UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery

Permalink

<https://escholarship.org/uc/item/7zx40314>

Author

Su, Michael

Publication Date

2019

Peer reviewed|Thesis/dissertation

Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery

By

Michael Su

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Wenjun Zhang, Chair
Professor Matthew Traxler
Professor David Schaffer

Spring 2019

© Copyright 2019
Michael Su
All Rights Reserved

Abstract

Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery

By

Michael Su

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor Wenjun Zhang, Chair

The alkyne functionality has attracted much interest due to its diverse chemical and biological applications. However, the underdeveloped alkyne biosynthetic toolbox severely limits the applications of this functionality in drug discovery and chemical biology. *De novo* biosynthesis is a recently developing and one of the most promising strategies for installing the terminal alkyne functionality into various molecular scaffolds, including diverse bioactive natural products, through the *in situ* generation and incorporation of the terminal alkyne functionality. We recently elucidated an acyl carrier protein (ACP)-dependent alkyne biosynthetic pathway comprised of JamABC from *Moorea producens* JHB that produces a terminal alkyne-tagged polyketide using hexanoic acid as a starting substrate for terminal alkyne generation. However, little is known about ACP interactions with the other alkyne biosynthetic enzymes in the JamABC pathway, including an acyl-ACP ligase (JamA) and a membrane-bound bi-functional desaturase/acetylenase (JamB). In this study, we demonstrated that JamB has a more stringent interaction with ACPs than JamA and subsequently pursued two different strategies to pinpoint key protein-protein interactions between JamB and several non-cognate ACPs. We first employed rational engineering of a non-cognate ACP (TtuC from *Teredinibacter turnerae* T7901) through site-directed mutagenesis, which significantly improved its compatibility with JamB and suggested a possible electrostatic interaction at the ACP-JamB interface. We also utilized an error-prone PCR and screening protocol towards a second non-cognate ACP (PeACP from *Pseudomonas entomophila* L48) to identify two hot spots on that ACP that are important for interacting with JamB and yielded mutants which were better recognized by JamB than the wild-type ACP. This study thus not only provides insights into the ACP interactions in alkyne biosynthesis, but it also potentially aids in future combinatorial biosynthesis of alkyne-tagged metabolites for chemical and biological applications.

Table of Contents

Abstract	1
Table of Contents	i
List of Figures	iii
List of Tables	iv
Acknowledgments	v
Chapter 1. Introduction	1
1.1 Significance of Terminal Alkynes	1
1.2 Terminal Alkyne Biosynthesis	4
1.3 Significance of Protein-Protein Interactions in Enzyme Biosynthetic Machinery	8
1.4 Motivations for Optimizing Protein-Protein Interactions in Terminal Alkyne Biosynthetic Pathways	10
Chapter 2. <i>In vivo</i> and <i>In vitro</i> Characterization of ACP Recognition by JamA and JamB	12
2.1 Introduction	12
2.2 Results	14
2.3 Discussion	18
2.4 Materials and Methods	19
Chapter 3. Improving ACP Recognition by JamB: Rational Engineering	24
3.1 Introduction	24
3.2 Results	24
3.3 Discussion	29
3.4 Materials and Methods	30
Chapter 4. Improving ACP Recognition by JamB: Error-Prone PCR and Screening	38
4.1 Introduction	38
4.2 Results	39
4.3 Discussion	45
4.4 Materials and Methods	46
Chapter 5. Conclusion	51

References	53
Appendices.....	57
A. SDS-PAGE analysis of proteins purified from <i>E. coli</i>	57
B. Deconvoluted masses of acyl-ACPs in HRMS analysis.	58
C. Relative JamA <i>in vitro</i> activities towards different ACPs.	60
D. HRMS of compounds 1-4	61
E. Relative JamB <i>in vivo</i> activities towards ACPs.	62

List of Figures

Figure 1-1. Overview of the various uses of alkyne tagging and the different biologically important molecules that can be labeled through alkyne tagging	1
Figure 1-2. Examples of terminal alkyne natural products with high bioactivity.....	3
Figure 1-3. The JamABC biosynthetic pathway for the C ₆ alkynoic starter unit generation and incorporation into jamaicamide B.....	4
Figure 1-4. Substrate specificity of JamB	5
Figure 1-5. The BesABCDEFG biosynthetic pathway for production of L-β-ethynylserine from L-lysine	6
Figure 1-6. Carrier proteins interact with different catalytic partners involved in fatty acid, polyketide, and non-ribosomal peptide biosynthesis	7
Figure 1-7. Abbreviated arrangements of PKS and NRPS to illustrate the key roles of CP domains towards each of these two synthases	8
Figure 2-1. <i>E. coli-hspks1</i> platforms for terminal alkyne biosynthesis	12
Figure 2-2. LC-HRMS chromatograms depicting the traces of the target terminal alkyne polyketides detected from the products of experimental <i>E. coli</i> strains	13
Figure 2-3. Early <i>in vivo</i> JamA loading assay data monitoring the formation of fatty acyl-JamC catalyzed by JamA, with hexanoic and 5-hexenoic acid as the substrate	14
Figure 2-4. JamA loading assay data towards the native ACP JamC and the foreign ACP TtuC	15
Figure 2-5. Relative JamB <i>in vivo</i> activities towards ACPs with significant JamA activity	18
Figure 3-1. Comparison of JamC and TtuC.....	25
Figure 3-2. Modelled coulombic surfaces of JamC and TtuC	25
Figure 3-3. Relative JamB <i>in vivo</i> activities towards TtuC mutant ACPs.....	26
Figure 3-4. Amino acid sequence alignments between JamC and either BpACP and PeACP	27
Figure 3-5. Relative JamB <i>in vivo</i> activities towards the PeACP and BpACP mutants created by rational engineering	28
Figure 4-1. Findings from applying the fluorogenic assay towards various controls in several types of experiments	40
Figure 4-2. Typical results from applying the fluorogenic assay towards a 96-well plate containing 88 PeACP mutants created from the error-prone PCR procedure	41
Figure 4-3. Relative JamB <i>in vivo</i> activities towards PeACP mutants. Error bars represent standard deviations from at least three experiments	41
Figure 4-4. The amino acid sequence alignment of PeACP and JamC	42
Figure 4-5. Modelled structures of PeACP and JamB	43
Figure 4-6. Relative JamA <i>in vitro</i> activities towards the wild-type PeACP and five PeACP mutants	43
Figure 4-7. The modelled coulombic surface of PeACP	44
Figure 4-8. Screening of the 44 potential AntD mutants	45

List of Tables

Table 2-1. DNA templates used for PCR amplification of ACP genes	16
Table 2-2. Average holo/apo ratios of the ACPs used in this study.....	17
Table 2-3. Plasmids used in this study	19
Table 2-4. Primers used in this study	20
Table 2-5. Strains used in this study	22
Table 3-1. DNA templates used for PCR amplification of ACP genes	30
Table 3-2. Plasmids used in this study	30
Table 3-3. Primers used in this study	32
Table 3-4. Strains used in this study	35
Table 4-1. DNA templates used for PCR amplification of ACP genes	46
Table 4-2. Plasmids used in this study	46
Table 4-3. Primers used in this study	47
Table 4-4. Strains used in this study	49

Acknowledgments

I would like to first of all thank Professor Wenjun Zhang, for her continual support and mentorship over the past five years of my graduate school experience, as well as her continual applications towards securing much of the funding towards the research detailed in this dissertation. It is because of her top-quality guidance that my scientific and experimental abilities were able to fully develop at the University of California, Berkeley. I sincerely appreciate her patient guidance and support towards furthering my academic, professional, and personal advancements. Without her, I would not have been able to achieve all of the work presented in this dissertation. I also would not have been able to learn the necessary skills for conducting world-class research beyond graduate school.

I would also like to thank my longtime mentor, Xuejun Zhu, for the extensive amount of effort that she made towards helping me learn all of the necessary techniques to carry out the projects that are described in this dissertation. She was also instrumental in helping me correct the many mistakes that I made in the lab while conducting the research that established the findings described in this dissertation, and I owe her a monumental degree of gratitude for this.

I also wish to credit the NIH for funding a great portion of this research as well as the entities that granted me the Chancellor's Fellowship for Graduate Study to fund my first three years of my research at the University of California, Berkeley.

I also wish to thank the members of the Deuber Lab for their assistance towards the use of their 96-well plate readers, which was of crucial importance for conducting the fluorogenic assay-based portions of this dissertation's research.

I would like to thank Carlet Altamirano for her overall guidance towards helping me make sufficient progress in the University of California's Ph.D. program.

I would like to thank all the current and former members of the Zhang Lab, most notably Joyce Liu who was also a patient, responsible, and crucial mentor during my first eighteen months in the lab. Nico Herman and Seong Jong Kim were also helpful in the early months of getting myself oriented and thus also deserve my thanks. I would also like to thank Frederick Twigg, Nick Harris, Will Skyrud, Jeffrey Li, Jeremy Seidel, and every other member of the Zhang Lab for their expertise and assistance in the many years of my tenure.

Finally, I would like to thank my father, Ching-Jong Su, for his long-distance moral support over the years in San Jose, as well as my sister, Anne Su, for her continuous assistance on many miscellaneous occurrences outside of the classroom and the lab. I also want to provide my greatest thanks to my late mother, whose death due to LAM lung disease in 2012 was a major catalyst for my personal drive towards conducting the type of research that is detailed in this dissertation.

Chapter 1. Introduction

Parts of this chapter have been adapted from the following with permission:

Su, M., Zhu, X. & Zhang, W. Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery. *AIChE Journal* **64**, 4255-4262 (2018).

1.1 Significance of Terminal Alkynes

The alkyne ($C\equiv C$) is a readily derivatized functionality valued for its diverse applications in nearly all areas of modern chemistry and biology.^{1,2} In particular, alkynes are widely employed in the alkyne-azide Huisgen cycloaddition reaction (known as a click reaction), which has allowed modular assembly of complex chemical structures for both drug development and material synthesis.³ Click chemistry also exhibits a key role towards bio-orthogonally tagging macromolecules with unique chemical handles that enable their visualization, enrichment, quantification, and mode of action study. Macromolecules that can metabolically or chemically install alkynes for these purposes include glycans, proteins, nucleic acids, lipids, and natural products (Figure 1-1). The mode of actions and cellular dynamics of these alkyne-tagged molecules can be further investigated through click chemistry to introduce an analytical handle,^{4,5} or directly using stimulated Raman scattering microscopy.⁶ These alkyne-tagged molecules can also incorporate ligands other than analytical handles for mutational biosynthesis, a function that is highly valued since many of them are structurally complex and have limited accessibility. We will limit the scope of our discussion of alkynes to their use in the production and development of natural products, which possess an irreplaceable role in the history of clinical and pharmaceutical research.

Figure 1-1. Overview of the various uses of alkyne tagging and the different biologically important molecules that can be labeled through alkyne tagging.

In the period between 1981 and 2014, studies have projected that ~67% of small-molecule anti-infective drugs and ~83% of anti-cancer small molecule treatments come from either natural products, their derivatives, or their mimics.⁷ However, transforming medicinally active natural products into effective drugs is difficult due to a lack of transforming technology for purification, diversification, and mode of action analysis of these compounds.⁸ One key problem is the time-consuming, labor-intensive preparation, and often inefficient process of obtaining natural products through biological fermentation.⁹ Because of this, sufficient quantities frequently cannot be acquired to characterize the chemical and biological properties of many active natural products. Another shortcoming comes from the complicated chemical syntheses which are heavily relied upon for natural product diversification also limit the types of natural product libraries that can be generated for drug screening.⁴ Organic synthesis often cannot synthesize these natural product-based libraries due to the structural complexity and chemical lability of these compounds. Finally, most therapeutically active natural products exert inhibitory functions on specific protein targets, the identification of which is essential for successful drug development. Unfortunately, these natural products frequently have weak specificities and affinities towards their protein targets, forcing the use of sensitive methods to identify these targets.⁹

Fortunately, click chemistry through the use of the alkyne-azide Huisgen cycloaddition reaction is an emerging technology that has addressed all of the problems discussed in the previous paragraph to varying degrees. Natural products can be tagged with clickable functionalities through various methods such as chemical synthesis, precursor-directed biosynthesis, and most recently, *de novo* alkyne biosynthesis. After tagging these natural products with clickable handles through one of these methods, azide or alkyne functionalized resins can be utilized to isolate and enrich these tagged natural products, significantly improving natural product yields for further compound discovery studies. Terminal alkynes also can greatly facilitate natural product diversification through their participation in mutational biosynthesis pathways, or mutasynthesis. By using alkyne substituted biosynthetic intermediates, commonly known as mutasynthons, and feeding them to genetically engineered producer strains, natural products can be generated that contain terminal alkyne groups that can easily be modified through click chemistry. One such example where such a diversification strategy yielded promising results is the Kirschning group's development of alkyne-modified ansamitocins by mutasynthetic supplementation of *Actinosynnema pretiosum* mutants with alkyne-substituted aminobenzoic acids.¹⁰ In this particular study, linker-based thiol groups were introduced to the aromatic moiety of ansamitocin P-3 using Huisgen-type cycloaddition to create tumor-targeting folic acid/drug conjugates. These conjugates selectively target the folate receptor (FR) in cancer tissues and bind to the FR with a high affinity ($K_d = 10^{-10}$ M), even when conjugated to a cytotoxin such as maytansin.^{11, 12} These thiol linkers that connect the drug to these tumor-specific ligands can be commonly designed such that a release mechanism of the cytotoxin is part of the conjugate's molecular architecture.¹³ Bioactivity tests conducted in the Kirschning study demonstrated that the ansamitocin derivatives containing either an alkyne or ester side chain possessed strong antiproliferative activity. Terminal alkynes can also help identify protein targets of medicinally active natural products by synthesizing alkyne-tagged analogs of these compounds and using them for activity-based protein profiling (ABPP) experiments. Upon binding to their protein targets, these analogs are then coupled to appropriate fluorescent reporters using click chemistry. Fluorescence screening protocols and further spectrometry analysis can then be used to identify the specific proteins that were targeted by these natural products. As a specific instance where alkynes have been used in ABPP studies, Sieber's group synthesized alkyne-tagged β -lactams as selective chemical probes to pinpoint the specific bacterial enzymes involved in

antibiotic resistance in *Staphylococcus aureus* strains.^{14, 15} Being stable enough to resist β -lactamase hydrolysis, these artificial lactam probes successfully detected and monitored the activities of several resistance-associated proteins through fluorescence scanning and mass spectrometry analysis.

Figure 1-2. Examples of terminal alkyne natural products with high bioactivity.

Finally, many natural products contain a terminal alkyne functionality which is crucial for bioactivity of these compounds (Figure 1-2).^{16, 17} One example is carmabin A, produced from the marine cyanobacterium *Moorea producens* (formerly classified as *Lyngbya majuscula*), which possesses antimalarial potency against the W2 chloroquine-resistant malaria strain ($IC_{50}=4.3 \mu M$).^{18, 19} Hantupeptin A from *Lyngbya majuscula* is another such potent compound, displaying brine shrimp toxicity at 10 ppm and strong cytotoxicity against the leukemia MOLT-4 cell line ($IC_{50}=32 \text{ nM}$).²⁰ Finally, caryoynencin, a polyene secondary metabolite from the carnation pathogen *Burkholderia caryophylli*, exhibits a potent antibacterial activity against a wide variety of infectious microorganisms, especially against several classes of *Trichophyton* fungi ($MIC=0.02 \mu g/ml$ for *Trichophyton interdigitale* and *Trichophyton rubrum*).^{21, 22} Considering the importance of alkynes in both chemistry and biology, a synthetic biology route for alkyne synthesis and installation is highly desired. In the next section, we will discuss such a route that we have developed over the past several years that does not encounter the drawbacks of other widely used methods of manufacturing terminal alkyne-tagged natural products for drug discovery and other relevant fields of research.

Figure 1-3. The JamABC biosynthetic pathway for the C₆ alkynoic starter unit generation and incorporation into jamaicamide B.

1.2 Terminal Alkyne Biosynthesis

In contrast to the broad applications of alkynes, few biosynthetic enzymes have been characterized for alkyne synthesis. This section will discuss the developments of several types of biosynthetic pathways capable of producing alkyne-tagged polyketides and alkyne-tagged non-ribosomal natural products, as well as the limitations of engineering these pathways for maximizing and diversifying these terminal alkyne products.

1.2.1 Biosynthesis of Terminal Alkyne as a Starter Unit for PKS

We recently elucidated an acyl carrier protein (ACP)-dependent terminal alkyne biosynthetic pathway in which a fatty acyl-ACP ligase activates a fatty acid using ATP and further attaches it to a cognate ACP, followed by the sequential terminal dehydrogenation of the fatty acyl moiety catalyzed by a membrane-bound bi-functional desaturase/acetylenase. The resulting ACP-bound terminal alkyne moiety is then incorporated by a promiscuous extracellular polyketide synthase (PKS) to synthesize a terminal alkyne-tagged polyketide. In particular, two homologous pathways, JamABC from *Moorea producula* JHB and TtuABC from *Teredinibacter turnerae* T7901, have been identified which activate hexanoic acid and decanoic acid, respectively. The first pathway, JamABC, consists of proteins expressed from the gene cluster identified to be responsible for the synthesis of jamaicamide B, a terminal alkyne-bearing PK/NRP hybrid from *Moorea producula* (Figure 1-3).^{23, 24} JamA, the fatty acyl-ACP ligase in this pathway, prefers loading C₆ fatty acids such as 5-hexenoic acid and 5-hexynoic acid to the JamC ACP over loading other saturated fatty acids the ACP-PP_i exchange assay.²³ However, JamA still loads fatty acids such as with heptanoic acid and butanoic acid to JamC with a moderate amount of activity. In our

first studies of reconstituting the JamABC pathway *in vitro*, however, we found that JamB, the pathway's bi-functional desaturase/acetylenase, has a high specificity for C₆ substrates that have a terminal alkene functionality.²⁵ Specifically, JamB cannot convert alkenoyl-JamCs other than 5-hexenoyl JamC to their corresponding alkynoyl-JamC forms. Also, JamB does not recognize hexenoyl-JamC with different unsaturation patterns such as 2-hexenoyl, 3-hexenoyl and 4-hexenoyl JamC. JamB can, however, recognize hexanoyl-JamC, but is only capable of converting that moiety to 5-hexenoyl-JamC (Figure 1-4). Having established that the JamABC system exclusively incorporates C₆ substrates for its starter units, we next investigated more than 80 gene operons across diverse bacterial genera that we identified as homologous to *jamABC* through genome-mining in an attempt to expand our scarce enzyme inventory of the alkyne biosynthetic toolbox. A phylogenetic analysis on these operons allowed us to select 11 representative *jamABC* cluster homologs from different clades for further biochemical analysis and subsequent *in vivo* reconstitution. These subsequent efforts allowed us to discover and characterize the second pathway, TtuABC system, which specifically recognizes C₁₀ fatty acids to produce a new type of terminal alkyne-bearing polyketide.²⁶ After characterizing the JamABC and TtuABC pathways, we further demonstrated that omitting the ACP, either JamC or TtuC in the respective pathway, completely abolished or significantly impaired alkyne synthesis,^{25, 26} illustrating the necessity of ACP in terminal alkyne biosynthesis.

