ABSTRACT

YE, ZHIXIA. New Chemical Biology Approaches for Engineering Natural Product Assembly Lines. (Under the direction of Dr. Gavin Williams).

Natural products have played widespread and critical roles in human well-being for the past 50 years. However, there is a need to produce new natural product analogues that overcome pandemic microbial resistance and to provide new leads with improved pharmacological properties. Discovering new natural products, generating new natural product-like scaffolds, and diversification of existing natural product represent three major approaches. Polyketides are a large class of natural products that have enjoyed particular clinical success. Combinatorial biosynthesis offers a promising approach for engineering polyketide diversity. However, disrupted acyl carrier protein (ACP) interactions with noncognate interacting domains (for example, ketosynthase (KS) and acyltransferase (AT)) of polyketide biosynthetic pathways are one of the major contributing factors to low product yields in combinatorial biosynthesis. One of the major goals of this dissertation is to unveil the molecular details regulating ACP:KS and ACP:AT interactions. An unnatural amino acid based photocrosslinking approach was developed and applied to probe AcpP:FabF interactions. This is the first time an unnatural amino acid has been utilized to study protein-protein interactions within natural product biosynthesis. The interaction epitope on FabF was comprehensively mapped and the key residues responsible for the interaction were identified. In addition, the effect of the AcpP conserved motif and of acyl substrate on AcpP:FabF interaction was probed. To investigate the interaction between a uniquely acyl-CoA promiscuous trans-acting AT (KirCII) and its presumed cognate ACP (ACP5), an acyltransferase coupled fluorescent strain-promoted cycloaddition assay was developed. The
study unveiled the molecular details regulating the KirACP5:KirCII interaction with the KirACP5 interaction epitope extensively mapped and key residues identified. A computational docking model was derived and was corroborated by the KirACP5 epitope and KirCII mutagenesis results. Moreover, this knowledge led to reprogrammed KirACP10:KirCII interactions via a single amino acid substitution. Taken together, understanding the molecular details regulating ACP:KS and ACP:AT interactions enables knowledge-based design of combinatorial biosynthesis. Reprogrammed ACP:AT interactions could enable direct introduction of diversity into natural product structure.

Posttranslational modification to attach a 4’-phosphopantetheine (Ppant) to the conserved serine of carrier proteins (CP’s) is required to convert CP’s into active forms. CP’s are responsible for delivering growing chains and extender units between active sites in natural product assembly lines. As a result, Ppant interacts extensively with the other domains in the pathway. We propose to install Ppant analogs onto CP’s via unnatural amino acid mutagenesis and bioorthogonal conjugation reactions. It is hoped that Ppant analogs that can broaden the substrate promiscuity of domains in natural product biosynthesis will be identified. So far, copper(I)-catalyzed azide-alkyne cycloaddition and ketone-hydrazide conjugation reactions have been evaluated with optimal conditions identified. Conjugation handles including azide, alkyne, and ketone groups have been site-specifically incorporated into ACP through unnatural amino acid mutagenesis, respectively. Small molecules with alkyne, azide, and hydrazide handles have been successfully conjugated onto ACP’s. Overall, the ketone-hydrazide reaction has been determined as the better conjugation approach. Ppant analogs with hydrazide handle are to be synthesized and analyzed for their biological function.
New Chemical Biology Approaches for Engineering Natural Product Assembly Lines

by
Zhixia Ye

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Chemistry

Raleigh, North Carolina

2014

APPROVED BY:

__________________________________  ____________________________________
Gavin J. Williams                      Stefan Franzen
Committee Chair

__________________________________  ____________________________________
Edmond F. Bowden                       Reza A. Ghiladi
BIOGRAPHY

The author, Zhixia Ye, was born in Chengdu, Sichuan Province, P. R. China, which is regarded as one of the best cities to live in China. Zhixia attended Beijing Institute of Technology (BIT), where she was awarded a B.S. in Chemistry in 2005. Her undergraduate research focused on developing an ethanol-based solvothermal route to synthesize LiMnO$_4$ nanocrystals for lithium ion batteries, under the supervision of Professor Lin Ren. In 2009, Zhixia moved to the US to pursue her Ph.D. studies at North Carolina State University under the supervision of Dr. Gavin J. Williams. Her Ph.D. research focused on development and application of various chemical biology approaches to unveil the molecular details regulating key protein-protein interactions in natural product biosynthesis, as well as the impact of prosthetic arm analogs in the biosynthetic process.
ACKNOWLEDGMENTS

The past five years has been a very important phase in my life and there are a lot of people I’d love to show my gratitude to. Firstly, I would like to acknowledge my advisor Dr. Gavin J. Williams for his guidance in my research and mentorship during my graduate study. Joining his laboratory is one of the best decisions I’ve made during my graduate study and I highly appreciate what I learned over the five years. I would also like to thank my committee members, Dr. Bowden, Dr. Franzen, Dr. Ghiladi and Dr. Wallace for their support in my graduate study and critical advice on my dissertation.

I am grateful for every member of the Williams’ laboratory for making the lab a nice and comfortable place to work, and for their generous help when I have problems. I really enjoyed the time in the lab. I want to thank my collaborators, Dr. Ewa M Musiol and Dr. Tilmann Weber (Interfaculty Institute of Microbiology and Infection Medicine, Eberhard Karls University of Tübingen, Tübingen, Germany) for providing the opportunity for me to work on a truly interesting project.

Most importantly, I want to thank my parents for their love and support, for raising me up and providing me with really good educational environment. I want to thank my boyfriend Lifu Wang for his love, support, patience and great cooking. It is a great blessing to share my life with him. Last but not least, I want to thank everyone that I came across during the past five years for making it so wonderful and meaningful.
TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................... ix
LIST OF FIGURES ........................................................................................................ x
LIST OF SCHEMES ...................................................................................................... xiv
CHAPTER 1. Introduction .............................................................................................. 1
  1.1 Polyketide biosynthesis ....................................................................................... 1
  1.2 Combinatorial biosynthesis using modular polyketide assembly lines .......... 4
  1.3 Protein-protein interactions within polyketide synthases ............................... 9

CHAPTER 2. Establishment of a photocrosslinking approach to probe protein-protein interactions ................................................................. 15
  2.1 Introduction .......................................................................................................... 15
  2.2 Materials and methods ....................................................................................... 18
    2.2.1 Gene cloning ................................................................................................ 18
    2.2.2 Protein expression ......................................................................................... 21
    2.2.3 Circular dichroism measurement ................................................................. 23
    2.2.4 Photocrosslinking ......................................................................................... 23
  2.3 Results and discussion ......................................................................................... 23
    2.3.1 Unnatural amino acid mutagenesis ............................................................. 23
    2.3.2 Circular dichroism ....................................................................................... 27
    2.3.3 Photocrosslinking ......................................................................................... 28
    2.3.4 Photocrosslinking specificity .......................................................... ..................... 30
    2.3.5 Probing AcpP:FabF interaction at single residue resolution ................. 32
    2.3.6 Impact of opposite charge on crosslinking efficiency ............................... 36
  2.4 Conclusions .......................................................................................................... 38

CHAPTER 3. Diazirine lysine based photocrosslinking to investigate protein-protein interactions ................................................................. 39
  3.1 Introduction .......................................................................................................... 39
  3.2 Materials and methods ....................................................................................... 40
    3.2.1 Protein expression ....................................................................................... 40
    3.2.2 Circular dichroism ....................................................................................... 41
    3.2.3 Photocrosslinking ....................................................................................... 41
  3.3 Results and discussion ......................................................................................... 41
    3.3.1 Incorporation of lysine derivatives into AcpP ............................................... 41
3.3.2 Diazirine-based photocrosslinking........................................44
3.4 Conclusion..............................................................................47

CHAPTER 4. Mapping a ketosynthase:carrier protein interaction epitope via unnatural amino acid-mediated photocrosslinking........................................48
4.1 Introduction ............................................................................48
4.2 Materials and methods ..........................................................50
  4.2.1 Cloning and expression of ACP’s .........................................50
  4.2.2 Expression of Sfp ................................................................51
  4.2.3 Mutagenesis ........................................................................52
  4.2.4 Alanine scanning mutagenesis of FabF .................................53
  4.2.5 Circular dichroism ..............................................................53
  4.2.6 Sfp catalyzed holo ACP and acyl ACP formation .................53
  4.2.7 Photocrosslinking ..............................................................54
4.3 Results and discussion ............................................................54
  4.3.1 Incorporation of unnatural amino acid into ACP’s ...............54
  4.3.2 Testing orthogonality of ACP:KS interactions with a panel of ACP’s .................54
  4.3.3 Probing contribution of a conserved carrier protein motif to ACP:KS protein interactions ..................................................63
  4.3.4 Probing the contribution of the ACP acyl moiety to ACP:KS interactions ......65
  4.3.5 Mapping the FabF:AcpP interaction epitope ...........................69
4.4 Conclusion ..............................................................................76

CHAPTER 5. Reprogramming acyl carrier protein interactions of an extender unit promiscuous trans-acting acyltransferase ........................................78
5.1 Introduction ............................................................................78
5.2 Materials and methods ..........................................................83
  5.2.1 Cloning and expression of ACP’s .........................................83
  5.2.2 Expression of AT’s .............................................................85
  5.2.3 Alanine scanning mutagenesis of KirACP5 .........................85
  5.2.4 Site-directed mutagenesis and chimeragenesis of ACP’s ..........86
  5.2.5 Site-directed mutagenesis and chimeragenesis of KirCII ..........87
  5.2.6 Expression of DEBS holo-M6TE and DEBS holo-M6TE-AT° ........87
  5.2.7 M6TE chimeragenesis and mutagenesis ...............................88
  5.2.8 Circular dichroism .............................................................88
5.2.9 Chemo-enzymatic synthesis of azidoethylmalonyl-CoA ........................................89
5.2.10 General procedure for KirCII assay .................................................................89
5.2.11 General procedure for strain-promoted azide-alkyne cycloaddition assay .......89
5.2.12 KirCII titration assay .........................................................................................90
5.2.13 KirCII time course assay ..................................................................................90
5.2.14 Homology model of KirACP5 and KirCII .........................................................90
5.2.15 M6TE assay ......................................................................................................90
5.2.16 Trans-AT complementation of M6TE-ATo ......................................................91

5.3 Results and discussion .........................................................................................91
5.3.1 ACP specificity of KirCII ..................................................................................91
5.3.2 Mapping the KirCII:KirACP5 interaction epitope by alanine scanning mutagenesis ........................................................................................................100
5.3.3 Computational docking .....................................................................................108
5.3.4 Engineering non-cognate ACP’s to be recognized by KirCII .........................120
5.3.5 KirCII chimeragenesis ......................................................................................129
5.3.6 KirCII complementation of DEBS M6TE ..........................................................131

5.4 Conclusion ...........................................................................................................135

CHAPTER 6. Probing the cis FluAT1:ACP interaction ............................................136
6.1 Introduction ...........................................................................................................136
6.2 Materials and methods ........................................................................................141
6.2.1 Cloning and expression of FluKS1AT1, FluACP1..............................................141
6.2.2 Cloning and expression of FluAT1 ....................................................................141
6.2.3 Cloning and expression of E. coli FAS proteins ..............................................142
6.2.4 Alanine scanning mutagenesis and expression of FluACP1 .............................143
6.2.5 Site-directed mutagenesis and expression of FluAT1 .......................................143
6.2.6 Mutagenesis of AcpP mutants ...........................................................................144
6.2.7 General procedure for FluAT1 assay ...............................................................144
6.2.8 General procedure for strain-promoted azide-alkyne cycloaddition assay .....144
6.2.9 In vitro fatty acid production ............................................................................144
6.2.10 M6TE complementing ....................................................................................145

6.3 Results and discussion ........................................................................................145
6.3.1 FluKS1AT1 specificity towards ACP’s ..............................................................145
6.3.2 Mapping the FluKS1AT1:KirACP5 epitope by alanine scanning mutagenesis .........................................................147
6.3.3 Construction of standalone FluAT1 .........................................................................................................................152
6.3.4 M6TE complementing ............................................................................................................................................154
6.3.5 Mapping the FluAT1:FluACP1 epitope by alanine scanning mutagenesis ....155
6.3.6 Computational docking ................................................................................................................................................158
6.3.7 Engineering non-cognate ACP’s to be recognized by FluAT1 ..........159
6.3.8 In vitro fatty acid production ..................................................................................................................................161
6.4 Conclusion.....................................................................................................................................................................166

CHAPTER 7. Probing the impact of prosthetic arm analogues on the activity and specificity of natural product biosynthesis ............................................168
7.1 Introduction .................................................................................................................................................................168
7.2 Materials and methods ....................................................................................................................................................174
  7.2.1 Synthesis of para-propargyloxyphenalanine (pPrF) .........................................................174
  7.2.2 Unnatural amino acid mutagenesis ..............................................................................................174
  7.2.3 Cu (I)-catalyzed alkyne-azide cycloaddition (CuAAC) reaction .................................175
  7.2.4 Synthesis of Ppant analog .................................................................................................................176
  7.2.5 General procedure for capping Ppant free thiol by iodoacetamide......176
  7.2.6 General procedure for ketone-hydrazide reaction .................................................................176
7.3 Results and discussion....................................................................................................................................................177
  7.3.1 Unnatural amino acid mutagenesis ........................................................................................177
  7.3.2 Click reaction of ACP2-pAzF with alkyne-rhodamine ......................................................177
  7.3.3 Synthesis of para-propargyloxyphenalanine (pPrF) .................................................179
  7.3.4 Incorporation of pPrF into ACP2......................................................................................180
  7.3.5 Click reaction of ACP2-pPrF with azide-rhodamine .....................................................181
  7.3.6 Synthesis of Ppant analog ........................................................................................................184
  7.3.7 Incorporation of pAzF into AcpP .....................................................................................184
  7.3.8 CuAAC reaction between Ppant analog and pAzF .....................................................185
  7.3.9 CuAAC reaction between Ppant and AcpP-pAzF ......................................................189
  7.3.10 Small molecules based ketone-hydrazide reaction optimization .......................190
  7.3.11 Incorporation of acetyl phenylalanine onto AcpP .............................................194
  7.3.12 Ketone-hydrazide reaction between AcpP-AcF and biotin hydrazide......195
7.3.13 Ketone-hydrazide reaction between AcpP-AcF and texas red hydrazide

7.4 Conclusions

CHAPTER 8. Future work

8.1 Incorporating non-natural extender unit into kirromycin biosynthesis

8.2 Combinatorial biosynthesis of polyketide via trans-AT swapping

8.3 Expanding the conventional repertoire of fatty acid biosynthesis

8.4 Probe deacetylase involved interactions via unnatural amino acid based approach

REFERENCES
LIST OF TABLES

Table 1: Comparison of chemical crosslinkers and photocrosslinkers........................................ 16
Table 2: Vectors and restriction sites for ACP’s and KS’s genes.................................................. 20
Table 3: Primer sequences for PCR amplification and site-directed mutagenesis. .................... 20
Table 4: Primer sequences for site-directed mutagenesis of FabF mutants.................................. 21
Table 5: Mass spectrometry analysis of WT and mutant ACP proteins......................................... 25
Table 6: Summary of photocrosslinking percentage between AcpP and FabF over time...... 29
Table 7: Mass spectrometry analysis of AcpP-pBpa and FabF crosslinking reaction.............. 29
Table 8: Summary of photocrosslinking results between ACP2-pBpa, AcpP-pBpa and KS’s. .... 31
Table 9: MS summary for incorporation of lysine derivatives into AcpP.............................. 43
Table 10: Crosslinking summary of mutant AcpPs with KS’s.................................................... 45
Table 11: Primer sequences for actACP and ScFAS-ACP cloning....................................... 51
Table 12: Primer sequences for ACP mutagenesis................................................................. 53
Table 13: MS analysis of ACP mutants....................................................................................... 54
Table 14: MS analysis of FabF-R206pBpa. *............................................................................ 60
Table 15: Summary for photocrosslinking of FabF-R206pBpa with ACP’s........................... 62
Table 16: Summary for protein-protein interactions involving DSL motif.......................... 63
Table 17: Primers sequences for KirACP’s cloning................................................................. 84
Table 18: Primer sequences for construction of KirACP10/KirACP5 chimeras.................... 86
Table 19: Primer sequences for construction of KirCII/DSZS chimera.................................. 87
Table 20: Primer sequences for M6TE chimera and mutant.................................................... 88
Table 21: MS analysis of azido- and DIBO-KirACP5. .............................................................. 94
Table 22: MS analysis of KirACP’s........................................................................................... 98
Table 23: Identity of ACP residues involved in protein interactions with various domains. 106
Table 24: Predicted AT specificity for Fluvirucin B1............................................................... 1377
Table 25: Primer sequences for FluAT1’s.................................................................................. 1422
Table 26: Primer sequences for E. coli FAS genes cloning..................................................... 1422
Table 27: Amino acid alignment and transacylation activity summary for ACP’s.............. 16060
Table 28: MS analysis for Ppant analog.................................................................................... 1844
Table 29: MS analysis of Ppant-IAM......................................................................................... 1866
Table 30: MS analysis of trityl Ppant and pAzF CuAAC reaction......................................... 1899
Table 31: MS analysis of AcpP-S36AcF and biotin hydrazide conjugation........................... 1955
**LIST OF FIGURES**

Figure 1: Structures of representative polyketides and their biological activities. .......... 2  
Figure 2: Structure of the 6-deoxyerythronolide B synthase (DEBS) .......................... 4  
Figure 3: Approaches for combinatorial biosynthesis of PKS’s ................................ 7  
Figure 4: AT engineering approaches. ................................................................. 8  
Figure 5: Representative protein-protein interactions in natural product biosynthesis. .......... 10  
Figure 6: Epitope mapping for chain elongation and transfer steps in DEBS. ................. 11  
Figure 7: Mechanism-based crosslinker to trap ACP:KS interaction. ............................ 13  
Figure 8: Illustration of specific aim 1 ............................................................... 14  
Figure 9: Benzophenone based photocrosslinking assay ............................................ 18  
Figure 10: SDS-PAGE for ACP’s and ACP-pBpa. .................................................... 24  
Figure 11: Mass spectrometry results for ACP’s and ACP-pBpa. .................................. 26  
Figure 12: CD spectra of wild-type and mutant ACP’s. .............................................. 28  
Figure 13: SDS-PAGE analysis for photocrosslinking between AcpP/AcpP-pBpa and FabF. .... 29  
Figure 14: Mass spectrometry analysis of AcpP-pBpa and FabF crosslinking reaction. .......... 30  
Figure 15: Specificity of photocrosslinking between ACP and KS domains. ..................... 31  
Figure 16: Phylogenetic trees of KS’s and ACP’s. .................................................... 32  
Figure 17: Crystal structure of E. coli FabF (C163A) (PDB: 3I8P) ............................... 33  
Figure 18: CD spectra of FabF mutants. .............................................................. 34  
Figure 19: Crosslinking summary of FabF alanine mutants with AcpP-pBpa. .................... 35  
Figure 20: CD spectra for FabF charged mutants. ................................................... 37  
Figure 21: Crosslinking summary of FabF opposite charge mutants with AcpP-pBpa .......... 37  
Figure 22: Structures of diazirine-based unnatural amino acids. .................................. 40  
Figure 23: Incorporation of lysine derivatives into AcpP .......................................... 42  
Figure 24: MS analysis of AK/DK incorporation into AcpP ....................................... 43  
Figure 25: CD of AcpP-DK and AcpP ................................................................. 44  
Figure 26: Photocrosslinking of mutant AcpP with KS’s ......................................... 45  
Figure 27: MS analysis of AcpP mutant crosslinking with KS’s .................................. 46  
Figure 28: Mapping interaction epitope with photocrosslinking approach. ...................... 50  
Figure 29: Photocrosslinking summary for ACP-pBpa with FabF .................................. 56  
Figure 30: Sequence alignment (A) and phylogenetic analysis (B) of ACP’s ................. 56  
Figure 31: Structure of AcpP (1T8K) .................................................................. 57  
Figure 32: CD spectra for AcpP-N24pBpa. ............................................................ 58  
Figure 33: Photocrosslinking of AcpP-N24pBpa with FabF ....................................... 58  
Figure 34: Probing ACP:KS protein interactions using unnatural amino acid mediated  
photocrosslinking ......................................................................................... 59  
Figure 35: MS analysis of FabF-R206pBpa ............................................................. 60  
Figure 36: CD spectra for FabF-R206pBpa ............................................................. 61  
Figure 37: Photocrosslinking of FabF-R206pBpa with ACP’s .................................... 62  
Figure 38: CD spectra for AcpP alanine mutants. ...................................................... 64  
Figure 39: Summary for photocrosslinking of FabF-R206pBpa with apo-AcpP mutants. .......... 65
Figure 79: KirCII-sdDSZS catalyzed transacylation onto ACP’s. ........................................... 1311
Figure 80: holo-M6TE catalyzed triketide lactone formation. ........................................... 1322
Figure 81: MCAT complementation of holo-M6TE-AT\textsuperscript{o} for triketide lactone production.1323
Figure 82: Triketide lactone production by Sfp acylated apo-M6TE-AT\textsuperscript{o}-ACP6-F38L.... 1334
Figure 83: KirCII complementing holo-M6TE-AT\textsuperscript{o}-KirACP5 chimera. ............................ 1345
Figure 84: Fluvirucin B1 biosynthesis ....................................................................................... 1366
Figure 85: Acyl-CoA specificity for FluKS1AT1. ......................................................................... 1388
Figure 87: Proposed FluAT1 applications. .................................................................................. 14040
Figure 88: An acyltransferase coupled fluorescent cycloaddition assay to probe
FluKS1AT1:ACP interaction........................................................................................................... 14040
Figure 89: ACP specificity of FluKS1AT1. .................................................................................. 1477
Figure 90: Mapping the FluKS1AT1:KirACP5 interaction epitope by alanine scanning
mutagenesis. ................................................................................................................................. 1499
Figure 91: Trans-acylation activities of KirACP5 alanine mutants mapped onto KirACP5
homology model. .......................................................................................................................... 15050
Figure 92: Comparison of KirCII:KirACP5 and FluKS1AT1:KirACP5 interaction epitope.
....................................................................................................................................................... 15050
Figure 93: FluKS1AT1 assay on ACP mutants. ........................................................................... 1511
Figure 94: Homology model of FluKS1AT1............................................................................. 1522
Figure 95: Design of standalone FluAT1 ...................................................................................... 1533
Figure 96: Summary of truncated FluAT1-catalyzed transacylation ........................................... 1544
Figure 97: Homology model of FluACP1 .................................................................................... 1555
Figure 98: Design of FluACP1 alanine mutants. ........................................................................... 1566
Figure 99: Summary for FluACP1 alanine scanning mutagenesis results ................................. 1577
Figure 100: Computational docking of FluACP1 (grey) onto FluAT1 (green) ............................... 1599
Figure 101: FluAT1\textsubscript{PAT} assay on AcpP mutants. ................................................................. 1611
Figure 102: GC-MS analysis of fatty acid methyl ester (FAME) standard. .............................. 1655
Figure 103: GC-MS analysis of WT in vitro fatty acid production .................................................. 1666
Figure 104: ACP-based interactions that take place during each round of chain extension on
modular, type I PKS’s..................................................................................................................... 1688
Figure 105: Post-translational modification of CP’s. ................................................................. 17070
Figure 106: Mechanism-based Ppant analogs to trap ACP:KS interactions ............................... 1711
Figure 107: Click Ppant analogs onto CP’s ................................................................................... 1733
Figure 108: Comparison of the closest mimic of “crypto pantetheine” and Ppant group. 1733
Figure 109: Incorporation of pAzF into ACP2. .......................................................................... 1777
Figure 110: Click reaction between ACP2-pAzF and alkyne rhodamine (AR). ......................... 1788
Figure 111: MS analysis of click reaction. .................................................................................. 1799
Figure 112: NMR spectrum of pPrF. ......................................................................................... 18080
Figure 113: Incorporation of pPrF into ACP2. .......................................................................... 1811
Figure 114: Click reaction of ACP2s with rhodamine dyes. ....................................................... 1822
Figure 115: Structures of rhodamine dyes used for click reaction. ........................................... 1822
Figure 116: Click reaction between ACP2-pPrF and azide-rhodamine (AzR) with varying
ratios of protein to dye. .................................................................................................................. 1833
Figure 117: MS analysis of AcpP-S36pAzF................................................................. 1855
Figure 118: Structure for Ppant and trityl-protected Ppant analogs. ................................. 1866
Figure 119: Optimization of CuAAC reaction................................................................. 1888
Figure 120: Structures for TBTA and THPTA ................................................................. 1899
Figure 121: Effect of varying biotin hydrazide:AcF ratio on AcF and biotin hydrazide conjugation reaction................................................................. 1922
Figure 122: Effect of aniline on AcF-biotin hydrazide conjugation efficiency ....................... 1933
Figure 123: Effect of organic additives on AcF-biotin hydrazide conjugation efficiency. .... 1944
Figure 124: Incorporation of AcF into AcpP at S36 site. ................................................... 1955
Figure 125: Reaction optimization between AcpP-S36AcF and TRH ............................... 1966
Figure 126: Biosynthesis of leinamycin. ............................................................................. 200
Figure 127: Comparison of palmitic acid and palmitic acid analogs produced from conventional (upper) and engineered (lower) fatty acid biosynthesis. ...................... 2011
Figure 128: Type II PKS constituted by E. coli FAS proteins ........................................... 2011
Figure 129: Probe CobB involved interactions via unnatural amino acid based photocrosslinking approach. ................................................................. 2033
LIST OF SCHEMES

Scheme 1: *E. coli* FAS biosynthesis. ................................................................. 49
Scheme 2: *E. coli* FAS biosynthesis. ................................................................. 1633
Scheme 3: Para-propargyloxyphenalanine (pPrF) synthesis. ............................... 1744
Scheme 4: Synthesis of Ppant analog. ................................................................. 1766
Scheme 5: Capping the free thiol by iodoacetamide. ........................................... 1866
Scheme 6: Ketone-hydrazide reaction. ................................................................... 1911
CHAPTER 1. Introduction

1.1 Polyketide biosynthesis

Polyketides represent one of the most structurally diverse and medicinally important classes of natural products and have extensive therapeutic applications such as antibiotics, anticancer, anticholesterol and antiparasite agents (Figure 1). Polyketides are biosynthesized by polyketide synthases (PKS’s) through the successive condensations of acyl-Coenzyme A (CoA) thioesters.\(^1\) PKS’s can be classified into three types based on their structural organization. Type I PKS’s are modular, with distinct, non-iteratively acting domains organized into modules, whereby each module is responsible for incorporation of a two-carbon unit into the final product. Type II PKS’s are composed of discrete, usually monofunctional domains whereby the polyketide is furnished through iterative catalysis of each domain. Type III PKS’s, also referred to as the “chalcone synthase” family, are essentially condensing enzymes that catalyze iterative condensations and produce relatively simple aromatics with modest structural diversity.\(^2\)
Figure 1: Structures of representative polyketides and their biological activities.

6-deoxyerythronolide B synthase (DEBS) is one of the best-characterized type I PKS’s. DEBS is composed of seven modules, including one loading module, distributed across three polypeptides (Figure 2). One module minimally contains three domains for the elongation of the polyketide chain: acyl carrier protein (ACP), β-ketoacyl-acyl carrier protein synthase (KS) and acyltransferase (AT). For ACP’s to function, post-translational modification is required by attaching a 4’-phosphopantetheine (Ppant) prosthetic group onto the conserved Ser. Ppant is responsible for holding the growing polyketide chain and extender unit within each module, and delivering them to other domains in the assembly line during biosynthesis. AT’s charge ACP’s with extender units and are usually regarded as the “gate-keeper” domains of PKS’s due to their apparent strict substrate specificity. KS’s are responsible for the decarboxylative condensations between growing polyketide chains and extender units. Oxidation states and
stereocenters of polyketide products are determined by various tailoring domains including ketoreductase (KR), enoylreductase (ER) and dehydrogenase (DH). Finally, the thioesterase (TE) domain is responsible for catalyzing cyclization (or hydrolysis) of the polyketide chain, releasing the final polyketide product.³

Polyketide biosynthesis is completed unidirectionally from the loading domain through to the last module of the assembly line. It has been reported that the final polyketide structure can be predicted based on the identity and sequence of modules.⁴ This colinearity between polyketide structure and PKS domain organization enables polyketide structures to be designed and predicted based on PKS gene sequences. However, deviations from colinearity are observed in some cases, and in some bacterial type I PKS’s, domains may be used more than once or skipped during biosynthesis.¹
Figure 2: Structure of the 6-deoxyerythronolide B synthase (DEBS). Propionyl starter unit is incorporated in the assembly line through the loading domain (LD), and then the methylmalonyl extender unit is selected by the acyl transferase domain in each module. Biosynthesis is completed unidirectionally from LD through to module 6 and final product is then cyclized and cleaved by thioesterase (TE). The contribution to the final polyketide product is color-coded to the module responsible. The horizontal wavy lines on ACP and KS represent interpolypeptide linkers. Vertical wavy lines underneath ACP’s represent phosphopantetheine (Ppant) group.

1.2 Combinatorial biosynthesis using modular polyketide assembly lines

Microbial resistance to antibiotics has made developing new antibiotics of paramount importance. Polyketides have attracted immense attention due to their broad and potent biological activities. It is expected that by producing novel and diverse polyketides, polyketides with improved clinical values can be discovered.
A module is defined as: “a multi-component system that can be divided into smaller subsystems, which interact with each other and can be predictably interchanged for functional flexibility and variety”. Potential modularity of PKS’s makes combinatorial biosynthesis a remarkably attractive approach to obtain novel polyketides. Combinatorial biosynthesis is defined as the “genetic manipulation of biosynthetic pathways in order to produce new and altered structures”. Current approaches for combinatorial biosynthesis of PKS’s can be classified into four categories. The first comprises inactivation or introduction of reductive domains to adjust the oxidation states of polyketides. Stereocenters can be inverted by replacing or engineering reductive domains with those of different stereospecificity. The second approach is precursor engineering: by feeding the assembly line with “unnatural” starter units; the structure of polyketides can be modified. The third approach is module swapping. By mixing and matching modules within the same polypeptide or modules from other natural product assembly lines, it is potentially possible to produce almost any polyketide at will. Success of the module swapping approach heavily relies on the modularity of PKS’s, which requires minimal structural perturbation and efficient intramodular substrate channeling. Module swapping also has the potential of yielding the largest diversity of polyketides. Assuming the number of modules encodes different two carbon unit is $X$, and the number of two carbon units in polyketides is $n$, the number of polyketides produced in this situation is $X^1 + X^2 + X^3 + \ldots + X^n$. $X$ is still increasing thanks to the continuing discovery of new assembly lines and engineering of currently known modules. It is thus tantalizing to use combinatorial biosynthesis of PKS’s to produce libraries of polyketides with potential clinical applications. Notably, the polyketides produced this way don’t include non-natural extender modifications,
and to introduce the extender modifications, the last approach is required, which is extender unit engineering, involving engineering the substrate specificity of individual AT domains. By altering or expanding the substrate specificity of AT’s, a plethora of extender units could be incorporated into polyketides, enriching the functionalities of polyketide products. Several strategies have been applied in AT engineering for natural product analogs (Figure 4). The most commonly used one is AT domain swapping, whereby the cis-acting AT domain is replaced by an AT homolog with different extender unit selectivity. A variety of polyketide analogs have been obtained via this approach in the early years of AT engineering. However, a general problem associated with this approach is poor product yield, presumably due to disrupted AT:ACP interactions, interference of preexisting KS:ACP interactions, and/or incapability of downstream modules to process the unnatural extender unit. Another strategy is based on generation of hybrid AT’s, where putative extender unit specificity determining regions are replaced by the analogous regions from AT’s with different acyl-CoA specificity. So far, there have been a very limited number of examples for this approach, likely due to the same reasons that explain the failure of AT domain swapping. The strategy with least perturbation to modular structure and protein-protein interactions is site-directed mutagenesis of AT’s, where residues responsible for extender unit specificity are mutated, thus potentially allowing extender unit analogs to be incorporated. Nevertheless, research has shown that the incorporation of non-natural extender unit via this approach was due to reduced AT activity towards the cognate substrate, instead of a shift in substrate specificity. Finally, an alternative to AT mutagenesis involves complementation of the cis-acting AT-null module with a trans-acting AT located external to the PKS. Provided orthogonal substrate specificity between the
trans-AT and PKS cis-AT’s, non-natural extender units could be regioselectively incorporated. Success of this approach is dependent on (1) the availability of trans-AT’s with unique acyl-CoA specificity; (2) manipulation of trans-AT ACP specificity; and (3) capability of the AT-null module to process unnatural substrates.

