UNIVERSITY OF CALIFORNIA,
IRVINE

Structural and Functional Studies of Enzymes from Non-Reducing Iterative Type I and Type II Polyketide Synthases Using Atom-Replaced Probes

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Gabriel Otis Moreno

Dissertation Committee:
Professor Shiou-Chuan (Sheryl) Tsai, Chair
Professor Thomas L. Poulos
Professor Celia W. Goulding

2018

DEDICATION

I dedicate this dissertation to my wife, Caroline.

We did it Preciousa!

It has not been easy, but we both have grown so much through these years of graduate school. Thank you for your faithful love and for not giving up on me.

I dedicate this dissertation to my children, Esperanza, Nehemiah and Trustin. It has been amazing to watch you grow over these years and you have filled them with joy.

I dedicate this dissertation to all of our friends and family. You have always been there when we needed you. Your prayers and love have brought us here.

Lastly,

I dedicate this dissertation, in loving memory, to my wonderful papa

Fernando Castillo Moreno

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Professor Tsai, you have my sincerest gratitude—without you, this dissertation would not have been possible. Thank you for believing in me when I did not. Thank you Professors Poulos and Goulding for your support over these years of graduate school. Professor Goulding, I have always appreciated your glowing scientific enthusiasm. Professor Poulos, your undying passion of scientific inquiry continues to inspire me.

I would also like to thank all of the members of the Tsai Lab, both past and present. You are all the most amazing people I have ever met. I am honored to have been your colleague and friend. In particular, I would like to thank Dr. Jackson for your mentorship when I first joined the Tsai Lab. Your fearless scientific spirit still makes me smile. Dr. Barajas, your unfailing, humble and helpful spirit has left a lasting impression on me. Drs. Shakya, Nguyen, Festin, Bruegger, Valentic, Milligan, White, and soon to be doctor, Bryan Ellis, it was awesome working and hanging out with each of you. Dr. Schaub, your friendship has been invaluable.

I would also like to thank the undergraduates with whom I have had the pleasure to work—Wilson Bahn, Betty Lai and Hau Truong. Thank you for your dedication. Betty was instrumental in getting the DpsKR $C222_1$ crystal structure. Hau led the optimization of RemI and EncD crystals, bringing us to the verge of solving their structures. Hau, it has been wonderful working with you and I thank you for everything. I am so excited that you are going on to graduate school and continuing research in structural biology!

I would also like to thank all of the members of our neighboring labs: the Goulding, Poulos, Gross, Spitale and Bardwell Labs. Everyone is always ready to help. In particular, I would like to thank Alex Chao, Alec Follmer and Dail Chapman. I have truly enjoyed our conversations together.

In addition, I would also like to thank the National Science Foundation and the Minority Biomedical Research Program for supporting my research. In particular, I would like to thank Drs. Mota-Bravo and de la Cruz, for their caring guidance.

Lastly, I would like to thank God. Though I haven't acknowledged you very much in my life, I am convinced that the consummate beauty and design of life cannot have brought itself to be. Nothing happens by accident and so I thank you for your grace on my life.

CURRICULUM VITAE

EDUCATION

University of California, Irvine; Irvine, CA — September 2018
Ph.D., Biological Sciences

Northern Arizona University; Flagstaff, AZ — May 2008
B.S., Microbiology, Minor in Chemistry

RESEARCH EXPERIENCE

University of California, Irvine; Irvine, CA — March 2012 – September 2018
Graduate Student Researcher; PI – Professor Shiou-Chuan (Sheryl) Tsai, Ph.D.
Polyketide synthase structural biology

University of California, Irvine; Irvine, CA — January 2012 – March 2012
Graduate Student Researcher; PI – Professor Hartmut Lueke
Monomerization of the H. pylori membrane urea channel HpUreI

University of California, Irvine; Irvine, CA — September 2011 – December 2011
Graduate Student Researcher; PI – Professor Melanie Cocco
Elucidation of the structural determinates of membrane association by the neuronal outgrowth inhibitor Nogo

Northern Arizona University; Flagstaff, AZ — January 2005 – May 2008
Undergraduate Student Researcher; PI – Professor Jeffery Leid
Resistance mechanisms of biofilms against host-immune system defenses

TEACHING EXPERIENCE

University of California, Irvine; Irvine, CA — March 2013 – June 2018
Graduate Teaching Assistant
Courses: Biochemistry lecture (3x), Advanced Biochemistry lecture, Biochemistry lab (3x), Molecular Biology lecture (5x), Molecular Biology lab, Protein Structure and Function (graduate course), Microbiology lab, Pharmacology lab, General Chemistry lab (2x)

Northern Arizona University; Flagstaff, AZ — January 2008 – May 2008
Undergraduate Teaching Assistant
Microbiology lab

U.S. Army, Joint Language Center; Ft. Meade, MD — November 2003 – August 2004
Chief Military Language Instructor
Mandarin Chinese

PUBLICATIONS

Moreno, G.O., Jackson, D.R., Lai, H.Y., Crosby, J., Tsai, S.C. Chimera sheds light on substrate specificity of the daunorubicin C9 ketoreductase, DpsE. *Manuscript in preparation.*

Moreno, G.O., Bruegger, J.B., Barajas, J.F., Shakya, G., Burkart, M.D., Tsai, S.C. Probing and engineering the stereospecificity of ketoreductase using polyketide mimics. *Manuscript in preparation.*

Alhoshani, A., Vithayathil, R., Bandong, J., Chrunyk, K.M., **Moreno, G.O.**, Weiss, G.A., Cocco, M.J. Glutamate provides a key structural contact between reticulon-4 (Nogo-66) and phosphocholine. BBA Biomembranes 1838(9), 2350-56.

Leid, J.G., Kerr, M., Selgado, C., Johnson, C., **Moreno, G.O.**, Smith, A., Cope, E.K. Flagellum-mediated biofilm defense mechanisms of Pseudomonas aeruginosa against host-derived lactoferrin. Infection and Immunity 77(10), 4559-66.

HONORS

National Science Foundation Graduate Researcher Fellow	July 2013 – July 2018
Minority Biomedical Research Support Fellow	June 2012 – June 2013
Graduated summa cum laude, Northern Arizona University	May 2008
Minority Access to Research Careers Fellow	January 2006 – May 2008

CERTIFICATES

Data Science Certificate Program, UCI	2018

ABSTRACT OF THE DISSERTATION

Structural and Functional Studies of Enzymes from Non-Reducing Iterative Type I and Type II Polyketide Synthases Using Atom-Replaced Probes

By

Gabriel Otis Moreno

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2018

Professor Shiou-Chuan (Sheryl) Tsai, Chair

Secondary metabolites from plants, fungi and bacteria offer abundant and varied complex bioactive chemicals. Among these, polyketides have become one of the most important sources of therapeutic compounds, including antibiotics (such as erythromycin and tetracycline), anticancers (such as doxorubicin and epothilone B), immunosuppressants (rapamycin), anticholesterols (such as lovastatin), and antioxidants (such as resveratrol). The polyketides derive their potency from their considerable structural-, stereo- and functional group complexity. This complexity is wrought into the molecules through the cooperation of multiple enzymatic domains functioning together in a highly orchestrated assembly line. The medical value of polyketides has drawn much attention to the research of polyketide synthases (PKSs) with the objective of engineering these systems to produce tailor-made compounds. Researchers have met with many successes and more failures in their endeavors to exert complete control over the biosynthetic outcome of these pathways. The chief limitation is that the sequence-structure-function relationship of proteins is still poorly understood. This is true at all levels, including protein folding, substrate specificity, catalytic activity, and protein-protein interactions.

The subjects of this dissertation are enzymes from a group of PKSs called iterative PKSs, which produce aromatic polyketides such as bikaverin, aflatoxin, doxorubicin and actinorhodin. There are two types of polyketide synthases that fall within this group: the fungal iterative type I non-reducing PKS (NR-PKS) and the bacterial type II PKS. Among PKSs, these two types remain the least understood, due to the instability of the intermediates produced along their biosynthetic pathway. Unlike other types of PKSs, NR-PKSs generate a long carbon chain with ketone groups present on every other carbon. This arrangement is highly prone to undergo intra- and inter-molecular aldol/Claisen condensations, which makes it extremely difficult to study the native enzyme-substrate interactions in these systems. In this thesis, a suite of novel polyketone mimetics that maximally mimic the steric and electron structure of the unstable natural intermediate while being stable enough to be used as a probe in experimental studies are used to reveal the details of enzyme-substrate interactions of the product template domain from fungal NR-PKS and the reductase enzyme from bacterial type II PKS.

CHAPTER 1

Introduction

1.1 Polyketide natural products

Polyketides, a large family of structurally complex natural products, include compounds with important bioactivities (1). As one of the largest classes of bioactive natural products, polyketides exhibit a striking potential for drug discovery (> 0.3 % of current known polyketides) compared to other natural product or synthetic compounds (< 0.1 %) (2). For example, polyketides include antibiotics (such as erythromycin (3), actinorhodin (4), and tetracenomycin (5)), chemotherapeutics (such as doxorubicin (6), resistomycin (7), and mithramycin (8)), cholesterol lowering drugs (such as lovastatin (9)), antioxidants (such as resveratrol (10)), as well as toxins

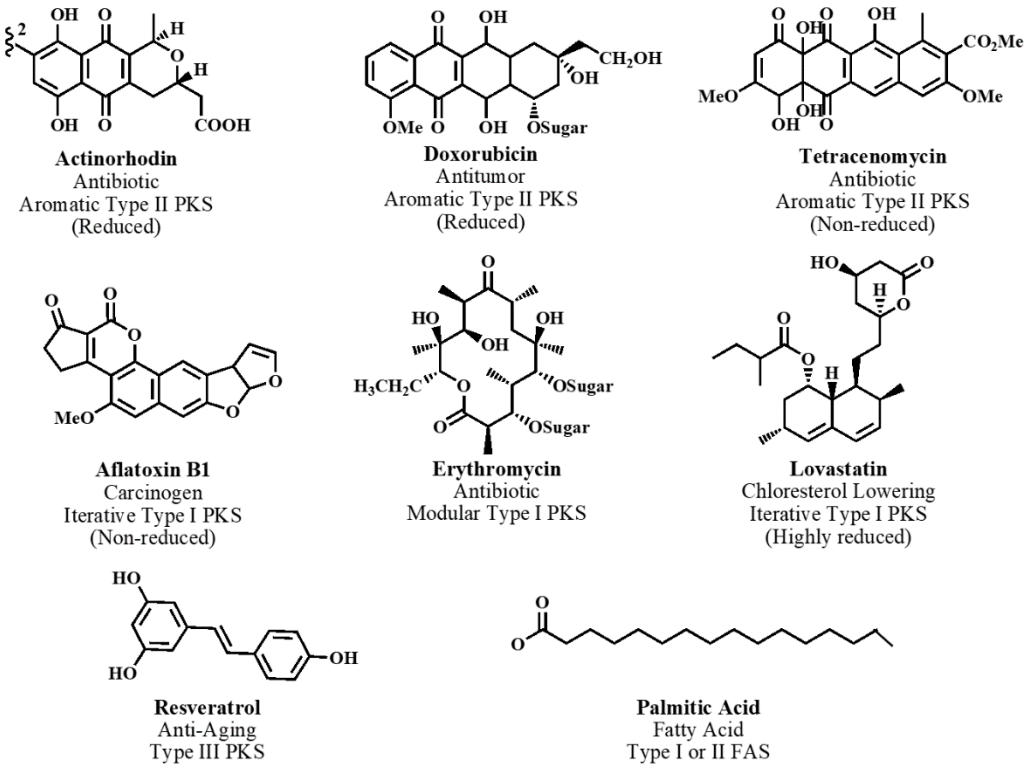


Figure 1.1 Examples of natural products from polyketide synthases and related fatty acid synthase.

(T-toxin) and carcinogens (aflatoxin B1 (11)) (Figure 1.1). Feeding experiments also identified that the polyketide building blocks are small acyl-CoAs such as acetyl- and propionyl-CoAs, and showed that polyketide and fatty acid biosyntheses are closely related (1, 12). At the center of polyketide biosynthesis is the polyketide synthase (PKS), a multi-domain enzyme complex that produces a huge variety of polyketides via a controlled variation of building blocks, chain lengths, chain reduction and cyclization. Considering the high sequence homology between PKS and fatty acid synthase (FAS), a fascinating question is how PKS and FAS exhibit such different regio-specificities, as reflected from the vastly different products these two mega-synthases yield.

The past two decades have witnessed significant advances in PKS structural biology for different types of PKSs that help visualize polyketide biosynthesis at all stages including chain initiation, elongation, reduction, cyclization, and chain termination. Using X-ray crystallography or nuclear magnetic resonance (NMR), these studies help correlate the PKS three-dimensional structures with observed substrate specificity. These studies also help in elucidating the sequence-function-structure relationship that predicts product outcome. This chapter reviews PKS structural biology, and provides a comparison between FAS and PKS, as well as enzyme domains across different types of PKSs and FASs.

1.2 Fatty acid synthase (FAS)

Fatty acid biosynthesis is catalyzed by a multidomain protein complex called fatty acid synthase (FAS) via the decarboxylative condensation of malonyl building blocks to produce a long chain fatty acid, typically 16-18 carbons in length (13). The FAS consists of 7 conserved protein domains that catalyze > 50 reactions en route to the final fatty acid product (Figure 1.2). Starting

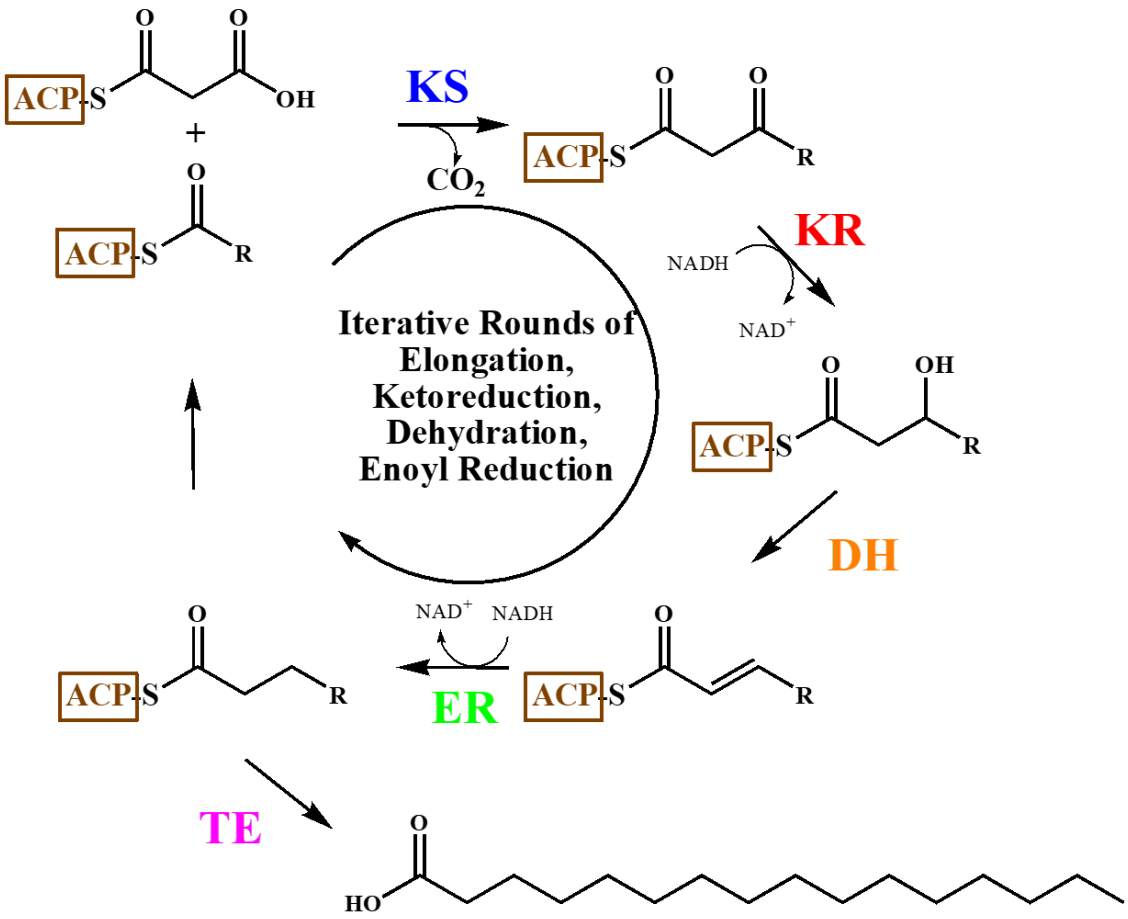


Figure 1.2 Fatty acid biosynthesis. Fatty acid biosynthesis is an iterative process that proceeds via decarboxylative condensations between malonyl-ACP and acyl-ACP catalyzed by the KS domain. Reductions and dehydration at each cycle produce an aliphatic fatty acid released by the TE domain.

with malonyl-CoA as the basic building block, malonyl-ACP transacetylase (MAT) primes the acyl-carrier protein (ACP) by transferring malonate from malonyl-CoA to the phosphopantetheine (PPT) group of ACP. The PPT is covalently linked to a highly conserved serine of ACP. Next, acetyl-CoA is transacylated to the ketosynthase (KS) domain to form acetyl-KS. This is followed by KS-catalyzed decarboxylative condensation between the KS-bound acyl group and malonyl-ACP, in which decarboxylation is a strong driving force for C-C bond formation. Following each elongation step, the ACP-bound chain is further reduced by NADPH-dependent ketoreductase (KR), dehydrated by the dehydratase (DH), and reduced by NADPH-

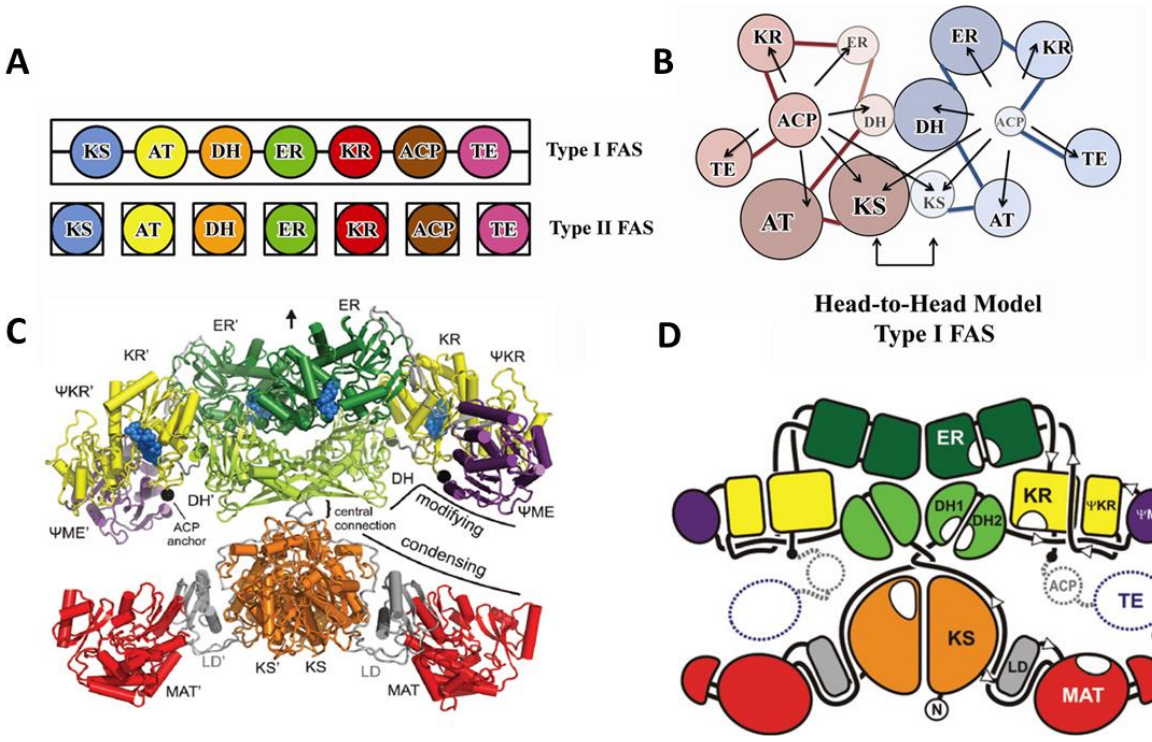


Figure 1.3 Structural overview of FAS. **(A)** Domain organization of Type I and Type II FAS. **(B)** The head-to-head model for Type I FAS complex formation. The central core is held together by the KS-KS dimer (head-to-head). Arrows highlight the ability of ACP to interact, not only with all domains from its own monomer, but also the KS from the other monomer. **(C)** Cartoon ribbon representation and **(D)** Domain illustration of the 3.2 Å porcine Type I FAS structure.

dependent enoyl reductase (ER) to afford a fully reduced aliphatic chain. The chain elongation and modification are conducted iteratively, until thioesterase (TE) hydrolyzes the fatty acyl-ACP thioester bond and releases the fatty acid product.

Based on the domain architecture, FAS can be classified as type I or type II FAS (Figure 1.3A). Type I FASs are found mainly in mammals and fungi, as well as certain bacteria. In type I FAS, the seven FAS catalytic domains are covalently linked and exist as a single polypeptide, where every active site except TE is used iteratively during fatty acid biosynthesis. In comparison, type II FASs are found in plants and bacteria, where each FAS catalytic domain exists as a stand-alone enzyme. Because of the importance of fatty acids, each FAS domain is highly conserved structurally and functionally from bacteria to mammals. Further, due to the importance of fatty

acids to basic survival, domains of FAS have been validated as drug targets against obesity, diabetes (14), cancer (15), as well as infectious diseases such as tuberculosis (16, 17). Crystal or NMR structures of FAS domains can serve as design templates for the development of new therapeutics.

The mammalian type I FAS is a homodimer (13), each 270 kDa monomer is active and consists of KS, AT, DH, ER, KR, ACP, and TE domains (Figure 1.3A). The 3.2 Å structure confirms the head-to-head model based on electron microscopy (18) and biochemical results (19, 20). The N terminal KS domains from both monomers interact extensively in the center of the megasynthase architecture (Figure 1.3 C-D) (21). The porcine FAS is separated into two portions: the lower condensing portion contains the condensing KS and MAT domains, while the upper portion consists of chain modifying domains DH, ER and KR (Figure 1.3C-D). Two additional nonenzymatic domains, termed ΨMe (an inactive methyltransferase) and ΨKR (a truncated KR) domains, are located at the periphery of the upper portion (Figure 1.3D). The central core of the X-shaped architecture of type I FAS consists of KS, DH and ER domains from both monomers, with an extensive dimer interface between KS, DH and ER domains. The dimer is held together by KS-KS, DH-DH and KS-DH interactions. The head-to-head arrangement also implies that ACP in either monomer (A or B) can interact with active sites of both monomers, consistent with previous biochemical studies (Figure 1.3B) (20, 22). The DH-KS interaction connects the top and the bottom of the X-shaped architecture, and strongly suggests that the DH fold is important in maintaining the FAS architecture. The AT and KR domains extend from the bottom and top of the protein, respectively, to form two asymmetric reaction chambers on either side of the central KS-DH-ER core. The flexible linker regions between KS-AT, AT-DH, DH-KR and KR-ER domains (Figure 1.3D, solid black lines) facilitate the opening and closing of each reaction chamber, thus

allowing the ACP-bound fatty acyl intermediate access to the active site of each domain. ACP and TE are invisible in the crystal structure, but these two C-terminal domains are presumed to be located near the center of each reaction chamber (Figure 1.3B-C) because ACP must present each catalytic domain with ACP-bound fatty acyl substrate.

The crystal structures of all type II FAS domains have been solved, including the type II KS, MAT, KR, ER, DH, ACP and TE domains (23-26). The reader can find an excellent review of type II FAS structural biology herein (27). The crystal structure of a type I FAS TE domain, as well as the NMR structure of a type I FAS ACP domain, have also been reported (28). One review comparing the type I FAS to the type I PKS can be found herein (29). These FAS single domain structures shed light on the molecular details of individual steps. However, in the absence of the larger complex structure, mechanistic details about protein-protein interactions are unavailable. The full length mammalian type I FAS solved to 3.2 Å has greatly expanded our knowledge about the complicated domain-domain interactions in the megasynthase (21), and it offers an excellent opportunity to correlate the FAS architecture with its close relative, the PKS. Although there is no solid proof, type II FAS is hypothesized to adopt a similar X-shaped architecture (29) due to the high degree of conservation between type I and II FAS domains. An excellent review of the type II FAS structural enzymology work can be found elsewhere (27).

1.3 Different Types of Polyketide synthases (PKS)

Based on the domain architecture, there are at least four distinct PKS types: modular type I, iterative type I, type II, and type III PKSs (Figure 1.4) (30). The modular type I PKSs are multi-modular complexes that often biosynthesize macrolides such as rifamycin, FK506, and erythromycin (31). Each "module" contains multiple catalytic domains that are typically used only

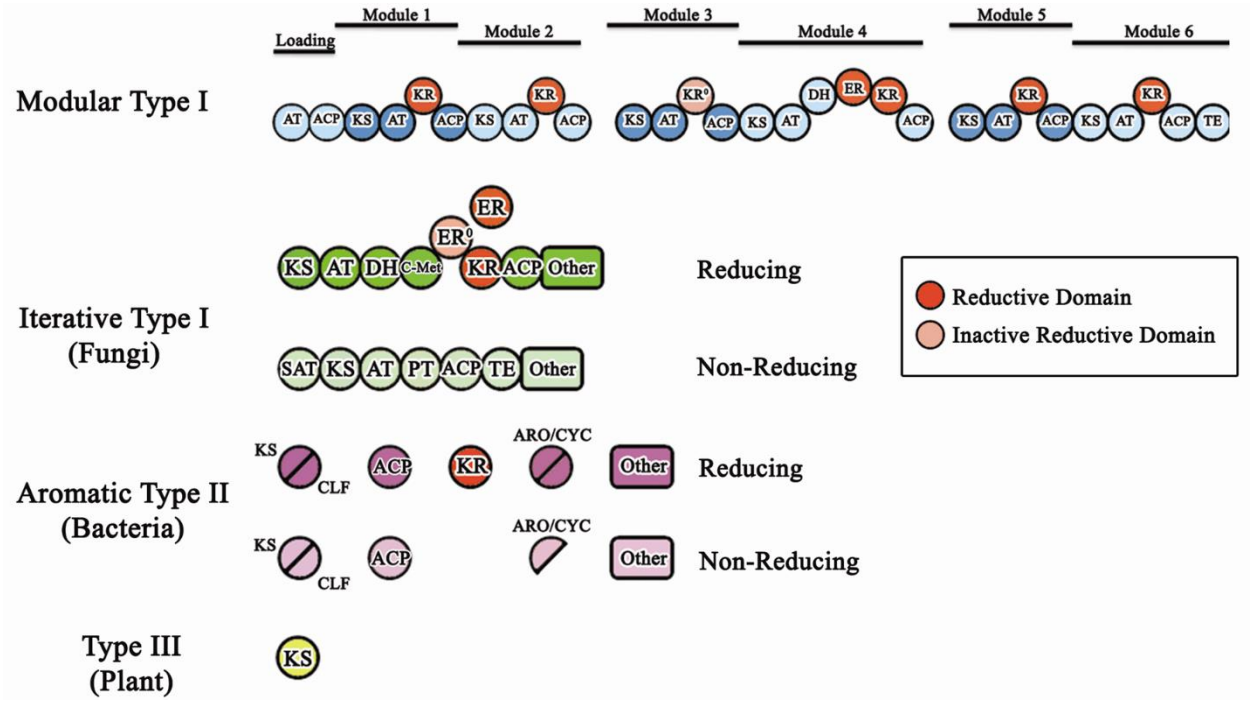


Figure 1.4 Different domain organization for modular Type I, iterative Type I, aromatic Type II, and Type III PKSs. The modular Type I PKS is a large multimodular complex where the β-carbonyl processing can be read directly from the linear order of domains (colinear logic). In contrast, iterative type I and aromatic Type II PKS use an interactive process (each active site used multiple times) resulting in polyketides containing cryptic β processing profiles that can only be inferred from the resulting polyketide. Both the iterative Type I and aromatic Type II PKSs can further be broken down into reducing and non-reducing types, depending on the presence or absence of KR domains.

once in an "assembly line" fashion. Iterative type I PKSs, often found in fungi, contain a single "module" that catalyzes polyketide chain elongation and modification in an iterative fashion. Type II PKSs, often found in *Streptomyces* species, consist of five to ten stand-alone domains that conduct iterative chain elongation, followed by reduction or cyclization to biosynthesize aromatic polyketides. Type II PKS and iterative type I PKS can be further divided into reducing and non-reducing PKSs depending on the presence or absence of reducing domains (Figure 1.4). Type III PKSs, often from plants, use a single domain to conduct a limited set of chain elongations followed by cyclizations to afford plant aromatic polyketides such as resverastrol. The reader can find an excellent review on type III PKSs herein (32), and this chapter will focus on the structural

enzymology of modular type I, iterative type I, and type II PKSs. An overview of each type of PKS is discussed below.

1.4 Modular type I PKS

1.4.1 Reaction overview

Modular type I PKSs are huge, multi-modular complexes that promote the biosynthesis of macrolides such as erythromycin (31) (Figure 1.4, top). The modular type I PKS resembles a molecular assembly line organized into multiple "modules". Each module contains multiple domains, often with the order KS-AT-DH-ER-KR-ACP-(TE), where AT is the abbreviation of acyltransferase. A "minimal module" contains KS-AT-ACP domains and catalyzes one round of chain elongation. The "minimal module" can optionally include chain modification domains such as DH, ER, KR, or MT (methyltransferase) domains. Domains of the same module are covalently linked. One module can be a single polypeptide, but different modules are often covalently linked into even longer polypeptides. During chain elongation and chain reduction step in modular PKS, the chain elongation and modification is catalyzed in tandem by different modules. Generally, each domain is used once in modular type I PKS, although exceptions exist and an excellent review on the loss of co-linearity by modular type I PKS can be found herein (33). As a result, the β-carbonyl processing of the resulting macrolide can be predicted directly from the linear order and number of domains present (termed "The Rule of Colinearity"). The linear logic of modular type I PKSs has been exploited in a "mix-and-match" approach to produce new polyketides with predicted reduction patterns (34, 35). Although this approach has been relatively successful in producing new macrolides, the lack of structural information about inter-domain and inter-module linker regions has resulted in the inactivity of many engineered hybrid complexes. Recently, computer

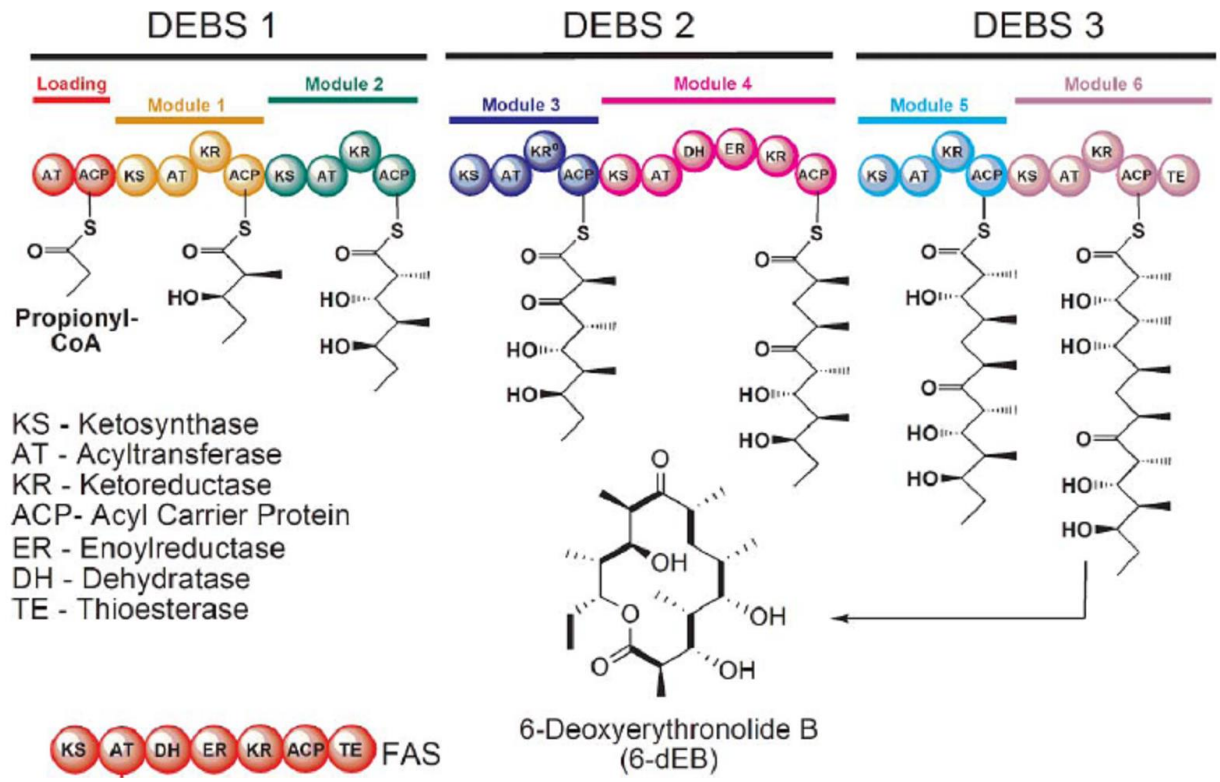


Figure 1.5 The DEBS Modular Type I PKS. Domain organization of the animal FAS and the DEBS modular PKS. In the FAS, the AT domain is responsible for loading both the starter and chain-elongating substrates, whereas in modular PKSs, these substrates are loaded by separate ATs in the loading and chain-extending modules. The intermediates formed by each of the DEBS modules are shown together with the final 14-membered ring product. The KR domain in module 3 of DEBS is inactive.

simulation and crystal structures of a modular KS-AT have provided valuable information about the KS-AT linker, and it suggests that at least the KS-AT region of each module adopts a similar quaternary structure as that of porcine FAS (36). The benefits and pitfalls of using type I PKSs as platforms for the development of new macrolides have been reviewed extensively (1, 37, 38).

The chemical reactions catalyzed by modular type I PKSs are very similar to type I FAS (Figure 1.2, Figure 1.5). Often, the N-terminal of a modular type I PKS consists of a "priming module" including the priming AT and ACP domains. The acyl substrate is attached to the priming ACP via a thioester bond with the phosphopantetheinyl group (PPT), which is covalently linked to the ACP, through the actions of the priming AT. The priming module is followed by the extending (chain elongating) modules. The primed acyl group is then trans-acylated to the

extending KS active site cysteine. The extending module typically contains an extending AT that loads malonyl, methylmalonyl or ethylmalonyl extender unit to the extending module ACP. KS then catalyzes decarboxylative Claisen condensation between the primer and extender units, followed by optional chain modification via KR, DH, ER or MT domains, using NADPH as the reduction cofactor. The elongated, modified polyketide chain is then passed to the next "module" until the completion of the final elongation modification, and the C-terminal TE domain catalyzes product release by macro-lactonization or thioester hydrolysis.

One of the best studied modular type I PKSs is the 6-deoxyerythronolide synthase (DEBS, Figure 1.5) (39). DEBS contains 28 domains that are organized into one loading-module (AT-ACP) and six extender modules. Besides the loading AT domain, all other ATs from modules 1-6 of DEBS are specific for methylmalonyl-CoA, which is the building block for six rounds of decarboxylative Claisen Condensation. Each round of chain elongation is carried out by different KS domains from the six modules. Ketoreduction (modules 1, 2, 4, 5, 6), dehydration (module 4), and enoylreduction (module 4) modify the linear polyketide chain in a highly regio- and stereo-specific manner. The terminal thioesterase (TE) then catalyzes an intra-molecular lactonization to biosynthesize the 14-membered ring product, 6-deoxyerythronolide (6-dEB), a precursor of the antibiotic erythromycin. An excellent review on the structure-function studies of DEBS can be found herein (39).

1.4.2 Similarities and differences between type I FAS and modular type I PKS

The type I FAS and PKS share many structural and functional similarities. First, the domain architectures are similar. For example, both type I FAS and PKS consist of multiple catalytic domains that are covalently linked to one (or several) polypeptides (Figure 1.5). Second, the KS

domain catalyzes chain elongation using 2-4 carbon building blocks in both FAS and PKS, with the growing chain attached to ACP. Third, in both type I FAS and PKS, the growing chain is modified by chain-modifying enzymes KR, DH, and ER, after each round of chain elongation. However, modular PKS differs significantly from type I FAS in the following chemical logic (Figure 1.2 versus Figure 1.5):

1. In FAS, MAT serves as both the priming and extending AT, with dual substrate specificity for acetyl- and malonyl-CoA (29). In comparison, in modular type I PKS, the priming and extending ATs are often located on different modules with distinct substrate specificities (40).
2. In type I FAS, the KS domain only accepts saturated acyl substrates for chain elongation. In modular type I PKS, the KS domains have a much broader substrate specificity and can accommodate acetyl, α-keto, α-hydroxyl, enoyl, or saturated acyl substrates (29).
3. In type I FAS, each round of chain extension and β-carbon processing utilizes the same KS, AT, KR, DH, ER and ACP domains. In modular type I PKS, the polyketide chain is processed by successive modules linked in assembly-line fashion, and the degree of β-carbon modification can differ at each module.
4. In type I FAS, the number of carbons (the "chain-length") of the product is determined by the KS and TE domains (hence the term "product specificity") (29). In modular type I PKS, the chain length of the polyketide product is dictated by the total number of priming and extending modules, as well as the presence of the thioesterase domain associated with the particular PKS.
5. Because the FAS biosynthesizes fatty acid iteratively, the same ACP transports the growing fatty acid chain, and reaches alternately into the active sites of KS, AT, KR, DH and ER, in a highly coordinated fashion. Generally, in modular type I PKS, each module has its own ACP and the growing chain is passed from one ACP in module n to the ACP in module n+1.

Therefore, to ensure the fidelity of polyketide biosynthesis, there must be highly specific inter-modular interactions through the linker regions (35).

1.4.3 Structural work on modular type I PKS

Extensive progress has been made on the structural biology of DEBS (39), including the crystal structures of KS3-AT3 (41), KS5-AT5 (36), KR1 (42), DH4 (43), TE (44), and the NMR structure of ACP2 (45). Here, the number indicates the module number, and this convention will be used throughout this chapter. KR1 of the tylosin PKS has been solved by Keatinge-Clay et al. (46), and its stereo-specificity critically evaluated (47). The TE domain from Pikromycin PKS (PIKS), has also been reported in apo and substrate analog bound forms (48-51). In addition, the NMR structure of intermodular linker region in DEBS has been reported (52). Finally, although they will not be discussed in this chapter, the crystal structures of several auxiliary enzymes of modular type I PKS have been reported, including the cytochrome P450 that are associated with pikromycin biosynthesis (53), as well as the series of enzymes that provide the methoxylmalonyl extender unit for FK506 and FK520 (54).

