UNIVERSITY OF CALIFORNIA SAN DIEGO

# Elucidation and Engineering of the Protein-Protein Interactions in Carrier Protein Dependent Pathways

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

in

Chemistry

by

Thomas G. Bartholow

Committee in charge:

Professor Michael D. Burkart, Chair

Professor Patricia A. Jennings

Professor Stanley J. Opella

Professor Charles L. Perrin

Professor Larissa M. Podust

2021

The Dissertation of Thomas G. Bartholow is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

# DEDICATION

To my family who has always supported and sometimes tolerated me

To my friends who continued to challenge me and keep me curious

To my parents for support and an inspiration for creativity

To my brother who gave me a path to aspire to

To the various dogs, cats, and rabbits who ensured my homes were never lonely

EPIGRAPH

Let’s think the unthinkable,

Let’s do the undoable.

Let us prepare to grapple with the ineffable itself,

And see if we may not eff it after all.

-Douglas Adams

TABLE OF CONTENTS

Dissertation Approval Page ..... iii
Dedication ..... iv
Epigraph ..... v
Table of Contents ..... vi
List of Figures and Tables ..... viii
Acknowledgements ..... xi
Vita ..... xii
Abstract of the Dissertation ..... xiv
Chapter 1 Elucidation of transient protein-protein interactions within carrier protein-dependent biosynthesis ..... 1
1.1 Introduction ..... 1
1.2 Results ..... 3
1.3 Discussion ..... 12
1.4 References ..... 15
1.5 Methods ..... 19
1.6 Supplemental Information ..... 28
Chapter 2 Control of unsaturation in de novo fatty acid biosynthesis by FabA ..... 60
2.1 Introduction ..... 60
2.2 Results ..... 63
2.3 Discussion ..... 69
2.4 References ..... 73
2.5 Methods ..... 77
2.6 Supplemental Information ..... 83
Chapter 3 Protein-protein Interaction based substrate control in the E. coli octanoic acid transferase, LipB ..... 98
3.1 Introduction ..... 98
3.2 Results ..... 100
3.3 Discussion ..... 105
3.4 References ..... 108
3.5 Methods ..... 110
3.6 Supplemental Information ..... 115
Chapter 4 The Effect of Substrate on the Protein-Protein Interactions of Acyl Carrier Protein Reductases ..... 136
4.1 Introduction ..... 136
4.2 Results ..... 138

4.3 Discussion ..... 146
4.4 References ..... 148
Chapter 5 Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production ..... 152
5.1 Introduction ..... 152
5.2 Results ..... 154
5.3 Discussion ..... 162
5.4 References ..... 165
5.5 Methods ..... 168
5.6 Supplemental Information ..... 171
Chapter 6 The Protein-protein interactions responsible for Ketosynthase specificity ..... 188
6.1 Introduction ..... 188
6.2 Results ..... 189
6.3 Discussion ..... 195
6.4 References ..... 197
Chapter 7 Possible evolutionary link and remaining acyl transferase activity of *E. coli* malonyl-CoA: acyl-carrier protein transacylase FabD ..... 199
7.1 Introduction ..... 199
7.2 Results ..... 200
7.3 Discussion ..... 210
7.4 References ..... 212

LIST OF FIGURES

Figure 1.1 The reactions and partners of saturated fatty acid biosynthesis ..... 2
Figure 1.2 1H-15N HSQC titration of C8-AcpP with FabF ..... 3
Figure 1.3 Docking workflow and generated models ..... 5
Figure 1.4 Combined data of AcpP interaction with FAB partner enzymes ..... 12
Figure 2.1 The type II fatty acid biosynthesis cycle in *E. coli* ..... 62
Figure 2.2 The chemical shift perturbations (CSPs) of C6, C8, and C10 after titration to saturation with FabA. ..... 63
Figure 2.3 Identifying structural features which facilitate specificity ..... 67
Figure 2.4 The poses of C10-AcpP mapped onto the AcpP-FabA model ..... 68
Figure 2.5 Detailed analysis of the energy landscapes of FabA docking ..... 71
Figure 3.1 Fatty acid biosynthetic cycle of E. coli and LipB transferase activity ..... 99
Figure 3.2 Titration NMR to probe the substrate selectivity of LipB ..... 101
Figure 3.3 Structural analysis of the AcpP•LipB surface ..... 106
Figure 4.1 The Fatty Acid Biosynthetic cycle ..... 138
Figure 4.2 Alignment of FabG homologs and structural effects of NADH substrate ..... 139
Figure 4.3 Alignment of FabI homologs and structural effects of NADH substrate ..... 140
Figure 4.4 The chemical shift perturbations resulting from titration of C8-AcpP with apo and NAD+ bound FabG ..... 143
Figure 4.5 The chemical shift perturbations resulting from titration of C8-AcpP with apo and NAD+ bound FabI ..... 145
Figure 5.1 Medium-chain fatty acid biosynthesis in Escherichia coli ..... 154
Figure 5.2 Acinetobacter baylyi thioesterase (AbTE) product profile and interface with Escherichia coli ACP ..... 156
Figure 5.3 Saturated fatty acid production of E. coli expressing Acinetobacter baylyi thioesterase (AbTE) mutants ..... 158
Figure 5.4 Protein expression, time course analysis, and MCFA detection in the producer supernatant by MCFA sensor ..... 161
Figure 5.5 Solution NMR analysis of molecular interactions between A. baylyi thioesterase (AbTE) and E. coli AcpP ..... 162
Figure 6.1 Titration of holo- and C8-AcpP with FabH ..... 190
Figure 6.2 Titrations of C8- and C10-AcpP with FabF and FabB ..... 193
Figure 6.3 Titrations of C10:1-AcpP with FabF and FabB ..... 194
Figure 6.4 Comparing a C8-AcpP•apo-FabF titration to a C8-AcpP•cerulenin-FabF titration ..... 195
Figure 7.1 Fatty acid biosynthesis in Escherichia coli ..... 202
Figure 7.2 Trapping E. coli malonyl-coA ACP transacylase acyl-carrier protein complex with covalent pantetheine probes ..... 203

Figure 7.3 Sequence alignment of FabD and homologs ..... 204

Figure 7.4 In vitro deacylation of E. coli acyl-carrier protein by malonyl-CoA ACP transacylase ..... 205

Figure 7.5 Continuous fluorescent assay to monitor E. coli malonyl-CoA ACP transacylase deacylation activity ..... 206

Figure 7.6 Titration NMR to probe the dual functions of the interface of FabD ..... 207

Figure 7.7 Most stable docked orientation of the C8-AcpP●FabD and holo-AcpP●FabD ..... 209

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Professor Michael D. Burkart. It has been a great privilege to have the freedom to explore my work and this beautiful enzyme system during my time. While being available for discussion or advisement, Mike still gave me the space to explore my work on my own. It was only with this space that I was able to develop as a scientist and complete the lofty goals I set for myself. Furthermore, the Burkart lab has been a tremendous place to spend my time as a researcher, the environment of camaraderie may be the secret that makes it so successful.

I would like to thank my committee members, Professor Patricia A. Jennings, Professor Charles L. Perrin, Professor Larissa M. Podust, and Professor Stanley J. Opella. Their feedback and guidance were critical to my development and I was truly blessed to have their attention. I would especially like to thank Professor Opella, his guidance and patronage of NMR at UCSD has truly been essential to the technique. As well as Professor Opella I would like to thank the group he has fostered. During meetings, their expertise was essential and without them I would have rarely had an expert eye on my science. Most especially Xuemei Huang, who maintained the instruments and gave the expertise that were essential to my work.

I have many friends that I need to thank for how much I've enjoyed my time in graduate school. To begin my long-time roommates Ryan, Jeff, and Ben for so much fun during a stressful time. Furthermore, all of the friends who have been so dear along the way: Srihita, Terra, Eunice, and Dolan. As well a special call out to the people who tolerated my hobbies for climbing and D&D. I also want to extend my thanks to friends back home from my Undergraduate and all the way back to childhood: Alex, Grant, Harith, Minesh, Jake, Brian, Thamo, and Allen.

I have been lucky to grow up with a large family. There just isn't enough time to thank everyone in particular here. But as well as my closest family: My father John, my mother Dawn, and my brother Johnny, I want to thank the Grandparents, aunts, uncles, and cousins living and passed who have made my life such a pleasure.

I finally want to thank any collaborators who I have worked with during my studies. In addition, the Burkart lab members who I did not work directly with but who I was so lucky to work with and know.

**Chapter 1**, is in full a reprint of the manuscript: Thomas G. Bartholow, Terra Sztain, Ashay Patel, D. John Lee, Megan A. Young, Ruben Abagyan, Michael D. Burkart “Elucidation of transient protein-protein interactions within carrier protein-dependent biosynthesis” Communications Biology 2020. The dissertation author is the primary author of the manuscript.

**Chapter 2**, is in full a reprint of the manuscript: Thomas G. Bartholow, Terra Sztain, Megan A Young, D. John Lee, Tony D. Davis, Ruben Abagyan, Michael D. Burkart, “Control of unsaturation in de novo fatty acid biosynthesis by FabA.” *Biochemistry*, (2021). The dissertation author is the primary author of the manuscript.

**Chapter 3**, is in full a reprint of the manuscript :Thomas G. Bartholow, Terra Sztain, Megan A Young, Tony D. Davis, Ruben Abagyan, Michael “Protein-protein Interaction based substrate control in the E. coli octanoic acid transferase, LipB.” D. Burkart, In review *RSC Chemical Biology*, (2021). The dissertation author is the primary author of the manuscript.

**Chapter 4**, is in print of the manuscript in preparation: Thomas G. Bartholow, Megan A Young, Michael D. Burkart “Atomic Details of the Cofactor Priming and Protein-Protein Interaction of E. coli Acyl Carrier Protein Reductases.” The dissertation author is the primary author of the manuscript.

**Chapter 5**, is in full a reprint of the manuscript: Stephen Sarria, Thomas G. Bartholow, Adam Verga, Michael D. Burkart, and Pamela Peralta-Yahya “Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production.” ACS Synthetic Biology, (2018). The dissertation author is the primary coauthor of the manuscript in collaboration with Dr. Stephen Sarria.

**Chapter 6** is in combination of two manuscripts in preparation Thomas G. Bartholow, Terra Sztain, Megan A. Young, Tony D. Davis Michael D. Burkart. “Selectivity in the initiating ketosynthase, FabH, of E. coli fatty acid biosynthesis.” And: Thomas G. Bartholow, Jeffery T. Mindrebo, Ashay Patel, Woojoo E. Kim, Aochiu Chen, Megan A. Young, Michael D. Burkart “A Protein-Protein Interaction based control step for the specificity of Unsaturated Fatty Acid elongation.” The dissertation author is the primary author of the manuscript.

**Chapter 7** is a print of the manuscript in preparation Thomas G. Bartholow, Tony D. Davis, Megan A. Young, Ruben Abagyan, Michael D. Burkart. “Possible evolutionary link and remaining acyl transferase activity of E. coli malonyl-CoA: acyl-carrier protein transacylase FabD.” The dissertation author is the primary author of the manuscript.

VITA

| | |
|---|---|
| 2014 | Bachelor of Science, University of Wisconsin Eau Claire |
| 2016 | Master of Science, University of California San Diego |
| 2021 | Doctor of Philosophy, Chemistry, University of California San Diego |

Publications

"Possible evolutionary link and remaining acyl transferase activity of E. coli malonyl-CoA: acyl-carrier protein transacylase FabD."
Thomas G. Bartholow, Tony D. Davis, Terra Sztain, Megan A. Young, Ruben Abagyan, Michael D. Burkart

"Selectivity in the initiating ketosynthase, FabH, of E. coli fatty acid biosynthesis"
Thomas G. Bartholow, Terra Sztain, Megan A. Young, Tony D. Davis Michael D. Burkart

"A Protein-Protein Interaction based control step for the specificity of Unsaturated Fatty Acid elongation."
Thomas G. Bartholow, Jeffery T. Mindrebo, Ashay Patel, Woojoo E. Kim, Aochiu Chen, Megan A. Young, Michael D. Burkart

"Protein-protein Interaction based substrate control in the E. coli octanoic acid transferase, LipB."
Thomas G. Bartholow, Terra Sztain, Megan A Young, Tony D. Davis, Ruben Abagyan, Michael D. Burkart, In review *RSC Chemical Biology*, (2021) in review

"Control of unsaturation in de novo fatty acid biosynthesis by FabA."
Thomas G. Bartholow, Terra Sztain, Megan A Young, D. John Lee, Tony D. Davis, Ruben Abagyan, Michael D. Burkart, *JACS*, (2021) in review

"Decoding the allosteric nature of acyl carrier proteins."
Terra Sztain, Thomas G. Bartholow, D. John Lee, Lorenzo Casalino, Andrew Mitchell, Megan A. Young, Jianing Wang, J. Andrew McCammon, Michael D. Burkart, *PNAS*, (2021) Accepted

"Elucidation of transient protein-protein interactions within carrier protein-dependent biosynthesis."
Thomas G. Bartholow, Terra Sztain, Ashay Patel, D. John Lee, Megan A. Young, Ruben Abagyan, Michael D. Burkart, *Communications Biology*, (2020)

"Gating mechanism of elongating-ketoacyl-ACP synthases."
Jeffrey T. Mindrebo, Ashay Patel, Thomas G. Bartholow, Woojoo E. Kim, Tony D. Davis, Aochiu Chen, James J. La Clair, J. Andrew McCammon, Joseph P. Noel, & Michael D. Burkart, *Nature Communications*, (2019)

"Shifting the Hydrolysis Equilibrium of Substrate Loaded Acyl Carrier Proteins."
Terra Sztain, Thomas G. Bartholow, J. Andrew McCammon, and Michael D. Burkart, *Biochemistry*, (2019)

"Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production."
Stephen Sarria, Thomas G. Bartholow, Adam Verga, Michael D. Burkart, and Pamela Peralta-Yahya, *ACS Synthetic Biology*, (2018)

ABSTRACT OF THE DISSERTATION

Elucidation and Engineering of the Protein-Protein Interactions in Carrier Protein Dependent Pathways

by

Thomas G. Bartholow
Doctor of Philosophy in Chemistry
University of California San Diego, 2021
Professor Michael D. Burkart, Chair

Protein-protein interactions are a fundamental principle of biochemistry and nature itself. This class of interactions span well understood principles such as electrostatic or Van der Waals interactions to complicated principles of protein dynamics and allostery. In this work I have applied my knowledge of these principles to a system I felt was perfect for a detailed characterization. Fatty acid biosynthesis is a simple system at a distance, with proteins whose function has been known for decades and relatively simple chemistries. However, this shallow observation betrays the subtle complications of the system. In reality, fatty acids are created through an elegant choreography of protein-protein interactions. With the central player, the acyl carrier protein, acting as messenger throughout. Like the cycle it is central to, the ACP is a simple protein at only 4 helices and under 100 amino acids. But in these studies, we have uncovered how such a small protein is able to facilitate this elegant system.

**Chapter 1.** Fatty acid biosynthesis (FAB) is an essential and highly conserved metabolic pathway. In bacteria, this process is mediated by an elaborate network of protein•protein interactions (PPIs) involving a small, dynamic acyl carrier protein that interacts with dozens of other partner proteins (PPs). These PPIs have remained poorly characterized due to their dynamic and transient nature. Using a combination of solution-phase NMR spectroscopy and protein-protein docking simulations, we report a comprehensive residue-by-residue comparison of the PPIs formed during FAB in Escherichia coli. This technique describes

and compares the molecular basis of six discrete binding events responsible for E. coli FAB and offers insights into a method to characterize these events and those in related carrier protein-dependent pathways.

**Chapter 2**. Carrier protein dependent biosynthesis provides a thiotemplated format for the production of natural products. Within these pathways, many reactions display exquisite substrate selectivity, a regulatory framework proposed to be con-trolled by protein-protein interactions (PPIs). In Escherichia coli, unsaturated fatty acids are generated within the de novo fatty acid synthase by a chain length-specific interaction between the acyl carrier protein AcpP and the isomerizing dehydratase FabA. To evaluate PPI-based control of reactivity, interactions of FabA with AcpP bearing multiple seques-tered substrates were analyzed through NMR titration and guided high-resolution docking. Through a combination of quantitative binding constants, residue-specific perturbation analysis, and high-resolution docking, a model for sub-strate control via PPIs has been developed. The in silico results illuminate the mechanism of FabA substrate selectivity and provide a structural rationale with atomic detail. Helix III positioning in AcpP communicates sequestered chain length identity recognized by FabA, demonstrating a powerful strategy to regulate activity by allosteric control. These studies broadly illuminate carrier protein dependent pathways and offer an important consideration for future inhibi-tor design and pathway engineering.

**Chapter 3**. Lipoic acid is an essential cofactor produced in all organisms by diverting octanoic acid derived as an intermediate of type II fatty acid biosynthesis. In bacteria, octanoic acid is transferred from the acyl carrier protein (ACP) to the lipoylated target protein by the octanoyltransferase LipB. LipB has a well-documented sub-strate selectivity, indicating a mechanism of octanoic acid recog-nition. The present study reveals the precise protein-protein interactions (PPIs) responsible for this selectivity in Escherichia coli through a combination of solution-state protein NMR titra-tion with high-resolution docking of the experimentally exam-ined substrates. We examine the structural changes of sub-strate-bound ACP and determine the precise geometry of the LipB interface. Thermodynamic effects from varying substrates were observed by NMR, and steric occlusion of docked models indicates how LipB interprets proper substrate

identity via allo-steric binding. This study provides a model for elucidating how substrate identity is transferred through the ACP structure to regulate activity in octanoyl transferases.

**Chapter 4**. Carrier Protein Reductases are an essential component of Fatty Acid and Polyketide synthases. The *E. coli* Enoylreductase and Ketoreductase perform the only reactions in FAS which do not have a redundant enzyme. As such they have been of high interest for the design of inhibitors, but development has been limited by poor structural information and high homology to human homologs. In this study an analysis is first performed to observe the homology of *E. coli* reductases, utilizing recent models to focus the on the AcpP•Reductase interface. Next $^{15}$N-HSQC NMR titrations were performed to compare substrate bound- and apo-Reductase Protein-Protein interactions. This demonstrated previously hypothesized differences in the interactions of reductases when they have no substrate loaded. The NMR data was compared to the analysis of structural differences and homology to create a picture of how and why the interface is different. This combination of data gives a first look at the impact of substrate on reductase function as well as a detailed analysis of the reductase structure.

**Chapter 5**. Medium-chain fatty acids (MCFAs) are key intermediates in the synthesis of medium-chain chemicals including α-olefins and dicarboxylic acids. In bacteria, microbial production of MCFAs is limited by the activity and product profile of fatty acyl-ACP thioesterases. Here, we engineer a heterologous bacterial medium-chain fatty acyl-ACP thioesterase for improved MCFA production in Escherichia coli. Electrostatically matching the interface between the heterologous medium-chain Acinetobacter baylyi fatty acyl-ACP thioesterase (AbTE) and the endogenous E. coli fatty acid ACP ( E. coli AcpP) by replacing small nonpolar amino acids on the AbTE surface for positively charged ones increased secreted MCFA titers more than 3-fold. Nuclear magnetic resonance titration of E. coli 15N-octanoyl-AcpP with a single AbTE point mutant and the best double mutant showed a progressive and significant increase in the number of interactions when compared to AbTE wildtype. The best AbTE mutant produced 131 mg/L of MCFAs, with MCFAs being 80% of all secreted fatty acid chain lengths after 72 h. To enable the future screening of larger numbers of AbTE variants to further improve MCFA titers, we show that a previously developed G-protein coupled receptor (GPCR)-based MCFA sensor differentially detects MCFAs secreted by E. coli expressing different AbTE variants. This work demonstrates that engineering the interface of heterologous

enzymes to better couple with endogenous host proteins is a useful strategy to increase the titers of microbially produced chemicals. Further, this work shows that GPCR-based sensors are producer microbe agnostic and can detect chemicals directly in the producer microbe supernatant, setting the stage for the sensor-guided engineering of MCFA producing microbes.

**Chapter 6.** Fatty acid biosynthesis is an essential metabolic pathway which uses well established mechanisms for the efficient production of primary metabolic products. These reactions are driven by the iterative elongation of fatty acids to create long carbon chains from repeating simple units. In *E. coli* these reactions are performed by three ketosynthases, each with a unique function. Using $^{15}$N-HSQC NMR titrations the protein-protein interactions which can serve as the first control step in elongation were characterized. The unique interactions which help facilitate each enzymes function are observed, further, the basic question of what interactions are observed was probed through titration of a AcpP with a partner protein bound to an inhibitor which fills the substrate pocket. These data taken together demonstrates the full breadth of interactions and regulations necessary for the efficient and accurate elongation of acyl chains. Gaining further knowledge of what drives the first reaction gives us more opportunities to engineer or inhibit the engines of carrier protein dependent synthases.

**Chapter 7**. Fatty acid and Polyketide synthases are responsible for the creation of essential products in primary and secondary metabolism. In *E. coli* the fatty acid synthase is driven by the FabD transacylase responsible for priming carrier proteins with the malonyl-AcpP necessary to fuel the synthase cycle. However, other species have noted high-homology partners of FabD which can further act as a transacylase for longer chain fatty acids. Herein, we report on the finding of medium and long chain fatty acid transacylase activity of FabD, first through mechanistic probes. The homology to established acyltransferases is explored as an explanation for this newly assigned activity. We follow this study with biophysical characterization of the activity through 15N-HSQC NMR titrations to establish the productive interaction of both holo- and acyl-AcpP with FabD. Finally, high resolution docking was utilized to explore if the newly assigned activity is likely to follow the previously seen crosslinked confirmation. This study serves as a demonstration of how chemical biology and biophysical techniques can be combined to discover new activities and characterize their details. But more importantly this could be an expansion of

our understanding of a well-established and deeply researched metabolic pathway, demonstrating there are still unknown activities in the *E. coli* cell to be discovered with the proper tools.

# Chapter 1 Elucidation of transient protein-protein interactions within carrier protein-dependent biosynthesis

## 1.1 Introduction

Carrier protein dependent synthases are responsible for the biosynthesis of a vast array of molecules, from primary metabolites to complex natural products [1,2]. These are generally organized as type I or type II enzymes, with the type I "megasynthases" containing multiple enzymatic domains and carrier proteins housed as large multi-domain proteins [3]. In contrast, type II synthases exist as discrete proteins that must recognize and associate with one another in solution through an organized choreography of metabolic steps (Fig. 1a). In *E. coli* fatty acid biosynthesis (FAB) more than 25 partner proteins (PPs) are known to functionally bind to the acyl carrier protein (AcpP)(Fig.1b) [4–6], a small, four-helix bundle protein that shuttles intermediates between both fatty acid biosynthetic enzymes and regulatory proteins [7]. AcpP must form specific protein•protein interactions (PPIs) with multiple partners, efficiently chaperoning intermediates through 30-35 discrete enzymatic steps to produce the membrane lipids that maintain homeostasis and facilitate cellular reproduction [8,9]. Simultaneously FAB generates fatty acid intermediates for cofactor biosynthesis and secondary metabolism[10]. This study further illustrates AcpP•PP recognition through unique PPIs with each of the FAB enzyme players (Fig. 1a) while presenting a combinatorial method to characterize these transient interactions, useful for both engineering and inhibitor design.

Throughout the iterative FAB cycle, substrates and intermediates are not only tethered to AcpP through a 4'-phosphopantetheine (PPant) thioester linkage[11,12], but they are sequestered within the hydrophobic pocket of the AcpP helices, protecting intermediates from non-specific reactivity[13–15]. Reactions are controlled through this sequestration and presentation of the substrates appended to 4'-phosphopantetheine, a process termed chain flipping [13,16]. The rapid doubling times and relatively narrow distribution of fatty acid products require an efficient, high fidelity FAB [17-*16*], suggesting that stochastic binding events of AcpP with its binding partners are unlikely. Activity studies and mechanism-based crosslinking experiments have demonstrated that acyl-AcpP binding and enzyme turnover are highly specific (Fig. 2c) [17,18]. A growing body of evidence suggests PPIs play an important role in the mechanism

of chain flipping and, therefore, the processivity of these pathways [21–24]. For example, recent studies have demonstrated that engineering enzyme specificity for a non-native AcpP-dependent enzyme can be accomplished by modifying the PPI residues for improved binding [25]. Even single atom changes in the identity of AcpP-bound cargo have been demonstrated to impart perturbations to the structure of acyl-AcpP [26].

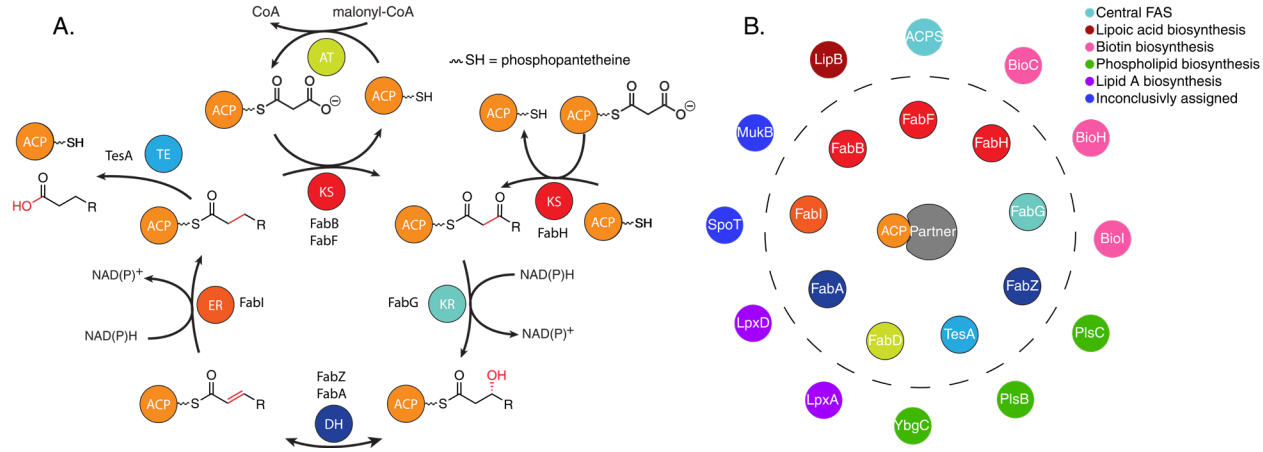



**Figure 1.1 The reactions and partners of saturated fatty acid biosynthesis. A) The FAB elongation cycle. ACP: acyl carrier protein, KS: ketosynthase, KR: ketoreductase, DH: dehydratase, ER: enoylreductase, TE: thioesterase, AT: acyltransferase. TesA is known to interact with AcpP but not a necessary part of *E. coli* FAB B) 21 examples of known AcpP interacting enzymes colored by function, the enzymes within the dotted line are those from FAB, whose color corresponds to the colors of Fig. 1A.**

Here we used $^{1}H$-$^{15}N$ HSQC NMR titration studies to collect residue-by-residue information for six *de novo* FAB partner enzymes to characterize each intrinsic PPI with the *E. coli* AcpP. Experiments were performed to study the interfaces of AcpP with elongating ketosynthases FabB and FabF, reductases FabG and FabI, dehydratase FabA, and thioesterase TesA. These spectroscopic data combined with a combinatorial docking protocol benchmarked with crosslinked structures of AcpP in complex with FabA, FabZ, FabB, and FabF, provide atomic-resolution information on which residues of AcpP mediate each step in iterative *de novo* FAB. This combinatorial method was able to appreciate the unique challenges of modular synthases, with substrate identity effecting carrier protein structure and each enzyme forming unique interactions with the carrier. Due to the high sequence homology of AcpP with carrier proteins from other species [25,21] and polyketide synthases [28], this protocol is expected to extend for characterization of ACP•PP interactions for engineering and drug design across multiple systems.

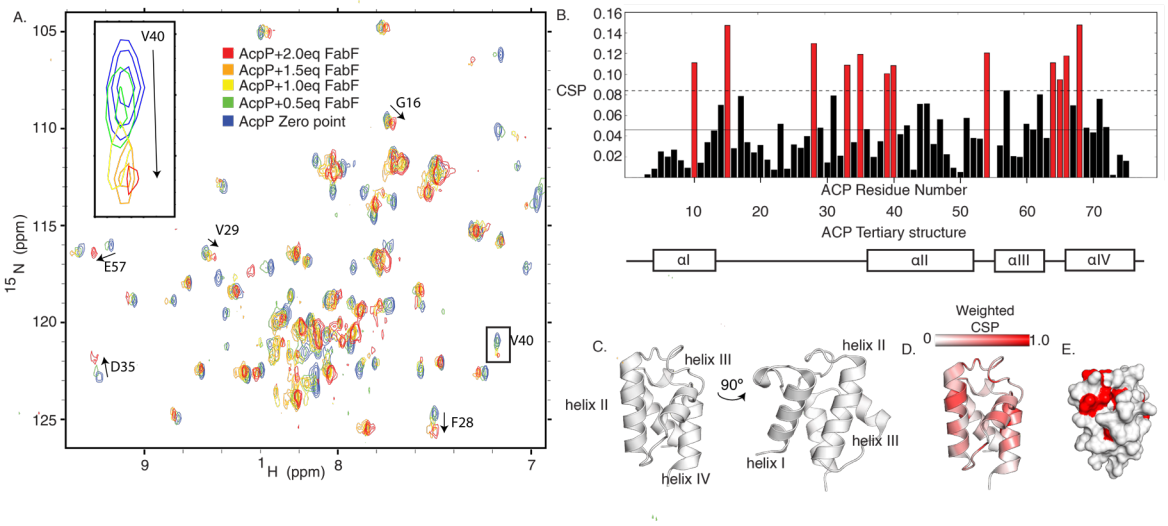



**Figure 1.2 $^{1}$H-$^{15}$N HSQC titration of C8-AcpP with FabF. A) 5 overlaid $^{1}$H-$^{15}$N HSQC spectra of $^{15}$N-C8-AcpP titrated with increasing concentrations of unlabeled FabF. The peak migration of V40 is enlarged as an example of a "titration curve." B) A bar chart of each AcpP residue's CSP with 2.0 molar equivalents of FabF. The mean is shown as a solid line and one standard deviation is the dashed line above. CSPs greater than this value are shown in red. The CSP equation used was: $CSP = \sqrt{\left(\frac{1}{2}\right)[\delta_H^2 + (\alpha \cdot \delta_N)^2}$. An alpha value of 0.2 was used. C) The tertiary structure of *E. coli* AcpP displaying the classical interface of helix II and IV, this orientation will be used throughout the paper when displaying the interacting face. This is rotated 90° to display the side face of the ACP. D) The cartoon structure colored by weighted CSP value, viewed from the interacting face. E. The surface of the AcpP interacting face with the CSPs one standard deviation above the mean colored in red.**

## 1.2 RESULTS

### NMR Titrations reveal dynamic AcpP interface

Previous work has established the utility of $^{1}$H-$^{15}$N HSQC-NMR titrations in the study of rapid and intricate PPIs [29–31]. In this study, uniformly labeled and perduterated $^{15}$N C8-AcpP (octanoyl-AcpP) (Fig. S6, Fig. S11) was subjected to NMR titration using increasing concentrations of unlabeled PPs to detect the residues on AcpP that experience chemical shift migration (Fig. 2a, S1-3). Once saturated with partner enzyme, the extent of peak migration was quantified using the chemical shift perturbation (CSP) calculations (Fig. 2b, S1-3b)[32]. The perturbed regions were then projected onto the amino acid sequence and 3D structure of the protein to identify regions affected by PP binding (Fig. 2c,d,4b). Furthermore, we utilized the TITAN NMR lineshape analysis program[33] to analyze our spectra and obtain thermodynamic and kinetic parameters (Table S9). Previous work has demonstrated the ability of chemical shift

perturbations to identify critical protein-protein interactions in carrier protein-mediated biosynthesis[34–36]. Combining our titration data on FabF, FabI, FabG, and TesA with previous titrations on FabB [34] and FabA [24] allowed us to compare the binding interface on AcpP dictating PP recognition, highlighting distinct AcpP residues involved in specific classes of enzyme binding and internal AcpP residues that are likely important to the mechanistic process of chain flipping.

**Docking to elucidate partner protein binding site**

The $^{1}H$-$^{15}N$ HSQC NMR titrations shown here can provide specific information of the AcpP residues involved in PP binding, however; the residues on the PP that mediate binding cannot be determined by this process. While titrations provide ample evidence of the carrier protein's interacting residues, no information is gained about the residues of the partner protein with which they are interacting. We have recently elucidated the x-ray crystal structures of several FAB enzymes crosslinked to AcpP, including FabA, FabZ, FabB, and FabF. These structures can indicate residues involved in PPIs on both proteins, however, each structure requires prior development of enzyme-specific crosslinking probes, which are not available in all cases. We sought to develop protein-protein docking protocols with Molsoft's ICM software to predict structures of the AcpP•FabI, AcpP•FabG, and AcpP•TesA complexes that have eluded experimental structural characterization [37,38]. Crystal structures of previously crosslinked AcpP-PPs were used to optimize this protocol, described more fully in the methods section. Briefly, it was identified that to accurately recreate complexes it was necessary to produce a water box in which the partner proteins were minimized. Docking simulations were carried out between the x-ray structural model of heptanoyl-AcpP-C7 (PDB 2FAD), to which a methylene was added to the acyl chain simulate C8-AcpP, and crystal structural models of binding partner proteins from which cofactors had been removed. Using expanded calculations to assist the general docking protocol (Fig. 3a), we were able to recapitulate crosslinked structure interfaces (Fig. 3b, Table S1) to sub 7 Å RMSD for the complex and sub 2 Å RMSD at the interface. Crosslinked structures were used in benchmarking as they give a learning set to examine which protocols and docking methods perform well, also demonstrating the ability of our docking method to appreciate ACP•PP interfaces. However, it must be noted that the comparison is imperfect, with the docked structures and NMR

representing interactions in solution while the crosslinked structures are crystallized and covalently bound in a catalytic conformation. For example, the AcpP structure 2FAD and the AcpP crosslinked to

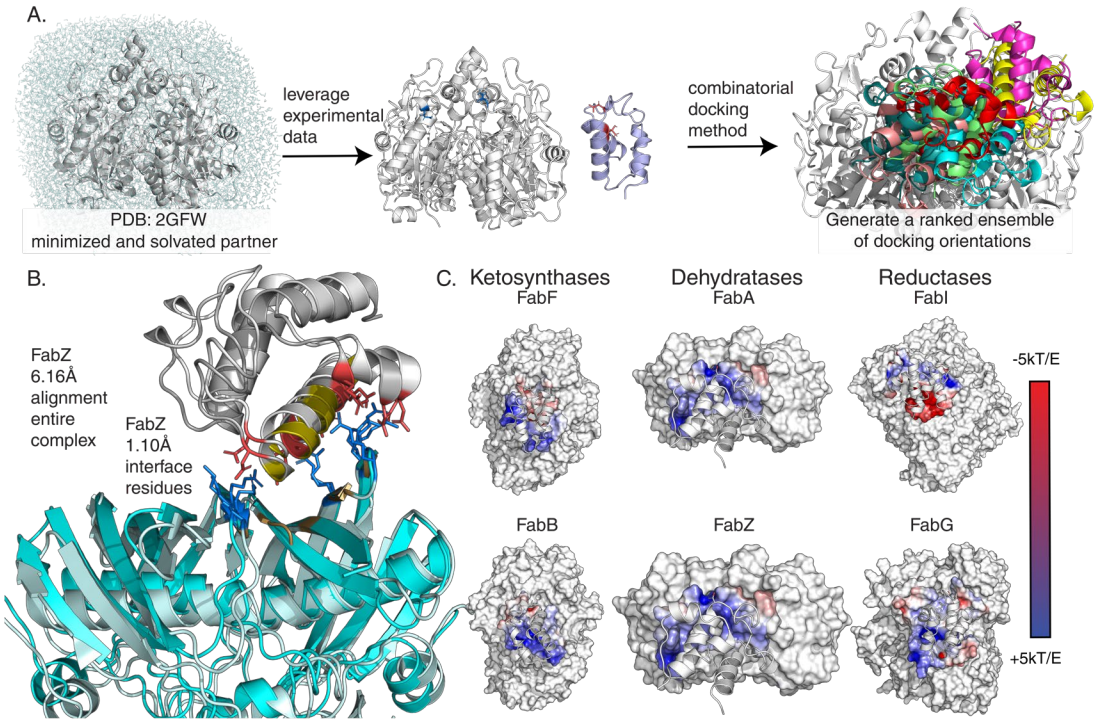



**Figure 1.3 Docking workflow and generated models. A) A Workflow showing water box generation and minimizations of FabZ (PDB: 6N3P) in purple docked to AcpP (PDB: 2FAD) in grey. The optimized docking protocol is described in more detail in the methods section. It was found to be essential to recapitulate the interface within 3 Å RMSD as shown in B) Comparison of the interfaces of the docked model (FabZ: light blue and ACP: dark gray) to the crosslinked crystal structure (FabZ: pale cyan and ACP: light gray). Charged residues within 5 Å of the interfaces are displayed as sticks with the negatively charged (red) and positively charged (blue). Hydrophobic residues are colored gold. C) The surfaces of 6 PPs from 3 families of enzymes with electrostatics on partner enzymes shown within 5 Å of the bound AcpP. Larger versions of these images are presented in Fig. S4.**

FabA have a ~2Å RMSD. Furthermore, it should be noted that the crosslinked and *apo* partner proteins have differing structural similarity. With the FabF structure 1.3Å RMSD between crosslinked[39] and uncrosslinked[40], FabB 3.5Å RMSD between crosslinked[34] and uncrosslinked[41], and FabA 4.6Å RMSD between crosslinked[35] and uncrosslinked[42] structures. The developed protocols were subsequently used to determine the binding interfaces of AcpP•FabI, AcpP•FabG, and AcpP•TesA (Fig. 3c, Fig. S4, Table S2, Fig. S4,S5) in conjunction with the NMR data. This methodology provides valuable context for matching the AcpP interactions to the PP structures, and the breadth of previously reported AcpP and PP activity and mutagenesis studies enable further validation of predicted AcpP•PP against past mutagenic experiments.

**A combinatorial method to characterize modular synthase PPIs**

To judge the ability of a combined NMR and docking method to accurately predict the structures of interacting enzymes, docking was expanded to include CSP information. The interface residues identified through CSPs were given as focus residues for AcpP binding and known interacting residues specified below were given for partner proteins. Docked models of the enzymes FabA, FabB, and FabF were compared with and without known residues supplied. All the enzymes tested had docked orientations which allowed for chain flipping of the substrate based on the position of Ser 36 and known active sites of the partners, demonstrating the ability of informed calculations to filter out nonproductive complexes. The average RMSD between the docked model and crosslinked crystal structures was: 9.29 Å in the informed calculation of FabF•AcpP binding and 20.08 Å (Table S7) without experimental knowledge. FabB models were 5.6 Å RMSD in the informed model and 10.5 Å in uninformed, and FabA models were 9.98 Å RMSD in the informed calculation and 9.89 Å in the uninformed. These data demonstrate that while docking alone was able to recreate the crosslinked structure, care must be taken to ensure the models are relevant. Leveraging experimental data ensured greater confidence in the models while yielding a suite of interface structures of relevant interacting orientations. To evaluate the ability of widely available online servers to recreate AcpP•partner protein interfaces the AcpP•FabF complex was docked using the Cluspro[43–45], HADDOCK[46,47], and Rosie[48,49] servers (Table S7,S8). It should be noted that with informed residues and properly prepared structures online servers perform well. The structures used below are taken from the most stable docked orientation for simplicity, though the full suite of orientations remains relevant. However, these methods open the door to more studies of possible preliminary encounter complex states which may facilitate the fully bound form seen in crosslinking.

**Ketosynthases: FabF and FabB**

Elongating ketosynthases iteratively extend acyl-AcpP by two carbon units using malonyl-AcpP as a carbon source via a decarboxylative thia-Claisen condensation (Fig. 1a) [50]. C8-AcpP was titrated with increasing concentrations of the FabF ketosynthase (Fig. 2a, Fig. S9) and compared to recently published data of C8-AcpP titrated with FabB[34]. An octanoyl acylation state was selected to maintain consistency with prior work [34,35] and was utilized for all titrations in this study. Upon titration with FabF, residues from C8-

AcpP display chemical shift perturbations beginning at the end of helix I, with very little interaction on loop I until nearly the beginning of helix II. Helix II displays perturbations throughout until there is a small loss at the end of helix II, signals assigned to residues nearly throughout C8-AcpP were perturbed until the end of helix IV where the perturbations drop off. The largest CSPs, Residues with CSPs greater than one standard deviation from the mean, in the FabF titration (Table S3) included I10 and L15 on helix 1 and F28 on loop 1 (Fig. 2b); D35, T39, V40 on helix 2; I54 on loop 2; and T64, V65, Q66, and A68 on helix 3. D35 and T39, the charged or polar residues which lie along the interface, appeared within interacting distance of N56' and Q63' (residue designators for the partner proteins in the complexes will be denoted by primes) [34,35,51] (Fig. 4c,d). Surprisingly, a large number of these residues (I10, L15, F28, I54, T64, V65, and A68) are located within the acyl pocket or far from the interface yet show large perturbations. We hypothesized that these interior perturbations represent internal hydrophobic rearrangements that occur upon chain flipping during the binding event. Titan analysis calculated a $K_d$ of 8.3 ± 9.8 μM with a $k_{off}$ of 3512 ± 3341 $s^{-1}$ and an approximately one to one stoichiometry(Table S9, Fig. S12).

FabB performs the same ketosynthase reaction as FabF, but performs the first unsaturated elongation step, making a case study in specificity [52]. The NMR titrations were previously performed, but the data will be restated here for comparison. Overall signals from the AcpP•FabB titration display a slightly less broad set of CSPs than those from the AcpP•FabF titration, starting with little perturbation until the top of helix I. Perturbations continue once more at the end of loop I with more sparse interactions on helix II, without perturbation at residues 40 and 41 as well as a drop off in perturbation at the end of helix II. There are 2 CSPs on loop II and perturbations span most of helix III and loop III. Finally, helix IV sees sparse perturbations at the top of helix IV. In FabB the most perturbed residues were L15 of helix 1; D35, S36, L37, D38, and L42 on helix 2; E60 on helix 3; and T63 on loop 3. FabB's effect on AcpP is largest on many acidic residues at the interface, D35, D38, and E60 are all among the most perturbed(Fig. 4). Residues D35 interacts with K62' and D38 forms a salt bridge with R65' on FabB. E60 interacts with K150' in the docked model a small number of hydrophobic residues, such as L15 and L42, exhibit CSPs[34]. Though FabF and FabB both perform the same fundamental chemical reaction, they appear to have distinct interfaces. FabF has a slightly larger interface (1023 $Å^2$) when comparing docked models with FabB (962 $Å^2$), perhaps consistent with its broader activity and the wider impact on CSPs compared to FabB[53]. This further agrees

with data demonstrating a tighter binding for FabF than the previous calculated FabB Kd of 37.6 ± 6.6μM. Though FabB and FabF share particularly similar structures and activity[37,39], their interactions with AcpP are unique.

**Reductases: FabG and FabI**

The condensation reaction performed by ketosynthases generates 3-oxoacyl-AcpP, which is subsequently reduced to 3R-hydroxyacyl-AcpP by FabG in a NADPH-dependent fashion (Fig. 1a, Fig S1)[54]. For NMR titrations, $NAD^+$ was added along with FabG, as previous studies showed a difference in AcpP binding efficiency in the presence and absence of $NAD^+$ (Fig. S7)[55]. The total perturbed residues span many residues across the AcpP. The perturbed residues begin at the end of helix I, with a few interactions across loop I. Helix II is perturbed to some degree across most of the AcpP, with only small regions seeing CSPs as low as background. Nearly all residues on loop II and through to helix IV are perturbed until the bottom of helix IV. The residues of C8-AcpP exhibiting the largest CSPs (Table S5, Fig. S2) were: N25 and F28 on loop 1; D35, S36, L37, T39, V40, E47, and F50 on helix 2; and T64, Q66, and A68 on helix 3. Interface residues D35 and E47 interact with R19' and R207' of FabG respectively, while L37, T38, and V40 form hydrophobic interactions at the interface. N25, F28, F50, T52, T64, Q66, and A68 were all positioned away from the interacting face in the model. The identified region of interaction is in agreement with the binding region previously identified by mutagenesis and activity assays [56,57]. Binding calculations demonstrated a $K_d$ of 52.3 ± 27.5 μM with a $k_{off}$ of 3559 ± 2061 $s^{-1}$ and a two to one stoichiometry(Table S9, Fig. S13).

The final step of each elongation cycle in (saturated) FAB is catalyzed by the enoylreductase, FabI, which produces a saturated acyl-AcpP through NADH–dependent reduction of enoyl-AcpP (Fig. 1a)[58]. FabI was also titrated with $NAD^+$ present. Upon interaction with FabI, NMR signals from residues throughout C8-AcpP exhibited CSPs (Table S4, Fig. S8). With perturbations beginning on helix I and showing a few sparse interactions through helix I and onto loop I. However, more interactions are seen on helix II, which shows interactions throughout only diminishing perturbation at the bottom of helix II. Finally, the loop II and helix III and IV see interactions fairly consistently until a drop in perturbations at the end of helix IV. The most perturbed residues of AcpP included I10 and L15 on helix 1; F28 in loop 1; D35, S36, L37, V43, M44, A45, and E47 in helix 2; A59 of helix 3; and Q66 and A68 of helix 4 are also highly perturbed. Similar to FabG,

salt bridges likely form at residues D35 and E47, with E47 likely binding K43' on FabI. And D35 interacting with R193'. Finally, the residues L37 and M44 on helix 2 form hydrophobic interactions with residues on the FabI interface, demonstrating a binding motif similar to FabG (Fig. 4). Uniquely, the perturbations and docked model of AcpP-FabI show not only the canonical AcpP helix II and III binding to the enzyme, but also additional interactions with helix 1. The identified binding region corresponds with previous mutational studies that first identified the AcpP–FabI interface [51]. Titan analysis calculated a $K_d$ of 1.7 ± 1.2 μM with a $k_{off}$ of 8500 ± 2700 $s^{-1}$ and approximately one to one stoichiometry(Table S9, Fig. S14).

**The TesA *E. coli* thioesterase**

Many organisms utilize a thioesterase to liberate fatty acids from the ACP. In *E. coli,* mature acyl-AcpPs are instead steered directly into other biosynthetic pathways via acyl transfer from AcpP, primarily for phospholipid biosynthesis. However, *E. coli* does possess the thioesterase TesA, which localizes in the bacterial periplasm[59]. Though TesA is not believed to be involved in the terminal step of *E. coli* FAB, it can hydrolyze acyl-AcpP *in vitro* and has been used as a tool for FAB engineering to increase free fatty acid titer when overexpressed within *E. coli* [60]. Upon titration with TesA, the pattern of C8-AcpP CSPs occurred predominantly in residues different from those perturbed by FAB (Table S6, Fig. S3, S5, S10). The perturbations are relatively minor throughout with small perturbations in helix and loop I. There are a larger number of perturbations on helix II, with more than half of the residues being perturbed over the background. Loop II and helix III show a diminished level of perturbation relative to helix II, this trend continues with few perturbations identified on loop III and helix IV. Overall, the titration by TesA appeared to affect signals from fewer C8-AcpP residues than the other proteins tested. The largest observed residues include loop 1 at S27 and D31; helix 2 at T42, M44, and A45; loop 2 at G52; and loop 3 at T63. Residue D31 appears to interact with R77' of TesA upon binding. Additionally, D35 appears to interact with the TesA backbone or sidechain at S43'. The internal AcpP residues A45 and L42, located within the central hydrophobic core, are both perturbed upon TesA binding. S27 lies in the loop following helix I and near the interface of AcpP and the enzyme, likely experiencing or stabilizing loop motions upon salt bridge formation by D31. M44 appears somewhat distal from the interface near the acyl cargo, although in the case of FabI is part of the interface. T63 appears in the docked model to be positioned to interact with the hydrophobic surface region

of TesA (Fig. 4). Titan analysis calculated a $K_d$ of 12.5 ± 7 μM with a $k_{off}$ of 9716 ± 820 $s^{-1}$ (Table S4, Fig. S15), though these data demonstrate more error due to the small number and small migration of peaks. The small number of interactions demonstrates that the TesA interface is not optimized for AcpP interactions, further suggesting that it could be engineered to provide a classical interface and increase the interactions and turnover.

**Elucidation of dynamic AcpP•PP interface throughout *E. coli* FAS**

Combining these NMR titrations and docked structures provides a powerful data set of functional PPIs in *E. coli* FAB (Fig. 1a, Fig S5). When compared against each other, these CSPs demonstrate two important concepts to shape our understanding of acyl carrier protein dependent synthases. Firstly, AcpP•PP interactions are commonly understood as electrostatic, with the acidic AcpP surface binding to a "positive patch" at the surface of the partner enzyme. However, the majority of the largest CSPs found in these studies correspond to hydrophobic residues (Fig. 3a,b) spanning the interface, acyl pocket, and back of the AcpP. But the data still suggests that electrostatic interface interactions are critical to the protein-protein binding event. Secondly, each enzyme enumerated above binds with AcpP transiently; the weak nature of these interactions is necessary for the "fast" or "fast-intermediate" exchange NMR chemical shifts and agrees with both our presented data and previously known AcpP binding affinities[32,56]. In both fast and fast-intermediate exchange, interactions between AcpP and partner proteins are occurring rapidly enough that residues resolve as a single migrating peak on the spectra, rather than two distinct peaks. The titrations effect on lineshape suggests that the interactions are not so rapid that the titrations are happening in "fast exchange". This is also reflected in the TITAN derived $k_{off}$ rates. These findings demonstrate that recognition between AcpP and its PPs are dynamic processes, driven both by the electrostatic interface and conformational dynamism of the AcpP.

Across the six elongating enzymes tested, half of the residues with perturbations one standard deviation above the mean were at the interface, while the other half of perturbed residues lied in the pocket of AcpP. This is most likely a result of the substrate chain flipping into the PP. Approximately one-third of the largest perturbations, just 10 of 29, are unique to a single partner. More perturbations are shared by three or more of the six enzymes examined than are unique. Each partner, excluding TesA, displays

perturbations at the “top” of the acyl pocket, at the start of helix 2 and the helix 3 to the beginning of helix 4. These interactions are likely those responsible for positioning S36 for substrate delivery. TesA is the only enzyme studied which is known to not be an AcpP FAB partner *in vivo* but has been demonstrated to have a low level of activity *in vitro*. Correspondingly, AcpP does not appear to form the interactions with TesA that are essential for efficient interactions. For other enzymes, it is not unreasonable that AcpP•PP interfaces would predominantly be shared sets of AcpP residues, with a few residues forming unique interactions that contribute to selectivity, given the small size of AcpP and the positively charged binding surfaces of PPs.

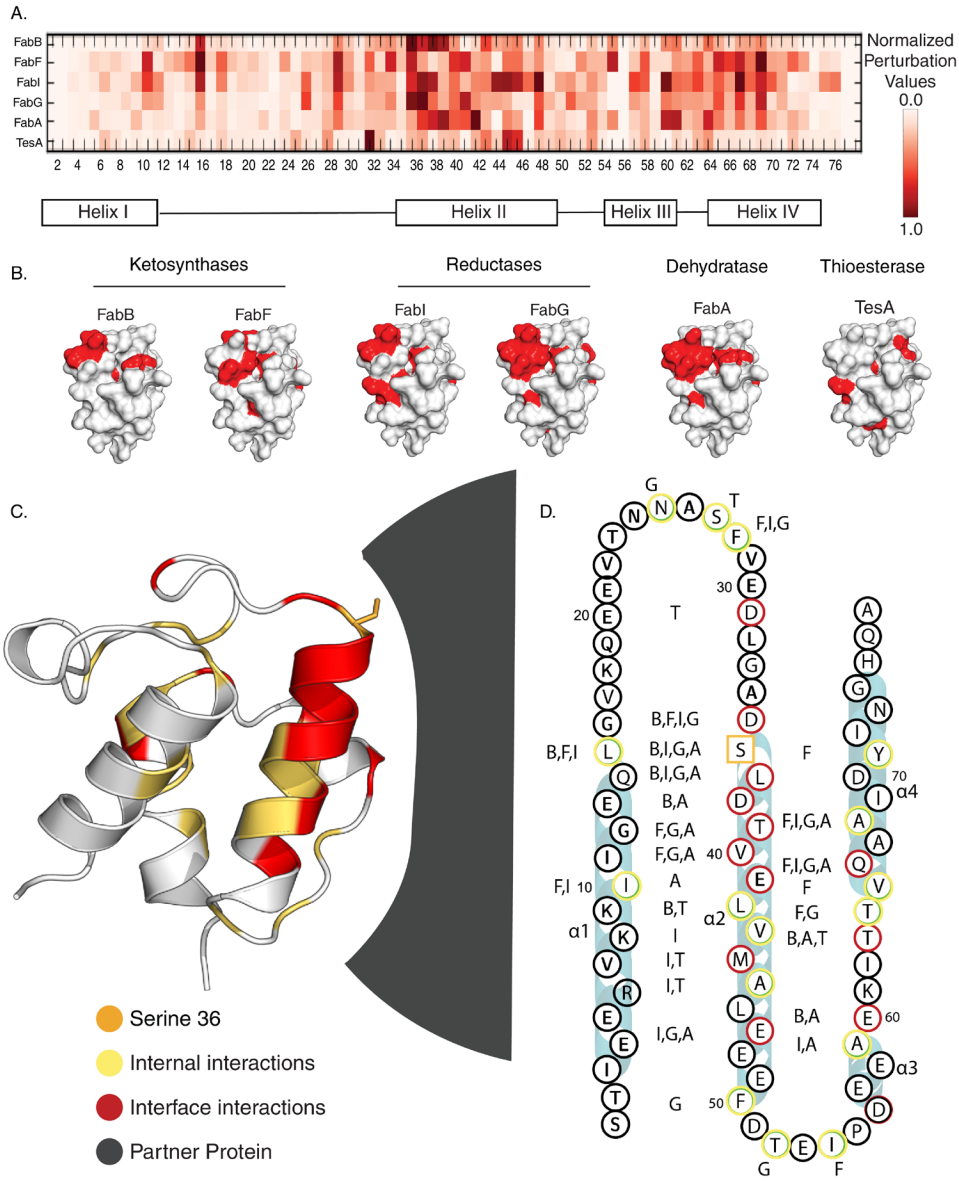



**Figure 1.4. Combined data of AcpP interaction with FAB partner enzymes. A) A heat map of CSPs, each titration is normalized and colored in proportion to the largest CSP in the titration. B) The most perturbed interface of 6 AcpP partners from 4 classes of enzyme, viewed from the interacting face. C) AcpP (PDB: 2FAD) displayed with the most perturbed residues colored by their interaction, the ACP is rotated 90° from the interacting face. A generic partner surface is shown for context. D) The C8 AcpP•partner PPIs with the largest CSPs colored based on their interaction with residues displayed in red (interface) or green (internal). S36 is displayed in orange. The PP responsible is labeled alongside the residue. B (FabB), F (FabF), I (FabI), G (FabG), A (FabA), and T (TesA) are used as shorthand.**

### 1.3 DISCUSSION

Our studies found that recreating the interface of ACP•PP was achievable across four examples, and the models were in excellent agreement with experimental structures. While our study utilized a system

that has a breadth of known information, the active sites of evolutionarily related polyketide synthases are largely conserved, making the inference of a partner protein's active site possible even without experimentally demonstrated residues. Furthermore, we believe the ability of NMR to appreciate the subtle differences in binding residues on the carrier protein is important to understand selectivity and substrate selection for inhibition or engineering. The combinatorial method synergizes the sensitivity and substrate accuracy of NMR experiments and the ability of informed docking to generate accurate models. This can be used to guide future engineering efforts, leveraging *in silico* screening for efficiency and economy. Furthermore, given the ability to recreate these interfaces, inhibitor screening should also be possible. The small size, simple electrostatic surfaces, and breadth of background knowledge of partner structures makes carrier protein dependent pathways ideal systems for computationally guided engineering and inhibition.

Taken together, these CSPs reveal a striking distinction between enzyme classes (Fig. 4a,b). Though AcpP contains only 77 amino acids, the residues involved in each binding are distinct, illuminating how one small protein can interact with dozens of partners. Each PP binding, excluding TesA, induces perturbations at the "top" of the acyl pocket, at the start of helix two and helix three to the beginning of helix four. Although AcpP•PP interactions are typically understood as predominantly electrostatic in nature, half of the largest CSPs correspond to hydrophobic residues (Fig. 4c,d). This represents an evolution in the understanding of type II FAS AcpP•partner recognition, demonstrating that unique residues are used for PPIs with different PPs. We propose a model wherein specific surface interactions are critical for creating allosteric movements within the central channel, triggering the chain flipping event. This may explain the stringent control of reactivity, yet broad range of substrates, necessary for FAB function. The disparate binding motifs found across this iterative pathway provides a compelling model for how a simple <10 kDa protein performs unique interactions for each of six enzymes, while still displaying similarities within classes. This basic model can be extended across the known AcpP interactome, currently at 27 proteins, each of which may demonstrate similarly unique PPIs. Although this constitutes a broad sampling of each enzyme, further study can provide detail into each protein's dynamics and allosteric control. This study provides a foundation with which to expand upon our understanding of PPI driven specificity, PPI redesign, and inhibitor development.

**Chapter 1**, is in full a reprint of the manuscript: Thomas G. Bartholow, Terra Sztain, Ashay Patel, D. John Lee, Megan A. Young, Ruben Abagyan, Michael D. Burkart “Elucidation of transient protein-protein interactions within carrier protein-dependent biosynthesis” Communications Biology 2020. The dissertation author is the primary author of the manuscript.

## 1.4 REFERENCES


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

2. Lai, J. R., Koglin, A. & Walsh, C. T. Carrier Protein Structure and Recognition in Polyketide and Nonribosomal Peptide Biosynthesis. *Biochemistry* **45**, 14869–14879 (2006).

3. Schweizer, E. & Hofmann, J. Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems. *Microbiol. Mol. Biol. Rev.* **68**, 501–517 (2004).

4. Gully, D. & Bouveret, E. A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in Escherichia coli. *PROTEOMICS* **6**, 282–293 (2006).

5. De Lay, N. R. & Cronan, J. E. In vivo functional analyses of the type II acyl carrier proteins of fatty acid biosynthesis. *J. Biol. Chem.* **282**, 20319–20328 (2007).

6. Gully, D., Moinier, D., Loiseau, L. & Bouveret, E. New partners of acyl carrier protein detected in Escherichia coli by tandem affinity purification. *FEBS Lett.* **548**, 90–96 (2003).

7. Roujeinikova, A., Simon, W. J., Gilroy, J., Rice, D. W., Rafferty, J. B., and 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.* **365**, 135–145 (2007).

8. Tan, Z., Black, W., Yoon, J. M., Shanks, J. V. & Jarboe, L. R. Improving Escherichia coli membrane integrity and fatty acid production by expression tuning of FadL and OmpF. *Microb. Cell Factories* **16**, (2017).

9. Shokri, A. & Larsson, G. Characterisation of the Escherichia coli membrane structure and function during fedbatch cultivation. *Microb. Cell Factories* **3**, 9 (2004).

10. Morris, T. W., Reed, K. E. & Cronan, J. E. Lipoic acid metabolism in Escherichia coli: the lplA and lipB genes define redundant pathways for ligation of lipoyl groups to apoprotein. *J. Bacteriol.* **177**, 1–10 (1995).

11. Majerus, P. W., Alberts, A. W. & Vagelos, P. R. Acyl Carrier Protein, IV. The Identification of 4′-phosphopantetheine as the Prosthetic Group of the Acyl Carrier Protein. *Proc. Natl. Acad. Sci. U. S. A.* **53**, 410–417 (1965).

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

13. Cronan, J. E. The chain-flipping mechanism of ACP (acyl carrier protein)-dependent enzymes appears universal. *Biochem. J.* **460**, 157–163 (2014).

14. Kim, Y., Kovriguine, E. & Eletr, Z. NMR Studies of Escherichia Coli Acyl Carrier Protein: Dynamic and Structural Differences of the Apo- and Holo-forms. *Biochem. Biophys. Res. Commun.* (2006).

15. Andrec, M., Hill, R. B. & Prestegard, J. H. Amide exchange rates in Escherichia coli acyl carrier protein: Correlation with protein structure and dynamics. *Protein Sci.* **4**, 983–993 (1995).

16. Colizzi, F., Masetti, M., Recanatini, M. & Cavalli, A. Atomic-Level Characterization of the Chain-Flipping Mechanism in Fatty-Acids Biosynthesis. *J. Phys. Chem. Lett.* **7**, 2899–2904 (2016).

17. Cronan, J. J. & Rock, C. O. Biosynthesis of Membrane Lipids. *EcoSal Plus* **3**, (2008).

18. Liu, T., Vora, H. & Khosla, C. Quantitative analysis and engineering of fatty acid biosynthesis in E. coli. *Metab. Eng.* **12**, 378–386 (2010).

19. 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.* **1**, 687–691 (2006).

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

21. Miyanaga, A., Ouchi, R., Ishikawa, F., Goto, E., Tanabe, G., Kudo, F., and Eguchi, T. Structural Basis of Protein–Protein Interactions between a trans-Acting Acyltransferase and Acyl Carrier Protein in Polyketide Disorazole Biosynthesis. *J. Am. Chem. Soc.* **140**, 7970–7978 (2018).

22. 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. U. S. A.* **113**, 1802–1807 (2016).

23. Weissman, K. J., Hong, H., Popovic, B. & Meersman, F. Evidence for a Protein-Protein Interaction Motif on an Acyl Carrier Protein Domain from a Modular Polyketide Synthase. *Chem. Biol.* **13**, 625–636 (2006).

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

25. Sarria, S., Bartholow, T. G., Verga, A., Burkart, M. D. & Peralta-Yahya, P. Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production. *ACS Synth. Biol.* **7**, 1179–1187 (2018).

26. Sztain, T., Patel, A., Lee, D. J., Davis, T. D., McCammon, J. A., and Burkart, M. D. Modifying the Thioester Linkage Affects the Structure of the Acyl Carrier Protein. *Angew. Chem. Int. Ed.* **0**,.

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

28. Lohman, J. R., Ma, M., Cuff, M. E., Bigelow, L., Bearden, J., Babnigg, G., Joachimiak, A., Phillips, G. N., and Shen, B. The crystal structure of BlmI as a model for nonribosomal peptide synthetase peptidyl carrier proteins. *Proteins* **82**, 1210–1218 (2014).

29. Rajagopal, P., Waygood, E. B., Reizer, J., Saier, M. H. & Klevit, R. E. Demonstration of protein-protein interaction specificity by NMR chemical shift mapping. *Protein Sci. Publ. Protein Soc.* **6**, 2624–2627 (1997).

30. Zuiderweg, E. R. P. Mapping Protein−Protein Interactions in Solution by NMR Spectroscopy. *Biochemistry* **41**, 1–7 (2002).

31. Pellecchia, M., Montgomery, D. L., Stevens, S. Y., Kooi, C. W. V., Feng, H. P., Gierasch, L. M., and Zuiderweg, E. R. P. Structural insights into substrate binding by the molecular chaperone DnaK. *Nat. Struct. Biol.* **7**, 298 (2000).

32. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).

33. Two-Dimensional NMR Lineshape Analysis | bioRxiv. https://www.biorxiv.org/content/10.1101/039735v1.

34. 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., and Tsai, S.-C. Molecular basis for interactions between an acyl carrier protein and a ketosynthase. *Nat. Chem. Biol.* **15**, 669 (2019).

35. 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., and Burkart, M. D.Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* **505**, 427–431 (2014).

36. ) Jaremko, M. J., Lee, D. J., Patel, A., Winslow, V., Opella, S. J., McCammon, J. A., and Burkart, M. D. Manipulating Protein-Protein Interactions in Nonribosomal Peptide Synthetase Type II Peptidyl Carrier Proteins.*Biochemistry*https://pubmed.ncbi.nlm.nih.gov/28920687/(2017) doi:10.1021/acs.biochem.7b00884.

37. Abagyan, R. & Totrov, M. Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins. *J. Mol. Biol.* **235**, 983–1002 (1994).

38. Abagyan, R., Totrov, M. & Kuznetsov, D. ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *J. Comput. Chem.* **15**, 488–506 (1994).

39. 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., and Burkart, M. D. Gating mechanism of elongating β-ketoacyl-ACP synthases. *Nat. Commun.* **11**, 1727 (2020).

40. Platensimycin is a selective FabF inhibitor with potent antibiotic properties | Nature. https://www.nature.com/articles/nature04784.

41. Identification and Analysis of the Acyl Carrier Protein (ACP) Docking Site on β-Ketoacyl-ACP Synthase III. https://www.jbc.org/content/276/11/8231.

42. 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* **4**, 253–264 (1996).

43. Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., Beglov, D., Vajda, S. The ClusPro web server for protein–protein docking. *Nat. Protoc.* **12**, 255–278 (2017).

44. Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., Kozakov, D. New additions to the ClusPro server motivated by CAPRI. *Proteins* **85**, 435–444 (2017).

45. Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., and Vajda, S. How good is automated protein docking? *Proteins* **81**, 2159–2166 (2013).

46. van Zundert, G. C. P., Rodrigues, J. P. G. L. M., Trellet, M., Schmitz, C., Kastritis, P. L., Karaca, E., Melquiond, A. S. J., van Dijk, M., de Vries, S. J., and Bonvin, A. M. J. J. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428**, 720–725 (2016).

47. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information | Journal of the American Chemical Society. https://pubs.acs.org/doi/10.1021/ja026939x.

48. Lyskov, S., Chou, F.-C., Conchúir, S. Ó., Der, B. S., Drew, K., Kuroda, D., Xu, J., Weitzner, B. D., Renfrew, P. D., Sripakdeevong, P., Borgo, B., Havranek, J. J., Kuhlman, B., Kortemme, T., Bonneau, R., Gray, J. J., and Das, R. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLOS ONE* **8**, e63906 (2013).

49. Lyskov, S. & Gray, J. J. The RosettaDock server for local protein–protein docking. *Nucleic Acids Res.* **36**, W233–W238 (2008).

50. Joshi, V. C. & Wakil, S. J. Studies on the mechanism of fatty acid synthesis: XXVI. Purification and properties of malonyl-coenzyme A—Acyl carrier protein transacylase of Escherichia coli. *Arch. Biochem. Biophys.* **143**, 493–505 (1971).

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

52. 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.* **255**, 11949–11956 (1980).

53. Yu, X., Liu, T., Zhu, F. & Khosla, C. In vitro reconstitution and steady-state analysis of the fatty acid synthase from Escherichia coli. *Proc. Natl. Acad. Sci.* **108**, 18643–18648 (2011).

54. Toomey, R. E. & Wakil, S. J. Studies on the mechanism of fatty acid synthesis XV. Preparation and general properties of β-ketoacyl acyl carrier protein reductase from Escherichia coli. *Biochim. Biophys. Acta BBA - Lipids Lipid Metab.* **116**, 189–197 (1966).

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

56. Zhang, Y.-M., Wu, B., Zheng, J. & Rock, C. O. Key Residues Responsible for Acyl Carrier Protein and β-Ketoacyl-Acyl Carrier Protein Reductase (FabG) Interaction. *J. Biol. Chem.* **278**, 52935–52943 (2003).

57. Ac, P., Ym, Z., Co, R. & Sw, W. Cofactor-induced Conformational Rearrangements Establish a Catalytically Competent Active Site and a Proton Relay Conduit in FabG. *Structure (London, England : 1993)* https://pubmed.ncbi.nlm.nih.gov/15016358/ (2004) doi:10.1016/j.str.2004.02.008.

58. Heath, R. J. & Rock, C. O. Enoyl-Acyl Carrier Protein Reductase (fabI) Plays a Determinant Role in Completing Cycles of Fatty Acid Elongation in Escherichia coli. *J. Biol. Chem.* **270**, 26538–26542 (1995).

59. Cho, H. & Cronan, J. E. 'Protease I' of Escherichia coli functions as a thioesterase in vivo. *J. Bacteriol.* **176**, 1793–1795 (1994).

60. Shin, K. S., Kim, S. & Lee, S. K. Improvement of free fatty acid production using a mutant acyl-CoA thioesterase I with high specific activity in Escherichia coli. *Biotechnol. Biofuels* **9**, 208 (2016).

61. NMRPipe: A multidimensional spectral processing system based on UNIX pipes | SpringerLink. https://link.springer.com/article/10.1007/BF00197809.

62. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinforma. Oxf. Engl.* **31**, 1325–1327 (2015).

63. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

## 1.5 Methods

### Materials

The $^{15}N$ ammonium chloride used in the labeled growth was purchased from Cambridge Isotopes laboratory. Deuterium oxide ($D_2O$) used in preparation of perdeuterated growth was purchased from Sigma Aldritch. All unlabeled proteins were grown on Luria broth from Teknova.

### General partner protein purification protocol

All partner proteins were generated through overexpression in *E. coli* BL21 (DE3) Cells and grown in LB media. Cells were grown in the presence of 50 mg/L kanamycin before induction with 1 mM IPTG at OD600 = 0.8 and incubated at 18°C for 12-18 hours. Cells were pelleted in a JLA-8.1 rotor at 800 RCF. Cells were re-suspended and lysed by sonication in 50 mM Tris-HCl (pH 7.4), 250 mM NaCl, and 10% glycerol. The lysate was then spun at 10,000 RCF in a JA-20 rotor for 1 hour to pellet the membrane and insoluble materials. Proteins were purified using Ni-IMAC (Bio-Rad) after a 30 minute batch binding time rotating at 4°C. The general protocol used 2 solutions, a wash of 40mL of lysis buffer followed by a wash of 40mL lysis buffer with 15mM imidazole added. This was followed by 3 5mL elutions with lysis buffer containing 250mM imidazole. Unless stated in the purification specifics below this was the method used in all purifications. After purification proteins were concentrated to ~2mL and purified using size exclusion chromatography on a Superdex 75 column. The partner protein was purified into NMR buffer and concentrated to the concentrations listed below before addition to the tube for NMR.

AcpP was grown on a His-tagged pET-22b vector in *E. coli* BL21 (DE3) cells. Cells were grown at 37°C in M9 minimal media containing 1g of $^{15}N$ $NH_4Cl$ and 8g of glucose. Perdeuteration was achieved by growing *E. coli* in increasing ratios of $D_20$. Starting by growing 5mL overnight in 25% $D_2O$/ 75% $H_2O$, this was used to inoculate 5mL cultures which were 50% $D_2O$/ 50% $H_2O$. This same technique was used to inoculate 75% $D_2O$, 90% $D_2O$, and finally 100% $D_2O$ starters. This starter was used to inoculate the liter of deuterated media. Once the growth reached an OD600 of 0.8 they were induced with 1mM IPTG and allowed to grow for an additional 4 hours at 37°C.

Following purification AcpP was dialized overnight into 50mM Tris, 250mM NaCl, and 1mM DTT buffer, in order to remove the imidazole before subsequent reactions. AcpP was first prepared as uniformly apo by reaction with *Pseudomonas aeruginosa* ACPH in a solution with 5mM $MgCl_2$, and 0.5mM $MnCl_2$, and 1mM DTT. This reaction was performed overnight at 37°C. Apofication was confirmed by conformationally sensitive UREA-PAGE. Following this, loading was performed using 3 *E. coli* biosynthetic enzymes CoaA, CoaD, and CoaE and the *Bacillus subtilis* SFP. The reaction is performed with 12.5mM $MgCl_2$, 10mM ATP, 0.1μM CoaA, 0.1μM CoaD, 0.1μM CoaE, 0.2μM Sfp, 0.02% Triton X, 0.01% Azide, and 0.1% TCEP. The reaction was performed overnight at 37°C, with loading confirmed by conformationally sensitive UREA-PAGE. Stable C8 acyl loaded ACP analogs were achieved through loading of an octanoyl pantethenamide probe.

**Purification and sample preparation of FabF**

ACP was concentrated to 3.87 mg/mL using Amicon Ultra-15 3kDa centrifugal filters, a Nanodrop was used to measure concentrations with the extinction coefficient 1490 $M^{-1}cm^{-1}$. FabF was concentrated to 10.1 mg/mL using Amicon Ultra-15 10kDa spin filters, using the extinction coefficient 25900 $M^{-1}cm^{-1}$. These were used to create a 0.042μM AcpP zero point sample and a 0.042μM AcpP 0.0837μM FabF saturated sample. FabF was purified for titration the day before the experiment, ensuring a “fresh” sample for maximum stability. Perdeuterated AcpP was used to boost signal from sensitivity lost due to the titrated partner

protein. This also lowers the concentration of partner protein necessary to accommodate stability concerns. The proteins were purified into a 10mM potassium phosphate pH 7.4 buffer with 0.5mM TCEP and 0.1% $NaN_3$ and any buffer added to the tubes was taken from the same FPLC buffer solution. The total volume of both the zero point and saturated samples was 450µL, 50µL of $D_2O$ was added to both tubes for locking.