Figure 3: Approaches for combinatorial biosynthesis of PKS’s. (A) Introduction of reductive domains, the number of reductive domains introduced can be varied and reductive domains initially present in the module can be deleted as well. (B) Precursor engineering, by inactivating the loading domain and module 1, the precursor can be changed from propionyl-CoA to NDK-SNAC analogs, functionality can be incorporated at the R position. NDK (natural diketide) is the natural substrate for module 2. SNAC, N-acetylcysteamine. (C) Module swapping, by replacing upstream or downstream modules with counterparts that synthesize different two carbon unit from the same or different assembly lines, polyketides structures can be expanded. (D) Extender unit engineering, rational design and directed evolution of AT’s could enable extender units of diverse structures to be incorporated. The horizontal arrows indicate genetic manipulations.
So far, all four combinatorial biosynthesis approaches enjoy varying degrees of success. However, a general problem encountered, especially for module swapping, is reduction in product yield. For example, Menzella et al gene synthesized 14 modules from 8 PKS’s and genetically constructed 154 bimodular systems, whereby two modules were connected with the same set of linker sequences. Among them, only about 50% produced the expected polyketide product. It was demonstrated that all 14 modules were catalytically competent in at least one bimodular combination and could process unnatural substrates. Thus, it was most likely that the absence of polyketide production in about half of the combinations was due to faulty protein-protein interactions at the module interface, i.e. ACP:KS interaction. They also demonstrated that by replacing the KS with the cognate KS of the upstream ACP, unproductive bimodular combinations can be rescued, to some extent. The importance of

Figure 4: AT engineering approaches. The domains from different biosynthetic pathway are color-coded. Mutagenesis is represented by red spheres in 3).
maintaining intermodular protein-protein interaction was further demonstrated with trimodular PKS systems, with 95% of them producing expected polyketide products.\textsuperscript{21} In addition, Khosla et al have reconstituted modular activity from separated domains of DEBS. KS3AT3 showed very low enzymatic activity towards ACP6 compared to ACP3, which is its natural ACP substrate.\textsuperscript{22} In all, there is an increasing recognition that protein-protein interactions play an integral role in polyketide biosynthesis.

1.3 Protein-protein interactions within polyketide synthases

To fully understand the compromised protein-protein interactions that arise as a consequence of various module swapping and AT engineering, it is beneficial to identify the epitopes that are responsible for the interactions. Based on the organization of PKS’s (Figure 5), the intermodular interaction between ACP from the previous module and KS from the downstream module and AT:ACP interaction should be the focus of investigation. Intermodular interactions can be either intrapolypeptide or interpolypeptide for type I PKS’s, depending on whether the interacting modules locate on the same polypeptide. Intrapolypeptide and interpolypeptide interactions are mediated by linker regions of proper lengths. Linker regions have been shown to play integral roles in successful combinatorial polyketide production.\textsuperscript{23} However, even in the presence of suitable linkers, the production of a given polyketide can still be quite low, highlighting the importance of intermodular interactions.\textsuperscript{11}
Very recently, various epitopes within the DEBS system have been characterized. Epitopes on ACP2 from DEBS responsible for interpolypeptide chain transfer and epitopes on ACP3 responsible for chain elongation with KS3-AT3 have been identified (Figure 6). A series of chimeric ACP’s designed based on secondary structures was prepared. The epitopes were mapped based on enzymatic activity of these chimeric ACP’s relative to wild-type activity.\textsuperscript{24} With NMR technology, the residues on ACP2 suspected to involve in interpolypeptide chain transfer with KS3AT3 were identified as well.\textsuperscript{4} These results have provided valuable information on the intermodular interactions in DEBS. However, looking beyond the DEBS system, these approaches are not general, and can’t be easily applied to other systems.
Figure 6: Epitope mapping for chain elongation and transfer steps in DEBS.\textsuperscript{24} Chain elongation and transfer were observed only for cognate ACP’s, upon chimeragenesis engineering, non-cognate ACP’s could be reprogrammed to be recognized by KS3AT3 during chain elongation and transfer. A back transfer assay where acyl group was transferred from KS to ACP was adopted to study chain transfer. The epitopes responsible for proper interactions are located on distinct regions of corresponding ACP’s, shown in orange and red oval presentations respectively.

A mechanism-based chemical crosslinker has been specifically developed to trap the weak and transient ACP:KS interaction.\textsuperscript{25} These crosslinkers are Ppant analogues and are able to covalently crosslink with the free thiol in KS active site (cysteine) (\textbf{Figure 7}). Such probes have been used successfully to report ACP:KS interactions from different biosynthetic
pathways. With this approach, it was also demonstrated that the ACP:KS interaction strength correlated with the acyl transfer rate between ACP and KS. This suggests that by identifying mutants with improved protein-protein interactions, the likelihood of identifying mutants with improved enzyme activity is higher. The mechanism-based crosslinkers were also applied in study of DH:ACP interaction from *E. coli* fatty acid biosynthesis. The 1.9Å crystal structure for the DH-ACP complex was characterized and snapshots of ACP in action were presented, whereby, the molecular details of the dehydration process was elucidated. However, lengthy synthetic routes are required to access these mechanism-based crosslinkers. In addition, installation of these Ppant analogs onto the apo-ACP requires the combined action of four enzymes. Moreover, the Ppant analogue structure needs to be carefully designed to be compatible with the substrate specificity of the PKS under investigation. Given the variety of PKS’s and their distinct substrate specificity, a general and high-throughput approach to probe protein-protein interactions in PKS’s is in great demand.
**Figure 7: Mechanism-based crosslinker to trap ACP:KS interaction.** Ppant analogs were installed onto apo-ACP through the combined action of four enzymes, CoaA/D/E and Sfp in the presence of ATP, yielding crypto-ACP. Upon coincubation of crypto-ACP and KS, ACP-KS complex was formed. Crosslinking results were analyzed by gel shift assay.

The overall goal of this dissertation is to improve the outcome of PKS combinatorial biosynthesis by improving our understanding of protein-protein interactions during PKS catalysis. This goal will be achieved through the completion of two specific aims. The **first** is to improve faulty protein-protein interactions that are a result of module swapping. In order to do this, I will a) develop general and high-throughput assays to probe ACP:KS and ACP:AT interactions; b) identify ACP and KS/AT epitopes that are responsible for cognate ACP:KS and ACP:AT interactions; and c) discover mutants with improved noncognate ACP:KS and ACP:AT interactions (**Figure 8**). The **second** specific aim seeks to probe the prosthetic arm requirement of natural product biosynthetic pathways. It is hoped that by identifying functional non-natural prosthetic arms, the substrate promiscuity or catalytic activity of natural product biosynthesis will be broadened.
Figure 8: Illustration of specific aim 1. (A) Identifying the epitopes responsible for cognate ACP:KS interaction. Shown here is the crystal structure of *E. coli* FabF (PDB: 3I8P), a ketosynthase from *E. coli* Fatty Acid Synthase (FAS). (B) Rescuing non-cognate ACP:KS interactions. Based on epitope information obtained from cognate ACP:KS interactions, KS mutants library will be designed, mutants with improved ACP:KS interaction will be identified through a proposed high-throughput assay, among the identified mutants, residues with improved enzymatic activity will be identified through secondary enzyme assay. The residues may be evolved to further improve catalytic activity between noncognate ACP:KS pairs through directed evolution. Shown here is the crystal structure of *E. coli* FabB (PDB: 1DD8), another ketosynthase in *E. coli* FAS, a single chain is shown, and residue colored in yellow is the active site of FabB.
CHAPTER 2. Establishment of a photocrosslinking approach to probe protein-protein interactions

This chapter was adapted from Ye et al, *Mol. BioSyst.* 2011, 7, 3152.

2.1 Introduction

The weak and transient nature of protein-protein interactions within PKS’s makes probing protein-protein interactions particularly difficult. To tackle this problem, a photocrosslinking approach was adopted. By forming a covalent linkage between interacting proteins, the weak interactions can be trapped and further characterization is enabled.

There are generally two classes of crosslinkers. The first one is chemical crosslinkers, like 1, 3-dibromopropanone (DBP). The spacer arm length and functional group specificity of chemical crosslinkers can be adjusted to meet specific requirements. However, nonspecific crosslinking is usually a problem for chemical crosslinkers, especially when applied *in vivo* (*Table 1*).27

The second class is photoreactive crosslinkers, which are usually incorporated through unnatural amino acid mutagenesis. Photoreactive crosslinkers enable better spatial and temporal control than chemical ones.28 There are three structural types of photocrosslinkers used most commonly, benzophenone, aryl azide and diazirine derivatives.
Table 1: Comparison of chemical crosslinkers and photocrosslinkers.

<table>
<thead>
<tr>
<th></th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical crosslinkers</td>
<td>Spacer arm length and functional group can be tailored to specific needs</td>
<td>Low specificity</td>
</tr>
<tr>
<td>Photocrosslinkers</td>
<td>Spatial and temporal control</td>
<td>Proteins may be damaged depending on the wavelength used.</td>
</tr>
</tbody>
</table>

Aryl azide has been genetically incorporated into proteins. Diazirine derivatives were either metabolically or genetically incorporated into proteins expressed in *E. coli* and mammalian cells. Both types have been applied in crosslinking experiments, however, some level of background crosslinking was observed, possibly due to the highly reactive nature of their excited states. A detailed discussion on diazirine derivaties is presented in chapter 3.

Para-benzoyl phenylalanine (pBpa) is a commercially available photocrosslinking unnatural amino acid. Compared to other photophores, like diazido esters, aryl azides and diazirines, benzophenones are chemically stable. Benzophenone will crosslink with nearby C-H bonds upon UV irradiation at 365 nm, this ambient light minimizes damage to the protein under investigation. The excitation of the benzophenone group is reversible, as a result, the extent of crosslinking can be maximized over time. The reaction volume of the benzophenone group is estimated as a sphere centered on the ketone oxygen with a radius of about 3.1 Å. Crucially, this strict spatial requirement by the benzophenone group potentially ensures that only protein pairs with cognate interactions will be crosslinked, since they usually come close to each other when interacting, whereas proteins pairs with non-cognate interaction will not.

pBpa-based photocrosslinking has been applied to probe protein-protein interactions at single residue level and high-resolution epitopes have been mapped. SecA and SecY are two
proteins involved in the process of protein translocation. pBpa was incorporated at different sites in SecY, and SecY mutants were allowed to crosslink in vivo with SecA, one of its interacting partners. By analyzing crosslinking efficiency with each mutant, high resolution epitopes on SecY responsible for interacting with SecA were mapped. pBpa was also used to identify the binding interface responsible for the strong interaction (nanomolar $K_D$) between the prototypical eukaryotic repressor protein Gal80 and the transcriptional activator Gal4. By introducing pBpa at different sites in the entire activation domain in Gal4, an extended binding interface between Gal4 and Gal80 was mapped, which wasn’t identified by conventional deletion and mutagenesis experiments. These results highlighted the utility of pBpa-based approaches to map protein-protein interaction epitopes at high resolution.

In the case of PKS’s, pBpa has not been previously employed to probe protein interactions. In fact, unnatural amino acid mutagenesis has not been previously reported for any biosynthetic machinery that employs ACP’s. Herein, pBpa will be incorporated into ACP’s to probe ACP:KS interactions in natural product biosynthetic pathways. The feasibility and specificity of pBpa-based crosslinking will first be established. Epitopes responsible for cognate ACP:KS interaction can then be mapped with this approach (Figure 9). Faulty ACP:KS interactions will be rescued with the information obtained. It is not expected that all of the mutants that exhibit improved ACP:KS interactions will also display improved catalytic activity. However, those that display improved enzymatic activity will be identified from the pool of mutants with improved ACP:KS interaction through secondary enzyme assays.
Figure 9: Benzophenone based photocrosslinking assay. (A) pBpa (X) is installed into ACP’s through unnatural amino acid mutagenesis, for cognate ACP:KS interaction, upon UV irradiation, ACP and KS will be covalently linked, while for non-cognate ACP:KS interaction, crosslinking should not be observed. (B) Gel shift assay to analyze crosslinking results, an upper band corresponding to ACP-KS complex should be observed upon UV irradiation for cognate ACP:KS interaction.

We will use unnatural amino acid mutagenesis to introduce pBpa into ACP’s. To site-specifically incorporate pBpa, the TAG amber stop codon needs to be introduced in the ACP gene through site-directed mutagenesis. pBpa will be incorporated through an aminoacyl-tRNA synthetase (aaRS)/suppressor tRNA pair that is specific for pBpa and orthogonal to endogenous aaRS/tRNA pairs in the host cells in response to the TAG codon.20,36

As initial proof of principle, the strategy was employed using the E. coli fatty acid synthase (FAS) system given that these proteins express more easily in E. coli than PKS’s. E. coli FAS is a type II system consisting of discrete proteins, responsible for the biosynthesis of fatty acids for lipid membrane formation.37

2.2 Materials and methods

2.2.1 Gene cloning

Genomic DNA of Saccharopolyspora erythraea was a kind gift from Prof Adrian T. Keatinge-Clay (The University of Texas at Austin) The ACP6 gene from DEBS was PCR
amplified from the genomic DNA of *Saccharopolyspora erythraea* using Phire Hot Start II DNA Polymerase (ThermoScientific, Inc.) and the oligonucleotides described in Table 3. The PCR reaction was carried out in a 50-µL reaction mixture consisting of 100 pmol each primer, 0.5 µg of plasmid, 0.25 mM each dNTP, and 1 units of Phire Hot Start II DNA Polymerase (ThermoScientific, Inc.) in the reaction buffer provided by the manufacturer. Amplification involved an initial denaturation step at 98°C for 5 min followed by cycling at 98°C for 1 min, 60°C for 1 min, and 72°C for 0.3 min (the extension rate was 1 min/kb) for 30 cycles, then a final extension for 10 min at 72°C. The PCR product was analyzed by 1% agarose gel and purified by using a Gel Extraction Kit (Biobasic, Inc.) as per the manufacturer’s instructions.

ACP2 gene from DEBS was synthesized and codon optimized by Blue Heron. Genes encoding acyl carrier protein (AcpP) and ketosynthases (FabF, FabB and FabH) from *E. coli* FAS were PCR amplified from genomic DNA of *E. coli* BL21(DE3) as described for DEBS ACP6. See Table 3 for a list of oligonucleotides used. AcpP, ACP2 and ACP6 genes were cloned into pET28a via NdeI and EcoRI sites respectively. FabB and FabH genes were cloned into pET28a via BamHI and SacI sites respectively. FabF was cloned into pETDuet via HindIII and NotI sites. The identities of all genes were confirmed by DNA sequencing from (GENEWIZ, Raleigh).

The conserved phosphopantetheinylation Ser sites of AcpP, ACP2 and ACP6 were mutated to the amber stop codon TAG using the QuikChange protocol (Agilent Technologies, Inc.). *Methanocaldococcus jannaschii* aminoacyl-tRNA synthetase (aaRS)/suppressor tRNA pairs developed and provided by Prof Peter G. Schultz (The Scripps Research Institute, San
were used to incorporate pBpa into AcpP, ACP2 and ACP6 in response to the amber stop codon TAG.

Site-directed mutagenesis was used to prepare FabF mutant genes, including alanine and opposite charge mutants. Identity of mutant genes was verified by DNA sequencing (GENEWIZ, Raleigh).

### Table 2: Vectors and restriction sites for ACP’s and KS’s genes.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Vector</th>
<th>Restriction sites (5' and 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcpP</td>
<td>pET28a</td>
<td>NdeI and EcoRI</td>
</tr>
<tr>
<td>ACP2</td>
<td>pET28a</td>
<td>NdeI and EcoRI</td>
</tr>
<tr>
<td>ACP6</td>
<td>pET28a</td>
<td>NdeI and EcoRI</td>
</tr>
<tr>
<td>FabB</td>
<td>pET28a</td>
<td>BamHI and SacI</td>
</tr>
<tr>
<td>FabH</td>
<td>pET28a</td>
<td>BamHI and SacI</td>
</tr>
<tr>
<td>FabF</td>
<td>pETDuet</td>
<td>HindIII and NotI</td>
</tr>
</tbody>
</table>

### Table 3: Primer sequences for PCR amplification and site-directed mutagenesis.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcpP-NdeI-FOR</td>
<td>AGCTTCGAGATGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>AcpP-EcoRI-REV</td>
<td>AGCTTGTATCGGATTTACCAGGCAGGAAGCGCG</td>
</tr>
<tr>
<td>AcpP-S36TAG-FOR</td>
<td>CGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>AcpP-S36TAG-REV</td>
<td>CGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>AcpP-S36TAG-REV</td>
<td>CGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>FabF-HindIII-FOR</td>
<td>AGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>FabF-NotI-REV</td>
<td>AGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>FabB-BamHI-FOR</td>
<td>AGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>FabB-SacI-REV</td>
<td>AGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>FabH-BamHI-FOR</td>
<td>AGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
<tr>
<td>FabH-SacI-REV</td>
<td>AGCTTTATCGAGATGAGCACTATCGAAGAAGCGCG</td>
</tr>
</tbody>
</table>
Table 4: Primer sequences for site-directed mutagenesis of FabF mutants.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>FabF-N57A-FOR</td>
<td>GGCTTAGAAGATTGTCGTGAGGACATTATCTCGCGC</td>
</tr>
<tr>
<td>FabF-N57A-REV</td>
<td>GCCGCAGATATGCTTCTCAGTACCCGAGAAGC</td>
</tr>
<tr>
<td>FabF-I92A-FOR</td>
<td>GTTCTGCTTCCCTTCAAGGCCAGAATC</td>
</tr>
<tr>
<td>FabF-I92A-REV</td>
<td>GGTGCGTTCTCTCCGTCCGGAACAAAC</td>
</tr>
<tr>
<td>FabF-L208A-FOR</td>
<td>GCGGCAGCTGCAAGGTCTCCAGCAATG</td>
</tr>
<tr>
<td>FabF-L208A-REV</td>
<td>CATTGCTGCCGGCAACGGGCG</td>
</tr>
<tr>
<td>FabF-E225A-FOR</td>
<td>GCCCGTGAGATAATGCCTGATGTTTCGTACTG</td>
</tr>
<tr>
<td>FabF-E225A-REV</td>
<td>CAGTACGAAACCATCAGCGCCGGCAGATATC</td>
</tr>
<tr>
<td>FabF-D227A-FOR</td>
<td>GCGCCAGGTAGGAAACCCGCAAGGGTACTTCT</td>
</tr>
<tr>
<td>FabF-D227A-REV</td>
<td>AGGCTGACGAAGGCTGGCG</td>
</tr>
<tr>
<td>FabF-T307A-FOR</td>
<td>CACCGTACTTCTGCGGCGGCGGATAAG</td>
</tr>
<tr>
<td>FabF-T307A-REV</td>
<td>CTCTATCGCCAGCCGCGAAGAGTACC</td>
</tr>
<tr>
<td>FabF-R206E-FOR</td>
<td>GCTTTGGCGCGGCGGCGGATATC</td>
</tr>
<tr>
<td>FabF-R206E-REV</td>
<td>GTTCCGGGCGGCTGAGC</td>
</tr>
<tr>
<td>FabF-E225K-FOR</td>
<td>GCCCGTGAGATAATGCCGCAAGGGTACTTCT</td>
</tr>
<tr>
<td>FabF-E225K-REV</td>
<td>CAGTACGAAACCATCAGCGCCGGCAGATATC</td>
</tr>
<tr>
<td>FabF-D227K-FOR</td>
<td>GCGCCAGGTAGGAAACCCGCAAGGGTACTTCT</td>
</tr>
<tr>
<td>FabF-D227K-REV</td>
<td>AGGCTGACGAAGGCTGGCG</td>
</tr>
<tr>
<td>FabF-R206E-REV</td>
<td>GTTCCGGGCGGCGGCGGATATC</td>
</tr>
</tbody>
</table>

2.2.2 Protein expression

Plasmid with target gene was transformed into E. coli BL21(DE3) for protein expression. Each protein (ACP’s, FabB, FabF and FabH) was expressed following published procedures. ACP-pBpa was expressed the same way as wild-type protein, and pBpa was included in the culture media at 0.2 mM final concentration. Plasmids harboring each ACP were co-transformed with plasmid pEVOL-pBpa into E. coli BL21(DE3). A single colony was then used to inoculate 3 mL of 2×YT (Sigma-Aldrich, Inc.) media containing 34 μg/mL chloramphenicol and 50 μg/mL kanamycin which was cultured overnight at 37 °C with shaking at 250 rpm. The overnight culture was used to inoculate 300 mL of 2×YT media containing the same antibiotics. 0.2 mM pBpa (final concentration) was dissolved in 1 M NaOH and added to the media. The culture was then incubated at 37 °C with shaking at 250
rpm until the OD$_{600}$ had reached ~0.6. Protein expression was induced by the addition of 0.5 mM IPTG and 0.02% (wt/v) L-arabinose (final concentrations) and was then incubated at 18 °C with shaking at 250 rpm for 18 h. ACP-pBpa was purified using the same procedure as WT ACP.

The expression of FabF mutant proteins differed from that of wild-type FabF. No expression of FabF mutants was observed under the expression condition for wild-type FabF. Plasmid encoding mutant FabF genes were each transformed into *E. coli* BL21(DE3) for protein expression. 3 mL overnight culture was diluted into 300 mL LB, incubated at 37 °C with shaking at 250 rpm until OD$_{600}$ reached 0.6, the cell culture was then cooled down to room temperature and induced with 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). The expression was continued at 18 °C for about 20 hours.

Cells were harvested by centrifugation at 5,000 g for 20 min, and resuspended in 20 mL of 100 mM Tris-HCl (pH 7.4) containing 300 mM NaCl and then lysed by sonication. Proteins were purified by Fast Protein Liquid Chromatography (FPLC) (Bio-Rad, Inc.) with 1 mL Histrap affinity columns (GE Healthcare Life Sciences, Inc.) using the following buffers: wash buffer [20 mM phosphate (pH 7.4) containing 0.5 M NaCl and 20 mM imidazole] and elution buffer [20 mM phosphate (pH 7.4) containing 0.5 M NaCl and 200 mM imidazole]. The purified protein was concentrated using an Amicon Ultra 10,000 MWCO centrifugal filter (EMD Millipore, Inc.). Protein purity was verified by SDS-PAGE and protein concentration was measured by Bradford Protein Assay Kit (Thermo Scientific, Inc.). The masses of mutant proteins were confirmed by electrospray ionization mass spectrometry. The protein samples
for MS were buffer exchanged into 100 mM ammonium acetate with zeba-spin desalting columns (Thermo Scientific, Inc.).

2.2.3 Circular dichroism measurement

Circular dichroism (CD) measurements were performed with a JASCO 810 CD Spectropolarimeter. Samples for CD were buffer exchanged into CD buffer\textsuperscript{42} (10 mM potassium phosphate, 50 mM sodium sulfate, pH 7.4) and the concentration of samples was 0.2 mg/mL. Spectra from 190 nm to 260 nm were scanned at a step of 0.5 nm at 20 °C in a 0.1 cm cuvette, with ten repeats. The scan speed was 100 nm/min.

2.2.4 Photocrosslinking

Photocrosslinking reactions were carried out on 96-well microtiter plate on ice. Reaction volume was 50 μL in PBS buffer; ACP’s were about five fold in excess of KS’s (8.8 μM final concentration for KS’s). Reaction mixture was irradiated at 365 nm with a handheld UV lamp (6 W) for specific lengths of time and then analyzed by SDS-PAGE (4~12%). Percent of crosslinking was determined by analyzing the pixels of bands on the SDS-PAGE using ImageJ\textsuperscript{43}. Percent of crosslinking for each lane on gel was defined as:

\[
\% \text{ crosslinking} = \frac{\text{pixels of crosslinked band}}{\text{pixels of crosslinked band} + \text{pixels of KS band}}
\]

2.3 Results and discussion

2.3.1 Unnatural amino acid mutagenesis

We speculated that the conserved Ser on ACP’s is a good site to install pBpa for our initial proof-of-principle study given that this site is where Ppant is installed. A molecular
docking study revealed Ser was located very close to the active site entrance of the KS.\textsuperscript{44} Thus, pBpa was incorporated into ACP2, ACP6 and AcpP at the conserved Ser sites, yielding ACP2-pBpa, ACP6-pBpa and AcpP-pBpa, which were expressed and purified to homogeneity (Figure 10). In each case, protein identity was confirmed by mass spectrometry (MS). The experimental masses closely matched the calculated masses (Table 5 and Figure 10), confirming that the unnatural amino acid was installed at a single position in each polypeptide. This is the first time that any unnatural amino acid has been incorporated into proteins from natural product biosynthetic pathways.

Figure 10: SDS-PAGE for ACP’s and ACP-pBpa. AcpP is known to run anomalously on SDS-PAGE due to its highly acidic nature, thus, its apparent mass on gel is about 15 kDa while its actual mass is about 10 kDa.\textsuperscript{25} The upper band in lane 6 was contaminant during the purification process for this batch of preparation. The ACP6-pBpa in lane 6 wasn’t used for photocrosslinking. M, marker; 1, AcpP-pBpa; 2, AcpP; 3, ACP2-pBpa; 4, ACP2; 5, ACP6; 6, ACP6-pBpa.
Table 5: Mass spectrometry analysis of WT and mutant ACP proteins. *: The calculated masses have been adjusted for the loss of N-terminal Methionine\textsuperscript{45} (130.194 Da), as observed by MS results.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Calculated mass* (Da)</th>
<th>Experimental mass (Da)</th>
<th>Δ (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP6</td>
<td>11289.98</td>
<td>11289.68</td>
<td>0.30</td>
</tr>
<tr>
<td>ACP6-pBpa</td>
<td>11453.87</td>
<td>11453.41</td>
<td>0.46</td>
</tr>
<tr>
<td>AcpP</td>
<td>10803.31</td>
<td>10802.75</td>
<td>0.56</td>
</tr>
<tr>
<td>AcpP-pBpa</td>
<td>10966.48</td>
<td>10967.13</td>
<td>-0.65</td>
</tr>
<tr>
<td>ACP2-pBpa</td>
<td>20521.73</td>
<td>20521.89</td>
<td>-0.16</td>
</tr>
</tbody>
</table>
Figure 11: Mass spectrometry results for ACP’s and ACP-pBpa. A, AcpP; B, AcpP-pBpa; C, ACP6; D, ACP6-pBpa; E, ACP2-pBpa.
Analysis of the MS results for each mutant ACP showed that incorporation of any canonical amino acid in place of pBpa did not take place. However, in the absence of UAA, full-length mutant ACP’s was observed, which indicated that there was some leaky incorporation; and the protein yield was about 10% of the yield for ACP-pBpa. Overall, this data suggests that the incorporation of natural amino acid at the Ser position was effectively suppressed in the presence of UAA. The yield of each ACP-pBpa correlated with the expression level of the corresponding WT ACP, the better ACP expressed, the more ACP-pBpa could be obtained. Generally, 2-5 mg of ACP-pBpa was obtained per liter of culture.

2.3.2 Circular dichroism

Structural conservation between mutant and wild-type proteins is very important when using mutant proteins to probe protein-protein interactions. Circular dichroism (CD) spectra in the far UV region provides information on the protein secondary structure, which can be utilized to confirm structural conservation between mutant and wild-type proteins.33 From Figure 12, the CD spectra of AcpP and AcpP-pBpa agreed well, indicating they had almost the same structures. Some degree of inconsistency in CD spectra of ACP2 and ACP2-pBpa was observed. However, secondary structure proportions calculated with SELCON33 using the same set of parameters showed the same distribution of secondary structures, suggesting that they shared similar structures.
2.3.3 Photocrosslinking

To demonstrate the feasibility of the benzophenone based photocrosslinking approach, ACP (AcpP-pBpa) and KS (FabF) from E. coli FAS system were subjected to photocrosslinking reaction conditions. Aliquots were taken at different time points and the results analyzed by SDS-PAGE. From Figure 13, after irradiation, a new band appeared for the AcpP-pBpa/FabF mixture. The new band was identified as the AcpP-FabF complex based on the following three observations. First of all, the apparent mass of the new band was about 52 kDa, which was about the sum of the apparent masses of FabF (~40 kDa) and AcpP-pBpa (15 kDa). Second, the presence of the new band was dependent on the presence of both AcpP-pBpa and FabF because no such band was detected when AcpP-pBpa or FabF was irradiated alone (results not shown). Finally, the new band was also dependent on UV irradiation, since the band wasn’t detected at time 0, and increased in intensity over time (Table 6).
Cumulatively, these observations indicated that the new band was the crosslinked product between AcpP-pBpa and FabF.

![Figure 13: SDS-PAGE analysis for photocrosslinking between AcpP/AcpP-pBpa and FabF. Samples were irradiated for 0, 30, 60, and 120 min.](image)

**Table 6: Summary of photocrosslinking percentage between AcpP and FabF over time.**

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>30</th>
<th>60</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Crosslinking</td>
<td>27</td>
<td>33</td>
<td>43</td>
</tr>
</tbody>
</table>

To further validate identification of the putative crosslinked product, the crosslinked product mixture of AcpP-pBpa and FabF was analyzed by MS. The experimental mass closely matched the calculated mass for AcpP-FabF (Table 7 and Figure 14), confirming the new band was the AcpP-FabF complex.

**Table 7: Mass spectrometry analysis of AcpP-pBpa and FabF crosslinking reaction.**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Calculated mass (Da)</th>
<th>Experimental mass (Da)</th>
<th>Δ (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcpP-FabF</td>
<td>56761.20</td>
<td>56760.72</td>
<td>0.48</td>
</tr>
</tbody>
</table>
Figure 14: Mass spectrometry analysis of AcpP-pBpa and FabF crosslinking reaction. The two spectra above showed the deconvoluted MS results. The two spectra below showed the MS results before deconvolution, with the one on the left showing m/z values and the one on the right showing charge state of each corresponding peak.

2.3.4 Photocrosslinking specificity

In order to accurately report interpolyptide interactions between ACP and KS, efficient crosslinking should occur between cognate ACP and KS pairs, while no or little crosslinking should be observed for non-cognate ACP and KS pairs. To demonstrate the photocrosslinking specificity, AcpP-pBpa and KS’s (FabB, FabF and FabH) from E. coli FAS were subjected to UV irradiation.

From Table 8 and Figure 15A, detectable crosslinking was observed for AcpP-pBpa with FabF (76%) and FabB (8%), while FabH didn’t show crosslinking with AcpP. Since FabB and FabF are known to use ACP-bound acyl donors, while FabH uses acetyl-CoA donor, complete absence of detectable crosslinking between AcpP and FabH provided some evidence that our photocrosslinking strategy faithfully reported interpolyptide interactions. Sequence
alignment\textsuperscript{46} of FabB, FabF and FabH (Figure 16A) showed that FabH was distantly related with FabB and FabF, confirming the observed crosslinking results.

