UNIVERSITY OF CALIFORNIA SAN DIEGO

Enzymology of Acyl Carrier Protein-Dependent Synthases in *Escherichia coli*

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Chemistry

by

Aochiu Chen

Committee in charge:

Professor Michael D. Burkart, Chair
Professor Eric E. Allen
Professor Joseph P. Noel
Professor Stanley J. Opella
Professor Emmanuel A. Theodorakis

2022

This dissertation of Aochiu Chen is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

To my family, friends, and loved ones who have supported me during this journey.

In memory of Tom Yang, a friend from middle school and a former UCSD student.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Acyl Carrier Protein ..... ACP

Acyl Carrier Protein Hydrolase ..... AcpH

E. coli Acyl Carrier Protein ..... AcpP

Acyl Carrier Protein Synthase ..... AcpS

8-Amino-7-Oxononanoate ..... AON

8-Amino-7-Oxononanoate Synthase ..... AONS

Aromatase ..... AR

Asymmetric Unit ..... ASU

Acyltransferase ..... AT

Adenosine Triphosphate ..... ATP

Bismaleimidoethane ..... BMOE

Benzenesulfonyl Fluoride ..... BSF

Chain-Length-Factor ..... CLF

Coenzyme A ..... CoA

Chemical Shift Perturbation ..... CSP

Cyclase ..... CYC

Dichloromethane ..... DCM

Dehydratase ..... DH

N, N-Diisopropylethylamine ..... DIEA

Diisopropylamine ..... DIPA

4-Dimethylaminopyridine ..... DMAP

Enoyl Reductase ..... ER

Electrospray Ionization ..... ESI

Fatty Acid ..... FA

Fatty Acid Biosynthesis ..... FAB

Fatty Acid Synthase ..... FAS

Type II Fatty Acid Synthase ..... FASII

Hydrogen-Bond Donor ..... HBD

High-Performance Liquid Chromatography ..... HPLC

High-Resolution Mass Spectrometry ..... HRMS

Heteronuclear Single Quantum Coherence ..... HSQC

Ketoreductase ..... KR

Ketosynthase ..... KS

Lithium Diisopropylamide ..... LDA

Malonylacyltransferase ..... MAT

Molecular Dynamics ..... MD

Molecular Weight Cut off ..... MWCO

Nuclear Magnetic Resonance ..... NMR

Polyacrylamide Gel Electrophoresis ..... PAGE

Protein Data Bank ..... PDB

Partner Enzyme ..... PE

Isoelectric Point ..... pI

Polyketide Biosynthesis ..... PKB

Polyketide Synthase ..... PKS

Type II Polyketide Synthase ..... PKSII

Pyridoxal 5'-Phosphate ..... PLP

Partner Protein ..... PP

4'-Phosphopantetheine ..... PPant

Protein-Protein Interaction ..... PPI

4'-Phosphopantetheinyl Transferase ..... PPTase

Protein-Substrate Interaction ..... PSI

Root-Mean-Square Deviation ..... RMSD

Site-Directed Mutagenesis ..... SDM

Sodium Dodecyl Sulphate ..... SDS

Sulfur (VI) Fluoride Exchange ..... SuFEx

Thioesterase ..... TE

Trifluoroacetic Acid ..... TFA

Tetrahydrofuran ..... THF

Type I Polyketide Synthase ..... TIPKS

Unsaturated Fatty Acid ..... UFA

Wild-Type ..... WT

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my advisor Professor Michael Burkart for his support and mentorship. With his help and guidance, I was able to perform high quality research with motivation and enthusiasm, which makes the journey enjoyable. I appreciate every opportunity he provided me especially getting me involved in a number of interesting projects. I would also like to thank Dr. Jim La Clair for sharing his wisdoms on every aspect of life. The mentorship that he and Professor Burkart gave has got me well prepared for a career in science, and I will be forever grateful for that.

I would like to express my sincere thanks to my coworker, co-author, mentor, and friend Dr. Jefferey Mindrebo. Many of my projects would not be possible without his pioneering work on the ketosynthases. His continuous mentorship also helps me not only on the science part, but more importantly on becoming a better person. My thanks also go to Dr. Eunice Kim and Dr. Thomas Bartholow, who had mentored me on doing research since the first day I joined the Lab. Thank you to Professor Tony Davis as well on providing me almost all the crosslinking probes I used in my study, and Dr. Ashay Patel whom I turned to for help on computational part. All the people mentioned here belong to the ketosynthase team, and I enjoyed every moment working with them.

To all the Burkart Lab members, thank you for all your support and hard work. You all have contributed to my amazing journey of studying abroad that I would have never dreamed of. Special thanks to Rebecca Re and Ziran Jiang who had worked with me closely on reviews and projects, and my undergraduate researchers Clara Duong and Kelley Tran on their help. I would also like to thank the Noel Lab, especially Professor Joseph Noel and

Dr. Gordon Louie, for the lessons in x-ray crystallography. Professor Noel along with my other committee members also provided me helpful feedbacks on my research, and I really appreciate that.

Finally, I want to thank my wife, Yi-Ning Huang, my parents, Cheng-Cheng Chen and Mei-Jing Huang, my brother and his wife, Hanchiu Chen and Jingying Lin, and their dog, Mochi, on their emotional support during the journey. My courage and inspiration largely come from my family and friends, and this has only become obvious to me since the moment I stepped out of the San Diego International Airport on August 16th, 2016.

Chapter 1, entitled Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions, in full, is a reprint of the material as it appears in: Chen A, Re RN, Burkart MD. "Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions", Natural Product Reports, vol. 35, 2018. The dissertation author is the primary author of this manuscript.

Chapter 2, entitled Enzymology of Standalone Elongating Ketosynthases, in full, is a reprint of the material as it appears in: Chen A, Jian Z, Burkart MD. "Enzymology of Standalone Elongating Ketosynthases", RSC Chemical Biology, under review. The dissertation author is the primary author of this manuscript.

Chapter 3, entitled Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases, in full, is a reprint of the material as it appears in: Mindrebo JT*, Chen A*, Kim WE, Re RN, Davis TD, Noel JP, Burkart MD. "Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases", ACS Catalysis, vol. 11,

2021. The dissertation author is the primary co-author of this manuscript along with Dr. Jefferey T. Mindrebo.

Chapter 4, entitled Utilizing Mechanism-Based Crosslinking Probes to Capture the *E. coli* Ketosynthase, FabB, in Different Catalytic States, in full, is a reprint of the material as it appears in: Chen A*, Mindrebo JT*, Davis TD, Kim WE, Katsuyama Y, Jiang Z, Ohnishi Y, Noel JP, Burkart MD. "Utilizing Mechanism-Based Crosslinking Probes to Capture the *E. coli* Ketosynthase, FabB, in Different Catalytic States", Acta Crystallographica Section D, submission in progress. The dissertation author is the primary co-author of this manuscript along with Dr. Jefferey T. Mindrebo.

VITA

2022 Doctor of Philosophy, Chemistry, University of California San Diego

2018 Master of Science, Chemistry, University of California San Diego

2015 Bachelor of Science, Chemistry, National Taiwan University

PUBLICATIONS

Chen A, Re RN, Burkart MD. "Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions.", Natural Product Reports, vol. 35, 2018.

Du D, Katsuyama Y, Horiuchi M, Fushinobu S, **Chen A**, Davis TD, Burkart MD, Ohnishi Y. "Structural basis for selectivity in a highly reducing type II polyketide synthase.", Nature Chemical Biology, vol. 16, 2020.

Mindrebo JT, Patel A, Kim WE*, Davis TD*, **Chen A***, Bartholow TG*, La Clair JJ, McCammon JA, Noel JP, Burkart MD. "Gating mechanism of elongating β-ketoacyl-ACP synthases.", Nature Communications, vol. 11, 2020.

Mindrebo JT*, **Chen A***, Kim WE, Re RN, Davis TD, Noel JP, Burkart MD. "Structure and mechanistic analyses of the gating mechanism of elongating ketosynthases.", ACS Catalysis, vol. 11, 2021.

Chen A, Jiang Z, Burkart MD. "Enzymology of standalone elongating ketosynthases.", RSC Chemical Science, under review.

Chen A*, Mindrebo JT*, Davis TD, Kim WE, Katsuyama Y, Jiang Z, Ohnishi Y, Noel JP, Burkart MD. "Utilizing mechanism-based crosslinking probes to capture the *E. coli* ketosynthase, FabB, in different catalytic states.", Acta Crystallographica Section D, submission in progress.

ABSTRACT OF THE DISSERTATION

Enzymology of Acyl Carrier Protein-Dependent Synthases in *Escherichia coli*

by

Aochiu Chen

Doctor of Philosophy in Chemistry

University of California San Diego, 2022

Professor Michael D. Burkart, Chair

To date, at least twenty enzymes in *E. coli* have been identified to functionally interact with the acyl carrier protein (ACP), a small, four-helix bundle that shuttles thioester-linked pathway intermediates to the active site of partner enzyme (PE). About half of these PEs are fatty acid synthases (FASs) that catalyze the primary metabolic pathway, fatty acid biosynthesis (FAB), to synthesize saturated and mono-unsaturated fatty acids. FAB goes through iterative cycles that condense 2-carbon unit at a time onto the growing acyl chain. During this process, FASs ensure product fidelity by applying accurate substrate recognition and maintain catalytic efficiency by forming productive protein-protein interactions (PPIs)

with ACP. The modular biosynthetic logic of FASs, and the evolutionarily related polyketide synthases (PKSs), makes them attractive metabolic engineering targets, but a thorough understanding of the underlying enzymology is needed to better achieve the goal.

Here, we utilize the PE-ACP crosslinking methodology to investigate the enzymology of three ACP-dependent synthases that catalyze carbon-carbon bond formation through decarboxylative condensation. These synthases are elongating ketosynthases (KSs), FabF and FabB, from FASs, and the 8-amino-7-oxononanoate synthase, BioF, from the biotin biosynthesis pathway. The first chapter provides an extensive review on the FASs and PKSs, with a focus on PPIs and the biochemical tools that are used to study these enzymes. The well-developed biochemical techniques, especially the mechanism-based crosslinking method, has allowed us to rapidly leverage our understanding of KS enzymology in the past three years (2019-2021), and this leads to the second chapter that provides a comprehensive overview on mechanisms governed by KSs. In the third chapter, a recently discovered double-loop gating mechanism was closely examined by mutational study of FabF. We also reported two FabF-ACP structures that capture the gate in different conformations. Although highly similar to FabF in function, FabB has a different substrate preference and is essential to unsaturated fatty acid (UFA) synthesis. In chapter four, crosslinkers that mimic specific stages of catalysis were used to probe the substrate specificity of FabB. The two reported FabB-ACP complex structures also delineate gating mechanisms utilized by the enzyme.

Finally, in the fifth chapter, we turn our focus to a PLP-dependent (Pyridoxal 5'-Phosphate) biotin synthase, BioF, that catalyzes the condensation between pimeloyl-ACP and L-alanine. We developed a SuFEx-based (Sulfur (VI) Fluoride Exchange) crosslinker

that successfully led to the BioF-ACP complex structure. From structural analysis and mutational study, we identified a BioF-ACP interface that was different from what *in silico* docking suggested.

Chapter 1: Type II Fatty Acid and Polyketide Synthases: Deciphering Protein-Protein and Protein-Substrate Interactions

Abstract

Metabolites from type II fatty acid synthase (FAS) and polyketide synthase (PKS) pathways differ broadly in their identities and functional roles. The former are considered primary metabolites that are linear hydrocarbon acids, while the latter are complex aromatic or polyunsaturated secondary metabolites. Though the study of bacterial FAS has benefitted from decades of biochemical and structural investigations, type II PKSs have remained less understood. Here we review the recent approaches to understanding the protein-protein and protein-substrate interactions in these pathways, with an emphasis on recent chemical biology and structural applications. New approaches to the study of FAS have highlighted the critical role of the acyl carrier protein (ACP) with regard to how it stabilizes intermediates through sequestration and selectively delivers cargo to successive enzymes within these iterative pathways, utilizing protein-protein interactions to guide and organize enzymatic timing and specificity. Recent tools that have shown promise in FAS elucidation should find new approaches to studying type II PKS systems in the coming years.

Introduction

Type II fatty acid synthase (FAS) and polyketide synthase (PKS) share much in common in terms of their core enzymatic components, but their products couldn't be more different. While fatty acids are metabolites primarily composed of one carboxyl group at the end of a linear, saturated, or mono-unsaturated hydrocarbon, the secondary metabolites of

type II PKSs, on the other hand, are highly functionalized polycyclic aromatic compounds, along with a few polyunsaturated natural products that have recently been identified. Among this class are first-line antibiotics, anticancer drugs, veterinary medicines, and agrochemicals (Figure 1.1).[1-3] An understanding of their biosynthesis offers the potential opportunity to redesign the pathways to produce novel aromatic and polyunsaturated compounds. However, while the type II FAS enzymatic pathway and protein structure have been extensively studied over the last century, many enzymes in type II PKS pathways remain mysterious, despite substantial interest in redesigning the pathways to produce custom designed molecules. Yet even the *Escherichia coli* FAS, the most well-studied acetate pathway, has remained recalcitrant to metabolic engineering and product control.[4] It has become clear that a more fundamental understanding of both pathways will be the key to the long-term goals of controlling their biosynthesis.

As evidenced by Figure 1.2, *Streptomyces coelicolor* produces several natural products from ACP dependent pathways, including the *de novo* fatty acid synthesis and actinorhodin biosynthesis, a type II PKS pathway. The two ACPs from these pathways are highly homologous, sharing 34% amino acid identity (65% similarity) such that a casual BLAST searcher would deduce that they served the same function. Yet these pathways do not become scrambled; *S. coelicolor* does not produce hybrid or varied compounds. How do two highly homologous acetate pathways achieve orthogonality? It can only occur through precisely controlled protein-protein interactions.

The similarities and differences in the mechanism, structure, organization, and timing of these two pathways have offered challenges to their understanding. Given the

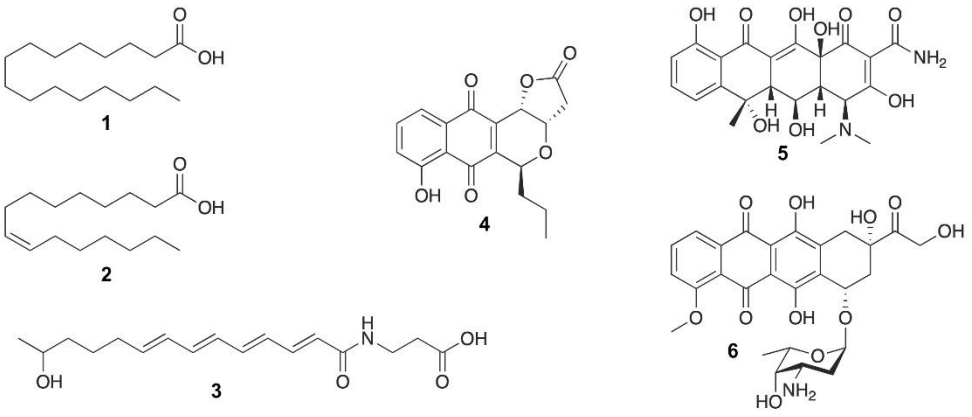


Figure 1.1: Representative metabolites from type II fatty acid synthase (FAS) and polyketide synthase (PKS) pathways. Fatty acids: palmitic acid (1) and palmitoleic acid (2). Polyene (type II) polyketide: ishigamide (3). Aromatic polyketides: agricultural antibiotic frenolicin B (4), antibiotic oxytetracycline (5), and anticancer doxorubicin (6).

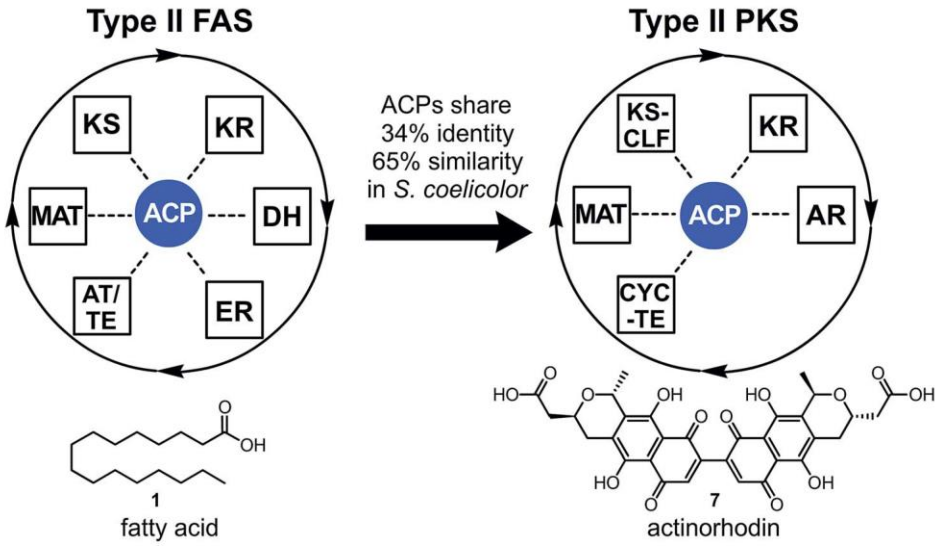


Figure 1.2: The central role of ACP in type II FAS and PKS pathways. Here, the actinorhodin type II PKS biosynthetic proteins in *Streptomyces coelicolor* are evolutionarily derived from type II FAS. Despite these similarities, the pathways remain isolated from each other - they do not scramble or exchange intermediates to provide chimeric products. This inherent fidelity, as well as known cross-reactivities, remains an important challenge for understanding and eventually reprogramming these pathways. ACP = acyl carrier protein; KS = β-ketoacyl-ACP synthase; MAT = malonylacyltransferase; KR = β-ketoacyl-ACP-reductase; DH = β-hydroxyacyl-ACP dehydratase; ER = enoyl reductase; AT/TE = acyltransferase/thioesterase; KS-CLF = ketosynthase-chain length factor; AR = aromatase; CYC-TE = cyclase-thioesterase.

evolutionary relationship between FAS and PKS,[5] lessons learned in more well-known systems, including decades of research on bacterial FAS, can be brought to bear for less characterized pathways in order to more fully understand their mechanisms and processivities. Indeed, new tools that have been designed to interact with any carrier protein dependent pathway are commonly first applied to bacterial FAS enzymes, particularly those of *E. coli*. In this review, we discuss the basic enzymes and mechanisms of both pathways and introduce efforts to date that elucidate the enzyme structures and activities, with a focus on new tools and methods. Chief among these is a newfound appreciation and focus on protein-protein interactions. It has become apparent that protein-protein interactions are critical to the orchestration of all carrier protein dependent biosynthetic pathways, yet the details of these essential phenomena remain mostly uncharacterized. While we understand that the carrier proteins of each synthase must functionally interact with the enzymatic partners through specific recognition elements, a considerable amount of work remains to be done to complete this characterization, even in bacterial FAS. Although much of what is described here focuses on FAS, we address the type II systems that have been studied and draw parallels where possible.

Type II fatty acid synthase and polyketide synthase enzymology

Initiation in type II FAS and PKS. Type II fatty acid biosynthesis is an iterative cycle that relies on ACP to transport fatty acid precursors and intermediates through interactions with its partner enzymes, each of which takes on a different role in the pathway (Figure 1.3, green path). We will focus on what is known about FAS from E. *coli* given that

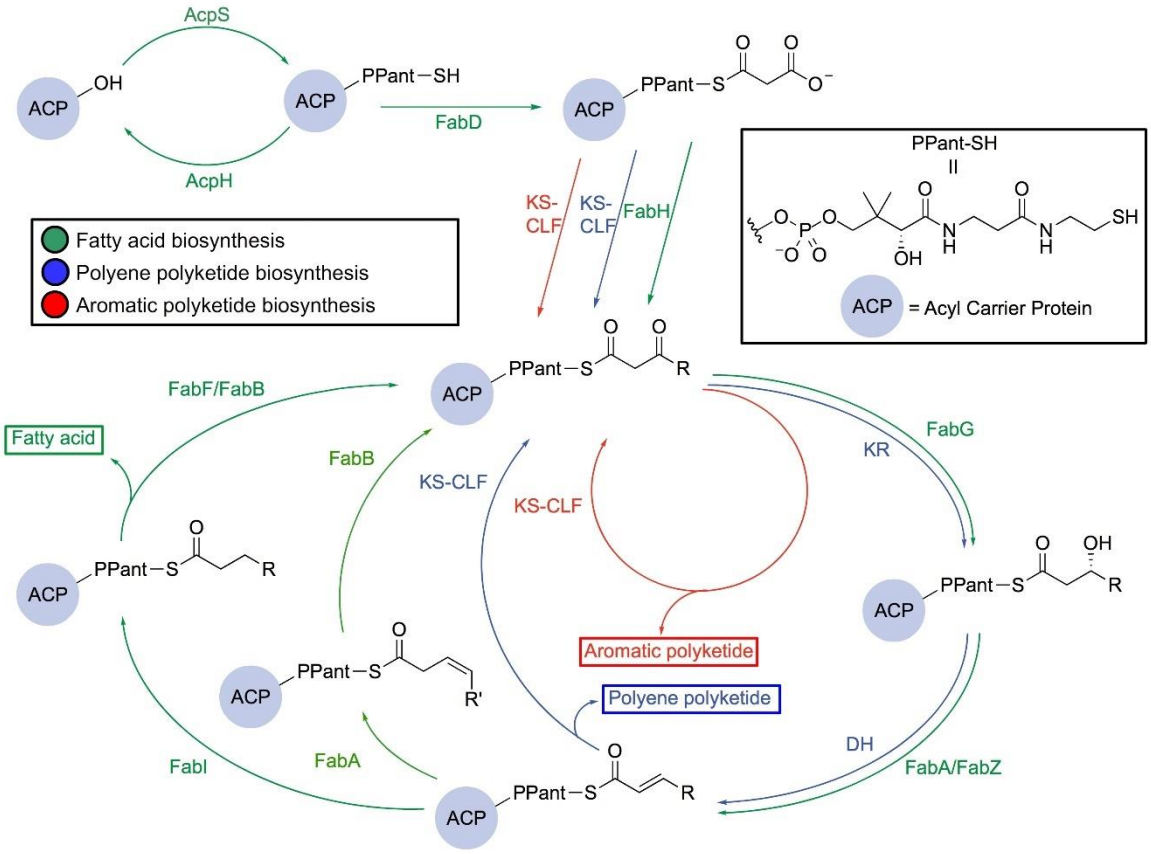


Figure 1.3: Comparison of the type II FAS and PKS pathways. Both the FAS and the PKS pathways begin with utilizing AcpS and FabD to generate its own malonyl-ACP. They start to diverge when entering the elongation cycle, in which the aromatic PKS pathway (red) only requires KS-CLF to elongate polyketones while the polyene PKS pathway (blue) and FAS pathway (green) involve a NADPH-dependent ketoreductase (KR, or FabG in FAS) to reduce the β-carbonyl and a dehydratase (DH, or FabA and FabZ in FAS) to dehydrate the β-hydroxyacyl. The enoyl intermediate of the FAS pathway can subsequently be either reduced by the NADH-dependent FabI (dark green) or isomerized to its *cis* form by FabA (light green) before being further elongated by a ketoacylsynthase (FabF or FabB). The latter pathway (light green) leads to the formation of unsaturated fatty acids.

The most biochemical, structural, and engineering data exist for this species. While several excellent reviews describe bacterial type II FAS in detail,[6,7] here we provide an abridged description as a way to parallel FAS activity with type II PKS. To begin, ACP synthase (AcpS), a 4'-phosphopantetheinyl transferase (PPTase), post-translationally transfers the pantetheine prosthetic group from coenzymeA (CoA) to convert *apo*-ACP into its activated form, *holo*-ACP.[8] This process has been found to be reversible with ACP hydrolase (AcpH), a Mn^{2+}-dependent phosphodiesterase, which hydrolyzes the pantetheine group from *holo*-ACP to regenerate *apo*-ACP,[9,10] presumably in the transition of the organism to stationary phase. Initially, malonyl-CoA is loaded onto the terminal sulfhydryl of *holo*-ACP *via* acyltransferase FabD, also known as the malonylacyltransferase (MAT) or malonyl-CoA:ACP transacylase, releasing CoA to generate malonyl-ACP. Next, FabH, the initiating β-ketoacyl-ACP synthase, a homodimer structure of the thiolase fold, drives the initiation of the elongation cycle by forming an acetyl-enzyme intermediate from acetyl-CoA and then condensing with malonyl-ACP via decarboxylative addition in the active site.[9] This Claisen-like condensation generates acetoacetyl-ACP and serves as the first elongation step. FabH is solely involved in the initiation of fatty acid synthesis, mainly showing activity towards fatty acid substrates with four carbons or fewer. From here, ACP shuttles the intermediates to each of the enzymes in the fatty acid cycle *via* a thioester linkage to the terminal thiol of the pantetheine arm of ACP.

To discuss type II PKS, we will focus on what is known about the *S. coelicolor* actinorhodin biosynthesis, since this is a canonical pathway for which most of the biochemical and structural information is known. The biosynthesis of type II PKSs follows

a similar initiation strategy to FAS, with ACP loading performed by a stand-alone malonyltransferase before condensation. However, the starting unit can be derived from different acyl-CoAs. In the actinorhodin biosynthesis, the acetate starting unit can be derived from the decarboxylation of malonyl-ACP, and other aromatic PKS pathways can incorporate alternative acyl starting units.

Interestingly, it has been found that type II PKSs borrow some of the catalytic machineries from the *de novo* FAS pathway. For instance, actinorhodin actACP is post-translationally modified by the endogenous AcpS PPTase.[11] Further, FabD has been shown to serve as an efficient malonyltransferase for *holo*-actACP.[12]

Elongation in type II FAS. In both FAS and PKS, the acyl ACP continuously proceeds through the iterative pathway, elongating the fatty acid or polyketide chain by two carbons in each full cycle until it has reached its full length. In FAS, the enzyme that acts on acetoacetyl-ACP is FabG, a β-ketoacyl-ACP reductase (KR), that utilizes NADPH to selectively reduce the β-ketoacyl attached to the ACP to form (*R*)-β-hydroxyacyl-ACP with exclusively *R* stereochemistry.[13] This β-hydroxyacyl-ACP is next dehydrated *via* elimination of water by a dehydratase (DH), either FabA or FabZ, the two key DHs in the *de novo* fatty acid synthesis cycle, to form a *trans*-enoyl-ACP. Their role is dependent on the carbon chain length in the pathway and will be discussed below. The *trans*-enoyl-ACP is reduced with NADH by enoyl reductase (ER), FabI. This yields an elongated, fully reduced acyl chain tethered to the ACP to complete the first turn of the FAS cycle.

For the second and subsequent turns, ACP delivers the nascent chain to an elongating ketosynthase (KS), FabB or FabF, thus entering the cycle again by condensing with a

malonate moiety from malonyl-ACP to produce another β-ketoacyl elongated intermediate. This product is then followed by ketoreduction (FabG), dehydration (FabZ or FabA), and enoylreduction (FabI) to complete another elongation cycle. Once the ketoacyl intermediate is reiterated through the cycle to reach its fully elongated length and undergoes enoyl reduction, the fatty acid is then cleaved and released by a thioesterase (TE) or transferred by an acyl transferase (AT) *via* thioesterification onto lipid precursors.[7,14]

Fatty acid chain elongation proceeds in two main routes to produce saturated or unsaturated fatty acids, with a couple of key differences between each path. Monounsaturated fatty acids are biosynthesized through the action of specialized DH (FabA) and KS (FabB) pairs in *E. coli*. The production of unsaturated fatty acids differs from their saturated counterparts through FabA's additional ability to preferentially isomerize *trans*-2-decenoyl-ACP to form *cis*-3-decenoyl-ACP.[15] In the case that a *cis*-enoyl-ACP is present in the chain, FabB acts after FabA to skip the ER step. Here, FabB catalyzes the condensation of malonyl-ACP with *cis*-3-decenoyl-ACP to retain the cis-β,γ-unsaturation. FabF, however, catalyzes the final condensation of the unsaturated fatty acid cycle, since FabF is active towards C4-C16 intermediates and FabB is only active towards C4-C14 fatty acid substrates.[16]

Elongation in type II PKS. Following initiation, the first step of the elongation in type II PKS pathways mimics that of FAS, but the main difference is the replacement of a homodimeric KS with a heterodimeric complex known as the ketosynthase-chain length factor (KS-CLF). The KS-CLF shows homology to FabF; however, the CLF does not bear a catalytic cysteine residue like the KS. The KS-CLF elongates the acyl-ACP to a β-ketoacyl-

ACP, and it is after this first elongation step that the major divergence from FAS is apparent. Rather than immediately reduce the nascent β-ketone after elongation, the KS-CLF instead catalyzes another elongation step (Figure 1.3, red path), an action that is iterated for the *in situ* production of polyketones. In the actinorhodin biosynthesis, seven cycles of KS-CLF elongation occur one after the other, such that a full-length octaketide intermediate is formed. This fundamental change in activity drastically differs from the FAS model, in that the elongating intermediates no longer visit the same active sites in the same way (Figure 1.4). Here, the elongating polyketone never needs to exit the active site of the KS-CLF. This highlights the subtle complexities of the KS-CLF mechanism in retaining and stabilizing inherently unstable polyketone intermediates, a topic that will be discussed below in the evaluation of structural biology.

Following elongation of the actinorhodin octaketide precursor, the first cyclization step occurs between carbons C7 and C12 in an aldol fashion (Figure 1.5). While the catalysis details are still unclear, this step likely occurs before activity by a ketoreductase (KR), which reduces the carbonyl at C9.[17-19] Further aromatization of this first ring is catalyzed by an aromatase (AR).[20] A second cyclization is next catalyzed by a bifunctional enzyme cyclase-thioesterase (CYC-TE), which concomitantly releases the substrate from ACP.[21] The resulting bicyclic intermediate is then modified by post-PKS tailoring enzymes to eventually form actinorhodin.

The elongation chemistries of FASs and PKSs differ significantly from each other due to the distinct traits of the intermediate species. While the acyl chains of elongated fatty acids are hydrophobic and chemically inert, the polyketide intermediates are amphiphilic

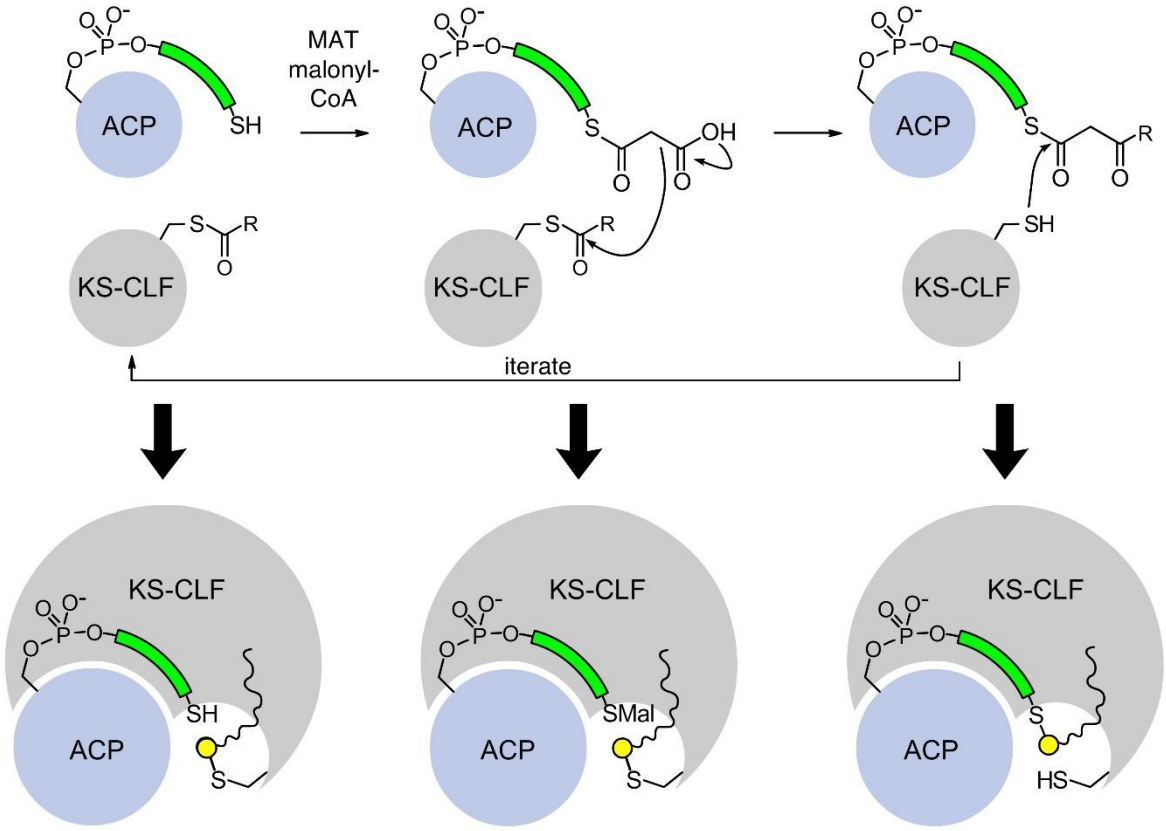


Figure 1.4: The state of polyketones during aromatic PKS iterative elongation. The ACP and KS-CLF only adopt three states during aromatic polyketide biosynthesis: (1) *holo*-ACP and acylated KS-CLF; (2) malonyl-ACP and acylated KS-CLF; (3) β-ketoacyl-ACP and *holo*-KS-CLF. Each of these states involves the elongated polyketone intermediate bound within the KS-CLF, indicating that the KS-CLF must stabilize these highly reactive elongated species *via* yet unknown mechanisms. The elongating polyketone above is depicted as R.

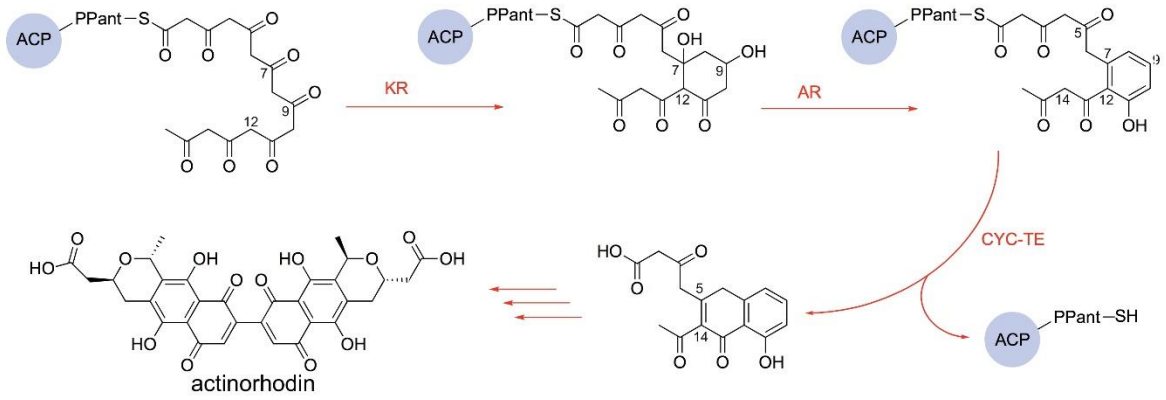


Figure 1.5: The actinorhodin biosynthetic pathway post-elongation. The fully elongated polyketone undergoes a C7–C12 cyclization, and the carbonyl at the C9 position is reduced by the ketoreductase (KR). The aromatase (AR) catalyzes the formation of the first aromatic ring followed by the catalysis of the dual functional cyclase-thioesterase (CYC-TE) that cyclizes the intermediate and unloads it from actACP. The bicyclic substrate is then modified by post-PKS enzymes to form actinorhodin.

and highly reactive. This implies a significant difference in the KS-CLF substrate binding pocket environment as well as the mechanisms of stabilization for the reactive polyketone intermediates to prevent unwanted reactions. While both systems have similar ACPs to shuttle the substrates and intermediates, organisms that possess one or more PKS pathways along with the ubiquitous FAS pathway have certain mechanisms to avert crosstalk between the systems. The protein-protein interactions between ACP and its partner proteins should play a crucial role here.

Type II PKSs had long been considered to only produce aromatic polyketides, but in recent years, pathways that produce polyenes have successively been discovered.[22-24] Similar to the highly reducing type I PKS pathway, the elongation cycle of the polyene PKSs involves a KS-CLF to produce a 1,3-diketone, a KR to reduce the carbonyl at the β position, and a DH to form the α,β-unsaturated ketone (Figure 1.3, blue path). These polyene PKSs are expected to have distinct substrate pockets from the typical type II PKSs. What is evident in the proposed biosynthesis of polyenes is that the mode of binding lies between that of type II FAS and classical aromatic type II PKS, in that the β-ketoacyl product of the KS-CLF must exit the pocket in order to be reduced by the KR and dehydrated by the DH at every turn of the cycle. How the channeling of the substrate differs from the other pathways remains to be explored.

Substrate sequestration. A previous understanding of type II fatty acid biosynthesis is the occurrence of substrate sequestration by the ACP,[25] whereby each acyl intermediate that is C6 or longer is securely bound within the hydrophobic core of the four-alpha helical bundle that makes up the ACP.[26] This phenomenon adds credence to the necessity for each

enzyme to form protein-protein interactions with the ACP as a means to release the nascent fatty acid intermediate into the active site of each enzyme. Upon elucidation of this activity, several groups set about determining the existence of sequestration activity (or lack thereof) in different carrier protein-dependent pathways, and some examples of ACP sequestration in *E. coli* and *S. coelicolor* are illustrated in Figure 1.6. As a result, it has been demonstrated that sequestration occurs primarily in type II pathways, including type II PKS and type II non-ribosomal peptide synthetases.[28,30,31] However, sequestration does not appear to occur in type I FAS or type I PKS, which have large catalytic complexes with different domains in contrast to the discrete proteins in type II systems.[32,33] This distinction makes intuitive sense, because type I synthases have the carrier proteins and catalytic domains within proximal reach, increasing the effective concentration through attachment within the module. Conversely, type II synthases require the ACP-bound substrate to encounter each appropriate enzyme within a pool of potential catalytic partners. Sequestration therefore serves the role of protecting elongating metabolites and delivering them to the appropriate enzyme at the proper time. How these pathways accomplish catalytic specificity remains unclear, and deciphering the role of protein-protein interactions will aid in our understanding of this phenomenon.

Current approaches to study protein-protein and protein-substrate interactions

Many chemical tools have been developed to probe protein-protein and protein-substrate interactions of FAS enzymes, ACP, and the fatty acid intermediates, many of which have been previously reviewed.[34] The same tools can also be applied when studying type II

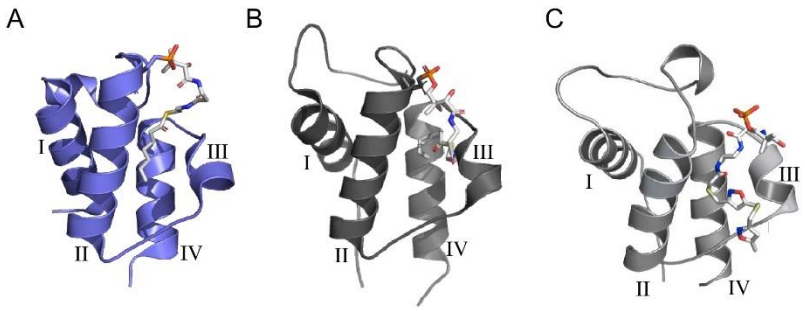


Figure 1.6: ACP sequestration of elongating intermediates. ACPs from both type II FAS and PKS pathways have been shown to sequester intermediates within the hydrophobic cavity at the center of the four-helix bundle. (A) *E. coli* heptanoyl-AcpP (PDB: 2FAD).[27] (B) *S. coelicolor* octanoyl-actACP (PDB: 2KGC).[28] (C) *S. coelicolor* actACP with tethered probe **11**, based on NMR (Figure 1.9B) and *in silico* docking.[29]

PKS to better understand the structure-function relationships of the enzymes involved in the pathway. Structural elucidation to visualize these interactions has also been enabled by using new chemical biology tools and techniques. In conjunction with these tools, manipulation of the enzymes involved in type II FAS and PKS can shed light onto previously undescribed protein-protein and protein-substrate interactions, which are summarized in this section.

Tools to visualize proteins. Structural biology tools such as X-ray crystallography, NMR, and computational methods are leading techniques to visualize protein structure and are regularly employed in studies to probe protein-protein interactions. While these tools are frequently used to complement biochemical studies, this section will focus on their methodology rather than on their application.

Crystallography. X-ray crystallography can be used to visualize proteins at the atomic level and thus serves as a primary tool to elucidate protein structure in structural biology. A concentrated protein sample of high purity is required to form fine crystals, and unpredictable crystallization conditions can cause barriers for protein crystallography. Many FAS and PKS protein structures have been solved in this manner, and additional co-crystallized structures have been solved with a bound substrate to highlight protein-substrate interactions. The information gained from crystallography lacks details about protein dynamics, but combined with other structural tools, such as NMR spectroscopy, valuable knowledge regarding protein structure and function can be obtained to better understand protein interactions.

NMR techniques. NMR serves as a powerful tool to determine protein structure, study the dynamics of small proteins, and investigate protein interactions without prior

knowledge of structural information. While X-ray crystallography relies on the formation of crystals to elucidate protein structure, NMR spectroscopy can yield structural information of a protein in solution or in its solid state, a major advantage to this technique. However, protein preparation and full elucidation of a structure with solution phase NMR can be laborious and require the use of NMR active isotopes. The use of 2D spectra, specifically heteronuclear single quantum coherence (HSQC) NMR spectroscopy, offers a convenient way to solely look at the N-H bond of a protein backbone and observe changes due to protein interactions with substrates or other protein partners.[35] 3D experiments can be performed to correlate the N-H signals with adjacent carbons of the protein backbone to assign the peaks within the HSQC spectra.[36] Often thought of as a protein "fingerprint", HSQC has been used as an ideal method to analyze the effect of titrating increasing concentrations of enzyme partners on the chemical environment of carrier proteins.[37] Here, calculating the chemical shift perturbation (CSP) tracks the HSQC changes for each residue upon binding with an interacting partner. This allows for the identification of key residues that play a role in these interactions, as well as giving insight into the conformational changes that occur. Salt bridge formation between the interacting proteins can also be studied by shifts in the titration curves, aiding in pinpointing the protein binding sites, which can be advantageous in engineering the pathways.[38] NMR spectroscopy can complement X-ray crystallography, because the protein sample conditions can be varied to observe the changes in chemical shifts and their effects on protein activity. This could reduce the barriers associated with crystallization by providing information needed to optimize the conditions for crystal formation.

Computational methods. Computational studies of protein activity are an increasingly important field for studying the FAS and PKS biosynthetic systems, as they can provide invaluable structural details about protein dynamics and predict critical interactions through techniques such as docking experiments and molecular dynamic (MD) simulations.[34] A docking approach is beneficial when simulations are needed to predict how a protein would interact with a substrate or inhibitor prior to its application and determines the favorable binding conformation of the substrate within an enzyme's binding site. Therefore, docking can provide structural clues that shed light onto protein-substrate interactions. MD is a powerful tool that computationally simulates the movements and interactions of atoms as a function of time.[39] By propagating Newton's laws of motion, the trajectories of atoms affected by protein-protein interactions can be evaluated to predict the binding and conformational energies of proteins. Coupled with structural data such as NMR or X-ray crystallography, MD simulations provide a unique glimpse into protein activity.

Manipulation of proteins to probe protein-protein and protein-substrate interactions. The complexity of the proteins in these type II FAS and PKS systems often makes it difficult to study the protein-protein and protein-substrate interactions, but through their manipulation, these interactions can be better understood.

Inhibition/mutagenesis studies. Though inhibition and mutagenesis studies are not new methods for researching proteins, they are continually leveraged to understand protein interactions. Targeting the enzymes of the fatty acid cycle with inhibitors can aid in understanding how the enzymes interact with pathway substrates. While several inhibitors have been reported for type II FAS systems,[40] the best characterized are inhibitors of E. coli

FAS, a model for the highly conserved type II FAS.[41] Further details of these inhibitors can be found in a previous review on the structure-function relationships in FAS.[34] Additionally, mutagenesis studies of specific residues of FAS and PKS enzymes can help to elucidate their role.

While most inhibitors mimic natural substrates or ligands in their interactions within enzyme active sites, covalent inhibitors have the unique ability to trap protein-substrate interactions, often through active site residues, which can then be observed by crystallography or NMR. Cerulenin is a well-studied covalent inhibitor of both PKS and FAS ketosynthases, where the attack by the KS active site cysteine upon the epoxide of cerulenin mimics the condensation transition state, forming an irreversible covalent complex with the enzyme. This interaction was used to illuminate the hydrophobic pocket at the dimer interface of FabF where the acyl chain of the inhibitor and natural substrate lie.[42] Another example is the diazaborine inhibitor of the FAS ER, which covalently bonds to the 2'-hydroxyl of the cofactor's ribose, mimicking the binding of the substrate through hydrogen bonding and hydrophobic interactions.[43] Competitive inhibitors can also be used to illuminate active site chemistry. Triclosan, a synthetic competitive inhibitor of the ER, also interacts with the 2'-hydroxyl of the ribose. However, this inhibitor forms a tight, non-covalent ternary complex with the bound FabI-NAD^+ through hydrogen bonding in the substrate site, demonstrating protein-substrate interactions.[44,45]

Previous studies of mutants of the ketosynthases, reductases, and dehydratases in the FAS cycle have provided insight into the residues that play important roles in the enzyme

mechanisms and their interactions with substrates.[34] Mutagenesis studies of the PKS KS-CLF have also provided understanding of important residues and confirmed the role of the CLF subunit as the determinant of the polyketide chain length.[46] In order to gain a more complete understanding of these systems, a combination of data from several different tools, including mutagenesis, inhibitors, and structural biology techniques, is required to highlight different protein-protein interactions.

Crosslinking methods to probe ACP-partner protein interactions. Bifunctional probes can be used to crosslink proteins with a binding protein as a means to capture protein-protein interactions. Such crosslinking probes are designed to tether enzymes together using many techniques: they can leverage mechanism-based inactivators, known inhibitors, or non-selective agents. Given that the ACP interacts and binds with its partner enzymes in each step of catalysis, crosslinking has been used to help evaluate protein-protein interactions in FAS, and the same methods can be applied when studying PKS. Used effectively, crosslinking acts as one of the main ways to capture these transient interactions.

Mechanism-based crosslinkers have led to advancements in understanding interactions of FAS proteins in complex with ACP. Analogues of the natural pantetheine arm appended to the ACP are prepared synthetically and then chemoenzymatically attached to the ACP using a "one pot" modification strategy.[47] Briefly, a pantetheine probe can be synthesized and converted by enzymatic transformation to a full coenzyme A analog, which is then coupled to a conserved serine of ACP by a phosphopantetheinyl transferase.[47,48] Commonly designed crosslinking probes contain a terminal warhead which, once tethered to ACP, can react with a partner enzyme to covalently bond and form a crosslinked complex.

Crosslinkers often mimic the natural substrate so that both protein-protein and protein-substrate interactions can be observed.

Different mechanism-based crosslinkers have been studied to probe ACP interactions with a few partner enzymes in the FAS system, but many FAS and PKS enzymes still lack elucidated crosslinked structures. Examples of these pantetheine analogues include epoxides based on the cerulenin inhibitor and simple acrylamide analogues, such as chloroacrylate (Figure 1.7A), that could undergo an addition-elimination reaction by a nucleophilic active site residue, such as the cysteine in KS.[49] Another example is the sulfonyl 3-alkynyl pantetheinamides (Figure 1.7A) that were used to crosslink ACP with the FAS DH, FabA. The active site residue, His70, of FabA is proposed to deprotonate the α-hydrogen of the sulfone to form an allene intermediate, which then interacts with the same residue to form the crosslinked complex (Figure 1.7D).[50,51] This crosslinker was first shown to best mimic the native substrate (*R*)-β-hydroxydecanoyl-ACP using fluorescent labeling experiments.[51] The interactions observed from this crosslinking will be discussed in detail in the "Protein-protein interactions" section.

Another recent study shows how fluorescent, solvatochromic fluorophores can be used to visualize ACP sequestration and released into partner enzyme active sites through dual fluorescent-crosslinking probes (Figure 1.8A).[52] While they do not mimic the native substrates, the small dyes fit into the hydrophobic pocket of ACP where cargo is generally sequestered. Once the probes were loaded onto ACP via the one pot chemoenzymatic reaction and reacted with the active site cysteine residue of FabF through a reaction with the α-bromo moiety of the probe, a crosslinked complex was produced. An increase in

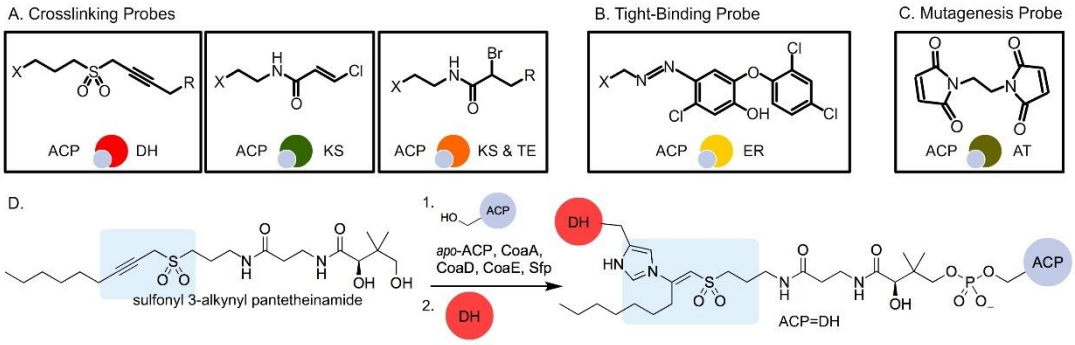


Figure 1.7: Probes for the investigation of PPIs. (A) Mechanism-based crosslinking probes. X = pantetheine analog; R = carbon chain. (B) The tight-binding probe for ER. Y = linker. (C) The BMOE probe. (D) Sulfonyl-3-alkynyl pantetheinamide undergoes chemoenzymatic one pot reaction and crosslink reaction to form crosslinked complex with DH.

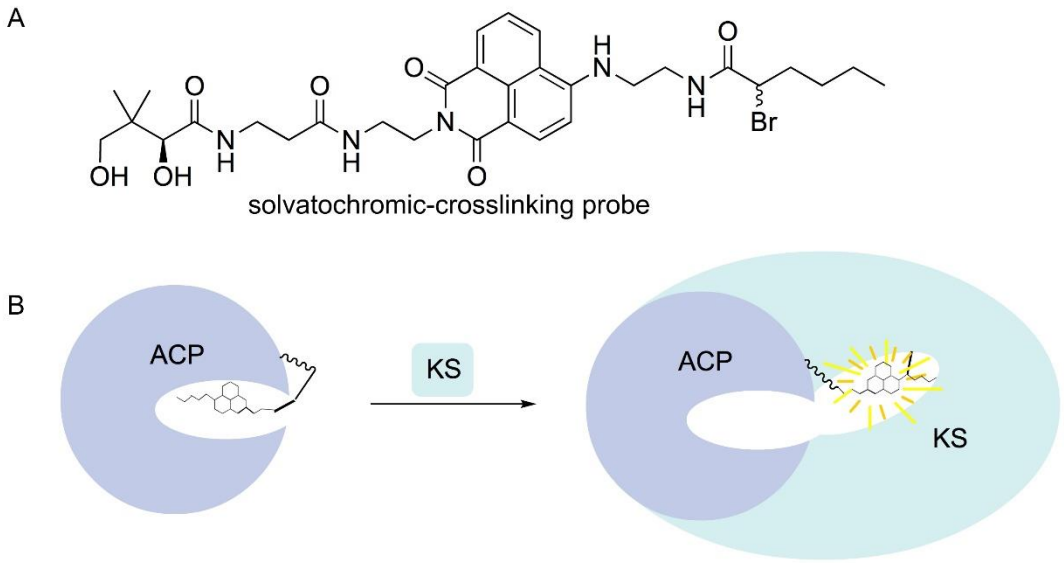


Figure 1.8: Solvatochromic-crosslinking probe. (A) The probe. (B) Flipping of the probe from the pocket of ACP into the pocket of KS occurred when the two proteins are bound, causing an increase in fluorescence. This is known as the chain-flipping mechanism.

fluorescence upon crosslinking of the *crypto*-ACP, which has a loaded pantetheine-probe, with FabF was believed to confirm the chain-flipping mechanism in which the probe leaves the ACP pocket and extends into the active site of FabF (Figure 1.8B). These solvatochromic pantetheine probes could be used for sequestration studies of the PKS ACP, as well as investigations of the protein-protein interactions with other enzymes in the FAS and PKS cycles.

Tight-binding probes have also been developed for interaction studies. Triclosan (Figure 1.7B) forms a tight complex with the FAS ER active site, and when appended to the pantetheine on ACP, it illuminates the protein-protein interactions. Details of these interactions will be further described in "The FabI-AcpP co-crystal structure" section.

Non-selective agents are another crosslinking tool that have been tested in type I PKS but can be applied to type II systems. Vicenistatin is a type I PKS product in which an in-trans AT is responsible for substrate loading. Crosslinking was employed using a bifunctional maleimide reagent, 1,2-bismaleimido-ethane (BMOE), to trap the VinK(AT)-VinL(ACP) complex (Figure 1.7C).[53] BMOE is a non-selective crosslinker that reacts with any free thiol. To trap VinK, Ser266 of VinK, a non-catalytic residue located at the base of the substrate-binding tunnel, was mutated to cysteine to enable the crosslinking reaction. The sulfhydryl of the mutated VinK cysteine reacts with one of the maleimide groups of BMOE, while the terminal sulfhydryl on the phosphopantetheine arm of VinL reacts with the other maleimide group to form a crosslinked complex. The crystal structure confirmed that residues Arg153, Met206, and Arg299 of VinK are essential for the salt bridges and hydrophobic interactions with VinL, similar to the interactions seen in FabA and ACP in the

FAS system.[54] This reagent could be useful when crosslinking different proteins to obtain protein-protein interactions due to its non-selective nature; however, it is important that a sulfhydryl group is present for crosslinking to occur. Cysteines present at other locations may be prone to reacting with this reagent and may need to be mutated to prevent this.

While few crosslinked complexes exist for the enzymes of FAS and PKS, we expect to see a lot of progress in the coming months and years to be made in this field to open a window into the protein-protein and protein-substrate interactions of these systems.

Atom replacement strategy for type II PKS. The inherent instability of polyketone intermediates in type II PKSs poses challenges for examining its mechanism. One way to overcome this is by using atom replacement strategies, in which polyketone surrogates that mimic the natural intermediates in length, polarity, and hydrophobicity are used to provide insight into the biosynthesis of the PKS products.[29, 55-57] Such strategies can be accomplished by substituting the carbonyl group with carbonyl bioisosteres that possess similar size or characteristics.[58]

A recent atom replacement approach uses an oxetane-based polyketide mimetic (**8**) that was demonstrated to probe the substrate binding and mechanism of the priming KS, DpsC, of the daunorubicin biosynthesis (Figure 1.9A).[56] In this study, oxetanes, a well-known carbonyl bioisostere, replaced the carbonyl of the thioester to mimic malonyl-phosphopantetheine.[56] Although oxetane is a slightly larger group compared to carbonyls, the lone pairs of the oxygen in both groups are positioned similarly, providing evidence for the use of the oxetane mimetic. Co-crystallization was performed with DpsC and the oxetane probe to characterize the substrate interaction with the priming KS. MD simulations on the

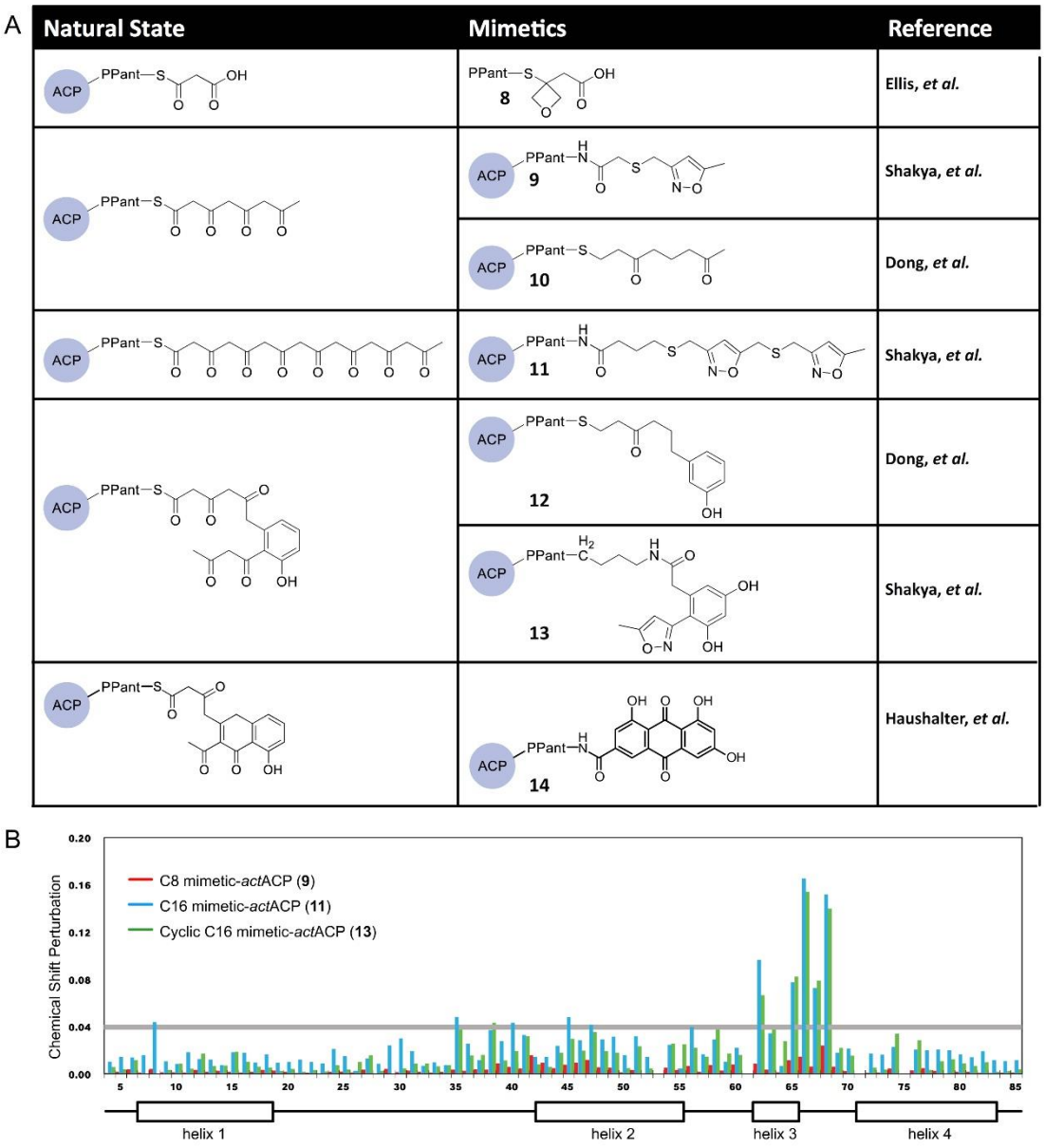


Figure 1.9: Atom replacement strategy. (A) Different carbonyl isosteres are utilized to generate stable mimetics in order to probe the protein-substrate interactions of PKSs. (B) Chemical shift perturbation plot of *crypto*-actACP using the chemical shift of *holo*-actACP as the base. The tetraketide mimetic shows no significant perturbation on actACP while the octaketide and cyclic polyketone mimetics show strong perturbation in helices II and III, suggesting that only the latter two substrates are sequestered by actACP.[41]

oxetane replacement probe and the natural substrate, malonyl-phosphopantetheine, show comparable binding affinities and movements of binding site residues, confirming that the mimetic is similar enough to provide mechanistic insights of DpsC.[56]

Another atom replacement strategy involves replacing the carbonyl and diketide units that are not involved in the cyclization steps with thioethers and isoxazoles, respectively. Figure 1.9A shows the linear and cyclic atom replacement probes that have been synthesized with these moieties to mimic the native substrates and loaded onto actACP *via* chemoenzymatic methods.[47] While no probe can exactly mimic the natural substrate, this atom replacement strategy acts as a useful tool to allow the study of protein-substrate interactions of ACP. Solution-phase protein NMR was applied to study how these isoxazole atom replacement probes that are loaded onto ^{15}N-labeled actACP (**9, 11, 13**) compare to *holo*-actACP. As shown in the CSP plot collected (Figure 1.9B), significant chemical perturbations of actACP are present with the octaketide and cyclic polyketone mimetics between helices II and III. This suggests sequestration of the substrates by actACP in a similar fashion to FAS ACP.[29,30,55] Docking experiments were also used to simulate the same atom replacement probes to confirm that sequestration only occurs with the fully elongated mimetics.[29]

In addition to the isoxazole atom replacement probes, other atom replacement strategies have been developed to probe the actPKS pathway, including strategically removing carbonyl units to produce mimetics of the linear octaketide and cyclic PKS substrates,[55] as well as using the tricyclic emodic acid to study the PKS pathway.[30] These mimetics are loaded onto ACP (Figure 1.9A) and all three replacement strategies confirmed

via HSQC NMR experiments that the linear polyketide and the bicyclic intermediate of actPKS are sequestered between helices II and III of actACP, demonstrating the efficacy of using this strategy.

Important findings from structural biology and mutational analysis

In the past two decades, extensive progress has been made on the structural study of type II FASs. The structures of almost all E. coli FAS enzymes have been studied by X-ray crystallography, which provides the base for mechanistic study.[13,42,59-64] Co-crystal structures of proteins with native substrates provide additional information of substrate sequestration and give more solid evidence of reaction mechanism. Excellent reviews by White *et al.* and Finzel *et al.* summarize these accomplishments.[6,34] In contrast to the FASs, few crystal structures of type II PKSs exist.[17,65-67] This is due in part to the instability of PKS proteins and the high reactivity of the PKS intermediates. This section will focus on what is known about the structural biology of type II PKSs as well as the findings from mutational studies.

The priming ketosynthase. In FAS, the elongation cycle starts with the formation of acetoacetyl-ACP from malonyl-ACP by a priming KS, FabH. However, there are three possible types of priming mechanisms in type II PKS, and only one of them involves a priming KS.[68] The two priming KSs that have been structurally characterized are ZhuH and DpsC, from the antibiotic R1128 and daunorubicin biosynthesis, respectively.[56,65] ZhuH and FabH are structurally similar to each other but have very distinct substrate specificities. Both ZhuH and ecFabH (from *Escherichia coli*) uptake substrates with C2-C4 chain lengths, but

only ZhuH can sequester branched chain substrates. More uniquely, the type II mtFabH (from *Mycobacterium tuberculosis*) uptakes long chain fatty acids generated by a type I FAS and can tolerate substrates with C8-C20 chain lengths.[69] Crystal structure overlays of ecFabH and ZhuH, as well as mtFabH sequestering its natural substrate lauroyl-CoA, reveal that the gatekeeper residue is Phe87' from another subunit in ecFabH and Met90 in ZhuH (Figure 1.10A).[65] Compared to the phenylalanine residue, the more flexible side chain of methionine may contribute to the ability of ZhuH to sequester branched acyl-CoA. In addition, the rotamer conformation of a highly conserved phenylalanine residue (Phe304 in ecFabH and its equivalent in other priming KSs) located at the base of the CoA binding cleft determines the shape of the pocket, thus contributing to the substrate specificity.[70]

The priming KS DpsC catalyzes the transfer of a propionyl group from propionyl-CoA to the active site serine. It subsequently interacts with malonyl-ACP to initiate the cycle. This process involves an intermediate state in which a malonate is in the proximity of the propionated serine but was challenging to characterize by X-ray crystallography or other structural biology methods. Previous attempts of visualizing the same intermediate state of ecFabH by co-crystallizing with malonyl-CoA failed due to the instability of the malonyl group.[71] To overcome this barrier, a recent work replaced the thioester carbonyl with an oxetane ring (Fig. 1.9A and 1.10B).[72] The resulting mimetic was successfully co-crystallized with DpsC, yielding the structure of the intermediate state. Given that DpsC has a Ser-His-Asp catalytic triad instead of the Cys-His-Asn triad in canonical priming KSs and that the structure of DpsC is quite unique, it is not a surprise to see that the substrate is oriented differently in the binding pocket compared to the hypothesized FabH orientation. A key

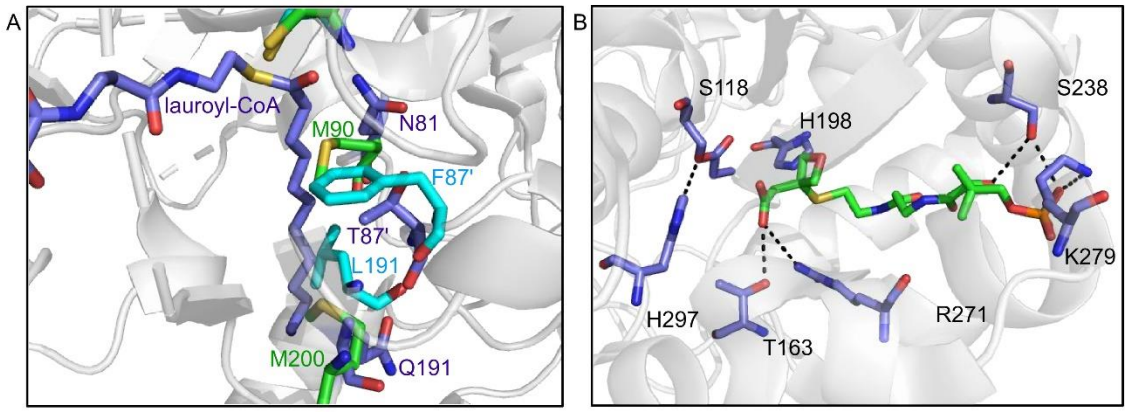


Figure 1.10: The binding pocket of priming KS. (A) Overlay of the crystal structures of ZhuH (green), ecFabH (light blue), and mtFabH with its substrate lauroyl-CoA (purple) reveals the gatekeeper residues M90 (ZhuH) and F87' (ecFabH). The phenylalanine rotamers that influence the pocket shape are also shown. (B) Malonyl-CoA mimetic (green, **8**) in the DpsC binding pocket. The oxetane is not involved in any hydrogen bond network. It is proposed that the thioester carbonyl will turn and form a hydrogen bond with H198 after decarboxylation.

difference is the lack of an oxyanion hole that stabilizes the thioester carbonyl during decarboxylation in FabH. It is hypothesized that the substrate reorients itself after decarboxylation to form a hydrogen bond with His198. The success of this carbonyl surrogate points out a new road to investigating the catalytic mechanism of type II PKSs.