1.5 The type II PKS

1.5.1 Reaction overview

The type II PKSs are found mainly in Actinomycetes and produce aromatic polyketides (55) (Figure 1.1), which are characterized by the presence of multiple aromatic rings fused in a linear arrangement. These aromatic polyketides include commonly-used antibiotics (such as actinorhodin and tetracycline) and anticancer agents (such as doxorubicin) that target the DNA/RNA and protein biosyntheses in bacteria and cancer cells. Similar to the type II FAS, the

type II PKS carries out a series of reactions catalyzed by individual soluble enzymes, each of which are encoded by a discrete gene and show considerable sequence conservation (Figure 1.6). Early in the study of PKSs, there were extensive advances in *Streptomyces* genetics and type II PKS biochemistry from the groups of Hopwood (55), Bristol University (lead by Simpson et al) (56), Hutchinson (57), Floss(58), Khosla (59, 60), Robinson (61), Moore (62) and Salas (63). Their

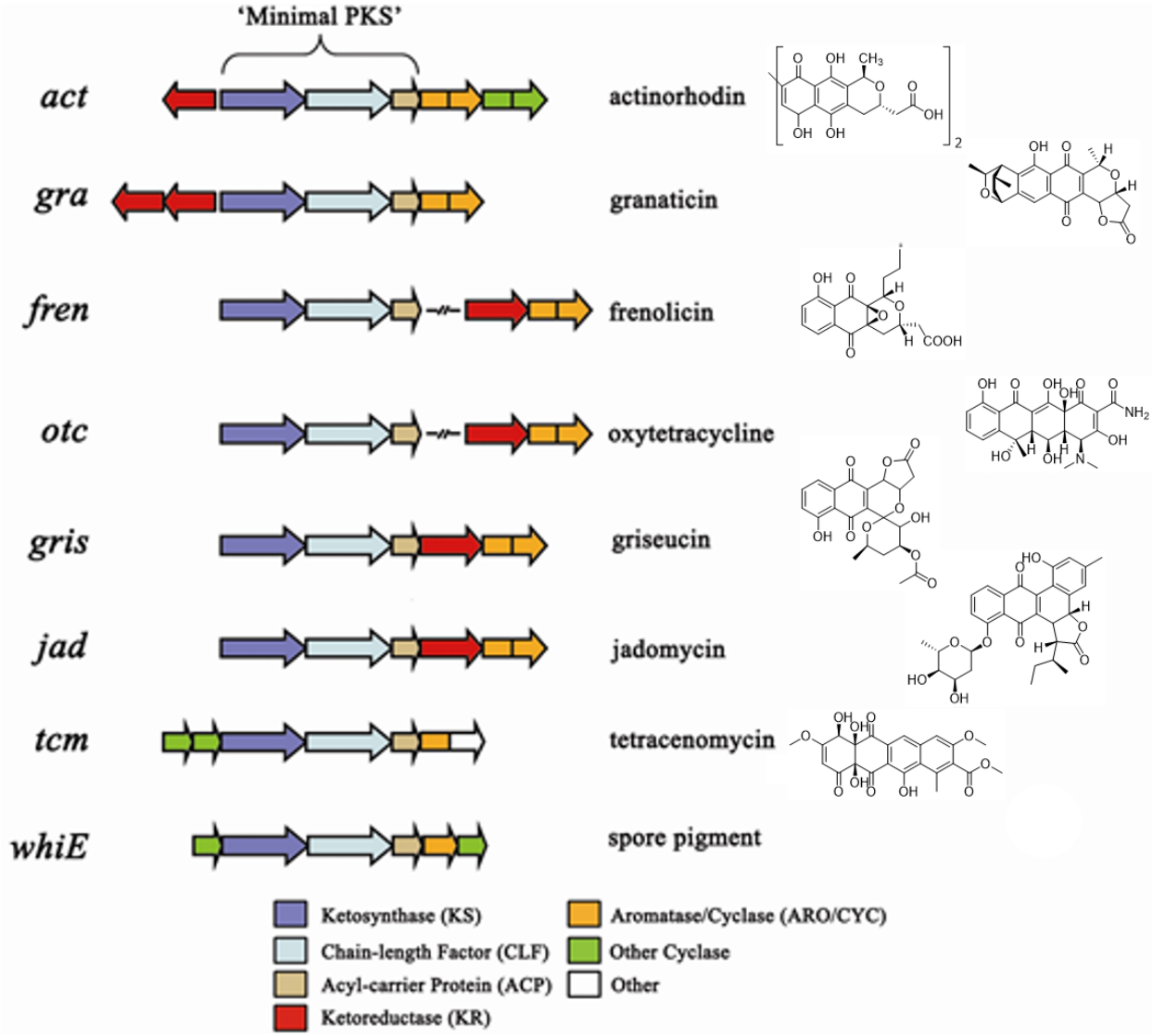


Figure 1.6 Organization of gene clusters in Type II PKSs. The 'minimal PKS' corresponds to the minimal protein machinery required for assembling the polyketide chain, including KS/CLF and ACP. Like the iterative Type I PKS, the Type II PKS can be grouped into reductive and non-reductive classes depending on the presence (red arrow, ***act-jad***) or absence (***tcm and whiE***) of a KR. The names and structures of the final product are shown on the right. The product for the whiE PKS is a spore pigment of unknown structure.

research supported a proposed mechanism (Figure 1.7), where the polyketide chain is assembled by a "minimal PKS" comprised of the ketosynthase (KS) / chain length factor (CLF) heterodimer and ACP. The CLF domain is in essence an inactivated KS where the active site Cys is replaced with a Gln. Chain elongation is then followed by first ring cyclization, chain reduction by KR, and subsequent ring aromatization/cyclization by aromatase /cyclase (ARO/CYC). The chain transfer among different catalytic domains is mediated by the phosphopantetheinyl group (PPT), which is

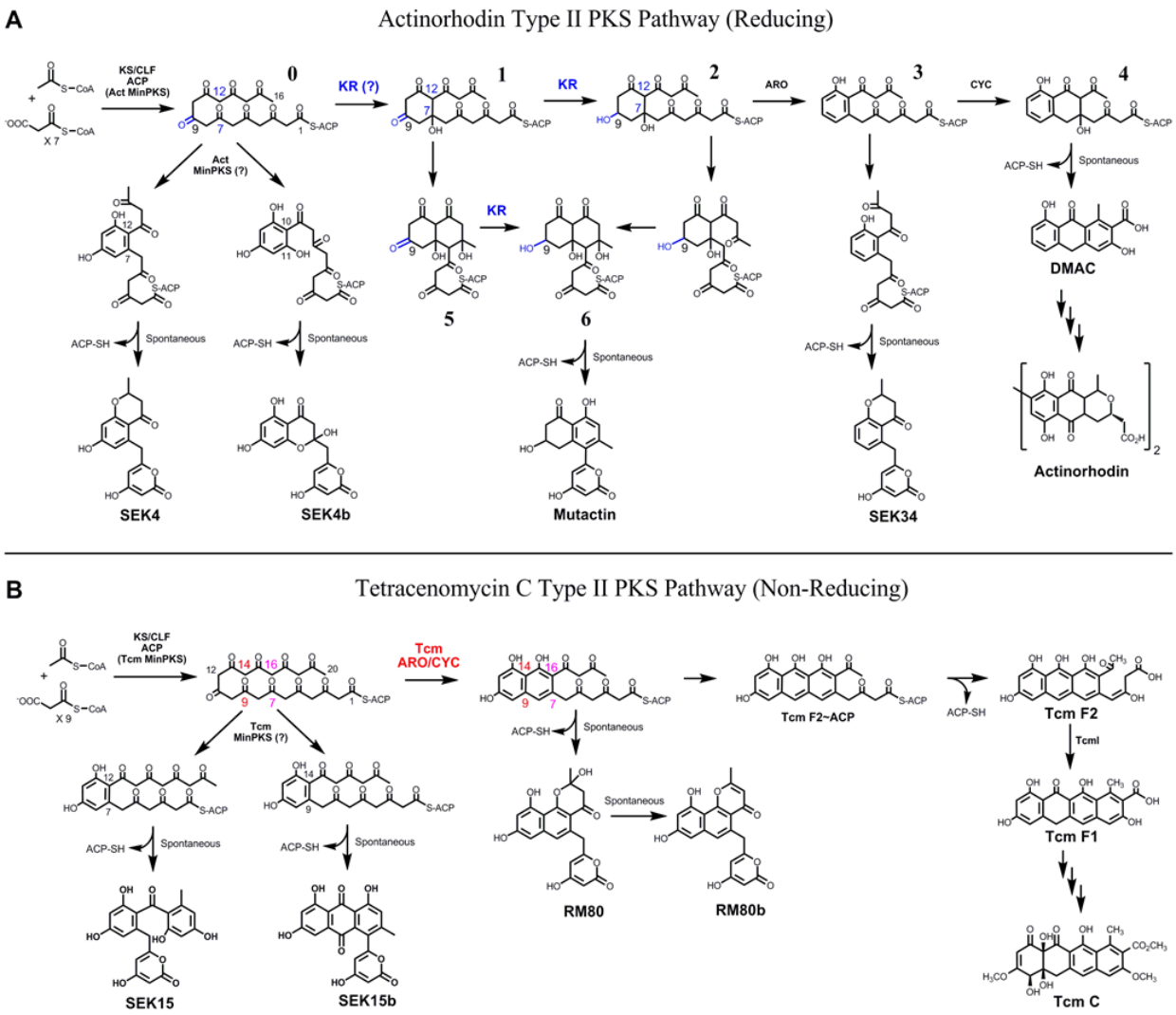


Figure 1.7 A comparison of the biosynthetic pathways for the actinorhodin (act) and tetracenomycin (tcm) PKSs. **(A)** The act PKS produces a 16C product and utilizes a KR to specifically reduce the C9 position (red) of a C7-C12 cyclized intermediate. **(B)** The tcm PKS produces a 20C product using ARO/CYC (blue) but lacks a KR, leading to C9-C14 first-ring cyclization. The various metabolites produced by the inclusion of only certain domains are boxed.

covalently linked to a conserved serine on ACP. Similar to type II FAS, each catalytic domain is a stand-alone domain in type II PKS.

The type II PKS can be further divided into two classifications: the reducing type II and non-reducing type II PKSs. The reducing type II PKS (Figure 1.7A, using the actinorhodin PKS as an example) contains a KR domain and typically reduces a C9-carbonyl group with a first ring C7-C12 cyclization. The non-reducing type II PKS (Figure 1.7B, using the tetracenomycin PKS as an example) lacks KR but contains a first ring ARO/CYC, biosynthesizes a C9-C14 first ring and C7-C16 second ring cyclized product. Similar to the approach exploited with modular type I PKSs, different type II PKSs have been combined in a "mix and match" fashion to biosynthesize over 100 new polyketides (64-74).

1.5.2 Similarities and differences between type II FAS and PKS

Past biochemical and genetic engineering experiments show that although the chemical reactions utilized by type II FAS and PKS are similar, there are many major differences:

(1) "Chain Length Factor": In type II FAS, there are three KS domains: the priming KS FabH, and the extending KSs, FabB and FabF, that catalyze decarboxylative Claisen condensations (75). The chain length controlled by each KS varies in different organisms. For example, in *E. coli*, FabH, FabB and FabF extend fatty acids with a chain length of 4, 4-16 and 16-18 carbons, respectively (27). In comparison, in *Mycobacterium tuberculosis*, where complex lipids are essential for pathogenic survival, the chain length can be as long as 56 carbons (76). Further, FabH, FabB and FabF are homodimers, whose chain length can be correlated with the size of the substrate binding pocket that is mainly defined by the dimer interface. In addition, one ACP is shared by all three KSs (75). In contrast, many type II PKSs do not have the priming KS, while the PKSs that

have a priming KS (such as the R1128 PKS) often contain two ACPs, one for the priming KS and the second for the extending KS (77). The biggest difference is the existence of chain length factor (CLF), which is essentially an inactivated KS domain, with the catalytic Cys mutated to Gln (69, 78). Domain swapping and mutational analyses support that in type II PKS, the polyketide chain length is controlled by the collective action of the priming KS and KS/CLF heterodimer (65).

(2) Ketoreduction. The ketoreduction by the ketoreductase (KR) is chemically identical to the corresponding fatty acid ketoreduction. However, the regio-specificity is drastically different between type II FAS and PKS. Whereas the type II FAS KR reduces every carbonyl group in each iterative round, the type II PKS KR typically reduces the C9-carbonyl group of a polyketide with C7-C12 first ring cyclization (79).

(3) Dehydration versus Aromatization. In type II FAS, two dehydratases, FabA and FabZ, catalyze the dehydration of the fatty acyl chain (80, 81). FabA and FabZ each have different substrate specificities, but both have the conserved His-Asp catalytic dyad. Type II PKS does not have a DH domain. Rather, the polyketide chain is pre-assembled by KS/CLF, followed by the action of the aromatase/cyclase (ARO/CYC), which produces the aromatic rings essential for the antibiotic and anticancer activity of type II polyketides (82). The same polyketide chain can be folded, in origami fashion, to different cyclization patterns. For example, doxorubicin and mithramycin are both cyclized from the same 20-carbon linear polyketide (6, 8) that are folded at C7-C12 versus C9-C14, respectively (83, 84). Further, the inclusion of KR often is related with a di-domain ARO/CYC with C7-C12 first ring cyclized product (67, 74, 84), while the inclusion of a mono-domain ARO/CYC (in the absence of KR) often produces C9-C14 first ring cyclized product (66, 84).

(4) Protein-protein interactions. During combinatorial biosynthesis, the activity of type II PKS is often abolished due to the loss of protein-protein interactions (35, 85). Further, the difference in chain elongation and chain modification also indicates that FAS and PKS have different protein-protein interactions (67). Finally, the stereospecificity during ketoreduction and cyclization specificity of ARO/CYC may also be related to protein- protein interactions between KS/CLF, ACP, KR and ARO/CYCs (86).

1.5.3 Structural work on type II PKS

Extensive progress has been made on the structural biology of the actinorhodin (act) PKS, including the crystal structures of MAT (87), KS/CLF (88) and KR (89-91) domains, and the NMR structure of holo and apo *act* ACP (55, 56). The NMR structures of two more ACPs, the *fren* ACP (92) and *otc* ACP (93), are also reported. The crystal structure of the priming ketosynthases, ZhuH and DpsC, of the R1128 and daunorubicin PKSs, respectively, have been solved (48, 94). The crystal structures of 5 ARO/CYC domains have been solved: the TcmN ARO/CYC (95), the WhiE ARO/CYC (96), the ZhuI ARO/CYC (97), the StfQ ARO/CYC and the BexL ARO/CYC (98). In addition, three fourth-ring cyclase structures have been reported (99-101).

1.6 Iterative type I PKS

Iterative type I PKS, generally found in fungi, consists of multiple catalytic domains that are covalently linked together. However, different from that of the modular type I PKS, domains in the iterative type I PKS are generally used in multiple rounds of chain elongation and modification. Similar to type II PKS, depending on the presence or absence of reducing domains KR and ER, the iterative type I PKS can be divided into reducing and non-reducing iterative type I PKS (102). The readers can find an excellent review on iterative type I PKS herein (11).

1.7 The ketoreductase (KR)

Three KR crystal structures have been reported: for modular type I PKS, crystal structures of the DEBS KR1 (EryKR1) (42) and KR1 of tylosin PKS (TylKR1) (46) have been solved (Figure 1.8A-B); for type II PKS, the actinorhodin KR (ActKR) crystal structure has been reported (89-91) (Figure 1.8D-E). EryKR1 and TylKR1 reduce a diketide substrate C=O to C-OH with the "2R, 3R" and "2R, 3S" stereo-chemistry, respectively (Figure 1.9). Therefore, these two KRs choose opposite diketide epimers (at the 2 position) for the reduction reaction. In comparison, the actinorhodin KR (ActKR) from *Streptomyces coelicolor* specifically reduces the C9 carbonyl group of a 16 carbon (octaketide) pre-assembled polyketide chain, which folds into C7-C12 first ring cyclized product mutactin (Figure 1.7A). The ketoreduction catalyzed by ActKR, as well as

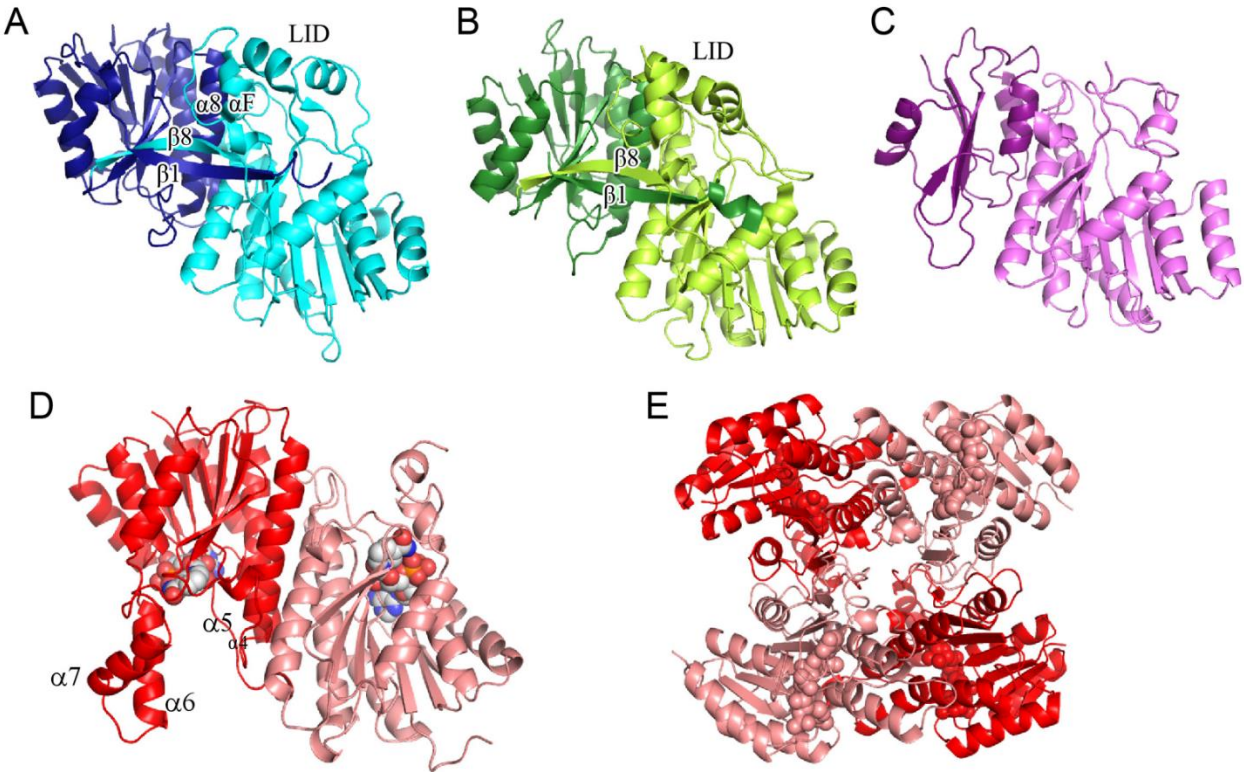


Figure 1.8 The crystal structures of **(A)** EryKR1 (structural and catalytic sub-domains in darker and lighter blue, respectively). (B) TylKR1 (structural and catalytic sub-domains in darker and lighter green, respectively). **(C)** Porcine FAS KR (structural and catalytic sub-domains in darker and lighter purple, respectively). **(D)** Act KR dimer (monomers A and B in red and pink, NADP in spheres). **(E)** The Act KR tetramer, colored as in **(D)**, NADP in spheres. **(A)** and **(C)** are in the same orientation.

by other type II PKS KR domains, is chemically identical to the corresponding fatty acid ketoreduction, yet with very different regio-specificities (79). Below, the molecular basis of ketoreduction specificity will be discussed for both modular type I and type II KRs in terms of sequence-function-structure relationships.

1.7.1 KR has a short-chain dehydrogenase/reductase (SDR) fold

In both FAS and PKS, the type I KR has two domains with the same protein fold: the catalytic sub-domain and a truncated, non-catalytic structural sub-domain. Both EryKR1 and TylKR1 are monomeric in solution and in crystal structures (Figure 1.8A-B, compared to porcine FAS KR domain in Figure 1.8C) (42, 46). In comparison, the type II KR exists as a tetramer (Figure

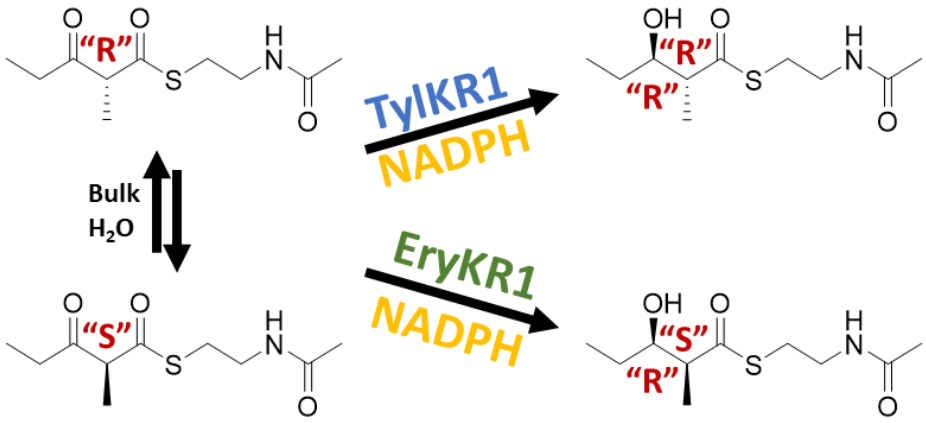


Figure 1.9 Reactions catalyzed by TylKR1 and EryKR1, sampling different diketide enantiomers.

1.8E) (89, 91), and each monomer contains a single domain. Each domain (or each sub-domain in type I KR) contains a short-chain dehydrogenase/reductase (SDR) fold, which consists of a highly conserved Rossmann fold with two right-handed α–β–α–β–α motifs that are connected by α3, and the core region consists of a 7-stranded β–sheet flanked by α-helices (Figure 1.8). The cofactor NADPH is bound at the junction of two α–β–α–β–α motifs in a highly conserved groove, containing a highly conserved pyridine-nucleotide binding-site motif (GxG/AxxG/AxxxG/A)

characteristic of a Rossmann fold (103). The polyketide substrate-binding pocket consists of a large cleft (on the average, 15 Å in width, 19 Å in length, 17 Å in depth) formed by helices α6–α7 and the loops between α4 and α5 (Figure 1.8D). The catalytic subdomains of EryKR1 and TylKR1, as well as ActKR, have the typical SDR motifs, such as the TGxxxGxG motif (##12-19), D63 and NNAG motif (##89-92), the active site tetrad Asn-Ser-Lys-Tyr, and the PG motif (##187-188) (103). The biggest difference between the type I and II PKS KRs is a long insertion between helices 6 and 7 for ActKR, and this may account for the different substrate specificities between type I and II PKS KRs.

The monomeric type I KR orients its two sub-domains (Figure 1.8A-B) very similarly to a type II KR dimer (Figure 1.8D), where there are extensive interactions, mainly hydrophilic in nature. The catalytic domains between EryKR1, TylKR1 and ActKR can be overlapped well with an RMSD of 1.5 – 1.8 Å, and the major differences reside in the "lid" region (Figure 1.8A-D) between α6–α7 and α4–α5. The structural sub-domains in EryKR1 and TylKR1 do not contain a cofactor-binding motif, and also lack the substrate-binding portion, thus rendering the structural sub-domain inactive (42, 46). The type I KRs have additional β1–β8 and α4–αΦ interactions that bridge the structural and catalytic sub-domains and stabilize the pseudo-dimeric KR protein fold. Based on the EryKR1 structure, Keatinge-Clay and Stroud proposed that, similar to modular type I PKS KR, the type I FAS KR may also consist of two sub-domains that are split and separated by ER and DH domains (42, 46)(Figure 1.3). Indeed, this hypothesis has been proven to be the case in the 3.2-Å porcine FAS crystal structure (Figure 1.8C) in which the KR also contains the structural and catalytic sub-domains that are separated by DH and ER. The major difference between the type I FAS and PKS KRs lies in the structural domain, which is truncated by 50 residues in type I FAS KR structure (Figure 1.8A-C).

1.7.2 An Ordered Bi Bi mechanism

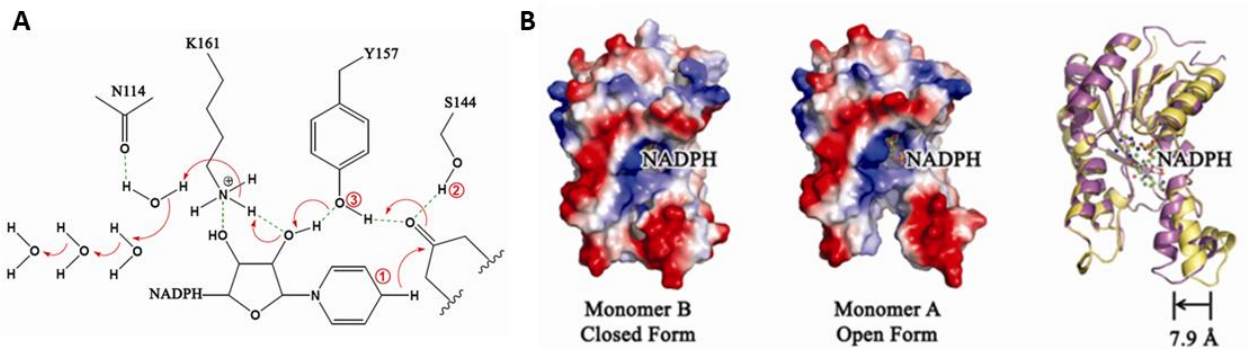


Figure 1.10 (A) Proposed proton-relay mechanism for act KR. **(B)** The asymmetric unit contains monomers A (**middle** panel) and B (**left** panel) in open and closed conformations, respectively. Surface potentials were colored from negative (red) to positive (blue). When the open and closed conformations are overlapped (right panel, open in yellow, closed in purple), the major conformational change is in the flexible a6-a7 region, with a 7.9 Å difference between the two monomers.

The ActKR active site tetrad consists of N114-S144-Y157-K161 (89-91). In comparison, the Asn and Lys positions are switched in EryKR1 and TylKR1 (K1776-S1800-Y1813-N1817) (42, 46). The KR tetrad is located near the nicotinamide ring of NADPH, where Tyr and Lys form hydrogen bonds with the NADPH ribose and nicotinamide ring. In ActKR, four crystalline water molecules form extensive hydrogen-bonds with N114 and K161 (Figure 1.10A). This extensive proton relay water network is very similar to the one observed in *E. coli* FabG-$NADP^+$ (104), leading to the hypothesis that the water-relay mechanism for FabG may also be applicable to ActKR: the ketone substrate is hydrogen bonded to S144 and Y157, whose sidechains form the oxyanion hole. Following hydride transfer from NADPH to the ketone substrate, the resulting alkoxide is stabilized by the oxyanion hole, while the tyrosyl proton is transferred to the alkoxide. An extensive proton relay then takes place to replenish the proton extracted from the tyrosyl-OH, sequentially including the 2-OH of NADPH ribose, lysine-NH, and then followed by the four water molecules (Figure 1.10A). A search through literature, Protein Data Bank and sequence

comparison proposed that the proton relay network may be a conserved feature in both FAS and PKS KRs.

Past genetic domain swapping experiments in a *Streptomyces* host suggested that ActKR may promote the polyketide first ring cyclization (Figure 1.7A, **0** to **1**) (84, 86, 105). However, it is unclear if the substrate of ActKR is linear (polyketide **0**) or cyclic (**1** or **5** Figure 1.7A). *In vitro* assays monitoring NADPH oxidation of a wide range of potential substrates, including linear, monocyclic and bicyclic polyketide mimics show that ActKR has a strong preference for bicyclic compounds, such as trans-1-decalone, trans-2-decalone and the aromatic tetralone (90). This result indicates that the type II PKS KRs have a different substrate specificity from that of FAS (106, 107) and modular type I PKS KRs (108), both of which are capable of reducing linear and mono-cyclic ketones. The strong preference of ActKR towards bicyclic polyketides supports that its natural substrate is likely a cyclic polyketide, such as **1** and **5** (Figure 1.7A). Using the bicyclic substrate trans-1-decalone and the tricyclic inhibitor emodin, detailed steady-state and inhibition kinetic studies further confirmed that KR proceeds through an Ordered Bi Bi mechanism (90), in which the cofactor NADPH binds KR prior to the substrate trans-1-decalone. Product release is most likely the rate-limiting step, as shown by viscosity dependence of k_{cat} and k_{cat}/K_m (unpublished data). The proposed mechanism is supported by mutations, inhibitor emodin-bound structures, and docking simulations (90, 109).

The timing of the first ring cyclization relative to ketoreduction has been a long-discussed issue. Extensive docking simulations of ActKR with polyketide substrates show that if the substrate is a linear polyketide (Figure 1.7A, **0**), ActKR can dock with every carbonyl group pointed towards the cofactor NADPH (90). This is inconsistent with the observed high specificity towards the C9-carbonyl group. Rather, docking simulations show that when a cyclic polyketide

(such as **5**, Figure 1.7A) is docked into ActKR, under the dual-constraints imposed by the ring structure of **5** and the three-point docking of the KR active site, the C9-carbonyl group of **5** is optimally positioned for ketoreduction for the C7-C12 cyclized polyketide. Therefore, both in vitro studies and docking simulations support that the cyclization event that leads to **5** (Figure 1.7A) may happen before ketoreduction. If so, does the first ring cyclization occur in KR or in KS/CLF?

The inhibitor emodin-bound ActKR structure also revealed that ActKR can exist with at least two different conformations (90): the open and closed forms that differ in the 10-residue loop region (residues 199-209) between helices 6-7 (Figure 1.10B). As a comparison for type I PKS KRs, EryKR1 and TylKR1 also show similar conformational changes and loop movement in this region (42, 46). Only the open conformation of ActKR is bound with the inhibitor emodin. The α6−α7 insert is a highly unique feature in type II PKS KRs (not observed in type I PKS or FAS KRs), and it also defines half of the substrate-binding pocket. The observed conformational flexibility in this loop region is proposed to influence KR substrate specificity. Separately reported for ActKR and TylKR1, the observed open and closed conformations may reflect different binding motifs during ketoreduction (46, 90).

1.7.3 The molecular basis of KR substrate specificity

1.7.3.1 Substrate specificity of modular type I PKS KR

The stereo-selective signature motifs for the modular PKS KRs were previously proposed to be "LDD" and "PxxxN" (110, 111), and the presence of these motifs predicts formation of the 3R stereomer (or "D stereomer") whereas the absence of these motifs predicts the 3S stereomer (or "L stereomer", Figure 1.11A). Keatinge-Clay and Stroud showed in the EryKR1 crystal structure that the 93-95 "LDD" motif is located in a loop adjacent to the active site (42). In DEBS

KR2, residues W141 and 93-95 “PQQ” are the counterparts of F141 and “LDD” in EryKR1. The importance of these motifs was demonstrated by mutational analyses of KR1 and KR2, in which the double mutation of KR1 (F141W, P144G) resulted in a switch of the alcohol stereochemistry (112, 113). Sequence alignment indicates that the “LDD” motif corresponds to “LRD” in mammalian FAS KR domains and is therefore predicted to produce the “R” stereochemistry. Consistent with this prediction, the mammalian FAS does produce a (3R)-hydroxyacyl moiety (114).

The 2 position stereochemistry is also affected by KR and its upstream KS domain: the extending ATs in DEBS only accept (2S)-methylmalonyl-CoA, but KS reverses its chirality during chain elongation to produce “R stereochemistry” at the α-position of the newly elongated chain (108, 115, 116). This is further complicated by the fact that some KS domains do not accept a 2S-methylmalonyl group. For example, studies of DEBS1-TE showed that the condensation of (2S)-methylmalonyl-CoA in DEBS module 2 proceeds with chirality inversion without cleavage of the C−H bond adjacent to the methyl group, corresponding to the inversion at 2S configuration as observed in FAS (117). In contrast, in DEBS module 1, the chain extension lost the hydrogen attached to C-2 of the methylmalonyl-CoA precursor, which implies that an additional obligatory epimerization step takes place in module 1. Four theories have been proposed to explain the determination of C2 stereochemistry: **(1)** Some KSs catalyze C2-epimerization after condensation to yield a mixture of (2S)- and (2R)-epimers, and a downstream enzyme selects one epimer of the enantio-mixture (118). **(2)** Spontaneous epimerization occurs in water, and a downstream enzyme selects one of the epimers (118). **(3)** KR catalyzes a cryptic epimerization, and hydride attacks the enol form of the polyketide intermediate; subsequently the 2 position chirality is determined by the relative orientation of C2 of the polyketide enol (119). **(4)** Some KRs epimerize the C2 position

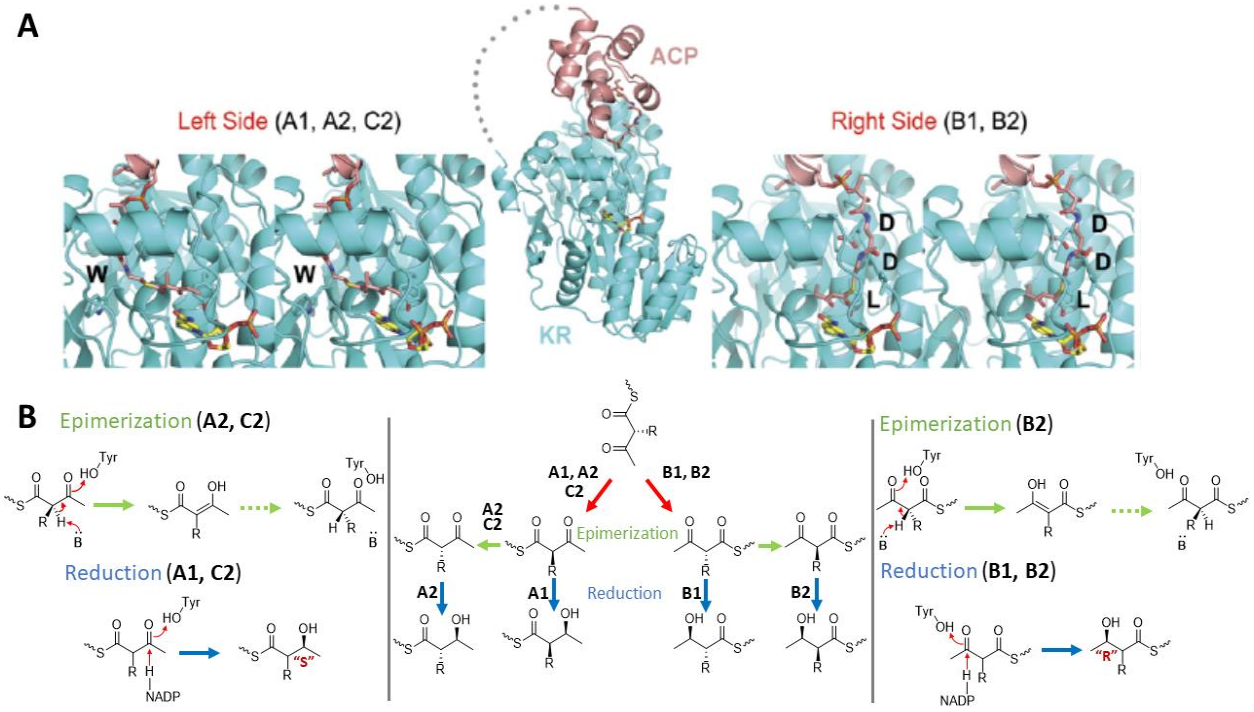


Figure 1.11 Epimerazation and stereospecifity of type I KRs. **(A)** The polyketide chain may be protected from epimerization within ACP until docking with KR. In A- and C2-type KRs, the polyketide can enter the active site from the left side. In B-type KRs, the PPT arm encounters the LDD motif, which prevents it from slipping behind the lid helix, and the polyketide enters the active site from the right side. Only the nicotinamide half of NADPH is shown. **(B)** Six types of Modular Type I KRs (see text for details). **Center**: through a combination of binding, epimerization, and reduction events, each KR type controls how a polyketide is processed. **Left** and **right**: catalysis proceeds similarly whether the polyketide enters a KR active site from the left side or from the right side. Epimerization is likely a combination of a catalyzed enolization and an uncatalyzed tautomerization back to the keto form. The catalytic base cannot catalyze enolization of the epimerized polyketide because the acidic hydrogen is inaccessible. In A2- and C2-type KRs, the base is either a lid helix residue or water; in B2-type KRs, the base may be the catalytic tyrosine. Reduction occurs through attack of the β-carbon by the NADPH hydride along with the transfer of a proton from the catalytic tyrosine to the carbonyl oxygen. Figure adapted from reference (47) with permission.

before catalyzing the reduction reaction. Based on extensive bioinformatic analysis guided by the crystal structures of EryKR1 and TylKR1, Keatinge-Clay further categorized the modular type I KRs into six types with the following explanations of their observed stereochemistry at the 2 and 3 positions (Figure 1.11A) (46): **(1)** A1-type (2R, 3S): The substrate enters the active site from the left side of KR (Figure 1.11B) as guided by the conserved tryptophan and a Gln or Leu interacts with unepimerized 2 (□ to the thioester carbonyl group) substituent. The reduction reaction yields a “3S” hydroxyl group. **(2)** A2-type KR (2S, 3S): Same as A1, but the polyketide substrate is

enolized when the acidic α-hydrogen is abstracted by a base (a lid helix residue or a water molecule), followed by an uncatalyzed tautomerization to an epimerized polyketide, which cannot be enolized again due to the inaccessibility of the enolization base caused by steric clashes with the side chains of a conserved histidine. The reduction yields a "3S" hydroxyl group. **(3)** B1-type (2R, 3R): The unepimerized polyketide enters the active site from the right side (Figure 1.11B), guided by the leucine of the LDD motif, and the aforementioned Gln, together with lid helix residues, can prevent epimerization at the 2 position. The reduction yields a "3R" hydroxyl group. Most KRs that produce a substrate for a DH are B1-type KRs, and they possess a highly conserved and slightly longer lid, as in the related FAS KRs, FabG and FabI (120, 121) that are also accompanied by a DH. **(4)** B2-type KRs (2S, 3R): The unepimerized polyketide enters the active site from the right side, guided by the leucine of the LDD motif. The catalytic tyrosine may serve as the enolization base, similar to the A2-type. In EryKR1, an active site Leu and loop helix Val create a hydrophobic pocket for the epimerized 2-methyl group. The reduction yields a "3R" hydroxyl group. **(5)** C1-type KRs are rare, nonfunctional KRs that lack the catalytic tyrosine, with no epimerization. **(6)** C2-type KRs, although inactive for ketoreduction, allow epimerization similar to the A2-type KRs. A protocol to assign substituent stereochemistry is also developed according to the above six groups (46).

1.7.3.2 Substrate specificity of type II PKS KR

Similar to two other SDR proteins, the fungal 1,3,8 reductase (T_3HNR) and 1,3,6,8-tetrahydroxynaphthalene (T_4HNR) (122-124), in vitro assays showed that ActKR (and by inference based on sequence alignment, other type II PKS KRs) strongly prefers bicyclic substrates and inhibitors (90). In the absence of downstream ARO and CYC domains, ActKR may reduce a

cyclic intermediate (**1** or **5**, Figure 1.7A). When PPT-tethered **1** and **5** were docked into the ActKR substrate pocket, the PPT group consistently docked into a positively-charged surface patch containing three arginines, with multiple carbonyl groups forming hydrogen-bonds with the active site residues. The substrate- and inhibitor-bound ActKR crystal structures, combined with in vitro assays, site-directed mutagenesis and sequence analysis (90), support that ActKR substrate specificity is defined by a combination of enzyme conformation, specific molecular interactions between the substrate and active site residues, as well as substrate and protein flexibility due to the dynamic nature of the binding cleft.

Nearly all type II KRs contain a unique motif corresponding to 94-XGG-96, which aligns with the "LDD" motif in modular type I PKS KRs (110, 111). In the case of ActKR, this motif corresponds to 94-PGG-96. However, ActKR may be one of the few SDR members that are non-selective during ketoreduction (125), producing a mixture of S- and R-mutactin (Figure 1.7A), based on previous reported inability to determine the C9-stereochemistry of mutactin (126). Systematic mutations of the PGG motif in ActKR confirmed in vitro that wildtype ActKR has a 3:1 preference for S- versus R-tetralol, when assayed from the reverse direction (alcohol oxidation by $NADP^+$)(127). In comparison, the single P94L mutant resulted in the complete abolishment of R-tetralol oxidation activity, while retaining full catalytic capacity for S-tetralol. Heterologous expression of act minimal PKS with wild type or P94L ActKR is consistent with the in vitro assay results. A comparison of emodin-bound wild type and P94L crystal structures show that when the P94 is changed to L94, steric interactions preclude oxidation of the R-isomer (Figure 1.12). This implies that the stereo-specific reduction of the ACP-bound polyketide is influenced at least partially by steric constraints imposed by pocket residues.

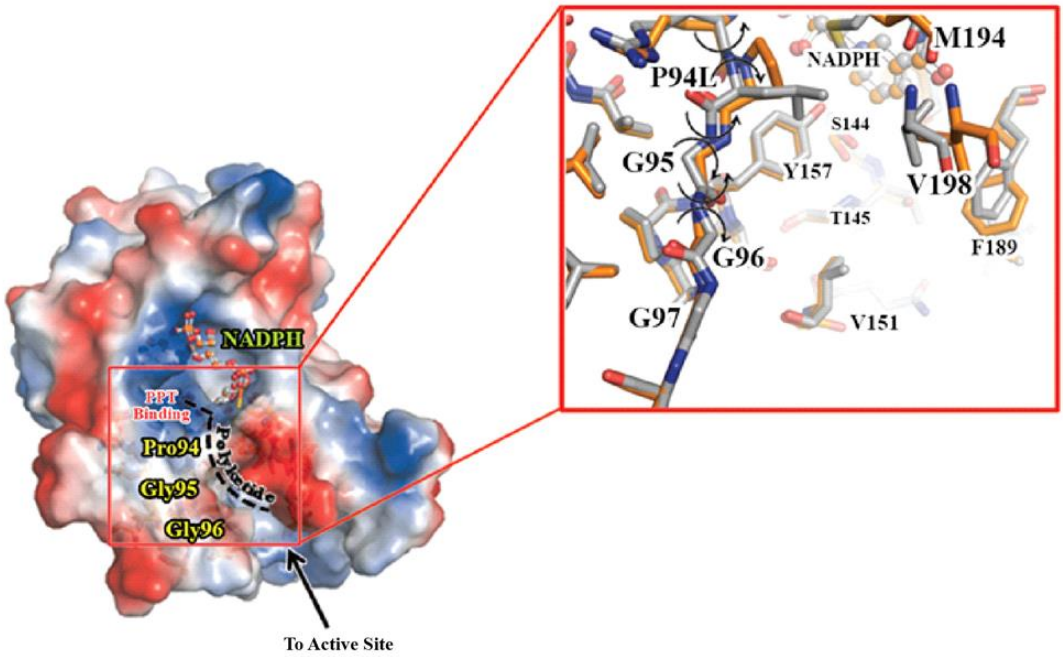


Figure 1.12 Overlay of the ActKR wildtype and P94L mutant structures. Left panel: Surface representation of the P94L mutant protein and the location of the PGG motif in relation to NADPH (orange ball and stick), the predicted PPT binding site (red letters), and the direction of polyketide binding into the active site (dashed line). Right panel: Close-up view of the overlaid wild-type (orange) and P94L (gray) structures. Curved arrows suggest that the P to L mutation allows the LGG loop more rotational freedom.

Previously, Moore, et al. proposed that ActKR accepts two possible substrates, the full length octaketide (16 carbons) or a truncated pentaketide (10 carbons) that undergoes subsequent rounds of elongation to produce a reduced hexaketide (128). However, docking simulations showed that a minimum of 12 carbons is necessary to position the C5=O in the oxyanion hole, assuming the PPT group is anchored in the arginine patch (R38, R65 and R93), as previously proposed (90). This region has also been identified in the homologous SCO1815 KR from *Streptomyces coelicolor* (129) and would restrict the attached polyketide to a single side of the pocket. This suggests that similar to KS/CLF (77, 88), ActKR physically counts the carbon chain-length, starting from the end of the PPT group (which is anchored to the same arginine patch), and the chain-length is determined by the physical distance between the PPT group and the KR active

site. This hypothesis helps explain how ActKR can reduce the C9 position, even when presented with longer substrates such as C18 or C20 (67).

1.8 Acyl Carrier Protein

1.8.1 Type I PKS ACP structure

The small size of ACP renders solution structures by nuclear magnetic resonance (NMR) spectroscopy an ideal methodology for visualizing the ACP structures. For type I FAS, the mammalian FAS apo ACP structures have been solved by NMR and X-ray crystallography (Figure 1.13A), and both structures revealed that ACP consists of a four-helix bundle stabilized by hydrophobic interactions (130, 131). In 2007, the first solved structure of an ACP from a type I modular PKS was also reported for DEBS ACP2 (45). Similar to the type I FAS ACP structure, the 10-kD DEBS ACP2 contains a three-helical bundle, and an additional short helix in the second loop also contributes to the core helical packing. The conserved serine residue in the universal "DSL" motif (where PPT is covalently attached) lies at the N-terminal end of helix-2. For both FAS and PKS, the type I and type II ACP structures are very similar, in which helix 2 is characterized by a highly conserved cluster of negatively charged residues and is regarded as a universal 'recognition helix' involved in interactions with other proteins (55, 92, 93, 130, 132) (Figure 1.13B).

Based on the solution structure of DEBS ACP2, homology models were constructed for five other DEBS ACP domains (ACP1-6, Figure 1.13D) (45). While the topologies of all six DEBS ACPs are similar, there are also major differences. A comparison of their steric and electrostatic surfaces at the putative interaction interface (centered on the "recognition helix" II) suggests a model for protein–protein recognition of ACP domains, consistent with the previously observed specificity that each ACP is highly specific for its corresponding KS domain (40). For example,

ACP6 is the least similar, both sterically (the lack of several hydrophobic residues conserved in other ACPs) and electrostatically (the only ACP with a neutral, rather than negative surface near helix II). The possible reason is because ACP6 does not need to transfer the chain to a downstream KS domain. Several distinguishing features could also be found for ACP3, such as an electropositive patch near the α2-α3 loop. Studies on type II ACPs using NMR indicate that the bound acyl moieties can adopt two different states; one involving direct interaction with the ACP through the cleft flanked by helices □□□and□□□□ the other allowing direct exposure to solvent (132). This two-state equilibrium may facilitate functional interaction of the bound acyl moieties with other catalytic domains. Consistent with the hypothesis that the ACP recognition site is near the "recognition helix II", mutagenesis indicates that, at least for the DEBS ACP2 domain, the specificity of ACP interactions with KS domains can be influenced by the amino acids located near the recognition helix (45).