To further test this observation, ACP2-pBpa from the type I PKS DEBS was incubated and irradiated, with FabB, FabF, and FabH, respectively. Gratifyingly, photocrosslinking between these combinations was not observed (Table 8 and Figure 15B). From Figure 16B, sequence alignment of ACP from type I and II natural product assembly lines showed these two types of ACP’s are not closely related with each other, supporting our crosslinking results.

Table 8: Summary of photocrosslinking results between ACP2-pBpa, AcpP-pBpa and KS’s. The experimental results correlated well with expected results. N.D.: not detected.

<table>
<thead>
<tr>
<th>AcpP-pBpa (experimental/expected)</th>
<th>ACP2-pBpa (experimental/expected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FabF 76%/+</td>
<td>N.D./-</td>
</tr>
<tr>
<td>FabH N.D./-</td>
<td>N.D./-</td>
</tr>
<tr>
<td>FabB 8%/+</td>
<td>N.D./-</td>
</tr>
</tbody>
</table>

Figure 15: Specificity of photocrosslinking between ACP and KS domains. Coomassie stained SDS-PAGE showing photocrosslinked products between (A) AcpP-pBpa and (B) ACP2-pBpa crosslinking with FabB, FabF and FabH.
2.3.5 Probing AcpP:FabF interaction at single residue resolution

There is some literature precedence of using pBpa to probe protein-protein interaction during protein translocation processes and interaction between a eukaryotic transcription activator and its suppressor, and high-resolution epitopes at single residue level have been mapped.\textsuperscript{25, 35}

It is our ultimate goal to identify the residues responsible for the specific ACP:KS interactions, and with this information in hand, rescue the faulty ACP:KS interactions that are often a result of module swapping. It is known from docking models that ACP docks onto KS near the active site entrance of KS, with the ACP Ppant attached to the conserved Ser inserted into the entrance.\textsuperscript{35} To probe the importance of residues located near the active site entrance, a small panel of FabF alanine mutants was designed (Figure 17). The FabF mutant residues were chosen based on their proximity to the active site entrance (R206, L208, E225, D227 and T307). Residues (N57 and I92) far away from the active site entrance were chosen as negative controls. All selected residues were in loop regions, with the hope that these mutants would
not have significant structural changes on protein structure. Alanine was chosen due to the relatively small size of its side chain, which was unlikely to have much interaction with other nearby side chains and should help delineate the role of each selected residue, particularly the contributions of polar, charged or hydrophobic side chains to protein-protein interactions. All FabF mutants showed similar CD spectra to that of WT, indicating that the mutant structures agreed well with that of WT (Figure 18).

**Figure 17: Crystal structure of E. coli FabF (C163A) (PDB: 3I8P).** Residues colored in blue show the locations of alanine-scanning mutants. N57 and I92 are negative controls; these two residues are far away from active site entrance of FabF, while the other residues were chosen due to their proximity to the entrance. The residue colored in yellow is location for active site of FabF (Cys163).
**Figure 18: CD spectra of FabF mutants.** The mutants generally showed similar CD spectra curves to that of wild-type FabF, indicating there was no structural perturbation in alanine FabF mutants.

It was expected that some of the FabF alanine mutants near the active site entrance would show decreased crosslinking compared to wild-type FabF, presumably because the ACP:FabF interface was disturbed, while others might show similar crosslinking percentages to the wild-type FabF (especially the negative controls) since they were not predicted to be involved in the ACP:FabF interaction. Alternatively, some of the mutants might show increased crosslinking than WT FabF.

Each FabF mutant was crosslinked with AcpP-pBpa under identical conditions, and the crosslinking efficiency of each FabF mutant was determined as a percentage of that with WT FabF, after 1 h of irradiation. From **Figure 19A**, most of the mutants showed similar crosslinking to WT FabF. However, D227A showed slightly lower crosslinking percentage in comparison to WT FabF, while R206A and L208A displayed no detectable crosslinking. Since
the secondary structures of the mutant and wild-type FabF were well conserved, as judged by CD spectroscopy, the absence of crosslinking for R206A and L208A likely indicates their involvement in critical AcpP:FabF interactions. In fact, there is literature evidence that positively charged and hydrophobic residues located near the active site entrance of *E. coli* FAS proteins are important for protein-protein interactions. Notably, the two negative controls had little impact on crosslinking efficiency, compared to mutation at R206 and L208. The clear relationship between FabF mutant position and crosslinking efficiency suggests that pBpa photocrosslinking accurately reports the AcpP:FabF interaction, and that this strategy could be used to fully map the interface by alanine scanning mutagenesis.

Figure 19: Crosslinking summary of FabF alanine mutants with AcpP-pBpa. (A) Crosslinking percentages of FabF mutants with AcpP-pBpa relative to wild-type FabF. (B) Crystal structure of FabF with alanine mutants color-coded to their impact on crosslinking, red/green indicated significant/minor influences, respectively. ND, not detected.
2.3.6 Impact of opposite charge on crosslinking efficiency

Of the initial mutant panel, R206, E225 and D227 were charged residues, and were subsequently mutated into oppositely charged residues (R206E, E225K and D227K) to probe the influence of introducing opposite charge at the potential ACP:KS interface. The expectation was that introduction of an oppositely charged residue at a position involved in the ACP:KS interaction would interfere with the electrostatic interaction between AcpP and FabF, resulting in further decrease in crosslinking percentages compared to their alanine counterparts. CD spectra for these charged mutants suggested conservation of secondary structures (Figure 20). Photocrosslinking was performed with each FabF charged mutant in the identical condition as previously described. No photocrosslinking was detected for R206E mutant, further emphasizing the strict requirement of arginine at 206 site on FabF. Interestingly, E225 and D227, when each was mutated to lysine, showed a modest increase in crosslinking percentage, relative to that of the corresponding alanine mutant (Figure 21). The increase in crosslinking percentage might be due to a more favorable electrostatic interaction between the E225K and D227K with the acidic AcpP, as was observed for some mutants of the ACP substrate for the trans-acting acyltransferase from the disorazole synthase. However, further analysis is necessary to explain the difference in crosslinking percentages of FabF mutants substituted at E225 and D227.
Figure 20: CD spectra for FabF charged mutants. The mutants generally showed similar CD spectra curves to that of wild-type FabF, indicating there was no structural perturbation in charged FabF mutants.

Figure 21: Crosslinking summary of FabF opposite charge mutants with AcpP-pBpa. (A) Crosslinking summary of FabF oppositely charged mutants with AcpP-pBpa relative to wild-type FabF. (B) Crystal structure of FabF with the charged residues highlighted and color-coded to their impact on crosslinking, red/green indicated significant/minor influences, respectively. ND, not detectd.
2.4 Conclusions

A benzophenone-based photocrosslinking approach to probe protein-protein interactions in natural product biosynthesis has been successfully established. The feasibility of crosslinking has been demonstrated with AcpP and FabF from *E. coli* fatty acid biosynthesis. Further, specificity of interactions between FabF and ACP was established by using cognate and non-cognate ACP’s. The interpolypeptide ACP:KS interaction was probed at single residue resolution, revealing that R206 and L208 at the FabF surface likely contribute to the AcpP:FabF interaction epitope. Currently, pBpa is placed at the conserved ACP phosphopanteithylation Ser site. However, pBpa can be placed at virtually any position on the ACP or KS, allowing ACP:KS interactions to be probed for apo-ACP, holo-ACP, acyl-ACP and substrate bound KS, thus enabling the importance of the Ppant group and acyl-substrate for ACP:KS recognition to be probed. This is difficult to achieve for the mechanism-based crosslinkers.\(^{25}\) By constructing a more comprehensive panel of FabF alanine mutants that target the AcpP:FabF interface, a more complete AcpP:FabF interaction epitope can be mapped.
CHAPTER 3. Diazirine lysine based photocrosslinking to investigate protein-protein interactions

3.1 Introduction

Diazirine is a photoreactive group that can be excited at ~350 nm, resulting in N₂ loss and carbene formation. Carbene is highly reactive and can insert into nearby C-H or heteroatom-H bonds.²⁸ Both aromatic and aliphatic amino acids have been modified with diazirines in order to photocrosslink interacting proteins (Figure 22). Compared to bulky aryl diazirine, like trifluoromethyl phenyl diazirine, the small size of alkyl diazirine ensures minimal perturbation to protein-protein interactions. Alkyl diazirine derivatives have been incorporated into proteins in the form of photo-leucine, photo-isoleucine, photo-methionine and diazirine modified lysine.³¹,³² Photo-leucine, photo-isoleucine and photo-methionine have been incorporated into membrane proteins with unmodified mammalian translational machinery based on the structural similarity of photo-Leu, photo-Ile and photo-Met to their natural counterparts, Leu, Ile and Met. With this approach, a previously unknown interaction between a membrane protein and a regulator was discovered. Diazirine modified lysine has been incorporated into proteins both in E. coli and mammalian cells using unnatural amino acid mutagenesis.³² Utility of this unnatural amino acid approach has also been demonstrated by photocrosslinking glutathione S-transferase (GST) monomers into dimers in E. coli, and by photocrosslinking a kinase and its substrate in mammalian cells.³²,⁴⁹
Figure 22: Structures of diazirine-based unnatural amino acids. Photo-Leu, photo-Ile and photo-Met are shown next to Ile, Leu and Met in order to show the structural similarity.

The aromatic side chain of pBpa, used in our earlier photocrosslinking validation experiments, is sterically demanding, and when applied in some other systems, could interfere or destroy particularly sensitive protein-protein interactions.\(^32\) To complement the pBpa-based crosslinking approach, diazirine modified lysine (DK) was explored as an aliphatic photocrosslinker alternative to pBpa.

### 3.2 Materials and methods

#### 3.2.1 Protein expression

\(N^\alpha\)-Allyloxy carbonyl-L-lysine (AK), diazirine modified lysine (DK), PBKPylS plasmid encoding wild-type pyrrolysyl-tRNA synthetase and NDBFKRS-3 plasmid encoding mutant pyrrolysyl-tRNA were kindly provided by the Dr. Deiters’ lab at the University of Pittsburgh.

AcpPS36TAG gene was cloned into pMyo4TAGpylT plasmid via NcoI and NdeI sites, replacing the myoglobin gene (Myo4TAG) with AcpPS36TAG gene, yielding
pAcpPS36TAGpylT. The AcpPS36TAG was appended with C-terminal His tag. pAcpPS36TAGpylT was cotransformed with PBKpysS and NDBFKRS-3 plasmids respectively into *E. coli* Top10 competent cells.

*N*<sup>e</sup>-Allyloxycarbonyl-L-lysine (AK) and diazirine modified lysine (DK) were individually incorporated into AcpP. AK acted as positive control for unnatural amino acid mutagenesis with the wild-type and mutant pyrrolysyl-tRNA synthetase, given its good incorporation efficiency (Deiters et al, unpublished results). AcpP-AK and AcpP-DK were expressed in the same way as that of wild-type AcpP. Protein expression was induced by 0.4% arabinose and 1 mM of AK and DK was used. Pure proteins were obtained and analyzed as previously described. Mutant protein identity was verified by MS.

3.2.2 Circular dichroism

Circular dichroism was performed as described in chapter 2, Section 2.2.3.

3.2.3 Photocrosslinking

Photocrosslinking was performed as described in chapter 2, Section 2.2.4.

3.3 Results and discussion

3.3.1 Incorporation of lysine derivatives into AcpP

It is known that the incorporation efficiency of DK with wild-type pyrrolysyl-tRNA synthetase is very low (Deiters et al, unpublished results). Fortunately, a mutant pyrrolysyl-tRNA synthetase NDBFKRS-3 was recently discovered that displayed increased incorporation efficiency for DK (Deiters et al, unpublished results). AK and DK were incorporated into AcpP
using the wild-type and mutant synthetase to compare the incorporation enhancement offered by the mutant synthetase in the case of AcpP.

![Image](image.png)

**Figure 23: Incorporation of lysine derivatives into AcpP.** (A) Structures for \(N^\varepsilon\)-Allyloxy carbonyl-L-lysine (AK) and diazirine modified lysine (DK). (B) SDS-PAGE of expressed AcpP-AK and AcpP-DK, S, soluble fraction; E, elution fraction; W/M, expression of the wild-type or mutant pyrrolysyl-tRNA synthetase, respectively.

From **Figure 23B**, increased AcpP yield was observed for the mutant synthetase over the wild-type, especially for DK. Densitometry analysis by ImageJ\(^{43}\) showed about two-fold expression increase for AK and about twenty-fold increase for DK. All four purified samples from **Figure 23B** were analyzed by MS (**Table 9** and **Figure 24**), revealing that in three cases, incorporation of the unnatural amino acid was detected. No AcpP-DK mass was detected when using the WT synthetase/tRNA pair, likely due to its poor incorporation efficiency towards DK.
Table 9: MS summary for incorporation of lysine derivatives into AcpP. W/M, incorporation with wild-type/mutant synthetase. Mass for AcpP-DK with wild-type synthetase was not detected (N.D.). The experimental mass closely matched the calculated mass, indicating AK and DK were successfully incorporated. *: The calculated masses have been adjusted for the loss of N terminal Methionine36 (130.194 Da), as observed by MS results.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Calculated mass* (Da)</th>
<th>Experimental mass (Da)</th>
<th>Δ (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcpP-AK/W</td>
<td>9645.12</td>
<td>9644.62</td>
<td>0.50</td>
</tr>
<tr>
<td>AcpP-AK/M</td>
<td>9645.12</td>
<td>9644.58</td>
<td>0.54</td>
</tr>
<tr>
<td>AcpP-DK/W</td>
<td>9687.17</td>
<td>N.D.</td>
<td>--</td>
</tr>
<tr>
<td>AcpP-DK/M</td>
<td>9687.17</td>
<td>9686.72</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Figure 24: MS analysis of AK/DK incorporation into AcpP. (A) AcpP-AK/W; (B) AcpP-AK/M; (C) AcpP-DK/W; (D) AcpP-DK/M.

Circular dichroism was performed to confirm the structural conservation between AcpP-DK and Acp (Figure 25). Gratifyingly, the CD spectra of AcpP and AcpP-DK agreed well, indicating that they shared similar secondary structure, even though the His-tag was at the N-terminus for AcpP and at the C-terminus for AcpP-DK.
Figure 25: CD of AcpP-DK and AcpP. The mutants generally showed similar CD spectra curves to that of wild-type AcpP, indicating there was no structural perturbation in AcpP-DK.

3.3.2 Diazirine-based photocrosslinking

Having confirmed the identity of each mutant protein, AcpP-DK, in parallel with AcpP-pBpa, was then crosslinked with KS’s from the *E. coli* FAS to compare the utility of the two photocrosslinking amino acids. Crosslinking was analyzed by SDS-PAGE (Figure 26) and the results are summarized in Table 10.
Figure 26: Photocrosslinking of mutant AcpP with KS’s. (A) AcpP-pBpa crosslinking with FabB, FabF and FabH. (B) AcpP-DK crosslinking with FabB, FabF and FabH.

<table>
<thead>
<tr>
<th>Table 10: Crosslinking summary of mutant AcpPs with KS’s.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>AcpP-pBpa</td>
</tr>
<tr>
<td>AcpP-DK</td>
</tr>
</tbody>
</table>

From Figure 26 and Table 10, FabF crosslinked with AcpP-pBpa to a greater extent than with FabB and FabH, which was in agreement with our previous experiment (Table 8). Interestingly, upon incubation of FabH with AcpP-pBpa, a new protein band was observed at ~80 kDa, which was not consistent with the anticipated FabH:AcpP crosslinked mass (~45 kDa). One of the possible explanations is more than one copy of AcpP-DK was crosslinked with FabH. Further investigation is required.

When AcpP-DK was crosslinked with FabF and FabB, a new band with apparent mass corresponding to that of the AcpP-KS complex appeared after UV irradiation. The crosslinking
percentages of AcpP-DK with FabB and FabF were about 3-fold less than that of AcpP-pBpa with FabB and FabF. In addition, crosslinking percentages didn’t increase over time. This is likely due to the highly reactive nature of the carbene group and the irreversible excitation of diazirines. Crosslinking was not observed when FabH was used as the crosslinking partner. Interestingly, in addition to the high molecular weight bands corresponding to AcpP crosslinked with FabB/FabF, an additional band was observed with a molecular weight slightly greater than that of AcpP-DK upon UV irradiation. This protein might be a result of carbene inserting into a heteroatom-H bond from small molecules in the reaction mixture.  

Crosslinking of FabF with AcpP-DK was analyzed by MS. Gratifyingly, the experimental mass corresponding to AcpP-DK-FabF complex closely matched the expected mass (Figure 27), indicating that the ~80 kDa band was indeed the AcpP-DK-FabF complex.

![Figure 27: MS analysis of AcpP mutant crosslinking with KS’s. The experimental mass (55453.21 Da) corresponding to AcpP-DK-FabF complex closely matched the expected mass (55453.42 Da).](image-url)
3.4 Conclusion

For the first time, diazirine modified lysine has been incorporated into an ACP. Preliminary crosslinking results using ACP and KS’s from the *E. coli* FAS indicated that diazirine is a viable photocrosslinker for probing protein-protein interactions. Because of the highly reactive nature of the carbene observed in these crosslinking experiments, diazirine-based photocrosslinkers may be better suited for probing protein-protein interactions that are stronger than that of the AcpP:KS investigated here, given strong interactions would be expected to be more specific.
CHAPTER 4. Mapping a ketosynthase:carrier protein interaction epitope via unnatural amino acid-mediated photocrosslinking

This chapter was adapted from Ye et al, Biochemistry, submited.

4.1 Introduction

In type II FAS’s and PKS’s, each catalytic enzyme is located on discrete proteins. The growing fatty acid and polyketide chains are covalently bound as thiol esters to an ACP, which shuttles the intermediates between each catalytic site. Type II FAS’s involve the iterative action of each enzyme to construct the final fatty acid in which every acyl building block is fully reduced. In type II PKS’s, building block acyl groups are not always fully reduced. Fatty acid biosynthesis in E. coli is depicted in Scheme I. The ACP interacts in trans with every other domain. Notably, interaction between ACP and ketosynthase (KS) domains enables chain elongation and translocation. Determining the molecular recognition details that govern such processes is critical to understanding the catalytic processes that occur during PKS and FAS biosynthesis, yet very little is known regarding the molecular interaction between ACP and KS domains of type II systems. Interaction between such domains is likely programmed via recognition of the acyl substrates and/or specific epitopes on the ACP. However, the requisite weak and transient nature of such interactions renders crystallographic analysis a difficult method to provide such information. Indeed, there are no reported crystal structures of complexes between ACP and KS from any type II PKS or FAS. Mutagenic studies have implicated helix II of ACP as being an important recognition motif for the ACP:KS interaction.
interaction, as well as interactions between other domains, and ACP:FabH docking studies have provided further evidence for the general involvement of this motif.

Moreover, type II fatty acid biosynthesis are a validated target for antibacterial drugs. Fatty acid biosynthesis genes are essential for bacterial survival. Isoniazid and triclosan targeting the enoyl-ACP reductase have long been in clinical use. Natural products like cerulenin and thiolactomycin targeting the condensing enzymes have served as the scaffolds for designing new and more specific keto-acyl ACP synthase inhibitors and were used extensively in the biochemical study of E. coli fatty acid biosynthetic pathway. Platensimycin, produced by Streptomyces platensis, displays strong, broad-spectrum Gram-positive antibacterial activity by selectively targeting FabF/FabB, inhibiting cellular lipid
biosynthesis. Platensimycin is now in preclinical trials. Cumulatively, these results suggest the significance of investigating the fatty acid biosynthesis process.

Previously, we reported the use of unnatural amino acid mutagenesis to install the photocrosslinking amino acid para-benzoyl-L-phenylalanine (pBpa) in place of the phosphopantetheine prosthetic arm of the ACP.\(^{61}\) This allows detection and quantification of ACP:KS interactions by SDS-PAGE analysis of photocrosslinked products (Figure 28). Here, we used this approach to a) demonstrate orthogonality of ACP:KS interactions using a comprehensive panel of carrier proteins, b) probe the contribution of carrier protein acyl identity and conserved DSL motif to ACP:KS interactions, and c) completely map the ACP:KS interaction epitope with single residue resolution.

Figure 28: Mapping interaction epitope with photocrosslinking approach. KS alanine mutants (e.g. 1, 2 and 3) are crosslinked with ACP-pBpa. Analysis by SDS-PAGE reveals varying degree of influences the residues have on ACP:KS interaction (highlighted with red, magenta, and yellow spheres).

4.2 Materials and methods

4.2.1 Cloning and expression of ACP’s

actACP and ScFAS-ACP were amplified from genomic DNA of *Streptomyces coelicolor* and were cloned into pET28a via NdeI and EcoRI sites with N-terminal His-tag. E. coli.
coli DH5α cells were used for plasmid amplification. DNA sequencing confirmed identity of genes. ACP2(Φ) was amplified from pET28a-ACP2 plasmid. PCR was carried out as described in Chapter 2.2.1.

Table 11: Primer sequences for actACP and ScFAS-ACP cloning.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>actACP-NdeI-FOR</td>
<td>AGCTTCCATATGGCAACCCTGCTGACCACC</td>
</tr>
<tr>
<td>actACP-EcoRI-REV</td>
<td>AGCTTCCGAATTTCATGCGCCCTCGGCC</td>
</tr>
<tr>
<td>ScFAS-ACP-NdeI-FOR</td>
<td>AGCTTCCATATGGCCGACTCAAGGAAGAGATCG</td>
</tr>
<tr>
<td>ScFAS-ACP-EcoRI-REV</td>
<td>AGCTTCCGAATTTCAGGCTGATCGCTTACAGGGGC</td>
</tr>
<tr>
<td>ACP2(Φ)-NdeI-FOR</td>
<td>AGCTTCCATATGCCTGCCTGATCGCTTACAGGGGC</td>
</tr>
<tr>
<td>ACP2(Φ)-EcoRI-REV</td>
<td>AGCTTCCGAATTCTTAGCCCCAATTCAGCGTCAGAAAGCC</td>
</tr>
</tbody>
</table>

Plasmids containing actACP or ScFAS-ACP genes were transformed into E. coli BL21(DE3). A single colony was then used to inoculate 3 mL of LB (Luria Burtani) medium containing 50 μg/mL kanamycin which was cultured overnight at 37 °C with shaking at 250 rpm. The overnight culture was used to inoculate 300 mL of LB media containing the same antibiotics and was then incubated at 37 °C with shaking at 250 rpm until the OD₆₀₀ had reached ~0.6. Protein expression was induced by the addition of 0.5 mM IPTG and was then incubated at 18 °C at 250 rpm for 18 h. actACP and ScFAS-ACP were purified following the same procedure as WT AcpP (described in Chapter 2).

4.2.2 Expression of Sfp

Sfp was expressed as previously described. briefly, pET28a-Sfp plasmid was transformed into E. coli BL21(DE3). A single colony was then used to inoculate 3 mL of LB (Luria Burtani) medium containing 50 μg/mL kanamycin which was cultured overnight at 37 °C with shaking at 250 rpm. The overnight culture was used to inoculate 500 mL of LB media
containing the same antibiotics and was then incubated at 37 °C at 250 rpm until the OD<sub>600</sub> had reached ~0.6. Protein expression was induced by the addition of 0.5 mM IPTG and was then incubated at 18 °C with shaking at 250 rpm for 18 h. Sfp was purified following the same procedure as WT FabF (as described in Chapter 2).

4.2.3 Mutagenesis

TAG codon was introduced into actACP and ScFAS-ACP at the conserved Ser site via site-directed mutagenesis. Mutations were confirmed by DNA sequencing. pBpa was incorporated in response to TAG codon as described in Chapter 2. actACP-S42pBpa and ScFAS-ACP-S41pBpa were expressed the same way as wild-type protein, and 0.2 mM pBpa was used.

TAG codon was introduced into AcpP at N24, D35 and L37 positions via site-directed mutagenesis. Mutations were confirmed by DNA sequencing. AcpP-N24pBpa, AcpP-D35pBpa and AcpP-L37pBpa were prepared following the same procedure as AcpP-S36pBpa.

AcpP mutants (D35A and L37A) were prepared by site-directed mutagenesis following the QuikChange protocol. Mutations were confirmed by DNA sequencing. Plasmids harboring each AcpP mutant were transformed into E. coli BL21(DE3). A single colony was then used to inoculate 3 mL of LB (Luria Burtani) medium containing 50 μg/mL kanamycin which was cultured overnight at 37 °C with shaking at 250 rpm. The overnight culture was used to inoculate 500 mL of LB media containing the same antibiotics and was then incubated at 37 °C at 250 rpm until the OD<sub>600</sub> had reached ~0.6. Protein expression was induced by the addition of 0.5 mM IPTG and was then incubated at 18 °C with shaking at 250 rpm for 18 h. AcpP alanine mutants were purified following the same procedure as WT AcpP.
Table 12: Primer sequences for ACP mutagenesis.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>actACP-S42TAG-FOR</td>
<td>GACATCGGGTACGACTAGCTCGCCCTGATGGAG</td>
</tr>
<tr>
<td>actACP-S42TAG-REV</td>
<td>CTCCATCAGGCGGAGCTAGTGATCACCAGTGC</td>
</tr>
<tr>
<td>ScFASACP-S41TAG-FOR</td>
<td>CCTGGACGTCGACTAGCTGCATGCTCGATCGG</td>
</tr>
<tr>
<td>ScFASACP-S41TAG-REV</td>
<td>CTCGACCAGGACAGCTAgTGCACTCGGACATCGG</td>
</tr>
<tr>
<td>AcpP-D35TAG-FOR</td>
<td>GAAGACCTGGGCGGCGTACCGTACACCGTGG</td>
</tr>
<tr>
<td>AcpP-D35TAG-REV</td>
<td>CAACGCGTCAAGAGACTACGCGCCAGGTGCTTC</td>
</tr>
<tr>
<td>AcpP-L37TAG-FOR</td>
<td>CCTGGGCGCGGATTCTGACACCGGAGG</td>
</tr>
<tr>
<td>AcpP-L37TAG-REV</td>
<td>GCTCAACGCTTGTCACCAGAATCCGCGCCAG</td>
</tr>
<tr>
<td>AcpP-N24TAG-FOR</td>
<td>GCAGGAAGAAGTGTTACCTAGAATGCTTCTTTTCGCCAAG</td>
</tr>
<tr>
<td>AcpP-N24TAG-REV</td>
<td>CTTCACCGAAGAGCATTCTAGGTAACCTCTTCCTGC</td>
</tr>
<tr>
<td>AcpP-D35A-FOR</td>
<td>GAAGACCTGGGCGGCGGCTCTTGACACCGG</td>
</tr>
<tr>
<td>AcpP-D35A-REV</td>
<td>CAACGCGTCAAGAGACTACGCGCCAGGTGCTTC</td>
</tr>
<tr>
<td>AcpP-L37A-FOR</td>
<td>CCTGGGCGCGGATTCTGACACCGGAGG</td>
</tr>
<tr>
<td>AcpP-L37A-REV</td>
<td>GCTCAACGCTTGTCACCAGAATCCGCGCCAG</td>
</tr>
</tbody>
</table>

4.2.4 Alanine scanning mutagenesis of FabF

Primers for alanine scanning mutagenesis of FabF were designed with lengths between 25 and 35 bp, while melting temperatures were between 60 and 75 °C. Site directed mutagenesis was performed following the QuikChange protocol (Agilent Technologies, Inc.). Mutations were confirmed by DNA sequencing. FabF mutant proteins were expressed and purified as described in Chapter 2.

4.2.5 Circular dichroism

Circular dichroism was performed as described in Chapter 2.

4.2.6 Sfp catalyzed holo ACP and acyl ACP formation

Sfp catalyzed holo- and acyl-ACP formation was performed in 100 μL of reaction mixture containing 50 mM Tris·HCL (pH 8.8), MgCl₂ (10 mM), DTT (5 mM), coenzyme A/malonyl-/octaonyl-coenzyme A (1 mM), apo-ACP (100 μM), and Sfp (20 μg) at 25 °C for 5 hours. Production of holo-, malonyl- and octaonyl-ACP was confirmed by ESI-MS.
4.2.7 Photocrosslinking

Photocrosslinking was performed as described in Chapter 2. Each crosslinking experiment was performed twice. % crosslinking for a specific FabF mutant was normalized based on % crosslinking for WT FabF with the same AcpP-pBpa mutant to obtain % WT.

4.3 Results and discussion

4.3.1 Incorporation of unnatural amino acid into ACP’s

pBpa was incorporated at the conserved Ser site in actACP and ScFAS-ACP, given this is where the prosthetic arm is installed and previous photocrosslinking results have demonstrated this to be a reasonable location. N24 site on AcpP was selected as a neutral control for ACP:KS photocrosslinking since N24 was distant from the DSL motif.\(^5\)

The experimental mass matched closely to the calculated mass of ACP mutant proteins in the MS analysis, confirming the successful incorporation of pBpa (Table 13).

**Table 13: MS analysis of ACP mutants.** *: The calculated mass is adjusted to include loss of N-terminal methionine (-130.194 Da).\(^5\) #: the mass presented is monoisotopic mass.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Calculated mass (Da)</th>
<th>Experimental mass (Da)</th>
<th>(\Delta) (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEBS ACP2-S50pBpa</td>
<td>20521.73*</td>
<td>20521.89</td>
<td>-0.16</td>
</tr>
<tr>
<td>DEBS ACP6-S45pBpa</td>
<td>11453.87*</td>
<td>11453.41</td>
<td>0.46</td>
</tr>
<tr>
<td>actACP-S42pBpa</td>
<td>11444.64*</td>
<td>11444.69</td>
<td>-0.05</td>
</tr>
<tr>
<td>ScFAS-ACP-S41pBpa</td>
<td>11113.46*</td>
<td>11113.42</td>
<td>0.04</td>
</tr>
<tr>
<td>EcFAS-AcpP-S36pBpa</td>
<td>10966.48*</td>
<td>10966.93</td>
<td>-0.45</td>
</tr>
<tr>
<td>EcFAS-AcpP-N24pBpa</td>
<td>10933.26**</td>
<td>10933.72#</td>
<td>-0.46</td>
</tr>
</tbody>
</table>

4.3.2 Testing orthogonality of ACP:KS interactions with a panel of ACP’s

Crosslinking was carried out using WT FabF and each pBpa-modified potential interaction partner ACP. Efficient photocrosslinking was detected between FabF and EcFAS-
AcpP-S36pBpa, while crosslinked bands were also detected for the carrier proteins DEBS ACP6-S45pBpa and ScFAS-ACP-S41pBpa. In contrast, photocrosslinking was not observed for DEBS ACP2-S50pBpa or actACP-S42pBpa (Figure 29). The results were in agreement with literature reports that detail ScFAS-ACP could serve as substrate for E. coli fatty acid synthase. Helix II (HII) has been suggested to be highly involved in the ACP:KS interaction, and this region appears to be poorly conserved among various ACP’s. For example, HII of EcFAS-AcpP and ScFAS-ACP contain several negatively charged residues (Figure 30A). In contrast, DEBS ACP6 presents an essentially neutral HII surface. Meanwhile, the HII surface for DEBS ACP2 are positively charged, and Arg59 is present at the site equivalent to AcpP Glu41. For actACP, the overall HII surface is neutral overall, with equal number of negative and positive charges, nevertheless, a serine and arginine was present at positions equivalent to Glu48 and Glu49 on AcpP. The presence of positively charged residues in DEBS ACP2 and actACP equivalent to glutamate residues in AcpP might account for the absence of photocrosslinking between FabF and DEBS ACP2-S50pBpa or actACP-S42pBpa. In addition, phylogenetic analysis showed DEBS ACP6 to be more closely related to EcFAS-AcpP, compared to DEBS ACP2 and actACP (Figure 30B).
Figure 29: Photocrosslinking summary for ACP-pBpa with FabF.