Chain length control by the extending ketosynthase. It is hypothesized that the reactive polyketone intermediates in the elongation cycle of the aromatic PKS pathway remain in the pocket of KS-CLF during all steps of elongation (Figure 1.4). The only crystal structure of KS-CLF to date is the actKS-CLF (Figure 1.11B).[17] This heterodimer is structurally similar to the homodimeric extending KS of FAS, such as FabF (Figure 1.11A). The substrate pocket can be identified in the structure, which extends from the catalytic Cys169 of the KS subunit to the Phe116 of the CLF subunit. Surface view of the tunnel shows a narrow and amphipathic substrate binding pocket, suggesting that the pocket has a high specificity for the substrate and that a strong sequestration interaction is expected. Despite the fact that the CLF can act as a decarboxylase,[73] it is primarily known to control the chain length of the substrate through its gatekeeper residues.[46] Mutations on these residues, suggested by the sequence alignment of several CLFs, have successfully engineered some KS-CLFs to produce substrates of different chain lengths. The F109A/F116A double mutant of actCLF produces decaketides (C24) that make up 96% of total polyketides *in vitro*, confirming that these phenylalanine residues (in gate 1 and 2 of Figure 1.11C) prevent the natural substrate, octaketides (C16), from extending further. In addition, when Gly116 (Figure 1.11C, gate 1) of tcmCLF is mutated into threonine, a larger residue, the mutant synthesizes more octaketides than nonaketides (C20). Combining these

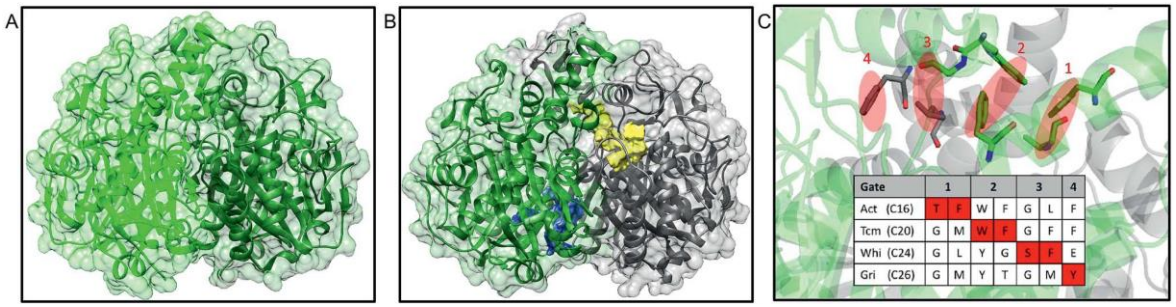


Figure 1.11: The structural information of extending ketosynthase. (A) Crystal structure of FabF (PDB: 2GFW). (B) Crystal structure of actKS-CLF with KS colored in gray and CLF colored in green (PDB: 1TQY). The substrate pocket (yellow) extends from the KS subunit into the CLF subunit. Another pocket with unknown purpose is found in the CLF subunit (blue). (C) Four gates in KS-CLF that determine the chain length, using actKS-CLF crystal structure as an illustration (KS in gray and CLF in green). The bulky gatekeeper residues are colored red in the protein sequence of different KS-CLFs. Act = actinorhodin, Tcm = tetracenomycin, Whi = WhiE spore pigment, and Gri = griseorhodin. The parenthesis after each KS-CLF indicates the chain length of the respective polyketide product. The substrate extends from the right to the left of the figure, thus, the earlier the gate is blocked, the shorter the product is.

results with the crystal structure of actKS-CLF reveals the gates for chain length control (Figure 1.11C).

The FAS ACP is known to sequester the acyl intermediates when shuttling them between the partner proteins. This sequestration serves as one of the driving forces of the elongation cycle. In the aromatic PKS system, on the other hand, ACP only interacts with one enzyme, KS-CLF, during elongation. actACP sequesters its substrate between helix II and III, and it has been shown that the sequestration happens only after the polyketone intermediate is fully elongated.[29,55] This result, in combination with the tight binding pocket of KS-CLF, supports the hypothesis that KS-CLF keeps the substrate inside its pocket during elongation.

The role of ketoreductase in the first ring cyclization. The KR of type II PKS shows regiospecificity by reducing only one carbonyl group of the elongating polyketone. In the case of actKR, such specificity is at the C9 carbonyl (Figure 1.5). Comparison of the actKR structure with FabG, the FAS KR that reduces each β-carbonyl group of a growing fatty acid chain, reveals a flexible region around helices VI and VII in FabG that contributes to the tolerance of substrates with different chain lengths.[18]

An important inquiry of the actinorhodin biosynthesis is the timing of the first ring cyclization. Early studies that led to the "design rule" of engineering type II PKSs show that reduction at the C9 position by actKR drives the first cyclization to occur between C7 and C12, implying that the cyclization happens after the reduction.[74] However, docking simulation of the full length linear substrate in the actKR binding pocket shows a loss of regiospecific reduction at C9, indicating that the cyclization is more likely to happen before

the reduction.[18] Further docking experiments of the cyclic substrate also suggest that constraints from the ring are necessary to position the substrate for C9 reduction.[75] These results show a direct causal relationship between C7 – C12 cyclization and C9 reduction, but with the former activity as the cause.

However, the first ring cyclization in aromatic PKSs can be mediated by either the KS-CLF or the KR. Given that the minimal actPKS (actACP and actKS-CLF)[76] only produces C7-C12 and C10-C15 cyclic compounds (SEK4 and SEK4b, respectively)[77] instead of the many other possible structures, it is proposed that the KS-CLF can catalyze the ring formation. The crystal structure of the actKS-CLF shows a water molecule in close proximity to the C7 carbonyl when a substrate is bound. It is possible that this water molecule donates a proton to the substrate and catalyzes the C7-C12 ring formation. The pocket also has room to accommodate the cyclized substrate.[17] However, it has been shown that the addition of actKR to the minimal tcmPKS affects the regiospecificity of cyclization, indicating that KR could play a role in the ring forming catalysis.[78] Given that the actACP sequesters the linear polyketone,[29] it is possible that the unstable intermediate can be protected and transferred to the actKR to be cyclized. Mutational studies of actKR reveal the highly conserved residue T145 that is necessary for mediating cyclization. It has been shown that T145 can form a hydrogen bond with the C11 carbonyl and subsequently facilitates the nucleophilic attack of C12.[78] The T145A mutant fails to direct cyclization, leading to the loss of ability to form mutactin, a molecule produced by co-expressing minimal PKS and actKR. As a result, the actKR catalyzes the first ring cyclization, whereas the actKS-CLF may contribute to it in some degree.

Protein-protein interactions

In almost all type II systems, the ACP is a 4-helical protein of about 10 kDa that shuttles substrate intermediates between each enzyme in the pathway. The remarkable ability for such a small protein to be recognized by and interact with 12 partner enzymes in the *E. coli de novo* fatty acid biosynthesis and at least 9 others from peripheral pathways belies a subtle communication strategy that we are only beginning to understand. The prevailing hypothesis is based upon specific salt bridges and hydrophobic interactions between the ACP and each enzyme, which serve to coordinate communication and control. Sequestration of substrates within the ACP results in small changes in the surface residues of the ACP, thus varying the binding specificity with pathway enzymes. It is believed that these protein-protein interactions accelerate the overall kinetics of the metabolic pathway by avoiding stochastic binding and sampling of each metabolite by every enzyme. At the same time, this mechanism prevents crosstalk from non-cognate enzymes and other ACPs.

To visualize such interactions *via* crystallography, it is necessary to trap the dynamic ACP as it interacts with partner enzymes. This can be accomplished by co-crystallization of unmodified ACP or by a crosslinking method. To date, only four crystal structures of ACP-bound FAS have been reported, and no such structures for type II PKSs exist. The following section will focus on the existing crosslinked/co-crystal structures of FASs and the information drawn from these studies about protein-protein and protein-substrate interactions.

The FabA=AcpP crosslinked complex. FabA, a DH of FAS, is known to control fatty acid chain length and saturation levels.[79] It catalyzes the dehydration of β-hydroxy-

decanoyl-ACP as well as the isomerization of the subsequent *trans*-2-decenoyl-ACP into *cis*-3-decenoyl-ACP. In *E. coli*, this is the only pathway for the biosynthesis of unsaturated fatty acids. The FabA catalysis involves a His-Asp dyad in which the His70 serves as a base to first remove the α-proton from C2. The enolate is delocalized *via* conjugation with the adjacent carbonyl, followed by elimination of a protonated β-hydroxyl group. This mechanism led to the discovery of the first reported mechanism-based inhibitor and subsequent design of the sulfonyl alkyne crosslinker.[80] Depicted in Figure 1.7D, this moiety targets the His70 residue of FabA, which catalyzes deprotonation and forms a covalent bond with the β-carbon. The crosslinked complex in Figure 1.12A, left, shows the FabA homodimer bound with two ACPs, one covalently bound to each active site His70.[54] Interestingly, the two FabA=AcpP protomers are not identical, possibly representing different stages in catalysis (Figure 1.13). NMR titration of increasing the ratio of FabA to octanoyl-AcpP shows a migration of cross peaks toward those peaks identified by the HSQC of the crosslinked complex, indicating that the bound conformations of ACPs represent the most accurate depiction of the natural, transient binding event.

The NMR titration study also identified important residues in helices II and III of AcpP that interact with FabA, which is in accurate agreement with those found in the crosslinked complex. Along with the interacting residues of FabA revealed by the crystal structure, a series of binding events can be proposed to mediate the transition of substrates (Figure 1.13). First, the arginine-rich positive residues (Figure 1.13, light blue) on the FabA surface interact with the anionic phosphopantetheine group attached to ACP. Subsequent mutagenesis studies of these surface residues show significantly lower efficiencies in

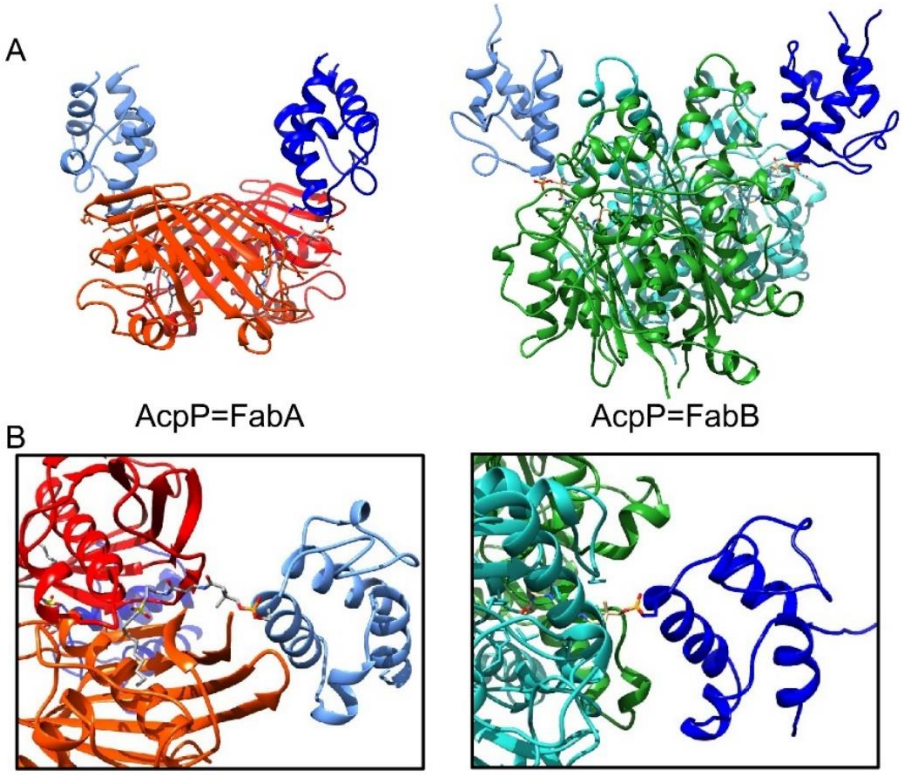


Figure 1.12: Crystal structures of crosslinked complexes. (A) Comparison of the FabA=AcpP (PDB: 4KEH) and FabB] AcpP (PDB: 5KOH) complexes. Both homodimers have two ACPs attached at the entrance to each active site, located at the interface of the monomers. However, the FabA=AcpP complex shows two different binding events and are distinguished by numbering. (B) Comparison of the protein-ACP interfaces.

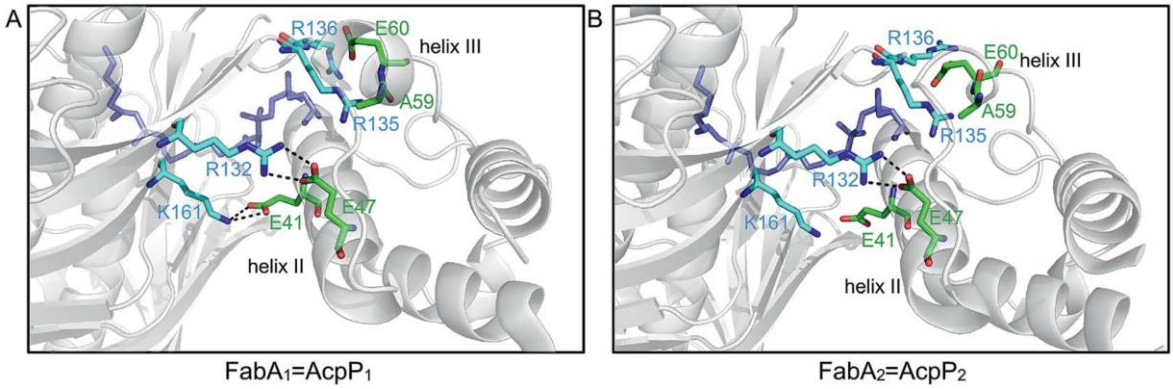


Figure 1.13: Important residues on the protein interface of the FabA=AcpP crosslinked complex. FabA residues (light blue) interact with the residues (green) in helices II and III of AcpP in order to transfer the substrate (purple). (A) The $FabA_1$=$AcpP_1$ protomer and (B) the $FabA_2$=$AcpP_2$ protomer show different interface interactions. It is proposed that they represent different stages in catalysis.

crosslinking, and thus incomplete binding.[81] Next, Arg132 and Lys161 of FabA form salt bridges with Glu41 and Glu47 of ACP, respectively, to stabilize the complex. Thirdly, helix III of ACP is pried open by the interaction of FabA surface arginines with the Ala59 and Glu60 residues of ACP. Finally, hydrophobic interactions between FabA and ACP helix II further anchor the complex. This process demonstrates the course of the chain-flipping mechanism, in which the protein-protein interactions flip the substrate out from sequestration within the ACP to enter the enzyme partner's active site for subsequent catalysis.

Further comparison of the crosslinked and the *apo*-FabA structures reveals two gatekeeper residues, Phe165 and Phe171, located at the entrance of the active site.[81] Mutations of these residues to alanine residues eliminate the ability of FabA to act selectively with substrates of different acyl chain lengths, confirming their gatekeeper role. Rotation of Phe171 can partially open the entrance to accept small molecule substrates, such as the inhibitor 3-decynoyl-N-acetyl cysteamine.[59] With the *crypto*-ACP bound, however, Phe165 also must turn away to fully open the gate for sequestration.

The FabB=AcpP crosslinked complex. Recently, the crystal structure of a crosslinked *E. coli* FabB=AcpP has been deposited (Figure 1.12A, right).[82] The crosslinking structure has been accomplished through the use of a chloroacrylate crosslinking probe (Figure 1.7A), with both monomers of the FabB homodimer capturing an ACP via covalent linkage to the active site Cys163. While a detailed study of this structure remains to be published, a simple comparison with the FabA=AcpP crosslinked structure can be made by fixing the positions of ACP in each (Figure 1.12B). While both structures represent

homodimers that are crosslinked to two ACPs, the binding motifs are surprisingly different. For instance, in FabA=AcpP, the FabA monomer that pries away helix III to release the substrate from sequestration is not the monomer that is covalently crosslinked to ACP. Conversely, in FabB=AcpP, the crosslinked monomer is the same, which engages helix III. Such surprising comparisons will likely be revealed as more crosslinked structures emerge.

The FabZ-ACP co-crystal structure from *Helicobacter pylori*. FabZ is the second DH from the bacterial FAS pathway shared with *E. coli* and lacks the ability to catalyze isomerization, but it can accept substrates with a larger variety of chain lengths. Unfortunately, efforts to overexpress FabZ from *E. coli* have failed to date to produce soluble protein. The recent co-crystal structure of a *H. pylori* FabZ homodimer with a single *holo*-ACP reveals that the binding of ACP induces the movement of the β-sheet layer, resulting in a Y-shaped tunnel that contains an additional sequestration site.[83] Further docking experiments using fatty acids with different chain lengths show that shorter substrates (C4-C12) prefer the original pocket I, whereas longer substrates (C14 and C16) prefer the temporal pocket II. This induced Y-shaped tunnel partially explains the difference in substrate specificity of FabZ and FabA, with FabA still possessing the narrow L-shaped tunnel even with ACP bound, thus sequestering substrates with chain lengths only up to C12. Docking of an unsaturated fatty acid shows that the Y-shaped tunnel sterically hinders a *cis*-substrate, explaining the lack of isomerization ability for FabZ.

The solution of the FabZ-ACP co-crystal structure with only one bound ACP offers a curious phenomenon in the company of several 2:2 structures. Further examination of the structure reveals that binding of an ACP induces the unbound monomer to form a closed

conformation that prevents the second ACP from binding. This leads to the proposal of the seesaw-like mechanism in which two subunits alternate to bind ACP, although crystal packing could also explain this unusual stoichiometry. In addition, Tyr100 is identified as a gatekeeper residue working similarly to those recognized in FabA.

The FabI-AcpP co-crystal structure. FabI, the ER of FAS, is responsible for reducing the enoyl-ACP to its acyl-ACP form, depending on its binding with the NADH cofactor. While no crosslinked complex yet exists for FabI, many tools, such as co-crystallization, inhibition, and MD simulations, have been employed to study the protein interactions with ACP. An earlier co-crystal structure of FabI with dodecanoyl-ACP showed its structure to be a tetramer with each monomer having a central β-sheet that contains eight helices surrounding seven β-strands.[84] The crystallography data showed that once FabI and ACP interact, FabI's structure does not change drastically, but the substrate binding loop does change, becoming more ordered. Additionally, the basic residues of FabI's helix VIII form hydrogen bonds with the acidic residues in helix II of ACP, stabilizing the structure. Mutations of these basic residues significantly decrease the enzyme's ability to reduce the dodecanoyl-ACP, confirming the importance of the electrostatic interactions of these basic residues with ACP's acidic residues. Additional stabilization is provided by hydrogen bonding of the ACP pantetheine with FabI.[84] The transient nature of this interaction caused the crystal structure to not be fully resolved, and so the same group turned to MD simulations to observe the interactions of the proteins and substrate. While the catalytic triad of FabI is made up of Tyr146-Tyr156-Lys173, the MD simulations, along with mutagenesis studies, suggest that the active site residue Tyr156 does not directly play a role in stabilizing the

substrate's carbonyl group through hydrogen bonding. Rather, Tyr146, due to its closer proximity to the substrate, helps to catalyze the reduction of the substrate by hydrogen bonding with its carbonyl group.[84] From these studies, it is inferred that the active site serine of ACP carries the substrate to the active site of FabI, consisting of helix VIII, the substrate binding loop, and a mobile loop of five residues. The latter loops move and become ordered to allow the phosphopantetheine arm of ACP to transfer the substrate to FabI's active site.[84,85]

A recent inhibition study using tight-binding probes to examine the FabI and ACP interactions was developed by chemoenzymatically modifying an *apo*-ACP to tether to triclosan, generating crypto-ACP. This selective inhibitor of bacterial FAS ERs was linked to the 4'-phosphopantetheine arm of ACP (Figure 1.7B).[85] Because of the known competitive inhibition of FabI by triclosan, the inhibitor was used as the ER specific motif to provide tight binding of the inhibitor within the enzyme active site. Data indicated that the binding of ACP to FabI only occurs when triclosan was present on the pantetheine arm and presented greater inhibition as a *crypto*-ACP probe than as a probe without ACP. This shows how important the binding of ACP to FabI is.[85] In the same study, the probe was loaded onto the PKS actACP to test E. coli FabI's specificity toward AcpP. The *crypto*-actACP with the triclosan probe shows no inhibition of FabI by itself, confirming the necessity of *crypto*-AcpP to observe the protein-protein interactions. This design can prove useful when studying other enzymes in the type II FAS and PKS systems to understand the key interactions of ACP with its partner proteins.

Future directions

The more we understand about protein-protein and protein-substrate interactions in type II FAS and PKS biosynthesis, the sooner we will be capable of designing these pathways for new applications. Understanding the fundamental principles underlying ACP interactions with substrates and enzyme partners will inform a set of rules guiding their processivity in these pathways. Such knowledge will direct efforts to modify and engineer these pathways for the production of novel metabolites tailored to specific applications.

FAS engineering provides an immediate example. To date, groups have struggled with metabolic engineering of type II FAS due to the lack of knowledge about protein interactions. Indeed, most published articles provided anecdotal rationales for success or failure upon expression of exogenous FAS enzymes in model organisms.[86] For instance, in an effort to engineer increased production of medium-chain fatty acids for improved biodiesel feedstocks, researchers expressed exogenous medium chain fatty acid TEs from plants into *E. coli*, resulting only in increased production of short chain fatty acids.[87] Despite multiple attempts, medium chain fatty acid production remained elusive until the recent rational mutagenesis of these TEs to better interact with the E. coli ACP.[88] Indeed, similar challenges have been faced in other organisms with type II FAS, such as micro-algae,[89] pointing to immediate opportunities to engineer the protein-protein interface of ACP with exogenous enzymes as a durable solution to the synthetic biology challenges in these pathways. We can expect to see significant contributions in this area very soon. Crosslinked complexes of the E. coli ACP and most, if not all, of its enzymatic partners will likely be elucidated within coming months and years, and this information will help inform next-

generation FAS engineering efforts. Likewise, type II PKS pathways are sure to experience a parallel increase in activity. Enabled by crosslinking and structural biology tools, the rules governing protein-protein interactions in aromatic PKSs will become clearer, and these pathways will become more available for modification and engineering.

Conclusion

The details of type II FAS and PKS biosynthetic pathways continue to be uncovered. New developments in structure, activity, and protein-protein interactions continue apace, as does the development of new tools to probe these pathways. We are hopeful that the coming decade will offer significant advances in our understanding of these pathways, where protein-protein and protein-substrate interactions will no doubt play a central role. With the underlying rules and mechanisms in place, manipulation of these pathways will rapidly follow.

Acknowledgement

Chapter 1, entitled Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions, in full, is a reprint of the material as it appears in: Chen A, Re RN, Burkart MD. “Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions”, Natural Product Reports, vol. 35, 2018. The dissertation author is the primary author of this manuscript.

References

(1) Griffin, M. O.; Fricovsky, E.; Ceballos, G.; Villarreal F. Tetracyclines: a pleitropic family of compounds with promising therapeutic properties. Review of the literature. *Am. J. Physiol.: Cell Physiol.* **2010**, *299*, C539–C548.

(2) Minotti, G.; Menna, P.; Salvatorelli, E.; Cairo, G.; Gianni, L. Anthracyclines: molecular advances and pharmacologic developments in antitumor activity and cardiotoxicity. *Pharmacol Rev.* **2004**, *56* (2), 185–229.

(3) Das, A.; Khosla, C. Biosynthesis of Aromatic Polyketides in Bacteria. *Acc. Chem. Res.* **2009**, *42*, 631–639.

(4) Handke, P.; Lynch, S. A.; Gill, R. T. Application and Engineering of Fatty Acid Biosynthesis in Escherichia Coli for Advanced Fuels and Chemicals. *Metab. Eng.* **2011**, *13* (1), 28–37.

(5) Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E. Evolutionary Implications of Bacterial Polyketide Synthases. *Mol. Biol. Evol.* **2005**, *22* (10), 2027–2039.

(6) White, S. W.; Zheng, J.; Zhang, Y.-M.; Rock, C. O. The Structural Biology of Type II Fatty Acid Biosynthesis. *Annu. Rev. Biochem.* **2005**, *74* (1), 791–831.

(7) Beld, J.; John Lee, D.; D. Burkart, M. Fatty Acid Biosynthesis Revisited: Structure Elucidation and Metabolic Engineering. *Mol. Biosyst.* **2015**, *11* (1), 38–59.

(8) Beld, J.; Sonnenschein, E. C.; Vickery, C. R.; Noel, J. P.; Burkart, M. D. The Phosphopantetheinyl Transferases: Catalysis of a Post-Translational Modification Crucial for Life. *Nat. Prod. Rep.* **2014**, *31* (1), 61–108.

(9) Thomas, J.; Cronan, J. E. The Enigmatic Acyl Carrier Protein Phosphodiesterase of *Escherichia Coli*: Genetic and Enzymology Characterization. *J. Biol. Chem.* **2005**, *280* (41), 34675–34683.

(10) Kosa, N. M.; Pham, K. M.; Burkart, M. D. Chemoenzymatic Exchange of Phosphopantetheine on Protein and Peptide. *Chem Sci* **2014**, *5* (3), 1179–1186.

(11) Cox, R. J.; Crosby, J.; Daltrop, O.; Glod, F.; Jarzabek, M. E.; Nicholson, T. P.; Reed, M.; Simpson, T. J.; Smith, L. H.; Soulas, F.; Szafranska, A. E.; Westcott, J. Streptomyces Coelicolor Phosphopantetheinyl Transferase: A Promiscuous Activator of Polyketide and Fatty Acid Synthase Acyl Carrier Proteins. *J. Chem. Soc. Perkin 1* **2002**, *14*, 1644–1649.

(12) Revill, W. P.; Bibb, M. J.; Hopwood, D. A. Purification of a Malonyltransferase from Streptomyces Coelicolor A3(2) and Analysis of Its Genetic Determinant. *J. Bacteriol.* **1995**, *177* (14), 3946–3952.

(13) Price, A. C.; Zhang, Y.-M.; Rock, C. O.; White, S. W. Structure of β-Ketoacyl-[Acyl Carrier Protein] Reductase from *Escherichia Coli* : Negative Cooperativity and Its Structural Basis. *Biochemistry* **2001**, *40* (43), 12772–12781.

(14) Spencer, A. K.; Greenspan, A. D.; Cronan, J. E. Thioesterases I and II of Escherichia Coli. Hydrolysis of Native Acyl-Acyl Carrier Protein Thioesters. *J. Biol. Chem.* **1978**, *253* (17), 5922–5926.

(15) Kass, L. R.; Bloch, K. On the Enzymatic Synthesis of Unsaturated Fatty Acids in Escherichia Coli. *Proc. Natl. Acad. Sci.* **1967**, *58* (3), 1168–1173.

(16) Garwin, J. L.; Klages, A. L.; Cronan, J. E. Beta-Ketoacyl-Acyl Carrier Protein Synthase II of Escherichia Coli. Evidence for Function in the Thermal Regulation of Fatty Acid Synthesis. *J. Biol. Chem.* **1980**, *255* (8), 3263–3265.

(17) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. An Antibiotic Factory Caught in Action. *Nat. Struct. Mol. Biol.* **2004**, *11* (9), 888–893.

(18) Korman, T. P.; Hill, J. A.; Vu, T. N.; Tsai, S.-C. Structural Analysis of Actinorhodin Polyketide Ketoreductase: Cofactor Binding and Substrate Specificity. *Biochemistry* **2004**, *43* (46), 14529–14538.

(19) Javidpour, P.; Bruegger, J.; Srithahan, S.; Korman, T. P.; Crump, M. P.; Crosby, J.; Burkart, M. D.; Tsai, S-C. The Determinants of Activity and Specificity in Actinorhodin Type II Polyketide Ketoreductase. *Chem. Biol.* **2013**, *20*, 1225-1234.

(20) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. Engineered Biosynthesis of Novel Polyketides: ActVII and ActIV Genes Encode Aromatase and Cyclase Enzymes, Respectively. *J. Am. Chem. Soc.* **1994**, *116* (24), 10855–10859.

(21) Taguchi, T.; Awakawa, T.; Nishihara, Y.; Kawamura, M.; Ohnishi, Y.; Ichinose, K. Bifunctionality of ActIV as a Cyclase-Thioesterase Revealed by in Vitro Reconstitution of Actinorhodin Biosynthesis in Streptomyces Coelicolor A3(2). *ChemBioChem* **2017**, *18* (3), 316–323.

(22) Du, D.; Katsuyama, Y.; Shin-ya, K.; Ohnishi, Y. Reconstitution of a Type II Polyketide Synthase That Catalyzes Polyene Formation. *Angew. Chem. Int. Ed.* **2018**, *57* (7), 1954–1957.

(23) Pohle, S.; Appelt, C.; Roux, M.; Fiedler, H.-P.; Süssmuth, R. D. Biosynthetic Gene Cluster of the Non-Ribosomally Synthesized Cyclodepsipeptide Skyllamycin: Deciphering Unprecedented Ways of Unusual Hydroxylation Reactions. *J. Am. Chem. Soc.* **2011**, *133* (16), 6194–6205.

(24) Bilyk, O.; Brötz, E.; Tokovenko, B.; Bechthold, A.; Paululat, T.; Luzhetskyy, A. New Simocyclinones: Surprising Evolutionary and Biosynthetic Insights. *ACS Chem. Biol.* **2016**, *11* (1), 241–250.

(25) Zornetzer, G. A.; Fox, B. G.; Markley, J. L. Solution Structures of Spinach Acyl Carrier Protein with Decanoate and Stearate. *Biochemistry* **2006**, *45* (16), 5217–5227.

(26) Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. Reversible Chemoenzymatic Labeling of Native and Fusion Carrier Protein Motifs. *Nat. Methods* **2012**, *9* (10), 981–984.

(27) 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. *J. Mol. Biol.* **2007**, *365* (1), 135–145.

(28) Evans, S. E.; Williams, C.; Arthur, C. J.; Płoskoń, E.; Wattana-amorn, P.; Cox, R. J.; Crosby, J.; Willis, C. L.; Simpson, T. J.; Crump, M. P. Probing the Interactions of Early Polyketide Intermediates with the Actinorhodin ACP from S. Coelicolor A3(2). *J. Mol. Biol.* **2009**, *389* (3), 511–528.

(29) 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. *J. Am. Chem. Soc.* **2014**, *136* (48), 16792–16799.

(30) Haushalter, R. W.; Filipp, F. V.; Ko, K.; Yu, R.; Opella, S. J.; Burkart, M. D. Binding and p K_a Modulation of a Polycyclic Substrate Analogue in a Type II Polyketide Acyl Carrier Protein. *ACS Chem. Biol.* **2011**, *6* (5), 413–418.

(31) Jaremko, M. J.; Lee, D. J.; Opella, S. J.; Burkart, M. D. Structure and Substrate Sequestration in the Pyoluteorin Type II Peptidyl Carrier Protein PltL. *J. Am. Chem. Soc.* **2015**, *137* (36), 11546–11549.

(32) Ploskoń, E.; Arthur, C. J.; Evans, S. E.; Williams, C.; Crosby, J.; Simpson, T. J.; Crump, M. P. A Mammalian Type I Fatty Acid Synthase Acyl Carrier Protein Domain Does Not Sequester Acyl Chains. *J. Biol. Chem.* **2008**, *283* (1), 518–528.

(33) Tran, L.; Broadhurst, R. W.; Tosin, M.; Cavalli, A.; Weissman, K. J. Insights into Protein-Protein and Enzyme-Substrate Interactions in Modular Polyketide Synthases. *Chem. Biol.* **2010**, *17* (7), 705–716.

(34) Finzel, K.; Lee, D. J.; Burkart, M. D. Using Modern Tools To Probe the Structure–Function Relationship of Fatty Acid Synthases. *ChemBioChem* **2015**, *16* (4), 528–547.

(35) Bodenhausen, G.; Ruben, D. J. Natural Abundance Nitrogen-15 NMR by Enhanced Heteronuclear Spectroscopy. *Chem. Phys. Lett.* **1980**, *69* (1), 185–189.

(36) Bax, A. Multidimensional Nuclear Magnetic Resonance Methods for Protein Studies. *Curr. Opin. Struct. Biol.* **1994**, *4* (5), 738–744.

(37) Macomber, R. S. An Introduction to NMR Titration for Studying Rapid Reversible Complexation. *J. Chem. Educ.* **1992**, *69* (5), 375.

(38) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

(39) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of Folded Proteins. *Nature* **1977**, *267* (16), 585–590.

(40) Wang, Y.; Ma, S. Recent Advances in Inhibitors of Bacterial Fatty Acid Synthesis Type II (FASII) System Enzymes as Potential Antibacterial Agents. *ChemMedChem* **2013**, *8*, 1589–1608.

(41) Heath, R. J.; White, S. W.; Rock, C. O. Inhibitors of Fatty Acid Synthesis as Antimicrobial Chemotherapeutics. *Appl. Microbiol. Biotechnol.* **2002**, *58* (6), 695–703.

(42) Moche, M.; Schneider, G.; Edwards, P.; Dehesh, K.; Lindqvist, Y. Structure of the Complex between the Antibiotic Cerulenin and Its Target, b-Ketoacyl-Acyl Carrier Protein Synthase. *J. Biol. Chem.* **1999**, *27* (12), 1574–1574.

(43) 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. *J. Biol. Chem.* **1999**, *274* (43), 30811–30817.

(44) Heath, R. J.; Rubin, J. R.; Holland, D. R.; Zhang, E.; Snow, M. E.; Rock, C. O. Mechanism of Triclosan Inhibition of Bacterial Fatty Acid Synthesis. *J. Biol. Chem.* **1999**, *274* (16), 11110–11114.

(45) Levy, C. W.; Roujeinikova, A.; Sedelnikova, S.; Baker, P. J.; Stuitje, A. R.; Slabas, A. R.; Rice, D. W.; Rafferty, J. B. Molecular Basis of Triclosan Activity. *Nature* **1999**, *398* (6726), 383–384.

(46) Tang, Y.; Tsai, S.-C.; Khosla, C. Polyketide Chain Length Control by Chain Length Factor. *J. Am. Chem. Soc.* **2003**, *125* (42), 12708–12709.

(47) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org Biomol Chem* **2006**, *4* (1), 44–46.

(48) Quadri, L. E. N.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry*, **1998**, *37*, 1585–1595.

(49) 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 Chem. Biol.* **2006**, *1* (11), 687–691.

(50) Ishikawa, F.; Haushalter, R. W.; Lee, D. J.; Finzel, K.; Burkart, M. D. Sulfonyl 3-Alkynyl Pantetheinamides as Mechanism-Based Cross-Linkers of Acyl Carrier Protein Dehydratase. *J. Am. Chem. Soc.* **2013**, *135* (24), 8846–8849.

(51) Ishikawa, F.; Haushalter, R. W.; Burkart, M. D. Dehydratase-Specific Probes for Fatty Acid and Polyketide Synthases. *J. Am. Chem. Soc.* **2012**, *134* (2), 769–772.

(52) Beld, J.; Cang, H.; Burkart, M. D. Visualizing the Chain-Flipping Mechanism in Fatty Acid Biosynthesis. *Angew. Chem. Int. Ed Engl.* **2014**, *53* (52), 14456–14461.

(53) Miyanaga, A.; Iwasawa, S.; Shinohara, Y.; Kudo, F.; Eguchi, T. Structure-Based Analysis of the Molecular Interactions between Acyltransferase and Acyl Carrier Protein in Vicenistatin Biosynthesis. *Proc. Natl. Acad. Sci.* **2016**, *113* (7), 1802–1807.

(54) Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Markwick, P. R. L.; Bruegger, J.; Caldara-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* (7483), 427–431.

(55) Dong, X.; Bailey, C. D.; Williams, C.; Crosby, J.; Simpson, T. J.; Willis, C. L.; Crump, M. P. Recognition of Extended Linear and Cyclised Polyketide Mimics by a Type II Acyl Carrier Protein. *Chem. Sci.* **2016**, *7* (3), 1779–1785.

(56) Ellis, B. D.; Milligan, J. C.; White, A. R.; Duong, V.; Altman, P. X.; Mohammed, L. Y.; Crump, M. P.; Crosby, J.; Luo, R.; Vanderwal, C. D.; Tsai, S.-C. An Oxetane-Based

Polyketide Surrogate to Probe Substrate Binding in a Polyketide Synthase. *J. Am. Chem. Soc.* **2018**, *140* (15), 4961–4964.

(57) Barajas, J. F.; Shakya, G.; Moreno, G.; Rivera, H.; Jackson, D. R.; Topper, C. L.; Vagstad, A. L.; La Clair, J. J.; Townsend, C. A.; Burkart, M. D. Polyketide Mimetics Yield Structural and Mechanistic Insights into Product Template Domain Function in Nonreducing Polyketide Synthases. *Proc. Natl. Acad. Sci.* **2017**, 201609001.

(58) Meanwell, N. A. Synopsis of Some Recent Tactical Application of Bioisosteres in Drug Design. *J. Med. Chem.* **2011**, *54* (8), 2529–2591.

(59) 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* (3), 253–264.

(60) Roujeinikova, A.; Baldock, C.; Simon, W. J.; Gilroy, J.; Baker, P. J.; Stuitje, A. R.; Rice, D. W.; Slabas, A. R.; Rafferty, J. B. X-Ray Crystallographic Studies on Butyryl-ACP Reveal Flexibility of the Structure around a Putative Acyl Chain Binding Site. *Structure* **2002**, *10* (6), 825–835.

(61) Arabolaza, A.; Shillito, M. E.; Lin, T.-W.; Diacovich, L.; Melgar, M.; Pham, H.; Amick, D.; Gramajo, H.; Tsai, S.-C. Crystal Structures and Mutational Analyses of Acyl-CoA Carboxylase β Subunit of Streptomyces Coelicolor. *Biochemistry* **2010**, *49* (34), 7367–7376.

(62) Qiu, X.; Janson, C. A.; Konstantinidis, A. K.; Nwagwu, S.; Silverman, C.; Smith, W. W.; Khandekar, S.; Lonsdale, J.; Abdel-Meguid, S. S. Crystal Structure of β-Ketoacyl-Acyl Carrier Protein Synthase III A KEY CONDENSING ENZYME IN BACTERIAL FATTY ACID BIOSYNTHESIS. *J. Biol. Chem.* **1999**, *274* (51), 36465–36471.

(63) Serre, L.; Verbree, E. C.; Dauter, Z.; Stuitje, A. R.; Derewenda, Z. S. The Escherichia Coli Malonyl-CoA:Acyl Carrier Protein Transacylase at 1.5-Å Resolution. CRYSTAL STRUCTURE OF A FATTY ACID SYNTHASE COMPONENT. *J. Biol. Chem.* **1995**, *270* (22), 12961–12964.

(64) 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. *J. Biol. Chem.* **2004**, *279* (50), 52593–52602.

(65) Pan, H.; Tsai, S.; Meadows, E. S.; Miercke, L. J. W.; Keatinge-Clay, A. T.; O'Connell, J.; Khosla, C.; Stroud, R. M. Crystal Structure of the Priming β-Ketosynthase from the R1128 Polyketide Biosynthetic Pathway. *Structure* **2002**, *10* (11), 1559–1568.

(66) Hadfield, A. T.; Limpkin, C.; Teartasin, W.; Simpson, T. J.; Crosby, J.; Crump, M. P. The Crystal Structure of the ActIII Actinorhodin Polyketide Reductase: Proposed Mechanism for ACP and Polyketide Binding. *Structure* **2004**, *12* (10), 1865–1875.

(67) Mori, T.; Awakawa, T.; Shimomura, K.; Saito, Y.; Yang, D.; Morita, H.; Abe, I. Structural Insight into the Enzymatic Formation of Bacterial Stilbene. *Cell Chem. Biol.* **2016**, *23*, 1468–1479.

(68) Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II Polyketide Synthases: Gaining a Deeper Insight into Enzymatic Teamwork. *Nat. Prod. Rep.* **2007**, *24* (1), 162–190.

(69) Choi, K.-H.; Kremer, L.; Besra, G. S.; Rock, C. O. Identification and Substrate Specificity of β-Ketoacyl (Acyl Carrier Protein) Synthase III (MtFabH) from Mycobacterium Tuberculosis. *J. Biol. Chem.* **2000**, *275* (36), 28201–28207.

(70) Gajiwala Ketan S.; Margosiak Stephen; Lu Jia; Cortez Joseph; Su Ying; Nie Zhe; Appelt Krzysztof. Crystal Structures of Bacterial FabH Suggest a Molecular Basis for the Substrate Specificity of the Enzyme. *FEBS Lett.* **2009**, *583* (17), 2939–2946.

(71) Qiu, X.; Janson, C. A.; Smith, W. W.; Head, M.; Lonsdale, J.; Konstantinidis, A. K. Refined Structures of β-Ketoacyl-Acyl Carrier Protein Synthase III. *J. Mol. Biol.* **2001**, *307* (1), 341–356.

(72) Ellis, B. D.; Milligan, J. C.; White, A. R.; Duong, V.; Altman, P. X.; Mohammed, L. Y.; Crump, M. P.; Crosby, J.; Luo, R.; Vanderwal, C. D.; Tsai, S.-C. An Oxetane-Based Polyketide Surrogate to Probe Substrate Binding in a Polyketide Synthase. *J. Am. Chem. Soc.* **2018**, *140* (15), 4961–4964.

(73) Bisang, C.; Long, P. F.; Corte´s, J.; Westcott, J.; Crosby, J.; Matharu, A.-L.; Cox, R. J.; Simpson, T. J.; Staunton, J.; Leadlay, P. F. A Chain Initiation Factor Common to Both Modular and Aromatic Polyketide Synthases. *Nature* **1999**, *401* (6752), 502–505.

(74) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. Rational Design of Aromatic Polyketide Natural Products by Recombinant Assembly of Enzymatic Subunits. *Nature* **1995**, *375* (6532), 549–554.

(75) Korman, T. P.; Tan, Y.; Wong, J.; Luo, R.; Tsai, S.-C. Inhibition Kinetics and Emodin Cocrystal Structure of a Type II Polyketide Ketoreductase. *Biochemistry* **2008**, *47* (7), 1837–1847.

(76) Carreras, C. W.; Khosla, C. Purification and in Vitro Reconstitution of the Essential Protein Components of an Aromatic Polyketide Synthase. *Biochemistry* **1998**, *37* (8), 2084–2088.

(77) Fu, H.; Hopwood, D. A.; Khosla, C. Engineered Biosynthesis of Novel Polyketides: Evidence for Temporal, but Not Regiospecific, Control of Cyclization of an Aromatic Polyketide Precursor. *Chem. Biol.* **1994**, *1* (4), 205–210.

(78) McDaniel, R.; Ebert-Khosla, S.; Fu, H.; Hopwood, D. A.; Khosla, C. Engineered Biosynthesis of Novel Polyketides: Influence of a Downstream Enzyme on the Catalytic Specificity of a Minimal Aromatic Polyketide Synthase. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91* (24), 11542–11546.

(79) Javidpour, P.; Bruegger, J.; Srithahan, S.; Korman, T. P.; Crump, M. P.; Crosby, J.; Burkart, M. D.; Tsai, S.-C. The Determinants of Activity and Specificity in Actinorhodin Type II Polyketide Ketoreductase. *Chem. Biol.* **2013**, *20* (10), 1225–1234.

(80) Helmkamp, G. M.; Bloch, K. β-Hydroxydecanoyl Thioester Dehydrase STUDIES ON MOLECULAR STRUCTURE AND ACTIVE SITE. *J. Biol. Chem.* **1969**, *244* (21), 6014–6022.

(81) Finzel, K.; Nguyen, C.; Jackson, D. R.; Gupta, A.; Tsai, S.-C.; Burkart, M. D. Probing the Substrate Specificity and Protein-Protein Interactions of the E. Coli Fatty Acid Dehydratase, FabA. *Chem. Biol.* **2015**, *22* (11), 1453–1460.

(82) Milligan, J. C.; Lee, D. J.; Jackson, D. R.; Schaub, A. J.; Beld, J.; Barajas, J. F.; Hale, J. J.; Luo, R.; Burkart, M. D.; Tsai, S.-C. Molecular basis for interactions between an acyl carrier protein and a ketosynthase. *Nat. Chem. Biol.* **2019**, *15* (7), 669-671.

(83) Zhang, L.; Xiao, J.; Xu, J.; Fu, T.; Cao, Z.; Zhu, L.; Chen, H.-Z.; Shen, X.; Jiang, H.; Zhang, L. Crystal Structure of FabZ-ACP Complex Reveals a Dynamic Seesaw-like Catalytic Mechanism of Dehydratase in Fatty Acid Biosynthesis. *Cell Res.* **2016**, *26* (12), 1330–1344.

(84) Rafi, S.; Novichenok, P.; Kolappan, S.; Zhang, X.; Stratton, C. F.; Rawat, R.; Kisker, C.; Simmerling, C.; Tonge, P. J. Structure of Acyl Carrier Protein Bound to FabI, the FASII Enoyl Reductase from *Escherichia Coli*. *J. Biol. Chem.* **2006**, *281* (51), 39285–39293.

(85) 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. *J. Am. Chem. Soc.* **2016**, *138* (12), 3962–3965.

(86) Fujita, Y.; Matsuoka, H.; Hirooka, K. Regulation of Fatty Acid Metabolism in Bacteria. *Mol. Microbiol.* **2007**, *66* (4), 829–839.

(87) Yan, H.; Wang, Z.; Wang, F.; Tan, T.; Liu, L. Biosynthesis of Chain-Specific Alkanes by Metabolic Engineering in Escherichia Coli. *Eng. Life Sci.* **2016**, *16* (1), 53–59.

(88) Feng, Y.; Zhang, Y.; Wang, Y.; Liu, J.; Liu, Y.; Cao, X.; Xue, S. Tuning of Acyl-ACP Thioesterase Activity Directed for Tailored Fatty Acid Synthesis. *Appl. Microbiol. Biotechnol.* **2018**, *102* (7), 3173–3182.

(89) Blatti, J. L.; Beld, J.; Behnke, C. A.; Mendez, M.; Mayfield, S. P.; Burkart, M. D. Manipulating Fatty Acid Biosynthesis in Microalgae for Biofuel through Protein-Protein Interactions. *PLoS ONE* **2012**, *7* (9), e42949.

Chapter 2: Enzymology of Standalone Elongating Ketosynthases

Abstract

The β-ketoacyl-acyl carrier protein synthase, or ketosynthase (KS), catalyses carbon-carbon bond formation in fatty acid and polyketide biosynthesis via a decarboxylative Claisen-like condensation. In Prokaryotes, standalone elongating KSs interact with the acyl carrier protein (ACP) which shuttles substrates to each partner enzymes in the elongation cycle for catalysis. Despite ongoing research for more than 50 years since KS was first identified in E. coli, the complex mechanism of KSs continue to be unravelled, including recent understanding of gating motifs, KS-ACP interactions, substrate recognition and delivery, and roles in unsaturated fatty acid biosynthesis. In this review, we summarize the latest studies, primarily conducted through structural biology and molecular probe design, that shed light on the emerging enzymology of standalone elongating KSs.

Introduction

The fatty acid biosynthesis (FAB) and the polyketide biosynthesis (PKB) are evolutionarily related biosynthetic pathways that synthesize primary and secondary metabolites, respectively. Utilizing the same fundamental enzymatic transformations, both FAB and PKB are of great interest for the production of valuable compounds such as biofuels, pharmaceutical leads, and agrochemicals. Fatty acid synthases (FASs) and polyketide synthases (PKSs) can be classified into type I, which multiple domains are linked together to form megasynthases, and type II, which possess discrete, primarily monofunctional enzymes. Despite variations of components across these assembly line pathways, the

ketosynthase (KS) domain is always present, serving as the key carbon-carbon bond catalysing enzyme in the FAB and PKB elongation cycle. Another crucial protein is the acyl carrier protein (ACP), a small 4-helical bundle that shuttles substrates and intermediates throughout the pathways, interacting with each catalytic domain. A functional ACP (holo-ACP) requires post-translational installation of a 4'-phosphopantetheine moiety (PPant) onto a serine residue in which PPant provides the thiol group for substrate tethering. In this review, we aim to delineate the enzymology of elongating KSs from type II systems by focusing on three topics: the KS gating mechanism, substrate control through the KS pocket, and mechanisms involving KS-ACP interactions.

The elongation mechanism of KS. The elongating KS active site consists of the cysteine-histidine-histidine catalytic triad and an oxyanion hole formed by the backbone amides of the catalytic cysteine and a conserved phenylalanine. The ping-pong mechanism of KS involves two half-reactions, the transacylation step and the condensation step, respectively. In the transacylation step, the acyl substrate gets transferred onto the active site cysteine with the oxyanion hole stabilizing the tetrahedral intermediate. Subsequently, the condensation step has the two histidine residues catalysing a decarboxylative enolate attack from a malonyl group to the thioester tethered substrate. The oxyanion hole continues to play the role stabilizing the oxyanion formed during the reaction. The overall catalysis condenses 2-carbon unit onto the acyl substrate to give a β-keto acyl product. Each half-reaction involves association and dissociation of ACP. Known to sequester the substrate, acyl-ACP delivers acyl substrate to the KS pocket under the chain-flipping mechanism (Figure

2.1A), and ACP is released as *holo*-ACP after the transacylation step. A malonyl-ACP then enters the cycle for the condensation step and leaves with the β-keto acyl product tethered (Figure 2.1B).

The detailed mechanism of the transacylation step is well understood. The active site cysteine locates at a sharp turn of the N-terminus of a long α-helix, also known as a nucleophilic elbow, which enhances the nucleophilicity and basicity of the thiol group by the α-helix dipole moment.[1, 2] The oxyanion hole also facilitates the reaction by accommodating the negative charge build-up on the carbonyl oxygen. The detailed mechanism of the condensation half-reaction, on the other hand, is still under debate. For type II FAS (FASII), a widely accepted version is a sequential mechanism that starts with the decarboxylation of the malonyl group to generate enolate and then the enolate attacks and undergoes condensation with the KS-bound acyl substrate. Experimental data from animal FAS shows that the decarboxylation by-product is bicarbonate while releasing carbon dioxide as the by-product is still possible in FASII.[3] At least two other alternative mechanisms have been proposed and reader is directed to the review written by Heil *et al.* for further detail.[4]

Standalone KSs from different pathways. As the canonical and most well-studied FASII system, *E. coli* FAS will lead the following introduction of the FAS KS. Two elongating KSs, the β-ketoacyl-ACP synthase II (FabF) and the β-ketoacyl-ACP synthase I (FabB), exist in the system. FabF serves as the primary elongating KS of FAB and is also the evolutional progenitor of the PKS KSs.[5] FabB specializes in the first elongation of mono-unsaturated fatty acid (UFA) while still capable of

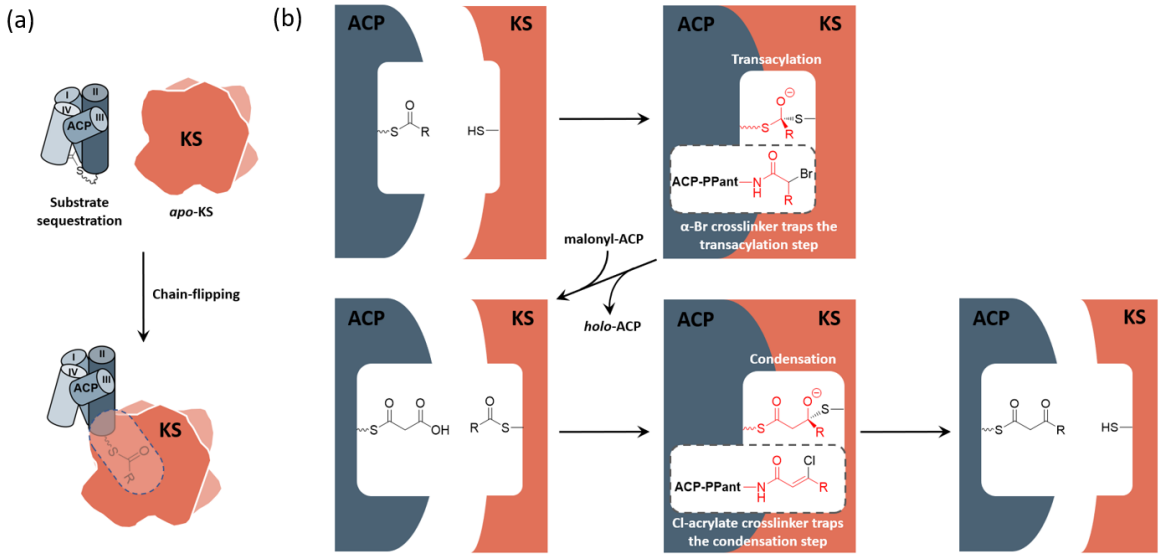


Figure 2.1: KS catalysis and mechanism-based crosslinkers: (A) Transition of the substrate from the ACP pocket to the KS pocket is described as a chain-flipping event. (B) The ping-pong mechanism of KS catalysis begins with the transacylation half-reaction to generate acyl-KS followed by the condensation half-reaction that synthesizes β-ketoacyl-ACP via decarboxylative condensation. The mechanism-based crosslinkers, α-bromo and Cl-acrylate, can trap KS-ACP interaction in the transacylation state and condensation state, respectively.

maintaining FAB by itself.[6-8] A lack of FabB in *E. coli* leads to UFA auxotroph. Both FabF and FabB are functional as homodimer, forming two substrate pockets along the dimer interface. A full cycle of saturated fatty acid (SFA) elongation starts with the KS condensing 2-carbon unit from a malonyl group to the acyl-chain followed by a ketoreductase (KR) reducing the β-keto group to a hydroxyl group, a dehydratase (DH) eliminating the hydroxyl group to generate enoyl intermediate, and finally an enoyl-reductase (ER) to reduce the enoyl double bond (Figure 2.2). In E. coli and related proteobacteria, the synthesis of UFA elongation occurs de novo through the same fashion, except for the particular cycle that retains the unsaturation the (2*E*)-enoyl intermediate is isomerized to (3*Z*)-enoyl by the DH that possesses isomerase activity, such as FabA in *E. coli*, and subsequently enters the next cycle of elongation.

The conventional type II PKB produces highly functional polycyclic aromatic compounds derived from the chemically flexible polyketone intermediates which are generated from skipping all the domains but KS in the elongation cycle (Figure 2.2). Given the high reactivity of the polyketone intermediates, the substrate remains protected inside the pocket during elongation until it reaches the maximum chain length.[9-12] Another class of PKS, the polyene PKS, that has only been studied less than a decade has every domain from the FAB elongation cycle except the ER domain, producing polyene[13-16] or aryl polyene products.[17-23] In terms of substrate shuttling, polyene PKS system is more similar to the FAS system where ACP is required to deliver substrates to multiple enzymes during elongation. Although various nomenclatures have been used to describe the two PKS systems mentioned above, for example by their product types (aromatic vs. polyene) or

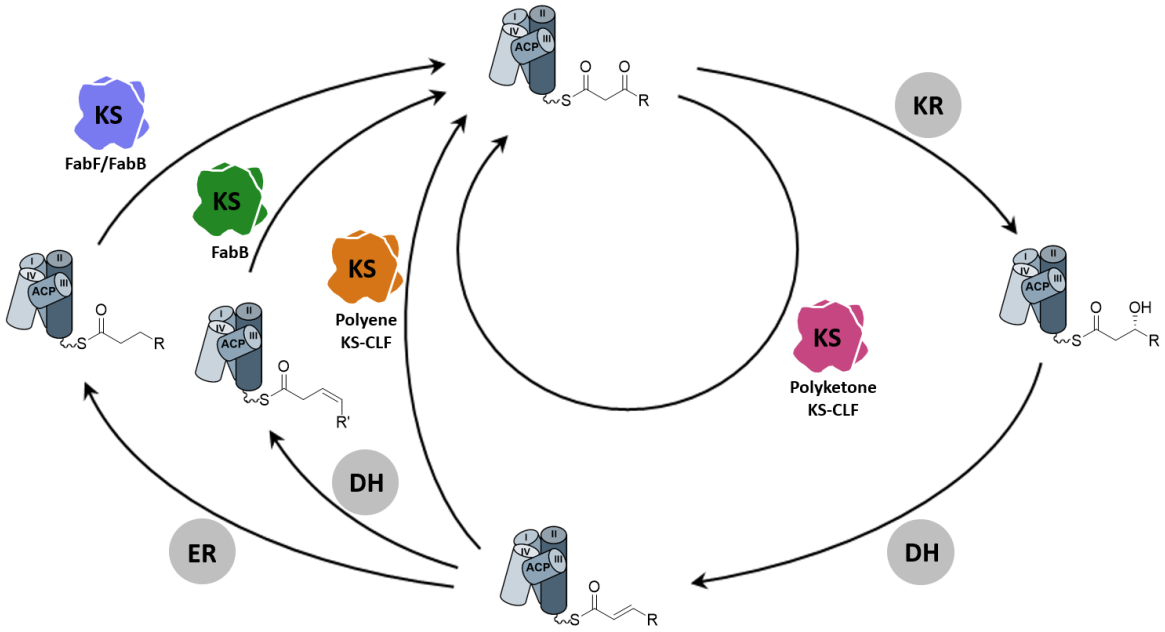


Figure 2.2: The elongation cycle of type II FAS and PKS: A full cycle of SFA elongation goes through KS-mediated elongation, ketoreduction, dehydration, and enoyl-reduction. The *cis*-double bond of mono-unsaturated fatty acid comes from the isomerization of (2E)-enoyl-ACP followed by the elongation of FabB (green KS) in a particular cycle. Consecutively skipping the reduction of enoyl-ACP yields polyene intermediates that are elongated by polyene KS-CLF (orange KS), and the elongation catalyzed individually by polyketone KS-CLF (magenta KS) yields chemically flexible intermediates that lead to polycyclic aromatic polyketides. (KR = ketoreductase, DH = dehydratase, ER = enoyl-reductase)

reduction level (non-reducing vs. highly reducing), in this review we will refer to them by their elongation intermediates, polyketone and polyene, respectively, for simplicity.

The type II PKS (PKSII) KSs are heterodimers formed by a KS_{α} and a KS_{β} subunits. The KS_{β} subunit has no condensation activity and lacks the canonical KS active site residues, but it plays a crirtical role in controlling the chain length of the resulting product and is thus often referred to as the chain-length-factor (CLF) subunit. There is one exception that the ApeR, one of the two KSs of aryl polyene biosynthesis, is instead a homodimeric KS.[21, 23] Historically, heterologous expression of polyketone KS-CLFs in *E. coli* has required co-expression of the pathway-specific malonyl-CoA:ACP transferase, which altogether is known as the "minimal PKS", to yield soluble proteins.[24] However, recent efforts have resulted in several successful cases of individually expressing polyketone KS-CLF in *E. coli*, lowering the barrier of studying the enzymology of FASII KS.[25-29] There are a number of organizational questions that remain to be answered. In organisms that possess both type II FAS and PKS systems, the mechanism that prevents crosstalk, or non-cognate protein catalysis, remains an important unanswered. In addition, how FabF and FabB, structurally highly similar, differ in substrate specificity has not been solved. Growing evidence has shown that the protein-protein interactions (PPIs) of KS and ACP, and changes imparted by substrate identity, play an important role in these mechanisms, which will be covered in this review. Type III PKS pathways such as chalcone synthase, are considered standalone KSs, but they are not within the scope of this review, which focuses on ACP-dependent systems.[30, 31]

Biochemical tools to investigate KS enzymology. During the elongation of an acyl chain, many mechanistic steps must be achieved through the close communication between KS and ACP. However, due to the conformational dynamics of the ACP and the rapid binding kinetics required to facilitate catalysis, fully understanding the protein-protein interactions (PPIs) and KS substrate recognition at the molecular level remains as an ongoing challenge. To address these issues, our lab and others have developed synthetic pantetheinamide probes to interrogate the ACP-KS PPIs and substrate-protein interactions structurally and biochemically.[32-35] It is noteworthy that amongst these chemical probes, mechanism-based crosslinking probes, or simply crosslinkers, have shown success trapping KS-ACP interactions in their catalytic states.[36-38] There are two major types of crosslinkers, α-bromo-pantetheinamide and chloro-acrylate-pantetheinamide, respectively, for KS-ACP crosslinking. According to the carbon count and the coordination of the active site, α-bromo crosslinkers trap KS-ACP complex in the transacylation state, while Cl-acrylate crosslinkers trap KS-ACP in the condensation state (Figure 2.1B). Crosslinking complexes have provided a deeper understanding of the enzymatic reaction mechanism of KS and the PPIs between KS and ACP.

In the study of PKS, because of the inherent chemical instability of the polyketone intermediates, different types of polyketides mimetic surrogates have been synthesized to gain more insight of the intermediate protection mechanism in the binding pocket.[10, 39-43] Additionally, the process of substrates sequestration in ACP and chain flipping were elucidated by vibrational spectroscopic and the

solvatochromic probes.[44, 45] Readers can be directed to the comprehensive reviews by Chen *et al.* and Sulpizio *et al.* for more information.[11, 46]

Gating mechanisms in elongating KS

A recently discovered KS gate involving two gating loops has been proposed to regulate substrate processing and catalysis.[38,47] Together with two other investigated KS gates, KS deploys a well-controlled system to orchestrate the complex 2-step catalysis. We will discuss the mechanism, influence, and universality of each gate. If not specifically mentioned, residue numbering in this section is based on *Ec*FabF.

The front gate: double drawbridge-like gating loops. The front gate consists of two stacking loops that move co-ordinately to control the entry of the substrates.[48] In *Ec*FabF, this contains loop 1 comprising G399-G402 (GFGG), and loop 2 comprising D265-N275 (DAYHMTSPPEN).[38] Transition from the closed to the open form of the gate requires a large conformational change of the loops, with the C_β of F400 moving 8.7 Å away from its original position (Fig. 2.3A). Such movement of F400 provides access to the substrate binding pocket and the active site cysteine, C163. In the open form, the oxyanion hole formed by the backbone amides of C163 and F400 is disrupted and therefore cannot stabilize KS reaction intermediates. Thus, the open form is believed to facilitate substrate delivery, whereas the closed form is the catalytically competent form with a well-ordered oxyanion hole.

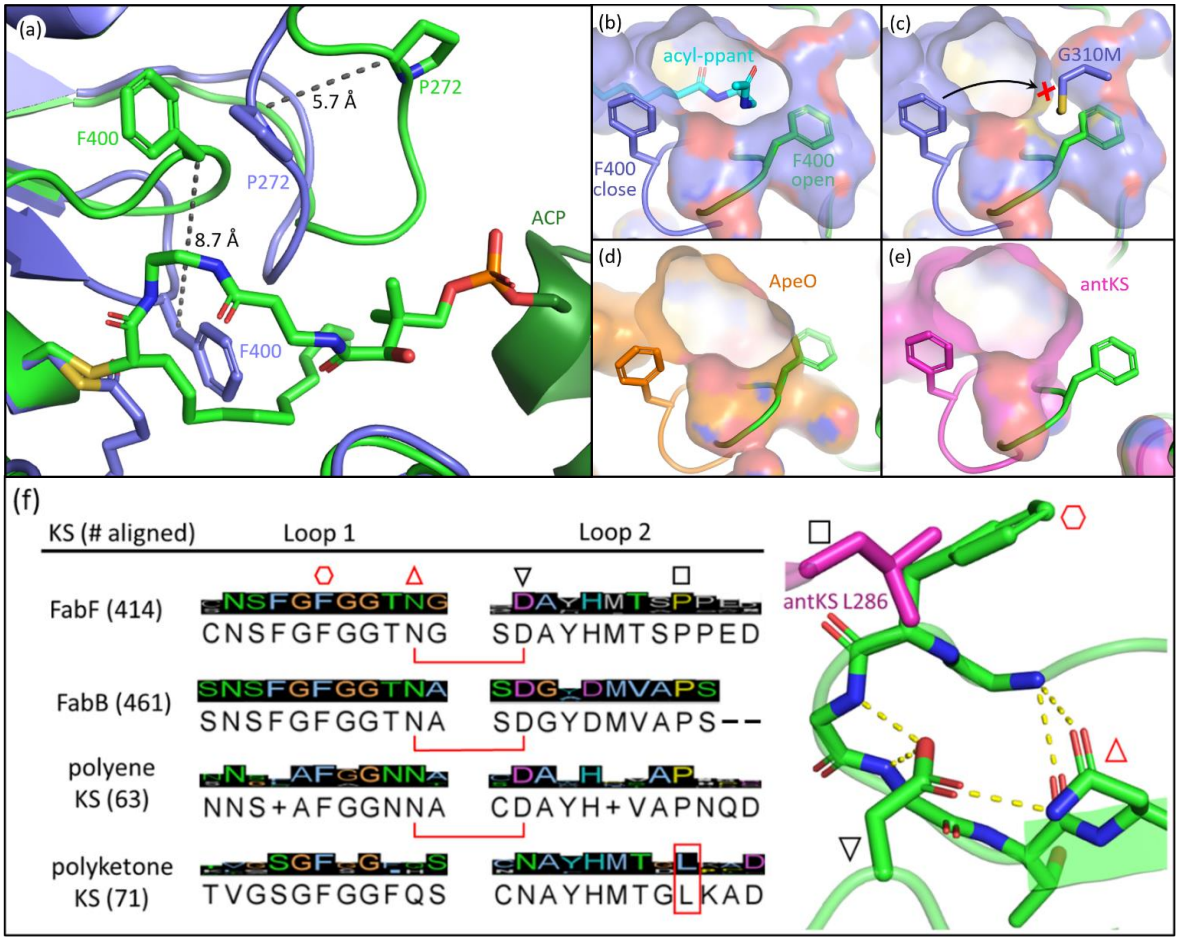


Figure 2.3: The gating loops: (A) structure overlay of C12-FabF (PDB ID: 2GFY, blue) in gate closed conformation and KS-ACP crosslink complex (PDB ID: 6OKG, green) in gate open conformation. Movements of F400 on loop 1 and P272 on loop 2 are measured. (B) surface plot of gate closed FabF (blue) overlaying with the open loop 1 demonstrates the side pocket that accommodates the open loop. (C) G310M mutant blocks the gate from opening. (D) surface plot of ApeO (PDB ID: 6QSP, orange), a KS from aryl polyene biosynthesis, overlaying with the open loop 1. (E) surface plot of a polyketone KS, antKS (PDB ID: 6SMO, magenta), overlaying with the open loop 1. The side pocket for the open gate is completely missing. (F) Over a thousand KS gating loop 1 and loop 2 sequences aligned (MUSCLE algorithm) showing the consensus sequence (left column). The coordination of open loop 1 is depicted on the right. The conserved Asn(N)-Asp(D) pair that is essential for the open gate is absent in polyketone KS. The conserved Leu (L286 in antKS) of polyketone KS clashes into the Phe(F) of open loop 1. This Leu is instead a conserved Pro in other types of KS.

