UNIVERSITY OF CALIFORNIA,
IRVINE

Synthetic Substrate Analogues as Tools for Structural Studies of Protein-Protein and Protein-Substrate Interactions in Microbial Biosynthesis

DISSERTATION

Submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Biological Sciences

by

Jacob Cole Milligan

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

2017

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

There are many people without whom this dissertation would not be possible. I would first like to express my sincere thanks to my advisor, Professor Sheryl Tsai. I am grateful for the opportunity that she gave this starry-eyed (read: comically naïve) first year student to work in her research group, and I am grateful for the independence that she afforded me in my projects. I gained invaluable experience as a scientist and as a person while working with her. I would also like to acknowledge the impact that she had in my search for employment after graduate school, and for this, I am forever grateful. I also want to thank the members of my dissertation committee, Professor Tom Poulos and Professor Celia Goulding. Their guidance and encouragement played a huge part in my development as a graduate student.

I would also like to thank all of the current and former graduate students in the Tsai research group whose time in the lab overlapped with my own. Every one of you contributed to an exceptional work environment that helped to make the challenges of graduate school a little bit easier to bear. Dave Jackson guided me through my first year in the Tsai lab, both at the bench and away from it. I couldn't have asked for a better mentor as a young Ph.D. student, and neither the respect that I have for him nor the impact that he had on my graduate career can be overstated. Jesus Barajas had the most selfless, helpful attitude as I constantly berated him with questions about X-ray crystallography, but his willingness to always answer my questions with a smile and to help me with whatever I needed serves as the model of a co-worker that I would like to become someday. Tim Valentic seemed to always have a way of providing whatever I happened to need at any particular moment, whether it was advice on my experiments, an inspiring YouTube video when things were tough, or a cold beer when things were *really* tough. Somehow, despite having three young children during graduate school, Gabe Moreno seemed to always be the most relaxed and stress-free member of the lab, an attitude that I always appreciated when I was feeling particularly stressed out myself. Alex White, a true chemical biologist, not only taught me the ways of transesterification but was also a good friend during a time when many lab members moved on all at once and we found ourselves to be among the few who were left. Andrew Schaub joined the lab at the same time as I did, so we went through a lot together, and I appreciate that he was always right there with me in the trenches. I would also like to acknowledge Chi Nguyen, a former graduate student in the lab who left shortly after I started. Her work in the lab set the stage for my own early projects, and my results would not have been possible without her laying the groundwork. I also want to thank Gaurav Shakya, a post doc in the lab during my first year. His positive and helpful presence was greatly missed when he moved on to bigger and better things. Finally, I would like to thank Alex Chao, the honorary Tsai lab member, for always keeping things interesting in lab.

I also want to thank the four undergraduate students who worked with me in the lab: Katrina Kelly, Stephanie Tu, Caitlyn Topper, and Pilar Altman. I was truly impressed by their hard work and dedication, and I will always be grateful for the long hours that they put in. In particular, I would like to thank Stephanie Tu and Pilar Altman who worked with me for the longest amount of time. Stephanie will soon be starting medical school at Yale University, and Pilar will be soon starting a Ph.D. program at the Scripps Research Institute. I hope they never forget about their old undergraduate research mentor once they make it big!

I would also like to acknowledge our collaborators and their lab members for all the opportunities that our joint ventures afforded me: Professor Mike Burkart at the University of California, San Diego; Professor Chris Vanderwal at the University of California, Irvine; Professor Ray Luo at the University of California, Irvine; and Professor Greg Challis at the University of Warwick. It is also important to acknowledge the National Institutes of Health and the UCI Chemical and Structural Biology Training Program for funding the work presented here.

I also want to thank my parents, Greg and Keri, and my sister, Kaley, for all their love and support (both emotional and financial) over the years. I truly could not have asked for a better family, and their encouragement gave me the confidence I needed to make it in this world.

Finally, I need to thank my wife, Jessy, for too many things to list here. Thank you for sticking by me despite long days in the lab when things were going well and I didn't want to leave; for being understanding despite long nights at the pub when things weren't going well and I didn't want to leave; for putting up with me despite my frustrations and my fears; for loving me when I gave you plenty of reasons not to. I truly wouldn't be where I am today without you. Thank you.

CURRICULUM VITAE

EDUCATION

University of California, Irvine; Irvine, CA — June 2017
Ph.D., Biological Sciences

Point Loma Nazarene University; San Diego, CA — May 2013
B.S., Chemistry

RESEARCH EXPERIENCE

University of California, Irvine - Irvine, CA — January 2014 – June 2017
Graduate Student Researcher; PI - Professor Shiou-Chuan (Sheryl) Tsai, Ph.D.
Structural studies of protein-protein and protein-substrate interactions in microbial biosynthesis

University of California, Irvine - Irvine, CA — September 2013 - December 2013
Graduate Student Researcher; PI - Professor Thomas Poulos, Ph.D.
Structural basis for human cytochrome P450-3A4 inhibition

Point Loma Nazarene University - San Diego, CA — May 2011 - May 2013
Undergraduate Student Researcher; PI - Professor Dale Shellhamer, Ph.D.
Kinetic and mechanistic studies on the reaction between chlorosulfonyl isocyanate and substituted alkenes

TEACHING EXPERIENCE

University of California, Irvine - Irvine, CA — September 2014 – June 2017
Graduate Teaching Assistant
Courses: Biochemistry lab (3x), Biochemistry lecture (2x), Literature in Molecular Biology and Biochemistry lecture (2x), Molecular Biology lecture, Protein Structure and Function lecture (graduate course), Nucleic Acid Structure and Function lecture (graduate course)

Point Loma Nazarene University - San Diego, CA — August 2012 - May 2013
Undergraduate Teaching Assistant
Courses: Organic Chemistry lecture, Instrumental Analysis lecture

WORKSHOPS

Maximizing Data Quality Workshop — April 2016
Advanced Light Source, Lawrence Berkeley National Lab; Berkeley, CA

RapiData: A Practical Course in Macromolecular X-Ray Diffraction Measurement — May 2015
Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Lab; Menlo Park, CA

Teaching Assistant Professional Development Program — September 2014
University of California, Irvine

HONORS AND AWARDS

NIH Chemical and Structural Biology Training Grant, University of California, Irvine	2016 - 2017
Cellular and Molecular Biosciences Fellowship, University of California, Irvine	2013
Graduated *cum laude*, Point Loma Nazarene University	2013
Dean's List, Point Loma Nazarene University	2010- 2013
Research Associates Scholarship, Point Loma Nazarene University	2011
Science Honors Scholarship, Point Loma Nazarene University	2010 - 2013

PROFESSIONAL SOCIETY MEMBERSHIPS

American Association for the Advancement of Science	2016
American Society for Microbiology	2016
American Chemical Society	2012

ADDITIONAL ACTIVITIES

NIH Graduate Professional Success in the Biomedical Sciences Program February 2015 – June 2017
Program Trainee; Member, Careers in Industry peer group; Member, Careers in Science Policy peer group

UCI Crystallography Club June 2015 - September 2015
Club Leader

PUBLICATIONS

Milligan, J. C.;* Jackson, D. R.;* Lee, D. J.;* Schaub. A, J,; Beld. J,; Barajas, J. F.; Hale, J.; Luo, R.; Burkart, M. D.; Tsai, S. C. Structure-based engineering of type II acyl carrier protein interactions. *Manuscript in preparation*.

Milligan, J. C.;* Ellis, B. D.;* White, A. R.; Duong, V.; Altman, P. X.; Luo, R.; Vanderwal, C. D.; Tsai, S. C. Oxetane-based polyketide surrogates to probe PKS enzyme-substrate complexes. *Manuscript in preparation*.

Schaub, A. J.; Xiao, L.; **Milligan, J. C.**; Follmer, A. H.; Thai, T.; Ho, T. N.; Tsai, S. C.; Luo, R. Carrier Protein Force Field: an AMBER forcefield for phosphopantetheine-mediated biosynthesis of fatty acids and polyketides. *Manuscript in preparation*.

Jackson, D. R.; Patel, A. B.; Valentic, T. R.; Mohammed, L. Y.; Vasilakis, K.; Wattana-Amorn, P.; **Milligan, J. C.**; Crump, M. P.; Crosby, J.; Tsai, S.C. Sequence-structure-function relationship of starter unit acceptance during daunorubicin biosynthesis. *Manuscript in preparation*.

Valentic, T. R.;* Ray, L.;* Miyazawa, T.;*, Song, L.; Takahashi, S.; **Milligan, J. C.**; Osada, H.; Tsai, S. C.; Challis, G. L. A crotonyl-CoA carboxylase/reductase independent pathway for unusual extender unit biosynthesis. *Nature Communications* 7, 13609 (2016).

Shellhamer, D. F.; Alexander, K. L.; Bunting, S. A.; Elwin, S. L.; Licata, C. J.; **Milligan, J. C.**; Robinson, R. D.; Shipowick, D. E.; Smith, L. B.; Perry, M. C. Improved Synthetic Utility of a Sluggish Electrophile: Reaction of Chlorosulfonyl Isocyanate with Unreactive and Reactive Alkenes. *Synthesis* 47, 1944–1950 (2015).

Shellhamer, D. F.; Bunting, S. A.; Hickle, K. R.; Horn, P. C.; **Milligan, J. C.**; Shipowick, D. E.; Smith, L. B.; Vandenbroek, D. J.; Perry, M. C.; Boatz, J. A. Kinetic Studies on the Reaction of Chlorosulfonyl Isocyanate with Monofluoroalkenes. *Journal of Organic Chemistry* 78, 246-52 (2013).

*These authors contributed equally to this work

PRESENTATIONS AND POSTERS

"Structure-Based Engineering of Type II Acyl Carrier Protein Interactions" 2016
Gordon Research Conference (Enzymes, Coenzymes, and Metabolic Pathways); Waterville Valley, NH. Poster presentation

"Structure-Based Engineering of Type II Acyl Carrier Protein Interactions" 2016
Gordon Research Seminar (Enzymes, Coenzymes, and Metabolic Pathways); Waterville Valley, NH. Poster presentation

"Structure-Based Engineering of Type II Acyl Carrier Protein Interactions" 2016
American Society for Microbiology Microbe Meeting; Boston, MA. Poster presentation

"Structure and Dynamics of an Acyl Carrier Protein-Ketosynthase Complex." 2015
Iterative Polyketide Synthase Symposium; Banff, Alberta, Canada. Seminar presentation.

"Synthetic Utility of CSI Reactions at Lower Temperatures." 2013
245th ACS National Meeting; New Orleans, LA. Poster presentation.

"Kinetic Studies on the Reaction of Chlorosulfonyl Isocyanate with Monofluoroalkenes." 2013
West Coast Undergraduate Biological Research Conference; San Diego, CA. Seminar presentation.

ABSTRACT OF THE DISSERTATION

Synthetic Substrate Analogues as Tools for Structural Studies of Protein-Protein and Protein-Substrate Interactions in Microbial Biosynthesis

By

Jacob Cole Milligan

Doctor of Philosophy in Biological Sciences

University of California, Irvine, 2017

Professor Shiou-Chuan (Sheryl) Tsai, Chair

Fatty acid synthases (FAS) and polyketide synthases (PKS) are closely related biosynthetic enzymes that form large, multi-domain enzyme complexes. FASs are a pillar of primary metabolism across all domains of life. These enzymes are responsible for the biosynthesis of the various fatty acid chains that comprise cellular membranes. PKSs are an important component in the secondary metabolism of many organisms. These enzymes biosynthesize a wide variety of bioactive natural products with many pharmaceutical and industrial applications. Both FASs and PKSs are widely studied as platforms for bioengineering, biofuel production, and biocatalysis.

One of the primary issues in the field is the difficulty of studying the protein-protein interactions in these pathways due to the transient nature of the interactions. This knowledge gap has greatly limited the effectiveness of combinatorial biosynthesis. Another primary issue in the field is the extreme reactivity of the enzymatic substrates and biosynthetic intermediates, which makes studying the protein-substrate interactions in these pathways nearly impossible. The lack of this information has hampered rational engineering of enzymatic active sites. The goal of this work is to overcome these obstacles through the utilization of synthetic substrate analogues. These

analogues were designed to trap otherwise transient protein-protein and protein-substrate complexes for X-ray crystallographic study.

In this dissertation, the application of a mechanism-based crosslinker is described which enabled the first structure determination of an acyl carrier protein (ACP) - ketosynthase (KS) complex from *E. coli*. This crystal structure, which is the only available structural data on the ACP-KS interface, was then used to design surface mutants of the KS that altered the fatty acid profile of *E. coli* cultures. This is the first reported case of successfully leveraging protein-protein interactions to modulate the product profile of a metabolic pathway. Also in this dissertation, the use of a polyketide substrate mimic is presented which allowed for the first structure determination of an acyl-KS intermediate in complex with an extender unit analogue. This crystal structure provided the only snapshot of a KS "in action" with all substrates positioned for catalysis and enabled a mechanism for the enzyme's unique activity to be proposed.

CHAPTER 1

Introduction

1.1 Fatty acids and their biosynthesis

Fatty acid biosynthesis is one of the most ancient and ubiquitous primary metabolic pathways in nature[1]. Fatty acids serve several critical cellular functions such as cellular membrane formation, signal transduction, and energy storage[2]. Fatty acids are biosynthesized by large, multi-domain enzyme complexes called fatty acid synthases (FAS).

FASs use acetyl-coenzyme A (CoA) as the building block for fatty acid production[3]. A single acetyl-CoA is typically used as a starting unit, and additional acetyl-CoA molecules are carboxylated to provide the extender unit malonyl-CoA[4]. Fully reduced fatty acids are generated by iterative rounds of decarboxylative Claisen condensation, ketoreduction, dehydration, and enoyl reduction. The growing fatty acid chain is tethered to the phosphopantetheine (PPT) cofactor of the acyl carrier protein (ACP) throughout biosynthesis[5] (Fig 1-1).

FASs can be categorized into two primary classes termed type I and type II. A type I FAS consists of several enzymatic domains that are covalently linked together to form one large "megasynthase,"[6] whereas a type II FAS consists of several discrete, mono-functional enzymes that must find each other in solution and form complexes[7].

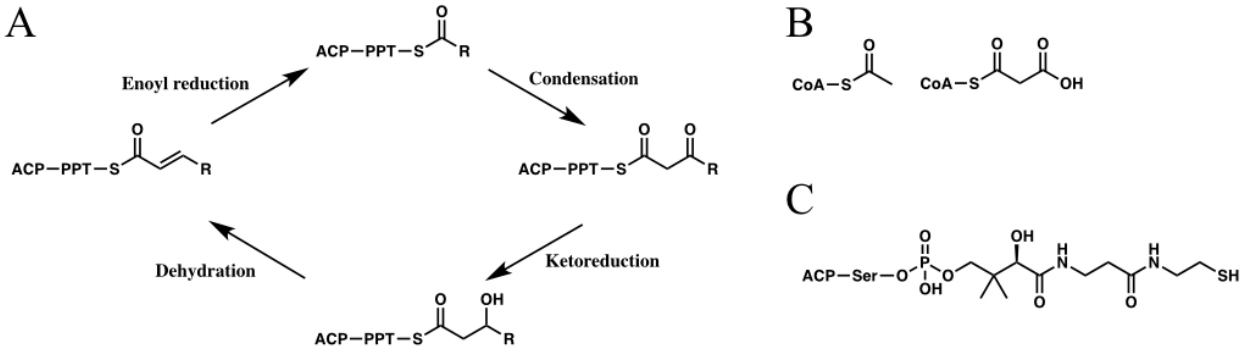

Figure 1-1. Overview of fatty acid biosynthesis. (A) Schematic of the reaction cycle. (B) Acetyl-CoA (left) and malonyl-CoA (right), typical building blocks used for fatty acid generation. (C) Phosphopantetheine prosthetic group used by the ACP during biosynthesis.

1.1.1 The structural biology of type I FASs

Eukaryotes such as yeast and mammals employ a type I FAS for the production of fatty acids. These giant, multifunctional enzymes carry out the entirety of biosynthesis within a single "megasynthase" complex[8]. Two different FAS architectures from yeast[9] and mammals[10] have been described (Fig 1-2). Though two separate complex types seem to exist in eukaryotes, they share the same types of enzymatic activity[11] and even employ the same underlying chemistry that powers the bacterial type II FAS[7]. Since the focus of this dissertation is on type II pathways, details about the activity of each catalytic domain type will be provided in the following section entitled "Type II FAS structure and function," while the current discussion will focus solely on the two architectures of the type I FASs.

The 8 domains of the yeast FAS are divided between two polypeptide chains that form a colossal 2.6 MDa $\alpha_6\beta_6$ heterododecamer whose size is on par with that of the ribosome[6]. The β-chain contains the acyl transferase (AT), enoyl reductase (ER), and dehydratase (DH) domains as well as a portion of the malonyl/palmitoyl transferase (MPT). The α-chain contains the rest of the MPT domain, the ACP, the ketoreductase (KR), the ketosynthase (KS), and the phosphopantetheine transferase (PPTase)[8].

The overall architecture of the yeast FAS resembles that of a barrel with a core comprised of 6 α-chains capped on either side with two dome-like structures comprised of 3 β-chains each (Fig 1-2A)[9]. The interior of this barrel harbors two reaction chambers that are sequestered from the cytoplasm. Interestingly, the yeast FAS devotes nearly 50% of its total sequence to an ordered, non-catalytic scaffold that facilitates proper domain arrangement[12].

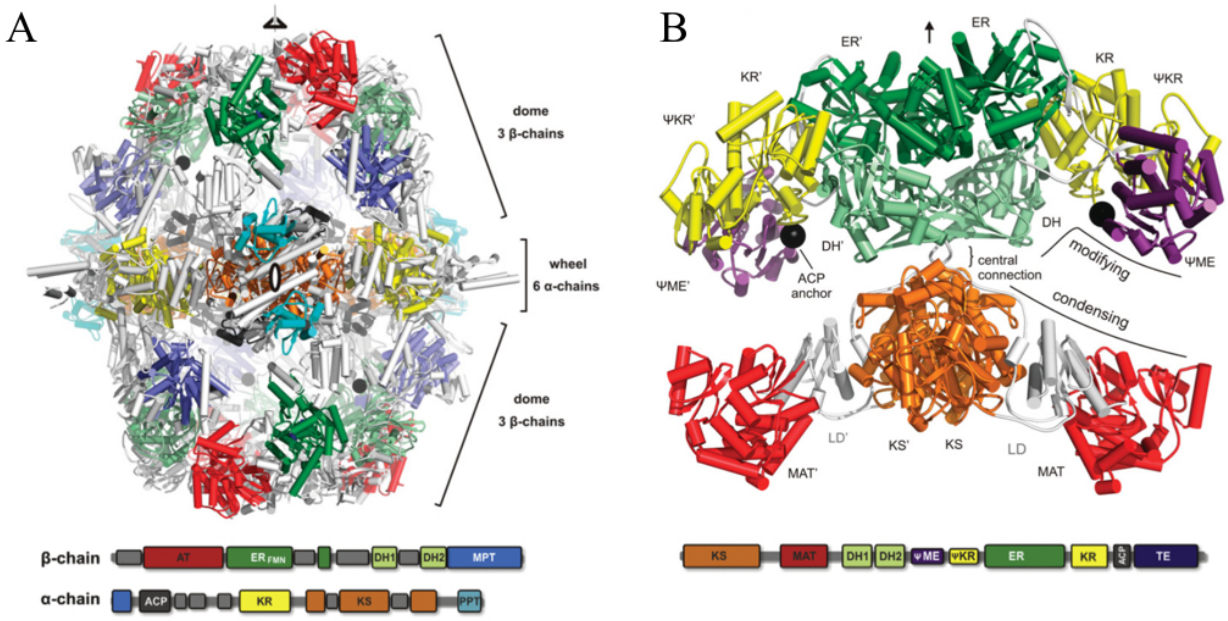

Figure 1-2. Architectures of the yeast and mammalian type I FAS. The overall structure and domain arrangement of the yeast (A) and porcine (B) FAS are shown, from reference 6. The domains in the structures match the colors of the domains in the cartoon representation below.

In contrast with the yeast FAS, all the domains of the mammalian FAS are present on a single polypeptide that forms a 540 kDa homodimer[6]. This polypeptide contains a KS domain, a bifunctional malonyl-acetyl-transferase (MAT) domain, two DH domains, the remnants of a methyl transferase (MT), a non-catalytic structural KR domain, an ER domain, a functional KR domain, the ACP, and a thioesterase (TE) that releases the final product.

The porcine FAS homodimer exhibits an "X-shaped" conformation that can be roughly divided into two halves: the lower half constitutes the condensing portion (KS and MAT) while the upper half serves as the reducing portion (KR, DH, and ER) (Fig 1-2B) (*10*). Most of the dimer interface area is at the junction between the KS domains in the bottom portion and the ER domains in the top portion[13].

There are a few differences between the yeast and mammalian FAS that are worth noting. First, in contrast with the yeast FAS, there is no major "scaffolding" present in the mammalian FAS to hold the domains in a rigid architecture; indeed, only about 9% of the total sequence is

dedicated to this purpose[13]. Electron microscopy studies have supported the hypothesis that the mammalian FAS is significantly more flexible than the yeast FAS, presumably due to the lack of a rigid scaffolding system[14]. Second, the mammalian FAS utilizes a single acyl transferase which transfers both acetyl and malonyl units to the ACP[6]. Third, the mammalian FAS has additional domains, the TE and PPTase domains, within the main FAS architecture; these domains exist *in trans* to the yeast FAS[6]. Finally, the fold and mechanism of the ER domains appears to be slightly different in both types of FAS[11,15].

It is worth noting that none of the aforementioned electron microscopy structures or X-ray crystal structure were able to visualize the highly dynamic ACP likely due to its flexibility and to the transient nature of its interactions with its various catalytic partners. There is only a single example in the literature of a type I FAS structure that captured an ACP stalled near the KS[16]. This structure did provide insight into the basis for ACP-KS recognition in the specific case of the yeast FAS, though the unique scaffolding of the yeast FAS suggests that this structure can provide no insight into ACP-KS binding in the mammalian FAS, prokaryotic type II FASs, or polyketide synthases (PKS) which are based on the architecture of these other two types[6]. Indeed, ACP-partner recognition in the yeast FAS is thought to be facilitated by the scaffolding network, which is in stark contrast to the free diffusion that governs ACP-partner interactions in all other FASs and PKSs; additionally, the ACP-KS interface in the yeast FAS is poorly conserved across other types of FASs and PKSs, suggesting a different interface in all other systems[6]. Thus, the structural basis for general ACP-partner recognition remains a major knowledge gap in the field.

1.1.2 Type II FAS structure and function

Contrary to yeast and mammals, bacteria and plants use a type II FAS to biosynthesize fatty acids[7]. Additionally, it has been shown that eukaryotic mitochondria also utilize a type II FAS[17]. The primary factor distinguishing type II from type I FASs is that the individual catalytic domains exist as discrete, mono-functional enzymes in type II systems. The domain types as well as the activity and mechanisms of the enzymes are almost identical between the two types; therefore, the primary difference is that the type II FAS enzymes are not physically tethered and must find each other in solution to form complexes. For this reason, the type I FAS is generally considered to be more efficient. However, the type I FAS is highly limited in its potential for product diversity. In contrast, the type II FAS is able to produce a wide array of different products that can be used in all aspects of cellular metabolism since the enzymes and intermediates are all diffusible intermediates that can be diverted into different biosynthetic pathways[7]. The model system for type II FAS structure and biochemistry is the *E. coli* FAS, which will be the focus of the following discussion[18].

The type II FAS contains two primary cycles termed the initiation cycle and the elongation cycle. The initiation cycle (Fig 1-3) begins with the post-translational modification of the ACP (AcpP) with a PPT prosthetic group by ACP synthase (AcpS). This active form of AcpP is central to the entire FAS pathway. It must carry the growing fatty acid chain not only to each of the catalytic domains in the pathway, but also to many other enzymes from different biosynthetic pathways. The AcpP must be able to recognize each of its catalytic partners in the proper sequence to ensure that the correct product is biosynthesized[7].

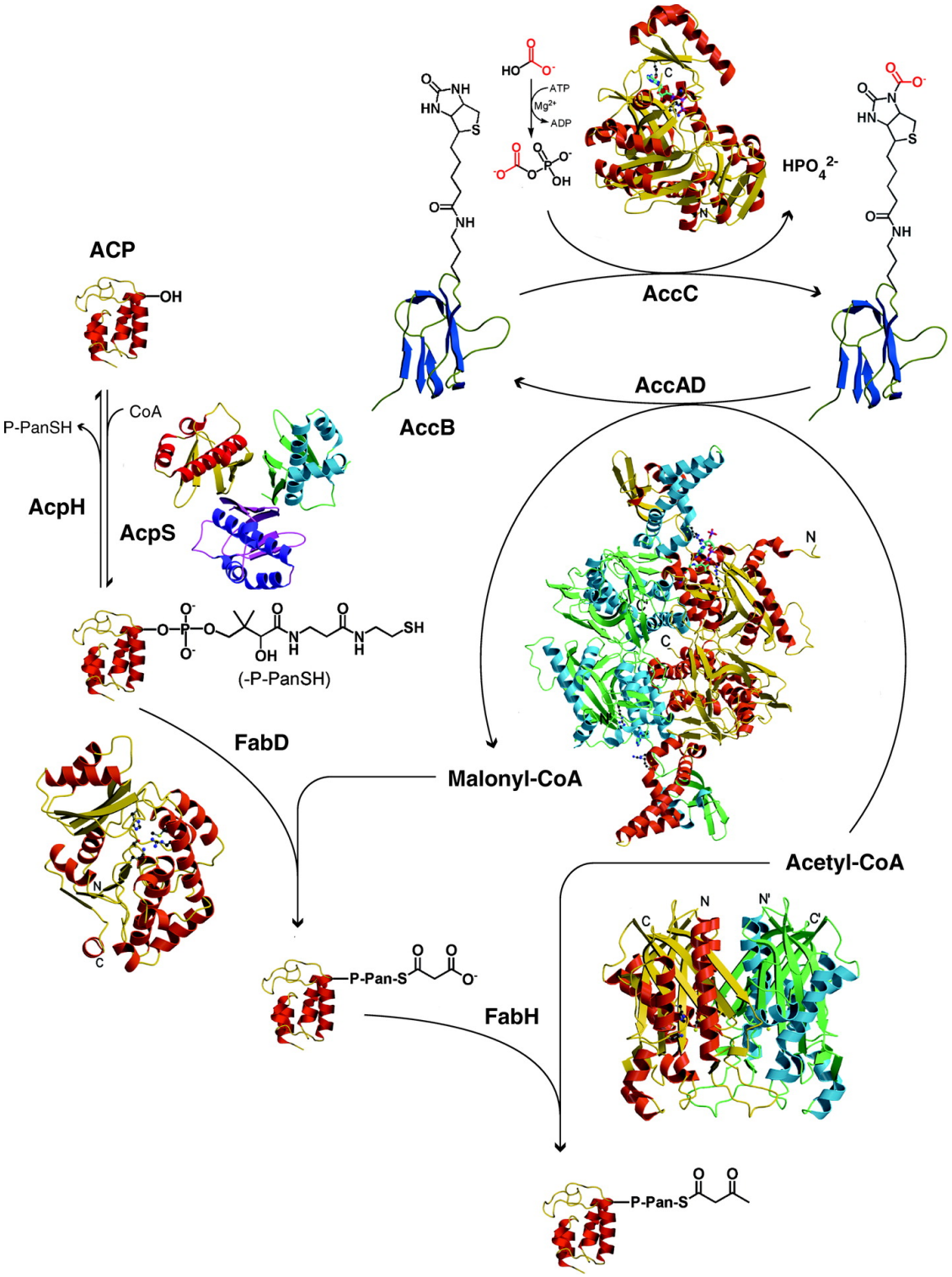

Figure 1-3. Overview of the initiation cycle of the *E. coli* type II FAS. AcpP is first post-translationally modified with a PPT prosthetic group by AcpS. Next, acetyl-CoA is carboxylated to malonyl-CoA by Acc. Then, the activated malonate is transferred to AcpP by FabD. Finally, FabH catalyzes the first round of elongation to form β-ketoacyl-AcpP. Figure from reference 7.

AcpP is a small (MW 8847 Da), acidic protein that forms a highly stable four-helix bundle[19] (Fig 1-4A). It is one of the most abundant proteins expressed in *E. coli*, typically comprising about 0.25% of the mass of all soluble proteins[20]. AcpP shares a high degree of sequence and structural similarity with other ACPs; indeed, ACPs from other bacteria, plants, and even mammals have been shown to be functional as replacements for AcpP in *E. coli*[21,22]. Thus, AcpP serves as a model system for the study of most ACPs. In particular, α-helix 2 displays a very high degree of sequence conservation of acidic residues. This helix is thought to be a "recognition helix" that AcpP uses to interact with its catalytic partners[23]. The growing fatty acid is tethered to AcpP as a thioester on the PPT prosthetic group[19] (Fig 1-4B). The nascent chain is typically sequestered within the interior cavity of AcpP[24], and upon docking of AcpP to a partner enzyme, the substrate is forced from the active site in a proposed "switchblade" mechanism[25] (Fig 1-4C).

Following activation of AcpP, the carboxylation of acetyl-CoA to malonyl-CoA—the first

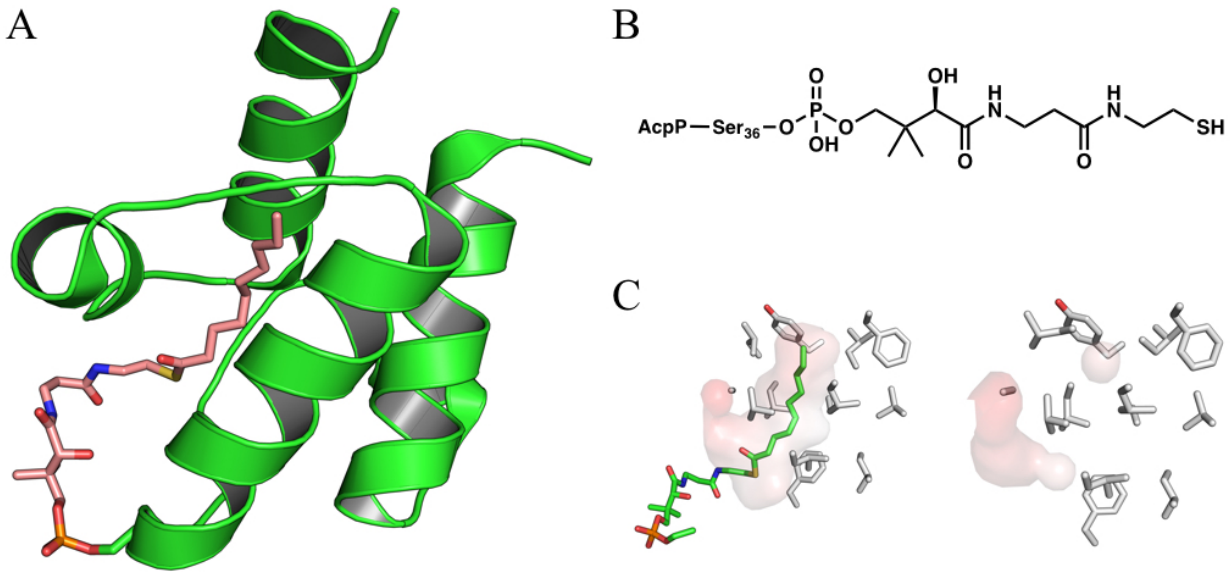

Figure 1-4 The structure of *E. coli* AcpP. (a) Overall structure of decanoyl-AcpP with the decanoyl-PPT group shown in pink sticks (PDB code: 2FAE). (b) chemical structure of the PPT group of AcpP. (c) The interior pocket of AcpP sequesters the substrate when AcpP is unbound (left), but upon binding to FabA, the pocket collapses, forcing the substrate out of the pocket and into the FabA active site (left) (PDB code: 4KEH).

committed step in fatty acid biosynthesis—is catalyzed by the biotin-dependent, multi-subunit enzyme acetyl-CoA carboxylase (ACC) (Fig 1-3)[26]. The activated malonate is then transferred to the PPT group of AcpP by the malonyl-CoA:ACP transacylase (FabD)[27]. Finally, the β-ketoacyl-ACP synthase III (KS III) FabH catalyzes the first round of elongation[28].

The structure of FabH resembles a typical thiolase fold (Fig 1-5A) with an activated cysteine nucleophile in the active site[29]. The enzyme utilizes an oxyanion hole formed by the side chains of H244 and N274 to stabilize the negative charge that forms upon enolate formation following the decarboxylation of malonyl-AcpP (Fig 1-5B). FabH is an essential enzyme in *E. coli* as it initiates the elongation cycle of fatty acid biosynthesis. It acts by first transferring an acetyl group from acetyl-CoA to the side chain of the active site cysteine. Then, upon dissociation of CoA, FabH binds to the malonyl-ACP generated by FabD and facilitates the decarboxylative Claisen condensation between these two acyl groups to release CO_2 and a β-ketoacyl-AcpP (Fig 1-5C).

