ABSTRACT

CASTELLI, KIMBERLY R. Efforts Towards an *in vivo* Platform for the Combinatorial Biosynthesis of Polyketides. (Under the direction of Dr. Gavin Williams).

Polyketides are a medically-relevant class of natural products that are produced by large modular proteins called polyketide synthases (PKSs). The loading modules of PKSs are responsible for initiating polyketide biosynthesis by priming the enzyme with monocarboxylic CoA-esters, or starter units. Efforts towards engineering polyketide diversity at the loading module are dependent upon the availability of acyl-CoA starter units, and new biosynthetic methods for the production of acyl-CoAs are desired. Herein, the substrate scope and kinetics of an acyl-CoA synthetase from *Pseudomonas chlororaphis*, AcsA, is presented. Mutagenesis of AcsA has provided insight into substrate selection by the enzyme active site, laying the foundation for future efforts in engineering the substrate scope of AcsA.

In addition, engineering polyketides in the natural host organisms is a long-term goal for the robust production of polyketides and polyketide analogs. The application of CRISPR-Cas9 as a tool for the *in vivo* mutagenesis of PKSs in *Actinomyces* is presented here. Preliminary analysis of fermentation extracts of polyketide-producing hosts confirms the production of FK520 and rapamycin, providing a starting point for future experiments.

Efforts Towards an *in vivo* Platform for the Combinatorial Biosynthesis of Polyketides

by
Kimberly R Castelli

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Chemistry

Chemistry

Raleigh, North Carolina

2016

APPROVED BY:

_____	_____
Dr. Gavin Williams Committee Chair	Dr. Reza Ghiladi
_____	
Dr. Chase Beisel	

BIOGRAPHY

Kimberly was born and raised in Carmel, New York. She moved to Raleigh, North Carolina in 2008 to attend North Carolina State University for her undergraduate studies. She graduated in 2012 with a Bachelor of Science in Molecular, Cellular, Developmental Biology, and a Bachelor of Science in Biochemistry. Kimberly took a job as a Chemistry Lab Analyst at Hospira after graduation, where she spent two years doing quality control testing on injectable pharmaceuticals. During this time, Kimberly knew she wanted to pursue a graduate degree to start her career in research. She was excited to receive an offer to return to North Carolina State University to work with Dr. Gavin Williams in the Department of Chemistry.

ACKNOWLEDGMENTS

I would first like to thank Dr. Gavin Williams for providing me with opportunities to grow as a scientist and allowing me to take on new and exciting projects in the the Williams Lab. In addition, I would like to thank my committee members, Dr. Reza Ghiladi and Dr. Chase Beisel.

I would also like to acknowledge the ongoing support that I received from the members of the Williams Lab as well as my friends and family. Having such a solid scientific and personal support system was indispensable for my success as a graduate student.

Finally, I would like to acknowledge and thank Ryan Leenay from the Beisel Lab, and Dr. Chase Beisel, himself, for their helpful discussions regarding CRISPR-Cas9.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Engineering the biosynthesis of polyketides

1.1 Introduction

Polyketides are a structurally diverse class of molecules that are produced as secondary metabolites by a variety of organisms (**Figure 1-1**)[1]. Polyketides often exhibit biological activities such as antibiotic, immunosuppressive, anti-tumor, and neuro-regenerative, which render this class of natural products very important with respect to drug discovery and development. In the case of the well-known polyketide antibiotic erythromycin, erythromycin-resistant organisms, such as *Staphylococcus,* emerged shortly after the clinical introduction of the drug in 1953[2]. One approach to battling microbial resistance to erythromycin and other important polyketide antibiotics is to produce analogs that could potentially evade the resistance mechanisms evolved by the resistant organisms[3]. While synthetic or semi-synthetic derivatives of erythromycin have been successfully made[4], a purely biosynthetic approach to producing erythromycin analogs has been the focus of decades of research[5–8].

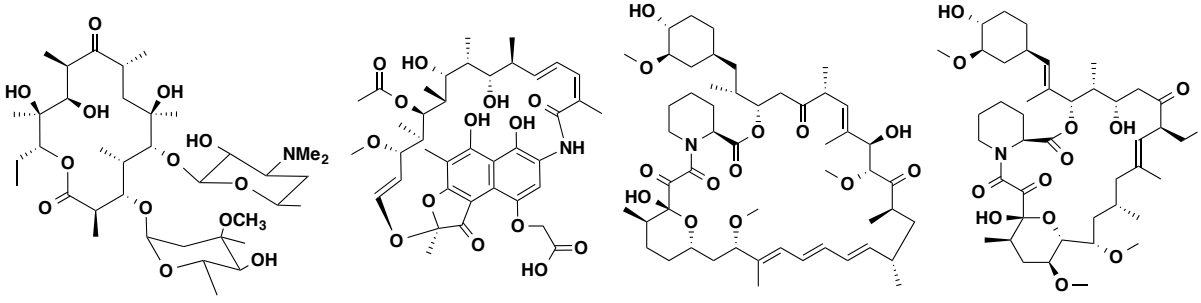


Figure 1-1: *Exemplary bacterial polyketides and their respective producing hosts.*

Erythromycin is biosynthesized by the type I polyketide synthase (PKS) 6-deoxyerythronolide B synthase (DEBS) from *Saccharopolyspora erythraea.* Modular type I PKSs like DEBS are large "assembly line" proteins made up of covalently linked catalytic domains that synthesize polyketides in a linear fashion[9]. Each step of polyketide biosynthesis is a result of the activity of a specific domain within the PKS. As such, every set of carbons within the polyketide structure can be assigned to a unique and specific PKS module. The DEBS PKS spans three genes in the genome of *S. erythraea*, encoding the proteins DEBS1, DEBS2 and DEBS3. Each gene encodes multiple catalytic domains that are organized into modules, beginning with the "loading module" and module 1 through module 6. The modular

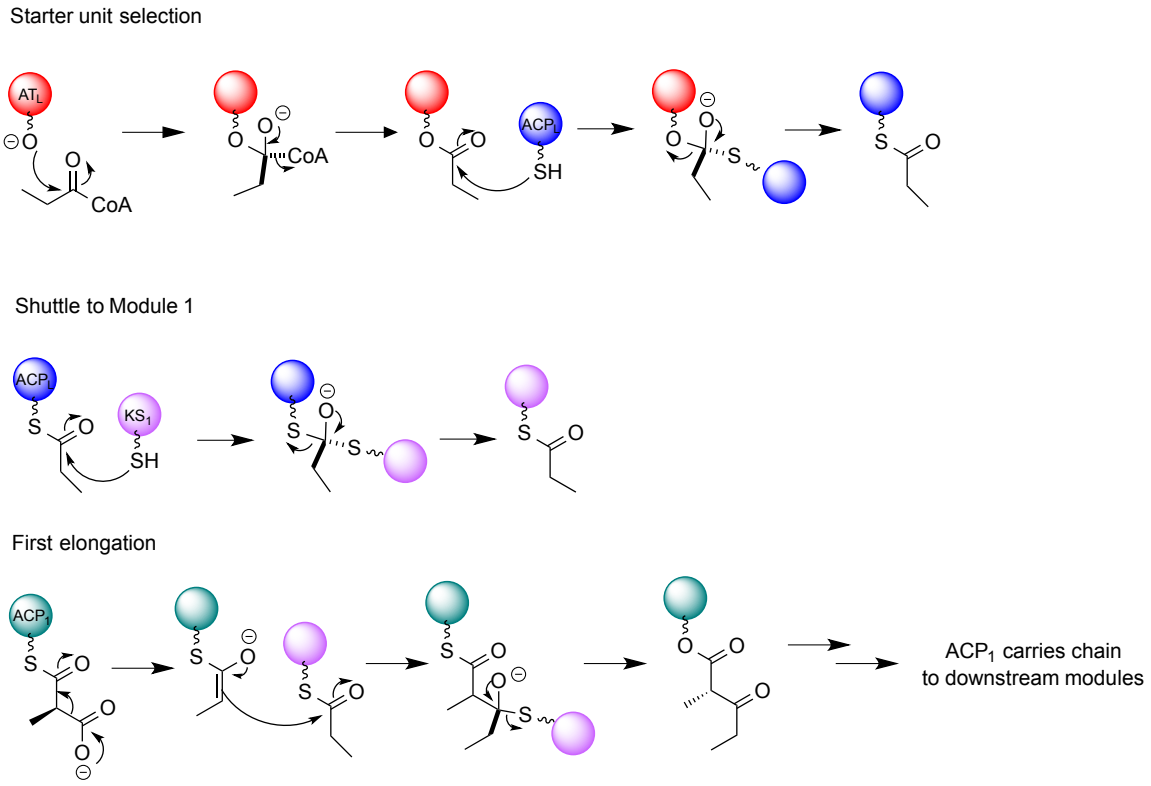


Figure 1-2. *Starter unit selection and decarboxylative condensation by Module 1 of DEBS. AT_L: loading domain acyltransferase; ACP_L: loading module acyl carrier protein; ACP_1: module 1 acyl carrier protein; KS_1: module 1 ketosynthase.*

organization of PKSs is the foundation for development of a combinatorial biosynthetic approach to diversifying polyketides[10,11].

Each module within a PKS consists minimally of an acyltransferase (AT), a ketosynthase (KS) and an acyl carrier protein (ACP). The AT acts as the "gatekeeper" for each module, selecting an activated thioester building block to load onto the acyl carrier protein (ACP), which acts as the shuttle that moves within and between the modules. In the first PKS module, the KS domain accepts the activated starter unit and catalyzes a Claisen-like decarboxylative addition of the building block from the ACP (**Figure 1-2**). Additional reductive domains may exist within a module to modify the product following each extension. Iterative extensions and modifications create a growing chain that is often cyclized after the final extension module. Many polyketides undergo post-PKS processing to render the fully biologically active natural product. Other polyketides, such as FK506/FK520 and rifamycin, are produced by hybrid PKS non-ribosomal peptide synthases (PKS-NRPSs). As well as introducing amino acids into the natural product scaffold, the NRPSs of these hybrid pathways are responsible for additional modifications like *O*-methylation, reduction, or isomerization[12].

Many acyl building blocks used in polyketide biosynthesis are generated during primary metabolism, or from the modification of primary metabolites. For example, the propionyl-Coenzyme A (CoA) starter unit used to initiate erythromycin biosynthesis is produced via decarboxylation of methylmalonyl-CoA, which is derived from the primary metabolite succinyl-CoA in *S. erythraea*[13].

In the DEBS pathway, the loading domain AT (AT_L) is responsible for priming the loading ACP with a propionyl-CoA, while the AT of DEBS modules 1-6 (AT_1, AT_2, etc.) prime the ACP with methylmalonyl-CoA. The six modules of DEBS catalyze iterative condensations of methylmalonyl-CoA to afford a chain which is then cyclized by the terminal thioesterase (TE) domain, forming the erythromycin aglycone core, 6-deoxyerythronolide B (6dEB – **Figure 1-3**). The order of the genes in the DEBS pathway generally corresponds with the order in which the modules catalyze a condensation event. Knowing the substrate specificity for each module allows for predicting the structure of the final polyketide product. Due to this collinearity of

PKS organization and polyketide biosynthesis, significant effort has been placed on developing a combinatorial biosynthetic approach to polyketide diversification.

While erythromycin biosynthesis has been the major focus of PKS research for many years, many other polyketides of interest have become increasingly popular targets for engineering. Rifamycin B, for example, is produced by *Amycolatopsis mediterranei* and exhibits antibiotic activity again *Mycobacteria* and is important for the treatment of tuberculosis (TB). Semi-synthetic derivatives of rifamycin B, such as rifampicin, have been made that exhibit higher potency against TB-causing strains of *Mycobacteria*. However, these semi-synthetic compounds were created by modifying the same accessible region of rifamycin B (carbons 3-4), and resistance to all of these derivatives has been observed[14–16]. Recently, a biosynthetic analog of rifamycin B, 24-desmethylrifamycin B, was produced via PKS

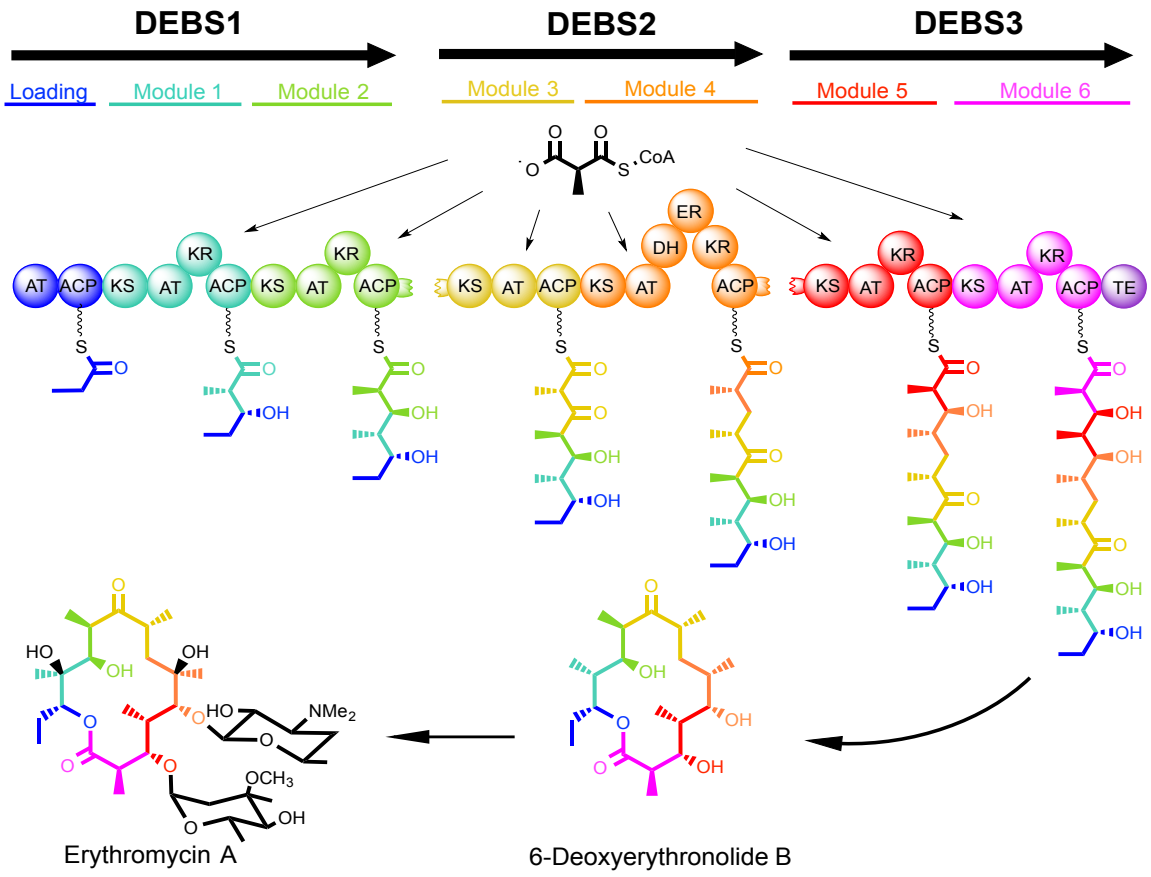


Figure 1-3: *Erythromycin biosynthesis.*

engineering. This compound was tested against rifampicin-resistant strains of *Mycobacterium tuberculosis* and demonstrated higher antibacterial activity[17]. The ability to alter regions of polyketide scaffolds that are currently synthetically-inaccessible highlights the utility of PKS engineering approaches to diversifying polyketide structures.

