Title

Engineering production of fluorinated polyketide natural product analogs

Author

Sirirungruang, Sasilada

Publication Date

2020

Peer reviewed|Thesis/dissertation

Engineering production of fluorinated polyketide natural product analogs

by

Sasilada Sirirungruang

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy in
Molecular and Cell Biology
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Michelle C. Y. Chang, Chair
Professor David F. Savage
Professor Evan W. Miller
Professor Wenjun Zhang

Summer 2020

Engineering production of fluorinated polyketide natural product analogs

by Sasilada Sirirungruang

Abstract

Engineering production of fluorinated polyketide natural product analogs

by

Sasilada Sirirungruang

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Michelle C. Y. Chang, Chair

Fluorination is an important tool for the discovery and production of synthetic pharmaceuticals, as it allows fine-tuning their pharmacokinetic properties and metabolic stability. However, unlike chemists, Nature rarely utilizes fluorine to enhance its molecules, and the few known naturally occurring organofluorines generally lack structural complexity. Despite advances in chemical fluorination methods, the regiospecific fluorination of complex natural products remains a challenge. Thus, we have taken a metabolic engineering approach to develop a novel platform for fluorine incorporation into polyketide natural products. Polyketide natural products are synthesized from one starter unit and multiple acyl-CoA extender units. We designed a modular system in which an inactivating mutation was introduced to the acyltransferase (AT) domain of interest to reduce the native activity of the module, while an alternative fluorine-selective AT enzyme was provided in *trans* to incorporate fluorine as a fluoromalonyl-CoA extender unit. A malonate transporter and a malonyl-CoA synthetase were introduced to support the intracellular accumulation of fluoromalonate and its enzymatic activation into fluoromalonyl-CoA extender unit, respectively. By engineering a single module construct of 6-deoxyerythronolide B (6dEB) synthase (DEBS), fluorinated triketide lactone was produced as the major polyketide product at mg/L levels in an *Escherichia coli* host. Extending these engineering concepts to a complete DEBS pathway allowed us to produce novel fluorine-containing 6dEB analogs in *E. coli* production host in a regiospecific manner. This platform can allow the precise incorporation of fluorine into many positions in the backbones of polyketides. These results open the door to producing a new class of drug-like molecules with potentially diverse and improved properties.

Table of Contents

List of Figures, Schemes, and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Abbreviations

6dEB	6 deoxyerythronolide B
ACP	acyl carrier protein
AT	acyltransferase
DEBS	Deoxyerythronolide B synthase
DH	dehydratase
DszAT	disorazole synthase acyltransferase
ER	enoylreductase
Fmal	Fluoromalonate or fluoromalonyl-CoA
F-TKL	2-fluoro-3-keto-4-methyl-5-hydroxy-n-heptanoic acid δ-lactone
H-TKL	3-keto-4-methyl-5-hydroxy-n-heptanoic acid δ-lactone
KR	ketoreductase
KS	ketosynthase
Mal	Malonate or malonyl-CoA
Memal	Methylmalonate or methylmalonyl-CoA
Mod	Module
PKS	polyketide synthase
SNAC	N-acetyl cysteamine (thioester)
TE	Thioesterase
TKL	2,4-dimethyl-3-keto-5-hydroxy-n-heptanoic acid δ-lactone
TTKP	4-hydroxy-6-(3-hydroxypentan-2-yl)-2H-pyran-2-one

Acknowledgments

First and foremost, I need to thank Professor Michelle Chang, who has made everything presented here possible. Besides the numerous scientific advice given every Monday, I would like to especially thank Michelle for her utmost patience and undeterred optimism. Over the past five years, most of which I had spent without a major positive result, Michelle had demonstrated to me the power of setting a daring goal and challenging one's self to it. Although the route was never straight or clear, it taught me so many creative problem-solving and scientific survival skills. Michelle, thank you very much for encouraging me along this journey and pushing me to achieve what I otherwise might have given up on long before it was fruitful.

Many other people also played a role in making this work possible- my thesis committee, my Chang lab colleagues, other colleages in MCB, in the College of Chemistry, and at UC Berekley, friends, MCB admin staff, loading dock staff, storeroom staff, and others. I would like to acknowledge them all for their help and support. Among these, a few stood out as whom I owe special thanks to.

First, I would like to acknowledge a vital collaborator of this work, Tom Privalski. It's funny how Tom keeps getting roped into the fluorine project, since he was an undergraduate in the Chang lab, until now that he is finishing his PhD study in the Khosla lab. I want to thank Tom for being willing to roll along with all my experimental ambitions, providing valuable opinions, and making 6dEB great again, despite me neglecting to add water to our master mix since the first experiment we carried out together. The COVID-19 pandemic forced me to finish our last planned experiment alone, so I hope we will have more opportunities to collaborate in the future.

Another person I would like to give a special thank to is Swetha Ramesh, a very dedicated and talented undergraduate who I had the priviledge to mentor. Swetha's hard work is instrumental to the construction of many production systems for fluorine incorporation we have now tested. Over the course of her time in the Chang lab, Swetha had developed herself immensely as a thoughtful scientist. That achievement became even more impressive when the admittedly unbearable amount of tedious work I assigned to her is considered, #reasonsforhiringundergrads. Many thanks to Swetha for sticking with me through the most labor-intensive year of my PhD and keeping me on my toes with your impressive work ethics.

Moreover, I would like to acknowledge Hongjun Dong, who seemingly has endless tricks up his sleeve when it comes to DNA manipulation. Hongjun was the only one who somehow was interested in my DEBS sequence cloning struggles and offered help. I had since repeated Hongjun's solution many times, and it is yet to fail. For that, Hongjun became my most favorite postdoc ever, even though he's no longer a postdoc- Hi, Professor Dong!

Besides scientific growth, PhD was a period of personal growth. For that, I would like to acknowledge Vivian Yu, who I find a tremendous support and a true friend. The past six years had forced me to let go of idealism and revise many believes and values I held dear. Turns out six years is long enough to find a flaw in just about everything I spent the period staring at–meritocracy, justice, equality, ethics, freedom. Being the resilient yet humble fighter that she is, Vivian had been for me a voice of reason and a voice of reality.

I would like to give special thanks to Vincent Blay for providing the most valuable emotional

and mental support. Thank you for being my rock each and everyday. Our paths crossed at the time when graduate school was taking its heaviest toll on me. Despite endless darkness that I felt was looming over my world, he chose to believe in me and made me believe in myself when no one else did. At the time when every day was an uphill battle, when everyday I asked myself why, when I saw no path forward, Vincent created a path for me. Vincent, many many thanks to you for being my biggest cheerleader of all time. You mean the world to me.

And lastly and most importantly, I would like to also acknowledge my family, who despite residing on the other side of the Earth, remains close at heart. I know my intellectual pursuit has had no small consequences for them too, so I would like to thank them for their understanding, patience, and support for every decision I make. Dad, Sister, and Nanny, I hope this is worth it and I hope I make you all proud. And for Mom, thank you for always answering my hopes and wishes; I know you are there, taking care of me.

Chapter 1: *Introduction*

1.1 Natural products and their analogs in modern medicine

Natural products are small molecules produced by living organisms. They are often termed secondary metabolites because natural products are not required for the survival of the host organisms under laboratory conditions, but undoubtedly provide ecological advantage to the producers in their native environment. Natural products are classified into different classes based on their biosynthetic pathway. Large classes include polyketides, non-ribosomal peptides, isoprenoids, phenylpropanoids, and alkaloids. Macromolecules such as DNA, RNA, and proteins are usually excluded from the definition of natural products.

The hallmarks of natural products are the complex structures that provide the basis for their bioactivity. Natural products use relatively small set of chemical functional groups; however, they are able to exert impressive specificity and potency through their intricate three-dimensional structures that have been selected for by evolution [1]. An analysis showed that natural products contain many more sp^3-hybridized bridgehead atoms stereocenters than synthetic molecules [2]. Natural products also contain more innovative aliphatic ring systems that provide suitable architecture for molecular positioning [1]. These features, however, often prevent convenient access to the molecules and their derivatives through total chemical synthesis.

Natural products as pharmaceuticals. Natural products have gathered a lot of attention from the scientific community due to their use as human therapeutics (*Figure 1.1*). Historically, natural products from plants were the major source of medicinal preparations. More recently, natural products and their derivatives continue to provide drug leads and become approved for clinical uses [3]. It is estimated that 34% of small molecule therapeutics approved by the US Food and Drug administration (FDA) form 1981 and 2019 were natural products or their direct derivatives [4]. Natural products treat wide spread of human conditions ranging from infectious diseases [5, 6], cancers [7, 8], cardiovascular diseases [9], immune conditions [10, 11], and neurodegenerative diseases [12, 13]. Some specific examples are acetylsalicylic acid (Aspirin), morphine, penicillin, vancomycin, paclitaxel (Taxol), atorvastatin, and tacrolimus (FK506).

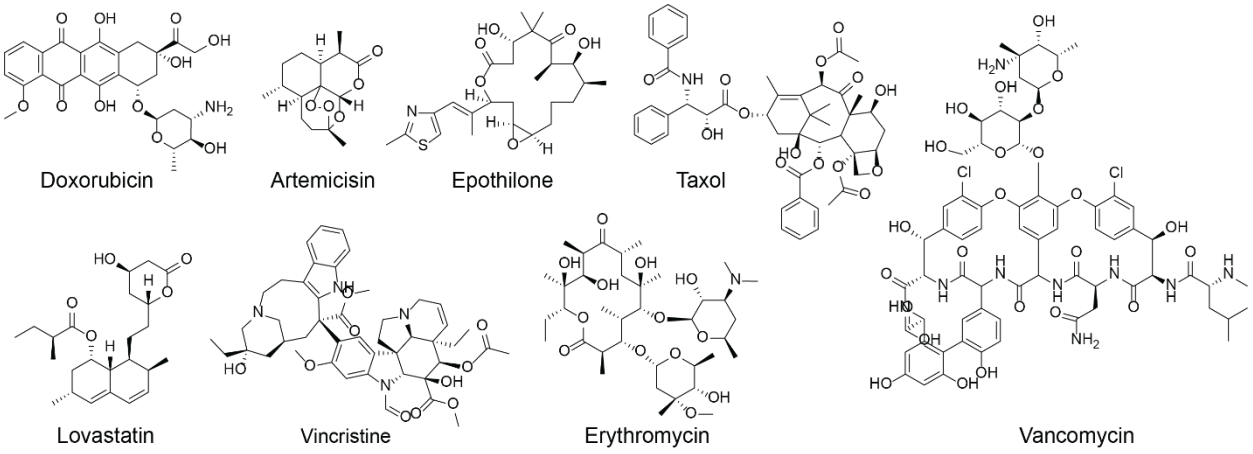


Figure 1.1 *Examples of natural product drugs*

One of the most well-known natural product success stories is that of statins, or hydroxymethylglutaryl coenzyme A (HMG-CoA) reductase inhibitors, in controlling human blood cholesterol levels and reducing morbidity and mortality rates of associated diseases. Applied daily at maximally recommended dose, lovastatin produces a mean reduction in LDL cholesterol of 40 %, a far greater reduction than could be obtained with any of the treatments available at the time of its introduction, with very few adverse effects [14]. Its semisynthetic analog simvastatin followed with even greater success - it was the eighth most commonly prescribed medication in the US in 2017, comprising more than 56 million prescriptions [15]. Statins are just one of many contributions natural products have made to modern medicine.

Yet, the most prominent example of successful implementation of natural product therapeutics is in the battle against infectious diseases. In the 19th century, infections were among the leading causes of mortality; however, they are mostly reduced to simple annoyances thanks to the introduction of antibiotic agents. Since the introduction of penicillin, the first widely used antibiotics isolated from *penicillium notatum*, the world observed a continuous decline in the incidence of acute rheumatic fever, syphilis, typhus, plaque, and tuberculosis [16, 17]. Currently, 6 out of 9 classes of antibiotics and 70 out of the 90 antibiotics marketed during 1982–2002 originated from natural products [18, 19].

For a time, new bioactive agents could be found from low-throughput culturing and screening methods with relative ease. Since then, the 'golden era' of antibiotics discovery had ended. The reservoir of easily accessible compounds from nature seem to have been exhausted, and natural products experienced diminished interests by major pharmaceutical companies. This decrease in interests had also been attributed to the United Nations Convention on Biological Diversity that promotes fair share of benefits from the use of natural products to the source country, and to the Nagoya Protocol that gives detailed suggestions on access and benefit sharing of use of traditional knowledge with indigenous and local communities [20]. In addition, concerns relating to repeated isolation of known compounds and to the manufacturing costs of such complex entities had shifted the focus of the pharmaceutical industry to synthetic small molecule libraries [21].

Synthetic combinatorial chemistry libraries have the advantage of being more compatible with high-throughput screening (HTS) techniques, whereas, natural products in forms of microbial extracts frequently interfere HTS implementations [22, 23]. Chemical libraries can also be selected based on target families and help in generating mechanistic hypothesis of disease phenotype modulation [24]. However, screening of synthetic compounds have sometimes yielded disappointing results, partially contributing to a decline in approvals of small molecule therapeutics [20]. The current hypothesis is that the structural complexity of natural products compared to synthetic compounds allows for a higher hit rate for selective binding to biological targets and that this structural space is important to preserve [24].

Renewed interest in natural product discovery. Recently, emerging challenges such as antibiotic-resistant infectious agents more than ever demand new development of therapeutics with novel modes of actions [18, 25]. It is estimated that 40-60 % of hospital-acquired *Staphylococcus aureus* infections in the US are methicillin-resistant [18]. Each year, more than 2.8 million antibiotic-resistant infections are confirmed in the US, resulting in more than 35,000 mortality [26].These demands, along with newly available data and technologies, had sparked a renewed interest in natural products as drug candidates. This renewed interest benefits from advances in four major areas that might contribute to novel discoveries.

First is the scope of genomic sequences in which new biosynthetic genes might be encoded had greatly expanded. Recently gathered genomic data revealed that many organisms maintain a much larger number of biosynthetic gene clusters than previously anticipated, suggesting many compounds are awaiting discovery [27, 28]. The search for new compounds had also recently expanded to marine organisms and other non-culturable living systems [27]. Metagenomic libraries are currently being explored for novel clusters.

The larger genomic scope explored is supported and enabled by advances in DNA techniques that allow cryptic clusters to be expressed. It is estimated that less than 10% of biosynthetic gene clusters are expressed in sufficient quantity to be observed by routine screening analysis [28]. However, improved molecular and synthetic biology techniques such as DNA synthesis, heterologous expression, regulators modification, modulations of transcription and translation processes allow rapid exploration of cryptic clusters or clusters from non-culturable organisms [28].

The development of improved culture, isolation, and characterization techniques also allow natural extracts to be more compatible with HTS techniques [21, 23]. Previous attempts of screening extracts suffer from the complexity of the extracts, the high rate of rediscovery of known compounds, and the low concentration of many components in the extracts [20]. Novel compounds might exist in concentrations too low to have measurable effects, or their effects might be confounded by interfering compounds. Prefractionation of extracts yield fractions of reduced complexity that can be prepared for screening, allowing for faster and smaller scale screening [20]. Fractionated extracts had been shown to increase the discovery rate and found compounds that did not show effects in complex extract [29].

Lastly, much has been learned regarding the significance of protein-protein interactions in the functions of megasynthases such as polyketide synthases and non-ribosomal peptide synthases. New structural knowledge is informing more effective rational design of new combinatorial clusters [21].

1.2 Organofluorines in medicinal and biological chemistry

Fluorine in drug discovery. Fluorine has played a pivotal role in drug discovery efforts. It is estimated that the number of approved drugs that contained fluorine had increased from 2% in 1970 to 20-30% currently [30, 31]. This dramatic increase relied on the continual development of new fluorination methods since the introduction of the first synthetic fluorinated drug, the antineoplastic agent 5-fluorouracil, in 1957 [32, 33]. Since then, fluorine substitution is commonly used in medicinal chemistry and over 150 fluorinated drugs have come to market [32, 33]. Some of the current best-selling drugs that contain fluorine are Prozac, Lipitor, Cipro, and Diflucan (*Figure 1.2*).

A substitution by a single fluorine atom or a trifluoromethyl group in key positions of a compound can result in profound beneficial effects. Due to its high electronegativity (3.98 on the Pauling scale) and small size (van der Waals radius of 1.47 Å), fluorine offers a unique opportunity to alter physical properties such as solubility, lipophilicity, conformation, and pK_a of a molecule. Fluorine's properties allow it to serve as isoteres for hydrogen and double-bonded oxygen. For example, fluorovinyl group (C=CHF) has been effectively used as isostere for peptide bond [32].

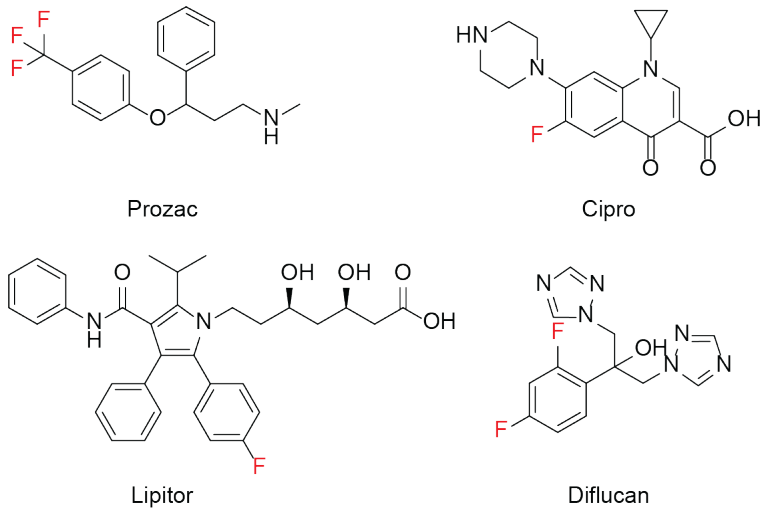


Figure 1.2 *Examples of fluorine-containing small molecule drugs*

In addition, C-F bond is relatively inert and is almost entirely foreign to biology. Thus, fluorine has been shown to improve bioavailability of drugs by improving pharmacokinetics properties and increasing metabolic stability.

Fluorine's high electronegativity alters the electron distribution of a molecule when fluorine is introduced. This effect can impact pK_a, dipole moment, chemical reactivity of neighboring functional groups, and binding affinity to target proteins [34]. High electron withdrawing capacity also allows fluorine to induce a conformational change in a molecule. A highly polarized C-F σ-bond has a low lying σ* orbital that is subjected to hyperconjugative electron donation [35]. Thus, in aliphatic systems, this effect results in a preference for vicinal functional group to align *gauche* to fluorine [35].

In terms of altering lipophilicity, strong electron withdrawing capability of fluorine decreases lipophilicity of alkyl groups when they undergo monofluorination or trifluoromethylation [33, 36]. In contrast, fluorination on or adjacent to atoms with π-bonds, and perfluorination increases lipophilicity [36]. Thus, fluorination of a target molecule at strategic sites could affect its membrane permeability and metabolic distribution.

Strategic fluorine substitution introduced at the site of metabolic attack could increase bioavailability due to the fact that C-F bond is more resistant to cleavage than C-H bond [34]. For example, replacing a susceptible methyl group with a fluorine atom of a CF_3 group has been shown to be effective against oxidative metabolism [37]. In aromatic systems, the presence of the fluorine can block oxidation at a specific position and decrease the rate of reaction of the π-system of the benzene ring with activated cytochrome P450(Fe^{IV}=O)$^{3+}$ [38].

Chemical fluorination methods. C-F bond could be created in many ways; however traditional methods involve highly hazardous reagents that require specialized facility to handle in large scale. Thus, development of safer, more convenient reagents had greatly expanded synthetic access to organofluorines.

Nucleophilic fluorination allows conversion of alcohols to fluorides. Traditional nucleophilic fluorination methods rely on hydrofluoric acid (HF). However, HF is among the most hazardous reagents used. Besides HF and its mixtures with hydrogen bond acceptors, alkali and sulfur fluorides such as LiF, KF, CsF, and SF_4 had been utilized for nucleophilic fluorination [30]. However, these reagents suffer from low solubility in organic solvents and side reactions. Diethylaminosulfur trifluoride (DAST) was developed to overcome these problems and has proven popular for converting a wide variety of alcohols to fluorides [39]. However, DAST generated safety concerns that prompted development of subsequent generations of deoxyfluorination agents such as Deoxo-Fluor[40], XtalFluor-E and XtalFluor-M [41, 42], Fluolead [43], PyFluor [44], and PhenoFluor [45].

On the other hand, electrophilic fluorination is traditionally achieved with elemental fluorine (F_2) [30]. However, its indiscriminatory reactivity and associated safety concerns let to the development of N-F fluorinating agents for fluorination of aromatics, alkenes, carbanions, and enol ethers [30]. Examples of N-F fluorinating agents include N-fluoropyridinium salts, N-fluorosulfonamides, and N-alkyl-N-fluorosulfonamides; most popular among them are SelectFluor [46], and N-fluorobenzenesulfonimide (NFSI) [47, 48].

The development of new reagents for safe, manageable fluorination continues to expand. However, currently challenges in regiospecific fluorination continue to restrict the chemical space accessible for drug discovery purposes. Most reagents are still too expensive for large-scale production [32]. 90% of fluorine-containing small molecule drugs introduced between 2001 and 2011 were synthesized from fluorine-containing precursors obtained from facilities with specialized equipment for economical fluorination [30].

Natural occurrence of organofluorines and biological fluorination. Organofluorines are the least abundant of naturally occurring organohalides [49]. Despite being the most abundant halogen in the earth crust, bioavailability of fluorine is very low [50, 51]. Fluoride ion exists at a very low concentration of 1.3 ppm in the oceans, compared to 20,000 ppm of chloride and 70 ppm of bromide, due to low solubility most inorganic fluoride salts [49, 51]. It also contains a tightly solvent shell with a high hydration energy, making it a poor nucleophile that requires a significant desolvation strategy in aqueous environment.

All known fluorometabolites are found in fluoroacetate-producing plant species, as part of nucleocidin biosynthetic pathway, as part of the fluorinase pathway, or in ω-fluorofatty acid biosynthesis (*Figure 1.3*). Many plants in the families Fabaceae, Rubiaceae, Bignoniaceae, Malpighiaceae and Dichapetalaceae had been described to accumulate fluoroacetate as defense molecule, due to its toxicity mediated by conversion to fluorocitrate, which inhibits aconitase enzyme [52]. Fluoroacetate had also been shown to be incorporated into ω-fluorofatty acids by the fatty acid synthase machinery [50]. Nucleocidin or 4'-fluoro-5'-O-sulfamoyladenosine is produced by *Streptomyces calvus* (ATCC13382) and had been demonstrated to have antibiotic activity against trypanosomes [53]. However, the synthetic pathways leading to these early identified organofluorines had not been fully elucidated [52, 54]. The most well understood organofluorine pathway is the fluorinase pathway (*Figure 1.4*) [55]. The fluorinase pathway, which was originally identified in *Streptomyces cattleya* and had since been found in a handful other bacteria, consists of the fluorinase enzyme, which produces 5'-fluoro-5'-deoxyadenosine (FDA) from S-adenosyl-L-methionine and inorganic fluoride [56, 57]. Subsequent transformations of FDA produces 4-fluorotheronine and fluoroacetate as the final products of the pathway [55].

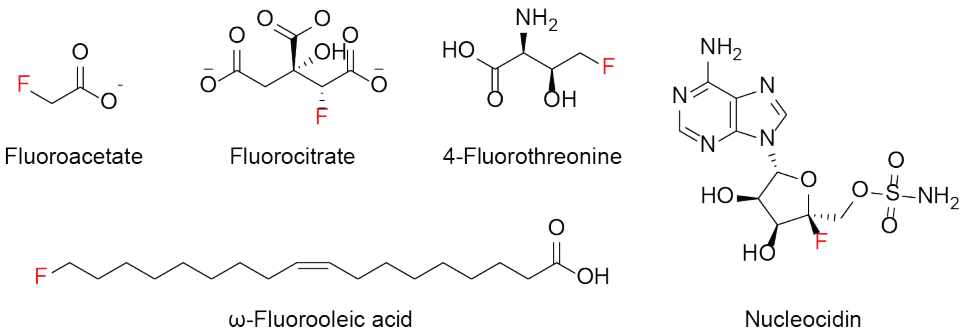


Figure 1.3 *Naturally occurring organofluorines*

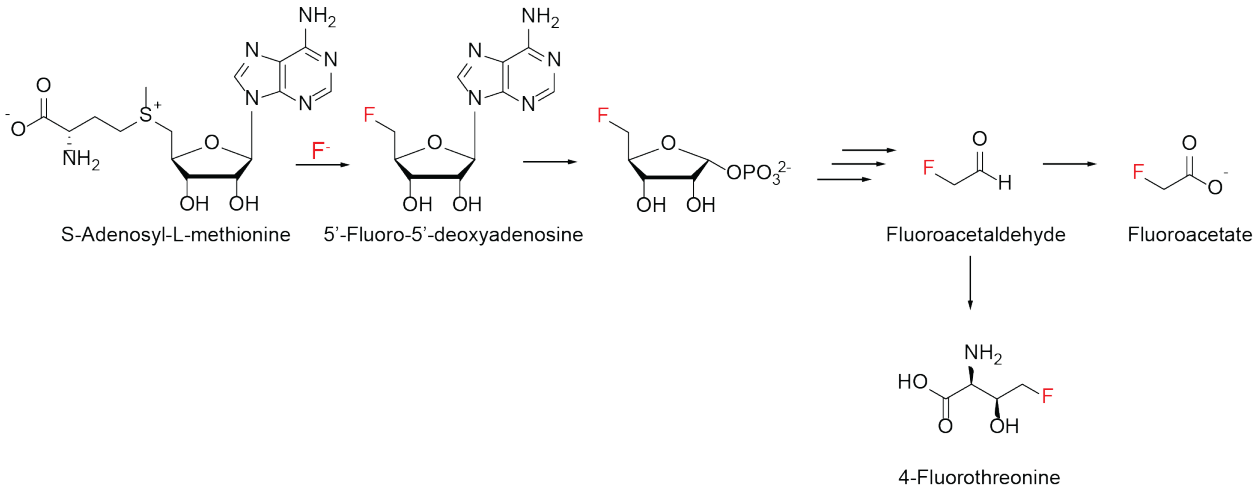


Figure 1.4 *Fluorinase pathway is the only known natural pathway to produce organofluorines from inorganic fluoride. The pathway consists of the fluorinase enzyme, which produces 5'-fluoro-5'-deoxyadenosine (FDA) from S-adenosyl-L-methionine and inorganic fluoride. FDA is then further transformed by a series of enzymatic reactions to form fluoroacetate and fluorothreonine.*

Thus, although natural products are extremely diverse, very few have leveraged the unique elemental properties of fluorine for their functions. There is no known example of the natural presence of fluorine in polyketides, non-ribosomal peptides, isoprenoids, phenylpropanoids, and alkaloids, for example. This lack of precedent highlights the gap between synthetic and natural compounds that could be bridge by introducing non-cognate chemical groups such as fluorine into natural products.

As a result of the scarcity of organofluorines in biological systems, most enzymes have not evolved to recognize fluorine substitution specifically. Notable exceptions are the FlA fluorinase and its homologs [58], the FlK thioesterase [59, 60], and the FthB fluorothreonyl-tRNA deacylase [61]. A number of defluorinase enzymes have also been identified [58, 62]. This provides us with the possibility to introduce fluorine into natural products as a new design element through incorporation of organofluorine precursors by biosynthetic enzymes.

1.3 Polyketide synthase engineering for drug discovery

Polyketide synthase function. Polyketide synthases (PKSs) are remarkable multifunctional enzymatic machinery responsible for the biosynthesis of diverse natural products. They assemble a starter unit with activated 2-carbon acyl-CoA extender units into large intricate structures with a broad range of bioactivities. Well known PKS products include erythromycin, tetracycline, lovastatin, rapamycin, and doxorubicin.

PKSs are organized into modules, which are further organized into domains [63]. The principal catalytic domains of PKSs are closely related to those of fatty acid synthase. The most important domain of each module is the ketosynthase (KS) domain. In addition, most PKS modules also consist of the acyltransferase (AT) domain and the acyl carrier protein (ACP). Each module might also maintain the ketoreductase (KR), the dehydratase (DH), and the enoyl reductase (ER) optional processing domains. A PKS might also consist of the thioesterase (TE) domain.

All domains of a PKS module function together to assemble an extender unit to the growing polyketide chain (*Figure 1.5*) [64]. A module's function starts when a growing polyketide intermediate is transferred to the KS domain and the AT domain selects for an extender unit. Most common extender units of PKSs are malonyl-CoA and methylmalonyl-CoA; however, ethylmalonyl-CoA, chloroethylmalonyl-coA, hydroxymalonyl-ACP, methoxymalonyl-ACP, and aminomalonyl-ACP have all been reported as extender units [65]. The AT domain catalyzes extender unit loading by transferring the carboxyacyl moiety from CoA carrier to the phosphopantetheinyl modification of the AT domain. The KS domain then catalyzes decarboxylative Claisen condensation reaction between the extender unit and the growing polyketide intermediate, elongating the intermediate by two carbons. The keto group in the resulting intermediate, which contains a β-keto thioester motif, might undergo a series of reduction reaction from keto to hydroxy group by KR domain, from hydroxy to alkene by DH domain, and from alkene to alkane by ER domain. Once the intermediate finishes being processed, it is passed on to the KS domain of the following module. Polyketide product is off-loaded from its synthase by either hydrolysis or cyclization reaction, both catalyzed by the TE domain.

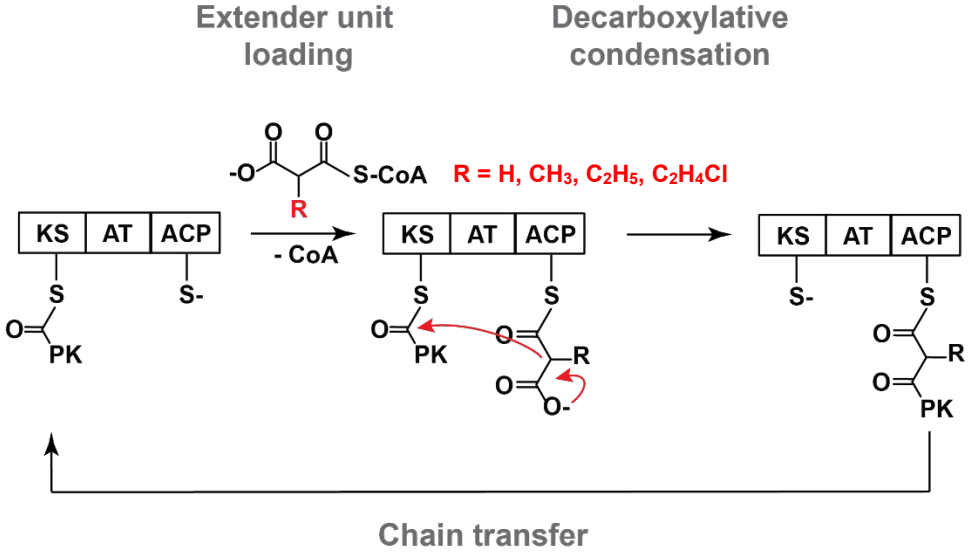


Figure 1.5 *Mechanism of polyketide synthases*

Polyketide synthase diversity and classification. PKS are classified into 3 types based on their domain architecture [66]. Since their initial discoveries during 1980-2000, many examples of all three types of synthases have been illustrated.

Type I PKSs comprise multifunctional polypeptide chains, that each contains multiple domains or modules arranged in tandem. Type I PKSs could be modular or iterative. In type I modular PKSs, individual active sites operate linearly and sequentially in an assembly line fashion on a vectorially transferred protein-bound intermediate [67]. In most cases, an active site within an assembly line only functions once to incorporate one extender unit in the overall pathway, and the order of modules dictate the sequence of biosynthetic events [66, 67]. The vectoral process is achieved by a turnstile-type coupling of elongation and closure of access to upstream module [68]. This domain architecture results in this type of synthases having an immense size. However, it consequently allows different modules to incorporate different types of extender units, affording structural diversity observed among product molecules [69]. Product structures are further diversified by presence or absence of processing domains in different modules [69].

Most type I modular PKSs have all functional domains encoded in large multifunctional proteins; however, a subset of this type of synthases lack cognate AT domains (*cis*-ATs). Instead, the missing AT activity is provided in *trans* by a standalone, discrete AT domain, giving these systems the name *trans*-AT or AT-less PKSs [70]. A *trans*-AT PKS system could consist of one or a few standalone AT domains, and each AT domain usually serves more than one module [71, 72]. As a result of this mechanism, each module of this type of PKS systems served by the same trans-AT incorporates the same extender unit malonyl-CoA and α-substitution of methyl group is afforded by methyltransferases instead [71]. In addition, *trans*-AT PKSs often contain aberrant domain architecture such as non-elongating module, dehydrating bimodules, and *O*-methylation modules, further diversifying product structures [73].

On the other hand, type I iterative PKSs use the same active sites to perform repeated installation of multiple extender units to the product molecule. A system could by fully iterative or partially iterative [74]. Type I iterative PKSs generate many complex ring structures through a programming pattern that defines which modifying domains are used in each cycle. They are found in fungi and bacteria [75]. Fungal type I iterative PKSs are responsible for generating most of the known toxic aromatic fungal metabolites and are classified into fully-reducing, partially-reducing, and non-reducing [76]. Bacterial versions fall into 5 clades, generating 5 class of products namely aromatics, polyunsaturated fatty acids, enediynes, polycyclid tetramate macrolactams, and those arise from *trans*-AT PKSs [74, 77].

Type II PKSs consist of multienzyme complexes, that are made up of individual proteins, each harboring a functional domain. They produce highly functionalized polycyclic aromatic or polyunsaturated compounds [78]. Type II PKS systems contain a KS-chain length factor (CHF) heterodimer and an ACP domain as their core enzyme complex [78, 79]. As most type II synthases do not contain an AT domain, it is thought that an FAS AT is recruited for acyltransferase function[78]. These domains act iteratively to extend the intermediate by malonyl-CoA extender units until a determined length, at which point, the poly-β-keto intermediate undergoes cyclization reaction [79, 80]. The folding of the intermediate is determined by additional subunits such as ketoreductases, cyclases aromatases, oxygenase, and transferases [80]. The cyclized aromatic core products are further tailored by redox, group transfer, hydrolysis, and rearrangement reactions, creating product diversity [78, 80].

Type III synthases are homodimers of KS domain subunits that performs all the functions typically played by KS, AT, and ACP domains in type I and II PKSs iteratively [81]. They do not use ACP domains – the KS domain of type III synthases utilize carboxyacyl-CoA extender unit directly – and do not possess any processing domains. Characterized type III synthases also overwhelmingly utilize only malonyl-coA extender unit [81]. Thus, the diversity of products arises from the starter unit [82], the number of the extender units incorporated [83], and the intramolecular cyclization termination reaction [84]. The iterative process is terminated by Claisen condensation, aldol condensation, or lactonization [81, 85]. Type III PKSs are found widely in plants but a handful have also been confirmed in fungi and bacteria [86].

Type I modular polyketide synthase extender unit selectivity and engineering. Type I modular PKSs have been subjected to many genetic, biochemical, and engineering studies due to their promise in rational redesigning to generate large diversity of combinatorial products. Of a particular significance has been the establishment of collinearity between synthase nucleotide sequence to polyketide structure [87, 88]. They have contributed numerous antibiotics to the market, and the search for novel ones intensified interests in manipulating their biosynthetic pathways. As the structural complexity of this type of molecules places tremendous synthetic challenges, efforts were focused on modifying their structures through biological implementations.

The AT domain received special attention in this matter as it dictates the identity of the extender unit incorporated into the polyketide product structures [89]. Many AT domains of type I modular PKSs share high sequence similarity that they can be identified computationally base on their sequences alone [90]. Interestingly, it was found that sequences of AT domains clade according to extender unit specificity, especially when only active site motifs are considered [90, 91]. These findings provide the basis for AT domain engineering approaches for increasing polyketide synthase products scope.

One of the most prominent and most common AT engineering approaches for PKSs is the substitution of the AT domain with an AT of a different substrate specificity, or of the module with a module of a different substrate specificity [89]. This approach stems from a vision that modular PKS domains and modules have lego-like modularity [92]. This process has yielded many new polyketide analogs [93]. However, the success often comes at the cost of decreased titers, thought to be due to disruptions to protein-protein interactions and the narrow specificity of downstream modules [92].

Appreciation of the protein-protein interactions in polyketide synthase processes and recent advances in bioinformatics led to site-specific mutagenesis approach to alter AT substrate specificity. Alignment analysis of amino acid sequences of AT domains found a short motif located approximately 100 residues C-terminal to the active site serine to be highly conserved and predictive of AT specificity [94, 95]. Methylmalonyl-coA specific AT domains have YASH whereas malonyl-coA specific AT domains have HAFH at this position [94, 95]. This finding let to AT engineering efforts that gave rise to many novel product analogs [96, 97]. However, a total switch of activity remains to be achieved.

The last commonly used strategy for modifying AT domain substrate specificity is complementation. In this strategy, the cognate AT domain (*cis*-AT) is inactivated through an active site mutation, and the missing AT activity is provided through an introduction of a discrete AT protein (*trans*-AT) [89, 92]. This approach has the potential to be minimally invasive, causing minimal perturbation to protein-protein interactions. This approach was encouraged by the finding

that many *trans*-AT domains have fast kinetics and are promiscuous towards ACP co-substrates [98, 99]. The possibility has been demonstrated both *in vitro* [99, 100] and *in vivo* [101] although the efficiency of *trans*-AT complementation is affected by the identity of the extender unit and the ACP.

The many approaches available for altering the substrate specificity of AT domains provide opportunities for introducing non-cognate extender unit such as fluorinated extender unit into polyketide molecules, which would further diversify chemical space accessible by polyketide biosynthetic machinery.

1.4 Thesis Organization

The goal of this work is to develop a method that enables the biosynthesis of fluorinated polyketide natural products. We aim to construct a generalizable platform that allows regioselective fluorine incorporation into all positions of all polyketide backbones. This work had built on previous works in the lab that described generation of fluorinated extender unit and acceptance of the fluorinated extender unit by polyketide synthase enzymes [102], mechanism of fluorinated extender unit selectivity and incorporation [103], and identification and implementation of fluorinated extender unit accumulation strategy in *E. coli* production host [104].

The following chapters report a contribution towards realizing such goals. In Chapter 2, characterization of fluorinated extender unit-specific *trans*-AT domain and its implementation *in vitro* to produce complex fluorinated polyketide molecules are described. Chapter 3 details a construction and characterization of model, single-modular fluorinated polyketide biosynthetic pathway in *E. coli* production host. Chapter 4 expands on work reported in chapter 3 by applying the pathway and concepts learned to biosynthesize fluorinated analogs of complex, multi-modular polyketide molecules. Chapter 5 explains conceptually a study of the fate of incorporated fluorinated extender unit when encountering downstream processing domains and in inter-modular transfer process.

1.5. References

1. Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. Counting on natural products for drug design. *Nat Chem* **2016**, *8*, 531-541.

2. Henkel, T.; Brunne, R. M.; Müller, H.; Reichel, F. Statistical Investigation into the Structural Complementarity of Natural Products and Synthetic Compounds. *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 643-647.

3. Patridge, E.; Gareiss, P.; Kinch, M. S.; Hoyer, D. An analysis of FDA-approved drugs: natural products and their derivatives. *Drug Discov. Today* **2016**, *21*, 204-207.

4. Newman, D. J.; Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770-803.

5. Moloney, M. G. Natural products as a source for novel antibiotics. *Trends in Pharmacological Sciences* **2016***, 37*, 689-701.

6. Wright, G. D. Opportunities for natural products in 21st century antibiotic discovery. *Nat. Prod. Rep.* **2017***, 34*, 694-701.

7. Dutta, S.; Mahalanobish, S.; Saha, S.; Ghosh, S.; Sil, P. C. Natural products: An upcoming therapeutic approach to cancer. *Food and Chemical Toxicology* **2019***, 128*, 240-255.

8. Cragg, G. M.; Pezzuto, J. M. Natural products as a vital source for the discovery of cancer chemotherapeutic and chemopreventive agents. *MPP* **2016***, 25*, 41-59.

9. Waltenberger, B.; Mocan, A.; Šmejkal, K.; Heiss, E. H.; Atanasov, A. G. Natural products to counteract the epidemic of cardiovascular and metabolic disorders. *Molecules* **2016***, 21*.

10. Mann, J. Natural products as immunosuppressive agents. *Nat. Prod. Rep.* **2001***, 18*, 417-430.

11. Dudics, S.; Langan, D.; Meka, R. R.; Venkatesha, S. H.; Berman, B. M.; Che, C.; Moudgil, K. D. Natural products for the treatment of autoimmune arthritis: their mechanisms of action, targeted delivery, and interplay with the host microbiome. *Int J Mol Sci* **2018***, 19*.

12. Di Paolo, M.; Papi, L.; Gori, F.; Turillazzi, E. Natural products in neurodegenerative diseases: a great promise but an ethical challenge. *Int J Mol Sci* **2019***, 20*.

13. Essa, M. M.; Vijayan, R. K.; Castellano-Gonzalez, G.; Memon, M. A.; Braidy, N.; Guillemin, G. J. Neuroprotective effect of natural products against Alzheimer's disease. *Neurochem. Res.* **2012***, 37*, 1829-1842.

14. Steinberg, D. Thematic review series: the pathogenesis of atherosclerosis. An interpretive history of the cholesterol controversy, part V: the discovery of the statins and the end of the controversy. *J. Lipid Res.* **2006***, 47*, 1339-1351.

15. Simvastatin: drug usage statistics, United States, 2007 - 2017. *https://clincalc.com/DrugStats/Drugs/Simvastatin* (accessed May, 2020).

16. Zaffiri, L.; Gardner, J.; Toledo-Pereyra, L. H. History of antibiotics from Salvarsan to Cephalosporins. *Journal of Investigative Surgery* **2012***, 25*, 67-77.

17. Adedeji, W. A. The treasure called antibiotics. *Ann Ib Postgrad Med* **2016***, 14*, 56-57.

18. Rossiter, S. E.; Fletcher, M. H.; Wuest, W. M. Natural products as platforms to overcome antibiotic resistance. *Chem. Rev.* **2017***, 117*, 12415-12474.

19. Peláez, F. The historical delivery of antibiotics from microbial natural products—Can history repeat? *Biochemical Pharmacology* **2006***, 71*, 981-990.

20. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov* **2015**, *14*, 111-129.

21. Li, J. W. -.; Vederas, J. C. Drug discovery and natural products: end of an era or an endless frontier? *Science* **2009**, *325*, 161-165.

22. Janzen, W. Screening technologies for small molecule discovery: the state of the art. *Chemistry & Biology* **2014**, *21*, 1162-1170.

23. Shen, B. A new golden age of natural products drug discovery. *Cell* **2015**, *163*, 1297-1300.

24. Sukuru, S. C. K.; Jenkins, J. L.; Beckwith, R. E. J.; Scheiber, J.; Bender, A.; Mikhailov, D.; Davies, J. W.; Glick, M. Plate-based diversity selection based on empirical HTS data to enhance the number of hits and their chemical diversity. *J Biomol Screen* **2009**, *14*, 690-699.

25. Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P & T : a peer-reviewed journal for formulary management* **2015**, *40*, 277-283.

26. Antibiotic-resistant Germs: New Threats. https://www.cdc.gov/drugresistance/biggest-threats.html (accessed May 11, 2020).

27. Lewis, K. New approaches to antimicrobial discovery. *Biochem. Pharmacol.* **2017**, *134*, 87-98.

28. Katz, L.; Baltz, R. H. Natural product discovery: past, present, and future. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 155-176.

29. Wagenaar, M. M. Pre-fractionated microbial samples--the second generation natural products library at Wyeth. *Molecules* **2008**, *13*, 1406-1426.

30. P, R. Fluorination methods for drug discovery and development. *Expert Opin Drug Discov* **2016**, *11*, 983-999.

31. Furuya, T.; Kamlet, A. S.; Ritter, T. Catalysis for fluorination and trifluoromethylation. *Nature* **2011**, *473*, 470-477.

32. Müller, K.; Faeh, C.; Diederich, F. Fluorine in pharmaceuticals: looking beyond intuition. *Science* **2007**, *317*, 1881-1886.

33. Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V. Fluorine in medicinal chemistry. *Chem Soc Rev* **2008**, *37*, 320-330.

34. Shah, P.; Westwell, A. D. The role of fluorine in medicinal chemistry. *J Enzyme Inhib Med Chem* **2007**, *22*, 527-540.

35. Gillis, E. P.; Eastman, K. J.; Hill, M. D.; Donnelly, D. J.; Meanwell, N. A. Applications of fluorine in medicinal chemistry. *J. Med. Chem.* **2015***, 58*, 8315-8359.

36. Smart, B. E. Fluorine substituent effects (on bioactivity). *Journal of Fluorine Chemistry* **2001***, 109*, 3-11.

37. Meanwell, N. A. Fluorine and fluorinated motifs in the design and application of bioisosteres for drug design. *J. Med. Chem.* **2018***, 61*, 5822-5880.

38. Park, B. K.; Kitteringham, N. R.; O'Neill, P. M. Metabolism of fluorine-containing drugs. *Annual Review of Pharmacology and Toxicology* **2001***, 41*, 443-470.

39. Middleton, W. J. New fluorinating reagents: dialkylaminosulfur fluorides. *J. Org. Chem.* **1975***, 40*, 574-578.

40. Banks, R. E.; Mohialdin-Khaffaf, S. N.; Lal, G. S.; Sharif, I.; Syvret, R. G. 1-Alkyl-4-fluoro-1,4-diazoniabicyclo[2.2.2]octane salts: a novel family of electrophilic fluorinating agents. *J. Chem. Soc. , Chem. Commun.* **1992**, 595-596.

41. L'Heureux, A.; Beaulieu, F.; Bennett, C.; Bill, D. R.; Clayton, S.; LaFlamme, F.; Mirmehrabi, M.; Tadayon, S.; Tovell, D.; Couturier, M. Aminodifluorosulfinium Salts: Selective fluorination reagents with enhanced thermal stability and ease of handling, *J. Org. Chem.* **2010***, 75*, 3401-3411.

42. Mohammadkhani, L.; Heravi, M. M. XtalFluor-E: A useful and versatile reagent in organic transformations. *Journal of Fluorine Chemistry* **2019***, 225*, 11-20.

43. Umemoto, T.; Singh, R. P.; Xu, Y.; Saito, N. Discovery of 4-tert-Butyl-2,6-dimethylphenylsulfur trifluoride as a deoxofluorinating agent with high thermal stability as well as unusual resistance to aqueous hydrolysis, and its diverse fluorination capabilities including deoxofluoro-arylsulfinylation with high stereoselectivity. *J. Am. Chem. Soc.* **2010***, 132*, 18199-18205.

44. Nielsen, M. K.; Ugaz, C. R.; Li, W.; Doyle, A. G. PyFluor: a low-cost, stable, and selective deoxyfluorination reagent. *J. Am. Chem. Soc.* **2015***, 137*, 9571-9574.

45. Neumann, C. N.; Ritter, T. Facile C–F bond formation through a concerted nucleophilic aromatic substitution mediated by the PhenoFluor reagent. *Acc. Chem. Res.* **2017***, 50*, 2822-2833.

46. Nyffeler, P. T.; Durón, S. G.; Burkart, M. D.; Vincent, S. P.; Wong, C. Selectfluor: mechanistic insight and applications. *Angew. Chem. Int. Ed. Engl.* **2004***, 44*, 192-212.

47. Differding, E.; Ofner, H. N-Fluorobenzenesulfonimide: a practical reagent for electrophilic fluorinations. *Synlett* **1991***, 1991*, 187-189.

48. Rostami, A. N-Fluorobenzenesulfonimide [(PhSO2)2NF] - A neutral N-F-containing electrophilic fluorinating agent. *Synlett* **2007***, 2007*, 2924-2925.

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

50. O'Hagan, D.; B. Harper, D. Fluorine-containing natural products. *Journal of Fluorine Chemistry* **1999***, 100*, 127-133.

51. Budisa, N.; Kubyshkin, V.; Schulze-Makuch, D. Fluorine-rich planetary environments as possible habitats for life. *Life (Basel)* **2014***, 4*, 374-385.

52. Leong, L. E. X.; Khan, S.; Davis, C. K.; Denman, S. E.; McSweeney, C. S. Fluoroacetate in plants - a review of its distribution, toxicity to livestock and microbial detoxification. *J Anim Sci Biotechnol* **2017***, 8*.

53. Thomas, S. O.; Singleton, V. L.; Lowery, J. A.; Sharpe, R. W.; Pruess, L. M.; Porter, J. N.; Mowat, J. H.; Bohonos, N. Nucleocidin, a new antibiotic with activity against trypanosomes. *Antibiot Annu* **1957**, 716-721.

54. Wu, L.; Maglangit, F.; Deng, H. Fluorine biocatalysis. *Current Opinion in Chemical Biology* **2020***, 55*, 119-126.

55. Deng, H.; Cross, S. M.; McGlinchey, R. P.; Hamilton, J. T. G.; O'Hagan, D. In vitro reconstituted biotransformation of 4-fluorothreonine from fluoride ion: application of the fluorinase. *Chemistry & Biology* **2008***, 15*, 1268-1276.

56. Deng, H.; O'Hagan, D. The fluorinase, the chlorinase and the duf-62 enzymes. *Current Opinion in Chemical Biology* **2008***, 12*, 582-592.

57. Lowe, P. T.; Cobb, S. L.; O'Hagan, D. An enzymatic finkelstein reaction: fluorinase catalyses direct halogen exchange. *Org. Biomol. Chem.* **2019***, 17*, 7493-7496.

58. Tong, W.; Huang, Q.; Li, M.; Wang, J. Enzyme-catalyzed C–F bond formation and cleavage. *Bioresources and Bioprocessing* **2019***, 6*.

59. Weeks, A. M.; Chang, M. C. Y. Catalytic control of enzymatic fluorine specificity. *PNAS* **2012***, 109*, 19667-19672.

60. Weeks, A. M.; Wang, N.; Pelton, J. G.; Chang, M. C. Y. Entropy drives selective fluorine recognition in the fluoroacetyl–CoA thioesterase from Streptomyces cattleya. *PNAS* **2018***, 115*, E2193-E2201.

61. McMurry, J. L.; Chang, M. C. Y. Fluorothreonyl-tRNA deacylase prevents mistranslation in the orgaofluorine producer Streptomyces cattleya. *PNAS* **2017***, 114*, 11920-11925.

62. Seong, H. J.; Kwon, S. W.; Seo, C.; Kim, J.; Jang, Y. Enzymatic defluorination of fluorinated compounds. *Applied Biological Chemistry* **2019***, 62*.

63. Bayly, C. L.; Yadav, V. G. Towards precision engineering of canonical polyketide synthase domains: recent advances and future prospects. *Molecules* **2017***, 22*.

64. Keating, T. A.; Walsh, C. T. Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis. *Curr Opin Chem Biol* **1999***, 3*, 598-606.

65. Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed. Engl.* **2009***, 48*, 4688-4716.

66. Shen, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* **2003***, 7*, 285-295.

67. Robbins, T.; Liu, Y.; Cane, D. E.; Khosla, C. Structure and mechanism of assembly line polyketide synthases. *Curr. Opin. Struct. Biol.* **2016***, 41*, 10-18.

68. 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*, 14-20.

69. Fischbach, M. A.; Walsh, C. T. Assembly-Line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **2006***, 106*, 3468-3496.

70. Helfrich, E. J. N.; Piel, J. Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat Prod Rep* **2016***, 33*, 231-316.

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

72. Cheng, Y.; Tang, G.; Shen, B. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **2003***, 100*, 3149-3154.

73. Jenner, M.; Frank, S.; Kampa, A.; Kohlhaas, C.; Pöplau, P.; Briggs, G. S.; Piel, J.; Oldham, N. J. Substrate specificity in ketosynthase domains from trans-AT polyketide synthases. *Angewandte Chemie International Edition* **2013***, 52*, 1143-1147.

74. Chen, H.; Du, L. Iterative polyketide biosynthesis by modular polyketide synthases in bacteria. *Appl. Microbiol. Biotechnol.* **2016***, 100*, 541-557.

75. Herbst, D. A.; Townsend, C. A.; Maier, T. The architectures of iterative type I PKS and FAS. *Nat Prod Rep* **2018***, 35*, 1046-1069.

76. Cox, R. J. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org. Biomol. Chem.* **2007***, 5*, 2010-2026.

77. Wang, B.; Guo, F.; Huang, C.; Zhao, H. Unraveling the iterative type I polyketide synthases hidden in Streptomyces. *PNAS* **2020***, 117*, 8449-8454.

78. Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* **2007***, 24*, 162-190.

79. Chen, A.; Re, R. N.; Burkart, M. D. Type II fatty acid and polyketide synthases: deciphering protein-protein and protein-substrate interactions. *Nat Prod Rep* **2018***, 35*, 1029-1045.

80. Zhang, Z.; Pan, H.; Tang, G. New insights into bacterial type II polyketide biosynthesis. *F1000Res* **2017***, 6*, 172.

81. Shimizu, Y.; Ogata, H.; Goto, S. Type III polyketide synthases: Functional classification and phylogenomics. *Chembiochem* **2017***, 18*, 50-65.

82. Jez, J. M.; Bowman, M. E.; Noel, J. P. Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. *PNAS* **2002***, 99*, 5319-5324.

83. Abe, I.; Oguro, S.; Utsumi, Y.; Sano, Y.; Noguchi, H. Engineered biosynthesis of plant polyketides: chain length control in an octaketide-producing plant type III polyketide synthase. *J. Am. Chem. Soc.* **2005***, 127*, 12709-12716.

84. Austin, M. B.; Bowman, M. E.; Ferrer, J.; Schröder, J.; Noel, J. P. An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. *Chemistry & Biology* **2004***, 11*, 1179-1194.

85. Austin, M. B.; Noel, J. P. The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* **2003***, 20*, 79-110.

86. Yu, D.; Xu, F.; Zeng, J.; Zhan, J. Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* **2012***, 64*, 285-295.

87. Minowa, Y.; Araki, M.; Kanehisa, M. Comprehensive analysis of distinctive polyketide and nonribosomal peptide structural motifs encoded in microbial genomes. *J. Mol. Biol.* **2007***, 368*, 1500-1517.

88. Callahan, B.; Thattai, M.; Shraiman, B. I. Emergent gene order in a model of modular polyketide synthases. *Proc. Natl. Acad. Sci. U. S. A.* **2009***, 106*, 19410-19415.

89. Musiol-Kroll, E. M.; Wohlleben, W. Acyltransferases as tools for polyketide synthase engineering. *Antibiotics (Basel)* **2018***, 7*.

90. Yadav, G.; Gokhale, R. S.; Mohanty, D. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *Journal of Molecular Biology* **2003***, 328*, 335-363.

91. Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E. Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.* **2005***, 22*, 2027-2039.

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

93. Weissman, K. J. Genetic engineering of modular PKSs: from combinatorial biosynthesis to synthetic biology. *Nat Prod Rep* **2016***, 33*, 203-230.

94. Del Vecchio, F.; Petkovic, H.; Kendrew, S. G.; Low, L.; Wilkinson, B.; Lill, R.; Cortés, J.; Rudd, B. A. M.; Staunton, J.; Leadlay, P. F. Active-site residue, domain and module swaps in modular polyketide synthases. *J. Ind. Microbiol. Biotechnol.* **2003***, 30*, 489-494.

95. 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*, 15464-15470.

96. Kushnir, S.; Sundermann, U.; Yahiaoui, S.; Brockmeyer, A.; Janning, P.; Schulz, F. Minimally invasive mutagenesis gives rise to a biosynthetic polyketide library. *Angew. Chem. Int. Ed. Engl.* **2012***, 51*, 10664-10669.

97. 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*, 443-450.

98. 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*, 95-102.

99. Kumar, P.; Koppisch, A. T.; Cane, D. E.; Khosla, C. Enhancing the modularity of the modular polyketide synthases: transacylation in modular polyketide synthases catalyzed by malonyl-CoA:ACP transacylase. *J. Am. Chem. Soc.* **2003***, 125*, 14307-14312.

100. 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*, 3796-3806.

101. Lopanik, N. B.; Shields, J. A.; Buchholz, T. J.; Rath, C. M.; Hothersall, J.; Haygood, M. G.; Håkansson, K.; Thomas, C. M.; Sherman, D. H. In vivo and in vitro trans-acylation by BryP, the putative bryostatin pathway acyltransferase derived from an uncultured marine symbiont. *Chem. Biol.* **2008***, 15*, 1175-1186.

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

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

104. Thuronyi, B. W.; Privalsky, T. M.; Chang, M. C. Y. Engineered fluorine metabolism and fluoropolymer production in living cells. *Angewandte Chemie International Edition* **2017***, 56*, 13637-13640.

Chapter 2: *Regioselective incorporation of fluorine functional group into polyketide molecules in vitro*

Portions of this work were performed in collaboration with the following persons: Thomas Privalski assisted with in vitro 6-dEB production experiments. Dr. Hongjun Dong assisted with introduction active-site mutations into the AT domains of DEBS pathway proteins.

2.1 Introduction

Despite the prevalence of fluorine in synthetic bioactive compounds, fluorinated natural products are quite rare with <25 characterized. We have focused on developing new methods for the site-selective incorporation of fluorine using the modular Type I polyketide synthase systems, exemplified by 6-deoxyerythronolide B synthase (DEBS), which is responsible for the biosynthesis of the erythromycin precursor. Previous work in the laboratory had shown that fluoromalonyl-CoA could be utilized as an extender unit for incorporation of fluorine into simple polyketide fragments, such as triketide and tetraketide lactones, through a complementation strategy in which a *trans*-AT protein is used to complement an inactivated *cis*-AT domain in the module of interest [1, 2].

The selectivity of a polyketide module is mainly determined by the selectivity of its AT domain [3-5], which selects the carboxyacyl-CoA extender unit and transfers it to the active site of the ACP domain to enable extender unit condensation while excluding other extender units by selective binding or hydrolysis [6-8]. The mutation in the *cis*-AT domain active site thus serves to suppress its activity and natural extender unit selectivity while preserving the structural integrity of the polyketide synthase protein. The *trans*-AT protein can then provide the missing transacylase activity by interacting with the intact ACP domain (*Figure 2.1*). For modules with active AT domains, the active *cis*-AT domain should trans-acylate the ACP domain at a much faster rate than a *trans*-AT due to proximity effects. Thus, the *trans*-AT protein should only successfully complement modules that can both interact with the *trans*-AT and do not contain a functional AT domain. As such, this kinetic resolution should allow for regioselective incorporation of fluoromalonyl-CoA only in positions where the native *cis*-AT had been inactivated.

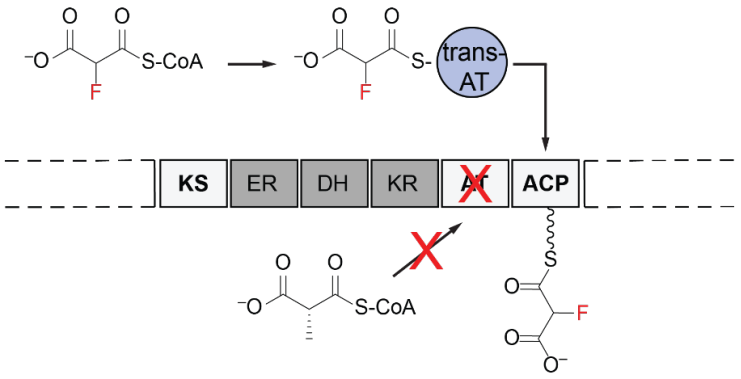


Figure 2.1. *Complementation strategy for altering the extender unit selectivity of a polyketide synthase module to accept fluoromalonyl-coA.*

By replacing a *cis*-AT domain with a *trans*-AT domain, the selectivity of a PKS module now lies physically outside of polypeptide chains that make up the megasynthase [6, 9, 10]. As *trans*-AT proteins are not physically a part of the assembly line polyketide megasynthases but a standalone protein of <50 kDa, it makes for a much simpler target for protein engineering. Using this strategy, we could engineer a single *trans*-AT protein to be selective for fluoromalonyl-CoA in order to produce different regioisomers of a fluorinated polyketide molecule. This enzyme would also need to exclude the major competing extender units, such as malonyl- and methylmalonyl-CoA, which are simultaneously present in the cell and needed to build the rest of the backbone of the product molecule. In the first-generation system, we were able to generate fluorine-containing polyketides using wild-type DszAT, the *trans*-AT from the biosynthetic pathway for disorazole encoded by *dszD* gene found in the myxobacterium *Sorangium cellulosum* So ce12 [11]. DszAT was chosen from many known *trans*-AT proteins due to its native preference

for malonyl-CoA but relaxed selectivity towards ACP partners [12]. DszAT loads all seven ACP domains in the disorazole biosynthetic pathways and has been shown to be able to load various DEBS ACP domains [12]. As fluoromalonyl-CoA is similar in size to malonyl-CcoA and DszAT is able to exclude methylmalonyl-CoA based on sterics [13], we reasoned that DszAT may be able to accept the fluoromalonyl-CoA extender unit and trans-acylate it onto ACP partners. Indeed, we observed a significant increase in fluorinated triketide lactone (F-TKL) yield when DszAT was used to complement inactivation mutation in $Mod3_{DEBS}$+TE(AT^0) and $Mod6_{DEBS}$+TE(AT^0) [2].

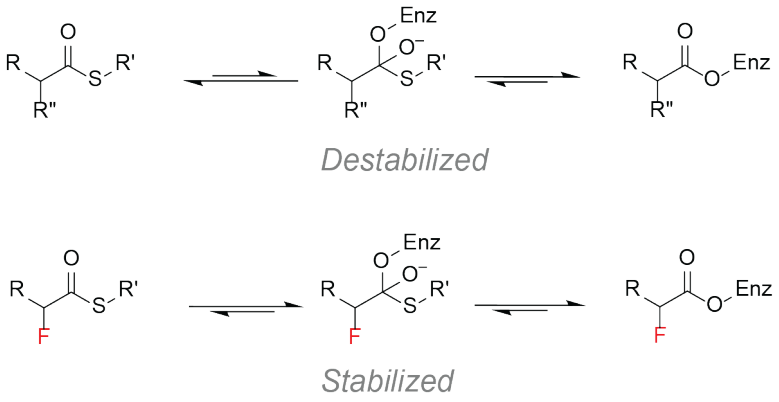


Figure 2.2. *Mechanism of transacylation reaction. The first step in transacylation reaction involves nucleophilic addition by active site serine to carbonyl carbon, leading to the formation of tetrahedral intermediate that is stabilized by fluorine at the α-position.*

Although wild-type DszAT acts on both malonyl-CoA and fluoromalonyl-CoA, we thought it may be possible to increase its selectivity towards fluoromalonyl-CoA through mutagenesis of the active site. The transacylation reaction catalyzed by DszAT is initiated by nucleophilic attack by an active-site serine and also goes through a tetrahedral intermediate, both of which are affected by the presence of fluorine atom at the α-position (*Figure 2.2*). For nucleophilic attack, the fluorine would be activating as the carbonyl should be more electrophilic and mutations in the active site that lower reactivity could have a more detrimental effect on the unactivated malonyl-CoA substrate. With regard to the tetrahedral intermediate, α-fluoro thioesters are more stable in the hydrated form (tetrahedral intermediate) than the keto form (substrate) due to the fluorine atom's preference to be adjacent to an sp^3 carbon over an sp^2 carbon [14, 15]. This is in contrast to normal thioesters, which are more stable in the keto form. Consequently, the activation energy of transcylation reaction of α-fluoro thioesters should be lower than that of unsubstituted thioesters, lending the reaction to occur more readily with fluorinated substrate.