Figure 1-4. Substrate specificity of JamB.

1.2.2 Biosynthesis of Terminal Alkyne-Containing Amino Acids

Developments have also been made towards expanding the alkyne biosynthetic toolbox to include distinct mechanisms. As an example, the Chang group recently characterized a unique pathway to produce a terminal alkyne-containing amino acid derived from *Streptomyces cattleya*.²⁷ This characterized pathway, BesABCDEF, takes L-lysine as a starting unit and utilizes halogenation, oxidative C–C bond cleavage, and triple bond formation through a putative allene intermediate to generate the terminal alkyne building block L-propargylglycine. The pathway then utilizes L-propargylglycine to synthesize the final terminal alkyne product, L- β -ethynylserine, with the help of glutamate and cellular hydrolases. Interestingly enough, the BesABCDEF pathway does not require an ACP to produce L- β -ethynylserine. Instead, the pathway's amino acid ligase,

BesA, connects the L-propargylglycine building block to glutamate to facilitate the production of the final terminal alkyne product (Figure 1-5). Even so, the necessary addition of a glutamyl group to L-propargylglycine for proper recognition by the hydroxylase BesE draws a parallel with the essential ligation of the ACP to fatty acids for proper desaturation of the fatty acids by either JamB or TtuB to occur within the JamABC and TtuABC pathways. The BesABCDE pathway could possibly be adopted for the production of halo-, alkene-, and alkyne-labelled proteins and natural products from glucose. The variety of downstream applications of these products include olefin metathesis or polymerization; halogen-dependent structural, functional, or physical property manipulation of key compounds; and click chemistry.

Figure 1-5. The BesABCDE biosynthetic pathway for production of L-β-ethynylserine from L-lysine.

Figure 1-6. Carrier proteins interact with different catalytic partners involved in fatty acid, polyketide, and non-ribosomal peptide biosynthesis. MAT: malonyl-CoA:ACP transferase; KS: ketosynthase; AT: acyl transferase; KR: ketoreductase; DH: dehydratase; ER: enoyl reductase; TE: thioesterase; C: condensation domain; A: adenylation domain.

1.2.3 Limitations of Engineering Terminal Alkyne Biosynthetic Pathways

Despite a variety of advancements towards producing terminal alkyne compounds, optimizing the underlying biosynthetic pathways is a difficult task. Although several of these pathways have been characterized to provide some insight into the order of tailoring reactions as well as the substrate specificity of participating proteins, utilizing this information to maximize expected product titers or diversify these pathways remains elusive. Much of the present inability to accomplish these critical goals of pathway engineering come from the lack of knowledge of the pathways' underlying protein-protein interactions. Attempts to swap biosynthetic enzymes with homologous proteins from other pathways often resulted in severely decreased or no production of the expected products, illustrating the momentous difficulty of engineering pathways of interest. Recently, heavy emphasis has been made towards understanding the interactions between the pathways' carrier proteins (CPs) and the other types of biosynthetic enzymes. In addition to being critical to alkyne biosynthesis, carrier proteins (CPs) in general are central hubs in the biosynthesis of fatty acids and medicinally important polyketide (PK) and non-ribosomal peptide (NRP) natural products²⁸ (Figure 1-6). In non-ribosomal peptide synthetases (NRPSs), which uses amino acids as building blocks, peptidyl carrier proteins (PCPs), often referenced as thiolation (T) domains, are analogous to a PKS's ACP domain. The main role of these PCPs involves tethering activated amino acid monomers activated by adenylation (A) domains through its 4'-phosphopantethiene arm. Condensation (C) domains catalyze peptide bond formation between peptidyl thioesters on the PCP domains, facilitating the incorporation of amino acid units into their respective growing natural products and further highlighting the importance of CPs in NRPSs. In both PKSs and NRPSs, the CP carries the bound final product to the terminal TE domain for the final product synthesis step, which involves intramolecular cyclization or hydrolytic release as a free acid (Figure 1-7). Thus, CPs play a major role in the production of a wide spectrum of clinically significant natural products, since CP-bound moieties are frequently the substrates that are

recognized by the pathways' downstream ketosynthase domains to synthesize the final product.²⁹ In the next section, we explore the different approaches towards understanding these crucial protein-protein interactions in biosynthetic machinery, especially those exhibited by ACPs.

Figure 1-7. Abbreviated arrangements of PKS and NRPS to illustrate the key roles of CP domains towards each of these two syntheses. The ACP domain is the CP domain in PKSs, while the T domain is the CP domain in NRPSs. The end products for the abbreviated arrangements for the PKS and NRPS illustrated here are 6-deoxyerythronolide B (Top) and tyrocidine A (Bottom), respectively. AT: acyl transferase; ACP: acyl carrier protein; KS: ketosynthase; KR: ketoreductase; TE: thioesterase; A: adenylation domain; T: thiolation domain; E: epimerization domain; C: condensation domain.

1.3 Significance of Protein-Protein Interactions in Enzyme Biosynthetic Machinery

Selective and programmed CP recognitions by multiple biosynthetic enzymes are essential for the precise assembly of complex natural product scaffolds, and combinatorial biosynthesis using simple mix-and-match strategy often yields significantly impaired assembly lines due to CP incompatibility. An example of this is a study by the Khosla group involving the use of a bimodular derivative of the 6-DEBS assembly line.³⁰ In that study, the Khosla group tested the influence of substrate-KS recognition on chimeric PKSs by replacing the PKS DEBS2 module of that

bimodular derivative with other PKS modules from the broader DEBS platform, as well as modules from the rifamycin and the rapamycin synthase assembly lines. Nearly all of the resulting chimeras had specific activities that below 10% of the reference bimodular derivative containing the DEBS2 module. The group then tested the replacement of the ketoreductase domain of the upstream module by a paralog that produced the enantiomeric ACP-bound diketide. The result was that changing the configuration of the diketide intermediate with the enantiomeric bound analog resulted in a 7-fold decrease in product turnover when applied to the reference bimodular derivative, but not for any of the developed bimodular chimeras. These results thus demonstrated that substrate recognition does not influence intermodular chain translocation to the same extent as protein-protein recognition does. Because combinatorial biosynthesis using simple mix-and-match strategy often yields significantly impaired assembly lines, there is thus a crucial demand to understand CP interactions with their catalytic partners.

Information about transient CP-enzyme interactions has emerged via X-ray crystallography, NMR, electron cryo-microscopy (cryo-EM) and crosslinking technology. For example, the Tsai and Burkhardt groups used X-ray crystallography and a sulfonyl-3-alkyne based probe that allows site-selective covalent crosslinking of the *E. Coli* ACP AcpP to the fatty acid 3-hydroxyacyl-ACP dehydratase FabA to solve the crystal structure of the crosslinked AcpP–FabA complex as a homo-dimer.³¹ In that study, the two groups discovered that AcpP has a negatively charged cleft between helices 2 and 3 that interacts with a positively charged arginine-rich patch on FabA. The two groups then identified residues at the AcpP–FabA interface using solution NMR techniques to provide an animated view of ACP in action during fatty acid dehydration, which primarily involved electrostatic interactions but also hydrophobic interactions from conserved residues. Combined with molecular dynamics simulations, the Tsai and Burkhardt groups demonstrated that FabA extrudes the sequestered acyl chain from the ACP binding pocket before dehydration by repositioning the ACP's helix 3. X-ray crystallography studies have also been conducted towards NRPS modules, such as the one conducted by Mahariel group towards solving the crystal structure of the 144 kDa termination module SrfA–C.³² In the solved crystal structure, the adenylation and condensation domains of SrfA–C associate closely to form a catalytic platform, with the PCP domain being flexibly tethered to this platform. The Mahariel group then proposed a model on how the PCP domain moves with its substrate-loaded 4'-phosphopantetheine arm between the adenylation domain's active site and the condensation domain's donor site to catalyze the addition of l-leucine to the hexa-peptide precursor as well as the cyclization of the resultant product to the lipoheptapeptide lactone surfactin, one of the most potent biosurfactants. NMR can also be employed to learn about transient CP-enzyme interactions when utilized in tandem with mutagenesis and mass spectrometry. The Dötsch group did just that to determine the structure of the ACP_I from the triplet ACP domains within the Curacin A PKS module as well as its interactions with the unique halogenase (Hal) domain embedded within the CurA module.³³ Standard NMR methods were used to solve the structure of the holo-ACP_I as well as the loaded form of ACP_I, which uses the 3-hydroxyl-3-methyl-glutaryl (HMG) moiety as a fundamental building block to synthesize the anti-proliferative compound Curacin A. However, NMR spectroscopy could not detect the interactions with HMG-ACP_I and Cur Hal due to the low population and transiency of the complex. Instead, the Dötsch group identified the interaction surfaces of holo-ACP_I and HMG-ACP_I using mutagenesis and MALDI-based product identification. The Smith and Skinotis groups utilized cryo-EM to determine the structures of a full-length PKS module from the pikromycin pathway of *Streptomyces venezuelae*.³⁴ The resulting structures revealed the ACP dynamics for sequential interactions with catalytic domains within the

reaction chamber responsible for carrying building blocks and intermediates between the acyltransferase, ketosynthase and ketoreductase active sites. The two research groups subsequently discovered a remarkable ACP localization dependent on the identity of the phosphopantetheine-linked polyketide intermediate or extender unit. Finally, the Burkhardt and Tsai groups developed mechanism-based crosslinkers for non-reducing polyketide synthases (NR-PKSs) in an attempt to understand ACP interactions with the product template (PT) domain that catalyzes polyketide cyclization.³⁵ By coupling mechanism-based crosslinking with *in vitro* assays, *in silico* docking, and bioinformatics, they identified the key ACP surface residues involved in ACP-PT recognition and utilized the subsequent findings to improve ACP compatibility with non-cognate PT domains. Protocols such as X-ray crystallography, NMR, cryo-EM, and crosslinking technology, thus provides a potential guide towards engineering ACPs to increasing their enzyme activities and improving their compatibility with other types of protein modules and mega-synthase domains. However, none of these methods at present help us provide a clear understanding on how ACPs selectively interact with biosynthetic enzymes such as those used to produce terminal alkyne products.

Despite the success of the protocols described in the previous paragraph towards probing transient CP-enzyme interactions, our understanding on the protein-protein interactions of ACPs towards alkyne biosynthetic enzymes, particularly the fatty acyl-ACP ligase and the bi-functional desaturase/acetylenase, is far from adequate. Preliminary studies of the substrate specificity of JamB, the acetylenase/desaturase from the JamABC pathway, implied strong protein-protein interactions between the acetylenase/desaturase and the dedicated ACP,²⁵ warranting the further investigation of the central role of ACP in alkyne synthesis. The flexibility of ACP in the alkyne biosynthetic machinery is particularly important for the application of these enzymes in engineered biosynthetic pathways to generate and incorporate this functionality into various PK/NRP scaffolds. This dissertation will try to extend our insights into ACP recognitions in terminal alkyne biosynthesis through protein engineering using JamABC as a model system.

1.4 Motivations for Optimizing Protein-Protein Interactions in Terminal Alkyne Biosynthetic Pathways

Despite the broad importance of the terminal alkyne functionality in modern chemistry and biology, much work still remains towards understanding and optimizing the biological routes to terminal alkyne compounds. Protein-protein interactions among the different types of catalytic domains of the biosynthetic pathways utilized to produce terminal alkynes are a dominant factor towards determining the pathways' overall production efficiencies. Unfortunately, these protein-protein interactions are not well understood, even after the characterization of several types of megasynthases.^{34, 36-39} What we do know is that the ACPs of the two terminal alkyne biosynthetic pathways that we previously characterized, JamABC and TtuABC, are the central hubs in the biosynthesis of their respective terminal alkyne polyketides. Elucidation of the protein-protein interactions exhibited by those ACPs will thus help augment the production of these terminal alkynes. Chapter 2 will discuss our selection of JamABC as the biosynthetic pathway to focus our protein-protein interaction studies on, our *in vivo* and *in vitro* characterization efforts of ACP recognition by JamA and JamB within the JamABC pathway, and how we concluded that the JamB-ACP interactions are the most stringent protein-protein interactions within JamABC. Having established that JamB-ACP interactions are the rate determining factor for terminal alkyne biosynthesis, we then turned our attention towards establishing which ACP amino acid residue

interactions are critical for JamB recognition. Chapter 3 discusses our applications of rational engineering towards improving the recognition of several non-cognate ACPs by JamB and the overall effectiveness of carrying out such attempts. Chapter 4 elaborates our use of an error-prone PCR and fluorogenic screening protocol toward improving the recognition of the non-cognate PeACP by JamB, and describes its potential to discover key ACP-JamB interactions that rational engineering could not reveal. Chapter 5 will discuss the future prospects of exploring protein-protein interactions within alkyne biosynthetic machinery and the broader impacts and advancements of doing so in various fields of chemistry and biology.

Chapter 2. *In vivo* and *In vitro* Characterization of ACP Recognition by JamA and JamB

Parts of this chapter have been adapted from the following with permission:

Su, M., Zhu, X. & Zhang, W. Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery. *AIChE Journal* **64**, 4255-4262 (2018).

2.1 Introduction

As mentioned in Chapter 1, we elucidated two homologous pathways, JamABC from *Moorea producens* JHB and TtuABC from *Teredinibacter turnerae* T7901, that respectively utilize hexanoic acid and decanoic acid as starting substrates to synthesize carrier protein-bound fatty acyl moieties with a terminal alkyne tag. We previously reported that a promiscuous type III polyketide synthase (PKS), HsPKS1, could incorporate these carrier protein-bound fatty acyl moieties into the polyketide backbone, and consequently generate terminal alkyne-tagged polyketides in engineered *E. coli* BAP1 strains that co-expressed *hspks1* and the genes encoding one of these two pathways (Figure 2-1).²⁵ However, little is known about ACP interactions with the other two types of alkyne biosynthetic enzymes, the acyl-ACP ligase (JamA, TtuA), and the membrane-bound bi-functional desaturase/acetylenase (JamB, TtuB).

Figure 2-1. *E. coli-hspks1* platforms for terminal alkyne biosynthesis. The balls/sticks on JamB and TtuB indicate that it is a transmembrane protein. In the JamABC+HsPKS1 pathway, the promiscuous type III polyketide synthase, HsPKS1, incorporates a carrier protein-bound fatty acyl moiety into the polyketide backbone, and consequently generates a terminal alkyne-tagged polyketide **1** in an engineered *E. coli* BAP1 strain that co-expressed *jamABC* and *hspks1*. The same mechanisms are utilized by HsPKS1 in the TtuABC+HsPKS1 pathway to generate a similar terminal alkyne-tagged polyketide an engineered *E. coli* BAP1 strain that co-expressed *ttuABC* and *hspks1*.

In deciding whether to focus on either the JamABC pathway or TtuABC pathway as a basis for ACP protein engineering, we considered each pathway's ability to accept foreign ACPs in place of the native ACP. If either pathway could still produce their corresponding terminal alkyne polyketides by a noticeable degree upon replacing the native ACP with a foreign ACP, the differences between the terminal alkyne titers produced from the resulting *E. coli* strains from ACP engineering would be smaller, making it difficult for compelling results to be made. To do this, we developed two experimental strains, XZ2 and XZ6, which either expressed *jamAB*, *ttuC*, and *hspks1*; or *ttuAB*, *jamC*, and *hspks1*. We then grew these strains in a 25 mL F1 media format, which was of a smaller scale than the 40 mL format utilized to evaluate the terminal alkyne production for the supermajority of engineered *E. coli* strains described in this dissertation. We then compared the alkyne titers of XZ2 and XZ6 to those achieved by from the strains expressing the entire biosynthetic pathway under the same scale as well as those lacking an expressed ACP. The results illustrated that TtuC had significant difficulty being recognized by JamA and JamB, as no traces of the terminal alkyne product could be found in XZ2 and the negative control. These findings also indicated that JamC had significant compatibility with TtuA and TtuB, since traces of the corresponding alkyne product could be found in all three strains, with the XZ6's LC-HRMS spectra having a moderate product peak (Figure 2-2). Thus, the JamABC+HsPKS1 system was quickly determined to be a more ideal system to work with for characterizing the alkyne biosynthetic enzymes and ACP engineering.

Figure 2-2. LC-HRMS chromatograms depicting the traces of the target terminal alkyne polyketides detected from the products of experimental *E. coli* strains. The target retention times for the product peaks are indicated as boxes for both sets of chromatograms. Top: Chromatograms with the JamABC terminal alkyne polyketide as the target. Bottom: Chromatograms with the TtuABC terminal alkyne polyketide as the target. The HRMS of the target compounds are found in Appendix D.

In this chapter, we elucidated the protein-protein interactions of JamA and JamB towards ACPs derived from a wide variety of biosynthetic pathways using a developed JamA loading assay and the full *E.coli-hspks1* platform, respectively. We then utilized the resulting findings to identify which of these non-cognate ACPs were possible candidates for further engineering in order to provide a better understanding of the key protein-protein interactions between the ACP and the other classes of alkyne biosynthetic enzymes.

2.2 Results

2.2.1 Developing the JamA Loading Assay

The acyl-ACP ligase JamA participates in the first step of the alkyne biosynthetic pathway by activating 5-hexenoic acid using ATP and loading it into the JamC ACP. This step is what allows the activated acid to be modified by JamB to form a terminal alkyne moiety through protein-protein recognition. Previous studies established that JamA has a high specificity for 5-hexynoic and 5-hexenoic acid against other types of fatty acids in the ATP-PP_i exchange assay.²³ To test the properties of JamA, we developed a loading assay that quantified the ability for the acyl-ACP ligase to activate starting acid substrates and ligate them to ACPs. In the first JamA loading assay studies, we discovered that the unsaturated hexanoic acid has a roughly equivalent activity with its alkene counterpart, 5-hexenoic acid, with a conversion of 0.290 vs. 0.257 respectively at the 5-minute time point (Figure 2-3). These results were consistent with findings from previous JamA and TtuA characterization studies.^{26, 40} All further applications of the JamA loading assay utilized 5-hexenoic acid as the loading acid substrate to retain consistency with previous studies.

