Title
Molecular insights into fluorine chemistry in living systems

Author
Weeks, Amy Marie

Publication Date
2013

Peer reviewed|Thesis/dissertation

Molecular insights into fluorine chemistry in living systems

by

Amy Weeks

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 J. Christopher Anderson

Fall 2013

Molecular insights into fluorine chemistry in living systems

by Amy Weeks

Abstract

Molecular insights into fluorine chemistry in living systems

by

Amy Weeks

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michelle C. Y. Chang, Chair

Based on its unique elemental properties, fluorine has emerged as an important design element in synthetic pharmaceuticals, but also a difficult substituent to incorporate into small molecules using synthetic chemistry. Synthetic biology approaches represent an attractive alternative that could allow us to harness the catalytic power and specificity of enzymes for fluorine incorporation into complex small molecules. However, fluorinated natural products are rare and derive from a single pathway for biosynthesis of fluoroacetate, an antimetabolite poison. The fluoroacetate-producing bacterium *Streptomyces cattleya* has evolved a fluoroacetyl-CoA thioesterase (FlK) that is involved in fluoroacetate resistance. Using FlK as a model system, we have focused on understanding enzymatic fluorine selectivity with the goal of defining principles for the design of fluorine-specific enzymes that can be applied in the development of new methods for preparation of fluorinated molecules.

We have demonstrated that FlK exhibits a remarkable 10^6-fold selectivity for hydrolysis of fluoroacetyl-CoA compared to acetyl-CoA, an abundant central metabolite and cellular competitor. This selectivity is based on a decrease in K_M (10^2) and an increase in k_{cat} (10^4) for the fluorinated substrate. Through x-ray crystallographic studies, we have identified a unique hydrophobic 'lid' in our crystal structure of FlK that shields the active site from water. Mutation of a key lid residue (Phe 36) led to a loss of fluorine-based binding specificity, which was correlated with the appearance of additional ordered water molecules at the active site in the mutant crystal structure. A crystal structure of the FlK product complex revealed that Phe 36 is dynamic, undergoing a conformational change when the active site is occupied with product. By studying the binding of a series of non-hydrolyzable substrate analogs, we have shown that the lid confers the fluorinated substrate with an entropic binding advantage that is lost in lid mutants, which may be related to water release from the 'polar hydrophobic' C-F unit. Kinetic studies revealed that Phe 36 controls the substrate off rate, providing the fluorinated substrate with a kinetic as well as a thermodynamic advantage.

We have also shown that catalytic specificity in FlK is based on utilization of an unusual reaction pathway initiated by deprotonation at the α-carbon for hydrolysis of fluoroacetyl-CoA but not acetyl-CoA. The enolate that is formed can then breakdown through a putative ketene intermediate to give the same product that would be produced by a canonical hydrolysis mechanism. FlK therefore represents a rare example in which the existence of two reaction pathways in the same active site controls substrate selectivity rather than product outcome. Although catalytic fluorine discrimination occurs mainly in one step of the reaction mechanism, we have demonstrated that fluorine is specifically recognized by the enzyme in both the acylation and deacylation steps of the reaction mechanism. These studies have defined the basis of fluorine selectivity in a naturally occurring enzyme-substrate pair, with implications for drug design and development of fluorine-selective biocatalysts.

Table of Contents

Appendices

List of Figures, Schemes, and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Abbreviations

SAM	*S*-adenosyl-L-methionine
SAH	*S*-adenosyl-L-homocysteine
P450s	cytochrome P450 monooxygenases
dNMP	deoxynucleotide monophosphate
dNDP	deoxynucleotide diphosphate
dNTP	deoxynucleotide triphosphate
dUMP	deoxyuridine monophosphate
RNR	ribonucleotide reductase
TCA	tricarboxylic acid
FAS	fatty acid synthase
RpoB	RNA polymerase subunit B
NAD^+	nicotinamide adenine dinucleotide
$NADP^+$	nicotinamide adenine dinucleotide phosphate
ATP	adenosine-5'-triphosphate
5'-FDA	5'-fluoro-5'-deoxyadenosine
5'-ClDA	5'-chloro-5'-deoxyadenosine
PET	positron emission tomography
5-FDRP	5-fluoro-5-deoxyribose-1-phosphate
5-FDRu1P	5-fluoro-5-deoxy-D-ribulose-1-phosphate
ACS	acetyl-CoA synthetase
AckA	acetate kinase
PTA	phosphotransacetylase
HAD	haloacid dehalogenase
ACP	acyl carrier protein
LB	Luria-Bertani
IPTG	isopropyl-β-D-thiogalactopyranoside
Tris-HCl	tris(hydroxymethyl)aminomethane hydrochloride
DTT	dithiothreitol
DTNB	5, 5'-dithiobis-2-nitrobenzoic acid
TEMED	*N*, *N*, *N*', *N*'-tetramethyl-ethane-1,2-diamine
TEV	tobacco etch virus
MBP	maltose binding protein
DMF	*N*, *N*-dimethylformamide

MES	2-(*N*-morpholino)ethanesulfonic acid
CAPS	*N*-cyclohexyl-3-aminopropanesulfonic acid
TFA	trifluoroacetic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
TCA	trichloroacetic acid
TCEP	tris(2-carboxyethyl)-phosphine hydrochloride
HMBC	heteronuclear multiple bond correlation
4-HBT	4-hydroxybutyryl-CoA thioesterase
KIE	kinetic isotope effect
HATU	1-[bis(dimethylamino)methylene-1H-1,2,3,-triazolo[4,5-b] pyridinium3-oxid hexafluorophosphate
PanK	pantothenate kinase
PPAT	phosphopantethiene adenylyltransferase
DPCK	dephospho-CoA kinase
ITC	isothermal titration calorimetry
CPMG	Carr-Purcell-Meiboom-Gill
HSQC	heteronuclear single quantum coherence
Bis-Tris	2-bis(2-hydroxyethyl)amino-2-(hydroxymethyl)-1,3-propanediol
Bis-Tris propane	1,3-bis[tris(hydroxymethyl)methylamino] propane
THF	tetrahydrofuran
EDTA	ethylenediaminetetraacetic acid
FAIDH	fluoroacetaldehyde dehydrogenase
ThrAld	fluorothreonine transaldolase
DMSO	dimethyl sulfoxide
TES	[tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid
MS	mannitol-soy flour
DNA	Difco nutrient agar
MM	minimal media
GYM	glucose, yeast extract, malt extract
SMM	supplemented liquid minimal media
PNP	purine nucleoside phosphorylase

Acknowledgments

There are many people to whom I owe a debt of gratitude for their contributions to my graduate education. I thank Prof. Michelle Chang for taking me in and for her mentorship and advice over the years, as well as for her help in editing this dissertation. Michelle's creativity, intensity, and scientific fearlessness have been an inspiration that I will carry with me going forward. I have been lucky to have Prof. Matt Francis, who has offered invaluable advice, pep talks, and scientific guidance, and Prof. Christopher Anderson as the members of my committee. Prof. Jennifer Doudna has been a source of optimism and encouragement. Dr. Stuart Licht, my undergraduate advisor, taught me about the excitement of discovery and encouraged me to come to Berkeley. Prof. JoAnne Stubbe stoked my interest in enzyme mechanism during my undergraduate years, and her knowledge, enthusiasm, and intution have had a strong impact on my graduate research. I also thank Prof. Judith Klinman, Prof. Jack Kirsch, and Prof. Michael Marletta for the insights and inspiration provided by both their work and their mentorship.

The research presented here has also benefited from a number of collaborations, both formal and informal. Scott Coyle, Dr. Martin Jinek, and Dr. Dipali Sashital taught me everything I know about x-ray crystallography, including both the technical and superstitious aspects. I am grateful to Dr. Brian Smith, Dr. Michael Winter, and Dr. Emily Weinert for assistance with the rapid chemical quench instrument, and to the other members of the Marletta lab for their hospitality and good nature with respect to the cacophony produced by the RCQ. Dr. Eric Underbakke provided helpful assistance and encouragement with initial mass spectrometry studies, and Dr. Lori Kohlstaedt performed LC-MS/MS analysis to identify the covalent intermediate in the FlK reaction. Dr. Ross Wilson provided valuable advice and training for the ITC experiments reported here. I also thank Dr. Christian Canlas for NMR training and Dr. Jeffrey Pelton for running protein NMR experiments.

I must also pay tribute to the members of the Chang lab, who have been a great intellectual resource and have made it fun to come to work. Mark Walker got the fluorine project up and running and laid the groundwork for the research described here, and should also be acknowledged for his encyclopedic knowledge of liquid chromatography, kindness to squirrels, and contributions in ensuring that anyone who walks into the Chang lab knows that we are serious about waffles. Maggie Brown has been a pleasure to work with and should be recognized for her hard work as safety officer, which has occasionally involved standing by the synthesis hood with a fire extinguisher just in case. Brooks Bond-Watts is to be commended for ensuring that things in the lab run smoothly and that we are never without a lab motto. He has also been a great friend, supporter, and dedicated believer that everything is going to be okay. I thank him for being who he is, and for never getting upset when I chew all of his gum. Stephanie Jones always maintains a sense of humor and is a model of calmness and composure under pressure. As a bonus, she also has excellent taste in music. I've very much enjoyed Ben Thuronyi's centrifuge-based humor over the years, and it is always a pleasure to talk science with him. I'm also grateful for his careful attention to selecting an appropriately adorable lab calendar each year. Although Matt Davis is sometimes quiet, he is probably the funniest person that I've ever met. I am also thankful for his generosity with baked goods, and confident that he will be victorious in the Chemical Biology Bake-Off one of these years. I've had lots of fun talking about the oddities of *Euglena* and the dialects of Pennsylvania with Mike Blaisse and have been inspired by his hard work and infectious enthusiasm. For this, he needs thanked. Jon McMurry is legendary for his limitless supply of interesting ideas and strange facts, and I am most grateful for his willingness to help out whenever he can. It's been great having him as a partner in crime on the fluorine project. Joe Gallagher is a fellow late-night lab stalwart whose boundless optimism and positive outlook on life are a breath of fresh air. I've often commiserated with Jeff Hanson about the experience of being a gradudate student in a young lab and about paper submission (and resubmission). Miao Wen has

provided lots of scientific guidance over the years, and is a shining example of postdoctoral competence. I've only just gotten to know Omer Ad, Jase Gehring, and Vivian Yu, but our brief overlap has been a pleasure and I wish them the best of luck with their projects. I am also grateful to Neil Lim, Kevin Wang, and Vincent Yeh, talented and enthusiastic undergraduates who have worked with me.

I am very lucky to have found many wonderful friends at Berkeley. Albert Lang has been a constant source of support since acting as my rotation mentor in the Klinman lab, and the sage advice that he gave me about my qualifying exam hangs above my desk to this day. I have many fond memories of playing softball, closing down 7th floor happy hour, watching A's games, and walking to the The Graduate with Katie Berry, Bryan Clarkson, Sandro Ataide, and Dipa Sashital. I will always be grateful for their friendship.

Most importantly, I would like to thank my family for their unwavering support. My mom, dad, and siblings Megan, Mary, and Rob have been a constant source of encouragement. I am also grateful to Kathy and Pat Coyle and Lizzie and Tyler Phelps for warmly welcoming me into their family. Finally, love and gratitude to my best friend and partner Scott Coyle, who has never failed to believe in me. Happiness is building our life together.

Chapter 1: *Introduction*

1.1 Organohalogen natural products

Introduction. Organohalogens comprise a diverse class of nearly 5,000 known natural products with a wide range of biological activities (*Figure 1.1*) [*1-3*]. Halogen modifications on secondary metabolites are observed at a variety of aliphatic and olefinic positions, as well as on aromatic and heterocyclic rings [*3*]. Organochlorines and organobromines are the most abundant halogenated natural products, with about 2000 of each known to date [*2*]. Although about equal numbers of chlorinated and brominated natural products are known, organochlorines are typically produced by terrestrial microorganisms, while organobromines are usually found in marine organisms [*4*]. At least 120 organoiodine natural products have also been identified, with

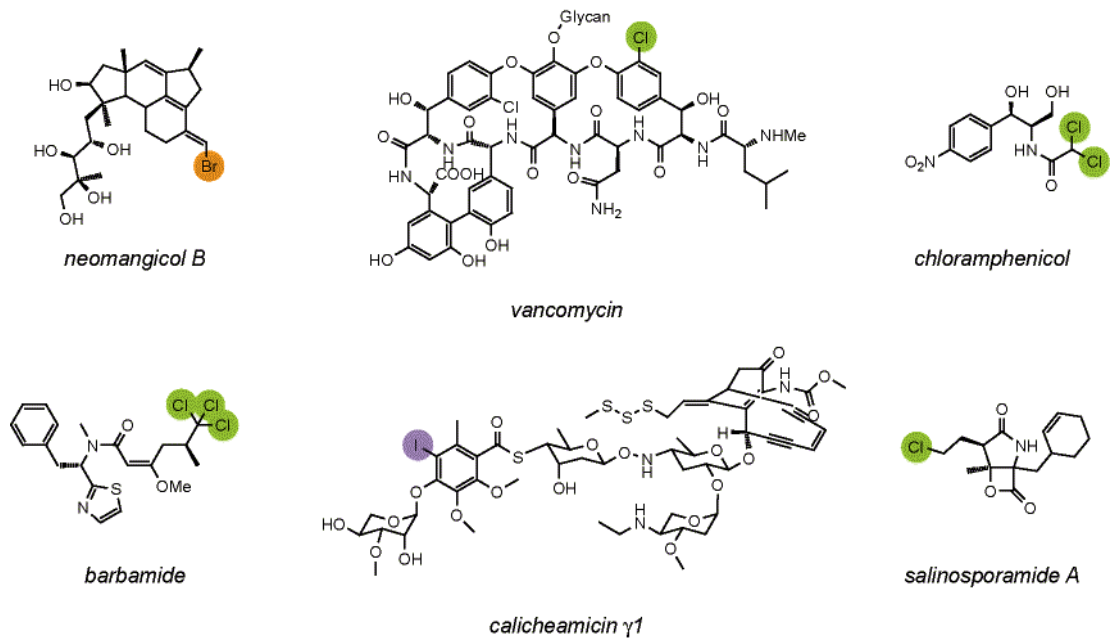


Figure 1.1. *Organohalogen natural products.*

their lower representation likely reflecting the low natural abundance of iodine [*2, 3*]. Despite fluorine's position as the most abundant halogen in the earth's crust, only 30 organofluorine natural products are known to date [*5*]. This may reflect the low bioavailability of fluoride, which is largely found in minerals, as well as the high energetic cost of desolvation or oxidation of fluoride [*5, 6*].

The vast majority of halogen modifications observed in natural products are installed using enzyme-catalyzed oxidative chemistry (*Figure 1.2*), which relies on oxygen or hydrogen peroxide as a co-substrate [*3*]. Indeed, halogenation enzymes appear to be evolutionarily derived from hydroxylation enzymes and share many mechanistic similarities with hydroxylation chemistry [*3, 7*]. The halide ion may be activated to either a formal X^+ equivalent for reactivity at an electron-rich site, or to an $X\cdot$ equivalent for reactivity at an unactivated carbon center [*3*]. The carbon-fluorine bond stands alone as the only natural product halogen modification that is formed exclusively non-oxidatively, as fluoride appears to be completely inert to biological oxidants (*Figure 1.2*) [*6, 8*]. The only known enzyme that catalyzes carbon-fluorine bond formation activates fluoride as a nucleophile and requires an activated electrophile as a co-substrate [*6, 9*].

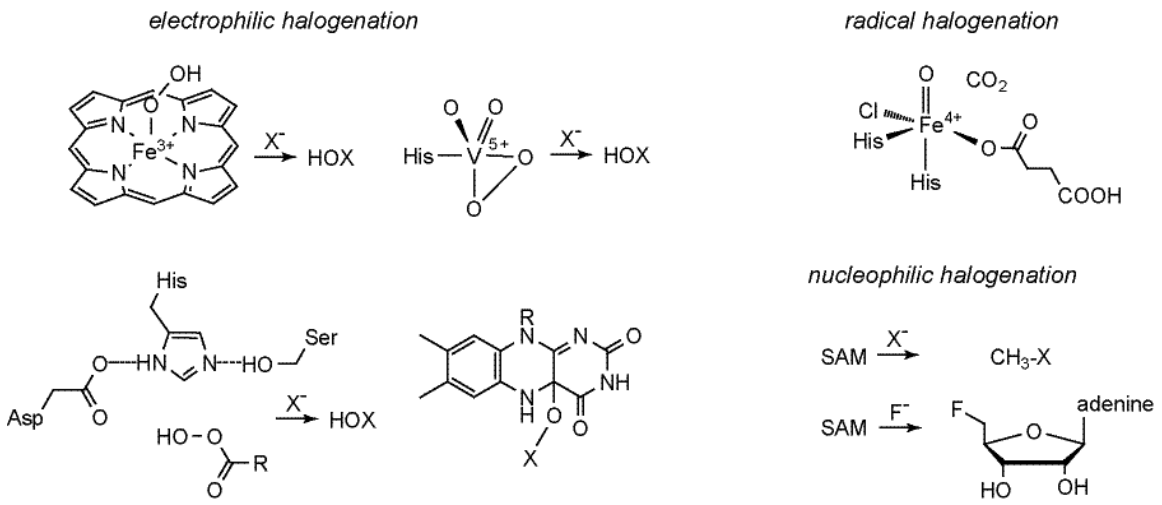


Figure 1.2. *Enzymatic halogenation strategies observed in nature. Halogenation may occur through electrophilic, radical, or nucleophilic mechanisms.*

Halogenation by X^+. The most common enzymatic halogenation strategy involves oxidation of the halide ion to form an X^+ equivalent in the form of an enzyme-bound hypohalite ion [*3, 10*] (*Figure 1.2*). All of the enzymes that generate X^+ equivalents require that the substrate to be halogenated be electron-rich, giving it the nucleophilic character needed to react with the halogen electrophile [*3*]. Formation of X^+ may be achieved using hydrogen peroxide to provide the oxidizing equivalents, as occurs in the heme-dependent haloperoxidases [*11*] and the vanadium-dependent bromoperoxidases [*12*]. Another strategy, observed in the flavin-dependent halogenases relies on molecular oxygen as the source of oxidizing equivalents [*13*]. In this case, it is believed that the X^+ equivalent is transferred from cofactor-generated hypohalite to an active site lysine, which ultimately serves as the X^+ donor [*14*]. The only example of electrophilic halogenation that occurs in the absence of a metal cofactor or prosthetic group has been observed in the perhydrolases, which catalyze perhydrolysis of a serine-bound acyl-enzyme intermediate to generate a peracid [*15, 16*]. Because the peracid is a strong oxidant, it can oxidize halide ions to generate hyophalous acids, which act as the proximal X^+ donor. These reactions lack stereo- and regiospecificity and maybe represent a biologically insignificant side reaction of hydrolytic enzymes [*17*].

Halogenation by X•. A more recently discovered mechanism of enzymatic halogenation, catalyzed by mononuclear non-heme iron enzymes that are dependent on molecular oxygen and α-ketoglutarate, involves one-electron oxidation of the halide ion to form an X· halogen donor (*Figure 1.2*) [*18-20*]. These enzymes are evolutionarily related to the non-heme iron, O_2, and α-ketoglutarate-dependent hydroxylases, whose iron is bound through a 2 His, 1 Asp/Glu 'facial triad' [*21*]. In the non-heme iron halogenases, the Asp/Glu ligand is missing, but is replaced by a halide ion [*22*]. The reaction mechanism for both the hydroxylases and halogenases involves formation of an Fe(IV)-oxo intermediate capable of abstracting H· from the unactivated carbon center of a methyl group on the substrate [*23, 24*]. In the hydroxylases, the substrate radical is captured through ·OH rebound [*23*], while in the halogenases, X· rebound occurs instead to give the halogenated product [*22, 24*].

Halogenation by X^-. Enzymatic halogenation mechanisms that utilize the halide ion as a nucleophile appear to be the least commonly employed in terms of the diversity of natural products generated (*Figure 1.2*) [*8*]. A common feature of all nucleophilic halogenases known to date is that *S*-adenosyl-L-methionine (SAM), a highly activated electrophile, acts as the substrate for halogenation.

The first nucleophilic halogenase discovered was methyl chloride transferase [*25*], an enzyme in which the halide reacts as a nucleophile with SAM to displace *S*-adenosyl-L-homocysteine (SAH) and generate methyl chloride (*Figure 1.2*). This enzymatic activity has been observed in fungi, algae, and plants [*25*], including the model plant *Arabidopsis thaliana* [*26*], and is believed to be the main source of biogenic methyl halides [*27*]. Kinetic studies have demonstrated that methyl halide transferases may exhibit specificity for chloride, bromide, or iodide [*28, 29*], and will also accept alternative nucleophiles such as thiocyanate [*29*].

A second class of nucleophilic halogenases utilizes a different mode of SAM reactivity, with the nucleophilic attack occurring on the 5' carbon of the adenosine moiety to displace L-methionine (*Figure 1.2*) [*6, 30, 31*]. The founding member of this class of enzymes is the fluorinase from *Streptomyces cattleya*, which catalyzes the formation of 5'-fluoro-5'-deoxyadenosine [*6, 30*]. At the present time, the fluorinase remains the only enzyme for carbon-fluorine bond formation, which is catalyzed by desolvation of the fluoride ion to enhance its nucleophilicity. A homologous chloride-specific enzyme was later discovered in *Salinispora tropica* that forms 5'-chloro-5'-deoxyadenosine [*32*], a precursor to the polyketide salinosporamide A [*33*].

Conclusion. Halogen modifications are a common motif in natural products and are typically installed through oxidative chemistry. Fluorination represents the only halogen modification that is formed exclusively through non-oxidative chemistry. Fluorine's inertness toward biological oxidants may account for the unique mechanism used for formation of carbon-fluorine bonds in natural products, but also the relative paucity of biogenic organofluorines.

1.2 Fluorine in drug design

Introduction. Natural products have historically had a strong impact on the development of new drugs [*34*]. Between 1981 and 2007, >50% of new pharmaceuticals were based on natural products [*34, 35*]. Of the thirteen natural product-derived drugs approved between 2005 and 2007, five of these were founding members of new classes of therapeutics [*35*]. In light of the underrepresentation of organofluorines among natural products [*3*], it seems surprising that >20% of pharmaceuticals contain at least one carbon fluorine bond [*36*]. However, introduction of a fluorine modification has been observed to confer favorable properties on drug molecules, such as enhanced stability and bioavailabilty and improved binding to macromolecular targets (*Figure 1.3*) [*37-40*]. The use of fluorine in pharmaceuticals has had an impact in a wide array of disease areas, including cancer, depression, viral infection, and bacterial infection, and is observed in blockbuster drugs including fluoxetine (Prozac®) and atorvastatin (Lipitor®) [*40*].

Fluorine effects on the physicochemical properties of small molecules. Fluorine's ability to tune the properties of drug molecules can be explained in terms of its unique position at the most electronegative element [*41*]. This high electronegativity results from fluorine's high nuclear charge in combination with the fact that its valence electrons are closely held in the 2*p* orbital [*42*]. This gives fluorine the smallest atomic radius of any period 2 element. Fluorine's

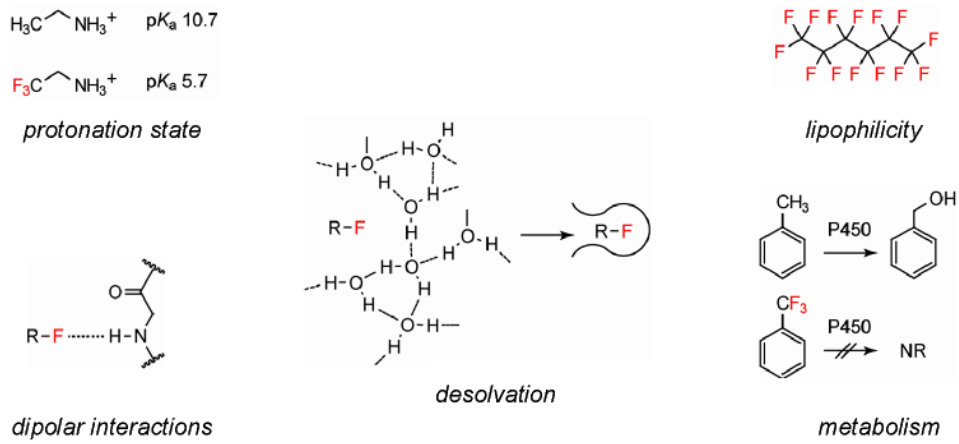


Figure 1.3. *Fluorine in drug design. The introduction of a fluorine substituent can favorably tune a number of drug properties such as protonation state, lipophilicity, protein-ligand interactions, and metabolic stability.*

high electronegativity means that removal of an electron from the fluorine atom is highly unfavorable (by more than 400 kcal/mol) while addition of an electron to produce fluoride is very favorable (by 78.3 kcal/mol) [*42*].

Because of its high electronegativity, fluorine forms the strongest single bond to carbon of any of the elements. This bond is highly polarized and its strength is related to electrostatic attraction between the $F^{\delta-}$ and $C^{\delta+}$ rather than more canonical covalent electron sharing [*42*]. The C-F bond is intermediate in length between C-H and C-O bonds [*38*], making it the most sterically conservative substitution for a C-H bond, but its polarization means that it can significantly alter the electronic properties of a molecule [*37, 38, 40*]. Based on the polarization of the C-F bond, fluorination can modulate the pK_a of adjacent acidic groups, which may have strong effects on binding affinity, bioavailability, and absorption. These effects may be potentiated through either ionization or changes in molecular conformation induced by pK_a shifts [*37, 38, 40*]. Although the C-F bond is highly polarized, it is a poor hydrogen bond acceptor because its electrons are tightly held and the fluorine atom itself is not polarizable [*42, 43*].

Despite the poor hydrogen bond accepting ability of the C-F bond, it is capable of participating in charge-dipole and dipole-dipole interactions [*38, 39, 42*]. Indeed, a survey of small molecule and protein crystal structures containing fluorinated ligands demonstrated the observation of very few C-F interactions consistent with hydrogen bonding, but many multipolar interactions with the C-F bond [*39*]. Among the most common C-F interactions observed in protein crystal structures with fluorinated ligands were C-F···C=O, C-F···H-C_α, and C-F···H-N (amide) interactions (*Figure 1.4*). Interactions between the guanidinium groups of Arg side chains and C-F bonds were also observed, with geometry compatible with dipolar interactions rather than hydrogen bonding (*Figure 1.4*). Dipolar interactions involving C-F bonds have generally been observed to make energetic contributions to binding affinity of 0.2-0.3 kcal/mol, equivalent to one-half to one-third of a neutral hydrogen bond [*39*].

Introduction of a C-F bond to an organic compound usually, but not always, increases the lipophilicity of that compound [*38, 40*]. This is based on the tendency of water molecules to order around the C-F unit, a property known as 'polar hydrophobicity' [*38*]. Increased lipophilicity is observed in fluorinated aromatic rings, per- or polyfluorinated alkyl chains, and in molecules with a fluorine substituent adjacent to an atom with a π-bond [*38, 40*]. In contrast,

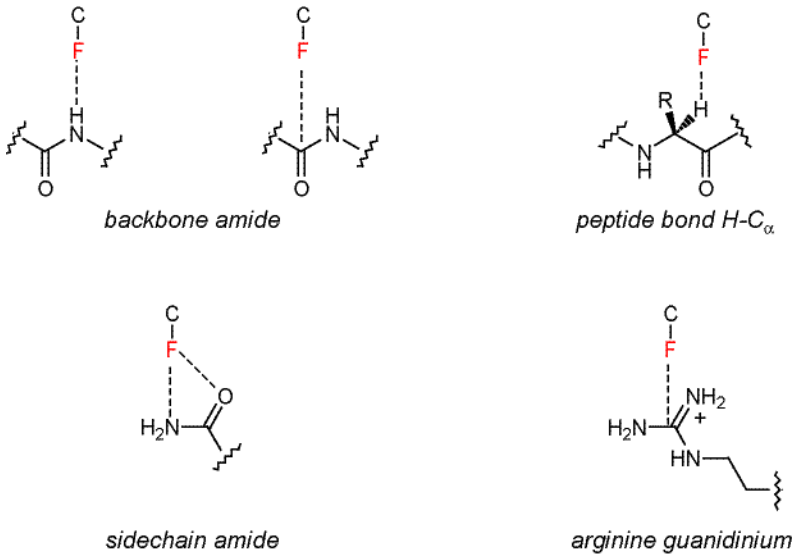


Figure 1.4. *Fluorine-protein interactions. Interactions with the peptide bond and interactions with the sidechain amide groups of asparagine and glutamine, as well as interactions with arginine guanidinium groups, are commonly observed in crystal structures with a bound fluorinated ligand in the Protein Data Bank.*

mono- or trifluorination of a saturated alkyl group has been observed to decrease lipophilicity, an effect that has been attributed to the introduction of a strong dipole to an otherwise hydrophobic group [*38, 40*]. As a consequence of the polar hydrophobicity of the C-F bond, fluorinated molecules tend to have negative entropies of aqueous solvation. Because the entropy of binding of hydrophobic molecules to proteins is typically large and positive, introduction of a fluorine modification has been exploited to increase the hydrophobic surface area of a ligand, leading to an increase in affinity for the protein receptor [*44*]. This effect was directly correlated with the total amount of hydrophobic surface area, regardless of the presence or absence of fluorine [*44*]. Fluorine-based changes in lipohilicity have also been used for tuning the hydrophobicity of small molecules such that they can pass through the lipid membrane without becoming trapped [*40*].

Fluorine effects on metabolic stability. Drug molecules can undergo a wide range of biotransformations subsequent to administration that promote inactivation and elimination from the body [*40*]. The rate at which these reactions occur impacts the efficacy of pharmacological agents as it impacts the duration and dose of the treatment [*40*]. Cytochrome P450 monooxygenases (P450s) are important players in drug metabolism in that they catalyze oxidative reactions that increase the hydrophilicity of drug molecules, facilitating their excretion [*45*]. Metabolic instability due to oxidation by P450s is a common problem in drug design that can be alleviated by fluorine substitution at or near the reactive sites [*40, 46, 47*]. Fluorination may directly block metabolism due the inertness of the C-F bond to biological oxidants, or may be used to deactivate moieties that are susceptible to oxidation, like aromatic rings, based on the electron-withdrawing nature of the C-F bond [*40, 46, 47*].

Fluorination has also proven useful in enhancing the hydrolytic stability of drug molecules. In the case of prostacyclin, a vasodilator and platelet aggregator, hydrolytic instability can be attributed to the presence of an electron-rich enol ether motif, leading to a short half-life of 10 min [*48*]. Fluorination adjacent to the enol ether withdraws electron density, extending the half-life to >1 month [*49*], while difluorination further extends the half-life to 90 days [*50*]. Similarly, fluorination of erythromycin to generate flurithromycin increases its acid stability compared to

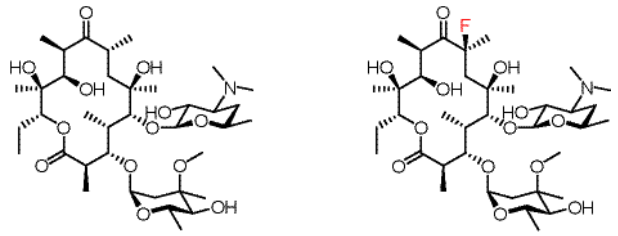


Figure 1.5. *Erythromycin and flurithromycin.*

the parent compound [*51, 52*]. While erythromycin (*Figure 1.5)* is unsuitable for treatment of the *Helicobacter pylori* infections that lead to peptic ulcers due to its inability to survive the acidic conditions of the stomach, flurithromycin has been used successfully for this purpose due to its longer half-life, higher bioavailability, and ability to accumulate to higher concentrations in tissues [*51, 52*].

Fluorine in mechanism-based inhibition. The introduction of fluorine into substrate analogs often results in unusual and unproductive reaction pathways that lead to target enzyme inactivation. In particular, the ease of fluoride elimination may result in formation of reactive species capable of covalent enzyme inactivation [*53*]. The reactivity of fluorinated compounds has therefore been exploited for the design of therapeutics that rely on mechanism-based inhibition of the target enzyme [*54*].

Several pharmaceuticals whose mode of action relies on mechanism-based inhibition are targeted toward disruption of DNA synthesis and replication. Many of these are prodrugs that are converted to deoxynucleotide mono-, di-, or triphosphates (dNMPs, dNDPS, or dNTPs) *in vivo*, allowing them to act as suicide substrates for their target enzymes. Trifluridine is an antiviral uridine analog that is converted to the dUMP analog, allowing it to act as a mechanism-based inhibitor of thymidylate synthase (*Figure 1.6*) [*55*]. Elimination of fluoride from the trifluoromethyl group of trifluridine results in formation of a Michael acceptor with which a nucleophilic enzyme side chain react, irreversibly modifying and inactivating the enzyme [*55*]. The anticancer agent 5-fluorouracil also targets thymidylate synthase for inhibition. The mechanism of thymidylate synthase involves formation of a covalent adduct between dUMP and the methylene tetrahydrofolate cofactor. With the natural substrate, the reaction is completed through abstraction of the C5 proton, followed by β-elimination of tetrahydrofolate [*56*]. 5-Fluorouracil undergoes the same reaction pathway as the natural substrate up until this point, but the unavailability of the C5 proton precludes elimination of the cofactor [*57*], leaving the enzyme irreversibly inactivated because it is covalently attached to both the substrate and the cofactor.

Fluorination may also interfere with DNA biosynthesis when it is used to replace the hydroxyl group nucleophile that is needed for chain elongation. This strategy has been used in the design of viral reverse transcriptase inhibitors like 2-fluoroadenosine that are used in the treatment of HIV and AIDS [*58*], as well as in anticancer agents that target DNA replication, like gemcitabine [*59*]. In addition to targeting DNA replication by acting as a chain terminator [*60*], gemcitabine is also a suicide substrate for ribonucleotide reductase (RNR) [*61*]. At the active site of RNR, fluoride elimination results in formation of an off-pathway radical (in the absence of reductant) or in formation of a covalent enzyme adduct (in the presence of reductant) [*62, 63*].

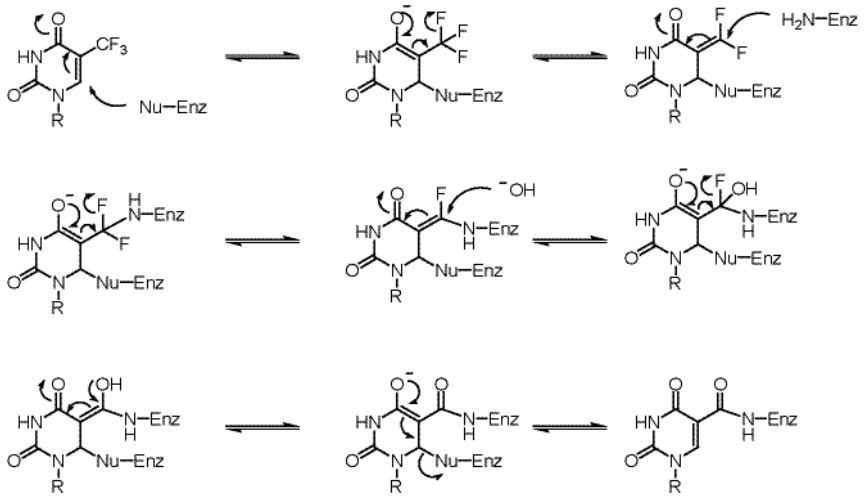


Figure 1.6. *Mechanism-based inhibition of thymidylate synthase by trifluridine. Nu, nucleophile; Enz, enzyme; H_2N, active site lysine; R, deoxyribose..*

Inhibition of ribonucleotide reductase results in a reduction of cellular dNTP levels, enhancing gemcitabine's ability to compete for incorporation into DNA and increasing its effectiveness as a chain terminator [*61*].

Conclusion. The fluorine substituent is a frequent feature of successful drug molecules, improving their pharmacokinetics, *in vivo* stability, and binding to macromolecular targets, and in some cases, facilitating the reaction pathways that are responsible for drug effectiveness. The effects are a result of fluorine's unique position as the most electronegative element in the periodic table, a property that explains its small size, lack of polarizability, and 'polar hydrophobicity'. Given the success of fluorine in pharmaceuticals, it is likely that the number of fluorinated drugs will continue to grow along with our understanding of how fluorine affects the chemistry and biochemistry of small molecules.

1.3 Fluorinated secondary metabolites

Introduction. In contrast to the near-ubiquity of fluorine in manmade drugs, fluorinated natural products are rare and mainly derive from a single pathway for production of fluoroacetate, an antimetabolite poison that targets the tricarboxylic acid (TCA) cycle (*Figure 1.7*) [*5*]. Fluorinated natural products have been observed mainly in the plant kingdom, but also in bacteria, and potentially in sponges [*5*]. In contrast to the complexity observed in other organohalogen natural products, the fluorometabolites known to date are very simple in terms of their chemical structures. However, they could potentially act as building blocks for more complex, as yet undiscovered (or engineered) secondary metabolites such as polyketides, terpenes, isoprenoids, and polypeptides of non-ribosomal or ribosomal origin.

Fluorinated secondary metabolites in plants. In 1943, the discovery of fluoroacetate as the toxic component of the poisonous African shrub *Dichapetalum cymosum* represented the first identification of a fluorinated compound of biogenic origin [*5*]. *D. cymosum* accumulates up to 2500 $\mu g\ g^{-1}$ dry weight fluoroacetate in its leaves, presenting a danger to grazing animals, particularly in early spring [*5*]. Subsequent to its initial discovery, fluoroacetate has been positively identified in a number of other Dichapetalaceae species [*64-66*], and toxicological

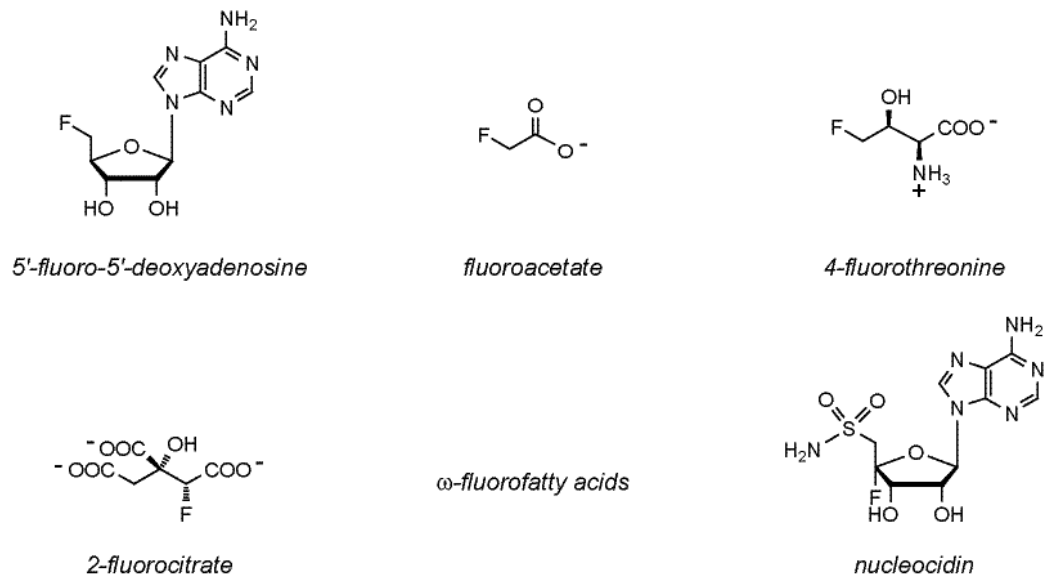


Figure 1.7. *Fluorinated secondary metabolites.*

evidence exists for its production in many others, primarily in Africa [*5*]. In Australia, several plants of the *Acacia* [*67*], *Gastrolobium* [*68*], and *Oxylobium* [*69*] genera have been demonstrated to accumulate fluoroacetate in their leaves. Although fluoroacetate-producing plants are less common outside Africa and Australia, *Palicourea margravii* [*70*] in Brazil and *Cyamopsis tetragonolobus* in India [*71*] have also been shown to biosynthesize fluoroacetate. Indeed, fluoroacetate has been identified in more than 40 species of plants that grow in tropical or subtropical regions on six different continents [*5*].

Although *D. toxicarum* is known to accumulate fluoroacetate in its leaves, its seeds are known as the most toxic part of the plant. The seeds contain high levels of organically bound fluorine, but almost no fluoroacetate. In 1959, ω-fluorooleic acid was identified as the toxic principle of *D. toxicarum* seeds and was shown to comprise up to 3% of the seed lipids [*72*]. Subsequent analysis of the lipids revealed the presence of a wide array of other ω-fluorofatty acids, albeit in smaller amounts, that appeared to be metabolites of ω-fluorooleic acid [*72*]. Because of the location of the fluorine substituent on the terminal (ω) carbon of all fluorinated lipids identified to date, biosynthesis of ω-fluorofatty acids is believed to involve utilization of fluoroacetate-derived fluoroacetyl-CoA as a starter unit by fatty acid synthases (FAS), an idea that seems plausible based on the known starter unit promiscuity of FAS [*5*].

Many efforts have been made to develop model systems for studying plant utilization of fluoride and fluorometabolite biosynthesis [*73, 74*]. Although both callus cultures and tissue cultures of fluoroacetate-producing plants have been produced, even the best systems produce an order of magnitude less fluoroacetate than intact plants [*73, 74*]. Thus, despite strong interest in plant fluorometabolite biosynthesis, little progress has been made in identifying the genes and proteins involved in fluoroacetate production and accumulation in plants.

Fluorinated secondary metabolites in sponges. In 2003, 5-fluorouracil and a series of derivatives were purified from the marine sponge *Phakellia fusca* from the South China Sea (*Figure 1.8*) [*75*]. Among the five compounds isolated and structurally characterized, two are known chemotherapeutic agents, while the others are novel 5-fluorouracil derivatives. While

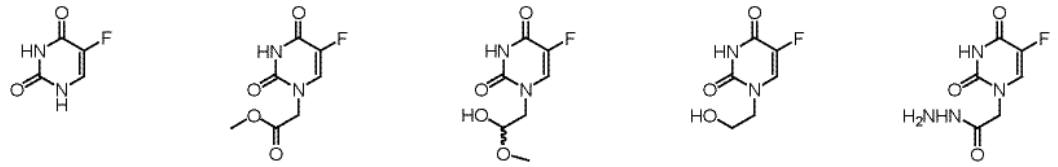


Figure 1.8. *5-Fluorouracil derivatives isolated from Phakellia fusca.*

these molecules could represent an exciting new class of fluorinated secondary metabolites, their relationship to clinically used pharmaceuticals raises the possibility that their presence in sponges is the result of industrial contamination. Additional experiments are therefore needed to rule out the possibility that these compounds are anthropogenic.

Fluorinated secondary metabolites in actinomycetes. Although fluoroacetate biosynthesis appears to be widespread in the plant kingdom, only two species of bacteria have been identified to date that can produce organofluorine molecules, both actinomycetes. The first, *Streptomyces calvus*, produces the broad-range antibiotic and anti-trypanosomal nucleocidin (*Figure 1.7*) [*76, 77*]. The second, *Streptomyces cattleya*, produces fluoroacetate as well as the fluorinated amino acid 4-fluorothreonine (*Figure 1.7*) [*78*].

Although nucleocidin was isolated from the fermentation broth of *Streptomyces calvus* in 1956, it was not until 1969 that its structure was discovered to contain a fluorine atom as part of a unique 4'-α-fluoro-*O*-sulfamoyl adenosine nucleoside (*Figure 1.7*) [*76*]. Although the mechanism of action of nucleocidin has been studied extensively, efforts to elucidate its biosynthesis have been hampered by a lack of reproducibility in the ability of *S. calvus* to produce this molecule under laboratory conditions. However, the discovery that specific mutations in the RNA polymerase β-subunit (*rpoB*) in certain streptomycetes can activate the expression of secondary metabolite biosynthesis genes has led to the generation of an *S. calvus* strain with improved nucleocidin production [*79*]. These mutations are thought to function in upregulating secondary metabolite biosynthesis by mimicking the guanosine tetraphosphate-(ppGpp) bound state of RpoB, which is associated with the onset of secondary metabolite biosynthesis. Future studies with the strain may therefore lead to elucidation of the nucleocidin biosynthesis pathway.

Streptomyces cattleya was discovered to produce both fluoroacetate and 4-fluorothreonine serendipitously during studies to optimize its production of the carbenapem antibiotic thienamycin (*Figure 1.7*) [*78*]. Three thienamycin-producing strains (NRRL 8057, MA 5176, and MA 5617) were tested for the ability to biosynthesize fluorometabolites, but only two of the three (NRRL 8057 and MA 5176) were observed to have this biosynthetic capacity. One of the strains (MA 5176) was further probed to examine which fluorine-containing compounds could serve as organofluorine precursors. Inorganic fluoride and 4-fluoroglutamate were shown to lead to production of both fluoroacetate and 4-fluorothreonine, while fluoroacetate and 4-fluorothreonine could be interconverted. Since the discovery of its ability to produce fluorometabolites, organofluorine biosynthesis in *S. cattleya* has been studied extensively. These studies are discussed in more detail in section 1.4.

Conclusions. Although fluorine is the most abundant halogen in the Earth's crust, fluorinated secondary metabolites remain rare among halogenated natural products and lack the complexity observed in other classes of natural products. Fluoroacetate, whose production has

been noted mainly in the plant kingdom, is the most widespread biogenic organohalogen. While plant model systems have proven intractable for understanding fluorometabolite biosynthesis, the discovery of fluoroacetate production in *Streptomyces cattleya* has opened the door to molecular genetic and mechanistic studies that could advance our understanding of this exotic chemical phenotype.

1.4 Organofluorine biosynthesis in *Streptomyces cattleya*

Introduction. The discovery that *Streptomyces cattleya* possesses the ability to biosynthesize fluorometabolites has provided a tractable system for production of these compounds under laboratory conditions [*78*]. The ability to monitor and control the production of fluoroacetate and 4-fluorothreonine, along with the convenient ^{19}F NMR handle of the metabolites involved has allowed for chemical, biochemical, and genetic studies of the biosynthetic origin of fluorometabolites, including their chemical precursors, the enzymatic activities responsible for converting them, and the genes that encode these activities (*Figure 1.9*). More recently, advances in DNA sequencing technology have allowed for genomic approaches to elucidating fluorometabolite biosythesis. A molecular understanding of the enzymatic processes that involved in fluorometabolite biosynthesis will provide insights into the fundamental question of how specialized metabolic functions evolve from pre-existing pathways as well as practical information about enzymatic fluorine specificity that can be leveraged for pathway and protein engineering.

Fluoroacetaldehyde as a common fluorometabolite precursor. Initial studies toward understanding the biosynthetic origins of fluoroacetate and 4-fluorothreonine focused on defining the relationship between the two metabolites. Although early studies showed that fluoroacetate could act as a precursor to 4-fluorothreonine [*78*], later experiments in which [$^2H_2$]-fluoroacetate was fed to *S. cattleya* cell suspensions demonstrated that fluoroacetate is not

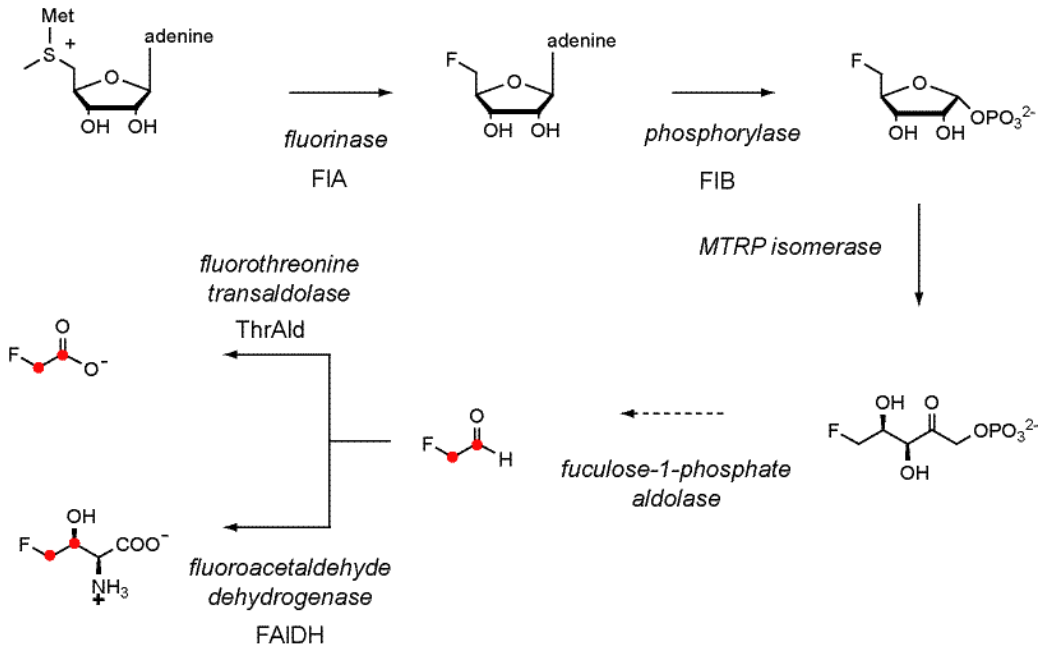


Figure 1.9. *Fluorometabolite biosynthesis in Streptomyces cattleya. The two-carbon unit shared among fluoroacetaldehyde, fluoroacetate, and 4-fluorothreonine is marked with red dots.*

directly converted to 4-fluorothreonine [*80*]. The observation of 4-fluorothreonine biosynthesis in the presence of fluoroacetate was attributed to defluorination of fluoroacetate, followed by *de novo* synthesis of 4-fluorothreonine, consistent with the absence of ^{2}H incorporation [*80*]. Fluoroacetate and 4-fluorothreonine thus do not have a precursor-product relationship.

To further probe the origins of fluoroacetate and 4-fluorothreonine, a variety of ^{13}C- and ^{2}H-labeled precursors were fed to *S. cattleya* cell suspensions, allow their fates to be traced using NMR and GC-MS. The extent of heavy isotope incorporation into fluoroacetate and 4-fluorothreonine was inconsistent with direct conversion of any of the precursor molecules to fluorometabolites. However, it was noted that the labeling patterns of C-1 and C-2 of fluoroacetate were identical to those of C-3 and C-4 of 4-fluorothreonine (*Figure 1.9*) [*81*], suggesting that both fluorometabolites arise from a common C_2 unit. This shared intermediate was proposed to be fluoroacetaldehyde (*Figure 1.9*) [*80, 81*], which could be converted to fluoroacetate in one oxidation step and to 4-fluorothreonine through an aldol condensation. Adding further fuel to this hypothesis, enzymatic activities were known that could effect these transformations on acetaldehyde, the nonfluorinated congener of fluoroacetaldehyde. Feeding studies with fluoroacetaldehyde demonstrated its competence as a precursor to both fluoroacetate and 4-fluorothreonine [*82*]. Additional studies showed that when cells were fed [1-$^{2}H_1$]-fluoroacetaldehyde, the ^{2}H labeled was retained in 4-fluorothreonine, ruling out oxidation of C-1 during conversion of the fluoroacetaldehyde precursor to the 4-fluorothreonine product [*82*].

With fluoroacetaldehyde firmly established as the common fluorometabolite precursor (*Figure 1.9*), efforts were made to isolate the enzymatic activities responsible for converting this intermediate to fluoroacetate and 4-fluorothreonine. A fluoroacetaldehyde dehydrogenase activity was isolated from *S. cattleya* lysate and shown to be supported by NAD^+ but not $NADP^+$ [*83*]. Incubation of this activity with other aldehydes showed that they could be oxidized with similar efficiencies, indicating that the enzyme is relatively promiscuous.

4-Fluorothreonine was initially postulated to be formed from fluoroacetaldehyde in a reaction analogous to that catalyzed by threonine aldolase, in which acetaldehyde and glycine undergo an aldol condensation to form threonine. However, isotopic labeling studies with ^{13}C-glycine were inconsistent with this molecule being directly incorporated into 4-fluorothreonine [*81*]. An unusual pyridoxal-5'-phosphate (PLP)-dependent activity was later purified from *S. cattleya* cell lysate [*84*] that catalyzes a transaldol reaction to convert fluoroacetaldehye and threonine to acetaldehyde and 4-fluorothreonine, using threonine rather than glycine as the condensing nucleophile.

Discovery and characterization of the fluorinase. Although isotopic labeling studies provided a wealth of information about the immediate precursor to fluoroacetate and 4-fluorothreonine, they failed to reveal the initial substrate for fluorination. Subsequent studies in which cell-free extracts were prepared under different conditions and supplied with inorganic fluoride and a variety of cofactors revealed that adenosine-5'-triphosphate (ATP) supports organofluorine biosynthesis [*30*]. Addition of SAM to the extracts also led to a similar distribution of organofluorines. Given the close relationship between SAM and ATP, which together with a molecule of L-methionine can be converted to SAM by SAM synthetase, 5'-fluoro-5'-deoxyadenosine (5'-FDA) was proposed as the initial fluorination product [*30*]. This was confirmed by comparison of the organofluorines produced by the protein extract to an authentic 5'-FDA standard. The 5'-FDA could be converted to fluoroacetate, and

supplementation of the extracts with exogenous 5'-FDA resulted in fluorometabolite production, further supporting its role as the initial product of fluorination.

The ability to track 5'-FDA synthase, or fluorinase, activity, allowed for purification of this enzyme from *S. cattleya* lysate [*9*]. Native purification demonstrated that the active enzyme is hexameric with a monomer molecular weight of 32 kDa and that it is competitively inhibited by SAH. Using N-terminal sequencing of the native protein and degenerate PCR primers designed based on this information, the fluorinase was cloned and recombinantly overepxressed, facilitating further structural and mechanistic studies [*6*].

The crystal structure of the fluorinase revealed that each monomer is consists of two domains that are connected by a 15-residue linker (*Figure 1.10*) [*6*]. The N-terminal domain consists of a seven-stranded β-sheet that is flanked by α-helices. The smaller C-terminal domain consists of two antiparallel β-sheets, one five-stranded, and one four-stranded. Electron density for bound

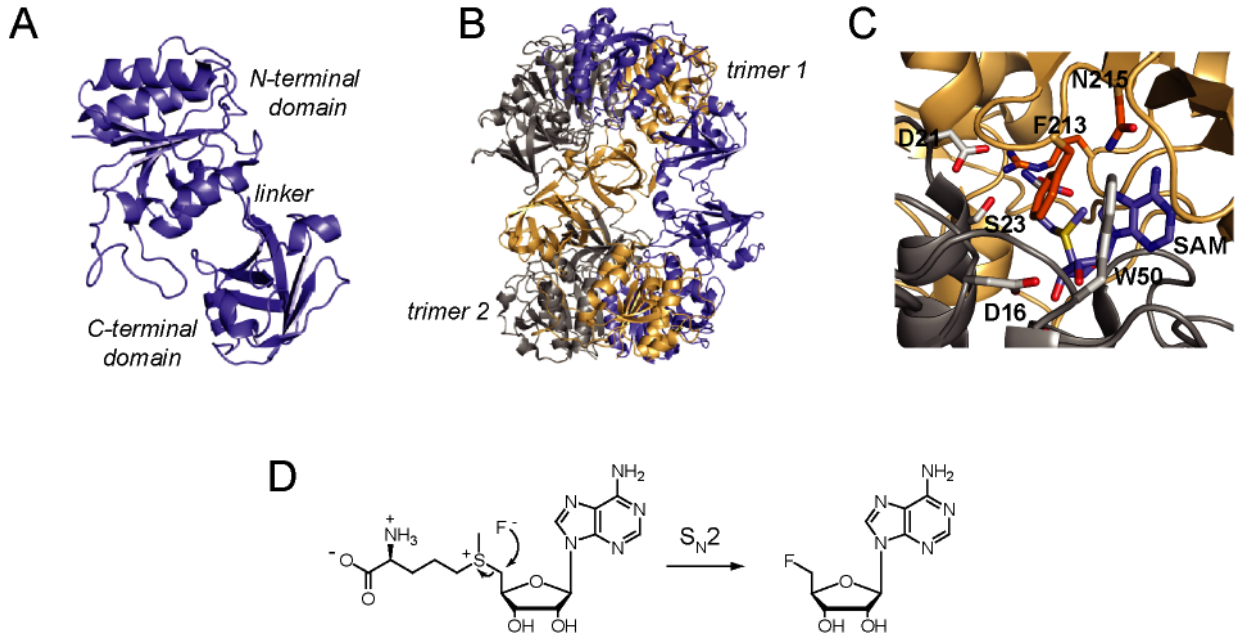


Figure 1.10. *Structure and mechanism of the fluorinase. (A) The fluorinase monomer. (B) The fluorinase hexamer, comprised of a dimer of trimers. (C) The active site of the fluorinase with bound SAM.*

SAM was observed between the C-terminal domain of one protomer and the N-terminal domain of a second protomer, indicating that this subtrate binds tightly and survived purification. The adenosine moiety, ribose ring, and methionine of SAM all make extensive contacts that span both protomers, consistent with the hexameric quaternary structure observed in the native protein. The ribose ring appears to bind in an unusual high-energy conformation, potentially promoting the reaction between fluoride and SAM. Fluoride binds in a pocket with a radius of 1.4-1.6 Å, potentially explaining selectivity for fluoride over larger anions. Notably, the area surrounding this pocket is completely devoid of ordered water molecules, and the fluoride ion appears to be fully desolvated. Based on these structural insights, the fluorinase appears to effect catalysis through a combination of organizing the ribose ring in a high-energy conformation and desolvating the fluoride ion to enhance its nucleophilicity.

Further biochemical studies of the fluorinase have defined its substrate specificity and aspects of its catalytic mechanism (*Figure 1.10*). The fluorinase is also capable of activating chloride as a nucleophile, but the 5'-chloro-5'-deoxyadenosine (5'-ClDA) product can only be observed then the reaction is monitored in a coupled-enzyme assay that shifts the equilibrium toward 5'-ClDA [*85*]. Despite the strong nature of the C-F bond, the fluorinase can also operate in the reverse direction at observable rates when provided with 5'-FDA and L-methionine [*85*]. Stereochemical analysis of the 5'-FDA products generated from istopically labeled precursors indicates that the fluorination reaction proceeds with inversion of configuration [*31*], consistent with an S_N2-type reaction mechanism. These advances have allowed for the development of a number of biotechnological applications of the fluorinase, including its use for synthesis of ^{18}F probes for positron emission tomography (PET) imaging [*86*] and in the production of fluorinated natural product analogs [*87*].

Intermediates linking 5'-FDA and fluoroacetaldehyde. Further efforts to identify the metabolic steps that link 5'-FDA and fluoroacetaldehyde have focused on what transformations might occur on 5'-FDA (*Figure 1.9*). The possible intermediacy of 5-fluoro-5-deoxyribose, the hydrolysis product of 5'-FDA, was explored, but incubation of this molecule with *S. cattleya* extract indicated that it is not biotranformed. An alternative route for further metabolism of 5'-FDA would be phosphorolysis of the nucleoside to generate 5-fluoro-5-deoxyribose-1-phosphate (5-FDRP) and adenine. Indeed, incubation of this proposed intermediate with *S. cattleya* extract led to accumulation of both fluoroacetate and 4-fluorothreonine [*88*]. A purine nucleoside phosphorylase activity capable of converting 5'-FDA to 5-FDRP was subsequently purified from *S. cattleya* extract [*88*].

Based on similarities between the first two steps of fluorometabolite pathway and the methionine salvage pathway [*89*], an analogy was drawn that suggested that the next step toward fluoroacetaldehyde might be isomerization to form 5-fluoro-5-deoxy-D-ribulose-1-phosphate (5-FDRu1P) as the next intermediate in the pathway (*Figure 1.9*) [*90*]. Feeding studies using 5-FDRP in the presence of EDTA (ethylenediaminetetraacetic acid), which is expected to inhibit aldolase enzymes by sequestering the required Zn^{2+} ion, revealed the accumulation of a new product, which was assigned as 5-FDRu1P based on comparison to an authentic standard [*90*]. An isomerase activity capable of catalyzing formation 5-FDRu1P was subsequently purified from *S. cattleya* [*91*].

Continuing with the methionine salvage pathway analogy [*89*], the next enzyme involved in fluorometabolite biosynthesis is proposed to be a fuculose-1-phosphate aldolase that would convert 5-FDRu1P to dihydroxyacetone phophate and fluoroacetaldehyde. While such an enzyme activity has not been purified or identified genetically in *S. cattleya*, a fuculose-1-phosphate aldolase has been isolated from *Streptomyces coelicolor* that can catalyze this tranformation and support biosynthesis of 4-fluorothreonine from fluoride *in vitro* [*91*].

Discovery of a fluorometabolite biosynthesis gene cluster. The availability of the fluorinase gene sequence made it possible to target the *flA* genomic locus for sequencing and led

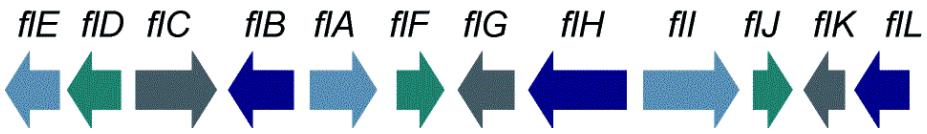


***Figure 1.11**. Organization of the fluorometabolite biosynthesis gene cluster.*

to the identification of a gene cluster involved in fluorometabolite biosynthesis, termed the *fl* locus (*Figure 1.11*) [*92*]. While the fluorinase was found to be adjacent to a gene encoding a purine nucleoside phosphorylase capable of converting 5'-FDA to 5-FDRP, none of the other genes in the cluster have obvious biosynthetic functions (*Table 1.1*). The remaining genes can be divided into classes based on their putative functions in resistance (*flD* and *flK*), transmembrane transport (*flC* and *flH*), and transcriptional (*flE*, *flF*, *flG*, and *flL*) and other types (*flI* and *flJ*) of regulation [*92*]. Of these genes, only *flK* and *flI* have been expressed and characterized *in vitro*. FlK has fluoroacetyl-CoA thioesterase activity [*92-94*], while FlI is an SAH hydrolase proposed to function in the relief of fluorinase inhibition by endogenous SAH [*92*].

Table 1.1. *Proposed functions of genes encoded in the fluorometabolite biosynthesis cluster of S. cattleya.*

gene	putative function/homology
flA	fluorinase
flB	5'-fluoro-5'-deoxyadenosine phosphorylase
flC	major facilitator superfamily permease
flD	dehalogenase/phosphatase
flE	DNA binding regulatory protein
flF	DNA binding regulatory protein
flG	DNA binding regulatory protein
flH	Na^+/H^+ antiporter
flI	*S*-adenosylhomocysteine hydrolase
flJ	regulatory protein
flK	hotdog-fold thioesterase
flL	DNA binding regulatory protein

Conclusion. The fluorinase from *Streptomyces cattleya* represents the only characterized naturally occurring enzyme for carbon-fluorine bond formation known to date. A pathway for conversion of inorganic fluoride to fluoroacetate and 4-fluorothreonine has been proposed, and several of the intermediates and enzymes involved have been identified. A clear analogy can be drawn between the initial steps of the fluorometabolite biosynthesis pathway and the methionine salvage pathway, providing insight into how enzymes with exotic functions may evolve from their well-studied counterparts. Despite the great progress that has been made in elucidating organofluorine biosynthesis, more work remains to be in cloning the remaining enzymes involved to complete the genetic basis for the pathway. This would provide new model systems for understanding molecular recognition of fluorine in an evolved system and could potentially allow the development of a fluoroacetate synthesis module for synthetic biology.

1.5 Mechanisms of fluoroacetate toxicity and resistance

Introduction. Despite its simple structure, fluoroacetate is a potent poison, with toxicity exceeding that of much more complex natural products that have been developed as chemotherapeutic agents, such as taxol [*95, 96*]. The high toxicity of fluoroacetate creates a difficult problem for the biosynthetic host, which must evade fluoroacetate toxicity while producing this molecule at high levels. Additionally, because fluoroacetate is widespread in the environment in certain parts of the world [*5*], selective pressure exists for organisms that come into contact with this toxin to evolve their own mechanisms of resistance. Unsurprisingly, despite fluoroacetate's universal mechanism of action, different modes of resistance have evolved in producers of fluoroacetate *versus* targets of fluoroacetate poisoning.

Lethal synthesis of fluorocitrate. Fluoroacetate's high toxicity arises from an antimetabolite mode of action in which the fluorine substitution cannot be discriminated by enzymes that normally metabolize acetate. In particular, fluoroacetate can be activated to fluoroacetyl-CoA by either acetyl-CoA synthetase (ACS) or the acetate kinase/phosphotransacetylase (AckA/PTA) system (*Figure 1.12*). Fluoroacetyl-CoA can then be

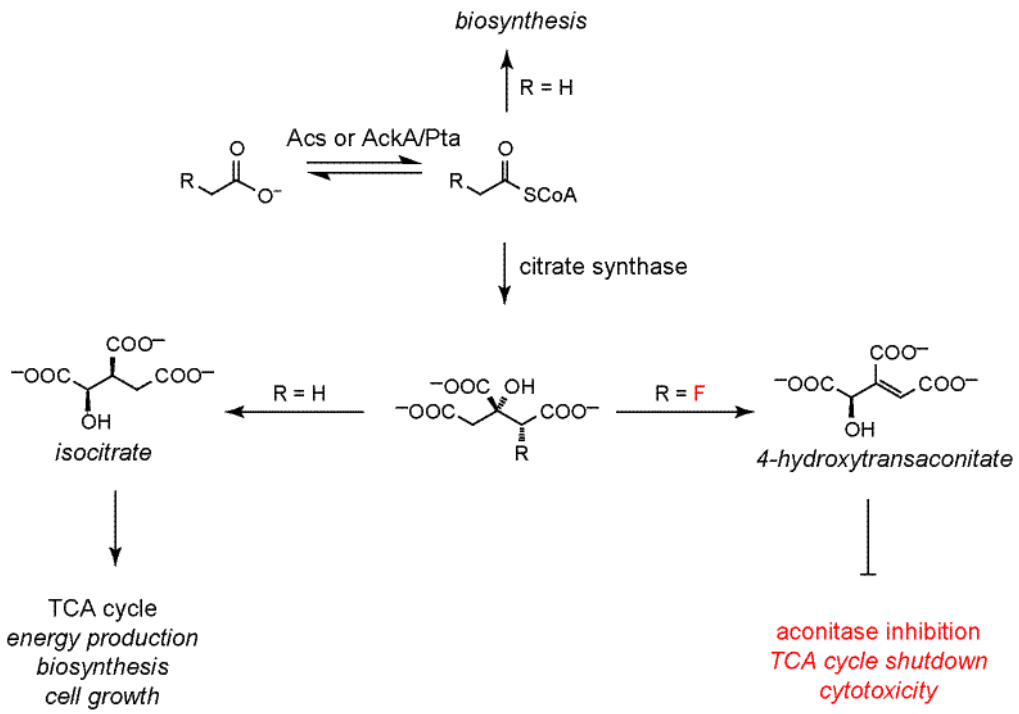


Figure 1.12. *Lethal synthesis of fluorocitrate. Fluoroacetate is utilized as a substrate by the acetate assimilation machinery, leading to formation of a potent noncovalent inhibitor of aconitase.*

utilized by citrate synthase in a condensation reaction with oxaloacetate to form 2-fluorocitrate [*97-99*]. Detailed studies of the stereochemical course of this reaction indicate that only the 2-*pro-S* hydrogen is abstracted from fluoroacetyl-CoA and that the condensation reaction proceeds with inversion of configuration at C-2 to general (2*R*, 3*R*)-2-fluorocitrate exclusively [*100*]. Early studies on the mechanism of fluorocitrate toxicity demonstrated that at least one consequence of fluorocitrate formation is aconitase inhibition (*Figure 1.12*) [*101*]. Based on the known catalytic mechanism of aconitase and the observation that inhibition was associated with fluoride release, it was proposed that the inhibitory molecule is 4-hydroxy-*trans*-aconitate, a hypothesis that was later confirmed through x-ray crystallographic studies of the aconitase-fluorocitrate complex [*102*].

Later investigations of fluoroacetate toxicity that examined the cellular fate of radiolabeled 2-fluorocitrate showed that the radiolabel becomes covalently associated with a citrate carrier protein [*103, 104*]. This covalent modification was proposed to cause irreversible inhibition of citrate transport into mitochondria independent of aconitase inhibition. The hypothesis that a defect in citrate transport is the proximal cause of 2-fluorocitrate's toxicity is attractive because it explains the irreversible nature of the toxicity as well as the observation of toxic effects at lower doses of 2-fluorocitrate (10 nM) than would be required to fully inhibit aconitase (60-200 μM) [*103, 104*]. While inhibition of citrate transport may be the major component of 2-fluorocitrate toxicity in eukaryotes, it is unlikely to explain toxicity in bacteria, which lack subcellular compartments.

Fluoroacetate resistance in animals. In areas of Western Australia where fluoroacetate-producing plants are endemic, it has been noted that a variety of indigenous animals appear to have evolved fluoroacetate resistance [*105*]. Particularly striking is the fact that in several animal

species including bush rats, brush-tailed possum, and grey kangaroo, populations in Western Australia are markedly more tolerant of fluoroacetate than populations in Eastern Australia, outside the range of fluoroacetate-producing plants. The levels of fluoroacetate tolerance observed also appear to be evolutionarily optimized based on the lifestyle of the particular animal, with herbivores exhibiting higher tolerance than omnivores, which exhibit higher tolerance than carnivores. Along these lines, seed-eating birds are more resistant to fluoroacetate than other animals that eat only leaves. Although patterns of fluoroacetate resistance have provided interesting insights into the evolution and speciation of Australian animals, the biochemical basis for this trait has yet to be elucidated.

Defluorination of fluoroacetate as a resistance mechanism. Fluoroacetate resistance is widespread among microbes, including plant pathogens and ruminal bacteria, as well as bacterial species that are more ubiquitous [*105, 106*]. In many cases, the genes and enzymes that confer

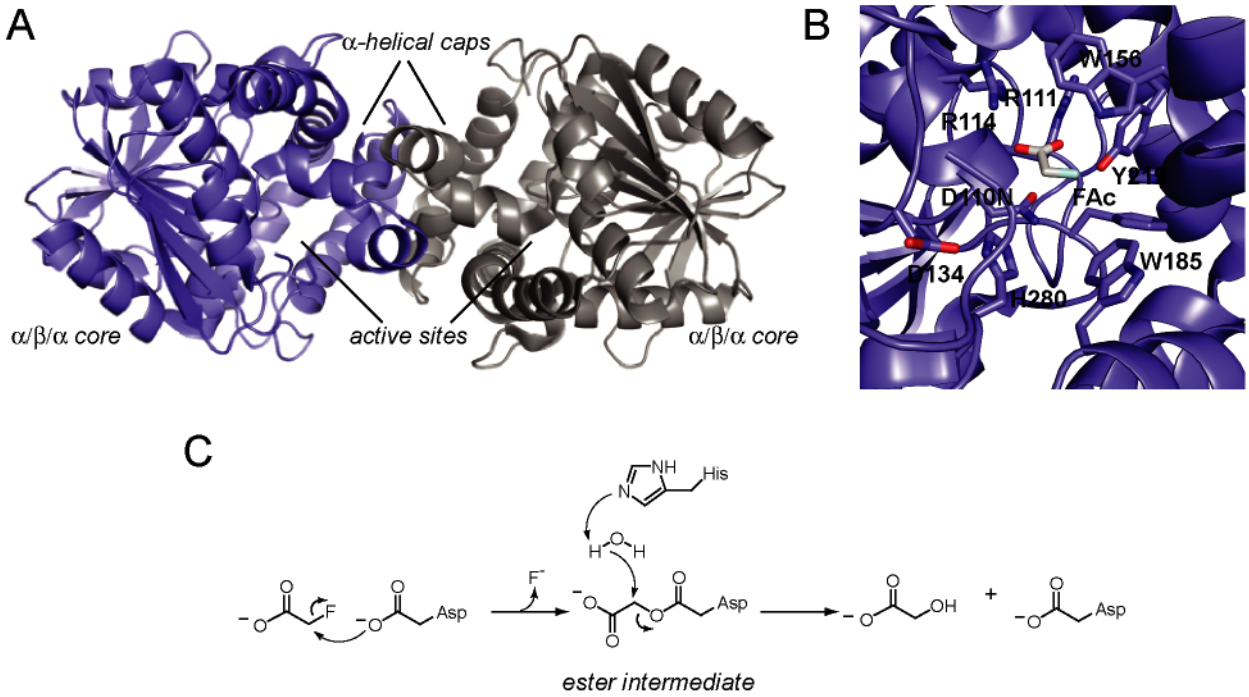


Figure 1.13. *Fluoroacetate dehalogenase structure and mechanism. (A) Homodimeric structure of fluoroacetate dehalogenase. (B) Active site of fluoroacetate dehalogenase. (C) Proposed mechanism for fluoroacetate dehalogenase.*

resistance are known [*106-108*]. The most common method of fluoroacetate detoxification involves hydrolytic cleavage of the C-F bond by a fluoroacetate dehalogenase enzyme (*Figure 1.13*) [*106*]. Because these enzymes are capable of cleaving the strongest bond to carbon with high specificity over carbon-chlorine or carbon-bromine bonds, interest in the structure and mechanism of fluoroacetate dehalogenases has been strong. As a result, a number of fluoroacetate dehalogenases are structurally and mechanistically characterized.

The canonical fluoroacetate dehalogenases belong to the α/β-hydrolase superfamily. Structural analysis indicates that fluoroacetate dehalogenases are homodimeric proteins, with each monomer comprised of an α/β/α core linked to an α-helical cap (*Figure 1.13*) [*109, 110*]. The conserved catalytic triad (Asp 110, His 280, Asp 134, *Rhodopseudomonas palustris* numbering) is almost completely buried at the dimer interface, accessible only by a narrow

channel [*110*]. The remainder of the active site environment is comprised of Phe 40, Arg 111, Arg 114, His 155, Trp 156, Trp185, and Tyr 219, which are conserved in the structurally characterized *Burkholderia* enzyme [*109, 110*].

The catalytic mechanism of fluoroacetate dehalogenases involves a double-displacement that forms glycolate and fluoride as the products (*Figure 1.13*) [*107*]. Asp 110 serves as the catalytic nucleophile, attacking the α-carbon of the substrate in an S_N2 fashion to displace fluoride with inversion of configuration. The resultant enzyme-bound ester intermediate is hydrolyzed by a water molecule that is activated by His 280, completing the catalytic cycle [*107*].

The basis of substrate specificity in fluoroacetate dehalogenase has been studied extensively with x-ray crystallography [*110*]. Comparison of the chloroacetate- and bromoacetate-bound structures of the enzyme from *Rhodopseudomonas palustris* shows that these larger substrates cannot be accommodate by the small fluoroacetate pocket. They bind further away from the catalytic triad with suboptimal geometry for catalysis to occur. It has been proposed that fluoroacetate dehalogenases provide additional activation of the C-F bond through three precisely placed hydrogen bonding residues (His, Trp, and Tyr) that interact with fluorine. These contacts are not observed in haloacid dehalogenases that lack fluoroacetate dehalogenase activity.

Although all of the structurally and mechanistically characterized fluoroacetate dehalogenases belong to the α/β hydrolase superfamily, a recent sequence- and activity-based screening effort identified four members of the haloacid dehalogenase (HAD) superfamily that hydrolyze the C-F bond of fluoroacetate [*111*]. Although the HAD superfamily harbors the L-2-haloacid dehalogenase family [*106*], which utilizes a similar mechanism to that of the fluoroacetate dehalogenases to dehalogenate 2-chloro-, 2-bromo-, and 2-iodo-carboxylic acids, fluoroacetate dehalogenase activity has never been previously observed in an HAD enzyme. It will be interesting to see whether these evolutionarily distant enzymes have arrived at the same solutions for C-F bond activation and specificity.

Fluoroacetate dehalogenase activity has been developed for several biotechnological applications. Because poisoning of livestock is problematic in areas where fluoroacetate-producing plants grow, a strain of ruminal bacteria transformed with a fluoroacetate dehalogenase gene has been engineered [*112*]. Inoculation of sheep with this strain markedly increases their resistance to fluoroacetate [*112*]. Interestingly, fluoroacetate-dehalogenating bacteria have been isolated from the rumen of Brazilian goats [*113*], suggesting that nature may have arrived at the same solution. Fluoroacetate dehalogenase activity has also been developed as selection marker for both bacteria [*114*] and yeast [*115*].

Hydrolysis of fluoroacetyl-CoA as a resistance mechanism. Like many poisonous secondary metabolites, fluoroacetate presents a self-toxicity problem for its biosynthetic host [*116*]. Most organisms that produce toxic secondary metabolites harbor resistance mechanisms that allow them to evade the toxicity of their natural products while producing them at high levels [*116*]. While fluoroacetate defluorination represents an attractive resistance mechanism for nonproducing organisms [*106*], it would be problematic for fluoroacetate producers as it would

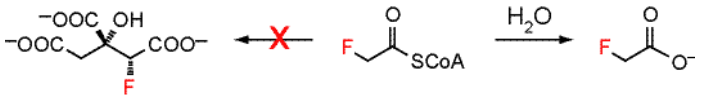


Figure 1.14. *Fluoroacetyl-CoA hydrolysis as a fluoroacetate resistance mechanism.*

destroy the metabolite and eliminate its toxicity to other organisms. Accordingly, only low levels of fluoroacetate defluorination activity have been detected in fluoroacetate producers, and it remains unclear whether these activities are specific or adventitious [*5*].

Because fluoroacetate requires further metabolic processing to effect its toxicity, one nondestructive resistance mechanism would be to block lethal synthesis of 2-fluorocitrate or its downstream effects. This could be achieved by expression of alternative acetate assimilation or TCA cycle enzymes that could discriminate between fluoroacetyl and acetyl groups, or by the existence of a fluoroacetate-insensitive aconitase. To date, none of these resistance mechanisms have been reported in fluoroacetate producers.

The existence of a fluoroacetyl-CoA hydrolase (*Figure 1.14*) activity to short-circuit the lethal synthesis of 2-fluorocitrate represents another possible resistance mechanism. This is a potentially attractive solution to the self-toxicity problem because it leaves the carbon-fluorine bond intact. To probe the hypothesis that fluoroacetyl-CoA hydrolysis could underlie fluoroacetate resistance in plants, crude mitrochondrial extracts from *Dichapetalum cymosum* were tested for this activity [*117*]. Strikingly, these extracts catalyzed coenzyme A release from fluoroacetyl-CoA but not acetyl-CoA, suggesting that the enzyme responsible possesses the specificity to eliminate fluoroacetyl-CoA while maintaining the high levels of acetyl-CoA required for cell growth. The fluoroacetyl-CoA hydrolase activity from *D. cymosum* has not been purified nor has the gene encoding this activity been cloned.

In light of the observation of fluoroacetyl-CoA hydrolase activity in *D. cymosum*, the observation of the *flK* gene encoding a hotdog-fold superfamily thioesterase in the fluorinase gene cluster of *Streptomyces cattleya* was particularly interesting [*92*]. This gene was cloned, overexpressed, and purified from *E. coli* and was shown to catalyze fluoroacetyl-CoA hydrolysis with high specificity over acetyl-CoA [*92*]. Based on this *in vitro* activity, FlK was proposed to play a role in fluoroacetate resistance in *S. cattleya*.

Based on sequence homology, FlK can be classified as a member of the hotdog-fold superfamily, which includes both thioesterases and hydratases/dehydratases (*Figure 1.15*) [*118*]. A core three-dimensional structure is common to both the thioesterases and the hydratases/dehydratases. This is comprised of a central α-helix 'hot dog' surrounded by an extended β-sheet 'bun'. The minimal functional unit of hotdog-fold enzymes is a dimer of identical subunits, although many superfamily members form higher-order structures such as tetramers and hexamers [*119-123*]. In both the thioesterases and the hydratases/dehydratases, the active site is located at the dimer interface and key residues are found on both interacting protomers [*119-124*].

The hot-dog fold thioesterases are divided into two clades based on the positioning of the catalytic Asp/Glu residue either on the hotdog helix (*Arthrobacter* clade) [*125*] or on a loop above the active site (*Pseudomonas* clade) [*126*]. Despite the differences in active site topology, extensive structural and biochemical evidence supports a common mechanism for these two groups of enzymes involving the formation of an enzyme-anhydride intermediate on the catalytic Asp/Glu residue concomitant with formation of free CoA (*Figure 1.15*) [*127, 128*]. This acyl-enzyme intermediate is then cleaved in a rate-limiting hydrolytic step to give the carboxylic acid product and regenerate the native enzyme. Hotdog-fold thioesterases are typically promiscuous and many accept a variety of medium-, long-chain, and aromatic acyl-CoAs with similar catatlytic efficiencies [*119, 129-131*]. Other hotdog-fold thioesterase substrates include fatty

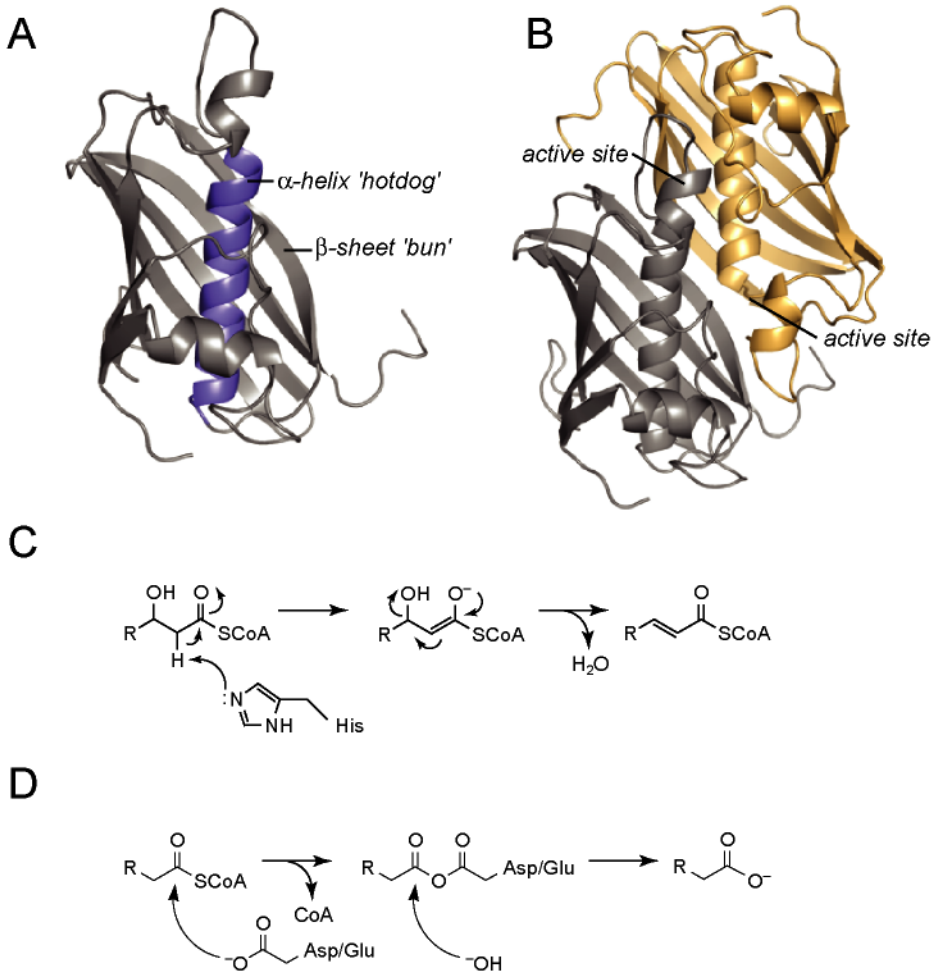


Figure 1.15. *Hotdog-fold superfamily structure and mechanism. (A) Hotdog fold monomer comprised of the 'hotdog' α-helix (blue) and β-sheet 'bun' (grey). (B) Minimal dimer unit of hotdog fold enzymes. Protomer 1, grey; protomer 2, orange. (C) Proposed mechanism for hotdog-fold dehydratases. (C) Proposed mechanism for hotdog-fold thioesterases.*

acyl-CoAs, acyl-ACPs in polyketide biosynthesis, peptidyl thioesters in nonribosomal polypeptide synthesis, and aromatic acyl-CoAs involved in catabolic pathways [*119, 129-131*].

Dehydratases of the hotdog-fold superfamily generally catalyze the *syn* elimination of water from fatty acyl-CoAs to give the α,β-unsaturated thioester (*Figure 1.15*) [*118*]. This reaction requires an absolutely conserved catalytic histidine residue and involves rate-limiting deprotonation of the thioester at the α-position, with the *pro*-2*S* proton being removed [*118*]. Hydratases catalyze the reverse reaction, and either direction may be physiologically relevant. The catalytic histidine is the only absolutely conserved catalytic residue, although all of the well-studied hydratases/dehydratases have an oxyanion hole comprised of variable amino acids, usually glycine or serine [*118, 124*]. Dehydratases/hydratases function mainly in fatty acid and polyketide biosynthesis.