Upon examination of the crystal structure of DszAT bound to malonate (PDB ID# 6APG) to determine potential sites for engineering [16], we found that DszAT possesses a phenylalanine residue at position 190, which has previously been found conserved in all malonyl-CoA-selective AT domains from many *cis*-AT PKSs and thus presumed to be important for selectivity against methylmalonyl-CoA (*Figure 2.3*) [17, 18]. Another residue that could be targeted is Leu87, whose backbone amide is found near and may play a role in positioning the substrate (*Figure 2.3*). Mutation of this leucine position is hypothesized to destabilize the formation of the tetrahedral intermediate, potentially

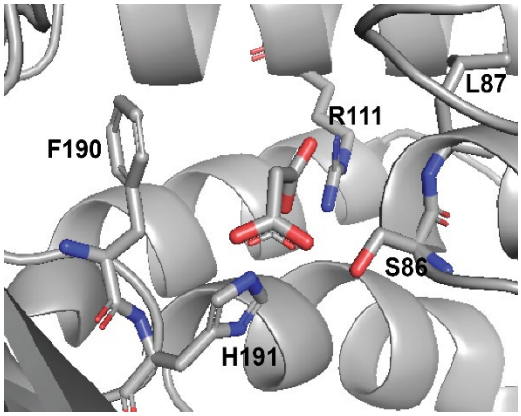


Figure 2.3. *Active site of DszAT bound to malonate. Residues near the active site that have been targeted for engineering are labeled.*

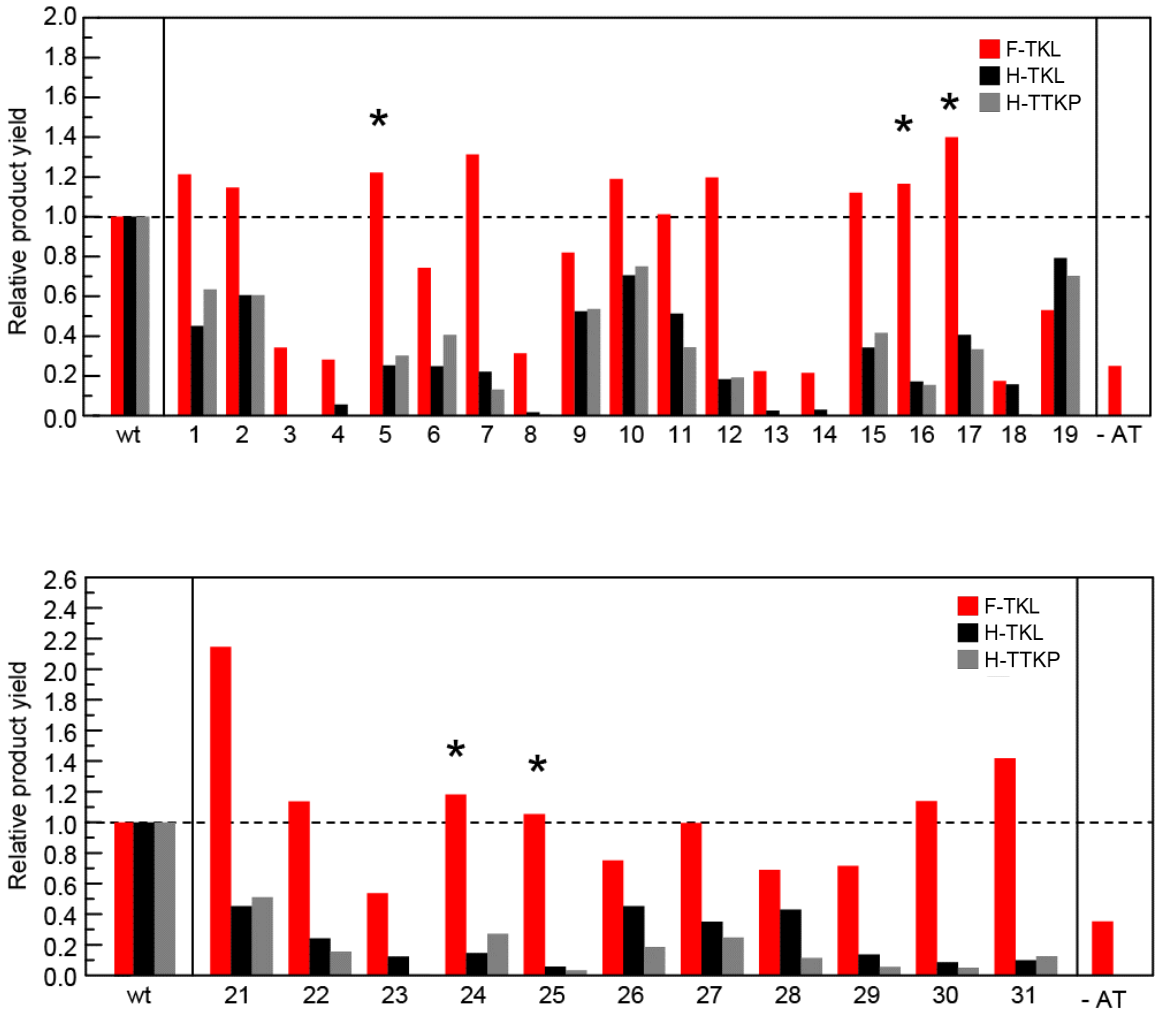


Figure 2.4. *DszAT mutant library screen. Lysate of E. coli expressing following mutants were screened through triketide lactone production assay with purified Mod3$_{DEBS}$+TE(AT^{0}): 1) F190A, 2) F190C, 3) F190D, 4) F190E, 5) F190G, 6) F190H, 7) F190I, 8) F190K, 9) F190L, 10) F190M, 11) F190N, 12)F190P, 13) F190Q, 14) F190R, 15) F190S, 16) F190T, 17) F190V, 18) F190W, 19) F190Y, 21) F190G L87V, 22) F190G L87A, 23) F190I L87V, 24) F190I L87A, 25) F190P L87A, 26) F190S L87V, 27) F190S L87A, 28) F190T L87V, 29) F190T L87A, 30) F190V L87V, 31) F190V L87A. The amount of F-TKL (red), H-TKL (black), and H-TTKP (grey) products are shown normalized by the amount of corresponding product produced by wild-type enzyme.*

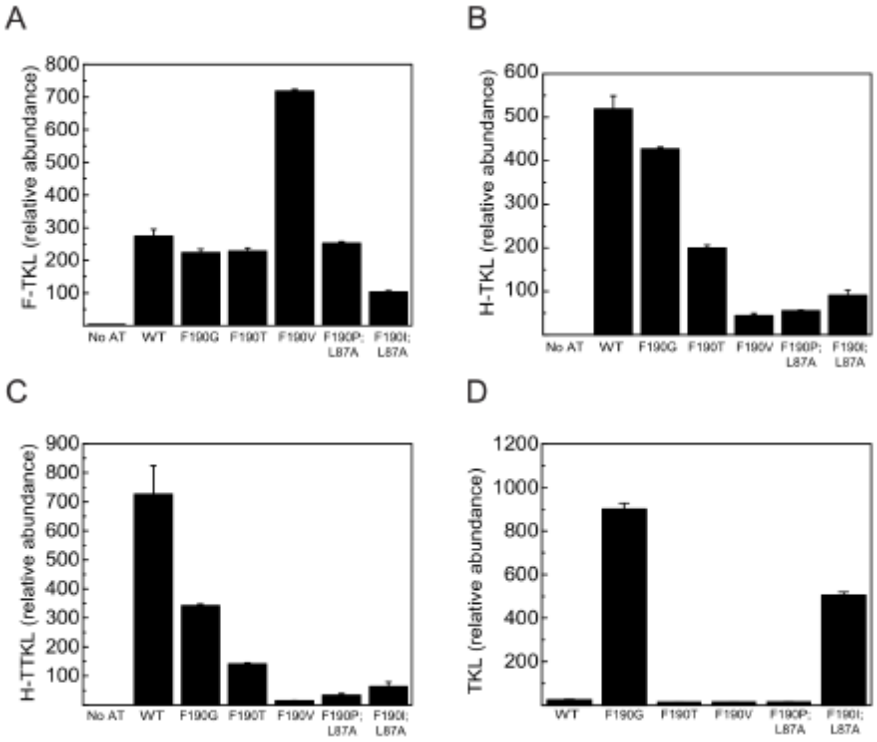


Figure 2.5. *Select DszAT mutants in vitro screen. Promising DszAT mutants from lysate screen, F190G, F190T, F190V, F190P L87A, and F190I L87A, were purified and assessed through in vitro triketide lactone assay with purified* $Mod3_{DEBS}$+TE(AT^0) *for production of A) F-TKL from fluoromalonate, B)H-TKL from malonate, C) H-TTKP from malonate, and D) TKL from methylmalonate. Data are mean ± s.e. (n=3).*

allowing to increase the natural catalytic selectivity towards a α-fluoro thioester substrate. In addition, double mutants containing mutations at His191 and the catalytic Ser86 were designed based on observations made on FlK [19, 20]. Specifically, a mutation to the histidine residue in the catalytic triad of FlK, analogous to H191 in DszAT, alters the catalytic power of the triad which resulted in less hydrolysis of the acyl-enzyme intermediate. Given increased hydrolysis activity of DszAT on fluoromalonyl-CoA, we though this strategy might limit unproductive substrate hydrolysis.

Previous students in the group constructed libraries of DszAT mutants based on these design criteria. A saturation mutagenesis library was constructed at position 190 along with select double mutants of positions H191 and S86 [21, 22]. After preliminary screening of these libraries in cell lysates using a triketide lactone (TKL) assay with purified $Mod3_{DEBS}$+TE(AT^0), a double mutant library was constructed with select F190 mutants that showed promise by adding mutations to L87 position [2]. The results of the lysate screen are reproduced here for reference showing the relative production of the products with fluoromalonyl-CoA (F-TKL) and malonyl-CoA (H-TKL and H-TTKP) compared to wild-type (*Figure 2.4).* Despite caveats that these products are not produced

at equimolar levels and that screening was carried out on a single replicate, several mutants with increased production levels of F-TKL and decreased production of H-TKL and H-TTKP were observed, displaying the increased selectivity for fluoromalonyl-CoA that was sought.

Following the initial lysate screen, mutants that showed promising activities, F190G, F190T, F190V, F190P L87A, and F190I L87A, were heterologously overexpressed, purified, and investigated through an *in vitro* triketide lactone assay using $Mod3_{DEBS}$+TE(AT^0) . This *in vitro* assay has an advantage over lysate screening because the concentration of DszAT mutants are controlled, allowing for more direct comparisons between variants. Selected mutants were tested with malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA substrates. The results are reproduced here for reference (*Figure 2.5*). From these experiments, F190V mutant emerged as the most promising DszAT mutant with increased production of F-TKL and decreased production of H-TKL compared to the wild-type enzyme, while maintaining a similar level of discrimination against TKL to the wild-type enzyme.

This chapter presents the work to follow up the preliminary results mentioned here. Specifically, we sought to characterize the F190V DszAT mutant through steady-state kinetic experiments to examine the nature of its apparent fluorine selectivity. Furthermore, we sought to demonstrate the benefits of such mutant when utilized in complex polyketide synthase system.

2.2 Materials and Methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), magnesium sulfate anhydrous, and glycerol were purchased from EMD Biosciences (Damstadt, Germany). Carbenicillin (Cb), glucose, isopropyl-β-D-thiogalactopyranoside (IPTG), sodium chloride, calcium chloride, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, dichloromethane, ethyl acetate, ethylene diamine tetraacetic acid disodium dihydrate (EDTA), sodium phosphate monobasic, potassium phosphate monobasic, ammonium chloride, and arabinose were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A trilithium sodium salt (CoA), malonyl-CoA, methylmalonyl-CoA, malonic acid, methylmalonic acid, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenolpyruvate (PEP), adenosine trisodium phosphate sodium salt (ATP), myokinase, pyruvate kinase, lactate dehydrogenase, poly(ethyleneimine) (PEI), β-mercaptoethanol (BME), thiamine pyrophosphate (TPP), α-ketoglutaric acid, sodium phosphate dibasic heptahydrate, cysteamine, 4-hydroxy-6-methyl-2-pyrone acetic anhydride, 1-ethyl-3-(-dimethylaminopropyl)carbodiimide hydrochloride (EDC), 4-dimethylaminopyridine (DMAP), reduced β-nicotinamide adenine dinucleotide 2′-phosphate (NADPH), acetonitrile, dimethyl sulfoxide (DMSO), sodium propionate, chloramphenicol, α-ketoglutarate dehydrogenase, and bovine serum albumin were purchased from Sigma-Aldrich (St. Louis, MO). Diethylfluoromalonate was purchased from Sigma-Aldrich (St. Louise, MO) or from Matrix Scientific (West Columbia, SC). Formic acid, dithiothreitol (DTT), and perchloric acid were purchased from Fluka (Charlotte, NC). 2-methyl-3-oxopentanoic acid ethyl ester was purchased from Fragmenta (Monmouth Junction, NJ). Mini-PROTEAN TGX 8-16%, and Bio-Rad Protein Assay Dye Reagent concentrate was purchased from Bio-Rad Laboratories (Hercules, CA).

Restriction enzymes, T4 DNA ligase, and Phusion DNA polymerase were purchased from New England Biolabs (Ipswich, MA). GoTaq Green Master Mix was purchased from Promega (Madison, WI). 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 water. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). PD-10 columns Sephadex G-25M was purchased from GE (Chicago, IL). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Amico Ultra 3,000 MWCO, 10,000 MWCO, 30,000 MWCO, and 100,000 MWCO centrifugal concentrators were purchased from EMD Millipore (Billerica, MA). Chloroform-d was purchased from Cambridge Isotope Laboratories (Andover, MA). ^{1}H-NMR spectra were collected at 25 ºC on Bruker AVB-400 or AVQ-400 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 Quadrupole-Time-of-Flight (QTOF) Accurate Mass spectrometer and a 6460 Triple Quadrupole (QQQ) purchased from Agilent Technologies.

Bacterial strains. *E. coli* DH10B-T1^{R} was used for DNA construction. *E. coli* BL21(DE3)-T1^{R}, BAP1 [23], and BAP1-T1^{R} [24] were used for heterologous protein expression. BAP1 and BAP1-T1^{R} were used for expression of DEBS modules and DEBS ACP domains as they require modification with phosphopantethiene. BL21(DE3)-T1^{R} was used for all other proteins.

Gene and plasmid construction. Standard molecular biology techniques were used to carry out plasmid construction. All PCR amplifications were carried out with Phusion High Fidelity DNA polymerase or with Taq DNA polymerase using GoTaqGreen master mix. For amplification of GC-rich sequences, PCR reactions carried out with Phusion High Fidelity DNA polymerase were performed in the GC buffer supplemented with 10 % (*v/v*) DMSO and 1 M betaine with primer annealing temperatures 4-8 °C below the T_m. DNA assembly was performed using the isothermal Gibson assembly protocol [25]. For analysis and isolation of large DNA fragments from agarose gel, 0.6 % gel and zymoclean large fragment DNA recovery kit (Zymoresearch) were used. All constructs were verified by sequencing (UC Berkeley DNA Sequencing Facility, Berkeley, CA; and MGH CCIB DNA Core, Center for Computational and Integrative Biology, Massachusetts General Hospital, Cambridge, MA). pFW3 [26], pBL12 , pBL13, pBL36, pFW98, and pFW100 [27] were gifted by the laboratory of Professor Chaitan Khosla (Stanford University).

pET16b-His$_{10}$Pres-ACP$_{DEBSMod6}$. The ACP domain of Mod6$_{DEBS}$ was amplified from pAYC-138 [26] using primers PresACP6_Fwd3 and PresACP6_Rev. The PCR product was assembled into the NdeI-BamHI site of a modified pET16b with the Factor Xa cleavage site replaced by the PreScission cleavage site via Gibson assembly.

pFW3_F190V. pFW3_F190V was constructed by amplification from pFW3 with two pairs of primers: DszAT F190 F1/DszAT F190V R1 and DszAT F190V F2/DszAT F190 R2. Primers DszAT F190V R1 and DszAT F190V F2 contain the F190V mutation to be introduced to the plasmid. The two PCR products were inserted into the XbaI-HindIII site of pFW3 via Gibson assembly.

pFW98_DEBS2(Mod3 AT^0). pFW98_DEBS2(Mod3 AT^0) encodes for DEBS2 with S653A mutation. pBP130_DEBS2(Mod3 AT^0) (*Section 4.2*) and pFW98 were digested with MauBI and Pfl23II. The 2 kB band from pBP130_DEBS2(Mod3 AT$^{0)}$ and the 14 kB band from pFW98 were gel-purified and ligated with T4 ligase. The resulting plasmids were screened via restriction mapping with KpnI. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing at MGH CCIB DNA Core, Center for Computational and Integrative Biology, Massachusetts General Hospital (Cambridge, MA).

pFW100_DEBS3(Mod5 AT^0). pFW100_DEBS3(Mod5 AT^0) encodes for DEBS3 with S642A mutation. The plasmid was cloned by first replacing the section to be mutated from the parent plasmid with a Cm^R marker and then removing it upon insertion of the piece bearing the desired mutation. This strategy allows for identification of mutants using antibiotic selection. Briefly, the Cm^R cassette was first amplified from pACYC184 with the primers pACYC184_CmOperon_pFW100_M5_Fwd_PacI/ pACYC184_CmOperon_pFW100_M5_Rev_SpeI and inserted into the SfiI-BsiWI site of pFW100 via Gibson assembly. *E. coli* cells transformed with the assembly mixture were selected on culture medium containing carbenicillin and chloramphenicol. Plasmids containing both resistance markers were further screened by restriction analysis with SacI. The pFW100 plasmid intermediates with the correct restriction pattern were confirmed by diagnostic sequencing. The target segments were amplified from pFW100 by two primer pairs to introduce the AT^0 mutation into module 5 of DEBS: pFW100_M5AT0_SfiI_F/pFW100_M5AT0_SfiI_R and pFW100_M5AT0_BsiWI_F/pFW100_ M5AT0_BsiWI_R. The PCR products were then inserted via Gibson assembly into the pFW100 intermediate, which was digested with PacI and SpeI to remove the Cm^R cassette. Resulting plasmids were selected for the loss of the Cm^R marker and screened by SacI restriction analysis. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing at MGH CCIB DNA Core, Center for Computational and Integrative Biology, Massachusetts General Hospital (Cambridge, MA)..

pFW100_DEBS3(Mod6 AT^0). pFW100_DEBS3(Mod6 AT^0) encodes for DEBS3 with S2107A mutation. The plasmid was constructed using a similar strategy. The Cm^R cassette was amplified from pACYC184 with primers pACYC184_CmOperon_pBP130_M6_Fwd_PacI/ pACYC184_CmOperon_pBP130_M6_Rev_SpeI and inserted into the BbvCI-AjuI site of pFW100 via Gibson assembly. *E. coli* transformed with the assembly mixture were selected on culture medium containing carbenicillin and chloramphenicol. Plasmids containing both resistant genes were further screened by restriction analysis with SacI. The pFW100 plasmid intermediates with the correct restriction pattern were confirmed by diagnostic sequencing. Segments of module 6 of DEBS were amplified by two primers pairs to introduce the AT^0 mutation into DEBS$_{Mod6}$. The M6TE-SA-M6-RP/M6TE-SA-M6-FP primer pair was used to amplify a segment of module 6 containing the inactivating mutation to the AT domain from pAYC138, whereas the M6TE-SA-130-FP/M6TE-SA-130-RP primer pair was used to amplify another contiguous segment from pBP130 [23]. These PCR products were inserted via Gibson assembly into the pFW100 intermediate, which was digested with PacI and SpeI to remove the Cm^R cassette. Resulting plasmids were selected for the loss of the Cm^R marker and screened by SacI restriction analysis. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing at

MGH CCIB DNA Core, Center for Computational and Integrative Biology, Massachusetts General Hospital (Cambridge, MA).

Expression of His-tagged proteins. For His_6-MatB, DszAT-His_6, DszAT F190V-His_6, His_{10}-$ACP_{DEBSMod6}$, $Mod3_{DEBS}$+TE(AT^0)-His_6, $Mod6_{DEBS}$+TE(AT^0)-His_6, plasmids encoding the proteins of interest were transformed into *E. coli* BL21(DE3)-$T1^R$ (His_6-MatB, DszAT-His_6, DszAT F190V-His_6) or *E. coli* BAP1-$T1^R$ for proteins that require phosphopantetheine modification of ACP domains (His_{10}-$ACP_{DEBSMod6}$, $Mod3_{DEBS}$+TE(AT^0)-His_6, $Mod6_{DEBS}$+TE(AT^0)-His_6). TB culture medium (1 L) with appropriate antibiotic (carbenicillin, kanamycin, chloramphenicol: 50 µg/mL; spectinomycin: 100 µg/mL) in a 2.5 L ultra-yield flask was inoculated with overnight culture of freshly transformed *E. coli* cells. Cells were grown at 37 ºC with shaking at 200 rpm to OD_{600} = 0.6-0.8, at which time, they were cold-shocked on ice for 20-40 min. Expression was induced by addition of IPTG (His_6-MatB, His_{10}-$ACP_{DEBSMod6}$, DszAT-His_6, DszAT F190V-His_6: 1 mM; $Mod3_{DEBS}$+TE(AT^0)-His_6, $Mod6_{DEBS}$+TE(AT^0)-His_6: 0.2 mM). Cells were then grown at 16 ºC with shaking at 200 rpm overnight and harvested by centrifugation at 8,000 × *g* for 5 min at 4 ºC. Cell pellets were flash-frozen in liquid N_2 and stored at -80 ºC until purification.

For His_6-PrpE, His_6-Epi, His_6-LDD_{DEBS}, $Mod1_{DEBS}$-His_6, $Mod2_{DEBS}$-His_6, DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6, DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6, and DEBS3(Mod6 AT^0)-His_6, expression plasmids were introduced into *E. coli* BL21(DE3) (His_6-PrpE and His_6-Epi) or *E. coli* BAP1 cells to allow phosphopantetheinyl modification of ACP domains (His_6-LDD_{DEBS}, $Mod1_{DEBS}$-His_6, $Mod2_{DEBS}$-His_6, DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6, DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6, and DEBS3(Mod6 AT^0)-His_6). Overnight seed cultures were used to inoculate 0.5 L of LB medium with 2% glucose containing the appropriate antibiotic (carbenicillin, kanamycin: 50 µg/mL) in 2.5 L ultra-yield shake flask. Cells were grown at 37 ºC with shaking at 200 rpm to an approximate OD_{600} of 0.3, at which point the temperature was slowly lowered. When cells reached OD_{600} of 0.6, at which point cells were at approximately 20 ºC, protein expression was induced with IPTG (His_6-PrpE, His_6-Epi, His_6-LDD_{DEBS}, $Mod1_{DEBS}$-His_6, $Mod2_{DEBS}$-His_6, DEBS2: 0.1 mM; DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6, DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6, and DEBS3(Mod6 AT^0)-His_6: 0.5 mM). Cells were then grown at 18 ºC with shaking at 200 rpm overnight and harvested by centrifugation at 8,000 × *g* for 5 min at 4 ºC. Cell pellets were flash-frozen in liquid N_2 and stored at -80 ºC until purification.

Purification of His-tagged proteins. Cell pellets were resuspended in Lysis Buffer A (200 mM sodium phosphate, 200 mM sodium chloride, 30% (*v/v*) glycerol, 2.5 mM DTT, pH 7.5) at 5 mL/g of cell pellet. Lysozyme (0.5 mg/mL) and cOmplete EDTA-free protease inhibitor cocktail (Roche, 1 eq) were then added before homogenization and sonication. Cleared cell lysates were then obtained after centrifugation at 8,000 × *g* for 20 min at 4 ºC. To precipitate DNA, poly(ethyleneimine) was added dropwise to the soluble fraction at a final concentration of 0.015% (*w/v*). Precipitated DNA was removed by centrifugation at 8,000 × *g* for 20 min at 4 ºC. Protein concentration was determined using A_{280} and protein extinction coefficients calculated from the sequence using the ExPASY ProtParam program [28].

His_6-MatB, DszAT-His_6, and DszAT F190V-His_6 [2]. The lysate was diluted 3-fold with Wash Buffer B (50 mM sodium phosphate, 300 mM sodium chloride, 20% (*v/v*) glycerol, 20 mM BME, pH 7.5) with 10 mM imidazole and incubated with Ni-NTA agarose resin (2-4 mL) for 45-60 min at 4 ºC to bind the His-tagged protein before loading onto the column by gravity flow. The column was washed with Wash Buffer B containing 10 mM imidazole until the eluate was negative for

protein content when tested by Bradford protein assay reagent (Bio-Rad). Column was further washed by wash buffer B containing 25mM imidazole until eluate was negative for protein content again. Protein was eluted from the column with Elution Buffer C (50 mM sodium phosphate, 50 mM sodium chloride, 20 mM BME, 20% (*v/v*) glycerol) containing 250 mM imidazole. Eluted protein was concentrated using a 10 kD MWCO Amicon Ultra spin concentrator (Millipore) before exchanging into Storage Buffer D (50 mM HEPES, 2.5 mM EDTA, 100 mM sodium chloride, 2.5mM DTT, 20% (*v/v*) glycerol) using a PD-10 desalting column containing Sephadex G-25 resin (GE Healthcare). Purified protein was flash-frozen in liquid N_2 before storing at -80 ºC.

His_6-$ACP_{DEBSMod6}$. The protein was purified as described above up to protein elution. After elution from the Ni-NTA column, the concentration of the protein solution was determined using Bradford protein assay reagent (Bio-Rad) against BSA as a standard. PreScission protease (Strain #1785) was then added to the protein solution at a ratio of 1 mg PreScission:25 mg protein substrate. The proteolysis reaction was dialyzed 2× 1:200 overnight at 4 ºC into Storage Buffer D without DTT. The mixture was then diluted 3-fold with Wash Buffer B without BME containing 10 mM imidazole. The solution was passed three times through Ni-NTA resin (2 mL) to remove PreScission protease and the cleaved His_6 tag before concentrating the eluate using a 3 kDa MWCO Amicon Ultra spin concentrator (Millipore). The purified protein was exchanged into Storage Buffer D without DTT using PD-10 desalting column with Sephadex G-25 resin (GE Healthcare). Thiol-containing reducing agents such as BME and DTT were excluded from buffers starting at the dialysis step as they interfere with transacylation assay. Purified protein was flash-frozen in liquid N_2 before storing at -80 ºC. Presence of the phosphopantethiene modification was checked by LC-TOF using positive ionization mode on an Agilent 6224 TOF MS. Samples were buffer exchanged into 10 mM sodium phosphate by 0.5 mL 7K Zeba Spin desalting column and analyzed on Proswift RP-4H HPLC column (1mm × 50 mm, room temperature) using a linear gradient from 10 to 100% acetonitrile over 3.5 min with 0.1% (*v/v*) formic acid as the aqueous mobile phase at flow rate of 0.3 mL/min. Resulting mass chromatograms were deconvoluted with Agilent MassHunter BioConfirm software.

$Mod3_{DEBS}$+$TE(AT^0)$-His_6, $Mod6_{DEBS}$+$TE(AT^0)$-His_6 [2]. The cleared lysate was incubated with Ni-NTA resin (2-4 mL) for 45-60 min at 4 ºC to bind His-tagged protein before loading onto the column by gravity flow. The column was washed with Wash Buffer B containing 10 mM imidazole until eluate was negative for protein content when tested by Bradford protein assay reagent (Bio-Rad). Protein was then eluted from the column with Elution Buffer C containing 100 mM imidazole until eluate was negative for protein content. Eluted protein was concentrated using a 30 kD MWCO Amicon Ultra spin concentrator and diluted 3-fold in Anion Exchange Buffer E (50 mM HEPES, 2.5 mM EDTA, 2.5 mM DTT, 20% (*v/v*) glycerol, pH 7.5) without sodium chloride. The diluted protein solution was loaded onto a 5 mL Hi-Trap Q HP column (GE Healthcare) using an NGC Medium-Pressure Liquid Chromatography System (Bio-Rad). The protein was eluted with a linear gradient from 0 to 1 M sodium chloride in Anion Exchange Buffer E over 30 col vol (150 mL). Fractions containing the target protein eluted ~350 mM sodium chloride were pooled and concentrated with a 30 kD MWCO Amicon Ultra spin concentrator. Purified protein was flash-frozen in liquid N_2 before storing at -80 ºC.

DEBS2-His_6, DEBS2(Mod3 AT^0)-His_6, DEBS3-His_6, DEBS3(Mod5 AT^0)-His_6, and DEBS3(Mod6 AT^0)-His_6 [27]. Cell pellets were resuspended in Lysis Buffer F (50 mM sodium phosphate, 10 mM imidazole, 450 mM NaCl, 20% (*v/v*) glycerol, pH 7.6) at 5 mL/g of cell pellet. Lysozyme (0.5 mg/mL) and cOmplete EDTA-free protease inhibitor cocktail (Roche, 1 eq) were

then added and incubated at 30 ºC for 30 min before sonication. Cleared cell lysates were obtained after centrifugation at 8,000 × *g* for 20 min at 4 ºC and loaded onto column containing Ni-NTA (10 mL/ 3 L cell culture) by gravity flow. The column was washed with Lysis Buffer F until eluate was negative for protein content when tested by Bradford protein assay reagent (Bio-Rad) The column was further washed with Wash Buffer G (50 mM sodium phosphate, 25 mM imidazole, 300 mM sodium chloride, 10% (*v/v*) glycerol, pH 7.5) before protein was eluted with Elution Buffer H (75 mM sodium phosphate, 500 mM imidazole, 20 mM sodium chloride, 10% (*v/v*) glycerol, pH 7.5) until eluate was negative for protein content. Eluted protein was loaded onto a 5 mL Hi-Trap Q HP column connected to an AKTA FPLC system. A gradient of 0 to 1 M sodium chloride was applied over 30 col vol (150 mL) of Anion Exchange Buffer I (50 mM sodium phosphate, 10% (*v/v*) glycerol) at flow rate of 5 mL/ min. Fractions (2.5 mL) were collected and analyzed by SDS-PAGE. Fractions containing the target proteins (*DEBS2-His$_6$, DEBS2(Mod3 AT^0)-His$_6$* eluting at ~330-430 mM sodium chloride; *DEBS3-His$_6$, DEBS3(Mod5 AT^0)-His$_6$, and DEBS3(Mod6 AT^0)-His$_6$* eluting at ~370-500 mM sodium chloride) were pooled and concentrated using a 100 kDa MWCO Amicon Ultra spin concentrator. Purified protein was flash-frozen in liquid N_2 before storing at -80 ºC.

His$_6$-LDD$_{DEBS}$, Mod1$_{DEBS}$-His$_6$, Mod2$_{DEBS}$-His$_6$, His$_6$-PrpE, and His$_6$-Epi [27]. Cell pellets were resuspended in 35 mL Lysis Buffer J (50 mM sodium phosphate, 10 mM imidazole, 450 mM NaCl, 10% (*v/v*) glycerol, pH 7.6) containing 1 EDTA-Free Protease Inhibitor tablet (Pierce) per L of cell culture. The cells were lysed by sonication and centrifuged at 20,000 × *g* for 60 min at 4 ºC. The cleared cell lysates were added to Ni-NTA agarose resin (1 mL/L cell culture) and incubated at 4 °C for 1 h to bind His-tagged protein. The resin was loaded into a Kimble-Kontes Flex column and washed with 20 col vol of Lysis Buffer J and 10 col vol of Wash Buffer G. Proteins were eluted with 8 col vol of Elution Buffer H. The eluate was further purified by anion-exchange on a Hi-Trap Q column using a gradient from 0 to 500 mM sodium chloride in Anion Exchange Buffer I on an AKTA FPLC system. Fractions were collected and analyzed by SDS-PAGE. Fractions containing the target protein were pooled and concentrated using a 100 kD MWCO Amicon Ultra spin concentrator. Purified protein was flash-frozen in liquid N_2 before storing at -80 ºC.

Synthesis of N-acetylcysteamine thioester of (2S,3R)-2-methyl-3-hydroxypentanic acid (NDK-SNAC) [2]. Ethyl-2-methyl-3-oxopentanoate (1.5 mmol) was dissolved in 10 mL of water and saponified in aqueous sodium hydroxide (10 M, 180 µL; 1.8 mmol). The reaction was stirred overnight at room temperature before pH was adjusted to <1 and reaction was extracted with 2 × 35 mL of ethyl acetate. The organic layers were combined, dried over anhydrous magnesium sulfate , and evaporated to dryness by rotary evaporation.

The resulting 2-methyl-3-oxopentanoic acid (0.85 mmol) was dissolved in 10 mL of dichloromethane on ice. After 3 mins, N-acetylcysteamine (86 µL, 0.8 mmol, 0.95 eq) and DMAP (8 pellets) were added to the reaction. After 5 mins, ethyl diisopropyl carbodiimide (244 mg, 1.275 mmol; 1.5 eq) was added. The reaction was removed from ice and stirred at room temperature overnight. The reaction was extracted with 2 × 20 mL saturated aqueous ammonium chloride and 2 × 20 mL and saturated aqueous sodium chloride. The aqueous layers were combined and back extracted with 3 × 20 mL dichloromethane. All organic fractions were combined, dried over anhydrous magnesium sulfate, and evaporated to dryness by rotary evaporation. The resulting 2-methyl-3-oxopentanoyl-SNAC product was confirmed by ^{1}H-NMR and stored in solid form at -80 ºC. ^{1}H NMR (400 MHz, $CDCl_3$=7.26 ppm) δ 5.79 (s, N-H), 3.77-3.82 (m, 1H), 3.42-3.49 (m,

2H), 3.03-3.12 (m, 2H), 2.50-2.62 (m, 2H), 1.97 (s, 3H), 1.20 (d, J = 6.8 Hz, 3H,), 1.07 (t, J = 7.2 Hz, 3H).

2-methyl-3-oxopentanoyl-SNAC was enzymatically reduced with the ketoreductase domain from the first module of the pikromycin synthase ($KR_{PIKSMod1}$) and NADPH. This reaction (3 mL) contained 50 mM 2-methyl-3-oxopentanoyl-SNAC (1 M stock in DMSO), 300 mM D-glucose, 100 µM $NADP^+$, 20 U/mL glucose-1-dehydrogenase, and 15 µM $Piks_{Mod1}KR$ in 300 mM HEPES, 100 mM sodium chloride, 10% (*v/v*) glycerol, pH 7.5. After stirring at room temperature overnight, thin-layer chromatography analysis of crude reaction mixture in 100% ethyl acetate solvent system revealed a product spot at $R_f \sim 0.3$. Saturated aqueous sodium chloride (3 mL) was then added to the reaction mixture, which was extracted with 3 × 9 mL of ethyl acetate. The organic layer was dried over anhydrous magnesium sulfate and evaporated to dryness by rotary evaporation. Product was confirmed by ^{1}H-NMR and stored as 1 M solution in 25% (*v/v*) DMSO at -20 ºC. ^{1}H NMR (400 MHz, $CDCl_3$=7.26 ppm) δ 3.81-3.84 (m, 1H), 3.43-3.47 (m, 2H), 2.99-3.05 (m, 2H), 2.69-2.75 (m, 1H), 2.62 (s, O-H), 1.97 (s, 3H), 1.43-1.57 (m, 2H), 1.23 (d, J = 7.2 Hz, 3H,), 0.98 (t, J = 7.2 Hz, 3H).

Synthesis of fluoromalonate [2]. Methanolic sodium hydroxide (2 M, 10.5 mL; 21 mmol) was added to 9:1 (*v/v*) dichloromethane:methanol (96 mL). Diethylfluoromalonate (1.5 mL, 9.6 mmol) was added and saponified with stirring at room temperature for 1-3 h. Disodium fluoromalonate was isolated by filtration through a Büchner funnel fitted with a filter paper and left to air-dry.

Synthesis of fluoromalonyl-CoA [2]. Fluoromalonyl-CoA was prepared enzymatically from fluoromalonate and coenzyme A using the malonyl-CoA synthetase from *Streptomyces ceolicolor* (MatB). The reaction mixture (5 mL) contained 20 mM fluoromalonate, 4 mM coenzyme A, 5 mM ATP, 20 mM phosphoenol pyruvate, 5 mM TCEP, 10 mM magnesium chloride, 144 U/mL pyruvate kinase, 80 U/mL myokinase, and 20 µM MatB in 200 mM sodium phosphate, pH 7.5. The reaction was incubated at 37 ºC for 1 h. Enzymes were removed by filtration through Nanosep with 10K Omega membrane spin column before reaction was lyophilized overnight. The residue from lyophilization was resuspended in water (2 mL) and stored frozen at -80 ºC until purification. The resuspended crude reaction residue (0.5 mL per injection) was purified on an Eclipse XDB-C18 column (5 µm, 9.4 × 250 mm, room temperature; Agilent) using a linear gradient from 0 to 10% acetonitrile over 30 min with 10 mM ammonium formate as the aqueous mobile phase at flow rate of 4 mL/min. Fractions (2 mL) were screened using an Agilent 1290 UPLC connected to a 6460 Triple Quadrupole (QQQ) mass spectrometer using multiple reaction monitoring (MRM) in positive ionization mode (transition, collision energy, fragmentation voltage: fluoromalonyl-coA (872 → 365, 135, 35)). Fractions were analyzed on a Poroshell HPH-C18 column (2.7 µm, 2.1 × 100 mm, room temperature; Agilent) with a linear gradient from 0 to 20 % acetonitrile over 6 min after an initial hold at 0 % acetonitrile for 30 s with 10 mM ammonium formate as the aqueous mobile phase at flow rate of 0.6 mL/min. Fractions containing pure fluoromalonyl-CoA were pooled and lyophilized overnight. The residue was re-dissolved in water and stored at -80 ºC. The concentration of fluoromalonyl-CoA was estimated using A_{260} (ε_{260} = 15,400 M^{-1} cm^{-1}).

Steady-state kinetic characterization of DszAT hydrolysis activity. The DszAT assay was adapted from a coenzyme A (CoA) release assay previously described in the literature [29-31]. Assays were performed in 96-well microplate (black, polystyrene, µ-clear, f-bottom, chimney well, med binding; Greiner Bio-One) by monitoring free CoA with DTNB at 412 nm using a

Synergy Mx microplate reader (Biotek). Reactions were monitored for 4 min using the minimum interval setting between measurements.

Assay components were prepared in three solutions in 50 mM sodium phosphate, 1 mM EDTA, 10% (*v/v*) glycerol, pH 7.5: Solution 1, 2.4 mM DTNB; Solution 2, fluoromalonyl-CoA at 4 × final concentration ; Solution 3, 20 µM *trans*-AT enzyme and 0.1 mg/mL BSA. Solution 1 (25 µL) and Solution 2 (25 µL) were added to the wells and pre-incubated at room temperature for 15-20 min to reduce background by allowing any free CoA contaminating the acyl-CoA substrate to be consumed. The reactions were then initiated with Solution 3 (50 µL). Final assay mixtures (100 µL) contained 10 µM DszAT (WT and F190V) and 0.6 mM DTNB in 50 mM sodium phosphate, 1 mM EDTA, 10% (*v/v*) glycerol, pH 7.5. The final concentrations of fluoromalonyl-CoA were 200, 150, 100, 50, 25, 12.5, 6.25, 3.13, 1.56, 0.78, 0.39, and 0.20 µM.

Absorbance values were converted to concentrations using the $\varepsilon_{412\ nm}$ = 14,150 $M^{-1}cm^{-1}$ and a pathlength of 0.3 cm. Initial velocities (v_0) were defined as the linear portion of the change of product concentration over time curve. The data were fit by non-linear curve fitting to the standard Michaelis-Menten equation using Origin 6.0 with initial values estimated by the Solver tool in Microsoft Excel:

$$v_0 = \frac{v_{max}[S]}{K_M + [S]}$$

where v_0 = initial velocity ; v_{max} = maximum velocity, [S] = substrate concentration, and K_M = Michaelis-Menten constant.

Steady-state kinetic characterization of DszAT transacylation activity.. The transacylase assay was adapted from a previously described method [7, 12, 32]. Assays were performed in 96-well microplate (half area, µCLEAR, black polystyrene, medium binding non-sterile, Greiner Bio-One) with NADH fluorescence (λ_{Ex} = 340 nm, λ_{Em} = 450 nm) monitored using a Synergy Mx microplate reader (Biotek; top optical position at 8 mm read height, gain = 100). Reactions were monitored for 5 mins at 30 ºC using minimum interval setting between measurements.

Assay components were prepared in three solutions in 50 mM sodium phosphate, 1 mM EDTA, 10% (*v/v*) glycerol, pH 7.5: Solution 1, 2.4 mU/µL α-ketoglutarate dehydrogenase, 1.6 mM NAD^+, 1.6 mM TPP, and 8 mM α-ketoglutarate; Solution 2, malonyl-CoA, methylmalonyl-CoA, or fluoromalonyl-CoA substrate at 4 × final concentration; Solution 3, 150 µM $ACP_{DEBSMod6}$, trans-AT enzyme (100 nM for malonyl-CoA and fluoromalonyl-CoA experiments; 20 µM for methylmalonyl-CoA), and 0.1 mg/mL BSA. Solution 1 (12.5 µL) and Solution 2 (12.5 µL) were added to the wells and pre-incubated at room temperature for 15-20 min to reduce background by allowing any free CoA contaminating the acyl-CoA substrate to be consumed. The reactions were then initiated with Solution 3 (25 µL).

Final assay mixtures (50 µL) contained 75 µM $ACP_{DEBSMod6}$, 0.6 mU/µL α-ketoglutarate dehydrogenase, 0.4 mM NAD^+, 0.4 mM TPP, 2 mM α-ketoglutarate, and 0.05 mg/mL BSA in 50 mM sodium phosphate, 1 mM EDTA, 10% (*v/v*) glycerol, pH 7.5. The final concentrations of malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA were 200, 150, 100, 50, 25, 12.5, 6.25, and 3.13 µM. The final concentrations of DszAT (WT and F190V) were 50 nM for malonyl-CoA and fluoromalonyl-CoA and 10 µM for methylmalonyl-CoA. Controls with no *trans*-AT added were run in parallel for all reactions.

Fluorescence was converted to concentration using an NADH standard curve collected externally in the same assay matrix with carboxyacyl-CoA substrate, ACP domain, and DszAT omitted. Initial velocities (v_0) were defined as the linear portion of the change of product concentration over time curve. Transacylation of methylmalonyl-CoA substrate data were fit by non-linear curve fitting to the standard Michaelis-Menten equation:

$$v_0 = \frac{v_{max}[S]}{K_M + [S]}$$

where v_0 = initial velocity ; v_{max} = maximum velocity, [S] = substrate concentration, and K_M = Michaelis-Menten constant. Transacylation of malonyl-CoA and fluoromalonyl-CoA substrates data were fit by non-linear curve fitting to the Hill equation with substrate inhibition:

$$v_0 = \frac{v_{max}[B]_0^n}{K_{0.5}^n + [B]_0^n(1 + \frac{[B]_0^n}{K_i^n})}$$

where v_0 = initial velocity, v_{max} = maximum velocity, $[B]_0$ = substrate concentration, n = Hill coefficient, $K_{0.5}$ =half-maximal concentration constant, and K_i = inhibition constant. Fitting were performed using Origin 6.0 with initial values estimated by the Solver tool in Microsoft Excel:

Assay for *in vitro* triketide lactone production. All assays (30 µL) contained 5 µM methylmalonyl-CoA epimerase (Epi), 20 µM MatB, 2.5 mM ATP, 15 U/mL PK, 10 U/mL myokinase, 50 mM phosphoenol pyruvate, and 10 mM NDK-SNAC in 400 mM sodium phosphate, 10 mM magnesium chloride, 5 mM TCEP, pH 7.5. When used, fluoromalonyl-CoA, malonyl-CoA, or methylmalonyl-CoA were added to 1 mM final concentration. Final enzyme concentrations were 10 µM for $Mod3_{DEBS}$+TE(AT^0), $Mod6_{DEBS}$+TE(AT^0), and DszAT variants. The following assay components were pre-incubated at 37 ºC for 30 min before reaction initiation: Epi, MatB, PK, myokinase, PEP, ATP, and carboxyacyl-CoA substrate. The reactions were then initiated with NDK-SNAC, DEBS module, and *trans*-AT and incubated for 1 h at 37 ºC. To quench the reaction, 25 µL of reaction mixture was mixed with 25 µL of 10% (*v/v*) perchloric acid and centrifuged for 10 min at 4 ºC at 20,817 × *g*. The supernatant was removed and stored at -80 ºC until analysis.

Quantification of triketide lactone products by LC-QQQ. Frozen samples were centrifuged for 10 min at 4 ºC at 20,817 × *g* to remove salts. The supernatant was removed and analyzed by LC-QQQ using MRM in positive ionization mode on an Agilent 6460 QQQ MS. Samples were analyzed on two tandem Ascentis Express RP-amide HPLC columns (2.7 µm, 2.1 mm × 10 cm, 25 ºC; Supelco) on an Agilent 1290 UPLC using an isocratic solvent system of 10% acetonitrile in 0.1% (*v/v*) formic acid as the aqueous mobile phase for 20 min after an initial ramp of 0 to 10% acetonitrile over 0.5 min at flow rate of 0.4 mL/min. Transitions for the TKLs and TTKPs products were determined by screening for optimal fragmentation conditions. Products were monitored with transitions as follows (parent ion *m/z* → product ion *m/z*, fragmentor, collision energy): TKL (171 → 153, 90,6); H-TKL (157 → 139, 90, 4); F-TKL (175 → 157, 90, 5); H-TTKP (199 → 111, 100, 9). The identity and the quantity of each compound was verified against synthetic standards using external standard curves.

Assay for in vitro 6dEB analog production. All assays (Mod6 AT^0 DEBS: 30 µL, Mod3 AT^0 DEBS and Mod5 AT^0 DEBS: 50 µL) contained 2 µM LDD_{DEBS}, 2 µM $Mod1_{DEBS}$, 2 µM $Mod2_{DEBS}$, 2 µM DEBS2 variant (WT or DEBS2(Mod3 AT^0)), and 2 µM DEBS3 variant (WT,

DEBS3(Mod5 AT^0), or DEBS3(Mod6 AT^0)), 2 μM MatB, 4 μM Epi, 2 μM PrpE, 1 mM CoA, 4 mM ATP, 2 mM methylmalonate, 0.5 mM NADPH, 0.5 mM propionate in 250 mM sodium phosphate, 10 mM magnesium chloride, 5 mM TCEP, pH 7.5. When used, 10 mM fluoromalonate, 2 mM malonate, and 6 μM DszAT variant (WT or F190V) were added. Reaction mixtures were incubated at room temperature for 3 h before quenching and extracting twice with 4 × reaction volume of ethyl acetate (Mod6 AT^0 DEBS: 120 μL, Mod3 AT^0 DEBS and Mod5 AT^0 DEBS: 200 μL). The combined organic layers were dried using a Speed-Vac SC110 (Savant). The residue was resuspended in methanol (Mod6 AT^0 DEBS: 30 μL, Mod3 AT^0 DEBS and Mod5 AT^0 DEBS: 25 μL). and centrifuged for 10 min at 4 ºC at 20,817 × *g*. The supernatant was collected and analyzed via LC-MS/MS.

Identification of 6dEB analogs by LC-QTOF. Products were analyzed by LC-QTOF high-resolution mass spectrometry in positive ionization mode on at Agilent 6530B QTOF MS. Samples were analyzed on Ascentis Express RP-amide HPLC columns (2.7 μm, 2.1 mm × 10 cm, 25 ºC; Supelco) on an Agilent 1290 UPLC using a linear gradient from 5 to 65% acetonitrile over 16.75 min with 0.1% (*v/v*) formic acid as the aqueous mobile phase after an initial increase of 0 to 5% acetonitrile over 30 s at flow rate of 0.25 mL/min. Fragmentation analysis was performed using the targeted MS/MS mode with fragmentation voltage of 185 and collision energy of 10. Masses observed corresponded to $[M+H^+-H_2O]^+$ of the target molecules as follows (molecular formula, expected mass, observed mass, Δppm(exp-obs)): 6dEB ($C_{21}H_{37}O_5^+$, 369.2636, 369.2639, 0.08 ppm), desmethyl 6dEB ($C_{20}H_{35}O_5^+$, 355.2479, 355.2476, 0.8 ppm), and fluorodesmethyl 6dEB ($C_{20}H_{34}FO_5^+$, 373.2385, 373.2384, 0.3 ppm).

From the fragmentation patterns, unique transitions were developed for detection with increased sensitivity with LC-QQQ using MRM in positive ionization mode on an Agilent 6460 QQQ MS with the same LC method. Products were monitored with transitions as follows (parent *m/z* → product *m/z*, fragmentor, collision energy): 6dEB (369.2→239.1, 50, 1), 2-desmethyl 6dEB and 4-desmethyl 6dEB (355→239.1, 50, 1), 8-desmethyl 6dEB (355.2→225.1, 50,1), 2-fluoro-2-desmethyl 6dEB and 4-fluoro-4-desmethyl 6dEB (373.2→275.1, 105, 0).

2.3 Results and Discussion

Hydrolysis activity of F190V DszAT. In order to begin characterizing the basis for the differential selectivity of our DszAT variants for fluoromalonyl-CoA and malonyl-CoA observed in end-product analysis, we first set out to carry out the steady-state kinetic characterization of the hydrolysis activity of DszAT. Hydrolysis is a side reaction of the desired transacylation activity that indicates that the enzyme can interact with the carboxyacyl-CoA, but also leads to its unproductive degradation if the reaction continues in the presence of the ACP substrate (*Figure 2.6A*). The hydrolytic activity of WT DszAT on fluoromalonyl-CoA had been previously observed [1]; however, the kinetic parameters describing this activity were not determined. Thus, we examined the hydrolytic activity of F190V DszAT compared to WT DszAT on a panel of the three different carboxyacyl-CoA substrates of interest: malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA. In order to assess the nature of their selectivity, wild-type and F190V DszAT were overexpressed and purified to near homogeneity (*Figure 2.6B*). Carboxyacyl-CoA hydrolysis was monitored as CoA release using DTNB to quantify free thiols in solution (*Figure 2.7A*).

We found that the hydrolytic activity of the wild-type and F190V mutant are similar. For both DszAT variants, the hydrolysis rates on malonyl-CoA and methylmalonyl-CoA were below the limit of detection of the assay. This result is consistent with the k_{cat}/K_M value of below 0.005 μM^{-1} min^{-1} previously reported for these substrates [12]. This result suggests that the F190V mutation does not introduce significant hydrolytic activity for either malonyl-CoA or methylmalonyl-CoA substrates. This observation is consistent with a model where the native substrate, malonyl-CoA, does not undergo unproductive hydrolysis whereas the excluded substrate, methylmalonyl-CoA, may not be an accepted substrate for transacylation. In contrast, we found that both wild-type and F190V DszAT demonstrates measurable hydrolytic activity on fluoromalonyl-CoA (*Figure 2.7B*). As fluoromalonyl-CoA is not naturally-occurring, this result is unlikely due to evolutionary selection pressure but rather the inherent chemical reactivity of fluoromalonyl-CoA where α-fluoro thioester are hydrolyzed more readily due to fluorine's ability to stabilize the tetrahedral intermediate and preference for sp^3 over sp^2 carbons. The F190V mutant was 3-fold less active than the wild-type enzyme arising from a 4.5-fold decrease in k_{cat} and a 2-fold decrease in the K_M.

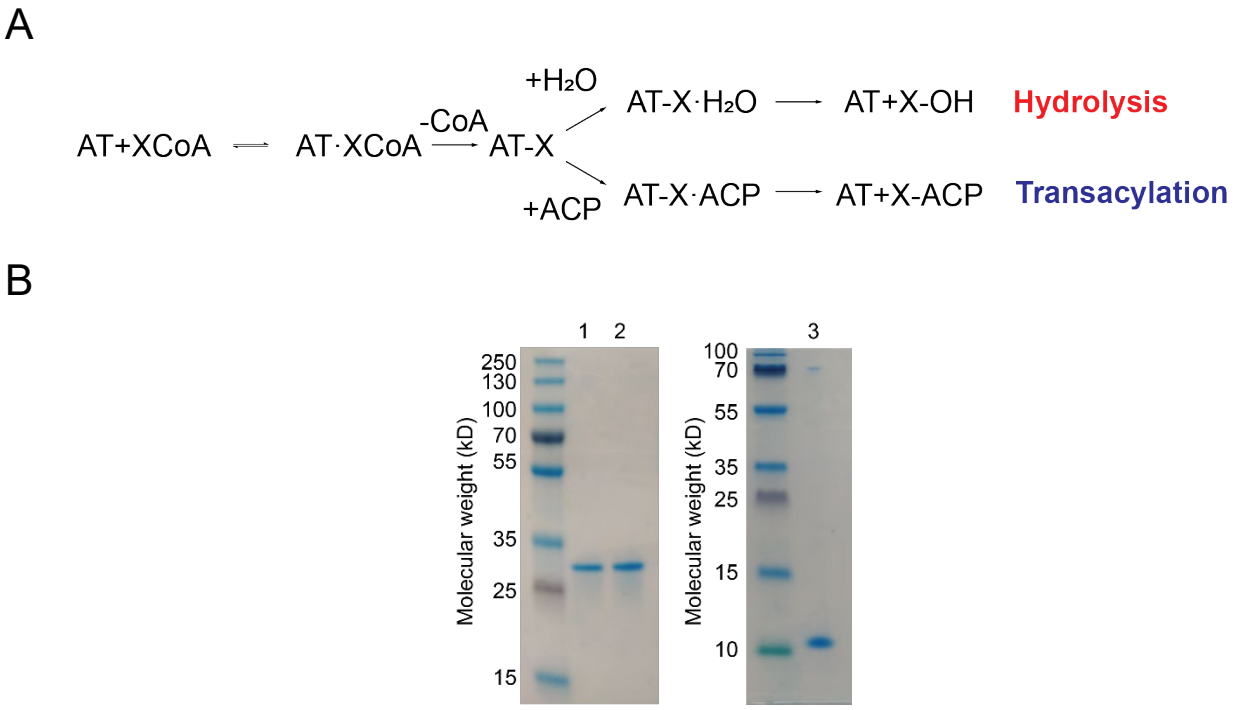
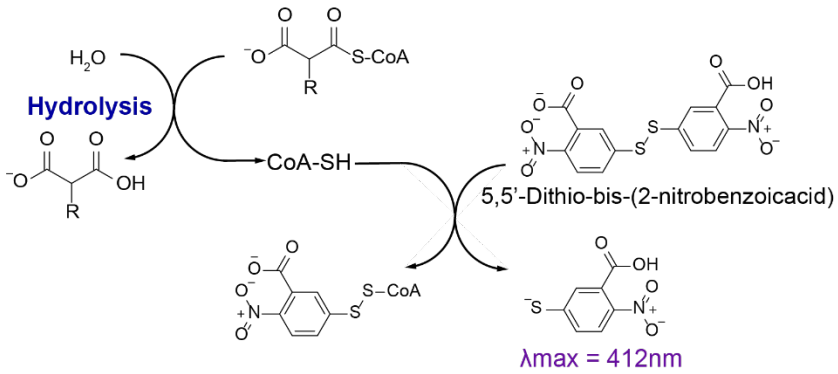
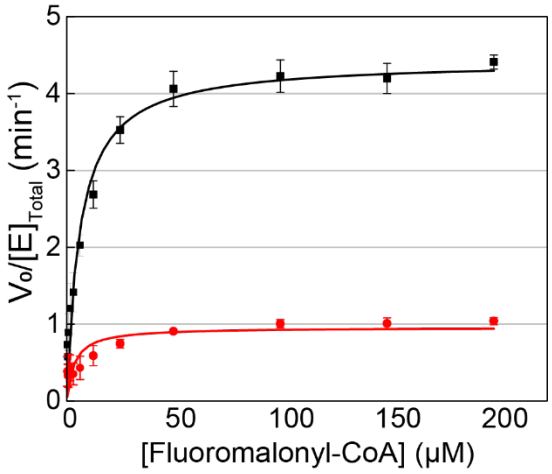


Figure 2.6. *Hydrolysis activity of DszAT variants. (A) Hydrolysis is a side reaction to the desired transacylation reaction performed by AT enzymes. In the absence of ACP partner, AT could interact with carboxyacyl-CoA substrate leading to its unproductive degradation. (B) SDS-PAGE of purified proteins used in steady-state kinetic characterization of DszAT variants. (lane 1) WT DszAT, (lane 2) F190V DszAT, and (lane 3) $ACP_{DEBSMod6}$.*

A



B



C

DszAT	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($min^{-1}μM^{-1}$)
Wild-type	4.4 ± 0.11	6.1 ± 0.70	0.72 ± 0.10
F190V	0.96 ± 0.061	3.5 ± 1.1	0.28 ± 0.11

Figure 2.7. *Steady-state kinetic characterization of fluoromalonyl-CoA hydrolysis by DszAT variants. (A) DTNB assay for measuring hydrolytic activity by free CoA release. (B) Michaelis-Menten curves for fluoromalonyl-coA hydrolysis by WT (black) and F190V DszAT (red). Data points are average ± s.d. of technical replicates (n=3). (C) Table contains k_{cat}, K_M, and k_{cat}/K_M calculated by non-linear curve fitting to the Michaelis-Menten equation. Error in k_{cat}/K_M is obtained by propagation from the individual kinetic terms. Data are mean ± s.d. of technical replicates (n=3).*

ACP transacylation activity of F190V DszAT. We focused next on characterizing the native ACP transacylation reaction performed by wild-type and F190V DszAT using malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA. Previous work examined the activity of DszAT on its cognate ACP partner, the ACP from module 1 of disorazole PKS, with various carboxyacyl-CoA substrates [12]. In these studies, DszAT was shown to be slightly more efficient with its

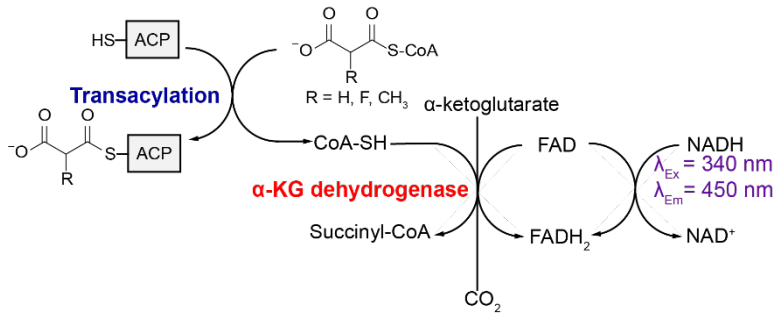
native ACP co-substrate but also accepted $ACP_{DEBSMod3}$ and $ACP_{DEBSMod6}$ with good catalytic efficiency.

As the phosphopantetheine group of holo-ACPs interferes with the DTNB assay, a α-ketoglutarate dehydrogenase-coupled transacylation assay [12] was used instead with $ACP_{DEBSMod6}$, which is more relevant to our studies (*Figure 2.8A*). The dose-response curves for transacylation of methylmalonyl-CoA by WT and F190V DszAT were fit with Michaelis-Menten equation. However, the dose-response curves for WT and F190V DszAT transacylation with malonyl- and fluoromalonyl-CoA exhibited sigmoidal behavior with evidence of inhibition at high substrate concentration and were therefore fit with the Hill equation modified for substrate inhibition. This observation suggests that transacylation does not follow the classic Michaelis-Menten model for enzyme catalysis, which is consistent with previous studies that have excluded the possibility of the ping-pong mechanism and noted that the transacylation reaction therefore likely undergoes a single displacement mechanism, based on the difference in k_{cat}/K_M values of hydrolysis and transacylation reactions (*Figure 2.9*) [12]. We similarly found that the k_{cat}/K_M values for the same enzyme-substrate pairs differed for hydrolysis and transacylation reactions.

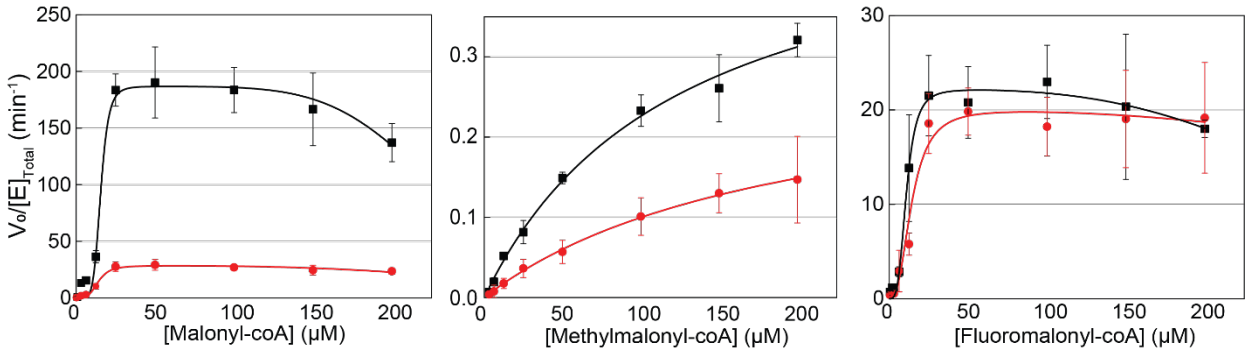
We found that WT DszAT displayed ~10^3-fold higher selectivity in k_{cat}/K_M with respect to malonyl-CoA compared to methylmalonyl-CoA (*Figure 2.8*). The observed discrimination against methylmalonyl-CoA is consistent with previous report that showed that the k_{cat}/K_M value differed by 5 orders of magnitude with the native ACP partner [12]. The difference in selectivity observed with $ACP_{DEBSMod6}$ appears to mainly arise from a decrease in k_{cat} with malonyl-CoA substrate when non-cognate ACP partner was used. In comparison, WT DszAT did not highly distinguish between malonyl-CoA and fluoromalonyl-CoA with only a 6-fold lower k_{cat}/K_M value for fluoromalonyl-CoA and comparable K_M values (*Figure 2.8BC*). This result supports our previous finding that WT DszAT is able to increase F-TKL yield in *in vitro* complementation experiments [2]. The similarity in the two observed K_M values further suggests that fluorine is sterically accommodated in the malonyl-CoA binding pocket.

When the F190V DszAT was compared to wild-type, we found that activity with fluoromalonyl-CoA is maintained with very similar values for K_M and k_{cat}/K_M, indicating that this mutation does not have much effect in this regard (*Figure 2.8BC*). However, activity measurably decreased with both malonyl-CoA and methylmalonyl-CoA, with a 6- and 2.5-fold respective decrease in k_{cat}/K_M. The decrease in k_{cat}/K_M with respect to malonyl-CoA mostly resulted from a decrease in k_{cat} value, while the decrease in k_{cat}/K_M with respect to methylmalonyl-CoA was derived from changes in both k_{cat} and K_M. This finding supports our hypothesis that the F190 residue does indeed play a key role in the substrate selectivity of DszAT. F190 packs against the α-carbon of malonyl-CoA to hold the substrate in place. Thus, replacing it with a smaller hydrophobic residue like valine could disproportionately affect its activity on malonyl-CoA. We reasoned that the activity on fluoromalonyl-CoA was only slightly affected in the F190V mutant as fluoromalonyl-CoA is activated by inductive effects and can more readily form the acyl-enzyme intermediate.

