be bacterial RNA polymerase (RNAP) (5). Recently, Str has generated considerable interest as a chemical tool for the inhibition of RNAP (6), and as a subject of biosynthetic study in isotope feeding experiments (7, 8). The structural feature of predominant biosynthetic interest in Str is an intriguing bicyclic ketal skeleton (Fig. 8-1) that is common to tirandamycin, tirandalydigin (**2**), BU-2313B (**3**), and nocamycin II (**4**), all of which also possess antimicrobial activity (1). Tirandamycin A (TirA, **5**) and tirandamycin B (TirB, **6**) were originally discovered in the 1970's from the fermentation broths of terrestrial *Streptomyces* species as anti-microbial agents (10, 11). In this current study, we have discovered three new tirandamycin derivatives including tirandamycin C (TirC, **7**), tirandamycin D (TirD, **8**), and tirandamycin E (TirE, **9**) together with known **5** and **6** from the marine-derived environmental isolate *Streptomyces* sp. 307-9. These five tirandamycins (**5-9**, Fig. 8-1) differ in the oxidative modification state of the bicyclic ketal moiety, and this

modification state was found to be a key determinant of potency against vancomycin-resistant *Enterococcus faecalis* (VRE). Additionally, the co-crystal structures of *Escherichia coli* and *Thermus thermophilus* RNAP complexed with Str display extensive contacts with the bicyclic ketal (12, 13). Taken together, these observations indicate that the formation and modification of the bicyclic ketal is critical for the production of a potent bioactive molecule, but to date there has been no genetic or biochemical insight into this process. We herein report the elucidation of the tirandamycin gene cluster (Tam) and the *in vitro* reconstitution of the complete oxidative pathway for structural diversification of tirandamycin, leading to formation of tirandamycin A-E via a cascade of oxidative modifications on the bicyclic ketal.

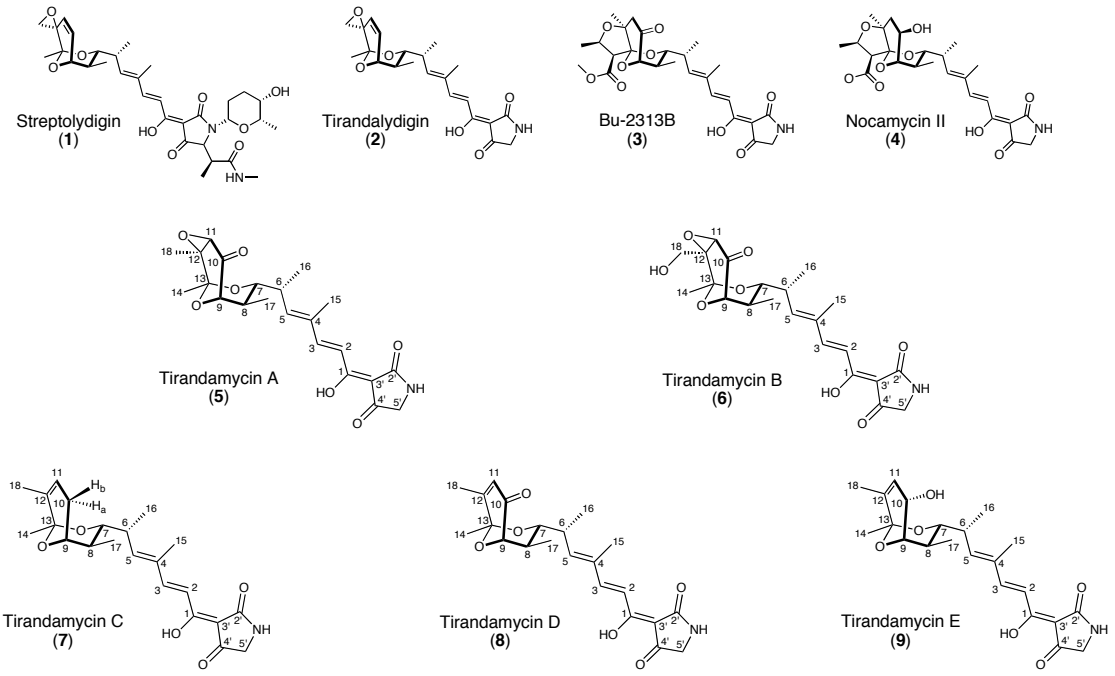


Figure 8-1. Tetramic acid natural products bearing a bicyclic ketal moiety.

8.2 Results and Discussion

8.2.1 Isolation and Structural Identification of TirC and TirD

During a screen to discover new natural products from marine-derived actinomycetes with activity against vancomycin-resistant *Enterococcus faecalis* (VRE), we isolated from marine sediments a bacterium of *Streptomyces*-like morphology

designated as strain *Streptomyces sp*. 307-9. Analysis of extracts from shake-flask fermentation identified the major anti-VRE components in the culture broth as the previously reported TirA (**5)** and TirB (**6**). Varying the duration of culture prior to extraction indicated a gradual change in the ratio of the two major products, with **6** accumulating over time at the expense of **5**. This suggests that **5** is a transiently secreted or diffusible intermediate *en route* to the biosynthesis of the final product **6**. Isolation of **5** was likely enabled by the capture of this intermediate from the culture broth via early harvest. We imagined that an adsorbent resin might be used to capture and observe previously undetected biosynthetic intermediates, protecting them from bioconversion and facilitating their isolation.

Table 8-1. Chemophysical properties of tirandamycins.

Compound	Appearance	$[\alpha]_D$	UV, λ_{max}	HRESIMS	
				calc.	***obs.***
TirA, **5**	Yellow	+9 (*c* 0.3, EtOH)	238, 356 nm	$[M+H]^+$ 418.1860	ND
TirB, **6**	Yellow	-14 (*c* 0.5, EtOH)	238, 356 nm	$[M+H]^+$ 433.1809	ND
TirC, **7**	Yellow	-59 (*c* 0.11, EtOH)	238, 356 nm	$[M+Na]^+$ 410.1944	$[M+Na]^+$ 410.1949
TirD, **8**	Yellow	-60 (*c* 0.02, EtOH)	238, 356 nm	$[M+Na]^+$ 424.1736	$[M+Na]^+$ 424.1737
TirE, **9**	Yellow	ND	238, 356 nm	$[M+H]^+$ 404.2068	ND
TirA*	Yellow	+4 (*c* 0.5, EtOH)	238, 356 nm	-	-
TirB*	Yellow	-8 (*c* 0.55, EtOH)	238, 356 nm	-	-

*Tirandamycins isolated from *S. tirandis*
ND: Not determined

We and others have previously used XAD-16 resin in microbial fermentation to increase the yield of a desired final product (sequestration of the final product presumably evades feedback repression mechanisms), facilitate convenient extraction from culture broth (14-17), or stabilize a desired fermentation product against spontaneous degradation and/or enzymatic conversion (18, 19). We added sterile XAD-16 resin bags to shake flask cultures of *Streptomyces* sp. 307-9 entering stationary phase and allowed fermentation to proceed for several days, after which the mesh bags were removed and the resin extracted. Reverse-phase HPLC analysis of the resulting crude extracts revealed two major new peaks that share the characteristic UV spectrum of **5** and **6** (Table 8-1). Notably, these two presumable intermediates (identified as new

tirandamycin derivatives TirC, **7** and TirD, **8**) were never observed when resin was omitted from the culture flask. Using the resin-trapping method, **7** and **8** were observed reproducibly and were purified by semi-preparative HPLC.

Molecular formulas for **7** and **8** were established by high-resolution ESI-MS (Table 8-1), which indicated that **8** ($C_{22}H_{27}NO_6$) differs from **5** ($C_{22}H_{27}NO_7$) by the absence of a single oxygen atom, and that **7** ($C_{22}H_{29}NO_5$) differs further by the absence of an additional oxygen and the presence of two additional protons. As shown in Table 8-2, ^{1}H NMR spectra of **7** and **8** supported a close structural relationship to **5** and **6**, demonstrated by the consistent appearance of the olefinic and methyl signals of the dienoyl acyl chain, and the geminal protons of the tetramic acid methylene carbon. 2D correlations observed in COSY, HSQC and HMBC experiments further confirmed the identical composition of this region in **5-8**, and highlighted structural differences in the pendant oxygenation patterns on the bicyclic ketal system. When compared to **5**, **8** was determined to contain a C-11/C-12 *cis* double bond in place of the epoxide, indicated by the chemical shift values for positions 11 (δ_H 6.08, δ_C 127.4) and 12 (δ_C 156.7), and the adjacent carbonyl at C-10 (δ_C 195.7). This assignment was further supported by 2D correlations including (a) a long-range COSY coupling between H-11 and H-18, and (b) HMBC correlations from H-11 to C-13 and C-18, H-18 to C-11 and C-12, and H-14 to C-12. **1** was also determined to contain this double bond, but presented a methylene carbon (δ_C 24.5) in place of the C-10 carbonyl of **2**, demonstrated by the appearance of new multiplet signals for H-10a (δ_H 1.96) and H-10b (δ_H 2.33) and the reduced deshielding at positions 9 (δ_H 3.90, d_C 71.4) and 11 (δ_H 5.70, δ_C 123.6). This feature was confirmed by 2D COSY correlations including (a) a geminal coupling between H-10a and H-10b and (b) couplings between the H-10 protons and the H-11 olefin signal, the H-9 methine signal (coupled to H-10b only) and a long range coupling to the H-18 methyl signal (See Supporting Information for a complete mapping of 2D correlations). Assignment of the diastereotopic protons of the C-10 center of **1** was achieved through analysis of a NOESY spectrum which indicated (a) proximity between H-10a and H-17, and (b) proximity between H-10b and H-9. With this information, H-10a was determined to be the *pro*-S proton and H-10b was assigned as *pro*-R. To confirm the absolute stereochemistry of **5-8**, we obtained optical rotation measurements for tirandamycins A and B from both

Table 8-2. ^{1}H, ^{13}C, and 2D NMR data for Tirandamycin A, C, D, and E in CD_2Cl_2.

	TirA (5)		Tir D (8)			TirC (7)				TirE (9)				
No.	δ_C[a]	δ_H (*J* Hz)	δ_C	δ_H (*J* Hz)	COSY[d]	HMBC[e]	δ_C	δ_H (*J* Hz)	COSY[d]	HMBC[e]	NOESY[f]	δ_C	δ_H (*J* Hz)	COSY[d]	HMBC[e]	NOESY[f]
1	175.5		175.5				175.5					175.5				
2	117.1	7.15, dd (15.8, 0.4)	116.7	7.15, dd (15.8, 0.4)	3	1, 4	116.1	7.12, dd (15.7, 0.4)	3	1	3, 15	116.3	7.12, d (15.7)	3	1, 4	
3	150.0	7.58, dd (15.8, 0.8)	150.2	7.61, dd (15.8, 0.7)	2	1, 5, 15	150.4	7.62, dd (15.7, 0.8)	2	1	2, 5	150.6	7.61, d (15.8)	2	1, 4, 5, 15	
4	135.3		135.5				134.9					134.8				
5	144.5	6.24, d (9.9)	145.8	6.30, d (10.1)	6, 15	3, 15	147.5	6.32, d (10.2)	6, 15	3, 15	3, 6, 8, 16	147.2	6.32, d (10.2)	6, 15	3, 15	
6	34.9	2.87, m	34.8	2.89, m	5, 16	4, 5, 16	34.9	2.83, m	5, 7, 16		5, 7, 15, 16, 17	35.0	2.89, m	5, 16	5, 16	7, 15, 16, 17
7	77.4	3.58[c]	77.7	3.44, dd (11.3, 2.1)	8	5	77.0	3.49, dd (11.2, 2.1)	6, 8		6, 8, 16, 17	77.2	3.72, dd (11.2, 1.6)	8	5	6, 8, 16, 17
8	35.0	1.97, m	33.9		7, 9, 17	7, 9	35.5	1.84, m	7, 9, 17		5, 7, 9, 17	36.8	1.96, m	7, 9, 17		6, 7, 9, 17
9	79.4	3.98, d (6.1)	79.5	3.97, d (5.8)	8	7, 8, 10, 13	71.4	3.90, bd (6.5)	8, 10b	7, 8, 11, 13	8, 10b, 17	73.9	3.91, dd (5.6, 5.6)	8, 10b	7, 8, 10, 11, 13	8, 10b, 17
10a	203.2		195.7				24.5	1.96, m	10b, 11, 18		10b, 11, 17					
10b							24.5	2.33, m	9, 10a, 11, 18		9, 10a	69.4	4.72, m	9, 18	11	9, 11
11	61.6	3.25, s	127.4	6.08, s	18	13, 18	123.6	5.70, bs	10a, 10b, 18		10a, 18	129.0	5.71, bs	18	13, 18	7, 10b, 18
12	57.5		156.7				133.2					134.0				
13	97.4		96.7				96.1					96.4				
14	22.8	1.53, s	24.6	1.54, s		12, 13	24.5	1.38, s		12, 13		24.6	1.38, s		12, 13	
15	12.5	1.91, d (1.3)	12.4	1.91, d (1.1)	5	3, 4, 5	12.4	1.91, d (1.3)	5	3, 4, 5	2,6	12.5	1.91, bs	5	3, 4, 5	6, 7, 17
16	17.2	1.14 d (6.8)	17.0	1.07, d (7.0)	6	5, 6, 7	17.2	1.05, d (7.0)	6	5, 6, 7	5, 6, 7	17.3	1.06, d (6.9)	6	5, 6, 7	6, 7, 17
17	11.6	0.71, d (7.0)	11.6	0.69, d (7.2)	8	7, 8, 9	13.2	0.68, d (7.0)	8	7, 8, 9	6, 7, 8, 9, 10a	13.2	0.94, d (7.5)	8	7, 8, 9	6, 7, 8, 9, 16
18	15.8	1.46, s	19.4	1.92, d (1.5)	11	11, 12, 13	18.3	1.61[c]	10a, 10b, 11	11, 12, 13	11	18.0	1.63, bs	10b, 11	11, 12, 13	11
1'				5.70, s	5'								5.76, bs			
2'	176.9		177.0				177.1					177.0				
3'	N/A[b]		N/A[b]				N/A[b]					100.5	N/A[b]			
4'	193.1		193.0				193.1					192.9				
5'	52.1	3.78, s	51.9	3.78, s	1'	2', 4'	51.9	3.78, s		2', 4'		52.1	3.78, s		2', 4'	

[a]Shift values for ^{13}C assignments taken from HSQC, HMBC, and ^{13}C data. [b]Signal not observed in HSQC and HMBC of **5**, **7**, **8**, and **9**, but seen in ^{13}C spectrum of **6** (not shown). [c]Signal obscured in ^{1}H spectrum but assigned by 2D spectra. [d]Correlations are between protons at positions indicated. [e]Correlations are from proton to indicated carbon positions. [f]NOESY data reported for 0.4-6.0 ppm.