**Purification and sample preparation of FabI**

ACP was concentrated to 2.38 mg/mL using Amicon Ultra-15 3kDa centrifugal filters, a Nanodrop was used to measure concentrations with the extinction coefficient 1490 $M^{-1}cm^{-1}$. FabI was concentrated to 22.1 mg/mL using Amicon Ultra-15 10kDa spin filters, using the extinction coefficient 15930 $M^{-1}cm^{-1}$. These were used to create a 0.0669µM AcpP zero point sample and a 0.0669µM AcpP 0.2689µM FabI saturated sample. This high equivalent concentration was used to ensure that there would be at least a 1:1 ratio of ACP: FabI tetramer. The FabI was purified the day before the experiment and concentrated to the high molarity necessary the morning of the experiment in order to ensure the sample was as stable as possible. Perdeuterated AcpP was used to boost signal form quenching. The proteins were purified into a 10mM potassium phosphate pH 7.4 buffer with 0.5mM TCEP, 0.5mM NAD+ and 0.1% $NaN_3$ and any buffer added to the tubes was taken from the same FPLC buffer solution. The total volume of both the zero point and saturated samples was 450µL, 50µL of $D_2O$ was added to both tubes for locking.

**Purification and sample preparation of FabG**

ACP was concentrated to 2.25 mg/mL using Amicon Ultra-15 3kDa centrifugal filters, a Nanodrop was used to measure concentrations with the extinction coefficient 1490 $M^{-1}cm^{-1}$. FabG was concentrated to 11.21 mg/mL using Amicon Ultra-15 10kDa spin filters, using the extinction coefficient 11460 $M^{-1}cm^{-1}$. These were used to create a 0.0538µM AcpP zero point sample and a 0.0538µM AcpP 0.2199µM FabG saturated sample. This high equivalent concentration was used to ensure that there would be at least a 1:1 ratio of

ACP: FabG tetramer. The FabG protein was purified the day before titration and concentrated the morning of the experiment to ensure a stable sample for the experiment. Perdeuterated AcpP was used to boost signal form quenching. The proteins were purified into a 10mM potassium phosphate pH 7.4 buffer with 0.5mM TCEP, 0.5mM NAD+ and 0.1% $NaN_3$ and any buffer added to the tubes was taken from the same FPLC buffer solution. The total volume of both the zero point and saturated samples was 450µL, 50µL of $D_2O$ was added to both tubes for locking.

**Purification and sample preparation of TesA**

ACP was concentrated to 6.5 mg/mL using Amicon Ultra-15 3kDa centrifugal filters, a Nanodrop was used to measure concentrations with the extinction coefficient 1490 $M^{-1}cm^{-1}$. TesA was concentrated to 3.2 mg/mL using Amicon Ultra-15 10kDa spin filters, using the extinction coefficient 40450 $M^{-1}cm^{-1}$. These were used to create a 0.0538µM AcpP zero point sample and a 0.0699µM AcpP 0.104µM TesA saturated sample. TesA was prepared the day before the experiment in order to have a fresh and stable sample for the experiment. Due to the small size of TesA, perduteration was likely not necessary but was maintained for consistency. The proteins were purified into a 10mM potassium phosphate pH 7.4 buffer with 0.5mM TCEP, and 0.1% $NaN_3$ and any buffer added to the tubes was taken from the same FPLC buffer solution. The total volume of both the zero point and saturated samples was 450µL, 50µL of $D_2O$ was added to both tubes for locking.

**NMR Methods**

Experiments were performed on a Bruker 800MHz spectrometer equipped with a cryo-probe at the UCSD Biomolecular NMR facility. Previous assignments[35] of the C8-AcpP were used to assign the backbone peaks on the HSQC for all experiments. Each experiment was performed at 37°C, with each titration

including at least 5 titration points in order to observe the full movement of peaks. CSPs were quantified using the formula

$$CSP = \sqrt{\frac{1}{2}[\delta_H^2 + (\alpha \cdot \delta_N^2)}$$ [32]

An α value of 0.2 was used in all CSP calculations. Titrations were performed by the preparation of an initial saturated and zero-point sample. To create intermediary samples for the titration these two samples were mixed. AcpP is an extremely stable protein, with no denaturation observed. However, in some cases a degree of partner protein crash was seen in the later titrations. To mitigate this when not in the spectrometer samples were kept refrigerated at 4°C and the spectra were collected one after another over the course of ~12 hours. Through this cautious approach and fresh partner protein preparation we were able to collect all saturated samples with no observed crashed partner, though in the case of FabF and FabG a small amount of crashed protein was observed for the final "middle" spectra's sample (3.0 equivalents in FabG and 1.0 equivalents in FabF). All HSQC were acquired with a 1.5 second recycle delay and 2048 data points in the spectra. Spectra were processed using in NMR Pipe[61] and NMRFAM-SPARKY[62]. Spectra were visualized and all NMR spectra figures were generated in SPARKY. CSP calculations and the CSP heatmap were generated in the Matplotlib python utility[63]. NMR titration data was further analyzed using the TITAN 2D lineshape analysis program. For the analysis in all cases a flexible stoichiometry model was used. Both to allow for examining the stoichiometry of the interactions and in order to allow the concentration of the AcpP to vary in the calculation. This was done because AcpP has a very low extinction coefficient, making quantifying exact AcpP concentrations difficult. In the analysis 5 titrations steps were used in each analysis. The calculations were performed by selecting each region of interest on the spectra. Following this the initial fitting was performed, with the parameters first estimated at 10μM with a $k_{off}$ rate of 5000$s^{-1}$. After the fitting each peak chosen was hand checked, in order to ensure that the peak had been properly fit. Though the TITAN program self fits peaks, in more crowded regions of the spectra the program can improperly fit the wrong peak. After the initial fitting jackknife error analysis was performed, this was done in order to verify the program ran without issue. Improperly fit peaks were identified in the more rapid bootstrap error analysis, peaks which had very small migrations or which migrated into other peaks displayed high error and were hand chosen to not be fitted. In this way Jackknife analysis was used to

identify user error and occasional problematic peaks before the much longer bootstrap analysis was run. After these problematic peaks were hand checked a final bootstrap analysis was performed, in all cases 300 steps were performed in the analysis. The results of this bootstrap analysis were used as the error in reporting values, a set of simulated and “real” peaks are presented for each analysis below. A selection of 2 contour plots and 2 3D contour plots for each analysis is supplied. It was observed that the FabG, FabF, and TesA resulted in greater error in the analysis. We suspect the error in FabG, FabF, and TesA is due to the instability of the partner protein seen in the final titration step. FabG in particular had all of the components at extremely high concentrations in order to ensure saturation.

**Docking Method**

In the case of FabF, FabB, FabI, FabG, FabA, and TesA structures for the partner protein were acquired from the protein data bank: 2GFW, 1G5X, 4CV3, 1Q7B, 1MKB, and 1IVN were used respectively. For FabZ the only structure available is 6N3P, a crosslinked crystal structure. As such this was used but AcpP was deleted. 2FAD was used as a starting crystal structure, with an additional carbon added in ICM to elongate the 7 carbon acyl chain and create a C8 acyl chain. All partner proteins were used in the subunit structure which it adopts in solution, especially given that in many cases AcpP binds multiple subunits of a dimer or tetramer. Specifically: FabI and FabG were docked as tetramers. FabA, FabZ, FabF, and FabB were docked as dimers. TesA was docked as a monomer.

**Preparation of PDB proteins for docking simulations**

Before docking the proteins were prepared by solvation and minimization. The AcpP and all partner proteins were solvated with the ICM quickflood procedure in order to generate a water box. After this the proteins were minimized in ICM to optimize side chain orientations and hydrogen bonding with the water box and any ligands. This was performed by running the optimizeHbonds and optimize HisProAsnGlnCys protocols

in order to form more solution relevant conformations of the residues. After this the cofactors, including the AcpP substrate, were deleted and the proteins were docked. However, the structural effects of the AcpP substrate remain reflected in the docking, with the acyl pocket remaining during the simulation.

**ICM docking – "Informed" and "uninformed" procedures**

Docking was performed both with and without focus residues using the ICM fast Fourier transform protein docking protocol. The "informed" docking procedure was performed by specifying experimentally known interacting residues on the AcpP and partner protein. Explicitly, the "uninformed" docking jobs are performed by docking the exact same input structures, without focus residues. In more detail, given the history of mutational study in *E. coli* FAB we leveraged studies which mutated the AcpP interface and saw diminished activity. As an example in the docking of FabG the work of Price et. al, 2004 was used. In their paper they identified the region of NADP binding and proposed an interface. R15 was identified at the edge of the NADP pocket. The R129 and R172 were not chosen because they perform a more complicated docking with the AcpP from the adjacent chain. However, the final model generated did dock such that these two identified residues were interacting with AcpP. This same methodology is very broadly applicable in carrier protein mediated biosynthesis, given the depth of the literature. For the calculations interface residues on the AcpP were chosen by selecting the largest perturbations whose position was such that they were likely hydrogen bonding at the interface. Docking poses were sorted by lowest energy, it was examined if scoring based on Van der Waals or electrostatic interactions specifically would yield more accurate structures. But we noted that the ICM energy scoring function performed best.. Table S7 displays the RMSD of the top 10 poses generated by the informed and uninformed docking jobs.

**Docking with ClusPro, HADDOCK, and Rosie**

It should be noted that the methodology used is not unique to ICM, as such we have performed 3 docking experiments with commonly available online docking servers. Docking was performed with 3 common servers: ClusPro[42-44], HADDOCK[45,46], and Rosie[47,48] with the partner protein FabF. The results are reported in table S8, comparing the top 3 structures generated from the servers. Models created by ICM, ClusPro, HADDOCK, and Rosie were compared to the crosslinked crystal structures through two methods. The first was a full alignment in Pymol, this yielded good values but often global changes in the partner protein and AcpP either upon crosslinking or due to differences in substrate or crystallization conditions appeared to have altered the backbone distant from the interface. To specifically look at the interfaces between the two enzymes the atoms within 5 Å of the interface were selected for both the docked and crystal structures. The two interfaces were then superimposed upon one another using the Pymol super command and the value was reported without deletion. Pymol was used in all structural visualizations. For the comparison study using online servers the input files used in all docking, post solvation and minimization were collected. For HADDOCK docking both the ICM prepared files and the raw PDB files (2FAD and 2GFW) the inputs were loaded into the server and active residues were defined as 65 and 616 on FabF and 35 and 39 on AcpP, the same focus residues in the ICM study. Passive residues were defined within 6.5Å of the active residues as was suggested by the program. As the study is meant to sample the most accessible components of the method the “EASY” access level account was used. Meaning all parameter settings were set to default. The results were ranked using the standard energy scoring metric in HADDOCK and benchmarked by alignment to the crosslinked structure. The Cluspro docking was the simplest method performed for this study the same input structures were used as the HADDOCK. After loading in the structures the attraction residues were set to be the same as in the HADDOCK and ICM jobs. Both the electrostatic and balanced scoring functions are presented. The Rosie server was run with the Docking2 refinement utility, as this utility requires an input structure the Cluspro best ranked “Electrostatic scoring” docked file was used. We felt this represented a second refinement step which could easily be taken by other groups after the rapid Cluspro docking. However, we would advise caution when performing refinements of ClusPro docking. Great care should be taken to ensure that the starting structures are an accurate starting structure and that Rosie is not optimizing an incorrect interface. We feel this demonstrates the ability of many utilities to recreate AcpP partner interfaces. As well it demonstrates the importance of

caution when evaluating docked structures from Cluspro. HADDOCK performed extremely well when supplied with properly prepared starting structures, but we feel this stands as additional evidence that preparation of the proteins to recreate a solvated structure is important. As like ICM HADDOCK performed poorly when provided with structures straight from the PDB.

**Statistics and Reproducibility**

NMR analysis was performed on single sets of $^{1}H$-$^{15}N$ HSQC experiments. Chemical shift perturbations were calculated using the supplied equation and α value. TITAN analysis was performed according to the method outlined by the developers. Calculation of error was performed using only the recommended bootstrapping method, with 300 steps of calculation. RMSDs were calculated using the Pymol align command, in all cases of alignment the 10 lowest energy states are reported. Starting structures were taken from publicly available repositories.

## 1.6 Supplementary Figures

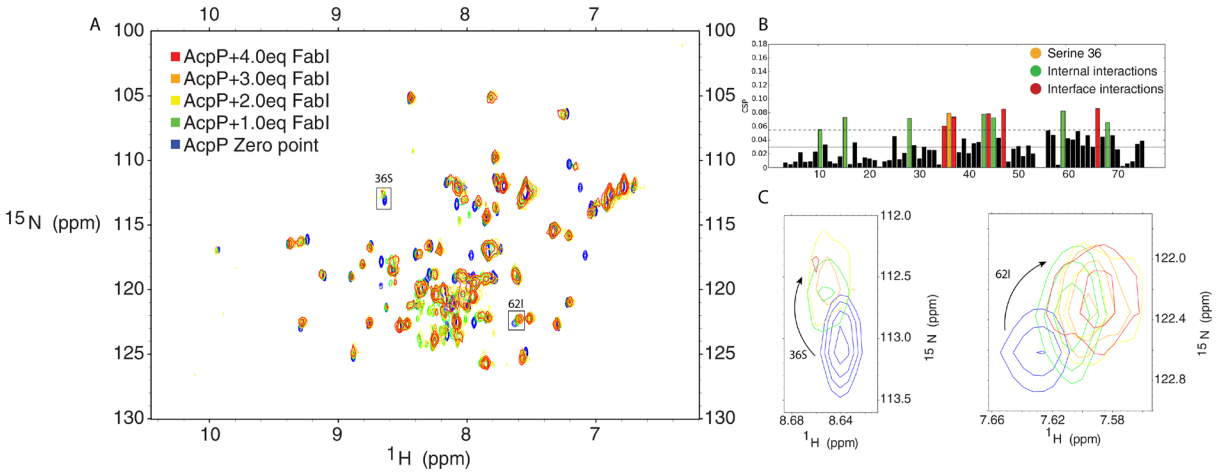



**Figure S1. $^{1}$H-$^{15}$N HSQC titration of C8-AcpP with FabI** A) 5 overlayed $^{1}$H-$^{15}$N HSQC spectra of $^{15}$N-C8-AcpP titrated with increasing concentrations of unlabeled FabI. B) Bar chart of each AcpP residue's CSP with 4.0 molar equivalents of FabI. The mean is shown as a solid line and one standard deviation is the dashed line above. CSPs greater than this value are shown in colors as described by the insert. C) Two examples of residues displaying "titration curves," zoomed in from the boxed residues on the full spectra.

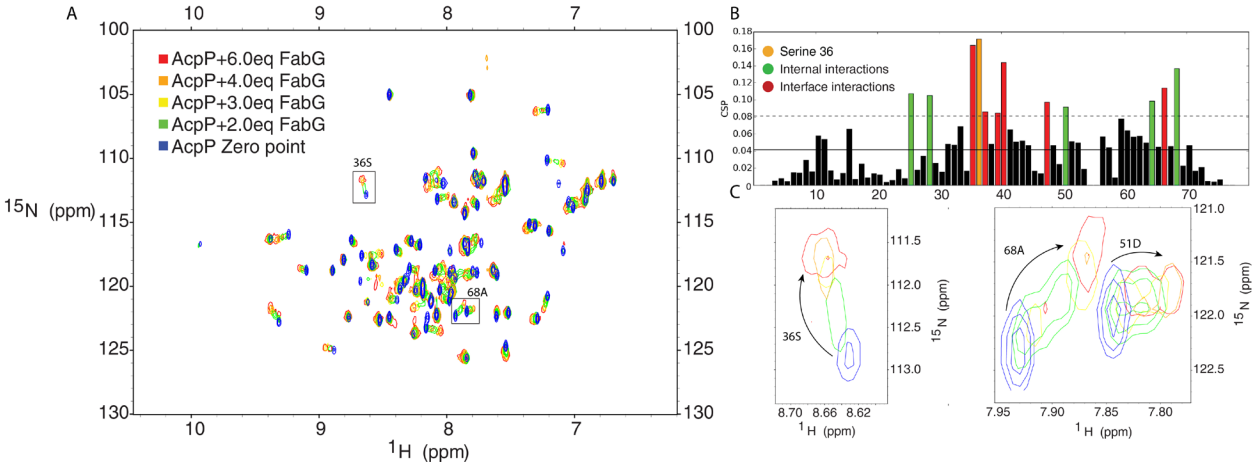



**Figure S2 $^{1}$H-$^{15}$N HSQC titration of C8-AcpP with FabG** A) 5 overlayed $^{1}$H-$^{15}$N HSQC spectra of $^{15}$N-C8-AcpP titrated with increasing concentrations of unlabeled FabG. B) Bar chart of each AcpP residue's CSP with 6.0 molar equivalents of FabG, The mean is shown as a solid line and one standard deviation is the dashed line above. CSPs greater than this value are shown in colors as described by the insert. C) Two examples of residues displaying "titration curves," zoomed in from the boxed residues on the full spectra.

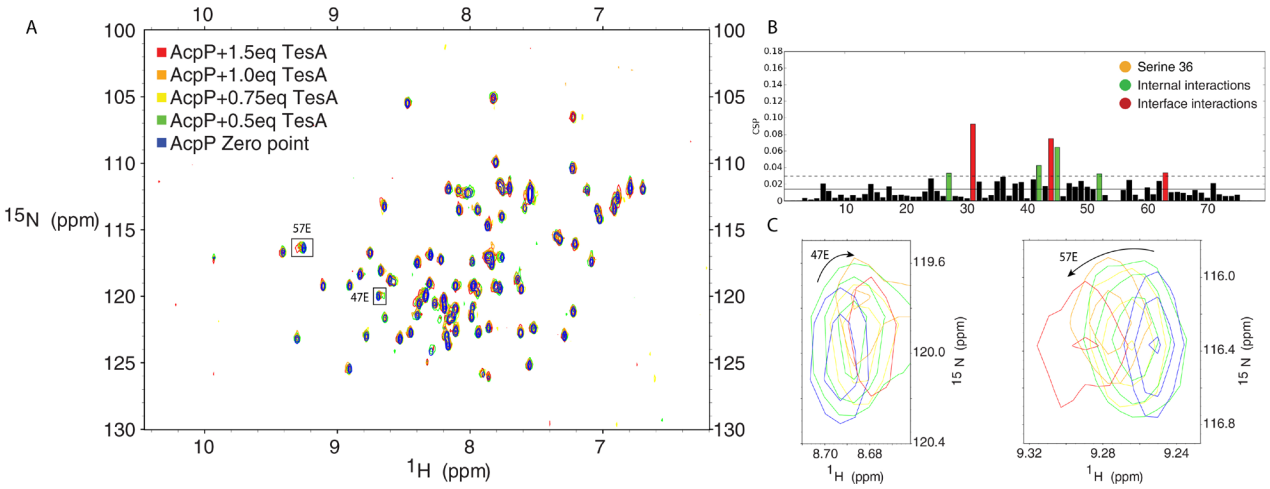



**Figure S3 $^{1}$H-$^{15}$N HSQC titration of C8-AcpP with TesA** A) 5 overlayed $^{1}$H-$^{15}$N HSQC spectra of $^{15}$N-C8-AcpP titrated with increasing concentrations of unlabeled TesA. B) Bar chart of each AcpP residue's CSP with 1.5 molar equivalents of TesA, The mean is shown as a solid line and one standard deviation is the dashed line above. CSPs greater than this value are shown in colors as described by the insert. C) Two examples of residues displaying "titration curves," zoomed in from the boxed residues on the full spectra.

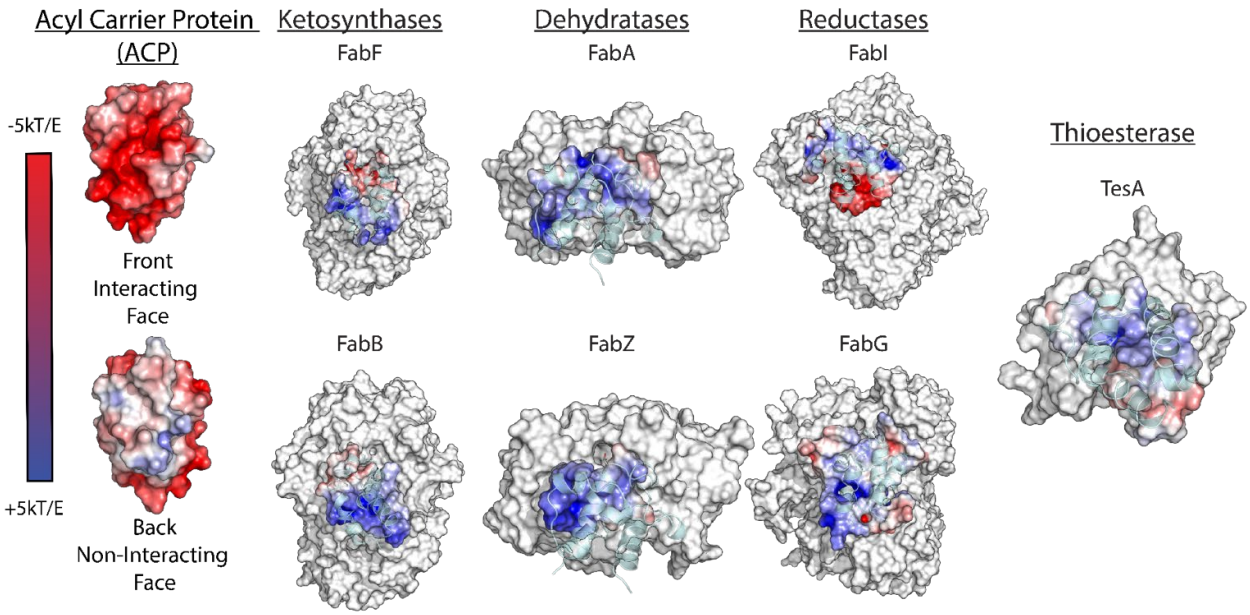



**Figure S4 Electrostatic binding surface of each partner enzyme**. The surfaces of each of the 7 partner proteins from 4 families of enzymes and the AcpP's interacting and non-interacting face. APBS electrostatics are shown for residues within 5 Å of AcpP in the docked model. The entire AcpP is displayed with the calculated electrostatics.

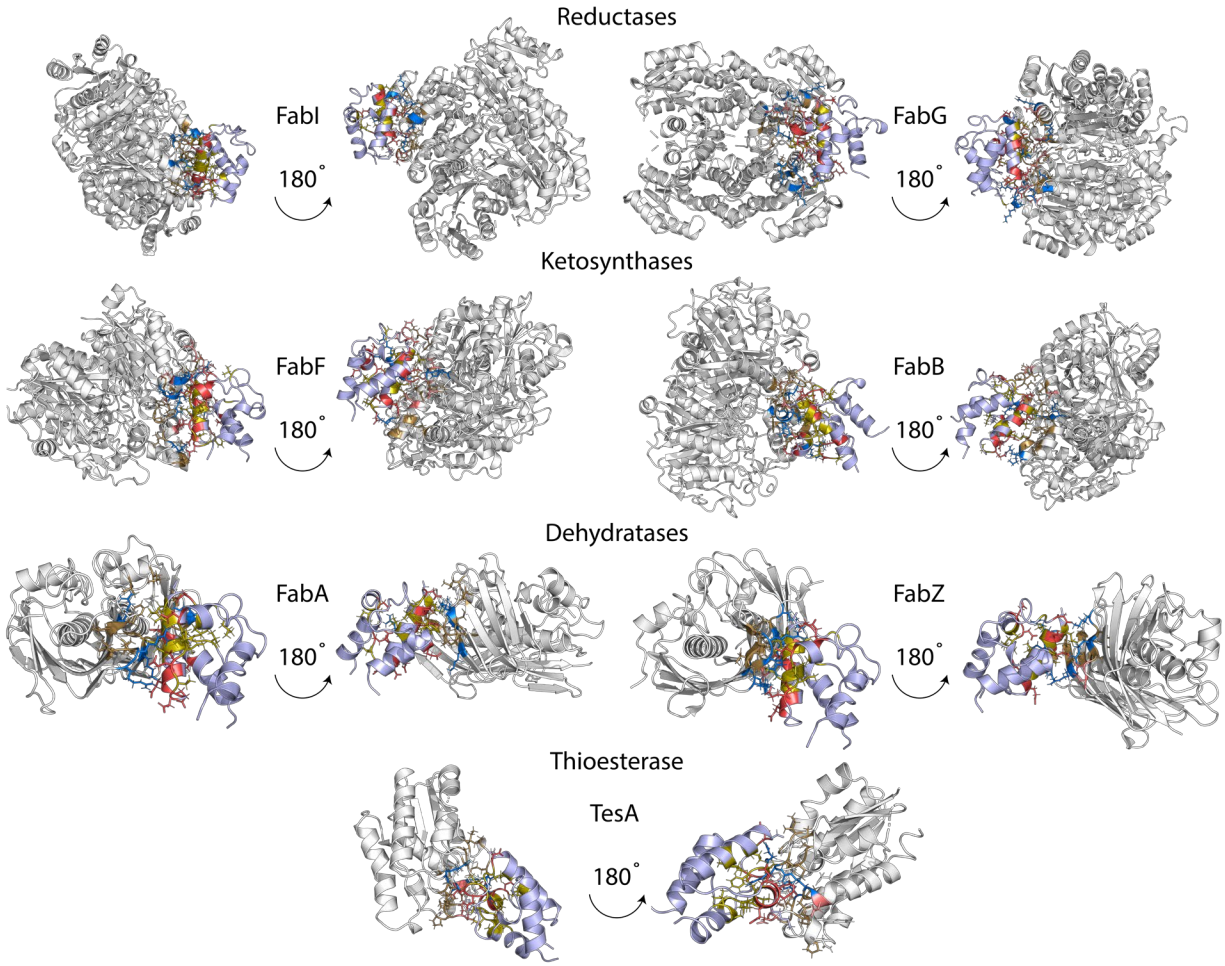



**Figure S5 Interfaces of each partner studied**. Two orientations of the interface of each enzyme studied in this work. Partner proteins are showed in gray with the AcpP in light purple. Negatively charged residues at the interface are colored red with positively charged residues showed in blue. Hydrophobic residues on the AcpP which lie at the interface are colored yellow while hydrophobic residues on the partner protein are tan.

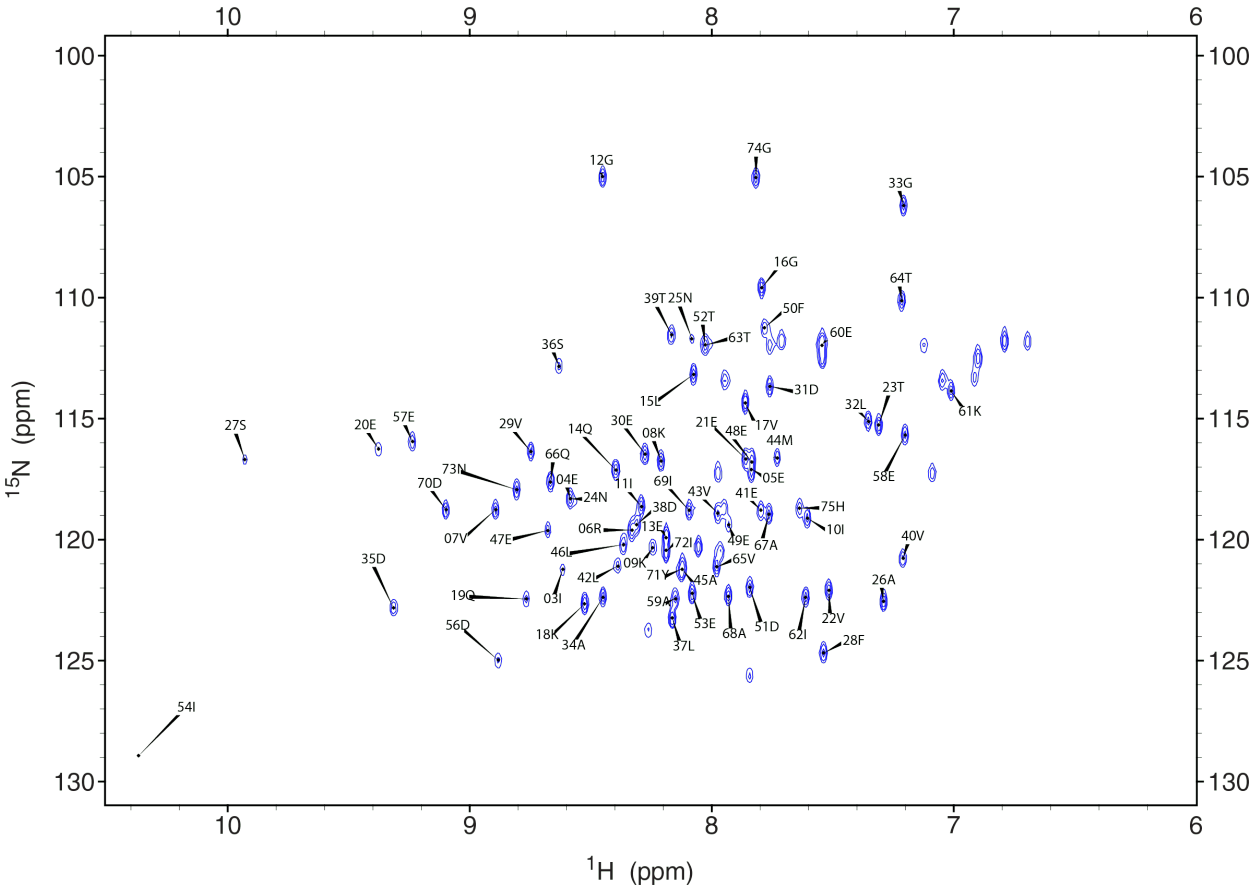



**Figure S6 Residue assignments of C8-AcpP**. Assignments were taken from a previous study performed in the Burkart lab[31] and used to assign the four titrations performed.

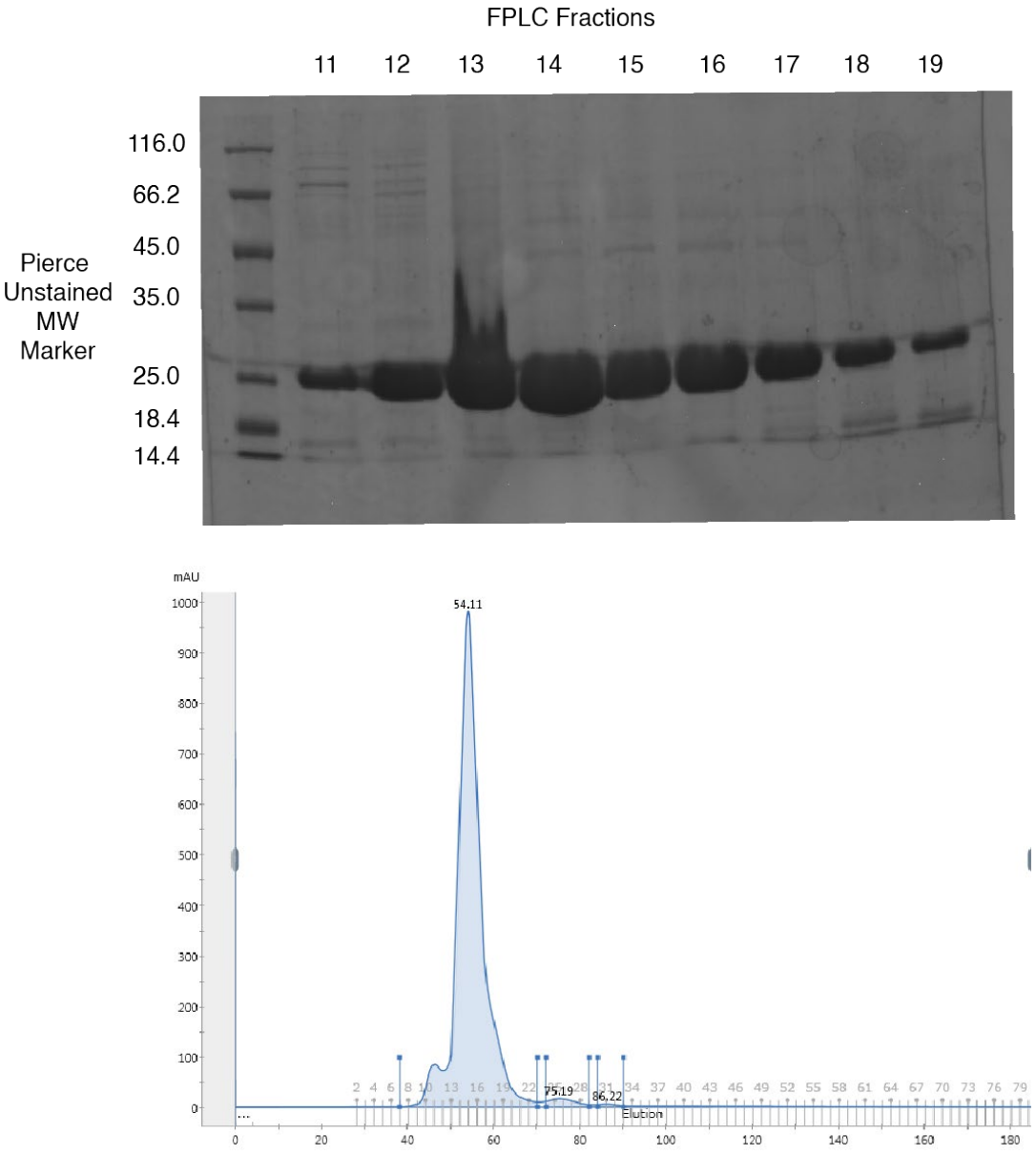



**Figure S7 Purification of FabG**. FabG was purified as described in the methods. A 12% acrylamide denaturing SDS gel was run on the primary peak identified by the FPLC to verify the identity of the FabG.

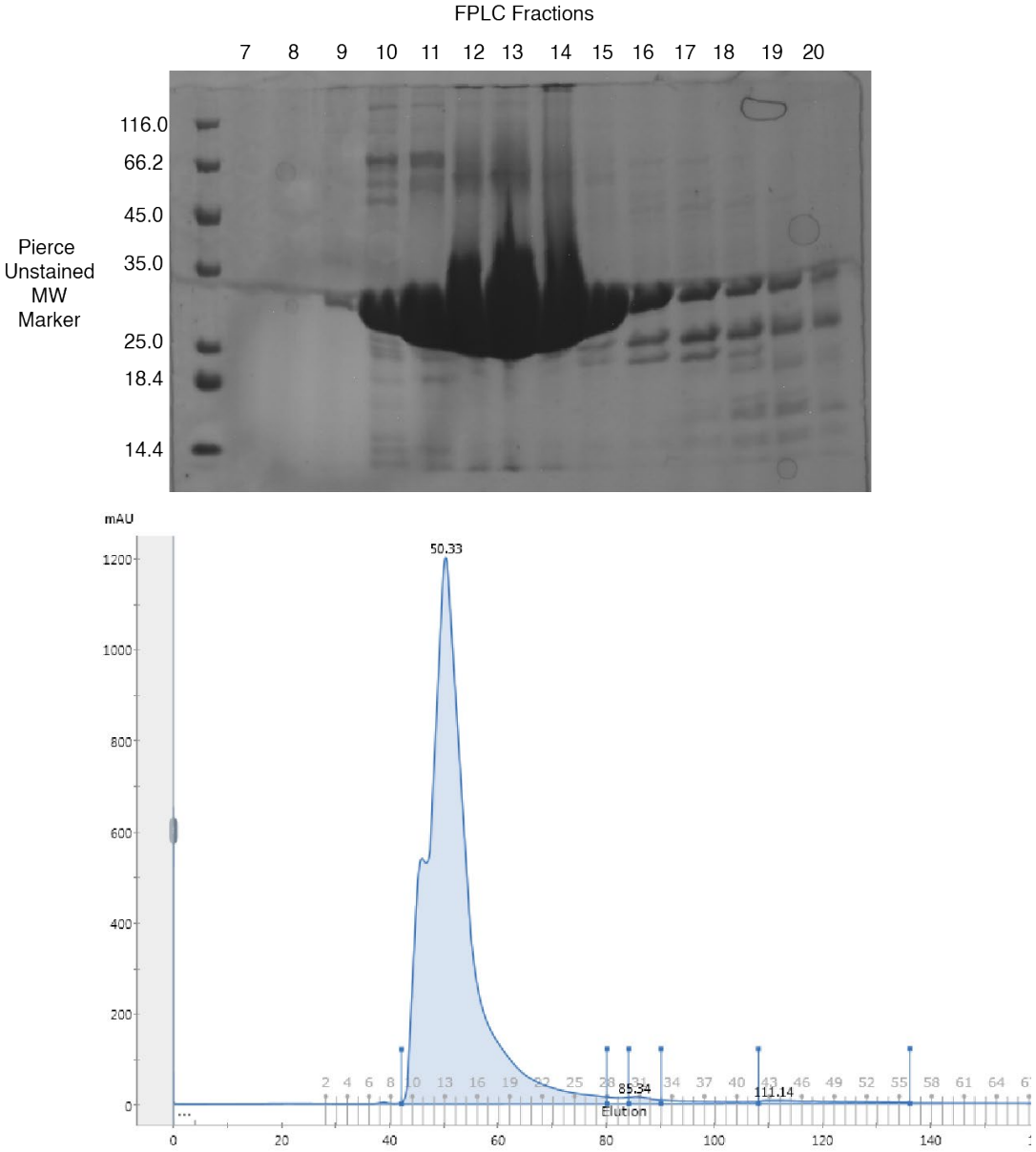



**Figure S8 Purification of FabI.** FabI was purified by the standard method described in the methods. After FPLCing samples were chosen to run on an SDS gel to verify the purity of the protein. A 12% acrylamide denaturing SDS gel was run, showing very large quantities of protein purified for the experiment. Fractions 12, 13, and 14 were taken as the center of the peak and due to the high concentration of protein already present.

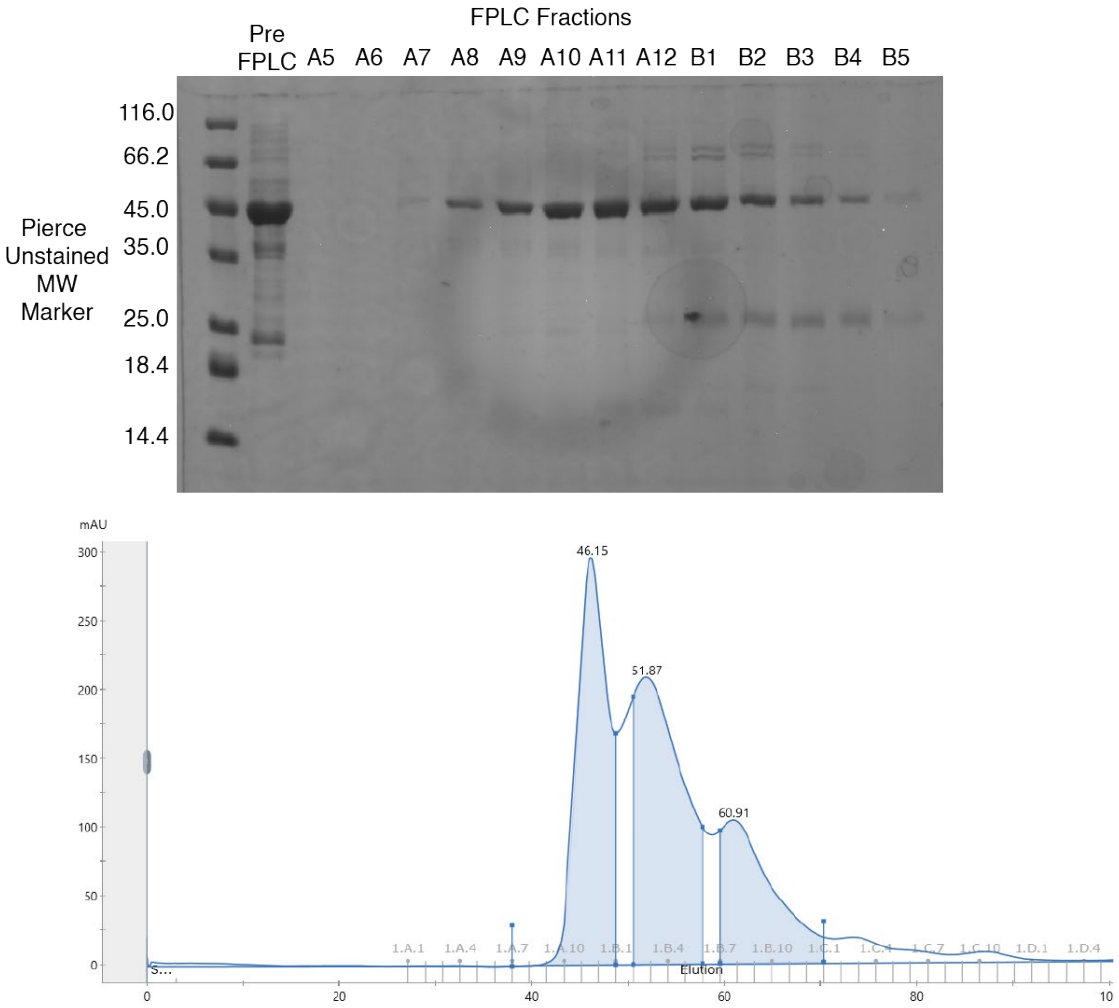



**Figure S9 Purification of FabF**. FabF was purified as described in the methods. The first peak on the FPLC was identified as the likely peak of interest. A sample of nickel purified FabF before the FPLC purification was run on the gel as a standard. A 12% polyacrylamide denaturing SDS gel was run in order to identify the useable fractions. Fractions A8-A12 were identified as sufficiently clean to concentrate for the experiment.

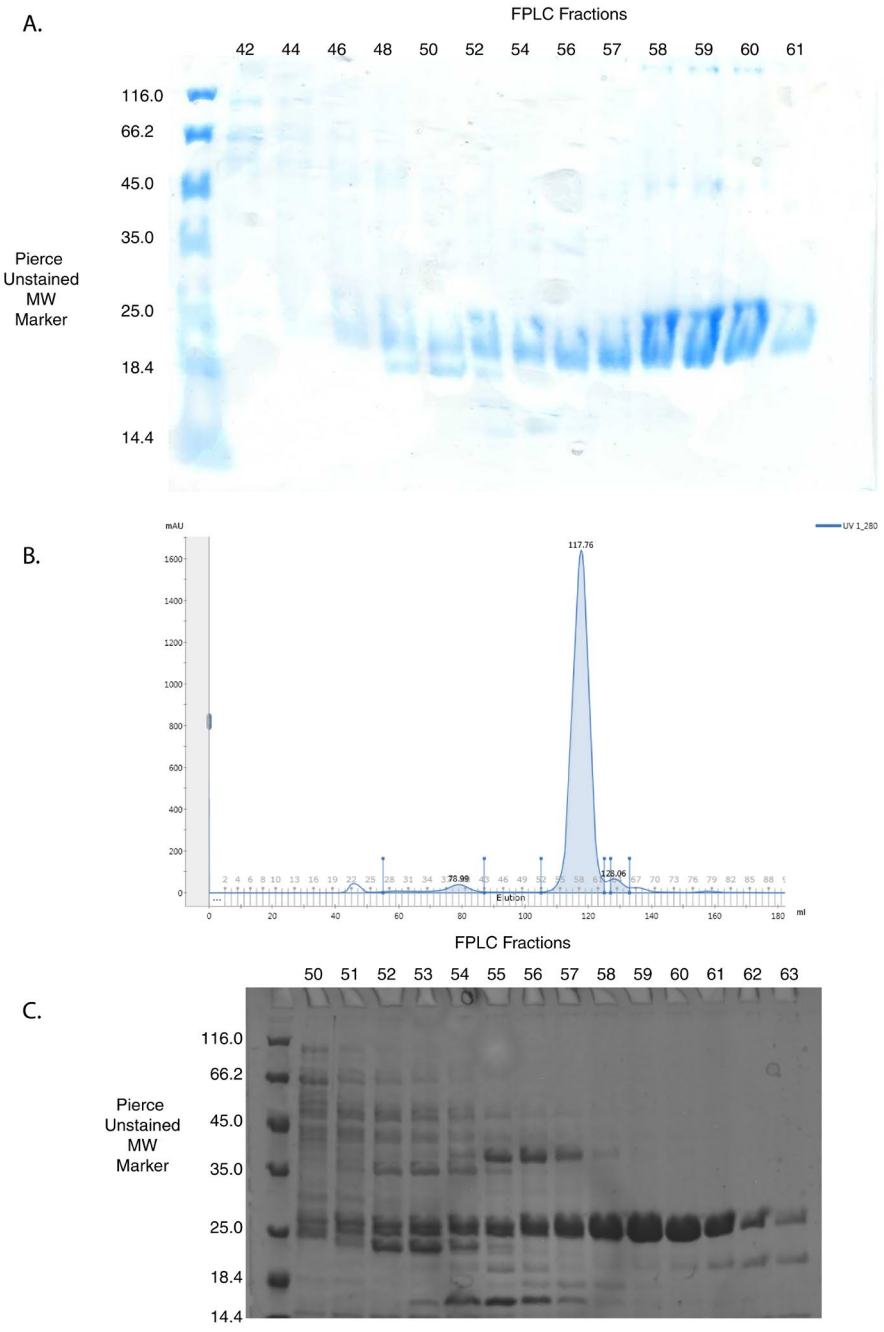



**Figure S10 Purification of TesA**. A/B) TesA was purified as described in the methods. The primary peak was sampled in order to ensure that the sample used was pure. A 12% polyacrylamide denaturing SDS gel was run focusing on the fractions making up the peak. Fractions 57, 58, 59, and 60 were collected and concentrated for the titration experiment. C) A test purification prior to the NMR experiment. This PAGE gel displayed an unsmeared example of the results expected for a TesA purification.

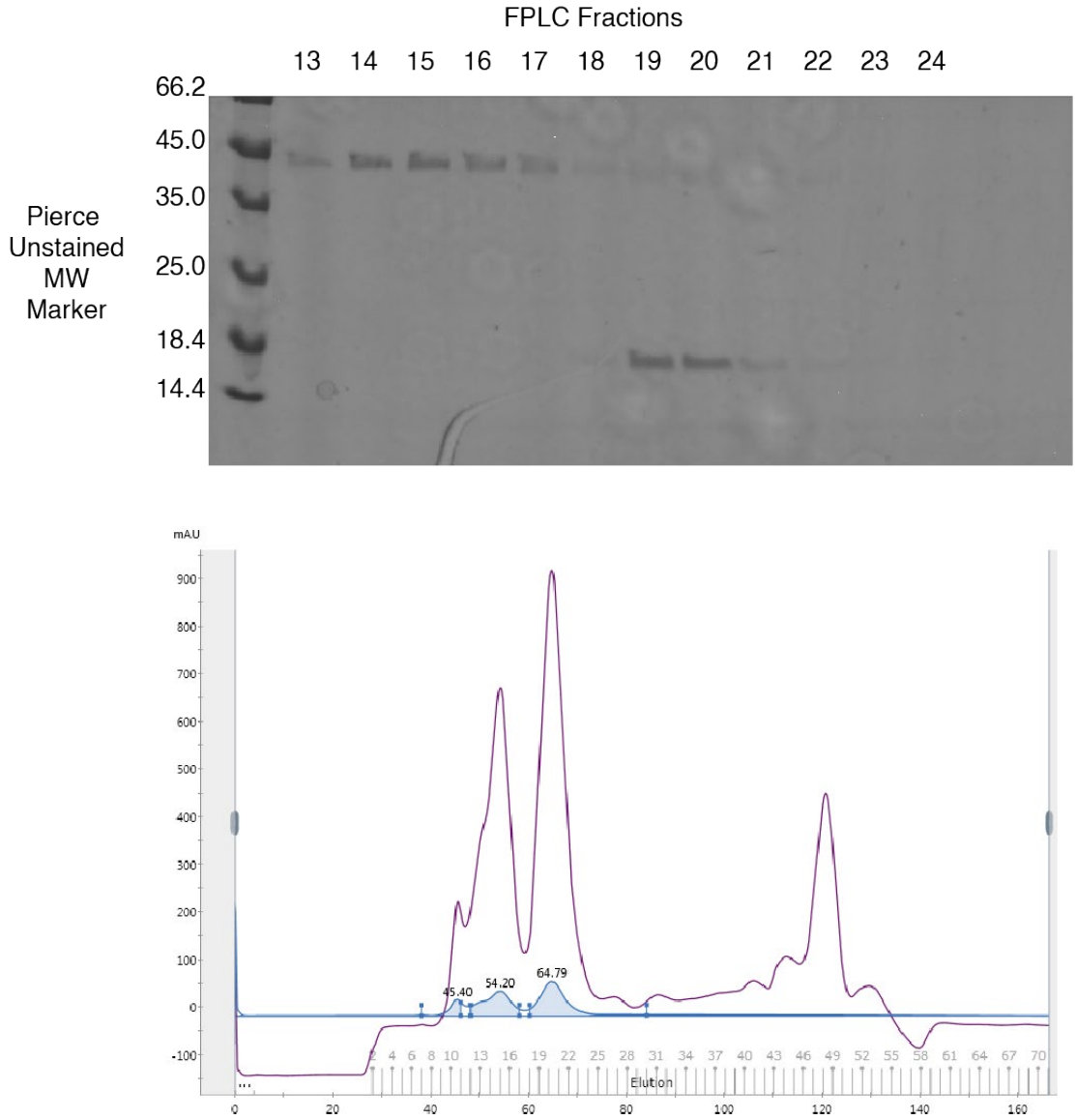



**Figure S11 Purification of AcpP**. An example of the purification of AcpP before an NMR experiment. Octanoyl-loaded AcpP was purified fresh for each titration. A 12% acrylamide denaturing SDS gel is shown. AcpP has very little absorbance at 280nm, as such the absorbance at 214 is also shown in magenta to help locate the AcpP's fractions. AcpP can be seen in fractions 19, 20, and 21, characteristically it travels unusually high on an SDS gel.

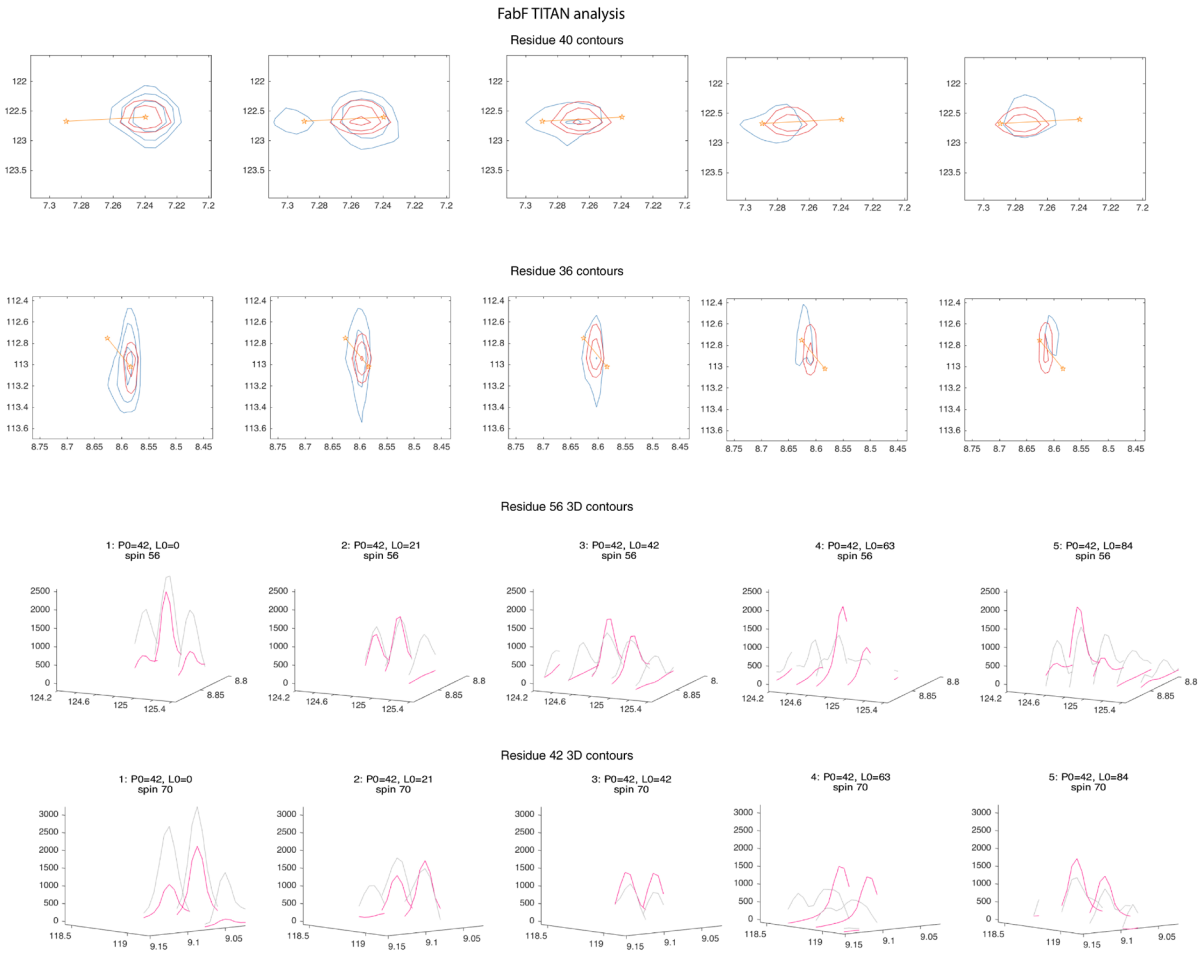



**Figure S12 Titan peak fitting of FabF titration data.** The FabF spectra and models were exported from the TITAN analysis program directly. Analysis was performed using the flexible docking method. Error analysis was performed using 300 steps of bootstrap error analysis. There was signal loss in the FabF experiment, likely due to the instability of the FabF protein and the presence of crashed partner protein in the titration.

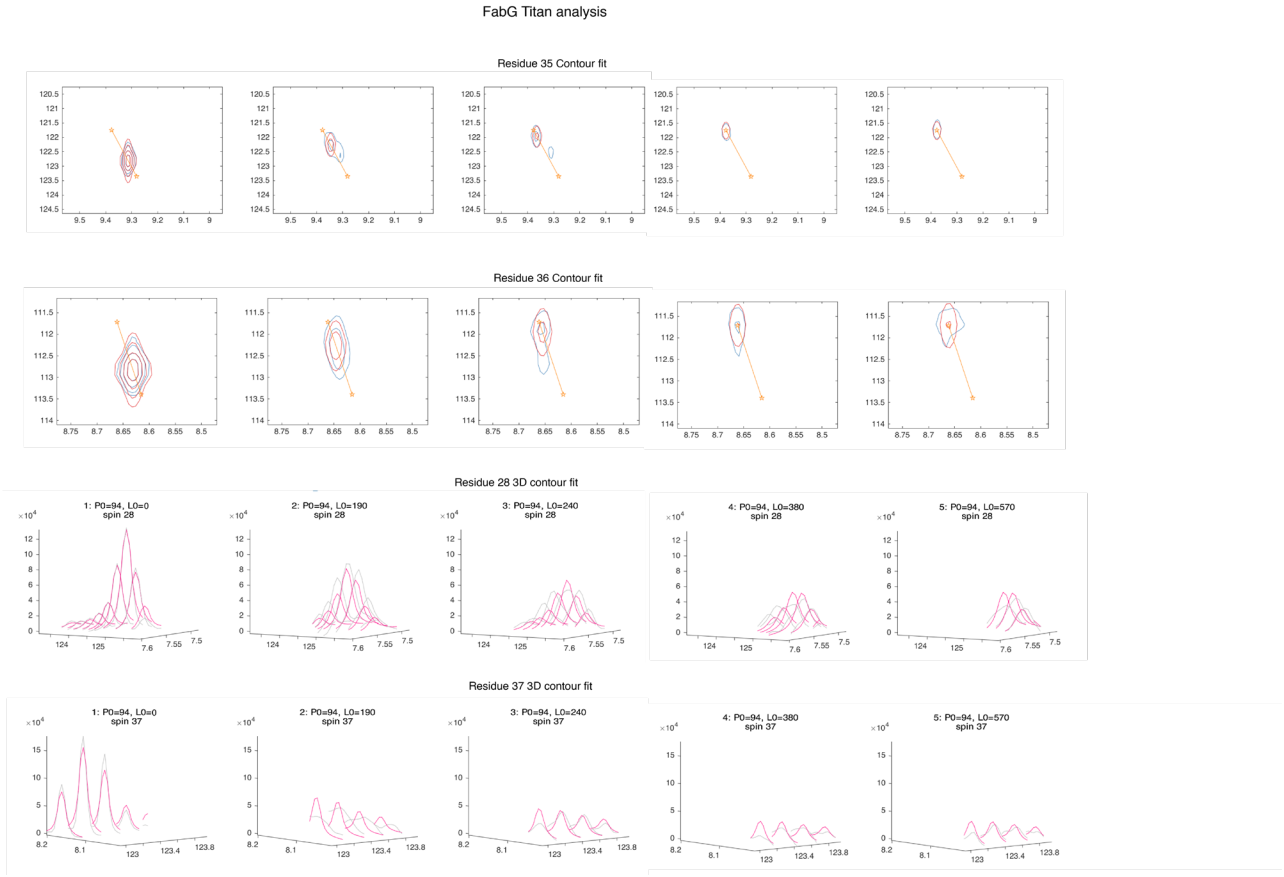



**Figure S13 Titan peak fitting of FabG titration data.** The FabG spectra and models were exported from the TITAN analysis program directly. Analysis was performed using the flexible docking method. Error analysis was performed using 300 steps of bootstrap error analysis.

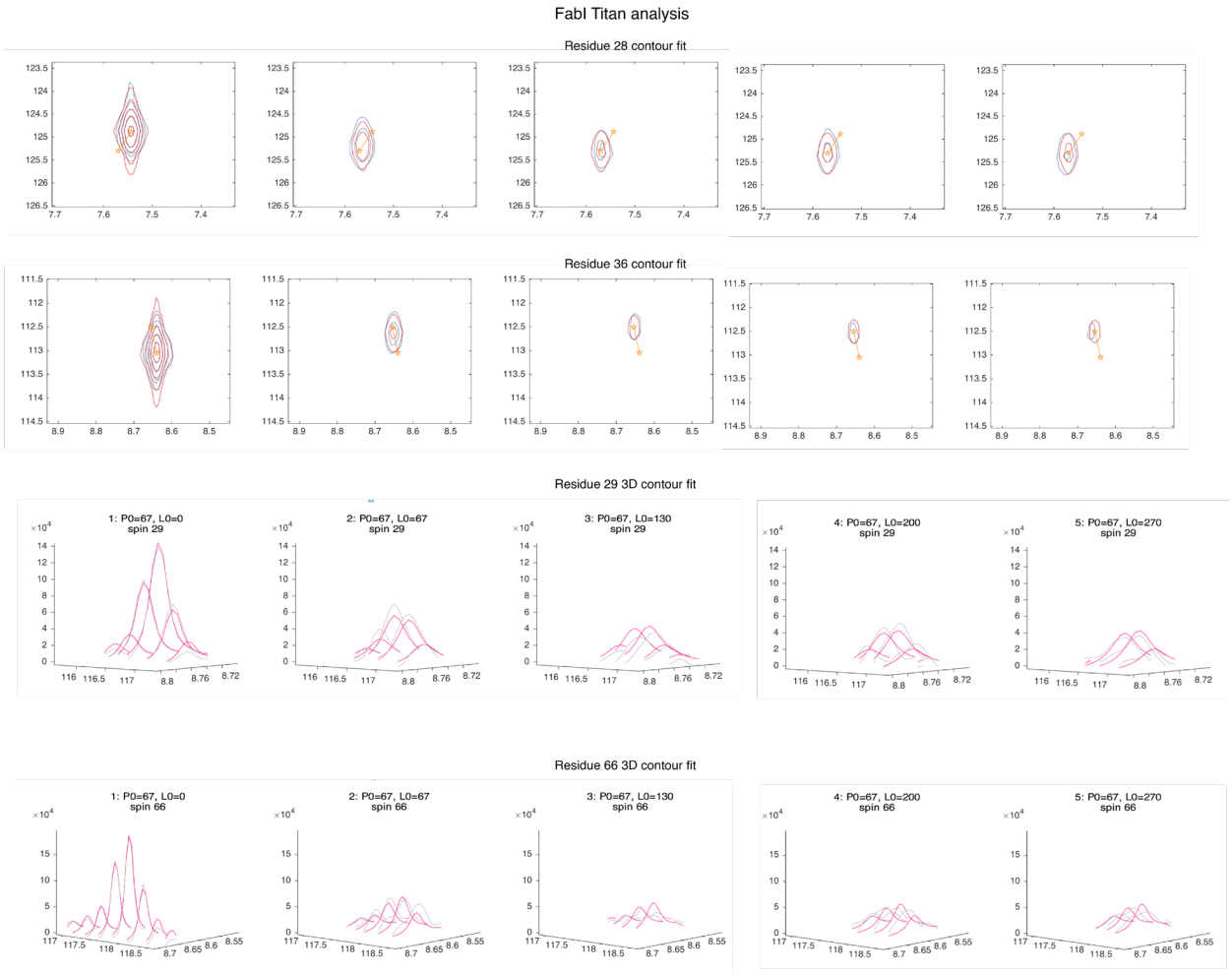



**Figure S14 Titan peak fitting of FabI titration data.** The FabI spectra and models were exported from the TITAN analysis program directly. Analysis was performed using the flexible docking method. Error analysis was performed using 300 steps of bootstrap error analysis

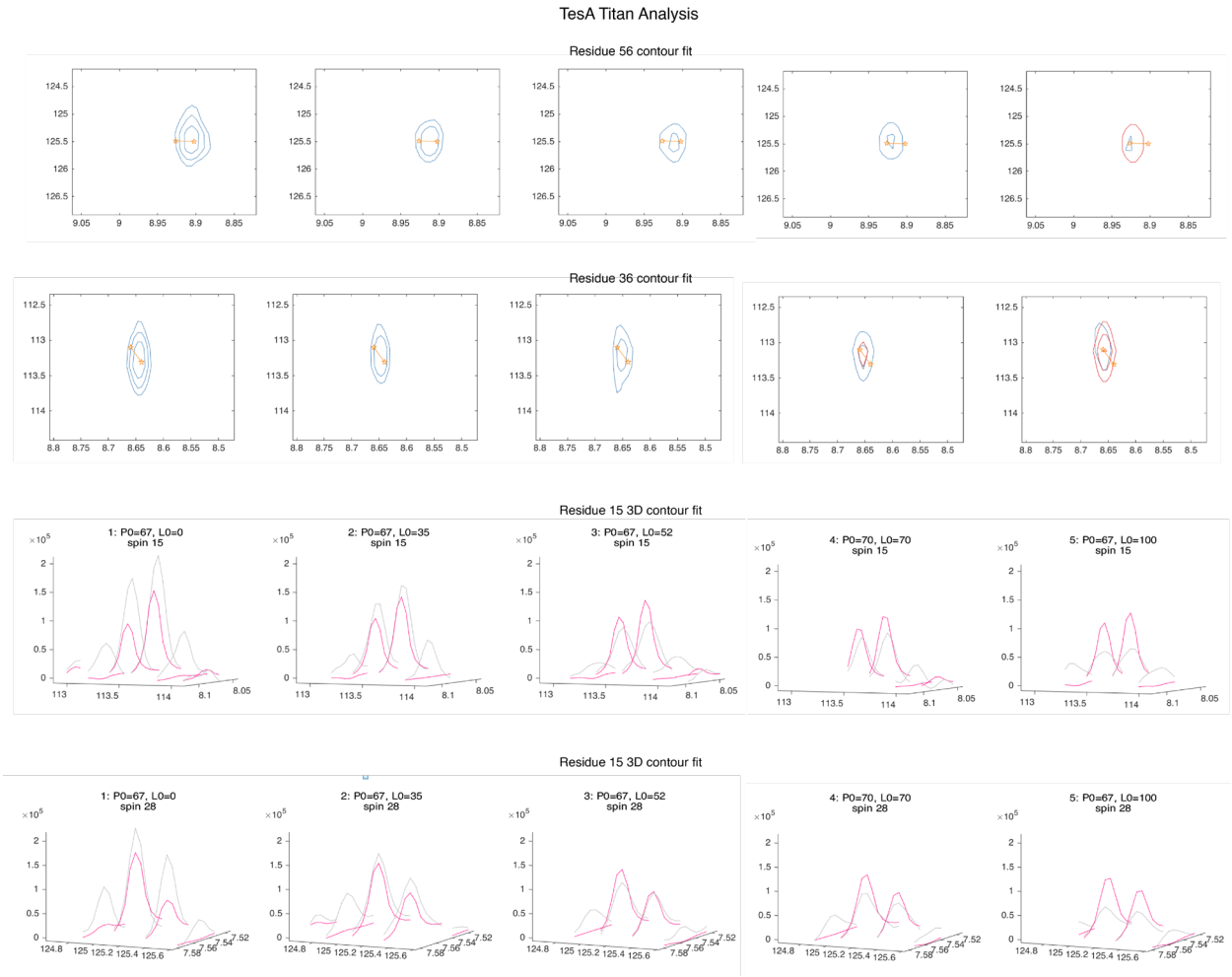



**Figure S15 Titan peak fitting of TesA titration data.** The TesA spectra and models were exported from the TITAN analysis program directly. Analysis was performed using the flexible docking method. Error analysis was performed using 300 steps of bootstrap error analysis

**Table S1 Interface and global alignment scores for docked models.** Alignments are generated from ICM docked models and aligned to crosslinked structures. The “informed” ACP residues are supplied form NMR titration data and the “informed” partner residues were taken from published data on essential electrostatic residues when possible.

| Partner | Interface alignment (Å) | Global alignment (Å) | ACP “informed” residues supplied | Partner “informed” residues supplied |
|---|---|---|---|---|
| **FabA** | 1.10 | 7.4 | 39,41,47 | 132,136 |
| **FabZ** | 1.19 | 5.39 | N/A | N/A |
| **FabB** | 0.84 | 8.5 | 35,38 | 61,62,65 |
| **FabF** | 1.58 | 6.7 | 35,39 | 65,212 |
| **FabI** | N/A | N/A | 35,44,47 | 201,204,205 |
| **FabG** | N/A | N/A | 35,39,47 | 14 |
| **TesA** | N/A | N/A | 31,44 | N/A |

**Table S2 Table of the docked partner characteristics.** Surface area is generated from the pymol "get area" command. Energetics values are taken directly from ICM.

| **Partner Enzyme** | Surface area ($Å^2$) | Total binding energetics | Van der waals energetics | Electrostatic energetics |
|---|---|---|---|---|
| **FabF** | 1022 | -41.4 | -42.8 | -13 |
| **FabB** | 963 | -49.32 | -39.9 | -12.7 |
| **FabI** | 1040 | -46.4 | -48.5 | -11.83 |
| **FabG** | 1040 | -53.8 | -59.6 | -8.5 |
| **FabA** | 704 | -41.1 | -28.5 | -16.0 |
| **FabZ** | 718 | -58.3 | -41.1 | -17.9 |
| **TesA** | 850 | -41.85 | -46 | -6.9 |

**Table S3: Peak list of zero point and saturated FabF titration points.** CSPs were calculated form these chemical shifts using the CSP equation: $CSP = \sqrt{\frac{1}{2}[\delta_H^2 + (\alpha \cdot \delta_N^2)}$

| RESIDUE | ZERO POINT $^{1}$H PEAK | ZERO POINT $^{15}$N PEAK | SATURATED $^{1}$H PEAK | SATURATED $^{15}$N PEAK |
|---|---|---|---|---|
| **03I** | 8.57 | 121.215 | 8.574 | 121.214 |
| **04E** | 8.532 | 118.385 | 8.522 | 118.413 |
| **05E** | 7.782 | 117.169 | 7.753 | 117.067 |
| **06R** | 8.271 | 119.589 | 8.272 | 119.731 |
| **07V** | 8.847 | 118.887 | 8.869 | 118.731 |
| **08K** | 8.152 | 116.799 | 8.152 | 116.683 |
| **09K** | 8.199 | 120.264 | 8.188 | 120.226 |
| **10I** | 7.561 | 119.173 | 7.566 | 118.388 |
| **11I** | 8.243 | 118.641 | 8.226 | 118.693 |
| **12G** | 8.399 | 105.015 | 8.352 | 105.058 |
| **13E** | 8.142 | 119.956 | 8.109 | 120.23 |
| **14Q** | 8.346 | 117.156 | 8.441 | 117.299 |
| **15L** | 8.024 | 113.236 | 8.012 | 112.196 |
| **16G** | 7.745 | 109.597 | 7.716 | 109.732 |
| **17V** | 7.812 | 114.458 | 7.808 | 113.903 |
| **18K** | 8.466 | 122.713 | 8.483 | 122.49 |
| **19Q** | 8.704 | 122.536 | 8.705 | 122.367 |
| **20E** | 9.329 | 116.26 | 9.324 | 116.48 |
| **21E** | 7.81 | 116.694 | 7.787 | 116.658 |
| **22V** | 7.464 | 122.103 | 7.474 | 122.087 |
| **23T** | 7.27 | 115.423 | 7.229 | 115.118 |
| **24N** | 8.532 | 118.385 | 8.522 | 118.413 |
| **25N** | 8.026 | 111.683 | 8.024 | 111.458 |
| **26A** | 7.236 | 122.607 | 7.273 | 122.529 |
| **27S** | 9.871 | 116.772 | 9.9 | 117.013 |
| **28F** | 7.498 | 124.698 | 7.494 | 125.616 |
| **29V** | 8.682 | 116.397 | 8.638 | 116.655 |
| **30E** | 8.218 | 116.529 | 8.214 | 116.429 |
| **31D** | 7.709 | 113.747 | 7.766 | 113.265 |

| | | | | |
|---|---|---|---|---|
| **32L** | 7.308 | 115.187 | 7.282 | 115.26 |
| **33G** | 7.158 | 106.181 | 7.303 | 106.436 |
| **34A** | 8.393 | 122.45 | 8.421 | 122.644 |
| **35D** | 9.237 | 122.792 | 9.256 | 121.955 |
| **36S** | 8.588 | 112.988 | 8.608 | 112.672 |
| **37L** | 8.114 | 123.253 | 8.13 | 123.138 |
| **38D** | 8.271 | 119.589 | 8.291 | 119.815 |
| **39T** | 8.116 | 111.653 | 8.055 | 112.295 |
| **40V** | 7.174 | 120.936 | 7.169 | 121.703 |
| **41E** | 7.725 | 119.111 | 7.784 | 119.1 |
| **42L** | 8.351 | 121.134 | 8.285 | 120.999 |
| **43V** | 7.925 | 118.949 | 7.934 | 118.975 |
| **44M** | 7.727 | 116.101 | 7.749 | 116.589 |
| **45A** | 8.059 | 120.675 | 8.016 | 120.219 |
| **46L** | 8.318 | 119.97 | 8.303 | 120.187 |
| **47E** | 8.581 | 119.522 | 8.506 | 119.389 |
| **48E** | 7.81 | 116.694 | 7.787 | 116.588 |
| **49E** | 7.871 | 119.406 | 7.884 | 119.412 |
| **50F** | 7.702 | 111.944 | 7.709 | 111.952 |
| **51D** | 7.809 | 122.137 | 7.777 | 121.763 |
| **52T** | 7.971 | 112.022 | 8.012 | 112.196 |
| **53E** | 8.049 | 122.387 | 8.037 | 122.645 |
| **54I** | 10.325 | 129.462 | 10.293 | 128.625 |
| **55P** | 0 | 0 | 0 | 0 |
| **56D** | 8.823 | 124.934 | 8.847 | 124.753 |
| **57E** | 9.172 | 115.993 | 9.26 | 116.394 |
| **58E** | 7.156 | 115.712 | 7.146 | 115.849 |
| **59A** | 8.114 | 123.253 | 8.13 | 123.138 |
| **60E** | 7.493 | 112.037 | 7.485 | 112.4 |
| **61K** | 6.969 | 113.744 | 7.023 | 113.929 |
| **62I** | 7.569 | 122.511 | 7.57 | 121.944 |
| **63T** | 7.971 | 112.022 | 8.012 | 112.196 |
| **64T** | 7.171 | 110.106 | 7.031 | 110.465 |
| **65V** | 7.931 | 121.163 | 7.911 | 120.502 |
| **66Q** | 8.616 | 117.698 | 8.458 | 117.439 |
| **67A** | 7.725 | 119.111 | 7.807 | 119.385 |
| **68A** | 7.89 | 122.432 | 7.82 | 121.446 |

| | | | | |
|---|---|---|---|---|
| **69I** | 8.053 | 118.918 | 8.005 | 118.68 |
| **70D** | 9.051 | 118.888 | 9.088 | 118.642 |
| **71Y** | 8.105 | 121.417 | 8.101 | 120.881 |
| **72I** | 8.122 | 120.58 | 8.1 | 120.905 |
| **73N** | 8.769 | 117.956 | 8.77 | 117.97 |
| **74G** | 7.779 | 104.714 | 7.749 | 104.743 |
| **75H** | 7.587 | 118.351 | 7.566 | 118.388 |