The open form of the gate was captured by crystal structures of *Ec*FabF-*Ec*ACP crosslink complex using the α-bromo-pantetheinamide crosslinking probes with various chain length of fatty acid mimetics (C12, C16, and C16:1), whereas using the C8 chloro-acrylate-pantetheinamide crosslinking probe yielded a gate-closed structure. Overlay of these structures reveals a "side pocket" that accommodates the open form of loop 1 (Figure 2.3B). More than a dozen *Ec*FabF mutants that interrogate the gating mechanism were developed and tested with a crosslinking gel-shift assay and two kinetic activity assays to validate the role of the gate. One particular interesting class of mutations that compromises FabF activity is the pocket block mutants, such as G310M (Figure 2.3C), that fills the side pocket to prevent the loop 1 from opening. The side pocket could be observed in structures of FAS and polyene KS that release substrates for each cycle of elongation (Figure 2.3B and 2.3D). However, for the polyketone KS, which is believed to keep the reactive polyketone intermediate in the pocket during elongation,[11] there is no space for loop 1 in the fully open conformation (Figure 2.3E). Sequence alignment over a thousand different KSs also reveal that polyketone KS lacks a pair of crucial residues (D265-N404 of *Ec*FabF) which stabilizes the open gate and is totally conserved in other KSs aligned. Furthermore, a conserved proline in loop 2 is instead a leucine in polyketone KS, which blocks the gate from opening (Figure 2.3F). Although the universality of the gating mechanism has yet to be confirmed, based on structure analysis and sequence alignment it seems to exist in KSs that actively accept and release substrates, an

activity absent in polyketone KSs, which retain the substrate during iterative rounds of elongation.

The KS active site and gating loops arrangement during different states of the catalysis are summarized in figure 2.4. Starting from the *apo*-KS (1), the gating loops transition to the open conformation (2) for substrate delivery from acyl-ACP. The gate then transitions back to reform the oxyanion hole to set up the substrate in state (3) for the transacylation reaction. After *holo*-ACP dissociation, the acyl-KS (4) reacts with malonyl-ACP to form the decarboxylative condensation intermediate depicted in state (5). The final product, β-ketoacyl-ACP, is generated (6) and released by opening the gate (7) to bring the KS back to the starting point. In the case of aromatic PKS, rather than releasing the substrate, a transacylation reaction should tailor state (6) to position the high energy polyketone intermediate back to the KS, depicted in (PK). The reaction undergoes an iterative cycle between (4), (5), (6), and (PK) until the substrate reaches the final chain length. The gate-open states (2) and (7) may not exist in this pathway. Most of the states depicted in figure 4 are supported by KS or KS-ACP crosslinked crystal structures.[12, 36-38]

The gating mechanism of the gating loops provides an important piece of puzzle that connects essential elements in KS catalysis. The loop 1 contains the residue F400 (in *Ec*FabF numbering) that plays a role in regulating the condensation half reaction. When ACP is bound, loop 2 sits in between loop 1 and ACP-bearing residues that interact with loop 1, ACP, and the phosphopantethein moiety, indicating a potential communication role that loop 2 might play. Swapping loop 2 of *Ec*FabF

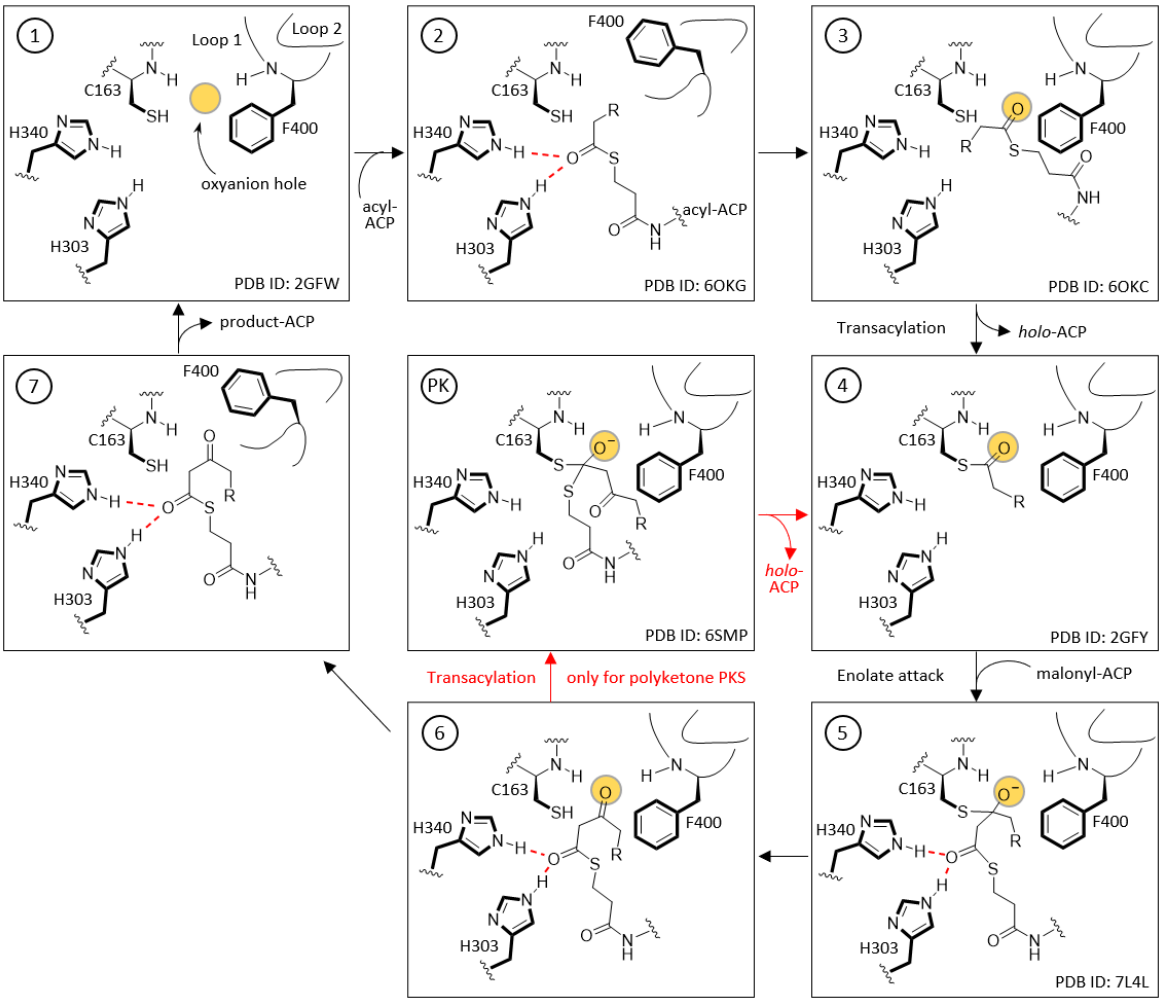


Figure 2.4: Detailed mechanism of KS elongation and the gating loops: FabF, FabB, and polyene KS goes through (1)-(7) followed by product-ACP release for one extension while polyketone KS goes through the smaller circle (4), (5), (6), and (PK), to keep the growing chain inside the KS pocket. (2) and (7) are considered the substrate delivery and released states, respectively, with the gating loops open and the catalytic oxyanion hole absent. Most depicted states are supported by at least one crystal structures with a representative of each labelled at the bottom-right corner of the box. The amino acid numbering is from EcFabF.

with the loop 2 of *Ec*FabB has detrimental effect on *Ec*FabF function,[47] suggesting that the gating loops are evolutionarily coupled in specific systems to function in a certain way. The large space on the other side of the substrate pocket, which has never before been assigned a purpose, is shown to accommodate the open form of loop 1 (Fig. 2.3B). Since this gate regulates substrate delivery, it could potentially play a role in KS substrate specificity. More functions of the gate have yet to be explored, but with the importance discovered so far, it is a factor that cannot be ignored in KS enzymology, particularly in the discussion of engineering efforts.

The gate for condensation half reaction: rotamers of F400. The conserved phenylalanine residue of gating loop 1 also regulates the condensation half-reaction through its rotamers. Acylation of the active site cysteine results in a ~60 degrees rotation of the phenyl group along the C_β-C_γ bond of the phenylalanine, setting up a malonyl substrate-binding pocket for the decarboxylative condensation half reaction (Fig. 2.5A).[49] Such a mechanism prevents KS from taking malonyl substrates before the transacylation happens, which is critical to the timing of the ping-pong mechanism. Mutating the cysteine to a glutamine, which gives an active site arrangement similar to that of acyl-KS, is reported to convert the KS to a malonyl decarboxylase,[50] while also significantly increasing KS's binding affinity towards platensimycin, a KS inhibitor that mimics the malonyl substrate.[51] A phenylalanine-to-alanine mutation almost eliminates condensation activity of KS *in vitro* (0.7-1.85 % activity of wild type) while also decreasing the transacylation activity as well as other activities that generate shunt products.[47, 51] This demonstrates that by simply removing the physical barrier formed by the phenylalanine gating residue does not improve

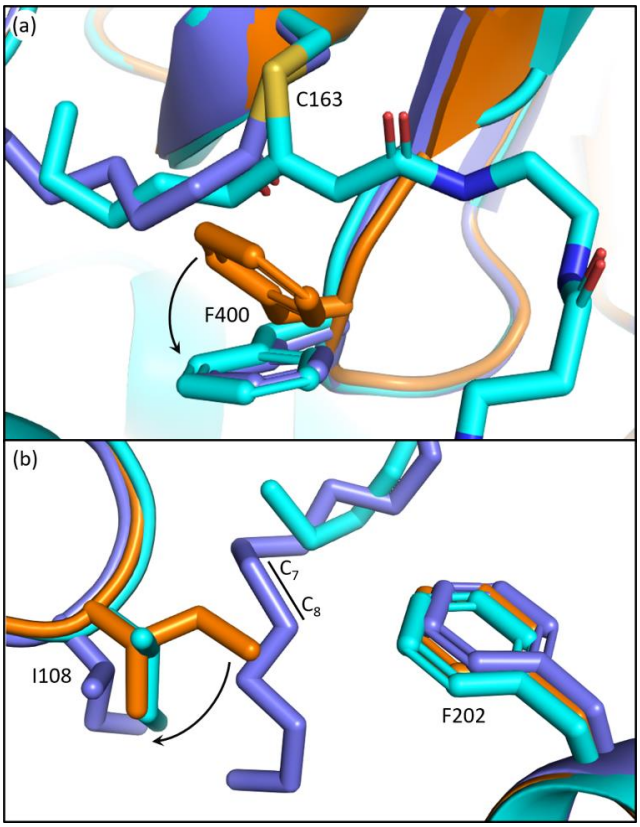


Figure 2.5: Gating residues of FabF: overlay of apoFabF (PDB ID: 2GFW, orange), C12-FabF (PDB ID: 2GFY, blue), and FabF-ACP crosslink structure (PDB ID: 7L4L, light blue). (A) Rotation of F400 triggered by substrate binding creates the malonyl binding pocket. (B) Rotation and translation of I108 when a long acyl-chain (more than 8-carbon) is bound.

the enzyme overall. These results support three functions for this residue: playing a role in the gating loops that regulate substrate delivery, controlling binding of malonyl-ACP substrates, and bearing the amide that takes part in forming the catalytic oxyanion hole, which might need the bulky side chain to fix the residue in the right position. Additionally, it has also been suggested that conformations of this phenylalanine may determine the protonation states of the active site residues in *Mycobacterium tuberculosis* KS.[52]

The phenyl ring of F400 is positioned squarely in front of the α carbon of the acyl chain in acyl-KS and may also contribute to the selectivity towards α-branched substrates. In type I PKS, which takes branched polyketones and α-branched malonyl extender units such as methylmalonyl-CoA (mmCoA), the phenylalanine is replaced either by a valine or an isoleucine, potentially creating the space to accommodate a branching alkyl group. Interestingly, *Ec*FabF F400V and F400I mutants, compared to the wild type (WT), suffer reduced condensation activities *in vitro* towards both mal-CoA and mmCoA,[47] yet another indication of the complex role of F400. Considering the position of the F400 side chain, it would be interesting to test these mutants' activities towards α-branched acyl substrates.

The back gate: I108-F202. In *Ec*FabF, I108 and F202 form a gate in the middle of the substrate pocket, approximately located at the C_7-C_8 position of a bound acyl chain. While FabB has the same loop 1 phenylalanine residue, I108 is highly conserved and unique to FabF and is replaced by a glycine residue in FabB. I108 has different rotamers across different *Ec*FabF structures (Figure 2.5B): In *apo*-FabF, the C_δ of I108 is only 3.7 Å away from the closest carbon of the F202 side chain, forming a gate-closed conformation. When a six-carbon chain is extended from the active site cysteine, like the ACP-FabF crosslinked

structure 7L4L (PDB ID), I108 adopts a different rotamer that extends the closest carbon-to-carbon distance with F202 to 5.8 Å. Finally, the lauric acid bound FabF structure (PDB ID: 2GFW) has the acyl chain penetrating through the gate, with I108 moving further away to create a 7.5 Å carbon-to-carbon gap. It has been shown that the I108F mutant predominantly elongates C6 acid to C8, resulting from a tighter gate that inhibits the loading of longer acids. Similar mutation idea has also been applied in engineering fungal FAS for short chain FA production.[53] I108F mutation also prevents the binding of cerulenin, a KS inhibitor that penetrates through the back gate among binding and grants the mutant cerulenin resistance in both *Ec*FabF and *Bacillus subtilis* FabF (*Bs*FabF).[54-56] I108M of *Bs*FabF has also been tested to show a two-fold higher cerulenin resistance than the I108F mutant.[57]

In human mitochondrial KS (*Hs*mtKS), a bimodal substrate preference exists that results mainly in producing C8 fatty acid, the precursor of lipoic acid biosynthesis, and C14-C16 fatty acid.[58] Structures of mtKSs reveal that M154, the same position as I108 of *Ec*FabF, must rearrange to allow longer acyl chain to bind, causing the bimodal substrate profile.[59] *Ec*FabF has yet not shown preference towards a certain medium chain length substrate. Whether I108 affects the substrate profile or not is still unclear. It is interesting to note that glutaryl-ACP methyl ester, equivalent to a C7 acid in length when transferred onto the active site cysteine, is elongated by KS to generate the precursor of biotin biosynthesis.[60] The methyl ester group of the dicarboxylate might be too bulky to penetrate through the back gate, thus preventing further elongation.

Substrate specificity and chain length control

The KS pocket has direct influence on the substrate via physical restrictions, polar contacts, and Coulomb-repulsion, thereby controlling the identity of the product. These interactions can most often be identified by analysing KS crystal structures, and on top of them are substrate bound KS structures that directly visualize substrate behaviour inside the pocket. In this section, we will discuss KS's mechanism being one of the key enzymes of product chain length control in FAS and PKS,[4] and its substrate specificity realized by some key elements in the substrate pocket.

The dichotomy of FabF and FabB. In the *E. coli* FAB, FabF and FabB have most of their SFA substrate preferences overlapped (C6-C12), with the exception of C14, which is preferentially elongated by FabF.[61] In terms of UFA biosynthesis, FabB catalyses the first elongation from C10:1 to C12:1 and then both KSs can elongate C12:1 up to C16:1.[6] The final elongation to C18:1 is exclusively catalysed by FabF.[62] While FabF pocket does not have an obvious end,[63] as the homodimer pockets meet at the interface, FabB has a clear 12-carbon pocket depth restricted by the coordinating E200-Q113 residues located at the bottom of the pocket.[64] However, both residues can adopt different conformations, dissociating the hydrogen bond between side chains, to expand the pocket depth, as seen in the C12-FabB crystal structure (PDB ID: 1EK4). This explains the existing but low activity of FabB toward C14 substrate (Figure 2.6).[65] Interestingly, this longer pocket cannot coexist in both subunits of the homodimer, because such conformation of the glutamine residues (Q113), located at the center of the dimer, would clash. This set of asymmetric pockets of FabB suggest a cooperativity relationship between the subunits on long chain substrates.[66] The possible

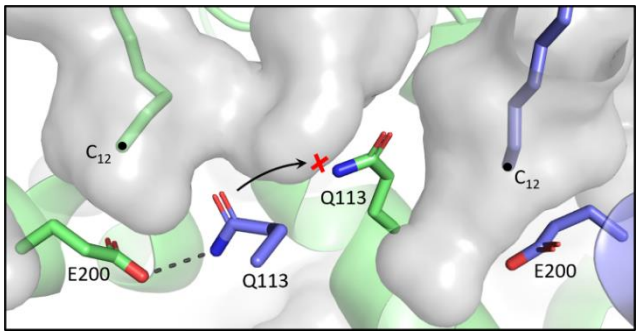


Figure 2.6: Asymmetric substrate pockets of FabB (PDB ID: 1EK4): E200-Q113 polar contact restricts the pocket size to 12-carbon on the left whereas the pocket with dissociated E200-Q113 on the right accommodates longer substrate. Q113 from each of the subunit of the homodimer cannot co-occupied the center space, leaving the pockets asymmetric.

cooperativity of the homodimeric type II KS has rarely been addressed, but cooperativity effect has recently been reported in the type I murine KS domain, as well as a type I PKS.[67,68]

Despite the lack of activity towards C16 SFA, *Ec*FabF elongates C16:1 UFA, producing the C18:1 acid that could be incorporated into membrane phospholipids to increase cell membrane fluidity.[69, 70] It has been demonstrated *in vitro* and *in vivo* that FabF plays a key role in the thermal regulation of *E. coli* by its substrate preferences under different temperature.[62, 71, 72] It is, however, not clear how the one double bond difference between C16 and C16:1 acid can result in such distinct FabF activities. The position of the double bond, which lies between C_7 and C_8, is reminiscent of the FabF back gate (I108-F202) (Figure 2.5B), but there has been no proof yet if they are related.[63] Interestingly, the two crosslink structures of FabF-ACP complex (gating loops in open form which mimics substrate delivery state) with C16 and C16:1 α-bromo probes, respectively, show distinct substrate conformation in which the *cis* double bond of C16:1 redirects the acyl chain back to the substrate pocket (Figure 2.7).[47] The kinked C16:1 substrate is likely more compact and provides a conformation that is more readily accommodated. This implies that the unsaturation may assist the process of C16:1 substrate delivery.

Another mysterious substrate preference is the necessity of FabB for elongating *cis*-3-decenoic acid (C10:1 UFA), the first UFA produced by FAB. It has been demonstrated that only FabB is capable of elongating this substrate *in vivo*[6] and in cell extract complementation assay.[73] However, *in vitro* study[74] shows that FabF has considerable activity on elongating the C10:1 substrate, and from comparing the

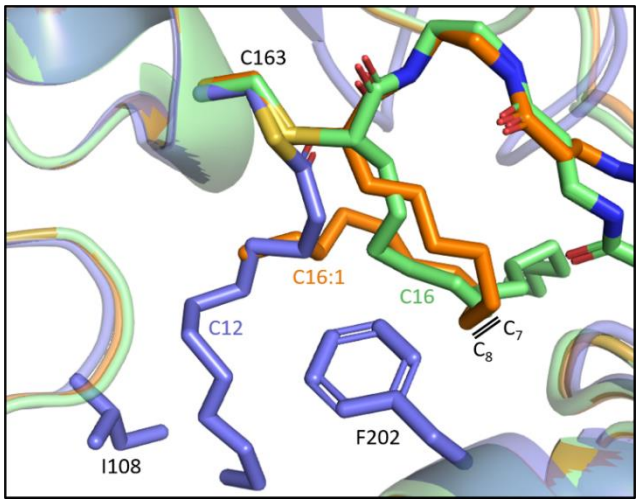


Figure 2.7: Comparison of 16-carbon SFA and UFA at substrate delivery state: overlay of C12-FabF (PDB ID: 2GFY, blue) and FabF-ACP complexes with crosslinkers mimicking C16 (PDB ID: 6OKG, green) and C16:1 (PDB ID: 7L4E, orange), respectively. The cis-double bond of C16:1 redirects the acyl chain to the substrate pocket.

pockets of the two KSs, there is no sign that FabF cannot catalyse such an elongation.[63] Therefore, FabF's may have negative selectivity towards C10:1 substrate before the substrate enters the pocket, and only under the competition of different chain length FAs. Such selectivity can potentially be achieved by allosteric regulation in combination with KS-ACP interactions. Several bacterial species do not contain a FabB orthologue, but rather have one, two, or more FabFs with at least one of them capable of complementing FabB's function, which provides clues to how the substrate preferences arise.[75-77]

Comparison of KS-CLFs and homodimeric KSs. In PKSII, the genes of the KS and CLF subunits are frequently clustered, with KS gene upstream and CLF gene downstream, and in many cases translationally coupled (i.e. overlapping stop and start codon).[28, 78] The substrate pocket of KS-CLF extends along the heterodimer interface and most often ended by one or more bulky residues of CLF.[11] Examples like F116' (The apostrophe denotes residue from the CLF subunit) of actinorhodin KS-CLF (PDB ID: 1TQY), W107' of AntD-AntE (PDB ID: 6SMP), and L125' of Iga11-Iga12 (PDB ID: 6KXF) have been demonstrated, by mutagenesis or computational method, to control the final chain length.[12, 37, 79] The KS-CLF of aryl polyene biosynthesis, ApeO-ApeC, has a wide pocket to accommodate the phenyl group of the substrate. Despite the difference in the width of the pocket, ApeO-ApeC still has its final product chain length determined by the depth of its substrate pocket.[21, 23]

In contrast to KS-CLFs, which normally have a well-defined pocket that leads to one single chain length final product, homodimeric KSs from the FAS have a chain

length profile resulting from complex factors such as indefinite pocket size, dimer cooperativity, gating mechanisms, and acyl-ACP recognition. Such chain length profile can often link to metabolic flux control.[80, 81] As previously discussed, the human mitochondrial FAB KS possesses bimodal substrate preferences that lead to two major chain length products, C8 and C14/C16. The C8 acid is the precursor of lipoic acid biosynthesis while the long chain acyl-ACP is an assembly factor for respiratory complexes as well as the iron-sulfur cluster synthesis complex.[82-86] Accumulating evidence has shown the role of mitochondrial FAB as a coordinator between acetyl-CoA level and metabolic state sensing,[87, 88] among other roles that have yet to be explored.

Protein-substrate interactions. The L-shape pocket of KSs consists of two segments, the PPant binding pocket and the acyl-chain binding pocket, respectively, that are roughly perpendicular to each other. The PPant binding pocket has two highly conserved threonine residues (T305 and T307 in *Ec*FabF) that form polar contact with the oxygen of the second amide of PPant, and potentially other carbonyl groups that pass by these threonine residues when entering the active site such as those from thioester and malonate. Such interactions anchor the PPant moiety on one side of the rather spacious PPant pocket, and this orientation seems to clear out the way for the movement of gating loop 1 (Figure 2.8A). One of these threonine residues is substituted by a serine in polyketone KSs, but the polar contact potentials provided by the side chains of these two residues exist among all KSs. To demonstrate their essential roles, mutation of either of these residues to alanine in both the IgaPKS (polyene) and AntPKS (polyketone) systems resulted in impaired enzyme

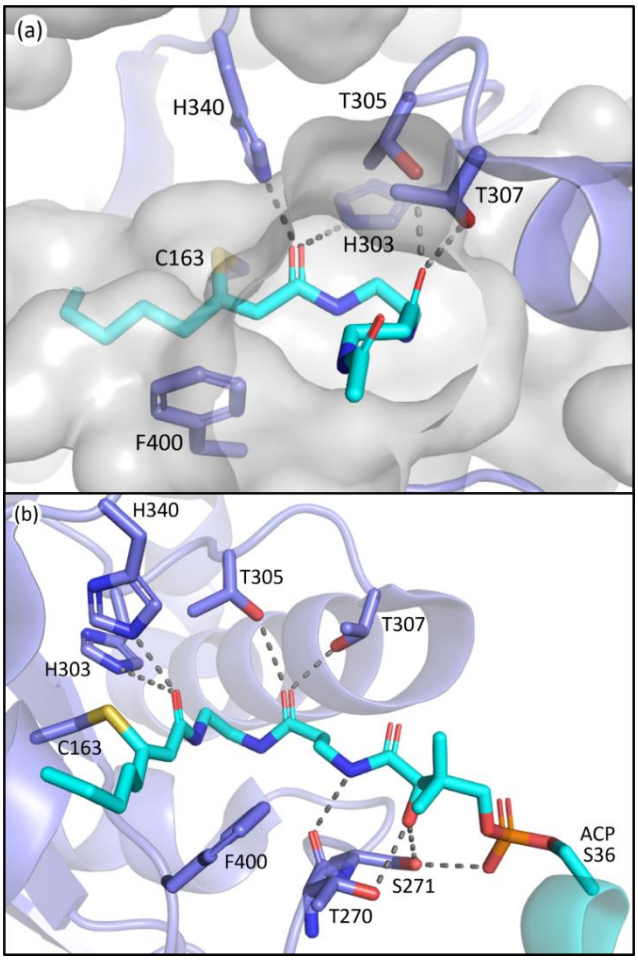


Figure 2.8: The PPant binding pocket of FabF: (A) coordination of the two threonine residues paves a hydrophilic path for PPant insertion. (B) Polar contact network of the PPant binding pocket. T270 and S271 belong to the gating loop 2.

activity.[12,37] The gating loop 2 also forms polar contacts with the phosphate and the first amide of the PPant (Figure 2.8B), potentially mediating between ACP and the gating loop 1.

In contrast to the high similarity of the ppant binding pocket, the substrate tunnels of different classes of type II KSs have distinct properties due to the diversity of substrates. The acidic side chain of the highly conserved residue D113 of IgaKS is reported to repulse the β-ketoacyl products from reloading, evidenced by the D113A mutant which produces triketone.[37] This repulsion mechanism is not observed in other kinds of KS including those of FAS as *Sp*FabF can produce triacetic acid lactone (TAL, product of triketone) *in vitro*,[89] and that the precursor of biotin biosynthesis, pimeloyl methyl ester, is also synthesized by FAS, indicating that a methyl ester group can fit into a KS substrate pocket.[60] The aryl polyene KSs of *Acinetobacter baumannii*, ApeO-ApeC and ApeR, have phenylalanine and tyrosine residues in the pockets that can stack with the phenyl ring of the substrate.[23] The hexaketide bound AntKS-CLF structures reveal the carbonyl orientations in the narrow pocket that leads to stable substrate binding predicted by simulations.[12] These examples demonstrate the customization nature of the substrate pocket, especially for PKS, in which the precision of the product is a first priority.

Protein-protein interactions of KS and ACP

KSs possess an Arg and Lys rich surface, known as the "positive patch", around the active site entrance that interacts with negatively charged residues on ACP. Although this is generally true for all KS-ACP interactions elucidated so far, molecular detail of these interactions can differ to prevent crosstalk between different KS-ACP pairs in the same organism.[11, 90, 91] This orthogonality is believed to reduce stochastic binding events and substrate sampling time, and thus accelerate the overall kinetics of catalysis. In this section, we will delineate the molecular basis of KS-ACP interactions and the mechanisms affected by them, such as chain-flipping and allosteric regulation.

Residues involved in PPIs. The high acidity of ACP and the positive patches on the partner enzymes are an obvious match for binding interface. In the past, the interface residues involve in type II KS-ACP interactions have only been addressed in a few studies using non crystallographic methods.[35, 92, 93] Since the first KS-ACP complex crystal structure utilizing mechanism based crosslinking probe was published in 2019, there are in total 10 KS-ACP complex structures released on the PDB, across 4 different KSs from FAS, polyene PKS, and aromatic PKS pathways.[12, 36, 38] By directly visualizing the protein-protein interfaces in these catalytic-relevant states, these structures allow us to assign PPIs at atomic resolution.

In figure 9, the interface polar contacts identified from four different KS-ACP pairs are aligned. Generally, most of the PPIs occur between KS' (the monomer that the catalytic site isn't involve, or CLF subunit in the case of PKSII) and the first two

helices of ACP. To facilitate discussion, PPIs are grouped into three regions. Region 1 involves a short helix on KS' (or CLF) that bears 2-3 Arg or Lys residues flanking ACP to interact with residues on C-terminal side of helix 1 and N-terminal side of helix 2 (Figure 2.9, first row). Region 2 covers interactions of helix 2 with residues on KS' (or CLF) as well as helix 3 with a single residue on KS (Figure 2.9, second row). Lastly, region 3 features the gating loop 2 of KS interacting with the N-terminal end of helix 2 and the PPant (Figure 2.9, third row). Some differences can be observed that may lead to unique PPIs. For example, region 1 of IgaPKS has only two distal basic residues on the CLF involved, when at least three residues from each side forming complicated polar contact network are seen in other KS-ACP pairs. AntPKS in region 2 also shows unique interactions, with the ACP bearing two interacting Lys residues, contradicting to the general wisdom that the ACP bears acidic residues that interact with the positive patch on the KS. *Ec*FabB in region 3 has the least interactions, with the highly conserved residues in *Ec*FabF, H268 and T270, substituted by D268 and V270, which are also highly conserved in FabB.

Alanine scan of interacting protein-protein interface residues of FabF shows that triple mutation removing all interactions in region 1 or region 2 has a greater impact on activity than a triple mutant containing a mix of residues from both regions, indicating that anchoring ACP at both regions is essential.[94] IgaKS-CLF was also subjected to the interface residues alanine scan and all of them show some degree of decline in crosslinking activity. Among them, R210A mutant stands out to totally abolish crosslinking activity because of the profound interactions R210 has with ACP

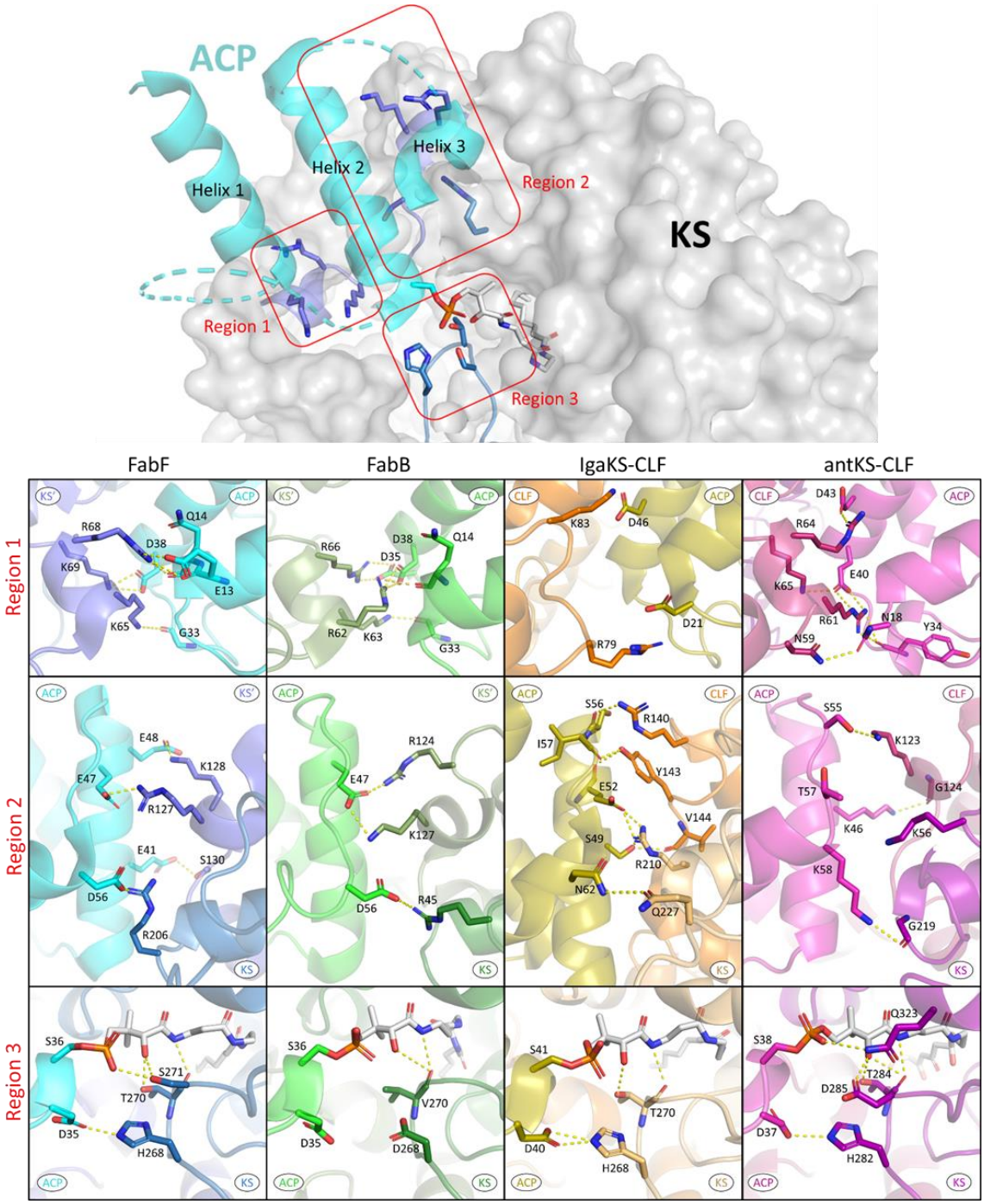


Figure 2.9: Protein-protein interactions of four different KS-ACP interfaces: PPIs are grouped into three regions as depicted on the top graph (ACP helix 4 and the inter-helices loops are not shown for simplicity). PPIs of each region from FabF (PDB ID: 7L4L, blue), FabB (PDB ID: 6OKC, green), polyene KS-CLF (PDB ID: 6KXF, yellow and orange), and polyketone KS-CLF (PDB ID: 6SMP, magenta) are aligned for comparison.

helix 2, forming polar contacts with the side chains of S49 and E52.[37] R61 of the AntCLF also plays a similar role in region 1 interacting with two ACP residues, and mutating this residue to alanine significantly reduced the enzyme activity, while mutant leads to the loss of one single polar contact between R64 of CLF and D43 of ACP retains the same level of activity.[12] Despite the ostensibly high similarity between KS-ACP interfaces, the variation of that one, or a few, crucial residues among different KSs can potentially be enough to insure the fidelity of ACP to its cognate partner.

Chain-flipping mechanism is initiated by PPIs. Carrier protein is known to be a flexible protein that can sequester substrate intermediates in type II pathways.[10, 39, 95-99] In *Ec*ACP, a hydrophobic cavity can be formed in between the four-alpha helical bundle to accommodate acyl substrates. Structural studies have shown that the pocket size can be altered through conformational movements of the helices with helix 3 and the two loops that connect helix 3 to helix 2 and helix 4, respectively, having the largest shift between the sequestered state and the unsequestered state.[97, 100] Among all the changes, the shifting of Isoleucine 62, located at the loop connecting helix 3 and 4, appears to have the biggest impact to the pocket size.

ACP delivers the substrate by de-sequestering and "flipping" the acyl chain into the KS pocket, a process termed chain-flipping. Structure analysis of KS-ACP complex reveals that the de-sequestration of ACP is induced by PPIs: By anchoring helix 2 and the C-terminal end of helix 1 and applying an attractive interaction to helix 3 resulted in shrinkage of the pocket size. Overlay of FabF-ACP structure to the helix

2 of acyl-ACP structure illustrates the mechanism mentioned above (Figure 2.10). With helix 2 and part of helix 1 well-anchored, the interaction of FabF-ACP (R206-D56) "flattens" ACP, tilting helix 1 and collapsing the pocket (Figure 2.10A). The structure overlay also features a 2.1-Å shift of I62, which is indicative of ACP switching from the sequestered state to the unsequestered state (Figure 2.10B). Mutational study on R206 reveals that mutating the residue to an alanine largely decreases the k_{cat} value of FabF while the K_M value remains the same.[94] This indicates that interaction with ACP helix 3 serves more of a catalytic role rather than strictly stabilizing the protein-protein interface. Interestingly, such interaction is quite unique across the four KSs that have KS-ACP structure solved (Figure 2.9, second row). In FabB, the corresponding position of FabF R206 is a glycine residue, and instead a R45 from neighboring loop forms polar contact with the same D56 of ACP. In IgaPKS, the interaction pair is (ACP)N62-Q227(KS), and in AntPKS, ACP becomes the side that possesses positively charged residue, K58, to interact with KS. The unique characteristic of this one single but important ACP(helix 3)-KS interaction may also contribute to the orthogonality of the type II pathways.

Allosteric regulation. ACP as a substrate carrier in the type II pathways provides an extra level of regulation before sampling substrates in the pocket of the partner enzyme. Such mechanism, termed allosteric regulation, describes indirect substrate recognition of the partner enzyme through interacting with ACP.[101, 102] Structural and computational studies on the substrate sequestration of acyl-ACP have long been supporting the feasibility of allosteric regulation because of the observed differences between ACPs bearing different

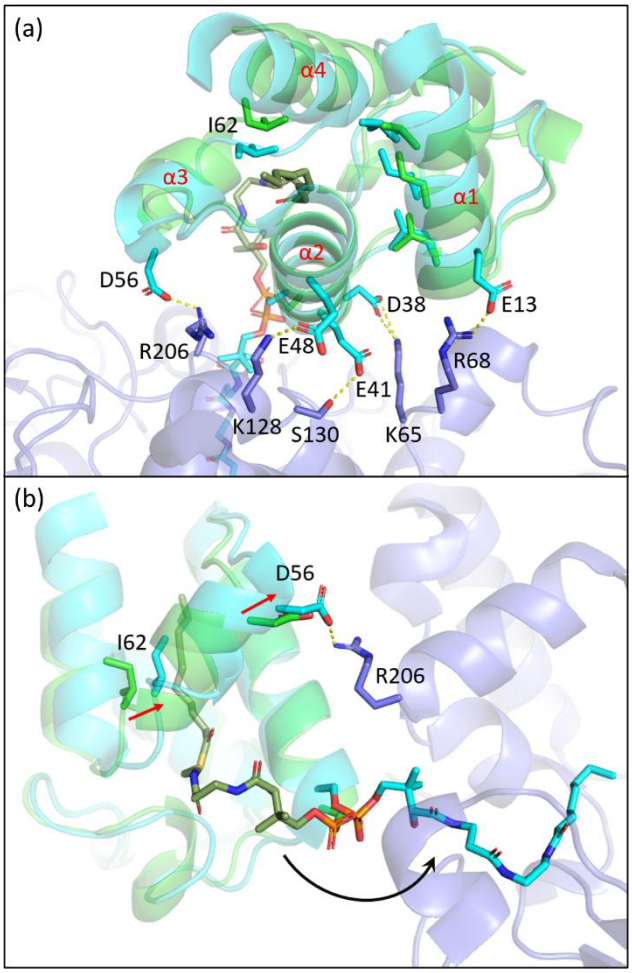


Figure 2.10: The chain-flipping mechanism: structure overlay of C10-ACP (PDB ID: 2FAE, green) to the ACP helix 2 of the FabF-ACP complex (PDB ID: 7L4L, blue) shows PPIs initiate chain-flipping. (A) PPIs anchor helix 2 and E13 and drag on helix 3 to twist the 4-helical bundle which shrinks the pocket. (B) Side view of the overlay. Movement of I62 shrinks the pocket and push the substrate out.

acyl chain length as well as other pathway intermediates.[98, 100, 101, 103-106] NMR titration study of ACP with six *E. coli* FASs reveals unique ACP-partner enzyme interacting interfaces, suggesting a complex role ACP may play in the pathway.[91] Recent works on LipB, an enzyme that directs C8-ACP to lipoic acid biosynthesis, elucidates an allostery that is chain length dependent.[101, 107] Acyl chain terminal carbon ^{13}C-labeled C8 and C12-ACP, which chain-flipping event can be tracked by HSQC NMR, are titrated with active site cysteine-mutated LipB, and only the cognate C8-ACP performs chain-flipping.[101] This study provides the first solid evidence of allosteric regulation of *Ec*ACP.

Since structural analysis of the pocket shows no clue on FabF's negative preference towards C10:1-ACP *in vivo*, allosteric regulation stands out to be a plausible explanation. Furthermore, the fact that FabF can elongate C10:1-ACP *in vitro* indicates that the preference occurs before the substrate enters the pocket,[65] and that C10:1-ACP could be competed out by other acyl-ACP, if present, causing the *in vivo* result. Interestingly, the depth of the ACP pocket is approximately 7 carbons, which is exactly the carbon count of the saturated tail of C10:1 as well as other UFAs. In the sequestered state, the extra carbons in the acyl chain and the ppant moiety are exposed to the ambient environment, even forming a hairpin when the chain gets longer. This exposure may have unique interaction with the gating loop 2 (Figure 2.9, region 3), potentially contributing to the allostery of ACP-KS. The *cis*-double bond of the C10:1 acyl chain may restrict the movement of the overhanging PPant resulting in the reduction of effective PPant conformations (those that interacts with the gating loop 2), while the same restriction may occur to other UFAs like C12:1 but overcome

by the longer chain length. Given that FabB has minimal polar contacts at region 3 (Fig. 9), it can be insensitive to the degree of exposure of the acyl-ppant, and thus the insensitivity towards distinguishing C10 and C10:1-ACP. While FabB has an extremely conserved gating loop 2 sequence, FabF loop 2, having a distinct sequence, is only partially conserved (Figure 2.3F). This can be explained by some organisms with UFA possess only FabF type of KS,[7] and some of these FabFs are reported to complement FabB activity in *E. coli*.[75-77] For example, *E. Faecalis* has no FabB but two FabFs, FabF1 and FabF2, respectively, and only FabF1 can complement *Ec*FabB *in vivo*.[75] Sequence alignment shows that *Ef*FabF2 has a more similar loop 2 sequence to *Ec*FabF than that of *Ef*FabF1. The gating loop 2 plays an important and complex role in the KS enzymology, interacting with both ACP and the substrate, and coordinating the movement of the gating loop 1. Swapping off the whole loop 2 of FabF with FabB loop 2 results in expressible FabF variant with no catalytic activity,[47] suggesting that loop 2 likely coevolved with other elements to fine-tune KS function.

Conclusion

Based on a conserved mechanism, elongating KSs have evolved for distinct pathways to produce specific product intermediates. In FASII, one or more elongating KSs with complicated gating mechanisms maintain the FA profile, which provides the precursor pool for membrane phospholipids and other essential metabolites, such as biotin and lipoic acid. PKSII polyketone KSs, on the other hand, have rigid pockets that precisely synthesize a single chain length of polyketones. Between them are the

PKSII polyene KSs that share some similarities with two other KS types but have some unique mechanisms to ensure product fidelity. KS-ACP interactions play an important role in substrate delivery, and these details will continue to emerge with more studies. Manipulating the KS-ACP interface has already shown success in engineering type I FAS and PKS pathways,[108, 109] and the recent engineering the FASII interface of *Ec*ACP and a non-cognate thioesterase indicates that this will prove valuable in PKSII as well.[110] These examples showcase the promising future of engineering FASII and PKSII pathways and would most certainly be facilitated with a thorough understanding of KS enzymology.

Acknowledgement

Chapter 2, entitled Enzymology of Standalone Elongating Ketosynthases, in full, is a reprint of the material as it appears in: Chen A, Jian Z, Burkart MD. "Enzymology of Standalone Elongating Ketosynthases", RSC Chemical Biology, under review. The dissertation author is the primary author of this manuscript.

References

(1) Hol, W. G. J.; van Duijnen, P. T.; Berendsen, H. J. C. The α-helix dipole and the properties of proteins. *Nature* **1978,** *273* (5662), 443-446.

(2) White, S. W.; Zheng, J.; Zhang, Y. M.; Rock, C. O. The structural biology of type II fatty acid biosynthesis. *Annu. Rev. Biochem.* **2005,** *74*, 791-831.

(3) Witkowski, A.; Joshi, A. K.; Smith, S. Mechanism of the β-Ketoacyl Synthase Reaction Catalyzed by the Animal Fatty Acid Synthase. *Biochemistry* **2002,** *41* (35), 10877-10887.

(4) Heil, C. S.; Wehrheim, S. S.; Paithankar, K. S.; Grininger, M. Fatty Acid Biosynthesis: Chain-Length Regulation and Control. *Chembiochem* **2019,** *20* (18), 2298-2321.

(5) Hillenmeyer, M. E.; Vandova, G. A.; Berlew, E. E.; Charkoudian, L. K. Evolution of chemical diversity by coordinated gene swaps in type II polyketide gene clusters. *Proceedings of the National Academy of Sciences* **2015,** *112* (45), 13952.

(6) Feng, Y.; Cronan, J. E. Escherichia coli unsaturated fatty acid synthesis: complex transcription of the fabA gene and in vivo identification of the essential reaction catalyzed by FabB. *J. Biol. Chem.* **2009,** *284* (43), 29526-35.

(7) Lu, Y. J.; Zhang, Y. M.; Rock, C. O., Product diversity and regulation of type II fatty acid synthases. *Biochem. Cell. Biol.* **2004,** *82* (1), 145-55.

(8) Magnuson, K.; Jackowski, S.; Rock, C. O.; Cronan, J. E., Jr. Regulation of fatty acid biosynthesis in Escherichia coli. *Microbiol. Rev.* **1993,** *57* (3), 522-42.

(9) Wang, Y.; Qian, G.; Li, Y.; Wang, Y.; Wang, Y.; Wright, S.; Li, Y.; Shen, Y.; Liu, F.; Du, L. Biosynthetic Mechanism for Sunscreens of the Biocontrol Agent Lysobacter enzymogenes. *PLOS ONE* **2013,** *8* (6), e66633.

(10) 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* (48), 16792-16799.

(11) Chen, A.; Re, R. N.; Burkart, M. D. Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions. *Nat. Prod. Rep.* **2018,** *35* (10), 1029-1045.

(12) Bräuer, A.; Zhou, Q.; Grammbitter, G. L. C.; Schmalhofer, M.; Rühl, M.; Kaila, V. R. I.; Bode, H. B.; Groll, M. Structural snapshots of the minimal PKS system responsible for octaketide biosynthesis. *Nature Chemistry* **2020,** *12* (8), 755-763.

(13) Bilyk, O.; Brötz, E.; Tokovenko, B.; Bechthold, A.; Paululat, T.; Luzhetskyy, A. New Simocyclinones: Surprising Evolutionary and Biosynthetic Insights. *ACS Chemical Biology* **2016,** *11* (1), 241-250.

(14) Du, D.; Katsuyama, Y.; Onaka, H.; Fujie, M.; Satoh, N.; Shin-Ya, K.; Ohnishi, Y. Production of a Novel Amide-Containing Polyene by Activating a Cryptic Biosynthetic Gene Cluster in Streptomyces sp. MSC090213JE08. *Chembiochem* **2016,** *17* (15), 1464-71.

(15) Du, D.; Katsuyama, Y.; Shin-ya, K.; Ohnishi, Y. Reconstitution of a Type II Polyketide Synthase that Catalyzes Polyene Formation. *Angewandte Chemie International Edition* **2018,** *57* (7), 1954-1957.

(16) Twigg, F. F.; Cai, W.; Huang, W.; Liu, J.; Sato, M.; Perez, T. J.; Geng, J.; Dror, M. J.; Montanez, I.; Tong, T. L.; Lee, H.; Zhang, W. Identifying the Biosynthetic Gene Cluster for Triacsins with an N-Hydroxytriazene Moiety. *ChemBioChem* **2019,** *20* (9), 1145-1149.

(17) Schöner, T. A.; Gassel, S.; Osawa, A.; Tobias, N. J.; Okuno, Y.; Sakakibara, Y.; Shindo, K.; Sandmann, G.; Bode, H. B. Aryl Polyenes, a Highly Abundant Class of Bacterial Natural Products, Are Functionally Related to Antioxidative Carotenoids. *ChemBioChem* **2016,** *17* (3), 247-253.

(18) Schöner, T. A.; Fuchs, S. W.; Reinhold-Hurek, B.; Bode, H. B. Identification and biosynthesis of a novel xanthomonadin-dialkylresorcinol-hybrid from Azoarcus sp. BH72. *PLoS One* **2014,** *9* (3), e90922.

(19) Schöner, T. A.; Fuchs, S. W.; Schönau, C.; Bode, H. B. Initiation of the flexirubin biosynthesis in Chitinophaga pinensis. *Microbial Biotechnology* **2014,** *7* (3), 232-241.

(20) Herman, N. A.; Kim, S. J.; Li, J. S.; Cai, W.; Koshino, H.; Zhang, W. The industrial anaerobe Clostridium acetobutylicum uses polyketides to regulate cellular differentiation. *Nature Communications* **2017,** *8* (1), 1514.

(21) Grammbitter, G. L. C.; Schmalhofer, M.; Karimi, K.; Shi, Y.-M.; Schöner, T. A.; Tobias, N. J.; Morgner, N.; Groll, M.; Bode, H. B. An Uncommon Type II PKS Catalyzes Biosynthesis of Aryl Polyene Pigments. *Journal of the American Chemical Society* **2019,** *141* (42), 16615-16623.

(22) Lee, W. C.; Choi, S.; Jang, A.; Son, K.; Kim, Y. Structural comparison of Acinetobacter baumannii β-ketoacyl-acyl carrier protein reductases in fatty acid and aryl polyene biosynthesis. *Scientific Reports* **2021,** *11* (1), 7945.

(23) Lee, W. C.; Choi, S.; Jang, A.; Yeon, J.; Hwang, E.; Kim, Y. Structural basis of the complementary activity of two ketosynthases in aryl polyene biosynthesis. *Scientific Reports* **2021,** *11* (1), 16340.

(24) Zhang, W.; Li, Y.; Tang, Y., Engineered biosynthesis of bacterial aromatic polyketides in Escherichia coli. *Proceedings of the National Academy of Sciences* **2008,** *105* (52), 20683.

(25) Cummings, M.; Peters, A. D.; Whitehead, G. F. S.; Menon, B. R. K.; Micklefield, J.; Webb, S. J.; Takano, E. Assembling a plug-and-play production line for combinatorial biosynthesis of aromatic polyketides in Escherichia coli. *PLOS Biology* **2019,** *17* (7), e3000347.

(26) Klein, J. G.; Wu, Y.; Kokona, B.; Charkoudian, L. K. Widening the bottleneck: Heterologous expression, purification, and characterization of the Ktedonobacter racemifer minimal type II polyketide synthase in Escherichia coli. *Bioorg. Med. Chem.* **2020,** *28* (20), 115686.

(27) Zhang, J.; Yuzawa, S.; Thong, W. L.; Shinada, T.; Nishiyama, M.; Kuzuyama, T. Reconstitution of a Highly Reducing Type II PKS System Reveals 6π-Electrocyclization Is Required for o-Dialkylbenzene Biosynthesis. *Journal of the American Chemical Society* **2021,** *143* (7), 2962-2969.

(28) Liu, X.; Hua, K.; Liu, D.; Wu, Z.-L.; Wang, Y.; Zhang, H.; Deng, Z.; Pfeifer, B. A.; Jiang, M. Heterologous Biosynthesis of Type II Polyketide Products Using E. coli. *ACS Chemical Biology* **2020,** *15* (5), 1177-1183.

(29) Park, D.; Swayambhu, G.; Pfeifer, B. A. Heterologous biosynthesis as a platform for producing new generation natural products. *Current Opinion in Biotechnology* **2020,** *66*, 123-130.

(30) Katsuyama, Y.; Ohnishi, Y. Chapter Sixteen - Type III Polyketide Synthases in Microorganisms. In *Methods in Enzymology*, Hopwood, D. A., Ed. Academic Press: **2012**; Vol. 515, pp 359-377.

(31) Yu, D.; Xu, F.; Zeng, J.; Zhan, J. Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* **2012,** *64* (4), 285-295.

(32) 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), 687-691.

(33) Worthington, A. S.; Hur, G. H.; Meier, J. L.; Cheng, Q.; Moore, B. S.; Burkart, M. D. Probing the compatibility of type II ketosynthase-carrier protein partners. *Chembiochem* **2008,** *9* (13), 2096-103.

(34) Acheampong, K. K.; Kokona, B.; Braun, G. A.; Jacobsen, D. R.; Johnson, K. A.; Charkoudian, L. K. Colorimetric Assay Reports on Acyl Carrier Protein Interactions. *Sci Rep* **2019,** *9* (1), 15589.

(35) Ye, Z.; Williams, G. J., Mapping a Ketosynthase:Acyl Carrier Protein Binding Interface via Unnatural Amino Acid-Mediated Photo-Cross-Linking. *Biochemistry* **2014,** *53* (48), 7494-7502.

(36) Milligan, J. C.; Lee, D. J.; Jackson, D. R.; Schaub, A. J.; Beld, J.; Barajas, J. F.; Hale, J. J.; Luo, R.; Burkart, M. D.; Tsai, S. C. Molecular basis for interactions between an acyl carrier protein and a ketosynthase. *Nat Chem Biol* **2019,** *15* (7), 669-671.

(37) Du, D.; Katsuyama, Y.; Horiuchi, M.; Fushinobu, S.; Chen, A.; Davis, T. D.; Burkart, M. D.; Ohnishi, Y. Structural basis for selectivity in a highly reducing type II polyketide synthase. *Nature Chemical Biology* **2020,** *16* (7), 776-782.

(38) Mindrebo, J. T.; Patel, A.; Kim, W. E.; Davis, T. D.; Chen, A.; Bartholow, T. G.; La Clair, J. J.; McCammon, J. A.; Noel, J. P.; Burkart, M. D. Gating mechanism of elongating β-ketoacyl-ACP synthases. *Nature Communications* **2020,** *11* (1), 1727.

(39) Evans, S. E.; Williams, C.; Arthur, C. J.; Płoskoń, E.; Wattana-amorn, P.; Cox, R. J.; Crosby, J.; Willis, C. L.; Simpson, T. J.; Crump, M. P. Probing the Interactions of early polyketide intermediates with the Actinorhodin ACP from S. coelicolor A3(2). *J. Mol. Biol.* **2009,** *389* (3), 511-28.

(40) Barajas, J. F.; Shakya, G.; Moreno, G.; Rivera, H.; Jackson, D. R.; Topper, C. L.; Vagstad, A. L.; La Clair, J. J.; Townsend, C. A.; Burkart, M. D.; Tsai, S.-C. Polyketide mimetics yield structural and mechanistic insights into product template domain function in nonreducing polyketide synthases. *Proceedings of the National Academy of Sciences* **2017,** *114* (21), E4142.

(41) Dong, X.; Bailey, C. D.; Williams, C.; Crosby, J.; Simpson, T. J.; Willis, C. L.; Crump, M. P. Recognition of extended linear and cyclised polyketide mimics by a type II acyl carrier protein. *Chemical science* **2016,** *7* (3), 1779-1785.

(42) Das, A.; Khosla, C. Biosynthesis of Aromatic Polyketides in Bacteria. *Accounts of Chemical Research* **2009,** *42* (5), 631-639.

(43) Ellis, B. D.; Milligan, J. C.; White, A. R.; Duong, V.; Altman, P. X.; Mohammed, L. Y.; Crump, M. P.; Crosby, J.; Luo, R.; Vanderwal, C. D.; Tsai, S.-C. An Oxetane-Based Polyketide Surrogate to Probe Substrate Binding in a Polyketide Synthase. *Journal of the American Chemical Society* **2018,** *140* (15), 4961-4964.

(44) Beld, J.; Cang, H.; Burkart, M. D. Visualizing the Chain-Flipping Mechanism in Fatty-Acid Biosynthesis. *Angewandte Chemie International Edition* **2014,** *53* (52), 14456-14461.

(45) Charov, K.; Burkart, M. D. A Single Tool to Monitor Multiple Protein-Protein Interactions of the Escherichia coli Acyl Carrier Protein. *ACS Infect. Dis.* **2019,** *5* (9), 1518-1523.

(46) Sulpizio, A.; Crawford, C. E. W.; Koweek, R. S.; Charkoudian, L. K. Probing the structure and function of acyl carrier proteins to unlock the strategic redesign of type II polyketide biosynthetic pathways. *Journal of Biological Chemistry* **2021,** *296*.

(47) Mindrebo, J. T.; Chen, A.; Kim, W. E.; Re, R. N.; Davis, T. D.; Noel, J. P.; Burkart, M. D. Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases. *ACS Catalysis* **2021,** *11* (12), 6787-6799.

(48) Gora, A.; Brezovsky, J.; Damborsky, J. Gates of Enzymes. *Chemical Reviews* **2013,** *113* (8), 5871-5923.

(49) Luckner, S. R.; Machutta, C. A.; Tonge, P. J.; Kisker, C. Crystal structures of Mycobacterium tuberculosis KasA show mode of action within cell wall biosynthesis and its inhibition by thiolactomycin. *Structure (London, England : 1993)* **2009,** *17* (7), 1004-1013.

(50) Witkowski, A.; Joshi, A. K.; Lindqvist, Y.; Smith, S., Conversion of a beta-ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine. *Biochemistry* **1999,** *38* (36), 11643-50.

(51) Wang, J.; Soisson, S. M.; Young, K.; Shoop, W.; Kodali, S.; Galgoci, A.; Painter, R.; Parthasarathy, G.; Tang, Y. S.; Cummings, R.; Ha, S.; Dorso, K.; Motyl, M.; Jayasuriya, H.; Ondeyka, J.; Herath, K.; Zhang, C.; Hernandez, L.; Allocco, J.; Basilio, A.; Tormo, J. R.; Genilloud, O.; Vicente, F.; Pelaez, F.; Colwell, L.; Lee, S. H.; Michael, B.; Felcetto, T.; Gill, C.; Silver, L. L.; Hermes, J. D.; Bartizal, K.; Barrett, J.; Schmatz, D.; Becker, J. W.; Cully, D.; Singh, S. B. Platensimycin is a

selective FabF inhibitor with potent antibiotic properties. *Nature* **2006,** *441* (7091), 358-61.

(52) Lee, W.; Engels, B. The protonation state of catalytic residues in the resting state of KasA revisited: detailed mechanism for the activation of KasA by its own substrate. *Biochemistry* **2014,** *53* (5), 919-931.

(53) Gajewski, J.; Pavlovic, R.; Fischer, M.; Boles, E.; Grininger, M. Engineering fungal de novo fatty acid synthesis for short chain fatty acid production. *Nature communications* **2017,** *8*, 14650-14650.

(54) Val, D.; Banu, G.; Seshadri, K.; Lindqvist, Y.; Dehesh, K. Re-engineering ketoacyl synthase specificity. *Structure* **2000,** *8* (6), 565-6.

(55) Schujman, G. E.; Choi, K. H.; Altabe, S.; Rock, C. O.; de Mendoza, D. Response of Bacillus subtilis to cerulenin and acquisition of resistance. *Journal of bacteriology* **2001,** *183* (10), 3032-3040.

(56) Moche, M.; Schneider, G.; Edwards, P.; Dehesh, K.; Lindqvist, Y. Structure of the complex between the antibiotic cerulenin and its target, beta-ketoacyl-acyl carrier protein synthase. *J. Biol. Chem.* **1999,** *274* (10), 6031-4.

(57) Trajtenberg, F.; Altabe, S.; Larrieux, N.; Ficarra, F.; de Mendoza, D.; Buschiazzo, A.; Schujman, G. E. Structural insights into bacterial resistance to cerulenin. *Febs j* **2014,** *281* (10), 2324-38.

(58) Zhang, L.; Joshi, A. K.; Hofmann, J.; Schweizer, E.; Smith, S. Cloning, expression, and characterization of the human mitochondrial beta-ketoacyl synthase. Complementation of the yeast CEM1 knock-out strain. *J Biol Chem* **2005,** *280* (13), 12422-9.

(59) Christensen, C. E.; Kragelund, B. B.; von Wettstein-Knowles, P.; Henriksen, A. Structure of the human beta-ketoacyl [ACP] synthase from the mitochondrial type II fatty acid synthase. *Protein Sci* **2007,** *16* (2), 261-72.

(60) Lin, S. W.; Hanson, R. E.; Cronan, J. E. Biotin Synthesis Begins by Hijacking the Fatty Acid Synthetic Pathway. *Nature chemical biology* **2010,** *6*, 682 - 688.

(61) Edwards, P.; Nelsen, J. S.; Metz, J. G.; Dehesh, K. Cloning of the fabF gene in an expression vector and in vitro characterization of recombinant fabF and fabB encoded enzymes from Escherichia coli. *FEBS Lett* **1997,** *402* (1), 62-6.

(62) Garwin, J. L.; Klages, A. L.; Cronan, J. E. Beta-ketoacyl-acyl carrier protein synthase II of Escherichia coli. Evidence for function in the thermal regulation of fatty acid synthesis. *Journal of Biological Chemistry* **1980,** *255* (8), 3263-3265.

(63) von Wettstein-Knowles, P.; Olsen, J.; Arnvig Mcguire, K.; Larsen, S. Molecular aspects of beta-ketoacyl synthase (KAS) catalysis. *Biochem. Soc. Trans.* **2000,** *28* (6), 601-7.

(64) 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 Lett* **1999,** *460* (1), 46-52.

(65) McGuire, K. A.; Siggaard-Andersen, M.; Bangera, M. G.; Olsen, J. G.; von Wettstein-Knowles, P. beta-Ketoacyl-[acyl carrier protein] synthase I of Escherichia coli: aspects of the condensation mechanism revealed by analyses of mutations in the active site pocket. *Biochemistry* **2001,** *40* (33), 9836-45.

(66) von Wettstein-Knowles, P.; Olsen, J. G.; McGuire, K. A.; Henriksen, A. Fatty acid synthesis. Role of active site histidines and lysine in Cys-His-His-type beta-ketoacyl-acyl carrier protein synthases. *FEBS J.* **2006,** *273* (4), 695-710.

(67) Rittner, A.; Paithankar, K. S.; Himmler, A.; Grininger, M. Type I fatty acid synthase trapped in the octanoyl-bound state. *Protein Science* **2020,** *29* (2), 589-605.

(68) Bagde, S. R.; Mathews, I. I.; Fromme, J. C.; Kim, C.-Y. Modular polyketide synthase contains two reaction chambers that operate asynchronously. *Science* **2021,** *374* (6568), 723-729.

(69) Mejía, R.; Gómez-Eichelmann, M. C.; Fernández, M. S. Escherichia coli membrane fluidity as detected by excimerization of dipyrenylpropane: sensitivity to the bacterial fatty acid profile. *Arch. Biochem. Biophys.* **1999,** *368* (1), 156-60.

(70) Gelmann, E. P.; Cronan, J. E. Mutant of Escherichia coli Deficient in the Synthesis of cis-Vaccenic Acid. *Journal of Bacteriology* **1972,** *112* (1), 381-387.

(71) Garwin, J. L.; Cronan, J. E., Jr. Thermal modulation of fatty acid synthesis in Escherichia coli does not involve de novo enzyme synthesis. *Journal of bacteriology* **1980,** *141* (3), 1457-1459.

(72) 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* (24), 11949-11956.

(73) Cronan, J. E., Jr.; Birge, C. H.; Vagelos, P. R. Evidence for two genes specifically involved in unsaturated fatty acid biosynthesis in Escherichia coli. *J Bacteriol* **1969,** *100* (2), 601-4.

(74) D'Agnolo, G.; Rosenfeld, I. S.; Vagelos, P. R. Multiple forms of beta-ketoacyl-acyl carrier protein synthetase in Escherichia coli. *J. Biol. Chem.* **1975,** *250* (14), 5289-94.

(75) Wang, H.; Cronan, J. E. Functional replacement of the FabA and FabB proteins of Escherichia coli fatty acid synthesis by Enterococcus faecalis FabZ and FabF homologues. *J. Biol. Chem.* **2004,** *279* (33), 34489-95.

(76) Morgan-Kiss, R. M.; Cronan, J. E. The Lactococcus lactis FabF fatty acid synthetic enzyme can functionally replace both the FabB and FabF proteins of Escherichia coli and the FabH protein of Lactococcus lactis. *Arch. Microbiol.* **2008,** *190* (4), 427-437.