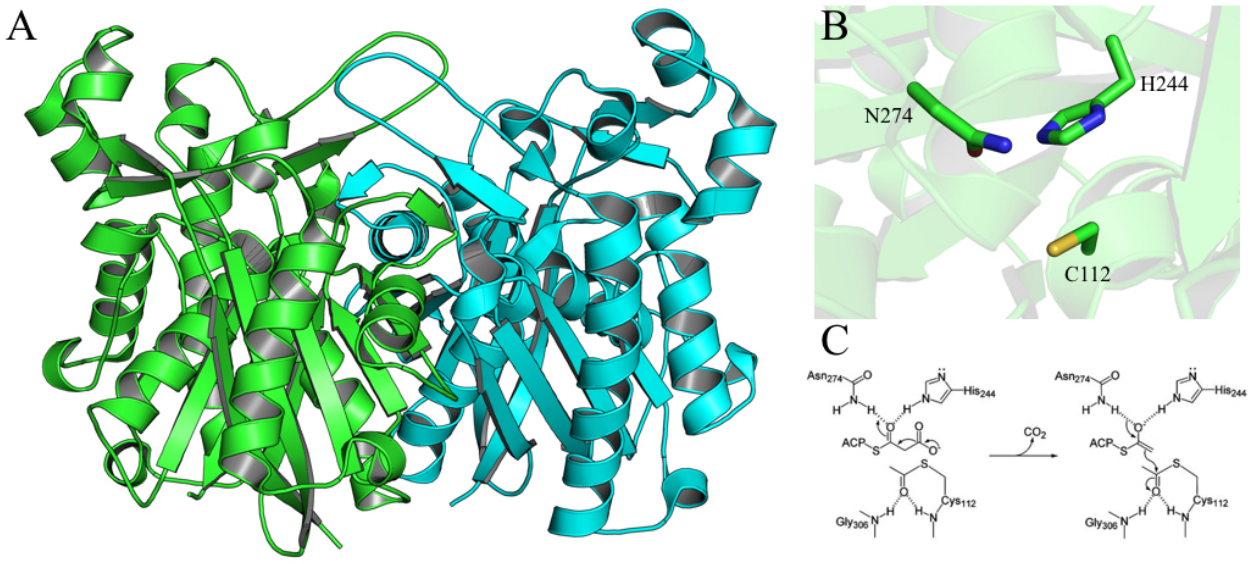

Figure 1-5 The structure of *E. coli* FabH. (a) Overall structure of FabH (PDB code: 1EBL). (b) FabH catalytic residues are shown in green sticks. C112 is the site of acylation, and H244/N274 form the oxyanion hole. (c) The mechanism of FabH-catalyzed decarboxylative Claisen condensation. Mechanism from reference 7.

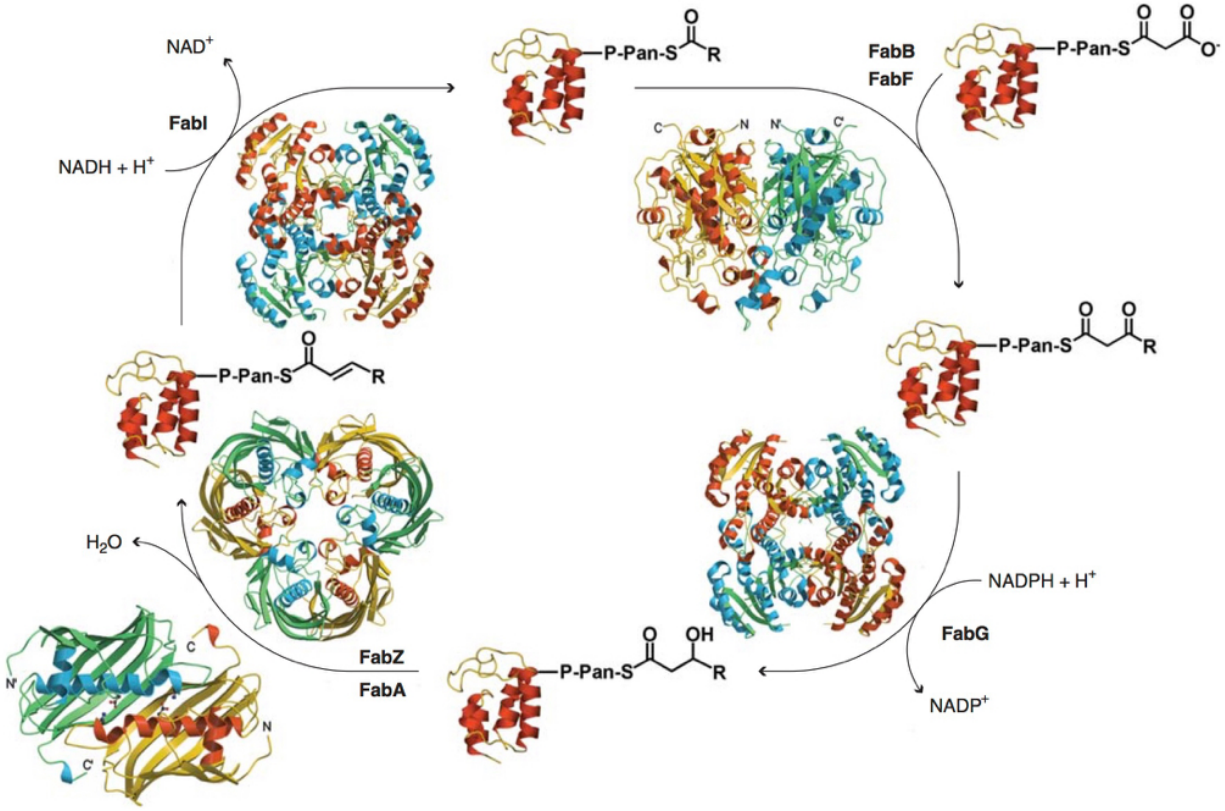

Figure 1-6 Overview of the elongation cycle of the *E. coli* type II FAS. Once acetoacetyl-AcpP enters the cycle, it undergoes iterative rounds of reduction and elongation. First, the β-ketoacyl-ACP reductase (KR) FabG reduces the di-keto intermediate using NADPH. In the next step, one of the two β-hydroxyacyl-ACP dehydratases (DH), either FabA or FabZ, facilitates dehydration. Finally, the enoyl-ACP reductase (ER) FabI catalyzes the final reduction using NADH to generate a fully reduced chain. From there, the cycle resumes with the action of the β-ketoacyl-ACP synthase I and II enzymes (FabB and FabF, respectively). Figure from reference 7.

The elongation cycle begins with the reduction of acetoacetyl-AcpP via the β-ketoacyl-ACP reductase (KR) enzyme FabG. Only one isozyme of FabG has been identified, and its status as an essential protein suggests that it may be a potential target for antibiotics[30]. This enzyme is a member of the short-chain dehydrogenase/reductase (SDR) superfamily, and it adopts a tetrameric α/β structure with a classic Rossman fold for binding its NADPH cofactor[31] (Fig 1-7A). Like other SDR enzymes, FabG contains active site tyrosine and lysine residues that

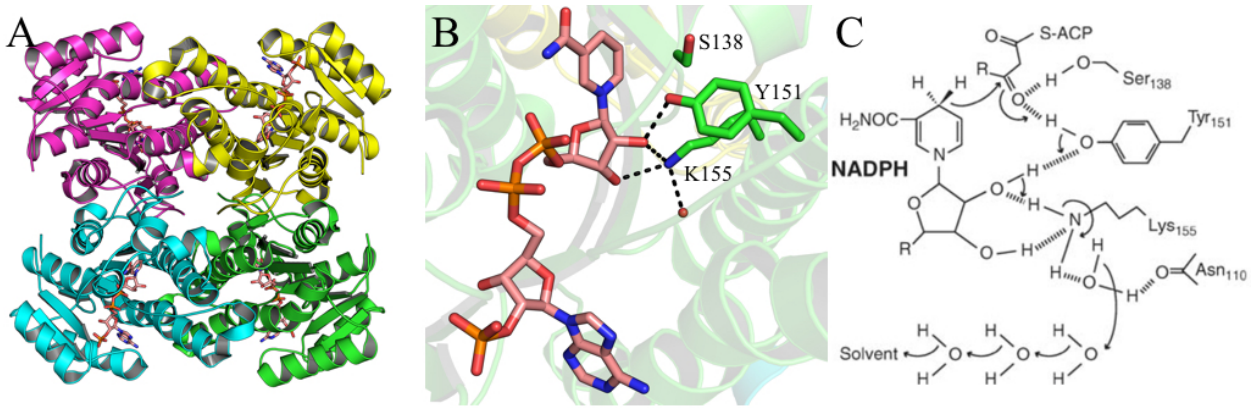

Figure 1-7 The structure of *E. coli* FabG. (a) Overall structure of FabG with the NADPH cofactor shown in pink sticks (PDB code: 1Q7B). (b) The active site residues of FabG are shown in green sticks. (c) The catalytic mechanism of FabG. Mechanism from reference 7.

participate in catalysis (Fig 1-7B). These side chains form an extensive hydrogen bond network leading to the solvent that facilitates the proton transfer (Fig 1-7C).

Following ketoreduction, the β-hydroxyacyl-ACP intermediate undergoes dehydration by one of the β-hydroxyacyl-ACP dehydratases[32] (DH). There are two enzymes that catalyze this reaction in *E. coli*: FabA and FabZ. The primary difference between the two isozymes is that FabA can also isomerize the *trans*-2-enoyl-ACP product to *cis*-3-enoyl-ACP; FabA has a particularly high affinity for a 10-carbon substrate with respect to this second reaction, likely due to the size of its substrate binding tunnel[33]. This reaction is critical for the formation of unsaturated fatty acids. Both FabA and FabZ adopt a β + α "hot dog" fold, and the overall structures of the two enzymes are essentially idential[34,35] (Fig 1-8A). The catalytic residues in FabA are H70, which serves as an active site based, and D84, which serves as an acid (Fig 1-8B and C). In the case of FabA, these residues also play a role in the isomerization reaction. While the lack of isomerase activity in FabZ is not fully understood, it is hypothesized that the subtle difference in active site structure does not allow FabZ to bind the *cis*-3-enoyl-ACP product[35].

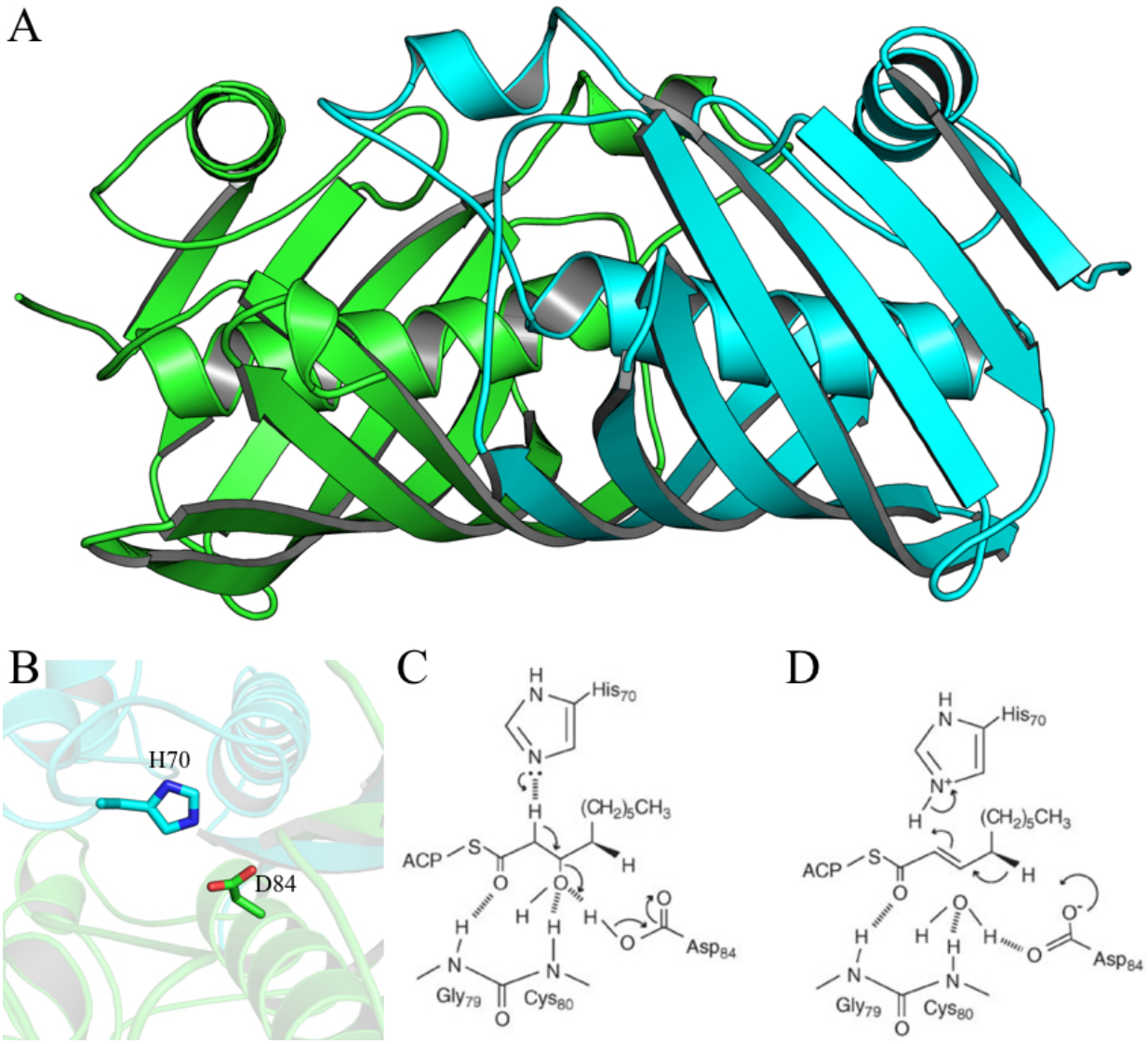

Figure 1-8 The structure of *E. coli* FabA. (a) Overall structure of FabA (PDB code: 1MKB). (b) The active site residues of FabA are shown in sticks. (c) The catalytic mechanism of FabA-catalyzed dehydration. (d) The catalytic mechanism of FabA-catalyzed isomerization. Mechanisms from reference 7.

The last step of the elongation cycle is the final reduction of the unsaturated acyl-ACP to a fully reduced product. This reaction is catalyzed by FabI, the enoyl-ACP reductase (ER). FabI plays a role in regulating fatty acid elongation as it facilitates the final step before further condensations can be carried out by the KSs[36]. It is the target for multiple clinical antibiotics, including isoniazid[37] and triclosan[38]. Like FabG, FabI is a member of the SDR superfamily,

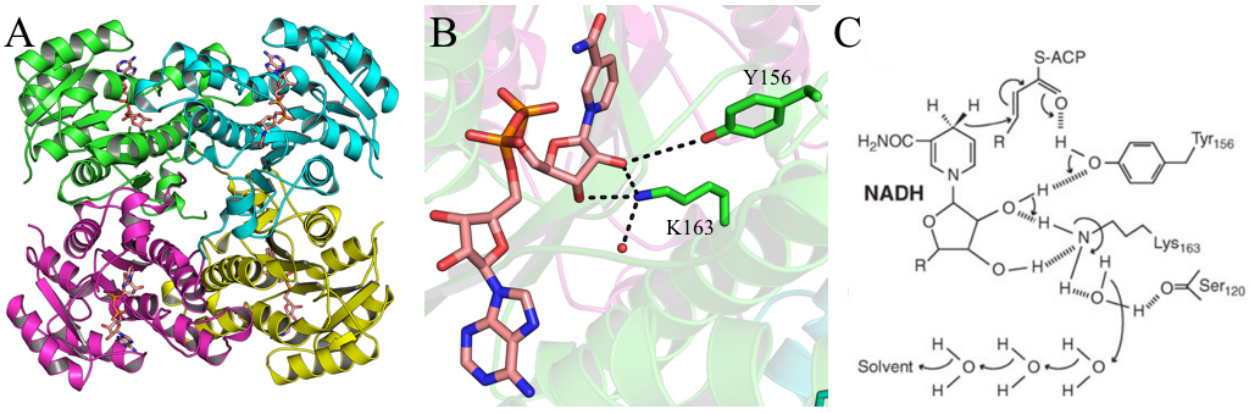

Figure 1-9 The structure of *E. coli* FabI. (a) The overall structure of FabI with the NADH cofactor shown in pink sticks (PDB code: 1DFI). (b) The active site residues of FabI are shown in green sticks. (c) The catalytic mechanism of FabI. Mechanism from reference 7.

though it utilizes an NADH cofactor as opposed to NADPH used by FabG. The structure of FabI shares many similarities with that of FabG[39] (Fig 1-9A), including active site tyrosine and lysine residues which form an extensive hydrogen bond network to facilitate proton transfer[40] (Fig 1-9B and C).

The elongation cycle begins again with the KS I and KS II enzymes, FabB and FabF, respectively. These enzymes both have thiolase folds, and they are structurally quite similar[41,42]. In addition, they both perform the same function, namely chain elongation. The primary difference between FabB and FabF is that FabB will efficiently elongate a *cis*-3-decanoyl-ACP substrate, whereas FabF will not; thus, FabB is critical for the biosynthesis of unsaturated fatty acids[43]. Also, the chain length preference is thought to be slightly different between FabB and FabF. FabB has higher activity when acting on shorter chain lengths, while the opposite is true for FabF. The decarboxylative Claisen condensation reaction catalyzed by these two enzymes is similar to the FabH reaction, though there are key differences. First, since FabH is primed with acetyl-CoA, it can only catalyze the first round of elongation. In contrast, the active site cysteines of FabB/F are primed with an acyl-ACP ready for subsequent rounds of elongation. Second, FabH uses the side

chains of one histidine and one asparagine residue within the active site to facilitate decarboxylation[29], whereas FabB/F both use two histidine side chains to fulfill this purpose[41,42].

The FabB/F reaction begins with the binding of an acyl-ACP followed by transfer of the acyl group onto the side chain of C163 in the active site. This sulfhydryl group is thought to be activated by the helix dipole effect generated at the N-terminus of a long α-helix[44]. Mutagenesis work suggests that neither of the histidine residues that are known to be important for catalysis affect the self-acylation reaction[7]. Certain details of the decarboxylative reaction remain unclear. One mechanism, based on crystal structures of FabB in complex with two covalent inhibitors cerulenin and thiolactomycin[45], predicted that both H298 and H333 (H303 and H337 in FabF) serve as hydrogen bond donors to stabilize the negative charge that builds up on the C1 carbonyl oxygen of the malonate extender unit. However, there are data to suggest that bicarbonate is released during the reaction, not CO_2[46]. This suggests that only one of the histidine side chains is serving as a hydrogen bond donor while the other serves to activate a water molecule for

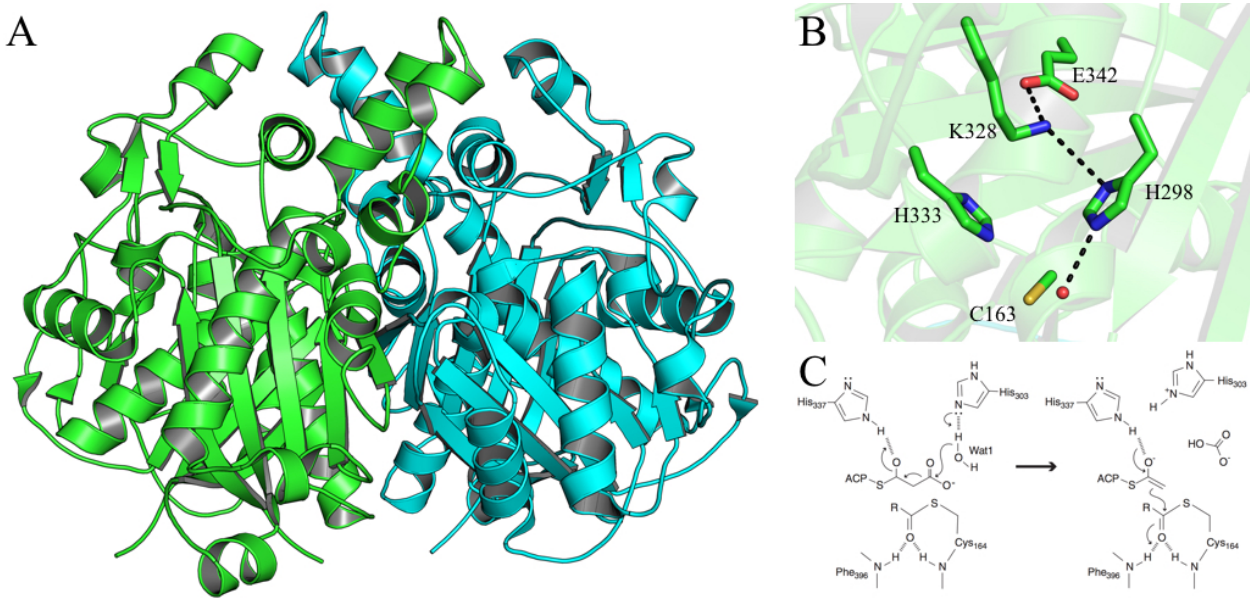

Figure 1-10 The structure of *E. coli* FabB. (a) The overall structure of FabB (PDB code: 1DD8). (b) The FabB catalytic residues are shown in green sticks. (c) The proposed mechanism of FabB decarboxylative Claisen condensation. Mechanism from reference 7.

nucleophilic attack at the C3 carbonyl carbon (Fig 1-10C). This mechanism is also consistent with the structural data. The Nδ nitrogen of H298 is quite close to the Nζ of K328 which is adjacent to the carboxylate of E342; this network could serve to alter the electronic state of the ring such that the histidine could abstract a proton from water and facilitate the release of bicarbonate. Additionally, ordered waters near H298 in the crystal structure also support this mechanism.

1.2 Polyketide natural products and their biosynthesis

Polyketide natural products are a large and diverse class of bioactive secondary metabolites produced in bacteria, fungi, and plants. A massive number of polyketides have been identified in nature with a myriad of chemical structures and biological functions, likely with many more yet to be discovered. Many of these compounds are of industrial interest. There are dozens of polyketides used as drugs, and their annual sales exceed $15 billion[47]. Notable examples include

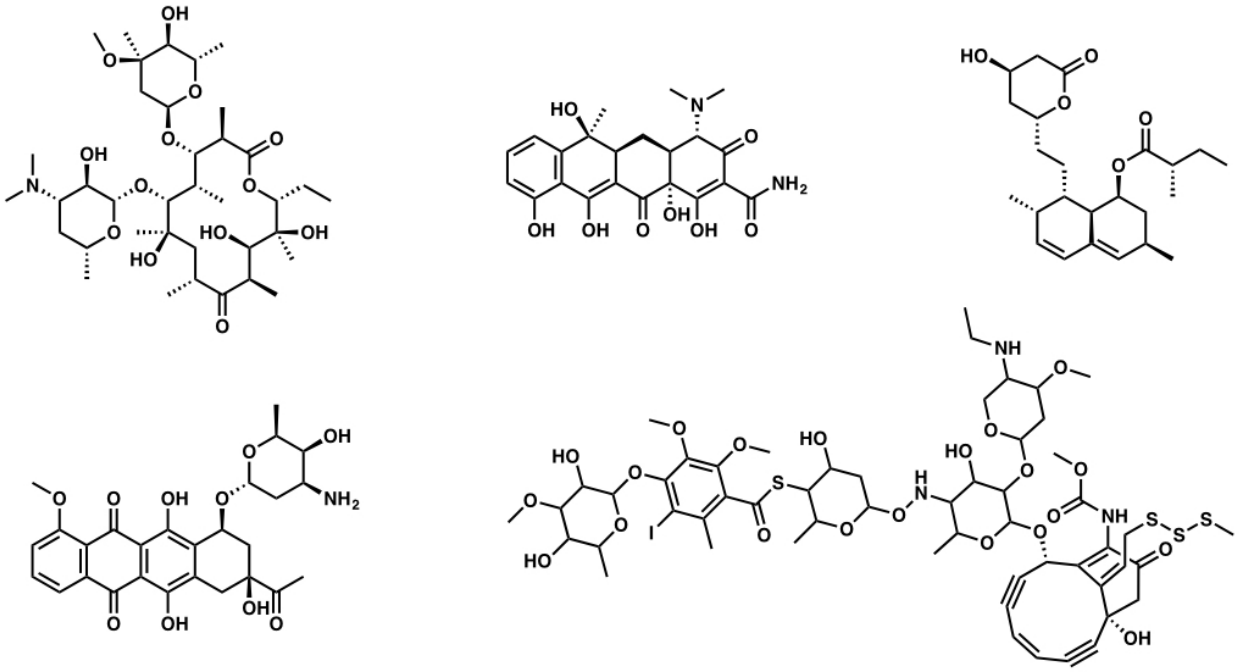

Figure 1-11 Notable examples of polyketides used as pharmaceuticals. Top left: erythromycin (antibiotic). Top center: tetracycline (antibiotic). Top right: lovastatin (cholesterol-lowering). Bottom left: daunorubicin (anti-cancer). Bottom right: calicheamicin (anti-cancer).

the antibiotics erythromycin and tetracycline, the cholesterol-lowering drugs lovastatin and simvastatin, and the anti-cancer drugs daunorubicin and calicheamicin (Fig 1-11). In addition, some polyketides are agricultural toxins. For example, the polyketide aflatoxin is a carcinogen that is linked to an increased risk of liver cancer in farmers and other agricultural workers exposed to the fungal aflatoxin producers *Aspergillus flavus* and *Aspergillus parasiticus*[48].

Polyketides are biosynthesized by large, multi-domain enzyme complexes termed the polyketide synthase (PKS). PKSs are closely related to FASs, and they use many of the same domain types, catalytic mechanisms, substrates, and biosynthetic strategies; thus, the following discussion will focus only on the structures of PKSs and the ways in which they diverge from their FAS counterparts. Like FASs, PKSs operate through iterative rounds of decarboxylative Claisen condensations to construct a carbon backbone; however, unlike in FASs, the reduction steps are "optional" in the case of PKSs. In some cases, almost no reduction occurs, and the result is the formation of a long poly-β-ketone chain. Also like FASs, PKSs can be divided into two primary types: type I and type II.

1.2.1 The structural biology of type I PKSs

Much like their evolutionary ancestor the FAS, type I PKSs are characterized by a small number of multi-domain enzymes that form very large complexes. There are two primary sub-types of type I PKSs: iterative and modular. Iterative type I PKSs (iPKS), found primarily in fungi, typically exist as a single polypeptide that uses a single set of domains in an iterative fashion to afford the full-length, final product[49] (Fig 1-12A). This biosynthetic strategy is most closely related to the mammalian FAS discussed previously. While there is a core set of obligatory domain types

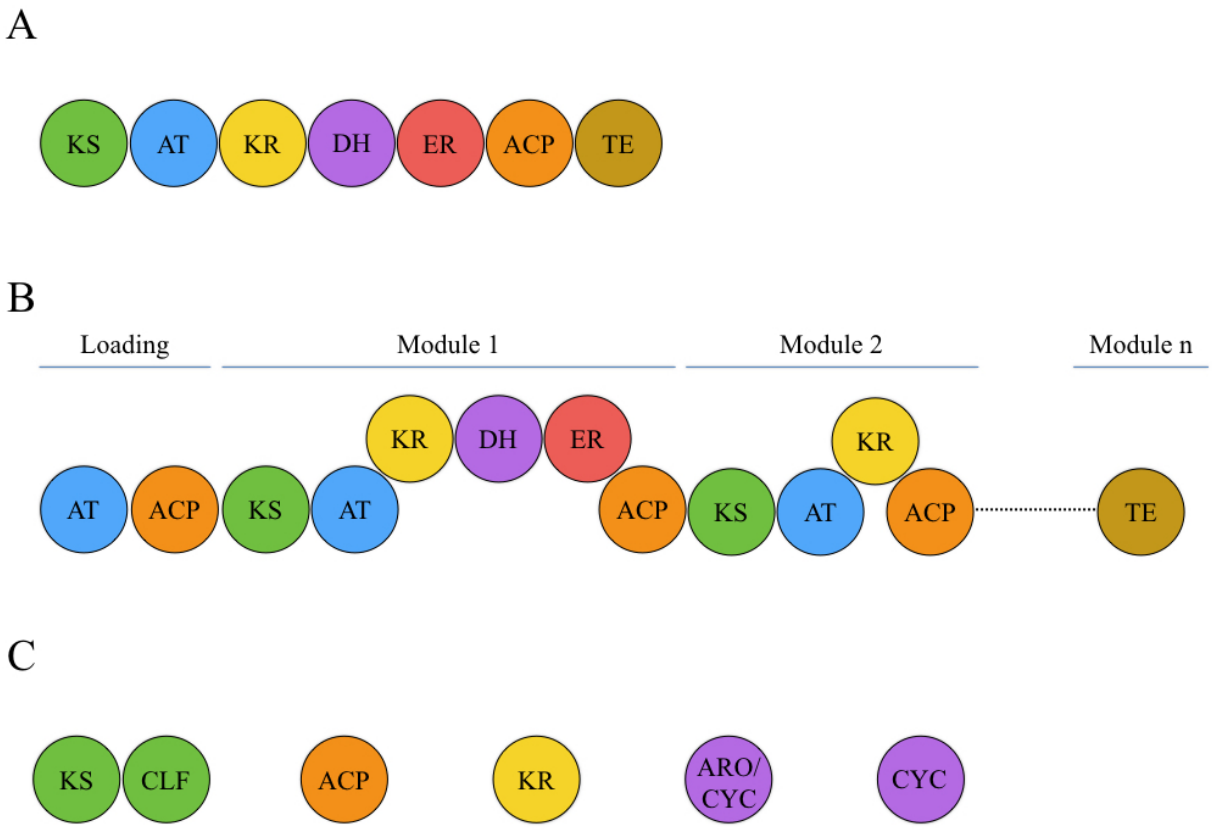

Figure 1-12 Overview of different PKS types. (a) A schematic of an iterative reducing type I PKS. (b) schematic of a *cis*-AT modular type I PKS. (c) A schematic of an iterative reducing type II PKS.

in every iPKS termed the "minimal PKS" (KS, AT, ACP), the presence of other domain types can vary across different biosynthetic pathways[50]. For example, an iPKS can be considered reducing or non-reducing depending on the presence or absence of KR, DH, and ER domains. Typically, the final product is released via a TE domain or a reductase (R) domain[51].

While structural characterization of full-length iPKSs has remained largely elusive likely due to their large size and inherent flexibility, a crystal structure of mycocerosic acid synthase (MAS), a reducing iPKS from *Mycobacterium smegmatis*, was recently published[52] (Fig 1-13). It bears a striking resemblance to the mammalian fatty acid synthase structure (Fig 1-2B), which is to be expected given its similar biosynthetic approach.

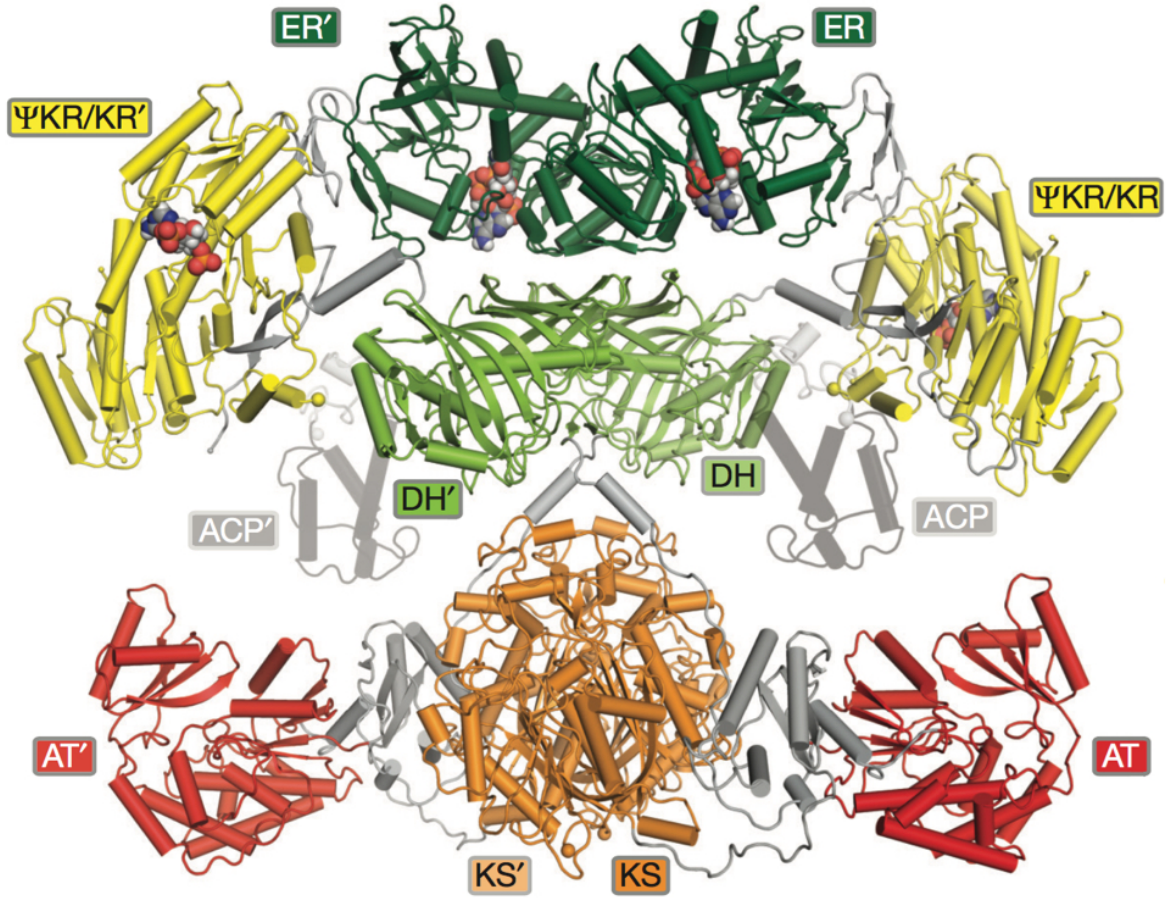

Figure 1-13 Crystal structure of *M. smegmatis* mycocerosic acid synthase (MAS). Figure from reference 52.

