Title
Engineered Biosynthesis of Fluorinated Polyketides

Author
Thuronyi, Benjamin Williams

Publication Date
2015

Peer reviewed|Thesis/dissertation

Engineered Biosynthesis of Fluorinated Polyketides

by

Benjamin Williams Thuronyi

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy in
Chemistry
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Michelle C. Y. Chang, Chair
Professor Matthew B. Francis
Professor John E. Dueber

Spring 2015

Engineered Biosynthesis of Fluorinated Polyketides

by Benjamin Williams Thuronyi

Abstract

Engineered Biosynthesis of Fluorinated Polyketides

by

Benjamin Williams Thuronyi

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michelle C. Y. Chang, Chair

Natural products use exquisite structural complexity to tailor a limited repertoire of functional groups for different molecular functions. These molecules have been adopted or adapted by humans to create a large fraction of current therapeutics. However, those adaptations have not always been able to combine the chemical diversity possible for totally synthetic compounds with the many advantages of biosynthesis. Extremely rare in nature but ubiquitous in synthetic pharmaceuticals, fluorine is a unique tool for molecular design, and an excellent candidate for expanding biosynthetic diversity. This dissertation describes our work to enable efficient fluorine incorporation into complex biosynthetic pathways, both *in vitro* and in the context of a whole organism, and our progress in enabling fluorinated polyketide biosynthesis. We hope to set the stage for a dramatic expansion in polyketide structural diversity *via* engineered biosynthesis, using a design element that could improve bioactive properties and furnish new therapeutics.

Table of Contents

Chapter 4: Engineered *trans*-acting acyl transferases: a strategy for efficient, regioselective fluorine incorporation into polyketides

Appendices

List of Figures, Schemes, and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Abbreviations

ACCase	acetyl-CoA carboxylase
AckA	acetate kinase
ACP	acyl carrier protein
ACS	acetyl-CoA synthetase
AT	acyltransferase
bp/kb	base pair/kilobase pair
CoA	coenzyme A
cryo-EM	cryo-electron microscopy
6-dEB	6-deoxyerythronolide B
DEBS	6-deoxyerythronolide B synthase
DH	dehydratase
DszAT/DszsAT	disorazole synthase acyltransferase
EI	electron ionization
ESI	electrospray ionization
ER	enoyl reductase
F-TKL/FTK	2-fluoro-3-keto-4-methyl-5-hydroxy-*n*-heptanoic acid δ-lactone
GC-MS	gas chromatography/mass spectrometry
HPLC	high performance liquid chromatography
H-TKL	3-keto-4-methyl-5-hydroxy-*n*-heptanoic acid δ-lactone
KR	ketoreductase
KS	ketosynthase
LC/MS	liquid chromatography/mass spectrometry
MS	mass spectrometry
NMR	nuclear magnetic resonance
nt	nucleotide
PHA	poly(hydroxyalkanoate)
PHB	poly(hydroxybutyrate)
PKS	polyketide synthase
Pta	phosphotransacetylase
RBS	ribosome binding site
SNAC	N-acetyl cysteamine (thioester)
TKL	2,4-dimethyl-3-keto-5-hydroxy-*n*-heptanoic acid δ-lactone

Acknowledgments

Science is only as fun as the people you get to do it with. I've been extremely lucky to do much of this work in close collaboration with Dr. Mark Walker, a superb scientist and a deeply caring person. Mark's support and encouragement made every challenge we faced together seem easier. Another vital collaborator whose development as a researcher and scholar I've been privileged to witness – and hopefully facilitate – over many years has been Tom Privalsky. Tom's dedication and curiosity have been an inspiration and he's taught me at least as much about mentoring as I ever taught him about science. The same goes for the many wonderful rotation students who have come through the lab. Their enthusiasm made them a pleasure to work with, and the speed with which they mastered often totally unfamiliar projects and methods never ceased to amaze me. Thanks especially to Ioana Aanei, Thom Brewer, Shion An, Jackie Blake-Hedges, Tammy Hsu and Akiko Carver.

Prof. Michelle Chang has made everything presented here possible. Her talent as a scientist and her unstinting efforts on behalf of myself and all of her group members have been immensely valuable, but her greatest accomplishment may be assembling such an extraordinary group of people to work together. Every one of my labmates has helped me to make progress, stay sane, and have a good time, even when science seemed to have a grudge against me, or half the equipment in the building seemed to be broken. They have all been counselors, cheerleaders and scientific advisors, and above all wonderful (and tolerant) friends, and I'll miss them greatly.

The work I present here has also benefitted enormously from the involvement of many collaborators outside our research group. Brian Lowry, Tom Robbins, Briana Dunn, Fong Wong and Prof. Chaitan Khosla were instrumental to our success. In particular, Dr. (now Professor) Lou Charkoudian has been an unfailing source of wisdom and encouragement. Many members of the Center for Sustainable Polymers, including Prof. Marc Hillmyer, Angelika Neitzel, Prof. Geoff Coates and Prof. Kechun Zhang, have provided valuable help and advice. Much of my research has involved analytical challenges, and would have been impossible if not for the efforts of Dr. Andreas Lingel, Dr. Zhongrui Zhou, Dr. Jeff Pelton and Dr. Chris Canlas.

Thanks, finally, to my friends and family for their love and support, and their willingness to share me with graduate school.

Chapter 1: *Introduction*

Portions of this work were published in the following scientific journal:

Thuronyi, B. W., Chang, M. C. Y. Synthetic biology approaches to fluorinated polyketides. *Accounts of Chemical Research* **2015**, 10.1021/ar500415c.

1.1. Natural products and their analogs in medicine

Antibiotics are a hugely important part of modern medicine. Bacterial infections that would once have meant a death sentence have become only an inconvenience. However, resistance to antibiotics has posed a problem since the first antibiotic was discovered. There is often a relatively short window between the introduction of a new antibiotic to the clinic and emergence of bacterial resistance to that antibiotic (*Figure 1.1*), and resistance mechanisms may be as ancient as antibiotics themselves [*1,2*]. This is at least in part because antibiotic use imposes a strong selective pressure on bacteria to either develop resistance *de novo* or enrich for resistance that already exists within a population. Multi-drug resistant bacteria are an increasingly common cause of death, and represent a growing public health crisis [*3*].

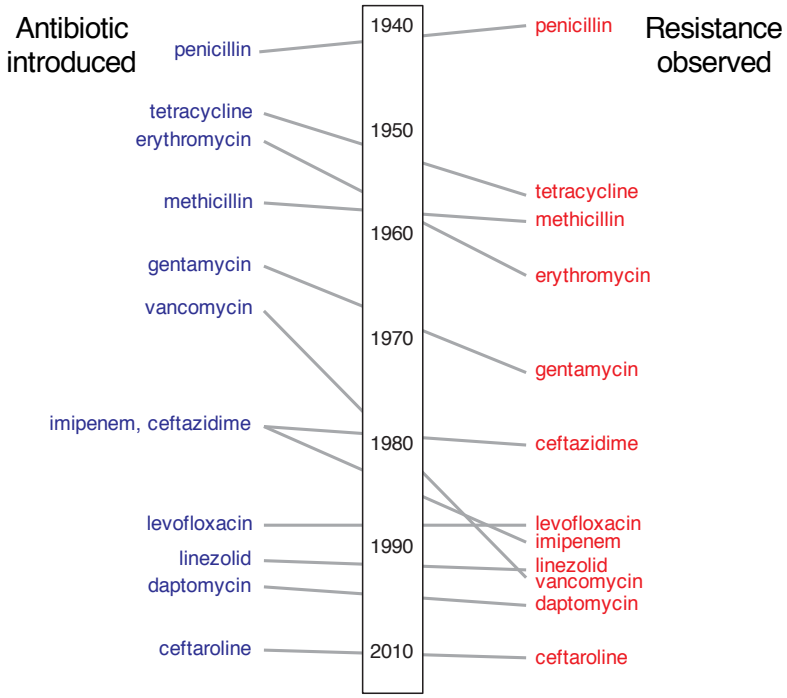


***Figure 1.1**. Timeline showing the dates of clinical introduction of major antibiotic classes and the first observation of resistant bacteria. Adapted from Antibiotic Resistance Threats in the United States, 2013, CDC.*

Antibiotic use places us in an arms race with bacteria, requiring us to use existing antibiotics judiciously, but also to continue introducing new ones, preferably with novel mechanisms of action. However, this process has not proven easy to maintain. Instead, antibiotic resistant strains are emerging more rapidly while introduction of new antibiotics has declined. The peak years for FDA antibiotic approvals were 1980-1995, followed by a steep drop (*Figure 1.2*).

One of the main causes of this decline is the nature of the source for the vast majority of antibiotics (and indeed, of a significant fraction of all human therapeutics): natural products and their derivatives (*Figure 1.2*) [*2,4*]. Screening microbial isolates from soil samples or other sources for antibiotic activity has revealed numerous classes of antibiotic natural products. However, continued screening efforts suffer from diminishing returns as a higher and higher percentage of hits represent known compounds. Both the difficulty and importance of natural

antibiotic sources are highlighted by the recent excitement over an ingenious method to make difficult-to-culture organisms accessible to screening, resulting in discovery of a new antibiotic [*5*]. Furthermore, considerable resources are being devoted to metagenomic sequencing and high throughput or *in silico* prediction of natural product biosynthetic clusters that may produce leads for new drugs [*6*-*8*].

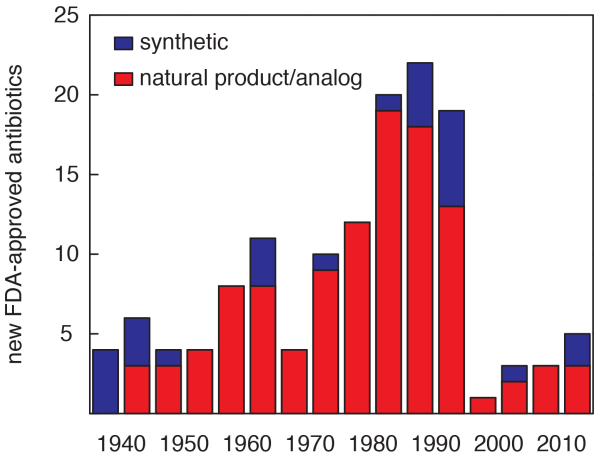


***Figure 1.2**. FDA antibiotic approvals over time. Antibiotics that are unmodified natural products, are close analogs of natural products, or incorporate a natural product pharmacophore are shown in red; synthetic antibiotics are in blue. Data is compiled from Wikipedia and [2].*

Along with discovery efforts, another critical source of new antibiotics comes from derivatization of natural products or generation of structural analogs. Strikingly small changes to drug structures can improve potency, increase effectiveness against drug-resistant bacteria, or alter distribution, bioavailability and stability of drugs. Therefore, making chemical modifications and exploring the structure-activity relationships of existing drugs – and natural products – is a rich source of new therapeutics.

However, methods to carry out such modifications are often limited. A hallmark of natural products is their structural complexity, including multiple stereocenters and diverse functional groups. These compounds are accessible by total synthesis but usually only at considerable effort and expense. The most common method to produce natural product analogs is to chemically derivatize the natural product itself, which is obtained from the producing organism (or sometimes a heterologous host), in a process known as semisynthesis. This strategy takes advantage of the relatively high efficiency and inexpensive feedstocks of native producer biosynthetic machinery, but it is constrained by the chemical reactivity of the natural product; not every desired transformation can be carried out selectively without unwanted side-reactions or degradation.

A desirable goal, therefore, is a method that combines the fine structural control available to total synthesis with the efficiency and selectivity of biosynthetic enzymes. Considerable progress has been made to engineer and adapt biosynthetic assembly lines to accomplish this. The natural product biosynthesis enzymes that are most amenable to engineering are modular assembly lines, where the natural product is a tailored polymer and each monomer unit is added by a distinct biosynthetic module. Many polyketides and all non-ribosomal peptides, which together represent more than 10^4 natural products, are made this way. This programmed biosynthesis can then be altered to make specific changes to the output structure.

Modular polyketide synthases (PKSs) specifically have been engineered in several ways that change the chemical composition of the backbone (*Figure 1.3*). Blocked assembly lines, where the initial modules have been inactivated, can be used to extend a completely synthetic precursor to a full-length natural product derivative [*9,10*]. The module that selects the organic acid starter unit, the monomer that initiates biosynthesis, is often promiscuous or can be replaced with a promiscuous module from another PKS [*11*]. To some extent, modules can be swapped and recombined, or specific tailoring domains inactivated, resulting in different chain tailoring and monomer arrangements [*12-14*]. However, such dramatic changes as module swapping almost invariably lead to dramatic reductions in assembly line function, and even blocking tailoring steps like ketoreduction or dehydration can interfere with downstream reactions in subtle but significant ways. Finally, if a unique chain extension monomer is present in the structure, its biosynthesis can be blocked and a chemically similar monomer unit fed as a substitute and incorporated into the final product [*15,16*].

However, thus far, replacement or modification of the non-unique extender units that make up the bulk of each polyketide's structure has not been feasible. The enzymatic domains that select these monomer units were until recently thought to be extremely selective for their native substrates, excluding non-native extender units from incorporation. However, recent work has shown that many extender selection domains can exhibit substrate promiscuity, particularly when presented with extender units that are not available in their native context [*17-20*].

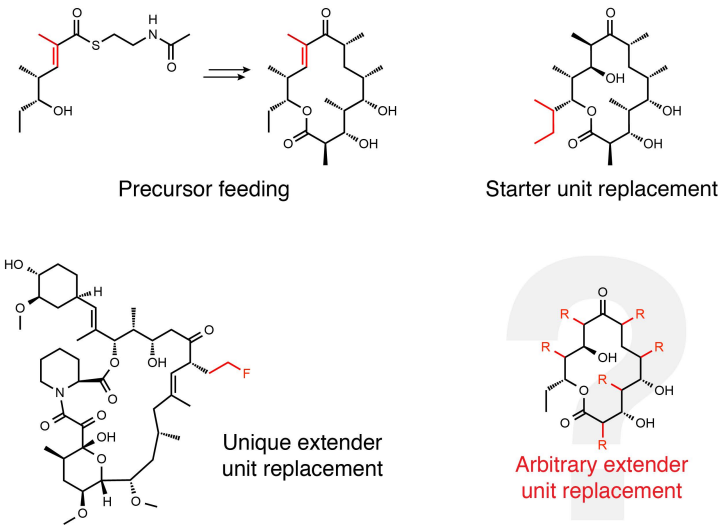


***Figure 1.3**. Strategies to modify biosynthesis of polyketides.*

A general strategy to introduce new monomers to natural product synthesis, particularly for natural product classes such as polyketides that rely on a relatively small native monomer pool, would dramatically expand natural product structural space. Furthermore, doing so in a modular, minimally perturbative manner could enable generation of large libraries of natural product analogs from diverse biosynthetic assembly lines. Monomer replacement has the potential to make large changes to the chemical composition of a natural product, but may avoid some of the difficulties associated with module rearrangement and modification if it can be accomplished in the context of a largely intact, native assembly line.

1.2. Fluorine in medicinal and biological chemistry

Fluorine-containing molecules are a rarity in nature, with fewer than 25 known examples including structural analogues, most of which stem from a single biosynthetic pathway found in *Streptomyces cattleya* (*Figure 1.4*) [*21*]. By contrast, other characterized organohalogens number in the tens of thousands. This is probably because fluorine is difficult both to access and to activate in an aqueous biological setting. Although environmental fluorine is not uncommon,

most of its mineral salts are highly insoluble, and solvated fluoride ion is a very poor nucleophile because of its tightly bound solvent shell. Unlike other halogens, it is extremely difficult to oxidize, excluding it from many of the pathways used in organohalogen biosynthesis. The sole characterized organofluorine synthetic enzyme, the fluorinase FlA, desolvates fluoride ion and forms 5′-fluorodeoxyadenosine by nucleophilic substitution [*22,23*].

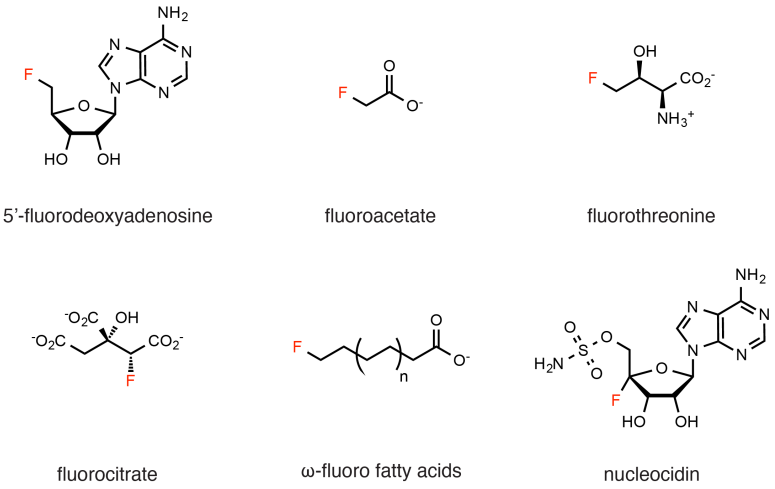


Figure 1.4*. A selection of naturally occurring organofluorines.*

In contrast to natural products, anthropogenic molecules make heavy use of fluorine. Its distinctive and dramatic properties make it useful in drug design, as well as a unique tool for studying enzyme mechanism and metabolism. Fluorine's small size and high electronegativity confer powerful effects on small molecule properties, and drug fluorination can improve bioavailability, reduce clearance, alter pK_a and lipophilicity, block metabolism, or increase potency. Fluorine is now found in 20-30% of therapeutics and agrochemicals.

Because organofluorines are so rare in biological systems, few enzymes have evolved that specifically recognize fluorine-containing substrates. The small organofluorine building blocks produced by the *S. cattleya* pathway, fluoroacetate and fluorothreonine, have not been found to be elaborated into more complex molecules, with the exception of omega-fluoro fatty acids, where fluoroacetate acts only as a chain initiator. Therefore, although natural products are extremely diverse and many are highly bioactive, almost none of them leverage the molecular properties that fluorine can confer. This element has been excluded from the evolution of biosynthesis, and introducing it to the pool of available design elements may produce new molecules with improved characteristics. This makes it an excellent candidate for a natural product monomer.

Organofluorine physical organic properties. The physical organic properties of fluorine provide the foundation (and rationale) for expanding its role in biology [*24-27*]. Fluorine is considered a sterically conservative substitution for hydrogen or oxygen; the C–F unit is similar in size to C=O [*25,27*]. This means it can be accommodated by many scaffolds without disrupting a molecule's fit in a binding site, an important qualification for being used to derivatize pharmaceuticals and other bioactive compounds.

The C–F bond is extremely strong, largely because it is so polarized that it has substantial ionic character ($C^{\partial+}$–$F^{\partial-}$) [*28*]. Aliphatic fluorine substituents are therefore fairly inert, with substitution reactions uncommon and elimination requiring an $E1_{CB}$ mechanism and full carbanion formation [*28*]. This can be useful for blocking metabolism at specific sites in a drug. C–F bonds are impervious to enzymes that oxidize organic compounds by hydrogen abstraction, and this can give fluorinated compounds much slower clearance rates.

Fluorine substitution has strong inductive effects on neighboring functional groups, activating electrophiles and increasing acidity. However, α-fluoro carbonyl compounds are a special case and are less acidic than would be expected. While a fluoroenolate is stabilized by induction, it is concomitantly destabilized because fluorine prefers to be bonded to an sp^3 rather than an sp^2 carbon; this is due to unfavorable overlap between the filled p orbitals on F and C-sp^2 and reduction in σ-electron density that can be polarized toward F [*27,28*]. A methyl group has the opposite hybridization preference, so these two substituents can confer similar acidity. By contrast, purely inductive effects make fluoroacetic acid over 2 pK_a units more acidic than propionic or acetic acid.

Organofluorines are primarily stabilized by dipolar electrostatic interactions [*25,28*]. Fluorine is so nonpolarizable that induced dipole effects are weak, and C–F is a poor hydrogen bond acceptor, about half as strong as C–OH [*27,28*]. However, the C–F dipole does interact favorably with centers of positive charge, as with the "fluorophilic" Arg side-chain, where the C–F bond adopts either a perpendicular or an out-of-plane parallel orientation to the plane of the guanidinium group, rather than the in-plane linear hydrogen bonding geometry [*25*]. It is also favorable to orient the C–F dipole toward the carbonyl carbon of a backbone amide or Asn/Gln side-chain (0.2-0.3 kcal/mol) [*29*], or toward the peptide C_α–H bond. [*25*]

Because the C–F dipole can only form strong, favorable interactions within significant orientational constraints, organofluorines are naturally suited to be ligands. Macromolecule binding reduces ligand entropy in most cases, and this cost can be offset by the enthalpic gain from a favorable C–F dipole arrangement [*25*]. Furthermore, the poor hydrogen-bonding ability of C–F and its low polarizability make solvation in water relatively unfavorable [*30*]. Since C−F is a poor replacement for H_2O, only a few arrangements of water molecules will have their dipoles properly oriented to interact favorably with C–F, leading to entropically unfavorable ordering of the solvent. As a result, when the C–F unit is desolvated upon macromolecule binding the release of ordered water provides an entropic driving force, even in the absence of specific C–F dipolar interactions in the binding site. A very polar C–F unit can therefore bind favorably in a nonpolar pocket, leading to the description of organofluorines as "polar hydrophobic." [*30*]

Fluorine as biochemical probe. Besides its pronounced and potentially useful effects on molecular properties, fluorine makes an excellent biochemical tool, either to elucidate how natural systems function or to develop new biochemistry [*31*]. Fluorine's small size means that many enzymes that accept H, OH, CH_3 or C=O substituents can tolerate fluorine as well. Fluorine can be used as a stereo- or regiochemical label, one that is much easier to install than ^2H, albeit a less perfect ^1H mimic. This approach has been used to investigate enzyme stereochemistry, since prochiral methylene units become chiral when one hydrogen is replaced by fluorine and otherwise cryptic enzyme stereospecificity can therefore be directly observed [*32,33*].

The high natural abundance (100%), sensitivity (γ = 252 rad s^{-1} T^{-1}, compared to 268 for ^1H or 67 for ^{13}C) and wide chemical shift range of ^{19}F, along with its virtual absence from natural biological systems, makes it a superb tool to probe biochemistry using NMR. Fluorination of the input molecule for a complex metabolic pathway makes it possible to track and identify all intermediates in that pathway without needing any *ab initio* information about them, as is often necessary when using mass spectrometry for the same purpose. The regiochemistry of transformations can be followed and the complete fate of the starting metabolite revealed, even

in a complex biological mixture such as cell lysate. Fluorine NMR has also been used for high-throughput screening in drug discovery. Many individual fluorine-tagged compounds in a single multiplexed assay mixture can be distinguished by chemical shift, and those that bind to a protein target identified based on perturbations in chemical shift or relaxation time [*34,35*]. Fluorinated amino acids can also be used to study protein dynamics or conformational change using NMR [*36*].

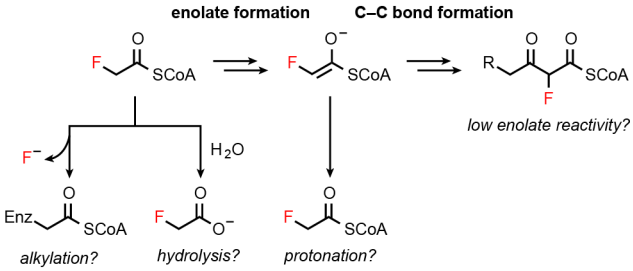


***Figure 1.5**. Possible derailment pathways for a fluorinated chain extender in polyketide biosynthesis.*

Reactivity of fluorinated building blocks. Use of fluorinated monomers for polyketide synthesis will involve fluorine's effects on reactivity as well as molecular properties. The presence of a fluorine substituent at the carbonyl α-position in a starter or extender unit creates unique challenges by altering the chemistry and reactivity of acid activation intermediates, chain extension units and the growing polyketide chain itself (*Figure 1.5*). Fluorine activates compounds as electrophiles, which makes thioester-tethered biosynthetic intermediates more

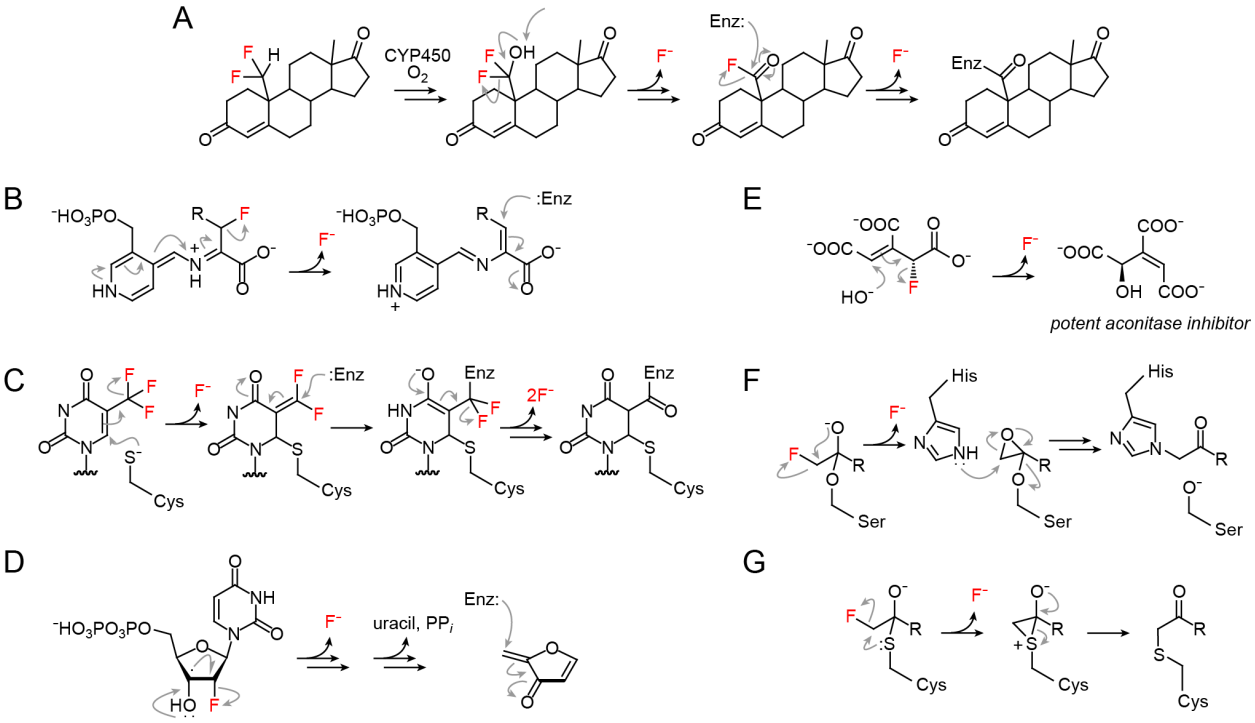


***Figure 1.6**. Reactions of enzymes with organofluorine substrates resulting in loss of fluoride or covalent enzyme inhibition. Only the key step(s) are shown. (A) Cytochrome P450 enzymes. (B) PLP-dependent enzymes. (C) Thymidylate synthase. (D) Ribonucleotide reductase. (E) Aconitase. (F) Serine proteases. (G) Cysteine proteases.*

labile to hydrolysis. For example, fluoroacetyl phosphate hydrolyzes 63-fold faster than the already highly activated acetyl phosphate at pH 7 [*37*]. Fluoroacetyl-CoA is hydrolyzed only 10-fold faster than acetyl-CoA at pH 7, but is up to 110-fold more susceptible when hydroxide ion is readily available at pH 11, a scenario that may mimic enzyme-catalyzed hydrolysis [*38*]. Therefore, the stability of the more labile intermediates in polyketide biosynthesis, particularly α-fluoro β-keto thioesters, will need to be addressed. Furthermore, fluorine substitution will

affect the kinetics and thermodynamics of extender unit decarboxylation, enolate generation and C–C bond formation during chain extension (*Section 1.3*).

A number of biocatalytic defluorination reactions have been documented, most of which form alkylating agents that act as covalent enzyme inhibitors (*Figure 1.6*) [*39*]. However, specific conditions – usually a source of electrons adjacent to fluorine, directly or via conjugation – are required for loss of fluoride, which is a poor leaving group especially in water. The α-fluorinated intermediates and reactions in polyketide synthesis should avoid such configurations. Cysteine proteases react with fluoromethylketones (FMKs), which resemble potential polyketide elongation intermediates, but FMKs are selective for this enzyme class and are very slow to alkylate serine proteases, suggesting that a specific configuration of active site residues is required for inhibition [*40,41*].

1.3. Polyketide synthases: mechanism and engineering

Polyketide synthases are powerful biosynthetic enzymes that produce structurally diverse natural products. They are classified based on architecture: type I PKSs include at least some multidomain, multifunctional proteins, while type II PKSs are complexes of exclusively monofunctional proteins. Both operate on thioester substrates that are covalently tethered to the enzyme during chain elongation. Type III PKSs have a single active site and operate exclusively on freely diffusing substrates. All type II and type III PKSs are iterative, carrying out multiple chain extensions using the same active site(s), but biosynthesis may still be programmed, with substrate specificities of different domains resulting in different chemical modifications or cyclizations at different elongation steps [*42,43*], and the same can be true of iterative type I PKSs (generally fungal PKSs are type I iterative) as well [*44-47*]. The basis for this substrate-directed programming provides many enzymological puzzles, but does not lend itself to facile reprogramming to generate new products.

Type I modular PKSs, by contrast, have been the subject of much study in part because they seem to present a clear paradigm for biosynthetic engineering. In typical type I modular systems, each module carries out a single chain extension event and any associated tailoring steps before passing the chain to the next module in the sequence. The biosynthetic outcome can therefore be read out of the primary sequence of the PKS, and in many cases it can be reprogrammed by changing the order, identity or functionality of domains or entire modules [*12,13*]. These PKSs are attractive platforms for monomer engineering, since individual steps in chain extension and the associated chain extender units can be targeted for modification.

The best understood type I modular PKS is the enzyme that synthesizes the structural core of erythromycin, the paradigmatic macrolide antibiotic. Erythromycin is derived from oxidation and glycosylation of 6-deoxyerythronolide B (6-dEB) (*Figure 1.7*), so the producing enzyme is called 6-**d**eoxy**e**rythronolide **B** **s**ynthase, or DEBS (*Figure 1.8*), a 2 MDa homodimeric complex composed of six modules (plus a loading didomain) spread across three polypeptides. The structure and function of DEBS has been subject to intensive study [*48,49*].

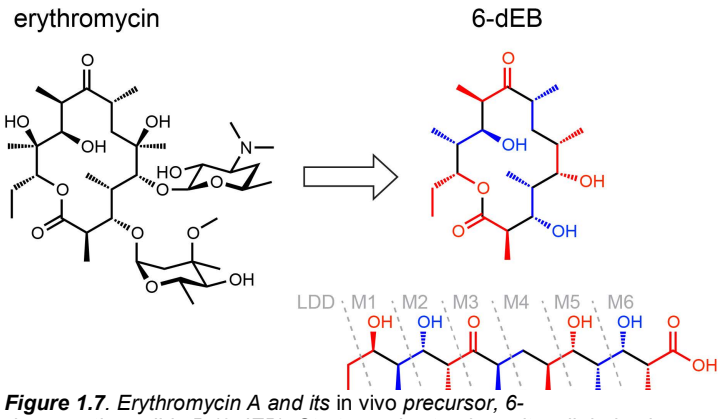


***Figure 1.7**. Erythromycin A and its in vivo precursor, 6-deoxyerythronolide B (6-dEB). Starter and extender units, all derived from propionic acid, are indicated in alternating colors. The module of DEBS that encodes each extender and ketone tailoring is shown at the lower right.*

Biosynthesis begins when a propionyl starter unit is transferred from propionyl-CoA to the acyl carrier protein (ACP) of the loading didomain (which lacks a ketosynthase [KS]), a reaction catalyzed by the acyl transferase (AT). ACP domains carry biosynthetic intermediates as thioesters using the thiol of a phosphopantetheine prosthetic group, a post-translational modification to a Ser residue. The starter unit is passed to Module 1, where it is transacylated from ACP_{LDD} to KS_1 on an active site Cys.

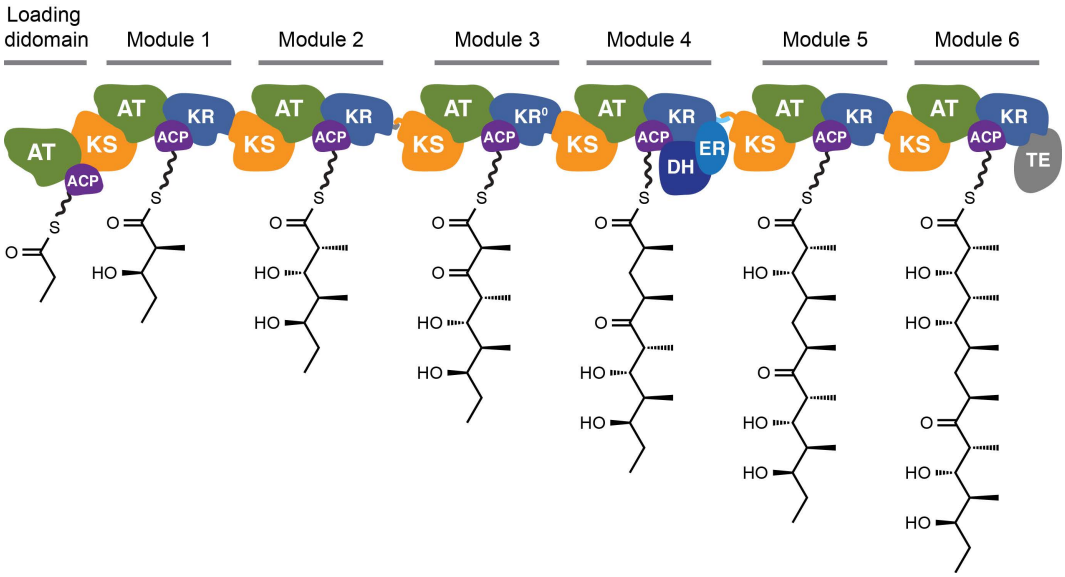


***Figure 1.8**. Modular architecture of the type I polyketide synthase DEBS. The biosynthetic intermediates for 6-dEB synthesis at each stage of chain extension are shown attached to each module's ACP, representing the step after chain extension and β-ketone tailoring steps carried out by that module have taken place. Phosphopantetheine groups are shown as wavy lines. Domain abbreviations: KS, ketosynthase. AT, acyl transferase. ACP, acyl carrier protein. KR, ketoreductase. DH, dehydratase. ER, enoyl reductase. TE, thioesterase.*

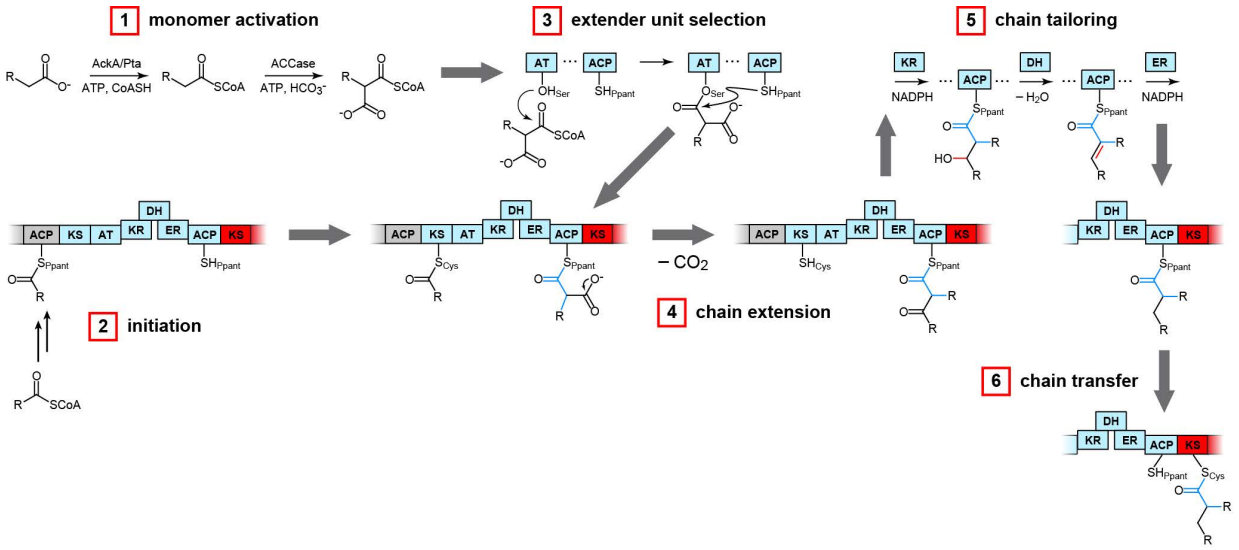


Figure 1.9. *Mechanistic steps in a single chain extension cycle catalyzed by a modular PKS (light gray, upstream module; light blue, active module; red, downstream module).*

At this point, chain extension begins using a similar series of steps for each module (*Figure 1.9*). An extender unit – an α-carboxylated acyl-CoA, most commonly biosynthesized either by an acyl-CoA carboxylase or crotonyl CoA reductase [*50*] – is transacylated from CoA onto the ACP by the AT domain, a reaction that proceeds through a covalent extender-AT intermediate. The extender-ACP then enters the KS active site, where a decarboxylative Claisen reaction leads to a new C–C bond. The growing chain is now attached to the ACP, and interacts with β-ketone processing domains, if present, which can carry out ketoreduction (ketoreductase, KR), dehydration (dehydratase, DH) and reduction to the alkane (enoyl reductase, ER). After any processing events take place, the ACP-linked chain is passed to the KS of the next module, and the cycle begins again. The successive reactions in this cycle do not rely on stochastic movement of the ACP from active site to active site. Instead, as stunning whole-module cryo-EM studies have recently revealed, the entire module is subject to conformational changes at each stage of chain elongation and processing [*51,52*]. These changes bias the ACP toward productive interactions, enforce unidirectional synthesis, and ensure that all processing events take place before the chain is passed to the next module – and go a long way toward explaining why chimeric or domain-inactivated modules do not work as efficiently as native ones.

Untangling the individual biochemical steps involved in polyketide synthesis, which is a concurrent process with many non-rate-limiting steps, has been difficult. DEBS, and other type I PKSs, can be dissected into single modules which catalyze one round of chain extension on substrates fed as N-acetyl cysteamine (SNAC) thioesters, making characterization easier. New tools, including sophisticated mass spectrometry techniques [*53,54*] and chemical probes [*55,56*], have made it possible to more directly observe intermediates in chain elongation.

These methods have recently been helping to overturn a long-held dogma in the polyketide biosynthesis field, that ATs are extremely specific for their cognate extender units. (The one notable exception is the epothilone PKS Module 3 AT, which has relaxed specificity between malonyl- and methylmalonyl-CoA [*57*].) Early studies that compared malonyl-CoA and methylmalonyl-CoA use by DEBS ATs showed extremely high levels of discrimination [*58,59*], and for some time no other substrates were tested. However, reports contemporaneous with the work presented here show that the observed specificity does not seem to extend to AT substrates

that are not natively available to DEBS, such as ethylmalonyl-CoA [*19*], and that completely unnatural extender units are surprisingly well-tolerated in some cases [*17,18*]. These findings may lead to a resurgence of interest in biosynthesis of polyketide analogues, because they provide an underexplored route to increasing structural diversity, potentially without compromising efficiency.

Despite these discoveries, prior to the work described here fluorinated extender units had not been examined in polyketide biosynthesis (except at specialized extender unit positions, see *Figure 1.3*). However, fluorinated derivatives of natural products have a rich history. Two fluorinated analogues of erythromycin, flurithromycin [*60*] and solithromycin, [*61*] have improved acid stability and potency, respectively, and numerous other fluorinated polyketides have been reported [*10,62-64*]. Other classes of natural product, such as terpenes, [*65*] non-ribosomal peptides and alkaloids [*66*] have been modified with fluorine as well, and fluorinated amino acids have made contributions to protein and peptide design [*67*]. A general route to incorporate fluorine into polyketides would expand these efforts.

1.4 Thesis organization

The motivation for this work is to build the tools needed for biosynthesis of fluorinated polyketides. Eventually, our goal is to enable fluorine incorporation at all possible backbone positions by any polyketide synthase, and to make it possible to do so with reasonable efficiency in a scalable *in vivo* system. This will require numerous enzymes from diverse sources, many of which will need to be adapted for fluorine tolerance or specificity. Host organism compatibility, monomer activation and polyketide synthesis must all be considered.

These aims are ambitious, and the results reported here represent only the first steps toward realizing them. Chapter 2 describes proof of concept studies showing that wild-type enzymes can activate fluoroacetate as a biosynthetic building block and that polyketide synthases can incorporate a fluorinated chain extender unit. It also presents preliminary evidence that these steps are compatible with *E. coli* metabolism. Chapter 3 expands on this by constructing and characterizing a model organofluorine biosynthetic pathway *in vivo* that closely resembles polyketide synthesis. Chapter 4 builds on another strategy introduced in Chapter 2, the use of *trans*-acyl transferase enzymes to introduce fluorinated extender units site-selectively and in a polyketide synthase-agnostic manner.

1.5 References

1. D'Costa, V. M., King, C. E., Kalan, L., Morar, M., Sung, W. W. L., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G. B., Poinar, H. N., Wright, G. D. Antibiotic resistance is ancient. *Nature* **2011**, *477* (7365) 457-461.

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

3. Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., Scheld, M., Spellberg, B., Bartlett, J. Bad bugs, no drugs: No ESKAPE! An update from the Infectious Diseases Society of America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2009**, *48* (1) 1-12.

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

5. Ling, L. L., Schneider, T., Peoples, A. J., Spoering, A. L., Engels, I., Conlon, B. P., Mueller, A., Schäberle, T. F., Hughes, D. E., Epstein, S., Jones, M., Lazarides, L., Steadman, V. A., Cohen, D. R., Felix, C. R., Fetterman, K. A., Millett, W. P., Nitti, A. G., Zullo, A. M., Chen, C., Lewis, K. A new antibiotic kills pathogens without detectable resistance. *Nature* **2015**, *517* (7535) 455-459.

6. Kersten, R. D., Yang, Y.-L., Xu, Y., Cimermancic, P., Nam, S.-J., Fenical, W., Fischbach, M. A., Moore, B. S., Dorrestein, P. C. A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nat Chem Biol* **2011**, *7* (11) 794-802.

7. Ribeiro, F. J., Przybylski, D., Yin, S., Sharpe, T., Gnerre, S., Abouelleil, A., Berlin, A. M., Montmayeur, A., Shea, T. P., Walker, B. J., Young, S. K., Russ, C., Nusbaum, C., MacCallum, I., Jaffe, D. B. Finished bacterial genomes from shotgun sequence data. *Genome Res* **2012**, *22* (11) 2270-2277.

8. Walsh, C. T., Fischbach, M. A. Natural products version 2.0: Connecting genes to molecules. *J Am Chem Soc* **2010**, *132* (8) 2469-2493.

9. Cane, D., Kudo, F., Kinoshita, K., Khosla, C. Precursor-directed biosynthesis:: Biochemical basis of the remarkable selectivity of the erythromycin polyketide synthase toward unsaturated triketides. *Chem Biol* **2002**, *9* (1) 131-142.

10. Eustáquio, A. S., Moore, B. S. Mutasynthesis of fluorosalinosporamide, a potent and reversible inhibitor of the proteasome. *Angew Chem Int Ed* **2008**, *47* (21) 3936-3938.

11. Marsden, A. F., Wilkinson, B., Cortes, J., Dunster, N. J., Staunton, J., Leadlay, P. F. Engineering broader specificity into an antibiotic-producing polyketide synthase. *Science* **1998**, *279* (5348) 199-202.

12. Menzella, H. G., Carney, J., Santi, D. Rational design and assembly of synthetic trimodular polyketide synthases. *Chem Biol* **2007**, *14* (2) 143-151.

13. Menzella, H. G., Reid, R., Carney, J. R., Chandran, S. S., Reisinger, S. J., Patel, K. G., Hopwood, D. A., Santi, D. V. Combinatorial polyketide biosynthesis by de novo design

and rearrangement of modular polyketide synthase genes. *Nat Biotechnol* **2005**, *23* (9) 1171-1176.

14. Petkovic, H., Lill, R. E., Sheridan, R. M., Wilkinson, B., McCormick, E. L., McArthur, H., Staunton, J., Leadlay, P. F., Kendrew, S. G. A novel erythromycin, 6-desmethyl erythromycin D, made by substituting an acyltransferase domain of the erythromycin polyketide synthase. *J Antibiot* **2003**, *56* (6) 543-551.

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

16. Mo, S., Kim, D. H., Lee, J. H., Park, J. W., Basnet, D. B., Ban, Y. H., Yoo, Y. J., Chen, S.-w., Park, S. R., Choi, E. A., Kim, E., Jin, Y.-Y., Lee, S.-K., Park, J. Y., Liu, Y., Lee, M. O., Lee, K. S., Kim, S. J., Kim, D., Park, B. C., Lee, S.-g., Kwon, H. J., Suh, J.-W., Moore, B. S., Lim, S.-K., Yoon, Y. J. Biosynthesis of the allylmalonyl-CoA extender unit for the FK506 polyketide synthase proceeds through a dedicated polyketide synthase and facilitates the mutasynthesis of analogues. *J Am Chem Soc* **2011**, *133* (4) 976-985.

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

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

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

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

21. Harper, D., O'Hagan, D., Murphy, C. Fluorinated natural products: Occurrence and biosynthesis. *Natural Production of Organohalogen Compounds* **2003**, 141-169.

22. Dong, C., Huang, F., Deng, H., Schaffrath, C., Spencer, J. B., O'Hagan, D., Naismith, J. Crystal structure and mechanism of a bacterial fluorinating enzyme. *Nature* **2004**, *427* (6974) 561-565.

23. Zhu, X., Robinson, D., McEwan, A., O'Hagan, D., Naismith, J. Mechanism of enzymatic fluorination in *Streptomyces cattleya*. *J Am Chem Soc* **2007**, *129* (47) 14597-14604.

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

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

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

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

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

29. Fischer, F. R., Schweizer, W. B., Diederich, F. Molecular torsion balances: Evidence for favorable orthogonal dipolar interactions between organic fluorine and amide groups. *Angew Chem Int Ed* **2007**, *46* (43) 8270-8273.

30. Biffinger, J. C., Kim, H. W., Dimagno, S. G. The polar hydrophobicity of fluorinated compounds. *ChemBioChem* **2004**, *5* (5) 622-627.

31. O'Hagan, D., Rzepa, H. S. Some influences of fluorine in bioorganic chemistry. *Chem Commun* **1997**, (7) 645-652.

32. Goldstein, J., Cheung, Y., Marletta, M., Walsh, C. T. Fluorinated substrate analogs as stereochemical probes of enzymatic reaction mechanisms. *Biochemistry* **1978**, *17* (25) 5567-5575.

33. Marletta, M., Walsh, C. T. Stereochemical outcome of processing of fluorinated substrates by ATP citrate lyase and malate synthase. *Biochemistry* **1981**, *20* (13) 3719-3723.

34. Vulpetti, A., Dalvit, C. Fluorine local environment: From screening to drug design. *Drug Discovery Today* **2012**, *17* (15-16) 739-746.

35. Vulpetti, A., Dalvit, C. Design and generation of highly diverse fluorinated fragment libraries and their efficient screening with improved 19F NMR methodology. *ChemMedChem* **2013**, *8* (12) 2057-2069.

36. Danielson, M. A., Falke, J. J. Use of 19F NMR to probe protein structure and conformational changes. *Annu Rev Biophys Biomol Struct* **1996**, *25* 163-195.

37. Marcus, A., Elliott, W. Fluoroacetyl phosphate; preparation and properties. *J Am Chem Soc* **1958**, *80* (16) 4287-4291.

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

39. Walsh, C. T. Suicide substrates: Mechanism-based enzyme inactivators. *Tetrahedron* **1982**, *38* (7) 871-909.

40. Imperiali, B., Abeles, R. H. Inhibition of serine proteases by peptidyl fluoromethyl ketones. *Biochemistry* **1986**, *25* (13) 3760-3767.

41. Powers, J., Asgian, J., Ekici, Ö., James, K. Irreversible inhibitors of serine, cysteine, and threonine proteases. *Chem Rev* **2002**, *102* (12) 4639-4750.

42. Austin, M. B., Bowman, M., Ferrer, J., Schroder, J., Noel, J. P. An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. *Chem Biol* **2004**, *11* (9) 1179-1194.

43. Belecki, K., Townsend, C. A. Environmental control of the calicheamicin polyketide synthase leads to detection of a programmed octaketide and a proposal for enediyne biosynthesis. *Angew Chem Int Ed* **2012**, *51* (45) 11316-11319.

44. Ames, B. D., Nguyen, C., Bruegger, J., Smith, P., Xu, W., Ma, S., Wong, E., Wong, S., Xie, X., Li, J. W.-H., Vederas, J. C., Tang, Y., Tsai, S.-C. Crystal structure and biochemical studies of the *trans*-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proc Natl Acad Sci USA* **2012**, *109* (28) 1-6.

45. Gao, Z., Wang, J., Norquay, A. K., Qiao, K., Tang, Y., Vederas, J. C. Investigation of fungal iterative polyketide synthase functions using partially assembled intermediates. *J Am Chem Soc* **2013**, *135* (5) 1735-1738.

46. Soehano, I., Yang, L., Ding, F., Sun, H., Low, Z. J., Liu, X., Liang, Z.-X. Insights into the programmed ketoreduction of partially reducing polyketide synthases: Stereo- and substrate-specificity of the ketoreductase domain. *Org Biomol Chem* **2014**, *12* (42) 8542-8549.

47. Vagstad, A. L., Newman, A. G., Storm, P. A., Belecki, K., Crawford, J. M., Townsend, C. A. Combinatorial domain swaps provide insights into the rules of fungal polyketide synthase programming and the rational synthesis of non-native aromatic products. *Angew Chem Int Ed* **2013**, *52* (6) 1718-1721.

48. Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat Prod Rep* **2012**, *29* (10) 1-24.

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

50. Chan, Y. A., Podevels, A. M., Kevany, B. M., Thomas, M. G. Biosynthesis of polyketide synthase extender units. *Nat Prod Rep* **2009**, *26* (1) 90-114.

51. Dutta, S., Whicher, J. R., Hansen, D. A., Hale, W. A., Chemler, J. A., Congdon, G. R., Narayan, A. R. H., Håkansson, K., Sherman, D. H., Smith, J. L., Skiniotis, G. Structure of a modular polyketide synthase. *Nature* **2014**, *510* (7506) 512-517.

52. 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* (7506) 560-564.

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

54. Hong, H., Leadlay, P. F., Staunton, J. The changing patterns of covalent active site occupancy during catalysis on a modular polyketide synthase multienzyme revealed by ion-trap mass spectrometry. *FEBS J* **2009**, *276* (23) 7057-7069.

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

56. Tosin, M., Betancor, L., Stephens, E., Ariel Li, W. M., Spencer, J. B., Leadlay, P. F. Synthetic chain terminators off-load intermediates from a type I polyketide synthase. *ChemBioChem* **2010**, *11* (4) 539-546.

57. Petkovic, H., Sandmann, A., Challis, I. R., Hecht, H.-J., Silakowski, B., Low, L., Beeston, N., Kuščer, E., Garcia-Bernardo, J., Leadlay, P. F., Kendrew, S. G., Wilkinson, B., Müller, R. Substrate specificity of the acyl transferase domains of EpoC from the epothilone polyketide synthase. *Org Biomol Chem* **2008**, *6* (3) 500-506.

58. Lau, J., Fu, H., Cane, D., Khosla, C. Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units. *Biochemistry* **1999**, *38* (5) 1643-1651.

59. Liou, G., Lau, J., Cane, D., Khosla, C. Quantitative analysis of loading and extender acyltransferases of modular polyketide synthases. *Biochemistry* **2003**, *42* (1) 200-207.

60. Toscano, L., Fioriello, G., Spagnoli, R., Cappelletti, L., Zanuso, G. New fluorinated erythromycins obtained by mutasynthesis. *J Antibiot* **1983**, *36* (11) 1439-1450.

61. Llano-Sotelo, B., Dunkle, J., Klepacki, D., Zhang, W., Fernandes, P., Cate, J. H. D., Mankin, A. S. Binding and action of CEM-101, a new fluoroketolide antibiotic that inhibits protein synthesis. *Antimicrob Agents Chemother* **2010**, *54* (12) 4961-4970.

62. Goss, R. J. M., Hong, H. A novel fluorinated erythromycin antibiotic. *Chem Commun* **2005**, *2005* (31) 3983-3985.

63. Goss, R. J. M., Lanceron, S., Deb Roy, A., Sprague, S., Nur-E-Alam, M., Hughes, D. L., Wilkinson, B., Moss, S. J. An expeditious route to fluorinated rapamycin analogues by utilising mutasynthesis. *ChemBioChem* **2010**, *11* (5) 698-702.

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

65. Ojima, I. Use of fluorine in the medicinal chemistry and chemical biology of bioactive compounds: A case study on fluorinated taxane anticancer agents. *ChemBioChem* **2004**, *5* (5) 628-635.

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

67. Salwiczek, M., Nyakatura, E. K., Gerling, U. I. M., Ye, S., Koksch, B. Fluorinated amino acids: Compatibility with native protein structures and effects on protein-protein interactions. *Chem Soc Rev* **2012**, *41* (6) 2135-2171.

Chapter 2: *Expanding the scope of polyketide biosynthesis using a fluorinated monomer*

Portions of this work were published in the following scientific journal:

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.

This work was performed in collaboration with the following persons:

All studies were performed in close collaboration with Mark Walker, PhD.