Streptomyces sp. 307-9 and *S. tirandis* NRRL 3689, the original strain from which tirandamycin A was isolated and characterized by X-ray crystallography (6, 20). $[\alpha]_D$

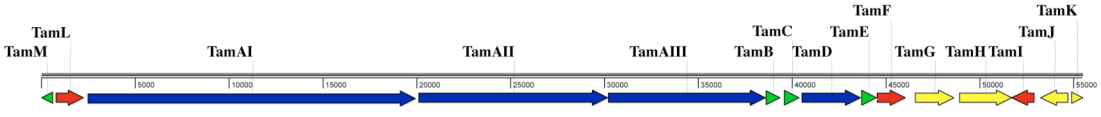
values of **5** and **6** from *Streptomyces* sp. 307-9 are consistent with those for tirandamycin A and B from *S. tirandis* (Table 8-1), indicating the same absolute configuration. Based on these data, the configuration of **7** and **8** was presumed to correspond to that of **5** and **6**.

8.2.2 Anti-VRE Activities of Tirandamycins

With purified **5-8** in hand, we compared their anti-VRE activities in a microtiter plate format with a consecutive two-fold dilution series of each compound spanning a concentration range of six orders of magnitude. **5** exhibited the highest activity with the minimal inhibitory concentration (MIC) of 2.25 μM, whereas all others presenting diminished inhibitory activity: **7**, MIC = 110 μM**, 6** MIC = 100 μM, **8** MIC > 9 μM. The MIC for **8** can only be expressed as a lower limit because of a minor presence of **3** in this sample that accounts for the apparent MIC. A comparison of these activities suggests that the C-10 ketone and C-11/C-12 epoxide (present in **5** but not in **7**) confer increased potency, but this affect can be attenuated by the hydroxyl group at C-18 (as in **6**). Previous studies have shown that inhibition of bacterial RNA polymerase (RNAP) is the primary mechanism behind the antibacterial activity of tirandamycin and streptolydigin (9, 21-23). The crystal structures of *E. coli* and *Thermus thermophilus* RNAP complexed with streptolydigin have been solved and revealed an abundance of key contacts in the streptolol ring system (structurally analogous to the bicyclic ketal system of tirandamycin) (12, 13). This is consistent with the bioassay results presented here which demonstrate that the substituents of this moiety are key determinants of potency, assuming that *in vivo* anti-VRE activity reflects inhibition of RNAP.

8.2.3 Elucidation of Tirandamycin Biosynthetic Gene Cluster

To understand the details of tirandamycin biosynthesis and especially the mechanism for varying oxidative modifications of the bicyclic ketal ring system, the tirandamycin biosynthetic gene cluster in *Streptomyces* sp. 307-9 has been elucidated (24). The contiguous 56 kb sequence assigned as the tirandamycin biosynthetic cluster (Tam) contains 15 open reading frames (ORFs) including three PKS (polyketide synthases) genes and one NRPS (nonribosomal polypeptide synthetase) gene, and several genes proposed to be involved in post-assembly line tailoring steps, self-resistance, and



ORF	Size[a]	Closest Homologue and Origin	%I[b]	%S[c]	Proposed Function
tamM	211	YP001106197 *Saccharopolyspora erythraea* NRRL 2338	52	65	Phosphopantetheinyl transferase
tamL	500	CAJ42334 *Streptomyces steffisburgensis*	58	71	FAD-dependent oxidase
tamAI	5828	NP822114 *Streptomyces avermitilis* MA-4680	45	57	PKS (4 modules)
tamAII	3365	AAZ94390 *Streptomyces neyagawaensis*	55	67	PKS (2 modules)
tamAIII	2801	AAX98191 *Streptomyces aizunensis*	57	68	PKS (2 modules)
tamB	258	NP821582 *Streptomyces avermitilis* MA-4680	53	68	Type II thioesterase
tamC	276	CAN89637 *Streptomyces collinus*	44	56	Conserved hypothetical protein
tamD	1051	CAN89638 *Streptomyces collinus*	49	62	NRPS (1 module)
tamE	276	YP946579 *Arthrobacter aurescens* TC1	35	49	Glycoside hydrolase
tamF	512	BAD86797 *Streptomyces* sp. KO-3988	39	55	Terpene synthase
tamG	697	NP825193 *Streptomyces avermitilis* MA-4680	63	74	DNA helicase
tamH	937	AAX98177 *Streptomyces aizunensis*	37	51	LuxR-type transcriptional activator
tamI	413	YP001107923 *Saccharopolyspora erythraea* NRRL 2338	55	66	Cytochrome P450 monooxygenase
tamJ	498	NP825134 *Streptomyces avermitilis* MA-4680	56	69	Efflux pump
tamK	210	CAJ89338 *Streptomyces ambofaciens* ATCC 23877	59	73	TetR-type transcriptional repressor

Figure 8-2. Organization of the tirandamycin biosynthetic gene cluster. Genes are classified by color for encoding assembly line enzymes (blue), tailoring enzymes (red), proteins involved in regulation and resistance (yellow), and proteins with accessory or unknown functions (green).

regulation of cluster expression (Fig. 8-2). The predictions of the building block specificity of the PKS/NRPS assembly and the stereochemistry of modifications during linear chain elongation based on the analysis of the conserved sequence motifs in PKS/NRPS are well consistent with the resultant tirandamycin structure (24). More importantly, two oxidative enzyme genes including the cytochrome P450 monooxygenase gene *tamI* and the flavin-dependent oxidase gene *tamL*, whose encoding proteins are likely responsible for the oxidative modifications of tirandamycin, were identified within the Tam cluster. However, mechanisms for the bicyclic ketal formation and the tetramic acid ring formation upon linear chain release remain obscure due to lack of apparent enzyme candidates.

8.2.4 Characterization of the oxidative cascade catalyzed by TamI and TamL

TirC (**7**), represents the least modified intermediate that has been identified from culture broth of *Streptomyces* sp. 307-9. Transformation into TirB (**6**), the most modified and final product, requires several oxidative modifications. We had originally anticipated that a cytochrome P450 enzyme would be capable of the epoxidation and hydroxylation steps, and indeed the presence of the P450 homolog TamI in the Tam cluster encouraged

this hypothesis. Sequence analysis of TamI revealed that it contains all of the highly conserved motifs of this enzyme class, including the Helix I, ExxR, and heme binding motifs (Fig. 8-3A) (25). We overexpressed and purified His_8-tagged recombinant TamI (Fig. 8-S1) in *E. coli* to yield an orange-red enzyme solution that was characterized by UV-visible spectroscopy using standard techniques (26). The sodium dithionite reduced enzyme solution displayed an absorbance peak at 420 nm with a 450 nm peak arising after bubbling of the solution with CO (Fig. 8-3C), though even after this CO treatment the enzyme preparation existed mainly as a presumably inactive P420 species, as was observed for the EpoK enzyme of epothilone biosynthesis (27). However, unlike EpoK, incubation with substrates failed to rescue the inactive P420 species of TamI, implying that only a small subpopulation of this enzyme is catalytically competent.

To test the hypothesis that TamI installs the C-11/C-12 epoxy and the C-18 hydroxy groups, the recombinant enzyme was incubated with the substrate tirandamycin A in reaction buffer containing NADPH, spinach ferredoxin, and spinach ferredoxin-$NADP^+$ reductase. LC-MS analysis of reaction supernatants revealed a small but reproducible conversion of TirA to TirB by hydroxylation at C-18. Under the same reaction conditions but with the substrate TirD, we observed a complete conversion of the

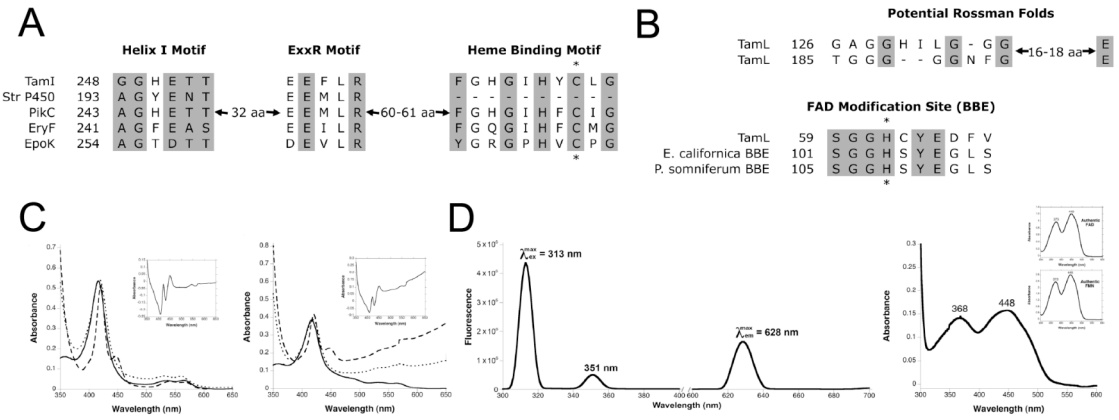


Figure 8-3. Sequence and spectral analysis of TamI and TamL. (A) TamI sequence motifs typical of cytochrome P450 enzymes. (B) TamL sequence motifs include Rossmann fold motifs (top) and a berberine bridge enzyme motif proposed to be the site of covalent attachment of the FAD cofactor. (C) UV-vis absorption spectrum for purified His_8-tagged TamI P450 (left) and His_6-tagged TamI-RhFRED fusion (right) in oxidized form (solid line), sodium dithionite reduced form (dotted line), and CO reduced form (dashed line); insets show CO-bound reduced difference spectra. (D) Fluorescence excitation and emission spectra (left) and UV-vis absorption spectrum (right) for purified His_8-tagged TamL flavoprotein; inset shows UV-vis absorption spectrum of pure FAD and FMN cofactors.

substrate predominantly to TirA and a small amount of TirB (Fig. 8-4A). In addition, TirB cannot be further modified by TamI. Taken together, these experiments demonstrate that the penultimate and final steps of tirandamycin biosynthesis are a C-11/C-12 epoxidation and a C-18 hydroxylation, respectively.

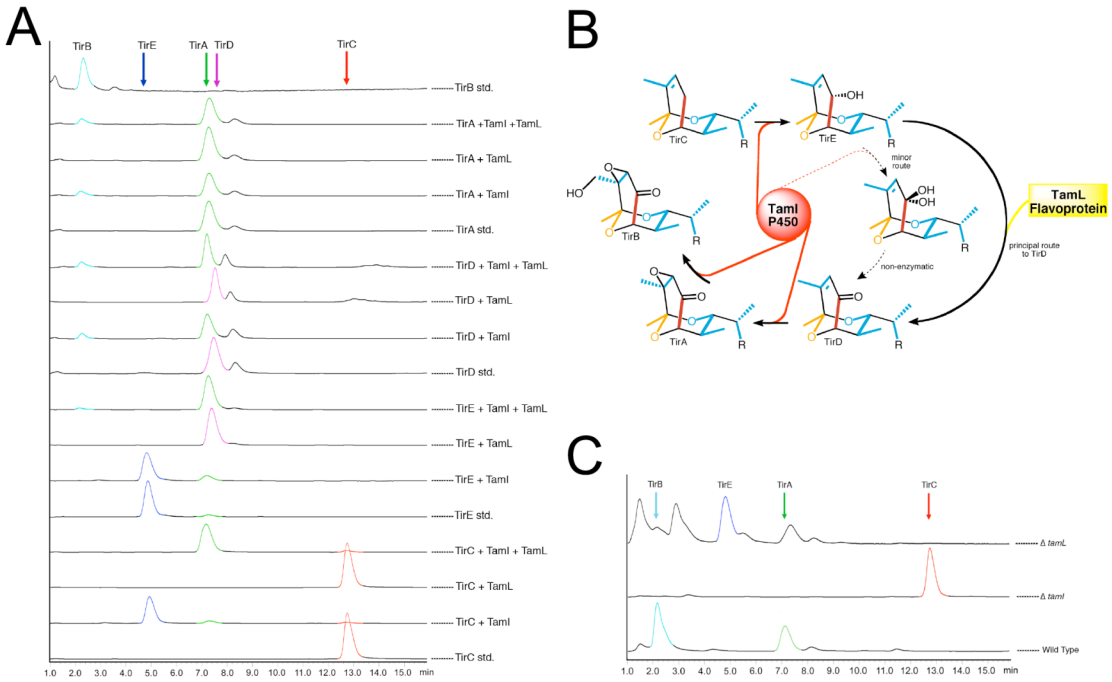


Figure 8-4. Oxidative cascade modifications of the tirandamycin bicyclic ketal. (A) *In vitro* reconstitution of TamI and TamL mediated oxidation steps. Peaks were identified by comparison to authenticated standards and MS detection of anticipated species. (B) The complete oxidative pathway as elucidated by biochemistry and mutagenesis studies; the predominant TamI P450 reactions are highlighted in red with the intervening TamL oxidation reaction highlighted in yellow; the dashed arrows represent a trace activity proposed to afford a minor TamL-independent route. (C) Metabolite profiles of wild type and mutant strains. Scales of each chromatogram have been zoomed as follows: Wild Type 4.1×, Δ*tamI* 1×, Δ*tamL* 17.5×.