**Table S4: Peak list of zero point and saturated FabI titration points.** CSPs were calculated form these chemical shifts using the CSP equation: $CSP = \sqrt{\frac{1}{2}[\delta_H^2 + (\alpha \cdot \delta_N^2)}$

| RESIDUE | ZERO POINT $^{1}$H PEAK | ZERO POINT $^{15}$N PEAK | SATURATED $^{1}$H PEAK | SATURATED $^{15}$N PEAK |
|---|---|---|---|---|
| **03I** | 8.627 | 121.2 | 8.619 | 121.18 |
| **04E** | 8.595 | 118.32 | 8.589 | 118.32 |
| **05E** | 7.838 | 117.15 | 7.833 | 117.1 |
| **06R** | 8.329 | 119.57 | 8.342 | 119.71 |
| **07V** | 8.903 | 118.83 | 8.897 | 118.88 |
| **08K** | 8.208 | 116.76 | 8.211 | 116.82 |
| **09K** | 8.261 | 120.12 | 8.24 | 120.24 |
| **10I** | 7.605 | 119.06 | 7.628 | 118.69 |
| **11I** | 8.293 | 118.63 | 8.246 | 118.64 |
| **12G** | 8.448 | 105 | 8.436 | 104.99 |
| **13E** | 8.188 | 119.89 | 8.18 | 119.9 |
| **14Q** | 8.4 | 117.11 | 8.378 | 117.13 |
| **15L** | 8.079 | 113.21 | 8.005 | 112.85 |
| **16G** | 7.793 | 109.61 | 7.788 | 109.59 |
| **17V** | 7.864 | 114.42 | 7.848 | 114.17 |
| **18K** | 8.526 | 122.69 | 8.523 | 122.65 |
| **19Q** | 8.765 | 122.47 | 8.759 | 122.37 |
| **20E** | 9.389 | 116.29 | 9.373 | 116.25 |
| **21E** | 7.864 | 116.7 | 7.851 | 116.66 |
| **22V** | 7.515 | 122.09 | 7.516 | 122.09 |
| **23T** | 7.328 | 115.36 | 7.33 | 115.3 |
| **24N** | 8.567 | 118.43 | 8.57 | 118.36 |
| **25N** | 8.085 | 111.71 | 8.076 | 112.03 |
| **26A** | 7.29 | 122.57 | 7.306 | 122.54 |
| **27S** | 9.929 | 116.71 | 9.947 | 116.83 |
| **28F** | 7.544 | 124.71 | 7.571 | 125.19 |
| **29V** | 8.746 | 116.41 | 8.754 | 116.63 |
| **30E** | 8.281 | 116.48 | 8.293 | 116.42 |
| **31D** | 7.767 | 113.67 | 7.787 | 113.48 |
| **32L** | 7.346 | 115.16 | 7.333 | 115.33 |
| **33G** | 7.225 | 106.21 | 7.259 | 106.26 |
| **34A** | 8.462 | 122.48 | 8.467 | 122.47 |

| | | | | |
|---|---|---|---|---|
| **35D** | 9.294 | 122.74 | 9.279 | 122.32 |
| **36S** | 8.64 | 112.89 | 8.664 | 112.34 |
| **37L** | 8.157 | 123.31 | 8.056 | 123.44 |
| **38D** | 8.329 | 119.57 | 8.342 | 119.71 |
| **39T** | 8.158 | 111.54 | 8.13 | 111.27 |
| **40V** | 7.221 | 120.84 | 7.2 | 120.74 |
| **41E** | 7.796 | 118.85 | 7.838 | 118.99 |
| **42L** | 8.383 | 121.05 | 8.434 | 121.01 |
| **43V** | 7.952 | 118.77 | 7.944 | 119.31 |
| **44M** | 7.74 | 116.68 | 7.851 | 116.66 |
| **45A** | 8.11 | 120.75 | 8.177 | 120.36 |
| **46L** | 8.377 | 120.16 | 8.362 | 119.87 |
| **47E** | 8.669 | 119.5 | 8.558 | 119.26 |
| **48E** | 7.838 | 117.15 | 7.833 | 117.1 |
| **49E** | 7.944 | 119.12 | 7.944 | 119.31 |
| **50F** | 7.788 | 111.22 | 7.764 | 111.4 |
| **51D** | 7.849 | 122.06 | 7.838 | 121.96 |
| **52T** | 8.039 | 111.93 | 8.07 | 112.09 |
| **53E** | 8.097 | 122.31 | 8.076 | 122.4 |
| **54I** | 0 | 0 | 0 | 0 |
| **55P** | 0 | 0 | 0 | 0 |
| **56D** | 8.884 | 125.03 | 8.884 | 124.65 |
| **57E** | 9.24 | 115.94 | 9.292 | 116.15 |
| **58E** | 7.207 | 115.63 | 7.208 | 115.66 |
| **59A** | 8.157 | 123.31 | 8.251 | 123.66 |
| **60E** | 7.549 | 112.13 | 7.545 | 112.43 |
| **61K** | 7.007 | 113.71 | 7.048 | 113.91 |
| **62I** | 7.628 | 122.43 | 7.59 | 122.11 |
| **63T** | 8.039 | 111.93 | 8.07 | 112.09 |
| **64T** | 7.207 | 110.13 | 7.162 | 110.37 |
| **65V** | 7.988 | 121.12 | 7.998 | 120.92 |
| **66Q** | 8.667 | 117.66 | 8.549 | 117.51 |
| **67A** | 7.796 | 118.85 | 7.853 | 118.99 |
| **68A** | 7.946 | 122.39 | 7.947 | 122.86 |
| **69I** | 8.096 | 118.94 | 8.056 | 118.67 |
| **70D** | 9.107 | 118.78 | 9.131 | 118.64 |
| **71Y** | 8.123 | 121.12 | 8.125 | 121.13 |

| | | | | |
|---|---|---|---|---|
| **72I** | 8.188 | 119.89 | 8.18 | 119.9 |
| **73N** | 8.81 | 117.79 | 8.809 | 117.85 |
| **74G** | 7.832 | 104.75 | 7.811 | 104.97 |
| **75H** | 7.644 | 118.42 | 7.628 | 118.69 |

**Table S5: Peak list of zeropoint and saturated FabG titration points.** CSPs were calculated form these chemical shifts using the CSP equation: $CSP = \sqrt{\frac{1}{2}[\delta_H^2 + (\alpha \cdot \delta_N^2)}$

| RESIDUE | ZERO POINT $^{1}$H PEAK | ZERO POINT $^{15}$N PEAK | SATURATED $^{1}$H PEAK | SATURATED $^{15}$N PEAK |
|---|---|---|---|---|
| **03I** | 8.616 | 121.22 | 8.623 | 121.22 |
| **04E** | 8.585 | 118.32 | 8.595 | 118.34 |
| **05E** | 7.836 | 117.12 | 7.841 | 117.13 |
| **06R** | 8.329 | 119.61 | 8.341 | 119.69 |
| **07V** | 8.893 | 118.76 | 8.899 | 118.86 |
| **08K** | 8.209 | 116.75 | 8.243 | 116.87 |
| **09K** | 8.244 | 120.33 | 8.232 | 120.23 |
| **10I** | 7.605 | 119.12 | 7.636 | 118.75 |
| **11I** | 8.291 | 118.63 | 8.236 | 118.89 |
| **12G** | 8.451 | 105 | 8.447 | 105.12 |
| **13E** | 8.188 | 119.93 | 8.197 | 119.92 |
| **14Q** | 8.396 | 117.13 | 8.382 | 117.2 |
| **15L** | 8.075 | 113.17 | 7.992 | 112.96 |
| **16G** | 7.794 | 109.6 | 7.789 | 109.64 |
| **17V** | 7.862 | 114.35 | 7.854 | 114.18 |
| **18K** | 8.525 | 122.66 | 8.536 | 122.63 |
| **19Q** | 8.766 | 122.45 | 8.775 | 122.37 |
| **20E** | 9.377 | 116.25 | 9.388 | 116.32 |
| **21E** | 7.859 | 116.68 | 7.861 | 116.66 |
| **22V** | 7.516 | 122.11 | 7.523 | 122.09 |
| **23T** | 7.31 | 115.27 | 7.334 | 115.31 |
| **24N** | 8.585 | 118.32 | 8.595 | 118.34 |
| **25N** | 8.082 | 111.7 | 8.065 | 110.94 |
| **26A** | 7.29 | 122.55 | 7.315 | 122.56 |
| **27S** | 9.929 | 116.69 | 9.938 | 116.93 |
| **28F** | 7.538 | 124.68 | 7.533 | 125.43 |
| **29V** | 8.747 | 116.37 | 8.735 | 116.54 |
| **30E** | 8.275 | 116.47 | 8.295 | 116.43 |
| **31D** | 7.76 | 113.68 | 7.793 | 113.38 |
| **32L** | 7.353 | 115.12 | 7.375 | 115.44 |
| **33G** | 7.207 | 106.2 | 7.304 | 106.25 |
| **34A** | 8.449 | 122.38 | 8.449 | 122.5 |

| | | | | |
|---|---|---|---|---|
| **35D** | 9.315 | 122.83 | 9.379 | 121.71 |
| **36S** | 8.631 | 112.84 | 8.657 | 111.64 |
| **37L** | 8.163 | 123.24 | 8.048 | 123.44 |
| **38D** | 8.311 | 119.38 | 8.341 | 119.69 |
| **39T** | 8.166 | 111.54 | 8.154 | 112.13 |
| **40V** | 7.21 | 120.78 | 7.243 | 121.79 |
| **41E** | 7.797 | 118.79 | 7.848 | 119.18 |
| **42L** | 8.388 | 121.09 | 8.445 | 120.88 |
| **43V** | 7.975 | 118.92 | 7.958 | 119.28 |
| **44M** | 7.729 | 116.63 | 7.771 | 116.38 |
| **45A** | 8.122 | 121.22 | 8.127 | 121.11 |
| **46L** | 8.365 | 120.21 | 8.374 | 120.13 |
| **47E** | 8.677 | 119.63 | 8.558 | 119.97 |
| **48E** | 7.834 | 116.81 | 7.841 | 117.13 |
| **49E** | 7.93 | 119.38 | 7.958 | 119.28 |
| **50F** | 7.782 | 111.24 | 7.731 | 111.84 |
| **51D** | 7.843 | 121.98 | 7.789 | 121.74 |
| **52T** | 8.027 | 111.95 | 8.084 | 111.75 |
| **53E** | 8.081 | 122.23 | 8.072 | 122.36 |
| **54I** | 0 | 0 | 0 | 0 |
| **55P** | 0 | 0 | 0 | 0 |
| **56D** | 8.883 | 124.98 | 8.953 | 124.78 |
| **57E** | 9.237 | 115.95 | 9.285 | 116.15 |
| **58E** | 7.201 | 115.67 | 7.189 | 115.68 |
| **59A** | 8.15 | 122.45 | 8.251 | 122.66 |
| **60E** | 7.545 | 111.96 | 7.548 | 112.42 |
| **61K** | 7.011 | 113.85 | 7.003 | 113.45 |
| **62I** | 7.612 | 122.39 | 7.606 | 121.99 |
| **63T** | 8.027 | 111.95 | 8.084 | 111.75 |
| **64T** | 7.215 | 110.14 | 7.086 | 110.4 |
| **65V** | 7.98 | 121.12 | 8.029 | 120.92 |
| **66Q** | 8.666 | 117.62 | 8.595 | 118.34 |
| **67A** | 7.764 | 118.95 | 7.818 | 119.13 |
| **68A** | 7.931 | 122.34 | 7.866 | 121.43 |
| **69I** | 8.094 | 118.79 | 8.069 | 118.69 |
| **70D** | 9.098 | 118.77 | 9.152 | 118.58 |
| **71Y** | 8.122 | 121.22 | 8.127 | 121.11 |

| | | | | |
|---|---|---|---|---|
| **72I** | 8.188 | 120.44 | 8.179 | 120.58 |
| **73N** | 8.806 | 117.94 | 8.812 | 117.92 |
| **74G** | 7.817 | 105.03 | 7.812 | 105.04 |
| **75H** | 7.637 | 118.7 | 7.636 | 118.75 |

**Table S6: Peak list of zero point and saturated TesA titration points.** CSPs were calculated form these chemical shifts using the CSP equation: $CSP = \sqrt{\frac{1}{2}[\delta_H^2 + (\alpha \cdot \delta_N^2)}$

| RESIDUE | ZERO POINT $^{1}$H PEAK | ZERO POINT $^{15}$N PEAK | SATURATED $^{1}$H PEAK | SATURATED $^{15}$N PEAK |
|---|---|---|---|---|
| **03I** | 8.641 | 121.46 | 8.638 | 121.45 |
| **04E** | 8.603 | 118.64 | 8.601 | 118.64 |
| **05E** | 7.837 | 117.36 | 7.836 | 117.34 |
| **06R** | 8.332 | 119.72 | 8.336 | 119.86 |
| **07V** | 8.907 | 119.03 | 8.903 | 118.95 |
| **08K** | 8.219 | 117.09 | 8.222 | 117.11 |
| **09K** | 8.269 | 120.41 | 8.26 | 120.39 |
| **10I** | 7.613 | 119.29 | 7.616 | 119.31 |
| **11I** | 8.308 | 118.87 | 8.299 | 118.89 |
| **12G** | 8.472 | 105.3 | 8.47 | 105.32 |
| **13E** | 8.198 | 120.09 | 8.197 | 120.04 |
| **14Q** | 8.401 | 117.38 | 8.388 | 117.25 |
| **15L** | 8.085 | 113.38 | 8.072 | 113.35 |
| **16G** | 7.807 | 109.75 | 7.802 | 109.73 |
| **17V** | 7.87 | 114.55 | 7.858 | 114.44 |
| **18K** | 8.528 | 123.03 | 8.527 | 122.99 |
| **19Q** | 8.781 | 122.87 | 8.783 | 122.82 |
| **20E** | 9.409 | 116.51 | 9.416 | 116.53 |
| **21E** | 7.87 | 116.91 | 7.864 | 116.89 |
| **22V** | 7.521 | 122.25 | 7.518 | 122.22 |
| **23T** | 7.325 | 115.54 | 7.325 | 115.46 |
| **24N** | 8.576 | 118.78 | 8.601 | 118.64 |
| **25N** | 8.085 | 111.91 | 8.073 | 111.96 |
| **26A** | 7.288 | 122.8 | 7.289 | 122.76 |
| **27S** | 9.929 | 116.91 | 9.931 | 117.15 |
| **28F** | 7.553 | 124.99 | 7.557 | 125.03 |
| **29V** | 8.753 | 116.61 | 8.757 | 116.64 |
| **30E** | 8.297 | 116.73 | 8.298 | 116.71 |
| **31D** | 7.793 | 114.38 | 7.759 | 113.74 |
| **32L** | 7.349 | 115.35 | 7.325 | 115.46 |
| **33G** | 7.223 | 106.4 | 7.225 | 106.38 |
| **34A** | 8.447 | 122.59 | 8.448 | 122.55 |

| | | | | |
|---|---|---|---|---|
| **35D** | 9.304 | 123.07 | 9.308 | 122.91 |
| **36S** | 8.646 | 113.13 | 8.654 | 112.93 |
| **37L** | 8.167 | 123.47 | 8.16 | 123.45 |
| **38D** | 8.332 | 119.72 | 8.336 | 119.86 |
| **39T** | 8.162 | 111.74 | 8.156 | 111.58 |
| **40V** | 7.222 | 120.99 | 7.224 | 120.97 |
| **41E** | 7.805 | 119.03 | 7.829 | 119.16 |
| **42L** | 8.396 | 121.29 | 8.399 | 120.99 |
| **43V** | 7.974 | 119.03 | 7.986 | 119.14 |
| **44M** | 7.758 | 116.92 | 7.864 | 116.89 |
| **45A** | 8.128 | 121.39 | 8.129 | 120.94 |
| **46L** | 8.383 | 120.4 | 8.389 | 120.38 |
| **47E** | 8.694 | 119.86 | 8.678 | 119.74 |
| **48E** | 7.843 | 116.94 | 7.864 | 116.89 |
| **49E** | 7.939 | 119.55 | 7.949 | 119.4 |
| **50F** | 7.779 | 111.41 | 7.767 | 111.52 |
| **51D** | 7.862 | 122.21 | 7.851 | 122.13 |
| **52T** | 8.018 | 112.08 | 8.063 | 112.12 |
| **53E** | 8.109 | 122.44 | 8.108 | 122.4 |
| **54I** | 0 | 0 | 0 | 0 |
| **55P** | 0 | 0 | 0 | 0 |
| **56D** | 8.906 | 125.3 | 8.923 | 125.33 |
| **57E** | 9.253 | 116.19 | 9.287 | 116.14 |
| **58E** | 7.21 | 115.89 | 7.201 | 115.91 |
| **59A** | 8.175 | 122.84 | 8.177 | 122.84 |
| **60E** | 7.549 | 112.19 | 7.551 | 112.3 |
| **61K** | 7.024 | 113.99 | 7.022 | 113.94 |
| **62I** | 7.62 | 122.63 | 7.62 | 122.46 |
| **63T** | 8.018 | 112.08 | 7.982 | 112.24 |
| **64T** | 7.231 | 110.21 | 7.217 | 110.22 |
| **65V** | 7.988 | 121.4 | 7.988 | 121.33 |
| **66Q** | 8.676 | 117.92 | 8.667 | 117.88 |
| **67A** | 7.775 | 119.23 | 7.782 | 119.2 |
| **68A** | 7.938 | 122.63 | 7.932 | 122.54 |
| **69I** | 8.111 | 119.03 | 8.103 | 118.98 |
| **70D** | 9.107 | 119.08 | 9.109 | 119.06 |
| **71Y** | 8.156 | 121.59 | 8.151 | 121.45 |

| | | | | |
|---|---|---|---|---|
| **72I** | 8.192 | 120.69 | 8.195 | 120.64 |
| **73N** | 8.832 | 118.2 | 8.825 | 118.18 |
| **74G** | 7.826 | 104.93 | 7.828 | 104.96 |
| **75H** | 7.64 | 118.56 | 7.643 | 118.6 |

**Table S7: Comparisons of the residue informed and uninformed top 10 docking poses.** The 10 most energetically favorable poses from otherwise identical “informed” and “uninformed” docking calculations.

| ***Pose*** | **FabF "informed" (Å)** | **FabF "uninformed" (Å)** | **FabB "informed" (Å)** | **FabB "uninformed" (Å)** | **FabA "informed" (Å)** | **FabA "uninformed" (Å)** |
|---|---|---|---|---|---|---|
| ***1*** | 6.7 | 14.8 | 8.5 | 8.5 | 7.4 | 7.4 |
| ***2*** | 9.2 | 17.7 | 2.9 | 19.9 | 12.9 | 12.9 |
| ***3*** | 5.8 | 26.2 | 6.7 | 4.5 | 7.8 | 7.8 |
| ***4*** | 7.4 | 26 | 4.2 | 11.7 | 6.7 | 6.7 |
| ***5*** | 10.4 | 10.4 | 4.1 | 12.2 | 6.8 | 6.8 |
| ***6*** | 8.5 | 21.7 | 4.5 | 17.5 | 12.8 | 12.8 |
| ***7*** | 8.1 | 19.1 | 6.2 | 7.8 | 8 | 13.2 |
| ***8*** | 4.2 | 22.3 | 6.3 | 6.8 | 13.2 | 5.8 |
| ***9*** | 9.3 | 22.4 | 4.7 | 8.5 | 5.8 | 18.4 |
| ***10*** | 10.2 | 20.2 | 7.6 | 7.9 | 18.4 | 7.1 |
| ***Average*** | 7.98 | 20.08 | 5.57 | 10.53 | 9.98 | 9.89 |

**Table S8: Comparisons of online docking resources' ability to recreate the FabF-AcpP crosslinked interface.** RMSD calculations were performed using the same methods as the ICM docked structures. All calculations were performed as described in the methods on free to use docking websites.

| *Software scoring rank* | ClusPro balanced scoring RMSD (Å) | ClusPro electrostatic scoring RMSD (Å) | Rosie: refining Cluspro electrostatic rank 1 structure RMSD (Å) | HADDOCK ICM prepared files RMSD (Å) | HADDOCK raw PDB files RMSD (Å) |
|---|---|---|---|---|---|
| ***1*** | 12.98 | 4.56 | 4.54 | 6.4 | 17.4 |
| ***2*** | 14.95 | 10.06 | 4.53 | 6 | 17.7 |
| ***3*** | 12.77 | 7.7 | 4.57 | 6.33 | 17.6 |

**Table S9: Thermodynamic and Kinetic parameters of the elongating enzymes and a thioesterase of *E. coli* Fatty Acid Biosynthesis binding an amide linked octanoyl AcpP.** Parameters were calculated using the TITAN lineshape analysis software.

| Protein | $K_d$ (μM) | $K_{off}$ ($s^{-1}$) | n |
|---|---|---|---|
| **FabF** | 6.9±8.5 | 3350±3310 | 1.3±0.3 |
| **FabB** | 37.6±6.6 | | |
| **FabI** | 1.6±1.2 | 6360±3260 | 1.2±0.1 |
| **FabG** | 52.3±27.5 | 3559±2061 | 2.2±0.4 |
| **TesA** | 13±8 | 9750±850 | monomer |

## Chapter 2. Control of unsaturation in de novo fatty acid biosynthesis by FabA

### 2.1 Introduction

Fatty acid biosynthesis (FAB), an essential primary metabolic pathway found across all domains of life, which produces not only the fatty acids required for membrane biogenesis and energy storage, is central to the biosynthesis of essential cofactors and cellular regulatory processes[1,2]. In *E. coli*, FAB exists as a precisely coordinated multi-enzyme pathway that must maintain control over both iterative and substrate-selective reactivity in order to maintain cellular homeostasis and environmental response[3–5]. This type II FAB expresses all of the protein components as stand-alone and freely diffusing enzymes/domains (Fig 1a), with *E. coli* FAB associated with more than 25 known participating enzymes[6–8]. Still, the membranes of *E. coli,* and by extension the resulting FAB products, are evolutionarily selected and dynamically tuned to the environment[9,10]. While the microbiological phenomena of fatty acid composition specificity have been observed[11,12], the molecular and enzymatic details of this control is just beginning to be unraveled. There remain multiple unknowns around how these enzymes can maintain selectivity over the dozens of possible substrates upon which they act to regulate such a complex and essential pathway.

An important mechanism may lie in the protein-protein interactions (PPIs) of the central FAB player, the acyl carrier protein (ACP), in its interactions with FAB enzymes[13,14]. The *E. coli* ACP, AcpP, is a small (~10 kDa), four-helical protein, with high homology to many other organisms' FAB systems. As such, AcpP has served as a model for other thiotemplated, or carrier protein-dependent, pathways[15,16].

Through each step, substrates that are covalently bound to the AcpP by a 4'-phosphopantheteine arm are accessed by partner enzymes through PPIs in order to elongate the cargo or perform tailoring reactions[17]. In FAB, elongation adds two carbon units to provide a β-ketone, which is fully reduced to the saturated alkane in three steps that involve formation of β-hydroxy and *trans*-α/β unsaturated intermediates. The *trans*-α/β unsaturated intermediate can undergo reduction to the fully reduced alkane. Alternatively, the *trans*-α/β-unsaturation can be isomerized to the *cis*-α/β-unsaturation and enter back into elongation in order to immortalize the *cis* double bond[18,19]. Throughout FAB elongation, ACP-bound intermediates can also be diverted to other enzymes to source fatty acids for the biosynthesis of molecules such as lipoic acid[20], acyl homoserine lactones[21], and lipid A[22].

ACPs carry the growing acyl chains covalently attached to a 4'-phosphopantetheine cofactor via a thioester linkage. The ACP has been found to sequester the acyl cargo inside of a hydrophobic pocket that is created by the four α-helices of the ACP structure (Fig 1b). Upon interaction with a partner enzyme, the cofactor and cargo are "chain-flipped" out of the pocket and into the active site of the partner enzyme[23,24]. After the reaction, the acyl chain is flipped back to sequester within ACP, thus protecting the substrate from hydrolysis. The PPI binding events have been previously demonstrated to occur rapidly, with the ACP and partner proteins forming transient interactions[25,26]. Published studies have demonstrated that substrate changes translate into structural modifications within the ACP, particularly upon helix III[27]. It has been thought that these structural changes effect partner enzyme interactions[28]. One of the most remarkable transformations within unsaturated FAB is the production of a C10:1 fatty acid. The first control step in unsaturation is catalyzed by 3-hydroxy-decanoyl dehydratase, FabA in *E. coli.* FabA's role in chain length-specific unsaturation has been established through both *in vitro* and metabolic flux experiments.[29,30] Specifically, FabA produces the C10:1 *cis* unsaturation by scavenging 3-hydroxy-decanoyl AcpP from the saturated FAB cycle (Fig 1a). For this, FabA performs a dual purpose. First, dehydrating to form *trans*-2-decenoate, then FabA performs an isomerization to *cis*-3-decenoyl AcpP. This substrate then continues to elongation by FabB. This process is the only entry to unsaturated fatty acids for *E. coli*. Previous studies have demonstrated that substrate analogs bearing a reactive chemical crosslinker loaded onto the AcpP will not react with partners when the probe does not mimic the natural ten-carbon substrate[31], suggesting a point of

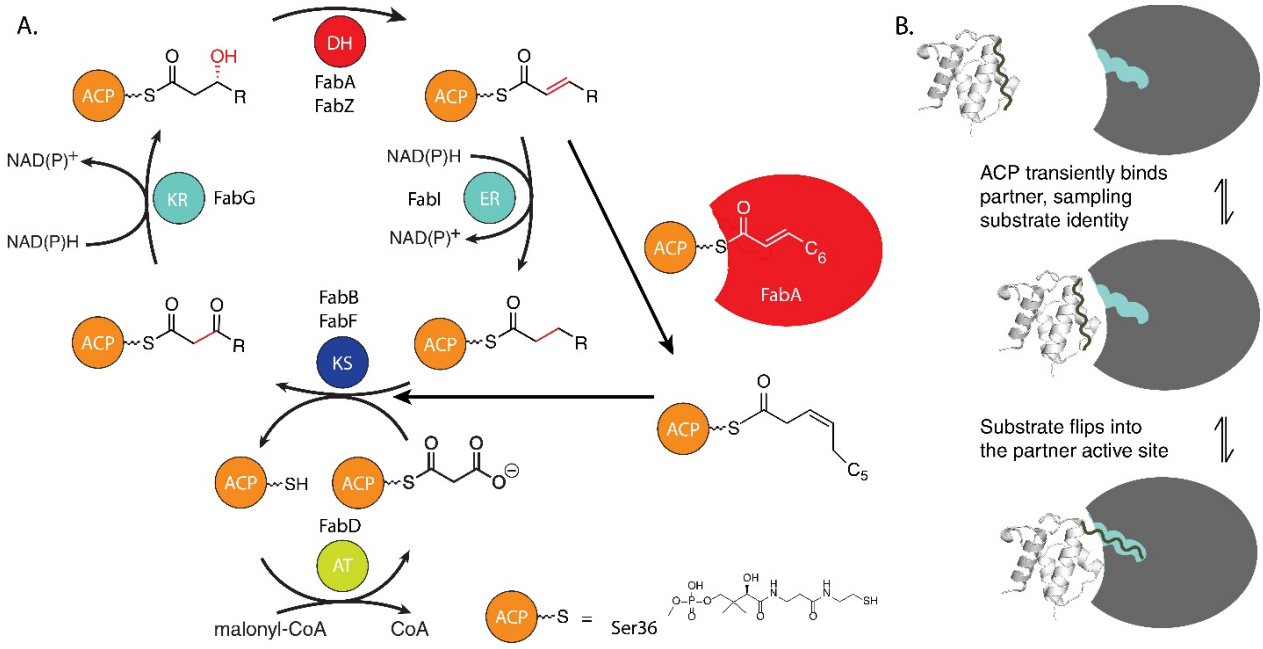



**Figure 2.1 A) The type II fatty acid biosynthesis cycle in *E. coli*. KS: ketosynthase, KR: Ketoreductase, DH: Dehydratase, ER: enoylreductase, TE: Thioesterase, AT: acyltransferase. The secondary role of FabA as an isomerase is specific to C10 acyl chains. B) AcpP as a 3 step process, wherein transient binding and unbinding can be performed separate from the chain flipping of substrates into a partner protein.**

substrate control preceding full chain flipping into the partner active site (Fig 1B). The first crosslinked crystal structure of *E. coli* AcpP was reported of the AcpP=FabA complex (PDB: 4KEH) using crosslinking probes attached to AcpP in order to trap the active site histidine residue[26]. The crosslinked structure established the catalytic conformation of the AcpP=FabA complex, and associated NMR titration experiments demonstrated the ability to probe dynamic PPIs to understand pre-catalytic interactions.

FabA has a unique and important secondary role as an isomerase, isomerizing *trans*-2-decenoyl-AcpP into *cis*-3-decenoyl-AcpP[29,30]. We recently elucidated the mechanism and specificity of isomerization through comparison of the crosslinked structure with molecular dynamics (MD) analysis[32]. This reaction has been demonstrated to be highly selective, and we have reported that chemically reactive probes demonstrate specificity for C10 acyl chain lengths[31] We observed significant PPI-based substrate control by FabA using crosslinking probes that mimicked C6, C8, and C10- chain lengths tethered to AcpP[33]. The corresponding magnitude and specificity implicit in this interaction indicated that selectivity surpassed active site recognition and implicated a substrate-controlled PPI. This inspired us to study the phenomenon in more detail, beginning with solution-state NMR.

Here we report the evaluation of FabA chain length specificity through an analysis of AcpP structure bearing three acyl chain lengths: six, eight, and ten carbons. By uniting NMR titration analysis, MD minimized acyl-AcpP structures, and high-resolution protein docking, we evaluate the role of chain length to modify AcpP structure and the ability of those structural modifications to regulate FabA activity. These data represent the first combination of structural and quantitative techniques to analyze chain length regulation in FAB.

## 2.2 Results

### FabA titration with hexanoyl, octanoyl, and decanoyl AcpP

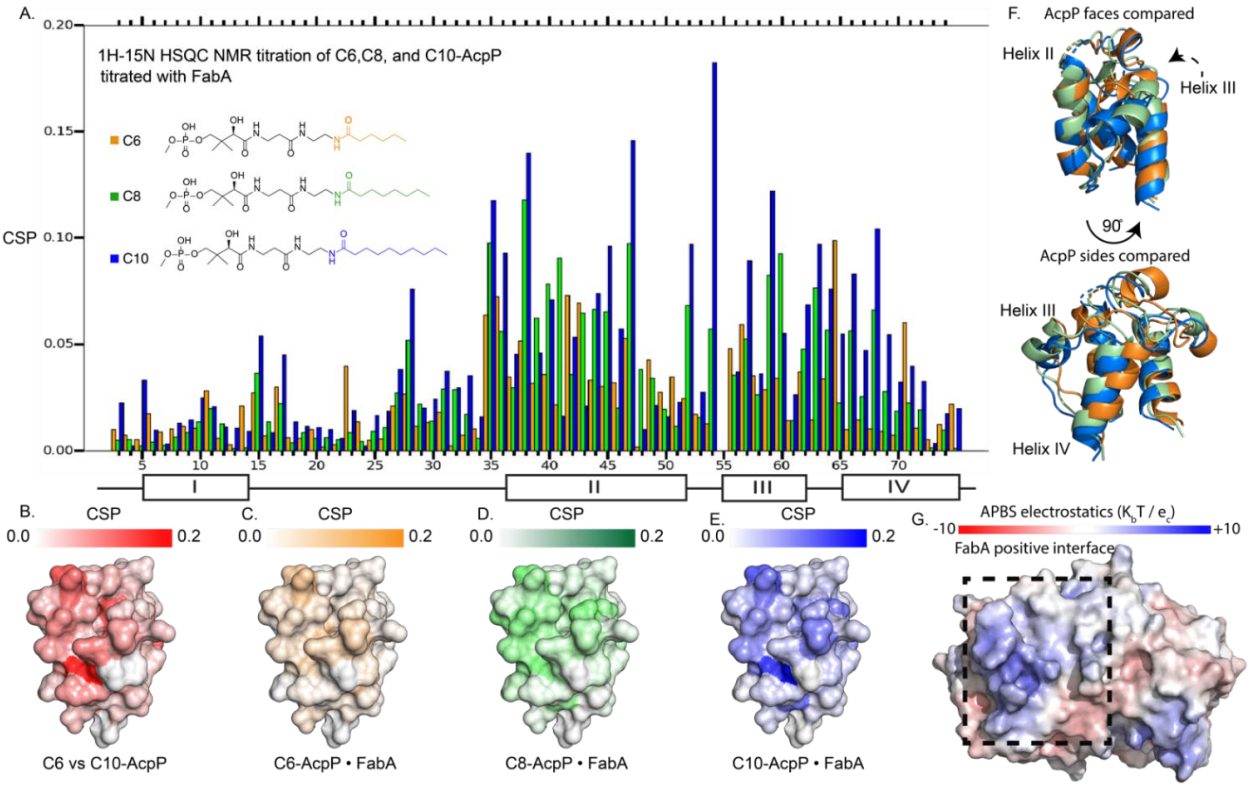



**Figure 2.2 A) The chemical shift perturbations (CSPs) of C6, C8, and C10 after titration to saturation with FabA. Each titration was performed to at least 1.5 molar equivalents of FabA. Individual perturbation graphs and spectra are presented in SI figures 1-4. B) The difference in residues between the C6 and C10-AcpP. C) The effect of FabA titration on the C6-AcpP D) The effect of FabA titration the C8-AcpP E) The effect of FabA titration C10-AcpP F) The starting MD derived structures of AcpP overlaid. Helix III exhibits the largest structural change between the chain lengths. G) The APBS electrostatics of the highly positive binding patch of FabA and negative face of AcpP. The left portion of the highlighted region is responsible for the majority of the binding interactions, binding helix III of AcpP.**

$^{1}$H-$^{15}$N HSQC NMR has found a role in the study of in solution interactions of AcpP with proteins [25,26], and we employed it here to observe the subtle differences in binding of the substrate bearing AcpPs with FabA. With the movement of peaks representing the average of the populations of states present in solution[34]. Uniformly labeled $^{15}$N-AcpP was prepared and loaded with the three different probes with C6, C8, and C10 attached through aminopantetheine linkage[33]. These were titrated with increasing concentrations of unlabeled FabA to observe the residues which experienced peak migration upon $^{1}$H-$^{15}$N HSQC NMR (Fig 2A). Saturated acyl chains were chosen as the tethered analogs to avoid reaction with FabA during experiments and remain consistent with prior crosslinking studies. However, future studies with catalytically inactive enzymes and the native substrate will be necessary to further study specificity.

The effect of the different chain lengths alone on the chemical shifts of AcpP led to the largest chemical shift perturbations (CSPs) at residues F28, D35, S36, E47, I54, D56, A59, I62, T64, and Q66 (Fig 2B, S1&2). These residues all reside along the hydrophobic pocket of AcpP, with many occurring at helix III or along the preceding loop. Across the three titrations we observed CSPs in similar regions of the AcpP (Fig 2C-E). The bulk of CSPs occurred in the helix II and III region and the helix II/III loop. However, the CSPs extended through the top of helix IV. Perturbation magnitude increased with each successive two-carbon increase in substrate size. From a mean CSP of 0.024 for C6, to 0.032 for C8, and 0.043 for C10 (Table 1).

**Table 2.1 The average CSPs, TITAN analysis, and docking energetics of the AcpP • FabA interaction.**

| **Chain Length** | **Average CSP** | **$K_d$ (μM)** | **$k_{off}$ ($s^{-1}$)** | **Electrostatics (kcal/mol)** | **Van der Waal's (kcal/mol)** | **Total energy (kcal/mol)** |
|---|---|---|---|---|---|---|
| **C6** | 0.0241 | 34.8±5.9 | 311±64 | -6.936 | -11.18 | -2.512 |
| **C8** | 0.0319 | – | – | -5.149 | -16.11 | -11.79 |
| **C10** | 0.0434 | 8.6±3.0 | 4535±647 | -8.065 | -28.13 | -22.36 |

The most perturbed residues of the C10-AcpP•FabA titration were located at D35, S36, D38, A45, E47, T52, I54, E57, A59, T63, Q66, and A68. The largest chemical shift differences between the C6 and C10-AcpPs alone correspond closely to these same most-perturbed residues in the titration with FabA. The most perturbed residues are neighboring or close to one another, except for F28, which remained one of the most perturbed residues. To compare the C6 and C10 titrations, the CSP values were normalized to the largest CSP within the data set (Fig S3). It was noted that CSPs unique to the C10-AcpP•FabA titration were I54, A59, T63, Q66, and A68, while the C6-AcpP•FabA titration had unique CSPs at residues L42, V43, V65, and Y71. Despite largely shared surface interactions, there are unique internal effects of FabA interacting with different cargo bearing AcpPs.

**Thermodynamic and kinetic parameters vary between AcpP substrates**

TITAN line shape analysis was utilized to generate quantitative data from the NMR titrations performed[35]. Characterization of the titrations of C6 and C10-AcpP with FabA found that the C6-AcpP bound with 34.8 ± 5.9 μM affinity and 311 ± 64 $s^{-1}$ off rate (Table 1, Fig S4). The C10-AcpP bound FabA with a 8.6 ± 3.0 μM Kd and 4535 ± 647 $s^{-1}$ off rate. This demonstrates a significant effect upon increasing the chain length. It is compelling that an addition of four carbons in cargo size could affect such a significant change to the binding affinity. A structural rationale for this difference was sought further through high resolution docking.

**AcpP•FabA 3D model**

To evaluate these observations structurally, high-resolution docking was performed in the ICM fast fourier transform docking protocol using NMR titration data to guide the docking algorithm[36–39]. To ensure that the most relevant acyl-AcpP structure is used for docking, MD derived structures of AcpP with tethered acyl chains of differing chain lengths were docked with explicit acyl cargo (Fig 2F). Next, the partner FabA enzyme was prepared through solvation of the reported crystal structure (PDB: 1MKB)[40], optimizing the hydrogen bonding and angles to the water molecules. The interfaces of the AcpP are highly acidic, and those of FabA are basic (Fig 2G), making accurate preparation of structures for hydrogen bonding important to accurate modeling. FabA was docked to the three different MD derived structures using identical methodologies except for varying the beginning AcpP inputs. The structures were sorted based on their RMSD to the post-catalytic crosslinked crystal structure 4KEH, and the most thermodynamically favorable

model within 5Å of the crosslinked structure was chosen for each chain length (Fig S6). This cutoff was chosen to accommodate the differences between a crosslinked complex and docked complex, both due to the unbound FabA and the un-chain flipped and precatalytic AcpP. Based on this analysis, we have noted that the energetics of the bound complexes agrees with the TITAN analysis and CSPs (Table 1). The AcpPs from the MD study were observed to vary most heavily and distinctly in the helix III region of AcpP (Fig 3a,b), matching the same regions identified by NMR titration studies above. Combined with the observation that the FabA interface is most heavily electropositive in the region which binds helix III, we identified an important role in recognition at helix III of AcpP.

Next it was examined whether the AcpPs would demonstrate different ability to bind FabA based on the docked thermodynamics. The energies of the three interactions were well in line with the known substrate preferences (Table 1), with C6 binding with an overall energy of -2.5 kcal/mol, C8 with an overall energy of -11.8 kcal/mol, and C10 with an overall energy of -22.4 kcal/mol[38]. These energies are not definitive alone, but they are useful for examining how well the AcpP conformations complement the surface of FabA. To appreciate exactly how these minor changes translate to differences in surface binding, we next examined the interface in detail, focusing on helix III.

**Identifying structural features which facilitate specificity**

To examine exactly which AcpP residues were most important to the interaction, the docked models were examined relative to one another (Fig 4, S4&7). There were several regions of the protein which were identified as important to the interactions, with the most significant occurring on helix III.

In the C10 docked structure, there appeared to be a coordinated network of interactions with three acidic residues, D56, E57, and E60, within a distance to form interactions with R136', the R137' backbone, and the R137' side chain, respectively. (Residues on FabA will be noted by a " ' " throughout the text.) These docked models likely represent a bound but non-chain-flipped encounter complex, representing an initial binding interaction. The C8 helix III structure demonstrates poorer binding, with helix III out of orientation and only one coordinated residue at E57 nearby R136'. E60 appears to be oriented such that there is no space for R137' to rotate and bind. C6 binds similarly, but with D56 successfully coordinating R136', reflecting how the orientation of helix III determines the creation of three critical interactions in C10. It appears that shorter chain lengths cannot form all three of these encounter complex interactions. Binding

would require structural rearrangement and hinder rapid binding. Two additional identified interactions lie on residues E41 and E47. E41 interacts with K161', and only C10 and C6 docked models, but not C8, are within approximate range (within ~3Å) to interact. E47 appears to be important in stabilizing the bottom

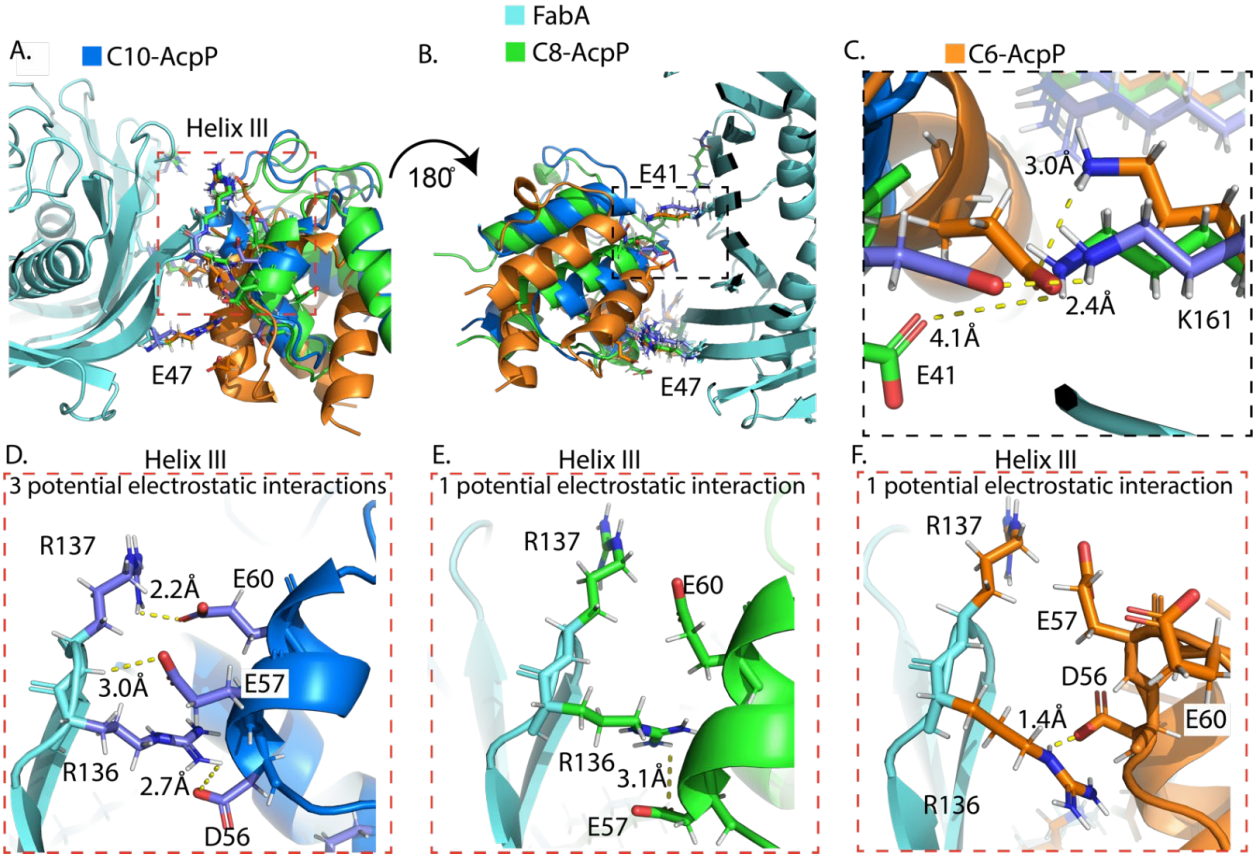



**Figure 2.3 A) The left face of the AcpP • FabA interaction. The helix III interactions are highlighted and focused on in panels D, E, and F. B) The right face of the AcpP • FabA interaction. E41 is highlighted and focused on in panel C. C) The interactions of E41 with K161' in C6, C8, and C10-AcpP. D) The interactions of helix III of C10-AcpP with FabA. Displaying the geometric complementarity of the C10-AcpP for the FabA residues E) The interactions of helix III of C8-AcpP with FabA. Displaying the non-complementarity of the binding region for the C8-AcpP helix III. F) The interactions of helix III of C6-AcpP with FabA. C6-AcpP has a helix III angled up and away from the region, such that only D56 at the base of helix III can form any interaction.**

of helix II, though it is out of interacting distance in the docked pose at 6Å, and it may interact upon full binding. Next, to complement these two studies, the CSP values were projected onto the docked AcpP structures (Fig 5). In addition to noting several strong internal perturbations, most significantly I54, two interactions were noted at D35 and D38. These residue bridges have been identified previously as participating in the stabilizing of the chain-flipped acyl chain[26]. However, it was noted that D35 was also within a proper hydrogen bonding distance to the backbone of A170'. These interactions may also anchor helix II, along with the important salt bridges identified at E41 and E47.

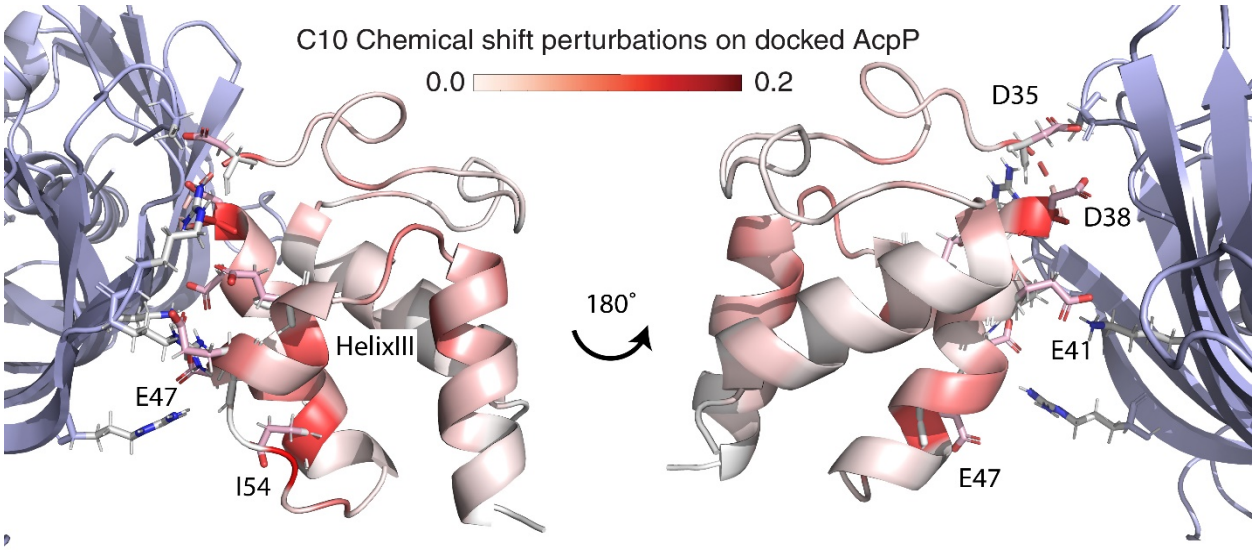



**Figure 2.4 The poses of C10-AcpP mapped onto the AcpP-FabA model. These residues are in agreement and further highlight the importance of I54, which most likely pushes the cargo out during chain flipping. Also, the importance of D35 and D38 in binding the poorly resolved loop region is shown in the CSPs, though not possible in the model.**

This interaction would have been impossible to identify in the crosslinked structure, where the loop with A170 was not resolved in the crystal structure. I54 has been identified as involved in chain flipping by examining its distance from interactions and positioning of the side chain directly into the acyl pocket. We have additionally examined the docked models of C6 and C8-AcpP in the Supplemental Information.

Taken together, these data create a compelling picture of FabA's chain length selectivity as determined by the PPIs with C6, C8, and C10 acyl chain sequestered AcpP. The CSPs reflect different surface interactions from the crosslinked crystal structure, while the magnitude of the CSPs with increasing chain length is in agreement with the TITAN analysis. This demonstrates a mechanism of substrate-dependent selectivity and regulation based upon the substrate-induced structure of AcpP. Here, the unique positioning of helix III is dictated by the respective influence of sequestered chain lengths. These structural differences, though subtle, are significant enough that the most energetically stable docking poses appear to be occluded in shorter chain lengths. This conclusion is further supported by the corroborating thermodynamic and CSP data.

**Expanding analysis of FabA docked states**

To explore the full structural space of FabA • AcpP interactions, docking was performed to observe secondary binding modes. Using the C6, C8, and C10 models examined above as reference structures, a

second docking calculation was performed with all docked poses examined based on their RMSD to the minimized model. The docked states were then graphed by RMSD vs energy (Fig 5) to observe other potential binding modes. In the C10 binding, a cluster of low RMSD conformations (~5Å) were observed (Fig 5A), along with a second cluster of stable secondary states at higher RMSD (~15Å) (Fig 5E). A second, lower energy state, was also observed at ~13Å RMSD from the preliminary model (Fig 5F). This model has FabA adopting a similar docked site (Fig S5), but with the AcpP directed more sharply into the FabA face. In addition, multiple low energy, higher RMSD states sampled features of the FabA interface. The C8 docked data displayed that the model closest aligned to the crosslinked-like structure was higher energy than a set of ~7Å RMSD models (Fig 5B,S5). The lowest energy of these adopts a similar structure to the crosslinked-like model examined. Lastly, C6 showed few low energy states, and low RMSD to the crosslinked-like state. The lowest energy model occurred at ~17Å RMSD from the docked model examined above (Fig 5C,S5). However, this state, as well as multiple similar energy states, docked "upside-down," likely with no ability to perform a functional interaction that could result in chain-flipping.

### 2.3 Discussion

Evaluation of this data demonstrates several key concepts. First, the "active" cross-link-like conformation, which is stable in the C10-AcpP docked model, is less energetically favorable for shorter chain lengths. The C8-AcpP docked model was observed to have a small set of favorable structures which were relatively similar to the active model; and C6's were significantly different. This could explain the observation that FabA crosslinking experiments with C8 analogs were more favorable than with C6 variants[31], given that C8 can form more stable PPI conformations. Finally, it has been long known that the AcpP • partner enzyme interface is a dynamic interaction[26,32,41]. We predict that the various energetically favorable bound states can constitute encounter complexes[42,43], or preliminary interactions that bind the AcpP transiently and allow rearrangement of the AcpP into the active state. This PPI mechanism would eliminate the need for many transient binding events, allowing the AcpP to associate with the interface in order to form a catalytic interaction from more than just one perfectly aligned binding event with a partner enzyme.

These experiments demonstrate the role of the AcpP sequestered substrate in the PPI-controlled catalysis of FabA. Here, the identity of the substrate is pivotal to the positioning of helix III to attain favorable

FabA binding. By merging our knowledge of the structural effects of chain length with observations of substrate specificity, we show how FabA, and presumably all other AcpP partner enzymes, maintain a control step over possible substrates tethered to AcpP before the chain flipping process commits the substrate into the active site. This PPI-based specificity mechanism has been hypothesized as an explanation for substrate specificity, although it had not been demonstrated for FabA[44]. While prior demonstration of selective crosslinking of FabA by a substrate-mimicking crosslinker favored C10-AcpP over C6 or C8[31,45], the phenomenon of substrate controlled PPI has not been specifically identified. This study presents a structural model for explaining this specificity, with additional work necessary to fully understand the mechanism. However, this can serve as a foundation to build the understanding of this phenomenon. This first step required leveraging dynamic studies by NMR to examine the solution interactions of acyl-AcpP with FabA, to inform docking simulations, and to provide confirmatory thermodynamics. Further, we observe the formation of several stable conformations of the acyl-AcpP • FabA interaction that are near or in the path of the crosslinked conformation. These states around the catalytic structure of AcpP may present an expanded paradigm for AcpP • partner interactions that capture encounter complexes that form prior to the chain-flipping event. Stable secondary interfaces could help lead acyl-AcpP, which may initially bind non-ideally, to the active  bound conformation required for chain-flipping. This explains the plasticity of the AcpP interface, with few studies identifying single mutations sufficient to abolish activity[46,47]. These findings also explain the remarkable efficiency of FAB, with the ability of acyl-AcpP to be “funneled” from improperly coordinated binding events into the required conformation.

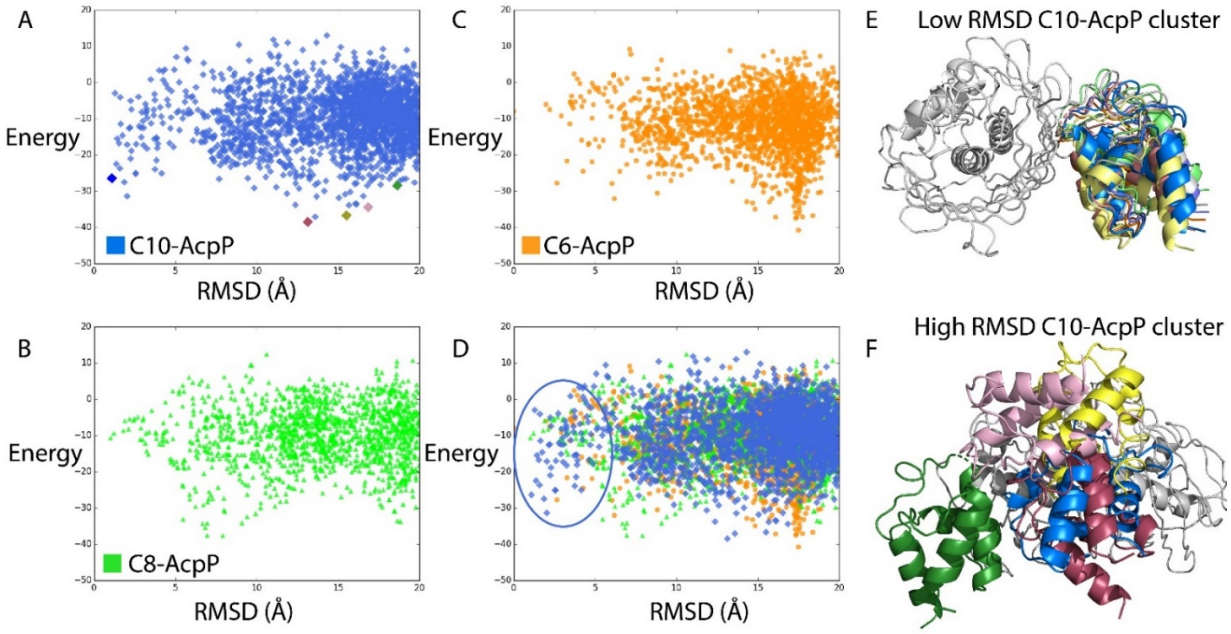



**Figure 2.5 A) The ensemble of conformations around the C10-AcpP • FabA interface. A narrowed RMSD range is presented in figure S7. Specific conformations at high RMSD are color coded blue, maroon, green, yellow, and pink panel F. B) The ensemble of conformations around the C8-AcpP • FabA interface. The conformations are presented in further detail in figure S8. C) The ensemble of conformations around the C6-AcpP • FabA interface. The conformations are presented in further detail in figure S8. D) An overlay of the conformations of the C6, C8, and C10-AcpP• FabA interface. The low RMSD displayed in panel E are circled in blue. E) The ensemble of low RMSD states of the C10-AcpP • FabA interface. F) A sample of low energy states of the C10-AcpP • FabA interface. Colors are matched to the model in panel A**

Previous microbiological and biochemical studies have demonstrated the role of FabA in scavenging acyl-AcpP pools to perform unsaturation. This work now presents a model for a first control step, maintaining the essential specificity by PPIs evolved to selectively recognize C10-AcpP structure. Unsurprisingly, these findings required a highly interdisciplinary methodology that relied on a fusion of experimental and computational analyses. This model for rapidly sampled chain length specificity must certainly extend beyond FabA, and likely accounts for acyl chain selectivity across some of the 25 known AcpP-dependent enzymes in *E. coli*, explaining both the speed and precision evolved into FAB. We suspect that these phenomena are involved in all thiotemplated pathways that sequester intermediates within the carrier protein, providing a powerful and fundamental control mechanism.

**Chapter 2**, is in full a reprint of the manuscript: Thomas G. Bartholow, Terra Sztain, Megan A Young, D. John Lee, Tony D. Davis, Ruben Abagyan, Michael D. Burkart, “Control of unsaturation in de novo fatty acid biosynthesis by FabA.” *Biochemistry*, (2021). The dissertation author is the primary author of the manuscript.

## 2.4 REFERENCES


(1) Crosby, J.; Crump, M. P. The Structural Role of the Carrier Protein – Active Controller or Passive Carrier. *Nat. Prod. Rep.* **2012**, *29* (10), 1111–1137. https://doi.org/10.1039/C2NP20062G.

(2) Lai, J. R.; Koglin, A.; Walsh, C. T. Carrier Protein Structure and Recognition in Polyketide and Nonribosomal Peptide Biosynthesis. *Biochemistry* **2006**, *45* (50), 14869–14879. https://doi.org/10.1021/bi061979p.

(3) Tan, Z.; Black, W.; Yoon, J. M.; Shanks, J. V.; Jarboe, L. R. Improving Escherichia Coli Membrane Integrity and Fatty Acid Production by Expression Tuning of FadL and OmpF. *Microb. Cell Factories* **2017**, *16* (1), 38. https://doi.org/10.1186/s12934-017-0650-8.

(4) Shokri, A.; Larsson, G. Characterisation of the Escherichia Coli Membrane Structure and Function during Fedbatch Cultivation. *Microb. Cell Factories* **2004**, *3*, 9. https://doi.org/10.1186/1475-2859-3-9.

(5) Rock, C. O. Opening a New Path to Lipoic Acid. *J. Bacteriol.* **2009**, *191* (22), 6782–6784. https://doi.org/10.1128/JB.01151-09.

(6) Gully, D.; Moinier, D.; Loiseau, L.; Bouveret, E. New Partners of Acyl Carrier Protein Detected in Escherichia Coli by Tandem Affinity Purification. *FEBS Lett.* **2003**, *548* (1–3), 90–96. https://doi.org/10.1016/s0014-5793(03)00746-4.

(7) Lay, N. R. D.; Cronan, J. E. In Vivo Functional Analyses of the Type II Acyl Carrier Proteins of Fatty Acid Biosynthesis. *J. Biol. Chem.* **2007**, *282* (28), 20319–20328. https://doi.org/10.1074/jbc.M703789200.

(8) Gully, D.; Bouveret, E. A Protein Network for Phospholipid Synthesis Uncovered by a Variant of the Tandem Affinity Purification Method in Escherichia Coli. *PROTEOMICS* **2006**, *6* (1), 282–293. https://doi.org/10.1002/pmic.200500115.

(9) Cronan, J. E.; Thomas, J. Bacterial Fatty Acid Synthesis and Its Relationships with Polyketide Synthetic Pathways. *Methods Enzymol.* **2009**, *459*, 395–433. https://doi.org/10.1016/S0076-6879(09)04617-5.

(10) Magnuson, K.; Jackowski, S.; Rock, C. O.; Cronan, J. E. Regulation of Fatty Acid Biosynthesis in Escherichia Coli. *Microbiol. Mol. Biol. Rev.* **1993**, *57* (3), 522–542.

(11) Sohlenkamp, C.; Geiger, O. Bacterial Membrane Lipids: Diversity in Structures and Pathways. *FEMS Microbiol. Rev.* **2016**, *40* (1), 133–159. https://doi.org/10.1093/femsre/fuv008.

(12) Rock, C. O.; Jackowski, S. Incorporation and Turnover of Fatty Acids in Escherichia Coli Membrane Phospholipids. In *Advances in Lipobiology*; Gross, R. W., Ed.; JAI, 1996; Vol. 1, pp 39–59. https://doi.org/10.1016/S1874-5245(96)80004-8.

(13) Chan, D. I.; Vogel, H. J. Current Understanding of Fatty Acid Biosynthesis and the Acyl Carrier Protein. *Biochem. J.* **2010**, *430* (1), 1–19. https://doi.org/10.1042/BJ20100462.

(14) Finzel, K.; Lee, D. J.; Burkart, M. D. Using Modern Tools to Probe the Structure-Function Relationship of Fatty Acid Synthases. *Chembiochem Eur. J. Chem. Biol.* **2015**, *16* (4), 528–547. https://doi.org/10.1002/cbic.201402578.

(15) Byers, D. M.; Gong, H. Acyl Carrier Protein: Structure–Function Relationships in a Conserved Multifunctional Protein Family. *Biochem. Cell Biol.* **2007**, *85* (6), 649–662. https://doi.org/10.1139/O07-109.

(16) Majerus, P. W.; Alberts, A. W.; Vagelos, P. R. Acyl Carrier Protein, Iv. the Identification of 4′-Phosphopantetheine as the Prosthetic Group of the Acyl Carrier Protein. *Proc. Natl. Acad. Sci.* **1965**, *53* (2), 410–417. https://doi.org/10.1073/pnas.53.2.410.

(17) Cronan, J. E. The Chain-Flipping Mechanism of ACP (Acyl Carrier Protein)-Dependent Enzymes Appears Universal. *Biochem. J.* **2014**, *460* (2), 157–163. https://doi.org/10.1042/BJ20140239.

(18) Feng, Y.; Cronan, J. E. Escherichia Coli Unsaturated Fatty Acid Synthesis. *J. Biol. Chem.* **2009**, *284* (43), 29526–29535. https://doi.org/10.1074/jbc.M109.023440.

(19) Cao, Y.; Yang, J.; Xian, M.; Xu, X.; Liu, W. Increasing Unsaturated Fatty Acid Contents in Escherichia Coli by Coexpression of Three Different Genes. *Appl. Microbiol. Biotechnol.* **2010**, *87* (1), 271–280. https://doi.org/10.1007/s00253-009-2377-x.

(20) Morris, T. W.; Reed, K. E.; Cronan, J. E. Lipoic Acid Metabolism in Escherichia Coli: The LplA and LipB Genes Define Redundant Pathways for Ligation of Lipoyl Groups to Apoprotein. *J. Bacteriol.* **1995**, *177* (1), 1–10. https://doi.org/10.1128/jb.177.1.1-10.1995.

(21) Parsek, M. R.; Val, D. L.; Hanzelka, B. L.; Cronan, J. E.; Greenberg, E. P. Acyl Homoserine-Lactone Quorum-Sensing Signal Generation. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (8), 4360–4365.

(22) Homma, J. Y.; Matsuura, M.; Kanegasaki, S.; Kawakubo, Y.; Kojima, Y.; Shibukawa, N.; Kumazawa, Y.; Yamamoto, A.; Tanamoto, K.; Yasuda, T. Structural Requirements of Lipid A Responsible for the Functions: A Study with Chemically Synthesized Lipid A and Its Analogues. *J. Biochem. (Tokyo)* **1985**, *98* (2), 395–406. https://doi.org/10.1093/oxfordjournals.jbchem.a135294.

(23) 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. https://doi.org/10.1002/anie.201408576.

(24) Cronan, J. E. The Chain-Flipping Mechanism of ACP (Acyl Carrier Protein)-Dependent Enzymes Appears Universal. *Biochem. J.* **2014**, *460* (2), 157–163. https://doi.org/10.1042/BJ20140239.

(25) 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. https://doi.org/10.1038/s41589-019-0301-y.

(26) 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. https://doi.org/10.1038/nature12810.

(27) 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. *Chem. Biol.* **2010**, *17* (7), 776–785. https://doi.org/10.1016/j.chembiol.2010.05.024.

(28) 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.* **2019**, *58* (32), 10888–10892. https://doi.org/10.1002/anie.201903815.

(29) Xiao, X.; Yu, X.; Khosla, C. Metabolic Flux Between Unsaturated and Saturated Fatty Acids Is Controlled by the FabA:FabB Ratio in the Fully Reconstituted Fatty Acid Biosynthetic Pathway of E. Coli. *Biochemistry* **2013**, *52* (46). https://doi.org/10.1021/bi401116n.

(30) Rock, C. O.; Tsay, J. T.; Heath, R.; Jackowski, S. Increased Unsaturated Fatty Acid Production Associated with a Suppressor of the FabA6(Ts) Mutation in Escherichia Coli. *J. Bacteriol.* **1996**, *178* (18), 5382–5387. https://doi.org/10.1128/jb.178.18.5382-5387.1996.

(31) 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. https://doi.org/10.1016/j.chembiol.2015.09.009.

(32) 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. https://doi.org/10.1073/pnas.1818686116.

(33) S. Worthington, A.; D. Burkart, M. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44–46. https://doi.org/10.1039/B512735A.

(34) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001.

(35) Waudby, C. A.; Ramos, A.; Cabrita, L. D.; Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Sci. Rep.* **2016**, *6*, 24826. https://doi.org/10.1038/srep24826.

(36) Abagyan, R.; Totrov, M. Biased Probability Monte Carlo Conformational Searches and Electrostatic Calculations for Peptides and Proteins. *J. Mol. Biol.* **1994**, *235* (3), 983–1002. https://doi.org/10.1006/jmbi.1994.1052.

(37) ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation - Abagyan - 1994 - Journal of Computational Chemistry - Wiley Online Library https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540150503 (accessed Aug 31, 2020).

(38) Neves, M. A. C.; Totrov, M.; Abagyan, R. Docking and Scoring with ICM: The Benchmarking Results and Strategies for Improvement. *J. Comput. Aided Mol. Des.* **2012**, *26* (6), 675–686. https://doi.org/10.1007/s10822-012-9547-0.

(39) Bartholow, T. G.; Sztain, T.; Lee, D. J.; Abagyan, R.; Burkart, M. D. Elucidation of Transient Protein-Protein Interactions within Carrier Protein-Dependent Biosynthesis. *Commun. Biol. Accept.*

(40) 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. https://doi.org/10.1016/S0969-2126(96)00030-5.

(41) Andrec, M.; Hill, R. B.; Prestegard, J. H. Amide Exchange Rates in Escherichia Coli Acyl Carrier Protein: Correlation with Protein Structure and Dynamics. *Protein Sci.* **1995**, *4* (5), 983–993. https://doi.org/10.1002/pro.5560040518.

(42) Kozakov, D.; Li, K.; Hall, D. R.; Beglov, D.; Zheng, J.; Vakili, P.; Schueler-Furman, O.; Paschalidis, I. C.; Clore, G. M.; Vajda, S. Encounter Complexes and Dimensionality Reduction in Protein–Protein Association. *eLife* **2014**, *3*, e01370. https://doi.org/10.7554/eLife.01370.

(43) Shen, Y.; Paschalidis, I. C.; Vakili, P.; Vajda, S. Protein Docking by the Underestimation of Free Energy Funnels in the Space of Encounter Complexes. *PLOS Comput. Biol.* **2008**, *4* (10), e1000191. https://doi.org/10.1371/journal.pcbi.1000191.

(44) Colizzi, F.; Masetti, M.; Recanatini, M.; Cavalli, A. Atomic-Level Characterization of the Chain-Flipping Mechanism in Fatty-Acids Biosynthesis. *J. Phys. Chem. Lett.* **2016**, *7* (15), 2899–2904. https://doi.org/10.1021/acs.jpclett.6b01230.

(45) Miyanaga, A.; Ouchi, R.; Ishikawa, F.; Goto, E.; Tanabe, G.; Kudo, F.; Eguchi, T. Structural Basis of Protein-Protein Interactions between a Trans-Acting Acyltransferase and Acyl Carrier Protein in Polyketide Disorazole Biosynthesis. *J. Am. Chem. Soc.* **2018**, *140* (25), 7970–7978. https://doi.org/10.1021/jacs.8b04162.

(46) 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. https://doi.org/10.1073/pnas.2009805117.

(47) 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. https://doi.org/10.1038/s41467-020-15455-x.

## 2.5 Experimental Methods

### Protein Purification and production Protocol

The FabA used in the study was generated through overexpression in *E. coli* BL21 (DE3), cells were grown in LB media. 1L of media was prepared with 50mg/L kanamycin, the growth was inoculated with a 5mL starter culture grown overnight. FabA was grown ~5 hours until an OD600= 0.8, then induced with 1mM IPTG and incubated 12-18 hours at 18 °C. After growth pelleting was performed on a Beckman floor centrifuge in a JLA-8.1 rotor at 800 RCF. Pelleted cells were flash frozen and stored until purification was performed prior to titration.

Labeled AcpP was grown from a pet-22b vector with a His-tag in *E. coli* BL21 (DE3) cells. In order to label the cells they were grown in $^{15}N$ supplemented M9 minimal media. 1g of $^{15}N$ $NH_4Cl$ and 8g of unlabeled glucose were added to 1L of M9 media. In order to achieve deuteration the media components were mixed in an oven dried glass graduated cylinder, followed by sterile filtration into a autoclaved and oven dried growth flask. Inoculating bacteria was carefully attenuated to the deuterated media, over the course of several growths. To begin BL21 cells were inoculated into a 25% $D_2O$/75% $H_2O$ unlabeled media, these were grown overnight at 37 °C. This growth was used to then inoculate another 50% $D_2O$/50% $H_2O$ media, which was grown overnight in the same conditions. This was then used to inoculate 75% $D_2O$/ 25% $H_2O$, which was grown and used to inoculate 90% $D_2O$ media. Finally the 90% $D_2O$ growth was used to inoculate a final starter with 100% $D_2O$ M9 media. This final 100% $D_2O$ starter was grown overnight at 37 °C and after confirming by eye that the media had become turbid with growth used to inoculate the labeled $D_2O$ M9 media. This was grown at 37 °C for ~16 hours until OD600=0.7. At this point 1mM IPTG was added for induction and the growth was left to grow for 4 hours at 37 °C. After induced growth the cells were spun down on a JLA-8.1 rotor at 800 RCF. Cells were spun for 1 hour and care was taken when harvesting cells to ensure there was no loss of material.

The $^{15}N$ ammonium chloride used in the labeled growth was purchased from Cambridge Isotopes laboratory. Deuterium oxide ($D_2O$) used in preparation of perdeuterated growth was purchased from Sigma Aldritch. All unlabeled proteins were grown on Luria broth from Teknova.

For purification cells were re-suspended in 50mM Hepes (pH 7.4), 250 mM NaCl, and 10% glycerol. The lysed cells were spun at 10,000 RCF in a Beckman floor centrifuge equipped with a JA-20 rotor. Spun protein was checked for full clarification after 1 hour, after confirming pelleting of membrane and insoluble materials the protein was taken for purification. Clarified lysate was mixed with 2mL bed volume of Bio-Rad Ni-IMAC resin and left on a rotator in a 4 °C cold room to batch bind for 30 minutes. Washing was performed with 2 40 mL washes with 50mM HEPES (pH 7.4), 250mM NaCl, and 10% glycerol, the first wash was performed with just buffer and the second with an added 15mM imidazole. After washing elution was performed with 3 5mL volumes of wash buffer with an added 250mM imidazole. Bradford reagent was used to test the purification for protein and at the end of elution to confirm no further protein was eluted. The same purification protocol was used for the AcpP, but out of caution for losing valuable labeled protein the wash volume was lowered to 30 mL and 10mM imidazole. After purification proteins were checked by 12% SDS-PAGE to confirm successful purification. Elutions were dialyzed overnight into 50mM Tris-HCl (pH 7.4), 250 mM NaCl, 10% glycerol, and 1mM DTT. For the AcpP purification the second wash was also dialyzed, but discarded once successful separation of the labeled AcpP was confirmed.

**ACP Chemoenzymatic loading**

After purification and dialysis the AcpP was made uniformly apo by reaction with *Pseudomonas aeruginosa* ACPH, with an added 5mM $MgCl_2$ and 0.5mM $MnCl_2$. Reaction was performed overnight at 37 °C on a rotator. Apofication was confirmed by conformationally sensitive Urea-PAGE. After confirmation that the AcpP was fully apo chemoenzymatic labeling was carried out. The loading was performed using 3 *E. coli* biosynthetic enzymes CoaA, CoaD, and CoaE plus the *Bacillus subtilis* SFP. The reaction contained 12.5mM MgCl2, 10mM ATP, 0.1μM CoaA, 0.1μM CoaD, 0.1μM CoaE, 0.2μM Sfp, 0.02% Triton X, 0.01 % Azide, 0.1% TCEP, and 0.1mM acyl mimic probe.

**Purification and preparation for titrations**

Samples were purified by the same means as previously published. After dialysis of the FabA or one pot chemoenzymatic loading of the AcpP the samples were collected and concentrated to 2mL on Amicon

Ultra-15 spin concentrators. 3kDa and 10kDa columns were used for the AcpP and FabA respectively. After concentration AcpP and FabA were purified by size exclusion chromatography on a Superdex 75 column, 10mM potassium phosphate pH 7.4, 0.5mM TCEP, and 0.1% azide buffer was prepared and used to purify the AcpP and FabA for the experiments. In order to assure consistency the same buffer was used for purifications and as buffer in the NMR experiments. In the case of the C6-AcpP experiment the carrier protein was taken from the FPLC and concentrated to 2.02 mg/mL and the FabA was concentrated to 24.5 mg/mL, final concentrations were 0.220 mM C6-AcpP and 1.226 mM FabA in the saturated sample and 0.220mM AcpP in the zero point sample. A high concentration of partner protein was found to be necessary to achieve full saturation in the case of C6-AcpP. In the C10-AcpP experiment the carrier protein was concentrated to a final concentration of 1.56 mg/mL and the FabA was concentrated to 33.6 mg/mL. The final concentrations were 0.0678 mM C10-AcpP and 0.0858 mM FabA in the saturated sample and 0.0678 mM C10-AcpP in the zero point sample. The ratios were chosen based on the previously published experiments with C8-AcpP, with a higher saturated molar ratio of 1.25 selected to ensure saturation.