A



B



C

Substrate	DszAT	k_{cat} (min^{-1})	K (μM)	k_{cat}/K (min^{-1} μM^{-1})	K_i (μM)	n
Malonyl-CoA	WT	187 ± 8.5	16 ± 1.5	12 ± 1.7	233 ± 16	6.2 ± 2.3
	F190V	29 ± 1.5	14 ± 1.1	2.0 ± 0.27	282 ± 37	4.0 ± 1.3
Methylmalonyl-CoA	WT	0.51 ± 0.04	123 ± 21	0.0041 ± 0.0010	n/a	n/a
	F190V	0.29 ± 0.084	193 ± 95	0.0015 ± 0.0012	n/a	n/a
Fluoromalonyl-CoA	WT	22 ± 1.6	11 ± 1.3	2.1 ± 0.40	304 ± 73	3.5 ± 1.4
	F190V	20 ± 2.1	14 ± 2.6	1.4 ± 0.41	542 ± 446	2.6 ± 1.0

Figure 2.8. *Steady-state kinetic characterization of ACP transacylation catalyzed by DszAT variants with malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA. (A) αKGDH-coupled kinetic assay for measuring transacylation activity by NADH fluorescence. (B) Transacylation of $ACP_{DEBSMod6}$ with malonyl-CoA (left), methylmalonyl-coA (center), and fluoromalonyl-CoA (right) by WT (black) and F190V (red) DszAT. Malonyl-CoA and fluoromalonyl-CoA reactions data were fit to Hill equation with substrate inhibition model, whereas methylmalonyl-CoA reactions data were fit to Michaelis-Menten model of enzyme catalysis. Data points are average ± s.d. of technical replicates (n=3). (C) Table contains k_{cat}, K, k_{cat}/K, Ki, and n calculated by non-linear curve fitting to the Hill equation with substrate inhibition (malonyl-CoA and fluoromalonyl-CoA, $K = K_{0.5}$, $k_{cat}/K = k_{cat}/K_{0.5}$) or Michaelis-Menten equation (methylmalonyl-CoA, $K = K_M$, $k_{cat}/K = k_{cat}/K_M$). Error in k_{cat}/K is obtained by propagation from the individual kinetic terms. Data are mean ± s.e. of technical replicates (n=3).*

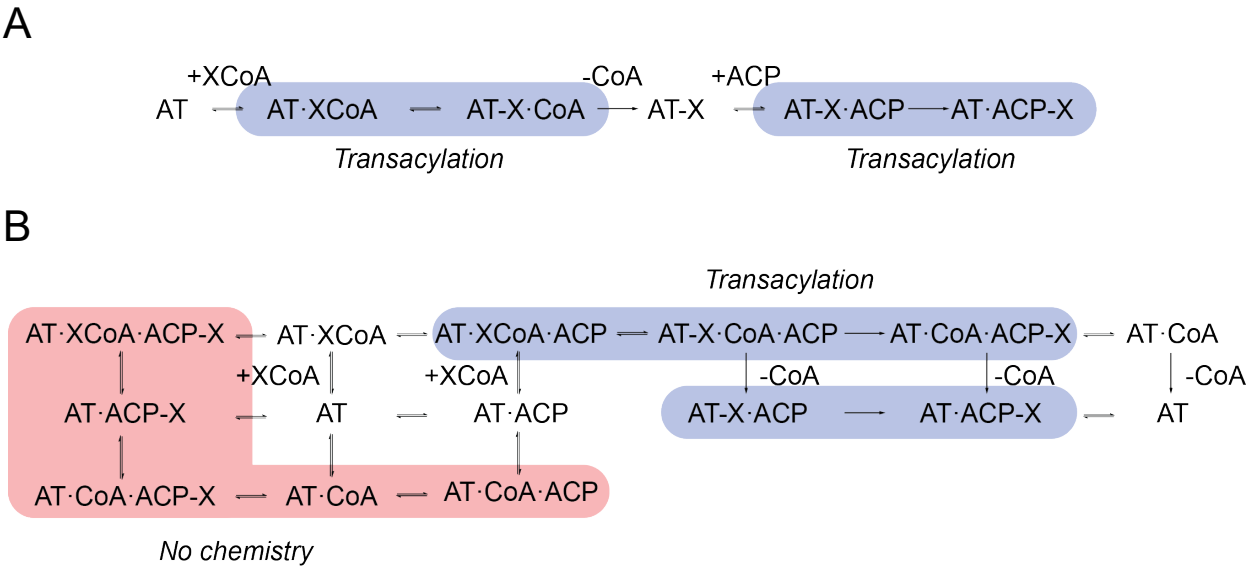


Figure 2.9 *Mechanism of ACP transacylation of DszAT. (A) Double displacement or ping-pong bi-bi reaction model for transacylation reaction. (B) Random order single displacement bi-bi reaction model for transacylation reaction. Substrate binding and product release are shown in all possible order combinations. Binding of products to free AT enzyme and AT enzyme complexes let to allosteric inhibition (red) to the transacylation reaction (blue).*

***In vitro* triketide lactone production with DszAT F190V.** With a more fluorine-selective variant of DszAT in hand, we sought to investigate the F190V mutant in triketide lactone production with $Mod3_{DEBS}$+TE(AT^0) and $Mod6_{DEBS}$+TE(AT^0) to assess how its differential activity on malonyl- vs fluoromalonyl-CoA could impact polyketide production. This assay utilizes MatB and methylmalonyl-CoA epimerase (Epi) in addition to the DEBS modules (*Figure 2.10*) and DszAT variants, along with an ATP regeneration system to ensure that the carboxyacyl-CoA extender can be maintained even should unproductive hydrolysis occur (*Figure 2.11*).

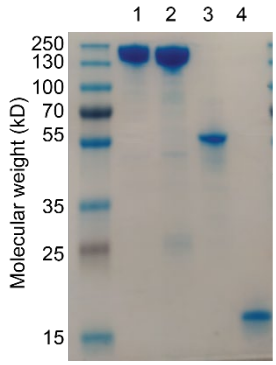


Figure 2.10. *SDS-PAGE of purified proteins used in the in vitro triketide lactone production assay. (lane 1) $Mod3_{DEBS}$+TE(AT^0), (lane 2) $Mod6_{DEBS}$+TE(AT^0), (lane 3) MatB, and (lane 4) methylmalonyl-coA epimerase (Epi).*

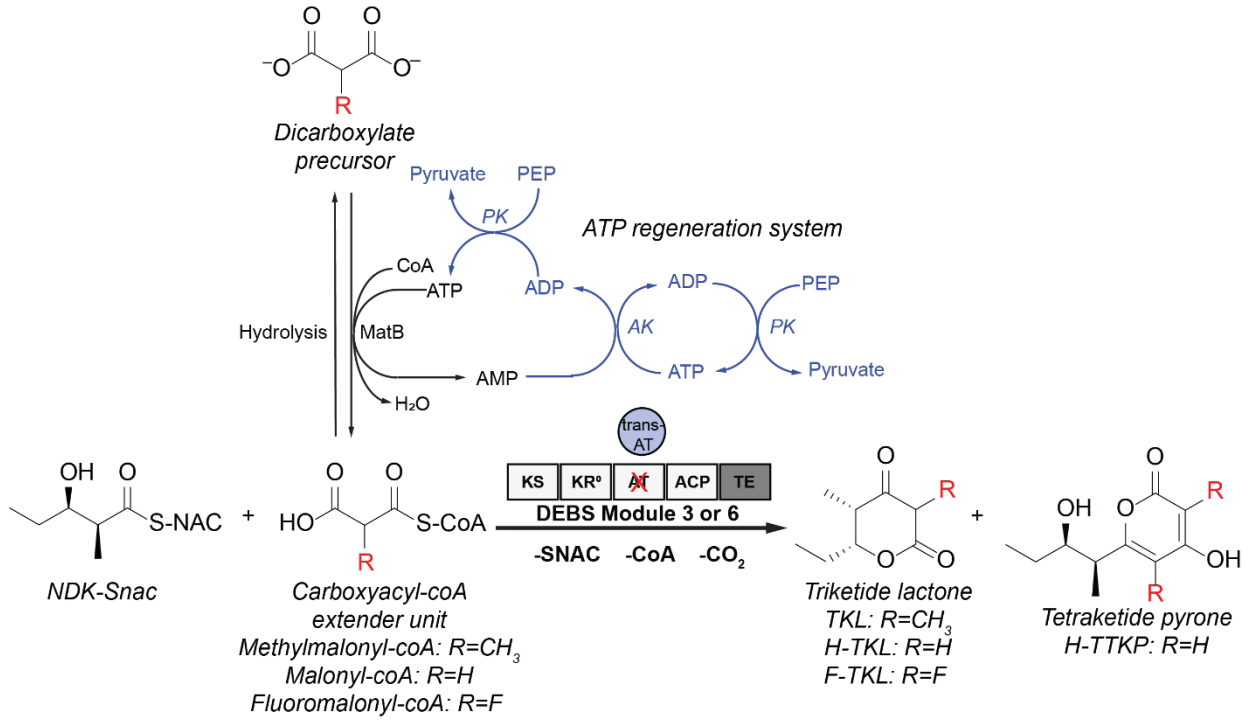


Figure 2.11. *The in vitro triketide lactone assay. Mod3$_{DEBS}$+TE(AT^0) or Mod6$_{DEBS}$+TE(AT^0) utilizes the N-acetylcysteamine thioester of (2S,3R)-2-methyl-3-hydroxypentanoic acid (Natural diketide-S-N-acetylcysteamine, NDK-SNAC) as a starter unit along with a carboxyacyl-CoA extender unit in a chain extension reaction to produce a triketide lactone. The reaction cycle is initiated by the acylation of the KS domain by NDK-SNAC. In AT^0 variants, the trans-AT (DszAT) selects the extender unit and transacylates the extender unit onto the phosphopantetheine arm of the ACP domain. Carbon-carbon bond formation catalyzed by the KS domain produces the triketide, which is cyclized by the TE domain to yield a triketide lactone. In some cases, the triketide intermediate can undergo another extension step, producing a tetraketide pyrone. The carboxyacyl-CoA extender unit can be provided directly either as the CoA thioester or is prepared in situ with the free acid and malonyl-CoA synthetase (MatB). In the regenerative system, the AMP produced by MatB is converted back to ATP through an enzyme cascade involving adenylate kinase (AK) and pyruvate kinase (PK), which derives energy from 2 molecules of phosphoenolpyruvate (PEP) to produce 1 molecule of ATP from AMP.*

When malonyl-CoA was provided as the sole extender unit to Mod3$_{DEBS}$+TE(AT^0), DszAT F190V was able to suppress the production of H-TKL by at least 90% compared to wild-type enzyme (*Figure 2.12*). This result is consistent with its decreased activity on the malonyl-CoA extender unit as shown by steady-state kinetic analysis. Moreover, DszAT F190V was able to moderately increase the amount of F-TKL produced in comparison to that produced by the wild-type enzyme (*Figure 2.12).* This result is also consistent with steady-state kinetic parameters that showed that the two DszAT variants maintained similar level of activity on the fluoromalonyl-CoA extender unit. Similar results were obtained for Mod6$_{DEBS}$+TE(AT^0) (*Figure 2.12*).

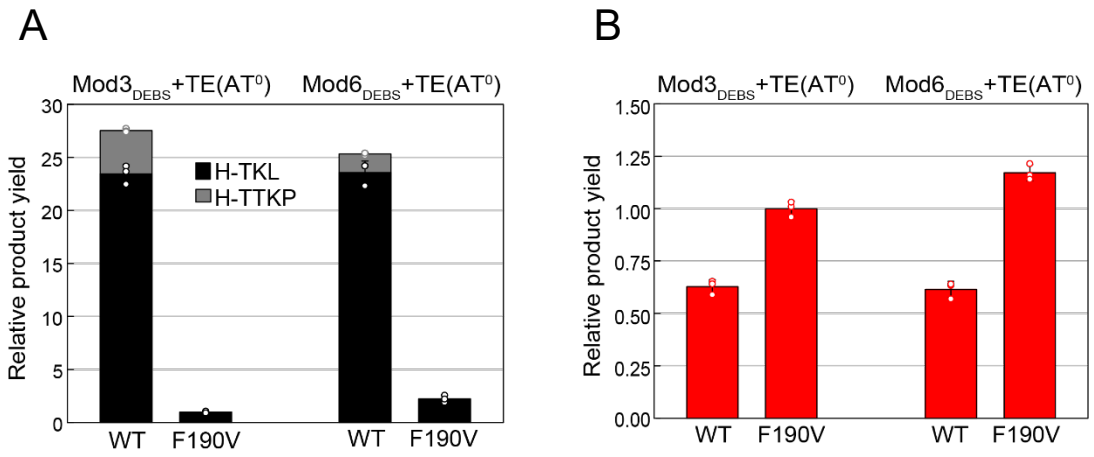


Figure 2.12. *In vitro triketide lactone assay for WT and F190V DszAT under pure carboxyacyl-coA extender unit conditions. Normalized representation of the amount of H-TKL (black), H-TTKP (grey), and F-TKL (red) products are shown for reactions containing 10 μM Mod3$_{DEBS}$+TE(AT^0) or Mod6$_{DEBS}$+TE(AT^0), 10 uM DszAT variant, 10 mM NDK-SNAC, and (A) 1 mM malonyl-CoA or (B) 1 mM fluoromalonyl-CoA under regenerative conditions. Reactions were quenched after 1 h at 37 ºC and analyzed by LC-QQQ. Data are mean ± s.d. of technical replicates (n=3).*

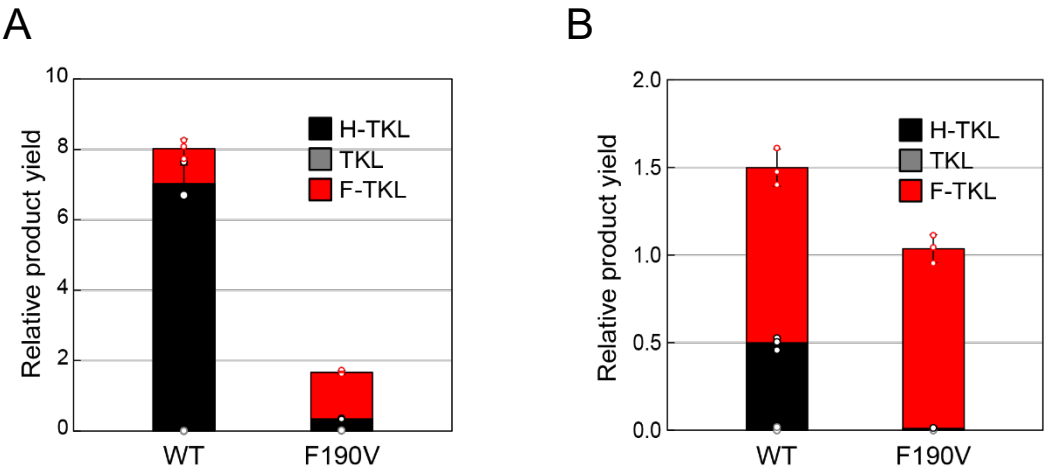


Figure 2.13. *In vitro triketide lactone assay for WT and F190V DszAT under mixed extender unit conditions. Normalized representation of the amount of H-TKL (black), H-TTKP (grey), and F-TKL (red) products are shown for reactions containing 10 μM Mod3$_{DEBS}$+TE(AT^0) or Mod6$_{DEBS}$+TE(AT^0), 10 uM DszAT variant, 10 mM NDK-SNAC, 1 mM methylmalonyl-CoA, 1 mM fluoromalonyl-CoA, and (A) 1 mM or (B) 0.1 mM malonyl-CoA under regenerative conditions. Reactions were quenched after 1 h at 37 ºC and analyzed by LC-QQQ. Data are mean ± s.d. of technical replicates (n=3).*

We then tested the two enzymes under mixed extender unit conditions. The two DszAT variants were compared at a 1 malonyl-CoA:1 methylmalonyl-CoA:1 fluoromalonyl-CoA extender unit ratio under regenerative conditions. We found that WT DszAT produced H-TKL as the major product and F-TKL as the minor product whereas the F190V variant was able to suppress the production of H-TKL and produced F-TKL as the major product (*Figure 2.13*). This result highlights the importance of the decreased activity of F190V mutant on malonyl-CoA for achieving regioselective fluoromalonyl-CoA incorporation as many polyketide molecules are produced from multiple extender units. Since malonyl-CoA is the most common extender unit used and is also at the highest concentration in the cell (~ 35 μM in *E. coli* [33]), selectivity between malonyl-CoA and fluoromalonyl-CoA is the most important for development of a more general platform for fluorinated polyketide production. To further assess the effect of carboxyacyl-CoA concentration on the regioselectivity of this system, TKL H-TKL and F-TKL production was assessed at a ratio of 1 malonyl-CoA:10 fluoromalonyl-coA. Under the increased fluoromalonyl-CoA condition, WT DszAT was able to produce F-TKL as the major product (*Figure 2.13*). Notably, under this condition, F190V DszAT variant made no detectable amount of H-TKL (*Figure 2.13*). Taken together these experiments showcase the ability of DszAT F190V to suppress malonyl-CoA extender unit incorporation and promote the regioselective incorporation of fluoromalonyl-coA into polyketide products.

***In vitro* 6dEB production with DszAT F190V.** Having successfully engineered and characterized a fluoromalonyl-CoA selective *trans*-AT, we were ready to proceed to incorporating fluorine into larger, more complex polyketide targets. A major challenge in moving from single-module systems that carry out an individual chain extension step is that the PKS will now contain multiple AT domains and utilize multiple extender units. In addition to fluoromalonyl-CoA, methylmalonyl-CoA and malonyl-CoA would be needed to construct the majority of the polyketide backbone. Therefore, we expect that competition between different extender units and AT domains could highly affect fluoromalonyl-CoA incorporation.

For these experiments, we decided to continue with the DEBS system [34] since the polyketide synthase has been entirely reconstituted *in vitro*, allowing for the production of the full-length erythromycin precursor 6-deoxyerythronolide B (6dEB) [27]. This system also has the advantage of not utilizing malonyl-CoA, enabling simpler analysis. The PKS consists of 6 modules split between 3 polypetides: DEBS1 contains the loading didomain (LDD_{DEBS}) along with modules 1 and 2, DEBS2 contains modules 3 and 4, and DEBS3 contains modules 5 and 6 (*Figure 2.14*). DEBS1-3 produce 6dEB through a series of nearly 30 enzymatic steps from 1 propionyl-CoA starter unit and 6 (2*S*)-methylmalonyl-CoA extender units (*Figure 2.14*). In the *in vitro* reconstitution system of DEBS, enzymatic functions contained within DEBS1 protein are separated into 3 different proteins - LDD_{DEBS}, DEBS module 1 ($Mod1_{DEBS}$), and DEBS module 2 ($Mod2_{DEBS}$) proteins – whereas DEBS2 and DEBS3 can be utilized in their native forms [27]. Thus, the *in vitro* reconstitution experiments contain a total of 5 separate polyketide synthase proteins (*Figure 2.15*).

In order to incorporate fluorine into the system, we needed to introduce the inactivating mutation to an AT domain of one of the DEBS modules. We initiated these experiments with generating the AT^0 mutation in module 6 as fluorine incorporation at the last module of the pathway would be the most direct assessment of fluorine incorporation since it requires the fewest number of downstream steps before offloading from the PKS and analysis as 6dEB. We therefore set out to introduce active site mutation (S2107A) to the AT domain of module 6 of DEBS3 using

a variety of different strategies. This effort is complicated both by the size of the DEBS1-3 proteins as well as the repetitive nature of the sequence as the modules are thought to arise by gene duplication events. As such, each AT domain displays high sequence identity to the other 5 AT domains in the system and is difficult to target. Furthermore, the repetitive nature of the sequence leads to issues with recombination where parts of the sequence can be lost. Many different strategies were tested. Ultimately, the strategy that yielded the best results was to introduce a selectable marker as intermediate. The targeted area was removed by digestion to insert a Cm^R cassette. The Cm^R cassette could then be removed and replaced by a segment containing the mutated AT domain. Intermediates were screened by antibiotic selection as well as restriction analysis. The final product was fully sequenced by complete plasmid sequencing in order to confirm that no other segments were lost or mutated.

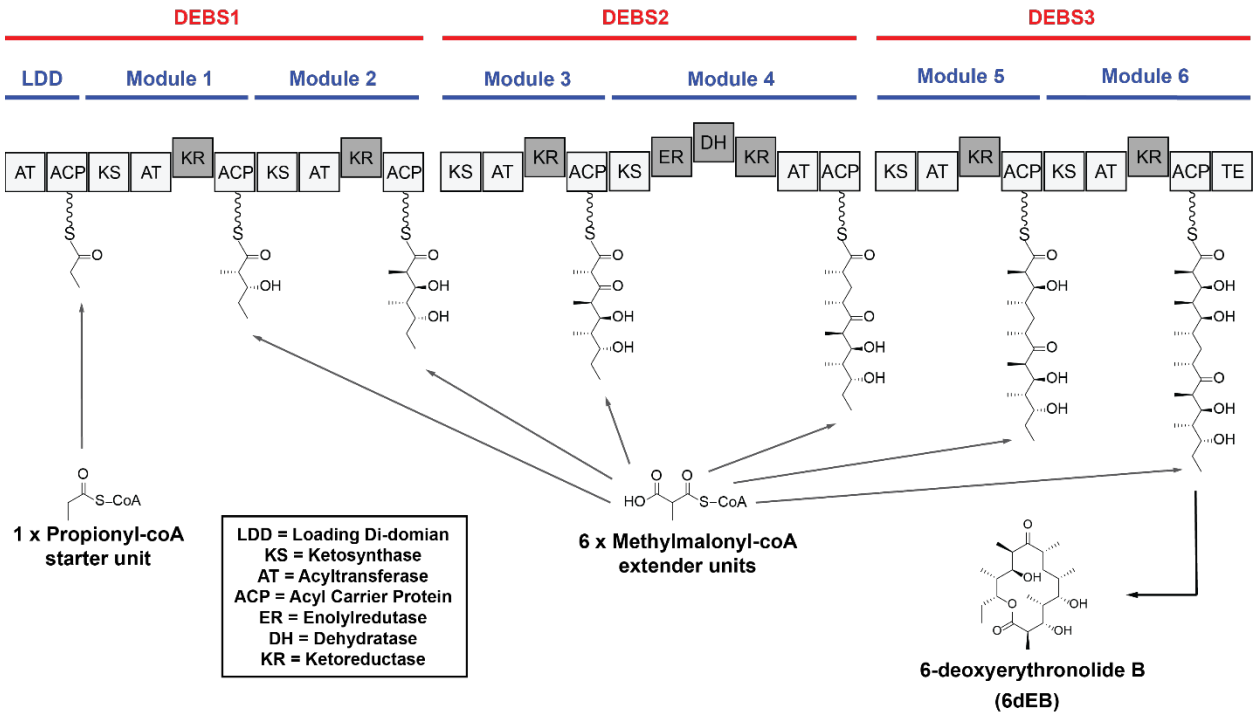


Figure 2.14. *The 6-deoxyerythronolide B synthase (DEBS) pathway. The DEBS pathway consists of the loading didomain (LDD) and 6 polyketide modules (module 1-6) split across 3 polypeptides: 1) DEBS1 containing LDD and modules 1-2; 2) DEBS2 containing module 3-4; and 3) DEBS3 containing module 5-6. DEBS1-3 produce 6-deoxyerythronolide B (6dEB), the aglycone precursor to an important antibiotic erythromycin, from 1 propionyl-CoA and 6 methylmalonyl-CoA. Each DEBS module consists of three necessary domains: ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP), and up to three optional tailoring domains: enoylreductase (ER), dehydratase (DH), and ketoreductase (KR).*

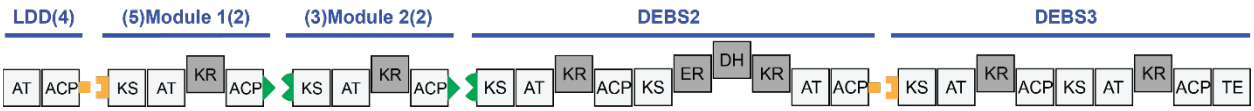


Figure 2.15. *The DEBS system used for In vitro reconstitution of 6dEB production. The complete in vitro system of DEBS consists of 5 proteins: LDD_{DEBS}, $Mod1_{DEBS}$, $Mod2_{DEBS}$, DEBS2, and DEBS3. LDD_{DEBS} was fused to the C-terminal docking domain of module 4 (yellow) while the module 1 protein was fused to the N-terminal docking domain of module 5 (yellow) to facilitate compatible interactions. Similarly, the module 1 protein was fused to the C-terminal docking domain of module 2 (green) while the module 2 protein was fused to the N-terminal docking domain of module 3 (green). The docking domains, both native and engineered, in the system are explicitly depicted for clarity.*

In addition to the DEBS proteins, this assay also requires PrpE, MatB, and methylmalonyl-coA epimerase (Epi), which together provide necessary propionyl-CoA starter unit, (2*S*)-methylmalonyl-CoA, and fluoromalonyl-CoA extender units from propionate, methylmalonate, and fluoromalonate, respectively. PrpE catalyzes the ATP-dependent ligation between coenzyme A and propionate to produce propionyl-CoA. MatB catalyzes the ATP-dependent ligation between coenzyme A and methylmalonate or fluoromalonate to produce (2R)-methylmalonyl-coA and fluoromalonyl-coA respectively. The methylmalonyl-CoA epimerase (Epi) then catalyzes the racemization of methylmalonyl-coA to form (2S)-methylmalonyl-coA. Enzymes used in this assay were overexpressed in *E. coli* and purified to near homogeneity (*Figure 2.16*).

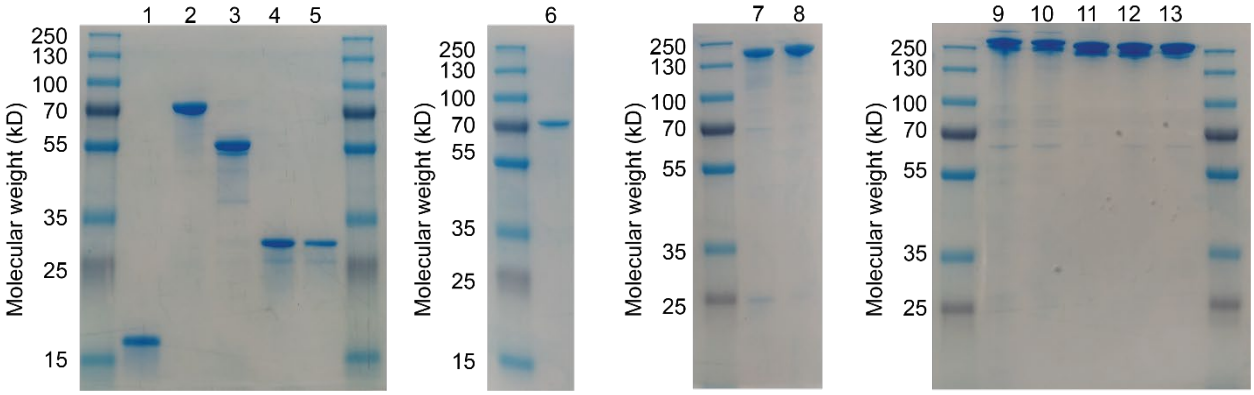
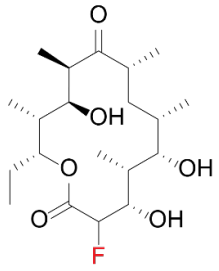


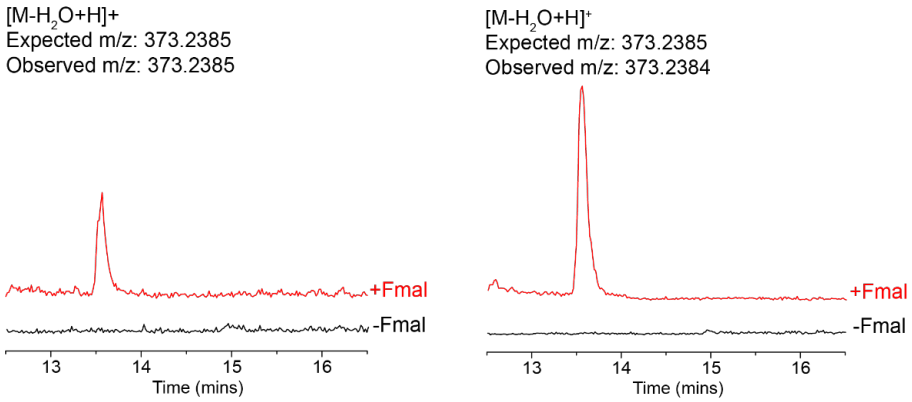
Figure 2.16 *SDS-PAGE of purified proteins used in the 6dEB in vitro production assay. (lane 1) methylmalonyl-coA , (lane 2) MatB, (lane 3) Prpe, (lane 4) WT DszAT, (lane 5) F190V DszAT, (lane 6) LDD, (lane 7) $Mod1_{DEBS}$, (lane 28) $Mod2_{DEBS}$, (lane 9) DEBS2, (lane 10) DEBS2(Mod3 AT^0), (lane 11) DEBS3, (lane 12) DEBS3(Mod5 AT^0), and (lane 13) DEBS3 (Mod6 AT^0).*

The *in vitro* production of 2-fluoro-2-desmethyl 6dEB was attempted by reconstituting DEBS pathway with active site mutation in the AT domain of module 6 (Mod6 AT^0 DEBS) with LDD_{DEBS}, $Mod1_{DEBS}$, $Mod2_{DEBS}$, DEBS2, and DEBS3(Mod6 AT^0) and providing propionate, methylmalonate, and fluoromalonate. When the Mod6 AT^0 mutation (DEBS3 S2107A) was complemented with either wild-type or F190V DszAT, a mass peak corresponding to desmethyl 6dEB containing a single fluorine substituent was observed (*Figure 2.17B*). The appearance of this peak was dependent on both the presence of fluoromalonate and DszAT (*Figure 2.17C*). Moreover, this compound was not detected when wild-type DEBS3 was utilized without the AT-inactivating mutation (*Figure 2.17C*). This result demonstrates that our complementation system is competent for incorporating fluoromalonyl-CoA into a full-length complex polyketide molecules assembled in the presence of other competing extender units.

A



B



C

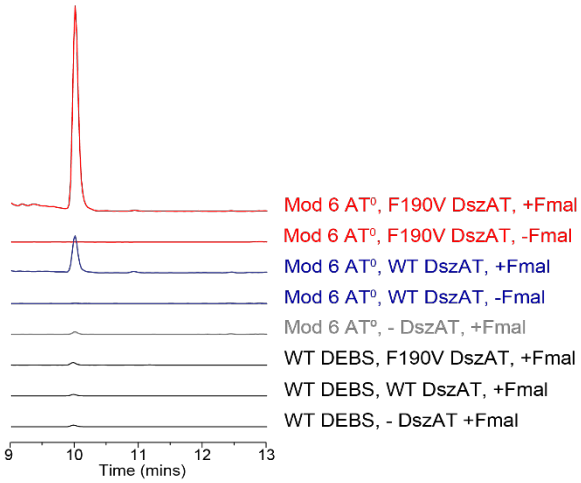


Figure 2.17. *In vitro reconstitution of fluorinated 6dEB analog production with Mod6 AT^0 DEBS. (A) Structure of 2-fluoro-2-desmethyl 6dEB, the expected product of fluoromalonyl-CoA incorporation by module 6 of DEBS. (B) Extracted ion chromatogram showing exact mass of in vitro production of monofluorinated desmethyl 6dEB by Mod6 AT^0 DEBS complemented with WT DszAT (left) or F190V DszAT (right). Reactions containing 2 μM each of LDD$_{DEBS}$, Mod1$_{DEBS}$, Mod2$_{DEBS}$, DEBS2, and DEBS3(Mod6 AT^0), 6 μM WT or F190V DszAT, 2 mM methylmalonate, and when used 10 mM fluoromalonate (+Fmal, red) were analyzed by LC-QTOF. Chromatograms are representative of at least 3 technical replicates. (C) Extracted ion chromatograms of reactions showing in vitro production of monofluorinated desmethyl 6dEB analyzed by LC-QQQ using MRM transition 373.2→275.1. Reaction conditions follow as in (B) with 2 μM DEBS3 used in place of DEBS3(Mod6 AT^0) in WT DEBS reactions (black). Chromatograms are representative of at least 3 technical replicates.*

To further elucidate regioselectivity of the system, we investigated the substitution pattern of the monofluorinated desmethyl-6dEB analog produced by both F190V and WT DszAT through fragmentation analysis. 6dEB had been shown to fragment into 4 major daughter ions (*Figure 2.18*) [35]. Each of the daughter ions arises from different parts of the parental ion carbon skeleton. Thus, the mass of each daughter ion reflects the substitutions residing on a specific location of 6dEB. Analysis of the mass shifts of all four daughter ions can then allow the location of the substitution to be narrowed down for a 6dEB analog of interest. In addition, each major daughter ion containing hydroxy groups undergoes successive losses of water to yield smaller daughter ions of the same carbon skeleton, allowing further confirmation of the daughter ion mass.

The fragmentation patterns obtained from the monofluorinated desmethyl 6dEB analog produced by complementing Mod6 AT^0 DEBS with either WT or F190V DszAT were both consistent with incorporation of fluoromalonyl-CoA extender unit by Module 6 of DEBS to produce 2-fluoro-2-desmethyl 6dEB (*Figure 2.19, 2.20*). Masses of ions in A, B, and D families match those expected for fluorine substitution in place of a methyl group on carbon 2, 4, 6, or 8 of 6dEB. Loss of HF from daughter ions expected to contain fluorine substitution was also observed in addition to successive losses of water due to hydroxy groups, which further supports the presence of fluorine in the molecule. Notably, the most prominent fragment in 6dEB MS/MS spectrum, the C ion (*Figure 2.18*), was entirely missing in the MS/MS spectrum of 2-fluoro-2-desmethyl 6dEB. This lack of detectable fragment prevents narrowing down the definite location of the fluorine substitution. However, the suppression of ion C was similarly observed when malonate was provided to the system in place of fluoromalonate (*Figure 2.21)* and is consistent with the reactivity of 6dEB analog with a modification near the C3 position. The lack of methyl substituent at C2 as in 2-fluoro-2-methyl 6dEB would reduce the reactivity of A with respect to the hydride shift from C3 to C9 necessary to form C ion (*Figure 2.18*).

We found that the ratio of 2-fluoro-2-desmethyl 6dEB analog to 6dEB was higher with F190V DszAT compared to wild-type (*Figure 2.22*). Steady-state kinetic experiments show that the main beneficial role of the F190V mutation should be in suppressing the incorporation of malonyl-CoA. However, in production assays 6dEB, the DszAT F190V mutant appears to also promote an increased level of fluoromalonyl-CoA incorporation. The molecular basis of this observation may be due to the decreased hydrolysis rate of fluoromalonyl-CoA exhibited by F190V mutant. The observation of native 6dEB in both systems is not surprising, as it has been previously observed that the other *cis*-AT domains of DEBS1-3 may be able to complement AT^0 mutations [12].

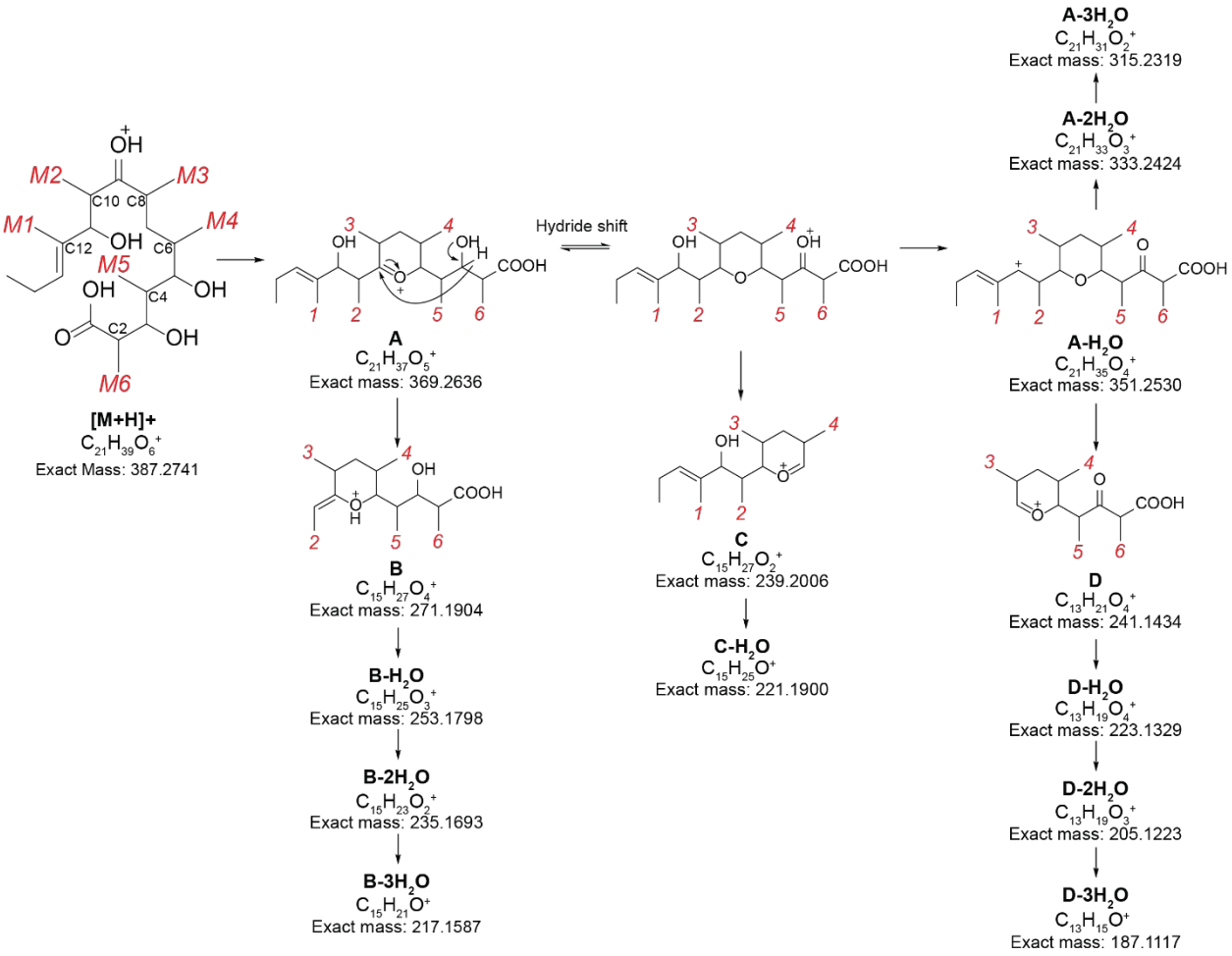
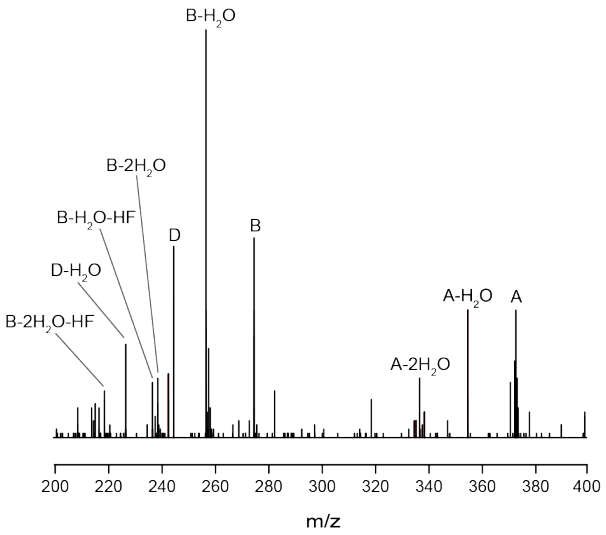


Figure 2.18. *Fragmentation analysis of 6dEB and analogs. 6dEB fragments into 4 daughter fragments shown as A, B, C, and D. Each daughter ion is derived from different parts of the carbon skeleton of the parent 6dEB. Module numbering (M1-M6, red) indicates the DEBS module responsible for installing the specified substituent. Carbon numbering (C2-C12, black) starts at the carboxylic acid with Module 6 being responsible for the substituent at C2. In addition, successive losses of water due to hydroxy groups are observed from all major daughter ions.*

A

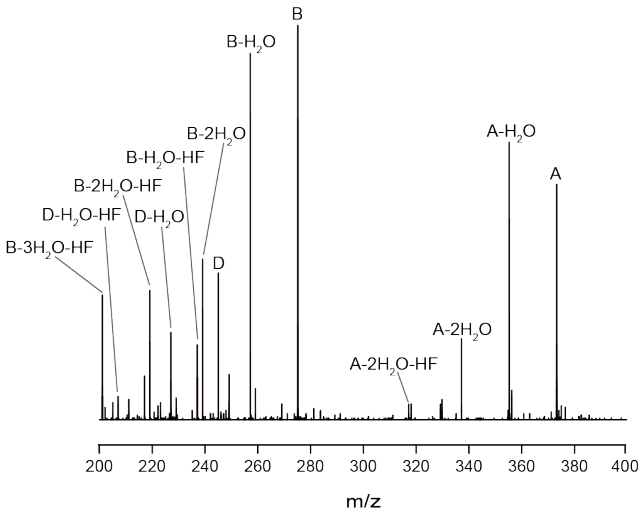


B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2372
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2305
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2124
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	Not observed
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1675
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1568
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.2423
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1471
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.138	219.1356
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	Not observed
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1186
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1082
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	Not observed
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	Not observed

Figure 2.19. *LC-QTOF fragmentation analysis of the monofluorinated desmethyl 6dEB analog produced in vitro by complementation of Mod6 AT^0 DEBS with WT DszAT and fluoromalonate. (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 technical replicates.*

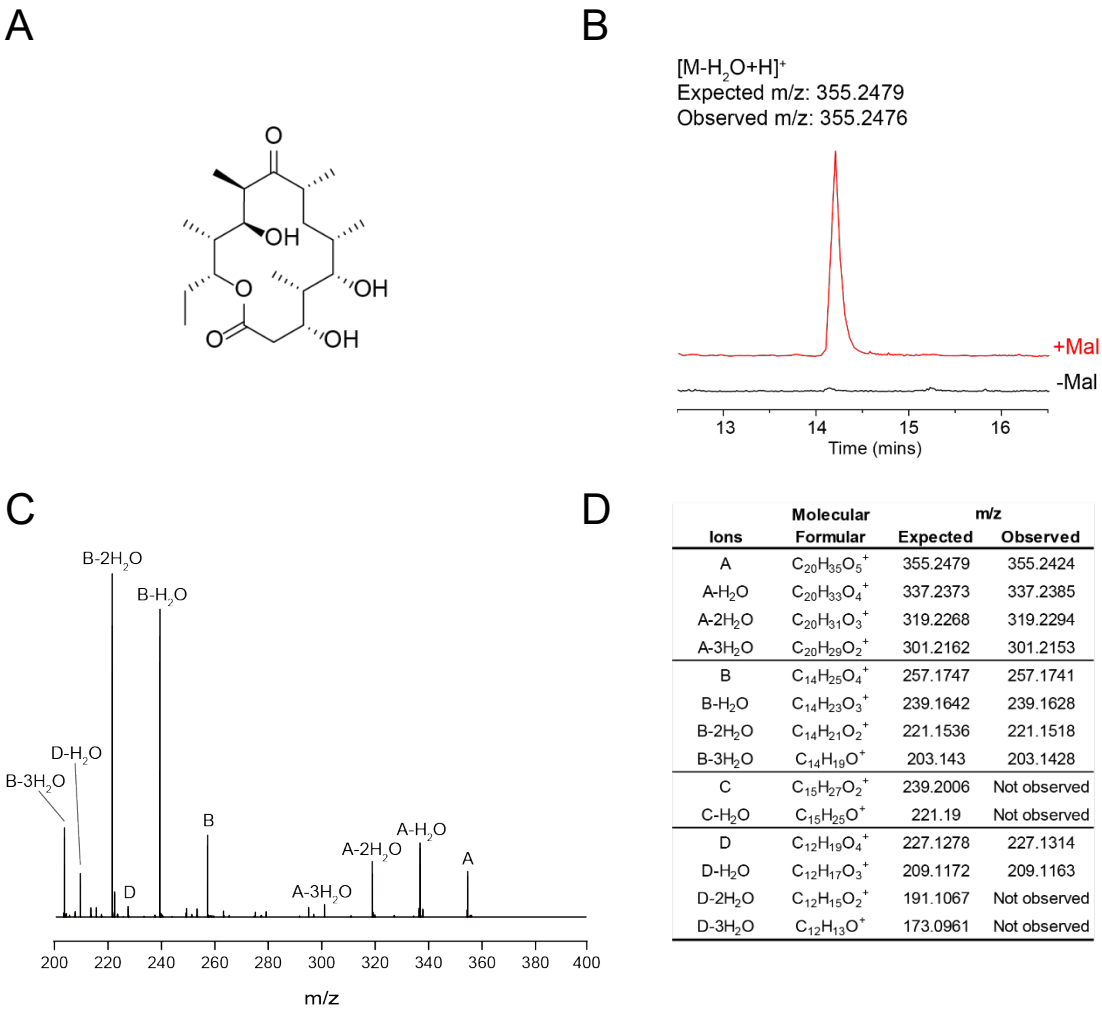
A



B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2373
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2258
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2156
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	317.2136
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1661
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1548
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.1466
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1471
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.138	219.1352
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1281
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1179
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1083
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	Not observed
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	Not observed

Figure 2.20. *LC-QTOF fragmentation analysis of the monofluorinated desmethyl 6dEB analog produced in vitro by complementation of Mod6 AT^0 DEBS with F190V DszAT and fluoromalonate. (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 technical replicates.*



Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{35}O_5^+$	355.2479	355.2424
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.2385
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2294
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	301.2153
B	$C_{14}H_{25}O_4^+$	257.1747	257.1741
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.1628
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1518
B-3H_2O	$C_{14}H_{19}O^+$	203.143	203.1428
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{19}O_4^+$	227.1278	227.1314
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.1163
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	Not observed
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	Not observed

Figure 2.21. *In vitro reconstitution of desmethyl 6dEB analog production with Mod6 AT^0 DEBS. (A) Structure of 2-desmethyl 6dEB, the expected product of malonyl-CoA incorporation by module 6 of DEBS. (B) Extracted ion chromatogram showing exact mass of in vitro production of desmethyl 6dEB by Mod6 AT^0 DEBS complemented with WT DszAT. Reactions containing 2 μM each of LDD$_{DEBS}$, Mod1$_{DEBS}$, Mod2$_{DEBS}$, DEBS2, and DEBS3(Mod6 AT0), 6 μM WT DszAT, 2 mM methylmalonate, and when used 10 mM malonate (+Mal, red) were analyzed by LC-QTOF. Chromatograms are representative of at least 3 technical replicates. (C) Fragmentation analysis of desmethyl 6dEB analog produced in vitro by complementation of Mod6 AT^0 DEBS with WT DszAT and malonate shows presence of ions in A, B, and D families. (D) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 technical replicates.*

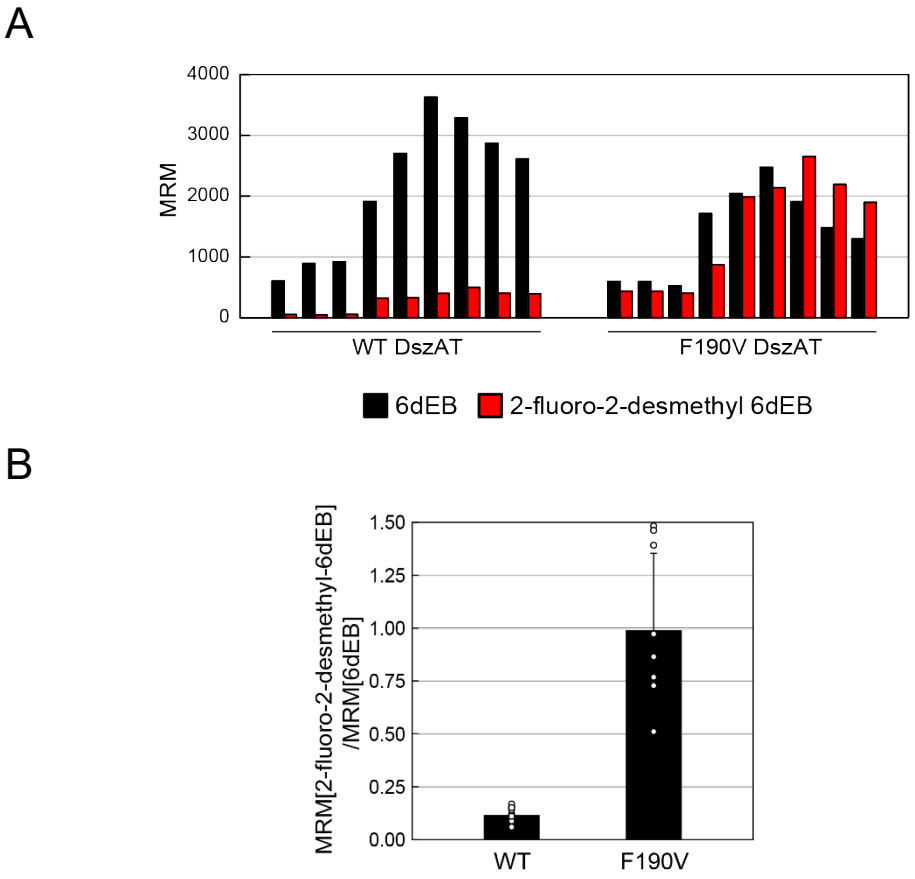


Figure 2.22. *Production of 2-fluoro-2-desmethyl 6dEB by complementation of DEBS3(Mod6 AT^0) with either WT or F190V DszAT. The in vitro reconstitution reaction was monitored by LC-QQQ using MRM specific for the product of interest. (A) Integrated MRM of two products shown for individual reactions (6dEB, black; 2-fluoro-2-desmethyl 6dEB, red). (B) Ratio between integrated MRM of 2-fluoro-2-desmethyl 6dEB to integrated MRM of 6dEB. Data are mean ± s.d. of 3 technical replicates of 3 biological replicates (n=9)*

With the initial success in incorporation of fluoromalonyl-CoA at module 6 position of 6dEB, we constructed additional mutant proteins with active site mutation in the AT domains of module 3 and 5 (DEBS2 S653A and DEBS3 S642A respectively). These proteins allowed reconstitution of Mod3 AT^0 DEBS and Mod5 AT^0 DEBS.

We attempted to produce 8-fluoro-8-desmethyl 6dEB with LDD$_{DEBS}$, Mod1$_{DEBS}$, Mod2$_{DEBS}$, DEBS2(Mod3 AT^0), and DEBS3 using propionate, methylmalonate, and fluoromalonate. However, we found no detectable signal corresponding to a monofluorinated desmethyl 6dEB analog by LC/QTOF analysis when the active site mutation was complemented by WT or F190V DszAT. In providing the same *in vitro* reconstitution system with methylmalonate and malonate, a mass peak corresponding to a desmethyl 6dEB analog could be detected when WT DszAT was provided (*Figure 2.23*). The peak was largely dependent on malonyl-CoA. Fragmentation analysis supported that a methyl substitution at C6 or C8 position was replaced by a hydrogen atom (*Figure 2.23*). This result is consistent with malonyl-CoA being incorporated by Mod3 AT^0 DEBS to produce 8-desmethyl 6dEB. This observation of a single incident of malonyl incorporation

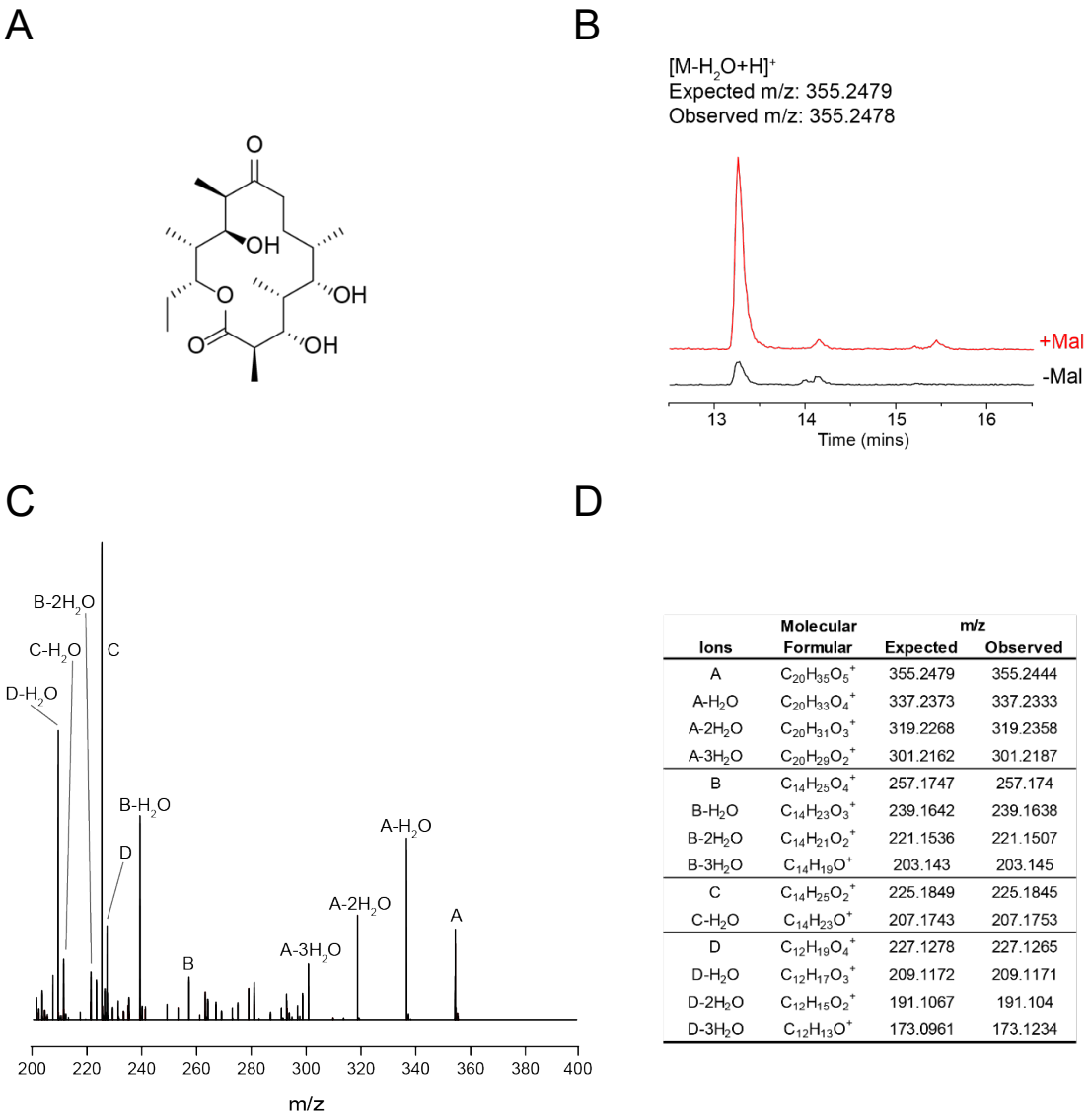
suggests a fully functional pathway reconstitution that exhibits a regiospecific behavior. As single modular experiments with $Mod3_{DEBS}$+TE(AT^0) suggests that fluoromalonyl-CoA is accepted by Module 3 of DEBS under pure and mixed extender unit conditions (*Figure 2.12-2.13*), the lack of observable fluoromalonyl-CoA incorporation into 6dEB could be due to either the competition by other *cis*-AT domains that prefer methylmalonyl-CoA or the rejection of the fluorine-containing polyketide intermediate by downstream enzymatic activities. It is foreseeable that as the number of downstream steps from unnatural extender unit incorporation increases, the overall efficiency of the system decreases.

We then attempted to reconstitute Mod5 AT^0 DEBS with LDD_{DEBS}, $Mod1_{DEBS}$, $Mod2_{DEBS}$, DEBS2, and DEBS3(Mod5 AT^0) to produce 4-fluoro-4-desmethyl 6dEB. When the system was provided with propionate, methylmalonate, and fluoromalonate, a mass peak corresponding to a monofluorinated desmethyl 6dEB was observed (*Figure 2.24*). This mass peak was observed when both WT and F190V DszAT were used to complement that AT^0 mutation. Furthermore, the compound was dependent on fluoromalonate and the AT^0 mutation as wild-type DEBS3 did not produce the peak (*Figure 2.24*). Notably, this mass peak has a different retention time from that observed with Mod6 AT^0 DEBS, suggesting a different isomer.

We investigated the regioselectivity of the monofluorinated desmethyl 6dEB product produced by Mod5 AT^0 DEBS complemented by both WT and F190V DszAT by fragmentation analysis. We found that this product exhibited an identical fragmentation pattern as observed for the monofluorinated desmethyl 6dEB molecule produced by Mod6 AT^0 DEBS (*Figure 2.25-2.26*). This similarity of fragmentation patterns suggests that the two isomers maintain fluorine substitutions in nearby positions on the carbon skeleton. Notably, loss of HF was observed from ions A, B, and D families, and ion C was not detectable. The suppression of ion C was similarly observed when malonate was provided to the Mod5 AT^0 DEBS system in place of fluoromalonate (*Figure 2.27)* and suggests a modification near C3 position. These characteristics are consistent with incorporation of fluoromalonyl-CoA by module 5, producing 4-fluoro-4-desmethyl 6dEB.

The successful production of 4-fluoro-4-desmethyl 6dEB is particularly interesting as it demonstrates further elongation from the incorporated fluoromalonyl-CoA. This result suggests the possibility of installing fluorine substitutions at various upstream positions on modular polyketide molecules.

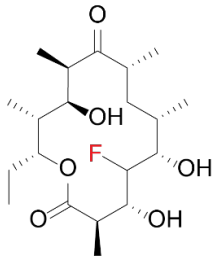
To determine selectivity of the system conferred by *trans*-AT proteins, we investigated the ratio between signals of 4-fluoro-4-desmethyl 6dEB analog to 6dEB when WT and F190V DszAT were used to complement AT^0 mutation in module 5 of DEBS. We found that the ratio was higher with F190V DszAT (*Figure 2.28*). This observed increased in selectivity is due to both an increase in MRM signal of 4-fluoro-4-desmethyl 6dEB and a decrease in that of 6dEB. This result is consistent with that observed with Mod6 AT^0 DEBS.