Figure 2-3. Early *in vivo* JamA loading assay data monitoring the formation of fatty acyl-JamC catalyzed by JamA, with hexanoic and 5-hexenoic acid as the substrate.

After determining that JamA did not have any substrate preference towards 5-hexenoic acid or its unsaturated counterpart, we then sought to evaluate the acyl-ACP ligase's initial loading capability of 5-hexenoic acid with TtuC in comparison to the native JamC at different time points. This was done in order to establish a first impression of how stringent JamA's protein-protein interactions are with ACPs and to determine the ideal time point to evaluate JamA compatibility with ACPs. Applying the JamA loading assay towards JamC and TtuC, the ACP homolog from the TtuABC system, indicated that the JamA loading assay activity of TtuC was 3.6 times smaller than that of the native JamC after 20 minutes of activity, with even greater differences in magnitude of 5.3 and 4.3 at the 5 and 10 minute marks respectively (Figure 2-4). We thus selected the 5-minute reaction time to measure JamA loading activity, because the reaction rate noticeably decreased after that time point. The results also indicated that while there was a decrease in JamA activity by replacing JamC with TtuC, TtuC was still able to be loaded with 5-hexenoic acid by a considerable rate. As a reminder, the preliminary *in vivo* experiments illustrated in Figure 2-2

revealed that the XZ2 strain had almost no terminal alkyne production at a small scale after 48 hours of F1 media expression. Together, these results suggest that JamA-ACP interactions did not seem to be the limiting factor in terminal alkyne production. It is worth noting that XZ2's apparent incapability of producing terminal alkynes was later disproved when future experiments were conducted at the standard 40 mL F1 media scale. However, these later experiments still demonstrated that JamA-ACP interactions did not have as much influence in terminal alkyne production as JamB-ACP interactions did.

Figure 2-4. JamA loading assay data towards the native ACP JamC and the foreign ACP TtuC. Each of the two curves monitors the calculated conversion from the active holo ACP to the 5-hexenoyl ACP for JamC and TtuC, using the data acquired from eight JamA loading assays. The error bars (which represent the standard deviation) for each time point are included for each ACP.

2.2.2 Characterization of JamA

Having developed a JamA loading assay and determined the assay's optimal parameters for studying ACP recognition by the acyl-ACP ligase, we proceeded to analyze the degree of ACP recognition by JamA towards various foreign ACPs. To do this, we first selected 15 ACPs from different natural product biosynthetic pathways to study the ACP recognition by JamA, the fatty acyl-ACP ligase. The ACPs we tested included JamC, seven JamC homologs²⁶ (TtuC, BpACP, PeACP, CyACP, PfACP, IlACP, SpACP) found in putative alkyne biosynthetic machinery, and seven ACPs involved in the biosynthesis of erythromycin (DEBS1-ACPL),⁴¹ daptomycin (DptF),⁴² mycosubtilin (MycA-ACPL),⁴³ A54145 (LptF),⁴⁴ CDA (Sco3249),⁴⁵ isonitrile lipopeptide (ScoB),⁴⁶ and antimycin (AntD-ACP),^{47, 48} respectively. The detailed information of these ACPs is provided in Table 2-1. We first cloned each individual ACP-encoding gene into an expression vector that encoded a C-terminal His₆-tag. The ACPs were then overexpressed in the *E. coli* BAP1 strain to ensure their posttranslational modifications to holo forms⁴⁹ (Appendix A). The ratios of the holo and apo forms that coexist for each purified ACP used in the study are listed in Table 2-2. We next purified JamA from *E. coli* and tested its abilities to activate and load 5-hexenoic acid onto different ACPs using *in vitro* ACP loading assays, followed by liquid chromatography–high-resolution mass spectrometry (LC-HRMS) analysis (Appendix B). Approximately 25% of JamC was converted to 5-hexenoyl-JamC after 5 minutes of the *in vitro* reaction with JamA, and this conversion was set to be a relative JamA activity of 100%. The results

suggested that JamA has relaxed substrate specificities with ACPs from the putative alkyne biosynthetic machinery: ScoB and all of the seven tested JamC homologs could be recognized by JamA to a varying degree (Appendix C). However, most of the tested ACPs from PK/NRP biosynthetic machinery could not be efficiently recognized by JamA. In particular, four ACPs exhibited a high degree of compatibility with JamA (TtuC, CyACP, BpACP, and PeACP with JamA activities of ~37%, ~41%, ~207%, and ~217% respectively), and they were therefore chosen for the study of protein-protein interactions with the subsequent enzyme JamB, the membrane-bound bi-functional desaturase/acetylenase.

Table 2-1. DNA templates used for PCR amplification of ACP genes.

ACP Gene	Full name of ACP	IMG Gene ID	DNA Template
<i>dptF</i>	Daptomycin biosynthetic gene F_ACP	645226825	Genomic DNA: <i>Streptomyces roseosporus</i> NRRL 15998
<i>mycA-acpL</i>	Mycosubtilin A_ACP	2621327559*	Genomic DNA: <i>Bacillus subtilis</i> NRRL B-354
<i>lptF</i>	A54145 biosynthetic gene F_ACP	651720807*	Genomic DNA: <i>S. fradiae</i> NRRL18158
<i>sco3249</i>	<i>Streptomyces coelicolor</i> _ACP	637267498	Genomic DNA: <i>S. coelicolor</i> A3(2)
<i>debs1-acpL</i>	6-deoxyerythronolide B synthase_ACP	640171466	Plasmid: pBPJW144
<i>scoB</i>	<i>Streptomyces coeruleorubidus</i> biosynthetic gene B_ACP	Unlisted–Genbank Accession: KJ950369	Plasmid: pET24b-sco100
<i>antD-acp</i>	Antimycin D_ACP	2541761205*	Plasmid: pET30-antD
<i>pfacp</i>	<i>Pseudomonas fluorescens</i> Pf-5_ACP	637317607	Plasmid: pXZ72
<i>peacp</i>	<i>Pseudomonas entomophila</i> L48_ACP	638004365	Plasmid: pXZ77
<i>cyacp</i>	Cyanothece sp. ATCC 51142_3_ACP	641678257	Plasmid: pXZ78
<i>bpacp</i>	<i>Burkholderia pseudomallei</i> K96243_ACP	637571469	Plasmid: pXZ79
<i>ttuc</i>	<i>Teredinibacter turnerae</i> T7901_ACP	644917234	Plasmid: pXZ80
<i>ilacp</i>	<i>Inquilinus limosus</i> DSM 16000_ACP	2523927252	Plasmid: pXZ81
<i>spacp</i>	<i>Schlesneria paludicola</i> DSM 18645_ACP	2550737270	Plasmid: pXZ82

*The actual gene is encoded within the gene identified by the IMG Gene ID.

Table 2-2. Average holo/apo ratios of the ACPs used in this study.

ACP	Average Holo/Apo Ratio
JamC	1.61
TtuC	14.75
PeACP	2.08
BpACP	1.24
AntD-ACP	No <i>apo</i> -ACP detected
DEBS1-ACPL	25.83
Sco3249	0.15
DptF	No <i>apo</i> -ACP detected
MycA	No <i>apo</i> -ACP detected
LptF	No <i>apo</i> -ACP detected
ScoB	No <i>apo</i> -ACP detected
CyACP	0.38
PfACP	1.26
IlACP	49.26
SpACP	0.65

2.2.3 Characterization of JamB

The bi-functional JamB desaturase/acetylenase was the most important of the proteins from the JamABC+HsPKS1 biosynthetic pathway to characterize, as preliminary experiments from previous studies as well as those described in the previous section suggested that its ACP recognition interactions make up the rate-limiting factor of terminal alkyne production. However, membrane-bound proteins such as JamB are difficult to work with *in vitro*, compelling us to probe JamB's possible interactions with various ACPs *in vivo* using the full *E. coli-hspks1* platform described at the beginning of this chapter.

We previously reported that a promiscuous type III polyketide synthase (PKS), HsPKS1, could incorporate a carrier protein-bound fatty acyl moiety into the polyketide backbone, and consequently generate a terminal alkyne-tagged polyketide **1** in an engineered *E. coli* BAP1 strain that co-expressed *jamABC* and *hspks1*, which will be named as strain XZ1 (Figure 2-1 and Appendix D).²⁵ Since HsPKS1 demonstrates no preference towards different carrier protein-bound acyl moieties,^{26, 50, 51} this *E.coli-hspks1* platform could be a feasible reporting system for the study of JamB-ACP interactions. We introduced the ACP-encoding genes into the engineered *E. coli* BAP1 strain that co-expressed *hspks1*, *jamA*, and *jamB*. The resulting *E. coli* strains were grown in shake flasks at 20 °C for ~2 days with the supplementation of IPTG and 5-hexenoic acid. The ACP encoding genes that were incorporated into the separate *E. coli* BAP1 strains were the native *jamC* gene and the four genes that encoded the ACPs demonstrating a considerable degree of JamA activity (*ttuC*, *cyACP*, *bpACP*, and *peACP*). Upon the co-expression with *jamC*, *ttuC*, or *cyACP*, terminal alkyne-tagged polyketide **1** could be detected in the culture extracts by LC-HRMS analysis. In addition, the fed alkene precursor also led to a major byproduct **2** with a terminal alkenoyl modification (Figure 2-1). The compounds **1** and **2** were produced in an approximately 1:5 ratio by the strain XZ1,^{25, 50} and this efficiency was set to be a relative JamB activity of 100%. JamB activities towards TtuC and CyACP were ~39% and ~38%, respectively (Figure 2-5). No

reliable production of **1** was observed from the culture extracts of the strains that co-expressed *bpACP* or *peACP*, suggesting weak/no recognition of JamB with these two ACPs (Figure 2-5).

Figure 2-5. Relative JamB *in vivo* activities towards ACPs with significant JamA activity. The compounds **1** and **2** were produced in an approximately 1:5 ratio by the strain XZ1, and this efficiency was set to be a relative JamB activity of 100%. Error bars represent standard deviations from at least three experiments.

2.3 Discussion

In our first series of experiments involving the JamABC pathway, we characterized the degree of ACP recognition of the JamA acyl-ACP ligase and JamB bifunctional desaturase/acetylenase. The *in vitro* characterization studies of JamA, which we carried out by utilizing a carefully designed JamA loading assay, revealed eight ACPs that were compatible with JamA by at least a marginal degree (7% of that of the native JamC). However, seven of the eight compatible ACPs are JamC homologs, suggesting that many of the necessary protein-protein interactions of JamA loading of 5-hexenoic acid loading are enabled by the structural similarities that these homologs share. It is not known why ScoB, the eighth compatible ACP, is somewhat compatible with JamA despite not being a JamC homolog. Possible clues might come from the fact that the ScoB ACP was involved in the biosynthesis of a family of isonitrile lipopeptides in a separate study that was conducted by our lab⁴⁶. These compounds possessed a isocyanide functional group, whose triple bond in one of its resonance structures might suggest hidden similarities between ScoB and JamC homologs involved in alkyne biosynthesis that would allow ScoB to have a moderate JamA activity of ~15% upon being tested by the JamA loading assay (Appendix C). Whatever the explanation for ScoB being the only non-JamC homolog with JamA activity, a parallel series of experiments revealed that ScoB could not be recognized by JamB due to the lack of compound **1** production upon co-expression of the corresponding gene with *hspks1*, *jamA*, and *jamB* (Appendix E). All of these findings advocate that being derived from a putative alkyne biosynthetic machinery pathway is a necessary condition for an ACP to be recognized by JamA, JamB, and HsPKS1 for terminal alkyne biosynthesis. All of the ACPs other than ScoB from

PK/NRP biosynthetic machinery either could not be efficiently recognized by JamA, suggesting that the structures of these proteins are too dissimilar to be recognized by the acyl-ACP ligase, which is somewhat expected due to the considerably different structures of the final products of the corresponding PK/NRP biosynthetic pathways when compared to the final product of the JamABC+HsPKS1 pathway, compound **1**.

While being a JamC homolog from putative alkyne biosynthetic machinery is a necessary condition for an ACP to be compatible with JamA, JamB, and HsPKS1 to produce the terminal alkyne polyketide compound **1**, it is not a sufficient condition for this compatibility to be achieved. This is because the ACP needs to be recognized by JamB, where the protein-protein interactions between JamB and the ACP are very precise and difficult to pinpoint using the currently analytical methods of doing so. Only two out of the seven JamC homologs with JamA activity, TtuC and CyACP, resulted in the production of compound **1** when incorporated in place of the native JamC ACP with the full *E.coli-hspks1* platform. Even if an ACP has an abnormally high JamA loading activity, it does not necessarily result in JamB recognizing that ACP, which was illustrated in the case of PeACP and BpACP, where neither ACP enabled the *E.coli-hspks1* platform to produce compound **1**, despite each of these two ACPs having a JamA loading activity of over 200%. These findings, along with the fact that the HsPKS1 is a promiscuous PKS, reinforce the conclusion that JamB exhibits more stringent protein-protein interactions with ACPs than JamA does. We subsequently worked to acquire a better understanding about the necessary protein-protein interactions that JamB must exhibit with the ACP in order for terminal alkyne biosynthesis to occur. To accomplish this goal, we engineered foreign ACPs that would result in higher titers of compound **1** when expressed with the rest of the *E. coli-hspks1* platform. The two engineering methods that we used to improve ACP recognition with JamB are rational engineering and error-prone PCR coupled with fluorogenic screening, both of which will be separately described in the following chapters.

2.4 Materials and Methods

2.4.1 Bacterial Strains and Plasmid Constructions

The full names of the ACP genes used in this study as well as the DNA templates used to PCR amplify them are listed in Table 2-1. Plasmid constructions were performed using standard protocols. Site-directed mutagenesis was performed using the standard QuikChange strategy using relevant templates. Plasmids were purified from *E. coli* XL1-Blue with Zyppy™ Plasmid Miniprep Kit and confirmed by DNA sequencing (UC Berkeley DNA Sequencing Facility). *E. coli* BAP1⁴⁹ and *E. coli* BL21 Gold (DE3) were used for protein expression. All primers used in this study were ordered from Integrated DNA Technologies. Complete lists of plasmids and primers used in this study are provided in Tables 2-3 and 2-4.

Table 2-3. Plasmids used in this study.

Plasmids	Derived From	Function
pXZ27	pCDFDuet-1	Coexpression of JamA and Hspks1
pXZ34	pETDuet-1	Expression of JamB
pXZ43	pETDuet-1	Coexpression of JamB and JamC
pXZ44	pETDuet-1	Expression of JamC
pXZ72	pET24b	Expression of PfACP

pXZ77	pET24b	Expression of PeACP
pXZ78	pET24b	Expression of CyACP
pXZ79	pET24b	Expression of BpACP
pXZ80	pET24b	Expression of TtuC
pXZ81	pET24b	Expression of IlACP
pXZ82	pET24b	Expression of SpACP
pXZ92	pCDFDuet-1	Coexpression of TtuA and Hspks1
pXZ101	pETDuet-1	Coexpression of TtuB and TtuC
pXZ107	pETDuet-1	Coexpression of TtuB
pXZ218	pETDuet-1	Coexpression of JamB and TtuC
pXZ219	pETDuet-1	Coexpression of TtuB and JamC
pXZ239	pETDuet-1	Coexpression of JamB and DEBS1-ACP _L (C-His Tag)
pXZ241	pETDuet-1	Coexpression of JamB and Sco3249 (C-His Tag)
pXZ242	pETDuet-1	Expression of DptF (C-His Tag)
pXZ244	pETDuet-1	Expression of mycA
pXZ246	pETDuet-1	Expression of LptF
pXZ252	pETDuet-1	Expression of AntD-ACP
pMS17	pETDuet-1	Coexpression of JamB and ScoB (C-His Tag)
pMS32	pETDuet-1	Coexpression of JamB and PeACP
pMS33	pETDuet-1	Coexpression of JamB and CyACP
pMS42	pETDuet-1	Coexpression of JamB and BpACP

Table 2-4. Primers used in this study.

Primer	Sequence (5'→3')	Function
pXZ80_F_NdeI	aaacatATGACAACCTGTAGAAACACT	Overexpression of TtuC mutants in pET24b
pXZ80_R_XhoI	aaactcgagTTTCGCGGCAATAAATTGCG	
pXZ172_F_NcoI	aaaccatgggcATGACAACCTGTAGAAACACT	Expression of TtuC in pETDuet
pXZ172_R_PstI	aaactgcagTTTCGCGGCAATAAATTGCG	
pXZ235_F_NcoI	AAAccatgggcATGgactggaacgcgccaccgca	Overexpression of DEBS1-ACP _L
pXZ239_R_BamHI	AAAggatcctcagtgggtggtggtggtggtgCTCGAGggtcccgcgcgagtgtt	
pXZ252_NcoI	AAAccatgggcATGgggacggggctcgcgccgc	Overexpression of AntD-ACP
pXZ240_R_BamHI	AAAggatcctcagtgggtggtggtggtggtgCTCGAGgtcctcggcggttcgccgg	
pXZ241_F_NcoI	aaaccatgggcATGAGTACGGACCCCAA GTC	Overexpression of Sco3249
pXZ241_R_BamHI	AAAggatcctcagtgggtggtggtggtggtgCTCGAGCGCCGCTTCCAGACCCGAGC	

pXZ242_F_NdeI	aaacatATGAACCCGCCgGAAGCGGT CAGCACGCCCAG	Overexpression of DptF
pXZ242_R_XhoI	AAActcgagtcagtggtggtggtggtgACCA GAGGTGCGGTTCGGCCAACTGGG	
pXZ244_F_NdeI	AAAcatatgTCGACGGAACCGGTACA GAC	Overexpression of MycA
pXZ244_R_XhoI	AAActcgagtcagtggtggtggtggtgACCA GATTGCTTGATTTCGGAATGGT	
pXZ246_F_NdeI	AAAcataTGGCCTTCCCGCCgCCgTC CCCGCCgGCCAGCCAGGACCCg	Expression of LptF-C-his
pXZ246_R_XhoI	AAActcgagtcagtggtggtggtggtgACCA GAGCGCACGGTGC GCGAGGCCA	
pMS17_For	aaaccatgggcATGCCTGCTCCCTCAC GCT	Overexpression of ScoB
pMS17_Rev	aaaggatcctcagtggtggtggtggtgCTCGA GTGCGGTGACATGGCCCCGCTC	
pMS26_BamHI	aaaggatcctcagtggtggtggtggtg	Overexpression of PeACP
pMS27_NcoI	aaaccatgggcATGGATAAGAAGACAA CGAT	
pMS42_NcoI	aaaccatgggcATGAAGAGCGCCACCTT GAA	Overexpression of BpACP
pMS42_BamHI	aaaggatcctcagtggtggtggtggtgctcagTC GTCCGGCCGACAGCTCGG	

*Restriction sites are underlined.