**NMR Experiments**

All spectra collected in this experiment were collected at the UCSD Biomolecular NMR facility on their Bruker 800MHz spectrometer. Previous assignments were used for the C8-AcpP backbone HSQC assignments[1]. The C6 and C10-AcpP HSQCs were assigned based on the C8-AcpP, due to the small differences between the two spectra. Assigned peaks are available to view on the BMRB. Experiments were performed at 37 °C, titrations had a total of 5 titration points. The chemical shift perturbations were quantified using the formula below with an α value of 0.2. This was in order to keep the data consistent with previous work in FAB.

$$CSP = \sqrt{\frac{1}{2}[\,\delta_H^2 + (\alpha \cdot \delta_N^2)]}$$

To perform the titrations two samples were prepared. A saturated sample and zero-point sample, buffers were prepared identically for both samples with only the presence of partner protein different between

samples. All three sets of HSQC experiments were acquired with a 1.5 second recycle delay and 2048 data points. Between experiments samples were stored at 4°C to maintain stability, no denaturation of the labeled AcpP was seen in the spectra and no visible crashed protein was observed in any sample. Processing was performed in NMRPipe 10.9[2] and visualization was performed in NMRFAM-SPARKY 3.115[3]. After processing all figures displaying spectra were generated in Sparky, chemical shift perturbation calculations and figure generation was performed using the Matplotlib python utility[4].

**Titan analysis**

Further analysis of the titrations was performed using the TITAN lineshape analysis program. 5 titrations steps were used in each analysis. Peaks were selected by hand across the titration before performing an initial fitting of the data. Fitting parameters were first estimated at 10μM with a $k_{off}$ of 5000 $s^{-1}$, following fitting the peaks were hand checked. Peaks were examined to be sure there was no errors in the cases of peaks which migrated into one another or crowded regions of the spectra which were incorrectly fit. After this an initial jackknife error analysis was performed, this gave a rough picture of the error of the calculations. After a final hand check that no peaks were fitted incorrectly the final error analysis was performed. In each titration data set 300 steps of bootstrap error analysis were performed, this took approximately 18 hours for each data set. Calculations were performed by the same protocol as previously published on fatty acid biosynthesis, a set of matched simulated and real peaks are presented.

**Docking Method**

The FabA structure used in docking was prepared from the 1MKB crystal structure. AcpPs used for docking were taken from previous MD simulations. FabA was kept as a dimer in the simulation, in order to best model the in-solution structure. Before docking the FabA structure was prepared by solvation and minimization. Any waters present in the crystal structure were kept, and the ICM quickflood procedure was performed to generate a water box. Following solvation the FabA was minimized to correctly orient the amino acid side chains for interaction with the AcpP. Optimization was performed on FabA by first running the ICM optimizeHbonds and optimize HisProAsnGlnCys protocols. Molecular dynamics derived AcpP

structures were used for the ACPs. The acyl chain and phosphopantetheine were preserved during the calculation to best mimic the different chain lengths. All docking was performed using the ICM – Molsoft FFT protein-protein docking algorithm.

Models of the FabA•AcpP interactions were chosen based on the most stable model under 5Å RMSD form the crosslinked crystal structure. This cut off was chosen due to the differences between crosslinked complexes and the docked complex. The chosen complexes were visualized against one another when comparing the most stable conformations. A second analysis was performed by using the stable conformation seen for the C6, C8, and C10-AcpP•FabA complex as a reference. The full global and refined docking was performed with the reference complex, yielding a set of models with RMSDs from the most stable complex. This data set of poses with deviation from the most stable state was used to map the FabA surface. Graphing of the energetics was done in Matplotlib, while visualizations were performed in Pymol[5].

**Supplementary analysis**

**Analyzing the CSPs of the C6-AcpP ● FabA and C8-AcpP ● FabA complex with docked complexes**

The structure of the C6 and C8-AcpP model with FabA displays an extremely similar structure to the C10. Similarly, aside from the magnitude of the perturbation the residues which are perturbed are remarkably similar. The largest area of difference lies in the ability to bind at helix III. Analyzing the normalized CSPs for the residues which appear to be unique for strong perturbations in C6 and C8-AcpP appears to identify alternate faces which are unusually perturbed. Care must be taken when consulting the normalized values from the CSPs as the much lower magnitudes of the shorter chain lengths may lead to background CSPs appearing more significant relative to the C10-AcpP's large perturbation. However, the largest relative CSPs of C8 appear to lie along helix II and at the top of helix III. While the largest CSPs of C6-AcpP are distal to the interacting face. The C8 perturbations occur at T39, E41, L42, and K61. Generally, it can be seen in the normalized values the C8 appears to have a greater relative interaction across helix II. This would appear to confirm the hypothesis that the structure of C8-AcpP is such that the structurally unaffected

helix II can bind appropriately to FabA. However, helix III which is more perturbed with the increase in chain length is comparably lower in relative CSPs. There is a single residue at K61 which is very perturbed, but this positive residues would seem to be unable to bind the important positive patch of the partner. Perhaps the poorly binding R137, unable to interact with the side chain of E60 as seen in the C10 model is binding the backbone amide of E60, thus perturbing the adjacent backbone amide. Similarly, the side chain of E60 is within reasonable distance to bind the backbone of R137 which may help position the side chain of R137. It is more difficult to hypothesize the source of the CSPs of C6. Caution must be taken as well with the CSPs being quite low. The best analysis is that the inability of C6 to form a stable or strong interface is such that the interface residues do not experience as large of an effect. It may be significant that the interface of C6 is noticeably more different with C10 than C8. This could suggest that C8 forms a complex which is more favorable and like the C10 interface. While C6's lack of a stable single conformation means that alternate modes of binding are more likely to be seen and these secondary interfaces cause the less similar CSP profile. However, the low CSPs make this a tenuous theory.

## 2.6 Supplemental Information

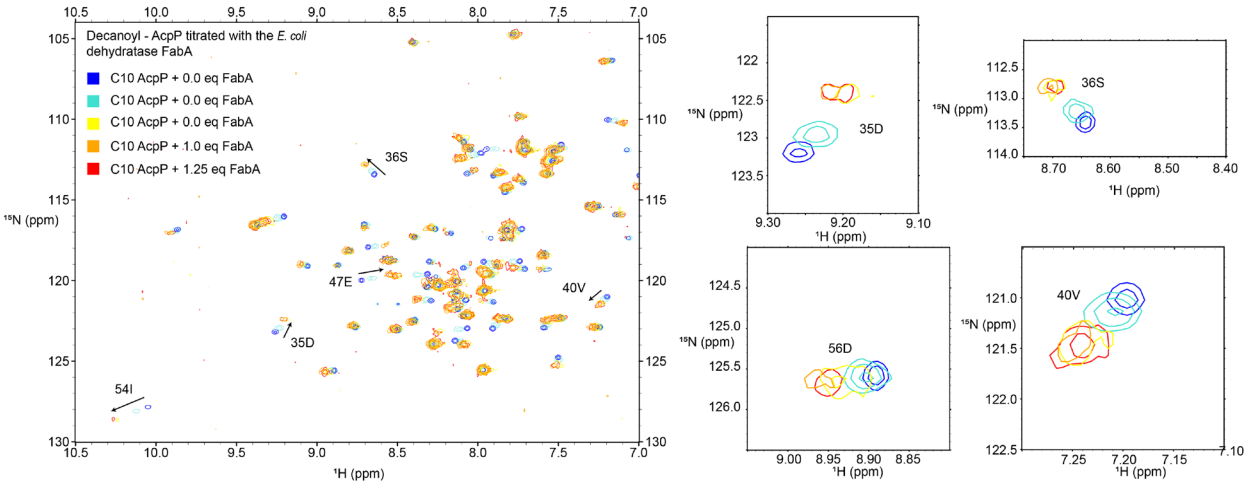



Figure S1 Titration of decanoyl -AcpP with the *E. coli* dehydratase. 5 1H-15N HSQC spectra were overlayed of the decanoyl AcpP interacting with increasing molar ratios of the dehydratase, FabA. The titration was observed to occur in fast exchange, with the bound and unbound state interchanging between bound and unbound rapidly and resolving as a single peak on the spectra.

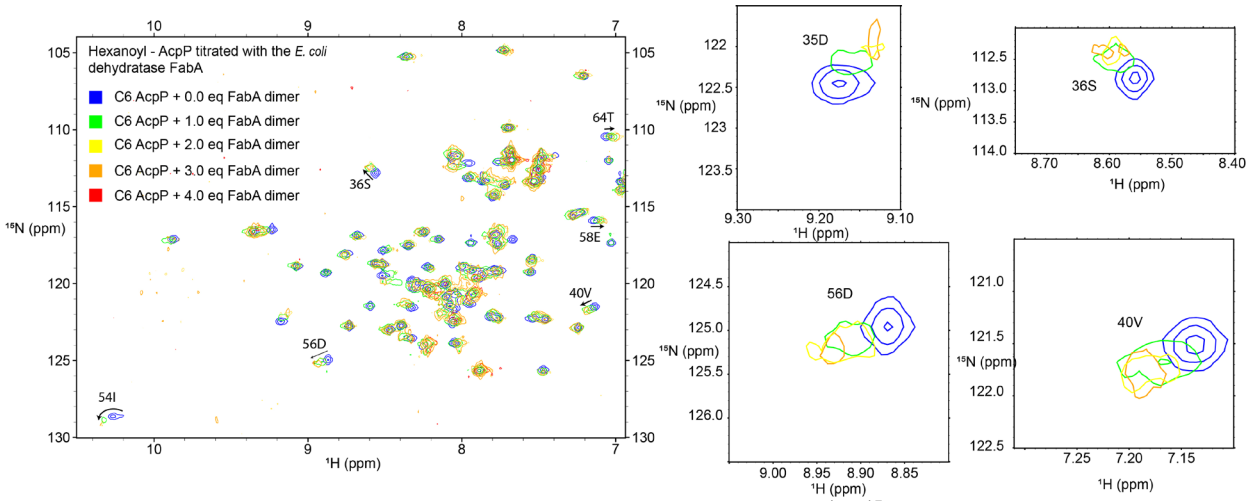



Figure S2 Titration of hexanoyl - AcpP with the *E. coli* dehydratase. 5 $^{1}H$-$^{15}N$ HSQC spectra were overlayed of the hexanoyl AcpP interacting with increasing molar ratios of the dehydratase, FabA. Higher molar ratios of FabA were required to insure full saturation of the AcpP. This has been attributed to the lower binding ability of the FabA with C6 AcpP. We have further characterized this observation with binding constants derived through TITAN.

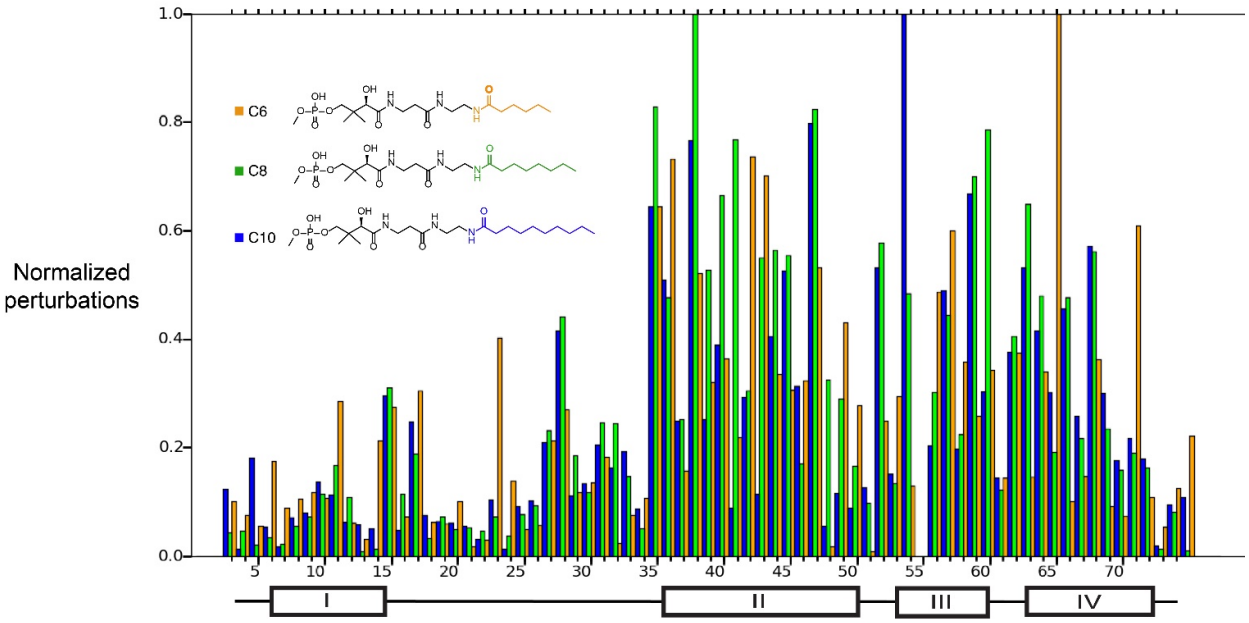



Figure S3. Normalized chemical shift perturbations of three AcpPs interacting with FabA. The perturbations are normalized within their own data set, setting the largest CSP at 1.0.

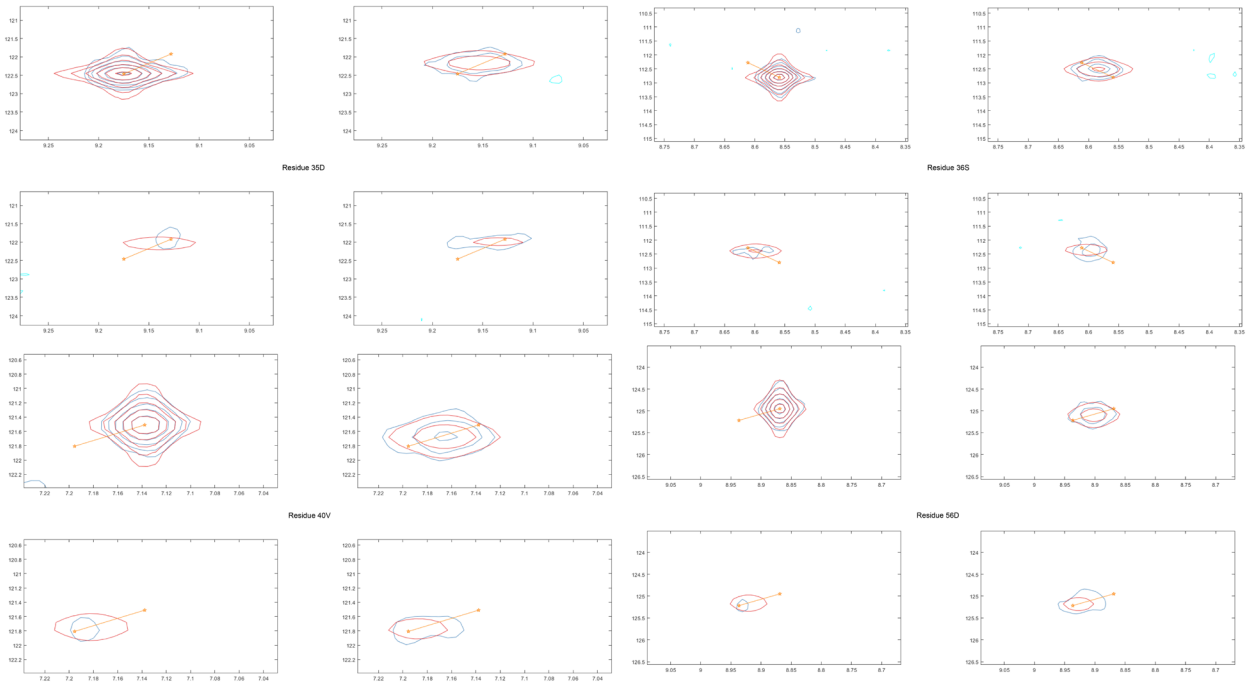
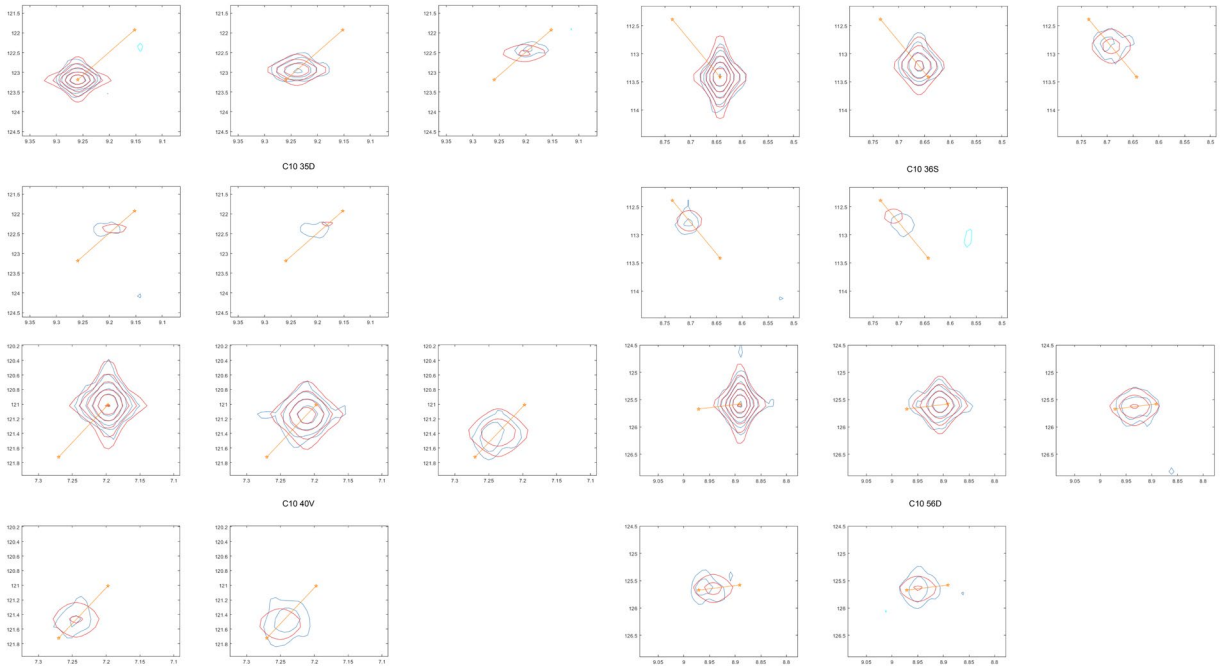



Figure S4 TITAN analysis of the C6-AcpP FabA titration. Real (red) and simulated (blue) titration peaks are shown for four selected residues of the TITAN analysis. The analysis was performed using the flexible docking method, allowing flexibility in the stoichiometry. The error was analyzed using 300 seps of bootstrap error analysis. Though there is significant signal loss in the real data, the peaks overlay well demonstrating a well fit model.




Figures S5 TITAN analysis of the C10-AcpP FabA titration. Real (red) and simulated (blue) titration peaks are shown for four selected residues of the TITAN analysis. The analysis was performed using the flexible docking method, allowing flexibility in the stoichiometry. The error was analyzed using 300 seps of bootstrap error analysis. Overall, the peaks overlay well between the real and simulated data.

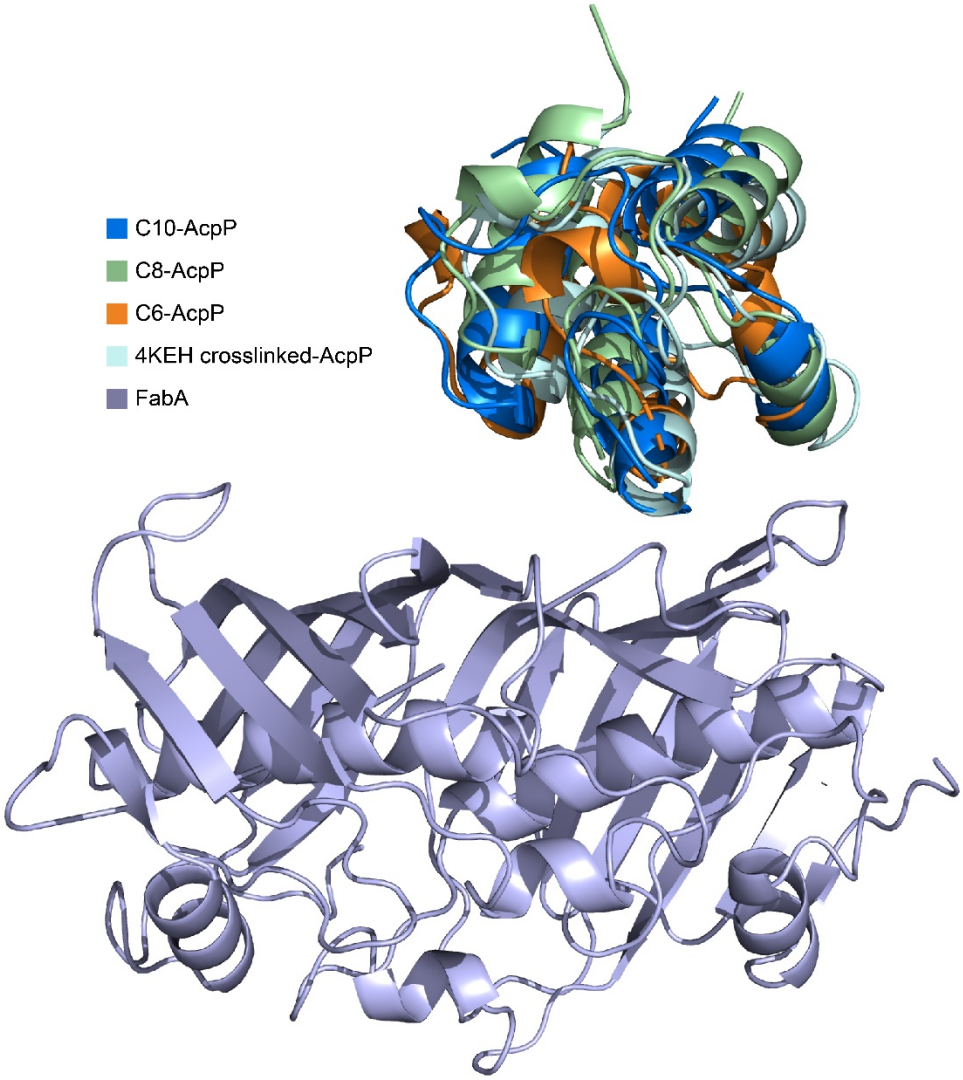



Figure S6 The most stable states of the FabA interaction with C6, C8, and C10-AcpP compared to the crosslinked AcpP. Showing the high similarity of all of the states, demonstrating the minor structural changes necessary to facilitate the different binding modes.

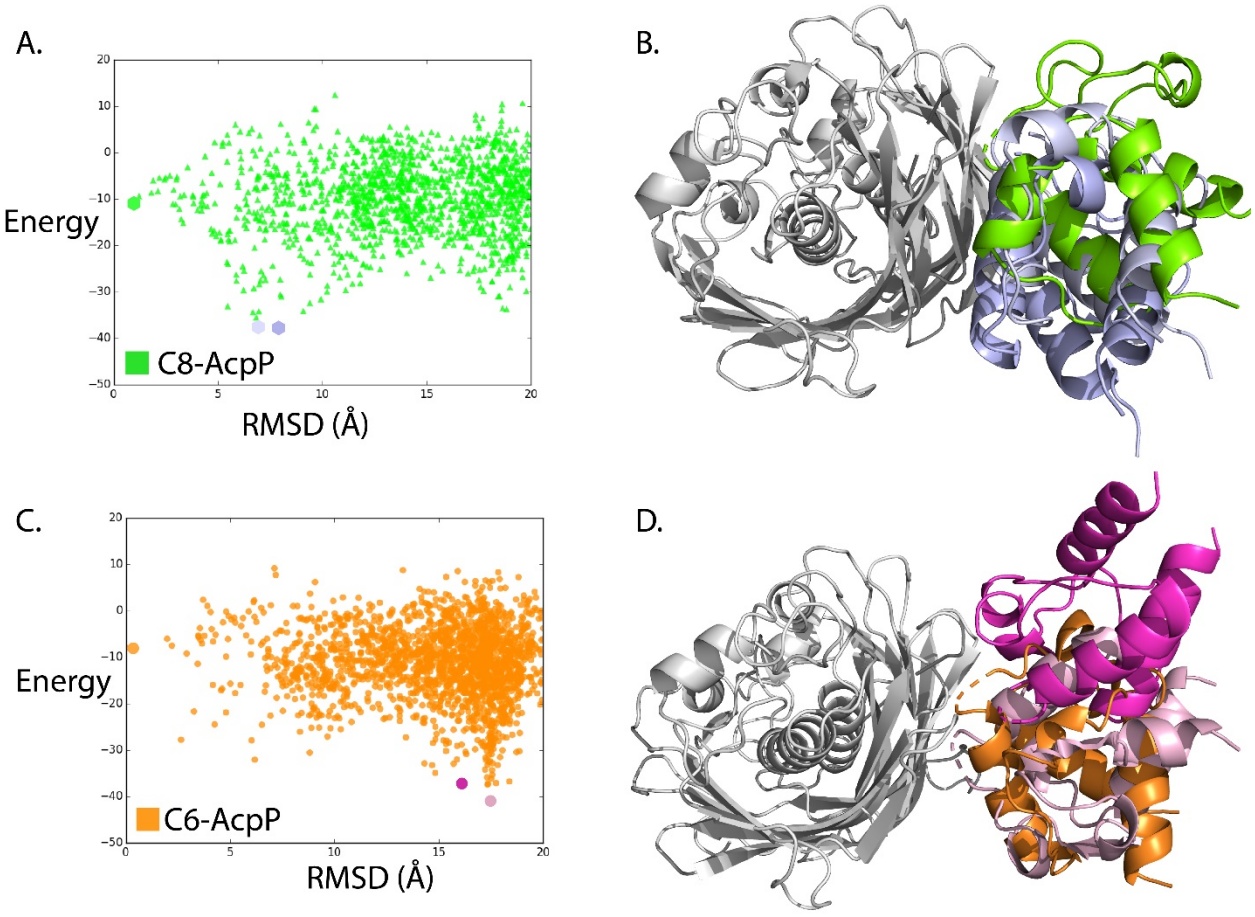



Figure S7 Docking details of the C6 and C8-AcpP with the *E. coli* dehydratase, FabA. The docked states of the AcpPs with FabA are shown with greater detail, for each chain length the most favorable low RMSD state is shown as well as the two most stable states. The more stable states associate in catalytically inactive states, with the phosphopanetheine too distant to chain flip into the active site.

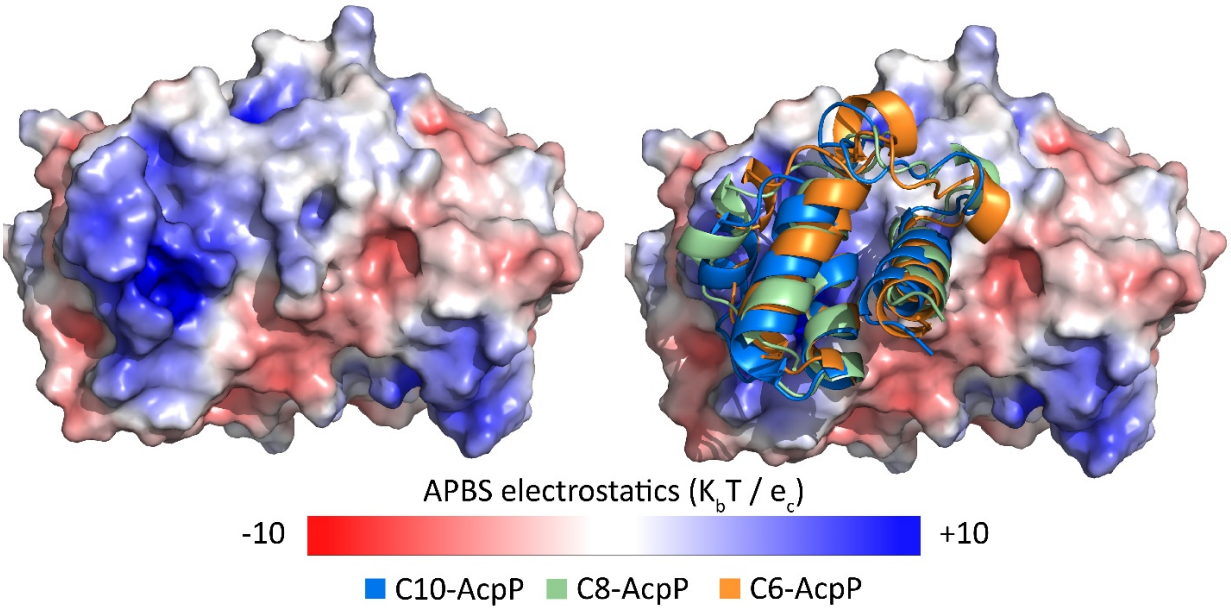
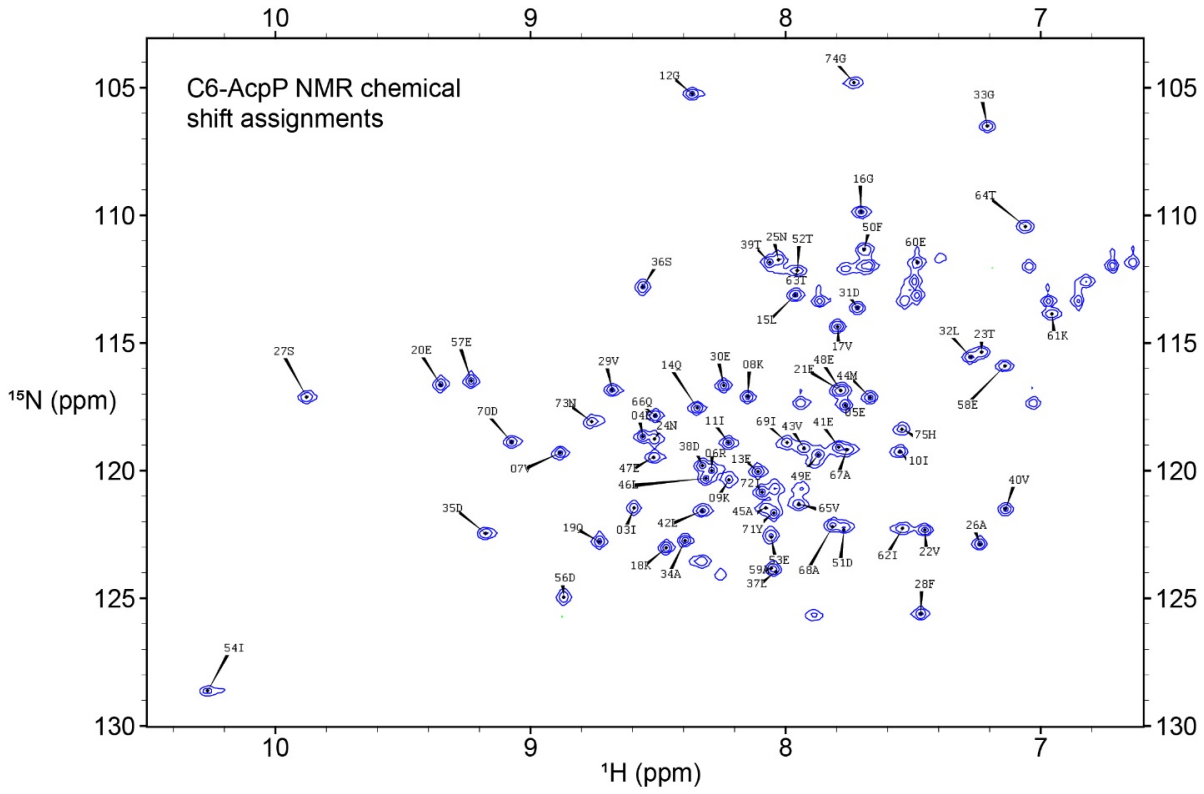



Figure S8 Electrostatics of the FabA interface. The interface of FabA is known to be very positively charged, with the AcpPs being very negatively charged. This simple electrostatic matching explains the plethora of stable states which can exist, also highlighting the remarkable substrate selectivity considering the seemingly simple interaction.




Figures S9 1H-15N HSQC assignments C6-AcpP. C8-AcpP assignments were taken from a previous study performed in the lab[26] and used to assign the zero-point spectra of the C6-AcpP. Given the small difference between C6 and C8-AcpP the spectra could be assigned by overlay.

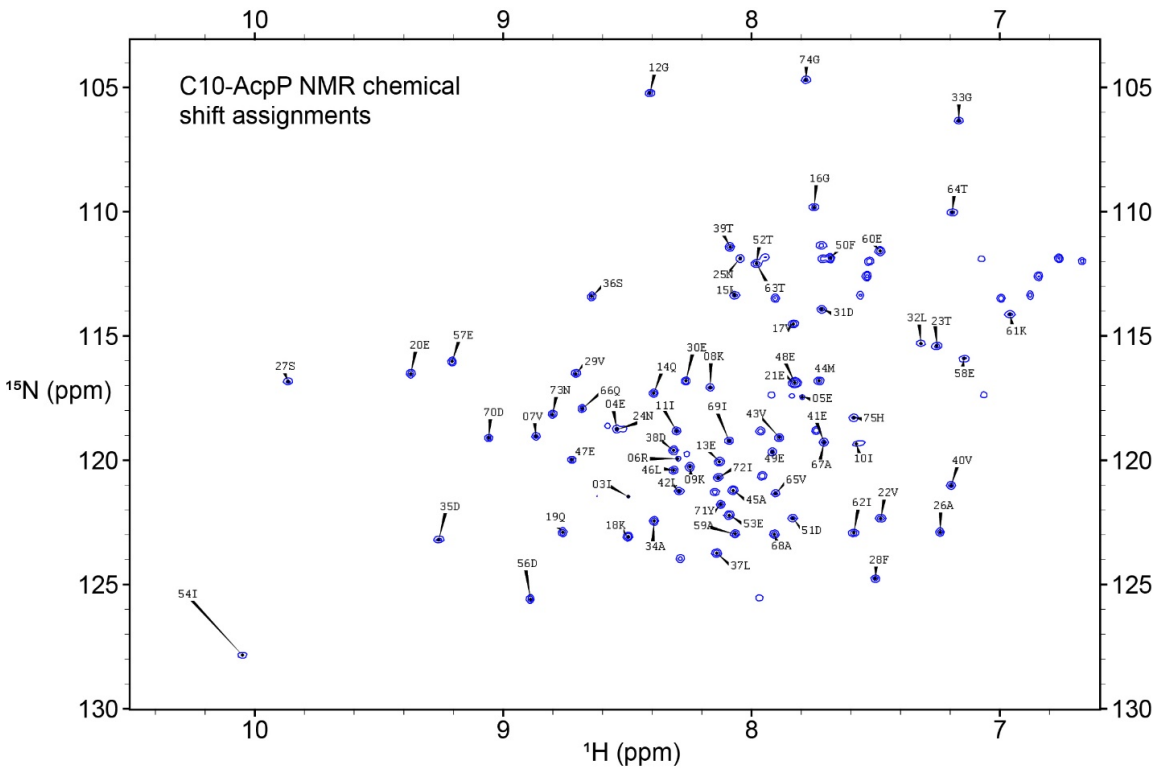



Figure S10 1H-15N HSQC assignments C10-AcpP. C8-AcpP assignments were taken from a previous study performed in the lab[26] and used to assign the zero-point spectra of the C10-AcpP. Given the small difference between C10 and C8-AcpP the spectra could be assigned by overlay.

Table S1 C6 titration chemical shifts

| Residue | Zero-point $^{1}$H peak | Zero-point $^{15}$N Peak | Saturated $^{1}$H peak | Saturated $^{15}$N Peak |
|---|---|---|---|---|
| **03I** | 8.594 | 121.468 | 8.582 | 121.505 |
| **04E** | 8.559 | 118.667 | 8.562 | 118.718 |
| **05E** | 7.766 | 117.437 | 7.761 | 117.408 |
| **06R** | 8.29 | 119.993 | 8.287 | 120.115 |
| **07V** | 8.883 | 119.305 | 8.875 | 119.257 |
| **08K** | 8.148 | 117.107 | 8.161 | 117.142 |
| **09K** | 8.221 | 120.35 | 8.206 | 120.383 |
| **10I** | 7.552 | 119.255 | 7.557 | 119.184 |
| **11I** | 8.223 | 118.909 | 8.195 | 119.051 |
| **12G** | 8.366 | 105.245 | 8.358 | 105.231 |
| **13E** | 8.108 | 120.033 | 8.111 | 120.018 |
| **14Q** | 8.346 | 117.546 | 8.34 | 117.4 |
| **15L** | 7.962 | 113.123 | 7.966 | 112.932 |
| **16G** | 7.704 | 109.866 | 7.694 | 109.873 |
| **17V** | 7.795 | 114.356 | 7.783 | 114.152 |
| **18K** | 8.467 | 123.019 | 8.47 | 122.978 |
| **19Q** | 8.73 | 122.775 | 8.726 | 122.739 |
| **20E** | 9.352 | 116.633 | 9.338 | 116.625 |
| **21E** | 7.784 | 116.86 | 7.785 | 116.849 |
| **22V** | 7.454 | 122.327 | 7.458 | 122.331 |
| **23T** | 7.231 | 115.351 | 7.274 | 115.533 |
| **24N** | 8.514 | 118.76 | 8.531 | 118.713 |
| **25N** | 8.028 | 111.736 | 8.025 | 111.705 |
| **26A** | 7.238 | 122.87 | 7.245 | 122.851 |
| **27S** | 9.877 | 117.115 | 9.882 | 117.261 |
| **28F** | 7.47 | 125.608 | 7.47 | 125.797 |
| **29V** | 8.679 | 116.848 | 8.679 | 116.93 |
| **30E** | 8.242 | 116.664 | 8.257 | 116.606 |
| **31D** | 7.718 | 113.616 | 7.724 | 113.492 |
| **32L** | 7.274 | 115.549 | 7.274 | 115.533 |
| **33G** | 7.209 | 106.502 | 7.218 | 106.474 |
| **34A** | 8.394 | 122.732 | 8.381 | 122.696 |
| **35D** | 9.175 | 122.461 | 9.178 | 121.988 |
| **36S** | 8.559 | 112.806 | 8.591 | 112.319 |
| **37L** | 8.038 | 123.952 | 8.017 | 123.985 |
| **38D** | 8.326 | 119.806 | 8.287 | 120.115 |
| **39T** | 8.062 | 111.832 | 8.025 | 111.705 |
| **40V** | 7.137 | 121.507 | 7.164 | 121.723 |
| **41E** | 7.791 | 119.081 | 7.779 | 119.221 |

| | | | | |
|---|---|---|---|---|
| **42L** | 8.325 | 121.573 | 8.318 | 121.058 |
| **43V** | 7.927 | 119.121 | 7.91 | 119.605 |
| **44M** | 7.667 | 117.139 | 7.696 | 117.324 |
| **45A** | 8.077 | 121.461 | 8.071 | 121.249 |
| **46L** | 8.312 | 120.304 | 8.287 | 120.115 |
| **47E** | 8.517 | 119.474 | 8.512 | 119.846 |
| **48E** | 7.784 | 116.86 | 7.785 | 116.849 |
| **49E** | 7.871 | 119.375 | 7.91 | 119.605 |
| **50F** | 7.692 | 111.335 | 7.679 | 111.518 |
| **51D** | 7.771 | 122.325 | 7.77 | 122.328 |
| **52T** | 7.953 | 112.169 | 7.962 | 112 |
| **53E** | 8.057 | 122.599 | 8.039 | 122.414 |
| **54I** | 10.265 | 128.629 | 10.268 | 128.718 |
| **56D** | 8.869 | 124.956 | 8.92 | 125.182 |
| **57E** | 9.232 | 116.489 | 9.238 | 116.908 |
| **58E** | 7.14 | 115.896 | 7.09 | 115.888 |
| **59A** | 8.055 | 123.825 | 8.019 | 123.839 |
| **60E** | 7.483 | 111.845 | 7.501 | 112.068 |
| **61K** | 6.956 | 113.854 | 6.965 | 113.944 |
| **62I** | 7.541 | 122.265 | 7.504 | 122.08 |
| **63T** | 7.953 | 112.169 | 7.94 | 112.248 |
| **64T** | 7.06 | 110.431 | 7.013 | 110.471 |
| **65V** | 7.948 | 121.318 | 7.919 | 120.633 |
| **66Q** | 8.511 | 117.845 | 8.507 | 117.913 |
| **67A** | 7.76 | 119.181 | 7.779 | 119.221 |
| **68A** | 7.816 | 122.195 | 7.8 | 121.954 |
| **69I** | 7.994 | 118.901 | 7.982 | 118.879 |
| **70D** | 9.074 | 118.872 | 9.071 | 118.822 |
| **71Y** | 8.046 | 121.657 | 8.071 | 121.249 |
| **72I** | 8.092 | 120.832 | 8.097 | 120.904 |
| **73N** | 8.76 | 118.077 | 8.757 | 118.111 |
| **74G** | 7.731 | 104.801 | 7.716 | 104.845 |
| **75H** | 7.544 | 118.371 | 7.541 | 118.525 |

Table S2 C10 titration chemical shifts

| Residue | Zero-point $^{1}$H peak | Zero-point $^{15}$N Peak | Saturated $^{1}$H peak | Saturated $^{15}$N Peak |
|---|---|---|---|---|
| **03I** | 8.495 | 121.464 | 8.502 | 121.308 |
| **04E** | 8.542 | 118.749 | 8.54 | 118.735 |
| **05E** | 7.795 | 117.452 | 7.79 | 117.219 |
| **06R** | 8.297 | 119.935 | 8.302 | 120 |
| **07V** | 8.868 | 119.046 | 8.872 | 119.037 |
| **08K** | 8.166 | 117.073 | 8.184 | 117.095 |
| **09K** | 8.248 | 120.269 | 8.233 | 120.34 |
| **10I** | 7.577 | 119.342 | 7.612 | 119.366 |
| **11I** | 8.302 | 118.825 | 8.282 | 118.932 |
| **12G** | 8.409 | 105.229 | 8.405 | 105.152 |
| **13E** | 8.129 | 120.057 | 8.144 | 120.058 |
| **14Q** | 8.393 | 117.301 | 8.394 | 117.237 |
| **15L** | 8.067 | 113.364 | 8.009 | 113.115 |
| **16G** | 7.748 | 109.812 | 7.736 | 109.807 |
| **17V** | 7.831 | 114.519 | 7.807 | 114.223 |
| **18K** | 8.496 | 123.08 | 8.509 | 123.008 |
| **19Q** | 8.76 | 122.896 | 8.765 | 122.818 |
| **20E** | 9.372 | 116.515 | 9.381 | 116.579 |
| **21E** | 7.826 | 116.878 | 7.812 | 116.866 |
| **22V** | 7.479 | 122.337 | 7.487 | 122.345 |
| **23T** | 7.253 | 115.408 | 7.278 | 115.359 |
| **24N** | 8.542 | 118.749 | 8.54 | 118.735 |
| **25N** | 8.045 | 111.875 | 8.054 | 111.766 |
| **26A** | 7.24 | 122.891 | 7.266 | 122.868 |
| **27S** | 9.866 | 116.837 | 9.898 | 117.055 |
| **28F** | 7.501 | 124.759 | 9.941 | 116.826 |
| **29V** | 8.707 | 116.514 | 7.529 | 125.278 |
| **30E** | 8.264 | 116.808 | 8.708 | 116.657 |
| **31D** | 7.717 | 113.924 | 8.284 | 116.668 |
| **32L** | 7.318 | 115.295 | 7.734 | 113.674 |
| **33G** | 7.165 | 106.34 | 7.278 | 115.359 |
| **34A** | 8.392 | 122.443 | 7.214 | 106.388 |
| **35D** | 9.259 | 123.196 | 8.4 | 122.548 |
| **36S** | 8.643 | 113.41 | 9.195 | 122.427 |
| **37L** | 8.14 | 123.745 | 8.69 | 112.796 |
| **38D** | 8.314 | 119.6 | 8.086 | 123.92 |
| **39T** | 8.086 | 111.412 | 8.511 | 119.702 |
| **40V** | 7.196 | 121.015 | 8.116 | 111.123 |
| **41E** | 7.708 | 119.277 | 7.238 | 121.472 |

| | | | | |
|---|---|---|---|---|
| **42L** | 8.292 | 121.244 | 7.723 | 119.189 |
| **43V** | 7.887 | 119.086 | 8.363 | 121.115 |
| **44M** | 7.727 | 116.796 | 7.871 | 118.961 |
| **45A** | 8.074 | 121.217 | 7.766 | 117.281 |
| **46L** | 8.313 | 120.4 | 8.165 | 121.722 |
| **47E** | 8.723 | 119.985 | 8.233 | 120.34 |
| **48E** | 7.826 | 116.878 | 8.543 | 119.483 |
| **49E** | 7.916 | 119.668 | 7.812 | 116.866 |
| **50F** | 7.682 | 111.868 | 7.94 | 119.578 |
| **51D** | 7.834 | 122.335 | 7.701 | 111.931 |
| **52T** | 7.98 | 112.08 | 7.806 | 122.417 |
| **53E** | 8.09 | 122.21 | 8.103 | 112.387 |
| **54I** | 10.05 | 127.841 | 10.261 | 128.585 |
| **56D** | 8.89 | 125.59 | 8.941 | 125.65 |
| **57E** | 9.205 | 116.025 | 9.316 | 116.328 |
| **58E** | 7.144 | 115.909 | 7.093 | 115.894 |
| **59A** | 8.066 | 122.969 | 8.236 | 123.12 |
| **60E** | 7.482 | 111.587 | 7.522 | 111.922 |
| **61K** | 6.959 | 114.128 | 6.993 | 114.204 |
| **62I** | 7.588 | 122.924 | 7.56 | 122.459 |
| **63T** | 7.98 | 112.08 | 8.103 | 112.387 |
| **64T** | 7.192 | 110.027 | 7.091 | 110.209 |
| **65V** | 7.903 | 121.339 | 7.961 | 121.079 |
| **66Q** | 8.682 | 117.93 | 8.568 | 117.784 |
| **67A** | 7.708 | 119.277 | 7.774 | 119.229 |
| **68A** | 7.907 | 122.98 | 7.874 | 122.261 |
| **69I** | 8.09 | 119.221 | 8.044 | 118.91 |
| **70D** | 9.058 | 119.105 | 9.095 | 118.972 |
| **71Y** | 8.123 | 121.783 | 8.119 | 121.503 |
| **72I** | 8.134 | 120.692 | 8.12 | 120.913 |
| **73N** | 8.801 | 118.155 | 8.799 | 118.132 |
| **74G** | 7.781 | 104.699 | 7.765 | 104.792 |
| **75H** | 7.588 | 118.288 | 7.587 | 118.428 |

Supplemental References


1. 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. https://doi.org/10.1038/nature12810.

2. NMRPipe: A multidimensional spectral processing system based on UNIX pipes | SpringerLink, (available at https://link.springer.com/article/10.1007/BF00197809).

3. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinforma. Oxf. Engl. 31, 1325–1327 (2015).

4. J. D. Hunter, Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95 (2007).

5. PyMOL. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

## Chapter 3. Protein-protein Interaction based substrate control in the E. coli octanoic acid transferase, LipB

### 3.1 Introduction

*De novo* lipoic acid biosynthesis occurs in all organisms as a branch point from type II fatty acid biosynthesis (FAB), and lipoic acid is the only known essential product of human mitochondrial FAB[1–5]. Octanoic acid is transferred from within the FAB onto a lipoylated target protein, whereupon thiol moieties are subsequently added through the activity of iron-sulfur cluster enzymes[1,6–8]. Control of octanoate transfer from the FAB must be maintained, as the role and structure of lipoic acid is reliant on the proper chain length and oxidative state of the fatty acid from which it is derived. In *E. coli*, the fidelity to select a single fatty acid from within the 30-35 potential acyl substrates attached to ACP poses a selectivity mechanism [9–11] that remains elusive [12,13].

FAB is an iterative, multi-enzyme pathway in which each reaction step is catalyzed upon a fatty acid precursor that is covalently attached to the acyl carrier protein (AcpP in *E. coli*)[14]. AcpP is a small, 77 amino acid protein with a four α-helical bundle structure[15,16]. The acyl substrates are carried on a 4'-phosphopantetheine cofactor attached to serine 36 of the AcpP, which carries each fatty acyl intermediate attached as a thioester. In solution, AcpP sequesters acyl cargo within a hydrophobic pocket between its α-helices, only presenting the hydrolyzable thioester once it favorably interacts with a partner protein through protein-protein interactions (PPIs)[17]. In the case of LipB, the sourcing of octanoyl-ACP must occur after enoyl reduction but before the substrate can re-enter the elongation cycle for another iteration (Figure 1A). Here we have pursued an understanding for how the LipB accomplishes this highly selective interaction, where LipB rapidly intercepts C8-AcpP with high fidelity.

LipB transfers octanoyl groups scavenged from AcpP to an active site cysteine 169' (residues of LipB will hereafter be labeled as prime), with the LipB's hydrophobic pocket sheltering the lipid tail before transferring it to E2 or other lipoyl domains. This creates an octanoyl-modified enzyme, freeing the LipB to scavenge more octanoic acid substrates[18]. LipB is required to source and attach octanoic acid from AcpP, as neither free octanoic acid nor octanoyl-CoA are substrates, requiring LipB to carefully select substrates attached to the AcpP or risk inactivating downstream enzymes[19].

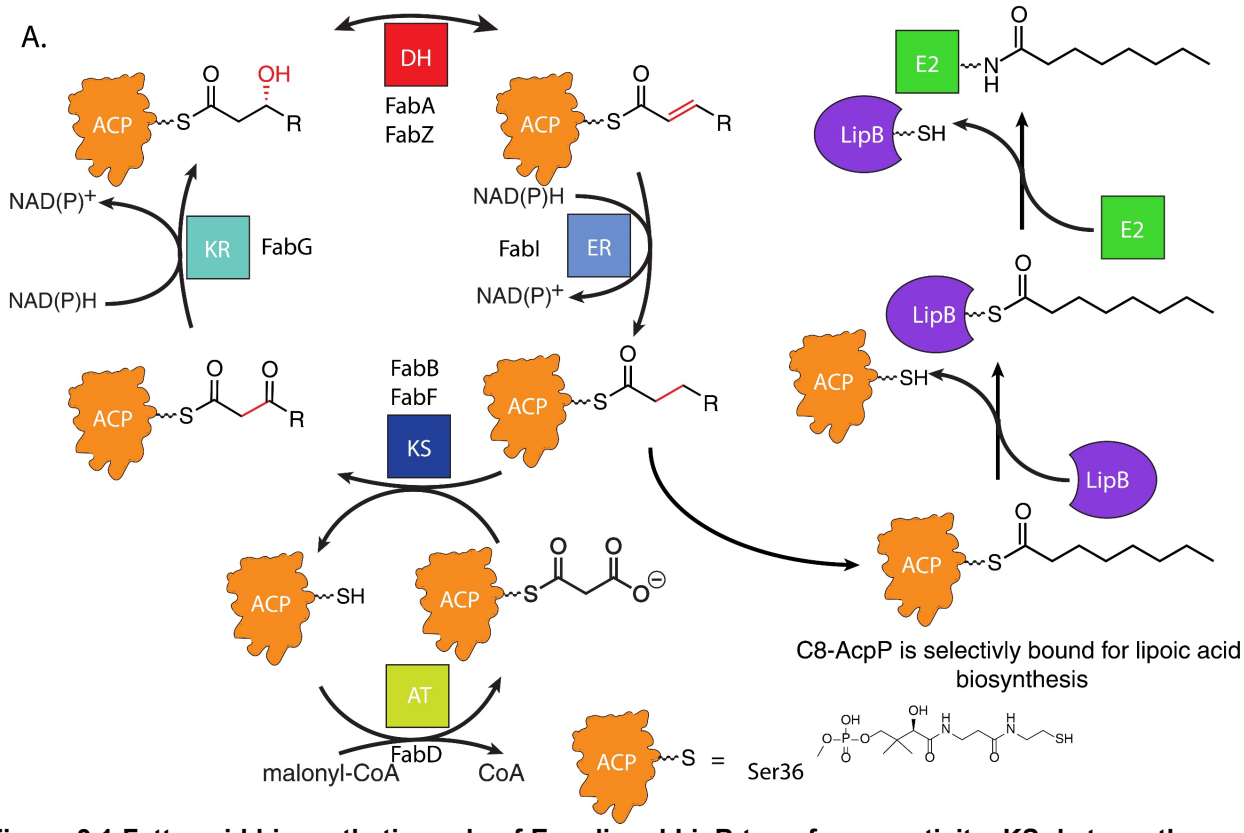



**Figure 3.1 Fatty acid biosynthetic cycle of E. coli and LipB transferase activity. KS: ketosynthase, KR: ketoreductase, DH: dehydratase, ER: enoylreductase, TE: thioesterase, AT: malonyl acyltransferase. The central steps occur, with ACPs carrying all of the substrates between partners. LipB must selectively bind a saturated eight carbon ACP. The acyl chain is transferred onto LipB to be carried to E2 domains and subsequent lipoic acid biosynthetic enzymes.**

Recent evidence has suggested that the AcpP • LipB interaction can exert allosteric control over the substrates it interacts with prior to catalysis through control of the "chain flipping" event[20]. Chain flipping is the term applied to the exit of the acyl chain from the carrier protein pocket[21]. This creates a "control" step which must occur prior to any catalysis. Prior studies have reported decanoic acid crystallized within the LipB active site, suggesting that it is possible for the C10-acyl chain to fit into the LipB pocket[22]. Further, we recently demonstrated that the acyl chain of dodecanoyl-AcpP does not chain flip when attached to the *E. coli* AcpP, but mutation of the LipB interface residue R145 can induce loss of chain length selectivity for chain flipping [20], indicating both a substrate selectivity by wild- type LipB and the importance of the proper protein-protein interface for this selectivity. To resolve this PPI-based control mechanism at atomic detail, we chose to perform NMR studies of AcpP with C6, C8, and C10 acyl chains titrated with LipB to observe interaction changes based on chain length. These chain lengths represent the known LipB substrate and two most similar chain lengths seen in the cell. This data was then used to guide high-resolution *in silico*

docking to identify the surface features responsible for the experimentally identified binding differences. We have shown that the implementation of NMR titration experiments to guide docking algorithms can accurately and reproducibly deduce PPI poses in ACP-dependent pathways[23].

It has been suggested that unique features imparted by identity of the acyl chain can likely serve as a source for binding discrimination by enzymes[24–26]. Furthermore, control of reactivity based on substrate has been seen by crosslinking[27] and NMR[20], but developing a structural model requires atomic level detail. Here we identify the structural features of LipB that allow interaction with C8-AcpP while inducing structural hindrance to C6- and C10-AcpP binding, further elucidating the mechanism and role of PPIs in carrier protein-dependent enzymes.

### 3.2 Results

#### NMR titration to examine residue-by-residue interaction of acyl-AcpP with LipB

It has been established that solution-state $^{15}N$-$^{1}H$ HSQC NMR spectroscopy can function in appreciating the transient dynamic interactions between AcpP and partner proteins[28,29]. Given the known specificity of LipB, we sought to elucidate how substrate specificity is conferred by PPI with AcpP carrying different acyl cargo. In order to detect small functional differences in the interactions of LipB, AcpPs of different chain lengths were prepared as C6-, C8-, and C10-linked 4'-pantetheinamide probes[30](Figure S1-S3, Table S1-S3). These AcpP species were titrated with increasing concentrations from none to beyond saturation, at 2 molar equivalents of LipB in the case of C6-AcpP, 1.5 equivalents of LipB in the case of C8-AcpP, and 2 molar equivalents of LipB in the case of C10-AcpP. The effect of increasing LipB concentration on each AcpP species was examined.

The first observation made when examining perturbations against one another is the difference in chemical shift perturbation (CSP) magnitude between the chain lengths (Figure 2, S4). It is immediately clear that the degree of perturbation is greatest in C8-AcpP, despite all experiments being titrated to saturation. Further, the CSPs of C8-AcpP reveal unique interactions. Whereas most titrations have been noted to have little effect on helix I of the AcpP, LipB effects strong CSP throughout the early residues and

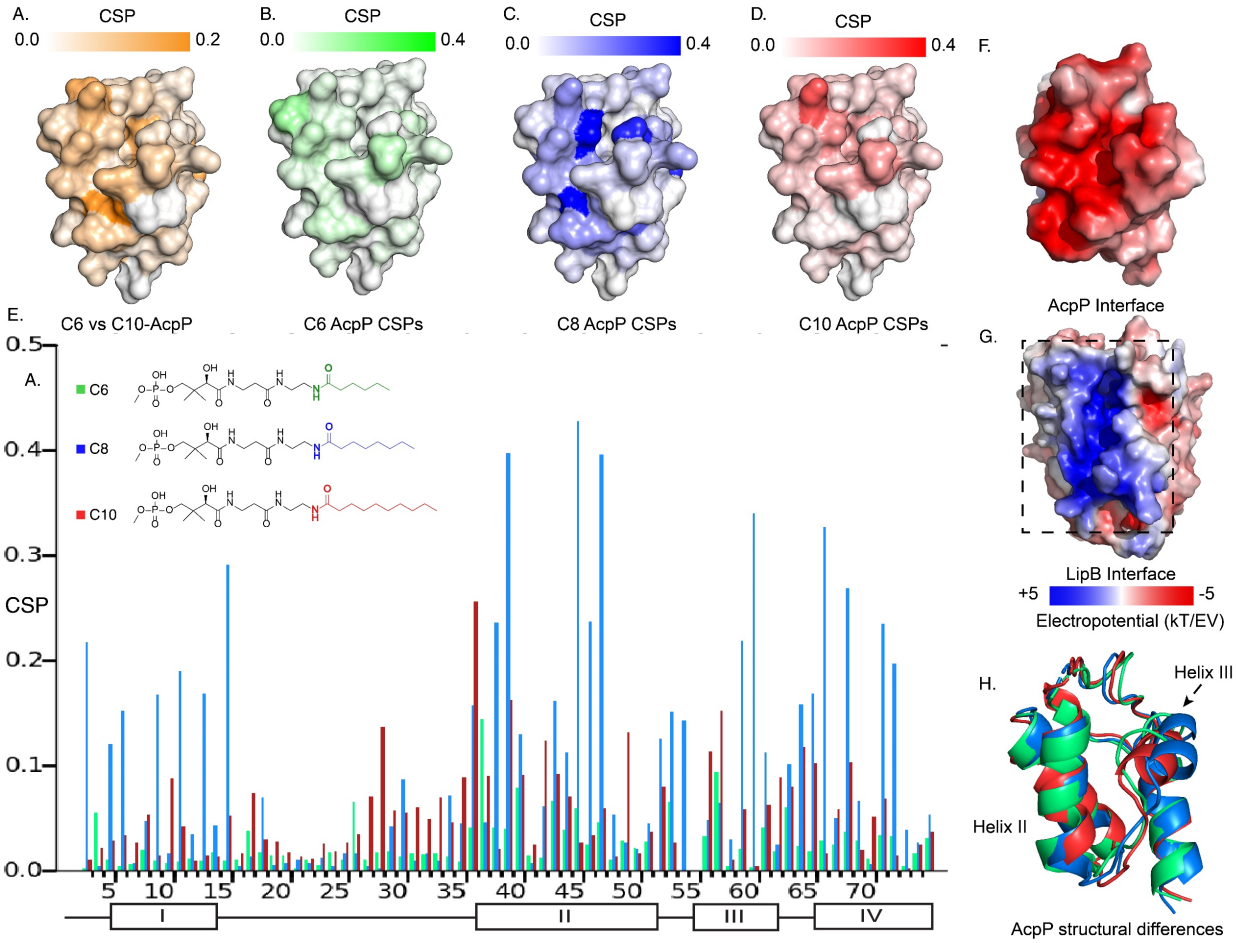



**Figure 3.2 Titration NMR to probe the substrate selectivity of LipB. A) Chemical shift difference between the C6- and C10-AcpP without LipB. CSPs are colored based on the magnitude of the CSP. Highlighting those residues which should be expected to be most different between the two proteins in solution. B) CSPs of C6-AcpP titrated with LipB. CSPs are colored based on the magnitude of perturbation. C) CSPs of the C8-AcpP titrated with LipB. D) CSPs of C10-AcpP titrated with LipB. E) Chemical shift perturbations (CSPs) of C6-, C8-, and C10-AcpP titrated with LipB. Individual spectra and bar charts are shown in Figure S1-3. F) Electrostatic surface of AcpP, showing the electronegative surface of AcpP. Values were calculated using the APBS electrostatics plugin in Pymol. G) Electrostatic surface of LipB, highlighting the positive surface for the AcpP interaction. H) The structures of C6 (Green), C8 (Blue), and C10 (Red) AcpP derived from molecular dynamics. The most significant differences are seen in the orientation of helix III, with C8 AcpP having the most space between helix III and helix II.**

lasting through to residue 18. After this, there is a drop-off in CSP through the end of loop 1 to residue 30. Small perturbations rise above background at residues 34 and 35, and CSPs continue consistently through helix 2. Next, there are consistent perturbations through loop 2, helix 3, and helix 4. The most unique region

of CSP is the very strong migrations occurring in helix I (Figure 2A). These perturbations appear consistent in these regions among C6-, C8- and C10-AcpP, with varying magnitudes.

**Binding thermodynamics and kinetics demonstrate specificity experimentally**

Though CSPs are not a quantitative measure of binding, C8-AcpP binding LipB displays considerably stronger perturbations than C6- or C10-AcpP. In order to deduce quatitative binding parameters of AcpP tethering the three chain lengths, we examined the NMR titration data by applying TITAN line shape analysis[31]. Here, C8-AcpP exhibited a 47.2 ± 5.1 μM Kd and a low off rate of 633 ± 98 $s^{-1}$. The C6-AcpP bound more poorly, with a 189.9 ± 15.21 μM Kd and off rate of 4237 ± 2544 $s^{-1}$. C10-AcpP bound slightly better than the C6-AcpP with a 134.8 ± 34.0 ±M Kd and off rate of 1521 ± 225 $s^{-1}$ (Figure S8-S10). This significant difference in off rate and less significant difference in binding constant are supported by the comparative magnitude of the respective CSP data. TITAN analysis quantitatively supports control LipB maintains over interactions with AcpPs tethering different acyl chains, both in terms of binding strength and associated off rate.

**Docking analysis to identify chain length specific interactions**

LipB was first modeled by homology modeling to the 2QHS *Thermus thermophilus* lipoyltransferase with ICM Homology [32,33](Figure S5). The *Thermus thermophilus* structure was chosen because it was the highest homology crystal structure available. AcpP structures were derived from molecular dynamics (MD) simulations and represented the highest population state seen in the simulation[20](Figure S7).

To perform *in silico* docking experiments, the ICM Fast Fourier Transform protocol was used to generate high quality docking poses and scores for the AcpP•LipB interface and used to sample AcpP conformations across the entire LipB protein surface. The resulting docking poses were organized based on RMSD from the previously published model of the C8-AcpP • LipB docked complex[20](Figure S6).

C8-AcpP adopted a low energy structure at 4.03Å RMSD, with an energy of -50.5 kcal/mol (Figure 3A) based on the ICM energy function. There was a second low energy docked conformation at 15.8Å RMSD, which was ruled out as inactive with serine 36  17Å from the LipB active site and the AcpP rotated away form the LipB pocket (Figure 3B,C). The C6-AcpP has significantly higher energy poses (Figure 3A) at low

RMSD, with the only similarly stable state at 20Å RMSD. Like C6-AcpP, C10-AcpP has most stable poses at 18.4Å and 24.5Å RMSD, with no similarly stable state near the C8-AcpPs.

The high RMSD states of C8-AcpP•LipB interaction are clearly not oriented for substrate delivery (Figure S5), with the 4'-phosphopantetheine positioned away from the active site. Therefore, lower RMSD states were examined to explore how the difference in AcpP structure translated to different energetics for binding. C8-AcpP binds tightly onto the LipB surface, with six arginine or lysine residues available for salt bridges with discrete AcpP residues. The most stable state of C8-AcpP has each of these residues coordinating closely, within 5Å. The C8-AcpP•LipB interactions begin at helix II, with E41 nearby R99', E47 with R142', and E49 with R93'. D51 and R145' remain 4.5Å apart, but side chain rotation can bring the residues within range of a salt bridge. E53 with R144' and E60 with K54' finishes the total salt bridges.

The C6-AcpP model displays a significantly poorer binding surface with LipB (Figure 3A,F), matching the results of the thermodynamic and CSP data. Specifically, E41, E48, and D51 lose the proper orientation of interaction (Figure S3). This is due to the structural effects of chain length shortening upon AcpP, with the structure most perturbed on helix II, loop II, and the orientation of helix III. There are two residues within range for a salt bridge in the C6-AcpP•LipB complex: E47 with the pair R142' and R144', and E57 with K54'. It is interesting to not how distinctly the binding surface of LipB can be effected by the small structural changes between C6 and C8-AcpP. The C10-AcpP model similarly binds more poorly to the LipB surface, displaying an ability to only bind four residues at the LipB surface. E41 and E47 appear to be out of any vicinity to interact with R99' or R142', but E48 and D51 are within interacting distance of R93'. E53 lies within 3.6Å of R142', and E60 is 3.8Å from K54' (Figure S4).

The disparities in possible interactions were further highlighted by aligning the C6- and C10-AcpP structures onto the C8 docked pose to reveal their structural differences. The most immediate difference between them is the orientation of helix III. In the C8-AcpP the helix is oriented outwards, creating space around the bottom of AcpP helix III for the residues 50'-55' of LipB (Figure 3F). For example, residue D56 lies nearby Q51' on LipB in the C8-AcpP, but the more closed structure of C6- and C10-AcpP place D56 in direct steric clash with LipB. This reflects a structural filtering mechanism, where the surface appears to be arranged to interact with the structural features of C8-AcpP. Other orientations of helix III disallow proper

binding for chain flipping from helix II. These regions of selectivity overlay with the differences in structure seen in the NMR data. Though the C6- and C10-AcpPs can likely relax their structure to better bind the LipB surface, this initial instability and necessary relaxation could slow the association, explaining the poor binding of non-substrate AcpPs. The effects of this selectivity are evident in the respective CSPs, thermodynamics, and conformational landscape of the docking calculations.

**Comparing CSPs to the docked model**

To combine both data sets the NMR titration of C8-AcpP was compared to the docked model. Beginning at helix I there are perturbations occurring through most of the helix, though they are most prominent at the end of the helix. These likely correspond to the interactions across the “right” side of the channel which binds helix II. There are several residues at the base of the helix which would likely experience movement while the ACP adopts a bound state, but the those residues at the end are close enough to likely be in direct contact. Leading to a higher degree of perturbation. The CSPs almost fully drop off until helix II, matching the pose where there is little partner in proximity. Helix II shows large perturbations through all of the helix until the bottom, this agrees with the model well. Especially the uncharged residues such as T39, V43, and A45 which usually see small CSPs would be participating in the buried interface. The perturbations end at the bottom of the helix, to be seen again at E53, which is seen binding R144. On helix III further down the ACP D56 has little perturbation, though it is usually an important residues in interactions. However this matches the model, where the acidic side chain is either binding a backbone or associating with Q68 either way a weaker than salt bridge interaction. However at the end of helix III strong perturbations continue with a salt bridge at E60. Helix IV sees perturbations which are likely linked to the movement of the helix upon chain flipping. Seen especially in V65, A64, Q66, and Y71 the perturbed residues at the top of helix IV may also be due to movement of the dynamic helix III upon binding. Especially when all the residues on the loop before and after helix III see some small perturbation.

The C6 and C10-AcpP perturbations are very different in their distribution. There are a limited number of electrostatic CSPs at D35, E49, D56, and E57 in the C10-AcpP titration. There were more larger perturbations of hydrophobic residues at I10, V17, S27, P28, T39, L42, V43, T63, T64, and V65. These perturbations are approximately consistent with C6-AcpP, though C6-AcpP has significantly smaller ones.

The largest effects are seen on helix II and III, seeming to indicate interactions at the interface. The only region which sees interactions greater than C8 is the loop preceding helix II and very top of helix II. The difference in hydrogen bonding matches the docked model for C10, where the helix III angle better aligns D56 and E57 to form interactions. This contrasts with C8-AcpP where D56 was not aligned for a strong interaction. This lower number of salt bridges and poorer interaction suggests a model where there are sets of interactions which must be sufficiently strong to induce chain flipping. Chain flipping is a large dynamic event, we propose that the deeply buried interface and multiple salt bridges are essential to drive this. With sufficient interactions shifting the residues such as Y71 and I54 to close the acyl pocket and initiate flipping. This stronger surface also explains the higher degree of CSP with weak interactions by C6 and C10-AcpP lowering the degree of structural perturbation upon binding. As well as decreasing the time spent in the "bound" state in solution, resulting in smaller shifts on the magnet.

### 3.3 Discussion

A significant kinetic advantage for any ACP partner protein, currently tallied at 27 known enzymes and regulatory proteins in *E. coli*[9–11], is the ability to discern acyl identity without the requirement of chain flipping. The ability to discern acyl chains based on the initial PPIs significantly increases the efficiency of this selection process and provides thermodynamic control to maintain the fidelity of lipoic acid biosynthesis. We have recently demonstrated how discrete salt bridge interactions at the protein interface can differentiate between C8- and C12-AcpP for chain flipping[20]. We have now determined the comparative binding constants and CSPs of LipB with acyl-AcpPs of both shorter and longer chain lengths with that of the natural C8 substrate, indicating a clear ability of LipB to select for interaction with C8-AcpP. This is accomplished by possessing a surface that can complement the specific shape of octanoyl-sequestered (C8-) AcpP, while deterring interactions with C6- and C10-AcpP. This selectivity is primarily reliant on the helix III perturbations in response to the sequestered acyl chain lengths, previously identified in numerous experimental and theoretical studies. Structurally, this model leverages unique conformational features of AcpP induced by the different chain lengths, an incredibly useful evolutionary feature when selectivity for a single chain length is required.

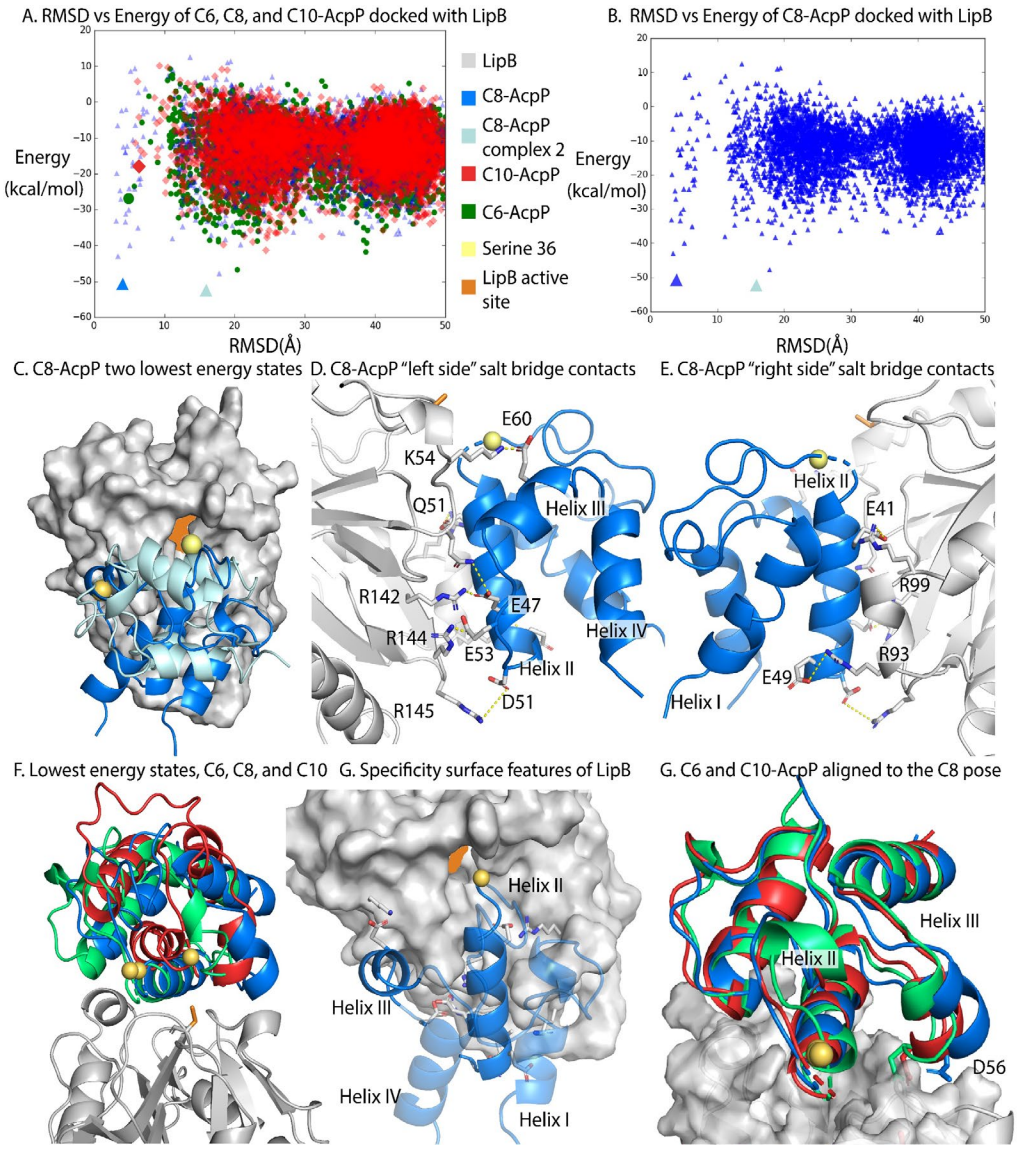



**Figure 3.3 Structural analysis of the AcpP•LipB surface. A)RMSD vs energy plot of the tested AcpPs. RMSD was determined against the previously published Cluspro model. Individual models examined in other panels are highlighted by larger colored icons. B) RMSD vs energy plot of the C8-AcpP with LipB. RMSD was determined against the previously published model. The two lowest energy states are examined in panel C. C) Comparison of the two lowest energy states of the C8-AcpP docked to LipB. The ~4Å RMSD pose (Lime green) is positioned with Ser36 in position for substrate delivery. The ~16Å pose (Light blue) can be seen binding with the substrate positioned far from the active site. D) Salt bridge contacts formed between C8-AcpP and LipB. The "left side" is presented here, representing the majority of the contacts. Individual contacts and distances are reported in SI Table 1. E) Salt bridge contacts formed between C8-AcpP and LipB on the "right side." A table of all contacts formed is reported in SI Table 1. F) "Active" energy states of the C6-, C8-, and C10-AcpP LipB docked poses. The most stable structure which was similar in RMSD to the low energy state in panel A was chosen. G) Architecture of the AcpP•LipB interaction. The interacting helices are shown with the precise orientation necessary to form a stable interaction. Helices I&IV are made transparent to better appreciate the surface of interaction of the other two helices. G) C6- and C10-AcpP are aligned to the docked orientation of C8-AcpP. This demonstrates the steric hinderance which makes C8 able to properly position into the active site, while the C6 and C10 must form the less stable poses seen in panel F.**

Understanding that the ACP•LipB reaction controls chain flipping by PPIs unlocks the potential to control these essential and sensitive interactions through inhibition or engineering. The high homology of LipB shared between bacteria implies that the observations made in *E. coli* will likely extend to other species. Furthermore, targeting the protein interface of AcpP • LipB in a pathogen could avoid potential side effects from activity against the human mitochondrial LipB (Figure S8). These data also suggest an important factor to consider when engineering FAB and related acetate pathway proteins. Where poor interface complementarity could lead to a loss of activity, understanding and optimizing these interactions may prove necessary. While these transient PPIs can be challenging to observe, we have developed an approach that merges new data with prior observations and formed a model that explains both the specificity and efficiency of lipoic acid biosynthesis.

