Title

Engineered production of fluorinated polyketides inspired by mechanistic studies of a polyketide synthase

Permalink

https://escholarship.org/uc/item/7fq3w7m3

Author

Ad, Omer

Publication Date

2017

Peer reviewed|Thesis/dissertation

Engineered production of fluorinated polyketides inspired by mechanistic studies of a polyketide synthase

by

Omer Ad

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michelle C. Y. Chang, Chair

Professor Christopher J. Chang

Professor Michiko E. Taga

Spring 2017

Engineered production of fluorinated polyketides inspired by mechanistic studies of a polyketide synthase

by Omer Ad

Abstract

Engineered production of fluorinated polyketides inspired by mechanistic studies of a polyketide synthase

by

Omer Ad

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michelle C. Y. Chang, Chair

Macrolide natural products have been a rich source of medicinally relevant molecules. Macrolides are assembled by large protein complexes known as polyketide synthases (PKSs) via polymerization reactions of simple acetate and propionate derived building blocks. While these enzymes are capable of generating molecules of impressive structural diversity, the introduction of nonnative and potentially bioactive functional groups into the backbone of these molecules remains challenging and is mostly limited to synthetic methods. The incorporation of fluorine atoms or fluorinated moieties into both synthetic and semi-synthetic medicinally relevant molecules has been shown to improve their bioactivity, bioavailability, and stability. However, the very properties that make fluorine a valuable component of pharmaceuticals also makes it challenging to incorporate via synthetic methods. Seeking to capitalize on the modular nature of PKSs, previous work in our lab has shown that fluorinated polyketides can be synthesized by using fluorinated analogs of PKS building blocks. The work described in this dissertation follows a mechanistic approach to fluorinated extender unit incorporation into polyketides. By examining the nature of fluorinated extender unit incorporation into the 6-deoxyerythronolide B synthase modules, we were able to build a platform showing suitable promise for the site-specific incorporation of fluorine into the backbone of medicinally relevant natural products.

1

Table of Contents

Chapter 1:   Introduction

Chapter 2:   Elucidating the mechanism of fluorinated extender unit loading for improved production of fluorine containing polyketides

Chapter 3:   Towards *in vivo* production of site-specifically fluorinated polyketides using a fluoromalonyl-CoA selective *trans*-AT

Chapter 4: Exploring extender unit promiscuity in polyketide synthases

Appendices

List of Figures, Schemes, and Tables

Chapter 2

Chapter 3

Chapter 4

List of Abbreviations

6-dEB	6-deoxyerythronolide B
ACCase	acetyl-CoA carboxylase
AckA	acetate kinase
ACP	acyl carrier protein
AT	acyltransferase
bp	base pair
Br-TKL	2-bromo-3-keto-4-methyl-5-hydroxy-*n*-heptanoic acid δ-lactone
CoA	coenzyme A
CRISPR	Clustered regularly interspaced short palindromic repeats
Cl-TKL	2-chloro-3-keto-4-methyl-5-hydroxy-*n*-heptanoic acid δ-lactone
DEBS	6-deoxyerythronolide synthase
DH	dehydratase
DszAT	disorazole synthase acyltransferase
ER	enoylreductase
ESI-MS	electrospray ionization mass spectrometry
Et-TKL	2-ethyl-3-keto-4-methyl-5-hydroxy-*n*-heptanoic acid δ-lactone
FCl-TKL	2-fluoro-3-keto-4-chloro-5-hydroxy-*n*-heptanoic acid δ-lactone
F-TKL	2-fluoro-3-keto-4-methyl-5-hydroxy-*n*-heptanoic acid δ-lactone
H-TKL	4-dimethyl-3-keto-5-hydroxy-*n*-heptanoic acid δ-lactone
H-TTKL	4-hydroxy-6-(3-hydroxypentan-2-yl)-2*H*-pyran-2-one
kDa	kilodaltons
KR	ketoreductase
KS	ketosynthase
LC-MS	liquid chromatography-mass spectrometry
LC-UV	liquid chromatography-ultraviolet visible spectroscopy
MatB/rpMatB	Rhodopseudomonas palustris MatB
Mod	module
Mod_{TE}	module with thioesterase
MS	mass spectrometry
NMR	nuclear magnetic resonance
PKS	polyketide synthase
Pta	phosphotransacetylase

SNAC   N-acetylcysteamine
TE   Thioesterase
TKL   2,4-dimethyl-3-keto-5-hydroxy-*n*-heptanoic acid δ-lactone
WT   wild-type

Acknowledgments

The M. Chang lab has been my second home for the last five years. I've been grateful for the way I was welcomed into in the lab the moment I arrived at Berkeley. I've had the distinct pleasure of working with fantastic people directly and indirectly every day and I hope they all learned as much from me as half of what I learned from them. Joel Sax is an incredibly bright undergraduate student and I'm so proud of what we were able to accomplish together. Thank you for going on this journey with me and being the "guinea-pig" as I developed a mentoring style. I was also very fortunate to be mentored by three terrific senior students; Amy, Mark, and Ben – I will be forever grateful for the diligence and patience and was greatly impacted by each of you. Kersh Theva arrived late during my time in the lab, but was a source of endless energy and engaging discussion. Kersh, thank you for teaching "an old dog new tricks." Jorge Marchand has been one of the most fun benchmates one could ask for. Jorge, your lack of hearing will always be hilarious to me. Jon McMurry was not directly acknowledged for any particular contribution. However, the work in this dissertation has been dramatically impacted by his presence during my time in the lab. Jon is the most supportive, generous, and knowledgeable scientist I have had the pleasure to work with and I am so thankful to have had the opportunity to go through these last five years with him. I was also fortunate enough to mentor some incredible rotation students. Jenna, Olga, Tammy, and Sasilada-it was great to get to know all of you and be able to take partial credit for your excellent careers. I would also like to thank Joyce Liu, Jeff Pelton, and Allegra Aron for being great collaborators and enriching my scientific repertoire with their knowledge.

None of this work would be possible without the dedication Michelle has for her students. In many ways, the work presented here is a departure from previous work in the lab. Michelle was always very supportive, and more importantly, trusting in my abilities to venture into this challenging research question. The greatest takeaway from my time working with Michelle is her ability to communicate science. I will carry the presentation and writing skills I imprinted from her for the rest of my career. Finally, I want to thank Michelle for always being able to introduce perspective into the lab. Michelle, you have been incredibly supportive whenever personal life intersected with graduate school and I will always appreciate your advice and understanding during that time.

My family and friends have sacrificed a lot as I went on this journey. My parents have been a source of support and constructive criticism throughout my time and it has been an honor to follow in their footsteps and take an academic path. Finally, I would like to thank my other half and soon-to-be-wife Christine. Christine has always been there for me - she supplied me with confidence, reassurance, and most-importantly, doses of reality. Thank you for your unwavering support. I am so lucky to have you in my life.

Chapter 1: *Introduction*

1.1 Secondary metabolites in modern medicine

Microbes have evolved to produce secondary metabolites as a defense mechanism against other pathogenic microbes, plants, and animals. Indeed, almost half of freshly isolated actinomycetes and filamentous fungi have been found to produce antibiotics [1]. Moreover, secondary metabolites produced by organisms of all taxa have been shown to also possess anticancer, antiviral, and antiparasitic properties [1,2]. These valuable chemical entities have been used in human healthcare as early as 2600 B.C, where clay tablets in Mesopotamia depicted the use of oils from myrrh, licorice, cypress trees to treat inflammation [2,3]. It was millennia later that the active ingredient in *Cupressus semperivens* (cypress) oils was found to be α-pinene [4,5], which has been shown to reduce inflammation in both human chondrocytes and mice [6,7]. The role of secondary metabolites in modern medicine expanded dramatically after the discovery of penicillin in 1929 when Alexander Fleming observed that *Penicillium notatum* was capable of killing *Staphylococcus aureus* cultures [8]. Strain optimization of the penicillin-producing *Penicillium chrysogenum* resulted in robust production of the first human therapeutic isolated from a microbe [1] and was responsible for launching a new era of medicine. Over the last few decades antibiotic resistant pathogens have become an increasing source of concern for human health, defining the discovery of novel antibiotics as one of the major health initiatives of the next century [9-11].

The high success rate of antibiotic discovery in the 20th century prompted a major shift in healthcare leading to the heavy reliance and largescale production of various secondary metabolites like erythromycin, cephalosporin C, and penicillins [9]. Specifically, between 1981 and 2005, 73% of newly discovered antibiotics consisted of penicillins, cephalosporins, macrolides, and quinolones [12,13]. However, the discovery of novel secondary metabolite drugs has dramatically stalled in the last twenty years [9-11] and has resulted in a healthcare crisis ushered by the horizontal gene transfer of antibiotic resistance genes [14]. As an alternative to secondary metabolite drugs, many research and industrial efforts began focusing on the production of synthetic bioactive molecules. The development of synthetic antibiotics is attractive given the uncommon occurrence of these molecules in nature, making antibiotic resistance to such molecules a low probability [12]. To date, synthetic antibiotics are largely outnumbered by their biosynthetic counterparts [12,13]. However, semi-synthetic derivatives have become a viable option for the production of new antibiotics and have supplied the major source of novel antibiotics over the last thirty years [3,13,15]. Specifically, the improved biochemical and biophysical understanding of antibiotic mechanisms of action make the design of synthetic and semi-synthetic antibiotics increasingly more likely to produce new broad and narrow spectrum antibiotics [16]. The structural complexity of secondary metabolites is extremely challenging to reproduce via synthetic methods [15] and as such, synthetic antibiotics are more simplistic in their structure and contain various functional groups that contribute to enhanced bioactivity (*Figure 1.1*) [17,18]. The structural divergence between secondary metabolite and synthetic drugs make the introduction of various functional groups into secondary metabolite scaffolds challenging, limiting the scope of diversity of semi-synthetic molecules produced [12].

As mentioned above, macrolide antibiotics are heavily used in the antibiotic arena. Specifically, macrolides comprise around twenty percent of currently prescribed antibiotics [19]. Macrolides are biosynthesized using large multi-domain complexes known as polyketide synthases (PKSs) [20-22]. Modular PKSs are responsible for the production of many bioactive macrolides and are primarily produced by bacteria [23]. These sizeable enzyme assembly lines

produce very large and complex structures through sequential C-C bond formations, reductive chemistry, and post-tailoring oxidation and glycosylation reactions (*Scheme 1.1*) [20-22,24]. For example, the antibiotic erythromycin is assembled by the formation of the polyketide 6-deoxyerythrolonide B (6-dEB). 6-dEB is an incredibly complex molecule boasting ten stereogenic centers [25] and becomes even more complex after tailoring reactions form the antibiotic erythromycin. While synthetic and semisynthetic efforts have resulted in the formation of erythromycin derivatives [12,15], the structural complexity of this molecule remains challenging to alter in a high throughput fashion [26].

As shown in *Scheme 1.1*, macrolides like 6-dEB are assembled using sequential polymerizations of an activated propionyl-CoA via a decarboxylative Claisen condensation of methylmalonyl-CoA extender units [20-22]. Macrolides and other PKS products are assembled by these condensation reactions using a variety of malonyl-CoA extender units [27,28]. The modular nature of these reactions initially inspired the strategy of module and domain swapping among various clusters in an effort to produce novel antibiotics [29,30]. However, the structural complexity of domain linkers, substrate specificity, and vectorial and turnstile nature of PKSs make domain swapping difficult and nonsystematic [30,31]. As a result of these challenges, recent
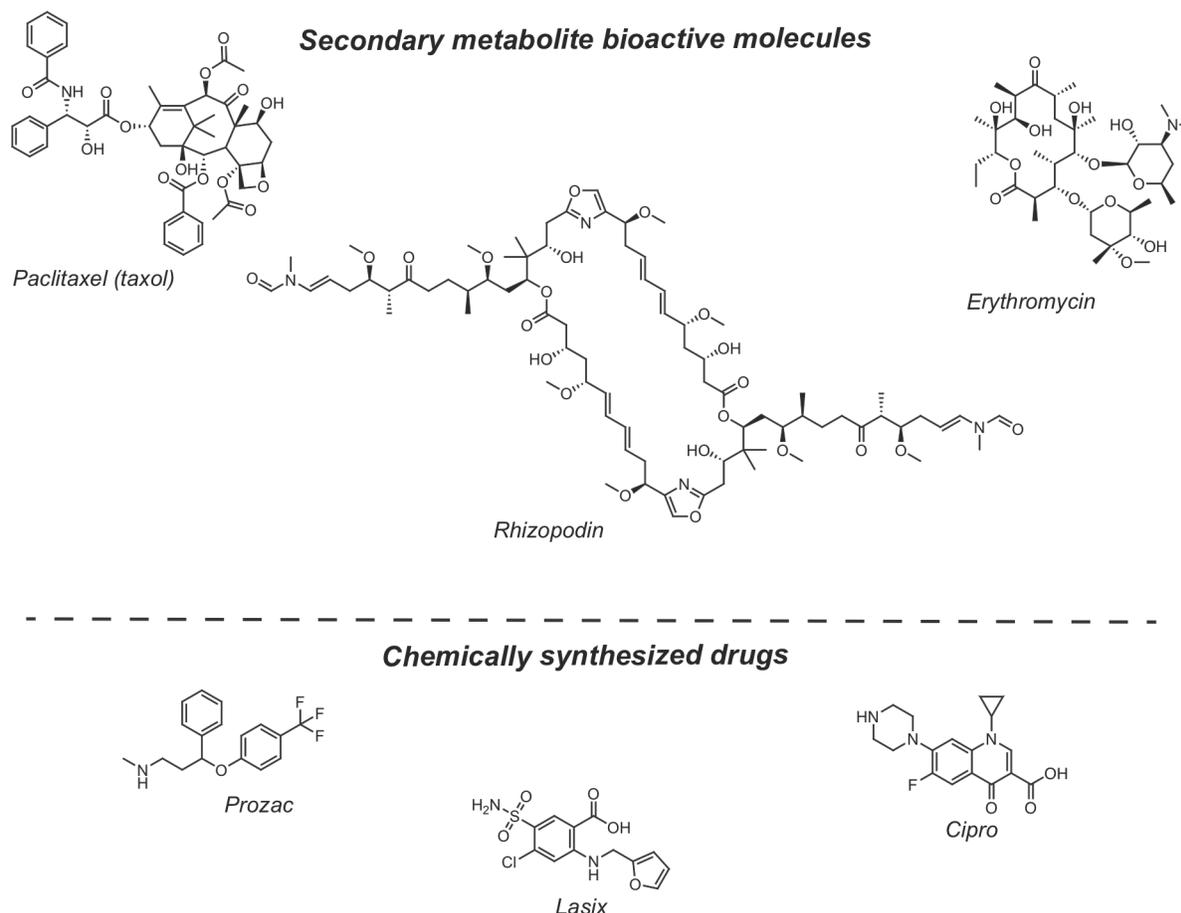
Figure 1.1: *Structures of secondary metabolites and chemically synthesized drugs. As observed, the secondary metabolite drugs are more complex in size and stereochemistry. Paclitaxel is a proposed anticancer agent, erythromycin is an antibiotic, and rhizopodin is a proposed antibiotic. Prozac is an antidepressant, Lasix is an edema medication, and Cipro is an antibiotic.*

efforts have focused on engineering specific PKS domains in order alter their substrate preference and produce novel antibiotics and other medicinally relevant molecules [*32-35*].

1.2 Engineering acyltransferases for the production of novel polyketides

Engineering *cis*-acyltransferases. C-C bond formation reactions in PKSs occur through decarboxylative Claisen condensations that are catalyzed by three domains in each PKS module. Initially, the polyketide growing chain is transferred to the active site cysteine of the ketosynthase (KS) via the upstream module's acyl-carrier protein (ACP) (*Scheme 1.2A*) [*20-22*]. The extender unit used for C-C bond formation is selected by the acyltransferase (AT) and transacylated onto the ACP. The AT has been shown to be the gatekeeper of the module, and ensures the acylation of a proper extender unit by selecting against unnatural extender units [*34,36*] or via hydrolysis of misacylated extender units [*36,37*]. The extender unit bound to the active-site serine of the AT is then transacylated onto the ACP and the ACP shuttles this phosphopantetheine-bound extender unit into the KS active-site, where C-C bond formation occurs between the extender unit and growing chain (*Scheme 1.2A*). This process is then repeated in the following modules, until the final product is synthesized. Type I modular PKSs are separated into two classes based on their domain architecture. Specifically, certain clusters contain *cis*-ATs that are covalently linked to the rest of the module, while the other class of enzymes contain a *trans*-AT that is not covalently attached to the other module domains and can act in *trans* on either a single ACP or ACPs from multiple modules (*Scheme 1.2B*) [*20,38*]. The molecular and biochemical understanding of *cis*-AT PKSs has initially been more extensive than that of *trans*-ATs, primarily due to the exhaustive
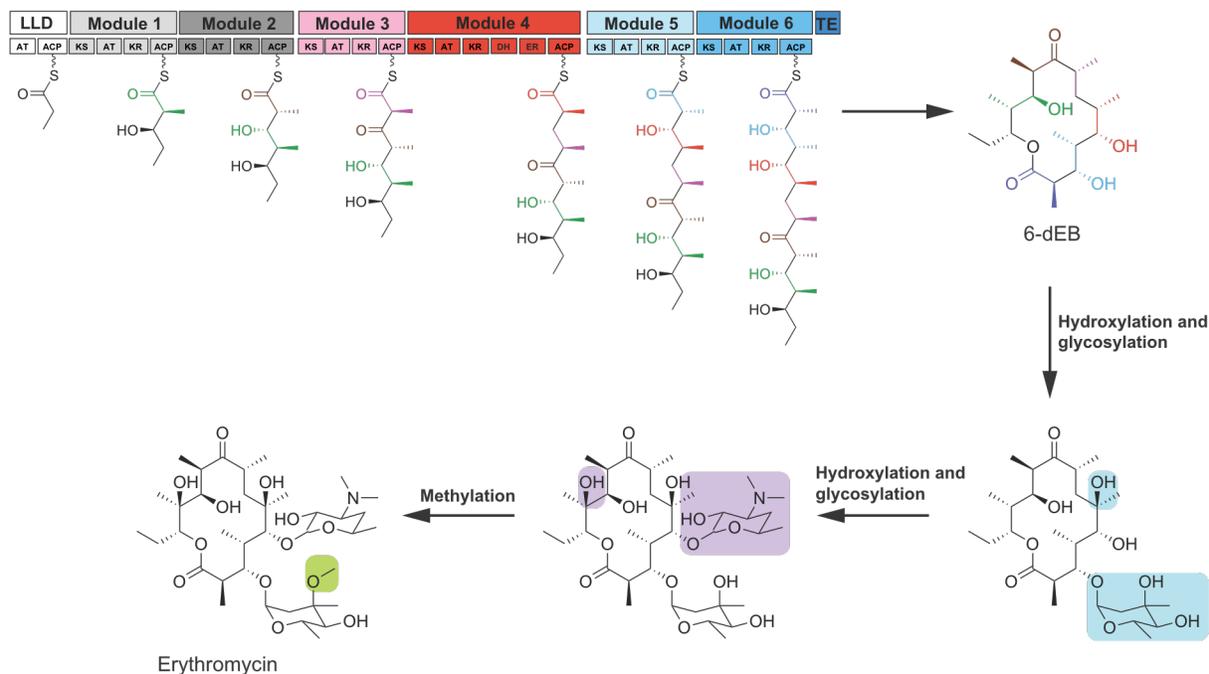
Scheme 1.1: *Biosynthesis of erythromycin. The polyketide backbone (6-dEB) is assembled using the DEBS PKS modules and is color-coded to represent monomer incorporation. Subsequent hydroxylation, glycosylation, and methylation results in the formation of the bioactive erythromycin.*

4

studies performed on the 6-dEB biosynthetic cluster (DEBS). As such, initial PKS engineering efforts centered around engineering the methylmalonyl-CoA specific cis-ATs of DEBS [20].

One of the first examples of AT engineering came in the form of mutagenesis to what were referred to as "hypervariable regions" of the DEBS ATs. Specifically, large regions of the AT sequence were changed in DEBS AT6 to those of the malonyl-CoA selective rapamycin biosynthesis cluster AT2 [39]. These mutations resulted in the incorporation of malonyl-CoA at the terminal position of DEBS and led to the production 2-desmethyl-6-dEB in *Streptomyces coelecolor* [39]. Another more conservative mutagenesis effort with module 4 AT of DEBS successfully introduced malonyl-CoA into full length 6-dEB. In this study by Kosan Biosciences, it was found that by changing the conserved YASH motif of the methylmalonyl-CoA selective DEBS ATs to the malonyl-CoA selective AT conserved motif HAFH, results in equivalent formation of 2-desmethyl-6-dEB and 6-dEB with only a 5-fold drop of product yield in *Streptomyces lividans* [40]. The ability to introduce malonyl-CoA into 6-dEB sparked interest in engineering *cis*-ATs to incorporate unnatural extender units, especially those with pharmaceutically relevant properties and synthetic chemical handles.

The advent of bioorthognal chemistry methods [41,42], specifically copper-click chemistry, inspired the desire to introduce propargylmalonyl-CoA into the backbone of polyketide scaffolds [35,43,44]. Propargylmalonyl-CoA was indeed incorporated into polyketide scaffolds both *in vitro* [35] and *in vivo [43,44]* via active-site residue changes that presumably expanded the active-site pocket and accommodated larger extender units [35]. In general, the incorporation of larger extender units into the backbone of polyketides with methylmalonyl-CoA specific ATs appears to be a promising approach for engineering structural diversity. Even extender units as large a 4-phenylethylmalonyl-CoA have been incorporated into DEBS polyketide products [45] and engineering efforts have also resulted in inversion of substrate selectivity in DEBS Module 6, resulting in a propargylmalonyl-CoA selective AT [35]. Although these engineered ATs provide an exciting opportunity for the production of unnatural polyketide scaffolds, the universality of these AT mutations remains to be seen. Additionally, the production of mutant libraries using this method is challenging from a genetic manipulation and engineering perspective, since each AT in
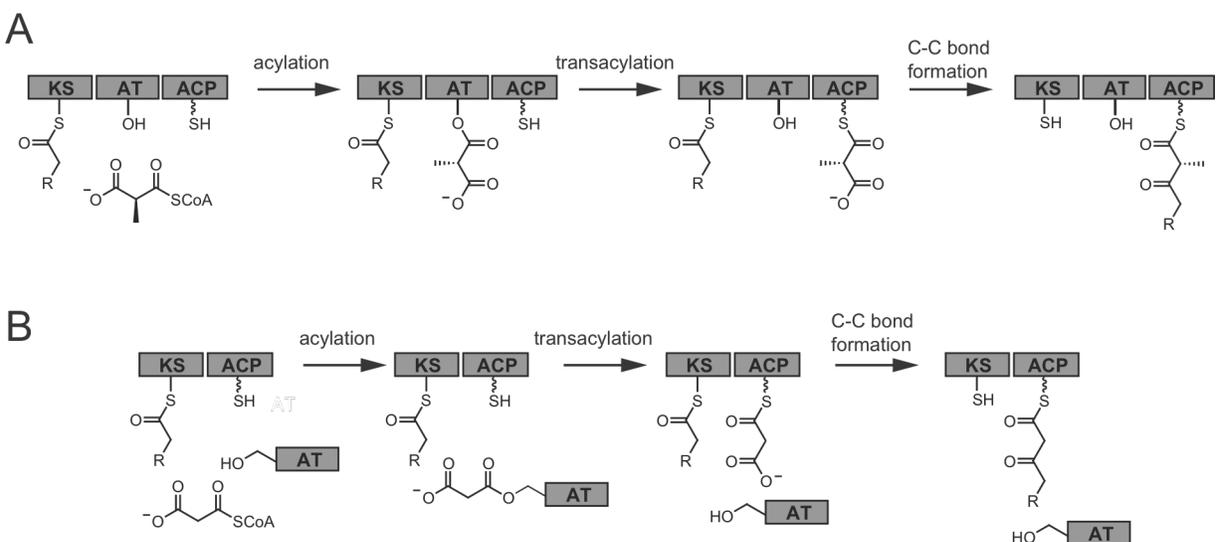
Scheme 1.2: *Cis- and trans-AT paradigms. (A) Scheme representing cis-AT type I modular PKS C-C bond formation paradigm. (B) Scheme representing trans-AT type I modular PKS C-C bond formation paradigm.*

5

the cluster needs to be engineered for every variant. This is particularly challenging given the high-GC content and large size of these genes, along with the fact that most of these secondary metabolites can only be produced at a large scale in actinomycete hosts [46,47]. The ability to engineer a single AT and complement it in *trans* with inactivated ATs has become a more attractive strategy for engineering novel polyketides given the low background of the natively produced product, as well as the universal success a single *trans*-AT affords in terms of substrate preference [27,32-34,48].

Engineering and complementation using *trans*-acyltransferases. Although only recently identified as a class of PKSs, type I modular *trans*-AT PKSs represent just shy of 40% of all modular PKS clusters [23,38]. *Trans*-AT PKSs are responsible for the biosynthesis of multiple antibiotics like mupiricin and the backbone of streptogramins [38]. While similar in architecture to the canonical type I PKSs such a DEBS, *trans*-AT PKSs have evolved independently from their *cis*-AT counterparts and are found mostly in proteobacteria, myxobacteria, and bacilli [27,49]. The most important difference of course arises in the domain architecture, where the modules do not contain an AT, but rather use an AT acting in *trans* to load ACPs with the proper extender unit (*Scheme 1.2B*). These *trans*-ATs can acylate multiple modules or a specific ACP [48,50,51]. Most *trans*-ATs are malonyl-CoA selective, however there are various ATs that can load other extender units such as ethyl- and methoxymalonyl-CoA [48,51,52]. Additionally, *trans*-AT PKSs feature various enzymes that catalyze more exotic chemical transformations then those normally found in type I modular PKSs like DEBS, but such enzymes are beyond the scope of this discussion [38,53].

The ability of *trans*-ATs to acylate multiple ACPs has inspired multiple engineering strategies seeking to exploit this promiscuity in order to load extender units onto ACPs from a nonative PKS pathway. Work with the *trans*-AT from the disorazole biosynthetic cluster (DszAT) has not only shown that can this AT acylate DEBS modules with malonyl-CoA, but it is able to do so at a rate faster than the native DEBS *cis*-AT is able to load those same modules with methylmalonyl-CoA [34,54]. The ability to use *trans*-ATs to acylate ACPs with nonative extender units becomes even more powerful once the native *cis*-AT is inactivated with a simple catalytic serine to alanine mutation (AT^0). The inability of the module to form methylmalonyl-ACPs results in malonyl-CoA incorporation and inversion of substrate selectivity in the desired module. More so, choosing the location of the AT^0 mutation enables the site-specific incorporation of malonyl-CoA into the polyketide backbone. This strategy was recently carried out with a reaction containing an *in vitro* reconstituted DEBS PKS with an AT^0 module 1. This system was complemented with DszAT and resulted in the formation of desmethyl-6-dEB [34]. The molecular mechanism of *trans*-AT acylation of either its native ACPs or those from other PKSs has yet to be determined, but will likely be investigated in order to optimize the complementation between the *cis*-AT^0 modules and the *trans*-AT of interest. Recent work by the Khosla lab has shown that the ethylmalonyl-CoA selective *trans*-AT KirCII is highly selective towards it's cognate ACP, but is more promiscuous when it comes to substrate preference, exhibiting acylation activity with methyl-, allyl-, and propargylmalonyl-CoA [34,48,50]. Conversely, the malonyl-CoA selective *trans*-AT DszAT exhibits promiscuity towards multiple ACPs, but demonstrates stricter extender unit selectivity than KirCII [32-34,50]. Findings such as these are of paramount importance for developing universal methods of *trans*-AT complementation for the production of novel molecules. Specifically, understanding ACP and substrate promiscuity in these ATs, as well as their gate-keeping mechanisms, can inform strategies for the design of chimeric *cis*- and *trans*-AT PKS clusters.

While it has been shown that *trans*-ATs can produce unnatural natural products *in vivo* [*48*], these examples are limited to production of novel molecules by inserting an unnatural substrate into the *trans*-AT's native PKS product. To date, the complementation of *trans*-ATs with nonnative PKS modules has only been explored *in vitro* [*32-34,48,51,54*]. However, the successful implementation of this strategy *in vitro* shows promise for this system as a tool for producing unnatural polyketides *in vivo*. This strategy is particularly exciting given its potential to incorporate the medicinally relevant fluorine atom [*55*] into the backbone of polyketides [*32,33*]. The similar size between fluorine and hydrogen makes fluoromalonyl-CoA a good substrate analog for malonyl-CoA [*56*]. It has been shown that while fluoromalonyl-CoA can be incorporated in model polyketide products using the methylmalonyl-CoA selective modules, C-C bond formation occurs via an ACP-independent mechanism, inhibiting the ability to produce full-length fluorinated polyketides [*32,33*]. However, the complementation of DszAT with AT^0 modules of DEBS results in restoration of the canonical mode of C-C bond formation and regiospecific incorporation into the backbone of polyketide products [*33*]. The ability to complement a malonyl-CoA selective *trans*-AT with a methylmalonyl-CoA module containing an AT^0 mutation has the potential to produce full-length fluorinated polyketides *in vivo*.

1.3   Fluorine in medicine

Properties of fluorine atoms and their impact on bioactivity. Only a handful of naturally occurring organofluorines have been characterized to date. Yet, as of 2008, 20% of pharmaceuticals on the market contain fluorine. [*57*]. As mentioned above, the structure of chemically synthesized pharmaceuticals varies substantially from secondary metabolite pharmaceuticals, and fluorine atoms play a central role in improving the bioactivity of these synthetic drugs [*17,18*]. This high abundance of fluorine atoms found in synthetic drugs suggest that the efficacy of fluorine in medicine is related to its physical properties [*57*] and is not simply related to bioisoterism [*17*]. A fluorine-to-hydrogen substitution is often referred to as a conservative substitution based on steric consideration. While this is true in many cases, the electronegative nature of fluorine actually results in the C-F bond bearing more similarity in size and dipole to C=O bonds than C-H bonds. [*17,58,59*]. The highly polarized nature of the C-F bond also increases the bond strength and results in improved interactions with polar functional groups, as well as strong interactions with hydrogen bond donors [*60*]. This electronegativity also has effects on the acidity of protons neighboring a C-F bond, ultimately leading to improved solubility, activity and potency, and pharmacokinetic properties [*58,61*]. Additionally, metabolically labile sites can be eliminated in molecules of interest by the addition of a fluorine moiety to that site [*57*]. This phenomenon is directly related to the fact that metabolic degradation is aided by cytochrome P_{450} catalyzed oxidation of drugs. When a fluorine atom is present at one of these metabolically labile sites, the reactivity towards cytochrome P_{450} oxidation is greatly reduced, avoiding changes to the lipophilicity of the molecule and ensuring improved drug longevity [*62,63*].

Fluorine incorporation in chemical synthesis. Fluorine chemistry has been extensively studied for over 100 years due to its relevance in the pharmaceutical arena [*64*]. For example, the semi-synthetic antibiotic solithromycin is a fluorinated analog of erythromycin currently under clinical trials [*65*]. Solithromycin has shown promise as a versatile intravenous and oral drug to fight against bacterial pneumonia and other bacterial infections and shows both enhanced potency and stability, as well as improved activity against antibiotic resistant strains [*65,66*]. Although
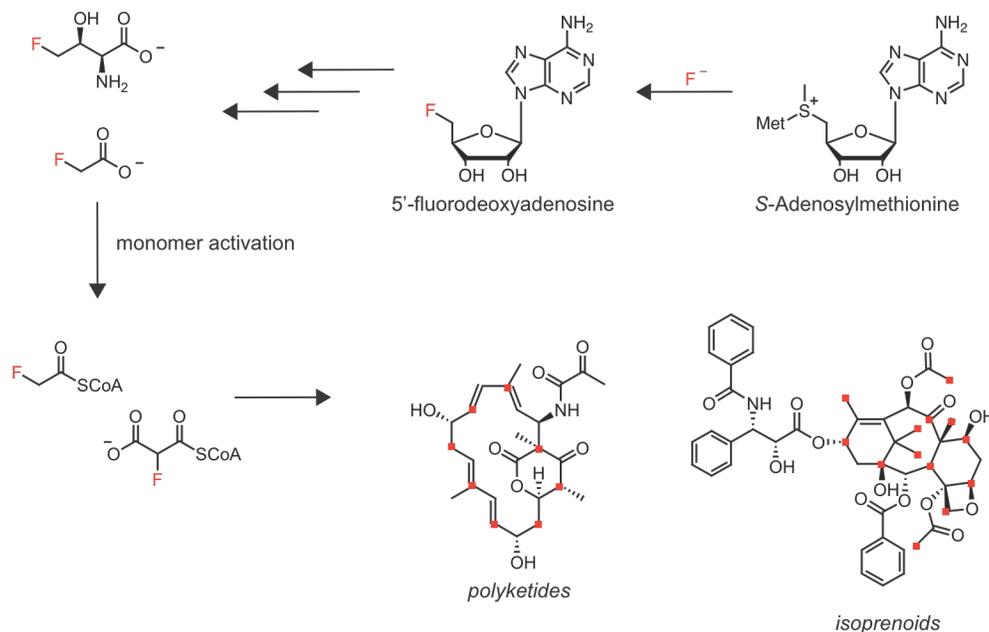
Scheme 1.3: *Proposed engineered biosynthetic scheme from fluoride to fluorinated secondary metabolites. Fluoroacetate and fluorothreonine biosynthesis starts with formation of 5'-fluorodeoxyadenosine from inorganic fluoride and S-adenosylmethionine. Fluoroacetate can then be activated to form fluoroacetyl-CoA and fluoromalonyl-CoA, which can both serve as building blocks for engineered fluorinated secondary metabolite biosynthesis. Potential fluorine incorporation sites are represented by red dots on the structures of the polyketide and isoprenoid molecules.*

fluorine chemistry has greatly improved over the past few decades, it is still limited in scope, making modification to complex molecules, such as secondary metabolites difficult to attain [*64,67*]. This challenge of late-stage fluorination has introduced an alternative synthetic method, where fluoro-alkyl building blocks are used as a tool to introduce fluorine into complex molecules, avoiding the fluorination of complex molecular scaffolds [*68,69*]. However, the use of fluorinated building blocks is not without its challenges. C-C bond formation and C-H functionalization reactions often occur at harsh conditions, suggesting that the introduction of fluorine in early-stage synthetic intermediates may negatively affect the synthetic strategy and diminish already poor reaction yields. The introduction of fluoroalkyl building blocks is very much a bioinspired strategy and is reminiscent of assembly line biosynthesis [*70*]. In fact, this approach shines a light on the fact that the introduction of fluoroalkyl groups into biosynthetic molecules via unnatural monomers ultimately has the potential to produce complex fluorinated molecules of interest, under mild reaction conditions [*20*].

Enzymatic incorporation of fluorine in bioactive molecules. The formation of a C-F bond is incredibly rare in natural systems. To date, only one C-F bond forming reaction has been identified and characterized [*71*]. The fluorinase is an enzyme first discovered in *Streptomyces cattleya* and forms a C-F bond via attack of an inorganic fluoride on *S*-adenosylmethionine (*Scheme 1.3*) [*72,73*]. After fluorination, 5'-fluorodeoxyadenosine undergoes multiple transformations to form both fluoroacetate and fluorothreonine [*71-73*]. Fluoroacetate can be converted into fluoroacetyl-CoA, which is the molecule ultimately responsible for fluoroacetate's antibiotic properties as a one-to-one inhibitor of the citric acid cycle [*74,75*]. The ability to form fluoroacetyl-CoA from fluoroacetate opens the door to fluorine incorporation into the backbone of
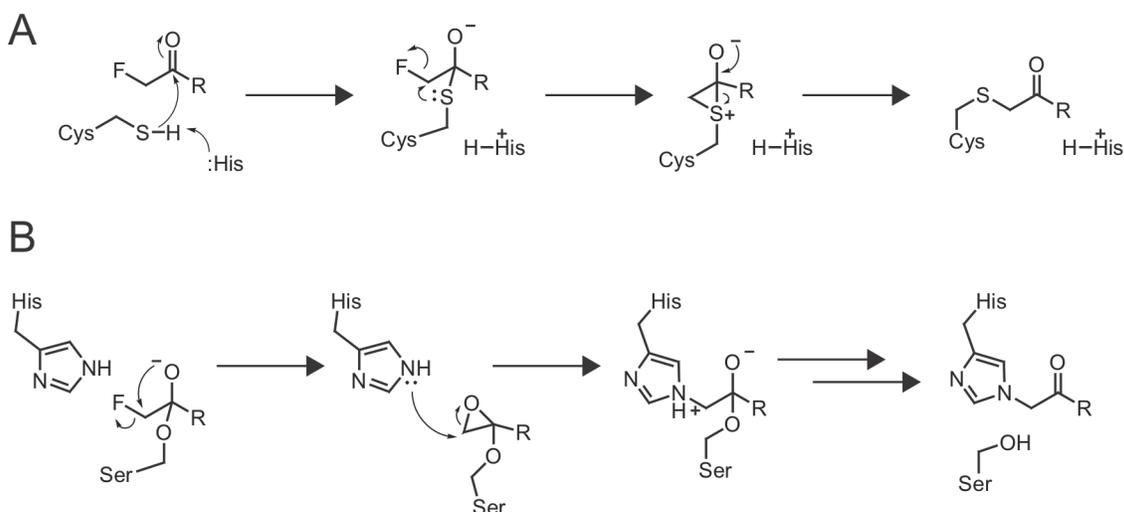
Scheme 1.4: *Fluoromethyl ketones as covalent active site inhibitors. (A) Covalent inhibition mechanism in the presence of an active site cysteine as present in a PKS KS. (B) Covalent inhibition mechanism of an active-site serine as present in a PKS AT.*

many different bioactive molecules, like polyketides and isoprenoids (*Scheme 1.3*) [*20,32*]. Aside from fluoroacetate, various fluorinated building block analogs like fluorinated amino acids such as 3,5-difluoro-β-hydroxytyrosine and alkaloid building blocks such as 5-fluorotryptamine have been used generate fluorinated peptide natural products [*76,77*] and fluorinated alkaloids [*78,79*], respectively.

Given that polyketides use activated propionyl-CoAs and acetyl-CoAs as building blocks for polyketide backbone assembly [*20,22*], the use of fluoroacetyl-CoA as a potential building block for polyketide biosynthesis is an exciting tool for the introduction of fluorine into complex molecules. An important consideration when using fluoroacetate derived building blocks is the potential covalent inhibition of enzymes by alkylation of active site nucleophiles. This is attributed to the flouromethyl ketone motif exhibited in fluoroacetate (*Scheme 1.4*) [*80-82*]. Fluoroacetyl-CoA was first incorporated into a polyketide product as a starter unit for a minimal actinorhodin PKS cluster *in vitro* (*Scheme 1.5*) [*83*]. The formation of this product illustrated that even though the fluoromethyl ketone motif is present in fluoroacetyl-CoA, it can still be incorporated into the minimal PKS shown in *Scheme 1.5*. Additionally, recent studies were also able to directly show that fluoromalonyl-CoA does not covalently inhibit DEBS modules [*32*]. The incorporation of a fluorinated PKS starter unit although exciting, was not entirely surprising given the substrate flexibility observed for polyketide starter units [*84*]. However, as mentioned above, AT domain gatekeeping activity makes incorporation of nonnative extender units difficult [*21,36,37*]. Recent work by our lab has shown that fluoromalonyl-CoA can be incorporated into model polyketide products both *in vivo* and *in vitro*
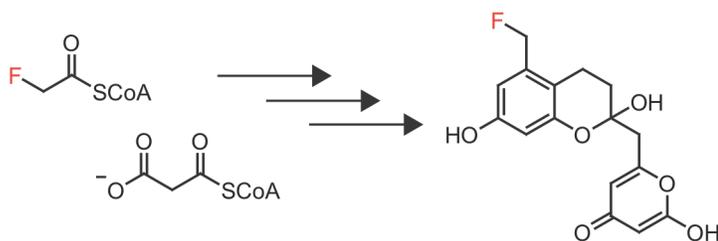
Scheme 1.5: *Biosynthesis of a fluorinated polyketide with fluoroacetyl-CoA. The incubation of a minimal actinorhodin PKS with fluoroacetyl-CoA and malonyl-CoA results in the formation of a fluorinated polyketide by using fluoroacetyl-CoA as starter unit instead of acetyl-CoA, along with seven malonyl-CoAs.*

[*32*]. While the mechanism of fluoromalonyl-CoA incorporation into these model polyketides was not extensively investigated, it laid the groundwork for studies described in this in the following chapters.

1.4   Thesis organization

The goal of this work was to develop methodology for site-specific incorporation of fluorine into the backbone of polyketides. While the immediate goal remained the *in vivo* production of full-length site-specifically fluorinated 6-dEB, this product was proposed as a model system for the site-specific incorporation of fluorine into the backbone of any polyketide of interest. Polyketide biosynthesis is complex and formation of products is not a clear indicator of successful method development. The work described here aims to isolate individual steps involved in PKS C-C bond formation with the ultimate goal of understanding the mechanism of unnatural extender unit incorporation.

To that end, Chapter 2 introduces a detailed study of the molecular mechanism of fluoromalonyl-CoA incorporation into both *cis*- and *trans*-AT type I modular PKSs as well as initial applications inspired by knowledge gained from the mechanistic study. Chapter 3 describes mechanistically driven mutagenesis screens in order to discover a fluoromalonyl-CoA selective *trans*-AT variant along with initial studies of potential *in vivo* heterologous host systems for the production of these novel macrolides. Chapter 4 capitalizes on the theoretical and experimental knowledge gained from the work in Chapter 2 and extends its applications to other halogenated extender units.

1.5   References

1.   Demain, A. L., Fang, A. The natural functions of secondary metabolites. *Adv Biochem Eng Biotechnol* **2000**, *69* 1-39.

2.   Dias, D. A., Urban, S., Roessner, U. A historical overview of natural products in drug discovery. *Metabolites* **2012**, *2* (2) 303-336.

3.   Cragg, G. M., Newman, D. J. Natural products: A continuing source of novel drug leads. *Biochim Biophys Acta* **2013**, *1830* (6) 3670-3695.

4.   Borchardt, J. K. The beginnings of drug therapy: Ancient mesopotamian medicine. *Drug News Perspect* **2002**, *15* (3) 187-192.

5.   Asgary, S., Naderi, G. A., Shams Ardekani, M. R., Sahebkar, A., Airin, A., Aslani, S., Kasher, T., Emami, S. A. Chemical analysis and biological activities of *Cupressus sempervirens var. Horizontalis* essential oils. *Pharm Biol* **2013**, *51* (2) 137-144.

6.   Rufino, A. T., Ribeiro, M., Judas, F., Salgueiro, L., Lopes, M. C., Cavaleiro, C., Mendes, A. F. Anti-inflammatory and chondroprotective activity of (+)-alpha-pinene: Structural and enantiomeric selectivity. *J Nat Prod* **2014**, *77* (2) 264-269.

7.   Quintao, N. L., da Silva, G. F., Antonialli, C. S., Rocha, L. W., Cechinel Filho, V., Ciccio, J. F. Chemical composition and evaluation of the anti-hypernociceptive effect of the

essential oil extracted from the leaves of ugni myricoides on inflammatory and neuropathic models of pain in mice. *Planta Med* **2010**, *76* (13) 1411-1418.

8. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *Brit J Exp Pathol* **1929**, *10* (3) 226-236.

9. Clatworthy, A. E., Pierson, E., Hung, D. T. Targeting virulence: A new paradigm for antimicrobial therapy. *Nat Chem Biol* **2007**, *3* (9) 541-548.

10. Renwick, M. J., Brogan, D. M., Mossialos, E. A systematic review and critical assessment of incentive strategies for discovery and development of novel antibiotics. *J Antibiot (Tokyo)* **2016**, *69* (2) 73-88.

11. Lewis, K. Platforms for antibiotic discovery. *Nat Rev Drug Discov* **2013**, *12* (5) 371-387.

12. Fair, R. J., Tor, Y. Antibiotics and bacterial resistance in the 21st century. *Perspect Medicin Chem* **2014**, *6* 25-64.

13. Newman, D. J., Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *J Nat Prod* **2007**, *70* (3) 461-477.

14. Blair, J. M., Webber, M. A., Baylay, A. J., Ogbolu, D. O., Piddock, L. J. Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol* **2015**, *13* (1) 42-51.

15. Wright, P. M., Seiple, I. B., Myers, A. G. The evolving role of chemical synthesis in antibacterial drug discovery. *Angew Chem Int Ed Engl* **2014**, *53* (34) 8840-8869.

16. Witting, K., Sussmuth, R. D. Discovery of antibacterials and other bioactive compounds from microorganisms-evaluating methodologies for discovery and generation of non-ribosomal peptide antibiotics. *Curr Drug Targets* **2011**, *12* (11) 1547-1559.

17. Patani, G. A., LaVoie, E. J. Bioisosterism: A rational approach in drug design. *Chem Rev* **1996**, *96* (8) 3147-3176.

18. Walsh, C. T., Wencewicz, T. A. Prospects for new antibiotics: A molecule-centered perspective. *J Antibiot (Tokyo)* **2014**, *67* (1) 7-22.

19. von Nussbaum, F., Brands, M., Hinzen, B., Weigand, S., Habich, D. Antibacterial natural products in medicinal chemistry--exodus or revival? *Angew Chem Int Ed Engl* **2006**, *45* (31) 5072-5129.

20. Staunton, J., Weissman, K. J. Polyketide biosynthesis: A millennium review. *Nat Prod Rep* **2001**, *18* (4) 380-416.

21. Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A., Cane, D. E. Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu Rev Biochem* **2007**, *76* 195-221.

22. Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat Prod Rep* **2012**, *29* (10) 1050-1073.

23. O'Brien, R. V., Davis, R. W., Khosla, C., Hillenmeyer, M. E. Computational identification and analysis of orphan assembly-line polyketide synthases. *J Antibiot (Tokyo)* **2014**, *67* (1) 89-97.

24. Weissman, K. J., Leadlay, P. F. Combinatorial biosynthesis of reduced polyketides. *Nat Rev Microbiol* **2005**, *3* (12) 925-936.

25. Khosla, C., Herschlag, D., Cane, D. E., Walsh, C. T. Assembly line polyketide synthases: Mechanistic insights and unsolved problems. *Biochemistry* **2014**, *53* (18) 2875-2883.

26. Kennedy, J. Mutasynthesis, chemobiosynthesis, and back to semi-synthesis: Combining synthetic chemistry and biosynthetic engineering for diversifying natural products. *Nat Prod Rep* **2008**, *25* (1) 25-34.

27. Dunn, B. J., Khosla, C. Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. *J R Soc Interface* **2013**, *10* (85) 20130297.

28. Ray, L., Moore, B. S. Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Nat Prod Rep* **2016**, *33* (2) 150-161.

29. Oliynyk, M., Brown, M. J., Cortes, J., Staunton, J., Leadlay, P. F. A hybrid modular polyketide synthase obtained by domain swapping. *Chem Biol* **1996**, *3* (10) 833-839.

30. Zhang, W., Liu, J. Recent advances in understanding and engineering polyketide synthesis. *F1000Res* **2016**, *5*.

31. Lowry, B., Li, X., Robbins, T., Cane, D. E., Khosla, C. A turnstile mechanism for the controlled growth of biosynthetic intermediates on assembly line polyketide synthases. *ACS Cent Sci* **2016**, *2* (1) 14-20.

32. Walker, M. C., Thuronyi, B. W., Charkoudian, L. K., Lowry, B., Khosla, C., Chang, M. C. Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* **2013**, *341* (6150) 1089-1094.

33. Ad, O., Thuronyi, B. W., Chang, M. C. Elucidating the mechanism of fluorinated extender unit loading for improved production of fluorine-containing polyketides. *Proc Natl Acad Sci U S A* **2017**, *114* (5) E660-E668.

34. Dunn, B. J., Watts, K. R., Robbins, T., Cane, D. E., Khosla, C. Comparative analysis of the substrate specificity of *trans-* versus *cis*-acyltransferases of assembly line polyketide synthases. *Biochemistry* **2014**, *53* (23) 3796-3806.

35. Koryakina, I., Kasey, C., McArthur, J. B., Lowell, A. N., Chemler, J. A., Li, S., Hansen, D. A., Sherman, D. H., Williams, G. J. Inversion of extender unit selectivity in the erythromycin polyketide synthase by acyltransferase domain engineering. *ACS Chem Biol* **2017**, *12* (1) 114-123.

36. Dunn, B. J., Cane, D. E., Khosla, C. Mechanism and specificity of an acyltransferase domain from a modular polyketide synthase. *Biochemistry* **2013**, *52* (11) 1839-1841.

37. Bonnett, S. A., Rath, C. M., Shareef, A. R., Joels, J. R., Chemler, J. A., Hakansson, K., Reynolds, K., Sherman, D. H. Acyl-CoA subunit selectivity in the pikromycin polyketide synthase PIKAIV: Steady-state kinetics and active-site occupancy analysis by FTICR-MS. *Chem Biol* **2011**, *18* (9) 1075-1081.

38. Helfrich, E. J., Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep* **2016**, *33* (2) 231-316.

39. 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* **1999**, *38* (5) 1643-1651.

40. Reeves, C. D., Murli, S., Ashley, G. W., Piagentini, M., Hutchinson, C. R., McDaniel, R. Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry* **2001**, *40* (51) 15464-15470.

41. Patterson, D. M., Nazarova, L. A., Prescher, J. A. Finding the right (bioorthogonal) chemistry. *ACS Chem Biol* **2014**, *9* (3) 592-605.

42. Sletten, E. M., Bertozzi, C. R. Bioorthogonal chemistry: Fishing for selectivity in a sea of functionality. *Angew Chem Int Ed Engl* **2009**, *48* (38) 6974-6998.

43. Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., Schulz, F. Enzyme-directed mutasynthesis: A combined experimental and theoretical approach to substrate recognition of a polyketide synthase. *ACS Chem Biol* **2013**, *8* (2) 443-450.

44. Bravo-Rodriguez, K., Klopries, S., Koopmans, K. R., Sundermann, U., Yahiaoui, S., Arens, J., Kushnir, S., Schulz, F., Sanchez-Garcia, E. Substrate flexibility of a mutated acyltransferase domain and implications for polyketide biosynthesis. *Chem Biol* **2015**, *22* (11) 1425-1430.

45. Koryakina, I., McArthur, J. B., Draelos, M. M., Williams, G. J. Promiscuity of a modular polyketide synthase towards natural and non-natural extender units. *Org Biomol Chem* **2013**, *11* (27) 4449-4458.

46. Jensen, P. R., Chavarria, K. L., Fenical, W., Moore, B. S., Ziemert, N. Challenges and triumphs to genomics-based natural product discovery. *J Ind Microbiol Biotechnol* **2014**, *41* (2) 203-209.

47. Li, J., Jaitzig, J., Lu, P., Sussmuth, R. D., Neubauer, P. Scale-up bioprocess development for production of the antibiotic valinomycin in escherichia coli based on consistent fed-batch cultivations. *Microb Cell Fact* **2015**, *14* 83.

48. Musiol-Kroll, E. M., Zubeil, F., Schafhauser, T., Hartner, T., Kulik, A., McArthur, J., Koryakina, I., Wohlleben, W., Grond, S., Williams, G. J., Lee, S. Y., Weber, T. Polyketide bioderivatization using the promiscuous acyltransferase KirCII. *ACS Synth Biol* **2017**, *6* (3) 421-427.

49. Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., Piel, J. Exploiting the mosaic structure of *trans*-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat Biotechnol* **2008**, *26* (2) 225-233.

50. Koryakina, I., McArthur, J., Randall, S., Draelos, M. M., Musiol, E. M., Muddiman, D. C., Weber, T., Williams, G. J. Poly specific *trans*-acyltransferase machinery revealed via engineered acyl-CoA synthetases. *ACS Chem Biol* **2013**, *8* (1) 200-208.

51. Ye, Z., Musiol, E. M., Weber, T., Williams, G. J. Reprogramming acyl carrier protein interactions of an acyl-CoA promiscuous *trans*-acyltransferase. *Chem Biol* **2014**, *21* (5) 636-646.

52. Zhao, C., Coughlin, J. M., Ju, J., Zhu, D., Wendt-Pienkowski, E., Zhou, X., Wang, Z., Shen, B., Deng, Z. Oxazolomycin biosynthesis in *Streptomyces albus* JA3453 featuring an

"acyltransferase-less" type I polyketide synthase that incorporates two distinct extender units. *J Biol Chem* **2010**, *285* (26) 20097-20108.

53. Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep* **2010**, *27* (7) 996-1047.

54. Wong, F. T., Chen, A. Y., Cane, D. E., Khosla, C. Protein-protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. *Biochemistry* **2010**, *49* (1) 95-102.

55. Muller, K., Faeh, C., Diederich, F. Fluorine in pharmaceuticals: Looking beyond intuition. *Science* **2007**, *317* (5846) 1881-1886.

56. Smart, B. E. Fluorine substituent effects (on bioactivity). *J Fluorine Chem* **2001**, *109* (1) 3-11.

57. Purser, S., Moore, P. R., Swallow, S., Gouverneur, V. Fluorine in medicinal chemistry. *Chem Soc Rev* **2008**, *37* (2) 320-330.

58. Hagmann, W. K. The many roles for fluorine in medicinal chemistry. *J Med Chem* **2008**, *51* (15) 4359-4369.

59. Bondi, A. Van der waals volumes + radii. *J Phys Chem-Us* **1964**, *68* (3) 441-&.

60. Berkessel, A., Adrio, J. A., Huttenhain, D., Neudorfl, J. M. Unveiling the "booster effect" of fluorinated alcohol solvents: Aggregation-induced conformational changes and cooperatively enhanced h-bonding. *J Am Chem Soc* **2006**, *128* (26) 8421-8426.

61. Morgenthaler, M., Schweizer, E., Hoffmann-Roder, A., Benini, F., Martin, R. E., Jaeschke, G., Wagner, B., Fischer, H., Bendels, S., Zimmerli, D., Schneider, J., Diederich, F., Kansy, M., Muller, K. Predicting and tuning physicochemical properties in lead optimization: Amine basicities. *ChemMedChem* **2007**, *2* (8) 1100-1115.

62. Bohm, H. J., Banner, D., Bendels, S., Kansy, M., Kuhn, B., Muller, K., Obst-Sander, U., Stahl, M. Fluorine in medicinal chemistry. *Chembiochem* **2004**, *5* (5) 637-643.

63. Park, B. K., Kitteringham, N. R. Effects of fluorine substitution on drug metabolism: Pharmacological and toxicological implications. *Drug Metab Rev* **1994**, *26* (3) 605-643.

64. Liang, T., Neumann, C. N., Ritter, T. Introduction of fluorine and fluorine-containing functional groups. *Angew Chem Int Ed Engl* **2013**, *52* (32) 8214-8264.

65. Fernandes, P., Martens, E., Bertrand, D., Pereira, D. The solithromycin journey-it is all in the chemistry. *Bioorg Med Chem* **2016**, *24* (24) 6420-6428.

66. Piccinelli, G., Fernandes, P., Bonfanti, C., Caccuri, F., Caruso, A., De Francesco, M. A. *In vitro* activity of solithromycin against erythromycin-resistant *Streptococcus agalactiae*. *Antimicrob Agents Chemother* **2014**, *58* (3) 1693-1698.

67. Campbell, M. G., Ritter, T. Late-stage fluorination: From fundamentals to application. *Org Process Res Dev* **2014**, *18* (4) 474-480.

68. Willis, N. J., Fisher, C. A., Alder, C. M., Harsanyi, A., Shukla, L., Adams, J. P., Sandford, G. Sustainable synthesis of enantiopure fluorolactam derivatives by a selective direct fluorination - amidase strategy. *Green Chemistry* **2016**, *18* (5) 1313-1318.