1.2 Engineering Polyketide Enzymatic Machinery

There are three main opportunities for diversifying polyketide structure using biosynthetic methods: building block selection, PKS tailoring, and post-PKS modification. However, the diversity of polyketide structures is limited in nature mainly by the availability and variety of starter and extender units selected by the PKS. Semi-synthetic approaches are also a viable route to polyketide analogs, but this is also limited by the availability of chemical handles on the core structure. Therefore, expanding the enzymatic toolbox via PKS engineering is an attractive means for creating polyketides with new chemical handles for further chemical or biosynthetic modification by the incorporation of non-natural building blocks.

One approach to PKS engineering focuses on chimeragenesis to produce truncated, elongated, or substituted polyketides. Replacing the AT_L of a truncated DEBS1 with the ATs from the soraphen PKS and oleandomycin PKS produced benzyl-substituted and acetyl-derived triketide lactones, respectively[18,19]. Insertion of the malonyl-CoA-specific second extension module from the rapamycin PKS between the first and second extension modules of DEBS1 produced DEBS products with extended core structures[20]. Complete derivatives of the erythromycin macrocyclic structure have been successfully biosynthesized by swapping the AT of an erythromycin extension module with ATs from other PKSs, producing substituted erythromycin derivatives[6,21,22]. Substituting ATs from DEBS into other PKSs has also led to a greater understanding of AT substrate specificity while producing polyketide analogs[23,24]. While chimeragenesis has its benefits for probing the scope of polyketide diversification, it is still limited to the substrates naturally accepted by enzymatic machinery and by the ability to navigate potentially unfavorable protein-protein interactions.

In contrast, engineering the AT domains within each module of the DEBS system has demonstrated a viable route to producing regioselectively modified polyketides. Single

mutations within highly conserved regions of DEBS allowed for discrimination between acetyl-CoA and propionyl-CoA starter units to selectively produce the respective triketide lactone products[25]. Additionally, point mutations in AT active sites of DEBS has allowed for incorporation of non-natural substrates such as ethylmalonyl-CoA, azidomalonyl-CoA, and propargylmalonyl-CoA to create substituted derivatives of 6dEB.

While erythromycin biosynthesis has been a large focus of PKS engineering, other polyketides such as rapamycin, FK506/520 and rifamycin have been receiving more attention due their therapeutic potential. A significant body of work has demonstrated that derivatives of rapamycin (Rap) and FK520 can be produced via precursor-directed biosynthesis, where alternative starter units are supplied and incorporated into the polyketide structure by the loading domain[26–29]. A different approach utilizing domain swapping produced the rifamycin B analogs, 32-desmethylrifamycin B and SV, through replacement of rifamycin (Rif) AT_6 with Rap AT_2. This AT replacement permitted the incorporation of an acetate in place of propionate (*via* malonyl-CoA vs. methylmalonyl-CoA, respectively) and led to production of semi-synthetic rifamycin derivatives that displayed effective antibiotic activity towards rifampicin-resistant *Mycobacterium tuberculosis*[30]. An important goal in PKS engineering is to easily and selectively modify polyketide scaffolds to improve the physical or chemical properties of the "drug" (e.g. improved antibiotic activity, increased water-solubility, increased target binding efficiency for decreased side-effects, etc.). In order to expand the scope of modifications that can be made to the polyketide core structures, a variety of non-natural starter and extender units to be utilized by engineered PKS machinery must be available.

1.3 Biosynthesis of polyketide extender units

Recent directed evolution and rational mutagenesis of a malonyl-CoA synthetase, MatB, from *Rhizobium trifolii,* produced mutant MatB enzymes capable of synthesizing a panel of CoA-linked extender units (**Figure 1-4**) [31,32]. This work has enabled the investigation of PKS substrate specificity beyond those found in nature. In addition, this work highlights the ability to alter enzyme substrate specificity through point mutations in the enzyme active site.

On a broader scale, the ability to couple polyketide engineering and non-natural starter unit biosynthesis points towards an efficient method for diversifying polyketide structure. Transferring engineered MatB genes into polyketide-producing bacterial hosts that house an engineered PKS that accepts non-natural substrates could allow for *in vivo* production of polyketide analogs. However, even as the panel of substrates that are accepted by mutant MatB grows, the product scope is still limited to the production of activated dicarboxylic acids. Many polyketide scaffolds require monocarboxylic acid building blocks, including the starter units for erythryomycin, FK520, rifamycin and rapamycin biosynthesis. The selective enzymatic synthesis of non-natural monocarboxylic acids *in vivo* could potentially provide a route for biosynthesis of novel compounds diversified at the carbons installed by the loading domains.

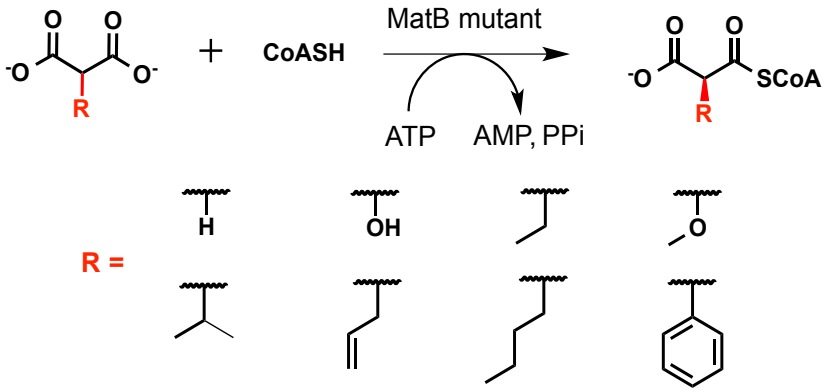


Figure 1-4: *Panel of accessible MatB synthesized extender units.*

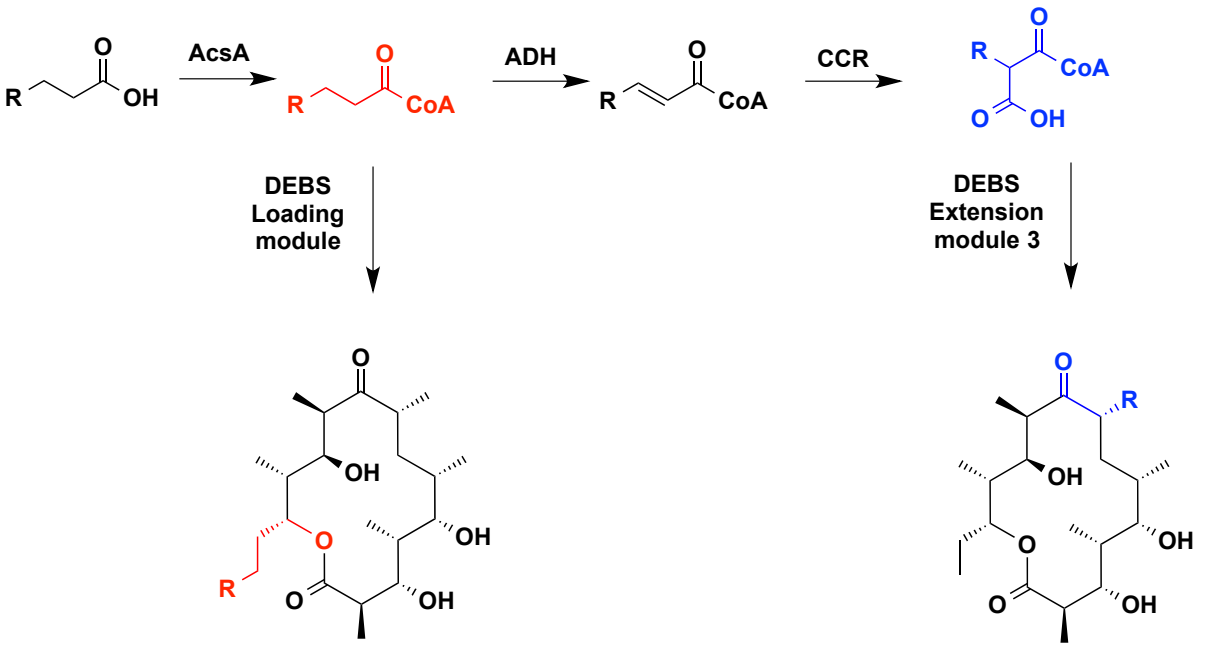


Figure 1-5: *Pathway to making starter and extender unit CoA esters using AcsA, an acyl dehydrogenase (ADH), and a carboxylase/reductase (CCR). New propionate-derived starter units could be accessed via AcsA, allowing for new functionalization at C1 position of erythromycin. New malonate-derivatives could be synthesized this way to functionalize C2, C4, C6, C8, C10, and C12 of erythromycin. This pathway could be applied to other PKS systems.*

1.4 Chemoenzymatic routes to acyl-CoA starter units

Nature has provided biosynthetic routes to a variety of interesting CoA-linked starter and extender units that are incorporated into polyketide pathways[33]. Advances in genetic analysis have led to the discovery of the enzymes responsible for the transformations of primary metabolites into esters. The enzymes responsible for this transformation are classified as adenylate-forming enzymes, and they synthesize CoA-thioesters by first activating the carboxylic acid with adenosine monophosphate (AMP)[34]. Recent work in characterizing the substrate scope of five different CoA ligases has established an enzymatic route to a relatively large panel of CoA thioesters. A malonyl-CoA synthetase, 4-coumarate-CoA ligase, benzoate-CoA ligase, cinnamyl-CoA ligase and penylacetyl-CoA ligase were tested against 123 different carboxylic acids[35]. Two malonyl-CoA synthetase (MCS) chimeras were

produced by swapping the "substrate specificity loop" of MCS with that of 4-coumarate-CoA ligase and benzoate-CoA ligase. However, neither chimera was shown to catalyze the ligation of CoA to any of the 123 tested carboxylic acids. The failure of this approach in producing a promiscuous CoA ligase leaves to question whether there is a more viable and conservative approach to altering the substrate scope and specificity of a CoA ligase. And while this work provides a starting point for the biosynthesis of a large number of acyl-CoA substrates, none of the enzymes studied provided viable routes to short-chain acyl-CoAs, such as propionyl-CoA and derivatives.

As an alternative to enzymatic ligation of CoA and various carboxylic acids, several synthetic methods for producing acyl-CoAs have been established, including the chemical activation of carboxylic acids by carboxyldiimidazole (CDI) and ethylchloroformate (ECF)[36]. CDI-activation of carboxylic acids led to the production of several saturated acyl-CoAs, ranging from 4 – 12 carbons in length, including isobutyryl-CoA and succinyl-CoA, in modest to good yields (68% and 86% respectively) in two synthetic steps. The products of this chemistry were then tested with three different acyl-CoA dehydrogenase (ADH) enzymes for the desaturation of acyl-CoAs. Many products were accessible with a conversion rate of 70% - 100% using this approach, namely acryloyl-CoA and crotonyl-CoA. Accordingly, in addition to providing a possible source of starter units, an enzyme for the biosynthesis of a variety of acyl-CoAs could be coupled with an ADH and crotonyl-CoA carboxylase/reductase (CCR) to produce polyketide extender units (**Figure 1-5**). The natural activity of CCR is either to reduce crotonyl-CoA to butyryl-CoA, or reductively carboxylate crotonyl-CoA to ethyl-malonyl-CoA (**Figure 1-6)**. This enzyme will also act upon acryloyl-CoA to produce either propionyl-CoA or methylmalonyl-CoA, and does so in a stereospecific manner[37]. Altogether then, promiscuous tailored acyl-CoA synthetases could provide simultaneous access to polyketide starter and extender units from simple, scalable starting materials that may not be available using current methods.

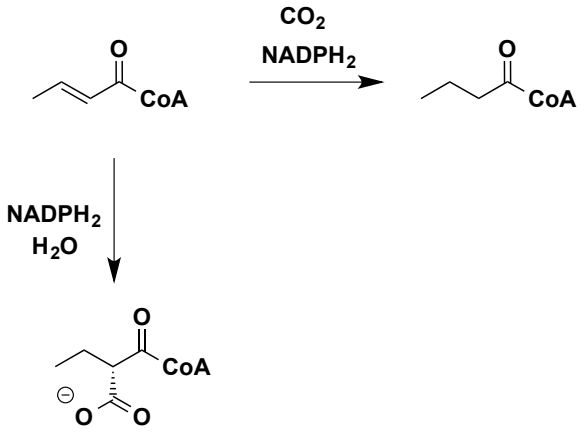


Figure 1-6: *CCR-catalyzed reduction or reductive carboxylation of crotonyl-CoA to butyryl-CoA and ethylmalonyl-CoA, respectively.*

1.5 Scope of this thesis

This thesis lays the foundation for future efforts in establishing an *in vivo* platform for generating polyketide analogs via AT engineering with non-natural starter/extender units. The goal of Chapter 2 was to establish the methods necessary for engineering an acyl-CoA synthetase to generate non-natural acyl-CoA starter units for use in polyketide biosynthesis. Presented in Chapter 2 is an absorbance-based end-point point assay for measuring the kinetics of acyl-CoA ligation reactions and for use as a high-throughput screening tool with saturation libraries. These methods were used to characterize a series of AcsA mutants. The goal of Chapter 3 was to begin to manipulate native polyketide-producing *Actinomyces* hosts for polyketide production and to evaluate CRISPR-Cas9 as a genome editing tool for AT engineering *in vivo*.

CHAPTER 2

Engineering the short chain acyl-CoA synthetase, AcsA from *Pseudomonas chlororaphis*

2.1 Introduction

Precursor-directed mutasynthesis has proven to be a viable route for the production of polyketide analogs, which takes advantage of the relaxed substrate specificity of the loading ATs of many different PKSs[26,38–44]. Engineering the ATs of naturally-promiscuous loading domains for altered substrate selectivity or substrate scope has yet to be fully explored. One limitation to this approach is the commercial or synthetic accessibility of acyl-CoAs. In terms of synthetic accessibility, *N*-acetylcysteamine (SNAC) can be used in place of CoA to activate starter and extender units for the incorporation by polyketides[45]. It has been shown *in vitro* and *in vivo* that synthetic acyl-SNACs can be successfully incorporated by PKS machinery to produce derivatives of polyketides, however there is a clear preference for CoA-linked starter and extender units[46]. Furthermore, once hydrolyzed, there is no way to regenerate the intact acyl-SNAC. Promiscuous fatty-acyl CoA ligases, such as FadD from *Mycobacterium tuberculosis*, have been shown to produce both acyl-CoAs and acyl-SNACs[47]. However, as these enzymes are part of fatty acid biosynthesis, the preferred substrates are typically characterized as long chain (> 6 carbons) acyl-thioesters.

The ability to enzymatically synthesize a panel of CoA-activated mono-carboxylic acid esters would be advantageous for further expanding the chemical diversity that can be introduced into polyketides. Ideally, such an enzyme could discriminate between primary metabolites and non-native or non-natural carboxylic acids that are fed to the cells via the media. This could allow for the generation of non-natural acyl-CoA starter units *in vivo* that could be incorporated into the polyketide by an engineered or naturally promiscuous PKS without consuming valuable resources from primary metabolism. While there are available routes to selectively activate dicarboxylic acids, such as malonate derivatives for use as extender units, there is no simple enzymatic route to activating a variety of short-chain monocarboxylic acids, such as propionate and derivatives, for use as starter units. Additionally,

there is no biosynthetic approach that allows for the selective production of activated acyl-CoAs.