1.8.2 The type II ACP structure

The NMR solution structure of actinorhodin acyl carrier protein (Act *apo*-ACP) from *S. coelicolor* represents the first high-resolution PKS structure reported in the literature (55). The Act ACP structure is composed of four helices and a large loop region between helices 1 and 2. The Act ACP protein fold is very similar to the *E. coli* FAS ACP (130). Later in 2003, the fren and otc ACP solution structures were solved (92, 93), which also have a similar fold (Figure 1.13). The Act apo-ACP has a disordered N-terminus and a longer flexible loop than the type II FAS ACP. The type II PKS ACPs also contain highly conserved hydrophilic Arg and Asn that may help stabilize the growing polyketide chain. In the same year, the first holo-ACP solution structure was solved for the frenolicin (Fren) ACP with the PPT group visible (92). The averaged Fren ACP

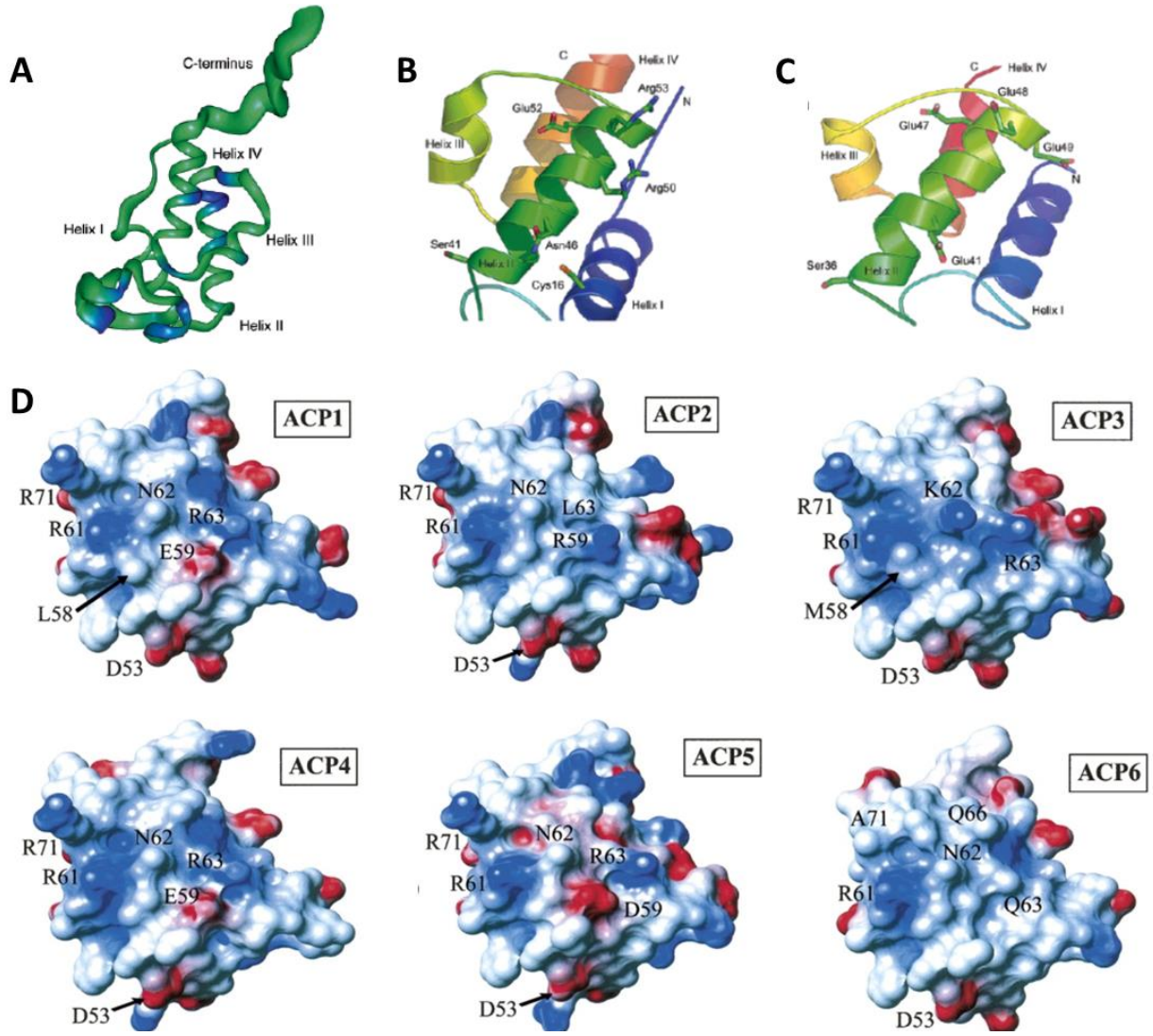


Figure 1.13 (A) Overall fold of ACP, the thickness corresponds to internal movement. **(B)** Negative and positive residues in Otc ACP that may interact with the incoming polyketide intermediate. **(C)** Negative residues in E. coli FAS ACP that may interact with the PPT group. **(D)** Electrostatic potential surface diagrams for DEBS ACP1-6. Positively charged residues are colored blue and negatively charged residues are colored red. Residue numbers for the labeled residues correspond to ACP2 numbering.

structure showed three long α-helices packed the core bundle, while three short helices are observed in intervening loops with one slowly exchanges between two conformations. A comparison of the *fren* ACP NMR structure with the crystal structure of butyryl-ACP from *E. coli* suggests that the conformational change may be triggered by polyketide binding to ACP. The solution structure of oxytetracycline (Otc) ACP further revealed that the long loop between helices I and II is flexible (Figure 1.13A) (93). Similar to the type I system, helix II is proposed to be the

"hot spot" for protein-protein interactions between ACP and downstream PKS domains. The hydrophilic groove is distinctly different between PKS and FAS ACPs: the former generally has both positively and negatively charged residues, as well as hydrophobic residues, thus the helical groove is capable of interacting with an incoming polyketide chain with both hydrophobic interactions and hydrogen bonds. In comparison, the FAS ACPs have a much more hydrophobic core groove, with negatively charged residues proposed to interact with the incoming fatty acids (130). Recent NMR experiments also revealed that the three longer helices in type II ACP support a rigid scaffold, and dynamic ACP conformation, necessary for substrate binding and transport, may be promoted by the flexible linker regions between helices that help harbor the highly active polyketide intermediates in the hydrophilic "recognition groove" between $\alpha 2$ and $\alpha 3$ (56).

As a concluding remark, the above observation can be further combined with the acyl-bound (hexanoyl, heptanoyl and decanoyl) crystal structures of the *E. coli* type II ACPs (133) (Figure 1.13C), which reveal that the acyl chains are sequestered in the hydrophobic cavity within the four-helix bundle, and the size of the cavity expands as the length of the bound acyl chain increases while the overall fold remains un-altered. This is also consistent with the "Switch Blade" theory based on the yeast FAS crystal structure (134), in which prior to each round of chain extension, the growing acyl chain (the "blade") switches its nestling cavity from the ACP to the KS (or other PKS enzymes) binding pocket, similar to a Swiss Army Knife. Therefore, the four-helix ACP structure must be quite dynamic, and the timing of "blade-switching" is closely related to the degree of exposure between the polyketide intermediate and the solvent, which not only depends on the ACP "recognition helix" property, but also on the chemical structure of a given polyketide intermediate. In a mega-synthase such as DEBS, besides specific interactions based on the electronic and steric properties of both ACP and substrates, the ACP may also conduct a

"random walk" until the correct substrate and downstream enzyme are matched.

1.9 Transport of polyketide intermediates between domains in type II PKS

The single domain studies reviewed above strongly supports the proposal about an intricate balance between different type II PKS domains, so that chain elongation, reduction and cyclization occur in a precisely timed fashion. Such an accurate timing requires extensive protein-protein interaction, as well as precise control in polyketide-ACP transportation between KS/CLF, KR and ARO/CYC. Presumably, such extensive protein-protein interactions require the formation of a multi-enzyme complex that is centered on the ACP. Protein sizing experiments showed that KS/CLF, KR, ARO/CYC and ACP exist as heterotetramer, homotetramer, monomer and monomer, respectively (55, 89, 93, 135, 136). Therefore, if there is a stable multi-domain complex, such as KS/CLF-KR-ACP, it should be detectable by protein sizing experiments. However, past studies by native gel and size-exclusion chromatography cannot detect any stable complex between apo or holo ACP and KS/CLF, KR or ARO/CYC (unpublished data). Similarly, no stable enzyme complex could be detected between KS/CLF and KR or ARO/CYC (unpublished data). The only exception is when the KS active site Cys is labeled with a 16-carbon acyl chain, and only then can a tight KS/CLF-ACP complex be detected by protein sizing experiments (Yi Tang, personal communication). The above results imply that the protein surface-surface interactions may not be sufficient to warrant a stable complex formation between type II PKS domains. Rather, the bound polyketide substrate may also be important in multi-enzyme complex formation in the type II PKSs. Of course, enzyme domains in the type II PKS may form a highly transient complex that evades the detection of protein sizing experiments. However, because of the availability of rate constants for the minimal PKS and ActKR, we feel that such a transient complex is a less

likely scenario—for an enzyme complex that turns over once every minute, the enzyme complex should, minimally, remain intact for one minute, which is sufficient time to allow detection via newer protein sizing techniques such as mass spectroscopy, similar to experiments on modular type I PKS (137). Therefore, in type II PKS, protein-protein interactions likely result from a combination of protein surface interacting "hot spots", as well as the polyketide intermediate that is attached to ACP via the PPT anchor. This also helps explain the precise timing observed in chain transport—only when the polyketide is "matured" would the next protein-protein interaction be promoted, resulting in the formation of a dynamic protein complex that may dissolve after the current reaction is completed.

With the above argument, two key questions remain: (1) Are there "hot spots" on each domain in type II PKS? And (2) can one trap the protein complex with a polyketide bound? The former question can be visualized via protein-protein docking simulation between ACP and other domains. In the protein-protein docking simulations, Arg/Lys rich, positively charged grooves were identified near the substrate pocket entrance of KS/CLF, KR and ARO/CYC (88, 89, 135). Also, the "arginine patch" surface generally complements well with the negatively charged surface of the ACP helical groove, where the polyketide is proposed to be stored in the "Switch Blade" theory (134). This also helps explain how the highly reactive polyketide intermediate may be transported from one domain to the next: in essence, ACP serves as the "time capsule" that alternately reaches in between one domain and its downstream partner domain, using its helical groove as the sheath to protect the reactive polyketide intermediate. Only when the polyketide chain is "mature" can a stable, albeit semi-transient complex be formed between the upstream domain, ACP, and the downstream domain (for example, KS/CLF-ACP-KR, or KS/CLF-ACP-ARO/CYC). The polyketide chain is transferred to the next domain as soon as the semi-transient

complex is formed. While this hypothesis may be attractive, it still requires experimental proof, preferentially in the form of a polyketide-bound complex of KS/CLF-ACP-KR or KS/CLF-ACP-ARO/CYC.

1.10 Design and application of stable isosteric polyketide mimetics for the study of type I iterative NR-PKS and type II PKS

Both type I iterative non-reducing (NR) PKS and type II PKS conduct multiple iterative rounds of chain elongation without reducing the β-carbonyl in between elongation cycles, resulting in a linear poly-β-ketone intermediate (Figure 1.14A). This intermediate is prone to spontaneous self-cyclization reactions (Figure 1.14B) preventing crystallographic studies of type I iterative NR-PKS and type II PKS enzymes-substrate interactions with the poly-β-ketone intermediate. In order to solve this problem, we developed stable atom-replaced mimetics that sterically and electronically mimic the unstable poly-β-ketone intermediates. These mimetics were applied with great success in the study of the product template (PT) domain from the fungal non-reducing polyketide synthase (NR-PKS), PksA, from aflatoxin biosynthesis and the C9 ketoreductase domain (KR) from actinorhodin biosynthesis.

1.11 Conclusion

The revelations from crystallographic analysis of type I and type II PKSs have raised awareness of the extraordinary architecture of these mega-synthases and offer a perspective in visualizing some of the unsolved questions concerning polyketide biosynthesis. The 3.2 Å porcine FAS crystal structure has provided a framework for the mega-synthase architecture that may also apply to type I and type II PKSs (21). In the arena of type II PKS, the detection of a multi-enzyme complex formation should help shed light on the question about how the ACPs gain access to each

of the PKS enzyme domains.

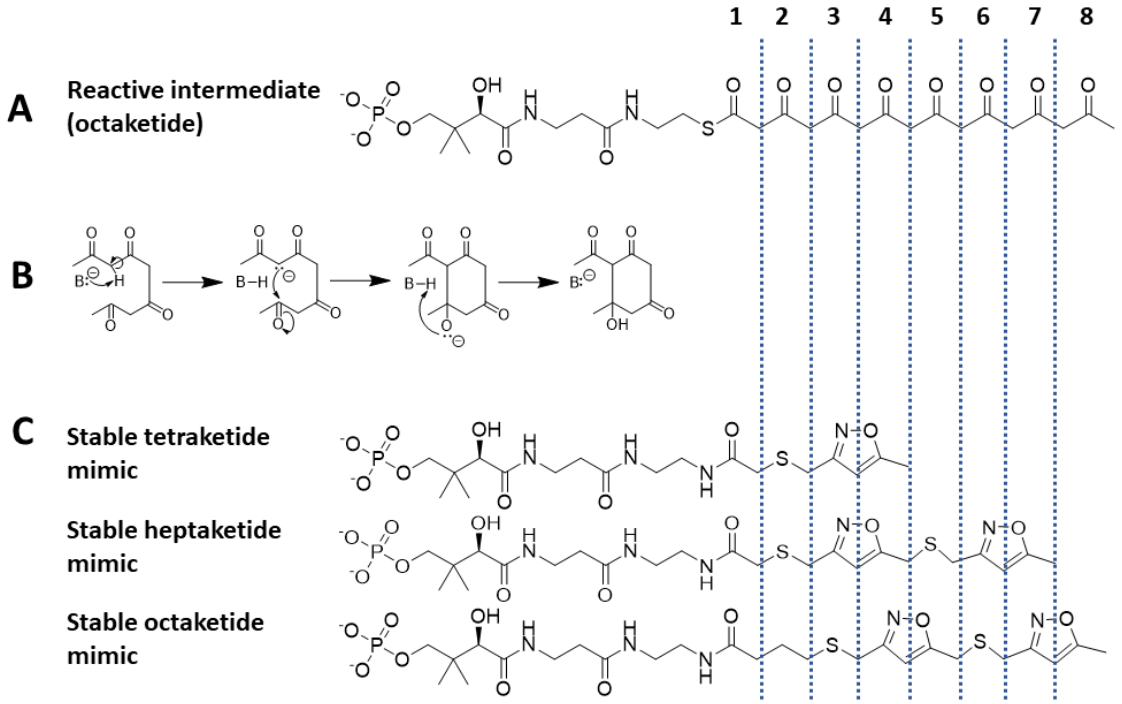


Figure 1.14 Rationale and design of stable poly-β-ketone mimetics. **(A)** The intrinsically unstable linear poly-β-ketone intermediate attached to the phosphopantetheinyl group (PPT). **(B)** Linear poly-β-ketones longer than three ketides are prone to spontaneous cyclizations through water catalyzed aldol condensation reactions. **(C)** Stable poly-β-ketones mimetics designed to mimic the unstable intermediate both sterically and electronically. These poly-β-ketone mimetics provide an opportunity to visualize near-native enzyme-substrate binding interactions through crystallographic studies.

In chapter two of this thesis, we apply our atom-replaced heptaketide mimetic to the study of the product template (PT) domain from a fungal non-reducing polyketide synthase (NR-PKS). PTs are responsible for controlling the aldol cyclizations of poly-β-ketone intermediates assembled during the catalytic cycle. They catalytically guide the regioselectivity of ring formation to provide specific products distinct from the thermodynamic outcome of random aldol self-condensations. Our ability to understand the high regioselective control that PT domains exert is hindered by the inaccessibility of intrinsically unstable poly-β-ketones for in vitro studies. We describe here the crystallographic application of 'atom replacement' mimetics in which isoxazole rings linked by thioethers mimic the alternating sites of carbonyls in the poly-β-ketone intermediates. We report

the 1.8 Å co-crystal structure of the PksA PT domain from aflatoxin biosynthesis with a heptaketide mimetic tethered to a stably modified 4´-phosphopantetheine (PPant), which provides important empirical evidence for a previously proposed mechanism of PT-catalyzed cyclization. Key observations support the proposed deprotonation at C4 of the nascent polyketide by the catalytic His1345 and the role of a protein-coordinated water network to selectively activate the C9 carbonyl for nucleophilic addition. The importance of the 4´-phosphate at the distal end of the pantetheine arm is demonstrated to both facilitate delivery of the heptaketide mimetic deep into the PT active site and anchor one end of this linear array to precisely meter C4 into close proximity to the catalytic His1345. Additional structural features, docking simulations, and mutational experiments characterize protein–substrate mimic interactions, which likely play roles in orienting and stabilizing interactions during the native multi-step catalytic cycle. These findings afford the first view of a polyketide 'atom-replaced' mimetic in a NR-PKS active site that could prove general for other PKS domains.

In chapter two of this thesis, we apply our atom-replaced octaketide mimetic to the crystallographic study of the C9 ketoreductase (KR) from type II actinorhodin PKS. During polyketide biosynthesis, the polyketide stereocenters are introduced upon ketoreduction, which is catalyzed by the NAD(P)H dependent KR. The KR domain has been extensively characterized in type I modular PKS. In contrast, the regio- and stereo-specificities of type II PKS KRs are less well understood. The actinorhodin KR (ActKR) specifically reduces the C-9 carbon of an octaketide, and current evidence suggests that KR is also involved in first ring cyclization. In addition, the ActKR has an unexplained substrate specificity towards longer substrates (pentaketide and octaketide), but not tetraketide. The instability of native polyketide substrates has been the major roadblock towards understanding the molecular basis of KR specificity. We apply

our novel mimetic strategy to great success in probing ActKR activity and specificity. We herein report two cocrystal structures of a fully-active ActKR mutant (H153Y H201G) bound with the atom-replaced mimetics corresponding to tetraketide and octaketide substrates. The octaketide cocrystal structure affords a first glimpse into the importance of phosphopantetheine for chain-length control.

In Chapter 4, we present the crystal structure of the C9 KR from the type II doxorubicin/daunorubicin PKS. Both doxorubicin and daunorubicin are produced by strains of *Streptomyces peucetius*. Doxorubicin, the product of the C13 hydroxylation of daunorubicin by the cytochrome P450 DoxA, has become one of the most important anticancer therapeutics of all time. The study of DpsKR crystal structures from two different crystal forms and two Dps/ActKR chimeras shed light on the role of the lid helices, α6 and α7, in type II KR-substrate interactions.

1.11 References

1. Hopwood DA (1997) Genetic Contributions to Understanding Polyketide Synthases. *Chem Rev* 97(7):2465-2498.
2. Watve MG, Tickoo R, Jog MM, & Bhole BD (2001) How many antibiotics are produced by the genus Streptomyces? *Archives of microbiology* 176(5):386-390.
3. Haight TH & Finland M (1952) The antibacterial action of erythromycin. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y* 81(1):175-183.
4. Malpartida F & Hopwood DA (1984) Molecular cloning of the whole biosynthetic pathway of a Streptomyces antibiotic and its expression in a heterologous host. *Nature* 309(5967):462-464.
5. Motamedi H & Hutchinson CR (1987) Cloning and heterologous expression of a gene cluster for the biosynthesis of tetracenomycin C, the anthracycline antitumor antibiotic of Streptomyces glaucescens. *Proc Natl Acad Sci U S A* 84(13):4445-4449.
6. Otten SL, Stutzman-Engwall KJ, & Hutchinson CR (1990) Cloning and expression of daunorubicin biosynthesis genes from Streptomyces peucetius and S. peucetius subsp. caesius. *J Bacteriol* 172(6):3427-3434.
7. Jakobi K & Hertweck C (2004) A gene cluster encoding resistomycin biosynthesis in Streptomyces resistomycificus; exploring polyketide cyclization beyond linear and angucyclic patterns. *J Am Chem Soc* 126(8):2298-2299.

8. Lombó F, Blanco G, Fernández E, Méndez C, & Salas J (1996) Characterization of Streptomyces argillaceus genes encoding a polyketide synthase involved in the biosynthesis of the antitumor mithramycin. *Gene* 172(1):87-91.
9. Manzoni M & Rollini M (2002) Biosynthesis and biotechnological production of statins by filamentous fungi and application of these cholesterol-lowering drugs. *Appl Microbiol Biotechnol* 58(5):555-564.
10. Schoppner A & Kindl H (1984) Purification and properties of a stilbene synthase from induced cell suspension cultures of peanut. *J Biol Chem* 259(11):6806-6811.
11. Schumann J & Hertweck C (2006) Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *Journal of Biotechnology* 124(4):690-703.
12. Seto H, Sato T, & Yonehara H (1973) Letter: Application of carbon-13 in biosynthetic studies; FT-13C nuclear magnetic resonance spectra of dihydrolatumcidin. *The Journal of antibiotics* 26(10):609-614.
13. Maier T, Jenni S, & Ban N (2006) Architecture of mammalian fatty acid synthase at 4.5 A resolution. *Science* 311(5765):1258-1262.
14. Clapham JC & Arch JR (2007) Thermogenic and metabolic antiobesity drugs: rationale and opportunities. *Diabetes, obesity & metabolism* 9(3):259-275.
15. Kridel SJ, Lowther WT, & Pemble CWt (2007) Fatty acid synthase inhibitors: new directions for oncology. *Expert opinion on investigational drugs* 16(11):1817-1829.
16. Wright HT & Reynolds KA (2007) Antibacterial targets in fatty acid biosynthesis. *Curr Opin Microbiol* 10(5):447-453.
17. Purohit HJ, Cheema S, Lal S, Raut CP, & Kalia VC (2007) In Search of Drug Targets for Mycobacterium tuberculosis. *Infectious disorders drug targets* 7(3):245-250.
18. Asturias FJ, *et al.* (2005) Structure and molecular organization of mammalian fatty acid synthase. *Nat Struct Mol Biol* 12(3):225-232.
19. Joshi AK, Rangan VS, & Smith S (1998) Differential affinity labeling of the two subunits of the homodimeric animal fatty acid synthase allows isolation of heterodimers consisting of subunits that have been independently modified. *J Biol Chem* 273(9):4937-4943.
20. Joshi AK, Witkowski A, & Smith S (1997) Mapping of functional interactions between domains of the animal fatty acid synthase by mutant complementation in vitro. *Biochemistry* 36(8):2316-2322.
21. Maier T, Leibundgut M, & Ban N (2008) The crystal structure of a mammalian fatty acid synthase. *Science* 321(5894):1315-1322.
22. Witkowski A, *et al.* (2004) Head-to-head coiled arrangement of the subunits of the animal fatty acid synthase. *Chem Biol* 11(12):1667-1676.
23. Leesong M, Henderson BS, Gillig JR, Schwab JM, & Smith JL (1996) Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure* 4(3):253-264.
24. Olsen JG, *et al.* (1999) The X-ray crystal structure of beta-ketoacyl [acyl carrier protein] synthase I. *FEBS Lett* 460(1):46-52.
25. Price AC, Zhang YM, Rock CO, & White SW (2001) Structure of beta-ketoacyl-[acyl carrier protein] reductase from Escherichia coli: negative cooperativity and its structural basis. *Biochemistry* 40(43):12772-12781.

26. Serre L, Verbree EC, Dauter Z, Stuitje AR, & Derewenda ZS (1995) The Escherichia coli malonyl-CoA:acyl carrier protein transacylase at 1.5-A resolution. Crystal structure of a fatty acid synthase component. *J Biol Chem* 270(22):12961-12964.
27. White SW, Zheng J, Zhang YM, & Rock (2005) The structural biology of type II fatty acid biosynthesis. *Annu Rev Biochem* 74:791-831.
28. Chakravarty B, Gu Z, Chirala SS, Wakil SJ, & Quiocho FA (2004) Human fatty acid synthase: structure and substrate selectivity of the thioesterase domain. *Proc Natl Acad Sci U S A* 101(44):15567-15572.
29. Smith S & Tsai SC (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat Prod Rep* 24(5):1041-1072.
30. Shen B (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* 7(2):285-295.
31. Caffrey P, Bevitt DJ, Staunton J, & Leadlay PF (1992) Identification of DEBS 1, DEBS 2 and DEBS 3, the multienzyme polypeptides of the erythromycin-producing polyketide synthase from Saccharopolyspora erythraea. *FEBS Lett* 304(2-3):225-228.
32. Austin MB & Noel JP (2003) The chalcone synthase superfamily of type III polyketide synthases. *Nat Prod Rep* 20(1):79-110.
33. Moss SJ, Martin CJ, & Wilkinson B (2004) Loss of co-linearity by modular polyketide synthases: a mechanism for the evolution of chemical diversity. *Nat Prod Rep* 21(5):575-593.
34. Castonguay R, He W, Chen AY, Khosla C, & Cane DE (2007) Stereospecificity of ketoreductase domains of the 6-deoxyerythronolide B synthase. *J Am Chem Soc* 129(44):13758-13769.
35. Gokhale RS, Tsuji SY, Cane DE, & Khosla C (1999) Dissecting and exploiting intermodular communication in polyketide synthases. *Science* 284(5413):482-485.
36. Tang Y, Kim CY, Mathews, II, Cane DE, & Khosla C (2006) The 2.7-Angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc Natl Acad Sci U S A* 103(30):11124-11129.
37. Walsh CT (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* 303(5665):1805-1810.
38. Weissman KJ & Leadlay PF (2005) Combinatorial biosynthesis of reduced polyketides. *Nat Rev Microbiol* 3(12):925-936.
39. Khosla C, Tang Y, Chen AY, Schnarr NA, & Cane DE (2007) Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu Rev Biochem* 76:195-221.
40. Chen AY, Schnarr NA, Kim CY, Cane DE, & Khosla C (2006) Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-deoxyerythronolide B synthase. *J Am Chem Soc* 128(9):3067-3074.
41. Tang Y, Chen AY, Kim CY, Cane DE, & Khosla C (2007) Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem Biol* 14(8):931-943.
42. Keatinge-Clay AT & Stroud RM (2006) The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. *Structure* 14(4):737-748.
43. Keatinge-Clay A (2008) Crystal structure of the erythromycin polyketide synthase dehydratase. *J Mol Biol* 384(4):941-953.

44. Tsai SC, *et al.* (2001) Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel. *Proc Natl Acad Sci U S A* 98(26):14808-14813.
45. Alekseyev VY, Liu CW, Cane DE, Puglisi JD, & Khosla C (2007) Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Sci* 16(10):2093-2107.
46. Keatinge-Clay AT (2007) A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem Biol* 14(8):898-908.
47. Castonguay R, *et al.* (2008) Stereospecificity of ketoreductase domains 1 and 2 of the tylactone modular polyketide synthase. *J Am Chem Soc* 130(35):11598-11599.
48. Pan H, *et al.* (2002) Crystal structure of the priming beta-ketosynthase from the R1128 polyketide biosynthetic pathway. *Structure* 10(11):1559-1568.
49. Tsai SC, Lu H, Cane DE, Khosla C, & Stroud RM (2002) Insights into channel architecture and substrate specificity from crystal structures of two macrocycle-forming thioesterases of modular polyketide synthases. *Biochemistry* 41(42):12598-12606.
50. Giraldes JW, *et al.* (2006) Structural and mechanistic insights into polyketide macrolactonization from polyketide-based affinity labels. *Nature chemical biology* 2(10):531-536.
51. Akey DL, *et al.* (2006) Structural basis for macrolactonization by the pikromycin thioesterase. *Nature chemical biology* 2(10):537-542.
52. Weissman KJ (2006) The structural basis for docking in modular polyketide biosynthesis. *Chembiochem* 7(3):485-494.
53. Sherman DH, *et al.* (2006) The structural basis for substrate anchoring, active site selectivity, and product formation by P450 PikC from Streptomyces venezuelae. *J Biol Chem* 281(36):26289-26297.
54. Watanabe K, Khosla C, Stroud RM, & Tsai SC (2003) Crystal structure of an Acyl-ACP dehydrogenase from the FK520 polyketide biosynthetic pathway: insights into extender unit biosynthesis. *J Mol Biol* 334(3):435-444.
55. Crump MP, *et al.* (1997) Solution structure of the actinorhodin polyketide synthase acyl carrier protein from Streptomyces coelicolor A3(2). *Biochemistry* 36(20):6000-6008.
56. Evans SE, *et al.* (2008) An ACP structural switch: conformational differences between the Apo and Holo forms of the actinorhodin polyketide synthase acyl carrier protein. *Chembiochem* 9(15):2424-2432.
57. Hutchinson CR (1997) Biosynthetic Studies of Daunorubicin and Tetracenomycin C. *Chem Rev* 97(7):2525-2536.
58. Floss HG (2001) Antibiotic biosynthesis: from natural to unnatural compounds. *J Ind Microbiol Biotechnol* 27(3):183-194.
59. Zawada RJ & Khosla C (1997) Domain analysis of the molecular recognition features of aromatic polyketide synthase subunits. *J Biol Chem* 272(26):16184-16188.
60. McDaniel R, Licari P, & Khosla C (2001) Process development and metabolic engineering for the overproduction of natural and unnatural polyketides. *Adv Biochem Eng Biotechnol* 73:31-52.
61. Robinson JA (1991) Polyketide synthase complexes: their structure and function in antibiotic biosynthesis. *Philos Trans R Soc Lond B Biol Sci* 332(1263):107-114.
62. Moore BS & Piel J (2000) Engineering biodiversity with type II polyketide synthase genes. *Antonie Van Leeuwenhoek* 78(3-4):391-398.

63. Mendez C & Salas JA (2003) On the generation of novel anticancer drugs by recombinant DNA technology: the use of combinatorial biosynthesis to produce novel drugs. *Comb Chem High Throughput Screen* 6(6):513-526.
64. Shen B & Hutchinson CR (1996) Deciphering the mechanism for the assembly of aromatic polyketides by a bacterial polyketide synthase. *Proc Natl Acad Sci U S A* 93(13):6600-6604.
65. Tang Y, Lee TS, & Khosla C (2004) Engineered biosynthesis of regioselectively modified aromatic polyketides using bimodular polyketide synthases. *PLoS Biol* 2(2):E31.
66. McDaniel R, Ebert-Khosla S, Hopwood DA, & Khosla C (1995) Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* 375(6532):549-554.
67. McDaniel R, Ebert-Khosla S, Fu H, Hopwood DA, & Khosla C (1994) Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc Natl Acad Sci U S A* 91(24):11542-11546.
68. Fu H, McDaniel R, Hopwood DA, & Khosla C (1994) Engineered biosynthesis of novel polyketides: stereochemical course of two reactions catalyzed by a polyketide synthase. *Biochemistry* 33(31):9321-9326.
69. McDaniel R, Ebert-Khosla S, Hopwood DA, & Khosla C (1993) Engineered biosynthesis of novel polyketides. *Science* 262(5139):1546-1550.
70. Meadows ES & Khosla C (2001) In vitro reconstitution and analysis of the chain initiating enzymes of the R1128 polyketide synthase. *Biochemistry* 40(49):14855-14861.
71. Carreras CW & Ashley GW (2000) Manipulation of polyketide biosynthesis for new drug discovery. *Exs* 89:89-108.
72. Shen Y, *et al.* (1999) Ectopic expression of the minimal whiE polyketide synthase generates a library of aromatic polyketides of diverse sizes and shapes. *Proc Natl Acad Sci U S A* 96(7):3622-3627.
73. Tang Y, Tsai SC, & Khosla C (2003) Polyketide chain length control by chain length factor. *J Am Chem Soc* 125(42):12708-12709.
74. Hertweck C, *et al.* (2004) Context-dependent behavior of the enterocin iterative polyketide synthase; a new model for ketoreduction. *Chem Biol* 11(4):461-468.
75. Heath RJ & Rock CO (2002) The Claisen condensation in biology. *Nat Prod Rep* 19(5):581-596.
76. Takayama K, Wang C, & Besra GS (2005) Pathway to synthesis and processing of mycolic acids in Mycobacterium tuberculosis. *Clinical microbiology reviews* 18(1):81-101.
77. Tang Y, Lee TS, Kobayashi S, & Khosla C (2003) Ketosynthases in the initiation and elongation modules of aromatic polyketide synthases have orthogonal acyl carrier protein specificity. *Biochemistry* 42(21):6588-6595.
78. Bisang C, *et al.* (1999) A chain initiation factor common to both modular and aromatic polyketide synthases. *Nature* 401(6752):502-505.
79. O'Hagan D (1993) Biosynthesis of fatty acid and polyketide metabolites. *Nat Prod Rep* 10(6):593-624.
80. Kimber MS, *et al.* (2004) The structure of (3R)-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from Pseudomonas aeruginosa. *J Biol Chem* 279(50):52593-52602.

81. Heath RJ & Rock CO (1996) Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in Escherichia coli fatty acid biosynthesis. *J Biol Chem* 271(44):27795-27801.
82. Rabbani A, Finn RM, & Ausio J (2005) The anthracycline antibiotics: antitumor drugs that alter chromatin structure. *Bioessays* 27(1):50-56.
83. Meurer G, *et al.* (1997) Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. *Chem Biol* 4(6):433-443.
84. Kantola J, *et al.* (1997) Folding of the polyketide chain is not dictated by minimal polyketide synthase in the biosynthesis of mithramycin and anthracycline. *Chem Biol* 4(10):751-755.
85. Burson KK, Huestis WH, & Khosla C (1997) Gene shuffling of bacterial aromatic polyketide synthases. *Abstr Pap Am Chem S* 213:83-Biot.
86. Zawada RJ & Khosla C (1999) Heterologous expression, purification, reconstitution and kinetic analysis of an extended type II polyketide synthase. *Chem Biol* 6(9):607-615.
87. Keatinge-Clay AT, *et al.* (2003) Catalysis, specificity, and ACP docking site of Streptomyces coelicolor malonyl-CoA:ACP transacylase. *Structure* 11(2):147-154.
88. Keatinge-Clay AT, Maltby DA, Medzihradszky KF, Khosla C, & Stroud RM (2004) An antibiotic factory caught in action. *Nat Struct Mol Biol*.
89. Korman TP, Hill JA, Vu TN, & Tsai SC (2004) Structural analysis of actinorhodin polyketide ketoreductase: Cofactor binding and substrate specificity. *Biochemistry* 43(46):14529-14538.
90. Korman TP, Tan YH, Wong J, Luo R, & Tsai SC (2008) Inhibition kinetics and emodin cocrystal structure of a type II polyketide ketoreductase. *Biochemistry* 47(7):1837-1847.
91. Hadfield AT, *et al.* (2004) The crystal structure of the actIII actinorhodin polyketide reductase: proposed mechanism for ACP and polyketide binding. *Structure (Camb)* 12(10):1865-1875.
92. Li Q, Khosla C, Puglisi JD, & Liu CW (2003) Solution Structure and Backbone Dynamics of the Holo Form of the Frenolicin Acyl Carrier Protein. *Biochemistry* 42(16):4648-4657.
93. Findlow SC, Winsor C, Simpson TJ, Crosby J, & Crump MP (2003) Solution Structure and Dynamics of Oxytetracycline Polyketide Synthase Acyl Carrier Protein from Streptomyces rimosus. *Biochemistry* 42(28):8423-8433.
94. Jackson DR, *et al.* (2018) Structural and Functional Studies of the Daunorubicin Priming Ketosynthase DpsC. *ACS Chem Biol* 13(1):141-151.
95. Ames BD, *et al.* (2008) Crystal structure and functional analysis of tetracenomycin ARO/CYC: Implications for cyclization specificity of aromatic polyketides. *Proceedings of the National Academy of Sciences* 105(14):5349-5354.
96. Lee MY, Ames BD, & Tsai SC (2012) Insight into the molecular basis of aromatic polyketide cyclization: crystal structure and in vitro characterization of WhiE-ORFVI. *Biochemistry* 51(14):3079-3091.
97. Ames BD, *et al.* (2011) Structural and biochemical characterization of ZhuI aromatase/cyclase from the R1128 polyketide pathway. *Biochemistry* 50(39):8392-8406.
98. Caldara-Festin G, *et al.* (2015) Structural and functional analysis of two di-domain aromatase/cyclases from type II polyketide synthases. *Proc Natl Acad Sci U S A* 112(50):E6844-6851.

99. Thompson TB, Katayama K, Watanabe K, Hutchinson CR, & Rayment I (2004) Structural and functional analysis of tetracenomycin F2 cyclase from Streptomyces glaucescens. A type II polyketide cyclase. *J. Biol. Chem.* 279(36):37956-37963.
100. Kallio P, Sultana A, Niemi J, Mantsala P, & Schneider G (2006) Crystal structure of the polyketide cyclase AknH with bound substrate and product analogue: implications for catalytic mechanism and product stereoselectivity. *J Mol Biol* 357(1):210-220.
101. Sultana A, *et al.* (2004) Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. *The EMBO journal* 23(9):1911-1921.
102. Kroken S, Glass NL, Taylor JW, Yoder OC, & Turgeon BG (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci U S A* 100(26):15670-15675.
103. Persson B, Kallberg Y, Oppermann U, & Jornvall H (2003) Coenzyme-based functional assignments of short-chain dehydrogenases/reductases (SDRs). *Chem Biol Interact* 143-144:271-278.
104. Price AC, Zhang YM, Rock CO, & White SW (2004) Cofactor-induced conformational rearrangements establish a catalytically competent active site and a proton relay conduit in FabG. *Structure (Camb)* 12(3):417-428.
105. Fu H, Ebertkhosla S, Hopwood DA, & Khosla C (1994) ENGINEERED BIOSYNTHESIS OF NOVEL POLYKETIDES - DISSECTION OF THE CATALYTIC SPECIFICITY OF THE ACT KETOREDUCTASE. *Journal of the American Chemical Society* 116(10):4166-4170.
106. Dutler H, Kull A, & Mislin R (1971) Fatty acid synthetase from pig liver. 2. Characterization of the enzyme complex with oxidoreductase activity for alicyclic ketones as a fatty acid synthetase. *Eur J Biochem* 22(2):213-217.
107. Joshi AK & Smith S (1993) Construction, expression, and characterization of a mutated animal fatty acid synthase deficient in the dehydrase function. *J Biol Chem* 268(30):22508-22513.
108. Ostergaard LH, *et al.* (2002) Stereochemistry of catalysis by the ketoreductase activity in the first extension module of the erythromycin polyketide synthase. *Biochemistry* 41(8):2719-2726.
109. Zhang YM & Rock CO (2004) Evaluation of epigallocatechin gallate and related plant polyphenols as inhibitors of the FabG and FabI reductases of bacterial type II fatty-acid synthase. *J Biol Chem* 279(30):30994-31001.
110. Caffrey P (2003) Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. *Chembiochem* 4(7):654-657.
111. Reid R, *et al.* (2003) A model of structure and catalysis for ketoreductase domains in modular polyketide synthases. *Biochemistry* 42(1):72-79.
112. Baerga-Ortiz A, *et al.* (2006) Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase. *Chem Biol* 13(3):277-285.
113. O'Hare HM, Baerga-Ortiz A, Popovic B, Spencer JB, & Leadlay PF (2006) High-throughput mutagenesis to evaluate models of stereochemical control in ketoreductase domains from the erythromycin polyketide synthase. *Chem Biol* 13(3):287-296.
114. Anderson VE & Hammes GG (1984) Stereochemistry of the reactions catalyzed by chicken liver fatty acid synthase. *Biochemistry* 23(9):2088-2094.

115. Marsden AF, *et al.* (1994) Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. *Science* 263(5145):378-380.
116. Siskos AP, *et al.* (2005) Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. *Chem Biol* 12(10):1145-1153.
117. Holzbaur IE, *et al.* (1999) Molecular basis of Celmer's rules: the role of two ketoreductase domains in the control of chirality by the erythromycin modular polyketide synthase. *Chem Biol* 6(4):189-195.
118. Holzbaur IE, *et al.* (2001) Molecular basis of Celmer's rules: role of the ketosynthase domain in epimerisation and demonstration that ketoreductase domains can have altered product specificity with unnatural substrates. *Chem Biol* 8(4):329-340.
119. Starcevic A, Jaspars M, Cullum J, Hranueli D, & Long PF (2007) Predicting the nature and timing of epimerisation on a modular polyketide synthase. *Chembiochem* 8(1):28-31.
120. Fisher M, *et al.* (2000) The X-ray structure of Brassica napus beta-keto acyl carrier protein reductase and its implications for substrate binding and catalysis. *Structure* 8(4):339-347.
121. Rozwarski DA, Vilcheze C, Sugantino M, Bittman R, & Sacchettini JC (1999) Crystal structure of the Mycobacterium tuberculosis enoyl-ACP reductase, InhA, in complex with NAD+ and a C16 fatty acyl substrate. *J Biol Chem* 274(22):15582-15589.
122. Simpson TJ & Weerasooriya MKB (2000) NMR studies of tautomerism in the fungal melanin biosynthesis intermediate 1,3,8-trihydroxynaphthalene. *J Chem Soc Perk T 1* (16):2771-2775.
123. Andersson A, Jordan D, Schneider G, & Lindqvist Y (1996) Crystal structure of the ternary complex of 1,3,8-trihydroxynaphthalene reductase from Magnaporthe grisea with NADPH and an active-site inhibitor. *Structure* 4(10):1161-1170.
124. Andersson A, Jordan D, Schneider G, & Lindqvist Y (1997) A flexible lid controls access to the active site in 1,3,8-trihydroxynaphthalene reductase. *FEBS Lett* 400(2):173-176.
125. Kallberg Y, Oppermann U, Jornvall H, & Persson B (2002) Short-chain dehydrogenase/reductase (SDR) relationships: a large family with eight clusters common to human, animal, and plant genomes. *Protein Sci* 11(3):636-641.
126. Zhang HL, *et al.* (1990) Mutactin, a Novel Polyketide from Streptomyces-Coelicolor - Structure and Biosynthetic Relationship to Actinorhodin. *Journal of Organic Chemistry* 55(5):1682-1684.
127. Javidpour P, Korman TP, Shakya G, & Tsai SC (2011) Structural and biochemical analyses of regio- and stereospecificities observed in a type II polyketide ketoreductase. *Biochemistry* 50(21):4638-4649.
128. Kalaitzis JA & Moore BS (2004) Heterologous biosynthesis of truncated hexaketides derived from the actinorhodin polyketide synthase. *J Nat Prod* 67(8):1419-1422.
129. Tang Y, *et al.* (2006) Structural and functional studies on SCO1815: a beta-ketoacyl-acyl carrier protein reductase from Streptomyces coelicolor A3(2). *Biochemistry* 45(47):14085-14093.
130. Reed MA, *et al.* (2003) The type I rat fatty acid synthase ACP shows structural homology and analogous biochemical properties to type II ACPs. *Org Biomol Chem* 1(3):463-471.
131. Bunkoczi G, Joshi, A., Papagrigoriu, E., Arrowsmith, C., Edwards, A., Sundstrom, M., Weigelt, J., Von Delft, F., Smith, S., Oppermann, U. (2006) Structure of aminoadipate-semialdehyde dehydrigenase-phosphopantetheinyl transferase in complex with cytosolic acyl carrier protein and CoA

132. Zhang YM, Wu B, Zheng J, & Rock CO (2003) Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction. *J Biol Chem* 278(52):52935-52943.
133. Roujeinikova A, *et al.* (2007) Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *J Mol Biol* 365(1):135-145.
134. Leibundgut M, Maier T, Jenni S, & Ban N (2008) The multienzyme architecture of eukaryotic fatty acid synthases. *Curr Opin Struct Biol* 18(6):714-725.
135. Ames BD, *et al.* (2008) Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides. *Proc Natl Acad Sci U S A* 105(14):5349-5354.
136. Dreier J, Shah AN, & Khosla C (1999) Kinetic analysis of the actinorhodin aromatic polyketide synthase. *The Journal of biological chemistry* 274(35):25108-25112.
137. Hicks LM, *et al.* (2006) Investigating nonribosomal peptide and polyketide biosynthesis by direct detection of intermediates on >70 kDa polypeptides by using Fourier-transform mass spectrometry. *Chembiochem* 7(6):904-907.