In contrast to iPKSs, modular type I PKSs (modPKS) work in an assembly line-like fashion in which the growing chain is passed through several multi-domain modules which each act on the nascent polyketide in order[53] (Fig 1-12B). The types of domains present in each module determine the chemical modifications that occur within. This domain-swapping utilized by nature allows for great diversity in the structure and function of the final natural products.

Recently, the first structure of a full length PKS module was solved using cryo-electron microscopy[54] (Fig 1-14). In contrast with the x-shape of the mammalian FAS and the iPKS, this modPKS structure displays more of an arch-like shape with a single reaction chamber within which the ACP has access to all of the catalytic active sites. Additionally, an outside entrance to this

internal reaction chamber was observed that presumably allows for an ACP from an upstream module to deliver the substrate to the subsequent module for further processing.

In contrast with these "*cis*-AT" PKSs, *trans*-acyltransferase (*trans*-AT) PKSs, a sub-group of modular PKSs that has recently emerged as a unique class unto itself[55], utilizes an assembly line system but also requires the participation of one or more AT enzymes that act *in trans* to the PKS itself[56]. *trans*-AT PKSs have been heavily studied as of late given their potential for unique extender unit incorporation via the standalone AT domains required for biosynthesis[57-59].

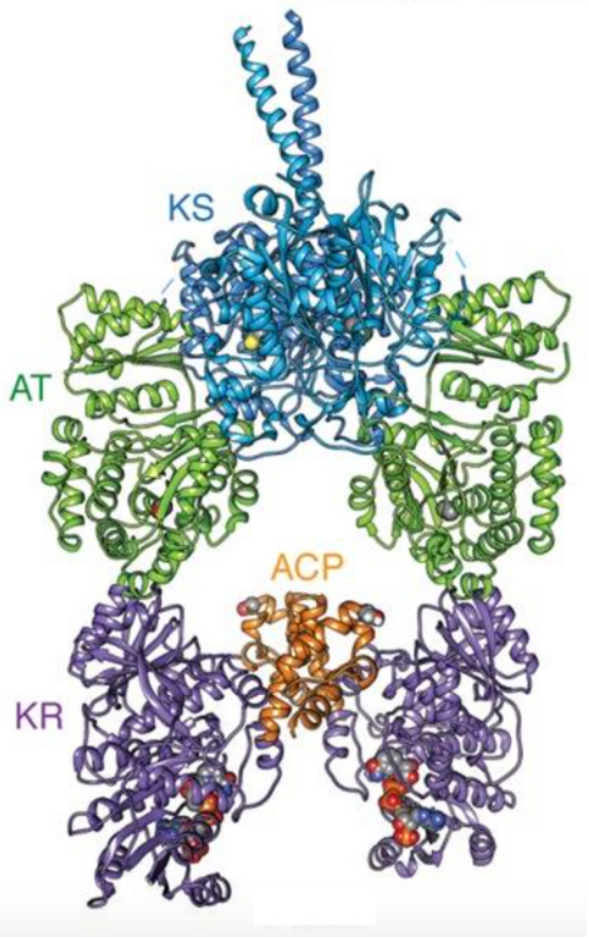

Figure 1-14 The cryo-EM structure of a full PKS module. This structure is of the third module in the pikromycin PKS. Figure from reference 54.

One trend that has emerged from the study of *trans*-AT PKSs is the presence of non-elongating modules, or modules within which no decarboxylative Claisen condensation occurs via the KS[60]. Instead, these modules often contain KS domains termed KS^0 that lack a conserved histidine required for the condensation reaction. These KS^0 domains typically facilitate the transfer of acyl groups from one module to the next. Bioinformatics analysis indicates that KS^0 domains are much more likely to be found in PKS-nonribosomal peptide synthetase (NRPS) hybrid systems, particularly at the interface site of the PKS and NRPS machinery[61].

1.2.2 Type II PKS enzymes

Much like the type II FAS in *E. coli*, type II PKSs are characterized by a set of discrete, mono-functional enzymes that must find each other in solution and form complexes. They are found primarily in bacteria of the genus *Streptomyces*, and the products are often polycyclic aromatic compounds[62]. These pathways can be thought of as a type II variant of an iPKS in that the minimal PKS enzymes act in an iterative fashion to produce a poly-β-ketone chain. Also like an iPKS, type II PKSs can be either reducing or non-reducing.

The enzymatic processes in type II polyketide biosynthesis are quite similar to those of type II fatty acid biosynthesis (Fig 1-15), though there are some key differences. First, the enzyme responsible for chain elongation in a type II PKS is a KS/chain length factor (CLF) heterodimer as opposed to the KS homodimer found in type II FASs[63]. This CLF is closely related to the KS in sequence and structure, though it is missing key catalytic residues. It is unique to type II PKSs, and its internal cavity determines the length of the growing poly-β-ketone[64]. Another key difference is the starter units employed by the two pathways. Type II FASs utilize a KS III such as FabH to select for an acyl-CoA starter unit, whereas only some type II PKSs utilize a KS III. Many others obtain an acetyl starter unit via the decarboxylation of malonyl-ACP within the KS/CLF active site[65]. Additionally, as in type I pathways, the reduction steps in type II PKSs are optional. The KR domains, if any are found in a given pathway, typically carry out only a single, regiospecific reduction[65]. DH and ER domains are typically not found in type II PKSs. One final feature unique to type II PKSs is the presence of aromatase/cyclase (ARO/CYC) enzymes that generate the final carbon backbone structure of the product through one or more cyclization reactions and aromatization[66].

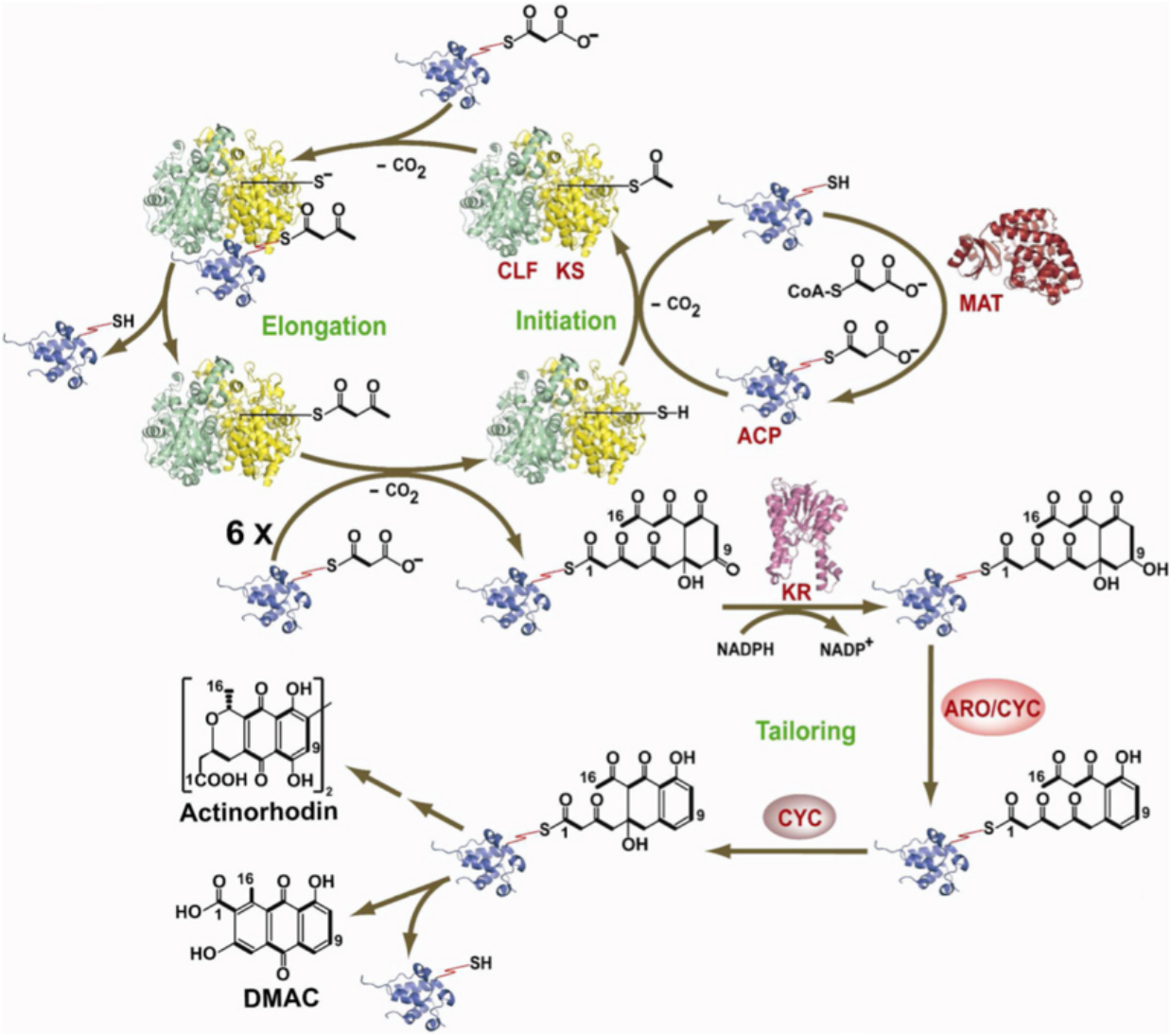

Figure 1-15 Overview of type II polyketide biosynthesis using the actinorhodin pathway as a model. The pathway begins with the priming of the KS/CLF heterodimer. A malonate from malonyl-CoA is transferred onto the PPT cofactor of ACP by the malonyl-CoA:ACP transacylase (MAT). This activated malonate is then transferred to the active site cysteine of the KS domain where it is decarboxylated to an acetyl moiety. Then, upon binding of a second malonyl-ACP to the KS/CLF, a decarboxylative Claisen condensation reaction is catalyzed within the KS active site to afford an acetoacetyl-ACP species. This group is transferred onto the KS active site cysteine once again, and this process is iterated until the full length product is generated. The final cyclization and tailoring steps can vary substantially between pathways, but in this case, the intermediate is first cyclized spontaneously within the KS/CLF, followed by reduction by a KR, aromatization and cyclization via an aromatase/cyclase (ARO/CYC), cyclization catalyzed by a cyclase (CYC), and further tailoring. Figure from reference 65.

1.3 Issues in the field and the use of synthetic substrate analogues

There are two primary issues in the study of FAS and PKS structural biology that have hindered understanding of these complex enzyme systems. First is the difficulty of studying the protein-protein interactions between the ACP and its partner enzymes[67]. The ACP sits at the center of fatty acid and polyketide biosynthesis, and it must interact with several different partners in the proper sequence in order to ensure that the correct product is generated. However, due to the transient nature of these interactions, capturing these ACP-partner complexes for crystallographic study has proven quite challenging. This lack of structural knowledge has largely precluded the successful use of combinatorial biosynthesis as envisioned by the pioneers of the PKS field[68].

One approach to solving multi-domain structures that has seen some success in recent years is the use of mechanism-based protein crosslinkers[25]. These crosslinkers, developed primarily in the lab of Michael Burkart at the University of California, San Diego, typically consist of a modified PPT cofactor that contains a reactive "warhead" specific to the unique active site chemistry of a particular FAS or PKS domain type[69,70]. This probe can be loaded onto *apo* (no PPT cofactor) ACP using an *in vitro* chemoenzymatic reaction to generate so-called *crypto* ACP. Since the crosslink will only take place once the ACP delivers the PPT mimic to the active site of the target domain, incubating *crypto* ACP with the target results in a covalent linkage that arises from the functional association of the ACP and its partner.

This crosslinking strategy enabled the first high resolution structure determination of a dehydratase domain in complex with an ACP, specifically the complex of FabA and AcpP from the *E. coli* type II FAS[25] (Fig 1-16). The crosslinker utilized a sulfonyl 3-alkyne analogue of the well-known DH suicide inhibitor 3-decynoyl-*N*-acetylcysteamine (Fig 1-16C). By loading this

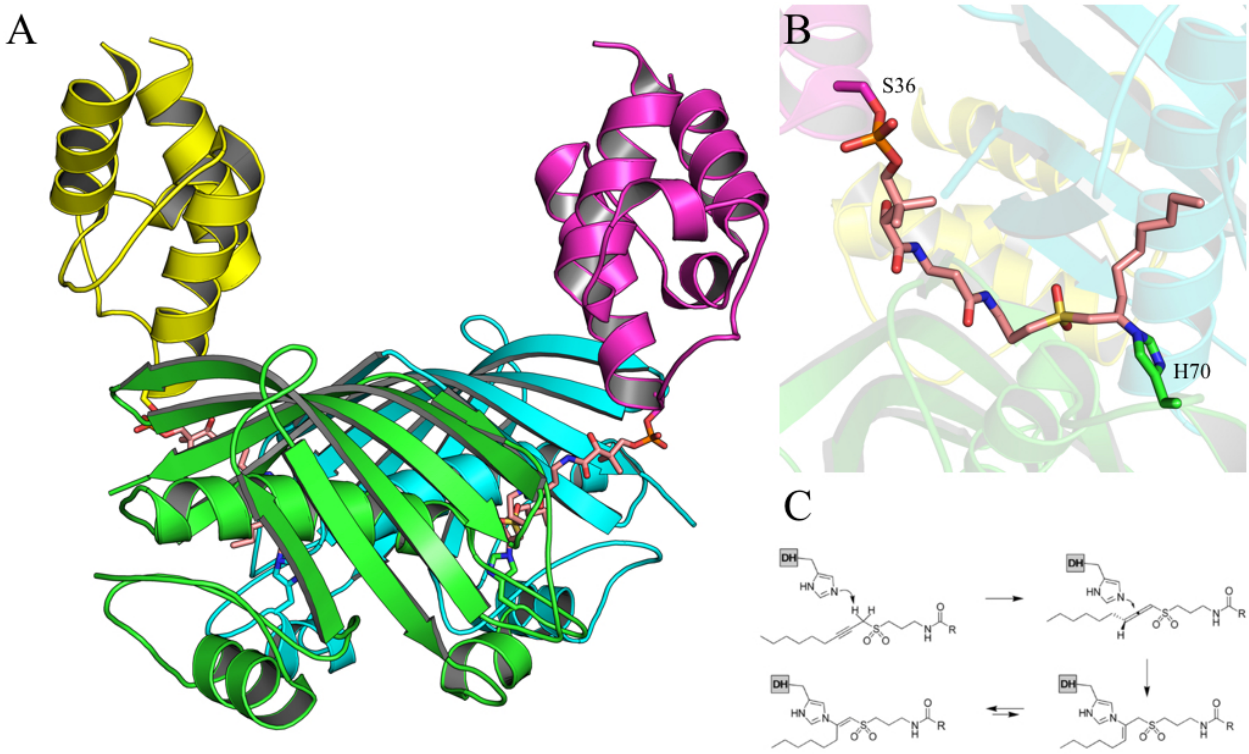

Figure 1-16 The crystal structure of the crosslinked complex between FabA and AcpP. (a) overall structure with the FabA monomers shown in green and cyan, the AcpPs shown in yellow and magenta, and the crosslinker shown in pink sticks (PDB code: 4KEH). (b) A close-up view of the FabA active site showing the covalent linkage between the crosslinker, which is covalently attached to S36 of AcpP, and the sidechain of FabA H70. (c) The mechanism-based crosslinking reaction. Mechanism from reference 25.

functionality onto the PPT cofactor of AcpP, a crosslinked complex was obtained and used for crystallographic study (Fig 1-16A). This work provided new insight not only into the protein-protein interactions between these critical players in fatty acid biosynthesis but also into AcpP's "switchblade" mechanism for substrate delivery. This strategy does have one disadvantage, however; since the crosslinkers are specific for certain domain types, this requires that a different crosslinker be developed for every FAS and PKS enzyme type. Currently, the only mechanism-based crosslinkers in the literature are specific for DH domains[69] and KS domains[70].

Another crosslinked complex structure between an AT domain (VinK) and an ACP (VinL) from the vicenistatin PKS was also determined recently[71] (Fig 1-17), though it did not use a

mechanism-based crosslinker. In this work, a non-catalytic serine residue within the active site pocket of the AT was mutated to a cysteine. This allowed for the use of a 1,2-bismeleimidoethane (BMOE) general thiol-to-thiol crosslinker which formed a covalent linkage between the non-native cysteine side chain within the active site and the terminal sulfhydryl of the ACP's PPT prosthetic group. While this strategy may be the best available option for difficult to target active sites, it is unclear how biologically relevant the resulting complex would be.

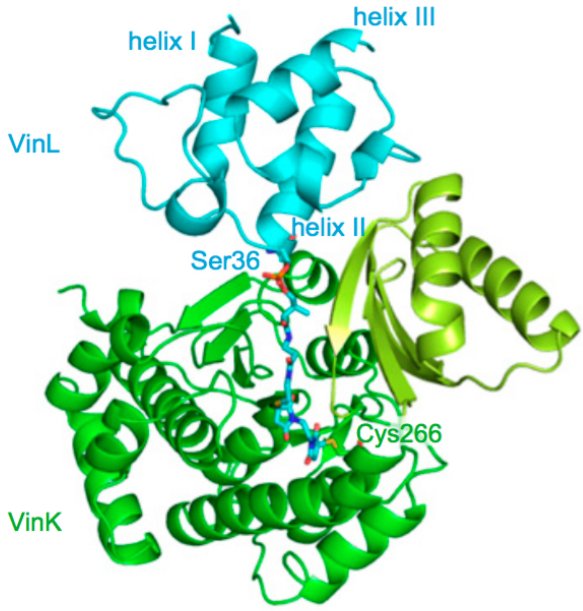

Figure 1-17 The crystal structure of the crosslinked complex between VinK and VinL. VinK is shown in green, VinL is shown in cyan, and the BMOE crosslinker is shown in cyan sticks. PDB code: 5CZD. Figure from reference 71.

In addition to protein-protein interactions, a major difficulty in the field of PKS structural biology is the difficulty of studying protein-substrate interactions. This problem is particularly difficult in the case of non-reducing PKSs such as type I non-reducing iterative PKSs and type II PKSs. The long poly-β-ketone chains that are typical intermediates in these pathways are highly susceptible to spontaneous, non-specific cyclization[72] (Fig 1-18). It is nearly impossible to trap these intermediates for crystallographic study, and, as a result, little is known about how the various PKS enzymes are able to bind to and orient their substrates for highly regiospecific and stereospecific catalysis. Without this structural knowledge, attempts to rationally engineer the specificity of PKS enzyme active sites has largely resulted in inactive proteins.

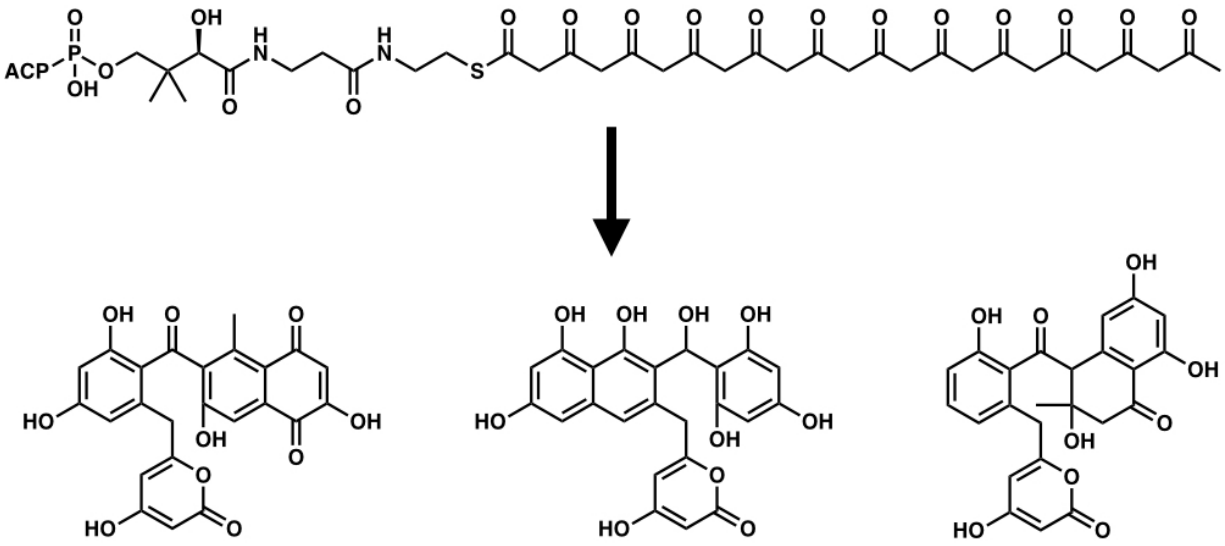

Figure 1-18 Highly reactive PKS intermediates are prone to spontaneous and non-specific intramolecular aldol condensations. This makes their characterization a unique challenge.

One approach that aims to solve this issue is the use of synthetic, inert substrate analogues. The class of probes developed by the Tsai and Burkart labs are constructed using a so-called "atom replacement" strategy (Fig 1-19) in which reactive ketide units are replaced with more stable functional groups[73]. In this case, isoxazole rings were used as stable replacements for diketide moieties. These probes enabled the structural characterization of the interior pocket of the ACP from the actinorhodin PKS[73] as well as co-crystal structures with multiple PKS enzymes (unpublished data). One obvious drawback to the use of these first generation probes is that they do not actually resemble a poly-β-ketone chain all that closely. In the future, new probes will need to be developed that more closely mimic the structure and physical properties of the natural substrates.

1.4 Introduction of dissertation chapters

Chapters 2 – 4 of this dissertation describe the results of my scientific efforts over the past four years in the lab of Sheryl Tsai at the University of California, Irvine. Chapter 2 describes the

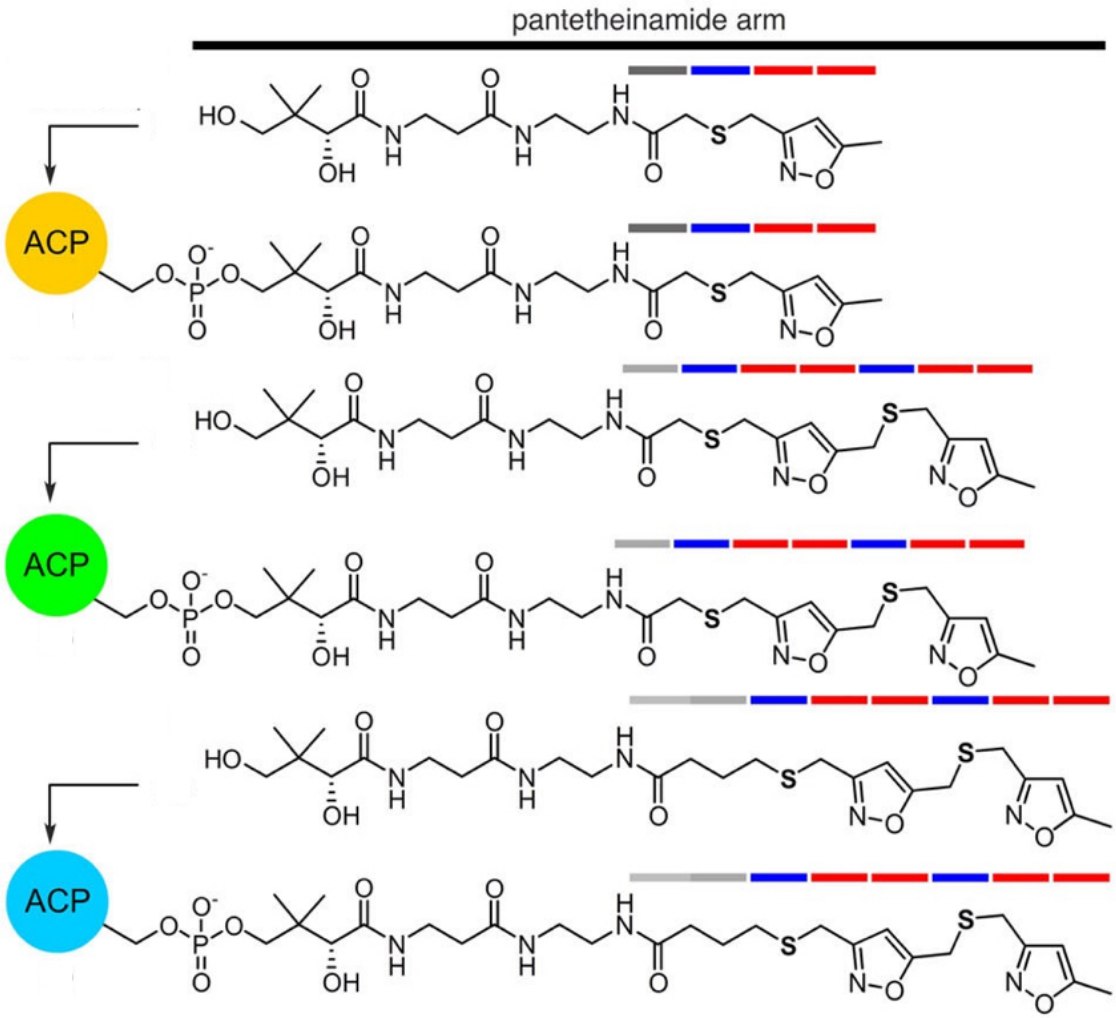

Figure 1-19 Isoxazole-based atom replacement probes for PKS structural studies. These synthetic probes can be used as substrate analogues for co-crystallization with PKS enzymes or chemoenzymatically loaded onto an ACP *in vitro*. The colored bars denote the number of ketide units. Figure from reference 73.

use of a mechanism-based crosslinker to determine the crystal structure of the complex between AcpP and FabB from the *E. coli* type II FAS. Chapter 3 details the development of an oxetane-based substrate isostere that enabled the structural characterization of the enzyme DpsC from a type II PKS. Chapter 4 outlines my efforts to structurally characterize two enzymes from the unique product release mechanism of a type I PKS. Finally, Chapter 5 offers conclusions, future directions, and some final thoughts on the state of FAS and PKS protein engineering efforts.

References

1. Biozentrum, University of Bazel. Biozentrum News, 2014. http://www.biozentrum.unibas.ch/news-events/news/news-details/article/the-factory-for-fatty-acids-is-ancient/ (accessed 2017).
2. Bukhari, H. S. T.; Jakob, R. P.; Maier, T. Evolutionary Origins of the Multienzyme Architecture of Giant Fungal Fatty Acid Synthase. *Structure* **2014,** *22* (12), 1775-1785.
3. Klein, H. P.; Lipmann, F. The relationship of coenzyme A to lipide synthesis I. Experiments with yeast. *Journal of Biological Chemistry* **1953,** *203,* 95-99.
4. Wakil, S. J. A malonic acid derivative as an intermediate in fatty acid biosynthesis. *Journal of the American Chemical Society* **1958,** *80,* 6465-6465.
5. Wakil, S. J. Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* **1989,** *28,* 4523-4530.
6. Maier, T.; Leibundgut, M.; Boehringer, D.; Ban, N. Structure and function of eukaryotic fatty acid synthases. *Quarterly Reviews of Biophysics* **2010,** *43* (3), 373-422.
7. White, S. W.; Zheng, J.; Zhang, Y.-M.; Rock, C. O. The Structural Biology of Type II Fatty Acid Biosynthesis. *Annual Reviews in Biochemistry* **2005,** *74,* 791-831.
8. Leibundgut, M.; Maier, T.; Jenni, S.; Ban, N. The multienzyme architecture of eukaryotic fatty acid synthases. *Current Opinion in Structural Biology* **2008,** *18,* 714-725.
9. Jenni, S.; Leibundgut, M.; Maier, T.; Ban, N. Architecture of a fungal fatty acid synthase at 5 Å resolution. *Science* **2006,** *311,* 1263-1267.
10. Maier, T.; Jenni, S.; Ban, N. Architecture of mammalian fatty acid synthase at 4.5 Å resolution. *Science* **2006,** *311,* 1258-1262.
11. Smith, S.; Witkowski, A.; Joshi, A. K. Structural and funtional organization of the animal fatty acid synthase. *Progress in Lipid Research* **2003,** *42,* 289-317.
12. Jenni, S.; Leibundgut, M.; Boehringer, D.; Frick, C.; Mikolasek, B.; Ban, N. Structure of fungal fatty acid synthase and implications for iterative substrate shuffling. *Science* **2007,** *316,* 254-261.
13. Maier, T.; Leibundgut, M.; Ban, N. The crystal structure of a mammalian fatty acid synthase. *Science* **2008,** *321,* 1315-1322.

14. Brink, J.; Ludtke, S. J.; Kong, Y.; Wakil, S. J.; Ma, J.; Chiu, W. Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis. *Structure* **2004,** *12,* 185-191.

15. Schweizer, E.; Hofmann, J. Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. *Microbiology and Molecular Biology Reviews* **2004,** *68,* 501-517.

16. Leibundget, M.; Jenni, S.; Frick, C.; Ban, N. Structural basis for substrate delivery by acyl carrier protein in the yeast fatty acid synthase. *Science* **2007,** *316,* 288-290.

17. Feng, D.; Witkowski, A.; Smith, S. Down-regulation of Mitochondrial Acyl Carrier Protein in Mammalian Cells Compromoses Protein Lipoylation and Respiratory Complex I and Results in Cell Death. *Journal of Biological Chemistry* **2003,** *278,* 20154.

18. Cronan, J. E.; Rock, C. O. *In Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology;* American Society for Microbiology: Washington, D.C., 1996; p 612.

19. Prescott, D. J.; Vagelos, P. R. Acyl carrier protein. *Advances in Enzymology and Related Areas of Molecular Biology* **1972,** *36,* 269-311.

20. Rock, C. O.; Cronan, J. E. The enigmatic acyl carrier protein phosphodiesterase of Escherichia coli. *Journal of Biological Chemistry* **1979,** *254,* 9778-9785.

21. Ohlrogge, J.; Savage, L.; Jaworski, J.; Voelker, T.; Post-Beittenmiller, D. Alteration of acyl-acyl carrier protein pools and acetyl-CoA carboxylase expression in Escherichia coli by a plant medium chain acyl-acyl carrier protein thioesterase. *Archives of Biochemistry and Biophysics* **1995,** *317,* 185-190.

22. Tropf, S.; Revill, W. P.; Bibb, M. J.; Hopwood, D. A.; Schweizer, M. Heterologously expressed acyl carrier protein domain of rat fatty acid synthase functions in Escherichia coli fatty acid synthase and Streptomyces coelicolor polyketide synthase systems. *Chemistry and Biology* **1998,** *5,* 135-146.

23. Zhang, Y. M.; Rao, M. S.; Heath, R. J.; Price, A. C.; Olson, A. J. Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III. *Journal of Biological Chemistry* **2001,** *276,* 8231-8238.

24. Roujeinikova, A.; Simon, W. J.; Gilroy, J.; Rice, D. W.; Rafferty, J. B.; Slabas, A. R. Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *Journal of Molecular Biology* **2007,** *365,* 135-145.

25. Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Marwick, P. R.; Bruegger, J.; Caldera-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, F.; O'dowd, B.; McCammon, J. A.; Opella, S. J.; Tsai, S. C.; Burkart, M. D. Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* **2014,** *505,* 427.

26. Cronan, J. E.; Waldrop, G. L. Multi-subunit acetyl-CoA carboxylases. *Progress in Lipid Research* **2002,** *41* (5), 407.

27. Ruch, F. E.; Vagelos, P. R. Characterization of a malonyl-enzyme intermediate and identification of the malonyl binding site in malonyl coenzyme A-acyl carrier protein transacylase of Escherichia coli. *Journal of Biological Chemistry* **1973,** *248,* 8095-8106.

28. Jackowski, S.; Murphy, C. M.; Cronan, J. E.; Rock, C. O. Acetoacetyl-acyl carrier protein synthase: a target for the antibiotic thiolactomycin. *Journal of Biological Chemistry* **1989,** *264,* 7624-7629.

29. Davies, C.; Heath, R. J.; White, S. W.; Rock, C. O. The 1.8 A crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from escherichia coli. *Structure* **2000,** *8,* 185-195.

30. Lai, C. Y.; Cronan, J. E. Isolation and characterization of β-ketoacyl-acyl carrier protein reductase (fabG) mutants of Escherichia coli and Salmonella enterica Serovar Typhimurium. *Journal of Bacteriology* **2004,** *186,* 1869-1878.

31. Price, A. C.; Zhang, Y. M.; Rock, C. O.; White, S. W. Cofactor-induced conformational rearrangements establish a catalytically competent active site and a proton relay conduit in FabG. *Structure* **2004,** *12,* 417-428.

32. Kass, L. R.; Bloch, K. On the enzymatic synthesis of unsaturated fatty acids in Escherichia coli. *Proceedings of the National Academy of Sciences* **1967,** *58,* 1168-1173.

33. Heath, R. J.; Rock, C. O. Roles of the FabA and FabZ beta-hydroxyacyl-acyl carrier protein dehydratases in Escherichia coli fatty acid biosynthesis. *Journal of Biological Chemistry* **1996,** *271,* 27795-27801.

34. Leesong, M.; Henderson, B. S.; Gillig, J. R.; Schwab, J. M.; Smith, J. L. Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure* **1996,** *4,* 253-264.

35. Kimber, M. S.; Martin, F.; Lu, Y.; Houston, S.; Vedadi, M.; Dharamsi, A.; Fiebig, K. M.; Schmid, M.; Rock, C. O. The structure of (3R)-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from Pseudomonas aeruginosa. *Journal of Biological Chemistry* **2004,** *279,* 52593-52602.

36. Heath, R. J.; Rock, C. O. Enoyl-acyl carrier protein reductase (fabI) plays a determinant role in completing cycles of fatty acid elongation in Escherichia coli. *Journal of Biological Chemistry* **1995,** *270,* 26538-26542.