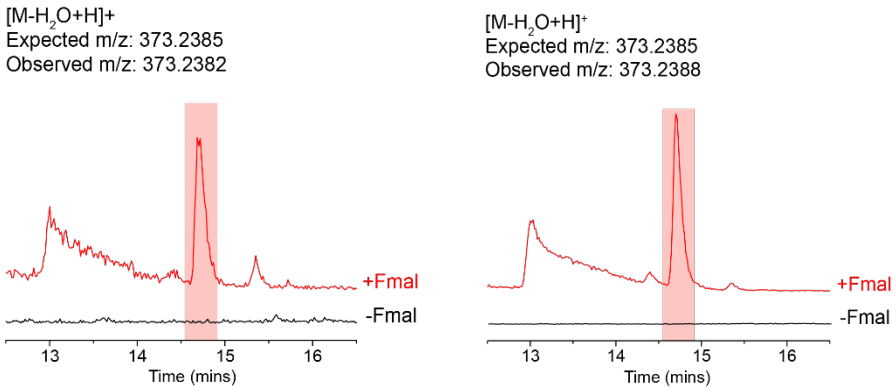
Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{35}O_5^+$	355.2479	355.2444
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.2333
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2358
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	301.2187
B	$C_{14}H_{25}O_4^+$	257.1747	257.174
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.1638
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1507
B-3H_2O	$C_{14}H_{19}O^+$	203.143	203.145
C	$C_{14}H_{25}O_2^+$	225.1849	225.1845
C-H_2O	$C_{14}H_{23}O^+$	207.1743	207.1753
D	$C_{12}H_{19}O_4^+$	227.1278	227.1265
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.1171
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	191.104
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	173.1234

Figure 2.23. *In vitro reconstitution of desmethyl 6dEB analog production with Mod3 AT^0 DEBS. (A) Structure of 8-desmethyl 6dEB, the expected product of malonyl-CoA incorporation by module 3 of DEBS. (B) Extracted ion chromatogram showing exact mass of in vitro production of desmethyl 6dEB by Mod3 AT^0 DEBS complemented with WT DszAT. Reactions containing 2 μM each of LDD$_{DEBS}$, Mod1$_{DEBS}$, Mod2$_{DEBS}$, DEBS2(Mod3 AT^0), and DEBS3, 6 μM WT DszAT, 2 mM methylmalonate, and when used 10 mM malonate (+Mal, red) were analyzed by LC-QTOF. Chromatograms are representative of at least 3 technical replicates. (C) Fragmentation analysis of desmethyl 6dEB analog produced in vitro by complementation of Mod3 AT^0 DEBS with WT DszAT and malonate shows presence of ions in A, B, and D families. (D) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 technical replicates.*

A



B



C

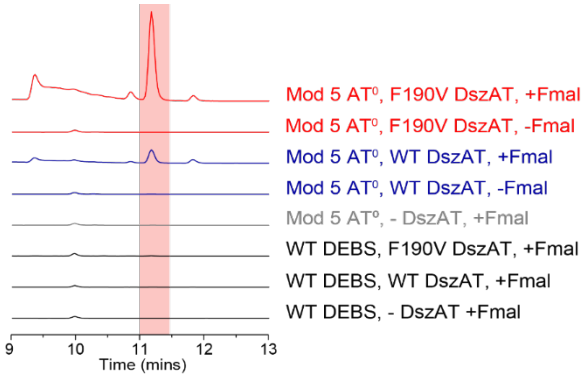
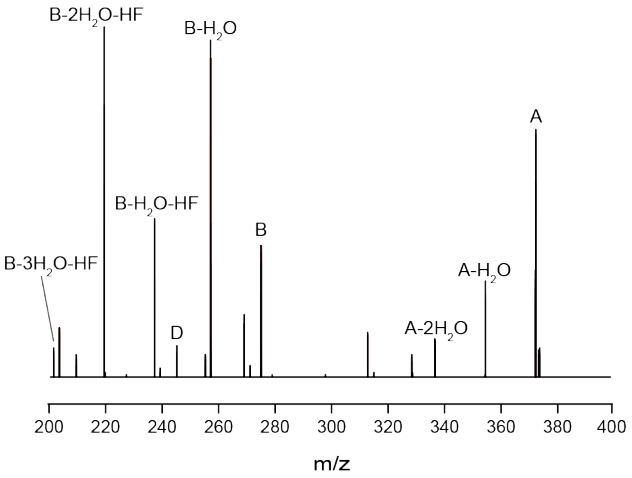


Figure 2.24. *In vitro reconstitution of fluorinated 6dEB analog production with Mod5 AT^0 DEBS. (A) Structure of 4-fluoro-4-desmethyl 6dEB, the expected product of fluoromalonyl-CoA incorporation by module 5 of DEBS. (B) Extracted ion chromatogram showing exact mass of in vitro production of monofluorinated desmethyl 6dEB by Mod5 AT^0 DEBS complemented with WT DszAT (left) or F190V DszAT (right). Reactions containing 2 μM each of LDD$_{DEBS}$, Mod1$_{DEBS}$, Mod2$_{DEBS}$, DEBS2, and DEBS3(Mod5 AT^0), 6 μM WT or F190V DszAT, 2 mM methylmalonate, and when used 10 mM fluoromalonate (+Fmal, red) were analyzed by LC-QTOF. Chromatograms are representative of at least 3 technical replicates. (C) Extracted ion chromatograms of reactions showing in vitro production of monofluorinated desmethyl 6dEB analyzed by LC-QQQ using MRM transition 373.2→275.1. Reaction conditions follow as in (B) with 2 μM DEBS3 used in place of DEBS3(Mod5AT^0) in WT DEBS reactions (black). Chromatograms are representative of at least 3 technical replicates.*

A

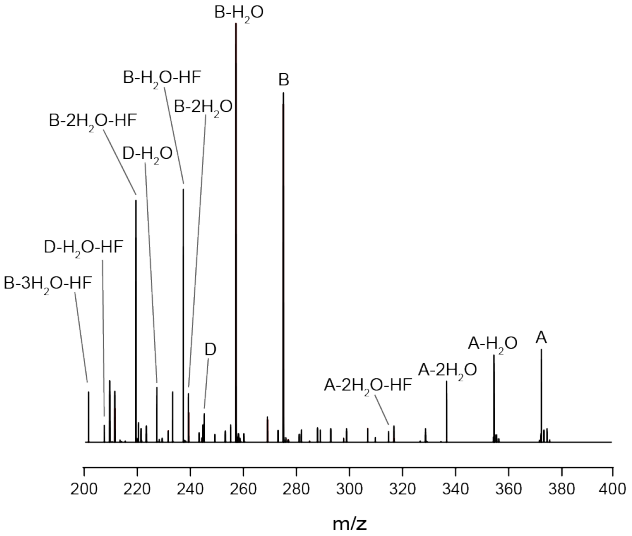


B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2372
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2259
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.1691
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	Not observed
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1638
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1515
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	Not observed
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1503
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.138	219.1399
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1129
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{18}FO_4^+$	245.1184	245.119
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	Not observed
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	Not observed
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	Not observed

Figure 2.25. *LC-QTOF fragmentation analysis of the monofluorinated desmethyl 6dEB analog produced in vitro by complementation of Mod5 AT^0 DEBS with WT DszAT and fluoromalonate. (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 technical replicates.*

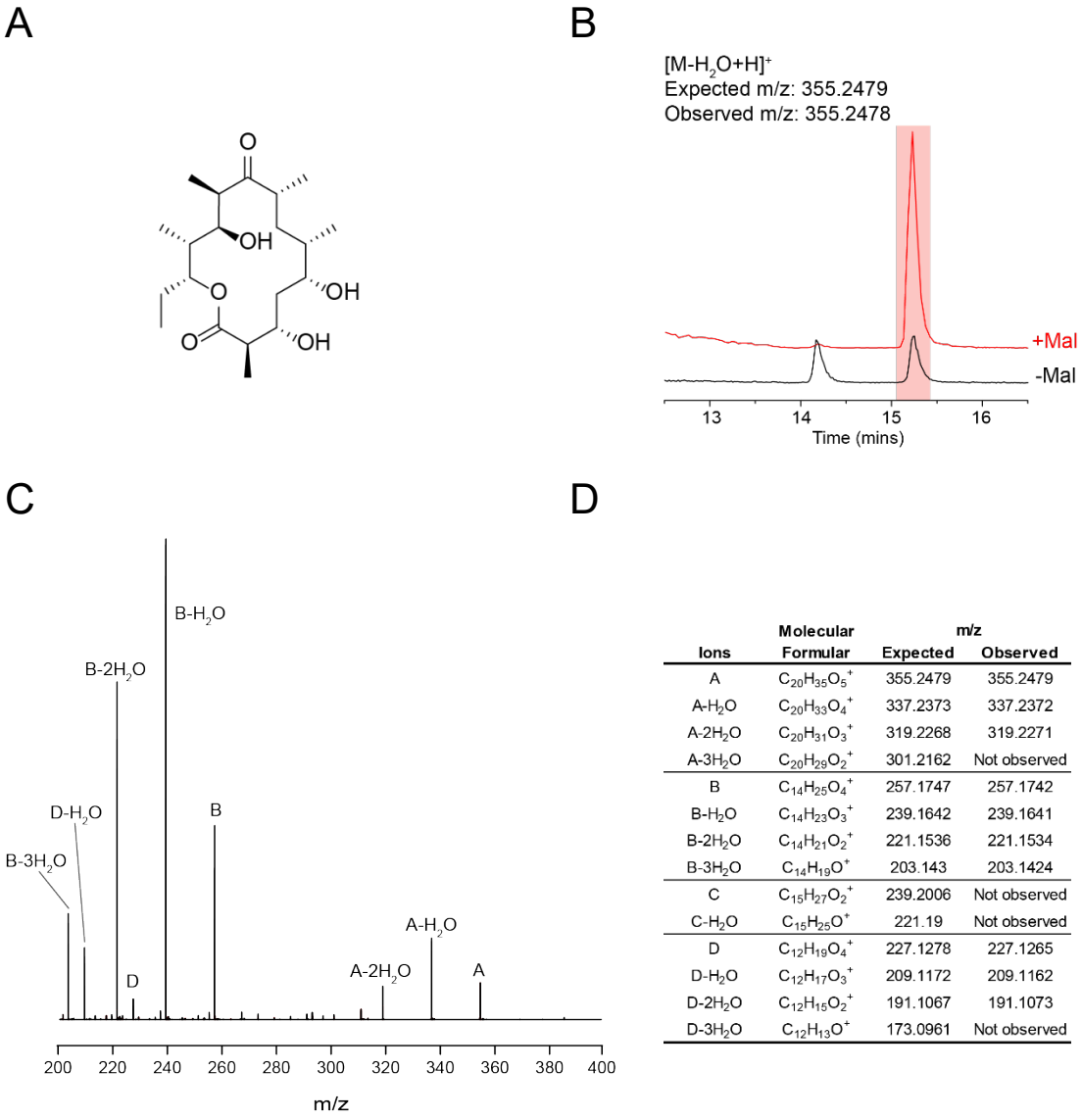
A



B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2383
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.225
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2116
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	317.2102
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1641
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1548
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.1445
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1481
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.138	219.1375
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1296
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1143
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1119
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	Not observed
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	207.0978

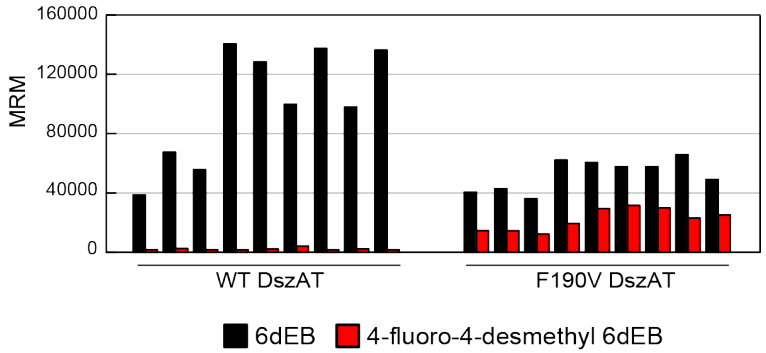
Figure 2.26. *LC-QTOF fragmentation analysis of the monofluorinated desmethyl 6dEB analog produced in vitro by complementation of Mod5 AT^0 DEBS with F190V DszAT and fluoromalonate. (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 technical replicates.*



Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{35}O_5^+$	355.2479	355.2479
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.2372
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2271
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	Not observed
B	$C_{14}H_{25}O_4^+$	257.1747	257.1742
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.1641
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1534
B-3H_2O	$C_{14}H_{19}O^+$	203.143	203.1424
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{19}O_4^+$	227.1278	227.1265
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.1162
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	191.1073
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	Not observed

Figure 2.27. *In vitro reconstitution of desmethyl 6dEB analog production with Mod5 AT^{0}DEBS. (A) Structure of 4-desmethyl 6dEB, the expected product of malonyl-CoA incorporation by module 5 of DEBS. (B) Extracted ion chromatogram showing exact mass of in vitro production of desmethyl 6dEB by Mod5 AT^0 DEBS complemented with WT DszAT. Reactions containing 2 μM each of LDD$_{DEBS}$, Mod1$_{DEBS}$, Mod2$_{DEBS}$, DEBS2, and DEBS3(Mod5 AT^0), 6 μM WT DszAT, 2 mM methylmalonate, and when used 10 mM malonate (+Mal, red) were analyzed by LC-QTOF. Chromatograms are representative of at least 3 technical replicates. (C) Fragmentation analysis of desmethyl 6dEB analog produced in vitro by complementation of Mod5 AT^0 DEBS with WT DszAT and malonate shows presence of ions in A, B, and D families. (D) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 technical replicates.*

A



B

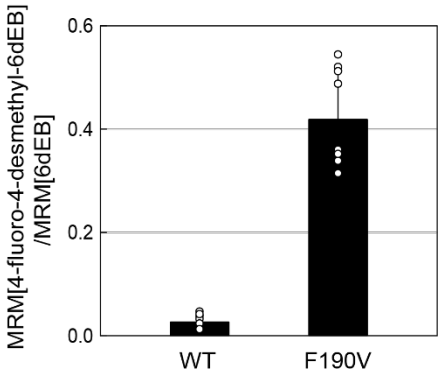


Figure 2.28. *Production of 4-fluoro-4-desmethyl 6dEB by complementation of DEBS3(Mod5 AT^0) with either WT or F190V DszAT. The in vitro reconstitution reaction was monitored by LC-QQQ using MRM specific for the product of interest. (A) Integrated MRM of two products shown for individual reactions (6dEB, black; 4-fluoro-4-desmethyl 6dEB, red). (B) Ratio between integrated MRM of 4-fluoro-4-desmethyl 6dEB to integrated MRM of 6dEB. Data are mean ± s.d. of 9 technical replicates.*

2.4 Conclusions

Through structure-guided protein engineering principles, we engineered a fluoromalonyl-CoA selective *trans*-AT protein from malonyl-coA-specific DszAT. Though steady-state kinetics characterization and single turnover end-point analysis, the engineered protein, DszAT F190V, was shown to have decreased activity on malonyl-CoA while maintaining a similar level of activity on fluoromalonyl-CoA compared to the wild-type enzyme. In addition, the native discrimination against methylmalonyl-CoA was preserved. Our kinetic data also support that transacylation of ACP substrates does not follow the ping-pong mechanism.

The DszAT F190V mutant was then used to complement an inactivating mutation in the AT domains of the fully reconstituted DEBS system to produce fluorine-containing 6dEB analogs *in vitro*. In order to develop this system, the S→A active site mutation was introduced to the AT domain of module 3 of DEBS2 and modules 5 and 6 of DEBS3. We found that both the WT and the F190V mutant of DszAT were able to support the production of 2-fluoro-2-desmethyl 6dEB and 4-fluoro-4-desmethyl analog, identified by exact mass and fragmentation pattern, when used

with Mod5 AT^0 DEBS and Mod6 AT^0 DEBS respectively. In these systems, the F190V DszAT was able to increase the selectivity of the system towards fluorine substitution by 5-10 fold.

2.5 References

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

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

3. Liou, G. F.; Khosla, C. Building-block selectivity of polyketide synthases. *Curr Opin Chem Biol* **2003***, 7*, 279-284.

4. Khosla, C. Structures and mechanisms of polyketide synthases. *J. Org. Chem.* **2009***, 74*, 6416-6420.

5. Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed. Engl.* **2009***, 48*, 4688-4716.

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

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

8. Bonnett, S.; Rath, C.; Shareef, A.; Joels, J.; Chemler, J.; Håkansson, K.; Reynolds, K.; Sherman, D. Acyl-coA subunit selectivity in the pikromycin polyketide synthase PikAIV: steady-state kinetics and active-site occupancy analysis by FTICR-MS. *Chemistry & Biology* **2011***, 18*, 1075-1081.

9. Musiol-Kroll, E. M.; Zubeil, F.; Schafhauser, T.; Härtner, 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*, 421-427.

10. Carpenter, S. M.; Williams, G. J. Extender unit promiscuity and orthogonal protein interactions of an aminomalonyl-ACP utilizing trans-acyltransferase from zwittermicin biosynthesis. *ACS Chem. Biol.* **2018***, 13*, 3361-3373.

11. 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.

12. 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*, 3796-3806.

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

14. Smart, B. E. Fluorine substituent effects (on bioactivity). *Journal of Fluorine Chemistry* **2001***, 109*, 3-11.

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

16. Mathews, I. I.; Allison, K.; Robbins, T.; Lyubimov, A. Y.; Uervirojnangkoorn, M.; Brunger, A. T.; Khosla, C.; DeMirci, H.; McPhillips, S. E.; Hollenbeck, M.; Soltis, M.; Cohen, A. E. The conformational flexibility of the acyltransferase from the disorazole polyketide synthase is revealed by an X-ray free-electron laser using a room-temperature sample delivery method for serial crystallography. *Biochemistry* **2017***, 56*, 4751-4756.

17. Starcevic, A.; Zucko, J.; Simunkovic, J.; Long, P. F.; Cullum, J.; Hranueli, D. ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res* **2008***, 36*, 6882-6892.

18. Yadav, G.; Gokhale, R. S.; Mohanty, D. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *Journal of Molecular Biology* **2003***, 328*, 335-363.

19. Weeks, A. M.; Chang, M. C. Y. Catalytic control of enzymatic fluorine specificity. *PNAS* **2012***, 109*, 19667-19672.

20. Weeks, A. M.; Coyle, S. M.; Jinek, M.; Doudna, J. A.; Chang, M. C. Y. Structural and biochemical studies of a fluoroacetyl-coA-specific thioesterase reveal a molecular basis for fluorine selectivity. *Biochemistry* **2010***, 49*, 9269-9279.

21. Ad, O. Engineered production of fluorinated polyketides inspired by mechanistic studies of polyketide synthase, University of California, Berkeley, 2017.

22. Sax, J. Engineering a trans-acyltransferase for fluorine incorporation into polyketides , University of California, Berkeley, 2016.

23. 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*, 1790-1792.

24. Dong, H.; Liffland, S.; Hillmyer, M. A.; Chang, M. C. Y. Engineering in vivo production of α-branched polyesters. *J. Am. Chem. Soc.* **2019***, 141*, 16877-16883.

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

26. 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*, 95-102.

27. Lowry, B.; Robbins, T.; Weng, C.; 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*, 16809-16812.

28. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. In *protein identification and analysis tools on the ExPASy server;* Walker, J. M., Ed.; The proteomics protocols handbook; Humana Press: New Jersey, 2005; pp 571-607.

29. Ushimaru, K.; Sangiambut, S.; Thomson, N.; Sivaniah, E.; Tsuge, T. New insights into activation and substrate recognition of polyhydroxyalkanoate synthase from Ralstonia eutropha. *Appl. Microbiol. Biotechnol.* **2013***, 97*, 1175-1182.

30. Jia, K.; Cao, R.; Hua, D. H.; Li, P. Study of class I and class III polyhydroxyalkanoate (PHA) synthases with substrates containing a modified side chain. *Biomacromolecules* **2016***, 17*, 1477-1485.

31. Wodzinska, J.; Snell, K. D.; Rhomberg, A.; Sinskey, A. J.; Biemann, K.; Stubbe, J. Polyhydroxybutyrate synthase: evidence for covalent catalysis. *J. Am. Chem. Soc.* **1996***, 118*, 6319-6320.

32. Molnos, J.; Gardiner, R.; Dale, G. E.; Lange, R. A continuous coupled enzyme assay for bacterial malonyl-CoA:acyl carrier protein transacylase (FabD). *Anal. Biochem.* **2003***, 319*, 171-176.

33. Rabinowitz, J. D.; Osterhout, R.; Bennett, B. D.; Gao, M.; Kimball, E. H.; Van Dien, S. J. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature Chemical Biology* **2009***, 5*, 593-599.

34. Khosla, C.; Tang, Y.; Chen, A. Y.; Schnarr, N. A.; Cane, D. E. Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annual Review of Biochemistry* **2007***, 76*, 195-221.

35. Ashley, G. W.; Carney, J. R. API-mass spectrometry of polyketides. II. Fragmentation analysis of 6-deoxyerythronolide B analogs. *J. Antibiot.* **2004***, 57*, 579-589.

Chapter 3: *Development of an E. coli production platform for fluorinated polyketides*

Portions of this work were performed in collaboration with the following persons: Swetha Ramesh cloned, purified, and helped characterize FabD protein.

3.1 Introduction

With the initial success in producing fluorine-containing polyketide molecules *in vitro*, we wanted to transfer this system to an *in vivo* fermentation system. *In vitro* experiments allowed us to initially test our ideas under highly controlled environments, allowing for elucidation of system parameters and their optimization. However, *in vitro* processes are challenging to adopt and scale, especially with regard to PKSs. Many PKS enzymes exist in large, multi-modular polypeptide chains that are not amenable to biochemical manipulation. Cloning PKS genes, as well as their overexpression and purification, is quite challenging. Typically, overexpression in the native hosts are low and heterologous expression suffers from issues of solubility. As such, these efforts often rely on the splitting of individual enzymes to develop a soluble system, which also requires determination of module and domain boundaries and the linker regions. Indeed, very few type I modular PKSs have been reconstituted *in vitro* to date [1, 2]. On the other hand, many PKS pathways have been fully reconstituted *in vivo* and shown to produce product in model, non-native polyketide-producing hosts [3-5]. Moreover, more recently developed molecular biology technologies would likely facilitate future PKS pathway reconstitution efforts [6-8]. Although product yields produced by PKS pathways are often low both in native and non-native producers, heterologous reconstitution still often leads to higher yields than *in vitro* reconstitution. Thus, we believe that the ability to produce fluorinated polyketide analogs *in vivo* would increase the variety of PKS pathways that could be manipulated through these methods.

The most popular hosts for reconstituting modular polyketide synthase pathways are *Escherichia coli* and *Streptomyces* spp. including model streptomycetes such as *S. ceolicolor, S. lividans, S. albus, and S. venezuelae* [9]. Many modular PKS pathways are found in soil-dwelling actinomyces species; therefore, well-characterized, fast-growing model streptomyces strains are a natural choice for reconstitution. They have the advantage of possessing secondary metabolism, allowing natural product production to be uncoupled from growth, and of maintaining a complex pool of metabolites including many precursors to polyketide molecules [10]. However, *E. coli* has more readily available genetic tools and better characterized cellular functions [11]. In addition, more biological parts such as promoters, RBSs, and terminators are available for controlling pathway expression [12, 13]. As we were starting from a model PKS pathway such as DEBS, which requires no specialized metabolites beyond propionyl-CoA and methylmalonyl-CoA, we decided that *E. coli* was an appropriate host for our work. However, we note the importance of investigating streptomyces strains as potential host organisms for fluorine incorporation into natural products as well.

Development of an *in vivo* system comes with multiple challenges as the precise control of each system component inside an *E. coli* cell is not straightforward. Although expression level of a single protein can typically be controlled by many different strategies, the level of each enzyme in a multiple enzyme pathway is influenced by other enzymes being expressed simultaneously and is subject to the sequence and identify of each enzyme. Similarly, precursors required for biosynthesis of polyketide molecules may or may not exist in the native metabolism of the heterologous host at a sufficient concentration [14]. While some of these precursors can be supplemented through the culture media, others will require a specialized transport system to enter the intracellular space. Moreover, when the biosynthesis requires a native metabolite that is heavily regulated by the host, the *in vivo* concentration may not be easily manipulated due to degradation or sequestration. Under these circumstances, further strain engineering to alter the native cellular homeostasis is likely required. Furthermore, in introducing foreign enzymes to produce foreign

metabolites which often have bioactivities, the toxicity of precursors, intermediates, and products must be considered. While bioactivities of some characterized natural products are known or could be inferred in the context of the production host, novel molecules and many pathway intermediates have little to no information regarding toxicity.

For our work, we decided to utilize the well-characterized DEBS pathway that we had developed *in vitro* in Chapter 2. We chose to start with the single-module DEBS system due to its relatively simple protein and precursor requirements. Beside the *in vitro* data showing that the F-TKL product is achievable through our complementation strategy, all proteins necessary for constructing this pathway, the single modular DEBS constructs and the *trans*-AT proteins, can all be functionally expressed under the same conditions in *E. coli* BAP1. We also know that feeding fluoromalonate to *E. coli* cells expressing $Mod6_{DEBS}$ has been shown to produce an observable amount of F-TKL from a relatively simple set of precursors, namely NDK-SNAC and fluoromalonyl-CoA [15]. The NDK-SNAC substrate can be supplemented directly to the culture media and can cross the cell membrane. Thus, the only substrate that would require a specialized transporter is fluoromalonyl-CoA. Challenges that we foresee and hope to address in this work include optimization of the relative levels of pathway proteins when concurrently expressed, transport of the fluoromalonyl-CoA precursor and of product, and pathway interference to growth and native metabolism of the host.

Our work presented in this chapter explores the potential of developing *E. coli* as a production platform for fluorine-containing polyketide molecules, using a single-module DEBS pathway as a model. We aim to establish the basic characters and optimize the conditions for production, in hope that it will establish the groundwork for related future endeavors.

3.2 Materials and Methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), magnesium sulfate anhydrous, and glycerol were purchased from EMD Biosciences (Damstadt, Germany). Carbenicillin (Cb), glucose, isopropyl-β-D-thiogalactopyranoside (IPTG), sodium chloride, calcium chloride, magnesium chloride hexahydrate, acetonitrile, dichloromethane, ethyl acetate, ethylene diamine tetraacetic acid disodium dihydrate (EDTA), sodium phosphate monobasic, potassium phosphate monobasic, ammonium chloride, and arabinose were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A trilithium sodium salt (coA), malonyl-CoA, methylmalonyl-CoA, malonic acid, methylmalonic acid, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenolpyruvate (PEP), adenosine trisodium phosphate sodium salt (ATP), myokinase, pyruvate kinase, lactate dehydrogenase, poly(ethyleneimine) (PEI), β-mercaptoethanol (BME), thiamine pyrophosphate (TPP), α-ketoglutaric acid, sodium phosphate dibasic heptahydrate, cysteamine, 1-ethyl-3-(-dimethylaminopropyl)carbodiimide hydrochloride (EDC), 4-dimethylaminopyridine (DMAP), acetonitrile, dimethyl sulfoxide (DMSO), biotin, thiamine hydrochloride, chloramphenicol, β-nicotinamide adenine dinucleotide (NAD^+), reduced β-nicotinamide adenine dinucleotide (NADH), β-nicotinamide adenine dinucleotide 2′-phosphate ($NADP^+$), α-ketoglutarate dehydrogenase, and bovine serum albumin (BSA) were purchased from Sigma-Aldrich (St. Louis, MO). Diethylfluoromalonate was purchased from Sigma-Aldrich (St. Louise, MO) or from Matrix Scientific (West Columbia, SC). Glucose dehydrogenase was purchased from Alfa Aesar (Lancashire, UK). Spectinomycin (Sp) was purchased from Chem-Impex International (Wood

Dale, IL). Formic acid, dithiothreitol (DTT), and perchloric acid were purchased from Honeywell-Fluka (Charlotte, NC). 2-methyl-3-oxopentanoic acid ethyl ester was purchased from Fragmenta (Monmouth Junction, NJ). Mini-PROTEAN TGX 8-16%, and Bio-Rad Protein Assay Dye Reagent concentrate was purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, T4 DNA ligase, and Phusion DNA polymerase were purchased from New England Biolabs (Ipswich, MA). GoTaq Green Master Mix was purchased from Promega (Madison, WI). Deoxynucleotides (dNTPs) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). PD-10 columns Sephadex G-25M was purchased from GE (Chicago, IL). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in water. Amico Ultra 3,000 MWCO, 10,000 MWCO centrifugal concentrators, and 30,000 MWCO centrifugal concentrators were purchased from EMD Millipore (Billerica, MA). High-resolution mass spectral analyses were carried out on an Agilent 6530 QTOF Accurate Mass spectrometer and an Agilent 6460 QQQ.

Bacterial strains. *E. coli* DH10B-T1^R was used for DNA construction. *E. coli* BAP1 [16], and BAP1-T1^R [17] were used for heterologous expression of DEBS modules and ACP proteins as they require phosphopantathiene modification and BL21(DE3)-T1^R was used for the expression of all other proteins. *E. coli* BAP1 and BAP1-T1^R were also used for *in vivo* production.

Gene and plasmid construction. Standard molecular biology techniques were used to carry out plasmid construction. All PCR amplifications were carried out with Phusion High Fidelity DNA polymerase or with Taq DNA polymerase using GoTaqGreen master mix. For amplification of GC-rich sequences, PCR reactions carried out with Phusion High Fidelity DNA polymerase were performed in the GC buffer supplemented with 10% DMSO and 1M betaine with primer annealing temperatures 4-8 °C below the T_m. DNA assembly was performed using the isothermal Gibson assembly protocol [18]. All constructs were verified by sequencing (UC Berkeley DNA Sequencing Facility; Berkeley, CA).

p15A-DszAT. The gene encoding DszAT was amplified from pFW3 [19] with the DszATF1/ T7TerminatorR1 primer pair. The *lac* operon was amplified from the same template with the LacCasetteLacI/LacCasetteLacO primer pair. The PCR products were inserted by Gibson assembly into the ClaI-PacI site of pJA4MCS2, which is pACYC184 [20] modified by replacing Tc^R cassette with a HindIII-XhoI-KpnI-XmaI-PacI-AflII-SFcI-PstI-AscI-AvrII-HincII linker.

p15A-DszATS86A. The S86A mutation was introduced to p15A-DszAT by amplifying p15A-DszAT with two sets of primers: DszATF1/S86AMutationR and S86AMutationF/DszATMutant_CTerm. S86AMutationR and S86AMutationF contain S86A mutation. The two PCR fragments were inserted by Gibson assembly into the NdeI-HindIII site of p15A-DszAT.

p15A-DszATF190V. F190V mutation was introduced to p15A-DszAT by amplifying p15A-DszAT with two sets of primers: DszATF1/F190VMutationR and F190VMutationF/DszATMutant_CTerm. F190VMutationR and F190VMutationF contain S86A mutation. The two PCR fragments were inserted by Gibson assembly into the NdeI-HindIII site of p15A-DszAT.

pET21c-FabD. FabD gene was amplified from *E. coli* BL21(DE3)-T1^R using the MalACP_F1/ MalACP_R1 primer set. The PCR product was inserted into the NdeI-EcoRI site of pET21c by Gibson assembly.

pET16b-His$_{10}$Pres-ACP$_{DEBSMod6}$. The ACP domain of module 6 of DEBS was amplified from pAYC138 [19] using the PresACP6_Fwd3/PresACP6_Rev primer set. The PCR product was inserted by Gibson assembly into the NdeI-BamHI site of a modified pET16b plasmid that has its FactorXa cleavage site replaced by Prescission protease cleavage site.

Expression and purification of His-tagged proteins. The expression and purification of His$_6$-MatB, His$_6$-Epi, DszAT-His$_6$, His$_{10}$-ACP$_{Mod6}$, Mod3$_{DEBS}$+TE(AT^0)-His$_6$, Mod6$_{DEBS}$+TE(AT^0)-His$_6$ were carried out as previously described in Chapter 2. For FabD-His$_6$, the protein was purified from *E. coli* BL21(DE3)-T1^R as previously described in Chapter 2 for DszAT-His$_6$. Purified proteins were flash-frozen in liquid nitrogen before storing at -80 ºC. Protein concentration was determined using A_{280} and protein extinction coefficients calculated from the sequence using the ExPASY ProtParam program [21].

Preparation of substrates. N-acetylcysteamine thioester of (2S,3R)-2-methyl-3-hydroxypentanic acid (NDK-SNAC) and fluoromalonate were synthesized as previously described in Chapter 2.

Growth Curves. Cultures were grown to OD_{600} = ~0.4 at 37 ºC with shaking at 200 rpm, at which point 200 µL of cultures were collect and supplemented with 1 mM IPTG, 0.2% arabinose, and when used 20 mM sodium propionate, 5 mM malonate, or 5 mM fluoromaloante final concentration. Cultures (100 µL) were then transferred to 96-well microplate (clear, polystyrene, U-bottom, sterile, Greiner Bio-One). Plate was incubated with continuous shaking uncovered at room temperature. Growth was monitored in absorbance mode at 600 nm using a Synergy Mx microplate reader with measurements taken every 2.5 min for 16 h. OD_{600} data were corrected for the path length of 100 µL liquid (0.3 cm).

***In vivo* production of triketide lactones in *E. coli*.** Procedure was adapted from previously described method [15]. LB (50 mL) containing appropriate antibiotics (Carbenicilin, Chloramphenicol: 50 µg/mL; Spectinocymin: 100 µg/mL) in a 250 mL baffled flask was inoculated to 2% (1 mL) of an overnight LB culture of *E. coli* BAP1 transformed with the plasmids of interest. Cultures were grown at 37 ºC with shaking at 200 rpm to OD_{600} = 0.4-0.6, at which point cultures were cooled on ice for 20 min and protein expression was induced with 1 mM IPTG and 0.2% (*w/v*) arabinose. The cultures were incubated at 16 ºC with shaking for 16-24 h. Cells (40 mL) were harvested by centrifugation at 1000 × *g* for 15 min at 4 ºC after recording the final OD_{600}. Cells were washed once with 1 vol (40 mL) of M9 medium and resuspended in M9 medium to OD_{600} ~ 100. The high-density cell suspension (50 µL) was added to a 1.7 mL microcentrifuge tube along with 5 mM fluoromalonate (1 M stock in water) and 1 mM NDK-SNAC (250 mM stock in 25% DMSO). Cell suspensions were incubated with the two substrates at 16 ºC with shaking at 200 rpm for 20-24 h. Afterwards, the mixture was centrifuged at 20,817 × *g* for 1 min at 4 ºC to pellet cells. The supernatant (35 µL) was removed and mixed with 10% (*v/v*) perchloric acid (35 µL) to precipitate protein. The acidified supernatant was then centrifuged at 20,817 × *g* for 10 min at 4 ºC and analyzed by LC-MS/MS.

***In vitro* assay for triketide lactone production.** All assays (30 µL) contained 5 µM methylmalonyl-CoA epimerase (Epi), 20 µM MatB, 2.5 mM ATP, 15 U/mL PK, 10 U/mL myokinase, 50 mM phosphoenol pyruvate, 1 mM NDK-SNAC, and 10 µM Mod3$_{DEBS}$+TE(AT^0),

in 400 mM sodium phosphate, 10 mM magnesium chloride, 5 mM TCEP, pH 7.5. For the FabD activity assay, 1 mM of fluoromalonate, malonate, or methylmalonate, 1 mM of coenzyme A , and 30 μM of *trans*-AT enzymes were added. For the *trans*-AT titration assay, 0.2, 1, or 5 mM of fluoromalonate and 1 mM of coenzyme A was used in place of fluoromalonyl-CoA. *Trans*-AT concentrations were 250, 100, 50, 25, 10, 5, 2.5, 1, 0.5, 0.25, 0.1, and 0 μM. The following assay components were pre-incubated at 37 ºC for 30 min before react ion initiation: Epi, MatB, PK, myokinase, PEP, ATP, and acyl-CoA substrate. The reactions were then initiated with NDK-SNAC, DEBS module, and *trans*-AT and incubated for 16-24 h at 37 ºC. To quench the reaction, 25 μL of reaction mixture was mixed with 25 μL of 10% (*v/v*) perchloric acid and centrifuged for 10 min at 4 ºC at 20,817 × *g*. The supernatant was removed and stored at -80 ºC until analysis.

Quantification of triketide lactone products by LC-QQQ MS. Frozen samples were centrifuged for 10 min at 4 ºC at 20,817 × *g* to remove salts. The supernatant was removed and analyzed by LC-QQQ using MRM in positive ionization mode on an Agilent 6460 QQQ MS. Samples were analyzed on two tandem Ascentis Express RP-amide HPLC columns (2.7 μm, 2.1 mm × 10 cm, 25 ºC; Supelco) on an Agilent 1290 UPLC using an isocratic solvent system of 10% acetonitrile in 0.1% (*v/v*) formic acid as the aqueous mobile phase for 20 min after an initial ramp of 0 to 10% acetonitrile over 0.5 min at flow rate of 0.4 mL/min. Transitions for the TKLs and TTKLs products were determined by screening for optimal fragmentation conditions. Products were monitored with transitions as follows (parent ion *m/z* → product ion *m/z*, fragmentor voltage, collision energy): TKL (171 → 153, 90,6); H-TKL (157 → 139, 90, 4); F-TKL (175 → 157, 90, 5). The identity and the quantity of each compound was verified against synthetic standards using external standard curves.

Synthetic standards were obtained as previously described [15]. The concentration of the 2-fluoro-2-desmethyltriketide lactone (F-TKL) synthetic standard was determined by ^{19}F-NMR using the ERETIC method [22] against an external standard of 5.0 mM 5-fluorouracil. The triketide lactone (TKL) standard was prepared enzymatically and the 2-desmethyltriketide lactone (H-TKL) standard was synthesized as described [23]. The concentrations of TKL and H-TKL were determined by ^{1}H-NMR in D_2O using the ERETIC method against an external standard of diethylfluoromalonate (75 mM).

Steady-state kinetic characterization of transacylation activity. The transacylation assay was developed from a previously described method [24-26]. Assays were performed in 96-well microplate (half area, μCLEAR, black polystyrene, medium binding non-sterile, Greiner Bio-One) with NADH fluorescence (λ_{Ex} = 340 nm, λ_{Em} = 450 nm) monitored using a Synergy Mx microplate reader (Biotek; top optical position at 8 mm read height, gain = 100). Reactions were monitored for 5 mins at 30 ºC using minimum interval setting between measurements.

Assay components were prepared in three solutions in 50 mM sodium phosphate, 1 mM EDTA, 10% (*v/v*) glycerol, pH 7.5: Solution 1, 2.4 mU/μL α-ketoglutarate dehydrogenase, 1.6 mM NAD^+, 1.6 mM TPP, and 8 mM α-ketoglutarate; Solution 2, 4× malonyl-CoA, methylmalonyl-CoA, or fluoromalonyl-CoA substrate; Solution 3, 150 μM $ACP_{DEBSMod6}$, 2× *trans*-AT enzyme, and 0.1 mg/mL BSA. Solution 1 (12.5 μL) and Solution 2 (12.5 μL) were added to the wells and pre-incubated at room temperature for 15-20 min to reduce background by allowing any free CoA contaminating the acyl-CoA substrate to be consumed. The reactions were then initiated with Solution 3 (25 μL).

Assays (50 µL) contained 75 µM $ACP_{DEBSMod6}$, 0.6 mU/µL α-ketoglutarate dehydrogenase, 0.4 mM NAD^+, 0.4 mM TPP, 2 mM α-ketoglutarate, and 0.05 mg/mL BSA in 50 mM sodium phosphate, 1 mM EDTA, 10% (*v/v*) glycerol, pH 7.5. The final concentrations of malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA were 200, 150, 100, 50, 25, 12.5, 6.25, and 3.13 mM. The final concentrations for FabD were 30 nM for malonyl-CoA, 50 nM for fluoromalonyl-CoA, and 100 nM for methylmalonyl-CoA. Controls with no *trans*-AT added were run in parallel for all reactions.

Fluorescence was converted to concentration using an NADH standard curve collected in the same assay matrix with acyl-CoA substrate, ACP domain, and DszAT omitted. Initial velocities (v_0) were defined as the linear portion of the change of product concentration over time curve. The data were fit by non-linear curve fitting to the Hill equation with substrate inhibition using Origin 6.0 with initial values estimated by the Solver tool in Microsoft Excel:

$$v_0 = \frac{v_{max}[B]_0^n}{K_{0.5}^n + [B]_0^n(1 + \frac{[B]_0^n}{K_i^n})}$$

where v_0 = initial velocity, V_{max} = maximum velocity, $[B]_0$ = substrate concentration, n = Hill coefficient, $K_{0.5}$ = half-maximal concentration constant, and K_i = inhibition constant.

3.3 Results and Discussion

Design of fluorinated triketide lactone biosynthetic pathway. Previous work in the laboratory has shown that feeding of NDK-SNAC and fluoromalonate to the culture medium of *E. coli* cells expressing $Mod6_{DEBS}$ led to low but detectable production of fluorinated triketide lactone (F-TKL) [15]. Using the complementation system developed in Chapter 2, we hope to improve the production yield of the desired fluorinated product (*Figure 3.1*). As the AT domain of module 6 of DEBS natively utilizes methylmalonyl-CoA and discriminates strongly against malonyl-CoA [27], we believe that it is likely that discrimination against fluoromalonyl-CoA may also occur given that is sterically more similar to malonyl-CoA than methylmalonyl-CoA. Thus the utilization of a more fluoromalonyl-CoA selective *trans*-AT protein should increase productivity of F-TKL and also decrease the amount of H-TKL that is made due to the presence of malonyl-CoA inside the cell.

In addition to utilizing a trans-AT with higher fluorine selectivity, increases in the intracellular concentration of the fluoromalonyl-CoA extender unit can also help to increase the yield of the desired fluorinated product. Since CoA thioesters are not typically transported across the membrane, we plan to use a two-enzyme system consisting of a malonate transporter and malonyl-CoA synthetase, MatB. The malonate transporter is expected to transport fluoromalonate in the culture medium across the cellular membrane into the cytoplasm where the malonyl-CoA synthetase would then produce fluoromalonyl-CoA [28] (*Figure 3.1*).

Thus, we designed a four-gene, three-plasmid system for the biosynthesis of F-TKL in *E. coli*. The system contained pAYC136 or pAYC138, a plasmid from p15A-DszAT series, and a plasmid from pFmal series (*Figure 3.2*). pACYC136 and pACYC138 encode for $Mod3_{DEBS}$+TE(AT^0) and $Mod6_{DEBS}$+TE(AT^0), respectively [19]. These proteins are functionally expressed in *E. coli* BAP1 and should provide the necessary DEBS pathway protein for cellular production for fluorine-

containing polyketide products. The p15A-DszAT series encode for variants of the *trans*-AT protein DszAT under the control of an ITPG-inducible T7 promoter.

The pFmal series, which contain a CloDF13 origin, encode a malonate transporter under an arabinose-inducible pBAD promoter and an ATP-dependent malonyl-CoA synthetase (MatB) from *Rhodopseudomonas palustris* under an ITPG-inducible T7 promoter [28]. pFmal is expected to support intracellular accumulation of fluoromalonyl-CoA. Representatives from three classes of prokaryotic malonate transporters, namely MdcF, MadLM, and MatC, have been identified, cloned, and inserted to pFmal backbone to produce pFmal(MdcF), pFmal(MadLM), and pFmal(MatC) respectively. MdcF is a putative malonate transporter found in an operon with malonate decarboxylase and 7 other genes involved in malonate metabolism in *Krebsiella pneumonia* [29]. MadLM is a characterized two-component sodium:malonate symporter from *Pseudomonas fluorescens* [30]. MatC is a putative malonate transporter found in an operon with malonyl-CoA synthetase and malonyl-CoA decarbozylase enzymes in *Streptomyces* coelicolor [31, 32]. The MatB from *Rhodopseudomonas palustris* has been shown to be able to catalyze the ligation of fluoromalonate to coenzyme A with a large catalytic efficiency of 1,250 ± 80 $M^{-1}s^{-1}$ [28, 33]. pFmal series have previously been shown to support the production of mM-levels of 2-fluoro-3-hydroxybutyrate and poly(2-fluoro-3-hydroxybutyrate) when co-expressed with the NphT7 ketosynthase and PhaB ketoreductase [28]. Thus, we believe that these constructs would also be able to support production of fluorinated triketide lactone due to the high flux of fluoromalonate though the pathway. However, the 2-fluoro-3-hydroxybutyrate pathway likely maintains a higher throughput than the single modular DEBS pathway. Thus, it is possible that a high steady-state concentration of fluoromalonate and fluoromalonyl-CoA would not be cleared in the polyketide production system, potentially leading to significant toxicity effect.

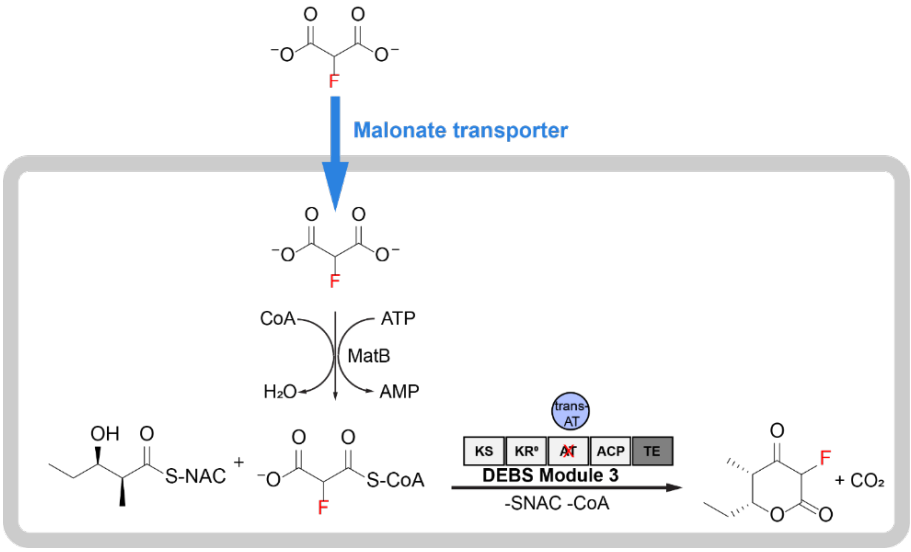


Figure 3.1. *Complementation strategy for fluorinated polyketide production in an E. coli production host. In addition to the polyketide synthase (PKS) and the trans-AT, the cellular production system requires a specialized transport system to import fluoromalonate into the intracellular space and activate to the fluoromalonyl-CoA extender unit. Cells natively maintain a high level of ATP, which provides the necessary ATP regeneration.*

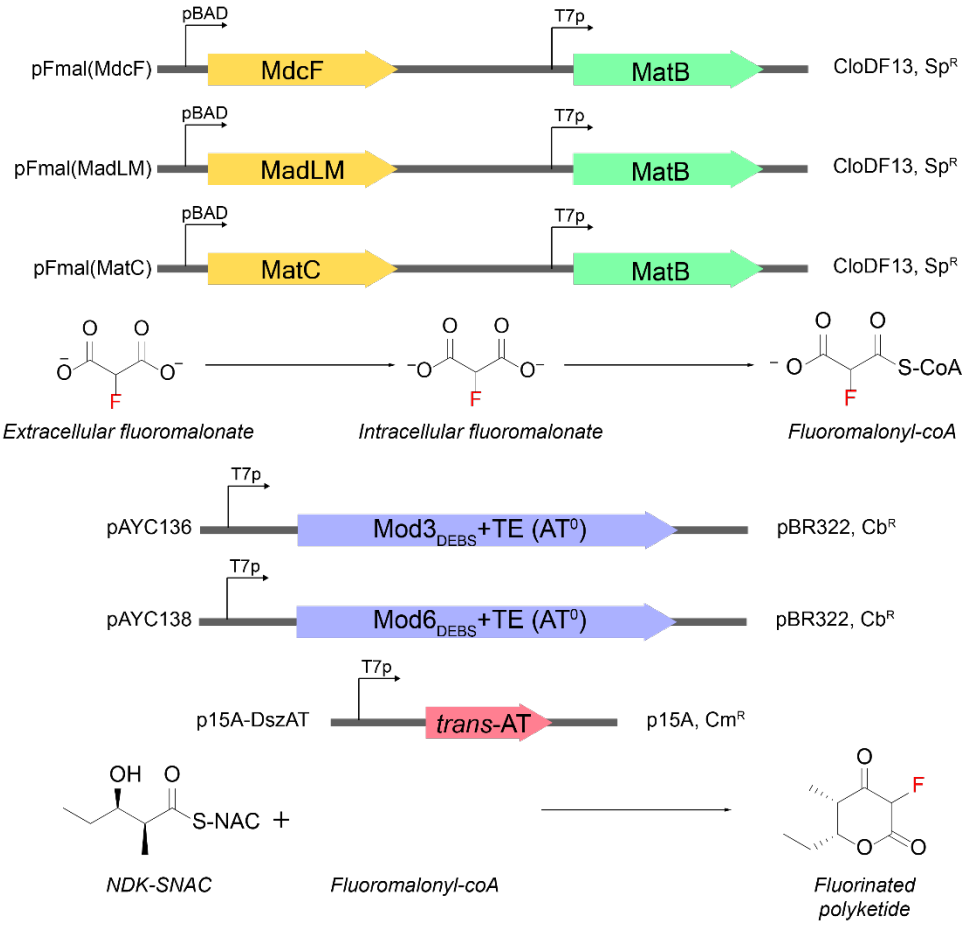


Figure 3.2. *Design of a pathway for fluorinated triketide lactone (F-TKL) production in E. coli. The system includes three plasmids with compatible origins and antibiotic markers, which together encode for four pathway proteins. pFmal (cloDF13 origin, Sp^R) series encode for a malonate transporter and malonyl-CoA synthetase, MatB. pAYC136 and 138 (pBR322 origin, Cb^R) encode for $Mod3_{DEBS}$+TE(AT^0) and $Mod6_{DEBS}$+TE(AT^0), respectively. p15A-DszAT (p15A origin, Cm^R) encodes for the trans-AT variant.*

We measured the growth curves of *E. coli* BAP1 harboring the pFmal plasmid and its derivatives. As these cells have no dedicated downstream pathway to utilize fluoromalonyl-CoA, we reasoned that the effects observed in this strain would represent the highest toxicity level potentially imposed on our system by fluoromalonate feeding. We found growth defects when many components of the fluoromalonyl-CoA system was present (*Figure 3.3)*. First, we observed that the expression of MadLM alone led to decreased cell growth rates, even in the absence of malonate or fluoromalonate. Since toxicity of transmembrane membrane proteins is commonly observed, we reasoned that high expression levels of MadLM might disturb the membrane integrity of the cells, leading to decreased cell growth. When MadLM or both MadLM and MatB were present, feeding fluoromalonate decreased cell growth. In contrast, no observable growth defect was found when cells were fed with fluoromalonate in the absence of MadLM or MatB. These results are consistent with a model where MadLM and MatB together with fluoromalonate enable accumulation of fluoromalonyl-CoA inside cells. However, it suggests that fluoromalonyl-CoA has some toxicity in the absence of a dedicated downstream pathway for its utilization. For

example, strains co-expressing of NphT7 and PhaB with MadLM and MatB show no evidence of a growth defect [28].

We then examined protein expression of *E. coli* BAP1 bearing the three-plasmid system by SDS-PAGE with Coomasie staining. In the protein gels, we were able to observe expression of all pathway proteins besides the malonate transporter, which was not unexpected since the expression of transmembrane protein is typically lower based on their toxicity (*Figure 3.4*). We noted that the expression of MatB when co-expressed with other pathway proteins was much lower than when it is expressed alone, which may mitigate its toxicity with regard to fluoromalonyl-CoA production. Thus, we decided to proceed with production experiments without further modification to the pFmal series.

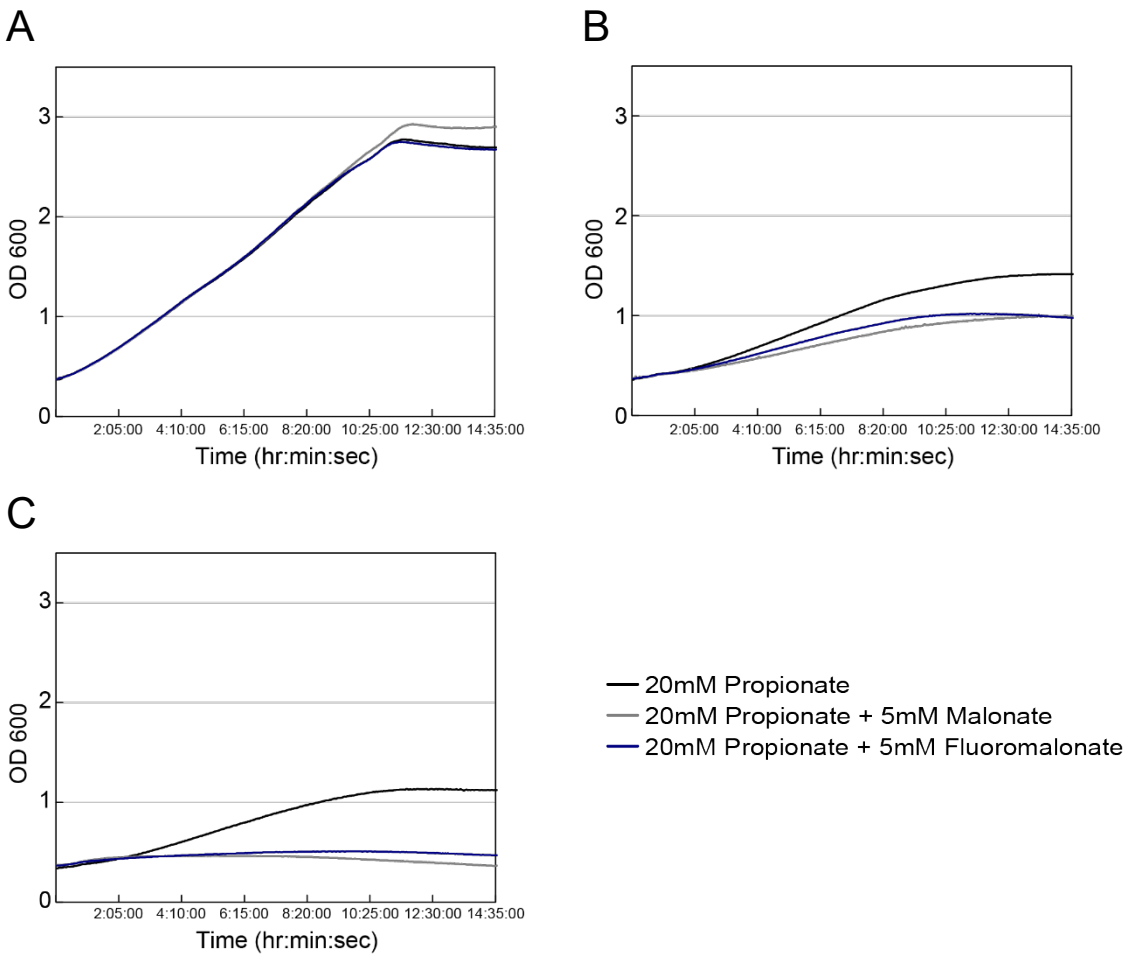


Figure 3.3. *Growth curves of E. coli BAP1 expressing various pFmal plasmids. E. coli BAP1 harboring (A) no plasmid control, (B) pFmal(MadLM, without MatB) [28] , or (C) pFmal(MadLM) were grown to OD_{600} = ~0.4 at 37 ºC with shaking at 200 rpm, at which point 100 µL of cultures were transferred to 96-well plate and supplemented with 1 mM IPTG, 0.2% arabinose, and when used 20 mM sodium propionate, 5 mM malonate, or 5 mM fluoromaloante. Growth was monitored at OD_{600} by microplate reader. Data are representative of 3 biological replicates.*

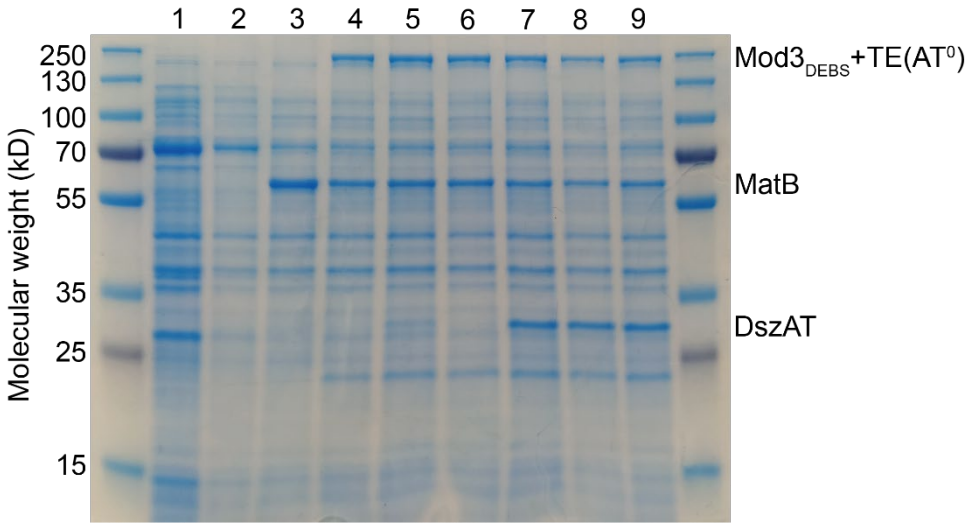


Figure 3.4. *SDS-PAGE analysis of E. coli cultures harboring the triple plasmid system for fluorinated triketide lactone production. Cells were grown at 37 ºC until OD_{600} = 0.6-0.8, at which time protein expression was induced by addition of IPTG and arabinose and cells were continued to grow at 16 ºC. Samples of E. coli BAP1 harboring (lane1) no plasmid control, (lane2), pFmal(MadLM, without MatB)[28], (lane3) pFmal(MadLM),(lane 4-6) pAYC136 and pFmal(MadLM), or (lane 7-9) pAYC136, pFmal(MadLM), and p15A-DszAT, were collected 20 h after protein expression was induced and analyzed by SDS-PAGE with Coomassie staining. Comparison between cells harboring pFmal with (lane 3) and without (lane2) MatB gene allows the MatB band to be identified. Staining is lower for this band in the strain expressing the full pathway (lane 7-9) compared to that expressing only pFmal(MadLM) (lane 3).*

Malonate transporters increased levels of fluorinated triketide lactone product. We developed an in vivo assay for F-TKL production using a concentrated cell suspension of E. coli BAP1 harboring the three-plasmid system for expression of a malonate transporter, MatB, trans-AT, and either $Mod3_{DEBS}$+TE(AT^0) or $Mod6_{DEBS}$+TE(AT^0). The cells were first induced for 16-24 h, concentrated to OD_{600} ~ 100 in M9, and fed NDK-SNAC and fluoromalonate. The resulting polyketide products were extracted and analyzed by LC-QQQ. Using this system, we compared the effect of the three different malonate transporters to when no transporter was expressed. We found that all three malonate transporters were able to increase the production titer of F-TKL (Figure 3.5). Similar results were obtained from both the $Mod3_{DEBS}$+TE(AT^0) and $Mod6_{DEBS}$+TE(AT^0) pathways, although overall productivity with $Mod3_{DEBS}$+TE(AT^0) is higher. MdcF co-expression increased F-TKL yield from 1- to 4-fold; however, both MatC and MadLM co-expression led to greater increases in F-TKL production up to 10-fold. Notably, under these conditions, H-TKL was observed at levels significantly below that of F-TKL and almost no detectable amount of TKL was observed in the production. These observations are consistent with the intrinsic intracellular pool of alternative extender units: E. coli maintains a very limited amount of malonyl-CoA (35µM) [34] and almost no methylmalonayl-CoA [35]. As MatC and MadLM were able to increase the F-TKL yield to similar level, but MadLM had been shown to be more versatile in terms of induction conditions [28], we decided to proceed with MadLM as the fluoromalonate transporter in our pathway.

We performed a titration experiment to determine the optimum concentration of fed substrates for production by altering both the amount of fluoromalonate and NDK-SNAC added to the cell suspension (*Figure 3.6*). We observed that F-TKL yield was strongly dependent on NDK-SNAC concentration, while fluoromalonate concentration seemed to only contribute significantly at low NDK-SNAC levels. We therefore chose assay conditions of 5 mM fluoromalonate and 1 mM

NDK-SNAC, which allows us to maintain good yields at more conservative NDK-SNAC feeding levels.

Assessing the role of DszAT in fluorinated triketide lactone *in vivo* production. Although we observe high selectivity of F-TKL production *in vivo* using the DszAT complementation, WT DszAT still strongly prefers malonyl-CoA. Therefore, we wanted to distinguish the impact of DszAT from the increase in intracellular fluoromalonyl-CoA resulting from MadLM and MatB co-expression, which could also be a source of the selective production of F-TKL. We therefore compared the effect of empty vector, an inactive DszAT variant (S86A DszAT) and F190V DszAT, which is shown *in vitro* to suppress malonyl-CoA incorporation (*Figure 2.12-2.13*), to WT DszAT. In this experiment, we found that F-TKL was produced in comparable yields under all production conditions for both $Mod3_{DEBS}$+TE(AT^0) or $Mod6_{DEBS}$+TE(AT^0) (*Figure 3.7A*), indicating that F-TKL production was independent of the presence of a functional DszAT. Further control experiments showed that DszAT was still being expressed in this system as the overexpression band was observed by SDS-PAGE (*Figure 3.7B*), suggesting that selective F-TKL production is predominantly caused by the high intracellular concentration of fluoromalonyl-CoA in this system.

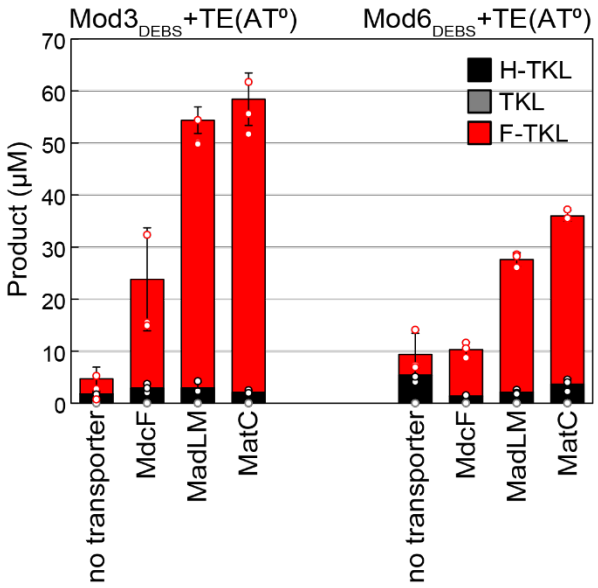


Figure 3.5. *Screening malonate transporters for in vivo F-TKL production. E. coli BAP1 cells were transformed with 1) pAYC136 (for $Mod3_{DEBS}$+TE(AT^0) expression, left) or pAYC138 (for $Mod3_{DEBS}$+TE(AT^0) expression, right); 2) p15A-DszAT, and 3) pFmal(mdcF), pFmal(MadLM), or pFmal(MatC). Concentrated cell suspension of E. coli BAP1 expressing $Mod3_{DEBS}$+TE(AT^0) (left) or $Mod6_{DEBS}$+TE(AT^0) (right), WT DszAT, MatB, and malonate transporter MdcF, MadLM, or MatC were provided with 1 mM NDK-SNAC and 5 mM fluoromalonate. Culture media were collected and analyzed by LC-QQQ after 24 h. Data are mean ± s.d. of biological replicates (n=3).*

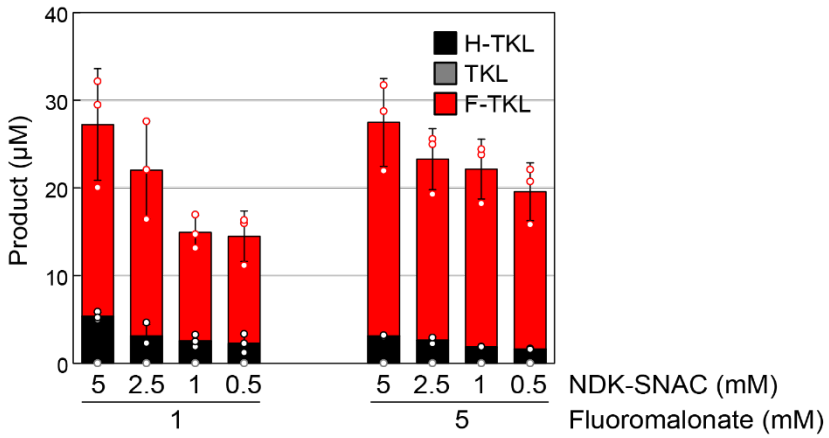


Figure 3.6. *Titration of NDS-SNAC and fluoromalonate for in vivo F-TKL production. E. coli BAP1 expressing Mod3$_{DEBS}$+TE(AT^0), WT DszAT, MatB, and malonate transporter MadLM were provided with varying concentrations of NDK-SNAC (0.5-5 mM) and fluoromalonate (1-5 mM). Culture media were collected and analyzed by LC-QQQ after 24 h. Data are mean ± s.d. of biological replicates (n=3).*

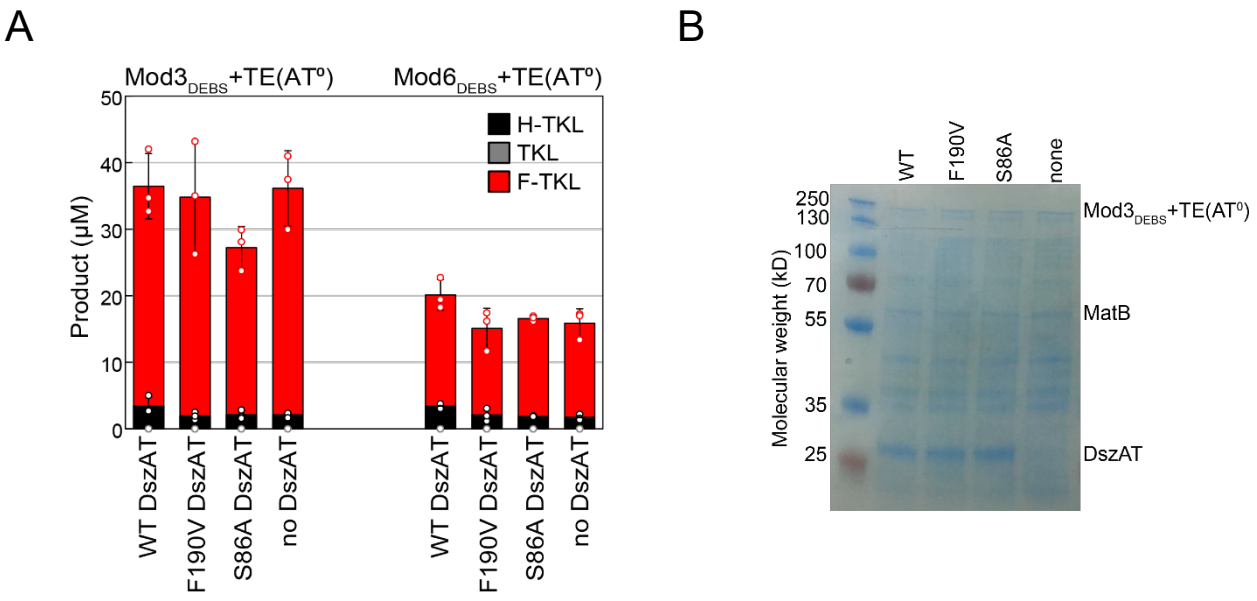


Figure 3.7. *DszAT-dependence for in vivo F-TKL production. (A) Triketide lactone production by E. coli BAP1 expressing Mod3$_{DEBS}$+TE(AT^0) (left) or Mod6$_{DEBS}$+TE(AT^0) (right), MatB, MadLM, and a trans-AT variant, namely WT DszAT, F190V DszAT, S86A DszAT, or no DszAT. Concentrated cell suspensions were provided with 1 mM NDK-SNAC and 5 mM fluoromalonate. Culture media were collected and analyzed by LC-QQQ after 24 h. Data are mean ± s.d. of biological replicates (n=3). (B) SDS-PAGE of E. coli production cultures showing expression of trans-ATs, along with Mod3$_{DEBS}$+TE(AT^0) and MatB.*

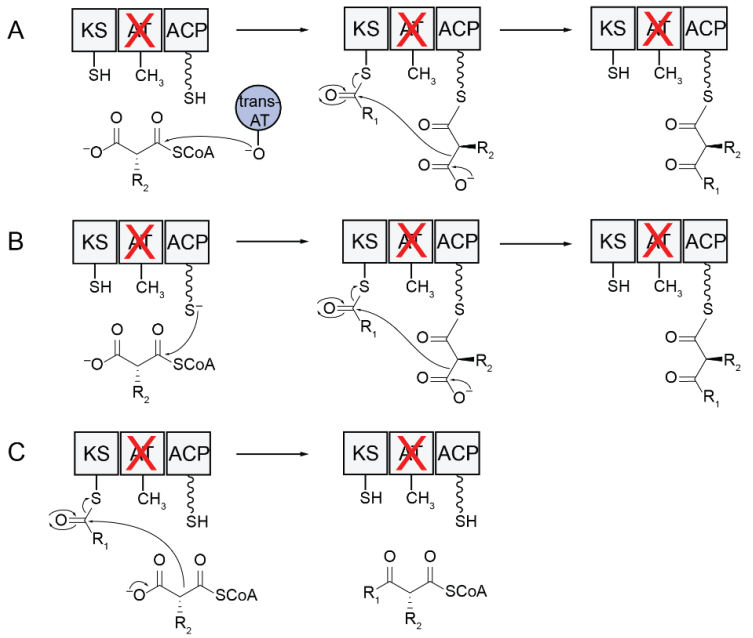


Figure 3.8. *Different mechanisms for cis-AT-independent chain elongation in modular PKS. (A) A canonical reaction cycle requires transacylation of the ACP phosphopantetheine group with a carboxyacyl-CoA substrate. In modules where the cis-AT is inactivated (AT^0), this can be achieved in trans with a trans-AT. (B) Self-malonylation could occur by chemical transacylation. Chemical acylation of the ACP phosphopantetheine group could also be accelerated with fluoromalonyl-CoA given its intrinsic reactivity. (C) The carboxyacyl-CoA extender unit can also be utilized directly by the KS domain, bypassing the formation of a carboxyacyl-ACP intermediate.*

Previous *in vitro* experiments showed that a functional *trans*-AT was necessary for complementing the AT^0 mutations of Mod3$_{DEBS}$+TE(AT^0) or Mod6$_{DEBS}$+TE(AT^0), since there are no other active DEBS *cis*-ATs to provide the required enzymatic activity in the single-module DEBS systems. One hypothesis that could explain this observation is that F-TKL can be made through an AT-independent mechanism. Unlike in the canonical mechanism in which an active AT domain is necessary for acylation of the ACP domain (*Figure 3.8A*), at least two other alternative mechanisms exist through which F-TKL could be produced independent of a functional AT domain. In one case, the activated nature of the fluoromalonyl-CoA extender unit lends itself to direct chemical acylation onto the ACP phosphopantathiene arm without assistance from an AT (*Figure 3.8B*). This intermediate can then undergo normal KS-catalyzed carbon-carbon bond formation to form product. Another possibility is an ACP- and AT-independent mechanism where the KS domain can condense the diketide substrate directly with the fluoromalonyl-CoA extender, bypassing the canonical covalently-tethered intermediate (*Figure 3.8C*) [36]. While self-malonylation (*Figure 3.8B*) has been reported for Type II PKS and Type II FAS ACPs [37, 38], it has been shown that this mechanism is both extremely slow (100-1000-fold defect) compared to the AT-dependent reactions with native substrates and also requires an acidic residue at the end of the Helix I, which was missing from ACP$_{DEBSMod3}$ and ACP$_{DEBSMod6}$ [38, 39]. Thus, we sought to differentiate between the canonical (*Figure 3.8A*) and the direct extender unit usage mechanism (*Figure 3.8B*), since formation of the tethered intermediate is important for continued chain elongation and direct extender unit usage would lead to chain termination of a growing polyketide

The canonical mechanism forming a tethered intermediate requires a functional ACP domain whereas the direct extender unit usage mechanism bypasses this domain. To differentiate between

these two mechanisms, we compared the production yields of *E. coli* cultures harboring Mod3$_{DEBS}$+TE (AT^0) or Mod3$_{DEBS}$+TE (AT^0 ACP^0) [36], which contain a functional or non-functional ACP domain, respectively. In these experiments, we found that the production of F-TKL was greatly reduced in the absence of a functional ACP domain (*Figure 3.9).* This result demonstrates that the majority of F-TKL is indeed produced through an ACP-dependent mechanism. Given the lack of dependence of F-TKL production on DszAT expression, we hypothesized that another endogenous host protein provided the necessary enzymatic activity to complement the AT^0 mutation on the DEBS module at a sufficient level to obviate the need for a heterologously-expressed *trans*-AT such as DszAT.