As such, the short chain acyl-CoA synthetase, AcsA, from *Pseudomonas chlororaphis*, was identified here as a primary target for engineering. This enzyme shows some activity with several short-chain monocarboxylic acids including propionic acid, butyric acid, isobutyric acid and valeric acid[48]. While the preferred substrate appears to be isobutyric acid, the ability of the enzyme to utilize other carboxylic acids indicates that this enzyme might be primed for engineering a broad substrate scope, and possible shifts in substrate specificity. Prior to engineering AcsA, the substrate scope must be further characterized.

2.2 Assessing the activity of wild-type AcsA

AcsA activates short-chain (< 6 carbons) carboxylic acids with CoA, first activating the acid with adenosine monophosphate (AMP) and releasing pyrophosphate (**Figure 2-1**). Given that the enzyme consumes CoA, an *in vitro* end-point assay was first developed that allows for the quantification of enzyme activity by directly measuring the amount of CoA remaining in the reaction. 5,5'-dithiobis(2-nitrobenzoic acid), commonly referred to as Ellman's reagent or DTNB, reacts stoichiometrically with CoA to form 2-nitro-5-thiobenzoate (TNB^{2-}). This product exhibits an absorbance maximum at 412 nm, which allows for a direct measurement of the concentration of CoA in the reaction mixture (**Figure 2-1**). Using Ellman's reagent to measure the depletion of CoA over time, we set out to determine the Michaelis-Menten kinetics of the wild-type AcsA (WT-AcsA) with a panel of substrates (**Figure 2-2**). The substrates were chosen to represent a variety of chemical moeties of interest, including short (< 6 carbons) chain and branched-chain alkyl groups (e.g. **1 – 4**), activated (**7**) and deactivated aromatic rings (**6**), short-chain dicarboxylic acids (**8 – 9**), and an amino acid (**5**). Products were detected by HPLC with substrates **1** – **4**, **6**, (**Appendix A**) and the masses confirmed via low-res LC-MS.

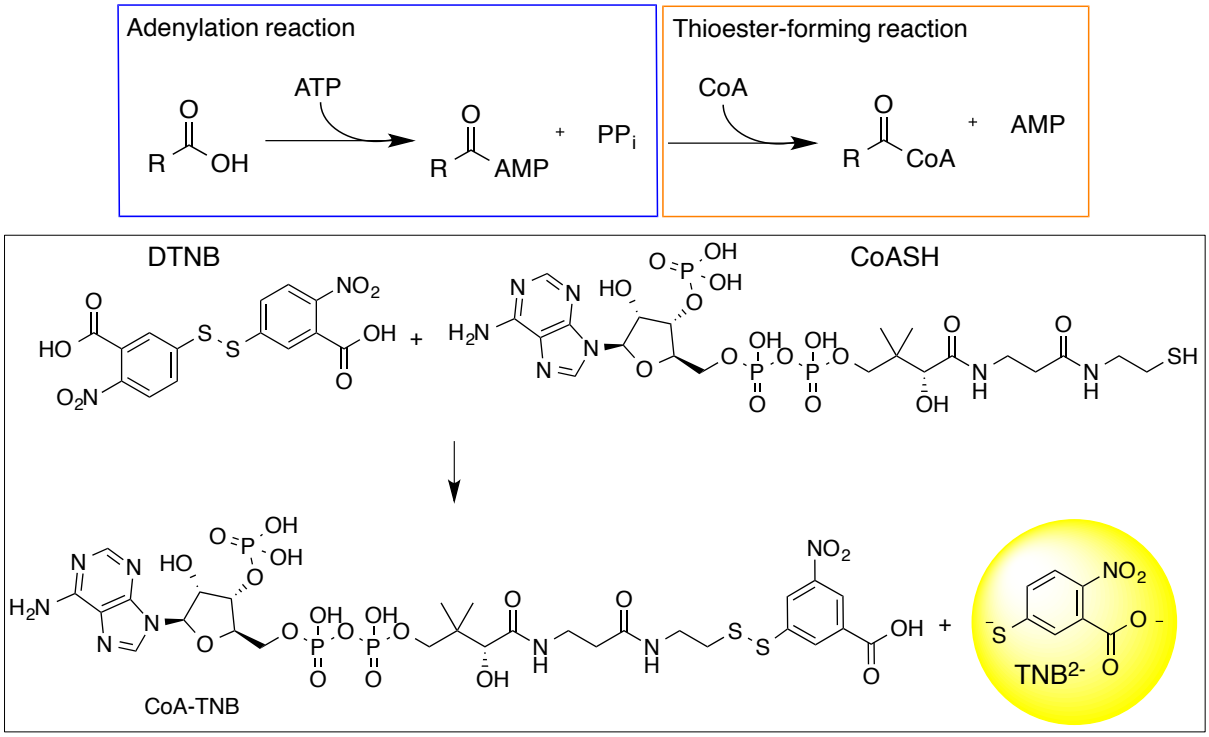


Figure 2-1: *Acyl-CoA ligation and the Ellman's Assay.* ***Top****: AcsA is an adenylate-forming enzyme that first activates the carboxylic acid with AMP. CoA is then added to the acid, releasing AMP.* ***Bottom****: The reaction of DTNB with free CoA produces a yellow-colored compound, TNB^{2-}. The formation of TNB^{2-} can be monitored at 412 nm which is inversely correlated to the formation of the acyl-CoA product.*

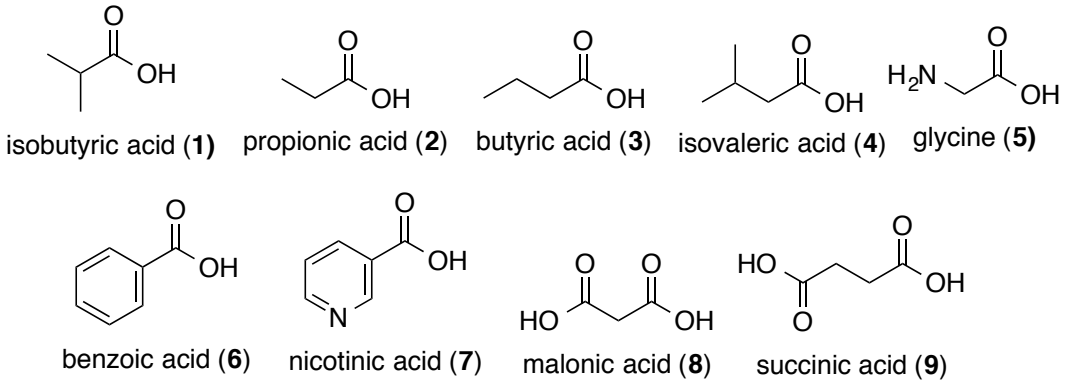


Figure 2-2: *Panel of substrates being tested with AcsA.*

2.2.1 Assay development

In order to determine the optimal conditions for the kinetic assay using DTNB, several experiments were set up varying each reaction component. First, the suitability of the assay was tested by measuring the absorbance at each concentration of CoA in a phosphate buffer. The relationship between the concentration of CoA and absorbance at 412 nm was linear between 0.1 mM to 2 mM when DTNB was provided in greater than 2-fold excess (data not shown).

The next step in developing this assay was to determine the working reaction conditions and characterize the limits of the assay with all reaction components necessary for the *in vitro* reactions of AcsA. First, a CoA standard curve was constructed in the presence of all reaction components (200 mM sodium phosphate, 2 mM $MgCl_2$, 100 mM $(NH_4)_2SO_4$, 5 mM ATP, 5 mM isobutyric acid, pH 7.2) without enzyme. Each reaction was quenched with an equal volume of DTNB in a solution of sodium phosphate and methanol to bring the sample to a total volume of 300 μL. The response was linear up to 3 mM CoA in this system (data not shown). To account for the need of methanol in the quenching solution (to deactivate AcsA) and to confirm that DTNB does not need to be provided in large excess with respect to CoA concentration, reactions were set up with varying ratios of CoA and DTNB in a phosphate buffered solution containing 50% methanol. The progress of the reaction was monitored by the absorbance at 412 nm over 5 minutes. The reaction proceeded at a linear rate over the entire five minutes, regardless of the ratio of the concentrations of CoA and DTNB (data not shown).

Based upon the limits of the plate reader detector, the maximum sample volume containing enzyme was determined to be 150 μL at a 1 mM concentration of CoA. At volumes and concentrations higher than this, the detector was fully saturated and the absorbance could not be measured. When concentrations of CoA needed to exceed 1 mM, new calibration curves were constructed and smaller reaction volumes were used.

2.2.2 Measuring Michaelis-Menten kinetics

Next, preliminary kinetic experiments at high concentrations of isobutyric acid established that 10 μg of AcsA provides a linear reaction velocity over 2 minutes, while 25 μg of AcsA provides a linear reaction over 5 minutes, as judged by the percent conversion of CoA to acyl-CoA determined via HPLC. A Michaelis-Menten curve for each substrate was then constructed (**Figure 2-3**) by carrying out the *in vitro* end-point assay at varying concentrations of each substrate (0.5 μM – 10 mM) with a fixed amount of enzyme (10 μg or 25 μg). The concentration of product formed during each reaction was determined by using a standard curve of the concentration of CoA vs. absorbance at 412 nm prepared under identical reaction conditions. The V_{max} and K_m were determined for each set of data using SigmaPlot. The k_{cat} for each substrate was calculated by dividing the V_{max} by the amount of enzyme, and the data are summarized in **Table 2-1.**

The wild-type enzyme appears to have highest specificity for isobutyric acid, based upon comparison of the K_m and k_{cat} with isobutyric acid to that of propionic acid and butyric acid. In comparison to similar enzymes like the malonyl-CoA synthetase MatB, which has a K_m of 2 μM for its natural substrate malonate[49], the affinity for isobutyric acid is rather low with a K_m 140 μM. Similar to AcsA, MatB also accepts several other substrates with lower efficiency (k_{cat}/K_ms for methylmalonate, ethylmalonate and allylmalonate are 175 $mM^{-1}*sec^{-1}$, 0.48 $mM^{-1}*sec^{-1}$and 0.87 $mM^{-1}*sec^{-1}$, respectively)[31]. Considering this, AcsA is a relatively slow enzyme that requires moderately high concentrations of substrate. Under the conditions used in the kinetics assay, kinetic parameters with isovalyric acid and benzoic acid were not determined because product was not detected.

There are some differences between the published values for the V_{max} and $k_{cat}/K_{m,}$ and the values determined here. The published work on AcsA describes the optimal enzyme activity resulting from reaction at 35 ℃ and 7.5 pH. The assay described here was performed in an incubator pre-set at 30 ℃, in solution buffered at 7.2 pH to accommodate the use of DTNB. Although the enzyme was reported to be stable at pH 6.5 – 9.0, only 85% activity was reported at 30 ℃. While the measured K_m of 0.14 mM is in complete agreement with the

reported K_m of 0.14 mM, the measured k_{cat} of 2.91 sec^{-1} is ~3-fold lower than the reported value of 10.64 sec^{-1}.

Table 2-1: Michaelis-Menten kinetics of WT AcsA

Substrate	Experimental			Published[48]		
	K_m (mM)	**k_{cat} (sec^{-1})**	**k_{cat}/K_m ($sec^{-1}*mM^{-1}$)**	**K_m (mM)**	**k_{cat} (sec^{-1})**	**k_{cat}/K_m ($sec^{-1}*mM^{-1}$)**
isobutyric acid	0.14 ± 0.03	2.91 ± 0.12	20.07	0.14 ± 0.07	10.64	76
propionic acid	1.03 ± 0.37	0.65 ± 0.07	0.63	1.10 ± 0.21	12.43	11.3
butyric acid	1.28 ± 0.19	0.46 ± 0.03	0.36	0.32 ± 0.04	15.04	47
ATP	-	-	-	0.22 ± 0.02	-	-
CoA	-	-	-	0.37 ± 0.02	-	-

A likely contributing factor for the reduced k_{cat} is the concentration of CoA that was used in the assay. While ATP was provided in excess at saturating concentrations (2 mM) based upon a K_m of 0.2 mM, CoA was not provided at fully saturating conditions (1 mM) based upon a K_m of 0.4 mM. The discrepancy in concentration of CoA is due to the limits of the assay and the reaction volume. In regards to the reaction mechanism, AcsA first catalyzes the carboxylic acid attack of ATP, loading AMP onto the acid and releasing pyrophosphate. Then, AMP is displaced by the nucleophilic attack of CoA, forming the acyl-CoA product. Under the conditions tested here, the measured K_m accurately reflects the recognition of the substrate by the enzyme. However, limiting the CoA could slow the second half of the reaction and cause a reduced k_{cat}.

In addition, there is evidence that nickel inhibits the activity of AcsA, as it was demonstrated that the activity of AcsA was reduced to 1.3% of the established activity in the presence of 1 mM $NiCl_2$[48]. AcsA was expressed in *E. coli* BL21 (DE3) with an N-terminal His-tag, and purified with Ni-NTA affinity chromatography. This method of expression and purification cannot be ruled out as a contributing factor for the overall decrease in reaction rate

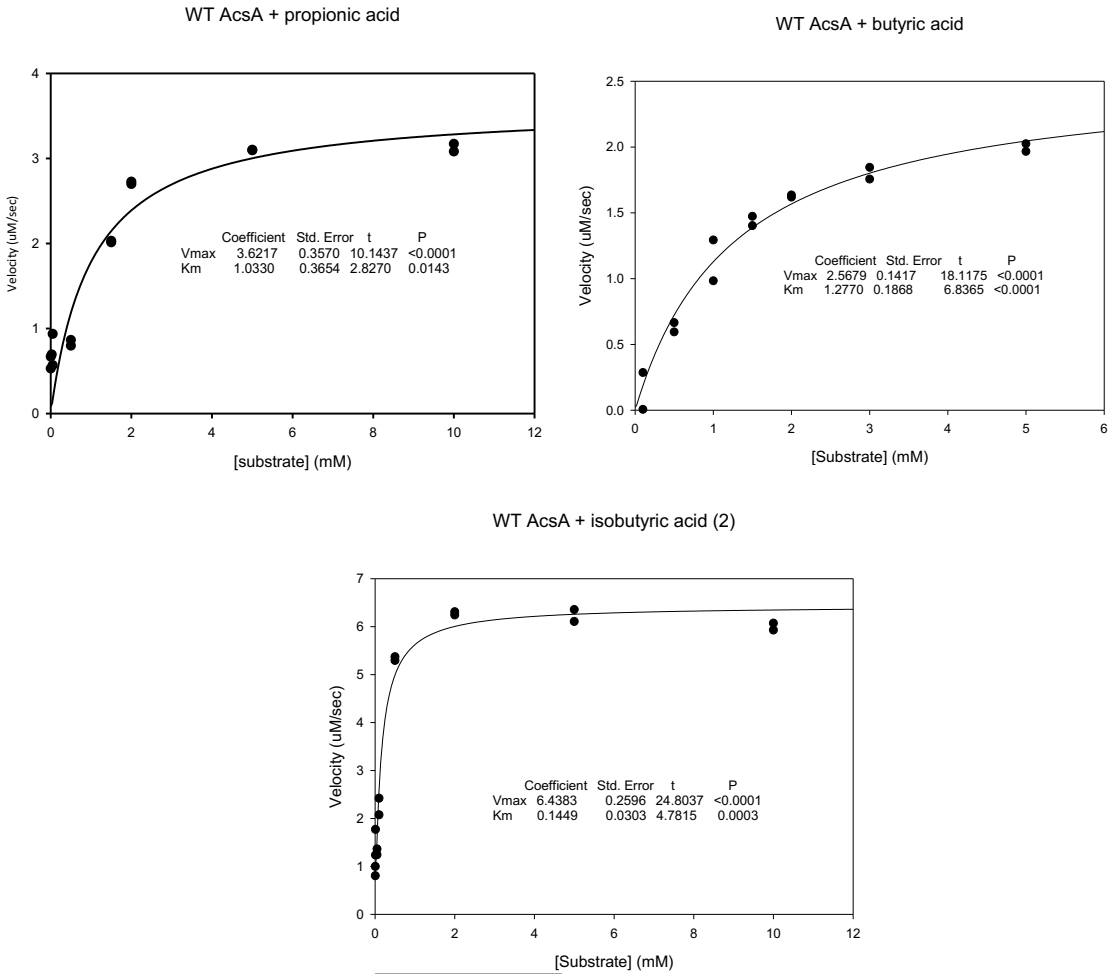


Figure 2-3: *Michaelis-Menten curves for the wild-type enzyme with top three substrates. The V_{max}, K_m and associate errors were calculated by SigmaPlot. Curves prepared for propionic acid, butyric acid and isobutyric acid.*

during kinetics studies with purified AcsA. The presence of the His-tag could also have a negative effect on the activity of AcsA. However, the effect of removing the His-tag was not tested.