2.1 Introduction

The catalytic diversity of biological systems provides enormous potential for creating new chemistry to allow the scalable production of pharmaceuticals, fuels and materials using living cells [*1-4*]. However, the scope of innovation of living organisms is typically limited to functions that confer a direct advantage for cell growth, maximizing biomass as the end product rather than a molecule or reaction of interest. In contrast, synthetic biology approaches allow us to disconnect some of these remarkable biochemical transformations from cell survival and reconnect them in new ways for the targeted synthesis of new classes of compounds. One particularly interesting area is the development of methods to introduce fluorine into complex small molecule scaffolds, which has become a powerful strategy for the design of synthetic pharmaceuticals: 20-30% of drugs, including many of the top sellers, contain at least one fluorine atom [*5-7*]. Recent innovations have expanded the scope of synthetic C–F bond forming methodologies, but the unusual elemental properties of fluorine that serve as the basis for its success also continue to restrict the range of molecular structures that can be accessed [*8-11*]. As such, the invention of new routes for the site-selective introduction of fluorine into structurally diverse molecules, particularly under mild conditions, remains an outstanding challenge.

Despite its effectiveness as a tool in human hands, fluorine has limited distribution in naturally occurring molecules. In fact, the only fluorinated natural products characterized to date consist of a small set of simple organofluorines related to the fluoroacetate pathway of

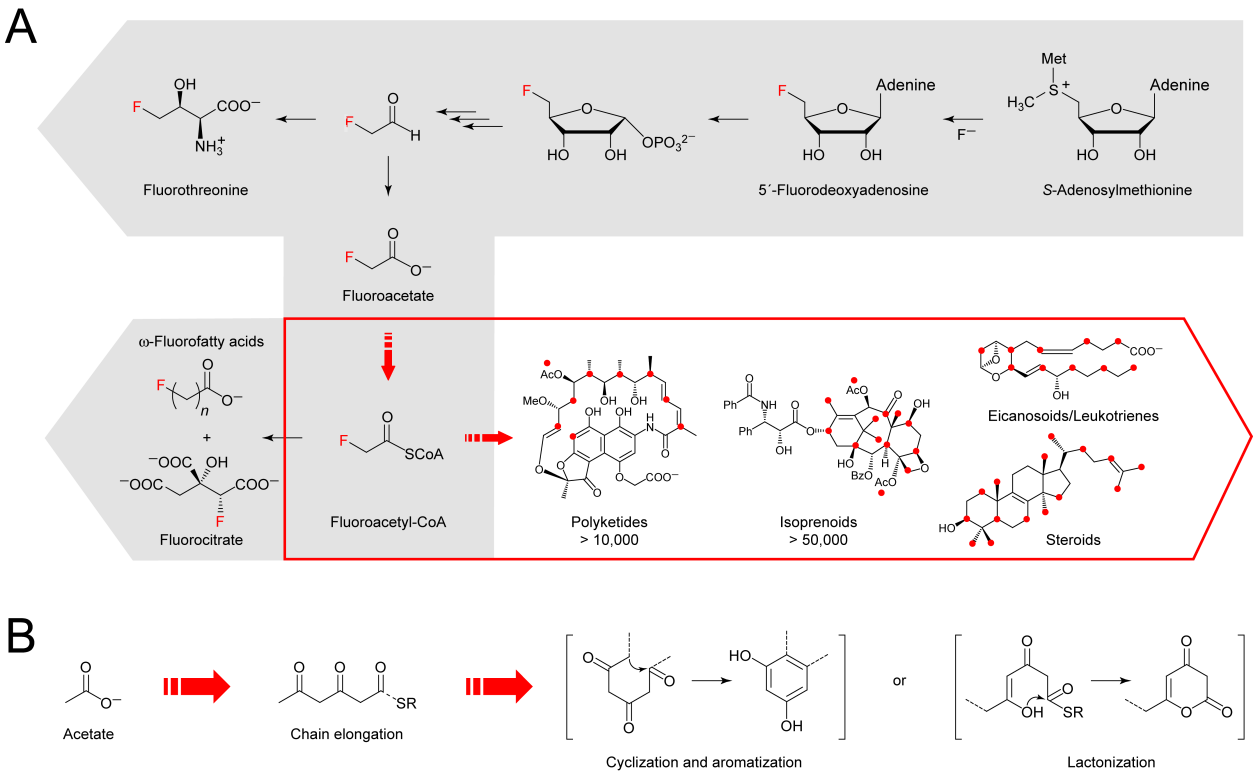


Figure 2.1. *Synthetic biology of fluorine. (A) The fluoroacetate pathway and its metabolites represent the known scope of biological fluorine chemistry, producing fluoroacetate and fluorothreonine as the end products (grey box). This scope could be greatly expanded by engineering downstream pathways and using fluoroacetate as an organofluorine building block for introducing fluorine site-selectively into large families of modularly-synthesized natural products constructed from acetate backbones (red box). A red dot represents a position that can be fluorinated with the incorporation of a fluoroacetate monomer without altering the carbon skeleton, including locations where fluorine would replace a methyl group derived from propionate or where downstream tailoring steps have occurred on the final structure. (B) Assembly of acetate units in the biosynthesis of polyketide natural products.*

Streptomyces cattleya, a soil bacterium that houses the remarkable ability to catalyze the formation of C–F bonds from aqueous fluoride (*Figure 2.1A*) [*12, 13*]. Although these compounds lack the intricacy typically expected of secondary metabolites, they represent a potentially rich source of organofluorine building blocks for the production of complex fluorinated natural products. The backbones of isoprenoids, steroids, alkaloids, eicosanoids, leukotrienes, and others are biosynthesized directly from the assembly and tailoring of simple acetate units (*Figure 2.1A*). Introduction of the fluoroacetate monomer in place of acetate would allow us to site-specifically incorporate fluorine into the backbone of these targets, combining the medicinal chemistry advantages of fluorine with the structural complexity and bioactivity of natural products to create new molecular function. In this Chapter, we show that the enzymatic synthesis of fluoromalonyl-CoA, a fluorinated analog of one of nature's most powerful carbon nucleophiles, can be engineered from fluoroacetate and harnessed as an extender unit for the production of fluorinated polyketides *in vitro* and *in vivo*.

Many acetate-based natural products, polyketides in particular, are generated through the iterative condensation of activated thioesters, resulting in reactive β-keto units that condense further to produce a wide range of structures (*Figure 2.1B*) [*14, 15*]. The structural diversity of these compounds is especially striking given that the vast majority of polyketides draw on only two monomers, acetate and propionate, as the extender units that form their carbon skeletons [*3, 14, 16*]. Although polyketide synthases have been observed to be promiscuous with regard to their starter units [*3, 14*], the encoding of extender units has been found to be quite selective and the large pool of substituted acyl-CoAs inside the cell are excluded from the backbone [*14, 16*]. Thus, while the specificity between the acetate and propionate extender units can be swapped using various approaches [*3, 14*], engineered extenders units have not yet been employed, except in cases where unusual extenders are naturally incorporated by a polyketide synthase [*17-19*]. In fact, fluoroacetate can be used as a starter unit to produce highly toxic ω-fluoro fatty acids (*Figure 2.1A*) [*13*], but fluorine has never been observed to date within the backbone, implying that chain extension reactions with the fluorinated acyl-CoA do not occur in these systems. The apparent inability of living systems to utilize fluorine for the biosynthesis of complex small molecules could arise from several sources, but is likely derived in part from the extreme properties of fluorine that affect biological as well as chemical synthesis. For example, the pK_a of the α-proton, electrophilicity of the carbonyl group, and the stability of the acyl-CoA and carbanion are all highly affected by fluorine substitution. Furthermore, the fluoroacetyl group bears a clear similarity to the fluoromethyl ketone motif used for the design of covalent enzyme inhibitors, which can lead to the irreversible alkylation of active-site nucleophiles [*20*]. Thus, the development of a system to incorporate novel fluorinated extenders could dramatically increase the range of complex structures that can be accessed but must also address the challenges involved in activating the fluoroacetate monomer for the C–C bond forming chemistry of chain extension reactions.

2.2 Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), yeast extract, malt extract, glycerol, and triethylamine (TEA) were purchased from EMD Biosciences (Darmstadt, Germany). Carbenicillin (Cb), isopropyl-β-D-thiogalactopyranoside (IPTG), phenylmethanesulfonyl fluoride (PMSF), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride, dithiothreitol (DTT), 4-(2-hydroxyethyl)-1-

piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA), dichloromethane, ethyl acetate and ethylene diamine tetraacetic acid disodium dihydrate (EDTA), were purchased from Fisher Scientific (Pittsburgh, PA). Sodium fluoroacetate, coenzyme A trilithium salt (CoA), acetyl-CoA, malonyl-CoA, methylmalonyl-CoA, diethylfluoromalonate, malonic acid, methylmalonic acid, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, lithium hexamethyldisilazide solution (LiHMDS), phosphoenolpyruvate (PEP), adenosine triphosphate sodium salt (ATP), nicotinamide adenine dinucleotide reduced form dipotassium salt (NADH), nicotinamide adenine dinucleotide phosphate reduced form (NADPH), myokinase, pyruvate kinase, lactate dehydrogenase, poly(ethyleneimine) solution (PEI), 5-fluorouracil, β-mercaptoethanol, sodium phosphate dibasic hepthydrate, chlorotrifluoromethane and N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED) were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate and ammonium persulfate were purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, T4 DNA ligase, Antarctic phosphatase, Phusion DNA polymerase, T5 exonuclease, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) and Platinum Taq High-Fidelity polymerase (Pt Taq HF) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in 10 mM Tris-HCl, pH 8.5, and stored at either 4°C for immediate use or -20°C for longer term use. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). O-(7-Azabenzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorophosphate (HATU), Amicon Ultra 3,000 MWCO and 30,000 MWCO centrifugal concentrators, 5,000 MWCO regenerated cellulose ultrafiltration membranes, and LiChroCART 250-4 Purospher RP-18e HPLC column were purchased from EMD Millipore (Billerica, MA). Deuterium oxide and chloroform-d were purchased from Cambridge Isotope Laboratories (Andover, MA). ^{19}F NMR spectra were collected at 25ºC on Bruker AVQ-400 or AV-600 spectrometers at the College of Chemistry NMR Facility at the University of California, Berkeley or on a Bruker Biospin 900 MHz spectrometer at the QB3 Central California 900 MHz NMR Facility or on a Bruker AV-600 spectrometer equipped with a QCI-CryoProbe at Novartis Institutes for Biomedical Research (Emeryville, California). Spectra were referenced to $CFCl_3$ (0 ppm) or 5-fluorouracil (D_2O: -168.33 ppm vs. $CFCl_3$). NMR assignments were made based on COSY, ^{13}C-^{1}H HSQC, ^{13}C-^{1}H HMBC and ^{19}F-^{1}H HMBC spectra where appropriate. High-resolution mass spectral analyses were carried out at the College of Chemistry Mass Spectrometry Facility.

Bacterial strains. *E. coli* DH10B-T1^{R} and BL21(de3)T1^{R} were used for DNA construction and heterologous protein production, respectively, except for DEBS modules, which were heterologously expressed in *E. coli* BAP1 [*21*].

Gene and plasmid construction. Standard molecular biology techniques were used to carry out plasmid construction. All PCR amplifications were carried out with Phusion or Platinum Taq High Fidelity DNA polymerases. For amplification of GC-rich sequences from *S. coelicolor*, PCR reactions were supplemented with DMSO (5%) using the standard buffer rather than GC buffer with primer annealing temperatures 8-10°C below the T_m. All constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA).

The synthetic gene encoding NphT7 was optimized for *E. coli* class II codon usage and synthesized using PCR assembly (*Appendix Table 1*). Gene2Oligo was used to convert the gene sequence into primer sets using default optimization settings (*Appendix Table 1*) [*22*]. To assemble the synthetic gene, each primer was added at a final concentration of 1 μM to the first PCR reaction (50 μL) containing 1 × Pt Taq HF buffer (20 mM Tris-HCl, 50 mM KCl, pH 8.4), $MgSO_4$ (1.5 mM), dNTPs (250 μM each), and Pt Taq HF (5 U). The following thermocycler program was used for the first assembly reaction: 95°C for 5 min; 95°C for 30 s; 55°C for 2 min; 72°C for 10 s; 40 cycles of 95°C for 15 s, 55°C for 30 s, 72°C for 20 s plus 3 s/cycle; these cycles were followed by a final incubation at 72°C for 5 min. The second assembly reaction (50 μL) contained 16 μL of the unpurified first PCR reaction with standard reagents for Pt Taq HF. The thermocycler program for the second PCR was: 95°C for 30 s; 55°C for 2 min; 72°C for 10 s; 40 cycles of 95°C for 15 s, 55°C for 30 s, 72°C for 80 s; these cycles were followed by a final incubation at 72°C for 5 min. The second PCR reaction (16 μL) was transferred again into fresh reagents and run using the same program. Following gene construction, the DNA smear at the appropriate size was gel purified and used as a template for the rescue PCR (50 μL) with Pt Taq HF and rescue primers under standard conditions. The resulting rescue product was inserted into pBAD33 and confirmed by sequencing, then amplified using the nphT7 F1/R1 primer set (*Appendix Table 1*) and inserted into the NdeI site of pET-16b using the Gibson protocol [*23*].

pET16b-His_{10}-AckA.EC and pET16b-His_{10}-Pta.EC were constructed by amplification from pRSFDuet-ackA.pta using the AckA.EC F/R and Pta.EC F/R primer sets (*Appendix Table 1*) and insertion into the NdeI-XhoI (Pta, Gene ID 12872491) or NdeI-BamHI (AckA, Gene ID 12874027) sites of pET16b. pET28a-His_6-MatB.SCo and pET28a-His_6-Epi.SCo were constructed by amplification from S. coelicolor A3(2) M145 (ATCC BAA-471) genomic DNA using the MatB.SCo F/R and epi.SCo F/R primer sets (*Appendix Table 1*) and insertion into the NdeI-XhoI sites of pET28a. pET16x-His_{10}-THNS was constructed by amplification out of S. coelicolor genomic DNA using the THNS F/R primer set (*Appendix Table 1*) and insertion into the NdeI-SpeI sites of pET16x. pCDFDuet-DszsAT.SCe-MatB.SCo and pCDFDuet-ø-MatB.SCo were constructed by amplification from pET28a-His_6-MatB.SCo and pFW3 (40) using the pCDF-MatB.SCo F/R and pCDF-DszsAT.SCe F/R primer sets (*Appendix Table 1*) and insertion of DszsAT.SCe and MatB.SCo into the NcoI-HindIII and NdeI-KpnI sites of pCDFDuet-1 respectively. pTRC33-NphT7-PhaB was constructed by amplifying NphT7 from pET-16b-NphT7 using the NphT7 G F/G R primer set (*Appendix Table 1*) and PhaB from pBT33-PhaABC (48) using the PhaB F/R primer set (*Appendix Table 1*) where each forward primer included the RBS from pET16b, then inserting both genes simultaneously into the BamHI-XbaI sites of pTRC33 using the Gibson protocol. pSV272-His_6-MBP-DEBSMod2 and pSV272-His_6-MBP-DEBSMod2/AT^0 (S2652A based on EryAI numbering) were constructed by amplification from pBP19 using the MBP-M2 F/R primer set (pSV272-MBP-DEBSMod2-His_6) (*Appendix Table 1*) or MBP-M2 F/MBP-M2ATnull R and MBP-M2ATnull F/MBP-M2 R (pSV272-MBP-DEBSMod2/AT^0-His_6) (*Appendix Table 1*) and insertion into the SfoI-HindIII sites of pSV272.1 using the Gibson protocol. pBAD33.BirA.EC was cloned by the QB3 Macrolab.

Expression of His-tagged proteins. TB (1 L) containing carbenicillin, kanamycin, and chloramphenicol (50 μg/mL) as appropriate in a 2.8 L Fernbach baffled shake flask was inoculated to OD_{600} = 0.05 with an overnight TB culture of freshly transformed *E. coli* containing the appropriate overexpression plasmid. The cultures were grown at 37°C at 250 rpm to OD_{600} = 0.6 to 0.8 at which point cultures were cooled on ice for 20 min, followed by induction of protein expression with IPTG (His_{10}-AckA, His_{10}-Pta: 1 mM; His_{10}-AccA/B/C/D

[*26*], $DEBS_{Mod6}/AT^0$+TE-His_6 (pAYC138 [*24*]): 0.4 mM; His_{10}-MatB, His_{10}-Epi, His_{10}-NphT7, His_6-DszAT (pFW3 [*24*]), His_6-$DEBS_{Mod6}$+TE-His_6 (pRSG54 [*27*]), $DEBS_{Mod3}$+TE-His_6 (pRSG34 [*28*]), $DEBS_{Mod3}/AT^0$+TE-His_6 (pAYC136 [*24*]), MBP-$DEBS_{Mod2}$-His_6, MBP-$DEBS_{Mod2}/AT^0$-His_6, DEBS1 : 0.2 mM) and overnight growth at 16°C. For His_{10}-AccB expression, pBAD33-BirA was co-expressed (L-arabinose, 0.2 %) and the medium was supplemented with 20 nM D-(+)-biotin at induction. For MBP-$DEBS_{Mod2}$-His_6 and MBP-$DEBS_{Mod2}/AT^0$-His_6 pRARE2 was co-expressed. Cell pellets were harvested by centrifugation at 9,800 × *g* for 7 min at 4ºC and stored at -80ºC.

Purification of His_{10}-AckA, His_{10}-Pta, His_{10}-MatB, His_{10}-Epi, His_{10}-AccA/B/C/D, His_6-DszAT and His_{10}-NphT7. Frozen cell pellets were thawed and resuspended at 5 mL/g cell paste with Buffer A (50 mM sodium phosphate, 300 mM sodium chloride, 20% glycerol, 20 mM BME, pH 7.5) containing imidazole (10 mM) for His_{10}-AckA, His_{10}-Pta, His_{10}-AccA/B/C/D, and His_{10}-NphT7 or Buffer B (200 mM sodium phosphate, 200 mM sodium chloride, 30% glycerol, 2.5 mM EDTA, 2.5 mM DTT, pH 7.5) for His_{10}-MatB and His_{10}-Epi. Complete EDTA-free protease inhibitor cocktail (Roche) was added to the lysis buffer before resuspension. The cell paste was homogenized before lysis by passage through a French Pressure cell (Thermo Scientific; Waltham, MA) at 14,000 psi. The lysate was centrifuged at 15,300 × *g* for 20 min at 4°C to separate the soluble and insoluble fractions. DNA was precipitated in the soluble fraction by addition of 0.15% (w/v) poly(ethyleneimine). The precipitated DNA was removed by centrifugation at 15,300 × *g* for 20 min at 4°C. The remaining soluble lysate was diluted three-fold with Buffer A containing imidazole (10 mM) and loaded onto a Ni-NTA agarose column (Qiagen, 1 mL resin/g cell paste) by gravity flow or on an ÄKTApurifier FPLC (2 mL/min; GE Healthcare; Piscataway, NJ). The column was washed with Buffer A until the eluate reached an $A_{280\ nm} < 0.05$ or was negative for protein content by Bradford assay (Bio-Rad).

His_{10}-AckA, His_{10}-Pta, His_{10}-MatB, His_{10}-AccB and His_{10}-AccD. The column was washed with 5 to 10 column volumes with Buffer A supplemented with imidazole (His_{10}-AckA, 40 mM; His_{10}-Pta, 35 mM; His_{10}-MatB, His_{10}-AccB, and His_{10}-AccD, 20 mM). The protein was then eluted with 300 mM imidazole in Buffer A.

His_{10}-Epi. His_{10}-Epi was eluted using a linear gradient from 0 to 300 mM imidazole in Buffer A over 30 column volumes.

His_{10}-AccA, His_{10}-AccC and His_{10}-NphT7. The column was washed with a linear gradient from 10 to 90 mM imidazole in Buffer A over 15 column volumes and then eluted with 300 mM imidazole in Buffer A.

Fractions containing the target protein were pooled by $A_{280\ nm}$ and concentrated using either an Amicon Ultra spin concentrator (3 kDa MWCO, Millipore) or an Amicon ultrafiltration cell under nitrogen flow (65 psi) using a membrane with an appropriate nominal molecular weight cutoff (Ultracel-5 or YM10, Millipore). Protein was then exchanged into Buffer C (50 mM HEPES, 100 mM sodium chloride, 2.5 mM EDTA, 20% glycerol, pH 7.5) with (His_{10}-AckA, His_{10}-Pta, His_{10}-AccA/B/C/D, His_{10}-NphT7) or without (His_{10}-MatB and His_{10}-Epi) DTT (0.5-1 mM) using a Sephadex G-25 column (Sigma-Aldrich, bead size 50-150 μm, 10 mL resin/mL protein solution), then concentrated again before storage.

Final protein concentrations before storage were estimated using the $\varepsilon_{280\ nm}$ calculated by ExPASY ProtParam as follows: His_{10}-AckA: 14.8 mg/mL ($\varepsilon_{280\ nm}$ = 24,860 $M^{-1}\ cm^{-1}$), His_{10}-Pta: 16.5 mg/mL ($\varepsilon_{280\ nm}$ = 37,360 $M^{-1}\ cm^{-1}$), His_{10}-MatB: 19.8 mg/mL ($\varepsilon_{280\ nm}$ = 33,920 $M^{-1}\ cm^{-1}$),

His_{10}-Epi: 18.5 mg/mL ($\varepsilon_{280\ nm}$ = 11,460 $M^{-1}\ cm^{-1}$), His_{10}-AccA: 33.2 mg/mL ($\varepsilon_{280\ nm}$ = 25,900 $M^{-1}\ cm^{-1}$), His_{10}-AccB: 23.0 mg/mL ($\varepsilon_{280\ nm}$ = 2,980 $M^{-1}\ cm^{-1}$), His_{10}-AccC: 32.6 mg/mL ($\varepsilon_{280\ nm}$ = 27,850 $M^{-1}\ cm^{-1}$), His_{10}-AccD: 4.5 mg/mL ($\varepsilon_{280\ nm}$ = 16,960 $M^{-1}\ cm^{-1}$), His_6-DszAT: 1.7 mg/mL ($\varepsilon_{280\ nm}$ = 17,420 $M^{-1}\ cm^{-1}$), His_{10}-NphT7: 0.4 mg/mL ($\varepsilon_{280\ nm}$ = 26,930 $M^{-1}\ cm^{-1}$). All proteins were aliquoted, flash-frozen in liquid nitrogen, and stored at -80°C.

Purification of $DEBS_{Mod6}$+TE-His_6, $DEBS_{Mod6}/AT^0$+TE-His_6, $DEBS_{Mod3}$+TE-His_6, and $DEBS_{Mod3}/AT^0$+TE-His_6. The His-tagged DEBS module with thioesterase ($DEBS_{Mod6}$+TE) construct was heterologously expressed in *E. coli* BAP1 pRSG54 as previously described and purified using a modified literature protocol [*29*]. Cleared cell lysates were prepared in Buffer B as described above, diluted three-fold with Buffer A, and passed over a Ni-NTA agarose column (Qiagen, approximately 1 mL/g cell paste) on an ÄKTApurifier FPLC. The column was washed with Buffer A until the eluate reached an $A_{280\ nm} < 0.05$. Protein was eluted with Buffer D (50 mM sodium phosphate, 50 mM sodium chloride, 20 mM BME, 20% glycerol, 100 mM imidazole, pH 7.5). The eluate was diluted two-fold with Buffer E (50 mM HEPES, 2.5 mM EDTA, 2.5 mM DTT, 20% glycerol, pH 7.5), loaded onto a HiTrap Q HP column (GE Healthcare, 5 mL), and eluted with a linear gradient from 0 to 1 M sodium chloride in Buffer E over 30 column volumes (4.5 mL/min). Fractions containing the target protein (eluted at ~350 mM sodium chloride) were pooled by $A_{280\ nm}$ and concentrated under nitrogen flow (65 psi) in an Amicon ultrafiltration cell using a YM10 membrane. The protein was flash-frozen in liquid nitrogen and stored at -80°C at a final concentration of 6 - 30 mg/mL, which was estimated using the calculated $\varepsilon_{280\ nm}$ ($DEBS_{Mod3}$+TE and $DEBS_{Mod3}/AT^0$+TE: 203,280 $M^{-1}cm^{-1}$; $DEBS_{Mod6}$+TE and $DEBS_{Mod6}/AT^0$+TE: 206,260 $M^{-1}\ cm^{-1}$).

Purification of MBP-$DEBS_{Mod2}$-His_6 and MBP-$DEBS_{Mod2}/AT^0$-His_6. Cleared lysates were produced as for other DEBS modules, diluted 2-fold with Buffer A + 10 mM imidazole, and bound in batch to Ni-NTA resin (2.5 mL resin per g cell paste) for 2h. The slurry was poured into a fritted column and washed with Buffer A + 10 mM imidazole until the $A_{280\ nm}$ of the eluent dropped below 0.05. The protein was then eluted with Buffer A + 300 mM imidazole. The protein was then concentrated to approximately 1 mg/mL in an Amicon ultrafiltration cell. The protein was then dialyzed overnight against Buffer E + 50 mM NaCl with 1 mg TEV per 100 mg protein to remove the MBP tag. The protein was then loaded onto a HiTrap Q HP column and eluted by a linear gradient from Buffer E to Buffer E + 500 mM NaCl over 20 column volumes. Fractions containing the desired protein were identified by SDS-PAGE (eluting at approximately 350 mM NaCl), pooled and concentrated in an Amicon Ultra spin concentrator to 20 - 25 mg/mL base on the calculated $\varepsilon_{280\ nm}$ of 158,360 $M^{-1}cm^{-1}$. Aliquots were flash frozen in liquid nitrogen and stored at -80°C.

ESI-MS screening method for acyl-CoAs. Preparative HPLC fractions were screened on an Agilent 1290 HPLC system using a Zorbax Eclipse Plus C-18 column (3.5 µm, 2.1 × 30 mm, Agilent) with a linear gradient from 0 to 65% acetonitrile over 2 min with 0.1% formic acid as the aqueous mobile phase (0.75 mL/min). Mass spectra were recorded on an Agilent 6130 single quadrupole MS with ESI source, operating in negative and positive ion scan mode.

Fluoromalonate. Diethylfluoromalonate (0.5 mL, 3.2 mmol) was saponified with methanolic sodium hydroxide (2 M, 3.5 mL) in dichloromethane and methanol (9:1 v/v, 32 mL) and the sodium salt isolated by filtration through a Büchner funnel with a fine porosity glass frit [*30*]. ^{19}F NMR (565 MHz, D_2O, 5-fluorouracil = -168.3 ppm): δ -176.43 (d, J=53 Hz).

Fluoromalonyl-CoA. Fluoromalonyl-CoA was prepared enzymatically from fluoromalonate and CoA using MatB and ATP. A myokinase/pyruvate kinase/PEP system was also used to regenerate ATP in order to avoid high concentrations of AMP that might inhibit MatB. The reaction mixture (10 mL) contained 100 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (5 mM), TCEP (2.5 mM), magnesium chloride (5 mM), fluoromalonate (2.5 mM), ATP (2.5 mM), pyruvate kinase (36 U), myokinase (20 U), CoA (2 mM) and MatB (5 μM). The mixture was incubated at 37°C for 6 h and then at room temperature for 16 h before lyophilizing overnight. The residue was dissolved in water (1.6 mL) and acidified to pH ~2 by addition of 70% perchloric acid (160 μL). Insoluble material was removed by centrifugation at 18,000 × *g* for 10 min. The supernatant was adjusted to pH 6 by addition of 10 M sodium hydroxide (100 μL) and desalted on an Agilent 1200 HPLC system using a Zorbax Eclipse XDB C-18 column (5 μm, 9.4 × 250 mm, Agilent) with a linear gradient from 0 to 10% methanol over 9 min with 50 mM sodium phosphate, 25 mM trifluoroacetic acid, pH 4.5 as the aqueous mobile phase (3 mL/min). Fractions eluting near the void volume, containing both fluoromalonyl-CoA and CoA, were lyophilized overnight, dissolved in water (1 mL), and purified using a Zorbax Eclipse XDB C-18 column (5 μm, 9.4 × 250 mm) with a linear gradient from 0 to 50% methanol over 45 min with 50 mM sodium phosphate, pH 4.5 as the aqueous mobile phase (3 mL/min). Fractions were screened by ESI-MS and those containing pure fluoromalonyl-CoA were lyophilized overnight, dissolved in water (1 mL), and desalted using a Zorbax Eclipse XDB C-18 column (5 μm, 9.4 × 250 mm) with a linear gradient from 0 to 15% acetonitrile over 30 min with water as the mobile phase (3 mL/min). The desalted fluoromalonyl-CoA was lyophilized and redissolved in water or D_2O. The fluoromalonyl-CoA solutions were stored at -20°C but are stable for at least 24 h at room temperature. During NMR measurements in D_2O, complete H-D exchange occurred at the fluorine-substituted carbon over the course of 48 h. ^{1}H NMR (600 MHz, D_2O, MeOH = 3.34 ppm): δ 8.55 (s, 1H, H_8), 8.27 (s, 1H, H_2), 6.17 (d, J=6.6 Hz, 1H, $H_{1'}$), 5.23 (d, J=50.3 Hz, 1H, O_2C-C<u>H</u>F-C=O), 4.89 – 4.79 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.59 (m, 1H, $H_{4'}$), 4.23 (m, 2H, $H_{5'}$), 4.00 (s, 1H, $H_{3''}$), 3.82 (dd, J=10.2, 4.6 Hz, 1H, pro-R-$H_{1''}$), 3.54 (dd, J=10.0, 4.4 Hz, 1H, pro-S- $H_{1''}$), 3.49 – 3.39 (m, 2H, $H_{5''}$), 3.38 – 3.29 (m, 2H, $H_{8''}$), 3.11-3.02 (m, 2H, $H_{9''}$), 2.42 (t, J=6.7 Hz, 3H, $H_{6''}$), 0.88 (s, 3H, $H_{10''}$), 0.74 (s, 3H, $H_{11''}$). ^{13}C NMR (226 MHz, D_2O, CH_3OH = 49.15 ppm): δ 196.92, 196.85 (d, J=27.5 Hz, <u>C</u>O_2), 175.02 ($C_{4''}$), 174.32 ($C_{7''}$), 169.52 (d, J=21.0 Hz, O_2C-CDF-<u>C</u>=O), 155.37 (C_6), 152.40 (C_2), 149.57 (C_4), 140.37 (C_8), 118.91 (C_5), 93.32 (td, J=25 Hz, 197 Hz, O_2C-<u>C</u>DF-C=O), 86.61 ($C_{1'}$), 83.81 (d, J=9 Hz, $C_{4'}$), 74.56 (d, J=5 Hz, $C_{3'}$ or $C_{1''}$), 74.35 ($C_{3''}$), 74.00 (d, J=5 Hz, $C_{2'}$), 72.16 (d, J=6 Hz, $C_{3'}$ or $C_{1''}$), 65.56 ($C_{5'}$), 38.66 ($C_{5''}$ or $C_{6''}$), 38.58 (d, J=8 Hz, $C_{8''}$) 35.58 (d, J=30 Hz, $C_{9''}$), 27.60 (d, J=3 Hz, $C_{5''}$ or $C_{6''}$), 21.16 ($C_{10''}$), 18.22 ($C_{11''}$). ^{19}F NMR (565 MHz, D_2O, CF_3CO_2H = -76.20 ppm): δ -182.11 (dd, J=7.8, 50.4 Hz, O_2C-CH<u>F</u>-C=O), -182.72 (m, O_2C-CD<u>F</u>-C=O). HR-ESI-MS $[M-H]^-$: calculated for $C_{24}H_{36}FN_7O_{19}P_3S$, *m/z*, 870.0989, found *m/z* 870.0991.

Enzyme assays. Kinetic parameters (k_{cat}, K_M) were determined by fitting the data using Microcal Origin to the equation: $v_o = v_{max}[S] / (K_M + [S])$, where v is the initial rate and [S] is the substrate concentration. Data are reported as mean ± s.e. (n = 3) unless otherwise noted with standard error derived from the nonlinear curve fitting. Error bars on graphs represent mean ± s.d. (n = 3). Error in k_{cat}/K_M is calculated by propagation of error from the individual kinetic parameters.

Acetyl-CoA carboxylase. ACCase activity was measured using a discontinuous HPLC assay. Assays were performed at 30°C in a total volume of 200 μL containing 50 mM HEPES, pH 7.5, TCEP (10 mM), bovine serum albumin (3 mg/mL), CoA (0.5 mM), ATP (2.5 mM), magnesium

chloride (10 mM), sodium bicarbonate (75 mM), acetate or fluoroacetate (10 mM), phosphoenolpyruvate (10 mM), pyruvate kinase (4 U), AckA (0.1 μM), Pta (10 μM) and ACCase (15 μM). The pH of the buffer remained unchanged after addition of sodium bicarbonate. ACCase stock solution was prepared by pre-mixing the protein subunits at equimolar ratio (85 μM) except for AccB, which was added at 1.5-fold molar excess. The reaction was initiated with addition of ATP. Aliquots (20 μL) were removed and quenched by the addition of 70% perchloric acid (1 μL). Insoluble material was removed by centrifugation and the supernatant was analyzed on an Agilent 1200 or 1290 HPLC system on a LiChroCART 250-4 Purospher RP-18e column (5 μm, 4.6 × 250 mm, Millipore) and monitored at $A_{260\ nm}$. For reactions with acetate, a linear gradient from 2 to 20% acetonitrile over 10 min with 5 mM sodium phosphate and 5 mM sodium citrate with 0.1% trifluoroacetic acid, pH 4.6 as the aqueous mobile phase (1 mL/min) was used to analyze the reaction. For reactions with fluoroacetate, a linear gradient from 2 to 15% acetonitrile containing 0.1% TEA over 15 min with 10 mM Tris, pH 8.0 containing 0.1% TEA as the aqueous mobile phase (1 mL/min) was used. Buffers containing TEA were made fresh daily and could be used for at least 6 h before significant change in chromatography was observed.

Malonyl-CoA synthetase. MatB activity was measured using a modified literature method [*31*]. The production of AMP was coupled to pyruvate formation by myokinase and pyruvate kinase, which in turn was coupled to NADH oxidation by lactate dehydrogenase. Assays were performed at 30°C in a total volume of 200 μL containing 100 mM HEPES, pH 7.5, TCEP (1 mM), ATP (2.5 mM), magnesium chloride (5 mM), phosphoenolpyruvate (1 mM), NADH (0.3 mM), myokinase (0.5 U), pyruvate kinase (3.6 U), lactate dehydrogenase (2.6 U), dicarboxylic acid (25 μM – 1 mM malonate, 50 μM – 10 mM fluoromalonate or 25 μM – 1.5 mM methylmalonate) and MatB (26 nM for malonate, 1 μM for fluoromalonate and 200 nM for methylmalonate). The reaction was initiated with addition of CoA (0.5 mM) and monitored at 340 nm in a Beckman Coulter DU-800 spectrophotometer.

Acetoacetyl-CoA synthase. NphT7 activity was measured using a NADPH-coupled assay with PhaB. Assays were performed at 30°C in a total volume of 500 μL containing 50 mM HEPES, pH 7.5, NADPH (160 μM), acetyl-CoA (200 μM), PhaB (0.05 mg/mL), NphT7 (0.2 μM for malonyl-CoA; 0.5μM for fluoromalonyl-CoA) and malonyl-CoA (5 – 150 μM) or fluoromalonyl-CoA (5 – 200 μM). Reactions were initiated with the addition of malonyl- or fluoromalonyl-CoA and monitored at 340 nm in an Agilent 8453 diode array spectrophotometer. The PhaB-coupled assay was tested both by doubling NphT7, which doubled the initial velocity with both the fluorinated and non-fluorinated substrates, and also by doubling the amount of PhaB, which led to no difference in initial velocity.

Acyl-CoA hydrolysis by DEBS. Hydrolytic activity of $DEBS_{Mod6}$+TE was measured by monitoring the reaction of free CoA with DTNB as described previously [*32*]. Assays were performed at 37°C in a total volume of 200 μL containing 400 mM sodium phosphate, pH 7.5, 500 μM DTNB, and $DEBS_{Mod6}$+TE (1 μM). Reactions were initiated by addition of acyl-CoA (0.5 mM) and monitored at 412 nm in a Beckman Coulter DU-800 spectrophotometer. Release of CoA was quantified by comparison to a standard curve (5–100 μM).

2-fluoro-3-hydroxybutyryl-CoA production using NphT7. As acetofluoroacetyl-CoA proved to degrade fairly rapidly under the assay conditions, 2-fluoro-3-hydroxybutyryl-CoA was isolated from a 10 mL reaction containing 100 mM HEPES, pH 7.5, fluoromalonate (10 mM), CoA (500 μM), NADPH (1 mM), ATP (1 mM), magnesium chloride (5 mM),

phosphoenolpyruvate (10 mM), pyruvate kinase (180 U), myokinase (100 U), MatB (40 μM), PhaB (7 μM) and NphT7 (2 μM) that was initiated by the addition of acetyl-CoA (0.5 mM, limiting reagent). The reaction was incubated at 30°C overnight followed by quenching by the addition of 70% perchloric acid (50 μL). 2-fluoro-3-hydroxybutryl-CoA was purified using a Zorbax Eclipse XDB C-8 column (5 μm, 9.4 × 250 mm, Agilent) with a linear gradient from 0 to 5% acetonitrile over 30 min (3 mL/min) with 50 mM sodium phosphate with 0.1% trifluoroacetic acid (pH 4.5) as the aqueous mobile phase. Fractions containing 2-fluoro-3-hydroxybutryl-CoA were identified by ESI-MS and lyophilized. The remaining solid was dissolved in water (1 mL) and purified a second time a Zorbax Eclipse XDB C-8 column (5 μm, 9.4 × 250 mm, Agilent) with a linear gradient from 0 to 5% acetonitrile over 30 min (3 mL/min) with 0.1% formic acid as the aqueous mobile phase. Fractions containing 2-fluoro-3-hydroxybutryl-CoA were identified by ESI-MS and lyophilized. Two diastereomers were observed by NMR in an approximately 2.5:1 ratio. ^{1}H NMR (600 MHz, D_2O, acetonitrile = 2.06 ppm): δ 8.54 (s, 1H, H_8), 8.30 (s, 1H, H_2), 6.08 (d, J=5.9 Hz, 1H, $H_{1'}$), 4.93 (dd, J=47.8, 2.6 Hz, 0.2H, HOCH-C<u>H</u>F-C=O minor diastereomer), 4.85 (dd, J=46.9, 2.2 Hz, 0.8H, HOCH-C<u>H</u>F-C=O major diastereomer), 4.77 – 4.77 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.46 (m, 1H, $H_{4'}$), 4.16-4.04 (m, 3H, HOC<u>H</u>-CHF-C=O, $H_{5'}$), 3.89 (s, 1H, $H_{3''}$), 3.72 (d, J=7.1 Hz, 1H, $H_{1''}$ *pro-R*), 3.46 (d, J=9.7 Hz, 1H, pro-S- $H_{1''}$ *pro-S*), 3.31 (t, J=6.5 Hz, 2H, $H_{5''}$), 3.26 – 3.20 (m, 2H, $H_{8''}$), 2.95 (m, 2H, $H_{9''}$), 2.30 (t, J=6.5 Hz, 2H, $H_{6''}$), 1.15 (d, J=6.7 Hz, 2H, <u>H</u>$_3$C-HOCH-CHF major diastereomer), 1.05 (d, J=6.4 Hz, 1H, <u>H</u>$_3$C-HOCH-CHF minor diastereomer), 0.79 (s, 3H, $H_{10''}$), 0.67 (s, 3H, $H_{11''}$). ^{19}F NMR (565 MHz, D_2O, CF_3CO_2H = -76.20 ppm): δ -198.62 (dd, J=48.24, 23.3 Hz, HOCH-CH<u>F</u>-C=O), -206.85 (dd, J=46.9, 27.5, HOCH-CH<u>F</u>-C=O). ESI-MS $[M+H]^+$: calculated for $C_{25}H_{42}FN_7O_{18}P_3S$, *m/z*, 872.2, found *m/z* 872.0.

Triketide lactone production using DEBS$_{Mod6}$+TE and MatB. Assay and preparative mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase (27 U/mL), myokinase (10 U/mL), CoA (0.5 mM), methylmalonyl-CoA epimerase (5 μM), MatB (40 μM or as specified) and fluoro- or methylmalonate (10 – 20 mM). Including NADPH resulted in only trace yields of reduced triketide product, even with the native methylmalonyl-CoA extender, so the cofactor was omitted. This mixture was incubated at 37°C for 30-45 min and then initiated by addition of the N-acetylcysteamine thioester of (2*S*,3*R*)-2-methyl-3-hydroxypentanoic acid (NDK-SNAC, 1 – 10 mM) [*33*] and DEBS$_{Mod6}$+TE (10 μM). Aliquots (35 μL) were removed and quenched by addition of 70% perchloric acid (1.75 μL). Samples were centrifuged at 18,000 × *g* to pellet the precipitated protein. The supernatant (33 μL) was removed and added to 1 M sodium bicarbonate (6.6 μL) bringing the final pH to 4-5. Excess salts were precipitated by freezing in liquid nitrogen and centrifuging at 18,000 × *g* until thawed. The supernatant was removed and analyzed on a Zorbax Eclipse XDB C-18 column (3.5 μm, 3 × 150 mm, 35°C, Agilent) using a linear gradient from 0 to 40% acetonitrile over 14 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 0% acetonitrile for 30 s (0.8 mL/min). Products were monitored using an Agilent G1315D diode array detector (TKL, $A_{260\ nm}$ or $A_{275\ nm}$; F-TKL, $A_{247\ nm}$; NDK-SNAC; $A_{260\ nm}$). The identity of each compound was verified using an Agilent 6130 single quadruple mass spectrometer in negative ion mode. For absolute quantification, each analyte was compared to an external standard curve. The concentration of the 2-fluoro-2-desmethyltriketide lactone (F-TKL) synthetic standard was determined by ^{19}F NMR using the ERETIC method [*34*] against an external standard of 5.00 mM 5-fluorouracil. The triketide lactone (TKL) standard was prepared enzymatically, and the 2-desmethyltriketide lactone (H-TKL) standard was

synthesized as described [*35*]. 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 75 mM diethylfluoromalonate.

Enzymatic preparation of methyl- and fluorotriketide lactones from methylmalonate and fluoromalonate. Reaction mixtures (TKL, 4 mL; F-TKL, 8 mL) containing NDK-SNAC (10 mM) were prepared as described above and incubated for 18 h. Protein was removed by the addition of 70% perchloric acid (0.05 volumes) and centrifuged at 18,000 × *g* for 10 min. The supernatant was removed and extracted extensively with dichloromethane (TKL, 5 × 15 mL; F-TKL, 5 × 30 mL) and the organic layers concentrated to 5-10 mL by rotary evaporation. The residue was transferred to a silanized glass vial (Sigmacote®, Sigma-Aldrich) and 50 mM sodium bicarbonate was added (1 mL). The dichloromethane was removed from the biphasic mixture by rotary evaporation to transfer the triketide into the aqueous phase. The aqueous solution of triketide was purified on a Zorbax Eclipse XDB C-18 column (5 μm, 9.4 × 250 mm, Agilent) using a linear gradient from 0 to 27.5% methanol with 50 mM sodium phosphate, pH 4.5 over 45 min as the aqueous mobile phase (3 mL/min). Fractions containing triketide were pooled and extracted with dichloromethane (4 × 3 volumes), and the combined organic layers were dried over magnesium sulfate and concentrated.

The TKL was purified further on a Zorbax Eclipse XDB C-18 (5 μm, 9.4 × 250 mm, Agilent) using a linear gradient from to 0 to 30% acetonitrile with 0.1% formic acid as the aqueous mobile phase over 45 min (3 mL/min) after transferring back into bicarbonate buffer as described above. Fractions containing TKL were combined and lyophilized for analysis. Due to the presence of a β-keto moiety, TKL was expected to be produced as a diastereomeric mixture, and was in fact isolated as a 100:7 mixture of (2*R*,4*S*,5*R*)-2,4-Dimethyl-3-oxo-5-hydroxy-*n*-heptanoic acid δ-lactone and its 2*S*-epimer. The observed NMR spectra are in agreement with the literature [*36*]. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.66 (ddd, J=8.4, 5.4, 2.9 Hz, 2*R* H_5), 4.48 – 4.42 (m, 2*S* H_5), 3.62 (q, J=6.6 Hz, 2*R* H_2), 3.25 (d, J=7.5 Hz, 2*S* H_2), 2.83 (dd, J=7.2, 4.7 Hz, 2*S* H_4), 2.63 (qd, J=7.6, 2.9 Hz, 2*R* H_4), 1.93 – 1.82 (m, 2*R* H_{6a}), 1.65 (dqd, J=14.8, 7.6, 5.4 Hz, 2*R* H_{6b}), 1.48 (d, J=7.5 Hz, 2*S* C_2-C$\underline{H}_3$), 1.37 (d, J=6.7 Hz, 2*R* C_2-C$\underline{H}_3$), 1.16 (d, J=6.4 Hz, 2*S* C_4-C$\underline{H}_3$), 1.12 (d, J=7.5 Hz, 2*R* C_4-C$\underline{H}_3$), 1.08 (t, J=7.5 Hz, 2*R* H_7), 1.01 (t, J=7.5 Hz, 2*S* H_7). ^{13}C NMR (226 MHz, $CDCl_3$, only the 2*R* epimer was detected): δ 205.59 (C_3), 170.21 (C_1), 78.68 (C_5), 50.56 (C_2), 44.52 (C_4), 24.19 (C_6), 10.09 (C_7), 9.90 (C_4-C$\underline{H}_3$), 8.40 (C_2-C$\underline{H}_3$). HR-ESI-MS $[M-H]^-$: calculated for $C_9H_{13}O_3$, *m/z* 169.0870, found *m/z* 169.0871.

The enzymatic F-TKL was compared against an authentic synthetic standard by LC-MS, HR-ESI-MS, ^{19}F-NMR, and GC-MS. ^{19}F-NMR (565 MHz, $CDCl_3$, $CFCl_3$ = 0 ppm): -171.95 (broad singlet), -210.32 (d, J = 45.8 Hz). HR-ESI-MS $[M-H]^-$: calculated for $C_8H_{10}FO_3$, *m/z* 173.0619, found *m/z* 173.0623.

Enzymatic preparation of F-TKL from fluoroacetate. One-pot reaction mixtures containing 200 mM HEPES, pH 7.5, TCEP (2 mM), bovine serum albumin (3 mg/mL), magnesium chloride (5 mM), fluoroacetate (10 mM), NDK–SNAC (10 mM), coenzyme A (2 mM), sodium bicarbonate (75 mM), ATP (2.5 mM), phosphoenolpyruvate (50 mM), pyruvate kinase (18 U/mL), myokinase (10 U/mL), AckA (10 μM), Pta (1 μM), ACCase (15 μM), MatB (40 μM), methylmalonyl-CoA epimerase (5 μM) and $DEBS_{Mod6}$+TE (10 μM) in a total volume of 1000 μL were incubated at 37°C for 1.5 hrs at which time sodium phosphate pH 7.5 (400 mM) was added to the reaction. The reaction was incubated at 37°C for a further 24 hrs. An aliquot (200 μL) was removed and prior to analysis, the aliquot was quenched by the addition of

70% perchloric acid (10 µL). F-TKL production was analyzed by LC-MS as described above using single ion monitoring at *m/*z 173 in negative ion mode. Telescope reaction mixtures containing 200 mM HEPES, pH 7.5, TCEP (2 mM), bovine serum albumin (3 mg/mL), magnesium chloride (5 mM), fluoroacetate (10 mM), coenzyme A (1 mM), sodium bicarbonate (75 mM), ATP (2.5 mM), phosphoenolpyruvate (50 mM), pyruvate kinase (18 U/mL), AckA (10 µM), Pta (1 µM) and ACCase (15 µM) in a total volume of 1000 µL were incubated at 37°C for 1.5 hrs. The reaction was then spun through an Amicon spin concentrator (MWCO 3 kDa) to remove proteins. 792 µL of the flow through was used to prepare a reaction with 400 mM sodium phosphate, pH 7.5, TCEP (2 mM), magnesium chloride (10 mM), NDK-SNAC (10 mM), phosphoenolpyruvate (50 mM), pyruvate kinase (18U/mL), myokinase (10 U/mL), MatB (40 µM), methylmalonyl-CoA epimerase (5 µM) and $DEBS_{Mod6}$+TE (10 µM) in a total volume of 1000 µL. The reactions were allowed to proceed for 24 h and were assayed as described for the one-pot reactions.

(2*S*,3*R*)-1-((*S*)-4-Benzyl-2-oxooxazolidin-3-yl)-2-methyl-1-oxopentan-3-yl 2-fluoroacetate (1). (*S*)-4-benzyl-3-((2*S*,3*R*)-3-hydroxy-2-methylpentanoyl)oxazolidin-2-one (171 mg, 0.585 mmol) was prepared as previously described [*33*] and combined with sodium fluoroacetate (70 mg, 0.703 mmol, 1.2 eq) and HATU (267 mg, 0.702 mmol, 1.2 eq) in a flame-dried round-bottom flask under a nitrogen atmosphere. Anhydrous THF (5.9 mL) and diisopropylethylamine (306 µL, 1.76 mmol, 3 eq) were added and the reaction was capped and stirred vigorously at room temperature for 44 h, during which time the white suspension turned orange-brown. The mixture was diluted with ethyl acetate and washed with saturated sodium bicarbonate, resulting in two clear layers. The orange-brown organic layer was washed again with saturated sodium bicarbonate, dried over $MgSO_4$, filtered through a plug of silica, and concentrated to give an orange-brown oil. The residue was purified by flash chromatography on silica (30 g) using a step gradient from 100% heptane to 25% ethyl acetate in heptane with the desired compound beginning to elute in 20% ethyl acetate. Fractions were concentrated to yield the product (173 mg, 84%) as a clear, colorless oil, R_f 0.35 (25% ethyl acetate/hexanes). ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.39 – 7.16 (m, Ph-H), 5.31 (ddd, J=7.9, 5.9, 3.2 Hz, $H_{3'}$), 4.87 (d, J=47.0, CH_2F-C=O), 4.60 (dddd, J=9.8, 7.7, 3.5, 2.3 Hz, H_4), 4.31 (ddd, J=8.7, 7.7, 0.8 Hz, H_5 *pro-R*), 4.19 (dd, J=8.9, 2.3 Hz, H_5 *pro-S*), 4.09 (qd, J=6.9, 3.2 Hz, $H_{2'}$), 3.28 (dd, J=13.4, 3.5 Hz, H_{6a}), 2.78 (dd, J=13.4, 9.8 Hz, H_{6b}), 1.79 – 1.64 (m, $H_{4'}$), 1.22 (d, J=6.9 Hz, C_2-CH_3), 0.95 (t, J=7.4 Hz, $H_{5'}$). ^{13}C NMR (151 MHz, $CDCl_3$): δ 173.95 ($C_{1'}$), 168.04 (d, J=22.0 Hz, CH_2F-C=O), 153.83 (N-C=O-O), 135.41 (C_{aryl}), 129.56 (C_{aryl}), 129.07 (C_{aryl}), 127.48 (C_{aryl}), 77.44 (d, J=182.3 Hz, CH_2F-C=O) , 76.40 ($C_{3'}$), 66.58 (C_5), 55.94 (C_4), 41.00 ($C_{2'}$), 38.05 (C_6), 25.16 ($C_{4'}$), 10.15 ($C_{2'}$-CH_3), 10.04 ($C_{5'}$). ^{19}F NMR (565 MHz, $CDCl_3$, $CFCl_3$ = 0 ppm): δ -230.44 (t, J=47.0 Hz). HR-ESI-MS $[M+Na]^+$: calculated for $C_{18}H_{22}FNO_5Na$, *m/z* 374.1374, found *m/z* 374.1381.

Preparation of (2S,4S,5R)-2-fluoro-4-methyl-3-oxo-5-hydroxy-n-heptanoic acid δ-lactone (F-TKL). Lactonization of 1 was carried out using literature methods [*35*]. 1 (160 mg, 0.455 mmol) was dried under vacuum in a pear-shaped flask then placed under nitrogen. In a flame-dried round-bottom flask, anhydrous THF (4.5 mL) and LiHMDS (1.0 M in THF, 1.366 mL, 3 eq) were combined, stirred and cooled to -78°C under nitrogen. The starting material was dissolved in anhydrous THF (3.5 mL), cooled to -78°C, and cannulated dropwise into the solution of base over 20 min. A rinse of anhydrous THF (1.5 mL) was also transferred by cannula. The reaction mixture was stirred for 3 h at -78°C and quenched by addition of saturated ammonium chloride/methanol/water (1:1:1 v/v/v, 13 mL). The mixture was then allowed to warm to room temperature while stirring. The pH of the quenched mixture was adjusted to 9

using 10 M NaOH and extracted with 3 × 40 mL ethyl acetate to remove the oxazolidinone auxiliary. The aqueous layer was adjusted to pH 2 using 12 M HCl and then extracted with 5 × 20 mL dichloromethane. The combined organic layers were concentrated to give a clear, colorless oil, which contained approximately 1 mol% starting material by ^{1}H NMR (36 mg, 45%). The product was further purified by flash chromatography on silica by washing extensively with dichloromethane (R_f < 0.05) then eluting with ethyl acetate (R_f ~ 0.4), and concentrated to yield a white, crystalline solid (13 mg, 16%). A mixture of enol (53%) and keto (47%) tautomers was observed in $CDCl_3$. The keto form was almost exclusively the 2*S* diastereomer as determined by ^{1}H NOESY and molecular modeling *(Figure 2.17)*. The doublet in the ^{19}F NMR spectrum in $CDCl_3$ at -205.96 ppm was assigned to the 2*R* keto diastereomer based on ^{1}H-^{19}F HMBC. ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.83 (d, J=45.7 Hz, *keto* H_2), 4.71 (ddd, J=8.4, 5.2, 3.0 Hz, *keto* H_5), 4.26 (ddd, J=8.8, 6.0, 3.3 Hz, *enol* H_5), 2.68 (qd, J=7.5, 3.0 Hz, *keto* H_4), 2.45 (qt, J=7.2, 3.8 Hz, *enol* H_4), 1.88 – 1.78 (m, *keto* H_{6a}), 1.75 (m, *enol* H_{6a}), 1.60 (m, *keto* H_{6b}), 1.55 – 1.44 (m, *enol* H_{6b}), 1.14 (d, J=7.5 Hz, *keto* H_8), 1.11 (d, J=7.1 Hz, *enol* H_8), 1.00 (t, J=7.4 Hz, *keto* H_7), 0.92 (t, J=7.5 Hz, *enol* H_7). ^{13}C NMR (151 MHz, $CDCl_3$): δ 198.58 (d, J=13.4 Hz, *keto* C_3), 164.15 (d, J=20.0 Hz, *keto* C_1), 162.67 (d, J=24 Hz, *enol* C_1), 156.94 (d, J=6.5 Hz, *enol* C_3), 130.00 (d, J=232.3 Hz, *enol* C_2), 89.82 (d, J=206.9 Hz, *keto* C_2), 80.27 (*enol* C_5), 77.89 (d, J=1.7 Hz, *keto* C_5), 43.97 (*keto* C_4), 35.84 (*enol* C_4), 24.15 (*keto* C_6), 23.96 (*enol* C_6), 10.38 (d, J=2.7 Hz, *enol* C_8), 10.11 (*keto* C_8), 9.93 (*keto* C_7), 9.72 (*enol* C_7). ^{19}F NMR (565 MHz, $CDCl_3$, $CFCl_3$ = 0 ppm): δ -172.36 (d, J=4.3 Hz, *enol*), -205.96 (d, J=44.9 Hz, 2*S keto*), -210.40 (d, J=45.6 Hz, 2*R keto*). ^{19}F NMR (565 MHz, 10% D_2O, 50 mM sodium phosphate pH 4.5): δ -178.66. ^{19}F NMR (565 MHz, 15% D_2O, 85 mM Tris pH 7.5, 5-fluorouracil = -168.3 ppm): δ -190.20. HR-ESI-MS $[M-H]^-$: calculated for $C_8H_{10}FO_3$, *m/z* 173.0619, found *m/z* 173.0617.