CHAPTER 2

Polyketide Mimetics Yield Structural and Mechanistic Insights into Product Template Domain Function in Non-Reducing Polyketide Synthases

2.1 Summary

Product template (PT) domains from fungal non-reducing polyketide synthases (NR-PKSs) are responsible for controlling the aldol cyclizations of poly-β-ketone intermediates assembled during the catalytic cycle. They catalytically guide the regioselectivity of ring formation to provide specific products distinct from the thermodynamic outcome of random aldol self-condensations. Our ability to understand the high regioselective control that PT domains exert is hindered by the inaccessibility of intrinsically unstable poly-β-ketones for *in vitro* studies. We describe here the crystallographic application of 'atom replacement' mimetics in which isoxazole rings linked by thioethers mimic the alternating sites of carbonyls in the poly-β-ketone intermediates. We report the 1.8 Å co-crystal structure of the PksA PT domain from aflatoxin biosynthesis with a heptaketide mimetic tethered to a stably modified 4´-phosphopantetheine (PPant), which provides important empirical evidence for a previously proposed mechanism of PT-catalyzed cyclization. Key observations support the proposed deprotonation at C4 of the nascent polyketide by the catalytic His1345 and the role of a protein-coordinated water network to selectively activate the C9 carbonyl for nucleophilic addition. The importance of the 4´-phosphate at the distal end of the pantetheine arm is demonstrated to both facilitate delivery of the heptaketide mimetic deep into the PT active site and anchor one end of this linear array to precisely meter C4

into close proximity to the catalytic His1345. Additional structural features, docking simulations, and mutational experiments characterize protein–substrate mimic interactions, which likely play roles in orienting and stabilizing interactions during the native multi-step catalytic cycle. These findings afford the first view of a polyketide 'atom-replaced' mimetic in a NR-PKS active site that could prove general for other PKS domains.

2.2 Introduction

Product template (PT) domains are a structural feature of the non-reducing class of iterative polyketide synthases (NR-PKSs). They both control the high reactivity of poly-β-ketone intermediates that are rapidly assembled upstream in the β-ketoacyl synthase (KS) domain (1) and catalyze their intramolecular closure to cyclic and fused cyclic products. How both of these tasks are carried out, and how competing self-condensations to more thermodynamically favored products are avoided, are central functional questions of NR-PKS catalysis. We address them here with the single PT domain for which an X-ray crystal structure exists and report the organized structure of a stable mimetic of the native, but synthetically inaccessible, poly-β-ketone substrate bound in the PksA PT domain.

Biosynthesis of environmental carcinogen aflatoxin B_1 is initiated by a fungal, iterative non-reducing polyketide synthase (NR-PKS) known as PksA (Figure 2.1B) (2, 3). This polypeptide consists of six covalently-linked enzyme domains, where the function of each has been characterized by its specialized role in the controlled polymerization of ketide units and cyclization to a particular product (Figure 2.1A) (2, 3).

Polyketide elongation is initiated by the starter-unit:acyl-carrier protein (ACP) (SAT) domain, which selectively primes the ACP with a hexanoyl starter unit (4). The ketosynthase (KS) then catalyzes seven iterative rounds of decarboxylative Claisen

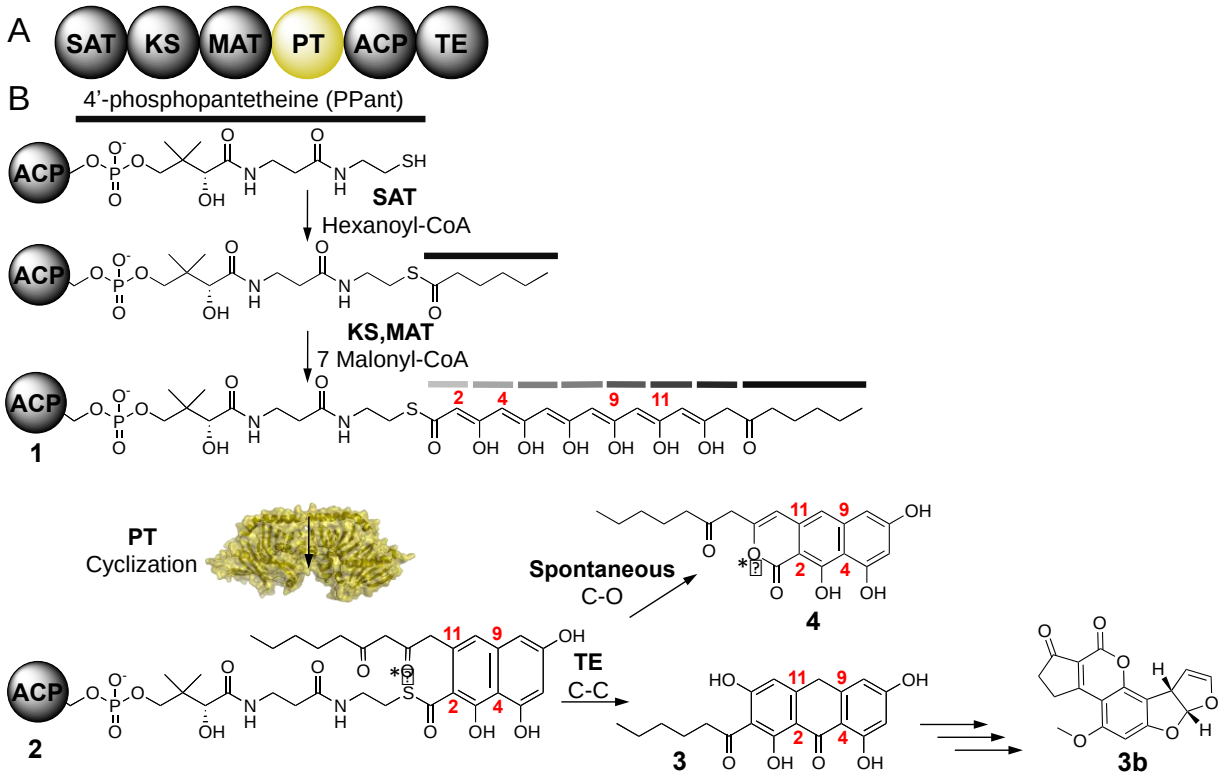


Figure 2.1 Aflatoxin biosynthesis. **(A)** Domain architecture of the aflatoxin NR-PKS. **(B)** Biosynthetic pathway of aflatoxin starts with a hexanoyl starter unit and undergoes seven rounds of elongation to yield a C_{20} linear intermediate **1**. Gray and black bars represent diketone groups. The product template (PT) domain cyclizes a C4-C9 and C2-C11 reaction **2** that is subsequently aromatized and released by the thioester (TE) domain **3**. Further biosynthetic processing yields aflatoxin B_1 **3b**. Absence of the TE yields pyrone product **4**.

condensations using malonyl building blocks supplied by the malonyl-CoA:ACP transacylase (MAT) domain to yield the mature 20-carbon poly-β-keto-intermediate **1** for which direct evidence has been obtained by mass spectrometry (Figure 2.1B) (2, 3). This reactive twenty-carbon intermediate **1** is then cyclized via two regioselective intramolecular aldol reactions by the product template (PT) domain to yield the C4-C9/C2-C11 cyclized intermediate **2** (5). Next, the thioesterase domain (TE) conducts a Claisen/Dieckmann cyclization to release the product from the 4'-phosphopantetheine

(PPant)-tethered ACP to afford noranthrone, **3** (6). The growing polyketide intermediate remains attached to the ACP-tethered PPant group until the release of noranthrone **3** from PksA by a TE domain. In the absence of the TE domain, a spontaneous cyclization gives rise to the pyrone **4** (6, 7). Further post-synthase tailoring of **3** yields aflatoxin B_1 (**3b)** (Figure 2.1B) (2).

The first-ring cyclization (from **1** to **3** in Figure 2.1B) is pivotal to the formation of stable polyketide intermediate **2**, which can be further processed towards the final product **3b**. To date, there are at least six classes of PT domains that differentially cyclize poly-β-ketones to unique ring patterns in a highly selective manner (8, 9). In order to achieve this selectivity, a high degree of control must be maintained between the PT and the ACP-tethered poly-β-ketone intermediate. To date, only the PT domain of PksA has yielded a crystal structure (PDB: 3HRQ, 3HRR) (5). Fortuitously, the structure was obtained with the methyl terminus of palmitate at its deepest recess and extending to the carboxylate in the PT reaction chamber. Using this 16-carbon chain as a template, superimposition of the 20-carbon polyketide intermediate **1** and energy minimization with limited constraints gave a revealing model of PPant–substrate binding and led to a proposal for selective activation for controlled cyclization (5). As unanticipated and appealing as this view of PT function was, it lacked experimental support.

Two problems exist that have blocked our interrogation of PT activity. First, the high intrinsic propensity to self-reaction of poly-β-ketones precludes them as practically useful substrates for structural studies on PKSs. Second, the high reactivity of the native 20-carbon poly-β-keto-intermediate **1** prevents its *de novo* synthesis. Recently, a general strategy for the design and synthesis of stable mimics was developed by using the concept

of ‘atom replacement’ (10). Conceptually, isoxazoles linked by thioethers are visualized to parallel the alternating pattern of methylene and carbonyl groups of NR-PKS intermediates. Furthermore, the native thioester within these substrates was replaced by an isosteric amide bond to further stabilize the pantetheine–substrate connection (10, 11). In this study, we prepared 4´-phosphorylated and unphosphorylated atom-replaced mimetics selected to represent substrate chain lengths of 8–16 carbons units of the corresponding linear and proposed monocyclic intermediates. Here, we report the 1.8 Å co-crystal structure of the PksA PT domain bound to a 4´-phosphopantetheinylated linear heptaketide mimetic in the first crystallographic application of isoxazole-thioether mimetics in PKS structural investigations.

2.3 Results and discussion

2.3.1 Design and rationale of the atom replacement mimetics

To understand the chain-length selectivity of the PksA PT, we designed stable atom-replaced mimics of different chain lengths and tethered them to a non-hydrolizable Pant group (Figure 2.2). Our current set of mimetics includes: a linear ring-opened **4c**-mimetic **5a**, and the corresponding linear isoxazole **4c**-mimetic **5b**; the bis-isoxazole **7c** and **8c**-mimetics **5c**, and **5d** (10)**.** In addition, to probe cyclization specificity, we also prepared the cyclized and aromatized polyketide mimics **5e** and **5f** (10) (Figure 2.2) to mirror hexaketide shunt products of PksA (Figure 2.1). Our selection of these materials was based on the fact that the tetraketides (**5a**-**5b**) with eight carbons are the shortest chain length that a fungal PT has been shown to accept (2). Each of these materials can be

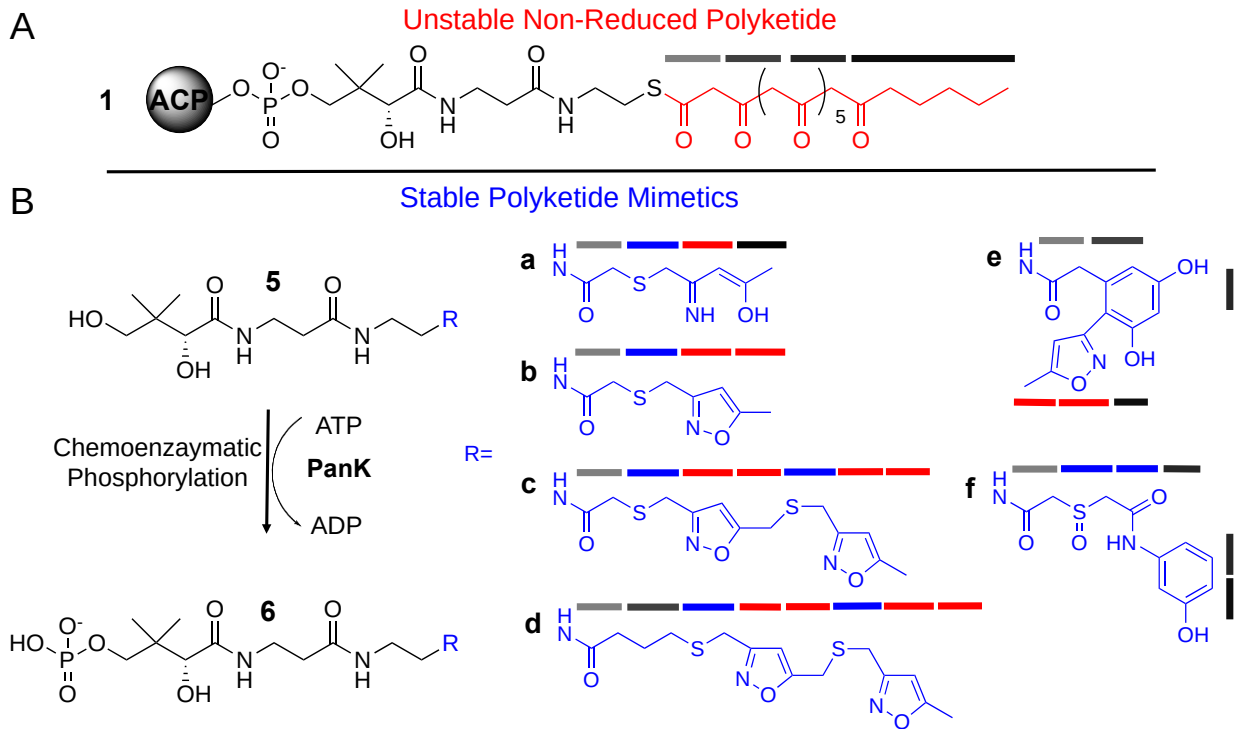


Figure 2.2 Atom replacement mimetics. **(A)** Structure of the ACP tethered natural substrate **1** containing an elongated polyketone tail (red) and PPant unit (black). **(B)** Structures of linear mimetics **5a**-**5d** and cyclic mimetics **5e**-**5f**. The corresponding phosphates **6a**-**6f** can be prepared chemoenzymatically from **5a**-**5d** using the pantetheine kinase CoaA. Various shaded gray bars denote natural ketide units (with the exception of the hexanoyl-starter unit, shown in black). Ketide mimetics containing N and S atom replacement are shown in blue bars and isoxazole derived ketides are shown in red bars.

optionally phosphorylated at the 4´-hydroxyl group of the pantetheine using a promiscuous pantetheine kinase (PanK) (Figure 2.2) as given by the respective conversions **5a**-**5f** into **6a**-**6f** (12-15).

2.3.2 Chemoenzymatic phosphorylation and isolation of pantetheine mimetics

To obtain the phosphorylated mimetics **6a**-**6f**, we incubated **5a**-**5f** with pantetheine kinase (PanK) from the pantothenate (vitamin B5) biosynthetic pathway in *E. coli* (15). Previous work on one-pot chemoenzymatic preparation of coenzyme A analogues provided a standard methodology for pantetheine phosphorylation (15). Since we only required the PanK for pantetheine phosphorylation and not the remaining CoA ligating and modifying enzymes, we first explored

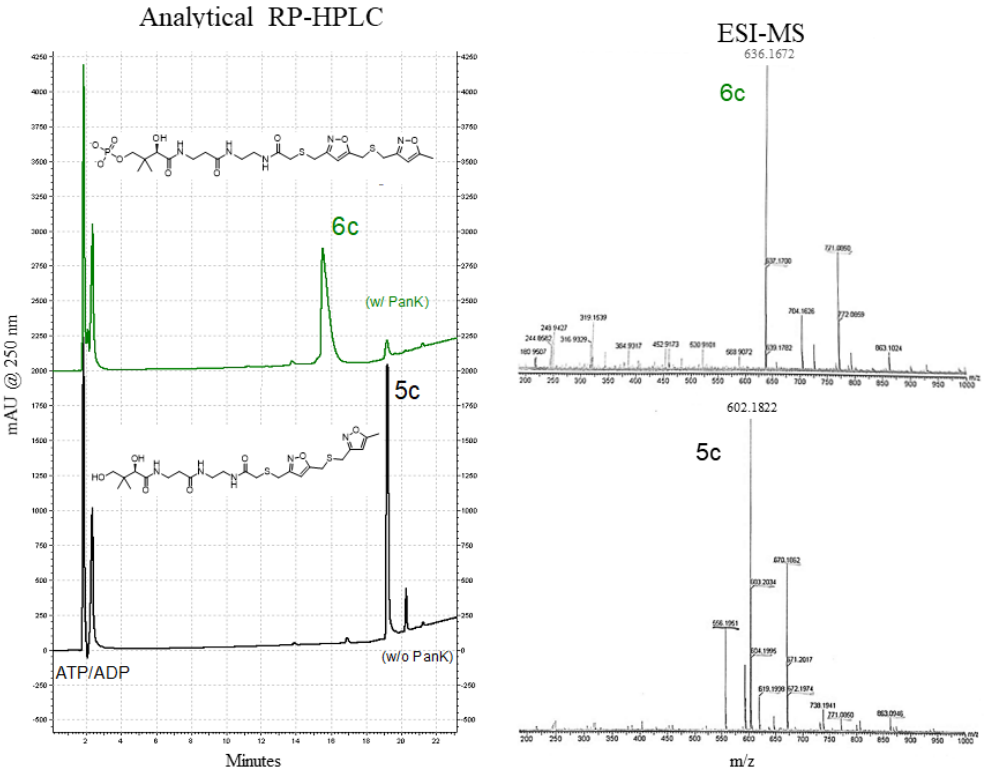


Figure 2.3 Analytical HPLC and ESI-MS analysis of **5c** and phosphorylated **6c**.

optimizing conditions to conduct pantetheine phosphorylation by evaluating varying concentrations of recombinant PanK and mimetics **5a**-**5f** in diverse buffers, pHs, concentrations of ATP, salts and temperatures. The phosphorylation of mimetics **5a**-**5f** into **6a**-**6f** was monitored by analytical HPLC analysis and ESI-MS. Time points were taken at 1, 3, 6, 12 and 24 h. Overall, no significant improvements were found in PanK phosphorylation. We found that the condition giving the highest yield of phosphorylated mimetics **6a**-**6f** was after incubating in 2 µM PanK, 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, and 8 mM ATP at 37 ºC for 3h. Once complete phosphorylation was observed, **6a**-**6f** were purified by semi-preparative HPLC (Figure 2.3 shows HPLC and ESI-MS data for mimetics 6c and 5c). This approach offered an effective protocol to prepare PPant mimics in purity and amounts sufficient for X-ray crystallography and enzyme assays.

2.3.3 The crystallization and crystal diffraction of PksA PT with phosphorylated and unphosphorylated atom-replaced mimetics: importance of chain length and PPant phosphate

To date, the crystallization of ligands with PksA PT had proven to be extremely difficult due to the presence of endogenous palmitate in the active site, which blocked substrate binding. We have found that the best way to perform this procedure involves first the incubation of PksA PT with XAD-2 or XAD-4 hydrophobic resins to remove the palmitate, followed by addition of the substrate or substrate mimetic at a concentration that does not cause immediate protein precipitation (5). To explore for the substrate binding and cyclization motifs of PT, we rigorously screened for conditions that allowed the co-crystallization of PksA PT and other PT constructs mimetics **5a**-**5f** and phosphorylated mimetics **6a**-**6f** (Figure 2.2). For PksA PT, we found that XAD-2 resins often resulted in protein precipitation, and hence, we screened 1200 conditions for each mimetic at 4 and 23 °C. From this set, we selected the most promising conditions and explored it for crystal growth in optimization trays. In addition to PksA PT, we also applied the screen to three other PT domains from NR-PKSs (Pks1, Pks4, and PksA) that represent different clades of PT (8).

2.3.4 Heptaketide-PT structure reveals extensive substrate-enzyme interactions in PPant binding region and cyclization chamber

The **6c**-bound PksA PT structure (**6c** • PksA PT) revealed a dimer with two active sites, with each monomer containing a double hotdog fold (Figure 2.4). Within each monomer the substrate-binding pocket contained a hexyl-binding region, a cyclization chamber and a PPant-binding region (Figure 2.5) (5). Mimetic **6c** interacted with residues in only the cyclization chamber and the PPant-binding region (Figure 2.5C-D). The PPant segment of **6c** served to orient

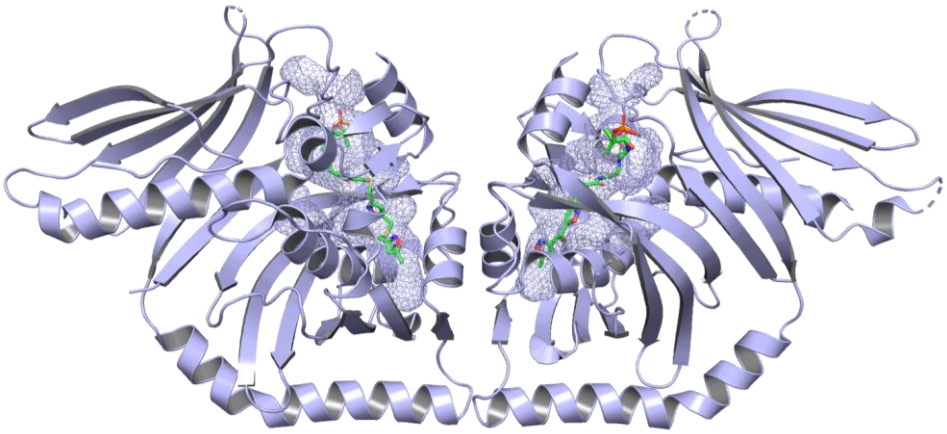


Figure 2.4 Cartoon of **6c** • PksA PT complex. The asymmetric unit contains a PT dimer with both active sites bound with the **6c** phosphorylated heptaketide mimetic.

the heptaketide mimetic by electrostatic and hydrophobic interactions (Figure 2.5C). Here, the terminal PPant phosphate of **6c** is anchored by strong electrostatic interaction with Arg1623 of PT near the entrance of the substrate-binding pocket (Figure 2.5C). Further, the *gem*-dimethyl group of PPant was oriented toward three Val and one Leu near the entrance of the catalytic pocket (Val1567, Val1625, Leu1630 and Val1633). These previously unidentified amino acids defined an important hydrophobic pocket within the PPant binding region (Figure 2.5C). In addition, the backbones of Leu1398 and Cys1353, along with a crystalline water, formed hydrogen bonds with the hydroxyl and carbonyl groups of **6c**. The water molecule was further coordinated to the backbone of Leu1398 and Cys1353, both of which are well conserved among most clades of PT domains (Figure 2.6) (8, 9). These observations suggest that the PPant binding region interacts extensively with the heptaketide substrate, and PPant binding is crucial for substrate orientation.

In the co-crystal structure of **6c** with PksA PT, the hexyl-binding region was largely unoccupied, which is consistent with its selectivity to bind the hexyl starter unit that is not present in **6c** (Figure 2.5B). A 7.0 Å gap between the back of the hexyl-binding region and the methyl

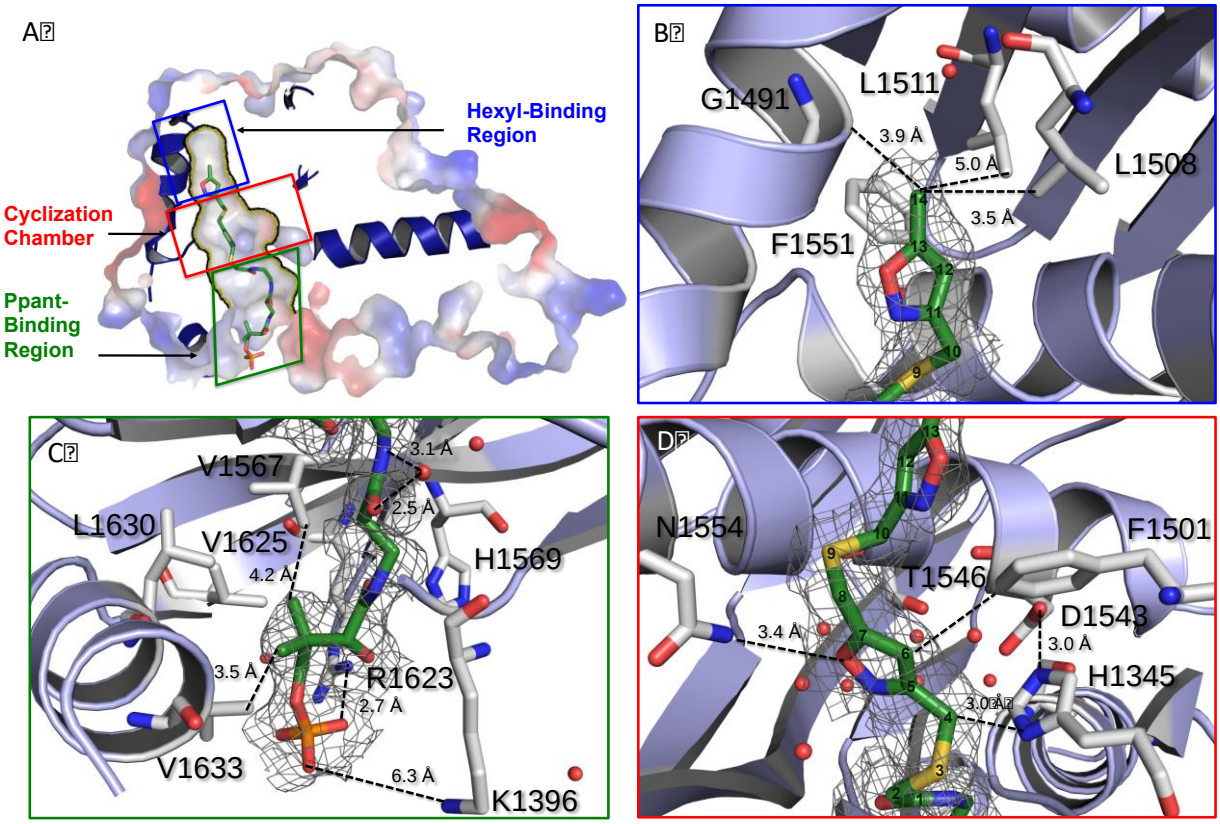


Figure 2.5 Structure of the **6c** • PksA PT complex. **(A)** A surface rendition depicting the hexyl-binding chamber, crystallization chamber and PPant-binding region within the 1.8 Å crystal structure of the **6c** • PksA PT complex. The PPant-tethered heptaketide interacts with various residues in the catalytic pocket. A schematic representation illustrates the close contacts in the **(B)** hexyl-binding region, **(C)** PPant-binding region and **(D)** cyclization chamber. Residue numbers and distances between the substrate and protein are provided.

terminus of the **6c** only contained one detectible water molecule. This observation is in accord with the hydrophobic nature of the hexyl-binding region, such that in the absence of the hexyl group, **6c** binds PT without contacting the hexyl-binding region (Figure 2.5B).

The fact that the structures of **6c** • PksA PT and the previously reported palmitate–PT were highly similar (RMSD = 0.154 Å) also supports the view that the lack of hexyl group in **6c** does not affect overall protein conformation of PksA PT.

The placement of the heptaketide mimetic **6c** in the cyclization chamber confirmed the importance of the cyclization chamber for polyketide orientation and catalysis, as proposed earlier (Figure 2.5D, Figure 4) (5). From the **6c** • PksA PT structure, we identified previously uncharacterized residues in the cyclization chamber that are important for polyketide recognition.

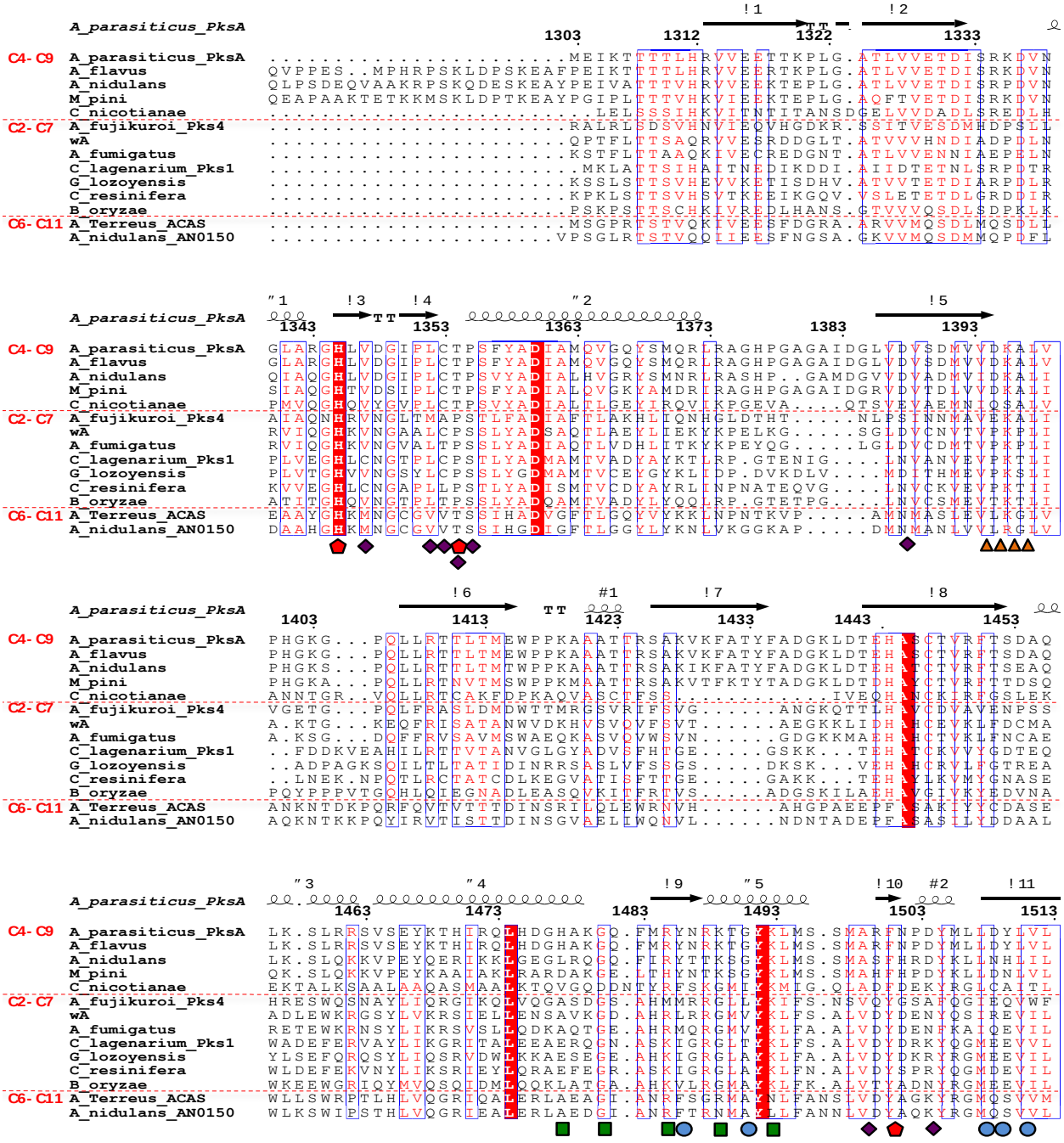
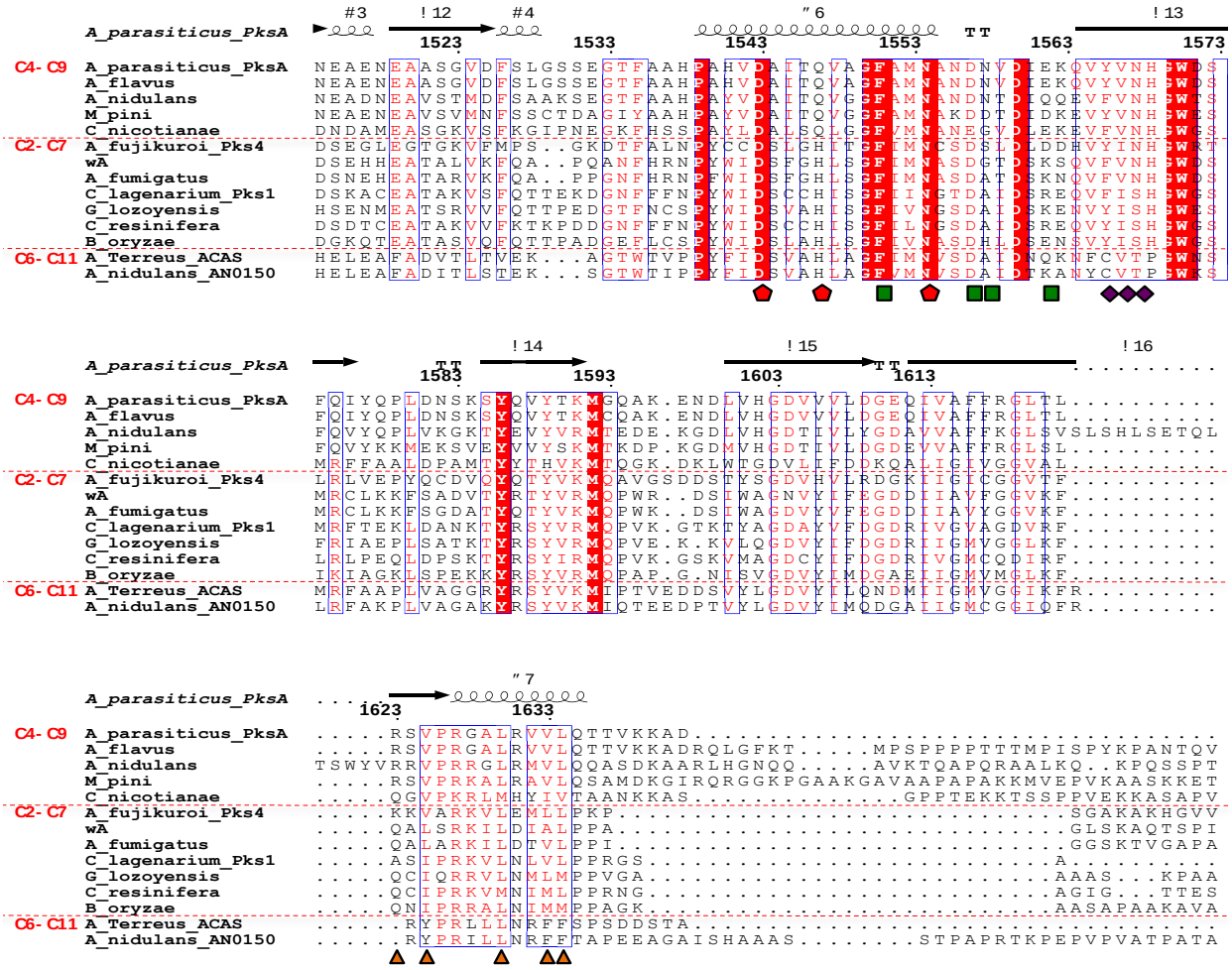


Figure 2.6A Multiple sequence alignment of PT domains. Red pentagons indicate catalytic residues. Orange triangles depict PPant-binding residues. Blue circles represent hexyl-binding residues. Purple diamonds represent cyclization chamber. Green squares depict dimeric interface. N-terminal alignment is on page 11 and C-terminal alignment is on page 12.

The catalytic dyad of PksA PT is composed of His1345 and Asp1543 (5). Regiospecific C4–C9 first-ring formation depends on the ability of His1345 to deprotonate C4 of **1** and initiate first-ring cyclization. In the **6c** • PksA PT structure**,** the catalytic His1345 is 3.0Å from C4 of **6c**, thus



strongly supporting C4 deprotonation and subsequent enolate formation for proper ring closure (Figure 3D). The oxyanion hole was located in the back of the cyclization chamber, constrained by Pro1355. It has been previously proposed that a network of organized water molecules bound to Ser1356, Asp1543, Asn1568 and Thr1546 stabilize the oxyanion and provide a highly structured, polar environment that favors aldol cyclization (Figure 4) (5, 9). Now, the **6c** • PksA PT structure provides concrete evidence for the maintenance of a polar environment through a water network that is, indeed, coordinated with Ser1356, Asp1543, Asn1568 and Thr1546 (Figure 3D). In the bis-isoxazole moiety of **6c** (Figure 2.2), a bridging sulfur atom replaces the C9 carbonyl of the actual polyketide intermediate **1** (Figure 1B). This weaker hydrogen bond acceptor shows no interaction with the water network, but is constrained intrinsically

to a 90° bond angle and was oriented away from the active site by its placement between the two rigid flanking isoxazole rings. Nonetheless, it was remarkable that the location of C4 near His1345 (3.0 Å) and the C9 carbonyl mimetic, the sulfur atom in **6c**, are in close proximity to the water network, which strongly support the proposed mechanism that PT activates a particular set of nucleophile/electrophile partners to catalyze a single regiospecific aldol cyclization between C4-C9.

2.3.5 Catalytic activity and mutagenesis of the PksA PT

We then turned our attention to identify residues important for substrate binding. Using our **6c** • PksA PT structure as a guide, we generated mutants in the PPant binding region, cyclization chamber, and hexyl-binding regions. The PT mutants were reconstituted with *holo*-PksA ACP, the tridomain PksA SAT-KS-MAT, hexanoyl-CoA, and malonyl-CoA. The PksA PT cyclizes the linear intermediate **1** to biosynthesize norpyrone **4** *in vitro*. Products of the enzymatic reactions determined using time course HPLC analysis to monitor the reaction course (Figure 2.8).

***Effects on pocket entrance*:** From these data, we were able to access the importance of Arg1623 and Lys1396, two positively-charged surface residues located at the entrance of the PPant binding region, which interact with the phosphate-terminus of **6c**. R1623A and K1396A mutants were 60% as efficient as wild-type PT in their production of norpyrone **4**, the *in vitro* product of PksA. The K1396Q-R1623Q double mutant also displayed a similar efficacy (60% efficiency as compare to wild-type PT). The moderate decrease in norpyrone formation for R1623A, K1396A and K1396Q-R1623Q indicates that the positively-charged side chains of Arg1623 and Lys1396 were not essential for product formation. Substitution of the negatively-charged residue, R1623E, did show further reduction of activity and generated 30% production of norpyrone. However, the mutational

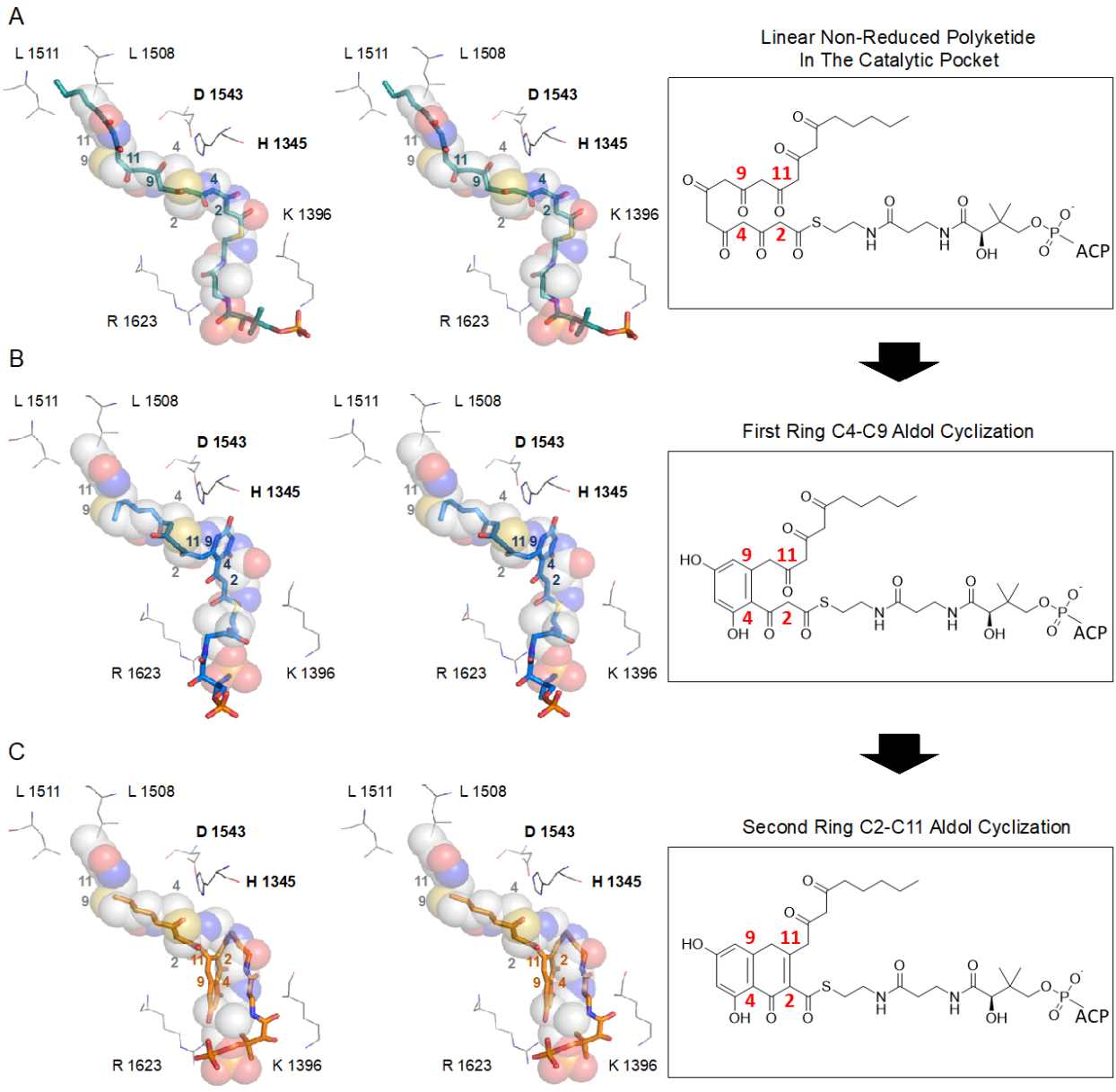


Figure 2.7 Docking studies. Stereoimages of **(A)** the docked linear polyketide intermediate (**1**), **(B)** the monocyclic C4-C9 intermediate and **(C)** the bicyclic C4-C9/C2-C11 (**2**). The docking ligands were rendered in stick format and heptaketide ligand in space filled representations. Colored numbers indicate the carbon registry of the docked linear polyketide intermediate while the gray numbers indicate the carbon registry for the **6c** ligand.