(77) Zhu, L.; Cheng, J.; Luo, B.; Feng, S.; Lin, J.; Wang, S.; Cronan, J. E.; Wang, H. Functions of the Clostridium acetobutylicium FabF and FabZ proteins in unsaturated fatty acid biosynthesis. *BMC Microbiol.* **2009,** *9*, 119.

(78) Huber, M.; Faure, G.; Laass, S.; Kolbe, E.; Seitz, K.; Wehrheim, C.; Wolf, Y. I.; Koonin, E. V.; Soppa, J. Translational coupling via termination-reinitiation in archaea and bacteria. *Nature Communications* **2019,** *10* (1), 4006.

(79) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. An antibiotic factory caught in action. *Nature Structural & Molecular Biology* **2004,** *11* (9), 888-893.

(80) Lowry, B.; Li, X.; Robbins, T.; Cane, D. E.; Khosla, C., A Turnstile Mechanism for the Controlled Growth of Biosynthetic Intermediates on Assembly Line Polyketide Synthases. *ACS Central Science* **2016,** *2* (1), 14-20.

(81) Srivastava, S.; Chan, C. Application of metabolic flux analysis to identify the mechanisms of free fatty acid toxicity to human hepatoma cell line. *Biotechnology and Bioengineering* **2008,** *99* (2), 399-410.

(82) Zhu, J.; Vinothkumar, K. R.; Hirst, J. Structure of mammalian respiratory complex I. *Nature* **2016,** *536* (7616), 354-358.

(83) Angerer, H.; Schönborn, S.; Gorka, J.; Bahr, U.; Karas, M.; Wittig, I.; Heidler, J.; Hoffmann, J.; Morgner, N.; Zickermann, V. Acyl modification and binding of

mitochondrial ACP to multiprotein complexes. *Biochim. Biophys. Acta Mol. Cell Res.* **2017,** *1864* (10), 1913-1920.

(84) Cory, S. A.; Van Vranken, J. G.; Brignole, E. J.; Patra, S.; Winge, D. R.; Drennan, C. L.; Rutter, J.; Barondeau, D. P. Structure of human Fe–S assembly subcomplex reveals unexpected cysteine desulfurase architecture and acyl-ACP–ISD11 interactions. *Proceedings of the National Academy of Sciences* **2017,** *114* (27), E5325.

(85) Boniecki, M. T.; Freibert, S. A.; Mühlenhoff, U.; Lill, R.; Cygler, M. Structure and functional dynamics of the mitochondrial Fe/S cluster synthesis complex. *Nature Communications* **2017,** *8* (1), 1287.

(86) Fox, N. G.; Yu, X.; Feng, X.; Bailey, H. J.; Martelli, A.; Nabhan, J. F.; Strain-Damerell, C.; Bulawa, C.; Yue, W. W.; Han, S. Structure of the human frataxin-bound iron-sulfur cluster assembly complex provides insight into its activation mechanism. *Nature Communications* **2019,** *10* (1), 2210.

(87) Masud, A. J.; Kastaniotis, A. J.; Rahman, M. T.; Autio, K. J.; Hiltunen, J. K. Mitochondrial acyl carrier protein (ACP) at the interface of metabolic state sensing and mitochondrial function. *Biochim. Biophys. Acta Mol. Cell. Res.* **2019,** *1866* (12), 118540.

(88) Van Vranken, J. G.; Nowinski, S. M.; Clowers, K. J.; Jeong, M. Y.; Ouyang, Y.; Berg, J. A.; Gygi, J. P.; Gygi, S. P.; Winge, D. R.; Rutter, J. ACP Acylation Is an Acetyl-CoA-Dependent Modification Required for Electron Transport Chain Assembly. *Mol. Cell.* **2018,** *71* (4), 567-580.e4.

(89) Zhang, Y. M.; Hurlbert, J.; White, S. W.; Rock, C. O. Roles of the active site water, histidine 303, and phenylalanine 396 in the catalytic mechanism of the elongation condensing enzyme of Streptococcus pneumoniae. *J. Biol. Chem.* **2006,** *281* (25), 17390-17399.

(90) Tang, Y.; Lee, T. S.; Kobayashi, S.; Khosla, C. Ketosynthases in the initiation and elongation modules of aromatic polyketide synthases have orthogonal acyl carrier protein specificity. *Biochemistry* **2003,** *42* (21), 6588-95.

(91) Bartholow, T. G.; Sztain, T.; Patel, A.; Lee, D. J.; Young, M. A.; Abagyan, R.; Burkart, M. D. Elucidation of transient protein-protein interactions within carrier protein-dependent biosynthesis. *Communications Biology* **2021,** *4* (1), 340.

(92) Zhang, Y.-M.; Rao, M. S.; Heath, R. J.; Price, A. C.; Olson, A. J.; Rock, C. O.; White, S. W. Identification and Analysis of the Acyl Carrier Protein (ACP) Docking

Site on β-Ketoacyl-ACP Synthase III. *Journal of Biological Chemistry* **2001,** *276* (11), 8231-8238.

(93) Beltran-Alvarez, P.; Arthur, C. J.; Cox, R. J.; Crosby, J.; Crump, M. P.; Simpson, T. J. Preliminary kinetic analysis of acyl carrier protein-ketoacylsynthase interactions in the actinorhodin minimal polyketide synthase. *Mol Biosyst* **2009,** *5* (5), 511-8.

(94) Mindrebo, J. T.; Misson, L. E.; Johnson, C.; Noel, J. P.; Burkart, M. D. Activity Mapping the Acyl Carrier Protein: Elongating Ketosynthase Interaction in Fatty Acid Biosynthesis. *Biochemistry* **2020,** *59* (38), 3626-3638.

(95) Zornetzer, G. A.; Fox, B. G.; Markley, J. L. Solution structures of spinach acyl carrier protein with decanoate and stearate. *Biochemistry* **2006,** *45* (16), 5217-27.

(96) Kosa, N. M.; Haushalter, R. W.; Smith, A. R.; Burkart, M. D. Reversible labeling of native and fusion-protein motifs. *Nature methods* **2012,** *9* (10), 981-984.

(97) Roujeinikova, A.; Baldock, C.; Simon, W. J.; Gilroy, J.; Baker, P. J.; Stuitje, A. R.; Rice, D. W.; Slabas, A. R.; Rafferty, J. B. X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site. *Structure* **2002,** *10* (6), 825-35.

(98) Płoskoń, E.; Arthur, C. J.; Kanari, A. L. P.; Wattana-amorn, P.; Williams, C.; Crosby, J.; Simpson, T. J.; Willis, C. L.; Crump, M. P. Recognition of Intermediate Functionality by Acyl Carrier Protein over a Complete Cycle of Fatty Acid Biosynthesis. *Chemistry & Biology* **2010,** *17* (7), 776-785.

(99) Haushalter, R. W.; Filipp, F. V.; Ko, K. S.; Yu, R.; Opella, S. J.; Burkart, M. D. Binding and "pKa" modulation of a polycyclic substrate analogue in a type II polyketide acyl carrier protein. *ACS Chem. Biol.* **2011,** *6* (5), 413-8.

(100) 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. *J. Mol. Biol.* **2007,** *365* (1), 135-45.

(101) Sztain, T.; Bartholow, T. G.; Lee, D. J.; Casalino, L.; Mitchell, A.; Young, M. A.; Wang, J.; McCammon, J. A.; Burkart, M. D. Decoding allosteric regulation by the acyl carrier protein. *Proceedings of the National Academy of Sciences* **2021,** *118* (16), e2025597118.

(102) Liu, J.; Nussinov, R., Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLoS Comput Biol* **2016,** *12* (6), e1004966.

(103) Zornetzer, G. A.; Tanem, J.; Fox, B. G.; Markley, J. L. The Length of the Bound Fatty Acid Influences the Dynamics of the Acyl Carrier Protein and the Stability of the Thioester Bond. *Biochemistry* **2010,** *49* (3), 470-477.

(104) Chan, D. I.; Stockner, T.; Tieleman, D. P.; Vogel, H. J., Molecular dynamics simulations of the Apo-, Holo-, and acyl-forms of Escherichia coli acyl carrier protein. *J Biol Chem* **2008,** *283* (48), 33620-9.

(105) Farmer, R.; Thomas, C. M.; Winn, P. J. Structure, function and dynamics in acyl carrier proteins. *PLOS ONE* **2019,** *14* (7), e0219435.

(106) Crosby, J.; Crump, M. P. The structural role of the carrier protein--active controller or passive carrier. *Nat. Prod. Rep.* **2012,** *29* (10), 1111-37.

(107) Bartholow, T. G.; Sztain, T.; Young, M. A.; Davis, T. D.; Abagyan, R.; Burkart, M. D. Protein–protein interaction based substrate control in the E. coli octanoic acid transferase, LipB. *RSC Chemical Biology* **2021,** *2* (5), 1466-1473.

(108) Rossini, E.; Gajewski, J.; Klaus, M.; Hummer, G.; Grininger, M. Analysis and engineering of substrate shuttling by the acyl carrier protein (ACP) in fatty acid synthases (FASs). *Chemical Communications* **2018,** *54* (82), 11606-11609.

(109) Klaus, M.; Ostrowski, M. P.; Austerjost, J.; Robbins, T.; Lowry, B.; Cane, D. E.; Khosla, C. Protein-Protein Interactions, Not Substrate Recognition, Dominate the Turnover of Chimeric Assembly Line Polyketide Synthases. *The Journal of biological chemistry* **2016,** *291* (31), 16404-16415.

(110) Sarria, S.; Bartholow, T. G.; Verga, A.; Burkart, M. D.; Peralta-Yahya, P. Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production. *ACS Synthetic Biology* **2018,** *7* (5), 1179-1187.

Chapter 3: Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases

Abstract

Ketosynthases (KSs) catalyze carbon-carbon bond-forming reactions in fatty acid synthases (FASs) and polyketide synthases (PKSs). KSs utilize a two-step ping-pong kinetic mechanism to carry out an overall decarboxylative thio-Claisen condensation that can be separated into the transacylation and condensation reactions. In both steps, an acyl carrier protein (ACP) delivers thioester-tethered substrates to the active sites of KSs. Therefore, protein-protein interactions (PPIs) and KS-mediated substrate recognition events are required for catalysis. Recently, crystal structures of Escherichia coli elongating type II FAS KSs, FabF and FabB, in complex with E. coli ACP, AcpP, revealed distinct conformational states of two active site KS loops. These loops were proposed to operate via a gating mechanism to coordinate substrate recognition and delivery followed by catalysis. Here, we interrogate this proposed gating mechanism by solving two additional high-resolution structures of substrate-engaged AcpP-FabF complexes, one of which provides the missing AcpP-FabF gate-closed conformation. Clearly defined interactions of one of these active site loops with AcpP are present in both the open and closed conformations, suggesting AcpP binding triggers or stabilizes gating transitions, further implicating PPIs in carrier protein-dependent catalysis. We functionally demonstrate the importance of gating in the overall KS condensation reaction and provide experimental evidence for its role in the transacylation reaction. Furthermore, we evaluate the catalytic importance of these loops using alanine-scanning mutagenesis and also investigate chimeric FabF constructs carrying elements found

in type I PKS KS domains. These findings broaden our understanding of the KS mechanism, which advances future engineering efforts in both FASs and evolutionarily related PKSs.

Introduction

Fatty acid biosynthesis (FAB) is an essential primary metabolic pathway conducive to metabolic engineering.[1-10] Fatty acid synthases (FAS) carry out biosynthesis by iteratively condensing and reducing malonyl-CoA derived 2-carbon units.[1,2] FAS can be organized as either large, single gene, multidomain megasynthases, type I, or as discrete, multiple gene, monofunctional enzymes, type II.[1,3,11-13] Despite differences in structural organization, the biochemistry of type I and II FAS is generally conserved (Figures S3.1 and S3.2).[2] Additionally, type II FASs are considered to be the evolutionary progenitors of polyketide synthases (PKSs), a class of enzymes known to produce a variety of structurally complex and biologically active compounds (Figure S3.1).[14-18] Ketosynthases (KSs) initiate each round of FAS by condensing a growing acyl-chain with 2-carbon units via a decarboxylative Claisen condensation with the 3-carbon malonyl-CoA substrate to produce a β -keto intermediate. Subsequent reactions catalyzed by the ketoreductase, dehydratase, and enoylreductase fully reduce the β -carbon before another round of chain extension or offloading of the mature fatty acid (FA). Central to this process is the small 9 kDa acyl carrier protein (ACP). The ACP carries thioester-tethered pathway intermediates to each enzyme active site using a post-translationally installed 4'-phosphopantetheine arm (PPant).[19,20] ACPs must form transient, productive protein-protein interactions (PPIs) with

each respective partner enzyme (PE) in order to deliver their substrates in catalytically competent forms.[21-30]

KS-mediated carbon-carbon bond formation is the primary driving force of FAB. These enzymes use a two-step kinetic ping-pong bi-bi mechanism that can be viewed mechanistically as transacylation and condensation reactions (Figure 3.1A).[30-33] In the transacylation step, acyl-ACP binds to the KS and transfers its thioester-bound cargo to the KS active site cysteine residue, producing an acyl-KS thio-intermediate.[34-36] In the condensation step, malonyl-ACP associates with the KS and undergoes decarboxylation to produce a keto-enolate tautomerization, wherein the carbanion tautomer condenses with the thioester-bound FA to produce the β-ketoacyl-ACP product (Figure S3.2). Mechanistic elements of the transacylation and condensation steps are still contested, and a detailed molecular understanding of KS substrate specificity is, therefore, lacking.[31-33] FA chain extension is energetically coupled to decarboxylation, making the forward reaction thermodynamically favored and essentially irreversible upon loss of CO_2. Therefore, KSs play a central role in controlling metabolic flux and FA chain length.[4,31] This point has been well demonstrated by recent successes in rationally retooling the product profile of the fungal FAS.[6,37-39] Future engineering efforts, in both FAS and PKS platforms, will benefit from a more in depth understanding of the reaction mechanisms and substrate specificities of functionally diverse KSs.[40,41]

Escherichia coli type II FAS has long served as a model system for understanding the structure, function, and mechanism of FASs.[1,42,43] *E. coli* possesses two related elongating KSs, FabB and FabF, that have broad overlapping substrate specificities and are

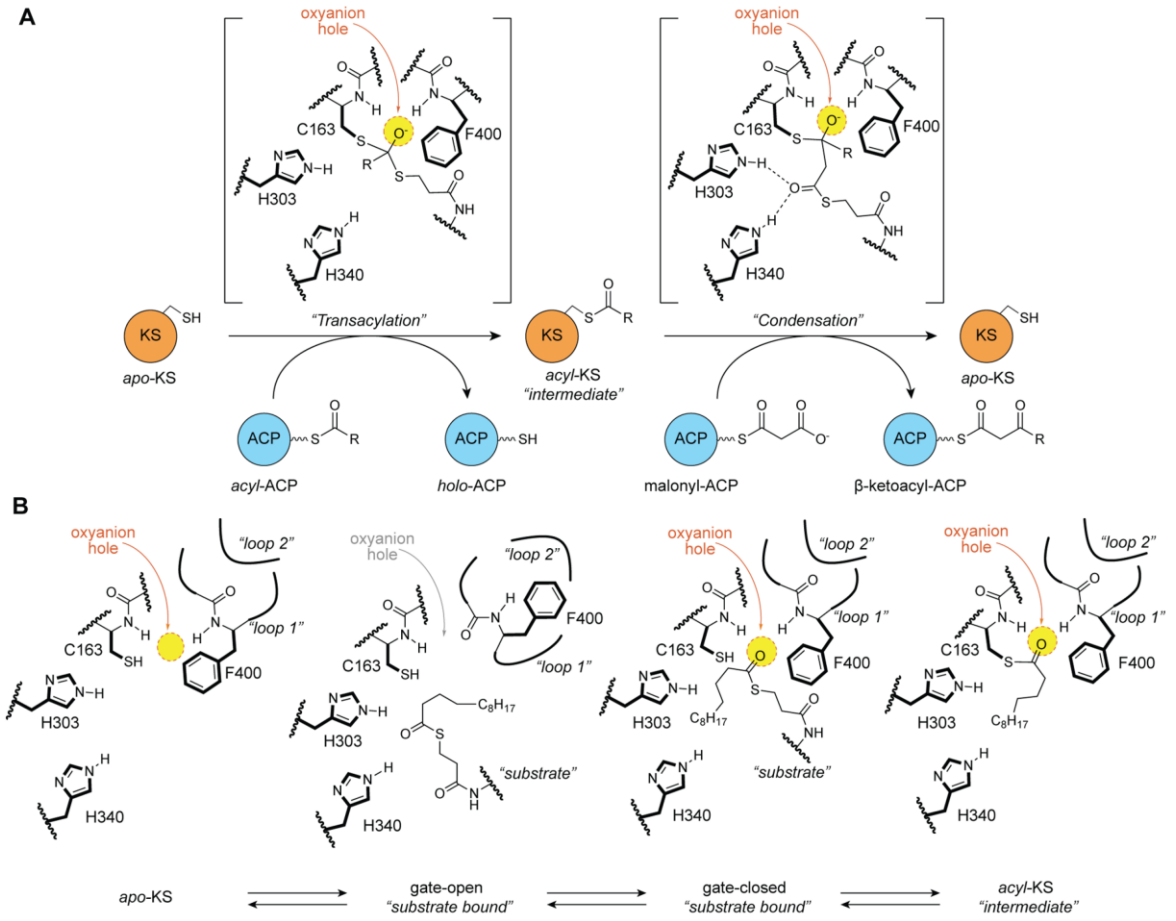


Figure 3.1: KS ping-pong reaction and gating mechanism overview. (A) Ketosynthase ping-pong mechanism with active site schematic showing oxyanion hole (yellow circle)-stabilized tetrahedral intermediates that are formed during the transacylation and condensation half-reactions. (B) Proposed ketosynthase gating mechanism coordinated by loops 1 and 2 and the Phe400 gating residue. Each active site represents a step forward during the transacylation half-reaction starting with the *apo*-KS active site and ending with the acyl-KS intermediate.

representative of specific KS families, KASI and KASII, respectively.[44,45] The active sites of elongating KSs are comprised of a Cys-His-His catalytic triad, where the cysteine, positioned at the N-terminus of a long α-helix, is the active site nucleophile and the two histidines participate in the decarboxylation step. The oxyanion hole formed by the backbone amides of the catalytic cysteine (Cys163) and a highly conserved phenylalanine gating residue (Phe400) stabilizes oxyanion tetrahedral intermediates formed during both half-reactions (Figure 3.1A).[34-36] Additionally, Phe400 is known to regulate access to the active site by sensing the acylation state of the catalytic cysteine.[32,34-36]

In our previous work, we determined crystal structures of FabB and FabF crosslinked to *E. coli* ACP (AcpP) loaded with either C12- or C16-α-bromo-pantetheinamide substrate mimetics.[24] These structures, along with biochemical and computational studies, led to the proposal of a double drawbridge-like gating mechanism[46] comprised of two active site loops, loop 1 and loop 2, that undergo conformational rearrangements to regulate KS substrate recognition and processing (Figure 3.1B). During substrate delivery, the gate opens to allow access to the KS active site, and in doing so, disrupts the oxyanion hole (Figure 3.1B). The active site machinery only reforms upon gate closure, thereby providing a means to regulate selectivity by organizing the oxyanion hole for acyl transfer when the correct substrate is bound. Loop 1 consists of a highly conserved GFGG motif, which includes the conserved phenylalanine gating residue, Phe400. Loop 2 is distal to the active site and abuts loop 1 in the gate-closed conformation. A conserved loop 2 aspartate residue, Asp265, stabilizes the gate-open state by coordinating a complex hydrogen-bonding interaction network with loop 1 (Figure 3.2A). Interestingly, much of the loop 2 motif appears to be only conserved within,

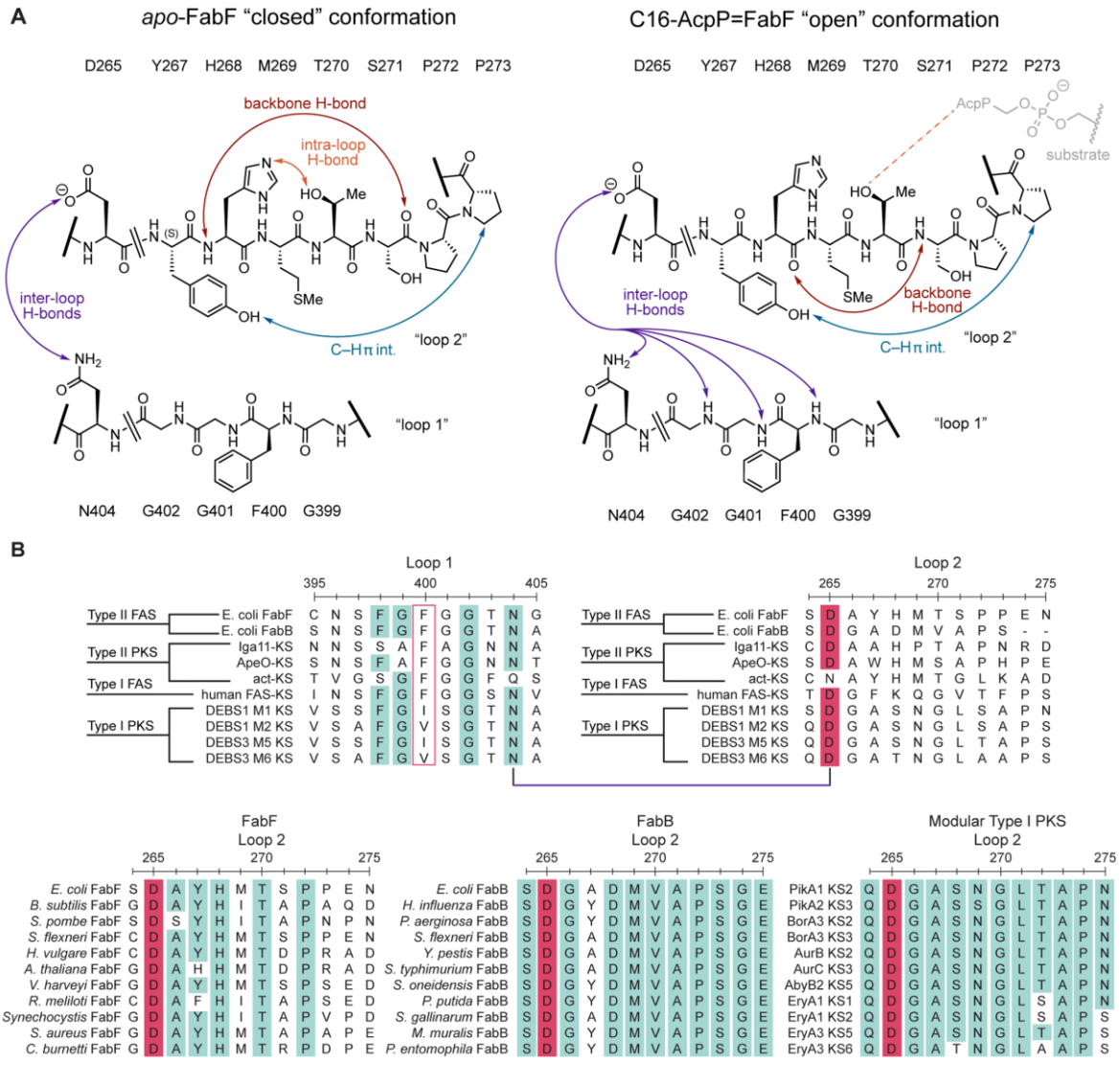


Figure 3.2: Gating loop interaction network and conservation. (A) Schematic for loop 1−loop 2 interactions in the closed (*apo*-FabF) and open (C16-AcpP=FabF) states. Interactions between specific residues are shown by arrows. FabF loop 2 sequence is above the loop 2 peptide schematic and the FabF loop 1 sequence is below the schematic. (B) Sequence alignment between FabF gating loops with the putative gating elements found in different classes of condensing enzymes. The loop 1 gating residue (Phe400) is outlined in a red box and the interaction between Asn404 and the conserved Asp265 is indicated by the connecting line segment. Loop 2 sequences from within three different classes of condensing enzymes (FabF, FabB, and modular type I PKS) are shown in separate alignments highlighting the conservation within, but not between, families of condensing enzymes. All residues above an identity threshold of 70% are highlighted.

but not between, different KS families (Figure 3.2B). In contrast, loop 1 shares similarities across other KS families, but the putative loop 1 gating residue notably diverges in cis-AT modular type I PKSs (TIPKS).[47-49] Instead of the conserved GFGG motif, TIPKSs often possess GVSG or GISG motifs (Figure 3.2B). TIPKS KSs utilize the same Cys-His-His catalytic triad and likely the same, or similar, catalytic mechanism as type I and type II FAS (Figure S3.2B). However, they are capable of accepting a variety of α-branched and β-branched acyl-substrates and often use bulkier α-branched extender units than FAS KSs, such as methylmalonyl-CoA and ethylmalonyl-CoA.[48,50] Therefore, the putative gating elements in PKSs may have diverged to facilitate catalysis with more structurally diverse substrates than FAS KSs.

Currently, the structure of the AcpP-bound gate-open conformation is available for FabF (PDB: 6OKG), while the structure of the AcpP-bound gate-closed conformation is available for FabB (PDB: 6OKC), making it difficult to directly compare the open and closed conformations in the same condensing enzyme family. Additionally, catalytic evidence that the proposed gating mechanism does indeed affect substrate turnover, and more specifically, the transacylation step, has not yet been demonstrated. In the work reported herein, we use mechanistic crosslinkers carrying different substrate mimetics to elucidate X-ray crystal structures of two additional acyl-AcpP=FabF (where "=" denotes a covalent crosslink) complexes, one of which is in the missing gate-closed conformation. Next, we employ two high-performance liquid chromatography (HPLC)-based kinetic assays to validate the importance of the KS gating machinery for the overall KS condensation reaction and transacylation half-reaction. We then explore the role of loop 2 using alanine-scanning

mutagenesis and demonstrate its conformational relevance to enzyme catalysis, providing additional evidence for an allosteric connection between AcpP-binding and gating events. Finally, we characterize a panel of FabF-TIPKS KS loop 1 chimeras, where FabF's GFGG motif is replaced by corresponding GVSG and GISG motifs found in the *cis*-AT modular TIPKS KS domains. Results from our study provide the missing AcpP-FabF gate-closed conformation structure as well as structural and biochemical support for the importance of gating in type II FAS KSs. These findings have broad implications for KS-directed metabolic engineering in FAS and provide additional insights into how these putative gating elements may operate in the evolutionarily related PKS condensing enzyme families.

Results and Discussion

Development and analysis of a C16:1-α-bromo-pantetheinamide crosslinking probe. In order to compare the open and closed states within a single family of condensing enzymes, we aimed to crystallize an acyl-AcpP=FabF complex in a catalytically relevant, gate-closed conformation. We first developed a pantetheinamide mechanistic crosslinker that we reasoned could trap the AcpP–FabF complex in a closed conformation. *E. coli* is the canonical paradigm for bacterial FAS and the two elongating KSs, FabF and FabB, perform different branchpoint reactions in FAS important to cell physiology. FabB, in concert with the dehydratase FabA, is responsible for elongation of *cis*-3-decenoyl-AcpP (*cis*-3-C10:1-AcpP) to *cis*-5-dodecenoyl-AcpP (*cis*-5-C12:1-AcpP) to establish the unsaturated fatty acid branch of FAS.[51-53] FabF is not able to perform this FabB specific extension in vivo, but previous studies regarding FabF's substrate specificity show that, unlike FabB, FabF is

responsible for elongation of *cis*-palmitoleoyl-ACP (C16:1) to *cis*-vaccenoyl-ACP (C18:1).[44,54] While cis-vaccenate is not essential for *E. coli* survival, this FabF product is responsible for *E. coli*'s rapid homeoviscous adaptive response to regulate membrane fluidity in response to reductions in temperature.[44,45,54-57] Therefore, we developed a synthetic route for a C16:1-α-bromo-pantetheinamide unsaturated crosslinking probe (Figure 3.3), designed to mimic FabF's privileged substrate, palmitoleoyl-ACP (Figure S3.3). As with other crosslinkers developed to trap ACP−PE interactions,[21,23-25,57-65] we replaced the labile thioester in the native substrate with an amide linkage, which is necessary to trap crosslinked ACP−PE complexes. However, recent NMR data from our group suggests that the identity of the linkage affects the overall structure of AcpP in solution.[66] These differences are mostly due to altered sequestration[67-69] of non-natively tethered substrates and likely minimally affect the conformation and structure in crosslinked complexes.

Using a one-pot chemoenzymatic method,[59] we loaded the C16:1-α-bromopantetheinamide crosslinking probe onto *apo*-AcpP to produce C16:1αBr-crypto-AcpP (C16:1-AcpP) (Figures 3.3 and S3.4). We then tested the crosslinking efficiency of C16:1-AcpP with wt FabF using a time course gel-based crosslinking assay. Similar to previously tested α-bromo crosslinkers, C16:1-AcpP crosslinks with FabF to completion in under 10 minutes across a range of pHs (Figure S3.5). Synthesis of different chain length, unsaturated crosslinkers may aid in elucidating structures of ACP−PE complexes with additional metabolic enzymes, such as *E. coli* FabB[51,53,70] or the FatA and FatB FAS thioesterases[71-73] from plant plastids, to elucidate their respective substrate specificities and recognition mechanisms.

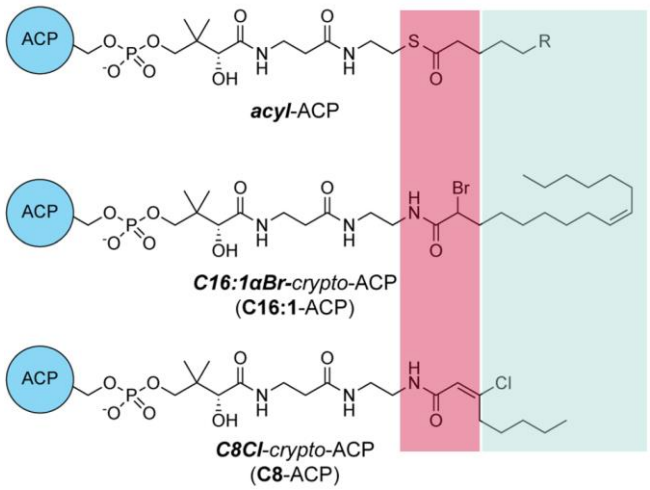


Figure 3.3: Crosslinking probes. Comparison of acyl-AcpP with crosslinker-loaded *crypto*-AcpPs utilized in this study to trap and crystalize AcpP=FabF complexes. An acyl-ACP substrate is shown to demonstrate the chemical structure of the natural substrate. The two crosslinking pantetheinamide probes have an amide instead of a thioester linking their acyl substrates as well as either an α-bromo or chloroacrylamide warhead. The region that has been modified to facilitate stable crosslinkers is shown in red, while the substrate mimetics are outlined in light green.

C16:1-AcpP=FabF captures distinct substrate conformation. Cross-seeding using C16-AcpP=FabF[24] crystals nucleated diffraction-quality crystals of C16:1-AcpP=FabF for X-ray data collection, structure determination, and refinement to a nominal resolution of 2.0 Å (Table S3.1, Figure S3.6). The complex crystallized in the same space group and unit cell as that of C16-AcpP=FabF (PDB: 6OKG), with an asymmetric unit containing a FabF monomer crosslinked to a single AcpP. Structural alignment with C16-AcpP=FabF (PDB: 6OKG) shows that these two complexes are nearly identical and overlay with a root-mean-square deviation (rmsd) of 0.124 Å. The gating loops in the C16:1-AcpP=FabF complex are in the open conformation, indicating that the gate-open conformation is favored by FabF in the presence of a crosslinked AcpP bearing long-chain α-bromo crosslinkers (Figures 3.4A−C and S3.7). This is not the case for the previously published C12- and C16-AcpP=FabB structures, which instead mimic the AcpP-bound transacylation reaction intermediate.[24] In these complexes, the carbonyl of the FA substrate analog is positioned in the oxyanion hole and the gating loops are found in the catalytically competent gate-closed conformation in a manner similar to the C12-FabF (PDB: 2GFY) and C12-FabB (PDB: 1EK4) acyl-enzyme intermediate complexes. Therefore, as in the C16-AcpP=FabF structure (PDB: 6OKG), the gating loops of C16:1-AcpP=FabF would need to return to the closed conformation to appropriately position the C16:1 FA substrate for transfer to the catalytic cysteine residue.

Closer analysis of the crosslinker indicates that the stereocenter of the α-carbon at the thioether crosslink is in the *R* configuration, while that of the previously solved structures was assigned to the *S* configuration (Figures 3.4A and S8). In the C16-AcpP=FabF complex,

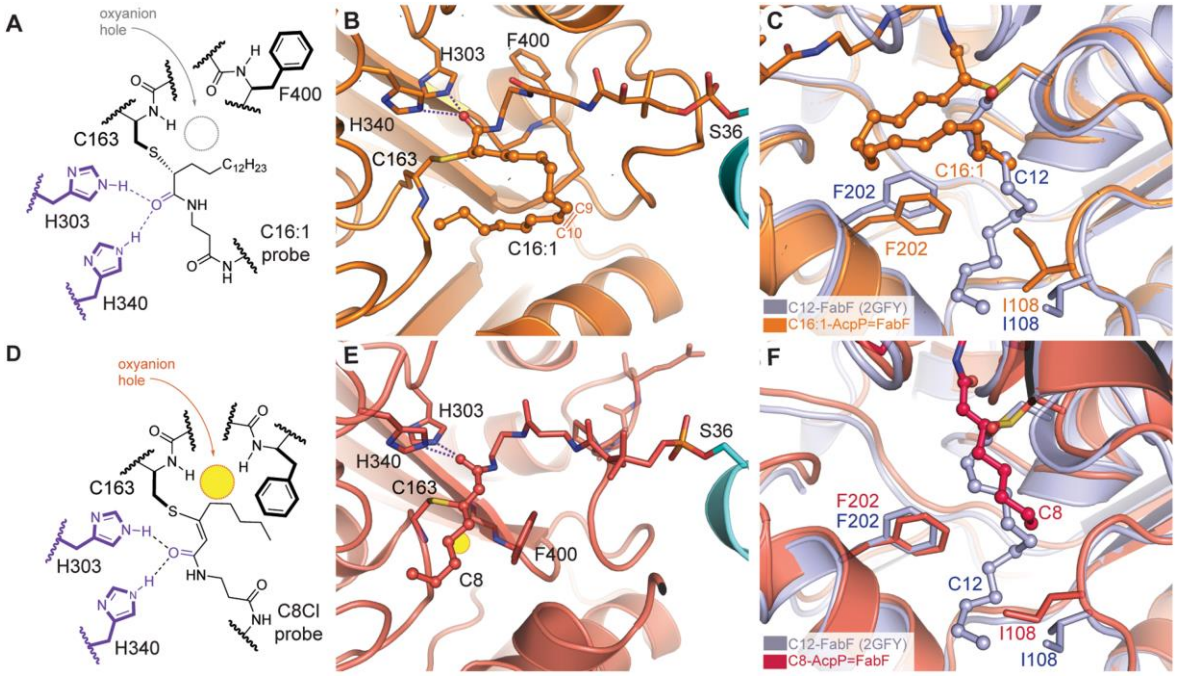


Figure 3.4: C16:1- and C8-AcpP=FabF active site organization. (A) 2D schematic of the C16:1-AcpP=FabF active site. The oxyanion hole is not formed and Phe400 is rotated away from the active site. (B) 3D rendering of the C16:1-AcpP=FabF active site and crosslinker. (C) Acyl-binding pocket of C16:1-AcpP=FabF compared to that of C12-FabF (PDB: 2GFY). (D) 2D schematic of the C8-AcpP=FabF active site. The oxyanion hole is formed and the gating loops are in the closed conformation. (E) 3D rendering of the C8-AcpP=FabF active site and crosslinker. (F) Acyl-binding pocket of C8-AcpP=FabF compared to C12-FabF (PDB: 2GFY). In panels C and F, structural alignments were performed over all atoms.

the C16 acyl chain extends away from the active site, whereas in C16:1-AcpP=FabF, the *cis*-double bond of the unsaturated probe is in a kinked conformation that redirects the polymethylene chain toward the acyl-binding pocket (Figures 3.4B, 3.4C and S3.7). The terminal methyl moiety of the acyl chain is positioned in front of the back gate, comprised of residues Ile108 and Phe202, of the acyl-binding pocket (Figure 3.4C).[35,74] When compared to the C12-bound acyl-enzyme intermediate structure of FabF (C12-FabF, PDB: 2GFY), the Ile108-Phe202 gate is in a closed conformation (Figure 3.4C). Despite the closed back gate, the tail of the kinked alkyl substrate is positioned at the entrance of the acyl-binding pocket. This may indicate that the Ile108-Phe202 back gate of FabF, which is replaced by Gly107-Phe201 in FabB, plays a role in substrate-intermediate processing. Previous studies have shown that altering the identity of this gating residue results in antibiotic resistance[75,76] and changes in KS chain length specificity in both bacterial and fungal FAS systems.[37,76]

The kinked C16:1 substrate favors a more compact conformation, which may be more readily accommodated and processed compared to the saturated C16 substrate. These observations, along with differences in the acyl-binding pocket and divergent loop 2 sequence found in the FabF condensing enzyme family (Figure 3.2B), may explain FabF's substrate preference for C16:1.[44,45] However, some FabF and FabB orthologs from related bacteria do not fit within the canonical *E. coli* FAS framework. For example, *Enterococcus faecalis*, *Lactococcus lactis*, and *Clostridium acetobutylicum* each contain only a single elongating FAS KS, which is from the FabF family, that is able to recover unsaturated FA auxotrophic growth defects in *E. coli* ΔfabB strains.[77-79] Additionally, the FabB from *Shewanella oneidensis* is capable of elongating C16:1 to C18:1 and its two endogenous fabF

genes are also capable of complementing ΔfabB growth defects.[80,81] Currently, no structures are available for these FabF orthologs, making it difficult to infer conclusions based on sequence alone. However, the primary sequence of the two *S. oneidensis* FabFs (NCBI Accession: WP_011074037 and WP_011074034), which are capable of complementing *E. coli* ΔfabB strains, notably diverge from that of FabF for loop 1 (AFGG vs GFGG), loop 2, and the gating residues regulating entrance to the FA-binding pocket. Instead of the canonical Ile108-Phe202 back gate, *S. oneidensis* FabF possesses a Thr108-Phe202 back gate at the equivalent position based on primary sequence analysis. Future structural, computational, and mutagenic work probing the Ile108-Phe202 back gate and/or loops 1 and 2 may provide insights into FabF's cryptic role in homeoviscous adaptation and its divergent function in other bacterial species.

C8-AcpP=FabF structure provides an ACP-bound gate-closed conformation. AcpP loaded with the *trans*-C8-chloroacrylate PPant probe, C8Cl-crypto-AcpP (C8-AcpP) (Figure 3.3), requires several hours to crosslink to FabF, but when loaded with a α-bromo crosslinkers or the smaller C3-chloroacrylate crosslinker that lacks an appended substrate analog, AcpP crosslinks in minutes.[24] This significant difference in crosslinking rates, and the propensity for α-bromo crosslinkers to trap FabF in the gate-open conformation, led us to hypothesize that crosslinking to *trans*-C8-chloroacrylate may require a transition to the gate-closed conformation to appropriately position the warhead for attack by the catalytic cysteine residue. Therefore, we obtained and collected data on diffraction quality crystals of C8Cl-crypto-AcpP=FabF (C8-AcpP=FabF), which upon structure elucidation and refinement, resulted in a 2.65 Å-resolution structure. The asymmetric unit of C8-AcpP=FabF

contains the FabF dimer crosslinked to two AcpPs. Analysis of the active sites shows reliable electron density for the PPant arm and the thiovinyl covalent crosslink between C_3 and Cys163 (Figure S3.9). The gating loops are positioned in the catalytically competent gate-closed conformation and an organized, but unoccupied, oxyanion hole is formed by the backbone amides of Cys163 and Phe400 (Figure 3.4D, 3.4E). The PPant moiety of the probe is anchored by polar contacts with Thr270, Ser271, Thr305, and Thr307. The carbonyl oxygen of the substrate mimetic coordinates to His303 and His340, with the latter histidine serving an essential role for the condensation reaction.[32,33] Similar probe-KS interactions were observed in the crystal structures of FabB−AcpP[23,24] and IgaKS-IgaACP,[62] indicating an evolutionarily conserved PPant-binding site.

The acylated C12-FabF structure (PDB: 2GFY) overlays with the KS domain from C8-AcpP=FabF with an rmsd of 0.314 Å (Figure 3.4F). The active sites of these complexes are in nearly identical conformations, including the gating loops, with the Phe400 gating residue rotated and translated away from Cys163 to form the presumptive malonyl-AcpP-binding pocket.[34,35,82] Interestingly, the chlorovinyl probe has the same carbon count as the putative condensation intermediate. The carbonyl group of the FA substrate mimetic interacts with His303 and His340, positioning the C_2 carbon in close proximity to the Cys163-bound thioester carbonyl carbon (Figure 3.5). With the gating loops closed, the oxyanion hole is organized, which, as seen when overlaid with C12-FabF, would be occupied by the carbonyl oxygen of the bound FA. Even with the altered Cys163 position in C8-AcpP=FabF compared to C12-FabF, angle measurements in the overlay place the C2 carbon within 30 degrees of the Burgi-Dunitz angle (ca. 105 °) (Figure 3.5). Given that the

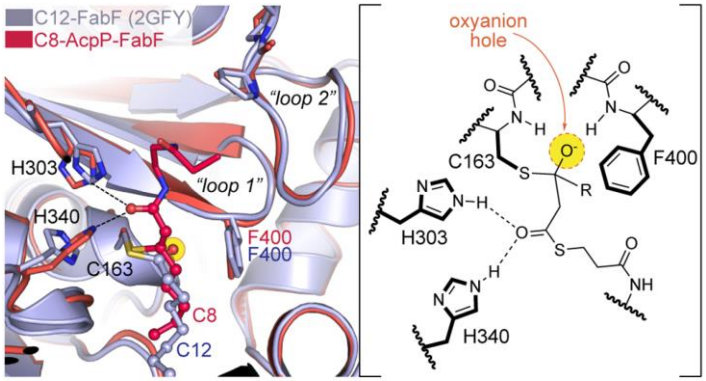


Figure 3.5: C8-AcpP=FabF mimics a condensation reaction intermediate. Superposition of C8Cl-AcpP=FabF and the C12-FabF acyl-enzyme intermediate (PDB: 2GFY) performed over all atoms, demonstrating that C8-AcpP=FabF mimics a condensation reaction intermediate between malonyl-AcpP and a hexanoate-bound FabF. The carbonyl of the C8Cl crosslinker is coordinated to the two catalytic histidines (dashed lines). The overlaid C12-FabF structure shows that the oxyanion hole (yellow circle) would be occupied by the carbonyl of the bound acyl substrate. Enolate attack on the bound FA would result in a tetrahedral intermediate stabilized by the oxyanion hole. A 2D schematic of the putative condensation half-reaction tetrahedral intermediate is provided for comparison.

observed probe behavior matches the current knowledge of the condensation mechanism, we propose that the C8 chlorovinyl probe mimics a condensation reaction intermediate between malonyl-AcpP and a FabF-bound hexanoic acid. When taken in consideration with our recently published structures and the previously solved structures of KS acyl-enzyme intermediates, C8-AcpP=FabF provides an additional snapshot into KS-mediated catalysis (Figure S3.10). However, future structural studies using stable malonyl-ACP or malonyl-CoA[83] mimetics would likely enhance our understanding of the decarboxylation reaction mechanism.

C8Cl-AcpP=FabF delineates a unique Loop 2 interaction network. The position of loop 2 in the gate-closed conformation suggests that movement of loop 2 is required to allow loop 1 to access the gate-open conformation. The interaction network that stabilizes the loop 2 closed conformation (*apo*-FabF PDB: 2GFW) is mediated by a $Ser271_{(O)C}$-$His268_{(H)N}$ main-chain hydrogen bond, a Thr270-His268 side-chain interaction, and a Pro273-Tyr267 C−H π interaction (Figures 3.2A and 3.6). Transition to the open conformation (C16-AcpP=FabF PDB: 6OKG) yields a distinct loop 2 interaction network, and interestingly, Thr270 swaps its interaction with His268 for Asp35 of the bound AcpP (Figures 3.2A and 3.6).

Analysis of the C8Cl-AcpP=FabF crosslinked complex reveals a loop 2 that, while in a similar conformation to that seen in *apo*-FabF and C12-FabF, nonetheless forms a distinct interaction network. The Pro273-Tyr267 C−H π interaction is maintained, but His268 instead interacts with Asp35 of AcpP and Thr270 and Ser271 form hydrogen-bonding interactions with the hydroxyl group of the PPant arm and the phosphoserine

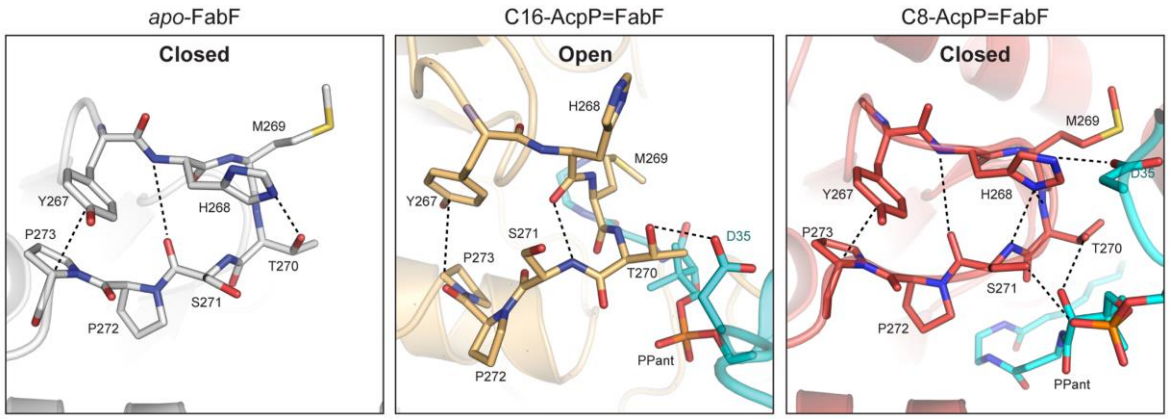


Figure 3.6: Loop 2 interaction networks: the apo-FabF, C16-AcpP-FabF, and C8-AcpP=FabF loop 2 interaction networks are distinct. Interactions between acyl-AcpP and loop 2 are present in the C8-AcpP=FabF (substrate bound, gate-closed) and C16-AcpP=FabF (substrate bound, gate-open) structures.

oxygen, respectively (Figure 3.6). Interestingly, similar loop 2 interaction networks are observed in the recently solved type II PKS ACP=KS-CLF complexes.[62,84] Collectively, these findings suggest that unique AcpP−loop 2 interactions are present in both the closed and open conformations, providing additional evidence for an allosteric role in substrate binding and cargo delivery during KS gating events.

Gating mechanism is important for the FabF condensation reaction. To confirm that FabF gating events are important for catalysis, we subjected the panel of FabF gating mutants used in our previous study[24] to an HPLC-based kinetic assay.[28] These mutants were designed to test gate function by blocking access to the gate open conformation (pocket-block), destabilizing the gate open conformation (destabilization), or limiting the flexibility of loop 1 (flex-reduction) (Figures S3.11-S3.14). Briefly, the assay monitors the rate of *holo*-AcpP formation in a reaction utilizing acyl-AcpP substrates as acyl-donors in the transacylation half-reaction while taking advantage of FabF's relaxed specificity for malonyl-CoA (extender unit) for the condensation half-reaction (Figures S3.15 and S3.16). Given that proper gate function would be required for substrate turnover, especially during the transacylation step, these assay parameters should reliably test gate function. Additionally, previous studies have utilized similar assay formats that take advantage of the relaxed specificity of KSs toward CoA-based substrates[28,36,85] or N-acetylcysteamine thioester substrates.[86-89] Using this assay format, we monitored the overall FabF condensation reaction for all previously tested gating mutants as well as an additional destabilization mutant, N404A, using both C6-AcpP and C12-AcpP substrates.

Using C6- or C12-AcpP, the apparent turnover rates of wt FabF in the tested conditions are 1.57 and 1.41 min^{-1}, respectively, which are comparable to previously reported values[28] (Table 3.1). Similar to a previous study by Zhang *et al.*, the F400A gate-removal mutant is 1-2% as active as wt FabF with both tested substrates.[32] We then evaluated the activity for the two pocket-block mutants, G310M and G310F, where mutation of glycine to a bulky hydrophobic residue fills the pocket occupied by Phe400 when in the open conformation (Figure S3.12). These mutants demonstrated reduced condensation rates for both C6 and C12-AcpP substrates with an overall activity roughly 50-fold lower than that of wt FabF.

The destabilization class of mutants was designed to disrupt the hydrogen-bonding network coordinated by Asp265, a highly conserved loop 2 residue that interacts with the backbone amides of loop 1 (in the gate-open conformation) and the side chain of Asn404 (Figures 3.2A and S3.13). The D265A mutant is 4.6% (0.07 min^{-1}) and 4.8% (0.06 min^{-1}) as active as wild-type with C6-AcpP and C12-AcpP, respectively. Interestingly, the more conservative D265N mutation is 13.8% (0.22 min^{-1}) as active as wild-type with C6-AcpP but only 6.4% (0.09 min^{-1}) as active as wild-type with C12-AcpP. To further validate the importance of this interaction network, we mutated the conserved Asn404 residue (Figure 3.2) to alanine and evaluated its activity with both substrates. As seen with D265A and D265N, the N404A mutant is a poor catalyst with rates 3.1% (0.05 min^{-1}) and 1.7% (0.03 min^{-1}) that of wt FabF with C6-AcpP and C12-AcpP, respectively.

The two flex-reduction mutants, G399A and G402A, were designed to inhibit gate function by limiting the conformational space available to the loop 1 GFGG motif (Figure

Table 3.1: Quantification of condensation and transacylation rates of FabF mutants.

FabF	C6-AcpP		C12-AcpP		transacylation	
	rate min^{-1}	% rel. to wt	rate min^{-1}	% rel. to wt	rate min^{-1}	% rel. to wt
wt	1.569 ± 0.104	100.00	1.414 ± 0.083	100.00	3.686 ± 0.200	100.00
F400A	0.024 ± 0.006	1.56	0.026 ± 0.004	1.85	0.165 ± 0.010	4.48
G310M	0.027 ± 0.002	1.72	0.026 ± 0.005	1.83	0.014 ± 0.026	3.73
G310F	0.041 ± 0.003	2.59	0.029 ± 0.002	2.08	-	-
D265A	0.072 ± 0.009	4.61	0.059 ± 0.008	4.18	0.018 ± 0.005	0.49
D265N	0.216 ± 0.008	13.76	0.091 ± 0.003	6.41	0.009 ± 0.003	0.24
N404A	0.049 ± 0.025	3.14	0.025 ± 0.002	1.74	0.049 ± 0.011	1.32
G399A	1.284 ± 0.098	81.83	1.622 ± 0.046	114.75	1.643 ± 0.006	44.58
G402A	0.301 ± 0.046	19.18	0.201 ± 0.013	14.18	0.039 ± 0.033	10.65
H268A	0.566 ± 0.039	36.10	0.475 ± 0.013	33.62	0.379 ± 0.014	10.29
M269A	1.771 ± 0.028	112.90	1.897 ± 0.047	134.18	-	-
T270A	0.209 ± 0.012	13.30	0.150 ± 0.013	10.60	0.021 ± 0.005	0.58
S271A	1.085 ± 0.012	69.15	1.153 ± 0.059	81.57	-	-
Y267A/P273A	1.240 ± 0.048	79.01	1.139 ± 0.074	80.53	-	-
loop swap	0.010 ± 0.001	0.64	0.008 ± 0.001	0.59	0.003[a] ± 0.001	0.08

[a]Estimated rate from total product formation at 50 min. [b]- indicates that the assay was not performed.

S3.14). The G399A mutant is as fast as wt FabF, while rates for the G402A mutant are 19.2% (0.301 min^{-1}) and 11.0% (0.201 min^{-1}) as active with C6- and C12-AcpP, respectively. These results are in line with our previous study as only G402A showed a reduction in crosslinking activity, while G399A was as active as wt FabF. Taken together, our results indicate that all FabF gating mutants, with the exception of G399A, significantly decrease FabF-catalyzed condensation rates.

Transacylation rates of FabF gating mutants. Results from our assays demonstrate that FabF gating mutants affect the overall condensation reaction. We next attempted to verify that the FabF gating mechanism is important for the transacylation half-reaction. To address this question, we modified our HPLC assay to monitor the transfer of the C12 FA from lauroyl-CoA to *holo*-ACP, effectively decoupling the transacylation and condensation reactions. Our proposed gating mechanism posits that gating is required for acyl substrates to access the FA-binding pocket and catalytic cysteine. Therefore, transfer of lauric acid from lauroyl-CoA to *holo*-AcpP would also require similar active site transformations. Furthermore, acyl-CoA to holo-ACP transferase assays have been used previously to assess the role of FabF active site residues in the transacylation step[32] and to evaluate substrate specificity and cooperativity in the murine FAS.[36] Using this assay design, we were able to reliably determine transacylation rates for wt FabF and FabF gating mutants by monitoring the formation of C12-AcpP (Table 3.1, Figures S3.15 – S3.17).

Type I and II FAS KSs both possess the conserved Phe400 gating residue and only accept acyl-ACP substrates without β-carbon modifications (i.e., β-hydroxy-acyl-ACPs). The Phe400 gating residue was recently proposed to serve as a β-carbon sensor in the murine

type I FAS KS domain.[36] This model suggests that Phe400 is responsible for ensuring partially processed FAS intermediates are not accepted. However, previous work on type II FAS KS domains suggests that Phe400 minimally affects transacylation and primarily regulates the condensation step by rotating away from the catalytic cysteine to form the malonyl-ACP-binding pocket after acylation of the catalytic cysteine.[32,34,34,74,90] Despite these previous reports, we determined a transacylation rate 4.5% that of wt for the F400A gate-deletion mutant (Table 3.1). Therefore, our results demonstrate the importance Phe400 for the transacylation half-reaction in addition to its well-established function in the condensation step.

The transacylation rate of the pocket block mutant, G310M, shows a similar reduction in transacylation activity as F400A, demonstrating that either blocking gate function or deleting the central gating residue decreases acyl-transfer rates. Flex reduction mutants, G399A and G402A, show a 2- and 10-fold drop in transacylation activity, respectively, suggesting that G399A may have a minor role during the transacylation step despite not reducing the overall condensation reaction rate in the condensation assay. Interestingly, the destabilization mutants, D265A and D265N, which are part of loop 2 and distal to the KS active site, have the largest decrease in activity with rates 200- and 300-fold lower than wt FabF, respectively (Table 3.1). The additional destabilization mutant, N404A, also showed a strong reduction in the transacylation rate, further indicating that disruption of this interaction network is detrimental to function. It is interesting that the rates of the destabilization class of mutants are more strongly reduced in the transacylation reaction than in the overall condensation reaction. However, differences in the assay format and the use

of non-native substrates affect absolute rates; therefore, direct comparisons between the condensation and transacylation rates should be avoided. Regardless, the proposed function of these residues is congruent with the reported results for both assays, indicating that the complex interaction network mediated by Asp265 is critical for gate function and the transacylation half-reaction.

Loops 1 and 2 were previously proposed to move in a coordinated manner to facilitate substrate processing during the transacylation half-reaction.[24] Results from our condensation reaction assay demonstrate that blocking gate function significantly reduces overall KS-mediated substrate turnover. However, an argument could be made that these mutations alter malonyl-CoA binding or positioning of the catalytic histidines, thereby only disrupting the condensation half-reaction. Additionally, analysis is further complicated as Phe400 of loop 1 coordinates the oxyanion hole that stabilizes chemically distinct tetrahedral intermediates for both half-reactions, and as discussed above, previous work has shown the importance of Phe400 for the condensation half-reaction. By implementing an assay that eliminates the condensation half- reaction, we have demonstrated that the observed reductions in the overall KS condensation reaction for FabF gating mutants are correlated with reductions in transacylation rates, thus demonstrating the importance of gating events for the transacylation half-reaction.

Importance of loop 2 residues for gate function. To quantitatively measure the importance of loop 2 for gate function, we subjected it to alanine-scanning mutagenesis and tested all resulting soluble protein constructs using our condensation and transacylation assays (Figure S3.11). In addition, we tested a FabF/FabB loop swap variant, where the loop

2 of FabB was swapped out for FabF's loop 2. The Y267A, P272A, and P273A mutants were insoluble, but we determined rates for alanine mutants of all remaining loop 2 residues as well as a Y267A/P273A double mutant. The overall KS condensation assay results indicate that only the conserved His268 and Thr270 (Figure 3.2B) are important for FabF activity. The H268A variant has a 3-fold reduction in activity, while the T270A mutation results in a 10-fold reduction. All other constructs are as active as wt FabF with the exception of the FabF/FabB loop swap variant, which has greater than a 150-fold drop in activity for both substrates (Table 3.1).

To determine if these mutations have any effect on the transacylation half-reaction, we monitored transacylation rates for H268A, T270A, and the FabF/FabB loop swap variants. The H268A variant has a transacylation rate 10.3% that of wt FabF while the T270A variant has a rate 0.6% that of wild-type. Interestingly, the transacylation rate for the FabF−FabB loop swap variant is too slow to be reliably measured (Table 3.1).

Results from the KS condensation and transacylation assays demonstrate that the conserved His268 and Thr270 residues, both of which interact with the bound acyl-AcpP in the closed and open states, respectively (Figure 3.6), are important for FabF activity. Mutation of these residues to alanine decreases KS transacylation rates, indicating that they likely play a role in gate function and that acyl-AcpP binding might participate in triggering or stabilizing gate transitions *via* loop 2 interactions, as suggested by previous molecular dynamics (MD) simulations.[24] Interestingly, replacing FabF's loop 2 with that of FabB results in an enzyme variant that is nearly inactive, providing additional evidence that loops 1 and 2 function in a coordinated manner, and may not be readily transferrable between

different KS families. Therefore, additional elements, which may include ACP-KS PPIs, likely coevolve with loop 2 to fine-tune gate function. A dual loop gating mechanism suggests a dynamic and complex catalytic process, which is potentially coupled to ACP binding, chain-flipping of acyl cargo into the KS active site, and KS-substrate recognition. Therefore, future studies analyzing the putative gating loops in FabB, type I FAS, and PKS KS domains will be essential to conclusively address these latter hypotheses.

Evaluating type II FAS-TIPKS loop 1 chimeras. Unlike FAS KSs, TIPKS KS domains readily accept α- and β-modified acyl-ACP substrates as well as a variety of α-branched extender units (Figures S3.1 and S3.2).[17,91] Comparisons between KS domains from TIPKS and type II FAS indicate that TIPKS KSs possess less-bulky GISG and GVSG loop 1 motifs (Figure 3.2B) that could provide additional space to accommodate β-carbon modifications or α-branched extender units. To evaluate this hypothesis, we modeled PKS-like acyl substrates into C12-FabF (PDB: 2GFY) or PKS-like extender units (2*S*-methylmalonyl-AcpP) into C8-AcpP=FabF. In both instances, potential clashes are identified with the native Phe400 FAS gating residue. However, mutating this gating residue into a PKS gating residue, such as isoleucine, could provide additional space for these PKS-like substrates (Figure S3.18). Therefore, the identity of the loop 1 gating residue may be tuned to enforce substrate specificity during the transacylation step by acting as a β-carbon sensor,[36] but it may also exert specificity during the condensation step by providing additional space for bulkier α-branched extender units.

To determine how TIPKS gating elements would alter FabF activity and substrate specificity, we generated chimeric FabFs containing loop 1 elements from TIPKS, where the

Phe400 gating residue was replaced with either isoleucine or valine. Additionally, we prepared two complete FabF-TIPKS KS loop 1 chimeras by mutating the FabF's GFGG motif to either GVSG (FabF-GVSG) or GISG (FabF-GISG). We first evaluated the overall condensation rates of our panel of loop 1 chimeras in assays containing either 0.25 or 1.0 mM malonyl-CoA (Table 3.2). The condensation rates for all mutants are significantly reduced compared to wt FabF for both concentrations of malonyl-CoA. The F400I and F400V mutations turn over substrate with rates 4.4 and 5.5% that of wt FabF, respectively, at 0.25 mM malonyl-CoA and 4.6 and 6.9% that of wt FabF, respectively, at 1.0 mM malonyl-CoA. These mutants are roughly 3- and 10-fold more active than the gate-deletion mutant, F400A, at 0.25 or 1.0 mM malonyl-CoA, respectively, indicating that reintroducing a gating residue recovers some activity. Condensation rates for the FabF-GVSG and FabF-GISG constructs are more significantly reduced with activities 1-2% that of wt FabF for both tested concentrations of malonyl-CoA.

To ascertain if the measured reductions in activity were related to extender unity identity, we tested our FabF loop 1 chimeras with assays containing 1 mM methylmalonyl-CoA. We measured an apparent turnover number of 0.28 min^{-1} for wt FabF, an 18-fold reduction compared to assays with 1 mM malonyl-CoA (Table 3.2). Extending our analysis to the chimeric FabF constructs, we again found the condensation rates for all variants are decreased compared to wt FabF. Interestingly, the relative-to-wt reduction in activity for these mutants is less pronounced than in assays using malonyl-CoA; however, this may be due to the decreased activity of wt FabF with methylmalonyl-CoA rather than any gain of function by the FabF loop 1 chimeras. The F400V and F400I variants are 14 and 25% as

Table 3.2: Quantification of condensation and transacylation rates of FabF gating mutants with malonyl-CoA (mCoA) or methylmalonyl-CoA (mmCoA) Substrates.

	0.25 mM mCoA		1 mM mCoA		1 mM mmCoA		transacylation	
FabF	rate min^{-1}	% rel. wt	rate min^{-1}	% rel. wt	rate min^{-1}	% rel. wt	rate min^{-1}	% rel. wt
wt	1.569 ± 0.104	100.00	5.011 ± 0.142	100.00	0.281 ± 0.056	100.00	3.686 ± 0.200	100.00
F400A	0.024 ± 0.006	1.56	0.027 ± 0.002	0.54	0.020 ± 0.009	7.14	0.165 ± 0.010	4.48
F400I	0.069 ± 0.003	4.44	0.232 ± 0.033	4.63	0.040 ± 0.002	14.33	0.371 ± 0.018	10.06
F400V	0.087 ± 0.013	5.53	0.346 ± 0.103	6.90	0.070 ± 0.034	25.03	0.400 ± 0.084	10.87
GISG	0.037 ± 0.009	2.34	0.084 ± 0.018	1.67	0.037 ± 0.007	13.09	0.016 ± 0.001	0.44
GVSG	0.023 ± 0.004	1.50	0.045 ± 0.005	0.90	0.012[a] ± 0.011	4.27	0.009 ± 0.002	0.25

[a]Estimated rate from total product formation at 40 min.

active as wt FabF, respectively, and the GISG mutant is 13% as active. We were unable to reliably detect activity for the GVSG mutant under the tested conditions.

The overall reaction rates of FabF gating chimeras are slower than wt FabF with both malonyl-CoA and methylmalonyl-CoA, but we cannot identify which half-reaction is more impacted by these mutations. To determine if these mutations would have any effect on the transacylation step, we evaluated the transacylation rates of our panel of FabF loop 1 chimeras (Table 3.2). The F400V and F400I variants have transacylation rates 10% that of wt FabF, again demonstrating that Phe400 is important for the transacylation half-reaction. Although, it is worth noting that the F400V and F400I constructs have transacylation rates 2.5-fold higher than F400A. The transacylation rates of the FabF-GVSG and FabF-GISG variants are found to be more significantly reduced, with activities 0.3% and 0.4% that of wt FabF, respectively. The more than 200-fold drop in transacylation activity for FabF-GVSG and FabF-GISG chimeras is likely correlated with their low activities in condensation assays with either malonyl- or methylmalonyl-CoA.

In summary, replacement of Phe400 with either valine or isoleucine results in an enzyme more active than the F400A gate-removal mutant, but still 20-fold less active than wild-type with malonyl-CoA substrates. Full replacement of FabF's loop 1 with similar sequences to those found in TIPKS KS domains results in even larger reductions in activity. These results may indicate that loop 1 elements from other KS families are not readily compatible with the gating machinery of FabF. Indeed, the FabF loop 2 is divergent from that of TIPKS KSs and may not be tuned to coordinate gating transitions with GVSG and

GISG loop 1 motifs (Figure 3.2B). This analysis is in agreement with our results from the FabF−FabB loop swap variant, suggesting any mutations to loop 1 may need to be evaluated in relationship to loop 2. Additionally, other residues in and near the acyl-binding pocket of the KS might provide favorable contacts to coordinate and position the substrate for proper gate closure and transfer. These residues are hydrophobic in FabF but are often replaced by more polar residues in type I and type II PKS KS domains. Finally, due to the pleiotropic effects of modifying KS gating elements, it seems difficult to determine how swapping loop 1 residues between different KS families, or even just the Phe400 gating residue, alters extender unit substrate specificity in the condensation step. Thus far, we have not found any instances of a TIPKS KS domain utilizing a phenylalanine gating residue, which instead is most often valine, isoleucine, leucine, or methionine. Additionally, there does not appear to be a direct relationship between the extender unit loaded by the TIPKS AT domain and loop 1 gating elements in the KS domain. The absolute conservation of Phe400 in FAS, and its general absence in TIPKS, may be linked to the FASs' role as a conserved, primary metabolic pathway that must efficiently and specifically synthesize FAs for cellular needs. PKSs, on the other hand, are part of specialized metabolism and have different evolutionary restrictions, suggesting that the gating elements of TIPKSs may have evolved to be more permissive to allow processing of diverse substrates in both the transacylation and condensation step. Duplications of these modules could then ensure at least some level of activity with the complex substrates processed in these assembly lines before evolutionary enhancement and fine tuning of turnover rates. Additional studies investigating the putative

gating elements in related KS domains will be essential to addressing these unanswered questions.