2.3 AcsA Mutagenesis

The data in the preceding section confirm the broad specificity of AcsA but also highlights that there is significant room for improvement. For example, the activity towards several carboxylic acids is too low for use in practical application. Moreover, in order to utilize AcsA as an *in vivo* tool for starter or extender unit generation, the specificity needs to be shifted away from substrates like propionic acid and isobutyric acid. In the absence of a published crystal structure, and a lack of information regarding specificity determinants in the AcsA active site, the first goal was to identify which AcsA residues are involved in determining substrate specificity.

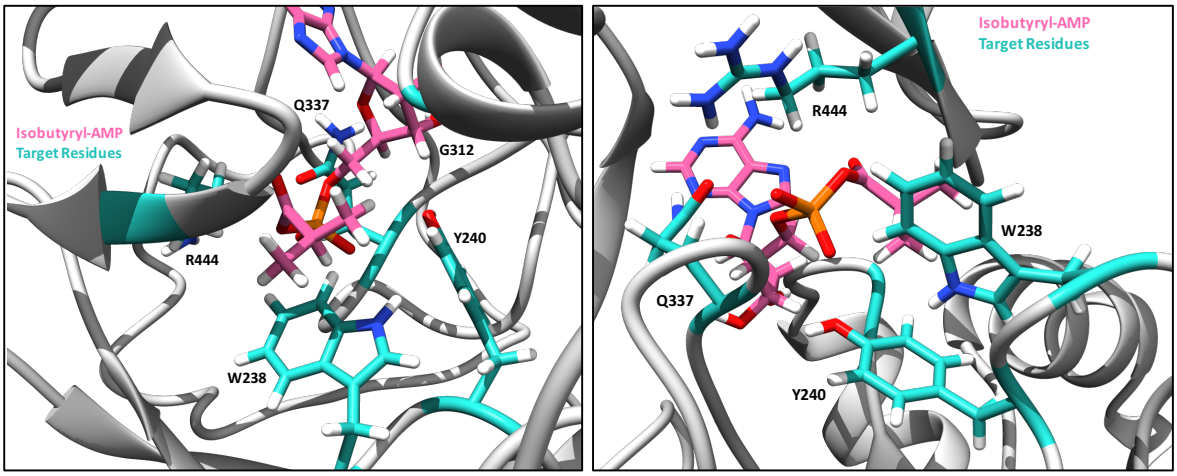


Figure 2-4: *Homology model of the AcsA active site with docked isobutyryl-AMP (in pink). Residues chosen for mutagenesis are highlighted in cyan.*

2.3.1 Constructing AcsA Mutants and saturation libraries

Using a homology model of AcsA (prepared using the I-TASSER server[50,51]), an isobutyryl-AMP intermediate was docked[52] to examine the potential active site conformation during the ligation portion of the reaction. In this configuration, several amino acids located within 4Å of the isobutyrate portion of isobutyryl-AMP were identified as targets for mutagenesis. These residues include: W238, Y240, G312, and R444 (**Figure 2-4**). Several other CoA ligases, including benzoate-CoA ligase, a human acyl-CoA synthetase, acetoacetyl-CoA ligase, succinyl-CoA ligase and malonyl-CoA synthetase, have been studied and characterized. In the interest of identifying trends to potentially support or refute whether the aforementioned residues play a role in substrate specificity or recognition in AcsA, the sequences of these other ligases were aligned with the sequence of AcsA (**Figure 2-5**). Residue G312 is well-conserved and likely provides the flexibility necessary for the catalytic loops. Residue R444 is conserved as a positively charged amino acid and is positioned proximal to the 5' phosphate of AMP, so it may serve as a catalytic residue. Residue W238 is conserved as an aromatic amino acid, and it appears to have the most interaction with the isobutyryl portion of isobutyryl-AMP.

To test the hypothesis that residues W238 and R444 are important in recognition of either isobutyric acid or ATP, a saturation library of each residue was created. From the W238X library, a W238A mutant was pulled out in the interest of assessing its role in substrate recognition and catalysis. As an alternative approach, the R444X library was screened using a modified version of the assay with DTNB to identify improved mutants or trends.

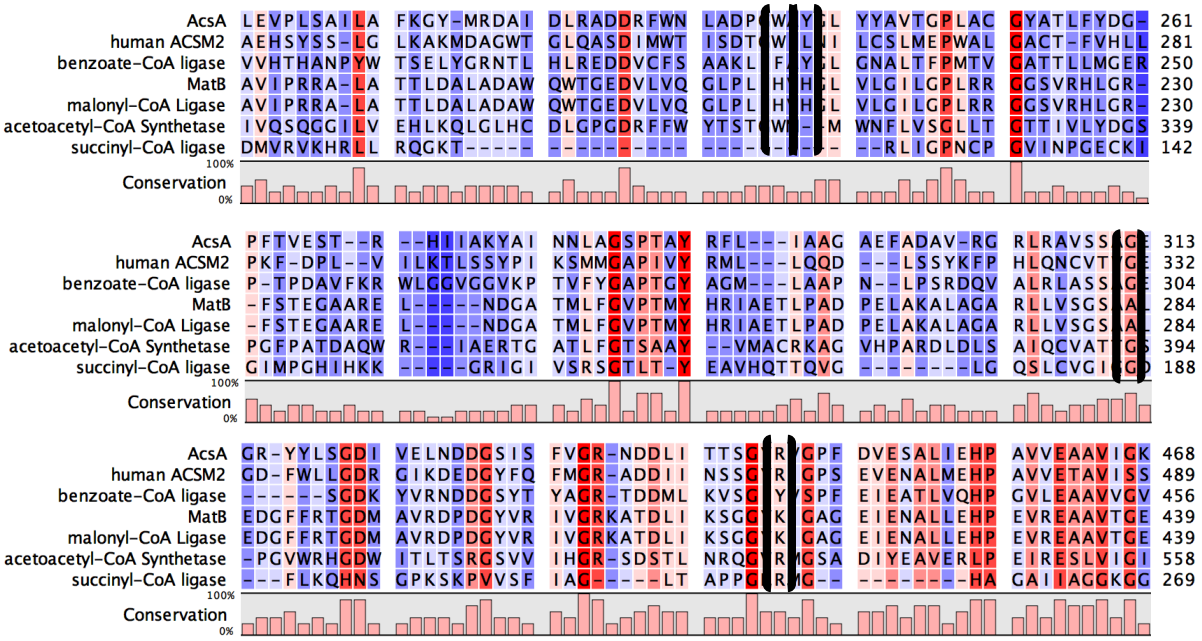


Figure 2-5: *Sequence alignments of AcsA and related CoA ligases. The sequence of AcsA is the top sequence in each block. Target residues are outlined in black brackets. From top to bottom: W238, Y240, G312, R444. Coloring indicates percent conservation of each residue, dark blue indicates 0%, dark red indicates 100% conservation.*

2.3.2 Analysis of the AcsA W238A mutant

The W238A mutant was tested against the panel of substrates alongside WT-AcsA. In a single set of reactions, W238A showed a negligible 1.7% increase in activity with isobutyric acid, but a seemingly significant reduction in activity with all other substrates in the panel, including propionic acid (**Figure 2-6**). This result was unexpected, as replacing a large aromatic tryptophan with a small an uncharged alanine residue would likely increase the size of the active site, which in theory would make room for larger substrates. The very large reduction in activity with benzoic acid is expected for this mutation, however, as any pi-pi stacking that could take place between the benzene and tryptophan is abolished. It is important to note that no error in the conversion rate is reported, and so there is no way to determine the significance of these results without repeating the experiment in replicate.

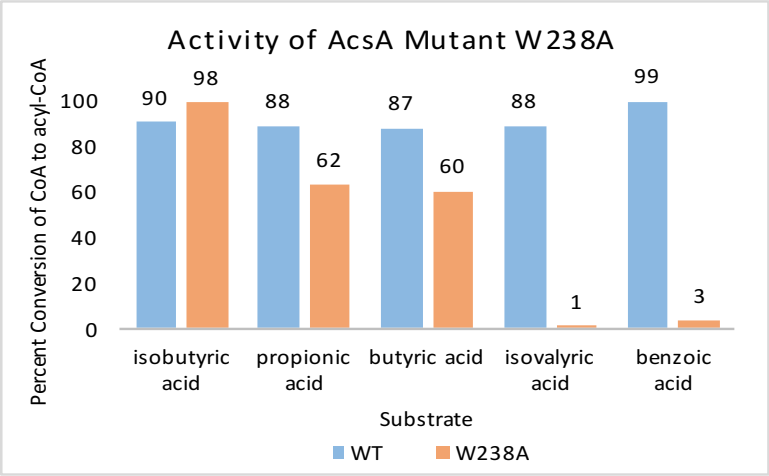


Figure 2-6: *Activity of wild-type vs mutant AcsA with select panel substrate. Percent conversion was calculated as the conversion of CoA to acyl-CoA via HPLC.*

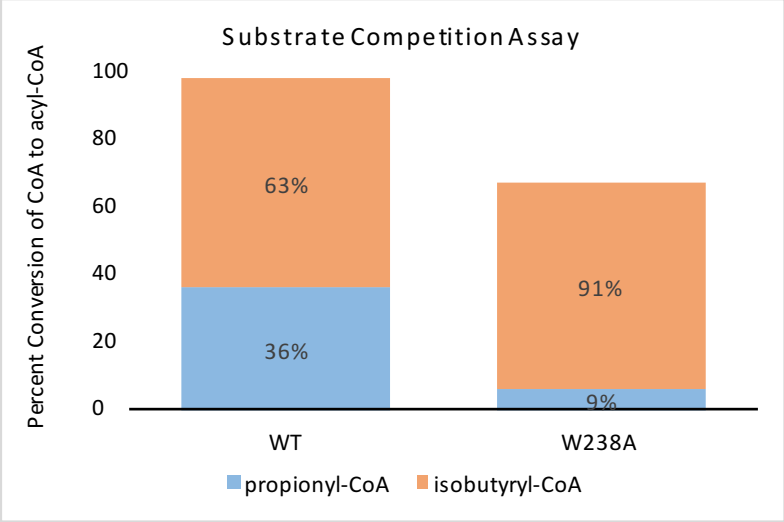


Figure 2-7: *Substrate competition assay with wild-type AcsA and W238A AcsA mutant. Y-axis features total percent conversion of CoA to acyl-CoAs, whereas the bar features the product ratio.*

To follow up this result, a substrate competition assay was performed with W238A, propionic acid and isobutyric acid. Using the experimentally-determined K_m for the WT-AcsA with propionic and isobutyric acid, each substrate was provided at saturating concentrations.

The reaction was stopped after 5 minutes to reflect the linear range established for the wild-type enzyme. However, in 5 minutes no products were detected with the W238A mutant. The experiment was repeated, and the reaction was stopped after one hour. While the wild-type reactions were approaching 100% conversion of CoA to acyl-CoA in one hour, the reactions with the W238A mutants were significantly slower. The WT AcsA converted 98% of CoA to acyl-CoA (37% propionyl-CoA, 63% isobutyric acid), while the mutant W238A converted only 68% of CoA (9% propionyl-CoA, 91% isobutyryl-CoA) (**Figure 2-7**). Based upon the reduced fraction of propionyl-CoA, the mutant clearly shows a shift in specificity away from propionic acid. The kinetics with isobutyric acid were determined using the method described previously. The K_m for isobutyric acid increased significantly to ~ 2.7 mM and the k_{cat} decreased significantly to ~ 0.1 sec^{-1}. The total conversion rate of 68% in one hour and the reduced k_{cat} and K_m with isobutyric acid is indicative that this specificity shift comes at the expense of the enzyme efficiency. The kinetics with propionic acid were not determined with W238A, as no significant product formation could be detected at the the lower concentrations of substrate used in the assay (0.05 mM – 10 mM). Therefore, not enough data points were available to construct a full Michaelis-Menten curve.

The apparent increase in selectivity for isobutyric acid vs. propionic acid was not an expected result from this experiment. However, the overall decrease in the reaction rate with both substrates was expected as the W to A mutation is expected to increase the volume of the active site. This likely increases the number of conformations the substrate can adopt during the course of the reaction due to a less restrictive active site architecture.

Table 2-2: Michaelis-Menten kinetics of W238A mutant

	W238A AcsA			**WT AcsA**		
Substrate	**K_m (mM)**	**k_{cat} (sec^{-1})**	**k_{cat}/K_m ($sec^{-1}*mM^{-1}$)**	**K_m (mM)**	**k_{cat} (sec^{-1})**	**k_{cat}/K_m ($sec^{-1}*mM^{-1}$)**
isobutyric acid	2.69 ±0.99	0.107	0.098 ±0.015	0.14 ±0.03	2.91	20.07 ±0.11

2.3.3 Screening an AcsA R444X library

The R444X saturation library was screened using a modified version of the assay using DTNB to measure free CoA. A total of 88 clones from the library were expressed in a 96-well plate and transferred to clear 96-well plates to be screened against the panel of substrates. Each well contained 75% cell lysate, 1 mM CoA, 5 mM acid, 2 mM ATP and the appropriate reaction buffer. The plates were incubated for 20 minutes at 30 ℃ to allow for the ligation reaction to proceed. The absorbance at 412 nm was recorded immediately after quenching each reaction with an equal volume of DTNB in a solution of sodium phosphate and methanol. The absorbance values for the mutants with butyric acid were compared to the absorbance values for the wild-type AcsA and a negative control (an empty pET28a plasmid) with butyric acid. This relatively poor substrate was chosen to test the ability of saturation mutagenesis at this active site residue to improve activity from a very low background.