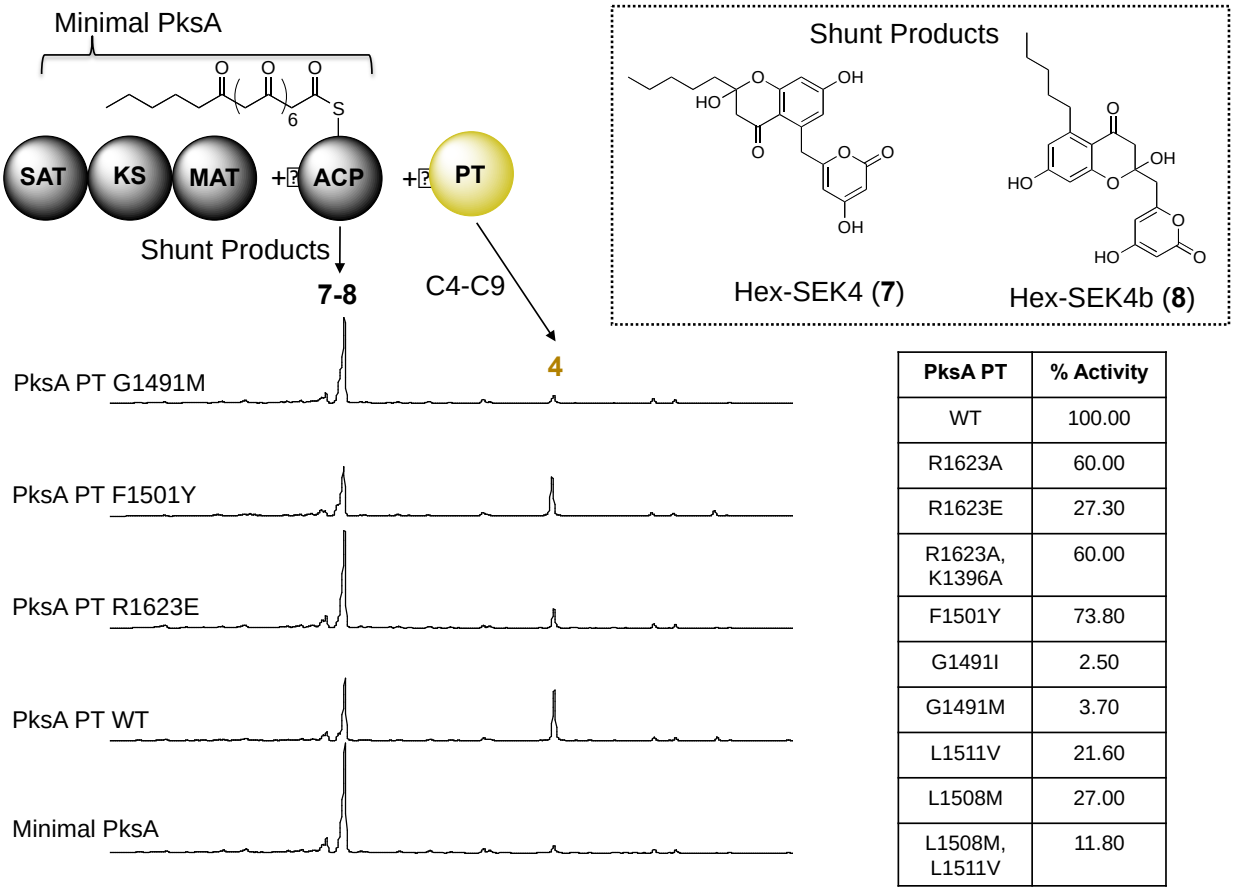
result showed that the positive charges of Arg1623 and Lys1396 are not the sole determinant for substrate binding. These observations were consistent with the **6c** • PksA PT structure where the PPant binding region interacts with PPant with both charge-charge and hydrophobic interactions.

***Interactions within the cyclization chamber*:** The Ala mutants of the active site dyad His1345 and Asp1535 completely abolished catalyzed norpyrone **4** formation (5). The result confirmed the crucial catalytic role of both residues for cyclization. Previous Ala mutations of Thr1546 and Asn1548 also resulted in a diminished activity (5), which is consistent with the observed interaction between **6c** and residues in the cyclization chamber (Figure 2.5D).

Interactions with the hexyl-binding region. Sequence alignment suggests that most PT domains, which typically incorporate shorter acetyl starter units, have a Met at the position Leu1508, located near the beginning of the hexyl-binding region. This observation suggested that the methionine residue may serve to block the hexyl-binding region and prevents the loading of a hexanoyl starter unit. Further, Gly1491 defines the bottom of the pocket, which is otherwise a bulky, hydrophobic residue in the non-hexyl loading PTs. To decrease the volume of the hexyl-binding region, we generated G1491I and G1491M. Both of these mutants returned less than 5% of the production of **4** when compared to of wild-type (5). The fact that truncation of the hexyl-binding region can greatly reduce product formation for PksA PT is in accord with the proposal that the overall size and shape of the substrate binding pocket was important in chain-length selectivity.

2.3.6 PT-polyketide docking simulations are consistent with the cocrystal structure

Each isoxazole moiety (Figure 2.2) is aimed to mimic a β-diketone/keto-β-enol in a linear polyketide (Figure 2.2). While the isoxazole ring within **5a-6f** serves to provide stability, it also restricts the number of potential enolization isomers sampled by the natural poly-β-ketone in **1**. We conducted *in silico* docking using GOLD (16) to further explore the potential issues arising from the selection of distinct enol-keto tautomers of **1**. Hexanoyl-containing polyketide intermediates with 14- and 20-carbons were docked into the substrate binding pocket of



PksA PT	% Activity
WT	100.00
R1623A	60.00
R1623E	27.30
R1623A, K1396A	60.00
F1501Y	73.80
G1491I	2.50
G1491M	3.70
L1511V	21.60
L1508M	27.00
L1508M, L1511V	11.80

Figure 2.8 *In vitro* reconstitution analysis of PksA PT WT and mutants. The minimal PKS consist of the basic enzymes required to generate linear polyketide products, which include the SAT-KS-MAT and the ACP. Minimal PKS generate various shunt products **7** or **8**. Addition of the PT generates a C4-C9/C2-C11 norpyrone **4**. PksA PT mutants display a decrease in **4** formation when compared to wild type. Mutants include: R1623A, R1623E, F1501Y, G1491I, G1491M, L1511V, L1508M and double mutants R1623A, K1396A, L1508M, L1511V.

PksA PT. We found that the binding motifs of the native mature 20-carbon or nascent 14-carbon intermediates are highly similar to the co-crystal structure of mimetic **6c** (Figure 2.7A), especially in the following features. First, the electrostatic interactions between the PPant phosphate and Arg1623/Lys1396 remained in close proximity, further supporting the importance of the PPant phosphate for substrate localization and binding. Second, the location and orientation of C4 and its proximity to the catalytic His1345 was highly similar in the docked and observed **6c** • PksA PT structure, supporting the view that the key residue for deprotonation is correctly placed in the co-crystal structure of **6c** • PksA PT (Figure 2.7A). Finally, in the catalytic chamber and hexyl-binding

region, the 20-carbon intermediate and 14-carbon mimetic had highly similar interactions (both polar and non-polar) with the active site residues, especially surrounding the carbonyl moieties flanking C4 and C2 that will be deprotonated to initiate the first- and second-ring cyclizations. The overall orientation of the PPant linked to the endogenous 20-carbon unreduced polyketide kinks inside the 30 Å PksA PT pocket. The previously identified Val/Leu PPant channel residues (5) along with the Arg/Lys surface residues help orient C4 toward His1345 (Figure 2.7A). Deprotonation and enolate formation at C4 by His1345, followed by aldol condensation at C9 generates the first C4–C9 ring intermediate (Figure 2.9).

In this mechanism, the C4–C9 ring is directed towards the cyclization chamber, allowing for a second His1345 deprotonation at C2, which undergoes reaction at C11. The second aldol

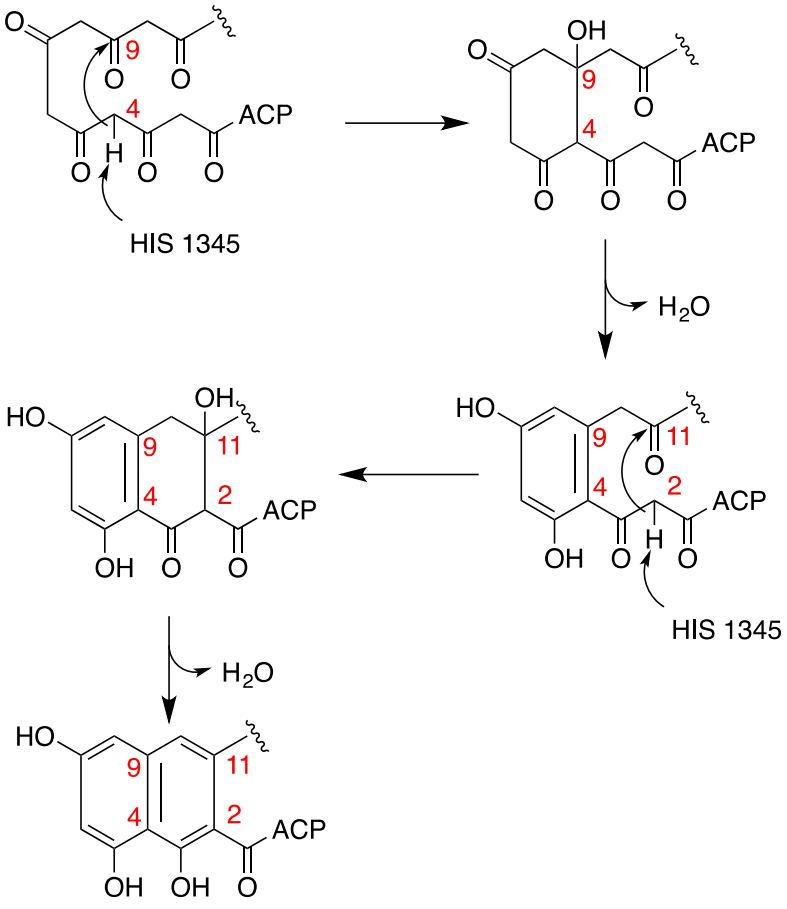


Figure 2.9 Proposed conversion of the linear polyketide **1** to **2** by the PksA PT domain.

reaction of C2–C11 generates intermediate **2**, which is displaced further into the cyclization chamber (Figure 2.7B-C). The proposed position of the bicyclic intermediate **2** in the cyclization chamber is further supported by the similar position observed for the HC8 cyclization analog used in the first PksA PT domain structure analysis (5). Overall, both the mature 20-carbon and nascent 14-carbon substrates align in orientations similar to the heptaketide mimetic **6c**, thus confirming that the 'atom-replacement' mimics are indeed a valid mimetic for the catalysis, specificity and mechanism of the PKS PTs.

2.4 Conclusions

We report the X-ray crystallographic application of atom-replacement mimetics to elucidate the cyclization specificity of fungal NR-PKS PT domains. Using a combination of structural biology, structure-based mutagenesis, and enzymatic assays, we have shown how these mimetics help identify key residues involved in substrate recognition that are not otherwise identifiable. The inherent instability and reactivity of linear, unreduced polyketide intermediates has hindered previous efforts to accurately visualize the interaction between a poly-β-ketone intermediate and a PKS. While no structure can exactly mimic the chemical properties of natural substrate, the atom-replacement mimetics have facilitated our ability to structurally visualize these linear intermediates in a PT domain by providing the polyketide mimicry while blocking their intrinsic reactivity. We demonstrate that attachment of the PPant moiety to atom-replaced mimetics can allow for further structural insight into the role of the PPant arm on the delivery and accurate placement of the reactive polyketide intermediate in the active site. The **6c** • PksA PT co-crystal structure provides critical visual confirmation for the role of

PT domains in regiospecific cyclization(s) during fungal polyketide biosynthesis. In these studies, the 14-carbon mimetic **6c** delivered high-resolution structural data that further validated the proposed substrate-binding residues. The binding region proximal to the PPant arm in **6c** contains several key residues previously unidentified that are responsible for coordinating the hydroxyl and stabilizing the *gem*-dimethyl moieties of the PPant arm. The nature of **6c**'s binding to catalytic chamber provided further evidence for residue flexibility to accommodate interactions with the rigid isoxazole rings. In an important observation, the heptaketide C4 carbon is positioned proximal to the catalytic His1345, providing further experimental evidence to the previously proposed mechanism of polyketide cyclization by PT. Such atom replacement mimetics can serve as powerful tools to further investigate distinct Type I, II and III PKS enzymes that would otherwise be inaccessible by conventional polyketides, which would be subject to rapid aldol and Claisen condensations in media used for X-ray crystallography. Overall, this study demonstrates that medicinal chemical techniques, such as structure mimicry, can have an immediate impact in the study of natural product biosynthesis. We found that the atom replacement concept not only offers sufficiently stable materials, but also offers materials with sufficient mimicry to support high-resolution analyses of structural biology.

2.5 Materials and methods

2.5.1 Protein expression and purification

The recombinant PksA PT monodomain with an *N*-terminal His_6-tag were produced in BL21 (DE3) *E. coli* cells (Novagen). Cells containing the PksA PT monodomain plasmid were grown to OD_{600} = 0.6 at 37 °C in LB medium containing 50µg/mL kanamycin. The cell cultures were cooled to 18 °C and expression was induced using 1mM IPTG. The cell cultures were incubated for an additional 16 h at 18 °C and harvested by centrifugation at 5,525 r.c.f. for 15 min. The cell pellets were resuspended in 50 mM Tris-HCl pH 7.5, 10% glycerol, 10 mM imidazole, 300 mM NaCl and 1mg/ml lysozyme. Resuspended cells were cooled on ice for 30 min and the cells were disrupted using sonication. The cell debris was cleared by centrifugation at 21,036 r.c.f. for 1 h. The supernatant was collected and batch bound to HisPur™ Cobalt Resin (Thermo Scientific) for 1 h at 4 °C. PksA PT was purified according to the manufacturer's instructions using an imidazole step-gradient. Fractions containing pure protein were determined by SDS-PAGE and fractions containing PksA PT were combined and dialyzed against 50 mM Tris-HCl pH 7.5, 10% glycerol, 300 mM NaCl at 4 °C for 12 h. Removal of the *N*-terminal His_6-tag was conducted by incubating the dialyzed PksA PT at 18 °C for 36 h with thrombin from bovine plasma (Sigma-Aldrich) at a concentration of 2 U/mg of PksA PT protein and 3.5mM $CaCl_2$. Removal of thrombin and further purification of PksA PT was conducted by anion exchange chromatography using HiTrap Q FF (GE Healthcare) according to the manufacturer's instructions. Purified PksA PT was dialyzed against crystallization buffer, which consisted of 25 mM Tris-HCl pH 7.5, 5% glycerol, 1 mM dithiothreitol for crystallographic studies.

2.5.2 Site-directed mutagenesis

The online program PrimerX was used to design primers for all PT mutants using the previously described cloned pET28b PT (5, 17). The following primers were used for mutagenesis. All mutations were confirmed though automated DNA sequencing.

Table 2.1 Mutagenic Primer Sequences.

Mutagenesis Primer	Sequence 5'-3'
G1491I_F	GGAAGACCATTTACAAGCTCATGAGCAGCATGGC
G1491I_R	CTTGTAAATGGTCTTCCTATTGTATCGCATGAACTGTCC
G1491L_F	GAAGACCCTGTACAAGCTCATGAGCAGCATG
G1491L_R	GAGCTTGTACAGGGTCTTCCTATTGTATCGCATG
G1491M_F	GACCATGTACAAGCTCATGAGCAGCATGGC
G1491M_R	GCTTGTACATGGTCTTCCTATTGTATCGCATGAAC
R1623A_F	CGCTGGCGTCAGTTCCTCGTGGTGC
R1623A_R	GAACTGACGCCAGCGTAAGGCCGCG
R1623E_F	CGCTGGAATCAGTTCCTCGTGGTGC
R1623E_R	AACTGATTCCAGCGTAAGGCCGCGG
K1396A_F	GGACGCGGCGTTGGTTCCCCATGGG
K1396A_R	CCAACGCCGCGTCCACCACCATGTC
F1501Y_F	CGGTATAATCCCGACTACATGCTCC
F1501Y_R	CGGGATTATACCGAGCCATGCTG
L1511V_F	GATTATGTGGTGCTGAACGAAGCAG
L1511V_R	CAGCACCACATAATCTAGGAGCATG
L1508M_F	CATGCTCATGGATTATCTGGTGCTG
L1508M_R	GATAATCCATGAGCATGTAGTCGGG

2.5.3 Protein x-ray crystallography

The native crystals for PksA PT were grown as previously described by Crawford *et al*. (5). The crystal structures for PksA PT showed a distinct electron density for the C_{16} fatty acid, palmitate, which presumably bound to the enzyme during overproduction and purification from *E. coli*. Palmitate was removed using 1:1 mixture of the lipid-binding resins, XAD-2 and XAD-4 Amberlite resins, followed by immediate incubation with the phosphopantetheine analogues with PT (10:1 ratio). XAD-2 and XAD-4 resin mixture was activated with methanol, washed with deionized water and equilibrated with crystallization buffer. Mimetics bound to PksA PT were

obtained by treating PksA PT with activated XAD resins for 2 h followed by addition of the mimetics. Co-crystals of PT grown in the same condition as the native PksA PT, 0.22 M ammonium acetate, 20% PEG 3350 overnight at 25°C. The crystals were transferred to a well solution containing the crystal well solution buffer with 30% PEG 3350 and flash frozen in liquid nitrogen. Single wavelength data from native crystals were collected to 1.74 Å on a CCD Q315 detector at beamline 8.2.1 of the Advance Light Source in the Lawrence Berkeley National Laboratory. Data was collected at a 0.99997 Å wavelength with a total of 180 frames at 0.5° oscillations. The space group for the heptaketide co-crystal structure was $P2_12_12_1$ with cell dimensions of a = 89.37 Å, b = 90.53 Å, c = 90.83 Å, $\alpha = \beta = \gamma = 90.00$. Phases were determined by molecular replacement using the SeMet PT model, as described previously (5). Electron density maps were generated using CNS, Phaser MR from PHENIX and the model was build using COOT (18). Refinement was carried out by CNS, PHENIX Refine and edited by visual inspection (19). Phosphopantetheine ligands were generated using the Elbow program in PHENIX, and the ligands were placed and validated using PHENIX LigandFit. The refined structures were evaluated with PROCHECK before submission (20). Structural analysis and figure generation used in the manuscript were done using PyMol (21). A summary of the crystallographic statistics is shown (Table 2.2).

Table 2.2 Crystallographic Statistics for **6c** • PksA PT Cocrystal Structure.

	PksA PT bound to **6c**
Data collection	
Wavelength (Å)	0.99997
Space group	P $2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	89.37, 90.53, 90.84
α, β, γ (°)	90, 90, 90
Resolution (Å)	50.00 – 1.74 (1.83 – 1.74)
R_{merge}	0.10 (0.53)
$I/\sigma(I)$	19.2 (2.57)
CC*	0.974
Completeness (%)	99.6 (95.2)
Redundancy	3.8 (3.4)
Refinement	
Resolution (Å)	45.42 – 1.80 (1.86 – 1.80)
No. reflections	68331 (6508)
R_{work}	0.1899
R_{free}	0.2189
No. atoms	
Protein	4923
Ligands	82
B factors	
Protein	27.40
Ligands	37.60
Water	34.10
R.m.s. deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.46

2.5.4 *In silico* docking

The docking program GOLD was used for docking between the PksA PT and mimetic **5a-6f** (22). Both protein and ligand were prepared for docking by removing waters, adding hydrogens, charges and converting the program Chimera (23). The PksA PT ligand-binding pocket was

defined as all residues within 15 Å of the nitrogen atom on His 1345. Docking was performed using the default settings with 100 docking trials performed. The docking solutions were ranked using the ChemPLP scoring functions.

2.5.5 Preparation and phosphorylation of the mimetics

The synthesis, purification and characterization (by NMR and HRMS) of the unphosphorylated AR mimetics **5a**-**5f** have been previously reported (10). Phosphorylation of **5a**-**5f** to **6a**-**6f**, respectively, was conducted biosynthetically using the first pantetheine kinase (PanK) from the pantothenate (vitamin B5) pathway (14, 15). PanK is promiscuous at catalyzing the phosphorylation of pantetheine and pantetheine analogs derived at the 4'-hydroxyl group.

General phosphorylation procedure: An aliquot of an atom-replacement mimetic was diluted into 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 μM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, the reaction was filtered using a 3,000 MWCO centrifugal filter (Millipore). The phosphorylated atom replaced mimetics were successfully separated by reverse phase liquid chromatography (Beckman Coulter HPLC System Gold) using a 5μ × 10mm × 15cm C-18 ODS column (Beckman Coulter Ultrasphere) with UV detection at 230-310 nm. Two solvents were used as given by CH_3CN (solvent A) and 0.1% HCO_2H in H_2O. The column was allowed to run at a flow rate of 4.75 mL/min using a linear gradient of 5-50% solvent A (95-50% solvent B) over 15 min followed by 50-95 % solvent A (50-5% solvent B) over 5 min. Under these conditions, the HPLC retention times were observed between 4-10 min, depending on the mimetic. Phosphorylated products were collected manually in a round-bottomed flask and subjected to rotary evaporation. The dried phosphorylated products were re-suspended in HPLC grade DMSO and stored at -20 ˚C. Samples were diluted in HPLC

grade CH_3OH and evaluated by ESI-LC-TOF (Micromass LCT 3). Briefly, 10 µL were injected and an *m/z* range of 200-1000 Da in negative mode was selected for data analysis. Results were analyzed using Masslynx 4.0 (Waters Mass Lynx 4.0 software.). The phosphorylated mimetics **6a-6e** were stored at 100 mM in DMSO at -80 °C.

*Phosphorylated-mimetic **6a**.* A 300 µL aliquot of 100 mM **5a** (13.4 mg, 29.9 µmol) was diluted into 12 mL of 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 µM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, mixture as purified by centrifugal filtration, dried, and subjected to HPLC purification (see general procedure) provided 7.2 mg (46%) of phosphorylated mimetic **6a**; HRMS (*m/z*): $[M+H]^-$ calcd for **6a** 527.16 (100%) and 528.15 (5%); found 528.5.

*Phosphorylated mimetic **6b**.* A 300 µL aliquot of 100 mM **5b** (12.9 mg, 30 µmol) was diluted into 12 mL of 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 µM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, mixture was purified by centrifugal filtration, dried, and subjected to HPLC purification (see general procedure) provided 6.9 mg (44%) of phosphorylated mimetic **6b**; HRMS (*m/z*): $[M+H]^-$ calcd for **6b** 509.15; found 509.19.

*Phosphorylated mimetic **6c**.* A 350 µL aliquot of 100 mM **5c** (19.5 mg, 35 µmol) was diluted into 14 mL of 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 µM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, mixture as purified by centrifugal filtration, dried, and subjected to HPLC purification (see general procedure) provided 12.5 mg (45%) of phosphorylated mimetic **6c**; HRMS (*m/z*): $[M+H]^-$ calcd for **6c** 636.16; found 636.16.

*Phosphorylated mimetic **6d***. A 300 µL aliquot of 100 mM **5d** (17.5 mg, 29.9 µmol) was diluted into 12 mL of 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 µM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, mixture as purified by centrifugal filtration, dried, and subjected to HPLC purification (see general procedure) provided 13.6 mg (68.4%) of phosphorylated mimetic **6d**; HRMS (*m/z*): $[M+H]^-$ calcd for **6d** 664.19; found 664.21.

*Phosphorylated mimetic **6e***. A 300 µL aliquot of 100 mM **5e** (14.7 mg, 29.9 µmol) was diluted into 12 mL of 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 µM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, mixture as purified by centrifugal filtration, dried, and subjected to HPLC purification (see general procedure) provided 9.3 mg (55%) of phosphorylated mimetic **6e**; HRMS (*m/z*): $[M+H]^-$ calcd for **6e** 571.18; found 571.13.

*Phosphorylated mimetic **6f***. A 300µL aliquot of 100 mM **5f** (15.0 mg, 29.9µmol) was diluted into 12 mL of 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 µM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, mixture as purified by centrifugal filtration, dried, and subjected to HPLC purification (see general procedure) provided 8.4 mg (49%) of phosphorylated mimetic **6f**; HRMS (*m/z*): $[M+H]^-$ calcd for **6f** 579.15; found 579.11.

2.5.6 *In vitro* reconstitution assay

To measure enzyme activity of both WT and mutant PksA PT constructs, we utilized the PksA system from *A. parasiticus* and performed *in vitro* reconstitution assays similar to that described by Newman *et al*.(7). Reaction mixtures (500 µL) containing 10 µM PksA SAT-KS-

MAT tridomain, 50 µM PksA ACP monodomain, 50 µM WT or mutant PksA PT monodomain, 5 mM malonyl-CoA and 0.5mM hexanoyl-CoA in 100mM potassium phosphate pH. 7.0, 1 mM Tris, 1 mM (2-carboxyethyl)phosphine (TCEP) and 10% glycerol were prepared. Reaction mixtures were incubated in the dark at 25 °C for 6 h. The reactions were quenched with 10 µL HCl and extracted three times by the addition of 200 µL EtOAc. The extractions were dried with sodium sulfate (Na_2SO_4 anhydrous), evaporated under vacuum and resuspended in 80% water, 20% CH_3CN at a final volume equivalent to the initial reaction volume. Products were analyzed by using reverse-phase high-performance liquid chromatography (HPLC) (Beckman Coulter). Two solvents were used as given by CH3CN (solvent A) and 0.1% HCO_2H in H_2O. Injections (100 µL) were separated on a Phenomenex Prodigy ODS3 column (4.6mm x 250mm, 5µm6 mm × 250 mm, 5 µm) by running a gradient from 5% to 85% solvent A (95% to 15% solvent B) over 30 min followed by 85% to 95% solvent A (15% to 5% solvent B) over 5 min at a flow rate of 1.0 mL/min. Product elution was monitored by diode-array detection at 280 nm.

2.5.7 Circular dichroism (CD)

All samples, both mutant and WT were prepared by diluting protein to 0.2 mg/mL in 20 mM Tris-HCl (pH 7.5). The CD data was collected using a Jasco J810 CD spectropolarimeter. Spectral scans were collected at 20°C from 190 nm to 260 nm using 0.5 nm steps with 5 repeats.

2.8 Author contribution

C.A. Townsend, M.D. Burkart and S.C. Tsai oversaw the design, execution and preparation of this manuscript; A. L. Vagstad provided the protein constructs. J. F. Barajas, and C.L. Topper were responsible for the expression, purification and crystallization studies; J. F. Barajas and G.

O. Moreno conducted chemoenzymatic phosphorylation and purification. G. O. Moreno assisted in data and figure preparation. J. F. Barajas generated homology models, sequence alignments and PT mutants, as well as the *in silico* docking and *in vitro* reconstitution analysis; and G. Shakya, H. Rivera Jr. and J. J. La Clair synthesized the mimetics used within this program.

2.8 References

1. Vagstad AL, Bumpus SB, Belecki K, Kelleher NL, & Townsend CA (2012) Interrogation of global active site occupancy of a fungal iterative polyketide synthase reveals strategies for maintaining biosynthetic fidelity. *J Am Chem Soc* 134(15):6865-6877.
2. Townsend CA (2014) Aflatoxin and deconstruction of type I, iterative polyketide synthase function. *Nat Prod Rep* 31(10):1260-1265.
3. Crawford JM*, et al.* (2008) Deconstruction of iterative multidomain polyketide synthase function. *Science* 320(5873):243-246.
4. Crawford JM, Dancy BC, Hill EA, Udwary DW, & Townsend CA (2006) Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc Natl Acad Sci U S A* 103(45):16728-16733.
5. Crawford JM*, et al.* (2009) Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature* 461(7267):1139-1143.
6. Korman TP*, et al.* (2010) Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis. *Proc Natl Acad Sci U S A* 107(14):6246-6251.
7. Newman AG, Vagstad AL, Storm PA, & Townsend CA (2014) Systematic domain swaps of iterative, nonreducing polyketide synthases provide a mechanistic understanding and rationale for catalytic reprogramming. *J Am Chem Soc* 136(20):7348-7362.
8. Li Y, Image, II, Xu W, Image I, & Tang Y (2010) Classification, prediction, and verification of the regioselectivity of fungal polyketide synthase product template domains. *J Biol Chem* 285(30):22764-22773.
9. Xu Y*, et al.* (2013) Rational reprogramming of fungal polyketide first-ring cyclization. *Proc Natl Acad Sci U S A* 110(14):5398-5403.
10. Shakya G*, et al.* (2014) Modeling linear and cyclic PKS intermediates through atom replacement. *J Am Chem Soc* 136(48):16792-16799.
11. Gehring AM, Lambalot RH, Vogel KW, Drueckhammer DG, & Walsh CT (1997) Ability of Streptomyces spp. acyl carrier proteins and coenzyme A analogs to serve as substrates in vitro for E. coli holo-ACP synthase. *Chemistry & biology* 4(1):17-24.
12. Haushalter RW, Worthington AS, Hur GH, & Burkart MD (2008) An orthogonal purification strategy for isolating crosslinked domains of modular synthases. *Bioorg Med Chem Lett* 18(10):3039-3042.

13. Ishikawa F, Haushalter RW, Lee DJ, Finzel K, & Burkart MD (2013) Sulfonyl 3-alkynyl pantetheinamides as mechanism-based cross-linkers of acyl carrier protein dehydratase. *J Am Chem Soc* 135(24):8846-8849.
14. Worthington AS & Burkart MD (2006) One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org Biomol Chem* 4(1):44-46.
15. Nazi I, Koteva KP, & Wright GD (2004) One-pot chemoenzymatic preparation of coenzyme A analogues. *Anal Biochem* 324(1):100-105.
16. Jones G, Willett P, & Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J Mol Biol* 245(1):43-53.
17. Lapid C (2003) PrimerX -Automated design of mutagenic primers for site-directed mutagenesis.
18. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
19. Echols N*, et al.* (2014) Automating crystallographic structure solution and refinement of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 70(Pt 1):144-154.
20. Laskowski RA (2001) PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res* 29(1):221-222.
21. L S (2003) The PyMOL Molecular Graphics System), 1.8.
22. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, & Taylor RD (2003) Improved protein-ligand docking using GOLD. *Proteins* 52(4):609-623.
23. Huang CC, Couch, G.S., Pettersen, E.F., and Ferrin, T.E. (1996) Chimera: An Extensible Molecular Modeling Application Constructed Using Standard Components. ed UCSF (Pacific Symposium on Biocomputing), p 724.

CHAPTER 3

Probing the Specificities of Type II Polyketide Ketoreductases Using Isosteric Polyketide Mimetic Probes

3.1 Summary

Polyketides are a class of natural products. During polyketide biosynthesis, the polyketide stereocenters are introduced upon ketoreduction, which is catalyzed by the NAD(P)H dependent ketoreductase (KR)(1). The KR domain has been extensively characterized in type I modular PKS (2-7). In contrast, the regio- and stereo-specificities of type II PKS KRs are less well understood. The actinorhodin KR (ActKR) specifically reduces the C-9 carbon of an octaketide, and current evidence suggests that KR is also involved in first ring cyclization (8-11). In addition, the ActKR has an unexplained substrate specificity towards longer substrates (pentaketide and octaketide), but not tetraketide (12). The instability of native polyketide substrates has been the major roadblock towards understanding the molecular basis of KR specificity. To probe ActKR activity and specificity, we developed a novel strategy that selectively replaces the reactive carbonyl groups of a polyketide substrate to more stable, isosteric chemical moieties such as C, N, S or O (13). This atom replacement approach has resulted in great success to probe the specificities of type II PKS (14). We herein report two cocrystal structures of a fully-active ActKR mutant (H153Y H201G) bound with the atom-replaced mimetics corresponding to tetraketide and octaketide substrates. The octaketide cocrystal structure affords a first glimpse into the importance of phosphopantetheine for chain-length control and substrate orientation in the type II KRs.

3.2 Introduction

Polyketides are an important class of secondary metabolites produced by a family of enzymatic assembly lines called polyketide synthases (PKSs) that are found in bacteria, plants and fungi. Polyketide natural products are characterized by their rich structural diversity, and they are an important source of bioactive pharmaceuticals. PKSs are related to fatty acid synthases (FASs), and they share the same biosynthetic logic centering on the growth of a carbon chain from simple activated carbon units, such as malonyl-CoA, through decarboxylative Claisen condensation and subsequent chain modifications. Each condensation reaction leaves a ketone group in the β position of the growing chain. In general, structural complexity is created in the chain through the programmed variation of 1) the regio- and stereospecificity—and varying degree—of ketone reduction, 2) regio-specificity of chain cyclization and sometimes aromatization, and 3) decoration of the resulting carbon skeleton chain by oxygenation, alkylation and glycosylation (15).

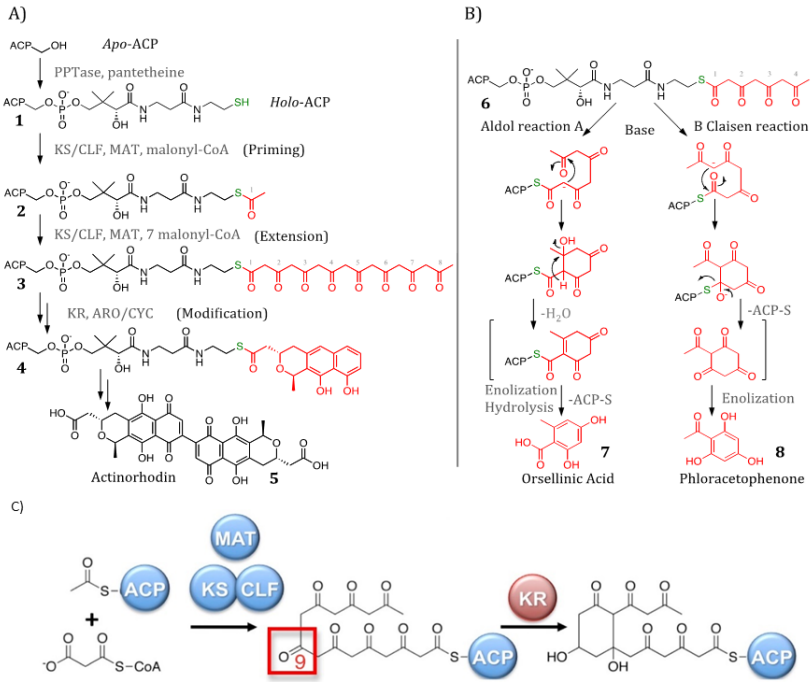


Figure 3.1 Actinorhodin biosynthesis and mechanisms of uncontrolled poly-β-ketide cyclization. **A)** Maturation of the intermediates leading to actinorhodin. **B)** Mechanisms of spontaneous cyclization of poly-β-ketones. **C)** The role of KR in catalyzing the C9-reduction and possible role in the first ring cyclization.

PKSs are generally divided into three types, based on how the constituent enzymatic domains are organized architecturally, with each type of domain organization producing a constitutionally distinct set of polyketides (16). Bacterial aromatic polyketides are produced by the type II PKSs in which each enzymatic activity of the assembly line leading to the natural product is expressed as a separate distinct polypeptide. This is contrast with type I PKSs, where all the domains responsible for a single elongation cycle are expressed as a single large polypeptide (Figure 3.1A). Type II PKSs can be further subdivided according to whether a C9-ketoreductase (KR) is present (Figure 3.1C). Ketoreduction is a key step that unlocks a vast diversity of chemical structure and pharmaceutical activity observed in polyketide natural products. The KR is responsible for regio- and stereospecific reduction of polyketide carbonyl groups. The structure and function of type I KRs have been vigorously studied over the past two decades, while the type II KRs have been investigated by our group in recent years (17-22). If we can control the regio- and stereo-specificity of ketoreduction, the chemical diversity of the resulting polyketide products would be expanded drastically. However, many key questions remain for type II KRs, such as the molecular basis of substrate specificity and stereospecificity. The lack of such information has hampered the efforts to engineer KR for new polyketide biosynthesis.

The actinorhodin (Act) PKS has long served as a model type II PKS for understanding type II polyketide biosynthesis. Its product, actinorhodin, is an antibiotic produced by *Streptomyces coelicolor*. The Act PKS contains 10 stand-alone enzymes that conduct iterative rounds of polyketide chain elongation, followed by chain modifications such as ketoreduction (Figure 3.1C). In the absence of an enzyme, linear unreduced poly-β-ketones longer than a triketide (**6**) undergo spontaneous intramolecular and intermolecular aldol/Claisen condensations, generating shunt products such as **7** and **8** (Figure 3.1B). However, in the presence of the actinorhodin minimal PKS

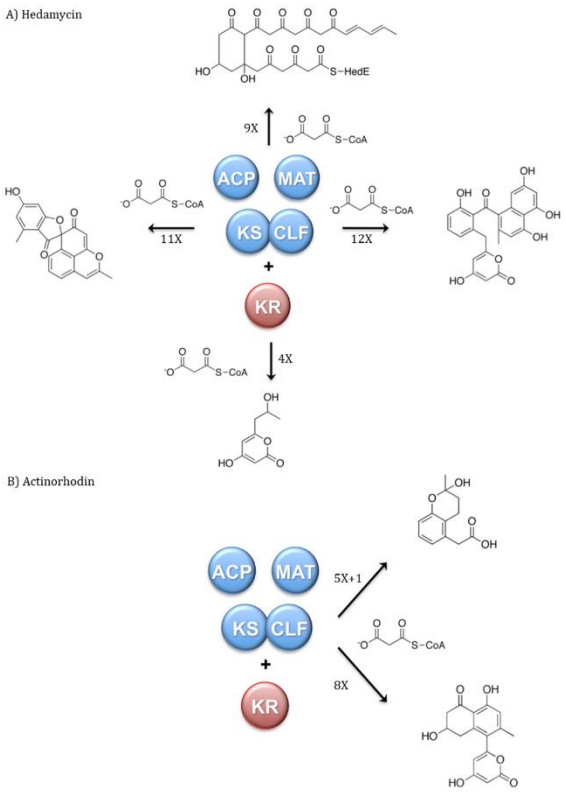


Figure 3.2 Differing chain length specificities of the HedKR and ActKR. **A)** HedKR recognizes substrates of widely varying length from four ketides to twelve ketides. **B)** ActKR recognizes pentaketide and octaketide substrates.

(Act minPKS; containing the ketosynthase (KS), chain length factor (CLF), acyl carrier protein (ACP) and malonyl-CoA:ACP transacylase (MAT)), the polyketide is elongated to a linear octaketide (Figure 3-1A), which is then reduced by the actinorhodin ketoreductase (ActKR) specifically at the C9 carbonyl group to a hydroxyl group (Fig. 3.1C). Biochemical data also suggest that the KR may serve a role in the cyclization of the first ring (Fig. 3.1C) (8-11).

Another aspect of ActKR that is not well understood is its chain-length specificity for the sixteen-carbon octaketide. During the biosynthesis of type I PKS, the KR reduces the β-carbonyl group after each round of polyketide chain elongation. In contrast, the type II KR reduces a polyketide intermediate only after it is fully elongated by the minimal PKS (Figure 3.1C). Past studies of ActKR show that it is highly specific for reduction of an octaketide, along with some

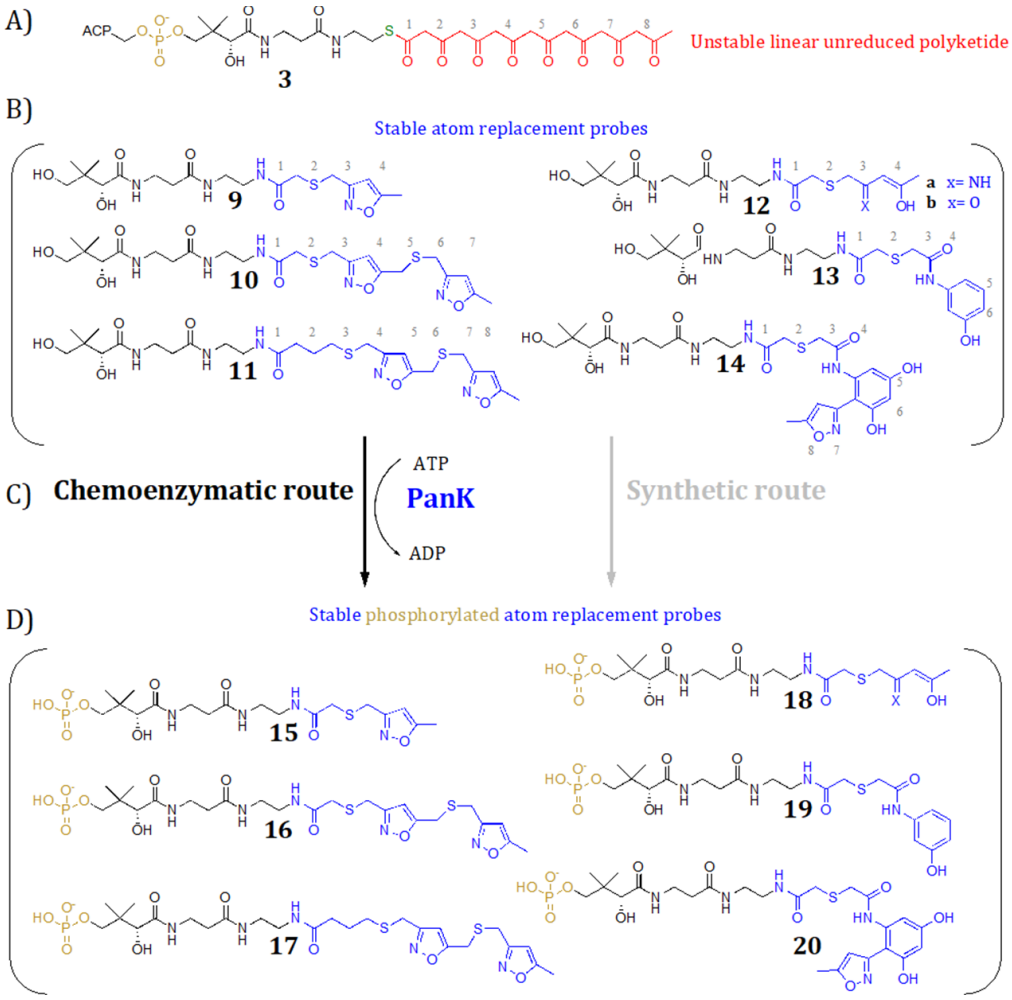


Figure 3.3 Suite of mimetics designed to mimic various intermediates in the aromatic polyketide pathway. **A)** Highly unstable natural intermediate linear poly-β-ketone. **B)** Suite of mimetics. Mimetics **9-12** mimic the unstable linear poly-β-ketone. **C)** In order to more closely resemble the natural substrates, the mimetics can be enzymatically phosphorylated using PanK. **D)** Enzymatically phosphorylated mimetics.

activity for a pentaketide, but at a much lower rate (Figure 3.2B). We have also previously solved another type II KR structure, the hedamycin KR (HedKR), which holds a high sequence homology (61 % sequence identity) with ActKR. However, HedKR is much more broad in its substrate specificity, and can reduce tetra-, octa-, nona-, undeca- and dodecaketides (Figure 3.2A). The α6-α7 loop of type II KR is likely involved in determining the substrate specificity because this is the least conserved region of KR. It remains a mystery how the KR recognizes an incoming polyketide substrate with a given chain length, and if KR is indeed responsible for first ring cyclization.