GC-MS analysis of F-TKL. Samples were dissolved in dichloromethane and BSTFA containing 1% trimethylsilyl chloride (Sigma-Aldrich, 0.1 volumes) was added. Samples were analyzed on a Trace GC Ultra (Thermo Scientific) coupled to a DSQII single-quadrupole mass spectrometer using an HP-5MS column (0.25 mm × 30 m, 0.25 μM film thickness, J & W Scientific). The injection volume was 1 μL and the oven program was as follows: 75°C for 3 min, ramp to 25°C at 25°C min^{-1}, ramp to 300°C at 50 °C min^{-1}, hold for 1 min.

Covalent inhibition assay for DEBS$_{Mod6}$+TE. Two triketide reaction mixtures (200 μL) were prepared as described above, one containing fluoromalonate (10 mM) and the other methylmalonate (10 mM). DEBS$_{Mod6}$+TE (10 μM) and NDK-SNAC (2.5 mM) were added to each and the reactions were incubated at 37°C for 18 h. The protein fraction was isolated from each mixture at room temperature by desalting on a Sephadex G-25 column (3 mL) using 400 mM sodium phosphate, pH 7.5. Fractions were pooled by Bradford assay and concentrated to 200 μL using Amicon Ultra spin concentrators (3 kDa MWCO). The isolated DEBS$_{Mod6}$+TE was assayed by adding TCEP (2.5 mM), methylmalonyl-CoA (1 mM) and NDK-SNAC (1 mM) to this mixture to give a final volume of 210 μL and incubating at 37°C for 3 h, then analyzed by HPLC as described above.

Triketide lactone production using AT^0 constructs. All assay mixtures contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase (27 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 μM) and NDK-SNAC (5 mM). Where appropriate, fluoromalonyl-CoA (1 mM), methylmalonyl-CoA (1 mM) or MatB (40 μM), fluoromalonate or

methylmalonate or malonate (20 mM), CoA (1 mM) and DszAT (5 μM) were added. Reactions were initiated by addition of the appropriate DEBS+TE construct (Mod6 and Mod6/AT^0, 10 μM; Mod3 and Mod3/AT^0, 5 μM in reactions containing DszAT and 8 μM otherwise) and incubated at 37°C for 18-20 h. Aliquots were removed, quenched, processed and analyzed as described above.

Tetraketide lactone production. All reactions contained 400 mM sodium phosphate (pH 7.5 for the 2,4-dimethyl- and 2-fluoro-4-methyl-tetraketide lactone reactions and pH 6 for the 2-methyl-4-fluoro-tetraketide lactone reaction), glycerol (20%), phosphoenolpyruvate (20 mM), TCEP (10 mM), magnesium chloride (5 mM), ATP (2.5 mM), pyruvate kinase (18 U/mL), myokinase (10 U/mL), methylmalonyl-CoA epimerase (5 μM), MatB (20 μM), CoA (1 mM), methylmalonyl-CoA (100 μM), NDK-SNAC (1 mM) and reduced nicotinamide adenine dinucleotide phosphate (NADPH; 5 mM).

The reaction to produce 2,4-dimethyl-tetraketide lactone also contained methylmalonate (5 mM), $DEBS_{Mod2}$ (10 μM) and $DEBS_{Mod3}$+TE (2 μM). The reaction to produce 2-fluoro-4-methyl-tetraketide lactone also contained fluoromalonate (5 mM), $DEBS_{Mod2}$ (10 μM), $DEBS_{Mod3}/AT^0$ (2 μM) and DszAT (2 μM). The reaction to produce 2-methyl-4-fluoro-tetraketide lactone also contained fluoromalonate (5 mM), $DEBS_{Mod2}/AT^0$ (10 μM) and $DEBS_{Mod3}$ (2 μM).

All reactions were initialized by the addition of $DEBS_{Mod2}$ or $DEBS_{Mod2}/AT^0$ and incubated at 37°C overnight. Reactions were then saturated with sodium chloride and the aqueous layer was acidified by the addition of 0.1 volumes of 70% perchloric acid and extracted four times into 2 volumes of chloroform. The chloroform layer was concentrated by vacuum centrifugation and the tetraketide lactones were resuspended in water for analysis. Tetraketide lactones were analyzed by LC-MS using a Phenomenex Kinetex XB-C18 1.7 μm 150 x 2.1 mm column with a mobile phase of ammonium acetate (50 mM) with a gradient from 0 to 60% acetonitrile over 15 min and detected on an Agilent single quadruple mass spectrometer in negative ion mode.

ESI-MS/MS analysis of tetraketide lactones. ESI LC/MS/MS spectra were collected using an LTQ FT (Thermo Scientific). Negative ions analyzed in linear ion trap mode. MS/MS spectra were collected with the following normalized collision energies: TKL 26, F-TKL 35, all tetraketides 26.

2.3 Results and Discussion

Chain elongation in polyketides and related fatty acid-based natural products relies on a separate pool of extender units formed by carboxylation of acyl-CoAs at the α-position. These malonyl-CoA derivatives are then used as masked enolates for C–C bond formation following decarboxylation. The fluorinated extender, fluoromalonyl-CoA, can be made through two routes: either a two-step activation of biogenic fluoroacetate or by a direct ligation of CoA to fluoromalonate.

We reasoned that the acetate kinase (AckA)–phosphotransacetylase (Pta) pair would be effective at fluoroacetate activation, as mutations in this gene locus have been shown to lead to fluoroacetate resistance in *Escherichia coli* [*37*]. The enzymes from *E. coli* were therefore overexpressed and characterized biochemically, confirming that AckA and Pta serve as an effective activation system to rapidly produce both acetyl- and fluoroacetyl-CoA at near quantitative yield (*Figures 2.2 and 2.3*).

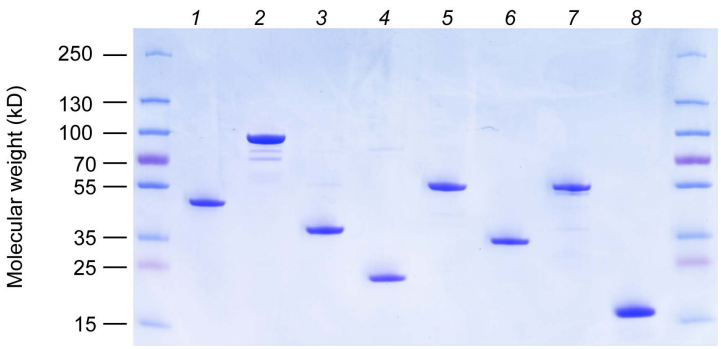


Figure 2.2. *Enzymes used in extender unit biosynthesis. (1, AckA; 2, Pta; 3, AccA; 4, AccB; 5, AccC; 6, AccD; 7, MatB; 8, methylmalonyl-CoA epimerase).*

Analysis of the kinetic parameters for these enzymes with respect to fluorinated substrates indicate that neither appears to be affected by the fluorine substituent beyond inductive effects that alter the nucleophilicity of the carboxylic acid (AckA) or electrophilicity of the carbonyl

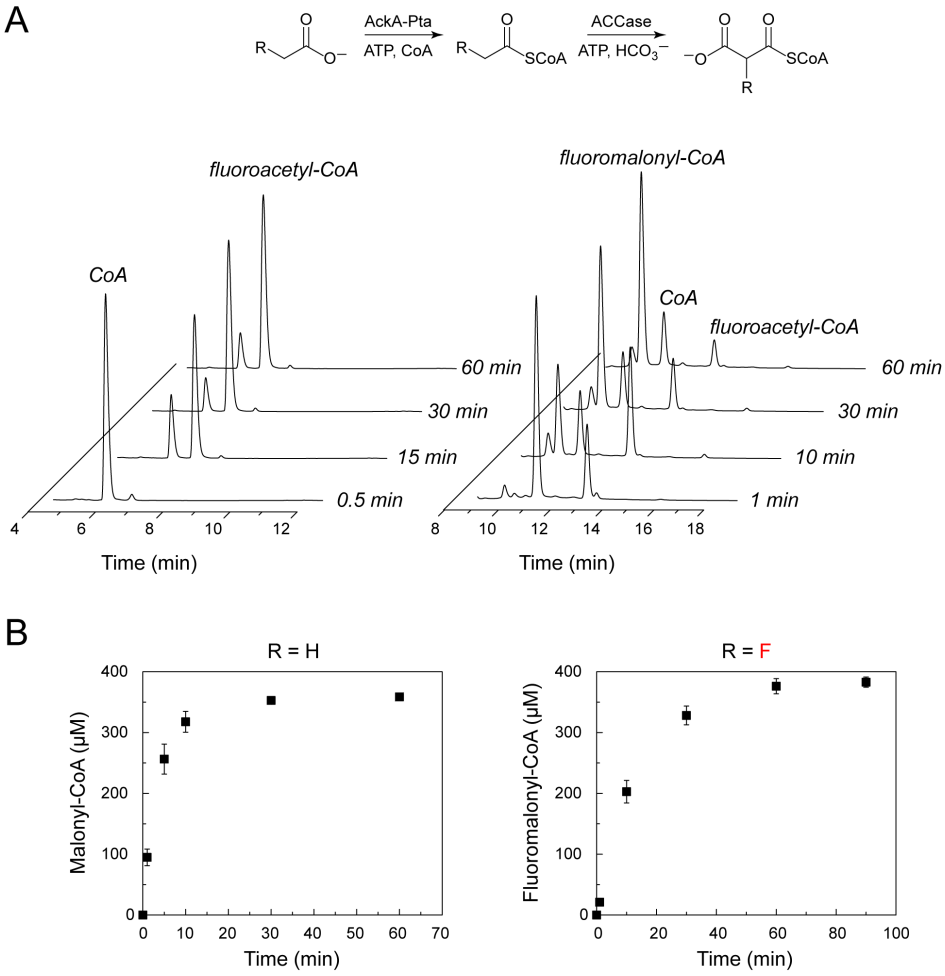


Figure 2.3. *Enzymatic synthesis of extender units from acetate and fluoroacetate. (A) Formation of fluoroacetyl-CoA (left) fluoromalonyl-CoA (right) monitored by RP-HPLC. (B) Extent of reaction from 500 μM CoA and either acetate (left) or fluoroacetate (right), utilizing AckA, Pta and ACCase.*

(Pta) [*38*].

Next, we purified the individual AccABCD subunits that make up the acetyl-CoA carboxylase (ACCase) from *E. coli* (*Figure 2.2*) and added these enzymes to the AckA–Pta system in order to carry out the carboxylation of fluoroacetate in a one-pot reaction to generate the fluoromalonyl-CoA extender unit (*Figure 2.3*). Under these conditions, the ligation of CoA with AckA–Pta to produce the acyl-CoA is rapid and production of the carboxylated product is limited by the ACCase. Although the rate of conversion is 2.5-fold slower for fluoroacetate compared to acetate, the overall extent of reaction is similar for both congeners and suggests that covalent inactivation of the ACCase by fluoroacetyl-CoA is not significant (*Figure 2.3*). In addition to the route from fluoroacetate, we also tested a malonyl-CoA synthetase (MatB) [*39*] for coupling CoA directly to fluoromalonate (*Figure 2.2*). Although MatB greatly prefers malonate over the fluorinated congener, fluoromalonyl-CoA is still produced at reasonable efficiency (*Figure 2.4*). Both of these systems also provide *in situ* regeneration capacity that can amplify product yields from polyketide synthases, and we found that either system increased polyketide production by tetrahydroxynaphthalene synthase compared with simple addition of malonyl-CoA (*Figure 2.5*).

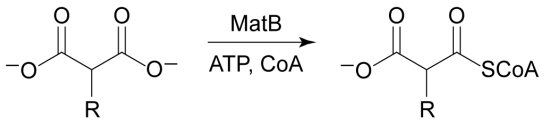
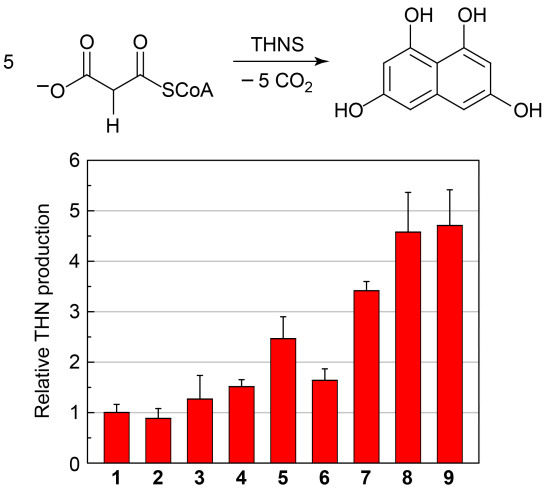


Figure 2.4. *Kinetic parameters for malonate activation (MatB, malonyl-CoA synthetase).*



1	Malonyl-CoA
2	Malonyl-CoA, acetyl-CoA
3	Acetyl-CoA, ACCase, ATP
4	Malonyl-CoA, ACCase, ATP
5	Malonyl-CoA, MatB, malonate, ATP, myokinase
6	Acetyl-CoA, AckA/PTA, ACCase, ATP
7	Acetyl-CoA, AckA/PTA, ACCase, ATP, 1 eq acetate
8	Acetyl-CoA, AckA/PTA, ACCase, ATP, 2 eq acetate
9	Acetyl-CoA, AckA/PTA, ACCase, ATP, 20 eq acetate

Figure 2.5*. Efficiency of polyketide production with tetrahydroxynaphthalene synthase (THNS) [Izumikawa et al. J Ind Microbiol Biotechnol (**2003**) 30, 510–515] using different extender regeneration systems. THNS uses only malonyl-CoA as both starter and extender unit. All samples contained a fixed amount of malonyl- or acetyl-CoA (0.5 mM), and relative THN production was monitored at A510 nm. Samples with no regeneration system (1, 2) were compared to those containing regeneration systems related to non-productive decarboxylation (3, 4) and hydrolysis (5), while also providing additional substrate (5-9) in situ. Values are reported as the mean ± s.d. (n = 3).*

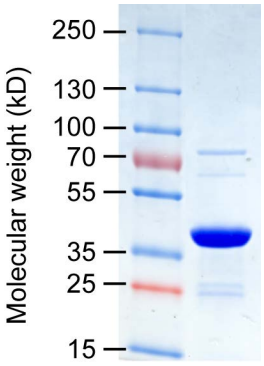


Figure 2.6. *Heterologous expression of NphT7.*

We next turned our attention to utilizing the fluoromalonyl-CoA monomer for downstream chain elongation reactions. To start, we examined the behavior of a simple polyketide synthase system with regard to one cycle of chain extension and ketoreduction. These steps are key to the function of larger multimodular systems because they often control downstream cyclization and rearrangements within the polyketide backbone [*3, 14*]. We constructed a synthetic gene encoding NphT7 from *Streptomyces sp. CL190* [*40*] and isolated the heterologously-expressed enzyme for biochemical characterization (*Figures 2.6 and 2.7*). NphT7

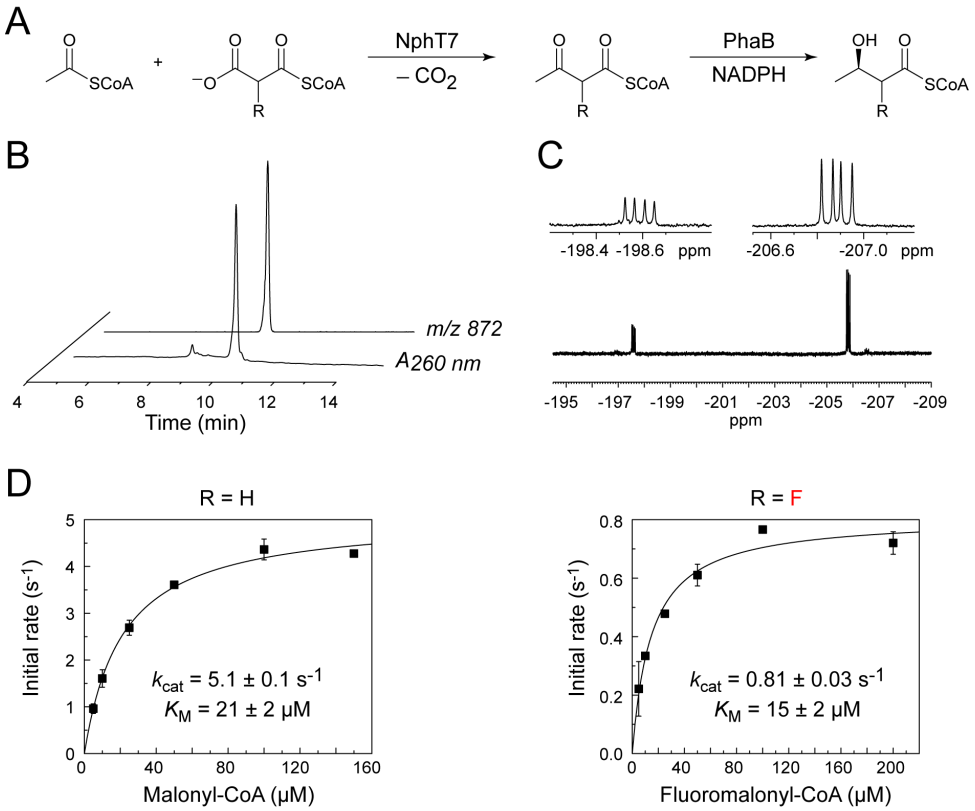


Figure 2.7. *Chain extension and keto reduction with a fluorinated extender. (A) Reactions catalyzed by NphT7 and PhaB. (B) Steady-state kinetic parameters for NphT7-catalyzed C–C bond formation measured using a coupled assay with PhaB. (C) HPLC trace of 2-fluoro-3-hydroxybutyryl-CoA isolated from enzymatic reaction mixtures (m/z = 872). (D) 19F NMR of the product shows that both diastereomers are produced.*

appears to be a free-standing ketosynthase that is related at the structural level to the ketosynthase domain of more complex polyketide synthases (*Figure 2.8*).

Using a coupled assay with an *R*-specific acetoacetyl-CoA reductase (PhaB), we found that NphT7 is competent to catalyze the formation of acetofluoroacetyl-CoA using an acetyl-CoA starter and fluoromalonyl-CoA extender with only a five-fold defect in catalytic efficiency (k_{cat}/K_M) derived from a drop in k_{cat} with the fluorinated substrate (*Figure 2.7D*). This lower turnover rate observed with the fluorinated substrate may be related to the reduced reactivity of the enolate species, which could be stabilized by the fluorine substituent. However, the overall yield was comparable for both fluorinated and nonfluorinated substrates, which shows that a decarboxylative Claisen condensation with fluoromalonyl-CoA can take place with a similar extent of conversion

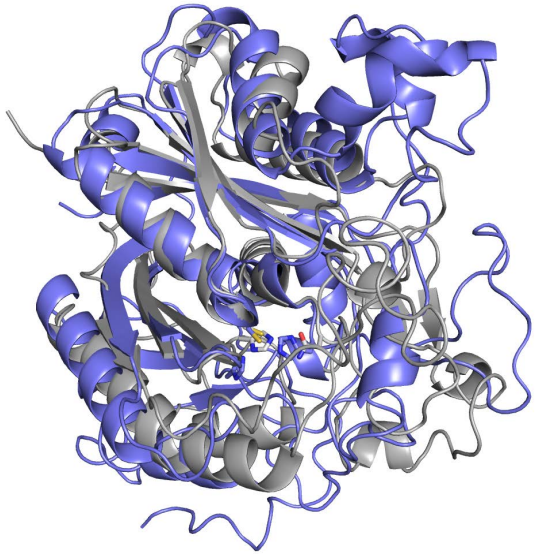


Figure 2.8. *Structural alignment of NphT7 and a DEBS ketosynthase domain. The NphT7 structure was predicted using Phyre2 [42] and based on a type III 3-oxoacyl-(acyl-carrier protein) synthase from Burkholderia xenovorans (PDB ID 4EFI). Despite low sequence identity (<20%), the predicted structure overlays well with the KS domain from $DEBS_{Mod5}$ [43]. Active site residues (C119, H334 H374 (N in NphT7); DEBS numbering) are highlighted.*

compared to malonyl-CoA. Furthermore, these experiments also show that the 2-fluoro-3-keto motif produced with the fluoromalonyl-CoA extender can be accepted by ketoreductases, as PhaB is capable of efficiently reducing the acetofluoroacetyl-CoA substrate (*Figure 2.7B*). The ^{1}H and ^{19}F NMR spectra of the reduced product indicate that both diastereomers are produced in this reaction (*Figures 2.6C and 2.9*), which may result either from lack of stereochemical preference of NphT7 with respect to the fluorine substituent or to racemization of the product prior to reduction by PhaB. Although PhaB does not appear to show diastereoselectivity with respect to the fluorine group, polyketide synthase ketoreductases are known to be selective with regard to their native α-substituent and could potentially carry out the stereochemical resolution of the fluorine modification upon reduction [*41*].

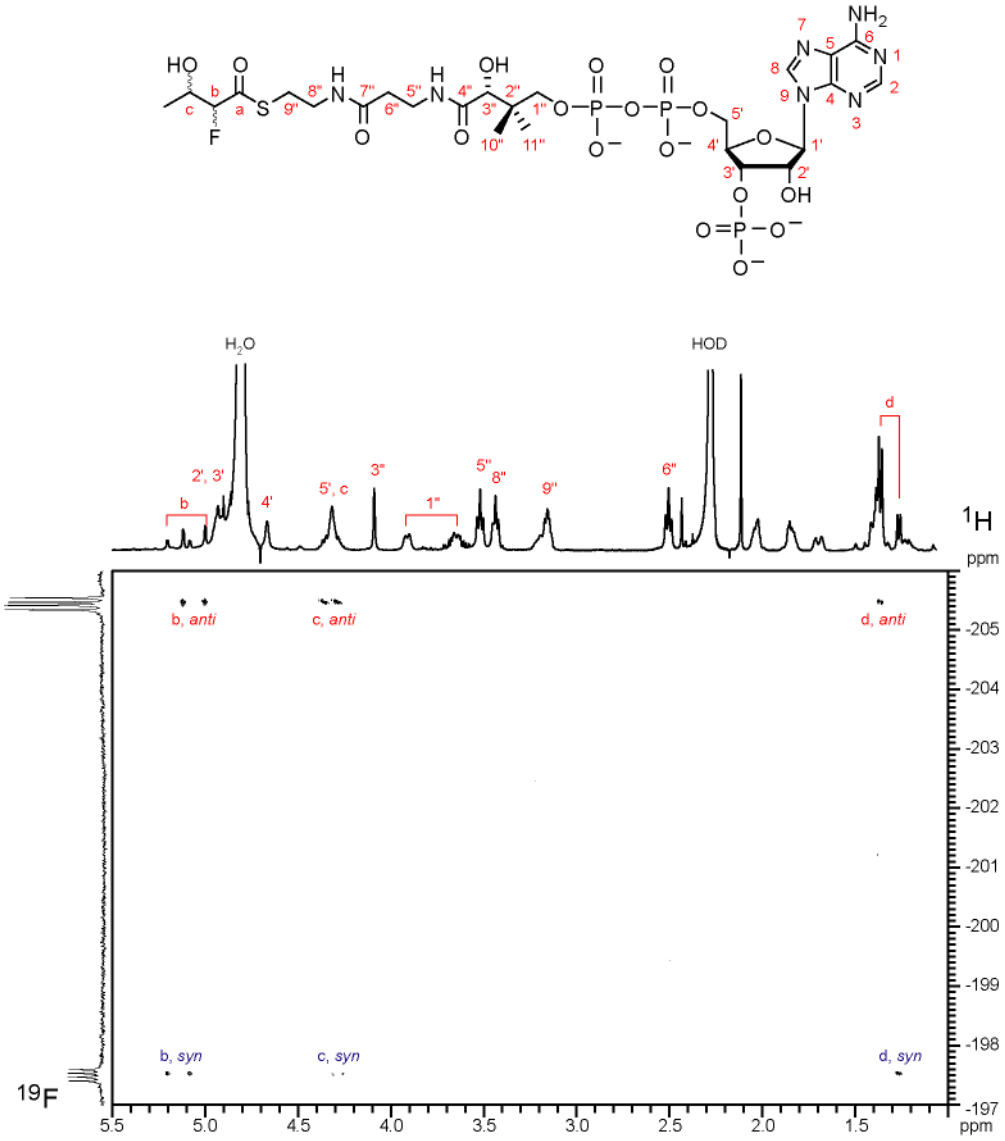


Figure 2.9. *^{1}H-^{19}F HMBC NMR analysis of enzymatically synthesized 2-fluoro-3-hydroxybutyryl-CoA. Based on data from other α-fluoroalcohols [44], the 19F resonance for the anti configuration of the fluorine and hydroxyl groups should be found upfield of the syn and was assigned as the major product. If PhaB maintains is native selectivity as an R-specific acetoacetyl-CoA reductase, the anti product is (2S, 3R)-2-fluoro-3-hydroxybutyryl-CoA.*

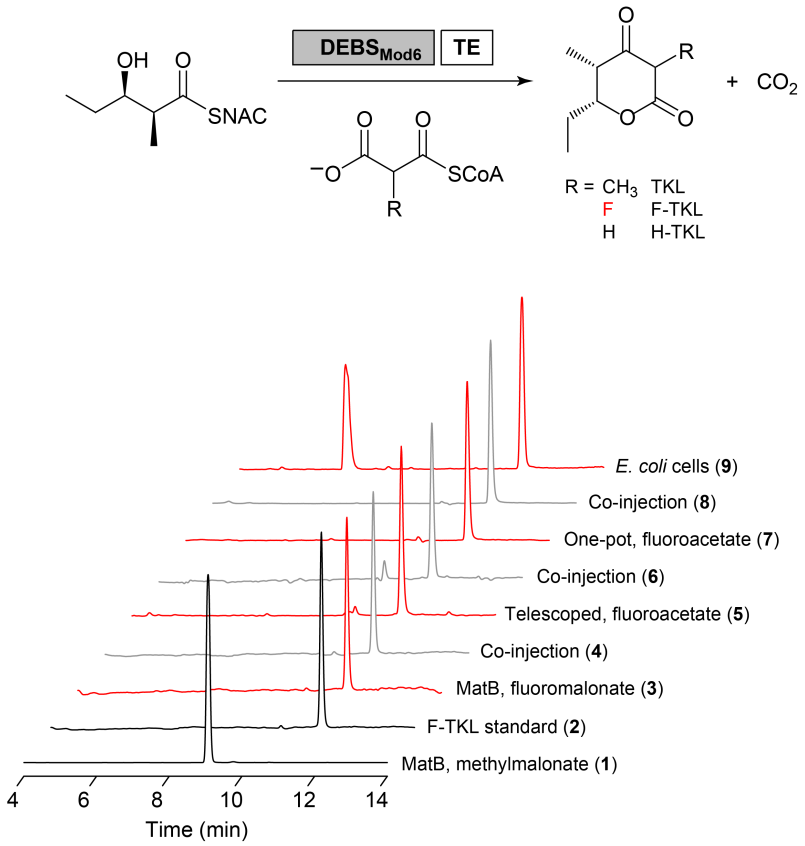


Figure 2.11. *Production of a fluorinated polyketide using DEBS$_{Mod6}$+TE. Chain extension by DEBS$_{Mod6}$+TE to form triketide lactones monitored by HPLC (TKL, m/z = 169; F-TKL, m/z = 173). CoA, ATP, and ATP regeneration system are included in all in vitro reactions. Data are normalized with respect to the TKL peak. Methylmalonate and MatB (1), synthetic F-TKL standard (2), fluoromalonate and MatB (3), co-injection of 2 and 3 (4), fluoroacetate, AckA–Pta, ACCase, and MatB were incubated before filtering through a 3 kDa MWCO membrane and adding DEBSMod6+TE and NDK-SNAC (5), co-injection of 5 and 2 (6), one-pot reaction using conditions from 5 (7), co-injection of 7 and 2 (8), (9), E. coli cells (see below).*

With this information in hand, we sought to extend our biosynthetic method for fluorine introduction to more complex polyketide synthase systems, which use the chain elongation reaction for the biosynthesis of many bioactive and clinically important natural products, such as erythromycin and rapamycin [*3, 14*]. Of these, the 6-deoxyerythronolide B synthase (DEBS), a multimodular polyketide synthase responsible for production of the erythromycin precursor, is likely the most well understood [*45*]. We therefore focused on studying chain extension by the sixth module of DEBS, including the terminal thioesterase (DEBS$_{Mod6}$+TE) (*Figure 2.10)* [*28*]. Using a diketide substrate (NDK-SNAC), DEBS$_{Mod6}$+TE can elongate the chain with its native methylmalonyl-CoA extender unit and cleave the tethered product to form methyltriketide lactone (TKL) (*Figure 2.11, R= CH3,* and *Figure 2.12*) [*27*].

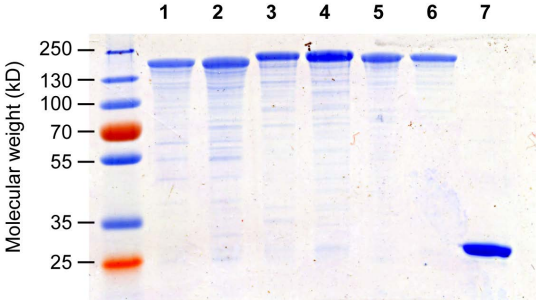
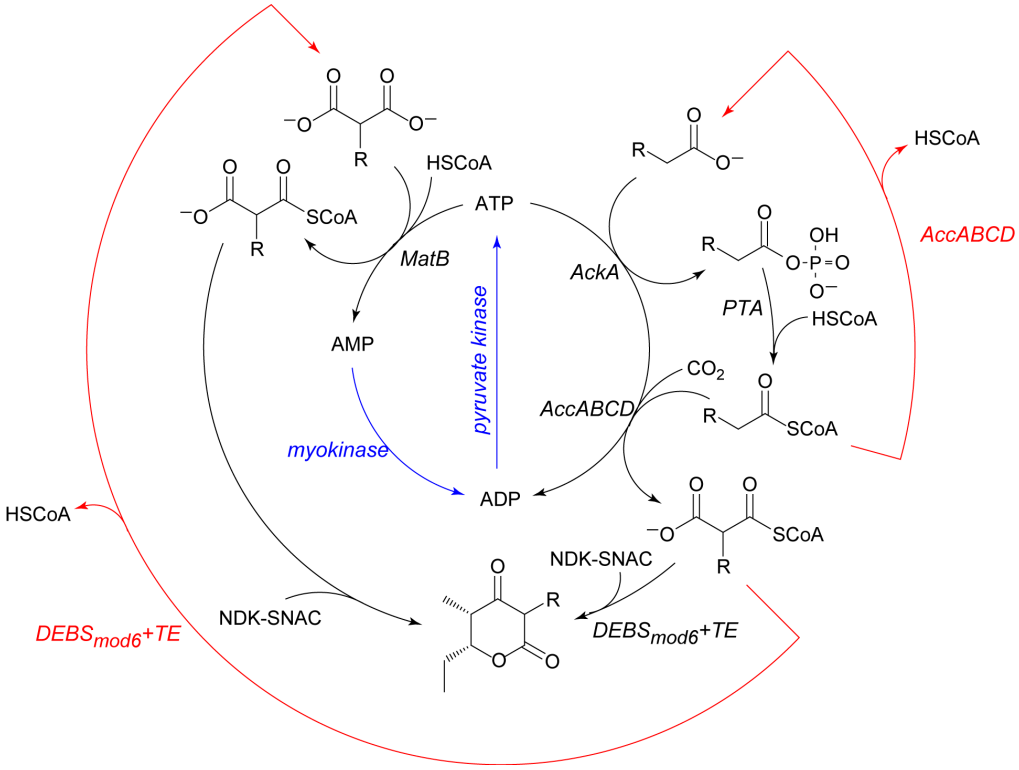


Figure 2.10. *Enzymes used in the production of model polyketides. (1, DEBS$_{Mod2}$; 2, DEBS$_{Mod2}$AT^{0}; 3, DEBS$_{Mod3}$+TE; 4, DEBS$_{Mod3}$AT^{0}+TE; 5, DEBS$_{Mod6}$+TE; 6, DEBS$_{Mod6}$AT^{0}+TE; 7, DszAT).*



Scheme 2.1. *Hydrolysis and regeneration reactions for F-TKL synthesis. Reaction scheme showing enzymes present in F-TKL forming reactions including observed non-productive hydrolysis reactions (red) and the ATP regenerating system (blue).*

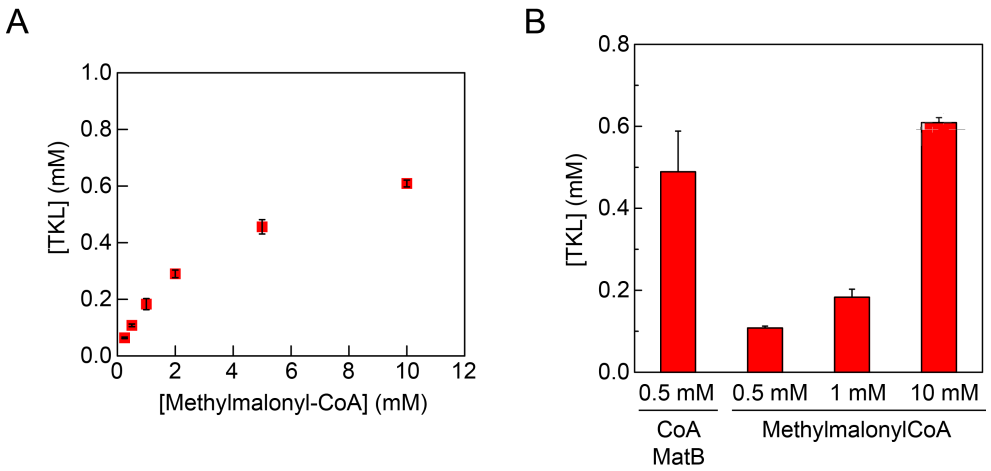


Figure 2.12. *Amplification of TKL formation using MatB. All reactions contained 400 mM sodium phosphate, pH 7.5, phosphoenolpyruvate (50 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase (27 U/mL), myokinase (10 U/mL), methylmalonyl- CoA epimerase (5 μM), methylmalonate (20 mM), NDK-SNAC (1 mM) and DEBSMod6+TE (10 μM). The source of extender unit was either methylmalonyl-CoA (0.5 – 10 mM) or MatB (40 μM) and CoA (0.5 mM). (A) Dependence of TKL formation on methylmalonyl-CoA. Data are average ± s.d. (n = 3) (B) Comparison of TKL yield with and without MatB regeneration. Values are reported as the mean ± s.d. (n = 3).*

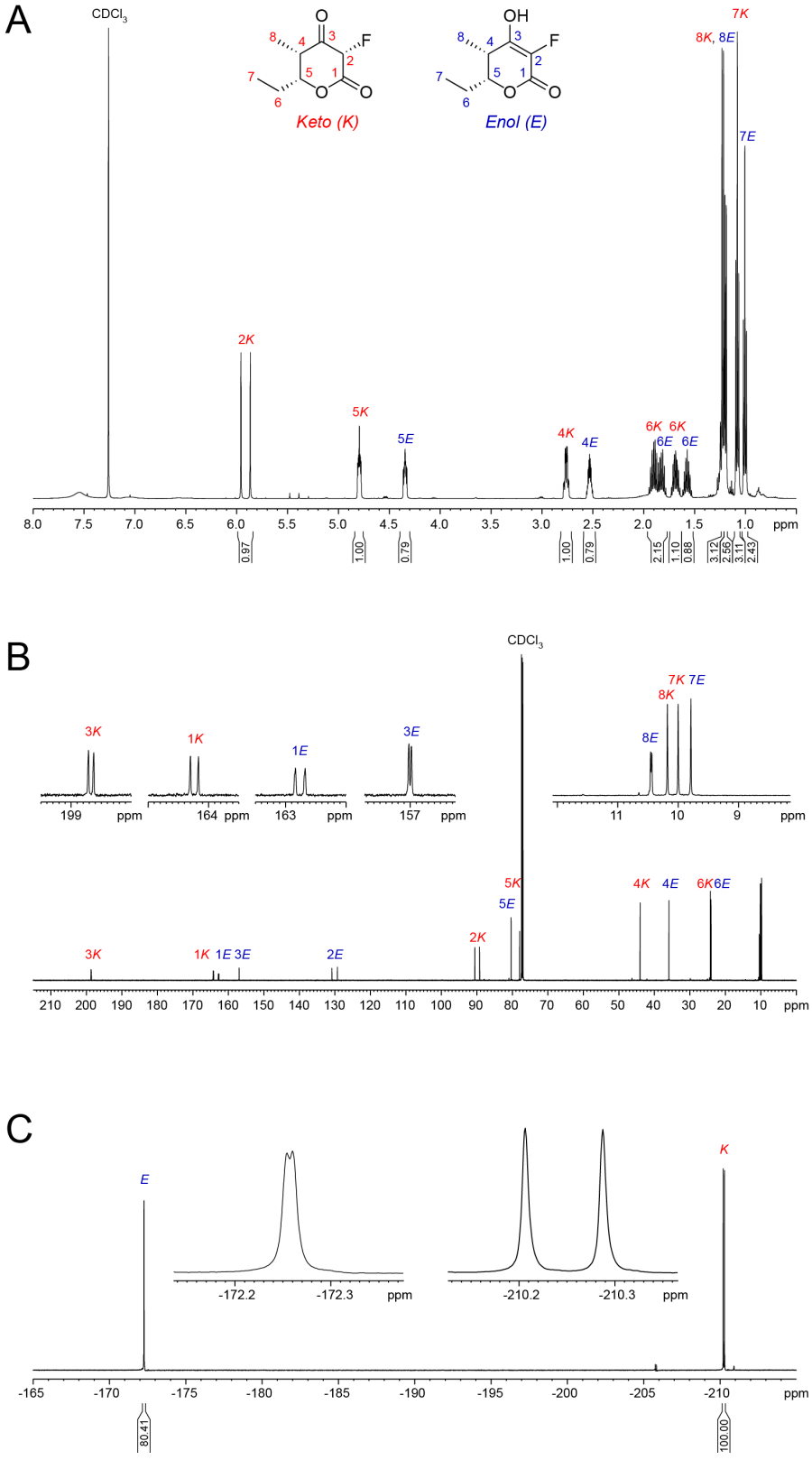


Figure 2.13. *1D-NMR spectra of synthetic F-TKL standard in $CDCl_3$. (A) ^{1}H NMR. (B) ^{13}C NMR. (C) ^{19}F NMR. The relative keto:enol ratio in $CDCl_3$ depends on concentration and increases with decreasing concentration.*

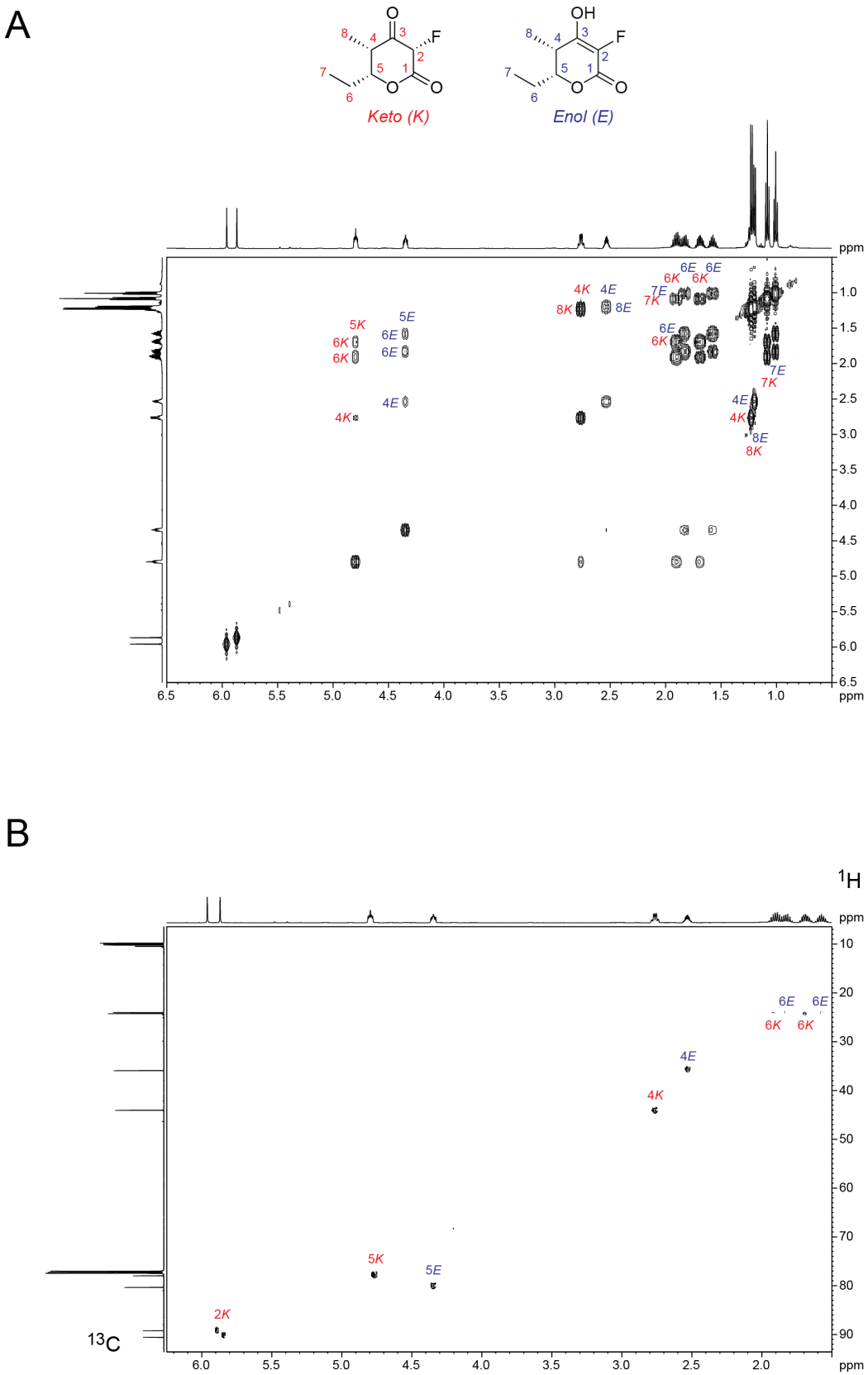
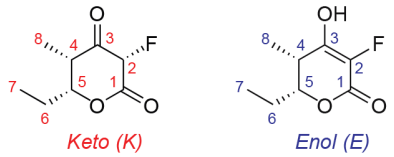
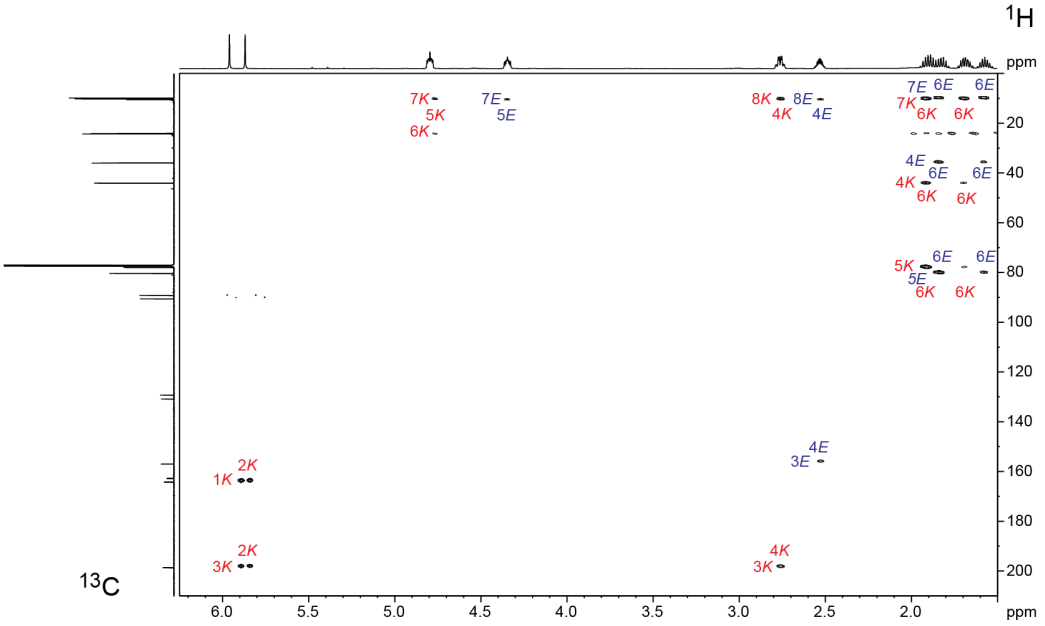


Figure 2.14. *2D-NMR spectra of synthetic F-TKL standard in $CDCl_3$. (A) COSY. (B) ^{1}H-^{13}C HSQC.*

C





D

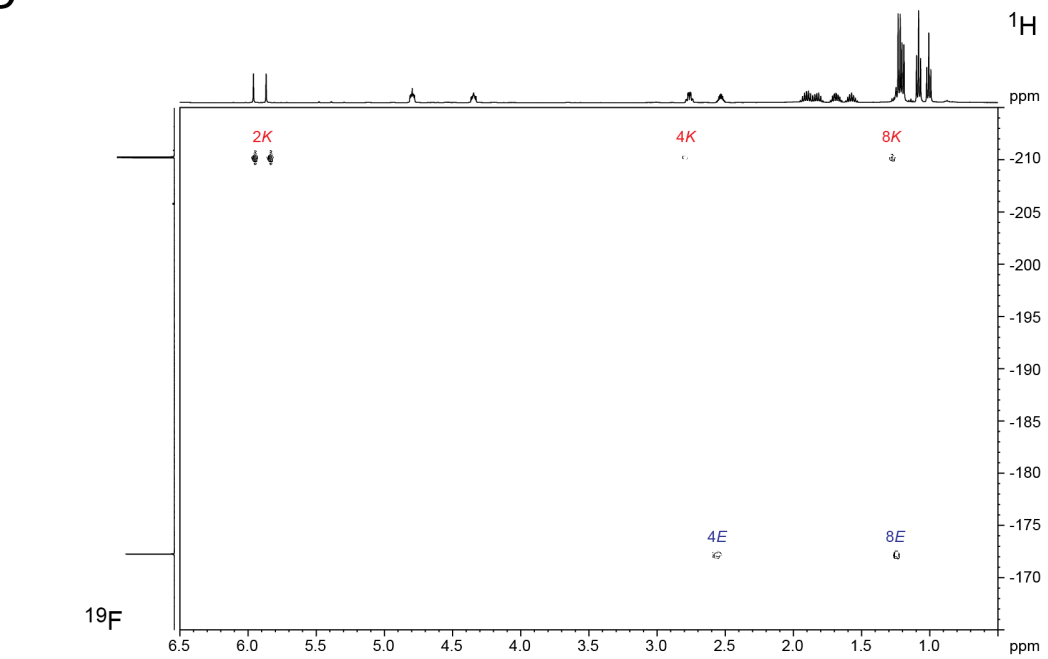


Figure 2.14. *2D-NMR spectra of synthetic F-TKL standard in $CDCl_3$, cont'd. (C) ^{1}H-^{13}C HMBC. (D) ^{1}H-^{19}F HMBC.*

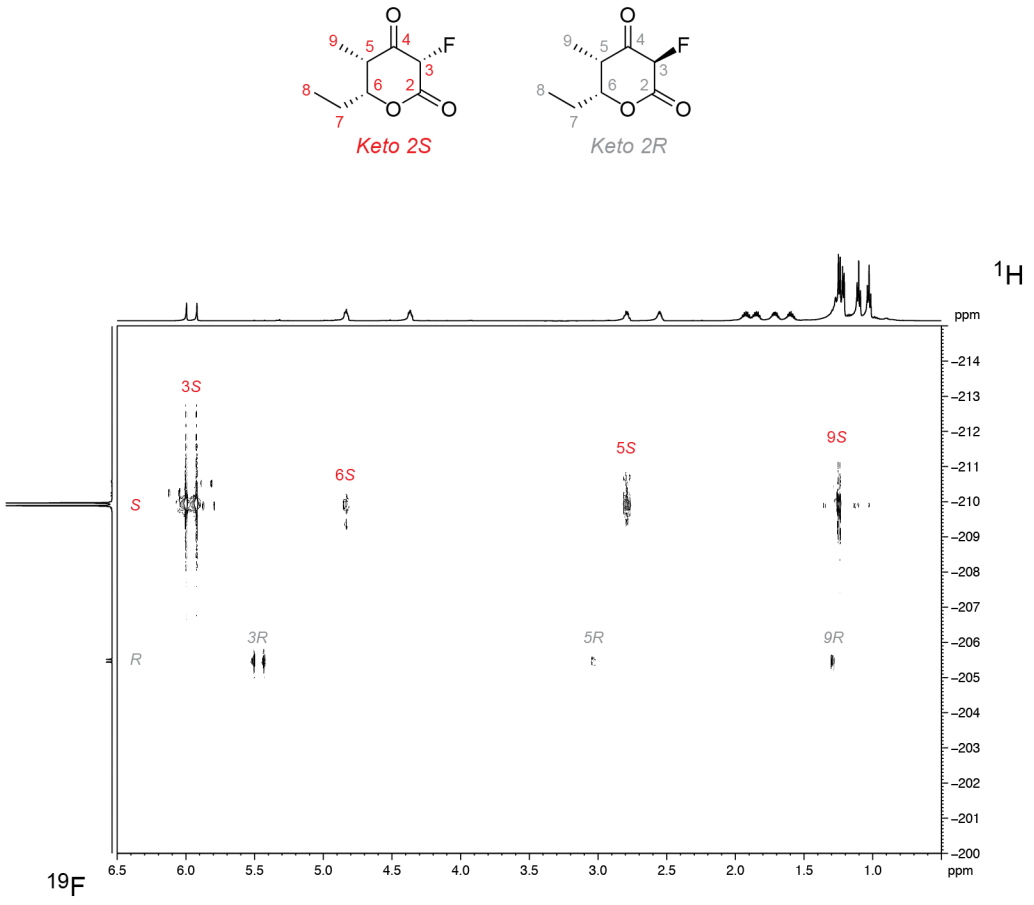


Figure 2.15. *^{1}H-^{19}F HMBC of keto isomer region of synthetic F-TKL standard in $CDCl_3$, showing crosspeaks for the major and minor epimers.*

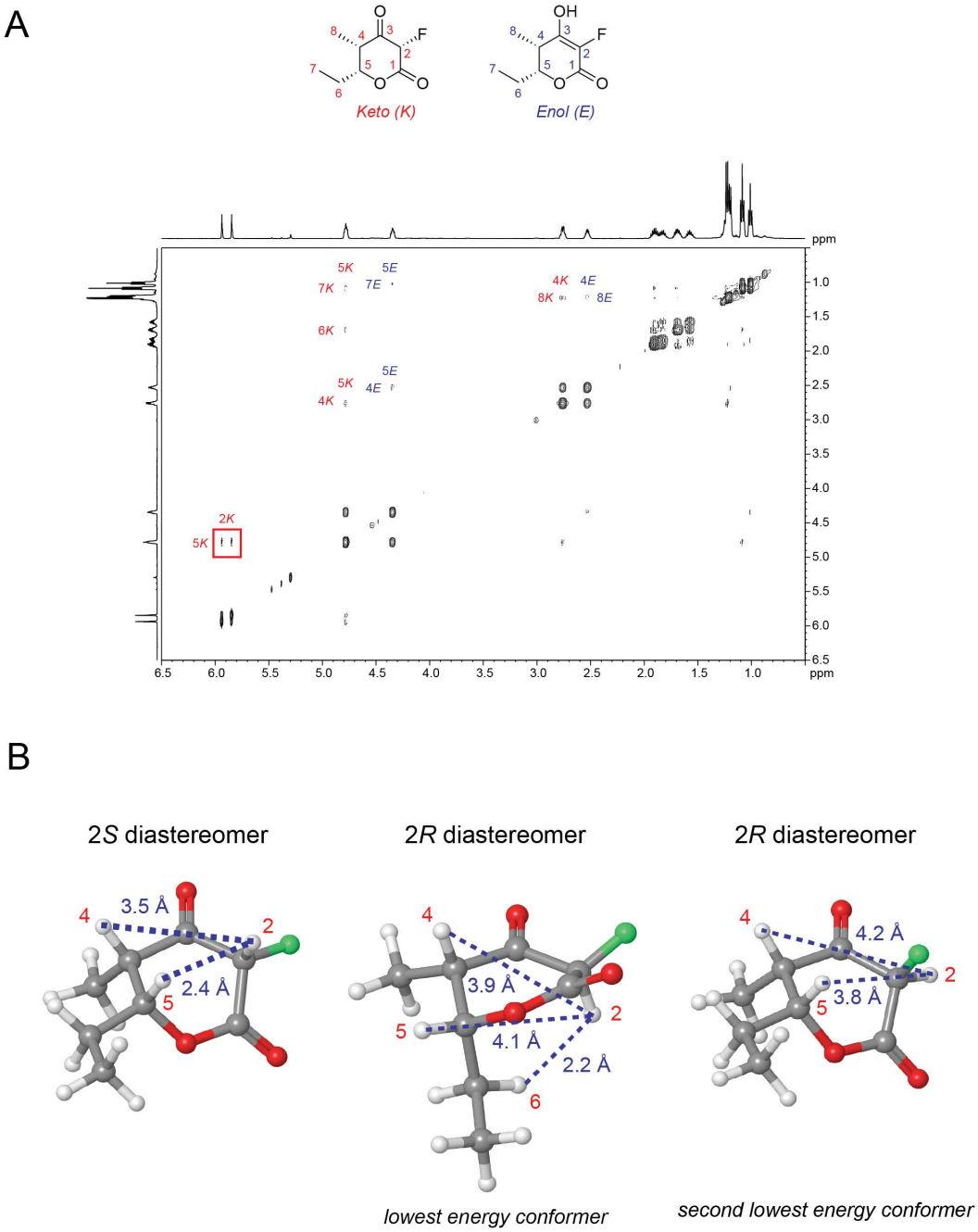


Figure 2.17. *Stereochemical analysis for F-TKL. (A) ^{1}H NOESY spectrum of synthetic F-TKL standard in $CDCl_3$. The same ratio between epimers is observed for enzymatically produced F-TKL. (B) Molecular modeling results for F-TKL. The lowest energy conformations of the two F-TKL keto diastereomers were selected based on a conformational search (Macromodel) using Maestro 9.3 (Schrödinger, Inc). Only the 2S epimer would be expected to show a single NOE coupling between H2 and H5, as observed.*

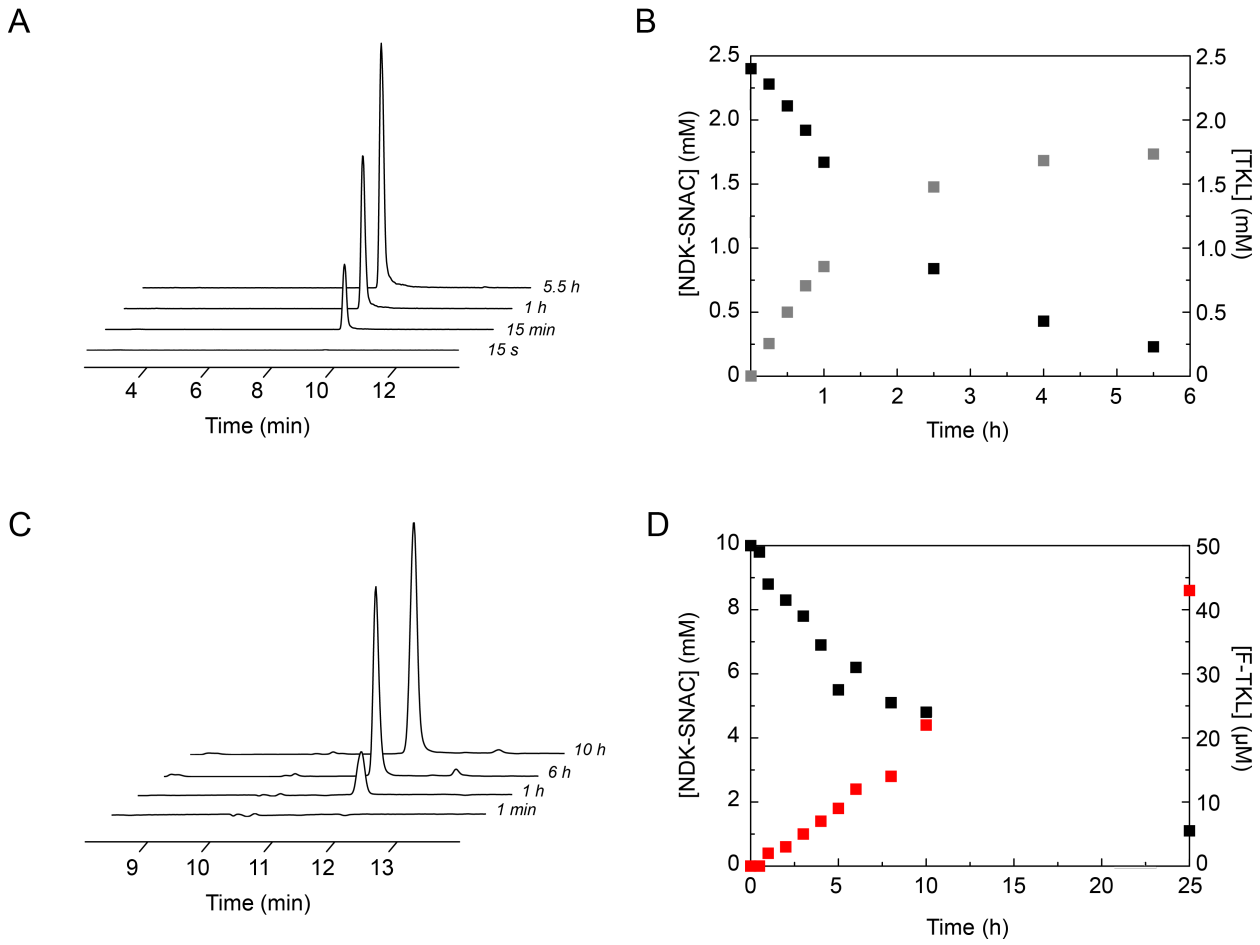


Figure 2.18. *Time-course for TKL and F-TKL formation by DEBSMod6+TE with substrate regeneration. (A) LC/MS traces monitoring TKL formation (m/z 169) from 2.5 mM NDK- SNAC. (B) Plot of NDK-SNAC and TKL concentrations. Initial rate: 1.5 min^{-1}. (C) LC/MS traces monitoring F-TKL formation (m/z 173) from 10 mM NDK-SNAC. (D) Plot of NDK-SNAC and F-TKL concentrations. Initial rate: 0.14 h^{-1}. (■, NDK-SNAC; ■, TKL; ■, F-TKL)*

We found that DEBS$_{Mod6}$+TE is also able to accept the fluorinated monomer to catalyze chain extension to form the 2-fluoro-2-desmethyltriketide lactone (F-TKL) and incorporate fluorine into the polyketide backbone (*Figure 2.11,* ***1-4***). The identity of the F-TKL was established by comparison to an authentic synthetic standard by reverse-phase HPLC monitored by ESI-MS and further confirmed by characterization of the isolated compound by high resolution MS, GC-MS, and ^{19}F NMR spectroscopy (*Figure 2.13-2.16*). Although the 2-*S* keto tautomer is generated in >94% diastereomeric excess (*Figure 2.15, Figure 2.17*), this ratio appears to be set by the compound's stereoelectronic factors rather than by the stereochemical preferences of DEBS$_{Mod6}$+TE, since the F-TKL is fully enolized in aqueous solution. The F-TKL can also be produced directly from fluoroacetate using the AckA–Pta/ACCase activation system in either a telescoped (*Figure 2.11,* ***5-6***) or single-pot reaction (*Figure 2.11,* ***7-8***) with DEBS$_{Mod6}$+TE at a similar yield to the MatB reaction, which allows us to connect fluorinated polyketide production directly to the biosynthetically available fluorinated building block (*Figure 2.1A, Scheme 2.1*).