Although the availability of numerous crystal structures and a wealth of biochemical data has allowed subdivision of the hotdog-fold superfamily into a number of individual families of dehydratases and thioesterases, FlK is not closely enough related to any of them to justify

categorization as one of their members. A BLAST search of the FlK sequence against the NCBI database revealed significant homology to many other hypothetical proteins in bacteria, archaea, and eukaryotes. Based on sequence conservation among these hypothetical protein sequences, a catalytic triad of Thr 42, Glu 50, and His 76 was proposed [*92*], similar to the known catalytic triad of Ser/Thr-His-Asp/Glu observed in many well-characterized hotdog-fold thioesterases. The identification of these conserved residues led to the proposal that FlK specificity for the fluorinated substrate might arise based on the attenuated reactivity of Thr as a nucleophile compared to Ser [*92, 93*]. However, recent evidence from other hotdog-fold thioesterases demonstrates that the Asp/Glu of the catalytic triad is the catalytic nucleophile and that the reaction mechanism involves an enzyme-anhydride intermediate (*Figure 1.15*) [*127, 128*]. Indeed, the catalytic Asp/Glu is the only absolutely conserved residue across all hotdog-fold thioesterases [*118*]. This suggests that other mechanisms of fluorine specificity are likely at play. Given the promiscuity typically observed in hotdog-fold thioesterases, a mechanistic understanding of FlK specificity could provide insights into molecular recognition of fluorine and fluorine's effects on catalysis in an evolved system.

Conclusion. Fluoroacetate toxicity arises from the lethal synthesis of 2-fluorocitrate. Despite the challenges presented by the need to discriminate the conservative fluorine substitution, fluoroacetate resistance is observed in nature in a wide variety of animals and microbes. Among the resistance mechanisms identified to date, a dichotomy exists between those observed in fluoroacetate nonproducers *versus* producers. While nonproducers typically harbor enzymes that catalyze fluoroacetate defluorination, fluoroacetate biosynthetic hosts appear to utilize fluoroacetyl-CoA hydrolysis, which alleviates toxicity while leaving the carbon-fluorine bond intact. Both the bacterial fluoroacetate dehalogenases and the fluoroacetyl-CoA thioesterase (FlK) from *S. cattleya* represent exciting model systems for understanding molecular recognition of fluorine. In addition, FlK provides a model system for a fluorine-specific acyl-transfer reaction, which could have implications for the use of fluoroacetyl-CoA as a building block in synthetic biology.

1.6 Specific aims and thesis organization

Studies of the interactions of manmade drugs with their protein receptors have provided considerable insights into molecular recognition of fluorine in biological systems. The enzymes involved in fluorometabolite biosynthesis and management in *Streptomyces cattleya* offer a unique opportunity to expand upon this knowledge by studying fluorine-protein interactions in a system that has evolved to utilize fluorinated substrates. In particular, the fluoroacetyl-CoA thioesterase (FlK) encoded in the fluorinase gene cluster represents an attractive model system for understanding how nature can achieve fluorine specificity in an acyl-transfer reaction, an enzymatic transformation that is near-ubiquitous in natural product biosynthesis. Elucidation of the principles of fluorine recognition is this system will not only provide insights into the evolution of an exotic chemical phenotype, but will also have implications for drug design, protein engineering, and synthetic biology.

Our approach relies upon *in vitro* structural, biochemical, and kinetic studies of FlK, as well as *in vivo* studies of fluoroacetate resistance, that will pinpoint the origins of fluorine specificity. This thesis is structured as follows: Chapter 2 describes structural studies of FlK and mutagenesis experiments based on this structural information. Chapter 3 explores the basis of fluorine discrimination at the level of chemical reactivity, while Chapter 4 describes the

thermodynamic and kinetic bases of fluorine-specific ligand binding in FlK. Chapter 5 presents studies designed to address how factors related to binding of the C-F bond to FlK influence the course of the hydrolysis reaction. Chapter 6 relates the *in vitro* studies of FlK to the biology of fluoroacetate resistance in *Streptomyces cattleya*, and also describes efforts toward molecular genetic studies of the other enzymes encoded in the fluorometabolite biosynthesis gene cluster. Finally, Chapter 7 describes the future directions and outlook for this project.

1.7 References

1. G. W. Gribble, The diversity of naturally produced organohalogens, In *Handbook of Environmental Chemistry*, Springer-Verlag: Berlin-Heidelberg, 2003, Vol. 3, pp. 1-15.
2. G. W. Gribble, Natural organohalogens: A new frontier for medicinal agents?, *J. Chem. Educ.* **2004**, *81*, 1441-1449.
3. F. H. Vaillancourt, E. Yeh, D. A. Vosburg, S. Garneau-Tsodikova and C. T. Walsh, Nature's inventory of halogenation catalysts: Oxidative strategies predominate, *Chem. Rev.* **2006**, *106*, 3364-3378.
4. G. W. Gribble, Naturally occurring organohalogen compounds--a comprehensive survey, *Fortschr. Chem. Org. Naturst.* **1996**, *68*, 1-423.
5. D. B. Harper and D. O'Hagan, The fluorinated natural products, *Nat. Prod. Rep.* **1994**, *11*, 123-133.
6. C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan and J. H. Naismith, Crystal structure and mechanism of a bacterial fluorinating enzyme, *Nature* **2004**, *427*, 561-565.
7. D. G. Fujimori and C. T. Walsh, What's new in enzymatic halogenations, *Curr. Opin. Chem. Biol.* **2007**, *11*, 553-560.
8. H. Deng and D. O'Hagan, The fluorinase, the chlorinase and the duf-62 enzymes, *Curr. Opin. Chem. Biol.* **2008**, *12*, 582-592.
9. C. Schaffrath, H. Deng and D. O'Hagan, Isolation and characterisation of 5'-fluorodeoxyadenosine synthase, a fluorination enzyme from *Streptomyces cattleya*, *FEBS Lett.* **2003**, *547*, 111-114.
10. K.-H. van Pée, Biosynthesis of halogenated metabolites by bacteria, *Annu. Rev. Microbiol.* **1996**, *50*, 375-399.
11. C. S. Neumann, D. G. Fujimori and C. T. Walsh, Halogenation strategies in natural product biosynthesis, *Chem. Biol.* **2008**, *15*, 99-109.
12. A. Butler and J. N. Carter-Franklin, The role of vanadium bromoperoxidase in the biosynthesis of halogenated marine natural products, *Nat. Prod. Rep.* **2004**, *21*, 180-188.
13. S. Keller, T. Wage, K. Hohaus, M. Holzer, E. Eichhorn and K. H. van Pée, Purification and partial characterization of tryptophan 7-halogenase (PrnA) from *Pseudomonas fluorescens*, *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 2300-2302.

14. E. Yeh, L. C. Blasiak, A. Koglin, C. L. Drennan and C. T. Walsh, Chlorination by a long-lived intermediate in the mechanism of flavin-dependent halogenases, *Biochemistry* **2007**, *46*, 1284-1292.

15. B. Hofmann, S. Tolzer, I. Pelletier, J. Altenbuchner, K. H. van Pee and H. J. Hecht, Structural investigation of the cofactor-free chloroperoxidases, *J. Mol. Biol.* **1998**, *279*, 889-900.

16. N. Itoh, T. Kawanami, J. Q. Liu, T. Dairi, M. Miyakoshi, C. Nitta and Y. Kimoto, Cloning and biochemical characterization of Co^{2+}-activated bromoperoxidase-esterase (perhydrolase) from *Pseudomonas putida* IF-3 strain, *Biochim. Biophys. Acta* **2001**, *1545*, 53-66.

17. O. Puk, P. Huber, D. Bischoff, J. Recktenwald, G. Jung, R. D. Sussmuth, K. H. van Pee, W. Wohlleben and S. Pelzer, Glycopeptide biosynthesis in *Amycolatopsis mediterranei* DSM5908: function of a halogenase and a haloperoxidase/perhydrolase, *Chem. Biol.* **2002**, *9*, 225-235.

18. E. Guenzi, G. Galli, I. Grgurina, D. C. Gross and G. Grandi, Characterization of the syringomycin synthetase gene cluster. A link between prokaryotic and eukaryotic peptide synthetases, *J. Biol. Chem.* **1998**, *273*, 32857-32863.

19. Z. Chang, P. Flatt, W. H. Gerwick, V. A. Nguyen, C. L. Willis and D. H. Sherman, The barbamide biosynthetic gene cluster: A novel marine cyanobacterial system of mixed polyketide synthase (PKS)-non-ribosomal peptide synthetase (NRPS) origin involving an unusual trichloroleucyl starter unit, *Gene* **2002**, *296*, 235-247.

20. F. H. Vaillancourt, J. Yin and C. T. Walsh, SyrB2 in syringomycin E biosynthesis is a nonheme Fe^{II} α-ketoglutarate- and O_2-dependent halogenase, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 10111-10116.

21. K. D. Koehntop, J. P. Emerson and L. Que, Jr., The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes, *J. Biol. Inorg. Chem.* **2005**, *10*, 87-93.

22. L. C. Blasiak, F. H. Vaillancourt, C. T. Walsh and C. L. Drennan, Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis, *Nature* **2006**, *440*, 368-371.

23. J. C. Price, E. W. Barr, B. Tirupati, J. M. Bollinger, Jr. and C. Krebs, The first direct characterization of a high-valent iron intermediate in the reaction of an α-ketoglutarate-dependent dioxygenase: a high-spin Fe^{IV} complex in taurine/α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*, *Biochemistry* **2003**, *42*, 7497-7508.

24. M. L. Matthews, C. M. Krest, E. W. Barr, F. H. Vaillancourt, C. T. Walsh, M. T. Green, C. Krebs and J. M. Bollinger, Substrate-triggered formation and remarkable stability of the C-H bond-cleaving chloroferryl intermediate in the aliphatic halogenase, SyrB2, *Biochemistry* **2009**, *48*, 4331-4343.

25. A. M. Wuosmaa and L. P. Hager, Methyl chloride transferase: A carbocation route for biosynthesis of halometabolites, *Science* **1990**, *249*, 160-162.

26. X. Ni and L. P. Hager, cDNA cloning of *Batis maritima* methyl chloride transferase and purification of the enzyme, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 12866-12871.

27. D. B. Harper, Halomethane from halide ion—a highly efficient fungal conversion of environmental significance, *Nature* **1985**, *315*, 55-57.

28. X. Ni and L. P. Hager, Expression of *Batis maritima* methyl chloride transferase in *Escherichia coli*, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3611-3615.

29. H. Toda and N. Itoh, Isolation and characterization of a gene encoding a Σ-adenosyl-l-methionine-dependent halide/thiol methyltransferase (HTMT) from the marine diatom *Phaeodactylum tricornutum*: Biogenic mechanism of CH_3I emissions in oceans, *Phytochemistry* **2011**, *72*, 337-343.

30. D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. Hamilton and C. D. Murphy, Biochemistry: biosynthesis of an organofluorine molecule, *Nature* **2002**, *416*, 279.

31. C. D. Cadicamo, J. Courtieu, H. Deng, A. Meddour and D. O'Hagan, Enzymatic fluorination in *Streptomyces cattleya* takes place with an inversion of configuration consistent with an S_N2 reaction mechanism, *Chembiochem* **2004**, *5*, 685-690.

32. A. S. Eustaquio, F. Pojer, J. P. Noel and B. S. Moore, Discovery and characterization of a marine bacterial SAM-dependent chlorinase, *Nat. Chem. Biol.* **2008**, *4*, 69-74.

33. A. S. Eustaquio, R. P. McGlinchey, Y. Liu, C. Hazzard, L. L. Beer, G. Florova, M. M. Alhamadsheh, A. Lechner, A. J. Kale, Y. Kobayashi, K. A. Reynolds and B. S. Moore, Biosynthesis of the salinosporamide A polyketide synthase substrate chloroethylmalonyl-coenzyme A from *S*-adenosyl-L-methionine, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 12295-12300.

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

35. A. L. Harvey, Natural products in drug discovery, *Drug Discov. Today* **2008**, *13*, 894-901.

36. A. M. Thayer, Fabulous fluorine, *Chem. Eng. News* **2006**, *84*, 15-24.

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

38. J. C. Biffinger, H. W. Kim and S. G. DiMagno, The polar hydrophobicity of fluorinated compounds, *Chembiochem* **2004**, *5*, 622-627.

39. K. Muller, C. Faeh and F. Diederich, Fluorine in pharmaceuticals: Looking beyond intuition, *Science* **2007**, *317*, 1881-1886.

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

41. L. Pauling, The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Structural Chemistry, Cornell University Press: Ithaca, NY, 1939.

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

43. J. Howard, V. Hoy, D. O'Hagan and G. Smith, How good is fluorine as a hydrogen bond acceptor?, *Tetrahedron* **1996**, *52*, 12613-12622.

44. J. Gao, S. Qiao and G. M. Whitesides, Increasing binding constants of ligands to carbonic anhydrase by using "greasy tails", *J. Med. Chem.* **1995**, *38*, 2292-2301.

45. P. B. Danielson, The cytochrome P450 superfamily: Biochemistry, evolution and drug metabolism in humans, *Curr. Drug Metab.* **2002**, *3*, 561-597.

46. M. Van Heek, C. F. France, D. S. Compton, R. L. McLeod, N. P. Yumibe, K. B. Alton, E. J. Sybertz and H. R. Davis, Jr., In vivo metabolism-based discovery of a potent cholesterol absorption inhibitor, SCH58235, in the rat and rhesus monkey through the identification of the active metabolites of SCH48461, *J. Pharmacol. Exp. Ther.* **1997**, *283*, 157-163.

47. S. B. Rosenblum, T. Huynh, A. Afonso, H. R. Davis, Jr., N. Yumibe, J. W. Clader and D. A. Burnett, Discovery of 1-(4-fluorophenyl)-(3*R*)-[3-(4-fluorophenyl)-(3*S*)-hydroxypropyl]-(4*S*)-(4 -hydroxyphenyl)-2-azetidinone (SCH 58235): A designed, potent, orally active inhibitor of cholesterol absorption, *J. Med. Chem.* **1998**, *41*, 973-980.

48. M. J. Cho and M. A. Allen, Chemical stability of prostacyclin (PGI2) in aqueous solutions, *Prostaglandins* **1978**, *15*, 943-954.

49. K. Bannai, T. Toru, T. Oba, T. Tanaka, N. Okamura, K. Watanabe, A. Hazato and S. Kurozumi, Synthesis of chemically stable prostacyclin analogs, *Tetrahedron* **1983**, *39*, 3807-3819.

50. C. S. Chang, M. Negishi, T. Nakano, Y. Morizawa, Y. Matsumura and A. Ichikawa, 7,7-Difluoroprostacyclin derivative, AFP-07, a highly selective and potent agonist for the prostacyclin receptor, *Prostaglandins* **1997**, *53*, 83-90.

51. S. Mabe, J. Eller and W. S. Champney, Structure-activity relationships for three macrolide antibiotics in *Haemophilus influenzae*, *Curr. Microbiol.* **2004**, *49*, 248-254.

52. M. T. Fera, M. Giannone, S. Pallio, A. Tortora, G. Blandino and M. Carbone, Antimicrobial activity and postantibiotic effect of flurithromycin against *Helicobacter pylori* strains, *Int. J. Antimicrob. Agents* **2001**, *17*, 151-154.

53. C. T. Walsh, Suicide substrates, mechanism-based enzyme inactivators: Recent developments, *Annu. Rev. Biochem.* **1984**, *53*, 493-535.

54. D. B. Berkowitz, K. R. Karukurichi, R. de la Salud-Bea, D. L. Nelson and C. D. McCune, Use of fluorinated functionality in enzyme inhibitor development: Mechanistic and analytical advantages, *J. Fluor. Chem.* **2008**, *129*, 731-742.

55. D. V. Santi and T. T. Sakai, Thymidylate synthetase. Model studies of inhibition by 5-trifluoromethyl-2'-deoxyuridylic acid, *Biochemistry* **1971**, *10*, 3598-3607.

56. J. E. Barrett, D. A. Maltby, D. V. Santi and P. G. Schultz, Trapping of the C5 methylene intermediate in thymidylate synthase, *J. Am. Chem. Soc.* **1998**, *120*, 449-450.

57. M. Malet-Martino, P. Jolimaitre and R. Martino, The prodrugs of 5-fluorouracil, *Curr. Med. Chem. Anticancer Agents* **2002**, *2*, 267-310.

58. R. Koshida, S. Cox, J. Harmenberg, G. Gilljam and B. Wahren, Structure-activity relationships of fluorinated nucleoside analogs and their synergistic effect in combination with phosphonoformate against human immunodeficiency virus type 1, *Antimicrob. Agents Chemother.* **1989**, *33*, 2083-2088.

59. M. Elnaggar, E. Giovannetti and G. J. Peters, Molecular targets of gemcitabine action: Rationale for development of novel drugs and drug combinations, *Curr. Pharm. Des.* **2012**, *18*, 2811-2829.

60. W. Plunkett, P. Huang and V. Gandhi, Preclinical characteristics of gemcitabine, *Anticancer Drugs* **1995**, *6 Suppl 6*, 7-13.

61. C. H. Baker, J. Banzon, J. M. Bollinger, J. Stubbe, V. Samano, M. J. Robins, B. Lippert, E. Jarvi and R. Resvick, 2'-Deoxy-2'-methylenecytidine and 2'-deoxy-2',2'-difluorocytidine 5'-diphosphates: potent mechanism-based inhibitors of ribonucleotide reductase, *J. Med. Chem.* **1991**, *34*, 1879-1884.

62. W. A. van der Donk, G. Yu, L. Perez, R. J. Sanchez, J. Stubbe, V. Samano and M. J. Robins, Detection of a new substrate-derived radical during inactivation of ribonucleotide reductase from *Escherichia coli* by gemcitabine 5'-diphosphate, *Biochemistry* **1998**, *37*, 6419-6426.

63. E. Artin, J. Wang, G. J. Lohman, K. Yokoyama, G. Yu, R. G. Griffin, G. Bar and J. Stubbe, Insight into the mechanism of inactivation of ribonucleotide reductase by gemcitabine 5'-diphosphate in the presence or absence of reductant, *Biochemistry* **2009**, *48*, 11622-11629.

64. J. Hutchinson and J. M. Dalziel, Flora of West Tropical Africa, In K. W. J. Keay, Crown Agents: London, 1958, Vol. 1.

65. B. Vickery and M. L. Vickery, Fluoride metabolism in *Dichapetalum toxicarum*, *Phytochemistry* **1972**, *11*, 1905-1909.

66. D. O'Hagan, R. Perry, J. M. Lock, J. J. Marion Meyer, L. Dasaradhi, J. T. G. Hamilton and D. B. Harper, High levels of monofluoroacetate in *Dichapetalum braunii*, *Phytochemistry* **1993**, *33*, 1043-1045.

67. P. B. Oelrichs and E. T. Mc, Isolation of the toxic principle in *Acacia georginae*, *Nature* **1961**, *190*, 808-809.

68. T. McEwan, Isolation and identification of the toxic principle of *Gastrolobium grandiflorum*, *Nature* **1964**, *201*, 827.

69. M. L. Baron, C. M. Bothroyd, G. I. Rogers, A. Staffa and I. D. Rae, Detection and measurement of fluoroacetate in plant extracts by ^{19}F NMR, *Phytochemistry* **1987**, *26*, 2293-2295.

70. M. M. de Oliveira, Chromatographic isolation of monofluoroacetic acid from *Palicourea marcgravii*, *Experientia* **1963**, *19*, 586-587.

71. T. Vartiainen and J. Gynther, Fluoroacetic acid in guar gum, *Food Chem. Toxicol.* **1984**, *22*, 307-308.

72. P. F. V. Ward, R. J. Hall and R. A. Peters, Fluoro-fatty acids in the seeds of *Dichapetalum toxicarum*, *Nature* **1964**, *201*, 611.

73. P. W. Preuss, L. Colavito and L. H. Weinstein, The synthesis of monofluoroacetic acid by a tissue culture of *Acacia georginae*, *Experientia* **1970**, *26*, 1059-1060.

74. R. A. Peters and M. Shorthouse, Fluorocitrate in plants and food stuffs, *Phytochemistry* **1972**, *11*, 1337-1338.

75. X. H. Xu, G. M. Yao, Y. M. Li, J. H. Lu, C. J. Lin, X. Wang and C. H. Kong, 5-Fluorouracil derivatives from the sponge *Phakellia fusca*, *J. Nat. Prod.* **2003**, *66*, 285-288.

76. G. O. Morton, J. E. Lancaster, G. E. Van Lear, W. Fulmor and W. E. Meyer, The structure of nucleocidin. 3. (A new structure), *J. Am. Chem. Soc.* **1969**, *91*, 1535-1537.

77. I. D. Jenkins, J. P. Verheyden and J. G. Moffatt, 4'-Substituted nucleosides. 2. Synthesis of the nucleoside antibiotic nucleocidin, *J. Am. Chem. Soc.* **1976**, *98*, 3346-3357.

78. M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch and E. Inamine, Biosynthesis of fluorothreonine and fluoroacetic acid by the thienamycin producer, *Streptomyces cattleya*, *J. Antibiot. (Tokyo)* **1986**, *39*, 259-265.

79. K. Fukuda, T. Tamura, Y. Segawa, Y. Mutaguchi and K. Inagaki, Enhanced production of the fluorinated nucleoside antibiotic nucleocidin by a *rif*^{R}-resistant mutant of *Streptomyces calvus* IFO13200, *Actinomycetelogica* **2009**, *23*, 51-55.

80. K. A. Reid, J. T. Hamilton, R. D. Bowden, D. O'Hagan, L. Dasaradhi, M. R. Amin and D. B. Harper, Biosynthesis of fluorinated secondary metabolites by *Streptomyces cattleya*, *Microbiology* **1995**, *141*, 1385-1393.

81. C. Schaffrath, C. D. Murphy, J. T. G. Hamilton and D. O'Hagan, Biosynthesis of fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*. Incorporation of oxygen-18 from [2-^{2}H, 2-^{18}O]-glycerol and the role of serine metabolites in fluoroacetaldehyde biosynthesis, *J. Chem. Soc., Perkin Trans. 1* **2001**, 3100-3105.

82. S. J. Moss, C. D. Murphy, D. O'Hagan, C. Schaffrath, J. T. G. Hamilton, W. C. McRoberts and D. B. Harper, Fluoroacetaldehyde: A precursor of both fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*, *Chem. Commun. (Camb.)* **2000**, 2281-2282.

83. C. D. Murphy, S. J. Moss and D. O'Hagan, Isolation of an aldehyde dehydrogenase involved in the oxidation of fluoroacetaldehyde to fluoroacetate in *Streptomyces cattleya*, *Appl. Environ. Microbiol.* **2001**, *67*, 4919-4921.

84. C. D. Murphy, D. O'Hagan and C. Schaffrath, Identification of a PLP-dependent threonine transaldolase: A novel enyme involved in 4-fluorothreonine biosynthesis in *Streptomyces cattleya*, *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 4479-4481.

85. H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson and J. B. Spencer, The fluorinase from *Streptomyces cattleya* is also a chlorinase, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 759-762.

86. H. Deng, S. L. Cobb, A. D. Gee, A. Lockhart, L. Martarello, R. P. McGlinchey, D. O'Hagan and M. Onega, Fluorinase mediated C-^{18}F bond formation, an enzymatic tool for PET labelling, *Chem. Commun. (Camb)* **2006**, 652-654.

87. A. S. Eustaquio, D. O'Hagan and B. S. Moore, Engineering fluorometabolite production: Fluorinase expression in *Salinispora tropica* Yields fluorosalinosporamide, *J. Nat. Prod.* **2010**, *73*, 378-382.

88. S. L. Cobb, H. Deng, J. T. Hamilton, R. P. McGlinchey and D. O'Hagan, Identification of 5-fluoro-5-deoxy-D-ribose-1-phosphate as an intermediate in fluorometabolite biosynthesis in *Streptomyces cattleya*, *Chem. Commun. (Camb.)* **2004**, 592-593.

89. M. Sauter, B. Moffatt, M. C. Saechao, R. Hell and M. Wirtz, Methionine salvage and *S*-adenosylmethionine: essential links between sulfur, ethylene and polyamine biosynthesis, *Biochem. J.* **2013**, *451*, 145-154.

90. M. Onega, R. P. McGlinchey, H. Deng, J. T. Hamilton and D. O'Hagan, The identification of (3*R*,4*S*)-5-fluoro-5-deoxy-D-ribulose-1-phosphate as an intermediate in fluorometabolite biosynthesis in *Streptomyces cattleya*, *Bioorg. Chem.* **2007**, *35*, 375-385.

91. H. Deng, S. M. Cross, R. P. McGlinchey, J. T. Hamilton and D. O'Hagan, In vitro reconstituted biotransformation of 4-fluorothreonine from fluoride ion: Application of the fluorinase, *Chem. Biol.* **2008**, *15*, 1268-1276.

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

93. M. V. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. Dry, P. F. Leadlay, J. B. Spencer and T. L. Blundell, Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK, *J. Biol. Chem.* **2010**, *285*, 22495-22504.

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

95. M. B. Chenoweth, Monofluoroacetic acid and related compounds, *J. Pharmacol. Exptl. Therap.* **1949**, *97*, 383-423.

96. E. K. Rowinsky, L. A. Cazenave and R. C. Donehower, Taxol: a novel investigational antimicrotubule agent, *J. Natl. Cancer Inst.* **1990**, *82*, 1247-1259.

97. P. Buffa and R. A. Peters, The *in vivo* formation of citrate induced by fluoroacetate and its signficance, *J. Physiol.* **1950**, *110*, 488-500.

98. P. Buffa, R. A. Peters and R. W. Wakelin, Biochemistry of fluoroacetate poisoning. Isolation of an active tricarboxylic acid fraction from poisoned kidney homogenates, *Biochem. J.* **1951**, *48*, 467-477.

99. J. F. Morrison and R. A. Peters, Biochemistry of fluoroacetate poisoning: The effect of fluorocitrate on purified aconitase, *Biochem. J.* **1954**, *58*, 473-479.

100. W. C. Stallings, C. T. Monti, J. F. Belvedere, R. K. Preston and J. P. Glusker, Absolute configuration of the isomer of fluorocitrate that inhibits aconitase, *Arch. Biochem. Biophys.* **1980**, *203*, 65-72.

101. R. Peters and R. W. Wakelin, Biochemistry of fluoroacetate poisoning; the isolation and some properties of the fluorotricarboxylic acid inhibitor of citrate metabolism, *Proc. R. Soc. Lond. B Biol. Sci.* **1953**, *140*, 497-507.

102. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme-inhibitor complex, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13699-13703.

103. E. Kun, E. Kirsten and M. L. Sharma, Enzymatic formation of glutathione-citryl thioester by a mitochondrial system and its inhibition by (-)-erythrofluorocitrate, *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 4942-4946.

104. E. Kirsten, M. L. Sharma and E. Kun, Molecular toxicology of (-)-erythro-fluorocitrate: Selective inhibition of citrate transport in mitochondria and the binding of fluorocitrate to mitochondrial proteins, *Mol. Pharmacol.* **1978**, *14*, 172-184.

105. L. E. Twigg and D. R. King, The impact of fluoroacetate-bearing vegetation on native Australian fauna: A review, *Oikos* **1991**, *61*, 412-430.

106. T. Kurihara and N. Esaki, Bacterial hydrolytic dehalogenases and related enzymes: Occurrences, reaction mechanisms, and applications, *Chem. Rec.* **2008**, *8*, 67-74.

107. J. Q. Liu, T. Kurihara, S. Ichiyama, M. Miyagi, S. Tsunasawa, H. Kawasaki, K. Soda and N. Esaki, Reaction mechanism of fluoroacetate dehalogenase from *Moraxella sp*. B, *J. Biol. Chem.* **1998**, *273*, 30897-30902.

108. H. Kawasaki, K. Tsuda, I. Matsushita and K. Tonomura, Lack of homology between two haloacetate dehalogenase genes encoded on a plasmid from *Moraxella sp*. strain B, *J. Gen. Microbiol.* **1992**, *138*, 1317-1323.

109. K. Jitsumori, R. Omi, T. Kurihara, A. Kurata, H. Mihara, I. Miyahara, K. Hirotsu and N. Esaki, X-Ray crystallographic and mutational studies of fluoroacetate dehalogenase from *Burkholderia sp*. strain FA1, *J. Bacteriol.* **2009**, *191*, 2630-2637.

110. P. W. Chan, A. F. Yakunin, E. A. Edwards and E. F. Pai, Mapping the reaction coordinates of enzymatic defluorination, *J. Am. Chem. Soc.* **2011**, *133*, 7461-7468.

111. W. Y. Chan, M. Wong, J. Guthrie, A. V. Savchenko, A. F. Yakunin, E. F. Pai and E. A. Edwards, Sequence- and activity-based screening of microbial genomes for novel dehalogenases, *Microb. Biotechnol.* **2010**, *3*, 107-120.

112. K. Gregg, B. Hamdorf, K. Henderson, J. Kopecny and C. Wong, Genetically modified ruminal bacteria protect sheep from fluoroacetate poisoning, *Appl. Environ. Microbiol.* **1998**, *64*, 3496-3498.

113. E. K. Camboim, A. P. Almeida, M. Z. Tadra-Sfeir, F. G. Junior, P. P. Andrade, C. S. McSweeney, M. A. Melo and F. Riet-Correa, Isolation and identification of sodium fluoroacetate degrading bacteria from caprine rumen in Brazil, *ScientificWorldJournal* **2012**, *2012*, 178254.

114. H. Kawasaki, H. Kuriyama and K. Tonomura, Use of haloacetate dehalogenase genes as selection markers for *Escherichia coli* and *Pseudomonas* vectors, *Biodegradation* **1995**, *6*, 213-216.

115. M. A. van den Berg and H. Y. Steensma, Expression cassettes for formaldehyde and fluoroacetate resistance, two dominant markers in *Saccharomyces cerevisiae*, *Yeast* **1997**, *13*, 551-559.

116. E. Cundliffe, Self-protection mechanisms in antibiotic producers, *Ciba Found. Symp.* **1992**, *171*, 199-208; discussion 208-214.

117. J. J. Marion Meyer, N. Grobbelaar, R. Vleggaar and A. I. Louw, Fluoroacetyl-coenzyme A hydrolase-like activity in *Dichapetalum cymosum*, *J. Plant Physiol.* **1992**, *139*, 369-372.

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

119. F. Song, Z. Zhuang, L. Finci, D. Dunaway-Mariano, R. Kniewel, J. A. Buglino, V. Solorzano, J. Wu and C. D. Lima, Structure, function, and mechanism of the phenylacetate pathway hot dog-fold thioesterase PaaI, *J. Biol. Chem.* **2006**, *281*, 11028-11038.

120. T. Yokoyama, K. J. Choi, A. M. Bosch and H. J. Yeo, Structure and function of a *Campylobacter jejuni* thioesterase Cj0915, a hexameric hot dog fold enzyme, *Biochim. Biophys. Acta* **2009**, *1794*, 1073-1081.

121. M. A. Willis, Z. Zhuang, F. Song, A. Howard, D. Dunaway-Mariano and O. Herzberg, Structure of YciA from *Haemophilus influenzae* (HI0827), a hexameric broad specificity acyl-coenzyme A thioesterase, *Biochemistry* **2008**, *47*, 2797-2805.

122. Z. Zhuang, F. Song, H. Zhao, L. Li, J. Cao, E. Eisenstein, O. Herzberg and D. Dunaway-Mariano, Divergence of function in the hot dog fold enzyme superfamily: the bacterial thioesterase YciA, *Biochemistry* **2008**, *47*, 2789-2796.

123. Z. Cheng, F. Song, X. Shan, Z. Wei, Y. Wang, D. Dunaway-Mariano and W. Gong, Crystal structure of human thioesterase superfamily member 2, *Biochem. Biophys. Res. Commun.* **2006**, *349*, 172-177.

124. T. Hisano, T. Tsuge, T. Fukui, T. Iwata, K. Miki and Y. Doi, Crystal structure of the (*R*)-specific enoyl-CoA hydratase from *Aeromonas caviae* involved in polyhydroxyalkanoate biosynthesis, *J. Biol. Chem.* **2003**, *278*, 617-624.

125. J. B. Thoden, Z. Zhuang, D. Dunaway-Mariano and H. M. Holden, The structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU, *J. Biol. Chem.* **2003**, *278*, 43709-43716.

126. M. M. Benning, G. Wesenberg, R. Liu, K. L. Taylor, D. Dunaway-Mariano and H. M. Holden, The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from *Pseudomonas sp*. strain CBS-3, *J. Biol. Chem.* **1998**, *273*, 33572-33579.

127. Z. Zhuang, J. Latham, F. Song, W. Zhang, M. Trujillo and D. Dunaway-Mariano, Investigation of the catalytic mechanism of the hotdog-fold enzyme superfamily

Pseudomonas sp. strain CBS3 4-hydroxybenzoyl-CoA thioesterase, *Biochemistry* **2012**, *51*, 786-794.

128. F. Song, J. B. Thoden, Z. Zhuang, J. Latham, M. Trujillo, H. M. Holden and D. Dunaway-Mariano, The catalytic mechanism of the hotdog-fold enzyme superfamily 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter sp*. strain SU, *Biochemistry* **2012**, *51*, 7000-7016.

129. F. Song, Z. Zhuang and D. Dunaway-Mariano, Structure-activity analysis of base and enzyme-catalyzed 4-hydroxybenzoyl coenzyme A hydrolysis, *Bioorg. Chem.* **2007**, *35*, 1-10.

130. D. Chen, R. Wu, T. L. Bryan and D. Dunaway-Mariano, *In vitro* kinetic analysis of substrate specificity in enterobactin biosynthetic lower pathway enzymes provides insight into the biochemical function of the hot dog-fold thioesterase EntH, *Biochemistry* **2009**, *48*, 511-513.

131. J. Cao, H. Xu, H. Zhao, W. Gong and D. Dunaway-Mariano, The mechanisms of human hotdog-fold thioesterase 2 (hTHEM2) substrate recognition and catalysis illuminated by a structure and function based analysis, *Biochemistry* **2009**, *48*, 1293-1304.

Chapter 2: *Structural and biochemical characterization of FlK*

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

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

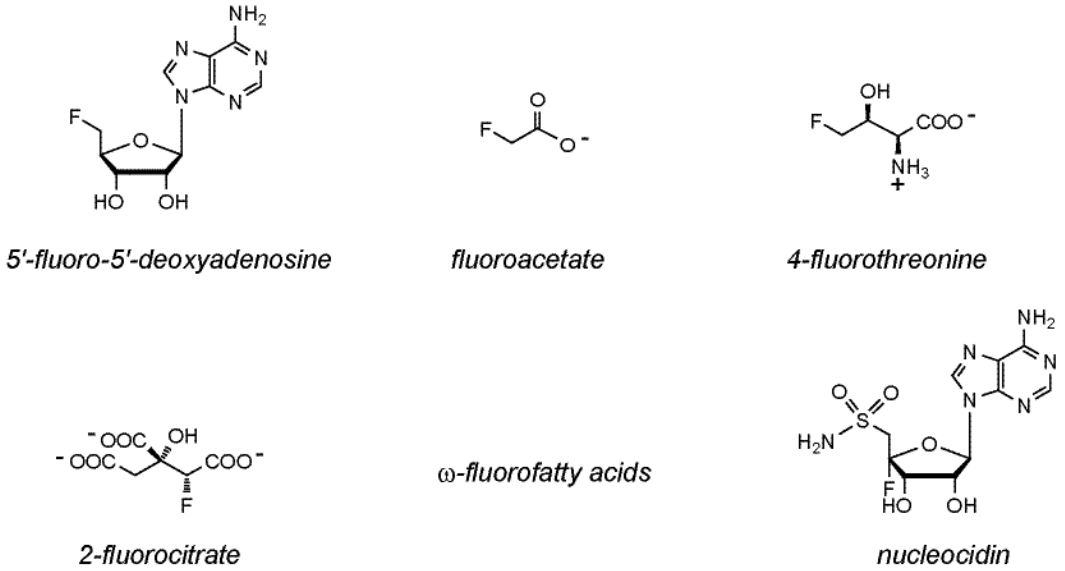
Portions of this work were performed in collaboration with the following persons:

X-ray crystallographic studies were assisted by Scott M. Coyle, Dr. Martin Jinek, and Dr. Jennifer A. Doudna.

2.1 Introduction

Fluoroacetate (*Scheme 2.1*) is a naturally occurring organofluorine that is surprisingly poisonous considering its simple structure, with toxicity rivaling some of the most potent anticancer therapeutics [*1, 2*]. Like Gemzar® and fluorouracil, it can be metabolized by the normal host machinery before generating a mechanism-based inhibitor that shuts down the target pathway upon fluoride elimination. Specifically, fluoroacetate can be activated to form fluoroacetyl-CoA, which is then converted to fluorocitrate via the TCA cycle [*3-6*]. While the aconitase-catalyzed formation of the fluoro-*cis*-aconitate intermediate parallels the reaction pathway of citrate, addition of water in the next step leads to stoichiometric fluoride elimination rather than formation of fluoroisocitrate [*7-10*]. The product of this reaction is a strong noncovalent inhibitor that remains bound in the aconitase active site with an extremely slow rate of displacement by citrate [*10, 11*].

Although fluoroacetate production has been mainly noted in the plant kingdom [*12-16*], our understanding of organofluorine biosynthesis has come exclusively from studies of *Streptomyces cattleya*, a soil microbe that is the only known genetic host for production of both fluoroacetate and fluorothreonine (*Scheme 2.1*) [*17-21*]. Similar to other antibiotic-producing hosts [*22-27*], *S. cattleya* is expected to contain pathways for detoxification in order to protect itself from its own



***Scheme 2.1**. Biogenic organofluorines found in the environment.*

production of fluoroacetate. A fluoroacetyl-CoA-specific thioesterase (FlK) was recently identified in *S. cattleya* that selectively hydrolyzes fluoroacetyl-CoA over acetyl-CoA [*28*]. The gene encoding FlK is found clustered with the C-F bond-forming fluorinase (*flA*) [*18*], raising the possibility that FlK-catalyzed hydrolysis of fluoroacetyl-CoA plays a role in fluoroacetate resistance in *S. cattleya* by preventing the entrance of fluoroacetyl-CoA into the TCA cycle.

The proposed biological function and reported selectivity of FlK led us to choose this enzyme as a model system for understanding the design principles that drive fluorine-based selectivity in protein-ligand interactions. In particular, FlK provides an interesting system for study because evolutionary selection against hydrolysis of the nonfluorinated substrate congener, acetyl-CoA, is likely important for maintaining normal cell growth.

2.2 Materials and methods

Commercial materials. Restriction enzymes, T4 DNA ligase and Phusion polymerase were purchased from New England Biolabs (Ipswich, MA). Platinum Taq HF polymerase and OneShot TOP10 chemically competent cells were purchased from Invitrogen (Carlsbad, CA). Oligonucleotides were synthesized by Integrated DNA Technologies (Coralville, IA). DNA sequencing was performed by Quintara Biosciences (Berkeley, CA). DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). Complete Mini EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany). Spectra Multicolor low range protein ladder was purchased from Fermentas (Glen Burnie, MD). Amicon Ultra 3,000 MWCO centrifugal concentrators and 5,000 MWCO regenerated cellulose ultrafiltration membranes were purchased from Millipore (Billerica, MA). Acrylamide/Bis-acrylamide (40%, 19:1 and 30%, 37:1) and Gel Filtration Standard were purchased from Bio-Rad Laboratories (Hercules, CA). Luria-Bertani (LB) Broth Miller and LB Agar Miller were purchased from EMD Biosciences (Darmstadt, Germany). Carbenicillin, isopropyl-β-D-thiogalactopyranoside (IPTG), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride and dithiothreitol (DTT) were purchased from Fisher Scientific (Pittsburgh, PA). Streptomycin sulfate, ammonium persulfate, oxalyl chloride, *N*,*N*-dimethylformamide (Sure-Seal), 5,5'-dithiobis-2-nitrobenzoic acid (DTNB), sodium fluoroacetate, coenzyme A trilithium salt (CoA), acetyl-CoA, *N*,*N*,*N*',*N*'-tetramethyl-ethane-1,2-diamine (TEMED), and tricine were purchased from Sigma-Aldrich (St. Louis, MO) .

Gene synthesis. A synthetic gene encoding flK was codon-optimized for *E. coli* using Gene Designer (DNA 2.0; Menlo Park, CA) and synthesized using PCR assembly. Gene2Oligo [*29*] was used to convert the gene sequence into primer sets using default optimization settings. Primers for gene assembly are listed in Table S1 and a primer map is shown in Figure S1. To assemble the synthetic gene, each primer was added at a final concentration of 1 μM to the first PCR (50 μL) containing Platinum Taq HiFi Buffer (20 mM Tris-HCl, 50 mM KCl, pH 8.4), $MgSO_4$ (1.5 mM), dNTPs (250 μM each), and 5 units Platinum Taq HF (Invitrogen). The following thermocycler program was used for the first assembly reaction: 95 ºC for 5 min; 95 ºC for 30 s; 55 ºC for 2 min; 72 ºC for 10 s; 40 cycles of 95 ºC for 15 s, 55 ºC for 30 s, 72 ºC for 20 s plus 3 s/cycle; these cycles were followed by a final incubation at 72 ºC for 5 min. The second assembly reaction (50 μL) contained 16 μL of the unpurified first PCR with standard reagents for Platinum Taq HF. The thermocycler program for the second PCR was: 95 ºC for 30 s; 55 ºC for 2 min; 72 ºC for 10 s; 40 cycles of 95 ºC for 15 s, 55 ºC for 30 s, 72 ºC for 80 s; these cycles were followed by a final incubation at 72 ºC for 5 min. The second PCR (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 amplification with Platinum Taq HF and rescue primers sFlK F100 and sFlK R100 (*Appendix 1*) under standard conditions. The PCR product was digested with Nde I and Bam HI and ligated into the Nde I-Bam HI sites of pET16b. The sequence of the synthetic gene was verified by sequencing

Construction of vectors for protein expression. The synthetic gene encoding FlK was PCR amplified from pET16b with Platinum Taq HF using the primers sFlK F13 and sFlK R15 (*Appendix 1*). The PCR product was digested with Sfo I and Xho I and ligated into the Sfo I-Xho I sites of pET23a vector modified to encode a His_{10} tag and a tobacco etch virus (TEV) protease cleavage site (Macrolab, UC Berkeley). The resulting plasmid was verified by sequencing.

Site-directed mutagenesis. Plasmids encoding mutant FlKs were prepared using oligonucleotide primers containing the desired mutation (*Appendix 1*). Each mutagenesis reaction contained forward and reverse primers (2 mM each), pET23a-His_{10}-Tev-sFlK template (50-100 ng), dNTPs (200 μM each), Phusion polymerase, and Phusion buffer. The reaction mixture was subjected to the following thermocycler program: 98 ºC for 3 min; 16 cycles of 98 ºC for 15 s, 55 ºC for 15 s, 72 ºC for 3 min 30 s; a final extension at 72 ºC for 7 min. Reactions were digested with Dpn I and transformed into OneShot TOP10 chemically competent cells. All mutants were then verified by sequencing of the flK gene. The E50Q mutant was subcloned into vector pSV272 (Macrolab, UC Berkeley) using primers sFlK F13 and sFlK R15 and using the Sfo I and Xho I restriction sites to generate a His_{10}-MBP-sFlK-E50Q expression construct.

***E. coli* viability assay**. Electrocompetent *E. coli* BL21(de3) cells were transformed with empty pET23a-His_{10}-Tev, pET23a-His_{10}-Tev-FlK, and pET23a-His_{10}-Tev-FlK-H76A and cells were grown on LB agar containing 50 μg/mL carbenicillin. Saturated cultures were diluted to $OD_{600\ nm}$ = 0.3 in LB and serial 10-fold dilutions were made in a 96-well plate. Dilutions (2 μL) were spotted onto LB agar containing either 50 μg/mL carbenicillin or 50 μg/mL carbenicillin and 20 mM sodium fluoroacetate. Plates were incubated at 37 ºC for 16 h and then photographed.

Expression and purification of FlK variants. LB containing carbenicillin (50 mg/L) was inoculated to OD600 nm = 0.05 with an overnight LB culture of freshly transformed *E. coli* BL21(de3) containing the overexpression plasmid. The cultures were grown at 37 °C at 200 rpm to $OD_{600\ nm}$ = 0.8, at which point they were cooled on ice for 20 min. Isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM and cells were incubated at 16 ºC for 12-16 h. Cells were harvested by centrifugation at 3,696 × *g* at 4 ºC for 15 min. The cell pellet was resuspended in 5 mL lysis buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 20 mM imidazole, 100 mM phenylmethanesulfonyl fluoride, one Complete Mini EDTA-free tablet per 50 mL buffer) per gram of cell pellet wet weight and lysed by one pass through a French pressure cell (Thermo Fisher) at 14,000 psi. Insoluble material was removed by centrifugation at 15,316 × *g* at 4 ºC for 20 min. A 10% (w/v) solution of streptomycin sulfate was added to the supernatant to a final concentration of 0.9% (w/v) and precipitated DNA was removed by centrifugation at 15,316 × *g* for 20 min. The supernatant was loaded onto a Ni-NTA agarose column (1 mL resin per liter cell culture) using an Äkta Purifier FPLC system (GE Healthcare). The column was washed with 10 column volumes of wash buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 20 mM imidazole) and 10 column volumes of wash buffer plus 8% elution buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 250 mM imidazole). The column was eluted using a gradient from 8%-100% elution buffer over 10 column volumes. Protein-containing fractions were pooled by their absorbance at 280 nm and concentrated to 10-15 mL using a 5,000 MWCO ultrafiltration membrane. TEV protease (purified as described previously [*30*]) was added at a 1:50 mass ratio and the mixture was dialyzed against TEV cleavage buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1 mM DTT) for 16 h. The dialysate was passed over a Ni-NTA column to remove the His_6-tagged TEV protease and the His_{10} peptide and the column was washed with 2 column volumes of wash buffer. The flow-through and wash were concentrated to 2 mL and loaded onto a Superdex 75 16/60 pg (GE Healthcare) column equilibrated with size-exclusion buffer (20 mM Tris-HCl, pH 7.6, 50 mM NaCl). FlK eluted from the column as a dimer (*Figure 2.2*) and was concentrated to 1-10 mg/mL and stored at 4 ºC. For long-term storage, 10% glycerol was added to the concentrated protein

and it was stored at -80 ºC. Both His_{10}-tagged and His_{10}-tagged MBP fusion proteins were purified following this protocol.

Crystallization and structure determination. Purified FlK in size-exclusion buffer was concentrated to 17 mg/mL. Crystals of apo-FlK were obtained using the hanging drop vapor diffusion method by combining equal volumes of protein solution and reservoir solution containing 0.1 M Tris-HCl, pH 7.6 and 25% polyethylene glycol 3350. Clusters of plates grew within 2 days. Single crystals were obtained by gently agitating the clusters. The crystals were cryoprotected by soaking briefly in a solution containing 0.1 M Tris-HCl, pH 7.6, 25% polyethylene glycol 3350, and 25% (v/v) ethylene glycol and flash cooled in liquid nitrogen. Data were collected at beamline 8.2.2 at the Advanced Light Source (Lawrence Berkeley National Laboratory). Data were processed with XDS and scaled with XSCALE [*31*]. Molecular replacement with Phaser [*32*] using a polyalanine model of *Thermus thermophilus* hypothetical protein TTHA0967 (PDB ID 2CWZ) identified two FlK monomers in the asymmetric unit arranged with typical hot dog-fold topology. Maps were improved by Phenix AutoSolve and AutoBuild [*33*]. The resulting maps were of sufficient quality to discard the original model and Arp/wARP [*34*] was used to build a near-complete chain trace. The remaining parts of the model were manually built using Coot [*35*]. The structure was solved at 1.9 Å resolution and refined to an R_{work}/R_{free} of 20.1%/23.1% using Phenix Refine.

Crystals of the open conformation of FlK were obtained by soaking apo-FlK crystals in a solution of 1 mM fluoroacetyl-CoA in mother liquor for 10 min followed by brief cryoprotection by soaking in mother liquor supplemented with 25% (v/v) ethylene glycol. Crystals were then flash-cooled and data were collected at beamline 8.3.1 (Advanced Light Source, Lawrence Berkeley National Lab). Molecular replacement performed with Phaser using the refined apo-FlK structure as a search model identified two monomers per asymmetric unit. The open FlK structure was manually rebuilt using Coot, and refined using Phenix Refine to R_{work} = 19.8% and R_{free} = 22.7% at 2.0 Å resolution. Crystals of the FlK-fluoroacetate complex were obtained by soaking apo-FlK crystals in 2.5 mM fluoroacetate. The initial model of the protein was obtained as for the open conformation of FlK. Unmodelled peaks in the map were identified using Coot and ligands corresponding to the density were modeled when possible. The structure was solved to 2.5 Å resolution and refined with Phenix Refine to R_{work}/R_{free} of 19.4%/24.0%

Crystals of FlK-F36A were obtained as for wild-type FlK. FlK-F36A crystals grew in a larger C2 crystal form that had a Matthews coefficient consistent with six monomers per asymmetric unit. Molecular replacement performed with Phaser using the refined apo-FlK structure as a search model unambiguously identified three dimers per asymmetric unit, in agreement with the Matthews coefficient. The structure was solved to 2.3 Å resolution, manually rebuilt using Coot, and refined to R_{work}/R_{free} of 21.1%/24.1%. Crystals of the FlK-F36A product complex were obtained in the same conditions as apo-FlK-F36A by cocrystallization with 1 mM fluoroacetyl-CoA and were found to contain three dimers per asymmetric unit. Maps were obtained and the model was built as for FlK-F36A. The structure was solved to 2.0 Å resolution and refined with Phenix Refine to R_{work}/R_{free} of 21.2%/25.0%.

Fluoroacetyl-CoA synthesis. Fluoroacetyl-CoA was synthesized as described previously [*28*] with some modification. Sodium fluoroacetate (100 mg, 1 mmol) was dried under vacuum in an oven-dried round-bottom flask equipped with a stir bar and a reflux condenser. Dry tetrahydrofuran (2 mL), dry *N*,*N*-dimethylformamide (100 mL), and oxalyl chloride (2 M in dichloromethane, 1 mmol) were added to the flask by syringe and the reaction mixture was

stirred and heated at 65 ºC for 2-3 hours under nitrogen atmosphere. After allowing the flask to cool, the reaction mixture (650 mL) was added to stirred solution of coenzyme A trilithium salt (50 mg, 0.06 mmol) dissolved in 1 mL 10% sodium bicarbonate. The mixture was stirred vigorously for 3 min under a stream of nitrogen and then immediately injected onto an Agilent Eclipse XDB-C18 column (9.4 x 250 mm, 5 mm) and purified by reverse-phase HPLC (0-100% B over 30 min at 3 mL/min; A: H2O, B: acetonitrile) using an Agilent 1200 binary pump coupled to a diode-array detector. Fractions (1.5 mL) were flash frozen in liquid nitrogen and lyophilized. The lyophilized fluoroacetyl-CoA was dissolved in water and characterized using an Agilent 6130 single-quadrupole electrospray ionization mass spectrometer (ESI-MS) (*m/z* (MH^+), 828.1; *m/z* (MH^+_{calc}), 828.1). The purified substrate was also assayed for contaminating coenzyme A by comparing the total coenzyme A absorbance at 260 nm ($\varepsilon_{260\ nm}$ = 13,100 M^{-1} cm^{-1}) with free thiol content measured by reaction with 5,5´-dithiobis-2-nitrobenzoic acid ($\varepsilon_{412\ nm}$ = 13,600 M^{-1} cm^{-1} for 2-nitro-5-thiobenzoate). Only fluoroacetyl-CoA of >95% purity was used for assays.

Kinetic measurements. The thioesterase activity of FlK was measured by spectrophotometrically monitoring the increase in absorbance at 412 nm due to reaction of enzymatically generated free coenzyme A with 5,5'-dithiobis-2-nitrobenzoic acid (DTNB) in a Beckman Coulter DU-800 spectrophotometer as described previously [*28*] with some modification. FlK activity was measured immediately after purification. Assays were performed at 25 ºC in a total volume of 500 mL containing 100 mM Tris-HCl, pH 7.6, 0.5 mM DTNB, the appropriate amount of fluoroacetyl-CoA or acetyl-CoA, and FlK (0.5 nM, fluoroacetyl-CoA; 10 µM, acetyl-CoA) or FlK mutant (50 nM-1 µM, fluoroacetyl-CoA; 10 µM for acetyl-CoA) in an amount sufficient to detect activity. Commercial acetyl-CoA was also tested for contaminating CoA as described above. A CoA standard curve from 0-50 µM was generated using the appropriate amounts of CoA in a solution containing 100 mM Tris-HCl, pH 7.6 and 0.5 mM DTNB. Kinetic parameters (k_{cat} and K_M) were determined by fitting the data using Kaleidagraph 3.51 (Synergy Software, Reading PA) to the equation:

$$v_o = V_{max}\,[S]/\,K_M + [S]$$

where v_o is the initial rate and [S] is the substrate concentration.

2.3 Results and discussion

Toxicity of fluoroacetyl-CoA and rescue by FlK. To test whether FlK is competent to confer fluoroacetate resistance in a heterologous bacterial system, we expressed the *flK* gene in *E. coli* and measured its viability by spotting 1:10 dilutions of culture onto solid media (*Figure 2.1*). In the presence of fluoroacetate, the viability of *E. coli* containing an empty plasmid or a plasmid encoding an FlK mutant with diminished activity was severely reduced, while *E. coli* harboring the plasmid encoding wild-type FlK exhibited viability similar to that of cultures spotted onto fluoroacetate-free media, which indicates that FlK is competent to confer fluoroacetate resistance in bacteria. As acetyl-CoA is an essential building block in the cell, these results are consistent with the conclusion that FlK demonstrates specificity towards fluoroacetyl-CoA over acetyl-CoA in vivo.

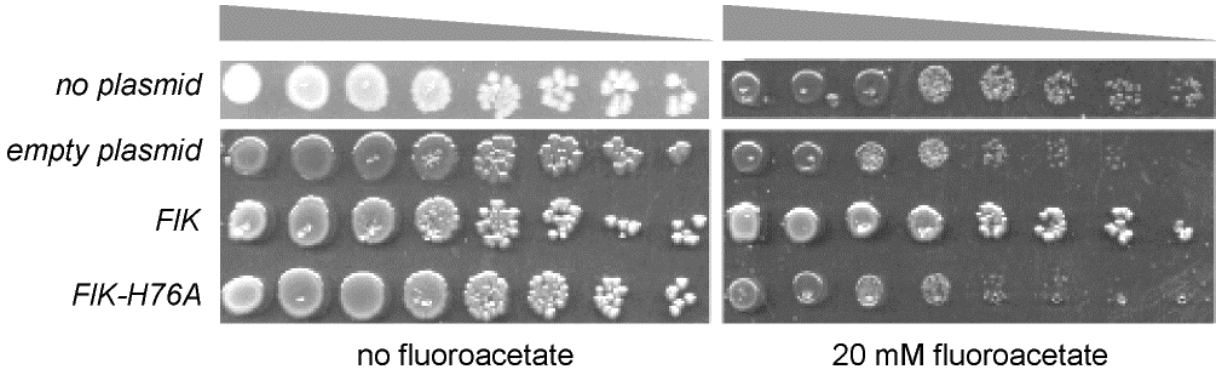


Figure 2.1. *FlK confers fluoroacetate resistance in vivo. The viability of E. coli BL21 (de3) was assayed by spotting 1:10 serial dilutions of culture onto solid media in the absence (left) or presence (right) of 20 mM fluoroacetate. E. coli harbored no plasmid (top row), empty pET23a-His$_{10}$-Tev (second row), pET23a-His$_{10}$-Tev-FlK (third row), or pET23a-His$_{10}$-Tev-FlK-H76A (fourth row).*

FlK purification strategy and determination of solution molecular weight. To further study FlK substrate specificity, we set out to purify recombinant FlK of suitable quality for crystallographic and kinetic studies. Because the high GC content of *Streptomyces cattleya* genes can make them difficult to amplify with PCR and can lead to poor expression in heterologous hosts, we synthesized an FlK gene that was codon-optimized for expression in *E. coli*. This synthetic gene was initially cloned into pET16b with an N-terminal His tag, expressed from a T7 promoter, and purified by Ni-NTA affinity chromatography followed by size-exclusion chromatography (*Figure 2.2*A). Although this His-tagged FlK expressed at high levels, SDS-PAGE analysis revealed that the purified protein could not be separated from a smaller band that was presumed to be a truncation product. The purified protein was nonetheless screened for crystallization, but produced no hits.

To improve the quality of FlK protein preparation, a second purification strategy was developed. FlK was cloned into a modified version of pET23a containing a N-terminal His-tag followed by a tobacco etch virus (TEV) protease site to allow for removal of the His-tag after Ni-NTA affinity purification, leaving GTGA-FlK. This FlK construct was purified by Ni-NTA affinity chromatography, cleaved with TEV protease, and then subjected to a second Ni-NTA purification to remove uncleaved FlK (*Figure 2.2*B). SDS-PAGE analysis of the flowthrough from the second Ni-NTA column showed this protein preparation to be homogeneous. In order to determine the solution molecular weight and to obtain FlK that was homogenous in terms of its oligomeric state, GTGA-FlK was further purified by size-exclusion chromatography (*Figure 2.2*C). This removed a small amount soluble aggregate that eluted in the void volume. By comparing the elution time of the major FlK peak with a set of gel filtration standards, FlK was determined to form a dimer in solution, consistent with observations for other hotdog-fold thioesterases.

Testing for FlK-catalyzed acetyl-CoA hydrolysis. Previous measurements of FlK activity suggested that this enzyme exhibits remarkable in vitro selectivity for fluoroacetyl-CoA over acetyl-CoA, with no hydrolysis observed at concentrations up to 1 mM acetyl-CoA in the presence of 50 ng FlK [*28*]. We can envision two major modes for recognition of the fluorinated substrate over the nonfluorinated substrate. One mode of substrate specificity could rely on structural recognition of the fluorine atom, which could be discriminated based on polarity or

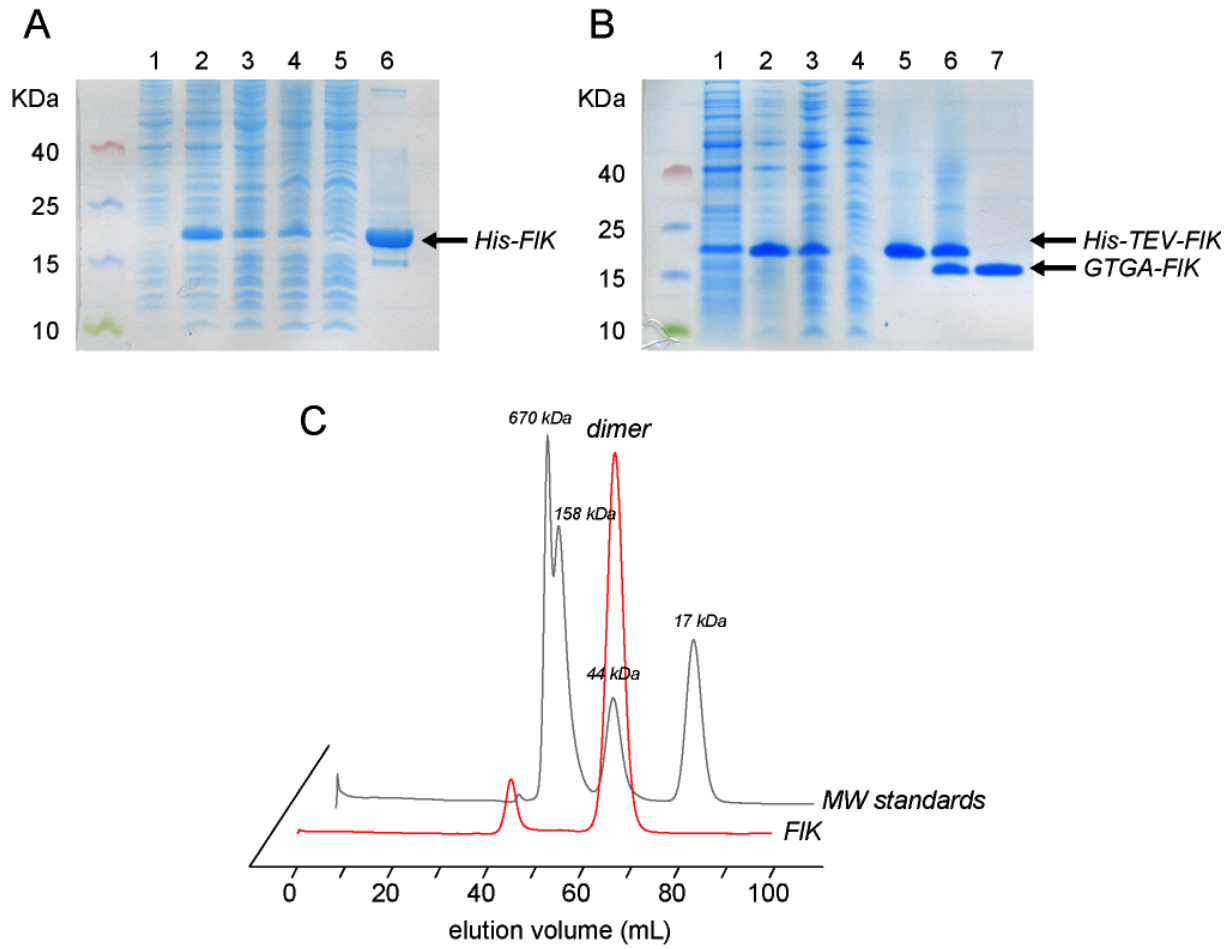


Figure 2.2. *Purification of wild-type FIK. (A) SDS-PAGE (10% tricine gel) analysis of FIK purified as a His-tag fusion. Pre-induction sample (lane 1), post-induction sample (lane 2), total soluble protein (lane 3), supernatant after streptomycin sulfate precipitation (lane 4), flow-through from Ni-NTA column (lane 5), Ni-NTA column eluant (lane 5). (B) Pre-induction sample (lane 1), post-induction sample (lane 2), total soluble protein (lane 3), flow-through from the first Ni-NTA column (lane 4), Ni-NTA column eluant (lane 5), FIK after treatment with TEV protease to remove the His_6-tag (lane 6), flow-through from the second Ni-NTA column (lane 7). Spectra Low-Range molecular weight standards (Fermentas) were used for comparison. (C) Molecular weight determination and separation of dimeric FIK from soluble aggregates by size-exclusion chromatography. The trace for FIK is shown in red and gel filtration standards (Bio-Rad Laboratories) are show in gray.*

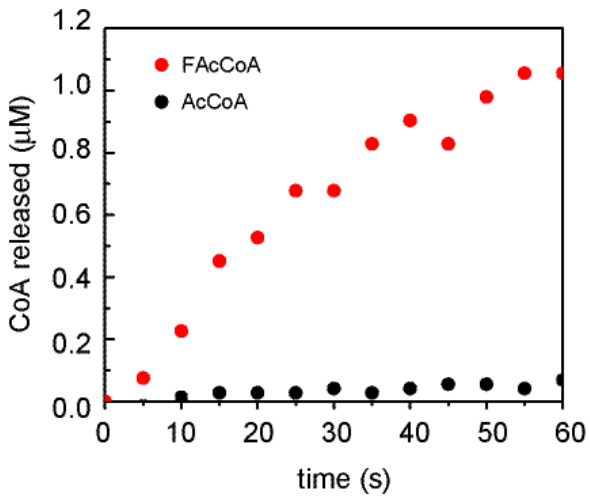


***Figure 2.3**. Uncatalyzed rates of fluoroacetyl-CoA (1 mM) and acetyl-CoA (1 mM) hydrolysis.*

size. A second possible component of selectivity is the differing chemical reactivities of fluoroacetyl-CoA and acetyl-CoA. The inductively electron-withdrawing fluorine atom activates the thioester C=O in fluoroacetyl-CoA toward nucleophilic attack. Indeed, the pseudo-first order rate constant for uncatalyzed hydrolysis of fluoroacetyl-CoA was measured to be an order of magnitude faster (1.4×10^{-4} s^{-1} at pH 7.6) than the hydrolysis of acetyl-CoA (1.3×10^{-5} s^{-1} at pH 7.6) under the same conditions (*Figure 2.3*). Thus, the active site nucleophile of FlK could be tuned towards lower reactivity, which would limit the hydrolysis of acetyl-CoA compared to fluoroacetyl-CoA.

To clarify whether the exclusion of acetyl-CoA binding plays a role in FlK selectivity, we performed a competition experiment in which the fluoroacetyl-CoA concentration was fixed at 20 μM and the rate of hydrolysis was measured in the presence of increasing concentrations of acetyl-CoA (*Figure 2.4*). No significant effect on the rate of free CoA release was observed except at the highest concentration (5 mM) of acetyl-CoA, suggesting that the absence of the fluorine modification severely diminishes substrate affinity. This result is surprising in light of previous studies that have demonstrated promiscuous substrate binding in other members of the hot dog-fold thioesterase superfamily, which have been shown to accept a range of diverse acyl-CoA substrates with physiologically relevant k_{cat}/K_M values [*36-41*]. Because fluoroacetyl-CoA is larger than acetyl-CoA (estimated total size: C-F, 2.82 Å; C-H, 2.29 Å [*42*]), we might have expected that the binding pocket of FlK would also accommodate acetyl-CoA as a substrate, especially considering that several of the substrate binding interactions with CoA-dependent enzymes involve the carbonyl group and the cofactor itself [*36, 40*].

Based on the results of the competition experiment, it seemed likely that the K_M for acetyl-CoA is above 1 mM, the highest concentration tested in previous studies [*28*]. After increasing both enzyme and substrate concentrations, acetyl-CoA turnover was detectable with a rate well above the non-enzymatic hydrolysis rate (*Figure 2.5*). Under these conditions, we measured a k_{cat} of 0.06 ± 0.001 s^{-1} and a k_{cat}/K_M of 30 M^{-1} s^{-1} for acetyl-CoA, which corresponds to a 10^6-fold decrease in catalytic efficiency compared to the fluoroacetyl-CoA substrate (k_{cat} = 390 ± 20 s^{-1}; k_{cat}/K_M = 5×10^7 M^{-1} s^{-1}). The combination of the slow rate of turnover and low pseudo-second order rate constant with

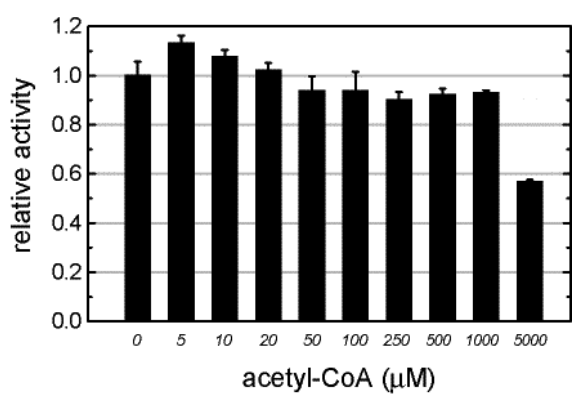


***Figure 2.4.** Rates of FlK-catalyzed hydrolysis of fluoroacetyl-CoA in the presence of varying concentrations of acetyl-CoA. The ability of acetyl-CoA to bind to FlK was tested using a competition assay in which the rate of hydrolysis of fluoroacetyl-CoA (fixed at 20 μM) was measured in the presence of increasing concentrations of acetyl-CoA. Data are mean ± SE (n = 3).*

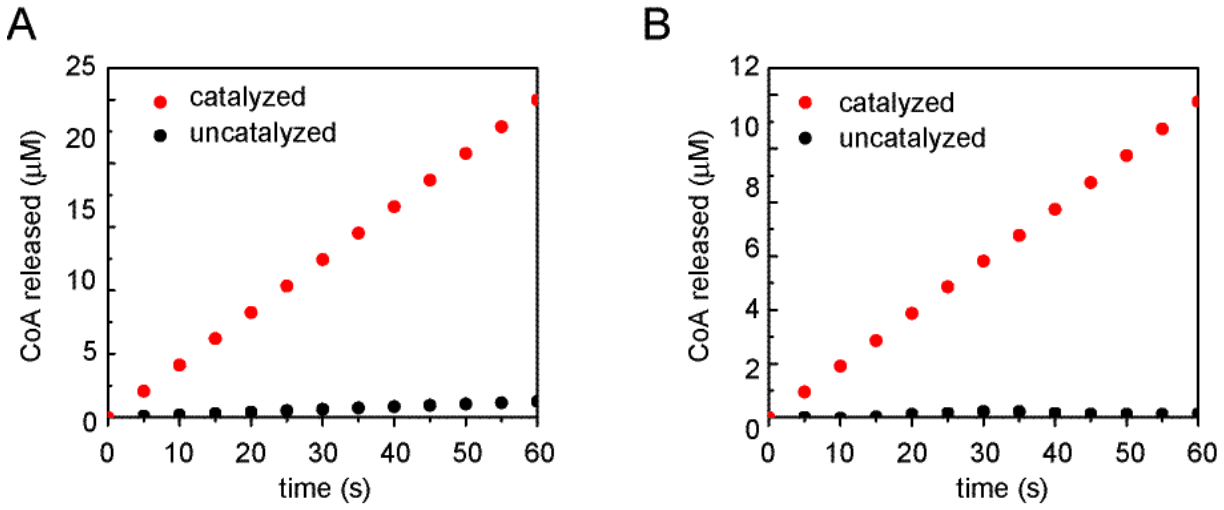


***Figure 2.5**. Comparison of catalyzed and uncatalyzed rates of hydrolysis. (A) Fluoroacetyl-CoA (150 μM in the presence or absence of 0.5 nM FlK). (B) Acetyl-CoA (1 mM in the presence or absence of 10 μM FlK).*

acetyl-CoA indicate that the nonfluorinated substrate binds poorly, and potentially in unproductive conformations, at the FlK active site. In contrast, the kinetic parameters measured for fluoroacetyl-CoA are similar to the values observed for other members of this superfamily for their substrates (k_{cat}, 0.56 to 190 s^{-1}; k_{cat}/K_M, 10^4 to 10^7 M^{-1} s^{-1}) [*36, 37, 39-41, 43-45*]. Taken together, the kinetic data for fluoroacetyl-CoA and acetyl-CoA are consistent with selection against acetyl-CoA as a substrate by eliminating recognition of common acyl-CoA features. Given that the magnitude of the discrimination exhibited by FlK is outside the window expected of single C-H to C-F substitutions that do not perturb ligand conformation or local pK_as [*42, 46, 47*], we turned to structural studies of FlK and FlK mutants in order to further elucidate the interactions at the active site that might promote binding and hydrolysis of the fluorinated substrate while excluding its nonfluorinated congener.

Crystal structure of FlK. A set of FlK structures were recently reported [*48*], with FlK found to be arranged as a dimer in a typical hot dog-fold thioesterase topology [*36, 37, 40, 45*]. We have also determined the crystal structure of the wild-type enzyme at 1.9 Å resolution by molecular replacement using a polyalanine model of *T. thermophilus* hypothetical protein TTHA0967 (PDB ID 2CWZ) as a search model (*Figure 2.6*, *Table 2.1*). The asymmetric unit contained two copies of the FlK dimer, in agreement with the apparent molecular weight measured by size-exclusion chromatography (*Figure 2.2*). A BLAST search [*49*] of FlK followed by multiple sequence alignment using MUSCLE [*50*] suggests that the FlK active site is composed of Thr 42, Glu 50, and His 76. Indeed, these residues are found at the dimer interface and arranged with a topology similar to the Arthrobacter clade of hot dog-fold thioesterases (*Figure 2.6*B) [*36*]. Although FlK adopts the fold that defines the superfamily, the structural data reveal some unique features not observed in structures of other superfamily members. The most striking difference is a "lid" structure formed by short helical segments (α1 and α2, residues 23-26 and 31-34, respectively) connected by flexible linkers, which is positioned directly over the active site of FlK and constrains the size of the active site pocket available for substrate binding (*Figure 2.6*C). Tertiary contacts within the lid are mediated by an edge-face interaction between Phe 33 and Phe 36 and by close packing of the side chains of Val 23, Leu 26, Val 39, and Phe 36. The surface of the lid that faces the substrate-binding cavity is

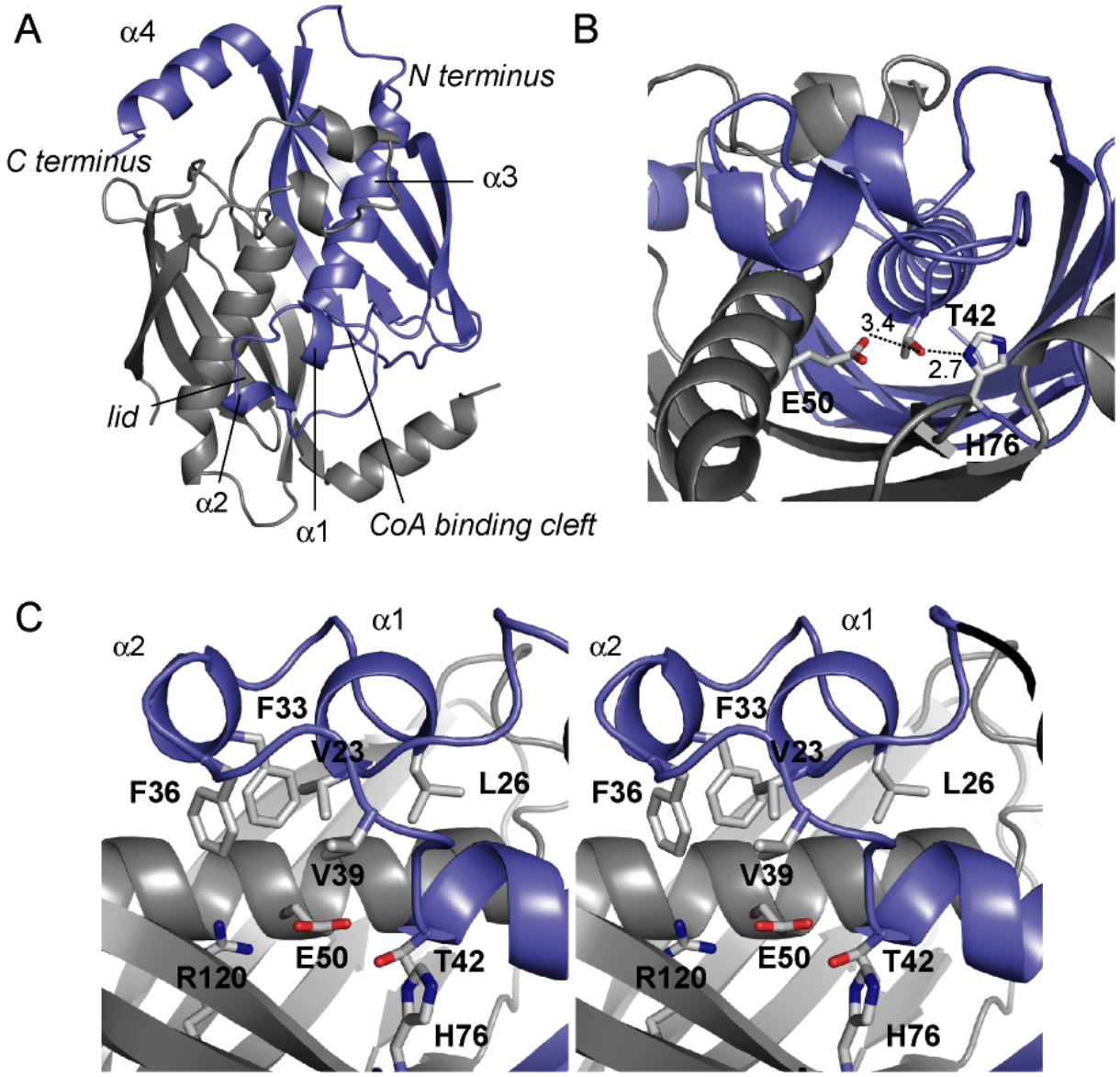


***Figure 2.6**. Crystal structure of FIK. (A) Cartoon representation of the FIK crystal structure looking down at the active site and lid structure formed by $\alpha 1$ and $\alpha 2$ (residues 23-26 and 31-34, repectively). Chain A is colored in blue and chain B is colored in gray ($\alpha 3$, residues 42-58; $\alpha 4$, residues 125-135). (B) View of the hydrogen bonding network at the putative active site showing Thr 42, His 76, and Glu 50 site water (D) Stereo view of the active site and the hydrophobic lid. (carbon, gray; nitrogen, blue; oxygen, red; sulfur, yellow).*

***Table 2.1**. Data collection and refinement statistics for wild-type FIK structures.*

	apo-FIK	FIK-FAc	FIK open
Data Collection			
X-ray source	ALS 8.2.2	ALS 8.3.1	ALS 8.3.1
Space group	C2	C2	C2
Cell dimensions			
a, b, c (Å)	48.4, 90.9, 62.0	62.3, 92.4, 50.3	62.8, 90.7, 52.3
α β γ (°)	90.0, 100.1, 90.0	90.0, 102.0, 90.0	90.0, 104.5, 90.0
Wavelength	0.9901	1.116	1.116
Resolution*	29.96-1.85 (1.85-1.90)	19.44-2.46 (2.50-2.46)	19.77-1.95 (1.95-2.00)
R_{merge}*	9.0% (60.1%)	10.6% (70.3%)	8.5% (60.4%)
I/σI*	12.25 (2.31)	12.94 (2.14)	14.01 (2.19)
Completeness*	99.7% (99.8%)	99.5% (100%)	99.7% (99.7%)
Redundancy*	4.1 (4.1)	4.1 (4.2)	3.7 (3.7)
Refinement			
Resolution	29.96-1.85	19.44-2.46	19.77-1.95
No. reflections	20926	10125	20623
R_{work}/R_{free}	0.201/0.231	0.194/0.240	0.198/0.227
No. atoms	2306	2195	2182
Non-solvent	2065	2103	2068
Solvent	241	92	114
B-factors			
Protein	29.2	38.3	23.4
Water	43.4	39.0	30.2
rmsd			
Bond lengths (Å)	0.014	0.005	0.013
Bond angles (°)	1.24	0.97	1.58

*Values in parentheses are for the highest resolution shell.

```
                                 1              10        20        30        40        50
S. cattleya (FlK)               -MKDGMRV------GERFTHDFVVPPHKTVRHLYPESPEFAEFPEVFATGFMVGLMEWAC
D. acidovorans                  -MKDAMAINHPIPLGLRHSQTLRVDDSLTVPAVSAAFTGFSDMPPVFATAYMVGFVEWAC
Geobacter sp. M18               -MKDTLAA------GIGTTLKFSVPVEKTVPCLYPESALFREMPEVFATGYLVGFIEWAC
G. bemidjiensis                 -MK-ELQV------GLKHTFSYLVPKERTVPFLYPESSYFQVMPEVFATGYMVGFMEWAC
M. acetivorans C2A              MDSSTLKP------GLAYEFRFKIPENKTVPYLYPESPEFQVMPKVFATGFMVGLFEWAC
Anaeromyxobacter sp.            -MKSTLAP------GVSLTFRYQVPETKTVPHVFPESPRFVEMPQVFATAFMVGLLEWAC
G. ferruginea                   -MKDTLKP------GIRFEHKYLVPANKTVPALYPESPEFLAMPEVFATGFMVGFLEWAC
S. lithotrophicus ES-1          -MKDTLKP------GIRYEHRFLVPSSKTVPALYPEAEEFLAMPEVFATGFLVGFLEWAC
R. eutropha JMP134              -MSPELRP------GLTFTWEYPVPPKATVPRLYDDIAMCTEMPDVLATGYMVGIMECAC
R. eutropha H16                 -MSPDLRP------GLTFSWQYTVPPKATVPRLYDDIPGCPEMPDVLATGYMVGIMECAC
C. taiwanensis                  -MSPDLRP------GLAFSWQYTVPPKATVPRLYDDIPGCPEMPDVLATGYLVGIMECAC
Ruminococcus sp. 18P13          -MKEIIA-------GIKGEAELKVSSNELAVNVGSGSL------EVFATPVMVMLMEKAA
Bacteroides sp. D4              -MMET---------GLTYTSTVVVSKENVAATMGSGDL------NVFATPAMVALMENAA
V. dispar                       -MVSA---------GQTATATVTVTESNIAKTMKSGSL------EVFATPAMCALMEEAA
                                     :        .        :     .  :            * **  :  :.*  .

                                     60         70        80        90        100       110
S. cattleya (FlK)               VRAMAPYLE-PGEGSLGTAICVTHTAATPPGLTVTVTAELRSVEGRRLSWRVSAHDGVDE
D. acidovorans                  IEALRPYLA-PSQRTVGTHVNLSHSAATPVGMQVTAEVELIEVEGKRLTFKVLCRDEVDV
Geobacter sp. M18               MEALAPYLE-EDERSVGTMINVTHSAATPPGMEVTAQVRCVEVTGKRTVWEIEVHDQADL
G. bemidjiensis                 MDALAPYLD-EGERTVGTMINVTHEAATPAGMEVTATVTLVEVDGKRTVWEIEARDEVEV
M. acetivorans C2A              IQAINPYLDFPAEQTVGTDVRLSHSAATPPGLTVTVKIKLEKIEGRKLTFSIIADDGVDK
Anaeromyxobacter sp.            IEAMQPHLD-GGEQSVGTGIWVTHGAATPPGFTVTVDVAVTKVEGRRLTFSVRAHDGVDA
G. ferruginea                   IMAIKPHLDWPEEQSVGTHINVSHEAATPPGLEVTASVELTIVDGRRLTFAVSAHDGVDT
S. lithotrophicus ES-1          IKCINPHIDWPAEQTVGTHINVSHQAATPPGLEVTALVELVEVDGRKLVFQVEAHDGVEV
R. eutropha JMP134              LQALRDYLDWPREQTLGTLVSFSHLAATPPGMTITVKGQLVEVDGRRLRFEVSAWDGEDK
R. eutropha H16                 LQMLREHLDWPREQSLGTLVSFSHLAPTPPGMTVTVKGELVAVDGRRLRFQLSAWDGEDK
C. taiwanensis                  LQMLREHLDWPREQSLGTLVSFSHLAPTPPGMTVTVKGELVEVDGRRLRFQLSAWDGEDK
Ruminococcus sp. 18P13          CKCVADYME-NDETTVGTEMNVKHLSASPAESDIRAEAELTEVNGRELVFSVKAYDNCGI
Bacteroides sp. D4              MNAVAGGLP-EGSTTVGAMMNTTHIKPSAVGDTVSATAVLKEVEGRKLTFEVRAQDSKGV
V. dispar                       QAAVQPYLE-DGEGTVGIALSITHEAPTPLGATVTAKATVSAVEGRKITFDIEASDGVGI
                                   :   :      ::*  :  .*  .:     :           .   : :   *

                                      120       130
S. cattleya (FlK)               IGSGTHERAVIHLEKFNAKVRQKTPAG------------
D. acidovorans                  ICEGRHERFIVEAQQFIRRVTSKGERG------------
Geobacter sp. M18               ISKGTHERFTIRLEQFKSRLKTKAEAAGIIIA-------
G. bemidjiensis                 IGRGRHERFVIDYEKFSKRVAAKGNK-------------
M. acetivorans C2A              ISEGTHERFIIDAAKFNSKAEAKAKNANN----------
Anaeromyxobacter sp.            ICEGTHERFVIDRARFDRKIQEKLAASTSC---------
G. ferruginea                   IARGTHERYVINKEKFDNKLRDKREKINQR---------
S. lithotrophicus ES-1          ISKGRHERFIINREKFEAKIGEKMRRSDT----------
R. eutropha JMP134              ITEGVHERHAIDAARFNEKVAAKAARAAG----------
R. eutropha H16                 ISEGVHERHLIDAGRFNEKVAAKAARAAG----------
C. taiwanensis                  ISEGVHERHLIDAGRFNQKVAAKAARAAG----------
Ruminococcus sp. 18P13          IGEGVHKRFLVFGGRFTEKAKAKLQK-------------
Bacteroides sp. D4              IGEGTHVRYIVDKEKFMSKLS------------------
V.dispar                        IGRGTHERFVINNEKFMAKVTSRAKSN------------
                                       *  .   *  :    *
```

FIgure 2.7. *Multiple sequence alignment of FlK and homologs identified by BLAST. Three groups of homologs are represented: (1) conserved amino acid identity at Phe 33 and Trp 51 (D. acidovorans to S. litotrophicus), (2) Phe 33 and Trp 51 substituted by cysteine (R. eutropha to C. taiwenensis), and (3) lid region missing (Ruminococcus sp. D4 to V. dispar). The lid in FlK is built from a loop containing α1 (residues 23-26), α2 (residues 31-34), and the neighboring flexible linker regions. At this time, all of these homologs are of unknown function and their substrate specificity is uncharacterized. Invariant amino acids are denoted with an asterisk, highly similar amino acids are denoted with a colon, and similar amino acids are denoted with a period. Residues investigated in this study using site-directed mutagenesis are colored red.*

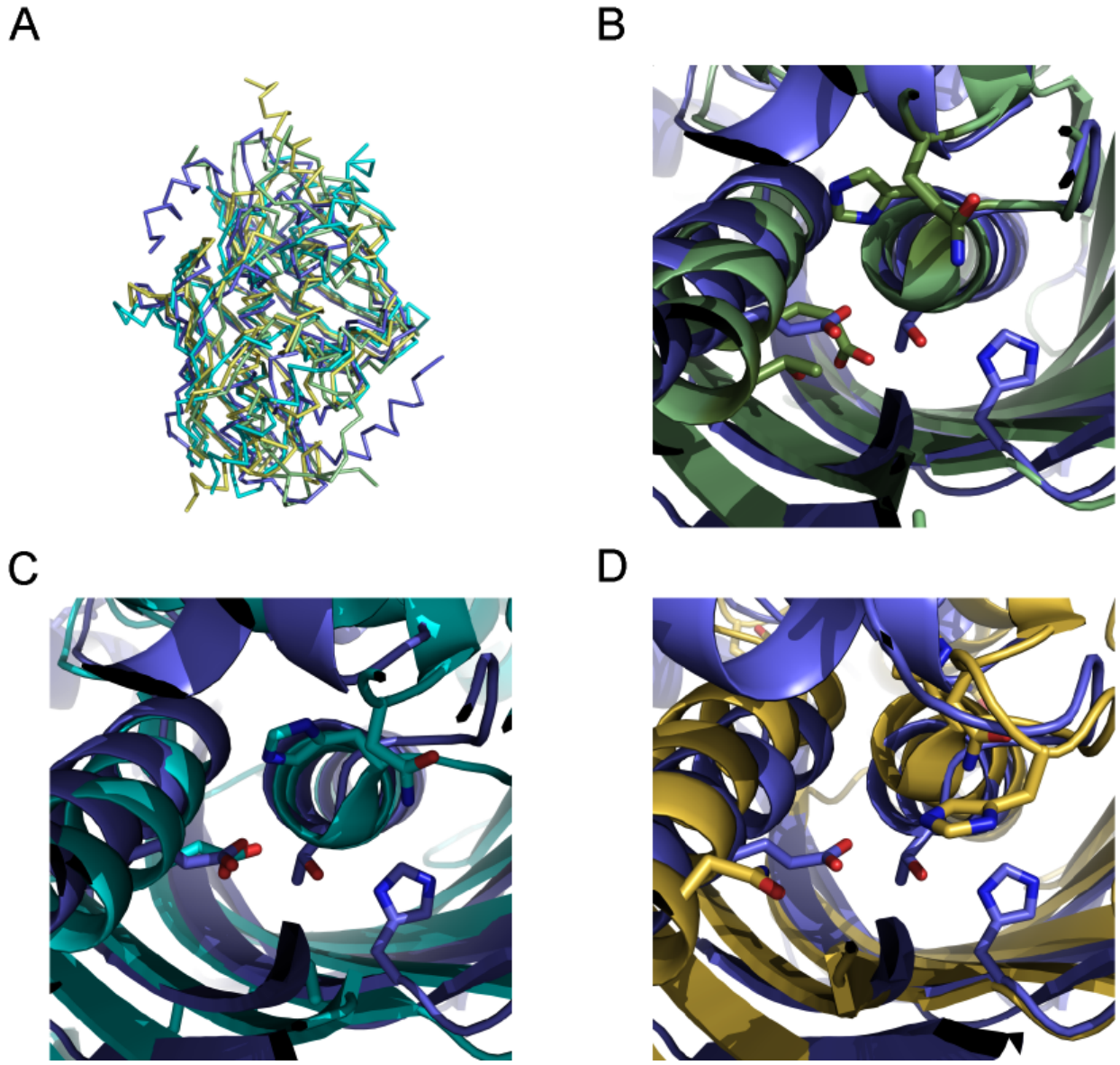


Figure 2.8*. Structural alignments of FlK, 4-hydroxybenzoyl-CoA thioesterase (PDB ID 1Q4S), hTHEM2 (PDB ID 3F5O), and PaaI (PDB ID 2FS2). (A) Ribbon diagram of the four aligned structures. FlK is shown in blue, 4-HBA-CoA thioesterase is shown in green, hTHEM2 is shown in teal, and PaaI is shown in yellow. (B) Alignment of FlK with 4-HBA-CoA thioesterase showing active site residues (T42, H76, and E50 for FlK; H64, E73, T77, and N58 for 4-HBA-CoA thioesterase. (C) Alignment of FlK with hTHEM2 showing active site residues (T42, H76, and E50 for FlK; H56, D65, N50, and S83 for hTHEM2). (D) Alignment of FlK with PaaI showing active site residues (T42, H76, E50 for FlK; H48, N46 and D61 for PaaI). (carbon, colored by chain; nitrogen, blue; oxygen, red)*

made up primarily of hydrophobic side chains (Val 23, Leu 26, Tyr 27, Phe 33, Phe 36, and Val 39). In combination with the side chains of Val 46 and Val 54 from helix α3 and Ile 72 from strand β2, the lid creates a chemical environment within the active site that is largely hydrophobic with the exception of the residues involved in nucleophilic/general base catalysis or water activation. The hydrophobic residues of the lid and those lining the bottom of the active site are well conserved in uncharacterized FlK homologs identified by sequence alignment (*Figure 2.7*), although one group of less closely related homologs lacks helix α2 and the second loop of the lid.