2.4.2 Protein Expression and Purification

All ACPs purified in this work contained C-terminus hexahistidine tags and their expression plasmids were transformed into *E. coli* BAP1. The cells were grown at 37 °C in 750 mL of LB medium with appropriate concentrations of antibiotics to an OD₆₀₀ of 0.4–0.6, followed by the induction with 0.12 mmol/L isopropyl-β-d-thiogalactopyranoside (IPTG) for 16 h at 16 °C. The cells were harvested by centrifugation (4424 g, 15 min, 4 °C), resuspended in 30 mL of lysis buffer (50 mmol/L HEPES, pH 8.0, 300 mmol/L NaCl, 10 mmol/L imidazole), and lysed by homogenization on ice. Cell debris was removed by centrifugation (15000 g, 1 h, 4 °C). Ni-NTA agarose resin was added to the supernatant (1.5 mL/L of culture), and the solution was nutated at 4 °C for 1 h. The protein–resin mixture was loaded onto a gravity flow column, and proteins were eluted with increasing concentrations of imidazole in buffer A (50 mmol/L HEPES, pH 8.0, 300 mmol/L NaCl). Purified proteins were concentrated and buffer exchanged into HEPES buffer (50 mmol/L HEPES, pH 8.0, 100 mmol/L NaCl) with Amicon Ultra centrifugal filters. The final proteins were flash-frozen in liquid nitrogen and stored at –80 °C. Protein concentrations were determined by NanoDrop. The fatty-acyl ACP ligase Jama was expressed and purified as previously described.²⁵

2.4.3 In vitro ACP Loading Assays

Assays were performed in 50 μL of 50 mmol/L HEPES (pH 8.0) containing 2 mmol/L MgCl₂, 2 mmol/L ATP, 1 mmol/L TCEP, 5 mmol/L fatty acids, and 50–100 μmol/L holo-ACP. The

reactions were initiated by the addition of 10 $\mu\text{mol/L}$ JamA and incubated at room temperature. At regular time intervals (5, 10, 20 min), 10 $\mu\text{mol/L}$ aliquots were quenched with 10 μL of 10% (v/v) formic acid. The mixture was diluted two-fold, and centrifuged at 4 $^{\circ}\text{C}$ to remove any precipitate. 10 μL of the resulting protein sample was analyzed by liquid chromatography-high resolution mass spectroscopy (LC-HRMS) with an Aeris 3.6 μm widepore XB-C18 column (250 x 2.1 mm). A linear gradient of 15–98% CH_3CN (v/v) over 16 min and 98% CH_3CN for a further 20 min in water supplemented with 0.1% (v/v) formic acid at a flow rate of 0.15 mL/min was used for analysis. MassHunter Qualitative Analysis software was used for data analysis, and the intact protein masses were obtained using a maximum entropy deconvolution algorithm from Agilent MassHunter BioConfirm Software. The parameters used in this algorithm: a mass range of 9000–17000 Daltons, a mass step of 1.0 Dalton, a signal-to-noise peak threshold of 30, a proton-based adduct, a 90% average mass peak height, an automatic isotope width, a minimum consecutive charge state of 5, and a minimum protein fit score of 8.

2.4.4 Biosynthesis of Terminal Alkyne-tagged Polyketides in *E. coli*

E. coli BAP1 strains were transformed with plasmids containing *jamA*, *jamB*, *hspks1*, and the gene encoding the appropriate ACPs. A BAP1 strain co-expressing *jamABC* and *hspks1* was developed in each in vivo experiment as a positive control. Table 2-5 provides a complete list of developed *E. coli* BAP1 strains used in this study. The *E. coli* cells were grown in 200 mL of LB medium with 100 mg/L carbenicillin and 100 mg/L spectinomycin at 37 $^{\circ}\text{C}$ to an OD_{600} of 0.4–0.45. Subsequently, these cells were harvested and concentrated 5-fold into 40 mL of fresh F1 medium (1L contains 3 g KH_2PO_4 , 6.62 g K_2HPO_4 , 4 g $(\text{NH}_4)_2\text{SO}_4$, 150.5 mg MgSO_4 , 5 g glucose, 1.25 mL trace metal solution, 100 $\mu\text{mol/L}$ $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$, and 10 mmol/L 100x vitamin solution) supplemented with 100 mg/L carbenicillin, 100 mg/L spectinomycin, 0.6 mmol/L IPTG, and 1 mmol/L 5-hexenoic acid. After 48 hours of growth and expression at 20 $^{\circ}\text{C}$, compounds were extracted from cell-free supernatant with ethyl acetate. The solvent was removed by rotary evaporation, and the residue was redissolved in methanol and analyzed by LC-HRMS with an Agilent Eclipse Plus C18 column (4.6 $\times$ 100 mm). A linear gradient of 20–50% CH_3CN (v/v) over 15 min, 95% CH_3CN for a further 3 min, and 20% CH_3CN for a final 5 min in H_2O supplemented with 0.1% (v/v) formic acid at a flow rate of 0.5 mL/min was used for LC-HRMS analysis. LC-HRMS analysis was performed on an Agilent Technologies 6510 Accurate Mass QTOF LC-MS.

Table 2-5. Strains used in this study.

Strains	Genes	Plasmids
XZ1	HsPKS1, JamA, JamB, JamC	pXZ27, pXZ43
XZ2	HsPKS1, JamA, JamB, TtuC	pXZ27, pXZ218
XZ3	HsPKS1, JamA, JamB	pXZ27, pXZ34
XZ4	HsPKS1, TtuA, TtuB, TtuC	pXZ92, pXZ101
XZ5	HsPKS1, TtuA, TtuB	pXZ92, pXZ107
XZ6	HsPKS1, TtuA, TtuB, JamC	pXZ92, pXZ219
XZ7	HsPKS1, JamA, JamC	pXZ27, pXZ44
MS1	HsPKS1, JamA, JamB, PeACP	pXZ27, pMS32
MS2	HsPKS1, JamA, JamB, BpACP	pXZ27, pMS42
MS3	HsPKS1, JamA, JamB, CyACP	pXZ27, pMS33

MS4	HsPKS1, JamA, JamB, DEBS-ACPL	pXZ27, pXZ239
MS5	HsPKS1, JamA, JamB, ScoB	pXZ27, pMS17

Chapter 3. Improving ACP Recognition by JamB: Rational Engineering

Parts of this chapter have been adapted from the following with permission:

Su, M., Zhu, X. & Zhang, W. Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery. *AIChE Journal* **64**, 4255-4262 (2018).

3.1 Introduction

Due to the weak or no recognition of the JamB desaturase/actylenase with non-cognate ACPs, we next tried to test if protein engineering of ACPs could improve their interactions with JamB. The crystal structure of a soluble stearyl-ACP desaturase in complex with an ACP (PDB: 2XZ0 and 2XZ1) revealed predominant electrostatic interactions between the two proteins: ACP contributed negatively charged residues while the soluble stearyl-ACP desaturase displayed positively charged residues at the interaction surface.⁵² Despite the membrane-bound nature of JamB, we hypothesized that electrostatic interactions might contribute to JamB-JamC interactions. Our structural modeling of JamB and JamC and protein-protein docking analysis suggested electrostatic interactions at the JamB-JamC interface.⁵⁰ To test this hypothesis, we set out to rationally engineer TtuC, BpACP, and PeACP by changing their electrostatic surfaces in an attempt to improve their interactions with JamB. To carry out rational engineering for these ACPs, we utilized amino acid sequence alignments and comparisons between JamC and the non-cognate ACPs to identify which residues, upon mutation to appropriately charged residues, could restore the protein-protein interactions necessary for JamB recognition and subsequently terminal alkyne production to occur. We then incorporated the genes encoding these chimeric ACPs into the rest of the *E.coli-hspks1* platform to evaluate the resulting strains' alkyne titers, which served to evaluate how successful each amino acid mutation was towards improving each non-cognate's ACP's interactions with JamB.

3.2 Results

3.2.1 Rational Engineering of TtuC

Structural and sequence comparisons between TtuC and JamC revealed a positively-charged eight-amino acid region on TtuC helix 3 while the corresponding region of JamC lacks these positively-charged residues (Figures 3-1 and 3-2). Individual amino acids in this region of TtuC were then mutated to the corresponding JamC amino acids using site-directed mutagenesis. The resulting *ttuC* mutants were then introduced into the *E. coli-hspks1* system to investigate their effects on compound **1** production. The introduction of K69T into *ttuC* led to an engineered *E. coli* strain (named as strain XZ2_2) with the compound **1** production efficiency comparable to XZ1 (expressing *jamC*); while the other mutations had no obvious effect (Figure 3-3). Since the K69T mutation had no effect on the TtuC expression or recognition by JamA based on SDS-PAGE and biochemical analysis (Appendixes A and C), we concluded that the improved alkyne production efficiency by XZ2_2 was due to the improved interaction between TtuC-K69T and JamB.

Figure 3-1. Comparison of JamC and TtuC. Top: The amino acid sequence alignment of TtuC and JamC. The helical regions predicted from the modelled structures are indicated. The residues on TtuC that were mutated in this study are underlined. TtuC's K69 residue is indicated with an asterisk. Bottom: Modelled ribbon diagrams of JamC and TtuC. JamC-T83 and TtuC-K69 are illustrated as sticks.

Figure 3-2. Modelled coulombic surfaces of JamC and TtuC.

Figure 3-3. Relative JamB *in vivo* activities towards TtuC mutant ACPs. Top: Activities towards mutant ACPs for the individual TtuC-to-JamC residue mutations within the TtuC 8 amino acid region identified in Figure 3-1. Bottom: Activities towards mutant ACPs for different types of point mutations of TtuC's K69 residue. The compounds **1** and **2** were produced in an approximately 1:5 ratio by the strain XZ1, and this efficiency was set to be a relative JamB activity of 100%. Error bars represent standard deviations from at least three experiments.

To better understand the effect of amino acids with different properties at K69 of TtuC, we generated additional single mutants of TtuC, including K69R, K69E, K69Q, K69S, K69C, K69V, and K69Y, through site-directed mutagenesis. Most of these mutants exhibited improved alkyne production efficiency and as expected, TtuC-K69R had no significant improvement towards alkyne production compared to the wild-type TtuC (Figure 3-3). Out of all of the single mutants, the TtuC-K69E mutant had the highest alkyne production efficiency, with SDS-PAGE and biochemical analysis of this mutant indicating that this mutation also had no obvious effect on the

TtuC expression or recognition by JamA (Appendixes A and C). These findings supported the sensitivity of JamB activity towards the electrostatic surface of TtuC and the critical role of the residue K69 on the Helix 3 of TtuC for interacting with JamB. Using a similar engineering strategy, however, we failed to improve the recognition of JamB with BpACP and PeACP; it is likely that engineering the electrostatic interactions were not sufficient for proper interactions between JamB and these ACPs. We describe our attempt towards these two ACPs in the next section.

Figure 3-4. Amino acid sequence alignments between JamC and either BpACP and PeACP. The helical regions predicted from the modelled structures are indicated. For the foreign ACPs BpACP (top) and PeACP (bottom), that N- and C-termini sections that were replaced with homologous amino acid sections of the JamC ACP are underlined with dotted lines. The residues of each foreign ACP that were mutated in this study are underlined with solid lines.

3.2.2 Rational Engineering Attempts Towards PeACP and BpACP

Attempts to do so for these two ACPs began with following a similar strategy that was employed with the TtuC ACP. In the subsequent structural and sequence comparisons between these two ACPs and JamC, the N- and C-termini of these ACPs were first replaced with homologous sections of the JamC ACP (Figure 3-4). The resulting chimeric ACPs were not significantly more compatible than their corresponding wild-type ACPs. In the smaller scale *in vivo* experiments (conducted in a 25 mL F1 media format) that documented these findings, replacing sections of both the N- and C-termini with JamC homologous sections was necessary for any sort of JamB activity to occur. Even then, none of these engineered ACPs exceeded more than 2% of the activity of that of the native JamC, deeming such a widespread strategy useless for identifying possible ACP residues to mutate to establish JamB compatibility. We thus moved on to another approach by selecting PeACP and BpACP residues which had significantly different

charges than the corresponding homologous JamC residues for mutation using site-directed mutagenesis (Figure 3-4).

Figure 3-5. Relative JamB *in vivo* activities towards the PeACP and BpACP mutants created by rational engineering. The compounds **1** and **2** were produced in an approximately 1:5 ratio by the strain XZ1, and this efficiency was set to be a relative JamB activity of 100%. Error bars represent standard deviations from at least two experiments.

Overall, we developed a total 14 individual mutant ACPs from these rational engineering efforts, with 6 of them being PeACP mutants and the other 8 being BpACP mutants. When we tested them in the *in vivo E.coli-hspks1* system, none of these mutant ACPs demonstrated significant increases in JamB activity compared to their corresponding wild-type counterparts. For the case of PeACP, the *E. coli* strain containing the wild-type *peACP* (referred to as the MS1 strain) had 1.7% of the compound **1** titer that XZ1 exhibited. The strain containing the most active of the 6 rationally engineered PeACP mutant genes, *peACP_T55K*, exhibited 6.7% of XZ1's compound **1** titer, which resulted in only a fourfold improved activity when compared to MS1 (Figure 3-5). A possible explanation on why the T55K mutation was the most effective of the 6 PeACP mutations originates from residue 55's position between two PeACP regions where the amino acid sequence was identical to the homologous positions of JamC. Changing that residue to providing a larger region of sequence identity between the two ACPs could have resulted in improved JamB-PeACP interactions. Interestingly, 3 of the 4 PeACP mutations that led to JamB activities higher than that of the wild-type were mutations towards more positively charged residues, with only the H65Y mutation being the only amino acid change towards a more negatively charged residue. It also appears that the end of PeACP helix 3 was the most effective spot PeACP rational engineering, since two of the four effective mutations (E64S, H65Y) were located in that area. However, the improvement in JamB activity was not statistically significant enough to conclude that electrostatic

interactions was a significant factor in JamB-ACP compatibility. The engineered BpACP mutant strains exhibited even smaller compound **1** titers than the PeACP mutant strains, with only the F67K mutation standing out in JamB activity among the 8 mutants that were tested. Even then, the engineered strain containing the appropriate mutant gene had only 2.6% of the compound **1** titer found in XZ1. As a point of reference, this was only about a twofold improvement compared to the strain containing the wild-type *bpACP*, which had 1.0% of XZ1's compound **1** titer. The only conclusion that can be drawn from this data was that mutating the foreign ACP's residue that is homologous to JamC's K78 residue to a lysine residue seems to produce some improved activity. Even then, doing so does not lead to results that are nearly as effective as the rational engineering efforts towards the TtuC ACP.

3.3 Discussion

Following our rational engineering studies towards the TtuC ACP in an effort to improve its compatibility with the JamB desaturase/acetylenase, we discovered that a single amino acid mutation in TtuC caused the non-cognate ACP to be almost as compatible with JamB as the native JamC was. Through a further series of *in vivo* experiments with TtuC mutants, we discovered that replacing the positively charged residue in question with a negatively charged residue allowed TtuC to have even more JamB activity than JamC possessed. We were thus able to conclude that electrostatic interactions from that residue were a significant factor for improving TtuC's compatibility with JamB. However, using electrostatic interactions as a basis for rational engineering is not a universally viable approach, as was illustrated when similar approaches were applied towards engineering two other non-cognate ACPs. Instead, we were only able to tentatively conclude that mutating the residue that is homologous to JamC's K78 residue to a lysine residue seems to result in improved JamB activity for the case of PeACP and BpACP. Even this conclusion does not provide us with as much of a future direction towards improving ACP-JamB interactions that rationally engineering TtuC did.

Rationally engineering ACPs, and other types of proteins for that matter, comes with a major drawback, as producing effective results from this approach requires detailed knowledge of the protein structure, especially at the tertiary and quaternary level. In the context of this study, designing ACPs to increase terminal alkyne production in the *E. coli-hspks1* biosynthetic pathways requires ample information about the location of key interactions within the ACP-JamB interface as well as the contributing factors (electrostatic interactions, hydrophobic interactions, etc.) behind these protein-protein interactions. However, our initial knowledge about these two types of essential background information was limited towards these three non-cognate ACPs, especially PeACP and BpACP. This is what caused our applied changes towards PeACP and BpACP to fail for the most part by utilizing this approach. As a result, additional clues of which regions and underlying contributing factors to focus on often need to be obtained through alternate methods. PeACP provides an example of how one such alternate method, error-prone PCR and screening, can provide further clues to improve the effectiveness rational engineering. Rational engineering efforts towards PeACP with the information that we had so far, yielded ineffective results, with the most active mutant strain having only 6.7% of XZ1's compound **1** titer. Applying the error-prone PCR and screening protocol that we describe in the next chapter towards the PeACP gene, however, would yield mutant strains that produced compound **1** titers that was at least 13% of the titer achieved by XZ1. The location of the mutations in these effective ACP mutants, as well as the type of amino acid changes involved in these mutations, can help provide a template on which

factors to focus future rational engineering studies towards in an attempt to identify which protein-protein interactions are important in terminal alkyne biosynthetic machinery as well as the best approaches to optimize such interactions.

3.4 Materials and Methods

3.4.1 Bacterial Strains and Plasmid Constructions

The full names of the ACP genes used in this study as well as the DNA templates used to PCR amplify them are listed in Table 3-1. Plasmid constructions were performed using standard protocols. Site-directed mutagenesis was performed using the standard QuikChange strategy using relevant templates. Plasmids were purified from *E. coli* XL1-Blue with Zippy™ Plasmid Miniprep Kit and confirmed by DNA sequencing (UC Berkeley DNA Sequencing Facility). *E. coli* BAP1⁴⁹ and *E. coli* BL21 Gold (DE3) were used for protein expression. All primers used in this study were ordered from Integrated DNA Technologies. Complete lists of plasmids and primers used in this study are provided in Tables 3-2 and 3-3.