69.   Tredwell, M., Tenza, K., Pacheco, M. C., Gouverneur, V. Synthesis of novel enantiopure fluorinated building blocks from acyclic chiral allylsilanes. *Org Lett* **2005**, *7* (20) 4495-4497.

70.   Saadi, J., Wennemers, H. Enantioselective aldol reactions with masked fluoroacetates. *Nat Chem* **2016**, *8* (3) 276-280.

71.   O'Hagan, D., Deng, H. Enzymatic fluorination and biotechnological developments of the fluorinase. *Chem Rev* **2015**, *115* (2) 634-649.

72.   O'Hagan, D., Schaffrath, C., Cobb, S. L., Hamilton, J. T., Murphy, C. D. Biochemistry: Biosynthesis of an organofluorine molecule. *Nature* **2002**, *416* (6878) 279.

73.   Schaffrath, C., Deng, H., O'Hagan, D. Isolation and characterisation of 5'-fluorodeoxyadenosine synthase, a fluorination enzyme from *Streptomyces cattleya*. *FEBS Lett* **2003**, *547* (1-3) 111-114.

74.   Brand, M. D., Evans, S. M., Mendes-Mourao, J., Chappell, J. B. Fluorocitrate inhibition of aconitate hydratase and the tricarboxylate carrier of rat liver mitochondria. *Biochem J* **1973**, *134* (1) 217-224.

75.   Villafranca, J. J., Platus, E. Fluorocitrate inhibition of aconitase. Reversibility of the inactivation. *Biochem Biophys Res Commun* **1973**, *55* (4) 1197-1207.

76.   Oldach, F., Al Toma, R., Kuthning, A., Caetano, T., Mendo, S., Budisa, N., Sussmuth, R. D. Congeneric lantibiotics from ribosomal *in vivo* peptide synthesis with noncanonical amino acids. *Angew Chem Int Ed Engl* **2012**, *51* (2) 415-418.

77.   Weist, S., Bister, B., Puk, O., Bischoff, D., Pelzer, S., Nicholson, G. J., Wohlleben, W., Jung, G., Sussmuth, R. D. Fluorobalhimycin - a new chapter in glycopeptide antibiotic research. *Angew Chem Int Ed Engl* **2002**, *41* (18) 3383-3385.

78.   Runguphan, W., Maresh, J. J., O'Connor, S. E. Silencing of tryptamine biosynthesis for production of nonnatural alkaloids in plant culture. *Proc Natl Acad Sci U S A* **2009**, *106* (33) 13673-13678.

79.   Walker, M. C., Chang, M. C. Natural and engineered biosynthesis of fluorinated natural products. *Chem Soc Rev* **2014**, *43* (18) 6527-6536.

80.   Powers, J. C., Asgian, J. L., Ekici, O. D., James, K. E. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev* **2002**, *102* (12) 4639-4750.

81.   Shaw, E., Angliker, H., Rauber, P., Walker, B., Wikstrom, P. Peptidyl fluoromethyl ketones as thiol protease inhibitors. *Biomed Biochim Acta* **1986**, *45* (11-12) 1397-1403.

82.   Thuronyi, B. W., Chang, M. C. Synthetic biology approaches to fluorinated polyketides. *Acc Chem Res* **2015**, *48* (3) 584-592.

83.   Hong, H., Spiteller, D., Spencer, J. B. Incorporation of fluoroacetate into an aromatic polyketide and its influence on the mode of cyclization. *Angew Chem Int Edit* **2008**, *47* (32) 6028-6032.

84.   Murli, S., MacMillan, K. S., Hu, Z., Ashley, G. W., Dong, S. D., Kealey, J. T., Reeves, C. D., Kennedy, J. Chemobiosynthesis of novel 6-deoxyerythronolide B analogues by

mutation of the loading module of 6-deoxyerythronolide B synthase. *Appl Environ Microbiol* **2005**, *71* (8) 4503-4509.

Chapter 2: *Elucidating the mechanism of fluorinated extender unit loading for improved production of fluorine-containing polyketides*

2.1   Introduction

Polyketide natural products represent a large class of bioactive compounds with a broad range of medicinal properties, including antibiotic and anticancer activity. Despite their diversity in structure, they are assembled with a common mechanism involving multiple Claisen condensations of acetate-derived extender units catalyzed by polyketide synthase (PKS) enzymes [1,2]. In addition to setting the oxidation state of the β-carbon at each polymerization step, modular PKSs can also select the α-substituent on the extender unit, which is predominantly derived from either acetate (R = H) or propionate (R = Me) [1,2]. The breadth of compounds (>10,000) that can be made from these two simple building blocks is quite considerable and many efforts have thus centered on increasing the types of extender units that can be incorporated into polyketides [1,3,4].

We have focused specifically on the insertion of fluorinated monomers by PKSs as an approach to merge the advantages of complex natural product scaffolds with the effectiveness of fluorine medicinal chemistry [5]. In this regard, the type I modular PKSs are an especially attractive target for engineering as their organization and structure lends itself to connecting sequence with product structure [1,2,6]. Previous studies have shown that the acyltransferase (AT) domain within the PKS (*cis*-AT) is essential to determining extender unit identity and can be engineered to enable insertion of alternative extender units [7-9]. Alternatively, the intrinsic selectivity of the *cis*-AT can be bypassed by inactivation and complementation with a standalone AT enzyme (*trans*-AT) as another approach to introducing non-native substituents [10-12]. While these methods are effective for altering product structure, product yield can often diminish in these engineered systems and thereby limit their scalability. As such, we set out to elucidate the mechanism by which alternative extender units are accepted or rejected in these systems in order to expand our understanding of the native process of extender unit selection as well as to define design principles for PKS engineering.

We initiated these studies using the 6-deoxyerythronolide B synthase (DEBS), which is responsible for the synthesis of the core of the antibiotic erythromycin, because it has been well characterized and is amenable to dissection to its modular units [2]. We observe that the fluoromalonyl-coenzyme A (CoA) extender unit can be accepted directly by the KS domain in modules where the *cis*-AT is inactivated, with a 2:1 ratio of C-C bond formation to nonproductive decarboxylation to form fluoroacetyl-CoA. Under these conditions, a single fluorine substituent can be introduced but further chain extension cannot easily occur when ligation of the extender unit to the acyl carrier protein (ACP) domain is bypassed. In contrast, complementation with a *trans*-AT leads to quantitative C-C bond formation with a fluorinated extender unit and restores the canonical PKS mechanism with 43% of wild-type (WT) yield. In this system, we can amplify single chain extensions with fluorinated monomers as well as enable further chain growth to produce polyketide products with multiple fluorine substitutions.

2.2   Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Carbenicillin (Cb), isopropyl-β-D-thiogalactopyranoside (IPTG), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride, dithiothreitol (DTT), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, dichloromethane, ethyl acetate, ethylene diamine tetraacetic acid disodium dihydrate

(EDTA), and restriction enzymes were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A sodium salt (CoA), malonyl-CoA, methylmalonyl-CoA, diethylfluoromalonate, malonic acid, diethylmethylmalonate, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenolpyruvate (PEP), adenosine triphosphate sodium salt (ATP), nicotinamide adenine dinucleotide reduced form dipotassium salt (NADH), nicotinamide adenine dinucleotide sodium salt hydrate (NAD$^+$), nicotinamide adenine dinucleotide phosphate sodium salt hydrate (NADP$^+$), nicotinamide adenine dinucleotide phosphate reduced form (NADPH), myokinase, pyruvate kinase, lactate dehydrogenase, α-Ketoglutarate Dehydrogenase, poly(ethyleneimine) solution (PEI), β-mercaptoethanol, thiamine pyrophosphate, alpha ketoglutaric acid, sodium phosphate dibasic hepthydrate, ^{13}C-iodomethane, cysteamine, acetic anhydride, (1-^{13}C)propionic acid (3-^{13}C)sodium propionate, (2-^{13}C)sodium acetate, 1-ethyl-3-(-dimethylaminopropyl)carbodiimide hydrochloride (EDC), 4-dimethylaminopyridine (DMAP), *o*-benzyl hydroxylamine, N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED), acetonitrile, dimethyl sulfoxide (DMSO), ammonium acetate, and ammonium formate were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate was purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, T4 DNA ligase, Phusion DNA polymerase, amylose resin, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 µM in 10 mM Tris-HCl, pH 8.5, and stored at either 4°C for immediate use or -20°C for longer term use. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Amicon Ultra 3,000 MWCO, 10,000 MWCO centrifugal concentrators, and 30,000 MWCO centrifugal concentrators were purchased from EMD Millipore (Billerica, MA). Deuterium oxide, DMSO-d_6, and chloroform-*d* were purchased from Cambridge Isotope Laboratories (Andover, MA). ^{19}F, ^{13}C, and ^{1}H NMR spectra were collected at 25ºC on Bruker AV-600 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley or on a Bruker Biospin 900 MHz spectrometer at the QB3 Central California 900 MHz NMR Facility NMR. Assignments were made based on literature president and reference spectra from authentic standards, where appropriate. High-resolution mass spectral analyses were carried out on a 6530 QTOF Accurate Mass spectrometer purchased from Agilent Technologies.

Bacterial strains. *E. coli* DH10B-T1^R and BL21(de3)T1^R were used for plasmid construction and heterologous protein expression, respectively. *E. coli* BAP1 [*13*] cells were used for heterologous expression of DEBS modules.

Gene and plasmid construction. Standard molecular biology techniques were used for plasmid construction. All PCR amplifications were done using Phusion High Fidelity DNA polymerase. For amplification of GC-rich sequences from *S. erythraea*, PCR reactions were supplemented with DMSO (10% *v/v*) using the standard buffer rather than GC buffer with primer annealing temperatures 6-8°C below the T_m with a maximum temperature of 72ºC. All constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA).

pET21c-His$_6$-Mod3$_{TE}$ACP^0-His$_6$ and pET21c-His$_6$-Mod3$_{TE}$AT^{0}ACP^0-His$_6$ were constructed by amplification from pRGS34 [14] and pAYC136 [15], respectively. For pET21c-His$_6$-Mod3$_{TE}$ACP^0-His$_6$, the S1430A mutation (based on EryAII numbering) was introduced by amplifying pRGS34 with Mod3$_{TE}$ACP^0 F1/R1 and pAYC136 with Mod3$_{TE}$ACP^0 F2/R2 (*Table S1*), which contained a 60 bp overlap. The two PCR fragments where inserted into the BsiWI-EcoRI sites of pRGS34 using the Gibson Protocol. For pET21c-His$_6$-Mod3$_{TE}$AT^{0}ACP^0-His$_6$, the S1430A (based on EryAII numbering) mutation was introduced by amplifying pAYC136 with Mod3$_{TE}$ACP^0 F1/R1 and Mod3$_{TE}$ACP^0 B F2/R2 (*Table S1*), which contained a 60 bp overlap. The two PCR fragments where inserted into the BsiWI-EcoRI sites pAYC136 using the Gibson protocol [16].

pET21c-KSAT^0-His$_6$ and pET21c-KS^{0}AT^0-His$_6$ and pET21c-KS^{0}AT-His$_6$ were constructed by amplification from pAYC02 [17]. For pET21c-KSAT^0-His$_6$, the S651A (based on EryAII numbering) mutation was introduced by amplifying pAYC02 with KS^{0}AT^0 F1 and KS^{0}AT^0 R2 (*Table S1*). The PCR fragment was inserted into the NdeI-BsiWI sites of pAYC02 using the Gibson Protocol. For pET21c-KS^{0}AT-His$_6$, the C202A (based on EryAII numbering) mutation was introduced by amplifying pAYC02 with KS^{0}AT^0 F1/R1 and KS^{0}AT^0 F2/R2 (*Table S1*), which contained a 60 bp overlap. The two PCR fragments where inserted into the NdeI-BsiWI sites pAYC2 using the Gibson protocol. For pET21c-KS^{0}AT^0-His$_6$, the S651A and C202A mutation were introduced by amplifying pET21c-KS^{0}AT-His$_6$ with KS^{0}AT^0 F1 and KS^{0}AT^0 R2 (*Table S1*). The PCR fragment was inserted into the NdeI-BsiWI sites of pAYC02 using the Gibson Protocol [16].

pET28a-His$_6$-KSAT2 was constructed by amplification from pSV272-His$_6$-MBP-DEBS$_{Mod2}$ using the primers pET28a KSAT2 F/R (*Table S1*). The fragment was insterted into the BsaI-BsaI site of pET28agg-RFP-NHis using the Golden Gate method [18].

pET28a-His$_6$-PIKSKR1 was constructed by ligating PIKS-KR1-block1 and PIKS-KR1-block3 into the BsaI-BsaI site of pET28agg-NHis using Golden Gate method [18]. Amplification of PIKS-KR1-block2 with PIKS_KR1 F/R (in order to correct for an error in the initial design of the gBlock) enabled the ligation of the linear product of the Golden Gate reaction with the amplified DNA using the Gibson Protocol [16].

Expression of His-tagged and MBP-tagged proteins. TB (1 L) in a 2.8 L Fernbach baffled shake flask was inoculated to OD$_{600}$ = 0.05 with an overnight TB culture of freshly transformed *E. coli* containing the appropriate overexpression plasmid. The cultures were grown at 37°C at 200 rpm to OD$_{600}$ = 0.6 to 0.8 at which point cultures were cooled on ice for 20 min, followed by induction of protein expression with 0.25 mM IPTG and overnight growth at 16°C. Cell pellets were harvested by centrifugation at 9,800 × *g* for 7 min at 4ºC and stored at -80ºC.

Purification of His$_6$-rpMatB, His$_6$-Epi,DszAT-His$_6$, and His$_6$-PiksKR1. Frozen cell pellets were thawed and resuspended at 5 mL/g cell paste with Buffer A (50 mM sodium phosphate, 300 mM sodium chloride, 20% (*v/v*) glycerol, 20 mM BME, pH 7.5) containing imidazole (20 mM). Complete EDTA-free protease inhibitor cocktail (Roche) was added to the lysis buffer before resuspension. The cell paste was homogenized before lysis by passage through a French Pressure cell (Thermo Scientific; Waltham, MA) at 14,000 psi. The lysate was centrifuged at 15,300 × *g* for 20 min at 4°C to separate the soluble and insoluble fractions. DNA was precipitated in the soluble fraction by addition of 0.015% (*w/v*) poly(ethyleneimine). The precipitated DNA was removed by centrifugation at 15,300 × *g* for 20 min at 4°C. The remaining soluble lysate was

diluted three-fold with Buffer A containing imidazole (20 mM) and loaded onto a Ni-NTA agarose column (Qiagen, 1 mL resin/g cell paste) by gravity flow or on an ÄKTA purifier FPLC (2 mL/min; GE Healthcare; Piscataway, NJ). The column was washed with Buffer A until the eluate reached an $A_{280\,nm} < 0.05$ or was negative for protein content by Bradford assay (Bio-Rad).

His$_6$-Epi and His$_6$-MatB, and DszAT-His$_6$. His$_6$-Epi, His$_6$-MatB, and DszAT-His$_6$ were eluted using a linear gradient from 0 to 300 mM imidazole in Buffer A over 30 column volumes.

His$_6$-PIKS KR1. The column was washed with 5 to 10 column volumes of Buffer A supplemented with 20 mM imidazole. The protein was then eluted with 250 mM imidazole in Buffer A.

Fractions containing the target protein were pooled by $A_{280\,nm}$ and concentrated using either an Amicon Ultra spin concentrator (3 kDa MWCO, Millipore) or an Amicon ultrafiltration cell under nitrogen flow (65 psi) using a membrane with an appropriate nominal molecular weight cutoff (Ultracel-5 or YM10, Millipore). Protein was then exchanged into Buffer C (50 mM HEPES, 100 mM sodium chloride, 2.5 mM EDTA, 2.5 mM DTT, 20% (*v/v*) glycerol, pH 7.5) using a Sephadex G-25 column (Sigma-Aldrich, bead size 50-150 μm, 10 mL resin/mL protein solution), then concentrated again before storage. Final protein concentrations before storage were estimated using the $\varepsilon_{280\,nm}$ calculated by ExPASY ProtParam as follows: His$_6$-MatB: 34.2 mg/mL ($\varepsilon_{280\,nm}$ = 40,340 M^{-1} cm^{-1}), His$_6$-Epi: 10.5 mg/mL ($\varepsilon_{280\,nm}$ = 11,460 M^{-1} cm^{-1}), DszAT-His$_6$: 13.3 mg/mL ($\varepsilon_{280\,nm}$ = 17,420 M^{-1} cm^{-1}), His$_6$-PIKS$_{Mod1}$ KR: 15.4 mg/mL ($\varepsilon_{280\,nm}$ = 73,450 M^{-1} cm^{-1}) All proteins were aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C.

Purification of His$_6$-KSAT2, KSAT3-His$_6$, KSAT3^0-His$_6$, KS^{0}AT3-His$_6$, KS^{0}AT3^0-His$_6$ and KSAT6-His$_6$. The KSAT didomains were heterologously expressed in *E. coli* BL21(de3) T1^R as described above. Frozen cell pellets were thawed and resuspended at 5 mL/g cell paste with Buffer A containing imidazole (10 mM). Complete EDTA-free protease inhibitor cocktail (Roche) was added to the lysis buffer before resuspension. The cell paste was homogenized before lysis by passage through a French Pressure cell (Thermo Scientific; Waltham, MA) at 14,000 psi. The lysate was centrifuged at 15,300 × *g* for 20 min at 4°C to separate the soluble and insoluble fractions. DNA was precipitated in the soluble fraction by addition of 0.015% (w/v) poly(ethyleneimine). The precipitated DNA was removed by centrifugation at 15,300 × *g* for 20 min at 4°C. The remaining soluble lysate was diluted three-fold with Buffer B containing 10 mM imidazole and loaded onto a Ni-NTA agarose column (Qiagen, 1 mL resin/g cell paste) by gravity flow. The column was washed with Buffer A until the eluate reached an $A_{280\,nm} < 0.05$ or was negative for protein content by Bradford assay (Bio-Rad). The column was washed with 20 column volumes with Buffer A supplemented with 25 mM imidazole. The protein was then eluted with Buffer D (50 mM sodium phosphate, 50 mM sodium chloride, 20% (*v/v*) glycerol, 20 mM BME,pH 7.5) containing 250 mM Imidazole. Fractions containing the target protein were pooled by $A_{280\,nm}$ and concentrated using an Amicon Ultra spin concentrator (30 kDa MWCO, Millipore). Protein was then exchanged into Buffer C (50 mM HEPES, 100 mM sodium chloride, 2.5 mM EDTA, 2.5 mM DTT, 20% (*v/v*) glycerol, pH 7.5) using a Sephadex G-25 column (Sigma-Aldrich, bead size 50-150 μm, 10 mL resin/mL protein solution), then concentrated again before storage. Final protein concentrations before storage were estimated using the $\varepsilon_{280\,nm}$ calculated by ExPASY ProtParam as follows: His$_6$-KSAT2: $\varepsilon_{280\,nm}$ = 93,390 M^{-1} cm^{-1}, KSAT3-His$_6$: $\varepsilon_{280\,nm}$ = 92,360 M^{-1} cm^{-1}) KSAT6-His$_6$: $\varepsilon_{280\,nm}$ = 71,390 M^{-1} cm^{-1}. All proteins were aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C with a final concentration of 10-30 mg/mL.

Purification of Mod3$_{TE}$-His$_6$, AT^0 Mod3$_{TE}$-His$_6$, AT^{0}ACP^0 Mod3$_{TE}$-His$_6$, and ACP^0 Mod3$_{TE}$-His$_6$, Mod6$_{TE}$-His$_6$, and AT^0 Mod6$_{TE}$-His$_6$. The His-tagged DEBS modules with thioesterase constructs were heterologously expressed in *E. coli* BAP1 as described above. Cleared cell lysates were prepared in Buffer B (200 mM sodium phosphate, 200 mM sodium chloride, 30% (*v/v*) glycerol, 2.5 mM EDTA, 2.5 mM DTT, pH 7.5) as described above, diluted three-fold with Buffer A, and passed over a Ni-NTA agarose column (Qiagen, approximately 1 mL/g cell paste) on an ÄKTApurifier FPLC. The column was washed with Buffer A until the eluate reached an A$_{280\ nm}$ < 0.05. Protein was eluted with Buffer D containing 100 mM imidazole. The eluate was diluted two-fold with Buffer E (50 mM HEPES, 2.5 mM EDTA, 2.5 mM DTT, 20% glycerol, pH 7.5), loaded onto a HiTrap Q HP column (GE Healthcare, 5 mL), and eluted with a linear gradient from 0 to 1 M sodium chloride in Buffer E over 30 column volumes (5 mL/min). Fractions containing the target protein (eluted at ~350 mM sodium chloride) were pooled by A$_{280\ nm}$ and concentrated with an Amicon Ultra 30,000 MWCO centrifugal concentrator. The protein was flash-frozen in liquid nitrogen and stored at -80°C at a final concentration of 20-30 mg/mL, which was estimated using the calculated $\varepsilon_{280\ nm}$ (Mod3$_{TE}$-His$_6$ constructs: 203,280 M^{-1}cm^{-1}; Mod6$_{TE}$-His$_6$ constructs: 206,260 M^{-1}cm^{-1}).

Purification of His$_6$-MBP-Mod2 and His$_6$-MBP-AT^0 Mod2. Cleared lysates were prepared as described for other DEBS modules with the exception buffer A being supplemented with Buffer F (20 mM TRIS-HCl, 300 mM Sodium Chloride, and 1 mM DTT, 20% (*v/v*) glycerol, pH 7.5). The clear lysate was diluted three-fold with Buffer F and loaded by gravity onto Amylose resin (2 mL/g cell paste) into a fritted column and washed with Buffer F until the eluate reached A$_{280\ nm}$ < 0.05. The protein was eluted with Buffer F containing 10 mM maltose and concentrated to ~1 mg/mL using an Amicon Ultra 30,000 MWCO centrifugal concentrator. The protein was then dialyzed overnight against Buffer E containing 50 mM NaCl with TEV protease (1 mg/50 mg protein substrate) to remove the MBP tag. The protein was loaded onto a HiTrap Q HP column and eluted by a linear gradient from 0 to 1 M NaCl in Buffer E over 30 column volumes. Fractions containing the desired protein were identified by SDS-PAGE (eluting at ~350 mM NaCl), pooled, and concentrated in an Amicon Ultra 30,000 MWCO centrifugal concentrators. Protein aliquots were flash-frozen in liquid nitrogen and stored at -80°C at a final concentration of 8-12 mg/mL, which was estimated using the calculated $\varepsilon_{280\ nm}$ (158,360 M^{-1}cm^{-1}).

Fluoromalonate and methylmalonate. Fluoromalonate and methylonate were prepared as described previously [10]. Diethylfluoromalonate and diethylmethylmalonate (0.5 mL, 3.2 mmol) were saponified with methanolic sodium hydroxide (2 M, 3.5 mL) in dichloromethane and methanol (9:1 v/v, 32 mL) and the sodium salt isolated by filtration through a Büchner funnel with a fine porosity glass frit [19].

Fluoromalonyl-CoA. Fluoromalonyl-CoA was prepared enzymatically from fluoromalonate and CoA using MatB and ATP as described previously [10]. The reaction mixture (10 mL) contained 100 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (5 mM), TCEP (2.5 mM), magnesium chloride (5 mM), fluoromalonate (10 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (36 U), myokinase (20 U), CoA (2 mM) and MatB (10 μM). The mixture was incubated at 37°C for 2 h and lyophilized overnight. The residue was dissolved in water (1.6 mL) and acidified to pH ~2 by addition of 70% (v/v) perchloric acid (160 μL). Insoluble material was removed by centrifugation at 18,000 × *g* for 10 min. The supernatant was adjusted to pH 6 by addition of 10 M sodium hydroxide (100 μL) and desalted on an Agilent 1200 HPLC system using a Zorbax Eclipse XDB C-18 column (5 μm, 9.4 × 250 mm, Agilent) with a linear gradient from 0

to 10% methanol over 9 min with 50 mM sodium phosphate, 25 mM trifluoroacetic acid, pH 4.5 as the aqueous mobile phase (3 mL/min). Fractions eluting near the void volume, containing both fluoromalonyl-CoA and CoA, were lyophilized overnight, dissolved in water (1 mL), and purified using a Zorbax Eclipse XDB C-18 column (5 μm, 9.4 × 250 mm) with a linear gradient from 0 to 10% acetonitrile over 30 min with 0.1% formic acid as the aqueous mobile phase (3 mL/min). Fractions were screened by ESI-MS and those containing pure fluoromalonyl-CoA were lyophilized overnight. The fluoromalonyl-CoA solutions were stored at -80°C.

ESI-MS screening method for fluoromalonyl-CoA. Fractions were screened as described previously [*10*]. Preparative HPLC fractions were screened on an Agilent 1260 HPLC system using a Kinetex 50 XB-C18 column (5 μm, 5.0 × 30 mm, Phenomenex) with a linear gradient from 0 to 100% acetonitrile over 3 min with 0.1% formic acid as the aqueous mobile phase (0.8 mL/min). Mass spectra were collected on an Agilent 6130 single quadrupole MS with ESI source, operating in negative and positive ion scan mode [*10*].

N-acetylcysteamine (SNAC). Cysteamine (2 g, 13 mmol) was added to a 250 mL Erlenmeyer flask containing 100 mL of water. Sodium bicarbonate (4.4 g, 5.2 mmol) and potassium hydroxide (~988 mg, 1 pellet) were added and the solution was left to stir at r.t for 5 min. Acetic anhydride (1.66 mL, 16.5 mmol) was added dropwise over the course of 2 min and the reaction was left stirring at r.t for 1 h. The solution was then adjusted to pH 7 with 6 M hydrochloric acid and the product was extracted with 3 × 200 mL of ethyl acetate. All organic washes were combined and dried with magnesium sulfate and solvent was removed by rotary evaporation to produce a clear oil at >99% yield. **^{1}H-NMR** (400 MHz, CDCl$_3$ = 7.26 ppm)**:** δ 1.36 (t, 1H, 8.5 Hz), 2.00 (s, 3H), 2.62-2.70 (m, 2H), 3.40-3.43 (m, 2H), 6.10 (s, N-H). **^{13}C-NMR** (101 MHz, CDCl$_3$ = 77.36 ppm) δ 23.3, 24.6, 42.5, 170.7.

N-acetylcysteamine thioester of 2-methyl-3-oxopentanoic acid (oxoNDK-SNAC). The protocol was adapted from a previous study [*20*]. 2-methyl-3-oxopentanoic acid ethyl ester (316 mg, 2 mmol) was saponified in 1 M sodium hydroxide at r.t for 2 h. The reaction was acidified to pH 1 and 2-methyl-3-oxopentanoic acid was extracted in 3 × 15 mL ethyl acetate. All organic washes were combined and dried with magnesium sulfate and solvent was removed by rotary evaporation. 2-methyl-3-oxopentanoic acid was then dissolved in 10 mL of dichloromethane in a 25 mL round bottom flask on ice. SNAC (310 mg, 2 mmol) and DMAP (40 mg, 0.3 mmol) where added and the reaction was allowed to stir for 5 min. EDC (422 mg, 2.2 mmol) was added and the reaction was stirred at room temperature for 24 h. The reaction was washed with 2 × 20 mL of saturated ammonium chloride. The aquous fractions were pooled and washed further with 3 × 20 mL ethyl acetate. All organic washes were combined and dried with magnesium sulfate and solvent was removed by rotary evaporation to produce a yellowish oil at 80% yield. **^{1}H-NMR** (400 MHz, CDCl$_3$ = 7.26 ppm)**:** δ 1.06 (t, 3H, 7.2 Hz), 1.38 (d, 3H, 7.1 Hz), 1.96 (s, 3H), 2.47-2.62 (m, 2H), 3.00-3.13 (m, 2H), 3.37-3.49 (m, 2H), 3.79 (q, 1H, 7.0 Hz), 5.89 (s, N-H). **ESI-MS [M-H]$^-$:** calculated for C$_{10}$H$_{17}$NO$_3$S, 230.1 *m/z*, found 230.1 *m/z*.

Enzymatic synthesis of the N-acetylcysteamine thioester of (2*S*,3*R*)-2-methyl-3-hydroxypentanoic acid (NDK-SNAC) using PIKS KR1. This protocol was adapted from a previous study [*21*]. Oxo-NDK-SNAC was dissolved in DMSO (1 M, 10 mL total volume) and added to a solution containing 300 mM HEPES pH 7.5, 10% glycerol, sodium chloride (100 mM), D-glucose (300 mM), and NADP$^+$ (0.1 mM). The reaction was initiated by the addition of glucose-1-dehydrogenase (20 U/mL) and PIKS KR1 (15 μM) and the reaction was allowed to stir at r.t for 3 h. The reaction was quenched with 1 reaction volume of 5 M sodium chloride (5 mL) and the

product was extracted with 3 × 20 mL ethyl acetate. The sample was loaded on a Biotage®SNAP cartridge KP-Sil 25 g column and run on a Isolera One by Biotage with ACI™ using a 0-10% methanol gradient with DCM as the second mobile phase. The fractions containing the final product were pooled together and rotovapped to dryness to yield the final product as a clear oil at 70% yield. **^{1}H-NMR** (600 MHz, CDCl$_3$ = 7.26 ppm)**:** δ 0.96 (t, 3H, 7.4 Hz), 1.20 (d, 3H, 7.0 Hz), 1.39-1.56 (m, 2H), 1.96 (s, 3H), 2.50 (s, 1H, O-H), 2.73 (qd, 1H, 7.1 and 3.7 Hz), 2.96-3.08 (m, 2H), 3.38-3.50 (m, 2H), 3.83 (dq, 1H, 8.4 and 4.3 Hz), 5.91 (s, 1H, N-H). **^{13}C-NMR** (151 MHz, CDCl$_3$ = 77.36 ppm)**:** δ 10.7, 11.4, 23.5 27.5, 28.9, 39.7, 53.3, 74.0, 170.6, 204.5. **ESI-MS [M+H]$^+$:** calculated for C$_{10}$H$_{19}$NO$_3$S, 234.1158 *m/z*, found 234.1157 *m/z*.

N-acetylcystamine thioester of propionic acid (Propionyl-SNAC). Propionic acid (150 mg, 2 mmol) was dissolved in 10 mL of dichloromethane on ice in a 25 mL round-bottom flask. SNAC (310 mg, 2 mmol) and DMAP (40 mg, 0.3 mmol) where added and the reaction was allowed to stir for 5 min. EDC (422 mg, 2.2 mmol) was added and the reaction was left stirring and was allowed to come to room temperature overnight. The reaction was washed with 2x20 mL of water. The water layers were washed further with 3 × 20 mL ethyl acetate. All organic washes were combined and dried with magnesium sulfate. After filtration, solvent was removed by rotary evaporation to produce a yellow oil at 10% yield. **^{1}H-NMR** (600 MHz, CDCl$_3$ = 7.26 ppm)**:** δ 1.14 (t, 3H, 7.7 Hz), 1.91 (s, 3H), 2.55 (q, 2H, 7.5 Hz), 2.98 (t, 2H, 6.6, Hz), 3.37 (q, 2H, 6.4 Hz), 6.21 (s, N-H). **^{13}C-NMR** (226 MHz, CDCl$_3$ = 77.36 ppm) δ 9.9, 23.4, 28.6, 37.7, 39.9, 170.7, 201.0. **ESI-MS [M-H]$^+$:** calculated for C$_7$H$_{13}$NO$_2$S, 176.0740 *m/z*, found 176.0739 *m/z*. (1-^{13}C)propionyl-SNAC was synthesized in the using the same protocol as above and was confirmed by ^{13}C NMR and **ESI-MS [M+H]$^+$:** calculated for C$_7$H$_{13}$NO$_2$S, 177.0774 *m/z*, found 177.0773 *m/z*.

N-acetylcystamine thioester of [3-^{13}C]-propionic acid (Propionyl-SNAC). [3-^{13}C]-sodium propionate (96 mg, 1 mmol) was dissolved in 10 mL of acetone on ice in a 25 mL round-bottom flask. SNAC (155 mg, 1 mmol) and DMAP (10 mg, 0.1 mmol) were added and the reaction was allowed to stir for 5 min. EDC (230 mg, 1.2 mmol) was added and the reaction was removed from ice and was left stirring at r.t for 24 h. The reaction mixture was partitioned with 30 mL of water and was washed with 2 × 20 mL of DCM. The combined organic washes were washed with 20 mL brine and dried with magnesium sulfate. After filtration, solvent was removed by rotary evaporation to produce a cloudy yellow oil. The oil was dissolved in ethyl acetate and was run over a 100 mL silica column with ethyl acetate as the mobile phase. The fractions containing the desired product were combined and solvent was removed by rotary evaporation. The final product was a yellow oil obtained at 15% yield. **^{1}H-NMR** (400 MHz, CDCl$_3$ = 7.26 ppm)**:** δ 1.01 and 1.33 (dt, 3H, J_{HH} = 7.5 Hz and J_{CH} = 129 Hz), 1.96 (s, 3H), 2.59 (qd, 2H, 7.5 Hz and 4.6 Hz), 3.01 (t, 2H, 6.4 Hz), 3.42 (q, 2H, 6.2 Hz), 5.92 (s, N-H). **^{13}C-NMR** (101 MHz, CDCl$_3$ = 77.36 ppm) δ 9.7, 23.2, 28.4, 37.3 and 37.6 (*Jcc* = 34), 39.7, 170.3, 200.9. **ESI-MS [M+H]$^+$:** calculated for C$_7$H$_{13}$NO$_2$S, 177.0774 *m/z*, found 177.0772 *m/z*. For absolute quantification, an ^{1}H NMR (in D$_2$O with a vanillin standard) was obtained for a 84 mM stock solution of propionyl-SNAC.

[4-^{13}C]-diethylmethylmalonate. The protocol was adapted from a previous study [*22*]. Sodium metal (Na^0, 200 mg, 8.7 mmol) was added to 15 mL of ethanol (on ice, in a 50 mL round bottom flask) over the course of 45 min. After 20 min, diethylmalonate (1.1 g, 6.67 mmol) was added and the sample was removed from the ice. The sample was put under nitrogen and ^{13}C-iodomethane (0.55 mL, 10 mmol) was added. The reaction was left stirring at r.t overnight. The sample was rotovapped to dryness and the precipitate was dissolved in 20 mL of saturated

ammonium bicarbonate. The product was extracted with 3 × 40 mL washes of diethyl ether and rotovapped. The sample was loaded on a Biotage® SNAP cartridge KP-Sil 10g column and run on a Biotage Isolera One with ACI™ using a 0-25% ethyl acetate gradient with hexanes as the second mobile phase. The fractions containing the product were screened using an iodine chamber. The fractions containing the final product were pooled together and rotovapped to dryness to yield the final product at 15% yield. **^{1}H-NMR** (400 MHz, CDCl$_3$ = 7.26 ppm): δ 1.20, 1.53 (dd, 3H, 7.4 Hz (J_{HH}) 131 Hz (J_{CH})), 1.23 (t, 6H, 7.16 Hz), 3.32-3.44 (m, 1H), 4.08-4.24 (m, 4H). **^{13}C-NMR** (101 MHz, CDCl$_3$ = 77.36 ppm) δ 13.7, 23.0, 46.23, 46.6, 61.6, 170.4.

Enzymatic synthesis of [4-^{13}C]-methylmalonyl-CoA. [4-^{13}C]-methylmalonate was generated by saponification of [4-^{13}C]-diethylmethylmalonate and used for the enzymatic preparation of [4-^{13}C]-methylmalonyl-CoA as described for fluoromalonyl-CoA. The reaction mixture (10 mL) contained 100 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (5 mM), TCEP (2.5 mM), magnesium chloride (5 mM), [4-^{13}C]-diethylmethylmalonate (7 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (36 U), myokinase (20 U), CoA (2 mM) and MatB (10 µM). The mixture was incubated at 37°C for 1 h and lyophilized overnight. The residue was dissolved in water (1.6 mL) and acidified to pH ~2 by addition of 70% (v/v) perchloric acid (160 µL). Insoluble material was removed by centrifugation at 18,000 × g for 10 min. The supernatant was adjusted to pH 6 by addition of 10 M sodium hydroxide (100 µL) and desalted on an Agilent 1200 HPLC system using a Zorbax Eclipse XDB C-18 column (5 µm, 9.4 × 250 mm, Agilent) with a linear gradient from 0 to 10% acetonitrile over 30 min with 0.1% formic acid as the aqueous mobile phase (3 mL/min). Fractions were screened by ESI-MS and those containing pure [4-^{13}C]-methylmalonyl-CoA were lyophilized overnight. **^{1}H-NMR** (900 MHz, D$_2$O = 4.79 ppm): δ 0.80 (s, 3H), 0.92 (s, 3H), 1.26 and 1.41 (d, 3H, 133 Hz (J_{CH} = 131)), 2.40 (t, 2H, 6.7 Hz), 3.01-3.08 (m, 2H), 3.34 (t, 2H, 6.2 Hz), 3.43 (td, 2H, 6.6 Hz and 2.6 Hz), 3.61 (dd, 1H, 9.8 Hz and 4.5 Hz), 3.86 (dd, 1H, 9.7 Hz and 4.5 Hz), 4.0 (s, 1H), 4.23-4.30 (m. 2H), 4.58 (dt, 1H, 2.4 and 2.6 Hz), 4.83-4.85 (m, 1H), 4.85-4.88 (m, 1H), 6.18 (d, 1H, 6.0 Hz), 8.41 (s, 1H), 8.63 (s, 1H). **^{13}C-NMR** (226 MHz, D$_2$O) δ 16.2, 21.0, 23.5, 31.1, 38.0, 41.0, 41.1, 56.2, 67.8, 74.7, 76.7, 76.8, 76.9, 76.9, 86.2, 90.2, 121.2, 145.2, 147.5, 151.2, 152.5, 176.1, 176.7, 177.5, 202.7. **ESI-MS [M+H]$^+$:** calculated for C$_{25}$H$_{40}$N$_7$O$_{19}$P$_3$S, 869.1424 m/z, found 869.1419 m/z.

2-propanone O-benzyl oxime. O-benzylhydroxylamine hydrochloride (199 mg, 1 mmol) was placed in a 10 mL scintillation vial and was dissolved in 10 mL of 50 mM sodium phosphate at pH 4.5. After 5 min, acetone (730 µL, 10 mmol) was added and the reaction was left stirring at r.t for 72 h. The reaction was washed with 3 × 20 mL DCM, dried with magnesium sulfate, filtered, and solvent was removed by rotatory evaporation resulting in a clear liquid at 90% yield. **^{1}H-NMR** (400 MHz, CDCl$_3$ = 7.26 ppm): δ 1.83 (s, 3H), 1.90 (s, 3H), 5.08, (s, 2H), 7.25-7.41 (m, 5H). **^{13}C-NMR** (101 MHz, CDCl$_3$ = 77.36 ppm) δ 16.1, 22.2, 75.6, 127.9, 128.2, 128.6, 138.6, 155.6. **ESI-MS [M+H]$^+$:** calculated for C$_{10}$H$_{13}$NO, 164.1070 m/z, found 164.1069 m/z. For absolute quantification, an ^{1}H NMR spectra (in D$_2$O with a vanillin standard) was obtained for a 604 mM stock solution of 2-propanone O-benzyl oxime in DMSO.

Triketide lactone production using Mod3$_{TE}$, KSAT3, Mod6$_{TE}$, and KSAT constructs. All assay mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (15 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 µM), CoA (1 mM), MatB (20 µM), methyl- or fluoromalonate (5 mM). The mixture was incubated at 37°C for 30-45 min and initiated by the addition of NDK-SNAC (5 mM) and DEBS protein (10 µM).

When used, DszAT (30 µM) was also added to the reaction mixture. The reaction was incubated at 37°C for 24 h (Mod3) and 16 h (Mod6). 50 µL aliquots were removed and quenched with 70% (*v/v*) perchloric acid (2.5 µL). Samples were centrifuged at 18,000 × *g* for 10 min at r.t to pellet the precipitated protein. The supernatant (50 µL) was removed and flash frozen. Excess salts were removed by centrifugation at 18,000 × *g* for 5 min at r.t. The supernatant was removed and analyzed on a Zorbax Eclipse XDB C-18 column (3.5 µm, 3 × 150 mm, 35ºC, Agilent) using a linear gradient from 0 to 40% acetonitrile over 14 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 30 s (0.8 mL/min). Products were monitored using an agilent G1315D diode array detector (TKL, $A_{260\ nm}$; F-TKL, $A_{247\ nm}$). For absolute quantification, standard curve were generated using F-TKL and TKL standards [*10*].

Initial rate of triketide lactone production using Mod3$_{TE}$ and KSAT3 constructs. All assay mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (15 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 µM), CoA (1 mM), MatB (20 µM), methyl- or fluoromalonate (5 mM). The mixture was incubated at 37ºC for 30-45 min and initiated by the addition of NDK-SNAC (5 mM) and DEBS protein (10 µM). When used, DszAT (30 µM) was also added to the reaction mixture. The reaction was incubated at 37°C. 50 µL Aliquots were removed over the course of 9 h, quenched, processed and analyzed as described above. For monitoring intial rate of product formation using malonate (5 mM) and AT^0 Mod3$_{TE}$, the same assay condtions as described above were carried out with the following changes. When DszAT (30 µM) was used, aliquots were removed and quenched over the course of 90 min. When DszAT was omitted, aliquots were removed and quenched over the course of 7 h. For quantification, the samples were analyzed on a Titan C-18 column (1.9 µm, 2.1 × 50 mm, r.t, Sigma) using a linear gradient from 0 to 40% acetonitrile over 5 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). H-TKL products were identified using an Agilent 6530 QTOF Accurate Mass Spectrometer in the negative mode. For reactions containing DszAT, H-TKL was quantified with a standard curve of an H-TKL authentic standard using an agilent G1315D diode array detector (H-TKL, $A_{256\ nm}$). For reactions in which DszAT was omitted, samples were quantified in MS negative mode using an H-TKL standard curve.

Triketide lactone production using Mod2 and KSAT2 constructs. All assay mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (20 mM), TCEP (5 mM), magnesium chloride (5 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (15 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 µM), CoA (0.1 mM or 1 mM), MatB (20 µM), and either methylmalonate or fluoromalonate (5 mM). Reactions were initiated by addition of NDK-SNAC (5 mM) and DEBS protein (10 µM) and incubated at 37°C for 16 h. Aliquots were removed, quenched, processed and analyzed as described above.

Quantification of fluorinated condensation products formed with Mod3$_{TE}$. All assay mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (15 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 µM), CoA (1 mM), MatB (20 µM), and fluoromalonate (5 mM). The mixture was incubated at 37ºC for 30-45 min and initiated by the addition of NDK-SNAC (5 mM) and DEBS AT^0 Mod3$_{TE}$ (10 µM). When used, DszAT (30 µM) was also added to the reaction mixture at this time. The reaction was incubated at 37°C for 16 h. 20 µL aliquots were removed and quenched with 70% (*v/v*) perchloric acid (1 µL). Samples

were centrifuged at 18,000 × g for 10 min at r.t to pellet the precipitated protein. The supernatant (20 μL) was removed and flash frozen. Excess salts were removed by centrifugation at 18,000 × g for 5 min at r.t. The supernatant was removed and analyzed on a Titan C-18 column (1.9 μm, 2.1 × 50 mm, r.t, Sigma) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). The triketide lactone was identified using an Agilent 6530 QTOF Accurate Mass Spectrometer in the negative mode. Products were monitored using an agilent G1315D diode array detector (F-TKL, $A_{254\,nm}$). For absolute quantification, standard curves were generated using an F-TKL standard [10]. To quantify the terminal ketone and α-fluorocarboxylic acid products formed as a result of triketide thioester hydrolysis and subsequent decarboxylation, 30 μL of the reactions were adjusted to ~pH 4.5 using 0.5 μL of 50% (v/v) formic acid solution and combined with 30 μL of O-benzylhydroxylamine hydrochloride (200 mM) in 20 mM sodium phosphate, pH 4.5 [23]. The reaction was incubated at room temperature for 48 h. Samples were centrifuged at 18,000 × g for 10 min at r.t to pellet the precipitated protein. The supernatant (55 μL) was removed and flash frozen. Excess salts were removed by centrifugation at 18,000 × g for 5 min at r.t. The supernatant was removed and analyzed on a Titan C-18 column (1.9 μm, 2.1 × 50 mm, r.t., Sigma). The fluoromethyl ketone product was analyzed using a linear gradient from 0 to 45% acetonitrile over 1 min with 5 mM ammonium formate, pH 6.5 as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). This was then followed by a 4 min isocratic phase at 45% acetonitrile (0.6 μL/min). The α-fluorocarboxylic acid product was analyzed using a linear gradient from 0 to 24% acetonitrile over 1 min with 5 mM ammonium formate, pH 6.5 as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). This was then followed by a 3.8 min gradient from 24 to 28% acetonitrile (0.6 μL/min). Products were identified using an Agilent 6530 QTOF Accurate Mass Spectrometer in the positive mode as well as monitored for quantification using an agilent G1315D diode array detector ($A_{256\,nm}$). For quantification, a standard curve was generated with 2-propanone O-benzyl oxime using an agilent G1315D diode array detector ($A_{256\,nm}$).

Triketide lactone production using propionyl-SNAC with the Mod2 + Mod3$_{TE}$ mini-PKS system. All reactions contained 400 mM sodium phosphate (pH 7.5), phosphoenolpyruvate (50 mM), TCEP (10 mM), magnesium chloride (5 mM), ATP (2.5 mM), pyruvate kinase (18 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 μM), MatB (20 μM), CoA (0.1 mM), methylmalonate (5 mM, omitted from reaction with both AT^0 constructs) and fluoromalonate (30 mM), propionyl-SNAC (5 mM) and reduced nicotinamide adenine dinucleotide phosphate (NADPH; 5 mM). Reactions were initiated by addition of Mod2 (10 μM), Mod3$_{TE}$ (10 μM), and DszAT (30 μM) and incubated overnight at 37ºC. Aliquots were removed, quenched, and processed as described above. The samples were analyzed on an EclipsePlus C-18 RRHD column (1.8 μm, 2.1 × 50 mm, r.t, Agilent) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). The triketide lactone products were identified using an Agilent 6530 QTOF Accurate Mass Spectrometer in the negative mode. Additionally, the samples were subjected to MS/MS analysis using an Agilent 6460 QQQ Mass Spectrometer. Samples were analyzed on a Poroshell 120 SB-Aq column (2.7 μm, 2.1 × 50 mm, r.t, Agilent) using the same gradient and flow rate as described above with a fragmentation voltage of 100 V and collision energies of 15 and 25 V.

Acyl-CoA hydrolysis by ACP^0 Mod3$_{TE}$. Acyl-CoA hydrolysis was determined using a fluorimetric coupled assay (detecting NADH fluorescence at 360 nm) that was adapted from an

acyltransferase hydrolysis assay described previously [24]. All reactions where composed of three solutions. Solution 1 (25 µL) contained sodium phosphate pH 7.5 (100 mM), α-ketoglutaric acid (1.6 mM), TPP (1.6 mM), NAD$^+$ (1.6 mM), and α-KGDH (1.6 U/mL). Solution 2 (25 µL) contained methylmalonyl-CoA (5-200 µM). Solution 3 (50 µL) contained sodium phosphate pH 7.5 (50 mM) and Mod3$_{TE}$ ACP^0 (400 nM). Solutions 1 and 2 were combined and the reaction was initiated with the addtion of solution 3. Combination of solutions 1 and 2 prior to intiation of the reaction allowed free-CoA in the acyl-CoA stock to be consumed.

NADH fluorescence was monitored using a Synergy Mx multimode microplate reader (Biotek) containing a High Energy Xenon Flash light source (double grating monochromators with 250-900 nm range). Assays were run in a 96-well microtiter plate (black polystyrene, flat bottom, half area, non-binding surface, Corning). Reactions were run for 10 min at 30ºC with the maximum number of measurements. Kinetic parameters (k_{cat}, K_M) were determined by fitting the data using Microcal Origin to the equation: $v_o = v_{max}$ [S] / (K_M + [S]), where v is the initial rate and [S] is the substrate concentration. Data are reported as mean ± s.e. (n = 3) with standard error derived from the nonlinear curve fitting. Error bars on graphs represent mean ± s.d. (n = 3). Error in k_{cat}/K_M is calculated by propagation of error from the individual kinetic parameters.

Determination of malonyl-ACP occupancy in Mod3$_{TE}$ with DszAT. All assay mixtures contained 50 mM sodium phosphate, pH 7.5, TCEP (5 mM), magnesium chloride (10 mM), methylmalonyl-CoA epimerase (1 µM), and either methylmalonyl-CoA (1 mM), malonyl-CoA (1 mM), or fluoromalonyl-CoA (1 mM). The reaction was initiated by the addition DEBS Mod3$_{TE}$ (1 µM). When used, DszAT (1 µM) was also added to the reaction mixture. The reaction was incubated at 37°C for 20 min, at which point 1.25 µL of trypsin (2.5 µg/µL) was added and the reaction was incubated at 37ºC for another 90 min. The reaction was quenched using liquid N_2 and was stored at -80ºC until analysis. Excess salts were removed by centrifugation at 18,000 × g for 5 min at r.t. The supernatant was removed and analyzed on an AdvanceBio Peptide Map (2.7 µm, 2.1 × 250 mm, 55ºC, Agilent) column using a linear gradient from 0 to 60% acetonitrile (90%) and 0.1% formic acid (10%) over 55 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 100% 0.1% formic acid for 5 min (0.4 mL/min). Appropriate peptides were identified using an Agilent QTOF 6530 mass spectrometer. Products were characterized using both MS1 (exact mass <2 ppm) and MS2 (phosphopantetheine ejection). The samples were analyzed in positive mode with a fragmentor voltage of 150 and collison energy of 35 V. The parent ions monitored were 683.7 m/z (*holo*-ACP, +3), 717.0 m/z (methylmalonyl-ACP, +3), 712.3 m/z (malonyl-ACP, +3), 718.3 m/z (fluoromalonyl-ACP, +3). For percent-occupancy, the MS1 extracted-ion-chromatograms for both the *acyl*-ACP and *holo*-ACP were integrated and calculated as follows:

$$\left(\frac{Acyl - ACP}{Acyl - ACP + Holo - ACP} \right) * 100\%$$

Hydrolysis of malonyl-CoA derivatives by DszAT. All assay mixtures contained 50 mM sodium phosphate, pH 7.5, TCEP (5 mM), magnesium chloride (10 mM), methylmalonyl-CoA epimerase (1 µM), and methylmalonyl-CoA (1 mM), malonyl-CoA (1 mM), or fluoromalonyl-CoA (1 mM). The reaction initiated by the addition DszAT (3 µM). The reaction was incubated at

28

37°C for 10 min and 10 µL aliquots were removed after 10 s, 1 min, 2 min, 3 min, 5 min, and 10 min. The aliquots were quenched by pipetting the 10 µL aliquot into an eppendorf tube containing 1 µL of 70% (*v/v*) perchloric acid. Samples were centrifuged at 18,000 × *g* for 10 min at r.t to pellet the precipitated protein. The supernatant (10 µL) was removed and flash frozen. Excess salts were removed by centrifugation at 18,000 x *g*. Samples containing fluoromalonyl-CoA and malonyl-CoA were analyzed on an EclipsePlusC18 RRDH column (1.8 µm, 2.1 × 50 mm, 27°C, Agilent) using a linear gradient from 0 to 5% acetonitrile over 4 min with 10 mM ammonium formate (not pH adjusted) as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). Samples containing methylmalonyl-CoA were analyzed on an EclipsePlusC18 RRDH column (1.8 µm, 2.1 × 50 mm, 27°C, Agilent) using a linear gradient from 0 to 10% acetonitrile over 8 min with 0.1% Formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). Products were monitored using an Agilent G1315D diode array detector (fluoromalonyl-CoA, methylmalonyl-CoA, malonyl-CoA, and CoASH $A_{260\ nm}$). Given the A_{260} absorbance of free-CoA and the malonyl-CoA derivatives, the CoA released could be calculated using the absorbance percentage free-CoA accounted for between the two compounds.

Monitoring [4-^{13}C]-methylmalonyl-CoA incorporation into AT^0 ACP^0 Mod3$_{TE}$ by ^{13}C-NMR. All assay mixtures contained 100 mM sodium phosphate, pH 7.5, TCEP (5 mM), magnesium chloride (10 mM), methylmalonyl-CoA epimerase (5 µM), and [4-^{13}C]-methylmalonyl-CoA (1 mM). The reaction was initiated by the addition of NDK-SNAC (5 mM) and DEBS Mod3$_{TE}$ or AT^0 ACP^0 Mod3$_{TE}$ (10 µM) to a final reaction volume of 300 µL. The reaction was incubated at 37°C for 18.5 h and the reactions were quenched with 70% (*v/v*) perchloric acid (15 µL). The internal standard, [2-^{13}C]-sodium acetate (final concentration, 2 mM in D$_2$O), was added to a final volume of 355 µL and the samples were centrifuged at 18,000 × *g* for 10 min at r.t to pellet the precipitated protein. The supernatant (350 µL) was removed and inserted into a Shigemi tube (D$_2$O) for ^{13}C-NMR. The products were quantified using an external standard solution containing [2-^{13}C]-sodium acetate (2 mM in D$_2$O), [4-^{13}C]-methylmalonyl-CoA (1 mM), [3-^{13}C]-sodium propionate (1 mM), [4-^{13}C]-methylmalonic acid (1 mM), and [3-^{13}C]-propionyl-SNAC (1 mM) in 100 mM sodium phosphate, pH 7.5 containing TCEP (5 mM) and magnesium chloride (10 mM). TKL and propionyl-CoA formed under these conditions were quantified using additional external standards of [2-^{13}C]-sodium acetate (2 mM in D$_2$O) and either TKL (1.3 mM) or propionyl-CoA (23 mM) in the same buffer. For solutions containing an unlabeled standard, the ratio between the standard and the internal standard ([2-^{13}C]-sodium acetate) was normalized to the appropriate isotopic ratio. All experimental data were collected on a Bruker Biospin 900 MHz spectrometer. For AT^0 ACP^0 Mod3$_{TE}$, the concentration of TKL was below the limit of detection by this method and was therefore determined using LCMS. Samples were separated by UHPLC on a Titan C-18 column (1.9 µm, 2.1 × 50 mm, r.t, Sigma) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% (v/v) formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 20 s (0.6 mL/min). TKL was identified using an Agilent 6530 QTOF Accurate Mass Spectrometer in negative ion mode and quantified using a calibration curve with a TKL standard.

^{19}F-NMR analysis of reactions with AT^0 Mod3$_{TE}$ and AT^0 ACP^0 Mod3$_{TE}$. All assay mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase/lactate dehydrogenase (15 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 µM), CoA (1 mM), MatB (20 µM), fluoromalonate (5 mM). The mixture was incubated at 37°C for 30-45 min and initiated by the

addition of NDK-SNAC (5 mM) and DEBS AT^0 $Mod3_{TE}$ (10 µM) or AT^0 ACP^0 $Mod3_{TE}$ (10 µM). When used, DszAT (30 µM) was also added to the reaction mixture (final volume, 500 µL). The reaction was incubated at 37°C for 14 hours and quenched with 70% (*v/v*) perchloric acid (25 µL). The internal standard TFA (8mM in D_2O) was added to a final volume of 600 µL. Samples were centrifuged at 18,000 × *g* for 10 min at r.t to pellet the precipitated protein and the sample was pipetted into an NMR tube. To quantify fluoroacetate formation, a solution containing 400 mM sodium phosphate, pH 7.5, TCEP (5 mM), magnesium chloride (10 mM), fluoroacetate (7 mM), and the TFA internal standard (8 mM in D_2O) was used. A control experiment was performed containing all of the reagents described above without DEBS and NDK-SNAC. All experimental data was collected on a Bruker AV-600 MHz spectrometer. To determine F-TKL concentrations, a portion of the reaction was removed and analyzed on a Zorbax Eclipse XDB C-18 column (3.5 µm, 3 × 150 mm, 35ºC, Agilent) using a linear gradient from 0 to 40% acetonitrile over 14 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 30 s (0.8 mL/min). Products were monitored using an agilent G1315D diode array detector (F-TKL, $A_{247\,nm}$) and compared against an F-TKL standard calibration curve.