Conclusions

To better understand the KS reaction mechanism, we sought to further characterize our recently proposed KS gating mechanism with additional structural and catalytic investigations. The data reported herein are congruent with a model wherein acyl-ACP binding and substrate transfer are mediated by the correlated movement of two KS active site loops, loops 1 and 2. We provide biochemical evidence that proper gate function is required for KS-mediated substrate turnover, and we further demonstrate its importance for the transacylation half-reaction. Importantly, our structural and biochemical data, along with previous MD simulation studies,[24] indicate a potential role for loop 2 in regulating loop 1 gating transitions. We observe distinct interaction networks for FabF's loop 2 in both the open and closed states, which include contacts to the bound acyl-AcpP molecule, suggesting an allosteric role connecting acyl-AcpP binding and substrate delivery to gate function. Loop 2 is conserved within, but not between KS families, and our kinetic data indicates that neither loop 2 nor loop 1 is readily transferrable between evolutionarily divergent KS families. Therefore, loop 2 may coevolve with loop 1, or other KS elements such as ACP-mediated PPIs or KS−substrate interactions, to fine-tune KS gate function.

Despite the central role elongating KSs play in FAS- and PKS-dependent metabolic pathways, there are relatively few successful KS engineering attempts reported in the

literature.[37,41,76,88,89] This challenge likely correlates with deficiencies in our fundamental understanding of the KS reaction mechanism. The manner in which KSs determine substrate preferences, for either the transacylation or condensation step, remains poorly understood. Using a combination of chemical biology, structural biology, and biochemistry, we provide compelling evidence in support of the KS gating mechanism and an additional structural framework governing its role in catalysis. Given the conserved nature of these gating elements in both FAS and PKS, the series of studies reported herein provide part of the necessary catalytic foundation to design and carry out future KS mechanistic studies. Moreover, this work also accelerates the development of strategies to engineer metabolic pathways in a predictable manner.

Acknowledgement

Chapter 3, entitled Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases, in full, is a reprint of the material as it appears in: Mindrebo JT*, Chen A*, Kim WE, Re RN, Davis TD, Noel JP, Burkart MD. "Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases", ACS Catalysis, vol. 11, 2021. The dissertation author is the primary co-author of this manuscript along with Dr. Jefferey T. Mindrebo.

Supplementary Information

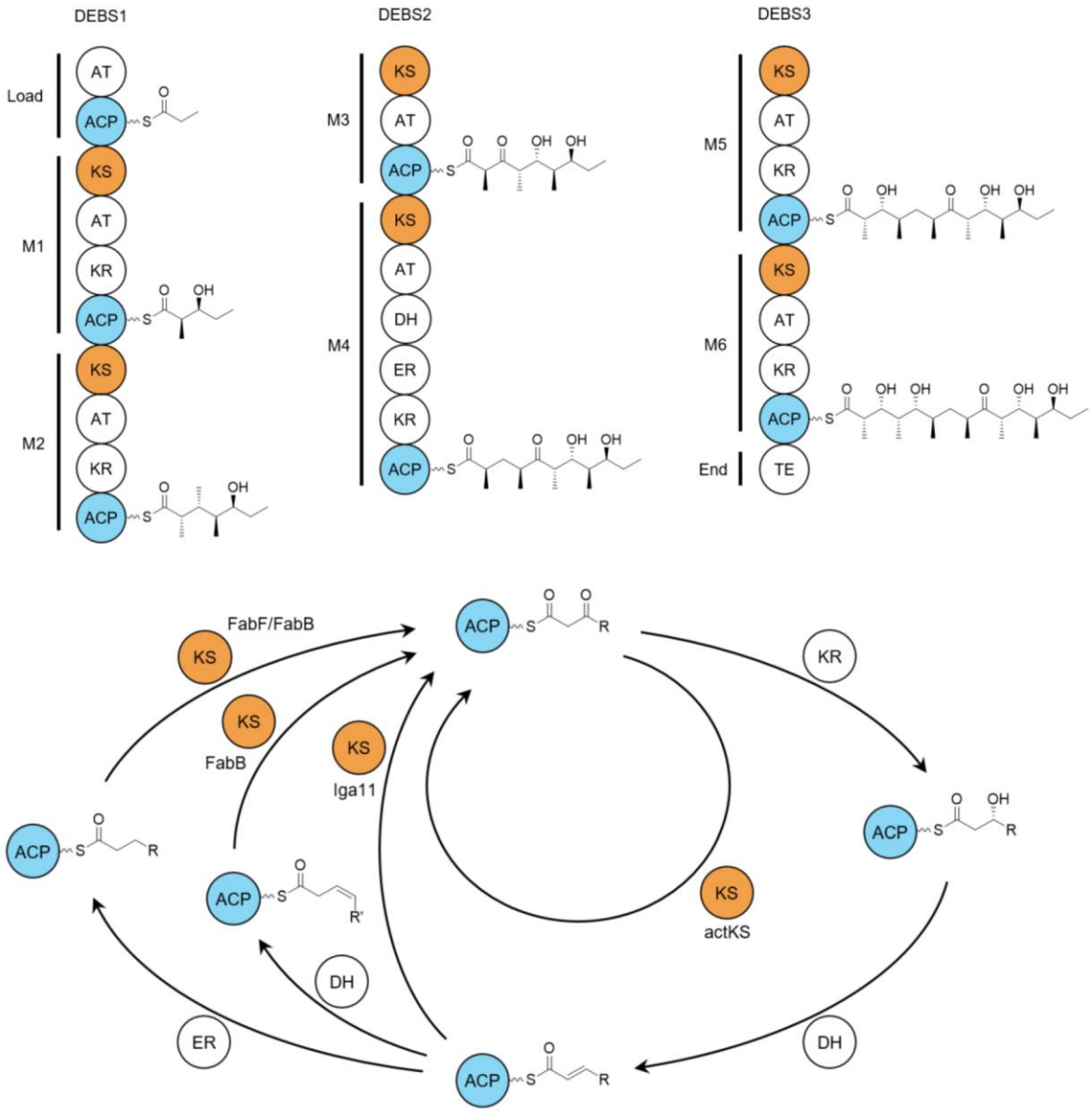


Figure S3.1: Reaction schemes of different modular synthases. (A) Organization of a type I PKS, deoxyerythronolide B synthase (DEBS). The KS domain elongates the substrate using extender unit incorporated by the AT domain from 2-(*S*)-methylmalonyl CoA. The growing intermediates are shown tethered to the ACP domains. M1-6, module 1-6; AT, acyltransferase; ACP, acyl carrier protein; KS, ketosynthase; KR, β-ketoreductase; DH, dehydratase; ER, enoylreductase; TE, thioesterase. (B) Simplified elongation cycle of various type II iterative synthases. Examples are given to different KSs: actKS and Iga11 belong to type II PKS while FabF and FabB belong to type II FAS.

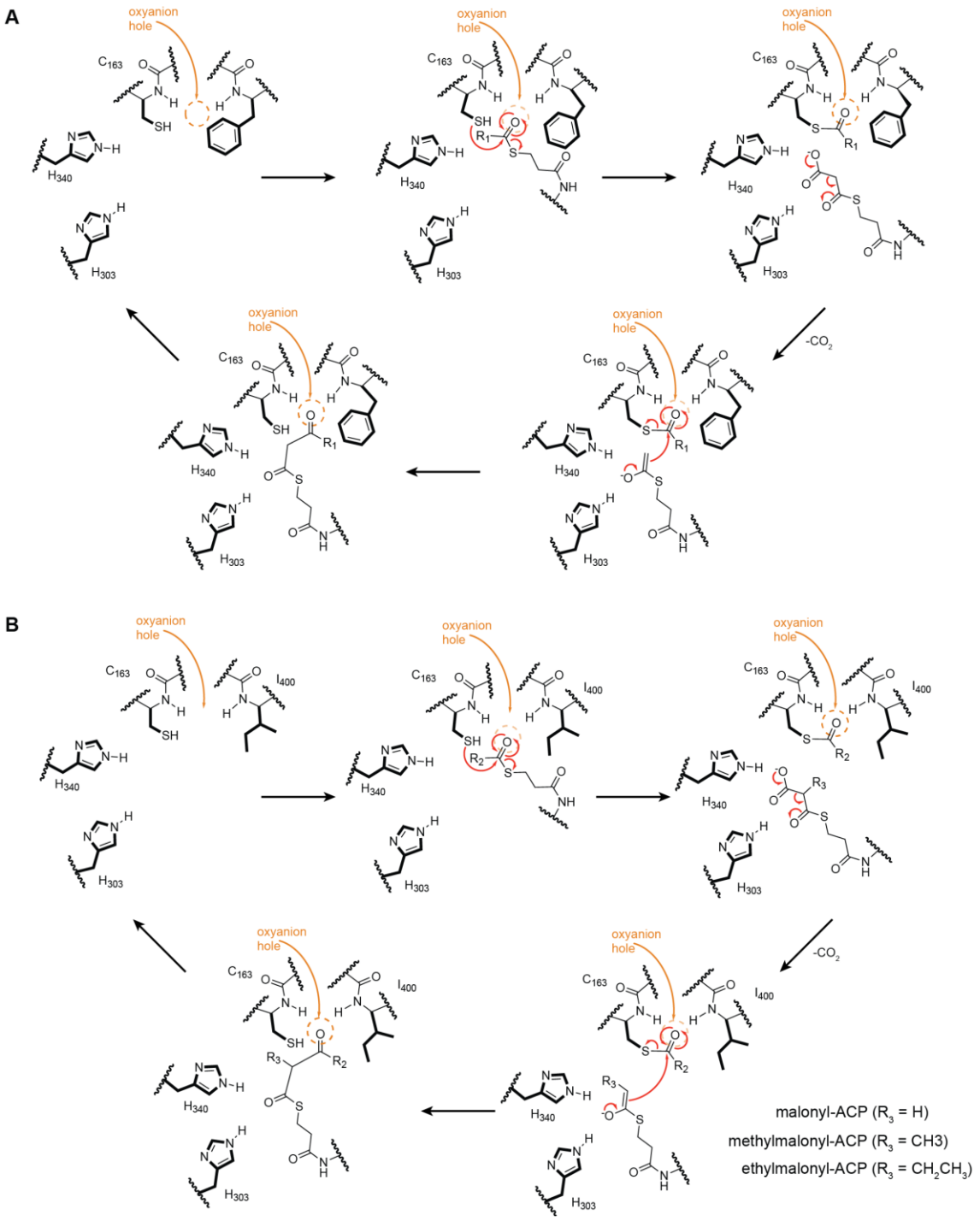


Figure S3.2: Catalytic mechanism of type II FAS KS and type I PKS KS. (A) General Type II FAS KS catalytic mechanism where R_1 represents saturated FAS substrate. (B) General Type I PKS catalytic mechains where R_2 represent PKS acyl substrate. PKS extension can occur using either malonyl-ACP, methylmalonyl-ACP, or ethylmalonyl-ACP.

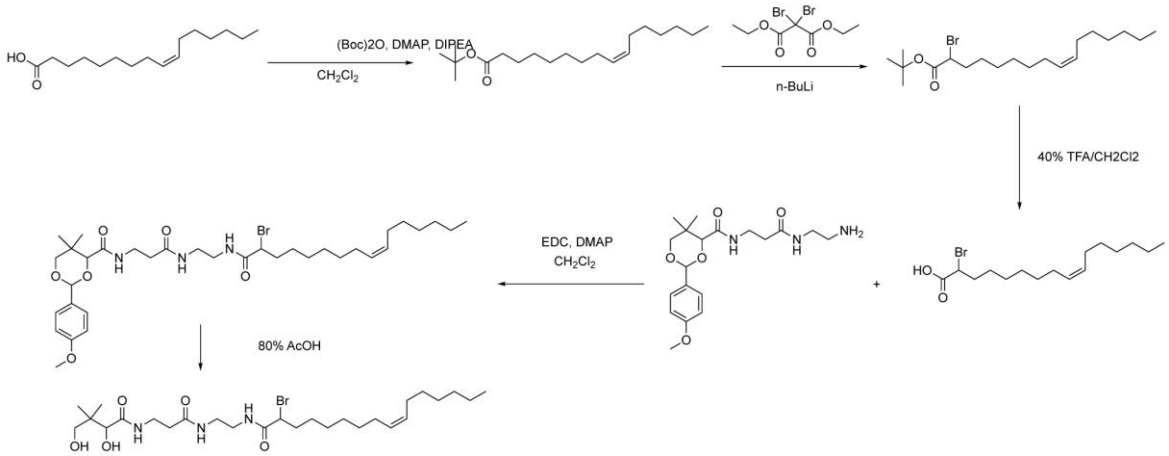


Figure S3.3: Synthetic route for the synthesis of the C16:1-α-bromo-pantetheinamide probe.

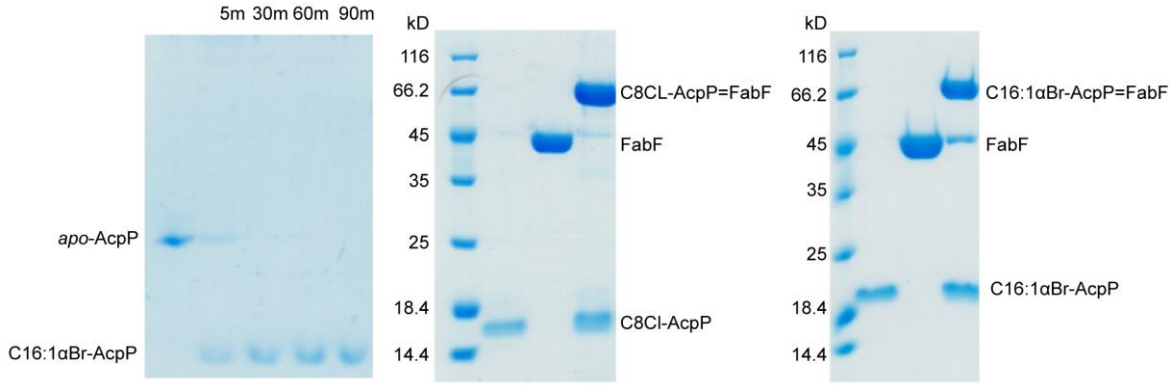


Figure S3.4: One-pot loading of C16:1-α-bromo pantetheineamide probe and crosslinking gels. A time course one-pot of apo-AcpP with the C16:1-αBr-pantetheinamide probe (left). Scaled up crosslinking for sample preparation of C8Cl-AcpP=FabF for crystal trials (middle). Scaled up crosslinking for sample preparation of C16:1αBr-AcpP=FabF for crystal trials (right).

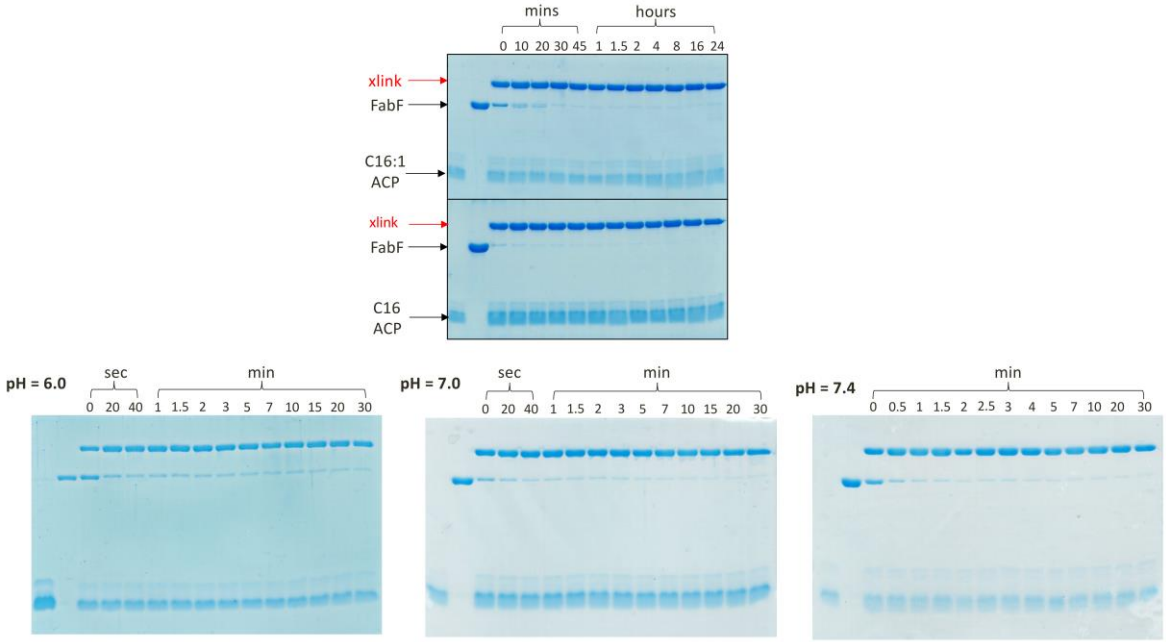


Figure S3.5: Time course crosslinking gels of C16:1-*crypto*-AcpP with FabF. The two gels on the top represent time-course crosslinking gels of C16:1- and C16-*crypto*-AcpP with FabF. The bottom three gels represent time-course crosslinking analysis of C16:1-crypto-AcpP with FabF at three different pHs (6.0, 7.0, and 7.4).

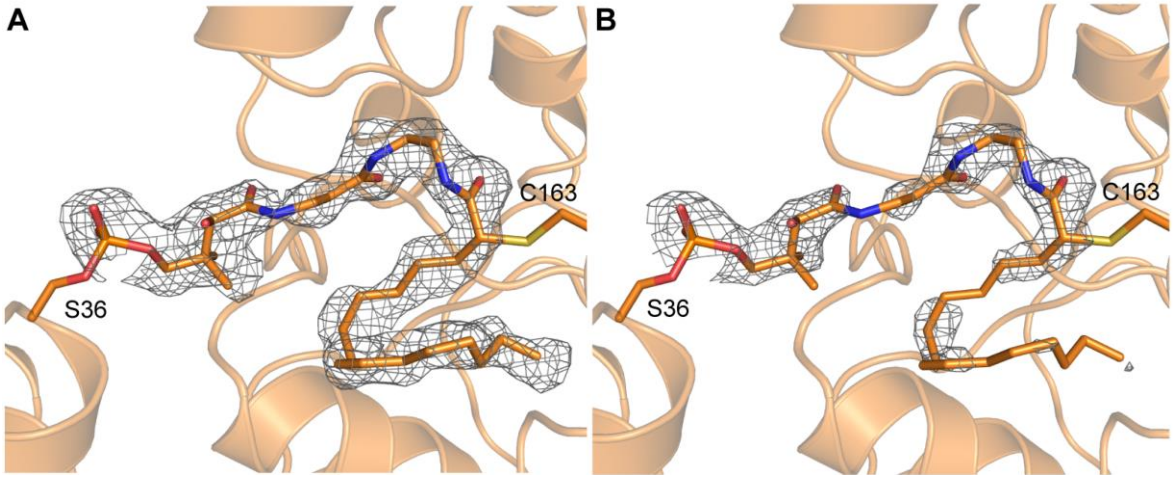


Figure S3.6: C16:1-AcpP=FabF substrate omit maps. (A) Fo-Fc polder omit map calculated by omitting PPant crosslinker and bulk solvent from the region defined by a 5 Å solvent exclusion radius. The Fo-Fc polder map is contoured at 3.0 σ and rendered using a carve radius of 1.6 Å. (B) Standard Fo-Fc omit map (no bulk solvent exclusion from omitted region) contoured at 3.0 σ and rendered using a carve radius of 1.6 Å.

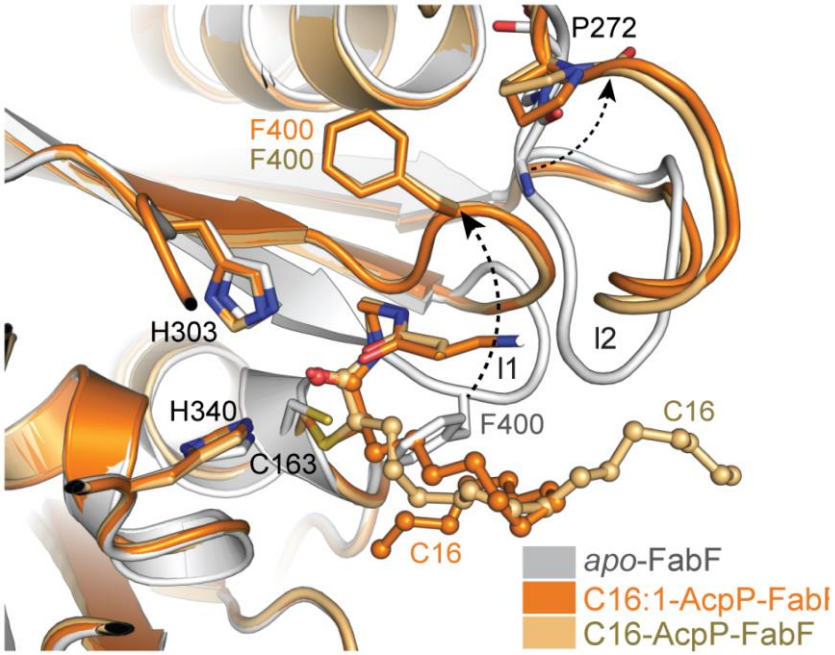


Figure S3.7: Overlay of *apo*-FabF (PDB: 2GFW), C16-AcpP=FabF (PDB: 6OKG), and C16:1-AcpP=FabF. The gating loops (l1 and l2) are found in the open conformation for both C16-AcpP=FabF and C16:1-AcpP=FabF. The arrows show the proposed trajectory of loop movement from the closed state (*apo*-FabF) to the open state. Apo-FabF, C16-AcpP=FabF, and C16:1-AcpP=FabF were structurally aligned over all residues (1 to 412) using all atoms. The fatty acid substrate mimetic of the C16 and C16:1 crosslinking probes are shown in ball and stick.

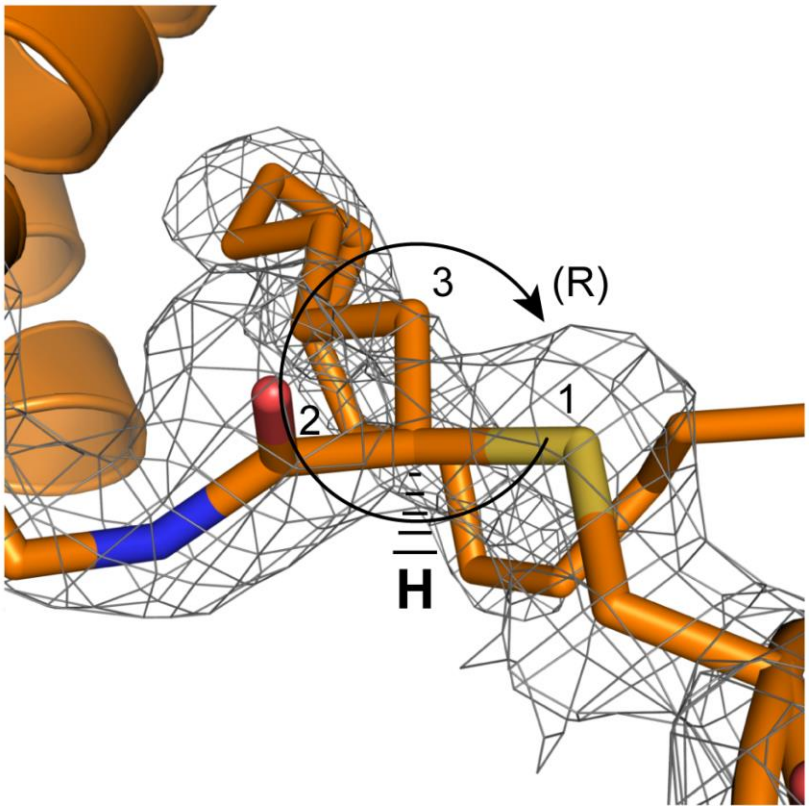


Figure S3.8: Stereochemistry at crosslinker α-carbon. The 2Fo-Fc electron density map show density for the thiovynyl crosslink site. The map is contoured at 1.0 σ and rendered using a carve radius of 1.6 Å. A hydrogen atom is added below to demonstrate that the substituent priorities (Cahn-Ingold-Prelog), based on our refinements, are arranged in the *R* configuration.

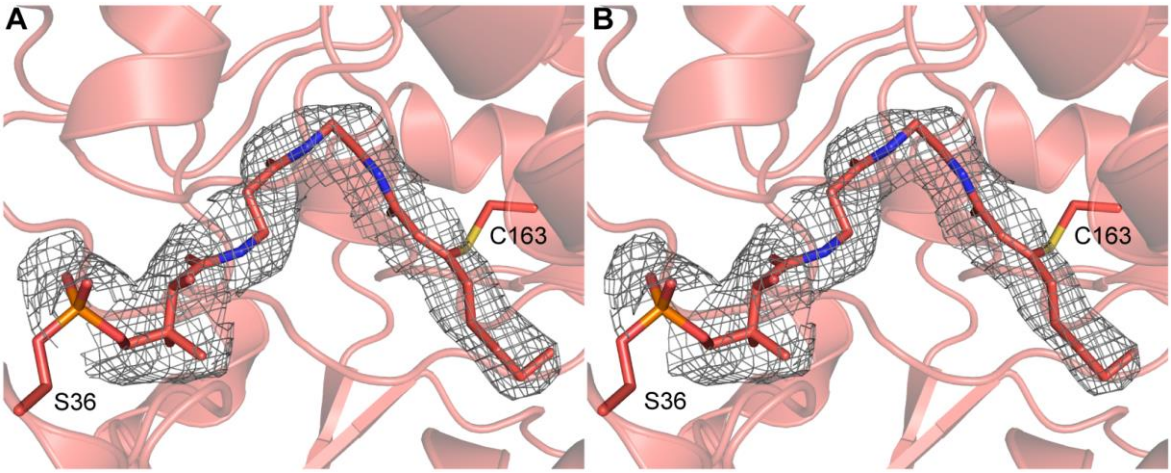


Figure S3.9: C8-AcpP=FabF substrate omit maps. (A) Fo-Fc polder omit map calculated by omitting PPant crosslinker and bulk solvent from the region defined by a 5 Å solvent exclusion radius. The Fo-Fc polder map is contoured at 3.0 σ and rendered using a carve radius of 1.6 Å. (B) Standard Fo-Fc omit map (no bulk solvent exclusion from omitted region) contoured at 3.0 σ and rendered using a carve radius of 1.6 Å.

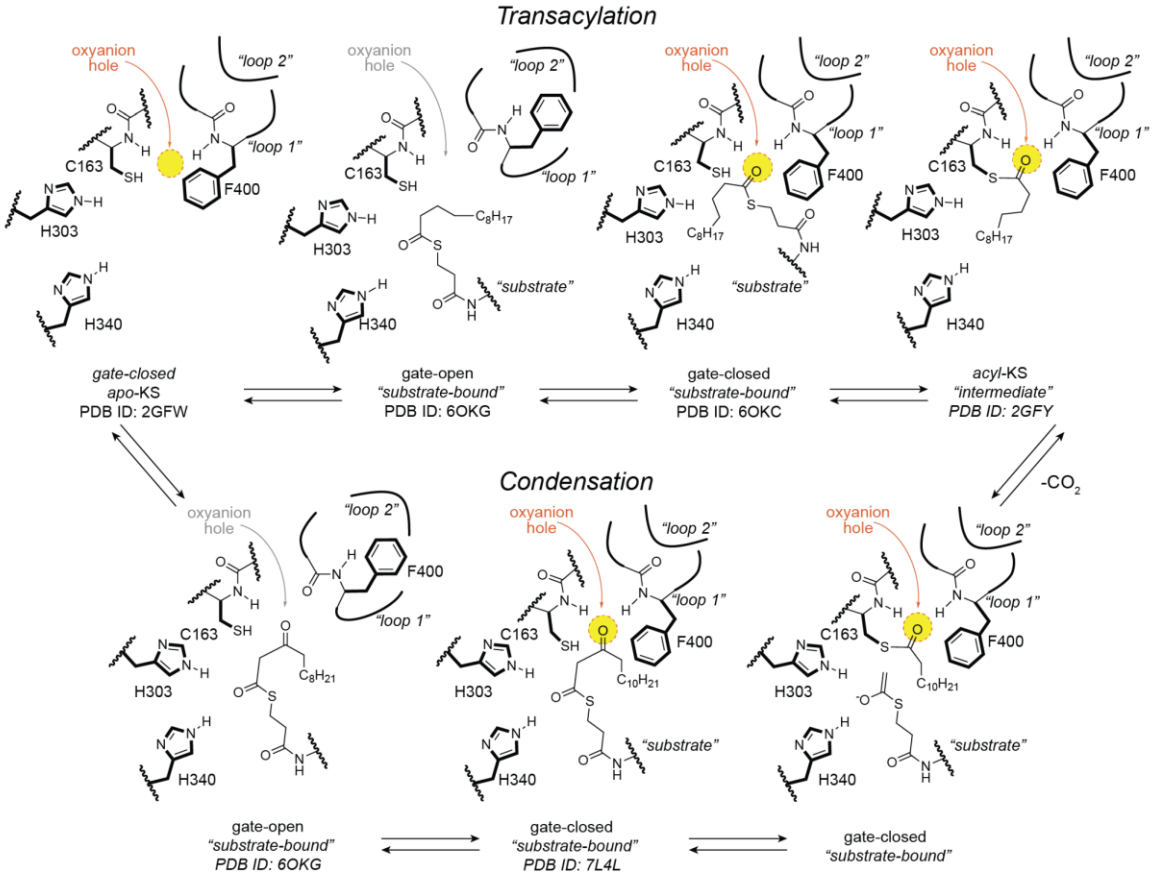


Figure S3.10: Trapped FAS KS catalytic intermediates throughout the KS catalytic cycle. Figure provides an overview of all currently available trapped intermediates of type II FAS KSs. All of the trapped states are derived from FabF structures, except for the third panel in the transacylation reaction (PDB ID: 6OKC), which is derived from the C12-AcpP=FabB structure.

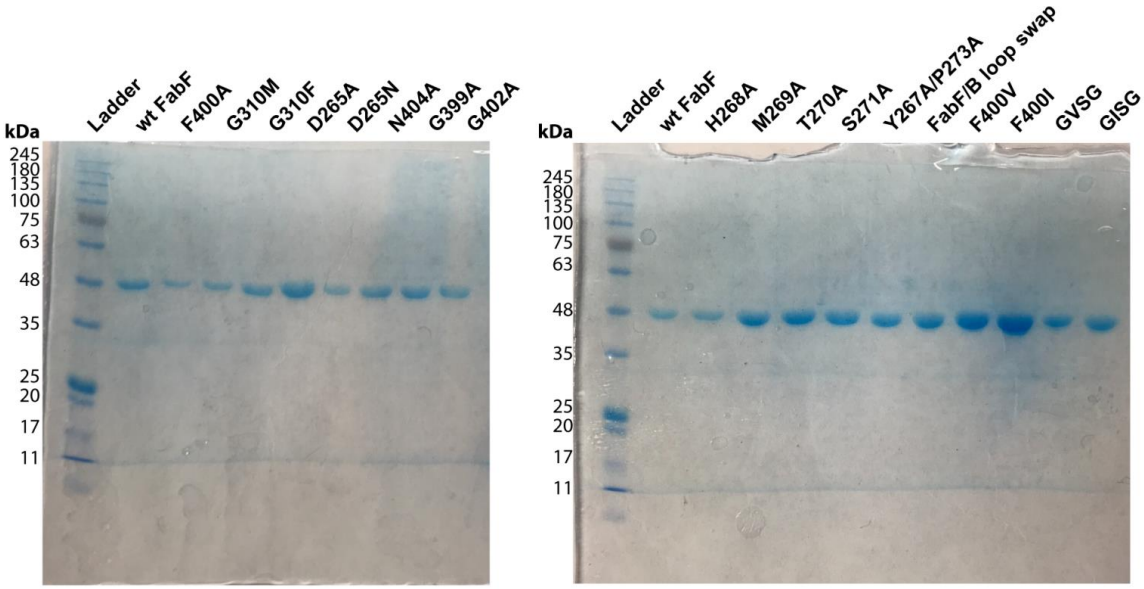


Figure S3.11: SDS-PAGE gels of all FabF constructs used in kinetic assays.

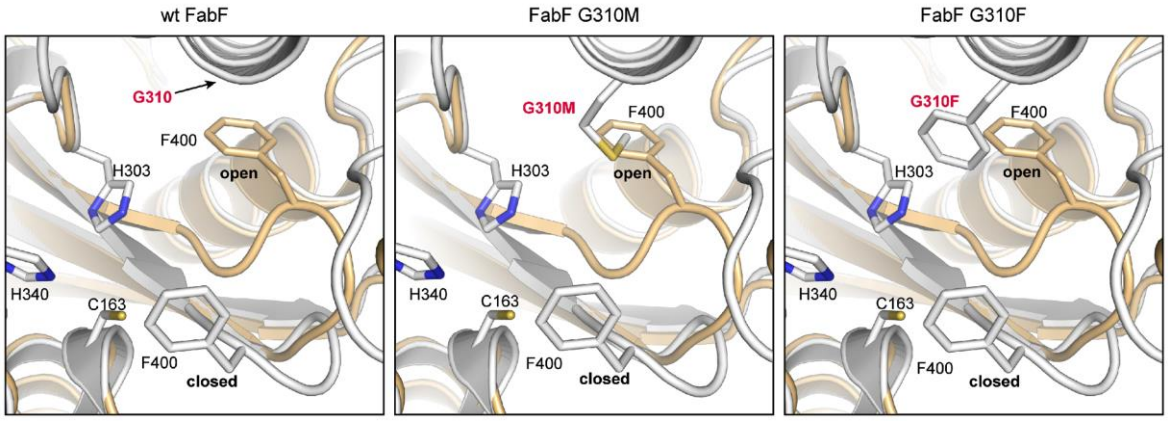


Figure S3.12: Rationale for FabF pocket block mutants. The open conformation of loop 1 positions Phe400 into a previously unoccupied pocket. Inserting bulky amino acids into this position should block access to the gate-open conformation, thereby limiting gate function.

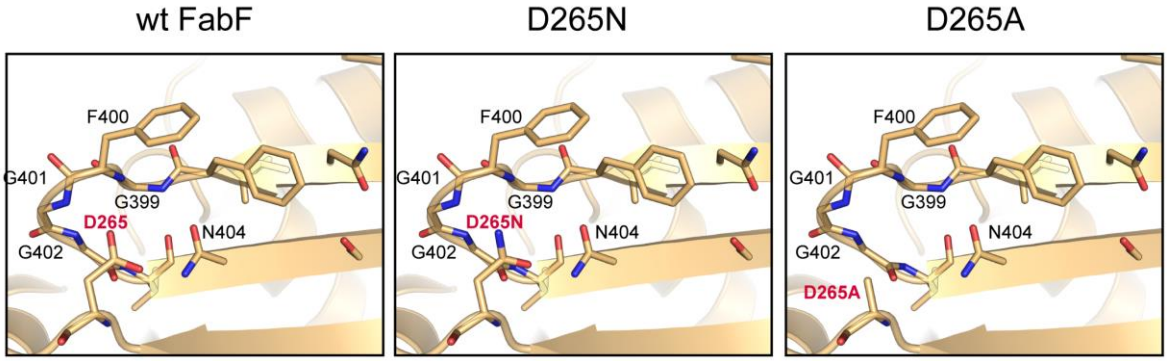


Figure S3.13: Rationale for FabF destabilization mutants. The open conformation of loop 1 is stabilized by a hydrogen bonding interaction network coordinated by the conserved D265 residue from loop 2. Mutation of this negatively charged aspartate residue for polar or non-charged amino acids should disrupt this network, thereby destabilizing the open conformation and inhibiting gate function.

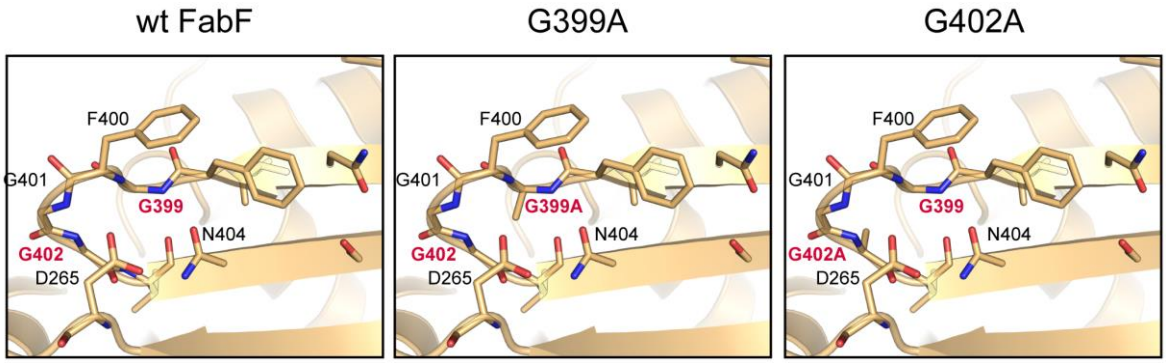


Figure S3.14: Rationale for FabF flex reduction mutants. Movement of loop 1 to the gate open conformation likely requires the conformational freedom of the conserved glycine residues from the GFGG motif. Mutation of these glycine residues (G399 and G402A) to alanine should limit the allowable backbone dihedral conformations, thus inhibiting gate function by hindering movement to the open conformation.

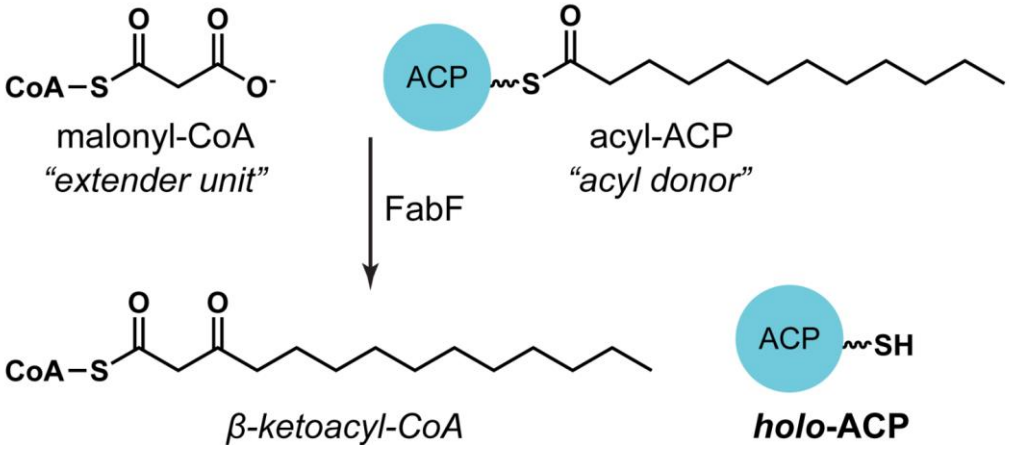


Figure S3.15: Overview of the KS catalyzed condensation reaction assay. The assay utilizes acyl-ACPs as acyl donors and malonyl-CoA as the extender unit to produce β-ketoacyl-CoA and *holo*-ACP. *holo*-ACP formation is monitored to follow the reaction.

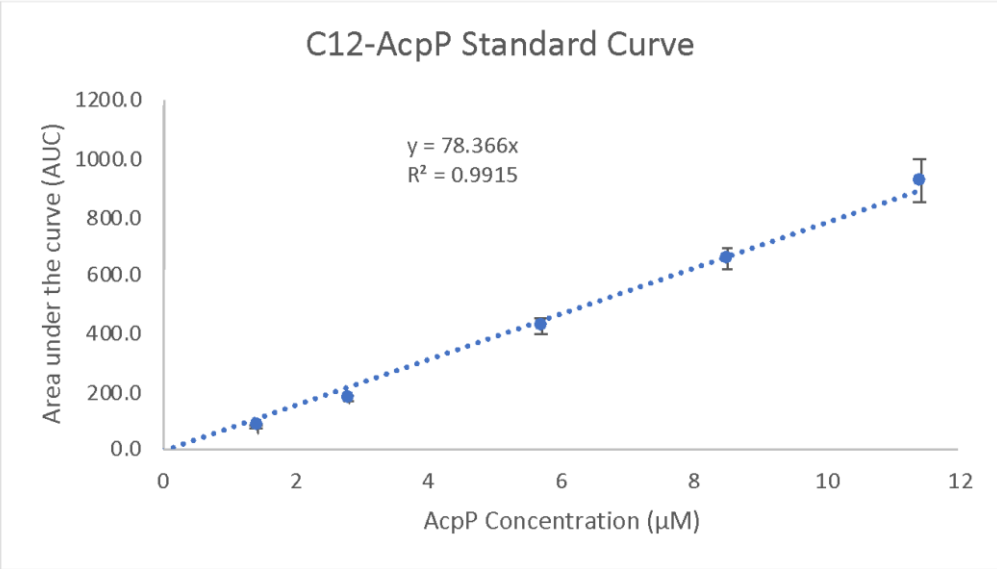


Figure S3.16: C12-AcpP calibration curve used for quantifying concentrations in kinetic assays.

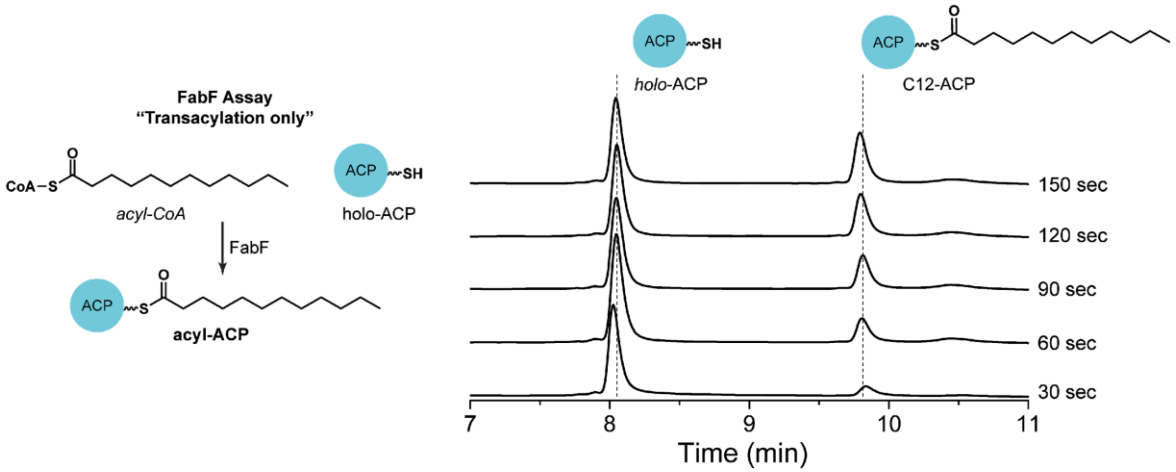


Figure S3.17: Overview of the KS catalyzed transacylation assay. The transacylation assay utilizes an acyl-CoA donor and *holo*-ACP acceptor and monitors the formation of the resulting acyl-ACP. A stacked HPLC chromatogram demonstrates the ability to monitor acyl-ACP formation over time, thereby allowing accurate rates to be determined by this assay format.

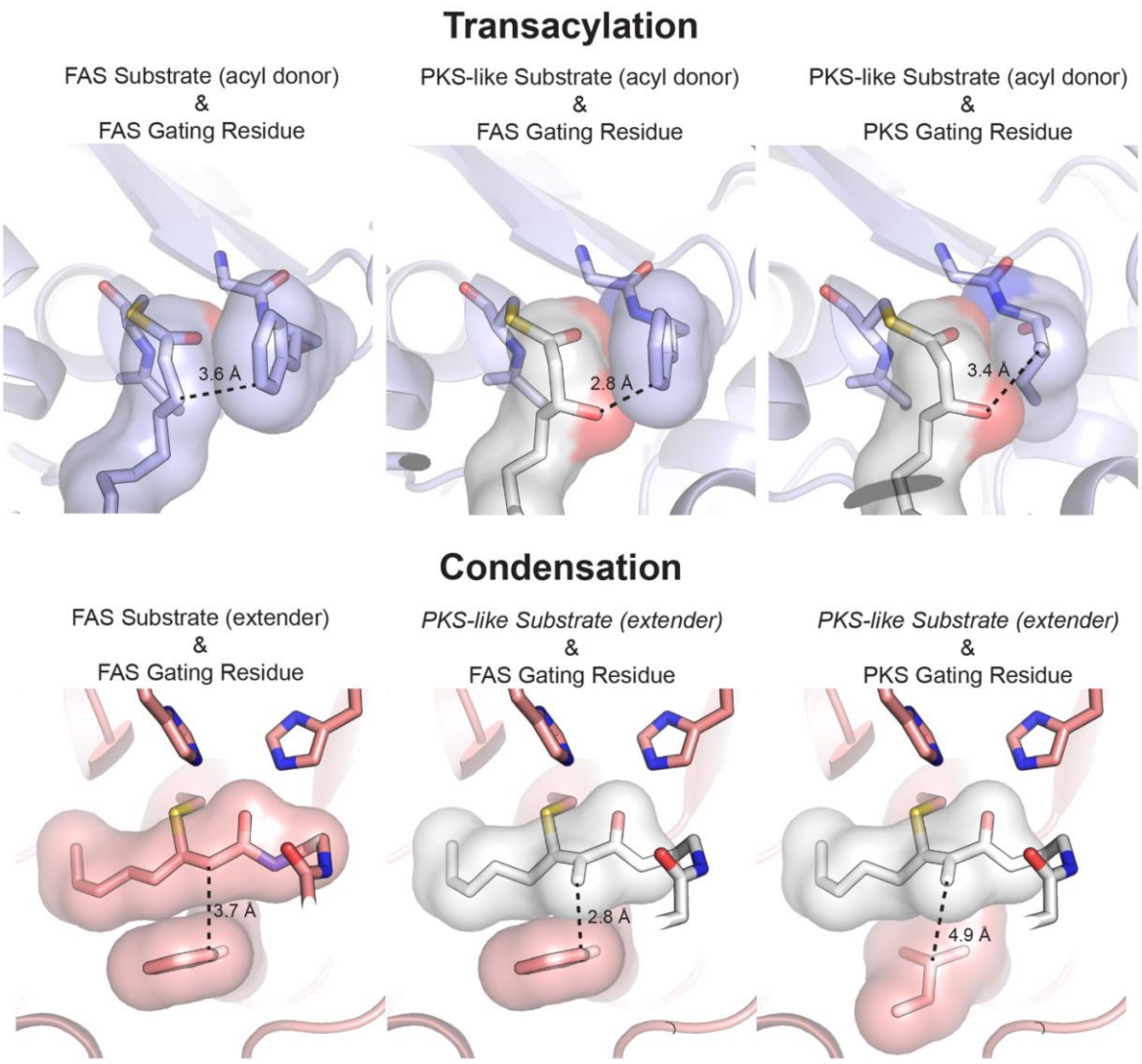


Figure S3.18: FAS vs TIPKS loop 1 gating residues. FAS and TIPKS have different loop 1 gating residues that may be better suited for turning over substrates in their respective pathways. In the transacylation reaction, the FAS-like substrate with FAS gating residue is derived from the C12-FabF acyl-enzyme intermediate structure (PDB: 2GFY). Modelling in a β-hydroxy-decanoic acid instead of decanoic acid (C12) demonstrates that there may be steric clashes between the β-position and the FAS gating residue. Replacement with Isoleucine could provide additional space for β-branched substrates. In the condensation reaction, represented by C8-AcpP=FabF (this study), the FAS extender unit is malonyl-AcpP. On the other hand, modelling in a methyl moiety that would be representative of a reaction with methylmalonyl-AcpP, a typical TIPKS extender unit, there may be a steric clash with the FAS gating residue. Replacement of the FAS gating residue with isoleucine could provide additional space for α-branched extender units typical of TIPKS systems.

Table S3.1: Table 1 x-ray crystallography data collection and refinement statistics.

	C8-AcpP-FabF	C16:1-AcpP-FabF
Wavelength (Å)	1	1
Resolution range (Å)	48.4 - 2.65 (2.745 - 2.65)	54.17 - 2.0 (2.071 - 2.0)
Space group	P 21 21 21	P 43 21 2
Unit cell (Å, º)	83.8 92.1 113.7 90 90 90	86.9 86.9 114.4 90 90 90
Total reflections	364173 (37926)	622926 (41223)
Unique reflections	26219 (2554)	30391 (2198)
Multiplicity	13.9 (14.8)	20.5 (18.8)
Completeness (%)	99.92 (100.00)	99.96 (99.93)
Mean I/sigma(I)	7.49 (1.72)	8.7 (1.4)
Wilson B-factor ($Å^2$)	36.83	24.31
R-merge	0.3958 (1.907)	.438 (4.396)
R-meas	0.411 (1.974)	.449 (4.519)
R-pim	0.11 (0.5101)	0.099 (1.037)
CC1/2	0.989 (0.713)	0.994 (0.549)
CC*	0.997 (0.912)	0.996 (0.562)
Reflections used in refinement	26203 (2554)	30316 (2972)
Reflections used for R-free	1299 (137)	1991 (196)
R-work	0.2221 (0.3027)	0.1671 (0.2756)
R-free	0.2664 (0.3773)	0.2068 (0.3442)
CC(work)	0.946 (0.833)	.945 (.759)
CC(free)	0.901 (0.752)	.913 (.650)
Number of non-hydrogen atoms	7266	3973
macromolecules	7070	3643
ligands	62	46
solvent	134	284
Protein residues	973	488
RMS (bonds)	0.004	0.004
RMS (angles)	0.97	1.02
Clashscore	6.15	3.14
Rama distribution Z-score	-1.70	-.41
Ramachandran Favored	917 (95.03%)	466 (96.2%)
Ramachandran Outlier	1 (.10 %)	0 (0%)
Average B-factor	47.16	32.36
macromolecules	47.42	31.61
ligands	44.53	51.18
solvent	34.98	38.92
Number of TLS groups	24	16

Table S3.2: Primers for site directed mutagenesis (SDM).

Name	Sequence
FabF-F400V-F	*TTCGGCGTTGGTGGCACTAATGGTTCTTTGATCTTTAAAAAGATC*
FabF-F400V-R	GCCACCAACGCCGAAGGAGTTACACAGAGTGTATTCCATTC
FabF-F400I-F	*TTCGGCATTGGTGGCACTAATGGTTCTTTGATCTTTAAAAAGATC*
FabF-F400I-R	GCCACCAATGCCGAAGGAGTTACACAGAGTGTATTCCATTC
FabF-GVSG-F	*TTCGGCGTTTCTGGCACTAATGGTTCTTTGATCTTTAAAAAGATC*
FabF-GVSG-R	GCCAGAAACGCCGAAGGAGTTACACAGAGTGTATTCCATTC
FabF-GISG-F	*TTCGGCATTTCTGGCACTAATGGTTCTTTGATCTTTAAAAAGATC*
FabF-GISG-R	GCCAGAAATGCCGAAGGAGTTACACAGAGTGTATTCCATTC
FabF-N404A-F	TTCGGTGGCACTGCGGGTTCTTTGATCTTTAAAAAGATCTAA
FabF-N404A-R	CGCAGTGCCACCGAAGCCGAAGGAGTTACACAGAGT
FabF-Y267A-P273A-F	GCTCATATGACGTCACCGGCAGAAAATGGCGCAGGCGCAGCTC
FabF-Y267A-P273A-F	TGCCGGTGACGTCATATGAGCAGCATCGCTGCTCATACCAAAGCC
FabF-H268A-F	TATGCGATGACGTCACCGCCAGAAAATGGCGCAGGCGCA
FabF-H268A-R	CGGTGACGTCATCGCATAAGCATCGCTGCTCATACCAAAGCC
FabF-M269A-F	TATCATGCGACGTCACCGCCAGAAAATGGCGCAGGCGCA
FabF-M269A-R	CGGTGACGTCGCATGATAAGCATCGCTGCTCATACCAAAGCC
FabF-S271A-F	TATCATATGACGGCGCCGCCAGAAAATGGCGCAGGCGCA
FabF-S271A-R	CGGCGCCGTCATATGATAAGCATCGCTGCTCATACCAAAGCC
FabF-FabB-LoopSwap-F	GTTGCTCCGTCTGGCGCAGGCGCAGCTCTG
FabF-FabB-LoopSwap-R	CATGTCTGCACCATCGCTGCTCATACCAAAGCCG

Supplementary note S3.1 Biological protocols

Protein expression and purification: *E. coli* AcpP and *E. coli* FabF were grown according to previously reported protocols.[24,28]

Site Directed Mutagenesis: All site directed mutagenesis constructs were produced using the method developed Liu and Naismith.[92] The primers used to create mutations specifically for this study are provided in Table S3.2. Primers for all other FabF mutants not found in Table S3.2 were reported previously.[24] The FabF/FabB loop swap construct was created using the around-the-horn method from NEB with overhangs that coded for the loop 2 region to be inserted into FabF.

Loading AcpP with probe molecules (one-pot method): Preparation of all *crypto*-ACPs via the one-pot method was done as previously described.[24,28]

Preparation of AcpP=FabF crosslinked complexes: Preparation of C16:1-*crypto*-AcpP=FabF and C8Cl-crypto-AcpP=FabF complexes was done as previously described.[24]

Crystallization, data collection, and refinement: Crystals of the C16:1-*crypto*-AcpP=FabF and C8Cl-*crypto*-AcpP=FabF crosslinked complexes were grown by cross-seeding using crystals of C16-*crypto*-AcpP=FabF. 1.5 μL of crosslinked complex (8 mg•mL^{-1}) was mixed with 1.0 μL of the corresponding mother liquor and 0.5 μL of seed stock and the mixture was placed inverted over 500 μL of the well solution (hanging-drop method) at 6 °C. The same crystal conditions were used as previously described (26–30% PEG 8 K, 0.1 M sodium cacodylate pH 6.5, and 0.3 M NaOAc). All data were collected at the Advanced Light Source (ALS) synchrotron at Berkeley. Data were indexed using iMosflm[93] then scaled and merged using the aimless program from the CCP4 software suite.[94] Scaled and

merged reflection output data was used for molecular replacement and model building in PHENIX.[95] Phases were solved using molecular replacement using C16-*crypto*-AcpP=FabF (PDB ID: 6OKG) as a search model with program Phaser from the PHENIX software suite. Model building was performed using Coot.[94] The parameter files for each of the covalently bonded 4'-phosphopantetheine analogs were generated using eLBOW in the PHENIX software suite.[96] Manually programmed parameter restraints using values provided by Jligand[97] were used to create the associated covalent bonds between 4'-phosphopantetheine to Ser36 and Cys92 during refinement.

Crosslinking assays with C16:1-*crypto*-ACP and FabF: Each reaction was set up to contain 40 μM of *crypto*-AcpP, 20 μM of FabF, and buffer (25 mM Tris, 150 mM NaCl, 10% glycerol, 0.5 mM TCEP, pH depends), as previously described.[24] All sample stocks were originally in pH 7.4 buffer. Prior to the pH 6.0 and pH 7.0 crosslinking assays, samples were buffer exchanged *via* mini spin column to the corresponding pH. To stop the reaction at the designated time point, 2.5 μL of the reaction solution was mixed with 10 μL of 3X SDS dye. Samples were analyzed by 12% SDS-PAGE.

Preparation of *holo*-AcpP and acyl-AcpP substrates: Preparation of *holo*-ACP and acyl-AcpP substrates was performed as previously described.[24,28]

Ketosynthase condensation assay: Ketosynthase condensation assays were carried out as previously described.[28] In assays run with 1000 μM malonyl-CoA, concentrations of all other assay components were held constant. Scouting reactions to evaluate effective time points for accurate rate determination were performed for each FabF mutant before running

assays in triplicate. All assays were performed in biological triplicate and the reported error is the standard deviation on the mean.

Ketosynthase transacylation assay: The FabF transacylation assay mixture contained 50 mM sodium phosphate (pH 6.8), 50 mM NaCl, 0.5 mM TCEP, 150 μM lauroyl-CoA, 25 μM holo-AcpP (wt), and 1 μM FabF (0 μM FabF for the control reaction) at 28 °C. Twenty microliters of each reaction mixture was quenched by the addition of 10 μL of 1.0% TFA (final TFA concentration of 0.33%) at five different time points to ensure accurate linear regression with at least four data points. Product formation was monitored by chromatographic separation using an Agilent 1100 series high-performance liquid chromatography (HPLC) instrument equipped with an Ascentis Express Peptide ES-18 column (15 cm × 4.6 mm, 2.7 μm) with a 1 mL/min flow rate and a detection wavelength of 210 nm. A gradient method was used [A: doubly distilled H2O with 0.1% (v/v) TFA; B: HPLC grade CH3CN with 0.1% (v/v) TFA]: 25% B, 0 min; 25% B, 2 min; 75% B, 12 min; 95% B, 13 min; 95% B, 15 min; 25% B, 17 min; 25% B, 20 min. Chromatographic traces of the wt-FabF transacylation assay can be found in Figure S3.13. C12-AcpP concentrations were calculated from a standard curve and used to determine initial reaction rates. All assays were performed in biological triplicate and the reported error is the standard deviation on the mean.

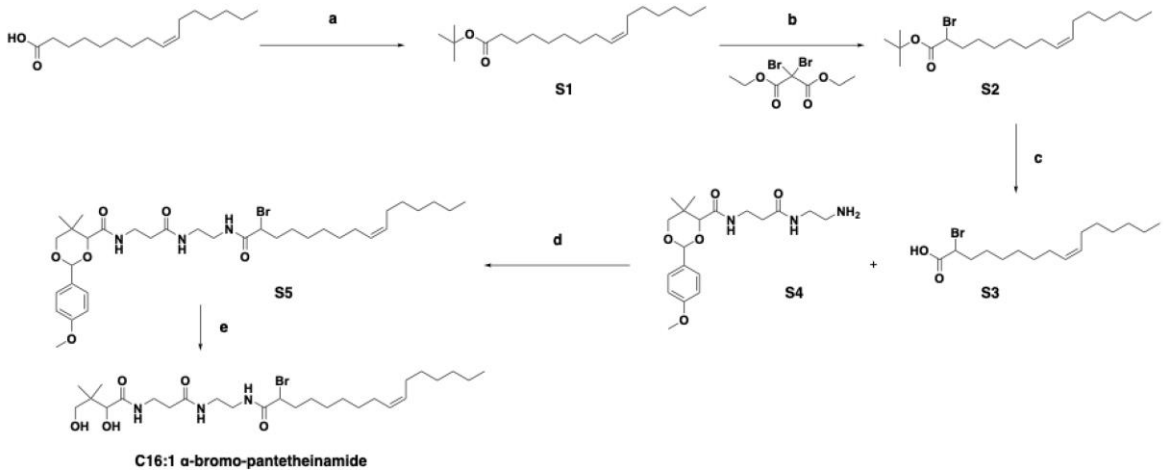
Analysis of kinetic data: Kinetic rate data were analyzed with excel and Prism. Data points were fit by linear regression to give rates for substrate turnover. Data were converted molar rates using the standard curve in figure S16 and divided by enzyme concentration to give

apparent turnover numbers. The averages of triplicates were calculated, and the standard deviation is reported for each biological triplicate.

Supplementary note S3.2 Synthetic protocols

Synthesis of C8Cl-chlorovinylacrylamide-pantetheinamide crosslinker: Preparation and synthesis of the C8Cl-chlorovinylacrylamide-pantetheinamide crosslinker was done as previously reported.[24,61]

Synthesis of C16:1-α-bromo-pantetheinamide crosslinker:



Reagents and conditions: [a] $(BOC)_2O$, DMAP, DIEA, CH_2Cl_2, rt, 38%; [b] n-BuLi, DIPA, THF, -78 °C 63%; [c] 40% TFA/CH_2Cl_2, quantitative yield; [d] EDC, DMAP, CH_2Cl_2, rt, 40%; [e] 80% aq. acetic acid, rt, 37%.

General synthetic methods: All commercial reagents were used as provided unless otherwise indicated. Compound **S4** is a known compound. This compound was prepared according to published literature procedures.[60] All reactions were carried out under an atmosphere of nitrogen in dry solvents with oven-dried glassware and constant magnetic stirring unless otherwise noted. ^{1}H-NMR spectra were recorded at 500 MHz. ^{13}C-NMR spectra were recorded at 125 MHz on JEOL NMR spectrometers and standardized to the NMR solvent signal as reported by Gottlieb.[98] Multiplicities are given as s = singlet, d =

doublet, t = triplet, q = quartet, p = pentet, dd = doublet of doublets, ddd = doublet of doublet of doublets, dt = doublet of triplets, tt = triplet of triplets, dq = doublet of quartets, dp = doublet of pentets, qt = quartet of triplets, ddt = doublet of doublet of triplets, ddq = doublet of doublet of quartet, and m = multiplet using integration and coupling constant in Hertz. TLC analysis was performed using Silica Gel 60 F254 plates (Merck) and visualization was accomplished with ultraviolet light (λ = 254 nm) and/or the appropriate stain [phosphomolybdic acid, iodine, ninhydrin, and potassium permanganate]. Silica gel chromatography was carried out with SilicaFlash F60 230-400 mesh (Silicycle), according to the method of Still.[99]

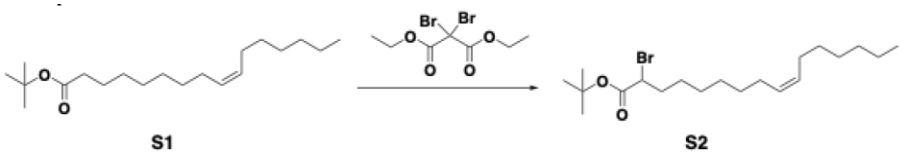
Synthesis of C16:1-α-bromo-pantetheinamide

Chemical Synthesis of S1. Compound name in bold refers to the structures shown above.

S1

$(BOC)_2O$ (197 mg, 0.9 mmol), DIEA (112 mg, 0.86 mmol), and DMAP (29 mg, 0.24 mmol) were added to a solution of the starting material, palmitoleic acid, (200 mg, 0.79 mmol) in CH_2Cl_2 (5 mL). The solution was stirred at room temperature overnight. The reaction mixture was diluted with CH_2Cl_2 and washed with H_2O three times, dried over Na_2SO_4, filtered, and evaporated to dryness. The resulting red oil was purified by flash chromatography (hexane:EtOAc = 9:1) to afford compound **S1** as an oil (92 mg, 38%). ^1H NMR (500 MHz, CDCl 3): δ NMR

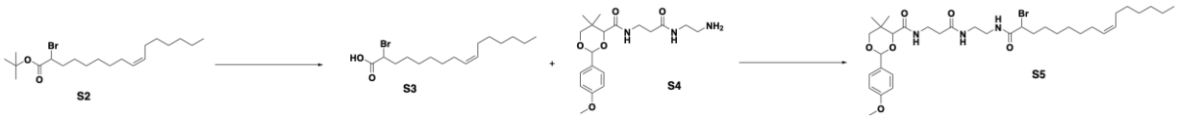
Chemical synthesis of S2.



S1 S2

A 1.6 M solution of n-BuLi in hexane (0.6 mL, 0.96 mmol) was added to a solution of DIPA (0.15 mL, 1.04 mmol) in THF (2 mL) at -78 °C under inert conditions and stirred for 15 minutes to prepare LDA *in situ*. The compound **S1** (270 mg, 0.87 mmol) was dissolved in THF (3 mL) in a different vial at -78 °C and LDA (0.13 mL, 0.96 mmol) was cannulated into the solution. The reaction mixture was stirred at room temperature for 1 hour. In a different vial, diethyl dibromomalonate (0.6 mL, 2.17 mmol) was dissolved in THF (3 mL) and cooled to -78 °C and this solution was cannulated into the mixture containing the compound **S1** and LDA, which was then stirred for 2 h at -78 °C. The reaction mixture was quenched with 2% HCl, diluted with brine and hexane, washed with H_2O three times, dried over Na_2SO_4, filtered, and evaporated to dryness. The resulting oil was purified by flash chromatography (hexane:DCM gradient from 10:1 to 8:1 to 6:1) to give compound **S2** as an oil (213 mg, 63%). ^{1}H NMR (500 MHz, CDCl 3): δ NMR

Chemical synthesis of S5.