Table 3-1. DNA templates used for PCR amplification of ACP genes.

ACP Gene	Full name of ACP	IMG Gene ID	DNA Template
<i>peacp</i>	<i>Pseudomonas entomophila</i> L48_ACP	638004365	Plasmid: pXZ77
<i>cyacp</i>	Cyanothece sp. ATCC 51142_3_ACP	641678257	Plasmid: pXZ78
<i>bpacp</i>	<i>Burkholderia pseudomallei</i> K96243_ACP	637571469	Plasmid: pXZ79
<i>ttuc</i>	<i>Teredinibacter turnerae</i> T7901_ACP	644917234	Plasmid: pXZ80

Table 3-2. Plasmids used in this study.

Plasmids	Derived From	Function
pXZ27	pCDFDuet-1	Coexpression of JamA and Hspks1
pXZ34	pETDuet-1	Expression of JamB
pXZ43	pETDuet-1	Coexpression of JamB and JamC
pXZ44	pETDuet-1	Expression of JamC
pXZ80	pET24b	Expression of TtuC
pXZ218	pETDuet-1	Coexpression of JamB and TtuC
pXZ278	pETDuet-1	Coexpression of JamB and TtuC-I68P
pXZ279	pETDuet-1	Coexpression of JamB and TtuC-K69T
pXZ280	pETDuet-1	Coexpression of JamB and TtuC-S70I
pXZ281	pETDuet-1	Coexpression of JamB and TtuC-I71L
pXZ282	pETDuet-1	Coexpression of JamB and TtuC-R72Y
pXZ283	pETDuet-1	Coexpression of JamB and TtuC-K73S
pXZ284	pETDuet-1	Coexpression of JamB and TtuC-N74Y
pXZ285	pETDuet-1	Coexpression of JamB and TtuC-N75P

pXZ291	pETDuet-1	Expression of TtuC-I68P
pXZ292	pETDuet-1	Expression of TtuC-K69T
pXZ293	pETDuet-1	Expression of TtuC-S70I
pXZ294	pETDuet-1	Expression of TtuC-I71L
pXZ295	pETDuet-1	Expression of TtuC-R72Y
pXZ296	pETDuet-1	Expression of TtuC-K73S
pXZ297	pETDuet-1	Expression of TtuC-N74Y
pXZ298	pETDuet-1	Expression of TtuC-N75P
pMS26	pETDuet-1	Coexpression of JamB and the <i>N</i> -terminus region-modified PeACP (to JamC)
pMS27	pETDuet-1	Coexpression of JamB and the <i>C</i> -terminus region-modified PeACP (to JamC)
pMS28	pETDuet-1	Coexpression of JamB and the <i>N</i> and <i>C</i> -terminus region-modified PeACP (to JamC)
pMS32	pETDuet-1	Coexpression of JamB and PeACP
pMS42	pETDuet-1	Coexpression of JamB and BpACP
pMS43	pETDuet-1	Coexpression of JamB and the <i>N</i> -terminus region-modified BpACP (to JamC)
pMS44	pETDuet-1	Coexpression of JamB and the <i>C</i> -terminus region-modified BpACP (to JamC)
pMS45	pETDuet-1	Coexpression of JamB and the <i>N</i> - and <i>C</i> -terminus region-modified BpACP (to JamC)
pMS46	pETDuet-1	Coexpression of JamB and TtuC-K69E
pMS47	pETDuet-1	Coexpression of JamB and TtuC-K69R
pMS48	pETDuet-1	Coexpression of JamB and TtuC-K69S
pMS49	pETDuet-1	Coexpression of JamB and TtuC-K69V
pMS50	pETDuet-1	Coexpression of JamB and TtuC-K69C
pMS51	pETDuet-1	Coexpression of JamB and TtuC-K69Y
pMS52	pET24b	Expression of TtuC-K69E
pMS53	pET24b	Expression of TtuC-K69R
pMS54	pET24b	Expression of TtuC-K69S
pMS55	pET24b	Expression of TtuC-K69V
pMS56	pET24b	Expression of TtuC-K69C
pMS57	pET24b	Expression of TtuC-K69Y
pMS58	pETDuet-1	Coexpression of JamB and TtuC-K69Q
pMS59	pETDuet-1	Coexpression of JamB and PeACP-T34R
pMS60	pETDuet-1	Coexpression of JamB and PeACP-T73E
pMS61	pETDuet-1	Coexpression of JamB and BpACP-I45N/E46R
pMS62	pETDuet-1	Coexpression of JamB and BpACP-K56S
pMS63	pET24b	Expression of TtuC-K69Q
pMS64	pETDuet-1	Coexpression of JamB and PeACP-D29S
pMS65	pETDuet-1	Coexpression of JamB and PeACP-T55K
pMS66	pETDuet-1	Coexpression of JamB and PeACP-E64S

pMS67	pETDuet-1	Coexpression of JamB and PeACP-H65Y
pMS68	pETDuet-1	Coexpression of JamB and BpACP-G36E
pMS69	pETDuet-1	Coexpression of JamB and BpACP-P37E
pMS70	pETDuet-1	Coexpression of JamB and BpACP-D53A
pMS71	pETDuet-1	Coexpression of JamB and BpACP-E62G
pMS72	pETDuet-1	Coexpression of JamB and BpACP-F67K
pMS73	pETDuet-1	Coexpression of JamB and BpACP-S70D

Table 3-3. Primers used in this study.

Primer	Sequence (5'→3')	Function
pXZ80_F_NdeI	aaacatATGACAACCTGTAGAAACACT	Overexpression of TtuC mutants in pET24b
pXZ80_R_XhoI	aaactcgagTTTCGCGGCAATAAATTGCG	
pXZ172_F_NcoI	aaaccatgggcATGACAACCTGTAGAAACACT	Expression of TtuC in pETDuet
pXZ172_R_PstI	aaactgcagTTTCGCGGCAATAAATTGCG	
pMS26_BamHI	aaaggatcctcagtggtggtggtggtggtggt	Expression of PeACP in pETDuet
pMS27_NcoI	aaaccatgggcATGGATAAGAAGACAAACGAT	
pMS26_FInt1	agaggagaagaatggaccacagtagatgaggttcagagTGGCTGGTTGCCTATATCGC	Expression of PeACP mutants in pETDuet (<i>N</i> - and <i>C</i> -terminus modified mutants)
pMS26_FInt2	tagaaacccaaaacctgactagtttagataatggaagtatagaggagaagaatggacca	
pMS26_NcoI	aaaccatgggcacatggaaaacttaaccgtagaaacccaaacactgact	
pMS27_RInt	atcctcagtggtggtggtggtggtggtgctcgagtgcaccaaATGGGTAGCCAATGCCTCGA	
pMS42_NcoI	aaaccatgggcATGAAGAGCGCCACCTTGAA	Expression of BpACP in pETDuet
pMS42_BamHI	aaaggatcctcagtggtggtggtggtggtgctcgagTCGTCCGGCCGACAGCTCGG	
pMS43_Int	aacccaaaacctgactagtttagataatggaagtatagagGAAATCGAAGCAAGAAGCCA	Expression of BpACP mutants in pETDuet (<i>N</i> - and <i>C</i> -terminus modified mutants)
pMS44_Int	ggtggtgctcgagtgcaccaaagtgtctgcatagcttcGATCGTCGGATACTGATACG	
K69Q_For	GTGGCGATAGGTATTcaaTCCATCCGCAAAAAT	TtuC Site-Directed Mutagenesis for K69Q
K69Q_Rev	ATTTTTGCGGATGGAttgAATACCTATCGCCAC	
K69R_For	GTGGCGATAGGTATTcgtTCCATCCGCAAAAAT	TtuC Site-Directed Mutagenesis for K69R
K69R_Rev	ATTTTTGCGGATGGAacgAATACCTATCGCCAC	

K69E_For	GTGGCGATAGGTATTgaaTCCATCC GCAAAAAT	TtuC Site-Directed Mutagenesis for K69E
K69E_Rev	ATTTTTGCGGATGGAttcAATACCTA TCGCCAC	
K69S_For	GTGGCGATAGGTATTtctTCCATCCG CAAAAAT	TtuC Site-Directed Mutagenesis for K69S
K69S_Rev	ATTTTTGCGGATGGAagaAATACCT ATCGCCAC	
K69C_For	GTGGCGATAGGTATTtgtTCCATCCG CAAAAAT	TtuC Site-Directed Mutagenesis for K69C
K69C_Rev	ATTTTTGCGGATGGAacaAATACCT ATCGCCAC	
K69V_For	GTGGCGATAGGTATTgttTCCATCCG CAAAAAT	TtuC Site-Directed Mutagenesis for K69V
K69V_Rev	ATTTTTGCGGATGGAaacAATACCT ATCGCCAC	
K69Y_For	GTGGCGATAGGTATTtatTCCATCCG CAAAAAT	TtuC Site-Directed Mutagenesis for K69Y
K69Y_Rev	ATTTTTGCGGATGGAataAATACCTA TCGCCAC	
pMS59_For	GATGTTCCGTTcAGCcgtTTTGGCCT TGATTCG	PeACP Site-Directed Mutagenesis for T34R
pMS59_Rev	CGAATCAAGGCCAAAacgGCTGAAC GGAACATC	
pMS60_For	ATCGAGGCATTGGCTgagCATGCCT GGGAAGCC	PeACP Site-Directed Mutagenesis for T73E
pMS60_Rev	GGCTTCCCAGGCATGctcAGCCAAT GCCTCGAT	
pMS61_For	GTGAACAGGCAGTTCaatcgtTACGG CCTGGATTCTG	BpACP Site-Directed Mutagenesis for I45N/ E46R
pMS61_Rev	CGAATCCAGGCCGTAacgattGAACT GCCTGTTTAC	
pMS62_For	TCGGCCGATGCGATGtctATGGTCTG GCGATCTC	BpACP Site-Directed Mutagenesis for K56S
pMS62_Rev	GAGATCGCCGACCATagaCATCGCA TCGGCCGA	
pMS64_For	GCTGAGGTGGCACAGagcGTTCCGT TCAGCACC	PeACP Site-Directed Mutagenesis for D29S
pMS64_Rev	GGTGCTGAACGGAACgctCTGTGCC ACCTCAGC	
pMS65_For	TCCGACTGGCTCAGCaaaGAAATCG ACCCACCC	PeACP Site-Directed Mutagenesis for T55K
pMS65_Rev	GGTGGGGTTCGATTTcttGCTGAGCC AGTCGGA	

pMS66_For	CCCACCGTTGTCTAT <u>agt</u> CACCCCAC CATCGAG	PeACP Site-Directed Mutagenesis for E64S
pMS66_Rev	CTCGATGGTGGGGT <u>Gact</u> ATAGACA ACGGTGGG	
pMS67_For	ACCGTTGTCTATGAG <u>tac</u> CCCACCAT CGAGGCA	PeACP Site-Directed Mutagenesis for H65Y
pMS67_Rev	TGCCTCGATGGTGGG <u>gta</u> CTCATAG ACAACGGT	
pMS68_For	CTGCCCCGCCGACAGC <u>gaa</u> CCGATCG ACGTGAAC	BpACP Site-Directed Mutagenesis for G36E
pMS68_Rev	GTTACACGTCGATCGG <u>ttc</u> GCTGTCGG CGGGCAG	
pMS69_For	CCCGCCGACAGCGGC <u>gaa</u> ATCGACG TGAACAGG	BpACP Site-Directed Mutagenesis for P37E
pMS69_Rev	CCTGTTCACGTCGAT <u>ttc</u> GCCGCTGT CGGCGGG	
pMS70_For	GGCCTGGATTTCGGCC <u>gcc</u> GCGATGA AGATGGTC	BpACP Site-Directed Mutagenesis for D53A
pMS70_Rev	GACCATCTTCATCGC <u>ggc</u> GGCCGAA TCCAGGCC	
pMS71_For	ATGGTCGGCGATCTC <u>ggg</u> GACTATC TCGGCTTC	BpACP Site-Directed Mutagenesis for E62G
pMS71_Rev	GAAGCCGAGATAGTC <u>ccc</u> GAGATCG CCGACCAT	
pMS72_For	GAGGACTATCTCGGC <u>aaa</u> GAGCTGT CCCCGAGC	BpACP Site-Directed Mutagenesis for F67K
pMS72_Rev	GCTCGGGGACAGCTC <u>ttt</u> GCCGAGAT AGTCCTC	
pMS73_For	CTCGGCTTCGAGCTG <u>gac</u> CCGAGCC TGCCGTAT	BpACP Site-Directed Mutagenesis for S70D
pMS73_Rev	ATACGGCAGGCTCGG <u>gtc</u> CAGCTCG AAGCCGAG	

*Restriction sites are underlined.

3.4.2 Protein Expression and Purification

All ACPs purified in this work contained C-terminus hexahistidine tags and their expression plasmids were transformed into *E. coli* BAP1. The cells were grown at 37 °C in 750 mL of LB medium with appropriate concentrations of antibiotics to an OD₆₀₀ of 0.4–0.6, followed by the induction with 0.12 mmol/L isopropyl-β-d-thiogalactopyranoside (IPTG) for 16 h at 16 °C. The cells were harvested by centrifugation (4424 g, 15 min, 4 °C), resuspended in 30 mL of lysis buffer (50 mmol/L HEPES, pH 8.0, 300 mmol/L NaCl, 10 mmol/L imidazole), and lysed by homogenization on ice. Cell debris was removed by centrifugation (15000 g, 1 h, 4 °C). Ni-NTA agarose resin was added to the supernatant (1.5 mL/L of culture), and the solution was nutated at 4 °C for 1 h. The protein–resin mixture was loaded onto a gravity flow column, and proteins were eluted with increasing concentrations of imidazole in buffer A (50 mmol/L HEPES, pH 8.0, 300

mmol/L NaCl). Purified proteins were concentrated and buffer exchanged into HEPES buffer (50 mmol/L HEPES, pH 8.0, 100 mmol/L NaCl) with Amicon Ultra centrifugal filters. The final proteins were flash-frozen in liquid nitrogen and stored at -80°C . Protein concentrations were determined by NanoDrop. The fatty-acyl ACP ligase JamA was expressed and purified as previously described.²⁵

3.4.3 *In vitro* ACP Loading Assays

Assays were performed in 50 μL of 50 mmol/L HEPES (pH 8.0) containing 2 mmol/L MgCl_2 , 2 mmol/L ATP, 1 mmol/L TCEP, 5 mmol/L fatty acids, and 50–100 $\mu\text{mol/L}$ holo-ACP. The reactions were initiated by the addition of 10 $\mu\text{mol/L}$ JamA and incubated at room temperature. At regular time intervals (5, 10, 20 min), 10 $\mu\text{mol/L}$ aliquots were quenched with 10 μL of 10% (v/v) formic acid. The mixture was diluted two-fold, and centrifuged at 4°C to remove any precipitate. 10 μL of the resulting protein sample was analyzed by liquid chromatography-high resolution mass spectroscopy (LC-HRMS) with an Aeris 3.6 μm widepore XB-C18 column (250 x 2.1 mm). A linear gradient of 15–98% CH_3CN (v/v) over 16 min and 98% CH_3CN for a further 20 min in water supplemented with 0.1% (v/v) formic acid at a flow rate of 0.15 mL/min was used for analysis. MassHunter Qualitative Analysis software was used for data analysis, and the intact protein masses were obtained using a maximum entropy deconvolution algorithm from Agilent MassHunter BioConfirm Software. The parameters used in this algorithm: a mass range of 9000–17000 Daltons, a mass step of 1.0 Dalton, a signal-to-noise peak threshold of 30, a proton-based adduct, a 90% average mass peak height, an automatic isotope width, a minimum consecutive charge state of 5, and a minimum protein fit score of 8.

3.4.4 Biosynthesis of Terminal Alkyne-tagged Polyketides in *E. coli*

E. coli BAP1 strains were transformed with plasmids containing *jama*, *jamB*, *hspks1*, and the gene encoding the appropriate ACPs. A BAP1 strain co-expressing *jamABC* and *hspks1* was developed in each in vivo experiment as a positive control. Table 3-4 provides a complete list of developed *E. coli* BAP1 strains used in this study. The *E. coli* cells were grown in 200 mL of LB medium with 100 mg/L carbenicillin and 100 mg/L spectinomycin at 37°C to an OD_{600} of 0.4–0.45. Subsequently, these cells were harvested and concentrated 5-fold into 40 mL of fresh F1 medium (1L contains 3 g KH_2PO_4 , 6.62 g K_2HPO_4 , 4 g $(\text{NH}_4)_2\text{SO}_4$, 150.5 mg MgSO_4 , 5 g glucose, 1.25 mL trace metal solution, 100 $\mu\text{mol/L}$ $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2$, and 10 mmol/L 100x vitamin solution) supplemented with 100 mg/L carbenicillin, 100 mg/L spectinomycin, 0.6 mmol/L IPTG, and 1 mmol/L 5-hexenoic acid. After 48 hours of growth and expression at 20°C , compounds were extracted from cell-free supernatant with ethyl acetate. The solvent was removed by rotary evaporation, and the residue was redissolved in methanol and analyzed by LC-HRMS with an Agilent Eclipse Plus C18 column (4.6 x 100 mm). A linear gradient of 20–50% CH_3CN (v/v) over 15 min, 95% CH_3CN for a further 3 min, and 20% CH_3CN for a final 5 min in H_2O supplemented with 0.1% (v/v) formic acid at a flow rate of 0.5 mL/min was used for LC-HRMS analysis. LC-HRMS analysis was performed on an Agilent Technologies 6510 Accurate Mass QTOF LC-MS.

Table 3-4. Strains used in this study.