Figure 30: Sequence alignment (A) and phylogenetic analysis (B) of ACP’s. Secondary structure for AcpP was highlighted above the sequences and H1 is denoted as α2, the numbering of AcpP residues is shown above AcpP sequence.
In order to provide further evidence that the photocrosslinking assay faithfully reports ACP:KS protein interactions, the site of pBpa installation in EcFAS-AcpP was moved to a different position not expected to result in photocrosslinking. Accordingly, pBpa was site-selectively installed at the N24 site, which is located on the face of EcFAS-AcpP distal to the conserved phosphopantetheinylation site (Figure 31). AcpP amino acid numbering was based on the pdb file, 1T8K. Although N24 was located in the middle of Loop I (L1), molecular dynamics studies showed N24 site displayed remarkable stability compared to the N-terminal and C-terminal of L1, and this stability was attributed to the backbone hydrogen bonding of residues N24, A26, and F28 with the backbone atoms of Q66, V65, and T63, respectively.\(^{63}\)

**Figure 31: Structure of AcpP (1T8K).** DSL motif and N24 are highlighted with sticks presentation. The AcpP structure on the right was rotated 90° towards the reader for better view.

Comparison of the EcFAS-AcpP-N24pBpa CD spectra to that of the wild-type AcpP indicated no significant structural changes upon introduction of the unnatural amino acid at N24 site (Figure 32). As expected, no crosslinking was detected after AcpP-N24pBpa was
irradiated in the presence of FabF (Figure 33). This result emphasizes the precision and accuracy with which specific ACP:KS interactions can be probed by carefully selecting the pBpa installation position.

Figure 32: CD spectra for AcpP-N24pBpa. The mutants generally showed similar CD spectra curves to that of wild-type AcpP, indicating there was no structural perturbation in AcpP-N24pBpa.

Figure 33: Photocrosslinking of AcpP-N24pBpa with FabF. No photocrosslinked band was detected after 0.5 h and 1 h irradiation.
Potentially, the ACP epitope that interacts with the cognate KS domain could be mapped by introducing a series of alanine mutations throughout the pBpa-modified ACP and subsequently determining photocrosslinking efficiency with the cognate WT KS. To circumvent production of each subsequent ACP alanine mutant as a pBpa-modified mutant protein, it would be more efficient to introduce pBpa into the KS domain, such that a single pBpa-modified KS could then be used to probe influence of various ACP conserved motifs and residues towards the KS:ACP interaction (Figure 34). This strategy comes with the additional advantage that now, various acyl-substrates, or Ppant analogues, could be installed at the conserved phosphopantetheinylation site of the ACP. In this way, for the first time, the role that the Ppant group and acyl-substrate plays in controlling the FabF:AcpP interaction epitope can be established.

![Figure 34: Probing ACP:KS protein interactions using unnatural amino acid mediated photocrosslinking.](image)

Previously, R206 in FabF was suggested to be the residue involved in covalent cross-linkage with AcpP-pBpa given that no crosslinking was observed between FabF-R206A and
EcFAS-AcpP-S37pBpa. Accordingly, R206 is an ideal site to install pBpa into FabF and the mutant FabF-R206pBpa was over-expressed and purified, and FabF-R206pBpa identity was confirmed by ESI-MS analysis (Table 14 and Figure 35). The secondary structure of FabF-R206pBpa was not significantly different from the WT FabF, as judged by CD-spectroscopy (Figure 36).

**Table 14: MS analysis of FabF-R206pBpa.** *: The calculated mass is adjusted to include loss of N-terminal methionine (-130.194 Da)⁴⁵.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Calculated mass (Da)*</th>
<th>Experimental mass (Da)</th>
<th>Δ (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FabF-R206pBpa</td>
<td>45887.99</td>
<td>45889.09</td>
<td>-1.1</td>
</tr>
</tbody>
</table>

**Figure 35: MS analysis of FabF-R206pBpa.**
Gratifyingly, upon photocrosslinking FabF-R206pBpa in the presence of apo-EcFAS-AcpP, a higher molecular weight protein band consistent with the photocrosslinked FabF:AcpP was observed (Figure 37). The percentage photocrosslinking for FabF-R206pBpa and apo-EcFAS-AcpP (33%) was less than half of that for AcpP-S36pBpa and FabF (76%). The difference in photocrosslinking efficiency might be attributed to the inherent flexibility of AcpP, which would allow it to present pBpa at a better position for photocrosslinking with FabF, while the FabF structure is likely more stable compared to AcpP. Upon incubation of FabF-R206pBpa with other members of the ACP panel, crosslinking was only detected with apo-ScFAS-ACP with comparable efficiency to apo-EcFAS-AcpP (Figure 37 and Table 15). In conclusion, this data demonstrates that FabF suitably modified with pBpa can be used to faithfully report orthogonal ACP interactions in the same way as the analogous strategy using
pBpa-modified ACP. This demonstration now enables identification of individual ACP residues that contribute to maintenance of the ACP:KS interaction.

**Figure 37: Photocrosslinking of FabF-R206pBpa with ACP’s.** DEBS ACP2(Φ) refers to ACP2 without C-terminal linker region. Aliquots were removed from photocrosslinking mixtures after 30 min and 1 hour of irradiation. Different electrostatic properties of ACP presumably contributed to different coomassie staining extent.

**Table 15: Summary for photocrosslinking of FabF-R206pBpa with ACP’s.** N.D., not detected.

<table>
<thead>
<tr>
<th>ACP’s</th>
<th>DEBS ACP2</th>
<th>DEBS ACP6</th>
<th>actACP</th>
<th>AcpP</th>
<th>ScFAS-ACP</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Photocrosslinking</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>33</td>
<td>29</td>
</tr>
</tbody>
</table>
4.3.3 Probing contribution of a conserved carrier protein motif to ACP:KS protein interactions

The ACP phosphopantetheinylation site is part of a highly conserved Asp-Ser-Leu (DSL) motif located at the N-terminus of helix II on the ACP surface. Since the Ppant delivers its acyl cargo to the cognate KS, this motif must be in close proximity to the KS during chain elongation transfer. Studies using a variety of approaches have highlighted the importance of the overall DSL motif in ACP involved interactions, as summarized in Table 16. For example, a computational docking study involving the *Streptomyces coelicolor* malonyl Coenzyme A-ACP transacylase (MCAT) and actACP showed that Leu from the DSL motif inserted into the hydrophobic pocket adjacent to the MCAT active site entrance, and Asp from the same motif formed a salt bridge with a conserved lysine on MCAT. In another study, mutation of Asp or Leu from the DSL motif of actACP to Ala showed reduced binding to *Streptomyces coelicolor* group I acyl carrier protein synthase (ACPS), particularly for Leu, which showed a ten-fold decrease in binding affinity. Nevertheless, few studies directly probed influence of DSL motif on KS:ACP interactions, and computational docking represents the major approach applied so far.

Table 16: Summary for protein-protein interactions involving DSL motif.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D35</td>
<td>D</td>
<td>D</td>
<td>N</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>S</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>S</td>
</tr>
<tr>
<td>S36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References to be included.
To evaluate if the pBpa-based photocrosslinking approach could be used to report the role of the DSL motif in ACP:KS interactions, the D35A and L37A mutants of EcFAS-AcpP were prepared. The secondary structures of D35A and L37A were not significantly different from the wild-type apo-EcFAS-AcpP, as judged by CD-spectroscopy (Figure 38).

Figure 38: CD spectra for AcpP alanine mutants. The mutants generally showed similar CD spectra curves to that of wild-type AcpP, indicating there was no structural perturbation.

Next, the apo-EcFAS-AcpP mutants D35A and L37A were each incubated with FabF-R206pBpa and were irradiated for photocrosslinking. Both D35A and L37A displayed very low photocrosslinking efficiency. Notably, D35A displayed ~5-fold lower crosslinking efficiency, compared to the WT apo-AcpP, while crosslinking of the L37A mutant could not be distinguished from the background of the assay (Figure 39). These results indicated that D35 and L37 played an intimate role in AcpP:FabF interaction. Our results are consistent with
literature evidence that indicates the DSL motif is involved in protein-protein interactions with other interaction partners, as summarized in Table 16.

![Figure 39: Summary for photocrosslinking of FabF-R206pBpa with apo-AcpP mutants.](image)
The photocrosslinking was performed twice. Average % photocrosslinking and standard deviation for each mutant and WT AcpP were calculated and plotted.

4.3.4 Probing the contribution of the ACP acyl moiety to ACP:KS interactions

FAS enzymes catalyze a series of reactions in a specific sequence within one biosynthetic cycle (Scheme I). It has been suggested that the nature of covalently attached intermediate induces a specific conformational change in the carrier protein that may help promote the correct protein-protein interactions.\(^5^0\) Significant efforts have been invested in studying the structural features for apo-, holo- and acyl-ACP’s. NMR studies using apo- and holo-AcpP showed that structures of these two forms were essentially the same, with perturbations present only in regions surrounding the phosphopanthetheine attachment site.\(^8^0\)
In addition, an NMR structural study of acyl-AcpP’s suggested the acyl chains were sequestered into the hydrophobic core formed by H_{II} and H_{IV} of AcpP and the hydrophobic core would expand to accommodate the elongating acyl chains. Similar ACP structural features were observed for ACP’s from PKS systems. Comparison of the {\textsuperscript{1}H-{\textsuperscript{15}}N} heteronuclear single quantum coherence (HSQC) spectra for holo- and malonyl-actACP showed very weak interaction between the malonyl phosphopantetheine moiety and actACP, and overlaying of holo-actACP with malonyl-actACP showed small root-mean-square deviation (RMSD) values. Additionally, the malonyl group remained solvent exposed. In stark contrast, significant chemical shift was observed for octanoyl-actACP in helix I, II, III and IV, and the octanoyl group was sequestered within actACP, with the thioester portion remained solvent exposed and available for transacylation by KS.

In order to probe the potential influence of phosphopantetheinylation and acyl-group identity on AcpP:FabF protein interactions, the broad specificity phosphopantetheinyl transferase Sfp was used to prepare holo-, malonyl-, and octanoyl-derivatives of the wild-type EcFAS-AcpP and each D35A/L37A mutant (Figure 40). Malonyl-CoA was chosen as it is the extender unit used in fatty acid biosynthesis, while octanoyl-CoA is a robust substrate for FabF and mimics to some extent an extended fatty acid intermediate. However, even with excess Sfp, it proved difficult to afford malonylated and octanoylated AcpP, AcpP-D35A and AcpP-L37A in quantitative yield (data not shown). Thus, photocrosslinking efficiency was compared only within each apo-, holo-, malonyl- and octaonyl- groups.
Figure 40: Sfp mediated holo-, malonyl- and octaonyl-AcpP formation. Sfp catalyzes the transfer of phosphopantetheine/acyl-phosphopantetheine from CoA/acyl-CoA onto apo-ACP, yielding holo-/acyl-ACP’s.

Photocrosslinking was performed with FabF-R206pBpa and apo-, holo-, malonyl- and octaonyl-WT AcpP, D35A and L37A mutants, respectively (Figure 41). For holo-AcpP, D35A and L37A mutant, the photocrosslinking efficiency for holo-AcpP-D35A mutant was comparable to holo-AcpP, while holo-AcpP-L37A showed one third the efficiency of holo-AcpP. This difference might be attributed to disruption of the hydrophobic interaction between Leu and the Ppant moiety, rendering the holo-AcpP-L37A mutant in an unfavorable state to crosslink with FabF-R206pBpa. In addition, L37 on AcpP might also be engaged in intimate hydrophobic interactions with FabF. A similar trend was observed for malonyl-AcpP’s, albeit with lower photocrosslinking efficiency, in agreement with an NMR study which showed holo- and malonyl-ACP shared similar structural features. Interestingly, significant differences were observed across the octanoyl-AcpP series. Octanoyl-AcpP-D35A displayed essentially the same photocrosslinking efficiency as octaonyl-AcpP, whereas the photocrosslinking
efficiency for octanoyl-AcpP-L37A was less than 5% that of octanoyl-AcpP. These results suggest that the structural changes that accompany octanoylation play a significant role in rescuing the poor interaction between AcpP-D35A mutant and FabF. The absence of improvement for octanoyl-AcpP-L37A was probably attributed to the disturbance of hydrophobic interaction between Leu and Ppant moiety, and the integral role that L37 plays in interacting with FabF. Overall, the results suggest that the hydrophobic interaction between Leu and Ppant moiety is very important for proper AcpP structural changes when modified with different acyl Ppant groups. This result might also suggest that hydrophobic interaction by Leu-37 is probably more important than electrostatic interaction by Asp-35 at the KS:ACP interface.

Figure 41: Summary for photocrosslinking of FabF-R206pBpa with AcpP mutants. The photocrosslinking was performed twice. Average % photocrosslinking and standard deviation for each mutant and WT AcpP were calculated and plotted.
4.3.5 Mapping the FabF:AcpP interaction epitope

Currently, identifying protein interaction epitopes involved in ACP interaction has been approached mainly via computational docking. To the best of our knowledge, no KS epitope for AcpP:FabF interaction has been experimentally mapped. To map the epitope on FabF that is involved in the AcpP:FabF interaction, a broader panel of FabF alanine mutants was first designed based on analysis of the FabF crystal structure (Figure 42). CD spectra for representative FabF mutants matched closely to the WT spectra, suggesting conservation of secondary structures (Figure 43). Next, each FabF mutant was incubated with AcpP-pBpa and irradiated at 365 nm for 1 hour and the photocrosslinked products analyzed by SDS-PAGE and ImageJ.

**Figure 42: Design of FabF mutants.** The crystal structure of FabF bound with platensimycin A1 (yellow sticks) is presented here (pdb file: 3I8P). The selected mutants were shown as spheres.
Figure 43: CD spectra for FabF mutants. The mutants generally showed similar CD spectra curves to that of wild-type FabF, indicating there was no structural perturbation.

Of the positions chosen for analysis, the majority of them displayed over 50% photocrosslinking efficiency compared to WT FabF (Figure 44). T210A, R211A, D213A and T307A displayed over 90% crosslinking efficiency compared to WT FabF. E225A and D227A displayed 80% and 70% photocrosslinking relative to WT FabF, respectively. To evaluate the influence of electrostatic properties at these two locations, opposite charge and conserved charge mutants were also designed for E225 and D227 sites of FabF. No significant change in crosslinking efficiency was observed for these mutants (E225D, E225K, D227E and D227K), implying that E225 and D227 probably lie at the periphery of the interaction epitope. Notably, two residues (R206 and L208) failed to yield a photocrosslinked product upon mutagenesis to alanine, and these two residues were the same as those identified in the preliminary analysis described in Chapter 2. To further probe the importance of these two FabF residues, each position was mutated to glutamine and lysine, respectively. No crosslinking could be detected.
for the R206E, R206K, L208E and L208K mutants, suggesting that Arg and Leu are strictly required at positions 206 and 208, respectively, for maintaining the FabF:AcpP interaction. Interestingly, the crosslinking efficiency for A45E and A45K differed from each other two-fold, indicating the requirement of a neutral or positively charged residue at the A45 site to maintain interaction with AcpP. FabF-S209A mutant displayed 60% photocrosslinking efficiency as WT FabF, which is likely attributed to its close location to R206 and L208, and R206 and L208 are presumed to be the key interaction sites. Taken together, the mutagenesis results helped portray a clear and specific interaction epitope for FabF:AcpP interactions (Figure 44).

Figure 44: Photocrosslinking summary for AcpP-S36pBpa with FabF mutants. The photocrosslinking was performed twice. Average % photocrosslinking and standard deviation for each mutant and WT AcpP were calculated and plotted. Photocrosslinking percentage for WT FabF was set to 100. The identified epitope (highlighted in red) is shown as surface presentation on the right.
FabF-catalyzed chain elongation occurs in two steps (Figure 45). FabF first catalyzes transacylation of the growing acyl chain from AcpP onto itself, followed by decarboxylative condensation between acyl-FabF and malonyl-AcpP. The epitopes on FabF involved in transacylation and condensation steps haven’t been characterized before and are thus investigated here. In addition, given identification of D35 and L37 as possible interaction locants on AcpP, the role that the acyl moiety plays in determining the interaction epitope with FabF could now be probed, since the phosphopantetheine prosthetic arm is available for acylation if pBpa is installed at D35 or L37.

![Figure 45: FabF transacylation and condensation scheme.](image)

Figure 45: FabF transacylation and condensation scheme. Octanoyl is presented as the acyl substrate. Octanoyl group is first transferred from octanoyl-AcpP to FabF during the transacylation process, producing octanoyl-FabF; FabF then catalyzes the condensation between malonyl-AcpP and octanoyl-FabF, yielding decaonly-AcpP.

AcpP-D35pBpa and AcpP-L37pBpa were prepared accordingly. CD spectra for AcpP-D35pBpa and AcpP-L37pBpa matched closely to WT AcpP, suggesting similar secondary structures (Figure 46).
Figure 46: CD spectra for AcpP-D35pBpa and AcpP-L37pBpa. The mutants generally showed similar CD spectra curves to that of wild-type AcpP, indicating there was no structural perturbation.

Next, the holo-, malonyl- and octaonyl-forms of AcpP-D35pBpa and AcpP-L37pBpa were prepared via Sfp-mediated phosphopantetheinylation, and photocrosslinking between apo-, holo- and acyl-AcpP-pBpa in the presence of FabF was performed (Figure 47). Interestingly, the crosslinking efficiencies for apo-, holo-, malonyl- and octanoyl-AcpP-L37pBpa were not significantly different from each other, and were close to AcpP-S36pBpa photocrosslinking efficiency (Figure 29), and were each at least two-fold better than that for AcpP-D35pBpa. In contrast, photocrosslinking was barely detectable for apo-AcpP-D35pBpa, and presence of the Ppant arm or malonyl-Ppant group improved the efficiency over 10-fold compared to apo-AcpP-D35pBpa. The highest crosslinking was obtained upon octanoylation of AcpP-D35pBpa, in agreement with the previous observation that octanoylation could rescue poor AcpP-D35A mutant:FabF interactions. The relatively conserved hydrophobic mutation
of Leu-37 to pBpa potentially contributes to the high crosslinking efficiency of the various AcpP-L37pBpa forms tested, while mutation of Asp-35 to pBpa might significantly disrupt the original electrostatic interaction formed by Asp-35. High crosslinking efficiency is required for accurate mapping of the interaction epitope, thus, malonyl- and octaonyl-AcpP-L37pBpa were chosen to map the FabF epitopes involved in transacylation and condensation.

Figure 47: Summary for photocrosslinking of apo-, holo-, malonyl- and octaonyl-AcpP-pBpa with FabF. The photocrosslinking was performed once.

Photocrosslinking between the previously described panel of FabF surface mutants (e.g. see Figure 44) and malonyl- and octaonyl-AcpP-L37pBpa was performed (Figure 48). Most FabF mutant displayed similar crosslinking efficiency for malonyl-AcpP-L37pBpa and octaonyl-AcpP-L37pBpa. However, there were some exceptions. C163 is the site on FabF responsible for holding the growing chain from AcpP during chain transfer. The
photocrosslinking efficiency for C163A and malonyl-AcpP-L37pBpa was approximately 3-fold higher than that for C163A and octanoyl-AcpP-L37pBpa, which was in agreement with FabF’s enzymatic mechanism. It has been reported that C163 is essential for the chain transfer process, but is dispensable for decarboxylation of malonyl-ACP. Another important observation was that R206K showed decent crosslinking for both malonyl-AcpP-L37pBpa and octanoyl-AcpP-L37pBpa, which was in contrast to AcpP-S36pBpa based photocrosslinking, implying that AcpP acylation was potentially capable of rescuing the faulty FabF mutant:AcpP interaction. Overall, the results presented here suggested that the FabF epitopes involved in transacylation and condensation steps are quite conserved, and the photocrosslinking results correlated with the mechanistic features of FabF, with the importance of the AcpP acyl group in FabF:AcpP interaction further demonstrated.
Figure 48: Summary for photocrosslinking of FabF mutants with malonyl- and octanoyl-AcpP-L37pBpa. The photocrosslinking was performed twice. Average % photocrosslinking and standard deviation for each mutant and WT AcpP were calculated and plotted. Photocrosslinking percentage for WT FabF was set to 100.

4.4 Conclusion

The pBpa-based photocrosslinking assay initially developed in Chapter 2 has been expanded to faithfully report a broader panel of KS:ACP interactions and ACP substrates. The importance of a conserved ACP DSL motif was evaluated, which confirms that aspartate and leucine play an integral role in AcpP:FabF interactions. The influence of the phosphopantetheine and acyl phosphopantetheine moiety on AcpP:FabF interaction was also investigated. Presence of phosphopantetheine and malonyl-phosphopantetheine on AcpP moderately improved AcpP:FabF photocrosslinking, in comparison to apo-AcpP. Modification of AcpP with octanoyl-phosphopantetheine resulted in a substantial enhancement in photocrosslinking efficiency, potentially attributed to the structural changes upon
octanoylation of AcpP. The FabF epitopes in AcpP:FabF interaction during chain transfer and elongation was probed. The epitopes are mainly conserved during chain transfer and elongation, and the photocrosslinking results agree with KS catalytic mechanism.
CHAPTER 5. Reprogramming acyl carrier protein interactions of an extender unit promiscuous trans-acting acyltransferase

This chapter was adapted from Ye et al, Chem Biol. 2014, in press.

5.1 Introduction

A growing number of PKS’s are being discovered that include trans-AT’s that are housed external to the PKS modules. Such trans-acting AT’s (trans-AT’s) at the genetic level could be easily transferred between biosynthetic systems. Subsequently, given their “stand-alone” nature, trans-AT’s might constitute a valuable alternative to strategies which involve cis-AT domain swapping, and there is much interest in developing trans-AT’s for use as tools in synthetic biology platforms aimed at the regioselective modification of polyketides. Yet, one critical barrier to the use of trans-AT’s as tools for polyketide diversification is our poor understanding of trans-AT extender unit and ACP specificity. For example, the native substrate for many trans-AT’s is malonyl-CoA (Figure 49), and these enzymes are unlikely to display activity towards other acyl-CoA’s. Furthermore, while some trans-AT’s that use malonyl-CoA install the extender unit at multiple positions within the resulting polyketide and are ACP promiscuous, others install unusual extenders at limited positions in the polyketide and are likely ACP specific (Figure 49). Thus, the ACP specificity of trans-AT’s needs to be better understood so that regioselectivity of extender unit installation can be manipulated. Although evidence of the ability to manipulate ACP interactions with other domains is emerging, little is known regarding trans-AT:ACP
interactions. Indeed, only two trans-AT structures are available,\textsuperscript{71, 96} in addition to that of a bi-functional trans-AT/decarboxylase.\textsuperscript{97} Docking models have helped propose the ACP interaction epitopes of two malonyl-CoA specific trans-AT’s.\textsuperscript{96, 71}

**Figure 49: Substrate specificity of trans-AT’s.** (A) Examples of polyketides biosynthesized via trans-AT’s. The contribution of each trans-AT, and the corresponding substrate are shown color-coded. (B) Examples of acyl-thioester substrates for trans-AT’s.

Recently, the promiscuity of several trans-AT’s towards a panel of acyl-CoA’s\textsuperscript{98} was probed by the Williams laboratory. One target, KirCII from kirromycin biosynthesis (**Figure 50**), is responsible for installation of the C28 ethyl moiety, via ethylmalonyl-CoA.\textsuperscript{93} KirCII has the distinction of being the only characterized trans-AT to naturally utilize a non-malonyl acyl-CoA substrate.\textsuperscript{99} All other unusual trans-AT extender units are introduced into polyketides linked to ACP’s (**Figure 49**). The Williams lab discovered that in addition to ethylmalonyl-
CoA, KirCII can utilize several other acyl-CoA’s, including ethyl-, propargylmalonyl-, allyl-, phenyl- and azidoethyl-CoA, making KirCII the most promiscuous trans-AT known. Kirromycin biosynthesis involves the action of another trans-AT (KirCI) responsible for malonyl installation onto ACP’s within 11 unique modules of the kirromycin biosynthetic assembly line. According to this biosynthetic logic, it is likely that the acyl-CoA and ACP specificity of KirCII is orthogonal to that of KirCI while KirACP5 has been identified as an ACP substrate for KirCII. These features would present a valuable opportunity to elucidate the molecular basis for ACP specificity of a uniquely acyl-CoA promiscuous trans-AT.

Figure 50: Biosynthesis of kirromycin. The putative biosynthetic contribution of KirCI (malonyl-CoA, red) and KirCII (ethylmalonyl-CoA, purple) to the final structure of kirromycin is highlighted.
Ultimately, the ability to manipulate trans-AT:ACP interactions could be harnessed by the recently discovered promiscuity of KS’s from the 6-deoxyerythronolide synthase (DEBS) towards diverse extender units, including propargylmalonyl-, phenyl-, propyl-, hydroxyl- and azidoethylmalonyl-CoA.\textsuperscript{101,102} KirCII does not utilize methylmalonyl-CoA, which is the natural and most efficient substrate for the DEBS cis-AT domains. Accordingly, KirCII could be used for site-specific incorporation of unnatural extender units into the polyketide structure by complementation of, for example, an AT-null mutant of module six of DEBS (M6TE-AT\textsuperscript{o}) (\textbf{Figure 51} and \textbf{Figure 52}).

\textbf{Figure 51: holo-M6TE catalyzed triketide lactone formation.} In the presence of diketide-SN\textsubscript{Ac} and methylmalonyl-CoA, holo-M6TE catalyzes the formation of triketide lactone. SN\textsubscript{Ac}, N-acetylcystamine.
Figure 52: KirCII complementing holo-M6TE-AT°. The catalytic serine in AT6 was mutated to alanine, yielding the AT° mutant, KirCII was introduced to complement the AT° mutant, for the production of triketide lactone analogs.

KirCII’s capability to transfer azidoethyl- and propargyl-CoA onto KirACP5 enabled the development of an acyltransferase coupled fluorescent click assay for the investigation of KirCII:KirACP5 interaction (Figure 53). Strain-promoted azide-alkyne cycloaddition (SPAAC) is emerging as an attractive alternative to the classic Cu(I)-catalyzed azide-alkyne cycloadditions (CuAAC) because toxic metal is not required and the reaction is highly efficient even in a very complex environment and proceeds efficiently at ambient temperatures. Density functional theory (B3LYP) calculations of the transition states of cycloadditions of phenyl azide with acetylene and cyclooctyne indicate that the fast rate of the strain promoted cycloaddition is due to a lower energy required for distorting the 1,3-dipole and alkyne into the transition-state geometry. 4-dibenzocyclooctynol (DIBO) has been shown to react fast with azido-containing compounds. Given that azidoethylmalonyl-CoA is a decent substrate for KirCII, the fluorescent Alexa Fluor® 647 DIBO alkyne (Life Technologies, Inc.) was
chosen as a secondary reporter for KirCII catalyzed acyl transfer. Molecular details regulating the KirCII:KirACP5 interaction could be revealed via this approach.

![Diagram showing the development of an acyltransferase coupled fluorescent cycloaddition assay to probe KirCII:KirACP5 interaction. KirCII first catalyzes the transfer of azidoethylmalonyl (AzEM) group onto KirACP5 from AzEM-CoA, a fluorescent cycloaddition reaction is then performed between AzEM-KirACP5 and Alexa Fluor® 647 DIBO. The assay was analyzed by SDS-PAGE and in-gel fluorescence.]

Figure 53: Development of an acyltransferase coupled fluorescent cycloaddition assay to probe KirCII:KirACP5 interaction. KirCII first catalyzes the transfer of azidoethylmalonyl (AzEM) group onto KirACP5 from AzEM-CoA, a fluorescent cycloaddition reaction is then performed between AzEM-KirACP5 and Alexa Fluor® 647 DIBO. The assay was analyzed by SDS-PAGE and in-gel fluorescence.

5.2 Materials and methods

5.2.1 Cloning and expression of ACP’s

The genes encoding kirromycin ACP’s were amplified from cosmids 1C24 and 2C23 using HotStar HighFidelity polymerase (Qiagen, Inc.) and the primers listed in Table 17. Fragments were cloned in pET30Ek/LIC (Novagen/Merck Millipore, Inc.) using the manufacturer’s protocol. Cloning was performed by our collaborator Ewa M Musiol (Interfaculty Institute of Microbiology and Infection Medicine, Eberhard Karls University of Tübingen, Tübingen, Germany).
### Table 17: Primers sequences for KirACP’s cloning.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACP0pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP0pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP1pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP1pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP2pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP2pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP3pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP3pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP4pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP4pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP5pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP5pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP7pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP7pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP8pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP8pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP9pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP9pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP10pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP10pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP11pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP11pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP12pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP12pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP13pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP13pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP14pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP14pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
<tr>
<td>ACP15pET30LIC-FOR</td>
<td>GACGACGACAAAGATGTGGGATGTGACTCCGTC</td>
</tr>
<tr>
<td>ACP15pET30LIC-REV</td>
<td>GAGGAGAAGCCCGGTTACGGATGGCCAGGCCGCCGCCGAGAA</td>
</tr>
</tbody>
</table>

Plasmids harboring each kirromycin ACP were co-transformed with plasmid pSU20-Sfp into *E. coli* BL21(DE3). A single colony was then used to inoculate 3 mL of LB (Luria Burtani) medium containing 34 μg/mL chloramphenicol and 50 μg/mL kanamycin which was cultured overnight at 37 °C with shaking at 250 rpm. The overnight culture was used to inoculate 300 mL of LB media containing the same antibiotics and was then incubated at 37 °C with shaking at 250 rpm until the OD$_{600}$ had reached 0.6. Protein expression was induced by the addition of 0.5 mM IPTG and was then incubated at 18 °C with shaking at 250 rpm for
18 h. MS analysis of initial expression trials indicated that *in vivo* phosphopantetheinylation yielded significant but incomplete conversion to the holo-form. To drive phosphopantetheinylation to completion, an *in vitro* conversion to the holo-form was performed using cleared cell extract from each ACP over-expression. Briefly, cleared cell extract prepared from 300 mL cell culture was mixed with 40 μL of 50 mM CoA and was incubated at room temperature for 4 h then kept at 4 °C overnight. The proteins were then purified by Ni-NTA chromatography largely as previously described, except the batch purification method was used. MS analysis showed full conversion to the holo-ACP for every ACP.

For the production of apo-KirACP5, only plasmid encoding KirACP5 was transformed into *E. coli* BL21(DE3) but was otherwise expressed and purified as described above for the holo-ACP.

5.2.2 Expression of AT’s

AT’s (MCAT, DSZS and KirCII) was expressed as previously described.

5.2.3 Alanine scanning mutagenesis of KirACP5

Plasmid pET30-KirACP5 was used as the template for mutagenesis of KirACP5. Primers for alanine scanning mutagenesis of KirACP5 were designed with lengths between 25 and 30 bp, while melting temperatures were between 70 and 75 °C. Site-directed mutagenesis was performed following the QuikChange protocol (Agilent Technologies, Inc.). 2 min/kb was used as the extension rate. *E. coli* DH5α was used for plasmid amplification. Each mutant was confirmed by DNA sequencing. KirACP5 mutant was expressed and purified in the same way as WT KirACP5.
5.2.4 Site-directed mutagenesis and chimeragenesis of ACP’s

Plasmid pET30-KirACP10 was used as the template for mutagenesis of KirACP10. For mutagenesis of DEBS ACP6, plasmid pET28a-ACP6 was used. ACP amino acid substitutions were performed using the QuikChange protocol (Agilent Technologies, Inc.). Primers were designed based on similar principle as alanine scanning mutagenesis of KirACP5. 2 min/kb was used as the extension rate. Multiple-template-based sequential PCR was used for construction of ACP chimeras. All mutations and chimeras were confirmed by DNA sequencing. Primer sequences used for construction of KirACP10/ KirACP5 chimeras are summarized in Table 18.