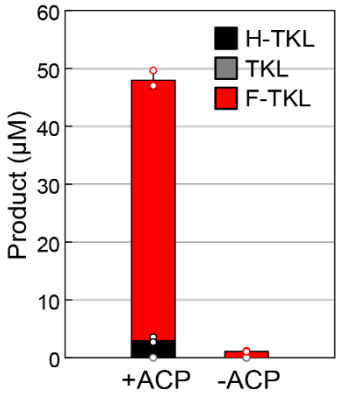


Figure 3.9. *Assessing the mechanism of F-TKL production. E. coli BAP1 expressing a Mod3$_{DEBS}$+TE variant with either an active (AT^0) or inactive (AT^0 ACP^0) ACP domain, WT DszAT, MatB, and MadLM were provided with 1 mM NDK-SNAC and 5 mM fluoromalonate. Culture media were collected and analyzed by LC-QQQ after 24 h. Data are mean ± s.d. of biological replicates (n=3).*

An endogenous *E. coli* protein complements the DEBS AT^0 mutation in F-TKL production. We next sought to identify endogenous *E. coli* proteins that could be candidates for supplementing the necessary *trans*-AT activity to our engineered PKS system as it would dominate the extender unit selectivity in our system. We therefore examined the genes annotated in the genome of *Escherichia coli* BL21(DE3) (NCBI Taxid ID 469008) as "acyltransferase" or "transacylase", which returned 13 proteins, of which 6 are unique genes (*Table 3.1*). From the list, the malonyl-CoA:ACP transacylase (GenBank AKN47232.1) stood out as it carried out the same enzymatic reaction as needed to complement the DEBS AT^0 mutations. Further investigation revealed that the malonyl-CoA:ACP transacylase is FabD, which functions in the fatty acid synthase pathway to load malonyl-CoA extender units onto AcpP for the elongation of straight-chain fatty acids. Given the promiscuity with fluoromalonyl-CoA that we have observed enzymes that natively utilize malonyl-CoA as a substrate, we hypothesized that FabD may be able to load the fluoromalonyl-CoA extender unit onto the ACP domain of DEBS modules through inter-molecular protein-protein interactions. In fact, FabD proteins from various *Streptomyces* spp. have been shown to function in Type II polyketide synthase pathways by loading the malonyl-CoA extender unit onto PKS ACP domains [40-42].

Table 3.1 *Acyltransferases and transcylases annotated in the E. coli BL21 (DE3) genome.*

Protein product	Region name	Length	GenBank	GI	Genomic Location
acyl transferase	PRK10191	162	AKN48056.1	850015287	complement(CP011938.1:2030501..2030989)
acyl transferase	WcaF	182	AKN48052.1	850015283	complement(CP011938.1:2026774..2027322)
malonyl-CoA-ACP transacylase	FabD	309	CAQ31613.1	242376895	AM946981.2:1152289..1153218
malonyl-CoA-[acyl-carrier-protein] transacylase	FabD	309	ACT42983.1	253977313	CP001509.3:1152295..1153224
predicted colanic acid biosynthesis acyl transferase	WcaB	162	CAQ32470.1	242377709	complement(AM946981.2:2031068..2031556)
predicted acyl transferase	WcaF	182	CAQ32466.1	242377705	complement(AM946981.2:2027341..2027889)
predicted acyl transferase	CaiE	196	CAQ30555.1	242375875	complement(AM946981.2:38976..39566)
putative acyl transferase	WcaB	162	ACT43815.1	253978145	complement(CP001509.3:2031074..2031562)
putative acyl transferase	WcaF	182	ACT43811.1	253978141	complement(CP001509.3:2027347..2027895)
malonyl CoA-ACP transacylase	FabD	309	AKN47232.1	850014463	CP011938.1:1151722..1152651
ACP S-malonyltransferase	FabD	309	WP_000191372.1	446113517	n/a
colanic acid biosynthesis acetyltransferase WcaB	WcaB	162	WP_000888740.1	446811484	n/a
glycerol-3-phosphate 1-O-acyltransferase PlsB	PlsB	807	WP_000017354.1	445939499	n/a

Both FabD and DszAT proteins belong to the αβ-hydrolase family. Sequence analysis of FabD revealed only 40% identity to DszAT (*Figure 3.10A*). FabD also only exhibits 25-30% sequence identity to other DEBS ATs (Module 1 AT = 25.75%; Module 2 AT = 30.2%; Module 3 AT = 26.81%; Module 4 AT =27.76%; Module 5 AT = 30.24%; Module 6 AT = 30%). The two proteins, however, share invariant residues that are essential for their functions [43, 44], namely (FabD/DszAT numbering) S92/S86, H201/H191, Q11/Q9, R117/R111. S92/S86 and H201/H191/Q9 represent the catalytic dyad for transacylation, while main-chain amide of Q11 forms oxyanion hole. R117/R11 is found at the base of the active site and interacts with β-carboxy group of enzyme substrate. Further structural analysis reveals significant homology between FabD and DszAT (*Figure 3.10B*). Overlaying the crystal structures of FabD (PDB ID 2G2Y) [43] and DszAT (PDB ID 6APG) [44] both bound to malonate shows that the proteins have a similar overall fold (rmsd = 1.7 Å) with both containing two subdomains with the active site in between. The small subdomain contains 4 β-sheets while the large subdomain consists of 10-12 α-helices. The active sites of the proteins also seem to have similar architecture with (FabD/DszAT numbering) S92/S86, H201/H191, Q11/Q9, R117/R111 overlaying well. Moreover, the substrate malonate is also bound in a similar orientation. The substrate pocket is largely polar. In both structures, malonate is bound by salt bridge interactions between its β-carboxy group and the side chain amine groups of R117/R111, positioning the carbonyl carbon near the active serine residues, which are found in nearly identical position in both proteins, for nucleophilic attack. Notably, in DszAT, F190 is found positioned near the Cα of malonate that may prevent binding of substrate with bulkier substituent such as a methyl group at the α-position. In FabD, this residue is replaced with S200, which may provide more flexibility in terms of substrate binding.

Interactions between AT and ACP domains has previously been shown to rely on a small number of charged amino acid residues located near the active sites of the two protein domains [19, 45]. R278 and R279 of DszAT interact with E70 of its native protein partner, DSZS ACP1. Similarly, K180 of DszAT interacts with D45 of DSZS ACP1 [45]. Sequence alignment shows that these positively-charged interface residues of DszAT, R278, R279, and K180, correspond to K276, R277, and R190 of FabD, maintaining the charge at the interface for interaction with an ACP partner (*Figure 3.10*). This finding suggests that FabD could potentially interact with a similar set of ACP partners as DszAT, including the ACP domains of DEBS.

A

```
FabD   1   MTQFAFVFPGQGSQTVGMLADMAASYPIVEETFAEASAALGYDLWALTQQGPAEELNKTW  60
DszAT  1   --MKAYMFPGQGSQAKGMGRALFDAFPALT---ARADGVLGYSIRALCQDDPDQRLSQTQ  55
               *::********: **   :  ::* :    * *...***.:.** *:.* : *.:*.

FabD   61  QTQPALLTASVALYRVWQQQGGKAPAMMAGHSLGEYSALVCAGVIDFADAVRLVEMRGKF  120
DszAT  56  FTQPALYVVNALSY-LKRREEEAPPDFLAGHSLGEFSALFAAGVFDFETGLALVKKRGEL  114
            ***** ....  * :...:   .* ::*******:***..***:**  .: **: **::

FabD   121 MQEAVPEGTGAMAAIIGLDDASIAKACEEAAEGQVVSPVNFNSPGQVVIAGHKEAVERAG  180
DszAT  115 MGDA---RGGGMAAVIGLDEERVRELLDQNG-ATAVDIANLNSPSQVVISGAKDEIARLQ  170
           * :*     *.***:****:  : :  :: . . .*. .*:***.****:* *: : *

FabD   181 AACKAAGAKRALPLPVSVPSHCALMKPAADKLAVELAKITFNAPTVPVVNNVDVK-CETN  239
DszAT  171 VPFEAAGAKKYTVLRVSAAFHSRFMRPAMVEFGRFLEGYDFAPPKIPVISNVTARPCKAD  230
           .. :*****.   * **.. *. :*.**  ::.  *    * .*.:**:.** .. *:::

FabD   240 GDAIRDALVRQLYNPVQWTKSVEYMAAQGVEHLYEVGPGKVLTGLTKRIVDTLTASALNE  299
DszAT  231 G--IRAALSEQIASPVRWCESIRYLMGRGVEEFVECGHGIVLTGLYAQI--------RRD  280
           *  ** **  *: .**.* :*: *: ..*** : * * * *****  .*         .:

FabD   300 PSAMAAALEL                                                  309
DszAT  281 AQPLVVA---                                                  287
           ...:..*
```

B

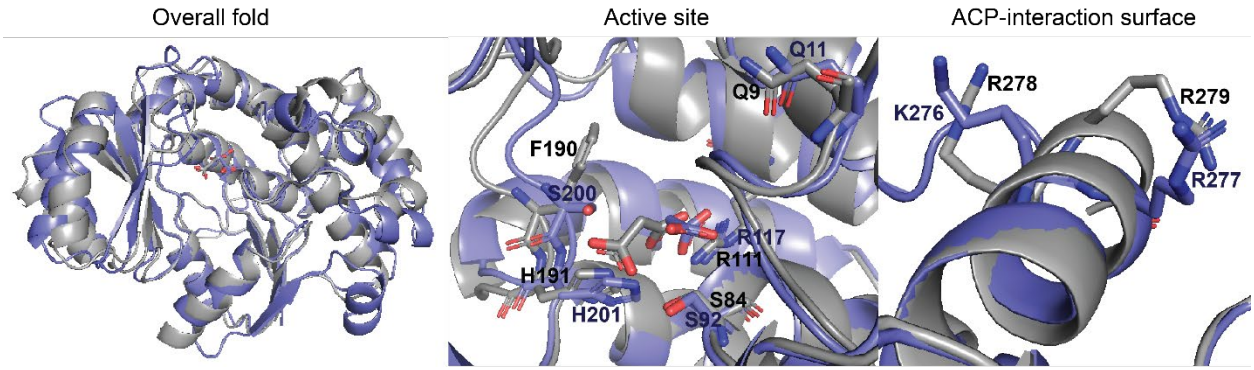


Figure 3.10 *Sequence and structural alignment between FabD and DszAT. (A) Sequence alignment of FabD and DszAT. Conserved functional residues (red), specificity-determining residue (blue), and ACP-interacting residues (green) identified from biochemical experiments are highlighted. (B) Structural comparison between FabD (blue) and DszAT (grey). The two proteins have the same overall fold (left), active site architecture (middle), and ACP-interaction surface (right). Malonate is bound to both proteins in a similar orientation. Active site, specificity-determining, and interface residues are shown in sticks for FabD (blue) and DszAT (grey).*

FabD accepts fluoromalonyl-coA. To investigate the ability of FabD to support F-TKL production, FabD was cloned, overexpressed, purified to homogeneity (*Figure 3.11*). Next, FabD was tested for its competence to produce triketide lactone *in vitro* with $Mod3_{DEBS}$+TE (AT^0). Using this assay, we found that FabD was able to produce F-TKL and H-TKL at a level comparable to DszAT when fluoromalonate and malonate were respectively provided (*Figure 3.12*). Interestingly, FabD is also able to produce TKL at a much higher level than DszAT (*Figure 3.12*). This result is somewhat surprising given its native role in straight-chain fatty acid biosynthesis.

To further characterize the activity of FabD with respect to polyketide formation, we examined the steady-state kinetic behavior of FabD with respect to transacylation of $ACP_{DEBSMod6}$ using the α-ketoglutarate dehydrogenase coupled assay [25]. In these assays, we found that FabD is active on fluoromalonyl-CoA and malonyl-CoA extender units with similar catalytic efficiency to that of DszAT (*Figure 3.13*). Moreover, FabD was also found to be active on the methylmalonyl-CoA extender unit. However, unlike DszAT, which discriminates against methylmalonyl-CoA with nearly 10^3-fold lower catalytic efficiency, FabD accepts methylmalonyl-CoA with comparable kinetic parameters to fluoromalonyl-CoA (*Figure 3.13*). This lack of discrimination is unexpected given the protein's native role in fatty acid synthesis, although it might be a result of the absence of methylmalonyl-CoA in the native metabolite pool of *E. coli.* This result is consistent with our triketide lactone assay data that showed similar levels of activity of FabD and DszAT on fluoromalonyl-CoA and malonyl-CoA, but increased level of FabD on methylmalonyl-CoA.

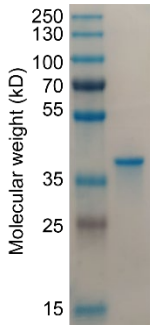


Figure 3.11 *SDS-PAGE of purified FabD.*

***In vivo* selectivity of *E. coli* FabD for triketide lactone production.** With information on the activity and the selectivity of FabD in hand, we aimed to determine its role in controlling the identity of the extender unit used *in vivo* for triketide lactone production. The *in vitro* studies confirmed that FabD is competent to interact with ACPs from DEBS and can complement AT^0 modules. Given its ability to accept all three extender units, FabD could affect the production of more complex polyketide targets that require the cellular availability of multiple extender units. Thus, we set out to investigate the selectivity of our cellular production platform by providing the cultures with alternative or mixed extender units, assuming that the distribution of triketide lactone product species would reflect the selectivity of the system.

In cells without DszAT, F-TKL is the dominant product when cells were provided with fluoromalonate (*Figure 3.14*). A similar result was found with malonate, where H-TKL becomes the major product. However, when cells were provided with methylmalonate, a mixture of TKL and H-TKL were produced with both $Mod3_{DEBS}$+TE (AT^0) and $Mod6_{DEBS}$+TE (AT^0), suggesting that methylmalonate is not able to suppress the production of H-TKL to the same extent as fluoromalonate. Under mixed extender unit conditions in which *E. coli* cultures were provided with equimolar fluoromalonate and either malonate or methylmalonate, we found that the product profile follows the precursor profile. More specifically, when cultures were provided with fluoromalonate and malonate, F-TKL and H-TKL were the major products; whereas, when cultures were provided with fluoromalonate and methylmalonate, F-TKL and TKL became the major products. This result suggests that the selectivity of the system *in vivo* is relaxed, likely due to the possibility of complementation of the DEBS AT^0 mutation by FabD, which is accepting of all three substrates; thus, the outcome seems mainly governed by the provided precursors.

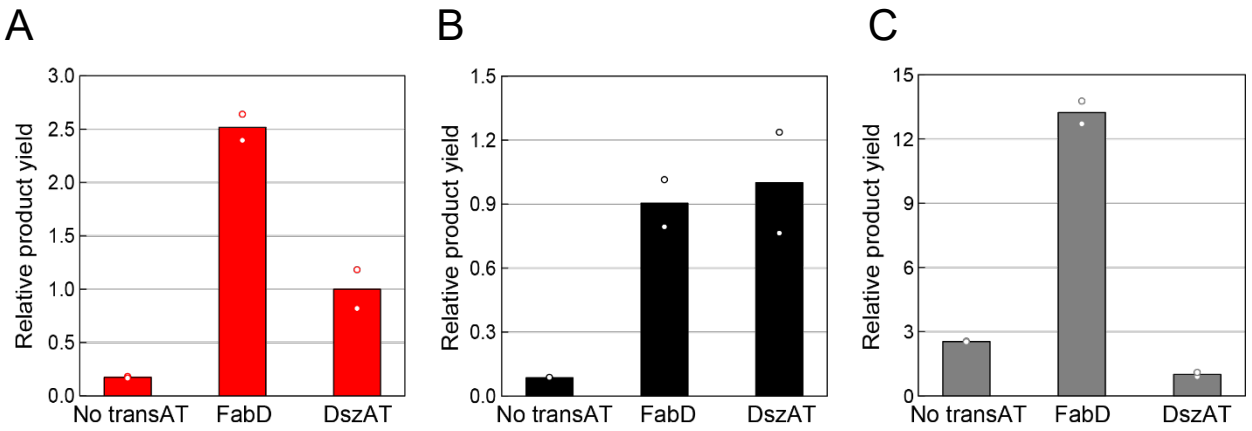
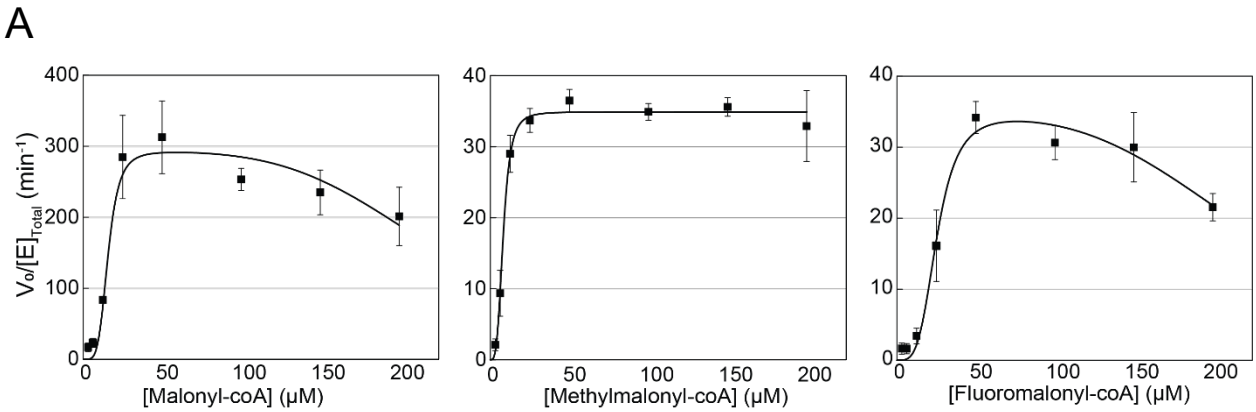


Figure 3.12. *In vitro triketide lactone assay for FabD complementation of Mod3$_{DEBS}$+TE(AT^{0}). Normalized representation of the amount of (A) F-TKL, (B) H-TKL, and (C) TKL products are shown for reactions containing 10 μM Mod3$_{DEBS}$+TE(AT^{0})), 30 uM FabD or WT DszAT, 1 mM NDK-SNAC, 1 mM (A) fluoromalonate, (B) malonate, or (C) methylmalonate, and 1 mM coenzyme A under regenerative condition. Reactions were quenched and analyzed by LC-QQQ after 24 h. Data are mean of technical replicates (n=2).*



B

Substrate	k_{cat} (min^{-1})	$K_{0.5}$ (μM)	$k_{cat}/K_{0.5}$ (min^{-1} μM^{-1})	K_i (μM)	n
Malonyl-CoA	293 ± 20	15 ± 1.5	20 ± 3.3	230 ± 19	4.3 ± 1.5
Methylmalonyl-CoA	36 ± 0.90	8.3 ± 0.42	4.2 ± 0.33	429 ± 92	3.4 ± 0.46
Fluoromalonyl-CoA	35 ± 2.7	25 ± 1.9	1.4 ± 0.21	230 ± 15	3.8 ± 1.2

Figure 3.13. *Steady-state kinetic characterization of ACP transacylation catalyzed by FabD with malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA. (A) Transacylation of ACP$_{DEBSMod6}$ with malonyl-CoA (left), methylmalonyl-coA (center), and fluoromalonyl-CoA (right) by FabD. Activity was measured using a αKGDH-coupled kinetic assay for measuring ACP$_{DEBSMod6}$ transacylation with different carboxyacyl-CoA extender units using NADH fluorescence. Data were fit to Hill equation with substrate inhibition mode.l Data points are average ± s.d. of technical replicates (n=3). (B) Table contains k_{cat}, $K_{0.5}$, $k_{cat}/K_{0.5}$, Ki, and n calculated by non-linear curve fitting to the Hill equation with substrate inhibition. Error in k_{cat}/K is obtained by propagation from the individual kinetic terms. Data are mean ± s.e. of technical replicates (n=3).*

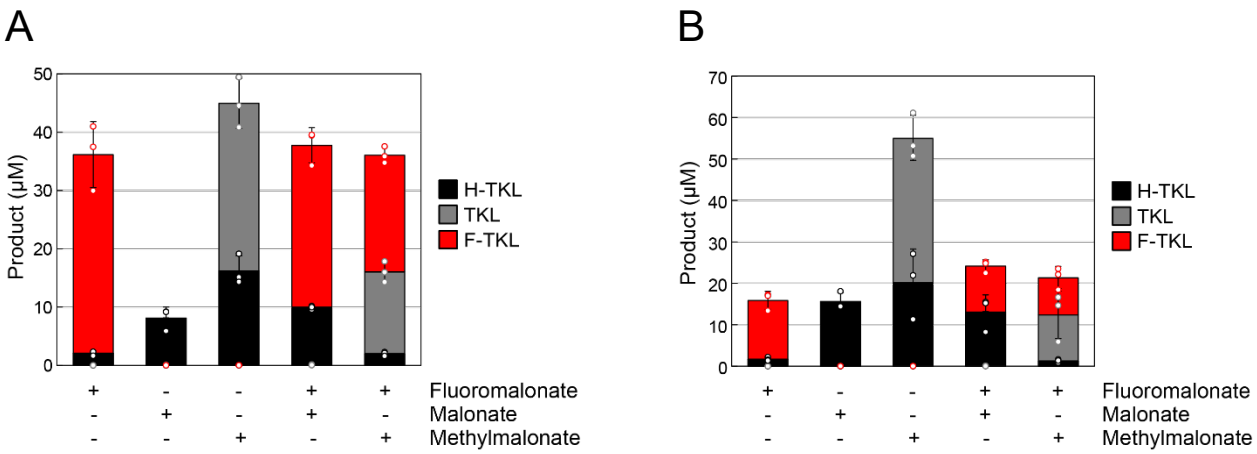


Figure 3.14. *Influence of extender unit availability on in vivo selectivity of chain elongation by single-modular DEBS constructs in engineered E. coli. Concentrated cell suspension of E. coli BAP1 expressing (A) Mod3$_{DEBS}$+TE(AT^0) or (B) Mod6$_{DEBS}$+TE(AT^0), MatB, and MadLM were provided with 1 mM NDK-SNAC and 5 mM fluoromalonate, malonate, and/or methylmalonate and analyzed by LC-QQQ after 24 h. Data are mean ± s.d. of biological replicates (n=3).*

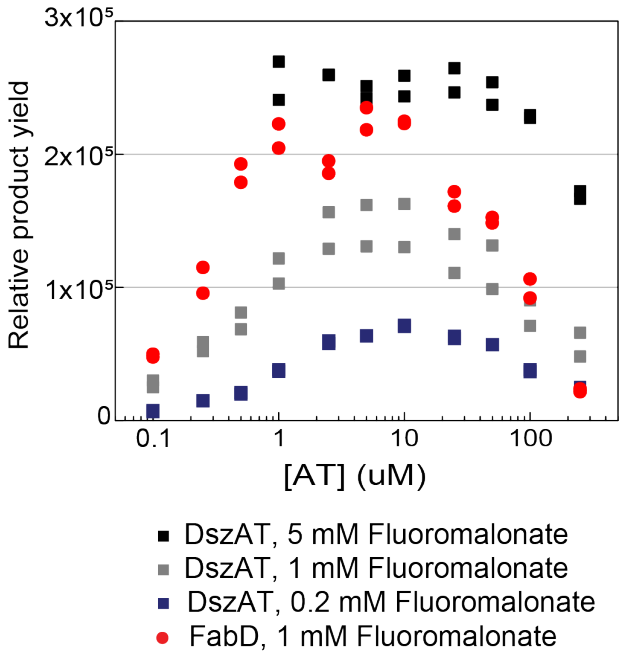


Figure 3.15. *Effect of trans-AT concentration on in vitro production of F-TKL. Relative amount of F-TKL are shown for reactions containing 10 μM DEBSMod3/AT0+TE; 0.1, 0.25, 0.5, 1, 2.5, 5, 10, 25, 50, 100, or 250 μM WT DszAT (square) or FabD (circle); 20 μM MatB; 0.2 (blue), 1 (grey, red), or 5 (black) mM fluoromalonate; 1 mM coenzyme A, and ATP regeneration system. Reactions were quenched after overnight incubation at 37 ºC and analyzed by LC-QQQ.*

Consequently, extender unit ratios will need to be optimized by feeding, transport, and activation separately for different PKS systems, depending on their differential usage of malonyl-CoA, methylmalonyl-CoA, and fluoromalonyl-CoA.

Effect of *trans*-AT on triketide production *in vitro*. In order to begin assessing whether increasing the overexpression levels of DszAT could allow it to outcompete endogenous FabD for DEBS ACP domains, we decided to carry out a titration of DszAT and FabD using the *in vitro* triketide production assay. Not surprisingly, we found that when the ratio between *trans*-AT: PKS is less than 1, increasing the amount of *trans*-AT allows for concomitant increases of F-TKL production (*Figure 3.15*). However, as the ratio of *trans*-AT:PKS is increased above 1, F-TKL production is suppressed. Similar trends were observed for both DszAT and FabD. Upon further examination of the effect of increased fluoromalonate, we observed that we could increase the overall production yield of F-TKL (*Figure 3.15*).

Based on these data, we hypothesized that when *trans*-AT is sub-stoichiometric compared to PKS module, PKS modules poised for transacylation remain unbound to the *trans*-AT. Thus, increasing the relative concentration of *trans*-AT protein led to increased ratio of bound PKS modules and a concomitant increase in F-TKL production. However, if the *trans*-AT levels become too high, it could remain bound to the PKS module after transacylation in an unproductive complex that prevents turnover of the PKS module for another catalytic cycle.

These results emphasize the importance of careful tuning of the relative levels of pathway proteins in the production host. It is also consistent with our finding that overexpression of DszAT in production cultures did not increase F-TKL yield, given that the host could utilize endogenous FabD to carry out DEBS ACP transacylation. Indeed, overexpressed DszAT may even shift the system into the regime where excess *trans*-AT could be inhibitory as DszAT is expressed at a significantly higher level than the DEBS proteins. Moreover, overexpressing DszAT possibly lowers the amount of DEBS module protein inside the cell due to expression limit.

3.4 Conclusions

We have developed an *in vivo* model system that can amplify the production of a simple fluorine-containing polyketide molecule to mg/L levels. The fluoromalonate transporters, MadLM and MatC, along the malonate:CoA ligase MatB supports the formation of sufficient intracellular concentration of fluoromalonyl-CoA extender unit for use by PKS pathways, leading to an increase in F-TKL yield. Cell growth experiments show that high levels of fluoromalonyl-CoA can cause growth defects, but the accumulation was not high enough with the downstream PKS expressed to be probematic. Experiments with different DzAT variants suggest that the relative ratio of carboxyacyl-CoA extender units in the cell mainly control the identity of the extender unit used to complement the AT^0 mutation on the PKS module. Studies suggest that one reason for this is that an endogenous *E. coli* acyltransferase protein, FabD, may present a similar ACP interaction surface as DszAT and thus be able to compete with DszAT under *in vivo* production conditions for DEBS complementation. This hypothesis is consistent with the observation of TKL, which arises from chain extension with methylmalonyl-CoA, even when this substrate is likely excluded by DszAT under cellular conditions with a 10^3-defect in catalytic efficiency. Despite FabD's role in straight-chain fatty acid synthesis in the host, *in vitro* transacylation studies show that it is able to transacylate $ACP_{DEBSMod6}$ with catalytic efficiency of only 5-fold lower than the native substrate,

malonyl-CoA. so role in polyketide production. However, it remains to be investigated whether FabD would be a factor in the complement of a broad range of PKS beyond DEBS.

3.5 References

1. Li, J.; Zhang, L.; Liu, W. Cell-free synthetic biology for in vitro biosynthesis of pharmaceutical natural products. *Synthetic and Systems Biotechnology* **2018***, 3*, 83-89.

2. Sattely, E. S.; Fischbach, M. A.; Walsh, C. T. Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Nat. Prod. Rep.* **2008***, 25*, 757-793.

3. Oßwald, C.; Zipf, G.; Schmidt, G.; Maier, J.; Bernauer, H. S.; Müller, R.; Wenzel, S. C. Modular construction of a functional artificial epothilone polyketide pathway. *ACS Synth. Biol.* **2014***, 3*, 759-772.

4. Tan, G.; Deng, K.; Liu, X.; Tao, H.; Chang, Y.; Chen, J.; Chen, K.; Sheng, Z.; Deng, Z.; Liu, T. Heterologous biosynthesis of spinosad: an omics-guided large polyketide synthase gene cluster reconstitution in streptomyces. *ACS Synth. Biol.* **2017***, 6*, 995-1005.

5. Zhang, J.; Wang, H.; Liu, X.; Hu, C.; Zou, Y. Heterologous and engineered biosynthesis of nematocidal polyketide–nonribosomal peptide hybrid macrolactone from extreme thermophilic Fungi. *J. Am. Chem. Soc.* **2020***, 142*, 1957-1965.

6. Morgado, G.; Gerngross, D.; Roberts, T. M.; Panke, S. In *synthetic biology for cell-free biosynthesis: fundamentals of designing novel in vitro multi-enzyme reaction networks;* Zhao, H., Zeng, A., Eds.; Synthetic Biology – Metabolic Engineering; Springer International Publishing: Cham, 2018; pp 117-146.

7. Chao, R.; Yuan, Y.; Zhao, H. Recent advances in DNA assembly technologies. *FEMS Yeast Res.* **2015***, 15*, 1-9.

8. Zhang, J. J.; Tang, X.; Moore, B. S. Genetic platforms for heterologous expression of microbial natural products. *Nat. Prod. Rep.* **2019***, 36*, 1313-1332.

9. Pfeifer, B. A.; Khosla, C. Biosynthesis of polyketides in heterologous hosts. *Microbiol Mol Biol Rev* **2001***, 65*, 106-118.

10. Bilyk, O.; Luzhetskyy, A. In *host organism: streptomyces;* Industrial Biotechnology; John Wiley & Sons, Ltd: 2016; pp 487-504.

11. Gao, X.; Wang, P.; Tang, Y. Engineered polyketide biosynthesis and biocatalysis in Escherichia coli. *Appl Microbiol Biotechnol* **2010***, 88*, 1233-1242.

12. O, H.; Kh, L.; W, V.; Tt, E.; Am, C. Assembly of standardized DNA parts using BioBrick ends in E. coli. *Methods Mol Biol* **2012***, 852*, 61-76.

13. Ehmoser-Sinner, E.; Tan, C. D. In *The Tools;* Ehmoser-Sinner, E., Tan, C. D., Eds.; Lessons on synthetic bioarchitectures: interaction of living matter with synthetic structural analogues; Springer International Publishing: Cham, 2018; pp 41-56.

14. Nielsen, J.; Keasling, J. Engineering cellular metabolism. *Cell* **2016***, 164*, 1185.

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

16. 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*, 1790-1792.

17. Dong, H.; Liffland, S.; Hillmyer, M. A.; Chang, M. C. Y. Engineering in vivo production of α-branched polyesters. *J. Am. Chem. Soc.* **2019***, 141*, 16877-16883.

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

19. 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*, 95-102.

20. Chang, A. C.; Cohen, S. N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* **1978***, 134*, 1141-1156.

21. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. In *protein identification and analysis tools on the ExPASy server;* Walker, J. M., Ed.; The proteomics protocols handbook; Humana Press: New Jersey, 2005; pp 571-607.

22. Akoka, S.; Barantin, L.; Trierweiler, M. Concentration measurement by proton NMR using the ERETIC method. *Anal. Chem.* **1999***, 71*, 2554-2557.

23. Hinterding, K.; Singhanat, S.; Oberer, L. Stereoselective synthesis of polyketide fragments using a novel intramolecular Claisen-like condensation/reduction sequence. *Tetrahedron Letters* **2001***, 42*, 8463-8465.

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

25. 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*, 3796-3806.

26. Molnos, J.; Gardiner, R.; Dale, G. E.; Lange, R. A continuous coupled enzyme assay for bacterial malonyl-CoA:acyl carrier protein transacylase (FabD). *Anal. Biochem.* **2003***, 319*, 171-176.

27. Khosla, C.; Tang, Y.; Chen, A. Y.; Schnarr, N. A.; Cane, D. E. Structure and mechanism of the 6-Deoxyerythronolide B synthase. *Annual Review of Biochemistry* **2007***, 76*, 195-221.

28. Thuronyi, B. W.; Privalsky, T. M.; Chang, M. C. Y. Engineered fluorine metabolism and fluoropolymer production in living cells. *Angewandte Chemie International Edition* **2017***, 56*, 13637-13640.

29. Hoenke, S.; Schmid, M.; Dimroth, P. Sequence of a gene cluster from Klebsiella pneumoniae encoding malonate decarboxylase and expression of the enzyme in Escherichia coli. *Eur. J. Biochem.* **1997***, 246*, 530-538.

30. Schaffitzel, C.; Berg, M.; Dimroth, P.; Pos, K. M. Identification of an Na+-dependent malonate transporter of Malonomonas rubra and its dependence on two separate genes. *J. Bacteriol.* **1998***, 180*, 2689-2693.

31. Leonard, E.; Yan, Y.; Fowler, Z. L.; Li, Z.; Lim, C.; Lim, K.; Koffas, M. A. G. Strain improvement of recombinant Escherichia coli for efficient production of plant flavonoids. *Mol. Pharm.* **2008***, 5*, 257-265.

32. Park, S. R.; Ahn, M. S.; Han, A. R.; Park, J. W.; Yoon, Y. J. Enhanced flavonoid production in Streptomyces venezuelae via metabolic engineering. *J. Microbiol. Biotechnol.* **2011***, 21*, 1143-1146.

33. Ha, C.; Kc, R.; I, R.; Jc, E. Structure-guided expansion of the substrate range of methylmalonyl coenzyme A synthetase (MatB) of Rhodopseudomonas palustris. *Appl Environ Microbiol* **2012***, 78*, 6619-6629.

34. Rabinowitz, J. D.; Osterhout, R.; Bennett, B. D.; Gao, M.; Kimball, E. H.; Van Dien, S. J. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nature Chemical Biology* **2009***, 5*, 593-599.

35. Haller, T.; Buckel, T.; Rétey, J.; Gerlt, J. A. Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by Escherichia coli. *Biochemistry* **2000***, 39*, 4622-4629.

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

37. Arthur, C. J.; Szafranska, A.; Evans, S. E.; Findlow, S. C.; Burston, S. G.; Owen, P.; Clark-Lewis, I.; Simpson, T. J.; Crosby, J.; Crump, M. P. Self-malonylation is an intrinsic property

of a chemically synthesized type II polyketide synthase acyl carrier protein, *Biochemistry* **2005***, 44*, 15414-15421.

38. Misra, A.; Sharma, S. K.; Surolia, N.; Surolia, A. Self-acylation properties of type II fatty acid biosynthesis acyl carrier protein. *Chem. Biol.* **2007***, 14*, 775-783.

39. Hitchman, T. S.; Crosby, J.; Byrom, K. J.; Cox, R. J.; Simpson, T. J. Catalytic self-acylation of type II polyketide synthase acyl carrier proteins. *Chem. Biol.* **1998***, 5*, 35-47.

40. Izumikawa, M.; Meluzzi, D.; Moore, B. S.; Cheng, Q.; Xiang, L. Enzymatic total synthesis of enterocin polyketides. *Nature Chemical Biology* **2007***, 3*, 557-558.

41. Florova, G.; Kazanina, G.; Reynolds, K. A. Enzymes involved in fatty acid and polyketide biosynthesis in Streptomyces glaucescens: Role of fabH and fabD and their acyl carrier protein specificity. *Biochemistry* **2002***, 41*, 10462-10471.

42. Carreras, C. W.; Khosla, C. Purification and in vitro reconstitution of the essential protein components of an aromatic polyketide synthase. *Biochemistry* **1998***, 37*, 2084-2088.

43. Oefner, C.; Schulz, H.; D'Arcy, A.; Dale, G. E. Mapping the active site of Escherichia coli malonyl-CoA-acyl carrier protein transacylase (FabD) by protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2006***, 62*, 613-618.

44. Mathews, I. I.; Allison, K.; Robbins, T.; Lyubimov, A. Y.; Uervirojnangkoorn, M.; Brunger, A. T.; Khosla, C.; DeMirci, H.; McPhillips, S. E.; Hollenbeck, M.; Soltis, M.; Cohen, A. E. The conformational flexibility of the acyltransferase from the disorazole polyketide synthase is revealed by an X-ray free-electron laser using a room-temperature sample delivery method for serial crystallography. *Biochemistry* **2017***, 56*, 4751-4756.

45. Kosol, S.; Jenner, M.; Lewandowski, J. R.; Challis, G. L. Protein–protein interactions in trans-AT polyketide synthases. *Nat. Prod. Rep.* **2018***, 35*, 1097-1109.

Chapter 4: *Cellular production of fluorine-containing polyketide pharmaceuticals*

Portions of this work were performed in collaboration with the following persons: Swetha Ramesh assisted with constructing and optimizing production system for 6dEB analogs in E. coli and streptomycetes. Dr. Hongjun Dong aided in constructing the production system for 6dEB analogs in E. coli. Dr. Edward Baidoo, Dr. Bashar Amer, and Nicholas Zill, assisted with mass spectrometric and metabolomic analysis. Navid Boozarpour assisted with constructing E. coli BAP1 ΔfadR.

4.1 Introduction

In this chapter, we describe efforts to expand our *in vivo* system for site-selective fluorine incorporation from single-module PKS systems to full-length PKSs to produce increasingly complex molecules using both *E. coli* and model streptomycetes as hosts. In the previous chapter, we had introduced a single-module PKS with inactivating mutation in AT domain (AT^0) into an *E. coli* production host and showed that the expected F-TKL product was produced as the major polyketide product. Here, we aim to extend this work with single modules to a full-length PKS that can biosynthesize molecules with the structural complexity used in pharmaceuticals.

In transferring this platform to multi-modular system, we must address multiple major challenges. One issue is that multi-modular PKS enzymes are extremely large (>200kD in size) and likely would push the limits of the protein expression power of our host organism; however, as there are examples of PKS pathways reconstituted in *E. coli* host, we believe this task is achievable [1, 2]. Large proteins are encoded by long DNA sequences and bring with them challenging cloning tasks. We expect that to introduce active site mutations into the AT domains of multi-modular PKS enzymes would involve advance DNA manipulation techniques [3-5].

Furthermore, to assemble large, multi-modular polyketide molecules requires many types of extender units to be simultaneously present. Most modular polyketide synthases are constructed from a variety of extender units that are differentially selected by different modules due to the selectivity of their AT domains [6, 7]. In order to incorporate fluorine functional group into polyketides, we had devised a system to accumulate and incorporate fluoromalonyl-CoA. However, to biosynthesize fluorine-containing complex polyketide molecules that require multiple chain elongation steps, other extender units need to also be available and would compete with fluoromalonyl-CoA for selection by the *trans*-AT. Our previous work has shown that the identity of the incorporated extender unit depends both on the relative concentrations of the various available extender units as well as the selectivity of the *trans*-AT.

In addition to cellular extender unit availability and *trans*-AT selectivity, a multi-modular PKS contains many functional *cis*-AT domains. Given the sequence conservation between modules and the observation that the multi-modular PKS pathway DEBS containing an AT^0 mutation maintains the ability produce its native product 6dEB [6], it is believed that other functional *cis*-ATs within the PKS can complement a module without a functional AT domain. In the context of our complementation strategy, our trans-AT would need to be able to outcompete other functional *cis*-AT domains within the synthase for transacylation of the AT^0 module, giving rise to a non-fluorinated product.

We decided to initiate these efforts with the DEBS PKS as it is one of the most well characterized PKS systems and we have already had success with *in vitro* complementation of the full-length system and *in vivo* complementation of single-module DEBS models. In both of these platforms, we have shown that fluoromalonyl-CoA can be incorporated into the 6dEB or triketide lactone product through DszAT complementation of the AT^0 module. The complete DEBS pathway has been reconstituted in *E. coli* and optimized for 6dEB production yield. The DEBS pathway was originally reconstituted in *E. coli* BAP1 on two plasmids utilizing the T7 promoter to drive heterologous expression: pBP144 encoding for DEBS1, and pBP130 encoding for DEBS2 and DEBS3 [8]. In addition, pBP130 also contains an operon encoding the propionyl-CoA carboxylase subunits, *pccAB*, which together function to provide the extender unit methylmalonyl-CoA from propionyl-CoA [8]. To date, both the pathway plasmids and the host have undergone

multiple iterations of re-design and engineering to optimize the yield of 6dEB product under different conditions [9-11]

Despite the advantages that *E. coli* provides, we also wanted to explore other host systems. As previously mentioned, *Streptomyces* spp. are often chosen as heterologous expression hosts for PKS pathways due to their close phylogenetic relationship to many native producers. In fact, more PKS pathways have been reconstituted in streptomycetes than in *E. coli*, including the DEBS pathway [12-14]. Streptomycetes provide many advantages over *E. coli* host. Among them is the secondary metabolism that occurs after the primary growth phase allows the organism to reach its optimal population density. During secondary metabolites, cells are not dividing and could devote their resources to biosynthesizing defense molecules such as polyketide products [15]. Secondary metabolism is characterized by an increased diversity of the metabolites pool and allows production to be uncoupled from growth. Furthermore, since streptomycetes naturally produce both malonyl-CoA and methylmalonyl-CoA, they may have evolved mechanisms to separate these pools and may not have as many challenges for crosstalk between these extender units compared to *E. coli*. Thus, we have also initiated studies to explore the possibility of developing a model streptomycetes as a host for fluorine incorporation.

4.2 Materials and Methods

Commercial Materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), tryptic soy broth, nutrient broth, magnesium sulfate anhydrous, and glycerol were purchased from EMD Biosciences (Damstadt, Germany). Carbenicillin (Cb), glucose, isopropyl-β-D-thiogalactopyranoside (IPTG), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), lithium acetate, sodium chloride, sodium acetate, calcium chloride, sucrose, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, ethyl acetate, ethylene diamine tetraacetic acid disodium dihydrate (EDTA), sodium phosphate monobasic, potassium phosphate monobasic, ammonium chloride, calcium carbonate, arabinose, yeast extract, were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A trilithium sodium salt (CoA), malonyl-CoA, methylmalonyl-CoA, malonic acid, methylmalonic acid, tris(2-carboxyethyl)phosphine hydrochloride (TCEP), adenosine trisodium phosphate sodium salt (ATP), β-mercaptoethanol, sodium phosphate dibasic heptahydrate acetonitrile, dimethyl sulfoxide (DMSO), sodium propionate, ammonium formate, chloramphenicol, PEG 3350, sorbitol, phenol:chloroform:isoamyl alcohol, 2,4-diaminopimelic acid, mannitol, soy bean flour, and Triton X-100 were purchased from Sigma-Aldrich (St. Louis, MO). Diethylfluoromalonate was purchased from Sigma-Aldrich (St. Louise, MO) or from Matrix Scientific (West Columbia, SC). Spectinomycin was purchased from Chem-Impex International (Wood Dale, IL). Formic acid, perchloric acid were purchased from Fluka (Charlotte, NC). 2-methyl-3-oxopentanoic acid ethyl ester was purchased from Fragmenta (Monmouth Junction, NJ). Sodium dodecyl sulfate (SDS), Mini-PROTEAN TGX 8-16%, and Bio-Rad protein assay dye reagent concentrate was purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, T4 DNA ligase, and Phusion DNA polymerase were purchased from New England Biolabs (Ipswich, MA). GoTaq Green Master Mix was purchased from Promega (Madison, WI). Deoxynucleotides (dNTPs) and salmon sperm DNA were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Bacto agar, peptone, casamino acid, and yeast nitrogen base without amino acids were

purchased from Becton, Dickinson and Company (Franklin Lakes, NJ). Malt extract was purchased from Merck (Kenilworth, NJ). SC powders were purchased from Sunrise science products (Knoxville, TN). 2xYT was purchased from HiMedia (West Chester, PA). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in water. DNA purification kits, Large Construct Kit, and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). Zymolyase, Frozen-EZ Yeast Transformation II Kit, and Zymoclean Large Fragment DNA Recovery Kit were Zymoresearch (Tustin, CA). Amico Ultra 3,000 MWCO concentrators were purchased from EMD Millipore (Billerica, MA). Chloroform-d was purchased from Cambridge Isotope Laboratories (Andover, MA). ^{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. High-resolution mass spectral analyses were carried out on either an Agilent 1290 connected to a 6530 QTOF Accurate Mass spectrometer or an Agilent 1290 connected to a 6460 QQQ mass spectrometer (Santa Clara, CA).

Bacterial and fungal strains. *E. coli* DH10B-T1^{R} was used for DNA construction. *E. coli* BAP1 [8], BAP1-T1^{R} [16], TB3 [17], and LF01 [18] were used for production. *E. coli* WM3064 was used for conjugal transfer of DNA into *Streptomyces albus* J1074. *Saccharomyces cerevisiae* VL6-48 [19] was used for transformation-associated recombination and gene editing. *Streptomyces albus* J1074 [20] and *Streptomyces lividans* K4-114 [21] were used for production.

Gene and plasmid construction. Standard molecular biology techniques were used to carry out plasmid construction. All PCR amplifications were carried out with Phusion High Fidelity DNA polymerase or with Taq DNA polymerase using GoTaqGreen master mix. For amplification of GC-rich sequences, PCR reactions carried out with Phusion High Fidelity DNA polymerase were performed in the GC buffer supplemented with 10% DMSO and 1 M betaine with primer annealing temperatures 4-8 °C below the T_m. DNA assembly was performed using the isothermal Gibson assembly protocol [22]. For analysis and isolation of large DNA fragments from agarose gel, 0.3-0.6 % gel and zymoclean large fragment DNA recovery kit (Zymoresearch) were used. All constructs were verified by sequencing (Sanger sequencing, UC Berkeley DNA sequencing facility, Berkeley; next generation sequencing, MGH CCIB DNA core, Cambridge, MA).

pBP(JW)130, pBP(JW)144 [8] and pDEBS [18] were gifts from the laboratory of Professor Blaine Pfeifer at the University at Buffalo. pCK7 [23] was a gift from the laboratory of Professor Chaitan Khosla at Stanford University. pCas, pTargetF [24], and pCap01 [25] were purchased from Addgene. pCas_PPhe-BsaI_NAT [26] was a gift from the laboratory of Professor Jamie Cate at the University of California, Berkeley.

pBP144_DEBS1(Mod1 AT^{0}). pBP144_DEBS1(Mod1 AT^{0}) encodes for DEBS1 with S1181A mutation. The plasmid was cloned by first replacing the section to be mutated from the parent plasmid with a Cm^{R} marker and then removing it upon insertion of the piece bearing the desired mutation. This strategy allows for identification of mutants using antibiotic selection. The Cm^{R} cassette was amplified with primers Cm-pT7-PhaEC-F/Cm-pT7-PhaEC-R from pT7-CapPhaEC [16] and inserted into the BsiWI-BstBI site of BP144 via Gibson assembly. *E. coli* containing resulting plasmids were selected on culture medium containing carbenicillin and chloramphenicol. Plasmids containing both resistance markers were further screened by restriction analysis by XhoI. Plasmids with the correct restriction pattern were confirmed by diagnostic Sanger sequencing. Segments of DEBS Module 1 were then amplified from pBP144 by two primer sets, DEBS1-1/DEBS1-2, and DEBS1-3/DEBS1-4. DEBS1-2 and DEBS1-3 contained the desired S1181A

mutation to be introduced to the active site of the AT domain of Module 1. The Cm^R-containing pBP144 intermediate was digested with BsiWI and BstBI to remove the Cm^R insert, which was replaced with the mutated Module 1 segments by Gibson assembly to generate a scarless construct with the AT^0 mutation in Module 1. The resulting transformants were screened for Cm sensitivity followed by restriction analysis of the isolated plasmids by XhoI. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing to ensure that no other changes had been introduced.

pBP130_DEBS2(Mod3 AT^0). pBP130_DEBS2(Mod3 AT^0) encodes for DEBS2 with S653A mutation. The Cm^R cassette was amplified with primers pACYC184_CmOperon_ pBP130_M3_Fwd_PacI/pACYC184_CmOperon_pBP130_M3_Rev_SpeI from pACYC184 and inserted into the MauBI-BaeI site of pBP130 using Gibson assembly. *E. coli* containing resulting plasmids were selected on culture medium containing carbenicillin and chloramphenicol. A segment of DEBS Module 3 was amplified with primers pET21-M3-RP/pET21-M3-FP2 from pAYC136 to introduce the desired S653A into the AT domain of DEBS Module 3. The Cm^R-containing pBP130 intermediate was digested with PacI and SpeI to remove the CmR cassette, which was replacred with the mutated Module 3 segments by Gibson assembly to generate a scarless construct with the AT^0 mutation in Module 3. The resulting transformants were screened for Cm sensitivity followed by restriction analysis of the isolated plasmids by XhoI. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing to ensure that no other changes had been introduced.

pBP130_Mod 5(AT^0). The segment of DEBS Module 5 containing the desired S642A mutation was amplified with primers M5_BbvCI/ M5_NsiI from pFW100_DEBS3(Mod5 AT^0) (***Section 2.2***) and gel purified (3.5 kB). pBP130 was digested with BbvCI and NsiI and the resulting 22 kB band was gel purified to provide the backbone. The backbone and insert were then ligated with T4 ligase. The plasmid candidates were screened by restriction analysis with NotI. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing to ensure that no other changes had been introduced.

pBP130_Mod 6(AT^0). pFW100_DEBS3(Mod6 AT0) (*Section 2.2*) was digested with AsiSI and BbvCI and the resulting 6 kB fragment containing the desired S2107A mutation was gel purified. pBP130 was digested with AsiSI and BbvCI and the resulting 20 kB band was gel purified to provide the backbone. The backbone and insert were then ligated with T4 ligase. The plasmid candidates were screened by restriction analysis with XhoI. Plasmids with the correct restriction patterns were confirmed by complete plasmid sequencing to ensure that no other changes had been introduced.

Preparation of fluoromalonate. Fluoromalonate were synthesized from diethylfluoromalonate as previously described in Chapter 2 [27].

Production of 6dEB analogs in *E. coli*. 6dEB analogs were produced in *E. coli* BAP1 or BAP1-$T1^R$ cultures grown in growth medium previously optimized for 6dEB production [28]. The 6dEB production medium (50 mL) containing appropriate antibiotics in a 250 mL baffled flask was inoculated with 1 mL (2% volume) of overnight culture of *E. coli* harboring appropriate plasmids in the same medium. Cultures were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point cultures were cooled on ice for 20-40 min before inducing protein expression (1 mM IPTG). Cultures were incubated at 22 ºC with shaking at 250 rpm for 16-24 h. Cells were then collected by centrifugation at 10,000 × *g* for 5 min at 4 ºC and resuspended in 1.5

mL of fresh media containing 1mM IPTG, 0.2% (*w/v*) arabinose, 5 mM fluoromalonate, 20 mM propionate, and appropriate antibiotics. High-density cultures were grown in 14-mL round-bottom Falcon test tubes for another 16-24 h at 22 ºC with shaking at 250 rpm, after which they were collected by centrifugation at 20,817 × *g* for 1 min 4ºC.

Supernatant (1.5 mL) was removed and extracted with 4 × 0.5 mL ethyl acetate by vortexing for 15-30 s and separating the layers by centrifugation at 13,000 × *g* for 3 min. The combined organic layers were dried by Speed Vac SC110 (Savant), resuspended in 30 µL of methanol, and analyzed by LC-QTOF or LC-QQQ. Alternatively, for semi-quantitative analysis, supernatant was acidified by perchloric acid to 5% (*v/v*) final concentration to precipitate proteins. Acidified culture medium was centrifuged 20,817 × *g* for 10 min at 4 ºC and the supernatant removed for analysis.

For large scale production of 2-fluoro-2-desmethyl 6dEB analog, the 6dEB production medium (12 × 500 mL) containing appropriate antibiotics in a 12 × 2.5 L ultrayield flask was inoculated with 5 mL (1% volume) of overnight culture of *E. coli* harboring pBP130_DEBS3(Mod6 AT^0), pBP144, and pFmal(MadLM) in the same medium. Cultures were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point cultures were cooled on ice for 20-40 min before inducing protein expression with 1 mM IPTG. Cultures were incubated at 22 ºC with shaking at 250 rpm for 16-24 h. Cells were then collected by centrifugation at 8,000 × *g* for 5 min at 4 ºC and each culture was resuspended in 20 mL of fresh media containing 1mM IPTG, 0.2% (*w/v*) arabinose, 5 mM fluoromalonate, 20 mM propionate, and appropriate antibiotics. High-density cell suspension was transferred to 250 mL baffled flask and incubated for another 16-24 h at 22 ºC with shaking at 250 rpm. Following incubation, cells were harvested by centrifugation at 8,000 × *g* for 5 min and resuspended in fresh media containing substrates, inducers, and antibiotics. Supernatant was collected for extraction.

Supernatant (40 mL) was extracted with 4 × 10 mL ethyl acetate by vortexing for 15-30 s and separating the layers by centrifugation at 8,000 × *g* for 5 min. The combined organic layers were dried by rotary evaporation.

Identification of 6dEB analogs. Products were analyzed by LC-QTOF high-resolution mass spectrometry in positive ionization mode on at Agilent 6530B Q-TOF MS. Samples were analyzed on Ascentis Express RP-amide HPLC columns (2.7 µm, 2.1 mm × 10 cm, 25 ºC; Supelco) on an Agilent 1290 UPLC using a linear gradient from 5 to 65% acetonitrile over 16.75 min with 0.1% (*v/v*) formic acid as the aqueous mobile phase after an initial increase of 0 to 5% acetonitrile over 30 s at flow rate of 0.25 mL/min. Fragmentation analysis was performed using the targeted MS/MS mode with fragmentation voltage of 185 and collision energy of 10. Masses observed corresponded to $[M+H^+-H_2O]^+$ of the target molecules as follows (molecular formula, expected mass, observed mass, Δppm(exp-obs)): 6dEB ($C_{21}H_{37}O_5^+$, 369.2636, 369.2641, 1.4 ppm), desmethyl 6dEB ($C_{20}H_{35}O_5^+$, 355.2479, 355.2482, 0.8 ppm), and fluorodesmethyl 6dEB ($C_{20}H_{34}FO_5^+$, 373.2385, 373.2388, 0.8 ppm).

From the fragmentation patterns, unique transitions were developed for detection with increased sensitivity using an Agilent 6460 QQQ UPLC/MS mass spectrometer using MRM in positive ionization mode with the same LC method. Products were monitored with transitions as follows (parent *m/z* → product *m/z*, fragmentor, collision energy): 6dEB (369.2→239.1, 50, 1), 2-desmethyl 6dEB (355.2→239.1, 50, 1), 8-desmethyl 6dEB (355.2→225.1, 50,1). 2-fluoro-2-desmethyl 6dEB (373.2→275.1, 105, 0).

^{19}F-NMR of fluorine-containing 6dEB analogs in *E. coli* culture medium. Culture extract (500 µL) was mixed with 100 µL of methanol and 100 µL of 1 µM 5-fluorouracil in D_2O. ^{19}F-NMR spectrum was collected at 298 K on a Bruker AV-600 spectrometer at the College of Chemistry NMR Facility. Tune and shimwere optimized. Spectrum was collected using the 19f.cocnmr.av600 pulse program with following parameters: rg = 512, o1p = -200, sw = 300, td =131072, d1 = 0, ds = 20, ns = 1000. Resulting spectrum was processed by MestReNova by backward linear prediction from 0 to 128 using Toeplitz method and phased manually. Spectrum was referenced by a set of doublet corresponding to fluorouracil at -169 ppm.

CoA extraction from *E. coli* and analysis. The OD_{600} of cultures to be analyzed was measured and a culture volume equivalent to 10 mL of sample at 1 OD/mL was collected. The samples were pelleted by centrifugation at 10,000 × *g* for 3 min at 4 ºC in a 15 mL conical centrifuge tube, decanted, and stored at -80 ºC until analysis. At the time of analysis, frozen pellets were centrifuged at 10,000 × *g* for 3 min at 4 ºC and any additional media was removed by pipet. Trichloroacetic acid (150 µL, 10% final concentration) containing 1 mM crotonyl-CoA and an internal standard was added to each cell pallet, which was then resuspended through the wash-board method by passing over a microcentrifuge rack 20 times. Resuspended cells were incubated on ice for 10 mins with vortexing for 15 s every 3 min. The cell suspension was centrifuged at 10,000 × *g* for 10 min at 4 ºC. The supernatant was removed by pipet, transferred to 1.7 mL microcentrifuge tube, centrifuged at 20,000 × *g* for 10 min at 4 ºC. The supernatant was collected and analyzed by LC-QQQ using an Agilent 1290 UPLC and Agilent 6460 Triple Quadrupole mass spectrometer using multiple reaction monitoring (MRM) in positive ionization mode. Samples were analyzed on a Poroshell HPH-C18 column (2.7 µm, 2.1 × 100mm, room temperature, Agilent) using a linear gradient from 0 to 20 % acetonitrile over 6 min (0.6 mL/min) after an initial hold at 0 % acetonitrile for 30 s with 10 mM ammonium formate as the aqueous mobile phase. Transitions used to monitor acyl-CoAs were as follows (transition, collision energy, fragmentation voltage): Coenzyme A (768 → 261, 135, 35); acetyl-CoA (810 → 303, 135, 35); propionyl-CoA (824 → 317, 135, 35); fluoroacetyl-CoA (828 → 321, 135, 35); crotonyl-CoA (836 → 329, 135, 35); malonyl-CoA (854 → 347, 135, 35); methylmalonyl-CoA (868 → 361, 135, 35); and fluoromalonyl-CoA (872 → 365, 135, 35). Ion counts were normalized by the crotonyl-CoA internal standard in each sample. Quantification was performed by comparing normalized ion counts to external standard curves built using commercial or synthetic standards.

For quantification of methylmalonyl-CoA, which has the same mass transition and similar retention time to succinyl-CoA, ^{13}C-propionate was used for confirmation. Cultures fed with ^{13}C-propionate were grown and processed as described above. ^{13}C-labeled metabolites were monitored as follows (transition, collision energy, fragmentation voltage): ^{13}C-propionyl-CoA (825 → 318, 135, 35); ^{13}C-methylmalonyl-CoA (869 → 362, 135, 35). The increase in ^{13}C-methylmalonyl-CoA in cultures fed with ^{13}C-propionate compared to unfed cultures was similar to an increase in ^{12}C-methylmalonyl-CoA in cultures fed with ^{12}C-propionate and was taken as the amount of methylmalonyl-CoA rather than succinyl-CoA.

λ-red-mediated genome editing of *E. coli*. The protocol used was adapted from a previously reported method [24]. LB culture medium (1 L) with 50 µg/mL kanamycin in 2.5 L ultra-yield flask was inoculated with overnight culture (10 mL) of *E. coli* BAP1 pCas. Cells were grown at 30 ºC with shaking at 200 rpm to OD_{600} = 0.1, at which time, λ-red was induced by the addition of arabinose (0.2% (*w/v*)). Cells were grown until OD_{600} = 0.3-0.4 at 30 ºC with shaking at 200 rpm. Cells were then made electrocompetent by washing in cold 10% (*v/v*) glycerol three times. Briefly,

cells were harvested by centrifugation at 8,000 × *g* for 5 min, and after decanting supernatant, resuspended at 50%, 50%, and then 5% culture volume. Cell density was adjusted to OD_{600} = 100 by resuspending in cold 10% (*v/v*) glycerol after the final wash and aliquots (40 µL) were stored at -80 ºC without flash freezing. An aliquot of electrocompetent cells was thawed on ice, mixed with 100 ng of pTargetF and 400 ng of repair oligos, and electroporated. Cells were recovered at 30 ºC for 1-2 h before plating of selection media containing kanamycin and spectinomycin. Edited cells form visible colonies after incubation at 30 ºC for 2 days. Targeted chromosomal regions were amplified by colony PCR and sequenced to confirm editing.

To insert guide sequences into pTargetF, the plasmid was digested with SpeI overnight. A primer pair containing 20 bp guide sequence flanked by Gibson overhangs was then inserted into the digested pTargetF plasmid via Gibson assembly. To construct pTargetF for knocking out FadR, primers FadR_gibsonF/FadR_gibsonR were used. To construct pTargetF for editing FabD, primers FabD_GibsonF/FabD_GibsonR were used.

Repair oligos were constructed by splice-by-overlap-extension (SOE) PCR from two overlapping PCR fragments that contain the mutation in the overlapping region introduced by primers. To knock out FadR, fragments were amplified by colony PCR of *E. coli* BAP1 with primer sets FadR_SOEF2/FadR_SOEmidR and FadR_SOEmidF/FadR_SOER2. To edit FabD, one fragment was amplified by colony PCR of *E. coli* BAP1 with the FabDRO_1F primer paired with either FabDRO_1R_Val, FabDRO_1R_Phe, or FabDRO_1R_Ala primers for introducing the S200V, S200F, and S200A mutations, respectively. The second fragment was amplified using the FabDRO_2R primer paired with either FabDRO_2F_Val, FabDRO_2F_Phe, or FabDRO_2F_Ala for introducing the S200V, S200F, and S200A mutations respectively. Silent mutations were further introduced into SOE PCR products by another round of SOE PCR from two fragments. One fragment was amplified using primers FabDRO_1F/FabDRO_g1_Mutate_R, whereas the second fragment was amplified using primers FabDRO_2R/FabDRO_g1_Mutate_F. Silent mutations were introduced to abolish recognition by Cas9 guide RNA of the edited sequence as the PAM site could not be mutated.