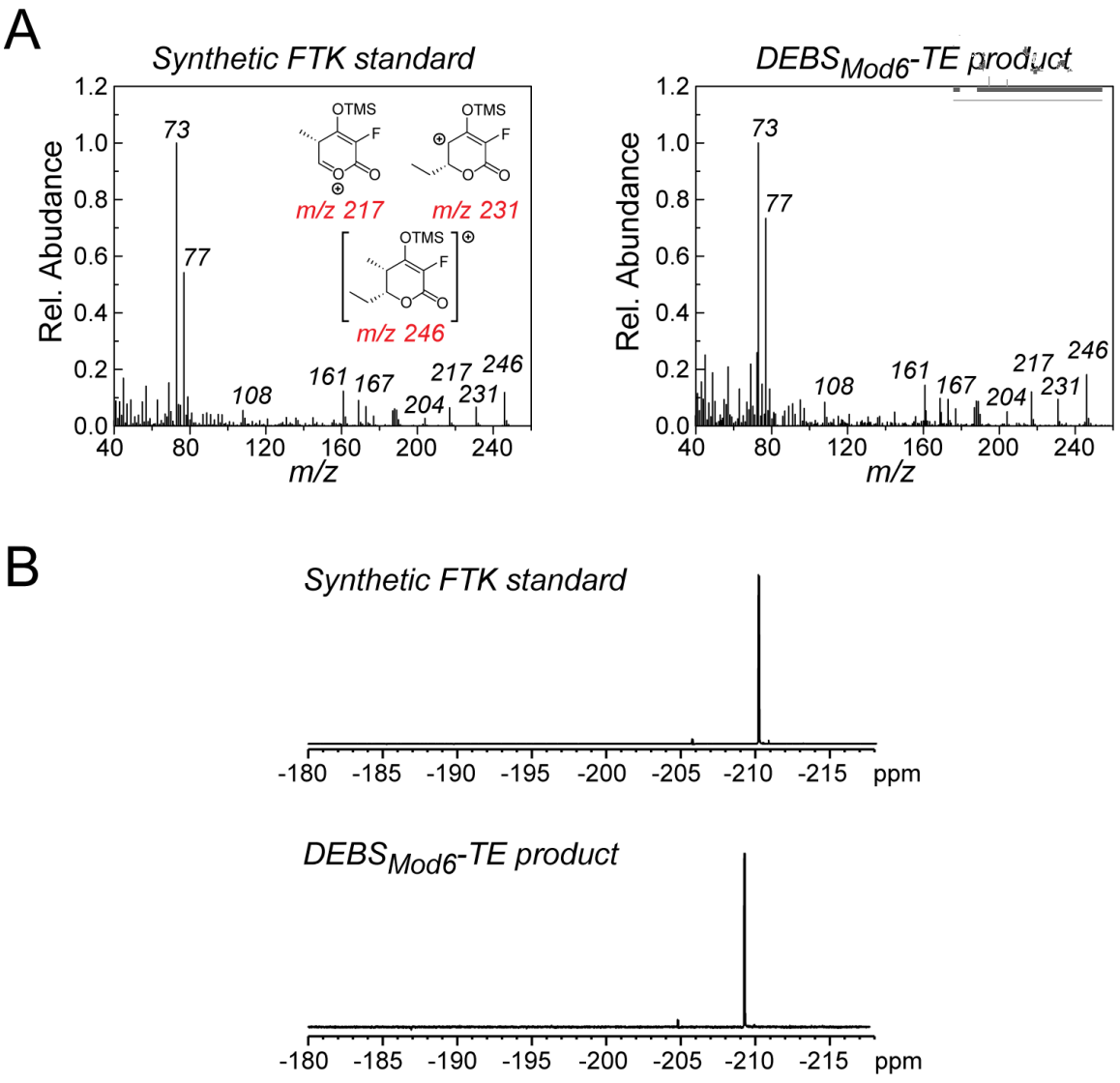


Figure 2.16. *GC-MS and ^{19}F NMR analysis of F-TKL. (A) Comparison of EI mass spectra of the standard (t_R = 8.51 min) compared to the enzymatic product (t_R = 8.56 min). (B) Comparison of ^{19}F NMR spectra in $CDCl_3$. The keto form is dominant at this concentration.*

In contrast to the chain extension reaction catalyzed by NphT7, $DEBS_{Mod6}$+TE does not incorporate the fluorinated extender unit into triketide lactone product as efficiently as its native methylmalonyl-CoA extender (*Figure 2.18*). Preliminary studies indicate that the reduced efficiency of $DEBS_{Mod6}$+TE with the fluorinated extender is not due to covalent inactivation of the enzyme (*Figure 2.19*), but rather to the more complex biochemistry of polyketide synthases with regard to monomer selection [*46*]. In fact, the acyltransferase domains of DEBS have been shown to be extremely selective for methylmalonyl-CoA, exhibiting a 100- to 1000-fold selectivity for their native substrate over malonyl-CoA [*47*]. The molecular mechanism underlying this selectivity remains an outstanding question, however, hydrolysis of incorrect extender units has been implicated as being involved [*46*]. Indeed, we observe extender unit hydrolysis even for the native substrate at a rate comparable to that of product formation (*Table 2.1*). This hydrolysis appears to limit fluoromalonyl-CoA incorporation, since substantially less F-TKL is observed using fluoromalonyl-CoA in the absence of MatB and ATP, and fluoromalonate remains the major

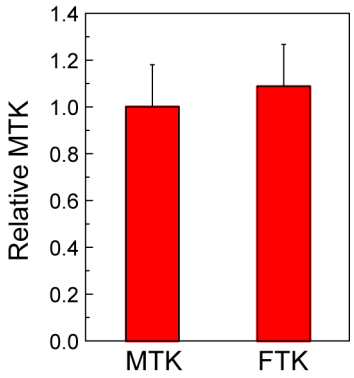


Figure 2.19. *Test for covalent inhibition of $DEBS_{Mod6}$+TE by fluoromalonyl-CoA. $DEBS_{Mod6}$+TE was incubated for 18 h in a F-TKL or TKL reaction. The enzyme was then isolated by Sephadex G-25 and tested for its ability to produce TKL.*

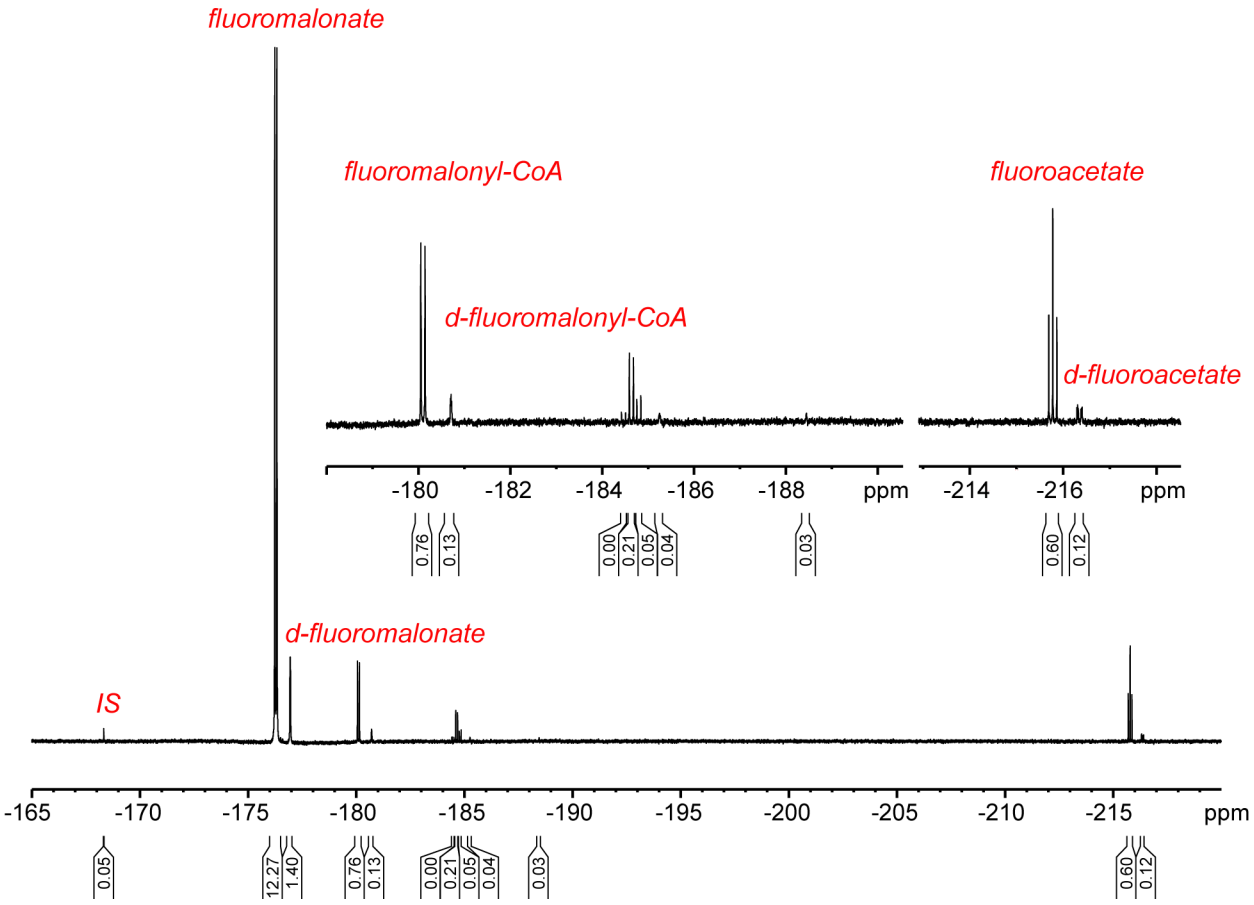


Figure 2.20. *^{19}F NMR analysis of F-TKL forming reaction with $DEBS_{Mod6}$TE. ^{19}F NMR analysis of the reaction mixture of $DEBS_{Mod6}$+TE with MatB indicates that the major pathway for loss of fluoromalonyl-CoA appears to be hydrolysis rather than unproductive decarboxylation. In addition, no detectable defluorination was observed. (IS, 5-fluorouracil, 50 μM)*

organofluorine species even in their presence (*Figure 2.20*). However, it is interesting to note that the fluoromalonyl-CoA extender is incorporated at higher efficiency than malonyl-CoA (R = H), which is reported to be naturally excluded by DEBS [47]. In fact, $DEBS_{Mod6}$+TE produces at least 10 times more F-TKL than H-TKL in a direct competition experiment with equimolar amounts (1 mM) of fluoromalonyl-CoA and malonyl-CoA (*Table 2.2*).

Table 2.1. Rates of acyl-CoA hydrolysis by $DEBS_{Mod6}$+TE. Steady-state hydrolysis rates were measured using 1 µM $DEBS_{Mod6}$+TE and 500 µM acyl-CoA. Values are reported as the mean ± s.d. (n = 4).

	v_0 (µM min^{-1})	Relative rate
Methylmalonyl-CoA	1.36 ± 0.05	1.0
Fluoromalonyl-CoA	3.5 ± 0.3	2.6
Malonyl-CoA	6.1 ± 0.4	4.5

To address the issue of site- or regioselective fluorine incorporation, we turned our attention to replacing the enzyme activity most likely associated with low fluoromalonyl extender incorporation, the AT. It was previously shown that an inactivated AT domain could be complemented *in trans* by another AT [*48*], and we hypothesized that a malonyl-specific AT might be more accommodating toward the fluoromalonyl extender than the methylmalonyl-specific ATs of DEBS. Therefore we set out to complement an AT-inactivated module with the *trans*-AT from the disorazole PKS, DszsAT [*49*, *50*]. When we prepared the catalytically compromised $DEBS_{Mod6}$+TE mutant (S2107A), we were surprised to find that F-TKL yield increased even in the absence of a complementing AT, while TKL yield dropped considerably as expected, but not to zero (*Figure 2.21*). Formation of these triketides must take place *via* an AT-independent mechanism that has not yet been characterized. We found that DszsAT did indeed accept fluoromalonyl-CoA and considerably enhanced F-TKL formation by the AT-null mutant (*Figure 2.21*).

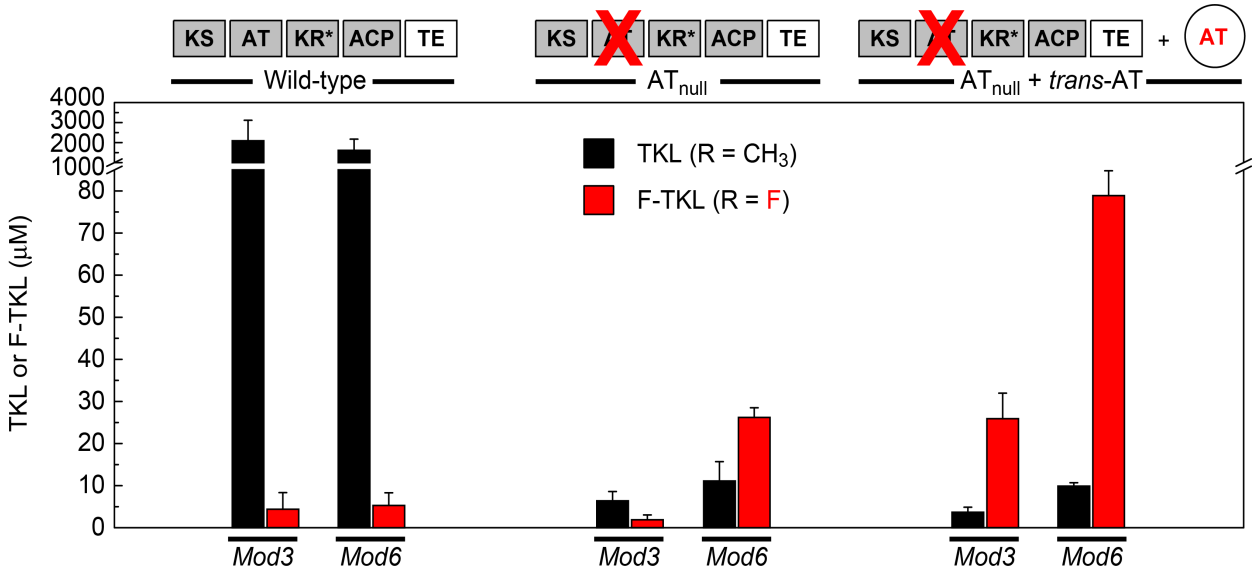


Figure 2.21. *Selectivity of $DEBS_{Mod6}$+TE and $DEBS_{Mod3}$+TE for the methylmalonyl-CoA versus fluoromalonyl-CoA extender unit, as monitored by TKL (m/z = 169) and F-TKL (m/z = 173) formation. Conditions include wild-type modules, AT^0 modules, and AT^0 modules in conjunction with the trans-AT from the disorazole PKS (DszsAT). Values are reported as the mean ± SD (n = 3). KR* denotes that the KR domain of Mod3 is inactive.*

Table 2.2. F-TKL and H-TKL production under competitive conditions. $DEBS_{Mod6}$+TE was incubated with equimolar amounts of malonyl-CoA and fluoromalonyl-CoA (1 mM). Without substrate regeneration, no detectable H-TKL was formed (<50 nM). MatB and regeneration enzymes were then included to amplify and quantify H-TKL formation. Values are reported as the mean ± s.d. (n = 3).

Condition	[F-TKL] (nM)	[H-TKL] (nM)	[F-TKL] / [H-TKL]
No regeneration	450 ± 60	< 50	> 9
MatB regeneration	7,390 ± 520	720 ± 50	10.3 ± 0.1

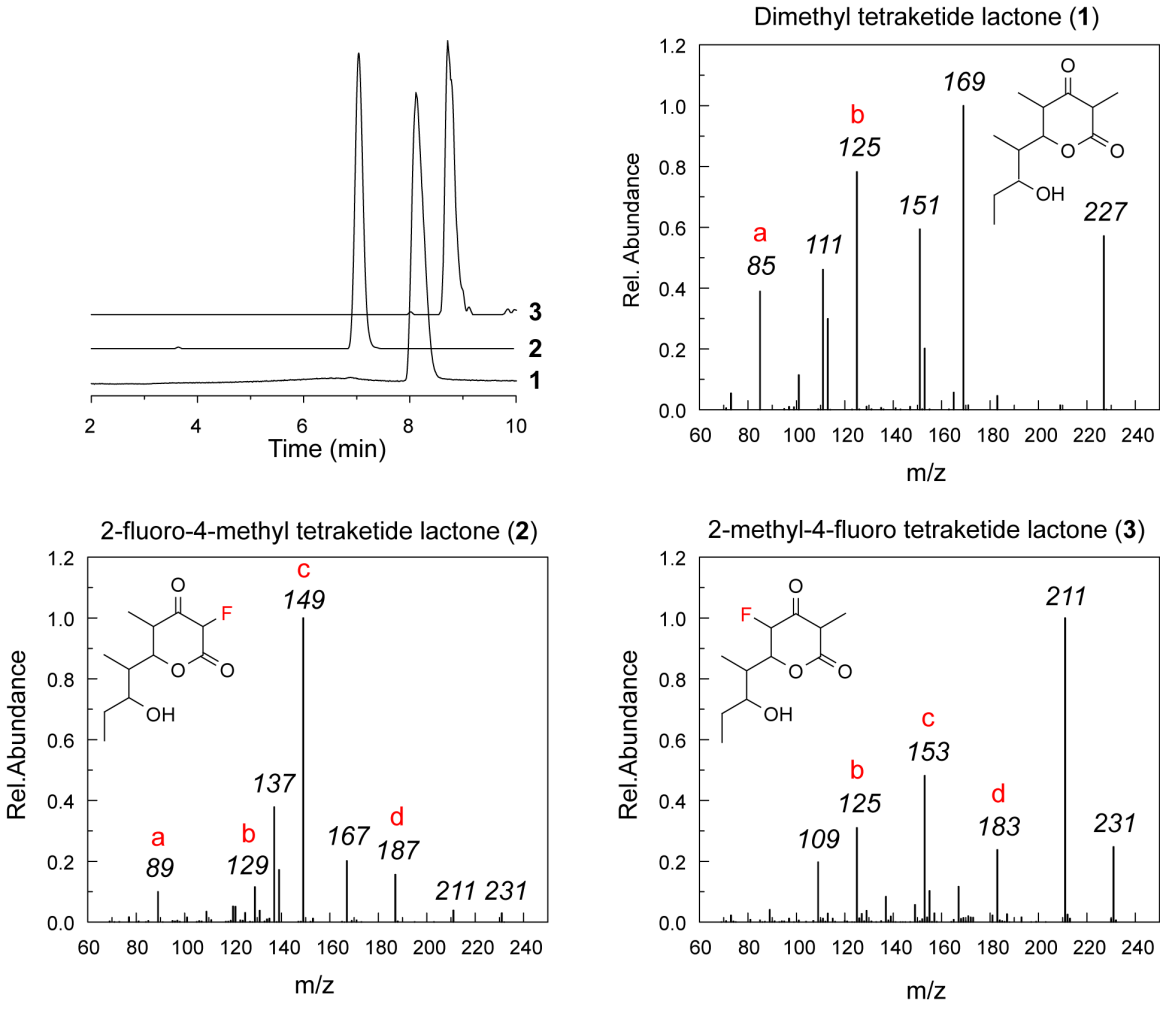


Figure 2.23. *Production of fluorinated tetraketide lactones. Production of 2,4-dimethyl tetraketide lactone (m/z-H: 227), 2-fluoro-4-methyl tetraketide lactone (m/z-H: 231), and 2-methyl-4-fluoro tetraketide lactone (m/z-H: 231). Chromatograms were collect by LC/MS by single ion monitoring in the negative channel. Mass spectra of the tetraketide lactones were obtained using LC/MS/MS on a LTQ-FT instrument in negative mode. Fragments that differ by 4 AMU, corresponding to a fluorine for methyl substitution, are labeled A – D.*

Using this approach, we began to explore the possibility of site-selective fluorine incorporation with a mini-PKS model system, consisting of $DEBS_{Mod2}$ and $DEBS_{Mod3}$+TE, that was designed to carry out two chain-extension reactions from the NDK-SNAC substrate [*25*]. Using the appropriate AT-null constructs, we were able to observe preferential production of either regioisomer of the fluoro-methyl tetraketide lactone (tetraKL) (*Figure 2.22*). The identity of the 2-fluoro-4-methyl tetraKL and 2-methyl-4-fluoro tetraKL was established by both high-resolution ESI-MS and LC-MS on the basis of their different retention times, as well as their mass fragmentation patterns, which are consistent with the incorporation of fluorine at the expected sites (*Figure 2.23*). These studies also indicate that further chain extension after fluorine insertion can be achieved and that fluorinated intermediates could potentially be tolerated in downstream reactions. This observation is consistent with previous work that has shown that intermediates with non-native substituents, including fluorine, can be extended and tailored to the final structure [*3, 14, 17-18, 51*] and gives promise that larger fluorinated polyketide targets may be accessible through this approach.

The observed selectivity for fluoromalonyl- over malonyl-CoA extender units suggests that polyketide chain extension reactions with fluoromalonyl-CoA could possibly be catalyzed *in vivo* in *E. coli*, which contains a sizable malonyl-CoA pool (~35 mM) [*52*] but almost no methylmalonyl-CoA [*21, 53*]. We carried out preliminary ^{19}F NMR studies of cells expressing MatB, NphT7, and PhaB and fed with nontoxic levels of fluoromalonate. Analysis of the media and cell extracts indicated that flux through fluoromalonyl-CoA could reach 100 μM to 1 mM, which is sufficient for use by PKSs in live cells (data not shown).

Next, we tested the ability of $DEBS_{Mod6}$+TE to catalyze chain elongation in cell lysates prepared from *E. coli* BAP1 coexpressing $DEBS_{Mod6}$+TE and MatB. Under these conditions, F-TKL is produced with no observable H-TKL (< 1 μM) upon addition of only NDK-SNAC, fluoromalonate, CoA, ATP, and the ATP regeneration system (*Figure 2.24A*). Negative controls with either no $DEBS_{Mod6}$+TE/MatB expressed or no NDK-SNAC substrate show no production of F-TKL (*Figure 2.24A*). These results demonstrate that the level of expression of the

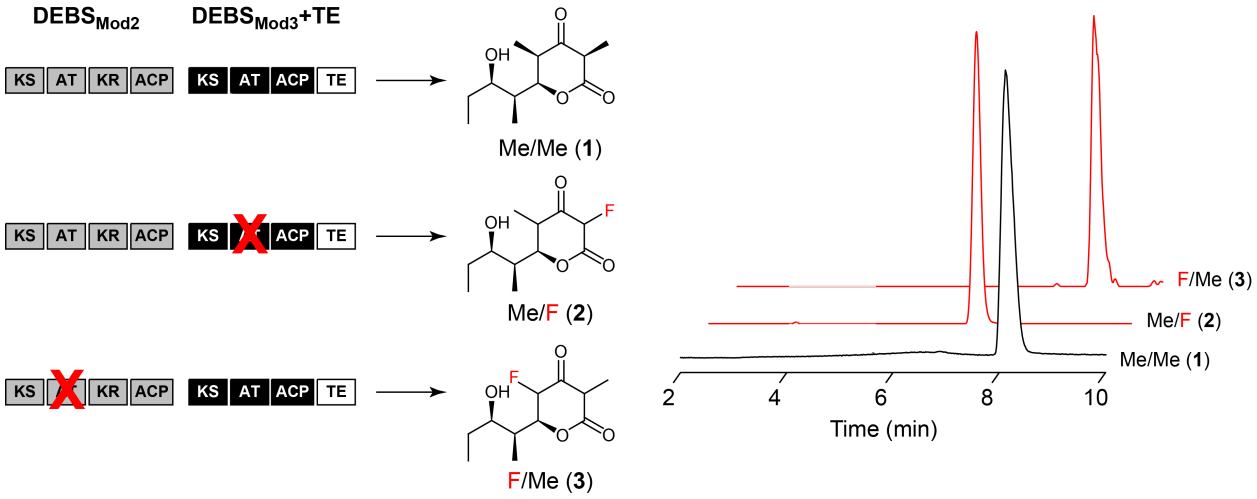


***Figure 2.22**. LC-MS traces showing regioselective tetraketide lactone formation using the DEBS mini-PKS consisting of $DEBS_{Mod2}$ and $DEBS_{Mod3}$+TE (Me/Me, 2-methyl-4-methyl-tetraketide lactone, m/z = 227; Me/F, 2-fluoro-4- methyl-tetraketide lactone, m/z = 231; F/Me, 2-methyl-4-fluoro-tetraketide lactone, m/z = 231). Me/Me was produced using $DEBS_{Mod2}$/$DEBS_{Mod3}$+TE and methyl-malonate (1). Me/F was produced using $DEBS_{Mod2}$/$DEBS_{Mod3}AT^{0}$+TE, DszsAT, methylmalonyl-CoA, and fluoromalonate (2). F/Me was produced using $DEBS_{Mod2}AT^{0}$/$DEBS_{Mod3}$+TE, methylmalonyl-CoA, and fluoromalonate (3). Data are normalized with respect to the Me/Me peak. All reactions contained MatB and the ATP regeneration system.*

$DEBS_{Mod6}$+TE and MatB enzymes is sufficient for the incorporation of the fluorinated extender unit. They also further imply that fluorine could be introduced into the polyketide backbone inside living cells, which are capable of generating ATP for monomer activation through normal metabolic processes.

Therefore, we cultured *E. coli* BAP1 coexpressing $DEBS_{Mod6}$+TE and MatB and harvested the intact cells after induction. These cells were then fed with the fluoromalonate precursor, which resulted in the production of F-TKL upon addition of NDK-SNAC (*Figure 2.11*, *Figure 2.24*). The identity of the F-TKL under these conditions was established by LC-MS, co-injection with an authentic standard, and high-resolution MS. Moreover, F-TKL can also be produced directly in cell culture with the simple addition of a mixture of both substrates to the media after induction of $DEBS_{Mod6}$+TE and MatB (*Figure 2.24C*), though in lower yield. Taken together, these studies show that the natural selectivity of the polyketide synthase allows for the site-selective introduction of fluorine over hydrogen into the polyketide backbone inside living cells.

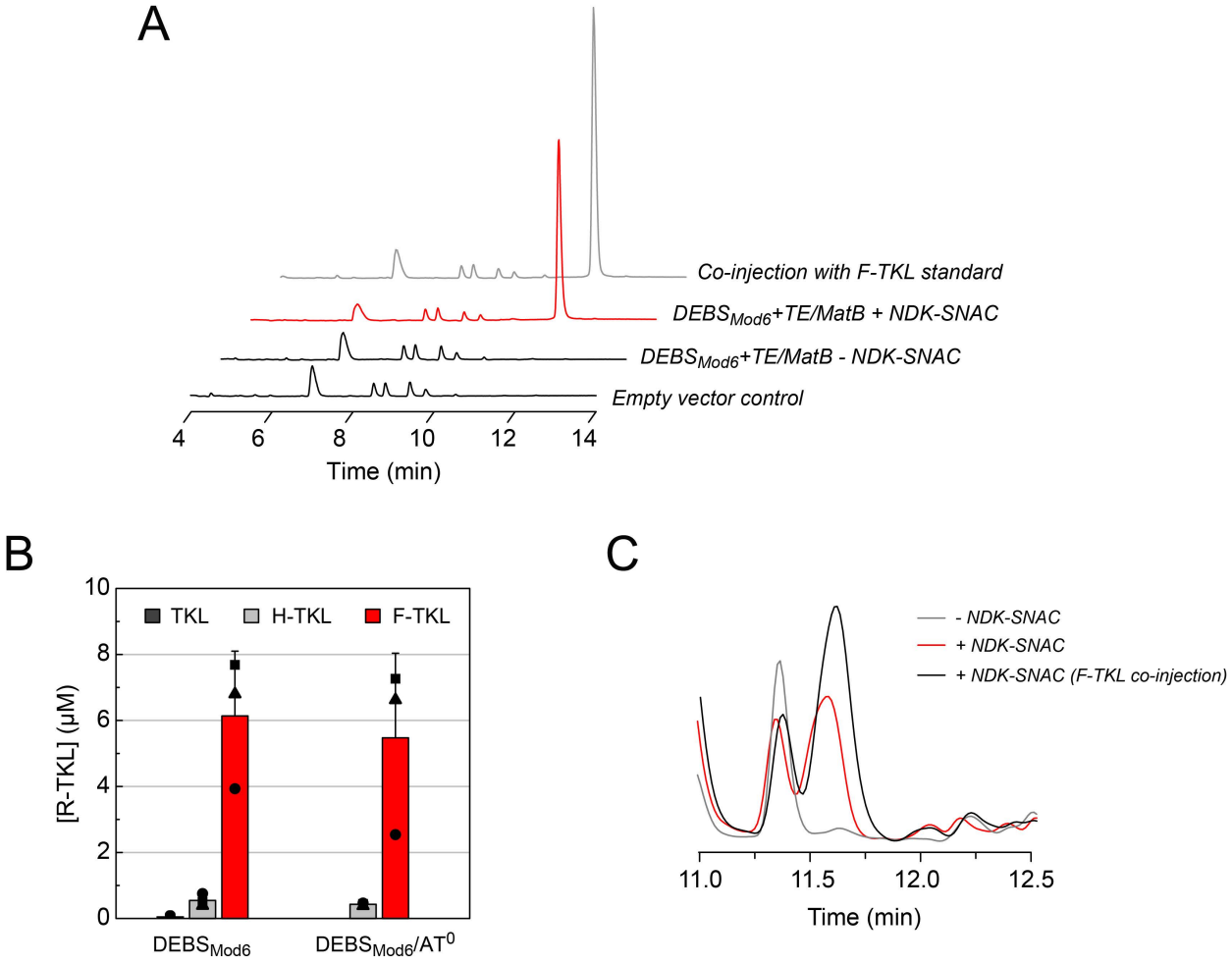


Figure 2.24. *F-TKL production in vivo. (A) LC/MS traces showing F-TKL formation (m/z 173) in E. coli cell lysate. (B) In vivo selectivity data showing F-TKL production compared to H-TKL and TKL in fluoromalonate-fed E. coli resting cells expressing either $DEBS_{Mod6}$+TE or $DEBS_{Mod6}$+TE/AT^0 and MatB. Bars represent mean ± s.d. (n = 3) with individual samples marked (■▲●). (C) LC/MS traces showing F-TKL formation (m/z 173) by E. coli cell culture upon feeding with NDK-SNAC.*

2.4 Conclusions

We have demonstrated that we can expand the fluorine chemistry of living systems with engineered pathways that use simple biogenic organofluorine building blocks to construct more complex fluorinated small molecule targets. To our knowledge, nature has not devised a strategy for generation of the fluoromalonyl building block or its utilization as a metabolic precursor for biosynthesis of complex natural products. Here, we show that the fluorinated extender unit, fluoromalonyl-CoA, can be produced enzymatically and used for downstream chain elongation reactions by polyketide synthases and that, at least in one case, subsequent α-ketoreduction is not inhibited by incorporation of a fluoroacetyl unit. Because of the modular nature of the biosynthetic pathways used to produce polyketides and related acetate-derived natural products, these findings open the door to general strategies for exploring the fluorine synthetic biology of complex natural products.

2.5 References

1. Ro, D. K., Paradise, E. M., Ouellet, M. O., Fisher, K. J., Newman, K. L., Ndungu, J. M., Ho, K. A., Eachus, R. A., Ham, T., Kirby, J., Chang, M. C. Y., Withers, S. T., Shiba, Y., Sarpong, R., Keasling, J. D., Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **2006,** *440* 940-943.

2. Atsumi, S., Hanai, T., Liao, J. C., Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **2008,** *451* (7174) 86-89.

3. Cane, D. E., Walsh, C. T., Khosla, C., Harnessing the biosynthetic code: Combinations, permutations, and mutations. *Science* **1998,** *282* (5386) 63-68.

4. Weeks, A. M., Chang, M. C. Y., Constructing *de novo* biosynthetic pathways for chemical synthesis inside living cells. *Biochemistry* **2011,** *50* (24) 5404-5418.

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

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

7. Furuya, T., Kamlet, A. S., Ritter, T., Catalysis for fluorination and trifluoromethylation. *Nature* **2011,** *473* (7348) 470-477.

8. Ball, N. D., Sanford, M. S., Synthesis and reactivity of a mono-σ-aryl palladium(IV) fluoride complex. *J. Am. Chem. Soc.* **2009,** *131* (11) 3796-3797.

9. Watson, D. A., Su, M., Teverovskiy, G., Zhang, Y., Garcia-Fortanet, J., Kinzel, T., Buchwald, S. L., Formation of ArF from LPdAr(F): Catalytic conversion of aryl triflates to aryl fluorides. *Science* **2009,** *325* (5948) 1661-1664.

10. Rauniyar, V., Lackner, A. D., Hamilton, G. L., Toste, F. D., Asymmetric electrophilic fluorination using an anionic chiral phase-transfer catalyst. *Science* **2011,** *334* (6063) 1681-1684.

11. Lee, E., Kamlet, A. S., Powers, D. C., Neumann, C. N., Boursalian, G. B., Furuya, T., Choi, D. C., Hooker, J. M., Ritter, T., A fluoride-derived electrophilic late-stage fluorination reagent for pet imaging. *Science* **2011,** *334* (6056) 639-642.

12. Dong, C., Huang, F., Deng, H., Schaffrath, C., Spencer, J. B., O'Hagan, D., Naismith, J. H., Crystal structure and mechanism of a bacterial fluorinating enzyme. *Nature* **2004,** *427* (6974) 561-565.

13. O'Hagan, D., Recent developments on the fluorinase from *Streptomyces cattleya*. *J. Fluorine Chem.* **2006,** *127* (11) 1479-1483.

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

15. Croteau, R., Kutchan, T. M., Lewis, N. G., Natural products (secondary metabolites). In *Biochemistry and molecular biology of plants*, Buchanan, R. B., Gruissem, W., Jones, R., ASPB: Rockville, MD, 2000, 1250-1318.

16. Chan, Y. A., Podevels, A. M., Kevany, B. M., Thomas, M. G., Biosynthesis of polyketide synthase extender units. *Nat. Prod. Rep.* **2009,** *26* (1) 90-114.

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

18. Mo, S., Kim, D. H., Lee, J. H., Park, J. W., Basnet, D. B., Ban, Y. H., Yoo, Y. J., Chen, S.-w., Park, S. R., Choi, E. A., Kim, E., Jin, Y.-Y., Lee, S.-K., Park, J. Y., Liu, Y., Lee, M. O., Lee, K. S., Kim, S. J., Kim, D., Park, B. C., Lee, S.-g., Kwon, H. J., Suh, J.-W., Moore, B. S., Lim, S.-K., Yoon, Y. J., Biosynthesis of the allylmalonyl-CoA extender unit for the FK506 polyketide synthase proceeds through a dedicated polyketide synthase and facilitates the mutasynthesis of analogues. *J. Am. Chem. Soc.* **2010,** *133* (4) 976-985.

19. Wilson, M. C., Moore, B. S., Beyond ethylmalonyl-coa: The functional role of crotonyl-coa carboxylase/reductase homologs in expanding polyketide diversity. *Nat. Prod. Rep.* **2012,** *29* (1) 72-86.

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

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

22. Rouillard, J. M., Lee, W., Truan, G., Gao, X., Zhou, X., Gulari, E., Gene2oligo: Oligonucleotide design for *in vitro* gene synthesis. *Nucleic Acids Res.* **2004,** *32* (Web Server issue) W176-W180.

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

24. 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* **2009,** *49* (1) 95-102.

25. Tsuji, S. Y., Cane, D. E., Khosla, C., Selective protein-protein interactions direct channeling of intermediates between polyketide synthase modules. *Biochemistry* **2001,** *40* (8) 2326-2331.

26. Yu, X., Liu, T., Zhu, F. & Khosla, C., In vitro reconstitution and steady-state analysis of the fatty acid synthase from Escherichia coli. *Proceedings of the National Academy of Sciences* **2011**, *108***,** 18643–18648.

27. Wu, N., Kudo, F., Cane, D. E., Khosla, C., Analysis of the molecular recognition features of individual modules derived from the erythromycin polyketide synthase. *J. Am. Chem. Soc.* **2000,** *122* (20) 4847-4852.

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

29. Kumar, P., Khosla, C., Tang, Y., Manipulation and analysis of polyketide synthases. *Method. Enzymol.* **2004,** 269-293.

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

31. Williamson, J. R., Corkey, B. E., Assays of intermediates of the citric acid cycle and related compounds by fluorometric enzyme methods. *Method. Enzymol.* **1969,** 434-513.

32. Huang, F., Haydock, S. F., Spiteller, D., Mironenko, T., Li, T.-L., O'Hagan, D., Leadlay, P. F., Spencer, J. B., The gene cluster for fluorometabolite biosynthesis in *Streptomyces cattleya*: A thioesterase confers resistance to fluoroacetyl-coenzyme A. *Chem. Biol.* **2006,** *13* (5) 475-484.

33. Cane, D. E., Tan, W., Ott, W. R., Nargenicin biosynthesis. Incorporation of polyketide chain elongation intermediates and support for a proposed intramolecular diels-alder cyclization. *J. Am. Chem.* **1993,** *115* (2) 527-535.

34. Akoka, S., Barantin, L., Trierweiler, M., Concentration measurement by proton NMR using the ERETIC method. *Anal. Chem.* **1999,** *71* (13) 2554-2557.

35. Hinterding, K., Singhanat, S., Oberer, L., Stereoselective synthesis of polyketide fragments using a novel intramolecular Claisen-like condensation/reduction sequence. *Tetrahedron lett.* **2001,** *42* (48) 8463-8465.

36. Luo, G., Pieper, R., Rosa, A., Khosla, C., Cane, D. E., Erythromycin biosynthesis: Exploiting the catalytic versatility of the modular polyketide synthase. *Bioorg. Med. Chem.* **1996,** 4 (7) 995-999.

37. Brown, T. D. K., Jones-Mortimer, M. C., Kornberg, H. L., The enzymic interconversion of acetate and acetyl-coenzyme a in *Escherichia coli*. *J. Gen. Microbiol.* **1977,** *102* (2) 327-336.

38. Walker, M. C., Wen, M., Weeks, A. M., Chang, M. C. Y., Temporal and fluoride control of secondary metabolism regulates cellular organofluorine biosynthesis. *ACS Chem. Biol.* **2012**, *7* (9) 1576–1585.

39. Hughes, A. J., Keatinge-Clay, A., Enzymatic extender unit generation for *in vitro* polyketide synthase reactions: Structural and functional showcasing of *Streptomyces coelicolor* MatB. *Chem. Biol.* **2011**, *18* (2) 165-176.

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

41. 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* (10) 1145-1153.

42. Kelley, L. A., Sternberg, M. J. E., Protein structure prediction on the web: A case study using the PHYRE server. *Nat. Prot.* **2009,** *4* (3) 363-371.

43. Tang, Y., Kim, C.-Y., Mathews, I. I., Cane, D. E., Khosla, C., The 2.7 Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Nat. Ac. Sc. U.S.A.* **2006,** *103* (30) 11124-11129.

44. Mohanta, P. K., Davis, T. A., Gooch, J. R., Flowers, R. A., Chelation-controlled diastereoselective reduction of α-fluoroketones. *J. Am. Chem. Soc.* **2005,** *127* (34) 11896-11897.

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

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

47. Liou, G. F., Lau, J., Cane, D. E., Khosla, C., Quantitative analysis of loading and extender acyltransferases of modular polyketide synthases. *Biochemistry* **2002,** *42* (1) 200-207.

48. Kumar P, Koppisch AT, Cane DE, 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* (47) 14307-14312.

49. 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.

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

51. Goss, R. J. M., Lanceron, S., Deb Roy, A., Sprague, S., Nur-E-Alam, M., Hughes, D. L., Wilkinson, B., and Moss, S. J., An expeditious route to fluorinated rapamycin analogues by utilising mutasynthesis. *ChemBioChem* **2010**, *11*, 698–702.

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

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

Chapter 3: *Development of a robust system for organofluorine biosynthesis* in vivo

Portions of this work were performed in collaboration with the following persons:

These studies were carried out in collaboration with Thomas Privalsky. Initial investigations into fluorinated poly(hydroxybutyrate) synthesis were performed by Mark Walker, PhD. Polymer properties were characterized by Angelika Neitzel (University of Minnesota). NphT7 expression vectors were cloned by Michael Blaisse.

3.1. Introduction

Natural products and organofluorines have each been important to the development of small molecule therapeutics, but they rarely intersect in structural space. Fluorine's small size and high electronegativity confer powerful effects on small molecule properties, and drug fluorination can improve bioavailability, reduce clearance, alter p*K*a and lipophilicity, block metabolism, or increase potency [*1,2*]. As a result, fluorine is now found in 20−30% of therapeutics and agrochemicals, and methods for generating diverse organofluorine structures continue to serve a growing number of applications [*5*]. Despite these advantages, fluorine is rarely observed in naturally occurring compounds. Using synthetic biology approaches to incorporate fluorine into natural product biosynthesis would combine some of the strengths of organic and biological chemistry and expand the structural space of natural products, which have provided the source for many human therapeutics.

We have previously shown that polyketide synthesis can accommodate fluorinated monomers as chain extenders *in vitro*, potentially leading to an enormous increase in natural product analogue diversity and new bioactive compounds [*3*]. However, polyketide synthesis, like all biosynthesis, is valuable mostly because it occurs *in vivo* using self-assembling, self-replicating catalysts, making it scalable and relatively inexpensive. By contrast, purified enzymes are not a practical commercial source of natural products. Several additional challenges arise from the transition between an *in vitro* and an *in vivo* pathway (*Figure 3.1*). Fluorinated building blocks must be either taken up by the cell from an exogenous supply, or biosynthesized from fluoride as in the natural *Streptomyces cattleya* pathway. The chemistry to convert building blocks to activated polyketide monomers must occur within the context of the thousands of reactions of native cellular metabolism. Likewise, pathway reactions must now select fluorinated substrates out of the pool of all other endogenous metabolites. Diversion of fluorinated intermediates from the synthetic pathway must be avoided, since this could waste flux by

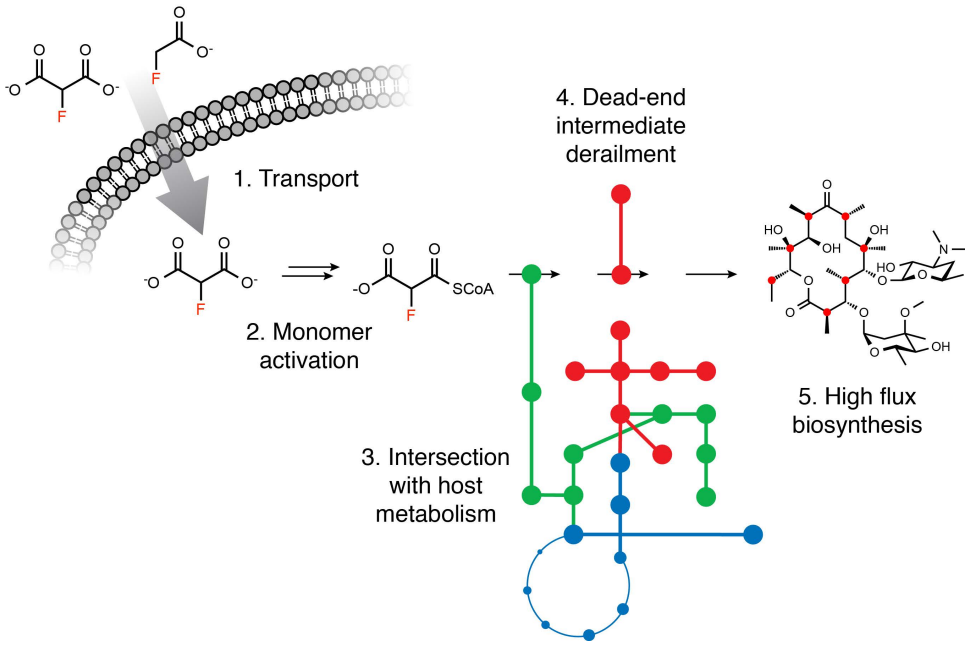


Figure 3.1. *Challenges associated with in vivo organofluorine metabolism.*

producing dead-end metabolites that cannot be reintroduced to the pathway, or interfere with background host metabolism resulting in toxicity. Finally, the synthetic pathway and fluorinated building block activation must be balanced with native energy and carbon resources needed to maintain the cell, produce necessary enzymes, and construct the non-fluorinated portions of the target natural product.

To investigate these issues, we aimed to construct a synthetic organofluorine metabolic pathway that would enable both high-flux biosynthesis and *in vivo* screening and evolution of biosynthesis-associated enzymes. Although native polyketide synthase modules can accept fluorinated monomers, their low activity and selectivity towards fluorine, as well as the necessity of supplying a synthetic diketide precursor to initiate chain extension, lead us to look elsewhere for a model system. We had previously characterized two enzymes, the acetoacetyl-CoA synthase NphT7 [*6*], which is essentially a stand-alone polyketide ketosynthase domain, and the acetoacetyl-CoA reductase PhaB, which mimics a polyketide ketoreductase. These enzymes effectively carry out a single round of polyketide chain extension and ketoreduction but use CoA-linked substrates, side-stepping any issues with extender loading onto an acyl carrier protein (Chapter 2). We planned to use this high-efficiency pathway to establish that a sufficient intracellular fluoromalonyl-CoA pool can be made available, that potentially labile intermediates and derailment products do not divert flux from the desired product or cause toxicity for the host, and that titers of the pathway product can be achieved that are relevant to natural product biosynthesis.

We selected *E. coli* as a chassis. Although it is not a host that is particularly amenable to polyketide synthase expression (with some exceptions [*7*]), its ease of genetic manipulation, fast growth and relatively well-characterized metabolism make it a useful platform. We hope to use *E. coli* to find generalizable principles for organofluorine metabolism that we can later transfer to natural product producing organisms such *Streptomyces* or *Saccharopolyspora*.