37. Banerjee, A.; Dubnau, E.; Quemard, A.; Balasubramanian, V.; Um, K. S.; Wilson, T.; Collins, D.; de Lisle, G.; Jacobs, W. R. inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. *Science* **1994,** *263,* 227-230.

38. Heath, R. J.; Yu, Y. T.; Shapiro, M. A.; Olson, E.; Rock, C. O. Broad spectrum antimicrobial biocides target the FabI component of fatty acid synthesis. *Journal of Biological Chemistry* **1998,** *273,* 30316-30321.

39. Roujeinikova, A.; Sedelnikova, S.; de Boer, G. J.; Stuitje, A. R.; Slabas, A. R.; Rafferty, J. B.; Rice, D. W. Inhibitor binding studies on enoyl reductase reveal conformational changes related to substrate recognition. *Journal of Biological Chemistry* **1999,** *274,* 30811-30817.

40. Rafferty, J. B.; Simon, J. W.; Baldock, C.; Artymiuk, P. J.; Baker, P. J.; Stuitje, A. R.; Slabas, A. R.; Rice, D. W. Common themes in redox chemistry emerge from the X-ray structure of oilseed rape (Brassica napus) enoyl acyl carrier protein reductase. *Structure* **1995,** *3,* 927-938.

41. Olsen, J. G.; Kadziola, A.; von Wettstein-Knowles, P.; Siggaard-Andersen, M.; Lindquist, Y.; Larsen, S. The X-ray crystal structure of beta-ketoacyl (acyl carrier protein) synthase I. *FEBS Letters* **1999,** *460,* 46-52.

42. Huang, W.; Jia, J.; Edwards, P.; Dehesh, K.; Schneider, G.; Lindqvist, Y. Crystal structure of beta-ketoacyl-acyl carrier protein synthase II from E.coli reveals the molecular architecture of condensing enzymes. *EMBO Journal* **1998,** *17,* 1183-1191.

43. Garwin, J. L.; Klages, A. L.; Cronan, J. E. Structural, enzymatic, and genetic studies of beta-ketoacyl-acyl carrier protein synthases I and II of Escherichia coli. *Journal of Biological Chemistry* **1980,** *255,* 19949-19956.

44. Kortemme, T.; Creighton, T. E. Ionisation of cysteine residues at the termini of model alpha-helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *Journal of Molecular Biology* **1995,** *253,* 799-812.

45. Price, A. C.; Choi, K. H.; Heath, R. J.; Li, Z.; White, S. W.; Rock, C. O. Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism. *Journal of Biological Chemistry* **2001,** *276,* 6551-6559.

46. Witkowski, A.; Joshi, A. K.; Smith, S. Mechanism of the beta-ketoacyl synthase reaction catalyzed by the animal fatty acid synthase. *Biochemistry* **2002,** *41,* 10877-10887.

47. Borchardt, J. K. Combinatorial biosynthesis: panning for pharmaceutical gold. *Modern Drug Discovery* **1999,** *2,* 22-29.

48. National Cancer Institute. Aflatoxins. https://www.cancer.gov/about-cancer/causes-prevention/risk/substances/aflatoxins (accessed May 6, 2017).

49. Cox, R. J. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Organic and Biomolecular Chemistry* **2007,** *5,* 2010-2026.

50. Campbell, C. D.; Vederas, J. C. Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* **2010,** *93,* 755-763.

51. Kopp, F.; Marahiel, M. A. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Natural Products Reports* **2007,** *24,* 735-749.

52. Herbst, D. A.; Jakob, R. P.; Zahringer, F.; Maier, T. Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **2016,** *531,* 533-537.

53. Keating-Clay, A. The uncommon enzymology of cis-acyltransferase assembly lines. *Chemical Reviews* **2017**.

54. Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structure of a modular polyketide synthase. *Nature* **2014,** *510,* 512-517.

55. Nguyen, T.; Ishida, K.; Jenke-Kodama, H.; Dittmann, E.; Gurgui, C.; Hochmuth, T.; Taudien, S.; Platzer, M.; Hertweck, C.; Piel, J. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nature Biotechnology* **2008,** *26,* 225-233.

56. Cheng, Y. Q.; Tang, G. L.; Shen, B. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proceedings of the National Academy of Sciences* **2003,** *100,* 3149-3154.

57. Musiol, E. M.; Härtner, T.; Kulik, A.; Moldenhauer, J.; Piel, J.; Wohlleben, W.; Weber, T. Supramolecular templating in kirromycin biosynthesis: the acyltransferase KirCII loads ethylmalonyl-CoA extender onto a specific ACP of the trans-AT PKS. *Chemistry and Biology* **2011,** *18,* 438-444.

58. Zhao, C. H.; Coughlin, J. M.; Ju, J. H.; Zhu, D. Q.; Wendt-Pienkowski, E.; Zhou, X. F.; Wang, Z. J.; Shen, B.; Deng, Z. X. Oxazolomycin biosynthesis in Streptomyces albus JA3453 featuring an "acyltransferase-less" type I polyketide synthase that incorporates two distinct extender units. *Journal of Biological Chemistry* **2010,** *285,* 20097-20108.

59. Menche, D.; Arikan, F.; Perlova, O.; Horstmann, N.; Ahlbrecht, W.; Wenzel, S. C.; Jansen, R.; Irschik, H.; R., M. Stereochemical determination and complex biosynthetic assembly of etnangien, a highly potent RNA polymerase inhibitor from the myxobacterium Sorangium cellulosum. *Journal of the American Chemical Society* **2008,** *130,* 14234-14243.

60. Helfrich, E. J. N.; Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural Product Reports* **2016,** *33,* 231-316.

61. Huang, Y.; Tang, G. L.; Pan, G.; Chang, C. Y.; Shen, B. Characterization of the ketosynthase and acyl carrier protein domains at the LnmI nonribosomal peptide synthetase−polyketide synthase interface for leinamycin biosynthesis. *Organic Letters* **2016,** *18,* 4288-4291.

62. Hertwick, C.; Luzhetsky, A.; Rebets, Y.; Bechthold, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Natural Product Reports* **2007,** *24,* 162-190.

63. Keatinge-Clay, A. T. . M. D. A.; Medzihradsky, K. F.; Khosla, C.; Stroud, R. An antibiotic factory caught in action. *Nature Structural and Molecular Biology* **2004,** *11,* 888-893.

64. Tang, Y.; Tsai, S. C.; Khosla, C. Polyketide chain length control by chain length factor. *Journal of the American Chemical Society* **2003,** *125,* 12708-12709.

65. Das, A.; Khosla, C. Biosynthesis of aromatic polyketides in bacteria. *Accounts of Chemical Research* **2009,** *42,* 631-639.

66. Lee, M. Y.; Ames, B. D.; Tsai, S. C. Inishgt into the molecular basis of aromatic polyketide cyclization: crystal structure and in vitro characterization of WhiE-ORFVI. *Biochemistry* **2012,** *51,* 3079-3091.

67. Bruegger, J.; Haushalter, R. W.; Vagstad, A. L.; Shakya, G.; Mih, N.; Townsend, C. A.; Burkart, M. D.; Tsai, S. C. Probing the selectivity and protein-protein interactions of a nonreducing fungal polyketide synthase using mechanism-based crosslinkers. *Chemistry and Biology* **2013,** *20* (9), 1135.

68. Crawford, J. M.; Townsend, C. A. New insights into the formation of fungal aromatic polyketides. *Nature Reviews Microbiology* **2010,** *8,* 879-889.

69. Ishikawa, F.; Haushalter, R. W.; Burkart, M. D. Dehydratase-specific probes for fatty acid and polyketide synthase. *Journal of the American Chemical Society* **2012,** *134,* 769-772.

70. Worthington, A. S.; Rivera, H.; Torpey, J. W.; Alexander, M. D.; Burkart, M. D. Mechanism-based protein cross-linking probes to investigate carrier protein-mediated biosynthesis. *ACS Chemical Biology* **2006,** *1,* 687-691.

71. Miyanaga, A.; Iwasawa, S.; Shinohara, Y.; Judo, F.; Eguchi, T. Structure-based analysis of the molecular interactions between acyltransferase and acyl carrier protein in vicenistatin biosynthesis. *Proceedings of the National Academy of Sciences* **2016,** *113,* 1802-1807.

72. Caldera-Festin, G.; Jackson, D. R.; Barajas, J. F.; Valentic, T. R.; Patel, A. B.; Aguilar, S.; Nguyen, M.; Vo, M.; Khanna, A.; Sasaki, E.; Liu, H. W.; Tsai, S. C. Structural and

functional analysis of two di-domain aromatase/cyclases from type II polyketide synthases. *Proceedings of the National Academy of Sciences* **2015,** *112,* E6844-E6851.

73. Shakya, G.; Rivera, H.; Lee, D. J.; Jaremko, M. J.; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdziowski, E. R.; Wong, F.; Tsai, S. C.; Burkart, M. D. Modeling linear and cyclic PKS intermediates through atom replacement. *Journal of the American Chemical Society* **2014,** *136,* 16792-16799.

CHAPTER 2

Mechanism-Based Crosslinking Enables Structure-Guided Engineering of Type II Acyl Carrier Protein Interactions

2.1 Abstract

The *E. coli* type II fatty acid synthase (FAS) has been widely studied as a platform for biofuel production given its ability to efficiently produce hydrocarbon chains. Although the individual enzymes in the pathway have been structurally and functionally characterized, uncovering the key protein-protein interactions, particularly between the acyl carrier protein (ACP) and its partner enzymes, has remained elusive due to the transient nature of the interactions. Without this structural information, protein engineering and synthetic biology efforts have proven unsuccessful. This chapter describes the application of a mechanism-based crosslinker to trap an ACP-ketosynthase (KS) complex and solve its crystal structure at 2.4 Å resolution. The structure of this ACP-KS complex reveals the key interactions between the ACP (AcpP) and one of the KS enzymes in the pathway (FabB). Structural analysis suggests that allosteric mechanisms exist to facilitate enzyme turnover by forcing the dissociation of one AcpP upon docking of a second AcpP to the opposite FabB monomer. Using this structural data, several surface mutants of FabB were generated and overexpressed in *E. coli* BL21 cells. As hypothesized, the lipid profile of the cultures was altered due to the modulation of the key protein-protein complex that affects chain length and saturation. This work represents the first snapshot of an ACP-KS interface as well as the first successful attempt to alter the product profile of a fatty acid synthase by specifically targeting protein-protein interactions.

2.2 Introduction

The *E. coli* type II fatty acid synthase (FAS) consists of 13 discrete enzymes that work together to biosynthesize fatty acids[1]. The FAS efficiently biosynthesizes hydrocarbons from simple building blocks, and therefore it is a major target for biofuel development[2]; additionally, these essential enzymes are targets for antibiotic discovery[3]. The *E. coli* acyl carrier protein (ACP), AcpP, is at the center of this pathway, sequestering unstable intermediates in a hydrophobic core for transport through the cytosol and delivery to appropriate reaction partners[1]. Rational engineering of FAS protein-protein interactions and modulation of fatty acid production requires structural information of how AcpP interacts with its partner enzymes. However, the transient nature of ACP-partner interactions has thus-far hindered structural studies of FAS complexes and precluded rational engineering of protein-protein interactions[4].

A mechanism-based crosslinking approach was recently applied to solve the crystal structure of AcpP in functional association with FabA, the *beta*-hydroxyl-ACP dehydratase, which gave new insight into AcpP's "switchblade" mechanism for substrate delivery[5]. A chloroacryl-pantetheine crosslinker **1** specific to the active site chemistry of ketosynthases (KS) has also been developed[6] (Fig 2-1), though its application to complex structure determination has not yet been described.

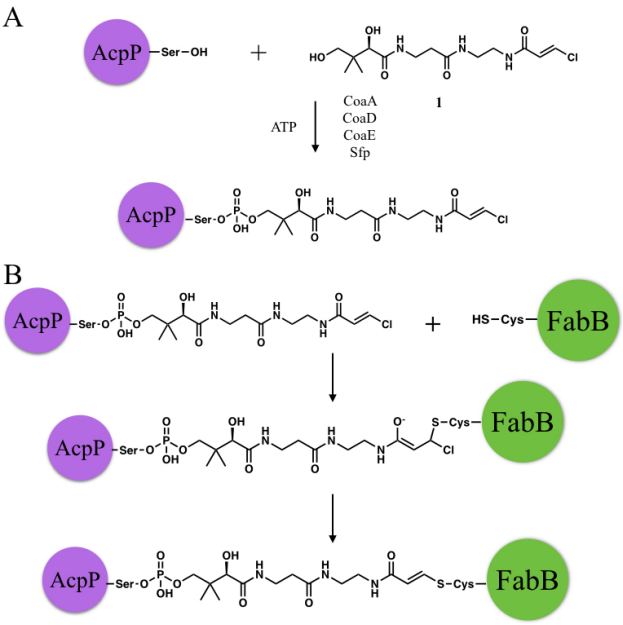

Figure 2-1 The chloroacryl-pantetheine crosslinker forms a covalent linkage with the FabB active site. (a) The crosslinker can be loaded onto recombinant AcpP using coenzyme A biosynthetic enzymes and a PPT transferase Sfp. (b) The mechanism of AcpP-FabB crosslinking.

In this work, crosslinker **1** was applied to obtain a stable complex between AcpP and FabB and to characterize the interactions between the two proteins using X-ray crystallography. The crystal structure revealed the phosphopantetheine (PPT) binding residues within the FabB active site, and an ACP-KS interface was structurally characterized for the first time. A comparison of the AcpP-FabA and AcpP-FabB complex structures identified unique AcpP residues that recognize each partner enzyme. Finally, this structural information was used to engineer altered fatty acid production *in vivo* by mutating FabB surface residues to disrupt its protein-protein interactions.

2.3 Results and discussion

2.3.1 Complex purification, crystallization, and X-ray diffraction analysis

AcpP and FabB were expressed separately in *E. coli* as His_6-tagged constructs and successfully purified using immobilized metal affinity chromatography (IMAC) (Figure 2-2). The purified AcpP was then successfully loaded at position S36 with crosslinker **1** (kindly provided by our co-worker Mike Burkart at the University of California, San Diego), as confirmed by MALDI-TOF mass spectrometry (Fig 2-3). Once the loaded AcpP was obtained, the conditions for the

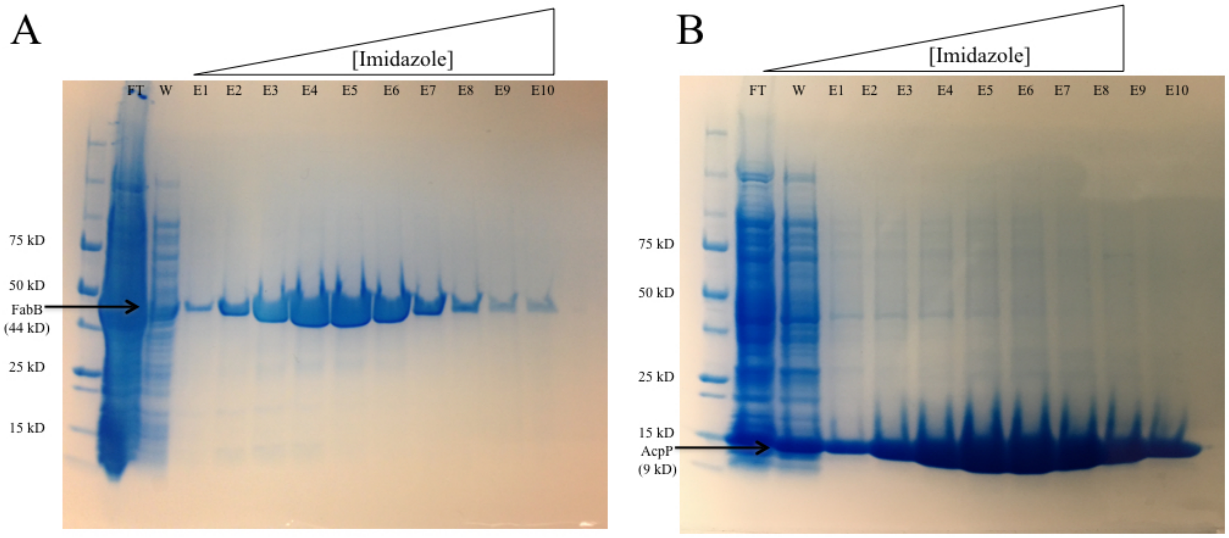

Figure 2-2 Purification of FabB and AcpP. (A) SDS-PAGE gel of fractions from IMAC purification of FabB. (B) SDS-PAGE gel of fractions from IMAC purification of AcpP.

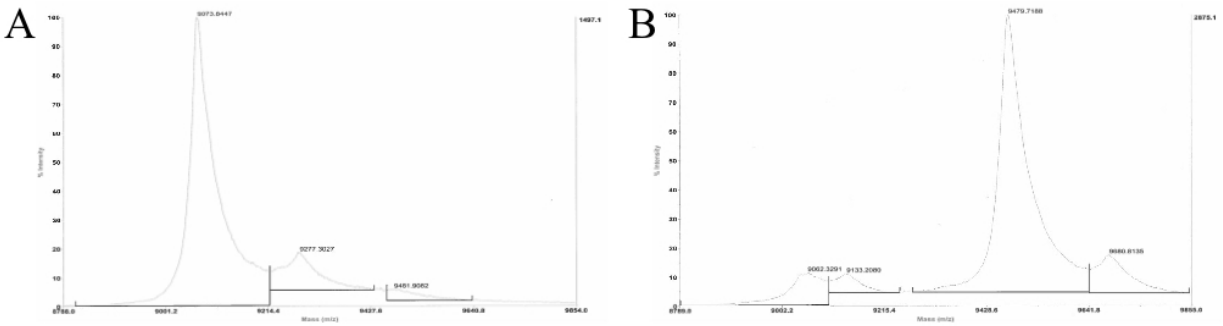

Figure 2-3. AcpP loading confirmed by MALDI-TOF mass spectrometry. (A) Mass spectrum for unloaded AcpP. (B) Mass spectrum for AcpP loaded with crosslinker **1**.

crosslinking reaction were optimized by varying buffer content and pH (data not shown). The optimal buffer conditions were found to be 100 mM Tris pH 7.5, 100 mM NaCl, 5 mM dithiothreitol (DTT), and 5 % glycerol. DTT was found to be particularly effective at pushing the reaction towards completion (Fig 2-4A).

The crosslinking reaction was scaled up under the identified optimal conditions, and the resulting complex was successfully purified using size exclusion chromatography (data not shown). The sample was screened for crystallization using high-throughput methods, and the optimal conditions for crystallization were found to be 0.1 M sodium acetate pH 5.4, 0.2 M ammonium acetate, and 20% PEG 4000 (Fig 2-4B). The diffraction pattern of the crystals was then

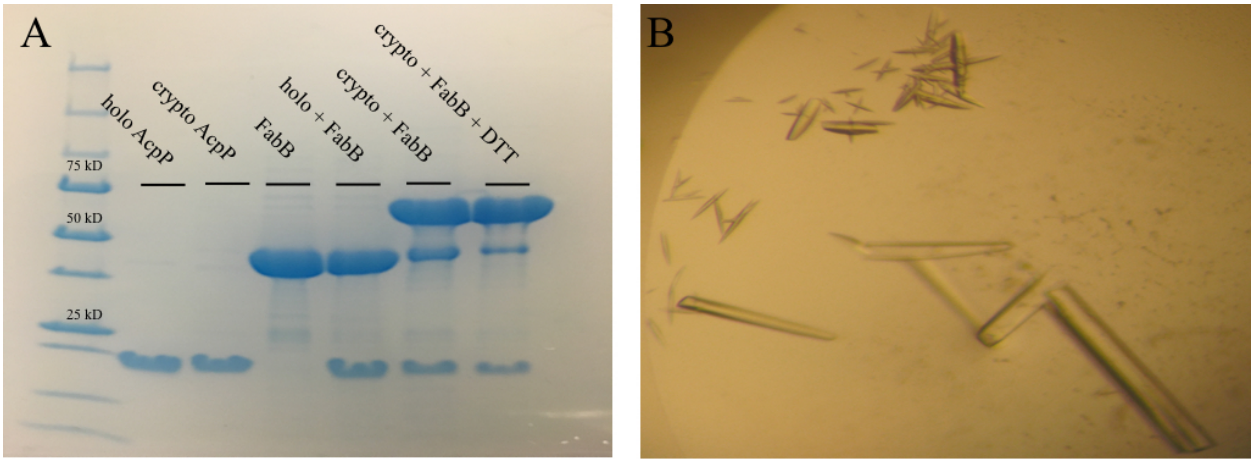

Figure 2-4 Covalent complex formation and crystallization. (A) SDS-PAGE gel demonstrating that crypto AcpP is required to form a covalent complex. "Holo" AcpP indicates AcpP loaded with its natural PPT cofactor, and "crypto" AcpP indicates AcpP loaded with crosslinker **1**. (B) AcpP-FabB complex crystals used for X-ray diffraction analysis.

	AcpP-FabB (5KOF)
Data collection	
Wavelength (Å)	1.0
Total reflections	72648 (7131)
Unique reflections	37073 (3680)
Space group	P 1 2_1 1
Cell dimensions	
a, *b*, *c* (Å)	59.33, 103.92, 83.4001
α, β, γ (°)	90, 110.61, 90
Resolution (Å)	39.2 – 2.4 (2.486 – 2.4)
R_{merge}	0.04978 (0.3953)
R_{meas}	0.0704 (0.559)
$I/\sigma(I)$	7.27 (1.76)
$CC_{1/2}$	0.996 (0.824)
CC*	0.999 (0.95)
Completeness (%)	100 (100)
Redundancy	2.0 (1.9)
Wilson B-factor	40.87
Refinement	
Resolution (Å)	39.2 – 2.4 (2.486 – 2.4)
No. reflections	37044 (3669)
R_{work}	0.2080
R_{free}	0.2254
No. atoms	
Protein	7110
Ligands	50
B factors	
Protein	61.77
Ligands	75.55
Water	54.44
Ramachandran	
Favored (%)	93.4
Allowed (%)	6.0
Outliers (%)	0.6
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.46

Table 2-1 Summary of X-ray crystallographic data collection and data refinement statistics.

measured, and the data were processed to a resolution of 2.4 Å. The phases were obtained by molecular replacement using a FabB crystal structure (PDB: 2VB9) as a search model (see Table 2-1).

2.3.2 Structural analysis

The AcpP-FabB complex crystallized in a 2 : 2 ratio (Fig 2-5A). FabB forms a tight homodimer in solution (*1*), and each of the monomers in the structure is crosslinked to an AcpP molecule. Well defined electron density for the crosslinker **1** can be observed stretching from S36 of AcpP to C163 in the active site of FabB (Fig 2-5B).

This crystal structure enabled the first visualization of PPT binding within the FabB active site (Fig 2-6). In addition to the covalent linkage with C163, the probe makes hydrogen bonds with H298 and H333 within the FabB pocket. These residues are known to be important for catalysis, presumably by serving as an "oxyanion hole" to stabilize the buildup of negative charge during

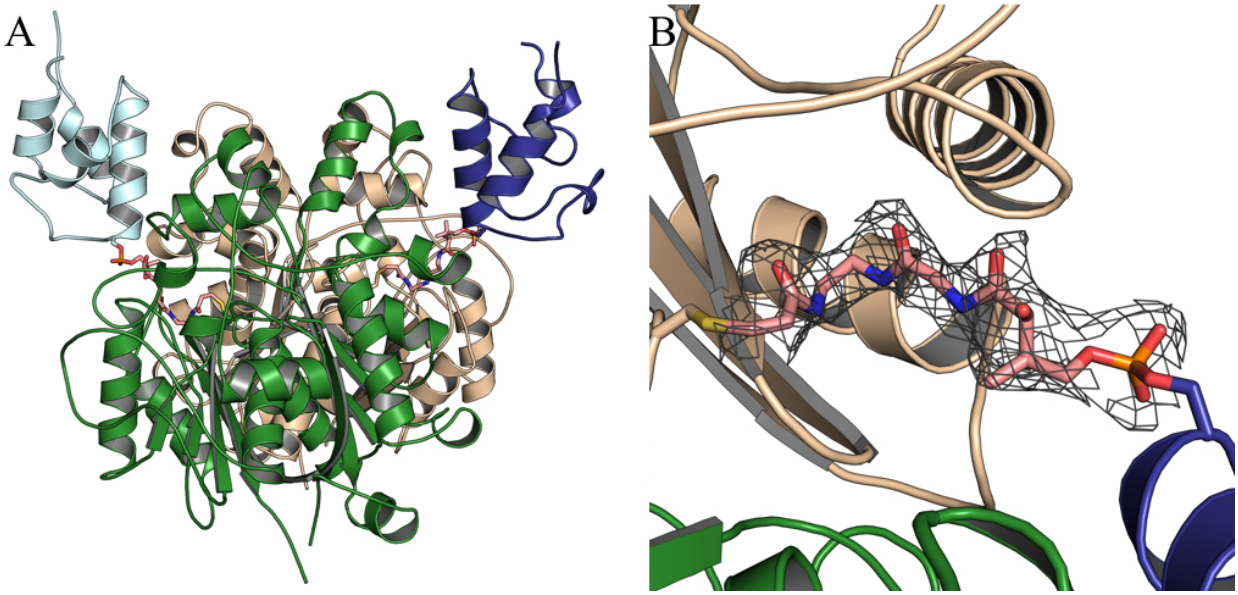

Figure 2-5 The X-ray crystal structure of the AcpP-FabB complex. (A) The overall AcpP-FabB structure with the FabB monomers shown in green and tan, the AcpP molecules shown in light blue and dark blue, and the crosslinker **1** shown in pink sticks. (B) The $2F_o - F_c$ electron density map for the crosslinker **1** within the FabB active site contoured at 1.0 sigma.

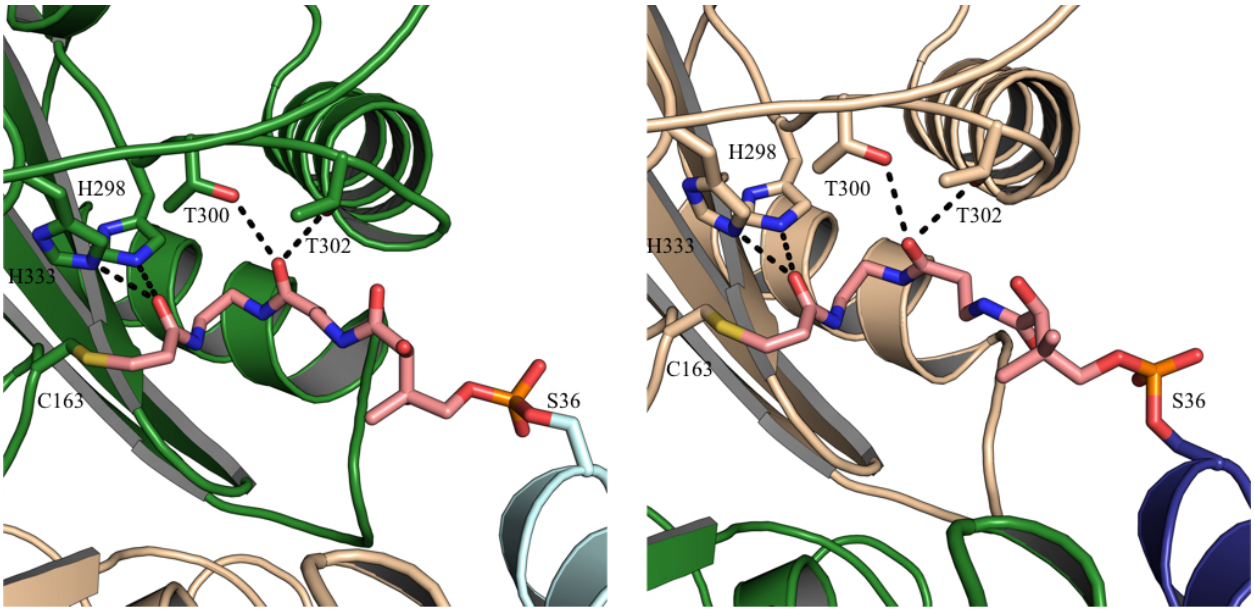

Figure 2-6 PPT binding within the FabB active site. The residues that form hydrogen bonds with the crosslinker **1** are shown for both AcpP-FabB pairs in the dimeric crystal structure.

the course of the reaction[7]; this structure is consistent with that hypothesis, as the two histidine side chains are in ideal positions within the PPT binding site to serve this purpose. Additionally, T300 and T302 also make hydrogen bonds with the probe, likely orienting the substrate for catalysis.

Another notable observation was made within the internal pocket of AcpP when bound to FabB. AcpP typically sequesters the nascent acyl chain within its hydrophobic core[8] (Fig 2-7A),

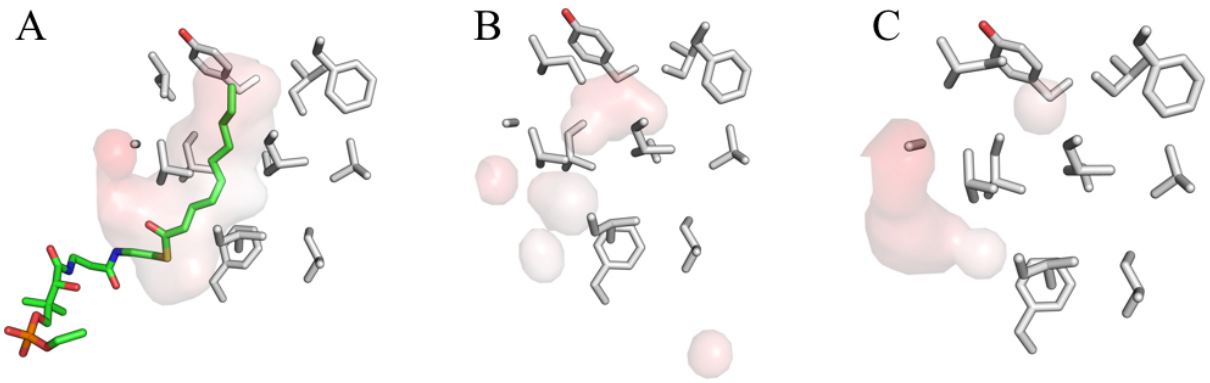

Figure 2-7 The interior pocket of AcpP collapses upon binding to partner enzymes. (A) AcpP sequestering a 10-carbon fatty acid within its hydrophobic core (PDB: 2FAE). The residues forming the pocket are shown in gray sticks, and the available pocket volume is shown. (B) AcpP pocket upon binding to FabB. (C) AcpP pocket upon binding to FabA (PDB: 4KEH).

though upon binding to FabB, the interior chamber of AcpP appears to collapse (Fig 2-7B). In particular, several hydrophobic residues that form this pocket shift towards the interior of the protein, forcing the pocket's collapse. This suggests that binding of AcpP to its partner enzymes forces a conformational change that facilitates substrate delivery to the nearby active site. The same effect was observed in the AcpP-FabA complex structure[5].

This work represents the first structural charaterization of an ACP in functional association with a KS. This snapshot of the ACP-KS interface revealed key insights into how AcpP is able to recognize and bind to one of its key partners. Consistent with the AcpP-FabA complex structure (*5*), AcpP docks with FabB primarily using α-helix 2 and, to a lesser extent, α-helix 3. The interface

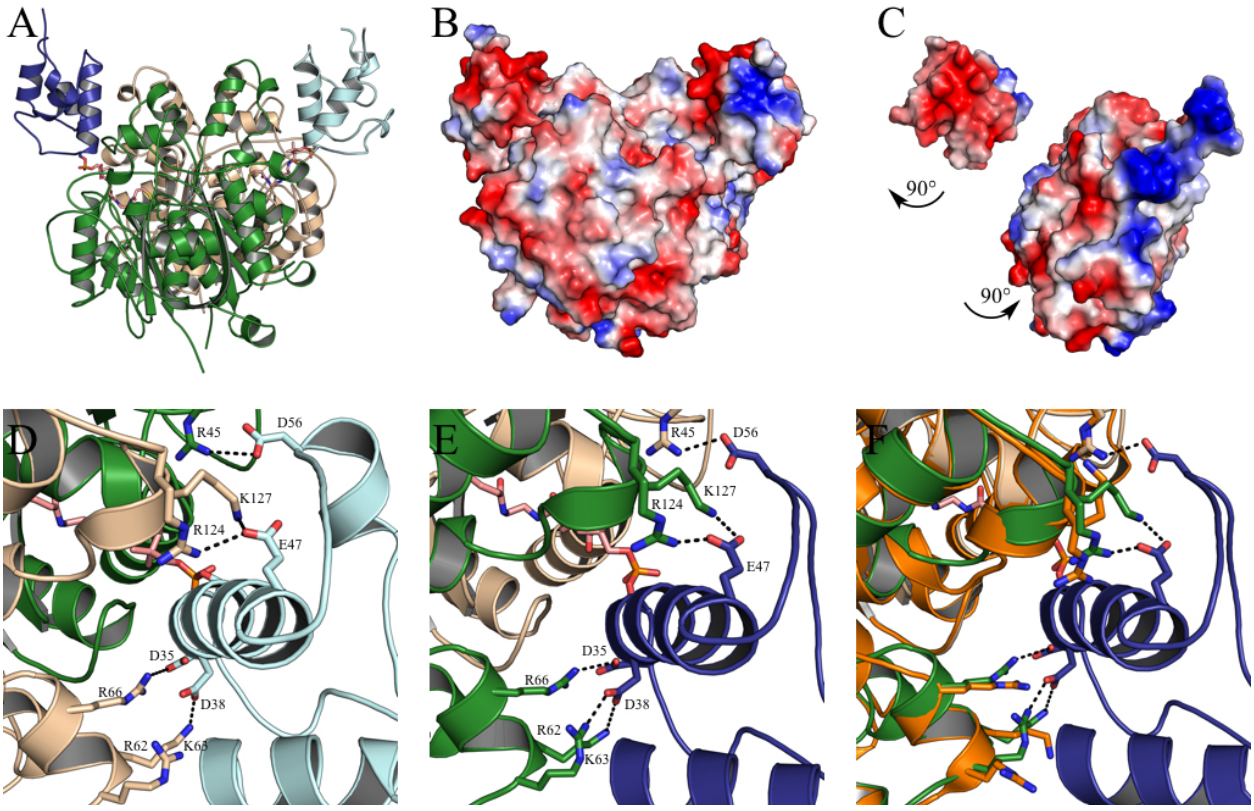

Figure 2-8 AcpP-FabB interface. (A) Overall AcpP-FabB structure. (B) AcpP-FabB structure represented as surface colored with electrostatic potentials (red for negative, blue for positive). (C) Surface electrostatics at the AcpP-FabB interface. (D) (E) AcpP-FabB interface residues from each of the AcpP-FabB pairs. (F) AcpP-FabB interface overlayed with unbound FabB (orange).

displays both charge and shape complementarity, as the highly acidic α-helix 2 of AcpP binds at a basic patch on the FabB surface (Fig 2-8C). The interface areas for both AcpP-FabB pairs are relatively small (451.1 $Å^2$ and 445.8 $Å^2$), consistent with the transient nature of AcpP-partner interactions.