**Chapter 3**, is in full a reprint of the manuscript :Thomas G. Bartholow, Terra Sztain, Megan A Young, Tony D. Davis, Ruben Abagyan, Michael “Protein-protein Interaction based substrate control in the E. coli octanoic acid transferase, LipB.” D. Burkart, In review *RSC Chemical Biology*, (2021). The dissertation author is the primary author of the manuscript.

## 3.4 REFERENCES


(1) Cicchillo, R. M.; Booker, S. J. Mechanistic Investigations of Lipoic Acid Biosynthesis in Escherichia Coli: Both Sulfur Atoms in Lipoic Acid Are Contributed by the Same Lipoyl Synthase Polypeptide. *J. Am. Chem. Soc.* **2005**, *127* (9), 2860–2861. https://doi.org/10.1021/ja042428u.

(2) Booker, S. J. Unraveling the Pathway of Lipoic Acid Biosynthesis. *Chem. Biol.* **2004**, *11* (1), 10–12. https://doi.org/10.1016/j.chembiol.2004.01.002.

(3) Schonauer, M. S.; Kastaniotis, A. J.; Kursu, V. A. S.; Hiltunen, J. K.; Dieckmann, C. L. Lipoic Acid Synthesis and Attachment in Yeast Mitochondria. *J. Biol. Chem.* **2009**, *284* (35), 23234–23242. https://doi.org/10.1074/jbc.M109.015594.

(4) Habarou, F.; Hamel, Y.; Haack, T. B.; Feichtinger, R. G.; Lebigot, E.; Marquardt, I.; Busiah, K.; Laroche, C.; Madrange, M.; Grisel, C.; Pontoizeau, C.; Eisermann, M.; Boutron, A.; Chrétien, D.; Chadefaux-Vekemans, B.; Barouki, R.; Bole-Feysot, C.; Nitschke, P.; Goudin, N.; Boddaert, N.; Nemazanyy, I.; Delahodde, A.; Kölker, S.; Rodenburg, R. J.; Korenke, G. C.; Meitinger, T.; Strom, T. M.; Prokisch, H.; Rotig, A.; Ottolenghi, C.; Mayr, J. A.; de Lonlay, P. Biallelic Mutations in LIPT2 Cause a Mitochondrial Lipoylation Defect Associated with Severe Neonatal Encephalopathy. *Am. J. Hum. Genet.* **2017**, *101* (2), 283–290. https://doi.org/10.1016/j.ajhg.2017.07.001.

(5) Bernardinelli, E.; Costa, R.; Scantamburlo, G.; To, J.; Morabito, R.; Nofziger, C.; Doerrier, C.; Krumschnabel, G.; Paulmichl, M.; Dossena, S. Mis-Targeting of the Mitochondrial Protein LIPT2 Leads to Apoptotic Cell Death. *PLOS ONE* **2017**, *12* (6), e0179591. https://doi.org/10.1371/journal.pone.0179591.

(6) Rock, C. O. Opening a New Path to Lipoic Acid. *J. Bacteriol.* **2009**, *191* (22), 6782–6784. https://doi.org/10.1128/JB.01151-09.

(7) Carreau, J.-P. [32] Biosynthesis of Lipoic Acid via Unsaturated Fatty Acids. In *Methods in Enzymology*; Vitamins and Coenzymes Part D; Academic Press, 1979; Vol. 62, pp 152–158. https://doi.org/10.1016/0076-6879(79)62212-7.

(8) Solmonson, A.; DeBerardinis, R. J. Lipoic Acid Metabolism and Mitochondrial Redox Regulation. *J. Biol. Chem.* **2018**, *293* (20), 7522–7530. https://doi.org/10.1074/jbc.TM117.000259.

(9) Gully, D.; Moinier, D.; Loiseau, L.; Bouveret, E. New Partners of Acyl Carrier Protein Detected in Escherichia Coli by Tandem Affinity Purification. *FEBS Lett.* **2003**, *548* (1–3), 90–96. https://doi.org/10.1016/s0014-5793(03)00746-4.

(10) Lay, N. R. D.; Cronan, J. E. In Vivo Functional Analyses of the Type II Acyl Carrier Proteins of Fatty Acid Biosynthesis. *J. Biol. Chem.* **2007**, *282* (28), 20319–20328. https://doi.org/10.1074/jbc.M703789200.

(11) Gully, D.; Bouveret, E. A Protein Network for Phospholipid Synthesis Uncovered by a Variant of the Tandem Affinity Purification Method in Escherichia Coli. *PROTEOMICS* **2006**, *6* (1), 282–293. https://doi.org/10.1002/pmic.200500115.

(12) Cronan, J. E.; Thomas, J. Bacterial Fatty Acid Synthesis and Its Relationships with Polyketide Synthetic Pathways. *Methods Enzymol.* **2009**, *459*, 395–433. https://doi.org/10.1016/S0076-6879(09)04617-5.

(13) Magnuson, K.; Jackowski, S.; Rock, C. O.; Cronan, J. E. Regulation of Fatty Acid Biosynthesis in Escherichia Coli. *Microbiol. Mol. Biol. Rev.* **1993**, *57* (3), 522–542.

(14) Shokri, A.; Larsson, G. Characterisation of the Escherichia Coli Membrane Structure and Function during Fedbatch Cultivation. *Microb. Cell Factories* **2004**, *3*, 9. https://doi.org/10.1186/1475-2859-3-9.

(15) Crosby, J.; Crump, M. P. The Structural Role of the Carrier Protein – Active Controller or Passive Carrier. *Nat. Prod. Rep.* **2012**, *29* (10), 1111–1137. https://doi.org/10.1039/C2NP20062G.

(16) Byers, D. M.; Gong, H. Acyl Carrier Protein: Structure–Function Relationships in a Conserved Multifunctional Protein Family. *Biochem. Cell Biol.* **2007**, *85* (6), 649–662. https://doi.org/10.1139/O07-109.

(17) 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. https://doi.org/10.1002/anie.201408576.

(18) Morris, T. W.; Reed, K. E.; Cronan, J. E. Lipoic Acid Metabolism in Escherichia Coli: The LplA and LipB Genes Define Redundant Pathways for Ligation of Lipoyl Groups to Apoprotein. *J. Bacteriol.* **1995**, *177* (1), 1–10. https://doi.org/10.1128/jb.177.1.1-10.1995.

(19) Zhao, X.; Miller, J. R.; Jiang, Y.; Marletta, M. A.; Cronan, J. E. Assembly of the Covalent Linkage between Lipoic Acid and Its Cognate Enzymes. *Chem. Biol.* **2003**, *10* (12), 1293–1302. https://doi.org/10.1016/j.chembiol.2003.11.016.

(20) Sztain, T.; Bartholow, T. G.; Lee, D. J.; McCammon, J. A.; Burkart, M. D. Decoding the Allosteric Nature of Acyl Carrier Proteins. *PNAS Rev.*

(21) Cronan, J. E. The Chain-Flipping Mechanism of ACP (Acyl Carrier Protein)-Dependent Enzymes Appears Universal. *Biochem. J.* **2014**, *460* (2), 157–163. https://doi.org/10.1042/BJ20140239.

(22) Ma, Q.; Zhao, X.; Eddine, A. N.; Geerlof, A.; Li, X.; Cronan, J. E.; Kaufmann, S. H. E.; Wilmanns, M. The Mycobacterium Tuberculosis LipB Enzyme Functions as a Cysteine/Lysine Dyad Acyltransferase. *Proc. Natl. Acad. Sci.* **2006**, *103* (23), 8662–8667. https://doi.org/10.1073/pnas.0510436103.

(23) Bartholow, T. G.; Sztain, T.; Lee, D. J.; Abagyan, R.; Burkart, M. D. Elucidation of Transient Protein-Protein Interactions within Carrier Protein-Dependent Biosynthesis. *Commun. Biol. Accept.*

(24) 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. *Chem. Biol.* **2010**, *17* (7), 776–785. https://doi.org/10.1016/j.chembiol.2010.05.024.

(25) 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. https://doi.org/10.1021/bi9014659.

(26) 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. https://doi.org/10.1016/j.jmb.2006.09.049.

(27) 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. https://doi.org/10.1016/j.chembiol.2015.09.009.

(28) 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. https://doi.org/10.1038/nature12810.

(29) 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. https://doi.org/10.1038/s41589-019-0301-y.

(30) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44–46. https://doi.org/10.1039/b512735a.

(31) Waudby, C. A.; Ramos, A.; Cabrita, L. D.; Christodoulou, J. Two-Dimensional NMR Lineshape Analysis. *Sci. Rep.* **2016**, *6*, 24826. https://doi.org/10.1038/srep24826.

(32) Neves, M. A. C.; Totrov, M.; Abagyan, R. Docking and Scoring with ICM: The Benchmarking Results and Strategies for Improvement. *J. Comput. Aided Mol. Des.* **2012**, *26* (6), 675–686. https://doi.org/10.1007/s10822-012-9547-0.

(33) ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation - Abagyan - 1994 - Journal of Computational Chemistry - Wiley Online Library https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540150503 (accessed Aug 31, 2020).

### 3.5 Experimental Methods

#### Protein Purification and production Protocol

LipB was grown through overexpression in *E. coli* BL21 (DE3), cells were grown in 1L of media with 50mg/L kanamycin. Growths were started through inoculation using a 5mL starter culture grown overnight. LipB was grown until it reached an OD600~ 0.6-0.8, then induced with 1mM IPTG and incubated overnight at 16°C. After growth pelleting was performed on a Beckman floor centrifuge in a JLA-8.1 rotor at 800 RCF. Pelleted cells were frozen and stored until the days prior to the titration for purification.

Labeled AcpP was grown from a pet-22b vector with a His-tag in *E. coli* BL21 (DE3) cells. In order to label the cells they were grown in $^{15}N$ supplemented M9 minimal media. 1g of $^{15}N$ $NH_4Cl$ and 8g of unlabeled glucose were added to 1L of M9 media. In order to achieve deuteration the media components were mixed in an oven dried glass graduated cylinder, followed by sterile filtration into an autoclaved and oven dried growth flask. Inoculating bacteria was carefully attenuated to the deuterated media, over the course of several growths. To begin BL21 cells were inoculated into a 25% $D_2O$/75% $H_2O$ unlabeled media, these were grown overnight at 37 °C. This growth was used to then inoculate another 50% $D_2O$/50% $H_2O$ media, which was grown overnight in the same conditions. This was then used to inoculate 75% $D_2O$/ 25% $H_2O$, which was grown and used to inoculate 90% $D_2O$ media. Finally the 90% $D_2O$ growth was used to inoculate a final starter with 100% $D_2O$ M9 media. This final 100% $D_2O$ starter was grown overnight at 37 °C and after confirming by eye that the media had become turbid with growth used to inoculate the labeled $D_2O$ M9 media. This was grown at 37 °C for ~16 hours until OD600=0.7. At this point 1mM IPTG was added for induction and the growth was left to grow for 4 hours at 37 °C. After induced growth the cells were spun down on a JLA-8.1 rotor at 800 RCF. Cells were spun for 1 hour and care was taken when harvesting cells to ensure there was no loss of material.

The $^{15}N$ ammonium chloride used in the labeled growth was purchased from Cambridge Isotopes laboratory. Deuterium oxide ($D_2O$) used in preparation of perdeuterated growth was purchased from Sigma Aldritch. All unlabeled proteins were grown on Luria broth from Teknova.

For purification, cells were re-suspended in 50mM Hepes (pH 7.4), 250 mM NaCl, and 10% glycerol. The lysed cells were spun at 10,000 RCF in a Beckman floor centrifuge equipped with a JA-20 rotor. Spun protein was checked for full clarification after 1 hour, after confirming pelleting of membrane and insoluble materials the protein was taken for purification. Clarified lysate was mixed with 2mL bed volume of Bio-Rad Ni-IMAC resin and left on a rotator in a 4 °C cold room to batch bind for 30 minutes. Washing was performed with 2 40 mL washes with 50mM HEPES (pH 7.4), 250mM NaCl, and 10% glycerol, the first wash was performed with just buffer and the second with an added 15mM imidazole. After washing elution was performed with 3 5mL volumes of wash buffer with an added 250mM imidazole. Bradford reagent was used to test the purification for protein and at the end of elution to confirm no further protein was eluted. The same purification protocol was used for the AcpP, but out of caution for losing valuable labeled protein the wash volume was lowered to 30 mL and 10mM imidazole. After purification proteins were checked by 12% SDS-PAGE to confirm successful purification. Elutions were dialyzed overnight into 50mM Tris-HCl (pH 7.4), 250 mM NaCl, 10% glycerol, and 1mM DTT. For the AcpP purification the second wash was also dialyzed but discarded once successful separation of the labeled AcpP was confirmed.

**ACP Chemoenzymatic loading**

After purification and dialysis, the AcpP was made uniformly apo by reaction with *Pseudomonas aeruginosa* ACPH, with an added 5mM $MgCl_2$ and 0.5mM $MnCl_2$. Reaction was performed overnight at 37 °C on a rotator. Apofication was confirmed by conformationally sensitive Urea-PAGE. After confirmation that the AcpP was fully apo chemoenzymatic labeling was carried out. The loading was performed using 3 *E. coli* biosynthetic enzymes CoaA, CoaD, and CoaE plus the *Bacillus subtilis* SFP. The reaction contained 12.5mM MgCl2, 10mM ATP, 0.1μM CoaA, 0.1μM CoaD, 0.1μM CoaE, 0.2μM Sfp, 0.02% Triton X, 0.01 % Azide, 0.1% TCEP, and 0.1mM acyl mimic probe.

**Purification and preparation for titrations**

Samples were purified by the same means as previously published. After dialysis of the LipB or one pot chemoenzymatic loading of the AcpP the samples were collected and concentrated to 2mL on Amicon

Ultra-15 spin concentrators. 3kDa and 10kDa columns were used for the AcpP and LipB respectively. After concentration AcpP and LipB were purified by size exclusion chromatography on a Superdex 75 column, 10mM potassium phosphate pH 7.4, 0.5mM TCEP, and 0.1% azide buffer was prepared and used to purify the AcpP and LipB for the experiments. In order to assure consistency, the same buffer was used for purifications and as buffer in the NMR experiments. For the first C8-AcpP titration the carrier protein and LipB were purified the day before the experiment on the FPLC. In order to assure stability of the partner protein the LipB was not concentrated until the morning of the experiment. The C8-AcpP was concentrated to 3.87 mg/mL and the LipB was concentrated to 6.1 mg/mL. These proteins were used to create a saturated NMR sample at 0.075 mM C8-AcpP and 0.113mM LipB. A zero-point AcpP sample was created with 0.075 mM C8-AcpP. In the case of C8-AcpP a 2.0 molar equivalents sample was prepared but had too poor signal to be useful. In the case of the C6-AcpP experiment the carrier protein was taken from the FPLC and concentrated to 4.1 mg/mL and the LipB was concentrated to 8.5 mg/mL, final concentrations were 0.105 mM C6-AcpP and 0.210 mM LipB in the saturated sample and 0.105 mM AcpP in the zero point sample. A 2 molar excess of partner protein was used to ensure full saturation in the non-substrate AcpP titration. In the C10-AcpP experiment the carrier protein was concentrated to a final concentration of 4.95 mg/mL and the LipB was concentrated to 7.78 mg/mL. The final concentrations were 0.1055mM C10-AcpP and 0.2112 mM LipB in the saturated sample and 0.1055 mM C10-AcpP in the zero point sample. Again the ratios were chosen to ensure full saturation at 2 molar equivalents. Approximately the same concentrations were chosen to make the experiment similar to the C6-AcpP.

**NMR Experiments**

All spectra collected in this experiment were collected at the UCSD Biomolecular NMR facility on their Bruker 800MHz spectrometer. Previous assignments were used for the C8-AcpP backbone HSQC assignments[1]. The C6 and C10-AcpP HSQCs were assigned based on the C8-AcpP, due to the small differences between the two spectra. Assigned peaks are available to view on the BMRB. Experiments were performed at 37 °C, titrations had a total of 5 titration points. The chemical shift perturbations were

quantified using the formula below with an α value of 0.2. This was in order to keep the data consistent with previous work in FAB.

$$CSP = \sqrt{\frac{1}{2}[\,\delta_H^2 + (\alpha \cdot \delta_N^2)]}$$

To perform the titrations two samples were prepared as described. A saturated sample and zero-point sample, buffers were prepared identically for both samples with only the presence of partner protein different between samples. All three sets of HSQC experiments were acquired with a 1.5 second recycle delay and 2048 data points. Between experiments samples were stored at 4°C to maintain stability, no denaturation of the labeled AcpP was seen in the spectra and no visible crashed protein was observed in any sample. Processing was performed in NMRPipe 10.9[2] and visualization was performed in NMRFAM-SPARKY 3.115[3]. After processing all figures displaying spectra were generated in Sparky, chemical shift perturbation calculations and figure generation was performed using the Matplotlib python utility[4].

**Titan analysis**

Further analysis of the titrations was performed using the TITAN lineshape analysis program. Peaks were selected by hand across the titration before performing an initial fitting of the data. Fitting parameters were first estimated at 10μM with a $k_{off}$ of 5000 $s^{-1}$, following fitting the peaks were hand checked. Peaks were examined to be sure there was no errors in the cases of peaks which migrated into one another or crowded regions of the spectra which were incorrectly fit. After this an initial jackknife error analysis was performed, this gave a rough picture of the error of the calculations. After a final hand check that no peaks were fitted incorrectly the final error analysis was performed. In each titration data set 300 steps of bootstrap error analysis were performed, this took approximately 18 hours for each data set. Calculations were performed by the same protocol as previously published on fatty acid biosynthesis, a set of matched simulated and real peaks are presented.

**Docking Method**

The LipB structure was prepared by homology modeling using the 2QHS *Thermus thermophilus* lipoyltransferase with ICM Homology.The *Mycobacterial* LipB 1W66 was also considered but 2QHS had a greater sequence homology. AcpPs used for docking were taken from previous MD simulations. Before docking the LipB structure was prepared by solvation and minimization. The ICM quickflood procedure was performed to generate a water box around the LipB. Following solvation the LipB was minimized to correctly orient the amino acid side chains for interaction with the AcpP. Optimization was performed on LipB by first running the ICM optimizeHbonds and optimize HisProAsnGlnCys protocols. Molecular dynamics derived AcpP structures were used for the ACPs. The acyl chain and phosphopantetheine were preserved during the calculation to best mimic the different chain lengths. All docking was performed using the ICM – Molsoft FFT protein protein docking algorithm.

Models of the LipB•AcpP interactions were chosen based on the most stable model under 10Å RMSD form the previously published docked model. This cutoff was chosen in order to give each chain length flexibility to adopt the most stable conformation. Over 10Å the docked conformation was so far from the active site that there was no chance for the conformation to be an active one. The chosen complexes were visualized against one another when comparing the most stable conformations. A second analysis was performed by using the stable conformation seen for the C6, C8, and C10-AcpP•LipB complex as a reference. This data set of poses with deviation from the previous model was used to map the LipB surface. Graphing of the energetics was done in Matplotlib, while visualizations were performed in Pymol[5].

### 3.6 Supplemental Information

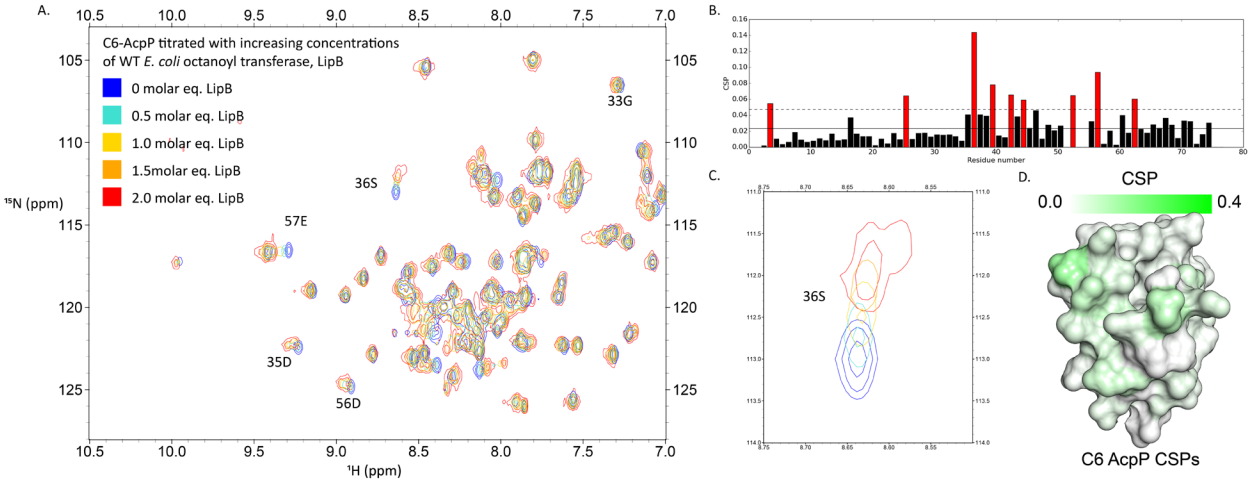



Figure S1 Titration of hexanoyl -AcpP with the *E. coli* LipB octanoyltransferase. 5 1H-15N HSQC spectra were overlayed of the C6 AcpP interacting with increasing molar ratios of the octanoyltransferase, LipB. The titration occurs in fast exchange, with the bound and unbound state interchanging between bound and unbound rapidly and resolving as a single peak on the spectra. A) The total NMR spectra with a selection of individual peaks highlighted. B) The chemical shift perturbations of each residue in the titration. One standard deviation above the mean is colored red to highlight the most perturbed residues. C) A focus on the important serine 36 of AcpP, it should be noted the difference between this shift in the C6-AcpP titration and the other chain lengths. D) The surface of the AcpP with the CSPs colored by magnitude.

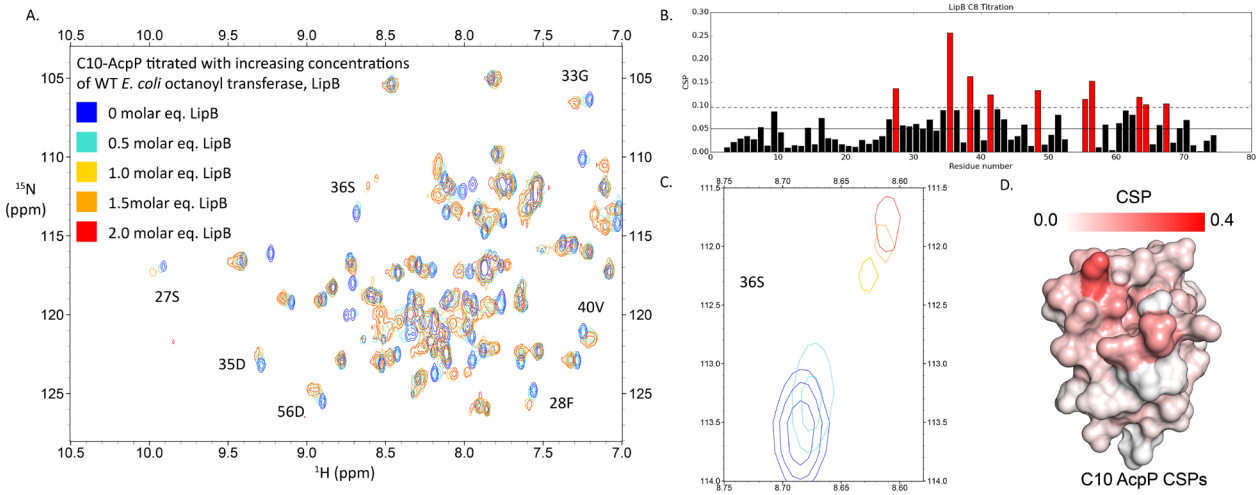



Figure S2 Titration of decanoyl -AcpP with the *E. coli* LipB octanoyltransferase. 5 1H-15N HSQC spectra were overlayed of the C10 AcpP interacting with increasing molar ratios of the octanoyltransferase, LipB. The titration occurs in fast exchange, with the bound and unbound state interchanging between bound and unbound rapidly and resolving as a single peak on the spectra. A) The total NMR spectra with a selection of individual peaks highlighted. B) The chemical shift perturbations of each residue in the titration. One standard deviation above the mean is colored red to highlight the most perturbed residues. C) A focus on the important serine 36 of AcpP, it should be noted the difference between this shift in the C10-AcpP titration and the other chain lengths. D) The surface of the AcpP with the CSPs colored by magnitude.

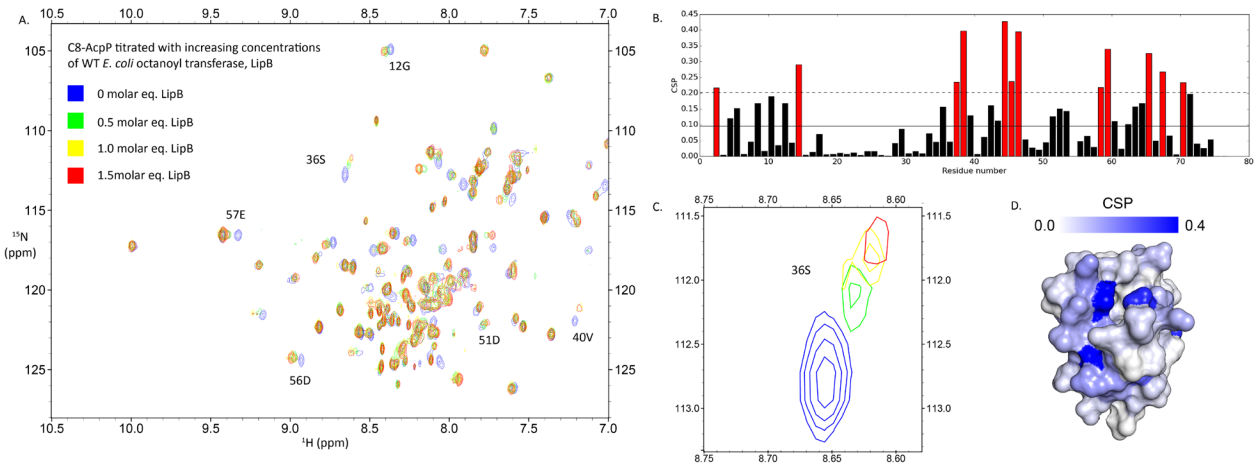



Figure S3 Titration of octanoyl -AcpP with the *E. coli* LipB octanoyltransferase. 4 1H-15N HSQC spectra were overlayed of the C8 AcpP interacting with increasing molar ratios of the octanoyltransferase, LipB. A fifth titration point was prepared but the signal was too weak to yield any useful data. The titration occurs in fast exchange, with the bound and unbound state interchanging between bound and unbound rapidly and resolving as a single peak on the spectra. A) The total NMR spectra with a selection of individual peaks highlighted. B) The chemical shift perturbations of each residue in the titration. One standard deviation above the mean is colored red to highlight the most perturbed residues. C) A focus on the important serine 36 of AcpP, it should be noted the difference between this shift in the C8-AcpP titration and the other chain lengths. D) The surface of the AcpP with the CSPs colored by magnitude.

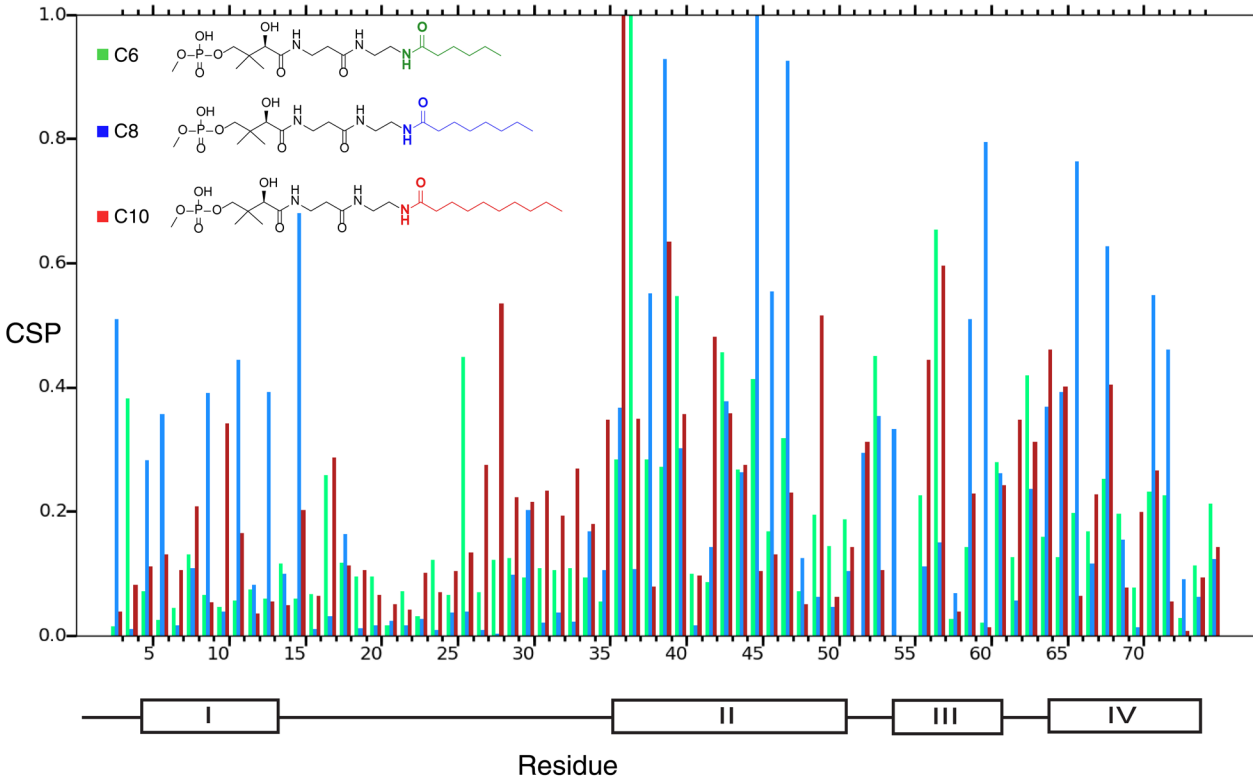



Figure S4. Normalized chemical shift perturbations of three AcpPs interacting with LipB. The perturbations are normalized within their own data set, setting the largest CSP at 1.0.

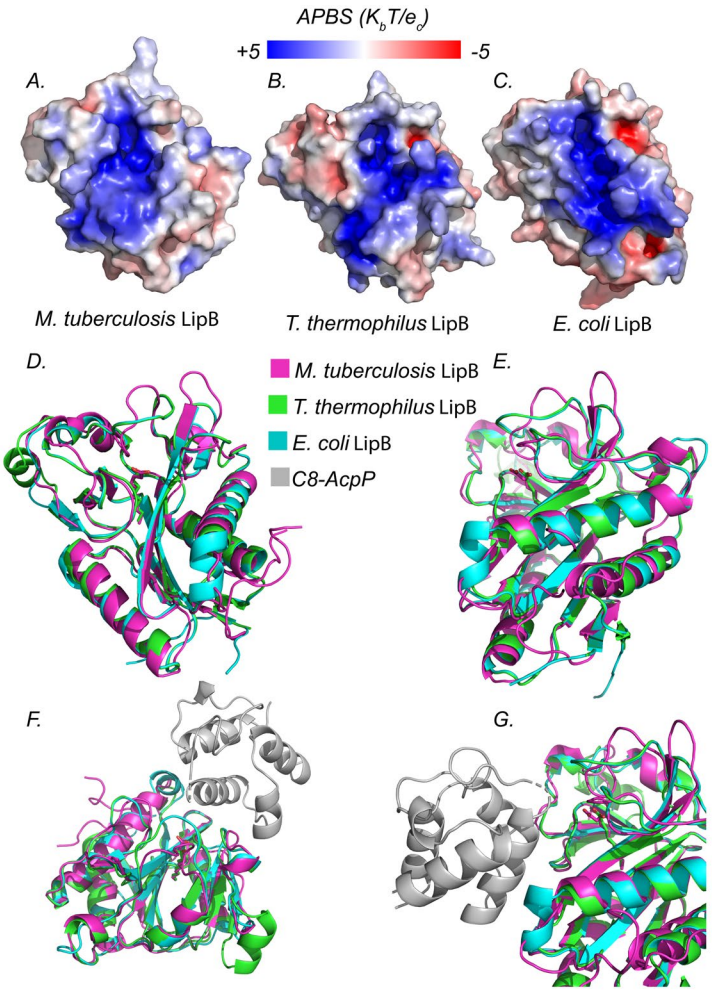



Figure S5 Comparison of the E. coli LipB model to the *M. tuberculosis* and *T. thermophilus* LipB. A) The *M. tuberculosis* LipB is shown with APBS coloring generated in Pymol. B) The *T. thermophilus* LipB shown with APBS coloring. C) The *E. coli* LipB model generated in this work shown with APBS coloring. It is interesting to not the similarities of the surfaces and electrostatics of the different LipBs. D&E) Overlays of the three LipBs, showing the similarities in structure between the species. The majority of the proteins overlay quite well, with only some loop regions showing large variations. The *T. thermophilus* LipB has a 51% similarity to E. coli and *M. tuberculosis* has a 53% similarity. F&G) The LipBs overlaid with an ACP to give context of the regions of the LipBs which are more different. It is promising that the AcpP binding surface appears to show very little variation, with the dissimilar loop beyond the binding site and active site.

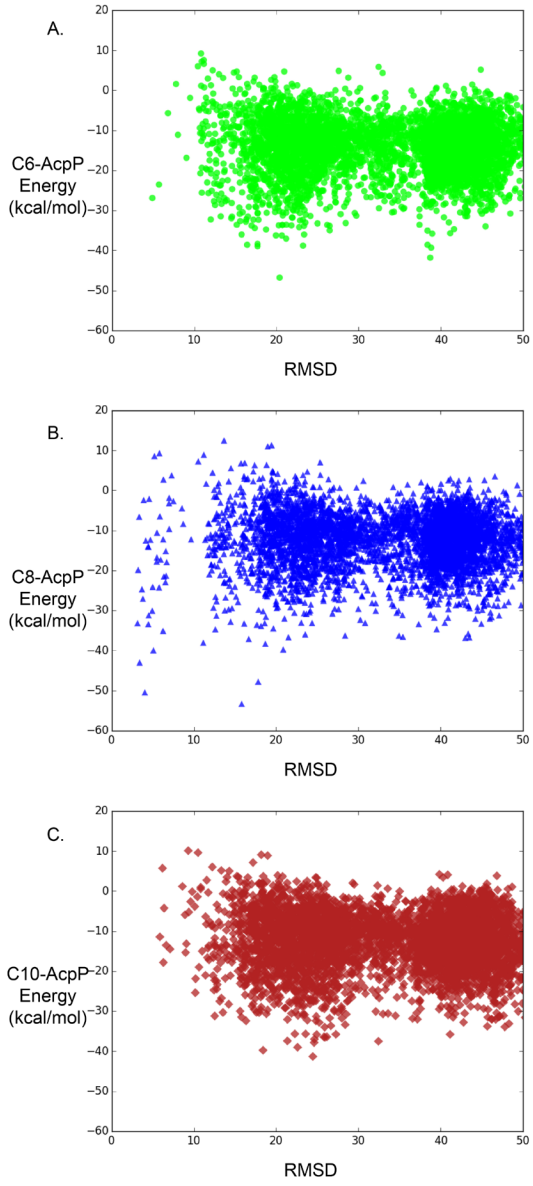



Figure S6 Docking details of the C6, C8, and C10-AcpP with the *E. coli* octanoyltransferase, LipB. The docked states of the AcpPs with LipB are shown with greater detail, for each chain length docked the full 50 angstrom RMSD surface is shown. A) The C6-AcpP docking to LipB RMSD vs energy plot. The RMSD is based on the previously published model as described in the methods. B) The C8-AcpP docking to LipB RMSD vs energy plot. The RMSD is based on the previously published model as described in the methods. C) The C10-AcpP docking to LipB RMSD vs energy plot. The RMSD is based on the previously published model as described in the methods.

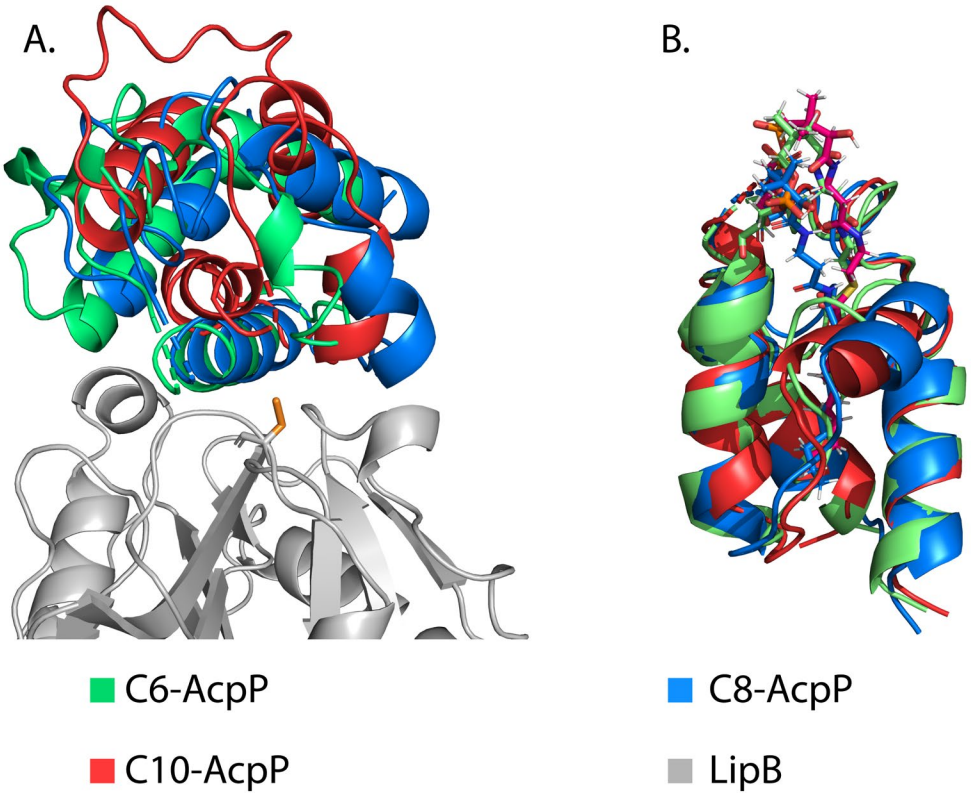



Figure S7 Comparison of the different chain lengths of AcpP. A) the most stable low RMSD (less than 5Å) state of the AcpP•LipB binding with C6, C8, and C10-AcpP. The complexes of C6 and C10 are significantly less stable than the C8-AcpP•LipB complex. B) The structures of the MD derived C6, C8, and C10-AcpP. The acyl chains were present during the simulations but in other figures they are not shown, as most of the chain is sequestered and it makes viewing the protein structures more difficult.

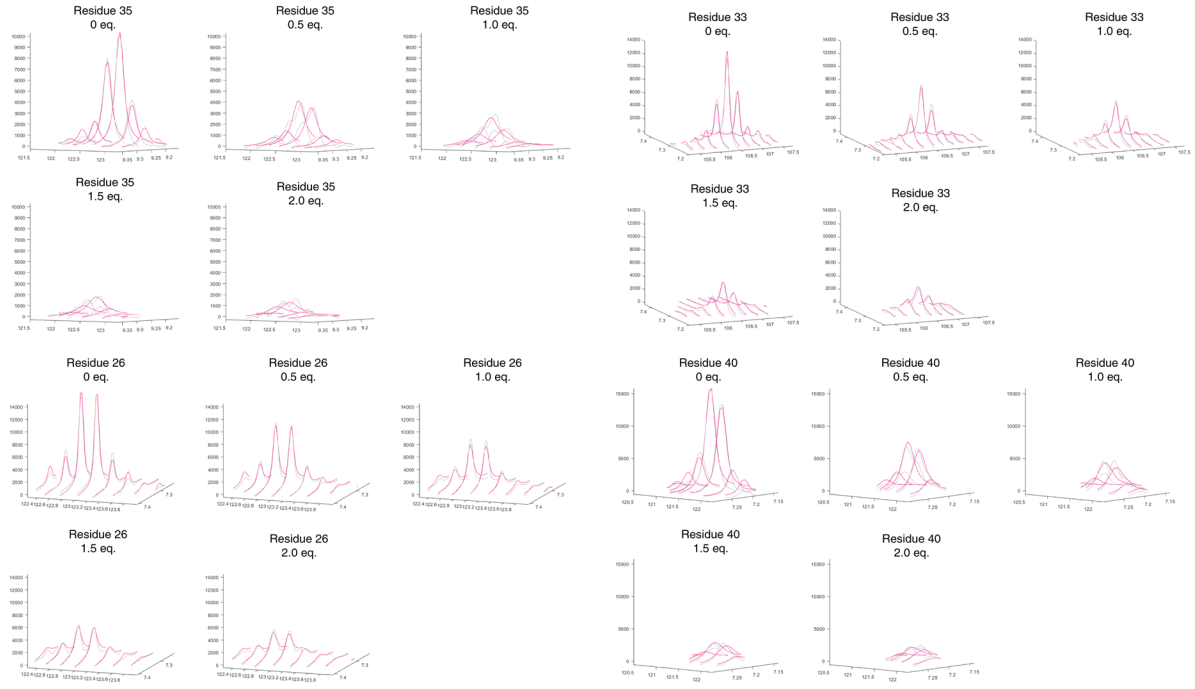



Figure S8 TITAN analysis of the C6-AcpP LipB titration. Real (red) and simulated (blue) titration peaks are shown for four selected residues of the TITAN analysis. The analysis was performed using the flexible docking method, allowing flexibility in the stoichiometry. The error was analyzed using 300 seps of bootstrap error analysis. Though there is significant signal loss in the real data, the peaks overlay well demonstrating a well fit model.

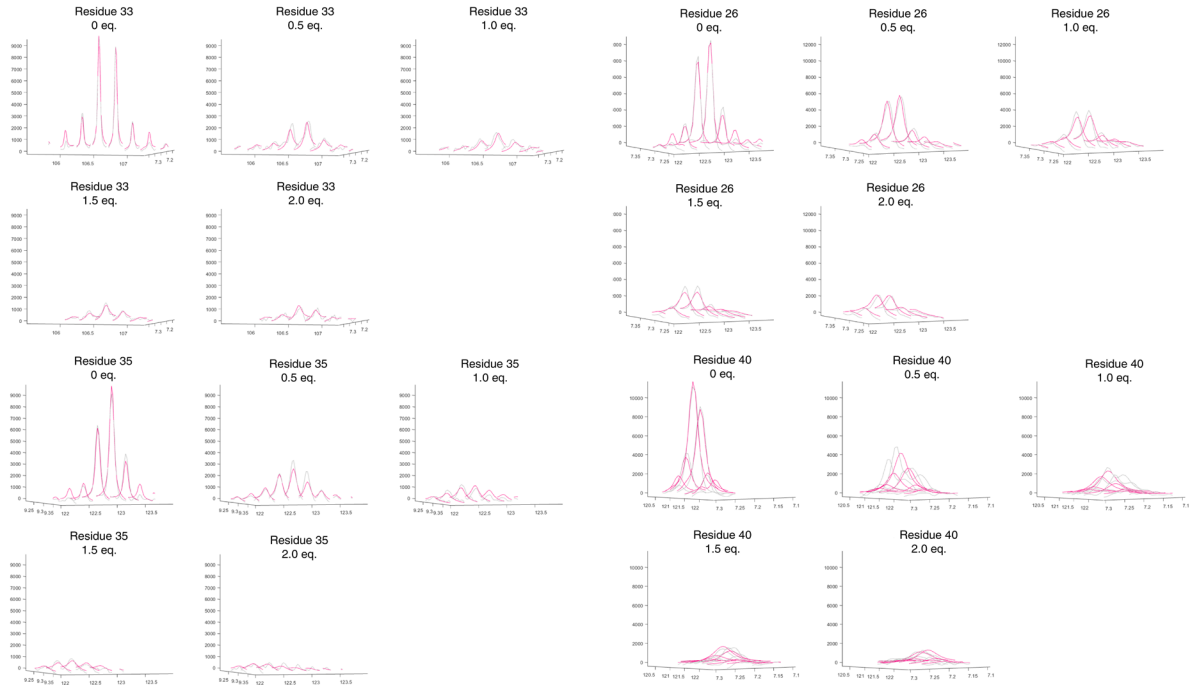



Figure S9 TITAN analysis of the C10-AcpP LipB titration. Real (red) and simulated (blue) titration peaks are shown for four selected residues of the TITAN analysis. The analysis was performed using the flexible docking method, allowing flexibility in the stoichiometry. The error was analyzed using 300 seps of bootstrap error analysis. Though there is significant signal loss in the real data, the peaks overlay well demonstrating a well fit model.

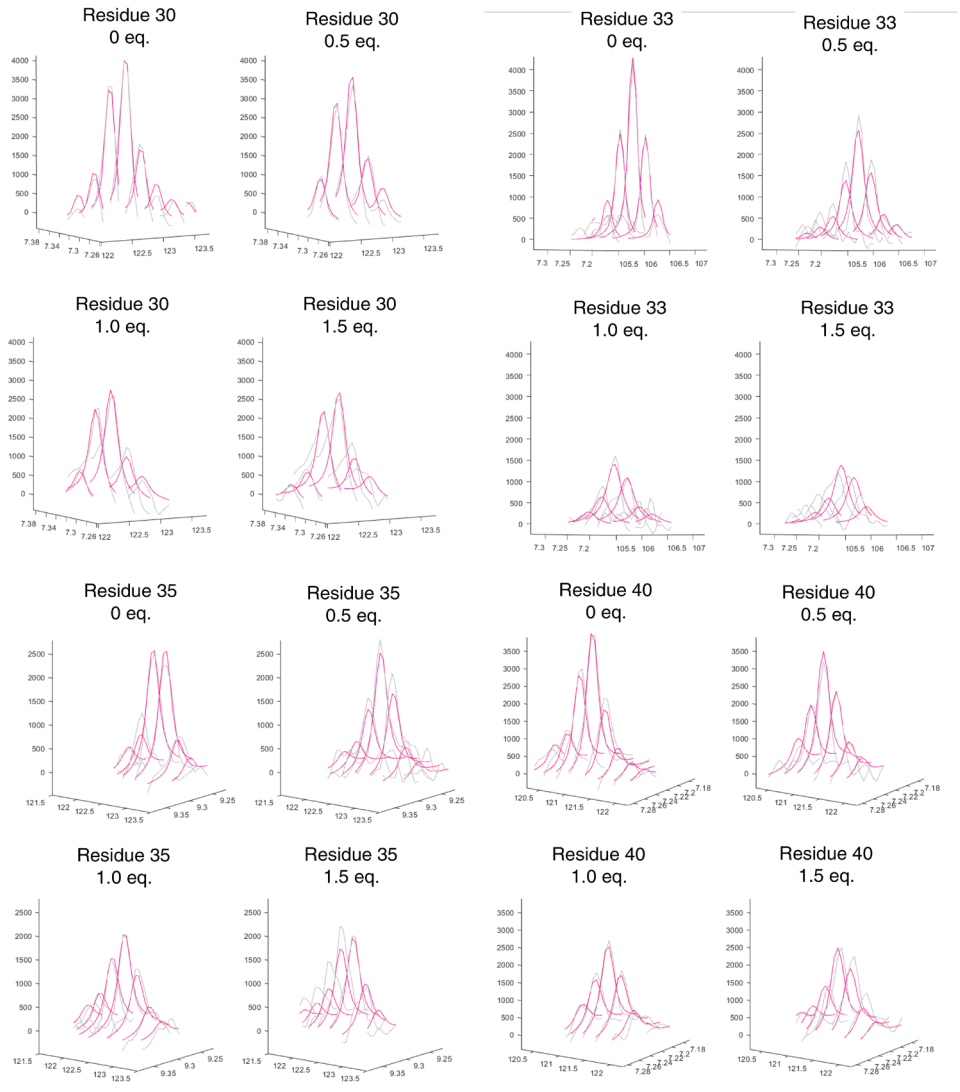



Figure S10 TITAN analysis of the C8-AcpP LipB titration. Real (red) and simulated (blue) titration peaks are shown for four selected residues of the TITAN analysis. The analysis was performed using the flexible docking method, allowing flexibility in the stoichiometry. The error was analyzed using 300 seps of bootstrap error analysis. Though there is significant signal loss in the real data, the peaks overlay well demonstrating a well fit model.

Table S1 C6-AcpP•LipB titration chemical shifts. The data have also been submitted to the BMRB for wide access.

| Residue Number | Residue | Nucleii | ZP chemical shift | Saturated chemical shift |
|---|---|---|---|---|
| **3** | I | H | 8.644 | 8.647 |
| **3** | I | N | 121.548 | 121.545 |
| **4** | E | H | 8.579 | 8.508 |
| **4** | E | N | 119.443 | 119.286 |
| **5** | E | H | 7.863 | 7.865 |
| **5** | E | N | 116.97 | 116.898 |
| **6** | R | H | 8.342 | 8.347 |
| **6** | R | N | 119.976 | 119.98 |
| **7** | V | H | 8.942 | 8.942 |
| **7** | V | N | 119.299 | 119.254 |
| **8** | K | H | 8.213 | 8.239 |
| **8** | K | N | 117.216 | 117.24 |
| **9** | K | H | 8.271 | 8.258 |
| **9** | K | N | 120.487 | 120.498 |
| **10** | I | H | 7.647 | 7.654 |
| **10** | I | N | 119.368 | 119.398 |
| **11** | I | H | 8.296 | 8.287 |
| **11** | I | N | 119.004 | 119.039 |
| **12** | G | H | 8.451 | 8.457 |
| **12** | G | N | 105.381 | 105.45 |
| **13** | E | H | 8.196 | 8.208 |
| **13** | E | N | 120.182 | 120.179 |
| **14** | Q | H | 8.414 | 8.426 |
| **14** | Q | N | 117.571 | 117.469 |
| **15** | L | H | 8.043 | 8.031 |
| **15** | L | N | 113.265 | 113.273 |
| **16** | G | H | 7.79 | 7.785 |
| **16** | G | N | 109.86 | 109.797 |
| **17** | V | H | 7.867 | 7.849 |
| **17** | V | N | 114.524 | 114.277 |
| **18** | K | H | 8.524 | 8.545 |
| **18** | K | N | 123.077 | 123.02 |
| **19** | Q | H | 8.776 | 8.777 |
| **19** | Q | N | 122.896 | 122.799 |
| **20** | E | H | 9.415 | 9.418 |
| **20** | E | N | 116.679 | 116.685 |
| **21** | E | H | 7.863 | 7.865 |
| **21** | E | N | 116.97 | 116.898 |
| **22** | V | H | 7.527 | 7.532 |
| **22** | V | N | 122.285 | 122.305 |
| **23** | T | H | 7.318 | 7.295 |
| **23** | T | N | 115.47 | 115.425 |
| **24** | N | H | 8.605 | 8.607 |
| **24** | N | N | 118.763 | 118.829 |
| **25** | N | H | 8.092 | 8.105 |
| **25** | N | N | 111.857 | 112.309 |
| **26** | A | H | 7.308 | 7.322 |
| **26** | A | N | 122.872 | 122.865 |
| **27** | S | H | 9.952 | 9.972 |
| **27** | S | N | 117.225 | 117.298 |

| | | | | |
|---|---|---|---|---|
| **28** | F | H | 7.56 | 7.559 |
| **28** | F | N | 125.623 | 125.75 |
| **29** | V | H | 8.729 | 8.727 |
| **29** | V | N | 116.847 | 116.941 |
| **30** | E | H | 8.299 | 8.32 |
| **30** | E | N | 116.744 | 116.709 |
| **31** | D | H | 7.784 | 7.8 |
| **31** | D | N | 113.742 | 113.671 |
| **32** | L | H | 7.347 | 7.368 |
| **32** | L | N | 115.69 | 115.725 |
| **33** | G | H | 7.273 | 7.292 |
| **33** | G | N | 106.513 | 106.505 |
| **34** | A | H | 8.46 | 8.471 |
| **34** | A | N | 122.803 | 122.809 |
| **35** | D | H | 9.23 | 9.254 |
| **35** | D | N | 122.441 | 122.178 |
| **36** | S | H | 8.638 | 8.626 |
| **36** | S | N | 113.001 | 111.985 |
| **37** | L | H | 8.131 | 8.08 |
| **37** | L | N | 123.803 | 123.667 |
| **38** | D | H | 8.304 | 8.347 |
| **38** | D | N | 119.806 | 119.98 |
| **39** | T | H | 8.132 | 8.177 |
| **39** | T | N | 111.942 | 111.434 |
| **40** | V | H | 7.221 | 7.215 |
| **40** | V | N | 121.58 | 121.677 |
| **41** | E | H | 7.84 | 7.857 |
| **41** | E | N | 119.312 | 119.33 |
| **42** | L | H | 8.387 | 8.471 |
| **42** | L | N | 121.529 | 121.332 |
| **43** | V | H | 8.006 | 8.029 |
| **43** | V | N | 119.194 | 119.44 |
| **44** | M | H | 7.755 | 7.733 |
| **44** | M | N | 117.244 | 116.839 |
| **45** | A | H | 8.149 | 8.147 |
| **45** | A | N | 121.461 | 121.291 |
| **46** | L | H | 8.384 | 8.429 |
| **46** | L | N | 120.025 | 120.258 |
| **47** | E | H | 0 | 0 |
| **47** | E | N | 0 | 0 |
| **48** | E | H | 7.863 | 7.865 |
| **48** | E | N | 116.97 | 116.898 |
| **49** | E | H | 7.941 | 7.961 |
| **49** | E | N | 119.504 | 119.675 |
| **50** | F | H | 7.762 | 7.768 |
| **50** | F | N | 111.603 | 111.746 |
| **51** | D | H | 7.874 | 7.836 |
| **51** | D | N | 122.146 | 122.157 |
| **52** | T | H | 8.017 | 8.105 |
| **52** | T | N | 112.328 | 112.309 |
| **53** | E | H | 8.12 | 8.147 |
| **53** | E | N | 121.728 | 121.291 |
| **54** | I | H | 10.368 | 10.368 |
| **54** | I | N | 128.931 | 128.931 |
| **56** | D | H | 8.909 | 8.934 |
| **56** | D | N | 124.839 | 124.646 |

| | | | | |
|---|---|---|---|---|
| **57** | E | H | 9.288 | 9.418 |
| **57** | E | N | 116.543 | 116.685 |
| **58** | E | H | 7.225 | 7.23 |
| **58** | E | N | 116.017 | 116.027 |
| **59** | A | H | 8.129 | 8.133 |
| **59** | A | N | 122.672 | 122.529 |
| **60** | E | H | 7.538 | 7.542 |
| **60** | E | N | 111.793 | 111.797 |
| **61** | K | H | 7.059 | 7.096 |
| **61** | K | N | 113.994 | 114.209 |
| **62** | I | H | 7.613 | 7.625 |
| **62** | I | N | 122.345 | 122.232 |
| **63** | T | H | 8.02 | 8.105 |
| **63** | T | N | 112.268 | 112.309 |
| **64** | T | H | 7.147 | 7.119 |
| **64** | T | N | 110.48 | 110.56 |
| **65** | V | H | 8.021 | 8.021 |
| **65** | V | N | 121.338 | 121.21 |
| **66** | Q | H | 8.57 | 8.545 |
| **66** | Q | N | 117.941 | 117.783 |
| **67** | A | H | 7.772 | 7.806 |
| **67** | A | N | 119.54 | 119.543 |
| **68** | A | H | 7.881 | 7.876 |
| **68** | A | N | 122.115 | 121.86 |
| **69** | I | H | 8.066 | 8.065 |
| **69** | I | N | 119.026 | 119.225 |
| **70** | D | H | 9.142 | 9.157 |
| **70** | D | N | 118.998 | 118.975 |
| **71** | Y | H | 8.21 | 8.257 |
| **71** | Y | N | 122.125 | 122.131 |
| **72** | I | H | 8.099 | 8.07 |
| **72** | I | N | 120.647 | 120.468 |
| **73** | N | H | 8.833 | 8.835 |
| **73** | N | N | 118.23 | 118.256 |
| **74** | G | H | 7.806 | 7.803 |
| **74** | G | N | 104.921 | 105.034 |
| **75** | H | H | 7.628 | 7.635 |
| **75** | H | N | 118.47 | 118.683 |

Table S2 C8-AcpP•LipB titration chemical shifts. The data have also been submitted to the BMRB for wide access.

| RESIDUE NUMBER | RESIDUE | NUCLEII | ZP CHEMICAL SHIFT | SATURATED CHEMICAL SHIFT |
|---|---|---|---|---|
| **3** | I | H | 8.561 | 8.687 |
| **3** | I | N | 122.666 | 121.263 |
| **4** | E | H | 8.661 | 8.655 |
| **4** | E | N | 118.452 | 118.452 |
| **5** | E | H | 7.84 | 7.84 |
| **5** | E | N | 117.526 | 116.674 |
| **6** | R | H | 8.437 | 8.34 |
| **6** | R | N | 121.248 | 120.288 |
| **7** | V | H | 8.979 | 8.971 |
| **7** | V | N | 119.3 | 119.269 |
| **8** | K | H | 8.228 | 8.199 |
| **8** | K | N | 116.872 | 117.166 |
| **9** | K | H | 8.337 | 8.247 |
| **9** | K | N | 121.898 | 120.805 |
| **10** | I | H | 7.624 | 7.602 |
| **10** | I | N | 118.974 | 118.937 |
| **11** | I | H | 8.388 | 8.259 |
| **11** | I | N | 119.968 | 118.791 |
| **12** | G | H | 8.369 | 8.408 |
| **12** | G | N | 104.892 | 105.041 |
| **13** | E | H | 8.248 | 8.158 |
| **13** | E | N | 120.95 | 119.853 |
| **14** | Q | H | 8.474 | 8.426 |
| **14** | Q | N | 117.519 | 117.339 |
| **15** | L | H | 8.105 | 7.982 |
| **15** | L | N | 114.809 | 112.847 |
| **16** | G | H | 7.718 | 7.724 |
| **16** | G | N | 109.869 | 109.871 |
| **17** | V | H | 7.863 | 7.853 |
| **17** | V | N | 114.035 | 113.954 |
| **18** | K | H | 8.488 | 8.567 |
| **18** | K | N | 122.864 | 122.57 |
| **19** | Q | H | 8.814 | 8.818 |
| **19** | Q | N | 122.342 | 122.311 |
| **20** | E | H | 9.427 | 9.428 |
| **20** | E | N | 116.615 | 116.566 |
| **21** | E | H | 7.87 | 7.856 |
| **21** | E | N | 116.696 | 116.717 |
| **22** | V | H | 7.531 | 7.532 |
| **22** | V | N | 122.302 | 122.252 |
| **23** | T | H | 7.39 | 7.401 |
| **23** | T | N | 115.44 | 115.501 |
| **24** | N | H | 8.604 | 8.606 |
| **24** | N | N | 118.551 | 118.576 |
| **25** | N | H | 8.104 | 8.119 |
| **25** | N | N | 111.373 | 111.455 |
| **26** | A | H | 7.352 | 7.358 |
| **26** | A | N | 122.885 | 122.775 |

| | | | | |
|---|---|---|---|---|
| **27** | S | H | 9.991 | 9.995 |
| **27** | S | N | 117.272 | 117.256 |
| **28** | F | H | 7.6 | 7.599 |
| **28** | F | N | 126.18 | 126.188 |
| **29** | V | H | 8.729 | 8.781 |
| **29** | V | N | 117.009 | 117.149 |
| **30** | E | H | 8.333 | 8.267 |
| **30** | E | N | 116.453 | 116.966 |
| **31** | D | H | 7.84 | 7.849 |
| **31** | D | N | 113.204 | 113.163 |
| **32** | L | H | 7.246 | 7.227 |
| **32** | L | N | 115.311 | 115.249 |
| **33** | G | H | 7.366 | 7.379 |
| **33** | G | N | 106.657 | 106.644 |
| **34** | A | H | 8.44 | 8.504 |
| **34** | A | N | 122.307 | 122.699 |
| **35** | D | H | 9.17 | 9.231 |
| **35** | D | N | 121.557 | 121.465 |
| **36** | S | H | 8.658 | 8.618 |
| **36** | S | N | 112.749 | 111.658 |
| **37** | L | H | 8.056 | 8.116 |
| **37** | L | N | 124.477 | 124.358 |
| **38** | D | H | 8.386 | 8.516 |
| **38** | D | N | 121.258 | 119.726 |
| **39** | T | H | 8.459 | 8.056 |
| **39** | T | N | 109.409 | 111.361 |
| **40** | V | H | 7.21 | 7.173 |
| **40** | V | N | 121.974 | 121.081 |
| **41** | E | H | 7.893 | 7.903 |
| **41** | E | N | 119.01 | 119.005 |
| **42** | L | H | 8.409 | 8.324 |
| **42** | L | N | 121.994 | 121.94 |
| **43** | V | H | 8.021 | 7.936 |
| **43** | V | N | 118.561 | 119.619 |
| **44** | M | H | 7.722 | 7.736 |
| **44** | M | N | 117.373 | 116.584 |
| **45** | A | H | 8.205 | 8.056 |
| **45** | A | N | 123.079 | 120.15 |
| **46** | L | H | 8.49 | 8.328 |
| **46** | L | N | 121.474 | 120.008 |
| **47** | E | H | 8.69 | 8.445 |
| **47** | E | N | 121.309 | 118.796 |
| **48** | E | H | 7.841 | 7.855 |
| **48** | E | N | 116.903 | 117.271 |
| **49** | E | H | 7.975 | 7.942 |
| **49** | E | N | 119.841 | 119.929 |
| **50** | F | H | 7.767 | 7.752 |
| **50** | F | N | 111.45 | 111.332 |
| **51** | D | H | 7.79 | 7.769 |
| **51** | D | N | 122.406 | 122.112 |
| **52** | T | H | 8.064 | 8.188 |
| **52** | T | N | 111.823 | 112.459 |
| **53** | E | H | 8.099 | 8.109 |
| **53** | E | N | 123.627 | 122.559 |
| **54** | I | H | 10.306 | 10.483 |
| **54** | I | N | 129.1 | 129.579 |

| | | | | |
|---|---|---|---|---|
| **56** | D | H | 8.927 | 8.974 |
| **56** | D | N | 124.477 | 124.236 |
| **57** | E | H | 9.327 | 9.415 |
| **57** | E | N | 116.569 | 116.455 |
| **58** | E | H | 7.22 | 7.195 |
| **58** | E | N | 115.923 | 115.758 |
| **59** | A | H | 8.119 | 8.082 |
| **59** | A | N | 124.257 | 122.727 |
| **60** | E | H | 7.476 | 7.635 |
| **60** | E | N | 111.355 | 113.623 |
| **61** | K | H | 6.95 | 7.082 |
| **61** | K | N | 113.709 | 114.143 |
| **62** | I | H | 7.592 | 7.583 |
| **62** | I | N | 121.943 | 121.779 |
| **63** | T | H | 8.07 | 8.069 |
| **63** | T | N | 112.103 | 111.389 |
| **64** | T | H | 7.21 | 7.008 |
| **64** | T | N | 110.371 | 110.841 |
| **65** | V | H | 7.822 | 8.015 |
| **65** | V | N | 121.291 | 120.602 |
| **66** | Q | H | 8.707 | 8.352 |
| **66** | Q | N | 117.966 | 116.488 |
| **67** | A | H | 7.892 | 7.848 |
| **67** | A | N | 119.649 | 119.92 |
| **68** | A | H | 7.971 | 7.809 |
| **68** | A | N | 122.795 | 121.082 |
| **69** | I | H | 8.143 | 8.108 |
| **69** | I | N | 119.712 | 119.282 |
| **70** | D | H | 9.191 | 9.197 |
| **70** | D | N | 118.43 | 118.403 |
| **71** | Y | H | 8.196 | 8.111 |
| **71** | Y | N | 122.498 | 120.896 |
| **72** | I | H | 8.26 | 8.151 |
| **72** | I | N | 122.164 | 120.882 |
| **73** | N | H | 8.811 | 8.842 |
| **73** | N | N | 117.741 | 117.968 |
| **74** | G | H | 7.797 | 7.778 |
| **74** | G | N | 104.838 | 104.999 |
| **75** | H | H | 7.583 | 7.598 |
| **75** | H | N | 118.17 | 118.535 |

Table S3 C10-AcpP•LipB titration chemical shifts. The data have also been submitted to the BMRB for wide access.

| Residue Number | Residue | Nucleii | ZP chemical shift | Saturated chemical shift |
|---|---|---|---|---|
| **3** | I | H | 8.638 | 8.65 |
| **3** | I | N | 121.531 | 121.495 |
| **4** | E | H | 8.597 | 8.611 |
| **4** | E | N | 118.707 | 118.837 |
| **5** | E | H | 7.841 | 7.881 |
| **5** | E | N | 117.486 | 117.486 |
| **6** | R | H | 8.344 | 8.346 |
| **6** | R | N | 119.755 | 119.991 |
| **7** | V | H | 8.893 | 8.923 |
| **7** | V | N | 119.026 | 119.141 |
| **8** | K | H | 8.203 | 8.203 |
| **8** | K | N | 117.172 | 117.547 |
| **9** | K | H | 8.268 | 8.249 |
| **9** | K | N | 120.456 | 120.478 |
| **10** | I | H | 7.626 | 7.643 |
| **10** | I | N | 119.423 | 118.813 |
| **11** | I | H | 8.334 | 8.286 |
| **11** | I | N | 118.94 | 119.115 |
| **12** | G | H | 8.464 | 8.466 |
| **12** | G | N | 105.384 | 105.447 |
| **13** | E | H | 8.187 | 8.204 |
| **13** | E | N | 120.25 | 120.197 |
| **14** | Q | H | 8.419 | 8.414 |
| **14** | Q | N | 117.339 | 117.425 |
| **15** | L | H | 8.113 | 8.051 |
| **15** | L | N | 113.507 | 113.315 |
| **16** | G | H | 7.803 | 7.78 |
| **16** | G | N | 109.807 | 109.817 |
| **17** | V | H | 7.87 | 7.845 |
| **17** | V | N | 114.711 | 114.21 |
| **18** | K | H | 8.524 | 8.545 |
| **18** | K | N | 123.155 | 122.979 |
| **19** | Q | H | 8.776 | 8.779 |
| **19** | Q | N | 122.993 | 122.803 |
| **20** | E | H | 9.403 | 9.415 |
| **20** | E | N | 116.551 | 116.653 |
| **21** | E | H | 7.872 | 7.866 |
| **21** | E | N | 116.999 | 116.912 |
| **22** | V | H | 7.52 | 7.534 |
| **22** | V | N | 122.311 | 122.284 |
| **23** | T | H | 7.303 | 7.326 |
| **23** | T | N | 115.591 | 115.451 |
| **24** | N | H | 8.575 | 8.579 |
| **24** | N | N | 118.803 | 118.926 |
| **25** | N | H | 8.078 | 8.106 |
| **25** | N | N | 112.013 | 111.888 |
| **26** | A | H | 7.281 | 7.327 |
| **26** | A | N | 122.921 | 122.849 |

| | | | | |
|---|---|---|---|---|
| **27** | S | H | 9.91 | 9.981 |
| **27** | S | N | 116.949 | 117.295 |
| **28** | F | H | 7.558 | 7.589 |
| **28** | F | N | 124.796 | 125.749 |
| **29** | V | H | 8.726 | 8.712 |
| **29** | V | N | 116.527 | 116.923 |
| **30** | E | H | 8.289 | 8.258 |
| **30** | E | N | 116.879 | 117.235 |
| **31** | D | H | 7.752 | 7.803 |
| **31** | D | N | 114.035 | 113.699 |
| **32** | L | H | 7.372 | 7.377 |
| **32** | L | N | 115.404 | 115.751 |
| **33** | G | H | 7.204 | 7.286 |
| **33** | G | N | 106.348 | 106.61 |
| **34** | A | H | 8.425 | 8.469 |
| **34** | A | N | 122.519 | 122.756 |
| **35** | D | H | 9.289 | 9.316 |
| **35** | D | N | 123.209 | 122.595 |
| **36** | S | H | 8.684 | 8.61 |
| **36** | S | N | 113.57 | 111.802 |
| **37** | L | H | 8.188 | 8.062 |
| **37** | L | N | 123.759 | 123.736 |
| **38** | D | H | 8.281 | 8.307 |
| **38** | D | N | 119.788 | 119.848 |
| **39** | T | H | 8.124 | 8.162 |
| **39** | T | N | 111.473 | 110.342 |
| **40** | V | H | 7.244 | 7.211 |
| **40** | V | N | 121.068 | 121.69 |
| **41** | E | H | 7.764 | 7.77 |
| **41** | E | N | 119.34 | 119.512 |
| **42** | L | H | 8.325 | 8.499 |
| **42** | L | N | 121.223 | 121.227 |
| **43** | V | H | 7.99 | 7.975 |
| **43** | V | N | 118.906 | 119.548 |
| **44** | M | H | 7.767 | 7.866 |
| **44** | M | N | 116.915 | 116.912 |
| **45** | A | H | 8.114 | 8.151 |
| **45** | A | N | 121.29 | 121.314 |
| **46** | L | H | 8.344 | 8.346 |
| **46** | L | N | 119.755 | 119.991 |
| **47** | E | H | 8.708 | 8.637 |
| **47** | E | N | 119.982 | 120.199 |
| **48** | E | H | 7.872 | 7.866 |
| **48** | E | N | 116.999 | 116.912 |
| **49** | E | H | 7.941 | 7.93 |
| **49** | E | N | 119.209 | 120.139 |
| **50** | F | H | 7.76 | 7.762 |
| **50** | F | N | 111.588 | 111.699 |
| **51** | D | H | 7.88 | 7.878 |
| **51** | D | N | 122.305 | 122.047 |
| **52** | T | H | 8.011 | 8.106 |
| **52** | T | N | 112.192 | 111.888 |
| **53** | E | H | 8.116 | 8.115 |
| **53** | E | N | 122.315 | 122.504 |
| **54** | I | H | 10.368 | 10.368 |
| **54** | I | N | 128.931 | 128.931 |

| | | | | |
|---|---|---|---|---|
| **56** | D | H | 8.899 | 8.956 |
| **56** | D | N | 125.505 | 124.756 |
| **57** | E | H | 9.229 | 9.415 |
| **57** | E | N | 116.112 | 116.653 |
| **58** | E | H | 7.204 | 7.218 |
| **58** | E | N | 116.007 | 116.014 |
| **59** | A | H | 8.106 | 8.115 |
| **59** | A | N | 122.914 | 122.504 |
| **60** | E | H | 7.548 | 7.553 |
| **60** | E | N | 112.456 | 112.458 |
| **61** | K | H | 7.026 | 7.105 |
| **61** | K | N | 114.239 | 114.424 |
| **62** | I | H | 7.638 | 7.638 |
| **62** | I | N | 122.998 | 122.371 |
| **63** | T | H | 8.011 | 8.106 |
| **63** | T | N | 112.192 | 111.888 |
| **64** | T | H | 7.249 | 7.11 |
| **64** | T | N | 110.084 | 110.54 |
| **65** | V | H | 7.951 | 7.97 |
| **65** | V | N | 121.409 | 120.692 |
| **66** | Q | H | 8.708 | 8.704 |
| **66** | Q | N | 118.007 | 117.894 |
| **67** | A | H | 7.764 | 7.825 |
| **67** | A | N | 119.34 | 119.615 |
| **68** | A | H | 7.948 | 7.985 |
| **68** | A | N | 122.947 | 123.653 |
| **69** | I | H | 8.119 | 8.106 |
| **69** | I | N | 119.205 | 119.327 |
| **70** | D | H | 9.096 | 9.15 |
| **70** | D | N | 119.228 | 118.992 |
| **71** | Y | H | 8.16 | 8.151 |
| **71** | Y | N | 121.792 | 121.314 |
| **72** | I | H | 8.187 | 8.204 |
| **72** | I | N | 120.25 | 120.197 |
| **73** | N | H | 8.833 | 8.835 |
| **73** | N | N | 118.276 | 118.284 |
| **74** | G | H | 7.83 | 7.804 |
| **74** | G | N | 104.989 | 105.097 |
| **75** | H | H | 7.647 | 7.643 |
| **75** | H | N | 118.557 | 118.813 |

Supplemental References


1. 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. https://doi.org/10.1038/nature12810.