Structural alignment of FlK to other structurally characterized thioesterases with the same active site topology revealed the conspicuous absence of a conserved Asn/Gln whose position in the tertiary structure of FlK approximately corresponds to Val 23 (*Figure 2.8*). In these other thioesterases, the carboxamide of the Asn/Gln is proposed to function in orientation and polarization of the thioester C=O and to assist in departure of the thiolate leaving group [*36, 37, 40*]. Based on the crystal structure of FlK, there does not appear to be another hydrogen bond donor to perform this function. The absence of the Asn/Gln-carbonyl oxygen interaction in FlK could contribute to the lowered catalytic efficiency for acetyl-CoA compared to fluoroacetyl-CoA both by decreasing binding interactions with a shared structural motif as well as by selecting against hydrolysis of the less reactive acetyl-CoA substrate.

Mutagenesis Studies of the Catalytic Residues of FlK. Although the detailed catalytic mechanism for the hot dog-fold thioesterases still requires further elucidation, Ser/Thr, Asp/Glu, His catalytic triads have been proposed and characterized by mutagenesis for many members of this superfamily [*36, 40, 44, 45*]. From their sequence conservation (*FIgure 2.7*) and location in the active site pocket of FlK (*Figure 2.6*B), FlK appears to contain a similar catalytic grouping comprised by Thr 42, Glu 50, and His 76. A previous report has excluded Glu 50 as a catalytic residue based on its location in active site pocket as well as an increase in k_{cat} observed for the E50A mutant, leading to a proposal of an alternative catalytic triad comprised of Thr 42, His 76, and an active site water [*48*]. However, the positioning of Glu 50 on the third turn of the hot dog helix (α3) is consistent with the position of the Asp/Glu in hTHEM2, 4-HBA-CoA thioesterase, and the phenylacetyl-CoA thioesterase from *E. coli* (PaaI) [*37*] (*Figure 2.9*) and is within hydrogen-bonding distance of the catalytic Thr.

To test whether Glu 50, His 76, and Thr 42, are involved in FlK catalysis, we measured the kinetic constants for the E50Q, H76A, T42A, T42S, and T42C mutants (*Table 2.2*). On the basis of mutagenesis studies on several members of the superfamily, the Asp/Glu and Ser/Thr residues have both been implicated as potentially involved in assisting water attack at the carbonyl [*36, 40, 44, 45*]. Mutation of Glu 50 in FlK to Gln resulted in a significant 3,000-fold decrease in k_{cat}, which is consistent with an important role for Glu 50 in catalysis. However, the rate of hydrolysis of fluoroacetyl-CoA catalyzed by FlK-E50Q remains over 2,000-fold higher that of the pseudo first-order rate of the uncatalyzed reaction at pH 7.6. We should note at this time that the E50Q mutant could only be solubly expressed as a MBP fusion protein. FlK-E50Q remained soluble after removal of the MBP tag, however a relatively large proportion (~60%) of the mutant eluted from a size-exclusion column in the void volume, suggesting the formation of a soluble aggregate. After protein concentration, we verified that the fraction corresponding to the dimer peak did not revert to soluble aggregate by size-exclusion chromatography (*Figure 2.9*).

Replacement of His 76 with Ala resulted in a substantial 10^5-fold decrease in k_{cat} to a rate acceleration of only 20-fold over the pseudo-first order rate for the uncatalyzed reaction. In

Table 2.2 *Kinetic constants measured for fluoroacetyl-CoA and acetyl-CoA hydrolysis by wild-type and mutant FlKs.*

	fluoroacetyl-CoA			*acetyl-CoA*		
	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
wild-type	390 ± 20	8 ± 1	(5 ± 1) x 10^7	0.06 ± 0.01	2100 ± 500	(3 ± 1) x 10^1
E50Q	0.13 ± 0.01	14 ± 1	(9 ± 1) x 10^3			
H76A	0.04 ± 0.002	4 ± 1	(1.0 ± 0.2) x 10^4	0.004 ± 0.0001	120 ± 10	(3 ± 0.3) x 10^1
T42A	0.45 ± 0.09	110 ± 40	(4 ± 2) x 10^3			
T42S	19 ± 4	60 ± 20	(3 ± 1) x 10^5	0.01 ± 0.001	61 ± 9	(1.6 ± 0.3) x 10^2
T42C	11 ± 2	360 ± 10	(3 ± 1) x 10^4	0.01 ± 0.001	390 ± 90	(1.2 ± 0.6) x 10^1
F33A	27 ± 4	1500 ± 400	(1.8 ± 0.5) x 10^4	nd	nd	
F36A	270 ± 20	710 ± 80	(3.7 ± 0.5) x 10^5	0.1 ± 0.02	3000 ± 1000	(2.9 ± 0.9) x10^1
V23A	150 ± 30	300 ± 150	(5 ± 3) x 10^5	0.01 ± 0.001	900 ± 100	(1.1 ± 0.2) x 10^1
L26A	6.7 ± 0.01	120 ± 20	(5.7 ± 0.8) x 10^4	nd	nd	

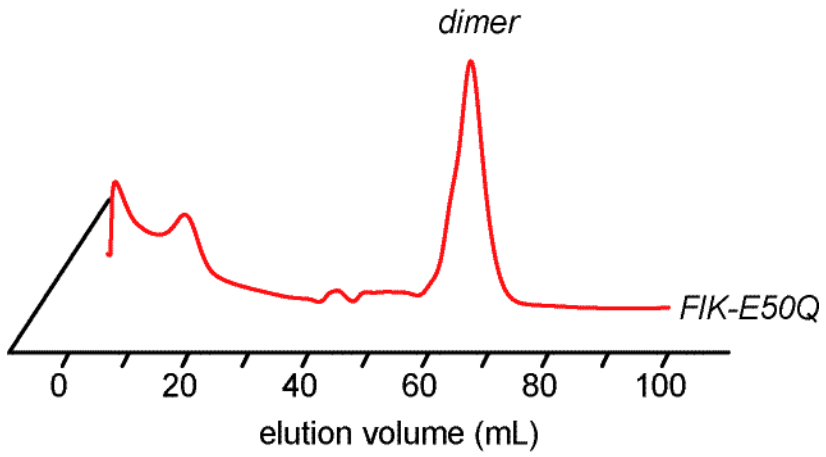


Figure 2.9. *Superdex 75 chromatogram of FlK-E50Q after purification and concentration. The soluble aggregates elute at the void volume of the column between 40-45 mL. The dimer was chromatographed after storage at 4ºC for several days followed by storage at -80ºC and found to remain stable.*

comparison, alanine replacement of His 65 in 4-HBA-CoA thioesterase [*36*] and of His 56 in hTHEM2 [*40*] resulted in only 4,000-fold and 2-fold reductions in k_{cat}, respectively. The effect of this mutation on the overall rate of the reaction is much greater than the effect of the E50Q mutation and may point towards a more important role for His 76 in FlK than in other characterized members of the superfamily. Consistent with the observed defect in k_{cat}, FlK-H76A was not competent to confer fluoroacetate resistance in a heterologous expression system (*Figure 2.1*). As discussed previously, His 76 is located at the dimer interface, but it is not within hydrogen-bonding distance of Glu 50. Based on a homology model of the FlK-substrate complex (*Figure 2.10*), His 76 appears to be positioned closer to the thioester sulfur atom (*Figure 2.6*B). Thus, it is possible that His 76 helps to replace the function of the conserved Asn/Gln found in other hot dog-fold thioesterases by assisting departure of the thiolate leaving group without carbonyl polarization in FlK. Because fluoroacetyl-CoA is more activated towards nucleophilic attack, this change in mechanism could contribute to the large observed difference in the k_{cat} parameter between fluoroacetyl-CoA and acetyl-CoA.

We next turned our attention to examining the role of Thr 42 in catalysis by substitution with Ala, which led to a 900-fold reduction in k_{cat}. Although the k_{cat}/K_M value measured for the T42A mutant is similar to the E50Q mutant, the 15-fold increase in K_M may indicate that this residue participates in both catalysis and recognition of the substrate. Due to speculation that FlK might achieve its selectivity by using a secondary alcohol nucleophile [*28, 48*], we also characterized the T42S and T42C mutants to further probe the role of nucleophilicity or basicity in the FlK reaction. The T42S and T42C mutants show 20- and 35- fold decreases in k_{cat}, respectively, but differ with regard to their response in k_{cat}/K_M. As k_{cat}/K_M is an order of magnitude lower for the T42C mutant, it is likely that the structural perturbations at the active site are greater in this mutant and do not allow us to conclusively determine the trend in reactivity for Thr 42 in catalysis.

To further probe the role of Thr 42, we compared the kinetic parameters with respect to both fluoroacetyl-CoA and acetyl-CoA for the T42S and T42C mutants (*Table 2.2*). In both mutants, the relative decrease in k_{cat} for acetyl-CoA was smaller than the decrease in k_{cat} observed with respect to fluoroacetyl-CoA. However, the most striking effect of these mutations is that neither

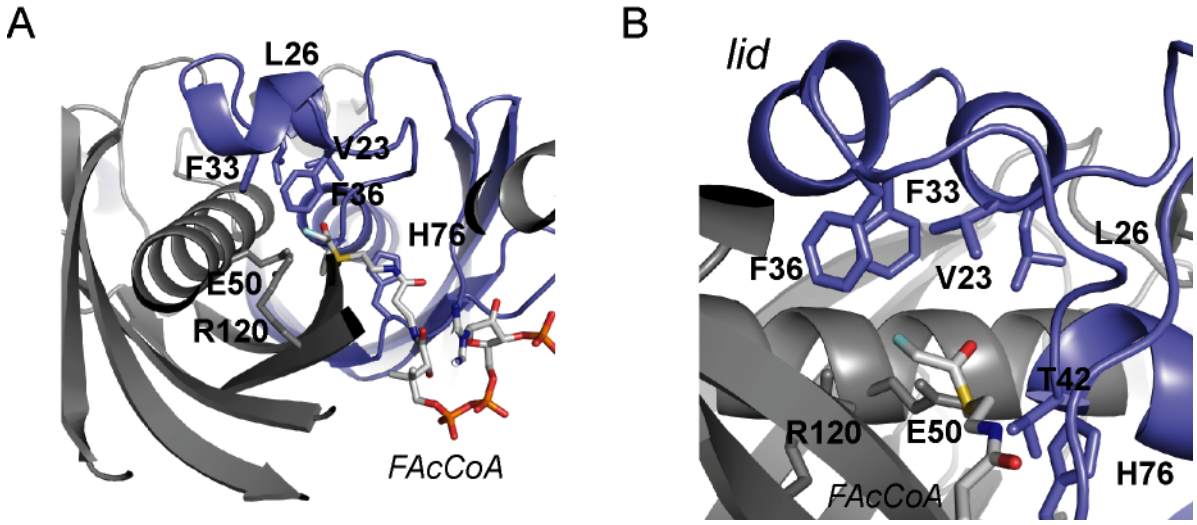


Figure 2.10. *Model of the FlK-fluoroacetyl-CoA complex. (A) Fluoroacetyl-CoA was modeled into the apo-FlK active site based on structural alignments (Figure 2.8) using Coot and is found to be consistent with available ligand-bound structures of hot-dog-fold thioesterases. Chain A is colored in blue and chain B is colored in gray. For clarity, only the modeled fluoroacetyl-CoA ligand is colored by atom (carbon, gray; nitrogen, blue; oxygen, red; fluorine, light green; sulfur, yellow). (B) A view of fluoroacetyl-CoA at the active site.*

the T42S nor the T42C mutant shows a substantial difference between fluoroacetyl-CoA and acetyl-CoA in the K_M parameter, within error. Interestingly, the K_M value for acetyl-CoA in the T42S mutant is within range of other characterized hot dog-fold thioesterases for their substrates. If we assume that the same kinetic mechanism is valid for both fluoroacetyl-CoA and acetyl-CoA, comparison of K_M values implies that discrimination in substrate binding has been lost in both the T42S and T42C mutants. In terms of k_{cat}/K_M, the discrimination between the fluorinated and nonfluorinated substrates is decreased from 10^6-fold in the wild-type enzyme to 200-fold for the T42S mutant and to 250-fold for the T42C mutant, and the observed discrimination is accounted for almost entirely by the difference in k_{cat} between the two substrates. One possible explanation for this is that the removal of steric bulk at this position in the T42S and T42C mutants could allow for structural rearrangements that introduce new interactions with acetyl-CoA, enhancing FlK's affinity for this substrate. In support of this hypothesis, it is interesting to note that the recently reported structure of FlK-T42S in complex with acetyl-CoA shows increased flexibility at the active site as determined by the observed population of two different rotamers of Ser 42 in the crystal structure [*48*]. Although no new interactions with acetyl-CoA were found in the structure, the observed flexibility raises the possibility that a second hydrogen bond donor is accessible to the acetyl-CoA carbonyl group that could account for the observed decrease in K_M.

Mutagenesis studies of the putative substrate binding pocket of FlK. We next turned our attention to examining the role of the hydrophobic lid residues identified in the putative substrate-binding pocket (*Figure 2.6*C). Based on structural alignments with ligand-bound forms of 4-HBA-CoA thioesterase from *Arthrobacter* sp. [*36*] and hTHEM2 [*40*] (*Figure 2.8*), the fluorine atom would be located in a largely hydrophobic environment surrounded by Phe 33, Phe 36, Val 23, Leu 26, Val 54, and Pro 37 (*Figure 2.10*), which are all conserved or conservatively substituted with bulky, hydrophobic residues in closely related FlK homologs of unknown

function (*Figure 2.7*). Mutation of Phe 36 to alanine led to a 100-fold decrease in k_{cat}/K_M with very little perturbation of k_{cat} for fluoroacetyl-CoA (*Table 2.2*). For the F33A, V23A, and L26A mutants, the magnitude in the decrease of k_{cat}/K_M (F33A, 2,800-fold; V23A, 100-fold; L26A, 900-fold) was larger than the impact on k_{cat} for fluoroacetyl-CoA (F33A, 14-fold decrease; V23A, 2.5-fold decrease; L26, 58-fold decrease). Interestingly, the kinetic constants with respect to acetyl-CoA for both the F36A and V23A mutants remained relatively unchanged, which is consistent with an important function for these residues in discrimination between the fluorinated and nonfluorinated substrates. Notably, all of these residues are part of the FlK-specific lid structure. Although the lid residues have been found to be conserved in several uncharacterized homologs, Phe 36 is unique and not found in other sequences published to date (*FIgure 2.7*).

We next looked to characterizing the role of hydrogen bonding and dipolar interactions in the substrate binding pocket. Arg 120 is the only polar residue found in the active site pocket besides the catalytic triad and was previously proposed to function as a hydrogen bond donor to the fluorine atom [*48*]. Mutation of Arg 120 to Ala, Gln, and Lys all led to formation of highly insoluble protein with no dimeric protein observed by size-exclusion chromatography, which implies that the 1-point hydrogen bond that this highly conserved residue shares with Glu 50 in apo FlK is important to the structural integrity of the dimer. Except for the conservative E50Q mutation, all other mutants tested of Glu 50 and Arg 120 resulted in insoluble protein that could not be purified in dimeric form. In addition to any enthalpic contributions of C-F---H-N dipolar interactions, it is also possible that only substrates that can replace the hydrogen-bonding interaction of Arg 120 with Glu 50, such as fluoroacetyl-CoA, can maintain the dimer interface and bind well in the active site pocket. Attempts to engineer a hydrogen bond donor for the carbonyl oxygen of the acyl-CoA substrate as observed in other members of the superfamily by replacement of Val 23 with Asn and Gln yielded protein with little or no activity toward either fluoroacetyl-CoA or acetyl-CoA. These results suggest that the specifics of active site packing and configuration may have diverged between FlK and other superfamily members without allowing us to directly address the role of the putative missing hydrogen bond.

In light of the kinetic constants measured for fluoroacetyl-CoA and acetyl-CoA for the F36A, F33A, and V23A mutants, a trend emerges in which removal of hydrophobic groups near the substrate binding pocket decreases the k_{cat}/K_M for fluoroacetyl-CoA by 2-3 orders of magnitude while leaving the k_{cat}/K_M for acetyl-CoA relatively unchanged. The magnitude of this decrease is consistent with both the difference in K_M with respect to fluoroacetyl-CoA and acetyl-CoA measured for wild-type FlK (*Table 2.2*) and the results from competitive inhibition studies (*Figure 2.4*). Although these data suggest that the hydrophobic groups are involved in specific recognition of the fluorine atom by FlK, a survey of the chemical environments of the fluorine atom in protein crystal structures with a bound fluorinated ligand has previously shown that binding of a fluorine atom in a lipophilic pocket does not typically confer advantages in terms of binding affinity on this scale [*46*]. It has been observed, however, that fluorinated compounds often have negative entropies of aqueous solvation due to the tendency of water molecules to order around the hydrophobic C-F unit [*42*]. In particular, monofluorination at the α-carbon of carbonyl-containing compounds typically increases their lipophilicity as determined by octanol-water and nonanol-water partition coefficients [*51*]. The hydrophobicity of the fluoroacetyl-CoA binding pocket could thus promote fluoroacetyl-CoA binding over acetyl-CoA binding by providing a more lipophilic environment into which fluoroacetyl-CoA can partition, resulting in the release of water and an entropic driving force for binding. Furthermore, any enthalpic dipolar interactions with the C-F unit would be favored in a more apolar environment because organic

fluorine does not compete well with stronger hydrogen bond acceptors. As it therefore seems inconsistent with previous studies that these residues are utilized mostly to make enthalpically favorable interactions with the fluorine substituent in the FlK binding pocket, we propose that the apolar, aromatic nature of the substrate binding pocket provides an entropic driving force for substrate binding by promoting release of ordered water molecules from the C-F unit.

Crystal structure of the FlK-fluoroacetate complex. We solved the structure of the FlK-fluoroacetate complex at 2.3 Å resolution to further investigate fluoroacetyl-CoA binding to FlK (*Table 2.3*). Six fluoroacetate molecules were trapped in the structure (*Figure 2.11*), all located either in the putative active site or along the channel where CoA is expected to bind based on homology (*Figure 2.10*). The fluoroacetate molecules bound closest to the putative catalytic residues are bound in different orientations at the two active sites, suggesting that their binding may differ from how the fluoroacetyl unit is bound in the substrate complex. Strikingly, a large conformational change occurs in the lid region of one active site in the product complex, involving the swinging aside of the unique Phe 36 and twisting of Phe 33 to open a channel in the protein betwen the active site and solvent that is blocked by Phe 36 in the apo structure (*Figure 2.11*). These movements are consistent with recently reported structures of the wild-type FlK bound to fluoroacetate and pantothenate-based substrate analogs [*48*]. Our structure is distinctive, however, in that we observe a 3.3 Å interaction between a fluoroacetate fluorine atom and Phe 33, suggesting a possible mechanism for disruption of the edge-face interaction between Phe 33 and Phe 36 and gate opening. The orientation of the fluoroacetyl group that leads to the disruption would not be possible in the substrate complex as it would force the CoA chain to occupy the same space as helix α3 from either chain A or chain B. This observation raises the possibility that the Phe 33-Phe 36 gate could be relevant not only for maintaining the hydrophobicity of the active site for fluoroacetyl-CoA binding but also potentially for release of fluoroacetate. The movement of Phe 33 and Phe 36 could facilitate fluoroacetate release by either opening up a channel though which fluoroacetate could exit, or by allowing the entrance of water into the active site and the displacement and hydration of fluoroacetate. We were also able to determine the structure of a fluoroacetyl-CoA soaked crystal at 2.0 Å resolution. Although electron density was clearly observed in the active site, it could not be unambiguously assigned to either substrate or products. The same "open gate" conformation was trapped in this structure, which supports the hypothesis that the observed conformational changes are relevant to the normal catalytic cycle of FlK.

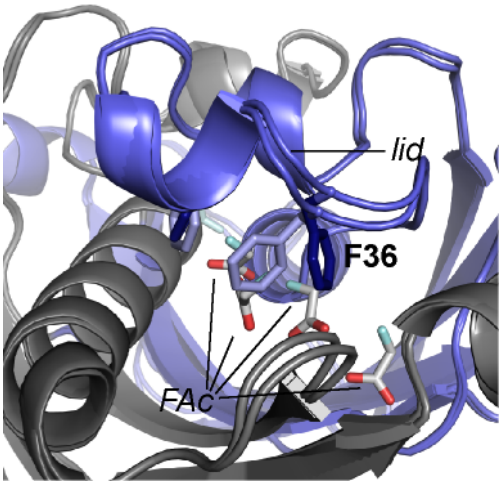


Figure 2.11. *Conformational changes in the fluoroacetate complex of FlK. Chain A of both the unliganded and fluoroacetate-bound structures is shown in blue and chain B of both structures is shown in gray. Phe 36 in shown in light blue for the closed conformation and dark blue for the open configuration. A total of six fluoroacetate molecules were found in the active site (carbon, gray; oxygen, red; fluorine, light green).*

Crystal structures of FlK-F36A and of the FlK-F36A product complex. To explore the role of ordered water and Phe 36 in substrate discrimination, we solved the crystal structures of

***Table 2.3**. Data collection and refinement statistics for FIK-F36A structures.*

	FIK-F36A	FIK-F36A-products
Data Collection		
X-ray source	ALS 8.3.1	ALS 8.3.1
Space group	C2	C2
Cell dimensions		
a, b, c (Å)	141.9, 89.3, 71.4	141.0, 88.6, 70.9
α, β, γ (°)	90.0, 118.0, 90.0	90.0, 118.0, 90.0
Wavelength	1.116	1.116
Resolution*	19.75-2.30 (2.48-2.30)	72.48-1.90 (1.90-1.95)
R_{merge}*	12.6% (40.3%)†	8.9% (66.3%)
I/σI*	6.0 (2.5)	13.37 (2.22)
Completeness*	99.7% (99.4%)	95.9% (90.9%)
Redundancy*	4.1 (4.2)	4.3 (4.1)
Refinement		
Resolution	19.75-2.30	72.48-2.00
No. reflections	34432	48041
R_{work}/R_{free}	0.211/0.241	0.212/0.250
No. atoms	6660	6351
Non-solvent	6092	6081
Solvent	563	270
B-factors		
Protein	33.7	32.6
Water	38.2	40.1
rmsd		
Bond lengths (Å)	0.018	0.006
Bond angles (°)	1.64	1.03

* Values in parenthesis are for the highest resolution shell.
†Due to high non-crystallographic symmetry, R_{pim} is reported.

apo-FlK-F36A at 2.3 Å resolution and of the FlK-F36A product complex at 2.0 Å resolution (*Figure 2.12*, *Table 2.3*). The overall folds of both structures do not appear to be significantly perturbed by the F36A mutation and are in good agreement with the corresponding wild-type structures with a core r.m.s.d. of 0.5 Å for both apo-FlK-F36A and the product complex of FlK-F36A. Notably, the channel that is normally occluded by Phe 36 is unblocked in the apo FlK-F36A mutant structure and correspondingly appears to have a more highly aquated active site compared to wild-type enzyme (*Figure 2.12*). In both apo structures, water molecules occupy the space that would typically be occupied with the substrate based on homology. However, additional ordered water molecules were observed at the active site of FlK-F36A and found to form a continuous hydrogen bonding network with the external waters that are usually blocked from the active site by Phe 36. One of these internal waters makes a 2.9 Å hydrogen bond with Arg 120, potentially competing with a favorable interaction with the fluorine atom. In contrast, three ordered water molecules were observed just outside of the Phe 36 "gate" in the wild-type FlK structure that were unable to interact with water molecules within the active site. We propose that the exclusion of water from the active site of the wild-type enzyme prevents solvation of residues that would otherwise make favorable interactions with the fluorine atom and promotes the release of ordered water molecules from the C-F unit, providing an entropic driving force for fluoroacetyl-CoA binding. The greater extent of the water network at the active site of FlK-F36A compared to wild-type supports the hypothesis that its large k_{cat}/K_M defect with respect to fluoroacetyl-CoA arises from a diminished ability to exclude water from the active site. We also solved the crystal structure of the FlK-F36A product complex in order to further examine the role of solvation of the fluorine atom itself in the increased kcat/KM of the F36A mutant (*Figure 2.13*). A 2.5 σ peak was observed in the F_{obs}-F_c solvent/ligand omit electron density map for fluoroacetate and for the first 8 atoms of the CoA chain. The orientation and position of the ligands differ significantly from the positions of pantothenate-derived fluoroacetyl-CoA analogs bound in two recently reported structures of FlK [*48*]. Dias et al. have speculated that the ligands are bound in positions that are irrelevant to catalysis due to the absence of the nucleotide moiety of CoA, which is believed to be important in correct binding of the substrate. However, the positions of fluoroacetate and CoA in our structure are in agreement with the positioning of substrates and products in structural homologs of FlK (Figure 2.9) and onfirm the location of the fluorine atom in the hydrophobic pocket, which is consistent with kinetic data implicating these lid residues in substrate recognition. The product-bound structure also reveals that the fluorine atom is indeed within 3.1 Å of the Arg 120 side chain and within 3.7 Å of the Gly 69 backbone amide. Consistent with previously identified fluorine arginine interactions, the plane of the C-F bond is oriented perpendicular to the plane of the guanidinium side chain in the product structure rather than in a linear hydrogen bond geometry. While hydrogen bonding with organic fluorine is considered rare or weak [*46, 52, 53*], these specific dipolar C-F…H-N interactions are abundant in protein structures with bound fluorinated ligands and are also thought to explain the "fluorophilic" character of arginine [*46, 54*]. Indeed, 32 structures were found in a search of the Protein Data Bank for C-F interactions with the guanidinium side chain [*46*]. As no analogous interactions would be possible for acetyl-CoA, these fluorine-specific interactions could make enthalpic contributions to the observed substrate selectivity. In addition to these interactions, the product structure reveals a 3.3 Å hydrogen bond between the backbone NH of Thr 42 and the thioester C=O and also shows that the second hydrogen bond equivalent to that provided by an Asn/Gln residue in structural homologs isindeed missing. In this structure, the fluorine atom appears to be solvated as it is within 2 Å of

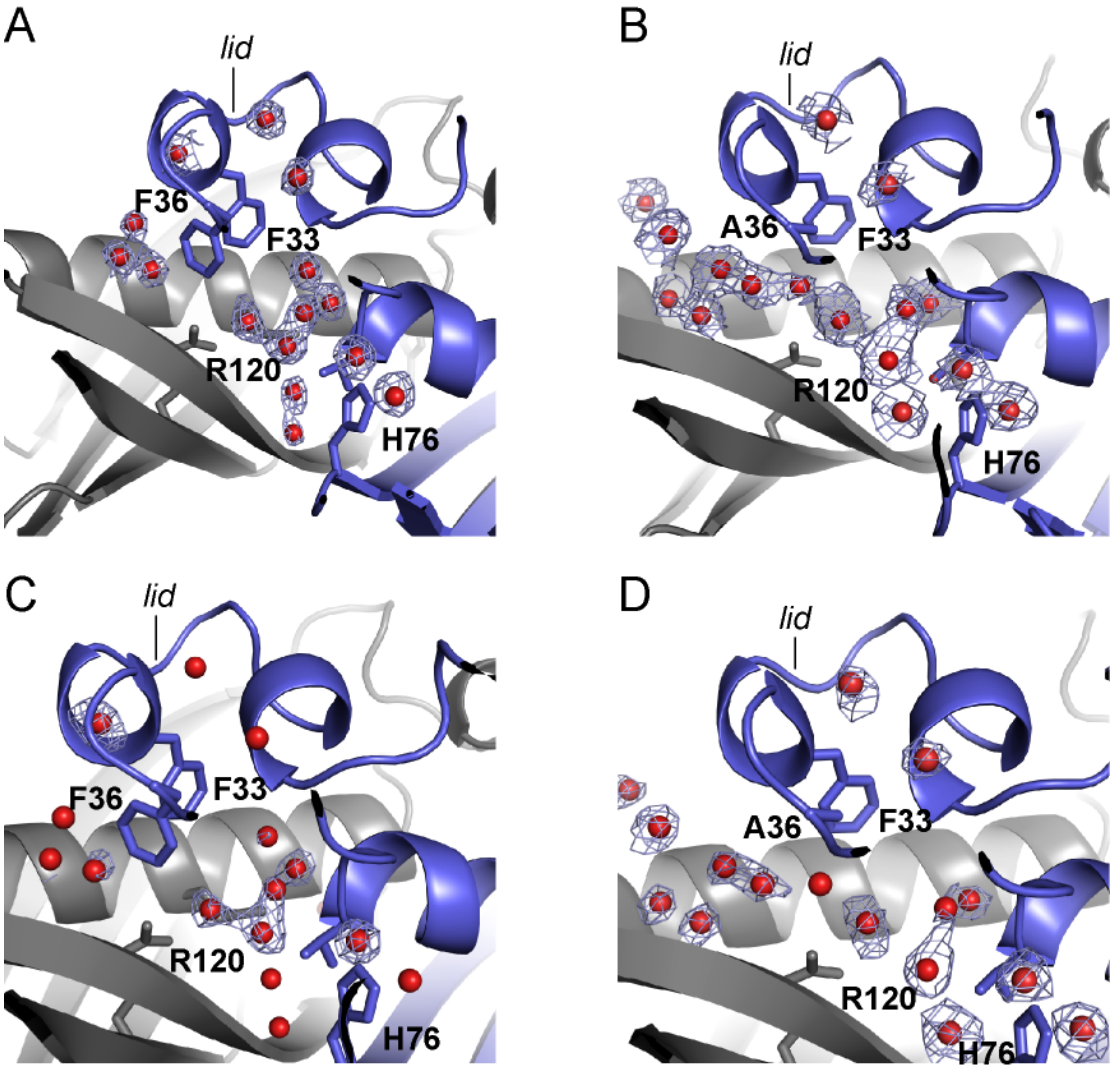


Figure 2.12. *Accessibility of the active site to water in FIK and FIK-F36A. Electron density maps are shown in light blue, chain A is colored in blue, chain B is colored in gray, and water molecules are shown as red spheres. (A) $2F_o$-F_c map contoured at 0.8 σ for water molecules at the active site of wild-type FIK. (B) 2Fo-Fc map contoured at 0.8 σ for water molecules at the active site of FIK-F36A. (C) F_o-F_c solvent omit map contoured at 2.5 σ for wild-type FIK overlaid with the refined FIK structure. (D) F_o-F_c solvent omit map contoured at 2.5 σ for FIK-F36A overlaid with the refined FIK-F36A structure.*

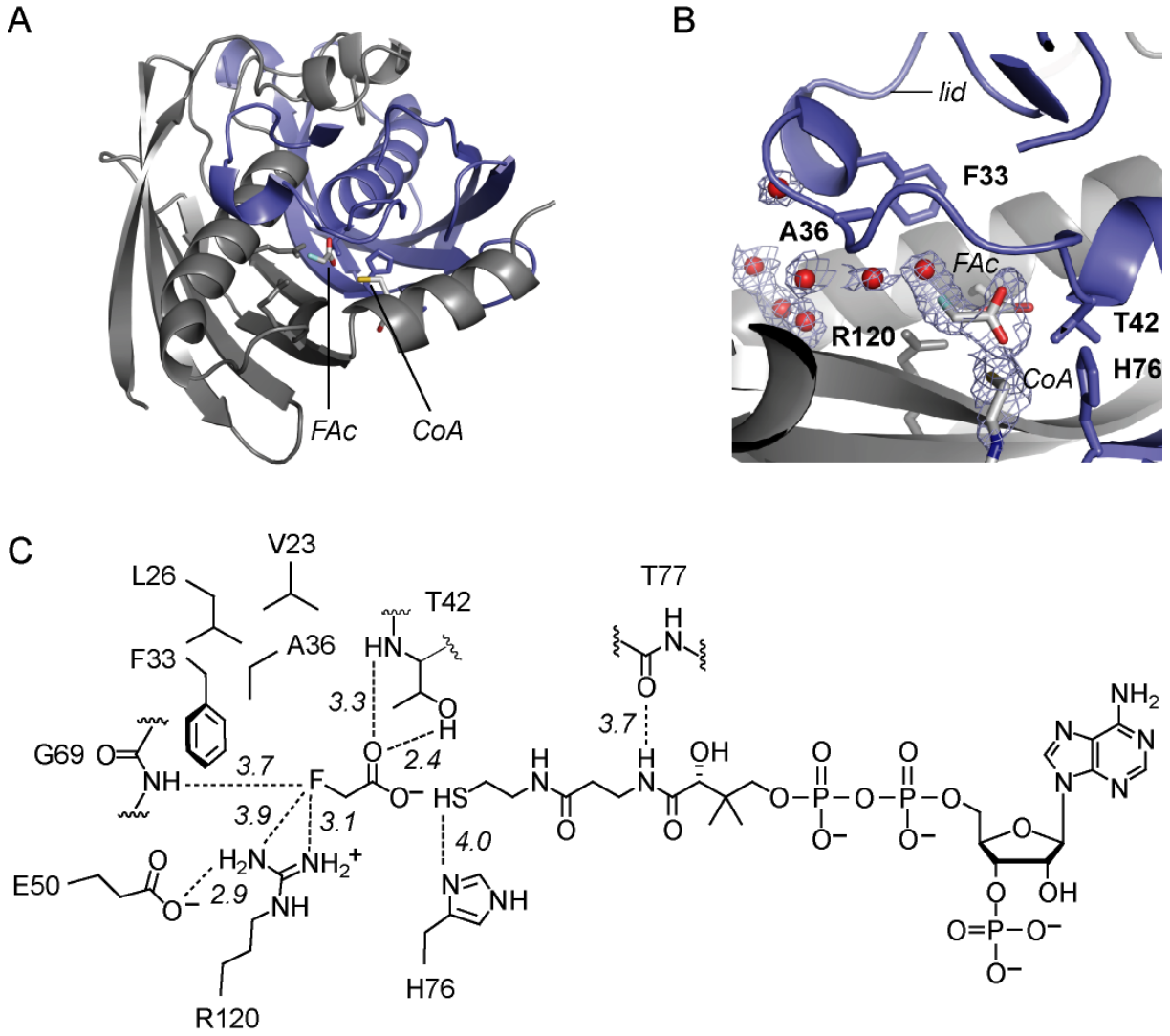


Figure 2.13 *FIK-F36A product complex. Chain A is colored in blue and chain B is colored in gray. For clarity, only the fluoroacetate and CoA ligands are colored by atom (carbon, gray; nitrogen, blue; oxygen, red; fluorine, light green; sulfur, yellow). (A) A view of the product bound at the FIK-F36A active site shown in the context of the overall fold. (B) F_o-F_c solvent/ligand omit map contoured at 2.5 σ overlaid with the refined FIK-F36A product complex structure. (C) ChemDraw figure showing interactions observed between FIK-F36A and the products.*

a chain of hydrogen bonded water molecules that would likely be excluded from the substrate binding pocket by Phe 36 in the closed conformation of FlK, potentially explaining the large observed defect in k_{cat}/K_M for this mutant. The product complex structure is thus consistent with a model in which the active site lid and Phe 36 in particular aid in discrimination between fluoroacetyl-CoA and acetyl-CoA by actively excluding water from the active site.

2.4 Conclusions

We have demonstrated that FlK is a fluoroacetyl-CoA-specific thioesterase that exhibits a 10^6-fold higher catalytic efficiency for fluoroacetyl-CoA compared to acetyl-CoA. Based on experiments in a heterologous *E. coli* host, the *in vitro* selectivity of FlK also appears to be relevant *in vivo* and is consistent with the estimated intracellular concentration of acetyl-CoA [*55*], which is lower than the measured K_M for acetyl-CoA. Based on kinetic measurements, the remarkable discrimination between the fluorinated and nonfluorinated substrates by FlK arises from both differences in substrate recognition (k_{cat}/K_M) and chemical reactivity (k_{cat}) related to the fluorine substituent. Using x-ray crystallographic and biochemical studies on wild-type and mutant FlKs, we have identified several components of this selectivity and explored their role in preferential hydrolysis of fluoroacetyl-CoA.

We propose that FlK has evolved a mode of substrate recognition that is distinct from the more promiscuous hot dog-fold thioesterases by minimizing interactions with the carbonyl unit common to all acyl-CoAs to select against acetyl-CoA while recouping the resulting decrease in binding energy through C-F specific properties of fluoroacetyl-CoA. This working hypothesis is supported by structural information that indicates that a key Asn/Gln hydrogen bond to the thioester carbonyl is missing in conjunction with biochemical data comparing fluoroacetyl-CoA and acetyl-CoA hydrolysis in the T42S and T42C mutants. Our data suggest that the specificity of FlK in binding fluoroacetyl-CoA is aided in part by hydrophobic interactions with the lipophilic binding pocket and by C-F---H-N dipolar interactions between the fluorine atom and Arg 120. Because neither of these enthalpic contributions would be expected to account for the observed 10^6-fold selectivity between the fluorinated and nonfluorinated substrates given the existing literature on protein/organofluorine binding affinities, we propose that entropic contributions related to ordered water play a larger role in fluorine-based substrate discrimination. In support of this model, we have biochemically and structurally characterized a mutant of a unique phenylanine (Phe 36) on the hydrophobic "lid" of FlK and have shown that defects in water exclusion from the active site lead to a 100-fold decrease in fluoroacetyl-CoA binding affinity without affecting k_{cat}. We have also identified key active site residues consisting of Thr 42, Glu 50, and His 76, and propose that His 76 may be able to assist in CoA departure with the more activated fluoroacetyl-CoA substrate in the absence of typical interactions that serve to polarize the thioester carbonyl.

2.5 Acknowledgments

We thank Dr. Dipali Sashital for protein purification and crystallization advice and George Meigs (Beamline 8.3.1, Advanced Light Source, Lawrence Berkeley National Laboratory) for assistance with x-ray data collection.

2.6 References

1. M. B. Chenoweth, Monofluoroacetic acid and related compounds, *J. Pharmacol. Exptl. Therap.* **1949**, *97*, 383-423.
2. E. K. Rowinsky, L. A. Cazenave and R. C. Donehower, Taxol: a novel investigational antimicrotubule agent, *J. Natl. Cancer Inst.* **1990**, *82*, 1247-1259.
3. P. Buffa and R. A. Peters, The *in vivo* formation of citrate induced by fluoroacetate and its signficance, *J. Physiol.* **1950**, *110*, 488-500.
4. P. Buffa, R. A. Peters and R. W. Wakelin, Biochemistry of fluoroacetate poisoning. Isolation of an active tricarboxylic acid fraction from poisoned kidney homogenates, *Biochem. J.* **1951**, *48*, 467-477.
5. J. F. Morrison and R. A. Peters, Biochemistry of fluoroacetate poisoning: The effect of fluorocitrate on purified aconitase, *Biochem. J.* **1954**, *58*, 473-479.
6. D. D. Clarke, Fluoroacetate and fluorocitrate: Mechanism of action, *Neurochem. Res.* **1991**, *16*, 1055-1058.
7. M. Brandt, S. Evans, J. Mendes-Morao and B. Chappell, Fluorocitrate inhibition of aconitate hydratase and the tricarboxylate carrier of rat liver mitochondria, *Biochem. J.* **1973**, *134*, 217-224.
8. J. J. Villafranca and E. Platus, Fluorocitrate inhibition of aconitase. Reversibility of the inactivation, *Biochem. Biophys. Res. Commun.* **1973**, *55*, 1197-1207.
9. R. Z. Eanes and E. Kun, Inhibition of liver aconitase isozymes by (-)-erythro-fluorocitrate, *Mol. Pharmacol.* **1974**, *10*, 130-139.
10. T. A. Kent, M. H. Emptage, H. Merkle, M. C. Kennedy, H. Beinert and E. Munck, Mossbauer studies of aconitase. Substrate and inhibitor binding, reaction intermediates, and hyperfine interactions of reduced 3Fe and 4Fe clusters, *J. Biol. Chem.* **1985**, *260*, 6871-6881.
11. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme-inhibitor complex, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13699-13703.
12. J. Marais, Isolation of the toxic principle "K cymonate" from "Gifblaar" *Dichapetalum cymosum*, *J. Vet. Sci. Anim. Ind.* **1943**, *18*, 203-206.
13. B. Vickery and M. Vickery, Fluoride metabolism in *Dichapetalum toxicarium*, *Phytochem.* **1972**, *11*, 1905-1909.
14. R. Hall, The distribution of organic fluorine in some toxic tropical plants, *New Phytologist* **1972**, *71*, 855-871.
15. L. Twigg and D. King, The impact of fluoroacetate-bearing vegetation on native Australian fauna: A review, *Oikos* **1991**, *61*, 412-430.
16. D. B. Harper and D. O'Hagan, The fluorinated natural products, *Nat. Prod. Rep.* **1994**, *11*, 123-133.
17. M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch and E. Inamine, Biosynthesis of fluorothreonine and fluoroacetic acid by the thienamycin producer, *Streptomyces cattleya*, *J. Antibiot. (Tokyo)* **1986**, *39*, 259-265.
18. C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan and J. H. Naismith, Crystal structure and mechanism of a bacterial fluorinating enzyme, *Nature* **2004**, *427*, 561-565.

19. D. O'Hagan and D. Harper, Fluorine-containing natural products, *J. Fluor. Chem.* **1999**, *100*, 127-133.

20. C. D. Murphy, C. Schaffrath and D. O'Hagan, Fluorinated natural products: The biosynthesis of fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*, *Chemosphere* **2003**, *52*, 455-461.

21. H. Deng, D. O'Hagan and C. Schaffrath, Fluorometabolite biosynthesis and the fluorinase from *Streptomyces cattleya*, *Nat. Prod. Rep.* **2004**, *21*, 773-784.

22. R. Benveniste and J. Davies, Aminoglycoside antibiotic-inactivating enzymes in actinomycetes similar to those present in clinical isolates of antibiotic-resistant bacteria, *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70*, 2276-2280.

23. E. Cundliffe, Mechanism of resistance to thiostrepton in the producing-organism *Streptomyces azureus*, *Nature* **1978**, *272*, 792-795.

24. J. F. Fierro, C. Hardisson and J. A. Salas, Involvement of cell impermeability in resistance to macrolides in some producer streptomycetes, *J. Antibiot. (Tokyo)* **1988**, *41*, 142-144.

25. A. S. Thiara and E. Cundliffe, Cloning and characterization of a DNA gyrase B gene from *Streptomyces sphaeroides* that confers resistance to novobiocin, *EMBO J.* **1988**, *7*, 2255-2259.

26. E. Cundliffe, How antibiotic-producing organisms avoid suicide, *Annu. Rev. Microbiol.* **1989**, *43*, 207-233.

27. C. Mendez and J. A. Salas, ABC transporters in antibiotic-producing actinomycetes, *FEMS Microbiol. Lett.* **1998**, *158*, 1-8.

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

29. J. M. Rouillard, W. Lee, G. Truan, X. Gao, X. Zhou and E. Gulari, Gene2Oligo: Oligonucleotide design for in vitro gene synthesis, *Nucleic Acids Res.* **2004**, *32*, W176-180.

30. J. E. Tropea, S. Cherry and D. S. Waugh, Expression and purification of soluble His_6-tagged TEV protease, *Methods Mol. Biol.* **2009**, *498*, 297-307.

31. W. Kabsch, Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants, *J. Appl. Crystallogr.* **1993**, *212*, 916-924.

32. A. J. McCoy, R. W. Grosse-Kunstsleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, Phaser crystallographic software, *J. Appl. Crystallogr.* **2007**, *40*, 658-674.

33. P. D. Adams, R. W. Grosse-Kunstsleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger, PHENIX: Building new software for automated crystallographic structure determination, *Acta Crystallogr. D* **2002**, *58*, 1948-1954.

34. A. Perrakis, R. Morris and V. S. Lamzin, Automated protein model building combined with iterative structure refinement, *Nat. Struct. Biol* **1999**, *6*, 458-463.

35. P. Emsley and K. Cowtan, Coot: Model-building tools for molecular graphics, *Acta Crystallogr. D* **2004**, *60*, 2126-2132.

36. J. B. Thoden, Z. Zhuang, D. Dunaway-Mariano and H. M. Holden, The structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU, *J. Biol. Chem.* **2003**, *278*, 43709-43716.

37. F. Song, Z. Zhuang, L. Finci, D. Dunaway-Mariano, R. Kniewel, J. A. Buglino, V. Solorzano, J. Wu and C. D. Lima, Structure, function, and mechanism of the phenylacetate pathway hot dog-fold thioesterase PaaI, *J. Biol. Chem.* **2006**, *281*, 11028-11038.

38. Z. Cheng, F. Song, X. Shan, Z. Wei, Y. Wang, D. Dunaway-Mariano and W. Gong, Crystal structure of human thioesterase superfamily member 2, *Biochem. Biophys. Res. Commun.* **2006**, *349*, 172-177.

39. Z. Zhuang, F. Song, H. Zhao, L. Li, J. Cao, E. Eisenstein, O. Herzberg and D. Dunaway-Mariano, Divergence of function in the hot dog fold enzyme superfamily: the bacterial thioesterase YciA, *Biochemistry* **2008**, *47*, 2789-2796.

40. J. Cao, H. Xu, H. Zhao, W. Gong and D. Dunaway-Mariano, The mechanisms of human hotdog-fold thioesterase 2 (hTHEM2) substrate recognition and catalysis illuminated by a structure and function based analysis, *Biochemistry* **2009**, *48*, 1293-1304.

41. D. Chen, R. Wu, T. L. Bryan and D. Dunaway-Mariano, *In vitro* kinetic analysis of substrate specificity in enterobactin biosynthetic lower pathway enzymes provides insight into the biochemical function of the hot dog-fold thioesterase EntH, *Biochemistry* **2009**, *48*, 511-513.

42. J. C. Biffinger, H. W. Kim and S. G. DiMagno, The polar hydrophobicity of fluorinated compounds, *Chembiochem* **2004**, *5*, 622-627.

43. N. Kunishima, Y. Asada, M. Sugahara, J. Ishijima, Y. Nodake, M. Miyano, S. Kuramitsu and S. Yokoyama, A novel induced-fit reaction mechanism of asymmetric hot dog thioesterase PAAI, *J. Mol. Biol.* **2005**, *352*, 212-228.

44. F. Song, Z. Zhuang and D. Dunaway-Mariano, Structure-activity analysis of base and enzyme-catalyzed 4-hydroxybenzoyl coenzyme A hydrolysis, *Bioorg. Chem.* **2007**, *35*, 1-10.

45. T. Yokoyama, K. J. Choi, A. M. Bosch and H. J. Yeo, Structure and function of a *Campylobacter jejuni* thioesterase Cj0915, a hexameric hot dog fold enzyme, *Biochim. Biophys. Acta* **2009**, *1794*, 1073-1081.

46. K. Muller, C. Faeh and F. Diederich, Fluorine in pharmaceuticals: Looking beyond intuition, *Science* **2007**, *317*, 1881-1886.

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

48. M. V. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. Dry, P. F. Leadlay, J. B. Spencer and T. L. Blundell, Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK, *J. Biol. Chem.* **2010**, *285*, 22495-22504.

49. S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, Basic local alignment search tool, *J. Mol. Biol.* **1990**, *215*, 403-410.

50. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.* **2004**, *32*, 1792-1797.

51. B. Smart, Fluorine substituent effects (on bioactivity), *J. Fluor. Chem.* **2001**, *109*, 3-11.

52. J. Howard, V. Hoy, D. O'Hagan and G. Smith, How good is fluorine as a hydrogen bond acceptor?, *Tetrahedron* **1996**, *52*, 12613-12622.

53. J. Dunitz and R. Taylor, Organic fluorine hardly ever accepts hydrogen bonds, *Chem. Eur. J.* **1997**, *3*, 89-98.

54. P. Zhou, J. Zou, F. Tian and Z. Shang, Fluorine bonding—How does it work in protein ligand interactions?, *J. Chem. Inf. Model* **2009**, *49*, 2344-2355.

55. M. Shimazu, L. Vetcher, J. L. Galazzo, P. Licari and D. V. Santi, A sensitive and robust method for quantification of intracellular short-chain coenzyme A esters, *Anal. Biochem.* **2004**, *328*, 51-59.

Chapter 3: *Catalytic control of enzymatic fluorine specificity*

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

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

3.1 Introduction

Living organisms have solved some of the most difficult challenges in catalysis by harnessing the exquisite selectivity and reactivity of enzyme active sites to build new and complex chemical behaviors [*1-5*]. Indeed, the plasticity of enzymes towards evolution of new function has allowed life to flourish in diverse biospheres by conferring a selective advantage to hosts that can demonstrate catalytic innovation and utilize unusual resources. The enzymes that form the molecular basis for these complex chemical phenotypes are a rich source of molecular diversity and provide an experimental platform for the continual search to gain insight into the mechanisms and dynamics of protein and organismal evolution [*6-9*]. One of the salient features of enzymes is their extremely high substrate selectivity, which allows them to choose the correct substrate over the thousands of other small molecules that are present in cells at any given time. Thus, the exploration of substrate specificity and its evolution is key for understanding both enzyme catalysis as well as the expansion of biodiversity at the molecular level [*7-12*].

In this context, a singular chemical trait found in the soil bacterium *Streptomyces cattleya* is its ability to catalyze the formation of carbon-fluorine bonds [*13, 14*]. Fluorine resides at one corner of the periodic table and its unique elemental properties make it highly effective as a design element for drug discovery but also quite challenging for the synthesis of fluorine-containing molecules [*15-17*]. Although fluorine has emerged as a common motif in man-made compounds, including top-selling drugs like atorvastatin and fluticasone, it has been underexploited in nature, with only a few biogenic organofluorine compounds (<20) identified to date despite the thousands of known natural products containing chlorine and bromine (~5,000) [*18, 19*]. In fact, the only fully characterized organofluorine natural products are derived from a single pathway that produces the deceptively simple poison, fluoroacetate. The high toxicity of fluoroacetate arises from its antimetabolite mode of action resulting from the substitution of one hydrogen in acetate, an important cellular building block, with fluorine. As a consequence of this conservative structural change, fluoroacetate can still enter normal acetate metabolism via activation to fluoroacetyl-CoA but then acts as a mechanism-based inhibitor of the tricarboxylic acid (TCA) cycle upon conversion to fluorocitrate and elimination of fluoride at a critical point [*20, 21*] (*Figure 3.1*).

The potency of fluoroacetate as a poison poses an especially difficult substrate selectivity problem at the cellular level for the host organism, where a single fluorine atom must be recognized very specifically over hydrogen in order to clear very low endogenous levels of toxic fluoroacetyl-CoA while simultaneously maintaining the high levels of acetyl-CoA required for cell growth and survival. Consequently, there exists a natural selective pressure on *S. cattleya* to meet this challenge and it has evolved a fluoroacetyl-CoA thioesterase (FlK) that catalyzes the

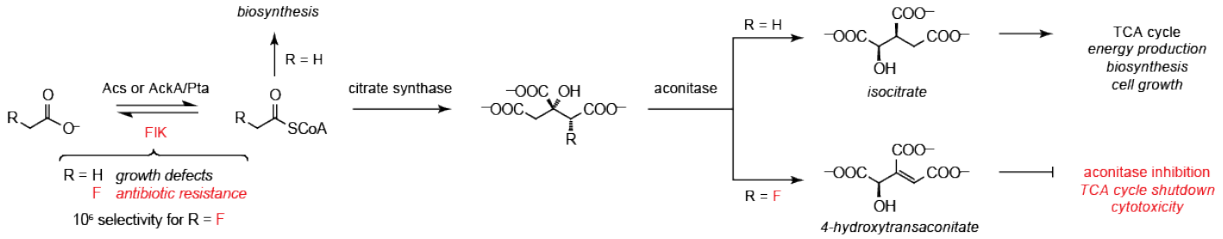


***Figure 3.1**. Molecular basis of fluoroacetate toxicity. FlK catalyzes the hydrolysis of fluoroacetyl-CoA, which reverses the lethal synthesis of an irreversible aconitase inhibitor.*

breakdown of fluoroacetyl-CoA to prevent lethal synthesis of fluorocitrate (*Figure 3.1*) [*22, 23*]. We have recently shown that FlK demonstrates an extremely high 10^6-fold selectivity for hydrolysis of fluoroacetyl-CoA over acetyl-CoA based on a decrease in K_M (10^2) and increase in k_{cat} (10^4) for the fluorinated substrate [*24*]. Although the K_M of FlK with respect to fluoroacetyl-CoA (8 μM) is significantly lower than that for acetyl-CoA (2 mM), it is interesting to note that FlK likely operates near saturation for both substrates *in vivo* given the high intracellular concentration of acetyl-CoA [*25*], with substrate selectivity thus governed by the rate of hydrolysis (k_{cat}) rather than substrate binding (K_D). We thus set out to explore the molecular origin of the catalytic selectivity of FlK with regard to fluorine and now report the existence of an unexpected C_α-deprotonation pathway for enzymatic thioester hydrolysis, through a putative ketene intermediate, that serves as the basis for fluorine discrimination.

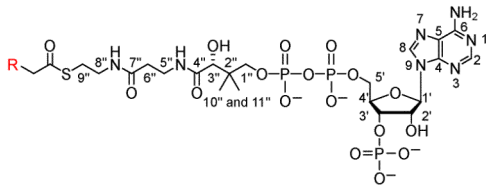
3.2 Materials and methods

Commercial materials. Acetyl coenzyme A sodium salt, *n*-propionyl coenzyme A lithium salt, acetoacetyl coenzyme A sodium salt hydrate, butyryl coenzyme A lithium salt hydrate, coenzyme A hydrate, coenzyme A trilithium salt, anhydrous *N*, *N*-dimethylformamide (DMF), chloroacetic anhydride, cyanoacetic acid, 2-(*N*-morpholino)ethanesulfonic acid (MES), *N*-cyclohexyl-3-aminopropanesulfonic acid (CAPS), sodium hydroxide, trifluoroacetic acid (TFA), sodium fluoroacetate, oxalyl chloride (2 M in dichloromethane), fluoropyruvic acid sodium salt, acetic anhydride-d_6 (99 atom % D), acetic acid-d_4 (99 atom % D), sulfuric acid-d_2 (96-98 wt % in D_2O, 99.5 atom % D), 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB), tris(hydroxymethyl) aminomethane hydrochloride (Tris-HCl), hydroxylamine hydrochloride, urea, iodoacetamide, and sodium citrate dihydrate were purchased from Sigma-Aldrich (St. Louis, MO). Bromoacetic anhydride, methanol, acetonitrile, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), monosodium phosphate monohydrate, disodium phosphate heptahydrate, triethylamine, hydrogen peroxide, sodium chloride, trichloroacetic acid (TCA), hydrochloric acid, tris(2-carboxylethyl)phosphine hydrochloride (TCEP), acetone, and D-sucrose were purchased from Fisher Scientific (Pittsburg, PA). 3,3,3-Trifluoropropionyl chloride was purchased from SynQuest (Alachua, FL). Deuterium oxide (100%) (D_2O) was purchased from Cambridge Isotope Laboratories (Andover, MA). Sequencing-grade chymotrypsin was purchased from Promega (Madison, WI).

Measurement of non-enzymatic hydrolysis rates. Non-enzymatic hydrolysis rates were measured by diluting either fluoroacetyl-CoA or acetyl-CoA into 1 mL water adjusted to the appropriate pH with either buffer (100 mM MES, pH 5.5; 100 mM MES, pH 6.5; 100 mM HEPES, pH 7.5; 100 mM CAPS, pH 10; 100 mM CAPS, pH 11) or sodium hydroxide (pH 12, 13, and 13.6). For experiments at <pH 13, aliquots (50 μL) were removed by hand at various times and quenched by mixing with 25 μL 50% TFA. For experiments at pH > 13, acyl-CoAs were mixed with water adjusted to the appropriate pH with sodium hydroxide using a KinTek Chemical Quench Flow Model RQF-3 and the reaction was stopped at various time points by mixing with 50% TFA. The pH values of the reaction mixtures were measured immediately after acyl-CoA addition and after hydrolysis to ensure that the pH remained constant during the experiment. Quenched samples were analyzed by HPLC on an Agilent Eclipse XDB-C18 (3.5 μm, 3.0 × 150 mm) using a linear gradient from 50 mM sodium phosphate, 0.1% TFA, pH 4.5 to methanol over 15 min at 0.6 mL/min with detection of CoA absorbance at 260 nm. Percent conversion was calculated based on the peak areas for the acyl-CoA and free CoA. Plots of CoA

released versus time were fit as pseudo-first-order reactions to the equation [CoA] = $k[S]_0t$, where k is the rate constant, $[S]_0$ is the initial acyl-CoA concentration, and t is time.

Synthesis of substrates. Solvents used in chemical synthesis were dried using a VAC Solvent Purifier System (Vacuum Atmospheres Company, Amesbury, MA) with the exception of anhydrous DMF, which was purchased from Sigma-Aldrich. HPLC purifications were performed using an Agilent Eclipse XDB-C18 column (9.4 x 250 mm, 5 μm) connected to an Agilent 1200 binary pump and Agilent G1315D diode-array detector (*Appendix 3*). Except where noted, a linear gradient from 0.2% aqueous TFA to 100% methanol at 3 mL/min was applied over 90 min. HPLC fractions were assayed for the desired compounds by liquid chromatography-mass spectrometry (LC-MS) using an Agilent 1290 binary pump coupled to an Agilent 6130 single-quadrupole electrospray ionization mass spectrometer. Fractions containing the desired compounds were pooled, methanol was removed on a rotary evaporator, and the pooled fractions were lyophilized. The lyophilizate was dissolved in water and acyl-CoAs were quantified by absorbance at 260 nm ($\varepsilon_{260\ nm}$ = 13,100 M^{-1} cm^{-1}). Acyl-CoA solutions were stored at -80°C until further use. High-resolution mass spectral analyses were performed at the QB3/Chemistry Mass Spectrometry Facility at the University of California, Berkeley. NMR spectra were collected at the College of Chemistry NMR Facility at the University of California, Berkeley. One-dimensional spectra were recorded on Bruker AV-600 or AVQ-400 NMR spectrometers at 298 K. Chemical shifts are expressed in parts per million (δ, ppm) downfield from tetramethylsilane (^{1}H, ^{2}H, and ^{13}C) or trichlorofluoromethane (^{19}F) and are referenced to the solvent signal. Heteronuclear multiple bond correlation (HMBC) experiments were performed on a Bruker AV-500 NMR spectrometer at 298 K (*Appendix 3*). Numbering for acyl-CoAs and ^{1}H/^{13}C HMBC crosspeaks is shown below. For spectra acquired in 90% H_2O/ 10% D_2O, the solvent signal was suppressed by direct pre-saturation.



Crosspeak numbering for HMBC spectra							
1	$H_{1''}C_{2''}$	**10**	$H_{10''}C_{11''}$	**19**	$H_{5''}C_{7''}$	**28**	H_2C_4
2	$H_{1''}C_{3''}$	**11**	$H_{10''}C_{2''}$	**20**	$H_{5''}C_{6''}$	**29**	H_2C_6
3	$H_{1''}C_{11''}$	**12**	$H_{10''}C_{1''}$	**21**	$H_{6''}C_{7''}$	**30**	H_8C_4
4	$H_{1''}C_{10''}$	**13**	$H_{10''}C_{3''}$	**22**	$H_{6''}C_{5''}$	**31**	H_8C_5
5	$H_{3''}C_{11''}$	**14**	$H_{11''}C_{10''}$	**23**	$H_{8''}C_{7''}$	**32**	$H_{1'}C_2$
6	$H_{3''}C_{10''}$	**15**	$H_{11''}C_{2''}$	**24**	$H_{8''}C_{9''}$	**33**	$H_{1'}C_4$
7	$H_{3''}C_{2''}$	**16**	$H_{11''}C_{1''}$	**25**	$H_{9''}C_{1RAc}$	**34**	$H_{1'}C_{4'}$
8	$H_{3''}C_{1''}$	**17**	$H_{11''}C_{3''}$	**26**	$H_{9''}C_{8''}$	**35**	$H_{2'}C_{1'}$
9	$H_{3''}C_{4''}$	**18**	$H_{5''}C_{4''}$	**27**	$H_{2RAc}C_{1RAc}$	**36**	$H_{3'}C_{4'}$

Fluoroacetyl-CoA *[22, 24]*. Sodium fluoroacetate (100 mg, 1 mmol) was dried under vacuum for 12-16 h in a round-bottom flask equipped with a stir bar and a reflux condenser. The flask was placed under nitrogen pressure and dry tetrahydrofuran (2 mL), DMF (100 μL), and oxalyl chloride (2 M in dichloromethane, 500 μL) were added by syringe. The reaction mixture was heated to 65°C and stirred for 3-4 h. After cooling, 300 μL of the reaction mixture were

added to a stirred, ice-cooled solution of coenzyme A hydrate (50 mg, ~0.06 mmol) and triethylamine (41.8 μL, 0.3 mmol) in DMF (1 mL). After 1 min, the reaction was quenched by addition of water (20 mL) and lyophilized. The lyophilizate was dissolved in water (1 mL) and purified three times by reverse-phase HPLC (0.022 mmol, 37%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.45 (s, 1H, H_8), 8.28 (s, 1H, H_2), 6.03 (d, J = 6 Hz, 1H, $H_{1'}$), 4.87 (d, J = 46 Hz, $C\underline{H_2}F$), 4.73 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.46 (s, 1H, $H_{4'}$), 4.16 (d, J = 15 Hz, 2H, $H_{5'}$), 3.86 (s, 1H, $H_{3''}$), 3.76 (m, 1H, $H_{1''}$), 3.55 (m, 1H, $H_{1''}$), 3.31 (t, J = 6 Hz, 2H, $H_{5''}$), 3.15 (t, J = 6 Hz, 2H, $H_{8''}$), 2.87 (t, J = 6.6 Hz, 2H, $H_{9''}$), 2.23 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.80 (s, 3H, $H_{10''}$), 0.69 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 199.86, 199.71 ($FCH_2\underline{C}O$), 174.54 ($C_{7''}$), 173.93 ($C_{4''}$), 149.65 (C_6), 148.29 (C_2), 144.69 (C_4), 142.34 (C_8), 118.31(C_5), 87.46 ($C_{1'}$), 85.41, 84.20 ($F\underline{C}H_2CO$), 82.95 ($C_{4'}$), 74.48 ($C_{2'}$), 74.45 ($C_{3''}$), 73.61 ($C_{3'}$), 72.24 ($C_{1''}$), 65.16 ($C_{5'}$), 38.31 ($C_{8''}$), 38.27 ($C_{2''}$) 35.38 ($C_{5''}$), 35.30 ($C_{6''}$), 26.87 ($C_{9''}$), 20.59 ($C_{10''}$), 18.31 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -226.1 (t, J = 49.2 Hz). HR-ESI-MS calcd (M-2H$^+$) *m/z* 412.5506, found (M-2H$^+$) *m/z* 412.5507.

Chloroacetyl-CoA *[26]***.** Triethylamine (41.8 μL, 0.3 mmol) was added to an ice-cooled stirred solution of coenzyme A hydrate (50 mg, ~0.06 mmol) in DMF (1 mL). Chloroacetic anhydride (30.8 mg, 0.18 mmol) was added and the reaction mixture was stirred for 3 min at room temperature. The reaction was quenched by addition of water (20 mL) and lyophilized. The lyophilizate was dissolved in water (1 mL) and purified by reverse-phase HPLC (0.028 mmol, 47%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.52 (s, 1H, H_8), 8.33 (s, 1H, H_2), 6.09 (d, J = 6 Hz, 1H, $H_{1'}$), 4.72 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.50 (s, 1H, $H_{4'}$), 4.29 (s, 2H, $C\underline{H_2}Cl$), 4.21 (d, J = 13.8 Hz, 2H, $H_{5'}$), 3.92 (s, 1H, $H_{3''}$), 3.81 (m, 1H, $H_{1''}$), 3.59 (m, 1H, $H_{1''}$), 3.37 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.27 (t, J = 6.6 Hz, 2H, $H_{8''}$), 2.98 (t, J = 6 Hz, 2H, $H_{9''}$), 2.35 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.91 (s, 3H, $H_{10''}$), 0.80 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 197.25 ($ClCH_2\underline{C}O$), 174.58 ($C_{7''}$), 173.97 ($C_{4''}$), 149.70 (C_6), 148.33 (C_2), 144.70 (C_4), 142.37 (C_8), 118.36 (C_5), 87.46 ($C_{1'}$), 83.04 ($C_{4'}$), 74.42 ($C_{2'}$), 74.15 ($C_{3''}$), 73.66 ($C_{3'}$), 72.24 ($C_{1'}$), 65.15 ($C_{5'}$), 47.86 ($Cl\underline{C}H_2CO$), 38.34 ($C_{8''}$), 38.20 ($C_{2''}$) 35.33 ($C_{5''}$), 35.18 ($C_{6''}$), 28.60 ($C_{9''}$), 20.65 ($C_{10''}$), 18.32 ($C_{11''}$). HR-ESI-MS calcd (M-2H$^+$) *m/z* 420.5361, found (M-2H$^+$) *m/z* 420.5363.

Bromoacetyl-CoA *[26]*. Triethylamine (41.8 μL, 0.3 mmol) was added to an ice-cooled stirred solution of coenzyme A hydrate (50 mg, ~0.06 mmol) in DMF (1 mL). Bromoacetic anhydride (46.8 mg, 0.18 mmol) was added and the reaction mixture was stirred for 3 min at room temperature. The reaction was quenched by addition of water (20 mL) and lyophilized. The lyophilizate was dissolved in water (1 mL) and purified by reverse-phase HPLC (0.013 mmol, 21%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.46 (s, 1H, H_8), 8.26 (s, 1H, H_2), 6.03 (d, J = 6 Hz, 1H, $H_{1'}$), 4.77 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.51 (s, 1H, $H_{4'}$), 4.12 (d, J = 13.8 Hz, 2H, $H_{5'}$), 4.06 (s, 2H $C\underline{H_2}Br$), 3.90 (s, 1H, $H_{3'}$), 3.73 (m, 1H, $H_{1''}$), 3.49 (m, 1H, $H_{1''}$), 3.49 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.29 (t, J = 6.6 Hz, 2H, $H_{8''}$), 2.90 (t, J = 6.6 Hz, 2H, $H_{9''}$), 2.27 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.79 (s, 3H, $H_{10''}$), 0.66 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 196.39 ($BrCH_2\underline{C}O$), 174.61 ($C_{7''}$), 173.99 ($C_{4''}$), 149.74 (C_6), 148.38 (C_2), 144.67 (C_4), 142.39 (C_8), 118.36 (C_5), 87.45 ($C_{1'}$), 83.09 ($C_{4'}$), 74.39 ($C_{2'}$), 74.13 ($C_{3''}$), 73.67 ($C_{3'}$), 72.20 ($C_{1'}$), 65.11 ($C_{5'}$), 38.33 ($C_{8''}$), 38.20 ($C_{2''}$) 35.33 ($C_{5''}$), 35.19 ($C_{6''}$), 33.55 ($\underline{C}H_2Br$), 29.06 ($C_{9''}$), 20.68 ($C_{10''}$), 18.28 ($C_{11''}$). HR-ESI-MS calcd (M-2H$^+$) *m/z* 442.5106, found (M-2H$^+$) *m/z* 442.5105.

Cyanoacetyl-CoA. Cyanoacetic acid (85 mg, 1 mmol) was dried under vacuum in an oven-dried round-bottom flask equipped with a stir bar. The flask was placed under nitrogen pressure

and dry dichloromethane (2.4 mL), dry DMF (30 μL), and oxalyl chloride (600 μL, 2 M in dichloromethane) were added by syringe. The reaction mixture was stirred for 3 h at room temperature. A portion (150 μL) was then added to a stirred, ice-cooled solution of coenzyme A hydrate (50 mg, ~0.06 mmol) and triethylamine (41.8 μL, 0.3 mmol) in dry DMF (1 mL). After 1 minute, the reaction was quenched by addition of water (20 mL) and lyophilized. The lyophilizate was dissolved in water (1 mL) and purified by reverse-phase HPLC (0.005 mmol, 8.3%). The resonance corresponding to the α-protons of cyanoacetyl-CoA was not of the expected intensity when the ^{1}H NMR spectrum was recorded in D_2O. ^{1}H/^{13}C HMBC experiments verified that the peak had been assigned correctly, but also revealed the presence of the cyanoacetyl-CoA enolate tautomer (*Appendix 3*). Comparison of the ^{1}H NMR spectrum of cyanoacetyl-CoA dissolved in D_2O to the spectrum of a sample dissolved in 90% H_2O/10% D_2O revealed that these protons readily exchange with solvent. ^{1}H NMR (600 MHz, 90% H_2O/10% D_2O, 25°C) δ (ppm): 8.50 (s, 1H, H_8), 8.28 (s, 1H, H_2), 6.06 (d, J = 5.4 Hz, 1H, $H_{1'}$), 4.74 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.46 (s, 1H, $H_{4'}$), 4.13 (d, J = 10.8 Hz, 2H, $H_{5'}$), 3.88 (s, 1H, $H_{3''}$), 3.73 (m, 1H, $H_{1''}$), 3.58 (s, 2H, $C\underline{H_2}CN$), 3.49 (m, 1H, $H_{1''}$), 3.31 (t, J = 6 Hz, 2H, $H_{5''}$), 3.24 (t, J = 6 Hz, 2H, $H_{8''}$), 2.95 (t, J = 6 Hz, 2H, $H_{9''}$), 2.31 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.84 (s, 3H, $H_{10''}$), 0.68 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 191.26 ($NCCH_2\underline{C}O$), 174.68 ($C_{7''}$), 174.12 ($C_{4''}$), 149.90 (C_6), 148.48 (C_2), 144.73 (C_4), 142.54 (C_8), 118.58 (C_5), 114.87 ($\underline{C}N$), 87.59 ($C_{1'}$), 83.27 ($C_{4'}$), 74.38 ($C_{2'}$), 74.26 ($C_{3''}$), 73.78 ($C_{3'}$), 72.22 ($C_{1'}$), 65.18 ($C_{5'}$), 38.37 ($C_{8''}$), 38.19 ($C_{2''}$) 35.39 ($C_{5''}$), 35.28 ($C_{6''}$), 28.98 ($C_{9''}$), 28.97 ($\underline{C}H_2CN$), 20.76 ($C_{10''}$), 18.38 ($C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 835.1278, found ($M+H^+$) *m/z* 835.1278.

[$^{2}H_2$]-Fluoroacetyl-CoA. Fluoropyruvic acid sodium salt (857 mg, 6.7 mmol) was dissolved in D_2O (10 g, 9 mL) and sulfuric acid-d_2 (1.4 mL) in an 80 mL microwave reaction vessel. The mixture was allowed to react at 120°C with 300 W microwave irradiation in a CEM Discover Labmate microwave reactor for 3 h. After cooling, the reaction mixture was extracted with anhydrous diethyl ether (10 × 10 mL). The ether was removed by rotary evaporator and the resultant [$^{2}H_2$]-fluoropyruvic acid was dissolved in D_2O (10 g, 9 mL) and sulfuric acid-d_2 (0.5 mL) and reacted for an additional 3 h under the same conditions. The reaction mixture was extracted with anhydrous diethyl ether (10 × 10 mL) and the solvent was removed on a rotary evaporator to yield [$^{2}H_2$]-fluoropyruvic acid as a yellow oil (660 mg, 91%). A single resonance was observed in the ^{19}F NMR spectrum, indicating the absence of monodeuterated and undeuterated fluoropyruvic acid. [$^{2}H_2$]-Fluoropyruvic acid (650 mg, 6 mmol) was dissolved in water (30 mL) and the solution was carefully adjusted to pH 7 with 2 M NaOH. Hydrogen peroxide (9 mL of a 30% solution) was added and the mixture was stirred at room temperature for 12 h. The mixture was acidified with dilute sulfuric acid and extracted with diethyl ether (10 × 10 mL). The ether was removed on a rotary evaporator and the resultant [$^{2}H_2$]-fluoroacetic acid was dissolved in water (10 mL), carefully neutralized with 2 M NaOH and lyophilized to yield sodium [$^{2}H_2$]-fluoroacetate (100 mg, 1 mmol, 17%) as a white powder. A single resonance was observed in the ^{19}F NMR spectrum, indicating the absence of monodeuterated and undeuterated sodium fluoroacetate. [$^{2}H_2$]-Fluoroacetyl-CoA was synthesized from sodium [$^{2}H_2$]-fluoroacetate and coenzyme A using the same protocol followed for synthesis of undeuterated fluoroacetyl-CoA. The lyophilized crude product was dissolved in water (1 mL) and purified three times by reverse-phase HPLC (0.013 mmol, 21%). [$^{2}H_2$]-Fluoroacetyl-CoA was judged to be >97% dideuterated based on the ratio of peak areas for undeuterated, monodeuterated, and dideuterated fluoroacetyl-CoA in the ^{19}F NMR spectrum. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.45 (s, 1H, H_8), 8.26 (s, 1H, H_2), 6.03 (d, J = 6 Hz, 1H, $H_{1'}$), 4.77 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.51 (s, 1H, $H_{4'}$),

4.13 (d, J = 12.6 Hz, 2H, $H_{5'}$), 3.83 (s, 1H, $H_{3''}$), 3.70 (m, 1H, $H_{1''}$), 3.52 (m, 1H, $H_{1''}$), 3.31 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.22 (t, J = 6 Hz, 2H, $H_{8''}$), 2.93 (t, J = 6.6 Hz, 2H, $H_{9''}$), 2.29 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.79 (s, 3H, $H_{10''}$), 0.67 (s, 3H, $H_{11''}$); ^{2}H NMR (92.1 MHz, D_2O, 25°C) δ (ppm): 5.0 (d, J = 34 Hz); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 199.98, 199.83 (FCD$_2$<u>C</u>O), 174.61 ($C_{7''}$), 174.00 ($C_{4''}$), 149.74 (C_6), 148.39 (C_2), 144.68 (C_4), 142.41 (C_8), 118.44 (C_5), 87.49 ($C_{1'}$), 83.07, 83.04 (F<u>C</u>D$_2$CO), 83.00 ($C_{4'}$), 74.46 ($C_{2'}$), 74.44 ($C_{3''}$), 73.64 ($C_{3'}$), 72.30 ($C_{1''}$), 65.14 ($C_{5'}$), 38.30 ($C_{8''}$), 38.27 ($C_{2''}$) 35.33 ($C_{5''}$), 35.17 ($C_{6''}$), 26.89 ($C_{9''}$), 20.64 ($C_{10''}$), 18.31 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -227.2 (m, J_{DF} = 7.5 Hz). HR-ESI-MS calcd (M-2H$^+$) *m/z* 413.5576, found (M-2H$^+$) *m/z* 413.5574.

[^{2}H$_3$]-Acetyl-CoA. Triethylamine (41.8 μL, 0.3 mmol) was added to an ice-cooled stirred solution of coenzyme A hydrate (50 mg, ~0.06 mmol) in DMF (1 mL). Acetic anhydride-d$_6$ (19 mg, 17 μL, 0.18 mmol) was added and the reaction mixture was stirred for 3 min at room temperature. The reaction was quenched by addition of water (20 mL) and lyophilized. The lyophilizate was dissolved in water (1 mL) and purified by reverse-phase HPLC (0.020 mmol, 33%). [^{2}H$_3$]-Acetyl-CoA was judged to be 99% deuterated based on the absence of a peak for the α-protons in the ^{1}H NMR spectrum. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.30 (s, 1H, H_8), 8.13 (s, 1H, H_2), 5.89 (d, J = 6 Hz, 1H, $H_{1'}$), 4.57 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.31 (s, 1H, $H_{4'}$), 4.03 (d, J = 13.2 Hz, 2H, $H_{5'}$), 3.71 (s, 1H, $H_{3''}$), 3.58 (m, 1H, $H_{1''}$), 3.39 (m, 1H, $H_{1''}$), 3.16 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.03 (t, J = 6.6 Hz, 2H, $H_{8''}$), 2.66 (t, J = 5.4 Hz, 2H, $H_{9''}$), 2.14 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.72 (s, 3H, $H_{10''}$), 0.54 (s, 3H, $H_{11''}$); ^{2}H NMR (92.1 MHz, D_2O, 25°C) δ (ppm): 2.15; ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 200.70 (CD$_3$<u>C</u>O), 174.52 ($C_{7''}$), 173.83 ($C_{4''}$), 149.62 (C_6), 148.27 (C_2), 144.67 (C_4), 142.32 (C_8), 118.28 (C_5), 87.44 ($C_{1'}$), 82.89 ($C_{4'}$), 74.49 ($C_{2'}$), 74.46 ($C_{3''}$), 73.58 ($C_{3'}$), 72.27 ($C_{1'}$), 65.15 ($C_{5'}$), 38.44 ($C_{8''}$), 38.30 ($C_{2''}$) 35.30 ($C_{5''}$), 35.11 ($C_{6''}$), 29.31 (1:1:1 t, J = 78 Hz, <u>C</u>D$_3$) 28.12 ($C_{9''}$), 20.58 ($C_{10''}$), 18.29 ($C_{11''}$). HR-ESI-MS calcd (M-2H$^+$) *m/z* 405.0645, found (M-2H$^+$) *m/z* 405.0646.

3,3,3-Trifluoropropionyl-CoA. 3,3,3-Trifluoropropionyl chloride (70 mg, 49 μL, 0.48 mmol) was added to a stirred solution of coenzyme A trilithium salt (50 mg, 0.06 mmol) and triethylamine (41.8 μL, 0.3 mmol) in water (1 mL). After 1 min, the reaction mixture was purified by HPLC using a linear gradient from 0-100% B over 30 min at 3 mL/min (A: H_2O, B: acetonitrile). Fractions were assayed for 3,3,3-trifluoropropionyl-CoA by LC-MS and were flash-frozen in liquid nitrogen and lyophilized (0.012 mmol, 20%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.44 (s, 1H, H_8), 8.22 (s, 1H, H_2), 5.98 (d, J = 6 Hz, 1H, $H_{1'}$), 4.71 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.42 (s, 1H, $H_{4'}$), 4.11 (d, J = 13.8 Hz, 2H, $H_{5'}$), 3.84 (s, 1H, $H_{3''}$), 3.72 (m, 1H, $H_{1''}$), 3.47 (q, J = 12 Hz, C<u>H</u>$_2$CF$_3$), 3.46 (m, 1H, $H_{1''}$), 3.26 (t, J = 6 Hz, 2H, $H_{5''}$), 3.18 (t, J = 6 Hz, 2H, $H_{8''}$), 2.89 (t, J = 6 Hz, 2H, $H_{9''}$), 2.24 (t, J = 6 Hz , 2H, $H_{6''}$), 0.77 (s, 3H, $H_{10''}$), 0.65 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 193.06 (CF$_3$CH$_2$<u>C</u>O), 174.62 ($C_{7''}$), 173.96 ($C_{4''}$), 149.68 (C_6), 148.31 (C_2), 144.78 (C_4), 142.30 (C_8), 118.27 (C_5), 87.37 ($C_{1'}$), 83.28 ($C_{4'}$), 74.24 ($C_{2'}$), 74.20 ($C_{3''}$), 73.67 ($C_{3'}$), 72.10 ($C_{1'}$), 65.10 ($C_{5'}$), 38.35 ($C_{8''}$), 38.20 ($C_{2''}$) 35.31 ($C_{5''}$), 35.18 ($C_{6''}$), 28.51 ($C_{9''}$), 20.75 ($C_{10''}$), 18.24 ($C_{11''}$). ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -62.17. HR-ESI-MS calcd (M-H$^+$) *m/z* 876.1045, found (M-H$^+$) *m/z* 876.1050.