We selected the relatively advanced fluoromalonate as the organofluorine input for our pathway. It has the advantage of being activated in one step by an enzyme, malonate:CoA ligase, that is not native to *E. coli*. Malonyl-CoA is a committed metabolite in *E. coli*, used only for fatty acid synthesis, and preliminary investigations have shown that fluoromalonyl-CoA does not enter that pathway (Xingye Yu, Ben Thuronyi, Chaitan Khosla & Michelle Chang, unpublished results), suggesting that it will be orthogonal to host metabolism and participate only in the synthetic pathway.

Perhaps most importantly, fluoromalonate use avoids the need for fluoroacetyl-CoA as a pathway intermediate. Formation of this metabolite is key to fluoroacetate toxicity, since its entry into the tricarboxylic acid cycle leads to aconitase inhibition [*8*]. Use of fluoroacetate and fluoroacetyl-CoA is an achievable goal, but will likely require acyl-CoA carboxylase engineering to ensure that fluoroacetyl-CoA is carboxylated efficiently and cannot build up. Alternatively, careful growth-phase-based regulation of fluoroacetyl-CoA use, a strategy used in the native fluoroacetate producer *S. cattleya* [*9*], could avoid concurrent TCA cycle and fluoroacetyl-CoA use and therefore prevent toxicity. These approaches could eventually allow relatively inexpensive fluoroacetate, or extremely inexpensive fluoride (via the *S. cattleya* fluorinase pathway [*10*]), to be used as the input.

The reduced diketidyl-CoA produced by our pathway is an α-fluorinated derivative of 3-hydroxybutyryl-CoA, which is the primary feedstock for a class of biopolymers called

poly(hydroxyalkanoates) (PHAs) [*11-13*]. These polyesters can accumulate at high levels (up to ~80% of dry cell weight) in native producers or transgenic *E. coli*. When an appropriate monomer composition is used, PHAs have materials properties similar to poly(propylene), and unlike petroleum-derived polymers they can be easily produced by fermentation of renewable feedstocks [*12*]. Naturally occurring PHAs derive their monomers either from fatty acid synthesis intermediates or from short-chain thiolases that are not known to produce 2-substituted acyl-CoAs. PHA biosynthesis can be engineered to derive monomers from exogenous fatty acids, but fatty acid catabolism is also not known to efficiently accept internally branched fatty acid chains that would result in 2-substituted 3-hydroxyacyl CoAs. Therefore, although the diversity of PHA monomers is enormous, it draws heavily on variable 3-position substituents; polymers with 2-substituents are an underexplored class. PHAs with some 2-methyl substituents have been reported, but incorporation by native producers is usually very low [*14-17*] except in one notable case, where fed 3-hydroxypivalic acid could be homopolymerized under specific growth conditions [*18*].

NphT7 has previously been shown to be capable of providing 3-hydroxybutyryl-CoA monomers for poly(hydroxybutyrate) [*19*]. We wondered whether, as we had found to be the case for polyketide biosynthesis, a 2-fluoro substituent might be acceptable to wild-type PHA synthases due to its small size, while simultaneously creating significant changes in PHA materials properties because of its electronegativity and unique physical properties. Because α-fluorinated thioesters are considerably more labile to hydrolysis and presumably also to retro-aldol decomposition than most other analogues would be, and off-pathway intermediates can easily be detected and characterized using ^{19}F NMR, the organofluorine-based pathway would also serve as a stringent test case for identifying side-reactions and bottlenecks that might impede PHA derivative engineering. Therefore the organofluorine-based pathway could serve as a starting point for enzyme and pathway modifications to develop diverse 2-substituted PHAs. Accordingly, we investigated pathways that incorporated a poly(hydroxyalkanoate) synthase to attempt to produce poly(3-hydroxybutyrate-*co*-2-fluoro-3-hydroxybutyrate).

3.2 Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB) and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Carbenicillin (Cb), isopropyl-β-D-thiogalactopyranoside (IPTG), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride, dithiothreitol (DTT), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile, dichloromethane, ethyl acetate and ethylene diamine tetraacetic acid disodium dihydrate (EDTA) were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A trilithium salt (CoA), acetyl-CoA, malonic acid, methylmalonic acid, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenolpyruvate (PEP), adenosine triphosphate sodium salt (ATP), myokinase, pyruvate kinase, β-mercaptoethanol, sodium phosphate dibasic heptahydrate and N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED) were purchased from Sigma-Aldrich (St. Louis, MO). Diethylfluoromalonate was purchased from Sigma-Aldrich (St. Louis, MO) or from Oakwood Chemical (West Columbia, SC). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate and ammonium persulfate were purchased from Bio-Rad Laboratories (Hercules,

CA). Restriction enzymes, T4 DNA ligase, Phusion DNA polymerase, T5 exonuclease, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) and Platinum Taq High-Fidelity polymerase (Pt Taq HF) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in 10 mM Tris-HCl, pH 8.5, and stored at 4°C. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). 10,000 MWCO regenerated cellulose ultrafiltration membranes were purchased from EMD Millipore (Billerica, MA). The Rezex ROA-Organic Acid H+ (8%) HPLC column was purchased from Phenomenex (Torrance, CA). Deuterium oxide and chloroform-d were purchased from Cambridge Isotope Laboratories (Andover, MA).

Bacterial strains. *E. coli* DH10B-T1^R was used for DNA construction. Except where otherwise noted, production and polymer formation experiments were conducted in BL21(de3)T1^R. Strains DH5α(de3), BW25113, W3110(de3), DH10B-T1^R and DH1(de3) were also tested in production experiments. Strains from the Keio collection [*20*] derived from BW25113 were used to test the influence of native genes on the pathway.

Gene and plasmid construction. Plasmids were constructed using the isothermal Gibson assembly method [*21*]. All PCR amplifications were carried out with Phusion or Platinum Taq High Fidelity DNA polymerases. For amplification of GC-rich sequences, PCR reactions were supplemented with DMSO (5%). All constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA).

The synthetic gene encoding *Rhodopseudomonas palustris* MatB was reverse translated using *E. coli* class II codon usage and synthesized using PCR assembly. Gene2Oligo was used to convert the gene sequence into primer sets using default optimization settings (*Appendix Table 2*) [*22*]. To assemble the synthetic gene, each primer was added at a final concentration of 1 μM to the first PCR reaction (50 μL) containing 1 × Pt Taq HF buffer (20 mM Tris-HCl, 50 mM KCl, pH 8.4), $MgSO_4$ (1.5 mM), dNTPs (250 μM each), and Pt Taq HF (5 U). The following thermocycler program was used for the first assembly reaction: 95ºC for 5 min; 95ºC for 30 s; 55ºC for 2 min; 72ºC for 10 s; 40 cycles of 95ºC for 15 s, 55ºC for 30 s, 72ºC for 20 s plus 3 s/cycle; these cycles were followed by a final incubation at 72ºC for 5 min. The second assembly reaction (50 μL) contained 16.5 μL of the unpurified first PCR reaction with standard reagents for Pt Taq HF. The thermocycler program for the second PCR was: 95ºC for 30 s; 55ºC for 2 min; 72ºC for 10 s; 40 cycles of 95ºC for 15 s, 55ºC for 30 s, 72ºC for 80 s; these cycles were followed by a final incubation at 72ºC for 10 min. The second PCR reaction (16.5 μL) was transferred again into fresh reagents and run using the same program. Following gene construction, the DNA smear at the appropriate size (1.5 kb) was gel purified and used as a template for the rescue PCR (50 μL) with Pt Taq HF and rescue primers r_palu matB F1/R1 under standard conditions. The resulting rescue product was inserted into the NdeI site of pET16b and confirmed by sequencing, then amplified using primers pCDFDuet R.p_matB F1/R1 and inserted between the NdeI and KpnI sites of pCDF-Duet1 to yield pCDF-Duet1.*Rpa*matB.

pET23c.His_6-TEV-NphT7 was constructed by PCR amplifying NphT7 from pET16b.nphT7 [*3*] using primers NphT7 23a F2/R1 and inserting into EcoRI-digested pET23a-His-TEV-SfoI.

pXHB vectors. The ribosome binding site and spacing from pET16b was inserted into pBAD33 by Gibson insertion of the overlapping primers pBAD33-pET.RBS GF1/GR1 between the KpnI and XbaI sites, removing KpnI in the process, to yield pBAD33.pET-RBS. To construct pXHB.ø.*Rpa*matB, the following fragments were combined in a one-pot Gibson assembly: 1) An araC/pBAD/rrnB fragment was PCR amplified from pBAD33-pET.RBS using primers pBAD GF1 and pBAD-pCDF GR1; 2) A lacI fragment was PCR amplified from pCDF-Duet1 using primers lacI GF1 and lacI-pBAD GR1; 3) plasmid pCDF-Duet1.*Rpa*matB was digested with XbaI and HindIII, removing lacI and one of the T7 promoters. This assembly yielded pXHB.ø.*Rpa*matB.

pXHB.ø.*Rpa*matB was used to construct variants as follows: Transporters were inserted after the pBAD promoter into XbaI/HindIII-digested plasmid by Gibson assembly. Ligases were inserted after *Rpa*matB into KpnI/XhoI-digested plasmid.

Transporters: MadLM (gene IDs 3480185 and 3480184) was PCR amplified using a colony of *Pseudomonas fluorescens Pf-5* as template with primers Pfl.madLM GF1/GR1. MatC (gene ID 1097877) was PCR amplified from pBAD18-Cb.500*Sco*matC (which was obtained from by PCR amplifying matC from a neighborhood PCR of *Streptomyces coelicolor A3(2)* M145 genomic DNA) using primers pCDF-ara-*Sco*MatC GF1/R1. MdcF was reverse translated based on the protein sequence from *Klebsiella pneumoniae 342* (gene ID 6935833) using the Integrated DNA Technologies *E. coli* codon optimization algorithm and purchased as two overlapping gBlocks (*Appendix Table* 2) which were inserted into pXHB.ø.*Rpa*matB in a 3-piece Gibson assembly.

Ligases: *Eco*ACS was PCR amplified from pCDF.P(Tet)-acs.accBC.dtsR1 (acs having been originally PCR amplified from *E. coli* K-12 DNA) using primers pCDF-ACS GF1/GR1. *Eco*PrpE was PCR amplified from pET16b.prpE (prpE having been originally PCR amplified from *E. coli* DH10b DNA) using primers PrpE GF1/GR1. *Sco*AlkK was PCR amplified from *Streptomyces coelicolor* genomic DNA using primers Sco.AlkK GF1/GR1. *Pol*AlkK was PCR amplified from pCDF3-sPhaZ.alKk (which contains a synthetic, *E. coli* codon optimized gene for the protein from *Pseudomonas oleovorans*,) using primers Spol.alkK GF1/GR1.

pXHB.*Pfl*madLM.ø was constructed by digesting pXHB1 with KpnI/NdeI, gel purifying the vector backbone and incubating with Gibson isothermal assembly mix without an insert. The selected clone had a deletion beginning 13 nt before the T7 promoter and extending to 13 nt after the XhoI site as compared to pXHB1.

pXHB3 was constructed by digesting pXHB1 with NdeI/NotI (which removes the T7 promoter). pT5 was PCR amplified from pT5T33-Bu2 using primers pXHB3.GF1/GR1 and inserted by Gibson assembly.

pPOL vectors. To construct pPOL1, NphT7 was PCR amplified from pET16b.nphT7 using primers FPHB-NphT7 F1/R1, and PhaB.PhaC was PCR amplified using primers FPHB-PHAB F1/PHAC R1 from pBT33-PhaABC [*23*] (PhaB was a synthetic gene [*23*] corresponding to PhaB1 from *Ralstonia eutropha H16*, gene ID 4249784, and PhaC was originally PCR amplified from *Ralstonia eutropha H16* genomic DNA, phaC1, Gene ID 4250156). The two fragments were inserted between the NcoI and XbaI sites of pTRC99a in a 3-piece Gibson assembly, yielding pPOL1. To construct pPOL1-nb, NphT7.PhaB was PCR amplified from pPOL1 using primers NphT7 G F3/PhaB G R1, and inserted between the NcoI and XbaI sites of pTRC99a by

Gibson assembly. To construct pPOL1-n, pPOL1-nb was digested with EcoRI/XbaI and the gap repaired by insertion of the oligos pPOL1-n GF1/GR1.

To construct pPOL4.1 and 4.2, *Rhodococcus ruber* PhaC (locus ID WP_017681423) (4.1) and *Rhodococcus opacus* PD630 PhaC (locus ID WP_005237952) (4.2) were reverse translated using the Integrated DNA Technologies *E. coli* codon optimization algorithm and purchased as two overlapping gBlocks (*Appendix Table 2*) which included the ribosome binding site from pET16b. pPOL1-nb was digested with XbaI and HindIII and the PhaC gene was inserted by Gibson assembly. The final clone selected for pPOL4.2 had a silent mutation at V254 (GTG > GTA).

FlK vectors. pET28a.flK and pET28a.flk-H76A were constructed by PCR amplification of flK from pET16b-His-sFlK or pET23a-His-Tev-sFlK-H76A using primers pET28a flK GF1/GR1 and insertion into XbaI/NcoI-digested pET28a. To construct pET28a.pTet.flK and pET28a.pTet.flK(H76A), these vectors were digested with XbaI and BglI (which removes lacI and pT7. TetR and pTet were amplified from pCDF.P(Tet)-acs.accBC.dtsR1 using primers pTet-pET28a.flk GF1/GR1 and inserted into the digested vectors.

Expression of His-tagged proteins. TB (3x1 L) containing carbenicillin (50 μg/mL) in 2.5 L Ultrayield baffled shake flasks (Thomson Instrument Company,) was inoculated to OD_{600} = 0.05 with an overnight TB culture of freshly transformed *E. coli* BL21(de3)T1^{R} containing the appropriate plasmid. The cultures were grown at 37°C at 200 rpm to OD_{600} = 0.6 to 0.8 at which point cultures were cooled on ice for 20 min, followed by induction of protein expression with IPTG (1 mM) and overnight growth at 16°C. Cell pellets were harvested by centrifugation at 10,000 × *g* for 5 min at 4ºC and stored at -80ºC.

Purification of His_{10}-*Rpa*MatB. *Rpa*MatB was purified as described [*3*] for *Sco*MatB.

Purification of His_6-TEV-NphT7. Frozen cell pellets (20 g) were thawed and resuspended at 5 mL/g cell paste with Buffer A (50 mM sodium phosphate, 500 mM sodium chloride, 20% glycerol, 20 mM BME, pH 7.5) containing imidazole (10 mM) and Tween 20 (1%). Complete EDTA-free protease inhibitor cocktail (Roche) was added to the lysis buffer before resuspension (1 pellet/100 mL lysis buffer). The cell paste was lysed by sonication (5 s on, 25 s off, 10 min processing time). The lysate was centrifuged at 15,300 × *g* for 40 min at 4°C to separate the soluble and insoluble fractions. The soluble lysate was applied to Ni-NTA Agarose (Qiagen, 1 mL/7 g cell paste) pre-equilibrated with Buffer A and allowed to batch bind for 1 h at 4°C on a rocking agitator. The resin was transferred to a column by gravity flow. The column was washed with Buffer A + 10 mM imidazole, then Buffer A + 20 mM imidazole, each time until the eluate was negative for protein content by Bradford assay (Bio-Rad), then eluted with Buffer A + 250 mM imidazole. The eluate was diluted 7-fold with Buffer B (50 mM HEPES, 2.5 mM EDTA, 1 mM DTT, 20% glycerol, pH 7.5) and loaded onto a HiTrap Q HP column (GE Healthcare, 5 mL) using an ÄKTApurifier FPLC (GE Healthcare Life Sciences; Piscataway, NJ). Protein was eluted with a linear gradient from 0 to 1 M sodium chloride in Buffer E over 50 column volumes (4 mL/min). Fractions containing the target protein (eluted at ~100 mM sodium chloride) were pooled by $A_{280\ nm}$ and concentrated under nitrogen flow (60 psi) in an Amicon ultrafiltration cell using a YM10 membrane. The protein (~0.3 mg) was flash-frozen in liquid nitrogen and stored at -80°C at a final concentration of 0.3 mg/mL, which was estimated using the calculated $\varepsilon_{280\ nm}$ (26,930 $M^{-1}\ cm^{-1}$).

Preparation of sodium fluoromalonate. Fluoromalonate was prepared either as described [*3*] or as follows. Diethylfluoromalonate (19 mmol) was hydrolyzed using sodium hydroxide (38 mmol, 2.1 eq) in water (40 mL). The mixture was stirred at room temperature (without controlling the moderate exotherm) for 15 min until the pH was basic and no organic phase was visible. The resulting solution was transferred to a round-bottom flask with water and concentrated by rotary evaporation in a 65 °C water bath to dryness (to remove ethanol). The resulting white, waxy solid was redissolved in deionized water to a final volume of 20 mL (1 M sodium fluoromalonate) and frozen in 2 mL aliquots at -20 °C. The exact concentration was determined by quantitative ^{19}F NMR as described below.

***In vitro* NphT7 assay.** Organofluorine products of NphT7 were observed by ^{19}F NMR. The reaction mixture (600 µL) contained 50 mM HEPES, pH 7.5, trifluoroacetate (2 mM), D_2O (15%), phosphoenol pyruvate (10 mM), pyruvate kinase (18 U/mL), myokinase (2.5 U/mL), ATP (2.5 mM), magnesium chloride (2.5 mM), coenzyme A (1 mM), acetyl-CoA (1 mM), fluoromalonate (5 mM) and *Rpa*MatB (10 µM). The mixture was preincubated at 298 K and formation of fluoromalonyl-CoA verified by NMR. The reaction was initiated by addition of NphT7 (5 µM) and observed by ^{19}F NMR at intervals essentially as described below (high-throughput NMR analysis).

***In vivo* production of 2-fluoro-3-hydroxybutyrate and poly(2-fluoro-3-hydroxybutyrate-*co*-3-hydroxybutyrate).** Chemically competent cells of the appropriate strain were transformed with either pXHB or pPOL. A colony was picked, grown in LB with 1% glucose and the appropriate antibiotic (carbenicillin and streptomycin, or spectinomycin for Sm^R strains), and cells were harvested at OD_{600} 0.3-0.5 and made chemically competent. These cells were freshly transformed with the second plasmid, pXHB or pPOL. Colonies were picked and grown for 12-18h at 37°C in LB (5 mL) with 1% glucose, carbenicillin (50 µg/mL) and streptomycin (50 µg/mL) or spectinomycin (25 µg/mL). Overnight cultures were diluted 50-fold into fresh medium (50 mL in a 250 mL baffled flask) and grown at 37°C until either OD_{600} 0.2-0.4 (early induction) or OD_{600} 0.7-0.9 (late induction). Culture (5 mL) was transferred to 40 mL culture tubes containing IPTG (1 mM final), arabinose (0.2% final) and sodium fluoromalonate (5 mM final unless otherwise noted). Cultures were grown at 30°C for approximately 48h before sampling. 2 mL culture was removed in duplicate and centrifuged at 23,000 × *g* for 5 minutes in plastic 2 mL screw-top tubes with O-ring seals (Thomas Scientific, Swedesboro, NJ). The supernatant was decanted and saved for ^{19}F NMR and LC-MS analysis. For polymer-producing strains, the tubes were centrifuged again and the remaining medium removed by aspiration; cell pellets were frozen at -20°C. For polymer isolation, polymer-producing strains were induced in 50 mL medium in 250 mL baffled flasks, the entire culture was collected by centrifugation at 10,000 × *g*.

Growth curves. Cultures prepared as above for production experiments were grown at 37°C to OD_{600} 0.8-1.2, then used to inoculate fresh medium (25 mL) to OD_{600} 0.10. Cultures were grown at 37°C and ODs monitored until they reached 0.25-0.3 (1-2 h). Culture was transferred to 40 mL tubes (5 mL per condition) containing inducing agents and fluoromalonate as appropriate, and 200 µL was removed from each tube and transferred to a 96-well plate. The plate was incubated with continuous shaking uncovered at 30°C in a SpectraMax M2 plate reader (Molecular Devices, Sunnyvale, CA) and OD_{600} measurements taken every 5 min for 10 h. An internal blank of uninoculated medium containing 200 µg/mL chloramphenicol was included.

The tubes were incubated at 30°C at 200 rpm and sampled as described above after 48 h. OD_{600} data were corrected for the path length of 200 μL liquid (0.6 cm).

Large-scale polymer production. Overnight cultures of *E. coli* BL21(de3)$T1^R$ containing pXHB1 and pPOL1 were prepared as described above and diluted 50-fold into LB (1L) containing 1% glucose, carbenicillin (50 μg/mL) and streptomycin (50 μg/mL). Cultures were grown to OD_{600} 0.7-0.9 and induced with 1 mM IPTG and 0.2% arabinose, and sodium fluoromalonate or sodium malonate (1 mM) was added. Cultures were grown for 48h at 30°C, then cell pellets were harvested by centrifugation (10,000 × *g*).

Polymer purification from cell pellets. Polymer was isolated by hypochlorite/sodium dodecyl sulfate (SDS) lysis of cells and oxidation of cell debris. This procedure is known to have a modest effect on the molecular weight of isolated polymer, since poly(3-hydroxybutyrate) is not completely resistant to hypochlorite oxidation [*4*]. Cell pellets were resuspended at room temperature in a solution of 5% sodium hypochlorite (from household bleach) and 1% SDS (from a 10% stock solution solubilized at 55°C) at approximately 20 mL/g cell paste. Polymer was pelleted by centrifugation at 10,000 × *g* and the supernatant was discarded. The polymer pellet was resuspended in water, then pelleted again; this was repeated a total of four times. Polymer was then washed 2x with methanol, 1x with ethyl acetate and 1x with hexanes. The polymer was resuspended in hexanes and transferred as a suspension to a tared 10 mL screw-top tube with a rubber seal. Residual hexanes were evaporated under a stream of nitrogen and chloroform (or chloroform-d) was added at ~1 mL/20 mg polymer. The mixture was heated in a sand bath at 90°C overnight and vortexed periodically until the polymer dissolved. The resulting solution was used for NMR characterization. For determination of materials properties, polymer was precipitated from chloroform by dropwise addition into >100 volumes of rapidly stirring methanol. The precipitated polymer was collected manually and dried under vacuum.

Analysis of polymer properties. SEC was performed using a Hewlett-Packard series 1100 liquid chromatography system ($CHCl_3$, 35 °C, 1 mL/min), equipped with a Hewlett-Packard 1047A RI detector and three PLgel 5 μm MIXED-C columns. Apparent molar masses were reported versus polystyrene standards (Polymer Laboratories). Thermogravimetric analysis (TGA) was performed on a Perkin Elmer Diamond TGA/DTA. Samples were heated in aluminum pans under air at a rate of 10°C min^{-1}.

Analysis of organofluorines in culture media by high-throughput ^{19}F NMR analysis. Culture medium samples for NMR analysis were prepared as described above. Culture medium (595 µL) was mixed with TFA internal standard (140. mM, 10 µL) and D_2O (95 µL) and approximately 600 µL of the solution were transferred to a 7" Grade 2, 5 mM NMR tube (Kimble Chase Life Science, Vineland NJ). ^{19}F NMR spectra were collected at 298 K on a Bruker AV-600 spectrometer at the College of Chemistry NMR Facility at the University of California. Tuning, shimming and 90° pulse time were optimized for each sample. The typical 90º pulse was ~16 µs at -3 dB power. Two spectra per sample (*Figure 3.8*) were collected consecutively using the zg30 pulse program as follows: Spectrum 1 observed TFA (referenced to -76.4 ppm) and fluoride (-121 ppm; not observed except as a contaminant in media) and used the following parameters: rg = 512, o1p = -100 ppm, sw = 60 ppm, td = 11750 (aq = 0.34 s), d1 = 0 s, ds = 20, ns = 100; Spectrum 2 observed fluoromalonate (-178.4 ppm), both diastereomers of 2-fluoro-(*R*)-3-hydroxybutyrate (-197.4 and -199.9 ppm; tentatively assigned as 2-(*R*)-3-(*R*) and 2-(*S*)-3-(*R*) respectively) and fluoroacetate (-217.9 ppm) and used the following parameters: rg = 512, o1p = -198 ppm, sw = 79 ppm, td = 15360 (aq = 0.34 s), d1 = 0 s, ds = 20, ns = 1000 (for quantitation to ~50 µM). Background signal from fluoropolymer in the probe was removed by backward linear prediction (ME_mod = LPbc, ncoef = 32, tdoff = 64) and the spectrum was phased and baseline-corrected automatically (apk, abs). Peaks were manually integrated and absolute integrals were normalized to TFA. To account for incomplete relaxation between scans (*Figure 3.2*), a standard curve was prepared for each analyte by collecting spectra of dilutions of a single sample in culture medium (*Figure 3.3*) and a linear scaling factor was applied to the normalized integrals. The concentrations of analytes in the 1x sample were determined by

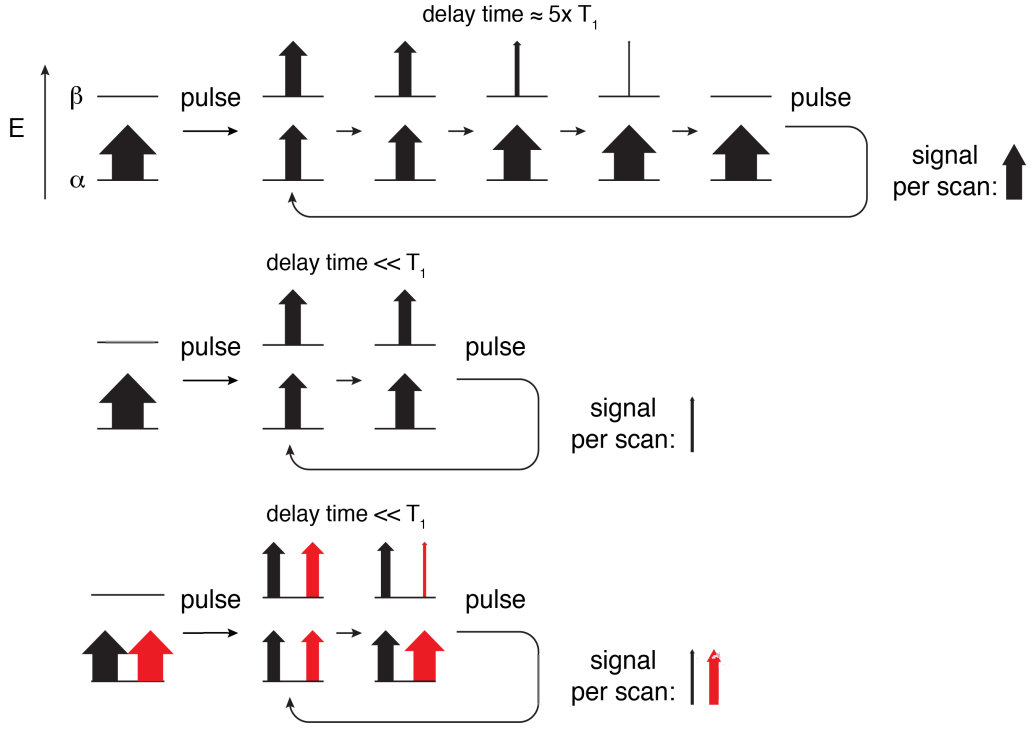


Figure 3.2. *Diagram showing how incomplete relaxation affects NMR integrations. Top panel: quantitative integrals result from complete relaxation. Middle panel: short relaxation times reduce signal. Bottom panel: species with different relaxation times give disproportionate amounts of signal under incomplete relaxation conditions.*

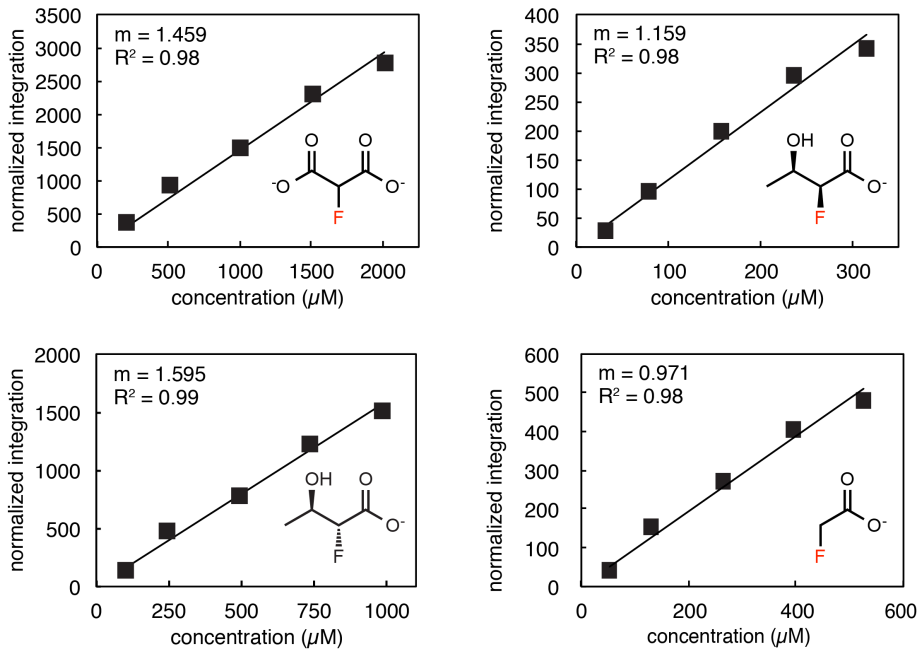


Figure 3.3. *^{19}F NMR standard curves for organofluorine analytes.*

comparison to the 2 mM TFA standard under quantitative conditions (d1 > 5 × the largest T1).

Analysis of 2-fluoro-3-hydroxybutyrate and 3-hydroxybutyrate in culture medium by LC-MS. Culture medium samples were diluted 100x into a solution of 10 µM adipic acid (internal standard) in water. Samples were analyzed using an Agilent 1290 HPLC system equipped with a needle wash system (50% methanol/water) on a Rezex ROA-Organic Acid H+ (8%) column (5 µm, 4 × 150 mm, Phenomenex) maintained at 55°C using 0.5% formic acid as the mobile phase and an isocratic method (0.6 mL/min) for 10.5 minutes. The injection volume was 2 µL. Products were monitored using a triple quadrupole mass spectrometer (Agilent 5130) operating in multiple reaction monitoring mode. 2-fluoro-3-hydroxybutyric acid (*m*/*z* 121 > 101), 3-hydroxybutyric acid (103 > 59) and adipic acid (145 > 83) were detected in negative ion mode using the following parameters: ΔEMV 400, fragmentor voltage 70, collision energy 10 (adipic acid) or 5 (hydroxybutyric acids), cell accelerator voltage 7, dwell time 120 ms. Crotonic acid (87 > 69), resulting from polymer acidolysis (see below), was detected in positive ion mode using the following parameters: ΔEMV 400, fragmentor voltage 20, collision energy 8, cell accelerator voltage 4, dwell time 200 ms. The ESI source with Jetstream sheath gas was configured as follows: gas temperature 285 °C, gas flow 10 L/min, nebulizer 35 psi, sheath gas temperature 300 °C, sheath gas flow 12 L/min. 2-Fluorocrotonate was not observed by MS, and only a trace amount was detected by ^{19}F NMR (as a doublet with J = 36 Hz, about -127 ppm) even when polymer was acidolyzed in neat concentrated sulfuric acid and diluted with basic D_2O at room temperature, presumably because the equilibrium constant for hydration of this species to 2-fluoro-3-hydroxybutyrate is large. *Figure 3.4* shows an example chromatogram.

Analysis of poly(2-fluoro-3-hydroxybutyrate-*co*-3-hydroxybutyrate) in whole cell acidolysate by LC-MS. Cell pellets, collected in 2 mL screw-top tubes as described above, were resuspended in 50 µL 16 mM adipic acid (internal standard) by adding ~20-30 glass beads (1 mm diameter) and vortexing vigorously. Concentrated sulfuric acid (150 µL) was added and the mixture incubated at 95 °C in a heat block for ~15 h, resulting in a uniform dark brown

suspension. Samples were vortexed and diluted with water (1.4 mL), then heated at 95 °C again for 2-6 h to allow partial hydration of crotonic acids formed during the depolymerization. Tubes were centrifuged at 21,850 × *g* to remove insoluble debris and supernatant was removed and diluted 50-fold for LC-MS analysis as described above for media samples. The final concentration of adipic acid after dilution was 10 µM.

For analyte quantitation, a standard solution of poly(2-fluoro-3-hydroxybutyrate-*co*-3-hydroxybutyrate) acidolysate was prepared by acidolysis of pure, isolated polymer (7.8 mg) as described above. The monomer ratio was determined by ^{1}H NMR before acidolysis. The concentration of compounds derived from 3-hydroxybutyrate was determined using independent standard curves of sodium hydroxybutyrate and crotonic acid, and these values were used along with the previously determined monomer ratio to calculate the concentration of 2-fluoro-3-hydroxybutyrate. Serial dilutions of this acidolysate solution were used as calibration standards; analyte responses were normalized to the internal standard.

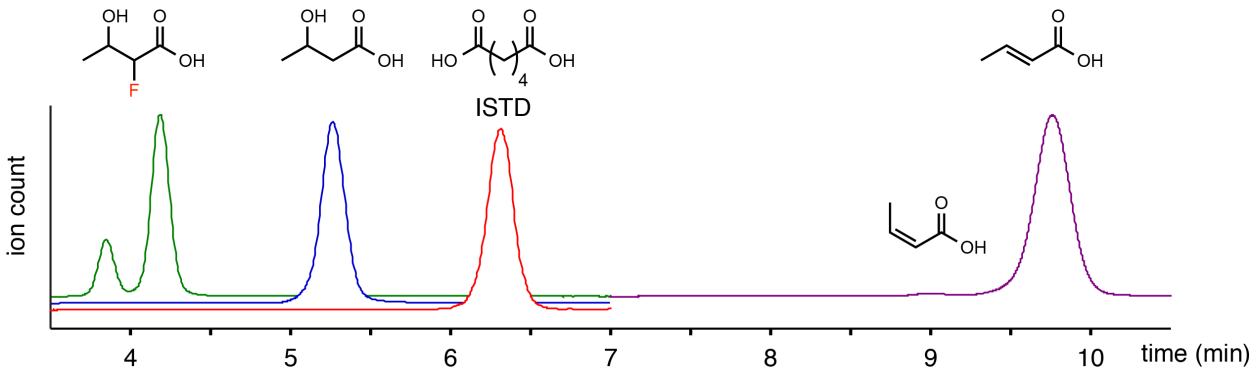


Figure 3.4. *Example MRM chromatogram for media/polymer analytes.*

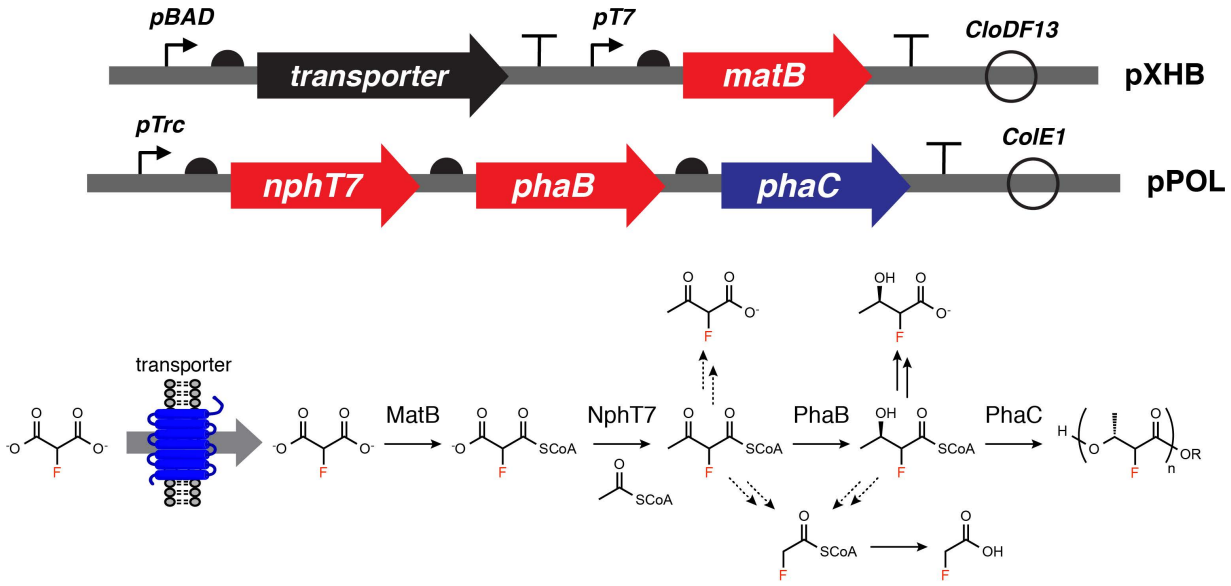


Figure 3.5. *Plasmids and pathway for 2-fluoro-3-hydroxybutyrate and poly(2-fluoro-3-hydroxybutyrate) synthesis from fluoromalonate. Formation of possible derailment products is shown with dotted arrows.*

3.3 Results and discussion

Design of a fluorodiketide biosynthetic pathway. We constructed a four-gene, two-plasmid system for biosynthesis of 2-fluoro-3-hydroxybutyryl-CoA from exogenous fluoromalonate (*Figure 3.5*). Plasmid pXHB encodes enzymes for formation of fluoromalonyl-CoA, while pPOL encodes ketosynthase, ketoreductase and polymerase enzymes.

We chose to express MatB using a T7 promoter from pXHB, which has a CloDF13 origin, giving a copy number approximately twice that of pPOL (ColE1 origin). We expected this combination to emphasize fluoromalonyl-CoA formation over diketide synthesis, since if fluoromalonyl-CoA is indeed a committed metabolite in *E. coli*, its accumulation would drive flux through the pathway without resulting in toxicity or off-pathway reactions. We previously found that the MatB enzyme from *Streptomyces coelicolor* was able to catalyze the ATP-dependent ligation of fluoromalonate to CoA with a catalytic efficiency of 530 ± 40 $M^{-1}s^{-1}$ [*3*]. However, a recently reported homolog from *Rhodopseudomonas palustris* was reported to have greater promiscuity toward alternative substrates, e.g. methylmalonate [*24*]. We tested the activity of this homolog on fluoromalonate and found a larger catalytic efficiency of 1,250 ± 80 $M^{-1}s^{-1}$ (*Figure 3.6*), including a 7-fold higher k_{cat} parameter, and accordingly used it for fluoromalonate activation in the current study.

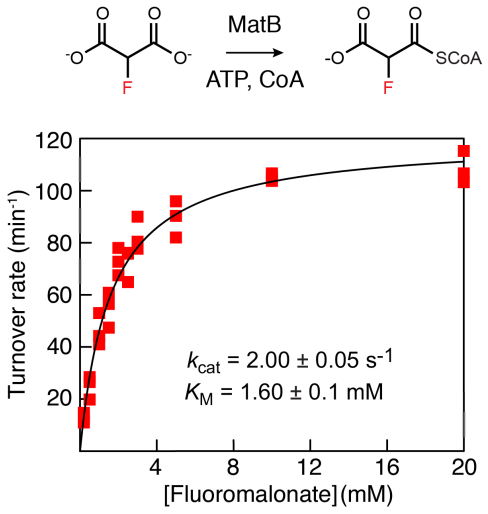


Figure 3.6. *Dose response curve for Rhodopseudomonas palustris MatB activation of fluoromalonate, fit to the Michaelis-Menten equation.*

To provide MatB with fluoromalonate, we incorporated a malonate transporter (see below)

into pXHB on a separate promoter. Although previous work in our group found that fed fluoromalonate could be converted to fluoro-triketide using a polyketide synthase module in *E. coli* cell culture, the optimal conditions required very high cell density (OD ~100) and high (up to 50 mM) fluoromalonate concentrations to obtain product titers on the order of 1 μM [*3*]. We hypothesized that because of its two ionizable groups and low pK_a values, diffusion of fluoromalonate across the cell membrane might be difficult, and that a malonate transporter might provide a solution if it had activity on fluoromalonate as well.

We ordered the NphT7 and PhaB genes as shown in pPOL in part because NphT7 catalyzes an effectively irreversible step (energetically equivalent to hydrolysis of ATP and producing CO_2), and would therefore be expected to have a larger flux control coefficient in the pathway than PhaB, making higher expression advantageous. In addition, the *phaCAB* operon in *Ralstonia eutropha* has a similar arrangement. We did not expect PhaB activity to limit pathway flux: Although it has not been characterized with respect to activity on acetofluoroacetyl-CoA, its activity on its native substrate acetoacetyl-CoA is high (k_{cat}/K_M 1.80 x 10^7) [*25*] and reduction of an α-fluoro ketone would be more energetically favorable.

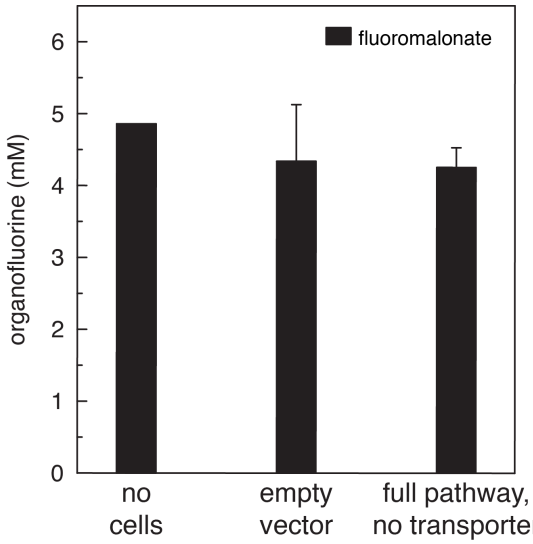


Figure 3.7. *Organofluorine levels observed by ^{19}F NMR in the absence of a malonate transporter. Following induction, cultures were incubated for 48 hrs at 30°C, then supernatant samples analyzed by NMR. Error bars show sample standard deviation, n=3 biological replicates, except for the no cell condition where n=1.*

A fluoromalonate transporter allows mM pathway flux. *E. coli* BL21(de3) cells harboring pXHB.ø.*Rpa*matB, which does not contain a malonate transporter, and pPOL1-nb did not convert fed fluoromalonate (5 mM) to any other fluorinated species within the limit of detection (50 μM, 1%) of our ^{19}F NMR assay (*Figure 3.7*). We therefore tested one representative of each of three classes of architecturally distinct prokaryotic malonate transporters that we identified using the Transporter Classification Database [*26*]. MadLM is a characterized two-component sodium: malonate symporter [*27*] from *Malonomonas rubra*; we used a homolog from *Pseudomonas fluorescens*. MatC and MdcF, from *Streptomyces coelicolor* and *Klebsiella pneumoniae* respectively, are putative malonate symporters found in operons with malonate:CoA ligase and malonyl-CoA decarboxylase genes involved in exploitation of malonate as a carbon source. None of these proteins share significant sequence identity (except minor similarity between MdcF and MadM, 26% identical over 80 amino acids, blastp expect value 0.013).

Surprisingly, we observed little toxicity (data not shown) when expressing any of the transporters at full induction (0.2% arabinose) even with a strong canonical RBS, suggesting that either translation or membrane insertion is not very efficient for these proteins. We found that all three transporters were able to accept fluoromalonate, leading to formation of two new organofluorines in the culture medium: 2-fluoro-3-hydroxybutyrate and fluoroacetate (*Figure 3.8*). Both MatC and the two-component transporter MadLM enabled mM levels of flux to fluorohydroxybutyrate (*Figure 3.9*). The maximal flux observed in this experiment was 30 µM/hr for MatC and 50 µM/hr for MadLM.

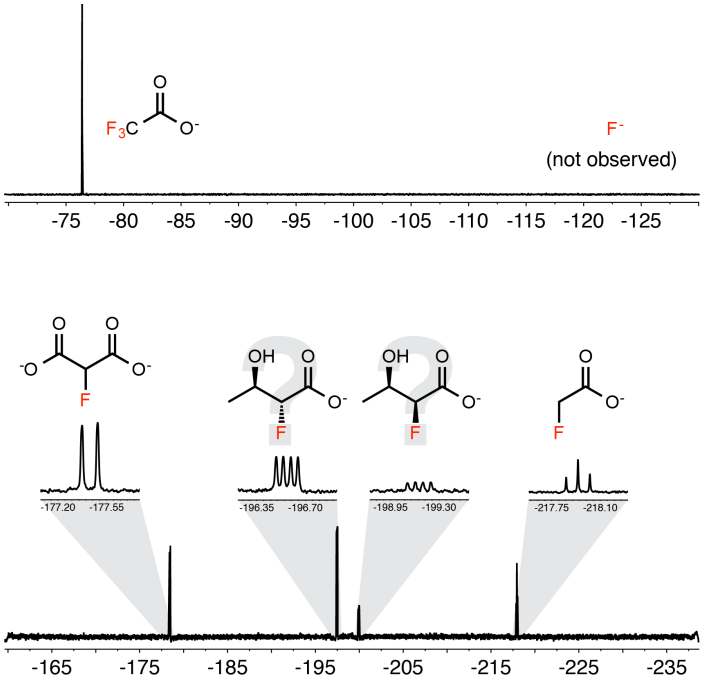


Figure 3.8. *Organofluorines in culture medium detected from full pathway expression. Diastereomers of 2-fluoro-3-hydroxybutyrate are tentatively assigned based on chemical shift (as in [3]). 2 mM TFA is used as an internal standard.*

We evaluated the response of each transporter to fed fluoromalonate concentration at low and high induction ODs (*Figure 3.10*) and found that MadLM achieved titer saturation at lower fluoromalonate levels and also was more versatile in that it performed similarly with early or late induction, whereas MatC required early induction to achieve high titers. We carried out the remaining studies using the MadLM-containing plasmid pXHB1.

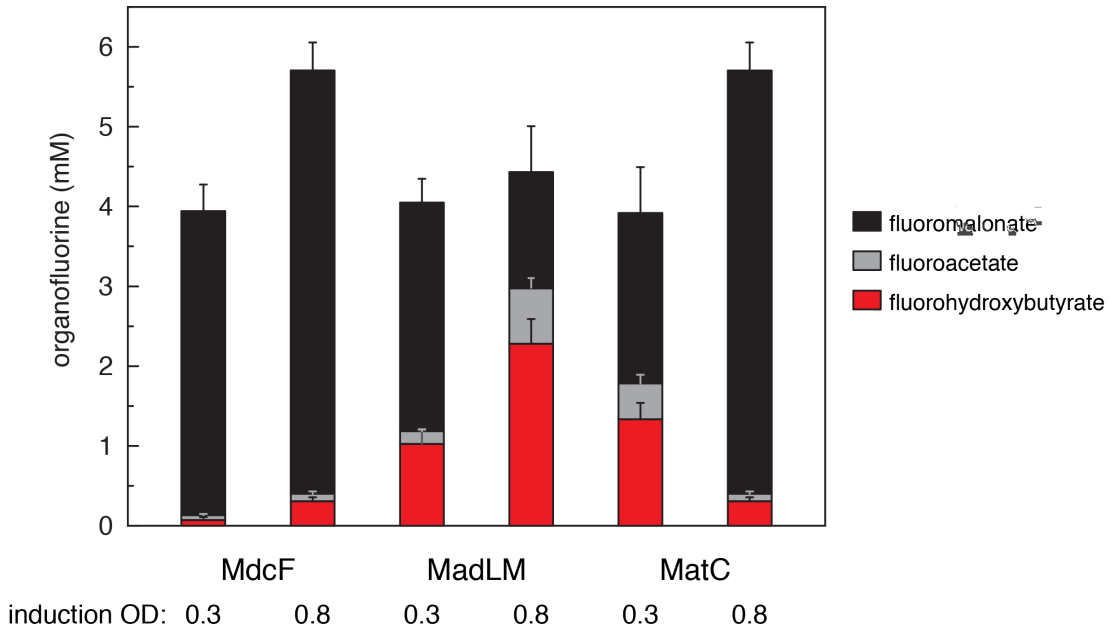


Figure 3.9. *Organofluorine levels observed by ^{19}F NMR with expression of malonate transporters at early or late induction times. Performance of MadLM was somewhat higher in this experiment than was typical. Error bars show sample standard deviation, n=3 biological replicates.*

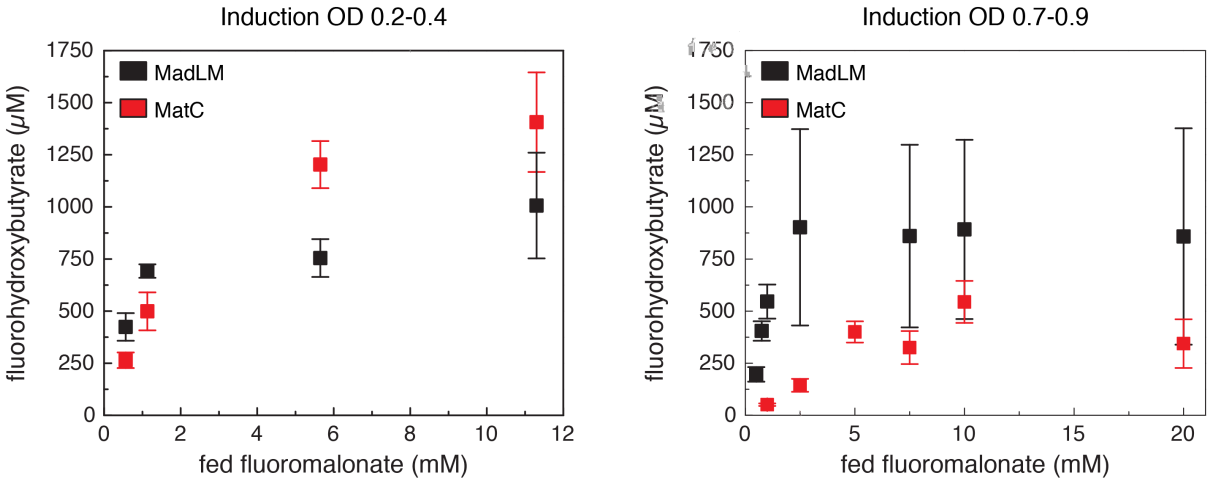


Figure 3.10. *Fluorohydroxybutyrate titers resulting from varying fluoromalonate feeding levels. Error bars show sample standard deviation. Three replicate cultures for each transporter and induction OD condition were grown and induced, then each culture was split and portions were incubated with various concentrations of fluoromalonate. Large variability with MadLM cultures at late induction OD appears to be due to culture-to-culture variation independent of fluoromalonate concentrations (that is, production was consistently high or low from a given culture) rather than variable responses to fluoromalonate level.*

We compared pathway flux in 6 *E. coli* host strains, B-type strain BL21(de3) and K-12-type strains BW25113, W3110, DH1, DH5α(de3) and DH10b. Because not all strains were readily available as *de3* lysogens, we used pXHB3 for this comparison, in which MatB is expressed from a T5 promoter rather than T7. Although the pathway was functional in all strains, only BL21(de3) and DH5α(de3) produced significant titers of fluorohydroxybutyrate (*Figure 3.11*). BL21(de3) is a common strain for protein overexpression and lacks the Lon and OmpT proteases, which may be important to express sufficient levels of some or all of the pathway enzymes. However, DH5α produces even higher titers with those proteases intact. The performance of DH5α is surprising given its genotypic similarity to DH1 and DH10b, which are far less effective. Unfortunately, its growth rate with the pathway plasmids was slow and lag times after inoculation were 5–8 hours, so BL21(de3) was selected as a production strain.

Only fluoroacetate and fluorohydroxybutyrate accumulate. Irrespective of pathway variant, strain, growth and induction conditions, we observed only fluoromalonate, fluorohydroxybutyrate and fluoroacetate in the growth medium (*Figure 3.8*). Fluorohydroxybutyrate presumably originates from hydrolysis of fluorohydroxybutyryl-CoA. To evaluate whether this reaction might be catalyzed by an endogenous thioesterase, we compared organofluorine production by *tesA* and *tesB* knockout strains from the Keio collection [*20*] to their parent (again using pXHB3), but found no difference in production (data not shown), albeit in a background of very low pathway performance (*Figure 3.11*). Another

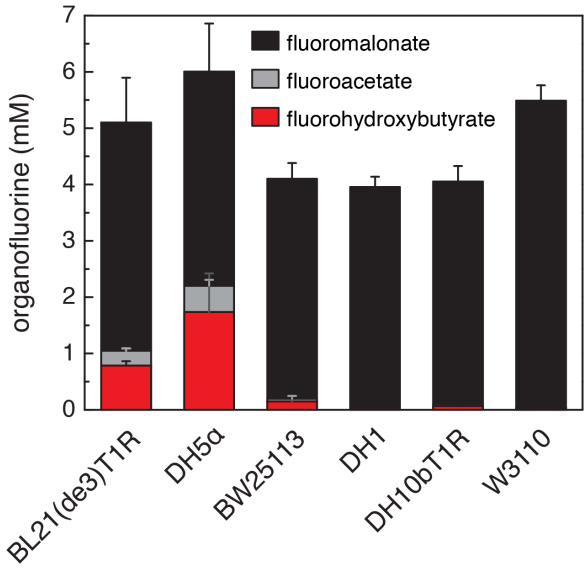


Figure 3.11. *Comparison of pXHB3/pPOL1-nb pathway organofluorines from six E. coli strains. Error bars show sample standard deviation, n=3 biological replicates.*

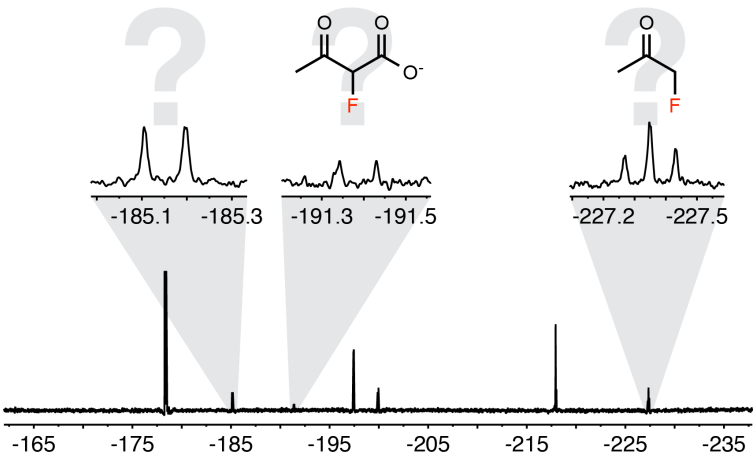


Figure 3.12. *^{19}F NMR spectrum of organofluorines produced using plasmids pXHB1 and pPOL1-n. 5-fold more scans were collected than for typical organofluorine quantitation spectra. Tentative assignments are based on chemical shift, multiplicity and by comparison with in vitro NphT7 products (Figure 3.14).*

enzyme may play a role, or the relative lability of these α-fluorinated acyl-CoAs may result in facile non-catalyzed hydrolysis. No organofluorines were detected by ^{19}F NMR in lysate from washed cells (intracellular concentrations < 60 μM) after 48h of growth, consistent with relatively low steady-state concentrations of fluoroacyl-CoA intermediates including fluorohydroxybutyryl-CoA.