Next, we tested the possible role of TamI in formation of the C-10 ketone by incubating TirC with recombinant P450 as above, and observed the formation of a very small amount of TirA (Fig. 8-4A), in addition to a new species that had not been identified previously. The new species, which we named tirandamycin E (TirE, **9**), possessed a molecular mass 16 amu greater than TirC and was of an increased polarity consistent with a tentative hydroxylation. To characterize TirE by NMR and identify the site of hydroxylation, we acquired several milligrams of product prepared by enzyme

reaction, which was cost prohibitive using the existing reaction setup with expensive spinach ferredoxin and ferredoxin-$NADP^+$ redox partners that are required to activate molecular oxygen. To overcome this limitation, we employed a recently developed strategy (28) to make a self-sufficient biosynthetic P450 reaction system by fusing TamI to the RhFRED (Fig. 8-S1), an FMN/Fe_2S_2 containing reductase domain, thus obviating the need for costly exogenous redox partners. Using this system we conducted a preparative scale conversion of TirC to TirE, purified the product by preparative HPLC and assigned the structure using extensive 1D and 2D NMR analysis (Table 8-2), which clearly indicated a C-10 hydroxylation. The stereochemistry was assigned based on the observation of a through-space deshielding effect between the C-10 hydroxyl and the H-7 and H-17 protons, and also by comparison of NOESY data between TirC and TirE which confirmed the disappearance of the *pro*-S proton at C-10.

These results demonstrate that the major activity of TamI on the C-10 center is a hydroxylation, but the minor conversion to TirA by TamI alone indicates that this enzyme also has a trace activity capable of installing the ketone. It is likely that this reaction proceeds through sequential hydroxylations at C-10 to form a gem-diol that exists in equilibrium predominantly in the ketone form of TirD. This route to carbonyl formation resembles that used by P450s to install a carboxylate at a terminal carbon, as in the biosyntheses of artemisinic acid (29). However, because this activity for TamI is so weak we speculated that another oxidative enzyme in the cluster might be responsible for the predominant route to TirD, and we therefore examined the ability of TamL, a predicted flavin-dependent oxidoreductase, to catalyze this reaction. Detailed sequence alignments revealed potential Rossmann fold nucleotide binding motifs (30) and a berberine sequence motif that bears the site of covalent attachment of FAD in some FAD-dependent enzymes (Fig. 8-3B) (31). We overexpressed and purified recombinant His_8-tagged TamL (Fig. 8-S1) in *E. coli* to yield a deep yellow protein solution with fluorescence excitation/emission and UV-visible absorption spectra consistent with a flavoprotein (Fig. 8-3D) (31, 32). We verified that the cofactor was covalently bound by boiling the protein solution and observing that the yellow color was retained with the denatured protein, and that the supernatant did not contain any released cofactor as determined by LC-MS. Acid hydrolysis treatment of the protein released AMP,

confirming FAD instead of FMN as the cofactor (Fig. 8-S2) (33), and MS analysis of the intact protein also supported a single FAD per enzyme molecule (Fig. 8-S3).

Incubation of recombinant TamL with purified TirE resulted in complete conversion to TirD, thus confirming the role of this enzyme in oxidizing the C-10 hydroxyl of TirE to a ketone (Fig. 8-4A). Furthermore, when reactions containing TamI P450 and TirC were amended with TamL, we observed nearly complete conversion to tirandamycin A with no residual TirE detected, indicating that oxidation of TirE by TamL affords the C-10 ketone with far greater efficiency than does TamI alone.

We were also able to confirm the roles of TamI and TamL *in vivo* by genetic disruption of the relevant ORFs which resulted in the accumulation of the anticipated intermediates in the fermentation broth. We developed a protocol for the intergenic conjugation between *Streptomyces* sp. 307-9 (acceptor) and *E. coli* S-17 (donor) (34), and used this system to transfer in suicide knockout vectors created using the REDIRECT system (35). Double-crossover allelic exchange mutants, which contained an in-frame replacement of the ORF by an apramycin resistance gene, were selected/screened on appropriate antibiotics, and verified by PCR screening of gDNA. Verified mutants were grown in fermentation media containing apramycin, and the broths were extracted and analyzed for tirandamycin metabolite profiles. As expected, wild-type cultures accumulated TirA and TirB whereas the P450 TamI knockout accumulated exclusively TirC (Fig. 8-4C), indicating that TamI is the sole P450 in this organism responsible for initiating the oxidative cascade. The flavoprotein TamL knockout accumulated TirE, which is consistent with *in vitro* results in which TamI alone was unable to efficiently catalyze the conversion beyond TirE.

The characterization of this oxidative cascade has revealed a remarkably versatile cytochrome P450 enzyme TamI that catalyzes at least two hydroxylations and one epoxidation at three distinct centers, which to our knowledge is the first reported example of a biosynthetic P450 with such versatile activity. These oxygenations occur in a precisely defined order with no detected promiscuity at a given step, but more notably these three reactions are not consecutive (Fig. 8-4B). Following the initial hydroxylation of TirC by TamI, TamL must oxidize this group to the ketone of TirD, which is the exclusively observed substrate for sequential epoxidation to TirA and hydroxylation to

TirB. The repeated exchange of substrates between these two enzymes represents a unique synergistic biosynthetic pathway, in which a P450 catalyzes multiple oxidations co-dependently with a second biosynthetic enzyme, in contrast to numerous examples of multi-step P450 reactions in which no intervening enzymatic steps are required.

8.3 Experimental Procedures

General Experimental Procedures — All NMR spectra were acquired on a Varian INOVA 400 MHz and a Varian INOVA 600 MHz spectrometer at the Center for Chemical Genomics, University of Michigan. High-resolution ESI-MS spectra were measured at the University of Michigan core facility in the Department of Chemistry using a Waters Micromass AutoSpec Ultima. Optical rotation measurements were obtained on an AUTOPOL III polarimeter. UV spectra were extracted from HPLC photo-diode array chromatograms. RP-HPLC purification was performed using Waters XBridge 5 μm ODS columns and a solvent system of MeCN and H_2O supplemented with 0.1% TFA. LC-MS analysis of HPLC fractions was performed on a Shimadzu 2010 EV ESI spectrometer. NMR spectra were processed using MestReNova software.

Biological materials — *Streptomyces* sp. 307-9 was isolated from marine sediments collected at a depth of 15 meters during a SCUBA expedition in Salt Cay, U. S. Virgin Islands. The procedure for the isolation of Actinomycetes from these samples has been described previously (16). Maintenance and propagation of cultures was performed using standard media and protocols (34). On 2xYT agar, this strain grew as substrate-adherent mycelia that secreted a dark pigment into the media, and produced a top layer of white and purple spores. Light microscopy images of stained cells from agar plate cultures revealed Gram-positive, filamentous cells as well as spore chains. These characteristics are consistent with morphology of *Streptomyces* species. A phylogenetic analysis was performed using a fragment of the 16S rRNA gene amplified from genomic DNA of *Streptomyces* sp. 307-9 using a previously described protocol (36). BLAST analysis revealed the highest sequence identity (97-99%) with previously characterized *Streptomyces* species. This strain is preserved and available from an in-house collection (database index DHS1549). See Supporting Information for microscopy images and 16S

rRNA gene sequence analysis. *Streptomyces tirandis* NRRL3689 was obtained from the USDA Agricultural Research Service.

Culture maintenance and fermentation — Seed cultures of 10 mL 2xYT media were inoculated with a loopful of vegetative cells from an ISP2 agar plate culture of *Streptomyces* sp. 307-9 and incubated with shaking (150 rpm) at 30 °C for 2 days. Six 2 L baffled flasks each containing 1 L of autoclaved production media Md2 (10 g dextrose, 2 g NZ-Amine A, 1 g yeast extract, 0.77 g meat extract, 30 g NaCl per 1 L H_2O) were each inoculated with 1 mL of seed culture and grown for 2 days under identical conditions. To each flask was then added two autoclaved resin bags containing 10 g XAD-16 resin (Supelco, Bellefonte, PA) enclosed in Unitherm 100 fabric (Midwest Filtration, Cincinnati, OH) by heat sealing. Production cultures were grown for an additional four days, after which the resin bags were collected for extraction.

Isolation and purification of tirandamycins — Resin bags were cut open and the free resin was rinsed with water, filtered, and batch extracted with 90% CH_2Cl_2/10% MeOH (3 x 1.5 L). The combined extract was washed twice with an equal volume of 1 M NaCl (adjusted to pH 4.0 with TFA). The dichloromethane layer was collected and dried by rotary evaporation to give a dark yellow oil that was dissolved into 2 mL DMSO. This extract was purified by RP-HPLC on a gradient of 30-100% MeCN and followed by UV/Vis photo-diode array detection. Compounds exhibiting the characteristic absorption spectrum of the tirandamycin tetramic acid moiety (λ_{max} 356 nm) were collected. The purest fractions (as determined by LC-MS analysis) were combined and dried to afford (in order of elution) **6** (3.0 mg), **5** (9.0 mg), **8** (1.4 mg), and **7** (7.0 mg).

Anti-VRE bioassay — 2 μL of 20 mM DMSO solutions of **5-8** were added to 100 μL LB broth (to give 400 μM final concentrations), and two-fold serial dilutions of this stock were prepared in LB. All cultures of Vancomycin-Resistant *Enterococcus faecalis* (VRE) were grown in LB supplemented with 10 μg/μL of vancomycin at 37 °C with shaking (150 rpm). An overnight seed culture of VRE was diluted to an OD_{600} of 0.05, grown to an OD_{600} of 0.4, back diluted to an OD_{600} of 0.005, and 50 μL of this dilution

was added to each well of a microtiter plate that contained 50 μL of a given dilution of compound in LB. Plate cultures were grown for 16 hours and OD_{600} measurements were taken. MIC values were taken as the lowest concentration of compound for which no growth was observed.

Cloning and preparation of recombinant enzymes — *tamI* and *tamL* were PCR amplified from *Streptomyces* sp. 307-9 genomic DNA using primers that append *EcoR*I and *Hind*III sites at the N and C termini, respectively. These PCR products were digested with *EcoR*I/*Hind*III and ligated into a pSJ2 overexpression vector that was similarly digested and dephosphorylated. pSJ2 is an in-house derivative of pET-21a containing a coding region for an N-terminal His_8-tag, and insertion of our gene cassettes creates an in-frame fusion to this tag. To create the TamI-RhFRED fusion protein, the *tamI* gene with stop codon removed was PCR amplified using primers that append *Nde*I and *Eco*RI sites and the N and C termini, respectively. The *Nde*I/*Eco*RI digested product was ligated into similarly digested pET28b-*pikC*-RhFRED as previously reported (28). The accuracy of all constructs was verified by DNA sequencing. These constructs were transformed into *E. coli* BL21(DE3) overexpression host and the resulting proteins were purified using nickel affinity chromatography by following a previously developed procedure (28).

CO-bound reduced difference spectrum of TamI — TamI was identified as an active P450 by obtaining the CO-bound reduced difference spectra using a UV-visible spectrophotometer 300 Bio (Cary). TamI in conversion buffer was reduced by adding several milligrams of sodium dithionite ($Na_2S_2O_4$) and a spectrum was recorded from 350 to 650 nm. After CO bubbling of the solution for 30-60 s, the spectrum of CO-bound reduced P450 species was recorded using the previous reduced spectrum as reference. This assay was also employed to determine the functional P450 concentration using the extinction coefficient of 91,000 $M^{-1} \cdot cm^{-1}$ (26).

Preparation of TirE (9) — TirE was obtained by preparative scale enzymatic conversion in 40 × 500 μl parallel reactions (overnight, 30°C), each containing

approximately 10 μM purified TamI-RhFRED, 0.15 mg TirC, 1 mM NADPH, and 5 μM glucose-6-phosphate and 2 Unit/ml glucose-6-phosphate dehydrogenase for NADPH regeneration. The reaction mixtures were combined and extracted by 3 × 40 mL of CH_2Cl_2, after which the extract was dried, redissolved in 2 mL of methanol, and purified by RP-HPLC to yield TirE. The structural assignment of TirE was based on analysis of ^{1}H, ^{13}C, COSY, HSQC, HMBC, and NOESY data sets.

Enzymatic Assays — The *in vitro* enzymatic conversions of tirandamycins were performed in a total volume of 100 μl of conversion buffer (50mM NaH_2PO_4, pH 7.3, 1 mM EDTA, 0.2 mM dithioerythritol, 10% glycerol) containing 2 μM TamL (or 1 μM TamI partnered by 3.5 μM spinach ferredoxin and 0.1 Unit/ml spinach ferredoxin-$NADP^+$ reductase), 200 μM substrate, and 500 μM NADPH. As negative controls, corresponding enzymes were omitted. After incubation at 30°C for 2 h, the reactions were quenched by extraction, using 2 × 200 μl of $CHCl_3$. The resulting organic extraction was dried, redissolved in 120 μl of methanol, and analyzed by LC-MS using 354 nm UV detection and positive/negative ion MS detection.

Identification of flavin cofactor of TamL — The UV-vis absorption spectrum of TamL was obtained by scanning a 400 μl sample of the purified enzyme (8.6 μM) in a cuvette. A FluoroMax-3 (HORIBA Jobin Yvon) fluorescence spectrophotometer was used to obtain the fluorescence excitation and emmision spectra. To determine the identity of flavin cofactor, following previously developed protocol with minor modifications (33), 200 μl of TirL (9.4 mg/ml) was boiled in presence and absence of 0.1 M HCl for 10 min. The denatured proteins were removed by centrifugation and 3 kDa size exclusion column (YM-3, Microcon), sequentially. The filtrates were analyzed by C18 reverse phase HPLC with established conditions (33).

***TamI and tamL in vivo* genetic disruption** — Knockout vectors were constructed by PCR targeted mutagenesis of appropriate cosmids as described previously (34). This REDIRECT system was used to completely replace *tamI* of cosmid 6G9 (24) and *tamL* of cosmid 2H2 (24) with a cassette bearing an apramycin resistance gene and an origin of

transfer for conjugation. The Apr^R mutagenized cosmid was introduced into *E. coli* S17-1 by transformation and then transferred to *Streptomyces* sp. 307-9 by conjugation. Intergeneric conjugation between *E. coli* S17-1 and *Streptomyces* sp. 307-9 was performed as described previously, with minor modifications (35). An overnight culture of the *E. coli* donor strain was diluted into fresh medium and incubated for 3 hours. The cells were harvested, washed twice and concentrated 10-fold in TS broth. Cells of *Streptomyces* sp. 307-9 were grown in 2xYT broth for several days, harvested by centrifugation, washed and re-suspended in a half volume of TS broth. These recipient cells were mixed with a half volume of *E. coli* donor cells, and 150 µL of the mixture was plated on MS agar. The plates were incubated at 30°C for 20 hours, and then covered with 1 mL water containing 500 µg of nalidixic acid to inhibit further growth of *E. coli* and 1 mg apramycin to select *Streptomyces* sp. 307-9 exconjugants. Incubation at 30°C was continued for 1–2 weeks to allow outgrowth of the exconjugants. Apr^R exconjugants were screened for chloramphenicol sensitivity to detect double-crossover allelic exchange. Exconjugants with the correct antibiotic resistance phenotype were confirmed for gene replacement by PCR and corresponding restriction digest of the PCR products. Metabolite profiles for mutants were determined by culturing and extraction as described earlier, but fermentation was in media supplemented with 50 µg/mL apramycin.