Steady-state kinetic experiments. Steady-state kinetic experiments were performed using 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) to detect release of free coenzyme A as described previously *[24]*. Enzymatic reactions were initiated by addition of FlK (5 nM – 10 μM) to reaction mixtures containing butyryl-CoA, propionyl-CoA, acetyl-CoA, acetoacetyl-CoA,

bromoacetyl-CoA, chloroacetyl-CoA, fluoroacetyl-CoA, or cyanoacetyl-CoA. When included, sucrose was added to a final concentration of 30% (w/v) from a 60% (w/v) stock solution. Each rate was measured in triplicate. Kinetic parameters (k_{cat} and K_M) were determined by fitting the data to the equation $V_0 = V_{max}[S]/K_M + [S]$, where V_0 is the initial rate and [S] is the substrate concentration, using Origin 6.0 (OriginLab Corporation, Northampton, MA). Pseudo-first order rates of non-enzymatic hydrolysis under assay conditions were measured by monitoring CoA release in the absence of enzyme. Data were fit to the equation $[CoA] = k[S]_0t$, where k is the rate constant, $[S]_0$ is the initial acyl-CoA concentration, and t is time. Taft plots were constructed by plotting log k_{uncat} or log k_{cat} versus the Taft polar substituent constant σ^* for the α-substituent of each acyl-CoA tested *[27, 28]*. Error bars shown on the plot for uncatalyzed hydrolysis represent the standard error for three measurements of the pseudo-first-order rate constants. Error bars shown on the plot for FlK-catalyzed hydrolysis are derived from the Michaelis-Menten fits of triplicate data sets. Taft plots were fit to the equation $\log k = \rho^*\sigma^*$, where ρ^* is the polar sensitivity factor.

Rapid quench kinetic experiments. Pre-steady-state kinetic experiments were performed using rapid chemical quench followed by HPLC separation of coenzyme A from unhydrolyzed acyl-CoA. FlK (25-100 μM in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl) was mixed with substrate (250-3000 μM in water) using a KinTek Chemical Quench Flow Model RQF-3. The reaction was stopped at various times by mixing with 50% TFA to achieve a final concentration of 17% TFA. Quenched samples were analyzed by HPLC on an Agilent Eclipse XDB-C18 (3.5 μm, 3.0 × 150 mm) using a linear gradient from 50 mM sodium phosphate, 0.1% TFA, pH 4.5 to methanol over 15 min at 0.6 mL/min with detection of CoA absorbance at 260 nm. Percent conversion was calculated based on the peak areas for substrate and product. Plots of coenzyme A released versus time were fit to the equation $[CoA] = A^*(1-\exp(-k_2^*t)) + V^*t$, where A is the burst amplitude, k_2 is the burst-phase rate constant, and V is the steady-state rate. In cases where data could not be fit with a burst phase, they were fit as pseudo-first-order reactions according to the equation $[CoA] = k[FlK]t + [CoA]_0$, where the y-intercept, $[CoA]_0$, is the concentration of CoA released before the first data point, k is the rate constant, and t is time.

Hydroxylamine trapping of the acyl-enzyme intermediate. For trapping experiments with acetyl-CoA and fluoroacetyl-CoA, FlK (100 μM) in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl (50 μL) was rapidly mixed with the appropriate substrate (acetyl-CoA, 6 mM; fluoroacetyl-CoA, 1 mM) using a KinTek Chemical Quench Flow Model RQF-3. The reaction was quenched after 300 ms (acetyl-CoA) or after 2 ms (fluoroacetyl-CoA) with 8 M urea in 50 mM citrate-phosphate buffer, pH 4.0 and collected into a tube containing 100 mM hydroxylamine in quench buffer. For trapping of the acyl-enzyme intermediate derived from 3,3,3-trifluoropropionyl-CoA, FlK (50 μM) in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl (50 μL) was allowed to react with 3,3,3-trifluoropropionyl-CoA (1 mM) for 15 min. The reaction was quenched by the addition of 8 M urea in 50 mM citrate-phosphate buffer, pH 4.0 (450 μL) and 1 M hydroxylamine hydrochloride, pH 7.0 (50 μL). A control sample without any acyl-CoA substrate was also prepared following the same procedure. For all samples, FlK was precipitated by addition of 100% (w/v) trichloroacetic acid (125 μL). After standing on ice for 1 h, the sample was centrifuged at 20,817 × *g* for 10 min at 4°C to pellet the precipitated protein. The pellet was washed three times with ice cold 0.01 M HCl/90% acetone, air dried, and then dissolved in 100 mM Tris-HCl, pH 8.0, 8 M urea (80 μL). TCEP was added to a final concentration of 5 mM and the sample was incubated at room temperature for 20 min before adding iodoacetamide (10 mM final

concentration) and incubating in the dark at room temperature for 15 min. The sample was then diluted with 100 mM Tris-HCl, pH 8.0 (720 μL) and sequencing-grade chymotrypsin (2 μg, Promega) was added. The digest was allowed to proceed overnight at room temperature and was then quenched by addition of formic acid to 5% (v/v). The sample was desalted on a C18 OMIX tip (Agilent). A nano LC column consisting of 10 cm of Polaris C18 5 μm packing material (Agilent) was packed into a glass capillary (100 μm inner diameter) using a pressure bomb. The column was extensively equilibrated with buffer A (5% acetonitrile/0.02% heptafluorobutyric acid (HFBA)) and directly coupled to an electrospray ionization source mounted on a Thermo-Fisher LTQ XL linear ion trap mass spectrometer. An Agilent 1200 HPLC equipped with a split line to deliver a flow rate of 30 nL/min was used for chromatography. Peptides were eluted using a gradient from 100% buffer A to 60% buffer B (80% acetonitrile/0.02% HFBA). The programs SEQUEST and DTASELECT *[29, 30]* were used to identify peptides from a database of the protein sequence and a database of common contaminants.

Phylogenetic analysis and sequence alignments. Representative sequences were retrieved from Pfam seed sequence sets from eight characterized families with the hotdog superfamily. The sequences, along with the sequences of FlK and its only known homolog with >60% sequence identity, were aligned using the PRALINE multiple sequence alignment server *[31]*, which incorporates information from PSI-BLAST pre-profile processing and PSIPRED secondary structure prediction. The alignment was manually adjusted where necessary to incorporate information from structural alignments. Using this alignment, a phylogenetic tree was constructed using the neighbor-joining method in MEGA 5 *[32]*. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the 78 sequences analyzed. Bootstrap values are indicated next to the branches. The tree is drawn to scale in units of number of amino acid differences per site. There were a total of 37 positions in the final dataset.

Kinetic isotope effect measurements. Kinetic isotope effects were measured by direct comparison of the kinetic constants for α-deuterated and undeuterated substrates. Fluoroacetyl-CoA and [$^2H_2$]-fluoroacetyl-CoA were purified three times by HPLC. Acetyl-CoA purchased from Sigma-Aldrich and [$^2H_3$]-acetyl-CoA were both purified by reverse-phase HPLC to ensure similar sample composition. Steady-state kinetic experiments were performed by mixing equal volumes of FlK (120 nM in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl for fluoroacetyl-CoA; 20 μM in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl for acetyl-CoA) and deuterated or undeuterated substrate (10-200 μM in water for fluoroacetyl-CoA; 500-6000 μM for acetyl-CoA) using a KinTek Chemical Quench Flow Model RQF-3. The reactions were stopped at various points on the second timescale by mixing with 50% TFA to achieve a final concentration of 17% TFA. Quenched samples were analyzed by HPLC on an Agilent Eclipse XDB-C18 (3.5 μm, 3.0 × 150 mm) using a linear gradient from 50 mM sodium phosphate, 0.1% TFA, pH 4.5 to methanol over 15 min at 0.6 mL/min with detection of CoA absorbance at 260 nm. Rates were determined by linear fitting of a plot of CoA released versus time using 12-15 data points for each substrate concentration. Error bars shown for individual rates represent the error derived from linear fitting. Kinetic constants were determined by fitting a plot of the initial rates (V_0) versus substrate concentration ([S]) to the equation $V_0 = V_{max}[S] / K_M + [S]$. Kinetic isotope effects were determined by calculating the ratios k_H/k_D and $^DV/K$. Errors for the individual parameters were propagated to calculate the errors for the kinetic isotope effect.

For pre-steady-state kinetic isotope effect experiments, FlK (50 μM in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl) was mixed with deuterated or undeuterated substrate (500 μM in water) using the chemical quench flow apparatus and reactions were quenched by mixing with 50% TFA. The plot of CoA released versus time for [$^2H_2$]-fluoroacetyl-CoA was best fit by the burst equation [CoA] = A*(1-exp(-k_2*t)) + V*t, where A is the burst amplitude, k_2 is the burst-phase rate constant, and V is the steady-state rate. For undeuterated fluoroacetyl-CoA, the plot was best fit by a line.

Time-dependent fluoride release and FlK inactivation with 3,3,3-trifluoropropionyl-CoA. *FlK-catalyzed turnover of 3,3,3-trifluoropropionyl-CoA*. Turnover of 3,3,3-trifluoropropionyl-CoA was measured using DTNB to detect release of free coenzyme A. Reactions were initiated by addition of FlK to reaction mixtures containing 3,3,3-trifluoropropionyl-CoA (100, 250, 500, or 1000 μM) and DTNB (0.5 mM) in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl. Time courses were fit to the equation [CoA] = A*(1-exp(-k_{obs}*t)) + V*t, where A is the concentration of CoA released during the first turnover of the reaction, k_{obs} is the rate constant for CoA release during the first turnover, and V is the rate for the pseudo-first-order phase of the reaction, which was found to correspond to the background rate of hydrolysis.

Inactivation of FlK by 3,3,3-trifluoropropionyl-CoA. Inactivation of FlK by 3,3,3-trifluoropropionyl-CoA was measured by testing the activity of FlK against fluoroacetyl-CoA at various time points after addition of 3,3,3-trifluoropropionyl-CoA. A solution of FlK (50 μM) was treated with 3,3,3-trifluoropropionyl-CoA (1 mM). At various time points after addition of 3,3,3,-trifluoropropionyl-CoA, an aliquot of the solution (0.5 μL) was removed and diluted 1:1000 into a reaction mixture containing fluoroacetyl-CoA (100 μM) and DTNB (0.5 mM) in 100 mM Tris-HCl, pH 7.6. Release of free coenzyme A was monitored by reaction of the thiol with DTNB. The rate constant for inactivation was measured by fitting a plot of activity remaining versus time after treatment with 3,3,3-trifluoropropionyl-CoA to the equation A = A_{final} + A_0*(1-exp(-k_{obs}*t)), where A is fraction of the initial activity remaining, A_{final} is the final activity, and A_0 is the initial activity.

FlK-catalyzed fluoride release from 3,3,3-trifluoropropionyl-CoA. Fluoride release was measured using a fluoride ion selective electrode (Mettler Toledo). FlK was dialyzed for 16 h against 2 L of 20 mM Tris-HCl, pH 7.6, 50 mM NaCl. The same dialysis buffer was then used for all subsequent protein and substrate dilutions and for dissolving sodium fluoride standards. FlK was diluted to 50 μM in dialysis buffer in total volume of 1 mL. Reactions were initiated by addition of 3,3,3-trifluoropropionyl-CoA (1 mM final concentration) and the change in fluoride electrode potential was recorded as a function of time. The background rate of fluoride release was measured by diluting 3,3,3-trifluoropropionyl-CoA to 1 mM in dialysis buffer in the absence of FlK. Electrode potential values were converted to fluoride concentration using a standard curve. The response time of the fluoride electrode was measured to be an order of magnitude faster than enzyme-catalyzed fluoride release (k_{obs} = 2 ± 0.1 min^{-1}). The background rate of fluoride release was subtracted and time courses of FlK-catalyzed fluoride release were fit to the equation [CoA] = A*(1-exp(-k_{obs}*t)) + F_0, where A is the final concentration of fluoride, k_{obs} is the rate constant, and F_0 is the initial amount of fluoride present in the solution.

Quantification of exchange of substrate α-protons with solvent. FlK storage buffer (20 mM Tris-HCl, pH 7.6, 50 mM NaCl) was prepared in H_2O and lyophilized. The buffer was dissolved in D_2O and lyophilized a second time. The lyophilizate was dissolved in D_2O and used

to equilibrate a Micro Bio-Spin column (Bio-Rad). A aliquot of concentrated FlK (50 μL) was passed over the column to exchange it into the deuterated solvent. A reaction mixture (1 mL) containing 3 mM acyl-CoA, 20 mM Tris-DCl, pD 7.6, 50 mM NaCl, and either 1 μM FlK (for fluoroacetyl-CoA) or 10 μM FlK (for acetyl-CoA) in D_2O was prepared. The reactions were allowed to proceed for 1 h at room temperature and were then lyophilized. The lyophilizate was dissolved in H_2O (200 μL) and analyzed by ^2H NMR (92.1 MHz, 25°C, 2000 scans) to look for deuterium exchange from solvent into the product α-protons. Standards containing either 150 μM [$^2H_3$]-acetic acid or 150 μM sodium [$^2H_2$]-fluoroacetate in the same buffer established that the limit of detection was <1%.

3.3 Results and discussion

Elucidating the kinetic basis for fluorine selectivity. The hydrolysis of thioesters typically involves direct attack of the nucleophile at the carbonyl group to form a tetrahedral intermediate, which subsequently collapses to produce the free carboxylic acid and thiol end products. The fluorine substitution itself activates the carbonyl moiety towards nucleophilic attack through inductive effects and provides a potential mechanism for discrimination between fluoroacetyl- and acetyl-CoA (*Figure 3.2*). However, the rate of chemical hydrolysis of fluoroacetyl-CoA at neutral pH is only 10-fold faster compared to acetyl-CoA rather than 10^4-fold, as occurs in the FlK active site [*24*]. Even at pH > 13, the largest difference in reactivity that we observe for chemical hydrolysis is 100-fold (*Figure 3.3*). These observations suggest that the large rate acceleration of FlK with respect to fluoroacetyl-CoA may originate from more than a simple increase in the electrophilicity of the carbonyl group upon fluorine substitution. In order to further probe the role of carbonyl activation in FlK substrate selectivity, we prepared a set of acyl-CoA analogs with different functional groups at the α-position, thereby accessing a wide range of values of the polar substituent constant (σ*) that describes the activation of the acyl-CoA carbonyl group towards nucleophilic attack [*27*]. By examining the relationship between σ* and the rate of reaction, this series of compounds can be used to assess the relative contribution of the α-substituent towards acyl-CoA hydrolysis. The polar substituent constant (σ*) is derived from comparison of the rates of base-catalyzed hydrolysis of a series of methyl esters to the rates of acid-catalyzed hydrolysis of the same esters to separate and quantify steric and electronic effects *[27]*. The methyl group is defined as the reference substituent and its σ* is defined as 0. A linear relationship between σ* and the logarithm of the rate constant indicates that a consistent amount of charge accumulates in the transition state for all substituents, while an abrupt change in the slope of the plot indicates a change in the amount of charge buildup, suggesting a change in rate-determining step, chemical mechanism, or both. The slope of the line, ρ*, is the polar sensitivity factor that compares the reaction under study to the reference reaction. A slope >1 indicates that the reaction is more sensitive to substituents than the reference reaction, while a slope <1 indicates that it is less sensitive.

We found a linear relationship between σ* and the logarithm of the pseudo-first order rate constant for chemical hydrolysis of the acyl-CoAs at neutral pH, implying that they share a common mechanism and transition state for the overall transformation under these conditions (*Figure 3.4, Figure 3.5*). Our analysis gave a ρ* value of 1.06 (R^2 = 0.99) for uncatalyzed hydrolysis and a ρ* of 2.3 (R^2= 0.98) for catalyzed hydrolysis, excluding fluoroacetyl-CoA and

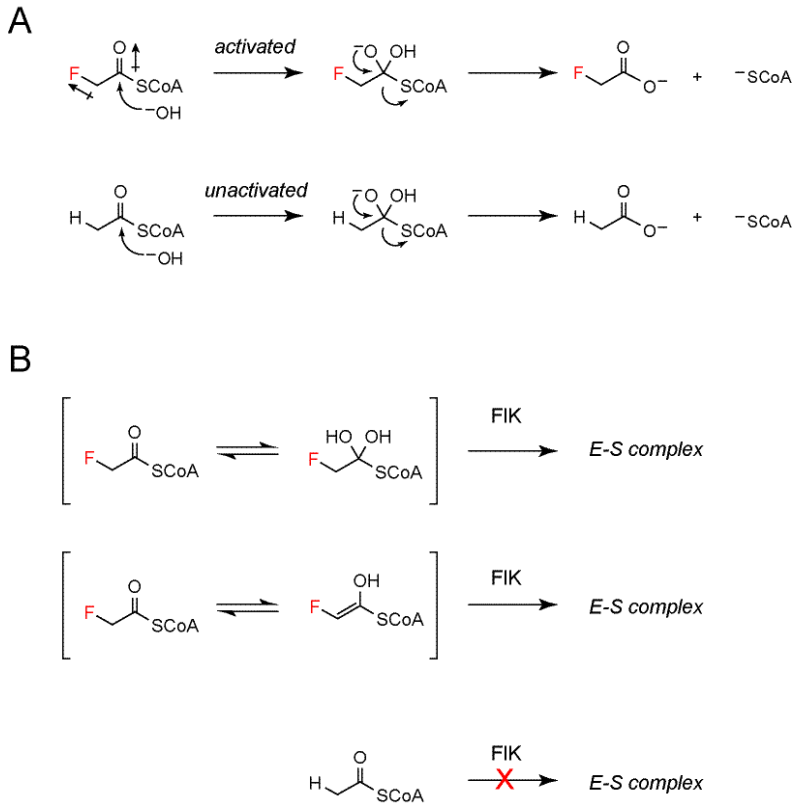
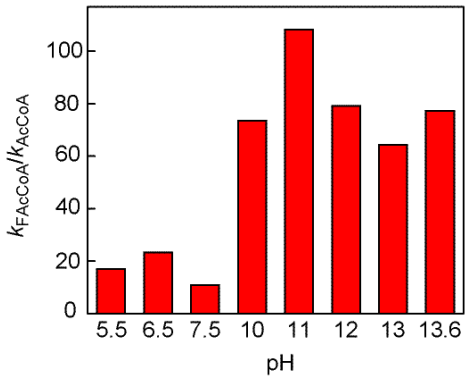


Figure 3.2. *Alternative mechanisms for discrimination of fluoroacetyl-CoA and acetyl-CoA. (A) Fluoroacetyl-CoA could be discriminated based on carbonyl activation related to the fluorine substitution. In this model, the rate of nucleophilic attack on the fluorinated substrate is faster than the nonfluorinated substrate based on inductive effects that increase the electrophilicity at the carbonyl group. (B) Fluoroacetyl-CoA has the potential to form a hydrate or an enolate in aqueous solution, which could be the form of the substrate recognized by FIK. In this scenario, acetyl-CoA would bind poorly to the enzyme because it is less likely to form either the hydrate or enolate, allowing for specific hydrolysis of the fluorinated substrate.*

A



B

pH	k_{AcCoA} (s^{-1})	k_{FAcCoA} (s^{-1})
5.5	5.8×10^{-7}	9.8×10^{-6}
6.5	7.1×10^{-7}	1.6×10^{-5}
7.5	1.7×10^{-6}	1.8×10^{-5}
10	2.1×10^{-5}	1.5×10^{-3}
11	8.1×10^{-5}	8.8×10^{-3}
12	7.0×10^{-4}	4.0×10^{-1}
13	3.0×10^{-2}	8.0×10^{-1}
13.6	2.4×10^{-1}	5.1

Figure 3.3. *pH dependence of non-enzymatic acyl-CoA hydrolysis. (A) Ratios of rate constants for fluoroacetyl-CoA and acetyl-CoA hydrolysis. (B) Pseudo-first-order rate constants for fluoroacetyl-CoA and acetyl-CoA hydrolysis.*

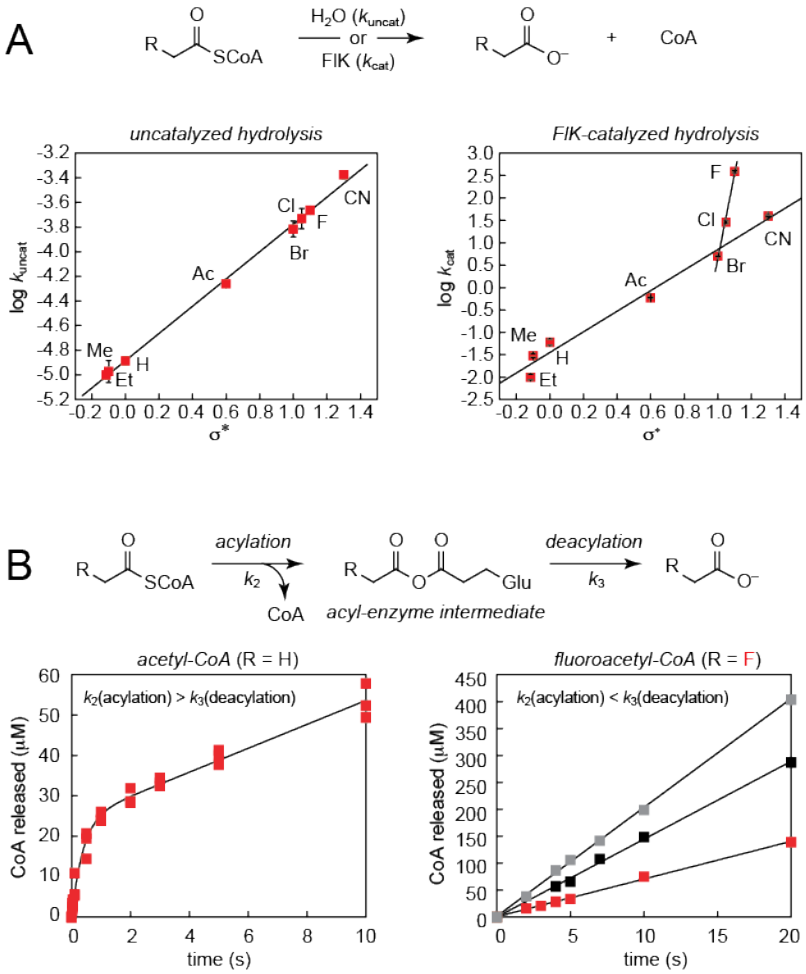


Figure 3.4. *Steady-state and pre-steady-state kinetics of FIK-catalyzed acyl-CoA hydrolysis. (A) Taft free energy relationship analysis of uncatalyzed (left) and FIK-catalyzed (right) acyl-CoA hydrolysis. Pseudo-first-order rate constants were determined by linear fitting (for uncatalyzed reactions) or by fitting a plot of initial rate versus substrate concentration to the Michaelis-Menten equation (for FIK-catalyzed reactions). (B) Pre-steady-state kinetics of FIK-catalyzed acetyl-CoA (left, 25 μM FIK) and fluoroacetyl-CoA hydrolysis (right; ■, 25 μM FIK; ■, 50 μM FIK; ■, 75 μM FIK).*

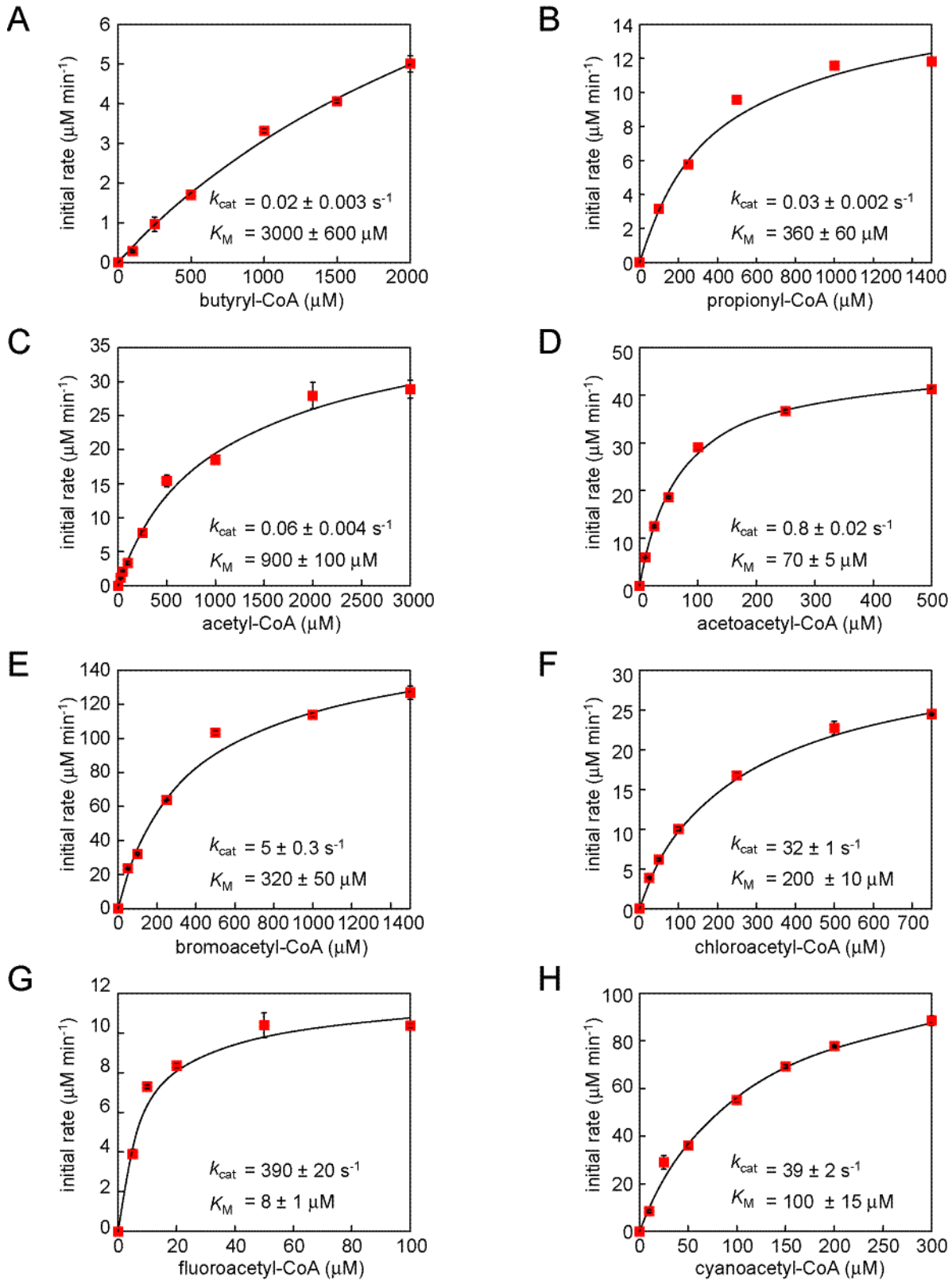


Figure 3.5. *Michaelis-Menten curves for substrates used in Taft analysis. (A) Butyryl-CoA. (B) Propionyl-CoA. (C) Acetyl-CoA. (D) Acetoacetyl-CoA. (E) Bromoacetyl-CoA. (F) Chloroacetyl-CoA. (G) Fluoroacetyl-CoA. (H) Cyanoacetyl-CoA.*

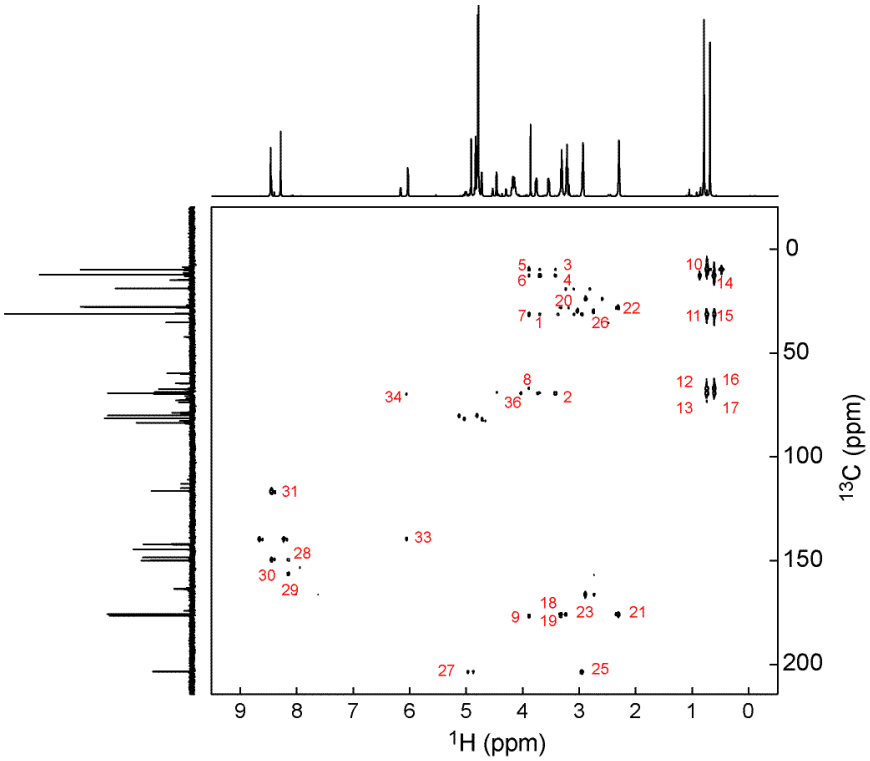


Figure 3.5. *$^{1}H/^{13}C$ HMBC spectrum of fluoroacetyl-CoA under FlK assay conditions. $^{1}H/^{13}C$ HMBC experiments show no evidence for formation of a fluoroacetyl-CoA hydrate or enolate within the limit of detection (see Materials and Methods for acyl-CoA numbering).*

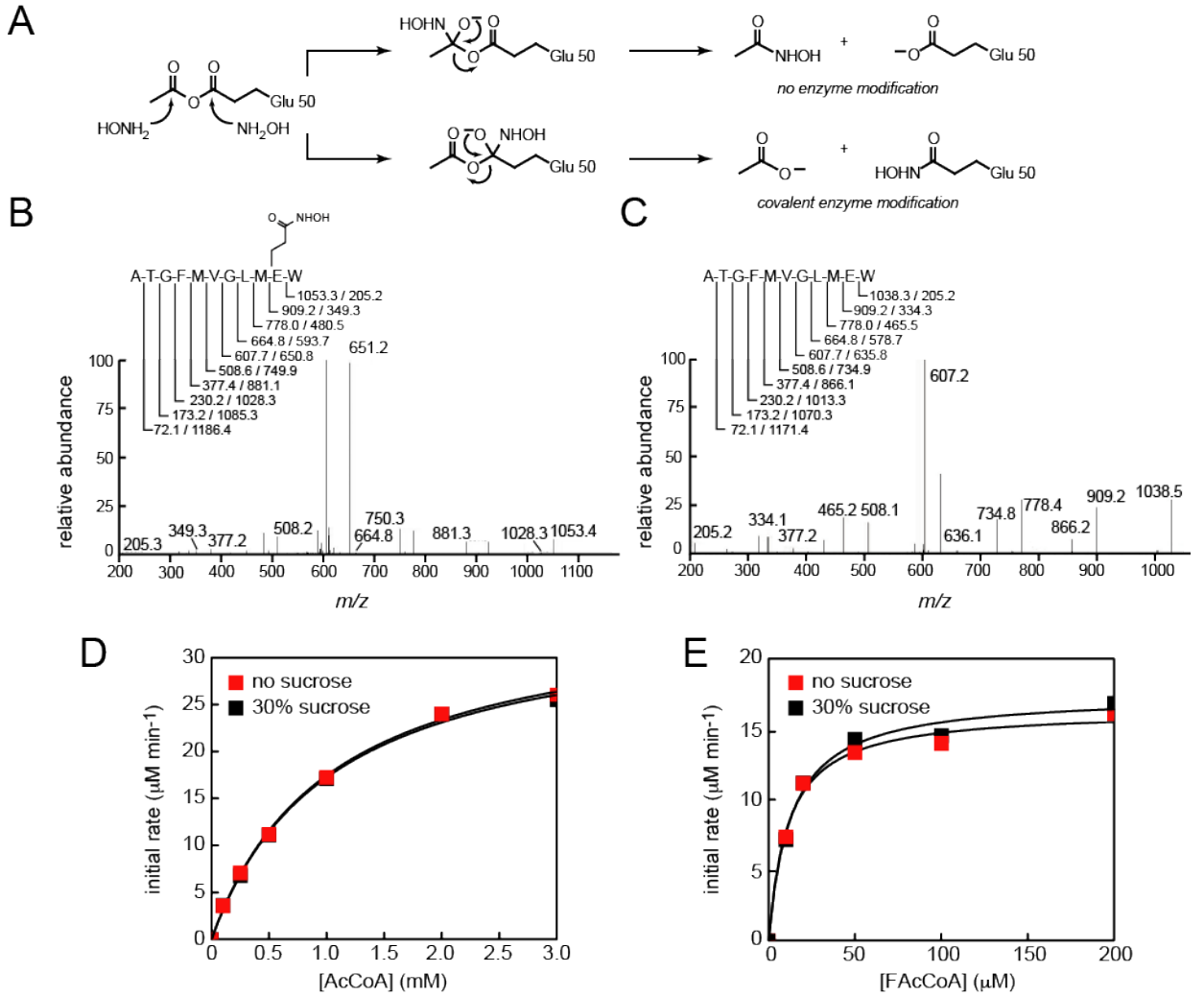
chloroacetyl-CoA from the linear fit. In comparison, the same plot for FlK-catalyzed acyl-CoA hydrolysis produces two intersecting lines, with the chlorinated and fluorinated substrates being hydrolyzed at rates faster than predicted by σ*, suggesting a change in transition state structure, rate-limiting step, or chemical mechanism for halogenated substrates compared to non-halogenated substrates (*Figure 3.4A, Figure 3.5*). An alternative explanation for this difference in behavior is that FlK actually binds a different form of the fluoroacetyl-CoA substrate, which has higher potential to form a hydrate or enolate in aqueous solution compared to other less polarized acyl-CoAs like acetyl-CoA (*Figure 3.2*). However, NMR studies showed no evidence for formation of a hydrate or enolate of fluoroacetyl-CoA under these conditions within the limit of detection (*Figure 3.6*). From the Taft analysis and these control experiments, it appears that the fluorine-dependent selectivity of FlK could then be related to a change in the rate-determining step or catalytic mechanism of hydrolysis between the two substrates.

We began to explore the basis of this difference in reactivity by comparing the pre-steady state kinetic behavior of FlK with respect to acetyl-CoA and fluoroacetyl-CoA. By quenching FlK-catalyzed hydrolysis reactions on a millisecond timescale, it is possible to monitor the rate of CoA release before the first turnover. With acetyl-CoA, one equivalent of free CoA is formed in a burst phase (k_2) during the first turnover before the steady state is reached ($k_3 = k_{cat}$), which indicates that there is a slow step in the catalytic cycle subsequent to CoA formation (*Figure 3.4B*). On one hand, a burst phase can result from rapid chemistry involved in the conversion of substrate to product followed by a rate-determining physical event, such as product release or some other conformational change required for catalysis [*33*]. The steady-state rate constant (k_{cat})

following the burst would then correspond mainly to the rate of this physical event rather than a chemical step in the catalytic mechanism. However, burst phases are also often observed in cases where a conformational change is not rate limiting and are derived instead from fast formation of an enzyme-bound intermediate prior to a rate-limiting catalytic step [*34*]. With regard to FlK, hydrolases represent a canonical example of covalent catalysis where the burst phase is derived from rapid enzyme acylation by the substrate followed by slower enzyme deacylation as the rate-determining step [*35*]. Furthermore, recent studies of a 4-hydroxybenzoyl-CoA thioesterase (4-HBT) from the FlK superfamily have shown that CoA is formed in a burst with exchange of protein-derived oxygen into the carboxylic acid, which provides strong support for the formation of an enzyme anhydride on the active site glutamate of the catalytic triad [*36-38*].

To test for the intermediacy of an enzyme-anhydride in FlK-catalyzed hydrolysis of acetyl-CoA, we carried out trapping experiments with hydroxylamine. Nucleophilic attack of hydroxylamine on the putative enzyme anhydride could occur either at the substrate-derived or enzyme-derived carbonyl group to regenerate the native enzyme or convert the carboxylate side chain to a hydroxamate, respectively (*Figure 3.7*). Notably, no covalent modification of the enzyme is expected for hydroxylamine trapping of other types of acyl-enzyme intermediates, such as esters, because the electrophilic carbonyl group in these cases are derived from the substrate with the native enzyme being re-formed upon reaction with nucleophiles. Tandem mass spectrometry analysis of FlK after incubation with an excess of acetyl-CoA followed by rapid quenching with urea in the presence of hydroxylamine revealed a 15 Da modification on E50 corresponding to conversion of the glutamate side chain to a hydroxamate (*Figure 3.7*). In contrast, no modification was detected in a control sample in which FlK was treated with urea and hydroxylamine in the absence of acetyl-CoA (*Figure 3.7*). Based on these results, we propose that the burst kinetics observed with acetyl-CoA result from the formation of a covalent acyl-enzyme intermediate on E50 concomitant with formation of free CoA. Although anhydrides are high in energy, these acyl-enzyme intermediates have been previously documented in biological catalysis, including in thioesterases of both the crotonase [*36*] and hotdog-fold [*37, 38*] superfamiles. The burst rate for CoA formation (k_2) for FlK then describes the rate of enzyme acylation, which is followed by the rate-limiting step (k_{cat}) in the catalytic cycle (*Figure 3.4B*). The dependence of k_{cat} on the electron-withdrawing ability of the α-substituent but lack of correlation with other parameters like size strongly suggests that the slow step in FlK turnover involves a chemical transformation of the acyl-enzyme intermediate rather than a rate-determining conformational change (*Figure 3.4A*). Although our previous crystallographic studies had suggested a role for opening of the F36 "gate" in product release, mutation of this residue to form a constitutively open enzyme does not affect the rate of hydrolysis of either acetyl-CoA or fluoroacetyl-CoA [*24*], suggesting that this motion is not rate-limiting. Furthermore, the insensitivity of k_{cat} to addition of a viscogen (30% sucrose) rules out the contribution of diffusional steps, such as product release, to rate limitation (*Figure 3.7D*). We therefore interpret the steady-state rate of hydrolysis of acetyl-CoA ($k_{cat} = k_3$) to represent a chemical step in FlK catalysis related to enzyme deacylation (*Figure 3.4B*).

In comparison, FlK shows no evidence of burst kinetics with fluoroacetyl-CoA based on the fact that the linear fits for product formation at different enzyme concentrations all intersect the y-axis at 0 (*Figure 3.4B*). For fluoroacetyl-CoA hydrolysis, plots of CoA released versus time cannot be fit by the burst equation, but are best described by a linear fit. The earliest time at which our instrument is able to quench, however, is 2 ms, leaving open the possibility that FlK does have a burst phase with respect to fluoroacetyl-CoA, but it occurs with a rate constant of



F

	% sucrose	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
acetyl-CoA	0	0.06 ± 0.002	1040 ± 80	60 ± 5
	30	0.06 ± 0.002	1030 ± 80	60 ± 5
fluoroacetyl-CoA	0	219 ± 8	11 ± 2	(2.0 ± 0.4) x 10^7
	30	233 ± 8	13 ± 2	(1.8 ± 0.2) x 10^7

Figure 3.7. *Characterization of FlK burst-phase kinetics. (A) Possible pathways of reaction between hydroxylamine and the acyl-anhydride intermediate. (B) LC-MS/MS analysis of hydroxylamine-quenched FlK reacted with acetyl-CoA showing the modification of the chymotryptic peptide containing Glu 50. (C) No modification was observed with FlK treated with hydroxylamine in the absence of acetyl-CoA substrate. Expected mass-to-charge (m/z) ratios for b-ions/y-ions are indicated. (D) Michaelis-Menten curve for FlK-catalyzed hydrolysis of acetyl-CoA in the presence (black) and absence (red) of sucrose. (E) Michaelis-Menten curve for FlK-catalyzed hydrolysis of fluoroacetyl-CoA in the presence (black) and absence (red) of sucrose. (F) Steady-state kinetic parameters for FlK in the presence and absence of sucrose.*

>500 s^{-1} and would therefore be over before the first quenched point that we are able to examine. To distinguish between the absence of a burst phase and the presence of a burst phase that is too fast to be measured by our assay, we examined the amount of free CoA present in the first quenched sample. If no burst phase had occurred, we would expect the CoA concentration to be given by [CoA] = k_{cat}[FlK]t, corresponding to a line with a y-intercept of zero. If a burst phase had occurred before the first quenched point, we would expect [CoA] = k_{cat}[FlK]t + [FlK], corresponding to a line with a y-intercept of [FlK]. For experiments conducted at three different FlK concentrations (25 μM, 50 μM, and 75 μM), we found that the data were best fit by lines whose y-intercepts were within error of zero. We therefore concluded that a step prior to or concomitant with CoA release is rate-limiting in the case of fluoroacetyl-CoA. The sucrose independence of the rate of hydrolysis of fluoroacetyl-CoA (*Figure 3.7D*) implies that a chemical step remains rate limiting and our kinetic data are thus consistent with two mechanisms. One possibility is that fluoroacetyl-CoA hydrolysis does not involve an acyl-enzyme intermediate and therefore has a single kinetic phase. A second possibility is that enzyme acylation is the slow step in the catalytic cycle ($k_{cat} = k_2$) and that the following steps are faster ($k_3 > k_2$). Although we were unable to observe hydroxylamine modification of FlK in trapping experiments in which fluoroacetyl-CoA was used as the substrate, we cannot rule out formation of an acyl-enzyme intermediate at this time. First, the single-phase kinetic behavior with respect to CoA formation suggests that the breakdown of any intermediate would be faster than its formation and prevent its accumulation for reaction with hydroxylamine. Additionally, the electron-withdrawing fluorine substituent may favor reaction of hydroxylamine at the substrate-derived carbonyl, which would mask the formation of an enzyme-anhydride. Although these two mechanistic possibilities cannot be distinguished at this point, the data are all consistent with a model in which fluorine selectivity arises from a change in the chemical mechanism or rate-limiting step of the reaction between acetyl-CoA and fluoroacetyl-CoA.

Identifying the chemical mechanism for fluorine selectivity in FlK-catalyzed thioester hydrolysis. We next turned our attention to using evolutionary relationships to gain insight into the chemical mechanisms of the different kinetic behaviors observed for each substrate. Interestingly, FlK appears at the sequence and structural levels to be a chimera of two hot-dog fold superfamilies – the 4-HBT thioesterases and the MaoC dehydratases (*Figure 3.8*, *Figure 3.9*, *Figure 3.10*). At a structural level, FlK possesses a distinctive hydrophobic lid (*Figure 3.10*) that is not found in any of the other crystallographically-characterized thioesterases but is shared with the dehydratase family [*39*]. Moreover, the lid is functionally important in FlK for molecular recognition of fluorine via a gatekeeper residue (F36) [*24*], which is conserved among dehydratases even though it is unique when compared to FlK's closest orthologs (*Figure 3.9*). However, FlK shares its catalytic dyad/triad (E50, H76, T42) and putative acyl-enzyme intermediate with the 4-HBT thioesterases. Despite its similarities to the 4-HBT thioesterases, the weaker oxyanion hole and key role of H76 in FlK hydrolysis are more reminiscent of the dehydratases, which utilize an active site histidine to remove a proton at the α-position of 3-hydroxyacyl-CoA substrates (*Figure 3.11*) [*39*]. Thus, we reasoned that FlK might utilize a dehydratase-like mechanism for hydrolysis initiated by C_α-deprotonation of either the fluoroacetyl-CoA substrate or the acyl-enzyme intermediate to discriminate between fluoroacetyl-CoA and acetyl-CoA based on the lowered pK_a of the protons on the fluorine-substituted carbon.

Our phylogenetic analysis places FlK in closer evolutionary proximity to dehydratases than thioesterases. At the structural level, the similarities between FlK and the MaoC dehydratases are apparent. Structural alignment of FlK [*23, 24*] and the (*R*)-specific enoyl-CoA hydratase (PhaJ) from Aeromonas caviae [*39*] shows that the overall folds of the two enzymes are very similar, with a core rmsd of 1.7 Å. The alignment shows the presence of an 'overhanging segment' or lid that covers the active site and closes it off from solvent in both enzymes, although this feature had previously been considered to distinguish the MaoC family from other hot dog families [*39*]. In both enzymes, the lid is functionally important. In PhaJ, it encompasses the hydratase 2 motif that contains the catalytic Asp and His, while in FlK, the lid contains the Phe 33-Phe 36 gate, which closes the active site off from solvent and is involved in substrate affinity [*24*]. Notably, the lid is absent in FlK's closest orthologs.

The active sites of PhaJ and FlK also share many key features. Both are mainly hydrophobic, with the exception of the catalytic residues, and both are able to accommodate short-chain acyl-CoAs. PhaJ's catalytic triad consists of Asp 31, His 36, and Ser 62 [*39*], while FlK's catalytic triad consists of Thr 42, Glu 50, and His 76 [*24*]. The Asp/Glu-His-Ser/Thr catalytic triad is common among many thioesterase and dehydratase members of the hot dog superfamily, and is often spatially permuted within the active site even between members with similar catalytic activities and chemical mechanisms. Site-directed mutagenesis studies of PhaJ and FlK demonstrate that mutation of the catalytic His to Ala leads to a 10^4-10^5-fold decrease in k_{cat} [*39*], while mutation of the catalytic His in hot dog-fold thioesterases tends to have a smaller (2-4000-fold) impact [*37, 40*]. Mutation of the catalytic Asp/Glu affects thioesterases and dehydratases similarly [*24, 37, 39, 40*], leading to a 10^4-10^5-fold decrease in k_{cat}. Mutation of Ser 62 in PhaJ and Thr 42 in FlK to Ala have similar impacts on catalytic activity (100-fold decrease in k_{cat}), and both residues have been proposed to be involved in positioning the other catalytic residues [*24, 39*].

If a dehydratase-like mechanism were operative in direct deprotonation of the fluoroacetyl-CoA substrate, we would expect k_{cat} to be sensitive to deuterium substitution on the α-carbon. On the other hand, if C_α-deprotonation occurred on an acyl-enzyme intermediate instead, the intrinsic kinetic isotope effect (KIE) would be masked by the slower rate of enzyme acylation with the fluorinated substrate. In this scenario, we would only be able to observe an apparent KIE with labeled substrate if deuteration were to slow down C_α-deprotonation such that it becomes at least partially rate-limiting. To test for C_α-deprotonation, we prepared [^{2}H]-labeled acetyl- and fluoroacetyl-CoA and measured the rates of their hydrolysis by FlK. For [$^2H_3$]-acetyl-CoA, deuterium substitution had no impact on the reaction rate (*Figure 3.11*). In contrast, we observed an apparent primary KIE of 2.4 ± 0.1 for hydrolysis of [$^2H_2$]-fluoroacetyl-CoA by FlK (*Figure 3.11B*). The magnitude of the KIE for fluoroacetyl-CoA hydrolysis is on the order of or larger than KIEs measured for other characterized enzymes that catalyze C_α-deprotonation [*41, 42*]. In comparison, enzymes that carry out nucleophilic attack at a carbonyl moiety typically show an inverse isotope effect, if any, with similarly labeled substrates [*43*]. To further confirm the ability of FlK to initiate C_α-deprotonation, we assayed the 3,3,3-trifluoropropionyl-CoA substrate and observed rapid formation of one equivalent of CoA followed by elimination of one equivalent of fluoride, suggesting formation of an enolate or carbanion species [*44*] (*Figure 3.12*). Time-dependent inactivation of the enzyme occurs on the same timescale as CoA formation, which is consistent with stoichiometric enzyme alkylation by the Michael acceptor produced in the active site after loss of fluoride (*Figure 3.12*). Taken together, these data support a model in which

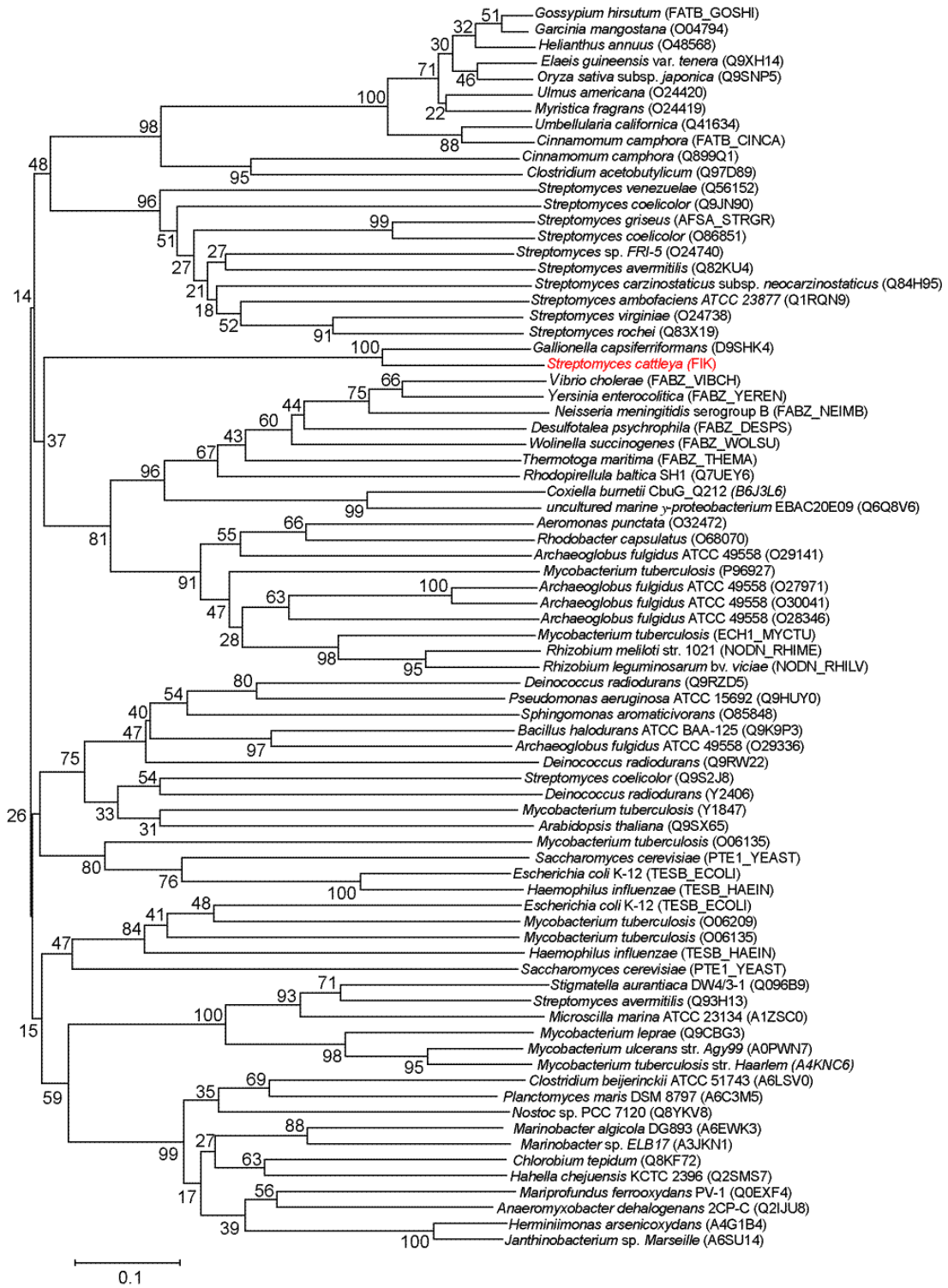


Figure 3.8. *Phylogenetic analysis of FIK homologs. Uncompressed phylogenetic tree for FIK. An alignment of Pfam seed sequences from each family within the hot dog clan was performed using PRALINE. A neighbor-joining tree was then constructed in MEGA 5.0 and tested with 500 bootstrap replicates. Bootstrap values are listed near each branch. InterPro accession numbers are listed in parentheses.*

```
Streptomyces cattleya          -MKDGMRVGERFTHDFVVPPHKTVRHLYPESPEFAEFPEVFATGFMVGLMEWACVRAMAP
Rhodospirillum rubrum          -MRDVLKVGMGETLSFEVPREKTVPFLYPESAEFQAIPEVFATGYMIGLMEWCCVRSLAP
Afipia sp. 1NLS2               -MKPSLTAGSTHRFTYRVPENKTVPHLFPEARDFQIMPHVFATGYMVGLMEWACMDMIRP
Oligotropha carboxidovorans    -MKPTLVAGATHRFTYRVPENKTVPDLFPEAHDFQIMPHVFATGYMVGLMEWACMDLIRP
Hyphomicrobium sp. MC1         -MKDTLHPGAKTQFTYRVPATKTVPHLYPEAHEFQLMPTVFATGYMVGLMEWTCLHIIAP
Hyphomicrobium denitrificans   -MKPSLQAGARTEFSFRVPATKTVPYLYPEAREFQLMPTVFATGFMVGLMEWTCLHILEE
Hyphomicrobium denitrificans   -MKPSLQAGARTQFSFRVPATKTVPHLYPEAREFQLMPTVFATGFMVGLMEWTCLHIIDP
Geobacter sp. M18              -MKDTLAAGIGTTLKFSVPVEKTVPCLYPESALFREMPEVFATGYLVGFIEWACMEALAP
Geobacter bemidjiensis         -MKE-LQVGLKHTFSYLVPKERTVPFLYPESSYFQVMPEVFATGYMVGFMEWACMDALAP
Anaeromyxobacter sp. Fw109-5   -MKSTLAPGVSLTFRYQVPETKTVPHVFPESPRFVEMPQVFATAFMVGLLEWACIEAMQP
Methanosarcina acetivorans     MDSSTLKPGLAYEFRFKIPENKTVPYLYPESPEFQVMPKVFATGFMVGLFEWACIQAINP
Desulfosporosinus youngiae     -MKSTLQSGLYYEFKFTVPENKTVPYLYPESEEFQAMPKVFGTGFMVGLFEWACIKAINP
Desulfosporosinus meridiei     -MKSTLQSGLSYEFKFTVPEDKTVPYLYPESEEFQAMPKVFGTGFMVGLFEWACIKAINP
Desulfosporosinus orientis     -MKSTLQSGLDYEFKFTVPDNKTVPYLYPEAEEFQAMPKVFGTGFMVGLFEWACIKAINP
Gallionella capsiferriformans -MKDTLKPGIRFEHKYLVPANKTVPALYPESPEFLAMPEVFATGFMVGFLEWACIMAIKP
Sideroxydans lithotrophicus    -MKDTLKPGIRYEHRFLVPSSKTVPALYPEAEEFLAMPEVFATGFLVGFLEWACIKCINP
                                    :  *      : :*  .**  ::**:  *  :* **.*.:::*::** *:  :

Streptomyces cattleya          YLE-PGEGSLGTAICVTHTAATPPGLTVTVTAELRSVEGRRLSWRVSAHDGVDEIGSGTH
Rhodospirillum rubrum          ALE-DGEGSLGIAINVSHLAPTPPGARVVVEAKIIAIDGRKVSWHVVARDEVDLIGEGTI
Afipia sp. 1NLS2               HLE-EGEGTLGTLINVNHTAATPPGLTIHVDVECTDVKGKLLRFNVKAHDGVDVIGEGRH
Oligotropha_carboxidovorans    HLE-DGEGTLGTLIDVNHTAATPPGLTINVDVECLEVKGKWSRFKVKAHDGIDVIGEGFH
Hyphomicrobium sp. MC1         HLD-KGEGSLGVHINVSHLAATVPGQTVTVDAECTKVAGRRLYFHVKAHDGIDLIGEGEH
Hyphomicrobium denitrificans   HLD-QGEGSLGVHINVSHVAATVPGQTVTVEAECTKVAGRRIFFHVKAHDGIDLIGEGEH
Hyphomicrobium denitrificans   HLD-PGEGSLGVHINVSHVAATVPGQTVTVEAECTKVAGRRIFFHVKAHDGIDLIGEGEH
Geobacter sp. M18              YLE-EDERSVGTMINVTHSAATPPGMEVTAQVRCVEVTGKRTVWEIEVHDQADLISKGTH
Geobacter bemidjiensis         YLD-EGERTVGTMINVTHEAATPAGMEVTATVTLVEVDGKRTVWEIEARDEVEVIGRGRH
Anaeromyxobacter sp. Fw109-5   HLD-GGEQSVGTGIWVTHGAATPPGFTVTVDVAVTKVEGRRLTFSVRAHDGVDAICEGTH
Methanosarcina acetivorans     YLDFPAEQTVGTDVRLSHSAATPPGLTVTVKIKLEKIEGRKLTFSIIADDGVDKISEGTH
Desulfosporosinus youngiae     HLDWPNEQTVGTDVKLSHLAATPPGLTVTVKLRLEKIEGKKLFFHVEAHDGIDLISEGTH
Desulfosporosinus meridiei     HLDWPNEQTVGTDVKLSHIAATPPGFTVTVKLRLEKIEGKKLLFHVEAHDGVDQISEGTH
Desulfosporosinus orientis     HLDWPNEQTVGTDVKLSHIAATPPGFTVTVKLRLEKIDGKRLLFHVEAHDGVDLISEGTH
Gallionella capsiferriformans HLDWPEEQSVGTHINVSHEAATPPGLEVTASVELTIVDGRRLTFAVSAHDGVDTIARGTH
Sideroxydans lithotrophicus    HIDWPAEQTVGTHINVSHQAATPPGLEVTALVELVEVDGRKLVFQVEAHDGVEVISKGRH
                                ::   * ::*  : :.* *.* .*  : .      : *.   : : . *  : *  *

Streptomyces cattleya          ERAVIHLEKFNAKVRQKTPAG------------
Rhodospirillum rubrum          GRAVVRWSSFTQRLAEKAATIAARRATKPSGTD
Afipia sp. 1NLS                ERYAVMWDKFTARVSKKAAKAGVAA--------
Oligotropha carboxidovorans    ERYAVMWDKFKARVSEKAVKAGVAA--------
Hyphomicrobium sp. MC1         QRMVVNWEKFEQRVNEKAKIARLAPITRGTV--
Hyphomicrobium denitrificans   ERVVVDWEKFEQRVNGKAKLARLSPITRGTG--
Hyphomicrobium denitrificans   QRMIVDWEKFEQRVNDKAKRARLSPITRGTG--
Geobacter sp. M18              ERFTIRLEQFKSRLKTKAEAAGIIIA-------
Geobacter bemidjiensis         ERFVIDYEKFSKRVAAKGNK-------------
Anaeromyxobacter sp. Fw109-5   ERFVIDRARFDRKIQEKLAASTSC---------
Methanosarcina acetivorans     ERFIIDAAKFNSKAEAKAKNANN----------
Desulfosporosinus youngiae     ERFVIDAAKFNEKVNRKSKTESVTE--------
Desulfosporosinus meridiei     ERFIINAAKFNEKVTRKSEIKA-----------
Desulfosporosinus orientis     ERFIIDAAKFNEKVTRKSEIERITE--------
Gallionella capsiferriformans ERYVINKEKFDNKLRDKREKINQR---------
Sideroxydans lithotrophicus    ERFIINREKFEAKIGEKMRRSDT----------
                                *  :    *  .   *
```

Figure 3.9. *Multiple sequence alignment of FIK and its closest homologs. The sequences of FIK's closest orthologs were retrieved by BLAST searching the NCBI database. Orthologs with >50% sequence identity were aligned to FIK using MUSCLE for sequence comparison. Residue 36 (FIK numbering) is shown in red.*

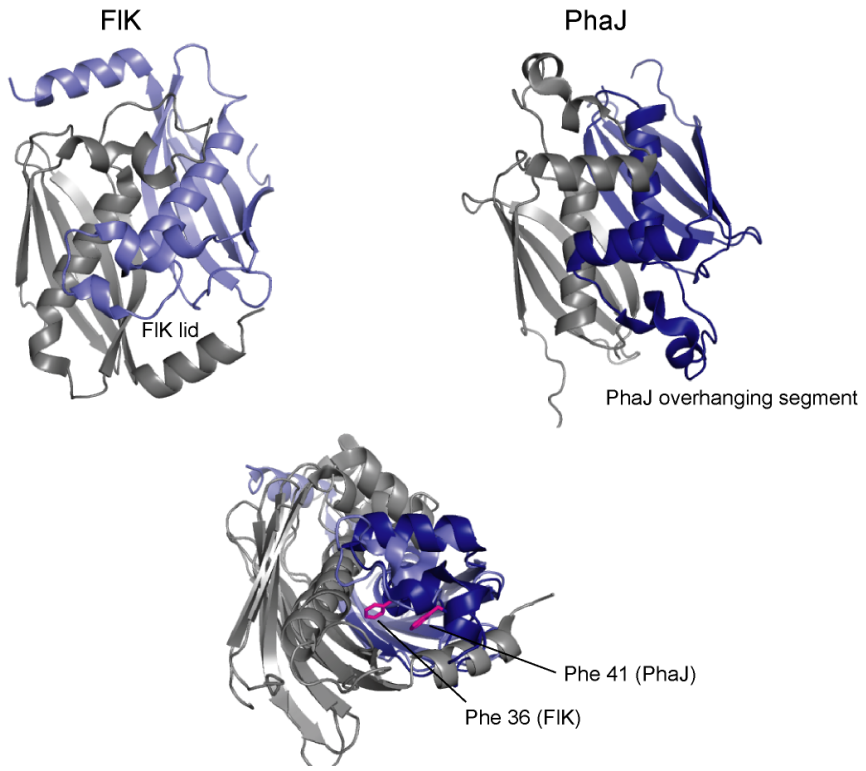


***Figure 3.10**. Structural comparison of FlK to PhaJ. Overall folds of FlK showing lid (top left; PDB ID: 3P2Q) and PhaJ showing overhanging segment (top right; PDB ID: 1IQ6). A structural alignment of FlK and PhaJ (bottom) was performed by secondary structure matching using Coot and shows a conserved phenylalanine in both the FlK lid and the PhaJ overhanging segment (FlK, light blue, Phe 36 magenta; PhaJ, navy, Phe 41, magenta).*

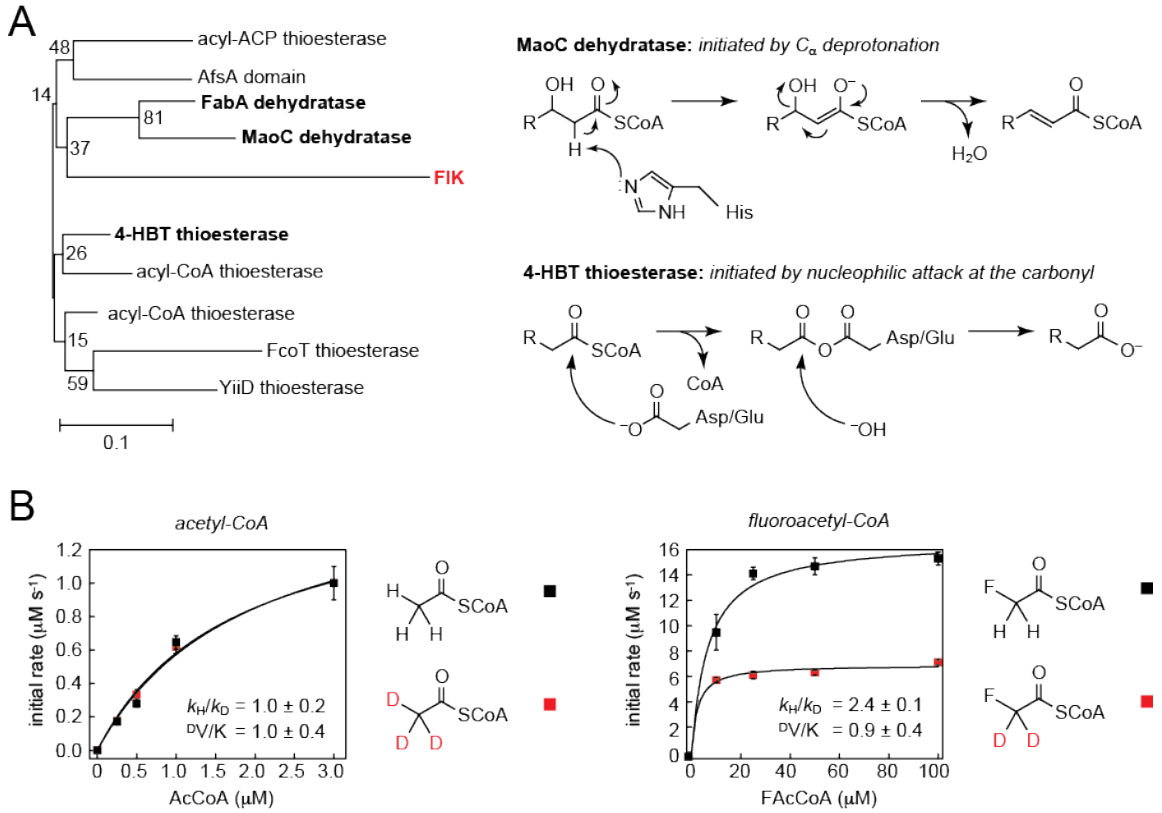


Figure 3.11. *Breakdown of the fluoroacetyl-enzyme intermediate is accelerated by C_α-deprotonation. (A) Structure-based sequence alignment and phylogenetic analysis of the hot dog superfamily indicate that FIK shares functional features with both the 4-HBT thioesterases, which utilize an enzyme anhydride intermediate in hydrolysis, and the MaoC dehydratases, which initiate dehydration by C_α-deprotonation using a catalytic histidine residue. (B) Deuteration at the α–position of fluoroacetyl-CoA leads to a 2.4-fold kinetic isotope effect (KIE) on k_{cat} for FIK-catalyzed fluoroacetyl-CoA hydrolysis, while deuteration at the α-position of acetyl-CoA has no impact on the rate constant.*

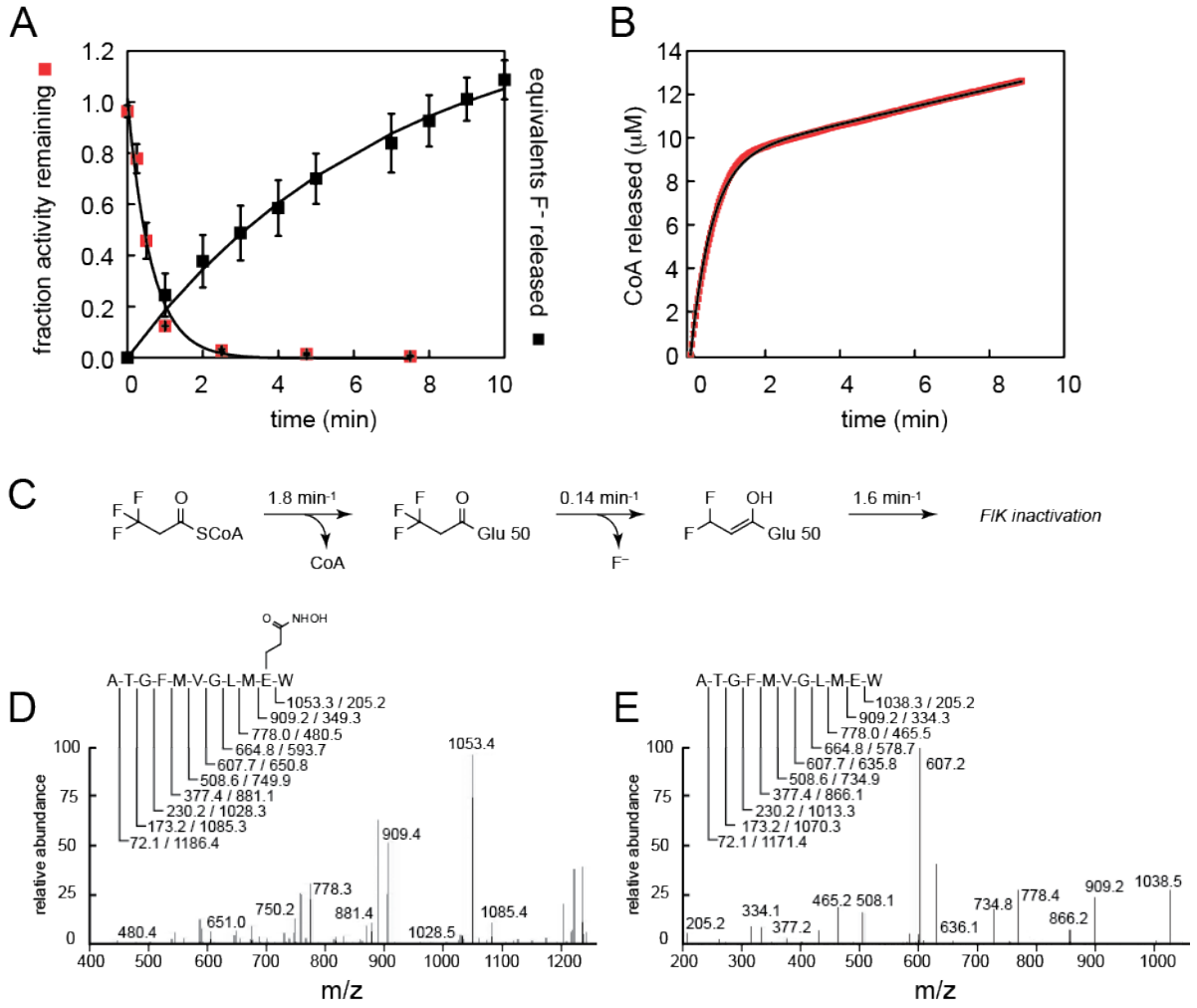


***Figure 3.12**. 3,3,3-Trifluoropropionyl-CoA hydrolysis, fluoride release, and FIK inactivation. (A) Fluoride release (black) and fraction FIK activity remaining after treatment with 3,3,3-trifluoropropionyl-CoA (red). Fluoride release was monitored using a fluoride electrode. Fraction activity remaining was determined by diluting an aliquot of 3,3,3-trifluoropropionyl-CoA-treated FIK into a reaction mixture containing fluoroacetyl-CoA at various times. (B) Time course for CoA release from 3,3,3-trifluoropropionyl-CoA (1 mM) catalyzed by FIK (10 μM) as monitored using DTNB. The time course shows rapid formation of 1 equivalent of CoA, followed by a much slower rate that is identical to the background rate of uncatalyzed hydrolysis, indicating inactivation of the enzyme after one turnover. Inactivation occurs with a rate constant identical to that of CoA formation (or enzyme acylation) as the active site cannot accommodate another substrate. (C) Modified E50 peptide resulting from hydroxylamine-trapping of FIK treated with the 3,3,3-trifluoropropionyl-CoA substrate. (D) Unmodified peptide from FIK treated with hydroxylamine in the absence of an acyl-CoA substrate. (E) Kinetic scheme showing rate constants derived from plots in (A) and (B).*

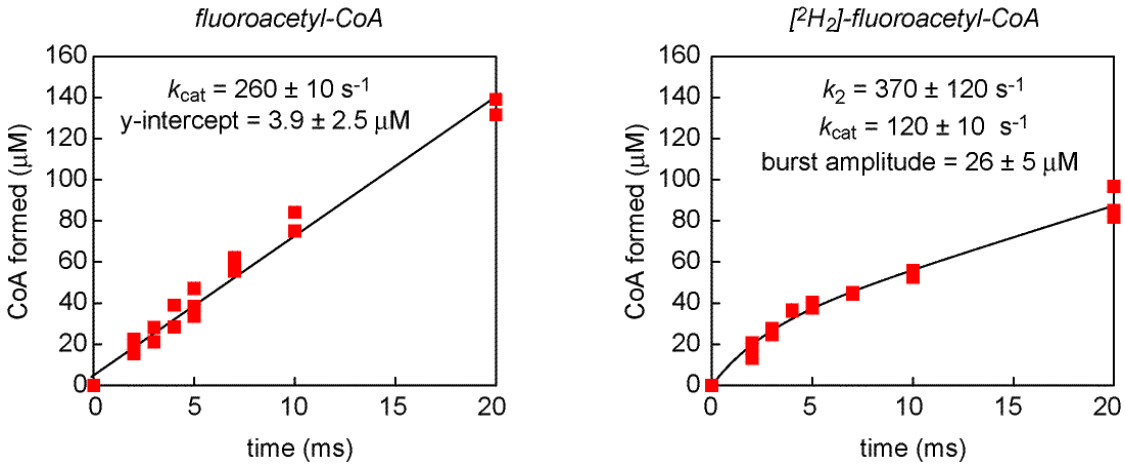


Figure 3.13. *Pre-steady-state kinetics of [$^2H_2$]-fluoroacetyl-CoA hydrolysis by FlK. Pre-steady-state measurements show that FlK (25 μM) develops burst behavior upon hydrolysis of [$^2H_2$]-fluoroacetyl-CoA, indicating both that an acyl-enzyme intermediate is formed and that enzyme deacylation becomes partially rate-limiting, consistent with the observed KIE.*

substrate selectivity is determined by the existence of a C_α-deprotonation pathway that is only available to activated substrates.

We further examined the pre-steady-state kinetic behavior of FlK with respect to the [$^2H_2$]-fluoroacetyl-CoA substrate and observed a burst phase of CoA formation during the first turnover of the reaction (*Figure 3.13*). The onset of burst kinetics with the deuterated substrate suggests that isotopic substitution slows steps in the catalytic cycle subsequent to the formation of CoA and that C_α-deprotonation occurs on an intermediate rather than directly upon the fluoroacetyl-CoA substrate itself. Indeed, the absence of a primary $^D(V/K)$ isotope on fluoroacetyl-CoA hydrolysis, within error, is consistent with a model where the isotopically-sensitive step takes place after the first irreversible step, which would be enzyme acylation and formation of free CoA. These behaviors are consistent with the observed KIE and potentially with the existence of an acyl-enzyme intermediate with the fluorinated substrate. These results imply that C_α-deprotonation of the fluorinated substrate occurs subsequent to formation of an enzyme anhydride and that the intrinsic KIE for this step may be larger than the apparent KIE measured with [$^2H_2$]-fluoroacetyl-CoA.

Another testable prediction from our model is that if H76 plays an analogous general base role in FlK as in the dehydratases (*Figure 3.11A*), selectivity for fluoroacetyl-CoA over acetyl-CoA should be lost when this residue is mutated because the pathway to C_α-deprotonation of the fluorinated substrate would be eliminated. Consequently, both substrates would react via the same mechanism of rate-limiting water attack at the substrate carbonyl, which we would observe experimentally as the disappearance of the KIE for fluoroacetyl-CoA hydrolysis. If an acyl-enzyme intermediate is formed with the fluorinated substrate, we would also expect to see burst kinetics with this mutant. Interestingly, our previous mutagenesis experiments showed that H76 is the most important residue in the catalytic triad and that mutation to alanine resulted in an unexpectedly large kinetic defect compared to other members of the hot-dog thioesterase superfamily [*24*]. The H76A mutant was assayed and confirmed to exhibit burst kinetics in the pre-steady-state with both substrates, which implies that an acyl-enzyme intermediate may also be formed with fluoroacetyl-CoA (*Figure 3.14B*). The insensitivity of k_{cat} to the presence of

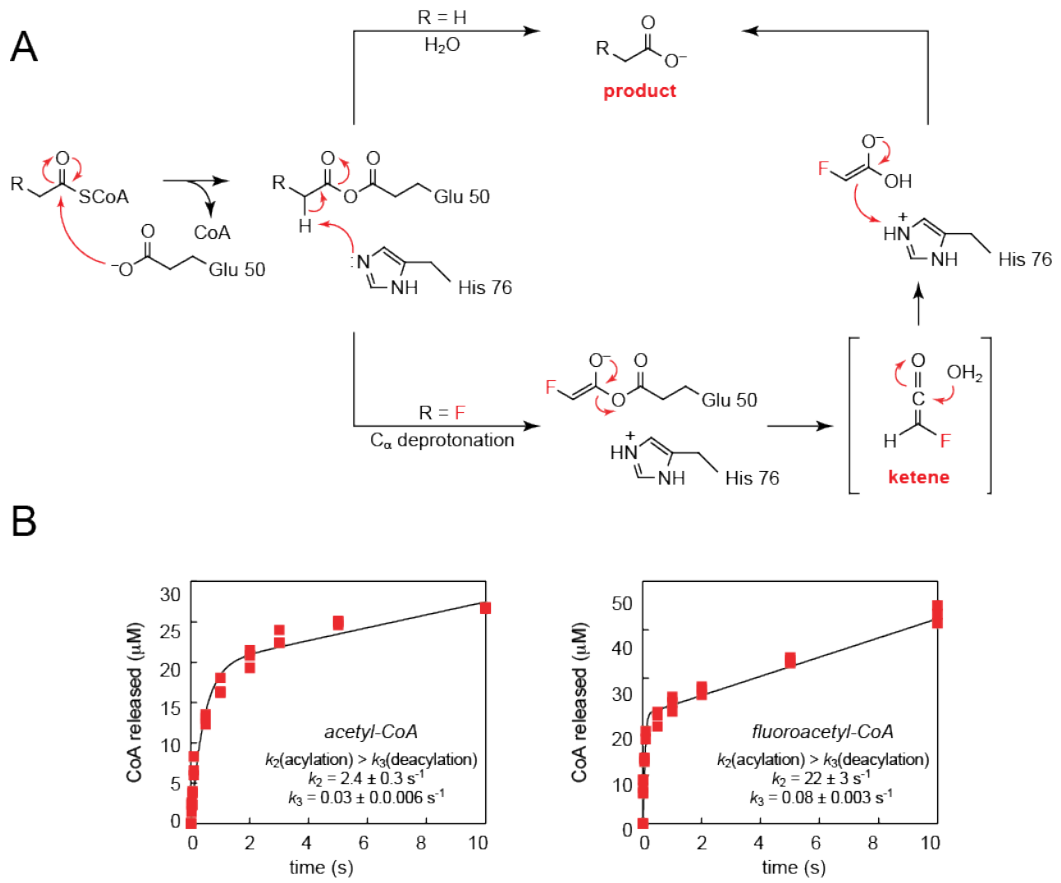
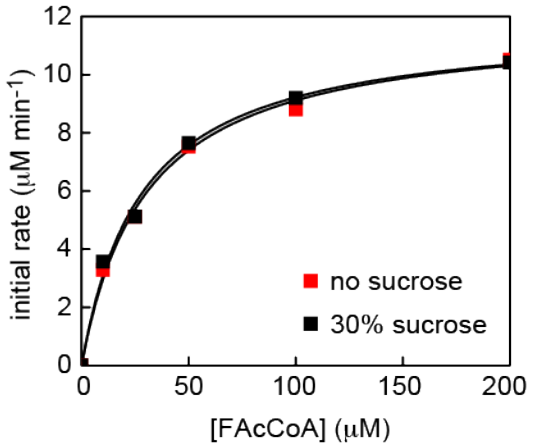


***Figure 3.14**. A change in catalytic mechanism accounts for FlK's specificity for fluoroacetyl-CoA. (A) Proposed enzymatic mechanisms for hydrolysis of fluoroacetyl-CoA versus acetyl-CoA. While acetyl-CoA is hydrolyzed through an acyl-enzyme intermediate, fluoroacetyl-CoA could be either directly deprotonated (direct) or deprotonation could occur from a fluoroacetyl-enzyme intermediate in either a stepwise or concerted fashion. (B) Pre-steady-state kinetic traces for hydrolysis of acetyl-CoA (top) and fluoroacetyl-CoA (bottom) by FlK-H76A.*

sucrose suggests again that the observed burst behavior arises from rate-limiting enzyme deacylation rather than a diffusion-controlled physical event (*Figure 3.15A*). Strikingly, this mutant demonstrates no selectivity beyond what is expected from polarization for the fluorinated substrate over the non-fluorinated substrate in k_{cat} or in the individual rates of enzyme acylation and deacylation, indicating that the selectivity for fluorine is directly related to the role of H76. Moreover, the KIE observed with [$^2H_2$]-fluoroacetyl-CoA is abolished in the H76A mutant, showing that H76 is directly involved in C_α-deprotonation (*Figure 3.15B*). Taken together, these experiments highlight the importance of H76 and the C_α-deprotonation pathway in controlling the fluorine specificity of FlK, which is conferred by catalytic selectivity of the enzyme that distinguishes between the two substrates.

A ketene intermediate in thioester hydrolysis. In FlK-catalyzed hydrolysis of the fluorinated thioester, the most likely mechanism that accounts for the observed rapid rate of enzyme deacylation following proton abstraction is the formation of a fluoroketene intermediate on FlK followed by rapid hydration (*Figure 3.14A*). Although unprecedented for enzymatic thioester hydrolysis of a physiological substrate, the utilization of a C_α-deprotonation strategy for breakdown of the enzyme-anhydride intermediate in FlK is reminiscent of well-characterized

A



	% sucrose	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
fluoroacetyl-CoA	0	0.04 ± 0.001	30 ± 4	1300 ± 200
	30	0.04 ± 0.001	28 ± 3	1300 ± 200

B

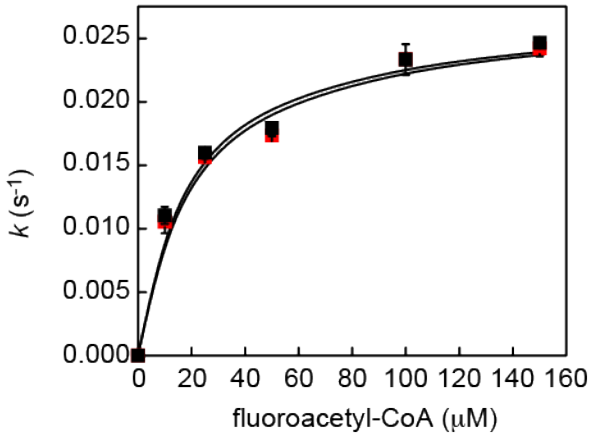


***Figure 3.15**. Characterization of FlK-H76A. (A) Michaelis-Menten curve for FlK-H76A-catalyzed hydrolysis of fluoroacetyl-CoA in the presence (black) and absence (red) of 30% sucrose. (B) FlK-H76A-catalyzed hydrolysis of fluoroacetyl-CoA (black) and [$^2H_2$]-fluoroacetyl-CoA (red). Upon introduction of the H76A mutation, the KIE indicating C_α-deprotonation is no longer observed.*

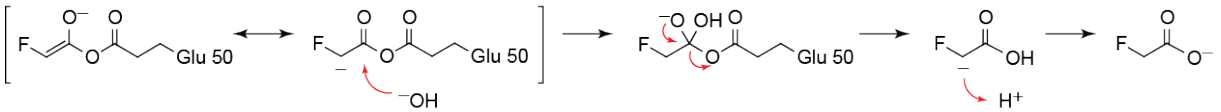


***Figure 3.16**. Alternative mechanism for enolate breakdown. An alternative mechanism for breakdown of the anhydride enolate intermediate would involve nucleophilic attack of hydroxide or water at the carbonyl of a carbanion or enolate, both which should be slower than direct attack at a neutral anhydride as observed for acetyl-CoA based on the rates of the hydrolysis of the corresponding thioester.*

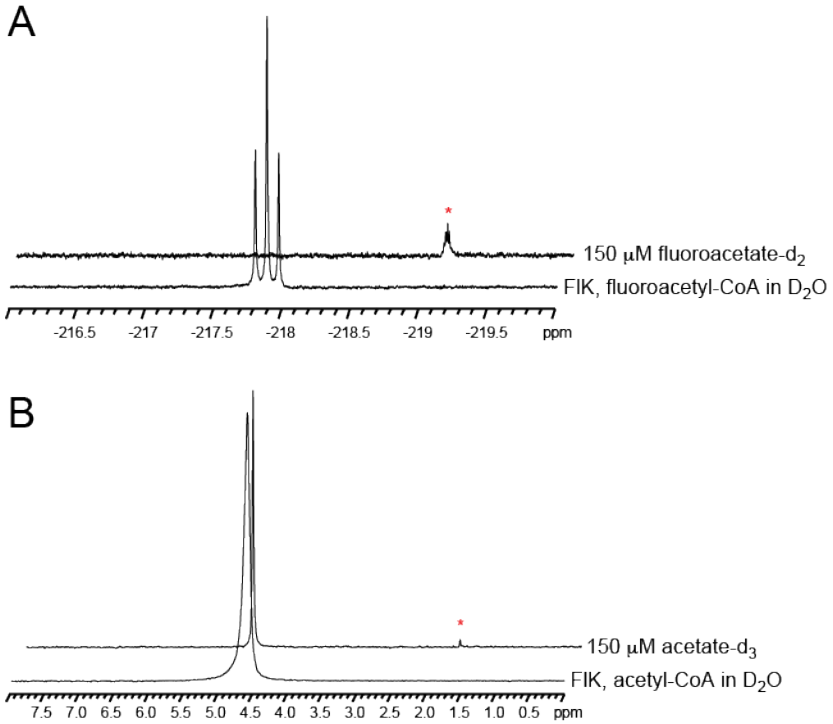


***Figure 3.17**. Quantification of exchange of α-protons with D_2O. Deuterium incorporation into fluoroacetate or acetate was analyzed by NMR following FlK-catalyzed hydrolysis of the corresponding thioester. (A) ^{19}F NMR analysis indicates that <1% of fluoroacetyl-CoA α-protons exchange with solvent as compared to a standard corresponding to 1% exchange. (B) ^{2}H NMR analysis indicates that <1% of acetyl-CoA α-protons exchange with solvent as compared to a standard corresponding to 1% exchange*

chemical model systems with acidic α-protons that are known to pass through a ketene intermediate [*45, 46*]. These systems demonstrate a characteristic acceleration in the rate of thioester hydrolysis under conditions where the C_α-deprotonation pathway is accessible, thereby outcompeting the rate of thioester hydrolysis by nucleophilic attack at the carbonyl group at high pH despite the increase in hydroxide concentration [*45-47*].