The average absorbance, the standard deviation and percent standard deviation were calculated for the controls on each plate. The average absorbance for the negative control with butyric acid is 2.47 ± 0.31 mAu and the average absorbance for the wild-type control is 2.58 ±

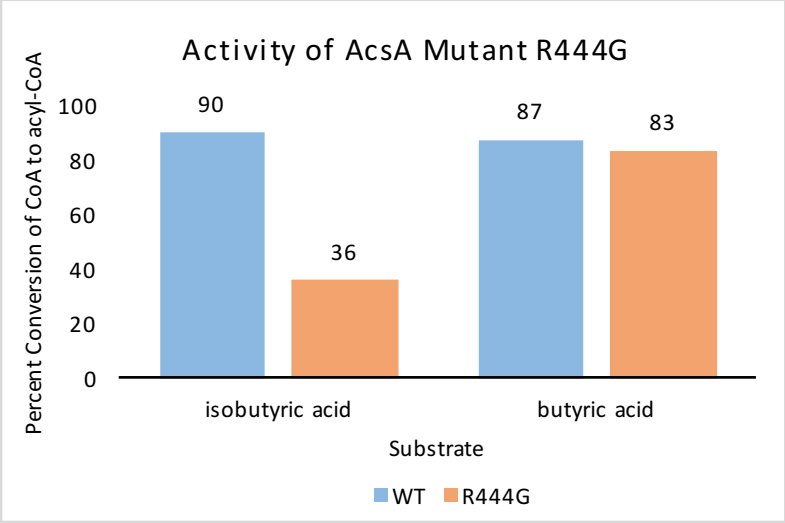


Figure 2-8: *Activity of R444G AcsA mutant vs. WT AcsA. The activity was measured as percent conversion of CoA to acyl-CoA via HPLC.*

0.18 mAu. Because the difference between the negative control and the wild-type control was within error for butyric acid, the activity of a mutant AcsA library member needs to be improved by several fold in order to be detected in the end-point assay. Despite this observation, two clones with the lowest absorbance (11A, 2.10 mAU; 2D, 2.20 mAu) were taken to a second round of screening to confirm the significance of the lower absorbance values and evaluate the accuracy of this screen.

In the second round of screening, clone 2D still exhibits a low absorbance value in comparison to the wild-type. This mutant was sequenced and identified as R444G. R444G was expressed, purified and used in overnight reactions (~16 hours) at 30 ℃ with 5 mM butyric acid or isobutyric acid, 2 mM CoA, 3 mM ATP and the appropriate buffers. Under these conditions, when the reaction is no longer under kinetic control, this mutant catalyzed 83% conversion of CoA to butyryl-CoA vs. 36% conversion of CoA to isobutyryl-CoA (**Figure 2-8)**. In contrast to the wild-type enzyme, which catalyzed 87% and 90% conversion, respectively, the R444G mutant lost nearly 2/3 activity with isobutyric acid but remained nearly as active with butyric acid. It is important to note that in contrast to the overnight reaction, the high-throughput screening in cell lysate was performed over a 20-minute time period and is likely under kinetic control based upon the wild-type data. Any mutants from the high-throughput screen should therefore have improved K_m for the substrate used in the screening. As R444G did not exhibit a lower absorbance value in the high-throughput screening against isobutyric acid, it is expected that this clone will be slower or less active with isobutyric acid. Despite these observations, no firm conclusions can be drawn about the R444G mutant without further kinetic evaluation.

2.4 Conclusions

The substrate scope of AcsA from *P. chlororaphis* has been probed with a variety of carboxylic acids, revealing the formation of acyl-CoA product with 5/9 panel members. Using an indirect method for measuring the production of acyl-CoA, the kinetics of the wild-type and W238A mutant have been calculated under a new set of reaction conditions.

The mutation W238A was shown to have an effect on the preference of AcsA for isobutyric acid and propionic acid as demonstrated by a substrate competition experiment. Although the W238A mutation negatively affected the reaction rate, the discrimination against propionic acid by the mutant AcsA supports the conclusion that AcsA can be engineered for new or different substrate specificities. The W238A mutation also resulted in a decreased catalytic efficiency of the enzyme, suggesting that this residue is in fact located in the active site and may interact with the carboxylic acid substrate during catalysis.

The residue R444 was also evaluated for its role in catalysis. The R444G mutation was pulled out of a saturation library using a high-throughput screening technique screening for activity with butyric acid. This mutant showed a substrate specificity shift, being less active with isobutyric acid and just as active with butyric acid in comparison to the wild type AcsA. No kinetic parameters for R444G were determined. Without follow-up studies, the hypothesis that residue R444 is involved in recognition and binding of ATP and acyl-AMP intermediate cannot be confirmed based up on the observed change in activity of R444G with isobutyric acid. The drastic reduction in size and charge of the R444G mutation makes it difficult to determine the cause of the reduced activity with isobutyric acid based upon the position of this residue in the active site. Further kinetic evaluation of the R444G mutant needs to be done.

These results have provided the preliminary evidence necessary for rational mutations to be made in AcsA to engineer new substrate specificities. Now that the importance of W238 has been confirmed with respect to carboxylic acid selection, mutants such as W238H, W238F, and W238Y should be explored to address the significance of aromatic residues at this position. Based upon the results of R444G, further mutagenesis needs to be performed in order to confirm the role of residue R444. Using the screening methods described, other residues such as G312 and Y240, can be easily screened with a panel of substrates.

2.5 Materials and methods

Materials. The following plasmids were used: pET28a, purchased from Novagen; pET-AcsA, constructed via subcloning of AcsA into pET28a. Gene "AcsA" (see Appendix B for nucleotide sequence) was synthesized by GenScript, USA. Unless otherwise noted, all

enzymes were purchased from NEB, and all chemicals were purchased from Fisher and Sigma-Aldrich. All primers were purchased from Integrated DNA Technologies (IDT).

Expression and purification of AcsA and mutants. Codon-optimized acyl-CoA synthetase (AcsA) from *Pseudomonas chlororaphis* was sub-cloned from pUC57-Kan into pET28a via *Eco*RI and *Hind*III restriction sites. The resulting plasmid pET-AcsA was confirmed by single-cut digestion and gel analysis, resulting in a single band at ~7kbp as expected. pET-AcsA was transformed into *E. coli* BL21 (DE3) cells via heat shock plated on Kan-selective LB agar media. One colony was selected and used to inoculate 10 mL LB media with 50 μg/uL Kan and incubated in an orbital shaker for ~15 hours at 37 ℃ at 270 rpm. The entire 10 mL culture was used to inoculate 1 L of 2xYT media (purchased from Fisher and prepared according to label) in a 2.8 L flask. The culture was incubated for 2 hours at 37 ℃ and 250 rpm in an orbital shaker and cooled to 28 ℃ over a 2-hour period. Then 1 mM IPTG was added to induce protein expression, and the culture was incubated at 28 ℃ at 250 rpm for ~18 hours. The culture was centrifuged for 30 minutes at 4 ℃ at 4000 rpm and the resulting pellet was used in SDS-PAGE analysis to confirm expression prior to purification.

Cells producing AcsA were lysed via sonication. The lysed cells were centrifuged at 15,000 rpm for 90 minutes. The clarified lysate was purified by fast-protein liquid chromatography, using a 1 mL nickel affinity column and 20 mM sodium phosphate/500 mM sodium chloride buffers at pH 7.8 with imidazole at either 20 mM (buffer A - equilibration) or 200 mM (buffer B - elution). The following protocol was used for purification: (flow rate = 2 mL/min) 5 mL 100 % buffer A, 5 mL sample injection with 100 % buffer A, 10 mL buffer A, 50 mL linear gradient from 100 % buffer A to 100 % buffer B, 10 mL 100 % buffer B, 10 mL linear gradient 100 % buffer B to 100 % buffer A, 10 mL buffer A re-equilibration. Purity and identify were assessed with SDS-PAGE. AcsA was identified as a single band at approximately 60 kDa. FPLC fractions containing single band at 60 kDa were pooled and concentrated to ~ 1 mL with a 30 kDa molecular weight cutoff spin filter by centrifugation at 4000 rpm. Concentrated protein was buffer exchanged into a protein storage buffer, containing 10 %

glycerol at pH 7. Protein concentration was determined using standard Bradford Assay, and aliquots of protein were stored at -80 ℃.

Constructing site-directed mutant and W238/R444 saturation libraries of AcsA. Using the wild-type AcsA in pET28a as template DNA, mutants were made using "round the horn" PCR with commercially phosphorylated primers. Each 10 μL reaction contained the following: ~10 ng template DNA, 2 mM mix dNTPs, 0.5 mM primer, 1X Phire buffer, 2 μL Phire II HF polymerase (NEB) and H_2O to volume. Cycling conditions were designed via the protocol available on NEB's website for Phire polymerase, adjusted for the annealing temperature of each primer set. After the cycling was complete, the entire reaction was run on a 0.7% agarose gel to purify the bands of interest (~7 kb). The bands of interest were cut out from the gel and the DNA extracted from the agarose. Multiple ligation reactions were set up using various volumes of DNA and T4 Ligase. Each ligation was left at room temperature for ~16 hours and ethanol-precipitated to purify. The precipitated DNA was chilled on ice before transforming into BL21 (DE3) electro-competent cells (Lucigen). Transformations were plated on kanamycin agar and incubated at 37 ℃ for ~16 hours.

Several single colonies were selected from each transformation and grown in 5 mL LB/kanamycin overnight at 37 ℃ for DNA sequencing. Of the 5 colonies grown and sequenced from the W238X library, two sequenced colonies corresponded to a W238A mutation. This mutant was expressed/purified in the same manner as described previously, and used in ligation reaction and competition assays to assess the importance of residue W238. The R444X library was sequenced in the same manner, and all 5 colonies picked for sequencing represented an individual mutant. This library was screened for activity against the panel of substrates using a high-throughput version of the Ellman's assay.

End-Point Assay for the determination of AcsA Kinetics. Ellman's reagent (5'5'-dithiobis(nitrobenzoic acid)) was prepared at a concentration of 7.8 mM in a 50/50 solution of 0.2 M sodium phosphate and methanol at pH 7.2. Carboxylic acid substrates were prepared via serial dilution at concentrations ranging from 0.005 mM to 10 mM. Each reaction contained 1 mM CoA, 3 mM ATP, 100 mM $(NH_4)_2SO_4$, 200 mM phosphate and 2 mM $MgCl_2$. In a 96-well plate, each concentration of carboxylic acid was added down one column (prepared in

duplicate). To each well in the column, a master mix containing the above reaction mixture plus 10 – 50 μg of enzyme was added to start the reaction in a total volume of 75 μL. The plate was incubated at 30 ℃ for an appropriate amount of time, as determined by a time course assay performed in the same manner using an approximated saturating concentration of carboxylic acid. After incubation, 75 μL of Ellman's reagent was used to quench each reaction well and the absorbance at 412 nm was read immediately (within 2 minutes) using Bio-Tek Hybrid Synergy 4 plate reader. A standard curve was prepared (in triplicate) on each plate using a master mix of the same reaction component, minus the carboxylic acid substrate. The absorbance at 412 nm was plotted against the concentration of CoA, and using this line, the amount of CoA remaining in each reaction was calculated. From the amount of CoA remaining, the product formation at each concentration of carboxylic acid was calculated and used to determine the product formation/time (velocity). This data was then plotted against the concentration of carboxylic acid to construct a Michaelis-Menten curve using SigmaPlot.

High-throughput lysate screening with Ellman's reagent. Validation of the lysate 96-well plate screening was performed using wild-type AcsA and an empty pET28a vector. Single colonies of BL21 (DE3) cells containing either AcsA in pET28a or empty pET28a were picked and grown in 10 mL LB/kan at 37 ℃ to generate seed cultures for expression. In a 2 mL 96-well, 490 μL of 2xYT broth was inoculated with 10 μL of starter culture. The plate was incubated at 350 rpm in an orbital shaker at 37 ℃ for 3 hours to reach an OD_{600} of 0.6 – 0.8 (reproducibility of culture conditions determined in a separate experiment, not shown).

Upon reaching OD, the shaker was lowered to a temperature of 28 ℃ and isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to each well at a concentration of 1 mM to induce protein expression. After 16 hours of expression at 28 ℃ and 350 rpm, the plate was centrifuged at 4000 rpm for 10 minutes to pellet cells. The supernatant was drained and cells were resuspended in a lysis buffer containing 2 mg/mL lysozyme (Sigma-Aldrich) and 0.3 μl/mL of Benzonase® Nuclease (Sigma-Aldrich).

The plate was subjected to a single freeze-thaw at -80 ℃ and centrifuged again at 4000 rpm for 15 minutes to pellet cell debris. The supernatant from each well was transferred (50

μL) to a clear 96-well plate for analysis with the following mix: 200 mM phosphate buffer, 2mM $MgCl_2$, 1mM CoA, 2mM ATP at pH 7.2. For validation purposes, 5mM isobutyric acid (the "preferred" substrate for AcsA) was used in each reaction well.

The plate was incubated at 30 ℃ for 15 minutes, at which time 75 μL of "Ellman's reagent" (prepared as described previously) was added to each well to quench the reaction. The absorbance at 412nm was read immediately. The standard deviation of all control readings was approximately 15% of the mean absorbance. An approximate 30% difference was observed between the average absorbance for the empty pET28a versus the wild-type AcsA.

To screen the R444X library, 88 single colonies were picked with a sterile toothpicks and used to inoculate 200 μL seed cultures in each well of a 96-well plate. The first column of the expression plate was reserved for 4 replicates of empty pET28a and wild-type AcsA as a control. The seed plate was prepared as described above. Three identical expression plates were prepared, transferring culture from each well of the seed plate to the corresponding well in each expression plate. Cultures were grown to the appropriate OD_{600}, induced with IPTG, expressed and otherwise handled as described previously.

To screen against the panel of substrates, 9 clear 96-well reactions plates were prepared, each containing 5 mM of the chosen substrate in all wells. Using the same reaction conditions described above, 50 μL of lysate was transferred to the corresponding well of each substrate plate. After quenching each well/plate, the absorbance at 412 nm was read immediately. The average and standard deviation of the control wells were calculated. Hits were identified by sorting the absorbance values from lowest to highest. Only 5 clones with the lowest absorbance values were used in a second round of expression/screening (performed in the same manner) to ensure accuracy. The hits that passed the secondary screening were then grown in 5 mL cultures and the DNA was isolated and sequenced.

AcsA ligation reactions and measuring percent conversion/enzyme activity. Each ligation reaction was set up and run overnight (~16 hours) at 30 ℃. Using the same concentration of $(NH_4)_2SO_4$, phosphate buffer and $MgCl_2$ at pH 7.2 as the kinetics assay, the reactions were run at 10 mM carboxylic acid, 10 mM CoA, 15 mM ATP and 100 μg of enzyme. After the incubation time, the reactions were quenched with equal volume of methanol and

centrifuged at 10,000 rpm for 10 minutes to clear insoluble material. Then, 25 μL of supernatant was analyzed via RP-HPLC-UV-Vis (Varian Pro-Star, Phenomenex C18 column) using the following separation: a linear gradient from 100% of 0.1% TFA in H_2O (Buffer A) to 46% methanol without TFA (Buffer B) over 15 minutes, 85% of Buffer B for 3 minutes, and a linear gradient returning to 100% Buffer A over 2 minutes. The absorbance at 254nm was monitored over the 20-minute method. A calibration curve of CoA (retention time = 8.6 minutes) was constructed using this separation method; the peak area at 8.6 minutes was plotted against CoA concentration, producing a linear curve up to 3mM CoA.