8.4 Supplementary Information

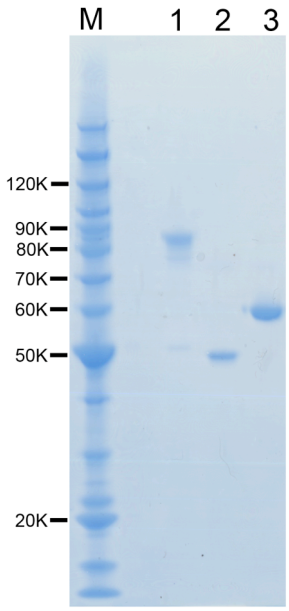


Figure 8-S1. SDS-PAGE analysis of purified recombinant enzymes. M, Protein marker; 1, His_6-tagged TamI-RhFRED; 2, His_8-tagged TamI; 3, His_8-tagged TamL.

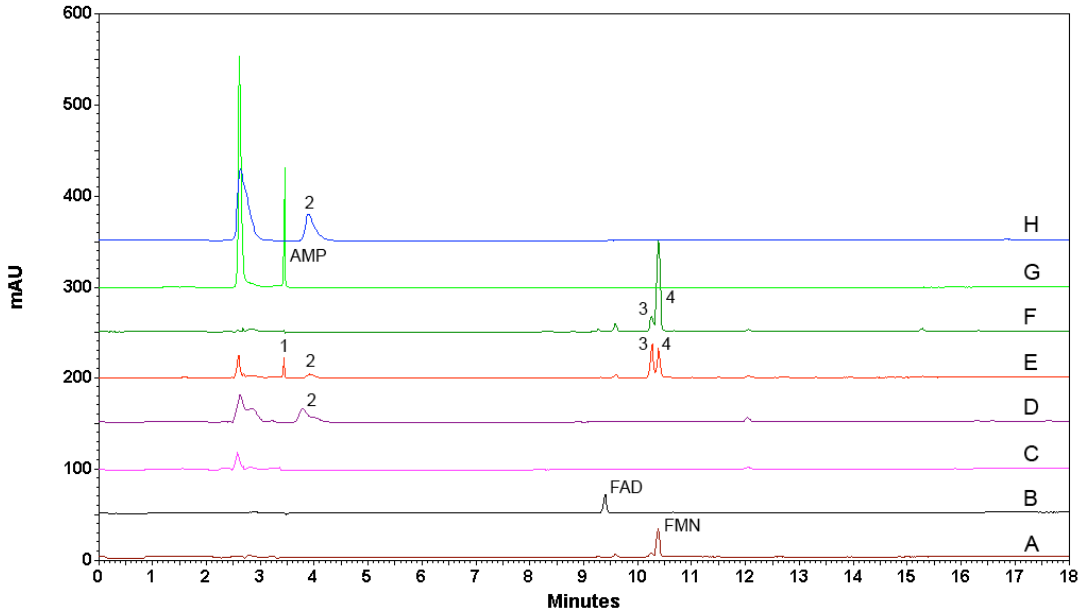


Figure 8-S2. Cofactor analysis of TamL. A. FMN standard; B, FAD standard; C, supernatant of the boiled TamL; D, supernatant of the boiled TamL in presence of 0.1 M HCl; E, boiled FAD in presence of 0.1 M HCl; F, boiled FMN in presence of 0.1 M HCl; G, 5'-AMP standard; H, boiled 5'-AMP in presence of 0.1 M HCl. Peak identities: **1**, 5'-AMP; **2**, decomposited AMP; **3**, phosphorylated riboflavin derivative due to acid hydrolysis of FMN; **4**, FMN.

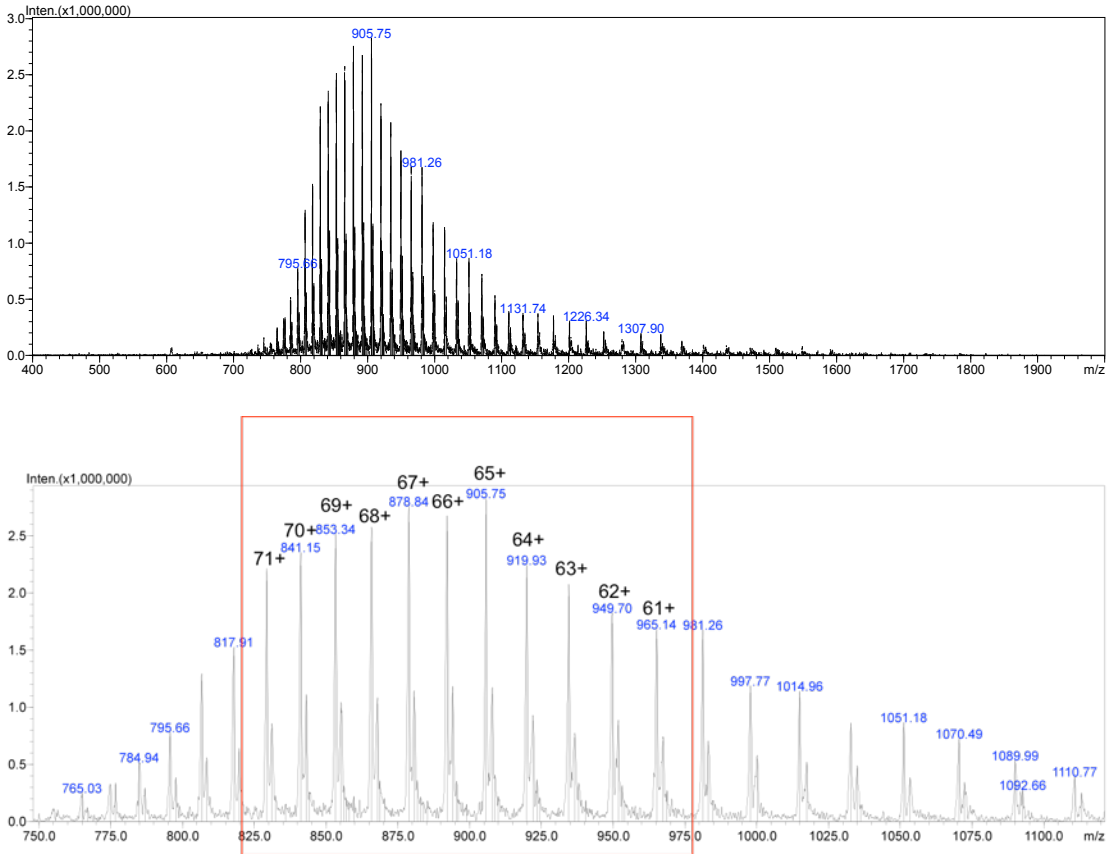


Figure 8-S3. LC-MS analysis of intact TamL showing all multiply charged species (top) and the zoomed in region (bottom) used to deconvolute the neutral mass (peaks in red box). From this data the calculated molecular mass of TamL is 58801.1 +/- 6 Da, which is 32 daltons off from an expected molecular mass of 58769.1 (TamL - Met + FAD - 2H) for a single covalently bound FAD molecule, suggesting the presence of 2 oxidized methionines in the observed species. The observed mass is 359.3 Da away from the expected mass for an FMN cofactor (expected mass 58441.8 Da), thus ruling out the presence of this cofactor in TamL.

8.5 References

1. Schobert, R., and Schlenk, A. (2008) *Bioorg. Med. Chem.* **16**, 4203-4221
2. Hazuda, D. *et al.* (1999) *Antivir. Chem. Chemother.* **10**, 63-70
3. Yu, F. *et al.* (2007) *Antimicrob. Agents Chemother.* **51**, 64-72
4. Rinehart, K. L. *et al.* (1963) *Antimicrob. Agents Chemother.* **161**, 346-348
5. Von Der Helm, K., and Krakow, J. S. (1972) *Nature New Biol.* **235**, 82-83
6. Vassylyev, D. G. *et al.* (2007) *Nature* **448**, 163-168
7. Chen, H., and Harrison, P.H. (2004) *Org. Lett.* **6**, 4033-4036
8. Chen, H., Olesen, S. G., and Harrison, P. H. (2006) *Org. Lett.* **8**, 5329-5332
9. Royles, B. J. L. (1995) *Chem. Rev.* **95**, 1981-2001
10. Meyer, C. E. (1971) *J. Antibiot.* **24**, 558-560
11. Hagenmaier, H. *et al.* (1976) *Arch. Microbiol.* **109**, 65-74
12. Temiakov, D. *et al.* (2005) *Mol. Cell* **19**, 655-666
13. Tuske, S. *et al.* (2005) *Cell* **122**, 541-552
14. Marshall, V. P., Mcwethy, S. J., Sirotti, J. M., and Cialdella, J. I. (1990) *J. Ind. Microbiol.* **5***,* 283-287
15. Lee, J. C., Park, H. R., Park, D. J., Lee, H. B., Kim, Y. B., and Kim, C. J. (2003) *Lett. Appl. Microbiol.* **37***,* 196-200
16. Magarvey, N. A., Keller, J. M., Bernan, V., Dworkin, M., and Sherman, D. H. (2004) *Appl. Environ. Microbiol.* **70***,* 7520-7529
17. Boot, C. M., Gassner, N. C., Compton, J. E., Tenney, K., Tamble, C. M., Lokey, R. S., Holman, T. R., and Crews, P. (2007) *J. Nat. Prod.* **70***,* 1672-1675
18. Lau, J., Frykman, S., Regentin, R., Ou, S., Tsuruta, H., and Licari, P. (2002) *Biotechnol. Bioeng.* **78***,* 280-288
19. Julien, B., and Shah, S. (2002) *Antimicrob. Agents Chemother.* **46***,* 2772-2778
20. Duchamp, D. J., Branfman, A. R., Button, A. C., and Rinehart, K. L. (1973) *J. Am. Chem. Soc.* **95***,* 4077-4078
21. Siddhikol, C., Erbstoeszer, J. W., and Weisblum, B. (1969) *J. Bacteriol.* **99***,* 151-155
22. Yang, X., and Price, C. W. (1995) *J. Biol. Chem.* **270***,* 23930-23933
23. Severinov, K., Markov, D., Severinova, E., Nikiforov, V., Landick, R., Darst, S. A., and Goldfarb, A. (1995) *J. Biol. Chem.* **270***,* 23926-23929
24. Carlson, J. C., Fortman, J. L., Anzai, Y., Li, S., Doug, B. A., and Sherman, D. H. Identification of the tirandamycin biosynthetic gene cluster from *Streptomyces* sp. 107-9. *Submitted to ChemBioChem*
25. Werck-Reichhart, D., and Feyereisen, R. (2000) *Genome Biol.* **1**, REVIEWS3003
26. Omura, T., and Sato, R. (1964) *J. Biol. Chem.* **239**, 2379-2385
27. Ogura, H. *et al.* (2004) *Biochemistry* **43**, 14712-14721
28. Li, S., Podust, L.M., and Sherman, D.H. (2007) *J. Am. Chem. Soc.* **129**, 12940-12941
29. Ro, D. K. *et al.* (2006) *Nature* **440**, 940-943
30. Kleiger, G., and Eisenberg, D. (2002) *J. Mol. Biol.* **323**, 69-76
31. Kutchan, T.M., and Dittrich, H. (1995) *J. Biol. Chem.* **270**, 24475-24481
32. Carter, C. J., and Thornburg, R.W. (2004) *Plant Physiol.* **134**, 460-469
33. Willie, A., Edmondson, D. E., and Jorns, M. S. (1996) *Biochemistry* **35**, 5292-5299
34. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2002) Practical streptomyces genetics (2nd Ed.) John Innes Foundation, Norwich, UK
35. Gust, B. *et al.*(2003) *Proc. Natl. Acad. Sci. U.S.A.* **100**, 1541-1546
36. Marchesi, J. R., Sato, T., Weightman, A. J., Martin, T. A., Fry, J. C., Hiom, S. J., Dymock, D., and Wade, W. G. (1998) *Appl. Environ. Microbiol.* **64***,* 795-9

Notes:

During preparation of this chapter, the first part of this work about identification of tirandamycin C and D has been submitted to ***J. Nat. Prod.*** as "Isolation and characterization of tirandamycins from a marine-derived *Streptomyces* sp.." Jacob, C. Carlson, **Li, S.** (Co-first author), Douglas, A. B., and Sherman, D. H. The second part of this work about characterization of the oxidative cascade will be submitted soon.

Author contributions:
Jacob C. Carlson, Shengying Li, and David H. Sherman designed the experiments;
Jacob C. Carlson, Shengying Li, Douglas, A. B., and Yojiro Anzai performed the experiments.

Chapter 9
Molecular Analysis of the Role of Tyrosine 224 in the Active Site of Streptomyces coelicolor RppA, A Bacterial Type III Polyketide Synthase

9.1 Introduction

Type III polyketide synthases (PKSs) represent the structurally simplest member of the PKS superfamily (1, 2), which also consists of the giant modular multifunctional type I PKS (3, 4) and the discrete multienzyme type II PKS (5, 6). Since the discovery of the first type III PKS, chalcone synthase (CHS), in the 1970s (7), a growing number of type III PKSs have been identified in plants (8–10), bacteria (11–13), and fungi (14). All members of the family characterized to date share some common structural and mechanistic features (2) including a homodimeric architecture, similar subunit molecular mass (40–47 kDa), analogous active site structure, a conserved Cys-His-Asn catalytic triad (thus a common reaction mechanism), and the utilization of acyl-coenzyme A (acyl-CoA) thioesters as substrates instead of the acyl-carrier protein-linked substrates employed by type I and II PKSs. To diversify polyketide product output, type III PKSs have evolved to use a wide range of natural acyl-CoA substrates. For example, the plant type III PKSs, which share 60–95% amino acid identity (2), utilize a variety of different substrates ranging from aliphatic-CoA to aromatic-CoA, from small acetyl-CoA (2-pyrone synthase in *Gerbera hybrida*) (15) to bulky *p*-coumaroyl-CoA (*Medicago sativa* CHS) (16) or from the polar malonyl-CoA (*Aloe arborescens* pentaketide chromone synthase) (17) to the relatively nonpolar isovaleroyl-CoA (*Humulus lupulus* phlorisovalerophenone synthase) (18).