<table>
<thead>
<tr>
<th>Table 18: Primer sequences for construction of KirACP10/KirACP5 chimeras.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer sequence</td>
</tr>
<tr>
<td>KirACP10-NA5-1F</td>
</tr>
<tr>
<td>KirACP10-NA5-1R</td>
</tr>
<tr>
<td>KirACP10-NA5-2F</td>
</tr>
<tr>
<td>KirACP10-NA5-2R</td>
</tr>
<tr>
<td>KirACP10-CA5-1F</td>
</tr>
<tr>
<td>KirACP10-CA5-1R</td>
</tr>
<tr>
<td>KirACP10-CA5-2F</td>
</tr>
<tr>
<td>KirACP10-CA5-2R</td>
</tr>
<tr>
<td>KirACP10-L0A5-1F</td>
</tr>
<tr>
<td>KirACP10-L0A5-1R</td>
</tr>
<tr>
<td>KirACP10-L0A5-2F</td>
</tr>
<tr>
<td>KirACP10-L0A5-2R</td>
</tr>
<tr>
<td>KirACP10-LIA5-1F</td>
</tr>
<tr>
<td>KirACP10-LIA5-1R</td>
</tr>
<tr>
<td>KirACP10-LIA5-2F</td>
</tr>
<tr>
<td>KirACP10-LIA5-2R</td>
</tr>
<tr>
<td>KirACP10-LIA5-3F</td>
</tr>
<tr>
<td>KirACP10-LIA5-3R</td>
</tr>
</tbody>
</table>
5.2.5 Site-directed mutagenesis and chimeragenesis of KirCII

Plasmid pET52-KirCII was used as the template for mutagenesis of KirCII. Primers for mutagenesis of KirCII were designed with lengths between 25 and 30 bp, while melting temperatures were between 70 and 75 °C. Site-directed mutagenesis was performed following the QuikChange protocol (Agilent Technologies, Inc.). 2 min/kb was used as the extension rate. KirCII/DSZS chimera (KirCII-sdDSZS) where the smaller subdomain of KirCII was replaced by that from DSZS, was prepared following the same procedure for KirACP10/KirACP5 chimeras using the following primers (Table 19). All mutations and chimeras were confirmed by DNA sequencing.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>KirCII-sdDSZS-1F</td>
<td>GATATACCATGGCAAGCTGAGCCACCC</td>
</tr>
<tr>
<td>KirCII-sdDSZS-1R</td>
<td>GATGAGCCGCGCCATCCCCGCCCCGCCCCGCCCCGCCCCGGCATC</td>
</tr>
<tr>
<td>KirCII-sdDSZS-2F</td>
<td>GGCGGGATGGCCGGG</td>
</tr>
<tr>
<td>KirCII-sdDSZS-2R</td>
<td>AGCTTCCTCGAGGACGACGGGGGCTGGGC</td>
</tr>
<tr>
<td>KirCII-sdDSZS-3F</td>
<td>GATATACCATGGCAAGCTGAGCCACCC</td>
</tr>
<tr>
<td>KirCII-sdDSZS-3R</td>
<td>GGCCTCTACGGCCGGAC</td>
</tr>
<tr>
<td>KirCII-sdDSZS-4F</td>
<td>GTCCCTGCGCGTGAGGCGCTCGACAGCCACTACGTCG</td>
</tr>
<tr>
<td>KirCII-sdDSZS-4R</td>
<td>AGCTTCCTCGAGTCCTGACCCGCGG</td>
</tr>
</tbody>
</table>

KirCII mutant and KirCII-sdDSZS was expressed and purified following the same procedure as KirCII.

5.2.6 Expression of DEBS holo-M6TE and DEBS holo-M6TE-AT°

DEBS apo-/holo-M6TE and apo-/holo-M6TE-AT° were expressed and purified as previously described. Holo-M6TE was prepared by using the K207 strain, which carries copies of sfp (phosphopantetheine transferase) gene for in vivo phosphopantetheinylation.
5.2.7 M6TE chimeragenesis and mutagenesis

M6TE chimera and mutant were prepared using the following primers. The M6TE-KirACP5 chimera was prepared following the same PCR procedure as KirACP10/KirACP5. M6TE-ACP6-F38L mutant was prepared following the QuikChange protocol (Agilent Technologies, Inc.), 3 min/kb was used as the extension rate. The chimera and mutation were confirmed by DNA sequencing, and were expressed and purified following the same procedure as WT M6TE.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>M6TE-KirACP5-1F</td>
<td>AGCTTCGGATCCCCGCGCGTTATGGGCGC</td>
</tr>
<tr>
<td>M6TE-KirACP5-1R</td>
<td>GCTGCAGGCTCGGGCTGAACAGCTGACTGCAACCC</td>
</tr>
<tr>
<td>M6TE-KirACP5-2F</td>
<td>GGGTGAGTACGCTGTTAGC CCGAGCGCAGCG</td>
</tr>
<tr>
<td>M6TE-KirACP5-2R</td>
<td>CGCTCGGGGCCTAGACCGCTCGGGGTGTTGGCC</td>
</tr>
<tr>
<td>M6TE-KirACP5-3F</td>
<td>CGCCACACCGAGCCG TCTAGTGACCCCCGCG</td>
</tr>
<tr>
<td>M6TE-KirACP5-3R</td>
<td>AGCTTCGATCCC CCGCTCCAGCCCCGCG</td>
</tr>
<tr>
<td>ACP6-F38L-FOR</td>
<td>CAGGACCAGCCGCGT G ACCGAGCTCGGCTTCC</td>
</tr>
<tr>
<td>ACP6-F38L-REV</td>
<td>GAAGCAGAGCTCGGTAGCCGGCTGCTTC</td>
</tr>
</tbody>
</table>

5.2.8 Circular dichroism

Circular dichroism (CD) measurements were performed with a JASCO 810 CD Spectropolarimeter. Samples for CD were buffer exchanged into CD buffer (10 mM potassium phosphate, 50 mM sodium sulfate, pH 7.4) and the concentration of samples was 0.2 mg/mL. Spectra from 190 nm to 260 nm were scanned at a step of 0.5 nm at 20 °C in a 0.1 cm cuvette, with ten repeats. The scan speed was 100 nm/min.
5.2.9 Chemo-enzymatic synthesis of azidoethylmalonyl-CoA

Chemo-enzymatic synthesis of azidoethylmalonyl-CoA was performed in 400 μL of reaction mixture containing 100 mM sodium phosphate (pH 7), MgCl₂ (2 mM), ATP (6 mM), coenzyme A (3 mM), azidoethyl malonate (15 mM), and mutant MatB₁⁰⁹ (20 μg) at 25 °C. Production of azidoethylmalonyl-CoA was confirmed by HPLC. Aliquots were removed after proper incubation time, quenched with an equal volume of ice-cold methanol, and centrifuged at 10,000g for 10 min and cleared supernatants were used for HPLC analysis on a Varian ProStar HPLC system. A series of linear gradients was developed from 0.1% TFA in water (A) to methanol (B) using the following protocol: 0-32 min, 20%-80% B; 32-35 min, 20%-100% A. The flow rate was 1 mL/min, and the absorbance was monitored at 254 nm using Pursuit XRs C18 column (250 mm × 4.6 mm, Varian Inc.). Chemo-enzymatic synthesis of other acyl-CoAs was performed in similar way using appropriate mutant MatB₁⁰⁹.

5.2.10 General procedure for KirCII assay

KirCII assay was performed in 10 μL of reaction mixture containing 50 mM Tris-HCl (pH 7.5), 50 mM MgCl₂, 300 μM azidoethylmalonyl CoA, 30 μM holo-KirACP5, and 3 μM KirCII at room temperature for 1 h.

5.2.11 General procedure for strain-promoted azide-alkyne cycloaddition assay

The strain-promoted azide-alkyne cycloaddition assay was performed in a total volume of 6 μL and contained 4 μL of KirCII assay mixture containing 20 μM KirACP5 and 2 μL of 600 μM DIBO. Reaction was incubated in dark at room temperature with gentle agitation for 1 h. Protein loading dye was then added and reactions were boiled for 5 min before analysis by SDS-PAGE. The gels were scanned using a Typhoon 7000 phosphoimager to determine the
intensity of DIBO labeled proteins bands. The bands were quantified by ImageQuant TL software (GE Life Sciences, Inc.) and rubber-band background subtraction was applied.

5.2.12 KirCII titration assay

KirCII titration assay was performed in 10 μL of reaction mixture containing 50 mM Tris-HCl (pH 7.5), 50 mM MgCl₂, 300 μM azidoethylmalonyl-CoA, 50 μM holo-KirACP5, and proper amount of KirCII at room temperature for 20 min. 4 μL of KirCII reaction mixture was then removed for strain-promoted azide-alkyne cycloaddition assay.

5.2.13 KirCII time course assay

KirCII time course assay was performed in 70 μL of reaction mixture containing 50 mM Tris-HCl (pH 7.5), 50 mM MgCl₂, 300 μM azidoethylmalonyl-CoA, 30 μM holo-KirACP5, and 1 μM KirCII at room temperature. Aliquots were removed at different time points and stored at -80 °C before used for strain-promoted azide-alkyne cycloaddition assay.

5.2.14 Homology model of KirACP5 and KirCII

Homology models for KirACP5 and KirCII were generated using the automated mode of SWISS-MODEL Workspace. The template used for KirACP5 was 1DV5, sequence identity between template and KirACP5 was 26%. The template for KirCII was 2QO3, sequence identity between template and KirCII was 34%.

5.2.15 M6TE assay

DEBS holo-Mod6TE reactions were performed following published protocol. 90 μg DEBS holo-Mod6TE was added to 35 μL of 50 mM Tris-HCl (pH 7) containing 2 mM MgCl₂, 5 mM diketide N-acetylcystamine, and 4 mM ethyl-CoA/azido-CoA (prepared by suitable
mutant MatB enzymes). Reactions were incubated overnight at room temperature and analyzed by RP-HPLC.

Reaction was quenched with an equal volume of ice-cold methanol, centrifuged at 10,000g for 10 min, and 25 μL used for HPLC analysis. HPLC analysis was performed on a Varian ProStar HPLC system. A series of linear gradients was developed from 0.1% TFA in water (A) to 0.1% TFA in acetonitrile (HPLC grade, B) using the following protocol: 0-40 min, 0-100% B; 40-42 min, 100% B; 42-43 min, 0-100% A; 43-45 min, 100% A. The flow rate was 1 mL/min, and the absorbance was monitored at 290 nm using Pursuit XRs C18 column (250 mm × 4.6 mm, Varian Inc.).

5.2.16 Trans-AT complementation of M6TE-AT°

Trans-AT complementing M6TE-AT° was performed as follows. 90 μg DEBS holo-Mod6TE-AT° was added to 35 μL of 50 mM Tris-HCl (pH 7) containing 2 mM MgCl₂, 5 mM diketide N-acetylcystamine, and 4 mM ethyl-CoA/azido-CoA (prepared by suitable mutant MatB enzymes), reaction was initiated by addition of 20 μg trans-AT. Reactions were incubated overnight at room temperature and analyzed by RP-HPLC using the same protocol as M6TE assay.

5.3 Results and discussion

5.3.1 ACP specificity of KirCII

Holo-ACP’s are required in the acyltransfer assay catalyzed by acyltransferases. Gratifyingly, only holo-KirACP5 was detected by MS after purification of KirACP5
coexpressed with Sfp (Figure 54). Apo-KirAC5 was expressed and purified for control assays (Figure 54).

![Fragmentor Voltage](image1)

**Figure 54: MS analysis of holo- and apo-KirACP5.** The expected masses for holo- and apo-KirACP5 are 17099.232 Da and 16758.89 Da, respectively. The experimental mass closely matched the expected mass for holo- and apo-KirACP5, respectively.

HPLC analysis of azidoethylmalonyl-CoA (AzEM-CoA) production showed full conversion after 7 h reaction (Figure 55). The reaction was typically completed after 3 h incubation. The MatB mutant used was T207S/M306I.
Figure 55: HPLC analysis of azidoethylmalonyl-CoA production. Aliquot of the reaction mixture was removed at 0 h and after 7 h. HPCL analysis showed full conversion after 7 h.

To validate that the KirCII-catalyzed acyl transfer coupled SPAAC assay was dependent on all reagents present, a set of control experiment was performed (Figure 56). The fluorescence intensity of the labeled ACP bands was quantified by in-gel fluorescence analysis. In the absence of KirCII, acyl-CoA, or cyclooctyne-fluorophore (DIBO), only background signal was detected. As expected, activity was not detected when apo-KirACP5 was used in place of holo-KirACP5. Furthermore, Sfp could be used to transfer the azido-functionalized phosphopantetheine moiety from azidoethylmalonyl-CoA to suitably provided apo-KirACP5, resulting in strong labeling with the cyclooctyne-fluorophore. The assay was analyzed by mass spectrometry (Table 21 and Figure 57). Azidoethylmalonyl-KirACP5 and DIBO-azidoethylmalonyl-KirACP5 mass were detected. To exclude the background labeling, a (-) KirCII control was performed for every ACP or ACP chimeras/mutants tested in the following assays and fluorescence intensity for (+) KirCII reactions was subtracted by that for the corresponding (-) KirCII controls.
Figure 56: Reagent control for KirCII assay. Results showed the fluorescence labeling of KirACP5 is dependent on the presence of all reagents.

Table 21: MS analysis of azido- and DIBO-KirACP5.

<table>
<thead>
<tr>
<th>ACP</th>
<th>Calculated mass (Da)</th>
<th>Experimental mass (Da)</th>
<th>Δ (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>azido-KirACP5/Sfp</td>
<td>17254.35</td>
<td>17255.34</td>
<td>-0.99</td>
</tr>
<tr>
<td>DIBO-azido-KirACP5/Sfp</td>
<td>18416.24</td>
<td>18416.40</td>
<td>-0.16</td>
</tr>
</tbody>
</table>

Figure 57: MS analysis of azido- and DIBO-KirACP5.
The assay was then biochemically characterized. A series of time course, titration and AT control experiments were performed. Time course analysis showed the fluorescence intensity was dependent on the KirCII reaction incubation time (Figure 58). Fluorescence intensity was in linear relationship with reaction time up to 2 hours.

![Time course analysis of KirCII assay](image)

**Figure 58: Time course analysis of KirCII assay.** Fluorescent labeling of KirACP5 was in linear relationship with reaction time. Fluorescence intensity was represented by arbitrary units. The assay was performed twice and experimental errors are shown as the standard deviation.

The fluorescence intensity was shown to be dependent on the KirCII concentration (Figure 59). Fluorescence intensity was in linear relationship with KirCII concentration in the range of 0.5 µM to 5 µM.
Figure 59: KirCII titration assay. [KirCII] was varied from 0.5 µM to 5 µM, fluorescence labeling of KirACP5 increased linearly as KirCII concentration increased. Fluorescence intensity was represented by arbitrary units. The assay was performed twice and experimental errors are shown as the standard deviation.

AT control was performed where KirCII was replaced by MCAT or DSZS in the assay (Figure 60). Malonyl-CoA-ACP transacylase (MCAT) is the acyltransferase in *Streptomyces coelicolor* A3(2) fatty acid biosynthesis. DSZS is the trans-AT from Disorazole Synthase in myxobacterium *Sorangium cellulosum* So ce12. MCAT and DSZS are known to load only malonyl-CoA onto holo-ACP and were not expected to utilize azidoethylmalonyl-CoA. Results showed only background fluorescence labeling when MCAT and DSZS were used in the AzEM-CoA transfer coupled click assay.
Figure 60: Azidoethylmalonyl-CoA transfer catalyzed by trans-AT’s. The transacylation was performed with KirACP5 as the substrate, fluorescence labeling was observed only for KirCII. Fluorescence intensity was represented by arbitrary units. The assay was performed twice and experimental errors are shown as the standard deviation.

Previous studies led to the conclusion that KirCII is responsible for installation of the C29-ethyl moiety in kirromycin. Given the presence of another trans-AT (KirCI) in the kirromycin biosynthetic gene cluster, whether KirCII displays specificity for KirACP5 was investigated. Accordingly, each kirromycin ACP was expressed in *E. coli*, purified, and converted to the holo-ACP by Sfp (Table 22).
Table 22: MS analysis of KirACP’s.

<table>
<thead>
<tr>
<th>ACP</th>
<th>Calculated mass (Da)</th>
<th>Experimental mass (Da)</th>
<th>Δ (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KirACP0/Sfp</td>
<td>17318.26</td>
<td>17319.33</td>
<td>-1.07</td>
</tr>
<tr>
<td>KirACP1/Sfp</td>
<td>18040.01</td>
<td>18041.00</td>
<td>-0.99</td>
</tr>
<tr>
<td>KirACP2/Sfp</td>
<td>18661.73</td>
<td>18662.72</td>
<td>-0.99</td>
</tr>
<tr>
<td>KirACP3/Sfp</td>
<td>17999.65</td>
<td>18000.33</td>
<td>-0.68</td>
</tr>
<tr>
<td>KirACP4/Sfp</td>
<td>18486.61</td>
<td>18487.68</td>
<td>-1.07</td>
</tr>
<tr>
<td>KirACP5/Sfp</td>
<td>17099.23</td>
<td>17100.27</td>
<td>-1.04</td>
</tr>
<tr>
<td>KirACP7/Sfp</td>
<td>17738.76</td>
<td>17739.67</td>
<td>-0.91</td>
</tr>
<tr>
<td>KirACP8/Sfp</td>
<td>18508.70</td>
<td>18509.45</td>
<td>-0.75</td>
</tr>
<tr>
<td>KirACP9/Sfp</td>
<td>16667.52</td>
<td>16668.46</td>
<td>-0.94</td>
</tr>
<tr>
<td>KirACP10/Sfp</td>
<td>16159.23</td>
<td>16160.31</td>
<td>-1.08</td>
</tr>
<tr>
<td>KirACP11/Sfp</td>
<td>18063.18</td>
<td>18064.29</td>
<td>-1.11</td>
</tr>
<tr>
<td>KirACP12/Sfp</td>
<td>17112.10</td>
<td>17112.80</td>
<td>-0.70</td>
</tr>
<tr>
<td>KirACP13/Sfp</td>
<td>17643.59</td>
<td>17644.52</td>
<td>-0.93</td>
</tr>
<tr>
<td>KirACP14/Sfp</td>
<td>17893.98</td>
<td>17894.71</td>
<td>-0.73</td>
</tr>
<tr>
<td>KirACP15/Sfp</td>
<td>16190.00</td>
<td>16190.66</td>
<td>-0.66</td>
</tr>
</tbody>
</table>

Each holo-ACP was tested as a substrate for KirCII-catalyzed trans-acylation using AzEM-CoA synthesized via an engineered MatB variant (Figure 61). In this assay, KirACP5 supported the highest activity with KirCII, although KirACP0, KirACP3, and KirACP7 supported basal activity at 15-25% of KirACP5. The remainder of the ACP’s tested displayed activities that were <10% of KirACP5. This data suggests that KirACP5 is the preferred interaction partner for KirCII, and other kirromycin ACP’s are poor substrates for KirCII. The ACP specificity of KirCII supports a role for programmed protein interactions in maintenance of kirromycin biosynthetic fidelity.

ACP’s from DEBS (DEBS ACP2) (type I PKS), E. coli fatty acid synthase (FAS) (AcpP) (type II FAS), actinorhodin biosynthesis (actACP65) (type II PKS) and Streptomyces FAS (ScFAS-ACP112) (type II FAS) were also examined as substrates for KirCII. All of these ACP’s displayed less than 10% transacylation activity compared to WT KirACP5. ACP
specificity of KirCII assay results agreed with an ACP phylogenetic analysis as shown in Figure 62. This observation further confirmed KirCII specificity towards its cognate ACP and suggested that the ACP specificity of KirCII would need to be overcome in order to be utilized in polyketide engineering.

Figure 61: ACP specificity of KirCII. KirCII assay was performed twice for each ACP, % transacylation activity was set to 100 for KirACP5, and the value for other ACP’s was calculated relative to KirACP5. Average % transacylation activity and standard deviation was plotted in the graph.
5.3.2 Mapping the KirCII:KirACP5 interaction epitope by alanine scanning mutagenesis

5.3.2.1 Homology modeling of KirACP5

A homology model for KirACP5 was constructed by Swiss-Model workspace\textsuperscript{110} automatic modelling mode in order to identify those residues predicted to be surface exposed (Figure 63A). The pdb file (1DV5) for D-alanyl carrier protein in D-Alanyl-lipoteichoic acid biosynthesis was selected by Swiss-Model as the template.\textsuperscript{113} Sequence identity between KirACP5 and the template is 26\% and the QMEAN Z-Score\textsuperscript{114} was -1.272, which is in the reliable range of modeling. The KirACP5 homology model represents a typical ACP structure, composed of a three helix bundle, connected by two long flexible loops. Notably, N- and C-terminal regions of KirACP5 weren’t included in the model, and secondary structure prediction
by PSIPRED\textsuperscript{115} predicted them to be flexible loop regions (\textbf{Figure 63B}). The secondary structure prediction agreed with the homology modeling results.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure63}
\caption{Homology model of KirACP5 (A) and secondary structure prediction of KirACP5 by PSIPRED (B). The secondary structures were labeled.}
\end{figure}

5.3.2.2 Design and preparation of KirACP5 alanine mutants

To map the epitope on KirACP5 involved in the KirCII:KirACP5 interaction, alanine scanning mutagenesis was adopted. A total of 61 non-alanine/glycine residues were then individually mutated to alanine, covering all the predicted solvent exposed regions in the KirACP5 homology model. The mutants were expressed and converted to the holo-form by Sfp-mediated phosphopantetheinylation, and purified to homogeneity. Notably, three of these mutants, D64A, L66A and I83A, could not be completely converted to the holo-form, while L52A, L58A, Y79A, F88A and T93A could not be expressed in soluble form. D64 and L66 are immediately adjacent to the presumed phosphopantetheinylation site (S65), and mutations at this position in other ACP’s are known to hinder modification by Sfp.\textsuperscript{116} The I83A mutant
is largely unfolded, as judged by analysis of the corresponding CD spectra. Thus, these eight mutants were omitted from further analysis. The CD spectra of a representative panel of the KirACP5 mutants matched very closely to that of WT KirACP5, indicating similar secondary structures (Figure 64).
Figure 64: CD spectra for representative KirACP5 mutants. The mutants generally showed similar CD spectra curves to that of wild-type KirACP5, indicating there was no structural perturbation.
5.3.2.3 KirCII assay

Each KirACP5 mutant was tested as a substrate for KirCII using azidoethylmalonyl-CoA and the strain-promoted cycloaddition assay (Figure 6A). Most of the mutants displayed similar activity to wild-type KirACP5, indicating that mutation to alanine at these positions does not significantly impact interaction with KirCII. However, alanine substitution at five positions (His44, Leu45, Arg51, Asp60, and Arg70) resulted in mutants that displayed activities <20% than that of the wild-type KirACP5 with KirCII. The locations of these five key positions were mapped onto the KirACP5 homology model (Figure 6B), which for the first time begins to define at least part of the KirACP5:KirCII interaction epitope. In addition to those residues identified from alanine scanning mutagenesis, the well-known and highly conserved ‘DSL motif’ (KirACP5 residues Asp64, Ser65, and Leu66) (see Table 23 for a list of carrier protein residues involved in various protein interactions) can be likely included in this epitope, given the phosphopantetheinylation site is embedded within this motif, and previous experimental evidence.116 Interestingly, this result contrasts with alanine scanning mutagenesis of the malonyl-CoA specific, yet ACP promiscuous trans-AT, DSZS,117 which revealed only two positions that were sensitive to substitution. From summary in Table 23, equivalent positions for His44, Asp60 and Arg70 on ACP’s are extensively involved in interactions with a variety of proteins, while no equivalent positions for Leu45 and Arg51 were involved in any interactions, suggesting both a conserved and unique feature of KirACP5:KirCII interaction.
Figure 65: Mapping the KirCII:ACP interaction epitope by alanine scanning mutagenesis. (A) Trans-acylation rates of holo-KirACP5 alanine mutants with KirCII. Rates are expressed as a percentage of the activity with wild-type KirACP5. Wild-type positions that were alanine were not mutated. Mutants that displayed <20% the activity of the wild-type holo-KirACP5 are highlighted red. (B) Trans-acylation activities of KirACP5 alanine mutants mapped onto an KirACP5 homology model ribbon diagram (left) and its computed surface (right). Red, <20% activity; green, 20-80% activity; yellow, >80% activity; grey, not determined.
Table 23: Identity of ACP residues involved in protein interactions with various domains. Only those residues important for interactions are shown. Key KirACP5 residues discussed in the text are color coded as follows: Orange, key electrostatic interactions from KirACP5 alanine scanning and docking studies; Red, positions of mutations in the HII’ region (discussed in text); Green, phosphopantetheinylation site.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D28</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P29</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E30</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S31</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A32</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V33</td>
<td>L</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R34</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D35</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A42</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A43</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H44</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L45</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A48</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R51</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L52</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P53</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P54</td>
<td>R</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D55</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R56</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V57</td>
<td>A</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L58</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S59</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D60</td>
<td>D</td>
<td>E</td>
<td>D</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V61</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G62</td>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V63</td>
<td>L</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D64</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S65</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L66</td>
<td>L</td>
<td>I</td>
<td>V</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G67</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R69</td>
<td>A</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R70</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 Continued.

<table>
<thead>
<tr>
<th>L71</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>S72</td>
<td>R</td>
</tr>
<tr>
<td>R73</td>
<td>R M M</td>
</tr>
<tr>
<td>R74</td>
<td>L A</td>
</tr>
<tr>
<td>V79</td>
<td></td>
</tr>
<tr>
<td>D82</td>
<td>R R</td>
</tr>
<tr>
<td>D83</td>
<td>I I</td>
</tr>
<tr>
<td>P84</td>
<td>E I</td>
</tr>
<tr>
<td>A85</td>
<td>A D</td>
</tr>
<tr>
<td>R86</td>
<td>T D</td>
</tr>
<tr>
<td>F88</td>
<td></td>
</tr>
<tr>
<td>G89</td>
<td>F Y</td>
</tr>
<tr>
<td>V90</td>
<td>D D T</td>
</tr>
<tr>
<td>Q92</td>
<td>D</td>
</tr>
<tr>
<td>T93</td>
<td></td>
</tr>
<tr>
<td>A96</td>
<td>Q</td>
</tr>
<tr>
<td>R97</td>
<td>R</td>
</tr>
<tr>
<td>A100</td>
<td>W</td>
</tr>
<tr>
<td>D103</td>
<td>T</td>
</tr>
<tr>
<td>K104</td>
<td></td>
</tr>
</tbody>
</table>

107
5.3.3 Computational docking

5.3.3.1 Homology modeling of KirCII

A homology model for KirCII was constructed by Swiss-Model workspace automatic modelling mode (Figure 66). Interestingly, the natively cis-acting KS3AT3 didomain from DEBS (34% sequence identity to KirCII) was selected by Swiss-Model as the template (pdb: 2QO3). Although KirCII functionally resembles the trans-AT’s DSZS and MCAT, the crystal structures of which are known, these two structures were not selected by Swiss-Model, likely due to poorer sequence identity with KirCII (27% and 22% sequence identity to KirCII, respectively). This is in agreement with phylogenetic analysis, which showed that KirCII is more closely related to cis-AT’s than to malonate-specific trans-AT’s (Figure 67). In addition to the large and small AT subdomains typical of trans-AT’s, the modeled KirCII structure also includes a KS-AT linker domain that is usually associated with cis-acting AT domains from type I PKS’s. This prediction is in agreement with the sequence length of KirCII, which is longer than that of DSZS and MCAT. Sequence alignment and secondary structure prediction also supported the presence of a KS-AT linker in KirCII (Figure 68 and 69). The KS-AT linker domain was included in our subsequent protein interaction studies, given the possibility that the linker region participates in ACP recognition for cis-acting AT domains. Furthermore, the KS-AT linker domain appears important for solubility of KirCII, given a truncated version of KirCII that lacks the linker could not be expressed in soluble form (data not shown).
**Figure 66: Homology model for KirCII.** The substructures on KirCII are color-coded.

**Figure 67: Phylogenetic analysis of AT’s.** AT’s from rapamycin (Rap-AT), erythronolide (DEBS-AT), disorazole (DSZS), kirromycin (KirCI, KirCII, KirAV1-AT1, KirAV1-AT2), *E. coli* fatty acid (FabD), and *S. coelicolor* fatty acid (MCAT) biosynthesis were examined. Tans-AT’s were highlighted with red boxes.
Figure 68: Amino acid alignment of the KirCII with DEBS AT’s. Sequence colored in green corresponds to KirCII’s KS-AT linker and sequences colored in yellow correspond to the DEBS AT’s KS-AT linker.
KirCII

CRMV6 TVASHSHVVEELDDLALR-P-LS-PGRYPYFSTVAAPVP-TDLGPAWYM 333
AT1
AKRLAVDYASHSSHSVETRDLALHAEL6EDPHPGLGFPYFSV7GRTQFDELDA6GYWR 348
AT2
ARAIPVDYASHTAHVEPVREDELVQALG-ITPRRAMEVPFSTL7G6DLG7TELDA6GYWR 348
AT3
VRDIDVYASHSPQIERVREELLETG6-1APPPAPARVTPHSTVESRSMDGT6LDASYWR 345
AT4
PRRAVRHSAPEVARIDRLLAEAG6-ITAVG6VPLHSTV76EVIDTSAMDSAYWR 342
AT5
ARRIPVDYASHSPQVESLRREELTELG6-TSPVSADVLYSTTT6QPDTATMD7AYWA 338
AT6
AKTDYASHSRHVEIRRTLADLG-1SARRAAATPVLH6ERRGADMGFPRWYD 345

KirCII

NLKVPVRILLAAAT6RLAEDEGHEIFVEVSTHPVLLSSLRLQIES6ARLPGEVLP6GFRR 393
AT1
NLKTVRFAD4VRALLAEQ6YSYRTLEVSAAHPLTAAIEEIGD6CA--------------- 393
AT2
NLKHFVEFSAVQALID6G6ATIFEVSPHPVLSAVQETLDIA--------------------- 391
AT3
NLKTVRFAD4VRALLAES6G6DAFIEVSPHPVYQAVEE6AAADG-------------- 390
AT4
NLKDFVLFQAVG6L6VEQGFDTFVEVSPHPVILLA6VEETAEH6A--------------- 387
AT5
NLKQEVRFQDATRQALEAG6FDAFVEVSPHPVLT6IEALDSALP---------------- 383
AT6
NLQVSQVRDEAVSAVADVG6HATFVEEMSHPHVT6AVQQEIA------------------ 385

KirCII

AMLSSL6TLFTY6RDHWP7SAPPAPALTLYQ4AVLA6RRPRPSPAAG6SG 445
AT1
---------------------------DLS---AIHLSMR6D6------------------- 466
AT2
-----------------------------ESDAAVLGTLRDAD------------------ 467
AT3
-----------------------------AED-AVV6SLHRD6D---------------- 496
AT4
-----------------------------EVY---CVPTRREE------------------ 398
AT5
-----------------------------A6AVG6TLRRDR6------------------ 399
AT6
-----------------------------AVIALG6SLRDTAE------------------ 400

:::::*
Figure 69: Secondary structure prediction for KS-AT linker in KirCII and DEBS AT3.
(A) Secondary structure prediction for KirCII linker using PSIPRED. (B) The predicted secondary structure was mapped onto the KS-AT linker structure on KirCII homology model. (C) Secondary structure prediction for the known AT3 linker of DEBS using PSIPRED. (D) The predicted secondary structure was mapped onto the KS-AT linker structure (pdb: 2QO3), which showed the prediction to be quite accurate.
5.3.3.2 Computational docking of KirACP5 onto KirCII

The PatchDock\textsuperscript{125} and the ClusPro\textsuperscript{126} servers were then used for molecular docking simulations using the KirACP5 and KirCII models. Distance constraints were imposed on Patchdock where distance between the Ser65 side chain oxygen on KirACP5 and Ser203 side chain oxygen on KirCII was set between 15 to 25 Å. No constraints were imposed on ClusPro docking.