An alternative but less probable mechanism would require attack of water directly on the enolate or carbanion species (*Figure 3.16*), which is predicted to be slower than direct attack of water on the neutral enzyme-anhydride as observed for acetyl-CoA based on work in chemical model systems [*48*] (*Figure 3.14A*). This alternative mechanism seems unlikely to explain the observed rate acceleration for FlK deacylation as the presence of carbanion character would be expected to deactivate the carbonyl toward nucleophilic attack based on charge repulsion between the approaching nucleophile and the carbanion. Indeed, in chemical model systems involving breakdown of an enolate through a carbanion intermediate, the rate of reaction is 1-2

orders of magnitude slower than the expected rate of a ketene-based mechanism [*48*]. The formation of a ketene during the FlK catalytic cycle is consistent with the lack of D_2O wash-in at the α-position of the fluoroacetate product (*Figure 3.17*), which may imply that hydration takes place more rapidly than proton exchange at H76 as well as the order of magnitude slower rate of turnover of cyanoacetyl-CoA, which chemical model systems suggest is hydrolyzed via nucleophilic attack on the carbanion species [*47, 48*].

3.4 Conclusion

Streptomyces cattleya is faced with an unusually difficult problem in substrate selectivity, where a single fluorine substituent must be recognized over hydrogen in order to detoxify low levels of endogenous fluoroacetyl-CoA while maintaining high concentrations of the cellular acetyl-CoA pool. The level of discrimination displayed by FlK is surprisingly high given the promiscuity of the other characterized members of the hot dog-fold thioesterase superfamily to which FlK belongs [*49, 50*] and the strong driving force for thioester hydrolysis ($\Delta G \sim -8$ kcal/mol) [*51*]. In this work, we show that the major mode of FlK substrate selectivity is based not on molecular recognition of fluorine but instead on exploitation of a unique chemical mechanism that is accessible only to the fluorinated substrate.

The selectivity for fluorine displayed by FlK is conferred by the lowered pK_a of the fluoracetyl-CoA α-protons that allow access to an unexpected but kinetically advantageous C_α-deprotonation pathway rather than the slower pathway provided by direct reaction at the carbonyl group, as is observed for acetyl-CoA. While C_α-deprotonation itself is observed in many systems [*41, 42, 52, 53*], especially those involved in acyl-CoA metabolism [*53*], its use in initiating thioester hydrolysis of a native biological substrate is unprecedented. In this regard, the unusual reactivity of FlK allows us to explore the existence of a new chemical species in enzyme catalysis as it strongly suggests a ketene intermediate in the FlK catalytic cycle based on similarities to related chemical model systems [*45-47*]. Although precedence for ketene intermediates exists for enzymatic reactions with mechanism-based inhibitors or non-physiological substrates [*54-61*], there is not yet any experimental evidence for their involvement in normal catalytic cycles. For example, a ketene-based mechanism has been proposed in the catalytic cycle of glycine reductase, but it represents only one of several mechanistic possibilities that remain to be distinguished experimentally [*62*]. FlK may therefore provide the first example of a physiological ketene intermediate derived from a native substrate in biological catalysis despite their rich history in organic chemistry and the strong interest in their discovery in enzyme catalysis.

Our kinetic and biochemical data suggest that an enzyme-anhydride intermediate is formed for the nonfluorinated substrate, which is consistent with the observation of a similar anhydride intermediate for the 4-HBT thioesterases [*37, 38*]. Although a simpler mechanistic possibility exists for the fluorinated substrate where C_α-deprotonation would occur directly on the fluoroacetyl-CoA substrate itself with concomitant formation of CoA, the burst phase kinetics observed with [$^2H_2$]-fluoroacetyl-CoA and the H76A mutant suggest that covalent catalysis is also utilized with the fluorinated substrate. The formation of an anhydride with the fluorinated substrate may be related to preserved evolutionary relationships with other thioesterases in the hot-dog fold superfamily but also serves to provide a more activated leaving group for ketene formation.

In summary, FlK provides a unique platform to examine the evolution of substrate selectivity in a native system where natural selection has been driven with respect to both substrates, which are estimated to be present simultaneously in the cell under saturating conditions. While there are many examples of catalytic selectivity in enzyme superfamilies where a shared intermediate can partition between divergent downstream pathways [*63, 64*], FlK utilizes reaction partitioning to distinguish between substrates rather than to control product distribution. Indeed, the same overall transformation is catalyzed by FlK with both substrates even though two distinct reaction mechanisms occur within the same active site. Interestingly, control over reaction partitioning by the substrate is also reminiscent of the function of mechanism-based inhibitors, many of which contain fluorine themselves [*65*]. However, these alternative substrates have been optimized by design [*65*] or evolution [*66*] to inactivate the target enzyme rather than providing a natural mechanism for exclusion of the incorrect substrate as is observed for FlK. Taken together, FlK provides insight into the role of catalytic diversity for controlling substrate selectivity at the cellular level rather than in the evolution of new metabolic transformations.

3.5 Acknowledgments

We thank the Marletta laboratory for use of their rapid chemical quench instrument and Lori Kohlstaedt (U.C. Berkeley Vincent J. Coates Proteomics Facility) for FlK mass spectrometry analysis. The College of Chemistry NMR Facility at the University of California, Berkeley is supported in part by the National Institutes of Health (1S10RR023679-01 and S10 RR16634-01). A.M.W. also acknowledges the support of a National Institutes of Health NRSA Training Grant (1 T32 GMO66698) and National Science Foundation Graduate Research Fellowship. This work was funded by generous support from U.C. Berkeley and a National Institutes of Health New Innovator Award (1 DP2 OD008696).

3.8 References

1. W. J. Albery and J. R. Knowles, Evolution of enzyme function and the development of catalytic efficiency, *Biochemistry* **1976**, *15*, 5631-5640.
2. M. I. Page and W. P. Jencks, Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect, *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*, 1678-1683.
3. J. G. Voet and R. H. Abeles, The mechanism of action of sucrose phosphorylase. Isolation and properties of a β-linked covalent glucose-enzyme complex, *J. Biol. Chem.* **1970**, *245*, 1020-1031.
4. F. C. Kokesh and F. H. Westheimer, A reporter group at the active site of acetoacetate decarboxylase. II. Ionization constant of the amino group, *J. Am. Chem. Soc.* **1971**, *93*, 7270-7274.
5. J. G. Zalatan and D. Herschlag, The far reaches of enzymology, *Nat. Chem. Biol.* **2009**, *5*, 516-520.
6. P. J. O'Brien and D. Herschlag, Catalytic promiscuity and the evolution of new enzymatic activities, *Chem. Biol.* **1999**, *6*, R91-R105.
7. J. Lee, M. Natarajan, V. C. Nashine, M. Socolich, T. Vo, W. P. Russ, S. J. Benkovic and R. Ranganathan, Surface sites for engineering allosteric control in proteins, *Science* **2008**, *322*, 438-442.

8. J. T. Bridgham, E. A. Ortlund and J. W. Thornton, An epistatic ratchet constrains the direction of glucocorticoid receptor evolution, *Nature* **2009**, *461*, 515-519.

9. K. A. Reynolds, R. N. McLaughlin and R. Ranganathan, Hot spots for allosteric regulation on protein surfaces, *Cell* **2011**, *147*, 1564-1575.

10. E. Fischer, Einfluss der configuration auf die wirkung der enzyme, *Chem. Ber.* **1894,** *27*, 2985-2993.

11. D. E. Koshland, Application of a theory of enzyme specificity to protein synthesis, *Proc. Natl. Acad. Sci. U. S. A.* **1958**, *44*, 98-104.

12. D. D. Boehr, R. Nussinov and P. E. Wright, The role of dynamic conformational ensembles in biomolecular recognition, *Nat. Chem. Biol.* **2009**, *5*, 789-796.

13. M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch and E. Inamine, Biosynthesis of fluorothreonine and fluoroacetic acid by the thienamycin producer, *Streptomyces cattleya*, *J. Antibiot. (Tokyo)* **1986**, *39*, 259-265.

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

15. K. Muller, C. Faeh and F. Diederich, Fluorine in pharmaceuticals: Looking beyond intuition, *Science* **2007**, *317*, 1881-1886.

16. J. C. Biffinger, H. W. Kim and S. G. DiMagno, The polar hydrophobicity of fluorinated compounds, *Chembiochem* **2004**, *5*, 622-627.

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

18. G. W. Gribble, The diversity of naturally produced organohalogens, In *Natural Production of Organohalogen Compounds: The Handbook of Environmental Chemistry*, G. W. Gribble, Ed., Springer: Berlin, **2003**, Vol. 3, pp. 1-15.

19. D. O'Hagan and D. Harper, Fluorine-containing natural products, *J. Fluor. Chem.* **1999**, *100*, 127-133.

20. D. D. Clarke, Fluoroacetate and fluorocitrate: Mechanism of action, *Neurochem. Res.* **1991**, *16*, 1055-1058.

21. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme-inhibitor complex, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13699-13703.

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

23. M. V. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. Dry, P. F. Leadlay, J. B. Spencer and T. L. Blundell, Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK, *J. Biol. Chem.* **2010**, *285*, 22495-22504.

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

25. M. Shimazu, L. Vetcher, J. L. Galazzo, P. Licari and D. V. Santi, A sensitive and robust method for quantification of intracellular short-chain coenzyme A esters, *Anal. Biochem.* **2004**, *328*, 51-59.

26. M. Yu, L. P. de Carvalho, G. Sun and J. S. Blanchard, Activity-based substrate profiling for Gcn5-related *N*-acetyltransferases: the use of chloroacetyl-coenzyme A to identify protein substrates, *J. Am. Chem. Soc.* **2006**, *128*, 15356-15357.

27. R. W. Taft, Jr., Polar and steric substituent constants for aliphatic and *o*-benzoate groups from rates of esterification and hydrolysis of esters, *J. Am. Chem. Soc.* **1952**, *74*, 3120-3128.

28. C. Hansch and A. Leo, *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley: New York, **1979**.

29. J. K. Eng, A. L. MacCormack and J. R. Yates, III, An approach to correlated tandem mass spectral data of peptides with amino acid sequences in a protein database, *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976-989.

30. D. L. Tabb, W. H. McDonald and J. R. Yates, III, DTASelect and Contrast: Tools for assembling and comparing protein identifications from shotgun proteomics, *J. Proteome Res.* **2002**, *1*, 21-26.

31. V. A. Simossis and J. Heringa, PRALINE: A multiple sequence alignment toolbox that integrates homology-extended and secondary structure information, *Nucleic Acids Res.* **2005**, *33*, W289-294.

32. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei and S. Kumar, MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.* **2011**, *28*, 2731-2739.

33. K. A. Johnson, Transient-state kinetic analysis in enzyme reaction pathways, *The Enzymes* **1992**, *20*, 1-61.

34. D. L. Purich, Covalent enzyme-substrate compounds: detection and catalytic competence, *Methods Enzymol.* **2002**, *354*, 1-27.

35. B. S. Hartley and B. A. Kilby, The reaction of *p*-nitrophenyl esters with chymotrypsin and insulin, *Biochem. J.* **1954**, *56*, 288-297.

36. B. J. Wong and J. A. Gerlt, Divergent function in the crotonase superfamily: an anhydride intermediate in the reaction catalyzed by 3-hydroxyisobutyryl-CoA hydrolase, *J. Am. Chem. Soc.* **2003**, *125*, 12076-12077.

37. Z. Zhuang, J. Latham, F. Song, W. Zhang, M. Trujillo and D. Dunaway-Mariano, Investigation of the catalytic mechanism of the hotdog-fold enzyme superfamily *Pseudomonas* sp. Strain CBS3 4-hydroxybenzoyl-CoA thioesterase, *Biochemistry* **2012**, *51*, 786-794.

38. F. Song, J. B. Thoden, Z. Zhuang, J. Latham, M. Trujillo, H. M. Holden and D. Dunaway-Mariano, The Catalytic Mechanism of the Hotdog-fold Enzyme Superfamily 4-Hydroxybenzoyl-CoA Thioesterase from *Arthrobacter* sp. Strain SU, *Biochemistry* **2012**, *51*, 7000-7016.

39. T. Hisano, T. Tsuge, T. Fukui, T. Iwata, K. Miki and Y. Doi, Crystal structure of the (*R*)-specific enoyl-CoA hydratase from *Aeromonas caviae* involved in polyhydroxyalkanoate biosynthesis, *J. Biol. Chem.* **2003**, *278*, 617-624.

40. F. Song, Z. Zhuang and D. Dunaway-Mariano, Structure-activity analysis of base and enzyme-catalyzed 4-hydroxybenzoyl coenzyme A hydrolysis, *Bioorg. Chem.* **2007**, *35*, 1-10.

41. B. J. Bahnson and V. E. Anderson, Isotope effects on the crotonase reaction, *Biochemistry* **1989**, *28*, 4173-4181.

42. D. R. Palmer, B. K. Hubbard and J. A. Gerlt, Evolution of enzymatic activities in the enolase superfamily: partitioning of reactive intermediates by (*D*)-glucarate dehydratase from *Pseudomonas putida*, *Biochemistry* **1998**, *37*, 14350-14357.

43. G. Lehrmann, D. Quinn and E. H. Cordes, Kinetic α-deuterium isotope effects for acylation of chymotrypsin by 4-methoxyphenyl formate and for deacylation of formylchymotrypsin, *J. Am. Chem. Soc.* **1980**, *102*, 2491-2492.

44. J. Stubbe and R. H. Abeles, Biotin carboxylations--concerted or not concerted? That is the question!, *J. Biol. Chem.* **1977**, *252*, 8338-8340.

45. B. Holmquist and T. C. Bruice, The carbanion (E1cB) mechanism of ester hydrolysis. I. Hydrolysis of malonate esters, *J. Am. Chem. Soc.* **1969**, *91*, 2993-3002.

46. B. Holmquist and T. C. Bruice, The carbanion mechanism of ester hydrolysis. II. *o*-Nitrophenyl α-cyano- and α-dimethylsulfonioacetate esters, *J. Am. Chem. Soc.* **1969**, *91*, 3003-3009.

47. R. F. Pratt, The carbanion mechanism (E1cB) of ester hydrolysis. III. Some structure-reactivity studies and the ketene intermediate, *J. Am. Chem. Soc.* **1970**, *92*, 5956-5964.

48. T. C. Bruice, A. F. Hegarty, S. M. Felton, A. Donzel and N. G. Kundu, Aminolysis of esters. IX. The nature of the transition states in aminolysis of phenyl acetates, *J. Am. Chem. Soc.* **1970**, *92*, 1370-1378.

49. Z. Zhuang, F. Song, H. Zhao, L. Li, J. Cao, E. Eisenstein, O. Herzberg and D. Dunaway-Mariano, Divergence of function in the hot dog fold enzyme superfamily: The bacterial thioesterase YciA, *Biochemistry* **2008**, *47*, 2789-2796.

50. N. Kunishima, Y. Asada, M. Sugahara, J. Ishijima, Y. Nodake, M. Miyano, S. Kuramitsu and S. Yokoyama, A novel induced-fit reaction mechanism of asymmetric hot dog thioesterase PAAI, *J. Mol. Biol.* **2005**, *352*, 212-228.

51. K. Burton, The free energy change associated with the hydrolysis of the thiol ester bond of acetyl coenzyme A, *Biochem. J.* **1955**, *59*, 44-46.

52. H. M. Miziorko and M. D. Lane, 3-Hydroxy-3-methylgutaryl-CoA synthase. Participation of acetyl-*S*-enzyme and enzyme-*S*-hydroxymethylgutaryl-SCoA intermediates in the reaction, *J. Biol. Chem.* **1977**, *252*, 1414-1420.

53. S. Thompson, F. Mayerl, O. P. Peoples, S. Masamune, A. J. Sinskey and C. T. Walsh, Mechanistic studies on β-ketoacyl thiolase from *Zoogloea ramigera*: identification of the active-site nucleophile as Cys89, its mutation to Ser89, and kinetic and thermodynamic characterization of wild-type and mutant enzymes, *Biochemistry* **1989**, *28*, 5735-5742.

54. A. L. Maycock, R. H. Suva and R. H. Abeles, Novel inactivators of plasma amine oxidase, *J. Am. Chem. Soc.* **1975**, *97*, 5613-5614.

55. R. B. Westkaemper and R. H. Abeles, Novel inactivators of serine proteases based on 6-chloro-2-pyrone, *Biochemistry* **1983**, *22*, 3256-3264.

56. P. R. Ortiz de Montellano and E. A. Komives, Branchpoint for heme alkylation and metabolite formation in the oxidation of arylacetylenes by cytochrome P-450, *J. Biol. Chem.* **1985**, *260*, 3330-3336.

57. J. N. Blankenship, H. Abu-Soud, W. A. Francisco, F. M. Raushel, D. R. Fischer and P. J. Stang, Mechanism-based inactivation of a bacterial phosphotriesterase by an alkynyl phosphate ester, *J. Am. Chem. Soc.* **1991**, *113*, 8560-8561.

58. R. Wang and C. Thorpe, The reductive half-reaction in acyl-CoA oxidase from *Candida tropicalis*: interaction with acyl-CoA analogues and an unusual thioesterase activity, *Arch. Biochem. Biophys.* **1991**, *286*, 504-510.

59. L. D. Thornburg and J. Stubbe, Mechanism-based inactivation of thymine hydroxylase, an α-ketoglutarate-dependent dioxygenase, by 5-ethynyluracil, *Biochemistry* **1993**, *32*, 14034-14042.

60. G. J. Poelarends, H. Serrano, W. H. Johnson, Jr. and C. P. Whitman, Inactivation of malonate semialdehyde decarboxylase by 3-halopropiolates: evidence for hydratase activity, *Biochemistry* **2005**, *44*, 9375-9381.

61. S. Gilch, M. Vogel, M. W. Lorenz, O. Meyer and I. Schmidt, Interaction of the mechanism-based inactivator acetylene with ammonia monooxygenase of *Nitrosomonas europaea*, *Microbiology* **2009**, *155*, 279-284.

62. J. R. Andreesen, Glycine reductase mechanism, *Curr. Opin. Chem. Biol.* **2004**, *8*, 454-461.

63. P. C. Babbitt, M. S. Hasson, J. E. Wedekind, D. R. Palmer, W. C. Barrett, G. H. Reed, I. Rayment, D. Ringe, G. L. Kenyon and J. A. Gerlt, The enolase superfamily: a general strategy for enzyme-catalyzed abstraction of the α-protons of carboxylic acids, *Biochemistry* **1996**, *35*, 16489-16501.

64. A. C. Eliot and J. F. Kirsch, Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations, *Annu. Rev. Biochem.* **2004**, *73*, 383-415.

65. D. B. Berkowitz, K. R. Karukurichi, R. de la Salud-Bea, D. L. Nelson and C. D. McCune, Use of fluorinated functionality in enzyme inhibitor development: Mechanistic and analytical advantages, *J. Fluor. Chem.* **2008**, *129*, 731-742.

66. M. Groll, R. Huber and B. C. Potts, Crystal structures of Salinosporamide A (NPI-0052) and B (NPI-0047) in complex with the 20S proteasome reveal important consequences of β-lactone ring opening and a mechanism for irreversible binding, *J. Am. Chem. Soc.* **2006**, *128*, 5136-5141.

Chapter 4: *Molecular recognition and kinetic discrimination of the C-F bond in FIK*

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

NMR studies were performed in collaboration with Dr. Jeffrey G. Pelton of the Central California 900 MHz NMR Facility.

4.1 Introduction

The impact of fluorination on lipophilicity, metabolic stability, *in vivo* distribution, and pharmacokinetic properties of small molecules is well understood [*1*, *2*]. In contrast, less is known about the role of organic fluorine in tuning binding affinity and selectivity of small molecules for their protein targets [*3*]. Fluorine's high electronegativity and 'polar hydrophobicity', as well as the highly polarized nature of the carbon-fluorine bond, all represent unique properties that may be exploited in protein-ligand interactions. Enzymatic inhibition studies in carboxypeptidase A [*4*], carbonic anhydrase [*5*], and chymotrypsin [*6*] have demonstrated that fluorine can have energetically significant effects on molecular recognition. In each of these cases, incorporation of fluorine into inhibitor scaffolds was shown to enhance binding affinity based on the larger hydrophobic surface area of the C-F bond compared to the C-F bond, a relatively non-specific effect that could be recapitulated using similarly sized alkyl groups [*7*]. Based on its high electronegativity, fluorine may also influence protein-ligand binding affinity by altering the preferred conformation of the small molecule ligand. Indeed, structure-activity relationship studies around fluoxetine (Prozac) suggest that the steric bulk of the trifluoromethyl group on its phenoxy ring allows adoption of a favorable conformation that promotes binding to the serotonin transporter [*8*, *9*]. In terms of specific recognition of the fluorine substituent itself, the high polarization of the C-F bond creates an opportunity for dipolar interactions to occur with protein-based functional groups. A systematic 'fluorine scan' of thrombin inhibitors has shown that these dipolar interactions may involve orthogonal interactions between the C-F bond and the amide bond or the side chains of arginine or asparagine [*10*, *11*]. Structure-activity relationship studies of quinolone antibiotics suggest that a favorable dipolar interaction may account for the increased potency of ciprofloxacin (Cipro) compared to analogs lacking the fluorine substituent [*12*, *13*].

Despite the strong interest in fluorine-protein interactions, few enzyme systems have evolved to utilize fluorinated substrates, and our understanding of fluorine recognition in evolutionarily optimized systems is less well developed [*3*]. The fluoroacetate-producing bacterium, *Streptomyces cattleya*, provides an interesting model system for exploring native fluorine recognition by protein targets that have been subject to an unusual evolutionary pressure to distinguish between fluorine and other similar substituents. The toxicity of fluoroacetate relies on its activation to fluoroacetyl-CoA and subsequent conversion to the fluorocitrate poison based on the inability of host primary metabolic enzymes to sufficiently discriminate against the conservative replacement of hydrogen with fluorine [*14*, *15*]. In comparison, the characterization of an *S. cattleya* fluoroacetate resistance protein, the fluoroacetyl-CoA thioesterase (FlK) [*16–18*], reveals a remarkably high 10^6-fold preference for the fluorinated substrate compared to acetyl-CoA, a cellularly abundant competitor [*17*]. This selectivity with regard to the α-carbon substituent also distinguishes FlK from other members of the hot dog-fold thioesterase superfamily, which are often found to be quite promiscuous [*19–22*]. Although a large part of this selectivity is derived from the difference in catalytic mechanism used by FlK to hydrolyze these two substrates [*23*], specific molecular recognition of fluorine still represents an important basis for its behavior.

We have therefore set out to define the molecular mechanism of fluorine recognition in a naturally-occurring enzyme that has evolved for this purpose. Using a set of substrate analogs, we show that FlK can specifically recognize fluorine using a hydrophobic 'lid' structure, which is unique among hotdog-fold thioesterases. These studies also show that a mobile residue on the

lid, Phe 36, is involved in providing an entropic driving force for fluorinated substrate binding and controlling the off rate of the substrate. Taken together, these results support a model in which conformational protein dynamics can play an important role in fluorine-based substrate discrimination by increasing the residence time of the fluorinated substrate in the active site.

4.2 Materials and methods

Commercial Materials. Acetyl coenzyme A sodium salt, trifluoroacetic acid (TFA), 5,5´-dithiobis(2-nitrobenzoic acid) (DTNB), magnesium chloride hexahydrate, tris(hydroxymethyl) aminomethane hydrochloride (Tris-HCl), streptomycin sulfate, calcium pantothenate, 2,2-dimethoxypropane, anhydrous *N*, *N*-dimethylformamide (DMF), ethyl chloroformate, ethanolamine, diisopropylethylamine, palladium on carbon (10 wt% loading), acetic anhydride, iodine, sodium thiosulfate, anhydrous magnesium sulfate, *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride, 2,2-dimethyl-1,3-dioxane-4,6-dione, 4-dimethylaminopyridine, sodium fluoroacetate, Selectfluor, and formic acid were purchased from Sigma-Aldrich (St. Louis, MO). Monosodium phosphate monohydrate, disodium phosphate heptahydrate, hexanes, ethyl acetate, methanol, anhydrous methanol, acetonitrile, toluene, triethylamine, adenosine-5'-triphosphate trisodium salt, isopropyl-β-D-thiogalactopyranoside (IPTG), sodium chloride, anhydrous sodium acetate, sodium bicarbonate, potassium carbonate, imidazole, and tris(2-carboxylethyl)phosphine hydrochloride (TCEP) were purchased from Fisher Scientific (Pittsburg, PA). 2-Fluoropropionic acid was purchased from Oakwood Products (West Columbia, SC). Benzyl-*N*-(2-aminoethyl)carbamate hydrochloride and *t*-butyl-γ-aminobutanoate hydrochloride were purchased from Bachem (Torrance, CA). 1-Hydroxybenxzotriazole hydrate and 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU) were purchased from Advanced ChemTech (Louisville, KY). Hydrochloric acid, potassium chloride, and KOD DNA polymerase were purchased from EMD (Darmstadt, Germany). Platinum Taq HF was purchased from Invitrogen (Carlsbad, CA). Complete Mini EDTA-free protease inhibitor was purchased from Roche Applied Science (Penzberg, Germany).

Expression and purification of FlK and variants. FlK and FlK-F36A were expressed and purified as described previously [*17*].

Synthesis of substrates. Fluoroacetyl-CoA was synthesized as described previously [*17*].

Pre-steady state kinetic analysis of FlK and FlK-F36A. Pre-steady-state kinetic experiments were performed using rapid chemical quench followed by HPLC separation of coenzyme A from unhydrolyzed acyl-CoA. FlK (50 μM in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl) was mixed with substrate (750 μM for fluoroacetyl-CoA or 6 mM for acetyl-CoA, in water) using a Chemical Quench Flow Model RQF-3 (KinTek). The reaction was stopped at various times by mixing with 50% TFA to achieve a final concentration of 17% TFA. Quenched samples were analyzed by HPLC on an Agilent Eclipse XDB-C18 (3.5 μm, 3.0 × 150 mm) using a linear gradient from 0 to 100% acetonitrile over 15 min (0.6 mL/min) with 50 mM sodium phosphate, pH 4.5 containing 0.1% TFA as the aqueous mobile phase. Conversion percentages were calculated based on the peak areas for substrate and product based on the CoA absorbance at 260 nm. Plots of coenzyme A released versus time were fit to the equation $[CoA] = A*(1-\exp(-k_2*t)) + V*t$, where A is the burst amplitude, k_2 is the burst-phase rate constant, and V is the steady-state rate. In cases where data could not be fit with a burst phase, they were fit as pseudo-

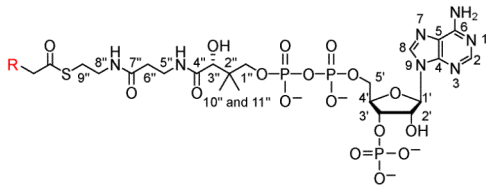
first-order reactions according to the equation $[CoA] = k[FlK]t + [CoA]_0$, where the y-intercept, $[CoA]_0$, is the concentration of CoA released before the first data point, k is the rate constant, and t is time.

Construction of vectors for protein expression. The genes encoding pantothenate kinase (*coaA*), phosphopantheine adenylyltransferase (*coaD*), and dephospho-CoA kinase (*coaE*) were amplified from a saturated overnight culture of *E. coli* BL21(de3)-T1^R using primers coaA F1 and R1, coaD F1 and R1, or coaE F1 and R1 (*Appendix 2*) using either Platinum Taq HF (*coaA* and *coaD*) or KOD polymerase (*coaE*) according to manufacturer protocol. The PCR products were digested with SfoI and XhoI and ligated into the SfoI-XhoI sites of a modified pET23a vector containing an N-terminal His_{10} tag and a tobacco etch virus (TEV) protease cleavage site. The resulting plasmids were verified by Sanger sequencing.

Expression and purification of CoA biosynthetic enzymes. LB containing carbenicillin (50 μg/mL) was inoculated to $OD_{600\ nm} = 0.05$ using an overnight culture of freshly transformed *E. coli* BL21(de3)-T1^R containing the appropriate overexpression plasmid. The cultures were grown at 37ºC at 200 rpm to $OD_{600\ nm} = 0.6$. Protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and cells were incubated at 30 ºC at 200 rpm for 4 h. Cells were harvested by centrifugation at 2,060 × *g* at 4ºC for 15 min. The cell pellet was resuspended in 10 mL/L cell culture of lysis buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 20 mM imidazole, 100 μM phenylmethanesulfonyl fluoride) with one Complete Mini EDTA-free protease inhibitor tablet per 50 mL buffer. Cells were lysed by passing through a French pressure cell at 14,000 psi. Insoluble material was removed by centrifugation at 18,500 × *g* at 4ºC for 20 min. Streptomycin sulfate (10% w/v solution) was added to the supernatant to a final concentration of 1% and precipitated DNA was removed by centrifugation at 18,500 × *g* at 4ºC for 20 min. The supernatant was incubated for 1 h at 4ºC on a nutating mixer with Ni-NTA resin (1 mL/L of cell culture) that had been equilibrated with wash buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 20 mM imidazole). The resin was then pelleted by centrifugation at 500 × *g* for 3 min and the supernatant was carefully decanted. The resin was resuspended in wash buffer, packed into a column, and washed wash buffer (10 column vol) followed by wash buffer containing 8% elution buffer (50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 250 mM imidazole; 10 column vol). Proteins were then eluted with elution buffer (10 column vol) in 5 mL fractions. Protein-containing fractions were pooled, dialyzed into storage buffer (50 mM HEPES, pH 8.0, 250 mM NaCl, 2 mM $MgCl_2$) [*24*], and stored at -80ºC until further use. Final protein concentrations before storage were estimated using the $\varepsilon_{280\ nm}$ calculated by ExPASY ProtParam as follows: His_{10}-PanK: 16 mg/mL ($\varepsilon_{280\ nm} = 45{,}380\ M^{-1}\ cm^{-1}$), His_{10}-PPAT: 10 mg/mL ($\varepsilon_{280\ nm} = 8{,}480\ M^{-1}\ cm^{-1}$); His_{10}-DPCK: 6 mg/mL ($\varepsilon_{280\ nm} = 16{,}960\ M^{-1}\ cm^{-1}$)

Synthesis of substrate analogs. Dichloromethane and tetrahydrofuran used in chemical synthesis were dried using a VAC Solvent Purifier System (Vacuum Atmospheres Company, Amesbury, MA). HPLC purifications were performed using Agilent Eclipse XDB-C18 or XDB-C8 columns (9.4 × 250 mm, 5 μm) connected to an Agilent 1200 binary pump and Agilent G1315D diode-array detector. Fractions were assayed for the desired compounds by liquid chromatography-mass spectrometry (LC-MS) using an Agilent 1290 binary pump coupled to an Agilent 6130 single-quadrupole electrospray ionization mass spectrometer. Fractions containing the desired compounds were pooled, organic solvents were removed on a rotary evaporator, and the pooled fractions were lyophilized. Substrate analogs were dissolved in water, quantified by

absorbance at 260 nm ($\varepsilon_{260\ nm}$ = 13,100 M^{-1} cm^{-1}), and stored at -80°C until further use. High-resolution mass spectra were collected at the QB3/Chemistry Mass Spectrometry Facility at the University of California, Berkeley. NMR spectra were recorded at the College of Chemistry NMR Facility at the University of California, Berkeley. One-dimensional spectra were collected on Bruker AV-600 or AVQ-400 NMR spectrometers at 298 K. Chemical shifts are expressed in parts per million (δ, ppm) downfield from tetramethylsilane (^{1}H and ^{13}C) or trichlorofluoromethane (^{19}F) and are referenced to the solvent signal.



Crosspeak numbering for HMBC spectra							
1	$H_{1''}C_{2''}$	**10**	$H_{10''}C_{11''}$	**19**	$H_{5''}C_{7''}$	**28**	H_2C_4
2	$H_{1''}C_{3''}$	**11**	$H_{10''}C_{2''}$	**20**	$H_{5''}C_{6''}$	**29**	H_2C_6
3	$H_{1''}C_{11''}$	**12**	$H_{10''}C_{1''}$	**21**	$H_{6''}C_{7''}$	**30**	H_8C_4
4	$H_{1''}C_{10''}$	**13**	$H_{10''}C_{3''}$	**22**	$H_{6''}C_{5''}$	**31**	H_8C_5
5	$H_{3''}C_{11''}$	**14**	$H_{11''}C_{10''}$	**23**	$H_{8''}C_{7''}$	**32**	$H_{1'}C_2$
6	$H_{3''}C_{10''}$	**15**	$H_{11''}C_{2''}$	**24**	$H_{8''}C_{9''}$	**33**	$H_{1'}C_4$
7	$H_{3''}C_{2''}$	**16**	$H_{11''}C_{1''}$	**25**	$H_{9''}C_{1RAc}$	**34**	$H_{1'}C_{4'}$
8	$H_{3''}C_{1''}$	**17**	$H_{11''}C_{3''}$	**26**	$H_{9''}C_{8''}$	**35**	$H_{2'}C_{1'}$
9	$H_{3''}C_{4''}$	**18**	$H_{5''}C_{4''}$	**27**	$H_{2RAc}C_{1RAc}$	**36**	$H_{3'}C_{4'}$

O,O′-isopropylidene-D-pantothenic acid (**2**) [*25*]. Calcium pantothenate (2 g, 4.2 mmol) and *p*-toluenesulfonic acid (1.6 g, 4.2 mmol) were suspended in DMF (40 mL). 2,2-Dimethoxypropane (62 mL) was added and the reaction was stirred at room temperature for 12 h. Toluene was added to form an azeotrope with DMF and the solvent was removed under vacuum on a rotary evaporator. The resultant tan residue was suspended in dichlormethane (40 mL) and precipitate was removed by vacuum filtration. The filtrate was evaporated to dryness to yield *O,O′*-isopropylidene-D-pantothenic acid as a tan solid (2.17 g, 98%). ^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ (ppm): 4.05 (s, 1H, $H_{3''}$), 3.62 (d, J = 11.6 Hz, 1H, $H_{1''}$), 3.45 (m, 2H, $H_{5''}$), 3.21 (d, J = 11.6 Hz, 1H, $H_{1''}$), 2.50 (t, J = 6 Hz, 2H, $H_{6''}$), 1.39, 1.36 (each s, 3H, $OC(C\underline{H_3})_2$), 0.96, 0.90 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (100.6 MHz, $CDCl_3$, 25°C) δ (ppm): 174.9, 170.2 ($C_{4''}$ and $C_{7''}$), 99.1 ($O\underline{C}(CH_3)_2$), 77.1 ($C_{3''}$), 71.4 ($C_{1''}$), 34.3 ($C_{5''}$), 33.0 ($C_{4''}$), 29.4 ($C_{2''}$), 22.1 ($OC(\underline{C}H_3)_2$), 18.9, 18.7 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M-H^+$) *m/z* 258.1347, found ($M-H^+$) *m/z* 258.1344.

O,O′-isopropylidene-oxa(dethia)-pantethiene (**3**) [*25*]. *O,O′*-isopropylidene-D-pantothenic acid **1** (836 mg, 3.2 mmol) and triethylamine (0.56 mL, 4 mmol) were dissolved in dry dichloromethane (10 mL). The reaction mixture was cooled to 0ºC and ethyl chloroformate (0.32 mL, 3.36 mmol) was added dropwise by syringe. After stirring for 30 min, ethanolamine (0.28 mL, 4.64 mmol) was added. The reaction mixture was stirred for 30 min at 0ºC and for 2 h at room temperature. The solvent was removed under vacuum and the resultant residue was suspended in ethyl acetate and purified by silica chromatography using a stepwise gradient from ethyl acetate to 60 methanol: 40 ethyl acetate to yield *O,O′*-isopropylidene-oxa(dethia)-pantethiene as a white solid (540 mg, 55%). ^{1}H NMR (600 MHz, MeOD, 25°C) δ (ppm): 4.05 (s,

1H, $H_{3''}$), 3.67 (d, J = 12 Hz, 1H, $H_{1''}$), 3.52 (t, J = 6 Hz, 2H, $H_{9''}$), 3.45 (m, 2H, $H_{5''}$), 3.42 (m, 2H, $H_{8''}$), 3.19 (d, J = 12 Hz, 1H, $H_{1''}$), 2.37 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.38 (s, 3H, OC(C$\underline{H}_3$)$_2$), 1.37 (s, 3H, OC(C$\underline{H}_3$)$_2$), 0.92, 0.90 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, MeOD, 25°C) δ (ppm): 174.2, 174.1 ($C_{4''}$ and $C_{7''}$), 100.5 (O$\underline{C}$(CH$_3$)$_2$), 78.5 ($C_{3''}$), 72.4 ($C_{1''}$), 61.7 ($C_{9''}$), 43.18 ($C_{8''}$), 36.5 ($C_{5''}$), 34.1 ($C_{4''}$), 29.9 ($C_{2''}$), 22.6 (OC($\underline{C}$H$_3$)$_2$), 19.5, 19.2 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd (M+H$^+$) *m/z* 303.1914, found (M+H$^+$) *m/z* 303.1915.

O,O'-isopropylidene-fluoroacetyl-oxa(dethia)-pantethiene (**4b**) [*25*, *26*]. Sodium fluoroacetate (55 mg, 0.55 mmol) and HATU (228 mg, 0.60 mmol) were dissolved in dry THF (5 mL). Diisopropylethylamine (0.26 mL, 1.5 mmol) was added, followed by *O,O'*-isopropylidene-oxa(dethia)-pantethiene **3** (150 mg, 0.5 mmol). The reaction mixture was stirred for 16 h under N_2 at room temperature. Solvent was removed under vacuum and the residue was purified on an Agilent Eclipse XDB-C18 column using a linear gradient from 0 to 100% acetonitrile over 30 min (3 mL/min) with 0.1% formic acid as the aqueous mobile phase. Fractions were screened for the desired compound by LC-MS, pooled, and lyophilized to yield *O,O'*-isopropylidene-fluoroacetyl-oxa(dethia)-pantethiene as a clear oil (38 mg, 21%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 4.99 (d, J = 46.2 Hz, 2H, COC$\underline{H}_2$F), 4.29 (t, J = 6 Hz, 2H, $H_{9''}$), 4.23 (s, 1H, $H_{3''}$), 3.83 (d, J = 12 Hz, 1H, $H_{1''}$), 3.48 (m, 4H, $H_{5''}$ and $H_{8''}$), 3.36 (d, J = 12 Hz, 1H, $H_{1''}$), 2.46 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.463 (s, 3H, OC(C$\underline{H}_3$)$_2$), 1.460 (s, 3H, OC(C$\underline{H}_3$)$_2$), 0.94, 0.93 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 174.1, 171.7 ($C_{4''}$ and $C_{7''}$), 170.4, 170.3 ($\underline{C}$OCH$_2$F), 99.8 (O$\underline{C}$(CH$_3$)$_2$), 78.5, 77.4 (CO$\underline{C}$H$_2$F), 77.0 ($C_{3''}$), 70.6 ($C_{1''}$), 64.2 ($C_{9''}$), 38.0 ($C_{8''}$), 35.2 ($C_{5''}$), 32.3 ($C_{4''}$), 28.3 ($C_{2''}$), 20.9 (OC($\underline{C}$H$_3$)$_2$), 18.2, 18.1 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd (M+H$^+$) *m/z* 363.1926, found (M+H$^+$) *m/z* 363.1926.

O,O'-isopropylidene-acetyl-oxa(dethia)-pantethiene (**4a**) [*25*]. Sodium acetate (45 mg, 0.55 mmol) and HATU (228 mg, 0.60 mmol) were dissolved in dry THF (5 mL). Diisopropylethylamine (0.26 mL, 1.5 mmol) was added, followed by *O,O'*-isopropylidene-oxa(dethia)-pantethiene **3** (150 mg, 0.5 mmol). The reaction was stirred for 16 h under N_2 at room temperature. Solvent was removed under vacuum and the residue was purified on an Agilent Eclipse XDB-C18 column using a linear gradient from 0 to 100% acetonitrile over 30 min (3 mL/min) with 0.1% formic acid as the aqueous mobile phase. Fractions were screened for the desired compound by LC-MS, pooled, and lyophilized to yield *O,O'*-isopropylidene-acetyl-oxa(dethia)-pantethiene as a clear oil (49 mg, 28%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 4.09 (s, 1H, $H_{3''}$), 4.01 (t, J = 5.4 Hz, 2H, $H_{9''}$), 3.70 (d, J = 12 Hz, 1H, $H_{1''}$), 3.31 (m, 4H, $H_{5''}$ and $H_{8''}$), 3.20 (d, J = 12 Hz, 1H, $H_{1''}$), 2.33 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.94 (s, 3H, COC$\underline{H}_3$), 1.329 (s, 3H, OC(C$\underline{H}_3$)$_2$), 1.326 (s, 3H, OC(C$\underline{H}_3$)$_2$), 0.79 0.76 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 175.1, 174.1 ($C_{4''}$ and $C_{7''}$), 171.7 ($\underline{C}$OCH$_3$), 99.8 (O$\underline{C}$(CH$_3$)$_2$), 77.0 ($C_{3''}$), 70.6 ($C_{1''}$), 38.5 ($C_{9''}$), 38.2 ($C_{8''}$), 35.4 ($C_{5''}$), 32.3 ($C_{4''}$), 30.2 (CO$\underline{C}$H$_3$), 28.3 ($C_{2''}$), 20.9, 20.4 (OC($\underline{C}$H$_3$)$_2$), 19.0, 18.2 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd (M+H$^+$) *m/z* 345.2020, found (M+H$^+$) *m/z* 345.2026.

O,O'-isopropylidene-N-Cbz-aza(dethia)-pantethiene (**6**). *O,O'*-isopropylidene-D-pantothenic acid **2** (836 mg, 3.24 mmol) and triethylamine (1.12 mL, 8 mmol) were dissolved in dry dichloromethane (10 mL). The reaction mixture was cooled to 0ºC and ethyl chloroformate (0.32 mL, 3.36 mmol) was added dropwise by syringe. After stirring the reaction mixture on ice for 30 min, benzyl-*N*-(2-aminoethyl)carbamate hydrochloride (822 mg, 4 mmol) was added. The reaction mixture was stirred at 0ºC for 30 min and at room temperature for 2 h. The solvent was removed on a rotary evaporator and the residue was purified by silica chromatography using 75

ethyl acetate: 25 methanol as the mobile phase. The resultant oil was further purified by silica chromatography using a stepwise gradient from ethyl acetate to 10 methanol: 90 ethyl acetate to yield the desired compound as a white solid (600 mg, 43%). ^{1}H NMR (600 MHz, MeOD, 25°C) δ (ppm): 7.32 (m, 5H, $C_6\underline{H_5}$), 5.12 (s, 2H, OCH_2Ph), 4.11 (s, 1H, $H_{3''}$), 3.74 (d, J = 12.6 Hz, 1H, $H_{1''}$), 3.41 (m, 2H, $H_{5''}$), 3.26 (m, 2H, $H_{8''}$), 3.23 (t, J = 6 Hz, 2H, $H_{9''}$), 2.38 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.54 (s, 3H, $OC(C\underline{H_3})_2$), 1.52 (s, 3H, $OC(C\underline{H_3})_2$), 0.98, 0.96 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, MeOD, 25°C) δ (ppm): 174.3, 172.4 ($C_{4''}$ and $C_{7''}$), 138.5, 129.6, 129.1, 129.03 ($\underline{C_6}H_5$), 100.5 ($O\underline{C}(CH_3)_2$), 78.5 ($C_{3''}$), 72.4 ($C_{1''}$), 67.6 ($O\underline{C}H_2Ph$), 40.5 ($C_{9''}$), 36.5 ($C_{8''}$), 34.1 ($C_{5''}$), 30.8 ($C_{4''}$), 29.9 ($C_{2''}$), 22.6 ($OC(\underline{C}H_3)_2$), 19.5, 19.2 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 436.2442, found ($M+H^+$) *m/z* 436.2441.

O,O'-isopropylidene-aza(dethia)-pantethiene (**7**). *O,O'*-isopropylidene-*N*-Cbz-aza(dethia)-pantethiene **6** (446 mg, 1 mmol) was dissolved in methanol (10 mL) in the presence of palladium (10 wt% on activated carbon, 50 mg). The flask was purged with N_2 and the reaction mixture was saturated with H_2 and stirred under H_2 atmosphere for 10 min. The reaction mixture was then filtered through a bed of Celite, which was washed extensively with methanol. Removal of the solvent *in vacuo* yielded *O,O'*-isopropylidene-aza(dethia)-pantethiene as a clear oil (260 mg, 86%). ^{1}H NMR (600 MHz, MeOD, 25°C) δ (ppm): 4.13 (s, 1H, $H_{3''}$), 3.74 (d, J = 12 Hz, 1H, $H_{1''}$), 3.46 (m, 2H, $H_{5''}$), 3.44 (m, 2H, $H_{8''}$), 3.20 (d, J = 12 Hz, 1H, $H_{1''}$), 2.73 (t, J = 6.6 Hz, 2H, $H_{9''}$), 2.45 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.45 (s, 3H, $OC(C\underline{H_3})_2$), 1.44 (s, 3H, $OC(C\underline{H_3})_2$), 0.99, 0.97 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, MeOD, 25°C) δ (ppm): 174.3, 172.3 ($C_{4''}$ and $C_{7''}$), 100.5 ($O\underline{C}(CH_3)_2$), 78.5 ($C_{3''}$), 72.4 ($C_{1''}$), 42.9 ($C_{9''}$), 42.1 ($C_{8''}$), 36.4 ($C_{5''}$), 34.1 ($C_{4''}$), 29.8 ($C_{2''}$), 22.6 ($OC(\underline{C}H_3)_2$), 19.5, 19.2 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 302.2074, found ($M+H^+$) *m/z* 302.2074.

O,O'-isopropylidene-fluoroacetyl-aza(dethia)-pantethiene (**8b**). Sodium fluoroacetate (36 mg, 0.36 mmol) and HATU (152 mg, 0.40 mmol) were dissolved in dry DMF (4 mL). Diisopropylethylamine (0.2 mL, 1.16 mmol) was added, followed by *O,O'*-isopropylidene-aza(dethia)-pantethiene **7** (100 mg, 0.33 mmol). The reaction mixture was stirred for 12 h at room temperature under N_2 atmosphere. Toluene was added to form an azeotrope with DMF and solvents were removed on a rotary evaporator. The residue was dissolved in water and purified on an Agilent Eclipse XDB-C8 column using a linear gradient from 0 to 100% methanol over 30 min (3 mL/min) with 0.1% formic acid as the aqueous mobile phase. Fractions were screened for the desired compound by LC-MS, pooled, and lyophilized to yield *O,O'*-isopropylidene-fluoroacetyl-aza(dethia)-pantethiene as a clear oil (16.4 mg, 14%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 4.99 (d, J = 46.2 Hz, 2H, $COC\underline{H_2}F$), 4.29 (t, J = 6 Hz, 2H, $H_{9''}$), 4.23 (s, 1H, $H_{3''}$), 3.83 (d, J = 12 Hz, 1H, $H_{1''}$), 3.48 (m, 4H, $H_{5''}$ and $H_{8''}$), 3.36 (d, J = 12 Hz, 1H, $H_{1''}$), 2.46 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.463 (s, 3H, $OC(C\underline{H_3})_2$), 1.460 (s, 3H, $OC(C\underline{H_3})_2$), 0.94, 0.93 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 174.1, 171.7 ($C_{4''}$ and $C_{7''}$), 170.4, 170.3 ($\underline{C}OCH_2F$), 99.8 ($O\underline{C}(CH_3)_2$), 78.5, 77.4 ($CO\underline{C}H_2F$), 77.0 ($C_{3''}$), 70.6 ($C_{1''}$), 64.2 ($C_{9''}$), 38.0 ($C_{8''}$), 35.2 ($C_{5''}$), 32.3 ($C_{4''}$), 28.3 ($C_{2''}$), 20.9 ($OC(\underline{C}H_3)_2$), 18.2, 18.1 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 363.1926, found ($M+H^+$) *m/z* 363.1926.

O,O'-isopropylidene-acetyl-aza(dethia)-pantethiene (**8a**). Sodium acetate (47 mg, 0.57 mmol) and HATU (238 mg, 0.63 mmol) were dissolved in dry DMF (6 mL). Diisopropylethylamine (0.3 mL, 1.82 mmol) was added, followed by *O,O'*-isopropylidene-aza(dethia)-pantethiene **7** (155 mg, 0.51 mmol). The reaction mixture was stirred for 16 h under N_2 at room temperature. Solvent was removed under vacuum and the residue was purified

directly on an Agilent Eclipse XDB-C18 column using a linear gradient from 0 to 100% acetonitrile over 30 min (3 mL/min) with 0.1% formic acid as the aqueous mobile phase. Fractions were screened for the desired compound by LC-MS, pooled, and lyophilized to yield *O,O′*-isopropylidene-acetyl-aza(dethia)-pantethiene as a clear oil (45 mg, 24%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 4.23 (s, 1H, $H_{3''}$), 3.83 (d, *J* = 12 Hz, 1H, $H_{1''}$), 3.47 (m, 2H, $H_{5''}$), 3.44 (m, 2H, $H_{8''}$), 3.36 (d, *J* = 11 Hz, 1H, $H_{1''}$), 3.27 (t, *J* = 6 Hz, 2H, $H_{9''}$), 2.45 (t, *J* = 6.6 Hz, 2H, $H_{6''}$), 1.96 (s, 3H, $COC\underline{H}_3$), 1.462 (s, 3H, $OC(C\underline{H}_3)_2$), 1.459 (s, 3H, $OC(C\underline{H}_3)_2$), 0.94, 0.93 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 174.3, 174.1 ($C_{4''}$ and $C_{7''}$), 171.7 ($\underline{C}OCH_3$), 99.8 ($O\underline{C}(CH_3)_2$), 77.0 ($C_{3''}$), 70.6 ($C_{1''}$), 38.7 ($C_{9''}$), 38.6 ($C_{8''}$), 35.3 ($C_{5''}$), 32.3 ($C_{4''}$), 28.3 ($C_{2''}$), 21.9, 20.9 ($OC(\underline{C}H_3)_2$), 18.2, 18.1 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 344.2180, found ($M+H^+$) *m/z* 344.2179.

O,O′-diacetyl-D-pantothenic acid (**10**) [*25*]. Calcium pantothenate (2 g) and I_2 (140 mg) were dissolved in acetic anhydride (40 mL). The reaction mixture was stirred at 0ºC for 2 h and at room temperature for 16 h. Toluene was added to the reaction mixture to form an azeotrope with acetic anhydride and the solvents were removed on a rotary evaporator. The resultant red oil was dissolved in dichloromethane (60 mL) and washed with 1 M sodium thiosulfate (60 mL). The organic layer was dried over $MgSO_4$ and the solvent was removed under vacuum to yield *O,O′*-diacetyl-D-pantothenic acid as a clear oil in quantitative yield. ^{1}H NMR (600 MHz, $CDCl_3$, 25°C) δ (ppm): 4.88 (s, 1H, $H_{3''}$), 4.01 (d, *J* = 11 Hz, 1H, $H_{1''}$), 3.80 (d, *J* = 11 Hz, 1H, $H_{1''}$), 3.55 (m, 2H, $H_{5''}$), 2.55 (t, *J* = 6.6 Hz, 2H, $H_{6''}$), 2.12 (s, 3H, $COC\underline{H}_3$), 2.01 (s, 3H, $CO\underline{CH}_3$), 1.04, 0.98 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, $CDCl_3$, 25°C) δ (ppm): 175.7 ($\underline{C}OOH$), 171.4, 170.2 ($C_{4''}$ and $C_{7''}$), 168.5 ($\underline{C}OCH_3$), 77.0 ($C_{3''}$), 69.5 ($C_{1''}$), 37.3 ($C_{5''}$), 34.8 ($C_{4''}$), 29.7 ($C_{2''}$), 21.5, 21.0 ($CO\underline{C}H_3$), 20.9, 20.7 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 304.1396, found ($M+H^+$) *m/z* 304.1391.

N-(O,O′-diacetyl-D-pantothenoyl)-γ-aminobutyric acid (**11**) [*25*]. *O,O′*-diacetyl-D-pantothenic acid **10** (480 mg, 1.6 mmol), *t*-butyl-γ-aminobutanoate hydrochloride (353 mg, 1.8 mmol), 1-hydroxybenzotriazole hydrate (243 mg, 1.8 mmol), and triethylamine (0.45 mL, 3.3 mmol) were dissolved in dry dichloromethane (10 mL) under N_2. The reaction mixture was cooled to 0ºC and *N*-(3-dimethylaminopropyl)-*N′*-ethylcarbodiimide hydrochloride (353 mg, 1.8 mmol) was added, and the mixture was stirred at 0ºC for 30 min and at room temperature for 16 h. The reaction mixture was then diluted with dichloromethane, washed with 0.1 M HCl (20 mL) and saturated sodium bicarbonate (20 mL), and the organic layer was dried over $MgSO_4$. Evaporation of the solvent yielded *N*-(*O,O′*-diacetyl-D-pantothenoyl)-γ-aminobutyric acid *t*-butyl ester (**12**) as a yellow oil (650 mg). This material was dissolved in dry dichloromethane and TFA and the reaction mixture was stirred at room temperature for 3 h. The solvent was removed by forming an azeotrope with toluene and evaporating under vacuum followed by lyophilization. *N*-(*O,O′*-diacetyl-D-pantothenoyl)-γ-aminobutyric acid was obtained as a yellow oil (620 mg, 85%). ^{1}H NMR (600 MHz, $CDCl_3$, 25°C) δ (ppm): 4.91 (s, 1H, $H_{3''}$), 4.01 (d, *J* = 11 Hz, 1H, $H_{1''}$), 3.80 (d, *J* = 11 Hz, 1H, $H_{1''}$), 3.55 (m, 2H, $H_{5''}$), 3.47 (m, 2H, $H_{8''}$), 2.55 (t, *J* = 6.6 Hz, 2H, $H_{6''}$), 2.41 (t, *J* = 6.6 Hz, 2H, $C\underline{H}_2COOH$), 2.28 (quintet, *J* = 8 Hz, 2H, $H_{9''}$), 2.20 (s, 3H, $COC\underline{H}_3$), 2.10 (s, 3H, $CO\underline{CH}_3$), 1.04, 1.03 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, $CDCl_3$, 25°C) δ (ppm): 178.5 ($\underline{C}OOH$), 173.4, 171.7 ($C_{4''}$ and $C_{7''}$), 159.9, 159.7 ($\underline{C}OCH_3$), 76.5 ($C_{3''}$), 75.3 ($C_{1''}$), 39.6 ($C_{8''}$), 37.4 ($\underline{C}H_2COOH$), 36.1 ($C_{5''}$), 35.3 ($C_{4''}$), 31.5 ($C_{2''}$), 28.1, 22.3 ($CO\underline{C}H_3$), 20.9, 20.8 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 389.1932, found ($M+H^+$) *m/z* 389.1927.

O,O'-diacetyl-malonyl-carba(dethia)-pantethiene methyl ester (**13**) [*25*]. *N*-(*O,O'*-diacetyl-D-pantothenoyl)-γ-aminobutyric acid **11** (620 mg, 1.6 mmol), 2,2-dimethyl-1,3-dioxane-4,6-dione (225 mg, 1.6 mmol), and 4-dimethylaminopyridine (225 mg, 1.8 mmol) were dissolved in dry dichloromethane (10 mL). The reaction mixture was stirred under N_2 atmosphere at room temperature for 10 min, and *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (335 mg, 1.8 mmol) was added. The reaction mixture was stirred at room temperature for 4 h. The solution was diluted with dichloromethane (10 mL) and washed with 0.1 M HCl (20 mL). The organic layer was dried over $MgSO_4$ and the solvent removed under vacuum to yield the desired compound as a yellow oil (590 mg, 70%). This material (590 mg, 1.1 mmol) was dissolved in dry methanol (20 mL) and heated to reflux for 18 h under N_2 atmosphere. The reaction mixture was cooled and methanol was removed on a rotary evaporator to yield *O,O'*-diacetyl-malonyl-carba(dethia)-pantethiene methyl ester as a yellow oil (514 mg, 99%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 4.80 (s, 1H, $H_{3''}$), 4.07 (d, J = 11 Hz, 1H, $H_{1''}$), 3.85 (d, J = 11 Hz, 1H, $H_{1''}$), 3.75 (s, 3H, Me), 3.70 (s, 1H, $COC\underline{H}_2CO$), 3.50 (m, 2H, $H_{5''}$), 3.15 (m, 2H, $H_{8''}$), 2.68 (t, J = 7.2 Hz, 2H, $H_{6''}$), 2.42 (m, 2H, $C\underline{H}_2$), 2.20 (s, 3H, $COC\underline{H}_3$), 2.09 (s, 3H, $CO\underline{CH}_3$), 1.78 (quintet, J = 7.2 Hz, 2H, $H_{9''}$), 1.03, 1.01 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 207.5 ($CH_2\underline{C}OCH_2$), 174.3, 173.5 ($C_{4''}$ and $C_{7''}$), 173.1, 170.6 ($\underline{C}OCH_3$), 170.1 ($\underline{C}$OOMe), 77.2 ($C_{3''}$), 69.5 ($C_{1''}$), 52.8 (Me), 40.0 ($COC\underline{H}_2CO$), 38.5 ($C_{8''}$), 36.6 ($C_{5''}$), 36.6 ($C_{4''}$), 35.8 ($C_{2''}$), 22.3, 20.7 ($CO\underline{C}H_3$), 20.2, 20.0 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 445.2544, found ($M+H^+$) *m/z* 445.2551.

O,O'-diacetyl-fluoromalonyl-carba(dethia)-pantethiene methyl ester [*25*, *26*]. *O,O'*-diacetyl-malonyl-carba(dethia)-pantethiene methyl ester **13** (50 mg, 0.1 mmol) and Selectfluor (40 mg, 0.1 mmol) were suspended in acetonitrile (500 μL). The mixture was heated to 82˚C under 100 W microwave irradiation for 10 min in a Discover Labmate microwave reactor (CEM). The reaction mixture was then diluted to 5 mL with water and purified on an Agilent Eclipse-XDB C8 column using a linear gradient from 0 to 100% methanol over 30 min (3 mL/min) with 0.1% formic acid as the aqueous mobile phase. Fractions containing the desired product as determined by LC-MS analysis were then pooled and lyophilized (12 mg, 25%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 5.68 (d, J = 47 Hz, 1H, $C\underline{H}F$) 4.80 (s, 1H, $H_{3''}$), 4.07 (d, J = 11 Hz, 1H, $H_{1''}$), 3.86 (s, 3H, Me), 3.85 (d, J = 11 Hz, 1H, $H_{1''}$), 3.50-3.40 (m, 2H, $H_{5''}$), 3.20 (m, 2H, $H_{8''}$), 2.85 (m, 2H, $H_{6''}$), 2.44 (m, 2H, $C\underline{H}_2$), 2.16 (s, 3H, $COC\underline{H}_3$), 2.10 (s, 3H, $CO\underline{CH}_3$), 1.77 (quintet, J = 7.2 Hz, 2H, $H_{9''}$), 1.03, 1.02 (each s, 3H, $H_{10''}$ and $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 211.0 ($CH_2\underline{C}OCHF$), 174.4, 173.6 ($C_{4''}$ and $C_{7''}$), 173.1 ($\underline{C}OCH_3$), 170.6 ($\underline{C}$OOMe), 92.0 ($\underline{C}$HF), 77.4 ($C_{3''}$), 69.5 ($C_{1''}$), 53.9 (Me), 38.4 ($C_{8''}$), 36.6 ($C_{5''}$), 36.6 ($C_{4''}$), 35.8 ($C_{2''}$), 21.7, 20.6 ($CO\underline{C}H_3$), 20.1, 19.9 ($C_{10''}$ and $C_{11''}$). HR-ESI-MS calcd ($M+H^+$) *m/z* 463.2086, found ($M+H^+$) *m/z* 463.2089.

General protocol for enzymatic synthesis of coenzyme A analogs from pantethiene analogs. Pantethiene analogs were deprotected as described below and dissolved in CoA synthesis buffer (50 mM Tris-HCl, pH 7.5, 20 mM KCl, 10 mM $MgCl_2$, 15 mM ATP) to a final concentration of 5 mM [*24*]. The reaction was initiated by addition of PanK, PPAT, and DPCK (2 mg each). For acid-labile oxa(dethia)-CoA analogs, the reaction mixture was lyophilized and the lyophilizate was dissolved in 5 mL water and 0.2 μm-filtered. For aza(dethia)-CoA analogs, the reaction mixture was lyophilized, the lyophilizate was dissolved in 5 mL water with 70% perchloric acid (250 μL, 0.05 volumes) was added to precipitate protein. Precipitated protein was removed by centrifugation and the solution was neutralized and filtered (0.2 μm). The filtered

solutions of crude CoA analogs were injected onto an Agilent Eclipse XDB-C18 equipped with an Agilent Eclipse XDB-C18 analytical guard column (4.6 × 12.5 mm, 5 μm). The column was washed with 0.2% TFA in water until the absorbance at 260 nm returned to baseline. A linear gradient was then applied from 0 to 100% methanol over 90 min (3 mL/min) with 0.2% TFA as the aqueous mobile phase. Fractions were screened by LC-MS, pooled, and lyophilized to yield pure CoA analogs as a white powder.

Oxa(dethia) and aza(dethia)-CoA analogs. The appropriate *O*, *O'*-isopropylidene-acyl-oxa(dethia)-pantethiene or *O*, *O'*-isopropylidene-acyl-aza(dethia)-pantethiene derivative (50 mg, 0.1 mmol) was dissolved in water (2 mL) and Dowex 50WX8 hydrogen form (50 mg) was added. The mixture was stirred at room temperature for 1 h, and then Dowex was removed by centrifugation and filtration of the reaction mixture. The filtrate was carefeully neutralized with 1 M NaOH and used without further purification for enzymatic synthesis of the acyl-oxa(dethia)-CoA or acyl-aza(dethia)-CoA.

Fluoroacetyl-oxa(dethia)-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.55 (s, 1H, H_8), 8.35 (s, 1H, H_2), 6.12 (d, *J* = 6 Hz, 1H, $H_{1'}$), 4.90 (d, *J* = 46 Hz, 2H, C$\underline{H}_2$F), 4.24 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.21 (t, *J* = 6 Hz, 2H, $H_{5''}$), 4.54 (s, 1H, $H_{4'}$), 4.21 (d, *J* = 13.8 Hz, 2H, $H_{5'}$), 3.93 (s, 1H, $H_{3'}$), 3.83 (m, 1H, $H_{1''}$), 3.60 (m, 1H, $H_{1''}$), 3.39 (m, 4 H, $H_{8''}$ and $H_{9''}$), 2.39 (t, *J* = 6 Hz , 2H, $H_{6''}$), 0.87 (s, 3H, $H_{10''}$), 0.76 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 174.63 ($C_{7''}$), 173.14 ($C_{4''}$), 170.4, 170.2 (FCH$_2$$\underline{C}$O), 149.75 ($C_6$), 148.39 ($C_2$), 144.67 ($C_4$), 142.41 ($C_8$), 118.45 ($C_5$), 87.49 ($C_{1'}$), 83.06 ($C_{4'}$), 78.53, 77.37 (F$\underline{C}H_2$CO), 74.43 ($C_{2'}$), 74.14 ($C_{3''}$), 73.62 ($C_{3'}$), 72.24 ($C_{1'}$), 64.10 ($C_{5'}$), 64.10 ($C_{9''}$), 38.29 ($C_{8''}$), 38.02 ($C_{2''}$) 35.33 ($C_{5''}$), 35.18 ($C_{6''}$), 20.61 ($C_{10''}$), 18.27 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -231.3 (t, *J* = 49.2 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 810.1319, found (M-H$^+$) *m/z* 810.1314.

Acetyl-oxa(dethia)-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.47 (s, 1H, H_8), 8.32 (s, 1H, H_2), 5.97 (d, *J* = 6 Hz, 1H, $H_{1'}$), 4.75 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.40 (s, 1H, $H_{4'}$), 4.10 (d, *J* = 14.4 Hz, 2H, $H_{5'}$), 3.90 (t, *J* = 6 Hz, 2H, $H_{5''}$), 3.79 (s, 1H, $H_{3''}$), 3.70 (m, 1H, $H_{1''}$), 3.41 (m, 1H, $H_{1''}$), 3.26 (m, 4H, $H_{8''}$ and $H_{9''}$), 2.26 (t, *J* = 6 Hz , 2H, $H_{6''}$), 1.85 (s, 3H, C$\underline{H}_3$CO), 0.75 (s, 3H, $H_{10''}$), 0.63 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 174.50 ($C_{7''}$), 174.08 ($C_{4''}$), 173.93 (CH$_2$$\underline{C}$O), 149.60 ($C_6$), 148.24 ($C_2$), 144.64 ($C_4$), 142.29 ($C_8$), 118.25 ($C_5$), 87.44 ($C_{1'}$), 82.83 ($C_{4'}$), 74.49 ($C_{2'}$), 74.06 ($C_{3''}$), 73.53 ($C_{3'}$), 72.28 ($C_{1''}$), 65.13 ($C_{5'}$), 63.22 ($C_{9''}$), 38.25 ($C_{8''}$), 38.20 ($C_{2''}$) 35.28 ($C_{5''}$), 35.08 ($C_{6''}$), 20.51 ($C_{10''}$), 20.18 ($\underline{C}H_3$CO), 18.22 ($C_{11''}$). HR-ESI-MS calcd (M-H$^+$) *m/z* 792.1413, found (M-H$^+$) *m/z* 792.1408.

Fluoroacetyl-aza(dethia)-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.50 (s, 1H, H_8), 8.32 (s, 1H, H_2), 6.08 (d, *J* = 6 Hz, 1H, $H_{1'}$), 4.83 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.74 (d, *J* = 51.6 Hz, 2H, FC$\underline{H}_2$CO), 4.70 (s, 1H, $H_{4'}$), 4.18 (d, *J* = 12 Hz, 2H, $H_{5'}$), 3.90 (s, 1H, $H_{3''}$), 3.79 (m, 1H, $H_{1''}$), 3.57 (m, 1H, $H_{1''}$), 3.34 (t, *J* = 6 Hz, 2H, $H_{5''}$), 3.26 (t, *J* = 6 Hz, 2H, $H_{8''}$), 3.20 (t, *J* = 6 Hz, 2H, $H_{9''}$), 2.35 (t, *J* = 6 Hz , 2H, $H_{6''}$), 0.85 (s, 3H, $H_{10''}$), 0.73 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 174.60 ($C_{7''}$), 174.16 ($C_{4''}$), 170.85, 170.74 (FCH$_2$$\underline{C}$O), 149.70 ($C_6$), 148.32 ($C_2$), 144.66 ($C_4$), 142.36 ($C_8$), 118.37 ($C_5$), 87.48 ($C_{1'}$), 84.50, 83.35 (F$\underline{C}H_2$CO), 82.98 ($C_{4'}$), 74.50 ($C_{2'}$), 74.16 ($C_{3''}$), 73.59 ($C_{3'}$), 72.30 ($C_{1''}$), 65.14 ($C_{5'}$), 38.42 ($C_{9''}$), 38.24 ($C_{8''}$), 38.20 ($C_{2''}$) 35.35 ($C_{5''}$), 35.23 ($C_{6''}$), 20.56 ($C_{10''}$), 18.28 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -227.9 (t, *J* = 48 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 809.1479, found (M-H$^+$) *m/z* 809.1470.

Acetyl-aza(dethia)-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.47 (s, 1H, H_8), 8.32 (s, 1H, H_2), 5.97 (d, J = 6 Hz, 1H, $H_{1'}$), 4.75 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.40 (s, 1H, $H_{4'}$), 4.10 (d, J = 14.4 Hz, 2H, $H_{5'}$), 3.90 (t, J = 6 Hz, 2H, $H_{5''}$), 3.79 (s, 1H, $H_{3''}$), 3.70 (m, 1H, $H_{1''}$), 3.41 (m, 1H, $H_{1''}$), 3.26 (m, 4H, $H_{8''}$ and $H_{9''}$), 2.26 (t, J = 6 Hz , 2H, $H_{6''}$), 1.85 (s, 3H, $C\underline{H_3}CO$), 0.75 (s, 3H, $H_{10''}$), 0.63 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 174.50 ($C_{7''}$), 174.08 ($C_{4''}$), 173.93 ($CH_2\underline{C}O$), 149.60 (C_6), 148.24 (C_2), 144.64 (C_4), 142.29 (C_8), 118.25 (C_5), 87.44 ($C_{1'}$), 82.83 ($C_{4'}$), 74.49 ($C_{2'}$), 74.06 ($C_{3''}$), 73.53 ($C_{3'}$), 72.28 ($C_{1''}$), 65.13 ($C_{5'}$), 63.22 ($C_{9''}$), 38.25 ($C_{8''}$), 38.20 ($C_{2''}$) 35.28 ($C_{5''}$), 35.08 ($C_{6''}$), 20.51 ($C_{10''}$), 20.18 ($\underline{C}H_3CO$), 18.22 ($C_{11''}$). HR-ESI-MS calcd (M-H$^+$) *m/z* 791.1573, found (M-H$^+$) *m/z* 791.1552.

Acyl-carba(dethia)-CoA analogs. The appropriate *O,O′*-diacetyl-malonyl-carba(dethia)-pantethiene methyl ester (50 mg) and potassium carbonate (50 mg) were dissolved in water (1 mL). The mixture was heated to 82˚C under 100 W microwave irradiation for 10 min in a CEM Discover Labmate microwave reactor to yield fluoroacetyl-carba(dethia)-pantethiene. After cooling, the mixture was neutralized with 1 M HCl and used without further purification.

Fluoroacetyl-carba(dethia)-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.53 (s, 1H, H_8), 8.33 (s, 1H, H_2), 6.10 (d, J = 6 Hz, 1H, $H_{1'}$), 4.95 (d, J = 47 Hz, 2H, $C\underline{H_2}F$), 4.85 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.52 (s, 1H, $H_{4'}$), 4.20 (d, J = 13.8 Hz, 2H, $H_{5'}$), 3.91 (s, 1H, $H_{3''}$), 3.80 (m, 1H, $H_{1''}$), 3.59 (m, 1H, $H_{1''}$), 3.37 (t, J = 7.2 Hz, 2H, $H_{5''}$), 3.05 (t, J = 7.2 Hz, 2H, $H_{8''}$), 2.44 (t, J = 7.2 Hz, 2H, $H_{6'}$), 2.37 (t, J = 7.8 Hz, 2H, $C\underline{H_2}CO$), 1.65 (quintet, J = 6.6 Hz, 2H, $H_{9''}$), 0.85 (s, 3H, $H_{10''}$), 0.74 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 210.0, 209.9 ($CH_2F\underline{C}O$), 174.6 ($C_{7''}$), 173.8 ($C_{4''}$), 149.7 (C_6), 148.4 (C_2), 144.7 (C_4), 142.4 (C_8), 118.5 (C_5), 87.5 ($C_{1'}$), 84.3 ($\underline{C}H_2F$), 83.0 ($C_{4'}$), 74.4 ($C_{2'}$), 74.1 ($C_{3''}$), 73.6 ($C_{3'}$), 72.2 ($C_{1''}$), 65.1 ($C_{5'}$), 38.5 ($C_{8''}$), 38.4 ($C_{2''}$), 38.3 ($\underline{C}H_2COCH_2F$), 35.4 ($C_{5''}$), 35.2 ($C_{6'''}$), 22.3 ($C_{9''}$), 21.8 ($C_{10''}$), 18.3 ($C_{11''}$). ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -229.5 (t, J = 46.3 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 808.1527, found (M-H$^+$) *m/z* 808.1505.

Acetyl-carba(dethia)-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.39 (s, 1H, H_8), 8.22 (s, 1H, H_2), 5.97 (d, J = 6 Hz, 1H, $H_{1'}$), 4.75 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.40 (s, 1H, $H_{4'}$), 4.10 (d, J = 12.6 Hz, 2H, $H_{5'}$), 3.80 (s, 1H, $H_{3''}$), 3.72 (m, 1H, $H_{1''}$), 3.49 (m, 1H, $H_{1''}$), 2.92 (t, J = 6 Hz, 2H, $H_{5''}$), 2.35 (t, J = 6 Hz, 2H, $H_{8''}$), 2.25 (t, J = 6 Hz, 2H, $H_{9''}$), 1.96 (s, 3H, $C\underline{H_3}CO$), 1.49 (t, J = 6 Hz , 2H, $H_{6''}$), 0.74 (s, 3H, $H_{10''}$), 0.64 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 215.4 ($CH_2\underline{C}O$), 174.50 ($C_{7''}$), 173.6 ($C_{4''}$), 149.60 (C_6), 148.24 (C_2), 144.64 (C_4), 142.28 (C_8), 118.25 (C_5), 87.44 ($C_{1'}$), 82.85 ($C_{4'}$), 74.46 ($C_{2'}$), 74.07 ($C_{3''}$), 73.50 ($C_{3'}$), 72.26 ($C_{1''}$), 65.14 ($C_{5'}$), 40.17 ($CH_2\underline{C}H_2COCH_3$), 38.48 ($C_{8''}$), 38.26 ($C_{2''}$) 35.33 ($C_{5''}$), 35.13 ($C_{6'''}$), 29.14 ($\underline{C}H_3CO$), 22.52 ($C_{9''}$), 20.52 ($C_{10''}$), 18.26 ($C_{11''}$). HR-ESI-MS calcd (M-H$^+$) *m/z* 790.1621, found (M-H$^+$) *m/z* 790.1598.

Competitive inhibition of FlK with substrate analogs. Steady-state kinetic experiments were performed using 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) to detect release of free coenzyme A as described previously. Enzymatic reactions were initiated by addition of FlK or FlK-F36A (5 nM for fluoroacetyl-CoA or 10 μM for acetyl-CoA) to reaction mixtures containing either fluoroacetyl-CoA (5 to 100 μM) or acetyl-CoA (100 to 1000 μM) as the substrate. When included, competitive inhibitors were added at a concentration at or above the K_I. Each rate was measured in triplicate. Kinetic parameters (k_{cat} and K_M) were determined by fitting the data to the equation $V_0 = V_{max}[S]/K_M + [S]$, where V_0 is the initial rate and [S] is the substrate concentration,

using Origin 6.0 (OriginLab Corporation, Northampton, MA). The K_I was calculated according to the equation $K_I = [I]/\alpha-1$, where [I] is the inhibitor concentration and α is $K_{M,app}/K_M$.

Isothermal titration calorimetry. ITC experiments were performed on a Microcal Auto-iTC200 (GE Healthcare). FlK and FlK-F36A were dialyzed against 100 mM Tris-HCl, pH 7.6 at 4°C for at least 16 h. Acyl-CoA analog ligands were dissolved in water, carefully neutralized with 1 M NaOH, and lyophilized. The lyophilized powder was dissolved in the same dialysis buffer used for protein buffer exchange immediately prior to the experiment. Titrations were performed under low c-value conditions [*27*] in which a large (10-100-fold) excess of ligand was titrated into the protein solution. Acyl-CoA analogs (7.5-25 mM) were titrated into the cell containing FlK or FlK-F36A (15-100 μM) using a pre-injection (0.5 μL) followed by thirteen additional injections (3.2 μL). The pre-injection point was discarded and data were baseline corrected, integrated, and fit to the one set of sites model using Origin 7.0 (Microcal). Due to the nature of the low c-value experimental design [*27*], stoichiometry of the interaction could not be fit and was held constant at 1:1 ligand:protein, a reasonable assumption based on the available structural information. Errors in the thermodynamic parameters are derived from nonlinear curve fitting.

Preparation of proteins for NMR spectroscopy. LB (10 mL) containing carbenicillin (50 μg/mL) was inoculated with a single colony of freshly transformed *E. coli* BL21(de3)-T1^R containing the appropriate overexpression plasmid for FlK or FlK-F36A. The cultures were grown at 37ºC at 200 rpm to for 4-6 h. For preparation of uniformly ^{15}N-labeled FlK, the culture was diluted into M9 media prepared with ^{15}N-ammonium chloride (per L: 12.8 g NaHPO4 · $7H_2O$, 3.0 g KH_2PO4, 0.5 g NaCl, 1.0 g $^{15}NH_4Cl$, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.4% glucose, 40 μg/mL thiamine, 5 mg $FeCl_3$, 0.5 mg $ZnCl_2$, 0.1 mg $CuCl_2$, 0.1 mg $CoCl_2 \cdot 6H_2O$, 0.1 mg HBO_3, 16 mg $MnCl_2$, and 0.1 mg each of D-biotin, folic acid, nicotinamide, calcium D-pantothenate, pyridoxal-HCl, riboflavin, lipoic acid, pyridoxine, and 4-aminobenzoic acid). For preparation of ^{15}N-Phe-labeled FlK or FlK-F36A, this media was prepared with unlabeled ammonium chloride and was supplemented with 100 mg/L of ^{15}N-Phe plus 200 mg/L of each of the other 19 amino acids. This culture was grown overnight and then diluted into 1:50 into 6 × 1 L of the same media in 2.8 L Fernbach flasks. These cultures were grown at 37ºC at 200 rpm to an OD_{600} = 0.7 - 0.8. The cultures were then cooled on ice for 20 min. Protein expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM and cells were incubated at 16 ºC at 200 rpm for 12-16 h. FlK or FlK-F36A was then purified as described previously [*17*].