Interestingly, the more divergent bacterial type III PKSs (with as low as 20% amino acid identity) utilize a much narrower pool of natural acyl-CoA substrates than

their relatives in plants (1, 2). Specifically, RppA from either *Streptomyces coelicolor* or *Streptomyces griseus* (19, 20), *Pseudomonas fluorescens* PhlD (21), and *Amycolatopsis mediterranei* DpgA (22, 23) coincidentally employ malonyl-CoA as their physiological starter unit, although alternative starter substrates are known to be accepted by these enzymes (24, 25). Such limited physiological substrate specificity may be misleading because of a paucity of information on functionally identified bacterial type III PKSs (12, 26, 27), a reflection of the short period of time since their discovery. Alternatively, the more limited substrate variability of bacterial type III PKSs might reflect convergent evolution within this family resulting from the more limited availability of acyl-CoA building blocks in prokaryotes compared with eukaryotes. With respect to the limited "building block" diversity in bacteria, it is intriguing that such seemingly divergent bacterial type III PKSs utilize a common starter unit. Understanding the convergence of bacterial type III PKS structure and function could provide important insights into substrate preferences and specificity, potentially enabling expansion of the repertoire of unnatural substrates and thus the creation of functionally diverse and novel natural products.

S. griseus RppA (Sg-RppA) has provided substantial fundamental information about the catalytic properties of the type III PKS subfamily as the first functionally characterized bacterial enzyme of its class (19, 28). The physiological reaction (Fig. 9-1A) is initiated by loading a malonyl-CoA molecule covalently onto the catalytic Cys138 of the RppA polypeptide. Four acetyl-CoA extender units, generated from decarboxylation of malonyl-CoA, are then added sequentially to the elongating linear intermediate via Claisen condensation. Finally, the linear pentaketide intermediate is cyclized and aromatized to form 1,3,6,8-tetrahydroxynaphthalene (THN), although the mechanism of this termination step remains unclear. During growth of the linear polyketide intermediate, the terminal carboxyl group is presumably derived from the starter malonyl-CoA and is prevented from premature decarboxylation, which would result in derailment to acetyl-CoA or triacetic acid lactone (TAL; 4-hydroxy-6-methyl-2-pyrone) via intramolecular cyclization (Fig. 9-1A).

The importance of the active site Tyr224 residue in the first step of this reaction relating to starter unit specificity of Sg-RppA was previously identified through

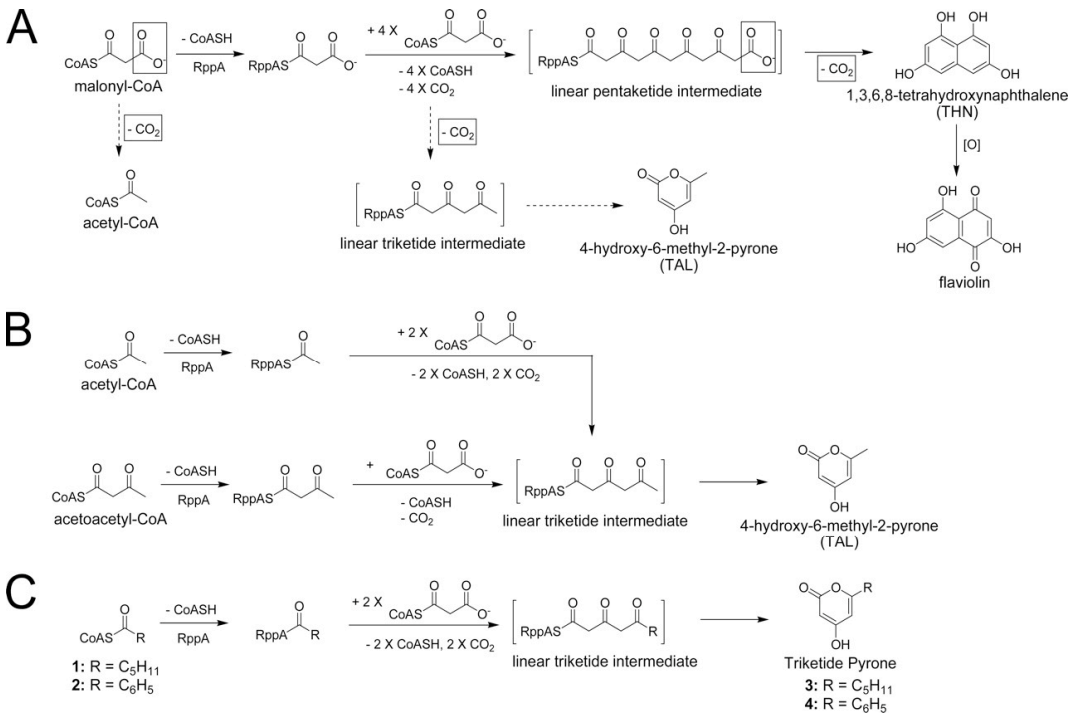


Figure 9-1. Reactions catalyzed by Sc-RppA. A, malonyl-CoA-initiated reaction. The dashed arrows indicate reactions leading to derailed products, and the boxed moieties derive from the starter malonyl-CoA. B, acetyl-CoA and acetoacetyl-CoA-primed reactions. C, hexanoyl-CoA 1- and benzoyl-CoA 2-initiated reactions.

analyzing the ability of a series of mutants to produce THN (28). It was revealed that although the aromatic amino acid Trp or Phe replacement of Tyr224 is tolerated by Sg-RppA, other Tyr224 mutants including Y224L, Y224G, and Y224C are inactive. In contrast, these mutants are functional when hexanoyl-CoA or phenylacetyl-CoA is presented as a nonphysiological starter unit. On the basis of these observations, Funa *et al.* (28) proposed that Tyr224 in Sg-RppA is critical for starter unit (malonyl-CoA) selection. Interestingly, this Tyr residue is also conserved in DpgA and PhlD (see Supplementary Information, Fig. 9-S1), the other two functionally identified bacterial type III PKSs that utilize malonyl-CoA as the physiological starter unit. The function of Tyr224 in RppA was further discussed by Austin *et al.* (29) upon analyzing the crystal structure of *S. coelicolor* RppA (Sc-RppA), which is the closest relative to *S. griseus* RppA (67% amino acid identity) and catalyzes an identical series of reactions. In the 2.22 Å crystal structure, Tyr224 was found at the surface of the active site pocket of Sc-RppA (8.5 Å away from the catalytic Cys138) and was believed to exert a horizontal steric constraint to the first entering malonyl-CoA through its bulky side chain, thereby

positioning the starter unit at a favorable location (near the catalytic Cys138) by limiting conformational freedom. Accordingly, the small amino acid Gly cannot play a similar role. Thus, when the position corresponding to Tyr224 is occupied by a Gly in CHS (Fig. 9-S1), this plant type III PKS cannot use malonyl-CoA as a starter unit. It is evident that the Gly residue provides additional space for accommodating the bulky substrate p-coumaroyl-CoA (30, 31). This volume-modulating model for substrate specificity is currently well accepted (2, 31). However, it has remained unclear how inactive Tyr224 mutants (Y224L, Y224G, and Y224C) enable selection of unnatural acyl-CoAs as the starter unit substrate in Sg-RppA.

Although Sc-RppA and Sg-RppA are highly homologous and possess nearly identical active site compositions, Sc-RppA was reported to have higher tolerance toward unnatural acyl-CoA starter units than Sg-RppA (25). For example, both acetyl-CoA and benzoyl-CoA were substrates for the Sc-RppA yet were not accepted by Sg-RppA (24). Thus, considering the broader flexibility, together with the corresponding x-ray crystal structure, Sc-RppA represents a compelling target for the study of substrate specificity because of its ability to accept various unnatural substrates, especially acyl-CoAs, which cannot be utilized by Sg-RppA.

In the current study we generated a series of point mutations at the Tyr224 residue in Sc-RppA and compared the resulting product profile to the wild type (WT) enzyme. To elucidate further how this residue determines substrate selectivity, acetoacetyl-CoA, acetyl-CoA, hexanoyl-CoA, and benzoyl-CoA, which, respectively, represent the spectrum of similar, smaller, and larger (aliphatic hexanoyl-CoA and aromatic benzoyl-CoA) substrates compared with malonyl-CoA, were examined as unnatural acyl-CoA starter units by both WT and Tyr224 mutant forms of Sc-RppAs. Moreover, the correlation between substrate loading and product formation of these enzymes was examined by radioactive substrate binding analysis. This investigation confirmed that the inability of Tyr224 mutants to synthesize THN results from failure to load malonyl-CoA onto the catalytic Cys residue as opposed to failure of subsequent chain elongation, cyclization, or product release steps.

9.2 Results

9.2.1 WT and Mutant RppA Overexpression and Purification

The substitution of Leu, Gly, Ala, Ser, Met, His, Cys, and Phe for Tyr224 was conducted to explore the effect of locally modifying steric and electronic/polarity properties on starter unit loading. The WT enzyme and Tyr224 series of RppA mutants were overexpressed in E. coli BL21 (DE3) and purified to homogeneity. As shown by SDS-PAGE (Fig. 9-2A), the monomer of each purified protein migrated with a molecular mass of ~ 43 kDa. Interestingly, previously reported purification efforts of the corresponding Sg-RppA Y224S and Y224H mutants were unsuccessful due to insolubility (28). This discrepancy might result from an intrinsic difference between these two RppAs, the type of cloning system or purification procedure used, or other variables. Nevertheless, each of the Sc-RppA mutant enzymes was expressed as a soluble polypeptide to enable more detailed analysis.

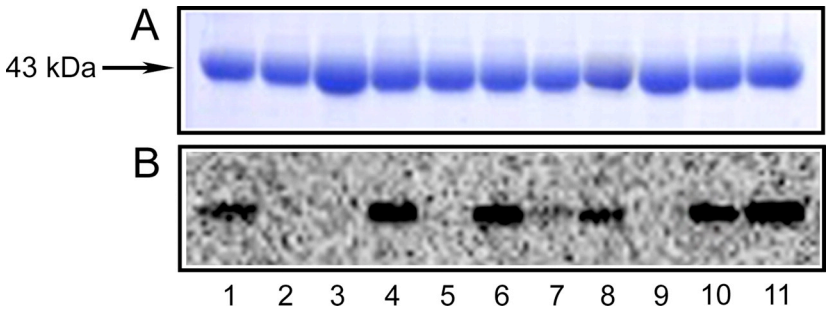


Figure 9-2. Radioactive substrate binding analysis of WT and mutant Sc-RppAs with [2-^{14}C]malonyl-CoA as radioactive label. A, Coomassie bluestained 15% SDS-polyacrylamide gel. B, autoradiograph of the dried SDS-polyacrylamide gel. Lanes: 1, WT Sc-RppA; 2, C138A; 3, C138S; 4, Y224L; 5, Y224G; 6, Y224A; 7, Y224S; 8, Y224M; 9, Y224H; 10, Y224C; 11, Y224F.

9.2.2 Mutant Activity Analysis

Using WT Sc-RppA as a positive control and the catalytic Cys mutants C138A and C138S as negative controls, we first examined the activities of the Tyr224 mutants of Sc-RppA to utilize malonyl-CoA. The reactivity of each mutant was assayed by measuring the amount of accumulated THN by HPLC analysis. Because authentic THN is not isolable due to its rapid conversion to flaviolin (Fig. 9-1A) (19), we elected to deduce the relative yield of THN by integration of the HPLC peak area using WT RppA enzyme as a reference. As shown in Fig. 3A, among the eight Tyr224 mutants, Y224G,

Y224S, and Y224H were unable to catalyze assembly of THN (as negative controls). The inactivity of Y224G is not surprising because Y224G of Sg-RppA was also an inactive mutant (28). For Y224S (whose counterpart in Sg-RppA was insoluble), this inactivation could be attributed to horizontal expansion of the active site through replacement of a smaller side chain. However, it is currently unclear why the RppA Y224H mutant cannot

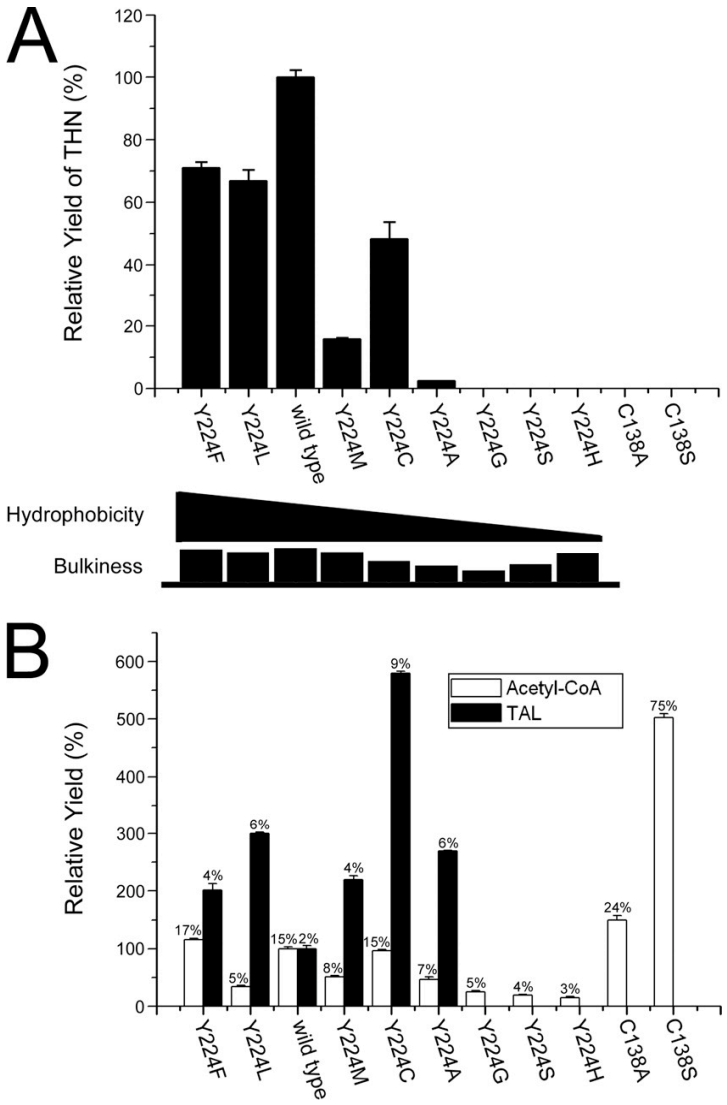


Figure 9-3. The product profile of malonyl-CoA-initiated reactions catalyzed by WT or mutant Sc-RppA. The product profile of Sc-RppA catalyzed reactions was analyzed by HPLC at 280 nm. All reactions were duplicated and the relative yields calculated from the integrated peak area of the HPLC trace using WT enzyme as reference (100%). A, the upper portion shows the THN production profile. The lower portion displays the amino acid hydrophobicity tendency and bulkiness (van der Waals volume). B, profile of the derailment products acetyl-CoA and TAL. The percentage of incorporated malonyl-CoA is shown above each bar.

synthesize THN from malonyl-CoA considering the size of the His side chain. In contrast, Y224F, Y224L, Y224C, Y224M, and Y224A retained 71, 67, 48, 16, and 3% of its activity, respectively, relative to WT Sc-RppA. A previous study reported Y224L, Y224C, and Y224A (but not Y224F) from Sg-RppA to be inactive toward malonyl-CoA (28). To further probe the basis for this discrepancy, we prepared WT Sg-RppA (from *S. griseus* NRRL B-2682) and its corresponding Y224L mutant to compare their ability to generate THN. Surprisingly, this Y224L mutant produced an ~50% level of THN compared with WT Sg-RppA (Fig. 9-S2). Therefore, this (re)examination of Y224L mutants indicates that the Leu replacement retains significant catalytic activity in both Sc-RppA and Sg-RppA.