2.3   Results and Discussion

ACP-independent chain extension with methyl- and fluoromalonyl coenzyme A (CoA) extender units. We were previously able to observe chain elongation by individual DEBS modules with a catalytically-deficient AT domain (AT^0), which we attributed to the greater number of enzyme turnovers possible when the extender unit pool is maintained using an *in situ* regeneration system for methyl- and fluoromalonyl-CoA [*10*]. However, end product analysis of single module systems does not necessarily distinguish between the two general mechanisms that could arise in this system. One possibility is that the extender unit could be loaded onto the ACP domain either by residual AT activity or by direct thiol-thioester chemical exchange to generate a canonical ACP-tethered intermediate. In this case, the polyketide product remains covalently attached to the enzyme after C-C bond formation and can be further processed effectively by the PKS (*Scheme 2.1A*). However, an alternative mechanism is that the extender unit could directly diffuse into the KS active site, bypassing both the AT and ACP domains to undergo a decarboxylative Claisen condensation to form a free polyketide product that is not covalently attached to the PKS (*Scheme 2.1B*). In the latter case, generation of full-length polyketides would be challenging once the covalent link between the PKS and the growing chain is lost since it amplifies substrate channeling from module to module by up to three orders of magnitude [*25*]. Therefore, we set out to determine which mechanism was dominant under our assay conditions for both the native and fluorinated extender units. In order to distinguish between the two proposed mechanisms, the ACP domain was inactivated by mutation of the serine of the phosphopantetheine arm modification site (S1430) to alanine (ACP^0). Since the canonical chain extension mechanism requires the phosphopantetheine arm to form the malonyl-ACP intermediate, we would expect that product formation in an AT^0 ACP^0 construct would be abolished if the malonyl-ACP intermediate is required for chain elongation (*Scheme 2.1A*) but should not be affected if the ACP can be bypassed (*Scheme 2.1B*). To generate these mutants, we started with DEBS Module 3 constructs fused to a thioesterase domain ($Mod3_{TE}$) and introduced the S1430A mutation into both a wild-type and AT^0 background where the catalytic Ser651 is mutated to alanine. A construct encoding for a truncated Mod3 protein containing only the KS and AT domains (KSAT didomain) [*17*] was also used as a control to eliminate the contribution of any remaining interactions with the ACP.
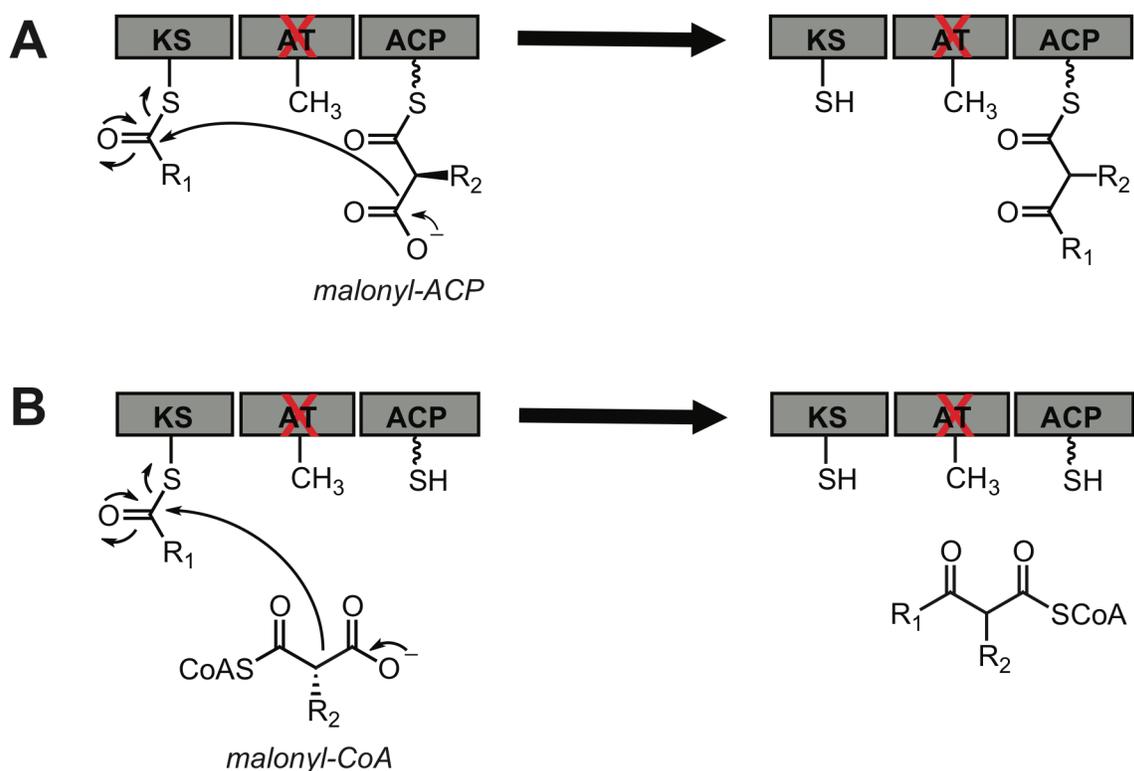
Scheme 2.1: *Proposed mechanisms of extender unit incorporation in a generalized Type I PKS module. For clarity, only the KS, AT, and ACP domains are shown. (A) Chemical transacylation of extender unit onto the ACP. (B) Direct diffusion of extender unit into the KS active site. (KS, ketosynthase; AT, acyltransferase; ACP, acyl carrier protein; R_1, growing polyketide chain; R_2 = H, Me, or F).*

These four protein constructs were heterologously expressed and purified as previously described (*Figure 2.1*) [*10*]. They were first assayed for their ability to carry out chain elongation on the natural diketide N-acetylcysteamine (NDK-SNAC) substrate surrogate (*Figure 2.2*) with their native methylmalonyl-CoA extender unit to yield the corresponding triketide lactone (TKL) product (*Figure 2.3A*). The samples were quantified under end-product and initial rate conditions using authentic standards by liquid chromatography monitored by UV-visible spectroscopy (LC-UV) and mass spectrometry (LC-MS) (*Figures 2.3AB and 2.4AB*). As consistent with previous studies [*10,26*], the production of TKL in the AT^0 module is dramatically decreased compared to production in the WT module, with a 110-fold drop in initial rate and 16- to 24-fold decrease in yield. However, assays with the ACP^0, AT^0ACP^0, and KSAT didomain yielded a similar catalytic defect as the AT^0 construct with regard to yield and initial rate. This observation implies that the ACP domain is not required for the residual chain elongation activity that is observed when the AT domain is inactivated. As an additional control, the KS^0, AT^0, and $KS^0 AT^0$ mutations were also introduced into the KSAT didomain, verifying that TKL formation was only observed in the presence of an active KS and not a result of unexpected activity from another domain in these mutants (*Figure 2.3B*). From these observations, we conclude that the observed TKL product arises from direct diffusion of the extender unit into the KS active site where C-C bond formation is catalyzed between the polyketide chain and methylmalonyl-CoA rather than methylmalonyl-ACP (*Scheme 1B*). As different modules may display diverse behavior, similar experiments were carried out on WT, AT^0, and KSAT didomain constructs of Mod2 and $Mod6_{TE}$ (*Figure 2.3*) to test

whether the results with Mod3$_{TE}$ could be generalized. Indeed, we observed that ACP-independent chain extension can also occur with Mod2 and Mod6$_{TE}$ as well, based on the comparison of activities of the WT, AT^0, and KSAT didomain constructs (*Figure 2.3C*). The relatively high yield of TKL in the AT^0 Mod2 construct compared to Mod3$_{TE}$ and Mod6$_{TE}$ may further indicate that the KS of Mod2 is more promiscuous with regard to non-native substrates. Experiments with Mod2 constructs also yielded additional information with regard to ACP-independent chain extension. We found that product formation with WT Mod2 occurred at an order of magnitude lower efficiency than WT Mod3$_{TE}$ and Mod6$_{TE}$, likely due to the absence of a thioesterase fusion. TKL is still formed by this construct because of the stability of the six-membered lactone ring but relies on uncatalyzed release from the ACP domain. However, the longer residence time of the polyketide on the ACP may create a bottleneck for formation of methylmalonyl-ACP. Under these conditions, the production of TKL by WT, AT^0, and KSAT Mod2 constructs is essentially identical (*Figure 2.3C*), suggesting that the ACP-independent mode of chain extension may be able to compete with canonical chain extension if the KS domain is saturated with methylmalonyl-CoA. The CoA pool was then reduced ten-fold to 0.1 mM to reduce saturation of the KS domain, resulting in reduction of the AT^0 and KSAT-didomain yields by up to 5.2-fold while WT only saw a two-fold loss in TKL formation (*Figure 2.3C*). These observations imply that saturation of the ACP-independent chain extension requires higher extender unit concentrations and that bias toward direct KS usage could occur in the WT module at high methylmalonyl-CoA concentration.

With a basic understanding of extender unit incorporation in hand, we set out to compare behavior with fluoromalonyl-CoA. Based on the presence of the α-fluorine substituent, which
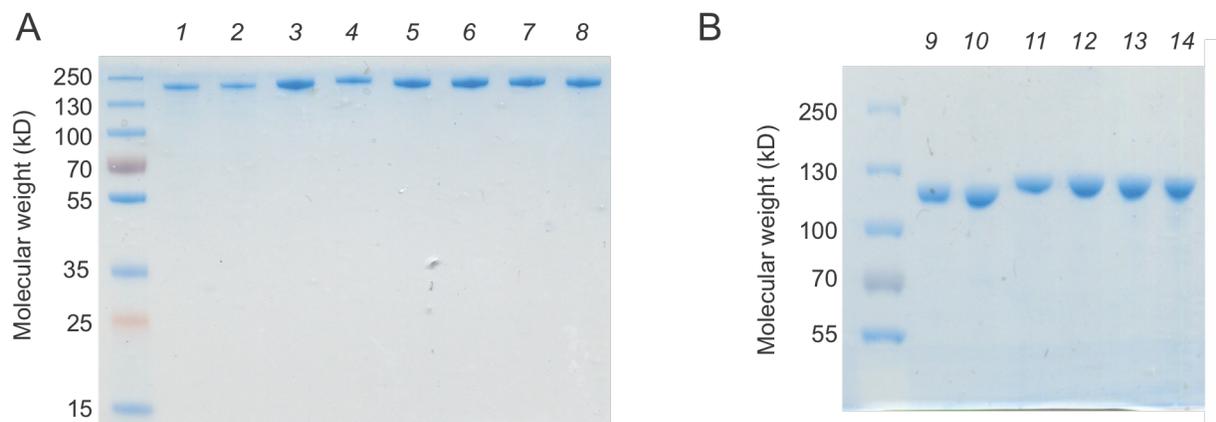
Figure 2.1: *SDS-PAGE gels of purified proteins DEBS polyketide protein. (A) Modules (**1**, Mod2; **2**, AT^0 Mod2; **3**, Mod3$_{TE}$; **4**, AT^0 Mod3$_{TE}$; **5**, AT^0 ACP^0 Mod3$_{TE}$; **6**, ACP^0 Mod3$_{TE}$; **7**, Mod6$_{TE}$; **8**, AT^0 Mod6$_{TE}$). (B) KSAT didomains (**9**, KSAT2; **10**, KSAT6; **11**, KSAT3; **12**, KSAT^{0}3; **13**, KS^{0}AT3; **14**, KS^{0}AT^0).*
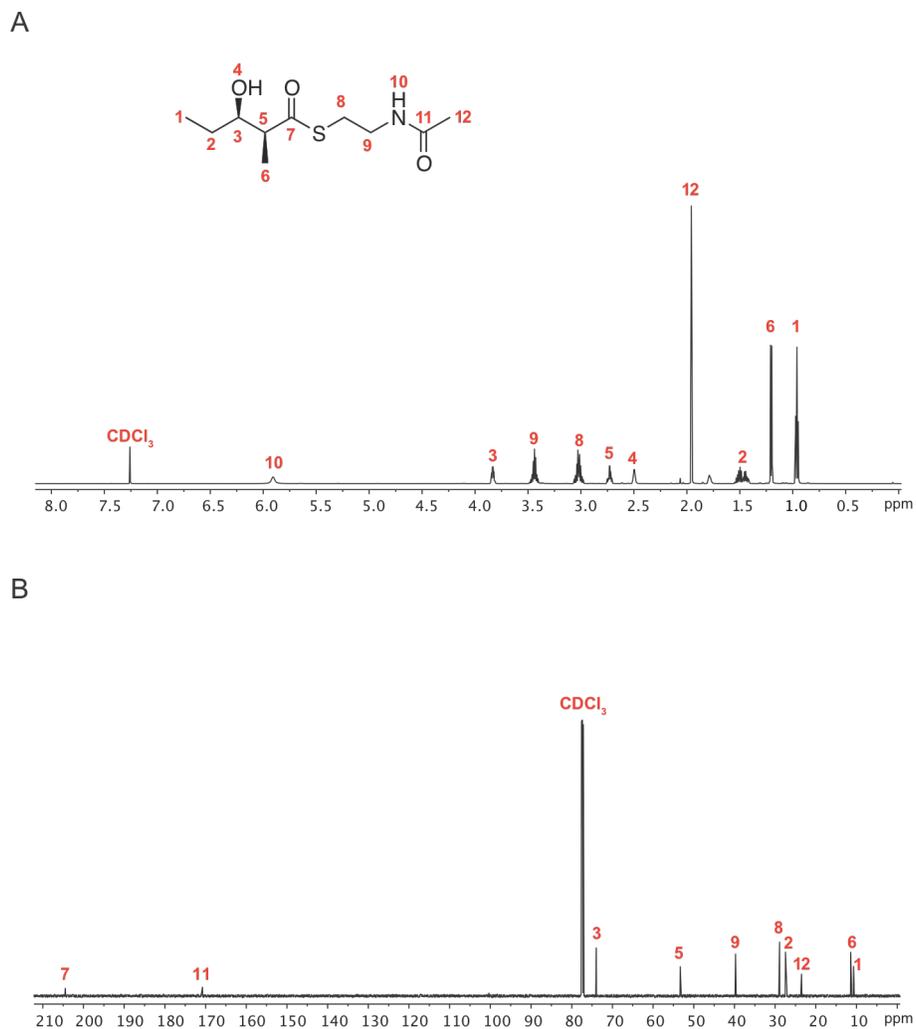
Figure 2.2: NMR spectra of NDK-SNAC. (A) ^1H NMR. (B) ^{13}C NMR.

activates this analog towards nucleophilic attack at the carbonyl, it is possible that fluoromalonyl-CoA would be more prone to undergo chemical transacylation to form the ACP-tethered intermediate (*Scheme 2.1A*). However, both the initial rate and end product analysis of fluorinated triketide lactones (F-TKLs) provide evidence that chain extensions carried out under these conditions occur independently of the ACP based on the similar activities measured for all five constructs (*Figures 2.3B and 2.4C*). As such, endpoint F-TKL formation experiments with Mod2 and Mod6$_{TE}$ constructs are consistent with this conclusion that product is mostly formed in these AT^0 variants without loading of the extender unit onto the ACP domain (*Figure 2.3C*). These initial rate and endpoint experiments provide evidence that malonyl-ACP intermediates are not required for C-C bond formation. However, chain extension with the CoA-linked extender unit leads to early chain termination since the covalent tether between the PKS and the growing polyketide chain is lost.

A

NDK-SNAC + (methylmalonyl/fluoromalonyl-SCoA) → Triketide lactone (TKL)

KS — AT — KR* — ACP — TE
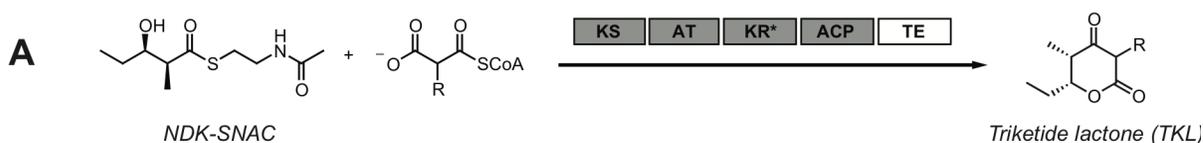
B

Construct	Methylmalonyl-CoA (R = Me)		Fluoromalonyl-CoA (R = F)	
	TKL (µM)	Initial rate (µM/h)	F-TKL (µM)	Initial rate (µM/h)
KS AT ACP	2540 ± 40	600 ± 20	9 ± 3	0.535 ± 0.007
KS ⨯AT ACP	109 ± 5	5.4 ± 0.3	33 ± 4	1.3 ± 0.2
KS AT ⨯ACP	119 ± 12	6.0 ± 0.7	9.9 ± 0.8	0.62 ± 0.06
KS ⨯AT ⨯ACP	136 ± 10	6.2 ± 0.3	38 ± 6	1.33 ± 0.16
KS AT	151 ± 18	6.4 ± 1.4	15 ± 1	0.96 ± 0.03
KS ⨯AT	154 ± 2	–	65 ± 6	–
⨯KS AT	n.d.	–	n.d.	–
⨯KS ⨯AT	n.d.	–	n.d.	–

C

Construct	Mod6$_{TE}$		Mod2			
	TKL (µM)	F-TKL (µM)	TKL (µM)		F-TKL (µM)	
	1.0	1.0	1.0	0.1	1.0	mM CoA
KS AT ACP	2650 ± 90	17.7 ± 0.6	250 ± 3	123 ± 3	31 ± 2	
KS ⨯AT ACP	22 ± 2	19 ± 1	215 ± 2	41.8 ± 0.4	45.0 ± 0.5	
KS AT	112 ± 3	37 ± 1	176 ± 18	56 ± 2	35 ± 3	

Figure 2.3: Analysis of TKL and F-TKL formation with constructs of DEBS Mod2, Mod3$_{TE}$, and Mod6$_{TE}$. (A) Reaction scheme for the TKL formation assay with Mod3$_{TE}$ (KR*, natively inactive ketoreductase domain). (B) Endpoint and initial rate quantification of TKL and F-TKL formation with 10 µM of DEBS Mod3$_{TE}$ constructs and KSAT3. (C) Endpoint quantification of TKL and F-TKL formation with 10 µM of DEBS Mod2 and Mod6$_{TE}$ constructs and KSAT6. Only KS, AT, and ACP domains are shown in the labels for Figure 2.2BC. All data are mean ± s.e (n=3) (n.d., not detected; –, data not collected)
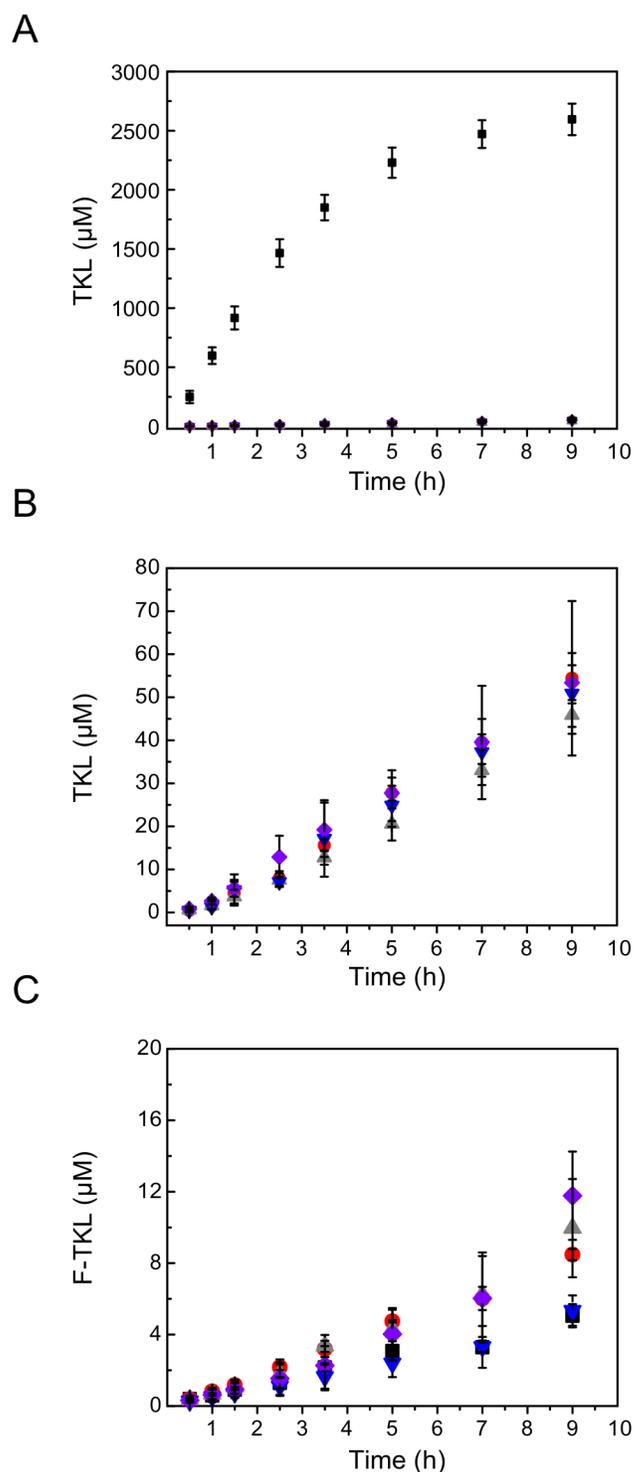
Figure 2.4: *Initial rate of TKL and F-TKL formation in DEBS module 3 protein (A) Initial rate of TKL formation with WT Mod3_{TE} (black), AT⁰ Mod3_{TE} (grey), ACP⁰ Mod3_{TE} (blue), AT⁰ ACP⁰ Mod3_{TE} (purple), and KSAT3 didomain (red). (B) Close-up of the initial rate of TKL formation for the Mod3_{TE} mutants and truncation products (AT⁰ Mod3_{TE} (grey), ACP⁰ Mod3_{TE} (blue), AT⁰ ACP⁰ Mod3_{TE} (purple), and KSAT3 didomain (red)). The rate of product formation is displayed in the manuscript text in Figure 1B. The error bars represent s.d. (n=3).*

Consequences of chain extension with CoA- versus ACP-linked extender units. We were interested in further examining the nature of ACP-independent chain extension and its behavior compared to canonical ACP-dependent chain extension so that we could better understand the issues that may arise when this mode of reactivity occurs. For example, we reasoned that the coupling between extender unit decarboxylation and C-C bond formation may be affected, leading to release of propoionate or fluoroacetate, which cannot be regenerated under our current assay conditions (*Scheme 1*). To address this question, we set out to characterize the partitioning between productive extender unit decarboxylation leading to TKL formation compared to nonproductive decarboxylation leading to propionate production (*Figure 2.5*). To detect the products of the reactions with methylmalonyl-CoA, [4-^{13}C]-methylmalonyl-CoA was chemoenzymatically synthesized from [4-^{13}C]-methylmalonic acid [22] (*Figure 2.6*) and used as a substrate with the WT and AT^{0}ACP^0 Mod3$_{TE}$ constructs. Monitoring the reaction of WT Mod3$_{TE}$ by quantitative ^{13}C-NMR spectroscopy allowed us to determine that for every ten extender unit turnovers, six are hydrolyses, three are C-C bond formations, and one is a nonproductive decarboxylation (*Figure 2.7*). These results indicate that hydrolysis of extender units by the AT domain and nonproductive decarboxylation by the KS domain are both competitive with product formation under *in vitro* conditions. Along these lines, the measured rate constants for methylmalonyl-CoA hydrolysis (1.0 min^{-1}, *Figure 2.8*) and triketide formation (1.0 min^{-1}, *Figure 2.3B*) are similar. In this particular case, the observation that extender unit hydrolysis is the predominant reaction observed could arise from the use of a non-polyketide starter substrate for this module [27]. In WT Mod3$_{TE}$, where chain extension presumably takes place with
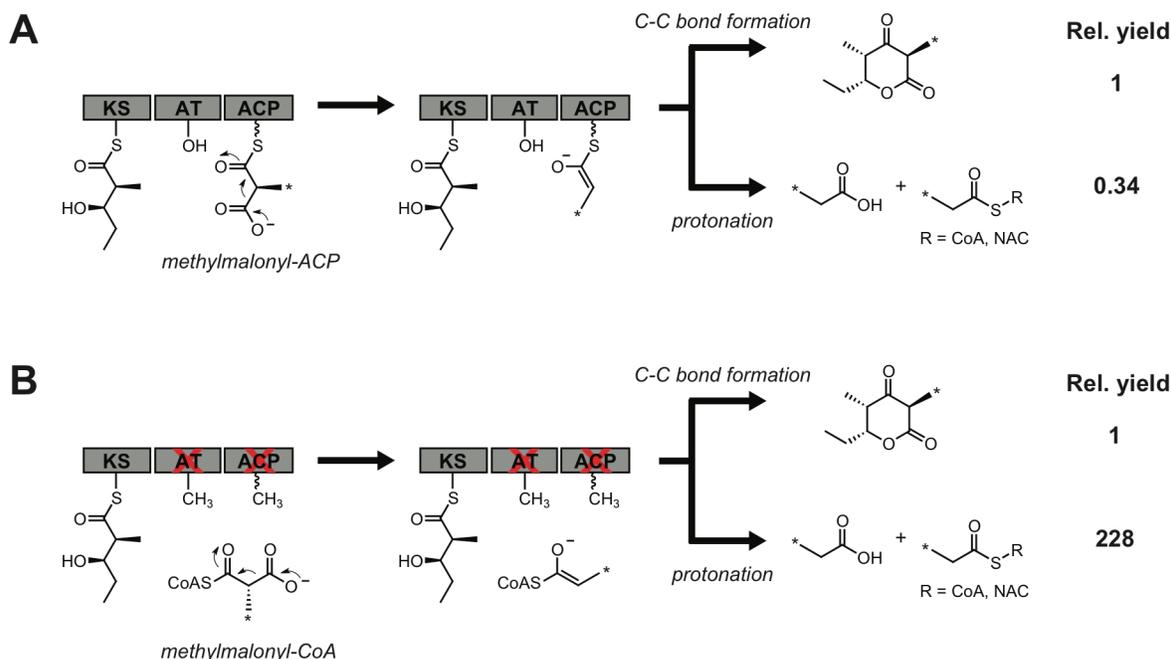
Figure 2.5 ^{13}C-NMR quantitation of products formed in the TKL assay with WT and AT^0 ACP^0 Mod3$_{TE}$ using the native methylmalonyl-CoA extender unit. For clarity, only the KS, AT, and ACP domains are shown with the KR* and TE domains omitted. (A) With WT Mod3$_{TE}$, extender unit decarboxylation leads to formation of TKL as the major C-C bond-coupled product rather than protonation to form propionate products. (B) With AT^0 ACP^0 Mod3$_{TE}$ AT^0 ACP^0, nonproductive decarboxylation becomes the major pathway. (*, site of ^{13}C label).

methylmalonyl-ACP, the ratio of productive to nonproductive decarboxylation was observed to be 3:1 (*Figures 2.5 and 2.7*). However, this ratio drastically shifted to 1:228 in $AT^0 ACP^0 Mod3_{TE}$, where we have proposed that TKL formation occurs using methylmalonyl-CoA directly, suggesting that the ratio of productive decarboxylation is increased ~685-fold when the methylmalonyl-ACP intermediate is formed (*Figures 2.5 and 2.7*). Thus, extender units that diffuse directly into the active site of the KS are lost to nonproductive decarboxylation but the fidelity of the growing polyketide chain is maintained. In contrast, analysis of reactions of $AT^0 ACP^0 Mod3_{TE}$ and fluoromalonyl-CoA by ^{19}F-NMR using the *in situ* extender unit regeneration system unexpectedly revealed a significantly higher ratio (2:1) for C-C bond formation compared to nonproductive decarboxylation to form fluoroacetate (*Figure 2.9*). The observed difference in behavior between methylmalonyl-CoA and fluoromalonyl-CoA in this experiment could be related to the stabilizing effects of the α-fluoro group on the corresponding enolate or a difference in how native and non-native extender units are treated by the enzyme. Given the inefficiency of ACP independent C-C bond formation and processivity with fluoromalonyl-CoA in DEBS modules bearing the AT^0 mutation, we turned our attention to examining the changes in chain extension mechanism that occur when $AT^0 Mod3_{TE}$ is complemented in *trans* with standalone ATs.

Efficient complementation of DszAT with fluorinated extender units. The formation of C-C bonds via ACP-independent reaction of the extender unit inhibits processive growth of the polyketide chain on the PKS to form the full-length product [25]. Previous work with the *trans*-AT of the disorazole biosynthetic cluster (DszAT) has shown that it can acylate DEBS ACPs with malonyl-CoA [11,15,28] and also amplifies chain extension with fluoromalonyl-CoA to produce F-TKL [10]. Given our observation that direct reaction of fluoromalonyl-CoA with the KS domain can lead to chain extension, we decided to further probe the behavior of the *trans*-AT in this system. We found that complementation of $AT^0 Mod3_{TE}$ with DszAT results in a 10-fold increase in F-TKL formation, as well as a 40-fold improvement in the initial rate of fluorinated extender unit incorporation (*Figures 2.10 and 2.11*). This enhancement is dependent on the presence of a functional ACP domain, as no yield or rate enhancements were observed in the control when $AT^0 ACP^0 Mod3_{TE}$ was complemented with DszAT (*Figures 2.10 and 2.11*). Moreover, no fluoroacetate was observed by ^{19}F-NMR when $Mod3_{TE} AT^0$ was complemented with DszAT, implying that nonproductive decarboxylation does not occur when the extender unit is properly attached to the ACP (*Figure 2.12*). These results are consistent with a mechanism of DszAT complementation by transacylation of the ACP with fluoromalonyl-CoA to form a covalently-tethered intermediate. As such, the restoration of the canonical PKS mechanism through this complementation can provide a successful solution to the challenge of chain transfer between upstream and downstream PKS modules.
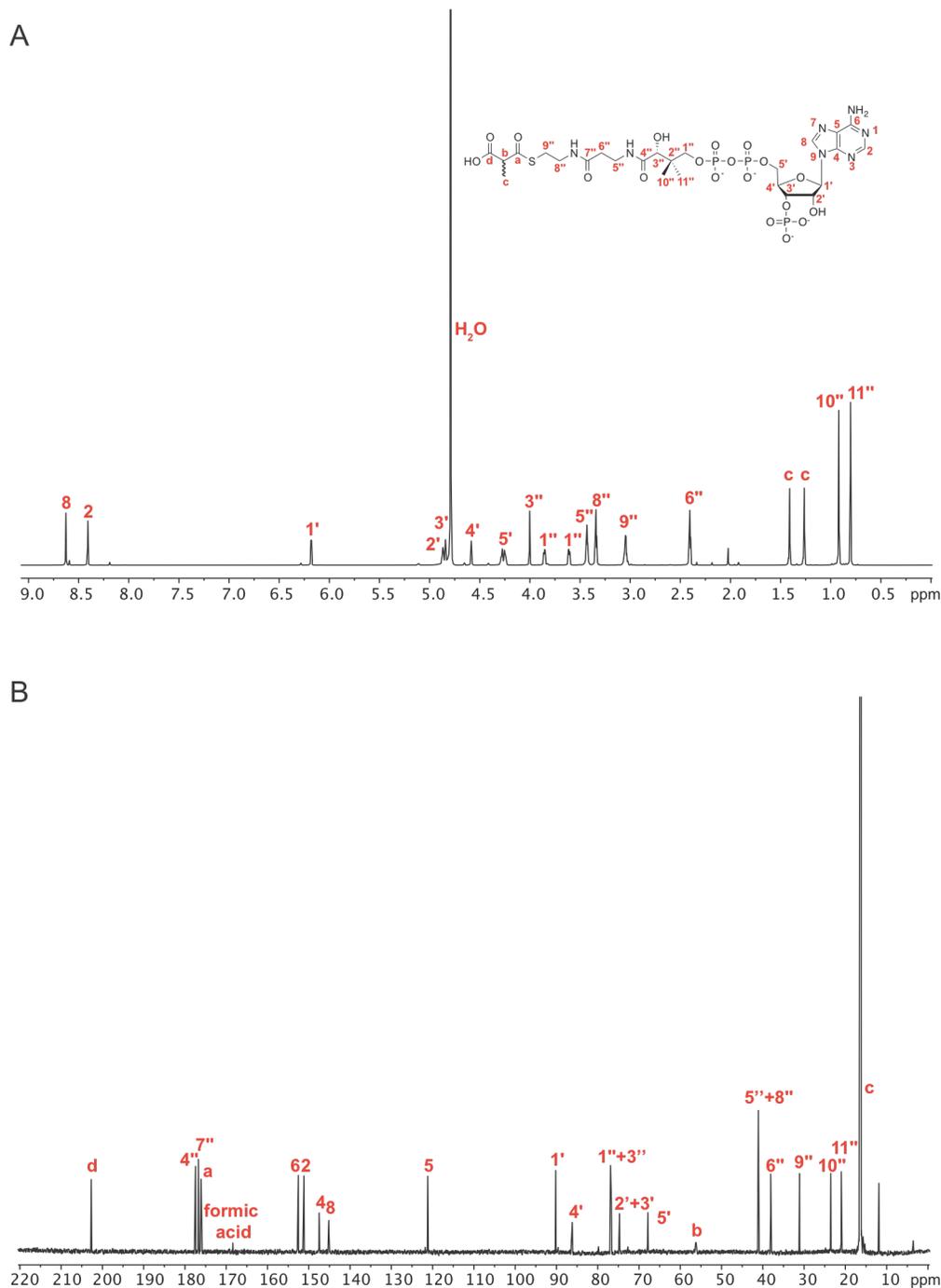
Figure 2.6. NMR spectra of [4-^{13}C]-methylmalonyl-CoA. (A) ^{1}H-NMR. (B) ^{13}C-NMR. The peak observed at 11.9 ppm on the ^{13}C-spectrum represents the ^{13}C-labeled carbon of propionyl-CoA, a known decomposition product of methylmalonyl-CoA. The integration ratios between this peak and the ^{13}C-labeled peak of the product is 0.005:1. Other quantitative NMR experiments described in this manuscript found that the integration of methylmalonyl-CoA is ~0.5 times the integration of an equivalent concentration of propionyl-CoA, suggesting that the product is of >99.5% purity.
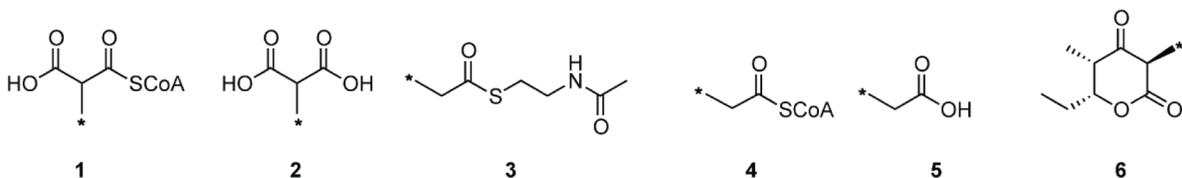
	Module3$_{TE}$ WT	Module3$_{TE}$ AT^0 ACP^0
Compound	Average ± s.d. (μM)	Average ± s.d. (μM)
1	230 ± 42	536 ± 47
2	290 ± 14	-16 ± 13
3	22 ± 6	83 ± 7
4	20 ± 6	69 ± 8
5	2 ± 5	76 ± 10
6	130 ± 3.4	0.86 ± 0.5
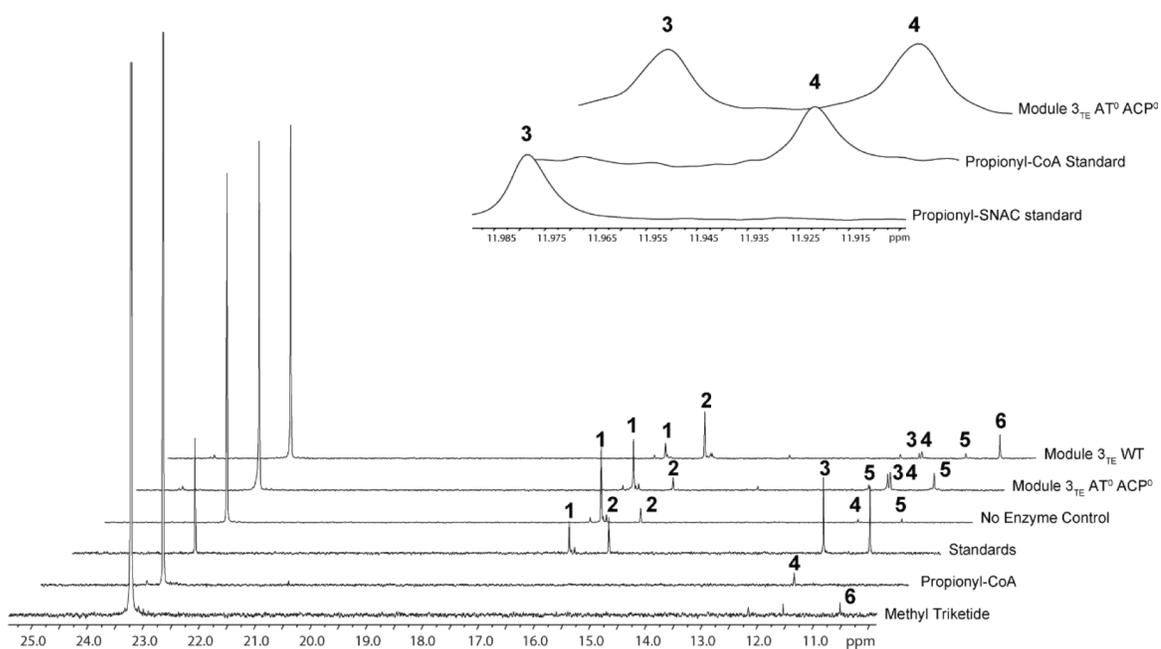
Figure 2.7: *^{13}C-NMR traces aligning the different reactions of the [4-^{13}C]-methylmalonyl-CoA stoichiometry experiment (845 μM of ^{13}C-labeled compounds). The peak at 23.24 ppm represents the [2-^{13}C]-sodium acetate standard. The concentration of the various products and substrates are tabulated in the chart below. The values in the chart have been normalized using the "No enzyme control" experiment. Mass balances for all experiments were observed to fall within 109 ± 5% of the expected value. As observed, the propionyl-SNAC and propionyl-CoA peaks can be resolved and integrated. Data are reported in mean ± s.d. (n=3).*
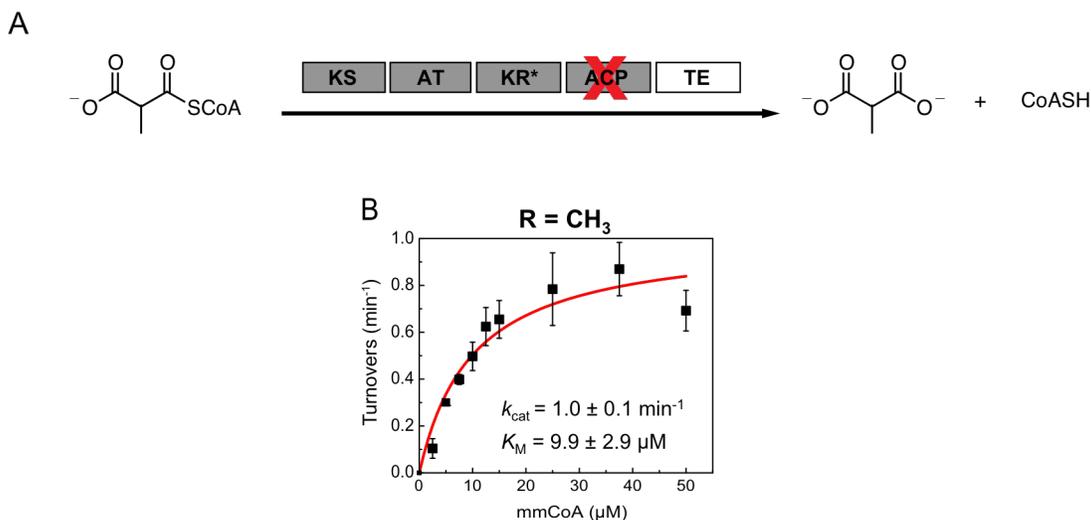
Figure 2.8: *Steady state kinetic analysis of ACP[0] Mod3$_{TE}$ hydrolysis of methylmalonyl-CoA. CoA release is quantified using NADH fluorescence as described on p. S15. (A) Scheme representing hydrolysis of methylmalonyl-CoA by ACP[0] Mod3$_{TE}$. (B) Michaelis-Menten curve for methylmalonyl-CoA. The error bars on the data points represent s.d. (n=3) and the kinetic parameters are reported as the mean ± s.e. determined from non-linear curve fitting.*

In order to confirm that fluoromalonyl-CoA could be effectively transacylated onto the ACP domain, we adapted a MS/MS spectrometry-based assay that allows us to directly detect malonyl-ACP intermediates and calculate a semi-quantitative occupancy of ACPs with different extender units (*Figures 2.13 and 2.14*) [*29-31*]. The different DEBS modules were first incubated with the appropriate extender unit and then subjected to tryptic digestion of the protein. To simplify analysis, the *in situ* regeneration system for the extender unit was not included. The parent ion of the peptide containing the phosphopantetheine arm of the ACP was then selected and further fragmented to eject the phosphopantetheine arm, confirming its identity and allowing us to measure the ratio of acyl- to *holo*-ACP (*Figure 2.13B*). When WT Mod3$_{TE}$ is incubated with methylmalonyl-CoA, the ACP is almost fully acylated (92%) as is expected (*Figure 2.13C*). However, no detectable acyl-ACP intermediates can be observed under the same conditions with AT[0] Mod3$_{TE}$. Similarly, basal levels of acylation by non-native extender units (malonyl-CoA and fluoromalonyl-CoA) can be observed in both the WT and AT[0] constructs (*Figure 2.13C*). We attribute this low level of acylation to an uncatalyzed thioester exchange between the CoA thioesters and the ACP thiol [*32*] because it is similar to the occupancy of acyl-ACPs for modules that are subjected to tryptic digestion before incubation with the various extender units, where acylation must occur chemically (*Figure 2.13C*). The similarity of the malonyl- and fluoromalonyl-ACP occupancies under all three conditions further implies that the AT domain does not acylate unnatural extender units onto the ACP, or does so at levels lower than uncatalyzed or chemical transacylation. As such, these data are consistent with our hypothesis that formation of products in modules with inactive AT and/or ACP domains is the result of KS-catalyzed C-C bond formation between the growing polyketide chain and a CoA-linked extender unit.
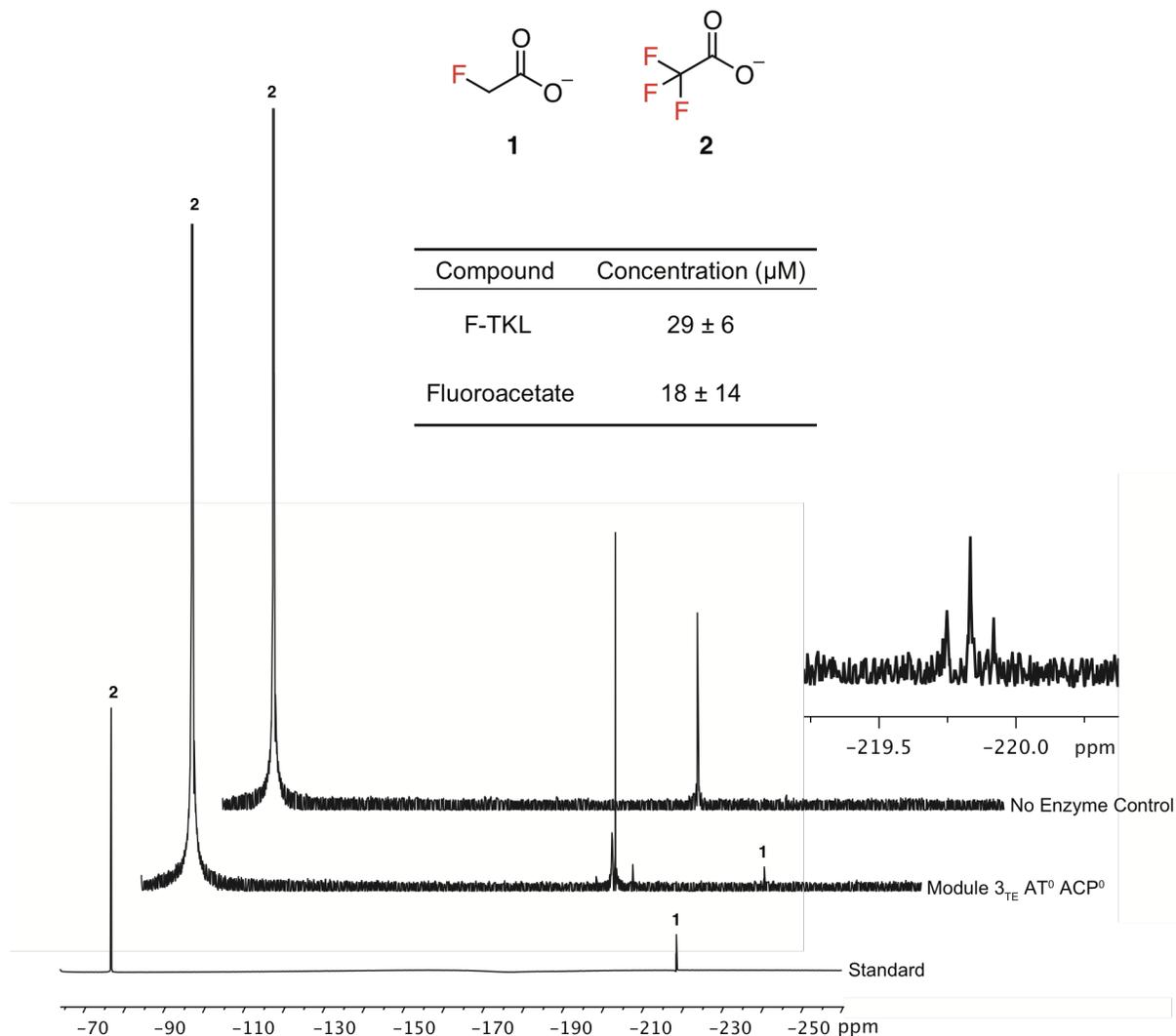
Figure 2.9: ^{19}F-NMR analysis of AT^0 ACP^0 $Mod3_{TE}$ F-TKL formation assay. AT^0 ACP^0 $Mod3_{TE}$ (10 μM), NDK-SNAC (5 mM), fluoromalonate (5 mM), and CoA (1 mM) incubated under in situ regeneration conditions and analyzed by ^{19}F NMR after 14 h. The control reactions omitted NDK-SNAC and AT^0 ACP^0 $Mod3_{TE}$. Fluoroacetate (**1**) was quantified using a trifluoroacetic acid (**2**) internal standard. F-TKL was quantified using LC-UV (A_{247}) using a standard curve with an authentic sample. The J_{HF} of the fluoroacetate standard is 48.3 Hz and was found to be 48.4 Hz in the fluoroacetate generated in the reaction. Additionally, no fluoroacetate is observed in the control reaction, suggesting that DEBS is responsible for the decarboxylation product.

A

B

Construct		Initial rate (mM/h)
KS AT ACP + AT		44 ± 2
KS AT ACP + AT		0.5 ± 0.2
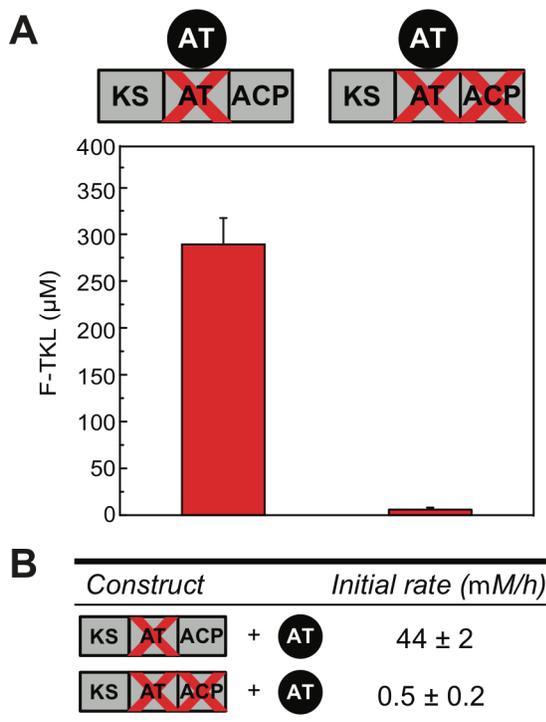
Figure 2.10: *Complementation of AT^0 Mod3$_{TE}$ constructs with DszAT. Reactions contain 10 μM DEBS Mod3$_{TE}$ and 5 mM NDK-SNAC 30 μM DszAT. For clarity, only the KS, AT, and ACP domains are shown. (A) Endpoint quantification of F-TKL formation. Data are mean ± s.d. (n=3). (B) Initial rate of F-TKL formation. Data are mean ± s.e. (n=3).*

When examining the reaction of AT^0 Mod3$_{TE}$ in the presence of DszAT, we found that 89% of the ACPs were acylated after a 20-minute incubation with malonyl-CoA, which is the native substrate of DszAT (*Figure 2.13C*). While 10% of ACPs were acylated when incubated with methylmalonyl-CoA, DszAT active sites have been previously shown to contain methylmalonyl-CoA acyl-enzyme intermediates [*33*] but the degree of acylation is sufficiently low that it does not appear to contribute to TKL production. Interestingly, only 28% occupancy of fluoromalonyl-ACP was observed (*Figure 2.13C*) despite the robust yield increase in F-TKL formation observed in the presence of DszAT (*Figure 2.10A*). However, the relatively low occupancy appears to be derived from rapid hydrolysis of fluoromalonyl-CoA by DszAT (*Figure 2.15*), which would not be yield-limiting under our *in situ* extender unit regeneration conditions where fluoromalonyl-CoA can be regenerated at a rate of 1.5 s^{-1} (unpublished data).

Since the direct observation of DszAT-catalyzed ACP acylation provides compelling evidence that F-TKL was being produced through the fluoromalonyl-ACP intermediate, we sought to assess the overall efficiency of this process using both the initial rate and yield of chain extension as markers. We thus characterized DszAT complementation of AT^0 Mod3$_{TE}$ with fluoromalonyl-CoA and using its native substrate, malonyl-CoA. With malonyl-CoA as the extender unit, the initial rate of desmethyl triketide lactone (H-TKL) observed was 10.8 ± 0.3 uM min^{-1} (*Table 2.1, Figure 2.16*). This value is within error of the initial rate of TKL formation measured with WT Mod3$_{TE}$ (10.0 ± 0.3 uM min^{-1}) with its native methylmalonyl-CoA extender unit (*Table 2.1*), which indicates that DszAT performs well in complementing the AT^0 mutation in *trans*. In comparison, the initial rate of F-TKL formation for AT^0 Mod3$_{TE}$ complemented with DszAT using fluoromalonyl-CoA had been determined to be 0.73 ± 0.03 uM min^{-1} (*Table 2.1*), which is 7.3% of the rate of the WT Mod3$_{TE}$ system. In this case, the upper limit for the initial rate of F-TKL production could potentially be set by the rate of fluoromalonyl-CoA regeneration (1.5 s^{-1}).

Figure 2.11: *Initial rate of F-TKL formation with AT^0 Mod3$_{TE}$ (black) and AT^0 ACP^0 Mod3$_{TE}$ (red) complemented with DszAT. Data are mean ± s.d. (n=3).*
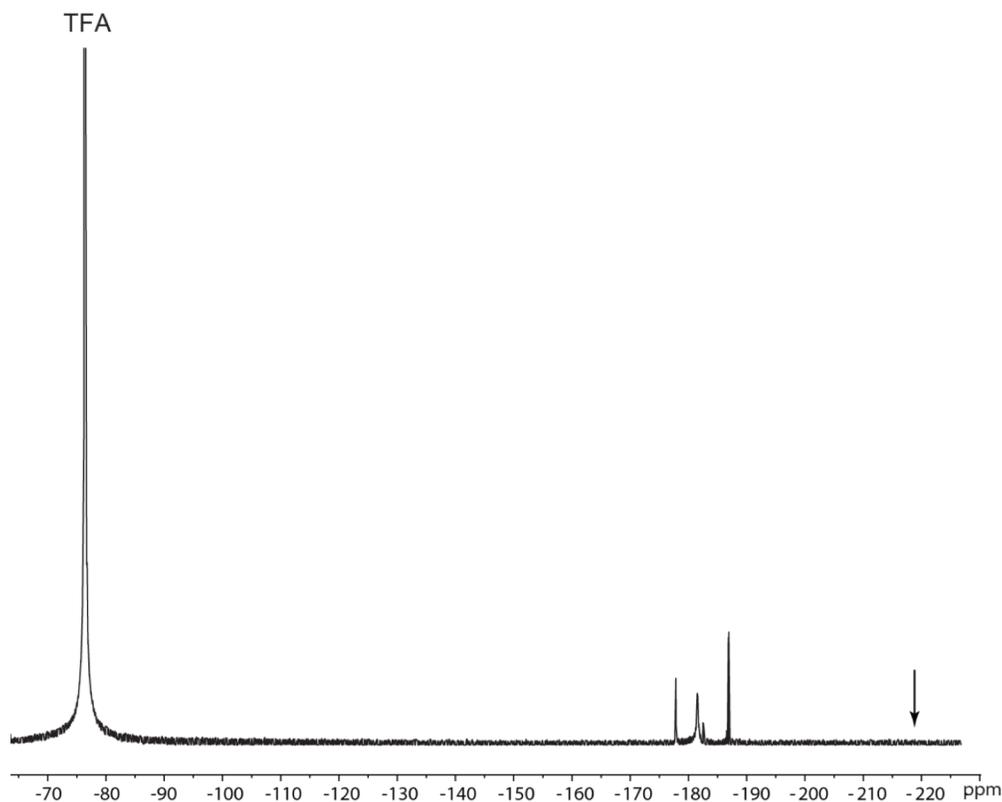
Figure 2.12: ^{19}F-NMR of F-TKL with AT^0 Mod3$_{TE}$ complemented with DszAT. ^{19}F-NMR spectrum of a reaction containing AT^0 Mod3$_{TE}$ (10 μM), DszAT (30 μM), NDK-SNAC (5 mM), fluoromalonate (5 mM), and CoA (1 mM) incubated under in situ regeneration conditions for 14 h. No fluoroacetate is observed above the NMR detection limit in this reaction (fluoroacetate appears as a triplet (-220 ppm, J_{HF} = 48.3 Hz), see arrow). This result indicates that complementation of AT^0 Mod3$_{TE}$ with DszAT also restores coupling between fluoromalonyl decarboxylation and C-C bond formation.