In order to critically evaluate the determinants of the regio-, stereo- and substrate specificity of KR, it is necessary to monitor the reduction using a polyketide substrate or mimetic. Application of the natural substrate to this task is challenging, because unreduced polyketide intermediates are too reactive to be isolated, let alone be used in an *in vitro* enzyme assay. Because polyketides are prone to intramolecular aldol cyclization, an unreduced polyketide longer than a triketide (six carbons) has never been synthesized, nor isolated from a biosynthetic reaction. Once these intermediates are biosynthesized by the minimal PKS, they will spontaneously cyclize (such as reactions depicted in Figure 3.1B), rendering any additional synthesis or testing impossible. To solve this problem, we have developed an atom replacement approach that replaces pre-defined positions of carbonyl groups with heteroatoms such as S, N, C or O (Figure 3.3). The atom-replaced substrate mimetics bear a strong resemblance to the actual polyketide substrate, yet they are stable enough to be synthesized and used in structural and functional studies. In order to assess the effect of chain length on KR-substrate binding and the potential role of ActKR in first-ring cyclization, poly-β-ketone mimetics of tetraketide (**9)** and octaketide (**11**) lengths were used in this study. Since the natural poly-β-ketone substrate is phosphorylated, the pantetheine kinase from vitamin B biosynthesis (PanK) was used to enzymatically phosphorylate **11**, yielding **17**.

Here, we report the versatility of the atom-replaced probes to critically analyze the structure and function of ActKR. Analysis of ActKR:tetraketide mimetic cocrystal structures led to the design of a double mutant (DM) that retains activity and bound our tetraketide mimetic, while the wildtype ActKR was not observed to do so. The double mutant also exhibits reversed stereospecificity of reduction. Together, the DM:tetraketide (hereafter referred to as DM:tetra) and DM:PPant-octaketide (hereafter referred to as DM:octa) mimetic cocrystal structures reveal the molecular basis for KR-substrate interactions. Furthermore, the DM:PPant-octaketide mimetic

cocrystal structure highlights the importance of interactions between the substrate phosphate group and the KR arginine patch. Thus the atom replacement approach is proven to identify enzyme features critical for chain length specificity that were unattainable before. Additionally, a limitation of the first generation of mimetics is noted, namely, the rigidity of the isoxazole rings, and has sparked the development of a new generation mimetic that retains the flexibility of the natural poly-β-ketone substrate. This new class of poly-β-ketone mimetics is poised to unlock many exciting discoveries in the biosynthesis of aromatic polyketides.

3.3 Results and Discussion

3.3.1 Results from the thesis of Joel Bruegger

A former graduate student, Dr. Joel Bruegger, first attempted the co-crystallization of wild type ActKR with a series of tetraketide mimetics (including **9** and **15)**. In these cocrystal structures, he misinterpreted the electron density of the N-terminal 6-His tag of a symmetrically-related monomer as that of the tetraketide and concluded that ActKR binds the tetraketide mimics (including **9** and **15)** on the exterior of ActKR. Thankfully our group has the tradition of having another graduate student validate the crystal structure prior to paper submission. I was assigned this task and found that the density on the surface of ActKR actually corresponded to that of the 6-His tag (Figure 3.4). A potential catastrophe was prevented.

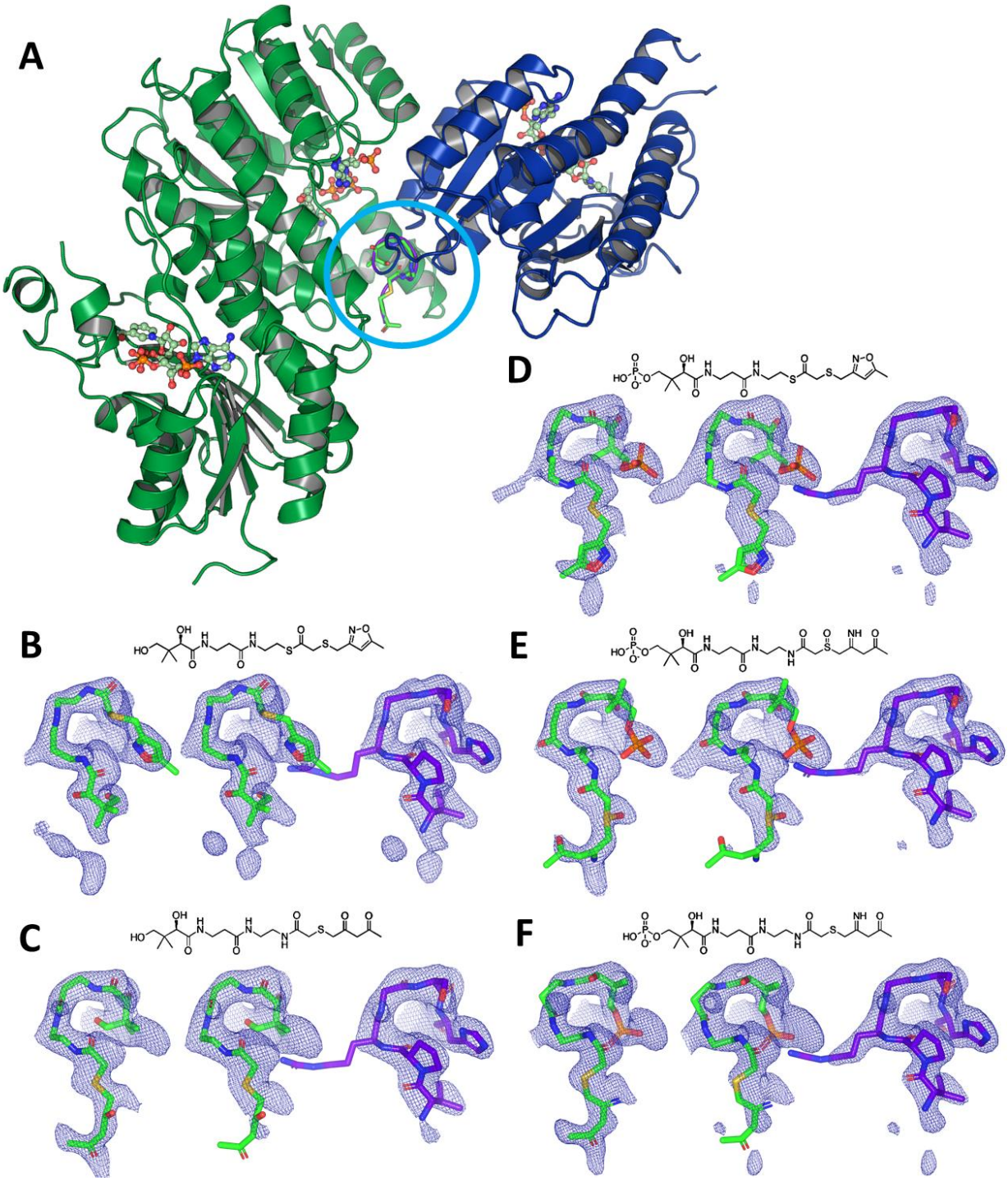


Figure 3.4 Density fit comparison with tetraketide mimetics versus N-terminal histidine tag. **A)** The N-terminal histidine tag (purple) of a neighboring symmetrically-related monomer (dark blue) interacts with the proposed pantetheine arm binding site of the KR in several crystal structures derived from crystals soaked with a series of tetraketide mimetics. **B-F)** Density fit comparisons. **Left side:** Original density and tetraketide model. **Center:** Overlay of the iterative build omit map and tetraketide model. **Right side:** Iterative build omit map fitted with model of the N-terminal histidine tag. Model bias can be seen in the lower left area of panels C-F.

3.3.2 The ActKR double mutant

Although Joel interpreted the 6-His tag electron density incorrectly, the observation that the tetraketide mimetics did not bind into the KR substrate pocket led to the hypothesis that the absence of binding may be a reflection of the ActKR's narrow chain-length specificity. This proposal inspired the mutation of two histidines that lie on either side of the entrance to the substrate pocket and that may play a role in KR-substrate interactions. We first identified the two His through a sequence alignment of different type II PKS KRs and found that these two His are the least conserved features between HedKR and ActKR. They were mutated to their HedKR counterparts, Tyr and Gly, to test whether these mutations would loosen ActKR's chain-length specificity to that of HedKR (Figure 3.2). Because HedKR accepts the tetraketide as its substrate, whereas ActKR does not, we hypothesized that the two His may be important for chain length specificity. Namely, in the case of ActKR, which is highly specific for the longer octaketide substrate, the two His can form hydrogen bonds with the carbonyl groups of a pre-mature tetraketide, leading to the "trapping" of the pre-mature tetraketide at the exterior of ActKR. Indeed, when co-crystallized under the same conditions as with the WT, the tetraketide mimetic is found to be bound inside the substrate pocket.

3.3.4 The DM:tetra cocrystal structure

Attempts were made to co-crystallize several different atom-replaced tetraketide mimetics with the WT ActKR (Figure 3.4). For all of the WT ActKR:tetraketide mimetic cocrystals, density corresponding the tetraketide mimetic was not observed in the substrate binding pocket. Dr. Joel Bruegger proposed that this may be a reflection of ActKR's previously observed narrow substrate

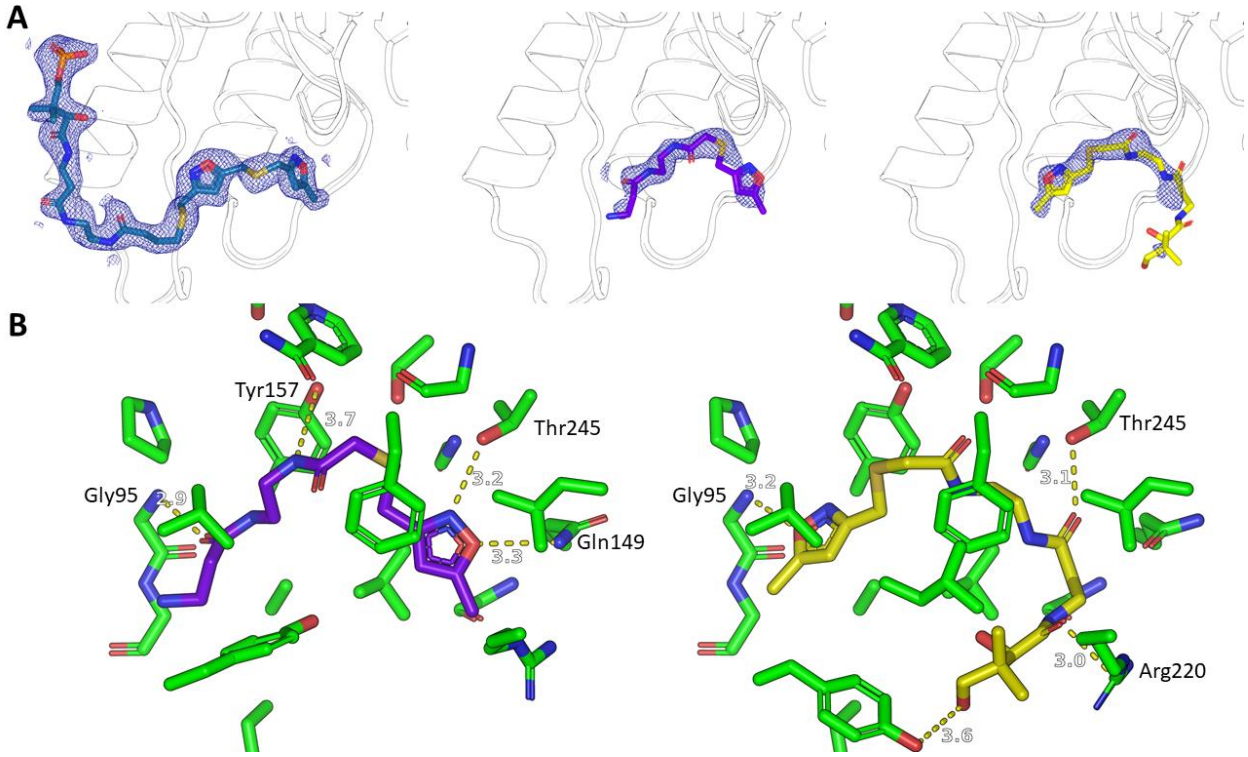


Figure 3.5 The tetraketide mimetic may be bound in one of two different orientations. **A)** Comparison of the 2mFo-DFc electron density maps for the octaketide mimetic and the two possible orientations of the tetraketide mimetic. The maps are contoured at 1 σ. **B)** Polar contacts possible with the two different candidate orientations of the tetraketide mimetic.

chain length specificity (12) and so designed a mutant based on the HedKR, which has a broader chain length specificity. In confirmation of this hypothesis, the double mutant cocrystallized with one of the tetraketide mimetics bound inside the substrate binding pocket. The DM:tetra condition crystallized in the space group $P4_12_12$ with two monomers in the asymmetric unit. With our current data, it is difficult to judge whether the tetraketide mimetic is bound in the forward direction (entering the substrate pocket from the front, as does the octaketide mimetic), or in the reverse direction (Figure 3.5A). This is not surprising, because the ActKR active site is not built to bind a tetraketide, thus cannot guarantee a productive binding motif of tetraketide in its substrate binding pocket. The $2mF_o$-F_c maps obtained after refining with both orientations are very similar. In addition, the number of polar contacts is the same between the two different orientations. The reverse orientation would be consistent with Dr. Joel Bruegger's stereochemical assays that show

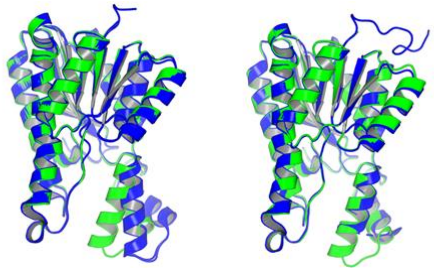


Figure 3.6 Structural alignments between the DM (green) and WT ActKR (blue). **Left:** Alignment between monomer H of the DM:octa structure and monomer A of the 1W4Z WT structure, which gives the highest RMSD of all DM monomer alignments with WT monomers. **Right:** Alignment between monomer B of the DM:octa structure and monomer B of the 1W4Z WT structure, which gives the lowest RMSD of all DM monomer alignments with WT monomers.

that the stereospecificity of the double mutant is completely switched from that of the wildtype—from (S)-dominant to (R)-dominant.

The only structural difference between the WT and DM ActKR lies in the conformation in the flexible lid region formed by helices α6 and α7 (Figure 3.6). The lid region of the DM is in the

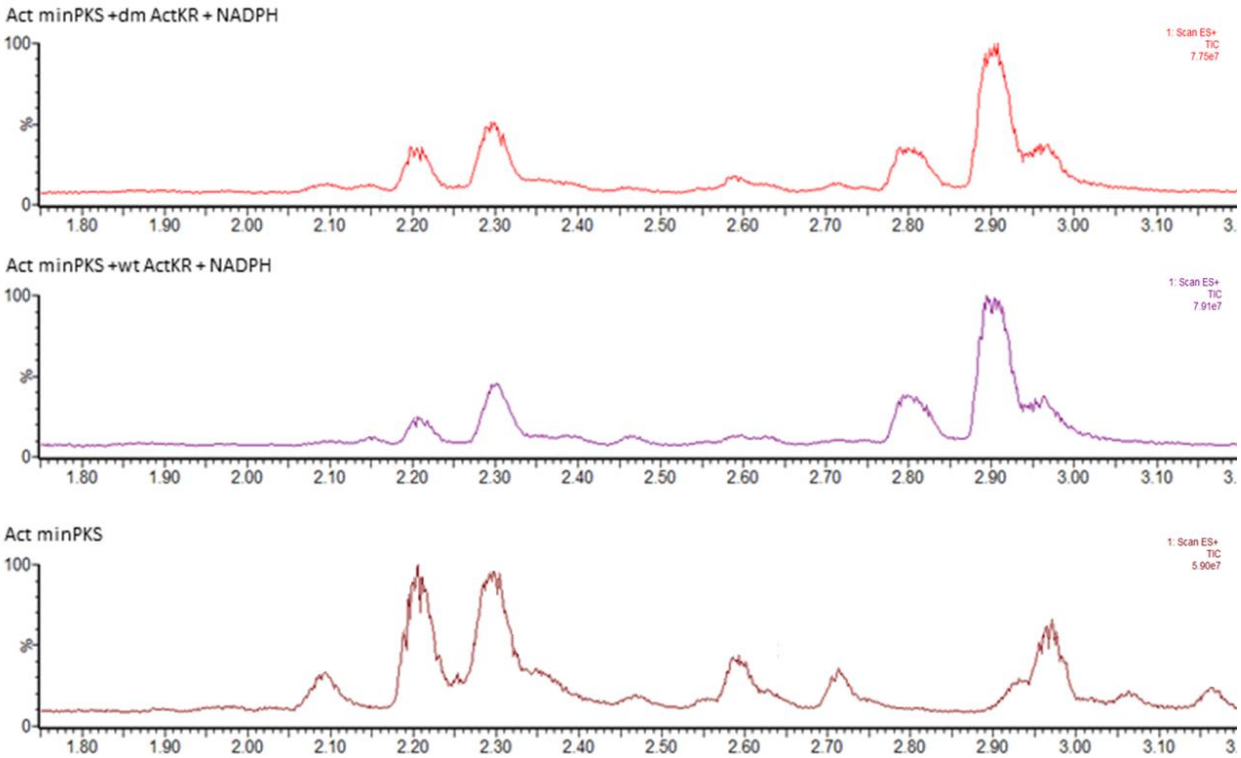


Figure 3.7 The ActKR H153Y H201G double mutation does not change ActKR activity within the Act minPKS. Shown are the LC-MS chromatograms from organic products extracted from: **(Top)** Act minPKS reconstituted with the double mutant. **(Middle)** Act minPKS reconstituted with WT ActKR and **(Bottom)** Act minPKS reconstituted with no KR. The product profiles with ActKR double mutant and wild-type are virtually identical.

closed position for all the 10 monomers of the two cocrystal structures. It is not known whether this conformation is a result of the double mutation or a result of having bound the mimetics. Monomer B from the DM:tetra structure aligns closely with monomer B from the WT 1W4Z structure, showing that the conformation seen in the DM is within the natural range of the KR (Figure 3.6). Consistent with this observation, the product profile detected by LC-MS is virtually identical between the Act minPKS reconstituted with the WT KR and the Act minPKS reconstituted with the DM KR (Figure 3.7). The above supports that chain length control is a combination of protein-protein interaction (preserved in DM KR) plus active site conformation.

3.3.3 The DM:PPant-octa cocrystal structure

The above cocrystal structure of DM:tetraketide is informative. However, the full-length substrate of ActKR should be an octaketide. In order to probe the substrate-protein interactions of ActKR, we co-crystallized the double mutant ActKR with the octaketide mimic. Further, to probe the importance of PPant group on substrate-protein interactions, this octaketide mimic includes the PPant moiety. The DM:PPant-octa crystallized in the space group P1 with two tetramers in the asymmetric unit (Figure 3.8A). The phosphate group on the pantetheinyl group of the mimetic binds to the arginine patch just above the substrate pocket entrance, highlighting the importance of this interaction for substrate binding. Clear density was seen for the PPant-octaketide mimetic only in monomers A and E, where this stabilizing interaction occurs (Figure 3.8B). In addition to the arginine patch, the mimetic is seen to interact with other important residues identified in previous work (18, 20, 22). In monomers A and E, the mimetic passes the 94-PGG-96 motif as it

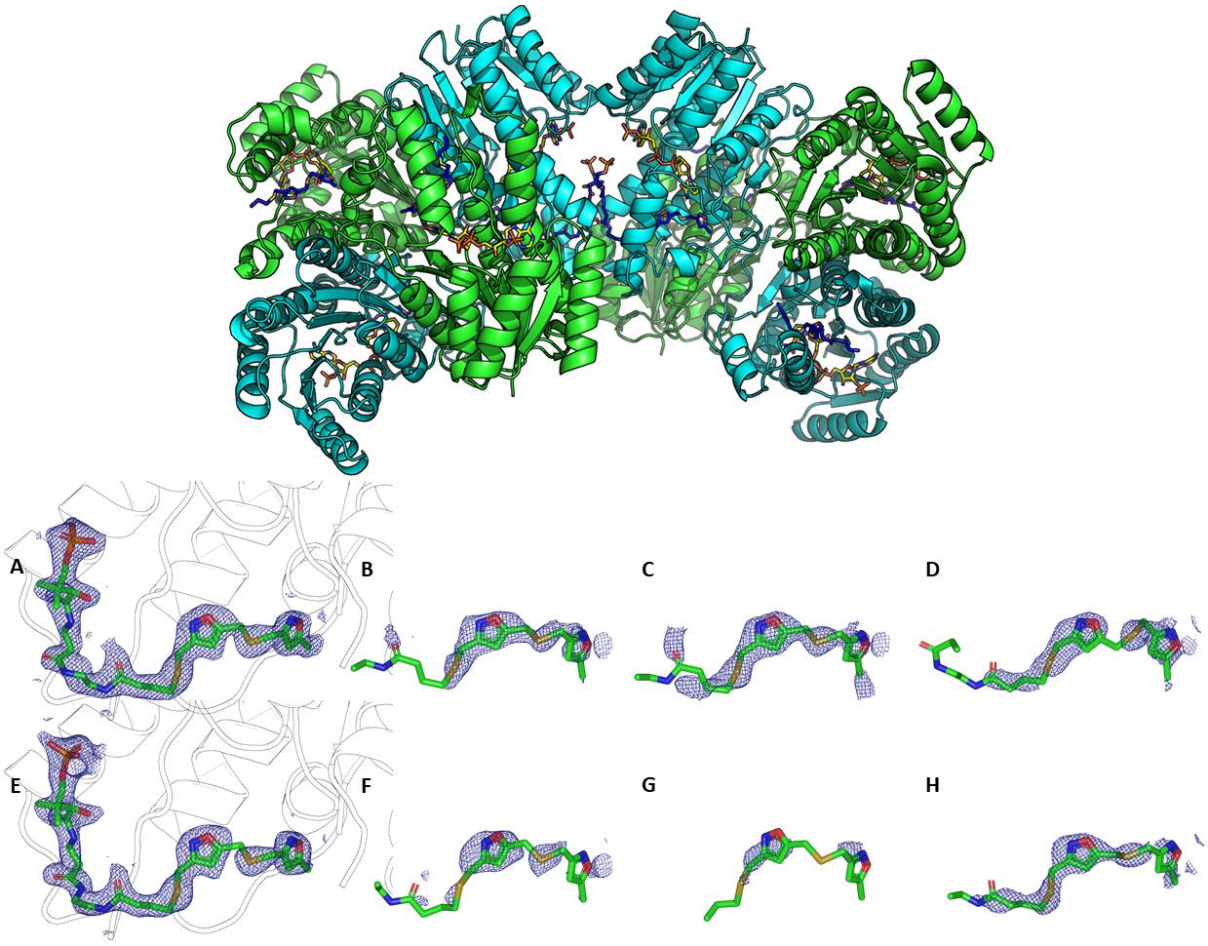


Figure 3.8 The DM:pPant-octa cocrystal structure. **A)** The asymmetric unit contains two DM tetramers (dimer of dimers). The cofactor NADPH is colored in yellow and the pPant-octaketide mimetic is colored in blue). Monomers A and E (in the middle and colored in cyan) pack against each in the region of the arginine patch proposed be the pPant-ACP binding site. The phosphate groups on the pPant-octaketide mimetics bound in monomers A and E interact with the proposed ACP-binding site as expected. Clear density for the pPant-octaketide mimetic is only seen in monomers A and E. Only density for the octaketide moiety is found in the substrate binding-site of the other monomers. **B)** $2mF_o$-DF_c electron density maps for the octaketide mimetic in each monomer substrate binding pocket contoured at 1σ. The models for the octaketide mimetic for monomers B-D and F-H were truncated where the $2mF_o$-F_c electron density fell below 0.7 σ.

enters the substrate pocket. The 94-PGG-96 motif has been shown to be important in setting the stereospecificity of reduction (20). The first polar interaction that occurs within the pocket is a hydrogen bond between the backbone amine group of Gly95 and the nitrogen on the first isoxazole ring (Figure 3.9). The second ring is held in place by the hydroxyl groups on Ser144 and Thr145 that hydrogen bond with the nitrogen and oxygen, respectively, on the second isoxazole ring. Another hydrogen bond donor within bonding distance of the nitrogen on the second ring is the

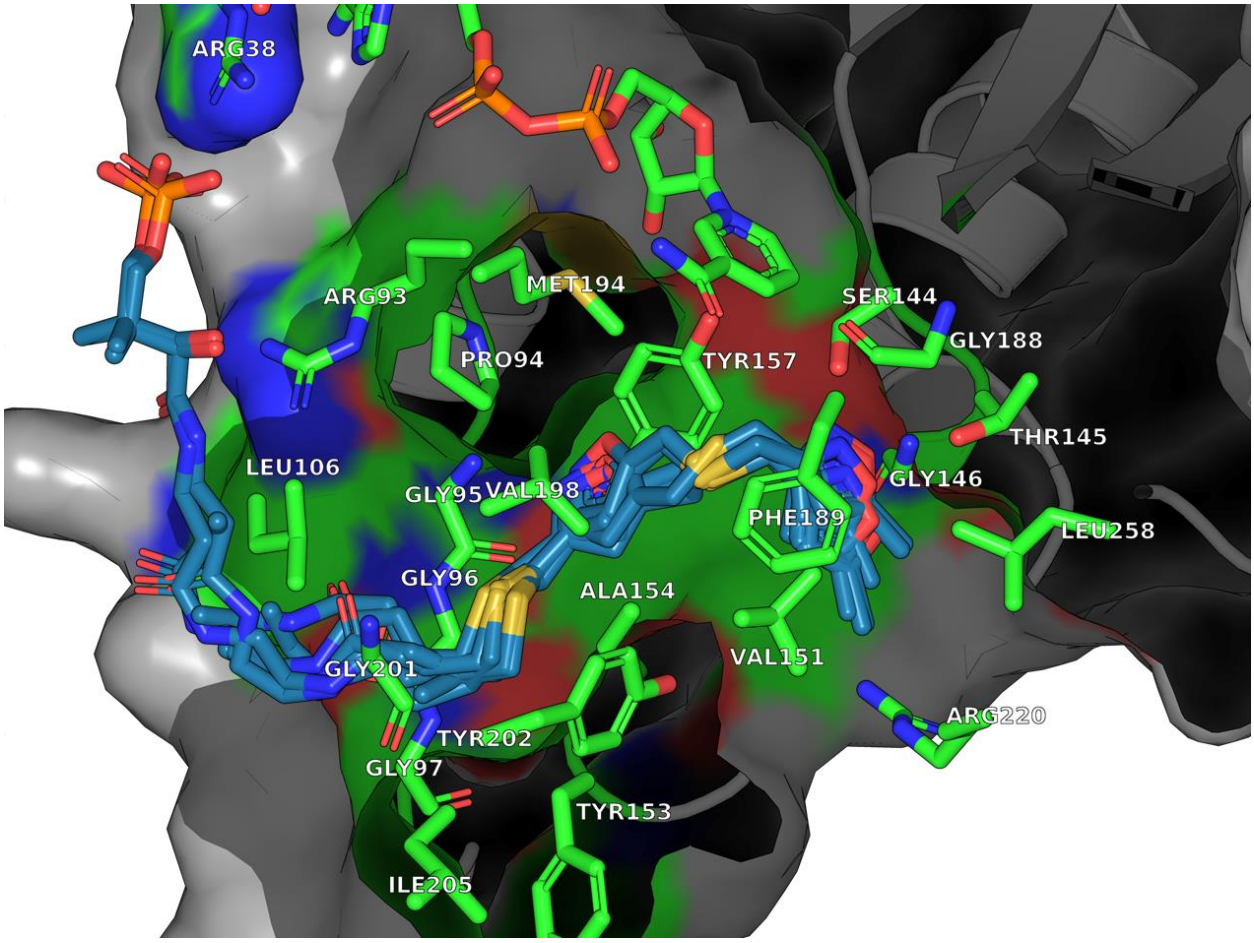


Figure 3.9 Overlay of the DM:pPant-octaketide monomers. Drawn are the residues within 4.5 A of the mimetic. Arg38 forms a salt bridge with the pantetheinyl phosphate group. The backbone amine group of Gly95 hydrogen bonds with the nitrogen on the first isoxazole ring. The hydroxyl groups on Ser144 and Thr145 hydrogen bond with the nitrogen and oxygen, respectively, on the second isoxazole ring. Phe189 appears to be involved in pi-pi interactions with the second isoxazole ring. Inner wall residues Pro94, Gly95, Gly96, Gly97, Gly 146, Val151, Ala154 and Tyr157, together with lid residues Gly 188, Phe189, Val198, Tyr202, Leu258 and Arg220, define the binding pocket and potential cyclization chamber. The models for the octaketide mimetic for monomers B-D and F-H were truncated where the $2mF_o$-F_c electron density fell below 0.7 σ.

backbone amine group of Gly146. Phe189 appears to be involved in pi-pi interactions with the second isoxazole ring. The majority of binding interactions seen in the pocket are hydrophobic in nature. Inner wall residues Pro94, Gly95, Gly96, Gly97, Gly146, Val151, Ala154 and Tyr157, together with lid residues Gly188, Phe189, Val198, Tyr202, Leu258 and Arg220, define the binding pocket and potential cyclization chamber (Figure 3.10). The vast majority of these residues are completely conserved across type II C9 ketoreductases and the remaining residues are semi-

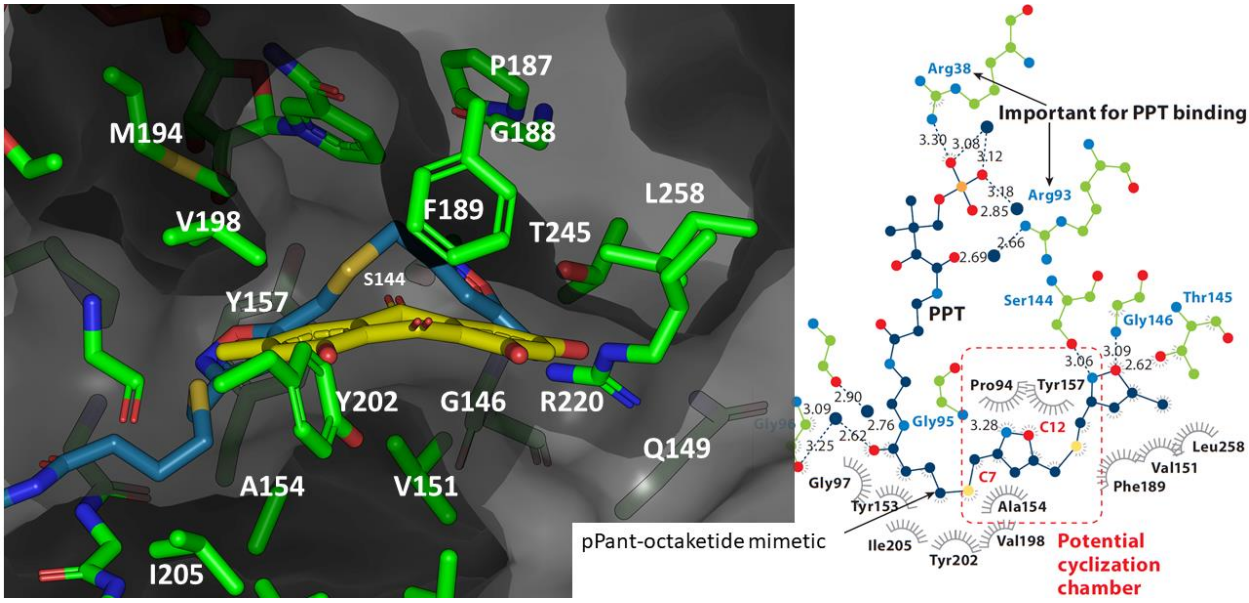


Figure 3.10 Potential cyclization chamber. Inner wall residues Pro94, Gly95, Gly96, Gly97, Gly146, Gln149, Val151, Ala154 and Tyr157, together with lid residues Pro187, Gly188, Phe189, Met194, Val198, Tyr202, Ile205, Leu258 and Arg220, define the binding pocket and potential cyclization chamber. All of these residues are highly conserved, if not completely conserved, across more than 20 type II C9 ketoreductases. The octaketide mimetic is colored in blue. In yellow is the emodin molecule from the ActKR:emodin cocrystal structure (PDB code 2RH4). The emodin and octaketide bend at the same place in the pocket, which may be caused by interactions with V198, L258 and V151, suggesting a role in orienting the substrate for reduction and possibly C7-C12 first ring cyclization.

conserved. Interestingly, when aligned with the ActKR:emodin cocrystal structure (PDB code 2RH4), the emodin and octaketide are seen to bend at the same place in the pocket, which may be caused by interactions with V198, L258 and V151, suggesting a role for these residues in orienting the substrate for reduction and possibly C7-C12 first ring cyclization. All three are conserved in 21 out of 22 type II C9 KRs referenced and semi-conserved in the other 1 out of 22.

3.3.5 Inspiration for the next generation of poly-β-ketone mimetics

The intermediates in early type I and II non-reducing PKS biosynthesis are long linear poly-β-ketones that are highly prone to spontaneous inter- and intramolecular aldol/Claisen condensations. In order to study these valuable classes of enzymes, our lab designed a series of atom-replaced mimetics that maximally retain the steric and electronic structure of the poly-β-ketone while being stable enough to use in crystallographic studies. These mimetics have proven

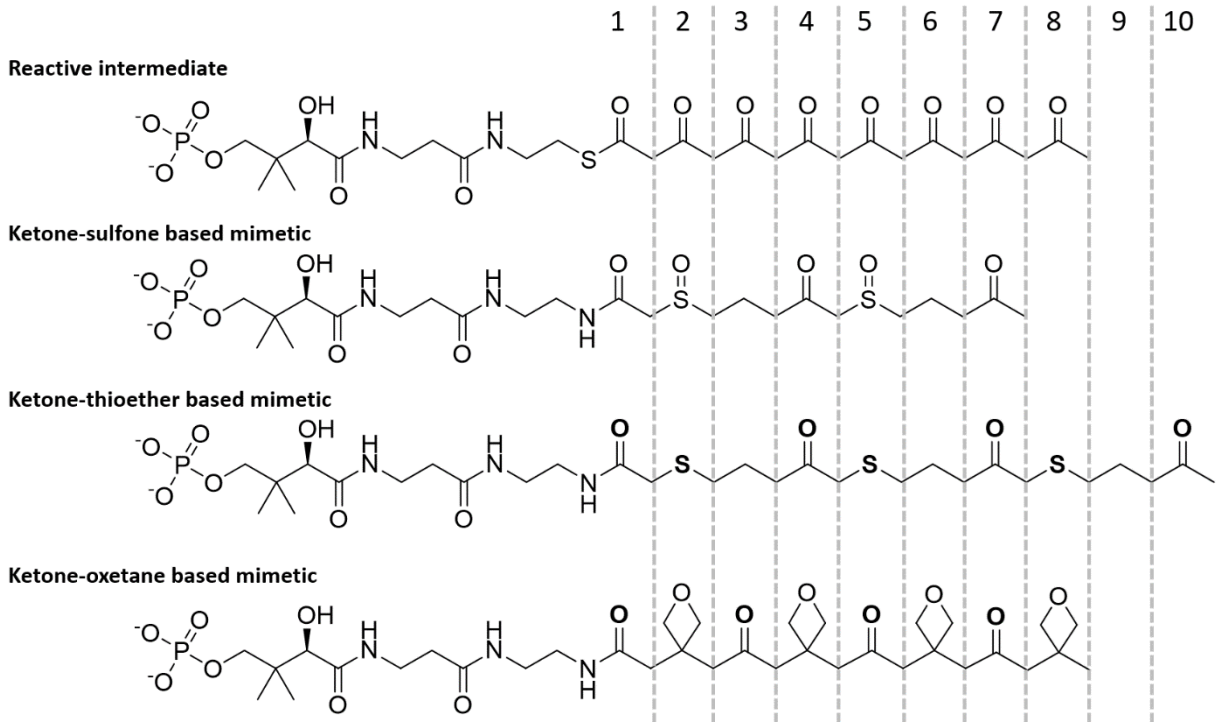


Figure 3.11 Designs of the next generation of poly-β-ketone mimetics.

to be fruitful tools in the study of these types of systems. However, the DM:PPant-octaketide mimetic cocrystal structure reveals a limitation of our first generation of mimetics—the rigidity of the isoxazole rings. In this structure, it appears that the mimetic wants to bend around Val151 but is unable to do so. In collaboration with the Vanderwal Lab at UCI, we are developing a new generation of poly-β-ketone mimetics that are highly flexible and promise to lead to a deeper understanding of type I and II non-reducing iterative PKSs (Figure 3.11).

3.4 Conclusions

In the Type II reducing PKSs (hitherto referred to simply as PKSs), a stereocenter is introduced upon the first modification step, the reduction of a carbonyl group, catalyzed by the NAD(P)H dependent Ketoreductase C9 (KR). The regio- and stereospecificities of type II PKS KRs is are less well understood. The actinorhodin KR (ActKR) specifically reduces the C9 carbon

of an octaketide and current evidence suggests that KR is also involved in first ring cyclization. However, the instability of the polyketide substrate, which is tethered to the shuttling protein domain--the acyl carrier protein (ACP)—via a 4'-phosphopantetheine arm prosthetic group, has hampered our understanding about enzyme mechanisms and specificity of KRs. To probe the activity and specificity of type II PKSs, we developed a novel strategy that selectively replaces the reactive carbonyl groups of a polyketide substrate to more stable chemical moieties such as C, N, S or O. This atom replacement approach has resulted in great success, and in this chapter, we applied the probes to the KR domain, leading to two cocrystal structures of ActKR bound with atom-replaced polyketide mimetics, which help identify residues essential for the substrate specificity of KR. The DM:PPant-octaketide cocrystal structure strongly supports the importance of the pantetheinyl phosphate for correctly orienting the poly-β-ketone intermediate, which in turn leads to a precise control of chain length and stereo-specificity in type II KR domains.

In addition, the rigidity of the isoxazole rings used in our first-generation atom-replaced probes was observed to limit its ability to fully enter the substrate pocket and inspired the development of a next generation of poly-β-ketone mimetics that also mimic the flexibility of the poly-β-ketone chain.

3.5 Materials and Methods

3.5.1 Expression and Purification of ActKR H153Y H201G

The actinorhodin ketoreductase H153Y H201G double mutant (DM) was previously generated from the ActKR/pET-28b(+) expression plasmid by Dr. Joel Bruegger. The ActKR(DM)/pET28b plasmid was transformed into *E. coli* BL21(DE3) competent cells and plated on LB-agar plates containing kanamycin (100 μg/mL). The plates were incubated overnight at 37

°C. Positive transformants were transferred to a 25 mL starter culture of Luria-Bertani (LB) broth containing kanamycin (75 µg/mL) and grown overnight at 37 °C with shaking, which was then used to inoculate two liters of TB with kanamycin (50 µg/mL). Cultures were grown at 37 °C until the A_{600} reached 1.0. The cells were then cooled to 18 °C, and 1 mM IPTG was added to induce protein expression. After 12-18 hours of incubation at 18 °C, the cells were harvested by centrifugation at 4000 x g for 10 minutes. The cell pellets were flash-frozen in liquid nitrogen and stored at -80 °C. The frozen cell pellets were thawed on ice and resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 % glycerol, 10 mM imidazole). The cell suspension was lysed using sonication (8 x 30 s cycles), and cellular debris was removed by centrifugation at 14000 rpm for 45 minutes. The lysate was incubated with 5 mL Ni-IMAC resin (BioRad) at 4 °C for one hour. The resin was poured into a fritted column and the flow-through fraction was collected. The resin was washed with 100 mL of lysis buffer then eluted with lysis buffer plus increasing amounts of imidazole (20-500 mM). The elutions were analyzed using SDS-PAGE, and the elutions containing the protein of interest were combined and dialyzed overnight into storage buffer (50 mM Tris pH 8.0, 100 mM NaCl, 10 % glycerol). The N-terminal histidine tag was removed from the wildtype and double mutant ActKR by adding 1 unit of bovine thrombin (Sigma) per mg of ActKR(DM) to the combined protein elutions prior to dialysis to cleave the N-terminal 6-His tag. Following the dialysis, ActKR(DM) was incubated with Ni-IMAC resin and passed over a fritted column to remove any uncleaved ActKR(DM) and the flow through was analyzed by MALDI-TOF MS to confirm cleavage of the His-tag (data not shown). The ActKR(DM) protein was then concentrated to 10 mg/ml, buffer exchanged into crystallization buffer (25 mM Tris, pH 7.5), and flash frozen in liquid nitrogen.