2. NMRPipe: A multidimensional spectral processing system based on UNIX pipes | SpringerLink, (available at https://link.springer.com/article/10.1007/BF00197809).

3. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinforma. Oxf. Engl. 31, 1325–1327 (2015).

4. J. D. Hunter, Matplotlib: A 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95 (2007).

5. PyMOL. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

## Chapter 4. The Effect of Substrate on the Protein-Protein Interactions of Acyl Carrier Protein Reductases

### 4.1 Introduction

Fatty acid biosynthesis is an essential and conserved system of chemical reactions, gated by an initial protein-protein interaction which serves as the carbon chain factory of the cell[1–3]. This conserved system creates fatty acids for membranes[4], natural products[5], and some signaling pathways[6] in virtually all species (Figure 1)[7,8]. Given the essential nature of the biosynthetic pathway it has been a popular choice for inhibition as a new antibacterial target[9–12]. However, inhibitors are limited by the high homology of the active site residues[13]. The same essential nature of FAS which makes it an attractive inhibition target means there is little variation in the mechanism of the reductases and their active site[14,15]. With the mechanism of human mitochondrial reductases the same as *E. coli*'s[16]. As such active site inhibitors such as triclosan have potent off target effects, with triclosan even an emerging environmental toxin[17,18]. This necessitates a new means of targeting these enzymes, which can take advantage of the more specific protein-protein interactions responsible for forming the reaction complex[19,20].

Reductases are a popular target of inhibitor design due to their non-redundancy in the fatty acid cycle[21]. The cycle begins with the condensation of acyl chains with malonate carried out by ketosynthases to form a 3-oxo fatty acid[22]. This 3-oxo species must be reduced by a 3-ketoacyl reductase to form a 3-hydroxy fatty acid[23]. 3-hydroxy fatty acids are dehydrated to a 2-3-unsaturated fatty acid[24]. In certain cases, the unsaturated fatty acid can be then accepted by a ketosynthase to elongate the unsaturated fat[25–27]. However, generally the final step is a final reduction by an enoyl-reductase to form a saturated fat[28,29]. This constitutes the full cycle that fats must move through to elongate, notably ketosynthases and dehydratases often have two homologous enzymes within the pathway making them more resilient to inhibition. However, both reductases are non-redundant and show high homology between bacteria, as well as maintaining the same mechanism of reduction[30].

Throughout the fatty acid elongation steps described above the acyl substrates are shuttled and protected by a small acyl carrier protein (ACP). ACPs are small at under 100 amino acids, typically with a four helical-bundle structure[31]. ACPs are highly acid on their surface with a hydrophobic central cavity, with a pantetheine cofactor functionalized onto a conserved serine which is sheltered by the hydrophobic

cavity[32]. The growing acyl substrates of FAS are held within the hydrophobic pocket of ACP, protecting reactive intermediates during the iterations of elongation[33,34]. The initial protein-protein interactions which precede the transfer of acyl intermediates into the partner, termed chain-flipping, may serve as a control point for substrates before the committed step of chain-flipping[35,36]. This ACP-Reductase interface interaction may serve as an alternative “druggable-site” with more specificity between the reductases of different species[37,38]. An inhibitor that could interfere with the protein-protein interactions of the reductases could accomplish the goal of stopping this critical step in elongation. We will demonstrate that substrate priming could be used as an alternative drug target in the FAS model system *E. coli* ACP, AcpP, with reductases requiring substrate for the full organization of the AcpP interface.

Numerous studies have been performed to attempt to design inhibitors for the reductase enzymes, especially FabI and its homologs in numerous species[39–44]. However, these studies primarily leverage the active site of enoyl and ketoreductases, suffering from the issues of specificity which have hampered drugs like triclosan. In this study we have expanded on previous experiments which developed computational models through a combination of NMR spectroscopy, structural analysis, and computational modeling. A set of analyses were performed to examine the structural changes due to cofactor binding in reductases. Further, novel experiments were performed to characterize the differences in protein-protein interactions between substrate-bound and apo FabI and FabG. The findings from this work were combined to examine sites and specific residues which lie along the ACP ● Partner interface which are promising as sites for inhibitors leveraging the interface of protein-protein interactions.

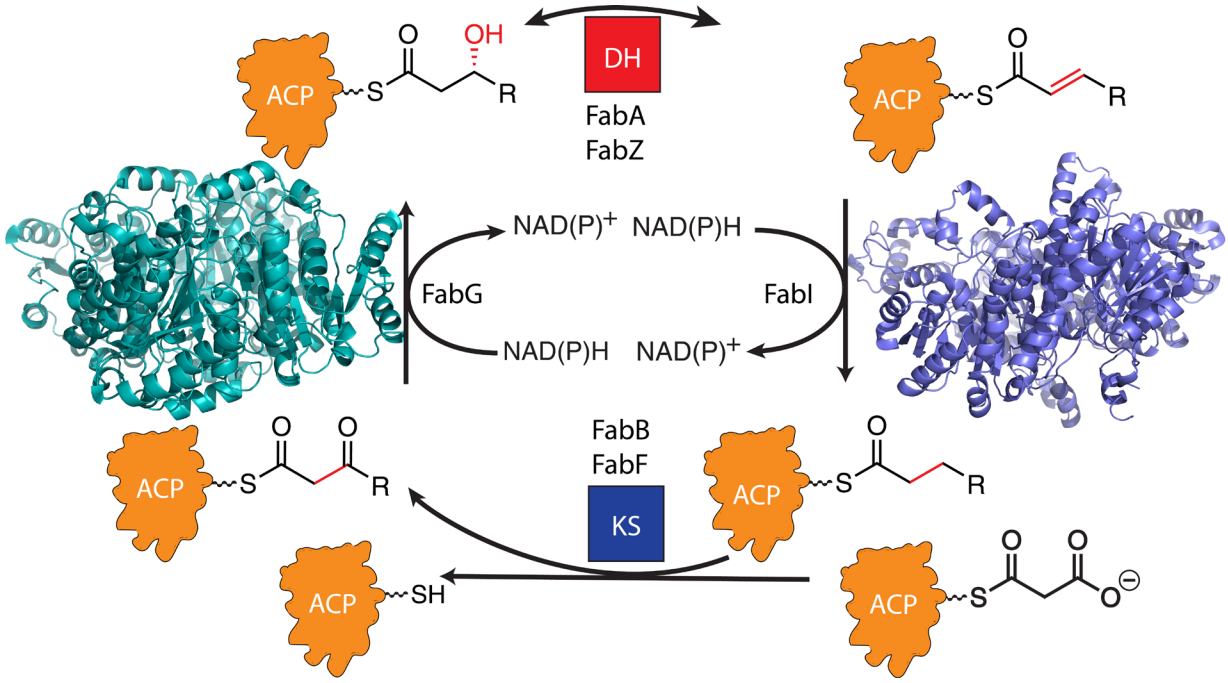



**Figure 4.1 The Fatty Acid Biosynthetic cycle. The Ketoreductase FabG is displayed next to its role deep teal. The Enoylreductase FabI is shown in deep blue. The ketosynthases (KS) are FabF and FabB, possessing similar but unique roles in *E. coli*. The dehydratases (DH) are FabA and FabZ, similarly they have the same mechanism but non-overlapping roles.**

## 4.2 Results

### Structural effects of cofactor binding

To judge the effects of substrate binding on the structure of the reductase the structures of FabI and FabG were aligned with and without the presence of substrate and colored by RMSD[45]. The bound AcpP from modeling was added into the structure to give context of the AcpP binding site[46].

FabG and its homologs have had significantly less studies into its interactions and drug development than the enoylreductase. We hope that this study aid in structurally guided drug design, with the models and druggable sites identified as a starting point. Structurally it was first noted that the largest region of variation between substrate bound (PDB:1Q7B[47]) and unbound (PDB:1I01[48]), is the region of residues 186-201 (Figure 2b). This region of the FabG partner protein has been shown to bind helix III of AcpP in docked models. The remainder of the large RMSD regions lie at the interface between the oligomers of FabG.

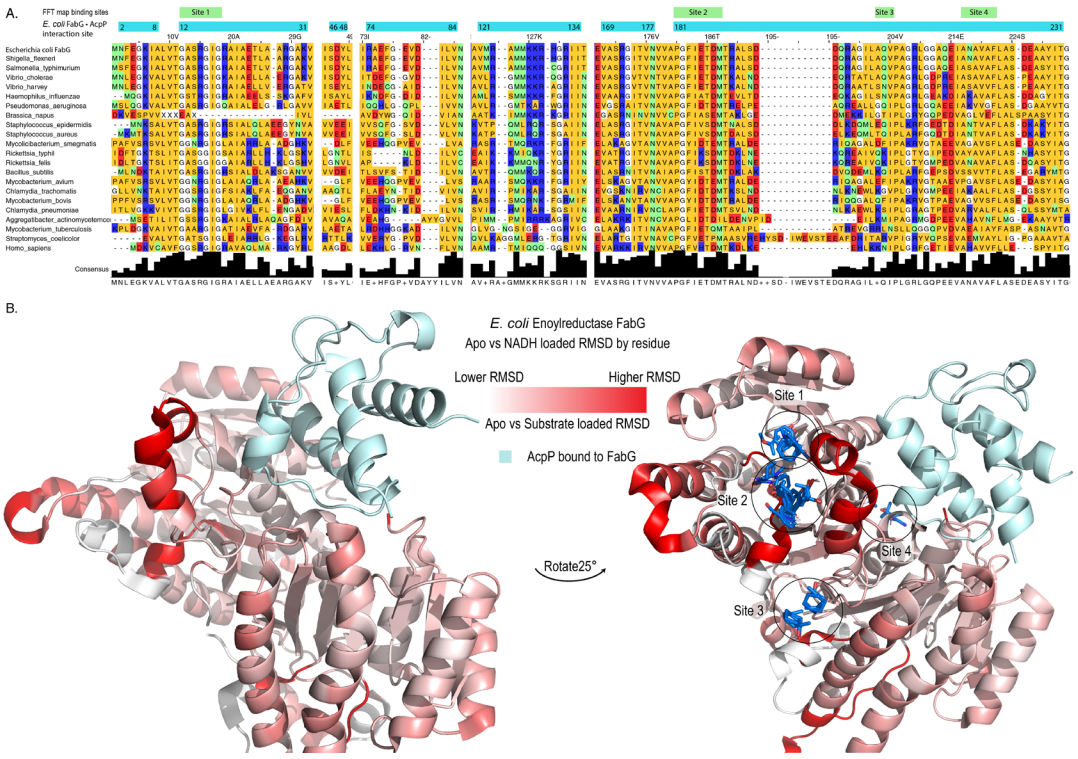



**Figure 4.2 Alignment of FabG homologs and structural effects of NADH substrate. A. Alignments of the high homology ketoreductase homologs and *H. sapiens* FabG. The entire sequence was cut to only the residues which are nearby the AcpP●FabG interface to focus on the interface homology. Sites identified by FFTMap are mapped onto the alignments. B. The structural RMSD between substrate-bound and apo FabG is shown in a red gradient. AcpP's previously identified bound orientation is shown in teal. The sites of FFTMap pockets identified are shown as blue sticks.**

Unlike FabG, FabI and its homologs are frequently computationally studied antibiotic targets, however the research that has taken place often does not account for the differences between substrate bound and unbound states. The structures of the NADH bound (PDB:4CV3[49]) and apo (PDB:5CFZ[50]) forms of FabI were compared. It was seen that there was an area of high RMSD at the ACP interface between residues 39-49 (Figure 3b), comprising a loop and top of helix which are nearby the NADH binding site.

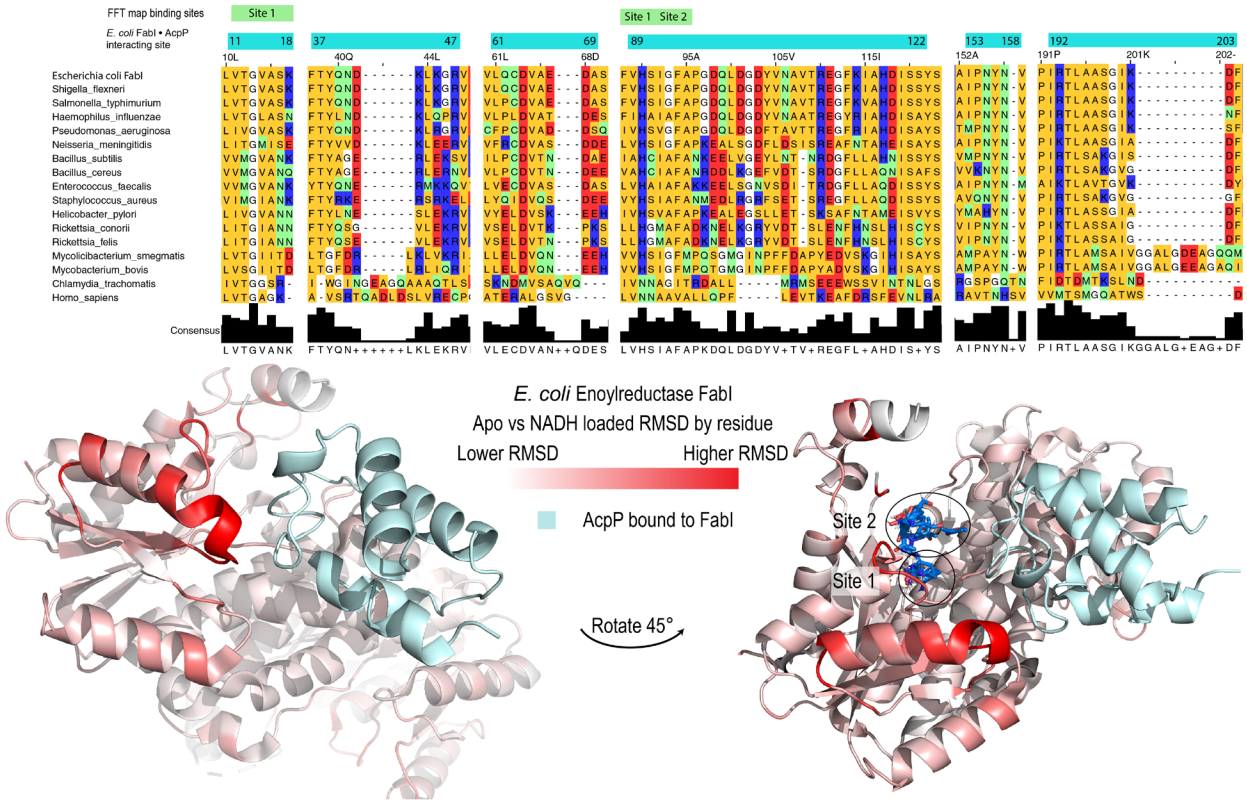



**Figure 4.3 Alignment of FabI homologs and structural effects of NADH substrate. A. Alignments of the high homology enoylreductase homologs and *H. sapiens* FabI. The entire sequence was cut to only the residues which are nearby the AcpP•FabI interface to focus on the interface homology. Sites identified by FFTMap are mapped onto the alignments. B. The structural RMSD between substrate-bound and apo FabI is shown in a red gradient. AcpP's previously identified bound orientation is shown in teal. The sites of FFTMap pockets identified are shown as blue sticks.**

### Alignments of reductases

To examine the common features of the reductase enzymes, genetic alignments were performed with the most similar genes to the *E. coli* FabI and FabG. Structures of available reductases were aligned to examine the structural similarities. The goal of this study was to judge if the structural observations made in this study would transfer to other prospective drug targets. It was observed that the regions which are adjacent to the AcpP interaction site exhibit high sequence homology, suggesting that findings affecting the protein-protein interactions in *E. coli* will serve as a model for other species. In the FabG alignment the areas of alignment begin at the N-termini with the first 30 residues (figure 2A), with most regions showing similar residues. This FabG interaction site lies near AcpPs primary binding surface, helix II. The most stark difference is seen at residue 23, with approximately one third of the species displaying a negatively charged

residue at this position, while another third shows oppositely charged residues, and another third of the species are chargeless at this position. In the docked model this residue lies nearby the top of helix II of AcpP, positioned on the FabG helix which lies against site 1 identified by FTMap which will be explored more later. The next two large sections of the reductases overlay with very high homology, between residues 74-84 and 121-134 there is more dissimilarity around residues 74 but after this section the proteins are all very similar. This site of interaction corresponds to an interaction between ACP and an interaction previously reported on the adjacent FabG monomer, binding the adjacent FabG monomer at loop II and the base of helix IV. The final section of interaction between ACP and FabG is set near the end of the protein on residues 169-231. This is also likely the most significant set of residues, coordinating the ACP at helix II and III and containing the most significant area of structural RMSD with substrate loading.

FabI shows even more similarity between the bacterial species' alignment, in addition the human enoyl reductase is very dissimilar. The *E. coli* binding site begins from residues 11 through 18 and 37 through 47 (Figure 3A). This comprises the alpha helices and loops that are of the highest RMSD between substrate bound and unbound FabI. In the docked model this region binds the ACP at the "top" of the enzyme with loop III and IV angled towards the region. The next region of binding is at residues 61 to 69, this is the second region which is pointed towards the top of the ACP. All three of these binding regions see overall high homology among bacteria, but the human enoyl reductase especially does not display high homology. The next large region of binding is between residues 89 and 122, this constitutes most of the surface near the ACP. With high homology this region appears to be the main ACP binding surface, which likewise shows little RMSD between the substrate bound and unbound states of FabI as evidence that ACP likely binds in this region. The ACP coordinates this region across helix III and at the serine 36 side of helix II. The region from 192 to 203 was a more complicated analysis. There is high homology between species and it is likely a highly specific region, but the loop and small helix is not resolved in many structures. It is also the region which can bind serine 36 of the ACP, suggesting a very important role in activity.

**FTMAP analysis for protein pockets nearby the interface**

Given the unique scope afforded by the reductase interaction models an FTMap analysis was performed on the residues within 20Å of the AcpP interface[51–54]. The FTMap results are collected in the

supplemental information. In the case of FabI two pockets were identified. These are displayed on figure 3, but the regions show high homology to other bacterial reductases and low homology to the *H. sapiens* homolog. The FabG analysis showed even more regions with 4 pockets identified by FTMap. The region on both proteins tended to lie nearby regions of high RMSD between the apo and substrate bound models.

**Binding differences tested by $^{15}$N-HSQC titration**

To experimentally observe the differences in interactions between substrate loaded and apo forms of FabI and FabG with AcpP, titrations were performed on apo FabI and FabG. These were compared to the published NADP+ loaded results[46] to see the differences caused by the structural changes observed above. Titrations were performed by collecting a set of $^{1}$H-$^{15}$N HSQCs and observing the changes in chemical shift affected by the addition of partner protein to AcpP[55]. C8-AcpP was prepared through chemoenzymatic loading as previously reported[56]. Unfortunately, the apo partner proteins experienced issues with stability, causing a loss of signal in experiments. However, after buffer and concentration optimization enough stability to reach saturation of AcpP with partner protiens was reached. CSPs were calculated with an α value of 0.14, as has been used previously in AcpP titrations[36].

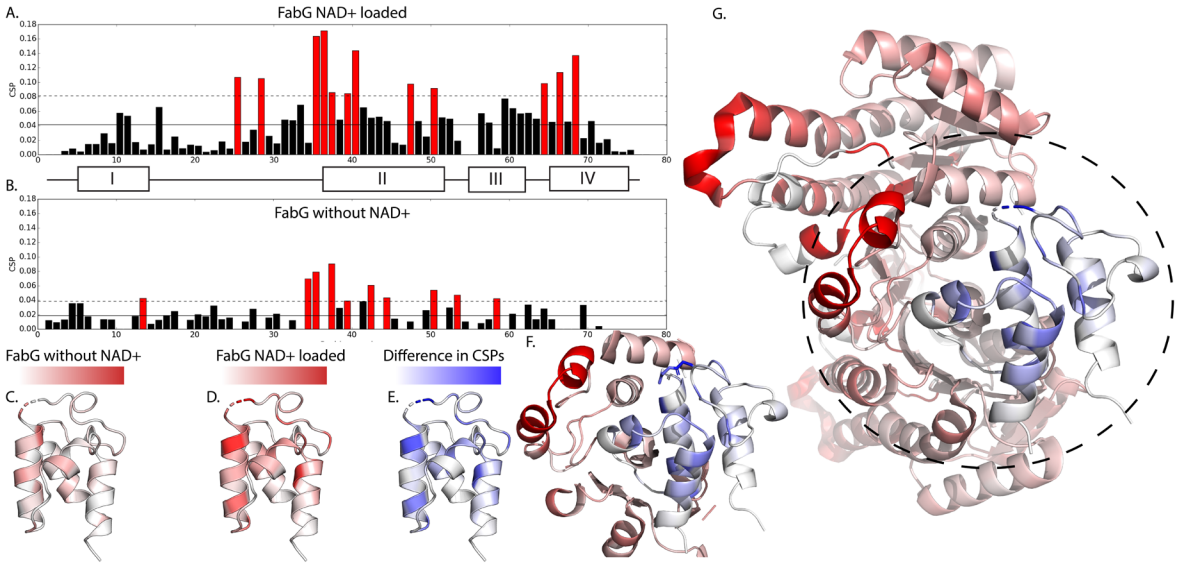



**Figure 4.4 The chemical shift perturbations resulting from titration of C8-AcpP with apo and NAD+ bound FabG. A. The perturbations resulting from the titration of C8-AcpP with NAD+ bound FabG. B. The perturbations resulting from the titration of C8-AcpP with apo FabG. C. Projection of the apo-FabG titration CSPs from section B onto the structure of AcpP. D. Projection of the NAD+ bound FabG titration CSPs from section A onto the structure of AcpP. E. Subtraction of the CSPs from the apo and NAD+ bound titrations, projected onto the AcpP structure. This highlights the regions which have perturbations that are unique to only one study. The FabG is colored in the same manner as figure 2. F. A zoom into the protein-protein interaction site of the C8-AcpP●FabG interface. G. the entire surface of the C8-AcpP●FabG interaction. The same RMSD coloring of the FabG is utilized as in figure 2.**

### Titration of AcpP with the 3-oxoacyl-carrier-protein Reductase, FabG

Titrations were performed and chemical shift perturbation (CSPs) were observed up to 2 molar equivalents of FabG without the addition of NADP+. The first observation that can be made when comparing the CSPs of AcpP interacting with both apo and NADP+ loaded FabG is the much lower magnitude of peak movement (Figure 4A&B). Furthermore, it was observed that the regions of interaction are more localized to the interface of the protein. The largest perturbations in the AcpP●apofabG interactions occur early on the AcpP, across helix II, and nearby helix III. The first set of AcpP●apoFabG interactions appear at residues V7, K8, and L15, these interactions lie near the top of the acyl pocket. The next region of the largest perturbations are residues S36, L37, and T39, these residues lie directly adjacent to a number of

salt bridges in the model. The region of interactions on FabG for these residues lies far from the region of high RMSD. It stands to reason that the interactions are still present in the AcpP●apoFabG interactions. The last region showing high RMSD are the residues M44, T52, E57, and I62. These are spread through mid-helix II, mid loop II, and helix III. These high CSP residues lie nearby regions of low RMSD between the apo and substrate loaded FabG. Furthermore, the highest CSPs are often relatively even mix of hydrophobic residues and acidic residues which perform the interaction. Of the set of largest perturbations only one is acidic. Taken together these results show how FabG is unable to coordinate its proper binding interface without the proper substrate.

To evaluate the differences between AcpP●FabG and AcpP●apoFabG titrations the CSPs from both interactions were subtracted from one another and visualized (Figure 4C). The first set of AcpP●FabG interactions to examine are the negative residues D35, E47, E48, and E60. D35, E47, and E48 all lie along the interacting face of helix II, constituting much of the electrostatic contacts. E60 is on helix III and appears to be angled towards the high RMSD region of the FabG interface (Figure 4D&E). The second set of AcpP●FabG interactions which were lost in the apo titration were hydrophobic residues within the acyl pocket. I10, V40, L42, P50, and A68 are all residues lining the acyl pocket which show a larger difference in perturbation between NADP+ and apo FabG titrations. Taken together these two sets of lost interactions suggest both a loss of the electrostatic interface and possibly a loss of chain flipping. With the large set of hydrophobic residues likely representing the effects of chain flipping the absence of these perturbations would seem to denote a lack of or diminished chain-flipping. However, these definitively demonstrate that an apo FabG is an inaccurate system for testing inhibition or interaction of FabG with AcpP.

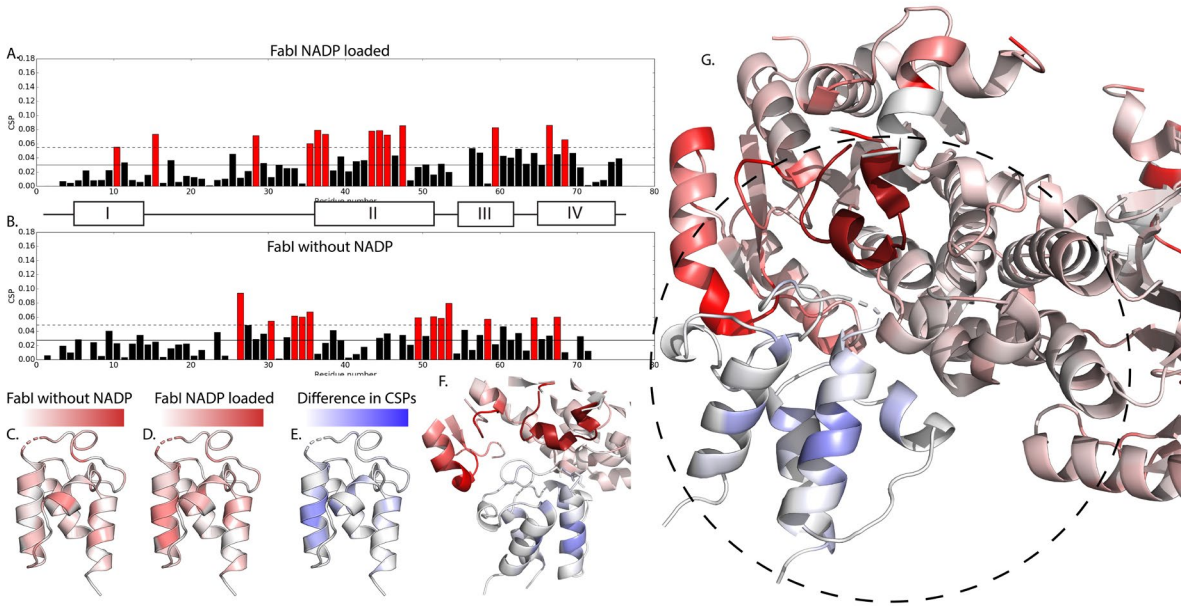



**Figure 4.5 The chemical shift perturbations resulting from titration of C8-AcpP with apo and NAD+ bound FabI. A. The perturbations resulting from the titration of C8-AcpP with NAD+ bound FabI. B. The perturbations resulting from the titration of C8-AcpP with apo-FabI. C. Projection of the apo-FabI titration CSPs from section B onto the structure of AcpP. D. Projection of the NAD+ bound FabI titration CSPs from section A onto the structure of AcpP. E. Subtraction of the CSPs from the apo and NAD+ bound titrations, projected onto the AcpP structure. This highlights the regions which have perturbations that are unique to only one study. The FabI is colored in the same manner as figure 3. F. A zoom into the protein-protein interaction site of the C8-AcpP●FabI interface. G. the entire surface of the C8-AcpP●FabI interaction. The same RMSD coloring of the FabI is utilized as in figure 3.**

### Titration of AcpP with the Enoyl-acyl-carrier-protein Reductase, FabI

Titrations were performed with FabI in the same manner as FabG. $^{1}H$-$^{15}N$ HSQC titrations were collected on increasing concentrations of FabI, protein concentration was increased in increments up to 2.0 molar equivalents. Interestingly the degree of CSP loss between apo- and substrate-bound FabI was less compared to FabG, though the AcpP●apoFabG and AcpP●apoFabI titrations have similar CSP magnitude (Figure 5A&B). The perturbations in the AcpP●apoFabI titration remain very low in magnitude through helix I and the first half of loop I. There are only a couple of small perturbations around I11 and L15. The first perturbation of interest occurs at residue P28, with additional smaller perturbations at A30 and L31. There are three large perturbations at D35, S36, and L37. There are four electrostatic residues with high

perturbation at D51, E53, D56, and E57, with a final three major hydrophobic perturbations at I62, A68, and Y71. The represent a spread set of perturbations, with many of the largest CSPs having other residues perturbed near them. The CSPs lie in regions nearby high RMSD regions of the FabI, with loop I within distance of interaction with the K43 helix of FabI. Secondly the region around S36 is directly adjacent to the region of FabI which is not resolved in the apo structure. In solution apoFabI appears to maintain interactions at this region suggesting that though it is unresolved in the crystal structure and likely quite dynamic, this region of FabI appears to still be capable of interacting with AcpP.

To observe where the differences in the apo and substrate loaded interaction lie the two sets of CSPs were subtracted from one another (Figure 5C). The CSPs which were larger in the substrate loaded titration are almost entirely hydrophobic or uncharged, spanning I10, L15, V43, M44, A45, A59, Q66, and A67. These residues are across all four helices, but only occur on helices. This may appear to suggest that there is still a difference in the effect on the substrate of the AcpP, though the interactions appear similar only the substrate loaded FabI seems to perturb the acyl pocket so much. There is one electrostatic interaction which is different at E60, right at the top of helix III and pointing into a low RMSD portion of FabI. Unlike the FabG titration, the FabI titration revealed residues with higher CSPs in the apo experiment (Figure 5D&E). Residues D51, E53, and Y71 are more perturbed in the apo titration, these residues are all located at the base of the AcpP. D51 and E53 are at the base of helix II and loop II and are not seen within distance for hydrogen bonding in the model. This interaction occurs on the region of FabI which is not resolved in the apo structure. This observation taken together with the hydrophobic perturbation differences seem to suggest that there is an alternate binding confirmation which AcpP adopts when interaction with apoFabI. However, given the lack of structural knowledge of the apoFabI structure it is difficult to conjecture what that interface might look like.

### 4.3 Discussion

Reductases remain a popular target for inhibition, with no substitutes for their activity and a well-known function with existing structures. However, researchers have been unable to develop inhibitors which are specific for targeting bacterial reductases but do not affect patients. With the common mechanism of reductases for both enzymes, an inhibitor which targets the pocket or binds the common NADH cofactor

always comes with a risk of affecting homologs. However, the ACP interfaces of many bacterial reductases share high homology and as discussed above the interfaces of the human reductases are dissimilar. This highlights an important detail for inhibitor design, the AcpP●Partner interface requires NADH to assemble. Without the cofactor AcpP loses significant contacts and appears unable to correctly bind reductases. Furthermore, studies should take care when designing protein-protein interaction inhibitors as the interface of FabI and FabG varies significantly with or without cofactor. But, these models and new observations can aid in the development of protein-protein interface inhibitors between ACP●Reductases.

Previous studies have heavily favored inhibitors designed to target FabI, with most studies designing active site inhibitors with mixed success. It is our belief that a detailed knowledge of the interface residues and function in the reductases could yield selective interface inhibitors. With the known interface inhibitors, the pockets nearby the carrier protein interface, and knowledge of the details of the interaction, inhibitor selectivity could be possible. The knowledge of the necessity of substrate priming for assembly of the interface could even be leveraged to target the dynamic motions necessary for creation of the interface. As computational capabilities grow and high throughput screening and small molecules libraries become more accessible, a detailed knowledge of protein function is more essential. The correct system and state of a protein is essential for inhibitor design. The high homology and essential nature of reductases make them an important target for bacterial inhibition.

**Chapter 4**, is in print of the manuscript in preparation: Thomas G. Bartholow, Megan A Young, Michael D. Burkart “Atomic Details of the Cofactor Priming and Protein-Protein Interaction of E. coli Acyl Carrier Protein Reductases.” The dissertation author is the primary author of the manuscript.

### 4.4 References


(1) Chan, D. I.; Vogel, H. J. Current Understanding of Fatty Acid Biosynthesis and the Acyl Carrier Protein. *Biochem. J.* **2010**, *430* (1), 1–19. https://doi.org/10.1042/BJ20100462.

(2) Cronan, J. E.; Rock, C. O. Biosynthesis of Membrane Lipids. *EcoSal Plus* **2008**, *3* (1). https://doi.org/10.1128/ecosalplus.3.6.4.

(3) Magnuson, K.; Jackowski, S.; Rock, C. O.; Cronan, J. E. Regulation of Fatty Acid Biosynthesis in Escherichia Coli. *Microbiol. Mol. Biol. Rev.* **1993**, *57* (3), 522–542.

(4) Rock, C. O.; Jackowski, S. Incorporation and Turnover of Fatty Acids in Escherichia Coli Membrane Phospholipids. In *Advances in Lipobiology*; Gross, R. W., Ed.; JAI, 1996; Vol. 1, pp 39–59. https://doi.org/10.1016/S1874-5245(96)80004-8.

(5) Carreau, J.-P. [32] Biosynthesis of Lipoic Acid via Unsaturated Fatty Acids. In *Methods in Enzymology*; Vitamins and Coenzymes Part D; Academic Press, 1979; Vol. 62, pp 152–158. https://doi.org/10.1016/0076-6879(79)62212-7.

(6) Parsek, M. R.; Val, D. L.; Hanzelka, B. L.; Cronan, J. E.; Greenberg, E. P. Acyl Homoserine-Lactone Quorum-Sensing Signal Generation. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (8), 4360–4365.

(7) Chakravarty, B.; Gu, Z.; Chirala, S. S.; Wakil, S. J.; Quiocho, F. A. Human Fatty Acid Synthase: Structure and Substrate Selectivity of the Thioesterase Domain. *Proc. Natl. Acad. Sci.* **2004**, *101* (44), 15567–15572. https://doi.org/10.1073/pnas.0406901101.

(8) Sohlenkamp, C.; Geiger, O. Bacterial Membrane Lipids: Diversity in Structures and Pathways. *FEMS Microbiol. Rev.* **2016**, *40* (1), 133–159. https://doi.org/10.1093/femsre/fuv008.

(9) Yao, J.; Rock, C. O. Bacterial Fatty Acid Metabolism in Modern Antibiotic Discovery. *Biochim. Biophys. Acta* **2017**, *1862* (11), 1300–1309. https://doi.org/10.1016/j.bbalip.2016.09.014.

(10) Campbell, J. W.; Cronan, J. E. Bacterial Fatty Acid Biosynthesis: Targets for Antibacterial Drug Discovery. *Annu. Rev. Microbiol.* **2001**, *55* (1), 305–332. https://doi.org/10.1146/annurev.micro.55.1.305.

(11) 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. https://doi.org/10.1007/s00253-001-0918-z.

(12) Zhang, Y.-M.; White, S. W.; Rock, C. O. Inhibiting Bacterial Fatty Acid Synthesis *. *J. Biol. Chem.* **2006**, *281* (26), 17541–17544. https://doi.org/10.1074/jbc.R600004200.

(13) Parsons, J. B.; Rock, C. O. Is Bacterial Fatty Acid Synthesis a Valid Target for Antibacterial Drug Discovery? *Curr. Opin. Microbiol.* **2011**, *14* (5), 544–549. https://doi.org/10.1016/j.mib.2011.07.029.

(14) Lu, X.; Huang, K.; You, Q. Enoyl Acyl Carrier Protein Reductase Inhibitors: A Patent Review (2006 – 2010). *Expert Opin. Ther. Pat.* **2011**, *21* (7), 1007–1022. https://doi.org/10.1517/13543776.2011.581227.

(15) Lu, H.; Tonge, P. J. Inhibitors of FabI, an Enzyme Drug Target in the Bacterial Fatty Acid Biosynthesis Pathway. *Acc. Chem. Res.* **2008**, *41* (1), 11–20. https://doi.org/10.1021/ar700156e.

(16) Liu, B.; Wang, Y.; Fillgrove, K. L.; Anderson, V. E. Triclosan Inhibits Enoyl-Reductase of Type I Fatty Acid Synthase in Vitro and Is Cytotoxic to MCF-7 and SKBr-3 Breast Cancer Cells. *Cancer Chemother. Pharmacol.* **2002**, *49* (3), 187–193. https://doi.org/10.1007/s00280-001-0399-x.

(17) Dann, A. B.; Hontela, A. Triclosan: Environmental Exposure, Toxicity and Mechanisms of Action. *J. Appl. Toxicol. JAT* **2011**, *31* (4), 285–311. https://doi.org/10.1002/jat.1660.

(18) Yueh, M.-F.; Tukey, R. H. Triclosan: A Widespread Environmental Toxicant with Many Biological Effects. *Annu. Rev. Pharmacol. Toxicol.* **2016**, *56*, 251–272. https://doi.org/10.1146/annurev-pharmtox-010715-103417.

(19) 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. https://doi.org/10.1039/C8NP00040A.

(20) Finzel, K.; Lee, D. J.; Burkart, M. D. Using Modern Tools to Probe the Structure-Function Relationship of Fatty Acid Synthases. *Chembiochem Eur. J. Chem. Biol.* **2015**, *16* (4), 528–547. https://doi.org/10.1002/cbic.201402578.

(21) Zhang, Y.-M.; Lu, Y.-J.; Rock, C. O. The Reductase Steps of the Type II Fatty Acid Synthase as Antimicrobial Targets. *Lipids* **2004**, *39* (11), 1055–1060. https://doi.org/10.1007/s11745-004-1330-3.

(22) 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. https://doi.org/10.1016/S0014-5793(96)01437-8.

(23) Yu, X.; Liu, T.; Zhu, F.; Khosla, C. In Vitro Reconstitution and Steady-State Analysis of the Fatty Acid Synthase from Escherichia Coli. *Proc. Natl. Acad. Sci.* **2011**, *108* (46), 18643–18648. https://doi.org/10.1073/pnas.1110852108.

(24) 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* (44), 27795–27801. https://doi.org/10.1074/jbc.271.44.27795.

(25) 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. https://doi.org/10.1016/j.chembiol.2015.09.009.

(26) Xiao, X.; Yu, X.; Khosla, C. Metabolic Flux between Unsaturated and Saturated Fatty Acids Is Controlled by the FabA:FabB Ratio in the Fully Reconstituted Fatty Acid Biosynthetic Pathway of Escherichia Coli. *Biochemistry* **2013**, *52* (46), 8304–8312. https://doi.org/10.1021/bi401116n.

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

(28) Bergler, H.; Fuchsbichler, S.; Högenauer, G.; Turnowsky, F. The Enoyl-[Acyl-Carrier-Protein] Reductase (FabI) of Escherichia Coli, Which Catalyzes a Key Regulatory Step in Fatty Acid Biosynthesis, Accepts NADH and NADPH as Cofactors and Is Inhibited by Palmitoyl-CoA. *Eur. J. Biochem.* **1996**, *242* (3), 689–694. https://doi.org/10.1111/j.1432-1033.1996.0689r.x.

(29) Heath, R. J.; Rock, C. O. Enoyl-Acyl Carrier Protein Reductase (FabI) Plays a Determinant Role in Completing Cycles of Fatty Acid Elongation in Escherichia Coli(∗). *J. Biol. Chem.* **1995**, *270* (44), 26538–26542. https://doi.org/10.1074/jbc.270.44.26538.

(30) Massengo-Tiassé, R. P.; Cronan, J. E. Diversity in Enoyl-Acyl Carrier Protein Reductases. *Cell. Mol. Life Sci. CMLS* **2009**, *66* (9), 1507. https://doi.org/10.1007/s00018-009-8704-7.

(31) Byers, D. M. B. M.; Gong, H. G. Acyl Carrier Protein: Structure–Function Relationships in a Conserved Multifunctional Protein Family. *Biochem. Cell Biol.* **2007**. https://doi.org/10.1139/O07-109.

(32) Kim, Y.; Prestegard, J. H. A Dynamic Model for the Structure of Acyl Carrier Protein in Solution. *Biochemistry* **1989**, *28* (22), 8792–8797. https://doi.org/10.1021/bi00448a017.

(33) Farmer, R.; Thomas, C. M.; Winn, P. J. Structure, Function and Dynamics in Acyl Carrier Proteins. *PLOS ONE* **2019**, *14* (7), e0219435. https://doi.org/10.1371/journal.pone.0219435.

(34) 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. https://doi.org/10.1073/pnas.1818686116.

(35) 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. https://doi.org/10.1002/anie.201408576.

(36) 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. https://doi.org/10.1038/nature12810.

(37) Ran, X.; Gestwicki, J. E. Inhibitors of Protein-Protein Interactions (PPIs): An Analysis of Scaffold Choices and Buried Surface Area. *Curr. Opin. Chem. Biol.* **2018**, *44*, 75–86. https://doi.org/10.1016/j.cbpa.2018.06.004.

(38) Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J. Recent Advances in the Development of Protein–Protein Interactions Modulators: Mechanisms and Clinical Trials. *Signal Transduct. Target. Ther.* **2020**, *5* (1), 1–23. https://doi.org/10.1038/s41392-020-00315-3.

(39) Cukier, C. D.; Hope, A. G.; Elamin, A. A.; Moynie, L.; Schnell, R.; Schach, S.; Kneuper, H.; Singh, M.; Naismith, J. H.; Lindqvist, Y.; Gray, D. W.; Schneider, G. Discovery of an Allosteric Inhibitor Binding Site in 3-Oxo-Acyl-ACP Reductase from Pseudomonas Aeruginosa. *ACS Chem. Biol.* **2013**, *8* (11), 2518–2527. https://doi.org/10.1021/cb4005063.

(40) Liu, N.; Cummings, J. E.; England, K.; Slayden, R. A.; Tonge, P. J. Mechanism and Inhibition of the FabI Enoyl-ACP Reductase from Burkholderia Pseudomallei. *J. Antimicrob. Chemother.* **2011**, *66* (3), 564–573. https://doi.org/10.1093/jac/dkq509.

(41) Lindert, S.; Tallorin, L. C.; Nguyen, Q. G.; Burkart, M. D.; McCammon, J. A. In Silico Screening for Plasmodium Falciparum Enoyl-ACP Reductase Inhibitors. *J. Comput. Aided Mol. Des.* **2015**, *29* (1), 79–87. https://doi.org/10.1007/s10822-014-9806-3.

(42) Taira, J.; Umei, T.; Inoue, K.; Kitamura, M.; Berenger, F.; Sacchettini, J. C.; Sakamoto, H.; Aoki, S. Improvement of the Novel Inhibitor for Mycobacterium Enoyl-Acyl Carrier Protein Reductase (InhA): A Structure–Activity Relationship Study of KES4 Assisted by in Silico Structure-Based Drug Screening. *J. Antibiot. (Tokyo)* **2020**, *73* (6), 372–381. https://doi.org/10.1038/s41429-020-0293-6.

(43) Junior, L. R. A.; Kronenberger, T.; Serafim, M. S. M.; Sousa, Y. V.; Franco, I. D.; Valli, M.; Bolzani, V. da S.; Monteiro, G. C.; Prates, J. L. B.; Kroon, E. G.; Mota, B. E. F.; Ferreira, D. dos S.; Oliveira, R. B. de; Maltarollo, V. G. Virtual Screening of Antibacterial Compounds by Similarity Search of Enoyl-ACP Reductase (FabI) Inhibitors. *Future Med. Chem.* **2019**. https://doi.org/10.4155/fmc-2019-0158.

(44) Ward, W. H. J.; Holdgate, G. A.; Rowsell, S.; McLean, E. G.; Pauptit, R. A.; Clayton, E.; Nichols, W. W.; Colls, J. G.; Minshull, C. A.; Jude, D. A.; Mistry, A.; Timms, D.; Camble, R.; Hales, N. J.; Britton, C. J.; Taylor, I. W. F. Kinetic and Structural Characteristics of the Inhibition of Enoyl (Acyl Carrier Protein) Reductase by Triclosan. *Biochemistry* **1999**, *38* (38), 12514–12525. https://doi.org/10.1021/bi9907779.

(45) Schrodinger, LLC. *The PyMOL Molecular Graphics System*.

(46) Bartholow, T. G.; Sztain, T.; Lee, D. J.; Abagyan, R.; Burkart, M. D. Elucidation of Transient Protein-Protein Interactions within Carrier Protein-Dependent Biosynthesis. *Commun. Biol. Accept.*

(47) Price, A. C.; Zhang, Y.-M.; Rock, C. O.; White, S. W. Cofactor-Induced Conformational Rearrangements Establish a Catalytically Competent Active Site and a Proton Relay Conduit in FabG. *Structure* **2004**, *12* (3), 417–428. https://doi.org/10.1016/j.str.2004.02.008.

(48) 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. https://doi.org/10.1021/bi010737g.

(49) Schiebel, J.; Chang, A.; Shah, S.; Lu, Y.; Liu, L.; Pan, P.; Hirschbeck, M. W.; Tareilus, M.; Eltschkner, S.; Yu, W.; Cummings, J. E.; Knudson, S. E.; Bommineni, G. R.; Walker, S. G.; Slayden, R. A.; Sotriffer, C. A.; Tonge, P. J.; Kisker, C. Rational Design of Broad Spectrum Antibacterial Activity Based on a Clinically Relevant Enoyl-Acyl Carrier Protein (ACP) Reductase Inhibitor *. *J. Biol. Chem.* **2014**, *289* (23), 15987–16005. https://doi.org/10.1074/jbc.M113.532804.

(50) Jordan, C. A.; Sandoval, B. A.; Serobyan, M. V.; Gilling, D. H.; Groziak, M. P.; Xu, H. H.; Vey, J. L. Crystallographic Insights into the Structure–Activity Relationships of Diazaborine Enoyl-ACP Reductase Inhibitors. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2015**, *71* (12), 1521–1530. https://doi.org/10.1107/S2053230X15022098.

(51) Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTMap Family of Web Servers for Determining and Characterizing Ligand Binding Hot Spots of Proteins. *Nat. Protoc.* **2015**, *10* (5), 733–755. https://doi.org/10.1038/nprot.2015.043.

(52) Brenke, R.; Kozakov, D.; Chuang, G.-Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-Based Identification of Druggable 'Hot Spots' of Proteins Using Fourier Domain Correlation Techniques. *Bioinformatics* **2009**, *25* (5), 621–627. https://doi.org/10.1093/bioinformatics/btp036.

(53) Kozakov, D.; Hall, D. R.; Chuang, G.-Y.; Cencic, R.; Brenke, R.; Grove, L. E.; Beglov, D.; Pelletier, J.; Whitty, A.; Vajda, S. Structural Conservation of Druggable Hot Spots in Protein–Protein Interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (33), 13528–13533. https://doi.org/10.1073/pnas.1101835108.

(54) Bohnuud, T.; Beglov, D.; Ngan, C. H.; Zerbe, B.; Hall, D. R.; Brenke, R.; Vajda, S.; Frank-Kamenetskii, M. D.; Kozakov, D. Computational Mapping Reveals Dramatic Effect of Hoogsteen Breathing on Duplex DNA Reactivity with Formaldehyde. *Nucleic Acids Res.* **2012**, *40* (16), 7644–7652. https://doi.org/10.1093/nar/gks519.

(55) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001.

(56) S. Worthington, A.; D. Burkart, M. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44–46. https://doi.org/10.1039/B512735A.

## Chapter 5 Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production

### 5.1 Introduction

Medium-chain fatty acids (MCFAs, C8-C12) have uses as antimicrobials and emulsifying agents and their derivatization using chain-tailoring enzymes results in an array of chemicals including alkenes, α-olefins, esters, ω-hydroxy-carboxylic acids, α,ω-dicarboxylic acids, alcohols, and ketones. Microbially, MCFAs are generated via hydrolysis of medium-chain fatty acyl-ACPs obtained from Type II fatty acid synthase (FAS)[1], or medium-chain fatty acyl CoAs produced via Type I FAS[2], or reverse β-oxidation[3]. Hydrolysis of fatty acyl intermediates is carried out by medium-chain acyl-ACP thioesterases (TEs) from bacteria, higher plants (FatBs), and mammals, or medium-chain acyl-CoA TEs from bacteria. While bacterial acyl-ACP TEs tend to hydrolyze C8-C16 acyl-ACPs[4], such as *Acinetobacter baylyi* TE[5], FatBs have a narrower substrate profile, such as *Cuphea palustris* TE (C8:0 acyl-ACPs)[6], and *Umbellularia californica* TE (C12:0 acyl-ACPs)[7]. The rat acyl-ACP TE has been show to produce primarily C6-C8 fatty acids[8]. On the acyl-CoA front, *E. coli* ydiI hydrolyzes C6-C10 acyl-CoAs[4].

MCFAs are not extensively synthesized by microbes and their engineered microbial production is hindered by the inefficient expression of heterologous plant thioesterases and the broad fatty acid profile of bacterial thioesterases. Previously long chain acyl-ACP thioesterases, which hydrolyze acyl-ACPs from fatty acid biosynthesis, have been engineered for improved MCFA product profile via active site mutagenesis and computationally-guided approaches. For example, by mutagenizing residues near the enzyme active site of *E. coli* TesA, TesA:L109P[1,9] was identified to preferentially hydrolyze C12:0 and C14:0 acyl-ACPs over C16:0 acyl-ACPs. Another example is *Pseudomonas aeruginosa* TesA:D17S/L462R, which has improved activity on C12:0 acyl-CoAs[10]. More recently, TesA was engineered using an iterative protein redesign and optimization algorithm to select TesA sequences that preferentially bind C8-C12 substrates. This approach resulted in the generation of TesA:S122K/Y145K/L146K and TesA:M141L/Y145K/L146K, which had a 1.8-fold improvement in C12 mole fraction and a 10-fold improvement in C8 mole fraction over wild type, respectively[11].

Using a medium-chain acyl-ACP TEs, such as *Acinetobacter baylyi* TE, as the engineering starting point, however, has the potential to result in a TE with an almost exclusive MCFA product profile and lead to potentially higher microbial yields of MCFAs. Nevertheless, engineering medium-chain acyl-ACP TEs has a unique set of challenges. First, medium-chain acyl-ACP-TEs are heterologous to *E. coli* and may have problems interfacing with the host machinery. Second, active site engineering of medium-chain acyl-ACP TEs may not prove to be as fruitful as mutations may need to be much more subtle, potentially in the second sphere, and overall more difficult to identify. We hypothesized that engineering the interface of a heterologous medium chain TE to better complement the surface of *E. coli* ACP may improve MCFA production in *E. coli* (**Fig. 1**). In Type II FAS, ACP is bound to the fatty acyl-chain and interacts with all the proteins in fatty acid biosynthesis. The fatty acyl chain is buried in the ACP hydrophobic core and protein-protein interactions between the ACP and partner enzymes release the acyl chain from the ACP core and into the partner enzyme's active site[12]. It is the tight binding of ACP to its protein partners, such as FAS subunits, that enables efficient fatty acid biosynthesis[13]. Indeed, crosslinking studies show that *E. coli* Type II FAS ACP binds more tightly to its cognate *E. coli* Type II FAS ketoacid synthase (KS) than *Streptomyces maritimus* Type II polyketide synthase KS.

In this work, we engineer the medium-chain acyl-ACP thioesterase from *Acinetobacter baylyi* (AbTE) to better interface with *E. coli* ACP to improve MCFA production. First, we docked *E. coli* ACP with the endogenous *E. coli* TE TesA and identify potential contact residues involved in stabilizing the ACP-TesA interaction. Next, we mutated the equivalent positions in AbTE to the amino acids found in *E. coli* TesA and measured its fatty acid profile. We find that mutation of just two residues on the AbTE surface, G17 and A165 to arginines, improves MCFA titers more than 3-fold when compared to expression of AbTE wild type in *E. coli*. To our knowledge, this is the first engineering of a heterologous thioesterase for improved MCFA production. More broadly, this work demonstrates that engineering the interface of heterologous enzymes to better couple with endogenous host enzymes may be a useful strategy to improve the microbial production of chemicals that require the expression of heterologous enzymes. Improving the microbial production of MCFAs is significant because MCFAs are key intermediates in the biosynthesis of medium-chain chemicals, including α-olefins, dicarboxylic acids and hydroxyacids, which are important targets in the polymer industry. Finally, using a previously developed G-protein coupled receptor (GPCR)-

based MCFA sensor[14], we detect different MCFA levels of produced by *E. coli* expressing the wild type and engineered thioesterases. Taken together, this work sets up the stage for sensor-guided engineering of MCFA producing microbes.

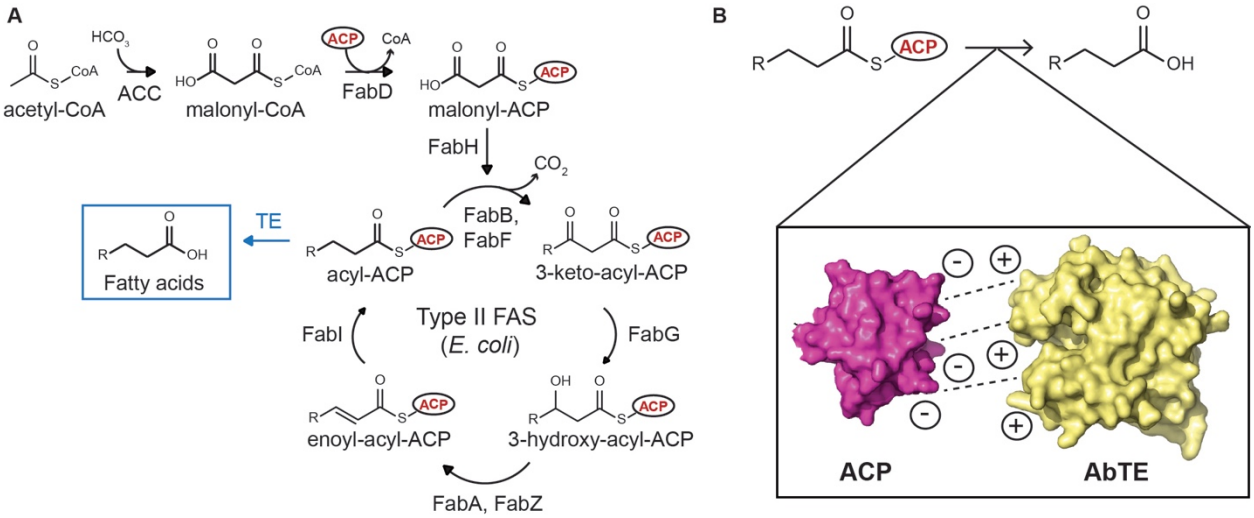



**Figure 5.1 Medium-chain fatty acid biosynthesis in *Escherichia coli*. A. *E. coli* Type II fatty acid synthase extends and reduces an acyl chain bound to acyl-carrier protein (ACP). Thioesterases (TEs) hydrolyze acyl-ACPs to free fatty acids of different chain lengths depending on their substrate specificity. B. Matching the surface interface between the native *E. coli* ACP and medium-chain heterologous thioesterases, such as *Acinetobacter baylyi* thioesterase (AbTE), improves medium-chain fatty acid production in *E. coli*.**

## 5.2 Results

***Screening medium-chain acyl-ACP thioesterases and E. coli hosts for MCFA production.*** To identify the acyl-ACP thioesterase (TE) that results in the highest MCFA production, we expressed the bacterial thioesterase from *Acinetobacter baylyi*[5], and the plant thioesterases from *Cocos nucifera*[6], *Cuphea palustris*[6], and *Umbellularia californica*[15] in *E. coli* MG1655. The percent sequence identity of these TEs to one another ranges from 15-17%, and *A. baylyi* TE has the highest percent identity with *E. coli* 'TesA at 38% (**Fig. SI1**). The thioesterases were expressed in *E. coli* MG1655 and analyzed for MCFA titers in the supernatant. We focus on fatty acids in the supernatant as secreted fatty acids could be continuously extracted from the cell culture, overcoming the need for lysing the producer cells to isolate the fatty acids, and potentially reducing the overall cost for MCFA microbial production. *A. baylyi* thioesterase (AbTE)

results in the highest MCFA titers in the supernatant at 25 mg/L, with its MCFA profile consisting of C8, C10, and C12 fatty acids (**Fig. 2A**). *C. nutifera* TE produced only C12 acid at 5 mg/L, while *U. californica* TE produced mostly C12 acid at 12 mg/L with some C10 acid at 3 mg/L. *C. palustris* TE on the other hand was very specific for C8 acid, producing 8 mg/L and almost no C10 and C12 acids. As our goal is to improve the TE leading to the highest MCFA production, we moved forward with AbTE. The *E. coli* genomic background has been shown to affect chemical production[16]. Therefore, we expressed AbTE in five different *E. coli* hosts: DH5α, BL21, DH10B, MG1655, and BW25113 Δ*fadE*, and measured the MCFA levels in the supernatant. We included the *fadE* deletion as it has been shown to improve fatty acid production in *E. coli*[17]. Surprisingly, *E. coli* hosts BL21 and MG1655 resulted in the highest MCFA productions at 26 mg/L and BW25113 Δ*fadE* produced only 12 mg/L (**Fig. 2B**). Based on these results, we moved forward with AbTE expressed in *E. coli* MG1655.

***Engineering Acinetobacter baylyi thioesterase for improved MCFA titers.*** Expression of the non-functional AbTE:S11A in *E. coli* produces exclusively saturated long-chain (C14-C18) fatty acids due to the presence of endogenous long chain thioesterases in *E. coli*, such as TesA. Expression of AbTE wild type (AbTE:WT) in *E. coli* produced ~29 mg/L of MCFAs (C8-C12) in the supernatant. When total fatty acids were measured, AbTE:WT expressed in *E. coli* produced ~48 mg/L of C8-C12 saturated fatty acids. Specifically, AbTE:WT produced C8, C10, and C12 acids at 9mg/L, 6, mg/L and 14 mg/L, respectively (**Fig. 2A**). In addition to saturated fatty acids, AbTE:WT also produced small levels of unsaturated C12-C16 fatty acids (**Fig SI2**). To identify the interface between *E. coli* TesA and ACP, we used ClusPro[18], which takes into account only the TesA-ACP protein interactions, to dock *E. coli* ACP (PDB ID: 2FAE) and *E. coli* TesA bound to octanoic acid (PDB ID: 1U8U) (**Fig. 2D**). We determined eight positions on TesA that were potentially part of the ACP-TesA interface: Y15, R16, R77, N112, R115, R116, D153, and R160. Structural alignment of TesA with a AbTE homology model revealed that all positions except for R16 (AbTE: G17), R115 (AbTE: T120), R116 (AbTE: A121), D153 (AbTE: N158) and R160 (AbTE: A165) had the same amino acids in these two proteins (**Fig. 2E, Fig. SI2**). Interestingly, four of the five amino acids that are different between TesA and AbTE are positively charged arginines, which could potentially form salt bridges with ACP amino acids on the surface resulting in stabilizing salt bridges during the transfer of the fatty acyl

substrate from ACP to TesA. These potential salt bridges could be the key to engineering a better interface between AbTE and *E. coli* ACP to improve MCFA production.

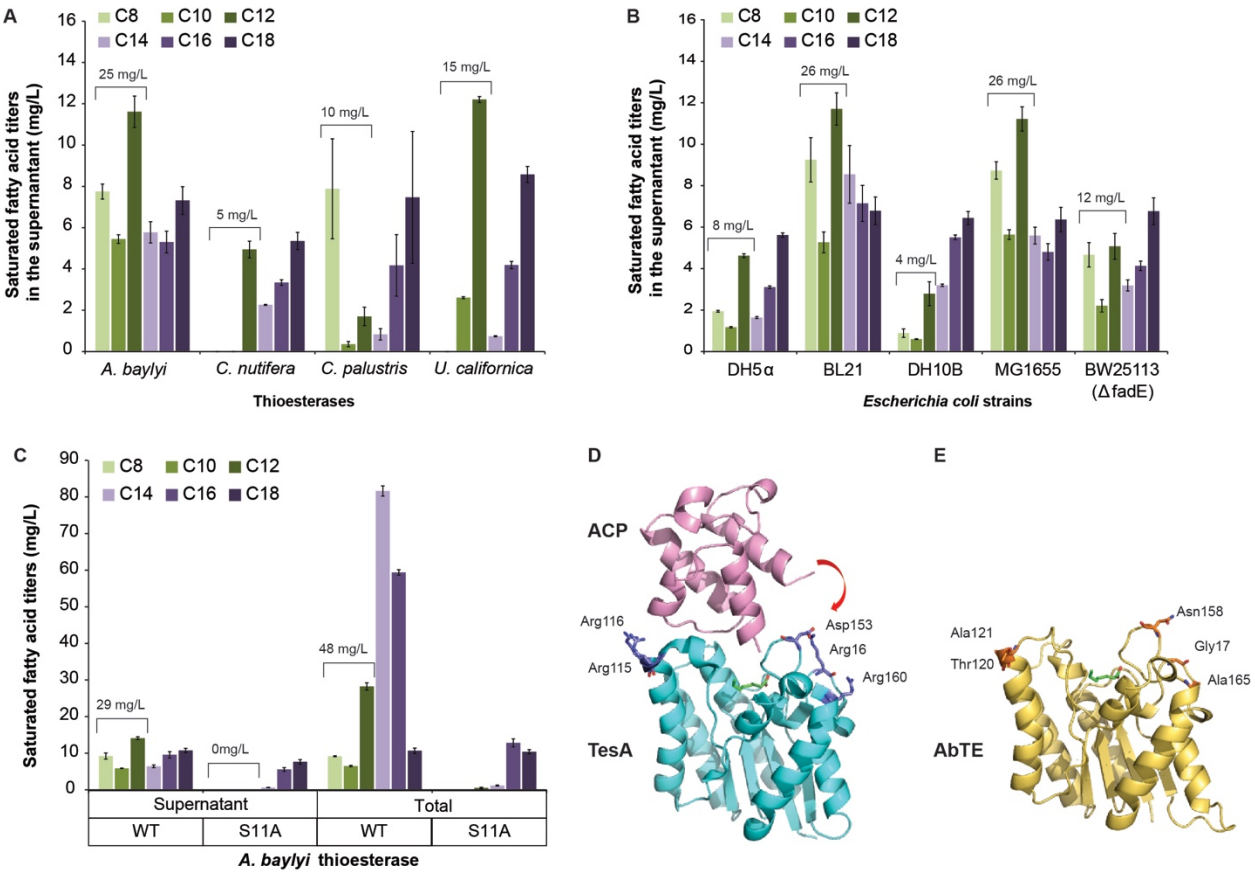



**Figure 5.2 *Acinetobacter baylyi* thioesterase (AbTE) product profile and interface with *Escherichia coli* ACP. A. Saturated fatty acid titers in the supernatant of *E. coli* (MG1655) expressing four different thioesterases. B. Saturated fatty acid titers in the supernatant of different *E. coli* strains expressing AbTE wild type. C. Total and secreted (supernatant) saturated fatty acid titers produced by *E. coli* when expressing wild type AbTE (WT) and inactive AbTE (S11A). The numbers above the brackets are the combined production of medium-chain fatty acids (C8-C12). The experiments were done in triplicate, and the error bars represent the standard deviation from the mean. D. Docking of *E.coli* ACP (magenta, PDB ID: 1FAE) and *E. coli* TesA (cyan, PDB ID 1U8U) identifies potential residues on the TesA surface that are important for interactions between TesA and ACP. E. Homology model of AbTE with surface residues equivalent to TesA highlighted.**

Based on the amino acids found on the TesA-ACP interface, we mutated positions 17, 120, 121, and 165 on AbTE to arginines to generate AbTE:G17R, AbTE:T120R, AbTE: A121R and AbTEA165R, and measured their fatty acid titers (**Fig. 3**). Expression of AbTE: G17R in *E. coli* more than doubled the AbTE wt MCFA titers in the supernatant to ~76mg/L, with octanoic, decanoic and dodecanoic acid produced at

30 mg/L, 18 mg/L and 28 mg/L, respectively (**Fig. 3A**). Expression of AbTE:A165R in *E. coli* resulted in slightly lower MCFA titers than AbTE wt. To determine if the effects of these mutations on MCFA titers were additive, we constructed all double mutants using AbTE:G17R as the starting point. Expression of AbTE:G17R/A165R in *E. coli* resulted in ~98 mg/L of MCFAs in the supernatant, a 29% increase in MCFA titers when compared to expression of the single mutant AbTE:G17R and a more than 3-fold improvement in MCFA titers when compared to AbTE wt (**Fig. 3B, Fig. SI3**). The major species in the AbTE:G17R/A165R saturated fatty acid profile was dodecanoic acid at 45 mg/L. The MCFA titers achieved by AbTE:G17R/A165R was unexpected as AbTE:A165R resulted in the lowest MCFA titers from all single mutants, demonstrating the value of testing all combinations of double mutants. Finally, we generated the triple mutants using AbTE:G17R/A165R as the starting point (**Fig. 3C, Fig. SI3**). At this stage, we also mutated the fifth position on the AbTE that varies from TesA, AbTE:N158D. This fifth position does not change the amino acid to a positively charged arginine, but a negatively charged aspartate. Nevertheless, aspartate could still form part of a stabilizing salt bridge. As Figure 3C shows, none of the triple mutants resulted in improved MCFA titers. It is worth noting that positions A121 and T120 are located on the other side of the AbTE binding pocket than G17 and A165. For completion, we also generated the remaining double mutants AbTE:T120R/A121R, AbTE:T120R/A121R, and AbTE:A121R/A165R and measured their saturated fatty acid titers in the supernatant (**Fig. 3D**). Two of these double mutants produced comparable MCFA titers with AbTE wild type, while AbTE:T120R/A121R produced higher levels at ~45 mg/L of medium chain fatty acids.

Although AbTE:G17R/A165R results in the highest MCFA titers (98 mg/L), AbTE:G17R had the highest percentage of MCFAs and lowest percentage of long chain fatty acids (**Table 1**). This trend is true whether analyzing the total fatty acid content or the fatty acid titers in the supernatant. We hypothesized that AbTE:G17R/A165R may express better than AbTE:G17R, thus explaining the higher MCFA titers produced by AbTE:G17R/A165R. Further, the fact that position 165 is located on the cell surface and that the mutation replaced a hydrophobic amino acid (alanine) with a charged positive amino acids (arginine) may also improve the solubility of AbTE:G17R/A165R over AbTE:G17R in the cytosol. An SDS-PAGE gel of *E. coli* expressing AbTE wt, AbTE:G17R or AbTE:G17R/A165R showed comparable soluble expression

of the three AbTE enzymes. The AbTE surface mutations did not affect the expression of the enzyme (**Fig. 4A**).

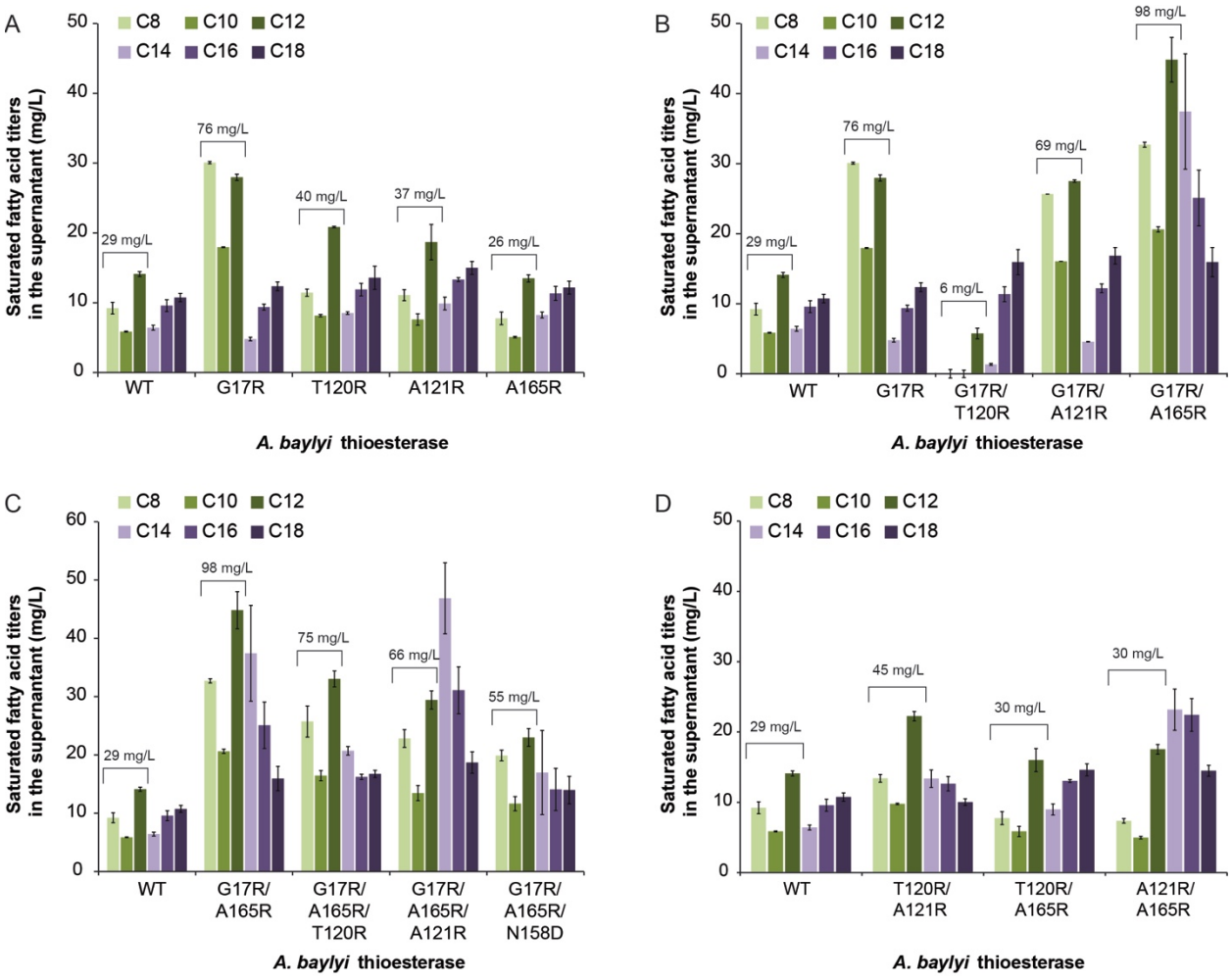



**Figure 5.3 Saturated fatty acid production of *E. coli* expressing *Acinetobacter baylyi* thioesterase (AbTE) mutants. Fatty acid production of *E .coli* expressing A. AbTE single mutants, B. AbTE double mutants, C. AbTE triple mutants, and D. Remaining double mutants. The numbers above the brackets are the combined production of medium-chain fatty acids (C8-C12). All experiments were done in triplicate, and the error bars represent the standard deviation from the mean.**

***Extending cultivation time to increase MCFA titers.*** Expecting to accumulate higher MCFA titers during a longer cultivation period, we increased the cultivation time from 24hrs to 72hrs before measuring fatty acid titers (**Fig. 4b**). We observed that although MCFA titers almost doubled between 24hrs and 72hrs when using AbTE wt, they remained the same when using AbTE:G17R and increased only 33% when using

AbTE:G17R/A165R. Interestingly, in the case of AbTE:G17R/A165R, the concentration of long chain fatty acids decreased almost by half between 24hrs and 72hrs. We speculate that long-chain fatty acids are more likely to be modified by endogenous FadD and be degraded via the β-oxidation pathway[19]. The low level increase in MCFA titers between 24hrs and 72hrs with the AbTE mutants when compared to AbTE wt led us to hypothesize that the AbTE mutants may lose most of their activity after 24hrs. It has been shown that adding solubility tags to heterologous proteins increase their viability inside the cell[20]. Therefore, we attached maltose binding protein (MBP) to the N-terminus of the AbTE mutants, the terminus at the opposite end of the ACP-thioesterase interface, and tested for fatty acid production. Expression of the AbTE:G17R or AbTE:G17R/A165R fused to MBT resulted in C16 and C18 fatty acid production, with no MCFAs being detected (**Fig. SI4**).