This crystal structure enabled the identification of specific residues that form the AcpP-FabB interface (Fig 2-8D-E). These interface interactions are exclusively electrostatic, in contrast with the AcpP-FabA interface which contained a mix of electrostatic and hydrophobic interactions[5]. Residue D35 at the bottom of α-helix 2 of AcpP forms a salt bridge with R66 on FabB, and D38 nearby on AcpP forms salt bridges with R62 and K63 on FabB. Additionally, E47 near the top of α-helix 2 on AcpP forms salt bridges with R124 and K127 on FabB. Finally, D56 on α-helix 3 of AcpP forms a salt bridge with R45 on FabB. Most of the interactions occur between pairs that are not crosslinked to each other; the only exception is the salt bridge between D56 and R45. This indicates that AcpP does not dock to the monomer that it actually delivers the substrate to. This observation is also consistent with the AcpP-FabA structure[5].

The AcpP-FabB complex structure suggests that there are some subtle conformational changes that occur in FabB upon AcpP binding. Almost all of the FabB side chains that form interactions with AcpP flip positions with respect to their locations in the free FabB crystal structure in order to form the proper interface structure (Fig 2-8F). This

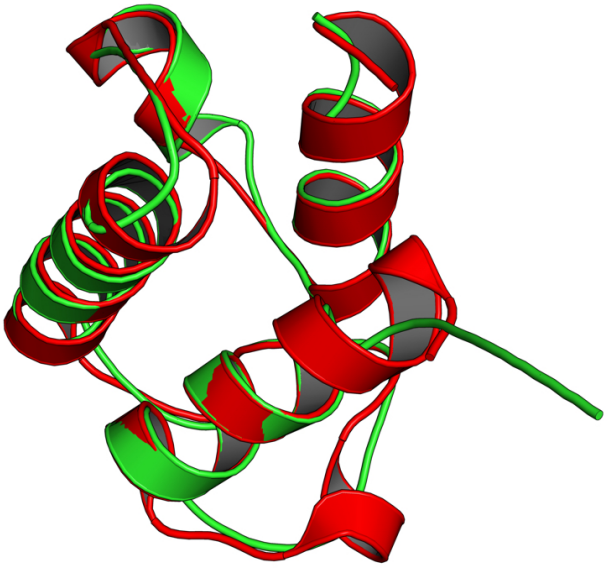

Figure 2-9 Comparison of AcpP bound to FabB and to FabA. The AcpP structure from the AcpP-FabB complex structure (green) and the AcpP-FabA complex structure (red) are overlaid.

also seems to require some slight movement in the FabB α-helices that sit at the interface. However, there do not appear to be any major conformational differences in AcpP when bound to either FabB or FabA, as both structures overlay quite well along α-helices 2 and 3 which form the interface (Fig 2-9).

A comparison between the two AcpP-FabB pairs in the crystal structure reveals some subtle differences between the two interfaces. The residues forming the two interfaces and their orientations seem to be quite similar, though it is curious that α-helix 3 of AcpP (where residue D56 resides) has become unstructured in one of the AcpPs (Fig 2-8D-E). Similarly, an examination of the crystallographic B-factors reveals that the same AcpP with a disordered α-helix 3 has significantly higher B-factors in this area than the corresponding region in the opposite AcpP. Finally, the FabB that this disordered, mobile AcpP is crosslinked to has a local spike in B-factors near its active site. Inspection of the crystal packing indicates that these observations are not likely to be artifacts of crystallization.

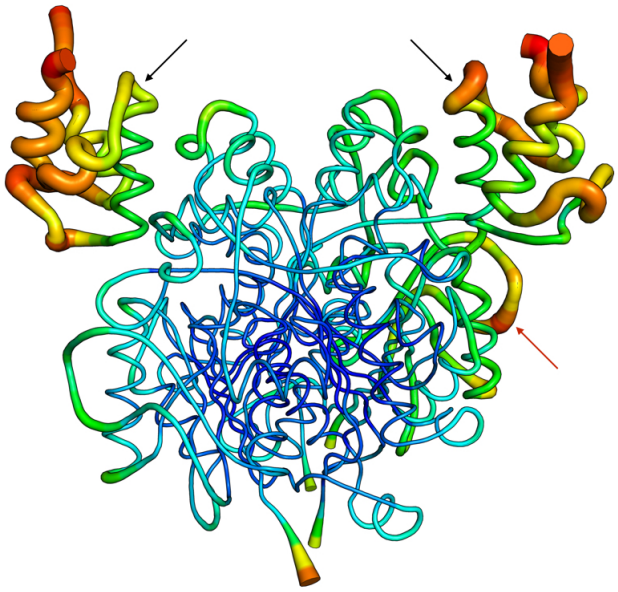

Figure 2-10 Crystallographic B-factors in the AcpP-FabB complex crystal structure. In this representation, thicker and darker tubes correspond to higher B-factors. The black arrows indicate the regions on each AcpP α-helix III that do not have similar B-factors, and the red arrow indicates the area near one of the FabB active sites with increased local B-factors.

These observations suggest that this crystal structure may have captured the complex in transition, namely the docking of one AcpP molecule and the dissociation of another. This second AcpP molecule is not able to leave in this situation since it is covalently tethered to its FabB

partner, but this thermal motion data seems to indicate that the FabB and AcpP are attempting to "shake free" of one another in this complex structure. This suggests that in the natural complex, only one AcpP is typically bound to the FabB dimer at one time; as one AcpP docks to one side of the FabB dimer, the AcpP on the other side dissociates. This could be a powerful mechanism to assist in enzyme turnover.

This hypothesis is consistent with the observation that pushing the crosslinking reaction to 50% completion is relatively simple, while pushing the reaction past 50% and to 100% is quite difficult. It is also consistent with the observation that there are subtle changes in the FabB surface upon AcpP binding, as previously mentioned (Fig 2-8F). These structural changes at one interface could also alter the opposite interface in a manner that results in negative binding cooperativity. Future kinetic and structural data will be required to confirm this hypothesis.

A comparison of AcpP's interface with FabB and with FabA reveals that AcpP uses different surface residues to bind to these two partner enzymes (Fig 2-11). In particular, AcpP uses two residues near the bottom of α-helix 2, D35 and D38, to iteract with FabB; however, this area

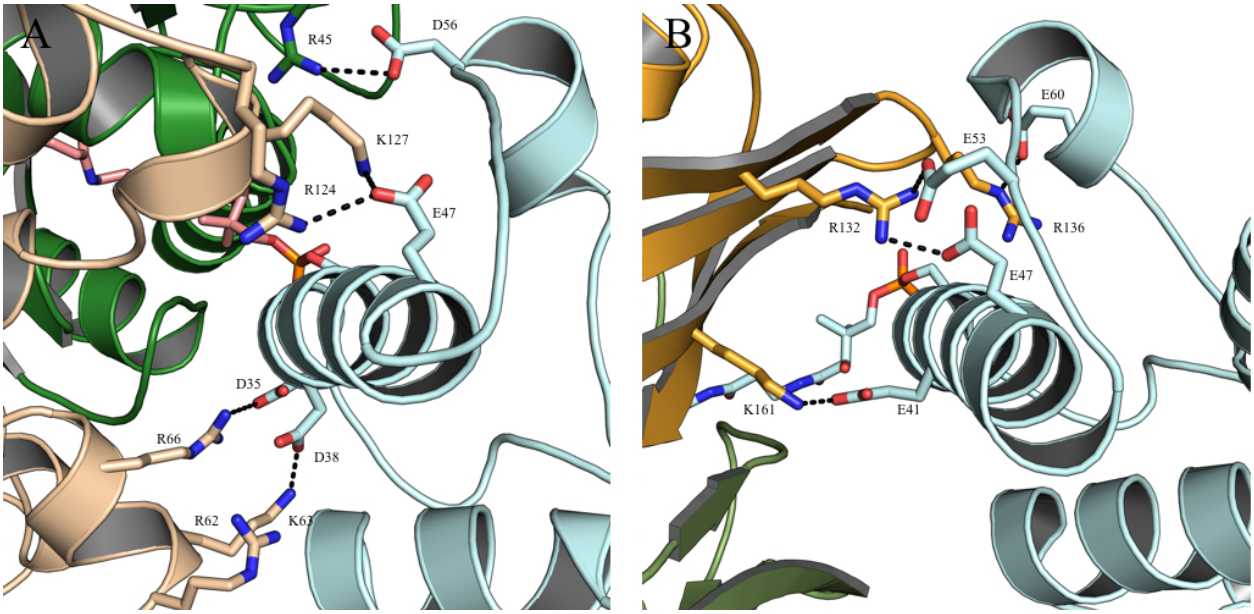

Figure 2-11 Comparison of the AcpP-FabB and AcpP-FabA interfaces. A view down α-helix 2 of AcpP in the AcpP-FabB (A) and AcpP-FabA (B) complex structures.

is not close enough to FabA to form any salt bridges due to the different binding mode of AcpP to FabA. Also, AcpP seems to use more residues near the top of α-helix 2 as well as residues on α-helix 3 to bind to FabA than it uses to bind to FabB. These results suggest that there is not a single binding "hot spot" on AcpP, though the interactions in these structures are localized to α-helices 2 and 3. AcpP-partner binding appears to be facilitated by a general, negatively-charged region of AcpP, but the specificity of AcpP-partner interactions seems to be promoted by fine-tuning residues at the interface of individual AcpP-partner complexes.

This observation highlights the importance of determining the structure of AcpP in complex with each of its catalytic partners, as there does not seem to be any way to predict general

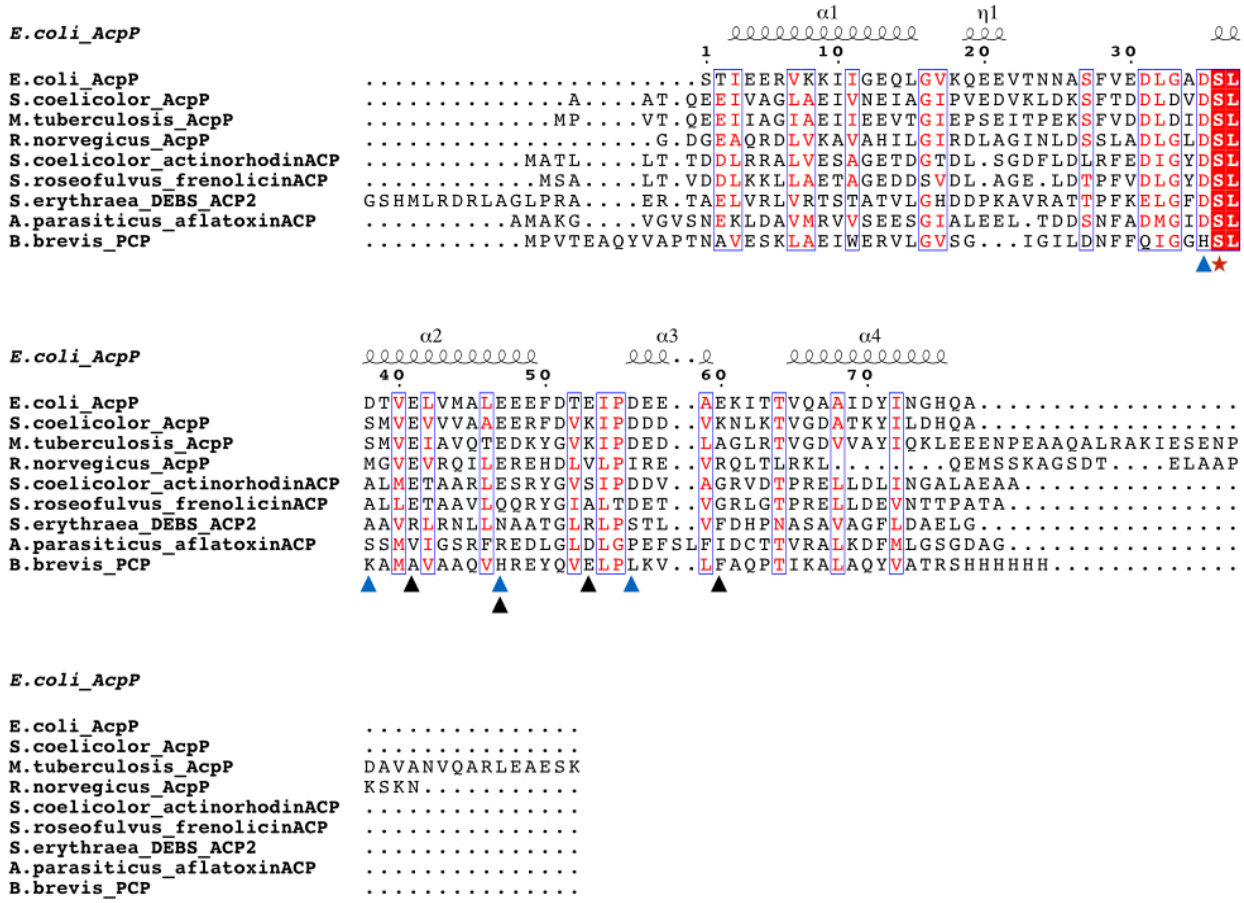

Figure 2-12 Sequence alignment of AcpP with other ACPs from other type II FASs, type I FASs, type II polyketide synthases (PKS), type I PKSs, and non-ribosomal peptide synthetases (NRPS). Blue triangles indicate residues at the FabB interface, black triangles indicate residues at the FabA interface, and a red star indicates the site of PPT attachment.

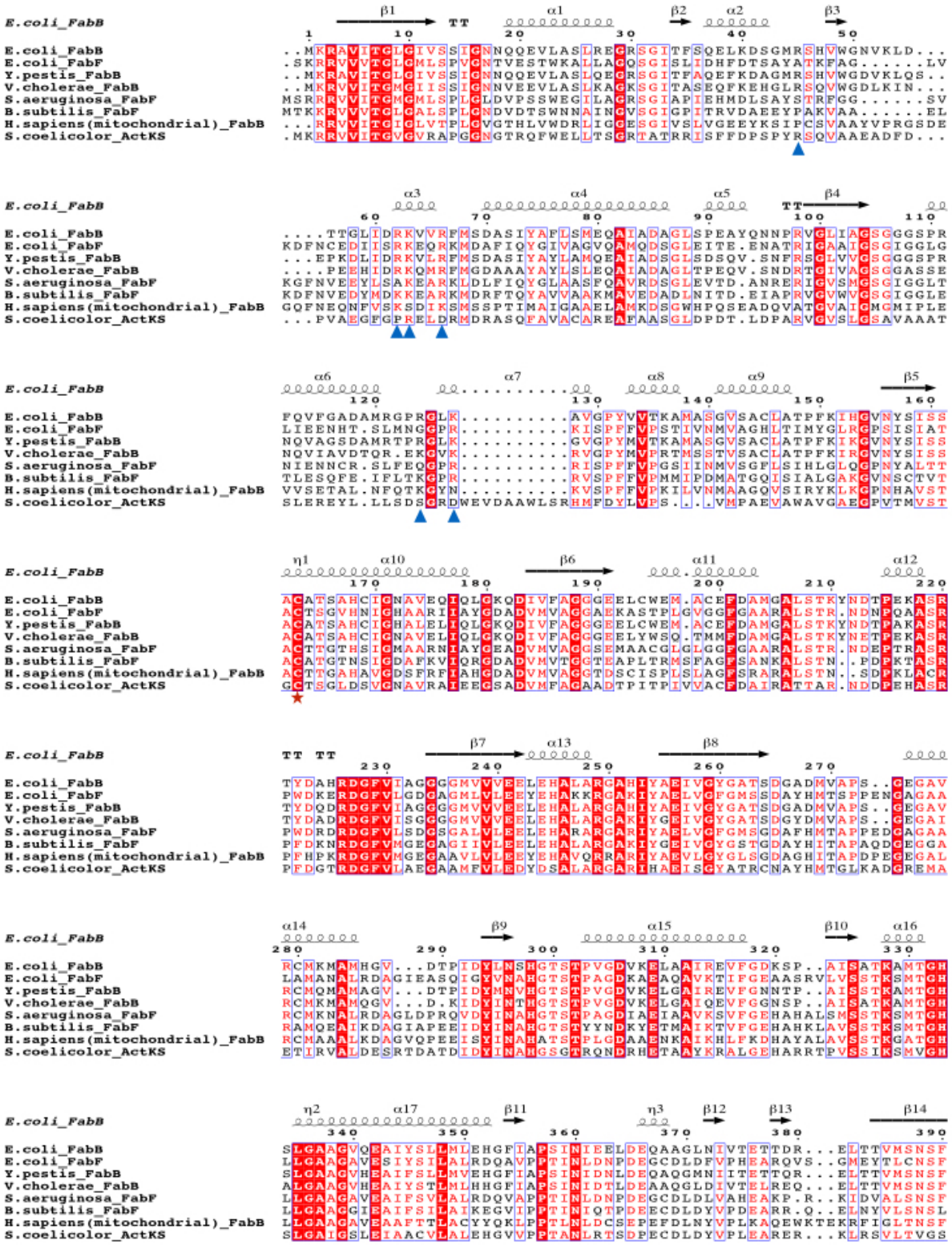

Figure 2-13 Sequence alignment of FabB with other KSs from other type II FASs, type I FASs, and type II polyketide synthases (PKS). Blue triangles indicate residues at the AcpP interface, and a red star indicates the active site cysteine.

AcpP interactions based on a single structure. A sequence alignment of AcpP with other ACPs from different organisms and different biosynthetic pathways also reveals that ACP-partner interactions are likely to be specific to each individual pathway, as many of the residues that AcpP uses to interact with its partners are not highly conserved (Fig 2-12). A similar trend is observed when comparing the FabB sequence with related KS enzymes from other organisms and biosynthetic pathways: several of the residues that form the interface with AcpP are not highly conserved (Fig 2-13).

2.3.3 FabB mutagenesis and lipid profile modulation in *E. coli*

With the knowledge of AcpP-FabB interactions in hand, the next goal was to use this information to modulate the lipid profile of the *E. coli* type II FAS *in vivo*. It has been shown that the relative abundance of FabB can affect the fatty acid product profile *in vitro*, namely the chain length and saturation level of the fatty acids[9]. This is due to FabB's unique ability to efficiently extend 10-carbon *cis*-unsaturated fatty acids generated in the dehydration step by FabA, as opposed to FabF—the other KS in the pathway—which will not efficiently extend these 10-carbon unsaturated fatty acids. Thus, it was hypothesized that introducing surface mutations in FabB at the interface with AcpP should weaken the interactions, thereby lowering the "effective concentration" of FabB *in vivo* and leading to a decrease in unsaturated fatty acid production and a decrease in the overall chain length of the fatty acids.

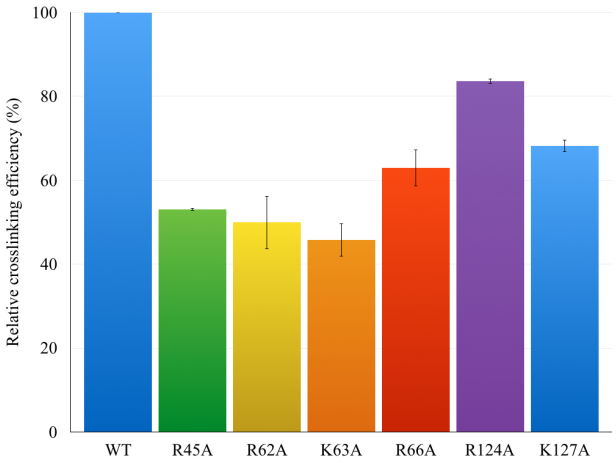

Figure 2-14 Relative crosslinking efficiency of FabB surface mutants with wild type AcpP.

In order to confirm that the AcpP-FabB interactions could be disrupted by mutagenesis, alanine mutants of each of the FabB residues at the AcpP interface (based on the crystal structure) were expressed and purified. The relative crosslinking efficiency of the mutants with wild type AcpP was then determined (Fig 2-14), an established method of measuring relative affinity[4]. As expected, the single alanine mutants resulted in less efficient crosslinking, suggesting that these FabB mutants did have lower affinity for AcpP *in vitro*.

The above hypothesis was then tested by overexpressing these mutants in *E. coli* cultures, an established method of probing the effect of type II FAS mutations on fatty acid production[10]. The culture lipids were then extracted, derivatized to fatty acid methyl esters, and anlyzed by gas

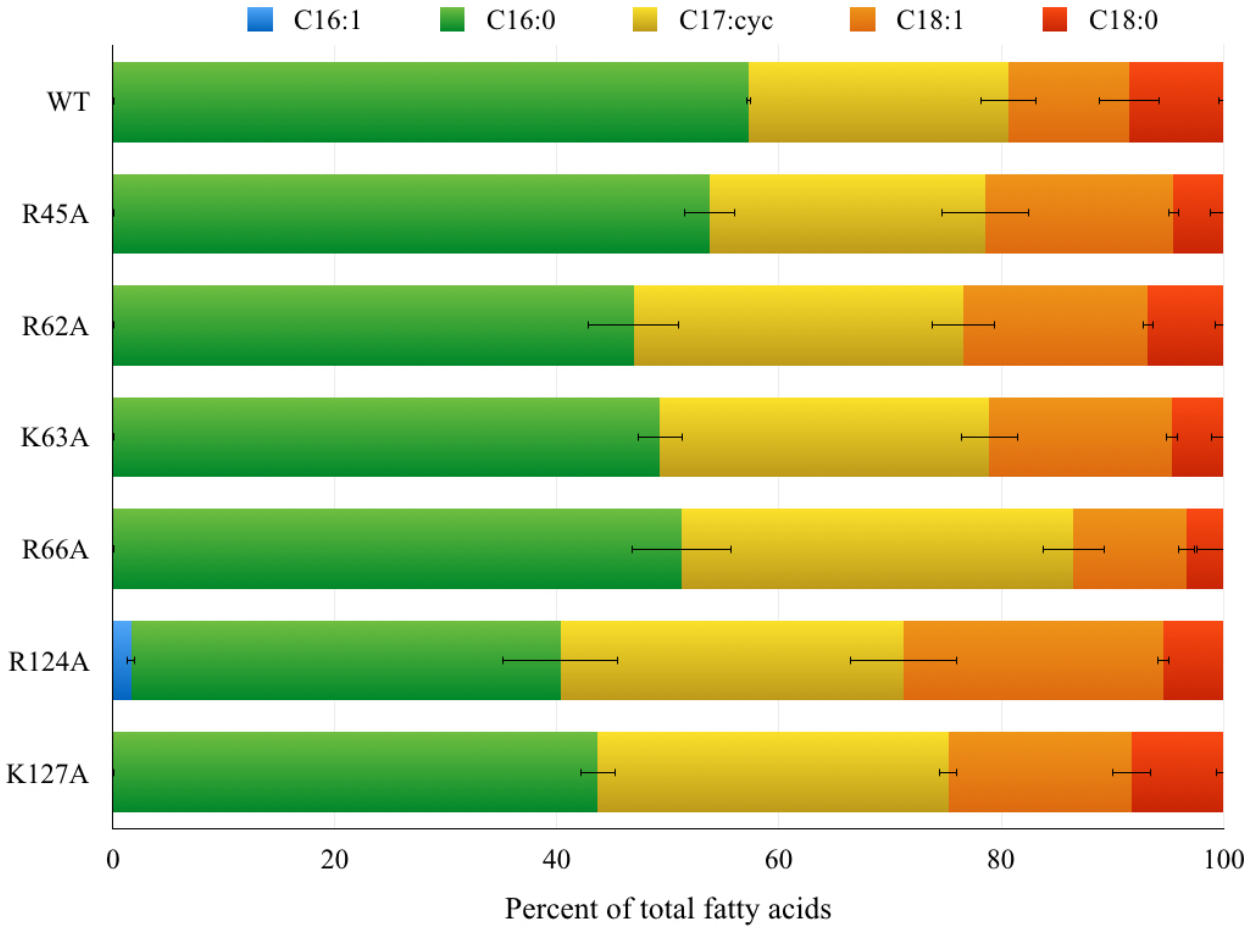

Figure 2-15 Product profile of the *E. coli* type II FAS bearing FabB surface mutations. The products are denoted by chain length and unsaturation level. All experiments were run in triplicate.

chromatography coupled with mass spectrometry (GC-MS). Surprisingly, instead of the expected result of less unsaturation and shorter chain lengths, the opposite result was observed: to differing degrees, most of the mutations actually resulted in an increase in unsaturated fatty acid prodution and an increase in chain length (Fig 2-15). Most notably, mutant R124A had a large increase in the production of the C18:1 product relative to wild type. In fact, this mutant even forced a small amount of production of a C16:1 product that is typically not produced at all under these conditions. Additionally, the R124A mutant facilitated much greater production of 18-carbon products relative to wild type.

When taken together, the results from the crosslinking efficiency and lipid profile experiments suggest that whatever the FabB mutants lost in affinity for AcpP, they more than made up for with increased dissociation and enzyme turnover. Lower affinity cuts both ways: though it may slow the association process, it could also help increase the rate of the dissociation step once the enzymatic reaction has taken place by removing a salt bridge that would otherwise need to be broken. Thus, instead of lowering the effective concentration of FabB, the mutations seem to have actually increased the effective concentration of FabB by increasing the amount of free, unbound FabB present at any given time. These results allow for the proposal of a new hypothesis that contrasts with the conventional wisdom in the community: that the rate limiting step is actually the dissociation step and enzyme turnover as opposed to protein association. Further kinetic experiments can be conducted to test this hypothesis.

2.4 Conclusion

In conclusion, the first crystal structure of an ACP-KS complex was solved through the use of mechanism-based crosslinking. This complex between AcpP and FabB from the *E. coli* type II

FAS provided the first structural characterization of the interface between a type II ACP and a KS, allowed for the identification of the PPT binding site residues within the FabB active site, and provided the first opportunity to compare ACP interations with different partners. Sequence comparisons with other related ACPs and KSs suggest that ACP-partner interactions may be specific to each system and not conserved across all pathways. The information gained from the AcpP-FabB crystal structure was then used to generate FabB mutants that altered the lipid profile of *E. coli* cultures. This work represents the first known example of specifically targeting protein-protein interactions to alter the product profile of a metabolic pathway.

2.5 Materials and methods

2.5.1 Protein expression and purification

pET28b 6x His-tagged constructs containing the genes for *E. coli* AcpP and FabB were separately transformed into *E. coli* BL21 (DE3) cells by heat shocking at 42 °C for 45 seconds and plated onto LB agar plates supplemented with 50 μg/mL kanamycin. Colonies were transferred to a 10 mL LB starter culture supplemented with 50 μg/mL kanamycin and shaken overnight at 37 °C. The starter culture was then transferred to 1 L of LB supplemented with 50 μg/mL kanamycin and shaken at 37 °C. Once the OD_{600} reached 0.6, expression was induced by addition of 1 mM IPTG, and the cells were shaken overnight at 18 °C. Cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 10 mM imidazole, 10 % glycerol), and flash frozen in liquid nitrogen for storage at -80 °C.

The resuspended cells were lysed by sonication, and the lysate was centrifuged at 21,000 rcf for 60 minutes to remove cellular debris. The lysate was then incubated with 5 mL of Ni-NTA resin (ThermoFisher Scientific) for 1 hour, and the resin was washed with lysis buffer to remove

unbound protein. The proteins of interest were eluted in fractions using an imidazole gradient. SDS-PAGE was used to analyze the fractions, and fractions containing the protein of interest were combined and dialyzed into dialysis buffer (25 mM Tris pH 7.5, 100 mM NaCl, 1 mM DTT, 5% glycerol). AcpP was further purified using a HiTrap Q FF anion exchange column (GE Healthcare Lifesciences).

2.5.2 Crosslinking, complex purification, and crystallization

The phosphopantetheine (PPT) prosthetic group was removed from AcpP by incubation with MBP-tagged AcpH, including 12.5 mM $MgCl_2$ and 2.5 mM $MnCl_2$. After removing the AcpH with amylose resin (New England Biolabs), *apo* AcpP was chemoenzymatically loaded with the chloracryl pantetheine crosslinker **2** to form *crypto* AcpP using CoaA, CoaD, CoaE, Sfp, 200 mM ATP, 10 mM MgCl, and a 1.5x molar excess of the crosslinker[11]. The loading was confirmed using MALDI-TOF mass spectrometry. AcpP was purified away from the loading enzymes using a second HiTrap Q FF column.

100 μM FabB was incubated with 200 μM *crypto* AcpP in 25 mM Tris pH 7.5, 100 mM NaCl, 5 mM DTT, and 5% glycerol overnight at 37 °C. The complex was then purified using a Superdex 200 (GE Healthcare Lifesciences) size exclusion column, concentrated to 6 mg/mL, and flash frozen for storage. The complex was crystallized at room temperature in 0.1 M sodium acetate pH 5.4, 0.2 M ammonium acetate, and 20% PEG 4000 using the sitting drop vapor diffusion method. The crystals were flash frozen in liquid nitrogen for storage. Diffraction data were collected using beamline 8.2.1 at the Advanced Light Source synchrotron facility and processed using Mosflm[12]. The structure was solved by molecular replacement using a FabB structure (PDB: 2VB9) as a search model and refined using the Phenix suite of programs[13]. All protein structure

figures were generated using PyMol. The diffraction data collection and refinement statistics are listed in table 2-1.

2.5.3 FabB mutagenesis and measurement of relative crosslinking efficiency

FabB mutant constructs were generated using standard polymerase chain reaction methods. The mutants were expressed and purified as described above for the wild type enzyme. The relative crosslinking efficiency was measured by incubating the mutant FabB with wild type AcpP for 1 hour at 37 °C. Aliquots were taken at various time points throughout the reaction and analyzed by SDS-PAGE. The intensity of the crosslinked and uncrosslinked bands were determined using the software ImageJ. These intensities were used to calculate the relative rates of the crosslinking reaction.

2.5.4 FabB mutant expression and fatty acid extraction/analysis

Mutant constructs were transformed into BL21 (DE3) cells by heat shocking at 42 °C for 45 seconds followed by shaking at 37 °C for 45 minutes. Cells were then plated on LB Agar plates supplemented with 50 μg / μL kanamycin and incubated overnight at 37 °C. Colonies were then transferred to 20 mL of LB media supplemented with 1 mM IPTG and shaken at 37 °C until stationary phase was reached. The cultures were then centrifuged, and the pellet was resuspended in 2.5 mL of water. Pentadecanoic acid was added as an internal standard (0.5 mg), and the cultures were then extracted with 2.5 mL of 1:1 chloroform/methanol with 2% acetic acid. The extracts were dried and resuspended in 1 M HCl in methanol. The mixture was refluxed for 2 hours followed by extraction with chloroform. The derivatized extracts were subjected to GC-MS

analysis (50 °C initial temperature, 10 °C increase / minute, 290 °C final temperature). Standards were used to confirm retention times and fragmentation profiles.

References

1. White, S. W.; Zheng, J.; Zhang, Y. M.; Rock, C. O. The structural biology of type II fatty acid synthases. *Annual Reviews in Biochemistry* **2005,** *74,* 791.
2. Yu, X.; Liu, T.; Zhu, F.; Khosla, C. In vitro reconstitution and steady-state analysis of the fatty acid synthase from Escherichia coli. *Proceedings of the National Academy of Sciences* **2011,** *108* (46), 18643.
3. Wright, H. T.; Reynolds, K. A. Antibacterial targets in fatty acid biosynthesis. *Current Opinions in Microbiology* **2007,** *10,* 447.
4. Bruegger, J.; Haushalter, R. W.; Vagstad, A. L.; Shakya, G.; Mih, N.; Townsend, C. A.; Burkart, M. D.; Tsai, S. C. Probing the selectivity and protein-protein interactions of a nonreducing fungal polyketide synthase using mechanism-based crosslinkers. *Chemistry and Biology* **2013,** *20* (9), 1135.
5. Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Marwick, P. R.; Bruegger, J.; Caldera-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, F.; O'dowd, B.; McCammon, J. A.; Opella, S. J.; Tsai, S. C.; Burkart, M. D. Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* **2014,** *505,* 427.
6. Worthington, A. S.; Rivera, H.; Torpey, J. W.; Alexander, M. D.; Burkart, M. D. Mechanism-based protein cross-linking probes to investigate carrier protein-mediated biosynthesis. *ACS Chemical Biology* **2006,** *1* (11), 687.
7. Price, A. C.; Choi, K. H.; Heath, R. J.; Li, Z.; White, S. W.; Rock, C. O. Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism. *Journal of Biological Chemistry* **2001,** *276* (9), 6551.
8. Roujeinikova, A.; Simon, W. J.; Gilroy, J.; Rice, D. W.; Rafferty, J. B.; Slabas, A. R. Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. *Journal of Molecular Biology* **2007,** *365,* 135.
9. Xiao, X.; Yu, X.; Khosla, C. Metabolic flux between unsaturated and saturated fatty acids is controlled by the FabA:FabB ratio in the full reconstituted fatty acid biosynthetic pathway of Escherichia coli. *Biochemistry* **2013,** *52,* 8304.