3.5.2 Expression and Purification of Act KS/CLF protein

Act KS/CLF was expressed in *S. coelicolor* CH999 containing the pRJC006 expression vector as previously described but with slight modifications (23). Fresh spores were used to inoculate 50 mL of Super YEME containing 50 μg/mL kanamycin. The 50 mL cultures were grown for 3 days at 30 °C with shaking at 250 rpm and then transferred to 2.5 L baffled flasks containing 500 ml of Super YEME and 50 μg/mL kanamycin and grown for an additional 2 days under the same conditions. Protein expression was induced with 10 μg/mL thiostrepton.

The cells were harvested 24 hours after induction by centrifugation (20 minutes at 5100 rpm) and the cell pellets were stored at -80 °C. Cell pellets from 2 L of cell culture were resuspended in 200 mL of lysis buffer (100 mM KP_i, 15 % glycerol, 300 mM NaCl, pH 7.5, with 2 tablets of EDTA-free protease inhibitor cocktail (Roche)) and lysed by sonication on ice (10 x 1 minute cycles). Cell debris was pelleted by centrifugation (1 hour at 14000 rpm) and DNA was precipitated using streptomycin sulfate (2 % final concentration) followed by centrifugation (45 min at 14000 rpm). The supernatant was filtered using a 0.45 μM filter and 30 % – 50 % $(NH_4)_2SO_4$ was used to precipitate Act KS/CLF overnight.

The precipitated protein was pelleted by centrifugation (30 min at 14000 rpm) and redissolved in Ni-binding buffer (50 mM KP_i, 10 % glycerol, 500 mM NaCl, 5 mM imidazole, pH 7.5). After resuspension, the solution was filtered using a 0.45 μM filter then bound to 2 ml of Ni-IMAC resin (Bio-Rad) pre-equilibrated with Ni-binding buffer and stirred at 4 °C for 1 hour. Act KS/CLF was eluted using Ni-binding buffer containing increasing amounts of imidazole. Fractions (2.5 mL each) containing between 375 mM and 750 mM imidazole were individually buffer exchanged into storage buffer (100 mM KP_i, 10 % glycerol, pH 7.5) using a PD-10 column (GE Healthcare) and flash frozen at -80 °C.

3.5.3 *In vitro* reconstitution assay

The *in vitro* reconstitution assays were incubated overnight at room temperature with 100 µL of 30 µM wild type or mutant ActKR, 7.5 µM actinorhodin ketosynthase/chain-length-factor (KS/CLF), 45 µM actinorhodin acyl carrier protein (ACP), 5 µM malonyl-CoA:ACP transacylase (MCAT), 20 uM malonyl-CoA synthetase (MatB), 100 mM malonate, 5 mM Coenzyme A, 20 mM ATP, 5 mM $MgCl_2$, and 2.6 mM NADPH in 100 mM Tris-HCl (pH 7.5). Compounds were extracted with 300 µL 94 % ethyl acetate, 5 % methanol and 1 % acetic acid. Solvent was evaporated with centrifugal SpeedVac and products were resuspended in 50 µL DMSO. Samples were run on reverse-phase HPLC with a C18 analytical column (Phenomenex, 4µm, 150 mm x 4.6 mm) and separated with a 5 % to 95 % Acetonitrile in 0.1 % formic acid gradient over 20 minutes with absorbance monitored by a UV-vis diode array detector tuned to 280 nm.

3.5.4 Enzymatic phosphorylation and purification of the octaketide mimetic 11

The synthesis, purification and characterization (by NMR and HRMS) of the unphosphorylated octaketide mimetic **11** have been previously reported (13). The phosphorylation of **11** to **17** was conducted chemoenzymatically using the first pantetheine kinase (PanK) from the pantothenate (vitamin B5) pathway (24-27). PanK is promiscuous at catalyzing the phosphorylation of pantetheine and pantetheine analogs derived at the 4'-hydroxyl group. A 300 µL aliquot of 100 mM **11** (17.5 mg, 29.9 µmol) was diluted into 12 mL of 25 mM potassium phosphate pH 7.5, 10 mM $MgCl_2$, 8 mM ATP, 2 µM PanK to a final concentration of 2.5 mM. After incubating at 37 °C for 3 h, mixture as purified by centrifugal filtration and dried. The phosphorylated pantetheine derivatives were successfully purified by reverse phase liquid chromatography (Beckman Coulter HPLC System Gold ®) with a C-18 ODS column (5µ x 10mm

x 15cm; Beckman Coulter Ultrasphere TM) with UV detection set at 250 nm. Derivatives were eluted with a flow rate of 4.75 ml/min using a linear gradient of 5-50 % MeCN/ 0.1% formic acid in a H2O/0.1 % formic acid mixture over 15 min followed by 50-95 % MeCN/0.1 % in a H2O/0.1 % formic acid mixture over 5 minutes. Under these conditions, the HPLC retention time of the phosphorylated derivative is around 2.5 minutes less than the non-phosphorylated mimetic. 13.6 mg (68.4 %) of the phosphorylated mimetic **17** was recovered; HRMS (*m/z*): $[M+H]^-$ calculated for **17** 664.71; found 664.21 (Figure 3.12).

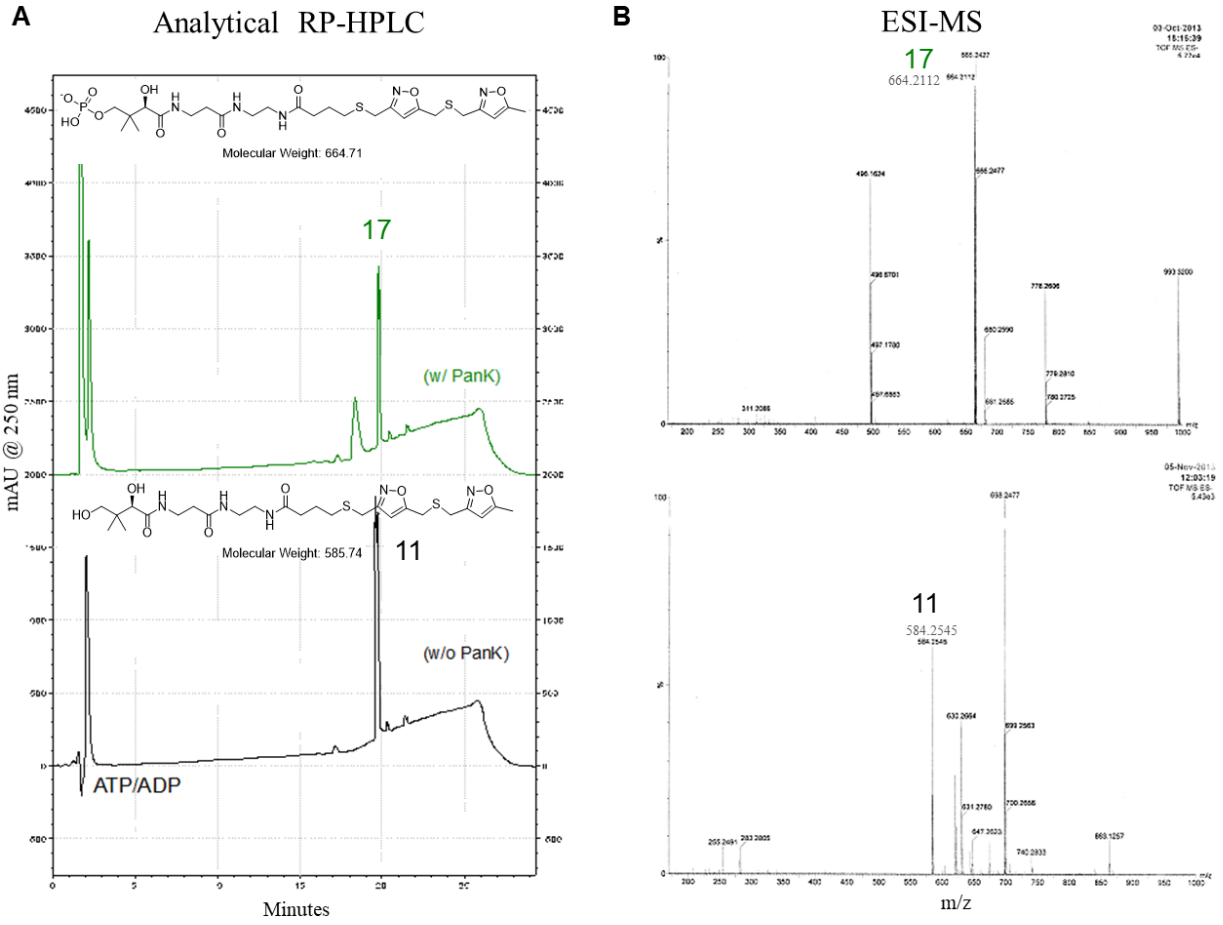


Figure 3.12 Purification and verification of the enzymatically phosphorylated octaketide mimetic.

3.5.5 Crystallization

To obtain the PPant-octaketide mimetic cocrystal structure, frozen aliquots of protein were thawed on ice then filtered with a 0.22 µM filter prior to crystallization. Both the wildtype and double mutant ActKR (6 mg/ml), both with and without the N-terminal histidine tag, were incubated with 5 mM **17** and 3 mM NADPH and screened extensively against several Qiagen NeXtal crystallization suites using sitting-drop vapor-diffusion. One condition was found to give crystals of mixed morphology with the wildtype ActKR and regular diamond plates with the double mutant ActKR. This condition was optimized to give: 0.24 M $MgCl_2$, 23 % PEG 3350, and 0.1 M Tris, pH 8.5. The double mutant was cocrystallized with the tetraketide mimetic by Dr. Joel Bruegger in the original WT ActKR crystallization condition (4-5 NaFormate) using 8 mg/ml protein, 2 mM NADP+ and 2 mM **15**.

3.5.6 Data Collection and Structure Refinement

The crystals were soaked with cryo-protectant composed of well solution and 30 % glycerol, flash frozen in LiN_2, and diffraction data was collected at beamlines 8.2.1 and 8.2.2 of ALS or beamline 12.2 of SSRL. HKL2000 (28) was used to index, integrate and scale the cocrystal diffraction data. The crystal structures were solved using molecular replacement with Phaser in Phenix suites (29), using wild type ActKR as the search template. Mimics **15** and **17** were initially fit into the electron density map with Phenix LigandFit and adjusted in Coot (30). Coot was also used to build in alpha helices 6 and 7 of the cocrystal structures. Multiple rounds of refinement were performed using Phenix Refine and Coot.

Table 3.1 Summary of X-ray crystallographic data collection and data refinement statistics.

	ActKR H153G H201Y:**15**	ActKR H153G H201Y:**17**
Data collection		
Wavelength (Å)	1.033	1.000
Total reflections	422739 (40324)	308038 (30174)
Unique reflections	59204 (5836)	116012 (11263)
Space group	P 4_12_12	P 1
Cell dimensions		
a, *b*, *c* (Å)	106.739, 106.739, 163.793	73.442, 73.93, 106.257
α, β, γ (°)	90, 90, 90	88.5, 86.605, 71.173
Resolution (Å)	48.61 - 2.117 (2.193 - 2.117)	36.29 - 2.128 (2.204 - 2.128)
R_{merge}	0.1914 (1.948)	0.07557 (0.2537)
R_{meas}	0.1995 (2.101)	0.09498 (0.3193)
$I/\sigma(I)$	8.81 (1.17)	9.02 (4.37)
$CC_{1/2}$	0.997 (0.246)	0.972 (0.903)
CC*	0.999 (0.629)	0.993 (0.974)
Completeness (%)	96.96 (74.82)	97.72 (95.07)
Redundancy	7.1 (6.8)	2.7 (2.7)
Wilson B-factor	39.36	21.18
Refinement		
Resolution (Å)	48.61 - 2.117 (2.193 - 2.117)	43.28 - 2.19 (2.268 - 2.19)
No. reflections	52920 (4023)	115794 (11263)
R_{work}	0.1885 (0.3114)	0.1646 (0.1892)
R_{free}	0.2093 (0.3484)	0.1936 (0.2380)
No. atoms		
Protein	3842	14976
Ligands	125	616
B factors		
Protein	50.80	30.38
Ligands	58.58	42.35
Water	52.23	37.71
Ramachandran		
Favored (%)	94.31	96.85
Allowed (%)	2.3	2.76
Outliers (%)	0.38	0.34
R.m.s. deviations		
Bond lengths (Å)	0.033	0.009
Bond angles (°)	1.64	1.6

3.6 References

1. Tsai SC & Ames BD (2009) Chapter 2 Structural Enzymology of Polyketide Synthases. *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates,* Methods in Enzymology), pp 17-47.
2. Annaval T, Paris C, Leadlay PF, Jacob C, & Weissman KJ (2015) Evaluating Ketoreductase Exchanges as a Means of Rationally Altering Polyketide Stereochemistry. *Chembiochem* 16(9):1357-1364.
3. Baerga-Ortiz A*, et al.* (2006) Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase. *Chem Biol* 13(3):277-285.
4. Bailey CB, Pasman ME, & Keatinge-Clay AT (2016) Substrate structure-activity relationships guide rational engineering of modular polyketide synthase ketoreductases. *Chem Commun (Camb)* 52(4):792-795.
5. Caffrey P (2003) Conserved amino acid residues correlating with ketoreductase stereospecificity in modular polyketide synthases. *Chembiochem* 4(7):654-657.
6. Castonguay R, He W, Chen AY, Khosla C, & Cane DE (2007) Stereospecificity of ketoreductase domains of the 6-deoxyerythronolide B synthase. *J Am Chem Soc* 129(44):13758-13769.
7. Castonguay R*, et al.* (2008) Stereospecificity of ketoreductase domains 1 and 2 of the tylactone modular polyketide synthase. *J Am Chem Soc* 130(35):11598-11599.
8. McDaniel R, Ebert-Khosla S, Fu H, Hopwood DA, & Khosla C (1994) Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc Natl Acad Sci U S A* 91(24):11542-11546.
9. Crump MP*, et al.* (1997) Solution structure of the actinorhodin polyketide synthase acyl carrier protein from Streptomyces coelicolor A3(2). *Biochemistry* 36(20):6000-6008.
10. McDaniel R, Ebert-Khosla S, Hopwood DA, & Khosla C (1995) Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* 375(6532):549-554.
11. Meurer G*, et al.* (1997) Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. *Chem Biol* 4(6):433-443.
12. Kalaitzis JA & Moore BS (2004) Heterologous biosynthesis of truncated hexaketides derived from the actinorhodin polyketide synthase. *J Nat Prod* 67(8):1419-1422.
13. Shakya G*, et al.* (2014) Modeling linear and cyclic PKS intermediates through atom replacement. *J Am Chem Soc* 136(48):16792-16799.
14. Barajas JF*, et al.* (2017) Polyketide mimetics yield structural and mechanistic insights into product template domain function in nonreducing polyketide synthases. *Proc Natl Acad Sci U S A* 114(21):E4142-E4148.
15. Hertweck C (2009) The Biosynthetic Logic of Polyketide Diversity. *Angewandte Chemie-International Edition* 48(26):4688-4716.

16. Shen B (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* 7(2):285-295.
17. Ames BD*, et al.* (2008) Crystal structure and functional analysis of tetracenomycin ARO/CYC: Implications for cyclization specificity of aromatic polyketides. *Proceedings of the National Academy of Sciences* 105(14):5349-5354.
18. Javidpour P*, et al.* (2013) The determinants of activity and specificity in actinorhodin type II polyketide ketoreductase. *Chem Biol* 20(10):1225-1234.
19. Javidpour P, Das A, Khosla C, & Tsai SC (2011) Structural and biochemical studies of the hedamycin type II polyketide ketoreductase (HedKR): molecular basis of stereo- and regiospecificities. *Biochemistry* 50(34):7426-7439.
20. Javidpour P, Korman TP, Shakya G, & Tsai SC (2011) Structural and biochemical analyses of regio- and stereospecificities observed in a type II polyketide ketoreductase. *Biochemistry* 50(21):4638-4649.
21. Korman TP, Hill JA, Vu TN, & Tsai SC (2004) Structural analysis of actinorhodin polyketide ketoreductase: Cofactor binding and substrate specificity. *Biochemistry* 43(46):14529-14538.
22. Korman TP, Tan YH, Wong J, Luo R, & Tsai SC (2008) Inhibition kinetics and emodin cocrystal structure of a type II polyketide ketoreductase. *Biochemistry* 47(7):1837-1847.
23. Matharu AL, Cox RJ, Crosby J, Bryom KJ, & Simpson TJ (1998) MCAT is not required for in vitro polyketide synthesis in a minimal actinorhodin polyketide synthase from Streptomyces coelicolor. *Chemistry & Biology* 5(12).
24. Haushalter RW, Worthington AS, Hur GH, & Burkart MD (2008) An orthogonal purification strategy for isolating crosslinked domains of modular synthases. *Bioorg Med Chem Lett* 18(10):3039-3042.
25. Ishikawa F, Haushalter RW, Lee DJ, Finzel K, & Burkart MD (2013) Sulfonyl 3-alkynyl pantetheinamides as mechanism-based cross-linkers of acyl carrier protein dehydratase. *J Am Chem Soc* 135(24):8846-8849.
26. Worthington AS & Burkart MD (2006) One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org Biomol Chem* 4(1):44-46.
27. Nazi I, Koteva KP, & Wright GD (2004) One-pot chemoenzymatic preparation of coenzyme A analogues. *Anal Biochem* 324(1):100-105.
28. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307-326.
29. Zwart PH*, et al.* (2008) Automated structure solution with the PHENIX suite. *Methods Mol Biol* 426:419-435.
30. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.

CHAPTER 4

Probing the Determinates of the Substrate Specificity of the Doxorubicin Ketoreductase

4.1 Summary

Doxorubicin, the product of the C13 hydroxylation of daunorubicin by the cytochrome P450 DoxA in the *Streptomyces peucetius* substrain ATCC 29050(1), has become one of the most important anticancer therapeutics of all time(2). Herein, we present the crystal structure of the C9 KR (DpsKR) from the type II doxorubicin PKS. The study of DpsKR crystal structures from two different crystal forms and two Dps/ActKR chimeras shed light on the role of the lid helices, α6 and α7, in type II KR-substrate interactions.

4.2 Introduction

In the 1950's and 60's, there was a mass effort to discover soil-bacteria that produce anticancer compounds. During that time, a Streptomyces strain producing that would lead to one of the most important anticancer drugs of all time, doxorubicin, was isolated from the soils around Castel del Monte, Italy. This first strain, *Streptomyces peucetius*, produced the red-pigmented anticancer drug daunorubicin. Later, a sub-strain producing the more powerful anticancer drug, doxorubicin, which is a hydroxylated version of daunorubicin, was discovered. Since then much work has been done towards understanding and optimizing the biosynthesis of doxorubicin.

Doxorubicin is the product of a type II polyketide synthase (PKS)(3). Type II PKSs are producers of a variety of polyaromatic compounds, such as the anthracyclines, of which doxorubicin is a member. The distinguishing architectural feature of type II PKSs is that, in

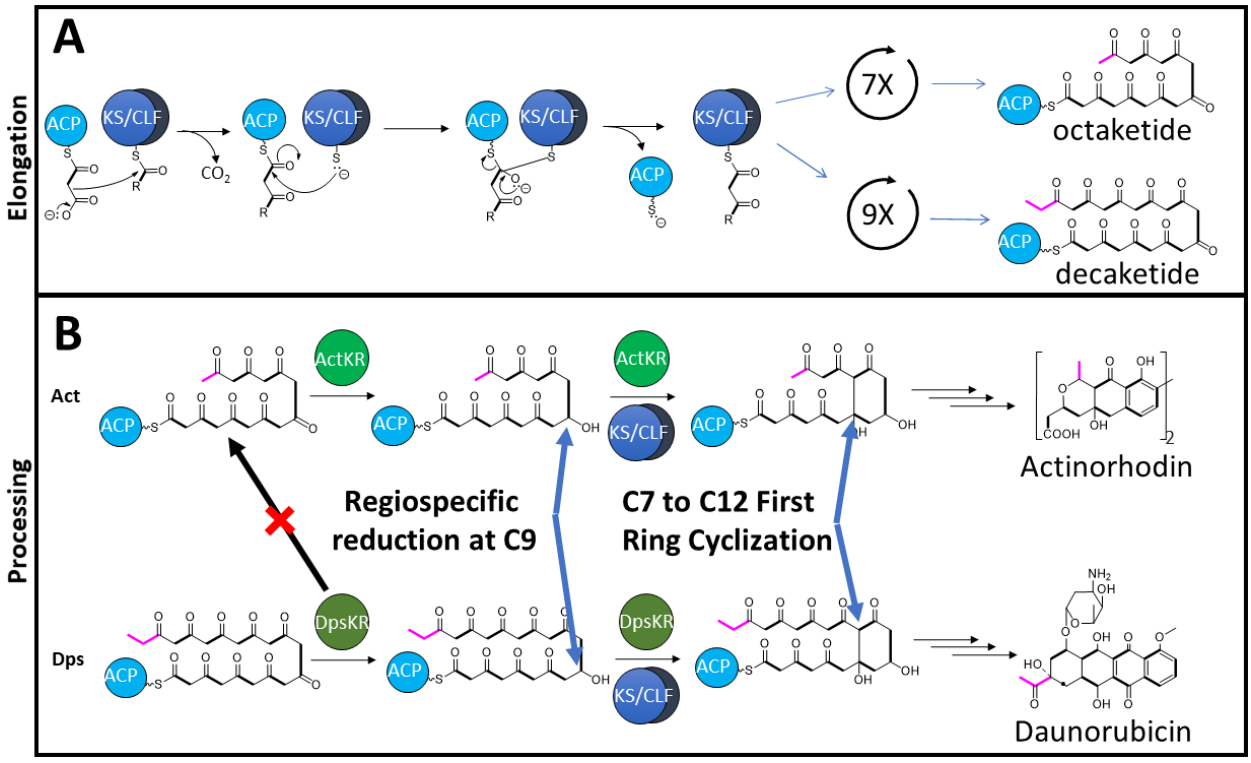


Figure 4.1 Early stages of actinorhodin and daunorubicin biosynthesis. **A)** Polyketide biosynthesis begins with starter unit selection (shown in magenta: acetyl group in the act system and propionyl group in the Dps system). The ACP then continues to deliver malonyl extender units to the KSCLF as it conducts iterative rounds of decarboxylative Claisen condensation, extending the poly-beta-ketone intermediate to its mature length (16 carbons by the act KSCLF and 21 carbons by the Dps KSCLF). **B)** Both the ActKR and DpsKR function in the same role in their respective pathways, reducing the C9 carbonyl and facilitating first ring cyclization. Importantly, the DpsKR has been shown, both in vivo and in vitro, to be unable to replace the ActKR in the Act PKS, while the reverse is possible (ActKR can replace DpsKR in the Dps PKS).

general, each step in the chemical transformations along the path from substrate to product is catalyzed by an enzymatic activity encoded on a single polypeptide. The biosynthesis of doxorubicin begins with the selection and loading of the propionate starter unit by DpsC, a beta-ketoacyl:acyl carrier protein synthase III (KSIII), onto the acyl carrier protein (ACP) (Figure 4.1). The ACP anchors the growing intermediate and shuttles it from one PKS enzyme to the next along the biosynthesis pathway. After the starter unit has been selected, the next step is to elongate it into a chain of mature length. The elongation is catalyzed by the ketosynthase/chain length factor (KSCLF) through iterative decarboxylative Claisen condensations of malonyl-CoA derived two-

carbon extender units which are delivered by the ACP. The result of the KSCLF-catalyzed elongation is a 21-carbon long polyketone intermediate. The next chemical transformation of the intermediate involves a C7-C12 first-ring cyclization and a ketoreduction of the carbonyl at C9. It is unclear which occurs first, the cyclization or the ketoreduction, but evidences suggest that the KR also influences the first ring cyclization in addition to catalyzing the regiospecific NADPH-dependent reduction of the ninth carbonyl (C9) group. The C7-C12 cyclized, C9 reduced beta-polyketone intermediate is then further processed by downstream enzymes into doxorubicin or its C14-unhydroxylated precursor, daunorubicin.

Despite considerable research effort invested in studying type II PKSs, and in particular, the doxorubicin PKS, there still remain many outstanding questions about the subject of this work, the DpsKR, and by extension, type II KRs in general. These questions include 1) to what extent do protein-protein interactions play a role in type II PKS biosynthesis? 2) What determines the substrate specificity of the type II C9 ketoreductase?

In this work, two crystal forms of DpsKR are presented, which show that the DpsKR is structurally similar to the ActKR in all major respects; therefore, a detailed structural comparison in the active site will allow us to critically evaluate the difference of substrate specificity in these two ketoreductases. In addition, a chimeric DpsKR sheds some light on the molecular determinates of substrate specificity and suggests a role for protein-protein interactions.

4.3 Results and discussion

4.3.1 The crystal structure of the doxorubicin C9 ketoreductase

In order to better understand the molecular determinates underlying the interesting behavior of DpsKR, we solved the DpsKR structure in two different crystal forms, the

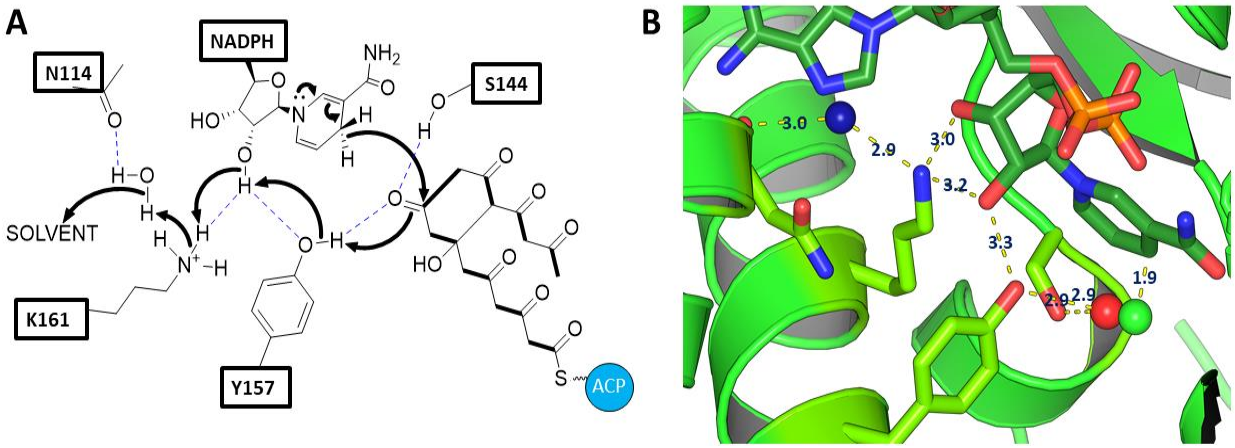


Figure 4.2 Proton wire in DpsKR. **A)** Scheme for reduction mechanism in type II KR. **B)** Proton wire in DpsKR. The green and red spheres are where the carbonyl carbon and oxygen, respectively, of the substrate need to be positioned for reduction.

first one without the cofactor (space group $P6_122$), and the second with 3 mM NADPH added (space group $C222_1$). Though NADPH was not added to the protein when it crystallized in $P6_122$, one of two monomers in the asymmetric unit is seen to be bound with NADPH (or possibly NADP+) picked up endogenously from the expression host *E. coli*. The structures reveal a prototypical type II PKS C9 ketoreductase built upon the Rossman fold, where the NAD(P)(H) cofactor binds, characteristic of all short-chain dehydrogenase/reductase (SDR) enzymes. The active site tetrad N114:S144:Y157:K161 is also conserved. Also seen is the water molecule hydrogen bonded to K161 that participates in the proton-wire responsible for reprotonating the active site tyrosine at the end of the reductive catalytic cycle (Figure 4.2). The DpsKR shares high sequence (58.2% and 63.7%, respectively) and structural homology with the two previously solved type II ketoreductases, ActKR and HedKR. The two crystal forms combined give 12 non-symmetrically related monomers. The mean, max and min RMSDs from all pair-wise alignments between all DpsKR monomers (12) and all ActKR (6) monomers are 0.47, 0.57 and 0.39 Å, respectively, while the mean, max and min RMSDs from all pair-wise

alignments between all DpsKR monomers and all HedKR (2) monomers are 0.42, 0.48 and 0.38 Å, respectively, showing that the DpsKR is slightly more similar structurally to the HedKR. The RMSD of alignment of monomers A and B in the $P6_122$ structure is 0.25 Å and the average RMSD of all pairwise alignments of the 10 monomers (A-J) in the $C222_1$ structure is 0.16 Å.

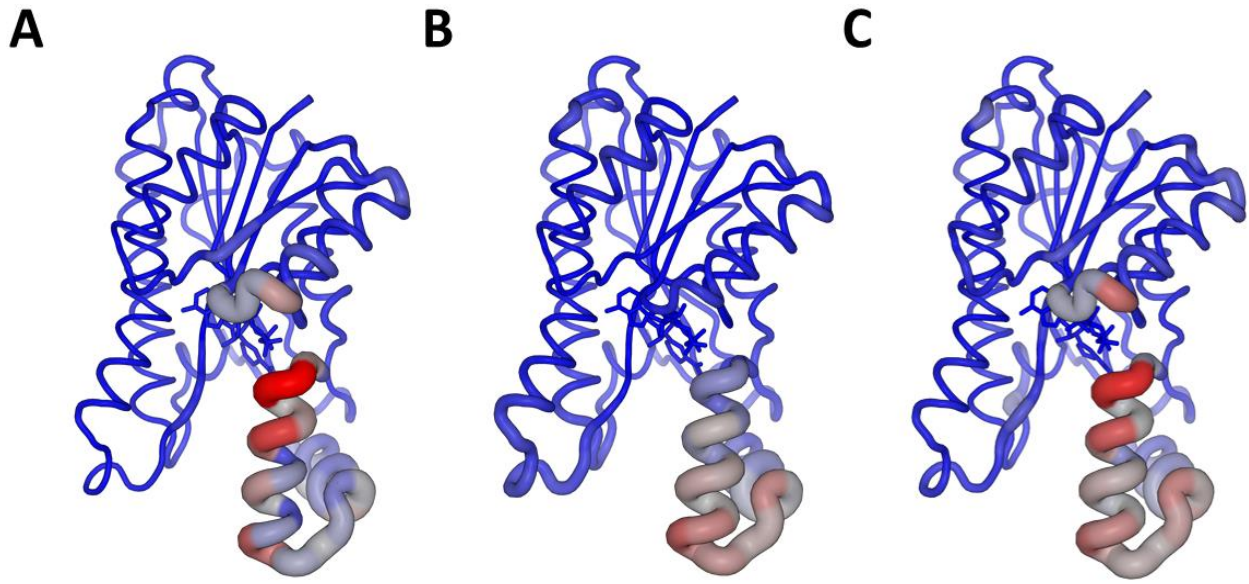


Figure 4.3 Structural variability among DpsKR monomers among **A)** the $P6_122$ crystal form, **B)** $C222_1$ crystal form and **C)** all monomers from both forms. Both crystal forms show significant variation in the alpha helices 6 and 7 (the 'lid' region). In addition, the $P6_122$ form (panel A) shows much higher variability at the top of helix 6 and the bottom of helix 2 which may cause by the binding of NADPH to monomer A. In panel C, we see that there is little structural variation between monomers of the two crystal forms, other than the flexible lid region and the top of The thickness of the chain and the color on a 'blue-white-red' spectrum are scaled according to displacement (panel A) or average deviation from the centroid (panes B and C) for equivalent backbone atoms of aligned monomers.

The top of helix 6 and the bottom of helix 2 in the $P6_122$ crystal form show a large displacement between the two monomers in the asymmetric unit (Figure 4.3A). This may represent a change in conformation due to cofactor binding, because the cofactor NADPH is bound in only one monomer of the dimer. In the $C222_1$ crystal form, the only significant variation seen between the 10 monomers in the asymmetric unit is in the flexible lid region. No significant variation was seen between monomers of the two different crystal forms besides those areas of variability already specifically mentioned (Figure 4.3C)

4.3.1. NADPH b-factors are correlated with lid b-factors and displacement

The asymmetric unit of the C2221 crystal form contains 10 monomers associated into two tetramers (ABCD and EFGH) and one dimer (IJ) (Figure 4.4D). The dimer IJ forms a tetramer with its symmetry related dimer *IJ (not shown). The flexible lid region of monomers B, F, I, and J, are completely free of packing contacts and monomer G has minimal packing contacts (Figure 4.4ABC). This offers an opportunity to investigate the relationship between NADPH-binding stability and the lid region conformation by comparing average b-factors in the absence of the influence of direct packing contacts. A general trend is observed where the magnitude of the average b-factor of the lid region is

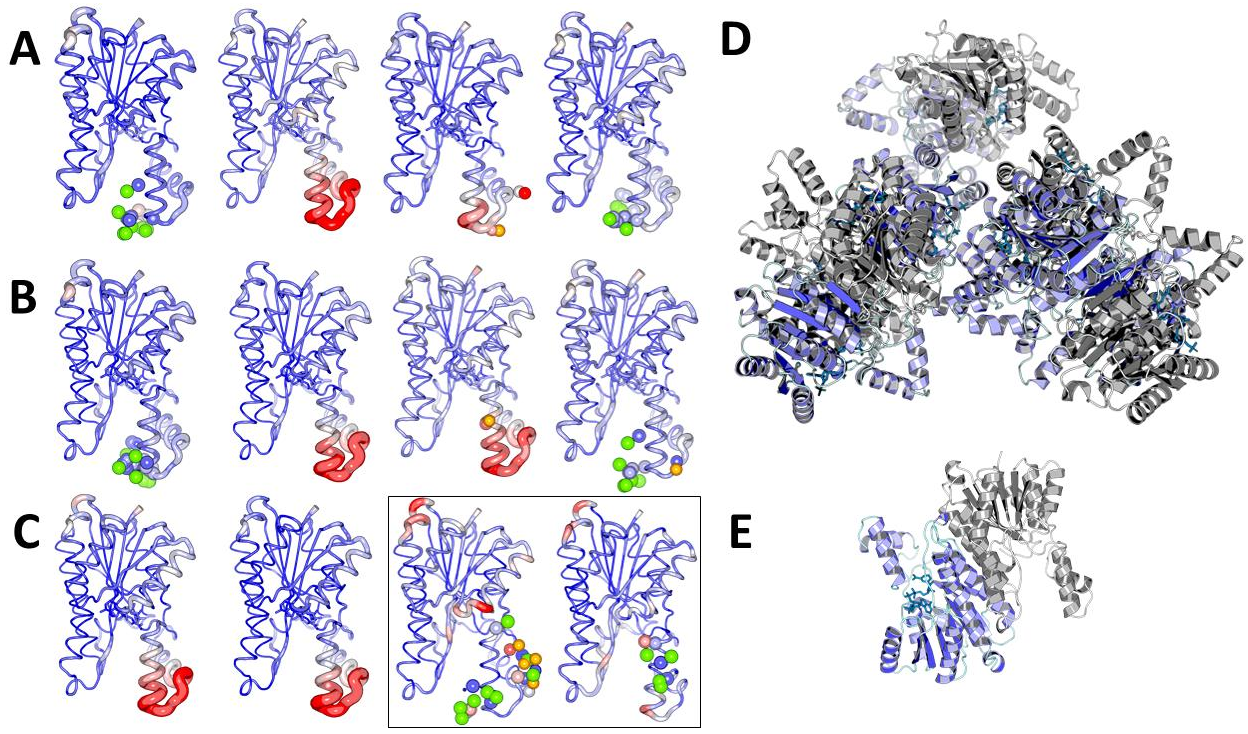


Figure 4.4 Comparing b-factors and packing contacts between the monomers in both crystal forms. A) Monomers A-D forming the first tetramer in C2221. B) Monomers E-H forming the first tetramer in C2221. C-unboxed) Monomers I and J in C2221, which pack with their symmetry related counterparts to form a tetramer. Note that monomers B, F, I and J are free of packing contacts, while monomer G has minimal contacts. C-boxed) Monomers A and B in the P6122 crystal form. The thickness of the chain and the color on a 'blue-white-red' spectrum are scaled according to backbone b-factors. The atoms involved in the contacts on the lid region of each monomer are shown as spheres. The spheres of the contacting molecule are colored according to the type of contact: green for carbon-carbon Van der Waals, orange for hydrogen bond and red for salt-bridge. The spheres of the lid region retain their color according to b-factor. D) The two and a half tetramers in the asymmetric unit of the C2221 crystal form. E) the two monomers in the P6122 crystal form.

directly proportional to the magnitude of the average NADPH b-factor (Figure 4.5A). In addition, the position of the lid region moves outward as its average b-factor increases, opening the substrate binding pocket (Figure 4.5B).

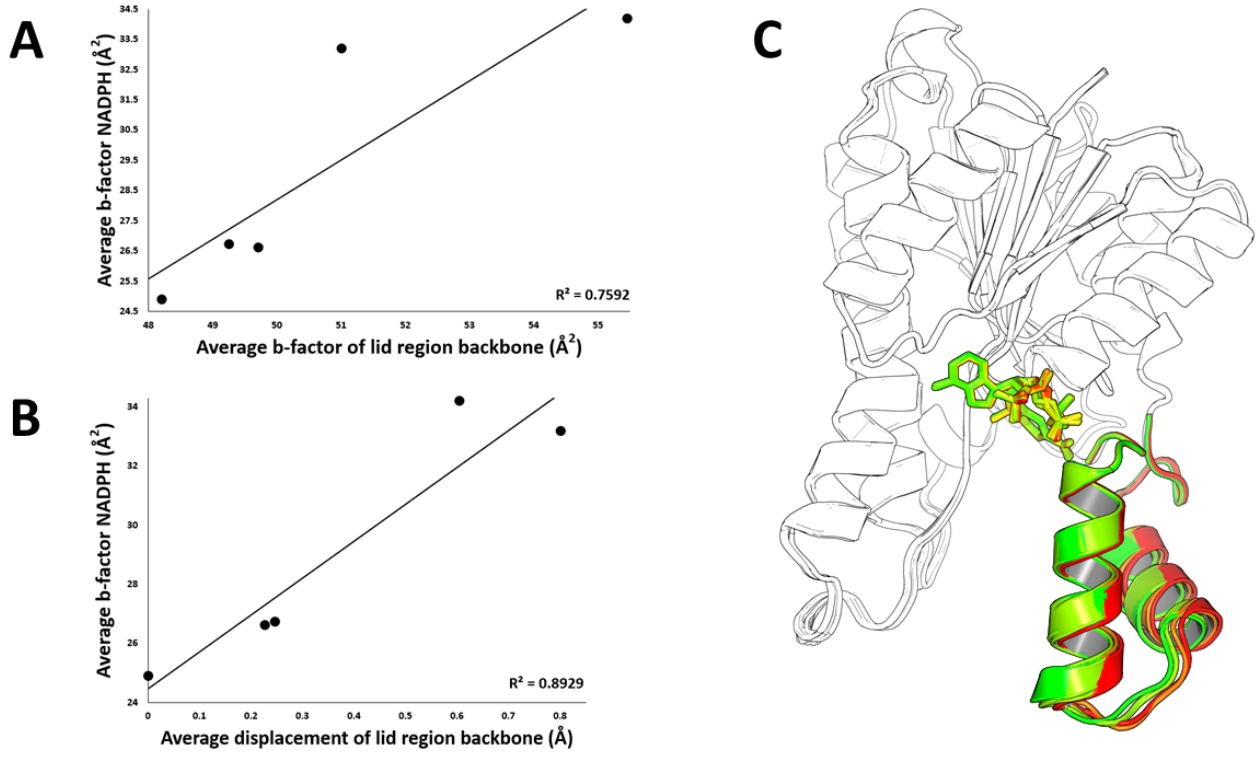


Figure 4.5 Comparing the NADPH cofactor b-factor for each monomer with the b-factor and displacement of the lid region. There is a high correlation between the b-factor magnitude of the NADPH with both the average b-factor (**panel A**) and the average displacement (**panel B**) outward of the lid region (R2 = 0.76 and 0.89, respectively). **C)** Cartoon showing the outward displacement of the lid region as the average b-factor increases. The cofactor NADPH is colored on a 'green_yellow_red' spectrum, scaled according to atom b-factor relative to other atoms in the cofactor. The lid region is colored similarly to the cofactor, but with colors scaled relative to other backbone atoms in the lid region.

This is consistent with the ordered 'bi-bi' mechanism proposed for SDR enzymes in which the cofactor NADPH binds before substrate binding, and then is released after product release (4). A conformational change in the substrate upon reduction would cause the ACP-bound product to disassociate from the KR. The now freed lid region would swing out and at the same time cause subtle destabilizing changes in the cofactor binding site, allowing the spent NADP+ cofactor to exit and be replaced by a fresh NADPH.

Binding of a new ACP-bound substrate would lock in a fresh NADPH and stabilize the productive conformation of the active site, beginning a new catalytic cycle.

It must be noted that the two complete tetramers in the asymmetric unit pack against each other at the lid region of monomers A and E in a two-fold pseudosymmetrical way (Figure 4.6), thus the b-factors in monomers B and F (which form tight dimers with A and E, respectively) may not be completely independent. Indeed, it would be quite interesting if this proved to be the case, as it might reveal potential allosteric interactions important to the catalytic cycle. This is a future direction that we will pursue.

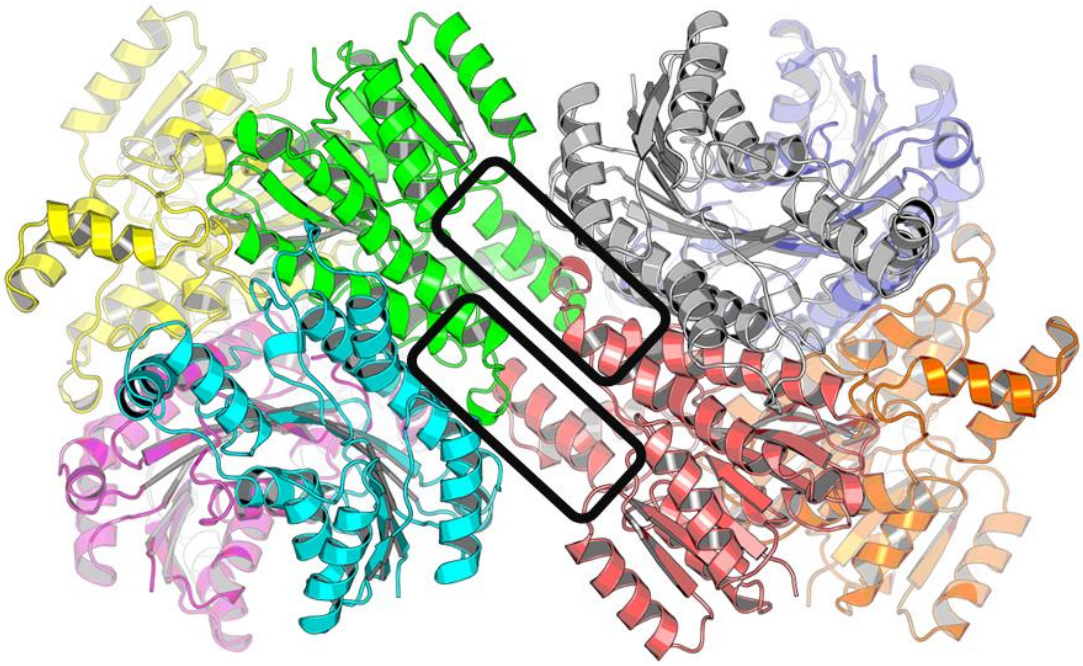


Figure 4.6 The two complete tetramers of the C2221 asymmetric unit (last two monomers, which break up the symmetry, are not shown). A highly symmetrical contact between monomers A and E is observed, which may artificially increase the correlations seen between cofactor b-factor and lid region b-factors and displacement.