Substrate	Mod3$_{TE}$ AT^0 (% acyl ACP)	Mod3$_{TE}$ WT (% acyl ACP)	Mod3$_{TE}$ AT^0 + DszAT (% acyl ACP)	Predigested Mod3$_{TE}$ AT^0 (% acyl ACP)
Malonyl-CoA	3.9 ± 0.2	4.5 ± 0.1	89 ± 1	5.6 ± 0.4
Methylmalonyl-CoA	n.d.	92 ± 1	10 ± 1	0.4 ± 0.2
Fluoromalonyl-CoA	5.7 ± 3	7.9 ± 0.2	28 ± 2	10 ±1
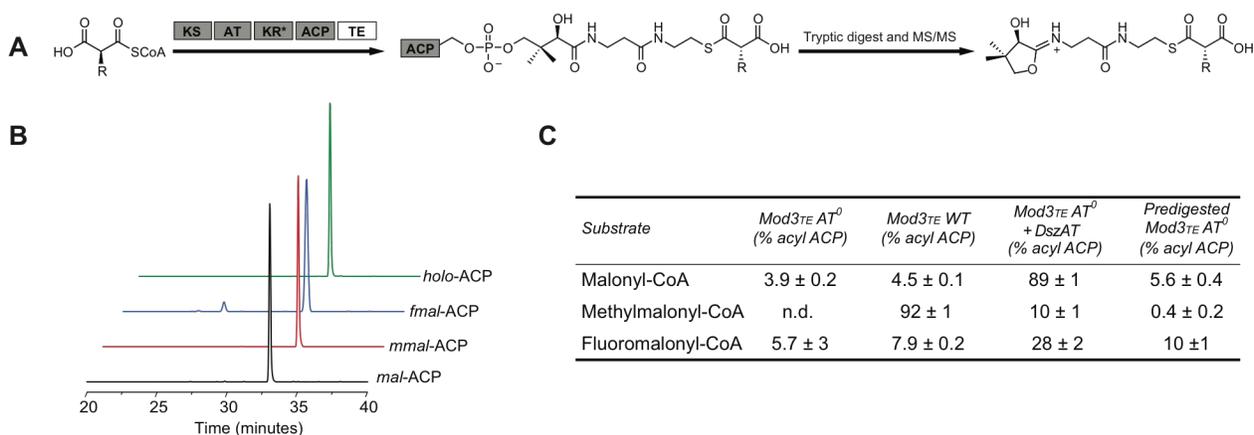
Figure 2.13: Semi-quantitative analysis of ACP acylation of Mod3$_{TE}$ constructs with malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA using LC-MS/MS. (KR*, natively inactive ketoreductase domain) (A) Scheme for phosphopantetheine arm ejection from the ACP. (B) Extracted ion chromatograms (EIC) of the malonyl-ACP fragments and holo-ACP. (C) Semi-quantitative ACP acylation occupancy calculated based on the MS peak area of acylated and unacylated phosphopantetheine arm. Data are mean ± s.d. (n=3).

43

Figure 2.14: *MS/MS characterization of phosphopantetheine ejection fagments. (A) MS/MS characterization of phosphopantetheine ejection fragments of holo-ACP, malonyl-ACP, methylmalonyl-ACP, and fluoromalonyl-ACP. The MS/MS spectrum of (B) malonyl-ACP, (C) methylmalonyl-ACP, (D) fluoromalonyl-ACP, (E) holo-ACP can be observed as well.*

We also characterized the product yield, which could be more relevant to engineering the production of fluorinated polyketides. Since the single modular systems have been found to produce other compounds in addition to the expected TKL, we quantified other major products resulting from C-C bond formation in order to better assess the stoichiometry of chain extension reactions. With fluoromalonyl-CoA, these products arise after the expected chain extension reaction when the TE releases the free acid rather than forming the lactone and still represent a positive C-C bond forming event (*Figure 2.17*). When the data are represented in terms of chain extension stoichiometry, we find that AT^0 $Mod3_{TE}$ can incorporate fluorinated extender units at 43% efficiency when complemented with

Figure 2.15: *Initial rates of extender unit hydrolysis by DszAT (3 µM). Data are mean ± s.d. (n=3).*

DszAT, compared to WT $Mod3_{TE}$ incubated with its native methylmalonyl-CoA extender unit (*Figures 2.17 and 2.18*). The endpoint yield relative to the native system is higher than expected based on the initial rate data, even taking into account other products that are formed. However, it is potentially consistent with the higher level of observed fluoromalonyl-ACP (*Figure 2.13*). This difference could result from the initial rate being limited by fluoromalonyl-CoA regeneration or by a slower rate of KS-catalyzed C-C bond formation with the α-fluoro substituent that is separate from the effectiveness of the ACP acylation reaction carried out by DszAT. Taken together, this data demonstrates that DszAT is sufficiently active for catalyzing ACP acylation with fluoromalonyl-CoA to produce a significant yield of fluorinated polyketide product compared to the wild-type system.

Table 2.1: *Comparison of initial rate of TKL formation with different extender units. Data for methylmalonyl-CoA and fluoromalonyl-CoA were taken from Figure 1B and 3B respectively and shown here for direct comparison. For clarity, only the KS, AT, and ACP domains are shown in the PKS schematic. Data are mean ± s.e. (n=3).*

Construct	DszAT	Extender unit	Initial rate (µM/min)
KS AT ACP	—	Methylmalonyl-CoA	10.0 ± 0.3
KS ~~AT~~ ACP	+	Malonyl-CoA	10.8 ± 0.3
KS ~~AT~~ ACP	+	Fluoromalonyl-CoA	0.73 ± 0.03
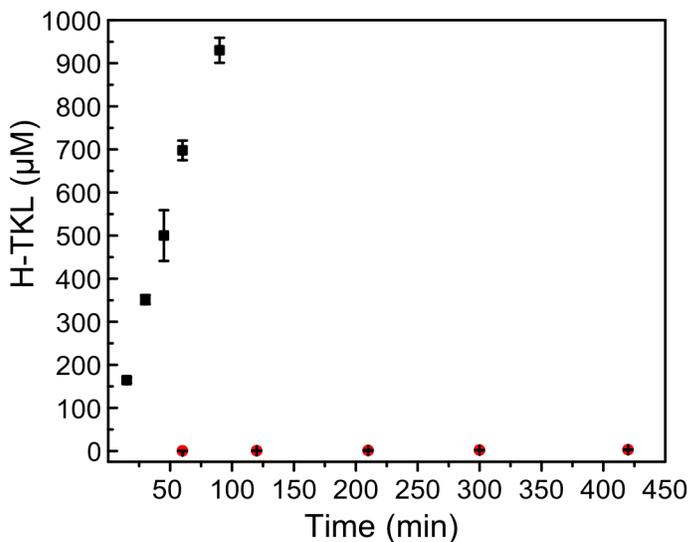
Figure 2.16:. *Initial rate of H-TKL formation with AT^0 Mod3$_{TE}$ complemented with DszAT (10.8 ± 0.3 µM min^{-1}, black) and AT^0 Mod3$_{TE}$ (0.006 ± 0.001 µM min^{-1}, red). Data are mean ± s.d. (n=3).*

Site-specific incorporation of fluorinated extender units into polyketide synthases. One major challenge for fluorine incorporation into full-length polyketides is the transfer of the growing chain from module to module. We have previously shown that a fluorinated extender unit can be incorporated into a polyketide nascent chain and undergo an additional chain extension with methylmalonyl-CoA to produce a monofluorinated product using a non-covalent di-modular "mini-PKS" system containing Mod2 and Mod3$_{TE}$ [*10*]. However, this system requires a 5:1 ratio of Mod2 AT^0 to Mod3$_{TE}$ to overcome the barrier to chain transfer, which is a stoichiometry that would be difficult to control under *in vivo* conditions. Given our ability to efficiently complement AT^0 Mod3$_{TE}$ with DszAT and fluoromalonyl-CoA, we decided to approach both the production of bis-fluorinated products and the improvement of the ratio of PKS modules to a more physiologically-relevant regime (1:1). We chose to use propionyl-SNAC (*Figure 2.20*) as the starter unit to produce F-TKLs in order to avoid the offloading that results from the propensity of triketidyl-thioesters to undergo cyclization of the polyketide chain after a single extension from NDK-SNAC (*Figure 2.19*) [*10*]. Additionally, the CoA-pool concentration was lowered from 1 mM to 0.1 mM, to lower competition between fluoromalonyl-CoA and fluoromalonyl-ACP for the KS active site that we observed with saturating concentrations of extender unit. Under these conditions, the complementation with the different AT^0 variants of the mini-PKS under regenerative methylmalonyl-CoA and fluoromalonyl-CoA conditions led to the formation of the expected site-specifically fluorinated products with a 1:1 ratio of Mod2 to Mod3$_{TE}$ (*Figure 2.19B*). To further substantiate the identity of these products, [1-^{13}C]-propionyl-SNAC was used as the starter unit, leading to the expected 1 amu shift in mass of the products. The two mono-fluorinated regioisomers were determined to be unique by different retention times, as well as distinctive fragmentation patterns (*Figures 2.19B, 2.21, and 2.22*). Control reactions with no added DszAT showed a greatly diminished F-TKL yield (up to 75- to 93-fold for monofluorinated products, while the bisfluorinated product could not be detected), indicating that DszAT complementation is crucial to achieve efficient production of an F-TKL formed from subsequent chain extensions (*Figure 2.19B*).
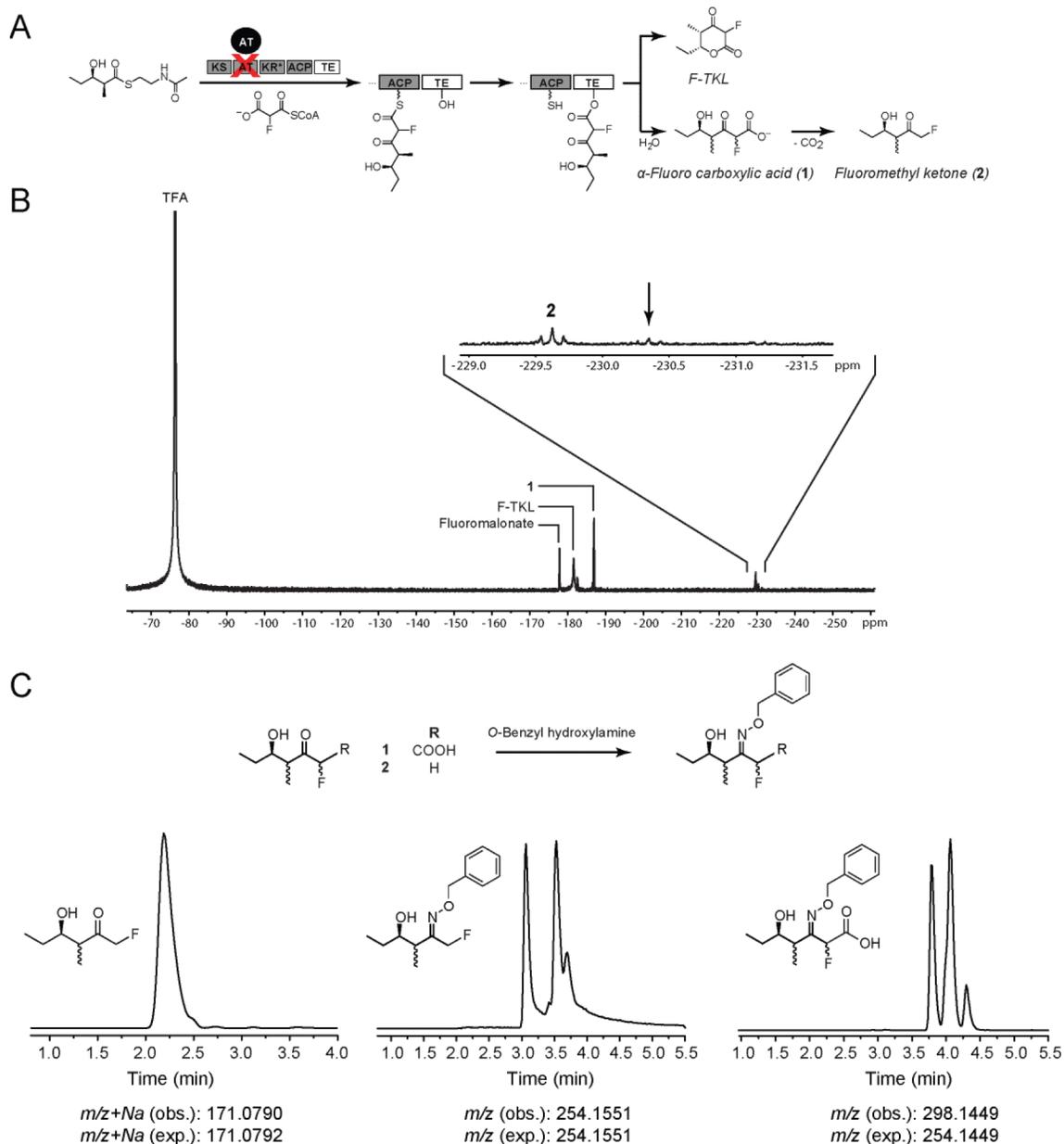
Figure 2.17: *Determining other major products formed in the chain extension reaction catalyzed by AT^0 $Mod3_{TE}$ and DszAT with fluoromalonyl-CoA. (A) Reaction scheme for formation of the α-fluorocarboxylic acid (**1**) and fluoromethyl ketone (**2**). (B) ^{19}F-NMR analysis of the reaction in 12.5% D_2O shows new peaks, one with a chemical shift consistent with a fluoromethyl ketone (**2**) and a triplet splitting pattern resulting from fluorine coupling to the two equivalent hydrogens bound to the ipso carbon (J_{HF} = 47.5 Hz). A second triplet is observed (arrow) and believed to be the result of epimerization of the methyl group. The α-fluoro carboxylic acid precursor (**1**) to **2** was then noted and appears as a doublet (J_{HF} = 49.9 Hz). (C) LC/QTOF analysis of **2** before and **1** and **2** after derivatization with O-benzyl hydroxylamine. **1** was not retained on the column before derivatization and the EIC is therefore unavailable.*
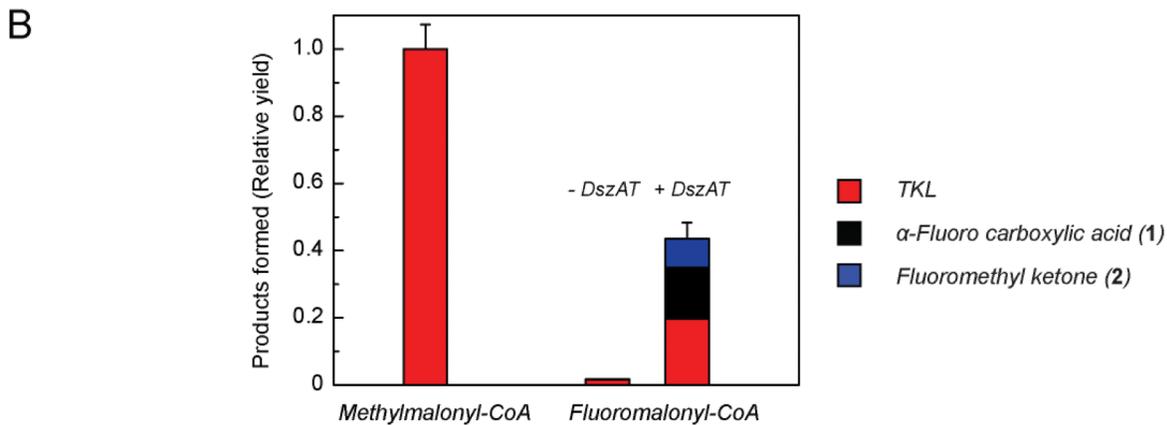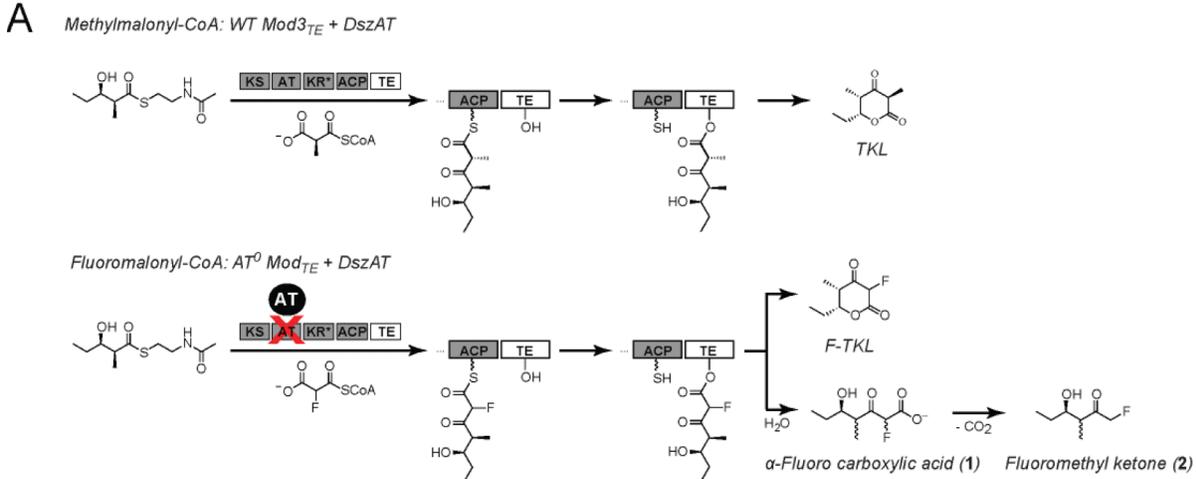
47

Figure 2.18: Comparison of C-C bond formation stoichiometry. (A) Reaction pathways for fluoromalonyl- and methylmalonyl-CoA. (B) Normalized bar graph for the relative yields of C-C bond products. Terminal ketones and carboxylic acids were included as a single C-C bond-forming event as they diverge from the F-TKL pathway after fluoromalonyl-CoA is incorporated. Data are represented average ± s.e. (propagated from the s.d. (n=3) of the individual products). The fluoromethyl ketone and α-fluorocarboxylic acid were quantified by LC-UV after derivatization with O-benxylhydroxylamine. (C) Tabulated data with concentrations of each observed product. Data are average ± s.d. (n.d., not detected; -, not measured)
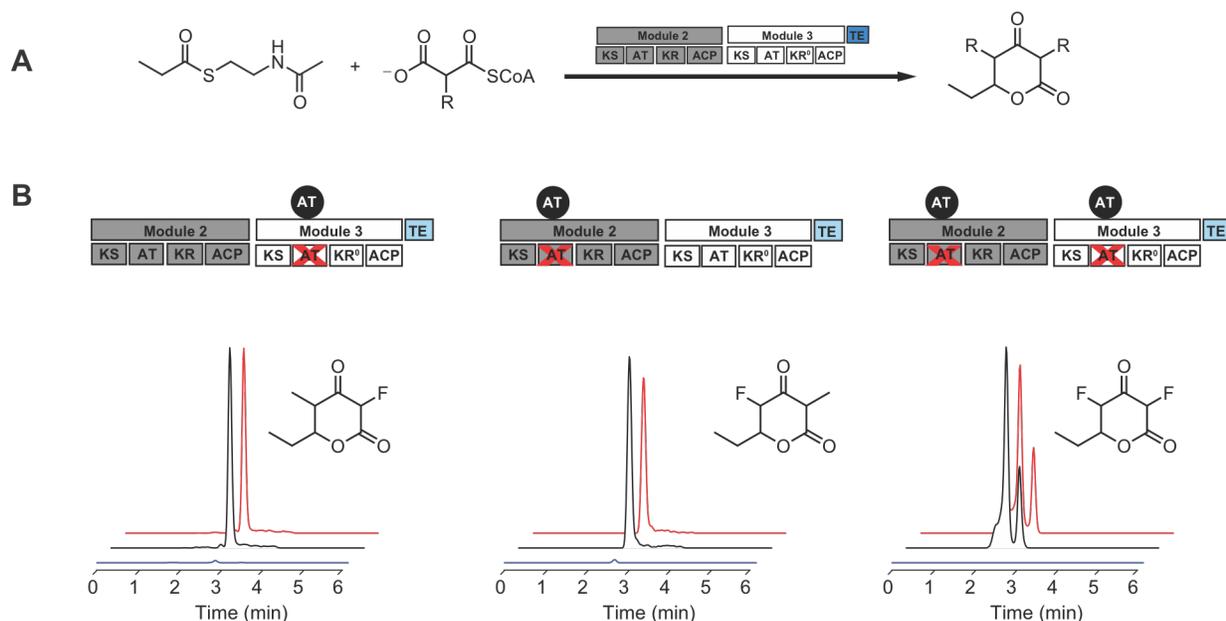
Figure 2.19: Formation of TKL products from propionyl-SNAC using a dimodular PKS system consisting of Mod2 and Mod3$_{TE}$. (A) Reaction scheme for TKL formation (R = Me and F). (*, natively inactive KR domain) (B) EICs of the two monofluorinated TKL isomers (natural abundance, 173.0619 m/z; ^{13}C-labeled product, 174.0653 m/z) and the bis-fluorinated TKL (natural abundance, 177.0369 m/z; ^{13}C-labeled product, 178.0403 m/z) corresponding to the insertion of one and two fluoromalonyl-CoA monomers, respectively (propionyl-SNAC + DszAT, black; ^{13}C-propionyl-SNAC + DszAT, red; propionyl-SNAC – DszAT, blue). The splitting of the peak observed for the bis-fluorinated TKL is likely related to the existence of tautomeric forms.

49
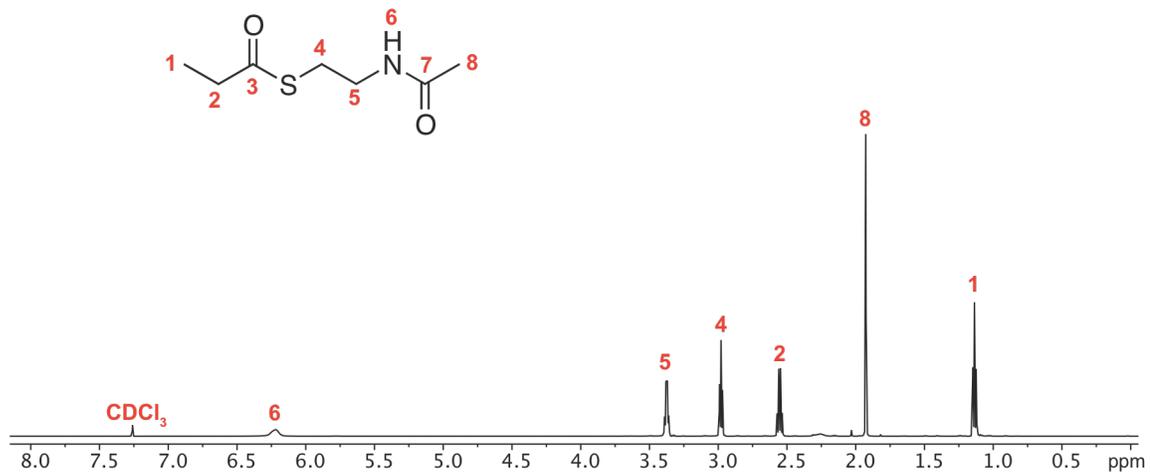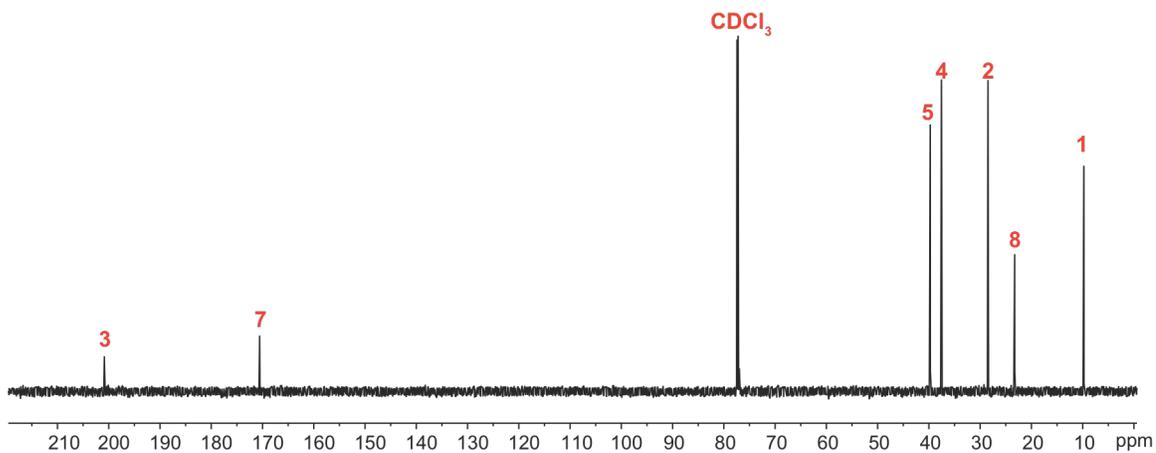
Figure 2.20: *NMR spectra of propionyl-SNAC. (A) ^{1}H-NMR. (B) ^{13}C-NMR*
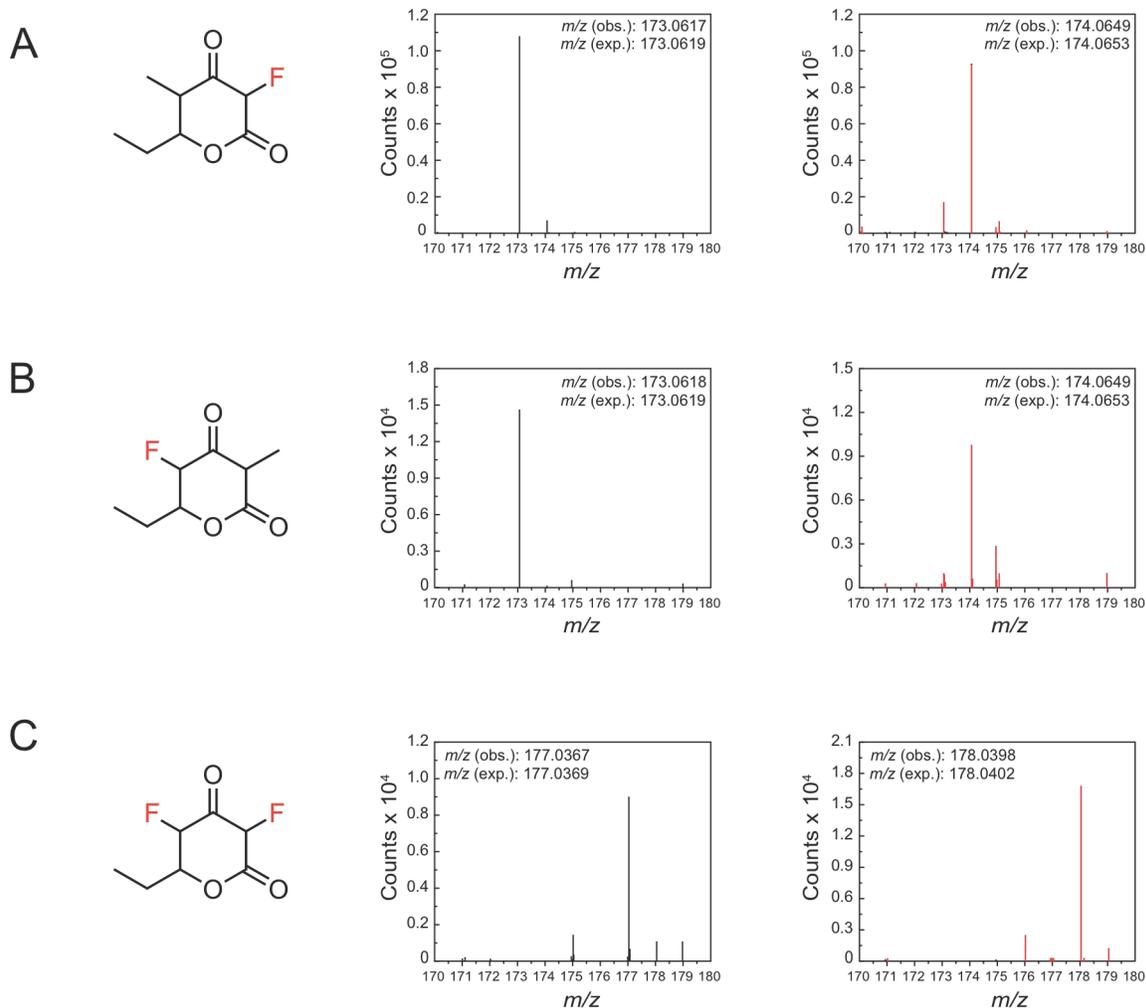
Figure 2.21: *Exact mass spectra for the triketides observed from the mini-PKS. (A) The Me/F product generated using Mod2 + AT[0] Mod3[TE] and DszAT with unlabled propionyl-SNAC (black) and ^{13}C-labled propionyl-SNAC (red). (B) The F/Me product generated using AT[0] Mod2 + Mod3[TE] and DszAT with unlabled propionyl-SNAC (black) and ^{13}C-labled propionyl-SNAC (red). (C) The F/F product generated using AT[0] Mod2 + AT[0] Mod3[TE] and DszAT with unlabled propionyl-SNAC (black) and ^{13}C-labled propionyl-SNAC (red).*
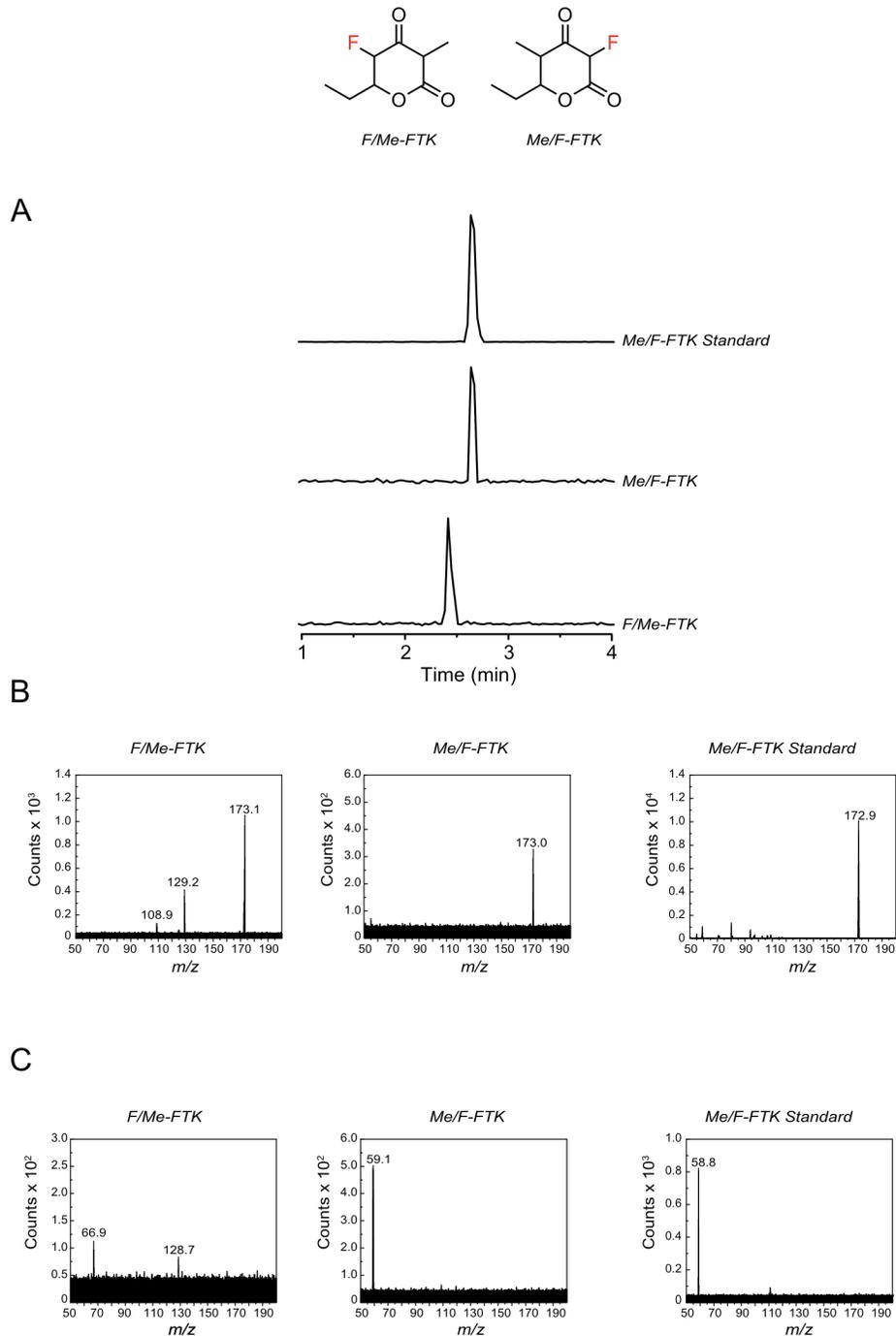
Figure 2.22: *Further characterization of the Me/F and F/Me TKL products. (A) Product ion chromatogram (173.1 m/z) for Me/F (Frag. voltage 100 V and collision energy 25 V) and F/Me (Frag. voltage 100 V and collision energy 15 V) reactions as well as an Me/F authentic standard (Frag. voltage 100 V and collision energy 25 V). As shown, the two TKLs have different retention times in this 4 min method. (B) The product ion spectra (Frag. voltage 100 V and collision energy 15 V) for the two TKLs and the chemical standard for the Me/F triketide. (C) The product ion spectra (Frag. voltage 100 V and collision energy 25 V) for the two TKLs and the chemical standard for the Me/F triketide. Note that under these conditions the F/Me and Me/F triketide have different fragmentation patterns.*

2.4   Conclusions

We set out to explore the molecular mechanism of fluorinated monomer incorporation by modular PKSs using DEBS as a model system with the goal of identifying limitations to be addressed in order to develop a robust system for production of site-selectively fluorinated polyketides. In this study, we found that inactivation of the AT domain, which is a strategy used to eliminate of the major selectivity filters for choosing the extender unit, initiates a mode of C-C bond formation that is independent of the ACP (*Scheme 2.2*). Although we had previously hypothesized that the more activated fluoromalonyl-CoA extender unit might be able to achieve ACP acylation with a catalytically-deficient AT domain, we observed instead that the KS domain can compete with the AT domain for both the native methylmalonyl-CoA unit and fluoromalonyl-CoA to carry out direct C-C bond formation while bypassing the canonical ACP-linked intermediate. This observation is consistent with previous studies demonstrating direct KS reactivity with non-hydrolyzable malonyl-SNAC mimics [*34,35*] as well as the *in vitro* detection of unnatural TKLs using Mod6$_{TE}$ heterologously expressed in *E. coli* without a phosphopantetheinyl transferase enzyme [*26*]. While this mode of reactivity does lead to a single chain extension, the covalent tether between the ACP and growing polyketide chain that ensures its transfer to the downstream module is lost, which leads to loss of processivity and ability to efficiently produce full-length polyketide products. Interestingly, we found that reaction partitioning between productive substrate decarboxylation to form the C-C bond and nonproductive pathways leading to release of propionate or fluoroacetate was also altered. For the native methylmalonyl-CoA, a ~685-fold improvement in this partition was observed when C-C
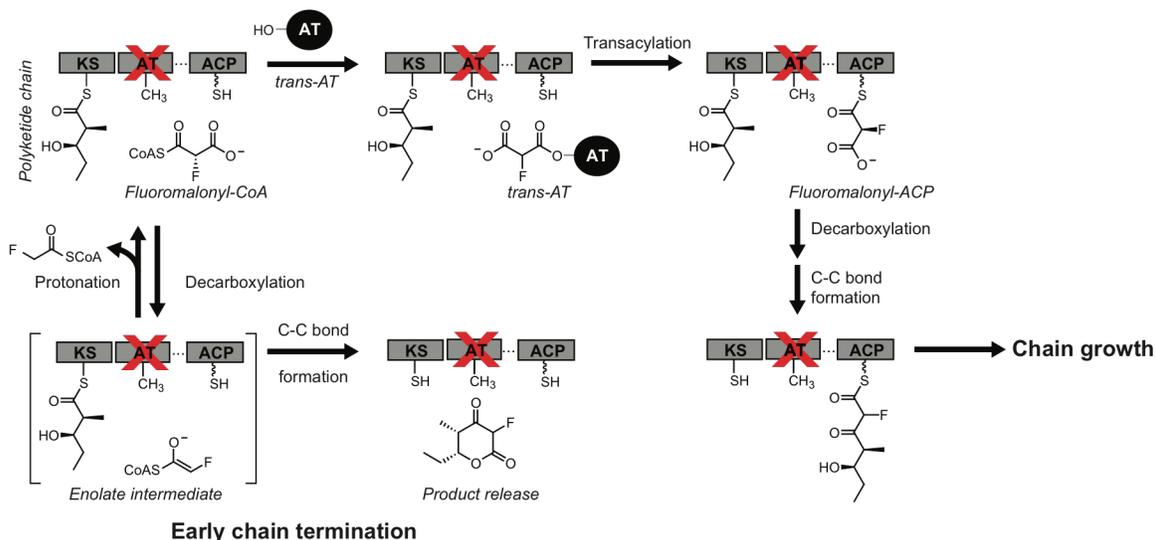
Scheme 2.2: *Possible reaction pathways for PKS-catalyzed chain extension when fluoromalonyl-CoA is used. Use of a PKS module with the AT inactivated (Ser → Ala) allows bypassing the main selectivity filter for native extender unit selection and enabling the use of fluoromalonyl-CoA. In the absence of a trans-AT, the extender unit can diffuse directly into the KS active site and undergo decarboxylation. In this mode of reactivity, decarboxylation is decoupled from chain extension and enolate protonation can lead to release of fluoroacetyl-CoA and degradation of the extender unit pool. Even when C-C bond formation occurs, the lack of an ACP-bound intermediate leads to early chain termination and release of the growing polyketide. In the presence of a trans-AT (DszAT), the AT^0 mutation can be complemented leading to ACP acylation with fluoromalonyl-CoA. When the canonical intermediate is formed, decarboxylation of fluoromalonyl-ACP is coupled to C-C bond formation with no loss of extender unit, resulting ultimately in chain growth and of formation of the full-length polyketide product.*

bond formation proceeded through the proper ACP-linked rather than CoA-linked extender unit. Although less dramatic, the partition for C-C bond formation with the fluoromalonyl-CoA extender also improves from 2:1 to quantitative.

Based on the results of these experiments, we carried out a detailed study of *trans*-AT complementation with DEBS modules in which we characterized the molecular mechanism of fluoromalonyl-CoA incorporation and rate enhancement upon the complementation of DszAT with DEBS modules. Using end product analysis, enzyme kinetics, and direct observation by MS with various DEBS constructs, we are able to show that complementation of AT^0 modules with a *trans*-AT, DszAT, enables formation of the canonical ACP-linked intermediate with fluoromalonyl-CoA. Under these conditions, C-C bond formation with a fluorinated monomer occurs at 43% yield compared to the wild-type system with the native methylmalonyl-CoA extender unit. Formation of fluoromalonyl-ACP has key implications for production of full-length polyketides, as the covalent linkage between the growing chain and PKS is maintained after the fluorinated monomer is incorporated. Furthermore, C-C bond formation is fully coupled to fluoromalonyl-ACP decarboxylation with no observed nonproductive loss to fluoroacetate.

We then sought to test these findings in a dimodular "mini-PKS" system, where we had previously required the usage of a 5:1 ratio of the two modules to achieve two rounds of chain extension without DszAT complementation [10]. It is in these multiple turnover systems where the identification of the two modes of C-C bond formation in AT^0 systems has dramatic implications on product yield and regiochemistry. Given that the ACP-independent mode of chain extension can be activated with high levels of extender unit, decreasing the concentration of CoA extender units below this critical amount should increase chain transfer efficiency by reducing the amount of chain termination occurring from direct diffusion of the extender unit into the KS active site. Using this information, we are able to show for the first time that mono- and bis-fluorinated products can be synthesized from the dimodular system at a 1:1 ratio of DEBS modules, indicating that channeling fluorinated growing chain from an upstream to a downstream module is possible. We further show that DszAT complementation increases titers by up to ~93-fold in this dimodular system. The ability to complement DszAT with two sequential AT^0 modules to insert two fluorinated extender units suggests that *trans*-AT complementation with PKS modules shows promise as a method for achieving robust production of site-specifically fluorinated full-length polyketide products.

2.5   References

1.   Staunton, J., Weissman, K. J. Polyketide biosynthesis: A millennium review. *Nat Prod Rep* **2001**, *18* (4) 380-416.

2.   Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A., Cane, D. E. Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu Rev Biochem* **2007**, *76* 195-221.

3.   Dunn, B. J., Khosla, C. Engineering the acyltransferase substrate specificity of assembly line polyketide synthases. *J R Soc Interface* **2013**, *10* (85).

4.   Ray, L., Moore, B. S. Recent advances in the biosynthesis of unusual polyketide synthase substrates. *Natural product reports* **2015**.

5.   Hagmann, W. K. The many roles for fluorine in medicinal chemistry. *J Med Chem* **2008**, *51* (15) 4359-4369.

6.   Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat Prod Rep* **2012**, *29* (10) 1050-1073.

7.   Bravo-Rodriguez, K., Klopries, S., Koopmans, K. R., Sundermann, U., Yahiaoui, S., Arens, J., Kushnir, S., Schulz, F., Sanchez-Garcia, E. Substrate flexibility of a mutated acyltransferase domain and implications for polyketide biosynthesis. *Chem Biol* **2015**, *22* (11) 1425-1430.

8.   Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., Schulz, F. Enzyme-directed mutasynthesis: A combined experimental and theoretical approach to substrate recognition of a polyketide synthase. *ACS Chem Biol* **2013**, *8* (2) 443-450.

9.   McDaniel, R., Thamchaipenet, A., Gustafsson, C., Fu, H., Betlach, M., Ashley, G. Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc Natl Acad Sci U S A* **1999**, *96* (5) 1846-1851.

10.   Walker, M. C., Thuronyi, B. W., Charkoudian, L. K., Lowry, B., Khosla, C., Chang, M. C. Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* **2013**, *341* (6150) 1089-1094.

11.   Dunn, B. J., Watts, K. R., Robbins, T., Cane, D. E., Khosla, C. Comparative analysis of the substrate specificity of *trans-* versus *cis*-acyltransferases of assembly line polyketide synthases. *Biochemistry* **2014**, *53* (23) 3796-3806.

12.   Koryakina, I., McArthur, J., Randall, S., Draelos, M. M., Musiol, E. M., Muddiman, D. C., Weber, T., Williams, G. J. Poly specific *trans*-acyltransferase machinery revealed via engineered acyl-CoA synthetases. *ACS chemical biology* **2013**, *8* (1) 200-208.

13.   Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **2001**, *291* (5509) 1790-1792.

14.   Gokhale, R. S., Tsuji, S. Y., Cane, D. E., Khosla, C. Dissecting and exploiting intermodular communication in polyketide synthases. *Science* **1999**, *284* (5413) 482-485.

15.   Wong, F. T., Chen, A. Y., Cane, D. E., Khosla, C. Protein-protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. *Biochemistry* **2010**, *49* (1) 95-102.

16.   Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **2009**, *6* (5) 343-345.

17.   Chen, A. Y., Schnarr, N. A., Kim, C. Y., Cane, D. E., Khosla, C. Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-deoxyerythronolide B synthase. *J Am Chem Soc* **2006**, *128* (9) 3067-3074.

18.   Engler, C., Kandzia, R., Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PloS one* **2008**, *3* (11) e3647.

19. Theodorou, V., Skobridis, K., Tzakos, A. G., Ragoussis, V. A simple method for the alkaline hydrolysis of esters. *Tetrahedron Lett* **2007**, *48* (46) 8230-8233.

20. Bonnett, S. A., Papireddy, K., Higgins, S., del Cardayre, S., Reynolds, K. A. Functional characterization of an NADPH dependent 2-alkyl-3-ketoalkanoic acid reductase involved in olefin biosynthesis in *Stenotrophomonas maltophilia*. *Biochemistry* **2011**, *50* (44) 9633-9640.

21. Piasecki, S. K., Taylor, C. A., Detelich, J. F., Liu, J., Zheng, J., Komsoukaniants, A., Siegel, D. R., Keatinge-Clay, A. T. Employing modular polyketide synthase ketoreductases as biocatalysts in the preparative chemoenzymatic syntheses of diketide chiral building blocks. *Chemistry & biology* **2011**, *18* (10) 1331-1340.

22. Rocha, D. F., Wouters, F. C., Machado, G., Marsaioli, A. J. First biosynthetic pathway of 1-hepten-3-one in *iporangaia pustulosa* (opiliones). *Scientific reports* **2013**, *3* 3156.

23. Palla, K. S., Witus, L. S., Mackenzie, K. J., Netirojjanakul, C., Francis, M. B. Optimization and expansion of a site-selective n-methylpyridinium-4-carboxaldehyde-mediated transamination for bacterially expressed proteins. *J Am Chem Soc* **2015**, *137* (3) 1123-1129.

24. Dunn, B. J., Cane, D. E., Khosla, C. Mechanism and specificity of an acyltransferase domain from a modular polyketide synthase. *Biochemistry* **2013**, *52* (11) 1839-1841.

25. Wu, N., Tsuji, S. Y., Cane, D. E., Khosla, C. Assessing the balance between protein-protein interactions and enzyme-substrate interactions in the channeling of intermediates between polyketide synthase modules. *J Am Chem Soc* **2001**, *123* (27) 6465-6474.

26. Koryakina, I., McArthur, J. B., Draelos, M. M., Williams, G. J. Promiscuity of a modular polyketide synthase towards natural and non-natural extender units. *Organic & biomolecular chemistry* **2013**, *11* (27) 4449-4458.

27. Jenner, M., Afonso, J. P., Bailey, H. R., Frank, S., Kampa, A., Piel, J., Oldham, N. J. Acyl-chain elongation drives ketosynthase substrate selectivity in *trans*-acyltransferase polyketide synthases. *Angewandte Chemie* **2015**, *54* (6) 1817-1821.

28. Wong, F. T., Jin, X., Mathews, I. I., Cane, D. E., Khosla, C. Structure and mechanism of the *trans*-acting acyltransferase from the disorazole synthase. *Biochemistry* **2011**, *50* (30) 6539-6548.

29. Meluzzi, D., Zheng, W. H., Hensler, M., Nizet, V., Dorrestein, P. C. Top-down mass spectrometry on low-resolution instruments: Characterization of phosphopantetheinylated carrier domains in polyketide and non-ribosomal biosynthetic pathways. *Bioorganic & medicinal chemistry letters* **2008**, *18* (10) 3107-3111.

30. Hagen, A., Poust, S., de Rond, T., Yuzawa, S., Katz, L., Adams, P. D., Petzold, C. J., Keasling, J. D. *In vitro* analysis of carboxyacyl substrate tolerance in the loading and first extension modules of borrelidin polyketide synthase. *Biochemistry* **2014**, *53* (38) 5975-5977.

31. Poust, S., Phelan, R. M., Deng, K., Katz, L., Petzold, C. J., Keasling, J. D. Divergent mechanistic routes for the formation of *gem*-dimethyl groups in the biosynthesis of complex polyketides. *Angewandte Chemie* **2015**, *54* (8) 2370-2373.

32. Bracher, P. J., Snyder, P. W., Bohall, B. R., Whitesides, G. M. The relative rates of thiol-thioester exchange and hydrolysis for alkyl and aryl thioalkanoates in water. *Orig. Life Evol Biosph* **2011**, *41* (5) 399-412.

33. Randall, S. M., Koryakina, I., Williams, G. J., Muddiman, D. C. Evaluating nonpolar surface area and liquid chromatography/mass spectrometry response: An application for site occupancy measurements for enzyme intermediates in polyketide biosynthesis. *Rapid Commun Mass Spectrom* **2014**, *28* (23) 2511-2522.

34. Tosin, M., Demydchuk, Y., Parascandolo, J. S., Per, C. B., Leeper, F. J., Leadlay, P. F. *In vivo* trapping of polyketide intermediates from an assembly line synthase using malonyl carba(dethia)-N-acetyl cysteamines. *Chemical communications* **2011**, *47* (12) 3460-3462.

35. Riva, E., Wilkening, I., Gazzola, S., Li, W. M. A., Smith, L., Leadlay, P. F., Tosin, M. Chemical probes for the functionalization of polyketide intermediates. *Angew Chem Int Edit* **2014**, *53* (44) 11944-11949.

Chapter 3: *Towards in vivo production of site-specifically fluorinated polyketides using a fluoromalonyl-CoA selective trans-AT*

58

3.1 Introduction

The production of site-specifically fluorinated polyketides using malonyl-CoA selective *trans*-ATs and DEBS modules containing inactive *cis*-ATs has shown potential as a method for site-specific incorporation of fluorine into the backbone of type I modular polyketide synthases [1-3]. This particular strategy builds on the steric similarity between fluoromalonyl-CoA and malonyl-CoA [4], as well as the highly selective nature of methylmalonyl-CoA selective ATs against malonyl-CoA [1,2,5]. Taken together, the inactivation a methylmalonyl-CoA selective cis-*AT*, along with the complementation of a malonyl-CoA selective AT, meets the initial requirements for site-specific incorporation of fluorinated extender units (*Scheme 3.1*) [2]. More importantly, with this tool in hand, the ability to produce site-specifically fluorinated full-length polyketides *in vivo* becomes increasingly more feasible. Using this strategy, a heterologous host expressing a *trans-AT*, along with a PKS cluster containing an inactivated *cis*-AT (AT^0) in the module corresponding to the desired site of fluorine incorporation can produce site-specifically fluorinated natural products (*Scheme 3.1*).
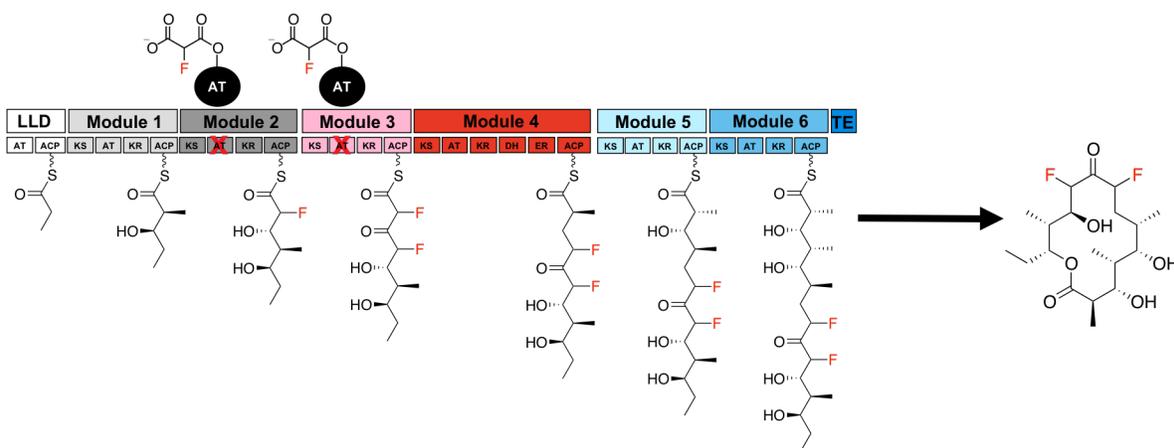
Scheme 3.1: *Representation of an engineered DEBS cluster complemented with DszAT resulting in the formation of a site-specifically fluorinated 6-dEB derivative.*

The first requirement for the production of fluorinated polyketides *in vivo* is the synthesis of the extender unit fluoromalonyl-CoA. Previous work has shown that fluoromalonyl-CoA can be synthesized from fluoromalonate by various malonyl-CoA synthetase (MatB) homologs [1,2]. Additionally, recent studies have shown that malonate derivatives can be taken up and activated by *Streptomyces* using endogenous synthetases, as well as a genomically integrated MatB variants [6]. The second requirement, and one of the major challenges for the production of fluorinated polyketides using this strategy, remains the preference of malonyl-CoA selective *trans*-ATs for malonyl-CoA acylation over fluoromalonyl-CoA [2]. Studies involving the malonyl-CoA selective *trans*-AT from the Disorazole biosynthetic pathway (DszAT) [7,8] show that DszAT exhibits a 10-fold selectivity towards malonyl-CoA over fluoromalonyl-CoA [2]. As such, *in vivo* complementation of DszAT with the appropriate DEBS construct will result in the formation of desmethlylated products, as opposed to the desired fluorinated derivatives given the presence of malonyl-CoA in the cell (*Scheme 3.2*). It is important to mention that while methylmalonyl-CoA is present *in vivo*, DszAT is highly selective against methylmalonyl-CoA and will likely not load methylmalonyl-CoA onto DEBS modules using DszAT [9]. In order to limit incorporation of malonyl-CoA, we set out to engineer a fluoromalonyl-CoA selective variant of DszAT. This

approach was selected over genetic modifications to host malonyl-CoA biosynthetic enzymes, since these methods may inhibit growth rate, as well as render the organism more susceptible to fluoroacetate toxicity [10], limiting the production of natural products.
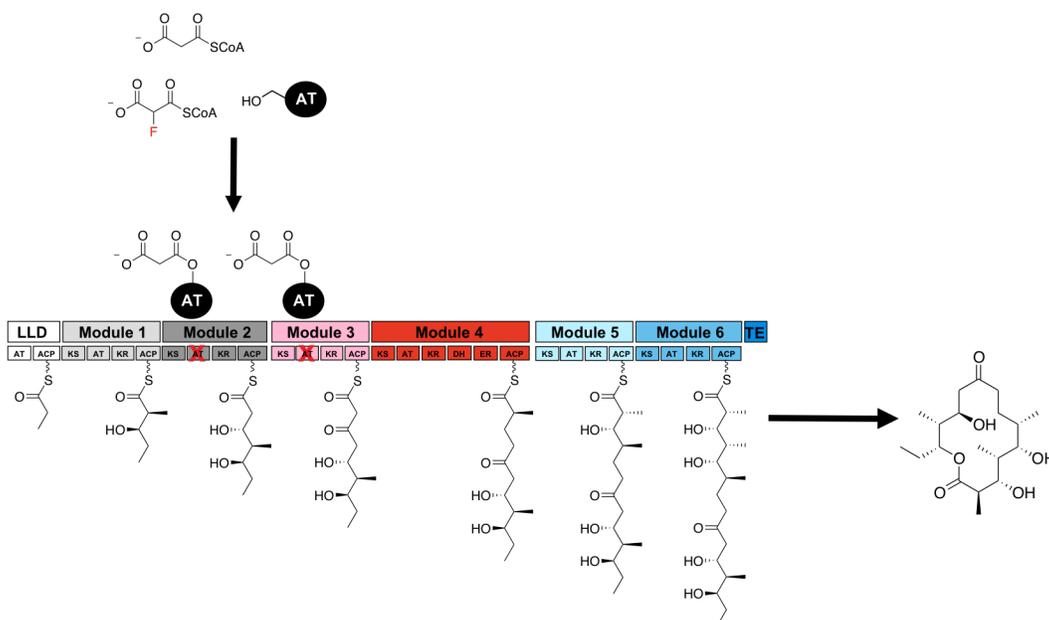
Scheme 3.2: *Proposed representation of an engineered DEBS cluster complemented with WT DszAT in vivo. DszAT will select for it's native substrate (malonyl-CoA) over fluoromalonyl-CoA, resulting in the formation of the didesmethylated 6-dEB derivative.*

Although examples of successful *cis-* or *trans*-AT engineering have been limited to mostly naturally occurring extender units [11-13], we set out to engineer a fluoromalonyl-CoA selective DszAT variant. Given the activated nature of fluoromalonyl-CoA and the fact that formation of F-TKLs is only an order of magnitude lower than formation of H-TKL [2], we focused specifically on identifying DszAT variants with reduced ability to acylate DEBS ACPs with malonyl-CoA without significantly impairing fluoromalonyl-CoA incorporation, rather than those that only demonstrate enhanced F-TKL production levels. The mutagenesis screens were designed based on the conserved malonyl-CoA specific AT residues, as well as guided by a crystal structure obtained for DszAT [8]. Ultimately, we found that 53% of the first round of mutants screened displayed equal or greater fluoromalonyl-CoA incorporation levels, suggesting that the high hit rate in the initial screen can be attributed to fluoromalonyl-CoA reactivity even in a compromised active site. The most selective variant was found to be F190V, which shows a 41-fold improvement in fluorine selectivity compared to WT DszAT with an 11-fold higher yield of F-TKL compared to malonyl-CoA-derived products. Furthermore, the F190V mutant did not exhibit higher activity with methylmalonyl-CoA compared to WT DszAT, implying that the increased selectivity is not related to a simple expansion of the active site pocket. Additional characterization has also shown that this fluorine-selective variant can be used in the presence of malonyl-CoA to produce fluorinated polyketide products.