The percent conversion of CoA to product was calculated for each product peak. The percent conversion of each panel member was compared for the wild type AcsA and each mutant of interest. The percent change in activity was calculated by determining the percent change in the conversion with each substrate.

Substrate competition assay. To assess the substrate preference of the W238A AcsA mutant, a competition assay was performed using isobutyric acid and propionic acid. Based upon the experimentally determined K_m values for isobutyric acid and propionic acid (0.73 mM and 1.03 mM, respectively), each substrate was run at an approximately saturating concentration (10 x K_m). The reaction was set up with the following components: 7 mM ATP, 5 mM CoA, 100 mM $(NH_4)_2SO_4$, 200 mM phosphate buffer, 2 mM $MgCl_2$, 50 μg of enzyme (wild type or mutant) and H_2O to a volume of 100 μL. Each reaction was incubated at 30 ℃ for 5 minutes, quenched with an equal volume of methanol, and centrifuged at 10,000 rpm for 10 minutes. The supernatant was analyzed via RP-HPLC using the aforementioned method and buffer system.

The total conversion of CoA to products was calculated using the total peak area of CoA (RT = 8.6 minutes), propionyl-CoA (RT = 12.4 minutes) and isobutyryl-CoA (RT = 15.3 minutes). Of the total percent conversion, the percent conversion of each substrate was calculated out of the total area of the two product peaks. At a reaction time of 5 minutes, the reaction with W238A mutant had nearly zero percent conversion. Thus, the experiment was repeated with a 1-hour incubation time. After a 1-hour reaction, a sufficient total conversion was achieved.

CHAPTER 3
In vivo polyketide engineering

3.1 Introduction

3.1.1 Use of native polyketide producing bacteria as hosts

The use of heterologous hosts, such as *Escherichia coli*, has enabled much of the work in PKS engineering through easy host manipulation and the availability of a wealth of genetic knowledge. In addition, metabolic engineering of alternative hosts has allowed for the heterologous expression of PKS enzymatic machinery to levels suitable for production of polyketides via fermentation[53–55]. However, the road to engineering alternative hosts for the production of polyketides is a long one, even with ever-improving advancements in genetics. This places time constraints on the expansion of PKS engineering beyond the traditionally studied systems to others, such as FK520/506, rifamycin B, and rapamycin. Many polyketides are produced by *Actinomyces,* a class of genetically diverse soil-dwelling Gram-positive bacteria. The life-cycle of most *Actinomyces* strains includes sporulation, which complicates the culturing process but allows for very easy storage and manipulation of the spores (**Figure 3-1**). The high GC content and number of repeats in the genomes of *Actinomyces* renders genetic manipulation using traditional methods, such as homologous recombination, inefficient or troublesome.

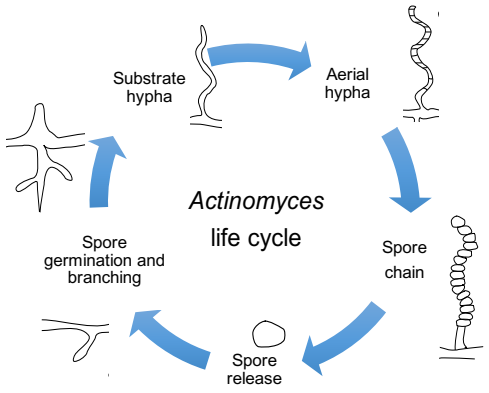


Figure 3-1: *Life cycle of sporulating Actinomyces.*

3.1.2 Genome editing with CRISPR-Cas9

Clustered regularly-interspaced short palindromic repeats (CRISPRs) are repetitive sequences of DNA found within the genomes of many members of Bacteria and Eukaryota, including *Actinomyces*. The bacterial CRISPR system functions in a manner analogous to that of the human immune system and antibody production: invading DNA is recognized, cut into small spacer sequences, and stored within the host genome in a series of repeats[56]. Upon future exposure of the cell to the foreign DNA, the spacer sequences are transcribed and processed into CRISPR RNAs (crRNAs) that recruit a series of CRISPR-associated (Cas) proteins which either contain nuclease domains (Cas9) or recruit a cascade of nucleases (Cas3/Cascade). In the Type II CRISPR-Cas9 system isolated from *Streptococcus pyogenes*, the hybrid CRISPR-Cas9/crRNA RNA-protein complex scans along the double-stranded DNA within the cell until the spacer RNA sequence base pairs with a homologous DNA sequence. If the DNA sequence contains the "trigger sequence" at the 5' end, which is known as the protospacer-adjacent motif (PAM), then the Cas9 endonuclease domains are activated, causing a double-stranded break in the DNA (**Figure 3-2**).

The protein Cas9 from *S. pyogenes* has proven a useful tool in genome engineering, as it can be targeted to any area of the genome by providing a spacer sequence that is immediately adjacent (5') to the "NGG" PAM sequence[57]. As long as the PAM is present and the 20 base

pair spacer sequence is hybridized with the target gene, Cas9 endonuclease will catalyze a double-stranded DNA break (DSB). In bacteria, a DSB in the DNA leads to cell death, as there typically are no active DNA repair pathways in bacterial cells. Therefore, when utilizing CRISPR-Cas9 as a tool for genome editing in bacteria, there is an inherent screening method to determine whether or not there is efficient targeting; a decrease in the viable cell count after transformation/induction of the CRISPR-Cas9 plasmid indicates Cas9 is actively targeting the genome. In order to utilize CRISPR-Cas9 as a tool for genome editing in bacteria, a homology-directed repair (HDR) template must provided to the cells to repair the DSB. Point mutations in the repair template are likely to be transferred to the host genome, allowing for the introduction of specific mutations to the targeted region. The highly specific, yet adaptable, nature of the CRISPR-Cas9 system permits its use for targeted genome editing of many cell types, potentially including polyketide-producing bacteria[58,59].

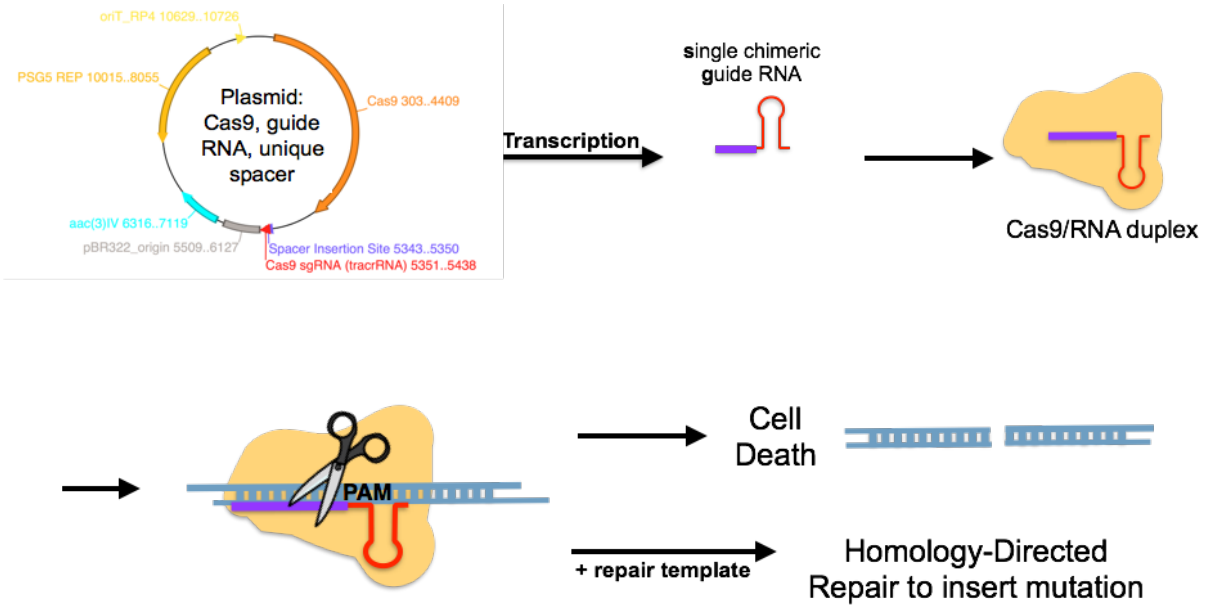


Figure 3-2: CRISPR-guided DNA cleavage for genome editing with Cas9.

3.2 Biosynthesis of FK520, rapamycin and rifamycin

Due to successes in engineering DEBS and Pik AT domains, our attention has turned towards applying AT engineering to other PKS systems. In the interest of drug discovery and polyketide analog generation, the following three polyketides were chosen: FK520, rapamycin and rifamycin B. These three polyketides are produced by three different *Actinomyces* strains that can be cultured in a laboratory setting, and for which the PKSs have been fully sequenced. Semi-synthetic derivatives of these molecules have been realized, some of which in the case of the synthetic rifamycin derivative, rifampicin, exhibit more favorable solubility or other properties.

Prior to engineering PKSs *in vivo* in the native hosts, the baseline of polyketide production by these organisms must be established and an extraction procedure must be optimized. The three strains must be cultured slowly to "turn on" secondary metabolite production of the mycelia in the stationary phase. From these cultures, the polyketide products of interest must be extracted, concentrated, and quantified. Information regarding the natural production of rifamycin, rapamycin and FK520 is essential before any *in vivo* PKS modifications can be made. Thus, before targeting the AT domains of the PKSs of interest with CRISPR-Cas9, fermentation analysis of wild-type *S. hygroscopicus subsp. ascomyceticus*, *S. rapamycinicus,* and *A. mediterranei* has been performed.

3.2.1 Analysis of fermentation extracts of *S. hygroscopicus subsp. ascomyceticus*, *S. rapamycinicus* and *A. mediterranei*

The goal of this work is to explore the possibility of making mutations in the PKS genes *in vivo* in the native FK520, rifamycin and rapamycin-producing hosts. In order to do so, the natural products must be successfully extracted from the fermentation culture and detected in a reliable and quantitative method. Therefore, the three polyketide-producing hosts were grown in the appropriate complex media for three days under conditions conducive to polyketide biosynthesis. After three days of growth, the mycelia were removed from the culture and the broth was extracted, concentrated and analyzed by LC-ESI-MS. Three commercial standards of rifamycin B, FK520 and rapamycin were also analyzed for method development.

The resulting mass spectra were analyzed for formation of FK520 $[M+NH_4]^+$, rapamycin $[M+Na]^+$ and rifamycin $[M+H]^+$ (**Figure 3-3**). The extracted ion count for the expected FK520 $[M+NH_4]^+$ was low when compared to the extracted ion count for the FK520 $[M+Na]^+$, and therefore the latter was used instead. The calculated mass of FK520 $[M+Na]^+$ is 814.4712 amu and the observed mass was 814.4714 amu, with a retention time of 17.1 minutes. The retention time and the ppm difference of 0.3 as compared to the standard, indicates that FK520 can be detected when extracted from fermentation cultures with confidence. The expected rapamycin $[M+Na]^+$ ion has a calculated mass of 936.5443 amu, and an observed mass of 936.5444 amu with a retention time of 17.97 minutes. The retention time and ppm difference of 0.1 as compared to the standard, indicates that rapamycin can be detected when extracted from fermentation cultures with confidence. The extracted ion count for the expected product rifamycin $[M+H]^+$ was not sufficient to definitively say that this product has been extracted and detected from the culture, although the retention time of 10.63 minutes provides some evidence for the product.

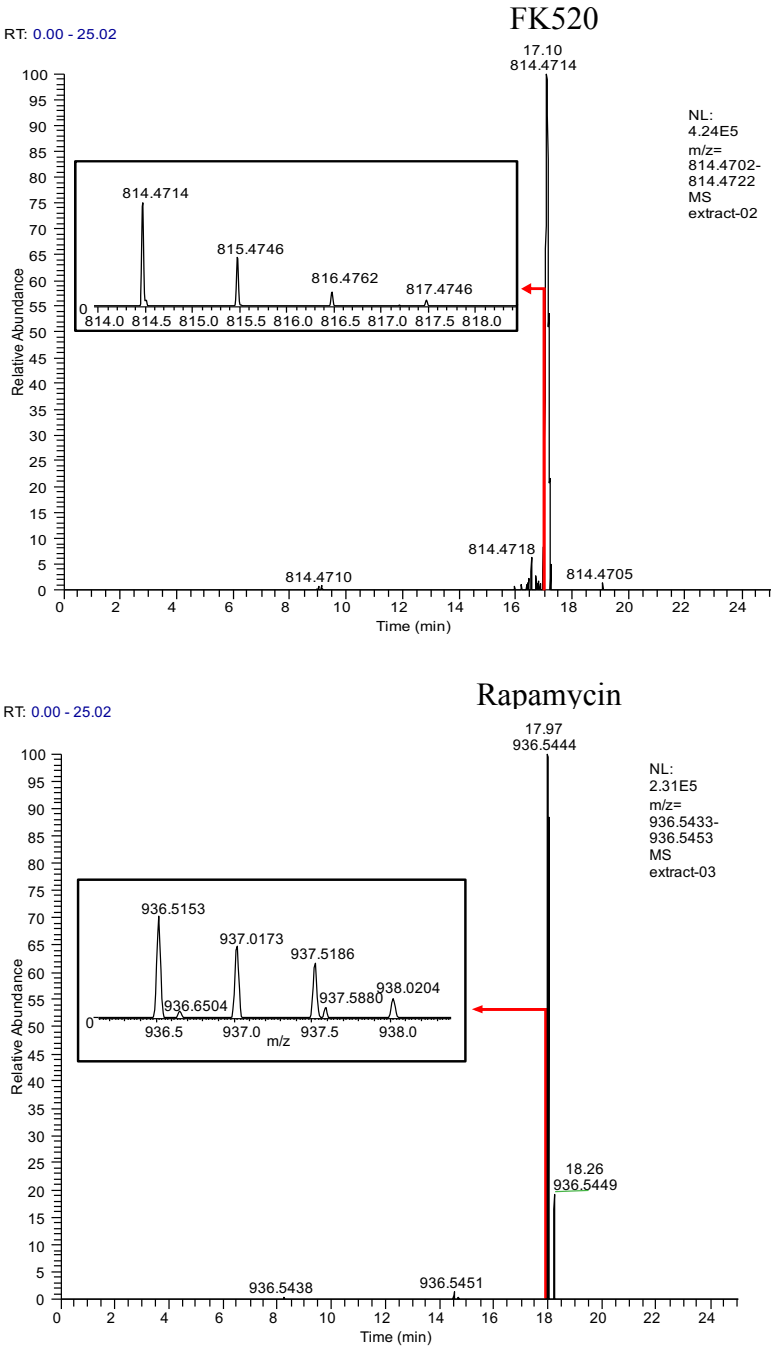


Figure 3-3: *LC-MS analysis of in tact masses of FK520 and rapamycin. Analysis was performed on a ThermoFisher Exactive Plus LC-MS. FK520[M+Na] calculated mass = 814.4712; observed mass = 814.4714; Δppm=0.3. Rapamycin[M+Na] calculated mass = 936.5443; observed mass = 936.5444; Δppm=0.1*

3.3 CRISPR-Cas9-mediated AT mutagenesis *in vivo*

Engineering PKSs in the natural polyketide-producing hosts is an attractive goal for the biosynthesis of polyketide analogs. While traditional genome editing methods, such as homologous recombination, can be used to edit the genomes of *Actinomyces*, the low rates of success and the lengthy processes are unattractive for making PKS mutations *in vivo*. The adaptable and highly specific CRISPR-Cas9 system is an attractive alternative for editing the genomes of polyketide-producing bacteria. Specifically, CRISPR-Cas9 can potentially be introduced into *S. hygroscopicus subsp. ascomyceticus, S. rapamycinicus,* and *A. mediterranei* to target the biosynthetic gene clusters for FK520, rapamycin and rifamycin B, respectively. CRISPR-Cas9 can introduce a DSB within the chosen region as dictated by a unique spacer sequence. The presence of a homology-directed repair template that contains a chosen DNA mutation should then affect a point mutation within the genes of interest. If CRISPR-Cas9 is targeted to the biosynthetic gene clusters in any of the three strains, a repair template could be used to introduce a stop codon or make a frame-shift mutation via nucleotide insertion/deletion. Doing so would likely decrease the final yield of the polyketide of interest significantly, as compared to the wild-type polyketide production. *Hypothesis: CRISPR-Cas9 can be used in vivo to introduce point mutations in the FK520, rapamycin and rifamycin B biosynethetic gene clusters. Targeting Cas9 to the polyketide biosynthetic gene cluster and introducing a stop codon or making a frame-shift mutation will cause significant reduction in the yield of the polyketide of interest.*

3.3.1 CRISPR-Cas9 induced glutamine auxotroph strain of *A. mediterranei*

To test the hypothesis that CRISPR-Cas9 can introduce site-specific DSBs within the genomes of *Actinomyces*, the essential glutamine synthetase gene, *glnA*, will be targeted in *A. mediterranei* using a 20 base pair (bp) spacer sequence[60]. The spacer sequence is designed to be homologous to a 20bp sequence found within the first 100bp of *glnA*, and followed by an NGG PAM sequence. The 20bp spacer sequence was assembled into the plasmid pCRISPomyces-2 to form pCRISPglnA, which also contains a codon-optimized *S. pyogenes* Cas9 gene, a temperature-sensitive replicon PSG5, apramycin resistance cassette, and *E.coli*

oriT to enable facile assembly and conjugation (**Figure 3-4**). Introduction of pCRISPglnA into *A. mediterranei* should produce a DSB in the *glnA* and reduce the number of surviving bacteria following transformation. Initial experiments using the pCRISPomyces-2 plasmid plus the spacer sequence but lacking a repair template did not provide sufficient evidence for a reduction in transformation efficiency or overall growth.