**Table 5.1 Acyl chain yields by mutant**

| Source | Thioesterase | *E. coli* strain | Media | C8:0 (%) | C10:0 (%) | C12:0 (%) | C14:0 (%) | C16:0 (%) | C18:0 (%) |
|---|---|---|---|---|---|---|---|---|---|
| Supernatant | AbTE WT | MG1655 | M9 | 16.5 | 10.5 | 25.2 | 11.5 | 17.1 | 19.2 |
| | AbTE:G17R | MG1655 | M9 | 29.3 | 17.5 | 27.3 | 4.7 | 9.1 | 12.1 |
| | AbTE:G17R/A165R | MG1655 | M9 | 18.5 | 11.7 | 25.4 | 21.2 | 14.2 | 9 |
| Total | AbTE WT | MG1655 | M9 | 4.7 | 3.3 | 14.4 | 41.7 | 30.4 | 5.5 |
| | AbTE:G17R | MG1655 | M9 | 16.14 | 9.92 | 25 | 29.4 | 13.9 | 5.7 |
| | AbTE:G17R/A165R | MG1655 | M9 | 9.4 | 6 | 18.8 | 38.7 | 18.6 | 8.5 |

***Detecting MCFAs in the E. coli supernatant via a GPCR-based sensor.*** Recently, we developed a G-protein coupled receptor (GPCR)-based MCFA sensor in yeast by expressing a mammalian olfactory receptor known to bind MCFAs, OR1G1, and coupling it to the yeast mating pathway resulting in green fluorescent protein expression upon MCFA detection on the cell surface[14] (**Fig. 4C**). Previously, we have shown that the MCFA sensor detects exogenously added MCFAs, but not long-chain (C16-C18) fatty acids[14]. Here, we set out to determine whether the MCFA sensor can detect different levels of microbially produced MCFAs in the supernatant of the producer microbe (**Fig. 4D**). The proposed MCFA detection has

the added difficulty that the sensor cell, *Saccharomyces cerevisiae*, and producer cell, *E. coli*, are different species, setting up the general question of whether a yeast whole cell biosensor can detect chemicals in the supernatant of other species. Further, since we envision using the MCFA sensor to screen chemical-producing microbes in a medium-throughput fashion, the MCFA production was carried out in 96-deep well plates rather than 5mL test tubes, resulting in slightly lower MCFA titers. To detect *E. coli* produced MCFA with the *S. cerevisiae* MCFA sensor we 1) lowered the *E. coli* supernatant pH from 7 to 4 as the MCFA sensor detects the protonated form of the MCFAs (pKa =4.5), and 2) diluted the *E. coli* supernatant 2-fold in fresh yeast media so that the microbially produced MCFA titers fall within the linear range of the sensor (C8: 2.7-36 mg/L; C10: 6-43 mg/L; C12: 0.2-50 mg/L). We see a linear correlation between MCFA sensor fluorescence and the MCFA titers produced by *E. coli* expressing the different AbTE variants (**Fig. 4D**). Interestingly, the fluorescent signal saturated between AbTE:G17R (C8: 24 mg/L; C10:11 mg/L; C12: 4 mg/L) and AbTE:G17R/A165R (C8: 34 mg/L; C10:19 mg/L; C12: 13 mg/L) even though concentrations of C8, C10 and C12 produced by AbTE:G17R and AbTE:G17R/A165R fall within the MCFA sensor range. The sensor is likely detecting the contribution of all three MCFAs together leading to early signal saturation.

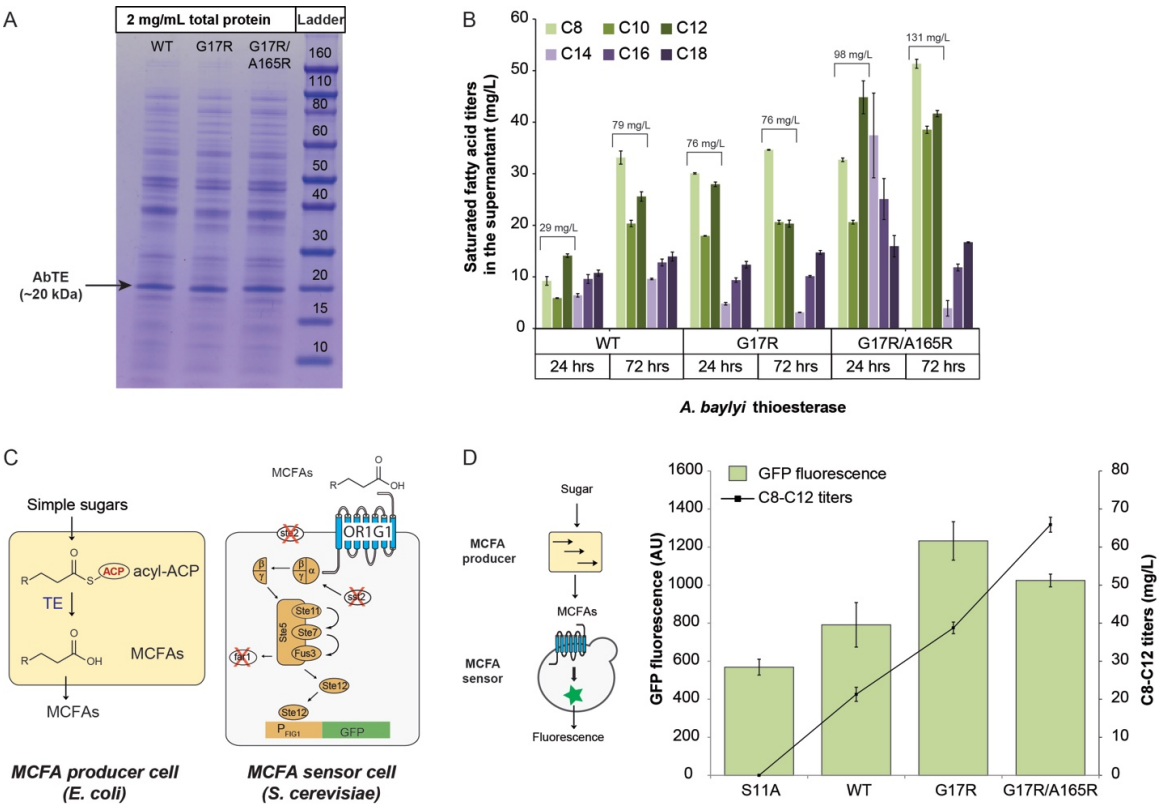



**Figure 5.4 Protein expression, time course analysis, and MCFA detection in the producer supernatant by MCFA sensor. A. SDS-PAGE gel of AbTE wt, AbTE:G17R and AbTE:G17R/A165R expressed in *E. coli.* The gel was equally loaded with 2 mg/ml of total protein. B. Time course of saturated fatty acid titers produced by *E. coli* expressing AbTE, AbTE:G17R, and AbTE:G17R/A165R. C. Schematic of MCFA producer cell and MCFA sensor cell. MCFA producer cell: glucose is fed to *E. coli* expressing a thioesterase (TE) to produce MCFAs that are secreted to the supernatant. MCFA sensor cell: *S. cerevisiae* expressing the GPCR OR1G1 (blue) detects MCFAs in the supernatant, transmits the signal via the yeast mating pathway (orange) resulting expression of green fluorescent protein (GFP). D. Detection of MCFAs by the MCFA sensor in the E. coli supernatant. Sensor GFP fluorescence and MCFA (C8-C12) titers as a function of AbTE variant expressed in *E. coli*.**

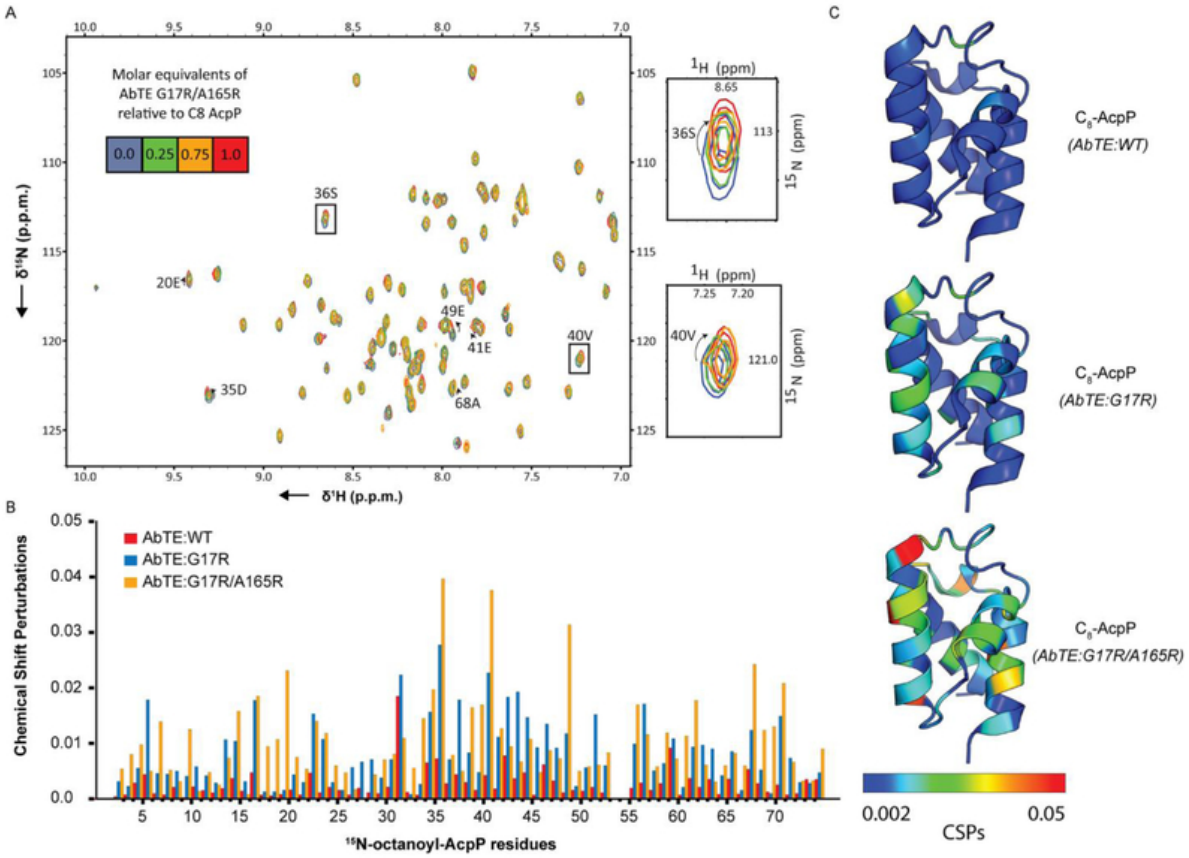



**Figure 5.5 Solution NMR analysis of molecular interactions between A. baylyi thioesterase (AbTE) and E. coli AcpP. A. HSQC titration of AbTE:G17R/A165R with E. coli 15N-octanoyl-AcpP. Right: Zoom of two selected cross-peaks 36S and 40V. B. CSPs were measured for each 5N-octanoyl-AcpP residue in the presence of AbTE:WT (red), AbTE:G17R (blue), or AbTE:G17R/A165R (orange) and plotted by residue number. C. CSPs are plotted onto the structure of AcpP (PDB ID = 2FAD) in the presence of AbTE:WT, AbTE:G17R, or AbTE:G17R/A165R. Shifts in 1H and 15N dimensions of 15N-octanoyl-AcpP upon titration with AbTE:WT.**

### 5.3 Discussion

This is the first example of a thioesterase engineered for improved MCFA product profile by matching the interface of the heterologous enzyme (thioesterase) to couple to the endogenous *E. coli* enzyme (ACP). Replacement of two small hydrophobic residues on the *A. baylyi* TE surface predicted to contact *E. coli* ACP with arginine, the amino acid found at the equivalent positions in *E. coli* TesA, resulted in more than 3-fold improvement in MCFA production. We speculate that improving the interface of AbTE and *E. coli* ACP enabled AbTE to more efficiently accept medium chain fatty acyl-ACPs, thus improving MCFA titers. As the SDS-PAGE gel shows, the improvement in MCFA titers does not come from improved

expression of AbTE mutants. Further improvement in MCFA titers can be generated by expressing the AbTE mutants in different *E. coli* hosts such as MG1655 Δ*fadD* or MG1655 Δ*fadE*, or using a more rich media.

In the future, engineering the interface of heterologous proteins to better match the interface of proteins they interact with in the production host could be applied to more distantly related proteins, such as plant thioesterases. Such an approach may prove even more beneficial for more distantly related enzymes, which may have only limited coupling with the endogenous proteins in the production host. However, the limited sequence identity between the heterologous and endogenous proteins may make it difficult to identify the key interfaces and residues to target for mutagenesis. Indeed, we attempted to use the interface engineering strategy to improve the product profile of a plant thioesterase, but the very low sequence identity of plant thioesterases to TesA or AbTE (**SI Fig. 1**) and the lack of crystal structure for plant thioesterases hindered this effort.

The MCFA yeast whole cell biosensor detected MCFAs in the supernatant of a different species, *E. coli*, setting up the stage for the sensor-guided engineering of MCFA producing microbes not only in *E. coli*, but potentially other industrially relevant hosts such as cyanobacteria and algae. All that was needed was dilution of the *E. coli* supernatant to set the MCFA levels within the linear range of the sensor, and dropping the pH of the supernatant to generate the protonated MCFA species detectable by the GPCR. Further, in this work, we have shown that MCFA production and detection can occur in 96-well plates, setting up the stage for the medium-throughput ($10^3$ samples per day) sensor-guided engineering of *E. coli* for improved MCFA production by screening large libraries of thioesterases coming from random or structure guided mutagenesis strategies. This screening throughput is one order of magnitude faster than current state of the art technology to detect microbially produced MCFA, i.e. gas chromatography/mass spectrometry (GC/MS). The MCFA sensor could be used to screen large libraries of MCFA producing microbes, to down select key mutants for testing using GC/MS. The medium-throughput screening of microbial MCFAs would be especially valuable for engineering of plant thioesterases that lack crystal structure and have very low sequence identity to bacterial thioesterases.

**Chapter 5**, is in full a reprint of the manuscript: Stephen Sarria, Thomas G. Bartholow, Adam Verga, Michael D. Burkart, and Pamela Peralta-Yahya “Matching Protein Interfaces for Improved Medium-Chain Fatty Acid Production.” ACS Synthetic Biology, (2018). The dissertation author is the primary coauthor of the manuscript in collaboration with Dr. Stephen Sarria.

## 5.4 References


1. Sarria S, Kruyer NS & Peralta-Yahya P (2017) Microbial synthesis of medium-chain chemicals from renewables. *Nat. Biotechnol* 35, 1158–1166.

2. Pfleger BF, Gossing M, and Nielsen J (2015) Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metabol. Eng* 29, 1–11.

3. Schweizer E, and Hofmann J (2004) Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems, *Microbiol. Mol. Biol. Rev* 68, 501–517.

4. Dellomonaco C, Clomburg JM, Miller EN, and Gonzalez R (2011) Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. *Nature* 476, 355–359.

5. McMahon MD, and Prather KL (2014) Functional screening and in vitro analysis reveal thioesterases with enhanced substrate specificity profiles that improve short-chain fatty acid production in Escherichia coli. *Appl. Environ. Microbiol* 80, 1042–1050.

6. Zheng Y, Li L, Liu Q, Qin W, Yang J, Cao Y, Jiang X, Zhao G, and Xian M (2012) Boosting the free fatty acid synthesis of Escherichia coli by expression of a cytosolic Acinetobacter baylyi thioesterase. *Biotechnol. Biofuels* 5, 76.

7. Jing F, Cantu DC, Tvaruzkova J, Chipman JP, Nikolau BJ, Yandeau-Nelson MD, and Reilly PJ (2011) Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity. *BMC Biochem*. 12, 44.

8. Pollard MR, Anderson L, Fan C, Hawkins DJ, and Davies HM (1991) A specific acyl-ACP thioesterase implicated in medium-chain fatty acid production in immature cotyledons of Umbellularia californica. *Arch. Biochem. Biophys* 284, 306–312.

9. Leber C, and Da Silva NA (2014) Engineering of Saccharomyces cerevisiae for the synthesis of short chain fatty acids. *Biotechnol. Bioeng* 111, 347–358.

10. Choi YJ, Lee SY, (2013) Microbial production of short-chain alkanes. *Nature* 502, 571–574.

11. Kovacic F, Granzin J, Wilhelm S, Kojic-Prodic B, Batra-Safferling R, and Jaeger KE (2013) Structural and functional characterisation of TesA - a novel lysophospholipase A from Pseudomonas aeruginosa. *PloS one* 8, e69125.

12. Grisewood MJ, Hernández-Lozada NJ, Thoden JB, Gifford NP, Mendez-Perez D, Schoenberger HA, Allan MF, Floy ME, Lai R-Y, Holden HM, Pfleger BF, and Maranas CD (2017) Computational Redesign of Acyl-ACP Thioesterase with Improved Selectivity toward Medium-Chain-Length Fatty Acids. *ACS Catal*. 7, 3837–3849.

13. Zornetzer GA, Tanem J, Fox BG and Markley JL (2010) The Length of the Bound Fatty Acid Influences the Dynamics of the Acyl Carrier Protein and the Stability of the Thioester Bond. *Biochemistry* 49, 470–477.

14. Worthington AS, Hur GH and Burkart MD (2011) Activity-guided engineering of natural product carrier proteins. *Mol. Biosyst* 7, 365–370.

15. Voelker TA and Davies HM (1994) Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of a plant medium-chain acyl-acyl carrier protein thioesterase. *J. Bacteriol* 176, 7320–7327.

16. Bond-Watts BB, Bellerose RJ & Chang MCY (2011) Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat. Chem. Biol* 7, 222–227.

17. Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, and Keasling JD (2010) Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* 463, 559–562.

18. Kozakov D, Beglov D, Bohnuud T, Mottarella SE, Xia B, Hall DR, and Vajda S (2013) How good is automated protein docking?. *Proteins: Struc., Func., Bioinf* 81, 2159–2166.

19. Kelley LA, Mezulis S, Yates CM, Wass MN & Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc* 10, 845–858.

20. Ford TJ, and Way JC (2015) Enhancement of E. coli acyl-CoA synthetase FadD activity on medium chain fatty acids. *PeerJ* 3, e1040.

21. Shiue E, and Prather KL (2014) Improving D-glucaric acid production from myo-inositol in E. coli by increasing MIOX stability and myo-inositol transport. *Metab. Eng* 22, 22–31.

22. Chan DI, and Vogel HJ (2010) Current understanding of fatty acid biosynthesis and the acyl carrier protein, *Biochem. J* 430, 1–19.

23. Zhang YM, Wu B, Zheng J, and Rock CO (2003) Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction. *J. Biol. Chem* 278, 52935–52943.

24. Zhang YM, Rao MS, Heath RJ, Price AC, Olson AJ, Rock CO, and White SW (2001) Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III, *J. Biol. Chem* 276, 8231–8238.

25. Rafi S, Novichenok P, Kolappan S, Stratton CF, Rawat R, Kisker C, Simmerling C, and Tonge PJ (2006) Structure of acyl carrier protein bound to FabI, the FASII enoyl reductase from Escherichia coli. *J. Biol. Chem* 281, 39285–39293.

26. Finzel K, Nguyen C, Jackson DR, Gupta A, Tsai SC, and Burkart MD (2015) Probing the Substrate Specificity and Protein-Protein Interactions of the E. coli Fatty Acid Dehydratase. FabA, *Chem. Biol* 22, 1453–1460.

27. Nguyen C, Haushalter RW, Lee DJ, Markwick PR, Bruegger J, Caldara-Festin G, Finzel K, Jackson DR, Ishikawa F, O'Dowd B, McCammon JA, Opella SJ, Tsai SC and Burkart MD (2014) Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* 505, 427–431.

28. Williamson MP (2013) Using chemical shift perturbation to characterize ligand binding. *Prog Nucl Mag Res Sp* 73, 1–16.

29. Feng YB, Wang YY, Liu J, Liu YH, Cao XP & Xue S (2017) Structural Insight into Acyl-ACP Thioesterase toward Substrate Specificity Design. *ACS Chem Biol* 12, 2830–2836.

30. Yu G, Nguyen TT, Guo Y, Schauvinhold I, Auldridge ME, Bhuiyan N, Ben-Israel I, Iijima Y, Fridman E, Noel JP, and Pichersky E (2010) Enzymatic functions of wild tomato methylketone synthases 1 and 2, *Plant Physiol*. 154, 67–77.

31. Kosa NM, Haushalter RW, Smith AR & Burkart MD (2012) Reversible labeling of native and fusion-protein motifs. *Nat Methods* 9, 981–984.

32. Haushalter RW, Worthington AS, Hur GH & Burkart MD (2008) An orthogonal purification strategy for isolating crosslinked domains of modular synthases. *Bioorg Med Chem Lett* 18, 3039–3042.

33. Torella JP, Ford TJ, Kim SN, Chen AM, Way JC, and Silver PA (2013) Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *PNAS* 110, 11290–11295.

## 5.6 Methods

### Methods

Plasmid construction.

Non-codon optimized Acinetobacter baylyi TE (AbTE), S. cerevisiae codon-optimized Cuphea palustris TE (CpTE), E. coli codon-optimized Umbellularia californica TE (UcTE) were commercially synthesized and cloned under PTRC in pMB1-PTRC-AgGPPS-(GSG)2-AgPS (pSS185) between NcoI/XmaI to generate pMB1-PTRC-AbTE (pSS192), pMB1-PTRC-CpTE (pSS183), and pMB1-PTRC-UcTE (pSS193). S. cerevisiae codon-optimized CnTE was amplified from pESC-LEU2-PTEF1-PHXT7-CnTE (pSS81) with primers SS455/SS456 and cloned under PTRC in pSS185 between NcoI/XmaI to generate pMB1-PTRC-CnTE (pSS174). AbTE mutants were generated using QuikChange protocol with some modifications (SI). Details on templates and primers used are in Table SI4. For protein expression, a C-terminal His6-tag was introduced into AbTE:WT, AbTE:G17R and AbTE:G17R/A165R using primers TB1/TB2 and cloned into pET-28b (amplified using primers TB3/TB4) to generate pET-28b -AbTE:WT, pET-28b -AbTE:G17R, and pET-28b -AbTE:G17R/A165R.

AbTE and ACP expression and purification.

pET-AbTE:WT, pET-AbTE:G17R, and pET-AbTE:G17R/A165R were transformed in E. coli BL21 (DE3) and grown in the presence of 50 mg/L of kanamycin. Cells were induced with 1mM IPTG at OD600=0.8 and grown for 16 to 18 hours at 16° C. Cells pellets were resuspended in 50 mM Hepes (pH 7.4), 250 mM NaCl, and 10% glycerol before lysis by sonication and clarification at 22,000 rcf. Clarified lysate was allowed to batch bind the Ni-NTA resin for 20 minutes followed by washing with buffer containing 25 mM imidazole. Final elution was performed with buffer containing 250mM imidazole, followed by dialysis into 50 mM Tris buffer (pH 7.4), 150 mM NaCl and 10% glycerol. After concentration to ~2mL, the enzymes were purified on a GE Superdex 200 gel filtration column and the fractions containing the desired protein were checked by UV trace and SDS PAGE before concentration. The same procedure was followed for the 15N-AcpP,

however, AcpP was grown in 1g 15N ammonium chloride, 4g of 12C glucose, 1L of M9 media, and 50mg of kanamycin.

15N-octanoyl-AcpP synthesis.

15N-AcpP was incubated overnight with Pseudomonas aeruginosa ACP hydrolase to generate pure apo 15N-AcpP as confirmed by a urea PAGE31. C8-pantethenamide32 was loaded onto apo 15N-AcpP to generate 15N-octanoyl-AcpP using 12.5mM MgCl2, 10mM ATP, 0.1μM E. coli CoaA, 0.1μM E. coli CoaD, 0.1μM E. coli CoaE, 0.2μM Bacillus subtilis surfactin phosphopantetheinyl transferase, 0.02% Triton X, 0.01% azide, and 0.1% tris(2-carboxyethyl)-phosphine (TCEP).

NMR titration experiments.

AbTE:WT, AbTE:G17R, AbTE:G17R/A165R and 15N-octanoyl-AcpP were purified separately on a Superdex 75 gel filtration column into 50 mM pH 7.4 potassium phosphate with 0.01% azide and 0.5 mM TCEP. Samples were prepared to a volume of 500μL with 50 μL of D2O. The AbTE:WT, AbTE:G17R, and AbTE:G17R/A165R experiments were performed with 0.75 mM, 0.101 mM, and 0.74 mM 15N-octanoyl-AcpP. The AcpPs were titrated to a final concentration of 1.5, 2.0, and 1 molar equivalents of AbTE: WT, AbTE :G17R, and AbTE G17R/A165R, respectively. Experiments were performed on a Bruker Avance 800MHz spectrometer, in 50 mM potassium phosphate pH 7.4 with 0.5mM TCEP and 0.1% sodium azide. Chemical shift perturbations were calculated using the equation $CSP=(12)[\delta 2H+(\alpha x \delta N)2----------\surd$ where α=0.2.

MCFA production, derivatization and quantification.

MCFA Production: Overnight cultures of E.coli MG1655 expressing AbTE:WT or AbTE variants were diluted 1:50 in 5 mL of M9 media (0.5% glucose, amp100) and grown at 37°C, 250 r.p.m. until reaching an OD600 = 0.3–0.4. The cells were then induced with 500 μM of IPTG (500 mM stock) and grown at 30°C, 250 r.p.m.

for 24 or 72 hrs. Fatty acid analysis: For secreted fatty acids, E. coli cultures were vortexed for 3 sec, 600 µL of culture removed and centrifuged for 10 min at 7354g. Next, 400 µL of the supernatant was removed for derivatization. For total fatty acids, 400 µL of culture was used for derivatization. Fatty acid derivatization: Fatty acids were derivatized to fatty acid methyl esters and analyzed via GC/MS as described in Torella et. al., 201333 with some modifications. To the 400 µL of sample, 50 µL of 10% (wt/vol) NaCl, 50 µL of glacial acetic acid, 20 µL of 90.5 mg/L nonanoic acid (internal standard), and 200 µL of ethyl acetate were added and the mixture was vortexed for 5 sec. The mixture was then centrifuged at 12,098g for 10 min. Methyl esters were generated by mixing 100 µL of the ethyl acetate layer with 900 µL of a 30:1 mixture of methanol and 37% (vol/vol) HCl in a 2 mL microcentrifuge tube, vortexed for 5 sec, and incubated at 50°C for 1 hr. After cooling to room temperature, 500 µL of water and 500 µL of hexanes were added. The mixture was vortexed for 5 sec, 100µL of the hexane layer was taken and mixed with 400 µL of ethyl acetate for analysis via GC-MS. FAME quantification: The samples were analyzed using Agilent 7890A/Agilent 5975 MS detector using a DB-5MS column. The inlet temperature was set to 300°C, flow at 1 mL/min, the oven at 70°C for 1 min, ramp at 30°C/min to 290°C, and held for 1 min at 290°C. Standard curves of C8–C18 fully saturated FAMEs (Alfa Aesar/TCI) were used for sample quantification.

### 5.7 Supplemental Information

**Table SI1.** Table of strains

| Strain # | Description | Reference |
|---|---|---|
| PPY11 | W303: MATa, leu2−3, trp1−1, can1−100, ura3−1, ade2−1, his3−11 | ATCC 208352 |
| PPY140 | PPY11 Δ*far1*, Δ*sst2*, Δ*ste2* | Mukherjee, 2015 |
| PPY643 | PPY140, pESC-His3-$P_{TEF1}$-OR1G1, pRS415-Leu2-$P_{FIG1}$-GFP | Mukherjee, 2015 |
| PPY252 | DH10B | Invitrogen |
| PPY251 | MG1655 | ATCC 47076 |
| PPY260 | DH5α | Invitrogen |
| PPY1151 | BW25113 Δ*fadE*739::kan | Keio collection |
| PPY1236 | PPY252, pMB1-$Amp^r$-$P_{TRC}$-AbTE:WT | This study |
| PPY1331 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:WT | This study |
| PPY1332 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:S11A | This study |
| PPY1333 | PPY251, pMB1-Ampr-$P_{TRC}$-AbTE:G17R | This study |
| PPY1334 | PPY251, pMB1-Ampr-$P_{TRC}$-AbTE:A165R | This study |
| PPY1335 | PPY251, pMB1-Ampr-$P_{TRC}$-AbTE:A121R | This study |
| PPY1336 | PPY251, pMB1-Ampr-$P_{TRC}$-AbTE:T120R | This study |
| PPY1337 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A121R | This study |
| PPY1338 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/T120R | This study |
| PPY1339 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:T120R/A121R | This study |
| PPY1340 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A165R | This study |
| PPY1341 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:A121R/A165R | This study |
| PPY1342 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:T120R/A165R | This study |
| PPY1396 | PPY251, pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A165R/T120R | This study |

| | | |
|---|---|---|
| PPY1397 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-AbTE:G17R/A165R/A121R | This study |
| PPY1398 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-AbTE:G17R/A165R/N158D | This study |
| PPY1403 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-CnTE | This study |
| PPY1404 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-CpTE | This study |
| PPY1405 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-UcTE | This study |
| PPY1406 | BL21 | NEB |
| PPY1407 | PPY260, pMB1-Amp$^r$-P$_{TRC}$-AbTE:WT | This study |
| PPY1408 | PPY1152, pMB1-Amp$^r$-P$_{TRC}$-AbTE:WT | This study |
| PPY1409 | PPY1461, pMB1-Amp$^r$-P$_{TRC}$-AbTE:WT | This study |
| PPY1503 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-AbTE:G17E | This study |
| PPY1504 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-AbTE:G17E/A165E | This study |
| PPY1505 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-AbTE:G17E/A165R | This study |
| PPY1506 | PPY251, pMB1-Amp$^r$-P$_{TRC}$-AbTE:G17R/A165E | This study |

**Table SI2.** Table of plasmids

| Strain # | Plasmid name | Description | Reference |
|---|---|---|---|
| PPY269 | pKM260 | pESC-His3-$P_{TEF1}$-OR1G1- $P_{ADH1}$ | Mukherjee, 2015 |
| PPY586 | pKM586 | pRS415-Leu2-$P_{FIG1}$-GFP | Mukherjee, 2015 |
| PPY1023 | pSS185 | pMB1-$Amp^r$-$P_{TRC}$-AgGPPS-$(GSG)_2$-AgPS | Sarria, 2014 |
| PPY1090 | pSS174 | pMB1-$Amp^r$-$P_{TRC}$-CnTE | This study |
| PPY1148 | pSS183 | pMB1-$Amp^r$-$P_{TRC}$-CpTE | This study |
| PPY1236 | pSS192 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:WT | This study |
| PPY1237 | pSS193 | pMB1-$Amp^r$-$P_{TRC}$-UcTE | This study |
| PPY1310 | pSS196 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R | This study |
| PPY1311 | pSS197 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:T120R | This study |
| PPY1312 | pSS198 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:A121R | This study |
| PPY1320 | pSS199 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:A165R | This study |
| PPY1321 | pSS200 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A165R | This study |
| PPY1322 | pSS201 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:A121R/A165R | This study |
| PPY1323 | pSS202 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:T120R/A165R | This study |
| PPY1326 | pSS203 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A121R | This study |
| PPY1327 | pSS204 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/T120R | This study |
| PPY1328 | pSS205 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:T120R/A121R | This study |
| PPY1329 | pSS206 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:S11A | This study |
| PPY1393 | pSS208 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A165R/T120R | This study |
| PPY1394 | pSS209 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A165R/A121R | This study |
| PPY1395 | pSS210 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A165R/N158D | This study |
| PPY1499 | pSS251 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17E | This study |
| PPY1500 | pSS252 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17E/A165E | This study |
| PPY1501 | pSS253 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17E/A165R | This study |
| PPY1502 | pSS254 | pMB1-$Amp^r$-$P_{TRC}$-AbTE:G17R/A165E | This study |
| BL136 | pTB1 | pET-28b -AbTE:WT | This study |
| BL137 | pTB2 | pET-28b –AbTE:G17R | This study |
| BL138 | pTB3 | pET-28b -AbTE:G17R/A165R | This study |

**Table SI3.** Table of primers

| Name | Sequence (codon change) |
|---|---|
| SS455 | TAACAATTTCACACAGGAAACAGACCATGGATGTTGCCAGATTGG TCTATG |
| SS456 | CCTGCAGGTCGACTCTAGAGGATCCCCGGGTTATTTAGATTCAGT TGGATGCAAACCC |
| SS547 | GAAAATACCACCAAATTATGGCACT**AGA**TATAGTCAGGCATTTG |
| SS541 | CGACAGTCTGAGTGCGGGTTAT**AGA**ATTAACCCCGAACAGGGCT GG |
| SS542 | CCAGCCCTGTTCGGGGTTAAT**TCT**ATAACCCGCACTCAGACTGTC G |
| SS545 | GCAAAATGACCAGATCCATCCAAAT**CGC**AAAGCCCAGTCAATCTT GCTAAATAACG |
| SS546 | CGTTATTTAGCAAGATTGACTGGGCTTT**GCG**ATTTGGATGGATCT GGTCATTTTGC |
| SS547 | GAAAATACCACCAAATTATGGCACT**AGA**TATAGTCAGGCATTTG |
| SS548 | CAAATGCCTGACTATA**TCT**AGTGCCATAATTTGGTGGTATTTTC |
| SS549 | GAAAATACCACCAAATTATGGC**AGG**GCCTATAGTCAGGCATTTG |
| SS550 | CAAATGCCTGACTATAGGC**CCT**GCCATAATTTGGTGGTATTTTC |
| SS583 | GAAAATACCACCAAATTATGGCAGG**AGA**TATAGTCAGGCATTTG |
| SS584 | CAAATGCCTGACTATA**TCT**CCTGCCATAATTTGGTGGTATTTTC |
| SS601 | CAAAACCATTCTTATCTTAGGCGAC**GCT**CTGAGTGCGGGTTATGG CATTAACC |
| SS602 | GGTTAATGCCATAACCCGCACTCAG**AGC**GTCGCCTAAGATAAGA ATGGTTTTG |
| SS608 | GGCTGGACACAAAAGTCTAATGCAA**GAT**GACCAGATCCATCCAA ATGC |
| SS609 | GCATTTGGATGGATCTGGTC**ATC**TTGCATTAGACTTTTGTGTCCA GCC |
| SS617 | CGACAGTCTGAGTGCGGGTTAT**GAA**ATTAACCCCGAACAGGGCT GG |
| SS618 | CCAGCCCTGTTCGGGGTTAAT**TTC**ATAACCCGCACTCAGACTGTC G |
| SS619 | GCAAAATGACCAGATCCATCCAAAT**GAA**AAAGCCCAGTCAATCTT GCTAAATAACG |
| SS620 | CGTTATTTAGCAAGATTGACTGGGCTTT**TTC**ATTTGGATGGATCT GGTCATTTTGC |
| TB1 | TGTTTAACTTTAAGAAGGAGATATACCatgggcaaaaccattcttatcttag |
| TB2 | CTTTCGGGCTTTGTTAGCAGCCGGATCttaatggtgatggtgatggtgtaaag |
| TB3 | ctttacaccatcaccatcaccattaaGATCCGGCTGCTAACAAAGCCCGAAAG |
| TB4 | ctaagataagaatggttttgcccatGGTATATCTCCTTCTTAAAGTTAAACA |

**Table SI4.** Site-directed mutagenesis primers and templates

| Strain # | Plasmid name | Mutation | Template | Mutagenesis primers |
|---|---|---|---|---|
| PPY1310 | pSS196 | G17R | PPY1236 (pSS192) | SS541/542 |
| PPY1311 | pSS197 | T120R | PPY1236 (pSS192) | SS549/550 |
| PPY1312 | pSS198 | A121R | PPY1236 (pSS192) | SS547/548 |
| PPY1320 | pSS199 | A165R | PPY1236 (pSS192) | SS545/546 |
| PPY1321 | pSS200 | G17R/A165R | PPY1310 (pSS196) | SS545/546 |
| PPY1322 | pSS201 | A121R/A165R | PPY1312 (pSS198) | SS545/546 |
| PPY1323 | pSS202 | T120R/A165R | PPY1320 (pSS199) | SS549/550 |
| PPY1326 | pSS203 | G17R/A121R | PPY1312 (pSS198) | SS541/542 |
| PPY1327 | pSS204 | G17R/T120R | PPY1311 (pSS197) | SS541/542 |
| PPY1328 | pSS205 | T120R/A121R | PPY1311 (pSS197) | SS583/584 |
| PPY1329 | pSS206 | S11A | PPY1236 (pSS192) | SS601/602 |
| PPY1393 | pSS208 | G17R/A165R/T120R | PPY1321 (pSS200) | SS549/550 |
| PPY1394 | pSS209 | G17R/A165R/A121R | PPY1321 (pSS200) | SS547/548 |
| PPY1395 | pSS210 | G17R/A165R/N158D | PPY1321 (pSS200) | SS608/609 |
| PPY1499 | pSS251 | G17E | PPY1310 (pSS196) | SS617/618 |
| PPY1500 | pSS252 | G17E/A165E | PPY1499 (pSS251) | SS619/620 |
| PPY1501 | pSS253 | G17E/A165R | PPY1321 (pSS200) | SS617/618 |
| PPY1502 | pSS254 | G17R/A165E | PPY1310 (pSS196) | SS619/620 |

| | *Escherichia coli* TesA | *Acinetobacter baylyi* TE |
|---|---|---|
| *Escherichia coli* TesA | | 38.3% |
| *Acinetobacter baylyi* TE | 38.3% | |
| *Cocos nucifera* TE | 19.0% | 16.9% |
| *Umbellularia californica* TE | 21.1% | 14.7% |
| *Cuphea palustris* TE | 16.7 | 16.9% |

**Figure SI1.** Percent sequence identity of thioesterases uses in this work. Bacterial thioesterases (blue boxes); plant thioesterases (green boxes).

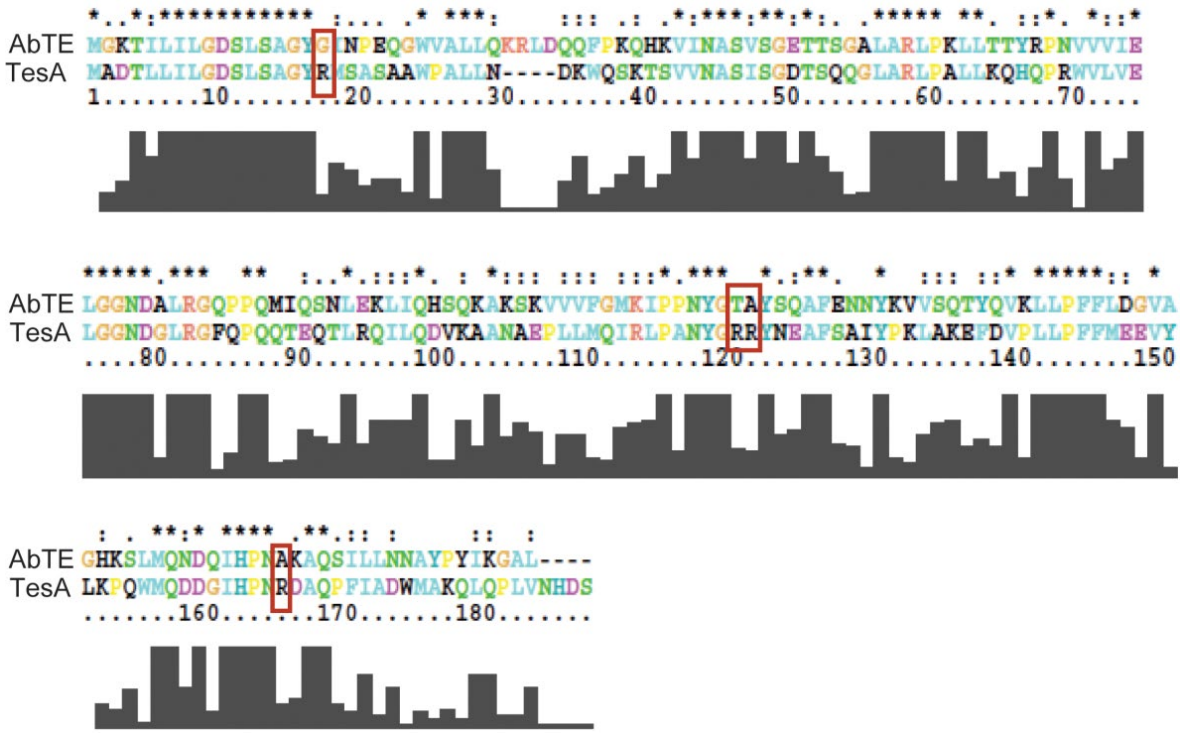



**Figure SI2.** Amino acid sequence alignment of AbTE wt and *E. coli* TesA with the signal peptide removed. Red boxes indicate amino acids targeted for mutagenesis.

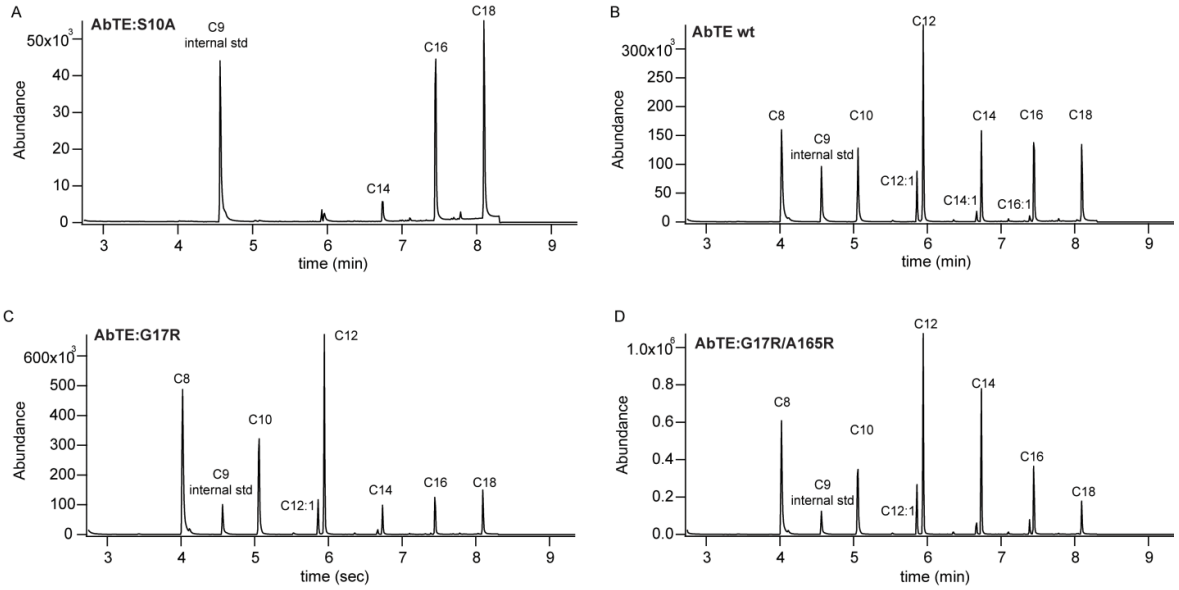



**Figure SI3.** Gas chromatograms of secreted fatty acids produced by *E. coli* expressing AbTE and AbTE variants. A) AbTE:S11A (inactive enzyme), B) Wild-type AbTE) C) AbTE:G17R D) AbTE:G17R/A165R. Single Ion Monitoring: 74 and 87.

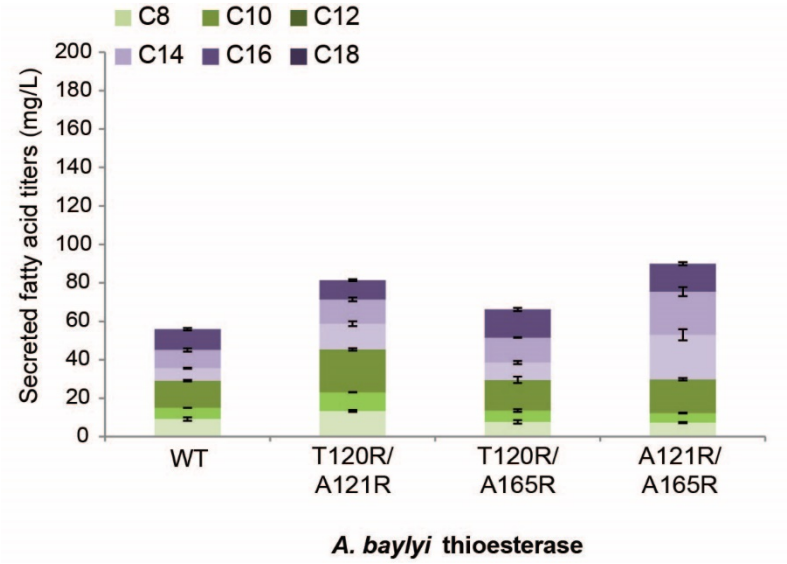



**Figure SI4.** Secreted fatty acid production of *E. coli* expressing AbTE:WT, AbTE: T120R/A121R, AbTE: T120R/A165R, AbTE: A121R/A165R.

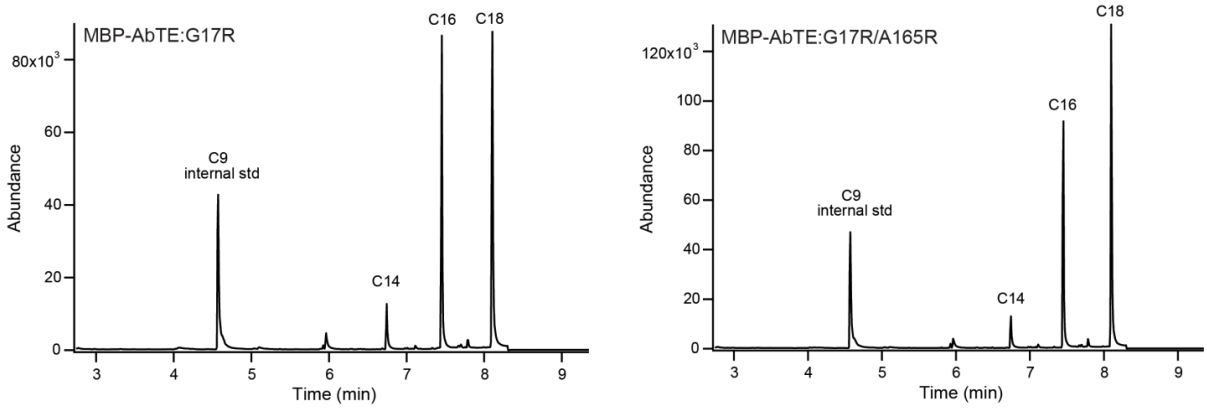



**Figure SI5**: Effect of fusing maltose binding protein (MBP) to AbTE mutants. Gas chromatograms of MBP-AbTE:G17R (left) and MBP-AbTE:G17R/A165R (right). Single Ion Monitoring: 74 and 87.

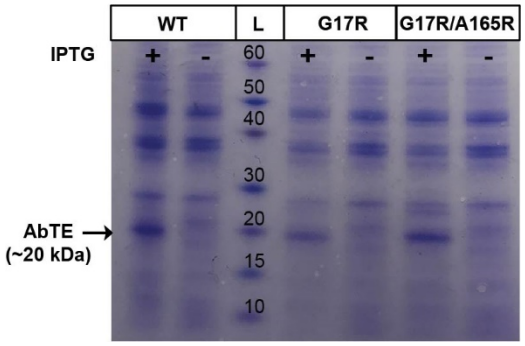



**Figure SI6**: Protein levels of Acitenobacter baylyi TE expressed in *E. coli.* Coomassie stained SDS-PAGE gel of induced (+IPTG) and uninduced (-IPTG) *E. coli* cultures expressing AbTE:WT, AbTE:G17R and AbTE:G17R/A165R.

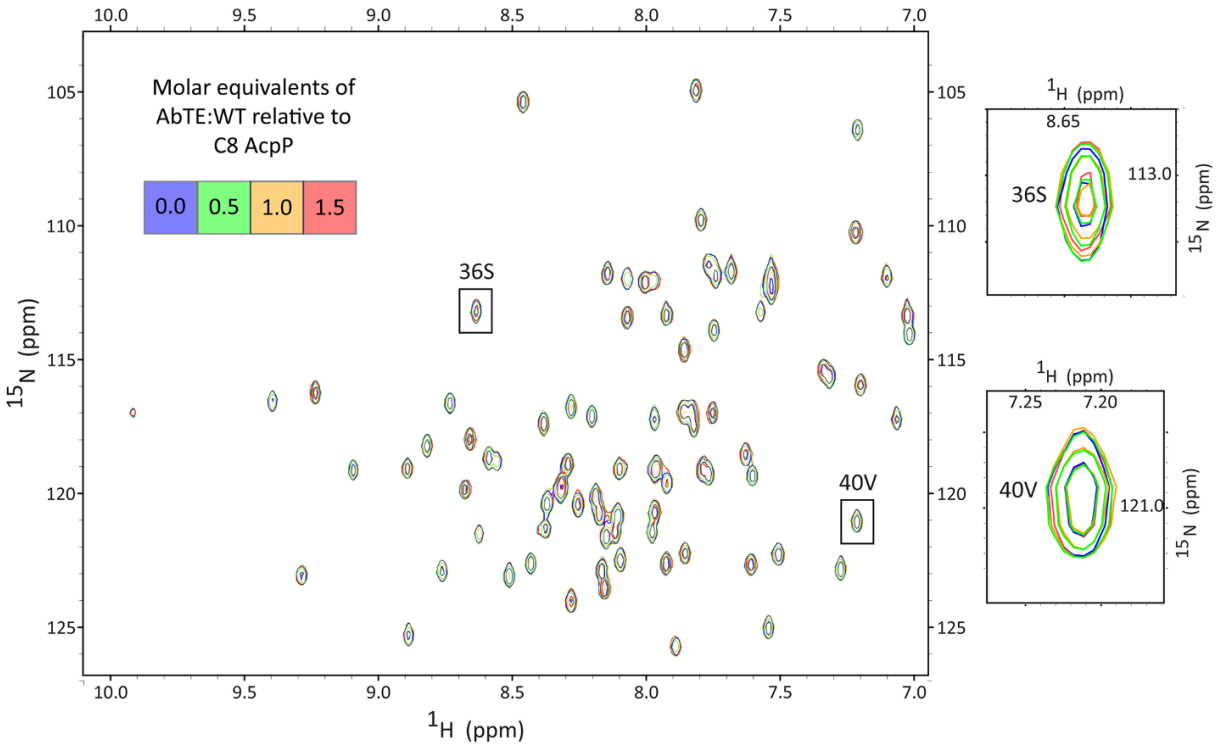



**Figure SI7**: HSQC titration of wild type AbTE with *E. coli* $^{15}$N-octonoyl-AcpP. The titration was performed up to a 1.5 molar ratio of ACP:TE in order to confirm that no significant perturbations occurred due to interaction.

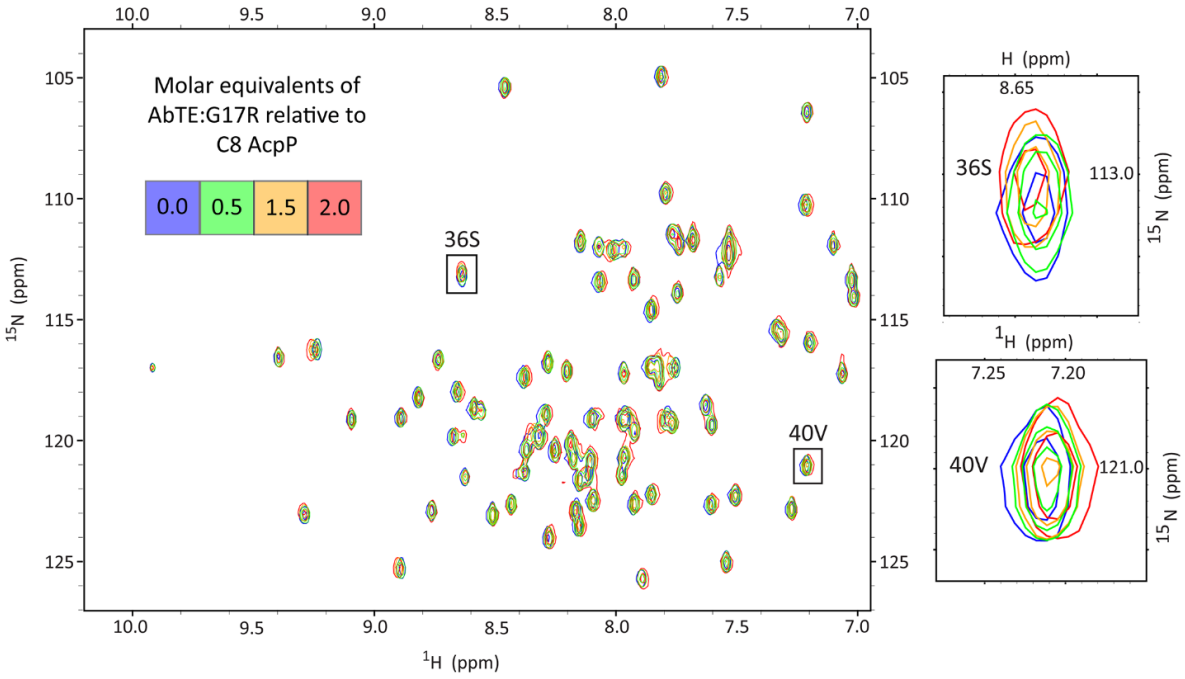



**Figure SI8**: HSQC titration of AbTE:G17R with *E. coli* $^{15}$N-octonoyl-AcpP. The titration was performed to 2 molar equivalents of partner protein.

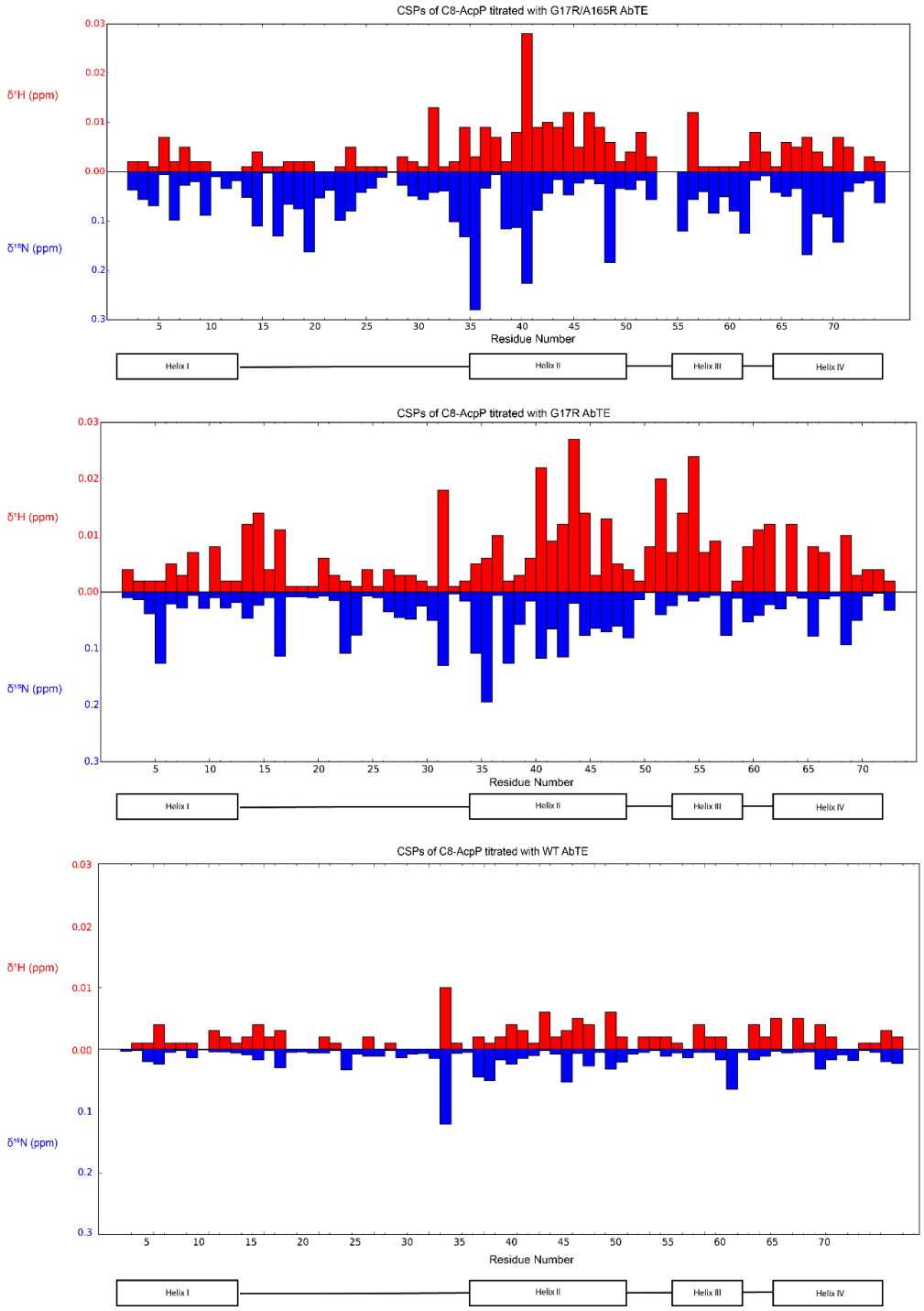



**Figure SI9**: HSQC titration of AbTE:G17R/A165R with *E. coli* $^{15}$N-octonoyl-AcpP perturbations separated by resonance in $^{1}$H and $^{15}$N dimension This shows the magnitude of each chemical shift before the averaging by the CSP equation: $CSP = \sqrt{\left(\frac{1}{2}\right)[\delta_H^2 + (\alpha x \delta_N)^2}$ where α=0.2.

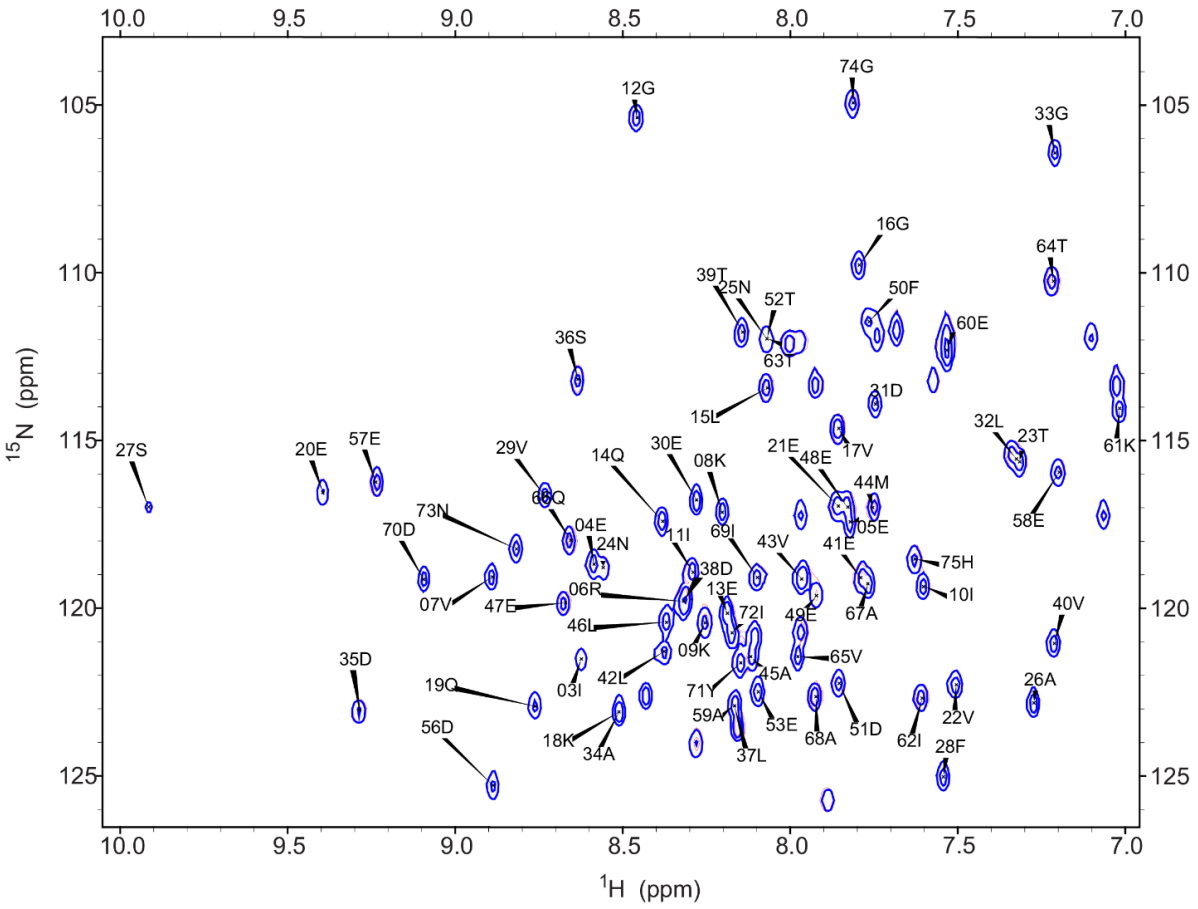



**Figure SI10:** The HSQC spectrum of $^{15}$N-octonoyl-AcpP. Peak assignments from previous with with C8-AcpP[3].

**Sequences**

***Acetobacter baylyi* TE (non-codon optimized, N-terminal signal peptide removed)**

ATGGGCAAAACCATTCTTATCTTAGGCGACAGTCTGAGTGCGGGTTATGGCATTAACCCCGAACAGG
GCTGGGTCGCTTTATTACAAAAACGTCTGGATCAACAATTTCCCAAGCAGCATAAAGTCATTAATGCC
AGTGTAAGTGGGGAAACCACCAGTGGTGCTTTAGCTCGTTTACCCAAACTACTTACTACTTATCGACC
TAATGTGGTGGTCATTGAGCTTGGTGGTAATGATGCATTAAGAGGACAACCGCCTCAAATGATTCAA
AGTAATCTGGAAAAATTAATCCAGCACAGCCAAAAGGCAAAATCTAAAGTCGTGGTGTTTGGAATGAA
AATACCACCAAATTATGGCACTGCCTATAGTCAGGCATTTGAAAATAATTATAAGGTAGTGAGTCAAA
CATATCAGGTTAAGTTGTTGCCATTTTTTCTTGATGGTGTGGCTGGACACAAAAGTCTAATGCAAAAT
GACCAGATCCATCCAAATGCCAAAGCCCAGTCAATCTTGCTAAATAACGCATACCCATATATTAAAGG
CGCTTTATAA

***Cocos nucifera* thioesterase FatB3 (*S. cerevisiae* codon-optimized, N-term signal peptide removed):**

ATGTTGCCAGATTGGTCTATGTTGTTGGCTGCTATTAGAACCATTTTCTCCGCTGCTGAGAAGCAATG
GACTTTGCTCGATTCTAAGAAGCGAGGTGCTGATGCTGTTGCTGATGCTTCTGGTGTTGGTAAGATG
GTTAAGAATGGCTTGGTCTACAGACAGAACTTCTCCATTAGATCCTACGAAATTGGTGTTGATAAGAG
AGCTTCCGTTGAGGCTTTGATGAATCATTTCCAAGAAACTTCTTTGAATCATTGTAAGTGTATTGGTTT
GATGCATGGTGGTTTCGGTTGTACTCCAGAAATGACTAGAAGAAATTTGATTTGGGTTGTTGCTAAGA
TGTTGGTTCATGTTGAAAGATACCCCTGGTGGGGTGATGTTGTTCAAATTAATACTTGGATTTCTTCTT
CTGGTAAGAATGGTATGGGTAGAGATTGGCATGTTCATGATTGTCAAACTGGTTTGCCAATTATGAGA
GGTACTTCTGTTTGGGTTATGATGGATAAGCATACTAGAAGATTGTCTAAGTTGCCAGAAGAAGTTAG
AGCTGAAATTACTCCATTCTTCTCTGAAAGAGATGCTGTTTTGGATGATAATGGTAGAAAGTTGCCAA
AGTTCGATGACGATTCTGCTGCTCATGTTAGAAGAGGTTTGACTCCAAGATGGCATGATTTCGATGTT
AATCAACATGTTAATAATGTTAAGTACGTTGGTTGGATTTTGGAATCTGTTCCAGTTTGGATGTTGGAT
GGTTACGAGGTTGCTACTATGTCTTTGGAGTACAGAAGAGAGTGTAGAATGGATTCTGTTGTTCAATC
TTTGACTGCTGTTTCTTCTGATCATGCTGATGGTTCTCCAATTGTTTGTCAACATTTGTTGAGATTGGA
AGATGGTACTGAAATTGTTAGAGGTCAAACTGAATGGAGACCAAAGCAACAAGCTAGAGATTTGGGT
AATATGGGTTTGCATCCAACTGAATCTAAATAA

***Cuphea palustris* thioesterase FatB1 (*S. cerevisiae* codon-optimized, N-term signal peptide removed):**

ATGAGGCCAAACATGTTGATGGATTCCTTCGGCTTGGAAAGAGTCGTCCAAGATGGTTTGGTCTTCA
GACAATCCTTCTCCATTAGATCCTATGAAATTTGTGCTGATAGAACTGCTTCCATTGAAACTGTCATGA
ACCATGTCCAAGAAACTTCCTTGAACCAATGTAAGTCCATTGGTTTGTTGGATGATGGTTTCGGTAGA
TCCCCAGAAATGTGTAAGAGAGATTTGATTTGGGTCGTCACTAGAATGAAGATTATGGTCAACAGATA
CCCAACTTGGGGTGATACTATTGAAGTCTCCACTTGGTTGTCTCAATCTGGTAAGATTGGTATGGGTA
GAGATTGGTTGATTTCTGATTGTAACACTGGTGAAATTTTGGTCAGAGCTACTTCCGTCTACGCTATG
ATGAACCAGAAGACGAGAAGATTCTCCAAGTTGCCACATGAAGTCAGACAAGAATTTGCTCCACATTT
CTTGGATTCCCCACCAGCTATTGAAGATAACGATGGTAAGTTGCAAAAGTTCGATGTCAAGACTGGT
GATTCCATTAGAAAGGGTTTGACTCCAGGTTGGTACGATTTGGATGTCAACCAACATGTCTCTAACGT
CAAGTACATTGGTTGGATTTTGGAATCTATGCCAACTGAAGTCTTGGAAACTCAAGAATTGTGTTCTT
TGACTTTGGAATACAGAAGAGAATGTGGTAGAGATTCTGTCTTGGAATCCGTCACTTCTATGGACCCA

TCTAAGGTCGGTGATAGATTCCAATACAGACATTTGTTGAGATTGGAAGATGGTGCTGATATTATGAA
GGGTAGAACTGAATGGAGACCAAAGAACGCTGGTACTAACGGTGCTATTTCTACTGGTAAGACTTAA

***Umbellularia californica*** **thioesterase FatB2 (*E. coli* codon-optimized, N-term signalpeptide removed):**

ATGACTCTAGAGTGGAAACCGAAACCAAAACTGCCTCAACTGCTGGATGATCACTTCGGTCTGCACG
GTCTGGTGTTTCGTCGTACTTTCGCAATTCGTTCTTATGAAGTGGGTCCAGATCGTTCTACCTCCATC
CTGGCCGTCATGAACCACATGCAGGAAGCCACCCTGAATCACGCGAAATCTGTTGGTATCCTGGGT
GATGGTTTCGGCACTACTCTGGAAATGTCTAAACGTGACCTGATGTGGGTAGTGCGTCGCACCCAC
GTAGCAGTAGAGCGCTACCCTACTTGGGGTGACACTGTGGAAGTCGAGTGTTGGATTGGCGCGTCC
GGTAACAATGGTATGCGTCGCGATTTTCTGGTCCGTGACTGTAAAACGGGCGAAATCCTGACGCGTT
GCACCTCCCTGAGCGTTCTGATGAACACCCGCACTCGTCGCCTGTCTACCATCCCGGACGAAGTGC
GCGGTGAGATCGGTCCTGCTTTCATCGATAACGTGGCAGTTAAAGACGACGAAATCAAGAAACTGCA
AAAACTGAACGACTCCACCGCGGACTACATCCAGGGCGGTCTGACTCCGCGCTGGAACGACCTGGA
TGTTAATCAGCATGTGAACAACCTGAAATACGTTGCTTGGGTCTTCGAGACTGTGCCGGACAGCATT
TTCGAAAGCCATCACATTTCCTCTTTTACTCTGGAGTACCGTCGCGAATGTACTCGCGACTCCGTTCT
GCGCAGCCTGACCACCGTAAGCGGCGGTTCTAGCGAGGCAGGTCTGGTCTGCGACCATCTGCTGC
AACTGGAAGGCGGCTCCGAAGTCCTGCGTGCGCGTACGGAGTGGCGTCCAAAGCTGACGGATTCTT
TCCGCGGCATCTCCGTAATTCCGGCGGAACCTCGTGTTTAA

**References**


1. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845-858.
2. Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., and Vajda, S. (2013) How good is automated protein docking?. Proteins: Struc., Func., Bioinf. 81, 2159-2166.
3. Nguyen, C., Haushalter, R. W., Lee, D. J., Markwick, P. R., Bruegger, J., Caldara-Festin, G., Finzel, K., Jackson, D. R., Ishikawa, F., O'Dowd, B., McCammon, J. A., Opella, S. J., Tsai, S. C. and Burkart, M. D. (2014) Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature 505*, 427-431.

## Chapter 6. The Protein-protein interactions responsible for Ketosynthase specificity

### 6.1 Introduction

Protein-protein interactions persist as a poorly understood but foundational element of enzymology, with the capacity to modulate reactivity for efficiency and control of produced products[1–5]. Facilitating functional interactions between two enzymes, PPIs have been increasingly identified as key components of enzyme function and specificity across systems of biology[6–8]. Metabolic biosynthesis is a key example of processes which require a combination of efficiency and specificity to produce the essential cellular components of life[9,10]. The speed and flexibility of this key step has made it both important to reactivity and difficult to characterize, leading to the development of new methods to interrogate its mechanism.

Fatty acid biosynthesis (FAB) is a conserved, essential pathway across all domains of life that produces the small carbon chains necessary for lipid production, cell membranes, and other metabolic purposes[11–14]. The iterative addition of activated keto units to a thioester cofactor is accomplished through a sequence of enzymes. In the type I FAB found in higher organisms, these enzymes compose large, multidomain megasynthases that relie on a bound carrier to move between each step[15]. Bacterial, mitochondrial, and chloroplast fatty acid biosynthesis is primarily performed by type II FAB[16,17]. Type II FAB enzymes are freely diffusing in the cystol and rely on protein-protein interactions and substrate pockets, rather than having an attached megasynthase, for substrate recognition. Both processes involve an acyl carrier protein (ACP) that shuttles the substrate between enzymes in the proper order for elongation[12,18]. AcpP is a small (~10kbp), highly soluble protein composed of four helices[19]. These helices form a hydrophobic encasement around a phosphopantetheine cofactor attached to a central serine residue. Throughout elongation the carrier protein protects substrates within the hydrophobic pocket, transferring the substrate to partners before returning to be shuttled to the next enzyme in the pathway. This process is known as chain flipping and allows the proper enzyme in the FAB pathway access to the thioester bond within ACP to perform its reaction[20–22]. The protein-protein interaction between ACP and each enzyme are a critical element of control in this pathway.

The first elongating protein-protein interaction in the FAB pathway occurs between ACP and the $\beta$-ketoacyl-ACP synthase III FabH[23,24]. FabH catalyzes the condensation reaction between malonyl-sequestered ACP and activated acetyl-CoA. Its specificity is demonstrated for catalyzing the entrance into the FAB cycle[25]. This provides the starting point to further elongation reactions leading to the development of saturated and unsaturated fatty acids[24]. Their production is essential to the understanding of antibiotic resistant bacteria and engineering FAB products[26,27]. One of the most remarkable aspects noted in recent work is the carrier protein's ability to demonstrate a control over reactivity based on the identity of its substrate. Previous studies have demonstrated that substrate orthologs bearing a reactive chemical crosslinker loaded onto the ACP will not react with partners when the probe does not mimic their substrate[1,28]. Expanding upon this work, we provide examples of the ability of the FabH enzyme to control reactivity based on PPIs with AcpP bearing the proper substrate. Uniquely through $^{1}$H-$^{15}$N HSQC NMR titrations, we can appreciate the transient interaction events which have previously been hypothesized to be the first method of control of FAB fidelity[29,30]. It is important to understand the control mechanism of such a process to manipulate it, both in the case of producing alternative products, or prevent it from producing harmful natural products[8,31]. Entrance into the FAB cycle is the first step in an essential pathway controlled by a single enzyme, making inhibition of this highly specific enzyme an opportunity for antibiotic innovation.

The remainder of the FAS cycle is driven by the elongating ketosynthases FabF and FabB[32]. The two ketosynthases in *E. coli* function by the same mechanism and often interchangeably. However, notably FabB is an essential enzyme while FabF is not[33]. This essential role has been identified as the elongation of unsaturated fatty acids, with FabB deleted *E. coli* an auxotroph for supplemented unsaturated fatty acids[33–35]. In order to characterize if this second step could have a protein-protein interaction-based control mechanism three different chain lengths of AcpP were titrated with FabF and FabB. Combined the study of these three proteins constitutes a holistic analysis of the protein-protein interactions of every ketosynthase in *E. coli*.