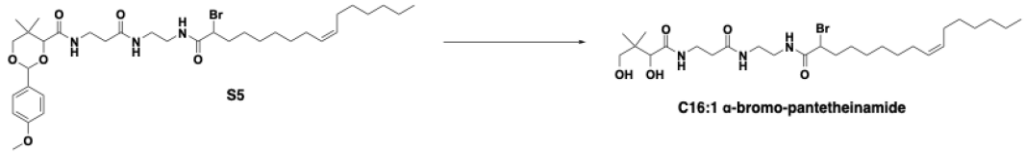


S2 S3 S4 S5

The compound **S2** (82 mg, 0.21 mmol) was dissolved in 40% TFA/CH_2Cl_2 and stirred for 1 h at room temperature. Solvent was evaporated by blowing air into the vial and reaction mixture was azeotrope with cyclohexane to give compound **S3** in quantitative yield. Compound **S4** (80 mg, 0.21 mmol), EDC (39 mg, 0.25 mmol), and DMAP (5 mg, 0.04 mmol)

were added to a solution of Compound **S3** (70 mg, 0.21 mmol) in DMF (5 mL). The solution was stirred at room temperature overnight. The mixture was diluted with water and extracted with DCM three times. The combined organic layers were washed with H_2O three times and then washed with brine, dried over Na_2SO_4, filtered, and evaporated to dryness. The resulting oil was purified by flash chromatography (DCM:methanol = 9:1, 1% TEA) to give compound **S5** as an oil (58 mg, 40%). ^{1}H NMR (500 MHz, CDCl 3): δ 7.43 (d, J = 8.8 Hz, 2H), 7.06 (m, 1H), 7.01 (m, 1H), 6.92 (d, J = 8.8 Hz, 2H), 6.56 (bs, 1H), 5.47 (s, 1H), 5.34 (m, 2H), 4.23 (dt, J = 8.4, 5.2 Hz, 1H), 4.10 (m, 1H), 3.82 (s, 3H), 3.69 (q, J = 11.2 Hz, 2H), 3.56 (m, 2H), 3.39 (m, 4H), 2.44 (t, J = 6.2 Hz, 2H), 2.05 (m, 6H), 1.27 (m, 16H), 1.10 (m, 6H), 0.88 (t, J = 6.9 Hz, 3H). 13 C NMR (125 MHz, $CDCl_3$) δ 172.0, 170.1, 169.9, 160.4, 130.2, 129.8, 127.7, 113.9, 101.5, 83.9, 78.6, 55.5, 51.4. 40.7, 39.5, 36.6, 35.8, 35.1, 33.3, 31.9, 31.7, 29.9, 29.2, 29.1, 28.9, 27.4, 27.4, 27.3, 22.8, 19.3, 14.3. HRMS (ESI) m/z $[M+Na]^+$ Calculated for $C_{35}H_{56}BrN_3O_6Na$: 716.3250; Found 716.3248.

Chemical synthesis of C16:1-α-bromo-pantetheniamide



Compound **S5** (50 mg, 72 mmol) was dissolved in aqueous 80% acetic acid. The reaction mixture was stirred at room temperature overnight. The solution was dried by blowing air into the vial. The resulting oil was purified by flash chromatography (DCM:methanol = 5:1) to give the final product as an oil (15 mg, 37%). ^{1}H NMR (500 MHz, CDCl 3): δ 7.00 (bs, 1H), 6.63 (bs, 1H), 6.48 (bs, 1H), 5.35 (m, 2H), 4.26 (m, 1H), 4.03 (d, J = 10.6 Hz, 1H), 3.70

(m, 1H), 3.50 (m, 5H), 3.36 (m, 2H), 2.45 (m, 2H), 2.34 (t, J = 7.6 Hz, 1H), 2.00 (m, 6H), 1.40 (t, J = 7.4 Hz, 1H), 1.27 (m, 16H), 1.02 (s, 3H), 0.93 (s, 3H), 0.88 (m, 3H). ^{13}C NMR (125 MHz, CDCl 3) δ 174.1, 173.9, 170.6, 130.3, 129.7, 76.4, 71.2, 51.2, 40.4, 39.4, 36.3, 35.6, 32.3, 32.1, 31.9, 29.9, 29.2, 28.9, 27.4, 27.4, 27.3, 23.6, 22.8, 21.9, 20.7, 14.3. HRMS (ESI) m/z $[M+Na]^+$ Calculated for $C_{27}H_{50}BrN_3O_5Na$: 598.2832; Found 598.2826.

References

(1) White, S. W.; Zheng, J.; Zhang, Y.-M.; Rock, C. O. The Structural Biology of Type II Fatty Acid Biosynthesis. *Annu. Rev. Biochem.* **2005**, *74*, 791 − 831.

(2) Beld, J.; Lee, D. J.; Burkart, M. D. Fatty acid biosynthesis revisited: structure elucidation and metabolic engineering. *Mol. Biosyst.* **2015**, *11*, 38 − 59.

(3) Maier, T.; Leibundgut, M.; Boehringer, D.; Ban, N. Structure and function of eukaryotic fatty acid synthases. *Q. Rev. Biophys.* **2010**, *43*, 373 − 422.

(4) Torella, J. P.; Ford, T. J.; Kim, S. N.; Chen, A. M.; Way, J. C.; Silver, P. A. Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 11290 − 11295.

(5) Handke, P.; Lynch, S. A.; Gill, R. T. Application and engineering of fatty acid biosynthesis in Escherichia coli for advanced fuels and chemicals. *Metab. Eng.* **2011**, *13*, 28 − 37.

(6) Fischer, M.; Grininger, M. Strategies in megasynthase engineering - fatty acid synthases (FAS) as model proteins. *Beilstein J. Org. Chem.* **2017**, *13*, 1204 − 1211.

(7) Zhu, Z.; Zhou, Y. J.; Krivoruchko, A.; Grininger, M.; Zhao, Z. K.; Nielsen, J. Expanding the product portfolio of fungal type I fatty acid synthases. *Nat. Chem. Biol.* **2017**, *13*, 360.

(8) Lu, X.; Vora, H.; Khosla, C. Overproduction of free fatty acids in *E. coli*: Implications for biodiesel production. *Metab. Eng.* **2008**, *10*, 333 − 339.

(9) Wu, H.; Karanjikar, M.; San, K.-Y. Metabolic engineering of *Escherichia coli* for efficient free fatty acid production from glycerol. *Metab. Eng.* **2014**, *25*, 82 − 91.

(10) Haushalter, R. W.; Kim, W.; Chavkin, T. A.; The, L.; Garber, M. E.; Nhan, M.; Adams, P. D.; Petzold, C. J.; Katz, L.; Keasling, J. D. Production of anteiso-branched fatty acids in *Escherichia coli*; next generation biofuels with improved cold-flow properties. *Metab. Eng.* **2014**, *26*, 111 − 118.

(11) Maier, T.; Leibundgut, M.; Ban, N. The Crystal Structure of a Mammalian Fatty Acid Synthase. *Science* **2008**, *321*, 1315 − 1322.

(12) Leibundgut, 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.

(13) Johansson, P.; Wiltschi, B.; Kumari, P.; Kessler, B.; Vonrhein, C.; Vonck, J.; Oesterhelt, D.; Grininger, M. Inhibition of the fungal fatty acid synthase type I multienzyme complex. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 12803.

(14) Ridley, C. P.; Lee, H. Y.; Khosla, C. Evolution of polyketide synthases in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4595 −4600.

(15) Jenke-Kodama, H.; Sandmann, A.; Mu¨ller, R.; Dittmann, E. Evolutionary Implications of Bacterial Polyketide Synthases. *Mol. Biol. Evol.* **2005**, *22*, 2027 − 2039.

(16) Herbst, D. A.; Townsend, C. A.; Maier, T. The architectures of iterative type I PKS and FAS. *Nat. Prod. Rep.* **2018**, *35*, 1046 − 1069.

(17) Hertweck, C. The Biosynthetic Logic of Polyketide Diversity. *Angew. Chem., Int. Ed.* **2009**, *48*, 4688 − 4716.

(18) Chen, A.; Re, R. N.; Burkart, M. D. Type II fatty acid and polyketide synthases: deciphering protein−protein and protein−substrate interactions. *Nat. Prod. Rep.* **2018**, *35*, 1029 − 1045.

(19) Vagelos, P. R.; Majerus, P. W.; Alberts, A. W.; Larrabee, A. R.; Ailhaud, G. P. Structure and function of the acyl carrier protein. *Fed. Proc.* **1966**, *25*, 1485 − 1494.

(20) Larrabee, A. R.; McDaniel, E. G.; Bakerman, H. A.; Vagelos, P. R. Acyl carrier protein. V. Identification of 4'-phosphopantetheine bound to a mammalian fatty acid synthetase preparation. *Proc. Natl. Acad. Sci. U.S.A.* **1965**, *54*, 267 − 273.

(21) Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Markwick, P. R. L.; Bruegger, J.; Caldara-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 − 431.

(22) Dodge, G. J.; Patel, A.; Jaremko, K. L.; McCammon, J. A.; Smith, J. L.; Burkart, M. D. Structural and dynamical rationale for fatty acid unsaturation in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 6775 − 6783.

(23) Milligan, J. C.; Lee, D. J.; Jackson, D. R.; Schaub, A. J.; Beld, J.; Barajas, J. F.; Hale, J. J.; Luo, R.; Burkart, M. D.; Tsai, S.-C. Molecular basis for interactions between an acyl carrier protein and a ketosynthase. *Nat. Chem. Biol.* **2019**, *15*, 669 − 671.

(24) Mindrebo, J. T.; Patel, A.; Kim, W. E.; Davis, T. D.; Chen, A.; Bartholow, T. G.; La Clair, J. J.; McCammon, J. A.; Noel, J. P.; Burkart, M. D. Gating mechanism of elongating β-ketoacyl-ACP synthases. *Nat. Commun.* **2020**, *11*, 1727.

(25) Misson, L. E.; Mindrebo, J. T.; Davis, T. D.; Patel, A.; Andrew McCammon, J.; Noel, J. P.; Burkart, M. D. Interfacial plasticity facilitates high reaction rate of *E. coli* FAS malonyl-CoA:ACP transacylase, FabD. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117*, 24224.

(26) Agarwal, V.; Lin, S.; Lukk, T.; Nair, S. K.; Cronan, J. E. Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17406 − 17411.

(27) Masoudi, A.; Raetz, C. R. H.; Zhou, P.; Pemble IV, C. W. Chasing Acyl-Carrier-Protein Through a Catalytic Cycle of Lipid A Production. *Nature* **2014**, *505*, 422 − 426.

(28) Mindrebo, J. T.; Misson, L. E.; Johnson, C.; Noel, J. P.; Burkart, M. D. Activity Mapping the Acyl Carrier Protein–Elongating Ketosynthase Interaction in Fatty Acid Biosynthesis. *Biochemistry* **2020**, *59*, 3626 − 3638.

(29) Zhang, Y.-M.; Rao, M. S.; Heath, R. J.; Price, A. C.; Olson, A. J.; Rock, C. O.; White, S. W. Identification and Analysis of the Acyl Carrier Protein (ACP) Docking Site on β-Ketoacyl-ACP Synthase III. *J. Biol. Chem.* **2001**, *276*, 8231 − 8238.

(30) Mindrebo, J. T.; Patel, A.; Misson, L. E.; Kim, W. E.; Davis, T. D.; Ni, Q. Z.; La Clair, J. J.; Burkart, M. D. Structural Basis of Acyl Carrier Protein Interactions in Fatty Acid and Polyketide Biosynthesis. In *Comprehensive Natural Products III: Chemistry and Biology*; Liu, H.-W., Begley, T. P., Eds.; Elsevier: Oxford, **2020**; pp 61 − 122.

(31) Heil, C. S.; Wehrheim, S. S.; Paithankar, K. S.; Grininger, M. Fatty Acid Biosynthesis: Chain-Length Regulation and Control. *ChemBioChem* **2019**, *20*, 2298 − 2321.

(32) Zhang, Y.-M.; Hurlbert, J.; White, S. W.; Rock, C. O. Roles of the Active Site Water, Histidine 303, and Phenylalanine 396 in the Catalytic Mechanism of the Elongation Condensing Enzyme of *Streptococcus pneumoniae*. *J. Biol. Chem.* **2006**, *281*, 17390 − 17399.

(33) von Wettstein-Knowles, P.; Olsen, J. G.; McGuire, K. A.; Henriksen, A. Fatty acid synthesis. Role of active site histidines and lysine in Cys-His-His-type beta-ketoacyl-acyl carrier protein synthases. *FEBS J.* **2006**, *273*, 695 − 710.

(34) Olsen, J. G.; Kadziola, A.; von Wettstein-Knowles, P.; Siggaard-Andersen, M.; Larsen, S. Structures of beta-Ketoacyl-Acyl Carrier Protein Synthase I Complexed with Fatty Acids Elucidate its Catalytic Machinery. *Structure* **2001**, *9*, 233 − 243.

(35) Wang, J.; Soisson, S. M.; Young, K.; Shoop, W.; Kodali, S.; Galgoci, A.; Painter, R.; Parthasarathy, G.; Tang, Y. S.; Cummings, R.; Ha, S.; Dorso, K.; Motyl, M.; Jayasuriya, H.; Ondeyka, J.; Herath, K.; Zhang, C.; Hernandez, L.; Allocco, J.; Basilio, A´.; Tormo, J. R.;

Genilloud, O.; Vicente, F.; Pelaez, F.; Colwell, L.; Lee, S. H.; Michael, B.; Felcetto, T.; Gill, C.; Silver, L. L.; Hermes, J. D.; Bartizal, K.; Barrett, J.; Schmatz, D.; Becker, J. W.; Cully, D.; Singh, S. B. Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* **2006**, *441*, 358.

(36) Rittner, A.; Paithankar, K. S.; Himmler, A.; Grininger, M. Type I fatty acid synthase trapped in the octanoyl-bound state. *Protein Sci.* **2020**, *29*, 589 − 605.

(37) Gajewski, J.; Pavlovic, R.; Fischer, M.; Boles, E.; Grininger, M. Engineering fungal *de novo* fatty acid synthesis for short chain fatty acid production. *Nat. Commun.* **2017**, *8*, 14650.

(38) Gajewski, J.; Buelens, F.; Serdjukow, S.; Janßen, M.; Cortina, N.; Grubmu¨ller, H.; Grininger, M. Engineering fatty acid synthases for directed polyketide production. *Nat. Chem. Biol.* **2017**, *13*, 363.

(39) Rossini, E.; Gajewski, J.; Klaus, M.; Hummer, G.; Grininger, M. Analysis and engineering of substrate shuttling by the acyl carrier protein (ACP) in fatty acid synthases (FASs). *Chem. Commun.* **2018**, *54*, 11606 − 11609.

(40) Klaus, M.; Grininger, M. Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* **2018**, *35*, 1070 − 1081.

(41) Klaus, M.; Buyachuihan, L.; Grininger, M. Ketosynthase Domain Constrains the Design of Polyketide Synthases. *ACS Chem. Biol.* **2020**, *15*, 2422 − 2432.

(42) Rock, C. O.; Cronan, J. E. Escherichia coli as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim. Biophys. Acta Lipids Lipid. Metabol.* **1996**, *1302*, 1 − 16.

(43) Cronan, J. E.; Thomas, J. Bacterial Fatty Acid Synthesis and its Relationships with Polyketide Synthetic Pathways. In *Methods in Enzymology*; Hopwood, D. A., Ed.; Academic Press: Cambridge, MA, **2009**; Chapter 17, pp 395 − 433.

(44) Edwards, P.; Sabo Nelsen, J.; Metz, J. G.; Dehesh, K. Cloning of the fabF gene in an expression vector and *in vitro* characterization of recombinant fabF and fabB encoded enzymes from *Escherichia coli*. *FEBS Lett.* **1997**, *402*, 62 − 66.

(45) 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*. *J. Biol. Chem.* **1980**, *255*, 11949 − 11956.

(46) Gora, A.; Brezovsky, J.; Damborsky, J. Gates of enzymes. *Chem. Rev.* **2013**, *113*, 5871 − 5923.

(47) Khosla, C. Structures and Mechanisms of Polyketide Synthases. *J. Org. Chem.* **2009**, *74*, 6416 − 6420.

(48) Robbins, T.; Liu, Y.-C.; Cane, D. E.; Khosla, C. Structure and mechanism of assembly line polyketide synthases. *Curr. Opin. Struct. Biol.* **2016**, *41*, 10 − 18.

(49) Weissman, K. J. Genetic engineering of modular PKSs: From combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.* **2016**, *33*, 203 − 230.

(50) Dunn, B. J.; Cane, D. E.; Khosla, C. Mechanism and specificity of an acyltransferase domain from a modular polyketide synthase. *Biochemistry* **2013**, *52*, 1839 − 1841.

(51) D'Agnolo, G.; Rosenfeld, I. S.; Vagelos, P. R. Multiple forms of beta-ketoacyl-acyl carrier protein synthetase in *Escherichia coli*. *J. Biol. Chem.* **1975**, *250*, 5289 − 5294.

(52) Cronan, J. E.; Birge, C. H.; Vagelos, P. R. Evidence for two genes specifically involved in unsaturated fatty acid biosynthesis in *Escherichia coli*. *J. Bacteriol.* **1969**, *100*, 601 − 604.

(53) Heath, R. J.; Rock, C. O. Roles of the FabA and FabZ β-Hydroxyacyl-Acyl Carrier Protein Dehydratases in *Escherichia coli* Fatty Acid Biosynthesis. *J. Biol. Chem.* **1996**, *271*, 27795 − 27801.

(54) de Mendoza, D.; Klages Ulrich, A.; Cronan, J. E. Thermal regulation of membrane fluidity in *Escherichia coli*. Effects of overproduction of beta-ketoacyl-acyl carrier protein synthase I. *J. Biol. Chem.* **1983**, *258*, 2098 − 2101.

(55) de Mendoza, D.; Garwin, J. L.; Cronan, J. E. Overproduction of *cis*-vaccenic acid and altered temperature control of fatty acid synthesis in a mutant of *Escherichia coli*. *J. Bacteriol.* **1982**, *151*, 1608 − 1611.

(56) Garwin, J. L.; Cronan, J. E. Thermal modulation of fatty acid synthesis in *Escherichia coli* does not involve *de novo* enzyme synthesis. *J. Bacteriol.* **1980**, *141*, 1457 − 1459.

(57) Dong, H.; Cronan, J. E. Temperature Regulation of Membrane Composition in the Firmicute, *Enterococcus faecalis*, Parallels That of *Escherichia coli*. *Environ. Microbiol.* **2021**, *23*, 2683.

(58) Ishikawa, F.; Haushalter, R. W.; Burkart, M. D. Dehydratase-specific probes for fatty acid and polyketide synthases. *J. Am. Chem. Soc.* **2012**, *134*, 769 − 772.

(59) Worthington, A. S.; Burkart, M. D. One-pot chemo-enzymatic synthesis of reporter-modified proteins. *Org. Biomol. Chem.* **2006**, *4*, 44 − 46.

(60) 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 Chem. Biol.* **2006**, *11*, 687.

(61) Worthington, A. S.; Hur, G. H.; Meier, J. L.; Cheng, Q.; Moore, B. S.; Burkart, M. D. Probing the compatibility of type II ketosynthase-carrier protein partners. *Chembiochem* **2008**, *9*, 2096 −2103.

(62) Du, D.; Katsuyama, Y.; Horiuchi, M.; Fushinobu, S.; Chen, A.; Davis, T. D.; Burkart, M. D.; Ohnishi, Y. Structural basis for selectivity in a highly reducing type II polyketide synthase. *Nat. Chem. Biol.* **2020**, *16*, 776.

(63) Gulick, A. M.; Aldrich, C. C. Trapping interactions between catalytic domains and carrier proteins of modular biosynthetic enzymes with chemical probes. *Nat. Prod. Rep.* **2018**, *35*, 1156 − 1184.

(64) Davis, T. D.; Michaud, J. M.; Burkart, M. D. Active site labeling of fatty acid and polyketide acyl-carrier protein transacylases. *Org. Biomol. Chem.* **2019**, *17*, 4720 − 4724.

(65) Corpuz, J. C.; Podust, L. M.; Davis, T. D.; Jaremko, M. J.; Burkart, M. D. Dynamic visualization of type II peptidyl carrier protein recognition in pyoluteorin biosynthesis. *RSC Chem. Biol.* **2020**, *1*, 8 − 12.

(66) Sztain, T.; Patel, A.; Lee, D. J.; Davis, T. D.; McCammon, J. A.; Burkart, M. D. Modifying the Thioester Linkage Affects the Structure of the Acyl Carrier Protein. *Angew. Chem., Int. Ed. Engl.* **2019**, *58*, 10888 − 10892.

(67) Epstein, S. C.; Huff, A. R.; Winesett, E. S.; Londergan, C. H.; Charkoudian, L. K. Tracking carrier protein motions with Raman spectroscopy. *Nat. Commun.* **2019**, *10*, 2227.

(68) Beld, J.; Cang, H.; Burkart, M. D. Visualizing the chain-flipping mechanism in fatty-acid biosynthesis. *Angew. Chem., Int. Ed. Engl.* **2014**, *53*, 14456 − 14461.

(69) Crosby, J.; Crump, M. P. The structural role of the carrier protein−active controller or passive carrier. *Nat. Prod. Rep.* **2012**, *29*, 1111 − 1137.

(70) Cronan, J. E.; Birge, C. H.; Vagelos, P. R. Evidence for Two Genes Specifically Involved in Unsaturated Fatty Acid Biosynthesis in *Escherichia coli*. *J. Bacteriol.* **1969**, *100*, 601 − 604.

(71) Eccleston, V. S.; Cranmer, A. M.; Voelker, T. A.; Ohlrogge, J. B. Medium-chain fatty acid biosynthesis and utilization in Brassica napus plants expressing lauroyl-acyl carrier protein thioesterase. *Planta* **1996**, *198*, 46 − 53.

(72) Salas, J. J.; Ohlrogge, J. B. Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Arch. Biochem. Biophys.* **2002**, *403*, 25 − 34.

(73) Voelker, T. Plant Acyl-ACP Thioesterases: Chain-Length Determining Enzymes in Plant Fatty Acid Biosynthesis BT. In *Genetic Engineering: Principles and Methods*; Setlow, J. K., Ed.; Springer US: Boston, MA, **1996**; pp 111 − 133.

(74) Moche, M.; Schneider, G.; Edwards, P.; Dehesh, K.; Lindqvist, Y. Structure of the Complex between the Antibiotic Cerulenin and Its Target, β-Ketoacyl-Acyl Carrier Protein Synthase. *J. Biol. Chem.* **1999**, *274*, 6031 − 6034.

(75) Trajtenberg, F.; Altabe, S.; Larrieux, N.; Ficarra, F.; de Mendoza, D.; Buschiazzo, A.; Schujman, G. E. Structural insights into bacterial resistance to cerulenin. *FEBS J.* **2014**, 281, 2324 − 2338.

(76) Val, D.; Banu, G.; Seshadri, K.; Lindqvist, Y.; Dehesh, K. Re-engineering ketoacyl synthase specificity. *Structure* **2000**, *8*, 565 − 566.

(77) Morgan-Kiss, R. M.; Cronan, J. E. The *Lactococcus lactis* FabF fatty acid synthetic enzyme can functionally replace both the FabB and FabF proteins of *Escherichia coli* and the FabH protein of *Lactococcus lactis*. *Arch. Microbiol.* **2008**, *190*, 427 − 437.

(78) Zhu, L.; Cheng, J.; Luo, B.; Feng, S.; Lin, J.; Wang, S.; Cronan, J. E.; Wang, H. Functions of the *Clostridium acetobutylicium* FabF and FabZ proteins in unsaturated fatty acid biosynthesis. *BMC Microbiol.* **2009**, *9*, 119.

(79) Wang, H.; Cronan, J. E. Functional replacement of the FabA and FabB proteins of Escherichia coli fatty acid synthesis by *Enterococcus faecalis* FabZ and FabF homologues. *J. Biol. Chem.* **2004**, *279*, 34489 − 34495.

(80) Luo, Q.; Li, M.; Fu, H.; Meng, Q.; Gao, H. *Shewanella oneidensis* FabB: A β-ketoacyl-ACP Synthase That Works with C16:1-ACP. *Front. Microbiol.* **2016**, *7*, 327.

(81) Li, M.; Meng, Q.; Fu, H.; Luo, Q.; Gao, H. Suppression of fabB Mutation by fabF1 Is Mediated by Transcription Read-through in *Shewanella oneidensis*. *J. Bacteriol.* **2016**, *198*, 3060 − 3069.

(82) Witkowski, A.; Joshi, A. K.; Lindqvist, Y.; Smith, S. Conversion of a beta-ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine. *Biochemistry* **1999**, *38*, 11643 − 11650.

(83) Stunkard, L. M.; Dixon, A. D.; Huth, T. J.; Lohman, J. R. Sulfonate/Nitro Bearing Methylmalonyl-Thioester Isosteres Applied to Methylmalonyl-CoA Decarboxylase Structure-Function Studies. *J. Am. Chem. Soc.* **2019**, *141*, 5121 − 5124.

(84) Bräuer, A.; Zhou, Q.; Grammbitter, G. L. C.; Schmalhofer, M.; Ruhl, M.; Kaila, V. R. I.; Bode, H. B.; Groll, M. Structural snapshots of the minimal PKS system responsible for octaketide biosynthesis. *Nat. Chem.* **2020**, *12*, 755 − 763.

(85) Borgaro, J. G.; Chang, A.; Machutta, C. A.; Zhang, X.; Tonge, P. J. Substrate recognition by β -ketoacyl-ACP synthases. *Biochemistry* **2011**, *50*, 10678 − 10686.

(86) Kapur, S.; Lowry, B.; Yuzawa, S.; Kenthirapalan, S.; Chen, A. Y.; Cane, D. E.; Khosla, C. Reprogramming a module of the 6-deoxyerythronolide B synthase for iterative chain elongation. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4110 − 4115.

(87) Kapur, S.; Chen, A. Y.; Cane, D. E.; Khosla, C. Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 22066 − 22071.

(88) Jenner, M.; Frank, S.; Kampa, A.; Kohlhaas, C.; Pöplau, P.; Briggs, G. S.; Piel, J.; Oldham, N. J. Substrate Specificity in Ketosynthase Domains from trans-AT Polyketide Synthases. *Angew. Chem., Int. Ed.* **2013**, *52*, 1143 − 1147.

(89) Murphy, A. C.; Hong, H.; Vance, S.; Broadhurst, R. W.; Leadlay, P. F. Broadening substrate specificity of a chain-extending ketosynthase through a single active-site mutation. *Chem. Commun.* **2016**, *52*, 8373 − 8376.

(90) 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. *J. Biol. Chem.* **2001**, *276*, 6551 − 6559.

(91) Robbins, T.; Kapilivsky, J.; Cane, D. E.; Khosla, C. Roles of Conserved Active Site Residues in the Ketosynthase Domain of an Assembly Line Polyketide Synthase. *Biochemistry* **2016**, *55*, 4476 −4484.

(92) Liu, H.; Naismith, J. H. An Efficient One-Step Site-Directed Deletion, Insertion, Single and Multiple-Site Plasmid Mutagenesis Protocol. *BMC Biotechnol.* **2008**, *8* (1), 91.

(93) 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 Crystallogr. D. Biol. Crystallogr.* **2011**, *67* (Pt 4), 271–281.

(94) 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 Crystallogr. D. Biol. Crystallogr.* **2011**, *67* (Pt 4), 235–242.

(95) Adams, P. D.; Afonine, P. V; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, 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 Crystallogr. D. Biol. Crystallogr.* **2010**, *66* (Pt 2), 213–221.

(96) Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. Electronic Ligand Builder and Optimization Workbench (ELBOW): A Tool for Ligand Coordinate and Restraint Generation. *Acta Crystallogr. D. Biol. Crystallogr.* **2009**, *65* (Pt 10), 1074–1080.

(97) Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V; Vagin, A. A.; Murshudov, G. N. JLigand: A Graphical Tool for the CCP4 Template-Restraint Library. *Acta Crystallogr. D. Biol. Crystallogr.* **2012**, *68* (Pt 4), 431–440.

(98) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **1997**, *62* (21), 7512–7515.

(99) Still, W. C.; Kahn, M.; Mitra, A. Rapid Chromatographic Technique for Preparative Separations with Moderate Resolution. *J. Org. Chem.* **1978**, *43* (14), 2923–2925.

Chapter 4: Utilizing mechanism-based crosslinking probes to capture the *E. coli* ketosynthase, FabB, in different catalytic states.

Abstract

Ketosynthases (KS) catalyse the essential carbon-carbon bond forming reactions in fatty acid biosynthesis using a two-step, ping-pong reaction mechanism. In *E. coli*, there are two homodimeric elongating KSs, FabB and FabF, which generally have broad overlapping substrate specificity. However, FabB is essential for the biosynthesis of unsaturated fatty acids (UFAs), which are required for cell survival in the absence of exogenous sources. Additionally, FabB has reduced activity towards substrates longer than 14 carbons, whereas FabF is efficient in catalysing the elongation of saturated C14 and unsaturated C16:1 acyl substrates. In this study, we solved two crosslinked crystal structures of FabB in complex with ACP functionalized with long-chain fatty acid crosslinking probes that approximate states that occur during catalysis. Both structures possess asymmetric substrate binding pockets, which suggest a cooperative relationship towards long substrates between two FabB monomers. In addition, these structures capture an unusual rotamer of the gating residue, F392, potentially representing an intermediate state before substrate release. These structures align well with our understanding of the KS reaction mechanism and substrate profile, demonstrating the utility of mechanism-based crosslinking methods to inform complex transformations occurring during enzyme catalysis.

Introduction

Fatty acid synthases (FASs) carry out a series of iterative biochemical transformations to produce cellular fatty acids (FAs). Depending on the organization, FASs can be classified as type I multidomain megasynthases, or type II, which are comprised of discrete and monofunctional enzymes. Interestingly, Type II FASs, commonly found in bacteria and plant plastids, are considered to be the evolutionary progenitors of polyketide synthases, which are known to produce complex, biologically active molecules.[1,2] FASs iteratively extend and reduce two-carbon malonyl-CoA derived units. A Ketosynthase (KS) begins each round of chain elongation by condensing a malonyl-CoA derived substrate with a growing fatty acid to generate a β-keto intermediate via a decarboxylative Claisen condensation reaction. Subsequently, a ketoreductase, dehydratase, and enoylreductase catalyse three additional reactions to fully reduce the β-ketone to a methylene. Central to the elongation cycle is the acyl carrier protein (ACP), which is required to shuttle thioester-tethered substrates to each enzyme active site.[3] The thiol group of the ACP is derived from 4'-phosphopantetheine (PPant), which is post-translationally installed onto a conserved serine residue. The protein-protein interactions (PPIs) of ACP and its partner enzymes are crucial for efficient substrate delivery, especially in type II systems possessing discrete domains.[4]

KSs catalyse the critical carbon-carbon bond forming reactions at the beginning of each FAS elongation cycle.[5] The KS two-step ping-pong kinetic mechanism can be understood as two half-reactions consisting of a transacylation and a condensation step. During the transacylation half-reaction, the acyl-chain from acyl-ACP is transferred to the

catalytic cysteine of the *apo*-KS to generate an acyl-KS intermediate. During the condensation half-reaction, malonyl-ACP (mal-ACP) binds to the KS and is subsequently decarboxylated to form an enolate ion that attacks the acyl-KS thioester to form the final product, β-ketoacyl-ACP. In *Escherichia coli*, a model system for type II FAS, there are two elongating KSs, FabB and FabF, belonging to the ketoacyl-ACP-synthase I (KASI) and ketoacyl-ACP synthase II (KASII) families, respectively. The two KSs have partially overlapping function, but FabB specifically elongates the *cis*-3-decenoyl acid (C10:1), the first unsaturated fatty acid (UFA) intermediate in *E. coli*, and strains lacking FabB are UFA auxotrophes.[6–8] In terms of the chain length spectrum, FabB has reduced activity towards substrates longer than 12 carbons, whether saturated or unsaturated, whereas FabF still retains robust activy on C14 and C16:1 FA substrates, but not C16.[9]

Interactions between ACP and its partner enzymes (PEs) are transient and the visualization of PE-ACP complexes approximating states formed during catalysis have proven challenging. In recent years, the application of phosphopantetheine analogs outfitted with mechanism-based crosslinking probes mimicking fatty acid substrates have advanced our understanding of the KS reaction mechanism and substrate specificity in both FAS and PKS.[10–14] These probes contain selectively reactive warheads that form a covalent bond with the catalytic cysteine of KSs, facilitating the formation of a covalently trapped KS-ACP complex for visualization using X-ray crystallography. Two classes of KS warheads, chloroacrylate and α-bromo,[15,16] have been successfully applied in previous structural studies. Chloroacrylate crosslinked complexes, which crosslink at the C_3 position of the acyl chain, mimic an intermediate formed during the condensation half reaction as demonstrated

by previously published ACP-KS crosslinked structures from FAS and PKS.[10,12,13] The FabB-ACP α-bromo crosslinked complexes, which crosslink at the C_2 position of the acyl chain, show an active site arrangement similar to the transacylation intermediate. However, FabF-ACP complexes using α-bromo probes are trapped in an intermediate substrate-delivery state typified by large conformational rearrangements of two KS active site loops.[11,13] The two loops, named loop 1 and loop 2, form a double drawbridge-like gate[17] at the entrance of the active site pocket that regulate substrate processing. Current models suggest this gate is likely triggered by the interaction of loop 2 and acyl-ACP and the subsequent chain-flipping[18] of ACP-sequestered cargo into the active site. The gate-open conformation has only been captured in crosslinked crystal structures, showcasing the utility of mechanism-based crosslinking probes.

In this study, we develop and apply two crosslinking probes bearing long-chain fatty acid substrate mimetics to investigate the molecular basis of FabB substrate preference. The first probe, C14-chloroacrylate (C14Cl), mimics the final favoured condensation reaction catalyzed by FabB in the elongation cycle, while the second probe, C16:1-α-bromo (C16:1Br), represents the acylation reaction of FabB with an unfavoured substrate. Two crystal structures, FabB=C14-ACP and FabB=C16:1-ACP (where the '=' sign denotes a crosslinked complex) were determined using C14Cl and C16:1Br as crosslinkers, respectively. Both structures possess a back gate comprised of a glutamine and a buried glutamate residue that can alternate between conformations to expand or limit the size of the fatty acid binding pocket. The FabB=C14-ACP structure possesses a balanced substrate binding pocket residue arrangement able to accommodate the C12 chain. However, the

FabB=C16:1-ACP has distinct, assymetric substrate binding pocket conformations in each of the protomers. Importantly, the pocket conformations affect thioacyl coordination and alter the conformation of key catalytic residues related to KS gate function, suggesting negative cooperativity towards long-chain acyl substrates. Insights from these structures correlate well with FabB function and provide new insights into *E. coli* FAS and KS-directed metabolic engineering.

Materials and methods

Synthesis of crosslinking probes. C16:1Br,[13] (*E*)-C3Cl,[15] and (*E*)-C8Cl[16] were prepared according to previous literature protocols. (*E*)-C14Cl and (Z)-C14Cl were prepared according to Figure S4.5. See Supplementary Information for complete synthetic details.

Protein expression and purification. The N-terminal His-tagged FabB recombinant protein was expressed in *E. coli* BL21(DE3) and cultured in LB broth containing 50 mg/L kanamycin. Cells were grown at 37 ^{o}C until OD_{600} reached ~0.7 and induced by 0.5 mM IPTG for 4 h under the same temperature. Cell pellets were collected by centrifugation, lysed by sonication in lysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol), and underwent another centrifugation at 17,400 g for 1 h to clear the lysate. The His-tagged protein was pull down by Ni-NTA resin (0.5 mL resin used per litre growth), washed with wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol) and 20 mM buffered imidazole, and eluted with 250 mM buffered imidazole. His-tag cleavage by bovine thrombin (2 U per 1 mg FabB) was done while dialyzing against the dialysis buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10% glycerol, 0.5 mM TCEP) for 16 h at 6 ^{o}C. The resulting

solution was flown through the Ni-NTA resin to remove the un-cleaved protein, and further purified by FPLC with HiLoad Superdex 200 (GE Biosciences) size exclusion column. Pure tag-free FabB is collected and concentrated to 2-4 mg/mL using Amicon Ultra Centrifuge Filters (MilliporeSigma) with 10 kDa molecular weight cut off.

Native ACP recombinant protein was expressed in *E. coli* BL21(DE3) and cultured in LB broth containing 100 mg/L ampicillin. Cells were grown at 37 ^{o}C until OD_{600} reached ~0.7 and induced by 0.5 mM IPTG for 16 h at 18^{o}C. Cell pellets were collected by centrifugation, lysed by sonication in lysis buffer (50 mM Tris pH 7.4, 150 mM NaCl, 10% glycerol), and underwent another centrifugation at 17,400 g for 1 h to clear the lysate. Taking advantage of high ACP stability under 50% isopropanol solution, irrelevant proteins were precipitated by dripping in same volume of isopropanol into the lysate at the speed of 0.1 mL per second and removed by centrifugation at 17,400 g for 1 h. The supernatant was purified by FPLC via HiTrap Q HP anion exchange column with a linear gradient of 0 M to 1 M NaCl buffer. ACP eluted around 0.3 M NaCl and was collected to go through the second round of FPLC purification with HiLoad Superdex 75 PG (GE Biosciences) size exclusion column. Pure ACP was collected and concentrated using Amicon Ultra Centrifuge Filters (MilliporeSigma) with 3 kDa molecular weight cut off.

ACP modification and crosslinking reaction. ACP is overexpressed in a 1-to-1 ratio of the *apo* and *holo* form. Since the crosslinking probes contain the PPant moiety, the *apo*-ACP is the desired form and thus the PPant on *holo*-ACP is removed by enzymatic reaction of the ACP hydrolase, AcpH. In detail, 2-10 mg/mL of ACP mixture can be totally converted to *apo*-ACP with 0.01 mg/mL of AcpH present in the reaction buffer (50 mM Tris

pH 7.4, 150 mM NaCl, 10% glycerol, 10 mM $MgCl_2$, 5 mM $MnCl_2$, 0.25% DTT) under 25 ^oC in 16 h. The resulting *apo*-ACP is purified by FPLC (HiLoad Superdex 75 PG SEC) and concentrated (Amicon Ultra Centrifuge Filters, 3kDa MWCO) to 2-10 mg/mL. To load the crosslinking probes onto *apo*-ACP, they were first converted to CoA analogues through the activities of three CoA biosynthetic enzymes (CoaA, CoaD, and CoaE) and transferred onto the serine residue (S36) of *apo*-ACP by the phosphopantetheinyl transferase, Sfp. The whole modification and loading can be done in a one-pot chemoenzymatic reaction with the following condition: 1 mg/mL *apo*-ACP, 0.04 mg/mL Sfp, 0.01 mg/mL each CoA enzymes, and 0.2 mM crosslinking probe in reaction buffer (50 mM potassium phosphate pH 7.4, 12.5 mM $MgCl_2$, 1 mM DTT) at 37^oC for 16 h. The stock solution of the probes was prepared by dissolving the probes in DMSO to a final concentration of 50 mM for C14Cl and 25 mM for C16:1Br because of the lower solubility resulted from the longer acyl chain. The probe-loaded ACP, or *crypto*-ACP, was then purified by FPLC (HiLoad Superdex 75 PG SEC) and concentrated (Amicon Ultra Centrifuge Filters, 3kDa MWCO) to 2-5 mg/mL. All the different forms of ACP can be tracked by the conformationally sensitive urea PAGE.

The crosslinking reaction was set up by simply mixing *crypto*-ACP with FabB in 3:1 ratio in reaction buffer (50 mM Tris pH 7.4, 150 mM NaCl) at 37^oC for 16 h. The reaction was monitored by 12% SDS PAGE and purified by FPLC (HiLoad Superdex 200 PG SEC) with minimal buffer (12.5 mM Tris pH 7.4, 25 mM NaCl). The crosslink complex was concentrated to 5-8 mg/mL using Amicon Ultra Centrifuge Filters with 30 kDa MWCO and flash-frozen to store under -80 ^oC if not immediately used for crystallography.

Protein crystallography, data processing, and structure refinement. The crystals of all crosslinked complexes were grown by vapor diffusion in a cold room kept at 6 °C. 1 μL of crosslinked complex (8–10 $mg \cdot mL^{-1}$) was mixed with 1 μL of corresponding crystallographic condition and the mixture was placed inverted over 500 μL of the well solution (hanging-drop method). The FabB=C14Cl-ACP and FabB=C16:1-ACP crystals were grown in 18-24% PEG 8 K, 0.2 M $Mg(OAc)_2$, 0.1 M sodium cacodylate pH 6.5. These conditions produced square plates and required two weeks for complete growth of the crystals. X-ray diffraction data (Table S4.1) were collected at the Advanced Light Source synchrotron at Berkeley. Data were indexed using iMosflm[19] then processed and scaled using aimless from the CCP4 software suite.[20] Scaled reflection output data were used for molecular replacement and model building in PHENIX.[21] Initial phasing via molecular replacement was performed by searching for the full FabB=ACP diprotomer complex using 5KOF as a search model. Subsequent refinement and model building were performed in PHENIX and Coot respectively. The parameter files for the covalently bonded 4′-phosphopantetheine were generated using Jligand.[22] Manually programmed parameter restraints were used to create the associated covalent bonds between 4′-phosphopantetheine to Ser36 and Cys163 during refinement.

Results and discussion

Rationale behind crosslinker choice and the KS stereospecificity for chloroacrylate probes. To understand the molecular basis of FabB substrate preference toward long chain FAs, we developed a 14-chain length chloroacrylate crosslinker, C14Cl,

that we hypothesized would mimic the condensation step between the longest favoured FabB substrate, C12, and malonyl-ACP. Probes containing two different configurations, *E* and *Z*-form, respectively, of the acrylate double bond were synthesized and tested (Figure S4.1). Only the *Z*-form shows crosslinking activity on FabB and the same applies to the other two tested KSs, FabF and actinorhodin KS (actKS), a type II polyketide synthase.[23] Given that there are published crosslinked structures of FabB-ACP and FabF-ACP utilizing (*E*)-C3Cl and (*E*)-C8Cl, respectively, we hypothesize that while *Z*-form is the preferred configuration, the selectivity only happens at longer chain length due to the restrictions in the substrate pocket. To test this hypothesis, we performed crosslinking assays using (*E*)-C8Cl and actKS, which has a narrower pocket more likely clash with the acyl substrate. Results from these assays showed no crosslinking activity. However, (*E*)-C3Cl was able to crosslink with our entire panel of KSs, indicating that actKS is active towards unsubstituted *E*-form chloroacrylate crosslinkers, generally. As a result, the data suggest stereospecificity of KSs toward the *Z*-form long-chain chloroacrylate crosslinker. After confirming the preferred stereochemistry, we loaded ACP with (*Z*)-C14Cl and performed large-scale crosslinking to obtain FabB=C14-ACP for structural analysis.

To explore mechanisms governing FabB substrate selectivity, we also obtained a crosslinked complex using the C16:1-α-bromo crosslinker, or C16:1Br (Figure S4.1), which would mimic the transacylation step of an unfavoured substrate, namely C16:1-ACP. Importantly, this substrate is readily accepted by FabF *in vitro* and *in vivo*, and its elongation to C18:1 is critical for *E. coli's* rapid homeoviscous adaptive response to regulate membrane fluidity due to changes in environmental temperature.[24,25] Previously, we utilized C16:1-

ACP to investigate the molecular mechanisms governing FabF's unique capacity to extend C16:1 to C18:1.[13] A FabB=ACP structure using the same crosslinker could potentially inform how FabB discriminates against longer acyl substrates.

FabB-ACP trapped in different catalytic states by the crosslinkers. The FabB=C14-ACP complex was crystallized and diffracted to a nominal resolution of 1.7-Å. The asymmetric unit contains the biologically relevant diprotomer structure (FabB-ACP)2 and posseses well-defined crosslinker density (Figure 4.1). The active site cysteine, C163, is covalently attached to the C3 position of the prosthetic acyl chain, which has the same carbon count as the tetrahedral intermediate formed during the condensation half reaction. The carbonyl group of the acyl chain coordinates with the two active site catalytic histidine residues, and this orientation has been observed in all other KS-ACP crosslink structures using the chloroacrylate crosslinkers.[10,12,13] Interestingly, the double bond of the acrylate is in the same trans configuration as observed in other structures despite the use of E form crosslinkers instead. This suggests a stereoselective addition-elimination reaction mechanism governed by the active site arrangement during reformation of the double bond. Overall, the two protomers of the structure align with an RMSD of 0.872 Å and show no obvious structural differences.

The FabB=C16:1-ACP complex was crystallized and diffracted to 2.20-Å resolution and the asymmetric unit contains the same diprotomer arrangement (Figure 4.1). The protomer that contains the chain A monomer of FabB, referred to as protomer A, has the acyl chain carbonyl positioned in the oxyanion hole formed by the backbone amides of C163 and F392. This coordination is essential to the transacylation half reaction.[26,27] The other

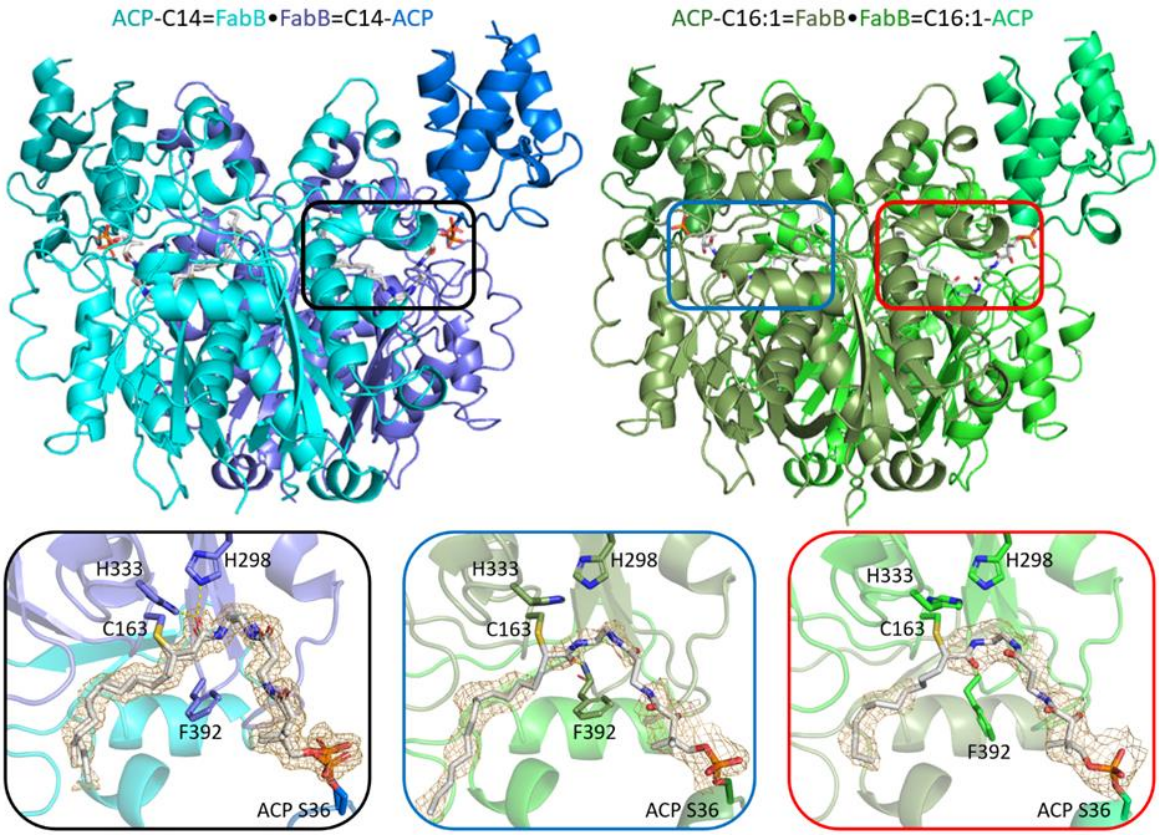


Figure 4.1: FabB trapped in different states by crosslinking with ACP. FabB=C14-ACP, blue structure on the left, possesses two identical active sites with the crosslinker solved in alternative conformations. One of its active sites is enlarged in the black frame, which depicts a condensation half-reaction arrangement. FabB=C16:1-ACP structure (green) on the right has two distinct protomers. Protomer A (blue frame) depicts a transacylation half-reaction arrangement while protomer B (red frame) features partially solved aliphatic chain and a catalytic incompetent F392 rotamer, potentially mimicking a substrate release state. The mesh surface depicts the Fo-Fc polder omit map calculated by omitting the crosslinker.

protomer of the complex, referred to as protomer B, captures a unique active site arrangement where the acyl carbonyl does not form polar contacts with any residue, but positions itself so that it repels F392 away from the active site, disassembling the oxyanion hole. Importantly, F392 is a crucial gating element and its rotamers regulate malonyl substrate entry.[28–30] Furthermore, it also plays a critical role in the KS gating mechanism during substrate delivery and undergoes an ~8 Å residue shift (Cα to Cα) when transitioning to the gate-open conformation. Due to the unfavourable substrate conformation, protomer B possesses a non-catalytic state (oxyanion hole not present) with an altered F392 gating residue position, likely representing a conformation primed for substrate release. An overlay of the two protomers with our recently published gate-open FabF=C16:1-ACP (PDB ID: 7L4E) structure suggests a potential trajectory for gate opening starting from the gate-closed protomer A to the disrupted protomer B and finally the gate-open FabF=ACP structure (Figure 4.2). The proposed transition features a 90^{o} rotation of the F392 side chain, likely driven by the acyl carbonyl flipping out from the oxyanion hole, which places the side chain underneath the PPant arm to facilitate the subsequent transition. The loop then undergoes a large conformational change to the gate-open form, with G391 and G394 as hinges, creating space for the substrate dissociation. During the gate transition, loop 2 undergoes a conformational change in concert with loop 1, as observed in FabF. The adjacent loop 2 residue, V270, abuts F392 and would sterically prohibit a transition to the gate open conformation. V270 from protomer B has a notably higher B-factor (135.5) than that of protomer A's V270 (93.1), indicating a less ordered loop 2 potentially primed for transition to the open state.

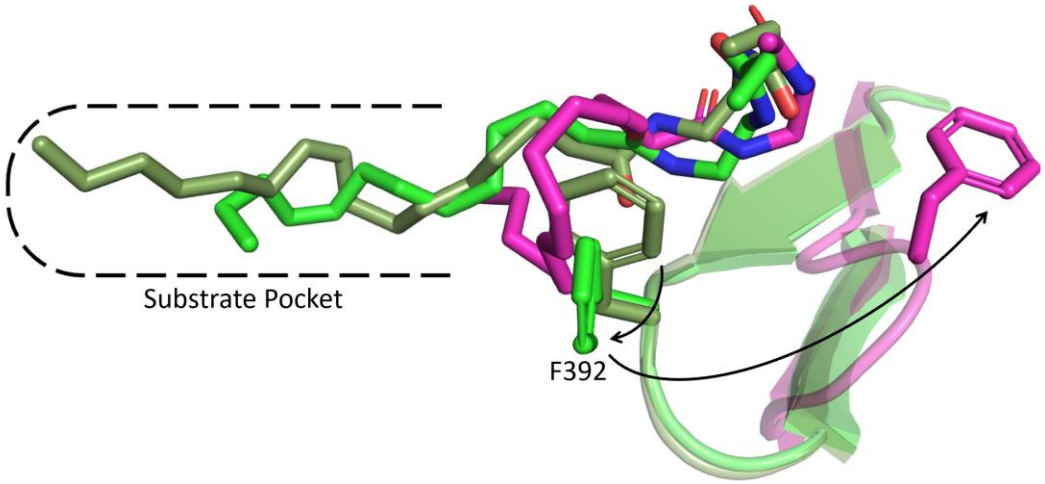


Figure 4.2: A potential trajectory for substrate de-sequestration is illustrated by overlaying three active sites from two crosslinked KS-ACP structures utilizing the same C16:1Br crosslinker. The arrow indicates the transition of the gating residue phenylalanine from protomer A (dark green, gate closed) to protomer B (light green), and finally the gate open form (magenta, FabF-ACP structure, PDB ID: 7L4E).

Asymmetric pockets of the FabB homodimer suggest negative cooperativity. The substrate pocket of FabB extends from the active site cysteine C163 to the homodimer interface. At the interface, E200 and Q113' (The apostrophe denotes residue from the other protomer) sit near the bottom of the pocket forming side chain polar contacts that determine the size of the binding pocket. However, these residues can also adopt different rotamers where the side chains turn away from each other in opposite directions, breaking the interaction and resulting in a pocket expansion. This double side chain rotation can be classified as a "swinging door enzyme gate" (Gora *et al*., 2013), and will hereafter be referred to as the "E-Q gate". In the FabB=C14-ACP structure, where a 12-carbon aliphatic chain extends into the binding pocket, two sets of alternative conformations were solved for both the acyl chains and E-Q gates. The E-Q gates from both pockets can be modelled in the open (expanded pocket) or closed (normal pocket) states. The open E-Q gate conformation places the Q113' residue away from the pocket towards the dimer interface and expands the size of the binding pocket when compared to the closed state. However, the modelled C12 substrate conformations are readily accommodated by both E-Q gate conformations (Figure 4.3A). Interestingly, there is only enough space at the dimer interface to accommodate Q113 from a single protomer, indicating that the two open E-Q gate conformations cannot coexist.

FabB's asymmetric pocket features are further illustrated in the FabB=C16:1-ACP structure, in which protomer A has an expanded pocket and protomer B does not (Figure 4.3B). In protomer A, the 16-carbon long aliphatic chain penetrates through the open E-Q gate with the last four carbons flanked by the side chains of the two gating residues. The expanded pocket extends to the α-helix (residue 110'-121') that contains Q113'. Protomer

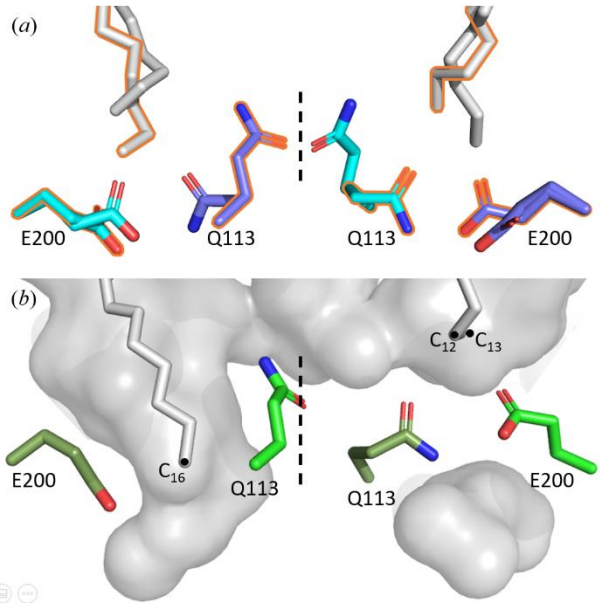


Figure 4.3: Asymmetric pockets and the E-Q gate. Residues with the same colour belong to the same FabB monomer. (*a*) FabB=C14-ACP possesses two sets (orange frame and no frame) of alternative conformations of the acyl chain and the E-Q gate. (*b*) FabB=C16:1-ACP has an expanded pocket (left, protomer A, open E-Q gate) and a normal pocket (right, protomer B, closed E-Q gate).

B, on the other hand, has a closed E-Q gate that distorts the oversized C16 chain, which prohibited the modelling of more than 13 carbons into the available density. As a result, a FabB binding pocket in the E-Q gate-closed conformation cannot comfortably accommodate an acyl chain longer than 12 carbons. Importantly, the sterically occluded chain in protomer B likely contributes to the distorted active site arrangement noted above. These observations suggest the inability to accommodate the acyl substrate results in improper positioning of the thioester moiety for catalysis, prohibiting the transfer of unfavourable substrates to the KS cysteine residue during the transacylation half-reaction.

FabB shows competent activity elongating FAs up to 14 carbons in length and is supported by the FabB=C14-ACP structural data as both protomers can accommodate 12 carbon substrates regardless of the E-Q gate conformation. However, only the expanded pocket can readily accommodate acyl substrates beyond 12 carbons in length. These results explain the low, but existing, FabB activity towards long-chain acyl substrates *in vitro* and *in vivo*.[9,31] Furthermore, since only one E-Q gate can occupy the open conformation, these data suggest a negative cooperativity mechanism toward long-chain acyl substrates, as observed in the FabB=C16:1-ACP structure where protomer B "struggles" to accept the substrate. Therefore, the observed pocket and E-Q gate mechanism align well with the reported activity of FabB.

The implication of the side pocket in KS mechanism. In FAS KSs, there is a "side pocket' located adjacent to the catalytic active site and opposite side of the substrate binding pocket. The PPant binding tunnel, acyl substrate pocket, and the side pocket form a "τ" shape space (Figure S4.2). The side pocket has a surface area of 180-Å^2 and a volume of 95-Å^3,

which is larger than the acyl substrate pocket that has a surface area of 145-Å^2 and a volume of 58-Å^3, as calculated by CASTp.[32] Despite its proximity to the active site, the side pocket has only been examined in passing in previous studies and has not been assigned a role in catalysis. The recently discovered gating loops reveal that the side pocket, which is partially formed by loop 1 and loop 2 when in the gate-closed conformation, is required to accommodate the loop 1 phenylalanine gating residue in the gate-open conformation (Figure S4.3). Introducing bulky residues by site-directed mutagenesis that fill in the available side-pocket space significantly impair FabF activity by restricting access to the gate-open conformation.[13]

Interestingly, the high resolution FabB=C14-ACP structure allows us to unambiguously assign 10 water molecules in the side pocket, forming a water network that connects loop 1, loop 2, and the PPant moiety (Figure 4.4). In the proximity of the catalytic triad, two water molecules, WAT1 and WAT2 (Figure S4.4), form a water network between T300, D306, and E309, which are absolutely conserved across 461 FabB sequences. Previous work suggests that a nucleophilic water molecule, activated by the catalytic histidine residue (H298 in FabB), is required for malonyl-ACP decarboxylation to form the enolate ion that subsequently attacks acyl-KS intermediate. In this model, a hydroxide ion attacks the malonyl carboxylate moiety to form bicarbonate as the final side product instead of CO_2.[28,33] The observed water network in the side pocket creates a hydrophilic and water-rich site next to the reaction chamber. This network might assist the decarboxylation of malonyl-ACP by providing a water molecule needed for bicarbonate formation as well as accommodating the bicarbonate by-product. Such an observation supports the bicarbonate-

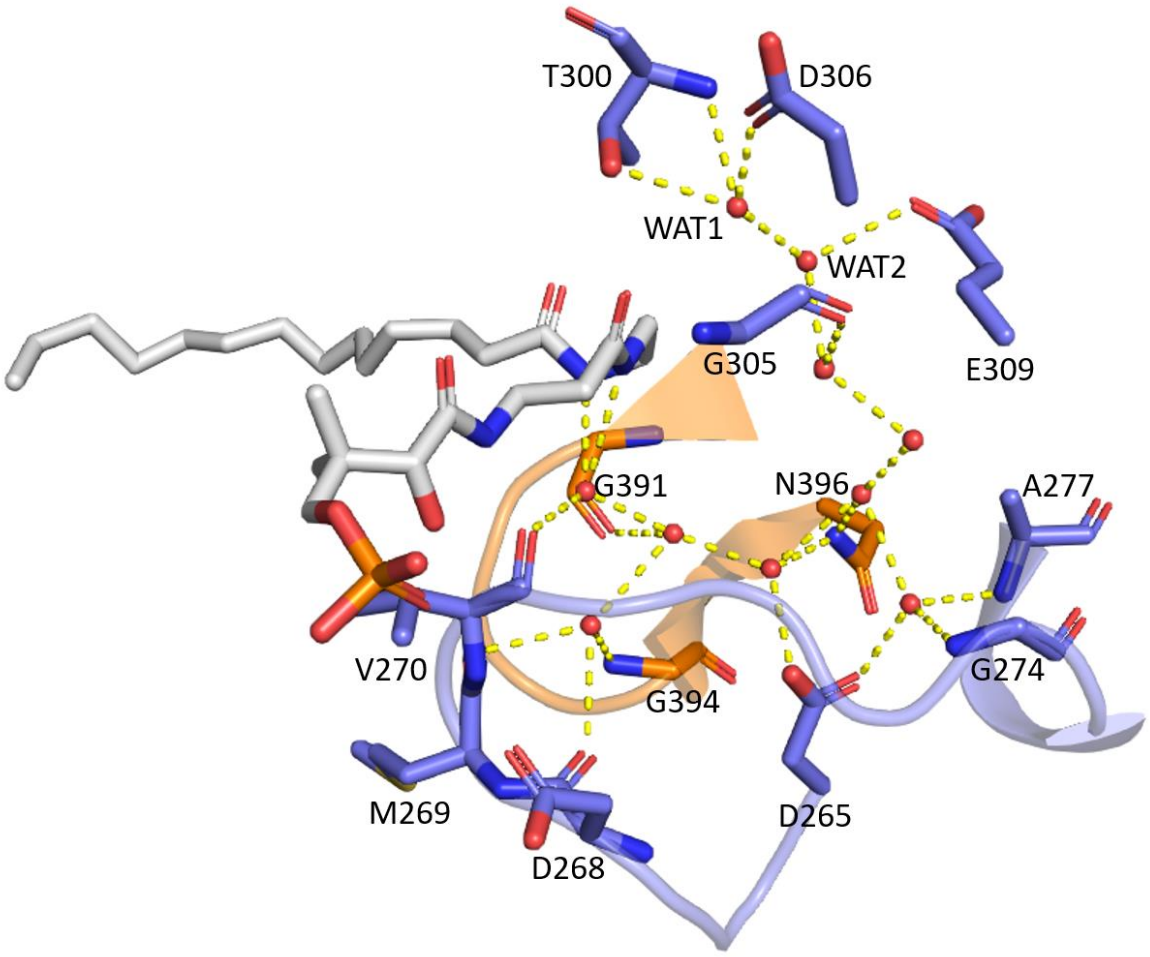


Figure 4.4: Ten water molecules in the side pocket form a water network in between highly conserved residues and the PPant arm. Cartoon representation is shown for the gating loop 1 (orange) and loop 2 (blue). All polar contacts (yellow dashed line) are within 3.2-Å.

forming mechanism, as opposed to the mechanism that releases carbon dioxide;[34] however, the KS decarboxylation mechanism is an ongoing debate and may differ between different KS families. More biochemical and structural studies are required to definitively address this long-standing question.[5]

Conclusion

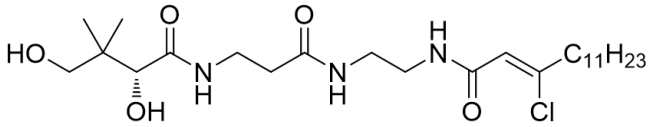
Since its discovery,[6] *E. coli* FabB has been studied for more than 50 years with at least 17 structures published on Protein Data Bank in various forms (*apo*, mutants, substrate-bound, inhibitor-bound, and ACP-crosslinked). Despite the extensive knowledge we have regarding this enzyme, some fundamental questions, such as substrate specificity and the decarboxylation mechanism, have yet to be fully addressed. In this study, we leveraged our understanding of the KS reaction mechanism and applied substrate-functionalized mechanism-based crosslinkers to elucidate two FabB=ACP crosslinked complexes in both favoured and unfavoured catalytic states. The C14Cl crosslinker captures FabB in on-pathway state representing the condensation reaction with a favoured substrate, C12, and malonyl-ACP, while complex trapped using the C16:1Br crosslinker represents the acylation reaction with an unfavoured long-chain acyl-ACP substrate. Analysis of these structures led to the identification of an E-Q gate at the back of the substrate binding pocket that enforces asymmetry between the two protomers. A 12-carbon substrate can be readily accommodated in both pockets. However, the E-Q gate must transition to the open conformation to accommodate longer fatty acid substrates to appropriately position the thioester moiety for the transacylation reaction. The asymmetric nature of the E-Q gate suggests negative

cooperativity of FabB towards long chain substrates, which may explain the reduced activity of FabB towards C14 and C16:1, two substrates readily extended by FabF. In addition to these mechanistic insights, we have characterized a water-rich side pocket that plays an important role in the dual-loop KS gating mechanism and supports the bicarbonate-forming decarboxylation mechanism. Results from these studies provide important insights into *E. coli* FAS and broadly inform mechanisms governing KS-mediated catalysis.

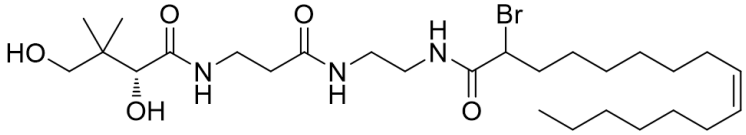
Acknowledgement

Chapter 4, entitled Utilizing Mechanism-Based Crosslinking Probes to Capture the *E. coli* Ketosynthase, FabB, in Different Catalytic States, in full, is a reprint of the material as it appears in: Chen A*, Mindrebo JT*, Davis TD, Kim WE, Katsuyama Y, Jiang Z, Ohnishi Y, Noel JP, Burkart MD. "Utilizing Mechanism-Based Crosslinking Probes to Capture the *E. coli* Ketosynthase, FabB, in Different Catalytic States", Acta Crystallographica Section D, submission in progress. The dissertation author is the primary co-author of this manuscript along with Dr. Jefferey T. Mindrebo.

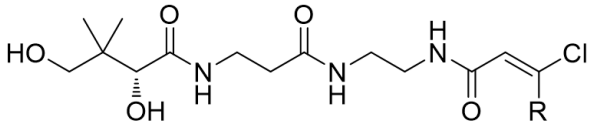
Supplementary Information



(*Z*)-C14Cl



C16:1Br



(*E*)-C3Cl: R = H
(*E*)-C8Cl: R = C_5H_{11}
(*E*)-C14Cl: R = $C_{11}H_{23}$

Figure S4.1: Crosslinkers used in this study. The top two crosslinkers were used to obtain the two FabB-ACP crosslinked structures while the tree *E* form crosslinkers were used for crosslinking activity study.