The presence of DNA repair mechanisms in *A. mediterranei* will likely allow some

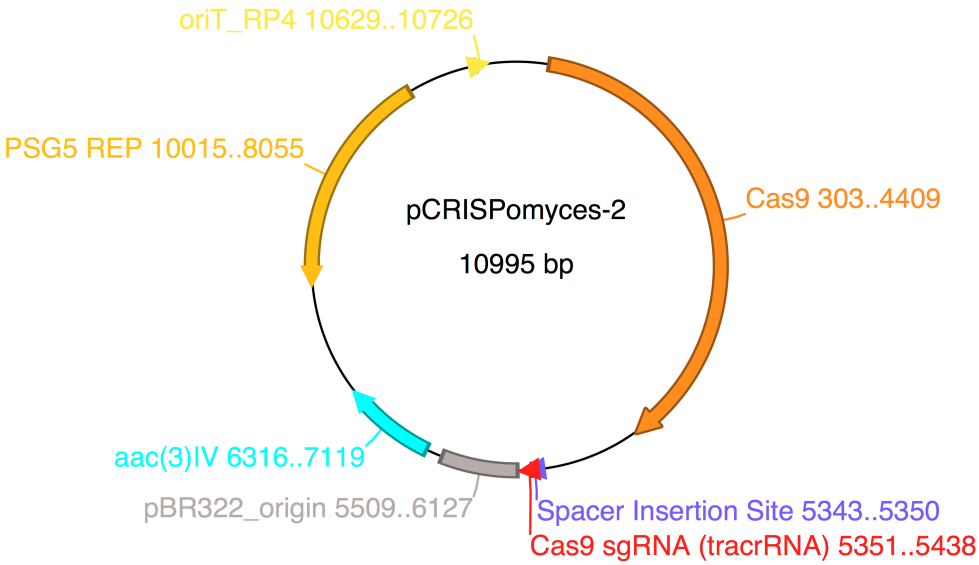


Figure 3-4: *Plasmid map of pCRISPomyces-2. The glnA spacer was inserted via golden-gate assembly at the "spacer insertion site" (in purple), to produce pCRISP-glnA.*

bacteria to survive the CRISPR-induced DSB in *glnA*. Therefore, a homology-directed repair template will be provided on the pCRISPglnA plasmid, which contains a TGA stop codon to insert into the upstream region of *glnA*, and will also mutate the PAM sequence to prevent repeated cleavage by Cas9. The stop codon in *glnA* will not only prevent transcription of a functional glutamine synthetase gene to create the glutamine auxotroph, but this will also demonstrate that a single base pair mutation (TAA->TGA) can be introduced using CRISPR-Cas9. As such, cells transformed with a plasmid containing both the CRISPR-Cas9 genes and

the glutamine synthetase repair template should only grow on media supplemented with glutamine.

3.4 Conclusions

The successful production and extraction of FK520 and rapamycin has been demonstrated from cultures of *S. hygroscopicus subsp. ascomyceticus* and *S. rapamycinicus*, respectively. Calibration curves with commercial standards of FK520, rapamycin and rifamycin B have been constructed via LC-ESI-MS (data not shown), and should allow for the reliable quantification of these compounds from fermentation extracts. Additional troubleshooting for the production and detection of rifamycin B from fermentation extracts of *A. mediterrani* needs to be performed.

A CRISPR-Cas9 shuttle plasmid for targeting *glnA* in *A. mediterranei* has been prepared for use. A mechanism for delivering the homology-directed repair template still needs to be realized, as cloning of the repair template into the CRISPR-Cas9 shuttle plasmid was unsuccessful. Altogether, this works lays the foundation upon which CRISPR-Cas9-mediated genome editing in polyketide-producing hosts can be built.

3.5 Materials and methods

Materials. The following microbial strains were used: *Streptomyces hygroscopicus subsp. ascomyceticus* (ATCC 55558), *Streptomyces rapamycinicus* (ATCC 29253), *Amycolatopsis mediterranei* (ATCC 21271), *Escherichia coli DH5α* (Lucigen), *Escherichia coli* BL21 (DE3) (Lucigen). Streptomyces were stored as spores in glycerol at -80°C and were prepared according to known methods[61]. The following plasmids were used: pCRISPomyces-2, purchased from Addgene[62]; Gene "glnA_repair" was synthesized by GenScript, USA.

Fermentation of Actinomyces. Overnight seed cultures were prepared by inoculating 5 mL of ISP medium I with 10 µL of spores from each glycerol stock. Cultures were heat shocked at 41 °C for 1 minute to germinate spores, and incubated in an orbital shaker at 30 °C, 250 rpm for ~15 hours. Entire 5 mL culture was added to 50 mL of fermentation media prepared in a 250 mL baffled flask for *S. hygroscopicus subsp. ascomycinicus* (1.5 g Tryptic

Soy Broth, 1 mL 0.5 g/mL Dextrose, pH 6), *S. rapamycinicus* (1.5 g Tryptic Soy Broth, 0.15 mL 0.5 g/mL Dextrose, pH 6), and *A. mediterranei* (1 g glycerol, 0.25 g yeast extract, 0.15 g beef extract, 0.25 g NZ amine A, 0.125 g peptone, 0.05 g malt extract, 2 mL 0.5 g/mL dextrose, pH 6). After 45 hours of growth in an orbital shaker at 28 °C at 250 rpm, cultures appeared fully opaque. Cultures fermented for a total of 69 hours before extraction of metabolites.

Extraction and mass spectroscopy analysis of FK520, rapamycin and rifamycin B. Fermentation cultures were centrifuged at 4000 rpm for 20 minutes to remove cell debris. Supernatant of *A. mediterranei* was acidified to pH 3.0 with HCl and extracted three times with equal volumes of ethyl acetate. Extracts were combined, washed with water, filtered through anhydrous $MgSO_4$ and evaporated to an oil. The supernatants of *S. hygroscopicus subsp. ascomycinicus* and *S. rapamycinicus* cultures were extracted 3 times with equal volumes of ethyl acetate. Combined extracts were washed once with water, filtered through anhydrous $MgSO_4$ and evaporated to a white powder. All extracts were stored in amber glass vials at -20 °C. Extracts were resuspended in 1 mL methanol and submitted for MS analysis.

Construction of plasmid pCRISPglnA. Two complementary oligonucleotides "pCRISP2-glnA-spacer-1" and "pCRISP-glnA-spacer-2" (synthesized by IDT) were annealed to create the following spacer sequence: 5'-gtccagttctgcgacctgcc-3'] with 4-nucleotide overlaps corresponding to the *Bbs*I restriction sites used for Golden Gate Assembly (GGA), performed via established methods[62]. Assembly was confirmed via sequencing using primer "pCRISP_GGA" from Genewiz, Inc. with the sequence 5'-CGTTCCTGCAATTCTTAG-3'. Multiple attempts to sub-clone the synthesized glutamine repair template, "glnA-rep", were made using a single *Xba*I restriction site.

REFERENCES

1. Wang, H., Sivonen, K. & Fewer, D. P. Genomic insights into the distribution, genetic diversity and evolution of polyketide synthases and nonribosomal peptide synthetases. *Curr. Opin. Genet. Dev.* **35,** 79–85 (2015).

2. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P T A peer-reviewed J. Formul. Manag.* **40,** 277–83 (2015).

3. Weisblum, B. Erythromycin resistance by ribosome modification. *Antimicrobial Agents and Chemotherapy* **39,** 577–585 (1995).

4. Seiple, I. B. *et al.* A platform for the discovery of new macrolide antibiotics. *Nature* **533,** 338–345 (2016).

5. Sharma, K. K. & Boddy, C. N. The thioesterase domain from the pimaricin and erythromycin biosynthetic pathways can catalyze hydrolysis of simple thioester substrates. *Bioorganic Med. Chem. Lett.* **17,** 3034–3037 (2007).

6. Stassi, D. L. *et al.* Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 7305–7309 (1998).

7. Ashley, G. W. *et al.* Preparation of Erythromycin Analogs Having Functional Groups at C-15. *J. Antibiot. (Tokyo).* **59,** 392–401 (2006).

8. Sundermann, U. *et al.* Enzyme-directed mutasynthesis: a combined experimental and theoretical approach to substrate recognition of a polyketide synthase. *ACS Chem. Biol.* **8,** 443–50 (2013).

9. Cortes, J., Haydock, S. F., Roberts, G. a, Bevitt, D. J. & Leadlay, P. F. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature* **348,** 176–178 (1990).

10. Weissman, K. J. & Leadlay, P. F. Combinatorial biosynthesis of reduced polyketides. *Nat. Rev. Microbiol.* **3,** 925–36 (2005).

11. Wong, F. T. & Khosla, C. Combinatorial biosynthesis of polyketides--a perspective.

Curr. Opin. Chem. Biol. **16,** 117–23 (2012).

12. Gatto, G. J., McLoughlin, S. M., Kelleher, N. L. & Walsh, C. T. Elucidating the substrate specificity and condensation domain activity of FkbP, the FK520 pipecolate-incorporating enzyme. *Biochemistry* **44,** 5993–6002 (2005).

13. Hsieh, Y. J. & Kolattukudy, P. E. Inhibition of erythromycin synthesis by disruption of malonyl-coenzyme A decarboxylase gene eryM in Saccharopolyspora erythraea. *J. Bacteriol.* **176,** 714–724 (1994).

14. Williams, D. L. *et al.* Contribution of rpoB mutations to development of rifamycin cross- resistance in Mycobacterium tuberculosis. *Antimicrob. Agents Chemother.* **42,** 1853–1857 (1998).

15. Vernon, A., Burman, W., Benator, D., Khan, A. & Bozeman, L. Acquired rifamycin monoresistance in patients with HIV-related tuberculosis treated with once-weekly rifapentine and isoniazid. *Lancet* **353,** 1843–1847 (1999).

16. Jenny-Avital, E. R. & Joseph, K. Rifamycin-resistant Mycobacterium tuberculosis in the highly active antiretroviral therapy era: a report of 3 relapses with acquired rifampin resistance following alternate-day rifabutin and boosted protease inhibitor therapy. *Clin. Infect. Dis.* **48,** 1471–1474 (2009).

17. Nigam, A. *et al.* Modification of Rifamycin Polyketide Backbone Leads to Improved Drug Activity Against Rifampicin-Resistant Mycobacterium tuberculosis*. *J. Biol. Chem.* (2014). doi:10.1074/jbc.M114.572636

18. Wilkinson, C. J., Frost, E. J., Staunton, J. & Leadlay, P. F. Chain initiation on the soraphen-producing modular polyketide synthase from Sorangium cellulosum. *Chem. Biol.* **8,** 1197–1208 (2001).

19. Long, P. F. *et al.* Engineering specificity of starter unit selection by the erythromycin-producing polyketide synthase. *Mol. Microbiol.* **43,** 1215–1225 (2002).

20. Rowe, C. J. *et al.* Engineering a polyketide with a longer chain by insertion of an extra

module into the erythromycin-producing polyketide synthase. *Chem. Biol.* **8,** 475–485 (2001).

21. Lau, J., Fu, H., Cane, D. E. & Khosla, C. Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units. *Biochemistry* **38,** 1643–1651 (1999).

22. Reeves, C. D. *et al.* Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* **40,** 15464–15470 (2001).

23. Reeves, C. D. *et al.* A new substrate specificity for acyl transferase domains of the ascomycin polyketide synthase in Streptomyces hygroscopicus. *J. Biol. Chem.* **277,** 9155–9 (2002).

24. Carney, J. R., Ashley, G. W., Arslanian, R. L. & Buchannan, G. O. Structure elucidation of new ascomycins produced by genetic engineering. *J. Antibiot. (Tokyo).* **58,** 715–721 (2005).

25. Del Vecchio, F. *et al.* Active-site residue, domain and module swaps in modular polyketide synthases. *J. Ind. Microbiol. Biotechnol.* **30,** 489–494 (2003).

26. Lowden, P. a S., Böhm, G. a, Metcalfe, S., Staunton, J. & Leadlay, P. F. New rapamycin derivatives by precursor-directed biosynthesis. *Chembiochem* **5,** 535–8 (2004).

27. Lechner, A. *et al.* Designed biosynthesis of 36-methyl-FK506 by polyketide precursor pathway engineering. *ACS Synth. Biol.* **2,** 379–83 (2013).

28. Ban, Y. H. *et al.* Mutational biosynthesis of a FK506 analogue containing a non-natural starter unit. *Mol. Biosyst.* **9,** 944–7 (2013).

29. Moss, S. J. *et al.* Novel FK506 and FK520 analogues via mutasynthesis: mutasynthon scope and product characteristics. *Medchemcomm* **4,** 324 (2013).

30. Nigam, A. *et al.* Modification of rifamycin polyketide backbone leads to improved

drug activity against rifampicin-resistant Mycobacterium tuberculosis. *J. Biol. Chem.* **289,** 21142–21152 (2014).

31. Koryakina, I. & Williams, G. J. Mutant malonyl-CoA synthetases with altered specificity for polyketide synthase extender unit generation. *ChemBioChem* **12,** 2289–2293 (2011).

32. Koryakina, I. *et al.* Poly specific trans -Acyltransferase machinery revealed via engineered acyl-coa synthetases. *ACS Chem. Biol.* **8,** 200–208 (2013).