4.3 Results and discussion

Phe 36, a conformationally dynamic residue at the FlK active site, is important for substrate binding and specificity. Although FlK adopts a canonical hot dog-fold, its structure contains an additional 'lid' that is unique among characterized hotdog-fold thioesterases [*17*, *26*]. While a number of proteins in the NCBI database also seem to contain this lid structure based on sequence conservation and homology modeling, Phe 36 is unique to *Streptomyces cattleya* FlK and an ortholog from *Streptomyces xinghaiensis*, whose genome also encodes a putative fluorinase enzyme that is 99% identical to the *S. cattleya* fluorinase (*Figure 4.1*). Interestingly, Phe 36 appears to be mobile based on a change in its conformation observed upon soaking FlK crystals with substrate or products (*Figure 4.2AB*) [*17*]. Interestingly, the observed

```
Streptomyces cattleya               -MKDGMRVGERFTHDFVVPPHKTVRHLYPESPEFAEFPEVFATGFMVGLMEWACVRAMAP
Streptomyces xinghaiensis           -MREGLLTGEKFTHRYRVPRDKTVPHLYRESPEFSTFPEVFATGFMVGLMEWTCVRAMEP
Anaeromyxobacter sp.Fw-109          -MKSTLAPGVSLTFRYQVPETKTVPHVFPESPRFVEMPQVFATAFMVGLLEWACIEAMQP
Methanosarcina acetivorans          MDSSTLKPGLAYEFRFKIPENKTVPYLYPESPEFQVMPKVFATGFMVGLFEWACIQAINP
Desulfosporosinus meridiei          -MKSTLQSGLSYEFKFTVPEDKTVPYLYPESEEFQAMPKVFGTGFMVGLFEWACIKAINP
Desulfosporosinus youngiae          -MKSTLQSGLYYEFKFTVPENKTVPYLYPESEEFQAMPKVFGTGFMVGLFEWACIKAINP
Gallionella capsiferriformans       -MKDTLKPGIRFEHKYLVPANKTVPALYPESPEFLAMPEVFATGFMVGFLEWACIMAIKP
Sideroxydans lithotrophicus         -MKDTLKPGIRYEHRFLVPSSKTVPALYPEAEEFLAMPEVFATGFLVGFLEWACIKCINP
Rhodovulum sp. PH10                 ------MPGLTHRFSYTVPETKTVPCLYPEAPEFSAMPKVFATGFMVGLMEWTCVRLMAP
Hyphomicrobium sp. MC1              -MKDTLHPGAKTQFTYRVPATKTVPHLYPEAHEFQLMPTVFATGYMVGLMEWTCLHIIAP
Hyphomicrobium zavarzinii           -MKATLQPGAKAQLVFTVPPEKTVPNLYPESPDFRAMPGVFATGFMVGLMEWACLEAMKP
                                            *      : :*  ***  :: *:  *  :* **.*.::**::**:*:  : *

Streptomyces cattleya               YLE-PGEGSLGTAICVTHTAATPPGLTVTVTAELRSVEGRRLSWRVSAHDGVDEIGSGTH
Streptomyces xinghaiensis           YLE-AGEGSLGTAICVAHTAATPPGFTVTVTAELLGIEGRRLKWQVTAHDGVHEIGAGTH
Anaeromyxobacter sp.Fw-109          HLD-GGEQSVGTGIWVTHGAATPPGFTVTVDVAVTKVEGRRLTFSVRAHDGVDAICEGTH
Methanosarcina acetivorans          YLDFPAEQTVGTDVRLSHSAATPPGLTVTVKIKLEKIEGRKLTFSIIADDGVDKISEGTH
Desulfosporosinus meridiei          HLDWPNEQTVGTDVKLSHIAATPPGFTVTVKLRLEKIEGKKLLFHVEAHDGVDQISEGTH
Desulfosporosinus youngiae          HLDWPNEQTVGTDVKLSHLAATPPGLTVTVKLRLEKIEGKKLFFHVEAHDGIDLISEGTH
Gallionella capsiferriformans       HLDWPEEQSVGTHINVSHEAATPPGLEVTASVELTIVDGRRLTFAVSAHDGVDTIARGTH
Sideroxydans lithotrophicus         HIDWPAEQTVGTHINVSHQAATPPGLEVTALVELVEVDGRKLVFQVEAHDGVEVISKGRH
Rhodovulum sp. PH10                 HLD-AGEGSLGVHVDVDHTAATPPGLTVTVDAECVAVDGRRLTFQVTAHDGVDTIGKGRH
Hyphomicrobium sp. MC1              HLD-KGEGSLGVHINVSHLAATVPGQTVTVDAECTKVAGRRLYFHVKAHDGIDLIGEGEH
Hyphomicrobium zavarzinii           HLD-AGEGSLGIHINVSHSAATVPGQTVTVDAECTSVAGRRVAFKVTAHDGVEVIGAGSH
                                    :::   * ::*  : : * *** **  **.      : *..: : : * **:  *  * *

Streptomyces cattleya               ERAVIHLEKFNAKVRQKTPAG-------------
Streptomyces xinghaiensis           ERAVIDVERFTTSLEKKIGRASVETGES------
Anaeromyxobacter sp.Fw-109          ERFVIDRARFDRKIQEKLAASTSC----------
Methanosarcina acetivorans          ERFIIDAAKFNSKAEAKAKNANN-----------
Desulfosporosinus meridiei          ERFIINAAKFNEKVTRKSEIKA------------
Desulfosporosinus youngiae          ERFVIDAAKFNEKVNRKSKTESVTE---------
Gallionella capsiferriformans       ERYVINKEKFDNKLRDKREKINQR----------
Sideroxydans lithotrophicus         ERFIINREKFEAKIGEKMRRSDT-----------
Rhodovulum sp. PH10                 ERVVVSWDKFNARTGEKAKAAGLASAEGE-----
Hyphomicrobium sp. MC1              QRMVVNWEKFEQRVNEKAKIARLAPITRGTV---
Hyphomicrobium zavarzinii           ERMIVPWDRFVGRVNEKAKRAGAAPIAAVESVGA
                                    :* ::   .*      *
```

Figure 4.1. *Sequence alignment of FIK and homologs. The homologs are the top ten hits from NCBI BLAST drawn from sequences that were available as of July 10, 2013. Sequences were aligned using MUSCLE. Phe 36 is shown in red, and the equivalent position in homologs lacking Phe 36 is shown in blue.*

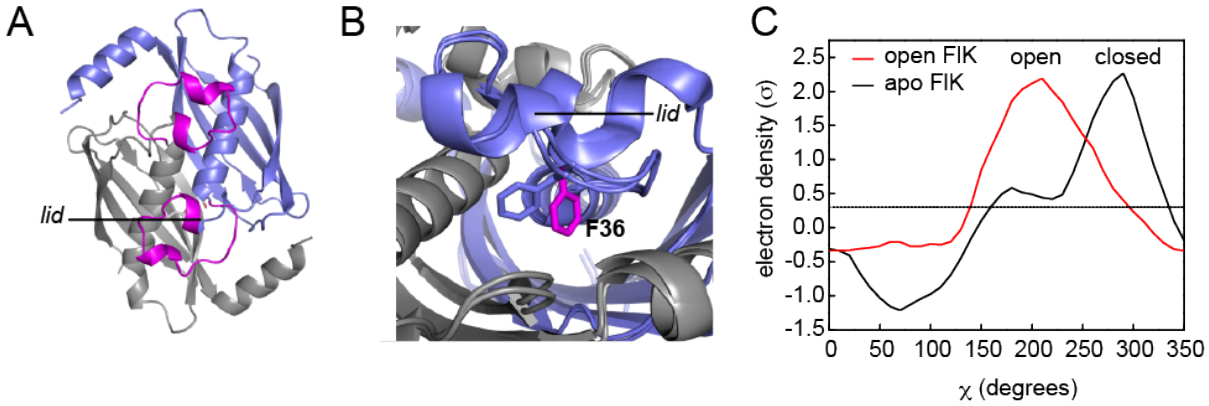


Figure 4.2. *Conformational change observed in FlK upon soaking with substrate or products. (A) The overall fold of FlK showing the locations of the lid (magenta) and the active site (shown in stick form). (B) Structural alignment of FlK in the presence and absence of products. Phe 36 is shown in blue for apo-FlK and in magenta for substrate-soaked FlK. (C) Ringer plot showing the Phe 36 x angle for FlK crystals in the presence (open) and absence (closed) of substrate. Figures are based on PDB entry 3P2Q for apo-FlK and 3P2S for open FlK.*

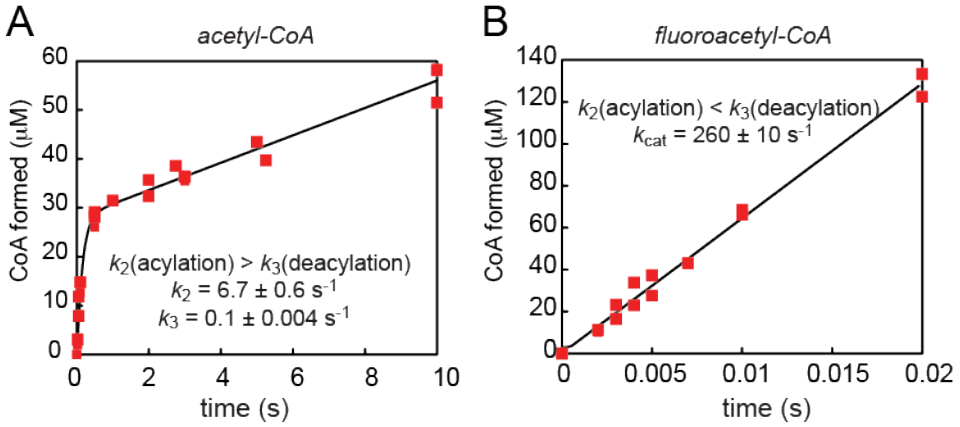


Figure 4.3. *Pre-steady-state kinetics of FlK-F36A. (A) Pre-steady-state time course for FlK-F36A hydrolysis of acetyl-CoA. (B) Pre-steady-state time course for FlK-F36A hydrolysis of fluoroacetyl-CoA.*

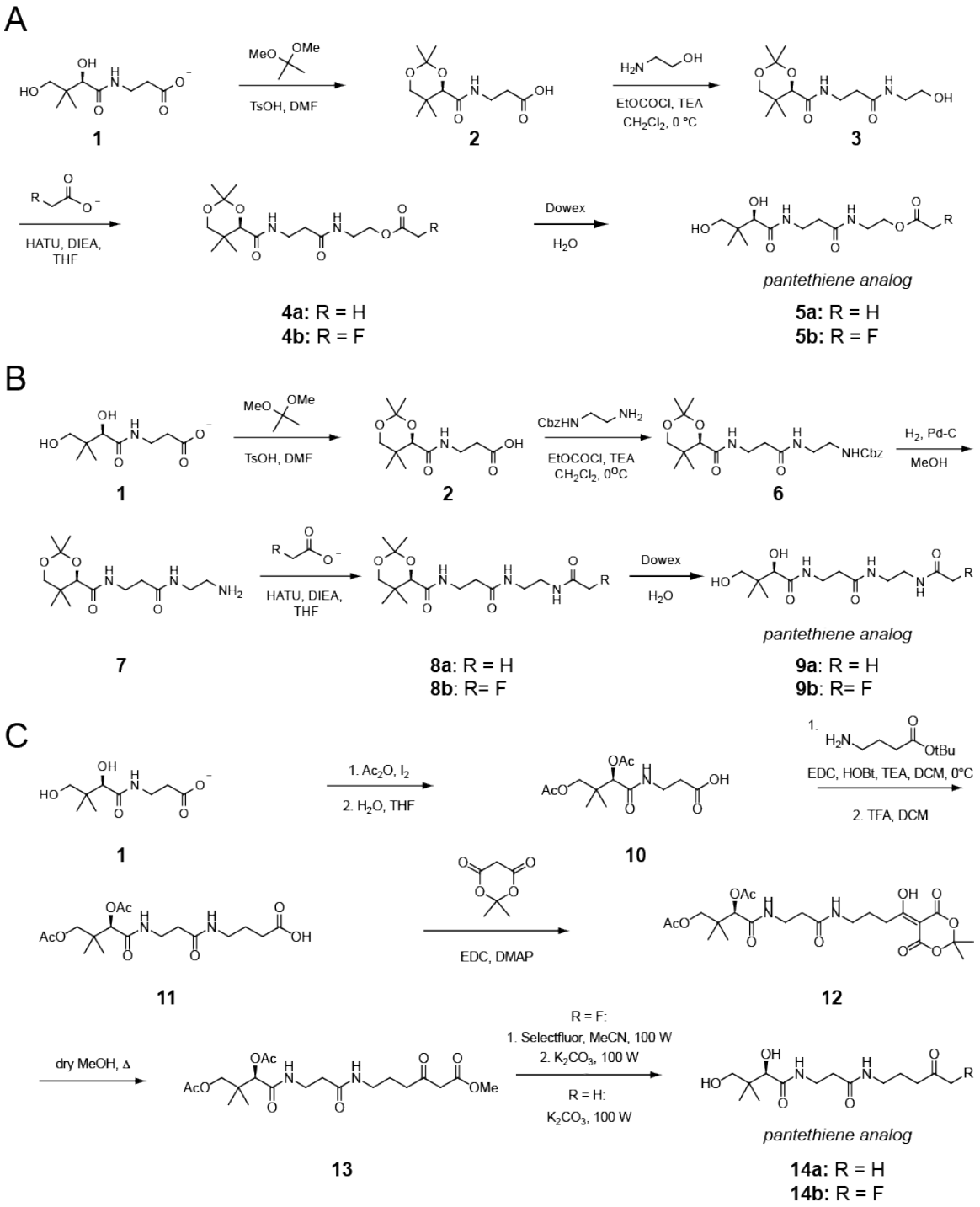
conformational change opens up a channel between the active site and solvent that is normally occluded by Phe 36, which potentially allowing water into the active site. To further explore this possibility, we analyzed the FlK electron density maps using Ringer, a program designed to detect molecular motions through electron density sampling [*28*]. The resulting plots of electron density (σ) versus the χ angle of Phe 36 reveal that apo-FlK appears to sample this alternative rotamer conformation as well, with about 10% occupancy under cryocrystallography conditions (*Figure 4.2BC*).

Biochemical characterization of the F36A mutant had shown that k_{cat}/K_M for the fluoroacetyl-CoA was decreased by 140-fold while leaving the kinetic parameters for acetyl-CoA unchanged [*17*]. As this effect was mainly derived from an increase in K_M for the fluorinated substrate, a role for Phe 36 in substrate binding can be suggested but cannot be directly shown because of the multistep chemical and kinetic mechanism for FlK acyl-CoA hydrolysis [*23*]. To further the

probe the role of Phe 36 in substrate recognition, we examined the pre-steady-state kinetics of the F36A mutant under saturating conditions with both acetyl-CoA and fluoroacetyl-CoA (*Figure 4.3*). For acetyl-CoA, burst-phase kinetic behavior similar to wild-type enzyme was observed, with one equivalent of CoA released rapidly during the first enzyme turnover followed by a slower steady-state release of CoA (*Figure 4.3A*). Based on mechanistic studies of the wild-type enzyme, we conclude that the F36A mutant also follows a two-step chemical mechanism for acetyl-CoA hydrolysis, with rapid formation an acyl-enzyme intermediate followed by slower intermediate breakdown. The rate constants determined for each of these steps by nonlinear curve fitting were identical, within error, to the wild-type enzyme [*23*]. For fluoroacetyl-CoA, we observed single-phase pre-steady-state kinetics, consistent with a change in rate-limiting step from intermediate breakdown to intermediate formation for the fluorinated substrate, also as observed for the wild-type enzyme. Based on linear fitting of the pre-steady-state kinetic data, the k_{cat} for hydrolysis of the fluorinated substrate by the F36A mutant is identical, within error, to that measured for the wild-type enzyme [*23*]. We therefore conclude that mutation of Phe 36 results in a change in K_M by affecting substrate binding, rather than chemical steps involved in substrate hydrolysis.

Fluoroacetyl-CoA and acetyl-CoA bind FlK with different affinities. To further examine the differences in binding affinity between fluoroacetyl-CoA and acetyl-CoA in FlK, we synthesized a series of acyl-CoA analogs in which the thioester sulfur atom was replaced by oxygen, nitrogen, or carbon to produce the less labile or non-hydrolyzeable ester, amide, or ketone congener. We used a synthetic strategy in which the desired acylated pantethiene analogs were first prepared chemically [*24*, *25*] and then converted enzymatically to the corresponding CoA analogs using purified pantothenate kinase, phosphopantethiene adenylyltransferase, and dephospho-CoA kinase from *E. coli* [*24*, *25*, *30*] (*Scheme 4.1*, *Figure 4.4*). Since the carbonyl polarization of these derivatives are different than the thioesters, they were also characterized by NMR to test whether the hydrate or an enolate form existed preferentially in solution (*Figure 4.5, Figure 4.6, Figure 4.7*). Within the limit of detection, none of these compounds showed evidence of any other forms besides the carbonyl species under these conditions. As an additional control, we tested whether FlK could catalyze hydrolysis of the ester or amide substrate analogs. Although esterases and amidases have not been identified in the hotdog-fold superfamily, these activities are observed in the α/β-hydrolase superfamily, which also includdes thioesterases. The ester analog was hydrolyzed very inefficiently, with a k_{cat} of 0.01 min^{-1} (*Figure 4.8A*), more than 500 times less than the rate constant measured for FlK-catalyzed fluoroacetyl-CoA hydrolysis [*17*, *23*], while no detectable hydrolysis was observed for the amide analog after 18 h in the presence of 100 μM FlK (*Figure 4.8B*).

Since hydrolysis of the ester, amide, and ketone substrate analogs by FlK was relatively inefficient, we decided to use them to isolate differences in molecular recognition between fluoroacetyl-CoA and acetyl-CoA since contributions to selectivity from the hydrolytic steps could be avoided. Using isothermal titration calorimetry, we measured the dissociation constant and the contributions made by entropy (ΔS) and enthalpy (ΔH) for each of the substrate analogs (*Table 4.1*, *Figure 4.9*). Based on the K_M measured for fluoroacetyl-CoA, it appears that all of these analogs bind to FlK with a defect of 1-2 orders of magnitude compared to the native substrate, with the ester analogs showing the highest affinity and the amide analogs binding with the lowest affinity. In comparison, studies of the hot dog-fold thioesterase PaaI from *A. evansii*, showed that the non-hydrolyzable substrate analog, 4-hydroxyphenyacyl-CoA, exhibited an



Scheme 4.1. *Synthesis of pantethiene analog precursors of non-hydrolyzable substrate analogs. (A) Synthetic scheme for acetyl- and fluoroacetyl-oxa(dethia)-pantethiene. (B) Synthetic scheme for acetyl- and fluoroacetyl-aza(dethia)-pantethiene. (C) Synthetic scheme for acetyl- and fluoroacetyl-carba(dethia)-pantethiene.*

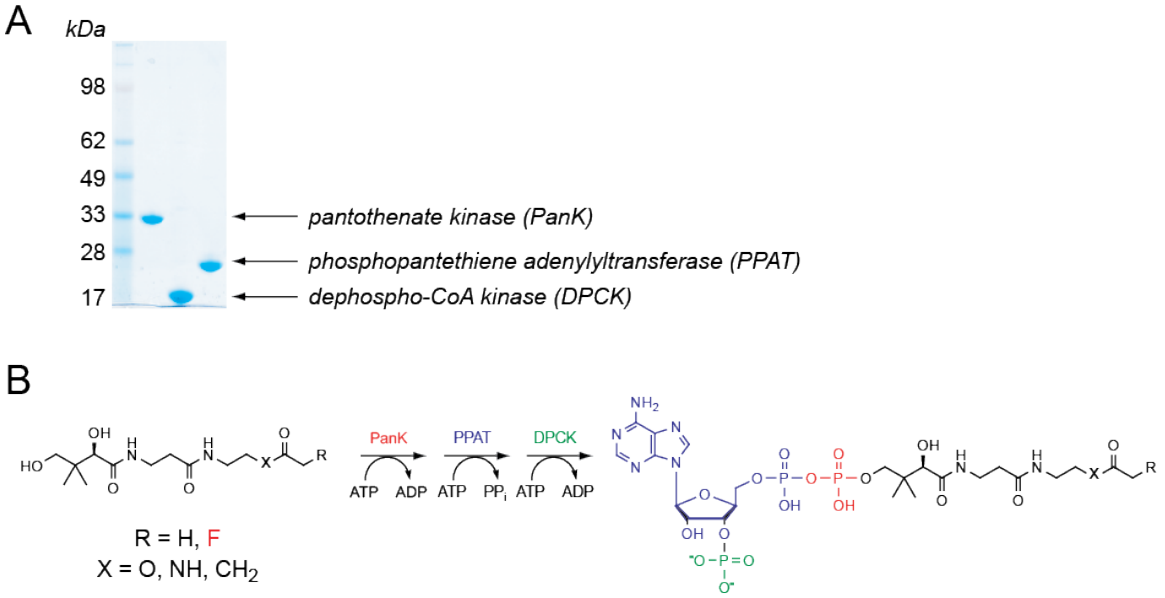


Figure 4.4. *Chemoenzymatic synthesis of non-hydrolyzable substrate analogs. (A) SDS-PAGE gel showing purified pantothenate kinase (PanK), phosphopantethine adenylyltransferase (PPAT), and dephospho-Co kinase (DPCK) from E. coli.*

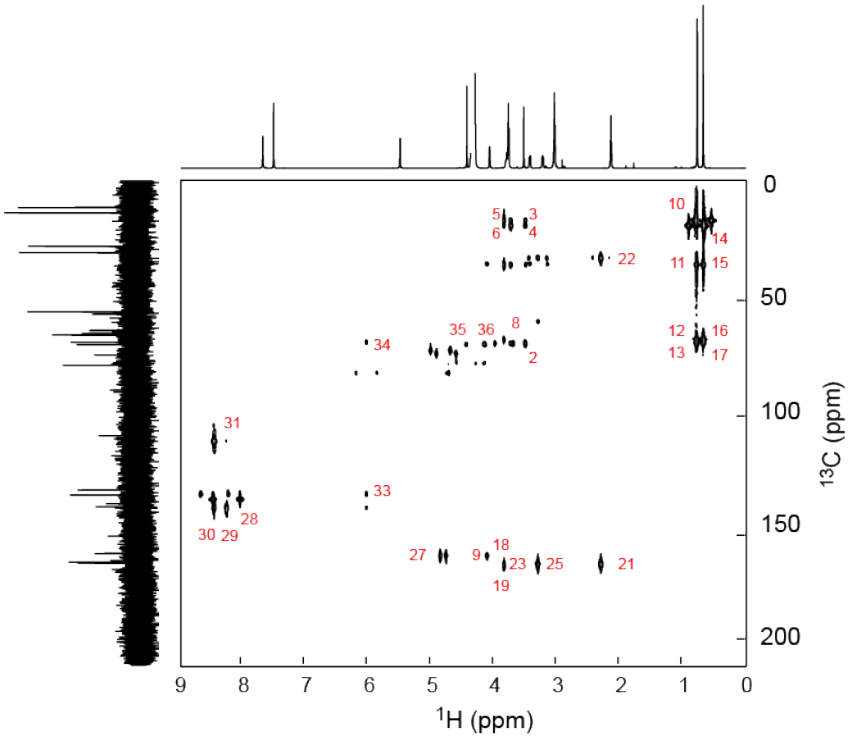


Figure 4.5. *$^{1}H/^{13}C$ HMBC spectrum of fluoroacetyl-oxa(dethia)-CoA. The carbonyl carbon and α-carbon chemical shifts are consistent with an ester.*

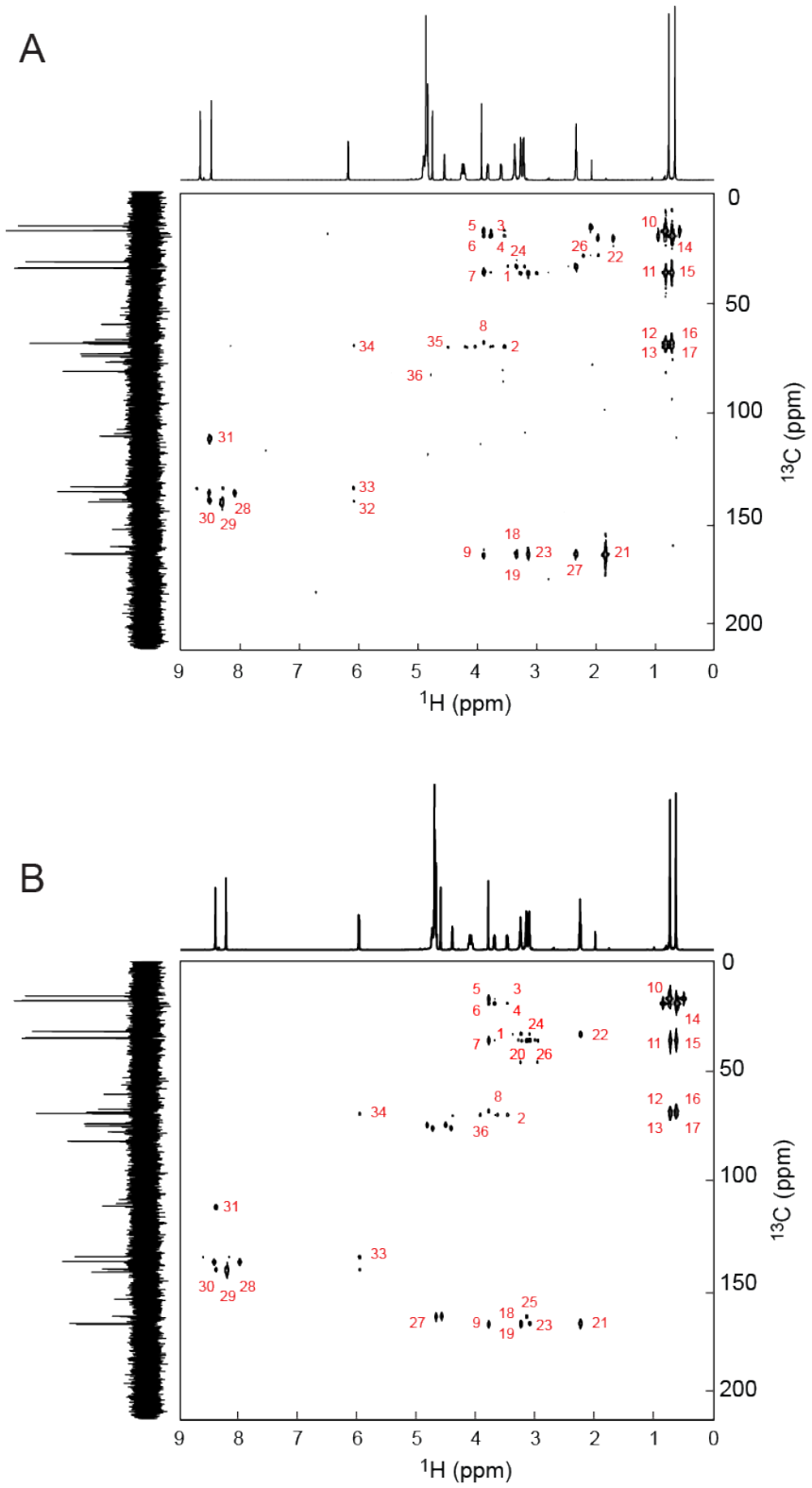


Figure 4.6. *^{1}H/^{13}C HMBC spectra of amide acyl-CoA analogs. (A) Acetyl-aza(dethia)-CoA. (B) Fluoroacetyl-aza(dethia)-CoA. The carbonyl carbon and α-carbon chemical shifts are consistent with an amide.*

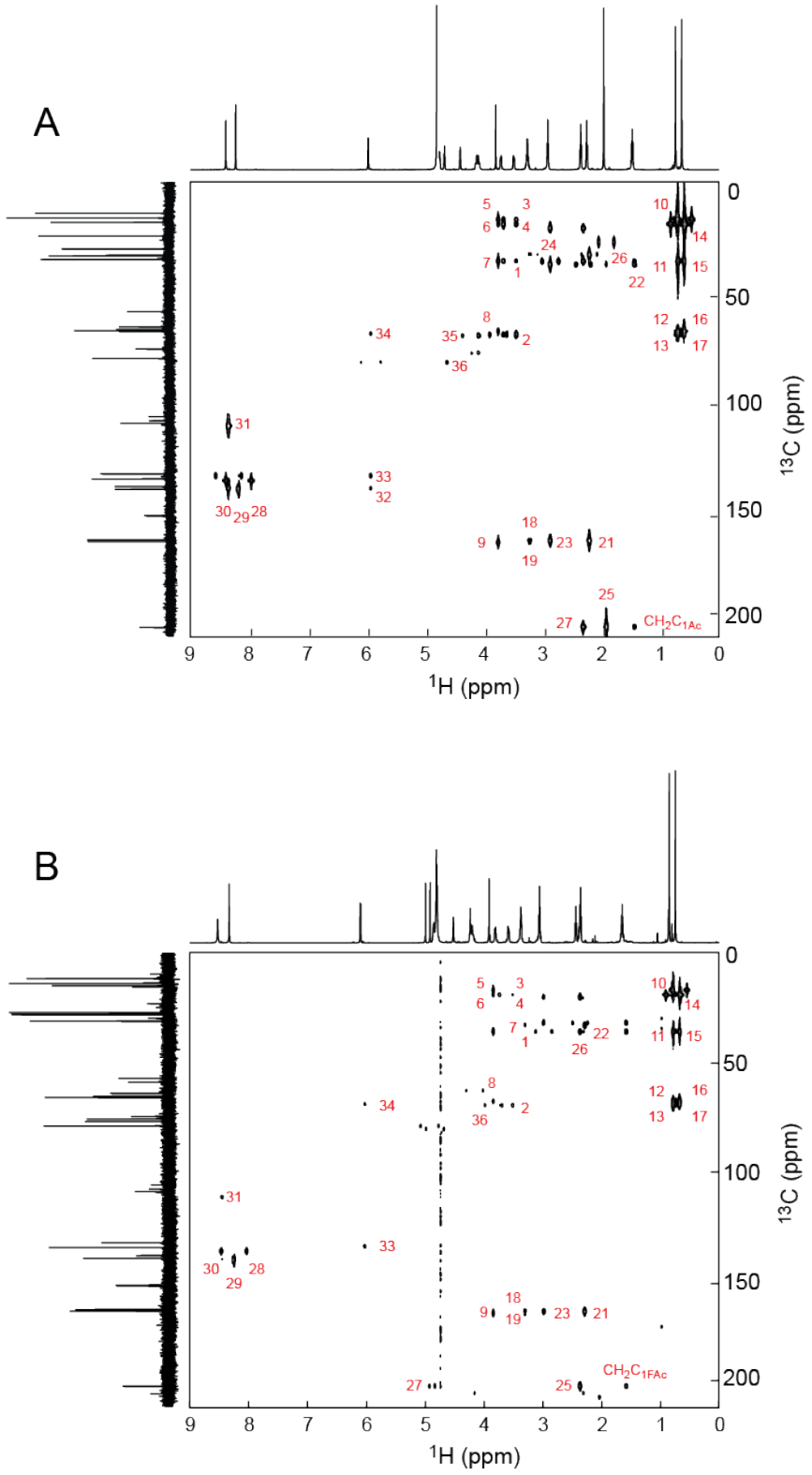


Figure 4.7. *^{1}H/^{13}C HMBC spectrum of ketone acyl-CoA analogs. (A) Acetyl-carba(dethia)-CoA. (B) Fluoroacetyl-carba(dethia)-CoA. The carbonyl carbon and α-carbon chemical shifts are consistent with a ketone.*

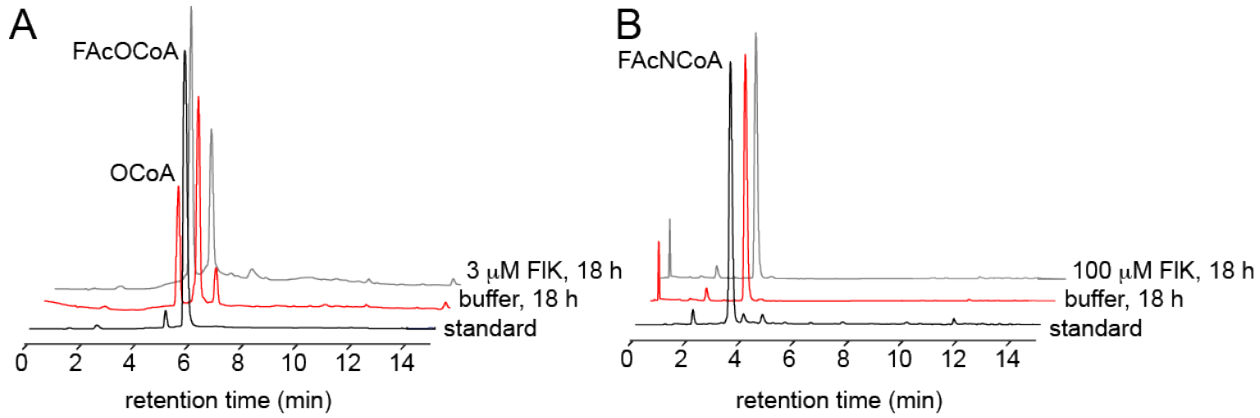


Figure 4.8. *Stability of non-hydrolyzable analogs of fluoroacetyl-CoA. (A) HPLC trace of a fluoroacetyl-oxa(dethia)-CoA standard (black), fluoroacetyl-oxa(dethia)-CoA after incubation in 100 mM Tris-HCl, pH 7.6 for 18 h (red), and after incubation in 100 mM Tris-HCl, pH 7.6 in the presence of 3 μM FlK for 18 h (grey). (B) HPLC trace of a fluoroacetyl-aza(dethia)-CoA standard (black), fluoroacetyl-aza(dethia)-CoA after incubation in 100 mM Tris-HCl, pH 7.6 for 18 h (red), and fluoroacetyl-aza(dethia)-CoA after incubation in 100 mM Tris-HCl, pH 7.6 in the presence of 100 μM FlK for 18 h.*

inhibition constant (K_I) similar to the K_M for the native substrate, 4-hydroxyphenylacetyl-CoA [*19*]. However, this system differs from FlK in two ways. First, PaaI is rather promiscuous, like the vast majority of hot dog-fold thioesterases characterized to date [*31*], which implies that the active site of PaaI may not highly optimized for recognition of a specific substrate. Additionally, this inhibitor differs from the substrate on the CoA side of the analog rather than on the acyl group side. Because FlK binds a single, highly polarized acyl group, it may be able to achieve some degree of substrate discrimination based on differences in carbonyl polarization of the substrate, which would lead to changes in the relative strength of hydrogen bonds or dipolar interactions with this moiety.

Within each fluorine/hydrogen analog pair, we measured a 5-10-fold lower affinity for the acetyl-CoA analogs compared to the fluoroacetyl-CoA analogs (*Table 4.1*, *Figure 4.9*), indicating that fluorine recognition contributes to FlK selectivity. The observed differences in K_D correspond to a ΔΔG of ~1 kcal/mol based on the single fluorine substitution. This magnitude of discrimination rivals the highest reported changes in affinity for binding of fluorinated synthetic inhibitors to their targets in cases where conformation or the pK_as of nearby groups are not the major contributors [*3*, *10*, *11*]. Depending on the inhibitor, such fluorine-based affinity changes may be either enthalpic or entropic in nature. Enthalpic changes based on fluorination have been attributed to either dipoloar interactions between the C-F bond and protein amides, carboxamides, or guanidinium groups, which typically contribute 0.2 – 0.3 kcal/mol per interaction to binding affinity [*3*]. On the other hand, entropic contributions from fluorine are accounted for by increases in hydrophobic surface area based on the polar hydrophobicity of the C-F bond [*3*]. When ΔH and ΔS values are compared within in each analog pair, we find that the fluorinated and nonfluorinated compounds have a similar enthalpy of binding (ΔH) while the fluorinated analogs demonstrate a more favorable entropy of binding (ΔS). This difference is sufficient to explain the overall ΔΔG related to the fluorine substituent, indicating that the energetic contribution of fluorine to binding (K_D) is mainly entropic in nature.

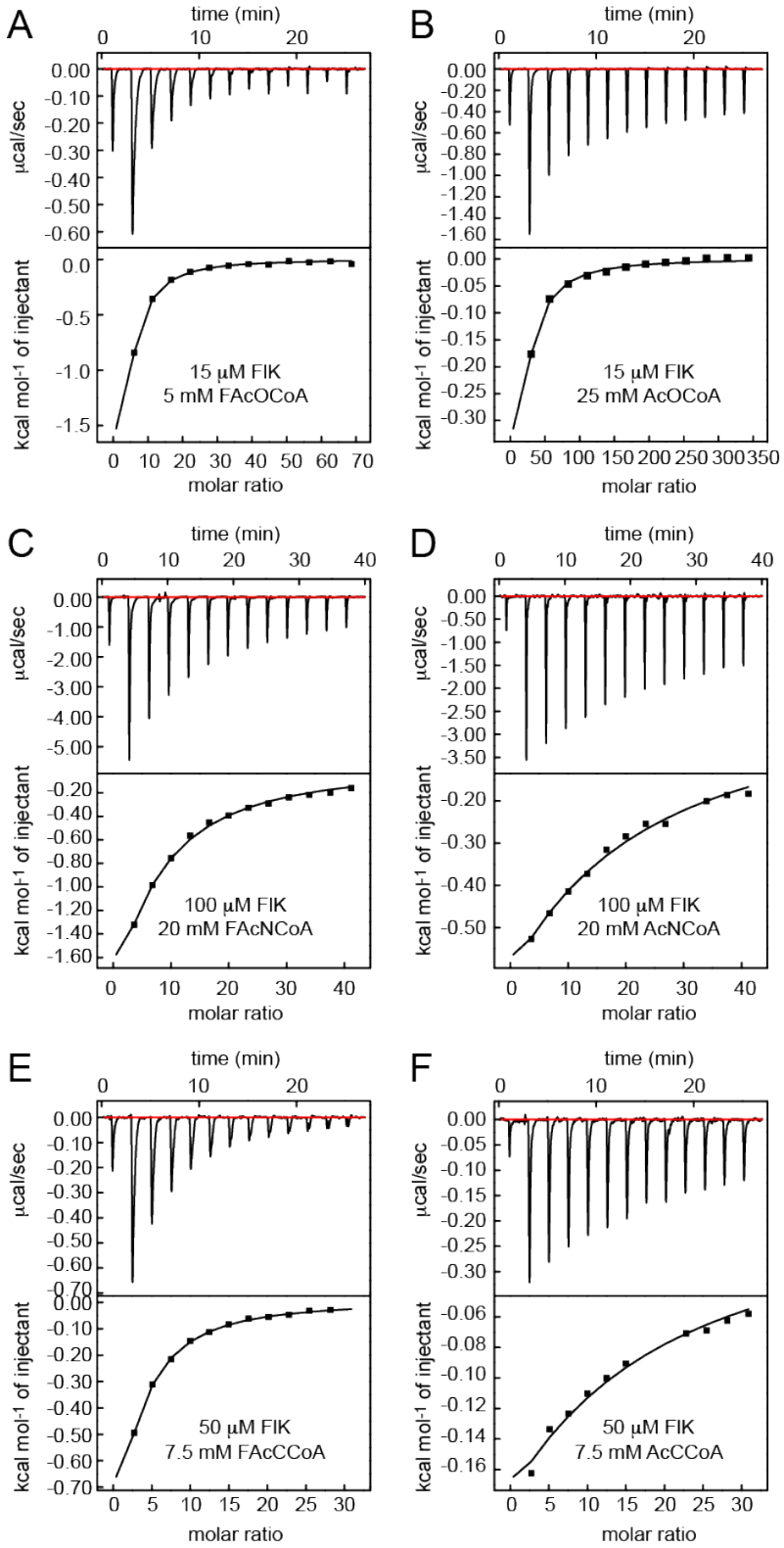


Figure 4.9. *ITC traces for binding of substrate analogs to FIK. (A) FAcOCoA, fluoroacetyl-oxa(dethia)-CoA. (B) AcOCoA, acetyl-oxa(dethia)-CoA. (C) FAcNCoA, fluoroacetyl-aza(dethia)-CoA. (D) AcNCoA, acetyl-aza(dethia)-CoA. (E) FAcCCoA, fluoroacetyl-carba(dethia)-CoA. (F) AcCCoA, acetyl-carba(dethia)-CoA.*

	wild-type FIK			FIK-F36A		
	K_D (μM)	ΔH (kcal/mol)	ΔS (kcal/mol)	K_D (mM)	ΔH (kcal/mol)	ΔS (kcal/mol)
FAcOCoA	90 ± 2	-12.1 ± 0.1	-6.5 ± 0.1	220 ± 30	-13.0 ± 0.1	-8.0 ± 0.1
AcOCoA	565 ± 45	-13.5 ± 0.3	-9.0 ± 0.1	500 ± 70	-9.9 ± 0.6	-6.3 ± 0.1
FAcNCoA	1700 ± 100	-7.4 ± 0.1	-3.7 ± 0.1	7500 ± 200	-7.3 ± 0.1	-4.4 ± 0.1
AcNCoA	9300 ± 500	-8.4 ± 0.3	-5.3 ± 0.1	6200 ± 300	-8.5 ± 0.1	-5.5 ± 0.1
FAcCCoA	440 ± 90	-5.6 ± 0.2	-1.0 ± 0.3	2700 ± 100	-5.6 ± 0.1	-2.1 ± 0.1
AcCCoA	7700 ± 230	-9.4 ± 0.2	-6.6 ± 0.1	n. d.	n. d.	n. d.

Table 4.1. *Thermodynamic parameters measure for substrate analog binding to FIK and FIK-F36A. Errors are derived from non-linear curve fitting. N. d., not determined.*

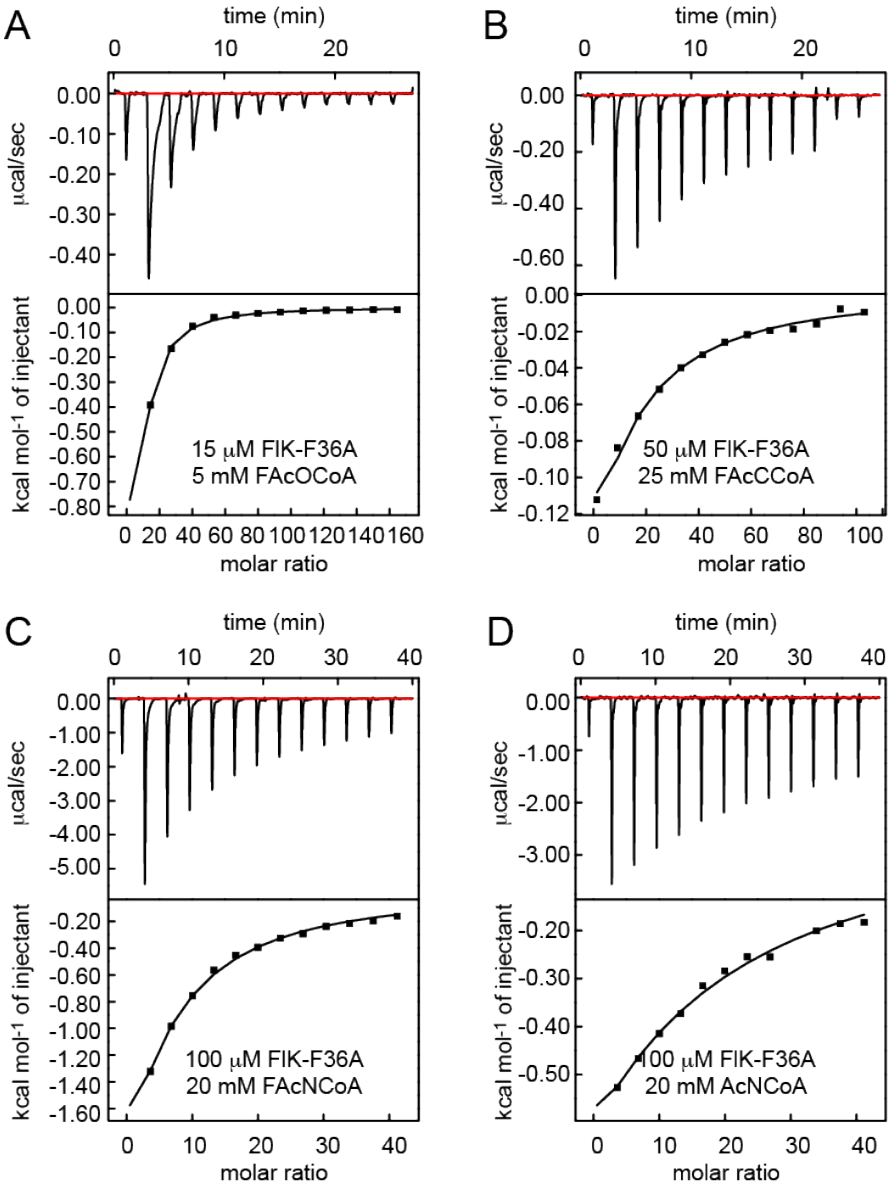


Figure 4.10. *ITC traces for substrate analog binding to FIK-F36A. (A) Fluoroacetyl-ox(dethia)-CoA. (B) Fluoroacetyl-carba(dethia)-CoA. (C) Fluoroacetyl-aza(dethia)-CoA. (D) Acetyl-aza(dethia)-CoA.*

Fluorinated substrate binding is entropically coupled to Phe 36. Because the origin of the binding selectivity for the fluorinated substrate is related to entropy and thus potentially the increase in hydrophobic surface area presented by the C-F group, we next set out to probe the role of the lid structure of FlK in substrate binding. The interior surface of the lid is lined with hydrophobic side chains [*17*, *26*], which appear to serve the dual roles of excluding water from the active site as well as creating a hydrophobic environment. Previous mutagenesis studies of the lid revealed that replacement of any of these residues with alanine led to an increase in K_M for the fluorinated substrate by 1-2 orders of magnitude while leaving the K_M for acetyl-CoA

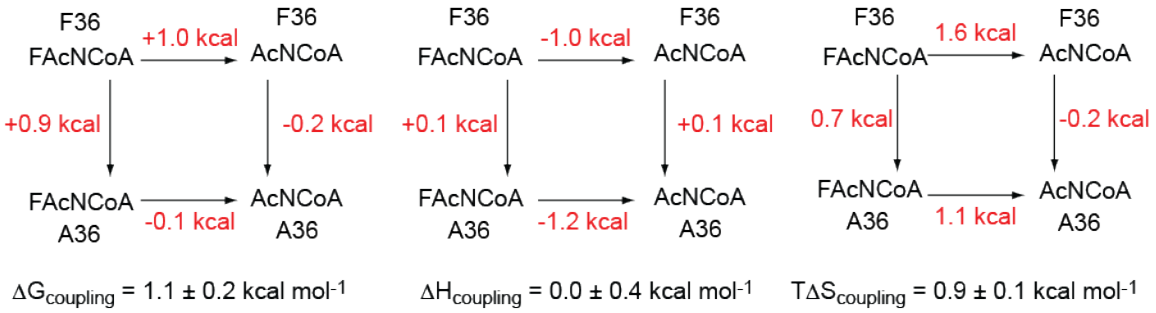
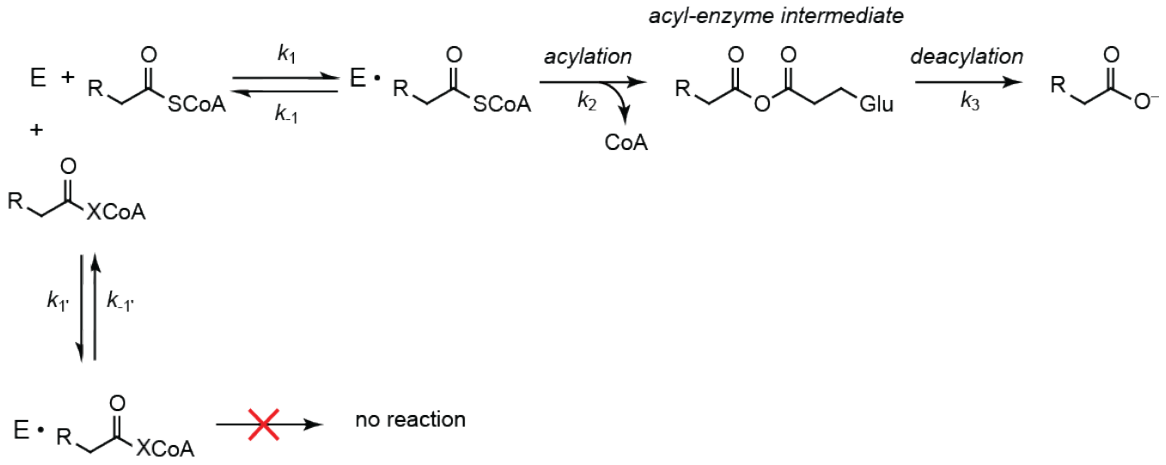


***Figure 4.11**. Double mutant cycle analysis for binding of aza(dethia)-CoA substrate analogs to FlK and FlK-F36A. Errors for coupling energies were propagated from the errors in the individual thermodynamic terms given in Table 4.1.*

relatively unchanged [*17*]. Since Phe 36 seems to both be particularly important in fluorine-selective binding and shows no significant catalytic defect when mutated, the F36A mutant was used to further probe the molecular basis of the entropic advantage on substrate binding provided by fluorine atom. Using ITC, the fluorine/hydrogen pair in each analog series bound to the FlK F36A mutant with similar affinity, demonstrating that fluorine selectivity is lost in the absence of the Phe 36 sidechain (*Table 4.1*, *Figure 4.10*). This behavior is caused by a loss of binding affinity for binding the fluorinated ligand rather than an increase in affinity for the nonfluorinated ligand. Further analysis of the data for the amide inhibitors by a chemical-enzymatic double-mutant cycle [*32*] shows that enhanced binding affinity of the fluorinated compound (ΔG) is energetically coupled to the Phe 36 side chain and entirely entropic in nature, within error (*Figure 4.11*). Phe 36 therefore participates directly in providing an entropic driving force for fluorinated substrate binding.

Phe 36 controls the substrate off-rate. We also examined the ability of our substrate analogs to act as competitive inhibitors of fluoroacetyl-CoA and acetyl-CoA hydrolysis by FlK (*Scheme 4.2*, *Table 4.2*). For FlK-catalyzed acetyl-CoA hydrolysis, all of the substrate analogs exhibited competitive behavior with inhibition constants (K_Is) within error of the dissociation constants measured by ITC. For FlK-catalyzed fluoroacetyl-CoA hydrolysis, all of the substrate analogs also behaved as competitive inhibitors but the measured inhibition constants were significantly higher than the dissociation constants measured by ITC. This observation indicates that binding of fluoroacetyl-CoA to FlK is not in rapid equilibrium and further suggests that the first chemical step (k_2) is faster than the substrate off rate (k_{-1}). Although that the assumption that $k_{-1} > k_2$ is typically made, it often only holds for poor substrates where k_2 is relatively slow but can break down for better substrates. Based on the results of our competitive inhibition experiments, FlK binding to acetyl-CoA is in rapid equilibrium, while the native fluoroacetyl-



Scheme 4.2. *Kinetic scheme for competitive inhibition of FIK by substrate analogs. R = H, F; X = O, NH, CH_2.*

		wild-type	F36A
inhibitor	substrate	K_I (μM)	K_I (μM)
AcOCoA	FAcCoA	2500 ± 1100	1500 ± 300
AcOCoA	AcCoA	510 ± 40	300 ± 30
FAcOCoA	FAcCoA	320 ± 50	260 ± 60
FAcOCoA	AcCoA	130 ± 20	330 ± 60
AcNCoA	FAcCoA	33000 ± 7000	2800 ± 700
AcNCoA	AcCoA	8400 ± 2100	4700 ± 500
FAcNCoA	FAcCoA	4300 ± 1100	4000 ± 500
FAcNCoA	AcCoA	1200 ± 400	7500 ± 700
AcCCoA	FAcCoA	11000 ± 2000	5100 ± 1100
AcCCoA	AcCoA	2600 ± 400	3600 ± 500
FAcCCoA	FAcCoA	900 ± 150	2100 ± 500
FAcCCoA	AcCoA	460 ± 60	4500 ± 800

Table 4.2. *Inhibition constants for competition of substrate analogs with fluoroacetyl-CoA and acetyl-CoA. Errors are derived from propagation of errors from nonlinear fitting of Michaelis-Menten curves in the presence and absence of inhibitor.*

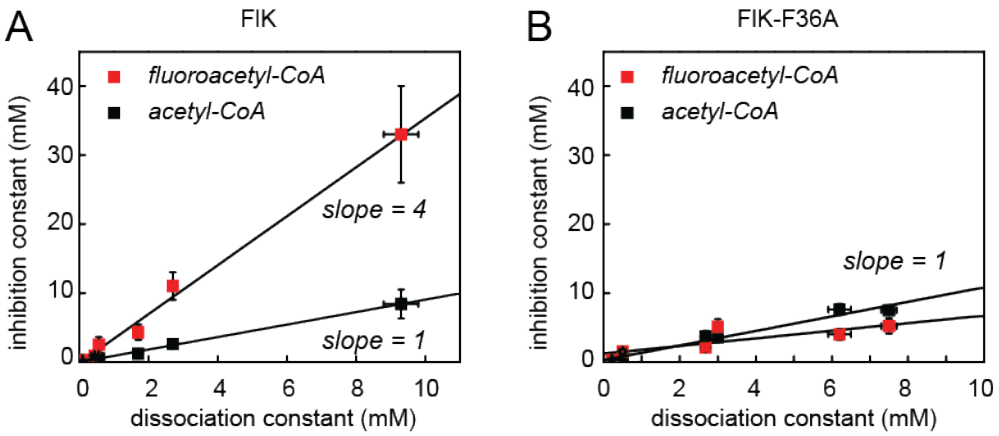


Figure 4.12. *Plots of inhibition constant versus dissociation constant. Fluoroacetyl-CoA substrate is in red and acetyl-CoA substrate is in black. (A) FlK. (B) FlK-F36A.*

CoA substrate carries out the first chemical step, enzyme acylation, faster than the substrate can dissociate. A plot of K_I versus K_D reveals that for both fluoroacetyl-CoA and acetyl-CoA, these two parameters are linearly related with a slope of $k_2/k_{-1} + 1$ derived from the kinetic model for competitive inhibition (*Figure 4.12A*). For acetyl-CoA, k_2 is small relative to the off rate, so k_2/k_{-1} approaches zero and the observed slope is equal to 1. For fluoroacetyl-CoA, the slope of this plot (slope = 4, *Figure 4.12B*) indicates that the enzyme acylation rate constant for this substrate (k_2) is three times faster than the substrate dissociation rate constant (k_{-1}), which explains why the substrate analogs are less effective at competing with the fluoroacetyl-CoA substrate.

The observed change in behavior with respect to the fluorine substituent could be caused either by slower acylation of the enzyme by acetyl-CoA (k_2 = 3 s^{-1}) compared to fluoroacetyl-CoA (k_2 = 270 s^{-1}), or a slower off rate related to the higher affinity of the fluorinated substrate ($K_D = k_{-1}/k_1$), or to a combination of both factors. To probe the role of enhanced binding of fluoroacetyl-CoA, we examined the ability of the non-hydrolyzable substrate analogs to competitively inhibit the F36A mutant (*Table 4.2*), which is compromised only in binding the fluorinated ligand. In this mutant, we again observed competitive behavior by all of the inhibitors with respect to both the fluoroacetyl-CoA and acetyl-CoA substrates. Interestingly, the K_I was within error of K_D and slopes of the plots of K_I versus K_D were equal to one regardless of the substrate used. Thus, k_2/k_{-1} approaches zero for both acetyl-CoA and fluoroacetyl-CoA, showing that rapid equilibrium is restored for fluoroacetyl-CoA in this mutant. Since the rate constant measured by rapid chemical quench for enzyme acylation by fluoroacetyl-CoA is unchanged, we can conclude that mutation of Phe 36 increases the substrate off rate (k_{-1}) such that it becomes faster than the first chemical step (k_2). We thus conclude that Phe 36 is directly involved in controlling the off rate for fluoroacetyl-CoA, allowing for kinetic discrimination of fluorine substituent in substrate binding.

Because the acetyl-CoA substrate is in rapid equilibrium for both wild-type FlK and FlK-F36A, we are unable to detect any potential changes in the off rate for this substrate based on competitive inhibition. However, ITC studies demonstrated that the acetyl-CoA analogs have identical dissociation constants for both wild-type FlK and FlK-F36A. Therefore, if the off rate is increased in the mutant, the on rate must increase by an equal amount to maintain the observed

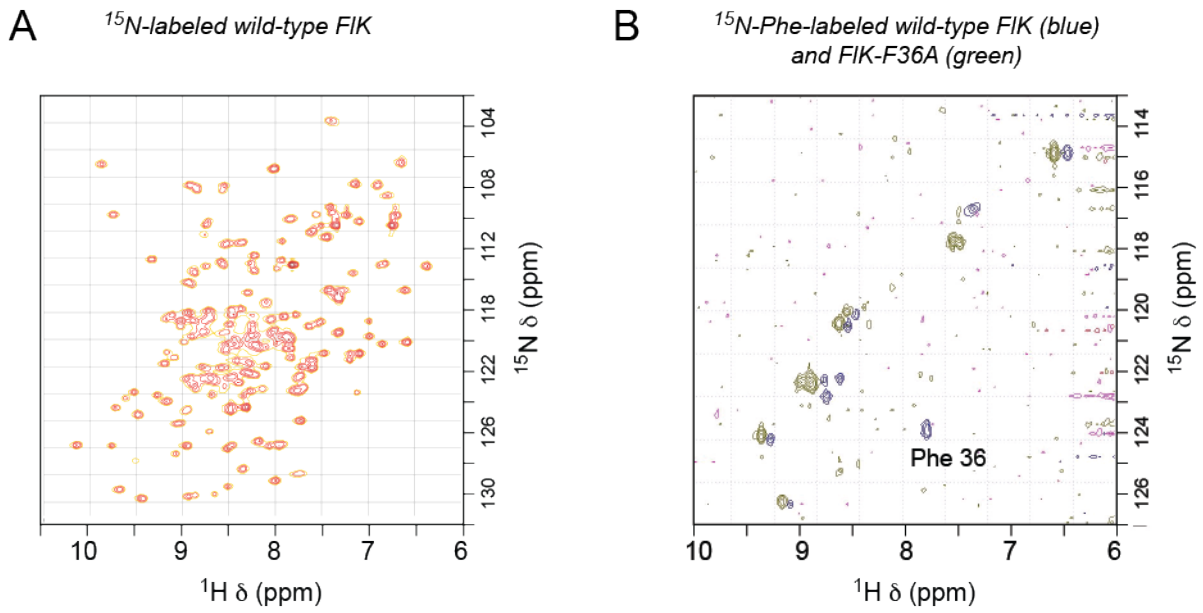


Figure 4.13. *^{1}H-^{15}N HSQC spectra of FlK and FlK-F36A. (A) HSQC spectrum of uniformly ^{15}N-labeled FlK. (B) HSQC spectra of ^{15}N-Phe-labeled FlK (blue) and FlK-F36A (green).*

K_D. The simplest explanation for the ITC data in combination with competitive inhibition studies is that mutation of Phe 36 increases the off rate for fluoroacetyl-CoA while leaving the off rate for acetyl-CoA unchanged.

Probing the role of Phe 36 dynamics in fluoroacetyl-CoA selectivity. Together, our kinetic and thermodynamic studies demonstrate that Phe 36 provides an entropic driving force for fluorinated substrate binding. This entropic favorability could be related to desolvation of the polar hydrophobic C-F unit based on the hydrophobicity of the active site in the Phe-closed state. Another possibility is that there is a favorable gain in conformational entropy upon ligand binding that is related to the crystallographically observed open to closed transition. To probe the role of Phe 36 dynamics in substrate discrimination, we plan to use Carr-Purcell-Meiboom-Gill (CPMG) NMR relaxation dispersion experiments [*33*], which can measure backbone and side chain motions that occur on a catalytic (μs-ms) timescale. Toward this end, we have prepared uniformly ^{15}N-labeled FlK, as well as ^{15}N-Phe-labeled FlK and FlK-F36A. The ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectrum for the uniformly ^{15}N labeled-protein is well dispersed (*Figure 4.13A*), demonstrating FlK's potential suitability for NMR studies. Because FlK contains only seven Phe residues, labeling only Phe with ^{15}N simplifies both the one-dimensional and two-dimensional NMR spectra, and also has the advantage of allowing for peak assignment based on mutagenesis. By comparison of the ^{15}N-Phe-FlK and ^{15}N-Phe-FlK-F36A spectra (*Figure 4.13B*), we have assigned the peak corresponding to Phe 36. Initial studies in which the longitudinal relaxation time T_1 and the transverse relaxation time T_2 were measured and compared did not provide any evidence for μs-ms timescale dynamics of Phe 36. However, it is possible that dynamics are only induced upon ligand binding, as conformational changes were only detected in FlK crystals in the presence of ligand, or that the ^{15}N label is not positioned such that its environment is sensitive to any conformational changes that do occur.

Future studies will be focused on ^{15}N-Phe-labeled samples in the presence of different FlK ligands. Additionally, we plan to employ a biosynthetic labeling strategy using 1-^{13}C-glucose, which will result in enrichment of the δ-position of Phe residues with ^{13}C [*34*, *35*].

4.4 Conclusion

The studies presented here demonstrate that Phe 36 is a key determinant of fluorine molecular recognition in the fluoroacetyl-CoA thioesterase FlK. Phe 36 is entropically coupled to fluorinated substrate binding and therefore provides a thermodynamic driving for preferential binding of fluoroacetyl-CoA compared to acetyl-CoA. Competitive inhibition studies also revealed that Phe 36 is involved in controlling the off rate for fluoroacetyl-CoA, which is increased when this residue is mutated. Based on binding and inhibition studies, this residue does not appear to play the same role in the case of acetyl-CoA. Based on the conformational changes in Phe 36 that were observed in crystallographic studies upon ligand binding, NMR relaxation studies are in progress to probe the role of Phe 36 dynamics in fluorine recognition and binding.

4.5 Acknowledgments

We thank Jared Bard for purification of uniformly ^{15}N-labeled FlK, Dr. Ross Wilson for assistance with ITC studies, and Dr. Jeffrey Pelton for assistance with protein NMR.

4.6 References

1. D. O'Hagan and H. S. Rzepa, Some influences of fluorine in bioorganic chemistry, *Chem. Commun.*, **1997**, 645–652.
2. B. E. Smart, Fluorine substituent effects (on bioactivity), *J. Fluor. Chem.*, **2001**, *109*, 3–11.
3. K. Müller, C. Faeh and F. Diederich, Fluorine in pharmaceuticals: Looking beyond intuition, *Science*, **2007**, *317*, 1881–1886.
4. J. A. Ippolito and D. W. Christianson, The contribution of halogen atoms to protein-ligand interactions, *Int. J. Biol. Macromol.*, **1992**, *14*, 193–197.
5. B. C. Finzel, E. T. Baldwin, G. L. Bryant, G. F. Hess, J. W. Wilks, C. M. Trepod, J. E. Mott, V. P. Marshall, G. L. Petzold, R. A. Poorman, T. J. O'Sullivan, H. J. Schostarez and M. A. Mitchell, Structural characterizations of nonpeptidic thiadiazole inhibitors of matrix metalloproteinases reveal the basis for stromelysin selectivity, *Protein Sci.*, **1998**, *7*, 2118–2126.
6. M. Cairi and J. T. Gerig, Fluorine recognition at the active site of [*N*-(4-fluorophenyl)-*N*-phenylcarbamoyl]-α-chymotrypsin, *J. Am. Chem. Soc.*, **1984**, *106*, 3640–3643.
7. J. Gao, S. Qiao and G. M. Whitesides, Increasing binding constants of ligands to carbonic anhydrase by using 'greasy tails', *J. Med. Chem.*, **1995**, *38*, 2292–2301.
8. D. T. Wong, F. P. Bymaster and E. A. Engleman, Prozac (fluoxetine, Lilly 110140), the first selective serotonin uptake inhibitor and an antidepressant drug: Twenty years since its first publication, *Life Sci.*, **1995**, *57*, 411–441.

9. D. L. Roman, C. C. Walline, G. J. Rodriguez and E. L. Barker, Interactions of antidepressants with the serotonin transporter: A contemporary molecular analysis, *Eur. J. Pharmacol.*, **2003**, *479*, 53–63.

10. J. A. Olsen, D. W. Banner, P. Seiler, U. Obst Sander, A. D'Arcy, M. Stihle, K. Müller and F. Diederich, A fluorine scan of thrombin inhibitors to map the fluorophilicity/fluorophobicity of an enzyme active site: evidence for C-F···C=O interactions, *Angew. Chem. Int. Ed Engl.*, **2003**, *42*, 2507–2511.

11. E. Schweizer, A. Hoffmann-Röder, K. Schärer, J. A. Olsen, C. Faeh, P. Seiler, U. Obst-Sander, B. Wagner, M. Kansy and F. Diederich, A fluorine scan at the catalytic center of thrombin: C–F, C–OH, and C–OMe bioisosterism and fluorine effects on pK_a and log D values, *ChemMedChem*, **2006**, *1*, 611–621.

12. K. Drlica and M. Malik, Fluoroquinolones: Action and resistance, *Curr. Top. Med. Chem.*, **2003**, *3*, 249–282.

13. J. Matsumoto, T. Miyamoto, A. Minamida, Y. Nishimura, H. Egawa and H. Nishimura, 1,4-Dihydro-4-oxopyridinecarboxylic acids as antibacterial agents. 2. Synthesis and structure-activity relationships of 1,6,7-trisubstituted 1,4-dihydro-4-oxo-1,8-naphthyridine-3-carboxylic acids, including enoxacin, a new antibacterial agent, *J. Med. Chem.*, **1984**, *27*, 292–301.

14. D. D. Clarke, Fluoroacetate and fluorocitrate: Mechanism of action, *Neurochem. Res.*, **1991**, *16*, 1055–1058.

15. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme-inhibitor complex, *Proc. Natl. Acad. Sci. U. S. A.*, **1996**, *93*, 13699–13703.

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

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

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

19. F. Song, Z. Zhuang, L. Finci, D. Dunaway-Mariano, R. Kniewel, J. A. Buglino, V. Solorzano, J. Wu and C. D. Lima, Structure, function, and mechanism of the phenylacetate pathway hot dog-fold thioesterase PaaI, *J. Biol. Chem.*, **2006**, *281*, 11028–11038.

20. D. Chen, R. Wu, T. L. Bryan and D. Dunaway-Mariano, *In vitro* kinetic analysis of substrate specificity in enterobactin biosynthetic lower pathway enzymes provides insight into the biochemical function of the hot dog-fold thioesterase EntH, *Biochemistry*, **2009**, *48*, 511–513.

21. M. A. Willis, Z. Zhuang, F. Song, A. Howard, D. Dunaway-Mariano and O. Herzberg, Structure of YciA from *Haemophilus influenzae* (HI0827), a hexameric broad specificity acyl-coenzyme A thioesterase, *Biochemistry*, **2008**, *47*, 2797–2805.

22. J. Cao, H. Xu, H. Zhao, W. Gong and D. Dunaway-Mariano, The mechanisms of human hotdog-fold thioesterase 2 (hTHEM2) substrate recognition and catalysis illuminated by a structure and function based analysis, *Biochemistry*, **2009**, *48*, 1293–1304.

23. A. M. Weeks and M. C. Y. Chang, Catalytic control of enzymatic fluorine specificity, *Proc. Natl. Acad. Sci. U. S. A.*, **2012**, *109*, 19667–19672.

24. I. Nazi, K. P. Koteva and G. D. Wright, One-pot chemoenzymatic preparation of coenzyme A analogues, *Anal. Biochem.*, **2004**, *324*, 100–105.

25. M. Tosin, D. Spiteller and J. B. Spencer, Malonyl carba(dethia)- and malonyl oxa(dethia)-coenzyme A as tools for trapping polyketide intermediates, *Chembiochem*, **2009**, *10*, 1714–1723.

26. M. V. B. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. V. Dry, P. F. Leadlay, J. B. Spencer and T. L. Blundell, Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK, *J. Biol. Chem.*, **2010**, *285*, 22495–22504.

27. W. B. Turnbull and A. H. Daranas, On the value of c: Can low affinity systems be studied by isothermal titration calorimetry?, *J. Am. Chem. Soc.*, **2003**, *125*, 14859–14866.

28. P. T. Lang, H.-L. Ng, J. S. Fraser, J. E. Corn, N. Echols, M. Sales, J. M. Holton and T. Alber, Automated electron-density sampling reveals widespread conformational polymorphism in proteins, *Protein Sci.*, **2010**, *19*, 1420–1431.

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

30. K. M. Clarke, A. C. Mercer, J. J. La Clair and M. D. Burkart, *In vivo* reporter labeling of proteins via metabolic delivery of coenzyme A analogues, *J. Am. Chem. Soc.*, **2005**, *127*, 11234–11235.

31. J. W. Labonte and C. A. Townsend, Active site comparisons and catalytic mechanisms of the hot dog superfamily, *Chem. Rev.*, **2013**, *113*, 2182–2204.

32. P. J. Carter, G. Winter, A. J. Wilkinson and A. R. Fersht, The use of double mutants to detect structural changes in the active site of the tyrosyl-tRNA synthetase (*Bacillus stearothermophilus*), *Cell*, **1984**, *38*, 835–840.

33. J. P. Loria, R. B. Berlow and E. D. Watt, Characterization of enzyme motions by solution NMR relaxation dispersion, *Acc. Chem. Res.*, **2008**, *41*, 214–221.

34. P. Lundström, K. Teilum, T. Carstensen, I. Bezsonova, S. Wiesner, D. F. Hansen, T. L. Religa, M. Akke and L. E. Kay, Fractional ^{13}C enrichment of isolated carbons using [1-^{13}C]- or [2-^{13}C]-glucose facilitates the accurate measurement of dynamics at backbone C_α and side-chain methyl positions in proteins, *J. Biomol. NMR*, **2007**, *38*, 199–212.

35. U. Weininger, M. Respondek and M. Akke, Conformational exchange of aromatic side chains characterized by L-optimized TROSY-selected ^{13}C CPMG relaxation dispersion,

J. Biomol. NMR **2012**, *54*, 9–14.

Chapter 5: *Determinants of specificity in the acylation and deacylation steps of FIK-catalyzed thioester hydrolysis*

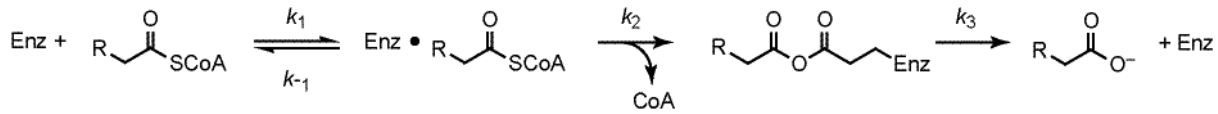
Portions of this work were performed in collaboration with the following persons:

Sodium (*R*)-$^{2}H_{1}$-fluoroacetate and sodium (*S*)-$^{2}H_{1}$-fluoroacetate were synthesized and provided by Dr. Neil S. Keddie, Rudy D. P. Wadoux, and Dr. David O'Hagan

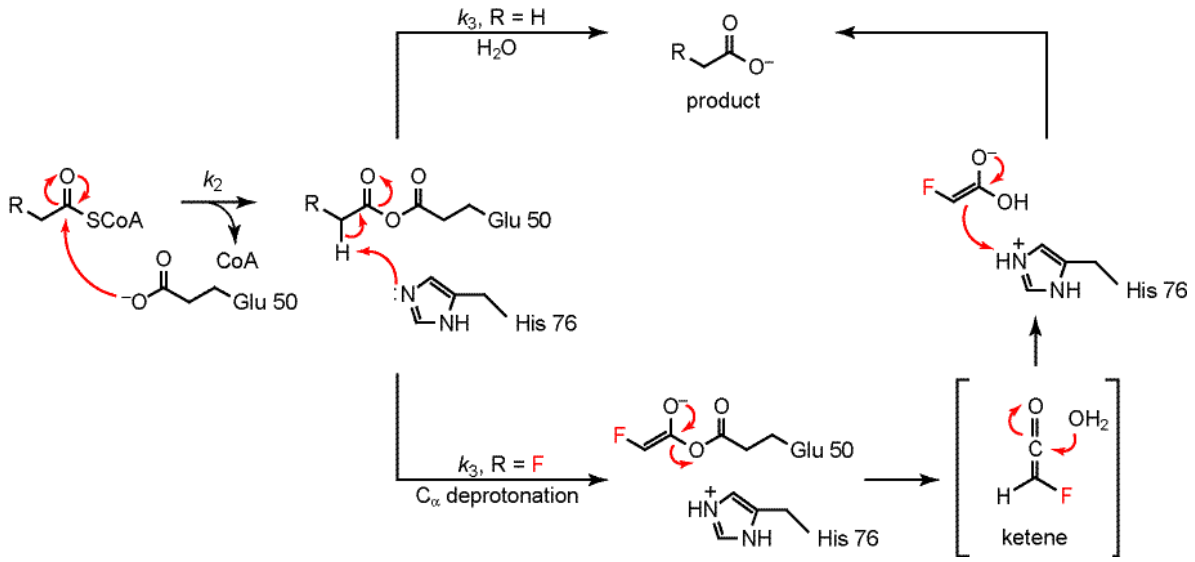
5.1 Introduction

The toxicity of fluoroacetate is attributable to its cellular activation to fluoroacetyl-CoA, a precursor to fluorocitrate, a mechanism-based inhibitor of aconitase that blocks the tricarboxylic acid cycle [*1*, *2*]. As one mechanism of biological resistance, the fluorometabolite-producing bacterium *Streptomyces cattleya* has evolved a fluoroacetyl-CoA thioesterase, FlK [*3*, *4*]. Remarkably, this enzyme exhibits a 10^6-fold preference for its cognate substrate, fluoroacetyl-CoA, compared to acetyl-CoA, an abundant central metabolite and cellular competitor that differs only by the absence of the fluorine substitution [*5*]. Based on its ability to exploit the unique properties of fluorine to achieve substrate specificity, FlK represents an ideal model system in which to query molecular recognition of fluorine and its influence on enzymatic reactivity, with implications for drug design and development of fluorine-selective biocatalysts.

We have previously shown that the hydrolytic mechanism of FlK involves a minimum of three kinetic steps: (i) formation of the enzyme substrate complex ($K_D = k_{-1}/k_1$); (ii) formation of an acyl-anhydride intermediate on Glu 50 (k_2); and (iii) breakdown of this intermediate to give the carboxylic acid product (*Scheme 5.1*) [*6*]. A large component of fluoroacetyl-CoA (R = F) specificity (10^4-fold) is based on a change in rate-limiting step from formation of this intermediate for fluoroacetyl-CoA (ii, $k_2 = k_{cat}$) to its breakdown for acetyl-CoA (R = H) (iii, $k_3 = k_{cat}$). Our previous mechanistic studies indicate that the chemical basis for the acceleration of intermediate breakdown for the fluorinated substrate is utilization of an unusual hydrolytic



Scheme 5.1. *Minimal kinetic mechanism for FlK-catalyzed acyl-CoA hydrolysis.*



Scheme 5.2. *Mechanistic basis for fluoroacetyl-CoA specificity in FlK. Both acetyl-CoA and fluoroacetyl-CoA form an acyl-anhydride intermediate. The fluorinated intermediate can breakdown through a pathway initiated by C_α-deprotonation, providing a chemical basis for fluorine specificity.*

mechanism initiated by C_α-deprotonation of the acyl-enzyme intermediate (*Scheme 5.2*) [*6*]. The resultant enolate can then breakdown through a putative ketene intermediate to give the same carboxylic acid product that would result from a more canonical hydrolysis mechanism.

The differences in chemical and kinetic mechanism between fluoroacetyl-CoA and acetyl-CoA provide a means for FlK to kinetically discriminate between two substrates that are both present at saturating concentrations under physiological conditions based on the high intracellular concentration of acetyl-CoA (2 mM) and the low K_M for fluoroacetyl-CoA (8 μM) [*6*]. However, it is interesting to note that this kinetic discrimination occurs after the first committed step in the reaction mechanism, which is enzyme acylation (*Scheme 5.1*, k_2), and that the specificity constant k_{cat}/K_M depends only on the binding and acylation steps of the reaction (*Scheme 5.1*) [*7*]. We therefore set out to further explore the active site chemistry involved in enzyme acylation. We first provide additional evidence supporting the formation of an acyl-enzyme intermediate during the hydrolysis of fluoroacetyl-CoA, and we show that the rate of formation of this intermediate does not depend on the key specificity-determining residue, His 76, which is involved in C_α-deprotonation. We also explore the role of Thr 42, a second member of the catalytic triad, on acylation and deacylation and argue that its function in catalysis to anchor the hydrogen bonding network between Glu 50 and His 76. We demonstrate that the primary determinant of specificity in enzyme acylation is carbonyl polarization.

In addition to investigating enzyme acylation, we also continue to examine the role of fluorine in enzyme deacylation. Although fluoroacetyl-CoA's α-protons are expected to have a lower pK_a than those of acetyl-CoA, it remains unclear why the C_α-deprotonation mechanism for deacylation (*Scheme 5.2*) is accessible to the fluorinated substrate, but not to substrates with other α-substituents. In particular, acetoacetyl-CoA is predicted to have a lower α-proton pK_a than fluoroacetyl-CoA based on its ability to form a resonance-stabilized enolate, but free-energy relationship analysis of its FlK-catalyzed hydrolysis suggests that it cannot access the C_α-deprotonation pathway [*6*]. Using substrates with stereogenic centers at the substrate α-carbon, we demonstrate that polarization alone is insufficient for reaction through the C_α-deprotonation pathway. We show that FlK utilizes chiral recognition of fluoroacetyl-CoA's prochiral fluoromethyl group to position the *pro*-R proton for preferential abstraction by His 76, indicating that specificity is based not only on the unique reactivity of fluorinated compounds, but also on fluorine molecular recognition. Taken together, our results show that fluorine recognition occurs in both the acylation and deacylation steps of the FlK reaction mechanism. While recognition of fluorine allows FlK to distinguish between two stereoisomers of an unnatural substrate, the acylation rate is primarily determined by polarization. In contrast, fluorine recognition is critical in the deacylation step, positioning the α-protons to allow a unique, specificity-determining reaction pathway that also requires fluorine-based substrate polarization.

5.2 Materials and Methods

Commercial Materials. Acetyl coenzyme A sodium salt, coenzyme A hydrate, coenzyme A trilithium salt, anhydrous *N*, *N*-dimethylformamide (DMF), 2-(*N*-morpholino)ethanesulfonic acid (MES), trifluoroacetic acid (TFA), oxalyl chloride (2 M in dichloromethane), 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB), methyl (*S*)-(-)-lactate, methyl (*R*)-(+)-lactate, methanesulfonyl chloride, magnesium chloride hexahydrate, 2-bis(2-hydroxylethyl)amino-2-(hydroxymethyl)-1,3-propanediol (Bis-Tris), 1,3-bis[tris(hydroxymethyl)methylamino] propane (Bis-Tris propane)

and tris(hydroxymethyl) aminomethane hydrochloride (Tris-HCl) were purchased from Sigma-Aldrich (St. Louis, MO). Monosodium phosphate monohydrate, disodium phosphate heptahydrate, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), hexanes, ethyl acetate, methanol, triethylamine, adenosine-5'-triphosphate trisodium salt, and tris(2-carboxylethyl)phosphine hydrochloride (TCEP) were purchased from Fisher Scientific (Pittsburg, PA). 2-Fluoropropionic acid was purchased from Oakwood Products (West Columbia, SC).

Protein expression and purification. FlK and FlK-T42S were expressed and purified as described previously [*5*]. Acetate kinase from *Escherichia coli* and phosphotransacetylase from *Streptomyces cattleya* were cloned, expressed, and purified as described previously [*4*].

Synthesis of substrates. Fluoroacetyl-CoA, chloroacetyl-CoA, and bromoacetyl-CoA were synthesized as described previously [*6*] Tetrahydrofuran (THF) used in chemical synthesis was dried using a VAC Solvent Purifier System (Vacuum Atmospheres Company, Amesbury, MA). Acyl-CoAs were purified by HPLC using an Agilent Eclipse XDB-C18 column (9.4 × 250 mm, 5 μm) connected to an Agilent 1200 binary pump and an Agilent G1315D diode-array detector, which was used to monitor coenzyme A absorbance at 260 nm (*Appendix 3*). Following sample loading, the column was washed with 0.2% aqueous TFA (3 mL/min) until the absorbance at 260 nm returned to baseline. A linear gradient from 0 to 100 % methanol over 90 min (3 mL/min) with 0.2% aqueous TFA as the aqueous mobile phase. Fractions were collected using an Agilent 1260 fraction collector and then assayed for the desired acyl-CoAs by liquid chromatography-mass spectrometry (LC-MS) using an Agilent 1290 binary pump coupled to an Agilent 6130 single-quadrupole electrospray ionization mass spectrometer. The fractions were analyzed using a Phenomenex Kinetex C18 column (4.6 × 30 mm, 2.6 μm) eluted with a linear gradient from 0 to 100% acetonitrile over 2 min (0.7 mL/min) with 0.1% formic acid as the aqueous mobile phase. Fractions containing the desired compound were pooled and lyophilized. Acyl-CoAs were then dissolved in water, quantified by their absorbance at 260 nm, and stored at -80°C until further use. High-resolution mass spectra were acquired at the QB3/Chemistry Mass Spectrometry Facility at the University of California, Berkeley. NMR spectra were recorded at the College of Chemistry NMR Facility at the University of California, Berkeley. One-dimensional spectra were recorded on Bruker AV-600 or AVQ-400 NMR spectrometers at 298 K. Chemical shifts are expressed in parts per million (δ, ppm) downfield from tetramethylsilane (^{1}H, ^{2}H, and ^{13}C) or trichlorofluoromethane (^{19}F) and are referenced to the solvent signal. Heteronuclear multiple bond correlation (HMBC) experiments were performed on a Bruker AV-500 NMR spectrometer at 298 K (*Appendix 3*).

(*RS*)-2-fluoropropionyl-CoA. 2-Fluoropropionic acid (92 mg, 1 mmol) was placed in an oven-dried round-bottom flask equipped with a stir bar and a reflux condenser. The flask was placed under nitrogen pressure and dry tetrahydrofuran (2 mL), DMF (100 μL), and oxalyl chloride (2 M in dichloromethane, 500 μL) were added by syringe. The reaction mixture was heated to 65°C and stirred for 3-4 h. After cooling, a portion of the reaction mixture (300 μL) was added to a stirred, ice-cooled solution of coenzyme A hydrate (50 mg, ~0.06 mmol) and triethylamine (41.8 mL, 0.3 mmol) in anhydrous DMF (1 mL). After 1 min, the reaction was quenched by addition of water (20 mL) and lyophilized. The lyophilizate was dissolved in water (1 mL) and purified by reverse-phase HPLC (15 μmol, 25%). ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.56 (s, 1H, H_8), 8.38 (s, 1H, H_2), 6.15 (d, J = 6 Hz, 1H, $H_{1'}$), 5.12 (dq, J_{HH} = 7.2, J_{HF} = 49.5 Hz, 1H, C<u>H</u>F), 4.84 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.56 (s, 1H, $H_{4'}$), 4.24 (d, J = 15 Hz, 2H, $H_{5'}$), 3.97

(s, 1H, $H_{3''}$), 3.85 (m, 1H, $H_{1''}$), 3.61 (m, 1H, $H_{1''}$), 3.41 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.31 (t, J = 6.6 Hz, 2H, $H_{8''}$), 3.00 (t, J = 6 Hz, 2H, $H_{9''}$), 2.39 (t, J = 6.6 Hz , 2H, $H_{6''}$), 1.46 (dd, J_{HH} = 6.6 Hz, J_{HF} = 24.9 Hz, 3H, C$\underline{H}_3$), 0.90 (s, 3H, $H_{10''}$), 0.78 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 203.2, 203.0 (FCH$\underline{C}$O), 174.6 ($C_{7''}$), 173.9 ($C_{4''}$), 149.8 (C_6), 144.7 (C_2), 142.7 (C_4), 142.4 (C_8), 118.5 (C_5), 87.9 ($C_{1'}$), 87.5, 86.8 (F$\underline{C}$HCO), 86.8 ($C_{4'}$), 74.4 ($C_{2'}$), 74.3 ($C_{3''}$), 74.2 ($C_{3'}$), 73.7 ($C_{1''}$), 65.1 ($C_{5'}$), 38.3 ($C_{8''}$), 36.9 ($C_{2''}$) 35.3 ($C_{5''}$), 35.2 ($C_{6''}$), 27.1 ($C_{9''}$), 20.7 ($\underline{C}H_3$), 18.3 ($C_{10''}$), 17.9 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -181.2 (sextet, J = 26.4 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 840.1247, found (M-H$^+$) *m/z* 840.1236.

(*S*)- and (*R*)-2-fluoropropionyl-CoA. Sodium (*S*)-2-fluoropropionate and sodium (*R*)-2-fluoropropionate were synthesized from methyl (*R*)-(-)-lactate and methyl (*S*)-(+)-lactate using literature methods [*8*]. 2-fluoropropionyl-CoAs were then synthesized from the corresponding carboxylic acids (1 mmol-scale) as described for (*RS*)-2-fluoropropionyl-CoA (10 μmol, 17%).

(S)-2-fluoropropionyl-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.54 (s, 1H, H_8), 8.33 (s, 1H, H_2), 6.10 (d, J = 6 Hz, 1H, $H_{1'}$), 5.08 (dq, J_{HH} = 7.2, J_{HF} = 49.5 Hz, 1H, C$\underline{H}$F), 4.87 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.52 (s, 1H, $H_{4'}$), 4.21 (d, J = 15 Hz, 2H, $H_{5'}$), 3.93 (s, 1H, $H_{3''}$), 3.81 (m, 1H, $H_{1''}$), 3.58 (m, 1H, $H_{1''}$), 3.37 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.27 (t, J = 6.6 Hz, 2H, $H_{8''}$), 2.97 (t, J = 6 Hz, 2H, $H_{9''}$), 2.35 (t, J = 6.6 Hz , 2H, $H_{6''}$), 1.41 (dd, J_{HH} = 6.6 Hz, J_{HF} = 24.9 Hz, 3H, C$\underline{H}_3$), 0.86 (s, 3H, $H_{10''}$), 0.75 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 203.1, 202.9 (FCH$\underline{C}$O), 174.6 ($C_{7''}$), 173.9 ($C_{4''}$), 149.6 (C_6), 145.3 (C_2), 143.9 (C_4), 141.6 (C_8), 118.3 (C_5), 87.9 ($C_{1'}$), 87.8, 86.8 (F$\underline{C}$HCO), 86.8 ($C_{4'}$), 74.8 ($C_{2'}$), 74.6 ($C_{3''}$), 74.4 ($C_{3'}$), 73.7 ($C_{1''}$), 65.1 ($C_{5'}$), 38.2 ($C_{8''}$), 37.5 ($C_{2''}$) 36.0 ($C_{5''}$), 35.3 ($C_{6''}$), 27.9 ($C_{9''}$), 21.1 ($\underline{C}H_3$), 20.1 ($C_{10''}$), 17.6 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -181.3 (sextet, J = 26.4 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 840.1247, found (M-H$^+$) *m/z* 840.1229.

(R)-2-fluoropropionyl-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.60 (s, 1H, H_8), 8.39 (s, 1H, H_2), 6.17 (d, J = 6 Hz, 1H, $H_{1'}$), 5.14 (dq, J_{HH} = 7.2, J_{HF} = 49.5 Hz, 1H, C$\underline{H}$F), 4.87 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.57 (s, 1H, $H_{4'}$), 4.25 (d, J = 15 Hz, 2H, $H_{5'}$), 3.98 (s, 1H, $H_{3''}$), 3.85 (m, 1H, $H_{1''}$), 3.62 (m, 1H, $H_{1''}$), 3.42 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.32 (t, J = 6.6 Hz, 2H, $H_{8''}$), 3.02 (t, J = 6 Hz, 2H, $H_{9''}$), 2.40 (t, J = 6.6 Hz , 2H, $H_{6''}$), 1.47 (dd, J_{HH} = 6.6 Hz, J_{HF} = 24.9 Hz, 3H, C$\underline{H}_3$), 0.90 (s, 3H, $H_{10''}$), 0.79 (s, 3H, $H_{11''}$); ^{13}C NMR (150.9 MHz, D_2O, 25°C) δ (ppm): 203.2, 203.0 (FCH$\underline{C}$O), 174.7 ($C_{7''}$), 174.0 ($C_{4''}$), 149.8 (C_6), 148.5 (C_2), 144.7 (C_4), 142.5 (C_8), 118.5 (C_5), 93.5, 92.3 (F$\underline{C}$HCO), 87.5 ($C_{1'}$), 83.9 ($C_{4'}$), 74.4 ($C_{2'}$), 74.2 ($C_{3''}$), 73.7 ($C_{3'}$), 72.2 ($C_{1''}$), 65.1 ($C_{5'}$), 38.34 ($C_{8''}$), 38.28 ($C_{2''}$), 35.3 ($C_{5''}$), 35.2 ($C_{6''}$), 27.1 ($C_{9''}$), 20.7 ($\underline{C}H_3$), 18.3 ($C_{10''}$), 17.9 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -181.2 (sextet, J = 26.4 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 840.1247, found (M-2H$^+$) *m/z* 840.1236.

(R)- or (S)-1-(methoxycarbonyl) ethyl methane sulfonate. Methyl lactate (1.0 g, 9.6 mmol) was dissolved in toluene before adding triethylamine (1.2 g, 1.6 mL, 11.5 mmol). The reaction mixture was cooled in an ice water bath and methanesulfonyl chloride (1.2 g, 0.82 mL, 10.6 mmol) was added dropwise over 15 min. The reaction mixture was stirred on ice for 1 h and then warmed to room temperature and filtered. The filtrate was concentrated under vacuum to yield a light pink oil, which was purified on a silica column using 80 hexanes: 20 ethyl acetate as the mobile phase to give (*S*)- or (*R*)-1-(methoxycarbonyl)-ethyl methane sulfonate as a clear oil (1.43 g, 84%).