Gratifyingly, both servers converged on almost identical docking results. One of the top models was chosen on the basis of the alanine scanning mutagenesis data described above. Overall, the docking model resembles that of the iterative type I PKS DynE8 AT and ACP,\textsuperscript{72} and also the AT3-ACP3 model from DEBS.\textsuperscript{91} KirACP5 is predicted to sit in a large cleft formed by parts of the KS-AT linker domain, as well as the small and large subdomains. Gratifyingly, this model places the ACP phosphopantetheine attachment site (S65) ~20 Å away from the presumed active site catalytic Ser (S203) of KirCII, and an active site tunnel can be identified that could accommodate the requisite ACP phosphopantetheine.

This model positions each of the five residues identified from alanine scanning mutagenesis (H44, L45, R51, D60, and R70) intimately at the interface between the ACP and KirCII. For example, the carboxylate side chain of D60 of KirACP5 is predicted to make a salt bridge with the terminus of the R410 side chain of KirCII (Figure 70). Interestingly, docking models of the DynE8 AT:ACP,\textsuperscript{72} DEBS AT4:ACP4,\textsuperscript{124} and fungal non-reducing PksA:ACP complexes,\textsuperscript{127} also implicate an Asp equivalent to D60 of KirACP5 at the interface (Table 23). Notably, D60, along with the nearby D64, appear to form a negatively charged group of residues in loop-I (L\textsubscript{1}) of KirACP5 that interacts with three consecutive Arg residues from
KirCII (R408-R410). Positively charged residues of other AT’s are often involved in ACP interactions with an Asp residue analogous to KirACP5 D64.\textsuperscript{69,72,123-124} In addition, KirACP5 R51 and R70, both judged as important for interaction with KirCII by alanine scanning mutagenesis, are positioned to interact with D98 and E395 of KirCII, respectively. Further, the main chain nitrogen of KirACP5 A48 is positioned for hydrogen bonding with the hydroxyl of KirCII Y426. Another key predicted interaction involves E245 and R69 from KirCII and KirACP5, respectively (\textbf{Figure 70}).
Figure 70: Docking model of the KirCII:KirACP5 interaction. Overview (A) and detailed analysis (B) of the docking model. KirCII large subdomain is shown in light orange, the small subdomain of KirCII is shown in light pink, KirCII linker domain is green, and KirACP5 is shown in teal. The H$_{11}$’ portion of L$_4$ of KirACP5 is highlighted red. Phosphopantetheinylation site of KirACP5 and the KirCII catalytic Ser are shown as spheres. For clarity, not every highlighted residue is labeled.
5.3.3.3 Mutant KirCII assay

To test the accuracy of this model, a series of KirCII mutants were constructed and assayed by the cycloaddition assay using KirACP5 as substrate. The secondary structure of each mutant matched closely to that of the wild-type KirCII, at least as judged by CD spectroscopy (Figure 71).

![CD spectra for KirCII mutants](image)

**Figure 71: CD spectra for KirCII mutants.** The mutants generally showed similar CD spectra curves to that of wild-type KirCII, indicating there was no structural perturbation.

KirCII assay was performed with each KirCII mutant and results were summarized in Figure 72. The KirCII mutants R408A, R409A, and R410A supported trans-acylation at slightly lower rates compared to WT KirCII with KirACP5 as substrate. Moreover, substitution of each Arg with Glu further reduced activity of each mutant. In addition, a triple mutant where
all three Arg residues (R408-R410) were mutated to Ala displayed less than 10% trans-acylation activity, compared to WT KirCII, while no activity was detected when all three arginines were mutated to Glu. These results suggest that KirCII R408-R410 plays a significant role in the interaction with KirACP5. Similarly, mutation of D98 of KirCII to Ala reduced activity to just 30% of the WT KirCII activity. Mutation of E245 to Ala reduced activity to 60% in comparison to WT KirCII, while activity could not be detected when the same residue was substituted with Arg. Substitution of KirCII E395 with Ala resulted in loss of activity, while substitution with Arg completely abolished activity. Similarly, the predicted importance of Y426 of KirCII was confirmed by mutagenesis. This data is consistent with the electrostatic contributions of these residues predicted on the basis of the docking model, and suggests that this set of residues are important for controlling interactions between KirCII and KirACP5. Finally, two surface residues more distant from the predicted interface were also targeted. Substitution of D123 with Ala resulted in a slight increase in activity with KirACP5 as substrate, while substitution of the same residue with Arg resulted in ~50% trans-acylation activity, compared to WT KirCII. Meanwhile, substitution of E170 with Ala resulted in small decrease in activity with KirACP5 as substrate, and substitution of E170 with Arg resulted in ~50% trans-acylation activity, compared to WT KirCII. Thus, substitution of negatively charged residues distal from the predicted interface with a neutral amino acid has a small effect compared to equivalent substitutions at predicted interface residues. Similarly, substitution at these distal sites with oppositely charged residues results in only ~2-fold decreased activity, compared to wild-type KirCII, whereas such substitutions at residues predicted to be involved in the KirCII epitope (e.g. R408E, R409E, R410E and E245R) lead to dramatic decreases in
activity, compared to wild-type KirCII. Cumulatively, this data suggests that several electrostatic interactions contribute to molecular recognition between KirCII and its cognate ACP.

**Figure 72: KirCII mutant assay summary.** The data is the average % transacylation, the error bars are the standard deviation (n=2).

Electrostatic surface potential maps of KirCII and KirACP5 were calculated using the PDB2PQR server. Electrostatic surface maps of KirCII and KirACP5 highlight the charge complementarity between several regions of the predicted interface (**Figure 73**).
Figure 73: Electrostatic surface potential maps of KirCII and KirACP5. Colors range from blue (positive) to white (neutral) to red (negative). Four key electrostatic contacts are shown boxed, the KirCII contacts are A (D98), B (R410) C (R409), and D (E245), while those of KirACP5 are A’ (R51), B’ (D60), C’ (D64), and D’ (R69). The asterisk indicates the position of the KirCII active site Ser and the KirACP5 phosphopantetheinylation site. The surfaces are represented so that if KirACP5 is rotated 180°, the letters from each surface would match.

5.3.4 Engineering non-cognate ACP’s to be recognized by KirCII

5.3.4.1 KirACP10 chimera

The docking model and mutagenesis data described above suggests that a portion of L1 and several C-terminal residues of helix-I (H1) and N-terminal residues of helix-II (HII) are involved in the KirCII interaction epitope, while loop-II (LII) and helix-III (HIII) of KirACP5 do not make significant contributions to the interaction with KirCII. To test this hypothesis, and to further define the ACP sequence elements required for ACP specificity of KirCII, chimeras between KirACP5 and a non-cognate ACP from kirromycin biosynthesis were
constructed (Figures 74). KirACP10 was selected to contribute to these chimeras, since this ACP shares the highest amino acid sequence identity (35%) to KirACP5 among the kirromycin ACP’s.

Notably, the entire N-terminal half (L₀-H₁-L₁) of KirACP5 is sufficient to provide robust recognition by KirCII when fused with the C-terminal portion (H₂-H₂-H₃-L₃) of KirACP10. The corresponding chimera, KirACP10-NA5, supported 50% activity compared to the WT KirACP5, a 4-fold improvement in efficiency compared to the non-cognate ACP (Figure 74). In contrast, the reverse chimera that includes the N-terminal portion of KirACP10 and the C-terminal portion of KirACP5 (KirACP10-CA5) displays no detectable activity with KirCII. This data suggests that the majority of the KirACP5:KirCII recognition features are located in the N-terminal L₀-H₁-L₁ portion of KirACP5, in agreement with the docking model and alanine scanning mutagenesis results.

In an attempt to further refine the putative interaction epitope, a series of chimeras was designed that exchanged individual secondary structure elements. Exchange of KirACP10 H₁ or L₁ with the corresponding structural element from KirACP5 (KirACP10-HIA5 and KirACP10-LIA5) each led to a 2-fold improvement in activity, compared to the WT KirACP10, and could not therefore recapitulate the full activity of the KirACP10-NA5 chimera that contained the entire N-terminal portion of KirACP5. Interestingly, preceding H₁ of each kirromycin ACP is a short sequence labeled L₀ that is predicted to form a random coil. As judged by sequence homology, this sequence element is not equivalent to the docking domains postulated to play some role in protein interactions between other trans-AT’s and ACP’s (results not shown). Substitution of the KirACP10 L₀ region with that from KirACP5
(KirACP10-L0A5) failed to significantly alter ACP specificity, while deletion of L₀ from KirACP5 resulted in an insoluble protein (results not shown).

**Figure 74: KirCII assay for KirACP10/KirACP5 chimeras.** Design (A) and acyl transfer assay (B) of KirACP10 chimeras. The data is the average % transacylation, the error bars are the standard deviation (n=2).
5.3.4.2 KirACP mutants

Interestingly, many residues within H, L, and H of KirACP5/KirACP10 are poorly conserved among kirromycin ACP’s, as judged by alignment of their amino acid sequences (Figure 75 and 76). Moreover, five of these poorly conserved positions correspond to residues that were previously highlighted by alanine scanning mutagenesis of KirACP5 (H44, L45, R51, D60, and R70). Accordingly, we hypothesized that some of these poorly conserved residues, perhaps including some from the alanine scanning mutagenesis set, might dictate the ACP specificity of KirCII. To test the contribution of these poorly conserved positions in controlling specificity, a panel of 13 mutants was constructed that substituted each selected residue within H, L, and H of KirACP10 with the corresponding residue from KirACP5 (Figure 75).

With respect to positions highlighted by alanine scanning mutagenesis, substitution at I37 and E53 in KirACP10 (equivalent to H44 and D60 in KirACP5) provided only small changes to the transacylation activity, compared to wild-type KirACP10 (see I37H and E53D). In contrast, the KirACP10 mutants A44R and N63R (corresponding to positions R51 and R70 in KirACP5) each display approximately 2-fold higher activity with KirCII, compared to wild-type KirACP10, underscoring the importance of these two previously highlighted positions for recognition by KirCII. Interestingly, substitution of KirACP10 Q62 with Arg (equivalent to R69 in KirACP5) failed to improve activity with KirCII, compared to WT KirACP10, even though this residue is adjacent to the equivalent of R70 in KirACP5. This result is entirely consistent with the alanine scanning mutagenesis data, which shows that of the R69/R70 pair in KirACP5, only substitution at R70 is detrimental to activity with KirCII. Presumably, R70
in KirACP5 contributes more to the interaction with KirCII than the neighboring R69, even though R69 is predicted to interact with KirCII via E245 (Figures 70). Perhaps mutation to Ala at R69 can be better compensated by other side chain rearrangements than Ala substitution at R70. Notably, aside from KirACP5, Arg is never found at residue 69 among the kirromycin ACP’s, and this residue is most frequently Leu or Val. Similarly, at residue 70, Arg is found only three times among all 15 kirromycin ACP sequences, and is most often the oppositely charged Glu.

The most active KirACP10 single amino acid exchanges were those at positions F51 and Y54 (equivalent to L58 and V61 in KirACP5), and do not correspond to positions highlighted by our alanine scanning mutagenesis. Substitution at these two positions with the equivalent residue from KirACP5, Leu and Val, respectively, supported activity 5-fold higher than WT KirACP10. Consistent with this data, mutagenesis of the KirACP5 residue corresponding to KirACP10 Y54 resulted in a mutant that showed only 21% the activity of WT KirACP5 (KirACP5-V61A, Figure 65). Intriguingly, the KirACP5 residue equivalent to KirACP10 F51 was not expressed in soluble form when mutated to Ala (KirACP5-L58A) as part of the previous surface mutagenesis, and was therefore not highlighted by the surface mutagenesis data. Interestingly, KirACP5 L58 and V61 are not well conserved among the kirromycin ACP’s (Figure 76). For example, residue 58 is most often Phe (as in KirACP10 F51), and is Leu in only four of the kirromycin ACP sequences. Residue 61 is usually Leu or Tyr (as in KirACP10 Y54) among the kirromycin ACP’s and is Val only in KirACP5. Combination of both F51L and Y54V in KirACP10 actually resulted in a slight decrease in trans-acylation activity compared to either the F51L or Y54V mutant. However, exchange of
a portion of a L₁ sequence in KirACP10 that included F51 and Y54 with that from KirACP5 (H₁’, corresponding to V57-D60 in KirACP5), resulted in a chimera with 5-fold the activity of wild-type KirACP10 (KirACP10-H*, corresponding to R56-V61 in KirACP5, Figure 75B). Thus, the precise identities of residues within this short helix are critical for maintaining maximal activity with KirCII. Together, this data suggests that in KirACP5, Arg residues at positions 51/70, and Leu/Val at residues 58/61, respectively, are required but not totally sufficient for recognition by KirCII.
Figure 75: Probing the ACP:KirCII interaction epitope by mutagenesis of a non-cognate ACP. (A) Sequence logo\textsuperscript{30} for the kirromycin ACP’s. Shown in grey and orange, respectively, are the amino acids at selected positions of KirACP5 and KirACP10. Asterisk indicates the phosphopantetheinylation site. Positions discussed in the text are highlighted. Boundaries between each secondary structure element are also shown. For brevity, N- and C-terminal loop regions (L\textsubscript{0} and L\textsubscript{III}) are not shown. L\textsubscript{0} is the region preceding H\textsubscript{I} and L\textsubscript{III} is the region after H\textsubscript{III}. (B) Scheme showing contribution of KirACP5 and KirACP10 to each chimera or mutant (left). Activities of WT, chimeric, and mutant kirromycin ACP’s with KirCII (right). Rates are expressed as a percentage of the activity with WT holo-KirACP5. The data is the average % transacylation, the error bars are the standard deviation (\(n=2\)).
Figure 76: Amino acid alignment of the ACP’s.
5.3.4.3 DEBS ACP6 mutants

Some trans-AT’s have been shown to be promiscuous towards the ACP substrate, and can utilize ACP’s from different biosynthetic pathways. The conclusion that Arg is required at positions 51/70 of KirACP5, and that a Leu and Val are required at positions 58 and 61, led us to predict whether ACP’s from other biosynthetic systems would serve as substrates for KirCII. For example, ACP6 from DEBS includes Ala and Gly at positions equivalent to 51/70 of KirACP5, while at residues equivalent to KirACP5 58 and 61, Phe and Leu are present in the DEBS ACP6, respectively. Accordingly, we predicted that DEBS ACP6 would not serve as an efficient substrate for KirCII. This hypothesis was tested by carrying out the cycloaddition assay using AzEM-CoA, KirCII, and holo-DEBS ACP6. As predicted, WT holo-DEBS ACP6 was not a detectable substrate for KirCII (Figure 77).

Given that the KirACP10 mutation F51L led to improved trans-acylation activity with KirCII, compared to the wild-type KirACP10, and that the equivalent residue in DEBS ACP6 is also Phe (F38), we proposed that introduction of leucine at position 38 of DEBS ACP6 would improve the transacylation efficiency of this non-cognate ACP with KirCII. To test this notion, the mutant DEBS ACP6-F38L was constructed, converted to the holo-form and assayed with KirCII. Remarkably, and in complete contrast to wild-type DEBS ACP6, the F38L mutant was a detectable substrate for KirCII, and this substitution improved trans-acylation activity at least 14-fold compared to WT DEBS ACP6 (Figure 77). To the best of our knowledge, this is the first time that the activity of a trans-AT has been rationally engineered towards a non-cognate carrier protein.
5.3.5 KirCII chimeragenesis

DSZS is a trans-AT from the disorazole synthase gene cluster from myxobacterium Sorangium cellulosum So ce12.\(^{111,48}\) Besides its cognate ACP’s, DSZS has been shown to catalyze acyl transfer onto DEBS ACP’s, albeit less efficiently.\(^{91}\) However, DSZS and MCAT possess strict specificity towards malonyl-CoA and the specificity may be difficult to change even through significant enzyme engineering.\(^{90}\) Computational docking studies have suggested that the smaller subdomain of trans-AT’s such as DSZS and MCAT is part of the docking site for ACP.\(^{48,96}\) If this is accurate, it might be possible to integrate the extender unit promiscuity of KirCII and the ACP promiscuity of DSZS, by constructing a KirCII-DSZS chimera whereby the smaller subdomain of KirCII was replaced with that from DSZS (Figure 78).

**Figure 77: Activities of WT and mutant DEBS ACP6 with KirCII.** KirCII assay for each ACP was performed twice. Transacylation activity for KirACP5 was plotted to be 100%, average % transacylation for WT and mutant ACP6 was calculated relative to WT KirACP5.
Figure 78: Chimeragenesis of KirCII/DSZS. KirCII structure is based on the homology model, pdb file for DSZS is 3RGI.

The corresponding fusion protein, KirCII-sdDSZS, was constructed via overlap extension PCR, expressed in *E. coli* and purified to homogeneity. The fusion protein then was tested for transacylation activity with a panel of ACP’s. Unfortunately, no activity was detected for any of the tested ACP’s, even for KirACP5 (Figure 79). There are two potential explanations contributing to this failure. The first is that the smaller subdomain on MCAT or DSZS is not the genuine/only ACP docking site. The second is that KirCII isn’t amenable to chimeragenesis, which might bring significant structural changes. Further investigation is required.
5.3.6 KirCII complementation of DEBS M6TE

MCAT has been shown to be capable of complementing holo-M6TE-AT° where the cis-AT6 was inactivated by mutating the catalytic Ser into Ala, with malonyl-CoA as extender unit. With this in mind, it was hoped that KirCII could be used to complement M6TE in which the ACP6 included the critical mutation, F38L, described above. As a positive control, the ability of M6TE to catalyze the formation of triketide lactone formation in the presence of diketide-SNAC and ethylmalonyl-CoA was first established (Figure 80). In addition, HPLC analysis of a reaction mixture that included MCAT, holo-M6TE-AT°, and malonyl-CoA resulted in the detection of the expected triketide lactone product (Figure 81), illustrating that the AT-null mutant of M6TE is competent to turnover malonyl-CoA when transacylated by MCAT.

Figure 79: KirCII-sdDSZS catalyzed transacylation onto ACP's. Each ACP was assayed by KirCII-sdDSZS and KirCII, the latter was performed as positive control. The results were plotted relative to KirCII transacylation onto KirACP5.
Figure 80: holo-M6TE catalyzed triketide lactone formation. Ethylmalonyl-CoA was used as the extender unit. Holo-M6TE-AT° catalyzed triketide lactone formation was performed as negative control.

Figure 81: MCAT complementation of holo-M6TE-AT° for triketide lactone production. Malonyl-CoA was used as the extender unit.

Previously, ACP6-F38L mutant displayed better transacylation activity compared to WT ACP6 in the KirCII assay. To investigate if this mutation could be utilized to diversify
polyketide structures, F38L was introduced into M6TE-AT° via site-directed mutagenesis. Mutant identity was confirmed by DNA sequencing. Triketide lactone product was detected when apo-M6TE-AT°-ACP6-F38L was acylated by Sfp to provide the ethylmalonyl-CoA extender unit, suggesting that the ACP6-F38L mutation didn’t disrupt the enzymatic activity of M6TE, or interaction with Sfp (Figure 82). However, no product was detected when KirCII was used in conjunction with holo-M6TE-AT°-ACP6-F38L. Inaccessibility of ACP6 or insufficient ACP6-F38L:KirCII interaction might be the causes, given that KirCII was much larger than MCAT, and although much better than WT DEBS ACP6, ACP6-F38L displayed about 15% transacylation activity compared to WT KirACP5, which may not be sufficient for efficient transacylation to occur.

![M6TE-AT°-F38L_Sfp_ethylmalonyl CoA](image)

**Figure 82: Triketide lactone production by Sfp acylated apo-M6TE-AT°-ACP6-F38L.** Sfp catalyzed the ethylmalonyl group transfer onto apo-M6TE-AT°-ACP6-F38L, which upon incubation with diketide-SNAc, produced the expected triketide lactone product.
To address the poor KirCII:ACP6-F38L interaction problem, M6TE-AT°-KirACP5 chimera was designed, where ACP6 in M6TE-AT° was replaced by KirACP5. Attempted complementation of holo-M6TE-AT°-KirACP5 with KirCII was performed using ethylmalonyl-CoA (Figure 83). However, no triketide lactone product was detected by HPLC analysis of the reaction mixture. Poor KS6:KirACP5 interaction was probably responsible for the failure, as it has been reported that specific motifs are required for intramodular KS:ACP interaction.\textsuperscript{24} A delicate balance between KS:ACP and AT:ACP interaction is highlighted by these results and minor perturbation to M6TE architecture is likely required in order to maintain the original cis-interactions. Cumulatively, in-depth understanding and engineering of cis-KS:ACP and trans-AT:ACP interactions are required for successful trans-AT complemention.

![Figure 83: KirCII complementing holo-M6TE-AT°-KirACP5 chimera. Ethylmalonyl-CoA was used as the extender unit.](image-url)
5.4 Conclusion

The ability of KirCII to utilize substrates modified with chemical handles for click chemistry was leveraged to provide a rapid and convenient assay for measuring the activity of KirCII. Accordingly, the activity of KirCII towards each kirromycin ACP was measured, providing evidence that KirCII interacts specifically with the expected cognate ACP. A combination of ACP surface mutagenesis, docking simulations, KirCII mutagenesis, and substitutions of amino acids and structural elements, led to the identification of several key electrostatic interactions that define the KirACP5:KirCII interaction epitope. Remarkably, this insight enabled engineering a completely non-cognate ACP into a robust substrate for KirCII via the introduction of a single amino acid substitution. The data presented here combined with previous studies highlight the importance of small helical turns located within loop regions that connect the larger helices that form the basis of all ACP structures. Dissapointingly, this current knowledge proved insufficient to successfully complement AT-null M6TE with KirCII, revealing that additional intrapolyptide protein interactions likely need to be overcome. Nevertheless, in the long term, an improved understanding of the features that allow recognition of an ACP by trans-AT’s can be used as a foundation for designing strategies that allow the diversification of polyketides.
CHAPTER 6.  Probing the cis-FluAT1:ACP interaction

6.1 Introduction

Fluvirucin B1 is a 14-membered macrolactam produced by *Actinomadura vulgaris* with moderate to good antifungal and antiviral activities.\textsuperscript{131,132} Fluvirucin B1 synthase is a type I PKS, composed of five modules, each harboring a set of catalytic domains responsible for the elongation and tailoring of two-carbon units (Figure 84). The final polyketide product is cleaved and cyclized by a thioesterase domain.

![Diagram of Fluvirucin B1 biosynthesis](image)

**Figure 84: Fluvirucin B1 biosynthesis.**

Fluvirucin AT’s were predicted to exhibit various acyl-CoA specificities (Table 24). FluAT1 and FluAT5 are predicted to utilize ethylmalonyl-CoA as the extender unit, FluAT3

136
uses methylmalonyl-CoA, while FluAT2 and FluAT4 utilize malonyl-CoA. Specificity for FluAT1, FluAT3 and FluAT5 has been experimentally verified.\textsuperscript{132}

### Table 24: Predicted AT specificity for Fluvirucin B1.

<table>
<thead>
<tr>
<th>Module</th>
<th>Predicted AT specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1</td>
<td>Ethylmalonyl-CoA</td>
</tr>
<tr>
<td>Module 2</td>
<td>Malonyl-CoA</td>
</tr>
<tr>
<td>Module 3</td>
<td>Methylmalonyl-CoA</td>
</tr>
<tr>
<td>Module 4</td>
<td>Malonyl-CoA</td>
</tr>
<tr>
<td>Module 5</td>
<td>Ethylmalonyl-CoA</td>
</tr>
</tbody>
</table>

Research from our lab (work by Irina Koryakina) has revealed that the didomain fragment FluKS1AT1 is capable of transferring a panel of extender units besides ethylmalonyl-CoA onto FluACP1 with varying degree of efficiency (Figure 85). It has also been demonstrated that FluKS1AT1 and FluKS5AT5 are capable of loading ethylmalonyl-CoA onto FluACP1, FluACP5 and DEBSACP6 with similar degree of efficiency, however, no ethylmalonyl-KirACP5 was detected (work by Irina Koryakina) (Figure 86).
Figure 85: Acyl-CoA specificity for FluKS1AT1. FluACP1 was used as ACP substrate.

Figure 86: ACP specificity for FluKS1AT1 and FluKS5AT5. Ethylmalonyl-CoA was used as the extender unit for this assay.

So far, there has been a very limited number of studies on cis-AT:ACP interactions. A crystal structure for AT\textsubscript{DYN10} from the iterative enediyne polyketide synthase have been solved
and its interaction with ACP investigated via computational docking.\textsuperscript{72} Similarly, the DEBS AT3:ACP3 interaction was probed by computational docking.\textsuperscript{91} AT\textsubscript{DYN10} and DEBS AT3 utilize malonyl-CoA and methylmalonyl-CoA as substrates, respectively. To the best of our knowledge, cis-AT:ACP interactions haven’t been previously studied with experimental approaches, let alone for cis-AT’s with substrate specificity other than malonyl- and methylmalonyl-CoA. With ethylmalonyl-CoA as the natural substrate and relative broad ACP specificity of FluAT1, this naturally cis-acting AT represents an attractive target to investigate. Understanding the FluAT1:ACP interaction could help complete our knowledge of cis-AT:ACP interaction profiles, which could potentially lead to more efficient cis-AT swapping results in polyketide combinatorial biosynthesis (Figure 87A). In addition, the relative broad ACP specificity of FluAT1 provides the potential of engineering FluAT1 to be trans-acting, which could have great application in complemention of AT-null modules (Figure 87B) and trans-AT swapping (Figure 87C), since most trans-AT’s identified so far use only malonyl-CoA as substrate. This approach could enable a new direction for the combinatorial biosynthesis of polyketides.
**Figure 87: Proposed FluAT1 applications.** (A), cis-AT swapping with a modular PKS’s; (B), trans-AT complementation for an AT-null module; (C), trans-AT swapping for PKS’s where extender units are provided by trans-AT’s.

Azidoethylmalonyl-CoA is a decent substrate for FluKS1AT1, thus an acyltransferase coupled fluorescent cycloaddition assay was employed to report the FluKS1AT1 transacylation activity (Figure 88).

**Figure 88: An acyltransferase coupled fluorescent cycloaddition assay to probe FluKS1AT1:ACP interaction.** FluKS1AT1 first catalyzes the transfer of azidoethylmalonyl (AzEM) group onto FluACP1 from AzEM-CoA, a fluorescent cycloaddition reaction is then performed. The assay is analyzed by SDS-PAGE and in-gel fluorescence.
6.2 Materials and methods

6.2.1 Expression of FluKS1AT1, FluACP1

FluKS1AT1 and FluACP1 plasmids were a kind gift from Nathan Schnarr (University of Massachusetts, Amherst). FluKS1AT1 were expressed and purified as described previously.\textsuperscript{132} Plasmids harboring FluACP1 were co-transformed with an Sfp encoding plasmid pSU20-Sfp into E. coli BL21 (DE3) for the \textit{in vivo} production of holo-ACP. A single colony was then used to inoculate 3 mL of LB (Luria Burtani) medium containing 34 µg/mL chloramphenicol and 100 µg/mL ampicillin which was cultured overnight at 37 °C with shaking at 250 rpm. The overnight culture was used to inoculate 300 mL of LB media containing the same antibiotics and was then incubated at 37 °C with shaking at 250 rpm until the OD\textsubscript{600} had reached ~0.6. Protein expression was induced by the addition of 0.5 mM IPTG and was then incubated at 18 °C with shaking at 250 rpm for 18 h. Proteins were purified by Ni-NTA chromatography as previously described.\textsuperscript{132}

6.2.2 Cloning and expression of FluAT1

Standalone FluAT1’s were PCR amplified from FluKS1AT1-pET21b vector using Phire Hot Start II DNA Polymerase (ThermoScientific, Inc.) following manufacturer’s protocols using the primers in Table 25. PCR was carried out as described in Chaper 2.2.1, except the annealing time was 1 min. FluAT1’s were cloned into pET28a via NcoI and NotI sites. Plasmids identities were confirmed by DNA sequencing. FluAT1’s were expressed and purified the same way as FluKS1AT1.
Table 25: Primer sequences for FluAT1’s.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>FluAT1-NcoI-FOR</td>
<td>AGCTTCCCATGGGGGTAGATAAACGCGAATCCTATACAAAGAGAAGACC</td>
</tr>
<tr>
<td>FluAT1-NotI-REV</td>
<td>AGCTTCCGCGCCGCCATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FluAT1_FAT-NcoI-FOR</td>
<td>AGCTTCCCATGGGGGTAGCAGGCTTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FluAT1_FAT-NotI-REV</td>
<td>AGCTTCCGCGCCGCCATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FluAT1_small-NcoI-FOR</td>
<td>AGCTTCCCATGGGGGTAGCAGGCTTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FluAT1_small-NotI-REV</td>
<td>AGCTTCCGCGCCGCTGGGTGATGCTATTGCTTCCCG</td>
</tr>
</tbody>
</table>

6.2.3 Cloning and expression of *E. coli* FAS proteins

*E. coli* FAS protein coding genes were amplified from *E. coli* genomic DNA using the following primers. PCR was carried out as described in Chaper 2.2.1, annealing time was adjusted based on gene size, 1 min/kb was used as the extension rate. The genes were cloned into pET28a using appropriate restrictions sites. Plasmids identities were confirmed by DNA sequencing.

Table 26: Primer sequences for *E. coli* FAS genes cloning.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>FabA-NcoI-FOR</td>
<td>AGCTTCCCATGGGGGTAGATAAACGCGAATCCTATACAAAGAGAAGACC</td>
</tr>
<tr>
<td>FabA-XhoI-REV</td>
<td>AGCTTCCGCGCCGCCCATTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FabD-NdeI-FOR</td>
<td>AGCTTCCGATAGACGCAATTTGCGATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FabD-EcoRI-REV</td>
<td>AGCTTCCGATAGACGCAATTTGCGATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FabG-NdeI-FOR</td>
<td>AGCTTCCGATAGACGCAATTTGCGATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FabG-EcoRI-REV</td>
<td>AGCTTCCGATAGACGCAATTTGCGATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FabI-NdeI-FOR</td>
<td>AGCTTCCATAGGGTTTTCTTTTCCGCTAAGCGCATCCTGG</td>
</tr>
<tr>
<td>FabI-EcoRI-REV</td>
<td>AGCTTCCGATAGACGCAATTTGCGATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>FabZ-NcoI-FOR</td>
<td>AGCTTCCCATGGGGGTAGATAAACGCGAATCCTATACAAAGAGAAGACC</td>
</tr>
<tr>
<td>FabZ-XhoI-REV</td>
<td>AGCTTCCGCGCCGCCCATTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>TesA-NcoI-FOR</td>
<td>AGCTTCCGATAGACGCAATTTGCGATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
<tr>
<td>TesA-XhoI-REV</td>
<td>AGCTTCCGATAGACGCAATTTGCGATTTCTGGTCACCGGCCGCCCCTGG</td>
</tr>
</tbody>
</table>

142
Plasmids harboring each *E. coli* FAS genes were transformed into *E. coli* BL21(DE3). A single colony was then used to inoculate 3 mL of LB (Luria Bertani) medium containing 50 μg/mL kanamycin which was cultured overnight at 37 °C with shaking at 250 rpm. The overnight culture was used to inoculate 500 mL of LB media containing the same antibiotics and was then incubated at 37 °C with shaking at 250 rpm until the OD₆₀₀ had reached ~0.6. Protein expression was induced by the addition of 0.5 mM IPTG and was then incubated at 18 °C with shaking at 250 rpm for 18 h. Proteins were purified by Ni-NTA chromatography following the same procedure as FabF, as described in Chapter 2.