10. Liu, X.; Hicks, W. M.; Silver, P. A.; Way, J. C. Engineering acyl carrier protein to enhance production of shortened fatty acids. *Biotechnology for Biofuels* **2016,** *9* (24).

11. Shakya, G.; Rivera, H.; Lee, D. J.; Jaremko, M. J.; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdziowski, E. R.; Wong, F.; Tsai, S. C.; Burkart, M. D. Modeling linear and cyclic PKS intermediates through atom replacement. *Journal of the American Chemical Society* **2014,** *136,* 16792-16799.

12. Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. IMosflm: a new graphical interface for diffraction image processing with MOSFLM. *Acta Crystallographica Section D* **2011,** *67,* 271.

13. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kuntsleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **2010,** *66,* 213.

CHAPTER 3

Oxetane-Based Polyketide Surrogates to Probe Polyketide Synthase Enzyme-Substrate Complexes

3.1 Abstract

Type II polyketides are a large class of bioactive natural products with a wide range of structures and functions. They are biosynthesized by type II polyketide synthases (PKS), which are characterized by several discrete, mono-functional enzymes that must find each other in solution and form complexes. One of the primary challenges when studying type II PKSs is the extreme reactivity of the enzymatic substrates and intermediates. This has hampered structural characterization of enzyme-substrate complexes, and, as a result, little is known about how these enzymes are able to position their unstable substrates for proper catalysis while displaying high levels of regio- and stereospecificity. This chapter describes the development and application of an oxetane-based PKS substrate mimic that enabled the first structure determination of the acyl-enzyme intermediate of a ketosynthase (KS) in complex with an inert extender unit mimic. This crystal structure enabled a mechanism for the unique activity of the KS enzyme DpsC to be proposed. The effectiveness of this substrate mimic suggests that this novel class of probes could have wide-ranging applications to the greater biosynthetic community.

3.2 Introduction

Type II polyketide natural products are a large and diverse class of secondary metabolites produced by bacteria[1]. Notable examples from the pharmaceutical industry include tetracycline, actinorhodin, and daunorubicin[2]. Polyketides are biosynthesized by a polyketide synthase (PKS), of which there are two primary types. Type II PKSs are characterized by several stand-alone enzymes that form complexes in solution. PKSs have been heavily studied due to their ability to efficiently biosynthesize complex small molecules and their potential to be engineered for custom biosynthesis and biocatalysis[3].

One of the primary challenges associated with investigating type II PKSs is the high reactivity of enzymatic substrates and intermediates. The poly-*beta*-ketone generated by most type II PKSs is highly susceptible to spontaneous, non-specific cyclization, which has made structural studies of PKS-substrate complexes extremely difficult[4] (Fig 3-1). Without this structural knowledge, attempts to engineer substrate specificity have primarily lead to inactive enzymes. Previously, synthetic isoxazole-based polyketide isosteres have been applied to the

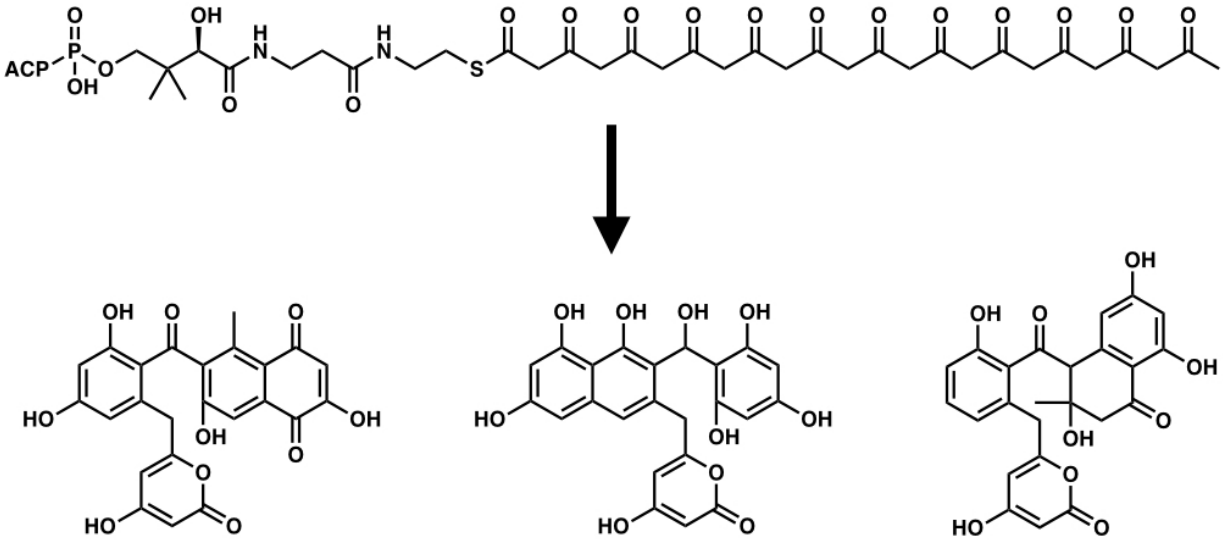

Figure 3-1 Highly reactive PKS intermediates are prone to spontaneous and non-specific intramolecular aldol condensations. This makes their characterization a unique challenge.

characterization of the interior pocket of an acyl carrier protein[5]. Given the substantial structural differences between these first-generation chemical probes and the natural poly-*beta*-ketonic substrates, a new class of probes were sought which were more closely related to the natural substrates.

The use of oxetane rings has undergone a renaissance in medicinal chemistry over the past decade through the work of Carreira, Müller, and co-workers[6]. One of its primary uses is as a stable replacement for gem-dimethyl groups that are readily degraded *in vivo*[7]. However, it is also frequently used as a stable carbonyl isostere. Though it is slightly larger in volume than a carbonyl, theoretical studies have predicted that its electronic properties as well as the positioning of its lone pairs are quite similar to that of a carbonyl[8]. The use of oxetanes as stable replacements for carbonyls has never been explored within the context of polyketide biosynthesis. Thus, it was proposed by our co-worker Chris Vanderwal that oxetanes should be able to replace carbonyl groups in the unstable poly-*beta*-ketonic intermediates produced by PKSs (Fig 3-2). This

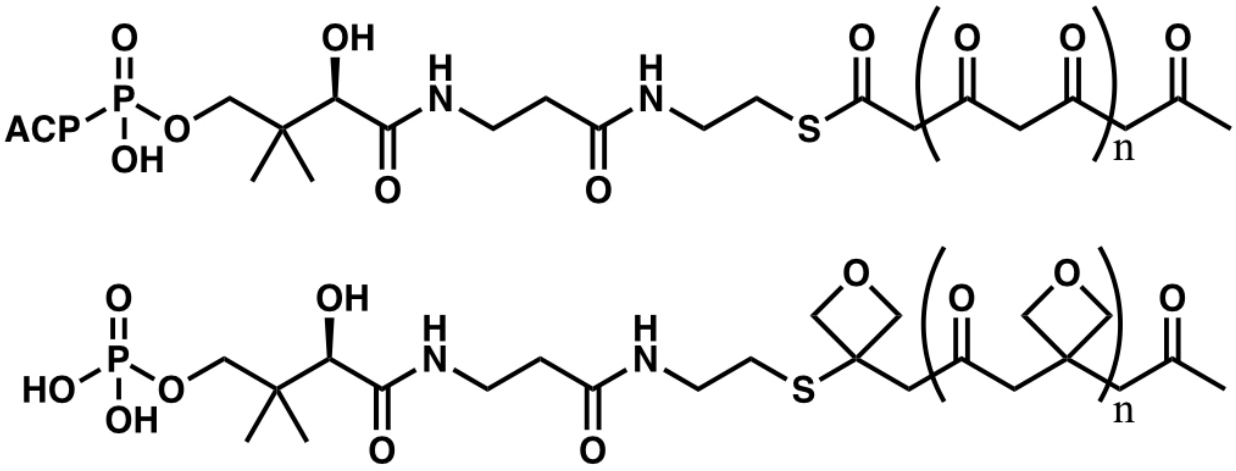

Figure 3-2 General example of proposed oxetane-based poly-β-ketone isosteric probes. A natural intermediate is shown at the top, and a general probe containing alternating carbonyls and oxetanes is shown below.

replacement should, in theory, yield relatively inert probes that could be used for co-crystallization with PKS enzymes.

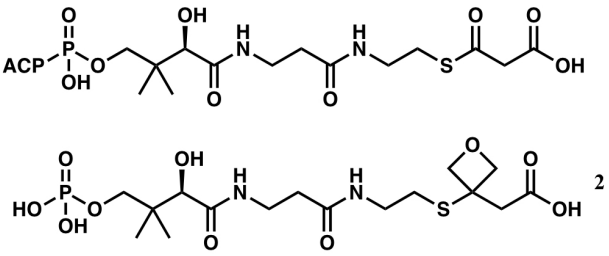

Figure 3-3 Oxetane-based malonyl-pantetheine analogue developed in the current work. The natural malonate extender unit is shown at the top, and probe **2** is shown at the bottom.

In this work, the development and application of oxetane-based PKS substrate analogue **2** is presented (Fig 3-3), and its applicability is demonstrated by co-crystallizing it with the acyl-enzyme intermediate of the protein DpsC from the daunorubicin type II PKS from *Streptomyces peucetius*[9,10]. DpsC is a unique enzyme with dual acyltransferase (AT) and ketosynthase III (KS III) activity[10,11] (Fig 3-4). Interestingly, DpsC does not contain any of the conserved active site residues known to be important for KS III activity, and the structural basis for this unique type of activity is unclear due to the lack of high resolution substrate-DpsC co-crystal structures. In particular, the decarboxylation and Claisen condensation steps of the enzymatic reaction are poorly understood (Fig 3-4). The results

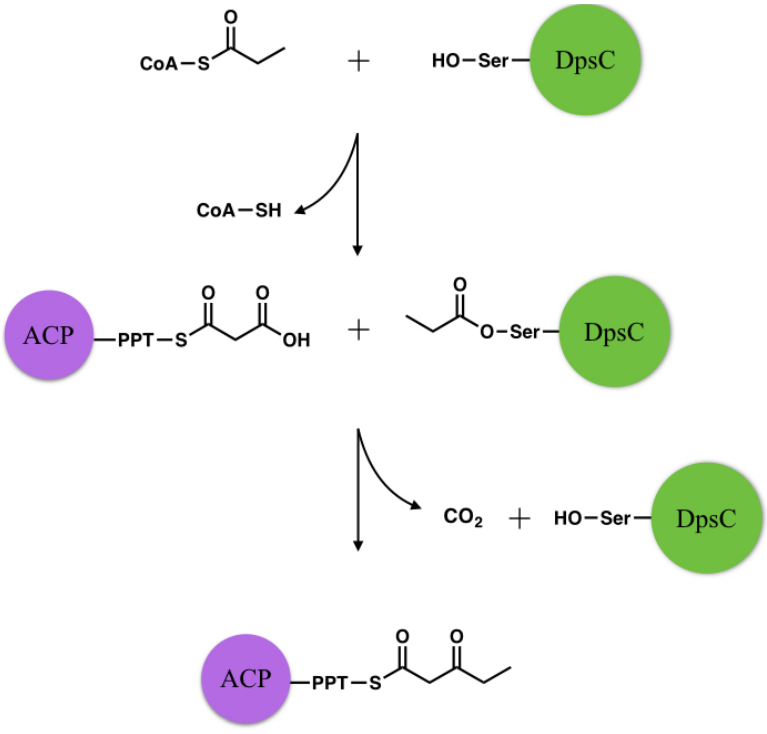

Figure 3-4 Overview of DpsC's KS III activity. In the first step, a propionyl group from propionyl-CoA is transferred to the active site serine nucleophile of DpsC. Then, an ACP loaded with a malonate extender unit binds to DpsC, and DpsC facilitates a decarboxylative Claisen condensation to afford the *beta*-ketoacyl-ACP product.

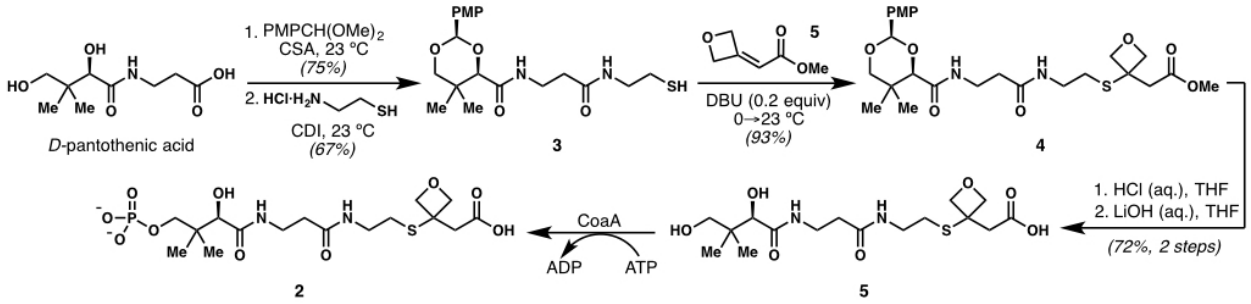

Figure 3-5 Synthesis of oxetane-based malonyl-PPT probe **2**. Figure courtesy of Bryan Ellis, Alex White, and Chris Vanderwal.

presented here allowed us to glean insights into the mechanism of substrate binding and activation prior to the condensation reaction catalyzed by this novel, dual-functioning enzyme.

*Author's note: all the data presented here (excluding the chemical synthesis of **5)** is entirely the work of the author of this dissertation, though it is important to acknowledge that the initial structure determination of apo and propionyl-DpsC was performed by David Jackson, a former graduate student in the Tsai group[12].

3.3 Results and disussion

3.3.1 Probe preparation

Compound **5** was synthesized and kindly provided by our co-workers Bryan Ellis, Alex White, and Chris Vanderwal at the University of California, Irvine. (Figure 3-5). In order to more closely mimic the natural substrate, **5** was chemoenzymatically phosphorylated using recombinant pantetheine kinase (CoaA) and ATP. The reaction was confirmed using ESI-MS, and the resulting product **2** was purified by HPLC (data not shown).

3.3.2 DpsC expression, purification, and crystallization

DpsC was expressed in *E. coli* as a His_6-tagged construct and successfully purified using immobilized metal affinity chromatography (IMAC). The sample was further purified using anion

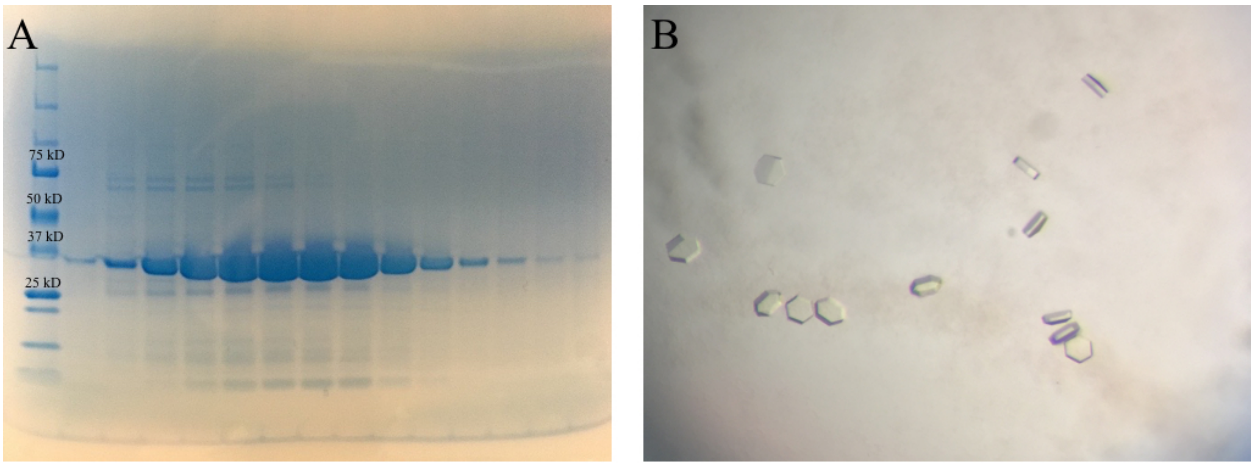

Figure 3-6 Purification and crystallization of DpsC. (a) SDS-PAGE analysis of size exclusion fractions. (b) DpsC crystals obtained after high-throughput screening and condition optimization.

exchange and size exclusion chromatography (Fig 3-6A). DpsC was then concentrated to 6 mg/mL and screened for crystallization using high-throughput methods, and, after optimization, the optimal conditions for crystallization were found to be 0.1 M MES pH 6.5, 0.1 M imidazole, 0.06 M $MgCl_2$, 15% PEG 4000, and 30% glycerol (Fig 3-6B).

In order to provide a source of anomalous diffraction, the crystals were soaked in a solution containing $GdCl_2$ prior to flash cooling in liquid nitrogen. The diffraction pattern of the crystals was then measured, and the data were processed to a resolution of 2.95 Å. The phases were obtained using the anomalous diffraction at a single wavelength (SAD), namely the L-I edge of gadolinium (Table 3-1).

3.3.3 Structural analysis of ligand-free DpsC

The DpsC crystal structure revealed a thiolase fold reminiscent of other KS IIIs (Fig 3-7A). DpsC crystallized as a homodimer, and each of the monomers are related by a two-fold axis of non-crystallographic symmetry. Two DpsC molecules (one dimer) were observed in the asymmetric unit.

	DpsC	propionyl-DpsC	propionyl-DpsC + **2**
Data collection			
Wavelength (Å)	1.48	1.0	1.0
Total reflections	34456 (3350)	71074 (6942)	85634 (7822)
Unique reflections	30794 (1674)	35537 (3471)	43341 (4164)
Space group	P 6_5 2 2	P 6_5 2 2	P 6_5 2 2
Cell dimensions			
a, *b*, *c* (Å)	91.2, 91.2, 314.6	91.1, 91.1, 316.1	91.6, 91.6, 316.1
α, β, γ (°)	90, 90, 120	90, 90, 120	90, 90, 120
Resolution (Å)	76.6 – 2.95	70.6 – 2.30	70.9 – 2.15
R_{merge}	0.177 (0.786)	0.024 (0.157)	0.021 (0.134)
R_{meas}	0.250 (0.111)	0.034 (.222)	0.030 (0.190)
$I/\sigma(I)$	11.32 (2.80)	20.76 (4.08)	16.26 (4.26)
$CC_{1/2}$	0.999 (0.999)	0.999 (0.935)	0.999 (0.978)
CC*	0.999 (0.999)	1.0 (0.983)	1.0 (0.994)
Completeness (%)	98.6 (95.7)	100 (100)	99 (95)
Redundancy	11 (11)	2.0 (2.0)	2.0 (1.9)
Wilson B-factor	47.45	26.44	35.15
Refinement			
Resolution (Å)	76.6 – 2.95	70.6 – 2.30	70.9 – 2.15
No. reflections	30794 (1674)	35537 (3471)	43233 (4160)
R_{work}	0.224	0.181	0.179
R_{free}	0.268	0.219	0.209
No. atoms			
Protein	5067	5115	5083
Ligands	14	20	60
B factors			
Protein	41.16	28.97	39.56
Ligands	94.05	25.89	95.06
Water	32.25	33.36	44.86
Ramachandran			
Favored (%)	92.5	95	96
Allowed (%)	6	3.5	3.5
Outliers (%)	1.5	1.5	0.5
R.m.s. deviations			
Bond lengths (Å)	0.009	0.009	0.02
Bond angles (°)	1.59	1.52	1.36

Table 3-1 Summary of X-ray crystallographic data collection and refinement statistics.

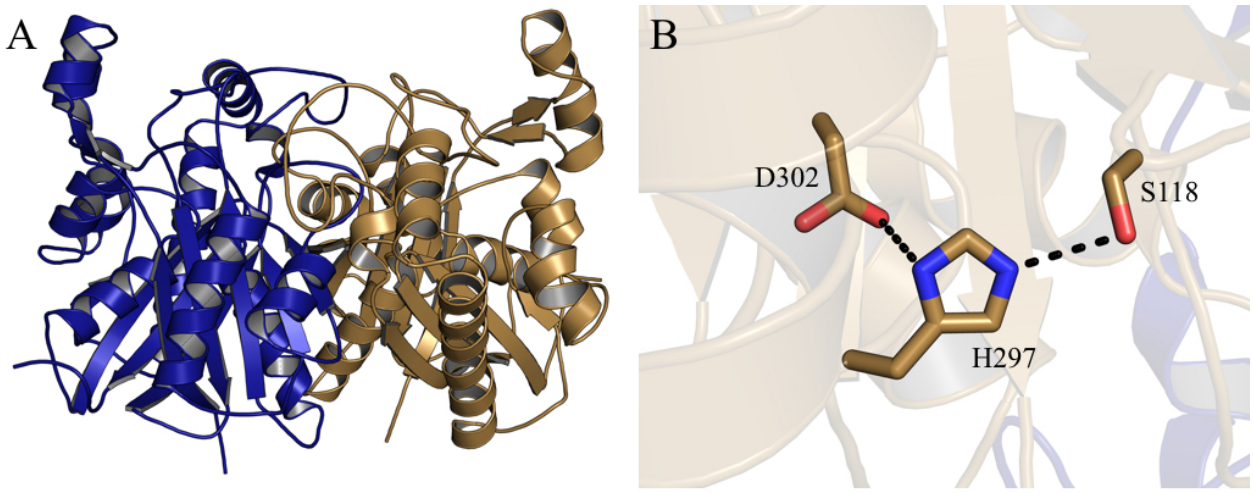

Figure 3-7 Ligand-free crystal structure of DpsC. (a) overall dimeric structure of DpsC determined using experimental phases. (b) active site catalytic triad of DpsC.

As expected from the sequence, DpsC contains an unusual active site architecture relative to other KSs. For example, DpsC contains an active site serine in place of the conserved cysteine found in all other known KSs (Fig 3-7B). Additionally, DpsC uses an Asp-His-Ser catalytic triad to enhance the nucleophilicity of S118. In a canonical KS, the cysteine nucleophile is thought to be activated by its location at the C-term end of a long helix through the helix-dipole effect[13]; however, this would not be sufficient to activate a serine side chain, necessitating the use of the observed catalytic triad.

A structural comparison between DpsC and FabH[14], the prototypical KS III from the *E. coli* type II fatty acid synthase (FAS), underscores the differences between DpsC and the other members of its KS III family (Fig 3-8). For example, the D-H-S catalytic triad found in DpsC is not present in FabH (Fig 3-8D). The structurally corresponding residues in FabH are S-N-C, none of which are in range to form hydrogen bonds with one another; additionally, the two key catalytic residues in FabH, H244 and N274, are not conserved in DpsC[14] (Fig 3-8E). These residues are known to form an oxyanion hole to stabilize the negative charge that builds up on the thioester carbonyl during the course of the decarboxylative Claisen condensation reaction. The

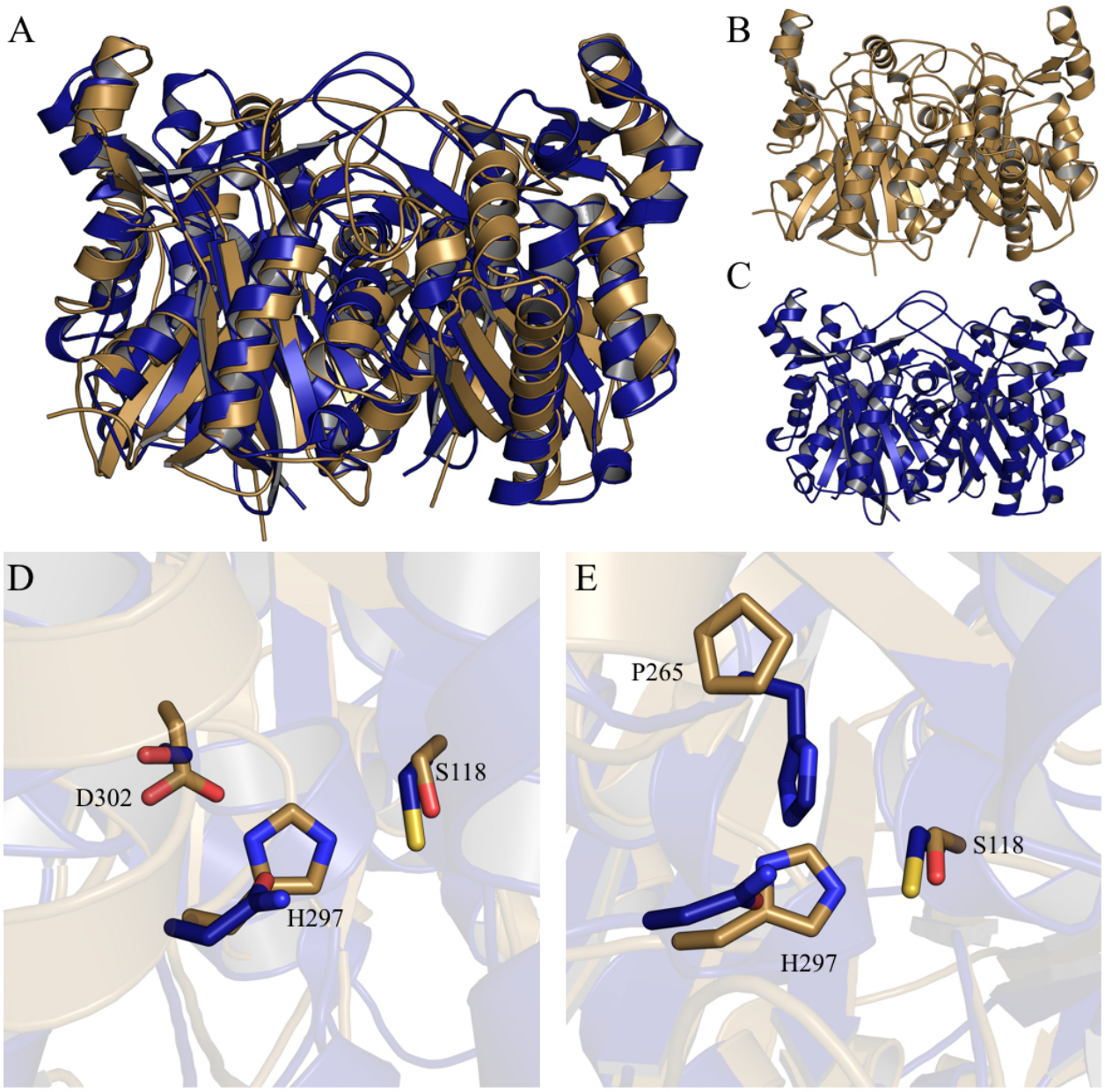

Figure 3-8 Structural comparison between DpsC and FabH. (a) overlay of DpsC (gold) and FabH (blue). The overall R.M.S.D. between these two structures is estimated to be 2.9 Å (b) DpsC structure. (c) FabH structure. (d) DpsC catalytic triad (gold) overlaid with corresponding residues in FabH (blue). (e) FabH oxyanion hole (blue) overlaid with corresponding residues in DpsC (gold).

corresponding residues in DpsC are P265 and H297. The side chain of P265 has no functionality to stabilize a negative charge, and H297 is already locked in the tightly H-bonded catalytic triad.

Given the nature of its sequence, it is not surprising that the active site architecture of DpsC is much more similar to an AT than to a KS; the activated serine side chain and the apparent

absence of an oxyanion hole are common features of AT domains. Thus, while this ligand-free structure provided some insight into the residues responsible for AT activity, the structural basis for DpsC's additional KS activity remains unclear from this structure.

3.3.4 Structural analysis of the propionyl-DpsC acyl enzyme intermediate

In order to determine the structure of the propionyl-DpsC acyl enzyme intermediate (Fig 3-4), native DpsC crystals were soaked in a solution containing propionyl-CoA before flash cooling in liquid nitrogen. The diffraction pattern of the crystals was then measured, and the data were processed to a resolution of 2.30 Å (Table 3-1). The phases were obtained by molecular replacement using the experimentally phased ligand-free structure of DpsC described above.

The overall structure of the propionyl-DpsC structure is nearly identical to the ligand-free structure (Fig 3-9). In this structure, clear electron density corresponding to the *O*-propionyl serine side chain could be observed (data not shown). This side chain remains locked into a tight hydrogen bond network with the side chains of H297 and D302.

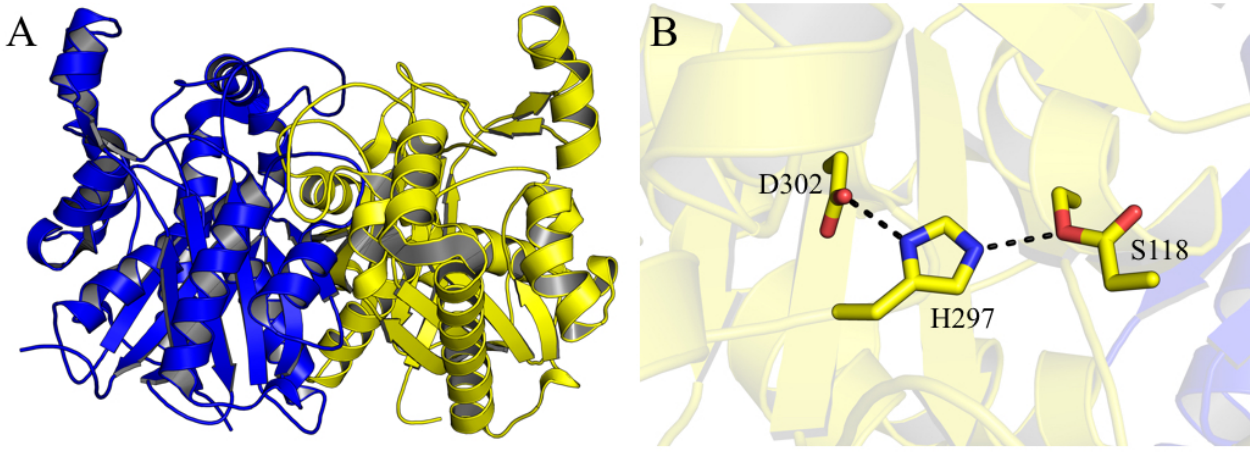

Figure 3-9 Crystal structure of the propionyl-DpsC acyl-enzyme intermediate. (a) overall dimeric structure of prop-DpsC. (b) catalytic triad of DpsC, including the *O*-propionyl-serine side chain.

In this acyl-enzyme intermediate structure, the ester carbonyl of the *O*-propionyl serine side chain forms a hydrogen bond with the backbone amide proton of A329 (Fig 3-10). This interaction likely serves to stabilize the negative charge that accumulates on the tetrahedral intermediate during the first acyl group transfer step of the AT and KS reactions. While this structure revealed further insight into the structural basis for AT activity as well as the first glimpse into the basis for KS activity, the residues that play a role in carbon-carbon bond formation remain unclear to this point without a co-crystal structure with an extender unit substrate or analogue.