The production of site-specifically fluorinated polyketides ultimately relies on a robust and genetically tractable expression system. The ability to integrate extender unit activating enzymes, as well as the introduction of AT^0 mutations remains challenging in *Streptomyces*. Even so,

Streptomyces remain the preferred host organism for the production of polyketides. Work in the latter part of this chapter will discuss the methods used for the introduction of point mutations into polyketides as well as the workflow required to engineering a heterologous *Streptomyces* host capable of producing site-specifically fluorinated polyketides.

3.2   Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), nutrient broth, malt extract, yeast extract, and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Carbenicillin (Cb), glucose, isopropyl-β-D-thiogalactopyranoside (IPTG), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride, dithiothreitol (DTT), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, dichloromethane, ethyl acetate, ethylene diamine tetraacetic acid disodium dihydrate (EDTA), glycine, sucrose, potassium sulfate, monopotassium phosphate, calcium chloride dihydrate, and restriction enzymes were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A sodium salt (CoA), malonyl-CoA, methylmalonyl-CoA, diethylfluoromalonate, malonic acid, diethylmethylmalonate, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenolpyruvate (PEP), adenosine triphosphate sodium salt (ATP), myokinase, pyruvate kinase, lactate dehydrogenase, poly(ethyleneimine) solution (PEI), mannitol, TES, β-mercaptoethanol, thiamine pyrophosphate, alpha ketoglutaric acid, sodium phosphate dibasic hepthydrate, cysteamine, 4-hydroxy-6-methyl-2-pyrone acetic anhydride, 1-ethyl-3-(-dimethylaminopropyl)carbodiimide hydrochloride (EDC), 4-dimethylaminopyridine (DMAP), *o*-benzyl hydroxylamine, N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED), reduced nicotinamide adenine dinucleotide phosphate (NADPH), acetonitrile, dimethyl sulfoxide (DMSO), Thiostrepton, PEG1000, sodium propionate, ammonium acetate, and ammonium formate were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate was purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, T4 DNA ligase, Phusion DNA polymerase, amylose resin, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 µM in 10 mM Tris-HCl, pH 8.5, and stored at either 4°C for immediate use or -20°C for longer term use. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Casaminoacids and peptone were purchase from Bekton Dickinson (Franklin Lakes, NJ). Amicon Ultra 3,000 MWCO, 10,000 MWCO centrifugal concentrators, and 30,000 MWCO centrifugal concentrators were purchased from EMD Millipore (Billerica, MA). Sodium fluoroacetate was purchased from Fluka-Honeywell (Mexico City, Mexico). Soy flour was purchased from Berkeley Bowl (Berkeley, CA) Chloroform-*d* were purchased from Cambridge Isotope Laboratories (Andover, MA). ^{13}C and ^{1}H NMR spectra were collected at 25°C on Bruker AV-600 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley or on a Bruker Biospin 900 MHz spectrometer at the QB3 Central California 900 MHz NMR Facility NMR. Assignments were made based on literature president and reference spectra from authentic standards, where

appropriate. High-resolution mass spectral analyses were carried out on a 6530 QTOF Accurate Mass spectrometer purchased from Agilent Technologies.

Bacterial strains. *E. coli* DH10B-T1^R and BL21(de3)T1^R were used for plasmid construction and heterologous protein expression, respectively. *E. coli* BAP1 [*14*] cells were used for heterologous expression of DEBS modules. *Actinomadura vulgaris* (ATCC 53715), *Streptomyces venezuelae* (ATCC 15439), and *Streptomyces lividans* (K4-114) were used for polyketide production experiments.

Gene and plasmid construction. Standard molecular biology techniques were used for plasmid construction (unless explicitly specified). All PCR amplifications were done using Phusion High Fidelity DNA polymerase or Q5 High Fidelity DNA polymerase. For amplification of GC-rich sequences from *S. erythraea*, PCR reactions with Phusion polymerase were supplemented with DMSO (10% *v/v*) using the standard buffer and PCR reactions with Q5 polymerase were supplemented with GC Enhancer solution. Primer annealing temperatures were 6-8°C below the T_m with a maximum temperature of 72°C. All constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA).

DszAT mutant libraries were generated by site-specific mutagenesis at selected catalytic residues within the active site of the enzyme. The libraries constructed full saturation at the amino acids at the F190 position (according to DszD numbering), a library at the S86 catalytic residue (S86C, S86D, S86E, S86A) and a double mutant S86; H191 library for the same mutants (S86C; H191A, S86D; H191A, and S86E; H191A) was also constructed. Finally, a double mutant library for the F190; L87 residues was generated for select F190 mutants from the single mutant screens a secondary screen for enhanced fluorine substrate selectivity (F190G; L87A, F190G; L87V, F190S; L87A, F190S; L87V, F190T; L87A, F190T; L87V, F190V; L87A, F190V; L87V, F190P; L87A, F190I; L87A, and F190I; L87V).

The mutant library was constructed by amplification from pFW3 [*15*]. The F190 mutations were introduced by amplifying pFW3 with two sets of primers. One fragment was amplified with DszAT F190 F1 and various F190 R1 primers (depending on mutation). The second fragment was amplified using various F190 F2 primers (depending on mutation) and DszAT F190 R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments where inserted into the XbaI-HindIII sites of pFW3 using the Gibson Protocol [*16*].

The L87 mutations were introduced by amplifying pFW3 [*15*] with two sets of primers. One fragment was amplified with DszAT L87 F1 and various L87 R1 primers (depending on mutation). The second fragment was amplified using various L87 F2 primers (depending on mutation) and DszAT L87 R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments where inserted into the XbaI-HindIII sites of pFW3 using the Gibson Protocol [*16*].

The H191A mutation was introduced by amplifying pFW3 [*15*]. One fragment was amplified with DszAT H191A F1/R1 and the second fragment was amplified DszAT H191A F2/R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments where inserted into the NdeI-EcoRI sites of pFW3 using the Gibson Protocol [*16*].

The S86 mutations were introduced by amplifying pET21c-DszAT H191A-His$_6$ or pFW3 with two sets of primers. One fragment was amplified with DszAT S86 F1 and various S86 R1 primers

(depending on mutation). The second fragment was amplified using various S86 F2 primers (depending on mutation) and DszAT S86 R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments where inserted into the XbaI-HindIII sites of pFW3 or pET21c-DszAT-H191A-His$_6$ using the Gibson Protocol [16].

All double mutants were designed as described above, with the exception of PCR template containing desired mutations from the initial round of screening.

The F190V mutation was introduced into pCDFDuet-DszAT-F190V-rpMatB-T103A by amplifying pCDFDuet1-DszAt.RpMatB-T103A. One fragment was amplified with pCDF-DszAT-F190V-MatB F1/R1 and the second fragment was amplified pCDF-DszAT-F190V-MatB F2/R2. The two fragments contained a 60 bp overlap. After treatment with DpnI, the two PCR fragments where inserted into the SacI-Aatll sites of pCDFDuet1-DszAt.RpMatB-T103A using the Gibson Protocol [16].

pSET152-ErmEP*-DszAT-MatB and was assembled by amplification of the DszAT and MatB genes from pUWL-DszAT-MatB. The genes encoding DszAT and MatB were introduced by amplification with the primer sets pSET152 DszAT MatB F1/R1 and pSET152 DszAT MatB F2/R2, which contained a 60 bp overlap. The PCR fragments were inserted into the NdeI-BamHI sites of pSET152-ErmEP* [17,18] using the Gibson Protocol [16].

pSET152-ErmEP*-F190V-DszAT-MatB was assembled by amplification of pSET152-ErmEP*-DszAT-MatB. The genes encoding DszAT and MatB were introduced by amplification with the primer sets pSET152 F190V DszAT F1/R1 and pSET152 DszAT MatB F2/R2, which contained a 60 bp overlap. The PCR fragments were inserted into the NdeI-HindIII sites of pSET152-ErmEP*-DszAT-MatB using the Gibson Protocol [16].

pSET152-ErmEP*-NphT7-PhaB was constructed in a sequential Gibson assembly. The NphT7 gene was inserted by amplification of the gene encodings for NphT7 from pET28a-NphT7 using the primers pSET152-NphT7-PhaB F1/R1 containing a 40 bp overlap with pSET152-ErmEP* [16]. The PCR fragment was inserted into the NdeI-BamHI sites of pSET152-ErmEP* using the Gibson Protocol [16]. After confirming the insertion of the NphT7 sequence, a PCR fragment of the PhaB gene was inserted into the BamHI site using the Gibson Protocol [16]. PhaB was amplified from pET28a-PhaB using the primers pSET152-NphT7-PhaB F3/R3 containing a 27 bp and 40 bp overlap, respectively. pSET152-NphT7-PhaB F3 contained the same RBS sequences like the one found before NphT7 between the backbone overlap sequence and the PhaB overlap sequence.

Introduction of point mutations into DEBS plasmids using sequential Cas9-Gibson cloning. The *in vitro* Cas9 digest protocol was adapted from a previous study [19]. The guide RNAs (sgRNA) were purchased from Synthego (San Jose, CA) and Cas9 nuclease from *Streptococcus pyogenes* was purchased from the UC Berkeley MacroLab (Berkeley, CA). The sgRNA contained an 80 bp scaffold (Synthego EZ Scaffold) preceded by the desired N20 sequence for the appropriate protospacer adjacent motif (PAM) cleavage site. Prior to initiating the reaction, a 300 nM solution of the sgRNA was prepared with nuclease-free water on ice and the desired DEBS plasmid was diluted into nuclease-free water on ice (30 nM final concentration). The ribonucleoprotein (RNP) formation reaction was assembled at room temperature and the reagents were added in the following order; nuclease free water (18.5 µL), 4x Cas9 reaction buffer pH 6.5 (7.5 µL, contains 64 mM HEPES, 32 mM sodium chloride, 0.4 mM magnesium chloride), sgRNA (3 uL, 30 nM final concentration), and Cas9 (1 µL, 30 nM final concentration). The RNP formation

63

reaction was incubated for 10 min at 37ºC. The plasmid DNA (30 nM) was added and the reaction was incubated for 12 h at 37ºC. The reaction was halted by heat-inactivation for 10 min at 70ºC. Some constructs required an additional digest using a restriction enzyme. When this was required, the Cas9 treated DNA was purified using standard molecular biology protocols and subjected to a restriction digest.

In cases of a single digest, the repair arm was assembled from two 60 bp oligonucleotides containing the desired mutations. The oligonucleotides were mixed together in an equimolar amount (100 µM final concentration) and incubated at 95ºC for 3 min and then placed on ice for 5 min. The repair arm (1 µL) and linearized plasmid (10-50 ng) were used to assemble the DEBS plasmid containing the desired point mutation using the Gibson protocol [16]. For example, pET21c-KSAT3 (pAYC02 [20]) was modified to KSAT^{0}3 (S651A, based on EryAII numbering) using the sgRNA M3 AT^0 and the primers M3 rep F1/R1, using the protocol described above.

In the cases of a double digest, the repair arm was amplified via PCR and the construct was assembled via Gibson assembly. For example, pCK7 M6 AT^0 was assembled by amplifying M6 AT^0 from pAYC138 [15] using the primers 2p pCK7 M6 AT0 F1/R1 that contained a 60 bp overlap with pCK7 [21]. After treatment with DpnI, the fragment was inserted into pCK7 using the PAM cleavage site corresponding to the sgRNA M6 AT^0 and NsiI restriction site using the Gibson protocol [16].

Expression and purification of His-tagged and MBP-tagged proteins. All of the enzymes used in this chapter were initially expressed and purified as previously described (Section 2.2).

Fluoromalonate and methylmalonate. Fluoromalonate and methylmalonate were prepared as previously described (Section 2.2).

Fluoromalonyl-CoA. Fluoromalonyl-CoA was prepared enzymatically from fluoromalonate and CoA using MatB and ATP as previously described (Section 2.2)

N-acetylcysteamine thioester of (2S,3R)-2-methyl-3-hydroxypentanoic acid (NDK-SNAC) using PIKS KR1. NDK-SNAC was prepared and purified as previously described (Section 2.2).

N-acetylcystamine thioester of propionic acid (Propionyl-SNAC). Propionyl-SNAC was prepared and purified as previously described (Section 2.2).

Triketide lactone production using Mod3TE/AT^0 and DszAT variants under regenerative extender unit conditions. *In vitro* TKL production assays were performed as previously described (Section 2.2).

Triketide lactone production using Mod3$_{TE}$/ AT^0 and DszAT variants with pure extender units. All reactions contained 100 mM sodium phosphate (pH 7.5), TCEP (5 mM), magnesium chloride (10 mM), methylmalonyl-CoA epimerase (5 µM), NDK-SNAC (2.5 mM, unless otherwise stated), fluoro-, methyl-, or malonyl-CoA (1 mM, unless otherwise stated), Mod3$_{TE}$ AT^0 TE (10 µM), and DszAT (WT or variants) (10 µM). The reactions were incubated for 16 hours at 37°C and analyzed as previously described (section 2.2).

Triketide lactone production using Mod3$_{TE}$ AT^0 and *E. coli* cell lysates containing overexpressed DszAT mutant variants. DszAT lysate prepared by inoculating 50 mL of TB (containg 50 µg/mL Carbenicillin) with freshly transformed *E. coli* harboring pFW3 and its derivatives (OD$_{600}$ = 0.05). The cultures were grown to OD$_{600}$ of 0.6-0.8 at 37°C and placed in ice

water for 10-20 min and subsequently induced with 400 µM IPTG and grown overnight (~20 hours) at 16°C. 10 mL aliquots of each culture were pelleted by centrifugation at 12,000 × g for 5 minutes at 4°C. The cell pellets were then normalized by weight and resuspended in 1 mg/mL in sodium phosphate (290mM, pH 7.5). The cells were lysed with 0.1 mm bacterial glass disruption beads using Mini-BeadBeater-24 (Biospec products, Bartlesville, OK) in 2 × 45 s exposures at maximum speed. The resulting homogenate was spun down at 20,000 × g at 4°C for 20 minutes and the supernatant was used as the source of DszAT variants in the assay. The lysate assays were performed with the protocol used for TKL production with the using the CoA regeneration system, as described above. Specifically, the assay mixture contained equal volume of cell lysate and regeneration system (100 µL total). Production of TKLs and tetraketide lactone (TTKLs) were normalized relative to production observed with WT DszAT samples analyzed in conjunction variants tested. Lysate from a strain containing an empty vector, pET51b+ (Invitrogen, Waltham MA), was analyzed as a control for background endogenous transacylation from *E. coli* (BL21(de3)T1^R) in lysate. Products were analyzed as described below.

Identification of products using Time-of-Flight exact mass spectrometry. TKLs and TTKLs were analyzed on an EclipsePlus C-18 RRHD column (1.8 µm, 2.1 × 50 mm, r.t, Agilent) or a Poroshell 120 SB-Aq column (2.7 µm, 2.1 × 50 mm, r.t, Agilent) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). TKLs and TTKLs products were confirmed via Agilent 6500 Series Accurate-Mass Quadrupole Time-of-Flight mass spectrometry analysis. Expected and observed masses were analyzed using negative channel. Masses are represented corresponding to neutral masses minus a proton (**M-H$^-$**) and are presented as (expected mass, observed mass; ratio of parts per million difference to observed mass): H-TKL ($C_8H_{11}O_3$): 155.0714, 155.0712; 1.29 ppm, F-TKL ($C_8H_{10}FO_3$): 173.0619, 173.0618; 0.58 ppm, and H-TTKL ($C_{10}H_{13}O_4$): 197.0819, 197.0821; 1.02 ppm. Products could also be quantified using an Agilent G1315D diode array detector (TKL, $A_{260\ nm}$; F-TKL, $A_{260\ nm}$, H-TKL, $A_{260\ nm}$, H-TTKL, $A_{285\ nm}$). For absolute quantification, standard curves were generated using the authentic standards. H-TTKL was quantified using a standard curve with the commercially available 4-hydroxy-6-methyl-2-pyrone.

Quantification of TKLs using QQQ multiple reaction monitoring (MRM). TKLs were separated on an EclipsePlus C-18 RRHD column (1.8 µm, 2.1 × 50 mm, r.t, Agilent) or a Poroshell 120 SB-Aq column (2.7 µm, 2.1 × 50 mm, r.t, Agilent) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). TKLs were analyzed using an Agilent 6400 series Triple Quadrupole LC/MS mass spectrometer using multiple reaction monitoring (MRM). Transitions for the TKLs and TTKLs products were determined by screening for optimal fragmentation conditions. The transitions appear as follows (parent m/z → product m/z, fragmentation voltage, collision energy): H-TKL: 155 → 97, 100, 5; F-TKL: 173 → 59, 135, 20; H-TTKL: 197 → 95, 135, 5; TKL: 169 → 111, 135, 10.

Triketide lactone production using propionyl-SNAC with the Mod2 + Mod3$_{TE}$ mini-PKS system. All reactions contained 100 mM sodium phosphate (pH 7.5), magnesium chloride (5 mM), TCEP (5 mM), methylmalonyl-CoA epimerase (5 µM), methylmalonyl-CoA (1 mM), malonyl-CoA (1 mM) and fluoromalonyl-CoA (1 mM), propionyl-SNAC (5 mM), and reduced nicotinamide adenine dinucleotide phosphate (NADPH; 5 mM). Reactions were initiated by

addition of AT^0 Mod2 (10 µM), WT Mod3$_{TE}$ (10 µM), and DszAT (10 µM) and incubated overnight at 37ºC. Aliquots were removed, quenched, and processed as previously described.

Pyrone production in *E. coli* resting cells. The following protocol was adapted from a previously established method [*1*]. LB (50 mL) containing glucose (2% *v/v*), Carbenicillin (50 µg/mL), Spectinomycin (50 µg/mL), and Chloramphenicol (25 µg/mL) in a 125 mL baffled shake flask was inoculated with 1 mL of an overnight LB culture (50 mL) of *E.coli* BAP1 freshly co-transformed with pBAD18-Cm.500matC, pAYC136 (AT^0 Mod3$_{TE}$), and pCDFDuet-DszAT-F190V-rpMatB-T103A or pCDFDuet1-DszAT.RpMatB-T103A. The cultures were grown at 37°C at 200 rpm to OD600 = 0.8-0.9, at which point cultures were cooled on ice for 10 min. Cultures were induced for protein expression with IPTG (1 mM final concentration) and L-arabinose (0.2% final concentration). The cultures were left shaking at r.t for 20-24 h following induction, shaking at 250 rpm on a benchtop shaker. Cells were collected by centrifugation 10414 × *g* at 4°C for 5 min. The cells were washed once with 100 mM potassium phosphate, pH 7.4, then resuspended in the same buffer at an OD$_{600}$ of 90-110. The reactions were set up in 0.6 mL tubes (50 µL cell suspension) and contained fluoromalonate (5 mM), malonate (1 mM), and methylmalonate (1 mM), and oxo-NDK-SNAC (stock, 100 mM solution in 10% DMSO; final, 7 mM). The cell suspensions were incubated with shaking at r.t for 22 h. The culture supernatant was collected by centrifugation at 18,000 × *g* for 10 min and analyzed on a Poroshell 120 SB-Aq column (2.7 µm, 2.1 × 50 mm, r.t, Agilent) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). Products were confirmed on an Agilent 6500 Series Accurate-Mass Quadrupole Time-of-Flight mass spectrometer. **ESI-MS [M+H]$^+$:** calculated for the fluorinated pyrone $C_8H_{10}FO_3$, 173.0608 *m/z*, found 173.0611 *m/z*. **ESI-MS [M+H]$^+$:** calculated for the desmethylated pyrone $C_8H_{11}O_3$, 155.0703 *m/z*, found 155.0702 *m/z*. **ESI-MS [M+H]$^+$:** calculated for the methylated pyrone $C_9H_{13}O_3$, 169.0859 *m/z*, found 169.0861 *m/z*.

Preparation of spore stocks for *Streptomces lividans*, *Streptomyces venezulae*, and *Actinomadura vulgaris*. A 4 mL mycelial suspension was grown in 50 mL culture tubes (with 4-10 glass beads) from either spores stocks (*S. lividans*, K4-114), frozen mycelial suspension (*A. vulgaris*, ATCC 53715), or lyophilized stock (*S. venezuelae*, ATCC 15439) in either nutrient broth (EMD Millipore, Billerica MA) (*S. lividans* and *A. vulgaris*) or GYM (4 g/L glucose, 4 g/L yeast extract, 10 g/L malt extract) (*S. venezuelae*) for either 48 h (*S. lividans* and *A. vulgaris*) or 24 h (*S. venezuelae*) at 30ºC. 100-200 µL of the mycelial suspensions were beaded on either soy-mannitol plates (20 g/L agar, 20 g/L mannitol, 20 g/L soy flour) (*S. lividans* and *A. vulgaris*) or GYM plates (*S. venezuelae*) and incubated at 30ºC for 5 days (*S. lividans* and *S. venezuelae*) or 21-30 days (*A. vulgaris*). Spores were wet-scraped into a 50 mL conical tube and pelleted by centrifugation (1160 × *g*, 30 min). The supernatant was decanted and the pellet was resuspended in 25% glycerol. Spores stocks were aliquoted into 25-100 µL and kept at -80ºC until use.

Preparation of *Streptomyces lividans* protoplasts. The protocol for the preparation of protoplasts was adapted from a previously established method [*22*]. An *S. lividans* (K4-114) spore stock (25 µL) was added to 25 mL of YEME media (3 g/L yeast extract, 5 g/L peptone, 3 g/L malt extract, 10 g/L glucose, 340 g/L sucrose, 5 mM magnesium chloride hexahydrate, 0.5% (*w/v*) glycine) in 125 mL baffled flask and left shaking at 200 rpm for 40 h at 30ºC. The cells were harvested by centrifugation at 1160 × *g* for 10 min in a 50 mL conical tube. Half of the volume of the supernatant was carefully decanted and 15 mL of 10.3% (*w/v*) sucrose solution (filter sterilized) were added. The sample was subjected to centrifugation at 1160 × *g* for 10 min. The entire volume

was decanted and the pellet was washed with another 15 mL of 10.3% (*w/v*) sucrose solution (filter sterilized). The sample was centrifuged at 1160 × g for 10 min and the supernatant was decanted. The pellet (mycelium) was suspended in 4 mL of P buffer (103 g/L sucrose, 0.25 g/L potassium sulfate, 2.02 g/L magnesium chloride hexahydrate, 0.2% (*v/v*) trace element solution [*22*], 0.05% (*w/v*) monopotassium phosphate, 3.68 % (*w/v*) calcium chloride dihydrate, 5.73% (*w/v*) TES buffer (pH 7.2)) containing 2 mg/mL lysozyme and incubated for 40 min at 30ºC while shaking (200 rpm). After addition of 5 mL P buffer, the protoplasts were freed from the mycelium by gently pipetting with a 5 mL pipette. The protoplasts were filtered through a 10 mL syringe containing a cotton ball into a 15 mL conical tube. The protoplasts were sedimented by gentle centrifugation at 1000 × g for 7 min. The supernatant was removed gently and the sedimented protoplasts were lightly resuspended in 1-2 mL of P Buffer. The protoplasts were aliquoted (20-100 µL) and left on ice for 20 min prior to being placed in -80ºC for storage.

***Streptomyces* protoplast transformation.** The protocol for protoplast transformations was adapted from a previously adapted method [*22*]. ~1 µg (in up to 20 µL of EB or TE buffer) of plasmid was added to a 15 mL conical tube. An aliquot of *S. lividans* (K4-114) (25-100 µL) protoplasts was added along with 0.5 mL of PEG 1000 (25% (*w/v*), microwave sterilized). After addition of 6 mL of P Buffer, the sample was quickly centrifuged for 7 min at 700 × g and decanted carefully. The sample left in the tube (100-200 µL) was plated on R5 plates (103 g/L sucrose, 0.25 g/L potassium phosphate, 10.12 g/L magnesium chloride hexahydrate, 10 g/L glucose, 0.1 g/L casaminoacids, 0.2% (*v/v*) trace element solution [*22*], 5 g/L yeast extract, 5.73 g/L TES buffer, 0.05% (*w/v*) monopotassium phosphate, 0.2 M calcium chloride dehydrate, 3% (*w/v*) L-proline, 0.07 N sodium hydroxide, 22 g/L agar) and left at 30ºC for 16 h for recovery. After 16 h, thiostrepton (10% DMSO) was overlayed on plates (~10 µg/mL final suspension concentration) and the plates were left incubating at 30ºC for a 7-14 d (colonies visible as soon as 5 d). In the case of pCK7 [*21*], the transformation was confirmed by observing the production of 6-dEB.

Production of 6-deoxyerythronolide B in *Streptomycese lividans*. An aliquot of spores containing *S. lividans* (K4-114) transformed with pCK7 [*21*] (25 µL) was added to 25 mL of nutrient broth in a 125 mL baffled flask with 8-20 glass beads. Thiostrepton (10 µg/mL final concentration, DMSO) and sodium propionate (5 mM final concentration) were added and the culture was incubated under shaking at 200 rpm for 30ºC. After 3-5 d, ~1 mL of the media was removed and centrifuged at 18,000 × g for 5 min. The supernatant was removed and filtered (2 µm) and analyzed for production using a Poroshell 120 SB-Aq column (2.7 µm, 2.1 × 50 mm, r.t, Agilent) using a linear gradient from 5 to 45% acetonitrile over 3.2 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 5% acetonitrile for 1 min (0.6 mL/min). This was followed by a 45 to 100% acetonitrile gradient over 0.8 min. 6-dEB was identified using Agilent 6500 Series Accurate-Mass Quadrupole Time-of-Flight mass spectrometry. **ESI-MS [M+H-H$_2$O]$^+$:** calculated for C$_{21}$H$_{37}$O$_5$, 369.2636 *m/z*, found 369.2636 *m/z*.

Production of monofluorinated Fluvirucin B$_1$ in *A. vulgaris* from fluoroacetate, fluoromalonate, or diethylfluoromalonate. An *A. vulgaris* spore stock (25 µL) was added to 25 mL of nutrient broth in a 125 mL baffled shake flask containing 8-20 glass beads. Either fluoroacetate (5 mM), fluoromalonate (5 mM), or diethylfluoromalonate (5 mM) was added and the culture was incubated under shaking (200 rpm) at 30ºC. After 10 d, ~1 mL of the media was removed and centrifuged at 18,000 × g for 5 min. The supernatant was removed and filtered (2 µm) and analyzed for production using a Poroshell 120 SB-Aq column (2.7 µm, 2.1 × 50 mm, r.t,

Agilent) using a linear gradient from 5 to 45% acetonitrile over 3.2 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 5% acetonitrile for 1 min (0.6 mL/min). This was followed by a 45 to 100% acetonitrile gradient over 0.8 min. Fluvirucin B_1 derivatives were identified using Agilent 6500 Series Accurate-Mass Quadrupole Time-of-Flight mass spectrometry. Fluvirucin B_1 **ESI-MS [M+H]$^+$:** calculated for $C_{24}H_{47}N_2O_5$, 443.3479 *m/z*, found 443.3481 *m/z*. Monofluorinated Fluvirucin B_1 derivatives; **ESI-MS [M+H]$^+$:** calculated for $C_{24}H_{46}FN_2O_5$, 461.3385 *m/z*, found 461.3384 and 461.3388 *m/z*.

The same protocol was carried out for the production of fluorinated pikromycin with one change; GYM as well as nutrient broth were used as production media. Fluorinated pikromycin was not observed in these experiments. Pikromycin was observed in the presence and absence of fluorinated building blocks **ESI-MS [M+H]$^+$:** calculated for $C_{28}H_{48}NO_8$, 526.3374 *m/z*, found 526.3372 *m/z*.

3.3   Results and Discussion

Assembly and screening of DszAT mutant library.   The main driving force behind the design of the mutagenesis screens was to find variants with a reduced level of malonyl-CoA incorporation efficiency and little to no reduction in the ability to incorporate fluoromalonyl-CoA. Mutagenesis sites were informed by the crystal structure of DszAT (*Figure 3.1*) [*15*]. The initial target residue was F190, since it is a highly conserved residue in malonyl-CoA selective ATs, and is presumed to be the major contributor for selectivity towards malonyl-CoA over larger substrates to [*23*]. Additionally, mutations where made at the leucine at position 87. This was done in an attempt to destabilize formation of the acyl-enzyme intermediate, reasoning that the α-fluoro group could help stabilize formation of the acyl-enzyme intermediate in the absence of a strong oxyanion hole. Such a mutation may create a larger differential activity between

Figure 3.1: DszAT active site as shown in the crystal structure obtained in [14]. PDB 3RGI.

fluoromalonyl-CoA and malonyl-CoA, which has a higher requirement for oxyanion hole catalysis. Finally, the double mutants containing an H191A mutation and mutations to the catalytic serine (S86) were inspired by results observed from studies using the fluoroacetyl-CoA thioesterase (FlK) [*10,24*]. Specifically, in FlK a histidine mutation to the catalytic dyad (much like H191), resulted in formation of an acyl-enzyme intermediate (fluoroacetyl-glutamate) that was very slow to hydrolyze. Given the hydrolytic activity of DszAT towards fluoromalonyl-CoA [*2*], it is reasonable to attempt to limit hydrolysis by mirroring the FlK active histidine mutation. Additionally, the substitution of the catalytic serine with more nucleophilic residues was implemented in order to compensate for the absence of the general base (H191).

The DszAT variants were cloned using the backbone pFW3 (pET21 vector containing a C-terminal 6x His tag and a T7 promoter) [*15*]. To avoid amplifying and sequencing long stretches of DNA from the plasmid backbone, the mutations were introduced using three-piece Gibson

A

Lyse cells → MatB, ATP+CoASH → Analyze by MS

B
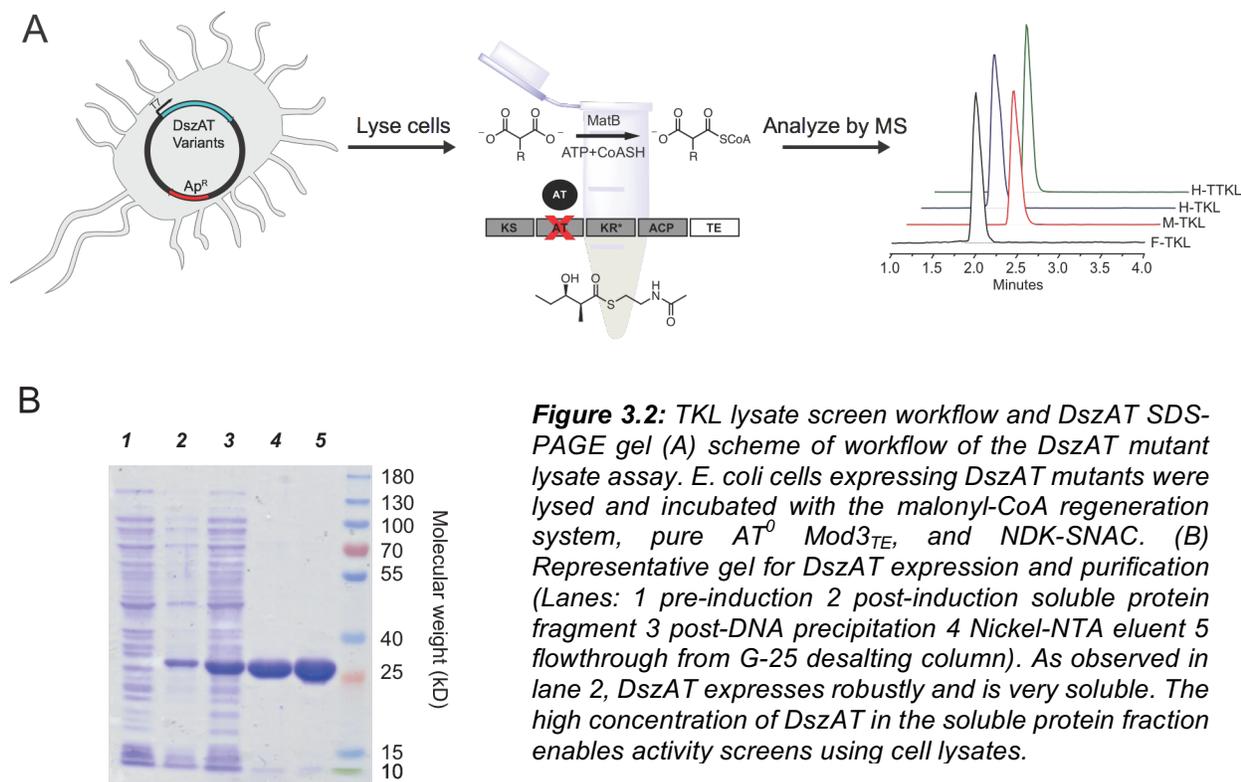
Figure 3.2: *TKL lysate screen workflow and DszAT SDS-PAGE gel (A) scheme of workflow of the DszAT mutant lysate assay. E. coli cells expressing DszAT mutants were lysed and incubated with the malonyl-CoA regeneration system, pure AT^0 $Mod3_{TE}$, and NDK-SNAC. (B) Representative gel for DszAT expression and purification (Lanes: 1 pre-induction 2 post-induction soluble protein fragment 3 post-DNA precipitation 4 Nickel-NTA eluent 5 flowthrough from G-25 desalting column). As observed in lane 2, DszAT expresses robustly and is very soluble. The high concentration of DszAT in the soluble protein fraction enables activity screens using cell lysates.*

reactions with overlapping primers containing the desired mutations [*15*]. These libraries were then transformed into *E. coli* (BL21(de3)T1[R]) and screened for activity using the TKL production experiment described previously using AT^0 $Mod3_{TE}$ and NDK-SNAC (Section 2.3). Cell cultures containing overexpressed DszAT were lysed and screened for C-C bond formation activity using purified AT^0 $Mod3_{TE}$, NDK-SNAC, and extender units formed under regenerative conditions (*Figure 3.2A*). As shown in *Figure 3.2B*, DszAT expression is robust, allowing activity to be screened in cell lysates without further purification. We chose to use the TKL formation assay to measure mutant activity because it affords us the opportunity to determine both AT acylation as well as transacylation activity in the context of product formation. While this assay could not provide us with AT acylation, AT hydrolysis, or ACP transacylation efficiencies directly, it offered an assay for rapidly measuring fluorine incorporation in a well-studied system [*2*]. As such, we felt confident that the enhanced ratio of F-TKL to malonyl-derived products represents sufficient information for screening of fluorine selective DszAT variants. As referenced in *Figure 3.2A*, the products monitored in these reactions were TKL, F-TKL, H-TKL, and H-TTKL (a tetraketide pyrone resulting from two chain extensions, *Figure 3.3A*). It is important to mention that after the completion of these screening experiments, we discovered that H-TKL is unstable in the reaction buffer ($t_{1/2} \sim 13$ h) (*Figure 3.4*), making H-TTKL appear to be the major product. However, the initial rate of H-TKL formation is 1.1 min^{-1} [*2*] compared to H-TTKL formation which is 0.08 min^{-1} (*Figure 3.3B*). The H-TKL decomposition products and mechanism remain unknown, but are most likely attributed to hydrolysis and subsequent decarboxylation, as well as a potential rearrangement. Since this decomposition occurs when H-TKL is incubated even in only phosphate buffer, it is reasonable to assume that this decomposition occurs at a pseudo-first order or zero-order rate model. Therefore, comparing H-TKL product formation remains a good indicator of

A

B
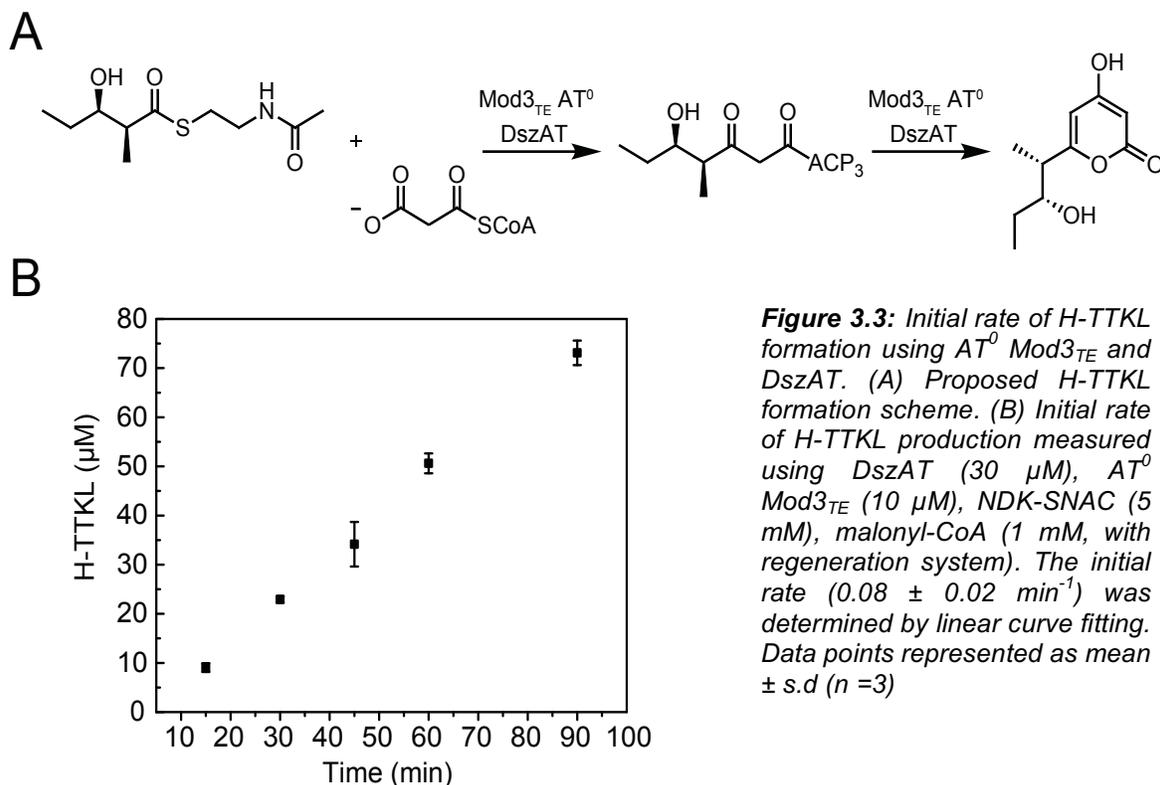
Figure 3.3: Initial rate of H-TTKL formation using AT^0 $Mod3_{TE}$ and DszAT. (A) Proposed H-TTKL formation scheme. (B) Initial rate of H-TTKL production measured using DszAT (30 µM), AT^0 $Mod3_{TE}$ (10 µM), NDK-SNAC (5 mM), malonyl-CoA (1 mM, with regeneration system). The initial rate (0.08 ± 0.02 min^{-1}) was determined by linear curve fitting. Data points represented as mean ± s.d (n =3)

DszAT mutant activity, since all the samples had the same incubation time and are always normalized to WT DszAT production.

We found that 53% of the F190 and F190/L87A mutants screened displayed equal or greater fluoromalonyl-CoA incorporation yields, suggesting that the high hit rate in the initial screen can be attributed to the activated fluoromalonyl-CoA substrate that enables catalysis even in a compromised active site (*Figure 3.5*). The catalytic residue mutations of the S86 and S86/H191A mutant library did not display *trans*-AT activity for any substrate, which is not surprising as these two residues are crucial for catalysis [*23*]. The ability to efficiently load fluoromalonyl-CoA in an active site that is inefficient at loading malonyl-CoA is not entirely surprising when considering the active site architecture of the only known enzyme whose native activity results in formation of an acyl-enzyme intermediate (FlK) with a fluorinated substrate (fluoroacetyl-CoA). When the active-site of FlK is compared to other hotdog fold thioesterases in its class [*10,24,25*], it becomes apparent that the oxyanion hole in FlK is weaker than its homologs [*10,24,25*]. Although FlK and DszAT are a different class of proteins, they both catalyze reactions that are proposed to form acyl-enzyme intermediates through an addition-elimination reaction and via formation of tetrahedral intermediate stabilized by an oxyanion hole. This suggests that perhaps the high hit rate in the DszAT mutant screen can be partially attributed to the activated nature of fluoromalonyl-CoA, which results in less of a need for oxyanion hole stabilization.

While the high success rate of this library could be related to reactivity of an extender unit containing α-fluoro moiety, this library can still be used to screen for DszAT variants selective for other non-natural extender units. The mutations made to DszAT in this discussion are not dissimilar from mutations made to DEBS *cis*-ATs, resulting in the selective incorporation of ethyl, propargyl, allyl, and azide containing extender units into the backbone of a DEBS derived

polyketide product [*11-13*]. Thus, new DszAT activities could potentially be found within these libraries.

Secondary screen of DszAT variants using purified protein. The initial lysate assay screens resulted in a few variants that showed promise as fluorine-selective enzymes. The F190G, F190T, F190V, F190P-L87A, and F190I-L87A mutants (*Figure 3.5*) were selected, expressed and purified. Once again, these variants were screened using the TKL formation assay and H-TKL, F-TKL, H-TTKL formation was normalized to production observed by WT DszAT. The TKL formation assay was carried out with pure-CoA substrates (*Figure 3.6*) in order to simplify extender unit stoichiometry. As observed in *Figure 3.6*, the F190V variant appears to produce more F-TKL than any other variant tested. Additionally, we observe that F190V DszAT exhibits the lowest level of production of the malonyl-derived products. It was also important to test for the methylmalonyl-CoA selectivity of these variants, as methylmalonyl-CoA will be present at high concentrations *in vivo*. The methylmalonyl-CoA incorporation observed F190V DszAT remained as low as that observed when WT DszAT was used (*Figure 3.6*) and is likely due to direct diffusion of the extender unit in the Mod3$_{TE}$ KS active site [*2*].

Figure 3.4: HPLC-UV (A_{254}) trace of a chemically synthesized H-TKL standard. After 16 h, decomposition of this product is observed and a second unknown peak appears. The mechanism of H-TKL decomposition remains unknown. Traces are normalized to H-TKL standard quenched after 1 min.

We also observed that the F190G variant possessed improved methylmalonyl-CoA incorporation activity compared to the other variants tested, while producing F-TKL and malonyl-derived products at levels similar to WT DszAT, suggesting that this mutation creates a more promiscuous DszAT variant (*Figure 3.6*).

Ultimately, the secondary screen with purified DszAT variants resulted in the identification the F190V mutant that shows promise as a fluoromalonyl-CoA selective *trans*-AT. The remaining questions regarding the efficacy of this variant resides in the scope of fluorine selectivity under competitive conditions, as well as in the context of production of more complex polyketide using multiple modules.

***In vitro* analysis and site-specific fluorine incorporation of the DszAT F190V mutant.** The identification of F190V DszAT provided an exciting opportunity for the use of a DszAT mutant as a method for fluorine incorporation *in vivo*. As shown in *Figure 3.7*, DszAT F190V displayed 41-fold improvement in fluorine selectivity compared to WT DszAT with an 11-fold higher yield of F-TKL compared to malonyl-CoA-derived products. The data in *Figure 3.7* shows that the nature of fluorine selectivity in the F910V mutant is driven by the selectivity against malonyl-CoA, as opposed improved levels of fluorine incorporation. This result is consistent with the initial hypothesis that the DszAT active site mutations have more dramatic effects on malonyl-CoA-based acyl-enzyme intermediates than the more activated fluoromalonyl-AT intermediates.
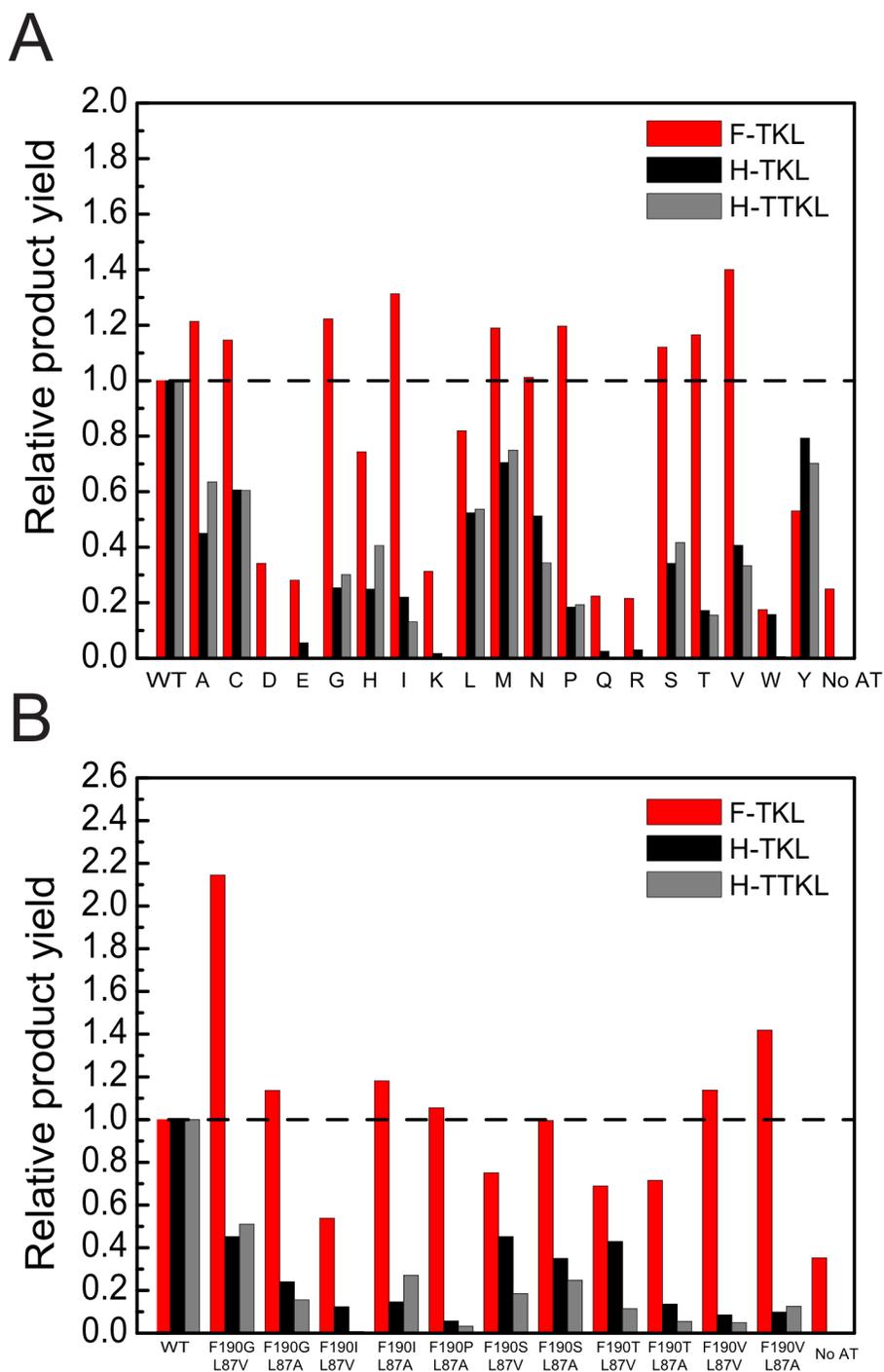
Figure 3.5: *F-TKL, H-TKL, and H-TTKL production of DszAT mutant library cell lysates with AT[0] Mod3$_{TE}$. C-C bond products formed in DszAT mutant screens are normalized to the production of F-TKL (red), H-TKL (black), H-TTKL (grey) by AT[0] Mod3$_{TE}$ and E. coli lysate containing WT DszAT (A) Production of C-C bond products using the F190 saturation mutagenesis library. (B) Production of C-C bond products using the oxyanion hole mutagenesis library.*
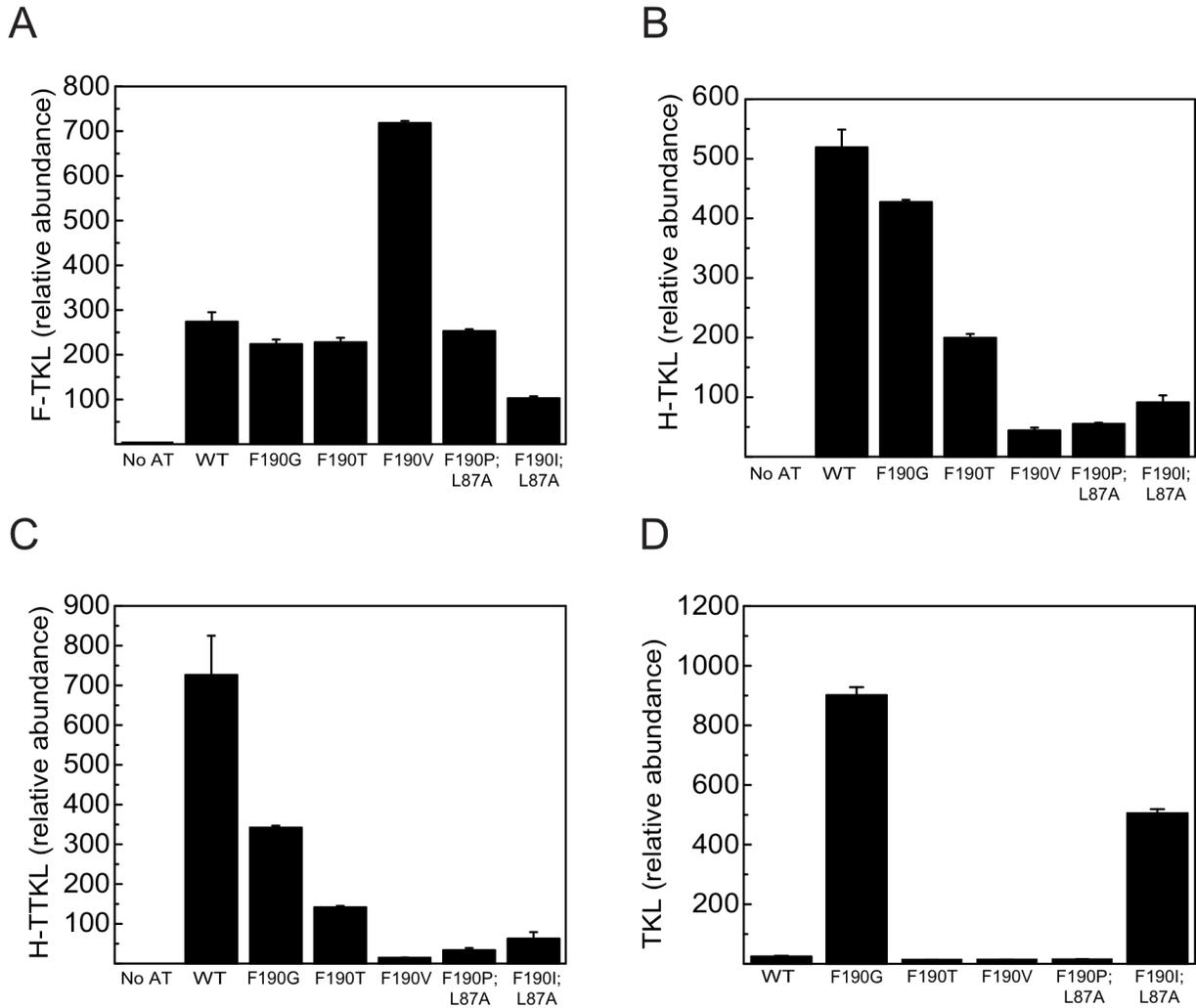
Figure 3.6: *F-TKL, H-TKL, H-TTKL, and TKL production using AT^0 Mod3$_{TE}$ and purified DszAT variants. Production of (A) F-TKL, (B) H-TKL, (C) H-TTKL, and (D) TKL with purified DszAT variants (10 µM), AT^0 Mod3$_{TE}$ (10 µM), fluoro-, methyl-, or malonyl-CoA (1 mM), and NDK-SNAC (2.5 mM). Data are shown as mean ± s.e (n=2) for F-TKL, H-TKL and H-TTKL and mean ± s.e (n=3) for TKL.*
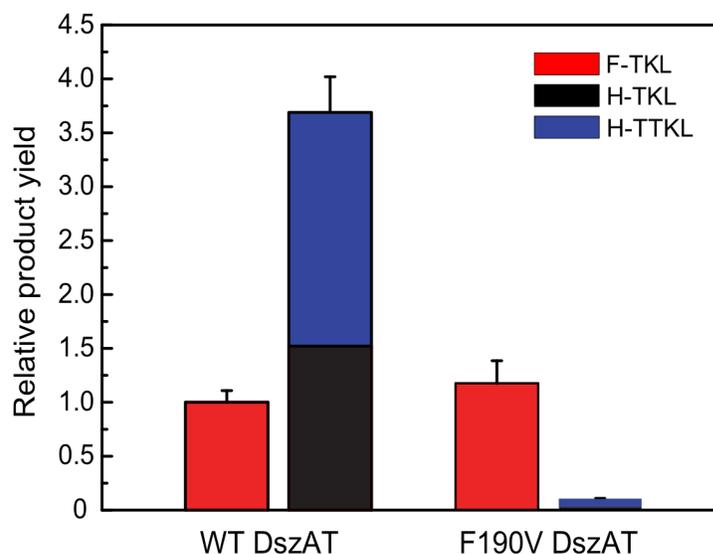
Figure 3.7: *Relative product yield of complementation of AT^0 Mod3$_{TE}$ with the F190V DszAT variant. Normalized representation of the relative yields of C-C bond products (F-TKL, H-TKL, and H-TTKL) formed in assays containing AT^0 Mod3$_{TE}$ (10 µM), DszAT (10 µM), fluoro- or malonyl-CoA (1 mM, regeneration system), and NDK-SNAC (2.5 mM). Data are represented average ± s.e. (propagated from the s.d. (n=3) of the individual products).*

Given that malonyl-CoA will be present in cells during the *in vivo* production experiments, it was important to determine the ability of both the WT and F190V DszAT variant to produce F-TKL under competitive conditions. When the TKL formation assay was carried out under competitive conditions with varying malonyl-CoA (0.25-5 mM) and a fixed concentration of fluoromalonyl-CoA (0.5 mM), the production of F-TKL by WT DszAT was poor even in the presence of 2-fold excess fluoromalonyl-CoA (*Figure 3.8A*). Additionally, as the ratio of malonyl-CoA to fluoromnalonyl-CoA increased, the formation malonyl-CoA derived products increased as well. However, when F190V DszAT was subjected to this assay, we observed that the formation of malonyl-derived product was quite poor (up to ~25-fold worse than observed in the WT DszAT experiment, *Figure 3.8*). More importantly, as the ratio of malonyl-CoA to fluoromalonyl-CoA increased, the production of malonyl-CoA derived products remained unaffected (*Figure 3.8B*). As expected, the F-TKL yields observed in the DszAT F190V assay were also higher than those observed in the WT DszAT competition experiment. While the F-TKL yields went down as malonyl-CoA levels were increased, it is unlikely that malonyl-CoA levels will be ever be more than equivalent to fluoromalonyl-CoA levels *in vivo*, if proper unnatural extender unit feeding protocols are followed [*6,12,13,26*].