Strains	Genes	Plasmids
XZ1	HsPKS1, JamA, JamB, JamC	pXZ27, pXZ43
XZ2	HsPKS1, JamA, JamB, TtuC	pXZ27, pXZ218

XZ2_1	HsPKS1, JamA, JamB, TtuC-I68P	pXZ27, pXZ278
XZ2_2	HsPKS1, JamA, JamB, TtuC-K69T	pXZ27, pXZ279
XZ2_3	HsPKS1, JamA, JamB, TtuC-S70I	pXZ27, pXZ280
XZ2_4	HsPKS1, JamA, JamB, TtuC-I71L	pXZ27, pXZ281
XZ2_5	HsPKS1, JamA, JamB, TtuC-R72Y	pXZ27, pXZ282
XZ2_6	HsPKS1, JamA, JamB, TtuC-K73S	pXZ27, pXZ283
XZ2_7	HsPKS1, JamA, JamB, TtuC-N74Y	pXZ27, pXZ284
XZ2_8	HsPKS1, JamA, JamB, TtuC-N75P	pXZ27, pXZ285
XZ2_9	HsPKS1, JamA, JamB, TtuC-K69E	pXZ27, pMS46
XZ2_10	HsPKS1, JamA, JamB, TtuC-K69R	pXZ27, pMS47
XZ2_11	HsPKS1, JamA, JamB, TtuC-K69S	pXZ27, pMS48
XZ2_12	HsPKS1, JamA, JamB, TtuC-K69V	pXZ27, pMS49
XZ2_13	HsPKS1, JamA, JamB, TtuC-K69C	pXZ27, pMS50
XZ2_14	HsPKS1, JamA, JamB, TtuC-K69Y	pXZ27, pMS51
XZ2_15	HsPKS1, JamA, JamB, TtuC-K69Q	pXZ27, pMS58
XZ3	HsPKS1, JamA, JamB	pXZ27, pXZ34
MS1	HsPKS1, JamA, JamB, PeACP	pXZ27, pMS32
MS1_12	HsPKS1, JamA, JamB, PeACP (<i>N</i> -terminus modified)	pXZ27, pMS26
MS1_13	HsPKS1, JamA, JamB, PeACP (<i>C</i> -terminus modified)	pXZ27, pMS27
MS1_14	HsPKS1, JamA, JamB, PeACP (<i>N</i> - and <i>C</i> -terminus modified)	pXZ27, pMS28
MS1_15	HsPKS1, JamA, JamB, PeACP-T34R	pXZ27, pMS59
MS1_16	HsPKS1, JamA, JamB, PeACP-T73E	pXZ27, pMS60
MS1_17	HsPKS1, JamA, JamB, PeACP-D29S	pXZ27, pMS64
MS1_18	HsPKS1, JamA, JamB, PeACP-T55K	pXZ27, pMS65
MS1_19	HsPKS1, JamA, JamB, PeACP-E64S	pXZ27, pMS66
MS1_20	HsPKS1, JamA, JamB, PeACP-H65Y	pXZ27, pMS67
MS2	HsPKS1, JamA, JamB, BpACP	pXZ27, pMS42
MS2_1	HsPKS1, JamA, JamB, BpACP (<i>N</i> -terminus modified)	pXZ27, pMS43
MS2_2	HsPKS1, JamA, JamB, BpACP (<i>C</i> -terminus modified)	pXZ27, pMS44
MS2_3	HsPKS1, JamA, JamB, BpACP (<i>N</i> - and <i>C</i> -terminus modified)	pXZ27, pMS45
MS2_4	HsPKS1, JamA, JamB, BpACP-I45N/E46R	pXZ27, pMS61
MS2_5	HsPKS1, JamA, JamB, BpACP-K56S	pXZ27, pMS62
MS2_6	HsPKS1, JamA, JamB, BpACP-G36E	pXZ27, pMS68
MS2_7	HsPKS1, JamA, JamB, BpACP-P37E	pXZ27, pMS69
MS2_8	HsPKS1, JamA, JamB, BpACP-D53A	pXZ27, pMS70
MS2_9	HsPKS1, JamA, JamB, BpACP-E62G	pXZ27, pMS71
MS2_10	HsPKS1, JamA, JamB, BpACP-F67K	pXZ27, pMS72

MS2_11	HsPKS1, JamA, JamB, BpACP-S70D	pXZ27, pMS73
--------	--------------------------------	--------------

3.4.5 Bioinformatic Analysis of Proteins

Amino acid sequence alignment of the ACPs was performed using the online program Clustal Omega.⁵³ Protein structures of JamB and the ACPs were modeled using the online program HHPred⁵⁴ on the basis of a membrane-bound desaturase SCD1 (PDB: 4ZYO/4YMK)^{55, 56} and CurA-ACP (PDB: 2LIU),³³ respectively. ACP-JamB models were created by docking the ACP to the electropositive surface of JamB near the substrate entrance using ClusPro.^{57, 58} The visualizations of each structure and model were created using the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco.⁵⁹

Chapter 4. Improving ACP Recognition by JamB: Error-Prone PCR and Screening

Parts of this chapter have been adapted from the following with permission:

Su, M., Zhu, X. & Zhang, W. Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery. *AIChE Journal* **64**, 4255-4262 (2018).

4.1 Introduction

As stated in the previous chapter, rational engineering is not always viable, since selecting the right residues to mutate often requires extensive knowledge of the ACP's structure and important contributing factors behind favorable ACP-JamB binding interactions. The experiments conducted in the previous chapter indicated the shortcomings of rational engineering when applied to the ACPs PeACP and BpACP. As a review, both of these non-cognate ACPs have very low JamB compatibility when each of them were used in place of the native JamC ACP in the JamABC+HsPKS1 alkyne biosynthetic pathway. Several different types of amino acid mutations were made towards PeACP and BpACP based upon electrostatic differences that certain residues of each ACP had towards the JamC residues at the corresponding homologous positions. The resulting mutants exhibited only minimal improvements in JamB compatibility, which was measured by the compound **1** titer of the corresponding *E. coli* strains that expressed these mutant ACP genes. An alternative engineering approach was thus needed to provide more guidance on which ACP regions were most likely to interact with JamB. In this chapter, we describe the development of an error-prone PCR and screening protocol to provide that more efficient engineering approach, and the protocol's application to evolve a heterologous ACP for improved recognition by JamB. One of the drawbacks of utilizing error prone PCR is that this protocol requires the generation of hundreds of mutants for a satisfactory number of hits to be acquired, which usually falls in the range of 10-15 successful amino acid mutations. The randomness of both the number and location of the mutations towards the target ACP gene also impedes the certainty of whether the locations and identity of the successful ACP mutations can provide enough of a trend to indicate which localized regions or contributing factors are the most important for optimizing JamB-ACP protein-protein interactions. Another problem also originates from the identified effective ACP mutants containing multiple amino acid mutations. Trying to locate which the most beneficial mutation from the group through further site-directed mutagenesis experiments can present significant difficulties due to possible synergistic interactions between the multiple mutated residues. Nevertheless, it is almost certain that enough information can be obtained to at least provide some future directions for ACP engineering if the error-prone PCR and screening protocol is properly applied to a mutant library of an appropriate size. This approach has previously been employed to rapidly improve the activity of chimeric assembly-line enzymes, including the recognition of non-cognate ACPs.^{60, 61} The proper parameters of executing this engineering strategy will be discussed in the following sections of this chapter.

4.2 Results

4.2.1 Selection of PeACP as the Subject ACP and Developing a PeACP Mutant Library

PeACP was chosen as the target for the error-prone PCR protocol since it could be readily loaded by JamA (Appendix C), but rational engineering failed to improve its interaction with JamB. We first constructed a library of PeACP variants by diversifying *peACP* (~250 bp) randomly using error-prone PCR such that each variant contains an average of two amino acid mutations. The mutant library of *peACP* was incorporated into a strain that co-expressed *jamA*, *jamB*, and *hspks1*. The strains expressing these mutants of PeACP were screened based on the titer of extracellular alkynes using a fluorogenic assay, which was previously established and employed for improving the activity of JamB.⁵⁰

4.2.2 Optimization of the Fluorogenic Assay

When conducting the fluorogenic assay to detect active ACP mutants, a ~4-fold difference in magnitude was typical between positive (JamC) and negative (wild-type PeACP) controls. This was in line towards the findings from the extensive and repeated applications of the fluorogenic assay in an earlier project involving the TtuABC+HsPKS1 pathway, where the average fluorescence ratio of the positive control to the negative control in that experiment ranged between 4 to 7 (Figure 4-1). A designed diagnostic experiment to evaluate the fluorogenic assay's ability to detect relative compound **1** titers produced from the mutant ACP strains revealed that applying the fluorogenic assay protocol to a strain that co-expressed *jamAC* and *hspks1* but not *jamB* (named as strain XZ7) resulted in a fluorescence signal that was ~20-fold weaker in magnitude than that from the XZ1 strain (Figure 4-1). This indicates that even a small amount of extracellular alkyne compound **1** was enough to elicit a noticeable fluorogenic response from the assay, as the MS1 strain expressing the wild-type *peACP* has 1.7% of the terminal alkyne production activity as the XZ1 strain expressing *jamC*. Because replacing the negative control strain type from the non alkyne-producing XZ7 to MS1 resulted in a considerable decrease from a 20-fold to a 4-fold difference in magnitude between the positive and negative controls, the wild-type for the subject ACP should have almost no JamB compatibility for the error-prone PCR protocol to be viable. Even with such a suitable subject like PeACP, a moderate need for repeated applications of the fluorogenic assay protocol towards entire groups of ACP mutant strains was encountered. This was because low compound **1** production from the XZ1 strains representing the positive controls sometimes occurred, resulting in an inability to detect active ACP mutants upon conducting the screening assay due to insufficient fluorogenic signal differences between the positive and negative controls. We also tested whether the fluorogenic response would be affected by whether the strain was grown from a frozen stock or from a fresh plate colony, as there were initial concerns whether terminal alkyne production would be severely hampered by growing the strains directly from the frozen stocks that the mutant library was stored to the small-scale format (each mutant was expressed in 300 μ L F1 media aliquots) that the screening assay protocol requires. The results from the diagnostic experiment suggested that there was no clear pattern on whether frozen or fresh stocks within the same strain type yielded a higher fluorescence response (Figure 4-1). We thus grew the PeACP mutant strains directly from the library's frozen stocks in our fluorogenic screening protocols to determine which PeACP mutants had higher JamB compatibility than the wild-type PeACP.

18993	3689
10900	2877
12844	2916
19938	3295
15438	2870
21256	2935
24542	3896

1655	1206	82	79	356	585	19576	307
2569	2329	86	81	444	435	16287	399
1579	1675	85	75	394	586	16626	333

Figure 4-1. Findings from applying the fluorogenic assay towards various controls in several types of experiments. The types of genes co-expressed in the mutant strains described below can be found in Table 4-4. Top: The results of applying the fluorogenic assay towards XZ4 strains which were fed different types of fatty acids. Cell-free supernatants from these strains were used in the fluorogenic assay reaction to detect the presence of terminal alkyne polyketides. The left column represents the positive control, as the strains were fed 9-decenoic acid, which resulted in the production of terminal alkyne compound **3**. The right column represents the negative control, as the strains were fed 7-octenoic acid, which did not result in any terminal alkyne product. Bottom: In the first six columns, cell-free supernatants from the strains in question were used in the fluorogenic assay reaction, while pure F1 media reaction mixtures were prepared for Columns 7-8. From left to right, each of the columns displayed fluorescence units exhibited from applying the assay to the following types of samples:

Column 1: XZ1 strain, in which the strain was derived from a fresh plate colony.

Column 2: XZ1 strain, in which the strain was derived from a frozen stock.

Column 3: XZ7 strain, in which the strain was derived from a fresh plate colony.

Column 4: XZ7 strain, in which the strain was derived from a frozen stock.

Column 5: MS1 strain, in which the strain was derived from a fresh plate colony.

Column 6: MS1 strain, in which the strain was derived from a frozen stock.

Column 7: Blank F1 media with 0.07 mM 5-hexynoic acid (a terminal alkyne precursor).

Column 8: Blank F1 media.

4.2.3 Applying the Fluorogenic Screening Assay to the PeACP Mutant Library

To define a mutant as worthy for further LC-HRMS confirmation, the mutant needed to exhibit at least 10% of the gain from the average fluorescence exhibited by the three MS1 strains tested (which represent the negative control) to the average fluorescence of the three XZ1 strains (which represent the positive control) that were tested in each 96-well plate that was tested by the fluorogenic assay. An example of the results of conducting fluorogenic assay screening towards part of the PeACP mutant library is illustrated in Figure 4-2. After screening the strains developed from our library of ~900 PeACP clones, 27 possibly active mutants were identified from the screening assay this way and were subjected to LC-HRMS confirmation by expressing these mutants in the full-scale *E. coli* biosynthetic platform as described in the section titled Biosynthesis of Terminal Alkyne-tagged Polyketides in *E. coli*.

<>	1	2	3	4	5	6	7	8	9	10	11	12
A	4035	3148	3521	3459	3516	3343	3551	3220	3686	3367	3885	3685
B	3930	3021	3368	3402	3049	3095	3117	3159	3385	2976	3826	3751
C	3545	3657	3102	3663	2939	3713	3556	4115	3220	2899	3576	3701
D	2947	3729	3142	3714	3106	3158	4039	3265	3619	3853	3730	3757
E	3304	2973	3489	3360	3207	3027	2852	2918	2890	3509	3455	3621
F	3001	2944	2926	3260	3240	2689	2772	3363	3788	3256	3018	3323
G	2877	3492	2929	3130	3070	2916	2811	3250	2670	3213	3287	3296
H	2974	3546	3056	3468	6550	8968	8931	2695	2869	3026	1221	3197

Figure 4-2. Typical results from applying the fluorogenic assay towards a 96-well plate containing 88 PeACP mutants created from the error-prone PCR procedure.

The controls are located on row H, with the following setup:

H5-H7: XZ1 strain

H8-H10: MS1 strain

H11-H12: XZ7 strain

PeACP mutant strains with significant fluorogenic gains from the wild-type MS1 (and thus represent possible JamB compatibility) are highlighted in yellow.

Figure 4-3. Relative JamB *in vivo* activities towards PeACP mutants. Error bars represent standard deviations from at least three experiments.

The subsequent LC-HRMS confirmation allowed us to identify six PeACP mutants (named as strains MS1_1 through MS1_6) with increased alkyne production efficiency (Figure 4-3). It is notable that although the best mutant strain only exhibited ~30% of the relative JamB activity compared to XZ1 (expressing *jamC*), it had ~16-fold improved activity compared to the strain expressing the wild-type *peACP*. Sequencing analysis of these six *peACP* mutants revealed that those from MS1_1, MS1_4, and MS1_6 contained one amino acid change, the mutants from MS1_3 and MS1_5 each had two mutations, and MS1_2's mutant possessed four amino acid mutations (Figure 4-3). Single mutants corresponding to several amino acid changes found in

MS1_3, MS1_5, and MS1_2 (except a few C-terminal residues) were constructed to evaluate the individual contributions of these mutations towards JamB recognition. None of these single mutants exhibited a significant positive effect (Figure 4-3), suggesting that some of these mutations have a synergistic effect on improving the relative JamB activity. It is worth noting that all five of the positive single mutations are either located right before PeACP helix 2 or at the beginning of PeACP helix 3, highlighting the hot spots on PeACP for protein engineering (Figures 4-4 and 4-5). SDS-PAGE and biochemical analysis of each of these five PeACP mutants revealed that none of these mutations significantly increased ACP expression or JamA activity (Appendix A and Figure 4-6), suggesting the improvements between the ACP and JamB recognition as the cause for the increase in compound **1** production. To better understand the molecular mechanism of these beneficial PeACP mutations, we mapped them to a modelled structure of the JamB-PeACP complex that was built using a previously reported method.⁵⁰ These residues were predicted to locate at the interface of the two proteins, but the beneficial mutations could not be easily explained by electrostatic interactions (Figures 4-5 and 4-7).

Figure 4-4. The amino acid sequence alignment of PeACP and JamC. The helical regions predicted from the modelled structures are indicated. The five PeACP residues as hot spots for providing JamB compatibility upon mutation are underlined.

Figure 4-5. Modelled structures of PeACP and JamB. Left: A modelled ribbon diagram of PeACP. The five PeACP residues as hot spots for providing JamB compatibility upon mutation are illustrated as sticks. Right: Modelled surface structure of the JamB-PeACP complex. The two hot spots of the JamB-PeACP binding interface that were identified in the error-prone PCR and screening are marked.

Figure 4-6. Relative JamA *in vitro* activities towards the wild-type PeACP and five PeACP mutants. When evaluating the activities of these mutants, the conversion of the wild-type PeACP to 5-hexenoyl-PeACP after 5 minutes of the *in vitro* reaction with JamA was set to be a relative JamA activity of 100%.

Figure 4-7. The modelled coulombic surface of PeACP. The ACP's orientation is the same as that of the ribbon diagram from Figure 4-5. Positively charged residues are portrayed on the surface as blue, while negatively charged residues are portrayed on the surface as red.

4.2.4 Limitations of the Error-prone PCR and Fluorogenic Screening Protocol

It is worth noting that the error-prone PCR and fluorogenic screening protocol should only be applied for ACPs that are incompatible with just one of the biosynthetic enzymes, which in this case, is JamB. If the ACP itself is not very soluble or incompatible with multiple proteins of the rest of the JamABC-Hspks1 biosynthetic pathway, then the engineering protocol discussed in this chapter will not be of much use here. This conclusion manifested itself in our attempts to apply the error-prone PCR and fluorogenic screening engineering method towards the AntD-ACP. AntD-ACP was originally chosen as an ACP subject because initial JamA loading assay data for that ACP indicated that AntD-ACP had high JamA activity. Later loading assay applications toward that ACP, though, revealed that AntD-ACP was actually incompatible with JamA, since very little traces of the loaded ACP were found in the resulting deconvolution spectra. However, since the disproving loading assay results were not available until much later, we mistakenly proceeded to create a mutant library of the AntD-ACP through error-prone PCR and apply the fluorogenic assay protocol to these generated ACP mutants. There were 44 AntD-ACP mutants out of the library of ~1000 mutants that created from error-prone PCR that initially seemed to have some promising activity upon a first round of application of the fluorogenic assay. A second round of fluorogenic assay screening, however, illustrated that none of these AntD-ACP mutant genes were actually capable of terminal alkyne production when co-expressed with *jamA*, *jamB*, and *hspks1*, as no mutants exceeded the requisite fluorescence signal gain from the wild-type AntD-ACP of at least 10% of the gain attained from the wild-type AntD-ACP to the native JamC (Figure 4-8). Later *in vivo* experiments towards the three AntD-ACP mutants with the highest fluorescence demonstrated that none of these mutant strains yielded any detectable amounts of compound **1**. It was concluded that AntD-ACP's low JamA compatibility and low solubility prevented any meaningful information to be obtained from these engineering efforts. These failures highlight that

trying to engineer ACPs to be compatible with the rest of the biosynthetic pathway using error-prone PCR and the fluorogenic screening assay requires that the subject ACP should be very soluble with an abundance of the *holo* form and have at least some JamA activity in order for fruitful results to be obtained.