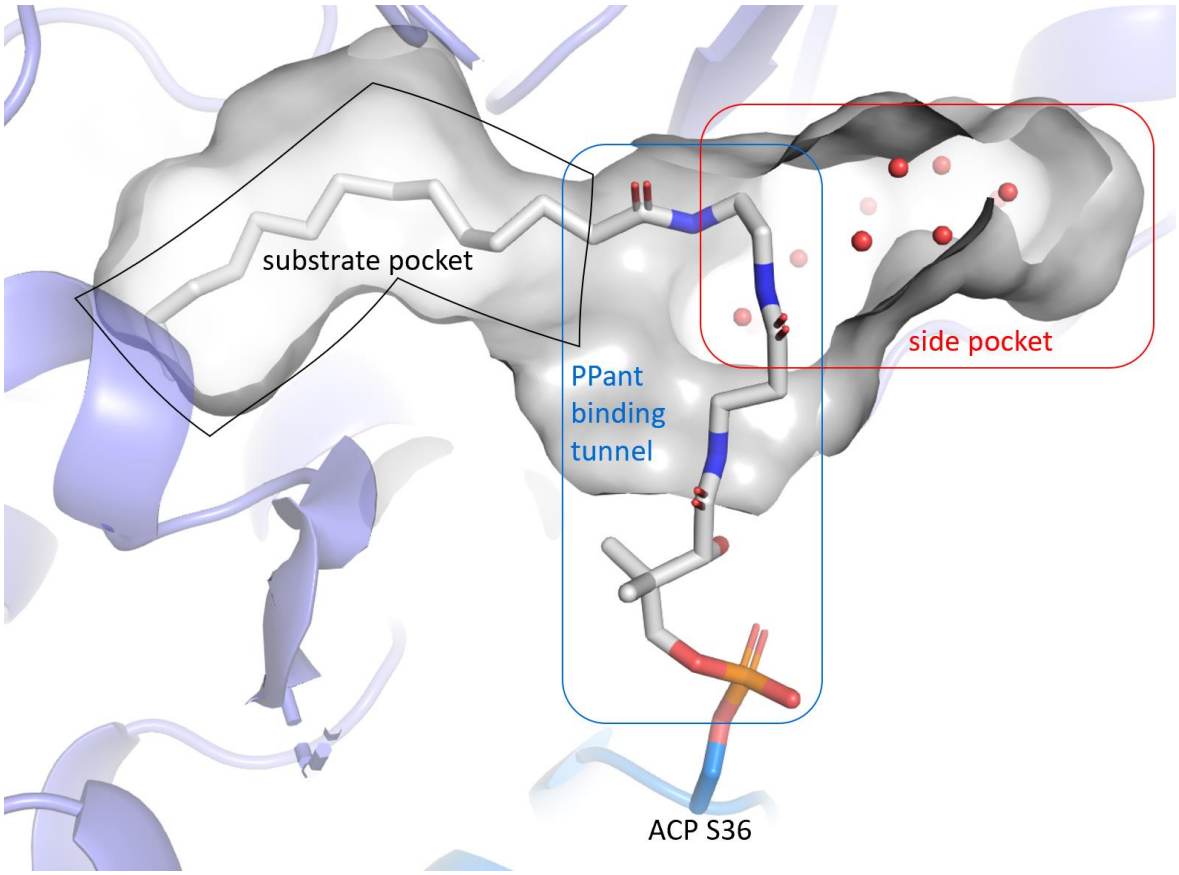


Figure S4.2: PPant binding tunnel, substrate pocket, and the side pocket form a "τ" shape space inside FabB.

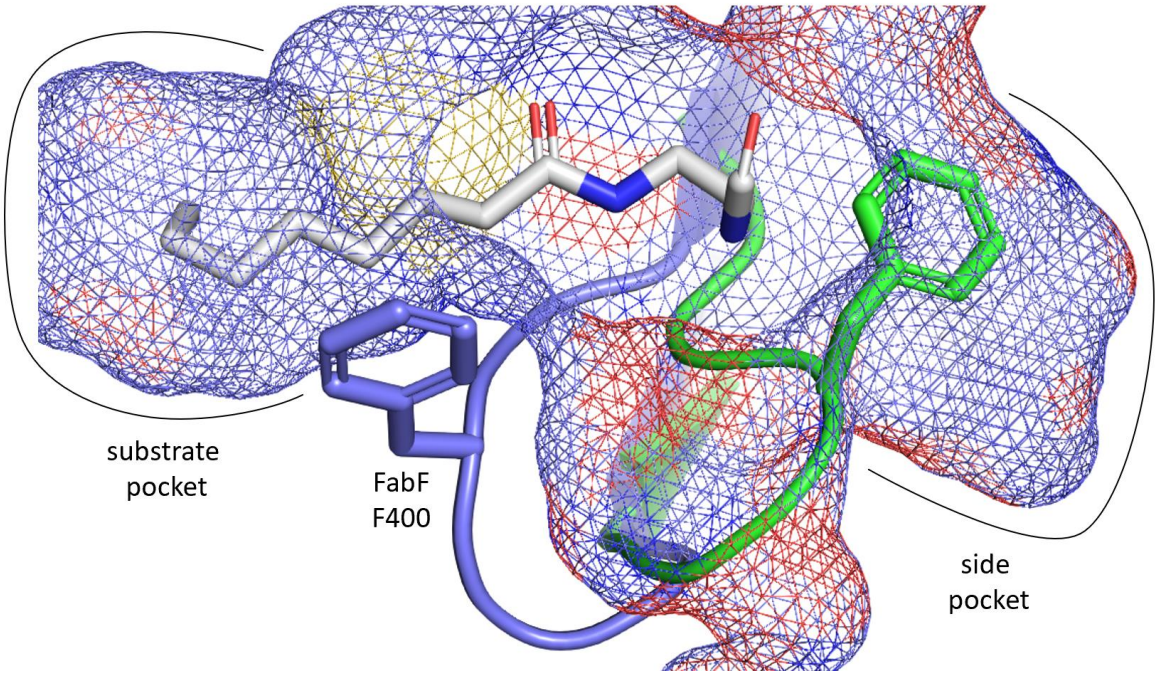


Figure S4.3: Overlay of the open form gating loop 1 (green, PDB ID: 6OKG) with the structure in closed form (blue, PDB ID: 7L4L). The side pocket accommodates the open loop.

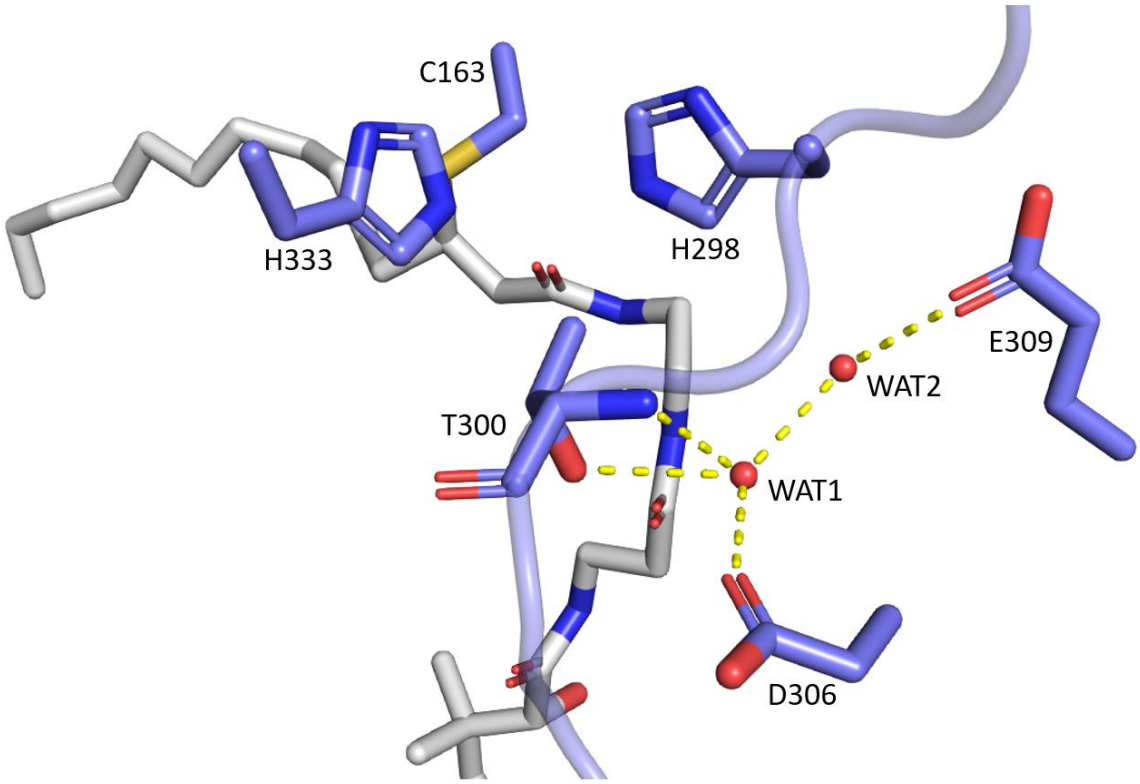


Figure S4.4: The two water molecules, WAT1 and WAT2, respectively, in proximity of the catalytic histidine residues. All residues shown are 100 % conserved over 461 FabB sequences aligned.

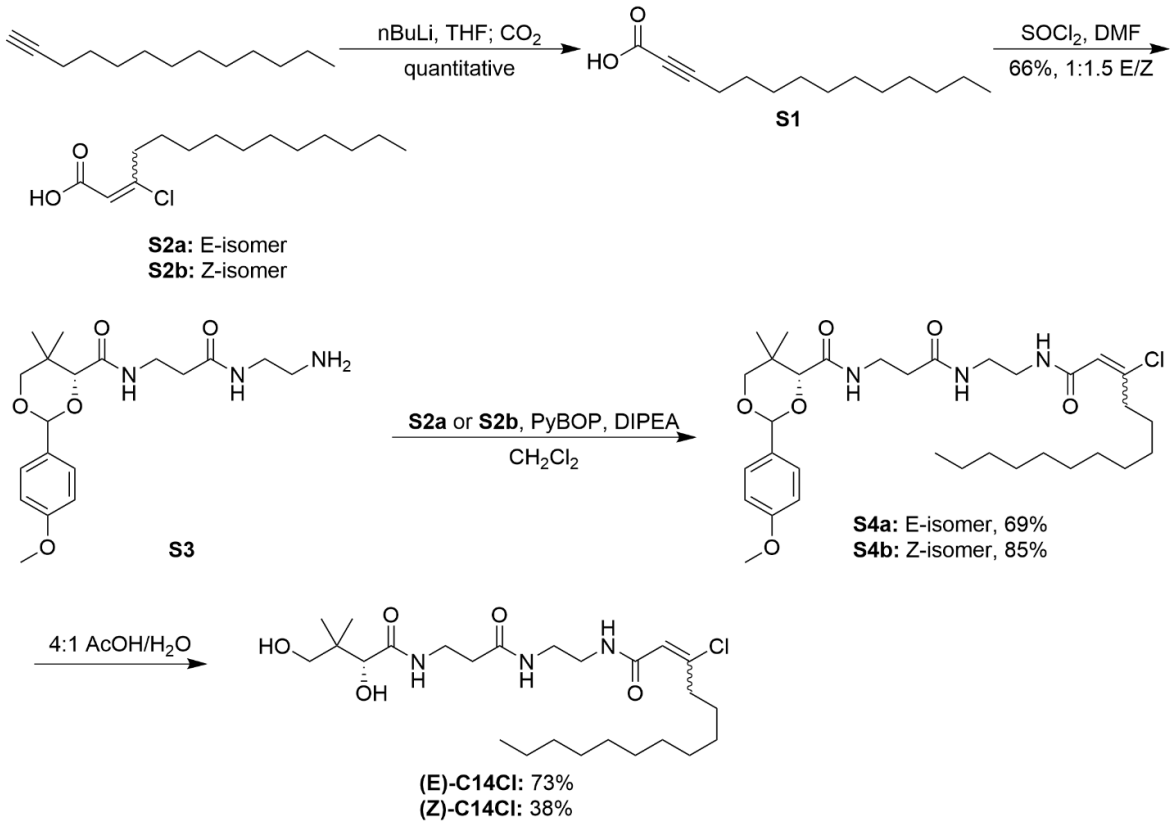


Figure S4.5: Synthesis of chloroacrylamide crosslinking probes. PyBOP= benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate; DIPEA=*N*,*N*-diisopropylethylamine; AcOH=acetic acid.

Table S4.1: Table 1 x-ray crystallography data collection and refinement statistics.

Structure (PDB ID)	FabB=C14-ACP (7SQI)	FabB=C16:1-ACP (7SZ9)
Wavelength (Å)	1	1
Resolution range (Å)	36.86 - 1.7 (1.761 - 1.7)	47.9 - 2.2 (2.279 - 2.2)
Space group	P 21 21 21	P 1 21 1
Unit cell (Å, º)	58.99 112.38 141.64 90 90 90	59.4505 103.58 79.2802 90 108.28 90
Unique reflections	104222 (10240)	38076 (2494)
Reflections used in refinement	103901 (10155)	37907 (2488)
Reflections used for R-free	5103 (466)	1808 (122)
R-work	0.1608 (0.3039)	0.2084 (0.3117)
R-free	0.1981 (0.3281)	0.2553 (0.3461)
Protein residues	965	957
Solvent	1069	165
RMS (bonds)	0.008	0.002
RMS (angles)	0.97	0.47
Ramachandran favoured (%)	96.33	94.84
Ramachandran allowed (%)	3.36	4.43
Ramachandran outliers (%)	0.31	0.74
Rotamer outliers (%)	0.80	1.46
Clashscore	3.80	4.84
Average B-factor	23.68	57.06

Supplementary Note S4.1: Synthesis of crosslinking probes

Chemical reagents were purchased from Acros, Fluka, Sigma-Aldrich, or TCI. Deuterated NMR solvents were purchased from Cambridge Isotope Laboratories. All reactions were conducted with vigorously dried anhydrous solvents that were obtained by passing through a solvent column exposed of activated A2 alumina. All reactions were performed under positive pressure of argon in flame-dried glassware sealed with septa and stirred with Teflon coated stir bars using an IKAMAG TCT-basic mechanical stirrer (IKA GmbH). Analytical Thin Layer Chromatography (TLC) was performed on Silica Gel 60 F254 precoated glass plates (EM Sciences). Visualization was achieved with UV light and/or appropriate stain (I_2 on SiO_2, $KMnO_4$, bromocresol green, dinitrophenylhydrazine, ninhydrin, or ceric ammonium molybdate). Flash column chromatography was carried out with Geduran Silica Gel 60 (40–63 mesh) from EM Biosciences. Yield and characterization data correspond to isolated, chromatographically, and spectroscopically homogeneous materials. ^{1}H NMR spectra were recorded on Varian Mercury 400, Varian Mercury Plus 400, or JEOL ECA500 spectrometers. ^{13}C NMR spectra were recorded at 100 MHz on Varian Mercury 400 or Varian Mercury Plus 400 spectrometers. Chemical shifts for ^{1}H NMR and ^{13}C NMR analyses were referenced to the reported values[35] using the signal from the residual solvent for ^{1}H spectra, or to the ^{13}C signal from the deuterated solvent. Chemical shift δ values for the ^{1}H and ^{13}C spectra are reported in parts per millions (ppm) relative to these referenced values, and multiplicities are abbreviated as s=singlet, d=doublet, t=triplet, q-quartet, m=multiplet, b=broad. All ^{13}C NMR spectra were recorded with complete proton decoupling. FID files were processed using MestreNova 10.0 (MestreLab Research).

Electrospray ionization (ESI) mass spectrometric analyses were preformed using a ThermoFinnigan LCQ Deca spectrometer. Spectral data and procedures are provided for all new compounds and copies of spectra have been provided.

2-tetradecynoic acid (S1). In a 50 mL round-bottomed flask, 1-tridecyne (0.12 mL, 0.52 mmol, 1.0 equiv.) and 5 mL THF were added. The vessel was cooled to 0 °C before the slow, dropwise addition of nBuLi (1.6 M in hexanes, 0.40 mL, 0.64 mmol, 1.1 equiv.). The reaction was stirred for 2 hours at 0 °C before flushing with carbon dioxide gas. After stirring for 3 hours at 0 °C, the reaction was warmed to room temperature (20 °C) and stirred overnight for 17 hours. The reaction was quenched with saturated aqueous NH_4Cl (10 mL) and extracted with ethyl acetate (4x25 mL). The combined organic extracts were dried ($MgSO_4$), filtered, and concentrated by rotary evaporation. Purification by silica flash chromatography (9:1 hexanes/ethyl acetate → 3:2 hexanes/ethyl acetate + 1% acetic acid) afforded carboxylic acid **S1** (96.7 mg, 83%, white solid).

TLC: R_f 0.21 (4:1 hexanes/ethyl acetate + 1% acetic acid). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ10.98 (s, 1H), 2.34 (t, *J* = 7.0 Hz, 1H), 1.81– 1.55 (m, 2H), 1.39 (s, 3H), 1.25 (s, 14H), 0.87 (t, *J* = 5.6 Hz, 3H).**^{13}C-NMR** (100 MHz, $CDCl_3$): δ 158.83, 92.92, 72.77, 32.03, 29.71, 29.53, 29.13, 28.94, 27.51, 22.81, 19.01, 18.88, 18.74, 14.19.

3-chloro-2-tetradecenoic acid (S2). In a 10 mL pear-shaped flask, 2-tetradecynoic acid **S1** (126.1 mg, 0.5621 mmol, 1.0 equiv.) and 5 mL DMF were added. The vessel was cooled to 0 °C before the addition of $SOCl_2$ (48.9 μL, 0.675 mmol, 1.2 equiv.). After stirring for 1 hour at 0 °C, the reaction was warmed to room temperature (20 °C). After 3 hours at room temperature, additional $SOCl_2$ was added (20 μL) and stirring was continued overnight.

After 15 hours, additional $SOCl_2$ was added (48.9 μL). After 2 hours, the reaction was concentrated by rotary evaporation, and the resulting liquid was poured over 12 grams of ice and extracted with diethyl ether (4x25 mL). The combined organic extracts were dried ($MgSO_4$), filtered, and concentrated by rotary evaporation. Purification by silica flash chromatography (9:1 hexanes/diethyl ether → 4:1 hexanes/diethyl ether) afforded (*E*)-3-chloro-2-tetradecenoic acid **S2a** (39.2 mg, yellow solid) and (*E*)-3-chloro-2-tetradecenoic acid **S2a** (57.2 mg, brown solid)

(*E*)-3-chloro-2-tetradecenoic acid (S2a). TLC: R_f 0.61 (1:1 hexanes/diethyl ether). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 6.09 (s, 1H), 2.98 (t, 2H), 1.63 (m, 2H), 1.26 (s, 14H), 0.95 – 0.80 (t, 3H).**^{13}C-NMR** (100 MHz, $CDCl_3$): δ 170.11, 161.06, 118.44, 36.11, 32.08, 29.78, 29.62, 29.51, 29.45, 28.90, 27.83, 22.86, 14.27.

(*Z*)-3-chloro-2-tetradecenoic acid (S2b). TLC: R_f 0.45 (3:2 hexanes/diethyl ether). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 6.05 (s, 1H), 2.46 (t, J = 7.4 Hz, 1H), 1.63 (m, 2H), 1.26 (s, 15H), 0.88 (t, 3H).**^{13}C-NMR** (100 MHz, $CDCl_3$): δ 169.52, 154.24, 115.60, 41.74, 32.05, 29.73, 29.58, 29.47, 29.38, 28.67, 27.35, 22.83, 14.25.

(4*R*)-*N*-(3-((2-((*E*)-3-chlorotetradec-2-enamido)ethyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (S4a). In a 10 mL pear-shaped flask, (4*R*)-*N*-(3-((2-aminoethyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide **S3** (57.0 mg, 0.105 mmol, 1.0 equiv.),[36] (*E*)-3-chloro-2-tetradecenoic acid **S2a** (39.2 mg, 0.1503 mmol, 1.0 equiv.), and 5 mL CH_2Cl_2 were added. To the solution was added PyBOP (94.3 mg, 0.181 mmol, 1.7 equiv.) and DIPEA (78.5 μL, 0.4509 mmol, 3.0 equiv.). After 20 hours, the volatiles were removed by rotary evaporation.

Purification by silica flash chromatography (3:1 hexanes/ethyl acetate → 39:1 ethyl acetate/methanol) afforded **S4a** (64.8 mg, 69%, white solid).

TLC: R_f 0.31 (ethyl acetate). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 7.41 (d, *J* = 8.2 Hz, 2H), 7.02 (s, 1H), 6.89 (d, *J* = 8.3 Hz, 3H), 6.75 (s, 1H), 5.97 (s, 1H), 5.44 (s, 1H), 4.06 (s, 1H), 3.79 (s, 3H), 3.65 (q, *J* = 11.4 Hz, 2H), 3.50 (m, 2H), 3.31 (s, 3H), 2.97 (t, *J* = 7.0 Hz, 2H), 2.38 (bs, 2H), 1.58 (bs, 2H), 1.24 (s, 15H), 1.05 (d, *J* = 6.4 Hz, 6H), 0.84 (t, *J* = 6.3 Hz, 3H). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ 172.05, 169.96, 164.95, 160.37, 153.67, 130.15, 127.66, 120.97, 113.83, 101.46, 83.89, 55.42, 39.84, 39.59, 36.36, 35.42, 35.05, 33.17, 32.01, 29.75, 29.73, 29.65, 29.54, 29.44, 29.05, 27.87, 22.78, 21.94, 19.21, 14.23.

(4*R*)-*N*-(3-((2-((*Z*)-3-chlorotetradec-2-enamido)ethyl)amino)-3-oxopropyl)-2-(4-methoxyphenyl)-5,5-dimethyl-1,3-dioxane-4-carboxamide (S4b). Prepared as described for **S4a** from **S3**[36] and **S2b** to afford **S4b** (116.3 mg, 85%, white solid).

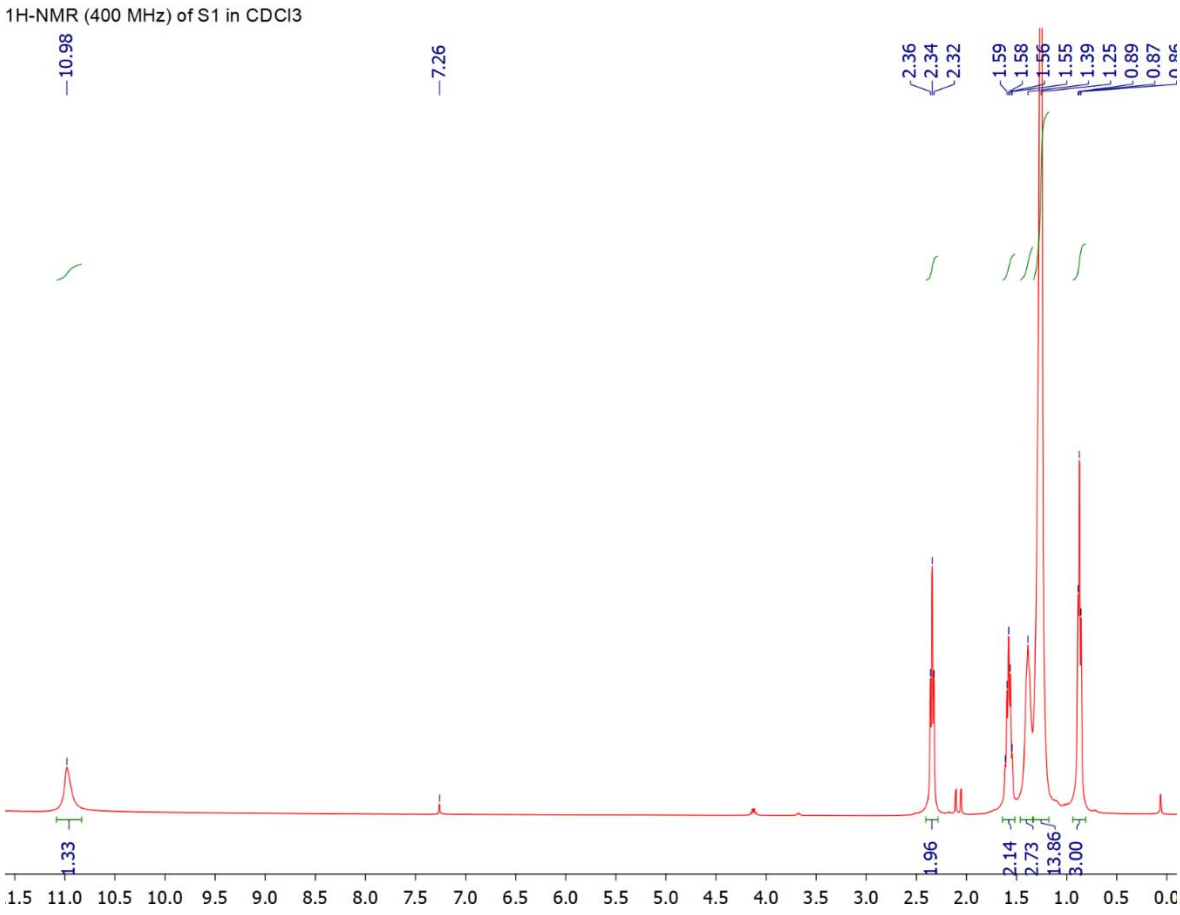
TLC: R_f 0.11 (ethyl acetate). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 7.39 (d, *J* = 8.7 Hz, 2H), 7.06 (t, *J* = 6.2 Hz, 1H), 6.95 (s, 2H), 6.87 (d, *J* = 8.7 Hz, 2H), 5.96 (s, 1H), 5.43 (s, 1H), 4.04 (s, 1H), 3.77 (s, 3H), 3.67 (m, 3H), 3.48 (m, 2H), 3.33 (s, 4H), 3.10 (m, 1H), 2.36 (m, 3H), 1.53 (s, 2H), 1.37 (q, *J* = 6.7, 6.1 Hz, 7H), 1.23 (s, 15H), 1.06 (d, *J* = 10.2 Hz, 6H), 0.83 (t, *J* = 6.6 Hz, 3H). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ 172.19, 169.68, 165.22, 160.25, 144.34, 130.18, 127.61, 119.50, 113.75, 101.34, 83.84, 78.43, 55.36, 55.28, 43.24, 40.83, 39.53, 35.99, 35.05, 33.07, 31.93, 29.66, 29.64, 29.53, 29.37, 28.70, 27.28, 22.72, 21.86, 19.15, 18.56, 17.13, 14.16, 12.61.

(*R*,*E*)-3-chloro-*N*-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)tetradec-2-enamide ((*E*)-C14Cl). In a 20 mL vial, **S4a** (64.8 mg, 0.104 mmol, 1.0 equiv.) and 1.0 mL 4:1 AcOH/H_2O were added. After 21 hours, the mixture was concentrated by rotary evaporation, and azeotroped from cyclohexane (3x10 mL) and benzene (3x10 mL). Purification by silica flash chromatography (ethyl acetate → 85:15 ethyl acetate/methanol) afforded **((*E*)-C14Cl** (38.2 mg, 73%, clear oil).

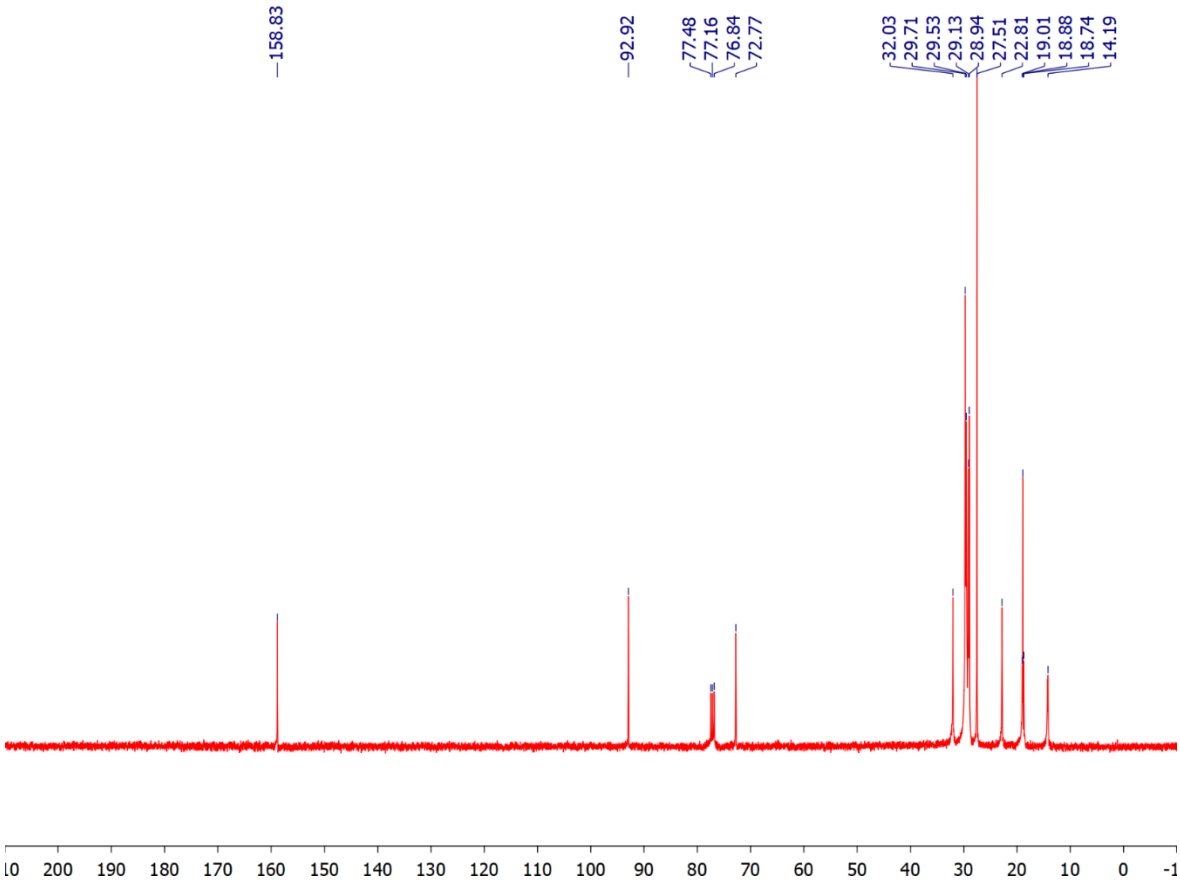
TLC: R_f 0.31 (9:1 ethyl acetate/methanol). **^{1}H-NMR** (400 MHz, $CDCl_3$): δ 7.54 (t, *J* = 5.1 Hz, 1H), 7.30 (s, 1H), 7.17 (s, 1H), 6.03 (s, 1H), 3.98 (s, 1H), 3.68–3.21 (m, 9H), 2.95 (t, 2H), 2.43 (s, 2H), 1.58 (s, 2H), 1.24 (s, 18H), 0.89 (d, *J* = 14.4 Hz, 6H), 0.84 (t, *J* = 6.6 Hz, 3H). **^{13}C-NMR** (100 MHz, $CDCl_3$): δ 174.38, 172.55, 165.33, 153.90, 120.96, 77.58, 77.16, 76.74, 70.73, 39.86, 39.49, 39.37, 36.10, 35.50, 32.04, 29.79, 29.76, 29.68, 29.58, 29.47, 29.12, 27.92, 22.81, 21.34, 20.82, 14.25.

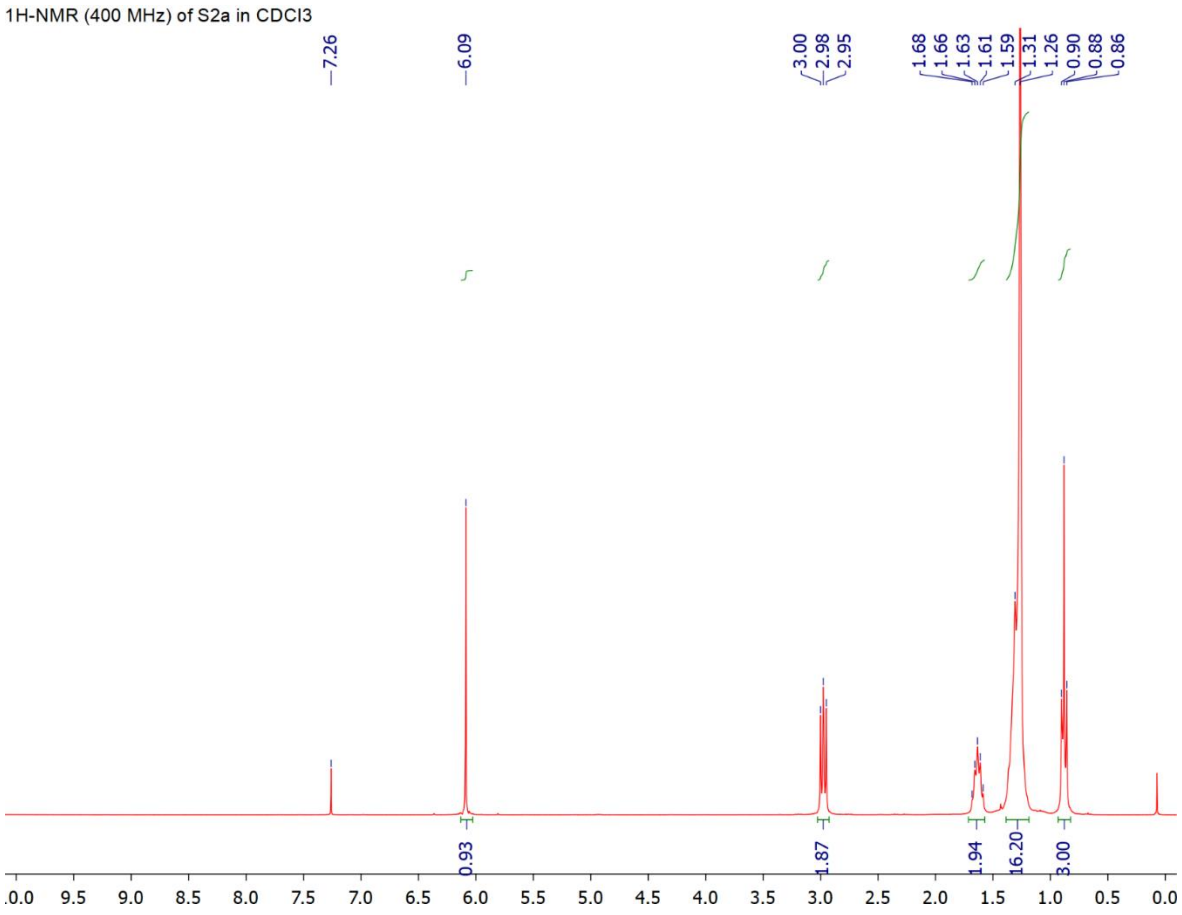
(*R*,*Z*)-3-chloro-*N*-(2-(3-(2,4-dihydroxy-3,3-dimethylbutanamido)propanamido)ethyl)tetradec-2-enamide ((*Z*)-C14Cl). Prepared as described for **(*E*)-C14Cl** from **S4b** to afford **((*Z*)-C14Cl** (35.9 mg, 38%, white solid).

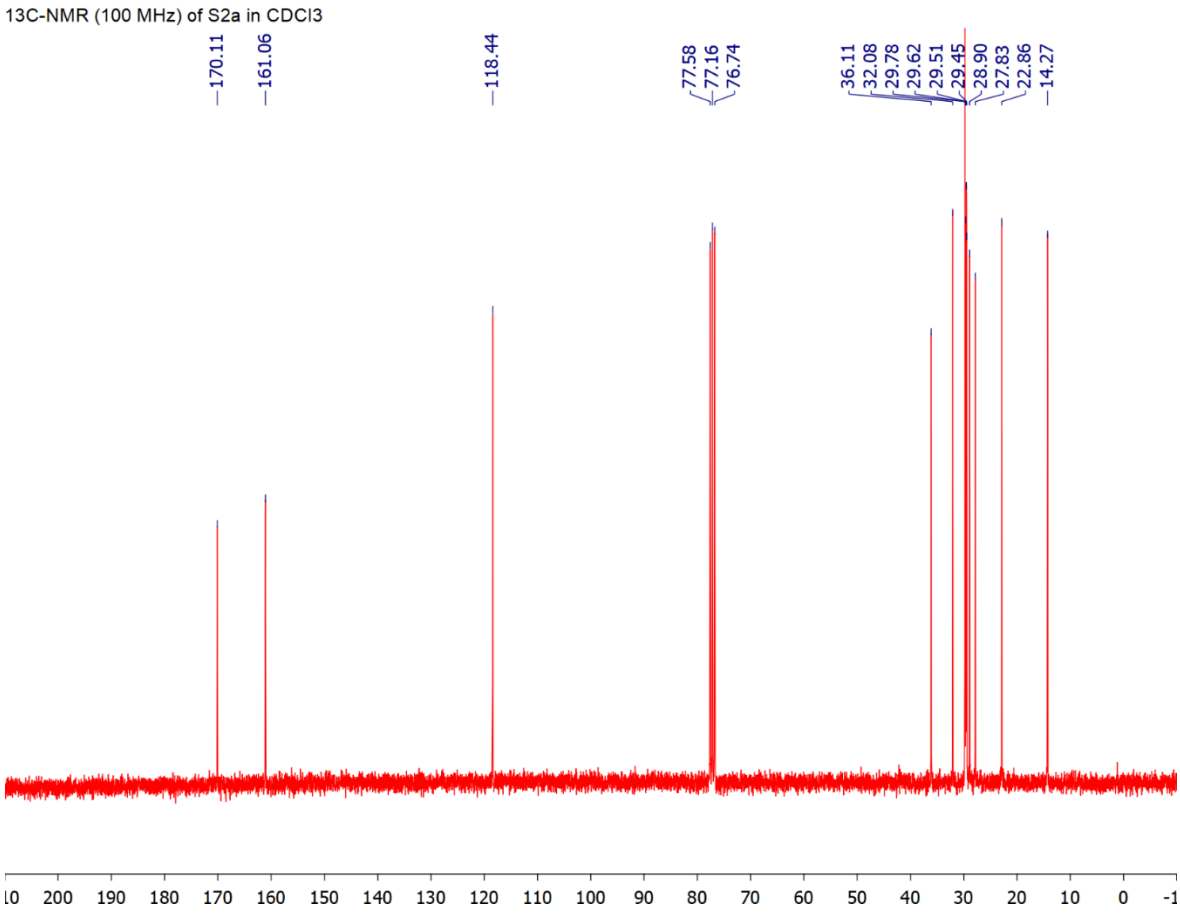
TLC: R_f 0.29 (9:1 ethyl acetate/methanol).


1H-NMR (400 MHz) of S1 in CDCl3
10.98
7.26
2.36
2.34
2.32
1.59
1.58
1.56
1.55
1.39
1.25
0.89
0.87
1.33
1.96
2.14
2.73
13.86
3.00
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

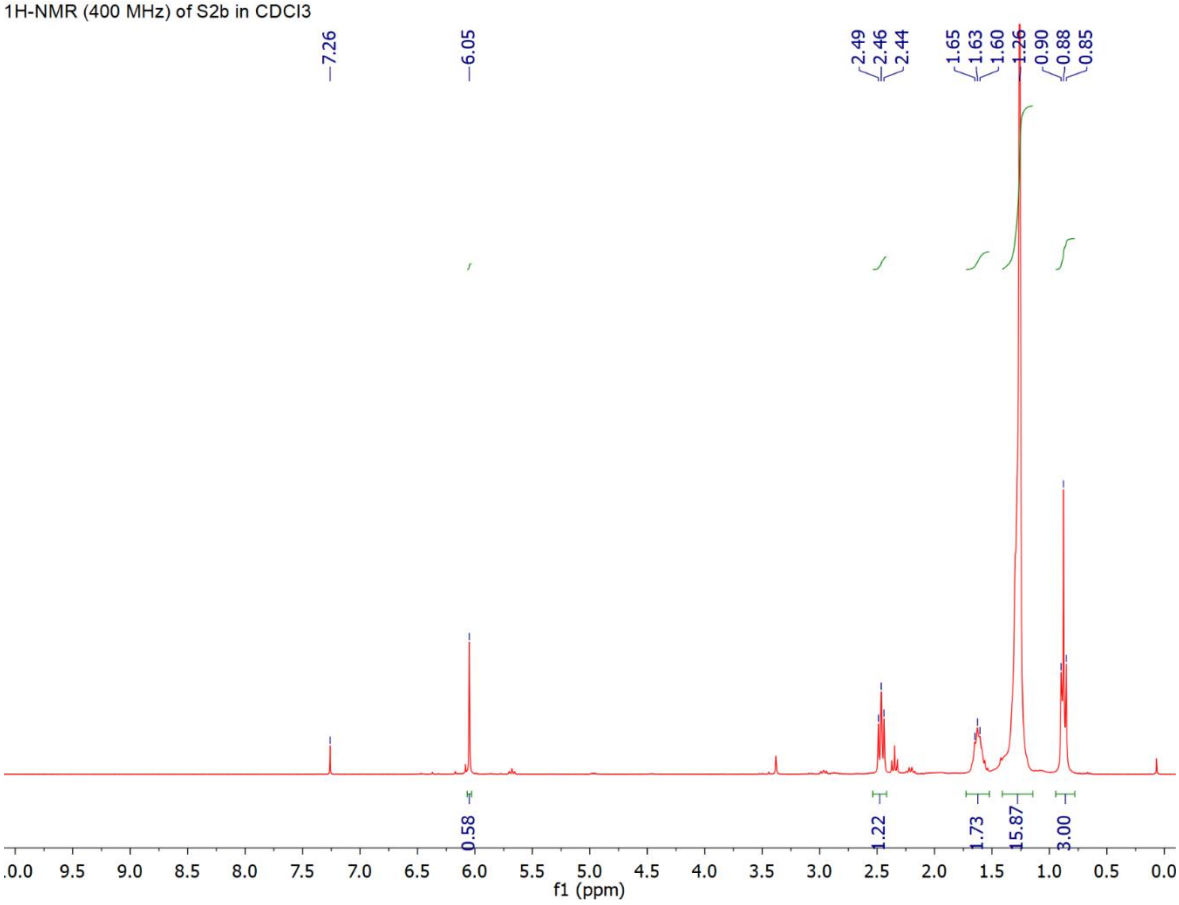
13C-NMR (100 MHz) spectra of S1 in CDCl3

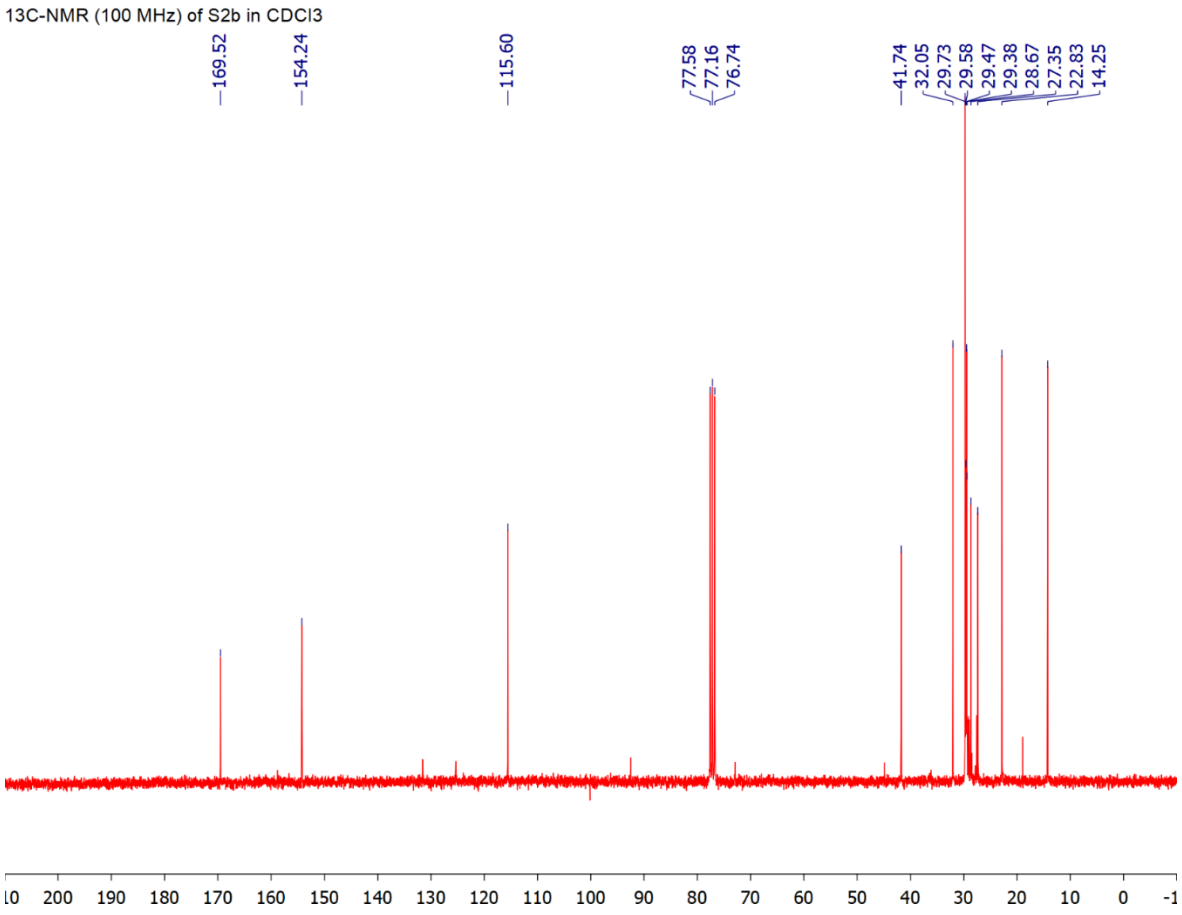



1H-NMR (400 MHz) of S2a in CDCl3
7.26
6.09
3.00
2.98
2.95
1.68
1.66
1.63
1.61
1.59
1.31
1.26
0.90
0.88
0.86
0.93
1.87
1.94
16.20
3.00
.0.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0

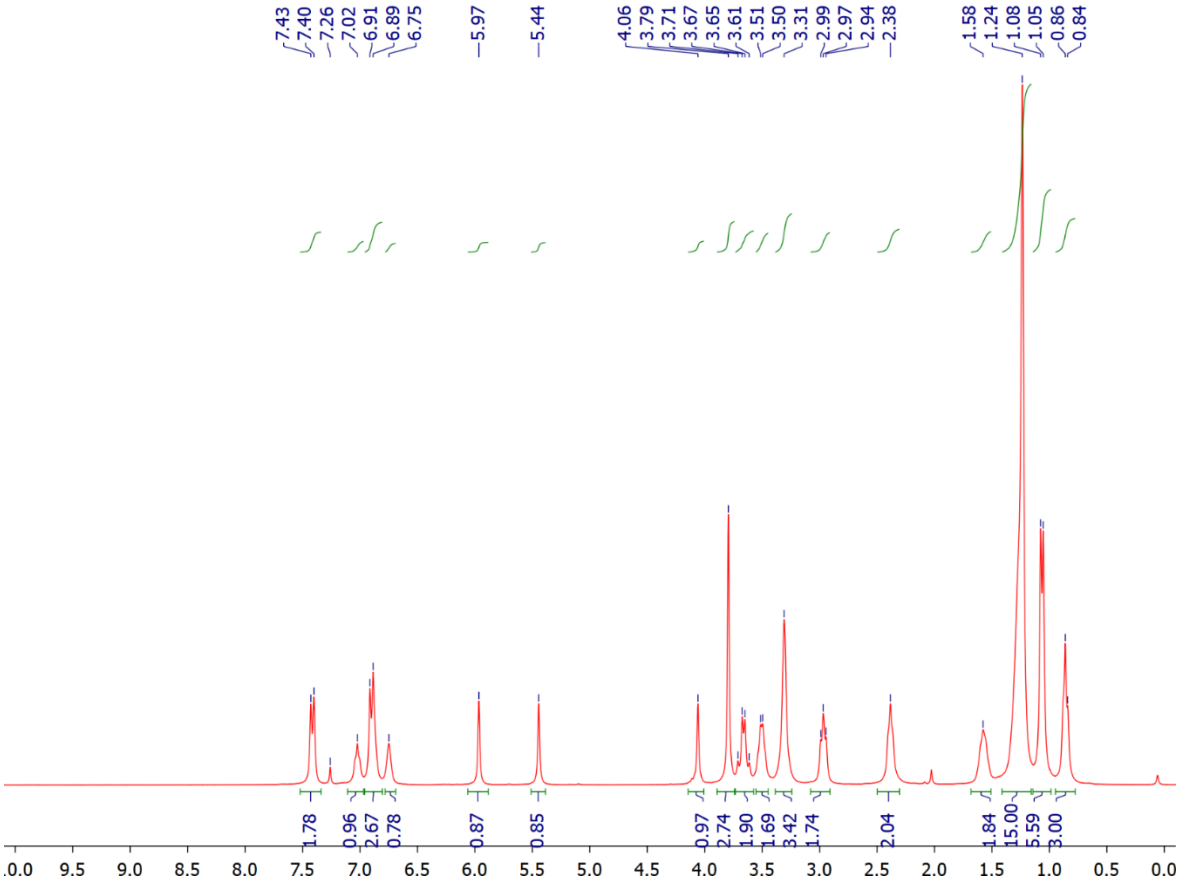

13C-NMR (100 MHz) of S2a in CDCl3
170.11
161.06
118.44
77.58
77.16
76.74
36.11
32.08
29.78
29.62
29.51
29.45
28.90
27.83
22.86
14.27
10 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1

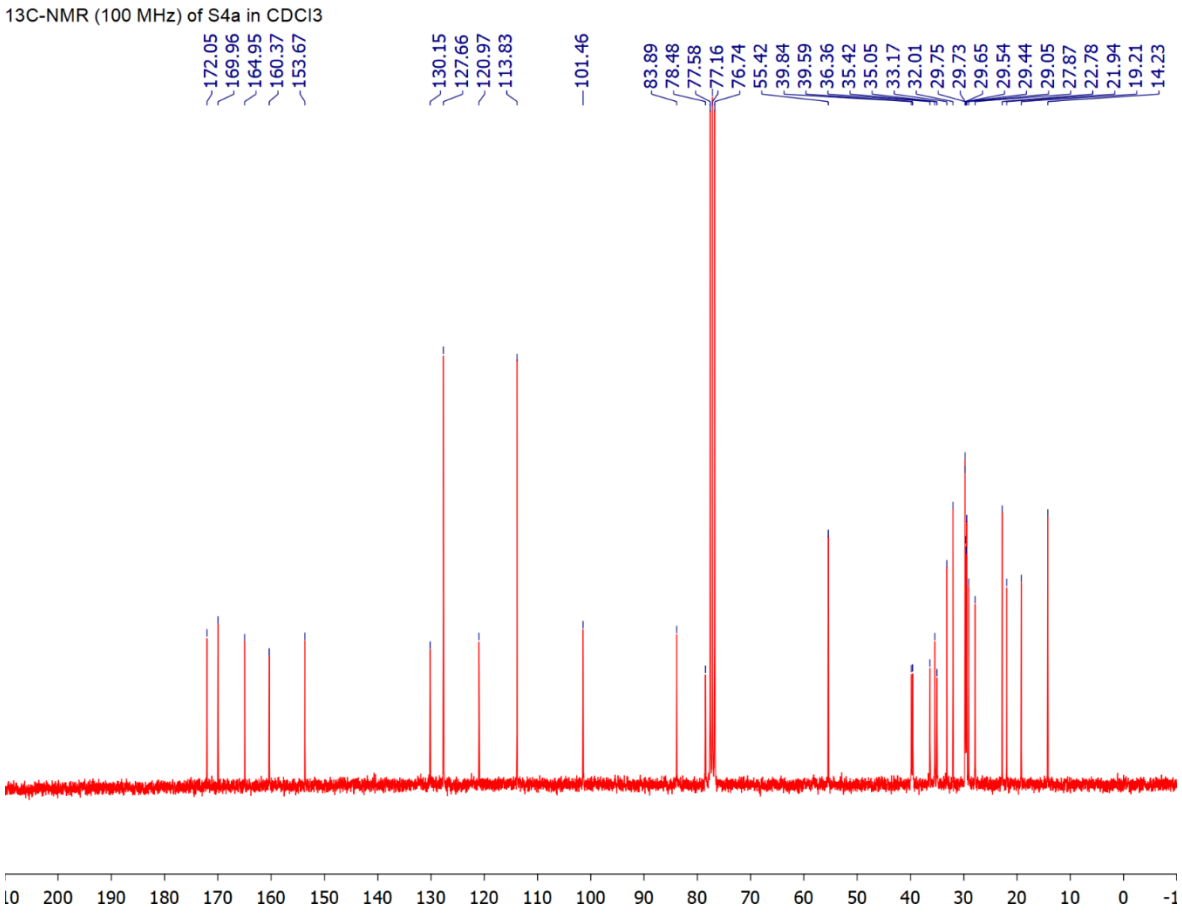
1H-NMR (400 MHz) of S2b in CDCl3

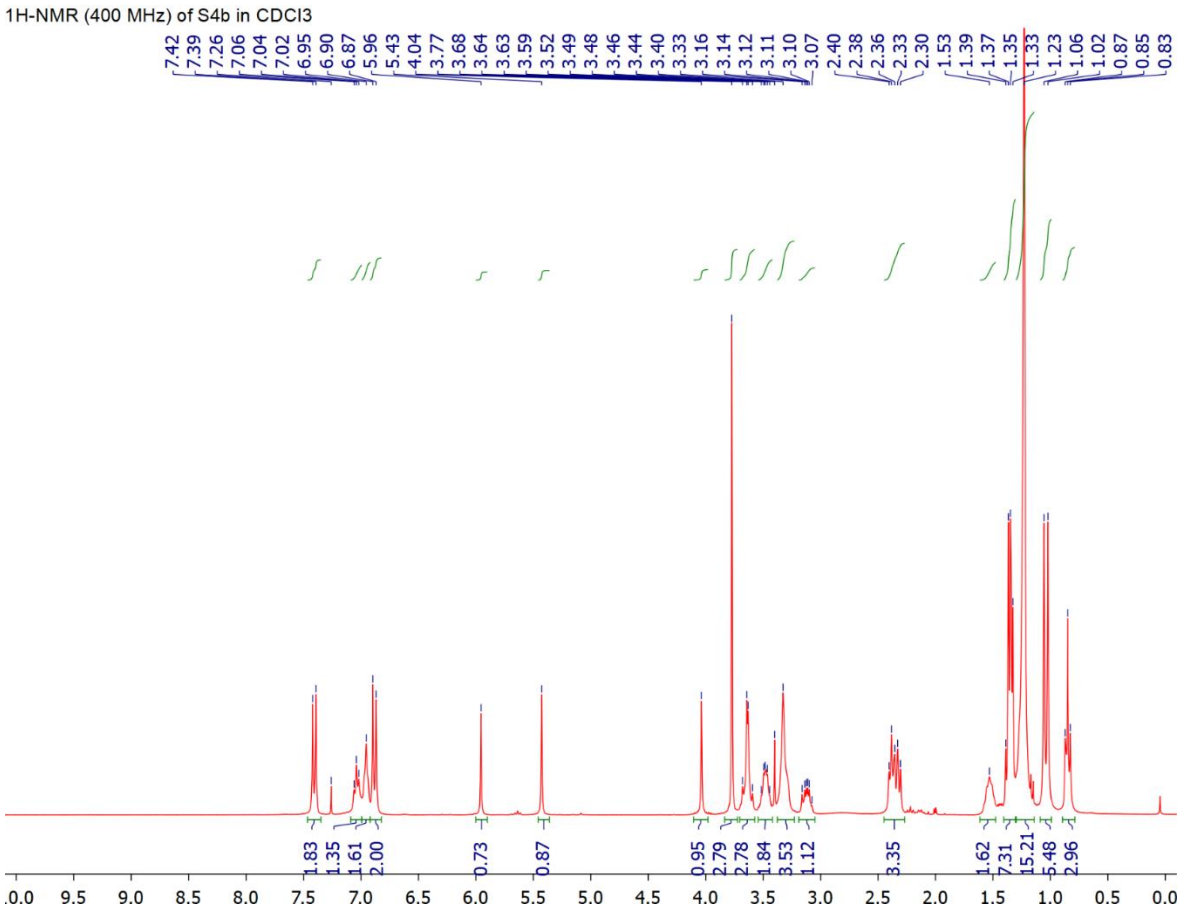


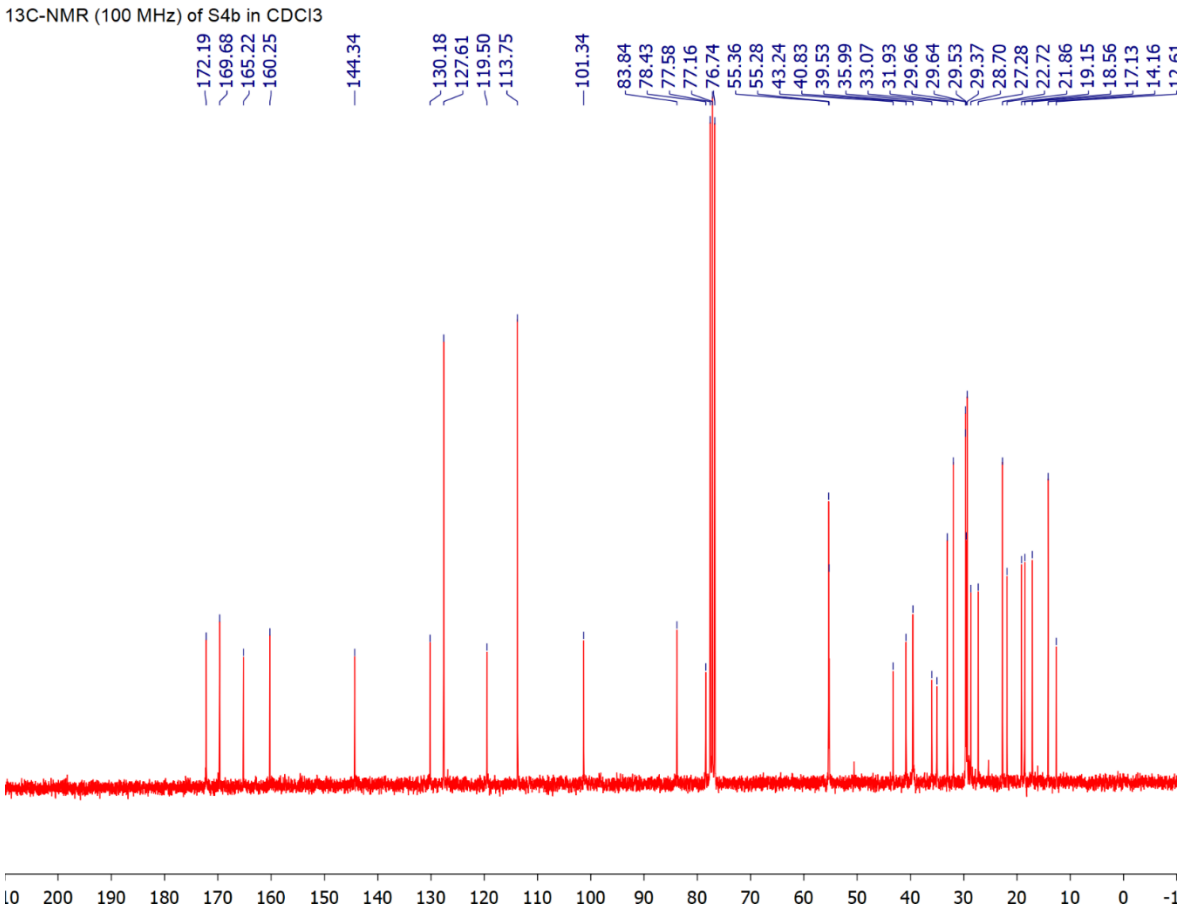

13C-NMR (100 MHz) of S2b in CDCl3
169.52
154.24
115.60
77.58
77.16
76.74
41.74
32.05
29.73
29.58
29.47
29.38
28.67
27.35
22.83
14.25
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

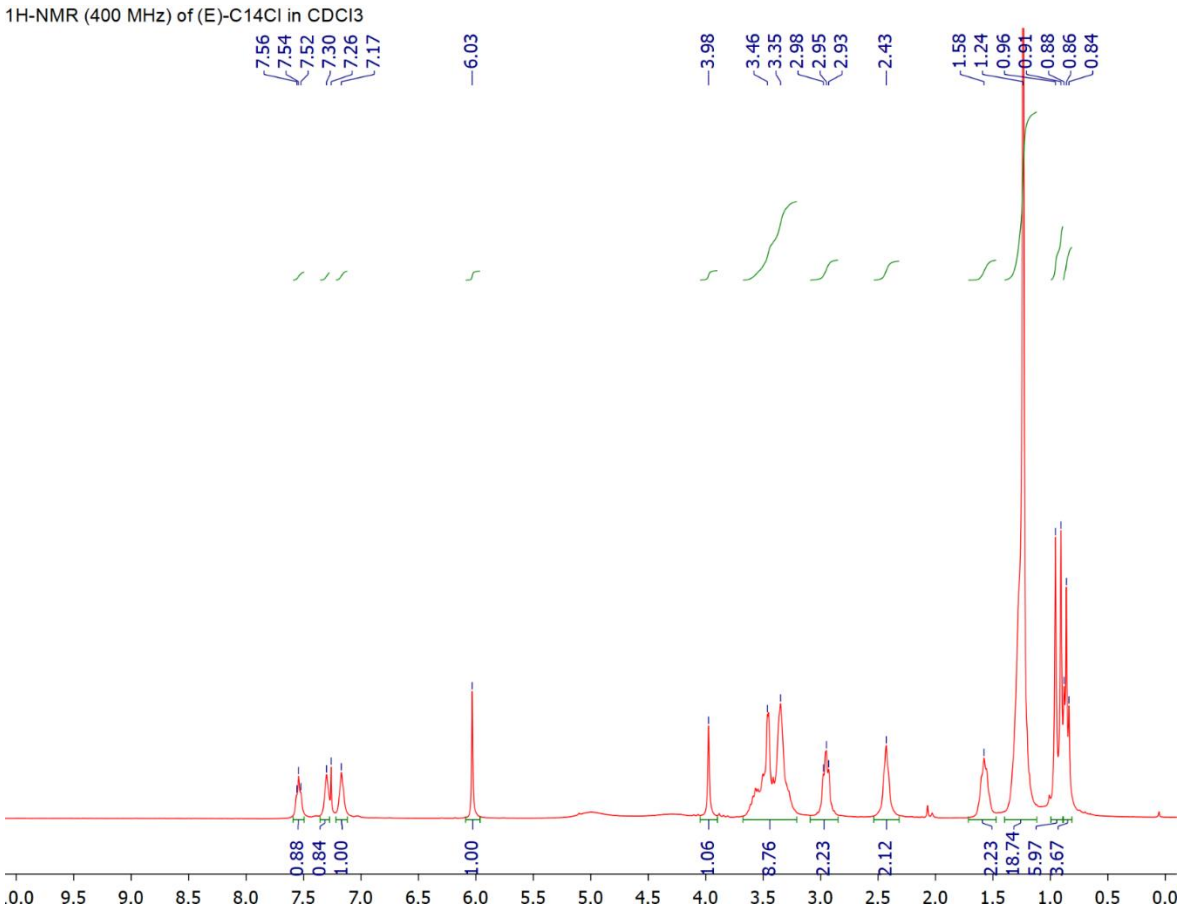
1H-NMR (400 MHz) of S4a in CDCl3

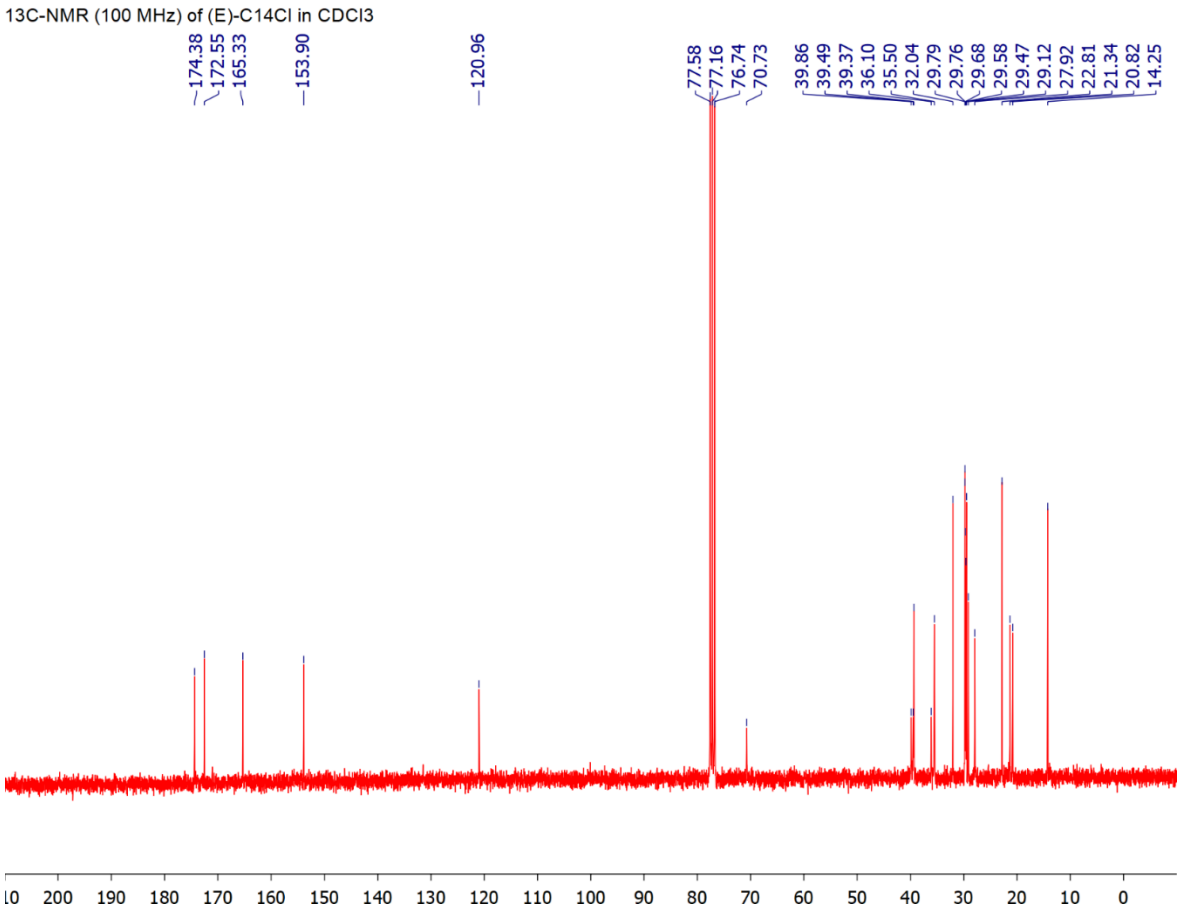



13C-NMR (100 MHz) of S4a in CDCl3
172.05
169.96
164.95
160.37
153.67
130.15
127.66
120.97
113.83
101.46
83.89
78.48
77.58
77.16
76.74
55.42
39.84
39.59
36.36
35.42
35.05
33.17
32.01
29.75
29.73
29.65
29.54
29.44
29.05
27.87
22.78
21.94
19.21
14.23
10 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1

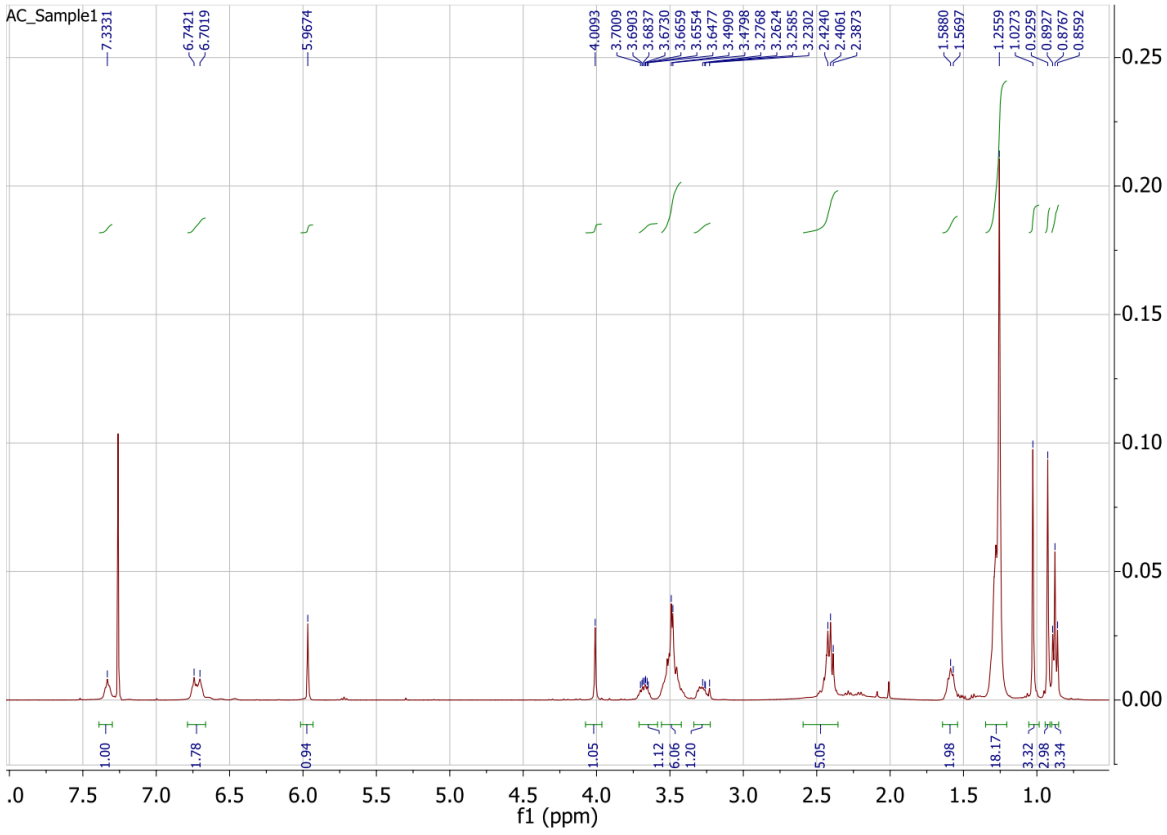

1H-NMR (400 MHz) of S4b in CDCl3
7.42
7.39
7.26
7.06
7.04
7.02
6.95
6.90
6.87
5.96
5.43
4.04
3.77
3.68
3.64
3.63
3.59
3.52
3.49
3.48
3.46
3.44
3.40
3.33
3.16
3.14
3.12
3.11
3.10
3.07
2.40
2.38
2.36
2.33
2.30
1.53
1.39
1.37
1.35
1.33
1.23
1.06
1.02
0.87
0.85
0.83
1.83
1.35
1.61
2.00
0.73
0.87
0.95
2.79
2.78
1.84
3.53
1.12
3.35
1.62
7.31
15.21
5.48
2.96
.0.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0


13C-NMR (100 MHz) of S4b in CDCl3
172.19
169.68
165.22
160.25
144.34
130.18
127.61
119.50
113.75
101.34
83.84
78.43
77.58
77.16
76.74
55.36
55.28
43.24
40.83
39.53
35.99
33.07
31.93
29.66
29.64
29.53
29.37
28.70
27.28
22.72
21.86
19.15
18.56
17.13
14.16
10 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -1

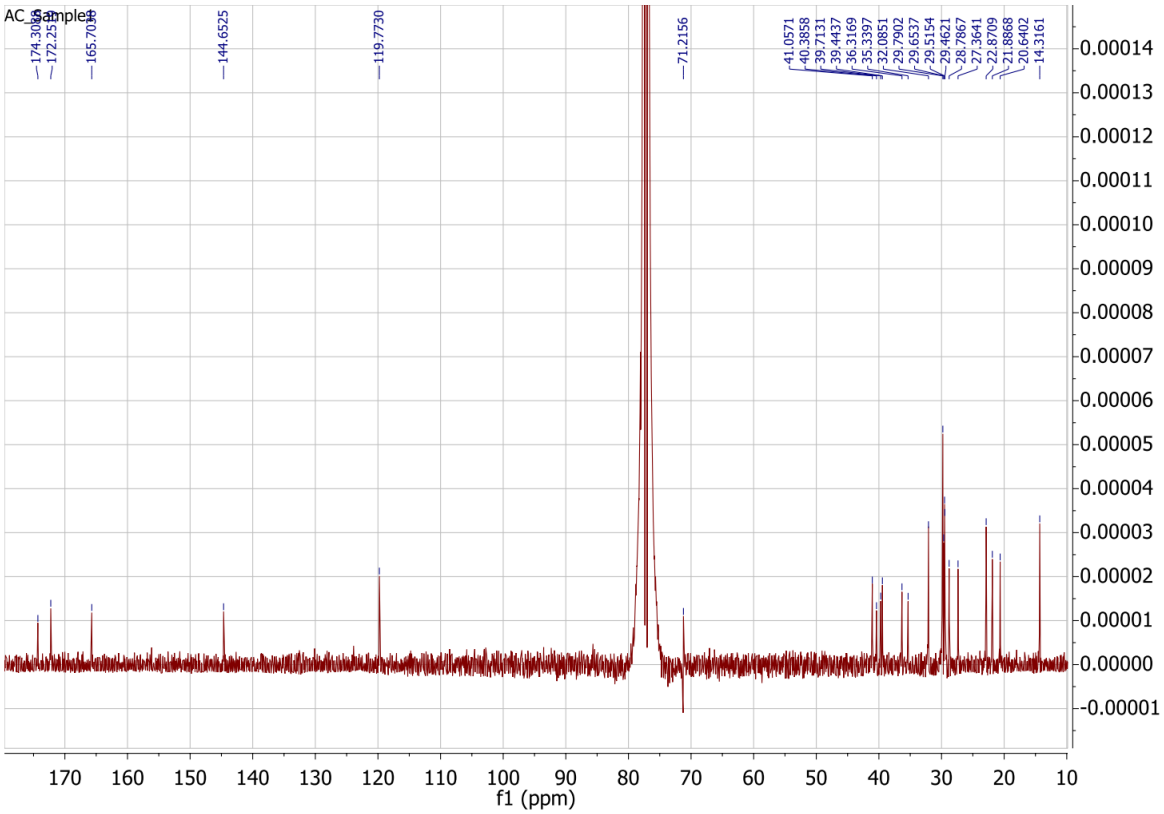

1H-NMR (400 MHz) of (E)-C14Cl in CDCl3
7.56
7.54
7.52
7.30
7.26
7.17
6.03
3.98
3.46
3.35
2.98
2.95
2.93
2.43
1.58
1.24
0.96
0.91
0.88
0.86
0.84
0.88
0.84
1.00
1.00
1.06
3.76
2.23
2.12
2.23
18.74
5.97
3.67


13C-NMR (100 MHz) of (E)-C14Cl in CDCl3
174.38
172.55
165.33
153.90
120.96
77.58
77.16
76.74
70.73
39.86
39.49
39.37
36.10
35.50
32.04
29.79
29.76
29.68
29.58
29.47
29.12
27.92
22.81
21.34
20.82
14.25
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0

^{1}H-NMR of (Z)-C14Cl in $CDCl_3$



^{13}C-NMR of (Z)-C14Cl in $CDCl_3$



References

(1) Chen, A.; N. Re, R.; D. Burkart, M. Type II Fatty Acid and Polyketide Synthases: Deciphering Protein–Protein and Protein–Substrate Interactions. *Nat. Prod. Rep.* **2018**, *35* (10), 1029–1045.