The TKL formation assay is a powerful tool for screening and confirming fluorine-selective DszAT variants. However, in order to better assess the viability of this mutant in the context of full-length product formation, the more complex mini-PKS system provides a more complete view [*1,2*]. In this system, we can monitor the regioselectivity of fluorine incorporation in the presence of two modules (*Figure 3.9A*). The mini-PKS also allows us to test for the formation TKL in the presence of methylmalonyl-CoA, along with the addition of fluoro- and malonyl-CoA. AT^0 Mod2 and WT Mod3$_{TE}$ were complemented with either WT DszAT or F190V DszAT, along with equimolar methyl-, fluoro-, or malonyl-CoA (1 mM) (*Figure 3.9A*). As previously observed [*2*], only the expected regioisomer was detected (FMe, *Figure 3.9B*) when both DszAT variants are used. The almost identical yield of this regioisomer among the two DszAT enzymes is expected
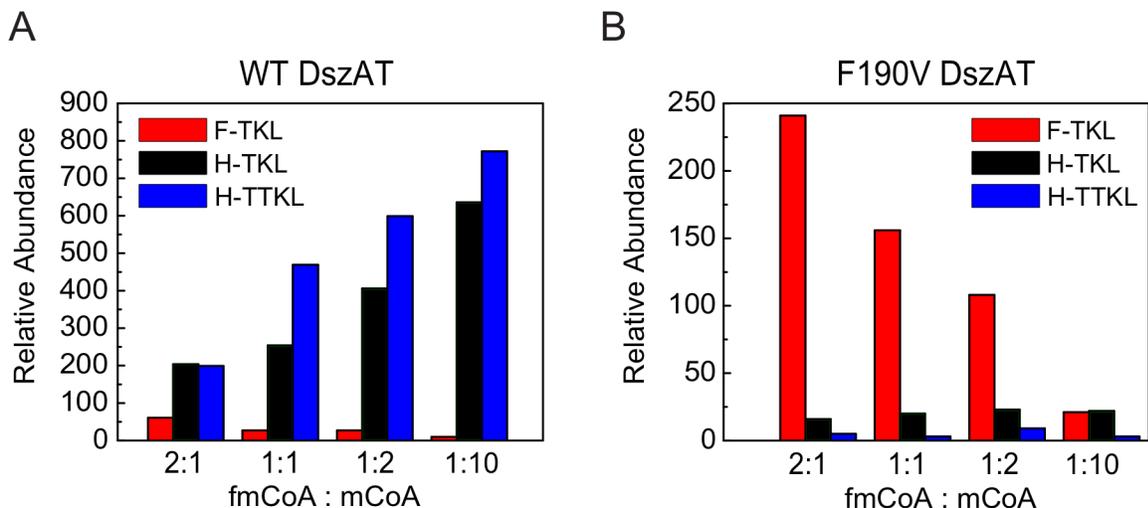
Figure 3.8: *Fluoro- and malonyl-CoA competition assay with DszAT variants. Competition assay with fixed fluoromalonyl-CoA concentration (0.5 mM) and varying malonyl-CoA concentrations (0.25-5 mM). The reaction mixture contained AT^0 $Mod3_{TE}$ (10 µM), DszAT (10 µM), and NDK-SNAC (2.5 mM). (A) Graph representing the competition experiment using WT DszAT. (B) Graph representing the competition experiment using F190V DszAT.*

based on the single chain extension experiments described above. It was exciting to see that the formation of the expected H-TKL (HMe, *Figure 3.9B*) was 31-fold lower in the presence of F190V DszAT compared to WT DszAT. This result suggests that when multiple PKS modules are required for fluorinated product formation, complementation of F190V DszAT with a module containing an inactive *cis*-AT can produce regiospecifically fluorinated products, even under competitive conditions.

Production of polyketides in *E. coli*. With the F190V DszAT mutant in hand, we then focused our attention towards the production fluorinated polyketides *in vivo*. In order to simplify analysis of the products formed, we decided to use the β-keto analog of NDK-SNAC (oxo-NDK-SNAC, *Figure 3.10A*), which forms the more stable triketide pyrone, as opposed to the triketide lactones (*Figure 3.10A*). In order to facilitate fluoromalonyl-CoA incorporation, a malonate transporter (MatC) was expressed [26]. *E. coli* resting cells containing plasmids expressing DszAT (WT or F190V), AT^0 $Mod3_{TE}$, MatC, and rpMatB were incubated with oxo-NDK-SNAC (7 mM) and with fluoromalonate (5 mM), malonate (1 mM, when used), and methylmalonate (1 mM, when used). As shown in *Figure 3.10B*, when resting cell cultures are incubated with only fluoromalonate and oxo-NDK-SNAC, the major product (by MS integration) for resting cells containing either WT or F190V DszAT is the fluorinated pyrone. However, the F190V mutant produces 2-fold less of the desmethylated pyrone than cultures harboring the WT DszAT variant. While this difference is not as dramatic as that observed in the *in vitro* experiments (*Figure 3.7*), it is still encouraging to see a production difference given the increased complexity of this assay mixture and factors such as cellular extender unit availability and activation. The assay ambiguity is further elaborated when malonate and methylmalonate are added to the cultures, as cultures with the two DszAT variants do not display statistically significant production levels. This provides further evidence that intercellular monomer availability and activation may play a large role in the observed product yields. It was also surprising to see large quantities of the methylated pyrone
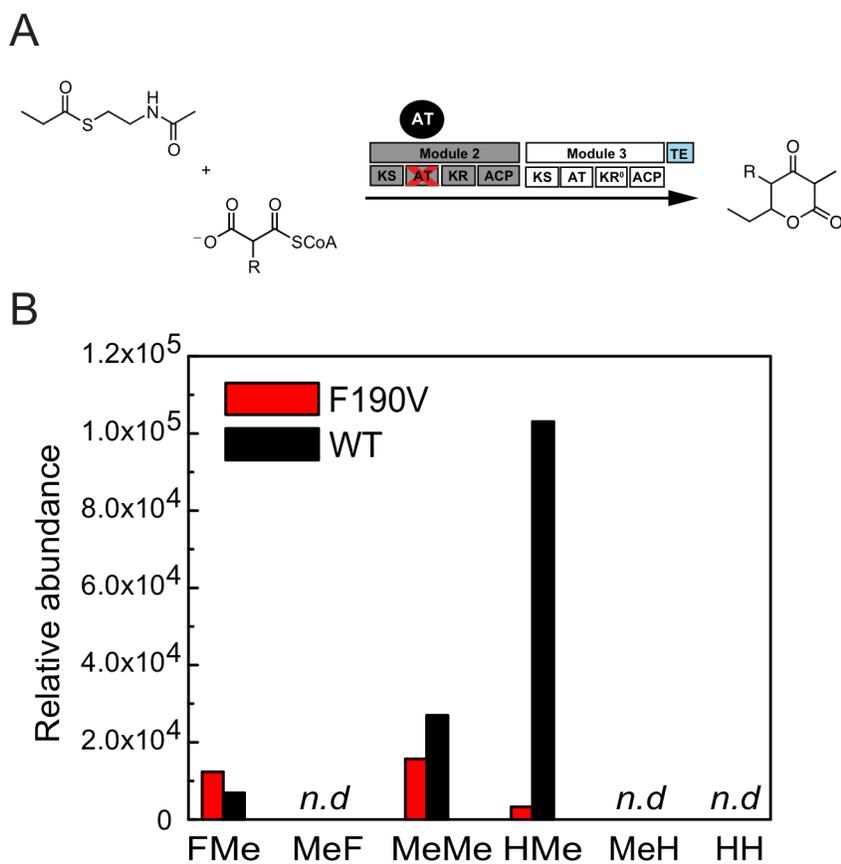
Figure 3.9: *Production using mini-PKS and DszAT variants. (A) scheme representing formation of TKLs with the mini-PKS. Reactions contained AT^0 Mod2 (10 µM), WT $Mod3_{TE}$ (10 µM), DszAT (10 µM), propionyl-SNAC (5 mM), methylmalonyl-CoA (1 mM), fluoromalonyl-CoA (1 mM), and malonyl-CoA (1 mM). (B) production of the various TKLs formed with either WT or F190V DszAT.*

formed, even in the absence of methylmalonyl-CoA, especially since previous resting cell production experiments using NDK-SNAC resulted in very low to undetectable yields of TKL [*1*]. However, we propose that these high yields methylated pyrone are attributed to decomposition of oxo-NDK-SNAC into two propionate units, which can be activated to form methylmalonyl-CoA in the cell.

Upon completion of these assays, we found that the construct containing DszAT contained the point mutation T103A (pCDFDuet1-DszAt.RpMatB-T103A). This mutation was also carried over to the F190V DszAT construct in the cloning process. While this residue is not directly involved in catalysis, it may have consequences to enzyme activity. Constructs containing the proper sequence have been cloned and their sequence was confirmed (Appendix 1). Additionally, an AT^0 variant of DszAT (S86A) was cloned as an added control. The resting cell production experiments will be repeated with the corrected constructs, along with experiments designed to specifically investigate the intercellular fluoromalonyl-CoA pool. The production of pyrones in *E. coli* lays the groundwork for future *in vivo* production experiments in both *E. coli* and anctinomycetes.

Towards the production of site-specifically fluorinated 6-dEB *in vivo*. Although the production of 6-dEB can be achieved through *E. coli* fermentation [*27*], the production of polyketides in a heterologous *Streptomyces* host remains a more viable option [*28,29*]. The Khosla lab at Stanford University gifted us with a *Streptomyces* expression vector (pCK7 [*20*]) containing
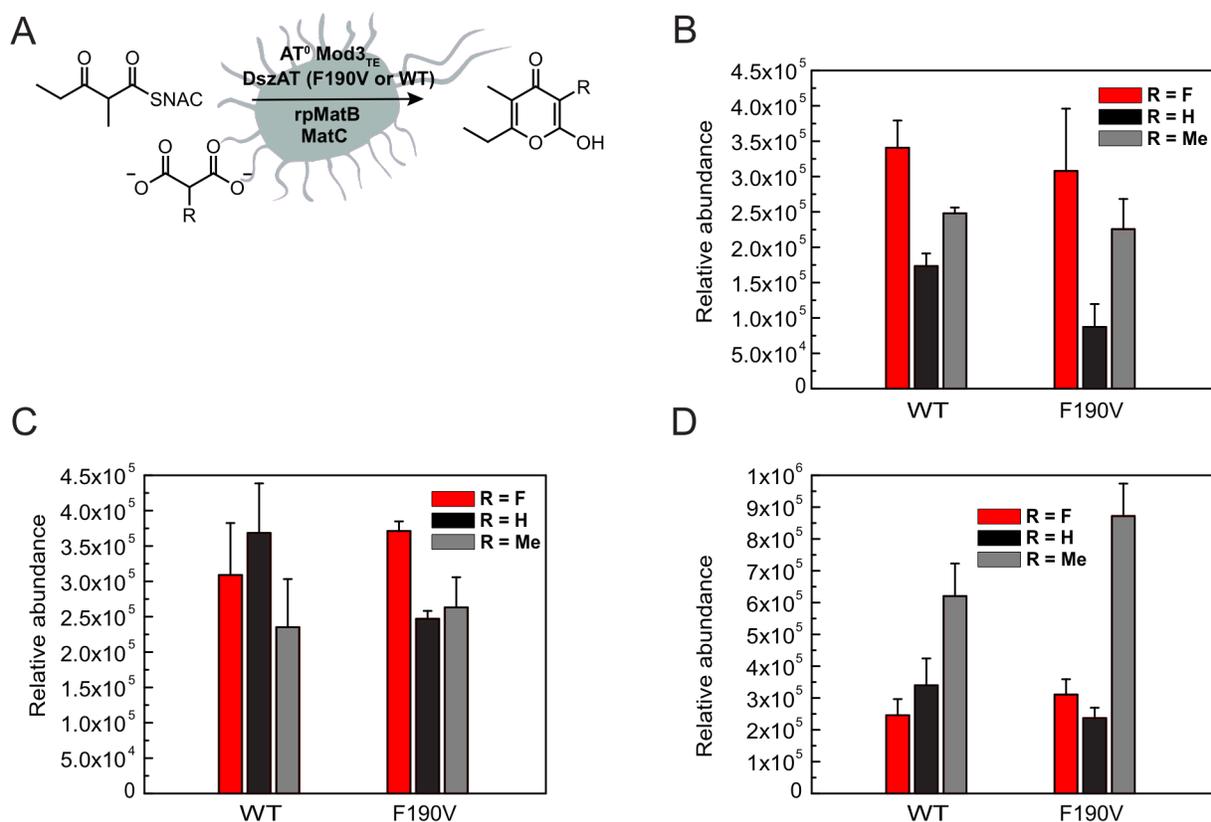
Figure 3.10: *In vivo production of triketide pyrones. (A) Scheme representing the substrates and enzymes used for in vivo production. (B) Production using fluoromalonate (5 mM) and oxo-NDK-SNAC (7 mM). (C) Production using fluoromalonate (5 mM), malonate (1 mM) and oxo-NDK-SNAC (7 mM). (D) Production using fluoromalonate (5 mM), malonate (1 mM), methylmalonate (1 mM) and oxo-NDK-SNAC (7 mM).*

the entire DEBS PKS cluster required for the production of 6-dEB. pCK7 contains a large *Streptomyces* origin (SCP2*, vector copy number 1-5 [30]) and a thiostrepton resistance cassette. We chose to use *S. lividans* (K4-114) as the host for 6-dEB production due to the dearth of literature precedence for *S. lividans* as a heterologous host for natural product biosynthesis [28,29]. Additionally, K4-114 has been used for the production of 6-dEB analogs in the past [31], and should provide a good starting point for full length polyketide production. In order to carry out our strategy for site-specific fluorine incorporation, a pCK7 plasmid containing an AT^0 mutation must be transformed into *S. lividans*. DszAT can then be integrated into the *S. lividans* genome through a bacterial conjugation [17,18] (*Scheme 3.3*). Based on *in vitro* studies with the F190V DszAT variant, the endogenous presence of malonyl-CoA should not impede fluoromalonyl-CoA incorporation into the module containing an AT^0 mutation. To this end, vectors encoding for either WT or F190V DszAT (pSET152-ErmEP*-DszAT-MatB) were assembled. The vector contained an additional ribosome binding site and the gene for *Rhodopseudomonas palustris* MatB (rpMatB). We chose to add rpMatB along with DszAT even though *S. lividans* possesses a MatB because rpMatB has been shown to have an ~7-fold higher catalytic power than *Streptomyces* MatBs when catalyzing fluoromalonyl-CoA formation, while maintaining the same catalytic power for malonyl-CoA formation [1,3,32]. The low copy number of pCK7 and the single copy of DszAT-

Scheme 3.3: Workflow for the engineered production of site-specifically fluorinated polyketides in S. lividans. (A) S. lividans protoplasts are transformed with pCK7 AT⁰. (B-D) DszAT and MatB are introduced to the genome of S. lividans via conjugation with E. coli. (E) Biosynthetic representation of a production experiment in the engineered S. lividans strain designed in (A-D). Sodium propionate is fed to produce propionyl-CoA and methylmalonyl-CoA and fluoromalonate is fed to produce fluoromalonyl-CoA.

MatB will result in a close to 1:1 stoichiometry between DEBS module and DszAT, affording optimal conditions for *trans*-AT complementation with only the desired AT[0] modules [*9*].

Production of 6-dEB in *Streptomyces lividans*. Prior to the biosynthesis of site-specifically fluorinarted 6-dEB analogs, it was important to establish optimal growth and production conditions for *S. lividans* cultures containing the DEBS biosynthetic pathway. pCK7 was transformed into *S. lividans* via a protoplast transformation [*22*] and spore stocks of this culture were grown in nutrient broth containing 10 µg/mL thiostrepton (in DMSO) supplemented with 5 mM sodium propionate (to facilitate production of propionyl-CoA and methylmalonyl-CoA). The production of 6-dEB was detected by analyzing the supernatant using TOF MS (*Figure 3.11*). The production of 6-dEB was measured after 3 d, and then the culture was sampled every 24 h for the next two days. The production of 6-dEB appeared to taper off after 5 d, and so culture supernatants were analyzed 5-7 d after inoculation. More importantly, given that 6-dEB could easily be detected in the supernatant (*Figure 3.11*), the fluorinated 6-dEB derivatives could be detected even if they were to be produced at 100-fold deficiency by supernatant analysis and >1000-fold deficiency if the fluorinated product is extracted from the supernatant with ethyl acetate.

Introduction of AT[0] mutations using Cas9 nuclease. The introduction of mutations into the DEBS cluster in pCK7 is challenging given the high-GC content, plasmid size, and repetitive nature of PKS gene sequences. Due the recent success of CRISPR-Cas9 systems as a genome manipulation tool in *E. coli*, the use of these techniques along with recombineering methods [*33,34*] provide an opportunity for streamlined engineering of large plasmids such as pCK7. Initial efforts to introduce AT[0] mutations into pCK7 in *E. coli* used the pTarget system [*35*] consisting of two vectors, pTargetF and pCas. pCas contained a constitutively expressed Cas9 nuclease, as well as an arabinose inducible promoter for the expression of a recombinase, and pTargetF contained the sgRNA template. The workflow consisted of transforming pTargetF, pCK7, and a single-stranded repair arm (lagging strand) [*33-35*] into *E. coli* cells containing pCas and expressed recombinase. Multiple efforts to optimize this protocol failed to result in successful insertion of the AT[0] mutation. Conditions such as repair arm length and sense, recovery media, variation in *E. coli* strain used, recovery time, recovery temperature, and incubation temperature all failed to produce the desired AT[0] mutation. Based on the absence of colonies in most experimental conditions, it is most reasonable to infer that the engineered DNA repair mechanism is responsible for the lack of success, as the plasmids cut by Cas9 would not be repaired. The inability to
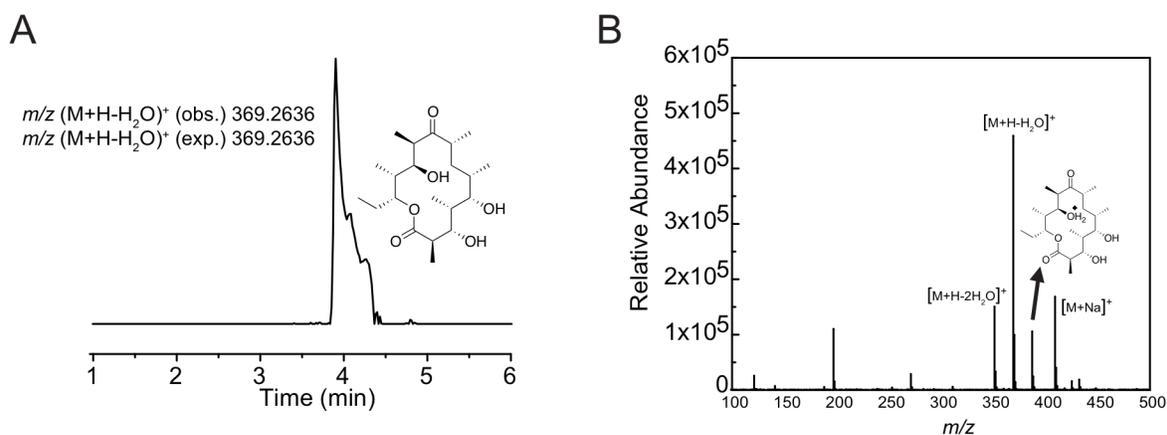
Figure 3.11: *Production of 6-dEB in Streptomyces lividans. (A) Extracted ion chromatogram of 6-dEB produced with S. lividans in nutrient broth. Mass spectrum of 6-dEB. The four characteristic 6-dEB ions can be observed ([M+H-2H₂O]⁺, [M+H-H₂O]⁺, [M+H]⁺, and [M+Na]⁺).*

79

Scheme 3.4: *Cas9-Gibson assembly plasmid editing protocol. Representation of a sequential Cas9-gibson procedure used to introduce AT^0 into PKS modules in plasmids. Cas9 nuclease is used as a restriction enzyme, providing a unique cut site near the desired modification site (in this case, the AT catalytic serine). The mutation is introduced via Gibson assembly [16] using an oligo or PCR product containing the desired mutation.*

introduce the AT^0 mutation using the pTargetF system inspired the use of Cas9 nuclease as a restriction enzyme. Studies have shown that RNPs can be used to linearize plasmids and enable the insertion of DNA using Gibson assembly [16,36]. Inspired by this result, we used synthetic sgRNAs to assemble RNPs *in vitro* [19] and digest pCK7 (*Scheme 3.4*). Once pCK7 is linearized, a repair arm containing the desired AT^0 mutation can be used to repair the double-strand break and form an AT^0 mutant variant of pCK7 (*Scheme 3.4*). It has been our experience that ligation efficiency decreases when large DNA fragments are used, so the vector pAYC02 [20] was used to test the viability of this protocol. pAYC02 is a vector containing the KSAT didomain of DEBS module 3, making it almost an order of magnitude smaller than pCK7 and a better substrate for optimization. As shown in *Figure 3.12*, when the RNP is assembled, pAYC02 is linearized. The gel bands are consistent with a double-strand break in close proximity to Ser651, as shown when an XbaI-treated sample produces the expected 2 and 6 kb fragments (*Figure 3.12*). The repair arm was assembled by annealing two 60 bp oligonucleotides containing the S651A mutation. The repair arm and Cas9-linearized pAYC02 were added to a Gibson assembly reaction and transformed into *E. coli*. Control reaction with either linearized vector or repair arm were also transformed into *E. coli*. No colonies were observed on the control plates and both of the colonies on the reaction plate contained the S651A mutation. While this protocol did not result in many transformants, all of the colonies produced the desired mutations. However, when this protocol was extended to introduction of AT^0 mutations into Modules 3 or 6 in pCK7, no AT^0 mutant transformants were
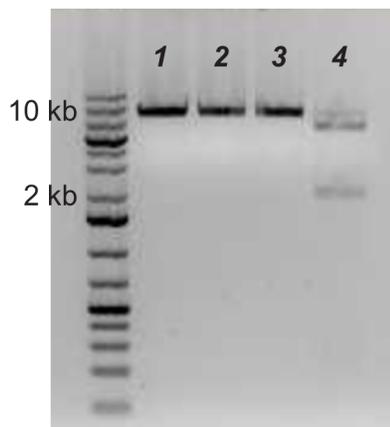
Figure 3.12: *Gel of in vitro Cas9 plasmid digest assay. DNA gel showing RNP (sgRNA M3 AT^0) activity on pAYC02 (KSAT didomain). Lanes: 1 mock treat 2 sgRNA only 3 Cas9 only 4 RNP. All samples were treated with XbaI prior to gel analysis to avoid plasmid supercoiling and to clarify digestion efficiency. As expected, treatment with RNP and XbaI results in expected bands at ~2 and ~6 kb.*

obtained and no more than 1-5 colonies were ever observed on transformation plates (including control plates). The small number of colonies suggested that the Cas9 catalyzed linearization was working efficiently but the repair mechanism needed to be optimized. To improve the repair efficiency, pCK7 was subjected to a "double-digest." This can be achieved using two different RNPs, each with a different PAM site. For our first effort, we chose to use the Module 6 AT^0 RNP along with the unique NsiI restriction site at the C-terminal of DEBS3, resulting in the excision of a 3.5 kb fragment. This digestion was performed sequentially, with the Cas9 digest occurring first, followed by the NsiI digest. When a PCR amplified 3.5 kb fragment containing the S2017A mutation (Based on EryAIII numbering) and appropriate pCK7 overlaps was incubated with the pCK7 double digest product in a Gibson assembly reaction, about half of the colonies screened contained the AT^0 mutation. The pCK7 M6 AT^0 construct will be available for use as soon as sequencing confirms that no undesired mutations were introduced during the Gibson reaction. This method appears to be the simplest way to introduce AT^0 mutations into these plasmids. It is fast, does not require multiple plasmids, and eliminates the complexity of large plasmid handling in *E. coli*. Additionally, as mentioned above, two RNPs can be used to perform a double digest and introduce AT^0 mutations into modules in the middle of the pathway, where unique restriction sites are not available.

With plasmids containing AT^0 mutations in hand, along with vectors for integration of DszAT variants into actinomycete genomes, we will soon have the ability to test for the production of fluorinated polyketides in *S. lividans*. Future work in the lab will focus on producing the gamut of monofluorinated analogs of 6-dEB. The only component yet to be explored in this system is the ability of DEBS β-processing enzymes to carry out their transformation in the presence of an α-fluoro substituent in the polyketide growing chain.

Production of full-length fluorinated polyketides using malonyl-CoA selective ATs. An often-overlooked facet important for the formation of full-length fluorinated polyketides is the ability of DEBS β-processing enzymes (KR, DH, ER) to catalyze transformations in the presence of a polyketide growing chain containing an α-fluoro moiety. These enzyme can convert a β-keto group to a methylene via the ketoreduction to a β-hydroxyl group, followed by formation of an
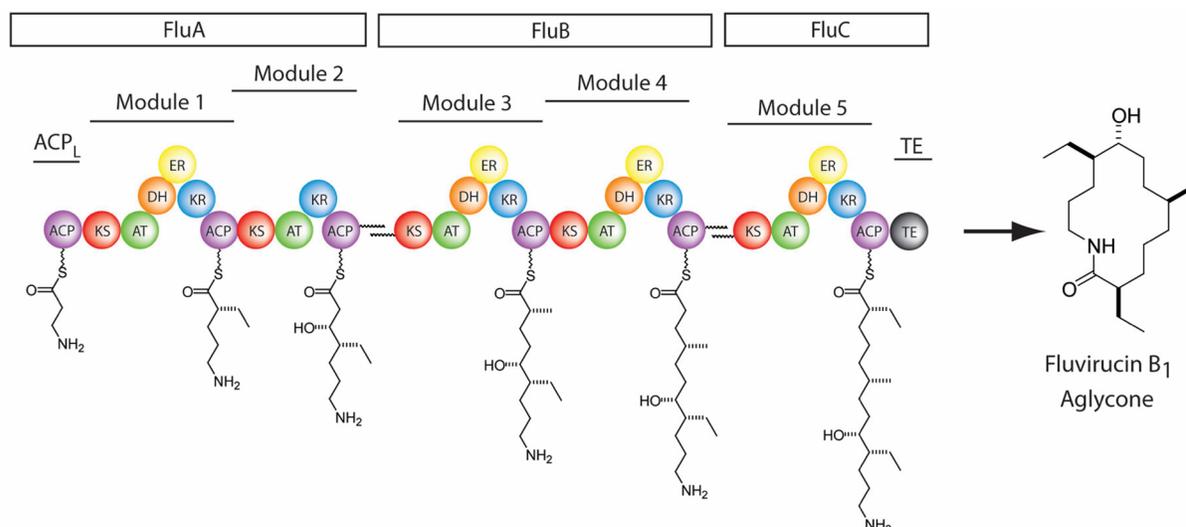
Scheme 3.5: *Fluvirucin B₁ PKS biosynthetic cluster (Flu). The 3-amino-3,6-L-talopyranose sugar is ligated onto the hydroxyl group after the lactam is assembled. Image obtained from [37]. Unlike DEBS, most of the modules in Flu are fully reducing.*

α,β-alkene that can be further reduced to form the saturated methylene. Given the ability of multiple *trans*-ATs to incorporate fluoromalonyl-CoA [*26*] and the relative success observed using DszAT to form F-TKLs [*2*], we felt that an effective way to test β-processing *in vivo* was to attempt to biosynthesize fluorinated polyketides using PKS pathways that contain malonyl-CoA specific ATs in modules that undergo β-processing (*Scheme 3.5*). However, the ability to transport and activate fluorinated extender units in actinomycetes has not been explored. Given this lack of information, it was worthwhile to test the *in vivo* incorporation of fluoromalonyl-CoA into malonyl-CoA specific modules of PKSs as a method to examine both the ability of actinomycetes to activate fluorinated extender units, as well as test the β-processing activity of fluorinated growing chains in these modules.

Fluvirucin B_1 is a proposed antiviral and antifungal that is produced by a five-module PKS in *Actinomadura vulgaris* (*Scheme 3.5*) [*37*]. The Fluvirucin B_1 PKS cluster (Flu) has two modules that incorporate malonyl-CoA (modules 2 and 4) (*Scheme 3.5*) [*37*]. As such, Flu is a good model system from studying β-processing of fluorinated polyketides. *A. vulgaris* spores were inoculated into nutrient broth containing either fluoroacetate (5 mM), fluoromalonate (5 mM), or diethylfluoromalonate (5 mM). The supernatant of these samples was analyzed after 10 days by TOF MS. It was exciting to see that all three fluorinated building blocks produced two compounds with exact masses corresponding to monofluorinated fluvirucin B_1 derivatives (*Figure 3.13B*). The presence a fluorine atom was confirmed using an MS/MS fragmentation indicative of a backbone fluorine atom (*Figure 3.13*). It is important to mention that module 4 in the Flu cluster contains all three β-processing enzymes, indicating that the complete set of reductions is possible with a fluorinated polyketide growing chain. The two isomers are also produced at similar levels (by MS integration, *Figure 3.13*), which could imply that ketoreduction (observed in module 2) and full β-processing (observed in module 4) can occur at similar efficiencies. Furthermore, the reasonably high relative yield of monofluorinated fluvirucin B_1 (8.3% ± 4.7%, by MS signal integration) suggests that intermodular chain transfer can take place and produce a full-length polyketide that is further elaborated by tailoring to the final product. Scaled-up production experiments will take place in the future in order to isolate and characterize fluvirucin B_1 by ^{19}F- ^{1}H and ^{13}C-NMR.

The production of monofluorionated fluvirucin B_1 under fluoroacetate feeding conditions indicates that fluoromalonyl-CoA may be formed through two biosynthetic pathways in *A. vulgaris*. While fluoromalonate is likely activated using a MatB homolog (*Scheme 3.6*), fluoromalonyl-CoA may also be biosynthesized from fluoroacetate by activation to fluoroacetyl-CoA, followed by a carboxylation with an acetyl-CoA carboxylase (ACCase, *Scheme 3.6*) [*1*]. The production of monofluorinated fluvirucin B_1 analogs is particularly interesting because when the same building blocks were fed to cultures of pikromycin-producing *Streptomyces venezuelae*, fluorinated Pikromycin could not be observed. The pikromycin cluster (PIKS) [*38*] contains a module (module 2) that is malonyl-CoA specific, yet neither monofluorinarted pikromycin or methymycin (minor product) were observed in the supernatant. This result suggests that either the Flu modules allow fluoromalonyl-CoA incorporation more readily than the PIKS module 2, or *A. vulgaris* can activate fluorinated building blocks more efficiently than *S. venezuelae*. If the latter is true, this would make *A. vulgaris* and great host for the production of fluorinated polyketides.

The inability of *S. venezuelae* to produce monofluorinated polyketide products provides a useful control experiment for understanding the source of fluoromalonyl-CoA incorporation into
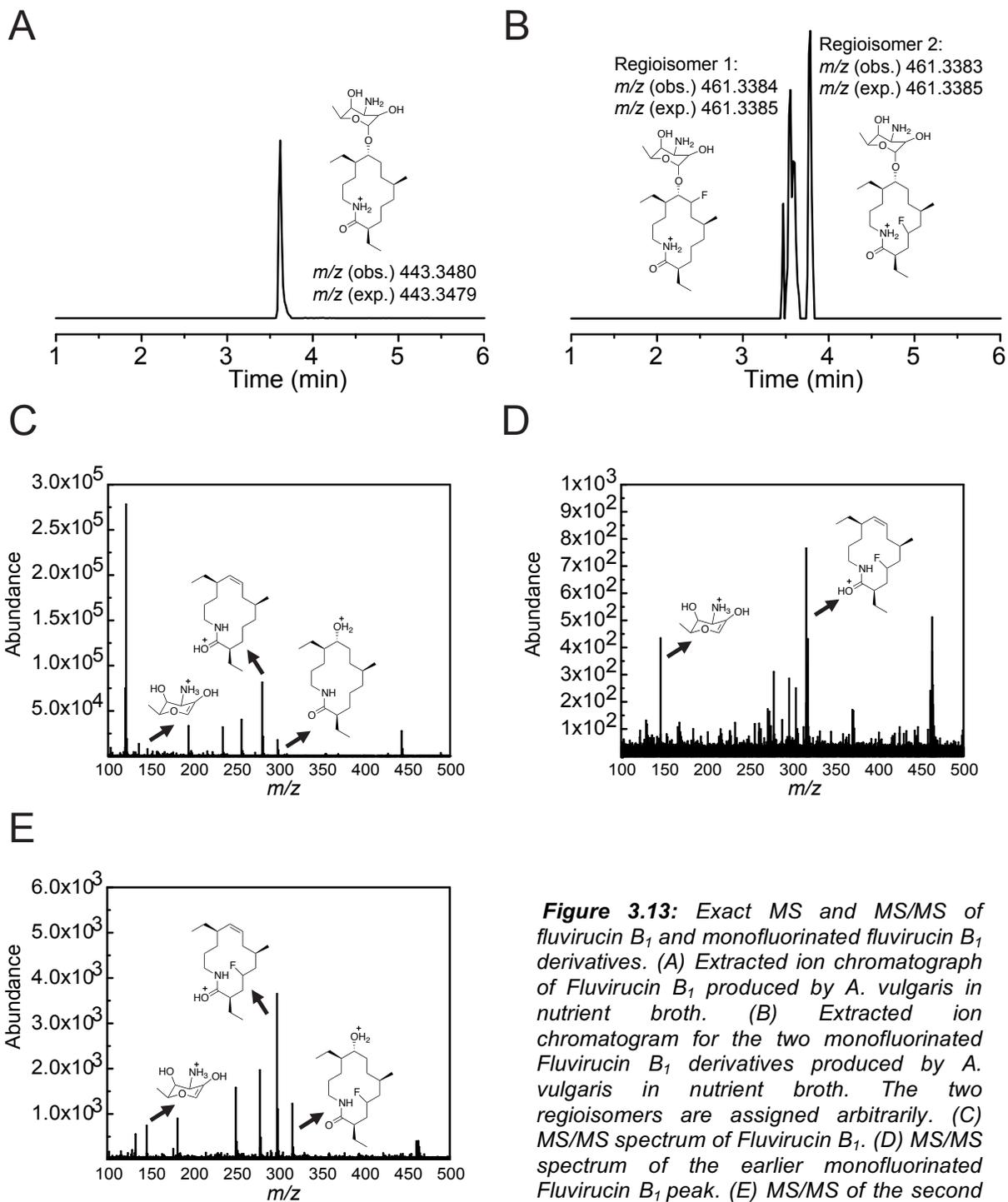
Figure 3.13: Exact MS and MS/MS of fluvirucin B₁ and monofluorinated fluvirucin B₁ derivatives. (A) Extracted ion chromatograph of Fluvirucin B₁ produced by A. vulgaris in nutrient broth. (B) Extracted ion chromatogram for the two monofluorinated Fluvirucin B₁ derivatives produced by A. vulgaris in nutrient broth. The two regioisomers are assigned arbitrarily. (C) MS/MS spectrum of Fluvirucin B₁. (D) MS/MS spectrum of the earlier monofluorinated Fluvirucin B₁ peak. (E) MS/MS of the second monofluorinated Fluvirucin B₁ peak.

fluvirucin B_1. One method to determine whether fluoromalonyl-CoA incorporation into fluvirucin B_1 is related to PKS module selectivity or monomer activation activity would be to introduce either the Flu cluster into *S. venezuelae* or to introduce the PIKS cluster into *A. vulgaris*. Unfortunately, these gene clusters are tens of kb in size, making this experiment difficult to execute. However, a simpler alternative would be to integrate a simple fluorine-permissive model PKS into the genome of both species. If the PKS product of this model enzyme is produced in both species, then the Flu ATs are responsible for fluorine incorporation. Conversely, if the model product is only observed in *A. vulgaris*, then fluorine incorporation is attributed to differential monomer activation between the two species. To address this, we propose using the free-standing KS NphT7 [39]. NphT7 has been shown to incorporate fluoromalonyl-CoA with ~5-fold deficiency compared to

Scheme 3.6: *Proposed fluoromalonyl activation pathways from the fluorinated building blocks fluoroacetate and fluoromalonate. Many organisms possess a malonate decarboxylase. If this class of enzyme is present in the A. vulgaris genome, then fluoroacetate might be the sole building block required for fluoromalonyl-CoA biosynthesis.*

malonyl-CoA, and is a good model system to test monomer activation since its activity towards fluoromalonyl-CoA is well characterized (*Scheme 3.7*) [1]. NphT7 catalyzes the decarboxylative Claisen condensation between acetyl-CoA and fluoromalonyl-CoA resulting in formation of 2-fluoroacetoacetate (*Scheme 3.7*). This product highly unstable, given its propensity to decarboxylate and form fluoroacetone [26], and thus NphT7 will be coupled with the KR PhaB to ultimately produce 2-fluorohydroxybutyrate (*Scheme 3.7*) [40,41]. 2-fluorohydroxybutyrate will not undergo terminal decarboxylation, making it easier to quantify by subjecting the culture supernatant to ^{19}F NMR analysis. The construct pSET152-ErmEP*-NphT7-PhaB was cloned and can be used to genomically integrate NphT7 and PhaB into *Streptomyces* genomes [17,18]. A conjugation with the *E. coli* strain GM272 containing pUZ8002 [22] and the actinomycetes *S. venezuelae*, *A. vulgaris*, and *S. lividans* (additional control), followed by the subsequent feeding of fluorinated building blocks will provide more insight into the nature of fluorine incorporation in observed in *A. vulgaris*.

The formation of these monofluorinated analogs of fluvirucin B_1 is an exciting opportunity to find an actinomycete host for fluorinated polyketide production. Very little is currently known about *A. vulgaris* and current efforts in the lab have been centered towards sequencing its genome (Chia-I Lin, unpublished data). Additionally, it will be important to develop protoplast
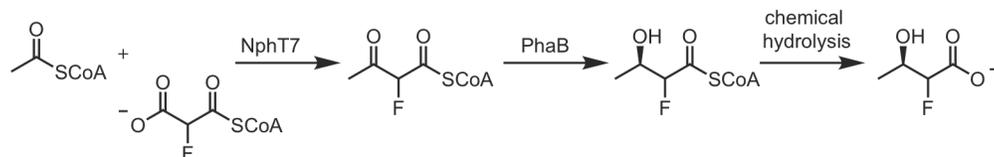
Scheme 3.7: *Formation of 2-fluorohydroxybutyrate using the chimeric NphT7-PhaB pathway. The Claisen condensation is catalyzed by NphT7 and the ketoreduction is catalyzed by PhaB. Chemical hydrolysis results in the final product, 2-fluorohydroxybutyrate.*

transformation protocols if conjugal transfer and genomic integration are not possible. Finally, it would be very informative to conduct RNAseq experiments with samples containing the various fluorinated building blocks, since these samples can provide insight into the biosynthetic machinery responsible for production or fluoromalonyl-CoA.

3.4 Conclusions

The complementation of DszAT with DEBS modules has shown promise as a source for both fluorinated product amplification, as well as site-specific introduction of fluoromalonyl-CoA into the backbone of model polyketides. Prior to the work discussed here, the complementation of DszAT in the presence of both fluoro- and malonyl-CoA has remained unexplored. Given DszAT's malonyl-CoA selectivity, the ability to load fluoromalonyl-CoA onto DEBS ACPs *in vivo* is directly related to the malonyl-CoA pool in the cells. We set out to screen for a fluoromalonyl-CoA selective DszAT variant via mutagenesis to active site residues rather than attempt to limit malonyl-CoA production in the host organism, which may have negative impact on the organism's growth. The mutagenesis screen ultimately resulted in the discovery of a fluoromalonyl-CoA DszAT variant containing a valine substitution of the active site phenylalanine (F190V). The F190V mutation removes the phenylalanine presumed to select for malonyl-CoA over larger extender units and has shown a 41-fold improvement in fluorine selectivity compared to WT DszAT.

Through the use of Cas9 nuclease, we were able to introduce an AT^0 mutation into *Streptomyces* DEBS expression vector capable of producing 6-dEB. After developing 6-dEB production conditions in *S. lividans*, we now have all of the components required for the production of site-specifically fluorinated 6-dEB *in vivo*. Future work in our lab will surely focus on production these fluorinated analogs.

Additionally, the ability to produce monofluorinated Fluvirucin B_1 derivatives in *A. vulgaris* has shown that type I modular PKS β-processing enzymes can carry out all of the reductive transformations in the presence of an α-fluoro substituent on the polyketide growing chain. This study also revealed that *A. vulgaris* is capable of biosynthesizing fluoromalonyl-CoA using two building blocks, fluoroacetate and fluoromalonate. Efforts to determine the molecular mechanism of monofluorinated Fluvirucin B_1 productions are ongoing and may have a significant implication on the discovery of optimal heterologous hosts for polyketide production.

3.5   References

1.   Walker, M. C., Thuronyi, B. W., Charkoudian, L. K., Lowry, B., Khosla, C., Chang, M. C. Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* **2013**, *341* (6150) 1089-1094.

2.   Ad, O., Thuronyi, B. W., Chang, M. C. Elucidating the mechanism of fluorinated extender unit loading for improved production of fluorine-containing polyketides. *Proc Natl Acad Sci U S A* **2017**, *114* (5) E660-E668.

3.   Thuronyi, B. W., Chang, M. C. Synthetic biology approaches to fluorinated polyketides. *Acc Chem Res* **2015**, *48* (3) 584-592.

4. Smart, B. E. Fluorine substituent effects (on bioactivity). *J Fluorine Chem* **2001**, *109* (1) 3-11.

5. Koryakina, I., McArthur, J. B., Draelos, M. M., Williams, G. J. Promiscuity of a modular polyketide synthase towards natural and non-natural extender units. *Org Biomol Chem* **2013**, *11* (27) 4449-4458.

6. Musiol-Kroll, E. M., Zubeil, F., Schafhauser, T., Hartner, T., Kulik, A., McArthur, J., Koryakina, I., Wohlleben, W., Grond, S., Williams, G. J., Lee, S. Y., Weber, T. Polyketide bioderivatization using the promiscuous acyltransferase KirCII. *ACS Synth Biol* **2017**, *6* (3) 421-427.

7. Carvalho, R., Reid, R., Viswanathan, N., Gramajo, H., Julien, B. The biosynthetic genes for disorazoles, potent cytotoxic compounds that disrupt microtubule formation. *Gene* **2005**, *359* 91-98.

8. Wong, F. T., Jin, X., Mathews, II, Cane, D. E., Khosla, C. Structure and mechanism of the *trans*-acting acyltransferase from the disorazole synthase. *Biochemistry* **2011**, *50* (30) 6539-6548.

9. Dunn, B. J., Watts, K. R., Robbins, T., Cane, D. E., Khosla, C. Comparative analysis of the substrate specificity of *trans-* versus *cis*-acyltransferases of assembly line polyketide synthases. *Biochemistry* **2014**, *53* (23) 3796-3806.

10. Weeks, A. M., Coyle, S. M., Jinek, M., Doudna, J. A., Chang, M. C. Structural and biochemical studies of a fluoroacetyl-CoA-specific thioesterase reveal a molecular basis for fluorine selectivity. *Biochemistry* **2010**, *49* (43) 9269-9279.

11. Koryakina, I., Kasey, C., McArthur, J. B., Lowell, A. N., Chemler, J. A., Li, S., Hansen, D. A., Sherman, D. H., Williams, G. J. Inversion of extender unit selectivity in the erythromycin polyketide synthase by acyltransferase domain engineering. *ACS Chem Biol* **2017**, *12* (1) 114-123.

12. Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., Schulz, F. Enzyme-directed mutasynthesis: A combined experimental and theoretical approach to substrate recognition of a polyketide synthase. *ACS Chem Biol* **2013**, *8* (2) 443-450.

13. Bravo-Rodriguez, K., Klopries, S., Koopmans, K. R., Sundermann, U., Yahiaoui, S., Arens, J., Kushnir, S., Schulz, F., Sanchez-Garcia, E. Substrate flexibility of a mutated acyltransferase domain and implications for polyketide biosynthesis. *Chem Biol* **2015**, *22* (11) 1425-1430.

14. Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **2001**, *291* (5509) 1790-1792.

15. Wong, F. T., Chen, A. Y., Cane, D. E., Khosla, C. Protein-protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. *Biochemistry* **2010**, *49* (1) 95-102.

16.   Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., 3rd, Smith, H. O. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **2009**, *6* (5) 343-345.

17.   Bibb, M. J., White, J., Ward, J. M., Janssen, G. R. The mrna for the 23s rRNA methylase encoded by the ErmE gene of *Saccharopolyspora erythraea* is translated in the absence of a conventional ribosome-binding site. *Mol Microbiol* **1994**, *14* (3) 533-545.

18.   Matsushima, P., Broughton, M. C., Turner, J. R., Baltz, R. H. Conjugal transfer of cosmid DNA from *Escherichia coli* to *Saccharopolyspora spinosa*: Effects of chromosomal insertions on macrolide A83543 production. *Gene* **1994**, *146* (1) 39-45.

19.   Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337* (6096) 816-821.

20.   Chen, A. Y., Schnarr, N. A., Kim, C. Y., Cane, D. E., Khosla, C. Extender unit and acyl carrier protein specificity of ketosynthase domains of the 6-deoxyerythronolide B synthase. *J Am Chem Soc* **2006**, *128* (9) 3067-3074.

21.   Kao, C. M., Katz, L., Khosla, C. Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science* **1994**, *265* (5171) 509-512.

22.   T. Kieser, M. J. B., M. J. Buttner, K. F. Chater, D. A. Hopwood *Practical streptomyces genetics* (Norwich: John Innes Foundation) **2000**.

23.   Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat Prod Rep* **2012**, *29* (10) 1050-1073.

24.   Weeks, A. M., Chang, M. C. Catalytic control of enzymatic fluorine specificity. *Proc Natl Acad Sci U S A* **2012**, *109* (48) 19667-19672.

25.   Labonte, J. W., Townsend, C. A. Active site comparisons and catalytic mechanisms of the hot dog superfamily. *Chem Rev* **2013**, *113* (3) 2182-2204.

26.   Thuronyi, B. W. (2015) Engineered biosynthesis of fluorinated polyketides. PhD (University of California at Berkeley, California).

27.   Lau, J., Tran, C., Licari, P., Galazzo, J. Development of a high cell-density fed-batch bioprocess for the heterologous production of 6-deoxyerythronolide B in *Escherichia coli*. *J Biotechnol* **2004**, *110* (1) 95-103.

28.   Stevens, D. C., Hari, T. P., Boddy, C. N. The role of transcription in heterologous expression of polyketides in bacterial hosts. *Nat Prod Rep* **2013**, *30* (11) 1391-1411.

29.   Rodriguez, E., Menzella, H. G., Gramajo, H. Heterologous production of polyketides in bacteria. *Methods Enzymol* **2009**, *459* 339-365.

30.   Fong, R., Vroom, J. A., Hu, Z., Hutchinson, C. R., Huang, J., Cohen, S. N., Kao, C. M. Characterization of a large, stable, high-copy-number streptomyces plasmid that requires stability and transfer functions for heterologous polyketide overproduction. *Appl Environ Microbiol* **2007**, *73* (4) 1296-1307.

31.   Ziermann, R., Betlach, M. C. Recombinant polyketide synthesis in streptomyces: Engineering of improved host strains. *Biotechniques* **1999**, *26* (1) 106-110.

32. Crosby, H. A., Rank, K. C., Rayment, I., Escalante-Semerena, J. C. Structure-guided expansion of the substrate range of methylmalonyl coenzyme A synthetase (MatB) of *Rhodopseudomonas palustris*. *Appl Environ Microbiol* **2012**, *78* (18) 6619-6629.

33. Sawitzke, J. A., Costantino, N., Li, X. T., Thomason, L. C., Bubunenko, M., Court, C., Court, D. L. Probing cellular processes with oligo-mediated recombination and using the knowledge gained to optimize recombineering. *J Mol Biol* **2011**, *407* (1) 45-59.

34. Thomason, L. C., Sawitzke, J. A., Li, X., Costantino, N., Court, D. L. Recombineering: Genetic engineering in bacteria using homologous recombination. *Curr Protoc Mol Biol* **2014**, *106* 1 16 11-39.

35. Jiang, Y., Chen, B., Duan, C., Sun, B., Yang, J., Yang, S. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol* **2015**, *81* (7) 2506-2514.

36. Wang, J. W., Wang, A., Li, K., Wang, B., Jin, S., Reiser, M., Lockey, R. F. CRISPR/Cas9 nuclease cleavage combined with gibson assembly for seamless cloning. *Biotechniques* **2015**, *58* (4) 161-170.

37. Lin, T. Y., Borketey, L. S., Prasad, G., Waters, S. A., Schnarr, N. A. Sequence, cloning, and analysis of the fluvirucin B_1 polyketide synthase from *Actinomadura vulgaris*. *ACS Synth Biol* **2013**, *2* (11) 635-642.

38. Kittendorf, J. D., Sherman, D. H. The methymycin/pikromycin pathway: A model for metabolic diversity in natural product biosynthesis. *Bioorgan Med Chem* **2009**, *17* (6) 2137-2146.

39. Okamura, E., Tomita, T., Sawa, R., Nishiyama, M., Kuzuyama, T. Unprecedented acetoacetyl-Coenzyme A synthesizing enzyme of the thiolase superfamily involved in the mevalonate pathway. *Proc Natl Acad Sci U S A* **2010**, *107* (25) 11265-11270.

40. Kim, J., Chang, J. H., Kim, E. J., Kim, K. J. Crystal structure of (r)-3-hydroxybutyryl-CoA dehydrogenase PhaB from *Ralstonia eutropha*. *Biochem Biophys Res Commun* **2014**, *443* (3) 783-788.

41. Bond-Watts, B. B., Bellerose, R. J., Chang, M. C. Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat Chem Biol* **2011**, *7* (4) 222-227.

Chapter 4: *Exploring extender unit promiscuity in polyketide synthases*

4.1 Introduction

The incorporation of unnatural extender units into polyketides has been a long-standing research area in the field of natural product biosynthesis [1-4]. Recent studies have shown that PKS modules with methylmalonyl-CoA specific ATs are able to incorporate larger extender units without any mutations to domain active sites [3,5]. Additionally, engineering efforts have resulted in change of substrate selectivity [6] as well as the production of unnatural polyketides *in vivo* [7,8]. Two extender unit classes that have yet to be explored thoroughly are α-halogenated extender units, as well an α-disubstituted extender units. These two classes of unnatural extender units are of interest to the fields of natural product biosynthesis and pharmacology, since α-halogenated extender unit incorporation can provide access to a wide variety of polyketide products via a synthetic handle on the polyketide backbone [9] and α-disubstituted positions on the backbone of various semi-synthetic drugs has shown potential for creating new antibiotics [10,11].

The incorporation of α-fluorinated extenders units has previously been studied and found to proceed via an ACP-independent mechanism of C-C bond formation [4]. However, the use of larger halogens like chlorine and bromine has the potential to be a conservative substitution for a methyl group given the steric similarity of chlorine and bromine atoms to a methyl group [12,13]. Even though chlorine and bromine moieties bear steric similarity, it is important to consider that they are great leaving groups and have the potential to covalently inhibit enzymes by converting the substrate into an electrophilic donor for nucleophilic residues in the active site [14]. When we tested incorporation of chloro- and bromomalonyl-CoA into the DEBS modules, we found that bromomalonate is a covalent inhibitor of DEBS modules, and chloromalonyl-CoA incorporation, while robust, proceeds through an ACP-independent mode of C-C bond formation. Nonetheless, the ability to introduce chloromalonyl-CoA in smaller polyketide scaffolds results in formation of chemically relevant chiral building blocks that contain a valuable synthetic handle.

The formation of polyketides containing substituted positions on backbone limits the β-processing enzyme activity at certain positions (*Scheme 4.1A*), but has been shown to be beneficial when the α-disubstituted extender unit is incorporated at the terminal position of the polyketide [10,11]. For example, the introduction of gem-dimethyl groups improves chemical stability bioactive molecules [15-17], since the α-disubstituted esters are non-enolizable and sterically obscured (*Scheme 4.1B*), resulting in improved pharmacokinetic properties [18,19]. Given the

A   **B**
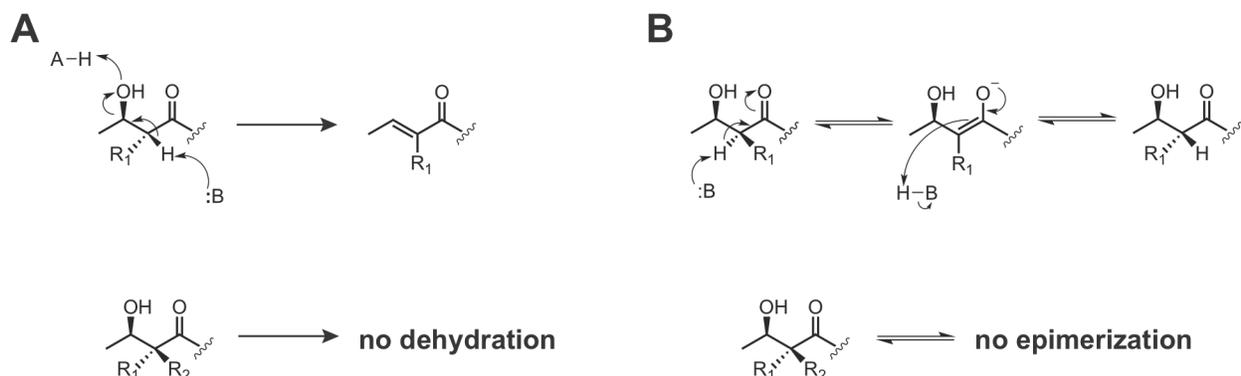
Scheme 4.1: *Biosynthetic consequences of α-disubstituted polyketides. (A) Scheme representing a dehydration reaction of a polyketide. The α-disubstituted polyketide cannot undergo a dehydration since there is no α-proton that can be abstracted by the base. (B) Scheme representing the epimerization of a polyketide. The α-carbon of the α-disubstituted polyketide cannot be epimerized.*

conservative substitution between fluorine and hydrogen, we tested the incorporation of 2-fluor-2-methylmalonyl-CoA into DEBS polyketides. Additionally, seeking to incorporate the synthetically challenging –CF$_2$– motif [20,21], we also attempted to synthesize and test 2,2-difluoromalonyl-CoA incorporation into DEBS modules.

Finally, the robust nature of ethylmalonyl-CoA incorporation into DEBS modules [5] inspired our interest to incorporate 4-fluoroethylmalonyl-CoA into DEBS proteins. 4-fluoroethylmalonyl-CoA has previously been biosynthesized using a chimeric *Salinospora tropica* chloroethylmalonyl-CoA biosynthetic pathway containing the *Streptomyces cattleya* fluorinase instead of the native *S. tropica* chlorinase [22]. In this study, 4-fluoroethylmalonyl-CoA was biosynthesized using a chimeric pathway containing enzymes from three different organisms [2,23,24]. *In situ* generated 4-fluoroethylmalonyl-CoA was incorporated into a DEBS module via an ACP-dependent mechanism of C-C bond formation, suggesting that this unnatural extender unit can be incorporated into full length polyketides *in vivo* through a solely biosynthetic pathway.

4.2 Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Carbenicillin (Cb), glucose, isopropyl-β-D-thiogalactopyranoside (IPTG), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride, dithiothreitol (DTT), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), kanamycin (Km), acetonitrile, dichloromethane, ethyl acetate, ethylene diamine tetraacetic acid disodium dihydrate (EDTA), and restriction enzymes were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A sodium salt (CoA), malonyl-CoA, methylmalonyl-CoA, ethylmalonate, acetyl-CoA, diethylfluoromalonate, diethylbromomalonate, diethylchloromalonate, malonic acid, diethylmethylmalonate, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenolpyruvate (PEP), adenosine triphosphate sodium salt (ATP), myokinase, pyruvate kinase, lactate dehydrogenase, poly(ethyleneimine) solution (PEI), mannitol, titanium tetrachloride in toluene (1M), TES, β-mercaptoethanol, thiamine pyrophosphate, alpha ketoglutaric acid, sodium phosphate dibasic hepthydrate, cysteamine, 4-hydroxy-6-methyl-2-pyrone acetic anhydride, 1-ethyl-3-(-dimethylaminopropyl)carbodiimide hydrochloride (EDC), 4-dimethylaminopyridine (DMAP), *o*-benzyl hydroxylamine, N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED), Selectfluor, reduced nicotinamide adenine dinucleotide phosphate (NADPH), reduced nicotinamide adenine dinucleotide (NADH), acetonitrile, dimethyl sulfoxide (DMSO), ammonium acetate, and ammonium formate were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate was purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, T4 DNA ligase, Phusion DNA polymerase, amylose resin, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in 10 mM Tris-HCl, pH 8.5, and stored at either 4°C for immediate use or -20°C for longer term use. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Amicon Ultra 3,000

MWCO, 10,000 MWCO centrifugal concentrators, and 30,000 MWCO centrifugal concentrators were purchased from EMD Millipore (Billerica, MA). Sodium fluoroacetate was purchased from Fluka-Honeywell (Mexico City, Mexico). Diethyl difluoromalonate was purchased from Matrix Scientific (Columbia, SC). Chloroform-*d* were purchased from Cambridge Isotope Laboratories (Andover, MA). ^{13}C and ^1H NMR spectra were collected at 25ºC on Bruker AV-600 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley. Assignments were made based on literature president and reference spectra from authentic standards, where appropriate. High-resolution mass spectral analyses were carried out on a 6530 QTOF Accurate Mass spectrometer purchased from Agilent Technologies.