<>	1	2	3	4	5	6	7	8	9	10	11	12
A	1989	2181	2484	2373	2013	1921	2144	1977	1996	1997	1993	2318
B	1727	1847	1927	1952	1855	1817	2023	1856	1757	2265	1876	2138
C	1669	1672	1585	1848	1910	2045	1989	2034	1946	1986	1968	2004
D	1757	1648	1723	1779	2027	1908	2133	1964	19670	1775	28	29

Figure 4-8. Screening of the 44 potential AntD mutants. The positive control is located at well D9, while the negative control is located at well D10.

4.3 Discussion

By applying the error-prone PCR and screening protocol towards the PeACP, we were able to identify two hot spots where important JamB-ACP protein-protein interactions occurred. We also discovered that the mutations on these two hot spots which cause the PeACP to be more compatible with JamB were not due to electrostatic interactions. Both of these conclusions were not able to be achieved from rational engineering of PeACP, as the only deductions that we were able to make from these were that the T55K mutation marginally improved JamB activity since that residue was between two amino acid regions that were conserved in the JamC sequence upon comparison between the two types of ACPs. Even co-expressing the *peACP_T55K* mutant gene along with *jamA*, *jamB*, and *hspks1*, however, resulted in a considerably smaller compound **1** titer than those achieved by instead expressing any of the 6 active PeACP mutant genes found through the error-prone PCR and fluorogenic assay screening protocol. These six PeACP mutants, yielded compound **1** titers that were two to four times as high as the titer achieved from PeACP_T55K, underscoring the overall effectiveness of such an engineering method when little is known about the obvious differences between the non-cognate ACP and JamC that could improve JamB recognition when corrected. In order for the error-prone PCR and screening protocol to work, however, the subject ACP should be readily soluble in the active *holo* form and have significant compatibility with JamA, as attempting to engineer non-cognate ACPs that do not match this criteria with this methods will not yield any active ACP mutants, as our unsuccessful attempts towards the AntD-ACP demonstrate.

Applying the error-prone PCR and fluorogenic screening protocol is also more time consuming and susceptible to errors. To illustrate the former obstacle, it took several weeks to develop the PeACP mutant strain library, two months to screen the entire library with the designed fluorogenic assay, and an additional month to test the 27 possibly active PeACP mutant strains under a full-scale 40 mL F1 media format. In contrast to executing this ACP engineering protocol, applying rational engineering to an ACP can produce conclusive results in as little as three weeks. The only time-consuming steps in rational engineering are designing the chimeric ACP genes, incorporating them into engineered strains that co-express *hspks1*, *jamA*, and *jamB*, and expressing these engineered strains. Another key shortcoming to using the engineering method described in this chapter notably originates from the occasional need to repeat the fluorogenic assay several times to the same ACP mutant batch due to the low fluorogenic signals exhibited from the cell-free supernatants of the tested strains, even from the XZ1 control strains. This irregular proneness for the fluorogenic assay to yield low signals to occur could possibly be caused by the small scale

of the protein expression, which can sometimes be unfavorable enough to severely hamper terminal alkyne production, or by the probes utilized in this assay being degraded due to excessive freeze-thaw cycles. Nevertheless, the error-prone PCR based engineering approach possesses great potential towards expanding the terminal alkyne biosynthetic toolbox that provides broad applications in synthetic and chemical biology. For instance, the location of the two hot spots of the PeACP that appear to be crucial for JamB interaction, as well the type of mutations that resulted in substantial improvements in terminal alkyne production, provides possible guidelines on which underlying factors are the most important in the protein-protein interactions between the ACP and the JamB bifunctional desaturase/acetylenase. Future rational engineering approaches that are based on the findings from the error-prone PCR experiments would have to be applied to PeACP and other suitable non-cognate ACPs before more definitive conclusions can be made on the best steps towards optimizing the production of alkyne tagged metabolites.

4.4 Materials and Methods

4.4.1 Bacterial Strains and Plasmid Constructions

The full names of the ACP genes used in this study as well as the DNA templates used to PCR amplify them are listed in Table 4-1. Plasmid constructions were performed using standard protocols. Site-directed mutagenesis was performed using the standard QuikChange strategy using relevant templates. Plasmids were purified from *E. coli* XL1-Blue with Zyppey™ Plasmid Miniprep Kit and confirmed by DNA sequencing (UC Berkeley DNA Sequencing Facility). *E. coli* BAP1⁴⁹ and *E. coli* BL21 Gold (DE3) were used for protein expression. All primers used in this study were ordered from Integrated DNA Technologies. Complete lists of plasmids and primers used in this study are provided in Tables 4-2 and 4-3.

Table 4-1. DNA templates used for PCR amplification of ACP genes.

ACP Gene	Full name of ACP	IMG Gene ID	DNA Template
<i>antD-acp</i>	Antimycin D_ACP	2541761205*	Plasmid: pET30-antD
<i>peacp</i>	<i>Pseudomonas entomophila</i> L48_ACP	638004365	Plasmid: pXZ77

*The actual gene is encoded within the gene identified by the IMG Gene ID.

Table 4-2. Plasmids used in this study.

Plasmids	Derived From	Function
pXZ27	pCDFDuet-1	Coexpression of JamA and Hspks1
pXZ43	pETDuet-1	Coexpression of JamB and JamC
pXZ44	pETDuet-1	Expression of JamC
pXZ77	pET24b	Expression of PeACP
pXZ92	pCDFDuet-1	Coexpression of TtuA and Hspks1
pXZ101	pETDuet-1	Coexpression of TtuB and TtuC
pXZ252	pETDuet-1	Coexpression of JamB and AntD-ACP (C-His Tag)
pMS32	pETDuet-1	Coexpression of JamB and PeACP
pMS75	pETDuet-1	Coexpression of JamB and PeACP-V26E

pMS76	pETDuet-1	Coexpression of JamB and PeACP-L37F (C-His Tag)
pMS77	pETDuet-1	Coexpression of JamB and PeACP-G44S
pMS78	pETDuet-1	Coexpression of JamB and PeACP-L53H
pMS79	pETDuet-1	Coexpression of JamB and PeACP-T60I (C-His Tag)

Note: All pMS32 mutants that were created from the error-prone PCR protocol had a C-His Tag.

Table 4-3. Primers used in this study.

Primer	Sequence (5'→3')	Function
pXZ252_NcoI	AA <u>Accatgggc</u> ATGgggacggggctcgccgc	Overexpression of AntD-ACP, Development of AntD-ACP Mutant Library
pXZ240_R_BamHI	AA <u>Aggatcctc</u> agtggtggtggtggtgCTCG AGgtcctcggcgggttcgccgg	
pMS26_BamHI	aaaggatcctcagtggtggtggtggtggt	Overexpression of PeACP, Development of PeACP Mutant Library
pMS27_NcoI	aa <u>accatgggc</u> ATGGATAAGAAGACAA CGAT	
pXZ172_F_NcoI	aa <u>accatgggc</u> ATGACAACCTGTAGAAAC ACT	Expression of TtuC in pETDuet
pXZ172_R_PstI	aa <u>actgcag</u> TTTCGCGGCAATAAATTG CG	
pMS75_For	ATCGATCAGGCTGAGgagGCACAGG ATGTTCCG	PeACP Site-Directed Mutagenesis for V26E
pMS75_Rev	CGGAACATCCTGTGCctcCTCAGCCT GATCGAT	
pMS76_For	TTCAGCACCTTTGGcttGATTCGGC TGCAGTC	PeACP Site-Directed Mutagenesis for L37F
pMS76_Rev	GACTGCAGCCGAATCaaaGCCAAAG GTGCTGAA	
pMS77_For	TCGGCTGCAGTCATCagcATGACGG GGGATTTG	PeACP Site-Directed Mutagenesis for G44S
pMS77_Rev	CAAATCCCCCGTCATgctGATGACT GCAGCCGA	
pMS78_For	GATTTGTCCGACTGGcacAGCACGG AAATCGAC	PeACP Site-Directed Mutagenesis for L53H
pMS78_Rev	GTCGATTTCCGTGCTgtgCCAGTCGG ACAAATC	
pMS79_For	ACGGAATCGACCCCatcGTTGTCT ATGAGCAC	PeACP Site-Directed Mutagenesis for T60I
pMS79_Rev	GTGCTCATAGACAACgatGGGGTCG ATTCCGT	

*Restriction sites are underlined.

4.4.2 Protein Expression and Purification

All ACPs purified in this work contained C-terminus hexahistidine tags and their expression plasmids were transformed into *E. coli* BAP1. The cells were grown at 37 °C in 750 mL of LB medium with appropriate concentrations of antibiotics to an OD₆₀₀ of 0.4–0.6, followed by the

induction with 0.12 mmol/L isopropyl- β -D-thiogalactopyranoside (IPTG) for 16 h at 16 °C. The cells were harvested by centrifugation (4424 g, 15 min, 4 °C), resuspended in 30 mL of lysis buffer (50 mmol/L HEPES, pH 8.0, 300 mmol/L NaCl, 10 mmol/L imidazole), and lysed by homogenization on ice. Cell debris was removed by centrifugation (15000 g, 1 h, 4 °C). Ni-NTA agarose resin was added to the supernatant (1.5 mL/L of culture), and the solution was nutated at 4 °C for 1 h. The protein–resin mixture was loaded onto a gravity flow column, and proteins were eluted with increasing concentrations of imidazole in buffer A (50 mmol/L HEPES, pH 8.0, 300 mmol/L NaCl). Purified proteins were concentrated and buffer exchanged into HEPES buffer (50 mmol/L HEPES, pH 8.0, 100 mmol/L NaCl) with Amicon Ultra centrifugal filters. The final proteins were flash-frozen in liquid nitrogen and stored at –80 °C. Protein concentrations were determined by NanoDrop. The fatty-acyl ACP ligase *JamA* was expressed and purified as previously described.²⁵

4.4.3 *In vitro* ACP Loading Assays

Assays were performed in 50 μ L of 50 mmol/L HEPES (pH 8.0) containing 2 mmol/L MgCl₂, 2 mmol/L ATP, 1 mmol/L TCEP, 5 mmol/L fatty acids, and 50–100 μ mol/L holo-ACP. The reactions were initiated by the addition of 10 μ mol/L *JamA* and incubated at room temperature. At regular time intervals (5, 10, 20 min), 10 μ mol/L aliquots were quenched with 10 μ L of 10% (v/v) formic acid. The mixture was diluted two-fold, and centrifuged at 4 °C to remove any precipitate. 10 μ L of the resulting protein sample was analyzed by liquid chromatography-high resolution mass spectroscopy (LC-HRMS) with an Aeris 3.6 μ m widepore XB-C18 column (250 x 2.1 mm). A linear gradient of 15–98% CH₃CN (v/v) over 16 min and 98% CH₃CN for a further 20 min in water supplemented with 0.1% (v/v) formic acid at a flow rate of 0.15 mL/min was used for analysis. MassHunter Qualitative Analysis software was used for data analysis, and the intact protein masses were obtained using a maximum entropy deconvolution algorithm from Agilent MassHunter BioConfirm Software. The parameters used in this algorithm: a mass range of 9000–17000 Daltons, a mass step of 1.0 Dalton, a signal-to-noise peak threshold of 30, a proton-based adduct, a 90% average mass peak height, an automatic isotope width, a minimum consecutive charge state of 5, and a minimum protein fit score of 8.

4.4.4 Biosynthesis of Terminal Alkyne-tagged Polyketides in *E. coli*

E. coli BAP1 strains were transformed with plasmids containing *jamA*, *jamB*, *hspks1*, and the gene encoding the appropriate ACPs. A BAP1 strain co-expressing *jamABC* and *hspks1* was developed in each in vivo experiment as a positive control. Table 4-4 provides a complete list of developed *E. coli* BAP1 strains used in this study. The *E. coli* cells were grown in 200 mL of LB medium with 100 mg/L carbenicillin and 100 mg/L spectinomycin at 37 °C to an OD₆₀₀ of 0.4–0.45. Subsequently, these cells were harvested and concentrated 5-fold into 40 mL of fresh F1 medium (1L contains 3 g KH₂PO₄, 6.62 g K₂HPO₄, 4 g (NH₄)₂SO₄, 150.5 mg MgSO₄, 5 g glucose, 1.25 mL trace metal solution, 100 μ mol/L Fe(NH₄)₂(SO₄)₂, and 10 mmol/L 100x vitamin solution) supplemented with 100 mg/L carbenicillin, 100 mg/L spectinomycin, 0.6 mmol/L IPTG, and 1 mmol/L 5-hexenoic acid. After 48 hours of growth and expression at 20 °C, compounds were extracted from cell-free supernatant with ethyl acetate. The solvent was removed by rotary evaporation, and the residue was redissolved in methanol and analyzed by LC-HRMS with an Agilent Eclipse Plus C18 column (4.6 x 100 mm). A linear gradient of 20–50% CH₃CN (v/v) over 15 min, 95% CH₃CN for a further 3 min, and 20% CH₃CN for a final 5 min in H₂O supplemented

with 0.1% (v/v) formic acid at a flow rate of 0.5 mL/min was used for LC-HRMS analysis. LC-HRMS analysis was performed on an Agilent Technologies 6510 Accurate Mass QTOF LC-MS.

Table 4-4. Strains used in this study.

Strains	Genes	Plasmids
XZ1	HsPKS1, JamA, JamB, JamC	pXZ27, pXZ43
XZ4	HsPKS1, TtuA, TtuB, TtuC	pXZ92, pXZ101
XZ7	HsPKS1, JamA, JamC	pXZ27, pXZ44
MS1	HsPKS1, JamA, JamB, PeACP	pXZ27, pMS32
MS1_1	HsPKS1, JamA, JamB, PeACP-G36S	pXZ27, pMS32 mutant (from library)
MS1_2	HsPKS1, JamA, JamB, PeACP-G44S/A72V/H74R/A75D	pXZ27, pMS32 mutant (from library)
MS1_3	HsPKS1, JamA, JamB, PeACP-V26E/L37F	pXZ27, pMS32 mutant (from library)
MS1_4	HsPKS1, JamA, JamB, PeACP-V61A	pXZ27, pMS32 mutant (from library)
MS1_5	HsPKS1, JamA, JamB, PeACP-L53H/T60I	pXZ27, pMS32 mutant (from library)
MS1_6	HsPKS1, JamA, JamB, PeACP-D38G	pXZ27, pMS32 mutant (from library)
MS1_7	HsPKS1, JamA, JamB, PeACP-V26E	pXZ27, pMS75
MS1_8	HsPKS1, JamA, JamB, PeACP-L37F	pXZ27, pMS76
MS1_9	HsPKS1, JamA, JamB, PeACP-G44S	pXZ27, pMS77
MS1_10	HsPKS1, JamA, JamB, PeACP-L53H	pXZ27, pMS78
MS1_11	HsPKS1, JamA, JamB, PeACP-T60I	pXZ27, pMS79
MS6	HsPKS1, JamA, JamB, AntD-ACP	pXZ27, pXZ252

4.4.5 Bioinformatic Analysis of Proteins

Amino acid sequence alignment of the ACPs was performed using the online program Clustal Omega.⁵³ Protein structures of JamB and the ACPs were modeled using the online program HHPred⁵⁴ on the basis of a membrane-bound desaturase SCD1 (PDB: 4ZYO/4YMK)^{55, 56} and CurA-ACP (PDB: 2LIU),³³ respectively. ACP-JamB models were created by docking the ACP to the electropositive surface of JamB near the substrate entrance using ClusPro.^{57, 58} The visualizations of each structure and model were created using the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco.⁵⁹

4.4.6 Error-Prone PCR and Construction of *peACP* Mutant Library

The error-prone PCR procedure was modified from established protocols.⁶² The reaction consisted of 0.35 mmol/L dATP, 0.4 mmol/L dCTP, 0.2 mmol/L dGTP, 1.35 mmol/L dTTP, 1.75 mmol/L MnCl₂, and 1 U Taq polymerase. The reaction mixture was submitted to 25 cycles of PCR: 95 °C for 1 min, 50 °C for 1 min, and 68 °C for 1 min. The resulting PCR products were digested with NcoI/BamHI and inserted via cohesive end ligation into pXZ34, a pETDuet-1 plasmid containing

jamB, that was pre-digested with NcoI/BamHI. The cohesive end ligation was executed by following the NEB Quick Ligation Kit's protocol. The ligation product was purified using spin column centrifugation, co-transformed with pXZ27, a pCDFDuet-1 plasmid containing *jamA* and *hspks1*, to an electrocompetent BAP1 strain, and plated on LB agar containing 100 mg/L carbenicillin and 100 mg/L spectinomycin.

4.4.7 Plate-Reader Fluorescence Measurements

The measurements were performed following a modified protocol established previously.⁵⁰ Reactions were run in a 96-well black plate, and measurements were performed using a Tecan Safire Microplate Reader. Each well had a 200 μ L sample containing 2.25 μ mol/L of azido probe di-pegOF, 5 mmol/L ascorbic acid, 500 μ mol/L CuSO₄, 100 μ mol/L BTTAA, and 190 μ mol/L of cultures. The reactions were performed in dark for 10 min at room temperature.

Chapter 5. Conclusion

Parts of this chapter have been adapted from the following with permission:

Su, M., Zhu, X. & Zhang, W. Probing the Acyl Carrier Protein-Enzyme Interactions within Terminal Alkyne Biosynthetic Machinery. *AIChE Journal* **64**, 4255-4262 (2018).

In summary, we have probed and improved the ACP-enzyme interactions within the alkyne biosynthetic machinery. Using both *in vitro* and *in vivo* analyses, we demonstrated that the bi-functional desaturase/acetylenase (JamB) has a more stringent interaction with ACP than the acyl-ACP ligase (JamA). The recognition of JamB with non-cognate ACPs was further improved through ACP engineering. Specifically, rational engineering of TtuC based on the electrostatic interaction with JamB was explored, and a single charged amino acid mutation on helix 3 of TtuC improved the JamB recognition to a level comparable to the cognate ACP (JamC). Furthermore, we improved the interaction between JamB and PeACP through error-prone PCR and screening, and identified hot spots on PeACP that are important for interacting with JamB. This work has thus provided new insights into the critical role of ACP in the alkyne biosynthetic machinery, and illustrated the promise of protein engineering for generating a compatible heterologous ACP in alkyne biosynthesis to produce alkyne-tagged metabolites.