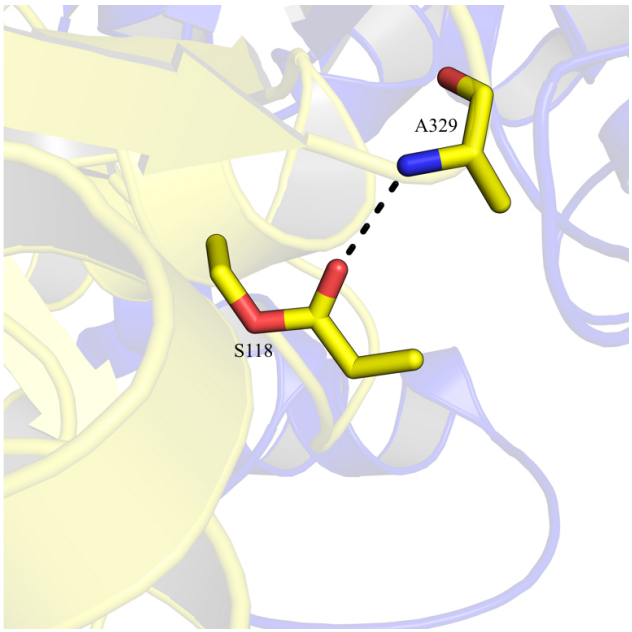

Figure 3-10 The *O*-propionyl serine side chain forms a hydrogen bond with the backbone amide proton of A329

3.3.5 Structural analysis prop-DpsC in complex with probe 2

In order to determine the structural basis for malonyl-ACP binding and decarboxylative Claisen condensation by DpsC, native crystals of DpsC were again soaked in a solution containing propionyl-CoA followed by soaking in a separate solution containing **2** prior to flash cooling in liquid nitrogen. The diffraction pattern of these crystals was then measured, and the data were

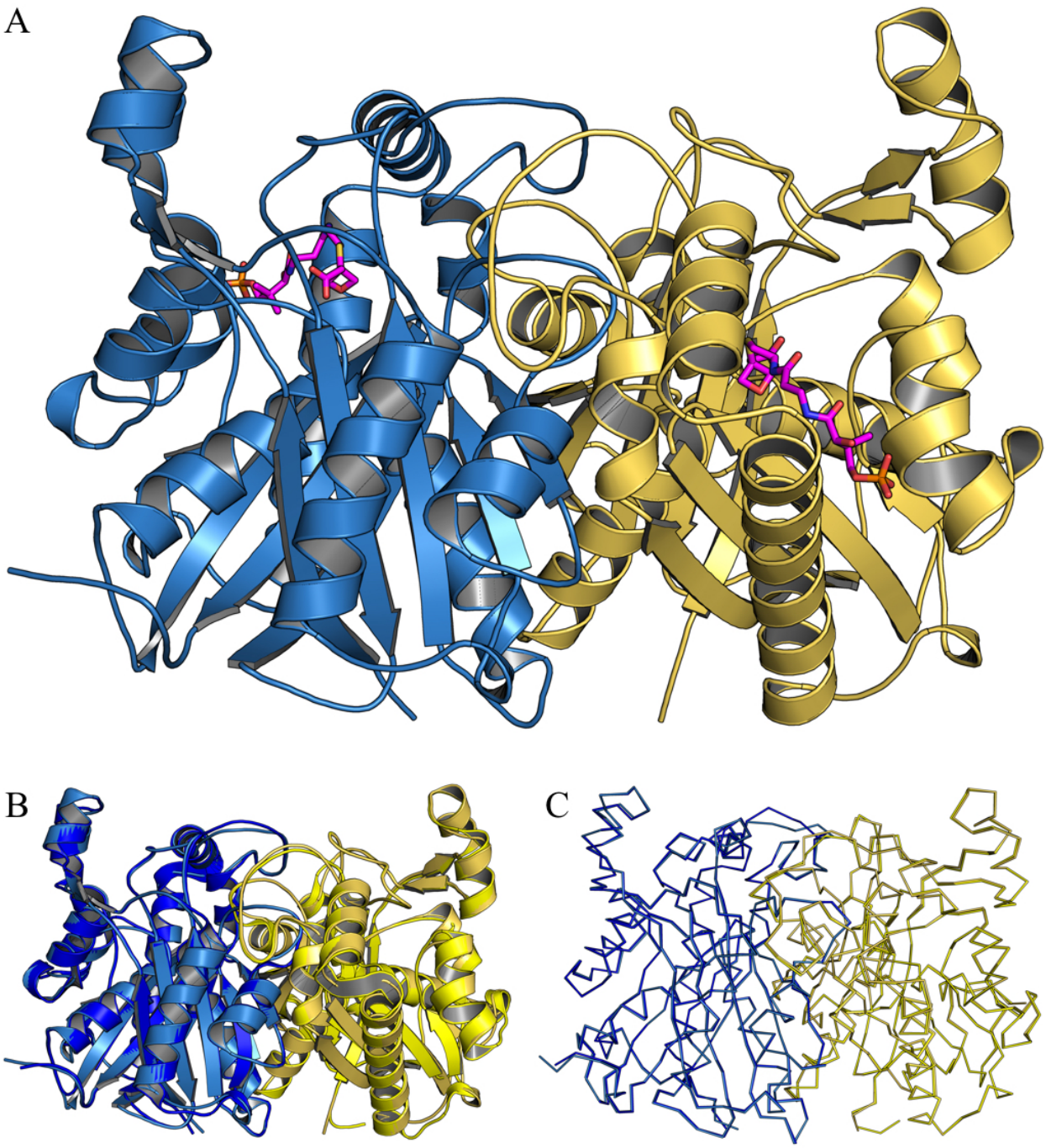

Figure 3-11 Crystal structure of the propionyl-DpsC acyl-enzyme intermediate in complex with malonyl-PPT mimic **2**. (a) The overall dimeric structure of propionyl-DpsC with **2** shown in magenta sticks. (b) Overlay of propionyl-DpsC and propionyl-DpsC + **2** shown in cartoon representation. The overall R.M.S.D. between these structures is estimated to be 0.226 Å. (c) Overlay of propionyl-DpsC and propionyl-DpsC + **2** shown in ribbon representation.

processed to a resolution of 2.15 Å. The phases were obtained by molecular replacement using the prop-DpsC structure described above.

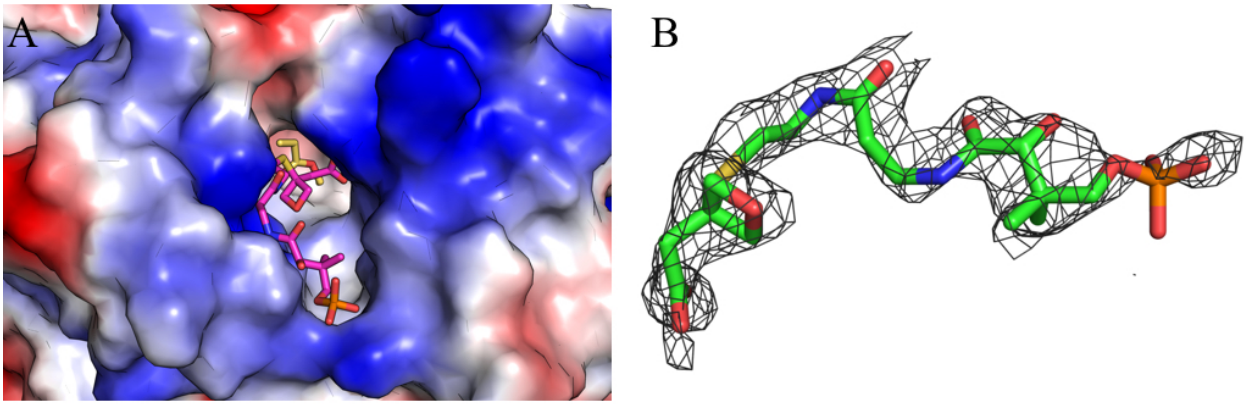

Figure 3-12 Probe **2** binds to the active site of prop-DpsC. (a) **2** binds in the active site pocket of DpsC, shown in an electrostatic surface charge representation. (b) The $2F_o$-F_c electron density map for **2** contoured at 0.8 σ.

Upon structure determination, well-defined electron density for a propionylated S118 side chain and for probe **2** were observed in the active site (Fig 3-12). The occupancy values for some of the atoms in **2** were determined to be relatively low, but the average occupancy of all atoms was calculated to be 79% in one monomer and 71% in the other monomer. This allowed us to unambiguously model every atom of **2** into the electron density.

In order to predict whether probe **2** and the natural malonyl-PPT extender unit would display similar properties within the active site of DpsC, molecular dynamics (MD) and molecular mechanics Poisson-Boltzmann Surface Area (MM-PBSA) calculations were conducted by our co-workers Vy Duong and Ray Luo at the University of California, Irvine. Simulations were conducted on both the co-crystal structure with **2** as well as a structure in which malonyl-PPT was modeled in place of **2**. The results predict that probe **2** and the natural substrate have very similar free energies of binding to DpsC (Table 3-2). Additionally, a comparison of the R.M.S.D. and R.M.S.F. values between the two MD trajectories indicate the probe **2** behaves similarly in the active site to malonyl-PPT

Ligand	Average ΔG	Std. dev.
Probe **2**	-15.42	4.57
Malonyl-PPT	-15.31	3.81

Table 3-2 Calculated binding free energies of probe **2** and malonyl-PPT. Data courtesy of Vy Duong and Ray Luo

in terms of dynamics and stability (data not shown). Taken together, these results further support the proposal that oxetane-based probes could be useful as polyketide surrogates.

This model provided new insights into the positioning of substrates immediately prior to catalysis (Fig 3-13). As expected, the D-H-S catalytic triad again forms a hydrogen bond network in this structure that results in an activated S118 nucleophile. The resulting *O*-propionyl serine side chain is in close proximity to the site of probe **2** binding. In particular, the two carbon atoms that would be joined in the reaction were observed to be at a reasonable distance (2.9 Å) and alignment for Claisen condensation (Fig 3-13A). Additionally, the carboxylate of **2** formed a charge-charge interaction with R271 as well as a hydrogen bond with the side chain of T163 (Fig 3-13B). The terminal phosphate formed a charge-charge interaction with K279, consistent with many other PKS enzymes that use positively charged surface residues to position the phosphate moiety of the phosphopantetheine cofactor[5]; S238 also formed a hydrogen bond with the phosphate.

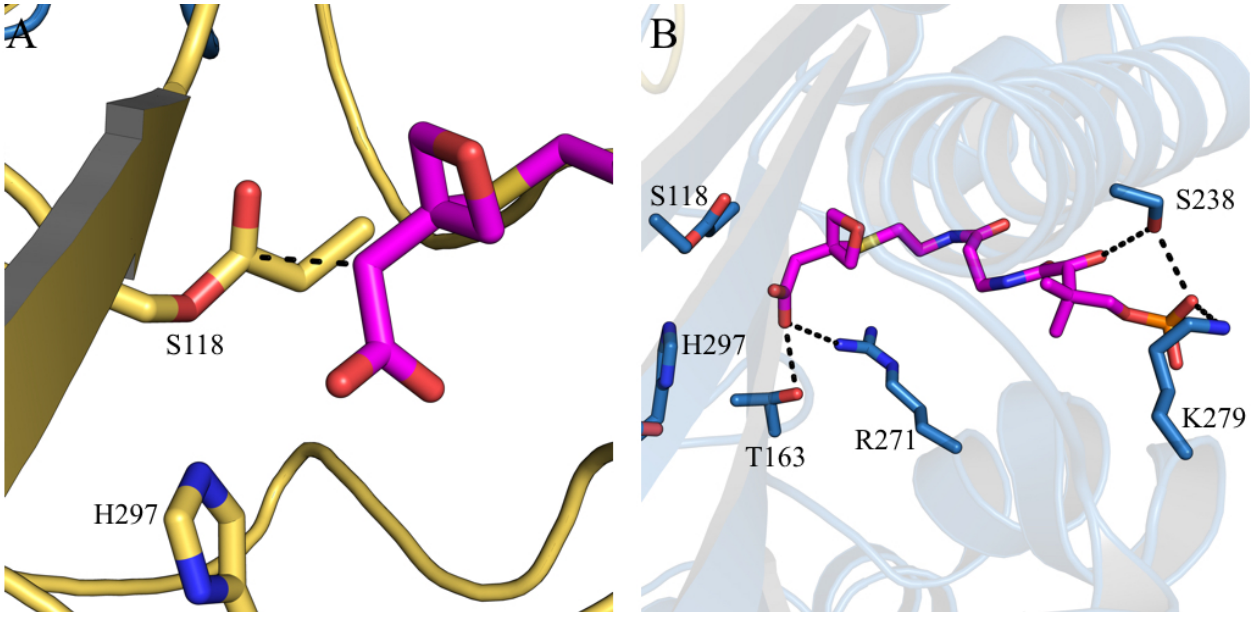

Figure 3-13 Interactions between probe **2** and the DpsC active site. (a) The two carbon atoms that would be joined in the reaction are indicated by black dashes. (b) hydrogen bonds and charge-charge interactions formed between **2** and DpsC.

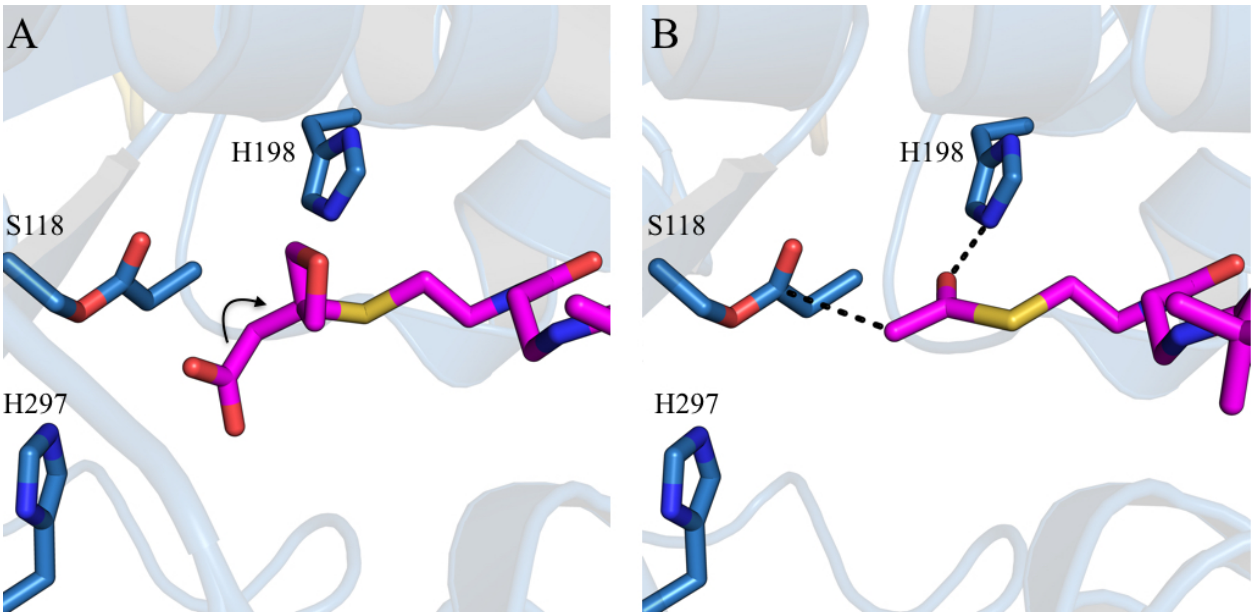

Figure 3-14 H198 may serve as an oxyanion hole. (a) the location of H198 within the DpsC active site is shown. The black arrow indicates the electrons transferred upon decarboxylation. (b) molecular modeling and docking simulations predict that the enolate formed as a result of decarboxylation could be stabilized by the side chain of H198.

Curiously, the oxetane oxygen atom did not appear to orient itself into a positively charged oxyanion hole within the DpsC active site. As previously mentioned, the conserved oxyanion hole residues that canonical KSIIIs use to stabilize the buildup of negative charge on the thioester carbonyl during decarboxylation are not present in DpsC; specifically, residues H244 and N274 in FabH are known to be essential for decarboxylation[14], but these residues are not conserved in DpsC. No other such oxyanion hole is implicated in the co-crystal structure. However, the side chain of H198 is within the active site and could potentially serve this purpose (Fig 3-14A). The oxetane oxygen atom is pointing away from this side chain in the crystal structure, but this is likely due to the lack of negative charge on probe **2**. However, as the substrate undergoes decarboxylation and the enolate intermediate is formed, the oxygen could easily reorient itself towards H198 through a simple bond rotation in order to stabilize the newly-formed negative charge. Molecular modeling and docking simulations further support this hypothesis (Fig 3-14B).

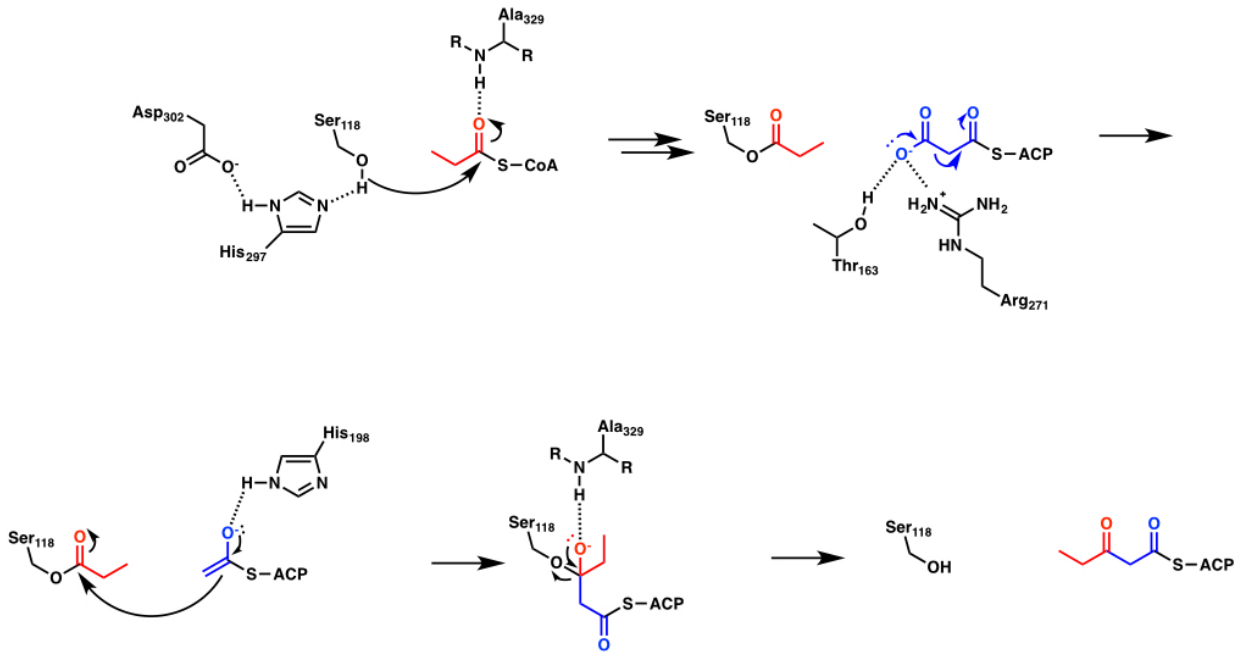

Figure 3-15 Overview of the proposed mechanism of KS activity by DpsC. The red atoms indicate atoms that derived from propionyl-CoA, and the blue atoms indicated atoms that derived from malonyl-ACP

This co-crystal structure enabled the first mechanism for DpsC's unique KS activity to be proposed (Figure 3-15). First, propionyl-CoA binds to the active site of DpsC. Then, the S118 side chain, whose nucleophilicity has been enhanced by H297 and D302, becomes propionylated with the help of the backbone amide proton of A329 which stabilizes the negative charge on the tetrahedral intermediate. Next, once the free CoA has dissociated from the active site, malonyl-ACP binds to T163 and R271. The decarboxylation is then facilitated by the stabilization of the negatively charged enolate via the side chain of H198. Finally, the Claisen condensation occurs, again with the help of A329. Future studies will be conducted to validate the proposed mechanism and the roles of each residue.

3.4 Conclusion

In conclusion, this work represents the first design and application of an oxetane-based polyketide substrate mimic. This probe allowed the only crystal structure of an acyl-KS intermediate in complex with an extender unit mimic to be solved. Using this structure, a mechanism for the unique KS activity of the dual-functioning enzyme DpsC was proposed. Without this probe, this novel snapshot of a KS and all of its substrates and intermediates positioned just prior to catalysis would not have been possible. Now that the use of oxetanes as PKS substrate mimics has been experimentally validated, this class of probes can be expanded to have many potential applications as tools for structural and biochemical studies of PKSs. For example, elongated probes with many oxetane-bearing "ketide" units could be used to study enzymes that act on longer poly-*beta*-ketone chains, and these probes could potentially even be used in feeding experiments to incorporate them into novel "unnatural" natural products.

*Author's note: this chapter is dedicated to my mom, Keri Milligan, who is alive and cancer-free today in-part due to the enzyme DpsC and the natural product it helps to produce (daunorubicin) which she took during her battle with breast cancer nearly 15 years ago.

3.5 Materials and methods

3.5.1 Expression and purification of DpsC

A pET28 expression vector coding for His_6-tagged DpsC was transformed into *E. coli* BL21(DE3) by heatshocking at 42 °C for 45 seconds. The transformed cells were plated on an LB agar medium supplemented with 50 μg/μL kanamycin and incubated at 37 °C for 18 hours. Cells were transferred to a 10 mL starter culture supplemented with 50 μg/μL kanamycin and shaken at 250 rpm for 18 hours at 37 °C. 10 mL of the starter culture was transfered to 1 L of LB

media supplemented with 50 μg/μL kanamycin. The cultures were shaken at 200 rpm at 37 °C until the OD_{600} reached 0.6; the expression was then induced by the addition of 1 mM IPTG, and the cultures were shaken for 18 hours at 18 °C. The cells were centrifuged at 5100 rpm for 10 minutes, resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 10% glycerol), and flash frozen in liquid nitrogen before storage at -80 °C.

The cells were lysed using a microfluidizer, and the lysate was centrifuged at 21,000 rcf for 1 hour to separate from cellular debris. The lysate was applied to a 5 mL HisTrap HP column (GE Healthcare) and eluted using an imidazole gradient via an Akta Purifier FPLC. The fractions were analyzed using SDS-PAGE, and the fractions containing DpsC were combined and concentrated to 10 mg/mL. The protein sample was further purified using a Hi-Trap Q column and a Superdex 200 column (GE Healthcare), and fractions were again analyzed using SDS-PAGE. The selected fractions containing DpsC were concentrated to 6 mg/mL and flash frozen in liquid nitrogen before storage at -80 °C.

3.5.2 Crystallization of DpsC and structure solution using SAD

DpsC was crystallized in a solution containing 0.06 M $MgCl_2$, 0.6 M $CaCl_2$, 0.1 M imidazole pH 7.0, 0.1 M MES pH 6.7, 15% PEG 4000, and 30% glycerol. The crystals were improved through multiple rounds of seeding using a Seed Bead (Hampton Research). The crystals were soaked in a solution containing 1 mM $GdCl_2$ prepared using mother liquor prior to flash cooling in liquid nitrogen.

The diffraction pattern of the crystals was measured at the Advanced Light Source using beamline 8.2.2. The diffraction data were collected at the L-1 edge of gadolinium (1.48 Å). The diffraction images were processed using Mosflm and scaled using Aimless[15]. The structure was

solved using AutoSol and the initial model built using AutoBuild (Phenix). The structure was further refined using the Phenix suite[16]. The diffraction data collection and refinement statistics are listed in Table 3-1.

3.5.3 Crystallization and structure solution of propionyl-DpsC and the probe 2 complex

DpsC was again crystallized in a solution containing 0.06 M $MgCl_2$, 0.6 M $CaCl_2$, 0.1 M imidazole pH 7.0, 0.1 M MES pH 6.7, 15% PEG 4000, and 30% glycerol. These crystals were also improved through multiple rounds of seeding using a Seed Bead (Hampton Research). The crystals were incubated in a 5 mM solution of propionyl-CoA prepared using mother liquor for 18 hours to form the propionyl-DpsC intermediate; in the case of the probe **2** complex, the crystals were then transferred to a drop containing 5 mM **2** for 3 hours, and flash frozen in liquid nitrogen.

The diffraction pattern of the crystals was measured at the Advanced Light Source using beamline 8.2.2. The diffraction images were processed using Mosflm[15] and scaled using Aimless. The structure was solved by molecular replacement using the prop-DpsC structure as a search model. The model was built and refined using the Phenix suite[16].

References

1. Crawford, J. M.; Townsend, C. A. New insights into the formation of fungal aromatic polyketides. *Nature Reviews Microbiology* **2010,** *8,* 879-889.
2. Tsai, S. C.; Ames, B. D. Structural Enzymology of Polyketide Synthases. *Methods in Enzymology* **2009,** *459,* 17-47.
3. Zhang, W.; Li, Y.; Tang, Y. Engineered biosynthesis of bacterial aromatic polyketides in Escherichia coli. *Proceedings of the National Academy of Sciences* **2008,** *105,* 20683-20688.
4. Caldera-Festin, G.; Jackson, D. R.; Barajas, J. F.; Valentic, T. R.; Patel, A. B.; Aguilar, S.; Nguyen, M.; Vo, M.; Khanna, A.; Sasaki, E.; Liu, H. W.; Tsai, S. C. Structural and functional analysis of two di-domain aromatase/cyclases from type II polyketide synthases. *Proceedings of the National Academy of Sciences* **2015,** *112,* E6844-E6851.
5. Shakya, G.; Rivera, H.; Lee, D. J.; Jaremko, M. J.; La Clair, J. J.; Fox, D. T.; Haushalter, R. W.; Schaub, A. J.; Bruegger, J.; Barajas, J. F.; White, A. R.; Kaur, P.; Gwozdziowski, E. R.; Wong, F.; Tsai, S. C.; Burkart, M. D. Modeling linear and cyclic PKS intermediates through atom replacement. *Journal of the American Chemical Society* **2014,** *136,* 16792-16799.
6. Burkhard, J. A.; Wuitschik, G.; Rogers-Evans, M.; Müller, K.; Carreira, E. M. Oxetanes as versatile elementsin drug discovery and synthesis. *Angewandte Chemie International Edition* **2010,** *49,* 9052-9067.
7. Wuitschik, G.; Rogers-Evans, M.; Müller, K.; Fischer, H.; Wagner, B.; Schuler, F.; Polonchuk, L.; Carreira, E. M. Oxetanes as promising modules in drug discovery. *Angewandte Chemie* **2006,** *118,* 7900-7903.
8. Wuitschik, G.; M., R.-E.; Buckl, A.; Bernasconi, M.; Märki, M.; Godel, T.; Fischer, H.; Wagner, B.; Parrilla, I.; Schuler, F.; Schneider, J.; Alker, A.; Schweizer, W. B.; Müller, K.; Carreira, E. M. Spirocyclic oxetanes: synthesis and properties. *Angewandte Chemie International Edition* **2008,** *47,* 4512-4515.
9. Di Marco, A.; Cassinelli, G.; and Arcamone, F. The discovery of daunorubicin. *Cancer Treatment Reports* **1981,** *65* (Suppl 4), 3-8.

10. Bao, W.; Sheldon, P. J.; Wendt-Pienkowski, E.; Hutchinson, C. R. The Streptomyces peucetius dpsC gene determines the choice of starter unit in biosynthesis of the daunorubicin polyketide. *Journal of Bacteriology* **1999,** *181,* 4690-4695.

11. Bao, W.; Sheldon, P. J.; and Hutchinson, C. R. Purification and properties of the Streptomyces peucetius DpsC beta-ketoacyl:acyl carrier protein synthase III that specifies the propionate-starter unit for type II polyketide biosynthesis. *Biochemistry* **1999,** *38,* 9752-9757.

12. Jackson, D. R. *Structural and Functional Studies of Enzymes Involved in Polyketide and Fatty Acid Biosynthesis;* University of California, Irvine: Irvine, CA, USA, 2015.

13. Kortemme, T.; Creighton, T. E. Ionisation of cysteine residues at the termini of model alpha-helical peptides. Relevance to unusual thiol pKa values in proteins of the thioredoxin family. *Journal of Molecular Biology* **1995,** *253,* 799-812.

14. Davies, C.; Heath, R. J.; White, S. W.; and Rock, C. O. The 1.8 A crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from Escherichia coli. *Structure* **2000,** *8,* 185-195.

15. Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. IMosflm: a new graphical interface for diffraction image processing with MOSFLM. *Acta Crystallographica Section D* **2011,** *67,* 271.

16. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kuntsleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **2010,** *66,* 213.

17. Hertwick, C. The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition* **2009,** *48,* 4688-4716.

CHAPTER 4

Progress Towards the Structural Characterization of the Product Release Mechanism in Enacyloxin IIa Biosynthesis

4.1 Abstract

Type I polyketide natural products are a large and diverse class of secondary metabolites produced primarily in bacteria and fungi. Polyketides are produced by large enzyme complexes known as the polyketide synthase (PKS). In contrast with type II PKSs, type I systems involve multi-domain "megasynthases" that can be hundreds to thousands of kilodaltons in molecular mass. One particular class of type I PKS, known as a *trans*-acyltransferase (*trans*-AT) modular PKS, has emerged recently as a new PKS paradigm with unique biosynthetic properties. One novel feature commonly present in *trans*-AT PKSs is the presence of non-elongating modules that contain ketosynthase (KS) domains incapable of decarboxylative Claisen condensations. These so-called "KS^0" domains frequently sit at the interface between PKS and nonribosomal peptide synthetase (NRPS) machinery in hybrid biosynthetic pathways, though the structural basis for their altered activity and their interface ability is not known. This chapter describes our progress towards the structural characterization of a KS^0 domain from the enacyloxin IIa PKS. This domain facilitates the transfer of a late-stage intermediate from the PKS to downstream NRPS proteins for product release. Crystals of $EnxKS^0$ were obtained in multiple conditions, though the poor diffraction quality and lack of phases precluded structure determination. Additionally, we present initial co-crystallization conditions for the enzyme responsible for product release, termed EnxCON, in complex with a synthetic peptide matching a portion of a peptide carrier protein (PCP) which delivers the substrate to the EnxCON. These crystals will require further optimization before X-ray diffraction analysis can be conducted.

4.2 Introduction

Type I polyketides are a large and structurally diverse class of natural products with a wide variety of bioactivities. Many have been used as pharmaceuticals or as drug leads, including the antibiotic erythromycin, the cholesterol-lowering drug lovastatin, and the immunosuppressant rapamycin[1]. Type I polyketides are biosynthesized by a type I polyketide synthase (PKS), which is characterized by a set of multi-domain enzymes that form very large complexes.

There are multiple flavors of type I PKSs that each have unique biosynthetic properties. Modular type I PKSs work in an assembly line-like fashion in which the growing chain is passed through several multi-domain modules which each act on the nascent polyketide in order[2] (Fig 4-1A). The types of domains present in each module determine the chemical modifications that occur within. In contrast, *trans*-acyltransferase (*trans*-AT) PKSs, a sub-group of modular PKSs which has recently emerged as a unique class unto itself[3], utilizes an assembly line system but also requires the participation of one or more acyltransferase enzymes which act *in trans* to the PKS itself[4] (Fig 4-1B).

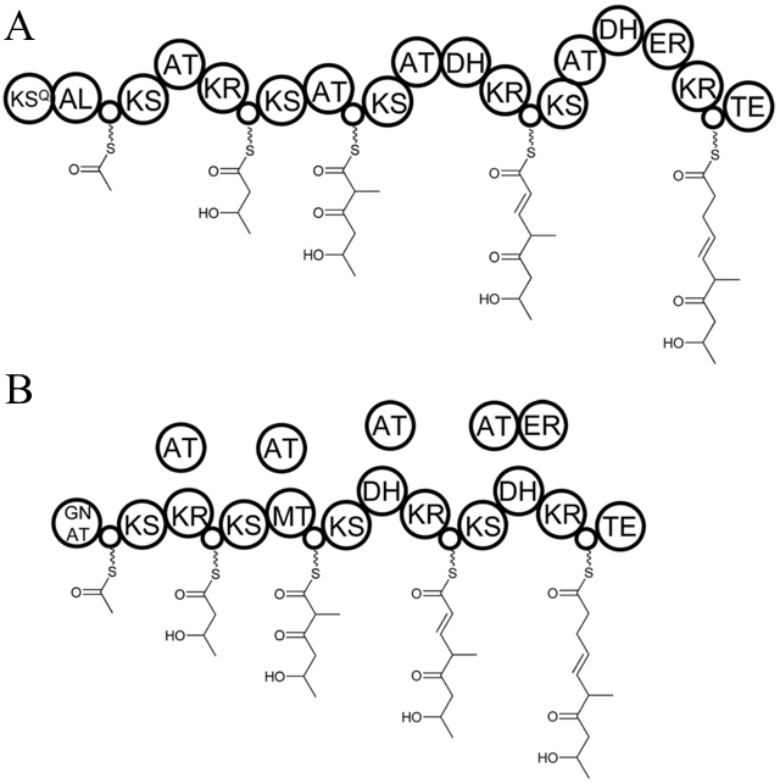

Figure 4-1 Type I PKS types. (a) canonical modular type I PKS architecture with an AT domain embedded in each module. Each domain type is labeled with an abbreviation. (b) *trans*-AT PKS with no AT embedded in each module; rather, a standalone AT is required for activity in each module. Figure adapted from reference 8.

trans-AT PKSs have been heavily studied as of late given their potential for unique extender

unit incorporation via the standalone AT domains required for biosynthesis[5-7]. One trend that has emerged from the study of *trans*-AT PKSs is the presence of non-elongating modules, or modules within which no decarboxylative Claisen condensation occurs via the ketosynthase[8] (KS). Instead, these modules often contain KS domains termed KS^0 that lack a conserved histidine required for the condensation reaction. These KS^0 domains typically facilitate the transfer of acyl groups from one module to the next. Bioinformatics analysis indicates that KS^0 domains are much more likely to be found in PKS-nonribosomal peptide synthetase (NRPS) hybrid systems, particularly at the interface between the PKS and NRPS machinery[9].

The antibiotic enacyloxin IIa (Enx) is an example of a natural product that is biosynthesized by a *trans*-AT PKS-NRPS hybrid system with a KS^0 domain[10]. Enx is biosynthesized by *Burkholderia ambifaria* and has been shown to be effective at targeting *Burkholderia multivorans*, a pathogen that causes devastating lung infections in humans with cystic fibrosis. Two KS domains are present in the final module of the Enx PKS; one has been shown to display normal KS activity while the other (EnxKS^0) is known to be essential for transfer of an end-stage intermediate to the downstream NRPS machinery for product release[10] (Fig 4-2). Specifically, EnxKS^0 transfers the polyketide chain from the acyl carrier protein (ACP) domain in the final module to a

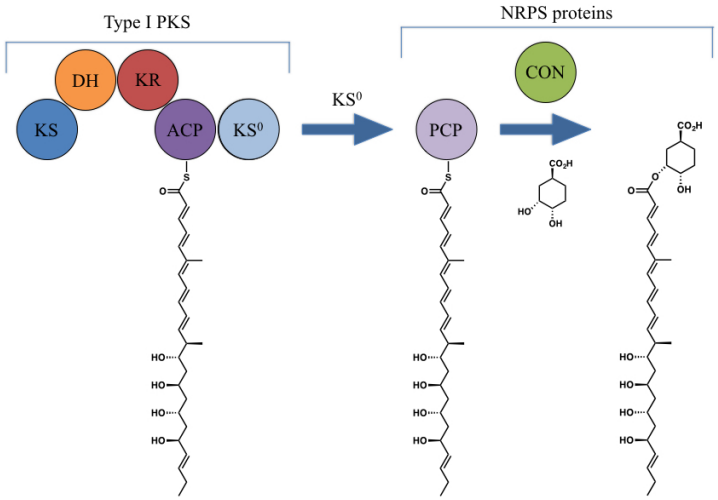

Figure 4-2 Product release in the Enx PKS. A schematic of the final module in the Enx PKS is shown. To release the product, the KS^0 domain transfers a late-stage intermediate from the final ACP to an NRPS PCP domain, which delivers this intermediate to the EnxCON for product release via coupling with DHCCA.