Bacterial strains. BL21(de3)T1^R cells were used for heterologous protein expression. *E. coli* BAP1 [25] cells were used for heterologous expression of DEBS modules.

Expression and purification of His-tagged and MBP-tagged proteins. All of the enzymes used in this chapter were initially expressed and purified in Chapter 2, with the exception of NphT7 (gifted by Ben Thuronyi), PhaB (gifted by Brooks Bond-Watts), PhaJ [23], and Ccr [23]). The protocols for expression and purifications of enzymes have been previously described (Section 2.2).

Purification of Ccr-His$_6$ and His$_{10}$-PhaJ. The enzymes were heterologously expressed in *E. coli* BL21(de3) T1^R as described previously (Section 2.2). Frozen cell pellets were thawed and resuspended at 5 mL/g cell paste with Buffer B (50 mM sodium phosphate, 300 mM sodium chloride, 20% (*v/v*) glycerol, 20 mM BME, pH 7.5) containing imidazole (10 mM). Complete EDTA-free protease inhibitor cocktail (Roche) was added to the lysis buffer before resuspension. The cell paste was homogenized before lysis by passage through a French Pressure cell (Thermo Scientific; Waltham, MA) at 14,000 psi. The lysate was centrifuged at 15,300 × g for 20 min at 4°C to separate the soluble and insoluble fractions. DNA was precipitated in the soluble fraction by addition of 0.015% (*w/v*) poly(ethyleneimine). The precipitated DNA was removed by centrifugation at 15,300 × g for 20 min at 4°C. The remaining soluble lysate was diluted three-fold with Buffer B containing 10 mM imidazole and loaded onto a Ni-NTA agarose column (Qiagen, 1 mL resin/g cell paste) by gravity flow. The column was washed with Buffer A until the eluate reached an A$_{280\ nm}$ < 0.05 or was negative for protein content by Bradford assay (Bio-Rad). The column was washed with 20 column volumes with Buffer A supplemented with 25 mM imidazole. The protein was then eluted with Buffer D (50 mM sodium phosphate, 50 mM sodium chloride, 20% (*v/v*) glycerol, 20 mM BME (pH 7.5)) containing Imidazole (250 mM). Fractions containing the target protein were then exchanged into Buffer C (50 mM HEPES, 100 mM sodium chloride, 2.5 mM EDTA, 2.5 mM DTT, 20% (*v/v*) glycerol, pH 7.5) by 2 x 3 h rounds of dialysis (24 mm dialysis tubing with 3.5-5 kDa MWCO from SpectrumLabs (Rancho Dominguez, CA)) and then concentrated before storage using an Amicon Ultra spin concentrator (3 kDa MWCO, Millipore). Final protein concentrations before storage were estimated using the ε$_{280\ nm}$ calculated by ExPASY ProtParam as follows: Ccr-His$_6$: ε$_{280\ nm}$ = 66,600 M^{-1} cm^{-1}. PhaJ does not have any Tryptophan residues, so protein quantification was obtained using a Bradford assay. All proteins were aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C.

Fluoromalonate and methylmalonate. Fluoromalonate and methylmalonate were prepared as previously described (Section 2.2).

Bromomalonate, chloromalonate, and 2,2-difluoromalonate. Bromomalonate, chloromalonate, 2,2-difluoromalonate were prepared in similar fashion to fluoromalonate (see

section 2.2), with the following modifications to the protocol. Diethylbromomalonate, diethylchloromalonate, or diethyldifluoromalonate (0.5 mL, 3.2 mmol) were saponified with methanolic sodium hydroxide (2 M, 3.5 mL) in dichloromethane and methanol (9:1 v/v, 32 mL). After stirring at r.t for 30 min, the cloudy mixture was rotovapped to dryness. The reaction was resuspended in 15 mL of water and washed with 15 mL of diethyl ether to remove any unreacted starting material. The aqueous phase was then acidified to pH 2-3 using hydrochloric acid (6 M) and the product was extracted by washing the aqueous layer with 2 x 30 mL of either ethyl acetate of dichloromethane. The organic washes were pooled and rotovapped to dryness. Bromomalonate ^{1}**H-NMR** (400 MHz, D_2O = 4.79 ppm)**:** δ 4.60 (s, 1H). Chloromalonic acid ^{1}**H-NMR** (400 MHz, $CDCl_3$ = 7.26 ppm)**:** δ 4.75 (s, 1H) ^{13}**C-NMR** (101 MHz, $CDCl_3$ = 77.36)**:** δ 55.1, 165.0. 2,2-difluorofluoromalonate ^{19}**F-NMR** (366 MHz, D_2O)**:** δ -107.4 (s, 2F).

2-fluoro-2-methylmalonate. The fluorination procedure was modified from a previous study [26]. Diethylmethylmalonate (170 µL, 1 mmol) was added to 10 mL of acetonitrile in a 50 mL round bottom flask with a stirbar. Selectfluor (780 mg, 2.2 mmol) was added to the reaction mixture along with 0.5 mL titanium tetrachloride (1M in toluene). The reaction was heated to 80ºC under reflux for 5 h and then partitioned with 40 mL of water. The reaction mixture was washed with 3 x 50 mL ethyl acetate and the organic layers were pooled and dried by rotary evaporation to produce a yellow liquid. The product saponified and extracted as described above. ^{1}**H-NMR** (400 MHz, D_2O = 4.79 ppm)**:** δ 1.55 and 1.61 (d, 3H, J_{HF} 22 Hz). ^{19}**F-NMR** (366 MHz, D_2O)**:** δ -143.2 (q, F, J_{HF} 22 Hz).

N-acetylcysteamine thioester of (2S,3R)-2-methyl-3-hydroxypentanoic acid (NDK-SNAC) using PIKS KR1. NDK-SNAC was prepared and purified as previously described (Section 2.2).

N-acetylcystamine thioester of propionic acid (Propionyl-SNAC). Propionyl-SNAC was prepared and purified as previously described (Section 2.2).

Malonyl-CoA synthetase kinetic characterization. rpMatB activity was measured using a modified literature method [2]. The production of AMP was coupled to pyruvate formation by myokinase and pyruvate kinase, which was then coupled to NADH oxidation by lactate dehydrogenase. Assays were performed at 30°C in a total volume of 200 µL containing HEPES (100 mM, pH 7.5), TCEP (1 mM), ATP (2.5 mM), magnesium chloride (5 mM), phosphoenolpyruvate (1 mM), NADH (0.3 mM), myokinase (0.5 U), pyruvate kinase (3.6 U), lactate dehydrogenase (2.6 U), chloromalonate (50 µM – 5 mM) or bromomalonate (1.7 µM – 3.7 mM), and MatB (20 nM for chloromalonate and 10 nM for bromomalonate). The reaction was initiated with addition of CoA (0.5 mM) and monitored at 340 nm in a Beckman Coulter DU-800 spectrophotometer. Kinetic parameters (k_{cat}, K_M) were determined by fitting the data using Microcal Origin to the equation: $v_0 = v_{max} [S] / (K_M + [S])$, where v_0 is the initial rate and [S] is the substrate concentration. Error in k_{cat}/K_M was calculated by propagation of error from the individual kinetic parameters.

Chloroacetyl-CoA analysis using HPLC and HPLC-MS. A small volume of a chloromalonyl-CoA formation reaction using rpMatB (50 µL) was set up as described above for fluoromalonyl-CoA (Section 2.2). The reaction was incubated for 2 h at 37ºC. Protein was removed using a NanoSep 10K omega spin concentrator (Pall New York, NY) and the sample analyzed on a Zorbax Eclipse XDB C-18 column (3.5 µm, 3 × 150 mm, r,t, Agilent) using a linear gradient from 2 to 25% acetonitrile over 8 min with a aqueous mobile phase containing sodium

93

phosphate (50 mM), sodium citrate (50 mM), and 0.2% TFA (pH 4.5) (0.8 mL/min). Reactions were monitored using an Agilent G1365D diode array detector (chloroacetyl-CoA, A_{260}). To confirm compound identity, the samples were analyzed on an Zorbax Eclipse XDB C-18 column (3.5 μm, 3 × 150 mm, r,t, Agilent) using a linear gradient from 0 to 10% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 30 s (0.8 mL/min). The CoA products were identified using an Agilent 6100 Mass Spectrometer in the negative mode. Chloroacetyl-CoA **ESI-MS [M-H]⁻** calculated for $C_{23}H_{36}ClN_7O_{17}P_3S$, 842.1 *m/z*, 843.1 *m/z*, 845.1 *m/z*, 846.1 *m/z*; found, 842.1 *m/z*, 843.1 *m/z*, 845.1 *m/z*, 846.1 *m/z*. Chloromalonyl-CoA $C_{24}H_{36}ClN_7O_{19}P_3S$, 886.1 *m/z*, 887.1 *m/z*, 888.1 *m/z*, 889.1 *m/z*; found 886.1 *m/z*, 887.1 *m/z*, 888.1 *m/z*, 889.1 *m/z*.

Formation of 2-fluoro-2-methylmalonyl-CoA using rpMatB. A small volume of a 2-fluoro-2-methylmalonyl-CoA formation reaction using rpMatB (50 μL) was set up as described above for fluoromalonyl-CoA (Section 2.2). The reaction was incubated for 90 min at 37ºC. The reaction was quenched with 70% perchloric acid (*v/v*) to a final concentration of 5% (*v/v*) and analyzed using a Zorbax Eclipse XDB C-18 column (3.5 μm, 3 × 150 mm, r,t, Agilent) using a linear gradient from 0 to 10% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 30 s (0.8 mL/min). 2-fluoro-2-methylamalonyl-CoA was identified using an Agilent 6100 Mass Spectrometer in the negative mode. **ESI-MS [M-H]⁻** calculated for $C_{25}H_{38}FN_7O_{19}P_3S$, 884.1 *m/z*, found 884.2 *m/z*. The reaction was also monitored using an Agilent G1315D diode array detector ($A_{260\ nm}$). Using the A_{260} of the compounds, it was determined that rpMatB converted ~40% of free-CoA to 2-fluoro-2-methylmalonyl-CoA.

Triketide lactone production using DEBS protein from modules 2,3, and 6. Chlorinated triketide lactones (Cl-TKLs) were produced and analyzed in the same manner as the TKLs described above (section 2.2). Cl-TKL production was quantified using an Agilent G1315D diode array detector (TKL, $A_{260\ nm}$). When quantification of Cl-TKL is referenced, a TKL standard curve was used instead of a Cl-TKL standard curve. Ethyltriketide lactones (Et-TKLs) were produced and analyzed in the same manner as Cl-TKL using ethylmalonate (5 mM). When quantification of Et-TKL is referenced, a TKL standard curve was used instead of a Et-TKL standard curve.

Triketide lactone production using propionyl-SNAC with the Mod2 + Mod3$_{TE}$ mini-PKS system. Fluoro-chlorotriketide lactones (FCl-TKL) were produced and analyzed as described in section 2.2. Chloromalonate and fluoromalonate were present in equimolar concentration (5 mM).

Covalent inhibition assay for WT Mod6$_{TE}$. Two TKL reactions (200 μL) containing the regeneration system were prepared as previously described (Section 2.2). One reaction contained either chloro- or bromomalonate (5 mM) and the other contained methylmalonate (5 mM). Mod6$_{TE}$ (10 μM) and NDK-SNAC (5 mM) were added to each assay mixture and the reactions were incubated at 37°C for 18 h. The proteins in the reaction were isolated using a Sephadex G-25 column (4 mL) using 400 mM sodium phosphate, pH 7.5. Protein containing fractions were pooled by Bradford assay and concentrated to 200 μL using NanoSep 10K omega spin concentrator (Pall New York, NY) at r.t. The isolated protein (containing WT Mod6$_{TE}$) was assayed by adding TCEP (2.5 mM), methylmalonyl-CoA (1 mM) and NDK-SNAC (1 mM) and incubated overnight at 37°C. The samples were analyzed using the conventional TKL analysis method (HPLC-MS and HPLC-UV) as previously described (Section 2.2). Additional experiments were carried out with bromomalonate (5 mM) in the absence of NDK-SNAC as well as bromomalonate (5 mM) without MatB and the CoA regeneration system.

Enzymatic Synthesis of 4-fluoroethylmalonyl-CoA. This protocol was inspired from methods previously established in the literature [2,23,24]. Reaction mixtures contained sodium phosphate (100 mM, pH 7.5), TCEP (5 mM), magnesium chloride (10 mM), sodium bicarbonate (10 mM), NADPH (0.5 or 5 mM), fluoro- or acetyl-CoA (0.1 or 1 mM), and malonyl-CoA (0.2 or 1 mM). The reactions were initiated by the addition of NphT7 (1 μM), PhaB (1 μM), PhaJ (1 μM), and Ccr (1 μM). The reaction was incubated for 2 h at 37ºC and quenched by adding 70% perchloric acid (*v/v*) to a final concentration of 5% (*v/v*). Samples were centrifuged at $18,000 \times g$ for 10 min at r.t to pellet the precipitated protein. The supernatant was removed and flash frozen. Excess salts were removed by centrifugation at $18,000 \times g$ for 5 min at r.t. The samples were analyzed on an EclipsePlus C-18 RRHD column (1.8 μm, 2.1 × 50 mm, r.t, Agilent) using a linear gradient from 0 to 7% acetonitrile over 4 min with 5 mM ammonium formate (not pH adjusted) as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). The various CoA products were identified using an Agilent 6530 QTOF Accurate Mass Spectrometer in the positive mode and monitored using an Agilent G1315D diode array detector (CoAs, $A_{260 \text{ nm}}$). Additionally, the samples were subjected to MS/MS analysis using an Agilent 6460 QQQ Mass Spectrometer. Samples were analyzed on a Poroshell 120 SB-Aq column (2.7 μm, 2.1 × 50 mm, r.t, Agilent) using the same gradient and flow rate as described above with a fragmentation voltage of 100 V and collision energy of 35 V.

Incorporation of 4-fluoroethylmalonyl-CoA into TKLs. Reaction mixtures contained sodium phosphate (100 mM, pH 7.5), TCEP (5 mM), magnesium chloride (10 mM), sodium bicarbonate (10 mM), NADPH (0.5 mM), fluoroacetyl-CoA (0.2 mM), and malonyl-CoA (0.1 mM). The reaction was initiated by the addition of NphT7 (1 μM), PhaB (1 μM), PhaJ (1 μM), and Ccr (1 μM). The reaction was incubated for 20 min at 37ºC at which point components of the TKL formation assay were added [2,4]. PEP (5 mM), magnesium chloride (4 mM), ATP (2.5 mM), pyruvate kinase (2 U/mL) and myokinase (1 U/mL) were added and the reaction was initiated by the addition of NDK-SNAC (2.5 mM), rpMatB (20 μM) and WT Mod3$_{\text{TE}}$ (5 μM). The reaction was incubated for 4 or 16 h at 37ºC and quenched by adding 70% perchloric acid (*v/v*) to a final concentration of 5% (*v/v*). Samples were centrifuged at $18,000 \times g$ for 10 min at r.t to pellet the precipitated protein. The supernatant was removed and flash frozen. Excess salts were removed by centrifugation at $18,000 \times g$ for 5 min at r.t. The samples were analyzed on an EclipsePlus C-18 RRHD column (1.8 μm, 2.1 × 50 mm, r.t, Agilent) using a linear gradient from 0 to 40% acetonitrile over 4 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 12 s (0.6 mL/min). This was followed by an additional step of 40%-100% acetonitrile in 1 min (0.6 μL/min). The products were identified using an Agilent 6530 QTOF Accurate Mass Spectrometer in the positive mode.

ACP occupancy of 4-fluoroethylmalonate and ethylmalonate in Mod3$_{\text{TE}}$. 4-fluoroethylmalonyl-CoA and ethylmalonyl-CoA were generated using the chimeric enzymatic pathway as described above. WT or AT^0 Mod3$_{\text{TE}}$ (1 μM) were added and the sample was incubated for 30 min at 37ºC. The tryptic digest and peptide analysis were carried out as previously described (Section 2.2). Specifically, appropriate peptides were identified using an Agilent QTOF 6530 mass spectrometer. Products were characterized using both MS1 (exact mass <2 ppm) and MS2 (phosphopantetheine ejection). The samples were analyzed in positive mode with a fragmentor voltage of 150 and collison energy of 35 V. The parent ions monitored were 683.7 *m/z* (*holo*-ACP, +3), 721.7 *m/z* (ethylmalonyl-ACP, +3), and 727.7 *m/z* (4-fluoroerthylmalonyl-ACP, +3).

4.3 Results and Discussion

Incorporation of chlorinated and brominated extender units into polyketides. Studies have shown that methylmalonyl-CoA specific ATs (i.e DEBS ATs) are fairly permissive for extender units containing α-substituents larger than a methyl group [3,5,6,27]. Inspired by these findings, we reasoned that bromomalonyl-CoA and chloromalonyl-CoA may serve as good candidate substrates to test DEBS AT promiscuity. A chlorine atom has an effective Van der

Table 4.1: *Effective Van der Waals radii of groups or atoms at the α-position of proposed extender units [12].*

Group or Atom	Effective Van der Waals radius (Å)[12]
Cl	1.73
CH$_3$	1.80
Br	1.86

Waals [12] radius slightly smaller than a methyl group [13] and a bromine atom has an effective Van der Waals radius slightly larger than a methyl group, suggesting that from a sterics perspective, the replacement of a methyl group with a chlorine or bromine atom is a conservative substitution (*Table 4.1*). Even though chlorine and bromine share a steric similarity with a methyl group, the two halogens are much better leaving groups than a methyl group, making these halogenated extender units potential electrophiles for active site nucleophilic residues [14]. From an engineering perspective, the ability to introduce chlorine and bromine atoms into the backbone of polyketides provides a valuable synthetic handle [9], enabling the production of polyketide libraries via a chemoenzymatic reaction.

Before testing the production of chlorinated and brominated polyketides, it was important to determine if rpMatB is capable of catalyzing the formation of chloro- and bromomalonate into their respective CoA thioesters. rpMatB has been shown to form malonyl-CoA derivatives larger than methylmalonyl-CoA [28] and has exhibited activity towards fluoromalonate [4]. As shown in *Table 4.2*, The catalytic efficiency of rpMatB formation of bromo- and chloromalonyl-CoA is similar to those of methylmalonate (for bromomalonyl-CoA) and ethylmalonyl-CoA (chloromalonyl-CoA) [28]. This result led us to attempt to directly isolate these two extender units and characterize them by MS and NMR. We were surprised to find that while chloromalonyl-CoA could be observed by MS, it appeared to decarboxylate in the reaction mixture (*Figure 4.1*). Additionally, bromomalonyl-CoA could never be detected, suggesting that bromomalonyl-CoA was too unstable (either hydrolytically or via decarboxylation) to analyze and isolate.

Table 4.2: *Tabulated kinetic parameters for both natural and unnatural extender units using rpMatB. Data for malonate, methylmalonate and ethylmalonate is referenced from [28] *and fluoromalonate parameters were obtained from unpublished data (Tom Privalsky).*

Extender unit	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
Chloromalonate	360 ± 60	0.56 ± 0.1	1.6 x 10^3
Bromomalonate	30 ± 4	6.0 ± 0.2	2.0 x 10^5
Fluoromalonate[*]	1.6 x 10^3 ± 100	2.0 ± 0.1	1.3 x 10^3
Malonate[28]	110 ± 8	18 ± 0.3	2.0 x 10^3
Methylmalonate[28]	81 ± 6	12 ± 0.2	1.0 x 10^5
Ethylmalonate[28]	1.3 x 10^3 ± 55	11 ± 0.2	9.0 x 10^3

The inability to detect bromomalonyl-CoA and the decomposition of chloromalonyl-CoA does not necessarily indicate that brominated and chlorinated polyketides could not be generated. If decomposition occurs at a slower time-scale than C-C bond formation, then chlorinated and brominated polyketides products could be synthesized. Additionally, the use of a malonyl-CoA regeneration system allows the addition of excess malonic acid derivatives while maintaining a controlled malonyl-CoA pool, making the decomposition of the extender units less costly. Chloromalonyl-CoA and bromomalony-CoA incorporation into polyketides was tested using the TKL formation assay for DEBS (Section 2.3). Surprisingly, chlorinated triketide lactones (Cl-TKLs) could be formed in all three DEBS modules tested, albeit at an

Figure 4.1: HPLC-UV (A_{260}) trace of a chloroacetyl-CoA standard and a chloromalonyl-CoA formation reaction with rpMatB and chloromalonate. As shown, chloromalonyl-CoA decomposes to form chloroacetyl-CoA.

order of magnitude lower yields than TKL (*Figure 4.2*) [2,4]. However, brominated triketide lactones (Br-TKLs) could not be observed. As mentioned above, these two extender units could potentially act as covalent inhibitors of active-site nucleophiles in DEBS modules. To test if covalent inhibition was responsible for the lack of Br-TKL formation and the low Cl-TKL yields, parallel TKL formations assays were set up with either chloromalonate, bromomalonate, or

Figure 4.2: Cl-TKL production with DEBS modules. (A) Extracted ion chromatograph and MS trace of enzymatically synthesized Cl-TKL. The MS trace shows the characteristic chlorine isotope distribution. (B) Production of Cl-TKL in modules 2,3, and 6 of DEBS (10 µM), chloromalonate (5 mM), CoA (1 mM with regeneration system), and NDK-SNAC (5 mM). When error bars are shown, data is represented as mean $\pm$ s.d (n =3). * Cl-TKL was quantified using a TKL standard curve (A_{260}).
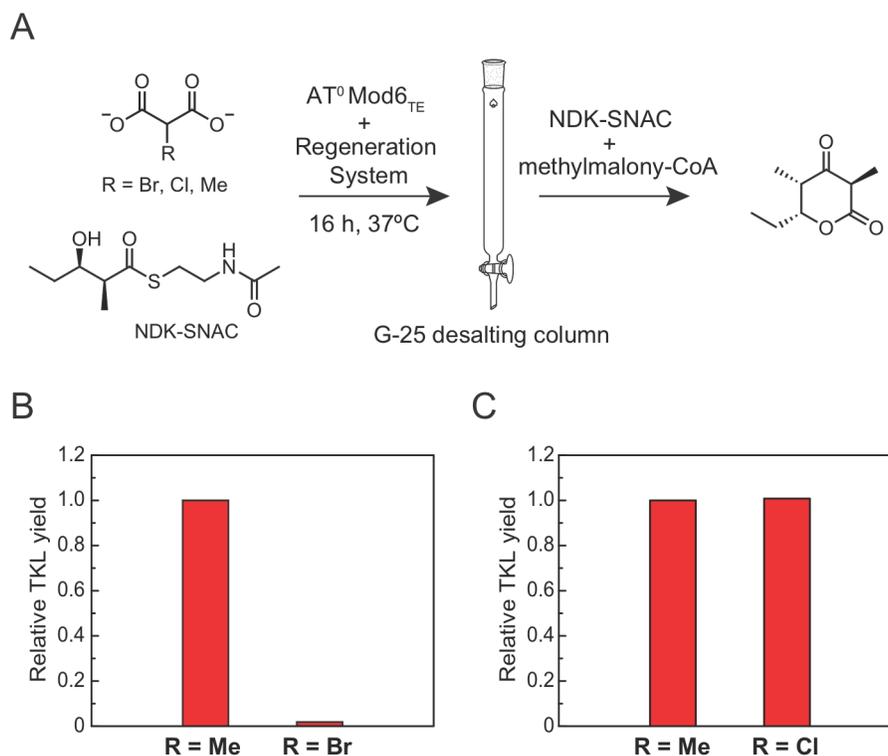
Figure 4.3: *Covalent inhibition assay of chloro- and bromomalonate with WT Mod6$_{TE}$. (A) Scheme representing the assay used to test chloro- and bromomalonyl-CoA inhibition of WT Mod6$_{TE}$. (B) A TKL assay containing bromomalonate (5mM) results in the covalent inhibition of WT Mod6$_{TE}$. (C) A TKL assay containing chloromalonate (5 mM) does not result in covalent inhibition of WT Mod6$_{TE}$.*

methylmalonate. After an overnight incubation, the WT Mod6$_{TE}$ protein was isolated using a G-25 desalting column and the protein was reincubated with NDK-SNAC and pure methylmalonyl-CoA overnight (*Figure 4.3A*) [*2*]. As observed in *Figure 4.3B*, the protein incubated with *in situ* generated bromomalonyl-CoA produced three orders of magnitude less TKL than the WT Mod6$_{TE}$ incubated with methylmalonyl-CoA. This data suggests that bromomalonyl-CoA acts as covalent inhibitor of DEBS. To better understand if this covalent inhibition was specific to an active-site residue, the same covalent inhibition reactions were carried out with bromomalonyl-CoA (presumably formed *in situ*) in the absence of NDK-SNAC, or simply with just bromomalonate. The covalent inhibition was observed under all three conditions, suggesting that bromomalonate was sufficient for covalent inhibition of DEBS protein. Ultimately, this data illustrated that bromomalonyl-CoA is not a good substrate (and may be an active-site inhibitor) for DEBS proteins given that bromomalonate or bromomalonyl-CoA appear and to alkylate and inactivate DEBS protein. Interestingly, when this covalent inhibition assay was carried out with chloromalonyl-CoA, the production of TKL after WT Mod6$_{TE}$ was isolated was essentially identical for protein preincubated with either chloromalonyl-CoA or methylmalonyl-CoA (*Figure 4.3C*). With knowledge that chloromalonyl-CoA does not inhibit DEBS proteins covalently, we set out to explore the nature of chloromalonyl-CoA incorporation into DEBS protein.

As shown in *Figure 4.2*, Cl-TKL yields appear to be almost equivalent in the presence of a module containing either an active or inactive AT (AT^0). Additionally, *Figure 4.4* demonstrates that Cl-TKL production in AT^0 Mod3$_{TE}$ is higher than TKL formation. This result is consistent with data observed previously [*2,4*], where fluoromalonyl-CoA incorporation levels appeared

unaffected by AT activity since the extender unit was able to directly diffuse into the KS active site and undergo C-C bond formation in an ACP-independent fashion. By examining Cl-TKL production with DEBS Mod3 proteins containing either inactive ATs or ACPs, it can be clearly determined that Cl-TKL formation occurs through this ACP-independent mechanism (*Table 4.3*). This behavior was also observed in modules 2 and 6, in similar fashion to fluoromalony-CoA incorporation [*4*]. The ACP-independent nature of chloromalonyl-CoA incorporation poses a major challenge when it comes to the generation of full-length chlorinated polyketides, since growing-chain offloading occurs after C-C bond formation, halting processive chain growth. However, it is important to mention that Cl-TKL formation appears to be produced at as much as 40-fold

Figure 4.4: *Relative yield of Cl-TKL and TKL production with chloro- or methylmalonate (5 mM), CoA (1 mM with regeneration system), NDK-SNAC (5 mM), and either AT^0 or WT Mod3$_{TE}$ (10 uM). * Cl-TKL was quantified using a TKL standard curve. Data is represented as mean ± error (error was propagated from individual standard deviations of each sample (n =3). ^{a} TKL production with Mod3$_{TE}$ was obtained from a previous study [4].*

higher yield than F-TKL formation [*4*]. The high yield of Cl-TKL through an ACP-independent mechanism is not too surprising given the steric similarity between methylmalonyl-CoA and chloromalonyl-CoA [*13*]. Additionally, the propensity of chloromalonyl-CoA to undergo a spontaneous decarboxylation (*Figure 4.1*) may explain the higher production of Cl-TKL compared to TKL via the ACP-independent mode of chain extension that occurs using AT^0 Mod3$_{TE}$.

The direct diffusion of chloromalonyl-CoA into the KS active site can be observed even when the dimodular mini-PKS is used. The incubation of chloromalonate and fluoromalonate with AT^0 Mod2, AT^0 Mod3$_{TE}$, and DszAT resulted in formation of a fluoro-chlorotriketide lactone (FCl-TKL) (*Figure 4.5*). It is important to mention that while the regiochemistry of FCl-TKL could not be confirmed directly, the lack of Cl-TKL yield enhancement in the presence of DszAT (*data not shown*) and the importance of ACP-dependent C-C bond formation in Mod2 of the mini-PKS [*4*], suggests that the connectivity shown in *Figure 4.5* is likely correct. The ability to incorporate chlorine into the backbone of TKLs, while not as immediately impactful as chlorine incorporation into the backbone of full-length polyketides, enables the use of Cl-TKLs as chiral building blocks for drug molecules of interest. For example, the backbone of the antitumor/antibiotic agent lankacidin C [*29,30*] is quite complex, requiring

Table 4.3: *Relative yield of Cl-TKL produced with various module 3 constructs (10 µM), chloromalonate (5 mM), CoA (1 mM, with regeneration system), and NDK-SNAC (5 mM). Data was obtained by integration of Cl-TKL peak at 260 nm.*

Construct	Relative yield
KS AT ACP	1.0
KS AT ACP	1.1
KS AT ACP	0.80
KS AT	0.84
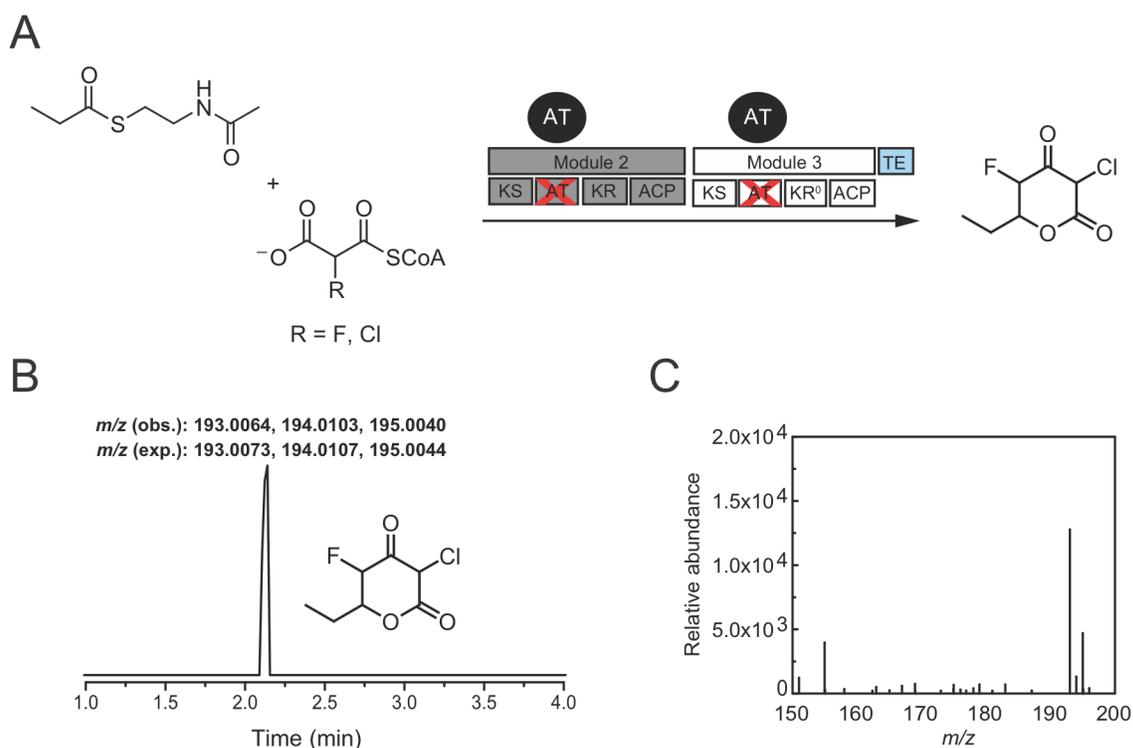
Figure 4.5: *FCl-TKL production using the mini-PKS (A) Scheme of mini-PKS experiment resulting in FCl-TKL formation. The reaction contained chloro- and fluomalonate (5 mM), CoA (0.1 mM, with regeneration system), propionyl-SNAC (5 mM), AT⁰ Mod2 (10 µM), AT⁰ Mod3_TE (10 µM), and DszAT (30 µM). (B) Extracted ion chromatogram of FCl-TKL. (C) Mass spectrum of FCl-TKL. The characteristic isotope distribution of the chlorine atom can be observed.*

four synthetic steps [*30*]. However, by chemoenzymatically generating a Cl-TKL containing a leaving group [*31*] in the PKS growing chain, this backbone can be assembled in one the enzymatic reaction with the mini-PKS (*Scheme 4.2*). Additionally, the ability to synthesize molecules such as FCl-TKL, enables the production of lankacidin C derivatives containing a fluorine substituent on its backbone.

The relatively high activity of ACP-independent chloromalonyl-CoA C-C bond formation may also prove a valuable tool in natural product identification and analysis. Recent studies have focused on the use of ethyl- protected malonyl carba(dethia)-*N*-acetyl cysteamines as nonhydrolyzable extender unit mimics that can off-load polyketide growing chains [*32-34*]. These mimics are also cell permeable, and can be used to determine the molecular structure and enzymes associated with poorly characterized polyketides in their native organism [*32*]. The off-loaded products can be analyzed using click-chemistry, if a handle is installed at the *N*-acetyl terminal, or using advanced metabolomics methods. However, if these analogs contained an α-chloro group, the unique isotope distribution of chlorine (*Figures 4.2 and 4.5*) can provide a biologically uncommon mass signature that can help identify the off-loaded intermediates. Additionally, the activated nature of chloromalonyl-CoA suggests that these chlorinated extender unit mimics may be more effective than the "WT" analogs, since they can compete with ACP-dependent C-C modes of bond formation.
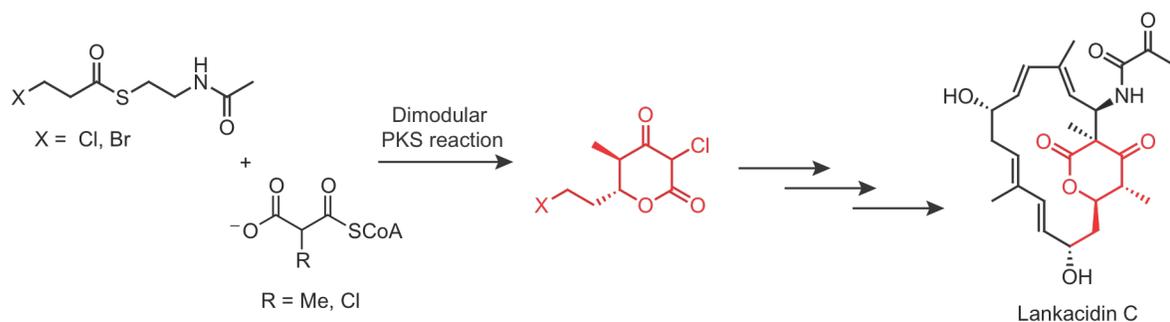
Scheme 4.2: *Proposed chemoenzymatic synthesis of lankacidin C using two PKS modules, chloromalonyl-CoA, and a starter unit containing a synthetic handle. The Cl-TKL derivative can serve as an upstream synthetic intermediate for lankacidin C total synthesis.*

Incorporation of α-disubstituted extender units into polyketides. Certain PKSs produce molecules containing germinal dimethylated moieties on their backbones, however, these functional groups are installed on acyl-ACPs using methyltransferases either prior [35] or after C-C bond formation [35,36]. The introduction of α-disubstituted extender units has remained unexplored in PKS such as DEBS since the incorporation of α-disubstituted extender units presents multiple challenges such as steric bulk and streochemical preferences. One set of substrates that have the potential to solve these two challenges are 2-fluoro-2-methylmalonayl-CoA and 2,2-difluoromalonyl-CoA. 2-fluoro-2-methylmalonayl-CoA can be considered a methylmalonyl-CoA analog, given the size similarity between hydrogen and fluorine atoms [37]. 2,2-difluoromalonyl-CoA may serve as an analog for malonyl-CoA since it lacks a chiral center at the α-carbon and contains two fluorine atoms instead of hydrogen atoms. The two fluorine atoms introduce more s-character to the carbon by pulling electrons from the p-orbital of the carbon, resulting in the two fluorine atoms having a bond angle that is even smaller than between two hydrogens [38]. As such, these two unnatural extender units show potential to be conservative substitutions for natural extender units, while enabling the insertion of pharmaceutically relevant fluorine atoms into the backbone of various polyketides. We initiated studies with 2-fluoro-2-methylmalonyl-CoA, given our extensive work with the methylmalonyl-CoA specific DEBS pathway [2,4]. The introduction of a geminal fluorine and methyl group enzymatically is of particular interest to the pharmaceutical community, as the polyketide backbone of the chemically synthesized solithromycin (currently
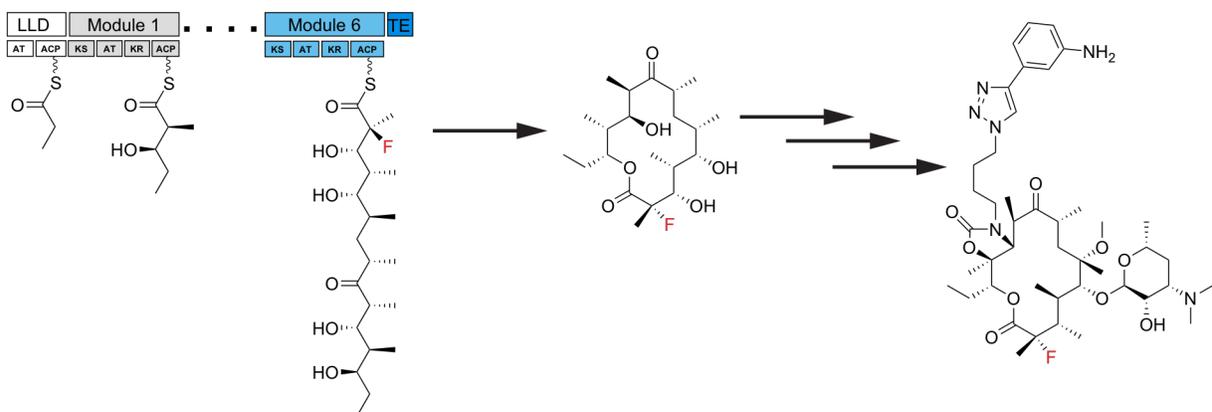
Scheme 4.3: *Proposed chemoenzymatic synthesis of solithromycin. By installing 2-fluoro-2-methylmalonate using the terminal module of DEBS, the polyketide backbone of solithromyin can be assembled. The rest of the functional groups can be installed using a combination of DEBS tailoring enzymes and chemical synthesis.*

A

m/z (obs.): 884.2
m/z (exp.): 884.1

2   4   6   8   10   12   14
Time (Min)
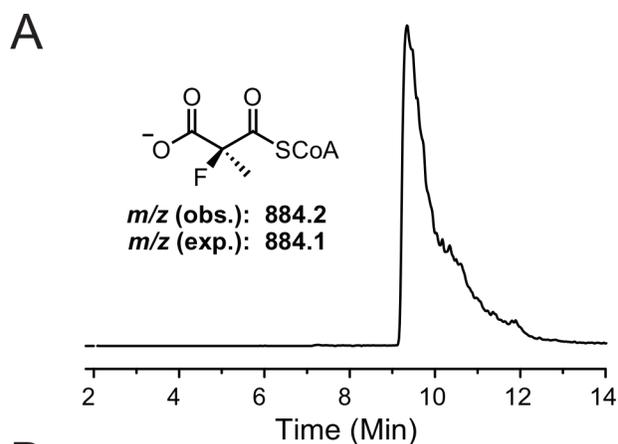
B

Extender Unit	Total product observed (%)
CoA-SH	56.5
2-fluoro-2-methylmalonyl-CoA	40.2
Methylmalonyl-CoA	3.3

Figure 4.6: *Formation of 2-fluoro-2-methylmalonyl-CoA. (A) Extracted ion chromatograph of 2-fluoro-2-methylmalonyl-CoA. (B) Tabulated data of conversion of 2-fluoro-2-methylmalonyl-CoA using rpMatB (20 μM), 2-fluoro-2-methylmalonate (5 mM), and CoA (1 mM) after a 90 min incubation at 37°C.*

under clinical trials) [10,11], can be biosynthesized using a DEBS biosynthetic cluster that incorporates 2-fluoro-2-methylmalonyl-CoA into its terminal module (*Scheme 4.3*). 2-fluoro-2-methylmalonate was synthesized from diethylmethylmalonate via a titanium tetrachloride catalyzed fluorination with Selectfluor [26] and was isolated after a saponification reaction. The formation 2-fluoro-2-methylmalonyl-CoA was then carried out using rpMatB. Given the promiscuous nature of rpMatB, it was reasonable to assay 2-fluoro-2-methylmalonyl-CoA formation, even though rpMatB activity has not been reported with α-disubstituted malonate derivatives. As observed in *Figure 4.6*, 2-fluoro-2-methylmalonyl-CoA can be synthesized using rpMatB, however the reaction is very slow, turning-over only 40% of the available CoA substrate in 90 min (full conversion is observed for methyl-, fluoro-, or malonyl-CoA in that time scale). *Figure 4.6B* also shows trace amounts of methylmalonyl-CoA is produced. This is consistent with the ^{1}H NMR analysis of 2-fluoro-2-methylmalonate, showing trace amounts of unreacted diethylmethylmalonate starting material. This trace amount of methylmalonate can then be rapidly converted into methylmalonyl-CoA using rpMatB, as expected [28]. Aside from the low conversion rate, the use of MatB to make 2-fluoro-2-methylmalonyl-CoA poses an additional challenge, since all MatB homologs produce (R)-methylmalonyl-CoA and PKSs use (S)-methylmalonyl-CoA. This is not a challenge with α-monosubstituted malonyl-CoA derivatives when the reaction is coupled with a methylmalonyl-CoA epimerase [39] or when the extender unit is incubated in the aqueous reaction mixture for a few hours, leading to chemical epimerization. However, the α-disubstituted nature of 2-fluoro-2-methylmalonyl-CoA makes this epimerization impossible, resulting in the formation of an extender unit with the incorrect KS substrate stereochemistry. Unsurprisingly, when the TKL formation assay was carried out using 2-fluoro-2-methylmalonate, rpMatB, and the CoA regeneration system, the expected TKL was not observed. The desired TKL could be observed, albeit at trace amounts, when racemic 2-fluoro-2-methylmalonyl-CoA was chemically synthesized (Monica, Neugebauer, unpublished data). This result suggests that if 2-fluoro-2-methylmalonyl-CoA can be generated with the proper stereochemistry, polyketides containing germinal fluorine and methyl substitutions may be biosynthesized. To this end, efforts to change the stereochemical preference of rpMatB could enable a fully enzymatic route to these targets. As an alternative, 2-fluoro-2-methylmalonyl-CoA can be synthesized with the desired stereochemistry through the conversion of 2-fluoropropionate into 2-fluoropropionyl-CoA using a propionyl-CoA synthetase

(PrpE) [*40*], followed by the propionyl-CoA carboxylase (Pcc) catalyzed carboxylation [*41*] to form 2-fluoro-2-methylmalonyl-CoA (*Scheme 4.4*).

2,2-difluoromalonate could easily be obtained through the saponification of the commercially available diethyldifluoromalonate. However, 2,2-difluoromalonyl-CoA could never be observed when 2,2-difluoromalonate was incubated with rpMatB. Whether this was due to the hydrolytic instability of 2,2-difluoromalonyl-CoA resulting from fluorine activation of the thioester, or the inability to attack the carbonyl carbon due to fluorine electrostatic interactions has yet to be determined. When 2,2-difluoromalonate was used in a TKL formation assay reaction mixture, the desired TKL was not observed, as expected. Based on these two results, it appears that 2,2-difluoromalonate may be a challenging substrate to use for the introduction of difluorinated extender units into the backbone of polyketides. The ability to load ACPs with 2,2-difluoromalonate using malonyl-specific ATs has not been explored, and may be a better setting for 2,2-difluoromalonate incorporation. However, this depends on whether or not 2,2-difluoromalonyl-CoA can be formed without hydrolyzing prior to forming an acyl-enzyme intermediate with the AT.
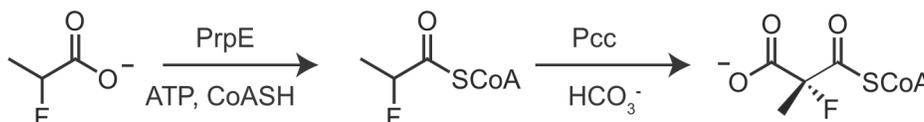
Scheme 4.4: *Alternate 2-fluoro-2-methylmalonyl-CoA biosynthetic route using a propionate synthetase (PrpE) and propionyl-CoA carboxylase (Pcc). Unlike formation 2-fluoro-2-methylmalonyl-CoA with MatB, this biosynthetic route results in formation of an extender unit with the proper stereochemistry for PKS C-C bond formation.*

Enzymatic synthesis of 4-fluoroethylmalonyl-CoA. 4-fluoroethylmalonyl-CoA has previously been incorporated into salinosporamide in place of the native 4-chloroethylmalonyl-CoA substrate [*22*]. This extender unit also has the potential to be a substrate for DEBS proteins, since DEBS modules have shown to incorporate ethylmalonyl-CoA efficiently [*5*]. We observed that ethylmalonyl-CoA incorporation using module 3 of DEBS proceeds via an ACP-dependent mode of C-C bond formation, since a dramatic drop in ethyltriketide lactone (Et-TKL) yield is observed when AT^0 or AT^0ACP^0 $Mod3_{TE}$ are used instead of WT $Mod3_{TE}$ (*Figure 4.7*). The promiscuity towards larger extender units, and ethylmalonyl-CoA in particular, may be attributed to the fact that *Saccharopolyspora erythraea* does not possess a crotonyl-CoA carboxylase (Ccr). The DEBS ATs have likely not evolved to select against ethylmalonyl-CoA, since it is not present in *S. erythraea* endogenously [*42*]. Seeking to capitalize on this promiscuity, we set out to synthesize 4-fluoroethylmalonyl-CoA and determine if it can be incorporated into TKLs produced using DEBS modules.

4-fluoroethylmalonyl-CoA was synthesized enzymatically using a chimeric pathway containing a thiolase (NphT7), KR (PhaB), DH (PhaJ), and an enoylreductase/carboxylase (Ccr) [*23*] (*Figure 4.8A*). We chose to use a thiolase that condenses malonyl-CoA and acetyl-CoA, instead of a thiolase that condenses two acetyl-CoA units [*24*], in order to control the regiochemistry of fluorine-incorporation and form exclusively 4-fluoracetoacetyl-CoA (*Figure 4.8A*). As shown in *Figure 4.8*, *in vitro* constitution of this pathway produces both ethylmalonyl-CoA and 4-fluoroethylmalonyl-CoA. Additionally, pathway intermediates are observed when the appropriate enzymes are omitted from the assay mixture (*Figure 4.8*). It is important to note that bicarbonate is required for Ccr catalysis, but has been reported to inhibit decarboxylative Claisen condensations [*43,44*]. A bicarbonate titration experiment found that 10 mM of bicarbonate can

produce sufficient acetoacetyl-CoA, while still allowing the downstream carboxylation reaction (*Figure 4.8F*). The ability to produce 4-fluoroethylmalonyl-CoA using this chimeric pathway suggested that 4-fluoroethylmalonyl-CoA can be generated *in situ* and may be incorporated into polyketides by incubating the reaction mixture with NDK-SNAC and a PKS module.