**6.2 Results**

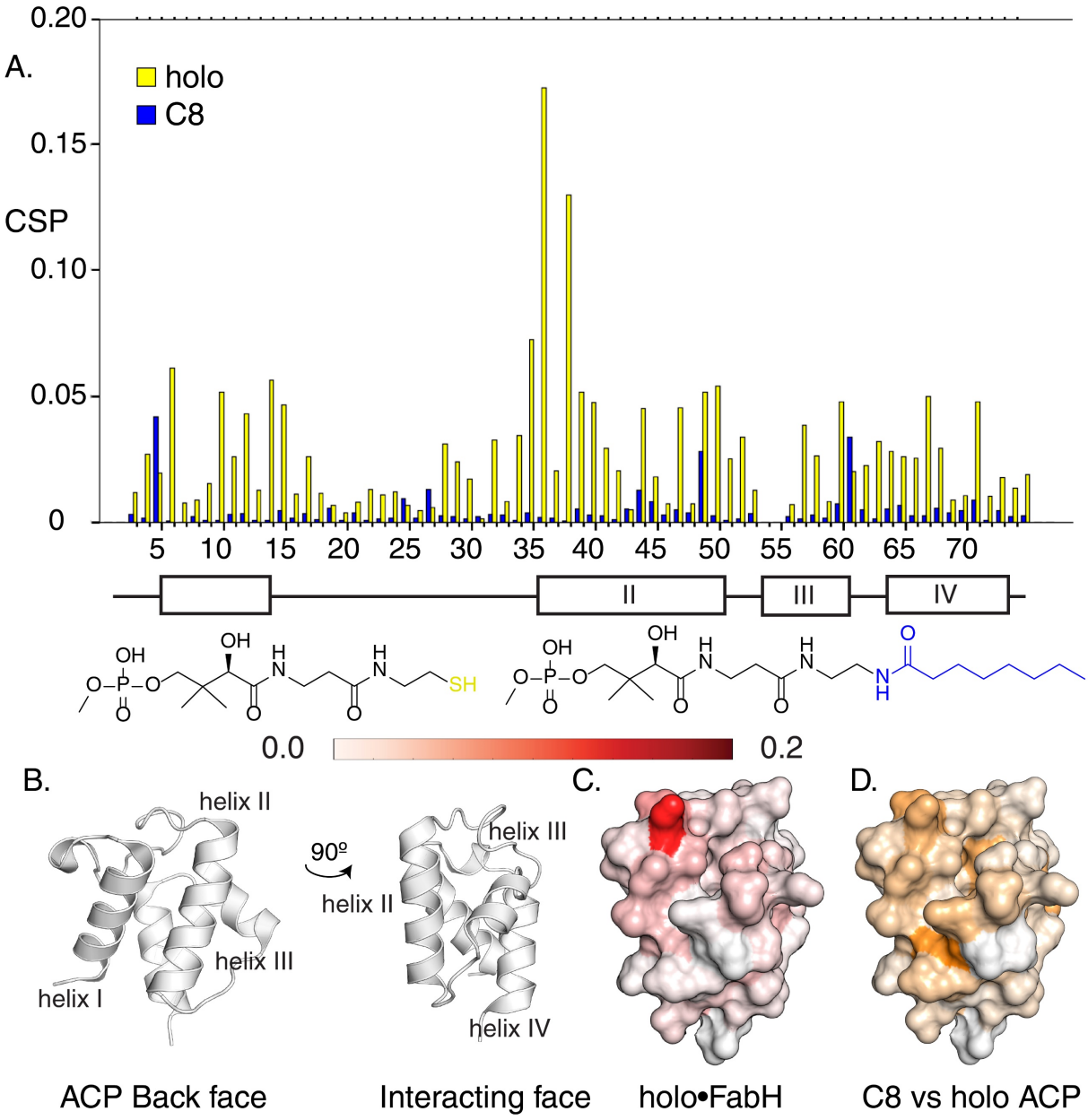



**Figure 6.1 Titration of holo- and C8-AcpP with FabH. A. Holo- (yellow) and C8-(blue) AcpP were titrated with FabH in order to observe the differential binding surfaces of the two ACP species. C8-AcpP demonstrates a precipitous drop in perturbations, demonstrating FabH's specificity for short chains. B. The secondary structure of the *E. coli* AcpP C. The surface of AcpP colored by intensity of CSP. D. Subtraction of the C8- CSPs from the holo-CSPs, demonstrating the regions which are unique to FabH.**

**Titration of FabH with holo and octanoyl AcpP**

The FabH ketosynthase is the first step in the fatty acid biosynthetic cycle. Catalyzing a ping-pong reaction wherein the FabH enzyme first reacts with acetyl-CoA and secondly with malonyl-AcpP to generate acetoacetyl-AcpP and $CO_2$[36,37]. As the initiation step in FAS it was hypothesized that FabH would maintain careful control over substrate specificity and reactivity, maintaining specificity for short chain AcpP and avoiding off target binding of medium and long chains. In order to observe this, titrations were performed to examine the interactions of holo and octanoyl ACP with FabH to observe any selectivity in protein-protein interactions. Octanoyl-AcpP has little to no interaction with FabH, with the mean CSP being 6.1 times smaller than to mean perturbation of holo AcpP. Even the most perturbed residues in the C8 titration are barely greater than the mean perturbation in holo. E5, E49, E61 are the most perturbed residues in the C8-AcpP•FabH titration, demonstrating that while there are still negatively charged residues interacting with the FabH the interactions are not near the phosphopantetheine appended serine 36. Furthermore, it was previously noted that internal perturbations are seen in functional partners, likely representing those residues responsible for chain flipping, no internal residue perturbations are present in the octanoyl titration. Demonstrating that FabH binds to a non-functional but still negative surface of C8-AcpP rather than the functional binding face.

Holo-AcpP titrated with FabH appears to have a more functional binding surface for AcpP interactions. The perturbations one standard deviation above the mean occur at the charged interface residues E35 and D38, positions beside the functional S36. Internal perturbations at R6, Q14, and F50 are hypothesized to represent the allosteric movements occurring upon partner binding to facilitate chain flipping. The largest perturbation is at serine 36, reflecting the interactions with the phosphopantethine substrate. The differences in the untitrated spectra of C8 and holo AcpP point to how reactivity can be controlled. Perturbations one standard deviation above the mean lie on residues: S27, F28, G33, D35, L37, D38, V40, M44, E57, E60, and A68. All these perturbations except A68 are on the binding face of Helix II & III or on loop 2 nearby the critical serine 36. Making it no surprise that proteins may have evolved mechanisms to recognize these two very different interfaces.

**Titration of FabF and FabB with C8-, C10-, and C10:1- AcpP**

To test the differences in FabF and FabB binding of three different AcpPs pantethenamide probes were prepared and loaded onto $^{15}N$ labeled ACP. It was seen that FabB retains the same magnitude of perturbation between C8- and C10-AcpP titrations. However, FabF saw a significant loss of perturbation, with the C8-AcpP having common magnitudes of perturbation but FabF losing significant magnitude when interaction with C10-AcpP. This data appears to suggest that there is an initial recognitions step which favors FabB binding C10, while FabF loses some interaction. This was seen to be a much smaller effect than the FabH preference, likely reflecting the significantly broader substrate profile of FabF.

Next titration of C10:1-AcpP with FabF and FabB was performed to test if the proteins showed a similar preference for substrates. The CSPs of C10:1-AcpP titrated with FabF and FabB were graphed next to one another. Like the previous case the most obvious difference is in the magnitude of CSPs. However, when they are examined in greater detail more can be elucidated. Beginning at residue 10 FabB shows much larger perturbations, with residue I10, I11, and L15 demonstrating large perturbations in FabB specifically. Further down the protein P28 and D31 show larger perturbation in FabB and proceed the important helix II. The DSL motif shows larger perturbations in the FabB titration with D35, S36, and L37. This would suggest that FabB has a more functional interaction, with the critical serine 36 which carries the fatty acid perturbed along with its neighbors. Further down helix II both proteins show similar interactions until the bottom of helix II. There residues L42 and V43 see significant interactions with one very large perturbation at E47. The perturbations are similar until helix III where FabB has a large perturbation at E57, while FabF has its first and only significantly larger perturbation at E60. After this FabB sees stronger interactions through helix IV at Q66, A69, and I69. The perturbations which are stronger for FabB over FabF predominantly are hydrophobic ones. This could imply that either chain flipping is not occurring in FabF, or that the complex is stronger and causing more allosteric effects in the FabB protein. However, either result implies that there is a mechanism to control the interaction. But, to verify this a second experiment was planned to interrogate the meaning behind CSPs.

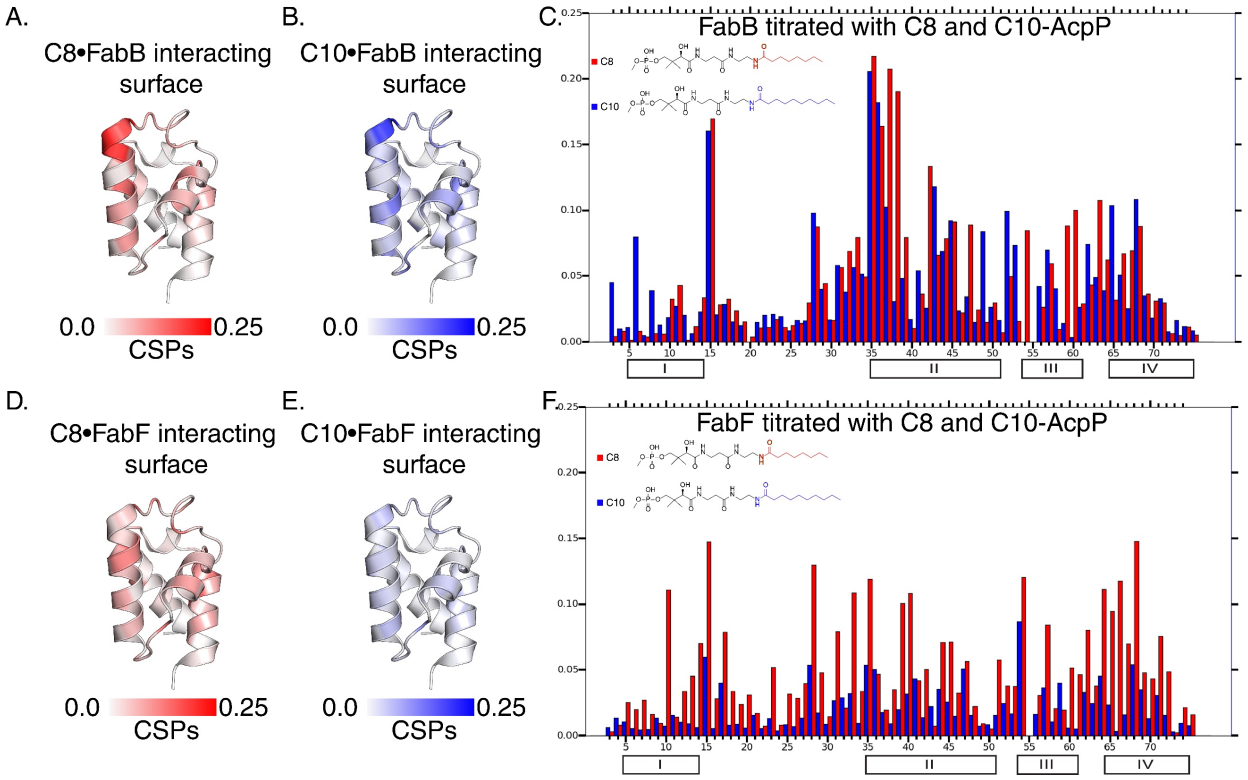



**Figure 6.2. Titrations of C8- and C10-AcpP with FabF and FabB. A. The secondary structure of AcpP colored by the magnitude of CSP from the C8-AcpP•FabB titration. B. The secondary structure of AcpP colored by the magnitude of CSP from the C10-AcpP•FabB titration. C. Titration of C8-(red) and C10-(blue) AcpP with FabB. D. The secondary structure of AcpP colored by the magnitude of CSP from the C8-AcpP•FabF titration. E. The secondary structure of AcpP colored by the magnitude of CSP from the C10-AcpP•FabF titration F. Titration of C8-(red) and C10(blue) AcpP with FabF.**

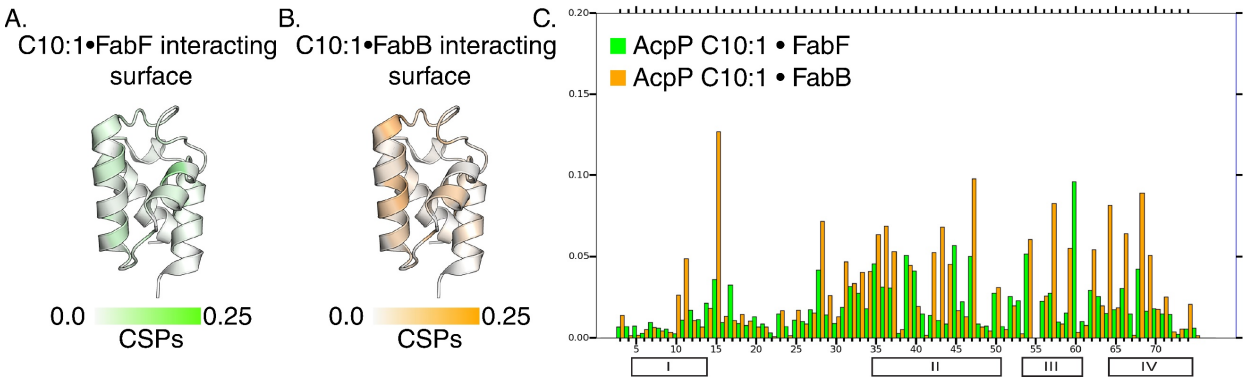



**Figure 6.3. Titrations of C10:1-AcpP with FabF and FabB. A. The secondary structure of AcpP colored by the magnitude of CSP from the C10:1-AcpP•FabF titration. B. The secondary structure of AcpP colored by the magnitude of CSP from the C10:1-AcpP•FabB titration. C. Titration of C10:1-AcpP with FabB(orange) or FabF(green).**

To characterize whether chain flipping is essential for the CSPs we see in titrations an experiment was performed with FabF preincubated with cerulenin. Cerulenin is a covalent inhibitor which fills the FabF pocket with an acyl tail. This creates a perfect test system, where the surface interactions of the AcpP•FabF are possible but no chain flipping can be accomplished. It has been previously identified that there are internal and surface perturbations, with the internal perturbations linked to chain flipping and the surface interactions AcpP•Partner interactions[30]. This experiment further reinforced this theory, with the cerulenin binding only affecting the internal hydrophobic interactions. CSPs showed a drop in magnitude in hydrophobic interactions internal to the interaction, while in some cases the electrostatic interactions on the surface increased in magnitude. Demonstrating that while the interface was retained only the residues which are perturbed when the acyl chain is outside the AcpP are affected by cerulenin. This also demonstrates that the CSPs in previous studies which are lost can be linked to a loss of interface interactions, with the pocket having little effect on the electrostatic interactions that are so important to the AcpP•Partner interface.

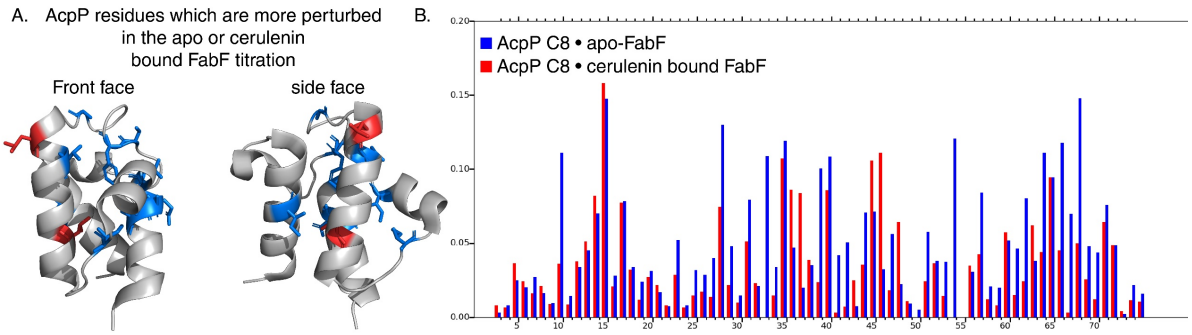



**Figure 6.4. Comparing a C8-AcpP•apo-FabF titration to a C8-AcpP•cerulenin-FabF titration. A. The secondary structure of AcpP with the residues which are uniquely perturbed in the apo-FabF titration (Blue) and cerulenin bound FabF (red) shown as colored sticks. B. Titration of C8-AcpP with apo-FabF (blue) or cerulenin bound FabF (red).**

### 6.3 Discussion

A better understanding of the mechanism of fatty acid biosynthesis is necessary due to its biochemical applications. Fatty acids can be utilized to replace diesel as a renewable fuel source. Studying fatty acid biosynthesis in *E. coli* bacterial cells could lead to production of efficient biofactories that produce fatty acids through genetic engineering. (10) The production of fatty acids in invasive bacterial species presents an interesting target for drug design. Many species of mycobacteria, especially those that develop antibiotic resistance, utilize fatty acids to produce protective biofilms. These greasy outer coatings become a barricade against host immune cells and antibiotics, allowing invasive bacteria to proliferate. (11) Soluble antibiotics cannot penetrate biofilms, but the importance of PPIs in FAB propose an alternative solution. The production of small molecules mimicking the proper ACP-substrate could be used to covalently bind to FAB enzyme active sites, halting the mechanism and the production of fatty acids.

**Chapter 6** is in combination of two manuscripts in preparation Thomas G. Bartholow, Terra Sztain, Megan A. Young, Tony D. Davis Michael D. Burkart. “Selectivity in the initiating ketosynthase, FabH, of E. coli fatty acid biosynthesis.” And: Thomas G. Bartholow, Jeffery T. Mindrebo, Ashay Patel, Woojoo E. Kim, Aochiu Chen, Megan A. Young, Michael D. Burkart “A Protein-Protein Interaction based control step for the

specificity of Unsaturated Fatty Acid elongation." The dissertation author is the primary author of the manuscript.

6.4 References


(1) 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. https://doi.org/10.1016/j.chembiol.2015.09.009.

(2) Ran, X.; Gestwicki, J. E. Inhibitors of Protein-Protein Interactions (PPIs): An Analysis of Scaffold Choices and Buried Surface Area. *Curr. Opin. Chem. Biol.* **2018**, *44*, 75–86. https://doi.org/10.1016/j.cbpa.2018.06.004.

(3) Jones, S.; Thornton, J. M. Principles of Protein-Protein Interactions. *Proc. Natl. Acad. Sci.* **1996**, *93* (1), 13–20. https://doi.org/10.1073/pnas.93.1.13.

(4) Diversity of protein–protein interactions | The EMBO Journal https://www.embopress.org/doi/full/10.1093/emboj/cdg359 (accessed May 31, 2021).

(5) Phizicky, E. M.; Fields, S. Protein-Protein Interactions: Methods for Detection and Analysis. *Microbiol. Rev.* **1995**, *59* (1), 94–123.

(6) Wendland, B.; Emr, S. D. Pan1p, Yeast Eps15, Functions as a Multivalent Adaptor That Coordinates Protein–Protein Interactions Essential for Endocytosis. *J. Cell Biol.* **1998**, *141* (1), 71–84. https://doi.org/10.1083/jcb.141.1.71.

(7) DeBello, W. M.; O'Connor, V.; Dresbach, T.; Whiteheart, S. W.; Wang, S. S.-H.; Schweizer, F. E.; Bertz, H.; Rothman, J. E.; Augustine, G. J. SNAP-Mediated Protein–Protein Interactions Essential for Neurotransmitter Release. *Nature* **1995**, *373* (6515), 626–630. https://doi.org/10.1038/373626a0.

(8) Protein–protein interactions within the Fatty Acid Synthase-II system of Mycobacterium tuberculosis are essential for mycobacterial viability - Veyron-Churlet - 2004 - Molecular Microbiology - Wiley Online Library https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2958.2004.04334.x (accessed May 31, 2021).

(9) Yamada, T.; Bork, P. Evolution of Biomolecular Networks — Lessons from Metabolic and Protein Interactions. *Nat. Rev. Mol. Cell Biol.* **2009**, *10* (11), 791–803. https://doi.org/10.1038/nrm2787.

(10) Portillo, F.; Vázquez, J.; Pajares, M. A. Protein-Protein Interactions Involving Enzymes of the Mammalian Methionine and Homocysteine Metabolism. *Biochimie* **2020**, *173*, 33–47. https://doi.org/10.1016/j.biochi.2020.02.015.

(11) Rock, C. O.; Jackowski, S. Incorporation and Turnover of Fatty Acids in Escherichia Coli Membrane Phospholipids. In *Advances in Lipobiology*; Gross, R. W., Ed.; JAI, 1996; Vol. 1, pp 39–59. https://doi.org/10.1016/S1874-5245(96)80004-8.

(12) Chan, D. I.; Vogel, H. J. Current Understanding of Fatty Acid Biosynthesis and the Acyl Carrier Protein. *Biochem. J.* **2010**, *430* (1), 1–19. https://doi.org/10.1042/BJ20100462.

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

(14) Parsek, M. R.; Val, D. L.; Hanzelka, B. L.; Cronan, J. E.; Greenberg, E. P. Acyl Homoserine-Lactone Quorum-Sensing Signal Generation. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (8), 4360–4365.

(15) Schweizer, E.; Hofmann, J. Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems. *Microbiol. Mol. Biol. Rev.* **2004**, *68* (3), 501–517. https://doi.org/10.1128/MMBR.68.3.501-517.2004.

(16) Schonauer, M. S.; Kastaniotis, A. J.; Kursu, V. A. S.; Hiltunen, J. K.; Dieckmann, C. L. Lipoic Acid Synthesis and Attachment in Yeast Mitochondria. *J. Biol. Chem.* **2009**, *284* (35), 23234–23242. https://doi.org/10.1074/jbc.M109.015594.

(17) 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. https://doi.org/10.1039/C8NP00040A.

(18) Byers, D. M.; Gong, H. Acyl Carrier Protein: Structure–Function Relationships in a Conserved Multifunctional Protein Family. *Biochem. Cell Biol.* **2007**, *85* (6), 649–662. https://doi.org/10.1139/O07-109.

(19) Kim, Y.; Prestegard, J. H. A Dynamic Model for the Structure of Acyl Carrier Protein in Solution. *Biochemistry* **1989**, *28* (22), 8792–8797. https://doi.org/10.1021/bi00448a017.

(20) 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. https://doi.org/10.1002/anie.201408576.

(21) Colizzi, F.; Masetti, M.; Recanatini, M.; Cavalli, A. Atomic-Level Characterization of the Chain-Flipping Mechanism in Fatty-Acids Biosynthesis. *J. Phys. Chem. Lett.* **2016**, *7* (15), 2899–2904. https://doi.org/10.1021/acs.jpclett.6b01230.

(22) Cronan, J. E. The Chain-Flipping Mechanism of ACP (Acyl Carrier Protein)-Dependent Enzymes Appears Universal. *Biochem. J.* **2014**, *460* (2), 157–163. https://doi.org/10.1042/BJ20140239.

(23) Davies, C.; Heath, R. J.; White, S. W.; Rock, C. O. The 1.8 Å Crystal Structure and Active-Site Architecture of β-Ketoacyl-Acyl Carrier Protein Synthase III (FabH) from Escherichia Coli. *Structure* **2000**, *8* (2), 185–195. https://doi.org/10.1016/S0969-2126(00)00094-0.

(24) Lai, C.-Y.; Cronan, J. E. Beta-Ketoacyl-Acyl Carrier Protein Synthase III (FabH) Is Essential for Bacterial Fatty Acid Synthesis. *J. Biol. Chem.* **2003**, *278* (51), 51494–51503. https://doi.org/10.1074/jbc.M308638200.

(25) Jeucken, A.; Molenaar, M. R.; van de Lest, C. H. A.; Jansen, J. W. A.; Helms, J. B.; Brouwers, J. F. A Comprehensive Functional Characterization of Escherichia Coli Lipid Genes. *Cell Rep.* **2019**, *27* (5), 1597-1606.e2. https://doi.org/10.1016/j.celrep.2019.04.018.

(26) Pérez-Castillo, Y.; Froeyen, M.; Nowé, A.; Cabrera-Pérez, M. Á. Chapter 5 - Bacterial FabH: Towards the Discovery of New Broad-Spectrum Antibiotics. In *Recent Advances in Medicinal Chemistry*; Atta-ur-Rahman, Choudhary, M. I., Perry, G., Eds.; Elsevier, 2014; pp 131–158. https://doi.org/10.1016/B978-0-12-803961-8.50005-1.

(27) Choi, K.-H.; Heath, R. J.; Rock, C. O. β-Ketoacyl-Acyl Carrier Protein Synthase III (FabH) Is a Determining Factor in Branched-Chain Fatty Acid Biosynthesis. *J. Bacteriol.* **2000**, *182* (2), 365–370.

(28) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44–46. https://doi.org/10.1039/b512735a.

(29) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001.

(30) Bartholow, T. G.; Sztain, T.; Lee, D. J.; Abagyan, R.; Burkart, M. D. Elucidation of Transient Protein-Protein Interactions within Carrier Protein-Dependent Biosynthesis. *Commun. Biol. Accept.*

(31) Nanson, J. D.; Himiari, Z.; Swarbrick, C. M. D.; Forwood, J. K. Structural Characterisation of the Beta-Ketoacyl-Acyl Carrier Protein Synthases, FabF and FabH, of Yersinia Pestis. *Sci. Rep.* **2015**, *5* (1), 14797. https://doi.org/10.1038/srep14797.

(32) Yu, X.; Liu, T.; Zhu, F.; Khosla, C. In Vitro Reconstitution and Steady-State Analysis of the Fatty Acid Synthase from Escherichia Coli. *Proc. Natl. Acad. Sci.* **2011**, *108* (46), 18643–18648. https://doi.org/10.1073/pnas.1110852108.

(33) Feng, Y.; Cronan, J. E. Escherichia Coli Unsaturated Fatty Acid Synthesis. *J. Biol. Chem.* **2009**, *284* (43), 29526–29535. https://doi.org/10.1074/jbc.M109.023440.

(34) 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* (22), 3060–3069. https://doi.org/10.1128/JB.00463-16.

(35) Xiao, X.; Yu, X.; Khosla, C. Metabolic Flux Between Unsaturated and Saturated Fatty Acids Is Controlled by the FabA:FabB Ratio in the Fully Reconstituted Fatty Acid Biosynthetic Pathway of E. Coli. *Biochemistry* **2013**, *52* (46). https://doi.org/10.1021/bi401116n.

(36) Gajiwala, K. S.; Margosiak, S.; Lu, J.; Cortez, J.; Su, Y.; Nie, Z.; Appelt, K. Crystal Structures of Bacterial FabH Suggest a Molecular Basis for the Substrate Specificity of the Enzyme. *FEBS Lett.* **2009**, *583* (17), 2939–2946. https://doi.org/10.1016/j.febslet.2009.08.001.

(37) Qiu, X.; Choudhry, A. E.; Janson, C. A.; Grooms, M.; Daines, R. A.; Lonsdale, J. T.; Khandekar, S. S. Crystal Structure and Substrate Specificity of the β-Ketoacyl-Acyl Carrier Protein Synthase III (FabH) from Staphylococcus Aureus. *Protein Sci. Publ. Protein Soc.* **2005**, *14* (8), 2087–2094. https://doi.org/10.1110/ps.051501605.

## Chapter 7. Possible evolutionary link and remaining acyl transferase activity of E. coli malonyl-CoA: acyl-carrier protein transacylase FabD

### 7.1 Introduction

Carrier protein mediated biosynthesis is implicated in both primary and secondary metabolism[1–4]. These pathways utilize conserved and iterative chemistries to build ketide or amino acid precursors into complex molecules vital to biology, pharmaceutics, and industry including fatty acids, polyketides, and nonribosomal peptides. Notably, fatty acids are primary metabolites that are critical for all organisms to maintain cellular homeostasis and structural integrity of the cellular membrane. Fatty acid biosynthesis is executed by fatty acids synthases (FASs), which consist of polypeptide megasynthases that contain distinct catalytic domains (type I) or discrete enzymes (type II) that polymerize and reduce two carbon ketide units. Importantly, engineering both type I and type II FASs has emerged as a strategy to produce shorter chain fatty acids as potential sources of biofuels and biologically important natural products[5–8]. Fatty acid biosynthesis is initiated by malonyl-CoA acyl carrier protein (ACP) transacylases (MAT), which charges the phosphopantetheine (PPant) arm of *holo*-ACP with malonate through transient malonylation of an active site serine residue (**Figure 1**)[9,10]. Notably, the MAT from *Escherichia coli*, FabD is a highly efficient enzyme that turns over substrate near the diffusion limit[10–12]. The high catalytic efficiency of *E. coli* FabD makes this enzyme a particularly attractive engineering target[13], as demonstrated by studies to expand acyl-CoA substrate tolerance[14].

Herein, we discover expanded activity for *E. coli* FabD beyond a MAT and demonstrate that this enzyme can recognize longer acyl-chain substrates. NMR and docking analysis were performed to follow up this discovery, finding that FabD forms two distinct yet active interfaces for both chain lengths. High-resolution docking revealed that there are alternate stable conformations which are unique to each chain length, yet both follow the precedent of conformations seen in homologous ACP-MAT complexes.

### 7.2 RESULTS

#### Expanded (Unusual) Substrate Tolerance for *E. coli* FabD Revealed Through Crosslinking

During our investigations on the structural elucidation of *E. coli* FabD with its cognate acyl-carrier protein (AcpP)[15], we synthesized pantetheine-based crosslinkers **1–12** as candidates that could trap transient interactions between these two enzymes (**Figure 2a**). We loaded *apo*-AcpP with **1–12** using our one-pot chemoenzymatic method[16] to produce *crypto*-AcpPs, which were incubated with wt FabD or FabD S92C. Surprisingly, no crosslinking was observed between wt FabD and AcpP, even with pantetheineamides **10–11**, which were inspired by our previous activity-based probes for ACP transacylase domains (**Figure 2b**)[17]. In contrast, FabD S92C displayed differential crosslinking abilities with each *crypto*-AcpP, with crosslinking being most efficient with short-chained **1**-*crypto*-AcpP and **8**-*crypto*-AcpP, which most closely resemble the small malonate core of malonyl-CoA. Intriguingly, we observed crosslinking between FabD S92C and medium-chained **6**-*crypto*-AcpP, whose efficiency was comparable to FabD S92C–**8**-*crypto*-AcpP (where the "–" denotes a covalent crosslink between the two proteins). In addition, FabD S92C–**9**-*crypto*-AcpP also formed, albeit to a lesser extent. These observations are in line with a previous study by Marcella and Barb, who demonstrated that an R117A variant of *E. coli* FabD displayed relaxed substrate specificity, recognizing short (≤4 carbons) linear saturated, branched, and unsaturated acyl-CoAs[14]. However, in our studies AcpPs loaded with α-bromo-pantetheineamides between 4–6 carbons did not undergo crosslinking. Taken together, our data suggests a function for *E. coli* FabD beyond a *bona fide* MAT, and we hypothesize that this enzyme could moonlight as general trans-AT or as a malonyl palmitoyltransferase (MPT) for FAS engineering.

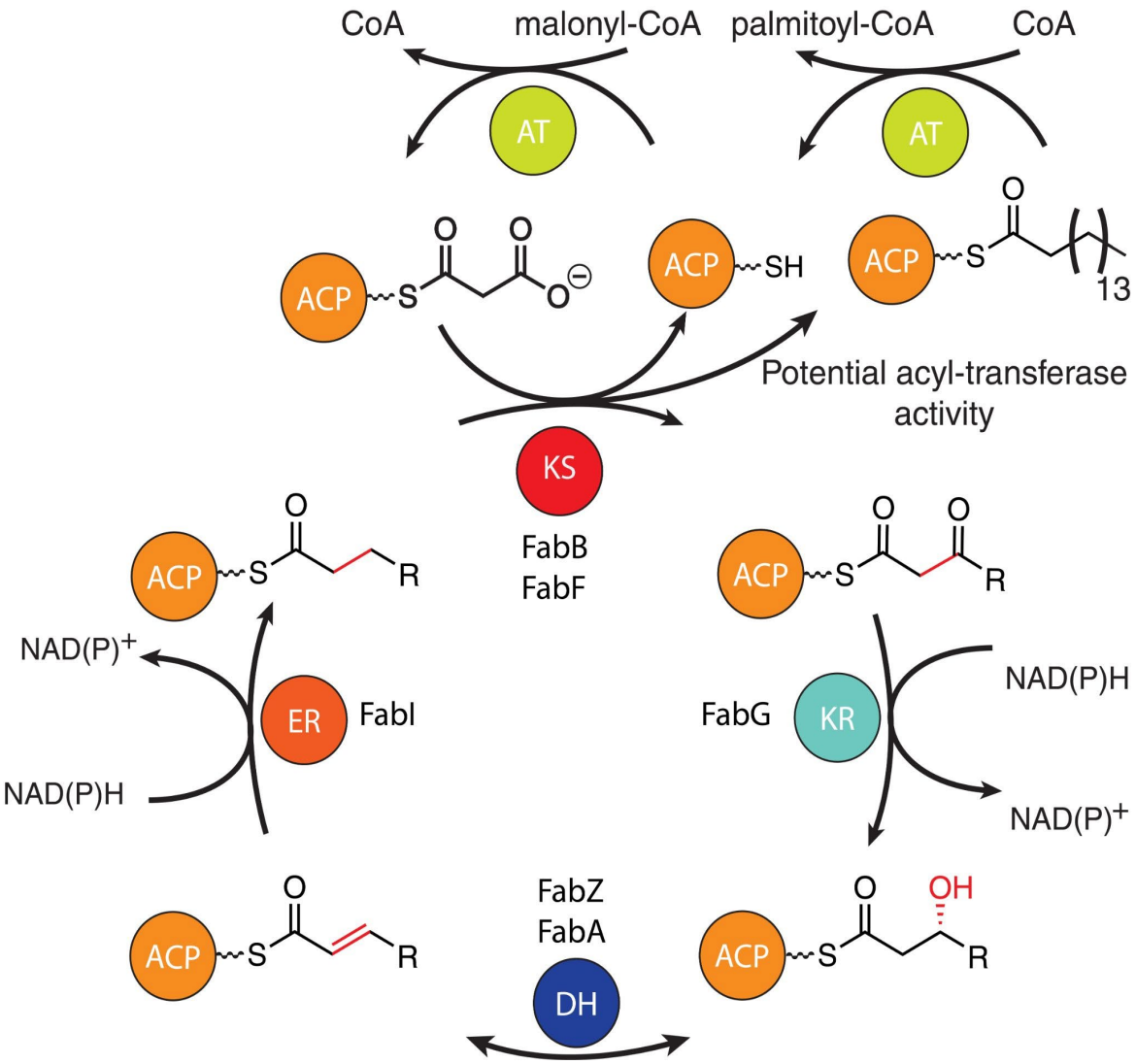



**Figure 7.1: Fatty acid biosynthesis in *Escherichia coli*.**

### Evolution of Malonyl-Palmitoyl Transferase from *E. coli* FabD

Pioneering investigations by Lynen, Schweizer, and colleagues on yeast type I FAS confirmed the presence of two transacylase domains, an AT that initiates biosynthesis using acetyl-CoA and an MPT that elongates the nascent fatty acid using malonyl-CoA and terminates biosynthesis to generate palmitoyl-CoA[18–22]. Intriguingly, the multifunctional MPT domain was believed to have evolved by covalent linkage of two smaller protein fragments[23].

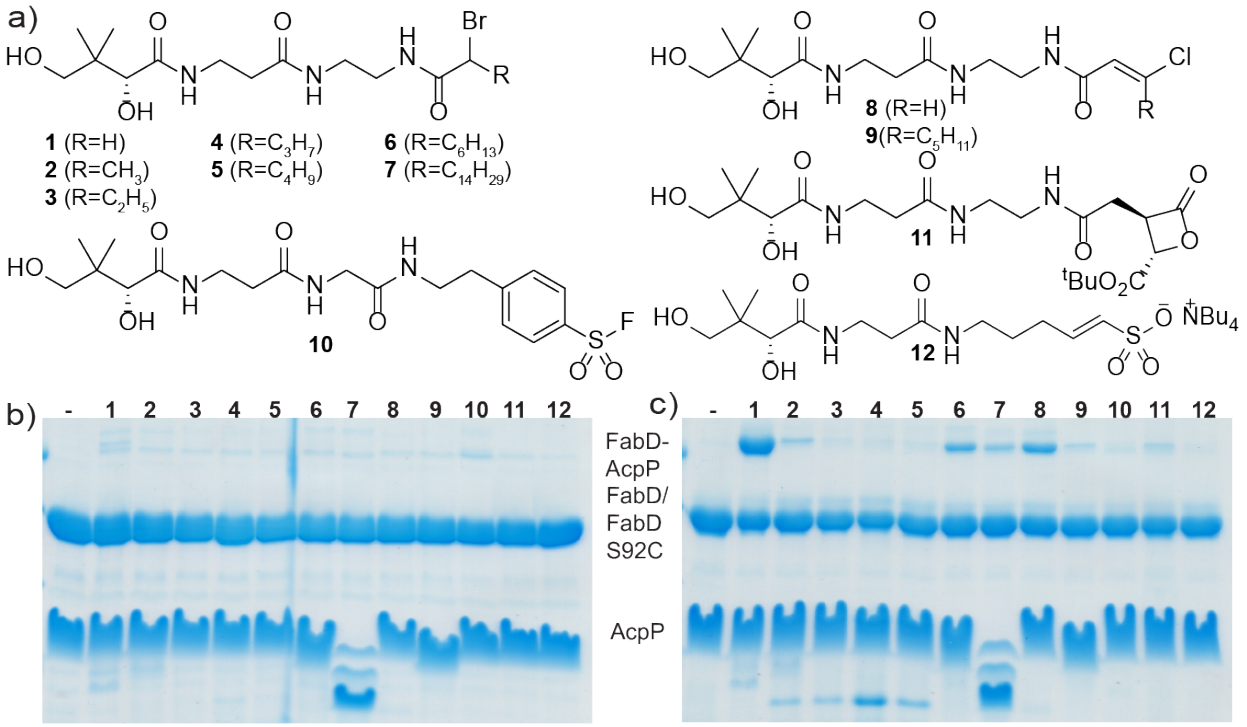



**Figure 7.2: Trapping E. coli malonyl-coA ACP transacylase acyl-carrier protein complex with covalent pantetheine probes. (a) Structures of pantetheineamide crosslinkers 1–11. SDS-PAGE analysis depicting crosslinking between 100 µM crypto-AcpP and (b) 20 µM FabD S92C or (c) 20 µM wildtype FabD in phosphate buffer pH 8 at 37 °C for 24 h.**

To probe the evolutionary link between MATs and MPTs, we conducted sequence and structural alignments of these proteins from bacteria and fungi. We aligned MATs from *E. coli* and *Mycobacterium tuberculosis* and MPTs from *M. tuberculosis, Thermomyces langinosus*, and *S. cerevisiae*. There was high sequence homology between *E. coli* FabD and MPTs from *S. cerevisiae* (44% identity, 53% similarity) and *T. lanuginosus* (40% identity, 53% similarity). Several common MAT motifs were identified across the aligned sequences, including the Ser-His catalytic dyad, the Gly-Xxx-Ser-Xxx-Gly sequence, Arg involved in malonyl-CoA substrate recognition, and Gln residues that form the oxyanion hole during catalysis (**Figure 3A**)

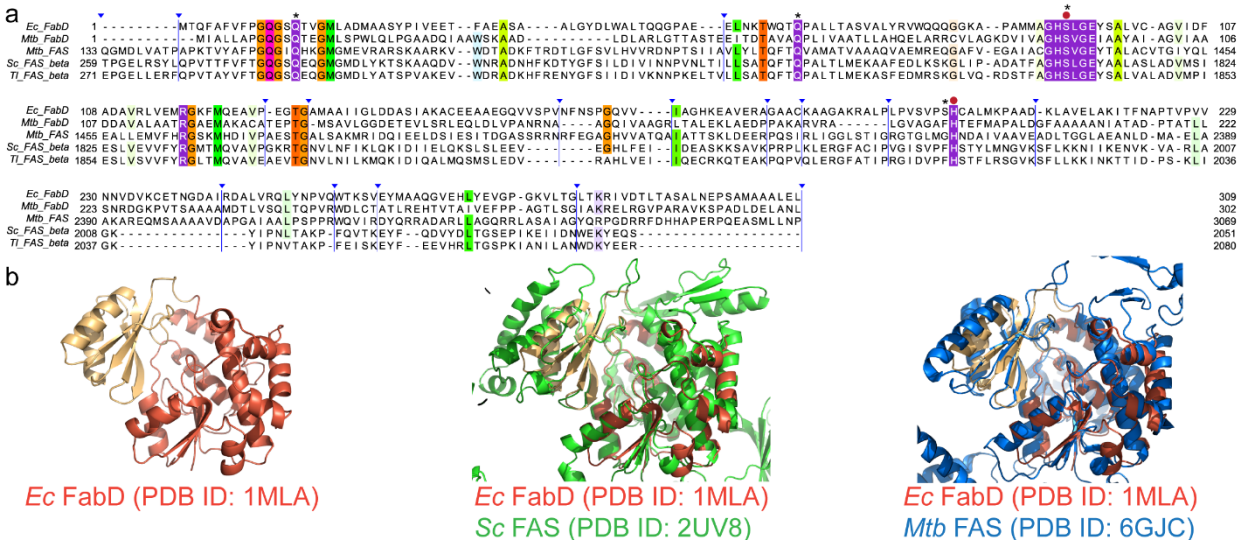



**Figure 7.3: (a) Sequence and (b) structural alignment of *Escherichia coli* FabD, *Mycobacterium tuberculosis* FabD, and malonyl-palmitoyl transferases (MPTs) from *Thermomyces langinosus* and *Saccharoymces cerevisiae* type I FASs, malonyl-acyl transferase (MAT) from *M. tuberculosis* type I FAS. Catalytic dyad (Ser 92 and His201) denoted with a red circle. Other conserved sequence motifs are boxed in purple, including: GxSxG motif, Gln11 and Gln63 (forms the oxyanion hole during malonyl transfer), Arg117 (recognizes carboxylate of malonyl-CoA substrate). Residues denoted with an asterisk are subject to mutagenesis for MPT activity assays).**

We examined the x-ray crystal structures of *E. coli* FabD (PDB ID: 1MLA) and the *S. cerevisiae* MPT domain (PDB ID: 2UV8) (**Figure 3B**). The overall architecture of each enzyme is similar, as each enzyme consists of a larger α/β hydrolase (ABH) subdomain and a smaller ferrodoxin-like (FL) subdomain. The active site of each enzyme is formed at the interface between the ABD and FL subdomains. One key difference between the two proteins is the entrance to the active site. In *E. coli FabD*, bulky residues (Glu54, Pro195) provide a narrow tunnel for malonyl-CoA. In contrast, in our structural and sequence alignments, these residues are substituted with glycine in *S. cerevisiae* and *T. lanuginosus*, enabling bulker substrates to enter the active site cavity.

### Evaluation of *E. coli* FabD MPT activity by conformationally-sensitive urea PAGE

To evaluate the MPT activity of *E. coli* FabD, we monitored the conversion of various acyl-AcpP to *holo*-AcpP by conformationally-sensitive urea polyacrylamide gel electrophoresis (PAGE). Acyl-AcpPs were chemoenzymatically synthesized from *apo*-AcpP using *Bacillus subtilis* phosophopantetheinyl transferase and coenzyme A to generate *holo*-AcpP, followed by treatment with *Vibrio harveyii* AasS (**Scheme 1**). Within 2 h of incubating palmitoleioyl-AcpP with FabD, the formation of holo-AcpP was prevalent, and after 24 hours of incubation, the ratio of holo-AcpP/palmitoleioyl-AcpP is ~1:1. No conversion

to holo-AcpP was observed after 24 hours of FabD incubation with C12-, C14-, or C18-AcpP, indicating chain transfer is specific for 16 carbons. Notably, the FabD S92C variant was not as efficient at C16 transfer. We could not monitor transfer of short to medium chain acyl groups from AcpP to CoA due to the instability of these acyl-AcpPs upon generation[24].

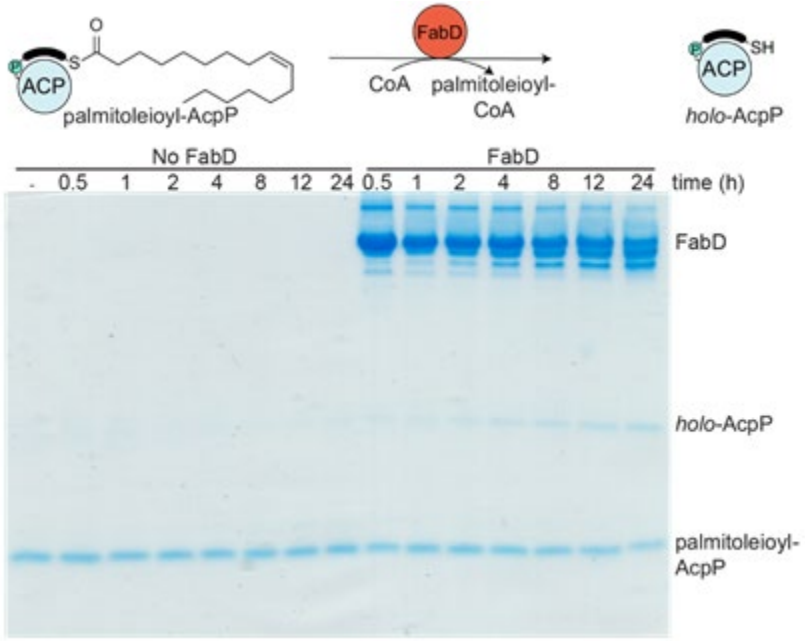



**Figure 7.4: In vitro deacylation of E. coli acyl-carrier protein by malonyl-CoA ACP transacylase. Conformationally sensitive Urea-PAGE analysis depicting deacylation of 20 μM palmitoleidoyl-AcpP with 1 mM coenzyme A and 20 μM FabD in phosphate buffer pH 8 at 37 °C for 0.5–24 h.**

**Evaluation of *E. coli* FabD deacylation activity by fluorescent assay**

Next, we turned our attention to a more quantitative analysis of the deacylation activity of FabD. We adapted the fluorogenic assay of the Bogyo group to monitor FabD deacylation kinetics. Fluorogenic substrate libraries have been widely used to probe substrate specificity of depalmitoylases[25], proteases[26], and hydrolases[27]. Thus, we synthesized a panel of 7-acyloxy-4-methylcoumarins of different chain lengths, including C2, C4, C8, and C16 acyl chains. We monitored the appearance of 7-hydroxy-4-methylcoumarin ($\lambda_{Ex}$=372 nm, $\lambda_{Em}$=445 nm) upon the addition of FabD, FabD S92C, and the malonyl palmitoyl transferase from *Mycobacterium tuberculosis*. Interestingly, neither FabD, FabD S92C, nor *Mtb* MPT cleaved C2 or C4 acylated coumarins above background. However, all three enzymes cleaved 7-octanoyl-4-methylcoumarin at comparable initial rates, and no cleavage of 7-hexadecanoyl-4-methylcoumarin was observed.

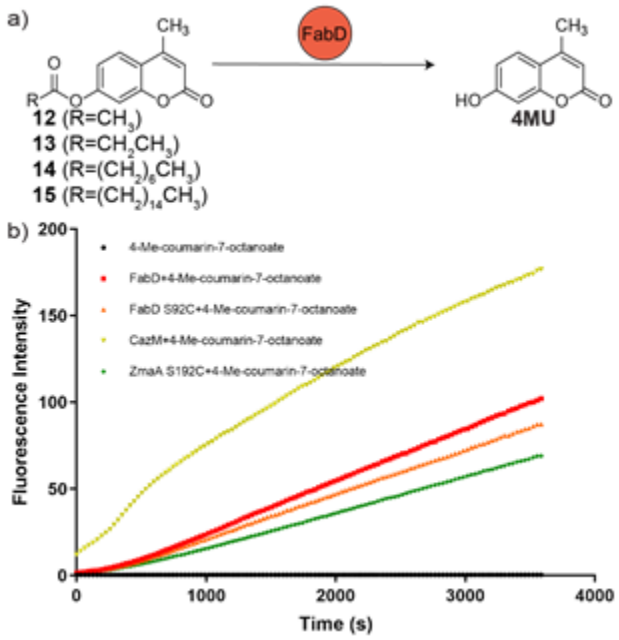



**Figure 7.5: Continuous fluorescent assay to monitor E. coli malonyl-CoA ACP transacylase deacylation activity. (a) Overall design of fluorogenic substrates 12–15 to detect MCAT deacylation activity. The substrates contain a 7-acyloxy group that is cleaved enzymatically to afford fluorescent 4-methylumbelliferone (4MU). (b) Measurement of 500 µM 4-Me-coumarin-7-octanoate hydrolysis in Tris buffer pH 7.4 by 1 µM FabD or FabD S92C. Incubation without enzyme is included as a negative control. Incubation with Mycobacterium tuberculosis malonyl-palmitoyltransferase, Chaetomium globosum CazM starter ACP transacylase, and Bacillus cereus ZmaA acetyltranferase are included as positive controls.**

### NMR studies of the FabD acyl carrier protein transacylase

To interrogate to protein-protein interaction mechanism of the dual functionality detected NMR titrations were performed. Previous studies have established the utility of NMR titrations to assess protein-protein interactions and identify the specific residues important to an interface[28–30]. To test the ability of FabD to interact with other chain length bearing ACPs we performed 1H-15N HSQC NMR titrations[31]. Both holo and octanoyl-AcpP were chosen to interrogate the ability of FabD to bind acyl-AcpPs. The residue by residue binding surface was interrogated through the titrations, allowing a view of any small differences between the binding of holo and medium chain length AcpPs. Both titrations were performed at 37°C with 1.5 molar equivalents of FabD to ensure full saturation of the AcpP.

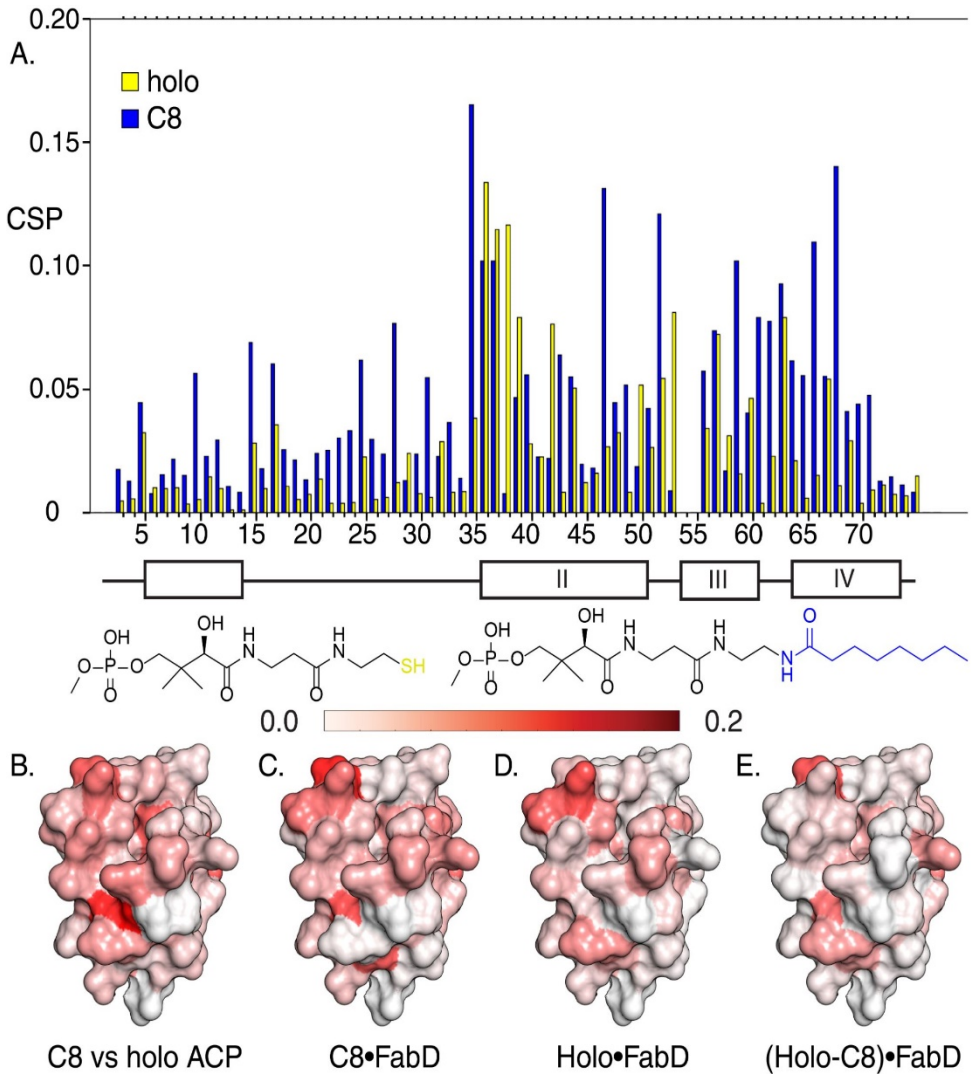



**Figure 7.6. Titration NMR to probe the dual functions of the interface of FabD A) Chemical shift perturbations of the holo (yellow) and C8 (blue) AcpP with FabD. B) Chemical shift differences between the C8 and holo AcpP before titration. This demonstrates where the differences in structure lie between the two chain lengths tested. C) The chemical shift perturbations from the titration of C8-AcpP with FabD. D) The chemical shift perturbations from the titration of holo-AcpP with FabD. E) The chemical shift perturbation differences between holo- and C8-AcpP titrated with FabD. The CSPs are directly subtracted from one another showing only the regions which are uniquely perturbed for either of the titrations. Predominantly C8 shows perturbations not seen in the holo-AcpP titration.**

Titration with holo-AcpP showed strong perturbations at the traditional interaction regions of AcpP. The top of helix II showed high CSP residus 36, 37, 38, 39, 42, and 44. The loop between helix II and III as well as helix III itself had strongs CSPs at residues 50, 52, and 53 on the loop and residues 57 and 60 on helix III. Finally, helix IV presented interactions at residues 63 and 68. Titration with C8-AcpP was expected

to show less interaction, since the established partner is holo-AcpP. However, the data agreed with the biochemical studies, showing that FabD not only interacts but forms unique contacts with FabD. These data are similar in difference to the CSPs seen when AcpP interacts with different enzymes, suggesting a unique binding mode[30]. The greatest CSPs were seen at the top and bottom of helix II at residues 35, 36, 37, and 47. The loop between helix II and III and helix III itself saw large CSPs at residues 52, 57, and 59. Finally large CSPs were seen on helix IV at residues: 63, 66, and 68.

To contrast the two titrations an analysis was performed to subtract the CSPs for the holo and C8 titrations from one another. The largest CSPs which were unique to holo-AcpP were located at residue 38 and 53, however unique perturbations are found across the top helix II and one is seen between helixes II and III. C8-FabD shows more interactions than does the holo-AcpP, with large CSPs throughout the entire AcpP. The largest of these are located at residues: 10 on helix I, residues 28, 31, and 35 lie on the loop very near the beginning of helix II. There are also perturbations at the bottom of helix II and along loop 2 at residues 43, 47,49 , and 52. The final region of CSPs it totally unique to C8-FabD and lie on the loop before and on helix IV at residues: 58, 60, 61, 66, 68, 70, and 71. This greater spread of the perturbed residues may be due to the lack of sequestration of the holo substrate. Furthermore, these large differences in CSP between the substrates is indicative of an alternative binding mode. Wherein the FabD partner binds the AcpP with alternative residues, thus explaining the very large difference in observed CSPs.

**Computational studies of the FabD acyl carrier protein transacylase**

To begin to interrogate the structural details of the AcpP•FabD interaction, high resolution docking experiments were performed using the ICM-Pro FFT docking protocol[32,33]. Briefly, holo and C8-AcpP structures were both docked against FabD[34], the ensemble of states was ranked according to the ICM energy function and the most favorable states were examined. It was first noted that the AcpP bound in two distinct orientations, both of which were approximately equally stable in their respective conformation. The two states were similar to published orientations from other transacylase structures[35]. However, no previous studies have noted alternative binding modes between substrates.

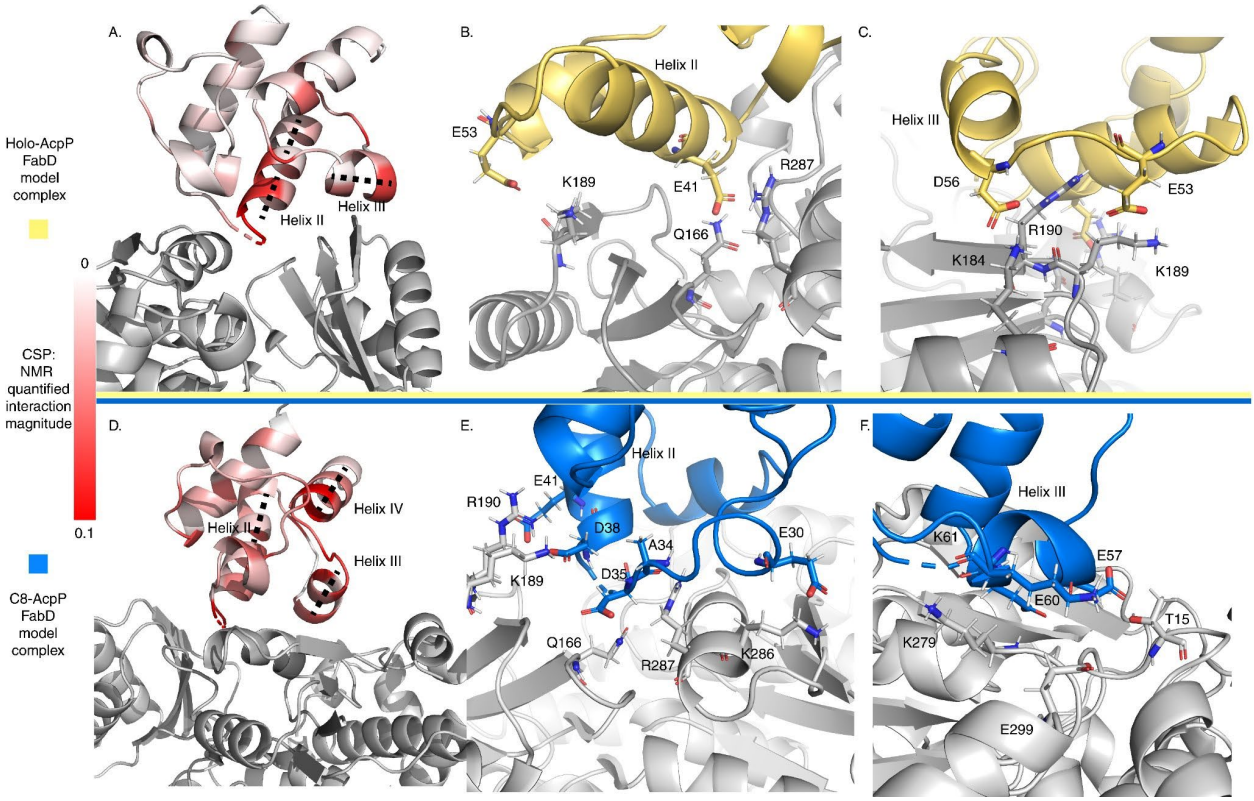



**Figure 7.7: Most stable docked orientation of the C8-AcpP●FabD and holo-AcpP●FabD. A) The CSPs of holo-AcpP titrated with FabD overlaid onto the docked structure. Demonstrating that the strongest perturbations seen in the NMR study align to the interactions in the model. B) The salt bridge interactions across helix II of AcpP (yellow) with FabD (Gray).C) The salt bridge interactions across helix III and loop II of AcpP (yellow) with FabD (Gray). D) The CSPs of C8-AcpP titrated with FabD overlaid onto the docked structure. Demonstrating that the strongest perturbations seen in the NMR study align to the interactions in the model However, C8 also sees more internal perturbations, likely due to chain flipping of the C8 substrate. E) The salt bridge interactions across helix II of AcpP (blue) with FabD (Gray). F) The salt bridge interactions across helix III and loop II of AcpP (blue) with FabD (Gray).**

The interactions of holo-AcpP were first examined by overlaying the CSPs onto the docked holo-AcpP structure. The first observation was the limited perturbed surface of the AcpP, with the majority of the perturbed region only helix II, helix III, and loop III. The most perturbed residues were examined for their specific electrostatic interactions. Beginning on helix II the region surrounding serine 36 is highly perturbed, likely due to movement of the comparatively more dynamic unacylated phosphopantetheine. Down the helix E41 appears positioned to interaction with both R287' (residues on FabD will be noted by apostrophes) and Q166', on loop III the large perturbation at E53 corresponds to a salt bridge with K189. The final interactions matching the strong perturbations appear to occur between D56 and K184', with R190' positioned towards the middle of helix II in alignment to bind the backbone of residue I54 and is also near E53. Overall the holo

state appears to suggest that the holo-AcpP•FabD interaction occurs with fewer contacts than the C8. This follows with the holo-AcpP having an unsequestered substrate, explaining the simple binding surface seen in the docking and NMR. Overall it is illuminating to see the matching of perturbations to bound interactions, suggesting that the major population seen experimentally is seen in the model.

The interactions of C8-AcpP are comparatively more spread than the holo-AcpP. There are strong perturbations across the entire region of the protein seen in proximity to the partner in the docked model. With large perturbations seen at loop II, the top of Helix II, loop III, across helix III ending in the middle of helix IV. The C8-AcpP•FabD model displays a significantly more complicated binding surface than the holo-AcpP model. Starting from the N-terminus the first salt bridge seen in the model is between E30 of the C8-AcpP and K286' of FabD, a second electrostatic interaction was next observed between the carbonyl backbone of A34 and R287'.The interactions on Helix II begin at the top of the helix between D35 and Q166', followed by another salt bridge at D38 to K189'. The last helix II interaction is seen between E41 and R190'. The second region of interaction lies near helix III, beginning with a hydrogen bond between E57 and T15' of FabD, this is followed next by a unique K61 which appears positioned to form a salt bridge with E299' of FabD. Next E60 and K279' appear both positioned to hydrogen bond the backbone of the other protein, with E60 positioned next to the amide of K279' and the side chain of 279' near the carbonyl of K61. These constitute significantly more interactions and a much more spread surface of interaction, matching the NMR data. Summed together the NMR and computational modeling agree that the FabD is forming two possible interactions, one holo matching the previously seen structure. The second the newly identified interaction is unique to this study in *E. coli*, but in agreement with the structures of similar complexes.

### 7.3 DISCUSSION

The work described here serves two important purposes. First, we identified an important secondary catalysis in the *E. coli* enzyme FabD and a possible evolutionary basis. In addition, this study also serves as an example of how alternate catalysis can be screened in biosynthetic enzymes. After discovering novel activity, follow-up assays, bioinformatic analysis, biophysical characterization, and

computational modeling can be used to characterize the newly discovered interaction. Classically, pantetheineamide probes have been used to interrogate protein-protein interactions[36]. Here, we have extended the use of these probes to infer information about the substrate preferences of enzymes. This led to bioinformatics studies which inferred more about the evolutionary basis of the activity than could have been speculated without experimental evidence to justify the secondary activity.

Following up on the screening, characterization of the AcpP●FabD interaction matched the identified function to a secondary binding mode, first seen in the unique CSP profile of the C8-AcpP•FabD titration when compared to holo-AcpP. This led us to performing high resolution crystal structures to identify if the protein surface features of the AcpPs could explain the binding modes. This led to the observation of a unique binding mode of the C8-AcpP computationally, with the significantly different structures of holo and C8-AcpP lending to different orientations of interaction. These two orientations follow with known acyltransferase complexes, though the dual binding mode is thus far unseen.

Given the long history of Fatty Acid Synthase research in *E. coli* it is unusual to discover alternate functions for long study proteins. However, as our tools to interrogate alternate or evolutionarily lost functionalities expand its possible the roles of even the core enzymes of cellular metabolism will be expanded. Given the complexity of the E. coli cell it is exciting to think that our understanding of its core systems continues to grow. We hope that studies like this one will follow asA the story of nature continues to grow.

**Chapter 7** is a print of the manuscript in preparation Thomas G. Bartholow, Tony D. Davis, Megan A. Young, Ruben Abagyan, Michael D. Burkart. “Possible evolutionary link and remaining acyl transferase activity of E. coli malonyl-CoA: acyl-carrier protein transacylase FabD.” The dissertation author is the primary author of the manuscript.

### 7.4 References


(1) Carreau, J.-P. [32] Biosynthesis of Lipoic Acid via Unsaturated Fatty Acids. In *Methods in Enzymology*; Vitamins and Coenzymes Part D; Academic Press, 1979; Vol. 62, pp 152–158. https://doi.org/10.1016/0076-6879(79)62212-7.

(2) Rock, C. O.; Jackowski, S. Incorporation and Turnover of Fatty Acids in Escherichia Coli Membrane Phospholipids. In *Advances in Lipobiology*; Gross, R. W., Ed.; JAI, 1996; Vol. 1, pp 39–59. https://doi.org/10.1016/S1874-5245(96)80004-8.

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

(4) Saruwatari, T.; Praseuth, A. P.; Sato, M.; Torikai, K.; Noguchi, H.; Watanabe, K. A Comprehensive Overview on Genomically Directed Assembly of Aromatic Polyketides and Macrolide Lactones Using Fungal Megasynthases. *J. Antibiot. (Tokyo)* **2011**, *64* (1), 9–17. https://doi.org/10.1038/ja.2010.130.

(5) Wu, J.; Zhou, P.; Zhang, X.; Dong, M. Efficient de Novo Synthesis of Resveratrol by Metabolically Engineered Escherichia Coli. *J. Ind. Microbiol. Biotechnol.* **2017**, *44* (7), 1083–1095. https://doi.org/10.1007/s10295-017-1937-9.

(6) Zhao, Y.; Wu, B.-H.; Liu, Z.-N.; Qiao, J.; Zhao, G.-R. Combinatorial Optimization of Resveratrol Production in Engineered E. Coli. *J. Agric. Food Chem.* **2018**, *66* (51), 13444–13453. https://doi.org/10.1021/acs.jafc.8b05014.

(7) Rahman, Z.; Sung, B. H.; Nawab, J.; Siddiqui, M. F.; Ali, A.; Geraldi, A.; Kim, S. C. Enhanced Production of Fatty Acid Ethyl Ester with Engineered FabHDG Operon in Escherichia Coli. *Microorganisms* **2019**, *7* (11). https://doi.org/10.3390/microorganisms7110552.

(8) 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. https://doi.org/10.1038/ncomms14650.

(9) Verwoert, I. I.; Verbree, E. C.; van der Linden, K. H.; Nijkamp, H. J.; Stuitje, A. R. Cloning, Nucleotide Sequence, and Expression of the Escherichia Coli FabD Gene, Encoding Malonyl Coenzyme A-Acyl Carrier Protein Transacylase. *J. Bacteriol.* **1992**, *174* (9), 2851–2857. https://doi.org/10.1128/jb.174.9.2851-2857.1992.

(10) Molnos, J.; Gardiner, R.; Dale, G. E.; Lange, R. A Continuous Coupled Enzyme Assay for Bacterial Malonyl-CoA:Acyl Carrier Protein Transacylase (FabD). *Anal. Biochem.* **2003**, *319* (1), 171–176. https://doi.org/10.1016/s0003-2697(03)00327-0.

(11) Marcella, A. M.; Barb, A. W. A Rapid Fluorometric Assay for the S-Malonyltransacylase FabD and Other Sulfhydryl Utilizing Enzymes. *J. Biol. Methods* **2016**, *3* (4). https://doi.org/10.14440/jbm.2016.144.

(12) Joshi, V. C.; Wakil, S. J. Studies on the Mechanism of Fatty Acid Synthesis. XXVI. Purification and Properties of Malonyl-Coenzyme A--Acyl Carrier Protein Transacylase of Escherichia Coli. *Arch. Biochem. Biophys.* **1971**, *143* (2), 493–505. https://doi.org/10.1016/0003-9861(71)90234-7.

(13) Marcella, A. M.; Barb, A. W. Acyl-Coenzyme A:(Holo-Acyl Carrier Protein) Transacylase Enzymes as Templates for Engineering. *Appl. Microbiol. Biotechnol.* **2018**, *102* (15), 6333–6341. https://doi.org/10.1007/s00253-018-9114-2.

(14) Marcella, A. M.; Barb, A. W. The R117A Variant of the Escherichia Coli Transacylase FabD Synthesizes Novel Acyl-(Acyl Carrier Proteins). *Appl. Microbiol. Biotechnol.* **2017**, *101* (23–24), 8431–8441. https://doi.org/10.1007/s00253-017-8586-9.

(15) 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. https://doi.org/10.1073/pnas.2009805117.

(16) Worthington, A. S.; Burkart, M. D. One-Pot Chemo-Enzymatic Synthesis of Reporter-Modified Proteins. *Org. Biomol. Chem.* **2006**, *4* (1), 44–46. https://doi.org/10.1039/b512735a.

(17) 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* (19), 4720–4724. https://doi.org/10.1039/c8ob03229g.

(18) Ayling, J.; Pirson, R.; Lynen, F. Participation of Covalently Linked Fatty Acyl Coenzyme A Products in the Action of Yeast Fatty Acid Synthetase. *Biochemistry* **1972**, *11* (4), 526–533. https://doi.org/10.1021/bi00754a007.

(19) Knobling, A.; Schiffmann, D.; Sickinger, H. D.; Schweizer, E. Malonyl and Palmityl Transferase-Less Mutants of the Yeast Fatty-Acid-Synthetase Complex. *Eur. J. Biochem.* **1975**, *56* (2), 359–367. https://doi.org/10.1111/j.1432-1033.1975.tb02241.x.

(20) Schreckenbach, T.; Wobser, H.; Lynen, F. The Palmityl Binding Sites of Fatty Acid Synthetase from Yeast. *Eur. J. Biochem.* **1977**, *80* (1), 13–23. https://doi.org/10.1111/j.1432-1033.1977.tb11850.x.

(21) Engeser, H.; Hübner, K.; Straub, J.; Lynen, F. Identity of Malonyl and Palmitoyl Transferase of Fatty Acid Synthetase from Yeast. 2. A Comparison of Active-Site Peptides. *Eur. J. Biochem.* **1979**, *101* (2), 413–422. https://doi.org/10.1111/j.1432-1033.1979.tb19734.x.

(22) Engeser, H.; Hübner, K.; Straub, J.; Lynen, F. Identity of Malonyl and Palmitoyl Transferase of Fatty Acid Synthetase from Yeast. Functional Interrelationships between the Acyl Transferases. *Eur. J. Biochem.* **1979**, *101* (2), 407–412. https://doi.org/10.1111/j.1432-1033.1979.tb19733.x.

(23) Lynen, F. On the Structure of Fatty Acid Synthetase of Yeast. *Eur. J. Biochem.* **1980**, *112* (3), 431–442. https://doi.org/10.1111/j.1432-1033.1980.tb06105.x.

(24) Sztain, T.; Bartholow, T. G.; McCammon, J. A.; Burkart, M. D. Shifting the Hydrolysis Equilibrium of Substrate Loaded Acyl Carrier Proteins. *Biochemistry* **2019**, *58* (34), 3557–3560. https://doi.org/10.1021/acs.biochem.9b00612.

(25) Chen, B.; Wu, X. Probing Substrate Preferences of Depalmitoylases. *Cell Chem. Biol.* **2019**, *26* (1), 3–6. https://doi.org/10.1016/j.chembiol.2018.12.008.

(26) Yoo, E.; Stokes, B. H.; de Jong, H.; Vanaerschot, M.; Kumar, T.; Lawrence, N.; Njoroge, M.; Garcia, A.; Van der Westhuyzen, R.; Momper, J. D.; Ng, C. L.; Fidock, D. A.; Bogyo, M. Defining the Determinants of Specificity of Plasmodium Proteasome Inhibitors. *J. Am. Chem. Soc.* **2018**, *140* (36), 11424–11437. https://doi.org/10.1021/jacs.8b06656.

(27) Lentz, C. S.; Ordonez, A. A.; Kasperkiewicz, P.; La Greca, F.; O'Donoghue, A. J.; Schulze, C. J.; Powers, J. C.; Craik, C. S.; Drag, M.; Jain, S. K.; Bogyo, M. Design of Selective Substrates and Activity-Based Probes for Hydrolase Important for Pathogenesis 1 (HIP1) from Mycobacterium Tuberculosis. *ACS Infect. Dis.* **2016**, *2* (11), 807–815. https://doi.org/10.1021/acsinfecdis.6b00092.

(28) Finzel, K.; Lee, D. J.; Burkart, M. D. Using Modern Tools to Probe the Structure-Function Relationship of Fatty Acid Synthases. *Chembiochem Eur. J. Chem. Biol.* **2015**, *16* (4), 528–547. https://doi.org/10.1002/cbic.201402578.

(29) 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. https://doi.org/10.1038/nature12810.

(30) 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. *Commun. Biol.* **2021**, *4* (1), 1–10. https://doi.org/10.1038/s42003-021-01838-3.

(31) Williamson, M. P. Using Chemical Shift Perturbation to Characterise Ligand Binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. https://doi.org/10.1016/j.pnmrs.2013.02.001.

(32) Neves, M. A. C.; Totrov, M.; Abagyan, R. Docking and Scoring with ICM: The Benchmarking Results and Strategies for Improvement. *J. Comput. Aided Mol. Des.* **2012**, *26* (6), 675–686. https://doi.org/10.1007/s10822-012-9547-0.

(33) ICM—A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation - Abagyan - 1994 - Journal of Computational Chemistry - Wiley Online Library https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.540150503 (accessed 2020 -08 -31).

(34) 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). https://doi.org/10.1073/pnas.2025597118.

(35) 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. *Nat. Chem.* **2020**, *12* (8), 755–763. https://doi.org/10.1038/s41557-020-0491-7.

(36) 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. https://doi.org/10.1016/j.chembiol.2015.09.009.