However, we also discovered important limitations when attempting to apply these two types of engineering methods towards non-cognate ACPs. To start with, both engineering methods require that the subject ACP is highly compatible with JamA, as neither method successfully established JamB recognition with non-cognate ACPs that had low JamA loading activity. For rational engineering, the ACP must be similar enough to the cognate ACP to allow a reasonable sequence alignment to be made. Rational engineering also requires a substantial knowledge of the subject ACP's key regions and the necessary properties that the amino acids in these areas must exhibit to establish JamB recognition. Such critical information is not always available, as was the case for the non-cognate ACPs PeACP and BpACP. This resulted in rational engineering efforts towards these ACPs to yield questionable findings. When applying error-prone PCR and fluorogenic screening procedures towards ACPs, a large number of ACP mutant genes needs to be created for a reasonable chance of obtaining several ACPs with improved JamB compatibility to occur. Although the fluorogenic screening assay can screen up to 88 ACP mutant strains per batch, doing so can often lead to low fluorogenic signals for every strain, including the positive controls, which make it impossible to detect which ACP mutant exhibits improved activity. These abnormally low signals often derive from the small-scale format of culturing the mutant strains in F1 media as well as from the necessity of individually adding the fluorogenic probes and the other reagents to the associated cell-free supernatants. When this problem occurs, the screening process must be repeated towards the entire batch of ACP mutant strains, the preparation of which takes at least four days to accomplish. As a result, it takes months to successfully apply the error-prone PCR and screening methods towards a single ACP, making this engineering approach not practical towards applying a wide variety of non-cognate ACPs within a reasonable period of time. In short, determining which of the two ACP engineering methods described in this dissertation is more suitable comes down to whether the principal objective is quickly obtaining results or ensuring productive findings from the engineering efforts. That's not to say that both rational engineering and error-prone PCR-based engineering can be applied to the same ACP to establish a comprehensive understanding about the ACP's important protein-protein interactions. Indeed,

more successful results can be achieved from rational engineering if the crucial residues were already identified from error-prone PCR and screening experiments applied towards the same ACP.

Attaining efficient pathway engineering is still an actively pursued goal, whose realization partially depends on optimizing the critical protein-protein interactions within the underlying biosynthetic pathways. As mentioned in the first paragraph of this chapter, the stringent recognition between ACPs and the bifunctional desaturases/acetylenases often becomes one of the limiting factors for engineered biosynthesis of alkyne-tagged natural products. The ACP engineering methods are not limited to improving JamB compatibility with non-cognate ACPs. Either of these methods, for example, can be applied to improve the production of another terminal alkyne pathway such as the TtuABC+HsPKS1 pathway that we characterized and described in the first two chapters. Attempting to optimize the compatibility of non-cognate ACPs towards that pathway, however, is more difficult due to the greater degree of flexibility that TtuA and TtuB possess towards non-cognate ACPs. A more accurate understanding of the crucial protein-protein interactions behind ACP-dependent biosynthetic pathways can be obtained when starting with an ACP that initially results in no alkyne production upon its replacement of the native ACP due to the non-cognate ACP's exclusive incompatibility with one of the other biosynthetic enzymes. Because the ACP is generally the central hub for the operation of multiple biosynthetic enzymes, improved understanding of the key ACP interactions with alkyne forming desaturases would help develop a more efficient platform for the *de novo* biosynthesis of alkyne tagged natural products. Another possible limiting factor is the ACP recognition by the downstream PKS/NRPS megasynthases. More information regarding ACP protein-protein interactions with PKS/NRPS, particularly the PKS-KS, NRPS-C, and PKS-AT domains, would be substantially helpful towards facilitating efficient pathway engineering. Depending on how incompatible these domains are towards non-cognate ACPs, rational engineering or error-prone PCR based engineering could provide more clues about the protein-protein interactions that are needed to occur.

As a final review, terminal alkynes possess broad applications in nearly all fields of modern chemistry and biology, which range from natural product discovery to biology and enzymology. Firstly, tagging PKs, NRPs and PK-NRP hybrids with terminal alkynes allows these products to be enriched and isolated using functionalized resins via click chemistry, substantially facilitating the purification process. Secondly, these tagged natural products can be easily visualized and traced, allowing the comprehensive assessment of these natural products' important characteristics. It was recently discovered that many natural products are adaptive compounds that are produced for specific physiological or social reasons.⁶³ Finally, the *de novo* biosynthesis of alkyne-tagged natural products may permit *in situ* detection and quantification of natural products in producing cell cultures through click chemistry. By optimizing the protein-protein interactions behind the relevant terminal alkyne biosynthesis pathways, we can develop a quantitative and high-throughput screening method that we can use to engineer the biosynthesis of natural products for overproduction and diversification.

References:

1. Binder, W.H. & Kluger, C. Azide/alkyne-"click" reactions: applications in material science and organic synthesis. *Curr Org Chem* **10**, 1791-1815 (2006).
2. Hein, C.D., Liu, X.M. & Wang, D. Click chemistry, a powerful tool for pharmaceutical sciences. *Pharm Res* **25**, 2216-2230 (2008).
3. Moses, J.E. & Moorhouse, A.D. The growing applications of click chemistry. *Chem Soc Rev* **36**, 1249-1262 (2007).
4. Thirumurugan, P., Matosiuk, D. & Jozwiak, K. Click Chemistry for Drug Development and Diverse Chemical–Biology Applications. *Chem Rev* **113**, 4905-4979 (2013).
5. Prescher, J.A. & Bertozzi, C.R. Chemistry in living systems. *Nat Chem Biol* **1**, 13-21 (2005).
6. Wei, L. *et al.* Live-cell imaging of alkyne-tagged small biomolecules by stimulated Raman scattering. *Nat Methods* **11**, 410-412 (2014).
7. Newman, D.J. & Cragg, G.M. Natural products as sources of new drugs from 1981 to 2014. *J Nat Prod* **79**, 629-661 (2016).
8. Clardy, J. & Walsh, C. Lessons from natural molecules. *Nature* **432**, 829-837 (2004).
9. Carlson, E.E. Natural Products as Chemical Probes. *ACS Chem Biol* **5**, 639-653 (2010).
10. Harmrolfs, K., Mancuso, L., Drung, B., Sasse, F. & Kirschning, A. Preparation of new alkyne-modified ansamitocins by mutasynthesis. *Beilstein J Org Chem* **10**, 535–543 (2014).
11. Salazar, M.D. & Ratnam, M. The folate receptor: what does it promise in tissue-targeted therapeutics? *Cancer Metastasis Rev* **26**, 141-152 (2007).
12. Low, P.S. & Antony, A.C. Folate receptor-targeted drugs for cancer and inflammatory diseases. *Adv Drug Deliv Rev* **56**, 1055-1058 (2004).
13. Lee, R.J., Wang, S. & Low, P.S. Measurement of endosome pH following folate receptor-mediated endocytosis. *Biochim Biophys Acta* **1312**, 237-242 (1996).
14. Staub, I. & Sieber, S.A. Beta-lactam probes as selective chemical-proteomic tools for the identification and functional characterization of resistance associated enzymes in MRSA. *J Am Chem Soc* **131**, 6271-6276 (2009).
15. Staub, I. & Sieber, S.A. Beta-lactams as selective chemical probes for the in vivo labeling of bacterial enzymes involved in cell wall biosynthesis, antibiotic resistance, and virulence. *J Am Chem Soc* **130**, 13400-13409 (2008).
16. Minto, R.E. & Blacklock, B.J. Biosynthesis and function of polyacetylenes and allied natural products. *Prog Lipid Res* **47**, 233-306 (2008).
17. Chai, Q.Y., Yang, Z., Lin, H.W. & Han, B.N. Alkynyl-containing peptides of marine origin: A review. *Mar Drugs* **14**, 216 (2016).
18. McPhail, K.L. *et al.* Antimalarial linear lipopeptides from a Panamanian strain of the marine cyanobacterium *Lyngbya majuscula*. *J Nat Prod* **70**, 984-988 (2007).
19. Hooper, G.J., Orjala, J., Schatzman, R.C. & Gerwick, W.H. Carmabins A and B, new lipopeptides from the Caribbean cyanobacterium *Lyngbya majuscula*. *J Nat Prod* **61**, 529-533 (1998).
20. Tripathi, A., Puddick, J., Prinsep, M.R., Lee, P.P.F. & Tan, L.T. Hantupeptin A, a cytotoxic cyclic depsipeptide from a Singapore collection of *Lyngbya majuscula*. *J Nat Prod* **72**, 29-32 (2009).

21. Yamaguchi, M., Park, H.J., Ishizuka, S., Omata, K. & Hiram, M. Chemistry and antimicrobial activity of caryophenols analogs. *J Med Chem* **38**, 5015-5022 (1995).
22. Ross, C., Scherlach, K., Kloss, F. & Hertweck, C. The Molecular basis of conjugated polyene biosynthesis in phytopathogenic bacteria. *Angew Chem Int Ed* **53**, 7794-7798 (2014).
23. Edwards, D.J. *et al.* Structure and biosynthesis of the jamaicamides, new mixed polyketide-peptide neurotoxins from the marine cyanobacterium *Lyngbya majuscula*. *Chem Biol* **11**, 817-833 (2004).
24. Jones, A.C. *et al.* Genomic insights into the physiology and ecology of the marine filamentous cyanobacterium *Lyngbya majuscula*. *P Natl Acad Sci USA* **108**, 8815-8820 (2011).
25. Zhu, X., Liu, J. & Zhang, W. *De novo* biosynthesis of terminal alkyne-labeled natural products. *Nat Chem Biol* **11**, 115-120 (2015).
26. Zhu, X., Su, M., Manickam, K. & Zhang, W. Bacterial Genome Mining of Enzymatic Tools for Alkyne Biosynthesis. *ACS Chem Biol* **10**, 2785-2793 (2015).
27. Marchand, J.A. *et al.* Discovery of a pathway for terminal-alkyne amino acid biosynthesis. *Nature* **567**, 420-424 (2019).
28. Mercer A.C. & Burkart, M.D. The ubiquitous carrier protein—a window to metabolite biosynthesis. *Nat Prod Rep* **24**, 750-773 (2007).
29. Twigg, F., Skyrud, D., Li, J. & Zhang, W. Engineering Enzymes for Natural Product Biosynthesis and Diversification. *Modern Biocatalysis: Advances Towards Synthetic Biological Systems* **10**, 261-286 (2018).
30. Klaus, M. *et al.* Protein–Protein Interactions, not Substrate Recognition, Dominates the Turnover of Chimeric Assembly Line Polyketide Synthases. *J Biol Chem* **291**, 16404-16415 (2016).
31. Nguyen, C. *et al.* Trapping the dynamic acyl carrier protein in fatty acid biosynthesis. *Nature* **505**, 427-431 (2014).
32. Tanovic, A., Samel, S.A., Essen, L.O. & Marahiel, M.A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **321**, 659-663 (2008).
33. Busche, A. *et al.* Characterization of molecular interactions between ACP and halogenase domains in the Curacin A polyketide synthase. *ACS Chem Biol* **7**, 378-386 (2012).
34. Whicher, J.R. *et al.* Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **510**, 560-564 (2014).
35. Barajas, J.F. *et al.* Structural and Biochemical Analysis of Protein-Protein Interactions Between the Acyl-Carrier Protein and Product Template Domain. *Angew Chem Int Ed Engl* **55**, 13005-13009 (2016).
36. Weissman, K.J. & Mueller, R. Protein-protein interactions in multienzyme megasynthetases. *ChemBiochem* **9**, 826-48 (2008).
37. Weissman, K.J. Uncovering the structures of modular polyketide synthases. *Nat Prod Rep* **32**, 436-453 (2015).
38. Drake, E.J. *et al.* Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235-238 (2016).
39. Dutta, S. *et al.* Structure of a modular polyketide synthase. *Nature* **510**, 512-517 (2014).
40. Dorrestein, P.C. *et al.* Activity screening of carrier domains within nonribosomal peptide synthetases using complex substrate mixtures and large molecule mass spectrometry. *Biochemistry* **45**, 1537-1546 (2006).

41. Lau, J., Cane, D.E. & Khosla, C. Substrate specificity of the loading didomain of the erythromycin polyketide synthase. *Biochemistry* **39**, 10514-10520 (2000).
42. Wittmann, M., Linne, U., Pohlmann, V. & Marahiel, M.A. Role of DptE and DptF in the lipidation reaction of daptomycin. *FEBS J* **275**, 5343-5354 (2008).
43. Hansen, D.B., Bumpus, S.B., Aron, Z.D., Kelleher, N.L. & Walsh, C.T. The Loading Module of Mycosubtilin: An Adenylation Domain with Fatty Acid Selectivity. *J Am Chem Soc* **129**, 6366-6367 (2007).
44. Miao, V. *et al.* The lipopeptide antibiotic A54145 biosynthetic gene cluster from *Streptomyces fradiae*. *J Ind Microbiol Biotechnol* **33**, 129-140 (2006).
45. Powell, A. *et al.* Engineered biosynthesis of nonribosomal lipopeptides with modified fatty acid side chains. *J Am Chem Soc* **129**, 15182-15191 (2007).
46. Harris, N.C. *et al.* Biosynthesis of isonitrile lipopeptides by conserved nonribosomal peptide synthetase gene clusters in Actinobacteria. *Proc Natl Acad Sci U S A* **114**, 7025-7030 (2017).
47. Sandy, M., Rui, Z., Gallagher, J. & Zhang, W. Enzymatic Synthesis of Dilactone Scaffold of Antimycins. *ACS Chem Biol* **7**, 1956-1961 (2012).
48. Yan, Y. *et al.* Biosynthetic pathway for high structural diversity of a common dilactone core in antimycin production. *Org Lett* **14**, 4142-4145 (2012).
49. Pfeifer, B.A., Admiraal, S.J., Gramajo, H., Cane, D.E. & Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **291**, 1790-1792 (2001).
50. Zhu, X., Shieh, P., Su, M., Bertozzi, C.R. & Zhang, W. A fluorogenic screening platform enables directed evolution of an alkyne biosynthetic tool. *Chem Commun (Camb)* **52**, 11239-11242 (2016).
51. Morita, H. *et al.* Synthesis of unnatural alkaloid scaffolds by exploiting plant polyketide synthase. *Proc Natl Acad Sci U S A* **108**, 13504-13509 (2011).
52. Guy, J.E. *et al.* Remote control of regioselectivity in acyl-acyl carrier protein-desaturases. *Proc Natl Acad Sci U S A* **108**, 16594-16599 (2011).
53. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 1-6 (2011).
54. Söding, J., Biegert, A. & Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**, W244-248 (2005).
55. Bai, Y. *et al.* X-ray structure of a mammalian stearyl-CoA desaturase. *Nature* **524**, 252-256 (2015).
56. Wang, H., *et al.* Crystal structure of human stearyl-coenzyme A desaturase in complex with substrate. *Nat Struct Mol Biol* **22**, 581-585 (2015).
57. Comeau, S.R., Gatchell, D.W., Vajda, S. & Camacho, C.J. ClusPro: A fully automated algorithm for protein-protein docking. *Nucleic Acids Res* **32**, W96-W99 (2004).
58. Comeau, S.R., Gatchell, D.W., Vajda, S. & Camacho, C.J. ClusPro: An automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* **20**, 45-50 (2004).
59. Pettersen, E.F. *et al.* UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).
60. Fischbach, M.A., Lai, J.R., Roche, E.D., Walsh, C.T. & Liu, D.R. Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. *Proc Natl Acad Sci U S A* **104**, 11951-11956 (2007).

61. Zhou, Z., Lai, J.R. & Walsh, C.T. Directed evolution of aryl carrier proteins in the enterobactin synthetase. *Proc Natl Acad Sci U S A* **104**, 11621-11626 (2007).
62. Gao, X. *et al.* Directed evolution and structural characterization of a simvastatin synthase. *Chem Biol* **16**, 1064-1074 (2009).
63. O'Brien, J. & Wright, G.D. An ecological perspective of microbial secondary metabolism. *Curr Opin Biotech* **22**, 552-558 (2011).

Appendices

A. SDS-PAGE analysis of proteins purified from *E. coli*.

Mini-PROTEAN TGX gels (4-15% precast, Biorad) were used. All proteins have a C-terminal hexahistidine tag. (A) From left to right the proteins are: TtuC-K69Y, TtuC-K69E, TtuC-K69C, TtuC-K69S, TtuC-K69V, TtuC-K69R. (B) From left to right: TtuC-I68P, TtuC-K69T, TtuC-S70I, TtuC-I71L, TtuC-R72Y, TtuC-K73S, TtuC-N74Y, and TtuC-N75P. (C) From left to right the proteins are: JamC, ScoB, DEBS1-ACP, Sco3249, DptF, no ACP control, MycA, and LptF. (D) The AntD-ACP protein. (E) The PeACP mutants. From left to right the PeACP mutants displayed are: L37F, T60I, D38G, G36S, and V61A. The presence of apo- and holo-ACPs was confirmed by LC-HRMS analysis and the ratio of holo- vs. apo-ACPs purified from *E. coli* was illustrated in Table 2-2.

B. Deconvoluted masses of acyl-ACPs in HRMS analysis.

	5-hexenoyl-ACP deconvoluted mass (amu)	
ACP	Theoretical	Experimental
JamC	12658.95	12659.17
TtuC	10822.00	10822.32
TtuC-K69T	10925.93	10926.31
TtuC-K69E	10822.94	10823.13
BpACP	12251.61	12252.59
ScoB	11114.31	11114.38
IlACP	10755.17	10755.73
SpACP	11082.29	11083.38
PeACP	10549.50	10550.31
PeACP-G36S	10636.77	10636.99
PeACP-L37F	10640.76	10640.95
PeACP-D38G	10548.70	10548.94
PeACP-T60I	10618.79	10619.06
PeACP-V61A	10578.69	10578.95
PfACP	13206.83	13207.07
CyACP	11792.54	11792.29

C. Relative JamA *in vitro* activities towards different ACPs.

Approximately 25% of JamC was converted to 5-hexenoyl-JamC after 5 minutes of the *in vitro* reaction with JamA, and this conversion was set to be a relative JamA activity of 100%. Error bars represent standard deviations from at least three experiments.

D. HRMS of compounds 1-4.

E. Relative JamB *in vivo* activities towards ACPs.

The compounds **1** and **2** were produced in an approximately 1:5 ratio by the strain XZ1, and this efficiency was set to be a relative JamB activity of 100%. Error bars represent standard deviations from at least three experiments.