Construction of pCap01_DEBS via transformation-associated recombination in *Saccharomyces cerevisiae*. Capture arms specific to the DEBS pathway and its actinorhodin promoter were added to pCap01 [25]. Capture arms to be inserted into pCap01 were constructed by SOE PCR from three overlapping gBlocks, CapActPDEBS1, CapActPDEBS2, and CapActPDEBS3, using primers pCAP01ActpDEBS1F1/pCAP01ActpDEBS1R1. The SOE PCR product and pCap01 were then digested with XhoI and KpnI and ligated using T4 ligase to yield pCap01_DEBSCaptureArms. The resulting plasmid was confirmed by sequencing.

To construct pCap01_DEBS, *S. cerevisiae* were made competent and transformed using Frozen-EZ Yeast Transformation II Kit (Zymoresearch) according to the manufacturers' protocol or through the lithium method [29].

To transform *S. cerevisiae* using the lithium method, an overnight culture was diluted to OD_{600} = 0.1 in 50 mL of YPD in a 250 mL unbaffled flask and grown to OD_{600} of 0.4-0.6. Cells (50 mL) were harvested by centrifugation at 1,000 × *g* for 5 min, washed once with 50 mL of 100 mM lithium acetate in buffer TE (10 mM Tris and 1 mM EDTA), and resuspended in 1 mL of 100 mM lithium acetate in buffer TE. Competent cells were aliquoted to 100 µL for each transformation and used fresh. To transform, 5-10 ng of pCap01_DEBSCaptureArms linearized by BamHI and 5-10-fold excess of pCK7 linearized by HindIII and EcoI were added to 100 µL of cell suspension.

Boiled and cooled 10 mg/mL salmon sperm DNA was add to 1/10 of the volume of cell suspension. Cells and DNA were mixed by vortexing before adding 750 µL of PEG mixture (40% (*w/v*) PEG 3350 and 100 mM lithium acetate in TE buffer). The cell suspension was vortexed again and incubated at 30 ºC for 1 h. At this time, DMSO (100 µL) was added and the mixture was heat shocked at 42 ºC for 15 min. Cells were then harvested by centrifugation for 15-30 s at 13,000 × g at room temperature, and the supernatant was removed by pipet. The cell pellet was resuspended in YPD and plated on synthetic complete medium without tryptophan. Plates were incubated at 30 ºC for 2 nights until colonies appeared.

CRISPR/Cas-mediated editing of yeast artificial chromosome. To insert guide sequences into pCas_PPhe-BsaI_NAT, the plasmid was digested with BsaI. Primer pairs containing 20 bp guide sequence flanked by Gibson overhangs were then inserted into the digested pCas_PPhe-BsaI_NAT plasmid via Gibson assembly. To construct pCas_PPhe-BsaI_NAT for editing the AT domain of Module 1,2,3, and 6 of DEBS, primer sets M1D1/M1D1R, M2D1/M2D1R, M3D1/M3D1R, and M6D1/M6D1R were used respectively.

Yeast cells containing YAC pCAP01_DEBS were co-transformed with 1 µg of pCas_PPhe-BsaI_NAT plasmid with appropriate guide sequence and 5 ug of repair template. Repair templates for Module 1, 2, 3, and 6 were constructed by amplifying ultramer primer pairs M1 repair oligo F/M1 repair oligo R, M2 repair oligo F/M2 repair oligo R, M3 repair oligo F/M3 repair oligo R, M6 repair oligo F/M6 repair oligo R, respectively. Cells were plated on synthetic complete medium without tryptophan and nourseothricin antibiotic.

Total DNA extraction from yeast. A yeast colony was picked and resuspended in 200 µL Smash-and-grab Buffer (10 mM Tris pH 8.0, 1 mM EDTA. 100 mM sodium chloride, 1 % SDS, and 2 % Triton X-100) in plastic 2 mL screw-top tubes with o-ring seals (Thomas Scientific). 1 mM glass beads (200 µL) were added to the cell suspension, and cells were lysed by bead beating for 5 × 30 s at maximum speed using a Mini-Beadbeater-24 (Biospec). Cell lysate was mixed with 200 µL of equilibrated phenol and mixed by vortexing. This mixture was extracted with 200 µL of water and centrifuged at 20,817 × *g* for 5 min at 4 ºC. The aqueous layer was transferred to a microcentrifuge tube and mixed with 1 mL of cold ethanol and incubated on ice to precipitate DNA. The DNA was pelleted by centrifugation at 20,817× *g* for 5 min at 4 ºC. After discarding the supernatant, the DNA pellet was washed with 500 µL of 100% ethanol and air-dried before resuspending in 10 µL of water.

Yeast artificial chromosome isolation. The yeast artificial chromosome was isolated based on a previously published protocol [30]. An overnight yeast culture (10 mL) was inoculated into 500 mL of appropriate synthetic complete medium in 2.5 L ultra-yield flask and grown to OD_{600} = 1.5, at which point cells were harvested by centrifugation at 5,000 × *g* for 10 min at 4 ºC. The cell pellets were resuspended in 100 mL of water, transferred to 50 mL Falcon tubes, and pelleted again by centrifugation. Cell pellets were stored at -20 ºC until DNA isolation.

Each cell pellet was resuspended in 40 mL of SPE Buffer (10 mM sodium phosphate, 1 M sorbitol, 10 mM EDTA, pH 7.5). Cells were spheroplasted by adding Zymolyase 100T (50 µL) and BME (40 µL) and incubated at 37 ºC for 1 h. Spheroplasting efficiency was measured by adding 200 µL of spheroplast suspension to 1.8 µL of 2 % (*w/v*) SDS or 1 M sorbitol and monitoring the OD_{600}. When the two OD_{600} readings were 4- to 5-fold different from each other, yeast spheroplasts were deemed ready and harvested by centrifugation at 1000 × *g* for 5 min at 4 ºC. The supernatant was discarded and spheroplasts were resuspended in 1 M sorbitol (1 mL).

Spheroplasts were lysed by adding 20 mL of Lysis Buffer (50 mM Tris-HCl, 20 mM EDTA, and 1% (*w/v*) SDS, pH 12.8) and incubating at 37 ºC for 30 min. DNA was extracted by adding room temperature phenol:chloroform:isoamyl alcohol (20 mL). The mixture was centrifuged at 5,000 × *g* for 20 min at room temperature. The aqueous layer was collected in a separate tube and mixed with 3 M sodium acetate (2 mL) and isopropanol (20 mL) to precipitate DNA, which was pelleted by centrifugation at 20,817 × *g* for 30 min at 4 ºC. The supernatant was discarded and the DNA pellet was rinsed with 10 mL of 70% ethanol.

The circular YAC was then isolated from total DNA using a Qiagen Large-Construct Kit. The DNA pellet was resuspended in TE buffer (1 mL) after which 3 µL of 10 mg/mL RNase A was added to degrade RNA. After incubating for 30 min at 37 ºC, the solution was diluted into 10 mL of Buffer EX (Qiagen) and incubated with 200 µL of ATP-dependent exonuclease and 300 µL of 100 mM ATP at 37 ºC for 30 min to degrade linear DNA. A Qiagen-tip 500 column was equilibrated with 10 mL of buffer QBT (Qiagen). The DNA sample was mixed with 12 mL of Buffer QS (Qiagen) and centrifuged at 5,000 × *g* for 10 min at 4 ºC to remove undissolved material. The sample was then applied to the equilibrated Qiagen-tip 500 column. The column was washed with 30 mL of Buffer QC (Qiagen), and DNA was eluted with 12 mL of Buffer QF (Qiagen) equilibrated to 55 ºC. The DNA was precipitated from the eluate by adding 15 µL of 10 mg/mL total yeast tRNA, 1.2 mL of 3 M sodium acetate, and 12 mL of isopropanol, and centrifuging at 5,000 × *g* for 1 h at 4 ºC. The DNA pellet was washed with 5 mL of 70% ethanol, air dried, and resuspended in water. The isolated YAC DNA was used for restriction patterning and transformation into *E. coli*.

***Streptomyces albus* J1074 spore suspension preparation.** A liquid culture of *S. albus* was plated on an MS agar plate and incubated at 30 ºC until a white aerial mycelium appeared (~3-7 d). Sterile water (10 mL) was added to the plate, and the surface was scraped with an inoculation loop. The crude spore suspension was then transferred to a 15 mL conical tube and vortexed vigorously to break up spores from mycelia. The spores were pelleted by centrifugation at 1,000 × *g* for 10 min at 4 ºC. Spores collected at the bottom of the tube while mycelia floats on top of water layer. The supernatant and mycelia were discarded by decanting. Spores from each plate were resuspended in water for immediate use or 20% (*v/v*) glycerol (500 µL) for storage at -20 ºC until use.

***Streptomyces albus* J1074 conjugation.** pCap01_DEBS was introduced into *S. albus* J1074 for genomic integration through conjugal transfer from *E. coli* WM3064 through either liquid culture or dry conjugation methods.

In liquid culture conjugation, *E. coli* WM3064 pCap01_DEBS was grown in LB supplemented with 0.3 mM of diaminopimelic acid (DAP) and 50 µg/mL kanamycin at 37 ºC with shaking at 200 rpm to OD_{600} = 0.6. Cells were harvested by centrifugation at 20,817 × *g* for 1 min at 4 ºC, washed twice with LB, and resuspended in 10% the original volume of LB supplemented with 0.3 mM of DAP. The spore suspension of *S. albus* J1074 (100 µL) was diluted in 400 µL of 2×YT and heat shocked at 55 ºC for 10 min. The cooled spore suspension was cooled to room temperature then mixed with 250 µL of *E. coli* WM3064 pCap01_DEBS before proceeding to plating as described below. In dry conjugation, *E. coli* WM3064 pCap01_DEBS was plated on LB supplemented with 0.3 mM of DAP and 50 µg/mL kanamycin. After overnight incubation, *E. coli* cells were scraped from plate and mixed with heat shocked *S. albus* spores freshly collected from plates as described above.

For both methods, the cell mixture was then centrifuged at 20,817 × *g* for 1 min at 4 ºC, plated on a non-selective MS plate, and incubated at 30 ºC for 14 h. The appropriate antibiotic (50 µg/mL kanamycin, 1mL) was overlaid onto the plate, which was incubated at 30 ºC until colonies sporulated (~3-7 d). Colonies were picked and streaked to single colonies on plates with antibiotic selection.

6dEB production in *Streptomyces albus* and *Streptomyces lividans*. Prepared growth media (50 mL), such as Tryptic soy broth (TSB), YEME, GYM, and nutrient broth (NB), supplemented with 5 mM sodium propionate and when used 2 mM fluoromalonate in a 250 mL baffled flask with 5-20 glass beads was inoculated with 50 µL of spore stock. *Streptomyces albus* cultures with the integrated DEBS pathway were supplemented with 50 µg/mL kanamycin and *Streptomyces lividans* cultures harboring pCK7 were supplemented with 10 µg/mL thiostrepton. Cultures were grown at 30 ºC with shaking at 200 rpm. After 7-13 d, culture media was harvested and cells pelleted by centrifugation at 20,817 × *g* for 5 min at 4 ºC. The supernatant was collected and acidified with 60% perchloric acid to 5% (*v/v*) final concentration. Proteins were precipitated by centrifugation at 20,817 × *g* for 10 min at 4 ºC. Cleared culture media were analyzed on an Agilent 1290 UPLC on a Poroshell 120 SB-Aq column (2.7 µm, 2.1 × 50 mm, 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) by LC-QTOF.

4.3 Results and Discussion

Design of fluorinated 6dEB biosynthetic pathway in *E.* coli host. To determine the best model system for our experiments, we obtained two published *E. coli* 6dEB production systems: 1) pBP(JW)130 and pBP(JW)144, and 2) pDEBS.

pBP(JW)130 and pBP(JW)144 together encode for the generation system of precursor methylmalonyl-CoA and the biosynthetic pathway of 6dEB [8, 31]. pBP(JW)130 has a pBR322 origin of replication and encodes DEBS2 and DEBS3 proteins in an operon controlled by an IPTG-inducible T7 promoter. pBP(JW)144 shares the origin of replication with pBP(JW)130 and encodes for DEBS1 protein along with propionyl-CoA carboxylase subunits, PccA and PccB. PccAB together catalyzes the carboxylation reaction of propionyl-CoA to methylmalonyl-CoA, allowing for the accumulation of the necessary precursor of 6dEB. Note that the (JW) notation designates changes in antibiotic selection marker in these plasmids. The two-plasmid production system was transformed into *E. coli* BAP1 [8] and TB3 [17] strains, both of which are derived from *E. coli* BL21(DE3) and carry the *sfp* and *prpE* genes on their chromosome under the control of the T7 promoter . Sfp is necessary for installing the phosphopantathiene modification necessary for turning apo-ACP into a functional holo-ACP domain, while PrpE facilitates the intracellular accumulation of the propionyl-CoA starter unit through its propionate:CoA ligase activity. *E. coli* TB3 contains an additional knockout of the *ygfH* gene, which encodes for a propionyl-CoA:succinyl-CoA transferase involved in propionate metabolism that interconverts propionyl-CoA and succinate with propionate and succinyl-CoA [17].

pDEBS is an bacterial artificial chromosome (BAC) that carries the DEBS pathway under control of a bi-directional inducible T7 promoter in which genes encoding DEBS1, DEBS2, and DEBS3 have been codon-optimized for *E. coli* [18]. In addition, pDEBS also bears the genes for PccAB, Sfp, and the GroES/EL chaperonin system from *Saccharopolyspora erythraea* [18]. pDEBS was originally designed to produce erythromycin from the 6dEB aglycone core and

therefore also carries the genes involved post-DEBS tailoring. pDEBS was transformed into *E. coli* LF01, which was derived from *E. coli* TB3 and further includes the knock-out of *vioAB*, *wzx*, and *wecDE* genes, which encode for sugar processing enzymes [18]. These knockouts were introduced to maintain flux from thymidine-diphosphate (TDP)-4-keto-6-deoxy-D-glucose to TDP-D-desosamine and TDP-L-mycarose, specialized sugars that are part of erythromycin A.

We compared the yield between the strains and found that *E. coli* BAP1 and TB3 containing pBP(JW)130 and pBP(JW)144 produced 6dEB levels much higher than *E. coli* LF01 pDEBS (*Figure 4.1*). We believe this difference in yield is due to the difference in copy number of the DEBS proteins-encoding plasmids as pBP130 and pBP144 are both high-copy number plasmids while pDEBS is a low-copy number BAC. We decided to proceed with *E. coli* BAP1 pBP130 pBP144 as our model system several reasons. Among them, BAP1 is one of the highest producers and a commonly used strain in our lab and the antibiotic resistance markers for pBP130 and pBP144 are compatible with the pFmal plasmids, which encode a malonate transporter with MatB. We do not believe the difference in antibiotic resistance between pBP and pBP(JW) exerts an effect on production. We also decided not to use *E. coli* TB3 as we believe the transferase knock out to maintain propionyl-CoA flux was not necessary, as we were able generate easily detectable amounts of 6dEB in *E. coli* BAP1.

We then tested this strain and plasmid combination in several different medias following the previously reported culturing protocol [28]. We found that a previously optimized media recipe [28] was able to generate higher amount of 6dEB compared to other common laboratory medium recipes (*Figure 4.2*). Therefore, we developed our production protocol based on this optimized media recipe.

With an initial model system and culturing condition at hand, we moved towards our complementation system to incorporate fluoromalonyl-CoA into polyketide molecules. To determine the effect of pFmal on the production of 6dEB, we co-transformed pBP130 and pBP144 with one of our pFmal plasmids encoding for malonyl-CoA synthetase MatB and malonate transporter MadLM (pFmal(MadLM)) into *E. coli* BAP1. The pFmal plasmids were previously characterized and shown to be able to support the production of fluorine-containing diketides and biopolymers in *E. coli* [32], as well as fluorinate triketides using single-module DEBS constructs (*Section 3.3*). We performed a cellular production experiment to investigate the effect of the pFmal(MadLM), on 6dEB titer and found that its inclusion slightly affected the 6dEB production system (*Figure 4.3*). This result was consistent with previous observations that the expression of MatB and MadLM could decrease cellular fitness, leading to a decrease in product yield. However, when looking at protein levels, we found that the expression of MatB and MadLM did not drastically lower the expression level of DBES proteins by SDS-PAGE and Western Blot analysis (*Figure 4.4*).

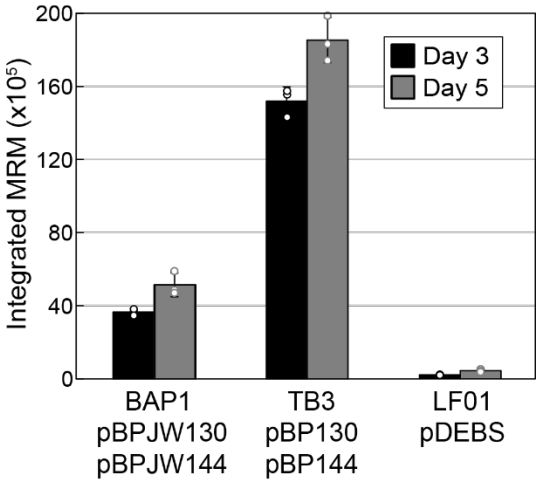


Figure 4.1. *6dEB production by various strain and plasmid combinations. E.coli BAP1, TB3, and LF01 transformed with pBPJW130/pBPJW144, pBP130/pBP144, and pDEBS respectively were grown in 6dEB production media. Cultures were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures provided with 20 mM sodium propionate. Cultures were incubated with substrate at 22 ºC with shaking at 250 rpm for 3-5 days, after which culture media were collected and analyzed by LC-QQQ. Data are mean ± s.d. of biological replicates (n=3)*

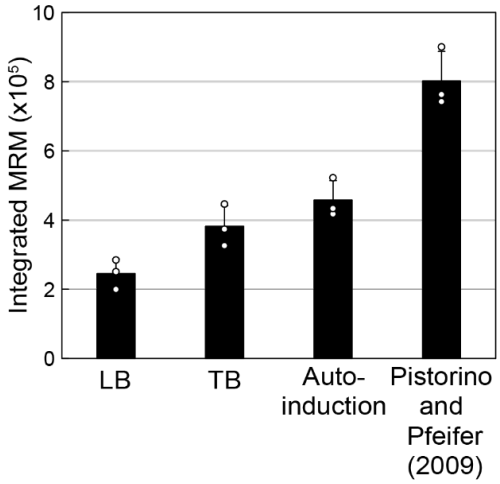


Figure 4.2. *6dEB production by E. coli BAP1 harboring pBP130 and pBP144 in various media. Cells cultured in LB, (Luria-Bertani), TB (Terrific Broth), and Pistorino and Pfeifer [28] were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and provided with 20 mM sodium propionate. Cultures were then incubated with substrate at 22 ºC with shaking at 250 rpm for 3 days. Cells cultured in Auto-induction medium [33] were grown at 37 ºC with shaking at 200 rpm for 3 h, before they were shifted to 22 ºC with shaking at 250 rpm for 3 days. Culture media were then collected and analyzed by LC-QQQ. Data are mean ± s.d of biological replicates. (n=3)*

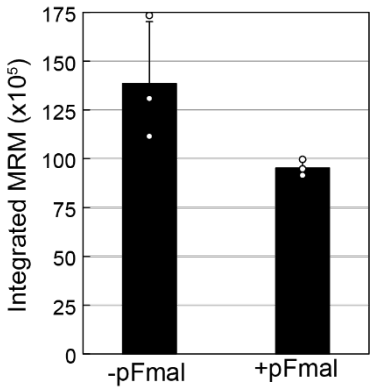


Figure 4.3. *Effects of MatB and MadLM expression on 6dEB production in E. coli BAP1 harboring pBP130 and pBP144. E. coli cells with or without pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose and provided with 20 mM sodium propionate. Cultures were then incubated with substrate at 22 ºC with shaking at 250 rpm for 1 day, after which, culture media were collected and analyzed by LC-QQQ Data are mean ± s.d of biological replicates (n = 3).*

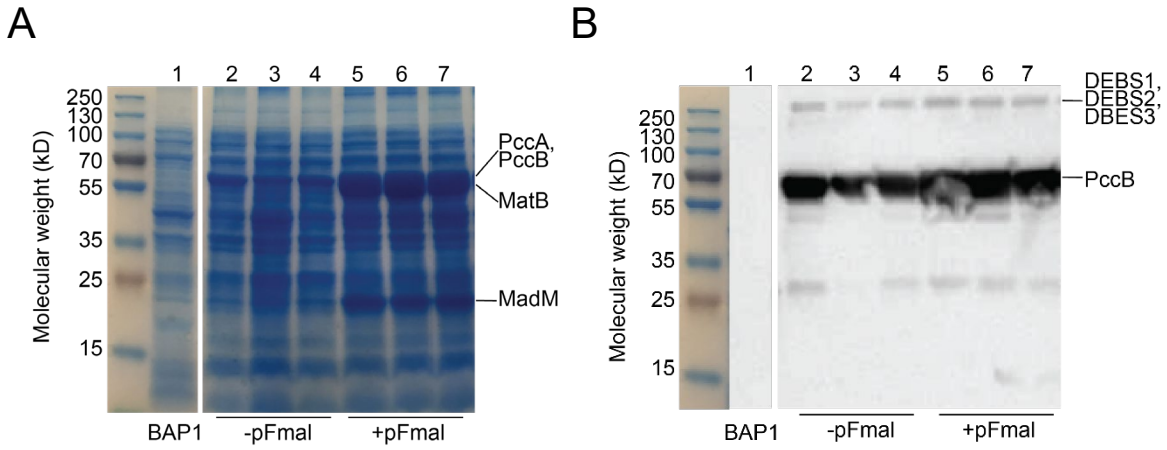


Figure 4.4 *Effects of MatB and MadLM expression on protein expression in E. coli BAP1 expressing Mod1 AT^0 DEBS pathway. (A) Total protein and (B) α-His Western analysis of cultures of E. coli BAP1 harboring pBP130 and pBP144_DEBS1(Mod1 AT^0) with or without pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose, and cultures were provided with 20 mM sodium propionate. Cultures were then incubated with substrates at 22 ºC with shaking at 250 rpm for 1 day, after which, cells were harvested by centrifugation and analyzed by SDS-PAGE. (lane 1) E. coli BAP1 control, (lane 2-4) E. coli BAP1 harboring pBP130 and pBP144_DEBS1(Mod1 AT^0), (lane 5-7) E. coli BAP1 harboring pBP130, pBP144_DEBS1(Mod1 AT^0), and pFmal(MadLM).*

Introducing inactivating mutation into the active site of DEBS *cis*-AT domains. The next step towards applying out complementation strategy was to introduce an inactivating mutation into the *cis*-AT domains of the DEBS proteins. We initially attempted to perform Gibson assembly to introduce active site mutations to the AT domains of DEBS modules encoded by pBP130 and pBP144 plasmids. However, we found that this strategy did not yield the correct plasmid due to loss of sections of the plasmid, likely through recombination. We believe that the truncated plasmids, which accounted for the majority of plasmid products of the Gibson reactions, were a result of the low efficiency of Gibson reaction when applied to large plasmid sizes, the repetitive nature of the DEBS sequence, and a bias towards smaller plasmids of *E. coli* chemical transformation process.

To increase the efficiency of the process by introducing a selection step, we turned to a two-step cloning strategy in which a selectable marker is introduced in the first step to allow for selection and counterselection *(Figure 4.5*). As pBP130 and pBP144 have carbenicillin and kanamycin resistance markers respectively, we chose chloramphenicol resistance cassette as the additional resistance cassette. In the first step, unique restriction sites flanking the mutation site were identified, and the plasmid was digested by the selected restriction enzymes. To do so, the chloramphenicol resistance cassette was amplified with primers that each contains a Gibson overhang, a restriction site new to pBPP130 or pBP144, and the annealing sequence in that order from 5′ to 3′. The chloramphenicol resistance cassette was then inserted into digested plasmid via Gibson assembly. This strategy inserted a removable chloramphenicol resistance cassette at the site of the desired mutation. The assembly product was then selected for by chloramphenicol resistance.

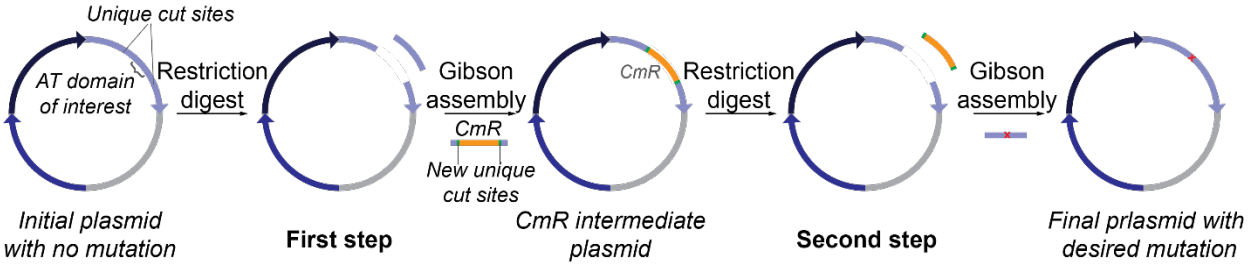


Figure 4.5. *Two-step Gibson assembly cloning strategy for introducing active mutation to AT domains on DEBS pathway enzymes. In the first step, the region of DNA to be mutated was replaced by a selectable Cm^R marker (orange) using unique restriction sites in the plasmid (pBP130 or pBP144). In the second step, the Cm^R cassette was removed using unique restriction sites introduced during the first stage and replaced with the target DNA sequence containing the mutation (red spot).*

In the second step, the product from the first step was digested with the unique restriction sites that had been introduced using the primers flanking chloramphenicol resistance cassette. The DNA piece containing the original DEBS sequence with a single point mutation in the active site of the AT domain was obtained either through amplifying a template that contains the mutation or through SOE PCR with the mutation introduced by primers. The final plasmid was assembled via Gibson assembly of digested intermediate plasmid and new DEBS sequence containing the desired mutation. This product was counter-selected for by chloramphenicol sensitivity. The final plasmids were fully sequenced by complete plasmid sequencing in order to confirm that no other unexpected

changes or recombinational events were introduced. Through this two-step process, active site mutation was successfully introduced to the AT domain of modules 1, 3, 5 and 6 of DEBS pathway.

Production of desmethyl-6dEB by mutant pathways. To determine the function of mutant pathways, we co-transformed wild-type and mutant pBP130 and pBP144 with pFmal(MadLM) into *E. coli* BAP1. All cultures fed with propionate produced levels of 6dEB product (*Figure 4.6*). However, all mutant pathways produced decreased levels of 6dEB compared to wild-type, suggesting that all the enzymatic activities in DEBS pathway proteins were functional in these mutant constructs and that complementation of the inactivating mutation of the active site of the AT domains had occurred, either by FabD or by other functional *cis*-AT domains in the PKS.

When fed with malonate along with propionate, cultures harboring mutant pathways produced increased levels of desmethyl-6dEBs (*Figure 4.7*). While the wildtype pathway seemed to produce a few different species of desmethyl-6dEB all at low levels, cells harboring pathways with an inactivating mutation in the AT domain of module 3, 5 or 6 produced predominantly a single species of desmethyl-6dEB.

We hypothesized that each of the desmethyl-6dEB species found was a result of the module with inactivated AT domain incorporating malonyl-CoA instead of methylmalonyl-CoA, possibly through complementation by FabD. In other words, the DEBS pathway with an inactivating mutation in Modules 3, 5 and 6 would produce 8-desmethyl-6dEB, 4-desmethyl-6dEB, and 2-desmethyl-6dEB respectively. To determine the substitution pattern on each of the observed desmethyl-6dEB species, we performed fragmentation analysis and found that the fragmentation patterns were consistent with our hypothesis (*Figure 4.8-4.10*). This result suggests that the mutant pathways exhibit regioselective behavior towards unnatural extender unit incorporation. Moreover, as fluoromalonyl-CoA is similar in size to malonyl-CoA, the regioselective incorporation of malonyl-CoA by mutant pathways suggests that fluoromalonyl-CoA incorporation is could also be achieved in these positions.

However, no dominant desmethyl-6dEB was found in cultures harboring inactivated AT domain in DEBS Module 1. As the expected molecule, 12-desmethyl-6dEB, had previously been observed through *in vitro* complementation experiments [6], we reasoned that the lack of detectable product from fermentation here was due to unbalanced ratio between the DEBS proteins and *trans*-AT protein. *In vitro* experiments showed that 12-desmethyl-6dEB was observed when DszAT was supplied at 10, 100, or 1000 nM and DEBS modules at 2 μM [6]. However, the precise amount of both DEBS proteins and FabD in *E. coli* cells are currently unknown. Alternatively, we reasoned that the intermediate that contains a non-native group must encounter a larger number of downstream enzymatic steps when malonyl-CoA was incorporated by Module 1 than when it is incorporated by Modules 3, 5 or 6, which occur later in 6dEB production. Thus, it is possible that it is processed at lower overall efficiency and produced below our detection limit.

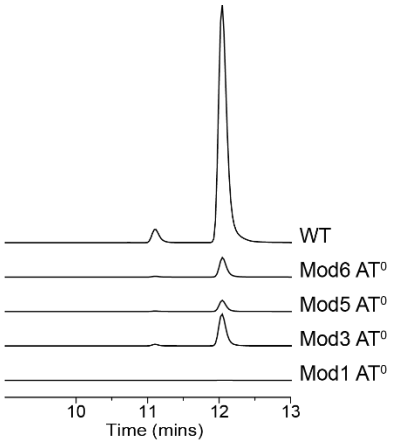


Figure 4.6. *Extracted ion chromatograms of reactions showing in vivo production of 6dEB by WT DEBS, Mod1 AT^0 DEBS, Mod3 AT^0 DEBS, Mod5 AT^0 DEBS, Mod6 AT^0 DEBS analyzed by LC-QQQ using MRM transition 369.2→239.1. E. coli BAP1 harboring pBP130 and pBP144 with AT^0 mutation in module 1, 3, 5, or 6 and pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose and cultures provided with 20 mM sodium propionate. Cultures were then incubated with substrate at 22 ºC with shaking at 250 rpm for 1 day, after which, culture media were collected and analyzed by LC-QQQ. Data are mean ± s.d of biological replicates (n = 3).*

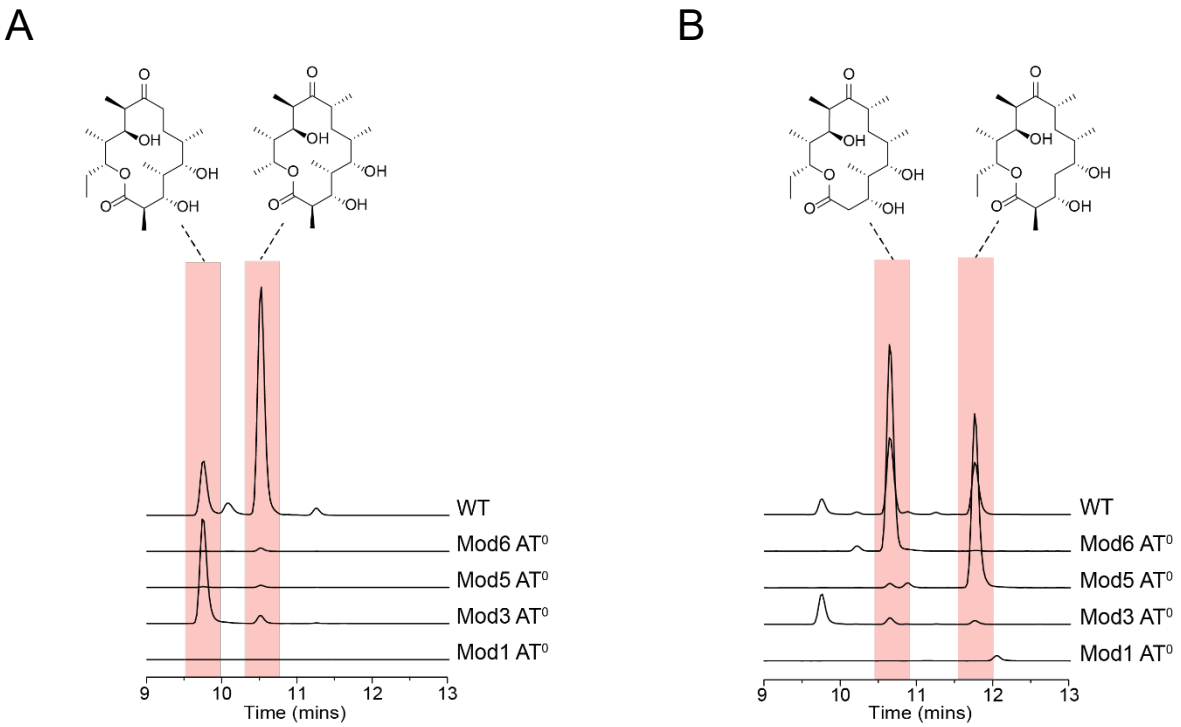
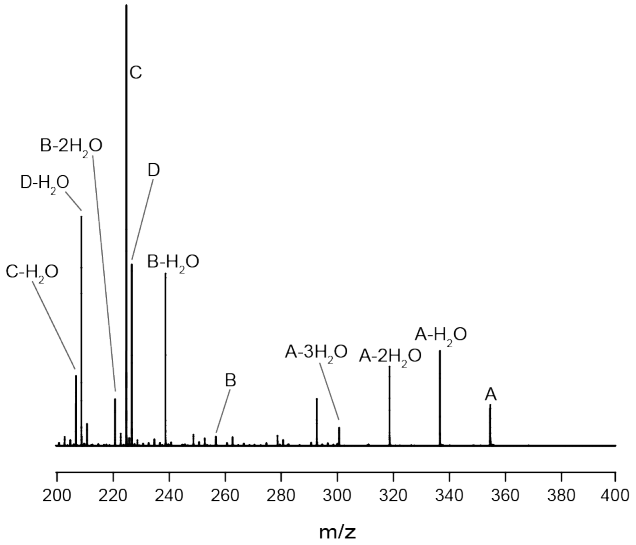


Figure 4.7. *Extracted ion chromatograms of reactions showing in vivo production of desmethyl 6dEB analogs by E. coli BAP1 expressing WT DEBS, Mod1 AT^0 DEBS, Mod3 AT^0 DEBS, Mod5 AT^0 DEBS, Mod6 AT^0 DEBS analyzed by LC-QQQ using MRM transition (A) 355.2→225.1 or (B) 355.2→239.1. E. coli BAP1 harboring pBP130 and pBP144 with AT0 mutation in module 1, 3, 5, or 6 and pFmal(MadLM) were grown to OD600 = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose and cultures provided with 20 mM sodium propionate. Cultures were then incubated with substrate at 22 ºC with shaking at 250 rpm for 1 day, after which, culture media were collected and analyzed by LC-QQQ. Data are mean ± s.d of biological replicates (n = 3).*

A

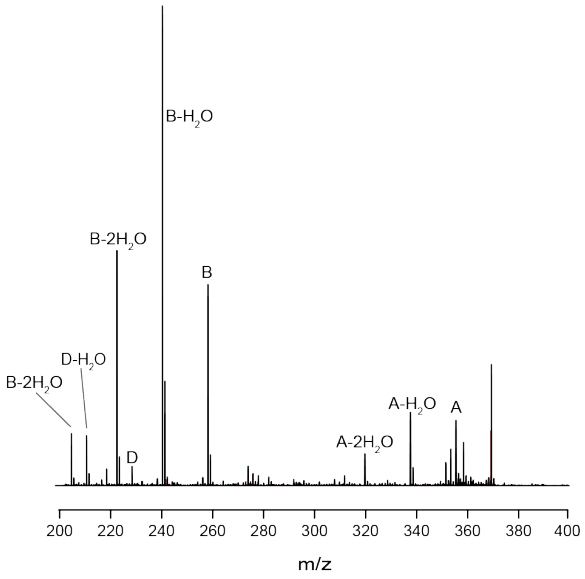


B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{35}O_5^+$	355.2479	355.2464
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.236
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2256
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	301.2152
B	$C_{14}H_{25}O_4^+$	257.1747	257.1739
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.1632
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1526
B-3H_2O	$C_{14}H_{19}O^+$	203.143	Not observed
C	$C_{14}H_{25}O_2^+$	225.1849	225.1839
C-H_2O	$C_{14}H_{23}O^+$	207.1743	207.1735
D	$C_{12}H_{19}O_4^+$	227.1278	227.1268
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.1162
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	191.1058
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	Not observed

Figure 4.8 *LC-QTOF fragmentation analysis of the desmethyl 6dEB analog produced by E. coli BAP1 expressing Mod3 AT^0 DEBS. E. coli BAP1 harboring pBP130 and pBP144_DEBS2(Mod3 AT^0) were grown to OD600 = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose and cultures provided with 20 mM sodium propionate. Cultures were then incubated with substrates at 22 ºC with shaking at 250 rpm for* 1 *day, after which, culture media were collected, extracted, and analyzed by LC-QTOF (A) Fragmentation pattern shows presence of ions in A, B, C, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 biological replicates.*

A

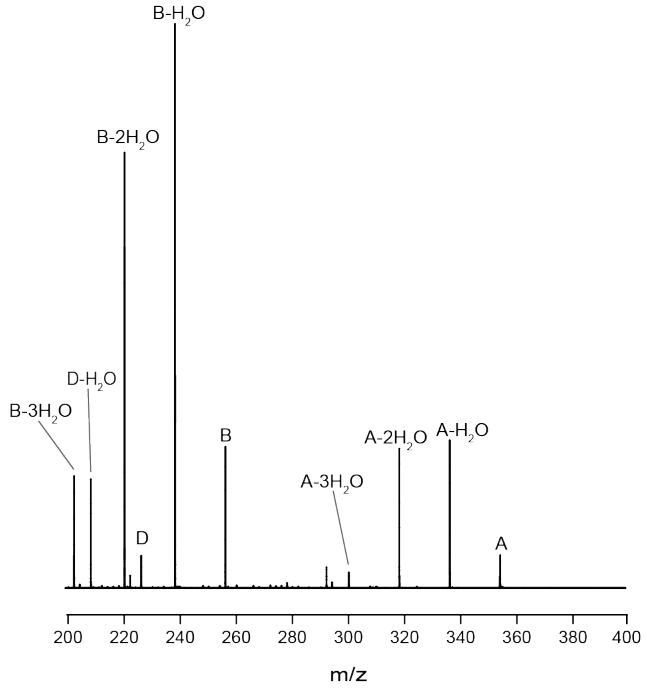


B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{35}O_5^+$	355.2479	355.2484
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.2367
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2261
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	Not observed
B	$C_{14}H_{25}O_4^+$	257.1747	257.1745
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.164
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1533
B-3H_2O	$C_{14}H_{19}O^+$	203.143	203.1434
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{19}O_4^+$	227.1278	227.1269
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.117
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	191.1069
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	173.0806

Figure 4.9. *LC-QTOF fragmentation analysis of the desmethyl 6dEB analog produced by E. coli BAP1 expressing Mod5 AT^0 DEBS. E. coli BAP1 harboring pBP130_DEBS3(Mod5 AT^0) and pBP144 were grown to OD600 = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose and cultures provided with 20 mM sodium propionate. Cultures were then incubated with substrates at 22 ºC with shaking at 250 rpm for 1 day, after which, culture media were collected, extracted, and analyzed by LC-QTOF (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 biological replicates.*

A



B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{35}O_5^+$	355.2479	355.2472
A-H_2O	$C_{20}H_{33}O_4^+$	337.2373	337.2366
A-2H_2O	$C_{20}H_{31}O_3^+$	319.2268	319.2256
A-3H_2O	$C_{20}H_{29}O_2^+$	301.2162	301.2158
B	$C_{14}H_{25}O_4^+$	257.1747	257.1738
B-H_2O	$C_{14}H_{23}O_3^+$	239.1642	239.1632
B-2H_2O	$C_{14}H_{21}O_2^+$	221.1536	221.1528
B-3H_2O	$C_{14}H_{19}O^+$	203.143	203.1424
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{19}O_4^+$	227.1278	227.1272
D-H_2O	$C_{12}H_{17}O_3^+$	209.1172	209.1163
D-2H_2O	$C_{12}H_{15}O_2^+$	191.1067	191.1056
D-3H_2O	$C_{12}H_{13}O^+$	173.0961	Not observed

Figure 4.10 *LC-QTOF fragmentation analysis of the desmethyl 6dEB analog produced by E. coli BAP1 expressing Mod6 AT^0 DEBS. E. coli BAP1 harboring pBP130_DEBS3(Mod6 AT^0) and pBP144 were grown to OD600 = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose and cultures provided with 20 mM sodium propionate. Cultures were then incubated with substrates at 22 ºC with shaking at 250 rpm for 1 day, after which, culture media were collected, extracted, and analyzed by LC-QTOF (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 biological replicates.*

Production of fluorodesmethyl-6dEB by DEBS pathways with inactivated *cis*-AT domains. We attempted to incorporate fluorine into 6dEB by feeding cells harboring the production system with fluoromalonate and propionate. The production strains express DEBS pathways with an active site mutation in one of the AT domains from pBP130 and pBP144 plasmids, and MatB and MadLM from pFmal(MadLM) (*Figure 4.11*). Initial attempts to produce fluorine-containing 6dEB analogs did not yield any desired products. However, we found that the level of 6dEB decreased when cultures expressing pFmal(MadLM) were provided fluoromalonate (*Figure 4.12*).

SDS-PAGE and α-His Western analysis of cells after feeding showed that feeding cells with fluoromalonate during log phase at the time of protein induction correlated with a decrease in global protein expression (*Figure 4.13*). We also found that cells that received fluoromalonate exhibited growth defects (*Figure 4.14*). These results are consistent with our previous finding that feeding fluoromalonate during log phase to cells expressing high amounts of MatB and MadLM let to growth defects (*Section 3.3*). Similar toxicity effects were similarly observed for cells that received malonate (*Figure 4.12-4.14*). Thus, we believe that the toxicity is not solely due to

intracellular accumulation fluoromalonyl-CoA, but rather related to increased concentration of malonate or malonyl-CoA. As malonate is a known inhibitor of succinate dehydrogenase [34], we hypothesized that increased high intracellular accumulation of malonate and fluoromalonate may lead to down regulation of the tricarboxylic acid (TCA) cycle. Thus, we attempted to rescue the cycle and growth by supplementing production cultures with fumarate, the product of succinate dehydrogenase. However, fumarate did not rescue growth of cells fed with malonate or fluoromalonate (*Figure 4.14*). Alternatively, we hypothesized that the high expression of MatB protein depleted the intracellular pool of free coenzyme A as the protein uses free coenzyme A to produce fluoromalonyl-CoA. However, an analysis of intracellular coenzyme A content revealed some level of free coenzyme A present throughout log-phase growth (*Figure 4.15*). The nature of this toxicity effect remains a topic of further investigations.

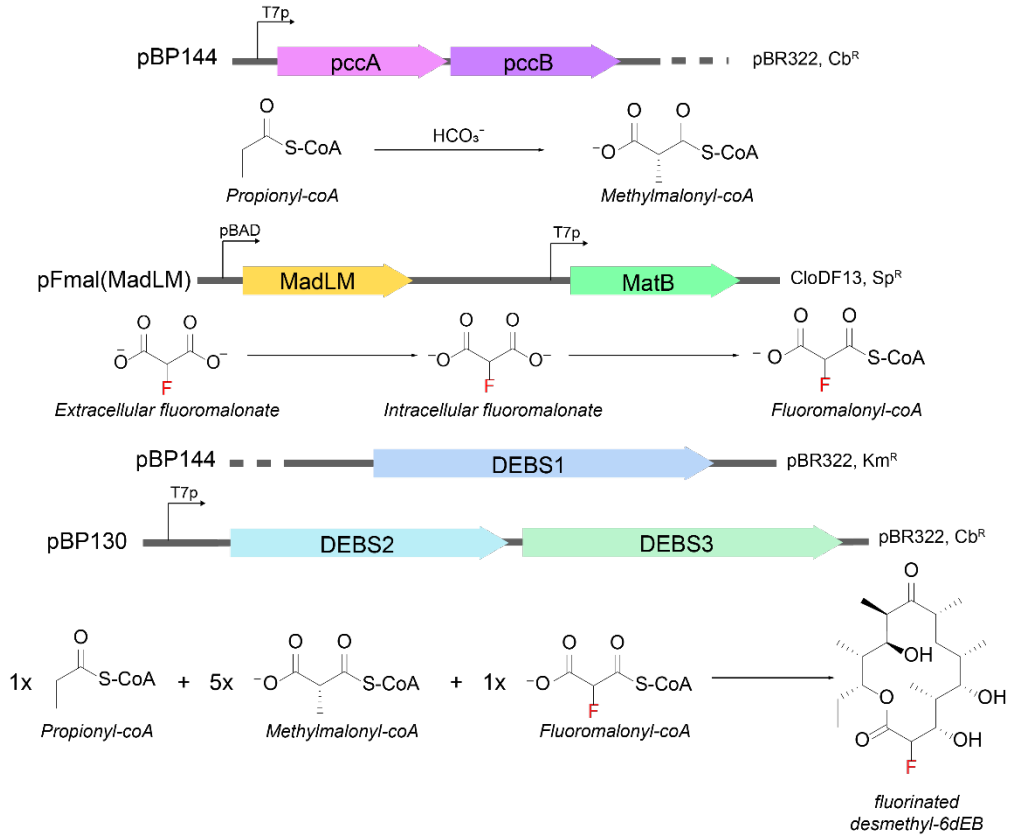


Figure 4.11. *Design of pathway for fluorine-containing 6dEB analogs production in E. coli. The system includes three plasmids with compatible antibiotic markers, which together encode for six pathway proteins. pFmal(MadLM) plasmid (cloDF13 origin, Sp^R) encodes for malonate transporter MadLM, and malonyl-CoA synthetase MatB. pBP130 (pBR322 origin, Cb^R) and pBP144 (pBR322 origin, Km^R) encode for DEBS1, DEBS2, and DEBS3 proteins. pBP144 also encodes for a propionyl-CoA carboxylase (pccAB).*

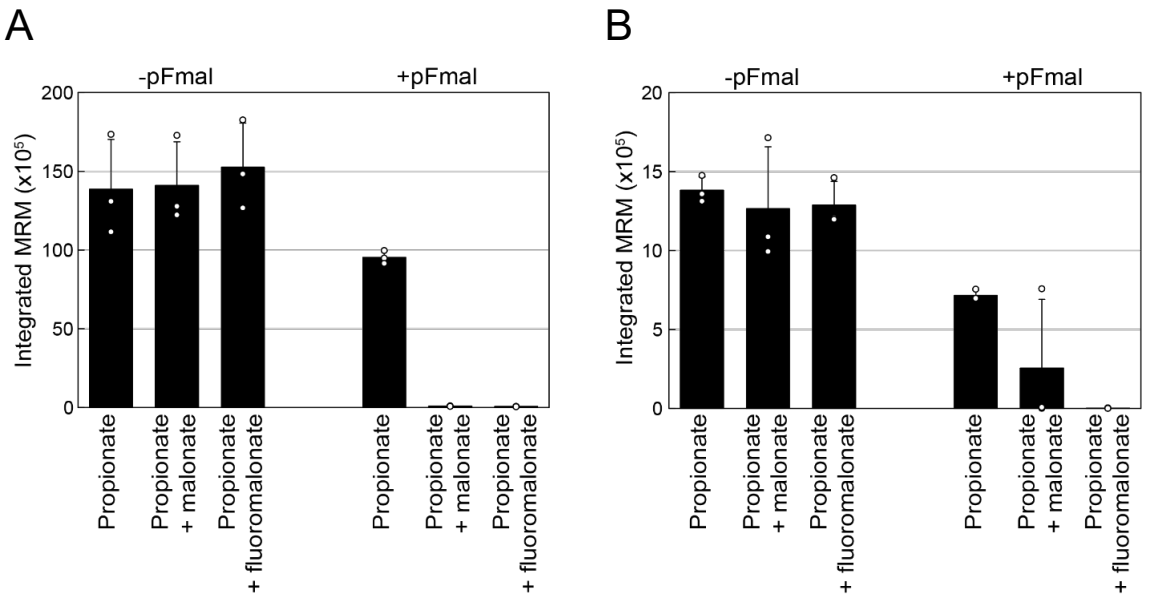


Figure 4.12. *Effects of malonate and fluoromalonate feeding on 6dEB production in E. coli BAP1 expressing (A) DEBS or (B) Mod1 AT^0 DEBS, MatB, and MadLM. E. coli BAP1 cells harboring (A) pBP130 and pBP144 or (B) pBP130 and pBP144_DEBS1(Mod1 AT^0) with or without pFmal(MadLM) were grown to OD$_{600}$ = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose, and cultures were provided with 20 mM sodium propionate, and 5 mM malonate or 5 mM fluoromalonate. Cultures were then incubated with substrates at 22 ºC with shaking at 250 rpm for 1 day, after which, culture media were collected and analyzed by LC-QQQ Data are mean ± s.d of biological replicates (n = 3).*

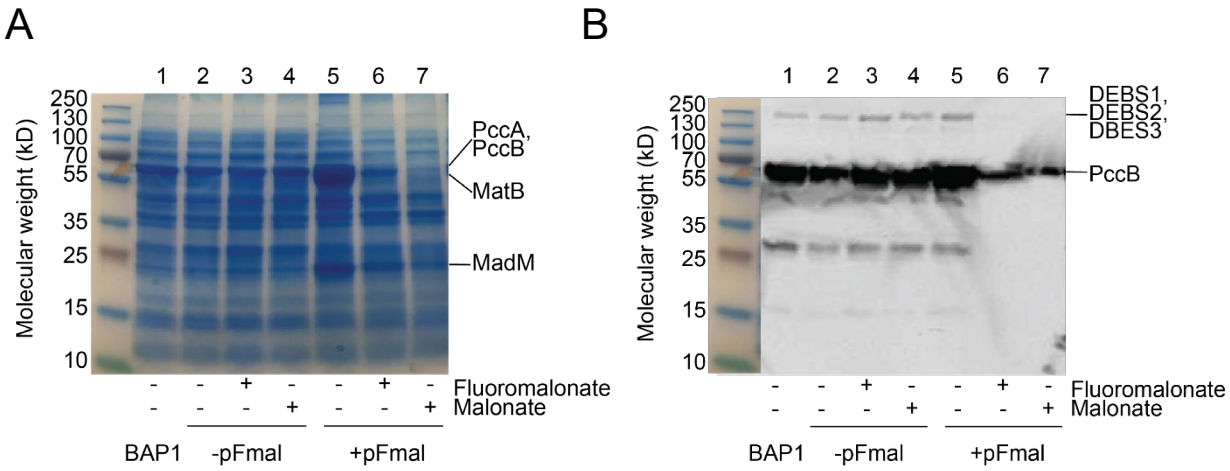


Figure 4.13. *Effects of malonate and fluoromalonate feeding on protein expression in E. coli BAP1 expressing Mod1 AT^0 DEBS pathway, MatB, and MadLM. (A) Total protein and (B) α-His Western analysis of cultures of E. coli BAP1 harboring pBP130 and pBP144_DEBS1(Mod1 AT^0) with or without pFmal(MadLM) were grown to OD$_{600}$ = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose, and cultures were provided with 20 mM sodium propionate, and 5 mM malonate or 5 mM fluoromalonate. Cultures were then incubated with substrates at 22 ºC with shaking at 250 rpm for 1 day, after which, cells were harvested by centrifugation and analyzed by SDS-PAGE. (lane 1) E. coli BAP1 control, (lane 2-4) E. coli BAP1 harboring pBP130 and pBP144_DEBS1(Mod1 AT^0), (lane 5-7) E. coli BAP1 harboring pBP130, pBP144_DEBS1(Mod1 AT^0), and pFmal(MadLM).*

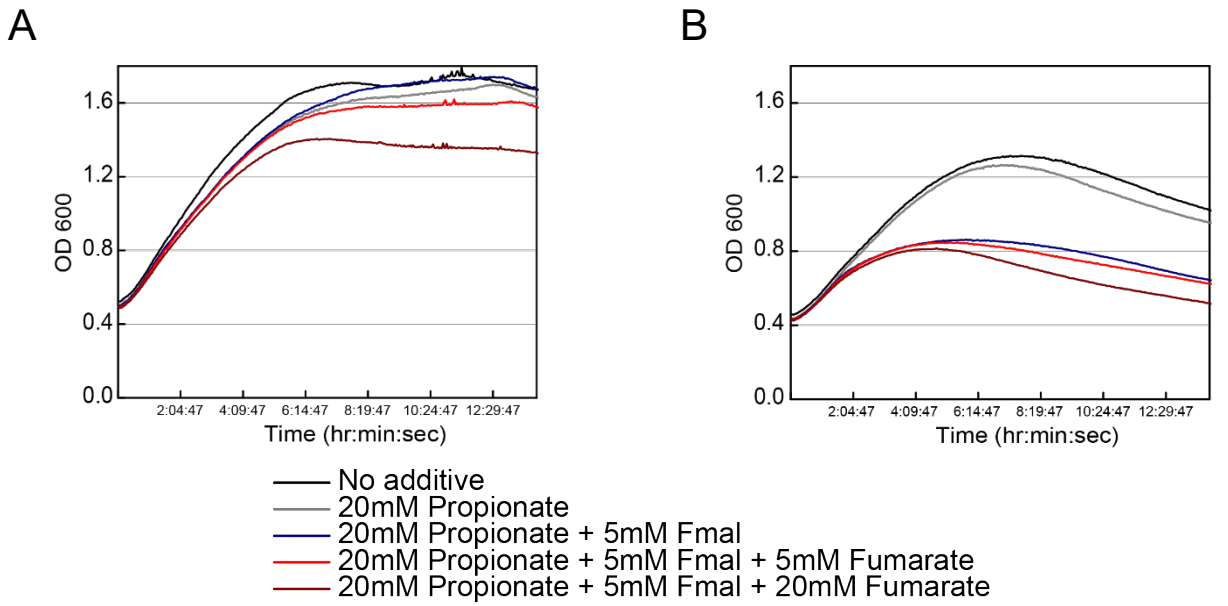


Figure 4.14. *Effects of malonate and fluoromalonate feeding on growth of E. coli BAP1 expressing Mod1 AT^0 DEBS pathway, MatB, and MadLM. E. coli BAP1 harboring pBP130 and pBP144_DEBS1(Mod1 AT^0) (A) without or (B) with pFmal(MadLM) were grown to OD_{600} = ~0.4 at 37 ºC with shaking at 200 rpm, at which point 100 µL of cultures were collected and supplemented with 1 mM IPTG, 0.2% arabinose, and when used 20 mM sodium propionate, 5 mM fluoromaloante, 5 or 20 mM fumarate. Growth was monitored at OD_{600} by microplate reader. Data are representative of 3 biological replicates.*

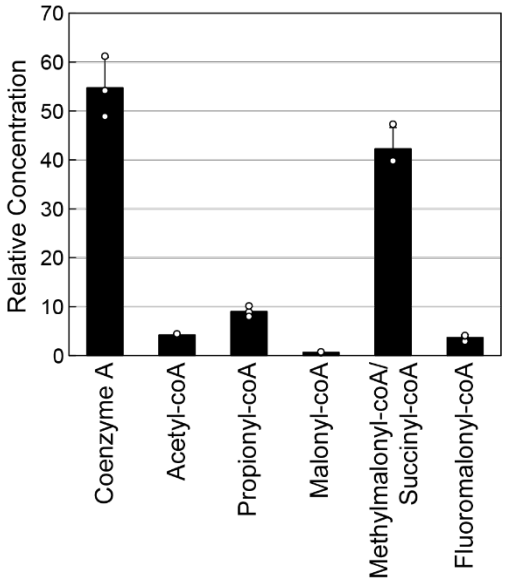


Figure 4.15. *Intracellular levels of acyl-CoA species of E. coli BAP1 expressing Mod6 AT^0 DEBS pathway, MatB, and MadLM. E. coli BAP1 harboring pBP130_DEBS3(Mod6 AT^0) and pBP144 with pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose, and cultures were provided with 20 mM sodium propionate and 5 mM fluoromalonate. Cultures were then incubated with substrates at 22 ºC with shaking at 250 rpm for 12 h, after which cells were harvested and acyl-CoA species were extracted and analyzed by LC-QQQ. Data are mean ± s.d of biological replicates (n = 3).*

To avoid toxicity, we looked for an alternative production procedure. To determine a possible improved procedure, we reasoned that a system competent to produce a fluorinated 6dEB requires: 1) expression of DEBS1-3 and PccAB and 2) accumulation of intracellular fluoromalonyl-CoA. We chose to use yield of 6dEB as a measure of the expression levels of the DEBS PKS and propionyl-CoA carboxylase proteins. Since fluoromalonyl-CoA causes a growth defect, we decided that the OD of the cultures should inversely correlate with the amount of intracellular fluoromalonyl-CoA. This assumption was supported by a fluoromalonate titration experiment that showed that feeding increasing amounts of fluoromalonate correlates with lowered OD values and decreased 6dEB product levels (*Figure 4.16*).

We then tested a variety of production procedures. We focused on altering the amount and timing of fluoromalonyl-CoA accumulation by changing the amount and timing of addition of propionate, fluoromalonate, and arabinose, which turns on the MadLM transporter expression. Under all experimental conditions, the precursors were added to the production cultures concomitant with arabinose induction. We found that when transporter induction and precursor feeding were delayed for 24 h after induction of DEBS1-3, increased level of 6dEB production and lowered culture OD were observed (*Figure 4.17*). We verified that the amount of intracellular fluoromalonyl-CoA was increased under this condition, relative to when precursors were added as all proteins were induced (*Figure 4.18*). As the machinery to produce 6dEB and intracellular fluoromalonyl-CoA are two parts of the system required to make fluorine-containing 6dEB analog, we found this culture condition to be promising for production. We also concentrated the biomass at the time of precursor feeding to 20 times the original cell density by harvesting and resuspending cells in 5% volume of fresh media. We reasoned that increasing cell density would effectively increase enzyme concentrations, leading to higher concentration of products. This new production procedure resembles more closely our fluorine-containing single-modular polyketide production system (*Section 3.2*).

Under the delayed precursor feeding condition, we found that cultures harboring the DEBS pathways with AT^0 mutation in either Module 5 or 6 each yielded a monofluorinated desmethyl-6dEB (*Figure 4.19*). We observed that the retention times of these molecules match those found produced *in vitro* with the mutation in the same module, suggesting that the regioselectivity of fluorine substitution is the same *in vivo* or *in vitro* (*Figure 4.19*). Thus, we believe the fluorodesmethyl-6dEB molecules observed from fermentation are the same as those observed *in vitro*.

To confirm the substitution pattern of the observed fluorodesmethyl-6dEB molecules, we performed fragmentation analysis on both molecules. We found that both molecules displayed the expected fragmentation pattern consistent with fluoromalonyl-CoA being incorporated by the module with inactivated AT domain (*Figure 4.20-4.21*). In other words, the observed fragmentation patterns were consistent with cultures harboring DEBS pathways with AT^0 mutation in Module 5 and 6 produced 4-fluoro-4-desmethyl-6dEB and 2-fluoro-2-desmethyl-6dEB respectively.