As described, in addition to the major product THN, a low level of malonyl-CoA would be incorporated into two other derailment products (acetyl-CoA and TAL) because of premature decarboxylation of the oligoketide chain (Fig. 9-1A). Because the synthesis of these by-products might provide new mechanistic insights into this type III PKS, we extended our analysis to the production of acetyl-CoA and TAL (Fig. 9-3B). Specifically, the five partially active mutants (Y224F, Y224L, Y224M, Y224C, and Y224A) accumulated corresponding increased levels of TAL relative to WT Sc-RppA, suggesting the dual function of Tyr224. In addition to the contribution to substrate specificity determination, Tyr224 potentially is able to influence the extent of polyketide chain elongation, thus affecting the partitioning of TAL formation. On the contrary, the mutants that lack an ability to synthesize THN were not able to accumulate TAL. This clearly indicates that the biochemical effect from the Tyr mutation occurs prior to formation of the triketide intermediate. Moreover, no CoASH was released (data not shown) strongly suggesting that these inactive mutants cannot load malonyl-CoA. Significantly, all Tyr224 mutants (together with the negative controls) were capable of producing acetyl-CoA, which indicates that this residue has no significant influence on the decarboxylative activity of Sc-RppA. It was also noteworthy that the decarboxylative activities of both the C138A and C138S mutants were significantly higher than WT enzyme. Approximately, 24 and 75% of the malonyl-CoA substrate was converted to the terminal metabolite acetyl-CoA in the C138A and C138S mutant forms of Sc-RppA, respectively. A similar

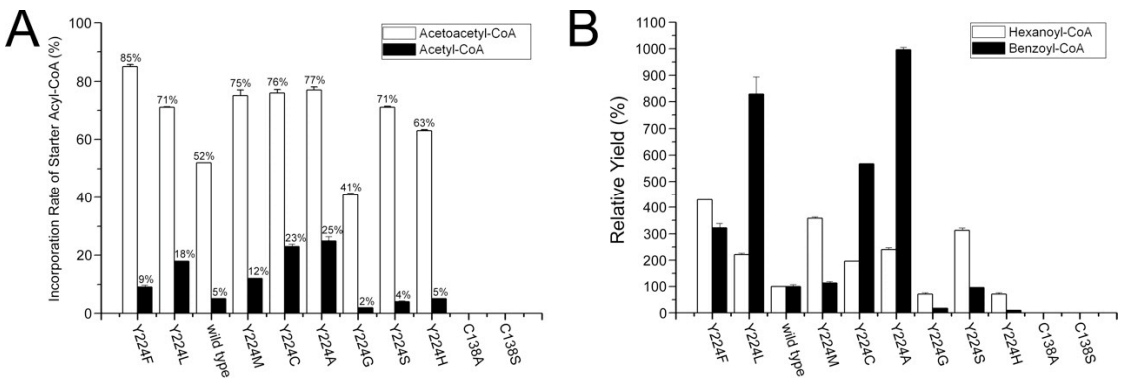


Figure 9-4. The product profile of diverse acyl-CoA-initiated reactions catalyzed by WT or mutant Sc-RppA. A, acetoacetyl-CoA- and acetyl-CoA-initiated reactions. Production of TAL from either acetoacetyl-CoA or acetyl-CoA was monitored at UV280 in duplicate. The numbers above each bar indicate the percentage of incorporated starter unit substrates that were calculated based on absolute yields of TAL. B, hexanoyl-CoA- and benzoyl-CoA-primed reactions. The relative yields of hexanoyl triketide pyrone 3 and benzoyl triketide pyrone 4 were measured at 280 and 320 nm, respectively, by using WT enzyme as reference (100%). All reactions were duplicated.

observation was reported previously in the C138A and C138S mutant forms of DpgA (23).

Next, using acetoacetyl-CoA or acetyl-CoA as the starter substrate and malonyl-CoA as the extender unit, we observed that all Tyr224 mutants (except for the negative controls C138A and C138S) were able to produce TAL from either acetoacetyl-CoA or acetyl-CoA (Fig. 9-4A) via intramolecular lactonization of the triketide linear chain elongation intermediate (Fig. 9-1B). The mutants with partial ability to generate THN produced significantly higher levels of TAL from either acetoacetyl-CoA or acetyl-CoA compared with WT Sc-RppA. We surmise that this was a result of the decreased loading efficiency of malonyl-CoA as a starter unit, resulting in a greater propensity for loading of alternative acyl-CoAs due to increased availability of the active site toward unnatural substrates. Moreover, these partially active Tyr224 mutants still accumulated THN, leading to a production profile similar to that shown in Fig. 9-3A (data not shown), although the corresponding yields were lower because of competition of unnatural acyl-CoA against malonyl-CoA as starter unit. More importantly, the inactive Tyr224 mutants including Y224H, Y224S, and Y224G were capable of accepting acetoacetyl-CoA and acetyl-CoA, resulting in a substantial accumulation of TAL with an incorporation rate ranging from 41 to 71% and 2 to 5% for acetoacetyl-CoA and acetyl-CoA, respectively.

It was evident that the incorporation rate of acetoacetyl-CoA (41–85%) was significantly higher than acetyl-CoA (2–25%). For instance, by calculating the absolute yield of TAL using authentic material as a standard, we determined that 52% of acetoacetyl-CoA was converted to TAL by WT Sc-RppA, whereas only 5% of acetyl-CoA was found to be incorporated as TAL product (similar results were also observed in Sg-RppA (24) and PhlD (32)).

Finally, when we utilized hexanoyl-CoA or benzoyl-CoA as alternate starter units, the corresponding product profiles were comparable with those of acetoacetyl-CoA and acetyl-CoA generated by different Sc-RppAs (Fig. 9-4). Specifically, the products hexanoyl triketide pyrone **3** and benzoyl triketide pyrone **4** of Y224F, Y224L, Y224M, Y224C, and Y224A were observed in significantly higher amounts. With respect to the inactive Tyr224 mutants (Y224H, Y224S, Y224G), we demonstrated that they are also capable of synthesizing the corresponding triketide pyrones from hexanoyl-CoA and benzoyl-CoA, albeit at relatively low levels. In this analysis, however, we were not able to determine the incorporation rates due to lack of standard compounds.

9.2.3 Steady-state Kinetics of RppAs

The HPLC-based product profile analysis only provided qualitative information on the activities of Tyr224 mutants. To better understand the effect of each mutation, steady-state kinetic parameters based on overall THN formation of WT Sc-RppA and the Tyr224 mutants were determined (Table 9-1). The kinetics of WT Sc-RppA (k_{cat} = 1.33 ± 0.05 min^{-1}, K_m = 3.9 ± 0.5 μM, k_{cat}/K_m = 5739 $M^{-1} \cdot s^{-1}$ were comparable with previous reports (20, 29). The mutants Y224F, Y224C, Y224L, and Y224M exhibited an ~2-, 13-, 32-, and 42-fold decrease in the k_{cat}/K_m value, respectively. These results provide direct support for the significance of Tyr224 on the substrate specificity of Sc-RppA. Concomitant with the increasing K_m value, the mutation of this Tyr residue also decreased the turnover number (k_{cat}), implicating its potential influence on the catalytic steps that follow substrate loading. Furthermore, we also determined the kinetic constants of WT Sg-RppA to be k_{cat} = 2.84 ± 0.07 min^{-1}, K_m = 1.1 ± 0.2 μM, k_{cat}/K_m = 44476 $M^{-1} \cdot s^{-1}$, showing that Sg-RppA is more reactive toward malonyl-CoA than is Sc-RppA.

Likewise, the Leu replacement of Tyr224 in Sg-RppA decreased the substrate specificity constant by 148-fold (Table 9-1).

Table 9-1. Steady-state kinetic constants for THN production in WT and Tyr224 mutants of Sc-RppA and Sg-RppA.

	Sc-RppA					Sg-RppA	
	WT	Y224F	Y224C	Y224L	Y224M	WT	Y224L
$k_{cat} \times 10^{-2}$ (min^{-1})	133.4 ± 4.8	110.0 ± 2.4	28.5 ± 0.8	26.4 ± 0.9	13.6 ± 0.7	284.2 ± 6.8	19.9 ± 0.6
K_m (μM)	3.9 ± 0.5	9.5 ± 0.3	12.0 ± 1.2	23.4 ± 2.0	17.1 ± 2.5	1.1 ± 0.2	11.1 ± 1.1
k_{cat}/K_m ($M^{-1}·s^{-1}$)	5,739	1,930	396	175	133	44,476	298

9.2.4 Radioactive Substrate Binding to WT and Mutant Sc-RppAs

All previous data support the hypothesis that Tyr224 in the active site of RppA is involved in substrate specificity determination (28, 29, 31). However, direct evidence that Tyr224 facilitates covalent linkage of the first malonyl-CoA to Cys138 has been lacking. Previous work that revealed the inability of mutant Sg-RppAs to synthesize THN from malonyl-CoA could also result from successful loading but with an inability to catalyze polyketide chain extension, cyclization, or release of final product. To exclude these possibilities, radioactive malonyl-CoA was employed to detect the covalent radiolabel on WT or mutant RppA. Again, the Sc-RppA mutants C138A and C138S were included as negative controls. As expected, WT Sc-RppA and all Tyr224 mutants capable of synthesizing THN (Y224F, Y224L, Y224A, Y224M, and Y224C) were successfully radiolabeled through the covalent acylation of Cys138 with [2-^{14}C]-malonyl-CoA (Fig. 9-2B). Together with negative controls, Y224H, Y224G, and Y224S were not labeled by radioactive malonyl-CoA. This substrate binding profile is consistent with the THN production profile (Fig. 9-3A), demonstrating the correlation between the starter substrate loading and product formation of Tyr224 mutants. Notably, compared with radiolabeled mutants, the radioactive intensity of acylated WT Sc-RppA was significantly lower. This might result from fast turnover of WT enzyme, thus leading to more released intermediates and products, both of which carry radioactivity. Moreover, WT and Y224L mutant Sg-RppA were also labeled by radioactive malonyl-CoA in a similar assay (data not shown).

9.3 Discussion

Recently, the first ligand-free crystal structure of the bacterial type III PKS Sc-RppA was reported (29). However, until co-crystal structures bearing substrate, product, and chain elongation intermediates become available, the answers to three key mechanistic questions will remain outstanding. These include substrate specificity determination, polyketide chain length control, and mode of acyl chain cyclization. Dissection of these parameters currently depends primarily on biochemical analyses and modeling from CHS plant type III PKS. In the current study, we decided to revisit the role of Tyr224 in the active site of Sc-RppA and used site-directed mutagenesis to replace it with a series of alternative amino acid residues to probe its effect on starter unit specificity.

Jez *et al.* (30) observed, via comparison of x-ray crystal structures, that a point mutation at Gly256 (homologous to Tyr224 in RppA) of CHS does not cause significant alteration in the conformation of the polypeptide backbone. Based on the similarity between CHS and RppA, we reasoned that our series of Tyr224 substitutions would change only the local active site properties without inducing conformational changes. Surprisingly, our mutagenesis analysis results were significantly different from previous studies on the highly related type III PKS Sg-RppA (28). A series of Tyr224 mutants of Sc-RppA including Y224F, Y224C, Y224L, Y224M, and Y224A retained the ability to load malonyl-CoA and produce THN, although the catalytic activities were reduced. The current study demonstrates a higher tolerance toward certain amino acid changes at Tyr224 (28). Furthermore, we have demonstrated that the Leu replacement of Tyr224 in Sg-RppA does not eliminate malonyl-CoA starter unit selectivity (and hence THN formation). Interestingly, this positional tolerance was also observed recently in two plant type III PKSs, including *A. arborescens* pentaketide chromone synthase (17) and octaketide synthase (33), both of which harbor a Leu residue at the site corresponding to Tyr224 in RppA (Fig. 9-S1) and utilize malonyl-CoA exclusively to initiate their physiological reactions. This demonstrates that plant type III PKSs employ a different strategy than bacterial enzymes to favor malonyl-CoA as the physiological starter unit. Quantitatively, the reported catalytic efficiencies (k_{cat}/K_m) of pentaketide chromone synthase and octaketide synthase with malonyl-CoA (k_{cat}/K_m = 104 and 16 $M^{-1} \cdot s^{-1}$ for

pentaketide chromone synthase and octaketide synthase, respectively) (17, 33) are dramatically lower than those of the RppAs (k_{cat}/K_m = 5739 and 44476 $M^{-1} \cdot s^{-1}$ for Sc-RppA and Sc-RppA, respectively). The magnitude of the discrepancy suggests that Tyr might be the more highly optimized residue to favor malonyl-CoA with high efficiency. Correspondingly, we hypothesized that the Leu residue in these two plant type III PKSs might represent a compromise between malonyl-CoA incorporative efficiency and product complexity (or diversity), because the presence of a bulky residue at position 224 (such as Tyr) would presumably limit the number of chain extensions, thus impairing chemical diversity (30). In addition, *Rheum palmatum* aloesone synthase (34) and *G. hybrida* 2-pyrone synthase (15, 31) possessing Leu at the homologous position (Fig. 9-S1) can also generate products solely from malonyl-CoA, although their physiological starter unit is acetyl-CoA. This alternate conversion was previously thought to begin by decarboxylation of malonyl-CoA to generate acetyl-CoA (15). The studies reported here suggest that these two enzymes can directly incorporate malonyl-CoA, which competes directly with acetyl-CoA. Furthermore, studies showing that G256L and G256F mutants of CHS are able to produce TAL solely from malonyl-CoA could be interpreted similarly (30).