### 6.2.4 Alanine scanning mutagenesis and expression of FluACP1

Primers for alanine scanning mutagenesis of FluACP1 were designed with lengths between 25 and 40 bp, while melting temperatures were between 70 and 75 °C. Site directed mutagenesis was performed following the QuikChange protocol (Agilent Technologies, Inc.). 2 min/kb was used as the extension rate. FluACP1-pET21b was used as template. FluACP1 mutants were expressed and purified under the same condition as WT FluACP1.

### 6.2.5 Site-directed mutagenesis and expression of FluAT1

Primers for site-directed mutagenesis of FluAT1 were designed with lengths between 25 and 35 bp, while melting temperatures were between 70 and 75 °C. Site directed mutagenesis was performed following the QuikChange protocol (Agilent Technologies, Inc.). 2 min/kb was used as the extension rate. FluAT1-pET21b was used as template. FluAT1 mutants were expressed and purified under the same condition as WT FluAT1.
6.2.6 Mutagenesis of AcpP mutants

Primers for site-directed mutagenesis of AcpP were designed with lengths between 25 and 40 bp, while melting temperatures were between 70 and 75 °C. Site directed mutagenesis was performed following the QuikChange protocol (Agilent Technologies, Inc.). 2 min/kb was used as the extension rate. AcpP-pET28a was used as template. AcpP mutants were expressed and purified as described in Chapter 4, Section 4.2.3.

6.2.7 General procedure for FluAT1 assay

FluAT1 assay was performed in 10 μL of reaction mixture containing 50 mM Tris-HCl (pH 7.5), 50 mM MgCl₂, 300 μM azidoethylmalonyl-CoA, 30 μM holo-FluACP1, and 3 μM FluAT1 at room temperature for 60 min.

6.2.8 General procedure for strain-promoted azide-alkyne cycloaddition assay

The strain-promoted azide-alkyne cycloaddition assay was performed in a total volume of 6 μL and contained 4 μL of FluAT1 assay mixture containing 20 μM FluACP1 and 2 μL of 600 μM of DBCO-fluor 488 (Sigma-Aldrich, Inc.). Reaction was incubated in dark at room temperature with gentle agitation for 1 h. Reactions were boiled after addition of protein loading dye for 5 min before analysis by SDS-PAGE. The gels were scanned using a Typhoon 7000 phosphoimager to determine the intensity of DBCO-fluor 488 labeled proteins bands. The bands were quantified by ImageQuant TL software (GE Life Sciences, Inc.) and rubber-band background subtraction was applied.

6.2.9 *In vitro* fatty acid production

*In vitro* fatty acid production was performed following published protocol. The assay was performed in 50 μL of reaction mixture containing 100 mM phosphate (pH 7.5), 1 mM
TCEP, 2.5 mM malonyl-CoA, 1 mM acetyl-CoA, 5 mM NADH, 5 mM NADPH, 1 µM FabA, 1 µM FabB, 1 µM FabD, 1 µM FabF, 1 µM FabG, 1 µM FabH, 10 µM FabI, 10 µM FabZ, 30 µM holo-AcpP and 30 µM TesA, at room temperature for 15 min.

Reaction was quenched by addition of 250 µL 4:1 (v:v) isopropanol:acetic acid and 200 µL water. 500 µL hexane was then added to the mixture in order to extract the free acids. After thorough mixing by vortexing, sample was centrifuged to expedite phase separation. 400 µL hexane layer was removed and dried under reduced pressure. The fatty acids were esterified by redissolving the dry sample in 500 µL 5% (v/v) H₂SO₄ in MeOH, incubating at 90 °C for 2 h after the reaction, 500 µL 0.9% (wt/vol) NaCl was added, fatty acid methyl esters were extracted using 100 µL hexane for subsequent GC-MS analysis.

GC-MS analysis was performed on Agilent 5975 GC/MS. The GC column is an Agilent HP-5ms, 30 m length, 0.25 mm diameter, 0.25 µm film. The GC gradient is as follows: holding at 100 °C for 4 min, the temperature was increased to 240 °C at a rate of 3 degrees per minute and holding at 240 °C for 10 min.

6.2.10 M6TE complementing

Attempted FluAT1 complementation of M6TE-AT° was performed as described in Chapter 5, Section 5.2.16.

6.3 Results and discussion

6.3.1 FluKS1AT1 specificity towards ACP’s

FluKS1AT1 has been shown to transfer ethylmalonyl-CoA extender units onto FluACP1, FluACP5 and DEBS ACP6 in our laboratory (work by Irina Koryakina). Here,
FluKS1AT1 activity towards a more comprehensive panel of ACP substrates was investigated (Figure 89) and azidoethylmalonyl-CoA was used as the substrate. Of the 20 ACP’s from the diverse PKS and FAS systems tested, 10 ACP’s showed over 20% activity of WT-FluACPI in the assay, all of which are PKS ACP’s. This data was in contrast with previous reports, where cis-AT6 from DEBS displayed a strong preference towards its cognate ACP, and about 10-fold decrease in transacylation activity was observed for non-cognate ACP’s. This data suggests that FluKS1AT1 displays a unique ACP specificity profile compared to DEBS AT6, even though FluAT1 displays 47% sequence identity to DEBS AT6. This observation presents FluAT1 as a potentially attractive target to be engineered into a trans-acting AT domain. Given its extender unit and ACP promiscuity, FluAT1 could potentially compliment KirCII-based approaches discussed in the previous chapter in the study of complementing AT-null modules. In addition, it would be intriguing to examine the molecular details regulating the FluAT1:ACP promiscuity and how this was different from the stringent KirCII:KirACP5 interaction described in Chapter 5.

Remarkably, KirACP5 displayed over 30% activity in the FluKS1AT1 assay where azidoethylmalonyl-CoA was used as substrate. This data contrasted with results by Irina Koryakina, where no ethylmalonyl-KirACP5 was detected in the FluKS1AT1 assay. This difference possibly suggested that FluKS1AT1’s ACP specificity is also dependent on the acyl-CoA substrate.
Figure 89: ACP specificity of FluKS1AT1. The assay for each ACP was performed twice. Average transacylation activity was plotted relative to WT FluACP1, which was set to 100%. Experimental errors were presented as standard deviations. Azidoethylmalonyl-CoA was used as the substrate.

6.3.2 Mapping the FluKS1AT1:KirACP5 epitope by alanine scanning mutagenesis

In a manner similar to that described for the KirACP5:KirCII interaction, to unveil the molecular basis for the ACP promiscuity of FluKS1AT1 and to compare the KirACP5 epitopes involved in the stringent KirCII:KirACP5 and relatively promiscuous FluKS1AT1:KirACP5 interactions, the KirACP5 epitope involved in the interaction with FluKS1AT1 was delineated by alanine scanning mutagenesis (Figure 90). KirACP5-V57A, R70A and D82A showed less than 20% activity of WT-KirACP5. These three positions correlate to positions known to be involved in multiple KS:ACP, AT:ACP and FabA:ACP interactions, as summarized in Table 23. Cumulatively, the epitope generally resembles that from the previously described KirACP5:KirCII interaction, which are mainly constituted of residues from L_I (V57) and H_{II} (R70) (Figure 91 and Figure 92). Notably, a residue from L_{II} (D82) was identified to be
contributing to FluKS1AT1:KirACP5 interaction, implying a unique feature of the FluKS1AT1:KirACP5 interaction.
Figure 90: Mapping the FluKS1AT1:KirACP5 interaction epitope by alanine scanning mutagenesis. Trans-acylation rates of holo-KirACP5 alanine mutants with FluKS1AT1 are expressed as a percentage of the activity with wild-type KirACP5. Mutants that displayed <20% the activity of the wild-type holo-KirACP5 are highlighted red.
Figure 91: Trans-acylation activities of KirACP5 alanine mutants mapped onto KirACP5 homology model. Ribbon diagram (left) and its computed surface (right). Red, <20% activity; green, 20-80% activity; yellow, >80% activity; grey, not determined.

Figure 92: Comparison of KirCII:KirACP5 and FluKS1AT1:KirACP5 interaction epitope. Magenta, KirCII:KirACP5 epitope; red, FluKS1AT1:KirACP5 epitope; R70 is involved in both epitope and is colored in dark red; S65 is shown as stick presentation.
KirACP10 mutants and chimeras described in Chapter 5 that displayed improved activity in the KirCII assay were re-examined here using FluKS1AT1 (Figure 93). Interestingly, over 3-fold improvements were observed for all of them compared to the WT KirACP10. In fact, KirACP10-H* and KirACP10-NA5 displayed levels of transacylation activity higher than wild-type FluACP1. This data indicates that the entire N-terminal portion of KirACP5 (L₀–H₁–L₄) or the KirACP5 helical turn H₁’ is sufficient for robust recognition by FluKS1AT1. The transacylation activity enhancement for KirACP10 chimeras and mutants was likely attributed to KirCII:KirACP5 and FluKS1AT1:KirACP5 interactions share similar interaction epitopes (Figure 92).

![Figure 93: FluKS1AT1 assay on ACP mutants](image)

**Figure 93: FluKS1AT1 assay on ACP mutants.** The assay for each ACP mutant/chimera was performed twice. Average transacylation activity was plotted relative to WT FluACP1, which was set to 100%. Error bars are standard deviations.
6.3.3 Construction of standalone FluAT1

To further investigate the FluAT1:ACP interaction, standalone FluAT1’s were designed based on a FluKS1AT1 homology model and sequence alignment with DEBS AT6. A homology model for FluKS1AT1 was constructed by Swiss-Model Workspace automated docking mode (Figure 94). 2QO3 was selected by Swiss-Model as as the template. 2QO3 is the pdb file for DEBS KS3AT3 didomain. Sequence identity between FluKS1AT1 and DEBS KS3AT3 is 55% and QMEAN Z-Score was -1.475, which is in the reliable range of modeling. Three versions of the standalone FluAT1 were constructed to evaluate the importance of linker regions. FluAT1 contains both N-terminal KS-AT linker and C-terminal post-AT linker, FluAT1_PAT contains only C-terminal post-AT linker and FluAT1_small contains none of the linkers (Figure 95).

Figure 94: Homology model of FluKS1AT1. FluKS1, FluAT1, KS-AT linker, and post-AT (PAT) linker are colored orange, green, magenta and cyan, respectively.
**Figure 95: Design of standalone FluAT1.** FluAT1 contains both KS-AT linker and PAT linker, FluAT1\textsubscript{PAT} contains only PAT linker and only the core AT part is present for FluAT1\textsubscript{small}.

All versions of FluAT1 were expressed and purified as soluble proteins. Expression level for FluAT1\textsubscript{PAT} and FluAT1\textsubscript{small} was at least five-fold better than FluAT1 (results not shown). Each standalone FluAT1 was assayed in the same way as FluKS1AT1, using FluACP1 and azidoethylmalonyl-CoA as substrates (Figure 96). FluAT1 displayed ~70% activity of the “full-length” FluKS1AT1, while FluAT1\textsubscript{PAT} showed ~90% transacylation activity compared to FluKS1AT1. This data indicates that the KS-AT linker plays a minor role in the FluAT1:FluACP1 interaction, which is in contrast to the role of the KS-AT in DEBS AT:ACP interactions whereby the KS-AT linker constitutes part of ACP docking site. Accordingly, deletion of the KS-AT linker from DEBS AT’s resulted in compromised DEBS AT:ACP interaction.\textsuperscript{91} FluAT1\textsubscript{small} displayed ~55% transacylation activity compared to FluKS1AT1, suggesting presence of the PAT linker is important but not mandatory for FluAT1 transacylation activity. Overall, the results suggest that FluAT1 does not require either N- and C-terminal linkers to interact with FluACP1 in the non-native trans-acting acyl-transfer assay, in contrast to standalone DEBS AT6, where deletion of either N- or C-terminal linker has a deleterious effect on both the specificity and activity.\textsuperscript{91}
6.3.4 M6TE complementation

In the FluKS1AT1-catalyzed transacylation assay, ACP6 was a decent substrate, displaying ~25% activity compared to WT FluACP1, it was thus intriguing to investigate if FluKS1AT1, FluAT1, FluAT1\textsubscript{PAT} and FluAT1\textsubscript{small} could be used to complement holo-M6TE-AT\textsuperscript{o}. Accordingly, FluKS1AT1, FluAT1, FluAT1\textsubscript{PAT} and FluAT1\textsubscript{small} complementation of holo-M6TE-AT\textsuperscript{o} was performed with ethylmalonyl-CoA as the extender unit. Unfortunately, no triketide lactone product was detected by HPLC (results not shown) for all the complementation assays. The failure was probably due to the insufficient FluAT1:ACP6 interaction. Future kinetic characterization of FluAT1 and MCAT acyltransfer onto ACP6, respectively, should help understand the kinetic requirements for trans-AT complementation.
6.3.5 Mapping the FluAT1:FluACP1 epitope by alanine scanning mutagenesis

Previous epitope mapping for FluAT1:KirACP5 suggest some shared features compared to that of the KirCII:KirACP5 interaction. To evaluate if the epitope location is common for FluAT1:ACP interactions involving other ACP substrates, the epitope for the cognate interaction between FluAT1 and FluACP1 was probed by alanine scanning mutagenesis. A FluACP1 homology model was constructed by Swiss-Model Workspace automated docking mode to identify the surface exposed residues (Figure 97). DEBS ACP2 (pdb: 2JU2\(^1\)) was selected by Swiss-Model as the template. Sequence identity between FluACP1 and DEBS ACP2 is 51\% and the QMEAN Z-Score was -0.407, which is in the reliable range of modeling. The FluACP1 homology model represents a typical ACP structure, composed of a three helix bundle, connected by two long flexible loops, the N-terminus region of FluACP1 is modeled as a flexible loop.

![Homology model of FluACP1](image)

**Figure 97: Homology model of FluACP1.** The secondary structural elements of FluACP1 are labeled accordingly.
Given the FluKS1AT1:KirACP5 interaction epitope is distributed across L1, HII and LII on KirACP5, FluACP1 alanine scanning mutagenesis was focused on this region, and 36 non-alanine surface exposed residues in loop I, helix II and loop II were individually mutated into alanine (Figure 98). The DSL motif wasn’t subjected to mutagenesis because previous results showed mutations in this motif compromise the post-translational modification of ACP’s (Chapters 4 and 5). Each mutant was expressed in E. coli and purified to homogeneity. Complete conversion from the apo-FluACP1 to holo-FluACP1 for all alanine mutants by treatment with Sfp and CoA was confirmed by MS analysis. Of the 36 alanine mutants constructed, seven mutants (L28A, G29A, I36A, F42A, G46A, F47A and L67A) weren’t expressed under the conditions tested. Mapping the location of these positions onto the FluACP1 structure reveals that L28, F42, F47 and L67 are potentially involved in hydrophobic interactions between nearby FluACP1 residues and could contribute to the stability of FluACP1.

Figure 98: Design of FluACP1 alanine mutants. (A), regions colored in cyan were subjected to alanine scanning mutagenesis, S49 is shown as stick presentation; (B), alanine mutants that didn’t express were shown in magenta sticks presentation.
Transacylation assay was performed using the remaining soluble FluACP1 mutants, using FluAT1\textsubscript{PAT} and azidoethylmalonyl-CoA as substrate (Figure 99). Surprisingly, almost all the mutants showed over 50% activity compared to WT FluACP1. Only D74A located in L\textsubscript{II} showed less than 50% activity compared to WT FluACP1. These mutagenesis results resemble those from a DSZS:DSZSACP1 epitope mapping study, where most alanine DSZSACP1 mutants showed over 50% activity of WT DSZSACP1.\cite{48} KS6AT6 and AT6 from DEBS showed similar level of transacylation activity towards DEBS ACP6, suggesting that presence of KS has a negligible role on AT transacylation.\cite{91} Taken together, the FluACP1 epitope differed from the KirACP5 epitope when interacting with FluAT1, and the difference wasn’t due to FluAT1\textsubscript{PAT} and FluKS1AT1 were used in the transacylation assay, respectively. Instead, the data implied that there is no universal interaction epitope for FluAT1:ACP interaction.

![Figure 99: Summary for FluACP1 alanine scanning mutagenesis results.](image)

The assay for each FluACP1 mutant was performed twice. Average transacylation activity were plotted relative to WT FluACP1, which was set to 100%. Error bars are standard deviations.
6.3.6 Computational docking

A homology model for the FluAT1 portion of the FluKS-AT1 homology model was obtained by deleting the FluKS1 sequence from the pdb file for the FluKS1AT1 homology model. Computational docking was performed with Patchdock. Constraints were imposed where distance between Ser49 side chain oxygen on FluACP1 and Ser185 side chain oxygen on FluAT1 was restricted between 15 to 25 Å.

The docking models were evaluated based on the FluACP1 alanine scanning mutagenesis results, where D74 on FluACP1 should be oriented close to FluAT1. One top-ranking model is presented in Figure 100. The distance between Ser49 side chain oxygen on FluACP1 and Ser185 side chain oxygen on FluAT1 is calculated to be 22.3 Å, and a tunnel could be identified that could accommodate the phosphopantethine arm. Consistent with the previous standalone FluAT1 assays, the KS-AT linker was not predicted to be involved in the FluAT1:FluACP1 interaction. Interestingly, few electrostatic/hydrophobic interaction pairs could be identified from the docking model. D74 on FluACP1 is positioned to interact with R186 and R188 on FluAT1. Additionally, the docking model presented resembles the trans-AT:ACP docking models for MCAT:actACP and DSZS:DSZSACP1, where few interactions pairs were identified. This lack of amino acid interaction pairs between FluACP1 and FluAT1 presumably explains the ACP promiscuity of FluAT1. This data highlights the underlying “trans-acting” potential for FluAT1, whereby the broad extender unit and ACP promiscuity could be leveraged for combinatorial biosynthesis.
Figure 100: Computational docking of FluACP1 (grey) onto FluAT1 (green). S49 and D74 on FluACP1 are shown in red and magenta sticks, respectively. R186 and R188 on FluAT1 are shown as blue sticks.

6.3.7 Engineering non-cognate ACP’s to be recognized by FluAT1

Although FluKS1AT1 was quite promiscuous towards a broad panel of ACP’s, a general trend was observed where FluKS1AT1 showed a strong preference for PKS ACP’s over FAS ACP’s. In particular, AcpP was the least active in the ACP panel assay. Sequence alignment revealed an intriguing correlation of ACP transacylation activity and amino acid identity at several ACP residues, highlighted grey in Table 27. Notably, these positions proved sensitive to alanine substitution (Section 6.3.5). For example, Asp and Thr are present in KirACP10 and AcpP, respectively, at an equivalent position to I36 from FluACP1. Phe on KirACP15 and Ala on AcpP are present at equivalent positions to G46 and F47 in FluACP1, respectively. Finally, Gly on KirACP10 and Lys on AcpP are present at a position equivalent
to D74 in FluACP1. Taken together, these amino acid identity differences might contribute to the poor AcpP:FluKS1AT1 interaction.

**Table 27: Amino acid alignment and transacylation activity summary for ACP’s.** Amino acids L28, G29, I36, F42, G46 and F47 correspond to locations on FluACP1 that didn’t express in alanine scanning mutagenesis. Some representative amino acids potentially contributing to the transacylation activity difference are highlighted in grey.

<table>
<thead>
<tr>
<th>ACP’s</th>
<th>Amino acid residues</th>
<th>% FluACP1 transacylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>FluACP1</td>
<td>L28 G29 I36 F42 G46 F47 D74</td>
<td>100</td>
</tr>
<tr>
<td>DEBS ACP2</td>
<td>L G V F G F D D</td>
<td>116.4</td>
</tr>
<tr>
<td>KirACP7</td>
<td>L Q E F G I E</td>
<td>60.2</td>
</tr>
<tr>
<td>KirACP0</td>
<td>L G T L G L Q</td>
<td>49.3</td>
</tr>
<tr>
<td>KirACP1</td>
<td>T R R L G V D</td>
<td>43.2</td>
</tr>
<tr>
<td>actACP</td>
<td>D G F F G Y R</td>
<td>41.0</td>
</tr>
<tr>
<td>KirACP5</td>
<td>L G P L G V V</td>
<td>35.5</td>
</tr>
<tr>
<td>KirACP4</td>
<td>A G T L G F G</td>
<td>31.1</td>
</tr>
<tr>
<td>KirACP3</td>
<td>S G D I G F E</td>
<td>30.6</td>
</tr>
<tr>
<td>KirACP13</td>
<td>L G I F G F D</td>
<td>23.8</td>
</tr>
<tr>
<td>KirACP14</td>
<td>L G I F G F D</td>
<td>17.8</td>
</tr>
<tr>
<td>KirACP8</td>
<td>L E W Y G V E</td>
<td>11.9</td>
</tr>
<tr>
<td>KirACP12</td>
<td>T G T F G L E</td>
<td>11.6</td>
</tr>
<tr>
<td>ScFAS ACP</td>
<td>A G K T D V N</td>
<td>11.2</td>
</tr>
<tr>
<td>KirACP2</td>
<td>A A V F G L E</td>
<td>8.4</td>
</tr>
<tr>
<td>KirACP11</td>
<td>H G V F G V E</td>
<td>8.4</td>
</tr>
<tr>
<td>KirACP9</td>
<td>T G D F G L A</td>
<td>5.9</td>
</tr>
<tr>
<td>KirACP10</td>
<td>L G D F G L G</td>
<td>5.2</td>
</tr>
<tr>
<td>KirACP15</td>
<td>L G I G F L E</td>
<td>4.8</td>
</tr>
<tr>
<td>AcpP</td>
<td>L G T V G A K</td>
<td>0.7</td>
</tr>
</tbody>
</table>

Accordingly, two AcpP mutants, A34F and K61D were designed in order to introduce residues equivalent to key FluACP1 positions (**Table 27**). Both mutants served as better substrates than WT AcpP in transacylation assays with FluAT1\textsubscript{PAT}, with A34F and K61D
showing close to 8-fold and 3-fold improvement over WT AcpP, respectively (Figure 101). This data suggested the importance of hydrophobic and charged residues in successful recognition by FluAT1. Similar or more significant enhancement in transacylation activity could potentially be observed if other mutations were introduced into AcpP, such as T23I.

**Figure 101: FluAT1PAT assay on AcpP mutants.** The assay for each AcpP mutant was performed twice. Average transacylation activity were plotted relative to WT FluACP1, which was set to 100%. Experimental errors were presented as standard deviations.

### 6.3.8 In vitro fatty acid production

Fatty acids are the major components of cell membranes and important metabolic intermediates in bacteria. Fatty acid synthases (FAS’s) can be classified into two types. Type I FAS’s are usually found in mammals, and are encoded on a single gene with different catalytic domains distributed in a large polypeptide. Type II FAS’s are found in bacteria, plants
and parasites. Type II FAS’s are encoded on multiple genes with each gene expressing a unique domain responsible for one catalytic step.\textsuperscript{54}

Fatty acid biosynthesis is initiated by acetyl-CoA carboxylase, which is a heterotetrameric enzyme encoded by four genes, accA, accB, accC and accD, converting acetyl-CoA into malonyl-CoA. Malonyl-CoA:ACP transacylase (FabD) then catalyzes the transfer of malonyl group from malonyl-CoA onto acyl carrier protein (ACP) to form malonyl-ACP. \(\beta\)-ketoacyl-ACP synthase III (FabH) catalyzes the condensation of malonyl-ACP with acetyl-CoA to form \(\beta\)-ketobutyryl-ACP and \(\text{CO}_2\). \(\beta\)-ketoacyl-ACP reductase (FabG) catalyzes reduction of \(\beta\)-ketoacyl-ACP to \(\beta\)-hydroxyacyl-ACP in an NADPH-dependent manner. \(\beta\)-hydroxyl-ACP is then dehydrated to produce trans-2-enoyl-ACP by either \(\beta\)-hydroxydecanoyl-ACP dehydratase/isomerase (FabA) or \(\beta\)-hydroxyacyl-ACP dehydratase (FabZ). Trans-2-enoyl-ACP is converted to acyl-ACP in an NAD(P)H-dependent reduction by enoyl-ACP reductase I (FabI). Subsequent elongation cycles are initiated by either \(\beta\)-ketoacyl-ACP synthase I (FabB) or \(\beta\)-ketoacyl-ACP synthase II (FabF) by condensing malonyl-ACP with the acyl-ACP, the condensed product undergoes the same reducing procedure until a saturated fatty acid (SFA) of 16 or 18 carbons is achieved. Both saturated (SFA) and unsaturated (UFA) fatty acids are produced in \textit{E. coli}. The ratio of SFA to UFA is determined by the expression level of FabB. UFA production is increased in response to elevated FabB expression.\textsuperscript{135} Fatty acid chain length is mainly regulated by the size of the acyl-intermediate binding site on the condensing enzymes.\textsuperscript{134}

All type II FAS’s have this basic set of enzymes for the generation of fatty acid, and diverse naturally occurring fatty acid structures can be attained by manipulating this basic
Limited natural modification to fatty acid structures has been observed, which is due to the finite variety of starter and extender units available.

**Scheme 2: *E. coli* FAS biosynthesis.** Fatty acid biosynthesis initiation is presented on the right panel, the growing fatty acyl chain is then iteratively condensed and reduced until desired fatty acyl chains are obtained, at which point it can be hydrolyzed by a thioesterase or undergo other necessary modifications.

Given that several AcpP mutants exhibited significantly improved activity in the FluAT1 transacylation assay (Figure 101), it was hypothesized that unnatural extender units could be incorporated into the fatty acid biosynthetic pathway, by replacing FabD with FluAT1, and using AcpP mutant in place of WT AcpP. The hypothesis was tested by *in vitro* fatty acid production via purified *E. coli* FAS proteins (FabA, FabB, FabD, FabF, FabG, FabH, FabI, FabZ, holo-AcpP and TesA). *E. coli* thioesterase I (TesA) is one of the two native thioesterases in *E. coli*. TesA is a periplasm-directed enzyme and upon deletion of the 5’ leader
peptide sequence is present in the cytosol where it hydrolyzes predominantly C14 and C16 fatty acid chains. Introducing of L109P mutation to TesA enables it to hydrolyze both short and long chain fatty acids. Accordingly, TesA was cloned without the leader sequence and the L109P mutation was introduced.

Commercial fatty acid methyl ester (FAME) standard was analyzed by GC-MS to identify the elution profile of various length products (Figure 102). In vitro fatty acid production by WT E. coli fatty acid synthase proteins using malonyl-CoA as extender unit was analyzed using the same protocol. Gratifyingly, C16 and C18 fatty acid products were detected (Figure 103). Having established feasibility of this approach, non-natural extender units will be incorporated by replacing FabD with FluAT1, and using AcpP mutants in place of WT AcpP.
Figure 102: GC-MS analysis of fatty acid methyl ester (FAME) standard.
Figure 103: GC-MS analysis of WT in vitro fatty acid production. *In vitro* fatty acid production was performed as described in the experimental section, the product was analyzed with identical protocol as FAME standard. The results were compared to FAME standard to identity each fatty acid product.

6.4 Conclusion

An acyltransferase coupled fluorescent cycloaddition assay was developed to probe the FluAT1:ACP interaction. FluAT1 displayed broad ACP promiscuity, particularly towards PKS ACP’s. The FluAT1:KirACP5 interaction epitope was mapped via alanine scanning mutagenesis, revealing that the location of KirACP5 epitope involved in the FluAT1:KirACP5 interaction generally resembles that for KirCII:KirACP5 interaction, suggesting a similar mode of interaction. The cognate FluAT1:FluACP1 interaction epitope on FluACP1 was also examined. Most FluACP1 alanine mutants displayed similar activity as WT FluACP1 in the transacylation assay, D74A displayed the most significant decrease in activity. Computational
docking models for FluAT1:FluACP1 was evaluated based on the FluACP1 alanine scanning mutagenesis. Few amino acid interaction pairs were identified in the selected docking model, which partly explains the ACP promiscuity of FluAT1. Based on sequence alignment analysis, AcpP mutants were designed and demonstrated significantly improved activity in FluAT1-catalyzed transacylation assay compared to WT AcpP, offering the opportunity to introduce non-natural extender units into fatty acid biosynthesis pathway. Overall, the results highlighted the trans-acting potential of FluAT1.
CHAPTER 7. Probing the impact of prosthetic arm analogues on the activity and specificity of natural product biosynthesis

7.1 Introduction

Carrier proteins (CP’s) are involved in all natural product biosynthetic pathways that utilize carrier protein bound substrates. CP’s are responsible for delivering growing chains and extender units to every active site of a type I or type II catalytic domain. Given this, for PKS’s, CP’s interact with every domain in the same module and the downstream KS for type I systems (Figure 104), and for type II systems, CP’s interact with every domain in the pathway.\(^{139}\)

![Figure 104: ACP-based interactions that take place during each round of chain extension on modular, type I PKS’s.](image)

Reducing domains are not shown here, but they are present in most modules and interact with the ACP. Paths a, b represent interactions within the same module. Path c represents interactions between different modules, in this case separate polypeptides (i.e. interpolypeptide protein:protein interactions). Horizontal wavy lines: interpolypeptide linkers, vertical wavy line: phosphopantetheine. R1 and R2 depend on the module number and the identity of the assembly line.
For CP’s to function, post-translational modification (PTM) is required to attach the phosphopantetheine group (Ppant group) from CoA onto the conserved serine site, converting CP’s from the apo to active holo form (Figure 105).\textsuperscript{139} As the prosthetic arm of CP’s, Ppant is responsible for holding the growing chains and extender units, and delivering them to the proper active sites of other domains. However, for most natural product assembly lines, including PKS’s and nonribosomal peptide synthases (NRPS’s), CP’s isolated from recombinant expression in \textit{E. coli} are usually in their apo form. Thus, additional PTM of CP’s is necessary and this is usually accomplished with a promiscuous phosphopantetheinyl transferase (PPTase), Sfp from \textit{Bacillus subtilis}.\textsuperscript{62}
**Figure 105: Post-translational modification of CP’s.** Ppant group is transferred to the conserved Ser on CP’s from CoA through the catalysis of phosphopantetheinytransferases (PPTase). Shown here is the crystal structure of AcpP (PDB: 1T8K).

There are several functional groups present in Ppant and the focus of previously described modifications is the terminal thiol position.\(^{25, 140, 141}\) Ppant analogs were made as mechanism-based crosslinkers to probe ACP:KS (Figure 7) and ACP:DH interactions. A small panel of them is shown in Figure 106.\(^{25, 141}\) The Ppant analogs showed varying degrees of efficiency towards trapping ACP:KS interactions. However, other than modification to the terminal thiol portion of Ppant, few modifications to the Ppant structure have been made to explore structural requirements for activity. In addition, the preparation and installation of such
Ppant analogs into ACP is highly complex, requiring the combined action of lengthy chemical synthesis and four enzymes.\textsuperscript{25}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure.png}
\caption{Mechanism-based Ppant analogs to trap ACP:KS interactions. Ppant analogs with efficient (A), moderate (B) and poor (C) crosslinking of cognate ACP:KS pair.}
\end{figure}

It is postulated that portions of the Ppant group provide a handle for recognition by the interacting domains, however, there has been very limited reports that probe the specific roles of individual Ppant functionality. Deletion of the Ppant geminal dimethyl group resulted in only a minor affect on the conformation of Ppant, as judged by NMR analysis.\textsuperscript{142} In contrast, the amide bond nearest the thiol group has been shown to be important for enzyme catalysis, and a strict distance requirement for the amide group and thiol was also established.\textsuperscript{143}

For a long time, the paradigm of bio-engineering involves amino acid mutagenesis of catalytic machinery. However, cofactor engineering provides an intriguing yet underexplored complement to traditional protein engineering. By preparing a diverse panel of pantetheine analogs, the role of the prosthetic arm in acyltransfer, condensation, reduction, dehydration
and cyclization steps in natural product biosynthesis can be studied. Ultimately, the structural requirements for a functional prosthetic arm could be unmasked. It is hoped that by installation of non-natural prosthetic arms, analogs will be found that broaden the promiscuity, or change substrate specificity of natural product biosynthetic domains.