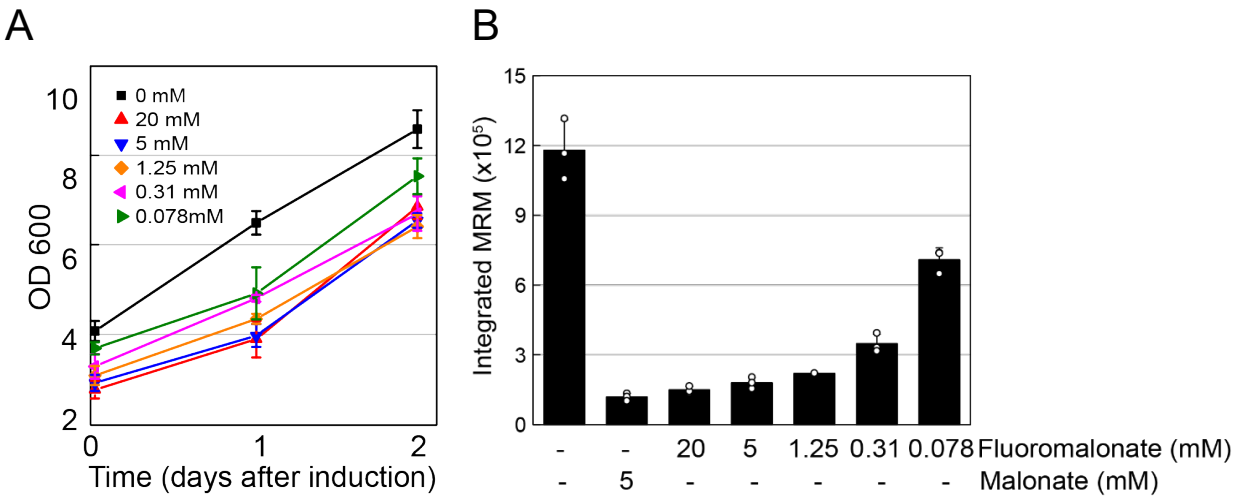


Figure 4.16. *Effects of the concentration of fluoromalonate in culture media on growth and production. (A) OD_{600} values of cultures of E. coli BAP1 expressing Mod6 AT^0 DEBS, MatB, and MadLM. E. coli BAP1 harboring pBP130 and pBP144_DEBS1(Mod1 AT^0), along with pFmal(MabLM), were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and 0.2% arabinose, and cultures were provided with 0, 0.078, 0.31, 1.25, 5, or 20 mM fluoromalonate and monitored OD_{600} for 2 days. Data are mean ± s.d of biological replicates (n = 3). (B) 6dEB production titers of respective cultures as in (A).*

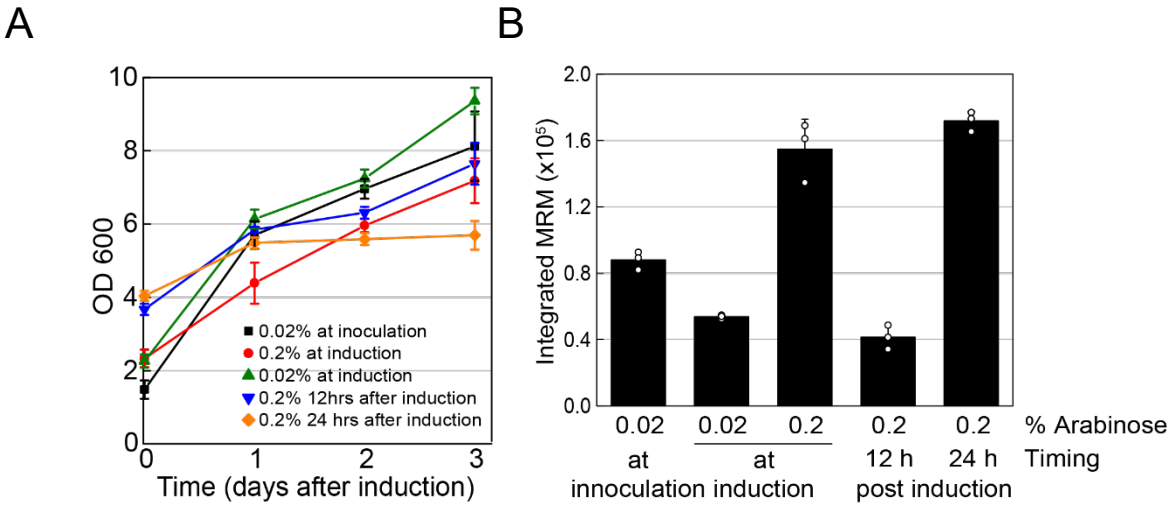


Figure 4.17. *Effects of fluoromalonate accumulation on growth and production (A) OD_{600} values of cultures of E. coli BAP1 expressing Mod6 AT^0 DEBS, MatB, and MadLM. E. coli BAP1 harboring pBP130 and pBP144_DEBS1(Mod1 AT^0), along with pFmal(MabLM), were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG. Arabinose inducer was added at different amounts and/or time. Cultures were provided with 5 mM propionate and 20 mM fluoromalonate at the time of arabinose addition and monitored OD_{600} for 3 days. Data are mean ± s.d of biological replicates (n = 3). (B) 6dEB production titers of respective cultures as in (A).*

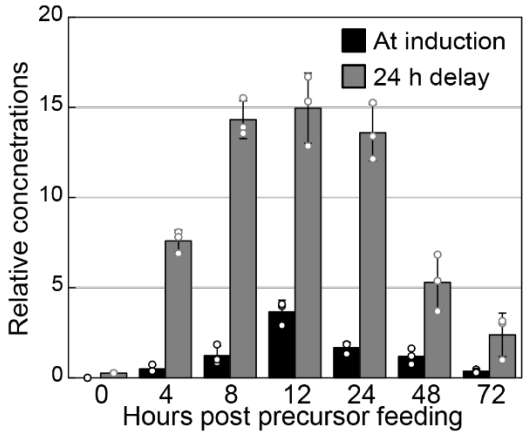


Figure 4.18. *Intracellular levels of acyl-CoA species of E. coli BAP1 expressing Mod6 AT^0 DEBS pathway, MatB, and MadLM. E. coli BAP1 harboring pBP130_DEBS3(Mod6 AT^0) and pBP144 with pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG. Cultures were provided with 0.2% arabinose, 20 mM sodium propionate and 5 mM fluoromalonate either at the time of DEBS protein induction (black) or 24 h post induction (grey). Cultures were incubated with substrates at 22 ºC with shaking at 250 rpm for 12 h, after which cells were harvested and acyl-CoA species were extracted and analyzed by LC-QQQ. Data are mean ± s.d of biological replicates (n = 3).*

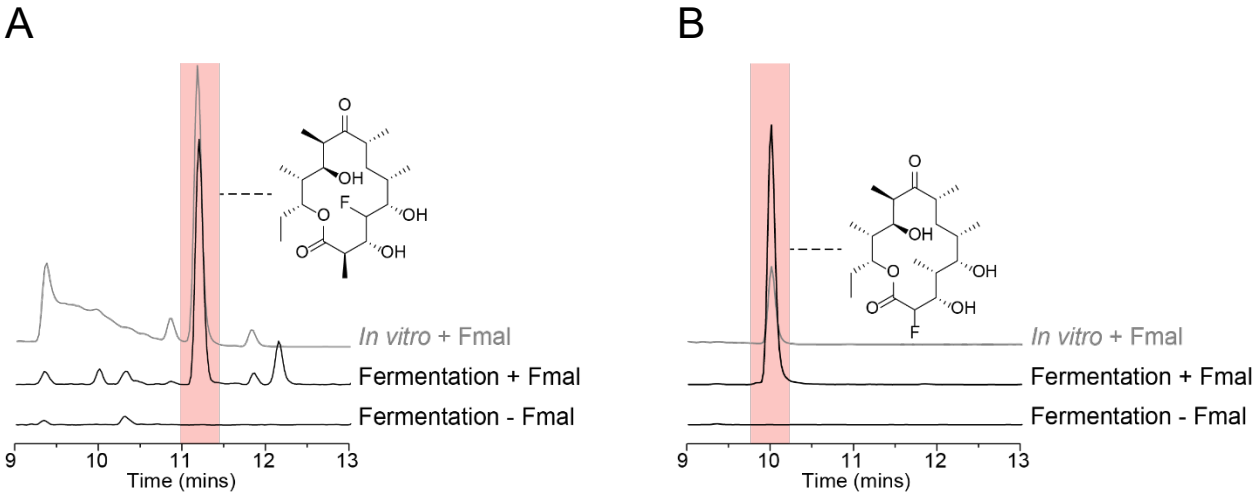
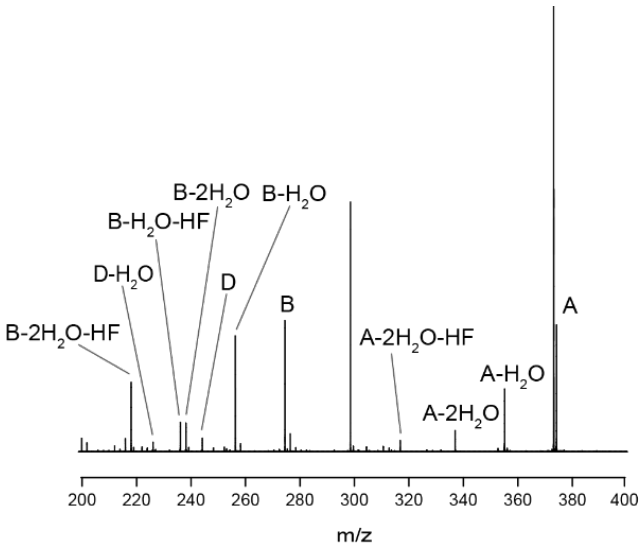


Figure 4.19. *Cellular production of fluorinated 6dEB analogs. Extracted ion chromatograms of cell culture extract of E. coli BAP1 expressing (A) Mod5 AT^0 DEBS and (B) Mod6 AT^0 DEBS for production of monofluorinated desmethyl 6dEB are shown. E. coli BAP1 harboring (A) pBP130_DEBS3(Mod5 AT^0) or (B) pBP130_DEBS3(Mod6 AT^0), pBP144, and pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate, and continued to grow at 22 ºC with shaking at 250 rpm for 24 h. Culture media were then collected and analyzed by LC-QQQ using MRM transition 373.2→275.1.*

A

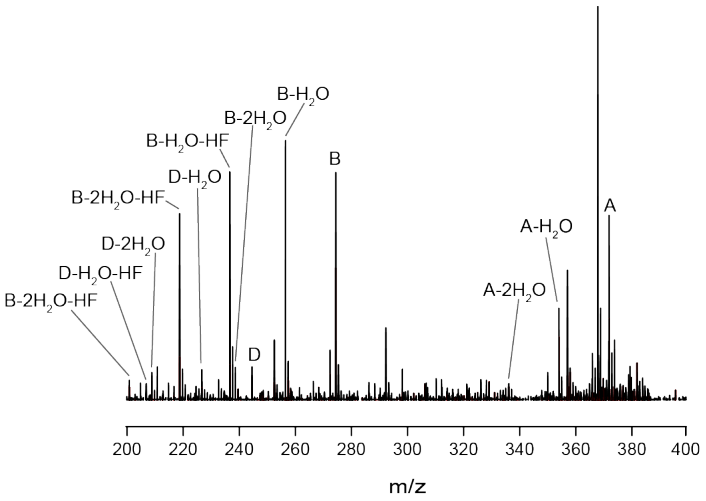


B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2366
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2268
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.217
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	317.2129
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1651
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1537
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.143
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.148
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.138	219.1373
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1259
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{18}FO_4^+$	245.1184	245.1157
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.107
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	Not observed
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	Not observed

Figure 4.20. *LC-QTOF fragmentation analysis of the monofluorinated desmethyl 6dEB analog produced by E. coli BAP1 expressing Mod5 AT^0 DEBS. E. coli BAP1 harboring pBP130_DEBS3(Mod5 AT^0), pBP144, and pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate. Culture media were then collected after 24 h, extracted, and analyzed by LC-QTOF (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 biological replicates.*

A



B

Ions	Molecular Formular	m/z Expected	m/z Observed
A	$C_{20}H_{34}FO_5^+$	373.2385	373.2386
A-H_2O	$C_{20}H_{32}FO_4^+$	355.2279	355.2284
A-2H_2O	$C_{20}H_{30}FO_3^+$	337.2173	337.2164
A-2H_2O-HF	$C_{20}H_{29}O_3^+$	317.2111	Not observed
B	$C_{14}H_{24}FO_4^+$	275.1653	275.1653
B-H_2O	$C_{14}H_{22}FO_3^+$	257.1547	257.1547
B-2H_2O	$C_{14}H_{20}FO_2^+$	239.1442	239.1451
B-H_2O-HF	$C_{14}H_{21}O_3^+$	237.1485	237.1487
B-2H_2O-HF	$C_{14}H_{19}O_2^+$	219.138	219.1377
B-3H_2O-HF	$C_{14}H_{17}O^+$	201.1274	201.1289
C	$C_{15}H_{27}O_2^+$	239.2006	Not observed
C-H_2O	$C_{15}H_{25}O^+$	221.19	Not observed
D	$C_{12}H_{18}FO_4^+$	245.1184	245.118
D-H_2O	$C_{12}H_{16}FO_3^+$	227.1078	227.1074
D-2H_2O	$C_{12}H_{14}FO_2^+$	209.0972	209.1535
D-H_2O-HF	$C_{12}H_{15}O_2^+$	207.1016	207.102

Figure 4.21. *LC-QTOF fragmentation analysis of the monofluorinated desmethyl 6dEB analog produced by E. coli BAP1 expressing Mod6 AT^0 DEBS. E. coli BAP1 harboring pBP130_DEBS3(Mod6 AT^0), pBP144, and pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate. Culture media were then collected after 24 h, extracted, and analyzed by LC-QTOF. (A) Fragmentation pattern shows presence of ions in A, B, and D families. (B) Tabulated expected and observed masses of daughter ions in all ion families. Data are representative of 3 biological replicates.*

We scaled up the production of 2-fluoro-2-desmethyl 6dEB in order to obtain enough materials to characterize by ^{19}F-NMR. Spectrum of crude culture extract revealed a set of signals at -194 to -195 ppm that showed a doublet-of-doublet splitting pattern consistent with that expected of 2-fluoro-2-desmethyl 6dEB (*Figure 4.22*). The chemical shift value is similar to that observed of other compounds with similar α-fluoro-β-hydroxy ester motif [35].

Variability of fluorodesmethyl-6dEB production. To determine the overall selectivity of the system, we aimed to compare the amount of 2-fluoro-2-desmethyl-6dEB to that 2-desmethyl-6dEB and 6dEB molecules produced by the same production culture. However, in this process, we found that the distribution of the three products observed from cells harboring the DEBS pathway with a AT^0 mutation in Module 6 and pFmal(MadLM) vary drastically among different transformants (*Figure 4.23*).

SDS-PAGE analysis of these cultures revealed that the expression level of MatB enzyme differs among cultures and correlates with product profile. MatB seems to be downregulated in cultures that produce high amounts of 6dEB and lower MRM signals of 2-fluoro-2-desmethyl 6dEB (*Figure 4.24*). This observation is consistent with the role of MatB in our system, which is to accumulate the fluoromalonyl-CoA precursor. Thus, lowered MatB expression possibly leads to decreased levels of intracellular fluoromalonyl-CoA and consequently decreased 2-fluoro-2-desmethyl-6dEB levels. This observation is also consistent with our previous observation that feeding fluoromalonate to cells expressing MatB and MadLM negatively affected cell growth. Although the nature of toxicity remains to be investigated in detail, it would not be surprising if the cells downregulate the expression of MatB and MadLM to alleviate toxicity.

Improving system selectivity. Despite large variations among transformants, we aimed to improve the selectivity of the system towards 2-fluoro-2-desmethyl-6dEB production. We attempted three strategies to overcome the undesired selectivity of FabD.

In the first strategy, we overexpressed F190V DszAT along with the DEBS pathway, MatB, and MadLM. *E. coli* BAP1 was co-transformed with pBP130_DEBS3(Mod6 AT^0), pBP144, pFmal(MadLM), and p15A-DszATF190V (*section 2.2*). We reasoned that as F190V DszAT was shown to be able to suppress malonyl-CoA incorporation and promote fluoromalonyl-CoA incorporation, a high expression level of F190V DszAT could decrease the effects of malonyl-CoA and methylmalonyl-CoA incorporation mediated by FabD. We found the distribution of products among different transformants vary greatly, similarly to cultures without F190V DszAT overexpression (*Figure 4.25*). We also observed that overexpressing F190V DszAT resulted in a marked decrease in the number of cultures the produced high MRM signals of 6dEB relative to 2-fluoro-2-desmethyl 6dEB. Correspondingly, the number of cultures that down regulated MatB expression was fewer compared to when no DszAT was expressed. These results suggest that overexpressing F190V DszAT may led to an overall increase in relative amount of 2-fluoro-2-desmethyl 6dEB compared to 6dEB.

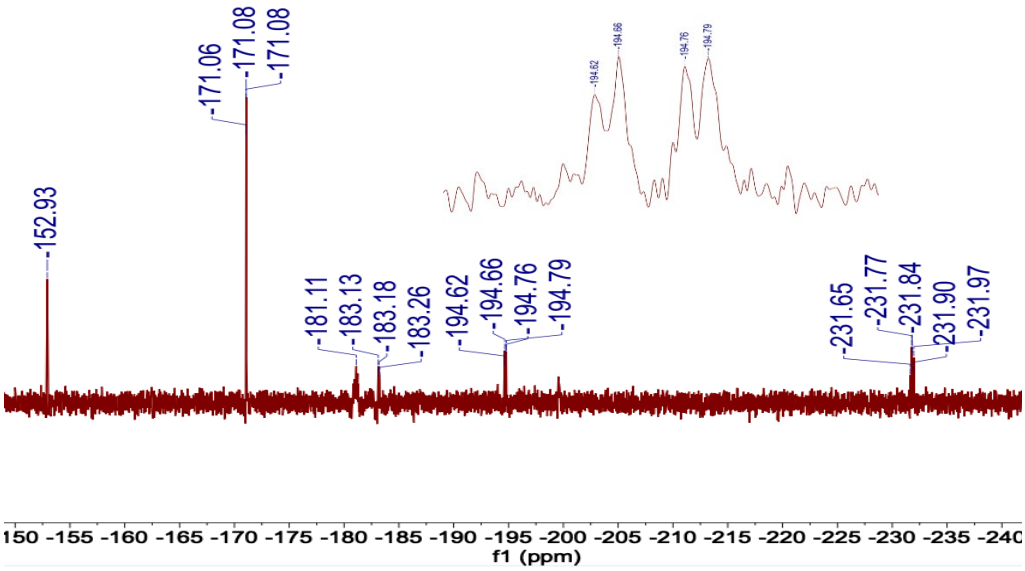


Figure 4.22. *^{19}F-NMR of extract of culture media of E. coli BAP1 expressing Mod6 AT^0 DEBS. Large-scale cultures of E. coli BAP1 harboring pBP130_DEBS3(Mod6 AT^0) and pBP144, along with pFmal(MadLM) were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate, and continued to grow at 22 ºC with shaking at 250 rpm for 24 h. Culture media were then collected, extracted, and analyzed by 19F-NMR in MeOH/D_2O.*

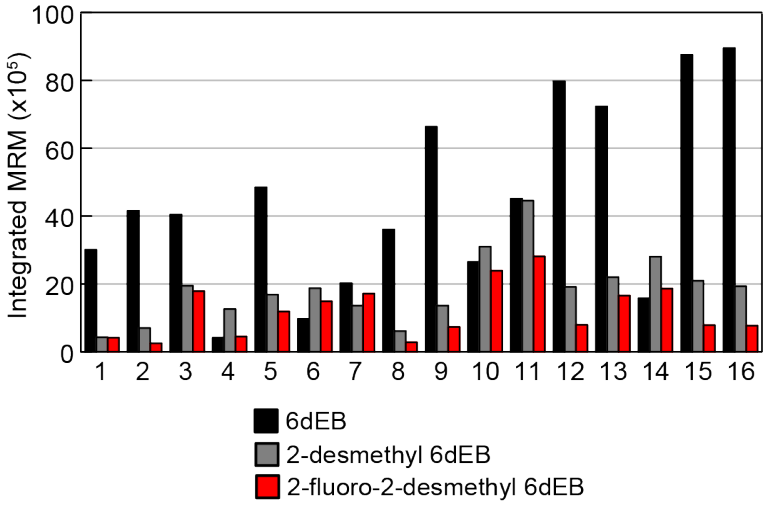


Figure 4.23. *Products of independent transformants of E. coli expressing Mod6 AT^0 DEBS. Relative signals of 6dEB (black), 2-desmethyl 6dEB (grey), and 2-fluoro-2-desmethyl 6dEB (red) produced by 16 independent transformants of E. coli BAP1 transformed with pBP130_DEBS3(Mod6 AT^0), pBP144, and pFmal(MadLM) are shown. Cells were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate, and continued to grow at 22 ºC with shaking at 250 rpm for 24 h. Culture media were then collected and analyzed by LC-QQQ.*

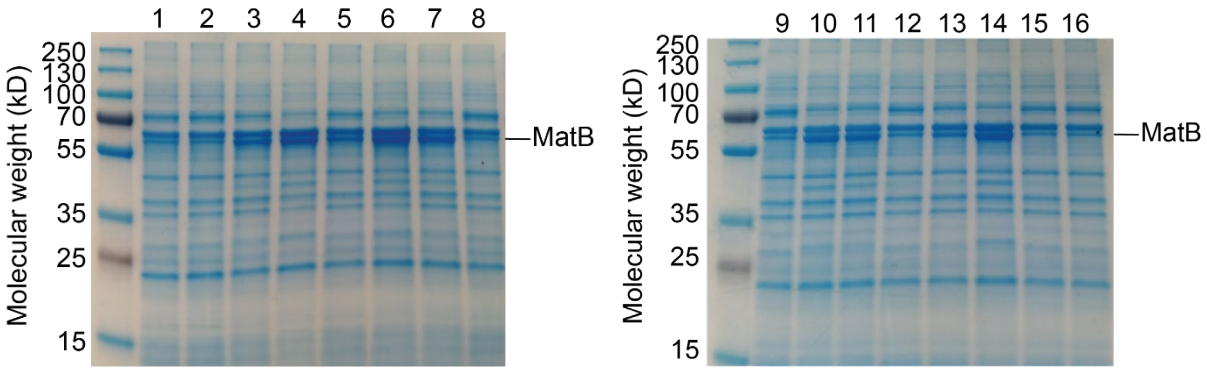


Figure 4.24. *Total protein analysis of independent transformants of E. coli expressing Mod6 AT^0 DEBS. E. coli BAP1 were transformed with pBP130_DEBS3(Mod6 AT^0), pBP144, and pFmal(MadLM). Cells were grown to OD$_{600}$ = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate, and continued to grow at 22 ºC with shaking at 250 rpm for 24 h. Cells (0.5 mL) were then collected and analyzed by SDS-PAGE.*

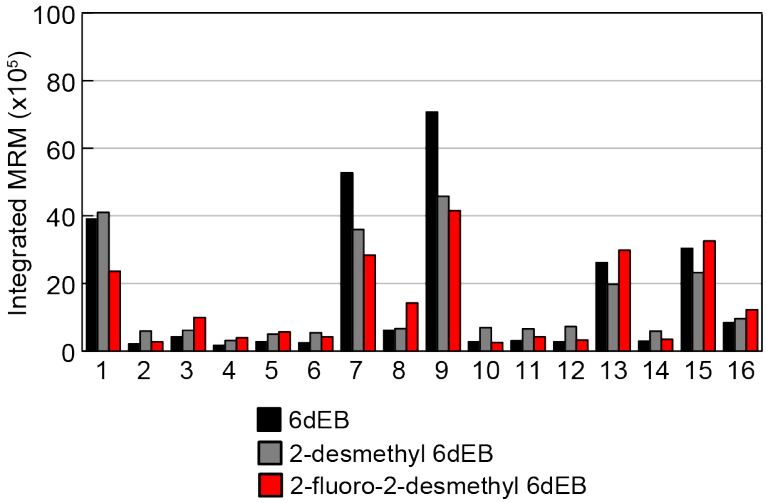


Figure 4.25. *Products of independent transformants of E. coli expressing Mod6 AT^0 DEBS and F190V DszAT. Relative signals of 6dEB (black), 2-desmethyl 6dEB (grey), and 2-fluoro-2-desmethyl 6dEB (red) produced by 16 independent transformants of E. coli BAP1 transformed with pBP130_DEBS3(Mod6 AT^0), pBP144, pFmal(MadLM), and p15A-F190VDszAT. Cells were grown to OD$_{600}$ = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate, and continued to grow at 22 ºC with shaking at 250 rpm for 24 h. Culture media were then collected and analyzed by LC-QQQ.*

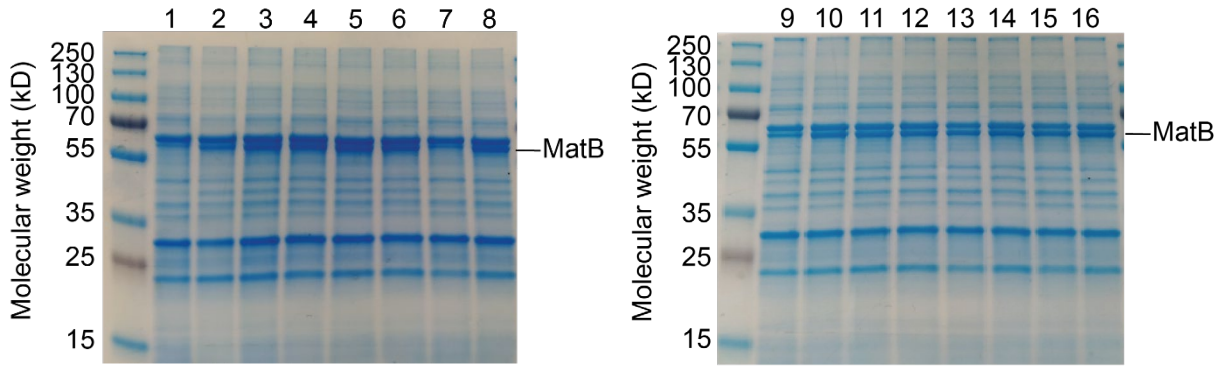


Figure 4.26. *Total protein analysis of independent transformants of E. coli expressing Mod6 AT^0 DEBS and F190V DszAT. E. coli BAP1 were transformed with pBP130_DEBS3(Mod6 AT^0), pBP144, pFmal(MadLM), and p15A-F190VDszAT. Cells were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate, and continued to grow at 22 ºC with shaking at 250 rpm for 24 h. Cells (0.5 mL) were then collected and analyzed by SDS-PAGE.*

In the second strategy, we aimed minimize FabD's effect on our engineered system. As FabD is an essential protein for E. coli [36] simple gene knockout is not possible. Thus, we aimed to lower the expression of fabD gene. FabD is located in the fab-acpP locus, which includes the fatty acid synthesis genes fabH, fabD, fabG, acpP, and fabF as well as three upstream genes plsX, rpmF, and yceD. The FabH promoter (FabHp) drives the expression of fabD along with fabH and fabG and is activated by the FadR regulator (Figure 4.27) [37]. In the absence of FadR, the activity of FabHp has been shown to be decreased [37]. Thus, in an effort to lower FabD expression level, we deleted fadR from E. coli BAP1 using a previously developed CRISPR/Cas system for E. coli genome editing [24]. We then attempted to produce 2-fluoro-2-desmethyl 6dEB using E. coli BAP1ΔfadR strain expressing the Mod6 AT^0 DEBS pathway, MatB, MadLM, and DszAT F190V. However, we found that the amount of 6dEB produced by E. coli BAP1ΔfadR was also much lower than that observed produced by the parental E. coli BAP1, and no 2-fluoro-2-desmethyl 6dEB was observed (Figure 4.28), suggesting an overall defect in polyketide production. Additional attempts were made to introduce a temperature sensitive mutation into FabD by replacing tryptophan at position 257 with glutamine residue using CRISPR/Cas system [38], however, the correct mutant was not isolated.

In the last strategy, we looked to alter the specificity of FabD protein. We mutated serine at position 200, which is thought to play a role in determining substrate selectivity based on alignment studies [39, 40]. The small size of S200 is expected to allow the protein to accept methylmalonyl-CoA substrate. Thus, we reasoned that mutating S200 to larger residues such as valine and phenylalanine could suppress FabD's activity on methylmalonyl-CoA, leading to higher incorporation of fluoromalonyl-CoA and thus more selective production of 2-fluoro-2-methyl 6dEB. As FabD is an essential protein and changes made to the protein might affect cell growth, we decided to forgo *in vitro* activity testing, but to directly introduce the mutations into genome and assess production. We found that while the serine to valine mutant was readily obtained, the serine to phenylalanine mutant exhibited some growth defects. As growth rates varied among different colonies of the same mutant, the fastest growing colony was selected for production. We found that both *E. coli* BAP1 *fabD*(S200V) and *E. coli* BAP1 *fabD*(S200F) expressing DEBS

pathway with active site mutation in the AT domain of Module 6, MatB, MadLM, and F190V DszAT produced very low levels 6dEB and no observable 2-fluoro-2-desmethyl 6dEB, similar to the Δ*fadR* strain (*Figure 4.28*).

To improve the selectivity of the *E. coli* production system of fluorine-containing polyketides through regioselective incorporation of fluoromalonyl-coA requires further investigation into understanding the role of FabD in complementing AT^0 mutations in co-expressed PKS systems as well as cross-talk between fatty acid and polyketide production. In addition, an understanding how fluoromalonate and fluoromalonyl-CoA impact *E. coli* physiolology could also aid in building a system with improved growth and function, while also increasing fluorine incorporation and system selectivity.

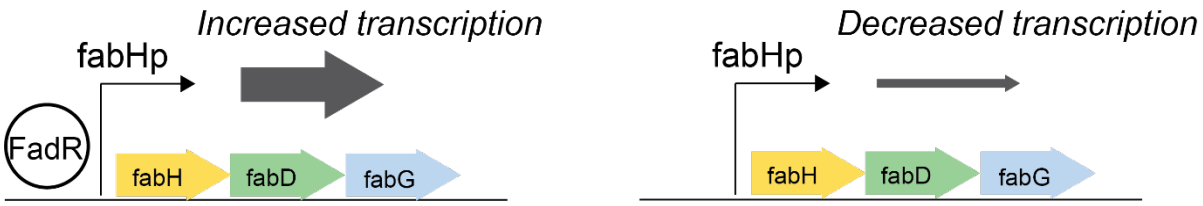


Figure 4.27. *Regulation of fabHDG operon by FadR protein. The activity of fabH promoter is regulated by FadR activator. In the presence of FadR protein, transcription of fabDHG open increases.*

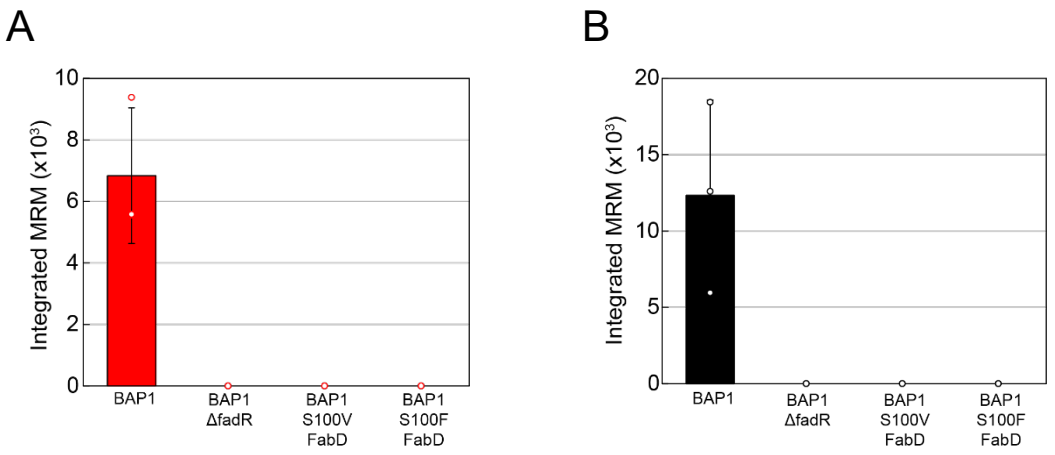


Figure 4.28. *6dEB analog production in E. coli BAP1 host variants. Relative amounts of (A) 6dEB and (B) 2-fluoro-2-desmethyl 6dEB produced by E. coli BAP1, BAP1 ΔfadR, BAP1 S100V FabD, and BAP1 S100F FabD harboring pBP130_DEBS3(Mod6 AT^0), pBP144, pFmal(MadLM), and p15A-F190VDszAT are shown. Cells were grown to OD_{600} = 0.4-0.6 at 37 ºC with shaking at 200 rpm, at which point protein expression was induced with 1 mM IPTG and cultures were grown at 22 ºC with shaking at 250 rpm. After 24 h, cultures were concentrated, provided with 0.2% arabinose, 20 mM sodium propionate, and 5 mM fluoromalonate, and continued to grow at 22 ºC with shaking at 250 rpm for 24 h. Culture media were then collected, extracted, and analyzed by LC-QQQ. Data are mean ± s.d of biological replicates (n = 3).*

Design and construction of fluorinated 6dEB biosynthetic pathway in a streptomycete host. With the initial success in the *E. coli* production system, we aimed to explore the possibility of developing a Streptomyces production system for fluorine-containing polyketide analogs as a complement to our *E. coli* system as many PKSs are derived from actinomycetes. Towards this goal, we devised two cloning strategies in order to introduce active-site mutations into polyketide biosynthetic gene cluster for expression in *Streptomyces* spp.

The first strategy makes use of pCRISPomyces-2, a CRISPR/Cas genome editing system that had been previously developed to introduce scar-less mutations in Streptomyces genomes (*Figure 4.29*) [41]. The pCRISPomyces-2 plasmid encodes for the Cas9 nuclease, single guide RNA (sgRNA), and repair template. A protospacer sequence, which is a 20-nucleotide genomic sequence that the Cas9 and sgRNA complex will target, is inserted to sgRNA via Golden gate assembly using BbsI restriction enzyme. It requires a protospacer adjacent motif (PAM) NGG, where N is any nucleotide, at its 3′-end. The homology-directed repair (HDR)-driven repair template can be inserted via Gibson assembly or simple restriction–ligation cloning using XbaI. The repair template is required for introducing defined mutations and prevent repair by non-homologous end joining (NHEJ), which introduces random deletions at the site of the double-stranded break created by Cas9. To introduce defined mutations with high efficiency, a 4-8 kb editing template is used to allow HDR between the chromosomal locus and the pCRISPomyces-2 plasmid. The repair template consists of the desired mutation and 2 homology arms of 2-4 kbp each, corresponding to upstream and downstream regions of the protospacer sequence. The repair template should also include a mutation to the PAM sequence to prevent Cas9 from further targeting the edited sequence. This strategy would be most suitable for pathways that have been previously validated in a streptomycete host but needs point mutations installed.

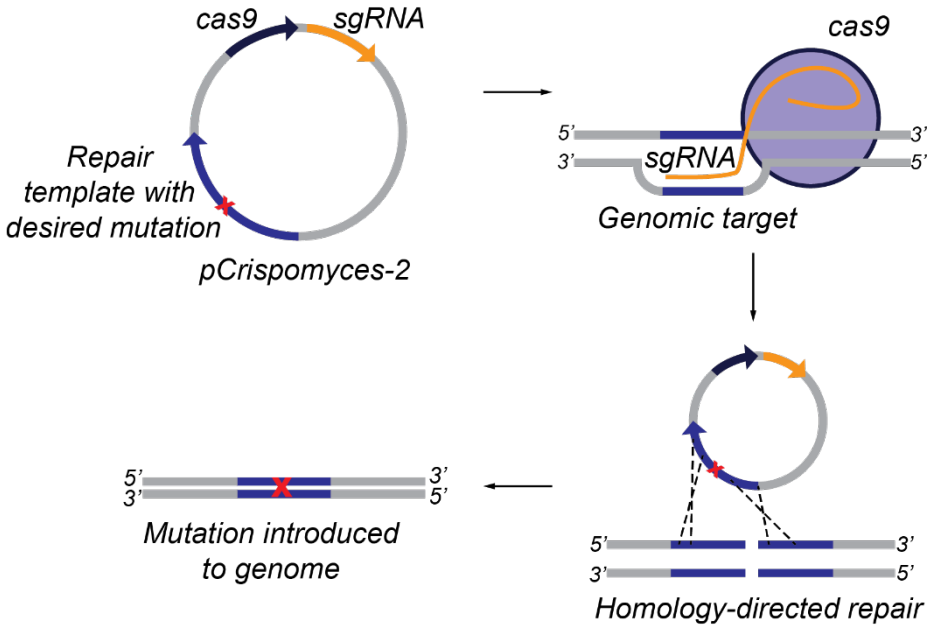


Figure 4.29 *Cloning strategy to introduce active-site mutation to AT domains of DEBS pathway based on pCrispomyces-2 editing system. Guide sequence and repair template are introduced to pCrispomyces-2, which is conjugated into Streptomyces host to perform editing of biosynthetic pathway previously integrated.*

We attempted to build a pCRISPomyces-2 construct targeting the active-site serine of the AT domain of DEBS Modules 1, 3, and 6. Although we successfully added protospacer sequences to the plasmids, we found constructing the repair templates to be tedious and inefficient. Initially, we attempted to introduce the repair templates through one-step Gibson assembly; however, the desired product was not achieved, likely due to the low efficiency of Gibson assembly at large plasmid sizes. We then turned to stepwise insertions of smaller fragments of repair template. However, successive rounds of performing and screening products of low efficiency Gibson assembly proved labor-intensive and time-consuming.

An alternative cloning strategy that we devised to avoid the tedious step of constructing pCRISPomyces-2 plasmids makes use of transformation-associated recombination (TAR) and CRISPR/Cas editing system in the yeast, *Saccharomyces cerevisiae*. In the first step, the biosynthetic gene cluster of interest would be assembled into a shuttle vector that is compatible with *E. coli*, *S. cerevisiae,* and *Streptomyces* spp. through TAR in yeast (*Figure 4.30*). Following the assembly step, gene editing would be performed in yeast to introduce the desired mutations into the pathway (*Figure 4.30*). Editing in *S. cerevisiae* host is preferred over in *E. coli* or *Streptomyces* spp. due to higher recombination efficiency and accuracy in *S. cerevisiae*. Then, the plasmid encoding for pathway genes with desired mutations would be isolated from yeast, transformed into *E. coli*, and conjugated into the Streptomyces host for genomic integration (*Figure 4.30*).

We chose the pCap01 shuttle vector for this purpose as it is single-copy number in *S. cerevisiae*, high-copy number in *E. coli*, and competent for conjugation and integration in streptomycetes [25]. The pCap01 plasmid has a CEN6/ARS4 origin and is thus a yeast artificial chromosome (YAC) [25]. It also consists of pUC origin for high-copy maintenance in *E. coli*, the RP4 origin of transfer for conjugation into streptomycetes, and the Φc31 *attP* site for chromosomal integration of the plasmid [25]. Capture arms that consist of sequences homologous to the 5′- and the 3′-ends of the gene cluster to be assembled into pCap01 need to be pre-installed on the plasmid to provide sites of homologous recombination, which can be done in the XhoI-KpnI site (*Figure 4.30*). Each capture arm can range in size from 10^1-10^3 bp, but it is thought that longer capture arms increase the efficiency of the assembly. Linearized pCap01 with specific capture arms for the cluster of interest can then be co-transformed with linearized gene cluster into yeast with the desired colonies selected on tryptophan drop-out medium.

We constructed the pCap01_DEBSCaptureArms plasmid, which is pCap01 vector equipped with specific capture arms for the DEBS cluster and its actinorhodin promoter (Actp) from pCK7, which encodes for DEBS1-3 under the control of Actp [23]. pCK7 possess a low-copy number truncated Streptomyces replicon, SCP2*, as well as a ColE1 replicon for genetic manipulation in *E. coli*. We used pCK7 linearized by HindIII and EcoI as a source of the DEBS and promoter sequences. After co-transforming both pieces of DNA into *S. cerevisiae* VL6-48 and selecting on tryptophan drop-out media, the resulting yeast colonies were grown and total DNA was extracted. However, we found that transformation of *E. coli* using total DNA extracted from yeast did not yield the desired construct. The number of resulting *E. coli* colonies from total yeast DNA transformation were usually small (<50 colonies/transformation) with none carrying the correct construct. We hypothesized that our initial selection and isolation procedure was not efficient enough due to the size of the plasmid and the repetitive nature of DEBS sequence. Thus, we devised an alternative assessment and isolation strategy to pre-screen *S. cerevisiae* transformants that allowed for extensive quality control throughout the process (*Figure 4.31*).

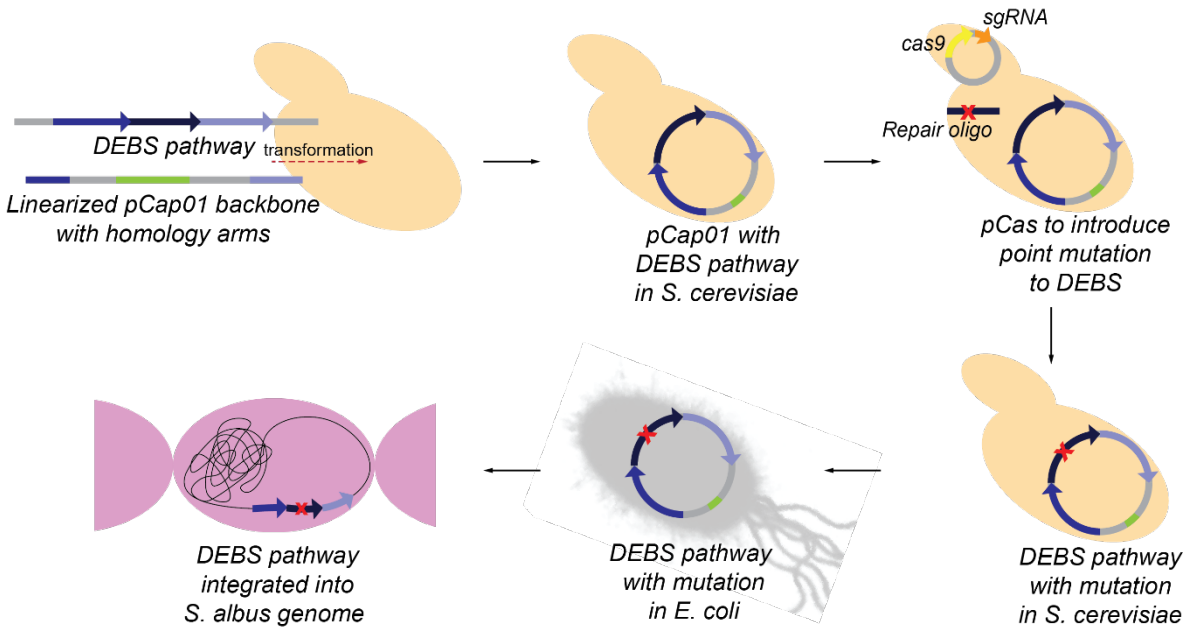


Figure 4.30 *Cloning strategy to introduce* an *active-site mutation* into *AT domains of* the *DEBS* PKS *based on TAR and CRISPR/Cas editing in S. cerevisiae.* The *DEBS pathway was assembled into yeast-compatible backbone through TAR and edited* using the *CRISPR/Cas platform. Pathways with correct mutations were isolated, transformed into E. coli, and conjugated into S. albus for genomic integration.*

In order to screen *S. cerevisiae* colonies that resulted from TAR assembly, we extracted total DNA from single yeast colonies and amplified representative regions of the plasmid. We designed primers to amplify 700-1000 bp from regions of regular intervals and used them to assess the completeness of the resulting plasmid assembly (*Figure 4.31*). We found that many yeast colonies, although TRP^+, contained plasmids that were missing large parts of the DEBS sequences. Thus, the additional PCR screen step allowed us to identify yeast colonies that were most likely to contain the correct plasmid assembly.

Yeast colonies that showed promise were expanded to large culture volume. We reasoned that transformation of total DNA extract from yeast did not yield many *E. coli* transformants because most of the DNA was chromosomal. In order to increase transformation efficiency, we opted to isolate our plasmid assembly from yeast chromosomal DNA. As the assembly product was a circular YAC of single-copy number, we needed a large starting culture volume to ensure enough plasmid and followed a previously developed isolation protocol that took advantage of the circular nature of YAC to isolate it from chromosomal DNA fragments [30]. Following large-scale isolation, YAC was transformed into *E. coli* for maintenance and conjugation (*Figure 4.31*). We found that electroporation was more efficient while chemical transformation suffered from bias against larger plasmid sizes. The assembly product was isolated from *E. coli* using a standard plasmid miniprep kit and assessed by restriction analysis. We found that individual transformants gave plasmid of different restriction patterns despite having been transformed with YAC isolated from a single yeast colony. This finding emphasizes the importance of extensive quality control procedure at every step in this process.

With the isolation and assessment procedure described here, we were able to assemble DEBS pathway and its actinorhodin promoter into pCap01 backbone to yield pCap01_DEBS. Following conjugation into *Streptomyces albus* J1074, we assessed the strain's production capacity in several commonly used growth media (*Figure 4.32*), in comparison to *Streptomyces lividans* harboring pCK7. We found that YEME and 1:1 YEME:TSB media gave the highest yield. We did find that *S. lividans* harboring pCK7 overall gave higher yield than *S. albus* J1074 with DEBS pathway integrated into the genome. We attributed the lower yield to the lower copy number of the pathway gene in *S. albus*. However, as *S. albus* has faster growth kinetics and smaller genome, and as genomically integrated pathway is more stable, *S. albus* may still offer some advantages for fluorinated polyketide production.

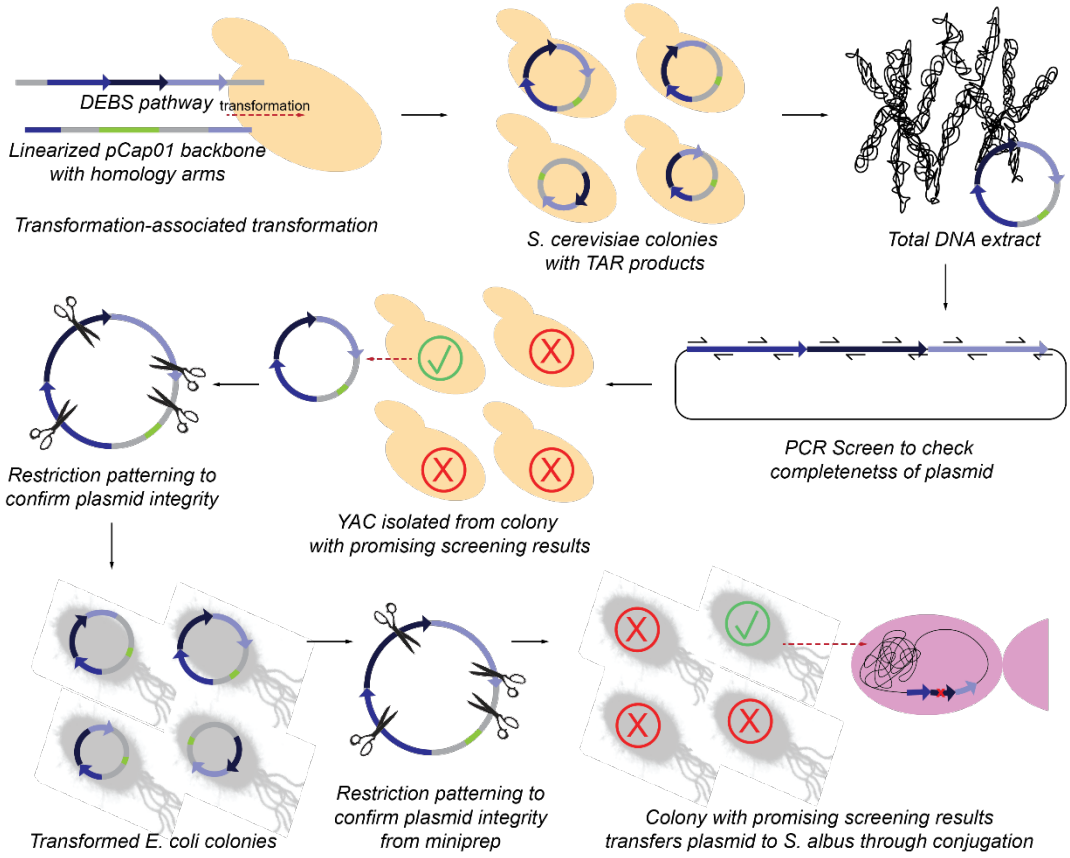


Figure 4.31 *YAC screening and isolation procedure to identify correct construct. This process includes PCR screen of total DNA extract from S. cerevisiae to identify correct clone before YAC isolation. Following transformation into E. coli, miniprep of E. coli transformants were screened by restriction patterning to determine correct clone before cull plasmid sequencing and conjugation into streptomyces.*

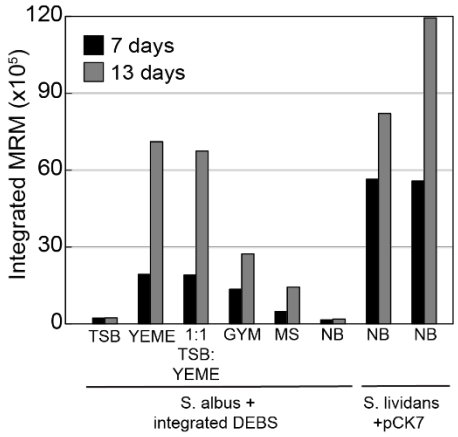


Figure 4.32 *6dEB production by Streptomyces albus J1074 with genomically integrated DEBS pathway in common laboratory media (TSB = tryptic soy broth; YEME = yeast extract, malt extract; GYM = glucose, yeast extract, malt extract; MS = soy, mannitol; NB = nutrient broth) supplemented with 5 mM sodium propionate in comparison to by Streptomyces lividans K4-114 harboring pCK7 at 7 (black) and 13 (grey) days after inoculation.*

In order to introduce active-site mutations into the *cis*-AT domains of DEBS proteins on pCap01_DEBS, we utilized a previously reported CRISPR/Cas genome editing platform using the pCas_PPhe-BsaI_NAT plasmid for *S. cerevisiae* to edit DEBS pathway encoded on the pCap01 YAC backbone [26]. The system was originally developed to edit the yeast genome, and to our knowledge, there was no previous report of CRISPR/Cas-mediated editing of artificial chromosomes. However, the system had been shown to be effective for editing polyploid yeast strains [26]. Thus, we reasoned that the platform should be applicable to yeast artificial chromosome, which should be maintained at single or low copy number.

We designed and constructed a few guide sequences for the AT domains of DEBS Modules 1, 2, 3, and 6. Through this editing approach and the described screening and isolation method, we successfully introduced the desired active-site mutation into the AT domain of DEBS Module 1. To check for successful editing, the target region was amplified from total yeast DNA extract and sequenced. Colonies with correctly edited plasmid that also showed completeness in PCR screen were expanded for YAC isolation. The resulting plasmid containing DEBS pathway with inactivating mutation in the AT domain of Module 1 was conjugated into *S. albus* J1074. However, no fluorine-containing 6dEB analog was found as a result of production.

As 6dEB had been observed in *S. albus* harboring the wild-type pathway, the DEBS proteins appear to be functionally expressed and methylmalonyl-CoA precursor is available. The absence of fluorinated 6dEB analogs suggest that further optimization is required. One important question is whether fluoromalonyl-CoA is available at a sufficient concentration, which could be measured using our acyl-CoA LC-MS/MS method. However, an alternative cell lysis and metabolite extraction method likely needs to be developed, as *S. albus* is a gram-positive bacterial strain that would withstand simple cell lysis method effective for *E. coli*. Another future direction would be to examine DEBS pathways with AT^0 mutation in the Module 5 or 6 AT domains, which have been shown to produce 4-fluoro-4-desmethyl 6dEB and 2-fluoro-2-desmethyl 6dEB respectively in *E. coli*. In addition to these studies, it will also be interesting to explore whether other

endogenous transacylation enzymes, whether housekeeping or involved in secondary metabolism, might be able to complement the AT^0 mutations in DEBS.

4.4 Conclusions

Work in this chapter expanded on that of the previous chapter to demonstrate *in vivo* fluorine incorporation into a model multi-modular polyketide molecule. We demonstrated a successful regiospecific incorporation of fluoromalonyl-CoA by mutant DEBS pathways to produce 2-fluoro-2-desmethyl 6dEB and 4-fluoro-4-desmethyl 6dEB analogs in an engineered *E. coli* host. The identity of these products has been confirmed by LC-MS/MS by comparison to molecules produced in *in vitro* enzymatic reactions. In the process, we investigated the toxicity of fluoromalonyl-CoA transport and activation system and its effects on fluorine-containing polyketide production. Thus, we were able to develop production procedure that minimized the negative toxicity effects. In addition, we started to develop a fluorine incorporation system to be used in a Streptomyces host. We explored different genetic manipulation methods necessary for introducing active site mutation to the AT domains of PKS, which is a necessary component for fluoromalonyl-CoA incorporation. Future work will continue to build other parts of the system to continue to tune extender unit pools as well as to better understand the interplay between our engineered system and endogenous metabolism.

4.5 References

1. Gao, X.; Wang, P.; Tang, Y. Engineered polyketide biosynthesis and biocatalysis in Escherichia coli. *Appl Microbiol Biotechnol* **2010***, 88*, 1233-1242.

2. Ahmadi, M. K.; Pfeifer, B. A. Recent progress in therapeutic natural product biosynthesis using Escherichia coli. *Current Opinion in Biotechnology* **2016***, 42*, 7-12.

3. Morgado, G.; Gerngross, D.; Roberts, T. M.; Panke, S. In *synthetic biology for cell-free biosynthesis: fundamentals of designing novel in vitro multi-enzyme reaction networks;* Zhao, H., Zeng, A., Eds.; Synthetic Biology – Metabolic Engineering; Springer International Publishing: Cham, 2018; pp 117-146.

4. Chao, R.; Yuan, Y.; Zhao, H. Recent advances in DNA assembly technologies. *FEMS Yeast Res.* **2015***, 15*, 1-9.

5. Zhang, J. J.; Tang, X.; Moore, B. S. Genetic platforms for heterologous expression of microbial natural products. *Nat. Prod. Rep.* **2019***, 36*, 1313-1332.

6. 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*, 3796-3806.

7. Musiol-Kroll, E. M.; Wohlleben, W. Acyltransferases as tools for polyketide synthase engineering. *Antibiotics (Basel)* **2018***, 7*.

8. 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*, 1790-1792.

9. Boghigian, B. A.; Zhang, H.; Pfeifer, B. A. Multi-factorial engineering of heterologous polyketide production in Escherichia coli reveals complex pathway interactions. *Biotechnol. Bioeng.* **2011***, 108*, 1360-1371.

10. 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*, 500-509.

11. Lee, H. Y.; Harvey, C. J. B.; Cane, D. E.; Khosla, C. Improved precursor-directed biosynthesis in E. coli via directed evolution. *The Journal of antibiotics* **2011***, 64*, 59-64.

12. Myronovskyi, M.; Luzhetskyy, A. Heterologous production of small molecules in the optimized Streptomyces hosts. *Nat Prod Rep* **2019***, 36*, 1281-1294.

13. Tan, G.; Deng, K.; Liu, X.; Tao, H.; Chang, Y.; Chen, J.; Chen, K.; Sheng, Z.; Deng, Z.; Liu, T. Heterologous biosynthesis of spinosad: an omics-guided large polyketide synthase gene cluster reconstitution in streptomyces. *ACS Synth. Biol.* **2017***, 6*, 995-1005.

14. Liu, X.; Liu, D.; Xu, M.; Tao, M.; Bai, L.; Deng, Z.; Pfeifer, B. A.; Jiang, M. Reconstitution of kinamycin biosynthesis within the heterologous host Streptomyces albus J1074. *J. Nat. Prod.* **2018***, 81*, 72-77.

15. Bilyk, O.; Luzhetskyy, A. In *host organism: streptomyces;* Industrial Biotechnology; John Wiley & Sons, Ltd: 2016; pp 487-504.

16. Dong, H.; Liffland, S.; Hillmyer, M. A.; Chang, M. C. Y. Engineering in vivo production of α-branched polyesters. *J. Am. Chem. Soc.* **2019***, 141*, 16877-16883.

17. Zhang, H.; Boghigian, B. A.; Pfeifer, B. A. Investigating the role of native propionyl-CoA and methylmalonyl-CoA metabolism on heterologous polyketide production in Escherichia coli. *Biotechnol. Bioeng.* **2010***, 105*, 567-573.

18. Fang, L.; Guell, M.; Church, G. M.; Pfeifer, B. A. Heterologous erythromycin production across strain and plasmid construction. *Biotechnol. Prog.* **2018***, 34*, 271-276.

19. Larionov, V.; Kouprina, N.; Graves, J.; Resnick, M. A. Highly selective isolation of human DNAs from rodent–human hybrid cells as circular yeast artificial chromosomes by transformation-associated recombination cloning. *PNAS* **1996***, 93*, 13925-13930.

20. Zaburannyi, N.; Rabyk, M.; Ostash, B.; Fedorenko, V.; Luzhetskyy, A. Insights into naturally minimised Streptomyces albus J1074 genome. *BMC Genomics* **2014***, 15*, 97.

21. Ziermann, R.; Betlach, M. C. Recombinant polyketide synthesis in Streptomyces: engineering of improved host strains. *BioTechniques* **1999***, 26*, 106-110.

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

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

24. 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*, 2506-2514.

25. Yamanaka, K.; Reynolds, K. A.; Kersten, R. D.; Ryan, K. S.; Gonzalez, D. J.; Nizet, V.; Dorrestein, P. C.; Moore, B. S. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *PNAS* **2014***, 111*, 1957-1962.

26. Ryan, O. W.; Skerker, J. M.; Maurer, M. J.; Li, X.; Tsai, J. C.; Poddar, S.; Lee, M. E.; DeLoache, W.; Dueber, J. E.; Arkin, A. P.; Cate, J. H. D. Selection of chromosomal DNA libraries using a multiplex CRISPR system. *Elife* **2014***, 3*.

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

28. Pistorino, M.; Pfeifer, B. A. Efficient experimental design and micro-scale medium enhancement of 6-deoxyerythronolide B production through Escherichia coli. *Biotechnology Progress* **2009***, 25*, 1364-1371.

29. Kawai, S.; Hashimoto, W.; Murata, K. Transformation of Saccharomyces cerevisiae and other fungi. *Bioeng Bugs* **2010***, 1*, 395-403.

30. Noskov, V. N.; Chuang, R.; Gibson, D. G.; Leem, S.; Larionov, V.; Kouprina, N. Isolation of circular yeast artificial chromosomes for synthetic biology and functional genomics studies. *Nat Protoc* **2011***, 6*, 89-96.

31. Zhang, H.; Wang, Y.; Wu, J.; Skalina, K.; Pfeifer, B. A. Complete biosynthesis of erythromycin A and designed analogs using E. coli as a heterologous host. *Chem. Biol.* **2010***, 17*, 1232-1240.

32. Thuronyi, B. W.; Privalsky, T. M.; Chang, M. C. Y. Engineered fluorine metabolism and fluoropolymer production in living cells. *Angewandte Chemie International Edition* **2017***, 56*, 13637-13640.

33. Fox, B. G.; Blommel, P. G. Autoinduction of protein expression. *Curr Protoc Protein Sci* **2009***, CHAPTER 5*, Unit-5.23.

34. Kotlyar, A. B.; Vinogradov, A. D. Interaction of the membrane-bound succinate dehydrogenase with substrate and competitive inhibitors. *Biochim. Biophys. Acta* **1984***, 784*, 24-34.

35. Ocampo, R.; Dolbier, W. R.; Abboud, K. A.; Zuluaga, F. Catalyzed Reformatsky reactions with ethyl bromofluoroacetate for the synthesis of α-fluoro-β-hydroxy acids. *J. Org. Chem.* **2002***, 67*, 72-78.

36. Zhang, Y.; Cronan, J. E. Transcriptional analysis of essential genes of the Escherichia coli fatty acid biosynthesis gene cluster by functional replacement with the analogous Salmonella typhimurium gene cluster. *J Bacteriol* **1998***, 180*, 3295-3303.

37. My, L.; Rekoske, B.; Lemke, J. J.; Viala, J. P.; Gourse, R. L.; Bouveret, E. Transcription of the Escherichia coli fatty acid synthesis operon fabHDG is directly activated by FadR and inhibited by ppGpp. *J. Bacteriol.* **2013***, 195*, 3784-3795.

38. Verwoert, I. I.; Verhagen, E. F.; van der Linden, K. H.; Verbree, E. C.; Nijkamp, H. J.; Stuitje, A. R. Molecular characterization of an Escherichia coli mutant with a temperature-sensitive malonyl coenzyme A-acyl carrier protein transacylase. *FEBS Lett.* **1994***, 348*, 311-316.

39. Starcevic, A.; Zucko, J.; Simunkovic, J.; Long, P. F.; Cullum, J.; Hranueli, D. ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res* **2008***, 36*, 6882-6892.

40. Yadav, G.; Gokhale, R. S.; Mohanty, D. Computational approach for prediction of domain organization and substrate specificity of modular polyketide synthases. *Journal of Molecular Biology* **2003***, 328*, 335-363.

41. Wang, Y.; Cobb, R. E.; Zhao, H. High-efficiency genome editing of streptomyces species by an engineered CRISPR/Cas system. *Meth. Enzymol.* **2016***, 575*, 271-284.

Chapter 5: *Biochemical elucidation of downstream processing of fluorine-containing polyketide intermediates*

5.1 Introduction

In the previous chapters, the main goal of study was to incorporate fluorine into modular polyketides via the use of fluoromalonyl-CoA precursor. Although, we successful showed incorporation in a few examples, we noticed that products derived from fluorine incorporation at an upstream module is less abundant than those derived from fluorine incorporation at a downstream module in comparable systems. This was observed with the incorporation of fluoromalonyl-CoA at DEBS Module 5 compared to Module 6 in 6dEB production system (*section 4.3*), as well as in Module 2 compared to Module 3 of the mini-PKS system consisting of Mod 2_{DEBS} and $Mod3_{DEBS}$ +TE with active mutation in the AT domain of either module [1]. These results are consistent with findings that engineered production of polyketide analogs generally yields lower titers than that of parental molecules [2-4]. These observations suggest that non-native substitutions on polyketide intermediates have significant effects on system turnover. Here, we aim to start looking into the effects of fluorine substitution on downstream processing and elongation from fluorine-containing polyketide intermediates.

Once selected by the AT domain, a carboxyacyl extender unit undergoes decarboxylation to form an enolate anion which then participates in Claisen condensation with the growing intermediate and elongates it by two carbons. After condensation, the elongated intermediate may be subjected to further processing, which includes reduction of the β-keto group by KR, epimerization of the α-substitution by KR [5], elimination of the β-hydroxy group by DH [6, 7], and reduction of alkene by ER [8, 9]. Furthermore, the intermediate is then transacylated to the active site of the KS domain of the following module. The efficiency of these enzymatic reactions could be affected by the reactivity of the substrate. Although many alkyl substitutions may be relatively inert, fluorine substitution at the α-position of a growing polyketide intermediate could enhance or hinder these reactions due to its high electronegativity. Furthermore, structure evidence showed that a polyketide mega-synthase adopts different conformations based on the identity of the intermediate tethered to its ACP domain [10]. Thus, the presence of a non-canonical substitution on the intermediate might have implications on the overall function of a PKS module.

In the series of enzymatic reactions that may happen after extender unit incorporation, ketoreduction and epimerization by KR domains are the first reactions to occur and thus the first focus of our investigation. The activities of KR domains determine the orientation of most saturated stereocenters on polyketide products [5]. In engineering fluorine-containing natural product analogs, the orientation of fluorine substitution might be of significant considerations given fluorine atom's effects on electronic distribution and stability of molecules. Thus, we seek to investigate the effects of fluorine substitution on the activities of PKS KR domains and to develop methods to control the orientation of fluorine substitution on polyketide products.

5.2 Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), glucose, and glycerol were purchased from EMD Biosciences (Damstadt, Germany). Carbenicillin (Cb), isopropyl-β-D-thiogalactopyranoside (IPTG), sodium chloride, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, dichloromethane, ethyl acetate, sodium phosphate monobasic, potassium phosphate monobasic, and ammonium chloride were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A trilithium sodium salt (CoA), tris(2-carboxyethyl)phosphine

(TCEP) hydrochloride, poly(ethyleneimine) (PEI), β-mercaptoethanol (BME), sodium phosphate dibasic heptahydrate, cysteamine, 1-ethyl-3-(-dimethylaminopropyl)carbodiimide hydrochloride (EDC), 4-dimethylaminopyridine (DMAP), reduced β-nicotinamide adenine dinucleotide 2′-phosphate (NADPH), acetonitrile, dimethyl sulfoxide (DMSO), and ethyl propionylacetate were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid, dithiothreitol (DTT), and perchloric acid were purchased from Fluka (Charlotte, NC). 2-methyl-3-oxopentanoic acid ethyl ester was purchased from Fragmenta (Monmouth Junction, NJ). Ethyl-2-fluoro-3-oxopentanoate was purchased from Santa Cruz Biotechnology (Dallas, TX). Mini-PROTEAN TGX 8-16%, and Bio-Rad Protein Assay Dye Reagent concentrate was purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, and Phusion DNA polymerase were purchased from New England Biolabs (Ipswich, MA). Glu-C was purchased from Promega (Madison, WI). Trypsin was purchased from QB3 MacroLab (Berkeley, CA). 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 water. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). PD-10 columns Sephadex G-25M was purchased from GE (Chicago, IL). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Amico Ultra 3,000 MWCO, and 10,000 MWCO centrifugal concentrators were purchased from EMD Millipore (Billerica, MA). Chloroform-d was purchased from Cambridge Isotope Laboratories (Andover, MA). 1H -NMR spectra were collected at 25ºC on Bruker AVB-400 or AVQ-400 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 Quadrupole-Time-of-Flight (QTOF) Accurate Mass spectrometer from Agilent Technologies.

Bacterial strains. *E. coli* DH10B-T1^R was used for DNA construction. *E. coli* BL21(DE3)-T1^R and BAP1 [11] were used for heterologous protein expression. BAP1 was used for expression of DEBS ACP domains when modification with phosphopantethiene is desired. BL21(DE3)-T1^R was used for when modification with phosphopantethiene is not desired and all other proteins.

Gene and plasmid construction. Standard molecular biology techniques were used to carry out plasmid construction. All PCR amplifications were carried out with Phusion High Fidelity DNA polymerase. All constructs were verified by sequencing (UC Berkeley DNA Sequencing Facility, Berkeley, CA). Sfp plasmid was gifted by the laboratory of Professor Chaitan Khosla (Stanford University).

pET28a-His$_6$TEV-KR$_{TYLSMod1.}$ pET28a-His$_6$TEV-KR$_{TYLSMod1}$ encodes for the KR domain from the first module of tylosin synthase. The plasmid was constructed by amplification from gblock Tylosin KR6698-8168, which encodes for the KR domain from module 1 of tylosin synthase codon optimized for *E. coli* with primer pairs TKR1Fwd/TKR1Rev. The PCR product was inserted into pET28a(gg)HisN-TEV-RFP via Golden Gate assembly using BsaI enzyme.

pET28a-His$_6$TEV-KR$_{DEBSMod1.}$ pET28a-His$_6$TEV-KR$_{DEBSMod1}$ encodes for the KR domain from the first module of DEBS according to previously described construct [12]. The plasmid was constructed by amplification from pCWori-DEBS1 (strain#1327) with primer pairs EryKR1Fwd/EryKR1Rev2. The PCR product was inserted into pET28a(gg)HisN-TEV-RFP via Golden Gate assembly using BsaI enzyme.

pET16b-His$_{10}$Pres-ACP$_{DEBSMod6}$. The ACP domain of Mod6$_{DEBS}$ was amplified from pAYC-138 [13] using primers PresACP6_Fwd3 and PresACP6_Rev. The PCR product was assembled into the NdeI-BamHI site of a modified pET16b with the Factor Xa cleavage site replaced by the PreScission cleavage site via Gibson assembly.

Expression and purification of His-tagged proteins. The expression and purification of His$_{10}$-ACP$_{Mod6}$ was carried out as previously described in Chapter 2. For Sfp-His$_6$, His$_6$-KR$_{PIKSMod1}$, His$_6$-KR$_{TYLSMod1}$, and His$_6$-KR$_{DEBSSMod1}$, the proteins were purified from *E. coli* BL21(DE3)-T1^R as previously described in Chapter 2 for His$_6$-MatB.