We did not observe acetofluoroacetate, the hydrolysis product of acetofluoroacetyl-CoA, from the full pathway, suggesting that PhaB activity meets or exceeds NphT7 activity. Only in a PhaB-deleted pathway, pXHB1 pPOL1-n, did we observe peaks in the ^{19}F NMR tentatively attributed to acetofluoroacetate and fluoroacetone, its decarboxylation product, along with another unidentified organofluorine (*Figure 3.12*). This strain still produced a modest amount of fluorohydroxybutyrate, presumably by action of an endogenous ketoreductase (*Figure 3.13*).

Fluoroacetate originates from off-pathway reactions and leads to minor toxicity. Toxic fluoroacetyl-CoA, and its hydrolysis product fluoroacetate, could potentially arise from several pathway intermediates. We found that fluoromalonate itself is stable in culture medium for >2 weeks and is not transformed by cells carrying empty vector, lacking a transporter, or lacking MatB (*Figure 3.7*). However, fluoromalonyl-CoA appears to be subject to decarboxylation at a low level even in the absence of downstream pathway enzymes, and expressing NphT7 further increases fluoroacetate formation (*Figure 3.13*). The cause of non-NphT7-catalyzed fluoromalonyl-CoA decarboxylation in *E. coli* remains to be determined.

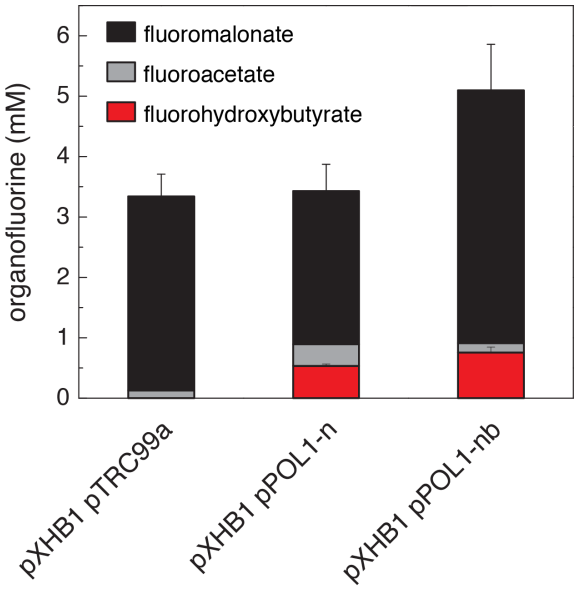


Figure 3.13. *Organofluorines produced by truncated pathway variants. Error bars show sample standard deviation, n=3 biological replicates.*

Like other ketosynthases, NphT7 has been shown [*6*] to decarboxylate extender units nonproductively, i.e. without chain extension. To verify this activity, we purified His$_6$-NphT7 and examined its activity condensing acetyl-CoA with fluoromalonyl-CoA (generated *in situ* using MatB) using ^{19}F NMR (*Figure 3.14*). New peaks tentatively assigned to acetofluoroacetate and acetofluoroacetyl-CoA formed rapidly, along with a small amount of fluoroacetyl-CoA. Non-quantitative NMR parameters were used to

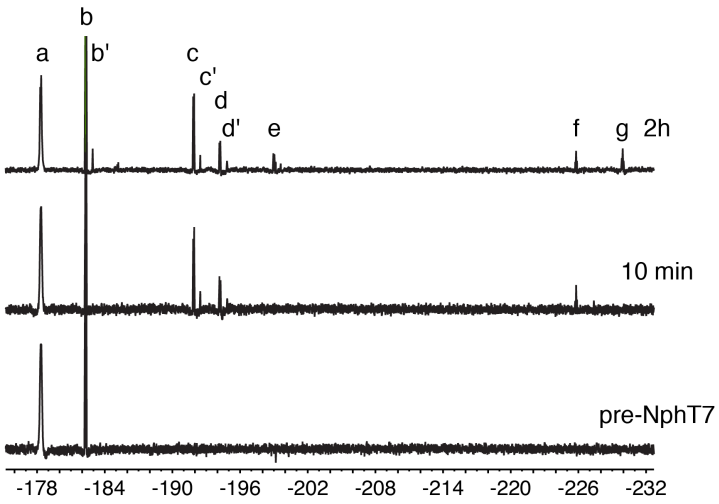


Figure 3.14. *NphT7-catalyzed condensation between fluoromalonyl-CoA and acetyl-CoA observed by ^{19}F NMR. Peaks with a prime are deuterated species arising from exchange with the solvent (15% D_2O/H_2O). Tentative assignments: a. fluoromalonate b. fluoromalonyl-CoA c. acetofluoroacetate d. acetofluoroacetyl-CoA e. unknown f. fluoroacetyl-CoA g. fluoroacetone*

achieve acceptable signal-to-noise, but the data allow an estimate of the ratio of productive (bond-forming) to nonproductive decarboxylation events at between 8 : 1 and 14 : 1. This ratio is roughly consistent with the increase in fluoroacetate production observed with NphT7 expression (*Figure 3.13*).

Fluoroacetyl-CoA is toxic to *E. coli* and many other organisms because of mechanism-based inhibition of aconitase and TCA cycle shutdown [*8*]. This toxicity could be a roadblock to achieving sustainable, high titers of fluoropolyketides. To evaluate toxicity, we compared the growth rates of pathway variants after induction in either the presence or absence of 5 mM fluoromalonate (*Figure 3.15*). Fluoromalonate-dependent toxicity was observed and tracked with 48 hr fluoroacetate production (except in the case of pXHB.*Sco*matC.*Rpa*matB, which showed

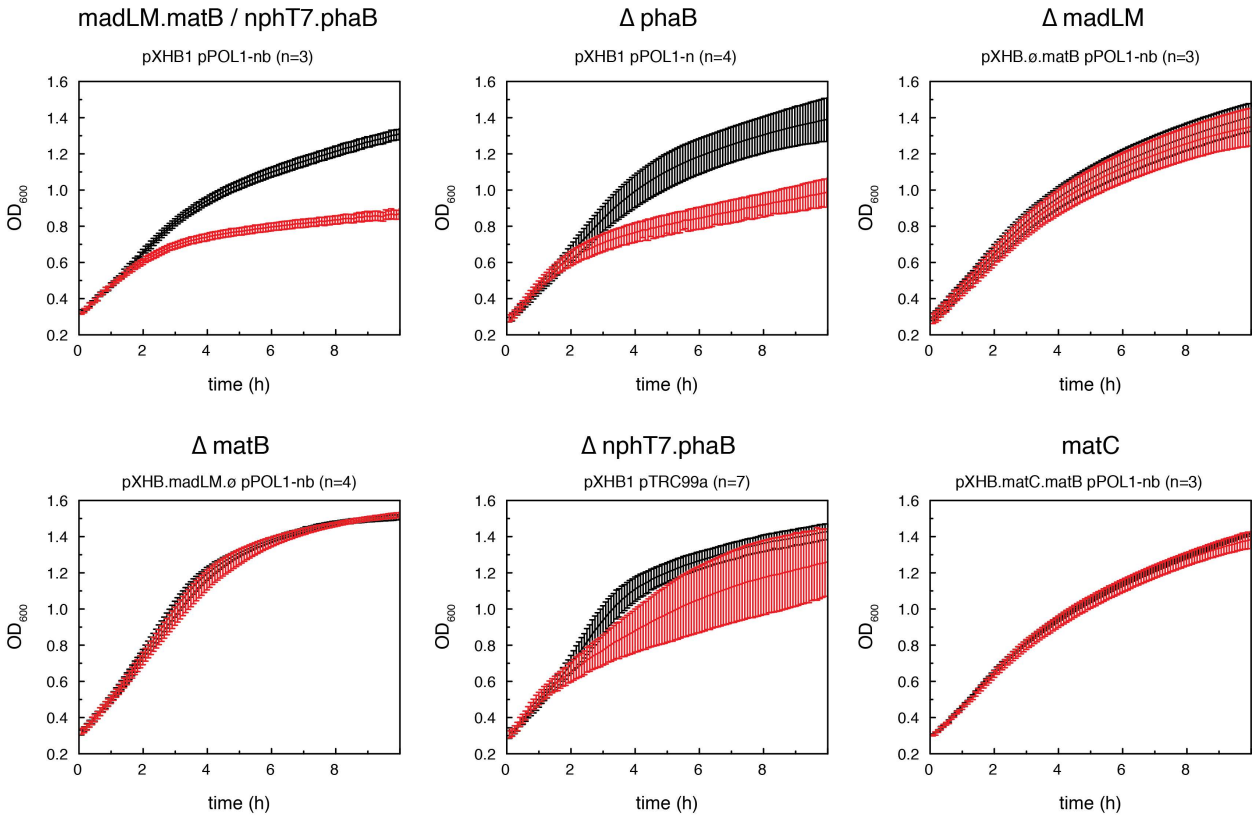


Figure 3.15. *Growth curves for indicated pathway variants after pathway induction in the presence (red) and absence (black) of 5 mM fluoromalonate. Cultures were grown in a 96-well plate with shaking at 30°C. Error bars show sample standard deviation, n as indicated.*

no fluoromalonate-dependent toxicity for reasons that are unclear, perhaps related to slower fluoromalonate uptake during log phase). However, the growth inhibition was relatively minor, resulting in saturation culture densities lower by at most 2 OD units (data not shown), possibly because the majority of fluorometabolite production occurs after the onset of stationary phase even when induction is during log phase growth.

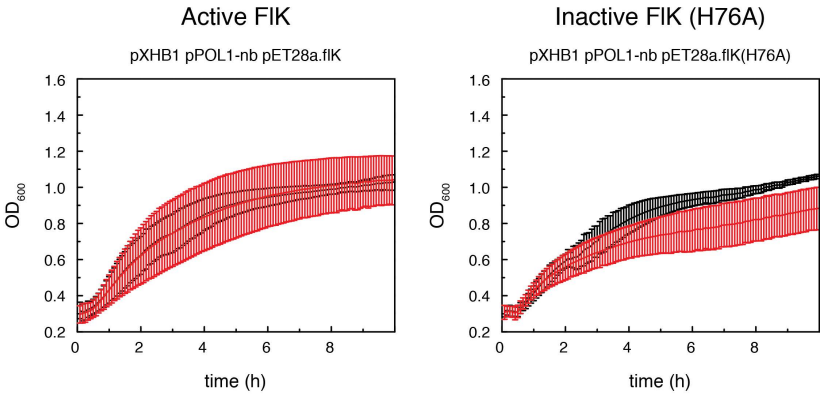


Figure 3.16. *Growth curves for strains expressing active or inactive FlK, a fluoroacetyl-CoA-specific thioesterase, from a T7 promoter in the presence (red) and absence (black) of 5 mM fluoromalonate. Cultures were grown in a 96-well plate with shaking at 30°C. Error bars show sample standard deviation, n=3 biological replicates.*

To verify that fluoroacetyl-CoA was responsible for the observed growth defects, we expressed the fluoroacetyl-CoA specific thioesterase from *Streptomyces cattleya*, FlK [*28-30*], along with pathway enzymes (a three-plasmid system). Wild-type FlK, but not a catalytically compromised mutant (H76A, catalytic efficiency reduced by a factor of 3×10^5) [*29*], rescued the fluoromalonate-dependent growth reduction when the full pathway was expressed (*Figure 3.16*), although expression of FlK from a T7 promoter was associated with a fluoromalonate-independent growth defect.

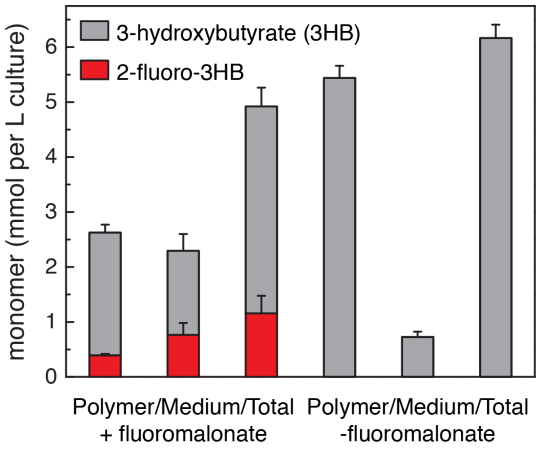


Figure 3.17. *Monomer content of polymer and growth medium for full pathway (+phaC) in the presence or absence of 5 mM fluoromalonate. Polymer was acidolyzed before LC-MS and medium assayed directly. The y-axis represents concentration in mM for media samples, and mmol per L culture for intracellular polymer. Error bars show sample standard deviation, n=3.*

We also observed an increased titer of fluoroacetate after 48 hrs in the active, but not an inactive FlK expression strain (n=3, 286 ± 80 µM for FlK vs. 219 ± 11 µM for FlK-H76A). This increase in titer is notable because it suggests that without a thioesterase, metabolism of fluoroacetyl-CoA is slow under these conditions, since it represents an apparent bottleneck in fluoroacetate formation. If *E. coli* can indeed maintain a pool of fluoroacetyl-CoA without severe toxicity or hydrolysis, expression of an appropriate acyl-CoA carboxylase could allow this off-pathway intermediate to be redirected to fluoromalonyl-CoA and desired downstream products.

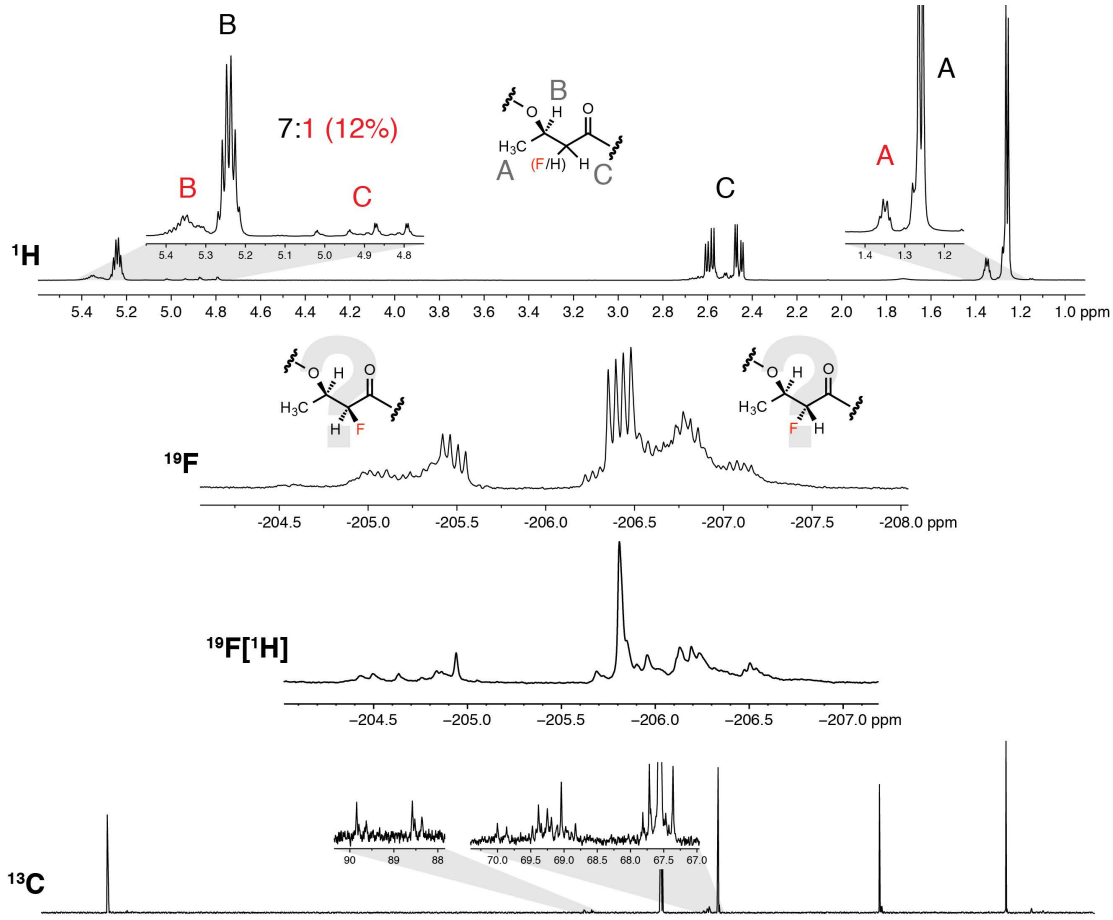


Figure 3.18. *1D NMRs of poly(2-fluoro-(R)-3-hydroxybutyrate-co-(R)-3-hydroxybutyrate). Diastereomers of the fluorinated monomer are assigned tentatively based on chemical shift. Red labels correspond to fluorinated monomers; black labels to non-fluorinated monomers. The ^{19}F and ^{19}F-^{1}H decoupled spectra were collected using different instruments.*

Fluorohydroxybutyrate can be incorporated into bioplastic. We next investigated whether expression of PhaC, a 3-hydroxyalkanoyl-CoA polymerase, would lead to formation of a novel 2-fluorinated polyester. To analyze intracellular polymer content, we carried out either

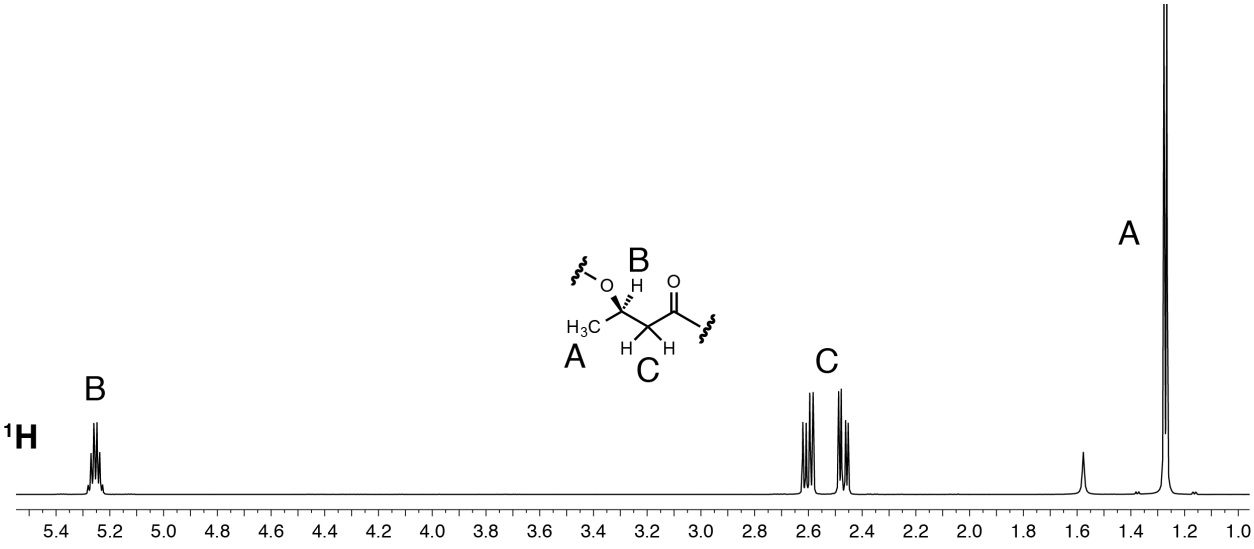


Figure 3.19. *^{1}H NMR spectrum of poly((R)-3-hydroxybutyrate) produced by the NphT7 pathway (without fluoromalonate feeding).*

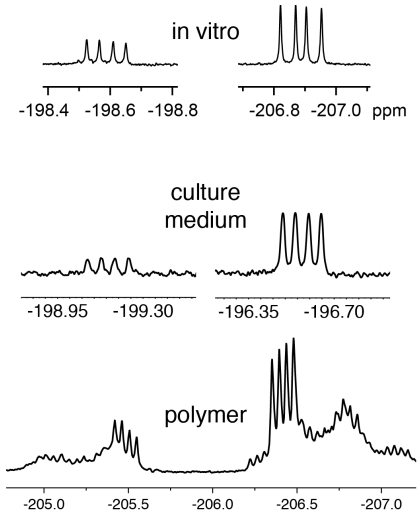


Figure 3.20. *Diastereomers of 2-fluoro-(R)-3-hydroxybutyrate from the in vitro reactions of NphT7 and PhaB (as the acyl-CoAs; reproduced from [3]), culture medium and polymer. The peak positions of the diastereomers are inverted in culture medium (note ppm scales), but since the peak ratios are identical, this is likely due to CoA thioesterification and/or differences in pH and ionic strength.*

whole-cell acidolysis in sulfuric acid to depolymerize PHAs [*31*] followed by LC-MS quantitation of monomer-derived species, or isolated and purified PHA by removing other cellular components with concentrated bleach and detergent, [*4*] followed by extensive washing. Expression of PhaC lead to robust formation of polymer and we found that fluorinated monomers were incorporated when fluoromalonate was fed (*Figure 3.17*). However, fluoromonomer content was relatively low (< 20%), and 3-hydroxybutyrate, which presumably arises from the endogenous pool of malonyl-CoA (*ca.* 35 µM [*32*]) supplying pathway enzymes with their native substrates, comprised the bulk of the polymer. We also found that overall polymer yields were reduced by 2-fold when fluoromalonate was included, and that only 34% of the total 2-fluoro-3-hydroxybutyryl-CoA produced was polymerized (in contrast to 88% of the 3-hydroxybutyryl-CoA polymerized when no fluoromalonate was fed). This reduction could be due either to preferential hydrolysis of the fluorinated acyl-CoAs by PhaC or an endogenous thioesterase, or to slower incorporation of fluorinated monomers by PhaC. Overall 2-fluoro-3-hydroxybutyrate production was similar whether PhaC was included in the pathway or not (compare *Figure 3.17* and *Figure 3.13*, pPOL1-nb), suggesting that PhaC itself does not contribute substantially to hydrolysis.

To verify fluorinated monomer incorporation, we purified polymer and analyzed it by NMR. The ^{1}H, ^{19}F and ^{13}C spectra all show peaks corresponding to fluorinated monomers, with characteristically large $J_{H\text{-}F}$ and $J_{C\text{-}F}$ coupling constants (*Figure 3.18*). Comparison of the ^{1}H

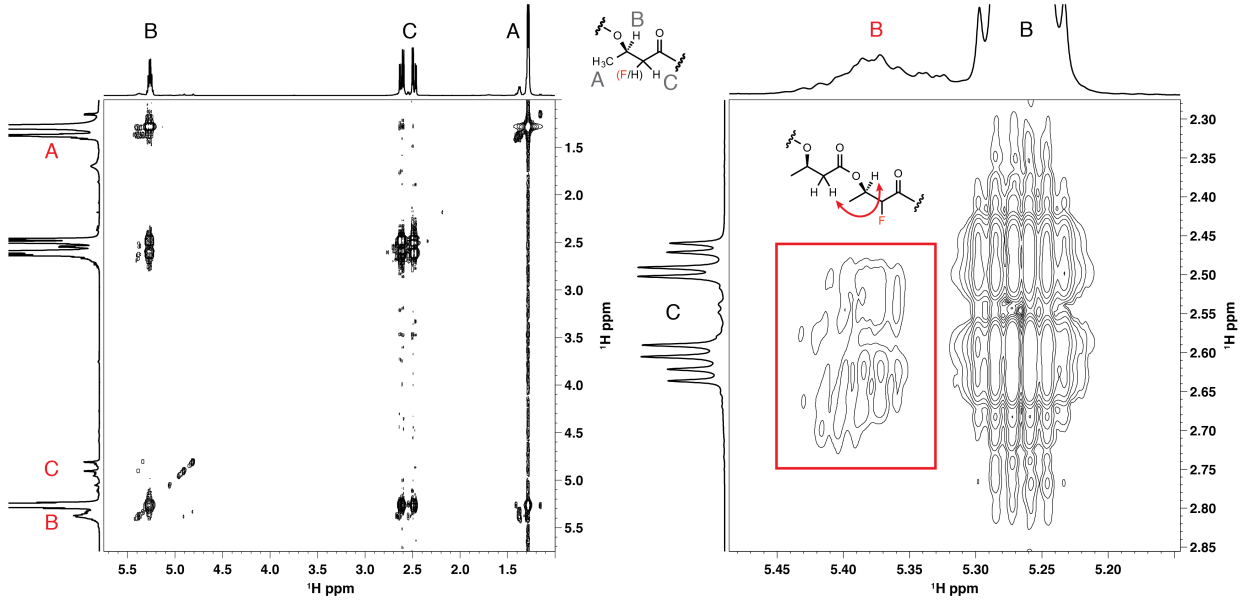


Figure 3.21. *Full COSY spectrum of poly(2-fluoro-3-hydroxybutyrate-co-3-hydroxybutyrate) and expansion showing evidence of co-polymerization.*

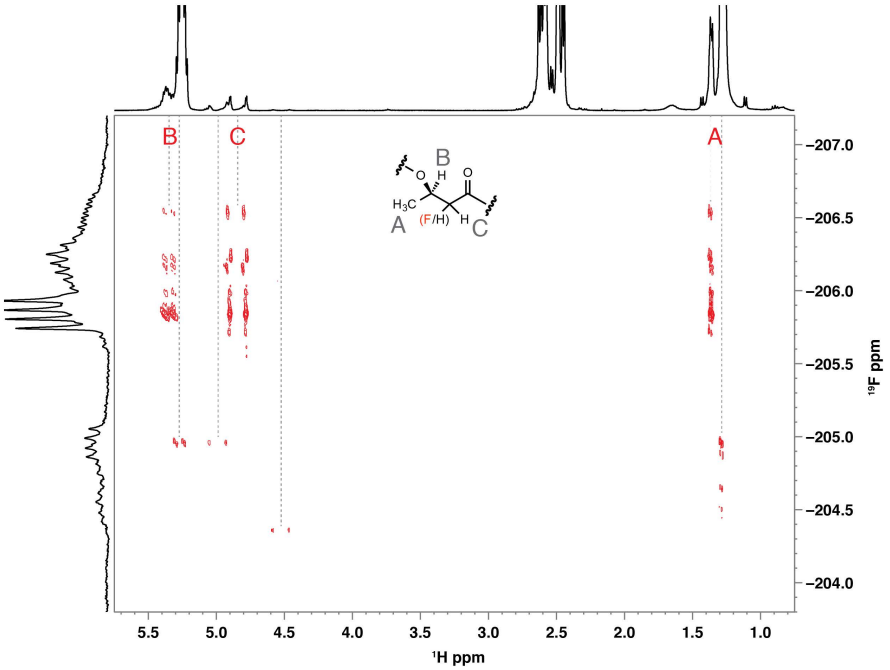


Figure 3.22. *^{1}H-^{19}F HMBC spectrum of poly(2-fluoro-3-hydroxybutyrate-co-3-hydroxybutyrate).*

spectrum of polymer isolated from cells cultured with (*Figure 3.18*) and without fluoromalonate (*Figure 3.19*) show identical poly(3-hydroxybutyrate) peaks. Notably, both diastereomers of 2-fluoro-(*R*)-3-hydroxybutyryl-CoA appear to be incorporated into polymer, with a ratio similar to that observed in culture medium and in the *in vitro* reactions of NphT7 and PhaB (*Figure 3.20*), suggesting that PhaC does not distinguish the stereochemistry of the 2-fluoro substituent. Because acetofluoroacetyl-CoA is expected to epimerize rapidly under biological conditions, resulting in a nearly 1:1 mixture of diastereomers (since the nearest stereocenter is 12 bonds away), the observed diastereomeric ratio presumably represents the substrate selectivity of PhaB. It is unlikely to be the result of epimerization of 2-fluoro-3-hydroxybutyryl-CoA or 2-fluoro-3-hydroxybutyrate, because the diastereomers of 2-fluoro-3-hydroxybutyric acid (two *syn* enantiomers and two *anti* enantiomers) are resolved under our LC conditions, and we observe a different ratio in polymer acidolysate (1.5 ± 0.2, mean ± standard deviation) compared to culture medium (3.1 ± 0.4) (which matches the ratio observed by ^{19}F NMR of 2.8 ± 0.4, suggesting that ionization and fragmentation efficiency is similar between the two diastereomers), and this would not be expected if epimerization is facile.

The fluoromonomer content is too low to show 1D ^{1}H or ^{13}C peaks corresponding to non-fluorinated monomers adjacent to fluorinated ones, but the COSY spectrum does have one such crosspeak (*Figure 3.21*), which indicates a fluorinated monomer used to extend a non-fluorinated chain. The ^{1}H-^{19}F HMBC spectrum also shows crosspeaks at ^{1}H chemical shifts very close to those of nonfluorinated monomers (*Figure 3.22*). Finally, the ^{19}F[^{1}H] spectrum (*Figure 3.18*), which shows every unique fluorine as a singlet, makes it clear that fluoromonomers exist in many distinct chemical environments within the polymer.

Runs of fluoromonomers seem unlikely, but the possibility exists that fluorinated monomers are chain-terminating, due either to slowness of subsequent extensions or the lability of the α-fluoro (thio)ester. To investigate this possibility, we tested the molecular weight of polymers

produced by our system that contained either no fluorinated monomers or 7% fluorinated monomers using GPC (*Table 3.1*). The molecular weights and dispersity indices were essentially identical, which given the considerable fluoromonomer content implies that chain termination is not significantly more likely following fluoromonomer incorporation. This is promising for the potential to synthesize higher fluoromonomer content materials. However, all other properties that we examined (*Table 3.1*) also showed no differences between the polymers (including thermal properties determined by DSC, data not shown), so higher fluoromonomer content is apparently required to assess whether poly(2-fluoro-3-hydroxybutyrate) has any characteristics of interest.

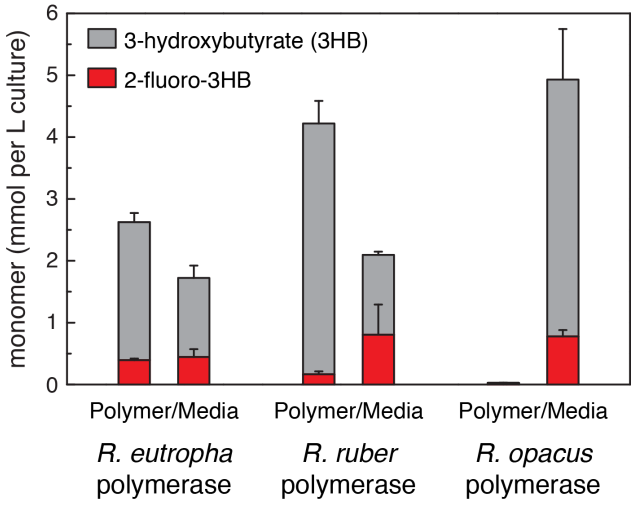


Figure 3.23. *Monomer content of polymer and growth medium using alternative polymerases. Error bars show sample standard deviation, n=3.*

To attempt to improve fluoromonomer incorporation, we synthesized genes for alternative polymerases, two PhaC homologs from organisms that had previously been reported to polymerize 2-substituted monomers [*18*]. However, neither performed better than the original PhaC from *Ralstonia eutropha* (*Figure 3.23*). We also added ligases in an operon after *Rpa*matB on pXHB1: two homologs of AlkK, an enzyme reported to ligate medium chain length 3-hydroxy acids to CoA in *Ralstonia* and related PHA-producing organisms [*33*], from *Streptomyces coelicolor* and *Pseudomonas oleovorans*; and acetyl-CoA synthetase (ACS) and propionyl-CoA synthetase (PrpE) from *E. coli*. Although some of these enzymes reduced media titers of 2-fluoro-3-hydroxybutyrate, a possible result of its reattachment to CoA, all of them reduced the fluoromonomer content of intracellular polymer, suggesting a deleterious effect on overall organofluorine production (data not shown).

Table 3.1. Properties of polymers produced using the NphT7-based pathway.

	Poly((*R*)-3-hydroxybutyrate)	**Poly(2-fluoro-(*R*)-3-hydroxybutyrate-*co*-(*R*)-3-hydroxybutyrate)**
Mol % fluoromonomer	-	7
Number average molecular weight (M_n, Da)*	5.4×10^5	6.0×10^5
Dispersity (M_n/M_w)	1.62	1.69
Decomposition temperature (°C)	260	260

** Bleach exposure during purification is known to slightly reduce polymer molecular weight [4].*

3.4 Conclusions

The model system we have developed is capable of supporting mM flux of organofluorine biosynthesis in *E. coli*. We used it to identify functional fluoromalonate transporters that should be portable to other organisms. We found that toxicity from the synthetic pathway is minimal, and can be attributed to fluoroacetyl-CoA formation. This metabolite can either be detoxified using FlK or potentially redirected to the pathway using e.g. an engineered ACCase. Endogenous metabolism still competes favorably with the organofluorine pathway, underlining the need for developing enzymes with improved fluorine tolerance and specificity. These results set the stage for biosynthesis of fluorinated analogues of complex, high-value compounds such as bioactive natural products.

3.5 References

1. Hagmann, W. The many roles for fluorine in medicinal chemistry. *J Med Chem* **2008**, *51* (15) 4359-4369.
2. Purser, S., Moore, P., Swallow, S., Gouverneur, V. Fluorine in medicinal chemistry. *Chem Soc Rev* **2008**, *37* (2) 320-330.
3. 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* (6150) 1089-1094.
4. Hahn, S. K., Chang, Y. K., Lee, S. Y. Recovery and characterization of poly(3-hydroxybutyric acid) synthesized in *Alcaligenes eutrophus* and recombinant *Escherichia coli*. *Appl Environ Microbiol* **1995**, *61* (1) 34-39.
5. Furuya, T., Kuttruff, C. A., Ritter, T. Carbon-fluorine bond formation. *Curr Opin Drug Discovery Dev* **2008**, *11* (6) 803-819.
6. Okamura, E., Tomita, T., Sawa, R., Nishiyama, M., Kuzumaya, T. Unprecedented acetoacetyl-coenzyme A synthesizing enzyme of the thiolase superfamily involved in the mevalonate pathway. *Proc Natl Acad Sci USA* **2010**, *107* (25) 11265-11270.
7. Jiang, M., Pfeifer, B. A. Metabolic and pathway engineering to influence native and altered erythromycin production through *E. coli*. *Metab Eng* **2013**, *19* (C) 42-49.
8. Lauble, H., Kennedy, M. C., Emptage, M. H., Beinert, H., Stout, C. D. The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme-inhibitor complex. *Proc Natl Acad Sci USA* **1996**, *93* (24) 13699-13703.
9. Walker, M. C., Wen, M., Weeks, A. M., Chang, M. C. Y. Temporal and fluoride control of secondary metabolism regulates cellular organofluorine biosynthesis. *ACS Chem Biol* **2012**, *7* (9) 1576-1585.
10. Walker, M. C., Chang, M. C. Y. Natural and engineered biosynthesis of fluorinated natural products. *Chem Soc Rev* **2014**, *43* (18) 6527-6536.
11. Brandl, H., Gross, R. A., Lenz, R. W., Fuller, R. C. Plastics from bacteria and for bacteria: Poly(beta-hydroxyalkanoates) as natural, biocompatible, and biodegradable polyesters. *Adv Biochem Eng Biotechnol* **1990**, *41* 77-93.

12. Lu, J., Tappel, R. C., Nomura, C. T. Mini-review: Biosynthesis of poly(hydroxyalkanoates). *Polym Rev* **2009**, *49* (3) 226-248.

13. Stubbe, J. J., Tian, J. Polyhydroxyalkanoate (PHA) homeostasis: The role of the PHA synthase. *Nat Prod Rep* **2003**, *20* (5) 445-457.

14. Agnew, D. E., Pfleger, B. F. Synthetic biology strategies for synthesizing polyhydroxyalkanoates from unrelated carbon sources. *Chem Eng Sci* **2013**, *103* (C) 58-67.

15. Bengtsson, S., Pisco, A. R., Reis, M. A. M., Lemos, P. C. Production of polyhydroxyalkanoates from fermented sugar cane molasses by a mixed culture enriched in glycogen accumulating organisms. *J Biotechnol* **2010**, *145* (3) 253-263.

16. Fuchtenbusch, B., Fabritius, D., Steinbüchel, A. A. Incorporation of 2-methyl-3-hydroxybutyric acid into polyhydroxyalkanoic acids by axenic cultures in defined media. *FEMS Microbiol Lett* **1996**, *138* (2-3) 153-160.

17. Pisco, A. R., Bengtsson, S., Werker, A., Reis, M. A. M., Lemos, P. C. Community structure evolution and enrichment of glycogen-accumulating organisms producing polyhydroxyalkanoates from fermented molasses. *Appl Environ Microbiol* **2009**, *75* (14) 4676-4686.

18. Fuchtenbusch, B., Fabritius, D., Waltermann, M., Steinbüchel, A. A. Biosynthesis of novel copolyesters containing 3-hydroxypivalic acid by *Rhodococcus ruber* NCIMB 40126 and related bacteria. *FEMS Microbiol Lett* **1998**, *159* (1) 85-92.

19. Matsumoto, K., Yamada, M., Leong, C. R., Jo, S.-J., Kuzumaya, T., Taguchi, S. A new pathway for poly(3-hydroxybutyrate) production in *Escherichia coli* and *Corynebacterium glutamicum* by functional expression of a new acetoacetyl-coenzyme A synthase. *Biosci Biotechnol Biochem* **2011**, *75* (2) 364-366.

20. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., Mori, H. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Mol Syst Biol* **2006**, *2* 2006.0008.

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

22. Rouillard, J.-M., Lee, W., Truan, G., Gao, X., Zhou, X., Gulari, E. Gene2oligo: Oligonucleotide design for *in vitro* gene synthesis. *Nucleic Acids Res* **2004**, *32* (Web Server issue) W176-180.

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

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

25. Matsumoto, K., Tanaka, Y., Watanabe, T., Motohashi, R., Ikeda, K., Tobitani, K., Yao, M., Tanaka, I., Taguchi, S. Directed evolution and structural analysis of NADPH-

dependent acetoacetyl coenzyme A (acetoacetyl-coa) reductase from *Ralstonia eutropha* reveals two mutations responsible for enhanced kinetics. *Appl Environ Microbiol* **2013**, *79* (19) 6134-6139.

26. Saier, M. H., Reddy, V. S., Tamang, D. G., Västermark, A. The transporter classification database. *Nucleic Acids Res* **2014**, *42* (Database issue) D251-258.

27. 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* (10) 2689-2693.

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

29. 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* (43) 9269-9279.

30. Weeks, A. M., Keddie, N. S., Wadoux, R. D. P., O'Hagan, D., Chang, M. C. Y. Molecular recognition of fluorine impacts substrate selectivity in the fluoroacetyl-CoA thioesterase FlK. *Biochemistry* **2014**, *53* (12) 2053-2063.

31. Karr, D. B., Waters, J. K., Emerich, D. W. Analysis of poly-beta-hydroxybutyrate in *Rhizobium japonicum* bacteroids by ion-exclusion high-pressure liquid chromatography and UV detection. *Appl Environ Microbiol* **1983**, *46* (6) 1339-1344.

32. Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., Rabinowitz, J. D. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat Chem Biol* **2009**, *5* (8) 593-599.

33. Satoh, Y., Murakami, F., Tajima, K., Munekata, M. Enzymatic synthesis of poly(3-hydroxybutyrate-co-4-hydroxybutyrate) with CoA recycling using polyhydroxyalkanoate synthase and acyl-CoA synthetase. *J Biosci Bioeng* **2005**, *99* (5) 508-511.

Chapter 4: *Engineered* trans*-acting acyl transferases: a strategy for efficient, regioselective fluorine incorporation into polyketides*

Portions of this work were performed in collaboration with the following persons:

Mark Walker, PhD helped to design screening strategies and carried out acyl transferase phylogenetic classification. Jacquelyn Blake-Hedges, Tammy Hsu and Omer Ad assisted with *trans*-acyl transferase expression and screening in cell lysate. Ningkun Wang cloned LkcD and VirI constructs.

4.1. Introduction

Trans-ATs are enzymes homologous to fatty acid synthase ATs that have evolved to act *in trans* as part of so-called *trans*-AT PKS clusters, [*4*] which are much less well-characterized than *cis*-AT clusters such as DEBS. In these type I systems, *cis*-AT function within each module has been lost (although fragments of the *cis*-AT domain may still be distinguishable), and all ACPs within the cluster are acylated with extender units by one or more *trans*-ATs. *Trans*-AT PKSs are also distinguished by other fascinating characteristics: modules sometimes possess tandem ACPs, which may or may not be essential for function; the order of chain extension events often diverges from the sequence of modules; and ketosynthase domains are more closely related by substrate specificity than by cluster of origin [*5,6*]. In addition, *trans*-AT clusters often include a domain (fused to a *trans*-AT or as a separate protein) homologous to *cis*-ATs, which is thought to function not as an acyl transferase but as an editing protein, hydrolyzing stalled or incorrect intermediates from the assembly line [*8*]. *Trans*-ATs may also be fused to a *trans*-acting enoyl-reductase domain, phylogenetically unrelated to *cis*-ERs, which functions analogously to *trans*-ATs [*10*].

Since *trans*-ATs have evolved to act *in trans*, they have compensated for loss of the high effective concentration that comes with physical tethering to their ACP substrates. These enzymes have catalytic efficiencies on the order of 10^3 to 10^4 M^{-1} s^{-1}. [*9,11*] while *cis*-ATs range from 10^1 to 10^2 M^{-1} s^{-1} when acting *in trans* [*9,12*]. This high efficiency makes them attractive targets for evolution: mutations that alter specificity often compromise catalytic activity, but *trans*-ATs can accommodate considerable activity loss while still remaining effective. *Trans*-ATs have been shown to efficiently acylate ACPs (or thiolation domains, the analogous component of non-ribosomal peptide synthetases) from other clusters, including *cis*-AT ACPs, and have been used to complement AT^0 modules [*11-14*]. We therefore aim to use *trans*-ATs to accomplish our goal of regioselective fluorinated extender incorporation by polyketide synthases (*Figure 4.1*).

Almost all known *trans*-ATs are malonyl-CoA specific. The exceptions are single examples of ATs that load rare extender units, ethylmalonyl [*15*] or methoxymalonyl [*16*]; no methylmalonyl-specific ATs have been characterized. KirCII, which loads ethylmalonyl units, does in fact have relaxed extender specificity that has been used in polyketide synthesis, but its kinetics are poor, especially compared to other *trans*-ATs, and its lack of specificity may actually be problematic [*17-19*]. Fortunately, our positive results using DszAT to load fluoromalonyl units onto DEBS ACPs (*Section 2.3*), compared to the lack of activity of methylmalonyl-specific DEBS *cis*-ATs on fluoromalonyl-CoA, suggest that malonyl-specific ATs are a promising area to explore.

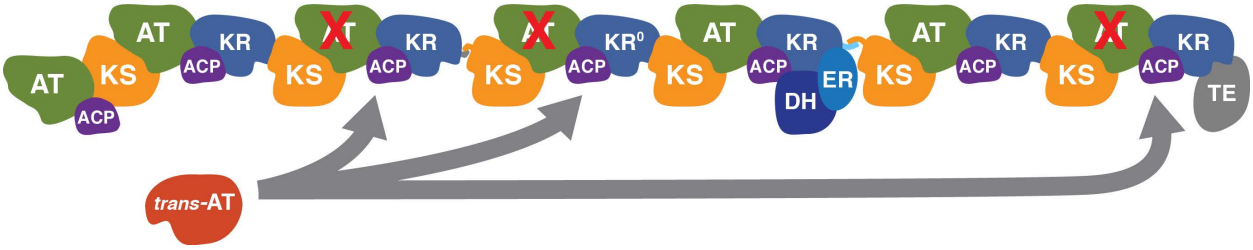


Figure 4.1. Trans-*AT complementation in the context of a full PKS assembly line.* Cis-*ATs load their native extender units, while modules with inactivated* cis-*ATs are loaded with fluoromalonyl extenders by the* trans-*AT.*

Rather than pursuing directed evolution studies on the first *trans*-AT we tested, we hope to explore existing sequence diversity in this enzyme family, which may result in differences in activity. By testing a small library of *trans*-ATs, we hope to find ATs with improved activity on fluoromalonyl-CoA, or a reduced preference for malonyl-CoA over fluoromalonyl-CoA, which may be more promising candidates for evolution than DszAT.

Only a handful of *trans*-ATs have been characterized biochemically. DszAT can acylate diverse ACPs, including those from DEBS, as described by our group and others [*9,14,17*]. FenF, which acts in mycosubtilin biosynthesis in *Bacillus*, can efficiently acylate ACPs and even peptidyl carrier proteins with K_m values on the order of 1 mM [*11*]. BryP-AT1 is a slower enzyme but it can also acylate heterologous ACPs, e.g. from bryostatin and pikromycin synthases [*7*]. KirCI-AT2 has been established as an active enzyme; it acylates all ACPs within its biosynthetic cluster except the one served by KirCII [*3*]. Finally, the AT from fatty acid synthesis in *Streptomyces coelicolor*, FabD, has also been shown to acylate DEBS ACPs [*20*].

Additional *trans*-ATs have not been characterized individually but have been confirmed to be part of an active *trans*-AT PKS. These include DifA [*21*], PedD [*22*], LkcD [*23*], VirI [*24*], LnmG [*25*], TaV-AT2 [*26,27*], MmpC [*28*], RhiG [*29*], OzmM [*16*], ChiA [*30*], PksE [*10*] and MlnA [*31*]. Metagenomic sequencing of environmental DNA has furnished a wealth of new secondary metabolite clusters, including *trans*-AT PKSs. These sequences are another source for library members, although their positive classification as PKS ATs is dependent on genetic context from correct cluster assembly, since they are homologous to fatty acid synthase ATs.

Our screening strategy for a *trans*-AT library could encompass both *in vitro* and *in vivo* testing. An *in vivo* approach requires more initial investment but has the advantages of being both more scalable and much more relevant, since it is essentially the situation in which these enzymes will be used. The work described in Chapter 3 should make it possible to generate sufficient fluoromalonyl-CoA for *in vivo* fluorinated extender use. For *in vivo* complementation assays, *trans*-ATs from the library will be expressed along with a full PKS that has AT active site mutation(s). DEBS again provides a convenient model system, since its efficient expression has been achieved in both the native producer [*32*] and a more easily genetically manipulated heterologous *Streptomyces* host, *S. coelicolor* [*1*]. We envision establishing 6-dEB production in such a host – strains of *S. lividans* [*33*], *S. coelicolor* [*34*] and *S. avermitilis* [*35*] have been developed that lack most biosynthetic clusters and therefore have enhanced secondary metabolite production potential and reduced background – and constructing all six possible AT mutants of DEBS (*Figure 4.2*). Expression of library members in all six backgrounds should provide a stringent test of *in vivo* complementation and reveal any differences in efficiency in forming the six fluorinated 6-dEB regioisomers arising from module, backbone position or adjacent site chemistry. These may include complementation of AT^0 mutant modules by other *cis*-ATs within DEBS [*13*], acylation at undesired positions [*17*], lack of *trans*-AT activity on specific ACPs, occlusion of ACPs by the full PKS homodimeric complex, or interference from native metabolism.

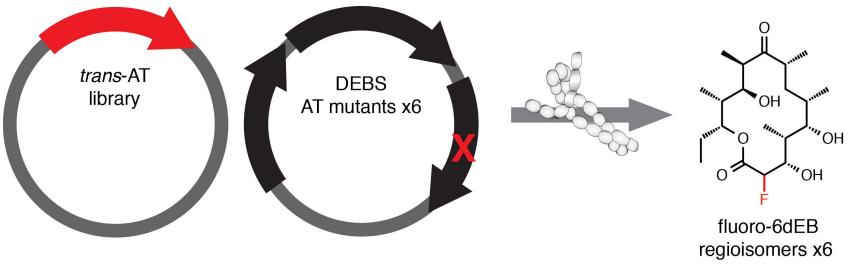


Figure 4.2. In vivo *library expression and combinatorial complementation scheme for fluorinated 6-dEB production and* trans-*AT screening.*

4.2 Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, Terrific Broth (TB), yeast extract, malt extract and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). Carbenicillin (Cb), isopropyl-β-D-thiogalactopyranoside (IPTG), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride, dithiothreitol (DTT), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, kanamycin (Km), acetonitrile and ethylene diamine tetraacetic acid disodium dihydrate (EDTA), were purchased from Fisher Scientific (Pittsburgh, PA). Coenzyme A trilithium salt (CoA), diethylfluoromalonate, malonic acid, methylmalonic acid, tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, phosphoenolpyruvate (PEP), adenosine triphosphate sodium salt (ATP), myokinase, pyruvate kinase, lactate dehydrogenase, poly(ethyleneimine) solution (PEI), β-mercaptoethanol, sodium phosphate dibasic hepthydrate and N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED) were purchased from Sigma-Aldrich (St. Louis, MO). Formic acid was purchased from Acros Organics (Morris Plains, NJ). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate and ammonium persulfate were purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, T4 DNA ligase, Antarctic phosphatase, Phusion DNA polymerase, T5 exonuclease, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) and Platinum Taq High-Fidelity polymerase (Pt Taq HF) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA), resuspended at a stock concentration of 100 μM in 10 mM Tris-HCl, pH 8.5, and at 4°C. DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). Complete EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Amicon 5,000 MWCO regenerated cellulose ultrafiltration membranes were purchased from EMD Millipore (Billerica, MA).

Bacterial strains. *E. coli* DH10B-T1^R and BL21(de3)T1^R were used for DNA construction and heterologous protein production, respectively, except for DEBS modules, which were heterologously expressed in *E. coli* BAP1 [*36*].

Gene and plasmid construction. Standard molecular biology techniques were used to carry out plasmid construction. All PCR amplifications were carried out with Phusion or Platinum Taq High Fidelity DNA polymerases. For amplification of GC-rich sequences, PCR reactions were supplemented with DMSO (5%) using the standard buffer rather than GC buffer with primer annealing temperatures 8-10°C below the T_m. All constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA).

Assembly of cloning vectors. To construct pCDF-pUWL-shuttle, a cassette encoding pTet.rfp was amplified from pWCD0941 (a gift from Will DeLoache, UC Berkeley) using primers pUWL-sgg CF1/CR1 and cloned into pCDF2.P(tac.tac)(gg) (a BsaI-acceptor vector) by BsaI Golden Gate (following the protocol below). To construct pSET152(gg), the same cassette was PCR amplified using primers pSET152gg CF1/CR2 and inserted between the EcoRV/BamHI sites of pSET152 [*37*] by Gibson assembly [*38*], resulting in a BsaI Golden Gate acceptor vector. To construct pSET152r(gg), *rop* was amplified from pET28a using primers pSET152-ROP-F1/R1 and inserted into the NheI site of pSET152(gg) by Gibson assembly. An analogous

procedure was used to convert pSET152-pc13 into pSET152r-pc13. To construct pET28a.HisN(sgg), the single backbone SapI site was first removed by SapI digestion and Gibson assembly insertion of the oligos pET.SapI_del F1/R1. The resulting vector, pET28a*SapI, was digested with NdeI/EcoRI and the NdeI/EcoRI fragment of pCDF-pUWL-shuttle, containing the SapI-acceptor RFP cassette, was inserted by ligation.

Assembly of trans-*AT library.* Genes were reverse-translated using IDT's codon optimization for *Streptomyces coelicolor*. Hairpins and high-G+C runs were removed manually as necessary to pass pre-synthesis QC. gBlocks (*Appendix Table 3, C*) were assembled by Golden Gate cloning into pCDF-pUWL-shuttle and pET28a.HisN(sgg) in parallel using SapI according to the J5 Manual protocol (j5.jbei.org). Reaction mixtures (15 µL scale) contained 10X NEB T4 DNA ligase buffer (1.5 µL), BSA (1 mg/mL, 1.5 µL), NEB SapI (10 U, 1 µL), NEB T4 DNA ligase (2,000 U, 1 µL) and ~30 ng/kb of each assembly piece. Reactions were routinely carried out at 7.5 µL scale. The assembly program was 20-40 cycles of (37ºC for 3 min and 16ºC for 4 min), followed by 50ºC for 5 min and 80ºC for 5 min. The entire mixture was transformed into KCM chemically competent cells.

ATs cloned into pCDF-pUWL-shuttle were sequence verified, then digested with XbaI, which mobilizes the AT with 25 bp homology to pUWL201PW [*39*] at both termini. The resulting fragments were inserted between the NdeI/HindIII sites of pUWL201PW by Gibson assembly.

*Sco*fabD was PCR amplified from *Streptomyces coelicolor A3(2)* M13 genomic DNA using primers FabD.GF1/GR1 and inserted between the NdeI/BamHI sites of pET16b by Gibson assembly.

Assembly of vectors for cloning eryAI-II-III. Initial attempts at a 4-piece Golden Gate assembly to build pSET152r-preEryAI+TE were unsuccessful, so pSET152-pc13, which lacks the cat-sacB cassette, was assembled first. PCR amplification of p*eryAI* and the 5' end of the eryAI gene failed, probably because of the short length of the fragment, so it was instead ordered as a gBlock. This fragment was assembled together with the 3' homology region to eryAIII, which was PCR amplified from pRSG64 using primers pSETEryAI F3b/R3, into pSET152(gg) by BsaI Golden Gate cloning. The cat-sacB cassette contained a single internal BsaI site, so this was removed by SOE PCR amplification from a colony of *E. coli* TUC01 [*2*] using primer pairs cat-sacB SOE F1/R1 (internal) and pSETEryAI F2b/R2b (external). The resulting fragment was inserted into the NheI site of pSET152r-pc13 by Gibson assembly, giving pSET152r-preEryAI+TE. Sucrose sensitivity conferred by the vector was verified during cloning; only about half of the candidates that incorporated the amplified cat-sacB cassette retained the phenotype.