33. Ray, L. & Moore, B. S. Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Nat. Prod. Rep.* **0,** 1–12 (2016).

34. Schmelz, S. & Naismith, J. H. Adenylate-forming enzymes. *Curr. Opin. Struct. Biol.* **19,** 666–671 (2009).

35. Go, M. K., Chow, J. Y., Cheung, V. W. N., Lim, Y. P. & Yew, W. S. Establishing a toolkit for precursor-directed polyketide biosynthesis: Exploring substrate promiscuities of acid-CoA ligases. *Biochemistry* **51,** 4568–4579 (2012).

36. Peter, D., Vögeli, B., Cortina, N. & Erb, T. A Chemo-Enzymatic Road Map to the Synthesis of CoA Esters. *Molecules* **21,** 517 (2016).

37. Erb, T. J., Brecht, V., Fuchs, G., Müller, M. & Alber, B. E. Carboxylation mechanism and stereochemistry of crotonyl-CoA carboxylase/reductase, a carboxylating enoyl-thioester reductase. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 8871–8876 (2009).

38. Kao, C. M., Luo, G., Katz, L., Cane, D. E. & Khosla, C. Manipulation of Macrolide Ring Size by Directed Mutagenesis of a Modular Polyketide Synthase. *J. Am. Chem. Soc.* **117,** 9105–9106 (1995).

39. Pieper, R., Luo, G., Cane, D. E. & Khosla, C. Remarkably broad substrate specificity of a modular polyketide synthase in a cell-free system. *J. Am. Chem. Soc.* **117,** 11373–11374 (1995).

40. Wiesmann, K. E. *et al.* Polyketide synthesis in vitro on a modular polyketide synthase.

Chem. Biol. **2,** 583–589 (1995).

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. Ritacco, F. V. *et al.* Production of Novel Rapamycin Analogs by Precursor Directed Biosynthesis. *Appl. Environ. Microbiol.* **71,** 1971–1976 (2005).

43. Jacobsen, J. R., Hutchinson, C. R., Cane, D. E. & Khosla, C. Precursor-directed biosynthesis of erythromycin analogs by an engineered polyketide synthase. *Science* **277,** 367–369 (1997).

44. Harvey, C. J. B., Puglisi, J. D., Pande, V. S., Cane, D. E. & Khosla, C. Precursor directed biosynthesis of an orthogonally functional erythromycin analogue: Selectivity in the ribosome macrolide binding pocket. *J. Am. Chem. Soc.* **134,** 12259–12265 (2012).

45. Hartung, I. V., Rude, M. a., Schnarr, N. a., Hunziker, D. & Khosla, C. Stereochemical assignment of intermediates in the rifamycin biosynthetic pathway by precursor-directed biosynthesis. *J. Am. Chem. Soc.* **127,** 11202–11203 (2005).

46. Klopries, S., Sundermann, U. & Schulz, F. Quantification of N-acetylcysteamine activated methylmalonate incorporation into polyketide biosynthesis. *Beilstein J. Org. Chem.* **9,** 664–674 (2013).

47. Arora, P., Vats, A., Saxena, P., Mohanty, D. & Gokhale, R. S. Promiscuous fatty acyl CoA ligases produce acyl-CoA and acyl-SNAC precursors for polyketide biosynthesis. *J. Am. Chem. Soc.* **127,** 9388–9389 (2005).

48. Hashimoto, Y. *et al.* Nitrile Pathway Involving Acyl-CoA Synthetase: OVERALL METABOLIC GENE ORGANIZATION AND PURIFICATION AND CHARACTERIZATION OF THE ENZYME. *J. Biol. Chem.* **280,** 8660–8667 (2005).

49. Pohl, N. L. *et al.* Remarkably Broad Substrate Tolerance of Malonyl-CoA Synthetase, an Enzyme Capable of Intracellular Synthesis of Polyketide Precursors. *J. Am. Chem.*

Soc. **123,** 5822–5823 (2001).

50. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9,** 40 (2008).

51. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5,** 725–738 (2010).

52. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **30,** 2785–2791 (2009).

53. Boghigian, B. a., Zhang, H. & Pfeifer, B. a. Multi-factorial engineering of heterologous polyketide production in Escherichia coli reveals complex pathway interactions. *Biotechnol. Bioeng.* **108,** 1360–1371 (2011).

54. Han, S. J., Park, S. W., Kim, B. W. & Sim, S. J. Selective production of epothilone B by heterologous expression of propionyl-CoA synthetase in Sorangium cellulosum. *J. Microbiol. Biotechnol.* **18,** 135–137 (2008).

55. Pfeifer, B. & Khosla, C. Biosynthesis of Polyketides in Heterologous Hosts. *Microbiol. Mol. Biol. Rev.* **65,** 106–118 (2001).

56. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* **11,** 181–190 (2010).

57. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. a. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507,** 62–7 (2014).

58. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **337,** 816–822 (2012).

59. Karvelis, T., Gasiunas, G. & Siksnys, V. Programmable DNA cleavage in vitro by Cas9. *Biochem. Soc. Trans.* **41,** 1401–6 (2013).

60. Peng, W. T. *et al.* Bacterial type I glutamine synthetase of the rifamycin SV producing

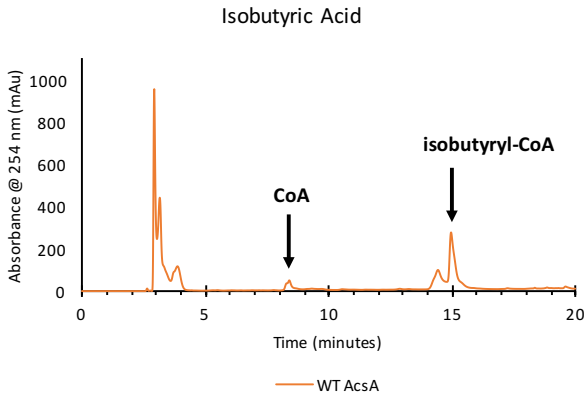
actinomycete, Amycolatopsis mediterranei U32, is the only enzyme responsible for glutamine synthesis under physiological conditions. *Acta Biochim. Biophys. Sin. (Shanghai)*. **38,** 821–830 (2006).

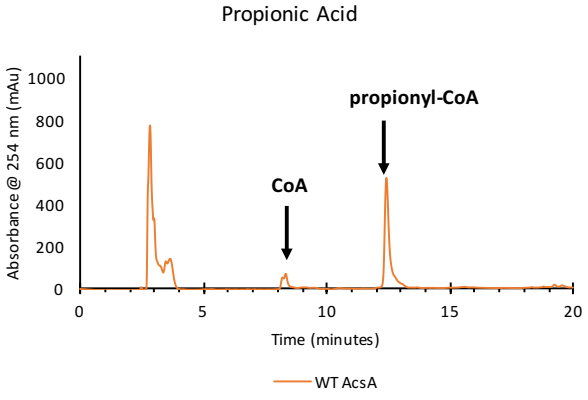
61. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. Practical Streptomyce Genetics. *John Innes Found. Norwich, UK* (2000).

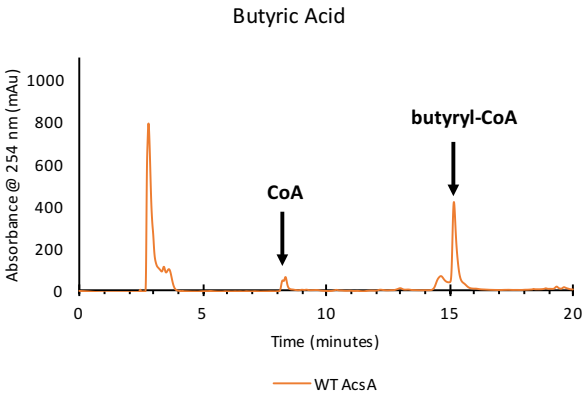
62. Cobb, R. E., Wang, Y. & Zhao, H. High-Efficiency Multiplex Genome Editing of Streptomyces Species Using an Engineered CRISPR/Cas System. *ACS Synth. Biol.* (2014).

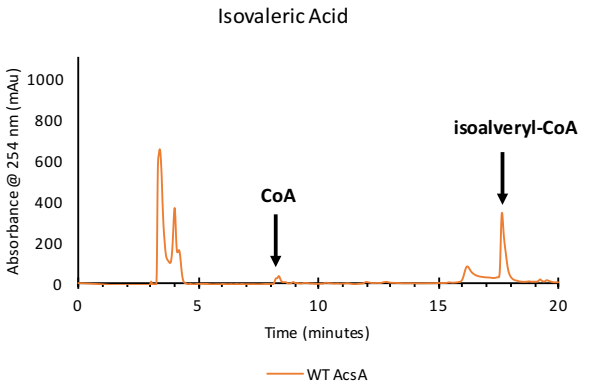
APPENDICES

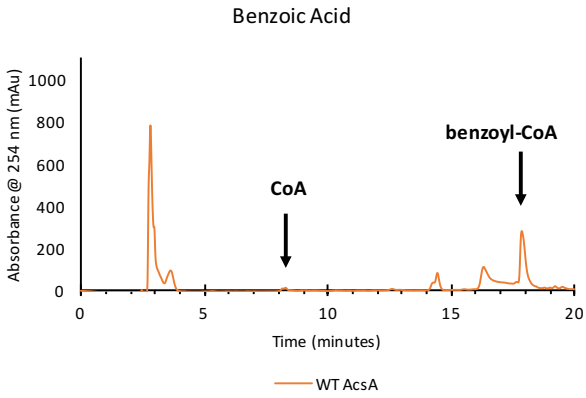
Appendix A


Isobutyric Acid
Absorbance @ 254 nm (mAu)
1000
800
600
400
200
0
0
5
10
15
20
Time (minutes)
CoA
isobutyryl-CoA
WT AcsA


Propionic Acid
Absorbance @ 254 nm (mAu)
1000
800
600
400
200
0
0
5
10
15
20
Time (minutes)
CoA
propionyl-CoA
WT AcsA


Butyric Acid
Absorbance @ 254 nm (mAu)
1000
800
600
400
200
0
0
5
10
15
20
Time (minutes)
CoA
butyryl-CoA
WT AcsA


Isovaleric Acid
Absorbance @ 254 nm (mAu)
1000
800
600
400
200
0
0
5
10
15
20
Time (minutes)
CoA
isoalveryl-CoA
WT AcsA


Benzoic Acid
Absorbance @ 254 nm (mAu)
1000
800
600
400
200
0
0
5
10
15
20
Time (minutes)
CoA
benzoyl-CoA
WT AcsA

Appendix B

Codon-optimized AcsA:

GAATTCATGCGGGACTATGAGCACGTGGTGGAGAGCTTCGACTATCTGCAGAGC
GCGACCCAGGACCTGCACGGCGAGCTGACCGCCCTGAACGCCTGCGTGGAGTGC
TGCGACCGTCACGCCCACGGCGAGGCCGTCGCGCTGTACTGCGAGGCCCAGGAC
GGCCACGCGGAGCGCTACCGCTTCCGCGACCTGCAGCGCCAGGCCGCGCGCTTC
GGCAACTTCCTGCGCGAGCAGGGCGTGAAGCCGGGTGACCGTGTGGCCGGCCTG
ATGCCGCGCACCGTGGAGCTGCTGATCGCCATCCTGGGTACCTGGCGTATCGGTG
CCGTCTACCAGCCGCTGTTCACCGCCTTCGGTCCGAAGGCCATCGAGCAGCGCCT
GAACTGCAGCAACGCCCGTTGGATCGTGACCGACCCGCACAACCGTCCGAAGCT
GGACGACGTCACCGACTGCCCGAGCATCGTGGTCACCGGCGGCGCCCCGCAGAA
CCCGGCCGACCACCACTTCTGGAGCGCCCTGAACCGCCAGGCCGACGACTGCGC
CCCGGTGCTGCTGGACGCCAGCGCCCCGTTCCTGCTGATGTGCACCAGCGGTACC
ACCGGTCCGGCCAAGCCGCTGGAAGTGCCGCTGAGCGCCATCCTGGCGTTCAAG
GGCTACATGCGTGACGCCATCGACCTGCGCGCCGACGACCGTTTCTGGAACCTG
GCGGACCCGGGTTGGGCGTACGGCCTGTACTACGCGGTGACCGGTCCGCTGGCC
TGCGGTTACGCCACCCTGTTCTACGACGGTCCGTTCACCGTGGAGAGCACCCGCC
ACATCATCGCCAAGTACGCGATCAACAACCTGGCGGGTAGCCCGACCGCCTACC
GCTTCCTGATCGCCGCGGGCGCGGAGTTCGCGGACGCCGTGCGTGGCCGCCTGC
GTGCCGTCAGCAGCGCGGGCGAGCCGCTGAACCCGCAGGTGGTCCGTTGGTTCG
CGGAGCAGCTGGGCGTGGTCATCCACGACCACTACGGCCAGACCGAGATCGGCA
TGGTGCTGTGCAACCACCACGGCCTGCGTCACCCGGTGCGTGAGGGTAGCGCCG
GCTACGCCGTCCCGGGTTACCGTATCGTGGTCCTGGACAAAGCCCACCGTGAGCT
GCCGGCCGGTCAGCCGGGCGTGCTGGCCGTCGACCGCGAGCGCAGCCCGCTGTG
CTGGTTCGACGGTTACCTGGGTATGCCGACCCAGGCCTTCGCCGGCCGCTACTAC
CTGAGCGGCGACATCGTGGAGCTGAACGACGACGGCAGCATCAGCTTCGTCGGC
CGCAACGACGACCTGATCACCACCAGCGGTTACCGCGTGGGTCCGTTCGACGTG
GAGAGCGCCCTGATCGAGCACCCGGCCGTGGTGGAGGCGGCCGTCATCGGCAAG
CCGGACCCGCAGCGCACCGAGCTGATCAAGGCCTTCGTGGTCCTGAACACCCCG
TACCTGCCGAGCCCGGAGCTGGCGGAGGAGCTGCGTCTGCACGTGCGTCAGCGT
CTGGCCGCGCACGCCTACCCGCGCGAGATGGAGTTCGTCGACCACCTGCCGAAG
ACCCCGAGCGGCAAACTGCAGCGGTTCATCCTGCGTAACCAGGAGATCGCCAAA
CAGCAGGCCCTGGGCTAAAAGCTT

Table A-1: List of primers

pCRISP2-glnA-spacer-1	ACGCGTCCAGTTCTGCGACCTGCC
pCRISP2-glnA-spacer-2	AAACGGCAGGTCGCAGAACTGGAC
pCRISP-GGA	CGTTCCTGCAATTCTTAG
W238NNK-RTH-F	/5PHOS/NNKGCGTACGGCCTGTACTACG
W238NNK-RTH-R	/5PHOS/ACCCGGGTCCGCCAG
R444NNK-RTH-F	/5PHOS/NNKGTGGGTCCGTTCGACG
R444NNK-RTH-R	/5PHOS/GTAACCGCTGGTGGTGATCAGGTC
T194S-RTH-F	/5PHOS/AGCAGCGGTACCACCGGT
T194S-RTH-R	/5PHOS/GCACATCAGCAGGAACGGG
AcsAMutSeq-1	AGCGCCCTGAACCGCCAG
AcsAMutSeq-2	TTCGCGGACGCCGTG
AcsAMutSeq-3	TTCGTCGGCCGCAACGAC