Synthesis of SNAC substrates. 2-methyl-3-oxopentanoyl-SNAC (oxo-NDK-SNAC) was synthesized as previously described in Chapter 2. 3-oxopentanoyl-SNAC (oxo-HDK-SNAC) and 2-fluoro-3-oxopentanoyl-SNAC (oxo-FDK-SNAC) were synthesized as previously described in Chapter 2 for oxo-NDK-SNAC from ethyl-3-oxopentanoate and ethyl-2-fluoro-3-oxopentanoate in place of ethyl-2-methyl-3-oxopentanoate respectively. Both aqueous saponification and N-acetylcysteamine (SNAC) coupling were shortened to 3 h at room temperature. Products were confirmed by ^{1}H-NMR and stored in solid form at -80 ºC. oxo-HDK-SNAC: ^{1}H-NMR (400 MHz, $CDCl_3$=7.26 ppm) δ 3.70 (s, 1H), 3.44 (q, J = 8 Hz, 2H), 3.9 (t, J = 8 Hz, 2H), 2.56 (q, J = 8 Hz, 2H), 1.98 (s, 3H), 1.08 (t, J =, 8 Hz, 3H). oxo-FDK-SNAC ^{1}H-NMR (400 MHz, $CDCl_3$=7.26 ppm) δ 5.28 (d, J = 52, 1H), 3.46 (q, J = 8, 2H), 3.13 (t, J=4, 2H), 2.57-2.83 (m, 2H) 1.97 (s, 3H), 1.11 (t, J=8, 3H); ^{19}F-NMR(400 MHz) δ 191.4 (d, J=52 Hz).

Synthesis of CoA substrates. The synthesis of 2-methyl-3-oxopentanoyl-CoA (oxo-NDK-CoA) and 3-oxopentanoyl-CoA (oxo-HDK-CoA) were developed from a previously described method [14]. oxo-NDK-CoA and oxo-HDK-CoA were prepared by coupling free acid to thiophenol, followed by transthioesterification with CoA and HPLC purification. In short, ethyl-2-methyl-3-oxopentanoate and ethyl-3-oxopentanoate were saponified as described in Chapter 2 for oxo-NDK-SNAC for synthesis of oxo-NDK-CoA and oxo-HDK-CoA respectively. The resulting 2-methyl-3-oxopentanoic acid or 3-oxopentanoic acid (1 mmol) was dissolved in 10 mL of ethylacetate on ice with stirring. After 3 min, thiophenol (82 µL, 0.8 mmol, 0.8 eq) and DMAP (catalytic) were added to the reaction. After 5 min, N,N'-diisopropyl carbodiimide (235 µL, 1.5 mmol; 1.5 eq) was added drop-wise. The reaction was stirred on ice for 2 h followed by 2 h at room temperature. The reaction was filtered, extracted with 3 ×10 mL of 1 M sodium bicarbonate. The organic layer was dried over anhydrous magnesium sulfate, and evaporated to dryness by rotary evaporation.

The resulting phenyl-NDKthioester or phenyl-HDKthioester (200 µmol) was dissolved in 235 µL of acetonitrile and added to 4.7 mL of 0.5 M sodium bicarbonate solution containing 100 µmol of CoA on ice. The reaction was stirred on ice for 2 h followed by 2 h at room temperature. The reaction was quenched by adding 50 % (*v/v*) formic acid (600 µL) to pH 3 and extracted twice with diethyl ether. Nitrogen was passed over the aqueous layer to remove residual ether, and the solution was stored at -80 ºC until purification.

The crude reaction was purified on an Eclipse XDB-C18 column (5 µm, 9.4 × 250 mm, room temperature; Agilent) using a linear gradient from 0 to 50 % acetonitrile over 24 min with 10 mM ammonium acetate as the aqueous mobile phase at flow rate of 4 mL/min after an initial hold at 0 % acetonitrile for 2 min. Fractions (2 mL) were screened using Agilent 1290 UPLC connected to QTOF using positive ionization mode, and those containing pure oxo-NDK-CoA or oxo-HDK-CoA were combined and lyophilized overnight and resuspended in water. Resuspended residue

was desalted using the same HPLC column with a linear gradient from 0 to 20 % acetonitrile in water as the aqueous mobile phase at flow rate of 4 mL/min after an initial hold at 0 % acetonitrile for 2 min. Desalted products were lyophilized and resuspended in water for biochemical assays.

2-Fluoro-3-oxopentanoyl-CoA (oxo-FDK-CoA) was prepared by coupling free acid to N-hydroxysuccinimide, following by thioesterification with CoA and HPLC purification. Ethyl-2-fluoro-3-oxopentanoate was saponified as described in Chapter 2 for oxo-NDK-SNAC. The resulting 2-fluoro-3-oxopentanoic acid (1 mmol) and N-hydroxysuccinimide (115 mg, 1 mmol, 1 eq) were dissolved in 5 mL of tetrahydrofuran on ice. Then, N,N'-diisopropyl carbodiimide (157 µL, 1 mmol; 1 eq) was added drop-wise. The reaction was stirred on ice for 30 min before it was dried under compressed air. Crude reaction residue was dissolved in cold 10 % acetonitrile in 0.5 M sodium phosphate pH 9 containing 100 µmol CoA (10 mL). The reaction was stirred on ice for 20 min before flash frozen and stored at -80 ºC until purification. oxo-FDK-CoA was purified following the same HPLC method as described above for oxo-NDK-CoA or oxo-HDK-CoA.

Assay for *in vitro* phosphopantethienylation. All assays (50 µL) 20 µM CoA substrate, 0.2 µM Sfp, and 20 µM ACP in contained 50 mM HEPES, 5 mM TCEP, 10 mM magnesium chloride pH 7.5. Reactions were incubated at room temperature for 30 min. Reactions were stored at -80 ºC until analysis.

Determination of acyl group of ACP by peptide mass spectometry. Before analysis, 700 ng of Glu-C protease (7 µL, 0.1 µg/µL) was added to 25 µL of reaction and incubated at 37 ºC for 2 h. Following Glu-C digestion, 700 ng of trypsin (1.4 µL, 0.5 µg/µL) was added and the reaction was incubated at 37 ºC for 2 h. To quench the reaction, reaction mixture was mixed with formic acid to 5 % (*v/v*) final concentration and centrifuged for 10 min at 4 ºC at 20,817 × *g*.

Resulting peptide mixture was analyzed by high resolution mass spectrometry in positive ionization mode on Agilent 6530B QTOF. Samples were analyzed on AdvanceBio Peptide Mapping C18 HPLC column (2.7 µm, 2.1 × 100 mm, Agilent) on an Agilent 1290 UPLC.

For $ACP_{DEBSMod6,}$ masses observed corresponded to $[M+1H]^{+,}$ $[M+2H]^{2+}$, and $[M+3H]^{3+}$ of peptide LGFDSLTAVGLR. $[M+2H]^{2+}$ is the most dominant ion and is counted as representative of the protein population. Masses observed of $ACP_{DEBSMod6}$ in various forms are as follows (expected mass): apo-ACP (624.8510), holo-ACP (794.8939), oxo-NDK-ACP (850.9201), oxo-HDK-ACP (843.9123), oxo-FDK-ACP (852.9076).

Steady-state kinetic characterization of ketoreduction activity. Assays were performed in 96-well microplate (black, polystyrene, µ-clear, f-bottom, chimney well, med binding; Greiner Bio-One) by monitoring NADPH fluorescence using a Synergy Mx microplate reader (Biotek; top optical position at 8 mm read height, gain = 50). Reactions were monitored for 6 mins at 30 ºC using minimum interval setting between measurements.

Assay components were prepared in two solutions. Solution 1, 300 mM HEPES pH 7.5, 10% (*v/v)* glycerol, 100 mM sodium chloride; solution 2 oxo-NDK-SNAC, oxo-FDK-SNAC, or oxo-HDK-SNAC substrate at 5× final concentration on 25% (*v/v)* DMSO. Immediately before analysis of each set of reactions, 9.9 µL of 20 mM NADPH and 19.8 µL of ketoreductase enzyme at 100× final concentration were mixed with 1554.3 µL of Solution 1, and 160 µL of Solution 1 containing NADPH and enzyme were transferred to each well. The reactions were then initiated with 40 µL of Solution 2.

Final assay mixtures (200 µL) contained 235.5 mM HEPES, 7.85% *v/v* glycerol, 78.5 mM sodium chloride, 100 µM NADPH, 5% *v/v* DMSO pH 7.5. The final concentrations of oxo-NDK-SNAC, oxo-FDK-SNAC, and oxo-HDK-SNAC were 100, 50, 25, 12.5, 6.25, 3.13, 156, 0.78 mM. The final concentrations of enzymes were as follows: 100 nM $KR_{PIKSMod1}$ for oxo-NDK-Snac, 500 nM $KR_{PIKSMod1}$ for oxo-HDK-Snac, 50 nM $KR_{PIKSMod1}$ for oxo-FDK-Snac, 2 µM $KR_{DEBSMod1}$ for oxo-NDK-Snac, 6.7 µM $KR_{DEBSMod1}$ for oxo-HDK-Snac, 200 nM $KR_{DEBSMod1}$ for oxo-FDK-Snac, 1 µM $KR_{TYLSMod1}$ for oxo-NDK-Snac, 200 nM $KR_{TYLSMod1}$ for oxo-HDK-Snac, 50 nM $KR_{TYLSMod1}$ for oxo-FDK-Snac.

Ketoreduction of oxo-NDK-SNAC and oxo-HDK-SNAC was monitored at λ_{Ex} = 340 nm and λ_{Em} = 450 nm, whereas ketoreduction of oxo-FDK-SNAC was monitored at λ_{Ex} = 360 nm and λ_{Em} = 450 nm to avoid absorbance at 340 nm by substrate or substrate contaminant. Fluorescence was converted to concentration using external NADPH standard curves collected in the same assay matrix at the same λ_{Ex} and λ_{Em}. Initial velocities (v_0) were defined as the linear portion of the change of product concentration over time curve. Ketoreduction data were fit by non-linear curve fitting to Michaelis-Menten equation modified with substrate inhibition:

$$v_0 = \frac{v_{max}[S]}{K_M + [S](1 + \frac{[S]}{K_i})}$$

where v_0 = initial velocity; v_{max} = maximum velocity, [S] = substrate concentration, K_M = Michaelis-Menten constant, and K_i = inhibition constant. Fitting were performed using Origin 6.0 with initial values estimated by the Solver tool in Microsoft Excel.

5.3 Results and discussion

Synthesis of fluorine-containing polyketide intermediate analogs. To be able to characterize the effects of fluorine substitution on downstream enzymatic activities, fluorine-containing polyketide intermediate analog that mimics the result of fluoromalonyl-CoA incorporation is necessary. Thus, we set out to synthesize 2-fluoro-3-oxopentanoyl-SNAC (oxo-FDK-SNAC) and 2-fluoro-3-oxopentanoyl-S-ACP substrates. Oxo-FDK-SNAC was successfully synthesized following the same protocol previously described for synthesizing oxo-NDK-SNAC (*section 2.2*).

To obtain 2-fluoro-3-oxopentanoyl-ACP (oxo-FDK-ACP), we planned to synthesize 2-fluoro-3-oxopentanoyl-CoA (oxo-FDK-CoA) and conjugate it to the active site serine of apo-ACP using Sfp enzyme (*Figure 5.1*) [15]. We attempted to synthesize oxo-FDK-CoA with the same procedure as previously described for oxo-NDK-CoA and oxo-HDK -CoA [14]. However, oxo-FDK-CoA could not be obtained to satisfactory amount. We found that thiophenol, which is the intermediate nucleophile used in the synthesis of oxo-NDK-CoA and oxo-HDK-CoA causes thialysis of oxo-FDK-CoA, leading to the formation of propionyl-CoA and fluoroacetyl-CoA by product. This observation is consistent with expected increased nucleophilicity of the β-keto group of oxo-FDK-CoA due to fluorine *substitution* at the α-position.

Thus, we set out to determine alternative synthetic route for the fluorinated diketide analog. We experimented with other thiol nucleophiles including SNAC, β-mercaptoethanol, dithiothreitol, and mercaptopropionic acid, but none of the reagents yielded the desired oxo-FDK-CoA product. We also attempted direct CoA coupling using HATU as activating agent, but no

product was obtained. We found that, however, shortening the reaction time and lowering the reaction temperature from 4 h total of thiophenol coupling and 4 h total of coenzyme A exchange to 20 min for each step at 0 ºC yielded the desired oxo-FDK-CoA product. In addition, using N-hydroxy succinimide as nucleophile instead of thiophenol yielded similar result through simpler reaction set up. We found the short reaction time at 0 ºC to minimize the degradation of oxo-FDK-CoA to propionyl-CoA and fluoroacetyl-CoA.

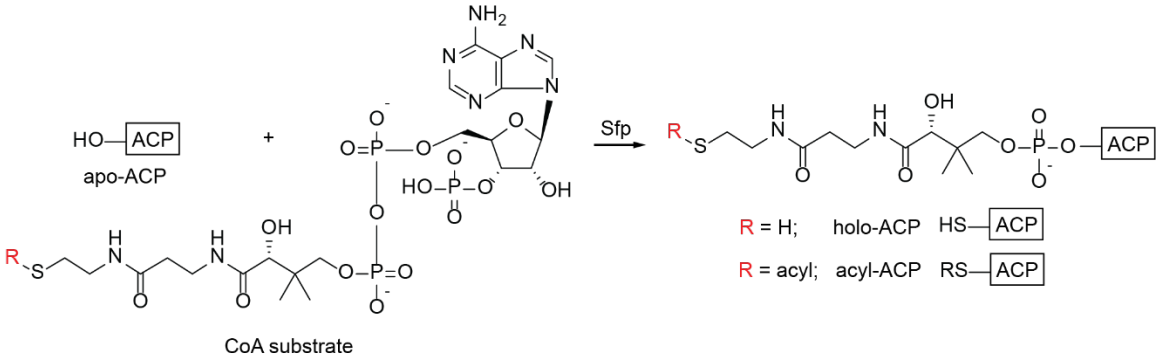


Figure 5.1. *Phosphopantetheinylation reaction. Sfp transfers CoA (R=H) or acyl-CoA (R=acyl) substrate to the active site Ser of apo-ACP protein to produce holo-ACP (R=H) or acyl-ACP (R=acyl).*

In order to generate ACP-linked substrates, acyl-CoA substrate would be added to ligated to apo-ACP by Sfp-mediated phosphopantetheination. Sfp protein was overexpressed and purified to near homogeneity (*Figure 5.2*). To determine reaction condition, we initially tested phosphopantetheination with apo-ACP and free coenzyme A. Sfp was able to transform apo-ACP almost entirely to holo-ACP within 30min (*Figure 5.3*). The appearance of holo-ACP was dependent on the presence of Sfp enzyme.

We then tested the activity of Sfp enzyme with diketide CoA substrates, namely oxo-NDK-CoA, oxo-HDK-CoA, and oxo-FDK-CoA. Note that in this case, the starting ACP was a mixture of apo-ACP and holo-ACP. Of three substrates tested, mass signals corresponding to acyl-ACP loaded with oxo-NDK and oxo-HDM moiety were detected, although signals corresponding to acyl-ACP loaded with oxo-FDK moiety was not observed (*Figure 5.4*). However, in all three cases, signals corresponding to apo-ACP decreased, suggesting that transformation of apo-ACP to other forms of the protein occurred. Thus, the lack of observable acyl-ACP signals in reaction including oxo-FDK-CoA was attributed to the degradation of oxo-FDK-ACP, rather than a lack of conversation from apo-ACP to acyl-ACP. As oxo-FDK-SNAC and oxo-FDK-CoA are sufficiently stable to be isolated, we hypothesize that the observed hydrolysis of the fluorinated diketide from the ACP carrier may be facilitated by Sfp, ACP, or the reaction condition. Further optimization is required to determine the appropriate condition for loading oxo-FDK substrate onto ACP and preventing degradation.

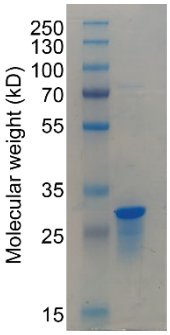


Figure 5.2 *SDS-PAGE of purified Sfp.*

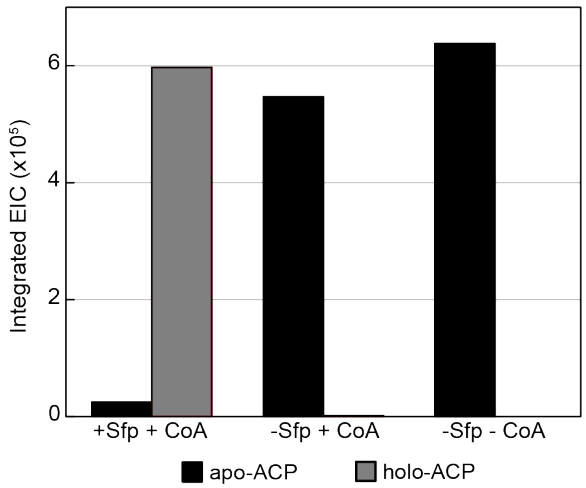


Figure 5.3. *In vitro phosphopantetheinylation of apo-ACP by Sfp. Integrated ion counts are shown for reactions containing 20 μM ACP substrate and when used 20 μM coenzyme A and 0.2 μM Sfp. After 30 min incubation, reactions were subjected to protease digestion by Glu-C and trypsin. Digestion reactions were quenched, and peptides were analyzed by LC-QTOF.*

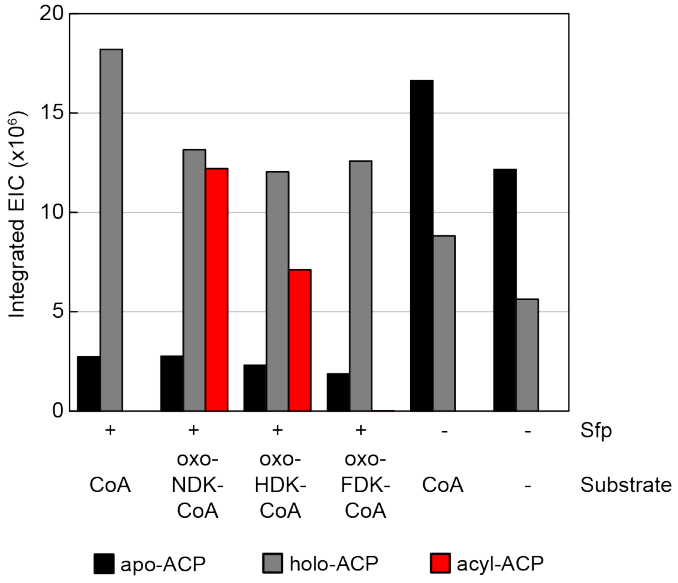


Figure 5.4. *In vitro substrate loading of apo-ACP by Sfp. Integrated ion counts are shown for reactions containing 20 μM ACP substrate and when used 0.2 μM Sfp, and 20 μM coenzyme A, oxo-NDK-CoA, oxo-HDK-CoA, or oxo-FDK-CoA. After 30 min incubation, reactions were subjected to protease digestion by Glu-C and trypsin. Digestion reactions were quenched, and peptides were analyzed by LC-QTOF.*

Structure and functions of ketoredutase domains. A KR domain functions to reduce β-keto group on polyketide intermediates to β-hydroxy group and to epimerize α-methyl substitution, determining two stereocenters of the polyketide molecule. KR domains are categorized based on the stereochemical outcome of the product (*Figure 5.5A*) [16]. A-type KR domains produce L-β-hydroxy whereas B-type KR domains produce D- β-hydroxy. C-type KR domains are defective in ketoreduction and produce β-keto product. KR type 1 maintains D-orientation of α-methyl substitution as produced by the KS domain in Claisen condensation whereas KR type 2 performs epimerization to create L-orientation of α-methyl substitution. Thus, each KR domain is designated with a letter A, B, or C, and a number 1 or 2.

Structural evidence shows the stereochemical outcome at the β-position is determined by the binding orientation of phosphopantetheine arm of the ACP domain [17]. A-type KRs possess a tryptophan residue near the active site that interacts with the phosphopantetheine arm while B-type KRs has the LDD motif that lets the phosphopantetheine arm to bind from the opposite site of the lid helix [18]. These structural features allow KR domains to catalyze reduction reaction to opposite faces of the substrate using one set of catalytic residues and NADPH binding motif. Epimerase activity has been shown to utilize the same active side residues as reductase activity [19] but is independent of NADPH co-factor [20].

Most studies on KR domains were carried out on α-methyl substrates which in most cases resemble native substrates with respect to the substituting group. Some KRs have been shown to maintain their stereoselectivity over substrates with non-native substitutions such as ethyl group and proton at the α-position [21]. However, no investigation had been done for substrates with altered electronic distribution such as those with fluorine substitution. Here, we aim to investigate the activities of PKS KR domains on α-fluoro substrates in order to characterize interactions between the incorporated fluorinated extender unit with downstream processing enzymes.

For this, we chose three representative KR domains that produce opposite stereochemistry of α-substitution: type B1 KR domain from module 1 of tylosin synthase ($KR_{TYLSMod1)}$, type B2 KR domain from module 1 of 6-deoxyerythronolide B synthase ($KR_{DEBSMod1),}$ and type B2 KR domain from module 1 of pikromycin synthase ($KR_{PIKSMod1)}$. All three enzymes have been characterized *in vitro* and shown to maintain their stereochemical outcome on diketide-SNAC thioesters, allowing for uncoupling the role of the carrier protein from the identity of substrate (*Figure 5.5B*) [21-23]. We aim to assess kinetic parameters and stereochemical outcome of the three enzymes' activities on native substrate as well as fluorinated acyl substrate analog carried by SNAC and carrier protein.

Ketoreduction activity on fluorine-containing polyketide intermediate. In order to determine effects of fluorine substitution on ketoreduction of polyketide intermediates, we set out to carry out the steady-state kinetic characterization of three representative ketoreductase domains. $KR_{PIKSMod1}$ and $KR_{DEBSMod1}$ of type B2 and $KR_{TYLSMod1}$ of type B1 were overexpressed and purified to near homogeneity (*Figure 5.6*). We examined the activity of all three enzymes on oxo-FDK-SNAC, a fluorine-containing racemic diketide analog, in comparison to their functions on methyl-containing diketide mimic oxo-NDK-SNAC and unsubstituted diketide mimic oxo-HDK-SNAC. Reduction was monitored by disappearance of NADPH as it is stoichiometrically consumed in the reaction.

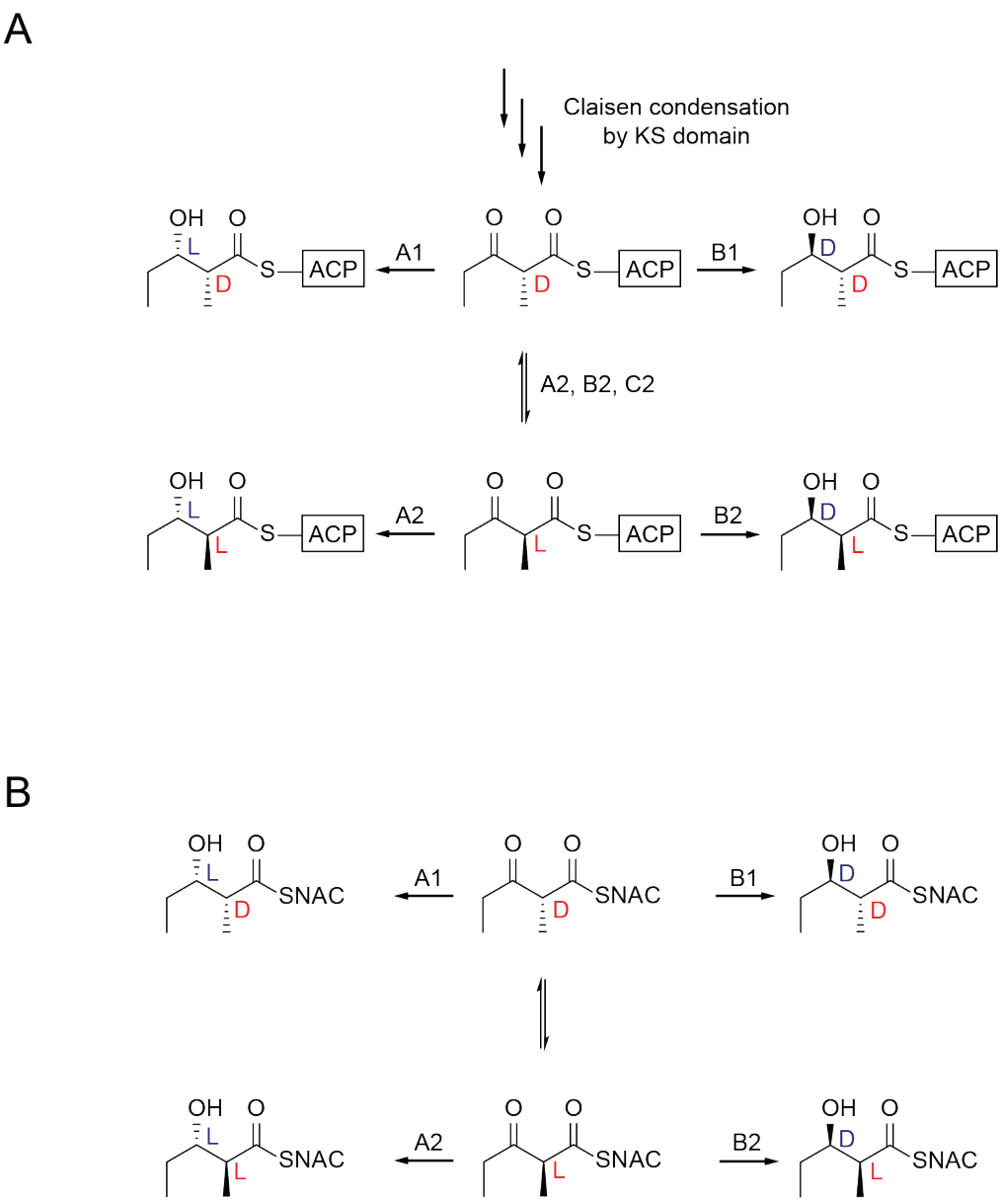


Figure 5.5. *Ketoreduction and epimerization activities of ketoreductase domains. (A) Activities of each type of PKS ketoreductase domains on their native α-methyl-β-ketoacyl-ACP substrates. Claisen condensation reaction catalyzed by KS domain produces α-D-substituted intermediate. This product may directly undergo reduction by type 1 KR (A1 and B1) domain. It may also undergo epimerization of α-position before reduction by type 2 KR (A2, B2, and C2). A-type and B-type reductases produce products of L and D stereochemistry at β-position respectively. (B) Reductases maintain their intrinsic stereospecificity and stereoselectivity on racemic SNAC thioester surrogate of diketide substrates.*

The dose-response curves for reduction reaction show evidence of inhibition at high substrate concentration and were therefore fit with Michaelis-Menten equation modified for substrate inhibition. We found that all three enzymes were active on the native substrate mimic oxo-NDK-SNAC. Determined kinetic parameters for $KR_{DEBSMod1}$ and $KR_{TYLSMod1}$ are comparable to those previously determined for different constructs of the same enzyme domains [23]. We similarly found that $KR_{TYLSMod1}$ displays higher k_{cat} and k_{cat}/K_M than $KR_{DEBSMod1}$. All three enzymes also exhibit activity on oxo-HDK-SNAC. This result is consistent with previous report that showed evidence of corresponding reactions by chiral chromatography [21]. As was previously shown, we observed $KR_{DEBSMod1}$ to be significantly slower than both $KR_{PIKSMod1}$ and $KR_{TYLSMod1}$ [21, 23].

In addition, we found that all three enzymes were able to turn over oxo-FDK-SNAC. In fact, they display more efficient behavior (10-400 higher k_{cat}/K_M) towards the fluorine-containing substrate compared to methyl-substituted and unsubstituted substrates. The difference appears to arise from both higher k_{cat} and lower K_M values. This result is consistent with the expected chemical behavior of fluorine substitution as the fluorine atom is expected to withdraw electron from the keto carbon, making it more susceptible for hydride attack. The lower K_M values further suggest that fluorine is accommodated in the substrate binding pocket. This may also reflect the altered electronics of the keto group as the keto oxygen is bound by hydrogen bonding in the substrate-binding pocket [24]. However, note that native substrates of KR domains are ACP-tethered polyketide chain and binding interactions involve protein-protein interface and interactions between KR and the phosphopantetheine prosthetic group of ACP.

Overall, the KR domains appear to display little discrimination against different substrate analogs. This observation may reflect the role of the domain in its native context as part of a mega-synthase. KR domain is typically presented with the product of the Claisen condensation reaction catalyzed by the KS domain of the same module between the growing chain and an extender unit selected by the AT domain, which usually displays strong discrimination against non-native extender units [25].

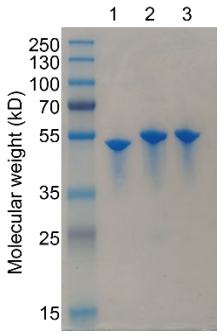
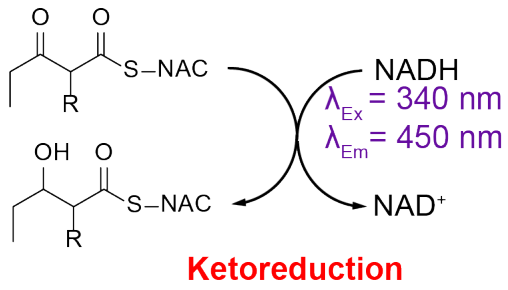
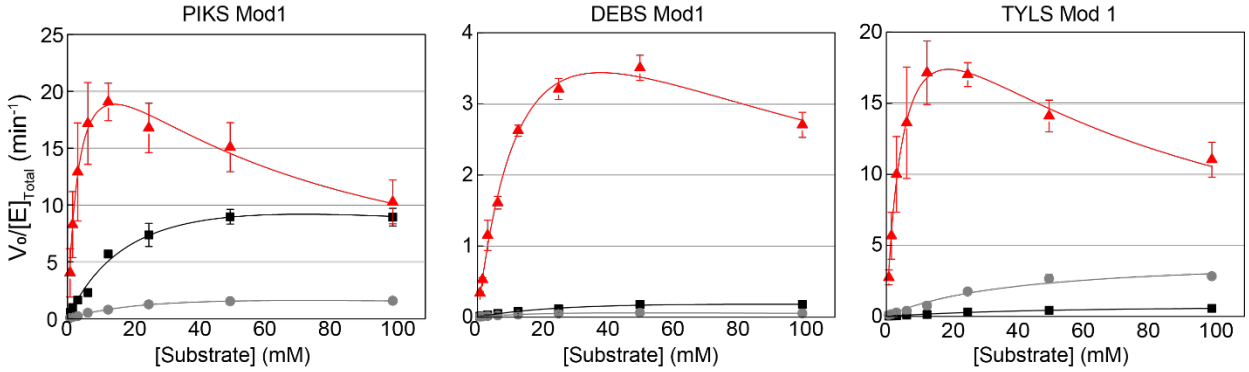


Figure 5.6. *SDS-PAGE of purified proteins used in steady-state kinetic characterization. (lane 1) $KR_{PIKSMod1}$, (lane 2) $KR_{DEBSMod1}$, (lane 3) $KR_{TYLSMod1}$*

A



B



C

Ketoreductase	Substrate	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})	K_i (mM)
PIKS Mod1	oxo-NDK-SNAC	15 ± 2.9	25 ± 7.3	0.62 ± 0.29	207 ± 120
	oxo-HDK-SNAC	3.1 ± 0.31	33 ± 4.8	0.093 ± 0.023	161 ± 42
	oxo-FDK-SNAC	29 ± 4.3	3.8 ± 1.2	7.6 ± 3.6	56 ± 20
DEBS Mod1	oxo-NDK-SNAC	0.34 ± 0.080	39 ± 13	0.0087 ± 0.0050	203 ± 139
	oxo-HDK-SNAC	0.12 ± 0.10	28 ± 33	0.0044 ± 0.0090	95 ± 158
	oxo-FDK-SNAC	6.6 ± 0.66	17 ± 2.8	0.38 ± 0.10	83 ± 18
TYLS Mod1	oxo-NDK-SNAC	1.0 ± 0.39	66 ± 32	0.016 ± 0.014	566 ± 1209
	oxo-HDK-SNAC	4.2 ± 0.32	40 ± 6.8	0.11 ± 0.027	fixed to 10^6
	oxo-FDK-SNAC	28 ± 4.0	5.8 ± 1.6	4.8 ± 2.0	63 ± 20

Figure 5.7. *Steady-state kinetic characterization of ketoreduction of oxo-diketide-SNAC by ketoreductases. (A) Assay for measuring ketoreductase activity by NADPH fluorescence. (B) Michaelis-Menten curves for ketoreduction of oxo-NDK-SNAC (black), oxo-HDK-SNAC (grey), and oxo-FDK-SNAC (red) by $KR_{PIKSMod1}$ (left), $KR_{DEBSMod1}$ (middle), $KR_{TYLSMod1}$ (right). Data points are average ± s.d. of technical replicates (n=3). (C) Table contains k_{cat}, K_M, k_{cat}/K_M, and K_i calculated by non-linear curve fitting to the Michaelis-Menten equation with substrate inhibition. Error in k_{cat}/K_M is obtained by propagation from the individual kinetics terms. Data are mean ± s.e. of technical replicates (n=3).*

5.4 Conclusions

Methods for the generation of fluorine-containing polyketide intermediate analog, which mimics the growing polyketide chain immediately after fluoromalonyl-CoA extender unit has been incorporated have been optimized, leading to successful synthesis of oxo-FDK-SNAC and oxo-FDK-CoA. Efforts to generate oxo-FDK-ACP from oxo-FDK-CoA and apo-ACP using Sfp enzyme suggest that the fluorine-containing acyl-ACP molecule is subjected to significant hydrolysis. As oxo-FDK-SNAC and oxo-FDK-CoA are sufficiently stable to be isolated, we hypothesize that hydrolysis of oxo-FDK-ACP may be facilitated by Sfp, ACP, or the reaction condition. Further optimization of the reaction is required to obtain ACP-tethered fluorine containing polyketide intermediate.

Steady-state kinetic characterization of ketoreduction of oxo-FDK-SNAC showed that the fluorine-substituted substrate analog could be turned over by the KR domain at rates faster than those observed for methyl-substituted and unsubstituted substrates. The increased rate is attributed to the increased reactivity of keto group adjacent to fluorine substitution. Similar analysis with oxo-FDK-ACP could reveal the role of protein-protein interaction in ketoreduction reaction.

The next immediate goals for our investigation would be to determine the stereochemical outcomes of the reduction of oxo-FDK-SNAC by different types of KR domains. Likely, the combination of chiral chromatography and NMR analysis of large-scale ketoreduction reaction would be useful for assigning the stereochemistry of the α-fluoro substitution as well as the β-hydroxy group.

5.5 References

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

2. Koch, A. A.; Schmidt, J. J.; Lowell, A. N.; Hansen, D. A.; Coburn, K. M.; Chemler, J. A.; Sherman, D. H. Probing selectivity and creating structural diversity through hybrid polyketide synthases. *Angewandte Chemie* **2020***, 132*, 13677-13682.

3. Xu, Y.; Zhou, T.; Zhang, S.; Espinosa-Artiles, P.; Wang, L.; Zhang, W.; Lin, M.; Gunatilaka, A. A. L.; Zhan, J.; Molnár, I. Diversity-oriented combinatorial biosynthesis of benzenediol lactone scaffolds by subunit shuffling of fungal polyketide synthases. *PNAS* **2014***, 111*, 12354-12359.

4. Weissman, K. J.; Leadlay, P. F. Combinatorial biosynthesis of reduced polyketides. *Nature reviews. Microbiology* **2005***, 3*, 925-936.

5. Kwan, D. H.; Tosin, M.; Schläger, N.; Schulz, F.; Leadlay, P. F. Insights into the stereospecificity of ketoreduction in a modular polyketide synthase. *Org. Biomol. Chem.* **2011***, 9*, 2053-2056.

6. Li, Y.; Dodge, G. J.; Fiers, W. D.; Fecik, R. A.; Smith, J. L.; Aldrich, C. C. Functional characterization of a dehydratase domain from the pikromycin polyketide synthase. *J Am Chem Soc* **2015***, 137*, 7003-7006.

7. Barajas, J. F.; McAndrew, R. P.; Thompson, M. G.; Backman, T. W. H.; Pang, B.; de Rond, T.; Pereira, J. H.; Benites, V. T.; Martín, H. G.; Baidoo, E. E. K.; Hillson, N. J.; Adams, P. D.; Keasling, J. D. Structural insights into dehydratase substrate selection for the borrelidin and fluvirucin polyketide synthases. *J. Ind. Microbiol. Biotechnol.* **2019***, 46*, 1225-1235.

8. Ames, B. D.; Nguyen, C.; Bruegger, J.; Smith, P.; Xu, W.; Ma, S.; Wong, E.; Wong, S.; Xie, X.; Li, J. W. -.; Vederas, J. C.; Tang, Y.; Tsai, S. Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *PNAS* **2012***, 109*, 11144-11149.

9. Roberts, D. M.; Bartel, C.; Scott, A.; Ivison, D.; Simpson, T. J.; Cox, R. J. Substrate selectivity of an isolated enoyl reductase catalytic domain from an iterative highly reducing fungal polyketide synthase reveals key components of programming. *Chem Sci* **2017***, 8*, 1116-1126.

10. Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R. H.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structural rearrangements of a polyketide synthase module during its catalytic cycle. *Nature* **2014***, 510*, 560-564.

11. 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*, 1790-1792.

12. Keatinge-Clay, A. T.; Stroud, R. M. The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. *Structure* **2006***, 14*, 737-748.

13. 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*, 95-102.

14. Blaisse, M. R.; Dong, H.; Fu, B.; Chang, M. C. Y. Discovery and engineering of pathways for production of α-branched organic acids. *J. Am. Chem. Soc.* **2017***, 139*, 14526-14532.

15. Golan, D. E.; Yin, J.; Walsh, C. T.; Lin, A. J. Site-specific protein labeling by Sfp phosphopantetheinyl transferase. *Nature Protocols* **2006***, 1*, 280-285.

16. Zheng, J.; T. Keatinge-Clay, A. The status of type I polyketide synthase ketoreductases. *MedChemComm* **2013***, 4*, 34-40.

17. Keatinge-Clay, A. T. A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem. Biol.* **2007***, 14*, 898-908.

18. Zheng, J.; Taylor, C. A.; Piasecki, S. K.; Keatinge-Clay, A. T. Structural and functional analysis of A-type ketoreductases from the amphotericin modular polyketide synthase. *Structure* **2010***, 18*, 913-922.

19. Xie, X.; Garg, A.; Keatinge-Clay, A. T.; Khosla, C.; Cane, D. E. Epimerase and reductase activities of polyketide synthase ketoreductase domains utilize the same conserved tyrosine and serine residues. *Biochemistry* **2016***, 55*, 1179-1186.

20. Garg, A.; Xie, X.; Keatinge-Clay, A.; Khosla, C.; Cane, D. E. Elucidation of the cryptic epimerase activity of redox-inactive ketoreductase domains from modular polyketide synthases by tandem equilibrium isotope exchange. *J. Am. Chem. Soc.* **2014***, 136*, 10190-10193.

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

22. Østergaard, L. H.; Kellenberger, L.; Cortés, J.; Roddis, M. P.; Deacon, M.; Staunton, J.; Leadlay, P. F. Stereochemistry of catalysis by the ketoreductase activity in the first extension module of the erythromycin polyketide synthase. *Biochemistry* **2002***, 41*, 2719-2726.

23. Siskos, A. P.; Baerga-Ortiz, A.; Bali, S.; Stein, V.; Mamdani, H.; Spiteller, D.; Popovic, B.; Spencer, J. B.; Staunton, J.; Weissman, K. J.; Leadlay, P. F. Molecular basis of Celmer's rules: stereochemistry of catalysis by isolated ketoreductase domains from modular polyketide synthases. *Chem. Biol.* **2005***, 12*, 1145-1153.

24. Liu, C.; Yuan, M.; Xu, X.; Wang, L.; Keatinge-Clay, A. T.; Deng, Z.; Lin, S.; Zheng, J. Substrate-bound structures of a ketoreductase from amphotericin modular polyketide synthase. *Journal of structural biology* **1990***, 203*, 135-141.

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

Appendix 1: *Plasmids and oligonucleotides*

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

A

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
pFW3	1186					Cb^R; DszAT from Wong, F. T., et al., Biochemistry 2009, 49 (1) 95-102.
pBL12	3567					Km^R; LDD from Lowry, B., et al. J Am Chem Soc 2013, 135 (45) 16809-16812
pBL13	3568					Cb^R; $Mod1_{DEBS}$ from Lowry, B., et al. J Am Chem Soc 2013, 135 (45) 16809-16812
pBL36	3569					Cb^R; $Mod2_{DEBS}$ from Lowry, B., et al. J Am Chem Soc 2013, 135 (45) 16809-16812
pFW98	3570					Cb^R; DEBS2 from Lowry, B., et al. J Am Chem Soc 2013, 135 (45) 16809-16812
pFW100	3572					Cb^R; DEBS3 from Lowry, B., et al. J Am Chem Soc 2013, 135 (45) 16809-16812
pET16b-His_{10}Pres-$ACP_{DEBSMod6}$	3602	PresACP6_Fwd3/PresACP6_Rev	NdeI/BamHI	pET16b	Gibson assembly	Cb^R; $ACP_{DEBSMod6}$ with N-terminal Prescission-cleavable His tag
pFW3_F190V	2330	DszAT F190 F1/DszAT F190 R1 and DszAT F190 F2/DszAT F190 R2	XbaI/HindIII	pFW3	Gibson assembly	Cb^R; F190V DszAT
pAYC136	1150					Cb^R; $Mod3_{DEBS}$+TE(AT^0) from Wong, F. T., et al., Biochemistry 2009, 49 (1) 95-102.
pFW98_DEBS2 (Mod3 AT^0)	3571	n/a	MauBI/Pfl23II	pFW98	Restriction -ligation	Cb^R; DEBS2 (Mod6 AT^0; S653A)
pFW100_DEBS3 (Mod5 AT^0)	3573	pACYC184_CmOperon_pFW100_M5_Fwd_PacI/ pACYC184_CmOperon_pFW100_M5_Rev_SpeI, then pFW100_M5AT0_SfiI_F/pFW100_M5AT0_SfiI_R and pFW100_M5AT0_BsiWI_F/pFW100_M5AT0_BsiWI_R.	SfiI/BsiWI, then PacI/SpeI	pFW100	2-step Gibson assembly	Cb^R; DEBS3 (Mod5 AT^0; S642)

pFW100_DEBS3 (Mod6 AT^0)	3551	pACYC184_CmOperon_pBP130_M6_Fwd_PacI/ pACYC184_CmOperon_pBP130_M6_Rev_SpeI, then M6TE-SA-M6-RP/M6TE-SA-M6-FP and M6TE-SA-130-FP/M6TE-SA-130-RP	BbvCI/AjuI, then PacI/SpeI	pFW100	2-step Gibson assembly	Cb^R; DEBS3 (Mod6 AT^0; S2107A)

B

Name	Sequence
PresACP6_Fwd3	CATCTAGAAGTGCTTTTTCAGGGCCCGCATATGGCGGCCCCGGCGCG GGAGATGACGTCGCAGGAGT
PresACP6_Rev	TCCTTTCGGGCTTTGTTAGCAGCCGGATCCTCAGAGCTGCTGTCCTAT GTGGT
DszAT F190 F1	GCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGAGGATCGAGATCT CGATCCCGCGAAA
DszAT F190V R1	CGGTCGCATGAAGCGGGAATGAACAGCGGCGCTCACGCGCAGGACT GTGTACTTCTTCGC
DszAT F190V F2	GCGAAGAAGTACACAGTCCTGCGCGTGAGCGCCGCTGTTCATTCCCG CTTCATGCGACCG
DszAT F190 R2	GCCAACTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTGG TGGTGGTGGTGG
pACYC184_CmOperon_pFW100_M5_Fwd_PacI	CGGCCGGGTCGCCTACTGCCTGGGCCTGGA TTAATTAA TTACGCCCCGCCCTGCCACT
pACYC184_CmOperon_pFW100_M5_Rev_SpeI	AGTCGACCTCCACGCCCTGCGCGTAC ACTAGT AACAGGAGGGACAGCTGATAGAAACAGA
pFW100_M5AT0_SfiI_F	TCGCCTACTGCCTGGGCCTGGA
pFW100_M5AT0_SfiI_R	ATCTCGCCCTG CGC GTGGCCGA
pFW100_M5AT0_BsiWI_F	TCGGCCAC GCG CAGGGCGAGAT
pFW100_M5AT0_BsiWI_R	ACCTCCACGCCCTGCGCGTAC
pACYC184_CmOperon_pBP130_M6_Fwd_PacI	CGGCGCGAGCCAGTACCGCT TTAATTAA TTACGCCCCGCCCTGCCACT
pACYC184_CmOperon_pBP130_M6_Rev_SpeI	ACCCGGCGCGTCCGAGCTGA ACTAGT AACAGGAGGGACAGCTGATAGAAACAGA
M6TE-SA-M6-RP	TCGGACACCTCCGGCGCGAG
M6TE-SA-M6-FP	GGTGGAGGCGTCGGCGTTGC
M6TE-SA-130-FP	AACCCAGCAGCACCGGGCCC
M6TE-SA-130-RP	CTGGTGCTGCTGGGCAGGCG

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

A

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
p15A-DszAT	3558	DszATF1/ T7TerminatorR1 and LacCasetteLacI /LacCasetteLacO	ClaI/PacI	pJA4MCS2	Gibson assembly	Cm^R; lacI; pT7 lacO DszAT T7-term; p15A ori

p15A-DszATS86A	3560	S86AMutationF /DszATMutant_ CTerm	NdeI/HindIII	p15A-DszAT	Gibson assembly	Cm^R; lacI; pT7 lacO S86A DszAT T7-term; p15A ori
p15A-DszATF190V	3559	F190VMutation F/DszATMutant _CTerm	NdeI/HindIII	p15A-DszAT	Gibson assembly	Cm^R; lacI; pT7 lacO F190V DszAT T7-term; p15A ori
pET21c-FabD	3561	MalACP_F1/ MalACP_R1	NdeI/EcoRI	pET21c	Gibson assembly	Cb^R; FabD from *E. coli*
pAYC136	1150					Cb^R; $Mod3_{DEBS}$+TE(AT^0) from Wong, F. T., et al., Biochemistry 2009, 49 (1) 95-102.
pET-$Mod3_{DEBS}$ AT^0ACP^0	1567					Cb^R; $Mod3_{DEBS}$+TE(AT^0ACP^0) from Ad, O., et al., Proc Natl Acad Sci 2017, 114 (5) E660-D668.
pAYC138	1045					Cb^R; $Mod6_{DEBS}$+TE(AT^0) from Wong, F. T., et al., Biochemistry 2009, 49 (1) 95-102.
pFmal(MdcF)	1830					SpR; lacI; AraC; pBAD33 mdcF rrnB T1 T2-term; pT7 lacO matB T7-term; CloDF13 ori from Thuronyi et al. Angew Chem Int Ed 2017, 56(44) 13637-13640.
pFmal(MatC)	1549					SpR; lacI; AraC; pBAD33 matC rrnB T1 T2-term; pT7 lacO matB T7-term; CloDF13 ori from Thuronyi et al. Angew Chem Int Ed 2017, 56(44) 13637-13640.
pFmal(MadLM)	1547					SpR; lacI; AraC; pBAD33 madLM rrnB T1 T2-term; pT7 lacO matB T7-term; CloDF13 ori from Thuronyi et al. Angew Chem Int Ed 2017, 56(44) 13637-13640.
pFmal((MadLM, without MatB)	1929					SpR; lacI; araC; pBAD33 madLM rrnB T1 T2-term; CloDF13 ori from Thuronyi et al. Angew Chem Int Ed 2017, 56(44) 13637-13640.

B

Name	Sequence
DszATF1	ATTTTGTTTAACTTTAAGAAGGAGATATACATATGAAAGCATACATGTTTCCCGGGCAAG
T7TerminatorR1	CCCTGCAGCTTAAGTTAATTAACCCGGGGCAAAAAACCCCTCAAGACCCGTTTAGAG
LacCasetteLacI	ATTCTCATGTTTGACAGCTTATCATCGATACTGCCCGCTTTCCAGTCGGG
LacCasetteLacO	CCCTTGCCCGGGAAACATGTATGCTTTCATATGTATATCTCCTTCTTAAAGTTAAACAAA
S86AMutationR	GAACTCGCCCAGAGCGTGGCCGGCCAG
S86AMutationF	CTGGCCGGCCACGCTCTGGGCGAGTTC
F190VMutationR	GAAGCGGGAATGAACAGCGGCGCTCAC
F190VMutationF	GTGAGCGCCGCTGTTCATTCCCGCTTC
DszATMutant_CTerm	TGGTGGTGGTGGTGCTCGAGT
MalACP-F1	TTTTGTTTAACTTTAAGAAGGAGATATACATATGACGCAATTTGCATTTGTGTTCCCTGG
MalACP-R1	GGTGGTGGTGCTCGAGTGCGGCCGCAAGCTTGTCGACGGAGCTCGAATTAAGCTCGAGCGC CGCTGCCATCGCTGAAGG
PresACP6_Fwd3	CATCTAGAAGTGCTTTTTCAGGGCCCGCATATGGCGGCCCCGGCGCGGGAGATGACGTCGC AGGAGT

PresACP6_Rev	TCCTTTCGGGCTTTGTTAGCAGCCGGATCCTCAGAGCTGCTGTCCTATGTGGT

Table 3. Plasmid constructs (A), and oligonucleotides (B) gBlocks (C) and strains (D) used in Chapter 4.

A

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
pBP130	3547, 3548					Cb^R; DEBS2 DEBS3; ColE1 ori from Pfeifer, B. A., et al. Science 2001, 291 (5509) 1790-1792
pBP144	3549, 3550					Km^R; DEBS1 pccAB; ColE1 ori from Pfeifer, B. A., et al. Science 2001, 291 (5509) 1790-1792
pBPJW130	3544					Am^R; DEBS2 DEBS3; ColE1 ori from Zhang, H., et al. Chem Bio 2010, 17(11) 1168-1169
pBPJW144	3544					Tc^R; DEBS1 pccAB; ColE1 ori from Zhang, H., et al. Chem Bio 2010, 17(11) 1168-1169
pDEBS	3554, 3555					Km^R; *E. coli* codon-optimized DEBS123, groES/EL, pccAB, sfp from Biotechnol Prog 2018, 34 (1) 271-276
pCK7	2066					ColE1 SCP2* Cb^R Ts^R pactI actIIORF4 DEBS123 from Kao, C. M., et al. Science 1994, 265(5171) 509-512
pCas	2636					For Cas9-mediated genome editing in *E. coli* from Jiang, Y., et. al., Appl Environ Microbiol 2015, 58 (4) 161-170
pTargetF	2637					For Cas9-mediated genome editing in *E. coli* from Jiang, Y., et. al., Appl Environ Microbiol 2015, 58 (4) 161-170
pCAP01	3541					Km^R; CEN6_ARS4; TRP1; RP4 ori; attP; pUC ori from Yamanaka, K. Proc Natl Acad Sci 2014, 111(5) 1957-1962
pCas_PPhe-BsaI_NAT	2046					For Cas9-mediated genome editing in *S. cerevisiae* from Ryan, O. W., et al. Elife 2014, 3:e03703
pBP144_DEBS1 (Mod1 AT^0)	3643	Cm-pT7-PhaEC-F/Cm-pT7-PhaEC-R, then DEBS1-1/DEBS1-2, and DEBS1-3/DEBS1-4.	BsiWI/BstBI	pBP144	2-step Gibson assembly	Km^R; DEBS1 S1181A pccAB; ColE1 ori

pBP130_DEBS2 (Mod3 AT^0)	3552	pACYC184_CmOperon_pBP130_M3_Fwd_PacI/pACYC184_CmOperon_pBP130_M3_Rev_SpeI, then pET21-M3-RP/pET21-M3-FP2	MauBI/BaeI	pBP130	2-step Gibson assembly	Cb^R; DEBS2 S653A DEBS3; ColE1 ori
pBP130_DEBS3 (Mod5 AT^0)	3557	n/a	BbvCI/NsiI	pBP130	Restriction -ligation	Cb^R; DEBS2 DEBS3 S642; ColE1 ori
pBP130_DEBS3 (Mod6 AT^0)	3553	n/a	AsiSI/BbvCI	pBP130	Restriction -ligation	Cb^R; DEBS2 DEBS3 S2107A; ColE1 ori
pCAP01.DEBS.Actp	3579	n/a	XhoI/KpnI, BamHI	pCAP01	TAR	DEBS123 from pCK7 on pACP01 backbone

B

Name	Sequence
Cm-pT7-PhaEC-F	GGGCGCTGTTCGAGGCGTACG AACAGGAGGGACAGCTGATAGAAACAGA
Cm-pT7-PhaEC-R	CGGCGTTGCGGACGCGTTCGAA TTACGCCCCGCCCTGCCACT
DEBS1-1	GGGCGCTGTTCGAGGCGTACGGGCGCGACCGCGAGC
DEBS1-2	GGCGATCTCGCCCTGCGCGTGCCCGATGAC
DEBS1-3	GTCATCGGGCACGCGCAGGGCGAGATCGCC
DEBS1-4	CGGCGTTGCGGACGCGTTCGAACGGCCCCGTCGCGACG
pACYC184_CmOperon_pBP130_M3_Fwd_ PacI	TGCACGAGCGGCCCCGCGCG TTAATTAA TTACGCCCCGCCCTGCCACT
pACYC184_CmOperon_pBP130_M3_Rev_ SpeI	TGGCGCGGCTGCACCACCCC ACTAGT AACAGGAGGGACAGCTGATAGAAACAGA
pET21-M3-RP	CAGGCGACCGAGGTGCACGA
pET21-M3-FP2	CCGACCGGGGCTGGGACCTG
M5_BbvCI	TCAGCTCCTCCCTCAGCTCG
M5_NsiI	CTGAGCATGCATGGGTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATA TACATATGAGCGGTGACAACGGCAT
FadR_gibsonF	GTCCTAGGTATAATACTAGTCTGGATTACAACATATTCCGGTTTTAGAGCTA GAAATAGC
FadR_gibsonR	GCTATTTCTAGCTCTAAAACCGGAATATGTTGTAATCCAGACTAGTATTATA CCTAGGAC
FabD_GibsonF	GTCCTAGGTATAATACTAGTTTCATCAGCGCACAGTGAGAGTTTTAGAGCTA GAAATAGC
FabD_GibsonR	GCTATTTCTAGCTCTAAAACTCTCACTGTGCGCTGATGAAACTAGTATTATA CCTAGGAC
FadR_SOEF2	CCACCATAAACATCAGTAACAGCAAG
FadR_SOEmidR	GGGTTTGCTCTTTAAACGGAAGGGAAGTGAGATTTCCATAACACAGCAAA
FadR_SOEmidF	TTTGCTGTGTTATGGAAATCTCACTTCCCTTCCGTTTAAAGAGCAAACCC
FadR_SOER2	CCAACGGAAGAAGACAAATACTGGTT
FabDRO_1F	CGGCGAGCTATCCAATTGTCGAAG
FabDRO_2F_Phe	AGCGTACCGTTTCACTGTGCGC
FabDRO_2F_Val	AGCGTACCGGTTCACTGTGCGC
FabDRO_1R_Ala	GCGCACAGTGAGCCGGTACGCT
FabDRO_1R_Phe	GCGCACAGTGAAACGGTACGCT
FabDRO_2F_Ala	AGCGTACCGGCTCACTGTGCGC
FabDRO_1R_Val	GCGCACAGTGAACCGGTACGCT
FabDRO_2R	GGTCGGTCACATTCAACATCAGACCTTT
FabDRO_g1_Mutate_R	GCTGCTGGCTTCATGAGTGCACAGTG

FabDRO_g1_Mutate_F	CACTGTGCACTCATGAAGCCAGCAGC
pCAP01ActpDEBS1F1	AGATCGACTAGTAACCTCGAGCCATGG
pCAP01ActpDEBS1R1	ATTTTTCTAAATACAGGTACCCTATGAATTCCCTCCG

C

Name	Sequence
CapActPDEBS1	AGATCGACTAGTAACCTCGAGCCATGGTCGTCATCGCCAAGGGCGCCTCGGCGGACCAGGTGCACGC AGCTCTGGAGACCGTGCCCGGCGTCATCGAGGTCGCACCACCACAGGTGAAGGACGGCCTCGCCTAT GTCGAGGCGACGCTCG
CapActPDEBS2	CGAGGATCACCGGAATGATCAGACCGCGGTCGCGGCTGGACGCCTCCCGCATGTCGTGCACGACGGC GCTGCTGCCGCCCACCCGTGCCTGCGCACCGTCCAGCCGGGCGAGTGTCTCGCGGGCCGCGGTGAC CGAACGCATGGCCGCCGGCGAGTCCGCCCCGGCGCCGAGCGTCGCCTCGACATAGG
CapActPDEBS3	GGTCTGATCATTCCGGTGATCCTCGCCGTGGTGTTCTGCATCCTCGCGCTGCTGCTGCGAGCACTGGT GGCGCCGCTGCTGCTGATCGCGAGCGTGGTGCTGTCCTTCTTCACCGCGCTCGGCCTCGCCGCCCTC TTCTTCAACCACGTCTTCGACTTCGCGGGAGCCGATTCGGCCTTCCCCCTGTGGGTCTTCGTCTTCCTG GTCGCCCTGGGCGTCGACTACAACATCTTCCTCGTCACCCGGATCAAGGAGGAGAGCGACCGGCTCG GCACCCGGCAGGGCGCGCTCAAGGGCCTCACCTCGACCGGTGGCGTGATCACCGCGGGGATCCGAG CTGACCGCCACGCTGTTCGACCGCGAGACGGTGCGGATGGACGACACCAGGCTCACCGCCCTGGGC GCCTACGACCGCCTCACCGGTCAGTGGCGACCCCGGGAAACCGGGCTGCCGACGCTGCTGGTCAGC GCCGGCGAGCCGATGGGTCCGTGGCCCGACGACAGCTGGAAGCCGACGTGGCCCTTCGAGCACGAC ACCGTCGCCGTCCCCGGCGACCACTTCACGATGGTGCAGGAACACGCCGACGCGATCGCGCGGCACA TCGACGCCTGGCTGGGCGGAGGGAATTCATAGGGTACCTGTATTTAGAAAAAT

D

Name	No.	Genotype
BAP1 *ΔfadR*	3644	F−*ompT hsdS*$_B$(r$_B$− m$_B$−) *gal dcm* (DE3) *ΔprpRBCD* (*sfp*) *ΔfadR*
BAP1 *fabD* S200V	3645	F−*ompT hsdS*$_B$(r$_B$− m$_B$−) *gal dcm* (DE3) *ΔprpRBCD* (*sfp*) *fabD (S200F)*
BAP1 f*abD* S200F	3646	F−*ompT hsdS*$_B$(r$_B$− m$_B$−) *gal dcm* (DE3) *ΔprpRBCD* (*sfp*) *fabD (S200V)*

Table 4. Plasmid constructs (A), and oligonucleotides (B) and gBlocks (C) used in Chapter 5.

A

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
pET28a-His6TEV-KR$_{TYLSMod1}$	3648	TKR1Fwd/ TKR1Rev	BsaI	pET28a(gg)-RFP	Golden gate assembly	Km^R; lacI, pT7 lacO codon optimized KR$_{TYLSMod1}$ T7-term; CoE1 ori
pET28a-His6TEV-KR$_{DEBSMod1}$	3647	EryKR1Fwd/ EryKR1Rev2	BsaI	pET28a(gg)-RFP	Golden gate assembly	Km^R; lacI, pT7 lacO KR$_{DEBSMod1}$ T7-term; CoE1 ori
pET28a.HisN.PIKS-KR1	1324					Km^R; lacI, pT7 lacO codon optimized KR$_{PIKSMod1}$ T7-term; CoE1 ori

B

Name	Sequence
TKR1Fwd	CACACCA GGTCTC A TATG ATGAGCCCTACCGACGCATGG
TKR1Rev	CACACCA GGTCTC A AAGC TCATTAGGCCGCCGTCAGC
EryKR1Fwd	CACACCA GGTCTC A TATG TCCACCGAGGTCGACGAGGTTTCCG

EryKR1Rev2	CACACCA GGTCTC A AAGC TCGTGG TTCGGCCGCCGCCT

C

Name	Sequence
Tylosin KR6698-8168	ATGAGCCCTACCGACGCATGGCGCTATCGTGTTACTTGGAAGGCGCTTACCGAATCCTCACCGGTCCG CCCGCATTCTATCGGCCGTTGCCTGCTGGTAGCTCCTCCAACGACCGATGGGGAGCTGCTGGACGGC CTTACAACCGTCCTTTCCGAACGTGGGGCAAGCGTAGCGCGCTTAGAAGTACCGATTGGCGCACGCCG GGCGGAAGTTGCCGAGTTACTGAAACCATCTATGGAATCAGCTGGTGAAGAAAATACGACCGTTGTCAG TTTGCTCGGCCTGGTGCCCTCAACTGACGCTGTGCGTACGAGCATTGCGCTGCTGCAGGCCGTGAGTG ACATTGGTGTACCGGCCGCCCGGGTTTGGGCGCTGACACGTCGTGCCGTCGCTGTCGTTCCAGGGGA AACCCCGCAGGATGCGGGCGCTCAACTCTGGGGATTTGGCCGTGTAGCTGCGCTTGAATTGCCTGATA TTTGGGGAGGACTGATTGATCTTCCTGAAACGGCGGAACTCACCCGTACCCCGGAGACAAGCCAGCCA CCCCAGACCCCTGAGCGCTTGCCGCAGACGCCCAACCGTCGGGCACTGGAACTGGCGGCCGCGGTC CTGGCGGGCCGCGATGGTGAGGATCAGGTTGCGGTTCGCGCATCGGGCATTTACGGGCGTCGTGTGT CTCGCGCGGCGGCAGCGGGTGCAGCCTCTTGGCAACCGTCTGGTACCGTGCTGATTACTGGCGGTAT GGGAGCCATCGGGCGCCGCTTGGCACGCCGCTTAGCGGCTGAGGGCGCTGAACGCTTGGTTTTGACC AGTCGGCGTGGTCCGGAGGCCCCGGGCGCAGCTGAATTAGCTGAAGAATTGCGTGGCCACGGATGCG AAGTTGTCCATGCTGCTTGCGATGTGGCCGAACGCGATGCCCTGGCGGCACTGGTGACCGCCTATCCA CCCAATGCCGTATTTCACACCGCCGGCATTCTGGATGACGCGGTTATTGATACGCTTTCCCCAGAATCC TTTGAAACCGTGCGTGGCGCCAAGGTTTGTGGCGCGGAGCTGCTGCACCAGCTCACCGCAGACATCAA GGGCCTGGATGCCTTTGTCTTGTTTTCGTCGGTTACCGGTACGTGGGGCAATGCCGGCCAGGGCGCGT ACGCCGCGGCCAACGCCGCATTAGACGCCTTGGCCGAGCGTCGGCGGGCTGCAGGCTTACCGGCAA CATCTGTAGCTTGGGGTCTGTGGGGTGGTGGCGGGATGGCTGCTGGGGCCGGCGAAGAATCATTGAG TCGCCGCGGCCTGCGCGCAATGGATCCAGATGCCGCGGTGGATGCCTTGTTAGGTGCTATGGGGCGC AATGATGTTTGCGTGACGGTGGTAGACGTGGATTGGGAACGTTTCGCCCCGGCGACAAACGCTATCCG TCCGGGACGTTTATTCGACACTGTCCCGGAGGCACGGGAAGCGCTGACGGCGGCCTAATGA