Expression and purification of proteins. *Rpa*MatB, DszAT, DEBS-Mod6AT^0+TE and methylmalonyl-CoA epimerase were expressed and purified as described above (*Section 2.2*).

Expression of *trans*-ATs in *E. coli*. TB (1 L) containing kanamycin (50 µg/mL) in a 2.8 L Fernbach baffled shake flask was inoculated to OD_{600} = 0.05 with an overnight TB culture of freshly transformed *E. coli* containing the appropriate overexpression plasmid. The cultures were grown at 37°C to OD_{600} = 0.6 to 0.8 at which point cultures were cooled on ice for 20 min, followed by induction of protein expression with IPTG (0.4 mM) and overnight growth at 16°C. Cell pellets were harvested by centrifugation at 9,800 × *g* for 7 min at 4ºC and stored at -80ºC.

Purification of FenF and PksC. FenF and PksC were purified essentially as described for DszAT (*Section 2.2*). Concentrations were estimated using extinction coefficients calculated using ExPASY ProtParam (reduced form).

Trans-AT lysate assays. Frozen cell pellets were thawed and resuspended at 500 µL/100 mg cell paste in 400 mM sodium phosphate, pH 7.5. The mixture was resuspended by pipetting and transferred to a screw-top vial. Glass beads (1 mM) were added (~250 µL) and the cells were lysed by bead beating for 2 x 30 s at maximum speed using a Mini-Beadbeater-24 (Biospec Products, Bartlesville, OK). The lysate was clarified by centrifugation at 23,000 × *g* for 20 min at 4°C. The clarified lysate constituted 50% of each reaction mixture. Assay mixtures (60-100 µL) contained 400 mM sodium phosphate (total concentration), pH 7.5, phosphoenolpyruvate (25 mM), TCEP (5 mM), magnesium chloride (10 mM), ATP (2.5 mM), pyruvate kinase (27 U/mL), myokinase (10 U/mL), CoA (1 mM), methylmalonyl-CoA epimerase (5 µM), *Rpa*MatB (20 µM) and fluoromalonate or malonate (10 mM). Purified *trans*-AT was added (5-10 µM) as indicated. Mixtures were incubated at 37°C for 15-30 min and then initiated by addition of the N-acetylcysteamine thioester of (2*S*,3*R*)-2-methyl-3-hydroxypentanoic acid (NDK-SNAC, 5 mM) and DEBS-Mod6AT^0+TE (5 µM). After incubation at 37°C for 18-22 h, aliquots (35 µL) were removed and quenched by addition of 70% perchloric acid (1.75 µL). Samples were centrifuged at 18,000 × *g* to pellet the precipitated protein. The supernatant (33 µL) was removed and added to 1 M sodium bicarbonate (6.6 µL) bringing the final pH to 4-5. Excess salts were precipitated by freezing in liquid nitrogen and centrifuging at 18,000 × *g* until thawed. The supernatant was removed and analyzed on a Zorbax Eclipse XDB C-18 column as described above (*Section 2.2*). Fluoro-triketide and desmethyl-triketide were detected using an Agilent single quadrupole mass spectrometer operating in single ion monitoring mode (negative ion mode) and quantified by comparison to external standard curves using standards prepared as described (*Section 2.2*).

4.3 Results and discussion

***Trans*-AT construct design and cloning.** As a preliminary step to construction of a full library of *trans*-ATs, we began with the set of *trans*-ATs from verified PKS clusters. Gene synthesis and cloning of these proteins is simplified by isolating the active AT domains – especially in cases where tandem ATs are of the proofreading type. We truncated the sequences to remove *trans*-ER domains and tandem AT domains based on a multiple sequence alignment (*Figure 4.3*), maintaining native termini where appropriate and following experimentally verified truncations when available. These proteins have percent identity ranging from 31 to 67% across their malonyl-specific AT domains and likely encompass a variety of activities.

Streptomyces genomes have high G+C content, which leads to extreme codon bias, especially at the wobble position. To make our library compatible with *Streptomyces* expression, we chose codons optimized for *S. coelicolor*. This introduced some difficulties for gene synthesis of these often repetitive, G+C-rich sequences. When needed, automated reverse translations were

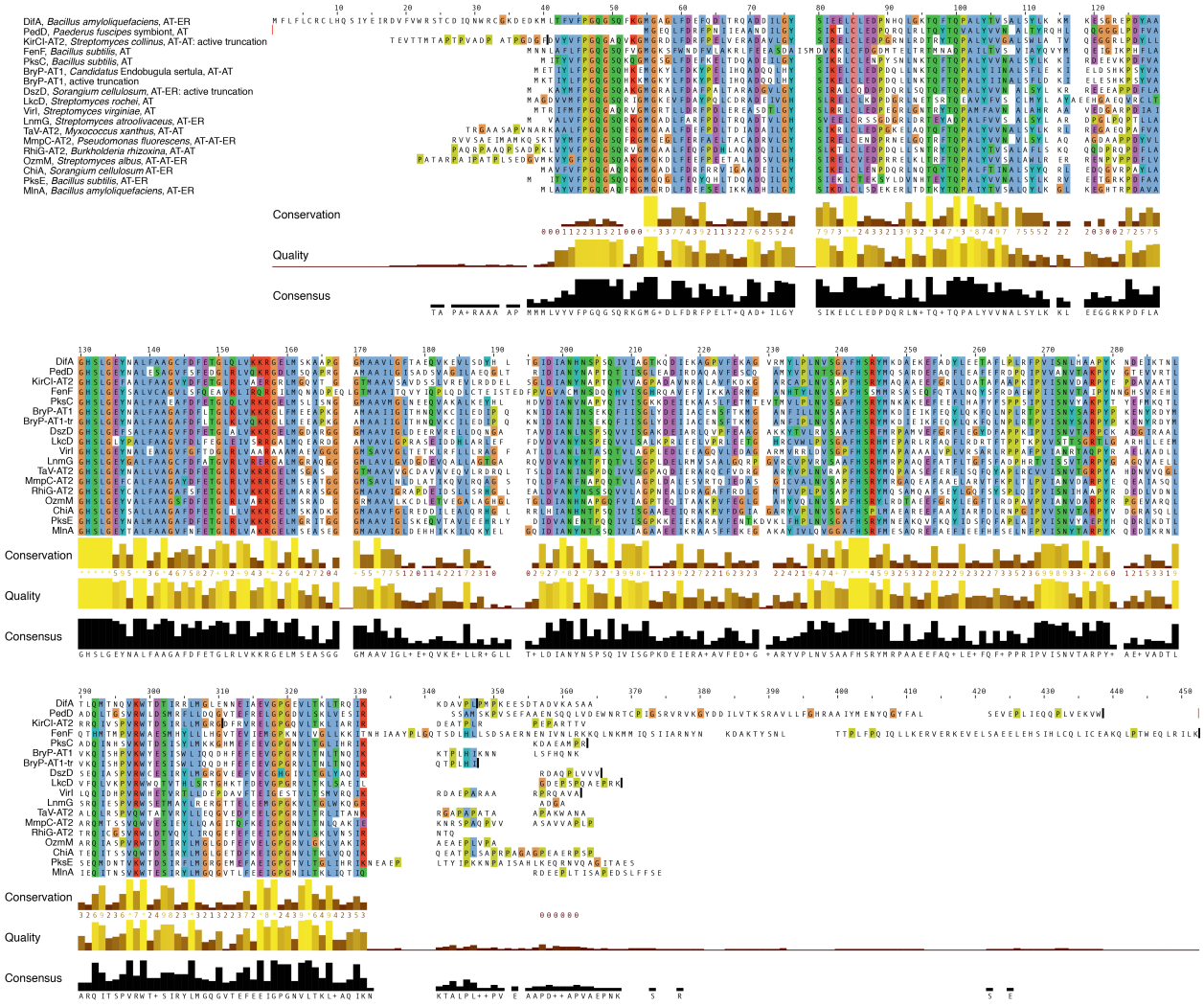


Figure 4.3. *Alignment of* trans*-AT sequences from verified PKS clusters produced using PRALINE with ClustalW coloring. The source protein architecture (e.g. stand-alone AT, AT-*trans*-ER fusion, tandem-AT) is indicated after the protein name and source organism, but only the FabD-homologous AT regions are shown. Experimentally characterized truncations are shown for KirCI [3], BryP [7] and DszD (DszAT) [9]. The start and end points of synthesized genes, where different from the regions shown, are indicated with vertical black bars.*

adjusted manually to remove long GC runs and stable predicted hairpins. The next most common codons were chosen when possible.

To clone the library, a two-stage pipeline was developed. Synthetic DNA fragments were first assembled into a shuttle vector by Golden Gate cloning. Most type IIs restriction enzymes (necessary for the Golden Gate method) have high G+C recognition sequences that are common in *Streptomyces* DNA, which includes both our back-translated genes and many available expression plasmids. Therefore we used SapI, which has a 7-nt recognition site that is also 43% T and therefore relatively rare in G+C rich DNA, for cloning. Its overhangs are only 3 nt long, but this potential drawback seems to be compensated for by higher enzyme activity than the traditional BsaI. Reaction mixtures on 7 µL scale containing ~25 nmol of each DNA fragment routinely gave hundreds of positive colonies.

Our destination vector, a replicating *Streptomyces* plasmid, did however contain >5 SapI sites. Therefore we designed the shuttle vector so that cloned inserts could be cut out by XbaI digest with 25 bp of homology to the destination vector appended on either end. These restriction fragments were then assembled with the cut destination vector using the Gibson method [*38*] without the need for PCR. For further characterization, AT genes were also cloned in parallel directly from synthetic DNA into a version of the pET28a expression vector adapted for SapI Golden Gate.

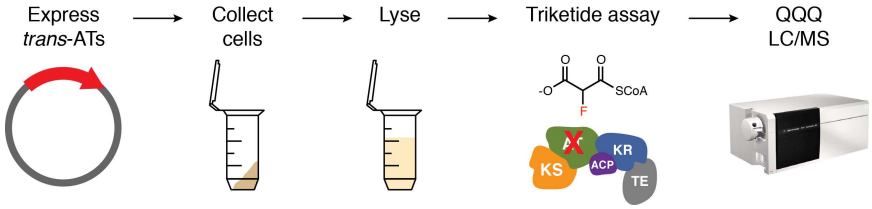


***Figure 4.4**. Workflow for* in vitro trans-*AT triketide formation assay in cell lysate.*

***In vitro* triketide assays.** We expressed a selection of library members in *E. coli* and tested complementation of a single DEBS AT^0 module in cell lysate (*Figure 4.4*). No AT purification is needed to see activity in cases where expression level and solubility are adequate (although this was not true for all ATs). Triketide formation is minimal in the *cis*-AT^0 background without a *trans*-AT, but adding pure *trans*-AT or cell lysate containing *trans*-AT, along with the fluoromalonyl-CoA regeneration system described in Scheme 2.1, allows fluoro-triketide formation (*Figure 4.5*).

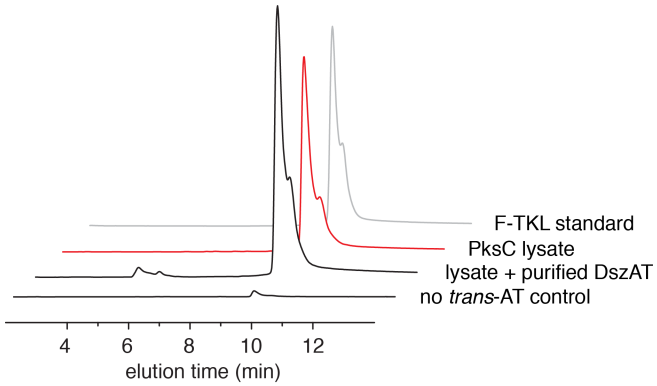


***Figure 4.5**. Example chromatograms showing LC-MS detection (single ion monitoring, m/z -173) of fluorotriketide lactone in cell lysate by complementation of DEBS-Mod6AT^0.*

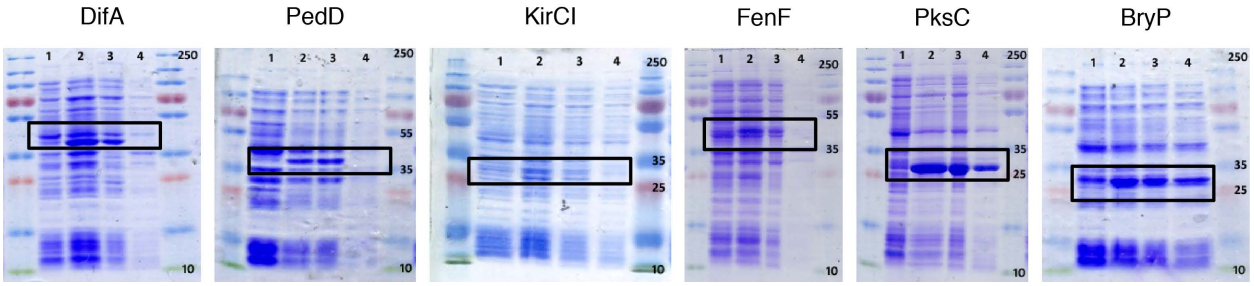


***Figure 4.6**. SDS-PAGE gels showing* trans*-AT expression and solubility in cell lysate. Lanes: 1 pre-induction 2 post-induction 3 insoluble 4 soluble. Boxes indicate the band corresponding to each AT.*

Expression of most of the library members we tested was robust despite the use of rare (in *E. coli*) codons, but in the *E. coli* host background many ATs were highly insoluble (*Figure 4.6*); this result may or may not carry over to a *Streptomyces* host. Even in cases of poor solubility, we were able to easily purify enough His-tagged *trans*-AT to assess activity (e.g. for FenF, *Figure 4.7* and *Figure 4.8*).

We used AT expression cell lysate to complement DEBS-Mod6AT^0 in a triketide-forming assay and compared fluoro-triketide and desmethyl-triketide (H-TKL) levels (*Figure 4.8*). Several *trans*-ATs were inactive above background, probably due to low solubility; purification of FenF allowed activity to be determined. PksC and BryP showed high activity on fluoromalonyl-CoA. Fluoro-triketide yields were apparently higher using these *trans*-ATs than desmethyl-triketide yields, a surprising observation given the native malonyl-CoA specificity of these enzymes. However, this assay data is complicated by the formation of additional products in the malonyl-CoA condition that were not quantified. When *trans*-AT activity is very high, a PKS module can catalyze *two* chain extensions on the diketide starter, which in the absence of NADPH results in a tetraketide pyrone (*Scheme 4.1*; Omer Ad, unpublished results). The exact mechanism for this process is as yet unknown, but modules have been made to iterate by replacement of interaction domains [*40*] presumably by intermolecular chain transfer to another copy of the module after the first elongation event. Intramolecular back-transfer of the chain to the ketosynthase Cys residue after elongation also cannot be ruled out. Regardless of its origin, formation of this tetraketide pyrone will need to be accounted for in future assays, since it represents two incorporations of malonyl extender. The data in *Figure 4.8* still provide a relative comparison of the ATs, but the fluoromalonyl/malonyl product ratio is not meaningful.

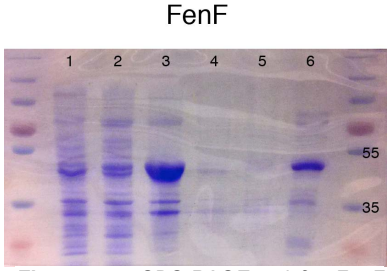


***Figure 4.7**. SDS-PAGE gel for FenF purification. Lanes: 1 pre-induction 2 post-induction 3 insoluble 4 soluble 5 post-DNA precipitation 6 purified*

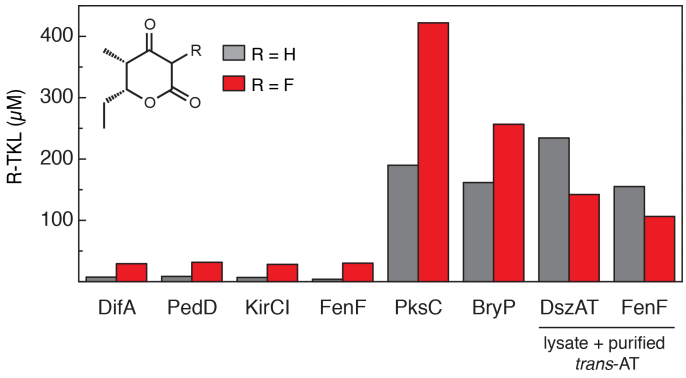
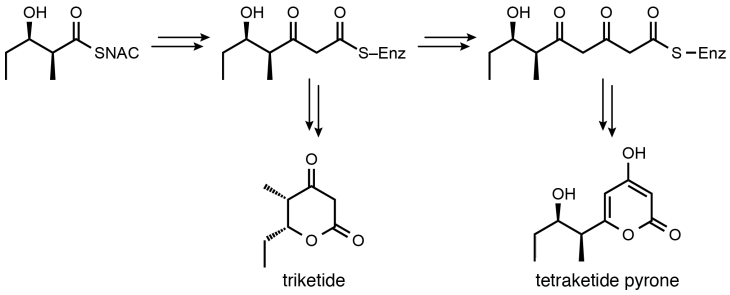


***Figure 4.8**. Cell lysate assay for selected* trans*-ATs expressed in* E. coli *complementing DEBS-Mod6AT^0. Triketides were formed from either malonyl- or fluoromalonyl-CoA in separate reaction mixtures. DszAT acts as a positive control. The selectivity for malonyl-derived vs. fluoromalonyl-derived triketides (R = H vs. R = F) cannot be inferred from these data because multiple chain elongation cycles with malonyl-CoA (tetraketide formation) were observed but could not be quantified.*

Design and plasmid construction for *in vivo* DEBS complementation. Construction of vectors for PKS expression, particularly of a full, multimodular assembly line, is an involved process because of the size of the genes involved (~30 kb for DEBS, which is encoded by eryAI-II-III) and their aforementioned high G+C content, which can make PCR amplification and even Sanger sequencing inefficient and/or error-prone. A vector with a *Streptomyces* replicative origin and resistance marker encoding DEBS was previously reported [*1*] and we took advantage of this for our design. This vector, pCK7, could itself be used for complementation, but it has some disadvantages: it is larger than necessary, with the backbone alone being over 10 kb due mostly to the SCP2* origin, making it difficult to handle and manipulate; it uses a heterologous promoter – p*actI*, which is positively regulated by transcription factor actII-ORF4



Scheme 4.1. *Triketide lactone and tetraketide pyrone formation resulting from one or two chain extension reactions with a malonyl extender.*

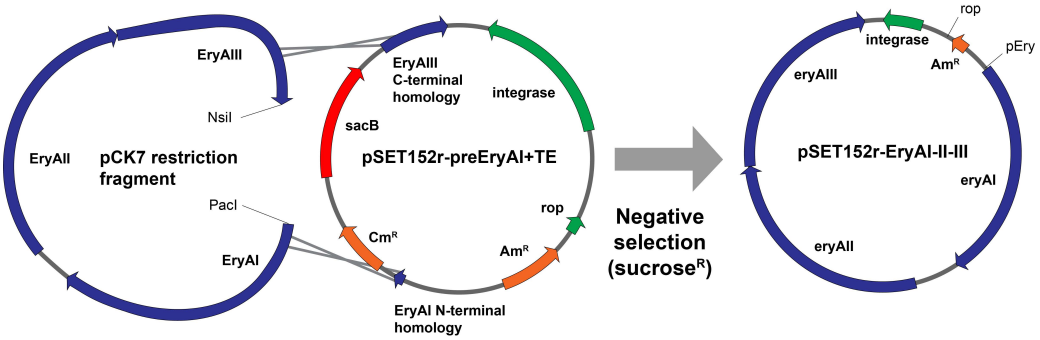


Figure 4.9. *Diagram showing recombination-based strategy for construction of an integrative DEBS expression plasmid for* Streptomyces *expression. Recombination (via λ-red, Gibson assembly or, with marker and origin addition, yeast) allows replacement of the cat-sacB cassette with the EryAI-II-III genes, derived by restriction digest from pCK7* [*1*]. *Recombinants can be selected by survival on 5% sucrose plates.*

during the onset of stationary phase – rather than the native p*eryAI*; and it has a replicative origin (albeit with low copy number), which may complicate adjusting the stoichiometry of DEBS to *trans*-ATs in our system. We therefore set out to create the integrative vector pSET152r-EryAI-II-III using a recombination-based strategy (*Figure 4.9*) that avoids PCR and the need for full sequencing of the 30 kb DEBS genes. Similar techniques have recently been reported to assemble large biosynthetic clusters [*41,42*].

We selected pSET152 as a minimal vector with site-specific integration capability (the øC31-phage-derived integrase catalyzes insertion of the complete vector backbone at one or more attB sites in a *Streptomyces* chromosome). To make *E. coli* passage of pSET152 derivatives with large inserts more reliable, we first reduced the copy number of the pUC origin by inserting the *rop* gene [*43*]. We next cloned pSET152r-preEryAI+TE by inserting long (50-800 bp) sequences homologous to the 5′ and 3′ ends of the eryAI-II-III cluster (including the native promoter), flanking a positive (chloramphenicol resistance) and negative (sucrose sensitivity) selectable cassette [*2*] which can be excised by NdeI digest. This should allow insertion of the PacI/NsiI restriction fragment of pCK7 by either *in vivo* λ-red recombination or *in vitro* Gibson assembly

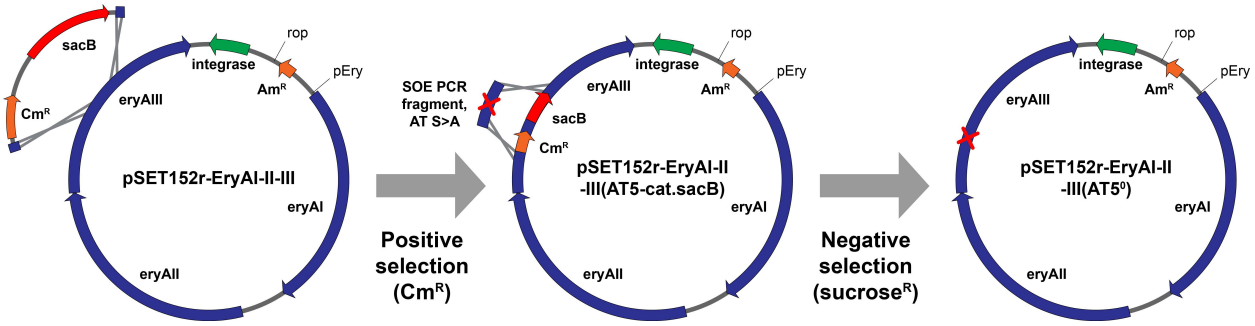


Figure 4.10. *Recombination-based strategy for introducing AT^0 mutations to the full DEBS cluster. The cat-sacB cassette can be introduced by PCR-targeted recombination[2] – potentially assisted by Cas9-based cleavage at the target site. The desired point mutant can then be prepared by 2-piece SOE PCR of a short region around the AT active site, and introduced scarlessly by a second recombination event.*

(which may be preferable because repeated sequences in the eryA locus could lead to undesired recombination events).

An analogous strategy can be used to prepare the desired set of AT^0 point mutations (*Figure 4.10*), which are difficult to produce by typical methods (e.g. QuikChange) for the same reasons discussed above. The selection cassette can be inserted at any locus in the *ery* cluster using PCR-targeted recombination, without the need for a unique, proximal restriction site [*44*]. A short, easily sequence-verified DNA fragment containing the desired point mutation can then be prepared by SOE PCR (with the mutation contained on the internal primers) and exchanged for the cassette by the same technique. Because the cassette can be counter selected using sucrose, the desired scarless point mutant can be selected for. Since this work was initiated, Cas9-based methods have been developed for *Streptomyces* that make point mutations possible using DNA oligos [*45*], and these provide an attractive alternative strategy.

4.4 Conclusions

The work outlined here is only the first steps toward construction of a *trans*-AT library and comprehensive screening. Remaining challenges include optimizing 6-dEB production in a heterologous host, including finding reproducible growth conditions, adjusting plasmid elements and achieving stable coexpression with *trans*-AT library members. The enzymes and systems identified through *E. coli* 2-fluoro-3-hydroxybutyrate production will need to be incorporated in and adapted to the *Streptomyces* host to achieve robust fluoromalonyl-CoA production. The library inaugurated here should be expanded to better cover *trans*-AT sequence space and a testing pipeline implemented. Finally, promising ATs should be selected and subjected to rational and/or random mutagenesis to improve fluoromalonyl selectivity. Testing *in vivo* with the full DEBS system will reveal mutations that improve AT expression, stability, ACP recognition and selectivity against endogenous extender units, in addition to fluorinated extender use, and can potentially be scaled [*32*] to accommodate large library sizes.

4.5 References

1. Kao, C. M., Katz, L., Khosla, C. Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science* **1994**, *265* (5171) 509-512.
2. Thomason, L., Court, D. L., Bubunenko, M., Costantino, N., Wilson, H., Datta, S., Oppenheim, A. Recombineering: Genetic engineering in bacteria using homologous recombination. *Curr Protocols Mol Biol* **2007**, *Chapter 1* Unit 1.16.
3. Musiol, E. M., Greule, A., Härtner, T., Kulik, A., Wohlleben, W., Weber, T. The AT_2 domain of KirCI loads malonyl extender units to the ACPs of the kirromycin PKS. *ChemBioChem* **2013**, *14* (11) 1343-1352.
4. Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat Prod Rep* **2010**, *27* (7) 996-1047.
5. 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. *Angew Chem Int Ed* **2012**, *52* (4) 1143-1147.
6. Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, T., Taudien, S., Platzer, M., Hertweck, C., Piel, J. Exploiting the mosaic structure of *trans*-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat Biotechnol* **2008**, *26* (2) 225-233.
7. 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* (11) 1175-1186.
8. Jensen, K., Niederkrüger, H., Zimmermann, K., Vagstad, A. L., Moldenhauer, J., Brendel, N., Frank, S., Pöplau, P., Kohlhaas, C., Townsend, C. A., Oldiges, M., Hertweck, C., Piel, J. Polyketide proofreading by an acyltransferase-like enzyme. *Chem Biol* **2012**, *19* (3) 329-339.
9. Wong, F. T., Chen, A. Y., Cane, D. E., Khosla, C. Protein−protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. *Biochemistry* **2010**, *49* (1) 95-102.
10. Bumpus, S. B., Magarvey, N. A., Kelleher, N. L., Walsh, C. T., Calderone, C. T. Polyunsaturated fatty-acid-like *trans*-enoyl reductases utilized in polyketide biosynthesis. *J Am Chem Soc* **2008**, *130* (35) 11614-11616.
11. Aron, Z. D., Fortin, P. D., Calderone, C. T., Walsh, C. T. FenF: Servicing the mycosubtilin synthetase assembly line in trans. *ChemBioChem* **2007**, *8* (6) 613-616.
12. Dunn, B. J., Cane, D. E., Khosla, C. Mechanism and specificity of an acyltransferase domain from a modular polyketide synthase. *Biochemistry* **2013**, *52* (11) 1839-1841.
13. Lowry, B., Robbins, T., Weng, C.-H., O'Brien, R. V., Cane, D. E., Khosla, C. In vitro reconstitution and analysis of the 6-deoxyerythronolide B synthase. *J Am Chem Soc* **2013**, *135* (45) 16809-16812.

14. 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* (6150) 1089-1094.

15. Musiol, E. M., Härtner, T., Kulik, A., Moldenhauer, J., Piel, J., Wohlleben, W., Weber, T. Supramolecular templating in kirromycin biosynthesis: The acyltransferase KirCII loads ethylmalonyl-CoA extender onto a specific ACP of the *trans*-AT PKS. *Chem Biol* **2011**, *18* (4) 438-444.

16. Zhao, C., Coughlin, J. M., Ju, J., Zhu, D., Wendt-Pienkowski, E., Zhou, X., Wang, Z., Shen, B., Deng, Z. Oxazolomycin biosynthesis in *Streptomyces albus* JA3453 featuring an "acyltransferase-less" type I polyketide synthase that incorporates two distinct extender units. *J Biol Chem* **2010**, *285* (26) 20097-20108.

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

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

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

20. 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* (47) 14307-14312.

21. Chen, X. H., Vater, J., Piel, J., Franke, P., Scholz, R., Schneider, K., Koumoutsi, A., Hitzeroth, G., Grammel, N., Strittmatter, A. W., Gottschalk, G., Sussmuth, R. D., Borriss, R. Structural and functional characterization of three polyketide synthase gene clusters in *Bacillus amyloliquefaciens* FZB 42. *J Bacteriol* **2006**, *188* (11) 4024-4036.

22. Piel, J. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of Paederus beetles. *Proc Natl Acad Sci USA* **2002**, *99* (22) 14002-14007.

23. Arakawa, K., Sugino, F., Kodama, K., Ishii, T., Kinashi, H. Cyclization mechanism for the synthesis of macrocyclic antibiotic lankacidin in *Streptomyces rochei*. *Chem Biol* **2005**, *12* (2) 249-256.

24. Pulsawat, N., Kitani, S., Nihira, T. Characterization of biosynthetic gene cluster for the production of virginiamycin M, a streptogramin type A antibiotic, in *Streptomyces virginiae*. *Gene* **2007**, *393* (1-2) 31-42.

25. Cheng, Y.-Q., Tang, G.-L., Shen, B. Type I polyketide synthase requiring a discrete acyltransferase for polyketide biosynthesis. *Proc Natl Acad Sci USA* **2003**, *100* (6) 3149-3154.

26. Calderone, C. T., Iwig, D. F., Dorrestein, P. C., Kelleher, N. L., Walsh, C. T. Incorporation of nonmethyl branches by isoprenoid-like logic: Multiple beta-alkylation events in the biosynthesis of myxovirescin A1. *Chem Biol* **2007**, *14* (7) 835-846.

27. Calderone, C. T. C., Kowtoniuk, W. E. W., Kelleher, N. L. N., Walsh, C. T. C., Dorrestein, P. C. P. Convergence of isoprene and polyketide biosynthetic machinery: Isoprenyl-s-carrier proteins in the PksX pathway of *Bacillus subtilis*. *Proc Natl Acad Sci USA* **2006**, *103* (24) 8977-8982.

28. El-Sayed, A. K., Hothersall, J., Cooper, S. M., Stephens, E., Simpson, T. J., Thomas, C. M. Characterization of the mupirocin biosynthesis gene cluster from Pseudomonas fluorescens NCIMB 10586. *Chem Biol* **2003**, *10* (5) 419-430.

29. Partida-Martinez, L. P., Hertweck, C. A gene cluster encoding rhizoxin biosynthesis in "Burkholderia rhizoxina", the bacterial endosymbiont of the fungus Rhizopus microsporus. *ChemBioChem* **2007**, *8* (1) 41-45.

30. Perlova, O., Gerth, K., Kaiser, O., Hans, A., Müller, R. Identification and analysis of the chivosazol biosynthetic gene cluster from the myxobacterial model strain Sorangium cellulosum So ce56. *J Biotechnol* **2006**, *121* (2) 174-191.

31. Schneider, K., Chen, X.-H., Vater, J., Franke, P., Nicholson, G., Borriss, R., Süssmuth, R. D. Macrolactin is the polyketide biosynthesis product of the Pks2 cluster of *Bacillus amyloliquefaciens* FZB42. *J Nat Prod* **2007**, *70* (9) 1417-1423.

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

33. Rückert, C., Albersmeier, A., Busche, T., Jaenicke, S., Winkler, A., Friðjónsson, Ó. H., Hreggviðsson, G. Ó., Lambert, C., Badcock, D., Bernaerts, K., Anne, J., Economou, A., Kalinowski, J. Complete genome sequence of *Streptomyces lividans* TK24. *J Biotechnol* **2015**, *199* 21-22.

34. Gomez-Escribano, J. P., Bibb, M. J. Engineering *Streptomyces coelicolor* for heterologous expression of secondary metabolite gene clusters. *Microbial Biotechnology* **2010**, *4* (2) 207-215.

35. Komatsu, M., Uchiyama, T., Omura, S., Cane, D. E., Ikeda, H. Genome-minimized *Streptomyces* host for the heterologous expression of secondary metabolism. *Proc Natl Acad Sci USA* **2010**, *107* (6) 2646-2651.

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

37. Bierman, M., Logan, R., O'Brien, K., Seno, E. T., Rao, R. N., Schoner, B. E. Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces spp*. *Gene* **1992**, *116* (1) 43-49.

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

39. Doumith, M., Weingarten, P., Wehmeier, U. F., Salah-Bey, K., Benhamou, B., Capdevila, C., Michel, J. M., Piepersberg, W., Raynal, M. C. Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in Saccharopolyspora erythraea. *Molecular and General Genetics* **2000**, *264* (4) 477-485.

40. Kapur, S., Lowry, B., Yuzawa, S., Kenthirapalan, S., Chen, A. Y., Cane, D. E., Khosla, C. Reprogramming a module of the 6-deoxyerythronolide B synthase for iterative chain elongation. *Proc Natl Acad Sci USA* **2012**, *109* (11) 4110-4115.

41. Su, C., Zhao, X.-Q., Wang, H.-N., Qiu, R.-G., Tang, L. Seamless stitching of biosynthetic gene cluster containing type I polyketide synthases using red/et mediated recombination for construction of stably co-existing plasmids. *Gene* **2015**, *554* (2) 233-240.

42. 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. *Proc Natl Acad Sci USA* **2014**, *111* (5) 1957-1962.

43. Lin-Chao, S., Chen, W. T., Wong, T. T. High copy number of the pUC plasmid results from a rom/rop-suppressible point mutation in RNA II. *Mol Microbiol* **1992**, *6* (22) 3385-3393.

44. Gust, B., O'Rourke, S., Bird, N., Kieser, T., Chater, K. F. Recombineering in *Streptomyces coelicolor*. *Norwich: The John Innes Foundation* **2003**.

45. Cobb, R. E., Wang, Y., Zhao, H. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synthetic Biology* **2014**.

Appendix 1: *Plasmids and oligonucleotides*

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

A

Name	No.	Primers	Restriction sites	Method	Description
pET28a-His_{10}-MatB	905	MatB.Sco F/R	NdeI/XhoI	Ligation	malonyl-CoA synthetase from *S. coelicolor*
pET28a-His_{10}-Epi	803	Epi.Sco F/R	NdeI/XhoI	Ligation	methylmalonyl-CoA epimerase from *S. coelicolor*
pSV272-MBP-$DEBS_{Mod2}$-His_6	1396	MBP-M2 F/R	SfoI/HindIII	Gibson	N-terminal MBP fusion of $DEBS_{Mod2}$. pBP19 as template
pSV272-MBP-$DEBS_{Mod2}AT^0$-His_6	1341	MBP-M2ATnull F/R (internal) MBP-M2 F/R (external)	SfoI/HindIII	Gibson	N-terminal MBP fusion of $DEBS_{Mod2}$ with S2652A mutation (introduced via SOE PCR)
pET16b-His_{10}-NphT7	988	NphT7 G F/R	NdeI	Gibson	Synthetic NphT7, acetoacetyl-CoA synthase from *S. sp. CL190* (primers for assembly are given in Table 1, B)
pRSG54	926				$DEBS_{Mod6}$+TE from Wu, N., et. al., *J. Am. Chem. Soc.* **2000,** 122 (20) 4847-4852.
pAYC138	1045				$DEBS_{Mod6}AT^0$+TE from Wong, F. T., et. al., *Biochemistry* **2009,** 49 (1) 95-102.
pRSG34	1123				$DEBS_{Mod3}$+TE from Gokhale, R. S., et. al., *Science* **1999,** 284 (5413) 482-485.
pAYC136	1150				$DEBS_{Mod3}AT^0$+TE from Wong, F. T., et. al., *Biochemistry* **2009,** 49 (1) 95-102.
pFW3	1186				Disorazole Trans AT from Wong, F. T., et. al., *Biochemistry* **2009,** 49 (1) 95-102.

B

Name	Sequence
nphT7 R1	aaacgaacgtcggtcatggtg
nphT7 F1	caccatgaccgacgttcgttttcgtatcattggcacgggt
nphT7 R2	gctccggcacgtacgcacccgtgccaatgatacga
nphT7 F2	gcgtacgtgccggagcgtattgtgtccaacgacgaggt
nphT7 R3	accagccggcgcacccacctcgtcgttggacacaatac
nphT7 F3	gggtgcgccggctggtgttgatgatgactggattacccgt
nphT7 R4	cgttgacgaatgccggtcttacgggtaatccagtcatcatcaac
nphT7 F4	aagaccggcattcgtcaacgtcgttgggcggcggac
nphT7 R5	tcggaggtcgcttggtcgtccgccgcccaacga
nphT7 F5	gaccaagcgacctccgacctggcaaccgcggcg
nphT7 R6	tcaacgccgcacgacccgccgcggttgccagg
nphT7 F6	ggtcgtgcggcgttgaaagcagcgggtattacgcc

Name	Sequence
nphT7 R7	gcaataaccgtcagttgctccggcgtaatacccgctgctt
nphT7 F7	ggagcaactgacggttattgcggtcgcaacgtccaccc
nphT7 R8	ggctgcggacggtccggggtggacgttgcgacc
nphT7 F8	cggaccgtccgcagccgccgacggcggcctac
nphT7 R9	cgcccagatgatgttgcacgtaggccgccgtcggc
nphT7 F9	gtgcaacatcatctgggcgcaaccggcaccgcggc
nphT7 R10	tgcacacagcgttaacatcaaatgccgcggtgccggttg
nphT7 F10	atttgatgttaacgctgtgtgcagcggcacggtttttgct
nphT7 R11	ccgccacgctggacagagcaaaaaccgtgccgc
nphT7 F11	ctgtccagcgtggcgggcacgctggtgtatcgtgg
nphT7 R12	caatgaccagtgcgtaaccgccacgatacaccagcgtgc
nphT7 F12	cggttacgcactggtcattggtgccgatctgtattcccgta
nphT7 R13	ggtccgccggattcagaatacgggaatacagatcggcac
nphT7 F13	ttctgaatccggcggaccgcaagaccgttgttctgtttgg
nphT7 R14	cgcacccgcgccgtcaccaaacagaacaacggtcttgc
nphT7 F14	tgacggcgcgggtgcgatggtgctgggtccgac
nphT7 R15	acccgtacccgtgctggtcggacccagcaccat
nphT7 F15	cagcacgggtacgggtccgatcgtccgtcgcg
nphT7 R16	caaacgtgtgcagggcaacgcgacggacgatcgg
nphT7 F16	ttgccctgcacacgtttggtggtctgaccgacctgatt
nphT7 R17	cacccgccggcacacgaatcaggtcggtcagaccac
nphT7 F17	cgtgtgccggcgggtggcagccgccaaccgct
nphT7 R18	tccaagccatccgtgtccagcggttggcggctgc
nphT7 F18	ggacacggatggcttggacgcgggtctgcaatacttcg
nphT7 R19	cctcgcgaccgtccatagcgaagtattgcagacccgcg
nphT7 F19	ctatggacggtcgcgaggtgcgtcgttttgttaccgaac
nphT7 R20	cctttaatcagttgcggcaagtgttcggtaacaaaacgacgca
nphT7 F20	acttgccgcaactgattaaaggtttcttgcacgaggcggg
nphT7 R21	gctaatatctgccgcatcgacacccgcctcgtgcaagaaa
nphT7 F21	tgtcgatgcggcagatattagccattttgtgccgcaccaagc
nphT7 R22	cgtccagcatgacaccgttcgcttggtgcggcacaaaatg
nphT7 F22	gaacggtgtcatgctggacgaggtctttggtgaactgcacc
nphT7 R23	atggtcgcacgcggcaggtgcagttcaccaaagacct
nphT7 F23	tgccgcgtgcgaccatgcaccgtaccgtcgaaacc
nphT7 R24	cgcacccgtattgccgtaggtttcgacggtacggtgc
nphT7 F24	tacggcaatacgggtgcggccagcattccgattacgatg
nphT7 R25	tgcacggactgctgcatccatcgtaatcggaatgctggc
nphT7 F25	gatgcagcagtccgtgcaggtagcttccgtccggg
nphT7 R26	gccagcaggaccagttcacccggacggaagctacc
nphT7 F26	tgaactggtcctgctggcgggttttggtggtggcatg
nphT7 R27	gcgcgaagctcgctgccatgccaccaccaaaaccc
nphT7 F27	gcagcgagcttcgcgctgatcgagtggtaagtcagcc
nphT7 R28	acccgctctagccgtcaggctgacttaccactcgatca
nphT7 F28	tgacggctagagcgggt
NphT7 G F	aatttcacacgagctcggtacccgggaggagatataccatgaccgacgttcgttttcg
NphT7 G R	gcgctgggtcattatatatctccttttcttaccactcgatcagcgcgaag
MatB.SCo F	tcgattgcacatatgtcctctctcttcccggccctct

Name	Sequence
MatB.SCo R	atcggatagctcgagtcagtcacggttcagcgcccgctt
Epi.SCo F	atcccgaatcatatgctgacgcgaatcgacca
Epi.SCo R	ttagtctggctcgagtcagtgctcaggtgactcaa
MBP-M2 F	gggatcgaggaaaacctgtattttcagggcatgagcggtgacaacggcatgaccgagg
MBP-M2 R	gcttgtcgacggagctcgaattcggggatcctcagtggtggtggtggtggtgctcgagtg
MBP-M2ATnull F	gttatcggtcacgcgcagggtgaaatcgcggccgcggtggtggcgggagcgttgtcgctg
MBP-M2ATnull R	cgcgatttcaccctgcgcgtgaccgataacggccgaaggaacggcaccgcaggcacgcca

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

A

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
pET16b.*Rpa*matB	1192	r_palu matB F1/R1	NdeI	pET16b	Gibson assembly	ColE1 Cb^R lacI pT7 His_{10}-*Rpa*matB T7-term
pCDF-Duet1.*Rpa*matB	1490	pCDFDuet R.p_matB F1/R1	NdeI/KpnI	pCDF-Duet1	Gibson assembly	CloDF13 Sp^R lacI pT7 *Rpa*matB T7-term
pET23c.His6-TEV-NphT7	2041	NphT7 23a F2/R1	EcoRI	pET23c.His-TEV-SfoI	Gibson assembly	ColE1 Cb^R lacI pT7 His_6-TEV-nphT7 T7-term
pBAD33.pET-RBS	1512	pBAD33-pET.RBS GF1/GR1	KpnI/XbaI	pBAD33	Gibson assembly	M13 Cm^R araC pBAD rrnB-term
pXHB.ø.*Rpa*matB	1514	pBAD GF1, pBAD-pCDF GR1 / lacI GF1, lacI-pBAD GR1	XbaI/HindIII	pBAD33.pET-RBS	Gibson assembly	CloDF13 Sp^R lacI araC pBAD rrnB-term pT7 *Rpa*matB T7-term
pXHB1 (pXHB.*Pfl*madLM.*Rpa*matB)	1547	Pfl.madLM GF1/GR1	XbaI/HindIII	pXHB.ø.RpamatB	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Pfl*madLM rrnB-term pT7 *Rpa*matB T7-term
pXHB.*Sco*matC.*Rpa*matB	1549	pCDF-ara-ScoMatC GF1/GR1	XbaI/HindIII	pXHB.ø.RpamatB	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Sco*matC rrnB-term pT7 *Rpa*matB T7-term
pXHB.*Kpn*mdcF.*Rpa*matB	1830	KpnmdcF block1/2	XbaI/HindIII	pXHB.ø.RpamatB	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Kpn*mdcF rrnB-term pT7 *Rpa*matB T7-term
pXHB1.*Eco*acs	1548	pCDF-ACS GF1/GR1	KpnI/XhoI	pXHB1	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Pfl*madLM rrnB-term pT7 *Rpa*matB *Eco*acs T7-term
pXHB1.*Eco*prpE	1630	PrpE GF1/GR1	KpnI/XhoI	pXHB1	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Pfl*madLM rrnB-term pT7 *Rpa*matB *Eco*prpE T7-term
pXHB1.*Sco*alkK	1618	ScoAlkK GF1/GR1	KpnI/XhoI	pXHB1	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Pfl*madLM rrnB-term pT7 *Rpa*matB *Sco*alkK T7-term
pXHB1.*Pol*alkK	1616	sPolAlkK GF1, sPolAlkK GR1	KpnI/XhoI	pXHB1	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Pfl*madLM rrnB-term pT7 *Rpa*matB *Pol*alkK T7-term
pXHB.*Pfl*madLM.ø	1929	n/a	KpnI/NdeI	pXHB1	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Pfl*madLM rrnB-term T7-term
pXHB3	1694	pXHB3.GF1/GR1	NdeI/NotI	pXHB1	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Pfl*madLM rrnB-term pT5 *Rpa*matB T7-term

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
pPOL1 (pPOL1-nbc)	1331	FPHB-NphT7 F1/R1 / FPHB-PHAB F1, FPHB-PHAC R1	NcoI/XbaI	pTRC99a	Gibson assembly	ColE1 Cb^R $lacI^q$ pTrc nphT7 phaB *Reu*phaC rrnB
pPOL1-nb	1589	NphT7 G F3, PhaB G R1	NcoI/XbaI	pTRC99a	Gibson assembly	ColE1 Cb^R $lacI^q$ pTrc nphT7 phaB rrnB
pPOL1-n	1791	pPol1-n GF1/GR1	EcoRI/XbaI	pPOL1-nb	Gibson assembly: oligo insertion	ColE1 Cb^R $lacI^q$ pTrc nphT7 rrnB
pPOL4.1	1693	pPOL4.1 block1/2	XbaI/HindIII	pPOL1-nb	Gibson assembly	ColE1 Cb^R $lacI^q$ pTrc nphT7 phaB *Rru*phaC rrnB
pPOL4.2	1783	pPOL4.2 block1/2	XbaI/HindIII	pPOL1-nb	Gibson assembly	ColE1 Cb^R $lacI^q$ pTrc nphT7 phaB *Rop*phaC rrnB
pET28a.flK	1927	pET28a flK GF1/GR1	XbaI/NcoI	pET28a	Gibson assembly	ColE1 Km^R lacI pT7 flK T7-term
pET28a. flK(H76A)	1925	pET28a flK GF1/GR1	XbaI/NcoI	pET28a	Gibson assembly	ColE1 Km^R lacI pT7 flK(H76A) T7-term
pET28a.pTet.flK	2084	pTet-pET28a.flk GF1/GR1	XbaI/BglI	pET28a.flK	Gibson assembly	ColE1 Km^R tetR pTet flK T7-term
pET28a.pTet. flK(H76A)	2085	pTet-pET28a.flk GF1/GR1	XbaI/BglI	pET28a.flK(H76A)	Gibson assembly	ColE1 Km^R tetR pTet flK(H76A) T7-term

B

Name	Sequence
Rhod_Palu_MatB.R2	CATgtgggcgggagagg
Rhod_Palu_MatB.F2	cctctcccgcccacATGAACGCGAATCTGTTTGCC
Rhod_Palu_MatB.R3	TCCAGTTTATCGAACAGGCGGGCGAAAAGATTCGCGTT
Rhod_Palu_MatB.F3	CGCCTGTTCGATAAACTGGACGATCCGCACAAACTGGC
Rhod_Palu_MatB.R4	AGCCGCGGTCTCGATGGCCAGTTTGTGCGGATCG
Rhod_Palu_MatB.F4	CATCGAGACCGCGGCTGGTGATAAGATCAGCTATGCCG
Rhod_Palu_MatB.R5	GCGCGAGCCACCAGCTCGGCATAGCTGATCTTATCACC
Rhod_Palu_MatB.F5	AGCTGGTGGCTCGCGCGGGCCGCGTCGCAAAC
Rhod_Palu_MatB.R6	CACGCGCCACCAGGACGTTTGCGACGCGGCCC
Rhod_Palu_MatB.F6	GTCCTGGTGGCGCGTGGTCTGCAAGTTGGTGACCG
Rhod_Palu_MatB.R7	GGTTTGCGCCGCAACACGGTCACCAACTTGCAGAC
Rhod_Palu_MatB.F7	TGTTGCGGCGCAAACCGAAAAGAGCGTTGAGGCACT
Rhod_Palu_MatB.R8	GTGGCCAGGTACAGAACCAGTGCCTCAACGCTCTTTTC
Rhod_Palu_MatB.F8	GGTTCTGTACCTGGCCACGGTTCGCGCTGGCGGC
Rhod_Palu_MatB.R9	CGTGTTCAGCGGCAGATAAACGCCGCCAGCGCGAACC
Rhod_Palu_MatB.F9	GTTTATCTGCCGCTGAACACGGCATATACCCTGCACGAACT
Rhod_Palu_MatB.R10	CGTCGGTGATGAAGTAGTCCAGTTCGTGCAGGGTATATGC
Rhod_Palu_MatB.F10	GGACTACTTCATCACCGACGCGGAGCCGAAAATTGTTGTC
Rhod_Palu_MatB.R11	CGTTTGCTCGGATCGCAGACAACAATTTTCGGCTCCG
Rhod_Palu_MatB.F11	TGCGATCCGAGCAAACGTGATGGCATTGCAGCCA
Rhod_Palu_MatB.R12	CCGACTTTCGCGGCGATGGCTGCAATGCCATCA
Rhod_Palu_MatB.F12	TCGCCGCGAAAGTCGGTGCAACCGTCGAGACCC
Rhod_Palu_MatB.R13	CGGCCATCCGGACCCAGGGTCTCGACGGTTGCA
Rhod_Palu_MatB.F13	TGGGTCCGGATGGCCGCGGCAGCCTGACGGAT
Rhod_Palu_MatB.R14	TCGCACCTGCCGCTGCATCCGTCAGGCTGCCG

Name	Sequence
Rhod_Palu_MatB.F14	GCAGCGGCAGGTGCGAGCGAAGCGTTCGCGAC
Rhod_Palu_MatB.R15	TGCACCGCGATCGATGGTCGCGAACGCTTCGC
Rhod_Palu_MatB.F15	CATCGATCGCGGTGCAGATGATCTGGCGGCGATT
Rhod_Palu_MatB.R16	GTGCCGCTGGTGTACAGAATCGCCGCCAGATCATC
Rhod_Palu_MatB.F16	CTGTACACCAGCGGCACGACGGGCCGCAGCAAA
Rhod_Palu_MatB.R17	GGCTCAACATCGCGCCTTTGCTGCGGCCCGTC
Rhod_Palu_MatB.F17	GGCGCGATGTTGAGCCACGACAATCTGGCGTCCA
Rhod_Palu_MatB.R18	CCACCAGGGTCAAGCTGTTGGACGCCAGATTGTCGT
Rhod_Palu_MatB.F18	ACAGCTTGACCCTGGTGGACTACTGGCGCTTCACCC
Rhod_Palu_MatB.R19	TGAATCAGCACATCGTCCGGGGTGAAGCGCCAGTAGT
Rhod_Palu_MatB.F19	CGGACGATGTGCTGATTCATGCGCTGCCGATCTATCA
Rhod_Palu_MatB.R20	CGAACAGACCGTGGGTGTGATAGATCGGCAGCGCA
Rhod_Palu_MatB.F20	CACCCACGGTCTGTTCGTGGCCAGCAACGTCAC
Rhod_Palu_MatB.R21	GGAGCCACGTGCAAACAAGGTGACGTTGCTGGCCA
Rhod_Palu_MatB.F21	CTTGTTTGCACGTGGCTCCATGATTTTCCTGCCGAAGTTCG
Rhod_Palu_MatB.R22	CAGATCCAGAATTTTATCCGGGTCGAACTTCGGCAGGAAAATCAT
Rhod_Palu_MatB.F22	ACCCGGATAAAATTCTGGATCTGATGGCGCGTGCCACGG
Rhod_Palu_MatB.R23	TCGGAACACCCATCAGGACCGTGGCACGCGCCAT
Rhod_Palu_MatB.F23	TCCTGATGGGTGTTCCGACGTTTTACACCCGCCTGC
Rhod_Palu_MatB.R24	AGGCGCGGGCTTTGCAGCAGGCGGGTGTAAAACG
Rhod_Palu_MatB.F24	TGCAAAGCCCGCGCCTGACCAAAGAGACCACGGGT
Rhod_Palu_MatB.R25	GAAATGAACAGGCGCATGTGACCCGTGGTCTCTTTGGTC
Rhod_Palu_MatB.F25	CACATGCGCCTGTTCATTTCCGGCAGCGCCCCGTTG
Rhod_Palu_MatB.R26	ACGATGGGTGTCAGCCAACAACGGGGCGCTGCCG
Rhod_Palu_MatB.F26	TTGGCTGACACCCATCGTGAGTGGAGCGCGAAAACG
Rhod_Palu_MatB.R27	TCCAGGACGGCGTGGCCCGTTTTCGCGCTCCACTC
Rhod_Palu_MatB.F27	GGCCACGCCGTCCTGGAACGCTACGGCATGACCGA
Rhod_Palu_MatB.R28	GGGTTGCTCGTATTCATGTTCGTCTCGGTCATGCCGTAGCGT
Rhod_Palu_MatB.F28	GACGAACATGAATACGAGCAACCCGTATGACGGCGATCGTGTGC
Rhod_Palu_MatB.R29	GGGCCGACTGCACCCGGCACACGATCGCCGTCATAC
Rhod_Palu_MatB.F29	CGGGTGCAGTCGGCCCGGCTCTGCCGGGCGTG
Rhod_Palu_MatB.R30	TCGGTGACACGGGCGCTCACGCCCGGCAGAGCC
Rhod_Palu_MatB.F30	AGCGCCCGTGTCACCGACCCGGAGACGGGTAAAGAG
Rhod_Palu_MatB.R31	ATGTCGCCACGCGGCAACTCTTTACCCGTCTCCGGG
Rhod_Palu_MatB.F31	TTGCCGCGTGGCGACATTGGTATGATCGAAGTTAAAGGTCC
Rhod_Palu_MatB.R32	GCCAGTAACCTTTGAAAACGTTCGGACCTTTAACTTCGATCATACCA
Rhod_Palu_MatB.F32	GAACGTTTTCAAAGGTTACTGGCGTATGCCGGAGAAAACCAAGT
Rhod_Palu_MatB.R33	CGTCGTCACGAAACTCGGACTTGGTTTTCTCCGGCATAC
Rhod_Palu_MatB.F33	CCGAGTTTCGTGACGACGGTTTTTTTATTACCGGCGACCTG
Rhod_Palu_MatB.R34	TAGCCACGTTCATCGATTTTACCCAGGTCGCCGGTAATAAAAAAAC
Rhod_Palu_MatB.F34	GGTAAAATCGATGAACGTGGCTATGTCCACATCCTGGGTCG
Rhod_Palu_MatB.R35	GTAATCACCAGGTCTTTGCCACGACCCAGGATGTGGACA
Rhod_Palu_MatB.F35	TGGCAAAGACCTGGTGATTACGGGTGGCTTCAATGTGTATCC
Rhod_Palu_MatB.R36	TCGATTTCGCTCTCAATCTCTTTCGGATACACATTGAAGCCACCC
Rhod_Palu_MatB.F36	GAAAGAGATTGAGAGCGAAATCGATGCCATGCCGGGTGTT
Rhod_Palu_MatB.R37	TCACGGCGCTTTCCACAACACCCGGCATGGCA