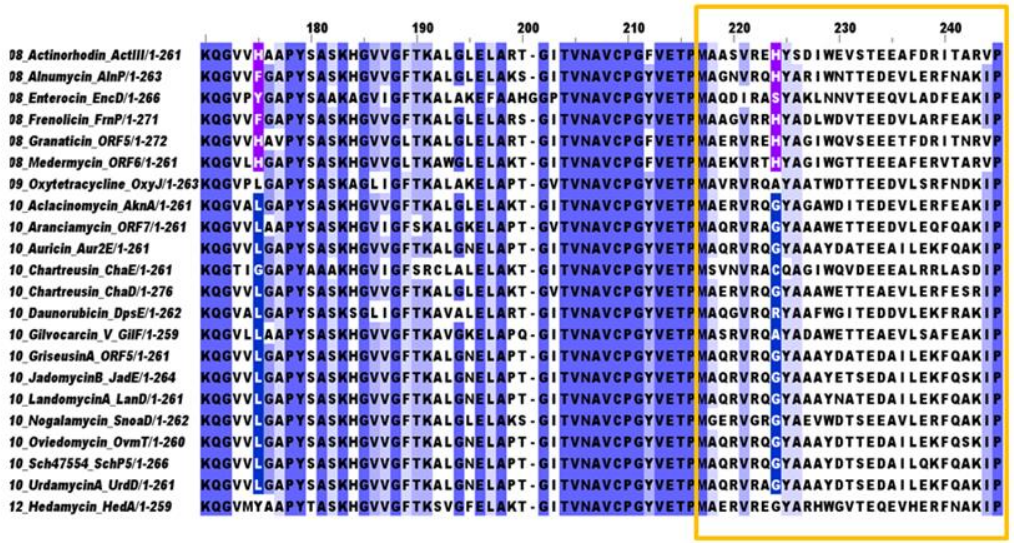


Figure 4.7 The sequence of the lid region is highly variable. Shown is the alignment of type II C9 KRs from PKSs producing product chain lengths from 8 to 12 ketides. The sequence corresponding to the lid region is boxed in yellow.

4.3.4 Substrate specificities of type II C9 ketoreductases

The type II C9 ketoreductase upholds a strict regiospecificity of ketoreduction, always reducing the ninth carbonyl from the thioester linkage to the ACP. While highly regiospecific, some KRs can accept substrates of a range of lengths, like the HedKR, while others, like the ActKR, have a narrower tolerance of varying chain lengths. A sequence alignment done with type II C9 KRs shows that the lid region—alpha helices 6 and 7—is highly variable (Figure 4.7), suggesting that this region may be involved in KR substrate specificity. Interestingly, it was discovered by Meurer et al that the DpsKR was incompetent when expressed together with the actinorhodin PKS *in vivo* (5). This was confirmed *in vitro* in the present work. These results suggest that the DpsKR cannot recognize octaketide substrates. There is also the possibility, however, that favorable interactions between the KR and KSCLF are important in the actinorhodin PKS and that the DpsKR does not harbor the required recognition patch. In order to better understand the molecular basis for DpsKR inability to function within the actinorhodin PKS, two chimeras were generated and assayed *in vitro*.

4.3.5 DpsKR:ActKR Chimeras exhibits significant activity relative to the ActKR

Past mutation studies and the variability of the lid region of type II KRs implicate this region as being important in substrate recognition (6). To understand how the lid region affects the specificity of KR, we underwent a protein engineering effort to switch the lid region. As a first approach, the lid region of DpsKR was replaced with the lid region of ActKR, generating the DpsActKR lid chimera, or simply, 'lid chimera' (Figure 4.8). When the actinorhodin minimal PKS (KSCLF, ACP, MCAT) was reconstituted with the lid chimera, the lid chimera exhibited 39 % activity relative to ActKR (Figure 4.9), strongly confirming the role of the lid region in controlling substrate specificity, but not completely explaining it. It is likely that the swapping of lid region somehow disrupted the overall protein folding of the chimeric DpsKR, leading to a reduced activity.

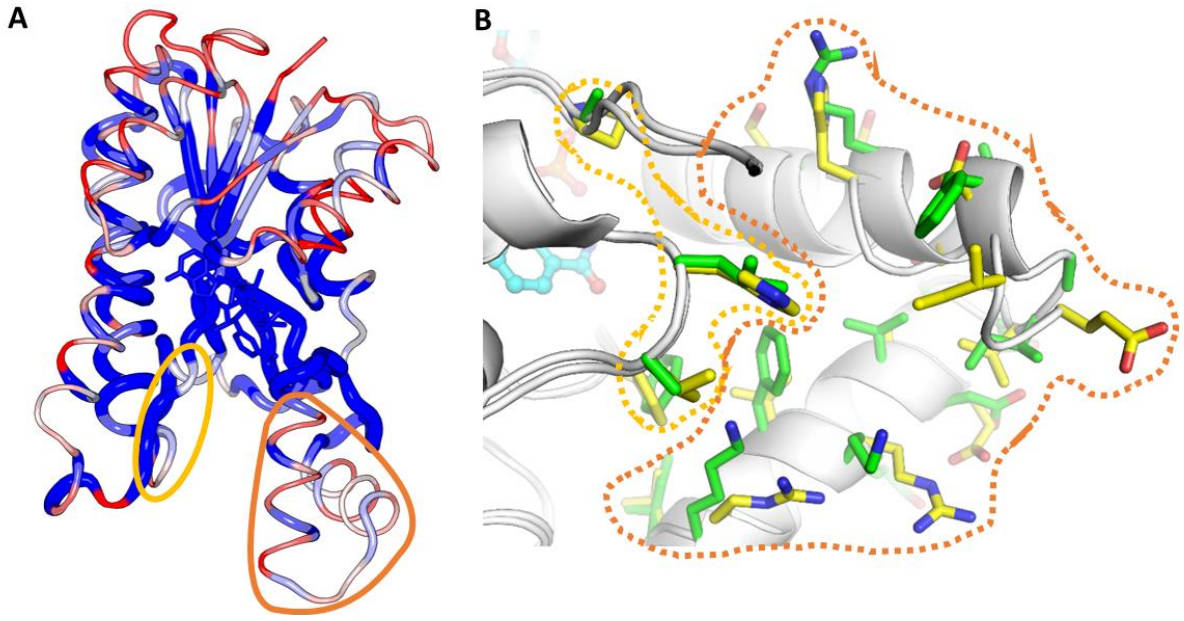


Figure 4.8 A) Cartoon representation of a type II C9 ketoreductase colored according to sequence conservation. The active-site (circled in yellow) is highly conserved, while the flexible lid region (circled in green) is not. **B)** Comparison of DpsKR (in green sticks) and ActKR residues (in yellow sticks) highlighting the regions changed in the chimera (lid circled in green and inner wall in red).

As a second approach, a structural overlay of the DpsKR and ActKR structures indicated that there were three additional residues on the side of the binding pocket opposite to the lid that were different between the two ketoreductases: G94, A152 and L153. The mutation of these last three different residues in the substrate binding pocket to their corresponding ActKR residues (P, V and H, respectively) generated the DpsActKR pocket chimera, or simply the 'pocket chimera'. When this pocket chimera was assayed with the actinorhodin minimal PKS, its activity was seen to leap to 89 % of ActKR activity. These results suggest that the inability of DpsKR to yield reduced products when reconstituted with the actinorhodin PKS, both *in vivo* and *in vitro*, is mostly an issue of substrate recognition by the substrate binding pocket. The remaining missing 10 % activity may be due to incorrect protein-protein interactions between the DpsKR and the Act

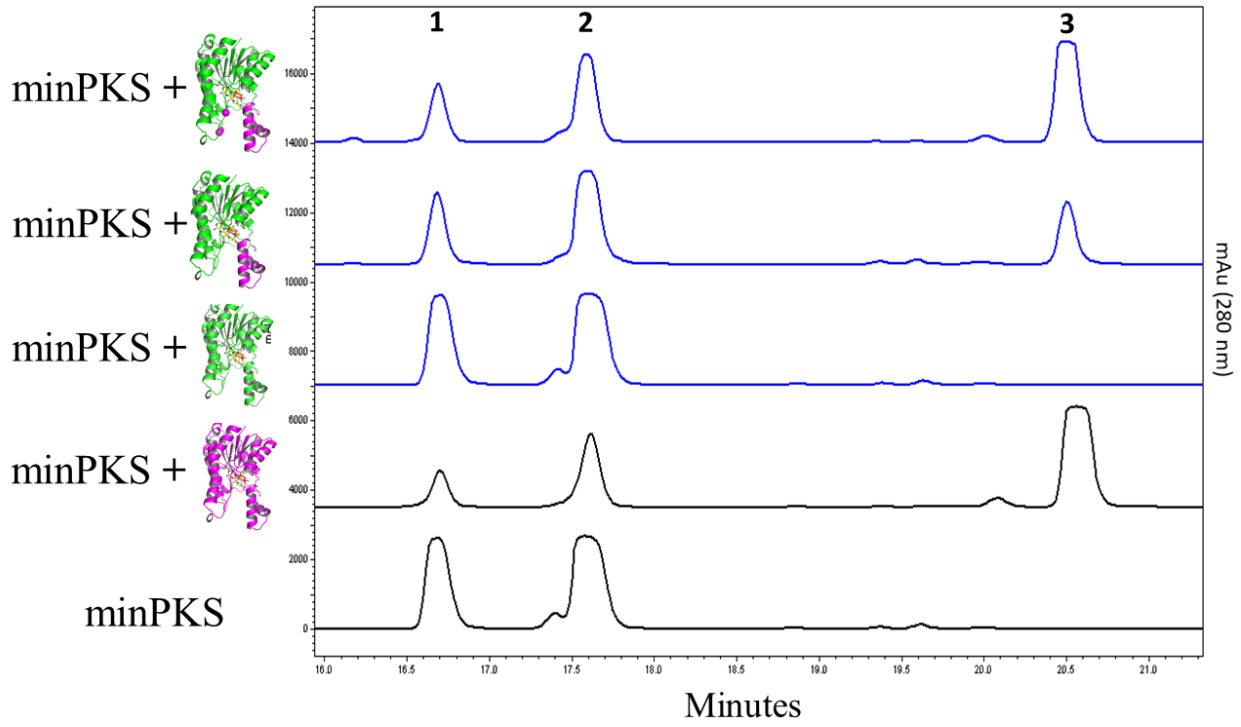


Figure 4.9 HPLC product profile of the Act minimal PKS (Act minPKS) when reconstituted with the ActKR, DpsKR, DpsActKR Lid Chimera and DpsActKR Full Pocket Chimera. Traces, from bottom to top: 1) Act minPKS alone, 2) Act minPKS with ActKR, 3) Act minPKS with DpsKR, 4) Act minPKS with DpsActKR Lid Chimera, and 5) Act minPKS with DpsActKR Full Pocket Chimera. The products peaks are: 1) sek4, 2) sek4b and 3) mutactin. Product peaks were identified by typical published retention times.

KS/CLF, or more likely, the missing activity represents a loss of structural fine-tuning and of a lack of dynamic structural cooperation between the native and foreign parts of the chimera, reducing its efficiency. To the best of our knowledge, this is the first success of protein engineering that changes the chain length specificity of a PKS KR domain.

4.4 Conclusions

The crystal structure of the DpsKR reveals that it shares high structural homology with the ActKR. It had previously been shown *in vivo* that the DpsKR (specific for a 21-carbon nonaketide) is inactive when reconstituted with the actinorhodin system (specific for a 16-carbon octaketide). This result was confirmed in the current study *in vitro*. Structural engineering studies show that the lid region plays a moderate role in the observed incompatibility. Replacing the lid region of DpsKR with the lid region of ActKR restores activity by 39 %. The modest stature of this increase may indicate a disruption of protein folding when we replaced the entire lid region. Rather, with the conversion of only three residues on the opposite side of the pocket, the substrate binding pocket is completely converted to the ActKR complement. This relatively small change resulted in a large jump in activity—to 89 % of the ActKR activity. Interestingly, the complete conversion of the substrate binding pocked did not grant complete activity. There are two possible explanations for this: 1) There are specific recognition patches on the surface of ActKR that are important for protein-protein interactions among the ActPKS enzymes. Since the arginine-patch, proposed to be the ACP binding site, is the same between the Dps and ActKR, the specific recognition patches would likely be recognized by the KSCLF. Interactions between the KR and the KSCLF have been

hypothesized, but not previously shown. 2) It may also be that the pocket transplant was indeed enough to change the substrate specificity, but at the same time disrupted subtle conformational communication across the KR that is important for facilitating efficient turnover. In conclusion, work presented in this chapter shows the first success of structure-based mutation of a polyketide KR that changes the chain-length specificity. This will have a high impact on future engineering work conducted for directed biosynthesis of polyketides.

4.5 Materials and methods

4.5.1 Expression and purification of DpsKR and ActKR

The recombinant N-terminal histidine tagged DpsKR and ActKR proteins were expressed in BL21 (DE3) *E. coli* cells (Novagen). Cells containing the KR-harboring plasmids were grown to OD_{600} 0.6 at 37 °C in LB medium containing 50 µg/ml kanamycin. The cell cultures were cooled to 18 °C and expression was induced using 1 mM IPTG. The cell cultures were incubated for an additional 16 hr. at 18 °C and harvested by centrifugation at 5,525 r.c.f. for 15 min. The cell pellets were resuspended in 50 mM Tris-HCl, pH 7.5, 10 % glycerol, 10 mM imidazole, 300 mM NaCl and 1 mg/ml lysozyme. Resuspended cells were cooled on ice for 30 min and the cells were disrupted using sonication. The cell debris was cleared by centrifugation at 21,036 r.c.f. for 1 hr. The supernatant was collected and batch bound to HisPur™ Cobalt Resin (Thermo Scientific) for 1 hr. at 4 °C. The proteins were purified according to the manufacturer's instructions using an imidazole step-gradient. Fractions containing pure protein were determined by SDS-PAGE and fractions containing the ketoreductase were combined and dialyzed against 50 mM Tris-HCl pH 7.5, 10 % glycerol, 300 mM NaCl at 4 °C for 12 h. Purified ketoreductase was concentrated to

10-12 mg/ml and buffer exchanged into crystallization buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 5 % glycerol) for crystallographic studies.

4.5.2 Crystallization and data collection

DpsKR (7 mg/ml) was screened at room temperature against several Qiagen NeXtal screening kits in 24-well sitting-drop vapor diffusion trays. One condition from the Classics I screen (0.1 M HEPES, pH 7.5, 0.2 M MgCl2, 30 % PEG 400) yielded a single well-formed crystal. The crystal was washed in well solution and flash frozen in liquid nitrogen. Diffraction data was collected out to 2 Å on Beamline 8.2.1 at the Advanced Light Source using monochromatic X-rays (0.99994 Å). DpsKR (7 mg/ml) was also screened with 3 mM NADPH at both room temperature and 4 ºC. One condition (0.1 M Tris, pH 8.5, 0.25 M NaC1, 31 % PEG 3350) yielded thin rods at 4 ºC. Several rounds of optimization and seeding yielded diffraction quality crystals, which were washed and flash frozen in liquid nitrogen. Diffraction data was collected out to 2.34 Å on Beamline 8.2.2 at the Advanced Light Source using monochromatic X-rays (1.0000 Å).

4.5.3 Phasing, model building and refinement

The DpsKR (endogenous NADP(H)) diffraction data was first indexed, integrated, scaled (using HKL2000 (7)) in space group P6122 and phased through molecular replacement with ActKR (PDB 1X7G) (using PHENIX Phaser) by Dr. Grace Festin, using R_{merge} to cutoff the resolution at 2.99 Å. Later, the data was reprocessed using $CC_{1/2}$ to guide resolution cutoff as recommended by Karplus and Diederichs (8). Keeping the higher resolution data (out to 2

Å) resulted in clearer electron density maps. The model was improved through iterative cycles of model building using COOT (9)and model refinement using PHENIX Refine (10).

The DpsKR (NADPH) diffraction data were indexed, integrated, and scaled in space group C2221 using HKL2000. The resolution was cutoff at 2.34 Å using CC1/2 as a guide. Initial phases were determined through molecular replacement with the DpsKR (APO) structure using PHENIX Phaser. The model was improved through iterative cycles of model building (using COOT) and model refinement (using PHENIX Refine) (Table 4.1).

Table 4.1 Summary of X-ray crystallographic data collection and data refinement statistics.

	DpsKR P 6_122	DpsKR C 222_1
Data collection		
Wavelength (Å)	1.000	1.000
Total reflections	1923057 (106737)	1524650 (117421)
Unique reflections	51635 (4984)	133922 (12875)
Space group	P 6_122	C 222_1
Cell dimensions		
a, *b*, *c* (Å)	102.902, 102.902, 242.009	87.401, 224.263, 360.897
α, β, γ (°)	90, 90, 120	90, 90, 90
Resolution (Å)	47.35 - 2.004 (2.076 - 2.004)	46.84 - 2.347 (2.431 - 2.347)
R_{merge}	0.3832 (8.291)	0.128 (0.6814)
R_{meas}	0.3885 (8.486)	0.1338 (0.7215)
$I/\sigma(I)$	10.02 (0.30)	14.66 (3.39)
$CC_{1/2}$	0.998 (0.119)	0.997 (0.842)
CC*	0.999 (0.461)	0.999 (0.956)
Completeness (%)	96.33 (64.61)	90.52 (87.73)
Redundancy	37.2 (21.4)	11.4 (9.1)
Wilson B-factor	41.89	28.4
Refinement		
Resolution (Å)	47.35-2.004 (2.076 -2.004)	46.84 - 2.347 (2.431 - 2.347)
No. reflections	49808 (3263)	133897 (12875)
R_{work}	0.1928 (0.3974)	0.1704 (0.2225)
R_{free}	0.2250 (0.4032)	0.1919 (0.2844)
No. atoms		

Protein	3735	18658
Ligands	48	480
B factors		
Protein	43.61	32.94
Ligands	47.39	29.25
Water	53.12	38.8
Ramachandran		
Favored (%)	98.04	97.28
Allowed (%)	1.77	2.48
Outliers (%)	0.2	0.24
R.m.s. deviations		
Bond lengths (Å)	0.008	0.01
Bond angles (°)	1.30	1.79

4.5.6 Chimera construction

The region between the C-terminal end of beta strand 6 and the N-terminal end of alpha helix 8 (Y189-M228) of DpsKR was replaced by the splicing in of the homologous region of ActKR. This corresponds to alpha helices 6 and 7—the lid region—and the two loops that connect them to the main KR body. The N-terminal boundary of the splice site was chosen because it is preceded by an 11-residue-long identity sequence shared between DpsKR and ActKR. A similarly shared stretch of residues near the C-terminal end of the splice region does not exist, so this boundary was chosen to be the junction between alpha helix 8 and the loop connecting to the C-terminal end of the lid region. The construction of the lid chimera was effected through a technique termed ‘multi-template sequential PCR’ (MTS-PCR) (Figure 4.10)(11). In this technique, the desired chimeric DNA molecule is built up step-wise in sequential PCR reactions. In this case, it was desired to insert a portion of ActKR into the middle of DpsKR, therefore, what we wanted was the N-terminal region of DpsKR, a central portion of ActKR, and the C-terminal portion of DpsKR. In the first PCR reaction, the 5’ end of dpsKR was amplified from a dpsKR template. The 3’ half of the reverse primer used in the reaction

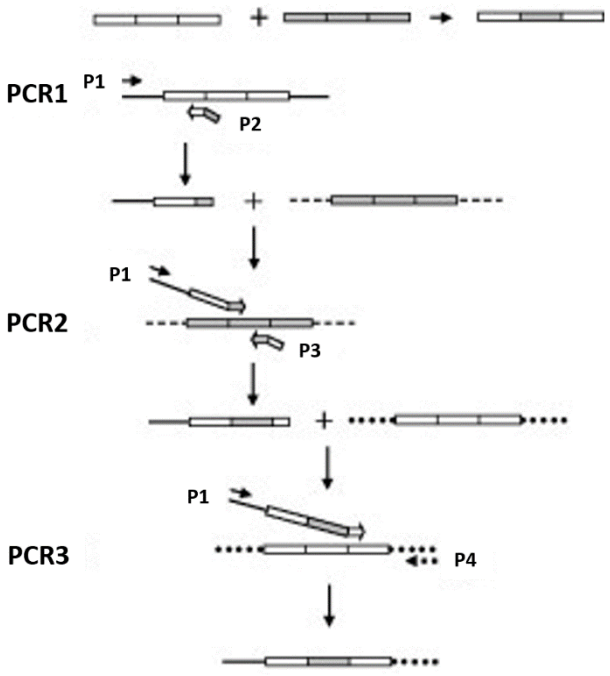


Figure 4.10 Scheme of multi-template sequential PCR. See text for details. Figure excerpted from a methods paper by Shan and Lynch (11).

was made up of a sequence complementary to the 5' end of the desired actKR insert, thus allowing the product of the first PCR to be used as a forward megaprimer in the second PCR. The megaprimer, containing the desired 5' region of dpsKR, which was generated in the first PCR reaction, was used as the forward primer (1:2500 dilution) in a second PCR reaction in which the actKR gene was used as the template. Again, as was the case in the first PCR reaction, the 3' half of the reverse primer used in the second PCR reaction was complementary to the desired 5' end of the dpsKR gene. The construction of the desired chimeric DNA molecule was completed in the third PCR reaction, in which the product of the second PCR reaction—now an even larger megaprimer containing the 5' portion of dpsKR and the central portion of actKR—was used as the forward primer (1:2500 dilution) in the third and last PCR reaction in which the forward primer was extended with the desired 3' region of dpsKR using dpsKR as the template. In order to maximize the amount of PCR product produced in the second and third PCR reactions, the forward primer used in the first PCR was also added to the second and third PCR reactions, in order to generate more forward primer as the reactions proceed. In order for this strategy to work, primer 1 must not bind anywhere on the templates used in the second and third PCR. Therefore, primer 1 must be made to specifically bind to the vector sequence upstream of the insert and not to the insert itself. Additionally, each gene

template had to be cloned into a different plasmid vector. The DpsKR gene was cloned into pUAB400 and pUC18 (using the NdeI and EcoRI cut sites), while the ActKR was left in the pET28 expression vector. The product of the third PCR—the completed 'lid chimera'—was digested with NdeI and EcoRI and ligated into pET28 for expression. The PCR products were able to be used directly without purification before use in subsequent PCRs or the ligation reaction. The additional three mutations required to convert the 'lid chimera' into the 'full pocket chimera' (G94P, A152V and L153H) were generated using site-directed PCR.

Table 4.2. Primers used for the construction of the DpsKR 'lid' and 'full' chimeras

MTS-PCR Primer 1	GACGGGAATCGAACCCGCGTAGCTAGTTTGGAAGAATGGG
MTS-PCR Primer 2	CTCGACGAATCCGGGGCAGACGGCGTTC
MTS-PCR Primer 3	CCACCTCCTCGGGCTGCACGTACCG
MTS-PCR Primer 4	GCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTG
G94P_Forward	GGCCGCCCCGGTGGAGGCGTGACCGCCGAGATCACG
G94P_Reverse	TCCACCGGGGCGGCCGGCGTTGTTGACGACGATGTCG
A152V_L153H_Forward	GAGTGGTCCACGGCGCCCCCTACTCCGCCTCCAAGAGC
A152V_L153H_Reverse	GCGCCGTGGACCACTCCCTGCTTGCCGCCGGTGG

4.5.7 Actinorhodin and daunorubicin KSCLF expression and purification

ActKS/CLF was expressed in *Streptomyces coelicolor* CH999 containing the pRJC006 expression vector as previously described, but with slight modifications (12). Fresh spores were prepared by spreading 100 ul spore stock onto a mannitol soy agar plate containing 50 ug/ml kanamycin. This plate was incubated at 30 C for 7-10 days. The fresh spores were then collected and used to inoculate 50 mL of Super YEME containing 50 μg/mL kanamycin. The 50 mL cultures were grown for 3 days at 30 °C with shaking at 250 rpm. The 50 mL cultures were transferred to 2.5 L baffled flasks containing 500 ml of Super YEME and 50 μg/mL kanamycin. The cultures were grown for an additional 2 days under the same conditions. Protein expression was induced with 10 μg/mL thiostrepton. Cells were harvested 24 hours after induction by centrifugation (20 minutes at 5100 rpm) and the cell pellets were stored at -80 °C. Cell pellets from 2 L of cell culture were resuspended in 200 mL of lysis buffer (100 mM KPi, 15 % glycerol, 300 mM NaCl, pH 7.5, with 2 EDTA-free protease inhibitor cocktail (Roche) tablets). The cell suspension was lysed by sonication on ice (10 x 1 minute cycles). Cell debris was pelleted by centrifugation (1 hour at 14000 rpm) and DNA was precipitated using streptomycin sulfate (2 % final concentration) followed by centrifugation (45 min at 14000 rpm). The supernatant was filtered using a 0.45 μM filter and 30 % – 50 % (NH4)2SO4 was used to precipitate ActKS/CLF overnight. Precipitated protein was pelleted by centrifugation (30 min at 14000 rpm) and redissolved in Ni-binding buffer (50 mM KPi, 10 % glycerol, 500 mM NaCl, 5 mM imidazole, pH 7.5). After resuspension, the solution was filtered using a 0.45 μM filter then bound to 2 ml of Ni-IMAC resin (Bio-Rad) pre-equilibrated with Ni-binding buffer and stirred at 4 °C for 1 hour. ActKS/CLF was eluted using Ni-binding buffer containing increasing amounts of imidazole. Fractions (2.5 mL each) containing between 375 mM and 750 mM imidazole were individually buffer exchanged into storage buffer

(100 mM KPi, 10 % glycerol, pH 7.5) using a PD-10 column (GE Healthcare) and flash frozen at -80 °C.

4.5.8 ActACP, DpsACP, MCAT, and MatB expression and purification

Histidine-tagged constructs of ActACP, DpsACP, *Streptomyces coelicolor* MCAT, and MatB were overexpressed from pET28 in *E. coli* BL21(DE3) cells. The expression and purification were conducted as described above for DpsKR and ActKR.

4.5.9 *In vitro* reconstitution assays

DpsKR and DpsKR chimera activity were measured relative to the ActKR using an in vitro assay based on the actinorhodin minimal PKS (12) (Figure 4.11). The DpsKR was also confirmed to be active with the Dps minimal PKS (Figure 4.11). Minimal PKS activity is

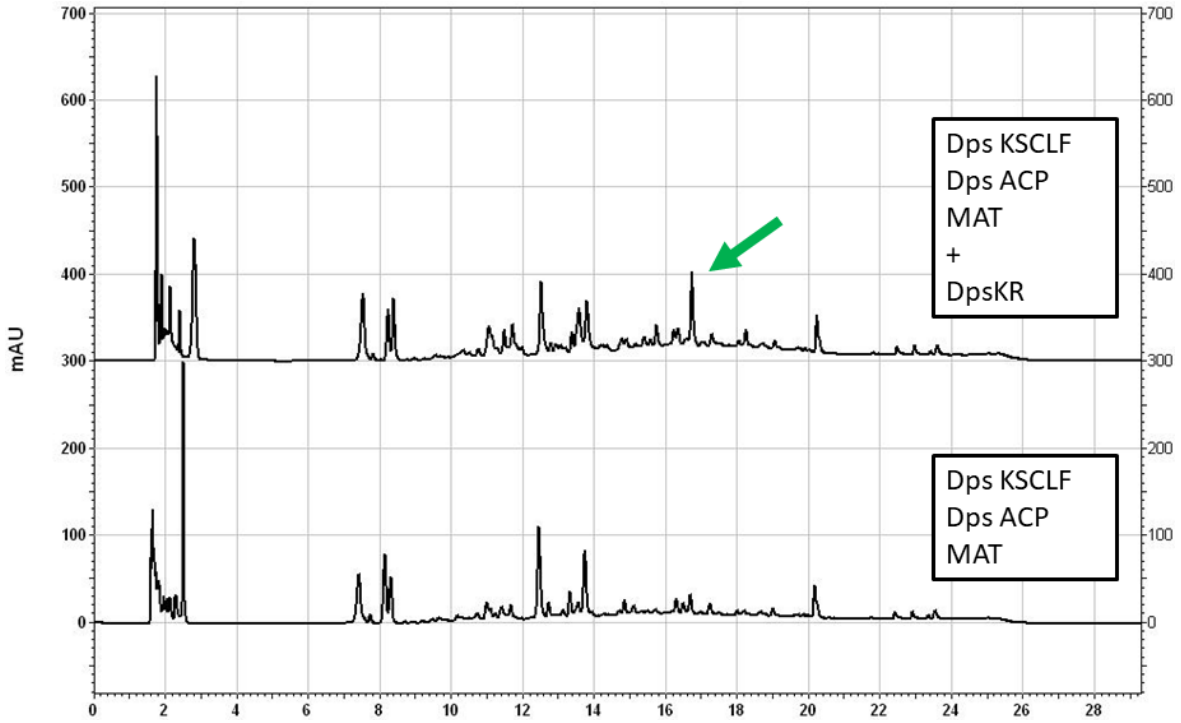


Figure 4.11 HPLC product profile of DpsKR reconstituted with the Dps minimal PKS (Dps KSCLF and ACP). **Top:** Trace of Dps minimal PKS with DpsKR showing the appearance of a new product with the addition of DpsKR. **Bottom:** Trace of minimal PKS alone with the DpsKR product absent.

reconstituted in vitro by combining the KS/CLF (7.5 μM) (responsible for chain extension), ACP (40 μM) (responsible for shuttling of the intermediates), and MCAT (5 μM) (responsible for extension unit selection). MatB (20 μM) is added in order to generate the activated extension unit, malonyl-CoA, from malonate and CoA. The reaction volume was 250 μL and also contained 30 μM KR, 2.6 mM NADPH (Sigma) (when KR is present), 100 mM malonate, 7.5 mM CoA, 10 mM ATP, and 5 mM $MgCl_2$ in 100 mM Tris, pH 7.5. The reactions were carried out at room temperature for 12 hours. The reaction mixtures were extracted three times with 300 μL of 94 : 5 : 1 of ethyl acetate : methanol : acetic acid and the organic layer was dried using Speed Vac Concentrator. The residual oil was resuspended in DMSO (Sigma) and subject to reverse-phase HPLC analysis using a Synergi Hydro-RP column (Phenomenex). The HPLC gradient was from 5 to 50 % MeCN in a H2O/0.1 % formic acid mixture over 15 minutes followed by 50 to 95 % MeCN in a H2O/0.1 % formic acid mixture over 5 min. SEK4, SEK4B, and mutactin were identified using relative retention times.

4.6 References

1. Lomovskaya N*, et al.* (1999) Doxorubicin Overproduction in Streptomyces peucetius: Cloning and Characterization of the dnrU Ketoreductase and dnrV Genes and the doxA Cytochrome P-450 Hydroxylase Gene. *Journal of Bacteriology* 181(1):305-318.
2. Tacar O, Sriamornsak P, & Dass CR (2013) Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. *J Pharm Pharmacol* 65(2):157-170.
3. Hutchinson CR (1997) Biosynthetic Studies of Daunorubicin and Tetracenomycin C. *Chemical Reviews* 97(7):2525-2535.
4. Kavanagh KL, Jornvall H, Persson B, & Oppermann U (2008) Medium- and short-chain dehydrogenase/reductase gene and protein families : the SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cell Mol Life Sci* 65(24):3895-3906.
5. Meurer G*, et al.* (1997) Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. *Chem Biol* 4(6):433-443.

6. Korman TP, Tan YH, Wong J, Luo R, & Tsai SC (2008) Inhibition kinetics and emodin cocrystal structure of a type II polyketide ketoreductase. *Biochemistry* 47(7):1837-1847.
7. Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307-326.
8. Karplus PA & Diederichs K (2015) Assessing and maximizing data quality in macromolecular crystallography. *Curr Opin Struct Biol* 34:60-68.
9. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
10. Zwart PH*, et al.* (2008) Automated structure solution with the PHENIX suite. *Methods Mol Biol* 426:419-435.
11. Shan Q & Lynch JW (2010) Chimera construction using multiple-template-based sequential PCRs. *J Neurosci Methods* 193(1):86-89.
12. Matharu AL, Cox RJ, Crosby J, Bryom KJ, & Simpson TJ (1998) MCAT is not required for in vitro polyketide synthesis in a minimal actinorhodin polyketide synthase from Streptomyces coelicolor. *Chemistry & Biology* 5(12).

CHAPTER 5

Conclusions and Future Directions

5.1 Conclusions

Polyketides represent an extremely important class of natural products whose number continues to grow with the discovery of new PKS gene clusters, which has increased with a vigorous pace as the metagenomic analysis of environmental DNA continues. The class of enzymes that produce them, PKSs, embody an elegant design paradigm that creates access to a seemingly inexhaustible pool of complex small molecules, with a significant fraction having bioactive properties. The production principle is based on the repeated use of simple enzymatic domains in a highly programmed and coordinated fashion. This type of modular system has great potential to be redesigned for the clean and efficient production of tailored chemicals, including new drugs. Although our understanding of the sequence-structure-function relationship of proteins is still many years from being complete enough to rationally guide the redesign of such systems, every structure solved, and every mutational analysis conducted, no matter how humble, brings us one hour closer to that day.

5.1.1 Application of a polyketone mimetic to the study of the product template domain (Chapter 2)

Fungal non-reducing polyketide synthases (NR-PKSs) conduct multiple rounds of condensing simple thioester-activated carbon extender units together through decarboxylative Claisen condensations without reducing the β-ketone in between elongation cycles. The resulting intermediate is a linear poly-β-ketone molecule that, if left on its own, can randomly cyclize in

many different ways through uncontrolled aldol condensations. The product template (PT) domain is responsible for controlling the aldol cyclizations of the intermediate. Our ability to understand the high regioselective control that PT domains exert is hindered by the inaccessibility of intrinsically unstable poly-β-ketones for in vitro studies. In chapter 2, we described the crystal structure of the PT domain cocrystallized with a stable 'atom replacement' mimetic in which isoxazole rings linked by thioethers mimic the alternating sites of carbonyls in the poly-β-ketone intermediate. The crystal structure supports the proposed deprotonation at C4 of the nascent polyketide by the catalytic His1345 and the role of a protein-coordinated water network to selectively activate the C9 carbonyl for nucleophilic addition. The importance of the 4′-phosphate at the distal end of the pantetheine arm was demonstrated to both facilitate delivery of the heptaketide mimetic deep into the PT active site and anchor one end of this linear array to precisely meter C4 into close proximity of the catalytic His1345. These findings afford the first view of a polyketide 'atom-replaced' mimetic in a NR-PKS active site that could prove general for other PKS domains.

5.1.2 Probing the Specificities of Type II Polyketide Ketoreductases Using Isosteric Polyketide Mimetic Probes (Chapter 3)

During polyketide biosynthesis, the polyketide stereocenters are introduced upon ketoreduction, which is catalyzed by the NAD(P)H dependent ketoreductase (KR) (1). In contrast to the type I KR domain, the regio- and stereospecificities of type II PKS KRs are not well understood. The actinorhodin KR (ActKR) specifically reduces the C-9 carbon of an octaketide, and current evidence suggests that KR is also involved in first ring cyclization (2-5). In addition, the ActKR has an unexplained substrate specificity towards longer substrates (pentaketide and

octaketide), but not tetraketide (6). The instability of native polyketide substrates has been the major roadblock towards understanding the molecular basis of KR specificity. In chapter 3, we described the cocrystal structure of a fully active ActKR H153Y H201G double mutant bound with our atom-replaced isosteric mimetic of the unstable natural octaketide substrate and the truncated tetraketide. The octaketide cocrystal structure affords a first glimpse into the importance of phosphopantetheine for chain-length control and substrate orientation in the type II KRs as well as gives a picture of enzyme-substrate interaction within the binding pocket.

5.1.3 Probing the Determinates of the Substrate Specificity of the Doxorubicin Ketoreductase (Chapter 4)

The type II polyketide doxorubicin is one of the most effective chemotherapeutics in our arsenal against cancer. As a type II C9 ketoreductase, the doxorubicin KR (DpsKR) is the first tailoring enzyme to act on the poly-β-ketone intermediate after it is fully extended by the minimal PKS in doxorubicin biosynthesis. Type II C9 KRs regiospecifically reduce the ninth carbon of the polyketide, counting from the thioester where it is joined to the ACP through the phosphor-pantetheinyl group. In addition, the KR plays a role in setting the cyclization of the first ring. In general, type II C9 KRs from one PKS can be replaced by type II C9 KRs from other pathways, without affecting the types of products produced by the PKS. The DpsKR, which naturally recognizes a 21-carbon long poly-β-ketone substrate, is an exception and has been shown both in vivo and in vitro to be inactive when paired with the actinorhodin minimal PKS, which produces a 16-carbon long poly-β-ketone intermediate. Until now, it was unknown whether this was a result of incompatible protein-protein interactions or a result of the DpsKR being unable to recognize the 16-carbon substrate. In chapter 4, the crystal structure of DpsKR is presented along with in

vitro reconstitution assays with a DpsKR/ActKR chimera that exhibits nearly the same activity as ActKR when reconstituted with the Act minimal PKS. These results rule out the possibility that incompatible protein-protein interactions prevent the DpsKR from reducing 16-carbon poly-β-ketone substrates produced by the Act minPKS.

5.2 Future Directions

In chapter 3, we saw that the rigidity of the isoxazole rings used in our first generation poly-β-ketone mimetics appeared to prevent the mimetic from fully entering the substrate binding pocket. In collaboration with the Vanderwal lab at UCI, we are currently developing a new generation of mimetics that alternate carbonyl groups with oxetane groups to achieve poly-β-ketone mimicry while retaining the flexibility of the natural intermediate. The application of this new generation of mimetics promises to shine a brighter light on enzyme-substrate interactions enzyme domains from non-reducing iterative PKSs, including the ActKR and DpsKR. Another question touched upon in chapter 3 was whether the closed conformation of the double mutant was a result of the mutations or a result of having the mimetic bound inside the pocket. In order to answer this question, two approaches will be taken: 1) solve the *apo* structure of the double mutant, and 2) conduct molecular dynamic studies of the ActKR WT and DM to analyze whether the conformation is somehow induced by the mutation.

In chapter 4, we saw that in the DpsKR $C222_1$ crystal structure, the magnitude of the NADPH b-factors correlated well with the outward displacement of the lid region in the absence of packing interactions. In the DpsKR $P6_122$ crystal structure, the N-terminal end of α6 (closest to the NADPH binding site) is displaced outward in the monomer where NADPH is absent. It would be enlightening to conduct molecular dynamics simulations on DpsKR to investigate if the binding

of NADPH in one monomer communicates a structural change to the neighboring monomer in the dimer, or whether there is communication from one dimer to the other dimer in the tetramer.

Lastly, the results of chapter 3 suggest that binding interactions between the ACP and KR are critical for correctly positioning the substrate for reduction; therefore, to fully understand the substrate specificity of KR, we need a more precise understanding of how the ACP binds the KR. Unlike the DH and KS domains (7, 8) , due to the nature of the KR reduction mechanism, it is difficult to design a mechanistic crosslinker that would crosslink the ACP and KR together for crystallographic studies. However, non-covalent complex crystal structure between the ACP and its binding partner have been obtained in a few cases (9-11). Complex cocrystal structures of the ACP:ActKR complex may be attainable by loading our next generation poly-β-ketone mimetic onto the ACP, incubating the loaded ACP together with KR and conducting high-throughput crystallization screening on the complex.

5.3 References

1. Tsai SC & Ames BD (2009) Chapter 2 Structural Enzymology of Polyketide Synthases. *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates,* Methods in Enzymology), pp 17-47.
2. McDaniel R, Ebert-Khosla S, Fu H, Hopwood DA, & Khosla C (1994) Engineered biosynthesis of novel polyketides: influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc Natl Acad Sci U S A* 91(24):11542-11546.
3. Crump MP, *et al.* (1997) Solution structure of the actinorhodin polyketide synthase acyl carrier protein from Streptomyces coelicolor A3(2). *Biochemistry* 36(20):6000-6008.
4. McDaniel R, Ebert-Khosla S, Hopwood DA, & Khosla C (1995) Rational design of aromatic polyketide natural products by recombinant assembly of enzymatic subunits. *Nature* 375(6532):549-554.
5. Meurer G, *et al.* (1997) Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. *Chem Biol* 4(6):433-443.
6. Kalaitzis JA & Moore BS (2004) Heterologous biosynthesis of truncated hexaketides derived from the actinorhodin polyketide synthase. *J Nat Prod* 67(8):1419-1422.
7. Worthington AS, Rivera H, Torpey JW, Alexander MD, & Burkart MD (2006) Mechanism-based protein cross-linking probes to investigate carrier protein-mediated biosynthesis. *ACS Chem Biol* 1(11):687-691.
8. Ishikawa F, Haushalter RW, & Burkart MD (2012) Dehydratase-specific probes for fatty acid and polyketide synthases. *J Am Chem Soc* 134(2):769-772.
9. Masoudi A, Raetz CR, Zhou P, & Pemble CWt (2014) Chasing acyl carrier protein through a catalytic cycle of lipid A production. *Nature* 505(7483):422-426.
10. Cryle MJ & Schlichting I (2008) Structural insights from a P450 Carrier Protein complex reveal how specificity is achieved in the P450(BioI) ACP complex. *Proc Natl Acad Sci U S A* 105(41):15696-15701.
11. Guy JE, *et al.* (2011) Remote control of regioselectivity in acyl-acyl carrier protein-desaturases. *Proc Natl Acad Sci U S A* 108(40):16594-16599.