(2) Hillenmeyer, M. E.; Vandova, G. A.; Berlew, E. E.; Charkoudian, L. K. Evolution of Chemical Diversity by Coordinated Gene Swaps in Type II Polyketide Gene Clusters. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (45), 13952–13957.

(3) White, S. W.; Zheng, J.; Zhang, Y.-M.; Rock, C. O. The Structural Biology of Type II Fatty Acid Biosynthesis. *Annu. Rev. Biochem.* **2005**, *74*, 791–831.

(4) Mindrebo, J.; Patel, A.; Misson Mindrebo, L.; Kim, W.; Davis, T.; Ni, Q. Z.; La Clair, J.; Burkart, M. Structural Basis of Acyl-Carrier Protein Interactions in Fatty Acid and Polyketide Biosynthesis; 2019. https://doi.org/10.1016/B978-0-12-409547-2.14662-1.

(5) Heil, C. S.; Wehrheim, S. S.; Paithankar, K. S.; Grininger, M. Fatty Acid Biosynthesis: Chain-Length Regulation and Control. *ChemBioChem* **2019**, *20* (18), 2298–2321.

(6) Cronan, J. E.; Birge, C. H.; Vagelos, P. R. Evidence for Two Genes Specifically Involved in Unsaturated Fatty Acid Biosynthesis in Escherichia Coli. *J. Bacteriol.* **1969**, *100* (2), 601–604.

(7) D'Agnolo, G.; Rosenfeld, I. S.; Vagelos, P. R. Multiple Forms of Beta-Ketoacyl-Acyl Carrier Protein Synthetase in Escherichia Coli. *J. Biol. Chem.* **1975**, *250* (14), 5289–5294.

(8) Feng, Y.; Cronan, J. E. Escherichia Coli Unsaturated Fatty Acid Synthesis: Complex Transcription of the FabA Gene and in Vivo Identification of the Essential Reaction Catalyzed by FabB. *J. Biol. Chem.* **2009**, *284* (43), 29526–29535.

(9) Edwards, P.; Sabo Nelsen, J.; Metz, J. G.; Dehesh, K. Cloning of the FabF Gene in an Expression Vector and in Vitro Characterization of Recombinant FabF and FabB Encoded Enzymes from Escherichia Coli. *FEBS Lett.* **1997**, *402* (1), 62–66.

(10) Milligan, J. C.; Lee, D. J.; Jackson, D. R.; Schaub, A. J.; Beld, J.; Barajas, J. F.; Hale, J. J.; Luo, R.; Burkart, M. D.; Tsai, S.-C. Molecular Basis for Interactions between an Acyl Carrier Protein and a Ketosynthase. *Nat. Chem. Biol.* **2019**, *15* (7), 669–671.

(11) Mindrebo, J. T.; Patel, A.; Kim, W. E.; Davis, T. D.; Chen, A.; Bartholow, T. G.; La Clair, J. J.; McCammon, J. A.; Noel, J. P.; Burkart, M. D. Gating Mechanism of Elongating β-Ketoacyl-ACP Synthases. *Nat. Commun.* **2020**, *11* (1), 1727.

(12) Du, D.; Katsuyama, Y.; Horiuchi, M.; Fushinobu, S.; Chen, A.; Davis, T. D.; Burkart, M. D.; Ohnishi, Y. Structural Basis for Selectivity in a Highly Reducing Type II Polyketide Synthase. *Nat. Chem. Biol.* **2020**, *16* (7), 776–782.

(13) Mindrebo, J. T.; Chen, A.; Kim, W. E.; Re, R. N.; Davis, T. D.; Noel, J. P.; Burkart, M. D. Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases. *ACS Catal.* **2021**, *11* (12), 6787–6799.

(14) Herbst, D. A.; Huitt-Roehl, C. R.; Jakob, R. P.; Kravetz, J. M.; Storm, P. A.; Alley, J. R.; Townsend, C. A.; Maier, T. The Structural Organization of Substrate Loading in Iterative Polyketide Synthases. *Nat. Chem. Biol.* **2018**, *14* (5), 474–479.

(15) 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 Chem. Biol.* **2006**, *1* (11), 687–691.

(16) Worthington, A. S.; Hur, G. H.; Meier, J. L.; Cheng, Q.; Moore, B. S.; Burkart, M. D. Probing the Compatibility of Type II Ketosynthase–Carrier Protein Partners. *ChemBioChem* **2008**, *9* (13), 2096–2103.

(17) Gora, A.; Brezovsky, J.; Damborsky, J. Gates of Enzymes. *Chem. Rev.* **2013**, *113* (8), 5871–5923.

(18) Cronan, J. E. The Chain-Flipping Mechanism of ACP (Acyl Carrier Protein)-Dependent Enzymes Appears Universal. *Biochem. J.* **2014**, *460* (2), 157–163.

(19) 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 Crystallogr. D Biol. Crystallogr.* **2011**, *67* (Pt 4), 271–281.

(20) 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 Crystallogr. D Biol. Crystallogr.* **2011**, *67* (Pt 4), 235–242.

(21) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, 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 Crystallogr. D Biol. Crystallogr.* **2010**, *66* (Pt 2), 213–221.

(22) Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. JLigand: A Graphical Tool for the CCP4 Template-Restraint Library. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68* (Pt 4), 431–440.

(23) Keatinge-Clay, A. T.; Maltby, D. A.; Medzihradszky, K. F.; Khosla, C.; Stroud, R. M. An Antibiotic Factory Caught in Action. *Nat. Struct. Mol. Biol.* **2004**, *11* (9), 888–893.

(24) 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. *J. Biol. Chem.* **1980**, *255* (24), 11949–11956.

(25) de Mendoza, D.; Cronan, J. E. Thermal Regulation of Membrane Lipid Fluidity in Bacteria. *Trends Biochem. Sci.* **1983**, *8* (2), 49–52.

(26) Olsen, J. G.; Kadziola, A.; von Wettstein-Knowles, P.; Siggaard-Andersen, M.; Larsen, S. Structures of β-Ketoacyl-Acyl Carrier Protein Synthase I Complexed with Fatty Acids Elucidate Its Catalytic Machinery. *Structure* **2001**, *9* (3), 233–243.

(27) Rittner, A.; Paithankar, K. S.; Himmler, A.; Grininger, M. Type I Fatty Acid Synthase Trapped in the Octanoyl-bound State. *Protein Sci. Publ. Protein Soc.* **2020**, *29* (2), 589–605.

(28) Zhang, Y.-M.; Hurlbert, J.; White, S. W.; Rock, C. O. Roles of the Active Site Water, Histidine 303, and Phenylalanine 396 in the Catalytic Mechanism of the Elongation Condensing Enzyme of Streptococcus Pneumoniae. *J. Biol. Chem.* **2006**, *281* (25), 17390–17399.

(29) Luckner, S. R.; Machutta, C. A.; Tonge, P. J.; Kisker, C. Crystal Structures of Mycobacterium Tuberculosis KasA Show Mode of Action within Cell Wall Biosynthesis and Its Inhibition by Thiolactomycin. *Struct. Lond. Engl.* **2009**, *17* (7), 1004–1013.

(30) Witkowski, A.; Joshi, A. K.; Lindqvist, Y.; Smith, S. Conversion of a Beta-Ketoacyl Synthase to a Malonyl Decarboxylase by Replacement of the Active-Site Cysteine with Glutamine. *Biochemistry* **1999**, *38* (36), 11643–11650.

(31) Mindrebo, J. T.; Misson, L. E.; Johnson, C.; Noel, J. P.; Burkart, M. D. Activity Mapping the Acyl Carrier Protein: Elongating Ketosynthase Interaction in Fatty Acid Biosynthesis. *Biochemistry* **2020**, *59* (38), 3626–3638.

(32) Tian, W.; Chen, C.; Lei, X.; Zhao, J.; Liang, J. CASTp 3.0: Computed Atlas of Surface Topography of Proteins. *Nucleic Acids Res.* **2018**, *46* (W1), W363–W367.

(33) Witkowski, A.; Joshi, A. K.; Smith, S. Mechanism of the Beta-Ketoacyl Synthase Reaction Catalyzed by the Animal Fatty Acid Synthase. *Biochemistry* **2002**, *41* (35), 10877–10887.

(34) Chisuga, T.; Nagai, A.; Miyanaga, A.; Goto, E.; Kishikawa, K.; Kudo, F.; Eguchi, T. Structural Insight into the Reaction Mechanism of Ketosynthase-Like Decarboxylase in a Loading Module of Modular Polyketide Synthases. *ACS Chem. Biol.* **2022**. https://doi.org/10.1021/acschembio.1c00856.

(35) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *J. Org. Chem.* **1997**, *62* (21), 7512–7515.

(36) Meier, J. L.; Burkart, M. D. Chapter 9 Synthetic Probes for Polyketide and Nonribosomal Peptide Biosynthetic Enzymes. In *Methods in Enzymology*; Complex Enzymes in Microbial Natural Product Biosynthesis, Part A: Overview Articles and Peptides; Academic Press, 2009; Vol. 458, pp 219–254.

Chapter 5: Probing the Protein-Protein Interactions of a Biotin Synthesis Enzyme, BioF, and Acyl Carrier Protein through Sulfur Fluoride Exchange Crosslinker.

Introduction

BioF, or 8-amino-7-oxononanoate synthase (AONS), catalyzes the decarboxylative condensation of L-alanine and pimeloyl group *via* the help of cofactor PLP to synthesize the product 8-amino-7-oxononanoate (AON) as an intermediate of biotin biosynthesis.[1] AON subsequently goes through the catalysis of BioA, BioD, and BioB enzymes to synthesize biotin, which is an essential enzyme cofactor also known as vitamin B_7.[2] In *E. coli*, the biotin biosynthesis pathway utilizes the endogenous fatty acid biosynthesis (FAB) pathway to generate the pimeloyl methyl ester precursor.[3] In this stage the pathway utilizes a small, 4-helical acyl carrier protein (ACP) to shuttle the thioester-tethered substrates to each partner enzyme (PE).[4] The pimeloyl-ACP methyl ester exits the FAB pathway and is first converted to pimeloyl-ACP by BioH before reacting with BioF to unload the ACP-tethered substrate, which makes BioF the last enzyme in biotin biosynthesis to accept substrate from the ACP. Historically, *Ec*BioF was thought to utilize pimeloyl-CoA as the diacid donor,[1] and in fact, organism like *Bacillus subtilis* utilizes a pimeloyl-CoA synthetase, BioW,[5,6] to generate pimeloyl-CoA as the primary diacid substrate accepted by *Bs*BioF rather than pimeloyl-ACP.[7] A better understanding of BioF's enzymology can inform us how different organisms adapt their systems to produce the same metabolite.

*Ec*ACP is known to sequester the tethered substrate during transportation.[8] The substrate delivery undergoes a chain-flipping mechanism,[9] which describes the process that the substrate flips out from the ACP pocket to the PE pocket, and is often triggered by

PPIs.[10,11] Characterization of the protein-protein interface is most often achieved by solving the structure of ACP in complex with PE, and to date, several *E. coli* PE-ACP complexes have been obtained, including FabD,[12] FabF,[13] FabB,[14] FabA,[15] and FabZ[16] from the FAB pathway, and BioH[17] from the biotin biosynthesis pathway. BioF is a more difficult enzyme to work with as the catalysis involves one cofactor and two substrates, and that the enzyme itself does not form direct linkage with ACP during the reaction, making it challenging to rationally design any crosslinker. Co-crystallization method had successfully yielded BioH-ACP structure, but the same approach failed on BioF.[7] In this study, we screened a library of crosslinkers and identified a warhead for BioF crosslinking, which ultimately led to the crystal structure of BioF-ACP complex. Analysis of the structure reveals the protein-protein interactions (PPIs) governing substrate specificity that shed light on the enzymology of BioF.

The crosslinker contains the benzenesulfonyl fluoride (BSF) warhead which is commonly used to target serine residue as seen in the inhibitor for serine protease.[18] The reactivity and mechanism regarding sulfur(VI) fluoride electrophile, being studied as early as 1927,[19] was revisited in 2014 by Sharpless *et al.* and was coined the term "Sulfur(VI) Fluoride Exchange", or SuFEx, as a new tool in click chemistry.[20] SuFEx is especially useful in covalent modification of protein because its electrophilicity is well-balanced to not react with nucleophilic amino acid unless there is one or more hydrogen bond donor (HBD) nearby.[21] This property makes the warhead only reacts with context-specific residues, most often exist in the protein active site, and residues like threonine, lysine, tyrosine, cysteine, and histidine have also been reported along with serine that got specifically targeted by the warhead.[22] Our application of SuFEx leads to a successful design of a PE-ACP crosslinker,

which is the first of its kind, further expand the utility of SuFEx. The BSF warhead also expand our toolbox on crosslinker development, especially with enzymes that catalyze complex reaction between multiple substrates.

Results

Crosslinker panel identified warhead for BioF-ACP crosslinking. The PLP-dependent BioF catalyzes decarboxylative condensation of pimeloyl-ACP and L-alanine to produce AON, *holo*-ACP, and carbon dioxide. The mechanism begins with BioF forming an internal aldimine with the PLP *via* the active site residue K236 and then the aldimine reacts with L-alanine to form an external aldimine that subsequently condenses with the pimeloyl group from pimeloyl-ACP (Figure S5.3A). Designing a mechanism-based crosslinking probes often relies on mimicking the reaction intermediate that covalently links PE with ACP.[23,24] Since BioF does not form covalent linkage with ACP during the catalysis, it is challenging to create a crosslinking probe by rational design. We thus turned our strategy into screening electrophilic warheads that may potentially react with BioF's active site residues. We performed the one-pot chemoenzymatic reaction[25] utilizing twelve PPant-containing crosslinking probes with five unique warheads to generate a panel of *crypto*-ACPs (Figure 5.1). These probes were used in previous studies with warheads targeting the active site of acyltransferase (Probe **1**, **10**, and **11**)[12,26] and ketosynthase (Probe **1**-**9**).[13,14,27,28] An *E. coi* KS, FabF, was used as a positive control to test the panel. All of the *crypto*-ACPs crosslinked, in which probe **1**-**9** that were utilized to study KS gave the highest yields (Figure 5.2, left). BioF was then subjected to the panel and only probe **10** shows obvious crosslinking

band (Figure 5.2, right). Probe **10** contains the warhead benzenesulfonyl fluoride (BSF) that has been used to target acyltransferase[26] and serine protease.[18] As a result, the panel successfully determined a potential warhead for crosslinking BioF and ACP. We also synthesized a shorter and a longer version of probe **10**, **10S** and **10L**, respectively (Figure 5.1), for crosslinking optimization.

One-pot and crosslinking optimization for the unstable BSF crosslinker. A typical PE-ACP crosslinking workflow involves *crypto*-ACP purification after one-pot reaction in order to eliminate residual small molecule crosslinker that might react with the PE. However, the BSF warhead is prone to hydrolysis and would lose activity within several hours in aqueous solution under room temperature. As a result, following a standard crosslinking procedure often led to a poor yield. To minimize the time of probe exposure in aqueous solution, we tested out a seamless procedure that skips the purification or desalting step after the one-pot reaction. Since one-pot reaction goes to completion under 6 hours, we set up four batches with probe **10S** to add in BioF for crosslinking at four different time points (0, 2-hour, 4-hour, and 6-hour one pot reaction time, respectively). Interestingly, the crosslink reaction almost reached to completion for the 2-hour and 4-hour time point trials (Figure S5.1). Given that a 2-hour one-pot reaction only converts half of the *apo*-ACP into **10S**-ACP, which means that there is more free probe than the **10S**-ACP in the solution, the high crosslinking yield implies that **10S**-ACP is more readily accepted by BioF than the free probe. This also suggests that the crosslinking reaction is mediated by ACP, potentially through a substrate delivery mechanism.

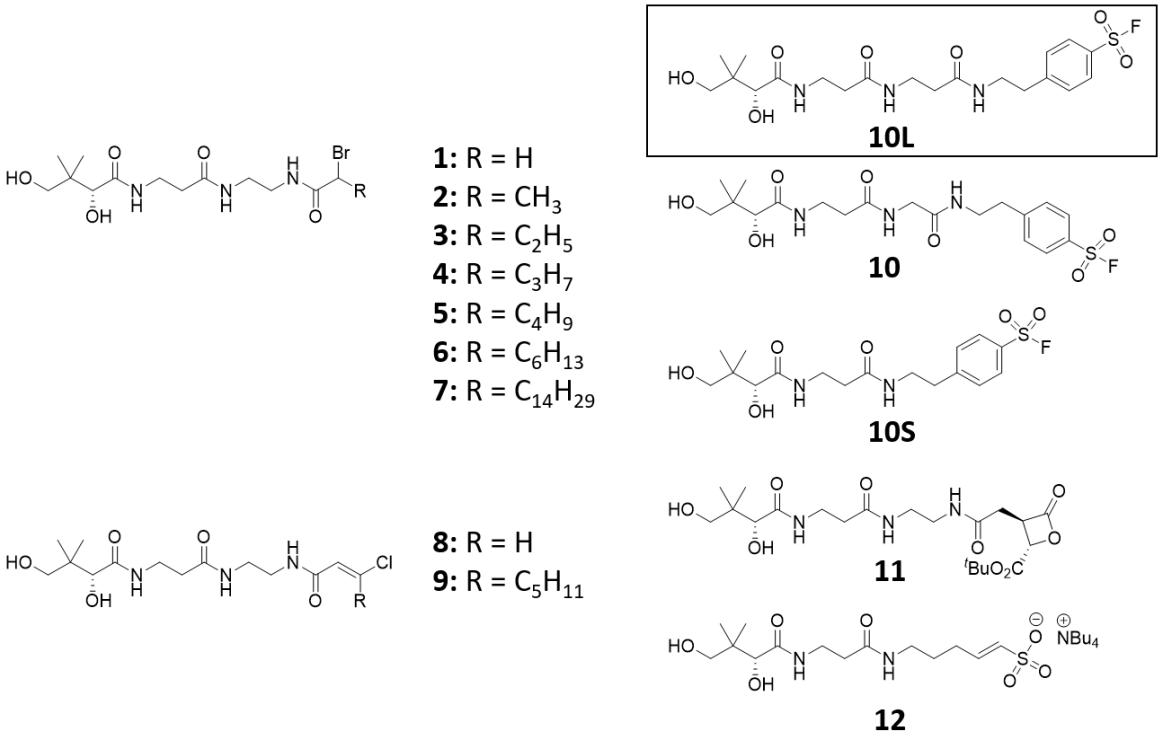


Figure 5.1: Crosslinking probes used in this study. Probe **1-12** were used in the crosslinker panel. Only **10** crosslinked BioF with ACP (Figure 5.2) and thus was subjected to linker length optimization. A longer (**10L**) and a shorter (**10S**) version of **10** were synthesized and tested through crosslinking reaction. The crosslinker **10L** ultimately yields the crystal structure of BioF=ACP.

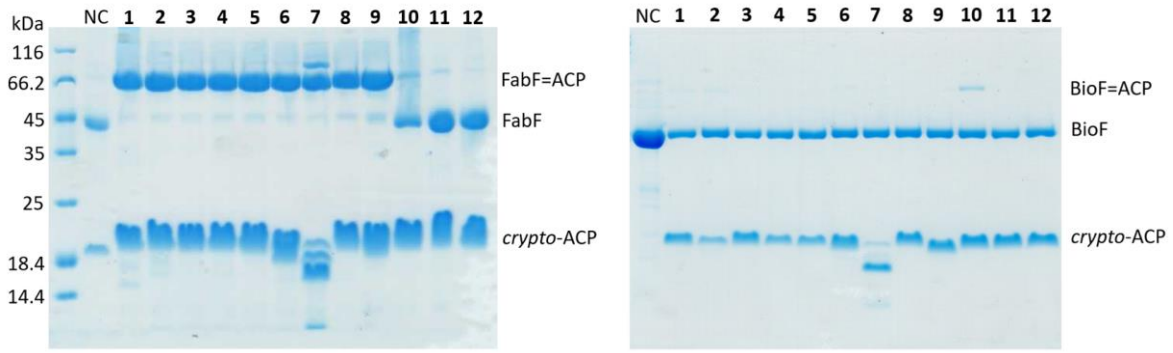


Figure 5.2: PE-ACP crosslink test using the crosslinker panel (Figure 5.1). twelve FPLC-pure *crypto*-ACPs were incubated with FabF (left gel) and BioF (right gel). While FabF reacts with all of the crosslinkers, BioF only reacts with **10**. (NC = negative control)

Different form of crosslinked complexes can be separated by an anion exchange column. A large scale of BioF and **10S**-ACP crosslinking was performed under the optimized protocol, which is adding in BioF after 2.5-hour of one-pot reaction. The resulting conversion rate is approximately 50%. Since BioF exists as a homodimer, it has two active sites that can theoretically crosslink with two ACPs at the same time. We term the complex as "single crosslinked" or "double crosslinked" for a BioF dimer crosslinks with one or two ACPs, respectively. A 50% crosslinking yield could mean that the product is a mixture of single crosslinked, double crosslinked, and non-crosslinked BioF. A BioF dimer has a molecular weight of ~83 kDa while ACP is only ~9 kDa, so the size difference between the single crosslinked (~92 kDa) and the double crosslinked (~101 kDa) complexes is too small to be separated by a size exclusion column. However, we notice that there is a notable isoelectric point (pI) difference between BioF (pI = 6.64) and ACP (pI = 3.98) that may provide the handle for separation. We subjected the product mixture to an anion exchange column (HiTrap Q Sepharose^{TM} HP from Cytiva) that distinguishes substances through their surface charge. With a linear gradient of NaCl concentration from 0 to 1.0 M, the BioF dimer eluted at 0.11 M, the single crosslinked complex eluted at 0.22 M, and the double crosslinked complex eluted at 0.30 M of NaCl concentration (Figure S5.2). There were no continuous fractions bridging the different elution regions indicating that the different crosslinked form is separated cleanly by the anion exchange column.

Linker length optimization. There are three candidates in the BSF crosslinker library, **10S**, **10**, and **10L**, respectively, in an ascending order of the linker chain length. Large scale crosslinking reaction with the three crosslinkers show no noticeable difference

on the yield, which is approximately 50%. We then analyzed the mechanism of BioF catalysis (Figure S5.3A) and found a tetrahedral intermediate that was formed by the condensation of the external aldimine with pimeloyl-ACP. This intermediate connects the active site bound PLP with ACP and can serve as a reference for optimizing crosslinker chain length. The linear alignment of the tetrahedral intermediate with the three crosslinkers shows that probe **10L** has the closest chain length (Figure S5.3B). Therefore, we chose **10L** as the crosslinker to proceed to obtain crystal structure.

Crystal structure of BioF-ACP crosslinked complex. The BioF=**10L**-ACP complex (where the '=' sign denotes the crosslinking bond) was crystallized and diffracted to a nominal resolution of 2.3-Å. The asymmetric unit (ASU) contains a monomer of BioF bound with an ACP, indicating that the complex is a diprotomer and thus is double crosslinked (Figure 5.3). This is called into question because the batch used for crystallography was from pure single crosslinked complex fractions. We performed an occupancy test on ACP to confirm if the ratio of BioF monomer to ACP is 1-to-1 in the crystal structure. In detail, the initial occupancy of ACP was set to be 0, 0.5, or 1.0, and the coordination was subjected to a few rounds of refinement without fixing the occupancy value. The result indicates that ACP occupancy is around 1.0, and so the crystal structure is indeed a double crosslinked complex. The reason of the change is unknown. One possible explanation is that the crosslinked protomer dissociates with BioF monomer and associates with another protomer during crystal packing.

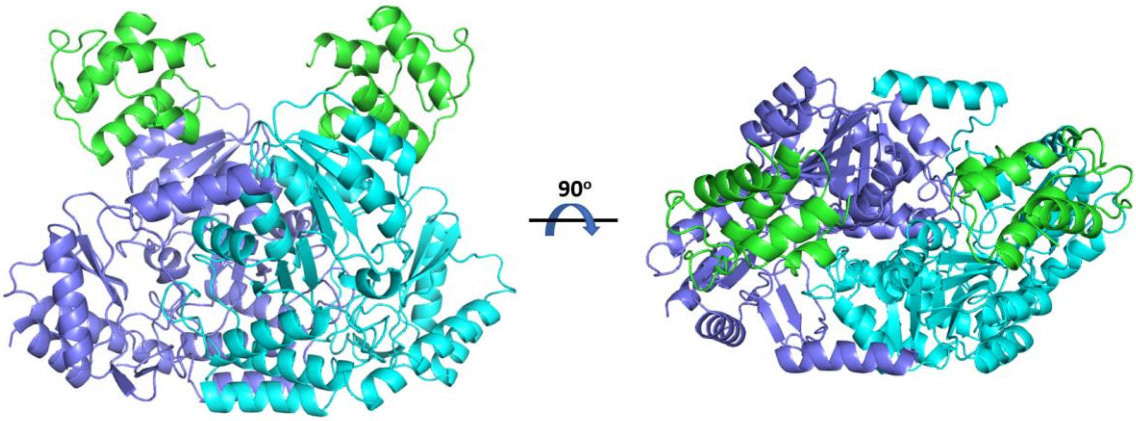


Figure 5.3: Top view and side view of the structure of BioF=ACP complex. The asymmetric unit (ASU) contains one BioF monomer (blue) and one ACP (green), and two ASUs are shown here to illustrate a double crosslinked complex.

BioF=**10L**-ACP structure aligns to the *apo*-BioF structure (PDB ID: 1BS0)[1] with an RMSD of 1.50 Å. The main difference occurs at region between residue 14 and 21, which locates at the C-terminal end of the first α-helix and the loop right after it. This conformation is unique to *apo*-BioF, as the *holo*-BioF (PLP-BioF, PDB ID: 1DJE) and the substrate bound BioF (AON-BioF, PDB ID: 1DJ9) structures has the same regional conformation with our crosslinked structure.[29] Indeed, a smaller RMSD value (1.05 Å) is observed from the alignment of BioF=**10L**-ACP and PLP-BioF. Another structural feature of BioF is a conformational transition at the C-terminal domain (between residue 320 and 361) likely triggered by the bound substrate, which was discovered in the structure of AON-BioF.[29] This domain forms a cover on the cleft that contains the active site at the bottom (Figure S5.4). When a substrate intermediate is in the cleft, the carboxylate of AON forms polar contacts with T352 that sits at the tip of the cover, and this contact is only made possible after the conformational transition. This transition is not observed in the BioF=**10L**-ACP structure likely because the crosslinker lacks the key element that the substrate has, such as the decarboxylate, to trigger the transition.

Determination of the crosslinking site. The linear distance of ACP S36 to the N_ζ of the active site K236 of BioF is 24.5-Å, a length that is within the reach of the crosslinker. However, the electron density is not sufficient to model in the crosslinker unambiguously, especially for the PPant moiety. The BSF warhead had been reported to crosslink with cysteine, tyrosine, lysine, serine, threonine, and histidine,[22] and the BioF active site chamber contains quite a few of them (Y49, H133, S135, S179, H207, T233, K236, and Y264', which the apostrophe denotes residue from the other monomer). The scarce electron density points

to S135 as the crosslinking site. A study on using benzene-1,2-disulfonyl fluoride as human neutrophil elastase inhibitor demonstrates that the BSF warhead crosslinked with a serine residue under the assist of a nearby histidine as a hydrogen bond donor (HBD).[30] This is reminiscent of the BioF S135 which the nearby H133 could play similar role (Figure S5.5). To investigate the crosslinking site, we performed site-directed mutagenesis[31] to obtain S135A and H133A mutants. A crosslinking assay was set up with the mutants and the wild-type (WT) BioF, and four time points (0, 1 h, 3 h, 16 h) were taken to be analyzed by gels. Both S135A and H133A reached about 30% crosslinking yield as that of the WT after 16 hours of incubation. In comparison, H207A mutant is as efficient as the WT while S179A mutant reached 72% crosslinking yield comparing to that of WT (Figure 5.4). This result suggests that S135 and H133 do work coordinately in the crosslinking reaction similar to the case reported in elastase.[30] However, the existing crosslinking yield also suggests that multiple crosslinking sites may exist, for example, the S179 residue. Interestingly, an active site residue N47 that serves as a HBD during decarboxylation (Figure S5.3A) also shows a decrease in crosslinking yield (59% of WT) when mutated to an alanine. Given that BSF warhead requires ambient HBD to react with an amino acid, this result implies that the adjacent Y49 might also be a crosslink target.

PPIs and the protein interface. The PPIs mainly occur between the helix 2 of ACP and the β6 of BioF (Figure 5.5). It is typical for ACP to have its helix 2 anchored on the PE interface when bound, which could be observed in other type II synthases.[10,11,14,15,17,27] R149 of BioF also form a salt bridge with E56 from the helix 3 of ACP. PPIs involving the dynamic helix 3 have been proposed to facilitate substrate delivery. In the case of an *E. coli* KS, FabF,

the R206 residue interacts with helix 3 and mutating the arginine to an alanine only affects the k_{cat} of FabF but not the K_M, implying that R206 plays more of a role in catalysis than ACP binding.[10] The (BioF)R149-E56(ACP) interaction may serves a similar role. Notably, R160 from the α6 of BioF could also plays a role in PPIs (Figure 5.5).

BioF from *Bacillus subtilis* (*Bs*BioF) utilizes pimeloyl-CoA, rather than pimeloyl-ACP, as the substrate while *E. coli* BioF can accept both.[7] Sequence alignment of the two BioF reveals that *Bs*BioF lacks the basic residues to interact with ACP. The two core interacting residues, R148 and R149, identified from the crosslinked structure are both replaced by valine in *Bs*BioF. R130 and R160 are also replaced by a glutamine and an asparagine, respectively. An *in silico* docking has been conducted by Manandhar *et al.* and they proposed a luminal surface of *Ec*BioF as the ACP binding site.[7] In addition to R130, R148, and R149, their docking result also suggests K6 and R14 from α1 participating in PPIs. However, the ACP binding site they suggest is not the same as what is presented in the crosslinked structure. In the structure, K6 and R14 are both more than 11-Å away from the closest ACP residue. Their data also does not include R160, which could be observed in the structure as well.

To further investigate the PPIs, we performed alanine-scan on the arginine residues to obtain R148A, R149A, and R160A mutants. Since ACP facilitates the crosslinking reaction through PPIs, and also that the crosslinker is as long as the native substrate to allow native ACP binding events, we hypothesize that the crosslinking efficiency would reflect the binding activity of BioF and ACP. We thus ran a crosslinking assay with these interface mutants to test their influence on PPIs. All of the mutants experienced a great reduction in

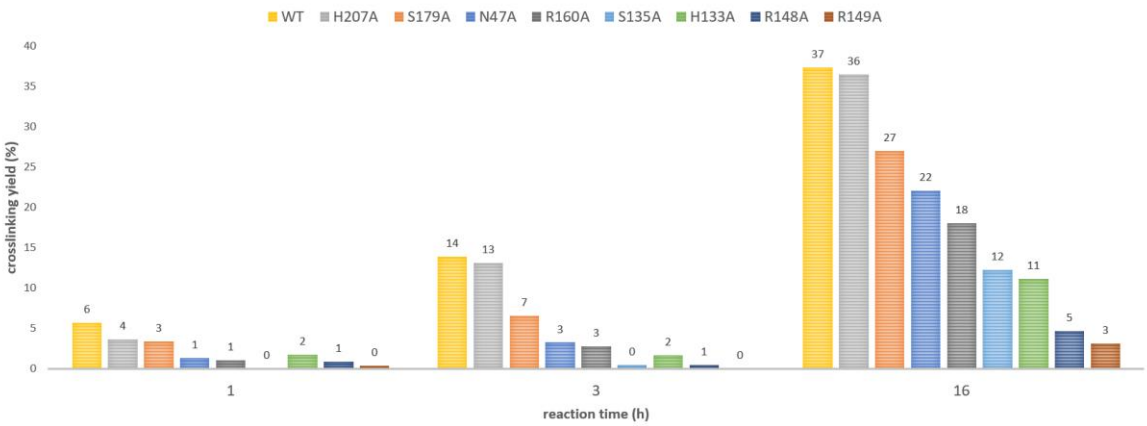


Figure 5.4: Crosslinking assay of the active site residues and interface residues mutants. H207, S179, N47, S135, and H133 are residues found in the active site chambers while R160, R148, and R149 are interface residues.

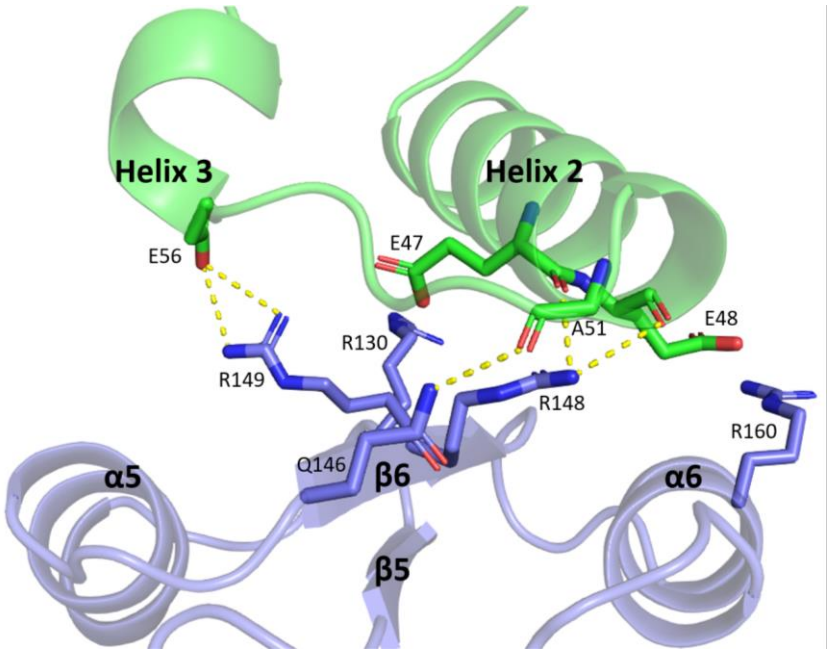


Figure 5.5: BioF-ACP interface and PPIs. The main interactions happen between β6 of BioF (blue) and the C-terminal of ACP (green) helix 2. The R149 interaction with helix 3 E56 can be important for substrate delivery.

crosslinking activity, especially R148A and R149A, which were only 12% and 8%, respectively, in crosslinking yield compared to that of WT after 16 hours of incubation, and for R160A it was 48% (Figure 5.4, S5.5). As a result, the assay suggests that all three arginine residues participate in PPIs, and the identified protein-protein interface in the crystal structure is real, rather than a result of crystal packing.

Conclusion

Utilizing protein crosslinking to investigate PPIs, and in general, the enzymology of ACP-dependent synthases has achieved much success in recent years. In this study, we expand the crosslinker pool by utilizing a SuFEx warhead that targets context-specific residue(s) in BioF. A new working procedure of one-pot and crosslinking was also introduced to fit the restriction of the crosslinker. The resulting BioF-ACP complex reveals a protein-protein interface that is not predicted by *in silico* docking. The subsequent mutagenesis study confirms two interface residues, R148 and R149, playing a key role in PPIs. This study would inform future PPIs investigation on enzymes that catalyze complex reaction.

Acknowledgement

I thank Prof. Tony Davis on probe synthesis, Marianne Bowman and Gordon Louie (Salk research institute) with the help on protein crystallography, Prof. Michael Burkart and Prof. Joseph Noel on project advising, and Kelley Tran with the help on mutagenesis study.

I also would like to thank Prof. Dominic Campopiano and Dr. David Clarke from University of Edinburgh (UK) on providing us the BioF plasmid and valuable suggestions.

Supplementary Information

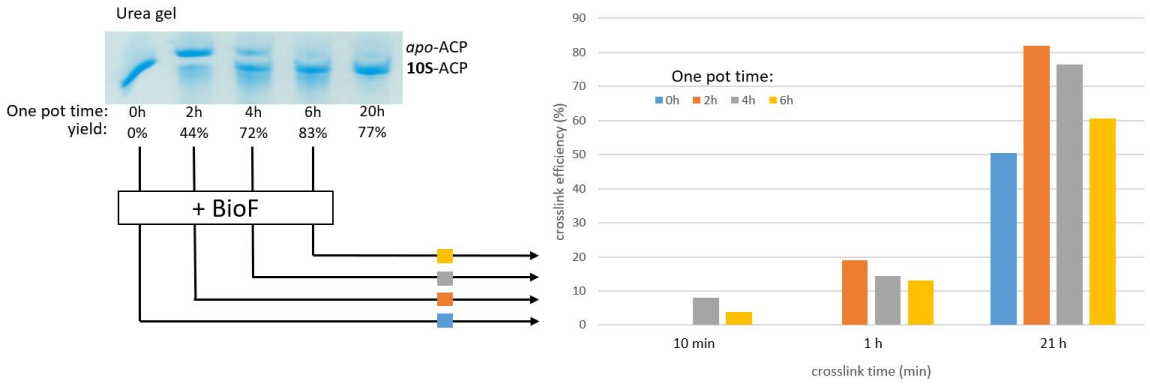


Figure S5.1: workflow optimization for the BSF crosslinker. Four batches of one-pot reaction (color-coded) were set up and BioF was added at certain time point (0, 2 h, 4 h, and 6 h, respectively, for the four batches) to initiate crosslinking reaction. The reactions were then monitored by gel to track the crosslinking yield.

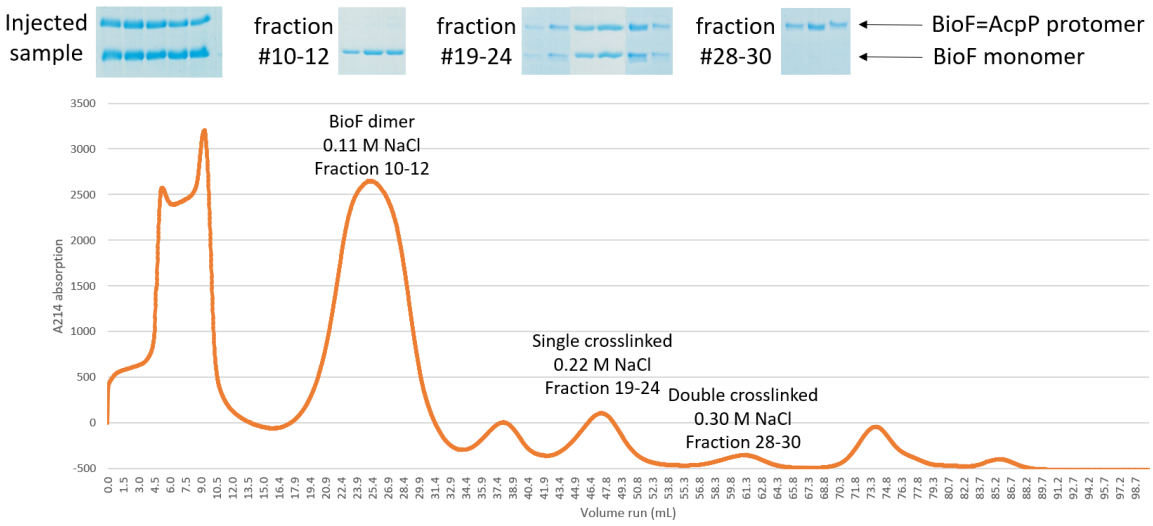


Figure S5.2: Anion exchange column can separate different form of BioF=ACP complex. The chromatogram of a FPLC purification using an anion exchange column is shown. Eluted fractions that contain protein were analyzed by SDS-PAGE and the result is shown on the top. Fraction 19-24 contains single crosslinked complexes (1 protomer + 1 monomer) while fraction 28-30 contains double crosslinked complexes (2 protomers).

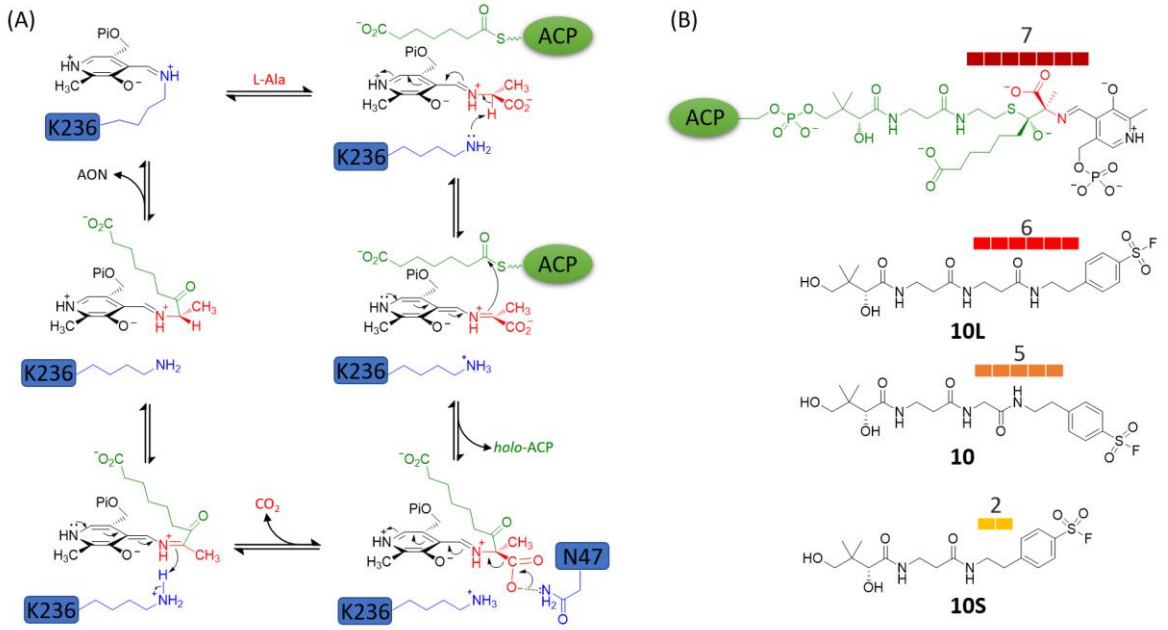


Figure S5.3: Mechanism of BioF catalysis and crosslinker design. (A) proposed reaction mechanism catalyzed by BioF.[29] (B) molecule on the top is an intermediate formed before the release of *holo*-ACP, and the three crosslinker candidates are listed underneath. The bars with number indicate the linear atom count from the PPant amide to the aromatic group. **10L** has the closest chain length to the intermediate.

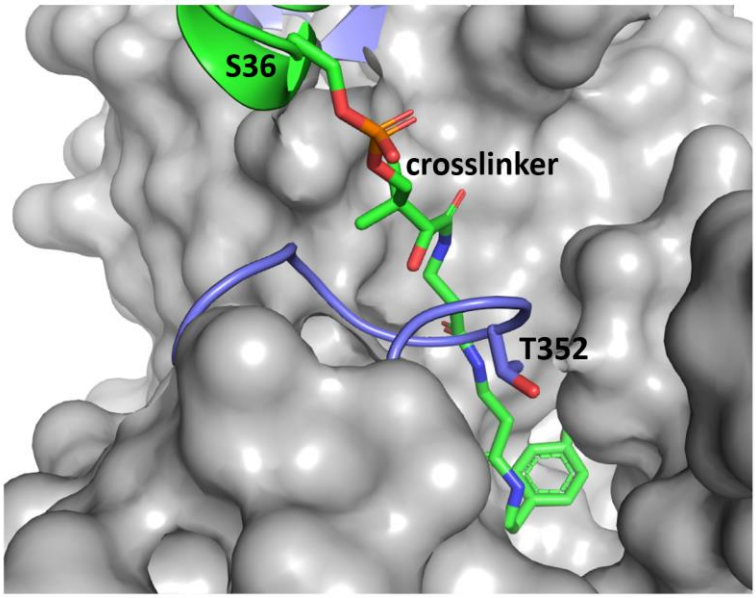


Figure S5.4: The cleft that contains active site and a cover formed by a C-terminal loop. The crosslinker extends from the S36 of ACP (green) to the active site. The key residue (T352) that interacts with the diacid substrate is shown.

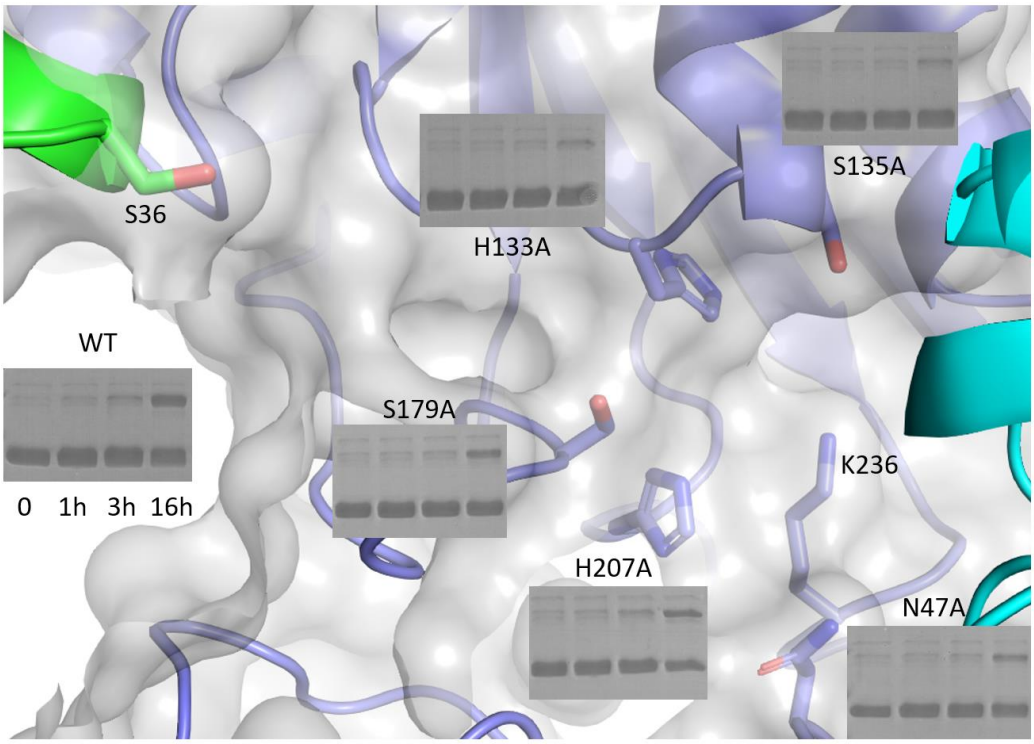


Figure S5.5: Potential crosslinking site in the active site chamber. Alanine mutants of these residues were tested with the crosslinking assay. S135A and H133A have significant reduction in crosslinking yield, which tiny crosslinking bands could only be observed after 16 h of incubation.

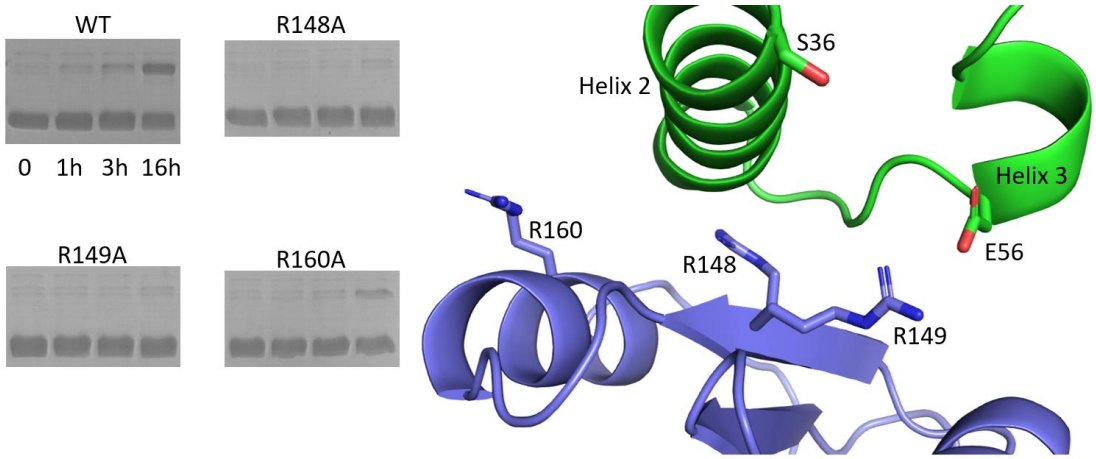


Figure S5.6: Mutagenesis study on interface residues. Alanine mutants of interface arginine residues were tested by crosslinking assay. R148A and R149A almost completely shut down crosslinking reaction.

References

(1) Alexeev, D.; Alexeeva, M.; Baxter, R. L.; Campopiano, D. J.; Webster, S. P.; Sawyer, L. The Crystal Structure of 8-Amino-7-Oxononanoate Synthase: A Bacterial PLP-Dependent, Acyl-CoA-Condensing Enzyme. *J. Mol. Biol.* **1998**, *284* (2), 401–419.

(2) Tong, L. Structure and Function of Biotin-Dependent Carboxylases. *Cell. Mol. Life Sci.* **2013**, *70* (5), 863–891.

(3) Lin, S.; Hanson, R. E.; Cronan, J. E. Biotin Synthesis Begins by Hijacking the Fatty Acid Synthetic Pathway. *Nat. Chem. Biol.* **2010**, *6* (9), 682–688.

(4) Lin, S.; Cronan, J. E. The BioC O-Methyltransferase Catalyzes Methyl Esterification of Malonyl-Acyl Carrier Protein, an Essential Step in Biotin Synthesis. *J. Biol. Chem.* **2012**, *287* (44), 37010–37020.

(5) Manandhar, M.; Cronan, J. E. Proofreading of Noncognate Acyl Adenylates by an Acyl-Coenzyme A Ligase. *Chem. Biol.* **2013**, *20* (12), 1441–1446.

(6) Manandhar, M.; Cronan, J. E. Pimelic Acid, the First Precursor of the Bacillus Subtilis Biotin Synthesis Pathway, Exists as the Free Acid and Is Assembled by Fatty Acid Synthesis. *Mol. Microbiol.* **2017**, *104* (4), 595–607.

(7) Manandhar, M.; Cronan, J. E. A Canonical Biotin Synthesis Enzyme, 8-Amino-7-Oxononanoate Synthase (BioF), Utilizes Different Acyl Chain Donors in Bacillus Subtilis and Escherichia Coli. *Appl. Environ. Microbiol.* **2018**, *84* (1), e02084-17.

(8) Roujeinikova, A.; Baldock, C.; Simon, W. J.; Gilroy, J.; Baker, P. J.; Stuitje, A. R.; Rice, D. W.; Slabas, A. R.; Rafferty, J. B. X-Ray Crystallographic Studies on Butyryl-ACP Reveal Flexibility of the Structure around a Putative Acyl Chain Binding Site. *Struct. Lond. Engl. 1993* **2002**, *10* (6), 825–835.

(9) Cronan, J. E. The Chain-Flipping Mechanism of ACP (Acyl Carrier Protein)-Dependent Enzymes Appears Universal. *Biochem. J.* **2014**, *460* (2), 157–163.

(10) Mindrebo, J. T.; Misson, L. E.; Johnson, C.; Noel, J. P.; Burkart, M. D. Activity Mapping the Acyl Carrier Protein: Elongating Ketosynthase Interaction in Fatty Acid Biosynthesis. *Biochemistry* **2020**, *59* (38), 3626–3638.

(11) Sztain, T.; Bartholow, T. G.; Lee, D. J.; Casalino, L.; Mitchell, A.; Young, M. A.; Wang, J.; McCammon, J. A.; Burkart, M. D. Decoding Allosteric Regulation by the Acyl Carrier Protein. *Proc. Natl. Acad. Sci.* **2021**, *118* (16), e2025597118.

(12) Misson, L. E.; Mindrebo, J. T.; Davis, T. D.; Patel, A.; McCammon, J. A.; Noel, J. P.; Burkart, M. D. Interfacial Plasticity Facilitates High Reaction Rate of *E. Coli* FAS Malonyl-CoA:ACP Transacylase, FabD. *Proc. Natl. Acad. Sci.* **2020**, *117* (39), 24224–24233.

(13) Mindrebo, J. T.; Patel, A.; Kim, W. E.; Davis, T. D.; Chen, A.; Bartholow, T. G.; La Clair, J. J.; McCammon, J. A.; Noel, J. P.; Burkart, M. D. Gating Mechanism of Elongating β-Ketoacyl-ACP Synthases. *Nat. Commun.* **2020**, *11* (1), 1727.

(14) Milligan, J. C.; Lee, D. J.; Jackson, D. R.; Schaub, A. J.; Beld, J.; Barajas, J. F.; Hale, J. J.; Luo, R.; Burkart, M. D.; Tsai, S.-C. Molecular Basis for Interactions between an Acyl Carrier Protein and a Ketosynthase. *Nat. Chem. Biol.* **2019**, *15* (7), 669–671.

(15) Nguyen, C.; Haushalter, R. W.; Lee, D. J.; Markwick, P. R. L.; Bruegger, J.; Caldara-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* (7483), 427–431.

(16) Dodge, G. J.; Patel, A.; Jaremko, K. L.; McCammon, J. A.; Smith, J. L.; Burkart, M. D. Structural and Dynamical Rationale for Fatty Acid Unsaturation in Escherichia Coli. *Proc. Natl. Acad. Sci.* **2019**, *116* (14), 6775–6783.

(17) Agarwal, V.; Lin, S.; Lukk, T.; Nair, S. K.; Cronan, J. E. Structure of the Enzyme-Acyl Carrier Protein (ACP) Substrate Gatekeeper Complex Required for Biotin Synthesis. *Proc. Natl. Acad. Sci.* **2012**, *109* (43), 17406–17411.

(18) Shannon, D. A.; Gu, C.; McLaughlin, C. J.; Kaiser, M.; van der Hoorn, R. A. L.; Weerapana, E. Sulfonyl Fluoride Analogues as Activity-Based Probes for Serine Proteases. *Chembiochem Eur. J. Chem. Biol.* **2012**, *13* (16), 2327–2330.

(19) Steinkopf, W. Über Aromatische Sulfofluoride. *J. Für Prakt. Chem.* **1927**, *117* (1), 1–82.

(20) Dong, J.; Krasnova, L.; Finn, M. G.; Sharpless, K. B. Sulfur(VI) Fluoride Exchange (SuFEx): Another Good Reaction for Click Chemistry. *Angew. Chem. Int. Ed.* **2014**, *53* (36), 9430–9448.

(21) Martín-Gago, P.; Olsen, C. A. Arylfluorosulfate-Based Electrophiles for Covalent Protein Labeling: A New Addition to the Arsenal. *Angew. Chem. Int. Ed.* **2019**, *58* (4), 957–966.

(22) Narayanan, A.; Jones, L. H. Sulfonyl Fluorides as Privileged Warheads in Chemical Biology. *Chem. Sci.* **2015**, *6* (5), 2650–2659.

(23) 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 Chem. Biol.* **2006**, *1* (11), 687–691.

(24) Ishikawa, F.; Haushalter, R. W.; Lee, D. J.; Finzel, K.; Burkart, M. D. Sulfonyl 3-Alkynyl Pantetheinamides as Mechanism-Based Cross-Linkers of Acyl Carrier Protein Dehydratase. *J. Am. Chem. Soc.* **2013**, *135* (24), 8846–8849.

(25) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44–46.

(26) D. Davis, T.; M. Michaud, J.; D. Burkart, M. Active Site Labeling of Fatty Acid and Polyketide Acyl-Carrier Protein Transacylases. *Org. Biomol. Chem.* **2019**, *17* (19), 4720–4724.

(27) Du, D.; Katsuyama, Y.; Horiuchi, M.; Fushinobu, S.; Chen, A.; Davis, T. D.; Burkart, M. D.; Ohnishi, Y. Structural Basis for Selectivity in a Highly Reducing Type II Polyketide Synthase. *Nat. Chem. Biol.* **2020**, *16* (7), 776–782.

(28) Mindrebo, J. T.; Chen, A.; Kim, W. E.; Re, R. N.; Davis, T. D.; Noel, J. P.; Burkart, M. D. Structure and Mechanistic Analyses of the Gating Mechanism of Elongating Ketosynthases. *ACS Catal.* **2021**, *11* (12), 6787–6799.

(29) Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. Mechanism of 8-Amino-7-Oxononanoate Synthase: Spectroscopic, Kinetic, and Crystallographic Studies. *Biochemistry* **2000**, *39* (3), 516–528.

(30) Zheng, Q.; Woehl, J. L.; Kitamura, S.; Santos-Martins, D.; Smedley, C. J.; Li, G.; Forli, S.; Moses, J. E.; Wolan, D. W.; Sharpless, K. B. SuFEx-Enabled, Agnostic Discovery of Covalent Inhibitors of Human Neutrophil Elastase. *Proc. Natl. Acad. Sci.* **2019**, *116* (38), 18808–18814.

(31) Liu, H.; Naismith, J. H. An Efficient One-Step Site-Directed Deletion, Insertion, Single and Multiple-Site Plasmid Mutagenesis Protocol. *BMC Biotechnol.* **2008**, *8* (1), 91.