(R)-1-(methoxycarbonyl) ethyl methane sulfonate: ^{1}H NMR (600 MHz, $CHCl_3$, 25°C) δ (ppm): 5.11 (q, J = 6.6 Hz, 1H, C$\underline{H}$), 3.78 (s, 3H, C$\underline{H}_3$O), 3.13 (s, 3H, C$\underline{H}_3$S), 1.59 (d, J = 7.2

Hz, C$\underline{H}_3$CH); ^{13}C NMR (150.9 MHz, $CHCl_3$, 25°C) δ (ppm): 170.1 ($\underline{C}$OOMe), 74.3 ($\underline{C}$H), 52.9 (O$\underline{C}H_3$), 39.2 (S$\underline{C}H_3$), 18.5 (CH$\underline{C}H_3$). HR-ESI-MS calcd (M+Na$^+$) *m/z* 205.0141, found (M+Na$^+$) *m/z* 205.0143.

(S)-1-(methoxycarbonyl) ethyl methane sulfonate: ^{1}H NMR (600 MHz, $CHCl_3$, 25°C) δ (ppm): 5.10 (q, J = 7.2 Hz, 1H, C$\underline{H}$), 3.77 (s, 3H, C$\underline{H}_3$O), 3.12 (s, 3H, C$\underline{H}_3$S), 1.58 (d, J = 7.2 Hz, C$\underline{H}_3$CH); ^{13}C NMR (150.9 MHz, $CHCl_3$, 25°C) δ (ppm): 170.1 ($\underline{C}$OOMe), 74.3 ($\underline{C}$H), 52.9 (O$\underline{C}H_3$), 39.2 (S$\underline{C}H_3$), 18.5 (CH$\underline{C}H_3$). HR-ESI-MS calcd (M+Na$^+$) *m/z* 205.0141, found (M+Na$^+$) *m/z* 205.0143.

Methyl (S)- or (R)-2-fluoropropionate. Formamide (3 mL) was placed in a two-neck round bottom flask equipped with a stir bar and heated to 60°C. Potassium fluoride (1.2 g, 22 mmol) was dissolved in the formamide with stirring. A short-path distillation head equipped with a thermometer and a receiving flask cooled with a dry ice-acetone bath was then attached. (*S*)- or (*R*)-1-(methoxycarbonyl) ethyl methanesulfonate (1 g, 5.5 mmol) was added slowly. The reaction mixture was then placed under vacuum and the product was continuously distilled during the course of the reaction to yield methyl (*R*)- or (*S*)-2-fluoropropionate as a clear oil (200 mg, 35%).

Methyl (S)-2-fluoropropionate: ^{1}H NMR (600 MHz, $CHCl_3$, 25°C) δ (ppm): 5.02 (dq, J_{HH} = 7.2 Hz, J_{HF} = 48.3 Hz, 1H, C$\underline{H}$), 3.80 (s, 3H, C$\underline{H}_3$O), 1.58 (dd, J_{HH} = 6.6 Hz, J_{HF} = 30.6 Hz); ^{13}C NMR (150.9 MHz, $CHCl_3$, 25°C) δ (ppm): 171.2, 171.1 ($\underline{C}$OOMe), 86.5, 85.3 ($\underline{C}$H), 52.6 (O$\underline{C}H_3$), 18.6, 18.4 (CH$\underline{C}H_3$); ^{19}F NMR (564.6 MHz, $CHCl_3$, 25°C) δ (ppm): -185.6 (sextet, J = 24 Hz).

Methyl (R)-2-fluoropropionate: ^{1}H NMR (600 MHz, $CHCl_3$, 25°C) δ (ppm): 5.03 (dq, J_{HH} = 7.2 Hz, J_{HF} = 48.6 Hz, 1H, C$\underline{H}$), 3.80 (s, 3H, C$\underline{H}_3$O), 1.59 (dd, J_{HH} = 6.6 Hz, J_{HF} = 30.6 Hz); ^{13}C NMR (150.9 MHz, $CHCl_3$, 25°C) δ (ppm): 171.2, 171.0 ($\underline{C}$OOMe), 86.4, 85.3 ($\underline{C}$H), 52.5 (O$\underline{C}H_3$), 18.5, 18.4 (CH$\underline{C}H_3$); ^{19}F NMR (564.6 MHz, $CHCl_3$, 25°C) δ (ppm): -185.6 (sextet, J = 24 Hz).

Sodium (S)- or (R)-2-fluoropropionate. Methyl (*R*)- or (*S*)-2-fluoropropionate (100 mg, 0.6 mmol) was dissolved in water and sodium hydroxide (0.6 mmol from a 10 M stock solution) was added. The reaction mixture was stirred at room temperature for 1 h, at which time it was judged to be ~80% complete based on ^{19}F NMR. The solution was flash frozen and lyophilized to yield sodium (*R*)- or (*S*)-2-fluoropropionate as a white solid (56 mg, 82%).

Sodium (S)-2-fluoropropionate: ^{1}H NMR (400 MHz, D_2O, 25°C) δ (ppm): 4.91 (dq, J_{HH} = 7.2 Hz, J_{HF} = 51.6 Hz, 1H, C$\underline{H}$), 1.49 (dd, J_{HH} = 6.8 Hz, J_{HF} = 17.4 Hz, 3H, CHC$\underline{H}_3$); ^{13}C NMR (100.6 MHz, D_2O, 25°C) δ (ppm): 179.2, 179.0 ($\underline{C}$OONa), 88.9, 87.7 ($\underline{C}$H), 18.3, 18.2 ($\underline{C}H_3$); ^{19}F NMR (375.6 MHz, D_2O, 25°C) δ (ppm): -174.6 (sextet, J = 26.4 Hz).

Sodium (R)-2-fluoropropionate: ^{1}H NMR (400 MHz, D_2O, 25°C) δ (ppm): 4.91 (dq, J_{HH} = 6.8 Hz, J_{HF} = 48 Hz, 1H, C$\underline{H}$), 1.48 (dd, J_{HH} = 6.8 Hz, J_{HF} = 18 Hz, 3H, CHC$\underline{H}_3$); ^{13}C NMR (100.6 MHz, D_2O, 25°C) δ (ppm): 179.3, 179.1 ($\underline{C}$OONa), 89.2, 87.4 ($\underline{C}$H), 18.2, 17.9 ($\underline{C}H_3$); ^{19}F NMR (375.6 MHz, D_2O, 25°C) δ (ppm): -172.8 (sextet, J = 28 Hz).

(*S*)- and (*R*)-$^2H_1$-fluoroacetyl-CoA. Sodium (*S*)-$^2H_1$-fluoroacetate and sodium (*R*)-$^2H_1$-fluoroacetate were obtained as a gift from David O'Hagan [*9*]. (*S*)- and (*R*)-$^2H_1$-Fluoroacetyl-CoA were synthesized from the corresponding sodium $^2H_1$-fluoroacetates enzymatically using

acetate kinase (AckA) from *Escherichia coli* and phosphotransacetylase (PTA) from *Streptomyces cattleya*. Reaction mixtures containing Tris-HCl, pH 7.5 (100 mM), $MgCl_2$ (5 mM), TCEP (2.5 mM), ATP (50 mM), coenzyme A (40 mM), sodium $^2H_1$-fluoroacetate (50 mM), AckA (10 μM), and PTA (5 μM) in a total volume of 1 mL were incubated at 37°C for 2 h. $^2H_1$-Fluoroacetyl-CoAs were then purified by reverse-phase HPLC. Fractions containing the desired compounds were pooled and lyophilized to yield (*S*)- or (*R*)-$^2H_1$-fluoroacetyl-CoA (6 μmol, 15%).

(S)-$^2H_1$-fluoroacetyl-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.62 (s, 1H, H_8), 8.41 (s, 1H, H_2), 6.19 (d, J = 6 Hz, 1H, $H_{1'}$), 4.98 (d, J = 46 Hz, 1H, $C\underline{H}_2F$), 4.87 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.58 (s, 1H, $H_{4'}$), 4.26 (d, J = 15 Hz, 2H, $H_{5'}$), 3.99 (s, 1H, $H_{3''}$), 3.85 (m, 1H, $H_{1''}$), 3.62 (m, 1H, $H_{1''}$), 3.43 (t, J = 6.6 Hz, 2H, $H_{5''}$), 3.35 (t, J = 6.6 Hz, 2H, $H_{8''}$), 3.07 (t, J = 6.6 Hz, 2H, $H_{9''}$), 2.42 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.92 (s, 3H, $H_{10''}$), 0.80 (s, 3H, $H_{11''}$); ^{13}C NMR (225 MHz, D_2O, 25°C) δ (ppm): 202.8, 202.7 ($FCH_2\underline{C}O$), 177.5 ($C_{7''}$), 176.8 ($C_{4''}$), 152.7 (C_6), 151.3 (C_2), 147.3 (C_4), 145.3 (C_8), 121.4 (C_5), 90.5 ($C_{1'}$), 86.9, 86.1 ($F\underline{C}H_2CO$), 83.0 ($C_{4'}$), 77.0 ($C_{2'}$), 76.9 ($C_{3''}$), 76.6 ($C_{3'}$), 76.5 ($C_{1''}$), 67.8 ($C_{5'}$), 45.0 ($C_{8''}$), 41.1 ($C_{2''}$) 38.1 ($C_{5''}$), 38.0 ($C_{6''}$), 29.7 ($C_{9''}$), 23.5 ($C_{10''}$), 21.0 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -226.6 (dt, J_{HF} = 49.2 Hz, J_{DF} = 7.8 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 827.1154, found (M-H$^+$) *m/z* 827.1151.

(R)-$^2H_1$-fluoroacetyl-CoA. ^{1}H NMR (600 MHz, D_2O, 25°C) δ (ppm): 8.64 (s, 1H, H_8), 8.42 (s, 1H, H_2), 6.20 (d, J = 6 Hz, 1H, $H_{1'}$), 4.99 (d, J = 46 Hz, 1H, $C\underline{H}_2F$), 4.87 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.59 (s, 1H, $H_{4'}$), 4.26 (d, J = 13 Hz, 2H, $H_{5'}$), 4.00 (s, 1H, $H_{3''}$), 3.87 (m, 1H, $H_{1''}$), 3.62 (m, 1H, $H_{1''}$), 3.44 (t, J = 6 Hz, 2H, $H_{5''}$), 3.36 (t, J = 5.4 Hz, 2H, $H_{8''}$), 3.08 (t, J = 6 Hz, 2H, $H_{9''}$), 2.43 (t, J = 6.6 Hz , 2H, $H_{6''}$), 0.92 (s, 3H, $H_{10''}$), 0.81 (s, 3H, $H_{11''}$); ^{13}C NMR (225 MHz, D_2O, 25°C) δ (ppm): 202.8, 202.7 ($FCH_2\underline{C}O$), 177.5 ($C_{7''}$), 176.8 ($C_{4''}$), 152.7 (C_6), 151.3 (C_2), 147.4 (C_4), 145.3 (C_8), 121.4 (C_5), 90.2 ($C_{1'}$), 86.8, 86.4 ($F\underline{C}H_2CO$), 83.0 ($C_{4'}$), 76.9 ($C_{2'}$), 76.7 ($C_{3''}$), 76.6 ($C_{3'}$), 76.5 ($C_{1''}$), 67.8 ($C_{5'}$), 45.0 ($C_{8''}$), 41.3 ($C_{2''}$) 38.1 ($C_{5''}$), 38.0 ($C_{6''}$), 29.7 ($C_{9''}$), 23.6 ($C_{10''}$), 20.9 ($C_{11''}$); ^{19}F NMR (564.7 MHz, D_2O, 25°C) δ (ppm): -226.6 (dt, J_{HF} = 48.6 Hz, J_{DF} = 7.2 Hz). HR-ESI-MS calcd (M-H$^+$) *m/z* 827.1154, found (M-H$^+$) *m/z* 827.1146.

Pre-steady state kinetic analysis. Pre-steady-state kinetic experiments were performed using rapid chemical quench followed by HPLC separation of coenzyme A from unhydrolyzed acyl-CoA. FlK or FlK-T42S (50 μM in 20 mM Tris-HCl, pH 7.6, 50 mM NaCl) was mixed with substrate diluted in water (chloroacetyl-CoA and $^2H_1$-fluoroacetyl-CoAs, 500 μM; cyanoacetyl-CoA and fluoroacetyl-CoA, 750 μM; bromoacetyl-CoA and (*R*)-2-fluoropropionyl-CoA, 1 mM; (*S*)-2-fluoropropionyl-CoA, 5 mM; acetyl-CoA, 6 mM) using a Chemical Quench Flow Model RQF-3 (KinTek). The reaction was stopped at various times by mixing with 50% TFA to achieve a final concentration of 17% TFA. Quenched samples were analyzed by HPLC on an Agilent Eclipse XDB-C18 (3.5 μm, 3.0 × 150 mm) using a linear gradient from 0 to 100% methanol over 15 min (0.6 mL/min) with 50 mM sodium phosphate, 0.1% TFA, pH 4.5 as the mobile phase. Conversion percentages were calculated based on the peak areas for substrate and based on the CoA absorbance at 260 nm. Plots of coenzyme A formed versus time were fit to the equation $[CoA] = A*(1-\exp(-k_2*t)) + V*t$, where A is the burst amplitude, k_2 is the burst-phase rate constant, and V is the steady-state rate. In cases where data could not be fit with a burst phase, they were fit as pseudo-first-order reactions according to the equation $[CoA] = k[FlK]t + [CoA]_0$,

where the y-intercept, $[CoA]_0$, is the concentration of CoA released before the first data point, k is the rate constant, and t is time.

For pre-steady state kinetics experiments at pH 6, FlK was exchanged into 100 mM Bis-Tris, pH 6.0 prior to the experiment using a MicroBiospin 6 column (Bio-Rad) that had been pre-equilibrated with the appropriate buffer. FlK was then quantified by absorbance and diluted to 50 μM with the same buffer prior to use in rapid quench experiments.

Taft free-energy relationship analysis of FlK acylation rates. Taft plots were constructed by plotting the log of the acylation rate constant (log k_2) versus the Taft polar substituent constant σ* for the α-substituent of each acyl-CoA tested [*10*]. Acylation rate constants and the corresponding error bars were derived from non-linear curve fitting of triplicate pre-steady state kinetic data sets. Taft plots were fit to the equation log k = ρ*σ* + c, where ρ* is the polar sensitivity factor and c is a constant.

Steady-state pH-rate profiles. Steady-state kinetic experiments for determining the pH-rate profile of FlK were carried out on a Beckman Coulter DU800 spectrophotometer by continuously monitoring the absorbance of the acyl-CoA thioester bond at 232 nm (ε = 4,500 M^{-1} cm^{-1}) in a 0.2 cm path length quartz cuvette at 25°C. Reaction mixtures contained the appropriate buffer (100 mM; Bis-Tris propane, pH 6.3-7; Tris-HCl, pH 7-9; Bis-Tris propane, pH 9), fluoroacetyl-CoA (10, 25, 50 or 100 μM), and FlK (1 nM) or FlK-H76A (1 μM). Each rate was measured in triplicate. Experiments at saturation (k_{cat} conditions) were repeated with similar results in a series of sulfonate buffers (MES, pH 6-7; HEPES, pH 7-8; TAPS, pH 7.5-9) to ensure the observed activity resulted from changes in pH and not buffer composition. Kinetic parameters (k_{cat} and K_M) were determined by fitting the initial rate data to the equation $V_0 = V_{max}[S]/K_M + [S]$, where V_0 is the initial rate and [S] is the substrate concentration, using Origin 6.0 (OriginLab Corporation, Northampton, MA). pH-rate profiles in which decreases in activity were observed at both low and high pH were fit to the equation:

$$k_{obs} = k_{max} / (1 + 10^{pKa1-pH} + 10^{pH-pKa2}),$$

where k_{obs} is the observed k_{cat} at a given pH, k_{max} is the pH independent rate constant, pK_{a1} is the pK_a of the group that ionizes at basic pH, and pK_{a2} is the pK_a of the group that ionizes at acidic pH. pH-rate profiles in which only one ionization at basic pH was observed were fit to the equation:

$$k_{obs} = k_{max} / (1 + 10^{pKa-pH}),$$

where k_{obs} is the observed k_{cat} at a given pH, k_{max} is the pH independent rate constant, pK_{a1} is the pKa of the group that ionizes at basic pH.

Steady-state kinetic analysis of 2-fluoropropionyl-CoAs. Steady-state rates of FlK-catalyzed 2-fluoropropionyl-CoA hydrolysis were measured by monitoring the increase in absorbance at 412 nm from the reaction of 5,5'-dithiobis-2-nitrobenzoic acid (DTNB) with enzymatically-produced coenzyme A in a Beckman Coulter DU800 spectrophotometer. Assays (100 μL) were performed at 25°C in Tris-HCl, pH 7.6 (100 mM) containing DTNB (0.5 mM), the appropriate 2-fluoropropionyl-CoA, and FlK (5 nM). Each rate was measured in triplicate. Absorbance values were converted to coenzyme A concentration using a standard curve. Kinetic parameters (k_{cat} and K_M) were determined by fitting the initial rate data to the equation $V_0 = V_{max}[S]/K_M + [S]$, where V_0 is the initial rate and [S] is the substrate concentration, using Origin 6.0.

5.3 Results and discussion

Probing the role of the catalytic residues in acylation specificity. Since acylation of Glu 50 represents the committed step in FlK-catalyzed hydrolysis, we were interested in further exploring the contribution of the other catalytic residues, His 76 and Thr 42, to this step. Our previous pre-steady state kinetic analyses indicated that the acylation rate constant differs by at least two orders of magnitude between the fluoroacetyl-CoA and acetyl-CoA substrates and that mutation of His 76 to Ala led to a loss of specificity in both the acylation and deacylation steps [*6*]. While the defect in deacylation can be attributed to the loss of the key base required for the C_α-deprotonation pathway (*Scheme 5.2*), the 10-fold reduction in the acylation rate constant for the H76A mutant is more difficult to interpret due to the participation of the His 76 sidechain in a hydrogen bond network involving Thr 42 and Glu 50 that has been shown to be important for their positioning (*Figure 5.1*) [*5*, *11*]. However, active roles for His 76 in determining substrate specificity for enzyme acylation beyond simple structural function are also possible. One potential role could involve deprotonation of the fluoroacetyl-CoA α-carbon by His 76 to initiate an E1Cb-like (elimination-addition) mechanism to generate a ketene, which could then be trapped by Glu 50, to form the acyl-enzyme intermediate in an analogous fashion to its breakdown, Another possibility is that His 76 could be involved in the deprotonation of the Glu 50 nucleophile to promote nucleophilic attack on the fluoroacetyl-CoA carbonyl group. We therefore set out to clarify whether general base catalysis involving His 76 is essential for controlling the observed specificity in acylation.

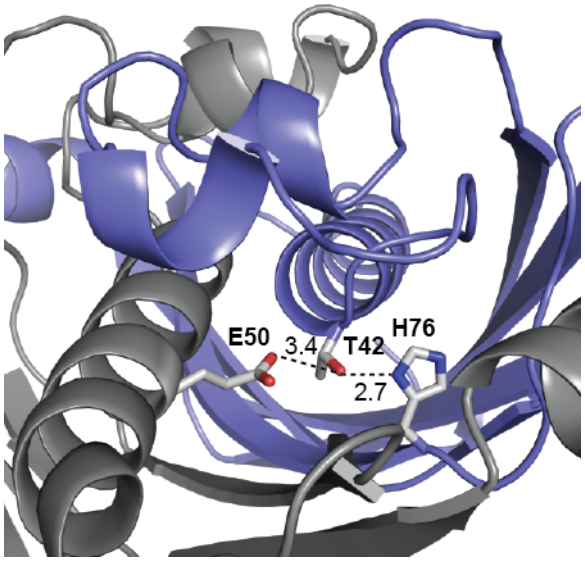


Figure 5.1. *Catalytic triad of FlK. The catalytic residues of FlK participate in a hydrogen bonding network, indicated by dashed lines. Distances are given in angstroms. Carbon, grey; oxygen, red; nitrogen, blue.*

We initiated these studies by examining FlK acylation under conditions in which His 76 is protonated and therefore unable to participate in general base catalysis, which should minimize structural perturbations in the active site compared to mutagenesis. To determine appropriate conditions, we first constructed a pH-rate profile of FlK and observed an ionization event with a pK_a of 6.6 ± 0.2 (*Figure 5.2A*). The absence of this ionization in the pH-rate profile of the FlK-H76A mutant is consistent with its assignment to His 76 (*Figure 5.2*B). Although a second

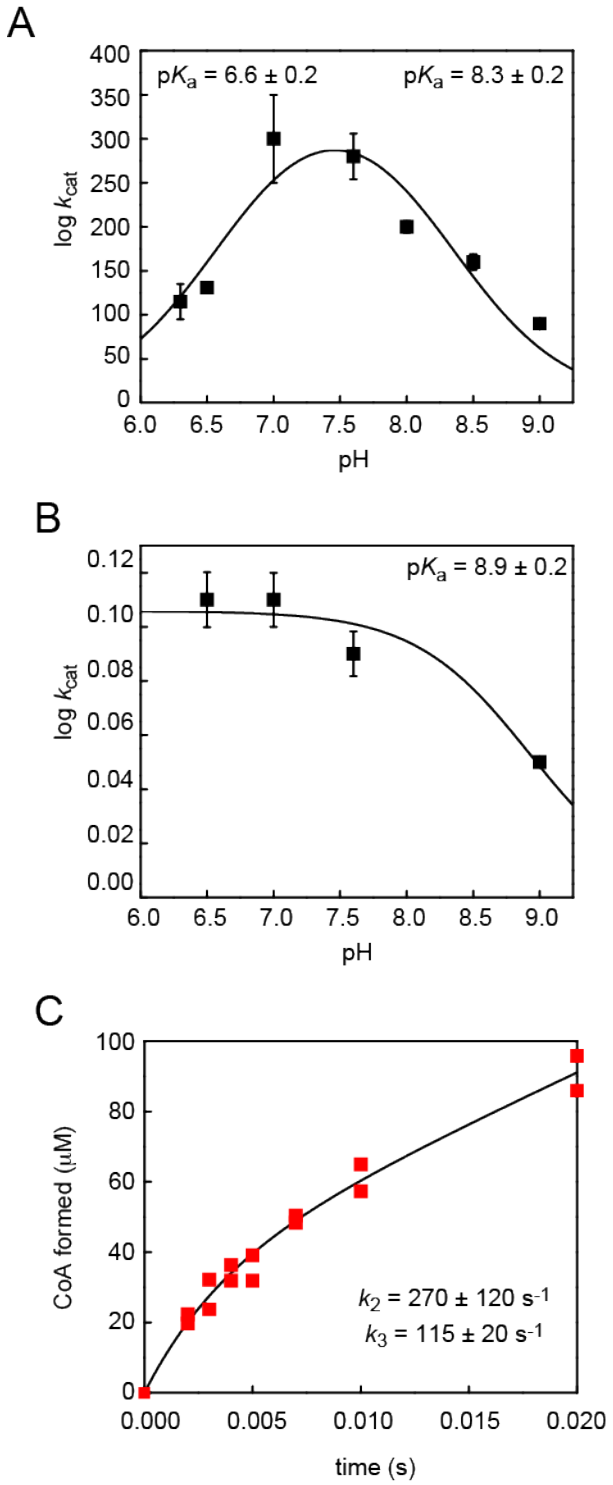


Figure 5.2. *pH dependence studies for exploring the role of His 76 in formation of the FIK acyl-enzyme intermediate. (A) The wild-type enzyme produces ionizations with pK_as of 6.6 ± 0.2 and 8.3 ± 0.2. (B) FIK-H76A produces only one ionization with a pK_a of 8.9 ± 0.2. (C) Pre-steady state kinetic analysis of FIK-catalyzed fluoroacetyl-CoA at pH 6, where His 76 is expected to be protonated.*

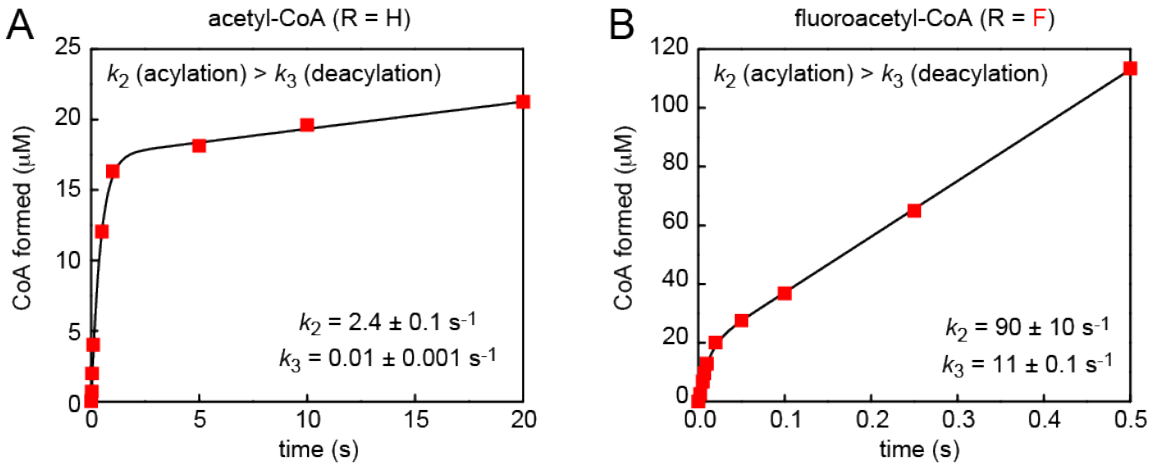


***Figure 5.3**. Pre-steady-state kinetic analysis of FIK-T42S. (A) FIK-T42S-catalyzed acetyl-CoA hydrolysis. (B) FIK-T42S-catalyzed fluoroacetyl-CoA hydrolysis.*

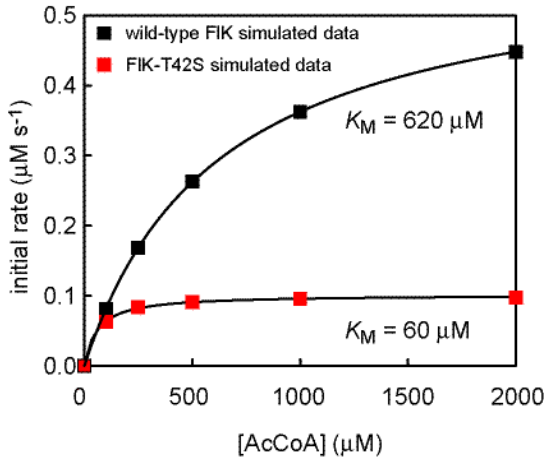


***Figure 5.4.** Simulated steady-state kinetic data for FIK-T42S-catalyzed acetyl-CoA hydrolysis. Data were simulated using the same K_D for both wild-type and mutant and using the acylation and deacylation rate constants measured using pre-steady-state kinetic analysis.*

ionization event with a pK_a of 8.3 ± 0.2 was also observed, it is not clearly attributable to any of the residues known to be involved in catalysis. The existence of a similar ionization in the H76A mutant (pK_a = 8.9 ± 0.2, *Figure 5.2B*), as well as in other hotdog-fold thioesterases [*12*], suggests that the origin of this deprotonation event is more complex and not directly related to His 76. We then performed pre-steady state kinetic studies of FlK at pH 6, where His 76 should be predominantly protonated. Under these conditions, we observed a burst phase of product formation consistent with enzyme acylation followed by a slower steady-state rate, which we interpret to be caused by the inaccessibility of the C_α-deprotonation pathway in the absence of general base catalysis by His 76 (*Figure 5.2C*). The observed burst phase rate constant (270 ± 120 s^{-1}) demonstrates the kinetic competence of the protonated His 76 in forming the acyl-enzyme intermediate. In comparison, no burst phase in observed in FlK at pH 7.6 and fluoroacetyl-CoA is turned over at a steady state rate of 220-270 s^{-1}. These results demonstrate that general base catalysis involving His 76 is not required for acylation activity or fluoroacetylation specificity.

Since His 76 appears to play a structural role in FlK acylation, we turned our attention to exploring the contribution of Thr 42 to Glu 50 acylation. Indeed, steady-state kinetics studies have shown that the FlK-T42S mutant retains only ~200-fold specificity for fluoroacetyl-CoA over acetyl-CoA, an effect that is partially attributable to a 30-fold decrease in the K_M for acetyl-CoA [*5*]. Using pre-steady state measurements, we determined that the acylation rate constant (k_2) for the T42S mutant for acetyl-CoA was within error of the rate constant measured for wild-type FlK while the deacylation rate constant (k_3) was reduced by 6-fold (*Figure 5.3A*). Kinetic modeling suggests that the observed decrease in K_M for acetyl-CoA in FlK-T42S can be explained by this defect in the deacylation rate in the absence of any changes in substrate affinity (*Figure 5.4*). For fluoroacetyl-CoA, the T42S mutation leads to a 3-fold decrease in acylation rate (k_2) and a 25-fold decrease in deacylation rate (k_3) (*Figure 5.3B*). From the different behavior of the two substrates with respect to acylation, it seems as if Thr 42 may make a small contribution to fluoroacetyl-CoA acylation specificity. Crystallographic studies have shown that Thr 42 is important in orienting Glu 50 and His 76 for reactivity [*11*]. In the crystal structure of the T42S mutant, the Glu 50 side chain is rotated by 90°, both Ser 42 and His 76 populate two rotamers, and the hydrogen bonding network among the catalytic residues is disrupted (*Figure 5.5*) [*11*].The observed acylation defect is therefore likely to be related to structural effects that result in misorientation of the catalytic residues. Furthermore, the lack of acylation defect for acetyl-CoA in both the H76A and T42S mutants suggests that the catalytic site is not optimized for utilization of this substrate.

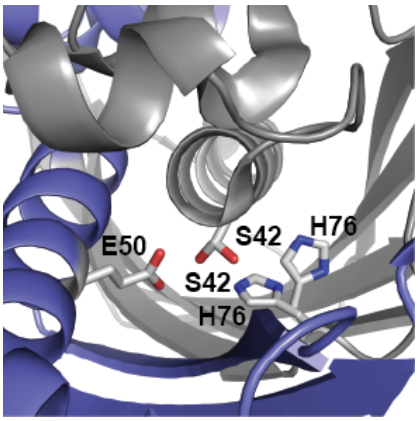


Figure 5.5. *The active site of FlK-T42S.In the T42S mutant, both Ser 42 and His 76 populate two different rotamers and Glu 50 is rotated relative to its position in the wild-type enzyme.*

Determining the impact of substrate polarization on the acylation rate. Based on our studies, the catalytic residues, His 76 and Thr 42, appear to play a limited role in controlling the specificity of Glu 50 acylation. We therefore next set out to probe the participation of the substrate itself in determining acylation specificity. Based on the electron-withdrawing nature of the fluorine atom, the fluoroacetyl-CoA carbonyl group is expected to be more activated towards

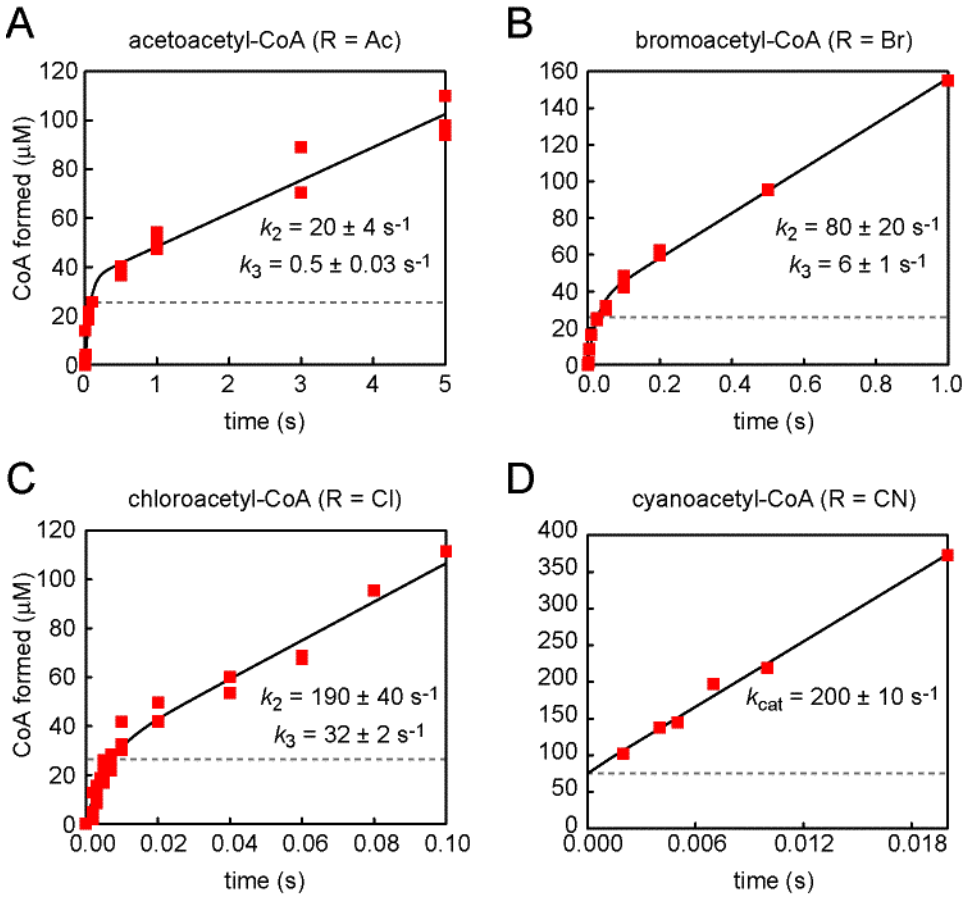


Figure 5.6. *Pre-steady-state kinetic analysis of a series of acyl-CoAs. Dotted line indicates one enzyme equivalent. (A) Acetoacetyl-CoA (R= Ac), (B) Bromoacetyl-CoA (R = Br), (C) Chloroacetyl-CoA (R = Cl), (D) Cyanoacetyl-CoA (R = CN).*

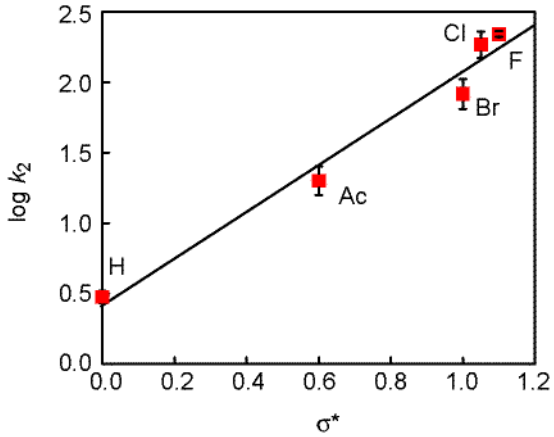


Figure 5.7. *Taft plot for the acylation step of FIK-catalyzed acyl-CoA hydrolysis. Acylation rate constants are derived from non-linear curve fitting of pre-steady-state kinetic time courses. The linear fit gave a ρ^* value of 1.7 ± 0.2 (R^2 = 0.986).*

nucleophilic attack compared to acetyl-CoA. We had previously compared the the rates of chemical hydrolysis of these two substrates and showed that the rate of fluoroacetyl-CoA hydrolysis is at least 10-fold higher than the rate for acetyl-CoA [*6*]. This ratio increases with increasing pH up to 100-fold, which is similar to the observed magnitude of discriminationbetween fluoroacetyl-CoA and acetyl-CoA in the acylation step. To probe the role of carbonyl polarization in acylation specificity, we examined the pre-steady state kinetics of FlK-catalyzed hydrolysis of a series of acyl-CoAs with different functional groups at the α-position in order to access a wide range of values of the Taft polar substituent constant (σ*) (*Figure 5.6*) [*10*]. By determining the relationship between σ* and the rate of acylation, we can define the impact of carbonyl polarization on substrate specificity for this particular step.

The rate of acylation for each substrate was measured using pre-steady state kinetic experiments (*Figure 5.6*). For acetyl-CoA, acetoacetyl-CoA, bromoacetyl-CoA, and chloroacetyl-CoA, a burst phase of product formation corresponding to one enzyme equivalent was observed, followed by a slower steady-state rate identical to the previously measured steady-state rate, indicating that hydrolysis of these substrates is likely to proceed through the same acyl-enzyme intermediate that was trapped for acetyl-CoA (*Scheme 5.1*). For FlK-catalyzed cyanoacetyl-CoA hydrolysis, plots of CoA formed versus time could not be fit with an equation describing burst-phase kinetic behavior and were instead best fit by a simple linear fit (*Figure 5.6D)*. However, the dead time of the rapid quench instrument is 2 ms, precluding the measurement of rate constants approaching or exceeding 500 s^{-1}. In order to distinguish between the absence of a burst phase and the presence of a burst phase that is over before the first quenched time point, we examined the amount of CoA formed after 2 ms. In the absence of a burst phase, the CoA concentration would be given by $[CoA] = k_{cat}[FlK]t$, corresponding to a line with a *y*-intercept of zero. On the other hand, if a burst phase had occurred before the first quenched time point, the CoA concentration would be given by $[CoA] = k_{cat}[FlK]t + [FlK]$, corresponding to a line with a *y*-intercept of [FlK]. Linear fitting of the data gave a *y*-intercept of 77 ± 7 μM, within error of the FlK concentration of 75 μM used in the assay (*Figure 5.6D*). We therefore conclude that cyanoacetyl-CoA is also hydrolyzed through a mechanism involving an acyl-enzyme intermediate, although we could not include this substrate in our free-energy relationship analysis because the acylation rate could not be accurately measured.

A plot of log k_2 versus σ* for the series of acyl-CoAs revealed a linear relationship (*Figure 5.7*), indicating that acylation of all of the substrates proceeds through a common mechanism and transition state [*10*]. Based on this linear relationship, specificity in FlK acylation appears to be based on carbonyl polarization rather than a change in mechanism, rate-limiting step, or transition state structure induced by the fluorine substituent. The slope of the line, ρ*, is the polar sensitivity factor that compares the reaction under study to the reference reaction, which is methyl ester hydrolysis [*10*]. A slope >1 indicates that the reaction under study is more sensitive to substituents than the reference reaction, while a slope <1 indicates that it is less sensitive. The Taft plot for acylation gave a ρ* value of 1.7 ± 0.2, indicating that FlK acylation is more sensitive to the presence of substituents that alter carbonyl polarization than chemical hydrolysis of methyl esters on which the Taft scale is based. This sensitivity may be attributable to the higher polarization of thioesters compared to esters, or to additional activation of the carbonyl group on the enzyme. In combination with the mutagenesis studies discussed above, our Taft analysis of FlK substrates suggests that specificity in the rate of enzyme acylation is determined mainly by the intrinsic polarization of the fluoroacetyl-CoA thioester compared to the acetyl-

CoA thioester rather than by features of the enzyme that specifically accelerate reaction with the fluorinated substrate.

Investigating the role of fluorine molecular recognition in FlK-catalyzed thioester hydrolysis. Although our studies suggest that the main selectivity determinant for FlK acylation is the intrinsic reactivity of the fluoroacetyl group, this finding does not exclude the possibility that the C-F bond is recognized in the enzyme active site. If so, interactions between FlK and the fluoroacetyl-CoA substrate could reduce free rotation of the fluoromethyl group such that each of the two prochiral α-protons would occupy a different position in three-dimensional space. In this case, we might expect that FlK could exhibit a kinetic preference for substrates with either an available pro-*R* or pro-*S* proton.. To test this hypothesis, we prepared (*RS*)-2-fluoropropionyl-CoA, a substrate in which one hydrogen at the α-carbon of a fluoroacetyl-CoA has been replaced by a methyl group. Upon incubation of this substrate with FlK, we observed two steady-state kinetic phases, each representing approximately 50% of the total substrate (*Figure 5.8A*). By separately calculating the rates from the fast and slow phases, we were able to construct a crude Michaelis-Menten curve for each phase (*Figure 5.8B*). Since both phases exhibit saturation behavior, we hypothesized that these two phases might correspond to different rates of hydrolysis for the different stereoisomers.

To test this idea, we prepared the individual (*S*)- and (*R*)-2-fluoropropionyl-CoA stereoisomers from the corresponding chiral methyl lactate and measured the steady-state kinetic parameters for each stereoisomer (*Figure 5.9*). Indeed, we observed that each stereoisomer was hydrolyzed at a different rate, with the k_{cat}s for the (*S*) and (*R*) isomers roughly corresponding to the observed rate constant for the fast and slow phase, respectively. Interestingly when these substrates are added to the Taft plot for log k_{cat} versus σ*, the (*S*) isomer falls more on the same line as fluoroacetyl-CoA and chloroacetyl-CoA while the (*R*) isomer appears to behave more similarly with the less polarized substrates (*Figure 5.10*). This result can be interpreted to indicate that there is a change in mechanism or rate-limiting step between the (*S*) and (*R*) isomers, and that hydrolysis of the (*S*) isomer proceeds through the same kinetic and chemical mechanism as fluoroacetyl-CoA, while hydrolysis of the (*R*) isomer proceeds through the same kinetic and chemical mechanism as acetyl-CoA. Therefore, fluorine-based polarization of the substrate alone is insufficient to dictate the reaction pathway through which thioester hydrolysis proceeds.

In addition to the difference in k_{cat}, the two stereoisomers diverge with regard to their K_M values, with the (*R*) isomer exhibiting a lower K_M than that measured for the (*S*) isomer. This observation could be explained by either preferential binding of the (*R*) isomer, faster acylation by the (*R*) isomer, or slower deacylation of the (*R*) isomer. To determine the basis of the observed K_M difference, we examined the pre-steady state kinetic behavior of each isomer under saturating conditions. For (*S*)-2-fluoropropionyl-CoA, we observed single-phase kinetics with a rate constant equal to k_{cat} for the (*S*) isomer, which indicates that acylation is rate liming for this substrate as it is for the native substrate (*Figure 5.11A*). In contrast, we observed a burst of one equivalent of CoA formed with (*R*)-2-fluoropropionyl-CoA, followed by a slower steady-state rate with a rate constant matching both the k_{cat} of the (*R*) isomer (*Figure 5.11B*). We interpret this kinetic behavior to demonstrate that rate-liming step for the (*R*) isomer has changed to breakdown of the acyl-enzyme intermediate ($k_3 = k_{cat}$). Taken together, the results of the pre-steady state kinetic experiments are also consistent with the behavior of isomer with regard to the

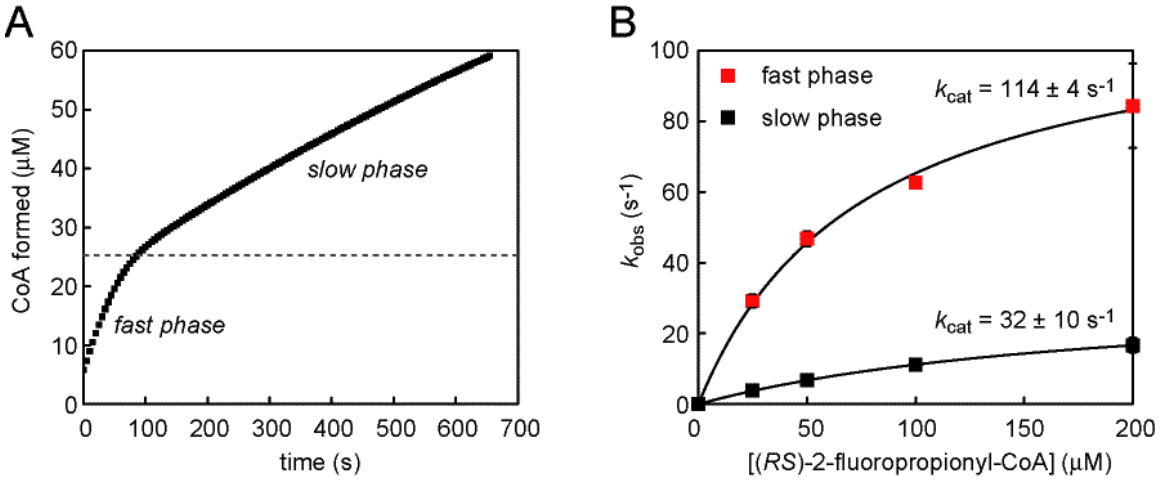


Figure 5.8. *Steady-state kinetic analysis of (RS)-2-fluoropropionyl-CoA. (A) Time course for hydrolysis of 50 mM (RS)-2-fluoropropionyl-CoA. Dotted line indicates 50% of the substrate. (B) Michaelis-Menten curves produced by calculating the observed rate constants (k_{obs}) for the slow phase and fast phase over a range of substrate concentrations.*

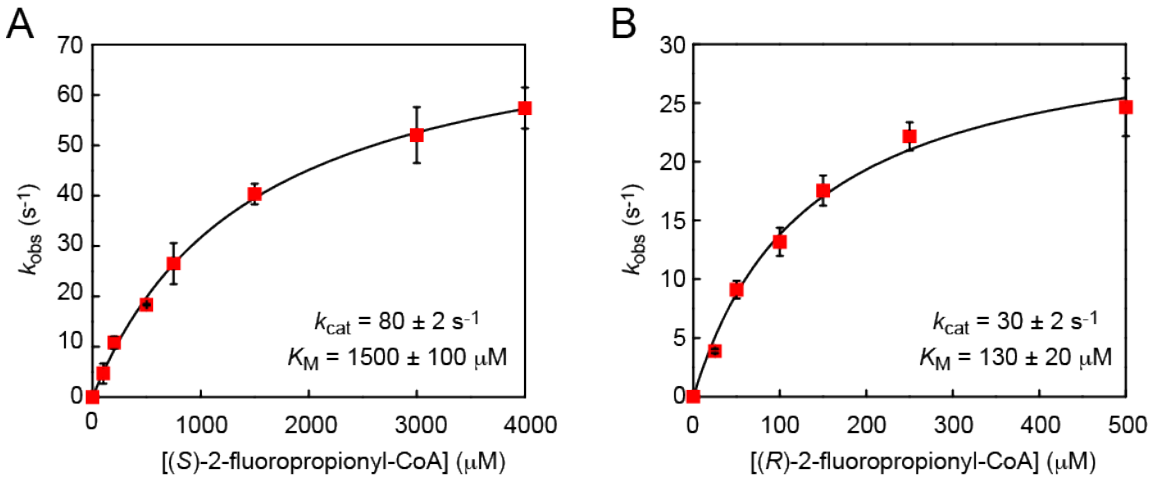


Figure 5.9. *Steady-state kinetic analysis of the individual (S)- and (R) stereoisomers of 2-fluoropropionyl-CoA. (A) (S)-2-fluoropropionyl-CoA. (B) (R)-2-fluoropropionyl-CoA.*

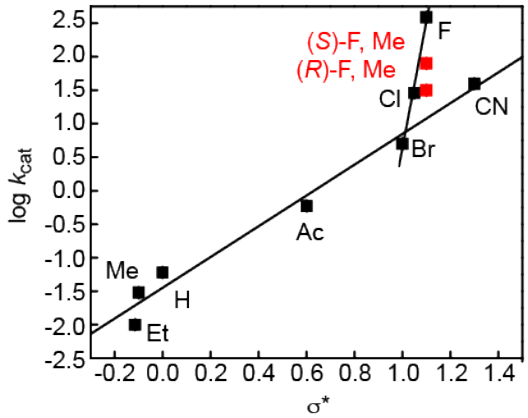


Figure 5.10. *Taft plot for k_{cat} of various acyl-CoA FIK substrates. 2-fluoropropionyl-CoA substrates are shown in red. Et, butyryl-CoA; Me, propionyl-CoA; H, acetyl-CoA; Ac, acetoacetyl-CoA; Br, bromoacetyl-CoA; Cl, chloroacetyl-CoA; F, fluoroacetyl-CoA; (S)-F, Me, (S)-2-fluoropropionyl-CoA; (R)-F, Me, (R)-2-fluoropropionyl-CoA; CN, cyanoacetyl-CoA.*

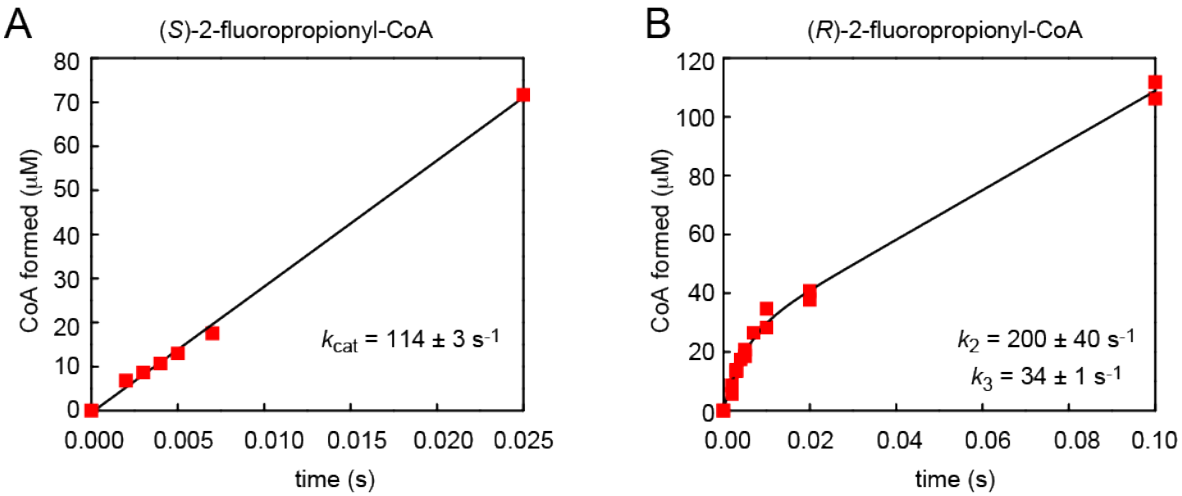


Figure 5.11. *Pre-steady-state kinetic analysis of FIK-catalyzed hydrolysis of (S)- and (R)-2-fluoropropionyl-CoA. (A) (S)-2-fluoropropionyl-CoA. (B) (R)-2-fluoropropionyl-CoA.*

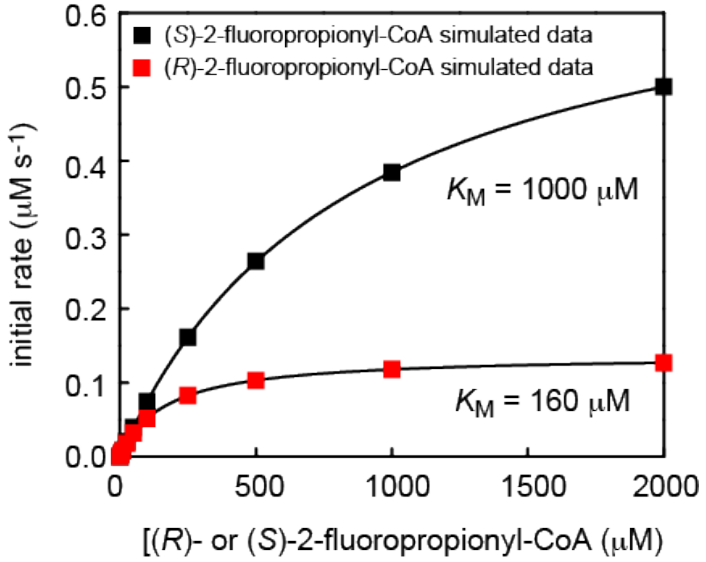


Figure 5.12. *Simulated steady-state kinetic data for FIK-catalyzed hydrolysis of (S)- and (R)-2-fluoropropionyl-CoA. At constant K_D, the changes in measured kinetic constants are sufficient to explain the difference in K_M between the two substrates.*

linear free-energy plot (*Figure 5.10*). Additional kinetic modeling using these rate constants shows that the differences in the rates of chemical catalysis between the two substrates are sufficient to explain their differing steady-state kinetics without the need to invoke differential binding affinities (*Figure 5.12*).

Because the only significant difference between the (*S*) and (*R*) isomers is expected to be the orientation of the α-substituents, the differences in the rate constants of the acylation (2-fold) and deacylation (4-fold) steps between the two stereoisomers suggest that the fluorine atom is specifically recognized by the enzyme and that the α-carbon does not freely rotate. Indeed, the difference observed in the deacylation rate for each isomer can be explained by a requirement for the α-proton to be accessible to His 76 in order for the substrate to access the faster C_α-deprotonation pathway (*Scheme 5.2*) [*6*]. Orientation of the substrate via the fluorine atom could result in the methyl group occupying the space where an α-proton would normally be poised for abstraction by His 76, which could lead to a change in mechanism from C_α-deprotonation to addition of water followed by elimination of the Glu 50 leaving group. Based on the kinetic data, we hypothesize the His 76 preferentially abstracts the *pro*-R proton, leading to a compromised deacylation rate when the methyl group occupies this position instead. However, the basis for the difference between the acylation rates is more difficult to explain. Based on an addition-elimination mechanism for enzyme acylation, the observed difference in the acylation rates is likely due to the increased steric bulk of the additional methyl substituent, which kinetic data suggest is positioned unfavorably for acylation in the (*S*) isomer. One possible source of this kinetic selectivity is that recognition of the fluorine substituent by FlK could force nucleophilic attack by Glu 50 to occur along a trajectory that is disfavored based on the Felkin-Anh model for carbonyl addition (*Figure 5.13*) [*13*, *14*].

Stereochemical course of the FlK reaction. To directly test the stereochemical course of acylation and deacylation without the need for the enzyme to accommodate an additional α-substituent in its active site, we prepared $^2H_1$-(*S*)-fluoroacetyl-CoA and $^2H_1$-(*R*)-fluoroacetyl-CoA from the corresponding carboxylic acids [*9*] Our previous kinetic studies with $^2H_2$-fluoroacetyl-CoA showed that there is a kinetic isotope effect (KIE) of 2.4 ± 0.1 on the deacylation step of the reaction mechanism, consistent with abstraction of an α-proton from the substrate in this step (*Scheme 5.2*) [*6*]. We therefore predict that if the enzyme utilizes chiral recognition of the prochiral fluoromethyl group in the deacylation step, one of the $^2H_1$-fluoroacetyl-CoA stereoisomers will have its ^2H positioned for proton abstraction and will exhibit a primary KIE, while the other stereoisomer will show a much smaller secondary KIE, which may be negligible compared to the error in the rates. In contrast, based on the addition-elimination mechanism that we have proposed for the acylation step, we would expect both substrates to exhibit a secondary inverse KIE on acylation, which may be too small to detect by direct comparison of rates.

To interrogate both the acylation and deacylation steps using the chiral deuterated substrates, we examined the pre-steady state kinetics of their FlK-catalyzed hydrolysis (*Figure 5.14*). Similar to the unlabeled fluoroacetyl-CoA, $^2H_1$-(*S*)-fluoroacetyl-CoA exhibited single-phase pre-steady state kinetic behavior with the observed k_{cat} within error of that for the undeuterated substrate. While this result does not conclusively demonstrate the presence of an inverse secondary KIE, it not inconsistent considering the propagated error. In contrast, $^2H_1$-(*R*)-fluoroacetyl-CoA exhibited burst-phase kinetic behavior. While the acylation rate constant for this substrate was within error of the k_{cat} observed for the undeuterated substrate, a kinetic

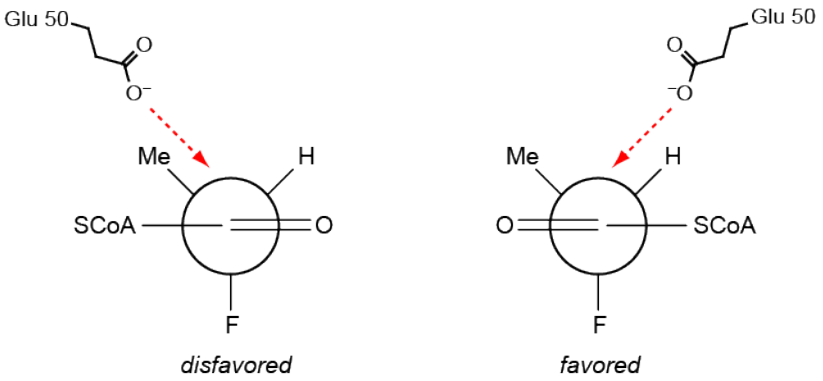


Figure 5.13. *Analysis of FlK acylation by 2-fluoropropionyl-CoA in terms of the Felkin-Ahn model. The steric and electronic preferences of the substrate in combination with the structural constraints of the enzyme may explain the defect in acylation rate for (S)-2-fluoropropionyl-CoA compared to (R)-2-fluoropropionyl-CoA.*

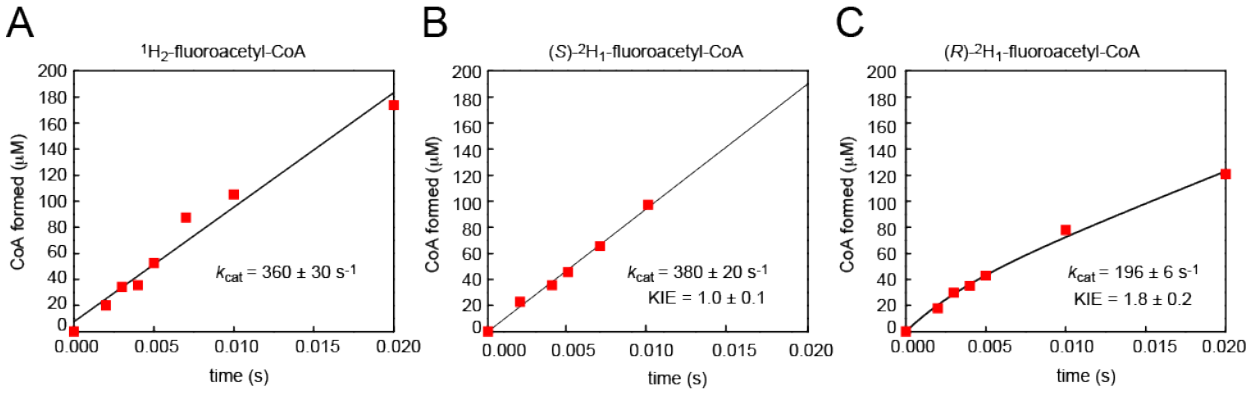
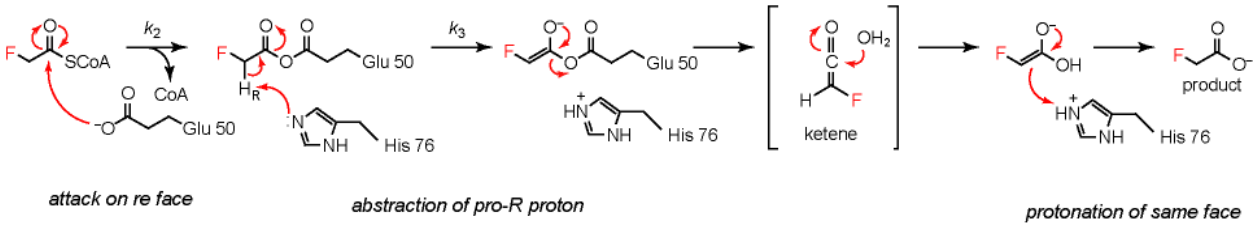


Figure 5.14. *Pre-steady-state kinetic analysis of monodeuterated fluoroacetyl-CoAs. FlK-catalyzed hydrolysis of (A) undeuterated fluoroacetyl-CoA (left) is compared to hydrolysis of (B) (S)-$^2H_1$-fluoroacetyl-CoA (center) and of (C) (R)-$^2H_1$-fluoroacetyl-CoA (right).*

isotope effect (KIE) of 1.8 ± 0.2 was observed on the steady-state rate, consistent with the previously reported KIE of 2.4 observed with the doubly deuterated $^{2}H_2$-fluoroacetyl-CoA substrate. This result provides further evidence that while the acylation rate is influenced mainly by polarization, consistent with an addition-elimination-type mechanism, deacylation requires chiral recognition of the prochiral fluoromethyl group to position the α-carbon for proton abstraction catalyzed by His 76, which preferentially abstracts the *pro*-R proton. Notably, the mechanistically-related hot dog-fold dehydratases, which utilize a mechanism initiated by C_α-deprotonation to catalyze elimination of water from 3-hydroxyacyl-CoAs, also carry out stereoselective deprotonation [*15*, *16*]. While the *pro*-2*S* proton is removed by dehydratases, it has the same relative stereochemistry as the *pro*-2*R*-proton in fluoroacetyl-CoA since the *S* designation is derived from a change in substituent priorities based on the Cahn-Ingold-Prelog rules. While the proton abstraction occurs on an acyl-enzyme intermediate in FlK instead of on the substrate itself, the commonality in mechanism between the hot dog-fold dehydratases and FlK is consistent with their surprisingly close evolutionary relationship [*6*], as well as the structural similarity of their active sites [*5*, *17*].

Proposed chemical, kinetic, and structural mechanism accounting for FlK's fluorine specificity. Our data suggest a chemical, kinetic, and structural mechanism for the selectivity of FlK for fluoroacetyl-CoA (*Scheme 5.3*). Following substrate binding, fluoroacetyl-CoA reacts with Glu 50 to acylate this side chain with an enhancement of two orders of magnitude on rate of acylation compared to acetyl-CoA provided by the instrinsic polarization of the fluoroacetyl carbonyl group. Based on the differences in acylation rate between the two 2-fluoropropionyl-CoA stereoisomers, we propose that the C-F bond is specifically recognized by FlK, allowing it to distinguish the two enantiotopic faces of the carbonyl. Following acylation, a proton is abstracted by His 76 to generate an enolate that can break down through a proposed ketene intermediate. Based on the differential rates of deacylation of the two 2-fluoropropionyl-CoA stereoisomers, as well as the observation of a primary KIE on only the (*R*)-$^{2}H_1$-fluoroacetyl-CoA stereoisomer, it appears that FlK also utilizes prochiral recognition of the prochiral fluoromethyl group in the deacylation step to specifically position the *pro*-R proton for abstraction. The resultant enolate can then breakdown to the carboxylic acid product through a proposed ketene intermediate. Based on the previously reported lack of exchange between the fluoroacetyl-CoA α-protons and solvent [*6*] and the pH-rate profiles of FlK and FlK-H76A, we propose that FlK utilizes a one-base mechanism to achieve this transformation, rather than separate catalytic residues that function in general acid catalysis and general base catalysis. Therefore, following hydration of the putative ketene intermediate, we propose that the resultant enolate is



Scheme 5.3. *Model for the influence of fluorine recognition of the course of the FlK-catalyzed reaction. Based on kinetic analysis of substrates with a stereogenic center at the α-carbon, we propose that Glu 50 attacks the fluoroacetyl-CoA thioester bond from the re face. The pro-R proton of the resultant intermediate is then removed by His 76. Following hydration of the putative intermediate, the carboxylic acid enolate is reprotonated on the same face from which the proton was removed.*

reprotonated by the His 76 imidazolium side chain on the same face from which deprotonation occured. This mechanistic proposal is similar to the one-base mechanism elucidated for the hot dog-fold dehydratases [*15*, *16*].

Taken together, our results suggest that molecular recognition of fluorine, rather than simple polarization and enolate stability, is involved in making the C_α-deprotonation pathway accessible to the fluorinated substrate. Polarization therefore appears to be necessary, but not sufficient, for rate acceleration. This provides a possible explanation for the inability of substrates like acetoacetyl-CoA, which is expected to form a more stable enolate than fluoroacetyl-CoA but is much larger in size, to access the C_α-deprotonation pathway.

5.4 Conclusion

Although fluorine's ability to initiate unusual enzymatic reaction pathways involved in mechanism-based inhibition is well known, FlK represents a rare example in which an enzyme exploits the unique properties of fluorine to enhance its reaction rate and to drive substrate specificity. The results reported here demonstrate that discrimination of the fluoroacetyl-CoA substrate relies not only on the enhanced polarization afforded by fluorine's high electronegativity, but also on molecular recognition of fluorine to position the substrate for optimum reactivity. These studies of FlK provide insights into how enzymes evolve to recognize unusual functional groups, with implications for fluorine molecular recognition and the design of mechanism-based inhibitors.

5.5 Acknowledgements

We thank Dr. Mark Walker and Benjamin Thuronyi for purification of acetate kinase and phosphotransacetylase and for advice on enzymatic synthesis of fluoroacetyl-CoA.

5.6 References

1. D. D. Clarke, Fluoroacetate and fluorocitrate: Mechanism of action, *Neurochem. Res.*, **1991**, *16*, 1055–1058.

2. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme-inhibitor complex, *Proc. Natl. Acad. Sci. U. S. A.*, **1996**, *93*, 13699–13703.

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

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

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

6. A. M. Weeks and M. C. Y. Chang, Catalytic control of enzymatic fluorine specificity, *Proc. Natl. Acad. Sci. U. S. A.*, **2012**, *109*, 19667–19672.

7. L. Hedstrom, Serine protease mechanism and specificity, *Chem. Rev.*, **2002**, *102*, 4501–4524.

8. E. Fritz-Langhals and G. Schutz, Simple synthesis of optically active 2-fluoropropanoic acid and analogs of high enantiomeric purity, *Tetrahedron Lett.*, **1993**, *34*, 293–296.

9. R. D. P. Wadoux, X. Lin, N. S. Keddie and D. O'Hagan, Chiral fluoroacetic acid: Synthesis of (*R*)- and (*S*)-[$^2H_1$]-fluoroacetate in high enantiopurity, *Tetrahedron Asymmetry*, **2013**, *24*, 719–723.

10. R.W. Taft, Polar and steric substituent constants for aliphatic and *o*-benzoate groups from rates of esterification and hydrolysis of esters, *J. Am. Chem. Soc.*, **1952**, *74*, 3120–3128.

11. M. V. B. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. V. Dry, P. F. Leadlay, J. B. Spencer and T. L. Blundell, Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK, *J. Biol. Chem.*, **2010**, *285*, 22495–22504.

12. F. Song, Z. Zhuang, L. Finci, D. Dunaway-Mariano, R. Kniewel, J. A. Buglino, V. Solorzano, J. Wu and C. D. Lima, Structure, function, and mechanism of the phenylacetate pathway hot dog-fold thioesterase PaaI, *J. Biol. Chem.*, **2006**, *281*, 11028–11038.

13. N. T. Anh, O. Eisenstein, J. M. Lefour and M. E. Tran Huu Dau, Orbital factors and asymmetric induction, *J. Am. Chem. Soc.*, **1973**, *95*, 6146–6147.

14. H. B. Burgi, J. D. Dunitz, J. M. Lehn and G. Wipff, Stereochemistry of reaction paths at carbonyl centres, *Tetrahedron*, **197.4**, *30*, 1563–1572.

15. J. M. Schwab, A. Habib and J. B. Klassen, A thorough study of the stereochemical consequences of the hydration/dehydration reaction catalyzed by β-hydroxydecanoyl thioester dehydrase, *J. Am. Chem. Soc.*, **1986**, *108*, 5304–5308.

16. J. W. Labonte and C. A. Townsend, Active site comparisons and catalytic mechanisms of the hot dog superfamily, *Chem. Rev.*, **2013**, *113*, 2182–2204.

17. T. Hisano, T. Tsuge, T. Fukui, T. Iwata, K. Miki and Y. Doi, Crystal structure of the (*R*)-specific enoyl-CoA hydratase from *Aeromonas caviae* involved in polyhydroxyalkanoate biosynthesis, *J. Biol. Chem.*, **2003**, *278*, 617–624.

Chapter 6: *Exploring organofluorine resistance, biosynthesis, and regulation in Streptomyces cattleya*

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

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

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

Dr. Mark C. Walker and Dr. Miao Wen

6.1 Introduction

Although fluoroacetate has been identified in more than 40 species of plants [*1*], *Streptomyces cattleya* represents one of only two bacterial species that possess the ability to biosynthesize fluorinated natural products [*2, 3*]. When grown in the presence of fluoride, *S. cattleya* produces fluoroacetate and 4-fluorothreonine [*3*], both potentially useful building blocks for incorporation of fluorine into natural products (*Figure 6.1*). Elucidating the enzymatic steps required for organofluorine biosynthesis, resistance, and regulation would not only provide insights into the evolution of an exotic chemical phenotype, but would also provide information about the minimal set of genes and proteins required to harness this phenotype for chemical synthesis.

In the absence of a sequenced *S. cattleya* genome, initial efforts toward identifying enzymes involved in biological fluorination focused on feeding of precursors and fractionation of biochemical activities from *S. cattleya* lysate. Addition of fluoride and *S*-adenosyl-L-methionine (SAM) to *S. cattleya* cell-free extract established 5'-fluoro-5'-deoxyadenosine (5'-FDA) as the initial product of fluorination, while feeding 5'-FDA to cell-free extract resulted in fluoroacetate production [*4*], establishing the chemical competence of this intermediate (*Figure 6.1*). The enzymatic activity responsible for catalyzing formation of 5'-FDA was subsequently purified from *S. cattleya* [*5*], providing a partial protein sequence that allowed for cloning of the fluorinase gene (*flA*) [*6*]. Extensive biochemical [*5, 7-10*], structural [*6*], and mechanistic [*6, 7*] studies showed that the fluorinase catalyzes an S_N2-type reaction between fluoride and SAM by desolvating fluoride to enhance its nucleophilicity.

Following identification of 5'-FDA as the initial fluorination product, 5-fluoro-5-deoxyribose-1-phosphate (5-FDRP) was identified as the second intermediate in the pathway (*Figure 6.1*) [*11*]. Incubation of this intermediate with *S. cattleya* extract led to accumulation of

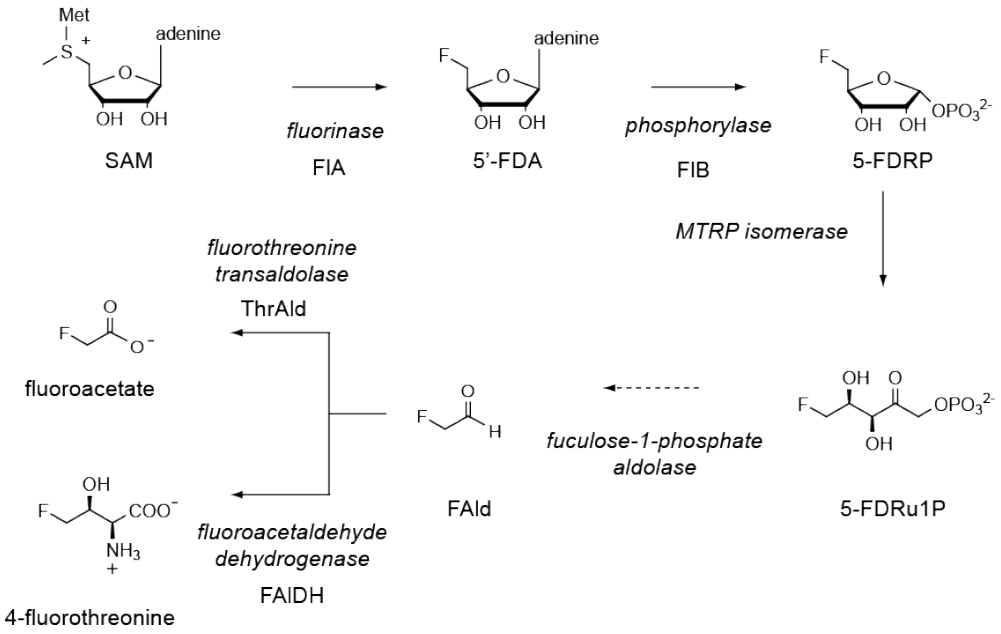


***Figure 6.1**. Proposed biosynthetic pathway for fluoroacetate and 4-fluorothreonine in Streptomyces cattleya. SAM, S-adenosyl-L-methionine; 5-FDRP, 5-fluoro-5-deoxyribose-1-phosphate; 5-FDRu1P, 5-fluoro-5-deoxyribulose-1-phosphate; FAld, fluoroacetaldehyde. Dotted line indicates a proposed activity that has not been purified or cloned.*

both fluoroacetate and 4-fluorothreonine [*11*]. A purine nucleoside phosphorylase activity capable of converting 5'-FDA to 5-FDRP was subsequently purified from *S. cattleya* extract [*11*]. Based on similarities between the first two steps of fluorometabolite pathway and the methionine salvage pathway (*Figure 6.1*) [*12*], an analogy was drawn that suggested that the next step might be isomerization to form 5-fluoro-5-deoxy-D-ribulose-1-phosphate (5-FDRu1P) as the next intermediate in the pathway [*13*]. Feeding studies using 5-FDRP in the presence of EDTA (ethylenediaminetetraacetic acid), which is expected to inhibit aldolase enzymes by sequestering the required Zn^{2+} ion, revealed the accumulation of a new product, which was assigned as 5-FDRu1P [*13*]. An isomerase activity capable of catalyzing formation 5-FDRu1P was subsequently purified from *S. cattleya* [*14*].

Continuing with the methionine salvage pathway analogy [*12*], the next enzyme involved in fluorometabolite biosynthesis is proposed to be a fuculose-1-phosphate aldolase that would convert 5-FDRu1P to dihydroxyacetone phophate and fluoroacetaldehyde (*Figure 6.1*), which has been proposed as the last common intermediate between fluoroacetate and 4-fluorothreonine based on isotopic labeling studies [*15*]. Purification of *S. cattleya* cell lysate has identified an NAD^+-dependent fluoroacetaldehyde dehydrogenase (FAlDH) activity that catalyzes formation of fluoroacetate [*16*], as well as an unusual PLP-dependent transaldolase (ThrAld) to convert fluoroacetaldehye and threonine to acetaldehyde and 4-fluorothreonine [*17*].

The availability of the fluorinase gene sequence made it possible to target the *flA* genomic locus for sequencing and led to the identification of a gene cluster involved in fluorometabolite biosynthesis, termed the *fl* locus (*Figure 6.2*) [*18*]. While the fluorinase was found to be adjacent to a gene encoding a purine nucleoside phosphorylase capable of converting 5'-FDA to 5-FDRP, none of the other genes in the cluster have obvious biosynthetic functions. The remaining genes can be divided into classes based on their putative functions in resistance (*flD* and *flK*), transmembrane transport (*flC* and *flH*), and transcriptional (*flE*, *flF*, *flG*, and *flL*) and other types (*flI* and *flJ*) of regulation [*18*]. Of these genes, only *flB*, *flK* and *flI* have been expressed and characterized *in vitro*. FlK has fluoroacetyl-CoA thioesterase activity [*18-20*], while FlI is an *S*-adenosylhomocysteine (SAH) hydrolase proposed to function in the relief of fluorinase inhibition by endogenous SAH [*18*].

In this study, we set out to characterize the functions of the *fl* locus genes in fluorometabolite biosynthesis, resistance, and regulation. This chapter reports the construction and preliminary characterization of *S. cattleya* strains in which the *fl* locus genes *flA, flB*, *flD*, and *flK* are knocked out. Progress toward construction of knockout strains for the genes proposed to be involved in transmembrane transport and regulation of fluorometabolite biosynthesis, as well as FAlDH and ThrAld is also summarized. Finally, experiments to test whether the 5'-untranslated region (5'-UTR) of the *flL* gene represents a fluoride riboswitch involved in posttranscriptional regulation of fluorometabolite biosynthesis are also described.

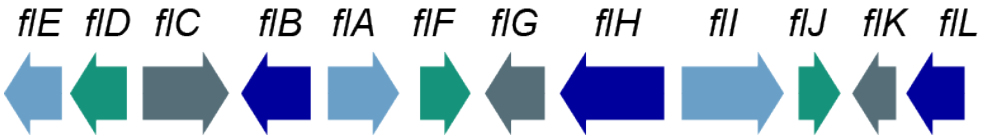


***Figure 6.2**. Organization of the fl locus in Streptomyces cattleya.*

6.2 Materials and Methods

Commercial Materials. Luria-Bertani (LB) Broth Miller and LB Agar Miller, yeast extract, malt extract, and nutrient broth were purchased from EMD Biosciences (Darmstadt, Germany). Agar and casaminoacids were purchased from Becton, Dickinson, and Company (Sparks, MD). Soy flour was manufactured by Arrowhead Mills (Boulder, CO). Carbenicillin (Cb), tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), sodium chloride, phenol:chloroform:isoamyl alcohol (25:24:1), magnesium chloride heptahydrate, kanamycin, ethylenediaminetetraacetic acid disodium dehydrate (EDTA), dimethyl sulfoxide (DMSO), ammonium sulfate, dibasic potassium phosphate, monobasic sodium phosphate, manganese chloride tetrahydrate, calcium chloride hexahydrate, and dextrose were purchased from Fisher Scientific (Pittsburgh, PA). Sodium fluoroacetate, sodium fluoride, 5-fluorouracil, apramycin sulfate, betaine, magnesium sulfate heptahydrate, iron sulfate heptahydrate, *N*-[Tris(hydroxymethyl)methyl]-2-aminoethanesulfonic acid (TES), PEG 6000, and Antifoam 204 were purchased from Sigma-Aldrich (St. Louis, MO). Deuterium oxide was purchased from Cambdrige Isotope Laboratories (Andover, MA). Deoxynucleotide triphosphates (dNTPs) and Platinum Taq HF were purchased from Invitrogen (Carlsbad, CA). Oligonucleotides were purchase from Integrated DNA Technologies (Coralville, IA). DNA purification kits were purchased from Qiagen (Valencia, CA). DNA sequencing was performed by Quintara Biosciences (Berkeley, CA).

Bacterial strains and yeast strains. *Streptomyces cattleya* NRRL 8057 (ATCC 35852) was purchased from the American Tissue Type Collection (Manassas, VA). *Saccharomyces cerevisiae* L4581 was a gift from the Madhani laboratory (University of California, San Francisco) and was used for plasmid assembly by *in vivo* homologous recombination. *E. coli* DH10B-T1^R was used for production of plasmid DNA. *E. coli* GM272 harboring the non-transmissible, oriT-mobilizing plasmid pUZ8002 was used for conjugative plasmid transfer into *S. cattleya*.

Construction of plasmids for gene disruption in *S. cattleya*. Plasmids for gene disruption contained a cassette consisting of an apramycin marker (Am) and an origin of transfer (oriT) flanked by *S. cattleya* genomic DNA sequence 2000 bp upstream and downstream of the open reading frame to be disrupted. Flanking DNA sequences immediately upstream (UF1 and UR1 primers) and downstream (DF1 and DR1 primers) of the open reading frame to be disrupted were amplified using either Platinum Taq with the addition of 5 M betaine or Taq with the addition of 10% DMSO. The Am^R/oriT cassette was amplified from pIJ773 (KF1 and KR1 primers). Primers for all amplification steps contained extensions at the 5′ and 3′ ends such that all PCR products to be assembled contained 50 nt of sequence identity with the preceding and following pieces. *Saccharomyces cerevisiae* L4581 was transformed with the three PCR products and Kpn I-linearized pRS316 to effect assembly by *in vivo* homologous recombination. Tranformants containing the circularized plasmid with the desired insert were selected on solid synthetic complete dextrose (SCD) media lacking uracil. DNA was recovered by phenol/chloroform extraction and dissolved in TE buffer (50 μL; 10 mM Tris-HCl, pH 7.5, 1 mM EDTA). *E. coli* DH10B-T1^R were transformed with the recovered DNA and transformants were selected on LB agar supplemented with carbenicillin (50 μg/mL) and apramycin (50 μg/mL). Plasmids were verified by Sanger sequencing and were used to generate *S. cattleya* knockout strains.

Preparation of *S. cattleya* growth media. Mannitol soy flour (MS) plates contained mannitol (20 g/L), soy flour (20 g/L), and agar (20 g/L). Difco nutrient agar (DNA) plates contained nutrient broth (8 g/L) and agar (10 g/L). Minimal media (MM) plates contained $(NH_4)_2SO_4$ (1 g/L), K_2HPO_4 (0.5 g/L), $MgSO_4 \cdot 7H_2O$ (0.2 g/L), $FeSO_4 \cdot 7H_2O$ (0.01 g/L) and agar (10 g/L). Sodium acetate and glucose stock solutions were autoclaved separately and added immediately before plates were poured to a final concentration of either 2 g/L or 10 g/L. GYM media contained glucose (4 g/L), yeast extract (4 g/L), and malt extract (10 g/L) and was adjusted to pH 5 or pH 7 as indicated. To prepare supplemented liquid minimal medium (SMM), the following solutions were prepared and sterilize separately and then mixed in the order and volumes given: 6.1% (w/v) PEG 6000 (81.9 mL); 2.5 M $MgSO_4$ (2.5 mL); 0.25 M TES, pH 7.2 (10 mL); 50 mM NaH_2PO_4 + 50 mM K_2HPO_4 (1 mL); 50% (w/v) glucose (2 mL); 1% (w/v) Antifoam 204 (1 mL); trace element solution (0.1 mL); 20% (w/v) Difco casaminoacids (1 mL); 20% glycine (2.5 mL). Trace element solution contained 0.1 g/L each of $ZnSO_4 \cdot 7H_2O$, $FeSO_4 \cdot 7H_2O$, $MnCl_2 \cdot 4H_2O$, $CaCl_2 \cdot 6H_2O$, and NaCl. This solution was sterilized by 0.2 μm filtration and was stored at 4°C for up to 1 month.

Conjugative transfer of plasmids from *E. coli* to *S. cattleya*. Plasmids for gene disruption were transferred into *S. cattleya* by conjugation with *E. coli* GM272 harboring the oriT-mobilizing plasmid pUZ8002 using a modified literature protocol. *E. coli* GM272 pUZ8002 were transformed with plasmids for gene disruption and transformants were selected on LB agar supplemented with kanamyin (50 μg/mL) and carbenicillin (50 μg/mL). Single colonies were inoculated into LB supplemented with the appropriate antibiotics and were grown overnight at 37ºC shaking at 200 rpm. The overnight *E. coli* culture was inoculated in LB supplemented with the appropriate antibiotics to $OD_{600\ nm}$ = 0.1. The culture (10 mL) was grown to an $OD_{600\ nm}$ of 04.-0.6 and cells were pelleted by centrifugation, washed twice with sterile water (10 mL), and resuspended in GYM media media (0.5 mL). *S. cattleya* spores (10^8 in 50 μL) were prepared for conjugation by heat shock at 55ºC for 5 min followed by addition of GYM (0.5 mL) and incubation for 4-5 h at 30ºC shaking at 200 rpm. *E. coli* (0.5 mL) were added to spores (0.5 mL) and the mixture was incubated at 30ºC shaking at 200 rpm for 1 h. The mixture was plated on MS media. The plate was incubated at 30ºC for 36 h, and then overlaid with a mixture of nalidixic acid (40 μL, 25 mg/mL) and 2.5 $MgCl_2$ (600 μL). The plate was incubated for an additional 16 h, after which it was overlaid with apramycin (20 μL, 50 mg/mL). Exconjugants appeared within 5-7 days. Gene disruption was verified by PCR amplification of the targeted locus followed by sequencing.

Characterization of organofluorine production in *S. cattleya* strains. *S. cattleya* glycerol stocks were streaked onto MS plates. Single colonies appeared within 2 d and were used to inoculate GYM (pH 7, 50 mL). After 2 d of growth, cultures were used to inoculate GYM (pH 5, 500 mL) to an OD_{600} of 0.05. After 24 h, sodium fluoride was added to 2 mM final concentration. At various times between 1 and 12 d after sodium fluoride addition, aliquots (5 mL) were removed from the cultures for analysis. Cells were pelleted by centrifugation and the supernatant was lyophilized. The resultant solid was dissolved in 100 mM Tris-HCl, pH 7.5 containing 5-fluorouracil (1 mM) as an internal standard (500 μL). Deuterium oxide (100 μL) was added for use as the lock solvent. ^{19}F NMR spectra were collected on a Bruker AV600 (400 scans; o1p, -190 ppm; d1, 11 s).

Characterization of fluoride and fluoroacetate sensitivity in *S. cattleya* wild-type and *ΔflK*. Spores of *S. cattleya* wild-type and Δ*flK*::Am^R (40 μL) were heat-shocked at 55ºC for 5

min and inoculated into GYM (pH 7, 50 mL). Cultures were grown at 30ºC, 200 rpm to OD600 = 1.5-2 and then were diluted to OD_{600} = 0.1 in GYM (pH 5, 50 mL) with and without sodium fluoride (2 mM) or sodium fluoroacetate (2 mM). Glass beads (4 per flask, 5 mm) were added to promote dispersed growth. Growth was monitored by measuring OD_{600} of samples (1 mL) withdrawn from the culture every 2 h until the cultures reached saturation.

Construction of reporter plasmids for riboswitch activity. A DNA fragment (*Appendix 2*) encoding the 5'-UTR and first eight amino acids of the *eriC*^{F} gene from *Pseudomonas syringae* [*21*] plus 25 bp of homology to the pRS414 vector on the 5' and 3' ends was synthesized (Integrated DNA Technologies). This fragment was cloned into EcoRI/BamHI-digested pRS414 using one-step isothermal *in vitro* recombination [*22*] to generate a promoterless *lacZ* transcriptional fusion, with the eighth codon of *eriC* fused to the ninth codon of *lacZ* [*21*]. The sequence of the resulting plasmid was verified by Sanger sequencing.