Name	Sequence
Rhod_Palu_MatB.F37	GTGGAAAGCGCCGTGATCGGTGTTCCGCACGC
Rhod_Palu_MatB.R38	CCGTAACACCCTCACCAAAATCCGCGTGCGGAACACCGA
Rhod_Palu_MatB.F38	GGATTTTGGTGAGGGTGTTACGGCAGTGGTCGTTCGTGATAAGG
Rhod_Palu_MatB.R39	GGCTTCATCAATCGTCGCACCCTTATCACGAACGACCACTG
Rhod_Palu_MatB.F39	GTGCGACGATTGATGAAGCCCAGGTGTTGCACGGTCT
Rhod_Palu_MatB.R40	TGAATTTTGCCAGTTGACCATCCAGACCGTGCAACACCTG
Rhod_Palu_MatB.F40	GGATGGTCAACTGGCAAAATTCAAAATGCCGAAGAAAGTGATCTTTG
Rhod_Palu_MatB.R41	TATTGCGCGGCAGATCATCGACAAAGATCACTTTCTTCGGCATTT
Rhod_Palu_MatB.F41	TCGATGATCTGCCGCGCAATACGATGGGTAAGGTCCAAAAAAATG
Rhod_Palu_MatB.R42	GTCTTTATAGGTTTCACGCAGGACATTTTTTTGGACCTTACCCATCG
Rhod_Palu_MatB.F42	TCCTGCGTGAAACCTATAAAGACATCTATAAGTAAggacgggagca
Rhod_Palu_MatB.F43	tgctcccgtccTTACTTATAGAT
r_palu matB F1	CAGCAGCGGCCATATCGAAGGTCGTCATatgaacgcgaatctgtttgcccg
r_palu matB R1	CTTTGTTAGCAGCCGGATCCTCGAGCATAttacttatagatgtctttataggtttcacgcaggaca
pCDFDuet R.p_matB F1	ttagttaagtataagaaggagatataCATatgaacgcgaatctgtttgcc
pCDFDuet R.p_matB R1	cagcggtttctttaccagactcgagggtaccttacttatagatgtctttataggtttcacgc
NphT7 23a F2	tcatcatcatgagaatctctacttccagggtatgaccgacgttcgttttcgtatcattgg
NphT7 23a R1	ccggatctcagtggtggtggtggtggtgctcgattaccactcgatcagcgcgaagctc
pBAD33-pET.RBS GF1	aaggagatatacaatctagAgtcgacctgcaggcatgcaagctt
pBAD33-pET.RBS GR1	ctagattgtatatctccttCgagctcgaattcgctagcccaaa
pBAD GF1	ttaccaattatgacaacttgacgg
pBAD-pCDF GR1	ttcgacttaagcattatgcggccgCgtctcatgagcggatacatatttg
lacI GF1	tgatcttttctactgaaccgctcaag
lacI-pBAD GR1	ccgtcaagttgtcataattggtaAgcataagggagagcgtcg
Pfl.madLM GF1	ttcgagctcgaaggagatatacaaAtgaggagggaagccgc
Pfl.madLM GR1	cttctctcatccgccaaaacagccaagcttTtagcccaccaggccg
pCDF-ara-ScoMatC GF1	ttcgagctcgaaggagatatacaatctatgtcccccgaactcatctcga
pCDF-ara-ScoMatC GR1	cttctctcatccgccaaaacagccaagcctacccgaagccgggca
pCDF-ACS GF1	aacctataaagacatctataagtaaggtAaagaggagaaaagatctAtgagcc
pCDF-ACS GR1	tttcgcagcagcggtttctttaccagactcttacgatggcatcgcgat
PrpE GF1	aagtaaggtaccaaagaggagaaaaGATCTatgTCTTTTAGCGAATTTTATCAGC
PrpE GR1	gttattgctcagcggtggcagcagcctaggCTACTCTTCCATCGCCTGG
ScoAlkK GF1	aagtaaggtaccaaagaggagaaaaGATCTatgTCGCCCCGGGAGGAC
ScoAlkK GR1	gttattgctcagcggtggcagcagcctaggTCAGAGCCGGGTGACGTCG
sPolAlkK GF1	gacatctataagtaaggtaccaaagCGAATTCGAGCTCGGTACC
sPolAlkK GR1	gttattgctcagcggtggcagcagcctaggCTTATTCGCAGACGCTGCTG
FPHB-NphT7 F1	cggataacaatttcacacaggaaacagaccaggagatatataatgaccgacgttcgttttcgtatcattggc
FPHB-NPHT7 R1	cgctgggtcattatatatctccttgaattcttaccactcgatcagcgcgaagctcgctgc
FPHB-PHAB F1	gaattcaaggagatatataatgacccagcgcatcgcttacgtaaccggtggcatgggtgg
FPHB-PHAC R1	aagcttgcatgcctgcaggtcgactctagatcatgccttggctttgacgtatcgcccagg
NphT7 G F3	aatttcacacgagctcggtacccgggAggagatataccATGACCGACGTTCGTTTTCG
PhaB G R1	GCCTGCAGGTCGACTCTAGAGGATCttagcccatgtgcaggccac
pPol1-n GF1	cgcgctgatcgagtggtaagaattctctagagtcgacctgcaggcatgcaa
pPol1-n GR1	ttgcatgcctgcaggtcgactctagagaattcttaccactcgatcagcgcg
pET28a flK GF1	gagcggataacaattcccctctagaagaaggagatatacatatgaaagatggtatgcgtgtaggc
pET28a flK GR1	tgatgatgatggctgctgcccatggttaacccgctggggttttctg

Name	Sequence
pTet-pET28a.flk GF1	cacgatcatgcgcacccgtggggccCgctttatgaatctaaagggtgg
pTet-pET28a.flk GR1	ctttcatatgtatatctccttcttctagaGgatcctgaagacgaaagggcctcg
pXHB3.GF1	caaatatgtatccgctcatgagacGgttttgcaccattcgatg
pXHB3.GR1	caggcgggcaaacagattcgcgttcaTatgtatatctccttgagctctgtg

C

Name	Sequence
pPOL4.1-block1	GCAGACTTTTCCCTGAACGGTGGCCTGCACATGGGCTAAACTAGTCCGGAGGTATAATTAatgctggatcacg tgcataaaaaactgaagtcaacactggaccctattggctgggggccagctgttacttccgtagccggtcgcgcagtgcggaatccgcaagcggtcaca gcagcgaccgccgagtatgccggccgcctggcgaaaatccccgcagcagctacgcgcgtgttcaatgcaaacgatcctgatgctcccatgccagtcg acccgcgtgatcgtcggttttcggataccgcgtggcaggaaaatcccgcatatttctcactgctccaatcgtacttagccactcgcgcctatgttgaagaact gaccgaagctggcagcggcgacccgttacaggatggcaaggctcggcagttcgctaacctgatgttcgatgccctcgcaccatccaatttttgtggaatc cgggtgtattgacgcgcgcatttgaaaccggaggtgcgtcgctgctgcgtggagctcgctacgccgcacatgatattttgaaccgcggtggcctgcctctg aaggtggattctgacgcgtttacggtaggtgagaacttagccgcgacgccgggcaaagtagtattccgcaatgatctgattgagctcatccaatacgctcc ccagacggagcaggtgcacgcggtgccgattctggcagctccaccctggatcaacaaatactacattctggacctggcgccggggcgcagcctcgcc gaatgggctgtgcagcatggtcggaccgtatttatgatttcgtaccgtaaccctgatgaaagtatgcggcatattaccatggatgattactatgtggacggg attgcaaccgccctggatgtagtcgaagagattaccggcagcccgaagat
pPOL4.1-block2	ccctggatgtagtcgaagagattaccggcagcccgaagattgaagtgctgagtatttgcctgggggtgcgatggcggcgatggcagccgcacgtgctt tcgccgtcggagataaacgtgttagcgcatttacgatgctgaatacgctcctggactacagccaagtgggggagctgggcttactgactgacccggctac ccttgatctggtagagttccggatgcgccagcagggcttcctttcaggcaaagagatggccggttcctttgacatgattcgtgccaaggatcttgtgtttaact attgggtatctcgctggatgaaaggggaaaagccggcagcgttcgacattctggcatggaatgaagattctacgagtatgcccgcagagatgcattcac attatctgcgcagtttgtatggccgcaatgaactggctgaaggtctgtacgttctggatgggcagcctctgaatcttcacgacatcgcttgtgatacgtatgttg tgggtgcgattaatgaccacatcgtgccttggacctcgtcgtatcaggcagtaaatcttctgggggcgatgtacgttatgttctgaccaacgggggtcacg tcgccggtgctgtgaatcccccgggtaaacgcgtgtggttcaaagcggttggcgcgccagacgccgagagtggcacgccgttacctgccgatccacaa gtatgggacgaggcagcgacccgttacgaacatagttggtgggaagattggacggcgtggtctaacaaacgtgccggcgaattagtagcgcctccgg ccatgggtagcacggcccaccctccgctcgaggatgcccctggtacatacgtttttagcTAAGGATCCAGTCGACCTGCAGGCATGC AAGCTTGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAG
pPOL4.2-block1	GCAGACTTTTCCCTGAACGGTGGCCTGCACATGGGCTAAACTAGTCCGGAGGTATAATTAatgtcagataccc cgctgccgacgattccggatgaactgaccgccccccttgacctgctgctgacctcaggttcccgctctgtagccgctcgtatgcttccggatagctcctgga cgcggctgggcgtcgggctggctggccgtccgggcaccgtggcacgtcgtggcggtgcgttggtgcgcgaactgggtgccattgcggcaggtaccagt gatcgcaccccggcaaaatcagataaacgctttggtgacgccgcctggcaacagaacccagctttgcgccgtgcaatgcaggcgtatctggcaacca gtcacaccgctgccgcgctgcttgatgatgcggagcttgactggcgcgatcacgagcggatgcggttcgttttagataacttggtggaaggcttatcaccta cgaataatcctctgctctctccgctgggttggaaagcgatggtagataccggcggcctttcggcggcgcgcggtgtgcgggccttagtacgcgatatgtta agcaaacctcgggttccagcgatggtagaaccggatgcgttcgcagttggcgaagacgttgccgctaccaagggcgcagtagtgctgcagacccgca catttgaactgatccactatacacctcaaaccgaaaaagtacatgccactccgttattgatcgtgccaccggtcatcaataaatattatatcttggatattgcg ccggggcgcagcttaatcgaatatctgctgcaacaaggccagcaagtgttcgcaatctcctggcgcaacccccacgctcgccaccgcgactgggatgc cgatacctatggctccgctattgtagaagccctcgataccgtacagtgcgttgcc
pPOL4.2-block2	cgctattgtagaagccctcgataccgtacagtgcgttgccggcactgattcagcccatgttctgggtacctgcagtggtggtatcctggccgccatggtggc ggcgcacttaacggagattggagaggggatcgtatcgcggggcttaccttagcggtgacggttctggaccaaactcaagctggcaccgcagcagca gtgatgtctgaacgggcggctgcggcggcaattcgtgatagcgctgcccgcggttatctggatggtcggaccctggctgaaatgtttgcgtggcttcggcc gagcgatcttgtgtggcgctactgggtgaacaactatgttcagggtcgcgcgcccgcagcgtttgatgttctgttctggaatagtgacaccacccgtatgac cgccgcgttgcatcgtgatctggtcctcctggggctgcgcaatgcactgacagctccgggggcagcgacaatgctgggcacaccagtagatctgtctac agttgccgcagacgcctatgtcgtaggcggtagtgcagatcacttgtgcccgtggcagtctacctatcgcagtgcacgtcttttagggtcgaaagattctcgt ttcgtcctgtcgagcaacggccatatcgcctcgctggttaacccaccgggcaatccccgcgcgtcctttcgcttcgggcagcccgtgccggaaacgcctg acgaatggttggctgcagcagagacggctagtgattcatggtggccagactacgcccgttggctggcggaacgtagtggcccggatgtcgatgctccgc acggcttaggagcgcgccagtttcctccactggctcccgccccgggcacatatgtgcatcactccTAAGGATCCAGTCGACCTGCAGGC ATGCAAGCTTGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAG

Name	Sequence
*Kpn*mdcF-block1	TGGGCTAGCGAATTCGAGCTCGAAGGAGATATACAAatgacctatgtgattattcatgcgctggcgccaatctttgtaatcatgctg ttgggattttgggcggggaaagccggaatggtggacaacaaaaatgtttcgctgcttaatatttttgttatggatttcgcactgccagcgaccttgttttcggcg accgtgcagacaccgtgggcaggcatcgttgctcagagcccgctcgtactcgtactgaccggagccatgtggattacctatgcggcaatctatttcctggc caccagcgtttttaagcgtaccccgcaggacgccgctgtattgacccttactgtagcattgcccaactatgctgcgctgggcctgccgattctcgggtctgtc ctgggcgagggcgcgagcacgagcttgtcggtcgcggtctcgattgcttgcgggtcagtccttatgaccccgttctgtct
*Kpn*mdcF-block2	cgattgcttgcgggtcagtccttatgaccccgttctgtctgttaatcctggagcgtgaaaaagcgcgggccgccggagaaaattctgggtctacgttagctat gctgccagtcttgatgtggcgttcagttaaaaaacctattgtgtggggacctctgttaggcgtggttctgagcgccatcgggatcaaaatgcctgagatgctt ctggcgtccattaaaccctaggcctggcagccactgcggcagctctgtttctcaccggcgtaattctgagcgcccgcaaactgcaactcaatgcgctgat cgcaaccagtaccattgtaaaactgcttgtgcaaccatttattgcttggggtttagtcatgctgctgggtttgcacgggtcgatcgctattaccgctattttaatg atcgcactggcggcgggctttttttggcgtggtgttcggcaatcgcttcggagtgcagtcgccggatgctgaagccgtgttgctgttaagcagcgttctctgtatt cttagcctgccattgttcatctctctgacctccggcctgtagAAGCTTGGCTGTTTTGGCGGATGAGAGAAG

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

A

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
pSET152	456					pUC Am^R oriT øC31-int *Gene* (**1992**) *116*, 43–49.
pSET152(gg)	1422	pSET152gg CF1/R1	EcoRV/BamHI	pSET152	Gibson assembly	pUC Am^R oriT øC31-int pTet RFP rrnB-term T7-term. BsaI GG acceptor
pSET152r(gg)	1423	pSET152-ROP-F1/R1	NheI	pSET152(gg)	Gibson assembly	pMB1 (pUC+rop) Am^R oriT øC31-int pTet RFP rrnB-term T7-term. BsaI GG acceptor
pUWL201PW	1421					pUC pIJ101 Cb^R Ts^R ermEp* *Mol Gen Genet* (**2000**) *264*, 477–485.
pSET152-pc13	-	pSETpreEryAI-piece1; pSETEryA1 F3b/R3	BsaI	pSET152(gg)	Golden Gate	pUC Am^R oriT øC31-int eryAI-N_{term} eryAIII-C_{term}
pSET152r-pc13	-	pSET152-ROP-F1/R1	NheI	pSET152r(gg)	Gibson assembly	pMB1 (pUC+rop) Am^R oriT øC31-int eryAI-N_{term} eryAIII-C_{term}
pSET152r-preEryAI+TE	1513	cat-sacB SOE F1/R1, pSETEryA1 F2b/R2b	BglII	pSET152r-pc13	Gibson assembly	pMB1 (pUC+rop) Am^R oriT øC31-int eryAI-N_{term} Cm^R sacB eryAIII-C_{term}
pCK7	2066					ColE1 SCP2* Cb^R Ts^R pactI actII-ORF4 eryAI eryAII eryAIII *Science* (**1994**) *265*, 509–512.
pCDF-pUWL-shuttle	2075	pUWL-sgg CF1/CR1	BsaI	pCDF2. P(tac.tac)(gg)	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac pTet RFP rrnB-term T7-term
pET28a*SapI	1510	pET.SapI_del F1/R1	SapI	pET28a	Gibson assembly	pMB1 lacI Km^R pT7 T7-term. Backbone SapI site removed (G>T)
pET28a.HisN(sgg)	1511	-	NdeI/EcoRI	pET28a*SapI	Ligation	pMB1 lacI Km^R pT7 pTet RFP rrnB-term 2xT7-term. SapI GG acceptor
pCDF-pUWL-shuttle.bryP	1496	BryP-block1/2	SapI	pCDF-pUWL-shuttle	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac bryP
pCDF-pUWL-shuttle.difA	1497	DifA-block1/2	SapI	pCDF-pUWL-shuttle	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac difA
pCDF-pUWL-shuttle.fenF	1498	FenF-block1/2	SapI	pCDF-pUWL-shuttle	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac fenF
pCDF-pUWL-shuttle.pedD	1499	PedD-block1/2	SapI	pCDF-pUWL-shuttle	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac pedD

Name	No.	Primers	Restriction sites	Backbone source	Method	Description
pCDF-pUWL-shuttle.pksC	1495	PksC-block1/2	SapI	pCDF-pUWL-shuttle	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac pksC
pCDF-pUWL-shuttle.lkcD	1860	LkcD-block1	SapI	pCDF-pUWL-shuttle	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac lkcD
pCDF-pUWL-shuttle.virI	1859	VirI-block1	SapI	pCDF-pUWL-shuttle	Golden Gate	CloDF13 Sp^R $lacI^q$ 2xpTac virI
pCDF-pUWL-shuttle.kirCI	1515	KirCI-block1/2; KirCI.SOEF1/R1/F2/R2	SapI	pCDF-pUWL-shuttle	Golden Gate, Gibson assembly	CloDF13 Sp^R $lacI^q$ 2xpTac kirCI
pUWL201PW.bryP	1506	-	XbaI; NdeI/HindIII	pUWL201PW	Gibson assembly	pUC pIJ101 Cb^R TsR ermEp* bryP
pUWL201PW.difA	1507	-	XbaI; NdeI/HindIII	pUWL201PW	Gibson assembly	pUC pIJ101 Cb^R Ts^R ermEp* difA
pUWL201PW.fenF	1508	-	XbaI; NdeI/HindIII	pUWL201PW	Gibson assembly	pUC pIJ101 Cb^R Ts^R ermEp* fenF
pUWL201PW.pedD	1509	-	XbaI; NdeI/HindIII	pUWL201PW	Gibson assembly	pUC pIJ101 Cb^R Ts^R ermEp* pedD
pUWL201PW.pksC	1505	-	XbaI; NdeI/HindIII	pUWL201PW	Gibson assembly	pUC pIJ101 Cb^R Ts^R ermEp* pksC
pET28a.HisN.bryP	1501	BryP-block1/2	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 bryP T7-term
pET28a.HisN.difA	1502	DifA-block1/2	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 difA T7-term
pET28a.HisN.fenF	1503	FenF-block1/2	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 fenF T7-term
pET28a.HisN.pedD	1504	PedD-block1/2	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 pedD T7-term
pET28a.HisN.pksC	1500	PksC-block1/2	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 pksC T7-term
pET28a.HisN.lkcD	1862	LkcD-block1	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 lkcD T7-term
pET28a.HisN.virI	1861	VirI-block1	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 virI T7-term
pET28a.HisN.kirCI	1521	KirCI.GGF2/GGR2	SapI	pET28a.HisN (sgg)	Golden Gate	pMB1 lacI Km^R pT7 kirCI T7-term
pET16b.HisN.*Sco*fabD	1518	FabD.GF1/GR1	NdeI/BamHI	pET16b	Gibson assembly	pMB1 lacI Cb^R pT7 *Sco*fabD T7-term

B

Name	Sequence
pSET152gg CF1	tatgacatgattacgaattcgatAtctaagtgagacctTccctatcagtgatagagattg
pSET152gg CR1	ggctgcaggtcgactctagaggatCcgagacCtataaacgcagaaaggccc
pSET152-ROP-F1	accgcgacgtatcgggccctggccaGCTAGCcttacacGGAGGcatcaGTG
pSET152-ROP-R1	cccggggacctgcaggtcgactctaTCAgaggttttcaccgtcatc
pSETEryA1 F3b	ATTATGGTCTCTcaccAgatctCATATGtcttcgaccacccgaac
pSETEryA1 R3	ATAATGGTCTCAgatCtcatgaattccctccgcc
cat-sacB SOE F1	Gtctattgctggactcggtac
cat-sacB SOE R1	gtaccgagtccagcaatagaC
pSETEryA1 F2b	gactgcacaacctggcatatgCaccTgtgacggaagatcacttcg
pSETEryA1 R2b	gcgttcgggtggtcgaagacatatgATCAAAGGGAAAACTGTCCATAT
pUWL-sgg CF1	ttataCATATGtGAAGAGCtccctatcagtgatagagattg
pUWL-sgg CR1	aacttAAGCTTaAGCaGAAGAGCtataaacgcagaaaggccc
pET.SapI_del F1	gaaaataccgcatcaggcGCTatTCcgcttcctcgctcactgactcgctgc
pET.SapI_del R1	GAatAGCgcctgatgcggtattttctccttacgcatctgtgcggtatttc
KirCI.SOEF1	cggataacaatttcacacaggaaaca

Name	Sequence
KirCI.SOER1	GACGGTGGCGAGCCAG
KirCI.SOEF2	ctggctcgccaccgtc
KirCI.SOER2	CCAAGGTAGTCGGCAAATAAAAGCTT
KirCI.GGF2	CGTATGTAGCTCTTCTatggacgtctacgtgttcccc
KirCI.GGR2	TTGGATAGCTCTTCAAGCtcacccccgtcccatc
FabD.GF1	agcggccatatcgaaggtcgtCatATGCTCGTACTCGTCGCTCCC
FabD.GR1	cgggctttgttagcagccggatccTCAGGCCTGGGTGTGCTCG

C

Name	Sequence
pSETpreEryAI-piece1	ATTATGGTCTCTtaaggacgaggcgatgtcaacctcttgatccttcctatattgttcgcccattgcgtggtcgtcgagtaggggacgcgtggcgga cctgtcaaagctctccgacagtcggactgcacaacctggcatatgcaccTGAGACCATTAT
BryP-block1	cgtatgtagctcttctatgaagaccatctacctttttccccgggcagggctcccagcacaagaagatgggcaaatacctcttcgacaagtaccccgagctg atccaccaggccgaccagcagctccactattccatcaaggagctctgtttggaggacccggaccagctgctcaacaagacccagttcacccagccgg cgctctacatcatcaacgccctctccttcctggacaagatcgagctggaatcccacaaaccgtcgtacgtcgccggccacagcctcggcgagtacaatg cgctgttcgcggccggggccttcgacttcctgacgggcctgaagcttgtgcagaagcggggtctcctcatggaggaggcccccaagggcgcgatggcc gccatcatcggcatcacgcataaccaggtgaagtgtatcctggaggacatcctgaagagctatccaa
BryP-block2	cgtatgtagctcttcttcccccagaagaacatcgacatcgccaacatcaactcggagaagcagttcatcatctcgggcctgtacgacgagatcatcgcgt gcgagaactcgttcaccaagatgggggccaacttcattccgctcaacgtgtccgccgccttccattcccgctacatgaaggatatcgaaatcaagttcga gcagtacctgcaaaagttccagctgaaccccctccggaccccggtcatcagcaactactccgcccggccgtaccccaaggagaactaccgggactac atggtcaagcagatctcgcaccccgtcaagtggtacgagtcgatctcctggctcatccagcaggaccacttcgaattcgaggaagtcggcccggggcgt gtcctcaccaacctgacgaaccagatcaagcagacgccgctccacatctgagcttgaagagctatccaa
FenF-block1	cgtatgtagctcttctatgaacaatctcgccttcctcttcccgggacagggctcgcagttcgtgggcatgggcaagtcgttctggaacgacttcgtgctcgcg aagcgactcttcgaggaggcctccgacgcgatctcgatggacgtgaagaagctgtgcttcgacggggacatgacggagctgacccgcaccatgaac gcgcagccggccatcctcaccgtctccgtgatcgcgtaccaggtgtacatgcaggagatcggtatcaagccgcacttcctggcgggtcactcgctgggc gagtactcggccctggtctgcgccggcgtgctctcgttccaggaagccgtgaaactgatccgccagcggggcatcctcatgcagaacgcggaccccga acagctgggaacgatggcagccatcacgcaggtctatatccagcccctgcaggatctgtgcacggagatcagcaccgaggacttccccgttggtgtcg cgtgcatgaactccgaccagcagcacgtgatcagcggtcaccgtcaggcggtagagttcgtcatcaagaaggcggagcgcatgggcgccaaccaca cgtacctgaatgtgagcgccccgttccactgaagagctatccaa
FenF-block2	cgtatgtagctcttctcactcctccatgatgcgaagcgccagtgagcagttccagacggcactgaaccagtactccttccgtgacgcggagtggccgatc atctccaacgtgacggccatcccctacaacaacggtcactccgtgcgggaacacctgcagacccacatgaccatgcccgtccgctgggccgagtccat gcactacctgctcctccacggagttaccgaggtcatcgagatgggtcccaagaacgtcctggtcgggctcctgaagaaaatcacgaaccacatcgctg cgtacccgctggggcagacgtccgacctgcacctcctgtccgactccgcggagcggaacgaaaacatcgtgaacctgcggaagaagcagctgaac aagatgatgatccagagcatcatcgcccgcaactacaacaaggacgcgaagacctacagcaacctgaccacacccctgttccctcagatccagctcc tgaaggagcgggtagagcgtaaggaagtggaactgagcgccgaggagctggagcacagcatccacctgtgccagctgatctgcgaggcgaagca gctgcccacgtgggagcagctcagaatcctgaagtgagcttgaagagctatccaa
PksC-block1	cgtatgtagctcttctatgatcacctacgtctttcccggccaggggtcccagaagcagggcatgggctcggggctgttcgatgagttcaaggagctgaccg accaagcggacgagatcttgggctactcgatcaagcgtctctgcctggagaacccgtactcgaacctgaacaagacccaattcacgcagccggcact gtacgtcgtgaacgccctctcgtacctcaagaagatccgggacgaggaggtcaagccggacttcgtggcgggccactccctgggcgagtacaacgca ctgttcgcggccgaggcgttcgactttgagaccggactccaactggtccgcaagcgcggcgagctgatgagccttatctcgaacggaggcatggcggc cgtcatggggctgaacgaggagcaagtcgccaaggcgctgaaggagtaccacctctgaagagctatccaa
PksC-block2	cgtatgtagctcttctctccacgacgtcgacatcgcgaacgtcaacgccccgtaccagatcgtgatctcgggcaagaaggacgagatcgagaaggcgg cgtccctcttcgaaaccatgaccgaggtcacgatggtcctgcccctgaacgtgtcgggagccttccactctcgctacatgaacaaggccaaggaggagt tcgaagagttcctccacgccttctacttctcgccgccgtccatccccgtaatctccaacgtgtacgccaagccctacacgtacgagttcatgaagcagacg ctggctgaccagatcaaccactccgtcaagtggaccgacagcatctcctacctgatgaagaagggtcatatggaattcgaggaggtgggcccgggga acgtcctcaccggactaatccaccgcatcaaaaaggatgcggaggctatgccgcgttgagcttgaagagctatccaa

Name	Sequence
PedD-block1	CGTATGTAGCTCTTCTatgggcgagcagcttttcgacaggtttcccaacatcatcgaggccgccaacgacatcctcgggtactcgatcaagac cctgtgcctggaggacccgcagagacagctccgcctcacccagtacacgcaggtggccctctacgtcgtgaacgccctgacctaccgtcagcacctgc agcagggaggcggtctgcccgacttcgtggcggggcactcgctgggagagtacaacgccctggagtcggccggcgtattttcgttcgaggacggcctc cggctggtgcagaagcggggcgacctgatgagccaggcgccgcgtggcgcgatggccgcgatcctcggcatcagcgccgactccgtggccgggat cctggcggagcaaggcctgacgcgcatcgacatcgcaaactacaacgcaccgacccagacgatcatctcgggtctcgaggccgacattcgtgacgc gcaagccgtgttcgaatcctgccaggcgatgtacgtcccgctcaacaccagcggcgccttccacTGAAGAGCTATCCAA
PedD-block2	CGTATGTAGCTCTTCTcactcgcgctatatgcagtccgcacgcgacgagtttgcgcagttcctggaggcgttcgagttccgggacccccagat ccccgtggtggcgaacgtgaccgcgaagccgtacgtcggcaccgaggtcgtccgcaccctggccgaccagctcacgggtagcgtccgatggctgga ctccatgcgcttcctgctcgatcagggcgtgacggagttccgtgaactgggacccggggacgtgctgagcaagctcgtggagagcatccgctcgagcg ccatgtccaagcccgtgtccgagttcgctgcggagaactcccagcagctggtcgacgagtggaaccgcacctgcccgatcggctcgcgcgtccgcgtc aaggggtacgacgacatactggtcacgaagtcgcgtgccgtgctgctgttcgggcaccgggccgccatctacatggaaaactaccagggttacttcgct ctgtccgaggtcgagcccctcatcgaacagcagcccctggtggagaaggtatggTGAGCTTGAAGAGCTATCCAA
DifA-block1	cgtatgtagctcttctatgttcctgttcctgtgccgctgcctgcaccagtcgatctacgagatccgagacgtcttcgtctggcgctccacctgcgacatccaga actggaggtgcggcaaggacgaggacaagatgctgacgttcgtcttcccggggcagggttcccagttcaagggcatgggtgcgggcctgttcgacga gttccaggacctgacccgccaggcggacgacatcctggggtacagcatcgaggagctgtgtctggaggaccccaaccaccagctgggcaagaccc agttcacccagccggcactgtacaccgtgtcggccctctcgtacctgaagaagatgaaggagagcgggcgtgagcccgactacgcggccggacact cgctgggcgagtacaatgccctgttcgcagccgggtgctttgatttcgagacaggactgcaacttgtcaagaagcggggagaactgatgtccaaggcg gcccccggcggatgaagagctatccaa
DifA-block2	CGTATGTAGCTCTTCTGGAatggcggccgtgctgggcttcaccgcagagcaggtgaaggaggtgctgtcagactaccacctcacgggtat cgacatcgccaaccacaacagcccgtcccagatcgtcatcgcggggaccaagcaggacatcgagaaggccggcccggtgttcgaaaaggcgggc gtgcggatgtacctccccctcaatgtgtcgggggcgttccattcgagatatatgaaggacgccgagaaggagttcgcggactacctggaagagacggc cttcctgcccctgcgcttccccgtcatctcgaacctgcacgcggccccgtacaagaacgacgagatcaagaccaacctcacgctgcagatgacaaacc aggtcaagtggacggacacgatccggcgtctcatgggactggagaacaacgagatcgcggaggtgggacccggcgaggtgctgacgaagctgac gcgtcaaatcaagaaggacgccgtgcccttgTGAGCTTGAAGAGCTATCCAA
KirCl-block1	cgtatgtagctcttctatggacgtctacgtgttccccggccaggggcgcaggtcaagggcatggggcgggacctcttcgaccgcttcccggagctcgtg gagcgggcggacgccgtcctgggctactcgatccgggagctctgcctggaggacccgggccgcaacctgcgcgacacccgctacacgcagccggc cctgtacgtggtcggggcgctgtcctggctcgccaccgtccaggagggcggtcggctcccggactacctgctgggccacagcctgggcgagttcgcgg ccctgttcgccgcaggcgtctacgacttcgaaacgggcctccggctggtcgccgagcggggccgcctgatgggccaggtcacgggtggcacgatggc cgccgtgtccgccgtggactcgagcctcgtgcgcgaggtcctgcgcgacgacgaactcagcggcctgtgaagagctatccaa
KirCl-block2	cgtatgtagctcttctctggacatcgccaactacaacgcgcctacccagacggtcgtagccgggccggcggacgctgtcaaccgggccctggcggtctt caaggacaagggtgcccgctgcgcgccgctgaacgtgagcgccccgttccacagccggtacatggcgcaggccgcggaggaattcggccggctcct ggatgcgactgccttcgcagcccccaagatcccggtaatctcgaatgtcgacgcgagaccctacgagcccgacgcggtggccgcgaccctccggcg ccagatcgtgtcgcccgtccgctggacggactccatccggctcctgatgggacgggggtgagcttgaagagctatccaa
VirI-block1	CGTATGTAGCTCTTCTatgacccgtatcttcatgttccccggacagggcgcacagcgtgtgggtatgggacggaccctcctcgaccgattccc ggacctggagcgggaagcctcggacacgctgggctactccctgcgccggctgtgcctcgaagacccggagggacggctgggtaacacgcggtaca cccagcccgcaatgttcgccgtcaacgccctggcccaccgtgcggcagtggaagacggtgcccgtccggacatcgccatcggtcacagcctcggcga gtacaacgcgctggaagccgcgggagtgttcggcttcacggacggcctgcgcctggtcgcagcccgtgccgcggctatggcggaagtcggcggagg cggtatgtcggcagtggtcggactcaccgaaacgaagctgcgcttcctgctgctgcgcgctggcttcgctaccctcgacctggcaaacctgaacaccgc gtcgcagaccgtcctggccggtcccctcgaagacctggaagaggcgggtcaggtcctggaagacgcgggagcccgcatggtacggcgactggacgt gtccggtcccttccattcccgttacatggcaccggctgcggccgcactggtacccctggttcggtccgcacggctgcgtccgccggcgttcccagtgatcg cgaaccggacagcccagccttaccgggccgaactcgctgccgacctcctcctgcagcaaatcgaccacccagtccgctggcacgaaaccgtccgga ccctgctcgacgaaccggacgcagtgttcacggagatcggtgaatcgacggtcctcacatcgatggtacgccagatcaagcgtgacgctgagcctgcc cgtgccgcacgccctcggcaggcagtggcaTGAGCTTGAAGAGCTATCCAA

Name	Sequence
LkcD-block1	CGTATGTAGCTCTTCTatggctggcgatgtcgtcatgttccccggccagggttcccagcgcatcgggatgggaaaggaggtcttcgacgcgta cccccagctgtgcgaccgggccgacgagatcgtgggccactcgctgcgggagctgtgcctcaaggacccggacggtcgcctcaacgagacgagcc gcacccaggaggccgtctacttcgtcagctgcctgatgtacctggcgtacgccgaggaacacggtgcggagcaggtgcgctgcctgaccggtcactcc ctgggtctgtaccccgcactgttcgccgcgggagtcttcgacctgttcgagggcctcgagatcgtgtcccgccgcggagcgctgatgcaggaggcacgg gacggtgccatggtggcagtcctgggaccgcgggcctccgagatcgacgaccacctggcacggctcgagttcttcgacgtggacgtcgcaaactaca actccccggagcaggtcgtcctcagcgcactcaagcctcgcctggaggagctggtgccgcgcctggaggaaacgggtcaccggtgcgtctggctccc cgtctccggagccttccactcccgccacatggagccggctcgcctgcgcttcgcacagttcctgcgcgaccggaccttcaccccgccgaccaagccggt cgtcagcacgaccagcggtcgcaccctgggggctcgccacctcctcgaggagatggtgttccagctcgtcaagcccgtccggtggtggcagacggtg acccacctgtcccgcaccggacacaagaccttcgacgaggtgggtcccgggcgtgtgctcaccaagctgtccgcggagatcctcggtgacgagccttc gccgcaggctgagccgcgtaagTGAGCTTGAAGAGCTATCCAA

Table 4. Plasmid constructs (A), oligonucleotides (B) and gBlocks (C) not described in Materials & Methods sections.

A

Name	No.	Primers	Restriction sites	Backbone	PCR template	Method	Description
pET16b. *Bsu*yngE	720	yngE.F1, yngE.R1	NdeI/BamHI	pET16b	*B. subtilis* colony	ligation	*B. subtilis* putative propionyl-CoA carboxylase carboxytransferase subunit
pET16b. *Bsu*yngH	732	yngH.F2, yngH.R1	NdeI/BamHI	pET16b	*B. subtilis* colony	ligation	*B. subtilis* putative propionyl-CoA carboxylase biotin carboxylase subunit
pCWori-HisN. *Bsu*yngHB. *Eco*birA	736	yngHB.F1, yngHB.R3	NdeI/SalI	pCWOri-AccA2.birA	*B. subtilis* colony	ligation	*B. subtilis* putative propionyl-CoA carboxylase biotin carboxy carrier protein subunit
pET16b. *Cgl*dtsR1	923	Cg.dtsR1.pET16b.F1, Cg.dtsR1.pET16b.R1	NdeI	pET16b	pCDF.P(Tet)-acs.accBC.dtsR1	Gibson assembly	*Corynebacterium glutamicum* ACCase beta (carboxytransferase) subunit, synthetic
pET16b. *Cgl*accE	924	Cg.accE.pET16b.F1, Cg.accE.pET16b.R1	NdeI	pET16b	pCDF.P(Tet)-acs.accBC.dtsR1	Gibson assembly	*Corynebacterium glutamicum* ACCase epsilon subunit, synthetic
pET16b. *Cgl*accBC	925	Cg.accBC.pET16b.F1, Cg.AccBC.pET16b.R1	NdeI	pET16b	pCDF.P(Tet)-acs.accBC.dtsR1	Gibson assembly	*Corynebacterium glutamicum* ACCase alpha (biotin carboxylase/biotin carboxy carrier protein) subunit, synthetic
pCDFDuet-1-dszAT-ø	1191	pCDFDuet_DszsAT.F1, pCDFDuet_DszsAT.R1	NcoI/HindIII	pCDFDuet-1	pFW3	Gibson assembly	CloDF13 Sp^R lacI pT7 dszAT pT7 T7-term. **DszAT has C-terminal TSLRPHNA appended.**
pET16b. *Eco*prpE	1196	prpE F2, prpE R2	NdeI/BamHI	pET16b	*E. coli* DH10b colony	Gibson assembly	pMB1 Cb^R lacI pT7 *Eco*prpE T7-term
pCDFDuet-1-ø-*SCo*.matB	1199	pCDF-Duet MatB F1, pCDF-Duet MatB R2	NdeI/KpnI	pCDFDuet-1	*S. coelicolor* gDNA	Gibson assembly	CloDF13 Sp^R lacI pT7 pT7 *Sco*matB T7-term
pCDFDuet-1-dszAT-*SCo*.matB	1200	pCDF-Duet MatB F1, pCDF-Duet MatB R2	NdeI/KpnI	pCDFDuet-1-dszAT-ø	*S. coelicolor* gDNA	Gibson assembly	CloDF13 Sp^R lacI pT7 dszAT pT7 ScomatB T7-term. **DszAT has C-terminal TSLRPHNA appended.**

Name	No.	Primers	Restriction sites	Backbone	PCR template	Method	Description
pCOLADuet-1-*Sco*accA2-*Sco*pccBE.*Eco*birA	1210	PCCase_birA.F1, PCCase_birA.R1	KpnI	pCOLADuet-1-accA2-pccBE	pBAD33.birA	Gibson assembly	*S. coelicolor* PCCase: pTrc AccA2 pT7 pccB pccE *Eco*birA
pCDFDuet-1-RBS500.dszAT.ø	1224	pCDF-DszAT-500-F1, pCDFDuet_DszsAT.R1	NcoI/HindIII	pCDFDuet-1	pFW3	Gibson assembly	CloDF13 Sp^R lacI pT7 calc.500-rate RBS dszAT pT7 T7-term. **DszAT has C-terminal TSLRPHNA appended.**
pET28a.*Sco*matB.500dszAT	1226	pET28a-matB-DszAT-R1, pET28a-matB-500DszAT-F1	XhoI	pET28a.*Sco*matB	pCDFDuet-1-RBS500.dszAT.ø	Gibson assembly	pMB1 Km^R lacI pT7 His6.*Sco*matB calc.500-rate RBS dszAT T7-term
pBAD18-Cb.500*Sco*matC	1254	MatC F1.474, MatC G R2, MatC NF1/ NR1	KpnI/HindIII	pBAD18-Cb	*S. coelicolor* gDNA; neighborhood PCR	Neighborhood PCR, Gibson assembly	Cb^R pBad calc.474-rate RBS *Sco*matC
pBAD18-Cm.500*Sco*matC	1255	MatC F1.474, MatC G R2, MatC NF1/ NR1	KpnI/HindIII	pBAD18-Cm	*S. coelicolor* gDNA; neighborhood PCR	Neighborhood PCR, Gibson assembly	Cm^R pBad calc.474-rate RBS *Sco*matC
pCDFDuet-1.nphT7-0	1280	nphT7 pCDF GF1/GR1	NcoI/HindIII	pCDFDuet-1	pTRC33-npht7-phaB	Gibson assembly	CloDF13 Sp^R lacI pT7 nphT7 pT7 T7-term
pCDFDuet-1.ø-phaB	1281	phaB pCDF GF1/GR1	NdeI/KpnI	pCDFDuet-1	pTRC33-npht7-phaB	Gibson assembly	CloDF13 Sp^R lacI pT7 pT7 phaB T7-term
pCDFDuet-1.nphT7-phaB	1282	phaB pCDF GF1/GR1	NdeI/KpnI	pCDFDuet-1.nphT7-ø	pTRC33-npht7-phaB	Gibson assembly	CloDF13 Sp^R lacI pT7 nphT7 pT7 phaB T7-term
pSV272.1-Mod6noTE	1283	MBP-M6 F1, MBP-M6 R1	SfoI	pSV272.1	PRSG54	Gibson assembly	His6-MBP-TEV-DEBS module 6 fusion, no thioesterase domain (cutoff from published HisN-ACP6 construct)
pSV272.1-Mod6AT^{0}noTE	1284	MBP-M6 F1, MBP-M6 R1	SfoI	pSV272.1	pAYC138	Gibson assembly	His6-MBP-TEV-DEBS module 6 fusion/Ser>Ala mutant, no thioesterase domain (cutoff from published HisN-ACP6 construct)
pET28a.HisN.PIKS-KR1	1324	PIKS_KR1 F2, PIKS_KR1 R2	BsaI	pET28a(gg)	PIKS-KR1-block2	Golden Gate / Gibson assembly	Pikromycin synthase module 1 KR (synthetic). To correct design error, GG to assemble vector, blocks 1 and 3; Gibson on PCR of block2 (sequential one pot)
pET28a.HisN.GDH	1325	GDH F1, GDH R1	BsaI	pET28a(gg)	*B. subtilis* colony	Golden Gate	*B. subtilis* glucose 1-dehydrogenase
pXHB.ø-*Rpa*matB.*Eco*acs	1519	pCDF-ACS GF1/GR1 (*Table 2, B*)	XbaI/HindIII	pXHB.ø-matB	pCDF.P(Tet)-acs.accBC.dts R1	Gibson assembly	Empty pBAD site
pXHB-*Sco*matC-*Rpa*matB.*Eco*acs	1550	pCDF-ara-ScoMatC GF1/GR1 (*Table 2, B*)	XbaI/HindIII	pXHB.ø-*Rpa*matB.*Eco*acs	pBAD18-Cb.500matC	Gibson assembly	CloDF13 Sp^R lacI araC pBAD *Sco*matC rrnB-term pT7 *Rpa*matB *Eco*acs T7-term
pPOL1-bc	1590	PhaB GF2 / FPHB PhaC R1 (*Table 2, B*)	NcoI/BamHI	pTRC99a	pPOL1	Gibson assembly	pPOL1 without nphT7
pPOL3	1627	TetR-pTet-pPOL GF1, TetR-pTet-nphT7 GR1 / nphT7 pCDF G F1, NphT7 pPOL GR1	NdeI	pPOL1	pCDF.P(Tet)-acs.accBC.dts R1	Gibson assembly	ColE1 Cb^R $lacI^q$ tetR pTet nphT7 pTrc phaBC rrnB-term

Name	No.	Primers	Restriction sites	Backbone	PCR template	Method	Description
pPOL2	1668	PhaC F10, rrnB R2	AatII/XbaI	pPOL1	pBT33-phaABCP-crt	Gibson assembly	pPOL1 but with synthetic phaP1 gene after phaC
pXHB3. *Eco*prpE	1695	pXHB3.GF1, pXHB3.GR1	NdeI/NotI	pXHB1.*Eco*prpE	n/a	Gibson assembly	pXHB1.EcoprpE but replaced T7 promoter with T5
pET21c. KSAT03.phaB	1967	KSAT3-PhaB GF1, KSAT3-PhaB GR1	EcoRI/HindIII	pET21c-KSAT^{0}3	pPOL1-nb	Gibson assembly	pMB1 CbR lacI pT7 DEBS-KS3AT3^0 phaB T7-term
pPOL1Cm-bc	2087	n/a	pPOL1-bc: ApaI/HindIII pTRC33: EcoRV/XbaI	pPOL1-bc / pTRC33	n/a	Gibson assembly	pPOL1-bc with backbone exchanged from pTrc99a to pTrc33

B

Name	Sequence
yngE.F1	attCATATGctgatggattatgaaaagg
yngE.R1	ggatccctcgagcccggggtcgacgagctcggtactctagttacaccggataaaccgg
yngH.F2	gaggtgaaCATATGtttacaaaagtactgatcgc
yngH.R1	attTCTAGAGGATCCtttataggtgctgtttcaaaaag
yngHB.F1	attCATATGacggttagcatacaaatg
yngHB.R3	attGTCGACGGATCCCCCGGGCTCGAGGAGCTCcaaattattgagtggaattgctc
Cg.dtsR1.pET16b.F1	cagcagcggccatatcgaaggtcgtCATATGACCATTTCCAGCCCG
Cg.dtsR1.pET16b.R1	CTTTGTTAGCAGCCGGATCCTCGAGcataTTACAGCGGCATATTACCATGC
Cg.accE.pET16b.F1	cagcagcggccatatcgaaggtcgtCATATGtccgaagagaccacgc
Cg.accE.pET16b.R1	CTTTGTTAGCAGCCGGATCCTCGAGcataTTAGAAAAAGTTCACGTTCTGAAACG
Cg.accBC.pET16b.F1	cagcagcggccatatcgaaggtcgtCATATGAGCGTTGAAACCCGC
Cg.AccBC.pET16b.R1	CTTTGTTAGCAGCCGGATCCTCGAGcataTTACTTAATCTCCAGCAGAACGAC
pCDFDuet_DszsAT.F1	gtttaactttaataaggagatataccATGAAAGCATACATGTTTCCCGGGC
pCDFDuet_DszsAT.R1	CTTAAGCATTATGCGGCCGCAAGCTTGTTACGACGACGAGGGGCTGGG
prpE F2	cagcagcggccatatcgaaggtcgtcatATGTCTTTTAGCGAATTTTATCAGCGTTCGATTAACGAA
prpE R2	ctttgttagcagccggatcctcgagcataCTACTCTTCCATCGCCTGGCGGATCT
pCDF-Duet MatB F1	ttagttaagtataagaaggagatataCATatgtcctctctcttcccggccctct
pCDF-Duet MatB R2	gtttctttaccagactcgagggtaccTCAGTCACGGTTCAGCGCCCG
PCCase_birA.F1	Gggccgcacctactggccgacctgaaggagatatacgatgaaggataacaccgtgccactgaaattg
PCCase_birA.R1	Cagcggtttctttaccagactcgagggtaccttatttttctgcactacgcagggatatttcaccgccc
pCDF-DszAT-500-F1	ccctgtagaaataattttgtttaactttaatCTCAAGCGTAAAGTTCCACAAGACGCAatgaaagcatacatgtttcccgggc
pET28a-matB-DszAT-R1	gatctcagtggtggtggtggtggtgTTACGACGACGAGGGGC
pET28a-matB-500DszAT-F1	ggcgctgaaccgtgactgactcgagCTCAAGCGTAAAGTTCCACAAG
MatC F1.474	Ggctagcgaattcgagctcggtaccgttttagccaccagttcgaaaaaatcaactaatatgtcccccgaactcatctc
MatC G R2	caggctgaaaatcttctctcatccgccaaaacagccaagcttCTACCCGAAGCCGGGCAC
MatC NF1	gacccttgacgtgatcc
MatC NR1	ccgggaagagagaggaca
nphT7 pCDF G F1	gtttaactttaataaggagatataccATGACCGACGTTCGTTTTCG
nphT7 pCDF G R1	cttaagcattatgcggccgcaagcttgTTAccactcgatcagcgcg
phaB pCDF G F1	ttagttaagtataagaaggagatataCATATGacccagcgcatcgcttac
phaB pCDF G R1	gtttctttaccagactcgagggtaccTTAgcccatgtgcaggccac
MBP-M6 F1	caacctcgggatcgaggaaaacctgtattttcagggcgccATGAGCGGTGACAACGGC

Name	Sequence
MBP-M6 R1	cgtttgatctcgagtgcggccgcAAGCTTtcaGAGCTGCTGTCCTATGTGGTC
PIKS_KR1 F2	AACGTTGGGCAGGTCTGGTTGATCTgcctgcccagccggatg
PIKS_KR1 R2	TTGCACGATGTGCACGACGAACCTGttccgctgtcagggtatcaac
GDH F1	ATTATGGTCTCTTATGTATCCGGATTTAAAAGG
GDH R1	ATAATGGTCTCAAAGCTTAACCGCGGCCTGCCTGG
PhaB G F2	cggataacaatttcacacaggaaacagaccaGgagatatataAtgacccagcgcatcgc
TetR-pTet-pPOL GF1	tacgcatctgtgcggtatttcacaccgcatCGCTTTATGAATCTAAAGGGTGG
TetR-pTet-nphT7 GR1	GTCATGGTATATCTCCTTATTAAAGggatcCtgaagacgaaagggcctcg
NphT7 pPOL GR1	TCAGAGCAGATTGTACTGAGAGTGCACCATctcgaGttaccactcgatcagcgcg
PhaC F10	ggactttgccgacacgg
rrnB R2	gaccgcttctgcgttctg
pXHB3.GF1	caaatatgtatccgctcatgagacGgttttgcaccattcgatg
Name	**Sequence**
pXHB3.GR1	caggcgggcaaacagattcgcgttcaTatgtatatctccttgagctctgtg
KSAT3-PhaB GF1	cgagctggcctaccgcgttgcttaaGAATTCAAGGAGATATATAATGACCCAGCG
KSAT3-PhaB GR1	gtggtggtgctcgagtgcggccgcaagcttTTAGCCCATGTGCAGGCCAC

C

Name	Sequence
PIKS-KR1-block1	ttatggtctcttatggcaaccggtgatgattggcgttatcgtattgattggaaacgtctgcctgcagcagaaggtagcgaacgtaccggtctgagcggtcgtt ggctggcagttacaccggaagatcatagcgcacaggcagcagcagttctgaccgcactggttgatgccggtgcaaaagttgaagttctgacagccggt gcagatgatgatcgtgaagcactggcagcacgtctgacagcgctgaccacaggtgatggttttaccggtgttgttagcctgctggatggtctggttccgca ggttgcatgggttcaggccctgggtgatgcaggtattaaagcaccgctgtggtctgttacccagggtgcagttagcgttggtcgtctggatactccggcaga tcctgatcgtgcaatgctgtggggtctgggtcgtgttgttgcactggaacatccggaacgttgggcaggtctggttgatctcgaatgagaccatta
PIKS-KR1-block2	Ttatggtctctcgaagcctgcccagccggatgcagcagcactggcccatctggttaccgcactgagcggtgccaccggtgaagatcagattgcaattcg taccaccggtctgcatgcacgtcgtctggcacgtgcaccgctgcatggtcgtcgtccgacccgtgattggcagccgcatggcaccgttctgattacaggtg gtacaggtgcactgggtagccatgcagcacgttggatggcacatcatggtgcagaacatctgctgctggttagccgtagcggtgaacaggcaccgggt gcaacccagctgaccgcagaactgaccgcaagcggtgcccgtgttaccattgcagcatgtgatgttgcagatccgcatgcaatgcgtaccctgctggac gcaattccggcagaaacaccgctgacagcagttgttcataccgcaggcgctctggatgatggtattgttgataccctgacagcggaattcctgagaccatt a
PIKS-KR1-block3	Attatggtctctttcccaggttcgtcgtgcacatcgtgcaaaagcagttggtgcaagcgttctggatgaactgacacgtgatctggatctggatgcatttgttct gtttagcagcgttagcagcacactgggtattccgggtcagggtaattatgcaccgcataatgcatatctggacgccctggcagcccgtcgtcgcgcaaca ggtcgtagcgcagtgagcgttgcctggggtccgtgggatggtggtggtatggcagccggtgatggtgttgccgaacgtctgcgtaatcatggtgttccggg tatggacccggaactggcgctggcagcgctggaaagcgcactgggacgtgatgaaaccgcaattaccgttgcagatatcgattgggatcgtttttatctg gcatatagcagcggtcgtccgcagccgctggttgaagaactgccggaagttcgtcgtattattgatgcacgtgattaagctttgagaccattat