Prior to the work reported here, the conservation and presumed role of Tyr224 in bacterial type III PKSs indicated its importance in substrate specificity determination (28, 29). However, the hypothesis for its function has been focused on steric factors. The bulky aromatic side chain is believed to be essential for recruiting malonyl-CoA to the thiol group of catalytic Cys through a horizontal steric constriction for the malonyl-CoA starter unit. In this study, we obtained experimental data that could not be explained clearly by this model. First, together with the active Tyr224 mutants of Sc-RppA (Y224F, Y224C, Y224L, Y224M, and Y224A), all mutants (Y224G, Y224S, and Y224H) that were not able to use the physiological starter unit (malonyl-CoA) retained function toward diverse unnatural substrates including acetyl-CoA, acetoacetyl-CoA, hexanoyl-CoA, and benzoyl-CoA. Based on incorporation rates, all of these mutants displayed good activities toward acetoacetyl-CoA. This is not surprising considering that acetoacetyl-CoA is similar in size to malonyl-CoA. However, it remains unclear how Y224G, Y224S, and Y224H mutants are able to differentiate acetoacetyl-CoA from the

isosteric malonyl-CoA. Second, we did not observe a correlation between the THN producing activity and the size of the replaced amino acid (Fig. 9-3A). Unexpectedly, the substitution of a series of nonconservative amino acids with significantly smaller side chains for Tyr224 (Y224C, Y224L, Y224M, and Y224A) was tolerated by Sc-RppA, which contradicts the described volume-modulating model (2, 31) for substrate specificity determination. For instance, in Y224C (one of the active mutants that continues to accommodate malonyl-CoA to produce THN), the small substituent Cys is isosteric to Ser; however the Ser mutation abrogated loading of the physiological starter unit (malonyl-CoA) based on radioactive substrate binding analysis. In contrast, modest modification of the active cavity volume by replacing Tyr224 with His completely disabled the malonyl-CoA loading ability of the mutant type III PKS. These results suggest that there might be an additional type of mechanism other than steric constraints that influences linkage of the first malonyl-CoA to the catalytic Cys. We initially considered that the most significant difference between acetoacetyl-CoA and malonyl-CoA was the terminal electronic property of the latter substrate. Based on observations reported here, it appears that the terminal negative charge of malonyl-CoA could be recognized by RppA. However, the polarity discrepancy is unable to be distinguished by the steric property of Tyr224 or other substituted residues at this position. Therefore, we propose that the hydrophobicity of Tyr224 could contribute to the preference for malonyl-CoA. Compared with other unnatural acyl-CoA substrates, malonyl-CoA carries a terminal negative charge at physiological pH (~7.5). Taking this electronic feature into account, when Tyr224 is replaced by a hydrophobic amino acid such as Leu, Cys, Met, or Ala (although missing the bulky side chain of Tyr), the hydrophobic interaction between the nonpolar side chain and the charged malonyl-CoA could limit the conformational freedom of this starter unit, thus decreasing the entropy penalty of covalent bond formation to active site Cys. However, hydrophobicity does not correlate perfectly with the THN productivity (Fig. 9-3A), indicating that steric factors remain important. Thus, the robustness of WT RppA likely derives from a combination of steric and electronic/polarity factors that favor malonyl-CoA. Additionally, the hydrophobic Y224A replacement generates only a slightly active mutant, possibly because of lack of steric effect to select the substrate. With respect to Y224H, the hydrophilic amino acid

substitution might counteract the steric effect of its bulky side chain. Specifically, it is possible that the protonated His side chain could form an ionic interaction with the terminal negative charge of malonyl-CoA to prevent it from being linked to Cys138.

Considering the current data, we propose that the ability of the malonyl-CoA-inactive mutants to accept unnatural acyl-CoAs that lack a terminal charge is governed by the hydrophobic interior of enzyme and the thermodynamically unfavorable conditions for delivering a formal charge there. Accordingly, an entropy compensation for substrate binding through limiting the conformational freedom of the substrate might be more stringent for the terminally charged malonyl-CoA than for other neutral unnatural acyl-CoAs.

9.4 Summary

In summary, we have demonstrated a surprising level of substrate tolerance involving a series of RppA Tyr224 mutations including both aromatic and hydrophobic amino acids with medium or large side chains. Both steric and electronic constraints exerted by Tyr224 in the active site are involved in facilitating the loading of the physiological substrate malonyl-CoA. In practical terms, re-engineering of plant and bacterial type III PKS to increase the preference for malonyl-CoA could be utilized to enable further product diversification. The identification of a number of active Tyr224 mutants with lower turnover numbers could provide potential candidates for subtle presteady-state kinetic studies, enzyme-linked intermediate trapping, or co-crystallization efforts. Moreover, the Tyr224 mutants that preferentially accept unnatural acyl-CoAs instead of malonyl-CoA could be applied for the production of novel pyrone products with higher yields.

9.5 Experimental Procedures

***S. coelicolor* RppA Gene Cloning** — Genomic DNA was prepared from *S. coelicolor* A3 (2) using a Wizard genomic DNA purification kit (Promega). The Sc-RppA gene was amplified by PCR under standard conditions with a pair of primers as follows: forward, 5'-AATCATATGGCGACTTTGTGC-3' (the underlined bases represent the introduced *Nde*I site for further cloning); reverse, 5'-

AATAAGCTTTCATGCCTGCCTGCCTCACCC-3' (the underlined letters indicate a *Hind*III restriction site for later cloning). The full-length cDNA encoding Sc-RppA was rescued by double digestion of *Nde*I and *Hind*III. The fragment containing Sc-RppA gene was ligated into previously *Nde*I/*Hind*III-digested pET-28b(+) (Novagen) to generate the recombinant plasmid pET-28b(+)-Sc-RppA. This construct was then transformed into *Escherichia coli* BL21 (DE3) for protein overexpression. The accuracy of the inserted Sc-RppA gene was confirmed by nucleotide sequencing.

Site-directed Mutagenesis, Protein Overexpression, and Purification — Site-directed mutagenesis was performed according to the QuikChange (Stratagene) protocol with pET-28b(+)-Sc-RppA as template. The PCR-amplified plasmids containing mutations were verified by DNA sequencing and then directly used for transforming the protein overexpression strain *E. coli* BL21 (DE3). The resulting transformants were grown at 37 °C overnight in Luria broth containing 50 μg/ml kanamycin, and 10 ml of culture was used to inoculate 1 liter of Terrific Broth containing 50 μg/ml of kanamycin. The Terrific Broth culture was incubated at 37 °C for 3~4 h until A_{600} reached 0.6–1.0. Then, gene overexpression was induced by adding isopropyl-β-$_D$-thiogalactopyranoside to a final concentration of 0.1 mM and cultured at 18 °C overnight. The culture was centrifuged at 4,000 g for 10 min to collect cells. The freeze-thaw cell pellet was resuspended in 40 ml of lysis buffer (250 mM NaCl, 50 mM NaH_2PO_4, 10 mM imidazole, 20% glycerol, pH 7.9), and the cell lysate was prepared by lysozyme digestion at 1 mg/ml for 30 min with subsequent sonication. Cell debris was removed by centrifugation at 15,000 g for 30 min, and the soluble fraction was mixed with 2 ml nickel-nitrilotriacetic acid-agarose (Qiagen) for 1 h at 4°C. The slurry was loaded onto an empty column. The column was then washed stepwise with 20 ml of lysis buffer and 40–60 ml of wash buffer (250 mM NaCl, 50 mM NaH_2PO_4, 40 mM imidazole, 20% glycerol, pH 7.9). The His_6 tag bound proteins were eluted with elution buffer (250 mM NaCl, 50 mM NaH_2PO_4, 200 mM imidazole, 20% glycerol, pH 7.9). The target protein was further purified and concentrated with YM-10 centrifuge filters (Amicon). Finally, the elution buffer was exchanged for storage buffer (250 mM NaCl, 50 mM NaH_2PO_4, 20% glycerol, pH 7.4) with a PD-10 column (GE Healthcare), and the resulting product concentrations were measured by Bio-Rad

protein assay kit using bovine serum albumin as standard. Purity and molecular mass were examined by standard SDS-PAGE analysis.

RppA Enzymatic Assay — The standard assay containing 50 mM sodium phosphate (pH 7.4), 0.2 mM acyl-CoA starter, and 3 μM WT or mutant RppA in a total volume of 100 μl was initiated by adding malonyl-CoA to a final concentration of 0.5 mM and was quenched with the addition of 10 μl of 6 M HCl after incubation at 30 °C for 1 h. Small amounts of enzyme were removed by centrifugation at 16,100 g for 10 min. The supernatant was used directly as the sample for HPLC analysis. Reverse-phase HPLC conditions were as follows, using an XBridge C18 5-μm column (4.6 x 250 mm): the mobile phase for the malonyl-CoA-, acetyl-CoA-, and acetoacetyl-CoA-primed reaction was linear from 5 to 35% B (A = deionized water + 0.1% trifluoroacetic acid; B = acetonitrile + 0.1% trifluoroacetic acid) over 20 min, and the mobile phase for the hexanoyl-CoA- and benzoyl-CoA-primed reaction was linear from 20 to 70% B over 20 min; the UV detection wavelength was 280 nm for the malonyl-CoA-, acetyl-CoA-, acetoacetyl-CoA-, and hexanoyl-CoA-primed reactions and 320 nm for the benzoyl-CoA primed reaction, at a flow rate of 0.8 ml/min. The peak identity in HPLC trace was determined by mass spectrometry and comparison with authentic compounds regarding HPLC retention time and UV spectrum.

Enzyme Kinetics — We exploited the reported coefficient constant for THN at 340 nm (ε = 18,000) (20) to obtain the steadystate kinetic parameters using a spectrophotometric method. The standard reaction buffered with 100 mM Tris-HCl (pH 7.5) containing the appropriate concentration of WT or mutant RppA in a total volume of 400 μl was initiated by adding a series of concentrations of malonyl-CoA ranging from 1 to 100 μM and was immediately monitored with a UV-visible spectrophotometer 300 Bio (Cary) at 30 °C. The initial velocity of THN accumulation was deduced from the absorbance curve within the linear range. All measurements were determined in triplicate, and velocities determined under different malonyl-CoA concentrations were fit to the Michaelis-Menten equation to calculate the kinetic constants k_{cat} and K_m.

Radioactive Substrate Binding Assay — The optimized reaction contained 50 mM HEPES buffer (pH 7.4), 50 μM [2-^{14}C]malonyl-CoA (55 mCi/mmol) and 15.0 μM WT or mutant enzymes in a total volume of 25 μl. Reactions were quenched by the addition of 25 μl of 2x SDS gel loading dye after a 5-min incubation at 30 °C. 10 μl of each reaction mixture was analyzed by SDS-PAGE (15%). After staining, destaining, and drying the protein gel on filter papers, the radioactivity bound to the protein was determined by autoradiography following a 4-day exposure using a PhosphorImager (Amersham Biosciences).

9.6 Supplementary Information

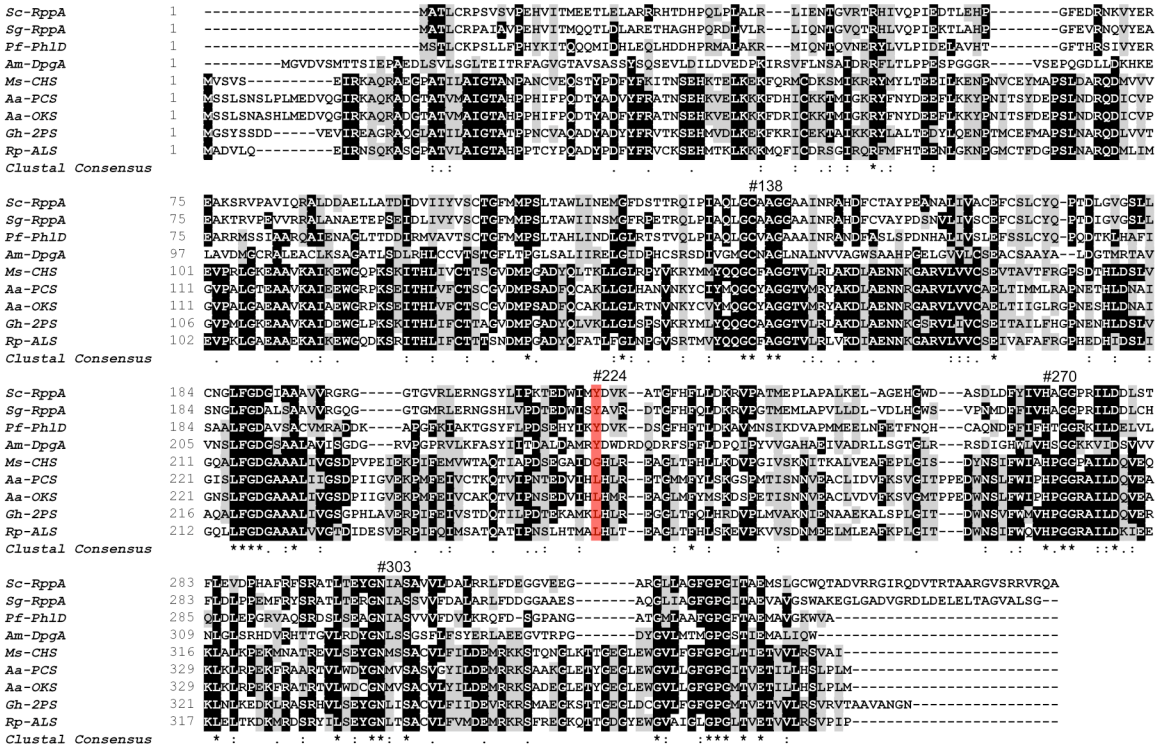


Figure 9-S1. Multiple alignment of amino acid sequences of heterologous type III PKSs. (Sc-RppA: *Streptomyces coelicolor* RppA; Sg-RppA: *Streptomyces griseus* RppA; Pf-PhlD: *Pseudomonas fluorescens* PhlD; Am-DpgA: *Amycolatopsis mediterranei* DpgA; Ms-CHS: *Medicago sativa* CHS; Aa-PCS: *Aloe arborescens* PCS; Aa-OKS: *Aloe arborescens* OKS; Gh-2PS: *Gerbera hybrida* 2-PS; Rp-ALS: *Rheum palmatum* ALS). The Tyr224 and catalytic triad Cys138, His270, and Asn303 are marked with # (numbering in Sc-RppA). The amino acids homologous to Tyr224 of Sc-RppA in other type III PKSs are highlighted within a red box.