To circumvent the lengthy chemical synthesis and enzyme-based installation of Ppant analogues, we will use bioorthogonal conjugation reactions to rapidly install pantetheine analogs onto ACP’s (Figure 107). Accordingly, an unnatural amino acid with bioorthogonal conjugation handles, for example, alkyne, azide and acetyl groups, will be incorporated into ACP’s at the conserved Ser site, yielding mutant ACP’s. Pantetheine analogs with terminal azide, alkyne or hydrazide groups will be synthesized and reacted with mutant ACP’s, yielding “crypto-ACP’s”, which are analogs of holo-ACP’s. Comparison of “crypto pantetheine” that is the closest mimic to Ppant and Ppant group is shown in Figure 108. Generally, the “crypto pantetheine” is less flexible than pantetheine at the phosphate end, and the negatively charged phosphate is replaced by a hydrophobic benzene ring. The geminal dimethyl group of pantetheine is replaced with a slightly larger triazole or similar sized hydrazide in the “crypto-pantetheine”. In addition, the amide group closest to the ACP polypeptide chain is absent for crypto-pantetheine via ketone-hydrazide reaction (Figure 108C).
Figure 107: Click Ppant analogs onto CP’s. Ppant analogs will be installed via unnatural amino acid mutagenesis and click reaction, para-azido-phenylalanine (pAzF) is incorporated at the conserved Ser site and pantetheine analog with terminal alkyne is clicked onto CP-pAzF, yielding crypto-CP. Shown here is the crystal structure of AcpP (PDB: 1T8K).

Figure 108: Comparison of the closest mimic of “crypto pantetheine” and Ppant group. (A) Phosphopantetheine. (B) Ppant analog introduced via azide-alkyne cycloaddition reaction. (C) Ppant analog introduced via ketone-hydrazide conjugation reaction. Wavy line, polypeptide chain.
7.2 Materials and methods

7.2.1 Synthesis of para-propargyloxyphenalanine (pPrF)

pPrF was synthesized as described in the published procedure.\textsuperscript{144} All reagents used were commercially available and were of the highest purity available. Proton NMR spectrum of the final product was taken to verify its structure.

\textbf{Scheme 3: Para-propargyloxyphenalanine (pPrF) synthesis.}

7.2.2 Unnatural amino acid mutagenesis

ACP2-pAzF and ACP2-pPrF were expressed using the pEVOL system developed by Peter G. Schultz and were expressed in the same way as that of ACP2-pBpa.\textsuperscript{29,41,145} pAzF and pPrF were included in the culture media at 1 mM. Pure mutant proteins were prepared as previously described in Chapter 2, Section 2.2.2. The mutant proteins were analyzed by MS.

AcpP-AcF was also expressed using the pEVOL system in the same way as that of AcpP-pBpa. AcF was included in the culture media at 1 mM. Pure mutant proteins were prepared as previously described in Chapter 2, Section 2.2.2. The mutant proteins were analyzed by MS.
7.2.3 Cu (I)-catalyzed azide-alkyne cycloaddition (CuAAC) reaction

CuAAC click reaction was performed essentially as described. Reactions were carried out in Eppendorf tubes. Reaction volume was 50 μL in PBS buffer (pH 7.4). For click reaction between ACP2-pAzF with alkyne rhodamine, 50 μM ACP2-pAzF, and 50 μM alkyne rhodamine dye was used, sodium ascorbate (1 mM final concentration) was then added into the tube, followed by tris-(benzyl triazolylmethyl)amine (TBTA) ligand (100 μM final concentration). The mixture was vortexed briefly and CuSO₄ (1 mM final concentration) was added. The mixture was vortexed again and centrifuged briefly to bring down all reagents from tube wall, and allowed to react for 1 hour at room temperature. After 1 hour, aliquots were taken and mixed with protein loading dye, boiled for 5 min at 95 °C and samples were then analyzed by SDS-PAGE.

Click reaction of ACP2-pPrF with azide rhodamine was carried out in similar ways.

CuAAC reaction between pAzF and Ppant analog was performed with CuSO₄ (0.1 mM), Ppant analog (50 μM), pAzF (500 μM), sodium ascorbate (5 mM), THPTA (500 μM) and 10% (v/v) DMSO, for 2 h under reduced pressure at room temperature. Copper source and concentration, ligand, reducing agent, % DMSO have been varied to identify the optimal reaction conditions. Small molecule CuAAC reactions were analyzed by HPLC on a Varian ProStar HPLC system. A series of linear gradients was developed from 0.1% TFA in water (A) to 0.1% TFA in acetonitrile (HPLC grade, B) using the following protocol: 0-5 min, 80% A; 5-25 min, linear gradient from 80% A to 20% A; 25-30 min, isocratic flow at 20% A; 30-31 min, linear gradient from 20% A to 80% A; 31-36 min, isocratic flow at 80% A. The flow
rate was 1 mL/min, and the absorbance was monitored at 220 nm using Pursuit XRs C18 column (250 mm × 4.6 mm, Varian Inc.).

7.2.4 Synthesis of Ppant analog

Synthesis of Ppant analogs was carried out by Zou Yan (Deiters lab, University of Pittsburgh). The synthesis followed the schemes described below. Proton NMR spectrum of the each intermediate product was taken to verify its structure. The final product was also analyzed by MS.

![Scheme 4: Synthesis of Ppant analog.](image)

7.2.5 General procedure for capping Ppant free thiol by iodoacetamide

100 µM Ppant analog was incubated with 10 mM iodoacetamide for 1 hour in the dark. Iodoacetamide solution was prepared fresh every time before use.

7.2.6 General procedure for ketone-hydrazide reaction

Ketone-hydrazide reaction was performed in 50 µL of reaction mixture containing 50 mM phosphate (pH 6.2), 50 mM NaCl, 200 µM AcF, 1 mM biotin hydrazide at room
temperature for 16 h. Ketone source (acetyl-phenylalanine and acetyl-lysine), pH (pH 4.6, pH 5.5, pH 6.2 and pH 6.5), temperature (room temperature, 30 °C and 37 °C) and organic additives (10% acetonitrile and 10% dimethylformamide) were investigated to identify the optimal reaction conditions.

7.3 Results and discussion

7.3.1 Unnatural amino acid mutagenesis

pAzF was incorporated into DEBS ACP2 at the conserved Ser site. The identity of the mutant protein was verified by MS. No incorporation of any natural amino acid at the Ser site was observed from the MS result (Figure 109).

Figure 109: Incorporation of pAzF into ACP2. (A) SDS-PAGE of ACP2 and ACP2-pAzF, M, marker; 1, ACP2; 2, ACP2-pAzF. (B) MS result for ACP2-pAzF. The experimental mass for ACP2-pAzF was 20458.62 Da, closely matching the expected mass, 20458.65 Da.

7.3.2 Click reaction of ACP2-pAzF with alkyne-rhodamine

As initial test of principle, click reactions between ACP2-pAzF/ACP2 with alkyne rhodamine, a fluorescent dye were performed. ACP2-pAzF was expected to click with alkyne rhodamine whereas ACP2 would not due to the absence of azide group in wild-type ACP2.
The click reactions were first analyzed by SDS-PAGE. In Figure 110A and 110B, the upper bands in lane 1 and 3 corresponded to alkyne rhodamine clicked with ACP2 and ACP2-pAzF respectively; the lower bands were unreacted alkyne rhodamine. Based on the intensity of upper bands (Figure 110C), click reaction worked efficiently, however, significant background labeling of ACP2 with alkyne rhodamine was observed.

![Figure 110: Click reaction between ACP2-pAzF and alkyne rhodamine (AR). (A) Direct visualization of the gel. (B) The gel under UV transluminator, upper fluorescent bands indicated rhodamine clicked on ACP2s, lower fluorescent bands represented unreacted rhodamine dye. (C) Coomassie stained gel. M, marker; 1, ACP2-pAzF clicked with AR; 2, ACP2-pAzF; 3, ACP2 “clicked” with AR; 4, ACP2.](image)

The click reactions were analyzed by MS to confirm the identity of the upper fluorescent bands. The experimental mass of ACP2-pAzF-AR complex closely matched the expected mass (Figure 111A), suggesting that the reaction worked successfully. Interestingly, the mass for “ACP2-AR” complex wasn’t detected (Figure 111B), which indicated that there was in fact no real “background labeling”, the presence of “clicked band” for ACP2 on SDS-PAGE might be due to the non-specific interaction between rhodamine and ACP2.
Figure 111: MS analysis of click reaction. (A) ACP2-pAzF clicked with AR, experimental mass of ACP2-pAzF-AR complex is 20925.92 Da, closely matching the expected mass 20926.16 Da, the 20458.94 Da peak corresponded to the mass for ACP2-pAzF. (B) ACP2 “clicked” with AR, 20357.83 Da represented mass for wild-type ACP2, however, the expected peak of ACP2-AR complex (20825.144 Da) wasn’t observed.

7.3.3 Synthesis of para-propargyloxyphenalanine (pPrF)

The background labeling observed on SDS-PAGE has been reported in the literature. The suggested solution to this problem was to swap the position of alkyne and azide, i.e. to have the azide group on the protein and have the alkyne group on the labeling reagent. To this end, an alkynyl unnatural amino acid (pPrF) was synthesized, as described in Scheme 3 with a final yield of 31%. NMR spectrum of pPrF (in D₂O) was in complete agreement with its structure (Figure 112).
7.3.4 Incorporation of pPrF into ACP2

pPrF was incorporated into DEBS ACP2 at the conserved Ser site. The identity of mutant protein was verified by MS. No incorporation of any natural amino acid at the Ser site was observed from the MS result (Figure 113).
7.3.5 Click reaction of ACP2-pPrF with azide-rhodamine

The intention of synthesizing an alkynyl unnatural amino acid and incorporating it into ACP2 was to compare the level of apparent background labeling between alkynyl dye and azido dye on SDS-PAGE. Thus, having prepared ACP2-pPrF and verified its identity, ACP2-pPrF was then clicked with azide-rhodamine. In addition, ACP2 was also incubated under click reaction condition with azide rhodamine in order to determine the extent of apparent background labeling in the absence of ACP-displayed alkyne. At the same time, the click reaction between ACP2-pAzF and alkyne rhodamine was repeated (Figure 114). The upper fluorescent bands indicated clicked rhodamine on ACP2’s, and lower fluorescent bands represented unreacted alkyne rhodamine dye. The lower fluorescent bands weren’t observed for azide-rhodamine, this may be related with the property of rhodamine dyes used.

Figure 113: Incorporation of pPrF into ACP2. (A) SDS-PAGE of ACP2-pPrF. (B) MS result for ACP2-pPrF. The experimental mass for ACP2-pPrF was 20471.79 Da, closely matching the expected mass, 20471.87 Da.
Figure 114: Click reaction of ACP2s with rhodamine dyes. (A) SDS-PAGE visualized under UV, upper fluorescent bands indicated clicked rhodamine on ACP2s, and lower fluorescent bands represented unreacted rhodamine dye. (B) Coomassie stained SDS-PAGE.

Figure 115: Structures of rhodamine dyes used for click reaction. (A) Alkyne rhodamine. (B) Azide rhodamine.

Background labeling was still observed for alkyne rhodamine. However, the click efficiency between ACP2-pPrF and azide rhodamine was very low, a very faint fluorescent band for ACP2-pPrF clicked with azide rhodamine could be observed when viewed directly under UV but difficult to recognize from the gel picture here. Thus it wasn’t safe to draw the
conclusion that background labeling was low for this pair. Click reaction between ACP2-pPrF and azide rhodamine was then optimized. To this end, click reactions of varying ratios of protein to dye were performed for ACP2-pPrF and azide rhodamine and analyzed by SDS-PAGE.

![Click reaction between ACP2-pPrF and azide-rhodamine (AzR) with varying ratios of protein to dye.](image)

**Figure 116:** Click reaction between ACP2-pPrF and azide-rhodamine (AzR) with varying ratios of protein to dye. (A) SDS-PAGE analysis of click reactions. (B) MS analysis of the click reaction for sample with protein to dye ratio of 1 to 50. Experimental mass of ACP2-pPrF-AzR complex was 20984.18 Da, closely matching the expected mass 20984.18 Da.

From **Figure 116**, as the ratio of protein to dye increased, the visibility of the fluorescent band increased as well. The reaction was analyzed by MS which revealed the experimental mass closely matched the expected mass, indicating that the reaction was
successful. However, even after optimization, the reaction efficiency between ACP-pAzF and alkyne rhodamine was still much better than that between ACP-pPrF and azide rhodamine, thus, ACP-pAzF was selected to be the click partner for Ppant analog.

7.3.6 Synthesis of Ppant analog

Accordingly, the alkynyl Ppant analog was synthesized by Yan Zou based on Scheme 4. The overall yield was 54%. Identity of Ppant analog was confirmed by NMR and MS analysis (Table 28).

<table>
<thead>
<tr>
<th>Table 28: MS analysis for Ppant analog.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Calculated mass (Da)</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Ppant analog</td>
</tr>
</tbody>
</table>

7.3.7 Incorporation of pAzF into AcpP

To facilitate the subsequent enzymatic analysis of crypto-ACP, pAzF was incorporated into AcpP at the conserved Ser site, since all the other enzymes in *E. coli* fatty acid biosynthesis are available in our lab. MS analysis showed pAzF incorporation into AcpP was successful (Figure 117).
Figure 117: MS analysis of AcpP-S36pAzF. The experimental mass (10903.81 Da) matched closely to the expected mass (10903.864 Da).

7.3.8 CuAAC reaction between Ppant analog and pAzF

MS analysis of initial attempts for CuAAC between AcpP-S36pAzF and alkynyl Ppant analog failed to provide evidence that the expected reaction product was present. To find the best reaction conditions, small molecule CuAAC optimization was performed with the alkynyl Ppant analog and pAzF. However, no click product was detected by HPLC even after optimization (results not shown). One potential reason was the free thiol on the alkynyl Ppant analog interfering with the CuAAC reaction. For example, it is known that azide-independent polypeptide labeling occurs through thiol-yne addition with reduced peptidylcysteines. To address this possibility, iodoacetamide (IAM) was used to cap the free thiol on the alkynyl Ppant analog before click reaction (Scheme 5). Ppant-IAM mass was detected after 1 hour IAM treatment with good abundance (Table 29). However, a mass ion consistent with Ppant was also detected. Elongation of IAM treatment time (up to 5 h) led to decomposition of Ppant-IAM and no Ppant-IAM mass was detected. The Ppant analog after 1 hour IAM treatment was
used to click with pAzF. However, no product mass was detected in the Ppant-IAM and pAzF click reaction, presumably due to the incomplete capping of free thiol by IAM.

![Scheme 5: Capping the free thiol by iodoacetamide.](image)

**Table 29: MS analysis of Ppant-IAM.**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Calculated mass* (Da)</th>
<th>Experimental mass (Da)</th>
<th>Δ (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ppant-IAM</td>
<td>285.4</td>
<td>286.1</td>
<td>-0.7</td>
</tr>
</tbody>
</table>

![Figure 118: Structure for Ppant and trityl-protected Ppant analogs.](image)

To address this issue, a trityl-protected Ppant analog was used in the click optimization (Figure 118). Trityl-protected Ppant analog was afforded as an intermediate for the Ppant analog synthesis (Scheme 4). The CuAAC reaction between trityl Ppant and pAzF was optimized by varying a series of reaction parameters. Trityl Ppant analog didn’t dissolve well in water, thus, DMSO was added in various concentrations to the CuAAC reaction mixture to help with solubility (Figure 119A). Control assays confirmed the product peak to elute at 20.2
min by HPLC analysis (results not shown). As a compromise between trityl Ppant solubility, protein stability and reactivity, 10% (v/v) DMSO was used in future click reactions.

Various Copper sources were also tested (Figure 119B), CuSO₄ (99.999%) marginally showed the highest activity and was used in future reactions (results not shown). In addition, the concentration of Cu was also varied, highest product yield was observed in the presence of 0.1 mM CuSO₄.

To evaluate the effect of ligand on the click reaction, Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) and Tris(3-hydroxypropyltriazolylmethyl)amine (THPTA) were examined (Figure 119C and Figure 120). THPTA has better water solubility compared to TBTA and has been demonstrated to be a more efficient ligand than TBTA. Under the same condition, better product yield was observed when THPTA served as the ligand.

The influence of reducing agent on click reaction efficiency was also investigated (Figure 119D). Sodium ascorbate and TCEP were analyzed. No product was detected when TCEP was used as reducing agent.
Figure 119: Optimization of CuAAC reaction. Influence of DMSO% (A), [Cu$^{2+}$] (B), ligands (C) and reducing agents (D) on CuAAC reaction efficiency was evaluated.
Figure 120: Structures for TBTA and THPTA.

The click reaction between trityl Ppant and pAzF was performed under the identified optimal condition and was analyzed by MS. The experimental mass matched closely to the calculated mass of the clicked product.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Calculated mass* (Da)</th>
<th>Experimental mass (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trityl Ppant-pAzF</td>
<td>676.28</td>
<td>677.28(^1)</td>
</tr>
<tr>
<td>Trityl Ppant</td>
<td>470.2</td>
<td>493.19(^2)</td>
</tr>
<tr>
<td>pAzF</td>
<td>206.08</td>
<td>207.09(^1)</td>
</tr>
</tbody>
</table>

7.3.9 CuAAC reaction between Ppant and AcpP-pAzF

Having identified the optimal reaction conditions for small molecule click chemistry, the CuAAC reaction was performed between trityl Ppant and AcpP-S36pAzF. Unfortunately, no product mass was detected by MS analysis. To verify if accessibility of pAzF on ACP was
a potential reason for this failure, pAzF was introduced in place of N24 - a position distal to the DSL motif of AcpP (Chapter 4, Section 4.3.2). Subsequent MS analysis of CuAAC reaction between trityl Ppant and AcpP-N24pAzF didn’t detect clicked product mass. Cumulatively, reactivity and accessibility of pAzF on AcpP was not perturbed based on click reactions of AcpP-S36pAzF with alkyne rhodamine (Section 7.3.2). Thus, nonspecific interaction between the Ppant analog and ACP might be one reason for the absence of click product in AcpP-pAzF and trityl Ppant click reaction, among other potential causes.

7.3.10 Small molecules based ketone-hydrazide reaction optimization

Although numerous efforts have been made in the CuAAC reaction, no product was detected in the protein click reaction. At the same time, a new bioorthogonal reaction was investigated. Ketone-hydrazide ligation was selected based on its mild reaction condition and high selectivity.\textsuperscript{150}

Acetyl-phenylalanine (AcF) and acetyl-lysine (AcK) were first investigated as surrogates for the ACP-linked ketone analog. Biotin hydrazide was selected as the reacting hydrazide partner. No product was detected for AcK and biotin hydrazide reaction. Conjugation product was detected for AcF and biotin hydrazide, however, the product yield was very low (results not shown).
**Scheme 6: Ketone-hydrazide reaction.** Reaction between biotin hydrazide and acetyl-phenylalanine is shown as example.

Ketone-hydrazide reaction parameters were varied to identify the optimal conditions. AcF and biotin hydrazide were used in the optimization reactions. The ketone-hydrazide reaction typically proceeds under acidic conditions. The lower the pH, the faster the reaction proceeds. To evaluate the effect of pH on reaction efficiency, pH 6.5 (10 mM phosphate, 50 mM NaCl), pH 6.2 (50 mM phosphate, 50 mM NaCl), pH 4.6 (50 mM sodium acetate) buffers were used. As expected, reactions proceeded more efficiently under lower pH, however, ACP’s precipitated out at pH 4.6 (results not shown). As a compromise between reactivity and protein stability, pH 6.2 was used in protein-based ketone-hydrazide reaction.

The reaction temperature was varied from room temperature, 30 °C to 37 °C. No product was detected when the reaction was performed at room temperature, and product was barely detectable when the reaction was performed at 30 °C. Good product yield was obtained at 37 °C (results not shown). Thus, ketone-hydrazide reaction was performed at 37 °C in future.

AcF and biotin hydrazide ratio was adjusted too. 1:1, 1:5, 1:10, 1:20 (AcF:biotin hydrazide) reaction ratio was analyzed. The product yield had a linear relationship with the AcF:biotin hydrazide ratio from MS analysis (**Figure 121**).
Figure 121: Effect of varying biotin hydrazide:AcF ratio on AcF and biotin hydrazide conjugation reaction. Reactions were analyzed by MS, product peak area was plotted against biotin hydrazide:AcF ratio.

Aniline has been demonstrated to be an efficient catalyst for aldehyde based hydrazone and oxime formation reactions. To test if aniline could serve as a catalyst in ketone-hydrazide reaction, 100 mM aniline was included in the reaction. Unfortunately, a remarkable decrease in product yield was observed in the presence of aniline, suggesting that instead of expediting the reaction, aniline actually hindered the ketone-hydrazide reaction.
Figure 122: Effect of aniline on AcF-biotin hydrazide conjugation efficiency. Product peak area from MS analysis was plotted. Presence of aniline resulted in a significant decrease in ketone-hydrazide product yield.

It was suggested that addition of organic solvent, such as 5% (v/v) acetonitrile (ACN) or 5% (v/v) dimethylformamide (DMF) could increase ketone-hydrazide conjugation efficiency. No improvement in product yield was observed in the presence of 5% ACN or 5% DMF for the biotin hydrazide and AcF conjugation reaction investigated here.
Figure 123: Effect of organic additives on AcF-biotin hydrazide conjugation efficiency. Product peak area from MS analysis was plotted. Addition of ACN and DMF to the ketone-hydrazide reaction mixtures resulted in product yield decrease.

Cumulatively, the optimized ketone-hydrazide reaction conditions are pH 6.2 (50 mM phosphate, 50 mM NaCl) reaction buffer, with the ratio of ketone:hydrazide being 1:10, and overnight incubation at 37 °C.

7.3.11 Incorporation of acetyl phenylalanine onto AcpP

Accordingly, acetyl phenylalanine was incorporated at the conserved Ser site on AcpP. Identity of the mutant proteins was verified by MS (Figure 124).
Figure 124: Incorporation of AcF into AcpP at S36 site. The experimental mass (10904.29 Da) matched closely to the expected mass (10905.55 Da).

7.3.12 Ketone-hydrazide reaction between AcpP-AcF and biotin hydrazide

The ketone-hydrazide reaction between AcpP-S36AcF and biotin hydrazide was performed. MS analysis detected the expected AcpP-S36AcF-hydrazide product, but the conversion efficiency was too low to be accurately quantified (Table 31).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Calculated mass* (Da)</th>
<th>Experimental mass* (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcpP-S36AcF-biotin hydrazide</td>
<td>11138.36</td>
<td>11138.68</td>
</tr>
</tbody>
</table>

7.3.13 Ketone-hydrazide reaction between AcpP-AcF and texas red hydrazide

To improve the conjugation efficiency of AcpP-AcF based ketone-hydrazide reaction, another round of optimization was performed with AcpP-AcF and texas red hydrazide (TRH) (LifeTechnologies, Inc.). Conjugation efficiency was analyzed by SDS-PAGE and
fluorescence imaging. Effect of buffer pH, TRH:AcpP-AcF ratio and organic additives were examined. 50 µM AcpP-S36AcF was used in all optimization reactions.

**Figure 125: Reaction optimization between AcpP-S36AcF and TRH.** The fluorescence intensity for each AcpP band on SDS-PAGE was plotted against the corresponding reaction condition.

In general, pH 6.2 buffer afforded higher product yield over pH 5.5 buffer. This was probably due to poor protein stability in the pH 5.5 buffer. In accordance with previous observation, in pH 6.2 buffer, AcpP-AcF labeling by TRH increased with increasing TRH:AcpP-AcF ratio, i.e. increasing texas red hydrazide concentration. Interestingly and in contrast to previous results, organic additives such as DMF and ACN significantly increased the labeling efficiency in pH 6.2 buffer. In conclusion, under pH 6.2, 10% DMF, 50 µM AcpP-S36AcF and 500 µM TRH constitutes the best reaction condition so far for ketone-hydrazide reaction. In future, this reaction condition will be used to introduce Ppant analogs onto ACP.
7.4 Conclusions

Unnatural amino acids with alkynyl, azido and ketone functionalities have been successfully incorporated into ACP. Small molecules with azido, alkynyl and hydrazide group have been “clicked” onto mutant ACP’s. CuAAC reaction between AcpP-pAzF and Ppant analog and ketone-hydrazide reaction have been extensively optimized, however, limited improvement was observed in CuAAC reaction efficiency, while optimal reaction condition was identified for AcpP-AcF and hydrazide reaction. The ketone-hydrazide reaction was identified to be better than the traditional CuAAC reactions and Ppant analogs with hydrazide functionality will be synthesized and introduced onto ACP via the ketone-hydrazide conjugation.
CHAPTER 8. Future work

8.1 Incorporating non-natural extender unit into kirromycin biosynthesis

Having established that KirACP10 from kirromycin could be engineered to be a robust substrate for KirCII, it would be interesting to determine if a non-natural extender unit could be introduced into kirromycin via the identified KirACP10 mutant. In this regard, transacylation activity/specificity of KirCI and KirCII towards KirACP10 mutant would be evaluated to establish the orthogonality of KirACP10 mutant. After establishing the orthogonality in vitro, the KirACP10 mutation will be introduced into the *Streptomyces collinus* Tü 365 genomic DNA. *In vivo* production of kirromycin analogs will be performed and analyzed. It is noteworthy that the modification is not limited to KirACP10, as it has been demonstrated that KirACP7-F56L (equivalent to the KirACP10-F51L mutation) showed similar level of improvement in the KirCII-catalyzed transacylation (results not shown), suggesting that this mutation represents a common feature for improving KirCII transacylation. Taken together, the site-specific modification of kirromycin scaffold is feasible by introducing corresponding KirACP mutation genetically.

8.2 Combinatorial biosynthesis of polyketide via trans-AT swapping

Being a hybrid peptide-polyketide natural product, leinamycin has displayed potent antitumor activity, especially against some cell lines that are resistant to clinical anticancer drugs. Leinamycin synthase from *Streptomyces atroolivaceus* S-140 is a hybrid NRPS-PKS system, and no cis-AT is encoded in the PKS gene cluster Instead, the malonyl-CoA
extender unit is incorporated into the pathway via a trans-AT encoded by the LnmG gene (Figure 127).\textsuperscript{86} Previously, FluAT1 has displayed broad ACP specificity in the transacylation assay, especially for PKS ACP’s. The hypothesis was that leinamycin ACP’s could be recognized by FluAT1 and by swapping out LnmG with FluAT1 in the leinamycin biosynthesis, non-natural extender units could be incorporated. Further modification of the leinamycin scaffold is possible thanks to the chemical handles on some extender units such as azidoethylmalonyl-CoA and propargylmalonyl-CoA. Such structural diversification could potentially broaden the biological activity of leinamycin. Noticeably, this trans-AT swapping approach can be applied to any trans-AT system, as long as the ACP’s could be recognized or engineered to be recognized by FluAT1.
Figure 126: **Biosynthesis of leinamycin.** The biosynthetic contribution of LnmG (malonyl-CoA, red) to the final structure of leinamycin is highlighted.

8.3 Expanding the conventional repertoire of fatty acid biosynthesis

The demonstration of *in vitro* fatty acid production and discovery of AcpP mutants with remarkably improved activity in FluAT1 transacylation assay prompted the possibility that non-natural extender unit could be incorporated into the *E. coli* fatty acid biosynthetic pathway. The fatty acids obtained would exhibit impressive diversity in structures compared to
conventional fatty acids (Figure 127). For example, industrially valuable branched-chain fatty acids with higher solubility, improved oxidative stability, and low viscosity\textsuperscript{154} could be biosynthesized via this approach. In addition, orthogonal AcpP mutants could be engineered to facilitate the site-specific incorporation of non-natural extender units. There have been reports of using PKS systems for production of fatty acid analogs but not vice versa.\textsuperscript{155} The hypothesis here is that the reductive domains could be manipulated/deleted from fatty acid production on the premise that the downstream domains could process the unsaturated substrates, in this way, unsaturated fatty acids, or “polyketides” could be biosynthesized using the \textit{E. coli} fatty acid synthase and the type II FAS is then transformed to a type II PKS (Figure 128). This proposed approach could enable polyketides to be produced at immensely higher rate\textsuperscript{133,156} in addition to the bonus of maintaining protein-protein interactions.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{palmitic_acid.png}
\caption{Comparison of palmitic acid and palmitic acid analogs produced from conventional (upper) and engineered (lower) fatty acid biosynthesis.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{type_ii_pks.png}
\caption{Type II PKS constituted by \textit{E. coli} FAS proteins. FabH, B, F constitute the ketosynthase, FluAT1 represents the acyltransferase, FabG, I, A, Z represent the reductive domains.}
\end{figure}
8.4 Probe deacetylase involved interactions via unnatural amino acid based approach

Next generation sequencing has facilitated advancement of epigenetic study greatly over the past decade\textsuperscript{157} and cancer biology has been investigated more thoroughly and comprehensively than ever. It has long been appreciated that aberrant post-translational modifications, i.e. methylation, phosphorylation and ubiquitylation etc., are frequently observed in cancer biology and neurodegenerative diseases.\textsuperscript{158} Deacetylases are one of the many types of cell biology regulators and are known to process a variety of substrates and interactions between deacetylases and acetylated peptides have been probed via several approaches.\textsuperscript{159,160} Drugs targeting deacetylases have entered clinical trials for cancer therapy.\textsuperscript{158}

Sir2 deacetylases are highly conserved throughout bacterial, archaea and mammalian cells,\textsuperscript{161} and Cob\textsubscript{B}\textsuperscript{162} is the counterpart of eukaryotic Sir2 family deacetylase in \textit{E. coli}. The central hypothesis is that when placed at proper sites, Cob\textsubscript{B}-pBpa mutant could photocrosslink with its interacting acetylated proteins (\textbf{Figure 126}). Acetyl coenzyme A synthetase (Acs)\textsuperscript{163} and CheY\textsuperscript{164} are two of the reported substrates for Cob\textsubscript{B} and are selected as the interacting partners for investigation. Acs synthesizes acetyl-CoA from acetate, ATP and CoA, and is integral to prokaryotic and eukaryotic metabolism. CheY is the response regulator of bacterial chemotaxis.
**Figure 129**: Probe CobB involved interactions via unnatural amino acid based photocrosslinking approach. AcK can be site-specifically incorporated into CheY via unnatural amino acid mutagenesis by utilizing an evolved pyrrolysyl-tRNA synthetase/pylT pair. CobB pdb file: 1S5P, CheY pdb file: 3CHY.

The long-term goal is to better modulate deacetylase activity with the design of more efficient drugs based on knowledge that will be learned from deacetylases:protein interactions. The overall objective here is to explore substrate promiscuity of CobB and map the interaction epitope on CobB via unnatural amino acid mutagenesis. Knowledge gained from study of *E. coli* CobB could be extended to that of mammalian Sir2 deacetylases, which could hopefully explain how a protein is selected as target for deacetylation, especially during aberrant posttranslational modifications. This information should help design more efficient and specific drugs for cancer treatment.
REFERENCES


34. Mori, H.; Ito, K., Different modes of SecY-SecA interactions revealed by site-directed in vivo photo-cross-linking. P Natl Acad Sci USA 2006, 103 (44), 16159-16164.


47. Zhang, Y. M.; Marrakchi, H.; White, S. W.; Rock, C. O., The application of computational methods to explore the diversity and structure of bacterial fatty acid synthase. *Journal of lipid research* 2003, 44 (1), 1-10.


210


68. Qiu, X.; Janson, C. A., Structure of apo acyl carrier protein and a proposal to engineer protein crystallization through metal ions. Acta crystallographica. Section D, Biological crystallography 2004, 60 (Pt 9), 1545-54.