Production of TKLs using 4-fluoroethylmalonyl-CoA. After the *in situ* generation of 4-fluoroethylmalonyl-CoA, WT $Mod3_{TE}$, NDK-SNAC, and rpMatB along with the ATP regeneration system were added to the reaction mixture. While the mass corresponding to fluoroethyltriketide lactone could not be observed, two masses corresponding to bicyclic triketide lactone products were observed (*Figure 4.9*). The propensity of five-membered-ring cyclization presumably leads to the elimination of fluoride, resulting in two alternate products. These products represent successful incorporation of 4-fluoroethylmalonyl-CoA into DEBS polyketides. It is important to note that cyclization may not necessarily occur in the context of a full-length polyketide, as 4-fluoroethylmalonyl-CoA incorporation does not result in salinosporamide cyclization [22]. If cyclization does occur in all instances of 4-fluoroethylmalonyl-CoA incorporation, the introduction of a rigid 5-membered ring into natural products and natural product sacaffolds affords the formation of a unique functional group that may result in molecules with improved bioactivity. Additionally, we found that C-C bond formation was dependent on the presence of an active AT (*Figure 4.9C*), as shown by the drastically decreased product yield observed when AT^0 $Mod3_{TE}$ is used in the TKL formation assay. We next turned our attention to the efficiency of 4-fluoroethylmalonyl-CoA ACP loading by directly measuring the level of 4-fluoroethylmalonyl-ACP (*Figure 4.10*). We were surprised to find that the percent occupancy of 4-fluoroethylmalonyl-ACP was only 9% (*Figure 4.10B*). However, we also measured acyl-ACP occupancy using *in situ* generated ethylmalonyl-CoA and found the occupancy to be 31%. Given that ethylmalonyl-CoA is known to be a good substrate for module 3 [5] (*Figure 4.7B*), we attribute the low level of occupancy using both substrates to be due to the fact that the substrates had to be generated by a four-enzyme pathway prior to incubating with the PKS module. As such, the 3-fold drop in occupancy is encouraging, suggesting that if the 4-

Figure 4.7: *Et-TKL production with Mod3$_{TE}$ variants (10 μM), ethylmalonate (5 mM), CoA (1 mM, with regeneration system), and NDK-SNAC (5 mM).*Et-TKL was quantified using a TKL standard curve (A$_{260}$). Production of Et-TKL likely higher than shown for WT Mod3$_{TE}$ due to UV signal saturation. Data reported as mean $\pm$ s.d (n =3).*

fluoromalonyl-CoA pool could be improved, the ACP occupancy will likely rise. Additionally, it was encouraging to see that when AT^0 $Mod3_{TE}$ was used, 4-fluoromalonyl-ACP could not be detected. This result suggests that the 9% occupancy observed with WT $Mod3_{TE}$ is catalyzed by the AT and is consistent with the production data shown in *Figure 4.9*.
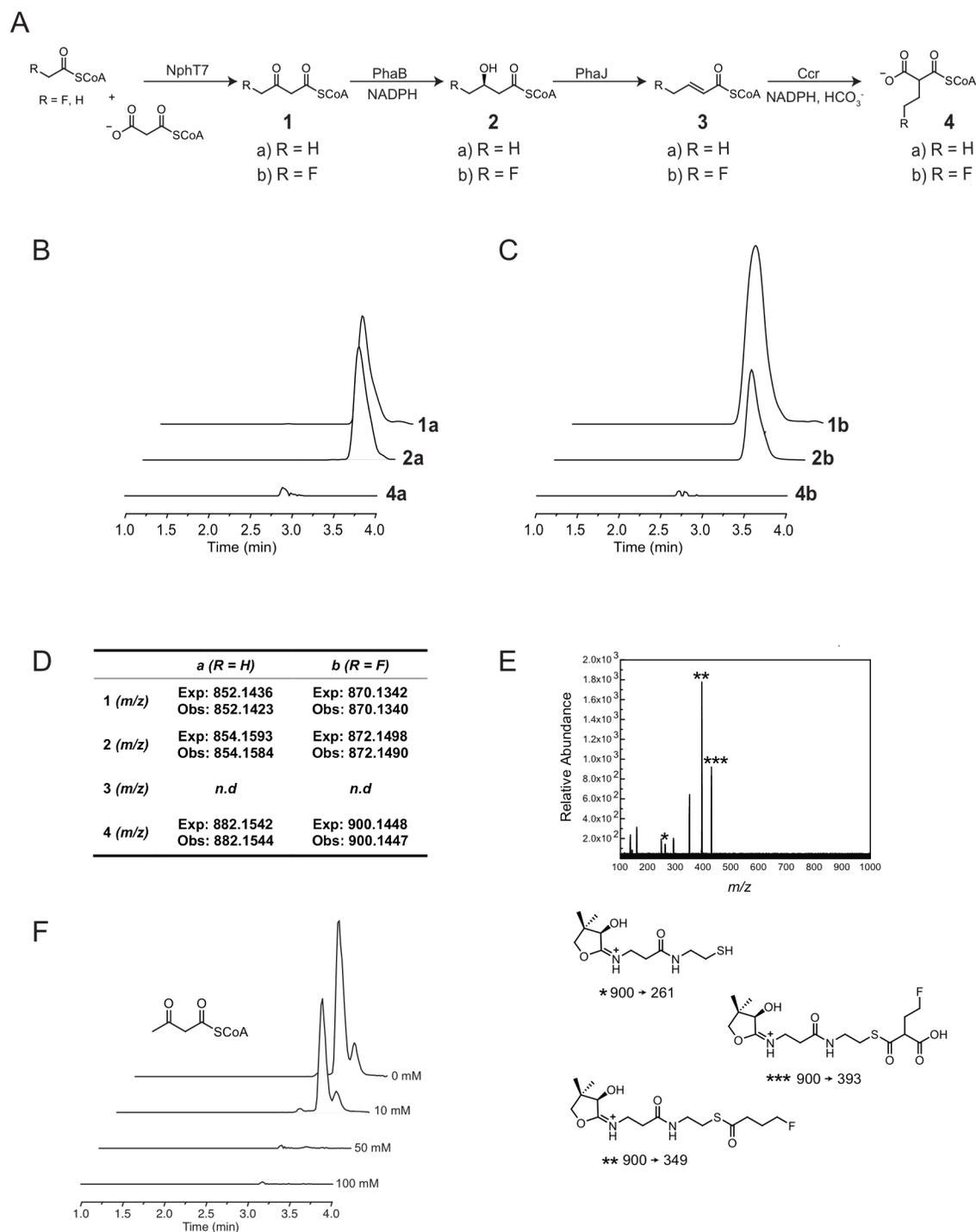
Figure 4.8: *Enzymatic synthesis of 4-fluoroethylmalonyl-CoA. (A) Scheme representing the chimeric pathway used to produce ethylmalonyl-CoA and 4-fluorethylomalonyl-CoA. The extracted ion chromatograms of pathway intermediates and the final product (B) ethylmalonyl-CoA and (C) 4-fluoroethylmalonyl-CoA were obtained by performing assays with only the enzymes required to form each individual product. Compounds 3a and 3b were not detected. (D) Exact masses of the final products and pathway intermediates. (E) Product ion spectra for 4-fluoroethylmalonyl-CoA. (F) Extracted ion chromatograms of acetoacetyl-CoA produced under varying bicarbonate concentrations. Peak height is normalized to the 0 mM bicarbonate extracted ion chromatogram. n.d = not detected*

The ability to generate 4-fluoromalonyl-CoA using a chimeric pathway and fluoroacetate, along with the capacity to incorporate this extender unit into DEBS polyketides, makes this pathway promising tool for the *in vivo* production of novel polyketides. The 4-fluoroethylmalonyl-CoA pathway can be integrated into a heterologous host with knocked-down methyl- and ethylmalonyl-CoA production pathways and form fluorinated or bicyclized full-length polyketides. 4-fluoroethylmalonyl-CoA is particularly exciting for applications involving ethylmalonyl-CoA selective *trans*-ATs such as KirCII [45,46]. KirCII is responsible for incorporating ethylmalonyl-CoA exclusively into module 5 of the RNA translation inhibitor kirromycin [45,46]. Propargyl- and allyl-kirromycins have been generated using propargyl- and allylmalonyl-CoA, so it is conceivable to predict that 4-fluoroethylmalonyl-CoA would serve as a great substrate for such systems.
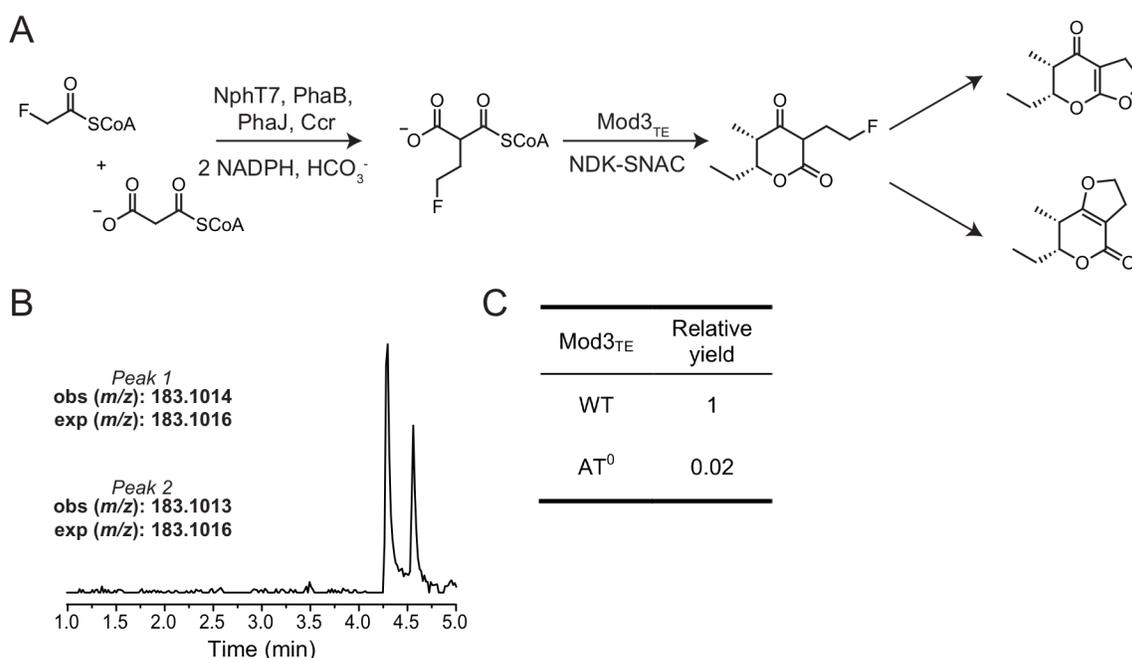
Figure 4.9: *Formation of polyketides with 4-fluoroethylmalonyl-CoA. (A) Scheme representing production of the cyclized TKLs using Mod3$_{TE}$ (5 µM), NDK-SNAC (2.5 mM), and in situ generated 4-fluoroethylmalonyl-CoA. (B) Extracted ion chromatogram of the two cyclized TKLs observed. (C) tabulated representation of product formation with AT^{0}or WT Mod3$_{TE}$. The 50-fold drop in production with the AT^0 mutant suggests that 4-fluoroethylmalonyl-CoA occurs via an ACP-independent mechanism.*

A

holo ACP
ethylmalonyl-ACP
4-fluoroethylmalonyl-ACP

20 22 24 26 28 30 32 34 36 38 40
Time (min)

B

Extender Unit	WT Mod3$_{TE}$	AT^0 Mod3$_{TE}$
Ethylmalonyl-CoA	31 %	-
4-fluoroethylmalonyl-CoA	9 %	n.d

C

1

2a R = H
2b R = F

3a R = H
3b R = F

	1 (m/z)	2 (m/z)	3 (m/z)
a	Obs: 261.1292 Exp: 261.1267	Obs: 331.1706 Exp: 331.1686	Obs: 375.1576 Exp: 375.1584
b	Obs: 261.1259 Exp: 261.1267	Obs: 349.1601 Exp: 349.1592	Obs: 393.1491 Exp: 393.1490
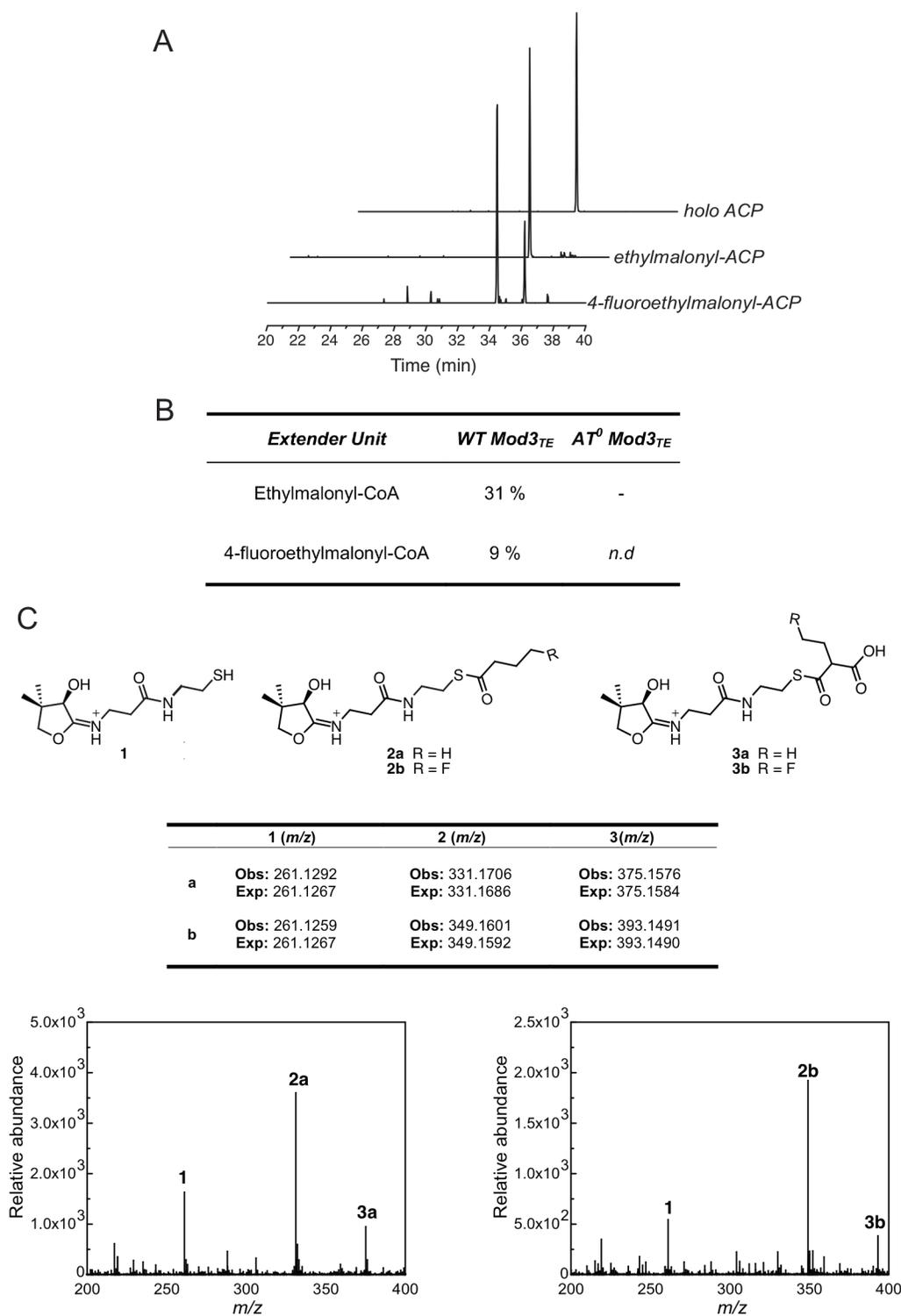
Figure 4.10: ACP acitive site occupancy with ethylmalonyl-CoA derivatives. (A) Extracted ion chromatograph of the holo and acyl-ACPs. (B) Active site occupancy of either ethylmalonyl-CoA or 4-fluoromalonyl-CoA (both generated in situ) and either AT^0 or WT Mod3$_{TE}$ (1 µM). Active site occupancy is calculated as described above (Section 2.2). (C) MS/MS characterization of acyl-ACP using the phosphopantetheine ejection method described above (Section 2.2). n.d = not detected, - = not tested

4.4 Conclusion

The production of Cl-TKL illustrated that chloromalonyl-CoA can be incorporated into DEBS modules. However, we found that C-C bond formation with this extender unit occurs in an ACP-independent fashion, suggesting that chloromalonyl-CoA incorporation into full-length polyketides is not feasible. However, robust production of ACP-independent C-C bond formation with chloromalonyl-CoA can introduce chlorine atoms into the terminal position of truncated polyketide products, resulting in the formation of potentially useful stereopure building blocks containing a synthetic handle on their backbone. Although bromomalonate failed to produce TKLs and was determined to be a covalent inhibitor of DEBS modules, its functionality is similar to that of a chlorine atom and similar chemistries can be used to chemically modify a molecule with either a chlorine or bromine synthetic handle [9]. Future chemo-enzymatic research efforts can use chlorinated polyketide building blocks to reduce complex synthetic steps in the total synthesis of medicinally relevant molecules.

The α-disubstituted extender units proved to be poor substrates for malonyl-CoA activating enzymes as well as DEBS modules. 2,2-difluromalonyl-CoA was never directly detected and could never be incorporated into TKLs. While rpMatB was able to generate 2-fluoro-2-methylmalonyl-CoA, the product formed possessed the incorrect stereochemistry preferred by DEBS KSs and failed to produce TKLs. Ultimately, the ability to generate 2-fluoro-2-methylmalonyl-CoA with MatB opens the door for engineering a variant that produces the correct stereoisomer for polyketide biosynthesis. Additionally, biosynthesis of methylmalonyl-CoA with PrpE and Pcc (*Scheme 4.4*) has the potential to produce the KS-preferred stereoisomer and merits investigation as a potential system for 2-fluoro-2-methylmalonyl-CoA biosynthesis.

The production of 4-fluoroethylmalonyl-CoA using a chimeric biosynthetic pathway and incorporation of this extender unit into Mod3$_{TE}$ ultimately resulted in the formation of bicyclic TKL through a fluoride elimination. The ability to generate 4-fluoroethylmalonyl-CoA in quantities sufficient for TKL production suggests that this chimeric pathway shows potential as a method for the generation and incorporation of a fluorinated PKS extender unit *in vivo*.

The unnatural extender units investigated in this study were all selected given their potential to introduce novel functionality into polyketides. Using chloro- and bromomalonate, we attempted to introduce synthetic handles directly on the backbone of the polyketide. The α-disubstituted extender units were studied in efforts to improve the pharmacokinetic properties and hydrolytic stability of the polyketides produced. Finally, 4-fluoromalonyl-CoA was studied since it can be synthesized enzymatically and has potential to enable the formation of bicyclic polyketides through the elimination of fluoride. The findings reported here lay the ground work for future engineering efforts involving the incorporation of α-halogenated and α-disubstituted extender units into the backbone of valuable natural product scaffolds.

4.5   References

1.   Khosla, C., Gokhale, R. S., Jacobsen, J. R., Cane, D. E. Tolerance and specificity of polyketide synthases. *Annu Rev Biochem* **1999**, *68* 219-253.

2.	Walker, M. C., Thuronyi, B. W., Charkoudian, L. K., Lowry, B., Khosla, C., Chang, M. C. Expanding the fluorine chemistry of living systems using engineered polyketide synthase pathways. *Science* **2013**, *341* (6150) 1089-1094.

3.	Koryakina, I., McArthur, J. B., Draelos, M. M., Williams, G. J. Promiscuity of a modular polyketide synthase towards natural and non-natural extender units. *Org Biomol Chem* **2013**, *11* (27) 4449-4458.

4.	Ad, O., Thuronyi, B. W., Chang, M. C. Elucidating the mechanism of fluorinated extender unit loading for improved production of fluorine-containing polyketides. *Proc Natl Acad Sci U S A* **2017**, *114* (5) E660-E668.

5.	Lowry, B., Robbins, T., Weng, C. H., O'Brien, R. V., Cane, D. E., Khosla, C. *In vitro* reconstitution and analysis of the 6-deoxyerythronolide B synthase. *J Am Chem Soc* **2013**, *135* (45) 16809-16812.

6.	Koryakina, I., Kasey, C., McArthur, J. B., Lowell, A. N., Chemler, J. A., Li, S., Hansen, D. A., Sherman, D. H., Williams, G. J. Inversion of extender unit selectivity in the erythromycin polyketide synthase by acyltransferase domain engineering. *ACS Chem Biol* **2017**, *12* (1) 114-123.

7.	Sundermann, U., Bravo-Rodriguez, K., Klopries, S., Kushnir, S., Gomez, H., Sanchez-Garcia, E., Schulz, F. Enzyme-directed mutasynthesis: A combined experimental and theoretical approach to substrate recognition of a polyketide synthase. *ACS Chem Biol* **2013**, *8* (2) 443-450.

8.	Bravo-Rodriguez, K., Klopries, S., Koopmans, K. R., Sundermann, U., Yahiaoui, S., Arens, J., Kushnir, S., Schulz, F., Sanchez-Garcia, E. Substrate flexibility of a mutated acyltransferase domain and implications for polyketide biosynthesis. *Chem Biol* **2015**, *22* (11) 1425-1430.

9.	Hutchinson, C. R., McDaniel, R. Combinatorial biosynthesis in microorganisms as a route to new antimicrobial, antitumor and neuroregenerative drugs. *Curr Opin Investig Drugs* **2001**, *2* (12) 1681-1690.

10.	Putnam, S. D., Castanheira, M., Moet, G. J., Farrell, D. J., Jones, R. N. Cem-101, a novel fluoroketolide: Antimicrobial activity against a diverse collection of gram-positive and gram-negative bacteria. *Diagn Microbiol Infect Dis* **2010**, *66* (4) 393-401.

11.	Putnam, S. D., Sader, H. S., Farrell, D. J., Biedenbach, D. J., Castanheira, M. Antimicrobial characterisation of solithromycin (CEM-101), a novel fluoroketolide: Activity against *Staphylococci* and *Enterococci*. *Int J Antimicrob Agents* **2011**, *37* (1) 39-45.

12.	Bott, G., Field, L. D., Sternhell, S. Steric effects - a study of a rationally designed system. *J Am Chem Soc* **1980**, *102* (17) 5618-5626.

13.	Patani, G. A., LaVoie, E. J. Bioisosterism: A rational approach in drug design. *Chem Rev* **1996**, *96* (8) 3147-3176.

14.	Prorok, M., Albeck, A., Foxman, B. M., Abeles, R. H. Chloroketone hydrolysis by chymotrypsin and n-methylhistidyl-57-chymotrypsin: Implications for the mechanism of chymotrypsin inactivation by chloroketones. *Biochemistry* **1994**, *33* (32) 9784-9790.

15. Shao, P. P., Ye, F., Chakravarty, P. K., Herrington, J. B., Dai, G., Bugianesi, R. M., Haedo, R. J., Swensen, A. M., Warren, V. A., Smith, M. M., Garcia, M. L., McManus, O. B., Lyons, K. A., Li, X., Green, M., Jochnowitz, N., McGowan, E., Mistry, S., Sun, S. Y., Abbadie, C., Kaczorowski, G. J., Duffy, J. L. Improved cav2.2 channel inhibitors through a gem-dimethylsulfone bioisostere replacement of a labile sulfonamide. *ACS Med Chem Lett* **2013**, *4* (11) 1064-1068.

16. Burkhard, J. A., Wuitschik, G., Rogers-Evans, M., Muller, K., Carreira, E. M. Oxetanes as versatile elements in drug discovery and synthesis. *Angew Chem Int Ed Engl* **2010**, *49* (48) 9052-9067.

17. Wuitschik, G., Carreira, E. M., Wagner, B., Fischer, H., Parrilla, I., Schuler, F., Rogers-Evans, M., Muller, K. Oxetanes in drug discovery: Structural and synthetic insights. *J Med Chem* **2010**, *53* (8) 3227-3246.

18. Piccinelli, G., Fernandes, P., Bonfanti, C., Caccuri, F., Caruso, A., De Francesco, M. A. *In vitro* activity of solithromycin against erythromycin-resistant *Streptococcus agalactiae*. *Antimicrob Agents Chemother* **2014**, *58* (3) 1693-1698.

19. Fernandes, P., Martens, E., Bertrand, D., Pereira, D. The solithromycin journey-it is all in the chemistry. *Bioorg Med Chem* **2016**, *24* (24) 6420-6428.

20. Hugenberg, V., Wagner, S., Kopka, K., Schober, O., Schafers, M., Haufe, G. Synthesis of geminal difluorides by oxidative desulfurization-difluorination of alkyl aryl thioethers with halonium electrophiles in the presence of fluorinating reagents and its application for ^{18}F-radiolabeling. *J Org Chem* **2010**, *75* (18) 6086-6095.

21. Prakash, G. K., Hu, J. Selective fluoroalkylations with fluorinated sulfones, sulfoxides, and sulfides. *Acc Chem Res* **2007**, *40* (10) 921-930.

22. Eustaquio, A. S., O'Hagan, D., Moore, B. S. Engineering fluorometabolite production: Fluorinase expression in *Salinispora tropica* yields fluorosalinosporamide. *J Nat Prod* **2010**, *73* (3) 378-382.

23. Bond-Watts, B. B., Bellerose, R. J., Chang, M. C. Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat Chem Biol* **2011**, *7* (4) 222-227.

24. Okamura, E., Tomita, T., Sawa, R., Nishiyama, M., Kuzuyama, T. Unprecedented acetoacetyl-Coenzyme A synthesizing enzyme of the thiolase superfamily involved in the mevalonate pathway. *Proc Natl Acad Sci U S A* **2010**, *107* (25) 11265-11270.

25. Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **2001**, *291* (5509) 1790-1792.

26. Mohr, J. T., Nishimata, T., Behenna, D. C., Stoltz, B. M. Catalytic enantioselective decarboxylative protonation. *J Am Chem Soc* **2006**, *128* (35) 11348-11349.

27. Dunn, B. J., Watts, K. R., Robbins, T., Cane, D. E., Khosla, C. Comparative analysis of the substrate specificity of *trans-* versus *cis-*acyltransferases of assembly line polyketide synthases. *Biochemistry* **2014**, *53* (23) 3796-3806.

28. Crosby, H. A., Rank, K. C., Rayment, I., Escalante-Semerena, J. C. Structure-guided expansion of the substrate range of methylmalonyl coenzyme a synthetase (MatB) of *Rhodopseudomonas palustris*. *Appl Environ Microbiol* **2012**, *78* (18) 6619-6629.

29. Cao, Z., Yoshida, R., Kinashi, H., Arakawa, K. Blockage of the early step of lankacidin biosynthesis caused a large production of pentamycin, citreodiol and epi-citreodiol in *Streptomyces rochei*. *J Antibiot (Tokyo)* **2015**, *68* (5) 328-333.

30. Kende, A. S., Liu, K., Kaldor, I., Dorey, G., Koch, K. Total synthesis of the macrolide antitumor antibiotic lankacidin-C. *Journal of the American Chemical Society* **1995**, *117* (31) 8258-8270.

31. Murli, S., MacMillan, K. S., Hu, Z., Ashley, G. W., Dong, S. D., Kealey, J. T., Reeves, C. D., Kennedy, J. Chemobiosynthesis of novel 6-deoxyerythronolide B analogues by mutation of the loading module of 6-deoxyerythronolide B synthase 1. *Appl Environ Microbiol* **2005**, *71* (8) 4503-4509.

32. Riva, E., Wilkening, I., Gazzola, S., Li, W. M. A., Smith, L., Leadlay, P. F., Tosin, M. Chemical probes for the functionalization of polyketide intermediates. *Angew Chem Int Edit* **2014**, *53* (44) 11944-11949.

33. Tosin, M., Smith, L., Leadlay, P. F. Insights into lasalocid a ring formation by chemical chain termination *in vivo*. *Angew Chem Int Edit* **2011**, *50* (50) 11930-11933.

34. Tosin, M., Demydchuk, Y., Parascandolo, J. S., Per, C. B., Leeper, F. J., Leadlay, P. F. *In vivo* trapping of polyketide intermediates from an assembly line synthase using malonyl carba(dethia)-N-acetyl cysteamines. *Chem Commun* **2011**, *47* (12) 3460-3462.

35. Poust, S., Phelan, R. M., Deng, K., Katz, L., Petzold, C. J., Keasling, J. D. Divergent mechanistic routes for the formation of gem-dimethyl groups in the biosynthesis of complex polyketides. *Angew Chem Int Ed Engl* **2015**, *54* (8) 2370-2373.

36. Miller, D. A., Luo, L., Hillson, N., Keating, T. A., Walsh, C. T. Yersiniabactin synthetase: A four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem Biol* **2002**, *9* (3) 333-344.

37. Smart, B. E. Fluorine substituent effects (on bioactivity). *J Fluorine Chem* **2001**, *109* (1) 3-11.

38. O'Hagan, D. Understanding organofluorine chemistry. An introduction to the C-F bond. *Chem Soc Rev* **2008**, *37* (2) 308-319.

39. Murli, S., Kennedy, J., Dayem, L. C., Carney, J. R., Kealey, J. T. Metabolic engineering of *Escherichia coli* for improved 6-deoxyerythronolide B production. *J Ind Microbiol Biotechnol* **2003**, *30* (8) 500-509.

40. Horswill, A. R., Escalante-Semerena, J. C. Characterization of the propionyl-CoA synthetase (PrpE) enzyme of *Salmonella enterica*: Residue Lys592 is required for propionyl-AMP synthesis. *Biochemistry* **2002**, *41* (7) 2379-2387.

41. Bramwell, H., Hunter, I. S., Coggins, J. R., Nimmo, H. G. Propionyl-CoA carboxylase from *Streptomyces coelicolor* A3(2): Cloning of the gene encoding the biotin-containing subunit. *Microbiology* **1996**, *142 (Pt 3)* 649-655.

42.   Stassi, D. L., Kakavas, S. J., Reynolds, K. A., Gunawardana, G., Swanson, S., Zeidner, D., Jackson, M., Liu, H., Buko, A., Katz, L. Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering. *Proc Natl Acad Sci U S A* **1998**, *95* (13) 7305-7309.

43.   Witkowski, A., Joshi, A. K., Lindqvist, Y., Smith, S. Conversion of a beta-ketoacyl synthase to a malonyl decarboxylase by replacement of the active-site cysteine with glutamine. *Biochemistry* **1999**, *38* (36) 11643-11650.

44.   Witkowski, A., Joshi, A. K., Smith, S. Mechanism of the beta-ketoacyl synthase reaction catalyzed by the animal fatty acid synthase. *Biochemistry* **2002**, *41* (35) 10877-10887.

45.   Musiol-Kroll, E. M., Zubeil, F., Schafhauser, T., Hartner, T., Kulik, A., McArthur, J., Koryakina, I., Wohlleben, W., Grond, S., Williams, G. J., Lee, S. Y., Weber, T. Polyketide bioderivatization using the promiscuous acyltransferase KirCII. *ACS Synth Biol* **2017**, *6* (3) 421-427.

46.   Ye, Z., Musiol, E. M., Weber, T., Williams, G. J. Reprogramming acyl carrier protein interactions of an acyl-CoA promiscuous *trans*-acyltransferase. *Chem Biol* **2014**, *21* (5) 636-646.

Appendix 1: *Plasmids and oligonucleotides*

Table 1. A) Plasmid constructs, B) oligonucleotides, and C) gBlocks used in Chapter 2.

A

Name	No.	Primers	Restriction sites	Method	Description
pET28a-His$_6$-KSAT2	1492	KSAT2 F1/R1	BsaI-BsaI	Golden gate	KSAT didomain of DEBS module 2.
pET21c-Mod3$_{TE}$ ACP^0-His$_6$	1566	Mod3TE ACP0 F1/R1 and F2/R2	BsiWI-EcoRI	Gibson	Based on pRSG34 with a S$\rightarrow$A mutation to ACP active site serine.
pET21c-Mod3$_{TE}$ AT^0 ACP^0-His$_6$	1567	Mod3TE ACP0 F1/R1 and F2/R2	BsiWI-EcoRI	Gibson	Based on pAYC136 with a S$\rightarrow$A mutation to ACP active site serine.
pET21c-KSAT^0-His$_6$	1904	KS^{0}AT^0 F1/R2	NdeI-BsiWI	Gibson	Based on pAYC02 with S$\rightarrow$A mutation in active site of AT.
pET21c-KS^{0}AT-His$_6$	1903	KS^{0}AT^0 F1/R1 and F2/R2	NdeI-BsiWI	Gibson	Based on pAYC02 with C$\rightarrow$A mutation in active site of KS.
pET21c-KS^{0}AT^0-His$_6$	1905	KS^{0}AT^0 F1/R1 and F2/R2	NdeI-BsiWI	Gibson	Based on pAYC02 with C$\rightarrow$A mutation in active site of KS and S$\rightarrow$A mutation to the active site of AT
pET28a-His$_6$-PIKSKR1	1324	PIKS_KR F1/R1	BsaI-BsaI	Golden gate/Gibson	Synthetic pikromycin KR 1. gblocks assembled using sequential GG to gibson.
pET16B-His$_{10}$-rpMatB	1192				Ben Thuronyi Thesis: Appendix 1
pFW3	1186				Disorazole *trans*-AT from Wong, F. T., et. al., *Biochemistry* **2009**, 49 (1) 95-102.
pSV272- His$_6$-MBP-DEBS$_{Mod2}$	1396				MBP-tagged Module 2 from Walker, M. C., et. al., *Science* **2013**, 341, 1089-1094.
pSV272- His$_6$-MBP-DEBS$_{Mod2}$	1341				MBP-tagged Module 2 AT^0 from Walker, M. C., et. al., *Science* **2013**, 341, 1089-1094.
pAYC138	1045				Module 6 + TE AT^0 from Wong, F. T., et. al., *Biochemistry* **2009**, 49 (1) 95-102.
pAYC136	1150				Module 3 + TE AT^0 from Wong, F. T., et. al., *Biochemistry* **2009**, 49 (1) 95-102.
pRSG54	926				Module 6 + TE from Wu, N., et. al., *J. Am. Chem. Soc.* **2000**, 122 (20) 4847-4852
pRSG34	1123				Module 3 +TE from Gokhale, R. S., et. al., *Science* **1999**, 284 (5413) 482-485
pAYC11	1044				KSAT didomain of module 6 from Chen, A. Y., et. al., *J. Am. Chem. Soc.* **2006**, 128 (9) 3067-3074

B

Name	Sequence
PIKS_KR1 F2	aacgttgggcaggtctggttgatctgcctgcccagccggatg
PIKS_KR1 R2	ttgcacgatgtgcacgacgaacctgttccgctgtcagggtatcaac
KSAT2 F1	attatggtctcttatggagccgatcgcgatcgtc
KSAT2 R1	ataatggtctcaaagcccggtagaaccagccgtc
KS^0AT^0 F1	cgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactataggg
KS^0AT^0 R1	cgccaggtgcagcgccaccagcgacgacgatgccgcggtgtcgacgctgatcgacggccc
KS^0AT^0 F2	gggccgtcgatcagcgtcgacaccgcggcatcgtcgtcgctggtggcgctgcacctggcg
KS^0AT^0 R2	gcgtcctccaacgtgagcgccccggcgacgtgcgcggcggcgatctcgccctgcgcgtgc
Mod3TE ACP0 F1	gggcctggaccgggtcgacgtggtgcagccggtgttgttcgcggtgatggtgtcgctggc
Mod3TE ACP0 R1	ttgcgcagggccatcgcgttgagcgcgtcgagtccgagctcgctgaacgcgcggcgcacg
Mod3TE ACP0 F2	cgtgcgccgcgcgttcagcgagctcggactcgacgcgctcaacgcgatggccctgcgcaa
Mod3TE ACP0 R2	ttgttagcagccggatctcagtggtggtggtggtggtgctcgagtgcggccgcaagcttg

C

Name	Sequence
PIKS-KR1-block1	ttatggtctcttatggcaaccggtgatgattggcgttatcgtattgattggaaacgtctgcctgcagcagaaggtagcgaacgtaccggtctgagcggtcgttggctggcagttacaccggaagatcatagcgcacaggcagcagcagttctgaccgcactggttgatgccggtgcaaaagttgaagttctgacagccggtgcagatgatgatcgtgaagcactggcagcacgtctgacagcgctgaccacaggtgatggtttaccggtgttgttagcctgctggatggtctggttccgcaggttgcatgggttcaggccctgggtgatgcaggtattaaagcaccgctgtggtctgttacccagggtgcagttagcgttggtcgtctggatactccggcagatcctgatcgtgcaatgctgtggggtctgggtcgtgttgttgcactggaacatccggaacgttgggcaggtctggttgatctcgaatgagaccatta
PIKS-KR1-block2	Ttatggtctctcgaagcctgcccagccggatgcagcagcactggcccatctggttaccgcactgagcggtgccaccggtgaagatcagattgcaattcgtaccaccggtctgcatgcacgtcgtctggcacgtgcaccgctgcatggtcgtcgtccgacccgtgattggcagccgcatggcaccgttctgattacaggtggtacaggtgcactgggtagccatgcagcacgttggatggcacatcatggtgcagaacatctgctgctggttagccgtagcggtgaacaggcaccgggtgcaacccagctgaccgcagaactgaccgcaagcggtgcccgtgttaccattgcagcatgtgatgttgcagatccgcatgcaatgcgtaccctgctggacgcaattccggcagaaacaccgctgacagcagttgttcataccgcaggcgctctggatgatggtattgttgataccctgacagcggaattcctgagaccatta
PIKS-KR1-block3	Attatggtctctttcccaggttcgtcgtgcacatcgtgcaaaagcagttggtgcaagcgttctggatgaactgacacgtgatctggatctggatgcatttgttctgtttagcagcgttagcagcacactgggtattccgggtcagggtaattatgcaccgcataatgcatatctggacgccctggcagcccgtcgtcgcgcaacaggtcgtagcgcagtgagcgttgcctggggtccgtgggatggtggtggtatggcagccggtgatggtgttgccgaacgtctgcgtaatcatggtgttccgggtatggacccggaactggcgctggcagcgctggaaagcgcactgggacgtgatgaaaccgcaattaccgttgcagatatcgattgggatcgtttttatctggcatatagcagcggtcgtccgcagccgctggttgaagaactgccggaagttcgtcgtattattgatgcacgtgattaagctttgagaccattat

Table 2. A) Plasmid constructs and B) oligonucleotides used in Chapter 3.

Name	No.	Primers	Restriction sites	Method	Description
pET21c-DszAT-F190V-His$_6$	2330	DszAT F190 F1/R2, DszAT F190V F2/R1	XbaI-HindIII	Gibson	F190V mutation using pFW3 as the template.
pET21c-DszAT-F190L-His$_6$	2077	DszAT F190 F1/R2, DszAT F190L F2/R1	XbaI-HindIII	Gibson	F190L mutation using pFW3 as the template.
pET21c-DszAT-F190S-His$_6$	2021	DszAT F190 F1/R2, DszAT F190S F2/R1	XbaI-HindIII	Gibson	F190S mutation using pFW3 as the template.
pET21c-DszAT-F190I-His$_6$	2012	DszAT F190 F1/R2, DszAT F190I F2/R1	XbaI-HindIII	Gibson	F190I mutation using pFW3 as the template.
pET21c-DszAT-F190P-His$_6$	2014	DszAT F190 F1/R2, DszAT F190P F2/R1	XbaI-HindIII	Gibson	F190P mutation using pFW3 as the template.
pET21c-DszAT-F190T-His$_6$	2074	DszAT F190 F1/R2, DszAT F190T F2/R1	XbaI-HindIII	Gibson	F190T mutation using pFW3 as the template.
pET21c-DszAT-F190A-His$_6$	2114	DszAT F190 F1/R2, DszAT F190A F2/R1	XbaI-HindIII	Gibson	F190A mutation using pFW3 as the template.
pET21c-DszAT-F190Y-His$_6$	2009	DszAT F190 F1/R2, DszAT F190Y F2/R1	XbaI-HindIII	Gibson	F190Y mutation using pFW3 as the template.
pET21c-DszAT-F190H-His$_6$	2020	DszAT F190 F1/R2, DszAT F190H F2/R1	XbaI-HindIII	Gibson	F190H mutation using pFW3 as the template.
pET21c-DszAT-F190Q-His$_6$	2013	DszAT F190 F1/R2, DszAT F190Q F2/R1	XbaI-HindIII	Gibson	F190Q mutation using pFW3 as the template.
pET21c-DszAT-F190N-His$_6$	2016	DszAT F190 F1/R2, DszAT F190N F2/R1	XbaI-HindIII	Gibson	F190N mutation using pFW3 as the template.
pET21c-DszAT-F190K-His$_6$	2018	DszAT F190 F1/R2, DszAT F190K F2/R1	XbaI-HindIII	Gibson	F190K mutation using pFW3 as the template.
pET21c-DszAT-F190D-His$_6$	2017	DszAT F190 F1/R2, DszAT F190D F2/R1	XbaI-HindIII	Gibson	F190D mutation using pFW3 as the template.
pET21c-DszAT-F190E-His$_6$	2019	DszAT F190 F1/R2, DszAT F190E F2/R1	XbaI-HindIII	Gibson	F190E mutation using pFW3 as the template.
pET21c-DszAT-F190C-His$_6$	2011	DszAT F190 F1/R2, DszAT F190C F2/R1	XbaI-HindIII	Gibson	F190C mutation using pFW3 as the template.
pET21c-DszAT-F190W-His$_6$	2073	DszAT F190 F1/R2, DszAT F190W F2/R1	XbaI-HindIII	Gibson	F190W mutation using pFW3 as the template.

pET21c-DszAT-F190M-His$_6$	2010	DszAT F190 F1/R2, DszAT F190M F2/R1	XbaI-HindIII	Gibson	F190M mutation using pFW3 as the template.
pET21c-DszAT-F190G-His$_6$	2015	DszAT F190 F1/R2, DszAT F190G F2/R1	XbaI-HindIII	Gibson	F190G mutation using pFW3 as the template.
pET21c-DszAT-F190R-His$_6$	1985	DszAT F190R F1/R1, DszAT F190R F2/R2	XbaI-HindIII	Gibson	F190R mutation using pFW3 as the template.
pET21c-DszAT-H191A-His$_6$	2232	DszAT F190R F1/R2, DszAT H191A F2/R1	NdeI-EcoRI	Gibson	H191A mutation using pFW3 as the template.
pET21c-DszAT-S86C-His$_6$	2243	DszAT S86 F1/R2, DszAT S86C F2/R1	XbaI-HindIII	Gibson	S86C mutation using pFW3 as the template.
pET21c-DszAT-S86D-His$_6$	2242	DszAT S86 F1/R2, DszAT S86D F2/R1	XbaI-HindIII	Gibson	S86D mutation using pFW3 as the template.
pET21c-DszAT-S86E-His$_6$	2241	DszAT S86 F1/R2, DszAT S86E F2/R1	XbaI-HindIII	Gibson	S86E mutation using pFW3 as the template.
pET21c-DszAT-S86D-H191A-His$_6$	2240	DszAT S86 F1/R2, DszAT S86D F2/R1	XbaI-HindIII	Gibson	S86D mutation using pET21c-DszAT-H191A-His$_6$ as the template.
pET21c-DszAT-S86E-H191A-His$_6$	2239	DszAT S86 F1/R2, DszAT S86E F2/R1	XbaI-HindIII	Gibson	S86E mutation using pET21c-DszAT-H191A-His$_6$ as the template.
pET21c-DszAT-F190G-L87V-His$_6$	2343	DszAT L87 F1/R1, DszAT L87V F2/R2	XbaI-HindIII	Gibson	L87V mutation using pET21c-DszAT-F190G-His$_6$ as the template.
pET21c-DszAT-F190G-L87A-His$_6$	2344	DszAT L87 F1/R1, DszAT L87A F2/R2	XbaI-HindIII	Gibson	L87A mutation using pET21c-DszAT-F190G-His$_6$ as the template.
pET21c-DszAT-F190I-L87V-His$_6$	2345	DszAT L87 F1/R1, DszAT L87V F2/R2	XbaI-HindIII	Gibson	L87V mutation using pET21c-DszAT-F190I-His$_6$ as the template.
pET21c-DszAT-F190I-L87A-His$_6$	2346	DszAT L87 F1/R1, DszAT L87A F2/R2	XbaI-HindIII	Gibson	L87A mutation using pET21c-DszAT-F190I-His$_6$ as the template.
pET21c-DszAT-F190P-L87A-His$_6$	2347	DszAT L87 F1/R1, DszAT L87A F2/R2	XbaI-HindIII	Gibson	L87A mutation using pET21c-DszAT-F190P-His$_6$ as the template.
pET21c-DszAT-F190S-L87V-His$_6$	2348	DszAT L87 F1/R1, DszAT L87V F2/R2	XbaI-HindIII	Gibson	L87V mutation using pET21c-DszAT-F190S-His$_6$ as the template.
pET21c-DszAT-F190S-L87A-His$_6$	2349	DszAT L87 F1/R1, DszAT L87A F2/R2	XbaI-HindIII	Gibson	L87A mutation using pET21c-DszAT-F190S-His$_6$ as the template.
pET21c-DszAT-F190T-L87V-His$_6$	2350	DszAT L87 F1/R1, DszAT L87V F2/R2	XbaI-HindIII	Gibson	L87V mutation using pET21c-DszAT-F190T-His$_6$ as the template.
pET21c-DszAT-F190T-L87A-His$_6$	2351	DszAT L87 F1/R1, DszAT L87A F2/R2	XbaI-HindIII	Gibson	L87A mutation using pET21c-DszAT-F190T-His$_6$ as the template.

Name	Size	Primers	Sites	Method	Notes
pET21c-DszAT-F190V-L87V-His$_6$	2352	DszAT L87 F1/R1, DszAT L87 F2/R2	XbaI-HindIII	Gibson	L87V mutation using pET21c-DszAT-F190V-His$_6$ as the template.
pCDFDuet1-DszAt.RpMatB-T103A	1491				Designed by Thomas Privalsky. Contains a point mutation at T103.
pCDFDuet-DszAT-F190V-rpMatB-T103A	2734	pCDF-DszAT-F190V-MatB F1/R1, pCDF-DszAT-F190V-MatB F1/R2	SacI-AatII	Gibson	F190V mutation inserted into pCDFDuet1-DszAt.RpMatB-T103A.
pBAD18-Cm.500matC	1255				Ben Thuronyi Thesis: Appendix 1
pCK7 (*E. coli*)	2066				ColE1 SCP2* Cb^R Ts^R pactI actII-ORF4 eryAI eryAII eryIII. *Science* **1994**, 265 509-512.
pCK7 (*S. lividans*)	2795				*S. lividans* spores containing pCK7 for production of 6-dEB.
pSET152-ermEp*	592				Genomic integration vector from Bibb, M. J., et. al., *Mol. Microbiol.* **1994**, 14 (3) 533-545.
pSET152-ermEp*-NphT7	2821	pSET152-NpHT7-PhaB F1/R1	NdeI-BamHI	Gibson	Insertion of NphT7 into the genomic integration vector pSET152-ermEp*.
pSET152-ermEp*-NphT7-PhaB	2822	pSET152-NpHT7-PhaB F3/R3	BamHI	Gibson	Insertion of RBS-PhaB into the genomic integration vector pSET152-ermEp*-NphT7.
pET28a-NphT7	1479				Assembled by Brooks Bond-Watts
pSET152-ermEp*-DszAT-MatB	2815	pSET152 DszAT MatB F1/R1, pSET152 DszAT MatB F2/R2	NdeI-BamHI	Gibson	Insertion of DszAT and RBS-MatB into pSET152-ermEp*.
pSET152-ermEp*-F190V-DszAT-MatB	2823	pSET152 F190V DszAT F1/R1, pSET152 DszAT MatB F2/R2	NdeI-HindIII	Gibson	F190V mutation into pSET152-ermEp*-DszAT-MatB.
pTargetF	2637				Used in *in vivo* Cas9-recombineering efforts. From Jiang, Y., et. al., *Appl. Environ. Microbiol.* **2015**, 58 (4) 161-170.
pCas	2636				Used in *in vivo* Cas9-recombineering efforts. From Jiang, Y., et. al., *Appl. Environ. Microbiol.* **2015**, 58 (4) 161-170.

B

Name	Sequence
DszAT F190 F1	gccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccgcgaaa
DszAT F190 R2	gccaactcagcttcctttcgggctttgttagcagccggatctcagtggtggtggtggtgg
DszAT F190L R1	cggtcgcatgaagcgggaatggagagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190I R1	cggtcgcatgaagcgggaatgaatagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190V R1	cggtcgcatgaagcgggaatgaacagcggcgctcacgcgcaggactgtgtacttcttcgc

DszAT F190S R1	cggtcgcatgaagcgggaatgagaagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190P R1	cggtcgcatgaagcgggaatgaggagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190T R1	cggtcgcatgaagcgggaatgagtagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190A R1	cggtcgcatgaagcgggaatgagcagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190Y R1	cggtcgcatgaagcgggaatgataagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190H R1	cggtcgcatgaagcgggaatgatgagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190L R1	cggtcgcatgaagcgggaatgggagagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190Q R1	cggtcgcatgaagcgggaatgttgagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190N R1	cggtcgcatgaagcgggaatgattagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190K R1	cggtcgcatgaagcgggaatgtttagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190D R1	cggtcgcatgaagcgggaatgatcagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190E R1	cggtcgcatgaagcgggaatgttcagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190C R1	cggtcgcatgaagcgggaatgacaagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190W R1	cggtcgcatgaagcgggaatgccaagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190M R1	cggtcgcatgaagcgggaatgtagagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190G R1	cggtcgcatgaagcgggaatgaccagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190L F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctctccattcccgcttcatgcgaccg
DszAT F190I F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctattcattcccgcttcatgcgaccg
DszAT F190V F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctgttcattcccgcttcatgcgaccg
DszAT F190S F2	gcgaagaagtacacagtcctgcgcgtgagcgccgcttctcattcccgcttcatgcgaccg
DszAT F190P F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctcctcattcccgcttcatgcgaccg
DszAT F190T F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctactcattcccgcttcatgcgaccg
DszAT F190A F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctgctcattcccgcttcatgcgaccg
DszAT F190Y F2	gcgaagaagtacacagtcctgcgcgtgagcgccgcttatcattcccgcttcatgcgaccg
DszAT F190H F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctcatcattcccgcttcatgcgaccg
DszAT F190Q F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctcaacattcccgcttcatgcgaccg
DszAT F190N F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctaatcattcccgcttcatgcgaccg
DszAT F190K F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctaaacattcccgcttcatgcgaccg
DszAT F190D F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctgatcattcccgcttcatgcgaccg
DszAT F190E F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctgaacattcccgcttcatgcgaccg
DszAT F190C F2	gcgaagaagtacacagtcctgcgcgtgagcgccgcttgtcattcccgcttcatgcgaccg
DszAT F190W F2	gcgaagaagtacacagtcctgcgcgtgagcgccgcttggcattcccgcttcatgcgaccg
DszAT F190M F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctatgcattcccgcttcatgcgaccg
DszAT F190G F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctggtcattcccgcttcatgcgaccg
DszAT F190R F1	atgcgtccggcgtagaggatcgagatctcgatcccgcgaaattaatacgactcactagggggaattgtgagcggataacaattcccctc
DszAT F190R R1	cggtcgcatgaagcgggaatggcgagcggcgctcacgcgcaggactgtgtacttcttcgc
DszAT F190R F2	gcgaagaagtacacagtcctgcgcgtgagcgccgctcgccattcccgcttcatgcgaccg
DszAT F190R R2	gccaactcagcttcctttcgggctttgttagcagccggatctcagtggtggtggtggtgg
DszAT S86 F1	ccacggggcctgccaccatacccacgccgaaacaagcgctcatgagcccgaagtggcgag
DszAT S86 R2	gccaactcagcttcctttcgggctttgttagcagccggatctcagtggtggtggtggtgg
DszAT S86C R1	aacagggcgctgaactcgcccagacagtggccggccaggaaatcgg
DszAT S86A R1	aacagggcgctgaactcgcccagacagtggccggccaggaaatcgg
DszAT S86D R1	aacagggcgctgaactcgcccagatcgtggccggccaggaaatcgg
DszAT S86E R1	aacagggcgctgaactcgcccagttcgtggccggccaggaaatcgg
DszAT S86C F2	ccgatttcctggccggccactgtctgggcgagttcagcgccctgtt
DszAT S86A F2	ccgatttcctggccggccacgctctgggcgagttcagcgccctgtt
DszAT S86D F2	ccgatttcctggccggccacgatctgggcgagttcagcgccctgtt
DszAT S86E F2	ccgatttcctggccggccacgaactgggcgagttcagcgccctgtt

119

DszAT L87 F1	ccacggggcctgccaccatacccacgccgaaacaagcgctcatgagcccgaagtggcgag
DszAT L87 R2	gccaactcagcttcctttcgggctttgttagcagccggatctcagtggtggtggtggtgg
DszAT L87V R1	aacagggcgctgaactcgccaaccgagtggccggccaggaaatcgg
DszAT L87A R1	aacagggcgctgaactcgccagccgagtggccggccaggaaatcgg
DszAT L87V F2	ccgatttcctggccggccactcggttggcgagttcagcgccctgtt
DszAT L87A F2	ccgatttcctggccggccactcggctggcgagttcagcgccctgtt
DszAT H191A R1	gaccatcgccggtcgcatgaagcgggaagcgaaagcggcgctcacgcgcaggactgtgta
DszAT H191A F2	tacacagtcctgcgcgtgagcgccgctttcgcttcccgcttcatgcgaccggcgatggtc
pCDF-DszAT-MatB F1	ccatgaaagcatacatgtttcccgggcaagggtctcaggcgaaggggatgggacgggcgc
pCDF-DszAT-MatB R2	gccgttctggtcgaggagctcgcgaacccgctcctcgtccagaccgatga
pCDF-DszAT-MatB F2	gaccatcgccggtcgcatgaagcgggaatgaacagcggcgctcacgcgcaggactgtgta
pCDF-DszAT-MatB R1	tacacagtcctgcgcgtgagcgccgctgttcattcccgcttcatgcgaccggcgatggtc
pSET152-NphT7-PhaB F1	ccggttggtaggatcgtctagaacaggaggccccatatgaccgacgttcgttttcgtatc
pSET152-NphT7-PhaB R1	agtgccaagcttgggctgcaggtcgactctagaggatccttaccactcgatcagcgcgaa
pSET152-NphT7-PhaB F3	ttcgcgctgatcgagtggtaaggatccaggaggccccatatgacccagcgcatcgcttac
pSET152-NphT7-PhaB R3	cagtgccaagcttgggctgcaggtcgactctagaggatccttagcccatgtgcaggccac
pSET152 DszAT MatB F1	ggtaggatcgtctagaacaggaggccccatatgaaaagcatacatgtttcccgggcaa
pSET152 DszAT MatB R1	acaaggtgacgttgctggccacgaacagaccgtgggtgtgatagatcggcagcgcat
pSET152 DszAT MatB F2	atgcgctgccgatctatcacacccacggtctgttcgtggccagcaacgtcaccttgt
pSET152 DszAT MatB R2	cttgggctgcaggtcgactctagaggatccttacttatagatgtctttataggttca
pSET152 F190V DszAT F1	cgtgccggttggtaggatcgtctagaacaggaggccccatatgaaaagcatacatgtt
pSET152 F190V DszAT R1	caggcgggcaaacagattcgcgttcatatgggggcctcctaagcttcgacgacgaggggct
M3 rep F1	cccgcgggcggtcgtgggggcacgcgcaaggcgagatcgccgccgcgcacgtcgccggggcg
M3 rep R1	cgccccggccgacgtgcgcggcggcgatctcgccttgcgcgtgccccacgaccgccgcggg
2p pCK7 M6 AT0 F1	gtacagccggtgttgttctccgtgatggtgtcgctggcgcggctgtggggcgcttgcgga
2p pCK7 M6 AT0 R1	ttaacctataaaaataggcgtatcacgaggccctttcgtcttcaagaatttatgcatctatgaattccctccgcccagccaggcgtcgat
sgRNA M3 AT^0	ggcggucgugggggcacucgc + Synthego EZ Scaffold
sgRNA M6 AT^0	ggccgucaucggccauucgc + Synthego EZ Scaffold

Table 3. Plasmid constructs used Chapter 4.

Name	No.	Primers	Restriction sites	Method	Description
pET23c.His$_6$-TEV-NphT7	2041				Ben Thuronyi Thesis: Appendix 1
pET16b-His.Phaj	855				Designed by Jeffrey Hanson.
pCWori-ccr.His	557				Brooks Bond-Watts Thesis: Appendix 2
pET16B-His-PhaB	234				Brooks Bond-Watts Thesis: Appendix 2

Table 4. A) Plasmid constructs, and B) oligonucleotides not described in materials and methods.

A

Name	No.	Primers	Restriction sites	Method	Description
pET28a-HisN-RFP-gg	1332	pET28a-RFP-gg-HISN F1/R1	NcoI-EcoRI	Gibson	RFP flanked by BsaI sites added to pET28a.
pET28a-NHis-gg-AT3	1344	M3 standalone F1/R1	BsaI-BsaI	Golden Gate	Inserts module 3 AT with into pET28a-HisN-RFP-gg.
pFW98	1565				DEBS2 protein (Module 3s and 4) from Lowry, B., et. al., *J. Am. Chem. Soc.* **2013**, 135 (45) 16809-16812.
pET28a-NHis-gg-AT30NP	2825	M3 standalone F1/R1	BsaI-BsaI	Golden Gate	Inserts module 3 AT^0 with into pET28a-HisN-RFP-gg.
pET28a-NHis-gg-KS	2824	M3KS standalone F2/R2	BsaI-BsaI	Golden Gate	Inserts standalone module 3 KS into pET28a-HisN-RFP-gg.
pSV272.1M3AT0NOTE	1994	M3TEL F1/R1, M3TEL F2/R2	NheI-AscI	Gibson	Inserted AT^0 mutation into pSV272-DEBMS (No. 1342).
pET21c-DEBS1$_{TE}$	2732	pCWORI DEBS1TE F1/R1	SexAI-HindIII/ NdeI-HindIII	Gibson/ligation	TE inserted in pCWOri-DEBS1 (No. 1327) using gibson. TE amplified from pRSG64. DEBS1$_{TE}$ was then inserted into pET21c via restriction ligation.
pSET152-DszAT-F190V ermEp*	2735	pSET152DszAT F1/R1	NdeI-BamHI	Gibson	F190V DszAT variant inserted into pSET152-ermEp*.
pUWL-DszAT	2739	pUWL DS F2/R2 ST	NdeI-HindIII	Gibson	DszAT inserted into pUWL201PW (No. 1421).
pUWL-DszAT-F190V	2738	pUWL-DszAT-F190V F1/R1, pUWL-DszAT-F190V F2/R2	NdeI-HindIII	Gibson	DszAT F190V inserted into pUWL201PW
pUWL-DszAT-MatB	2736	pUWL DS MB F1/R1	HindIII	Gibson	rpMatB inserted into pUWL-DszAT.
pUWL-DszAT-F190V-MatB	2737	pUWL DS MB F1/R1	HIndIII	Gibson	rpMatB inserted into pUWL-DszAT-F190V.
pCDFDuet-DszAT-rpMatB	2816	pCDF-DszAT-MatB F1/R2, pCDF-DszAT-A103T-MatB F2/R1	XmaI-SacI	Gibson	Corrected the T103A mutation in pCDFDuet1-DszAt.RpMatB-T103A.
pCDFDuet-DszAT-F190V-rpMatB	2817	pCDF-DszAT-MatB F1/R2, pCDF-DszAT-A103T-MatB F2/R1	XmaI-SacI	Gibson	Corrected the T103A mutation in pCDFDuet-DszAT-F190V-rpMatB-T103A.
pCDFDuet-DszAT-S86A-rpMatB	2818	pCDF-DszAT-S86A-MatB F2/R1, pCDF-DszAT-MatB F1/R2	XmaI-SacI	Gibson	Introduction of S86A mutation into pCDFDuet-DszAT-rpMatB.

B

Name	Sequence
pET28a-RFP-gg-HISN F1	cctggtgccgcgcggcagccatatgtgagacctccctatcagtgatagagattg
pET28a-RFP-gg-HISN R1	aagcttgtcgacggagctcgaattcttaagcagagacctataaacgcagaaaggccc
M3AT Standalone F1	attatggtctcttatggagcgcgagcaccgggagacc
M3AT Standalone R1	ataatggtctcaaagcacgcggtaggccagctcgtcgg
M3KS Standalone F2	attatggtctcttatggtgactgacagcgagaagg
M3KS Standalone R2	ataatggtctcaaagcaggagcttcctcgacgatca
M3TEL F1	gagcccacgccgcacgtcgactggtcgtccggcggcgtggcgctgctggcgggcaaccag
M3TEL R1	tcgtacggcgtggagcccgcggcggtcgtggggcacgcgcagggcgagatcgccgccgcg
M3TEL F2	cgcggcggcgatctcgccctgcgcgtgccccacgaccgccgcgggctccacgccgtacga
M3TEL R2	ttcgggctttgttagcagccggatctcagtggtggtggtggtggtgctcgacatcctcgg
pCWORI DEBS1TE F1	ctcaccgcggtgctggccggctcggaggaccaggtcgcggtgcgcgccgacgccgtgcgt
pCWORI DEBS1TE R1	catgtttgacagcttatcatcgataagcttcctccgcccagccaggcgtcgatgtgccgc
pSET152DszAT F1	aatcgtgccggttggtaggatcgtctagaacaggaggccccatatgaaagcatacatgtt
pSET152DszAT R1	cagtgccaagcttgggctgcaggtcgactctagaggatccctagttattgctcagcggtg
pUWL DS F2 ST	ttgtttaacattaaagaggagaaattacatatgaaagcatacatgtttcccgggcaaggg
pUWL DS R2 ST	tagtggatcccccgggctgcaggaattcgatatcaagcttcgacgacgaggggctgggcg
pUWL-DszAT-F190V F1	cgtggcaccgcgatgctgttgtggggcacaatcgtgccggttggtaggatcgatccactag
pUWL-DszAT-F190V R1	accatcgccggtcgcatgaagcgggaatgaacagcggcgctcacgcgcaggactgtgta
pUWL-DszAT-F190V F2	tacacagtcctgcgcgtgagcgccgctgttcattcccgcttcatgcgaccggcgatggt
pUWL-DszAT-F190V R2	cgcaaaccgcctctccccgcgcgttggccgattcattaatgcagagcttctagaac
pUWL DS MB F1	cgcccagcccctcgtcgtcgaagcttaggaggccccatatgaacgcgaatctgtttgccc
pUWL DS MB R1	aatgcagagcttctagaactagtggatccttacttatagatgtctttataggtttcacgc
pCDF-DszAT-MatB F1	ccatgaaagcatacatgtttcccgggcaagggtctcaggcgaaggggatgggacgggcgc
pCDF-DszAT-MatB R2	gccgttctggtcgaggagctcgcgaacccgctcctcgtccagaccgatga
pCDF-DszAT-A103T-MatB F2	ttcgccgcgggggtgttcgacttcgagaccggcctcgcgctggtgaagaa
pCDF-DszAT-A103T-MatB R1	ttcttcaccagcgcgaggccggtctcgaagtcgaacacccccgcggcgaa
pCDF-DszAT-S86A-MatB F2	aacagggcgctgaactcgcccagagcgtggccggccaggaaatcggggggagc
pCDF-DszAT-S86A-MatB R1	gctcccccgatttcctggccggccacgctctgggcgagttcagcgccctgtt