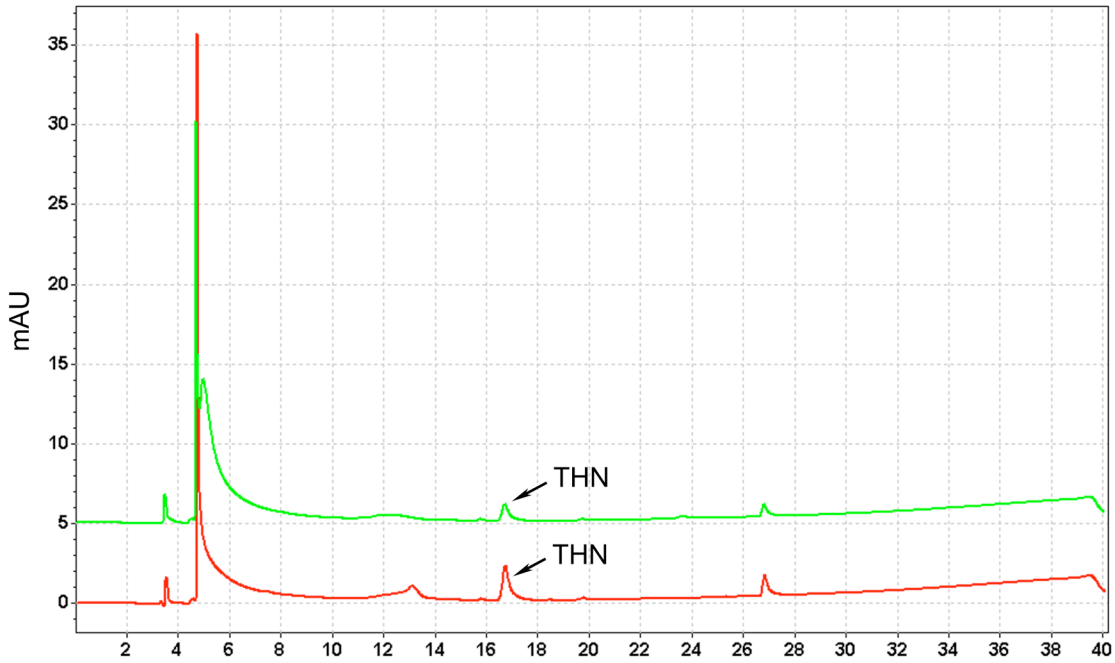


Figure 9-S2. HPLC analysis (UV_{280}) of THN production catalyzed by wt and Y224L mutant Sg-RppA. The reactions containing 200 µM malonyl-CoA and 1 µM enzyme in 100 µL of 100 mM Tris-HCl (pH 7.5) were incubated for 1 hr and quenched by adding 10µL of 6M HCl. HPLC gradient conditions: linear from 10% to 100% solvent B (acetonitrile + 0.1% trifluoroacetic acid) in solvent A (deionized water + 0.1% trifluoroacetic acid) over 30 min, then 100% solvent B for 10 min. The flow rate was 0.8 mL/min. The HPLC traces of wt Sg-RppA and Y224L mutant catalyzed reactions are colored in red and green, respectively. The peaks corresponding to THN product are labeled.

9.7 References

1. Moore, B. S., and Hopke, J. N. (2001) *Chembiochem* **2**, 35-38
2. Austin, M. B., and Noel, J. P. (2003) *Nat. Prod. Rep.* **20**, 79-110
3. Rawlings, B. J. (2000) *Nat. Prod. Rep.* **18**, 190-227
4. Katz, L., and Donadio, S. (1993) *Annu. Rev. Microbiol.* **47**, 875-912
5. Hutchinson, C. R., and Fujii, I. (1995) *Annu. Rev. Microbiol.* **49**, 201-238
6. Staunton, J., and Weissman, K. J. (2001) *Nat. Prod. Rep.* **18**, 380-416
7. Hahlbrock, K., and Kreuzaler, F. (1972) *Hoppe Seyler's Z. Physiol. Chem.* **353**, 1522-1526
8. Akiyama, T., Shibuya, M., Liu, H.-M., and Ebizuka Y. (1999) *Eur. J. Biochem.* **263**, 834-839
9. Schoppner, A., and Kindl, H. (1984) *J. Biol. Chem.* **259**, 6806-6811
10. Helariutta, Y., Elomaa, P., Kotilainen, M., Griesbach, R. J., Schröder, J., and Teeri, T. H. (1995) *Plant Mol. Biol.* **28**, 47-60
11. Li, T.-L., Choroba, O. W., Hong, H., Williams, D. H., and Spencer, J. B. (2001) *Chem. Commun. (Camb.)* **20**, 2156-2157
12. Sankaranarayanan, R., Saxena, P., Marathe, U. B., Gokhale, R. S., Shanmugam, V. M., and Rukmini, R. (2004) *Nat. Struct. Mol. Biol.* **11**, 894-900
13. Bangera, M. G., and Thomashow, L. S. (1999) *J. Bacteriol.* **181**, 3155-3163
14. Seshimea, Y., Juvvadia, P. R., Fujiib, I., and Kitamoto, K. (2005) *Biochem. Biophys. Res. Commun.* **331**, 253-260
15. Eckermann, S., Schröder, G., Schmidt, J., Strack, D., Edrada, R. A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpeläinen, I., Proksch, P., Teeri, T. H., and Schröder, J. (1998) *Nature* **396**, 387-390
16. Heller, W., and Hahlbrock, K. (1980) *Arch. Biochem. Biophys.* **200**, 617-619
17. Abe, I., Utsumi, Y., Oguro, S., Morita, H., Sano, Y., and Noguchi, H. (2005) *J. Am. Chem. Soc.* **127**, 1362-1363
18. Paniego, N. B., Zuurbier, K. W., Fung, S.-Y., Van der Heijden, R., Scheffer, J. J., and Verpoorte, R. (1999) *Eur. J. Biochem.* **262**, 612-616
19. Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y., and Horinouchi, S. (1999) *Nature* **400**, 897-899
20. Izumikawa, M., Shipley, P. R., Hopke, J. N., O'Hare, T., Xiang, L., Noel, J. P., and Moore, B. S. (2003) *J. Ind. Microbiol. Biotechnol.* **30**, 510-515
21. Achkar, J., Xian, M., Zhao, H., and Frost, J. W. (2005) *J. Am. Chem. Soc.* **127**, 5332-5333
22. Pfeifer, V., Nicholson, G. J., Ries, J., Recktenwald, J., Schefer, A. B., Shawky, R., Schroder, J., Wohlleben, W., and Pelzer, S. (2001) *J. Biol. Chem.* **276**, 38370-38377
23. Tseng, C. C., McLoughlin, S. M., Kelleher, N. L., and Walsh, C. T. (2004) *Biochemistry* **43**, 970-980
24. Funa, N., Ohnishi, Y., Ebizuka, Y., and Horinouchi, S. (2002) *J. Biol. Chem.* **277**, 4628-4635
25. Jeong, J.-C., Srinivasan, A., Grüschow, S., Bach, H., Sherman, D. H., and Dordick, J. S. (2005) *J. Am. Chem. Soc.* **127**, 64-65
26. Gross, F., Luniak, N., Perlova, O., Gaitatzis, N., Jenke-Kodama, H., Gerth, K., Daniela, G., Dittmann, E., and Müller, R. (2006) *Arch. Microbiol.* **185**, 28-38
27. Gerth, K., Pradella, S., Perlova, O., Beyer, S., and Müller, R. (2003) *J. Biotechnol.* **106**, 233-253
28. Funa, N., Ohnishi, Y., Ebizuka, Y., and Horinouchi, S. (2002) *Biochem. J.* **367**, 781-789
29. Austin, M. B., Izumikawa, M., Bowman, M. E., Udwary, D. W., Ferrer, J.-L., Moore, B. S., and Noel, J. P. (2004) *J. Biol. Chem.* **279**, 45162-45174

30. Jez, J. M., Bowman, M. E., and Noel, J. P. (2001) *Biochemistry* **40**, 14829-14838
31. Jez, J. M., Austin, M. B., Ferrer, J.-L., Bowman, M. E., Schroder, J., and Noel, J. P. (2000) *Chem. Biol.* **7**, 919-930
32. Zha, W., Rubin-Pitel, S. B., and Zhao, H. (2006) *J. Biol. Chem.* **281**, 32036-32047
33. Abe, I., Oguro, S., Utsumi, Y., Oguro, S., Sano, Y., and Noguchi, H. (2005) *J. Am. Chem. Soc.* **127**, 12709-12716
34. Abe, I., Utsumi, Y., Oguro, S., and Noguchi, H. (2004) *FEBS Lett.* **562**, 171-176

Notes:

This work has been published as "Molecular analysis of the role of tyrosine 224 in the active site of *Streptomyces coelicolor* RppA, a bacterial type III polyketide synthase." **Li, S.**, Grüschow, S., Dordick, J. S., and Sherman, D. H. ***J. Biol. Chem.*** 2007, 282 (17), 12765-12772.

Author contributions:
Shengying Li, Sabine Grüschow, Jonathan J. Dordick, and David H. Sherman designed the experiments;
Shengying Li performed the experiments.

Chapter 10
Future Work

In this Ph.D. dissertational study, I have been focused on the biochemical, structural, and bioengineering investigations of PikC, the cytochrome P450 enzymes with innate substrate flexibility and hydroxylation versatility from the macrolide antibiotic pikromycin biosynthetic pathway in *Streptomyce venezuelae*. In collaboration with the crystallographic group led by Dr. Larissa M. Podust and the synthetic group led by John Montgomery, the structural basis for the substrate flexibility of PikC has been elucidated; a more catalytically active mutant $PikC_{D50N}$ was identified and the mechanism for the improved activity has been understood via analysis of its crystal structure; a self-sufficient version of PikC enzyme was created by fusing this P450 enzyme to a heterologous reductase domain RhFRED; and the unique desosamine anchoring functionality observed in the substrate-enzyme co-crystal structures has been harnessed to achieve regioselective oxidation of a series of carbocyclic rings linked to the desosamine glycoside by using the engineered $PikC_{D50N}$-RhFRED. All these results indicate that PikC is a good model to help answer some fundamental questions of P450 catalysis and $PikC_{D50N}$-RhFRED possesses great potential to be developed into an applicable catalyst in synthetic chemistry. However, this study also leaves a number of unsolved questions awaiting for future investigations.

First, the observation of two different substrate binding modes clearly provides the structural basis for the substrate flexibility of PikC. However, the observed position and orientation of substrates in the active site does not well explain the C10 hydroxylation of YC-17 and the C12 hydroxylation of narbomycin since these two sites are substantially distant from the heme iron reactive center, leaving the possibilities including: 1) the substrate could be repositioned upon the dioxygen binding to the heme

iron during the catalytic cycle; 2) the genuine oxidative species in the C10 hydroxylation of YC-17 and the C12 hydroxylation of narbomycin might be the more extended ferric hydroperoxy species (Fe^{3+}–OOH, Compound 0) other than the ferryl-oxo species (Fe^{4+}=O, Compound I) proposed to be responsible for the C-12 hydroxylation of YC-17. There will be several options to test these hypotheses. First, we could investigate the hydroxylation pattern of YC-17 and narbomycin by using synthetic metalloporphyrins at different oxidative stages to mimic the chemistry of P450 "Compound 0" and "Compound I". Second, we could attempt to solve the crystal structure of the ferrous dioxygen adduct of PikC with substrate bound to see if the substrate repositioning occurs. Third, we may select a mutant PikC that displays a different hydroxylation pattern toward either YC-17 or narbomycin through random mutagenesis, and then compare its structure to that of wild type PikC.

Second, taking advantage of the unique desosamine anchoring functionality in the substrate binding process of PikC, considerable regioselectivity in oxidation of a group of desosaminyl derivatives has been achieved by using the engineered $PikC_{D50N}$-RhFRED, which is approximately 13-fold more active than wild type PikC. However, the stereoselective control was not successful. This will require more future work focusing on protein engineering such as directed evolution of $PikC_{D50N}$-RhFRED regarding its stereoselectivity against certain substrates, and/or substrate structural refinement.

Third, the presence of desosamine (or desosamine analogs) in diverse secondary metabolite substrates of different biosynthetic P450 enzymes suggest that the desosamine anchoring functionality might have been utilized extensively in nature. For example, in addition to PikC, MycCI, TylHI, and EryK, which are cytochrome P450s involved in tailoring of mycinamicin, tylosin, and erythromycin, are capable of hydroxylating different desosamine (or mycaminose, a desosamine analog, in the case of tylosin) containing substrates (mycinamicin VIII, 23-deoxy-*O*-mycaminosyltylonolide, and erythromycin D) at sites remote from the sugar linkage. These findings suggest that more natural product P450s could be further harnessed for a similar substrate engineering approach as we have done for PikC to selectively oxidize C-H bonds, and to facilitate further chemical diversification of both synthetic and natural product molecules of biological interest.

Finally, the knowledge and experience accumulated in the PikC project could be applied to more functional analysis of biosynthetic P450 enzymes in future, which is likely to result in new biochemistry, more insights into biosynthesis of natural products and biodegradation, more candidates for application in biotechnology, and more facile models for understanding CYP structures and catalytic mechanisms.