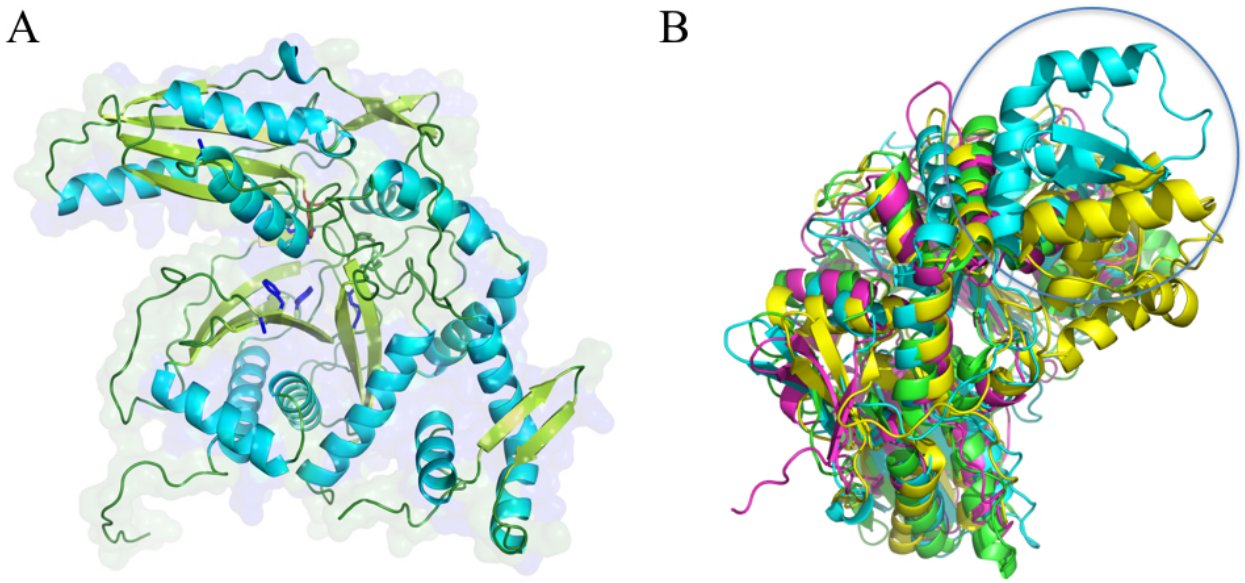

Figure 4-3 Crystal structure of EnxCON. (a) overall structure of EnxCON is shown with active site residues shown in blue sticks. (b) structural alignment of EnxCON (cyan) with several other NRPS condensation domains. The non-conserved docking domain has been circled. Figure courtesy of Tim Valentic.

standalone NRPS peptide carrier protein (PCP) which then interacts with a modified NRPS condensation domain (EnxCON) that releases the product by coupling it with 3,4-dihydroxycyclohexane-1-carboxylic acid (DHCCA).

The crystal structure of EnxCON was recently solved by Tim Valentic[11] (Fig 4-3A), a former graduate student in the Tsai group. In this structure, a unique sub-domain was observed that did not correspond to anything observed in other condensation domain structures (Fig 4-3B). *In vitro* experiments suggest that this domain (termed the "docking domain") is responsible for binding a unique "recognition helix" on the PCP.

Given the observation that a KS^0 domain generally sits at the interface between PKS and NRPS systems, we set out to structurally characterize $EnxKS^0$ to determine the structural basis for its ability to interface between otherwise orthogonal biosynthetic systems as well as its abnormal acyltransferase activity. The sequence of $EnxKS^0$ does not contain anything that appears analogous to the docking domain on EnxCON, so it is possible that $EnxKS^0$ uses another structural element

somewhere on its surface to bind the PCP. Additionally, in order to confirm that the docking domain on EnxCON is responsible for binding the recognition helix on the PCP, we set out to co-crystallize EnxCON with a synthetic eleven-residue peptide that matches the sequence of the PCP recognition helix (PCP-peptide).

Here we present the expression and purification of $EnxKS^0$ as well as its crystallization and preliminary X-ray diffraction analysis. The crystals formed in the space group H 3 2, and the data were processed to a resolution of 3.05 Å. The phases could not be obtained using molecular replacement despite a reasonable level of sequence similarity to published KS structures (~ 40%), and the anomalous diffraction from native sulfur atoms was not sufficient to obtain phases. Crystals of the selenomethionine derivative were also obtained, though the quality of their diffraction pattern was substantially weaker than that of the native crystals (~ 6.5 Å). Additionally, initial co-crystallization conditions for EnxCON in complex with the PCP-peptide are presented. These crystals will require further optimization before X-ray diffraction analysis can be conducted.

4.3 Results and discussion

4.3.1 Expression, purification, and crystallization of $EnxKS^0$

In order to increase the likelihood of crystallization, five constructs of $EnxKS^0$ were cloned from the full length gene (kindly provided by our co-worker Greg Challis at the University of Warwick) into a pET28 expression vector containing a thrombin-cleavable N-term His_6-tag. Each construct varied in the length of the unconserved N-term linker region between $EnxKS^0$ and its neighboring ACP domain.

Each $EnxKS^0$ construct was expressed in *E. coli* and successfully purified using immobilized metal affinity chromatography (IMAC) (Fig 4-4A). The sample was further purified

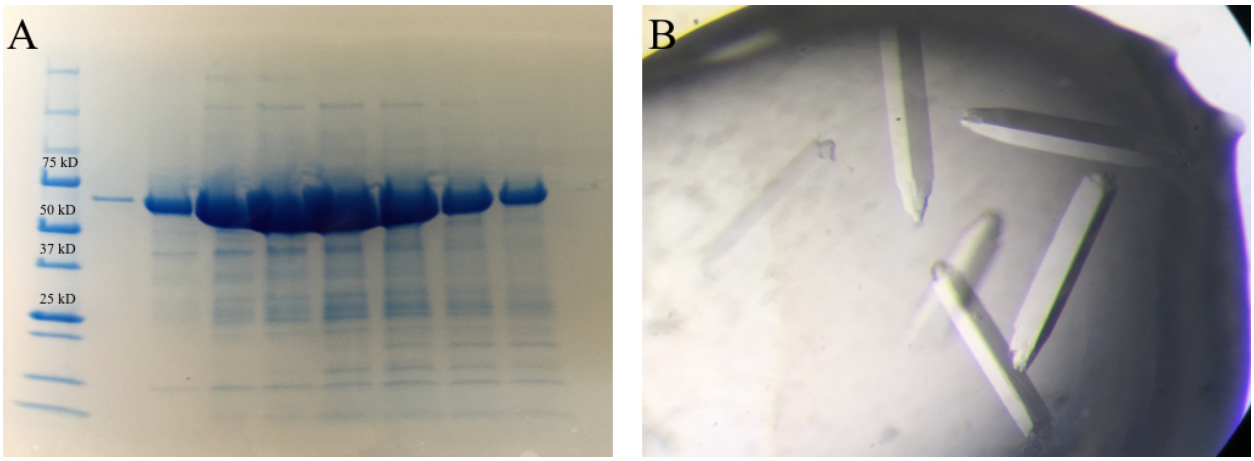

Figure 4-4 Purification and crystallization of EnxKS^0. (a) SDS-PAGE analysis of size-exclusion fractions. (b) crystals of EnxKS^0 obtained after screening and optimization.

using anion exchange chromatography and size exclusion chromatography. EnxKS^0 was then concentrated to 12 mg/mL and screened for crystallization using high-throughput methods, and multiple initial crystallization conditions were determined: 0.1 M HEPES pH 7.5, 0.2 M $CaCl_2$, 20% PEG 400; 0.1 M MES pH 6.5, 0.1 M $MgCl_2$, 20% PEG 400; and 0.1 M HEPES pH 7.5, 0.2 M $MgCl_2$, 20% PEG 400. After optimization, including microseeding and screening for additives, the condition that afforded the largest and most ordered looking crystals was found to be 0.1 M MES pH 8.5, 0.1 M $MgCl_2$, 0.3 M Non-detergent sulfobetaine #256 (Hampton Research), 30% PEG 400 (Fig 4-4B).

4.3.2 Initial X-ray diffraction analysis of EnxKS^0 crystals

EnxKS^0 crystals were flash frozen in liquid nitrogen, and the diffraction pattern of the crystals was measured using beamline 12-2 at the Stanford Synchrotron Radiation Lightsource. The diffraction quality of individual crystals was found to be relatively poor, but by combining several datasets from several different crystals, the data were able to be processed to a resolution of 3.05 Å (Table 4-1).

	EnxKS^0
Data collection	
Wavelength (Å)	1.0
Total reflections	4542001
Unique reflections	139794
Space group	H 3 2
Cell dimensions	
a, *b*, *c* (Å)	278.92 278.92 203.14
α, β, γ (°)	90, 90, 120
Resolution (Å)	93.63 – 3.05 (3.13 – 3.05)
R_{merge}	0.326 (8.175)
R_{pim}	0.082 (2.107)
$I/\sigma(I)$	9.6 (0.8)
$CC_{1/2}$	0.998 (0.368)
Completeness (%)	99.6 (99.2)
Redundancy	32.6 (31.2)
Wilson B-factor	111.5

Table 4-1 Summary of X-ray crystallographic data collection statistics.

Unfortunately, it was not possible to obtain the phases using molecular replacement. Several different published KS structures were used as search models, but the phases could not be obtained despite the fact that the search models shared approximately 40% overall sequence similarity with EnxKS^0. This is not entirely unexpected; type I KS domains typically have highly conserved core elements (>90% similarity) while also containing linker and surface regions with almost no sequence and structural similarity. This can, of course, be problematic for molecular replacement. Additionally, the anomalous diffraction from the native sulfur atoms was not sufficient to obtain the phases. A selenomethionine derivative of EnxKS^0 was also expressed in *E. coli* and purified in the same manner as the native protein, and, after screening for crystallization, it was found to crystallize in the same conditions as native EnxKS^0. However, the diffraction

quality of these crystals was significantly worse than that of the native protein; the data extended to approximately 6.5 Å at best, and these datasets could not be indexed.

4.3.3 Initial co-crystallization of EnxCON and PCP-peptide

A His_6-tagged EnxCON construct (also provided by our co-worker Greg Challis) was expressed in *E. coli* and purified using IMAC, anion exchange chromatography, and size exclusion chromatography (Fig 4-5A). The EnxCON sample was concentrated to 10 mg/mL and incubated with 5 mM PCP-peptide for 1 hour on ice before screening for crystallization.

Two initial crystallization conditions were obtained: 0.1 M Tris pH 8.5, 0.2 M $CaCl_2$, 20 % PEG 4000 and 0.2 M $CsCl_2$, 2.2 M $AmSO_4$ (Fig 4-5B). The morphology of the crystals in the first condition appeared cubic in nature, while the crystals in the second condition appeared more rod-like. In both conditions, however, the crystals were likely too small to be harvested for diffraction analysis; thus, the crystallization will need to be further optimized before their diffraction quality can be assessed.

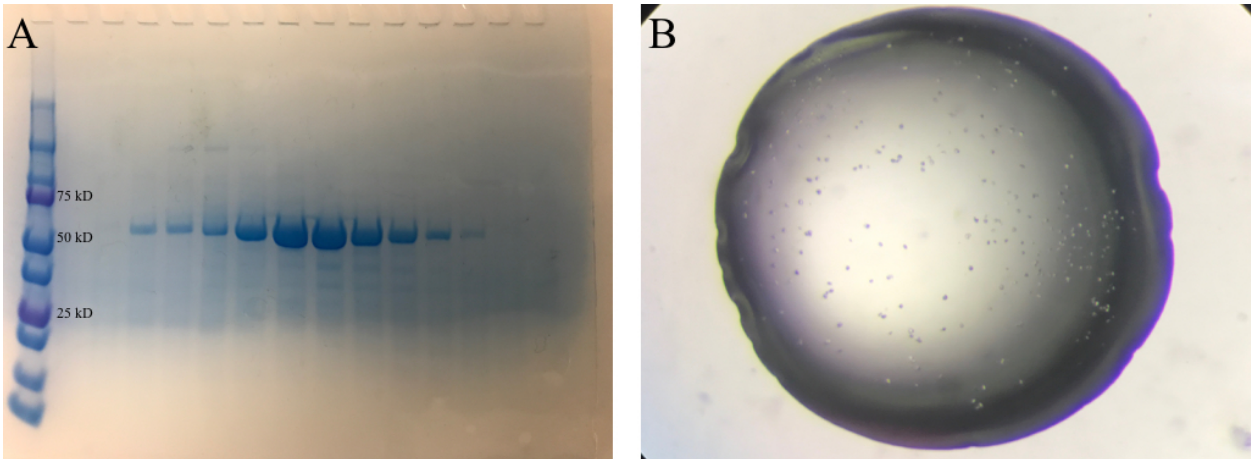

Figure 4-5 Purification and crystallization of EnxCON. (a) SDS-PAGE analysis of size-exclusion fractions. (b) initial microcrystals of EnxCON obtained after screening.

4.4 Conclusion

The results presented herein represent progress towards the structural characterization of the unique product release mechanism in the Enx PKS, though much remains to be done. Several constructs of EnxKS^0 were generated and successfully expressed and purified, and though the enzyme crystallized somewhat readily, the quality of the crystals was difficult to optimize. In addition, initial co-crystallization conditions for EnxCON and the PCP-peptide were obtained, though these crystals will require further optimization. We hope that in the future, the determination of these structures will elucidate the mechanism by which KS^0 domains form the interface between PKS and NRPS systems as well as provide insight into this product release machinery which has potential applications for custom biosynthesis and synthetic biology.

4.5 Materials and methods

4.5.1 Expression and purification of EnxKS^0

Five constructs of EnxKS^0 were cloned from the full length gene into a pET28b expression vector using standard cloning methods. The sequence of interest was amplified using PCR followed by digestion with NdeI and HindIII (restriction sites installed via PCR primers). Following gel extraction of the digested product, the gene was incubated with pre-digested pET28b in the presence of T4 DNA ligase overnight at 16 °C. Constructs were transformed into *E. coli* Top 10 by heatshocking at 42 °C for 45 seconds. The transformed cells were plated on an LB medium supplemented with 50 μg/μL kanamycin and incubated at 37 °C for 18 hours. Cells were transferred to a 10 mL starter culture supplemented with 50 μg/μL kanamycin and shaken at 250 rpm for 18 hours at 37 °C. The plasmid was then isolated from the cultures and sequenced to confirm the identity of the insert.

The pET28 expression vectors coding for His_6-tagged $EnxKS^0$ were transformed into *E. coli* BL21(DE3) by heatshocking at 42 °C for 45 seconds. The transformed cells were plated on an LB medium supplemented with 50 µg/µL kanamycin and incubated at 37 °C for 18 hours. Cells were transferred to a 10 mL starter culture supplemented with 50 µg/µL kanamycin and shaken at 250 rpm for 18 hours at 37 °C. 10 mL of the starter culture was transfered to 1 L of LB media supplemented with 50 µg/µL kanamycin. The cultures were shaken at 200 rpm at 37 °C until the OD_{600} reached 0.6; the expression was then induced by addition of 1 mM IPTG, and the cultures were shaken for 18 hours at 18 °C. The cells were centrifuged at 5100 rpm for 10 minutes, resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 10% glycerol), and flash frozen in liquid nitrogen before storage at -80 °C.

The cells were lysed using a microfluidizer, and the lysate was centrifuged at 21,000 rcf for 1 hour to separate from cellular debris. The lysate was applied to a 5 mL HisTrap HP column (GE Healthcare) and eluted using an imidazole gradient via an Akta Purifier FPLC. The fractions were analyzed using SDS-PAGE, and the fractions containing $EnxKS^0$ were combined and concentrated to 10 mg/mL. The protein sample was further purified using a Hi-Trap Q column and a Superdex 200 column (GE Healthcare), and fractions were again analyzed using SDS-PAGE. The selected fractions containing $EnxKS^0$ were concentrated to 12 mg/mL and flash frozen in liquid nitrogen before storage at -80 °C.

4.5.2 Crystallization of $EnxKS^0$ and X-ray diffraction analysis

$EnxKS^0$ was crystallized in three different conditions: 0.1 M HEPES pH 7.5, 0.2 M $CaCl_2$, 20% PEG 400; 0.1 M MES pH 6.5, 0.1 M $MgCl_2$, 20% PEG 400; and 0.1 M HEPES pH 7.5, 0.2 M $MgCl_2$, 20% PEG 400. After optimization, including microseeding and screening for additives,

the condition that afforded the largest and most ordered looking crystals was found to be 0.1 M MES pH 8.5, 0.1 M $MgCl_2$, 0.3 M Non-detergent sulfobetaine #256 (Hampton Research), 30% PEG 400. These crystals were then flash cooled in liquid nitrogen.

The diffraction pattern of the crystals was measured at the Stanford Synchrotron Radiation Lightsource using beamline 12-2. The diffraction images were processed using Mosflm[12], and multiple datasets were combined using the program Blend within the ccp4 suite of programs[13]. The resulting dataset was scaled using Aimless. The Phenix suite of programs was used for molecular replacement (Phaser-MR) and SAD phasing (AutoSol)[14].

4.5.3 Expression, purification, and crystallization of EnxCON

A pET151 expression vector coding for His_6-tagged EnxCON was transformed into *E. coli* BL21(DE3) by heatshocking at 42 °C for 45 seconds. The transformed cells were plated on an LB medium supplemented with 100 μg/μL carbenicillin and incubated at 37 °C for 18 hours. Cells were transferred to a 10 mL starter culture supplemented with 100 μg/μL carbenicillin and shaken at 250 rpm for 18 hours at 37 °C. 10 mL of the starter culture was transfered to 1 L of LB media supplemented with 100 μg/μL carbenicillin. The cultures were shaken at 200 rpm at 37 °C until the OD_{600} reached 0.6; the expression was then induced by addition of 1 mM IPTG, and the cultures were shaken for 18 hours at 18 °C. The cells were centrifuged at 5100 rpm for 10 minutes, resuspended in lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 10% glycerol), and flash frozen in liquid nitrogen before storage at -80 °C.

The cells were lysed using a microfluidizer, and the lysate was centrifuged at 21,000 rcf for 1 hour to separate from cellular debris. The lysate was applied to a 5 mL HisTrap HP column (GE Healthcare) and eluted using an imidazole gradient via an Akta Purifier FPLC. The fractions

were analyzed using SDS-PAGE, and the fractions containing EnxCON were combined and concentrated to 10 mg/mL. The protein sample was further purified using a Hi-Trap Q column and a Superdex 200 column (GE Healthcare), and fractions were again analyzed using SDS-PAGE. The selected fractions containing EnxCON were concentrated to 10 mg/mL and flash frozen in liquid nitrogen before storage at -80 °C.

EnxCON was co-crystallized with the PCP-peptide (GenScript) in two different conditions: 0.1 M Tris pH 8.5, 0.2 M $CaCl_2$, 20 % PEG 4000 and 0.2 M $CsCl_2$, 2.2 M $AmSO_4$.

References

1. Hertwick, C. The biosynthetic logic of polyketide diversity. *Angewandte Chemie International Edition* **2009,** *48,* 4688-4716.
2. Keating-Clay, A. The uncommon enzymology of cis-acyltransferase assembly lines. *Chemical Reviews* **2017**.
3. Nguyen, T.; Ishida, K.; Jenke-Kodama, H.; Dittmann, E.; Gurgui, C.; Hochmuth, T.; Taudien, S.; Platzer, M.; Hertweck, C.; Piel, J. Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nature Biotechnology* **2008,** *26,* 225-233.
4. Cheng, Y. Q.; Tang, G. L.; Shen, B. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proceedings of the National Academy of Sciences* **2003,** *100,* 3149-3154.
5. Musiol, E. M.; Härtner, T.; Kulik, A.; Moldenhauer, J.; Piel, J.; Wohlleben, W.; Weber, T. Supramolecular templating in kirromycin biosynthesis: the acyltransferase KirCII loads ethylmalonyl-CoA extender onto a specific ACP of the trans-AT PKS. *Chemistry and Biology* **2011,** *18,* 438-444.
6. Zhao, C. H.; Coughlin, J. M.; Ju, J. H.; Zhu, D. Q.; Wendt-Pienkowski, E.; Zhou, X. F.; Wang, Z. J.; Shen, B.; Deng, Z. X. Oxazolomycin biosynthesis in Streptomyces albus JA3453 featuring an "acyltransferase-less" type I polyketide synthase that incorporates two distinct extender units. *Journal of Biological Chemistry* **2010,** *285,* 20097-20108.
7. Menche, D.; Arikan, F.; Perlova, O.; Horstmann, N.; Ahlbrecht, W.; Wenzel, S. C.; Jansen, R.; Irschik, H.; R., M. Stereochemical determination and complex biosynthetic assembly of etnangien, a highly potent RNA polymerase inhibitor from the myxobacterium Sorangium cellulosum. *Journal of the American Chemical Society* **2008,** *130,* 14234-14243.
8. Helfrich, E. J. N.; Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases. *Natural Product Reports* **2016,** *33,* 231-316.
9. Huang, Y.; Tang, G. L.; Pan, G.; Chang, C. Y.; Shen, B. Characterization of the ketosynthase and acyl carrier protein domains at the LnmI nonribosomal peptide

synthetase−polyketide synthase interface for leinamycin biosynthesis. *Organic Letters* **2016,** *18,* 4288-4291.

10. Mahenthiralingam, E.; Song, L.; Sass, A.; White, J.; Wilmot, C.; Marchbank, A.; Boaisha, O.; Paine, J.; Knight, D.; Challis, G. L. Enacyloxins are products of an unusual hybrid modular polyketide synthase encoded by a cryptic Burkholderia ambifaria genomic island. *Chemistry and Biology* **2011,** *18,* 665-677.
11. Valentic, T. R. Structural studies of polyketide synthase biosynthetic machinery; University of California, Irvine: Irvine, CA, USA, 2016.
12. Battye, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. IMosflm: a new graphical interface for diffraction image processing with MOSFLM. *Acta Crystallographica Section D* **2011,** *67,* 271.
13. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D* **2011,** *67,* 235-242.
14. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kuntsleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **2010,** *66,* 213.

CHAPTER 5

Conclusions and Future Directions

5.1 Conclusions

Fatty acids and polyketide natural products are critical components in many aspects of microbial life including membrane structure and self- defense. These molecules have also become relevant to human industry as well. Fatty acids can be used as biofuels, which can increase our energy security and decrease our reliance on fossil fuels as well as our reliance on the nations that produce them[1]. Enzymes that biosynthesize fatty acids have also been targeted with therapeutics against infectious diseases. Polyketides have a variety of bioactivities that have proven particularly useful as pharmaceuticals; indeed, polyketide-based drugs have annual sales in excess of $15 billion[2].

These molecules are produced by closely related enzyme systems termed fatty acid synthases (FAS) and polyketide synthases (PKS). These massive enzyme complexes share great potential for human engineering and synthetic biology, though two primary challenges stand in the path to long-term success. First, the protein-protein interactions that are the key to biosynthesis remain poorly understood due to the transient nature of these interactions. Second, the extreme reactivity of polyketide substrates and intermediates has impeded our understanding of how PKS enzymes act on their substrates with high levels of specificity.

The work presented in this dissertation outlines the use of synthetic substrate analogues as a strategy to overcome these challenges. We hope that the tools developed and employed herein, mechanism-based crosslinking and the use of inert substrate mimics, will continue to illuminate the structure and function of these powerful enzymes in the future.

5.1.1 Mechanism-based crosslinking and AcpP-FabB structure determination (Chapter 2)

In order to determine the structural basis for acyl carrier protein (ACP)-ketosynthase (KS) recognition in the *E. coli* type II FAS, a mechanism-based crosslinking strategy was employed to solve the crystal structure of the AcpP-FabB complex. This structure enabled the first residue-level mapping of the complete interface between an ACP and a KS. Comparison of this structure with the AcpP-FabA structure, as well as bioinformatics analysis, suggests that AcpP use different parts of its surface to interact with different catalytic partners. This structural information was used to generate surface mutants of FabB that modulated the lipid profile of *E. coli* cultures. FabB is known to control properties such as chain length and saturation level, and these properties were both altered in the mutant cultures.

5.1.2 Oxetane-based polyketide isosteres and co-crystal structure with DpsC (Chapter 3)

In order to determine the structural basis for extender unit binding and catalysis, an inert, oxetane-based extender unit analogue was developed and co-crystallized with the KS III enzyme DpsC. This structure provided the first snapshot of a ketosynthase and all of its substrates in position just prior to catalysis. DpsC is a unique, bi-functional enzyme, and the basis for its KS III activity was not known due to its lack of conserved residues known to be necessary for catalysis. However, this structure allowed us to propose the first complete mechanism for this enzyme. We hope that this new class of probes will find use in the biosynthesis research community.

5.1.3 Towards the crystal structures of EnxKS^0 and EnxCON + synthetic PCP peptide

In order to determine the structural basis for a KS^0's ability to interface with both PKS and nonribosomal peptide synthetase (NRPS) machinery, we set out to solve the crystal structure of a

KS^0 domain ($EnxKS^0$) from the enacyloxin IIa PKS. This enzyme is part of a unique sub-group of KSs that does not catalyze decarboxylative Claisen condensation but rather catalyzes acyl transfer, typically between PKS and NRPS enzymes. Crystals were obtained, but the poor diffraction quality and lack of phases precluded structure determination. Future work is needed to complete this project. Additionally, to confirm the mode of binding between EnxCON and its cognate peptide carrier protein (PCP), EnxCON was co-crystallized with a synthetic peptide matching the sequence of its predicted binding site on the PCP. Putative co-crystals were obtained following crystallization screening, though they require further optimization before X-ray diffraction analysis can be conducted. While the use of synthetic peptides is certainly not novel, this work does reiterate the useful role that synthetic analogues can play.

5.2 Future Directions

5.2.1 Structure determination of additional ACP-partner complex structures

Thus far, the structures of two crosslinked ACP-partner structures have been determined, though there are several other complexes that remain to be characterized. For example, there are two other KS enzymes in the *E. coli* type II FAS, FabF and FabH, that can be targeted using the same crosslinker utilized in chapter 2 of this dissertation. While the mode of ACP binding to FabF may be similar to that of FabB, the interface with FabH is predicted to be quite different given that FabH is a priming KS as opposed to FabB and FabF which are elongating KSs. This structure could provide new insight into ACP-partner recognition patterns. Also, there are currently no published structures of an ACP-KS complex from the PKS field. A crystal structure of the complex between an ACP and a KS/CLF from a type II PKS would be an excellent place to begin as it would allow for the first direct comparison of this complex type between an FAS and a PKS.

In order to determine the structures of other types of ACP-partner complexes, new crosslinkers will need to be developed. Currently, the only mechanism-based crosslinkers in the literature are specific for dehydratase (DH) and KS domains[3,4]. This could present quite a challenge, however, for most of the other FAS and PKS domain types. For example, the other two enzymes in the *E. coli* type II FAS elongation cycle—FabG and FabI—utilize NADPH/NADH cofactors to perform the key chemical transformations. Their lack of active site nucleophilic side chains suggests that a new strategy will have to be pursued in order to target these enzymes. One approach, developed in the lab of Michael Burkart at the University of California, San Diego, utilized the tight-binding FabI inhibitor triclosan coupled to the phosphopantetheine (PPT) cofactor of AcpP to obtain an AcpP-FabI complex in solution[5] (Fig 5-1). The triclosan-loaded AcpP displayed nM affinity for FabI, and it was able to selectively pull down FabI in an affinity purification. If covalent linkage is not an option for certain domain types, tight binding inhibitors may be the net best option for future attempts at co-crystallization.

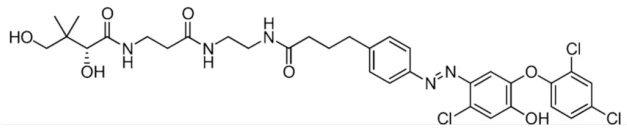

Figure 5-1 Tricolsan-based probe for trapping an AcpP-FabI complex. This probe, when loaded onto AcpP, displays nM affinity for FabI. Figure from reference 5.

5.2.2 Development of additional oxetane-based polyketide isosteres

While the malonyl-PPT mimic presented in chapter 3 is a great start, the possibilities for additional oxetane-based polyketide analogues are quite numerous. The next logical progression from the work presented herein would be the development of a poly-β-ketone mimic in which alternating ketide units are replaced with oxetane rings (Fig 5-2). This should significantly lower

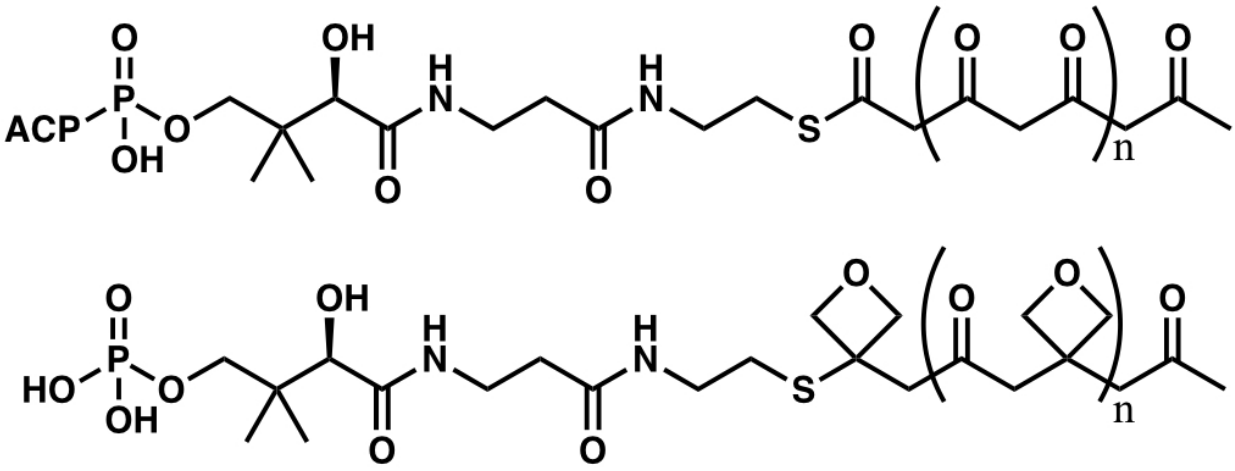

Figure 5-2 Proposed oxetane-based poly-β-ketone mimics. A general example of an oxetane-based PKS intermediate (bottom) would more closely resemble the natural substrate (top) than previous isoxazole-based probes.

the pK_a of the α-protons, resulting in much greater stability and an increased resistance to spontaneous aldol condensation. These probes are predicted to be much more similar to the natural substrates than the previous generation of isoxazole-based probes. Also, the earlier probes were quite limited in their flexibility, which further limited their effectiveness. The synthesis of these probes is currently being developed in the lab of Chris Vanderwal at the University of California, Irvine.

5.3 Final thoughts

If you've made it this far in my dissertation, then please humor me one last time as I allow myself a final soapbox moment to share my thoughts on the future of FAS and PKS enzyme engineering efforts. While structure-based, rational engineering of enzymes will continue to have a place in this field (and rightly so), I predict that directed evolution will play a much larger role in the future of all protein engineering efforts, including FAS and PKS enzymes. Directed evolution is not a particularly new or esoteric technique, so I find it peculiar that this powerful methodology is not utilized more widely in the PKS field especially since the few examples that

do exist in the literature have been quite successful. The whole message of this dissertation is that we do not yet have the structural information necessary to successfully engineer these enzymes, but directed evolution requires very little prior knowledge of the enzyme's structure or mechanism.

One of my main takeaways from the extensive literature I've read in the past four years is that there may not be a universal "code" of PKS protein-protein and protein-substrate interactions that, once unlocked, will enable universal engineering of PKS enzymes and the biosynthesis of designer molecules as envisioned by many in the field. Each PKS seems to be unique, and cracking the enzymology and dynamics of one pathway may or may not shed much light on a different pathway. Evolution doesn't require extensive information about what makes all PKSs tick; the only thing it cares about is what works.

References

1. Janssen, H. J.; Steinbuchel, A. Fatty acid synthesis in Escherichia coli and its applications towards the production of fatty acid based biofuels. *Biotechnology for Biofuels* **2014,** *7,* 7.
2. Borchardt, J. K. Combinatorial biosynthesis: panning for pharmaceutical gold. *Modern Drug Discovery* **1999,** *2,* 22-29.
3. Ishikawa, F.; Haushalter, R. W.; Burkart, M. D. Dehydratase-specific probes for fatty acid and polyketide synthase. *Journal of the American Chemical Society* **2012,** *134,* 769-772.
4. Worthington, A. S.; Rivera, H.; Torpey, J. W.; Alexander, M. D.; Burkart, M. D. Mechanism-based protein cross-linking probes to investigate carrier protein-mediated biosynthesis. *ACS Chemical Biology* **2006,** *1,* 687-691.
5. Tallorin, L.; Finzel, K.; Nguyen, Q. G.; Beld, J.; La Clair, J. J.; Burkart, M. D. Trapping of the enoyl-acyl carrier protein reductase-acyl carrier protein interaction. *Journal of the American Chemical Society* **2016,** *138,* 3962-3965.