***LacZ* reporter assay for riboswitch activity.** *E. coli* DH10B-T1^R transformed with the *eriC* riboswitch reporter construct were streaked onto LB agar supplemented with 50 μg/mL carbenicillin and 400 μg/mL X-gal in the presence or absence of 50 mM NaF. White colonies on the 0 mM NaF plate in combination with blue colonies on the 50 mM NaF were interpreted to indicate fluoride riboswitch expression.

6.3 Results and discussion

Of the 12 genes that comprise the *fl* locus (*Figure 6.2*), only two, *flA* and *flB*, have apparently biosynthetic functions as a fluorinase and a purine nucleoside phosphorylase, respectively [*18*]. Although *flD* has been proposed to encode a dehalogenase involved in fluorometabolite resistance [*18*], it contains the structural and catalytic motifs characteristic of a haloacid dehalogenase superfamily phosphatase as defined by the Structure-Function Linkage Database [*23*]. Because at least two of the known intermediates in fluorometabolite biosynthesis are phosphorylated (*Figure 6.1*), we considered *flD* to be a candidate biosynthetic gene.

To test the roles of *flA*, *flB*, and *flD* in fluorometabolite biosynthesis, we used a reverse molecular genetics approach in which the open reading frames encoding each gene were replaced through homologous recombination with an apramycin resistance gene. The resultant knockout strains were then characterized with respect to their ability to biosynthesize fluoroacetate and 4-fluorothreonine using ^{19}F NMR with an internal 5-fluorouracil standard to allow quantification of each fluorometabolite.

Organofluorine biosynthesis in *S. cattleya* Δ*flA*. Because the fluorinase is proposed to be the sole C-F bond-forming enzyme in *S. cattleya* [*4, 6*], the Δ*flA* strain was expected to be completely deficient in both fluoroacetate and 4-fluorothreonine biosynthesis. As predicted, ^{19}F NMR of culture supernatant of the Δ*flA* strain showed the absence of fluorometabolite production 3 d after sodium fluoride addition, while both fluoroacetate and fluorothreonine were apparent in wild-type *S. cattleya* grown under the same conditions (*Figure 6.3*). These results confirm that fluorinase-catalyzed fluoride attack on SAM represents the sole entry point for fluoride into the organofluorine biosynthesis pathway.

Organofluorine biosynthesis in *S. cattleya* Δ*flB*. While *flB* is proposed to encode a key enzyme for fluorometabolite biosynthesis [*11, 18*], the *S. cattleya* genome encodes at least two other purine nucleoside phosphorylases (PNPs) that could compensate for its absence in the Δ*flB*

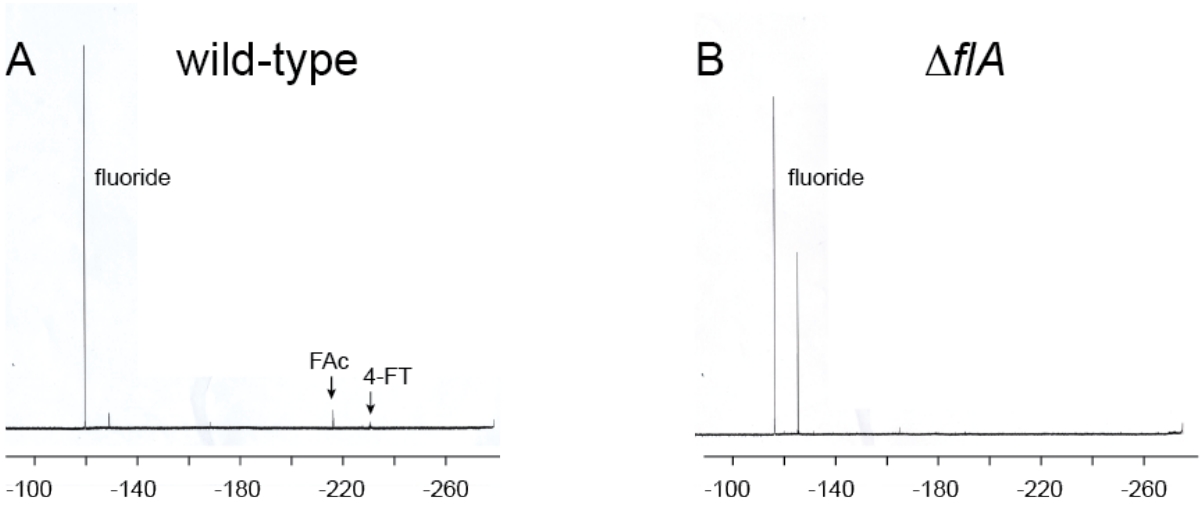


Figure 6.3. *^{19}F NMR spectrua of fluorometabolites from S. cattleya wild-type and ΔflA. (A) Wild-type culture produced both fluoroacetate (FAc) and 4-fluorothreonine (4-FT). (B) No fluorometabolites were detected in S. cattleya ΔflA. The peak at -130 ppm results from a reaction between fluoride and the silicate glass NMR tube.*

knockout strain depending on their substrate promiscuity [*24*]. Indeed, both fluoroacetate and fluorothreonine were observed in the culture medium from Δ*flB* both 3 d and 12 d after sodium fluoride was added (*Figure 6.4*), suggesting that other purine nucleoside phosphorylase orthologs can also convert 5'-FDA to 5-FDRP. Although fluoroacetate was present, the concentration in the culture supernatant was about half of that observed in the wild-type strain. This is consistent with analogous knockout studies in the chlorinase-harboring bacterium *Salinaspora tropica*, which showed that when the primary chlorinase pathway PNP is knocked out, other cellular PNP orthologs can accept the chlorinated substrate to generate 5-chlororibose-1-phosphate, albeit at reduced levels [*25*]. In another study of the FlB knockout in cell-free extract, no fluorometabolite biosynthesis was observed [*29*]. A possible explanation for this difference is that other factors needed for fluorometabolite biosynthesis or compensatory PNP activities did not survive cell lysis.

Remarkably, fluorothreonine levels in *S. cattleya* Δ*flB* were impacted very little or even increased compared to wild-type. There are several possible explanations for this result. One possibility is that replacement of the *flB* open reading frame leads to dysregulation of expression of another gene involved in fluoroacetate or fluorothreonine biosynthesis, causing a change in the relative levels of the two metabolites. Another explanation is that the pathways toward fluoroacetate and fluorothreonine diverge after the fluorinase-catalyzed step and that *flB* is not involved in fluorothreonine biosynthesis. Another possibility is that cellular PNPs with promiscuous activity toward 5'-FDA more effectively channel 5-FDRP to fluorothreonine. An additional possibility is that the FlB protein exerts an inhibitory effect on another protein involved in fluorothreonine biosynthesis and deletion of the gene relieves this inhibition. Examination of the product distribution from cells directly fed 5-FDRP could address the requirement for PNP activity in fluorothreonine biosynthesis, while overexpression of *flB* in wild-type cells might reveal any inhibitory effect that the gene product has on fluorothreonine biosynthesis.

Organofluorine biosynthesis in Δ*flD*. To test the possibility that *flD* plays a role in

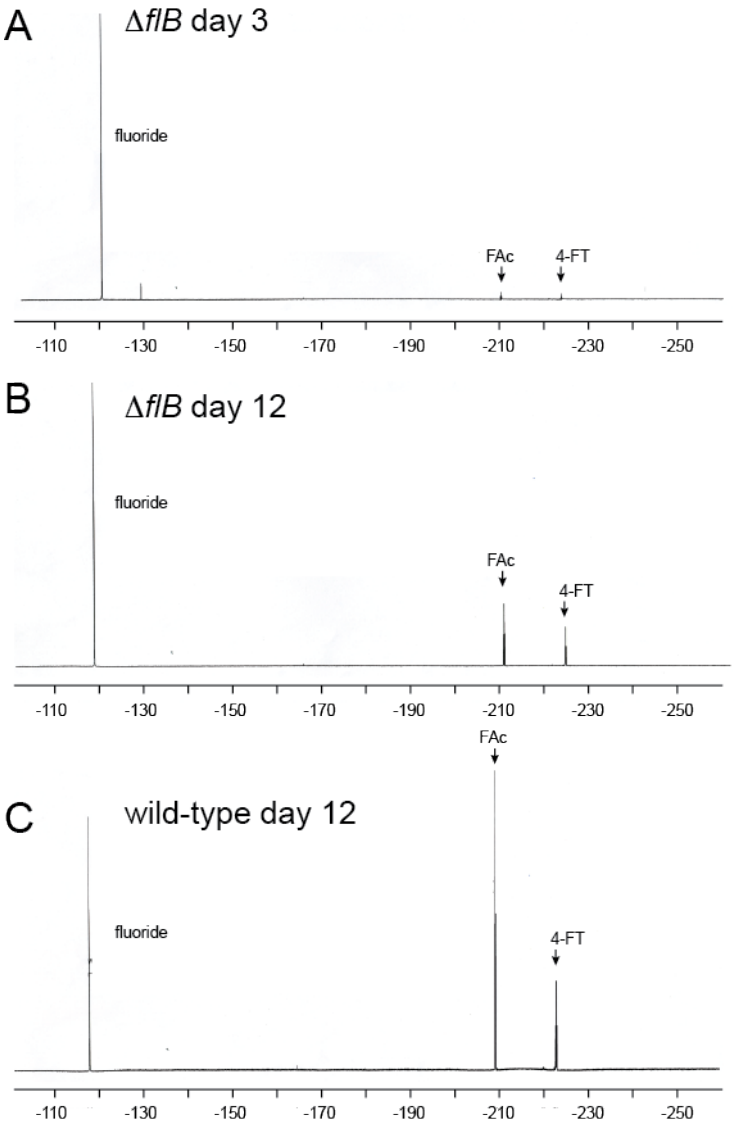


***Figure 6.4**. Fluorometabolite production in S. cattleya ΔflB. (A) ΔflB 3 days after fluoride addition. (B) ΔflB 12 days after fluoride addition. (C) Wild-type control 12 days after fluoride addition.*

fluorometabolite biosynthesis, the Δ*flD* strain was also constructed and assayed for fluoroacetate and fluorothreonine production using ^{19}F NMR. Both fluoroacetate and fluorothreonine were observed in the culture supernatant 3 d after fluoride addition at levels similar to wild-type *S. cattleya*, although the metabolites were not quantified. These results indicate that *flD* is not absolutely required for fluorometabolite biosynthesis and may play a regulatory role in organofluorine metabolism.

Examining the role of FlK in fluoroacetate resistance. Fluoroacetate is a potent toxin with an antimetabolite mode of action that exploits its close similarity to acetate, a key central metabolite [*26, 27*]. The conservative nature of the fluorine substitution allows fluoroacetate to be activated to fluoroacetyl-CoA by either the acetyl-CoA synthetase or the acetate kinase/phosphotransacetylase system. Fluoroacetyl-CoA can then be accepted by citrate synthase and converted to fluorocitrate, a mechanism-based inhibitor of the TCA cycle [*26, 27*]. The toxicity of fluoroacetate poses a problem for *S. cattleya*, suggesting that its genome is likely to encode enzymes that confer fluoroacetate resistance. A gene encoding a hotdog-fold thioesterase, *flK*, was observed downstream of *flA* (*Figure 6.2*) and was hypothesized to represent a fluoroacetyl-CoA-specific thioesterase that would hydrolyze fluoroacetyl-CoA to short circuit the lethal synthesis of fluorocitrate while preserving the high levels of acetyl-CoA required for cell growth [*18*]. Support for this hypothesis was provided by *in vitro* experiments using purified FlK that demonstrate its ability to hydrolyze fluoroacetyl-CoA with 10^6-fold selectivity over acetyl-CoA [*20*]. Additionally, overexpression of *flK* in a heterologous *E. coli* host demonstrated the competence of this gene to confer fluoroacetate resistance *in vivo* [*20*]. We therefore set out to address whether *flK* plays a role in fluoroacetate resistance in the context of fluorometabolite biosynthesis in *S. cattleya*.

To test the possibility that *flK* is involved in fluoroacetate resistance, we replaced its open

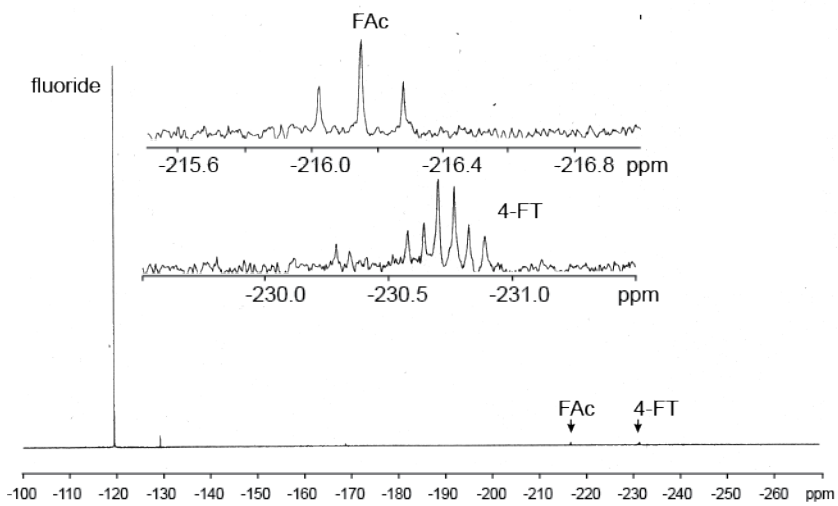


***Figure 6.5**. Fluorometabolite production in S. cattleya ΔflK. Supernatant was collected 3 d after fluoride addition.*

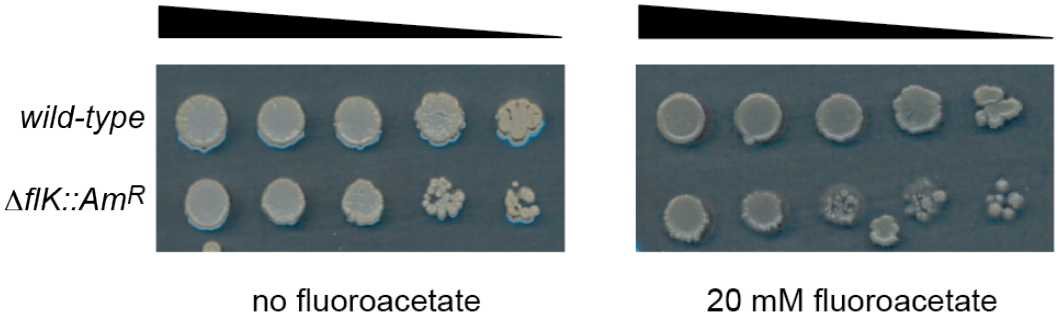


Figure 6.6. *Growth of S. cattleya wild-type and ΔflK on fluoroacetate-containing solid minimal medium.*

reading frame in the *S. cattleya* genome with an apramycin resistance marker. Assay of the culture medium from the resultant Δ*flK* strain by ^{19}F NMR demonstrated its ability to biosynthesize both fluoroacetate and 4-fluorothreonine at levels similar to wild-type (*Figure 6.5*), making this strain suitable for studies of fluoroacetate resistance. In contrast, no fluorometabolite biosynthesis was observed another study in which the FlK open reading frame was replaced [*29*[. However, these experiments were done in cell-free extract rather than in cel1 culture and it is therefore possible that other factors needed for fluorometabolite biosynthesis did not survive cell lysis. The growth rate and final optical density of the Δ*flK* strain were similar to wild-type under cell culture conditions, suggesting that alternative mechanisms of resistance might be at play.

Because the GYM medium typically used for fluorometabolite production is nutrient rich, metabolic pathways other than the TCA cycle may allow for normal growth under these conditions, even in the absence of *flK* expression. We therefore set out to test whether an *flK*-dependent growth defect might be apparent under more nutrient-limited conditions by serially diluting wild-type and Δ*flK* strains onto solid minimal medium (MM) in the presence and absence of 2 mM sodium fluoroacetate. No growth differences were apparent on MM with glucose as the sole carbon source (*Figure 6.6*). However, neither the wild-type nor the mutant strain showed any growth on MM with acetate as the sole carbon source, raising the possibility that *S. cattleya* is deficient in acetate uptake and assimilation. Indeed, bioinformatic analysis of the *S. cattleya* genome revealed the absence of an isocitrate lyase needed to form a functional glyoxylate shunt [*24*], which is usually required for growth on acetate.

The inability of *S. cattleya* to assimilate acetate made it necessary to test the role of *flK* under conditions in which fluoroacetate is produced intracellularly. To do this, we grew the wild-type and Δ*flK* strains in liquid supplemented minimal media (SMM) in the presence or absence of 2 mM sodium fluoride or 2 mM sodium fluoroacetate. We then monitored growth by periodically measuring the OD_{600} of the cultures until they reached saturation. Under these conditions, the Δ*flK* strain retains a slight growth defect, with a 1.2-fold increase in doubling time, suggesting a role for *flK* in fluoroacetate resistance (*Figure 6.7*). However, given the subtlety of the effect and its absence under other growth conditions, it is likely that this role is redundant. In contrast to the growth defect observed when sodium fluoride was added to the medium, Δ*flK* growth in the presence of 2 mM sodium fluoroacetate was indistinguishable from that of wild-type (*Figure*

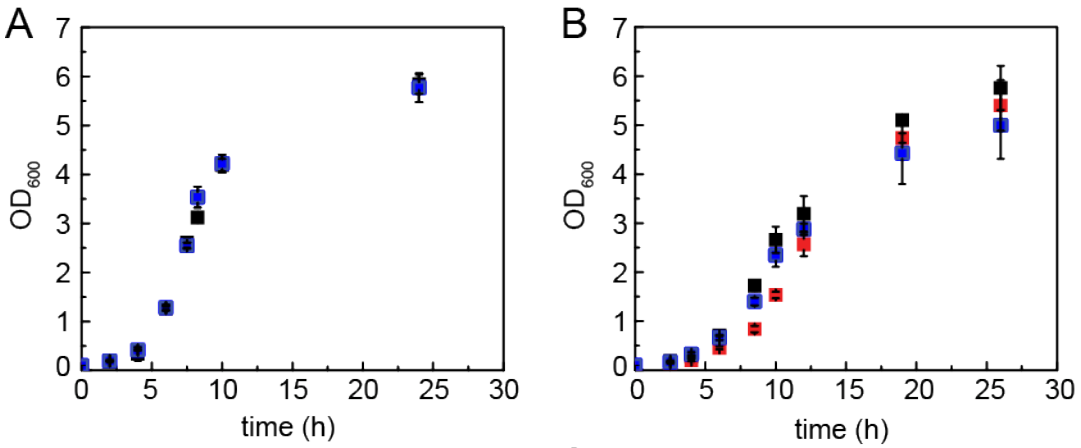


Figure 6.7. *Growth curves for S. cattleya (A) and S. cattleya DflK:Am^R (B) in GYM, pH 5 in the absence (black squares) or presence of 2 mM sodium fluoride (red circles) or 2 mM sodium fluoroacetate (blue triangles).*

6.8), suggesting that exogenously added fluoroacetate does not enter acetate metabolism and therefore does not exert the toxic effects of endogenously produced fluoroacetate.

FlK plays a redundant role in fluoroacetate resistance. Taken together, our studies of growth and organofluorine production suggest that *flK* may play a redundant role in fluoroacetate resistance, but that other factors may be more important. Recent transcriptomic studies [*24*] have revealed that the major mode of fluoroacetate resistance may be due to temporal coordination of organofluorine production and growth phase rather than the ability of the *S. cattleya* metabolic machinery to distinguish between fluorinated and nonfluorinated molecules. A comparison of expression levels of the TCA cycle genes over time in the presence and absence of fluoride revealed that the expression of these genes is highly responsive to growth phase but not to fluoride addition. In contrast, the gene encoding the fluorinase (*flA*) is upregulated both in response to fluoride and to the onset of stationary phase. Examination of the transcriptional response of other genes that are likely involved in fluoroacetate and fluorothreonine biosynthesis showed that the 5'-fluorodeoxyadenosine phosphorylase (*flB*), the fluoroacetaldehyde dehydrogenase (FAlDH) and the fluorothreonine transaldolase show the same behavior as *flA*, with stationary phase upregulation based on fluoride addition. It therefore appears that the fluorometabolite gene cluster and related genes are only expressed after exit from exponential growth and entrance into stationary phase results in a decrease in flux through the TCA cycle [*24*]. However, fluoride-dependent expression of many of the fluorometabolite biosynthetic genes demonstrates that there are many layers of regulation that control fluoroacetate and fluorothreonine production.

Progress toward knockout studies of other *fl* locus genes. In order to study the functions of the other genes in the *fl* locus, it is of interest to construct knockout strains that will allow for assessment of the phenotypic consequences of their absence. Additionally, the existence of knockout strains will provide a suitable background for expression of tagged versions of proteins of interest for purification and biochemical characterization and immunoprecipitation studies of protein-protein and protein-DNA interactions.

Toward this goal, we have constructed the plasmids required to disrupt *flC*, *flE*, *flG*, *flH*, *flI*, *flJ*, and *flL*, as well as the genes encoding the fluoroacetaldehyde dehydrogenase and the fluorothreonine transaldolase (*Appendix 2*). These plasmids contain an apramycin resistance

marker and oriT cassette flanked by 2000 bp of sequence derived from the regions immediately upstream and downstream of the gene of interest. Conjugative transfer of these constructs from *E. coli* to *S. cattleya* as described in Section 6.2 (Materials and methods) should allow for isolation of *S. cattleya* strains in which each gene is replaced with the apramycin resistance gene.

Testing the function of the *flL* 5′-UTR in organofluorine regulation. The observation of fluoride-dependent changes in transcription [*24*] raises the question of how *S. cattleya* is able to sense fluoride ions and translate this information into changes in gene expression. A fluoride-binding riboswitch associated with the *crcB* non-coding RNA motif was recently identified [*21*] and shown to discriminate against other halide ions [*21, 28*]. This fluoride riboswitch has been observed in bacteria and archaea and was found to activate expression of putative fluoride transporters, enzymes that are known to be inhibited by fluoride, and other genes of unknown function upon binding to fluoride [*21*]. Although the presently characterized fluoride riboswitches appear to play a role in alleviating fluoride toxicity, one could also envision their use as regulators in an organism that metabolizes fluoride.

Interestingly, the *fl* locus of *S. cattleya* encodes a putative transcription factor, *flL*, that has an unusually long 5′-UTR, which was observed to be transcribed in RNA-Seq experiments. It is tempting to speculate that this 5′-UTR represents a fluoride riboswitch that controls expression of a transcription factor, providing a mechanism to translate the presence of fluoride ions into changes in gene expression. As a first step toward testing whether *flL* expression is regulated by a fluoride riboswitch, a plasmid was generated to act as a positive control for riboswitch expression. This plasmid contained the known $eriC^F$ riboswitch from *Pseudomonas syringae* plus the first eight codons of the $eriC^F$ gene fused in frame to the ninth codon of *lacZ* to generate a promoterless *lacZ* transcriptional fusion. *E. coli* harboring this plasmid gave white colonies in the presence of X-gal and in the absence of NaF, but gave blue colonies in the presence of both X-gal and 50 mM NaF. This demonstrates fluoride-dependent gene expression indicating transcriptional regulation by a fluoride riboswitch. This *lacZ* assay will be suitable for future use in testing for the presence of fluoride riboswitches in the *S. cattleya* genome.

6.4 Conclusion

Using a molecular genetics approach, we have shown that the fluorinase (FlA) represents the sole entry point to organofluorine metabolism in *S. cattleya*. We have also demonstrated that FlB is involved in fluoroacetate biosynthesis, although it appears that other cellular purine nucleoside phosphorylases can also function in organofluorine metabolism. Interesting, FlB seems to be less important for 4-fluorothreonine biosynthesis, as the knockout strain produces similar levels of this metabolite compared to wild-type *S. cattleya*. An *S. cattleya* Δ*flD* strain has also been generated, and preliminary characterization indicates that it produces both fluoroacetate and 4-fluorothreonine, although the levels were not quantified. Plasmids were constructed for making knockout strains of the remaining *fl* locus genes, as well as the genes encoding the fluoroacetaldehyde dehydrogenase and the threonine transaldolase.

We have also explored the question of how *S. cattleya* is able to manage the toxic intracellular production of fluoroacetate, given its high effectiveness as an inhibitor of the TCA cycle. The fluoroacetyl-CoA specific thioesterase, FlK, has been proposed to be involved in resistance as it can selectively reverse the formation of fluoroacetyl-CoA with a 10^6-fold selectivity over acetyl-CoA and has been shown to confer fluoroacetate resistance to *E. coli*.

Although the Δ*flK* strain retains a slight growth defect, it does not demonstrate any appreciable change in organofluorine production or final optical density resulting from fluoride or fluoroacetate addition under organofluorine production conditions. In fact, the major fluoroacetate resistance mechanism was demonstrated to be temporal coordination of growth phase and secondary metabolite production through transcriptional regulation.

Finally, a previously reported assay for fluoride riboswitch activity was validated, demonstrating its usefulness for future studies of fluoride riboswitch regulation in the *S. cattleya* genome.

6.5 References

1. D. B. Harper and D. O'Hagan, The fluorinated natural products, *Nat. Prod. Rep.* **1994**, *11*, 123-133.
2. G. O. Morton, J. E. Lancaster, G. E. Van Lear, W. Fulmor and W. E. Meyer, The structure of nucleocidin. 3. (A new structure), *J. Am. Chem. Soc.* **1969**, *91*, 1535-1537.
3. M. Sanada, T. Miyano, S. Iwadare, J. M. Williamson, B. H. Arison, J. L. Smith, A. W. Douglas, J. M. Liesch and E. Inamine, Biosynthesis of fluorothreonine and fluoroacetic acid by the thienamycin producer, *Streptomyces cattleya*, *J. Antibiot. (Tokyo)* **1986**, *39*, 259-265.
4. D. O'Hagan, C. Schaffrath, S. L. Cobb, J. T. Hamilton and C. D. Murphy, Biochemistry: biosynthesis of an organofluorine molecule, *Nature* **2002**, *416*, 279.
5. C. Schaffrath, H. Deng and D. O'Hagan, Isolation and characterisation of 5'-fluorodeoxyadenosine synthase, a fluorination enzyme from *Streptomyces cattleya*, *FEBS Lett.* **2003**, *547*, 111-114.
6. C. Dong, F. Huang, H. Deng, C. Schaffrath, J. B. Spencer, D. O'Hagan and J. H. Naismith, Crystal structure and mechanism of a bacterial fluorinating enzyme, *Nature* **2004**, *427*, 561-565.
7. C. D. Cadicamo, J. Courtieu, H. Deng, A. Meddour and D. O'Hagan, Enzymatic fluorination in *Streptomyces cattleya* takes place with an inversion of configuration consistent with an S_N2 reaction mechanism, *Chembiochem* **2004**, *5*, 685-690.
8. S. L. Cobb, H. Deng, A. R. McEwan, J. H. Naismith, D. O'Hagan and D. A. Robinson, Substrate specificity in enzymatic fluorination. The fluorinase from *Streptomyces cattleya* accepts 2'-deoxyadenosine substrates, *Org. Biomol. Chem.* **2006**, *4*, 1458-60.
9. H. Deng, S. L. Cobb, A. R. McEwan, R. P. McGlinchey, J. H. Naismith, D. O'Hagan, D. A. Robinson and J. B. Spencer, The fluorinase from *Streptomyces cattleya* is also a chlorinase, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 759-762.
10. H. M. Senn, D. O'Hagan and W. Thiel, Insight into enzymatic C-F bond formation from QM and QM/MM calculations, *J. Am. Chem. Soc.* **2005**, *127*, 13643-13655.
11. S. L. Cobb, H. Deng, J. T. Hamilton, R. P. McGlinchey and D. O'Hagan, Identification of 5-fluoro-5-deoxy-D-ribose-1-phosphate as an intermediate in fluorometabolite biosynthesis in *Streptomyces cattleya*, *Chem. Commun. (Camb.)* **2004**, 592-593.

12. M. Sauter, B. Moffatt, M. C. Saechao, R. Hell and M. Wirtz, Methionine salvage and *S*-adenosylmethionine: essential links between sulfur, ethylene and polyamine biosynthesis, *Biochem. J.* **2013**, *451*, 145-154.

13. M. Onega, R. P. McGlinchey, H. Deng, J. T. Hamilton and D. O'Hagan, The identification of (3*R*,4*S*)-5-fluoro-5-deoxy-D-ribulose-1-phosphate as an intermediate in fluorometabolite biosynthesis in *Streptomyces cattleya*, *Bioorg. Chem.* **2007**, *35*, 375-385.

14. H. Deng, S. M. Cross, R. P. McGlinchey, J. T. Hamilton and D. O'Hagan, *In vitro* reconstituted biotransformation of 4-fluorothreonine from fluoride ion: Application of the fluorinase, *Chem. Biol.* **2008**, *15*, 1268-1276.

15. S. J. Moss, C. D. Murphy, D. O'Hagan, C. Schaffrath, J. T. G. Hamilton, W. C. McRoberts and D. B. Harper, Fluoroacetaldehyde: A precursor of both fluoroacetate and 4-fluorothreonine in *Streptomyces cattleya*, *Chem. Commun. (Camb.)* **2000**, 2281-2282.

16. C. D. Murphy, S. J. Moss and D. O'Hagan, Isolation of an aldehyde dehydrogenase involved in the oxidation of fluoroacetaldehyde to fluoroacetate in *Streptomyces cattleya*, *Appl. Environ. Microbiol.* **2001**, *67*, 4919-4921.

17. C. D. Murphy, D. O'Hagan and C. Schaffrath, Identification of a PLP-dependent threonine transaldolase: A novel enzyme involved in 4-fluorothreonine biosynthesis in *Streptomyces cattleya*, *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 4479-4481.

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

19. M. V. Dias, F. Huang, D. Y. Chirgadze, M. Tosin, D. Spiteller, E. F. Dry, P. F. Leadlay, J. B. Spencer and T. L. Blundell, Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK, *J. Biol. Chem.* **2010**, *285*, 22495-22504.

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

21. J. L. Baker, N. Sudarsan, Z. Weinberg, A. Roth, R. B. Stockbridge and R. R. Breaker, Widespread genetic switches and toxicity resistance proteins for fluoride, *Science* **2012**, *335*, 233-235.

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

23. S. C. Pegg, S. D. Brown, S. Ojha, J. Seffernick, E. C. Meng, J. H. Morris, P. J. Chang, C. C. Huang, T. E. Ferrin and P. C. Babbitt, Leveraging enzyme structure-function relationships for functional inference and experimental design: The structure-function linkage database, *Biochemistry* **2006**, *45*, 2545-2555.

24. M. C. Walker, M. Wen, A. M. Weeks and M. C. Chang, Temporal and fluoride control of secondary metabolism regulates cellular organofluorine biosynthesis, *ACS Chem. Biol.* **2012**, *7*, 1576-1585.

25. A. S. Eustaquio, F. Pojer, J. P. Noel and B. S. Moore, Discovery and characterization of a marine bacterial SAM-dependent chlorinase, *Nat. Chem. Biol.* **2008**, *4*, 69-74.

26. D. D. Clarke, Fluoroacetate and fluorocitrate: Mechanism of action, *Neurochem. Res.* **1991**, *16*, 1055-1058.

27. H. Lauble, M. C. Kennedy, M. H. Emptage, H. Beinert and C. D. Stout, The reaction of fluorocitrate with aconitase and the crystal structure of the enzyme-inhibitor complex, *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 13699-13703.

28. A. Ren, K. R. Rajashankar and D. J. Patel, Fluoride ion encapsulation by Mg^{2+} ions and phosphates in a fluoride riboswitch, *Nature* **2012**, *486*, 85-89.

29. C. Zhao, P. Li, Z. Deng, H. Y. Ou, R. P. McGlinchey and D. O'Hagan, Insights into fluorometabolite biosynthesis in *Streptomyces cattleya* DSM46488 through genome sequence and knockout mutants, *Bioorg. Chem.* **2012**, *44*, 1-7.

Appendix 1: *Published plasmids and oligonucleotides*

Table 1. Plasmid constructs.

name	location	primers	restriction sites	method
pET23a-His-Tev-SfoI	489	--	--	--
pET23a-His-Tev-sFIK	493	sFIK F101/R101	SfoI/XhoI	ligation
pET23a-His-Tev-sFIK-T42S	1388	sFIK T42S F1/R1	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-T42A	1390	sFIK T42A F1/R1	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-T42C	611	sFIK T42C F1/R1	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-H76A	793	sFIK H76A F2/R2	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-F36A	678	sFIK F36A F1/R1	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-F33A	1389	sFIK F33A F1/R1	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-V23A	1391	sFIK V23A F1/R1	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-V23Q	1392	sFIK V23Q F1/R1	SfoI/XhoI	Quickchange
pET23a-His-Tev-sFIK-L26A	1393	sFIK L26A F1/R1	SfoI/XhoI	Quickchange
pSV272.1	390	--	SfoI/XhoI	--
pSV272-His-MBP-TEV-sFIK-E50Q	798	sFIK F13/R15	SfoI/XhoI	ligation

Table 2. List of oligonucleotides used for sFlK gene assembly.

name	*sequence*
sFlK F1	CGGCGAGTTGCATATGAAAGATGGTATGCGTGTAGGCGAGC
sFlK R1	CATCTTTCATATGCCCCTCGCCC
sFlK F2	GCTTTACCCATGACTTCGTTGTTCCACCGCACAAAACCG
sFlK R2	ACAACGAAGTCATGGGTAAAGCGCTCGCCTACACGCATAC
sFlK F3	TGCGCCATCTGTATCCGGAATCTCCAGAGTTCGCAGAGT
sFlK R3	CCGGATACAGATGGCGCACGGTTTTGTGCGGTGGA
sFlK F4	TCCCGGAGGTGTTCGCTACGGGTTTCATGGTAGGCC
sFlK R4	GCGAACACCTCCGGGAACTCTGCGAACTCTGGAGATT
sFlK F5	TGATGGAATGGGCGTGCGTGCGCGCAATGGCCC
sFlK R5	GCACGCCCATTCCATCAGGCCTACCATGAAACCCGTA
sFlK F6	CGTACCTGGAACCGGGTGAAGGTTCTCTGGGCACC
sFlK R6	ACCCGGTTCCAGGTACGGGGCCATTGCGCGCAC
sFlK F7	GCGATCTGTGTGACTCACACCGCAGCAACCCCGCC
sFlK R7	TGTGAGTCACACAGATCGCGGTGCCCAGAGAACCTTC
sFlK F8	AGGTCTGACCGTGACCGTAACTGCGGAACTGCGTT
sFlK R8	CGGTCACGGTCAGACCTGGCGGGGTTGCTGCGG
sFlK F9	CCGTTGAGGGTCGTCGTCTGAGCTGGCGCGTATC
sFlK R9	ACGACGACCCTCAACGGAACGCAGTTCCGCAGTTA
sFlK F10	TGCTCACGACGGTGTTGACGAAATCGGTTCTGGCACC
sFlK R10	CAACACCGTCGTGAGCAGATACGCGCCAGCTCAG
sFlK F11	CACGAGCGTGCAGTTATCCACCTGGAGAAATTCAATGCAAAGGT
sFlK R11	TGGATAACTGCACGCTCGTGGGTGCCAGAACCGATTTCGT
sFlK F12	GCGTCAGAAAACCCCAGCGGGTTAAGGATCCgcgacgg
sFlK R12	CGCTGGGGTTTTCTGACGCACCTTTGCATTGAATTTCTCCAGG
sFlK R13	TTTTCGCCGGGGCCCCCCGTCCCGGATCCTTAACC
sFlK R14	GGGGCCCCGGCGAAAA
sFlK F100	AGCATCTTTCATATGAAAGATGGTATGCGTGTAGGCGAGCGC
sFlK R100	ATAGGATCCTTAACCCGCTGGGGTTTTCTGACGCACC
sFlK F13	TATAGGCGCCATGAAAGATGGTATGCGTGTAGGCGAGCGCTTTA
sFlK R15	TATACTCGAGTTAACCCGCTGGGGTTTTCTGACGCACCTTT

Figure 1. Primer map for sFlK gene synthesis. Pairs of primers are shown in alternating red and black.

```
                    Nde I
5'-CGGCGATTTG   CATATGAAAG   ATGGTATGCG   TGTAGGCGAG   CGCTTTACCC
3'-GCCGCTCAAC   GTATACTTTC   TACCATACGC   ACATCCGCTC   GCGAAATGGG

   ATGACTTCGT   TGTTCCACCG   CACAAAACCG   TGCGCCATCT   GTATCCGGAA
   TACTGAAGCA   ACAAGGTGGC   GTGTTTTGGC   ACGCGGTAGA   CATAGGCCTT

   TCTCCAGAGT   TCGCAGAGTT   CCCGGAGGTG   TTCGCTACGG   GTTTCATGGT
   AGAGGTCTCA   AGCGTCTCAA   GGGCCTCCAC   AAGCGATGCC   CAAAGTACCA

   AGGCCTGATG   GAATGGGCGT   GCGTGCGCGC   AATGGCCCCG   TACCTGGAAC
   TCCGGACTAC   CTTACCCGCA   CGCACGCGCG   TTACCGGGGC   ATGGACCTTG

   CGGGTGAAGG   TTCTCTGGGC   ACCGCGATCT   GTGTGACTCA   CACCGCAGCA
   GCCCACTTCC   AAGAGACCCG   TGGCGCTAGA   CACACTGAGT   GTGGCGTCGT

   ACCCCGCCAG   GTCTGACCGT   GACCGTAACT   GCGGAACTGC   GTTCCGTTGA
   TGGGGCGGTC   CAGACTGGCA   CTGGCATTGA   CGCCTTGACG   CAAGGCAACT

   GGGTCGTCGT   CTGAGCTGGC   GCGTATCTGC   TCACGACGGT   GTTGACGAAA
   CCCAGCAGCA   GACTCGACCG   CGCATAGACG   AGTGCTGCCA   CAACTGCTTT

   TCGGTTCTGG   CACCCACGAG   CGTGCAGTTA   TCCACCTGGA   GAAATTCAAT
   AGCCAAGACC   GTGGGTGCTC   GCACGTCAAT   AGGTGGACCT   CTTTAAGTTA

                                             Bam HI
   GCAAAGGTGC    GTCAGAAAAC   CCCAGCGGGT   TAAGGATCCG   CGACGG-3'
   CGTTTCCACG    CAGTCTTTTG   GGGTCGCCCA   ATTCCTAGGC   GCTGCCCCCC

   GGGGCCGCTTTT-5'
```

Table 3. List of oligonucleotides used for site-directed mutagenesis of sFlK.

name	*sequence*
sFlK E50Q F1	GGTAGGCCTGATGCAGTGGGCGTGCGTGC
sFlK E50Q R1	CCATCCGGACTACGTCACCCGCACGCACG
sFlK T42S F1	CCCGGAGGTGTTCGCTAGCGGTTTCATGGTAG
sFlK T42S R1	CTACCATGAAACCGCTAGCGAACACCTCCGGG
sFlK T42A F1	TTCCCGGAGGTGTTCGCTGCTGGTTTCATGGTAGG
sFlK T42A R1	CCTACCATGAAACCAGCAGCGAACACCTCCGGGAA
sFlK T42C F1	GTTCCCGGAGGTGTTCGCTTGCGGTTTCATGGTAGGC
sFlK T42C R1	CAAGGGCCTCCACAAGCGAACGCCAAAGTACCAATCCG
sFlK H76A F1	GCGATCTGTGTGACTGCCACCGCAGCAACCCC
sFlK H76A R1	GGGGTTGCTGCGGTGGCAGTCACACAGATCGC
sFlK F33A F1	TCCGGAATCTCCAGAGGCGGCAGAGTCCCGGAGG
sFlK F33A R1	CCTCCGGGAACTCTGCCGCCTCTGGAGATTCCGGA
sFlK F36A F1	CTCCAGAGTTCGCAGAGGCGCCGGAGGTGTTCGCTAC
sFlK F36A R1	GTAGCGAACACCTCCGGCGCCTCTGCGAACTCTGGAG
sFlK V23A F1	CACCGCACAAAACCGCGCGCCATCTGTATCC
sFlK V23A R1	GGATACAGATGGCGCGCGGTTTTGTGCGGTG
sFlK V23Q F1	CCACCGCACAAAACCCAGCGCCATCTGTATCC
sFlK V23Q R1	GGATACAGATGGCGCTGGGTTTTGTGCGGTGG
sFlK V23N F1	TTGTTCCACCGCACAAAACCAATCGCCATCTGTATCCGGAATC
sFlK V23N R1	GATTCCGGATACAGATGGCGATTGGTTTTGTGCGGTGGAACAA
sFlK L26A F1	ACAAAACCGTGCGCCATGCGTATCCGGAATCTCCAG
sFlK L26A R1	CTGGAGATTCCGGATACGCATGGCGCACGGTTTTGT

Appendix 2: *Unpublished plasmids, strains, and oligonucleotides*

Table 1. Plasmids for constructing *Streptomyces cattleya* knockout strains. Plasmid construction is described in detail in section 6.2. Primer sequences are listed in Table 2.

name	*location*	*primers*	*restriction site*	*method*
pRS316-ΔThrAld	1032	ThrAld K, U, and D	KpnI	yeast recombination
pRS316-ΔflL	1031	flL K, U, and D	KpnI	yeast recombination
pRS316-ΔflJ	1030	flJ K, U, and D	KpnI	yeast recombination
pRS316-ΔflH	1029	flH K, U, and D	KpnI	yeast recombination
pRS316-ΔflE	1028	flE K, U, and D	KpnI	yeast recombination
pRS316-ΔflC	1027	flC K, U, and D	KpnI	yeast recombination
pRS316-ΔflB	1026	flB K, U, and D	KpnI	yeast recombination
pRS316-ΔflA	1025	flA K, U, and D	KpnI	yeast recombination
pRS316-ΔFAlDH	922	FAlDH K, U, and D	KpnI	yeast recombination
pRS316-ΔflG	754	flG K, U, and D	KpnI	yeast recombination
pRS316-ΔflI	609	flI K, U, and D	KpnI	yeast recombination

Table 2. Primers used for construction of plasmids listed in Table 1.

name	*sequence*
ThrAld UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccagatgggcgtcgaagatccgggaaccggtcccggccgta
ThrAld UR1	aacttcgaagcagctccagcctacaagggctgggctcctcccatggcggtctcacggatcc
ThrAld KF1	accgccatgggaggagcccagcccttgtaggctggagctgcttc
ThrAld KR2	cggcggacgtcctcgatgagggcgaattccggggatccgtcgacct
ThrAld DF2	actgcaggtcgacggatccccggaattcgccctcatcgaggacgtccgccggatcaccggctga
ThrAld DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtacgagcaccgcgccgacggtgaacagcgcggcg
flL UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccacgccgccggctgcccgaactccccgcccatgaa
flL UR1	aacttcgaagcagctccagcctacagggtccctccggacgagggcgggggcc
flL KF1	cccccgccctcgtccggagggaccctgtaggctggagctgcttc
flL KR1	ggctttccggccgggtggcggccggattccggggatccgtcgacct
flL DF1	ctgcaggtcgacggatccccggaatccggccgccacccggccggaaagccacc
flL DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtacaggagtactggtggtgcacccggcaggcgctgacctgg
flJ UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccgcgccccgggtacgaaggtctgccacaccccgga
flJ UR1	aacttcgaagcagctccagcctacacttggcgttgaacttctccaggtggatcacggcgcgctcgtg
flJ KF1	ccacctggagaagttcaacgccaagtgtaggctggagctgcttc
flJ KR1	ccgatactgagggggacgcccgggattccggggatccgtcgacct
flJ DF1	ctgcaggtcgacggatccccggaatcccgggcgtccccctcagtatcggtagtggtccg
flJ DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtaccagcatcccgatgctgtagaaggcgccgagcacctcgc
flH UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccggagatggtggccatcccctccgcgcacctggcc
flH UR1	aacttcgaagcagctccagcctacaccctccaccgcggcgccgcgacggcc
flH KF1	gccgtcgcggcgccgcggtggagggtgtaggctggagctgcttc
flH KR1	cggcccggccttcgggaggtgcccgattccggggatccgtcgacct
flH DF1	ctgcaggtcgacggatccccggaatcgggcacctcccgaaggccgggccgaaaccgcat
flH DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtaccggcggtgtccaccgccaccaccatcaccggcag

Table 2 (cont'd).

name	*sequence*
flE UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccccagcagcaacccgtacgccgccccgagcaccac
flE UR1	aacttcgaagcagctccagcctacatcacttcgggaaattcggctcaagcgaggatgcccccgcg
flE KF1	gcttgagccgaatttcccgaagtgatgtaggctggagctgcttc
flE KR1	cccgcctcgggatcctcggacggcaattccggggatccgtcgacct
flE DF1	ctgcaggtcgacggatccccggaattgccgtccgaggatcccgaggcgggaccaggacct
flE DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtacgcccggcgtgcgcccgcgactccagcgtgatgtt
flC UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggccccttgaggtggtacgggtaggcgaggctggccgcgttgc
flC UR1	aacttcgaagcagctccagcctacaccgccgggccgccctcagccgcgc
flC KF1	ggcgcggctgagggcggcccggcggtgtaggctggagctgcttc
flC KR1	gcgggcgccggggccgcggccctggattccggggatccgtcgacct
flC DF1	ctgcaggtcgacggatccccggaatccagggccgcggccccggcgccc
flC DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtacggccctcggggaacgtttcggtgacgtactgatggccagc
flB UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccgcgtggaggaagaggtcgggggcgggtttgcc
flB UR1	aacttcgaagcagctccagcctacagggcggcccggcggtcagcgcggg
flB KF1	ccccgcgctgaccgccgggccgccctgtaggctggagctgcttc
flB KR1	atcagctacgaccggaggtgcgcgcattccggggatccgtcgacct
flB DF1	ctgcaggtcgacggatccccggaatgcgcgcacctccggtcgtagctgatcgcgg
flB DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtacgcctaggccctgtcgtcaacatgtcgcctgccc
flA UF1	ttcgatatcaagcttatcgataccgtcgacctcgagggggggcccggtacctcttcgcggggcggctgctggccggggtg
flA UR1	gtgaattcctcctgcgatcgctcctggtggatcagacccg
flA KF2	tatggtcgatcgggtctgatccaccaggagcgatcgcaggaggaattcactgtaggctggagctgcttc
flA KR1	attccggggatccgtcgacctcgcgttccggccgccccgcccgccggcgcg gacgggacggccggaaggggg
flA DF1	ccccttccggccgtcccgtccgcgcc
flA DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtactaccagaccctgcgggccatctcgctcgccttcttcgt

Table 2 (cont'd).

name	*sequence*
FAIDH UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggccctgatccggatgtccccggcgatcaccacgcccttgtccag
FAIDH UR1	aacttcgaagcagctccagcctacagccacagcctcctggtcggcgttccgggggcatttg
FAIDH KF2	ggaacgccgaccaggaggctgtggctgtaggctggagctgcttc
FAIDH KR1	acccgcgcggcgcccgcggtgtccaattccggggatccgtcgacct
FAIDH DF1	ctgcaggtcgacggatccccggaattggacaccgcgggcgccgcgcgggtcg
FAIDH DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtacccgaatcgccgccccttgaagcgggagcgagagca
flG UF1	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccccgcatgcgcgcacctccggtcgtagctgatcgcggat
flG UR1	aacttcgaagcagctccagcctacagccctgtcgtcaacatgtcgcctgccccgcgacgtctg
flG KF2	gcaggcgacatgttgacgacagggctgtaggctggagctgcttc
flG KR1	ccaccaggcgaaaggcggaaaccgcattccggggatccgtcgacct
flG DF1	actgcaggtcgacggatccccggaatgcggtttccgcctttcgcctggtggtctcggggtcc
flG DR1	tacgccaagctcggaattaaccctcactaaagggaacaaaagctgggtactcctcgggggtgccgtgcgggcccaccacca
flI UF2	aggaattcgatatcaagcttatcgataccgtcgacctcgagggggggcccgaggcgacgccgagcagcacgacgtggggc
flI UR1	ttcatgctcggccagcgtgatttccttgcgtccgaattccgc
flI KF1	ggaatacgccgggagccggccgctgtgtaggctggagctgcttc
flI KR1	attccggggatccgtcgacctgcagttcgaagttcctattctctag
flI DF1	gggggacgcccgggatggaagcgccc
flI DR1	atgattacgccaagctcggaattaaccctcactaaagggaacaaaagctggacagcggtggccggtcatgttggggacgtacc

Table 3. *Streptomyces cattleya* knockout strains. Construction of strains is described in detail in section 6.2.

name	*location*
S. cattleya Δ*flA*::Am^R	1354
S. cattleya Δ*flB*::Am^R	670
S. cattleya Δ*flD*::Am^R	1355
S. cattleya Δ*flK*::Am^R	1356

Table 4. Plasmids for heterologous expression of *fl* locus proteins in *E. coli.* Primer sequences are listed in Table 5.

name	*location*	*primers*	*restriction sites*	*method*
pET23a-His-Tev-flD	623	flD F1/R1	SfoI/XhoI	ligation
pET23a-His-Tev-flB	613	flB F3/R1	SfoI/XhoI	ligation
pSV272-His-MBP-Tev-flB	610	flB F3/R1	SfoI/XhoI	ligation
pET23a-His-Tev-flI	612	flI F1/R1	SfoI/XhoI	ligation

Table 5. Primers for constructing plasmids listed in Table 4.

name	*sequence*
flD F1	gccatgccgccccagctggtcatattcgactgcgacgg
flD R1	agtctatactcgagttaccccacccccaacaaccccggcaactccc
flB F3	gccatgcgggcacggaaatcggggaacgagcagcgga
flB R1	attctcgagtcagccgcgccgggcgaacccctt
flI F1	gccatgccggggctgatggcgatgcgggcggaatacgc
flI R1	atgctatactcgagttagtatcggtagtggtccggcttgtacgggccctcc

Table 6. Plasmids for heterologous expression of the asparagine oxygenase (AsnO) from *Streptomyces coelicolor* in *E. coli*. Primer sequences are given in Table 7.

name	*location*	*primers*	*restriction sites*	*method*
pCWori-HisN-AsnO	485	asnO F2/R1	NdeI/HindIII	ligation
pCWori-HisN-AsnO-E157A	484	asnO F2/asnO E157A R1; asnO E157A F1/asnO R1	NdeI/HindIII	SOE/ligation
pCWori-HisN-AsnO-E157G	483	asnO F2/asnO E157G R1; asnO E157G F1/asnO R1	NdeI/HindIII	SOE/ligation

Table 7. Primers for construction of plasmids listed in Table 6.

name	*sequence*
asnO F2	`atatatatcatatggctgcgaatgccgcgggaccggcgtc`
asnO R1	`aaataagcttttaggcgggctgcggggccgctccg`
asnO E157A F1	`gacgttccacaacgcgaacgccttccacg`
asnO E157A R1	`cgtggaaggcgttcgcgttgtggaacgtc`
asnO E157G F1	`ctgacgttccacaacggcaacgccttccacgagc`
asnO E157G R1	`gctcgtggaaggcgttgccgttgtggaacgtcag`

Table 8. Plasmids for overexpression of *E. coli* CoA biosynthetic genes.

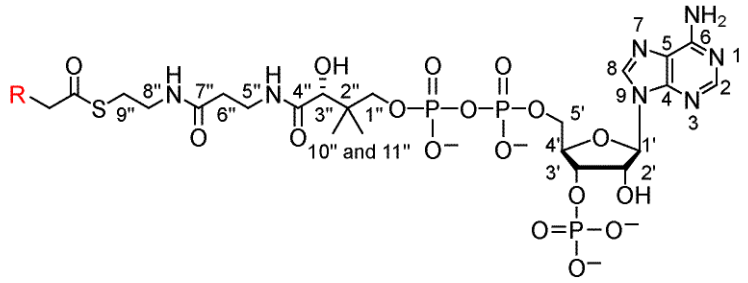
name	*location*	*primers*	*restriction sites*	*method*
pET23a-His-Tev-PanK	680	coaA F2/R2	SfoI/XhoI	ligation
pET23a-His-Tev-PPAT	776	coaD F2/R2	SfoI/XhoI	ligation
pET23a-His-Tev-DPCK	1307	coaE F2/R2	SfoI/XhoI	ligation

Table 9. Primers for construction of plasmids in Table 8.

name	*sequence*
coaA F2	`tataatcgaggcgccatgagtataaaagagcaaacgttaatgacgccttacctacagtttga`
coaA R2	`tcgattatactcgagttatttgcgtagtctgacctcttctaccgcatgattag`
coaD F2	`tataatcgaggcgccatgcaaaaacgggcgatttatccgggtactttcgat`
coaD R2	`tcgattatactcgagttacgctaacttcgccatcagcgcctgatg`
coaE F2	`tataatcgaggcgccatgaggtatatagttgccttaacgggaggcattgg`
coaE R2	`tcgattatactcgagttacggtttttcctgtgagacaaactgcgacgcaa`

Appendix 3: *2D NMR spectra and HPLC chromatograms for synthetic acyl-CoAs and analogs*

Figure 1. Crosspeak numbering for $^{1}H/^{13}C$ HMBC spectra.



Crosspeak numbering for HMBC spectra							
1	$H_{1''}C_{2''}$	**10**	$H_{10''}C_{11''}$	**19**	$H_{5''}C_{7''}$	**28**	H_2C_4
2	$H_{1''}C_{3''}$	**11**	$H_{10''}C_{2''}$	**20**	$H_{5''}C_{6''}$	**29**	H_2C_6
3	$H_{1''}C_{11''}$	**12**	$H_{10''}C_{1''}$	**21**	$H_{6''}C_{7''}$	**30**	H_8C_4
4	$H_{1''}C_{10''}$	**13**	$H_{10''}C_{3''}$	**22**	$H_{6''}C_{5''}$	**31**	H_8C_5
5	$H_{3''}C_{11''}$	**14**	$H_{11''}C_{10''}$	**23**	$H_{8''}C_{7''}$	**32**	$H_{1'}C_2$
6	$H_{3''}C_{10''}$	**15**	$H_{11''}C_{2''}$	**24**	$H_{8''}C_{9''}$	**33**	$H_{1'}C_4$
7	$H_{3''}C_{2''}$	**16**	$H_{11''}C_{1''}$	**25**	$H_{9''}C_{1RAc}$	**34**	$H_{1'}C_{4'}$
8	$H_{3''}C_{1''}$	**17**	$H_{11''}C_{3''}$	**26**	$H_{9''}C_{8''}$	**35**	$H_{2'}C_{1'}$
9	$H_{3''}C_{4''}$	**18**	$H_{5''}C_{4''}$	**27**	$H_{2RAc}C_{1RAc}$	**36**	$H_{3'}C_{4'}$

Figure 2. $^{1}H/^{13}C$ HMBC spectrum of fluoroacetyl-CoA in D_2O.

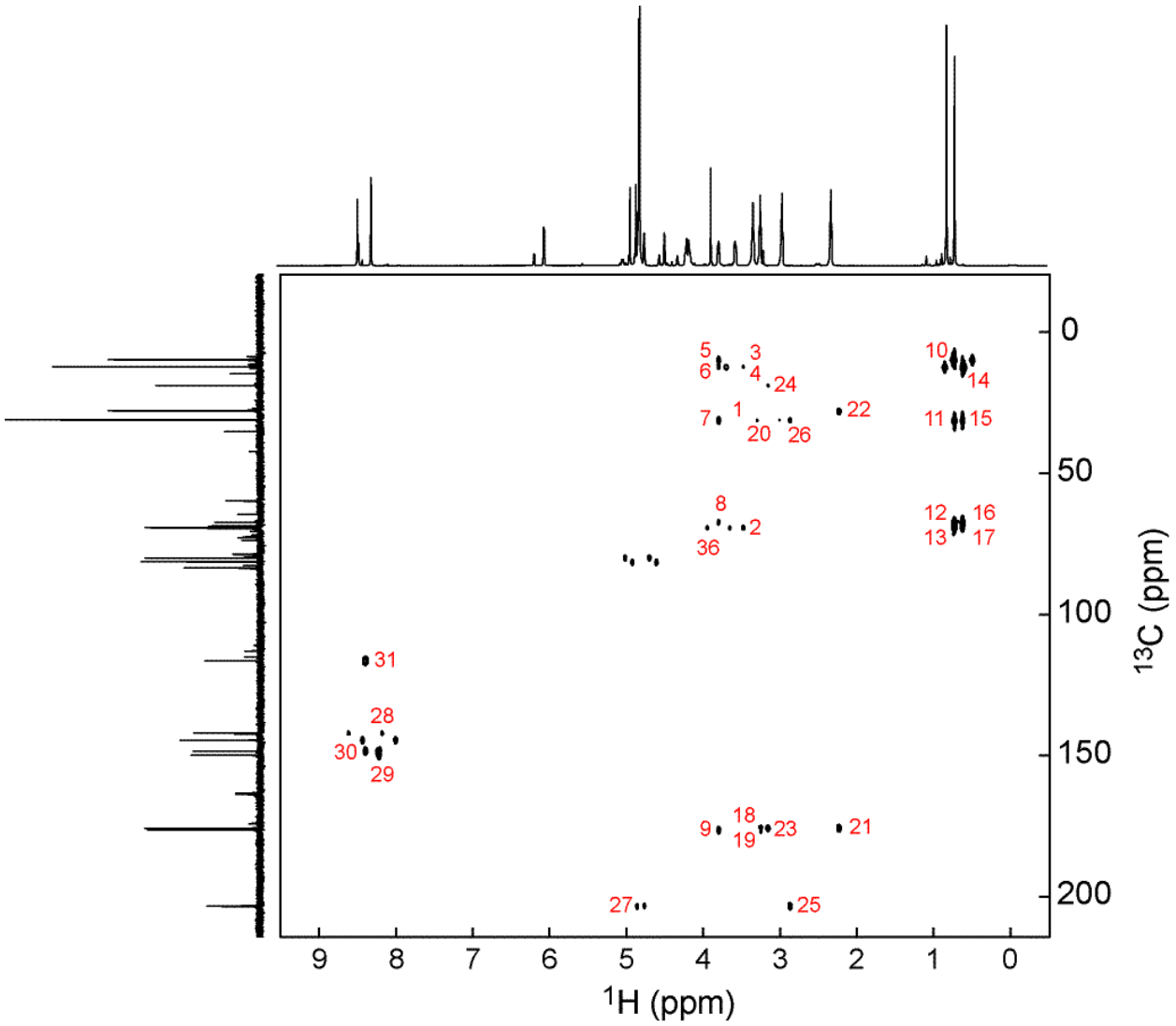


Figure 3. $^{1}H/^{13}C$ HMBC spectrum of chloroacetyl-CoA in D_2O.

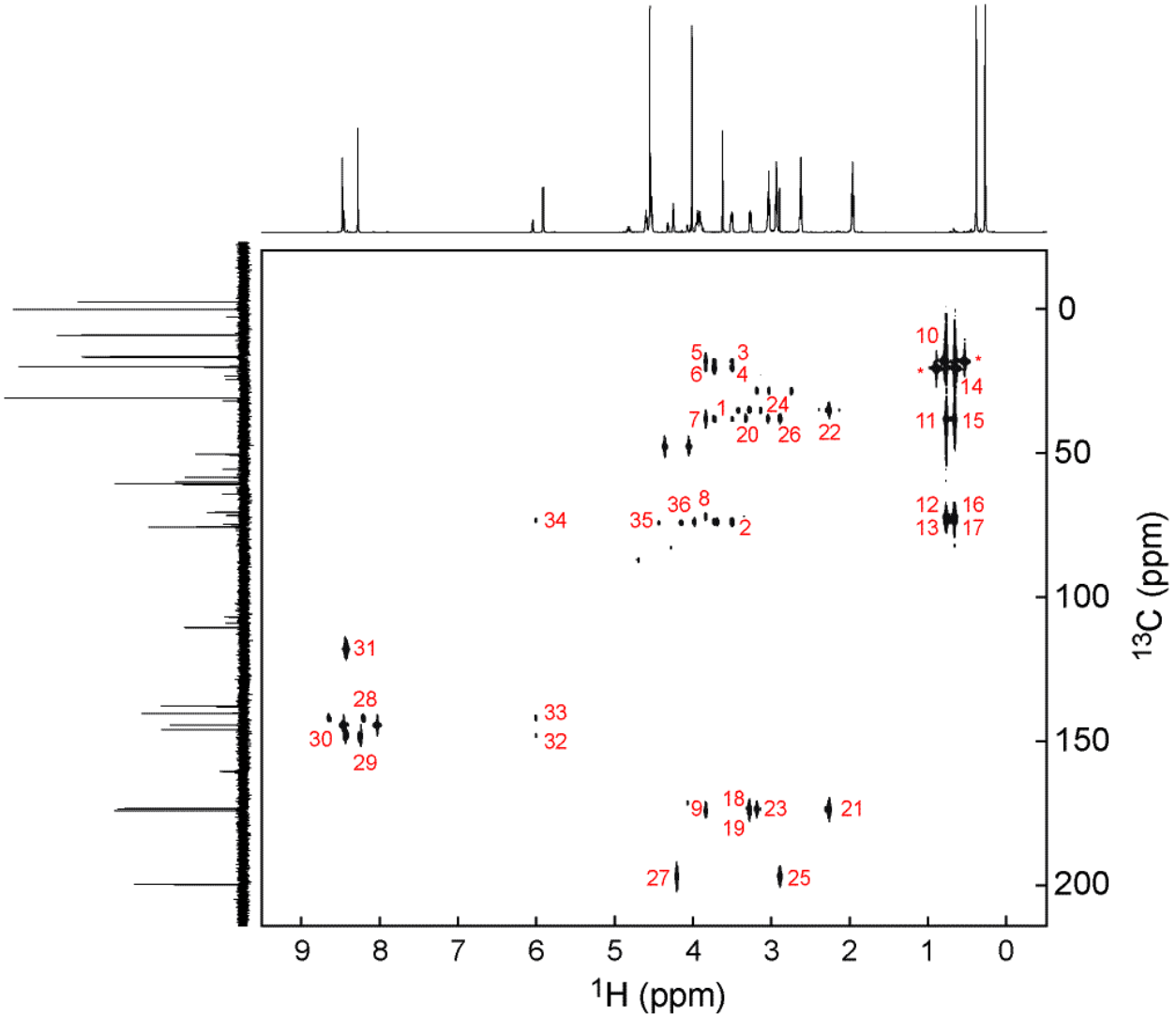


Figure 4. $^{1}H/^{13}C$ HMBC spectrum of bromoacetyl-CoA in D_2O.

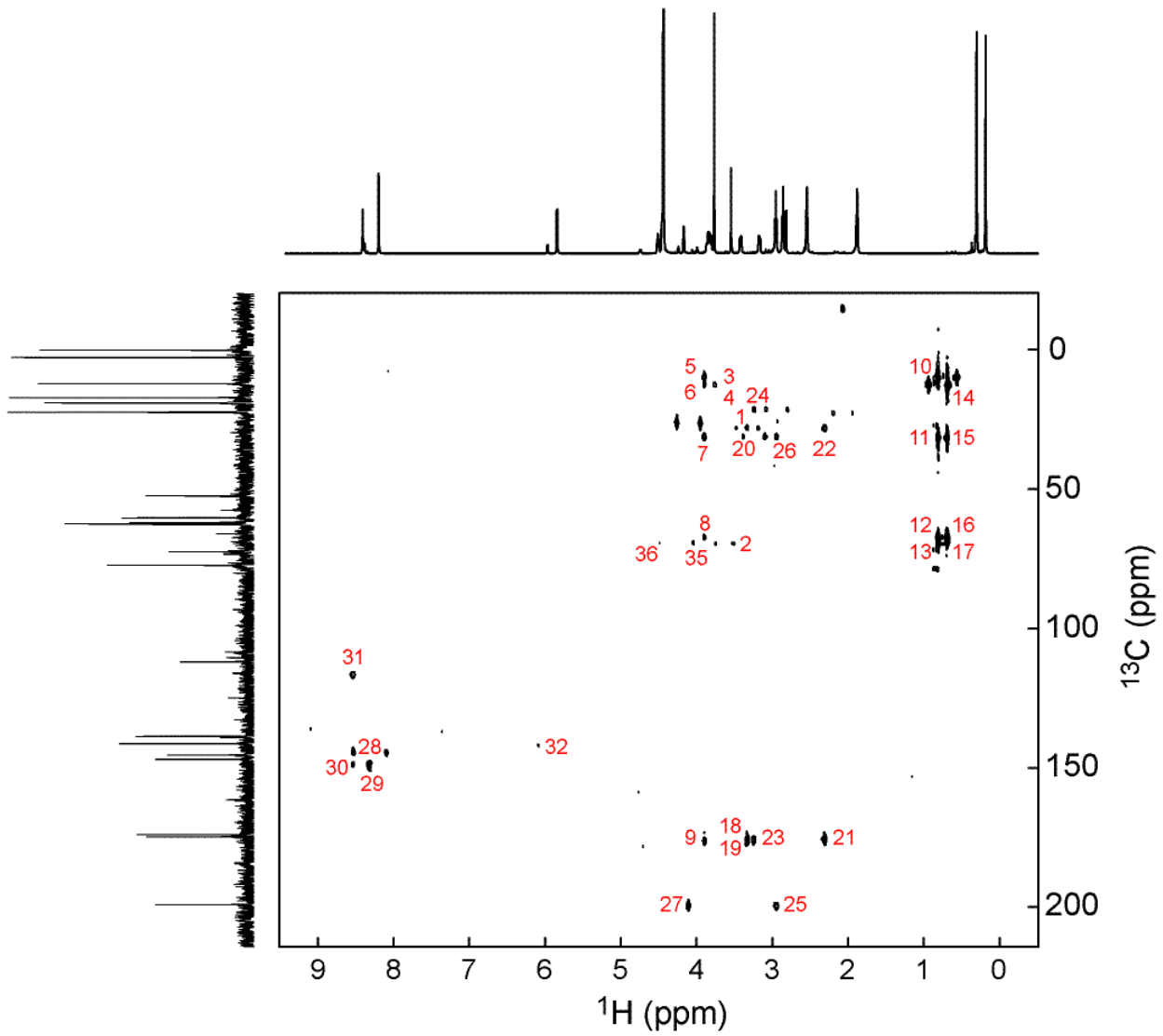


Figure 5. $^{1}H/^{13}C$ HMBC spectrum of cyanoacetyl-CoA in 90% H_2O/10%D_2O. $^{1}H/^{13}C$ HMBC indicates that the cyanoacetyl-CoA has a stable enolate form under these conditions. Crosspeaks corresponding to the enolate tautomer are circled. Peak 37 corresponds to the crosspeak between the α-protons and the carbon of the cyano group.

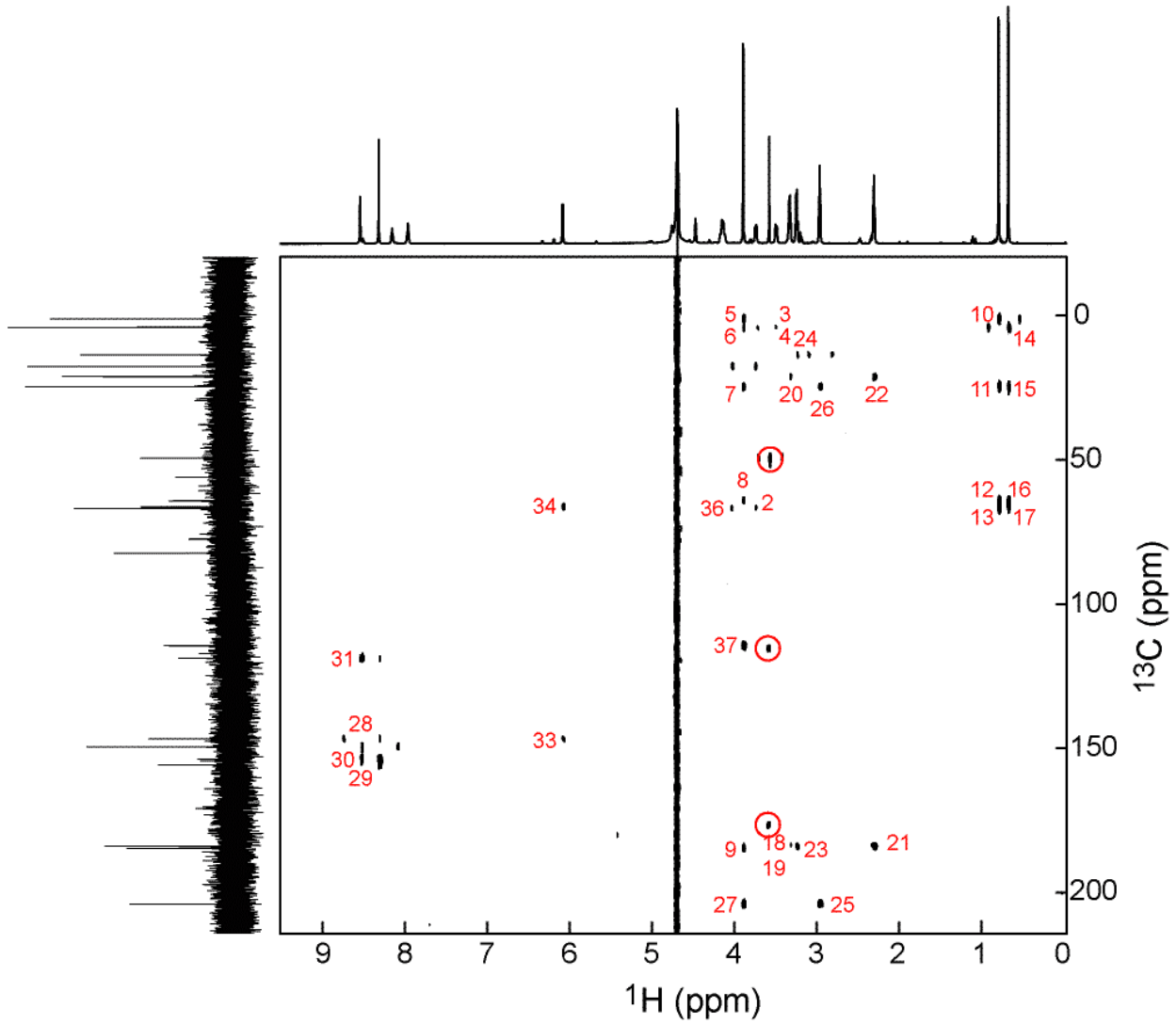


Figure 6. HPLC chromatograms for acyl-CoA FlK substrates. Acyl-CoA substrates were analyzed by RP-HPLC as described for non-enzymatic hydrolysis experiments in section 3.2. Absorbance was monitored at 260 nm.

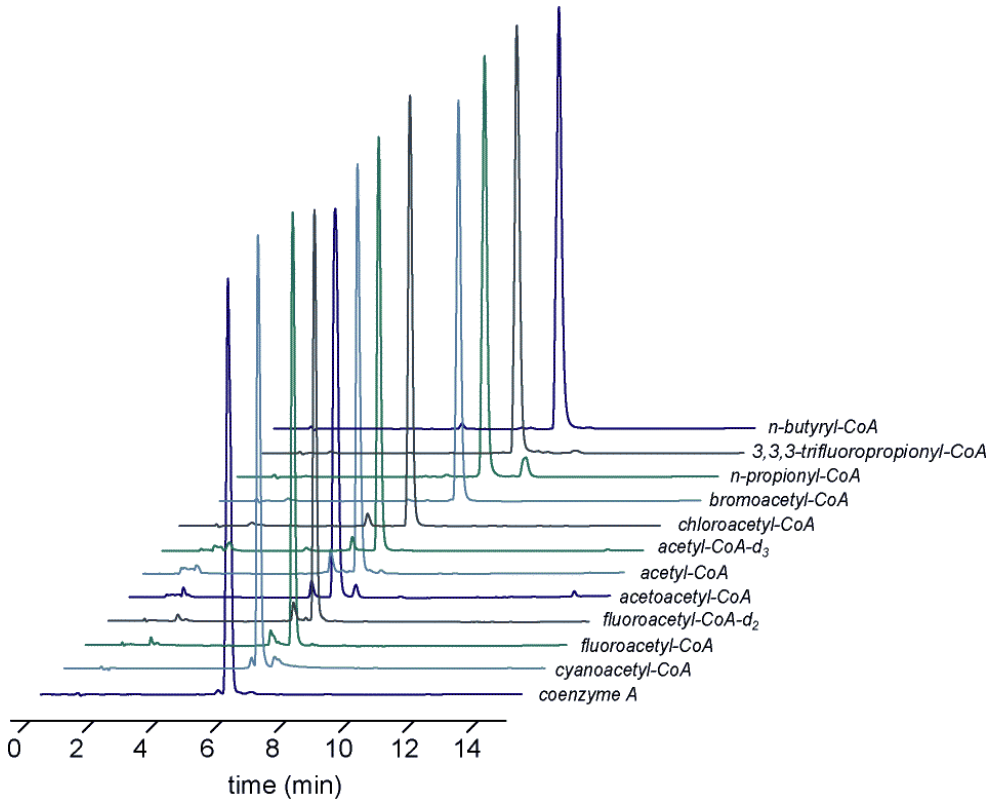


Figure 7. $^{1}H/^{13}C$ HMBC spectrum of fluoroacetyl-oxa(dethia)-CoA in D_2O.

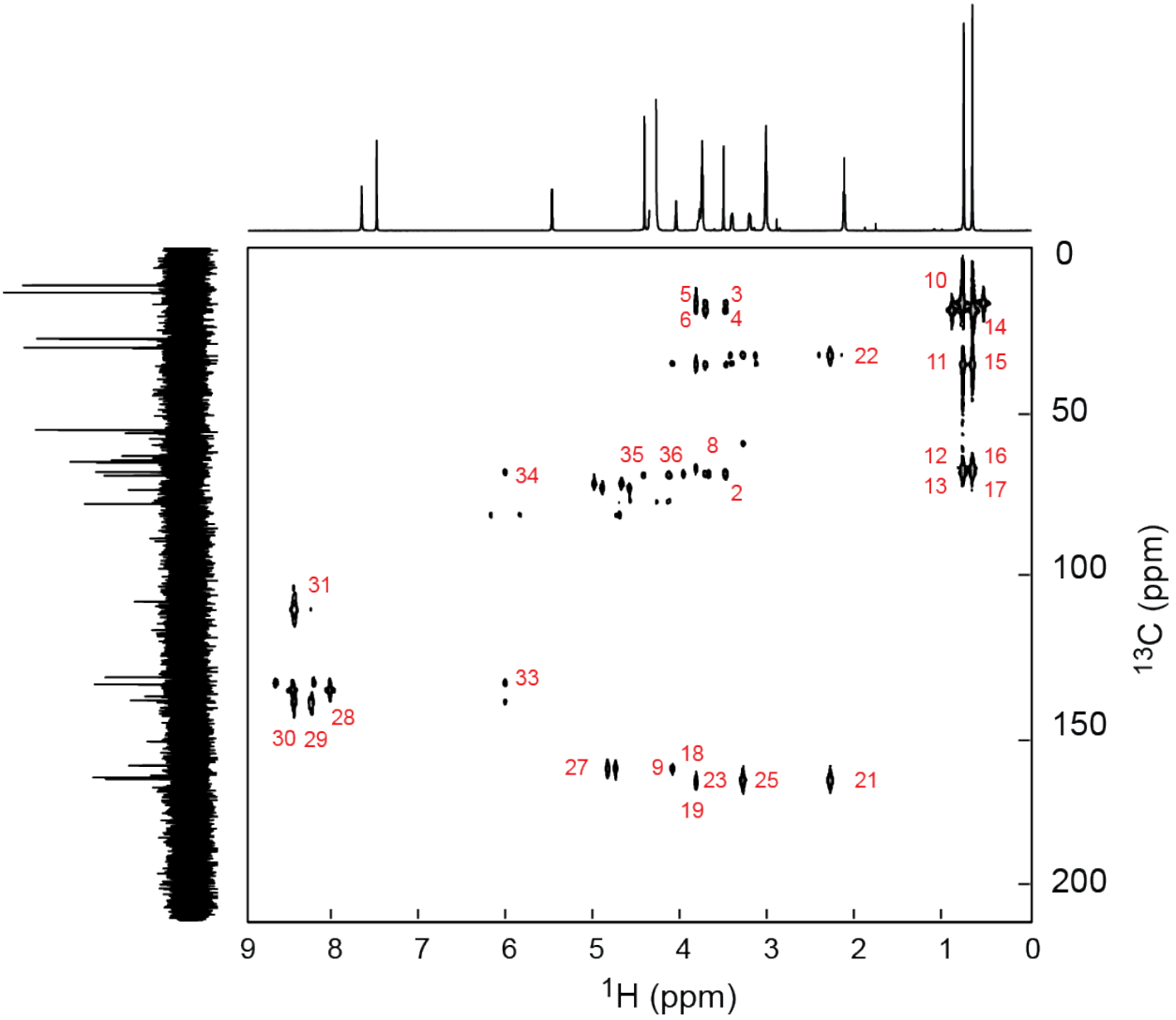


Figure 8. $^{1}H/^{13}C$ HMBC spectrum of fluoroacetyl-aza(dethia)-CoA in D_2O.

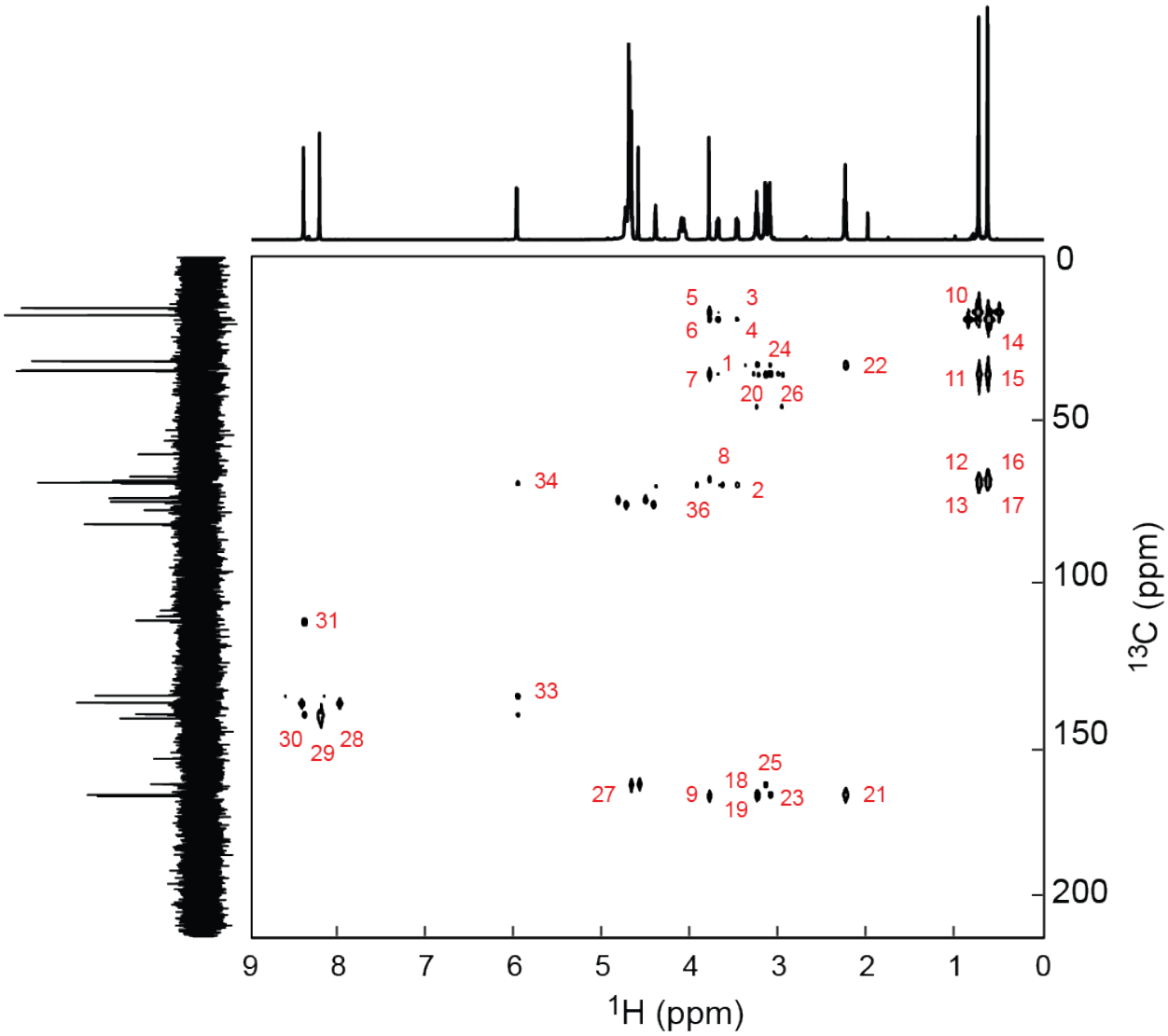


Figure 9. $^{1}H/^{13}C$ HMBC spectrum of acetyl-aza(dethia)-CoA in D_2O.

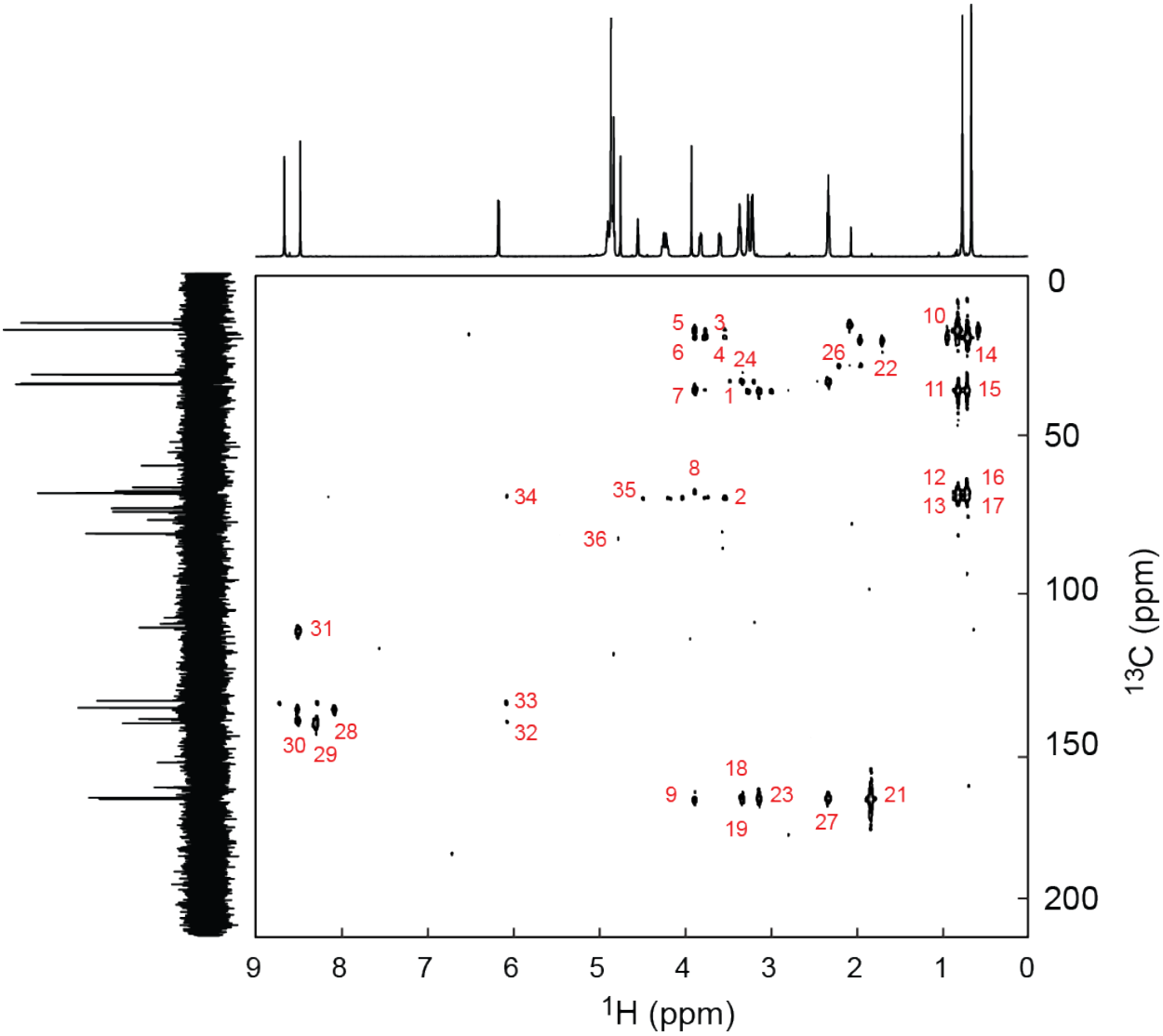


Figure 10. $^{1}H/^{13}C$ HMBC spectrum of fluoroacetyl-carba(dethia)-CoA in D_2O.

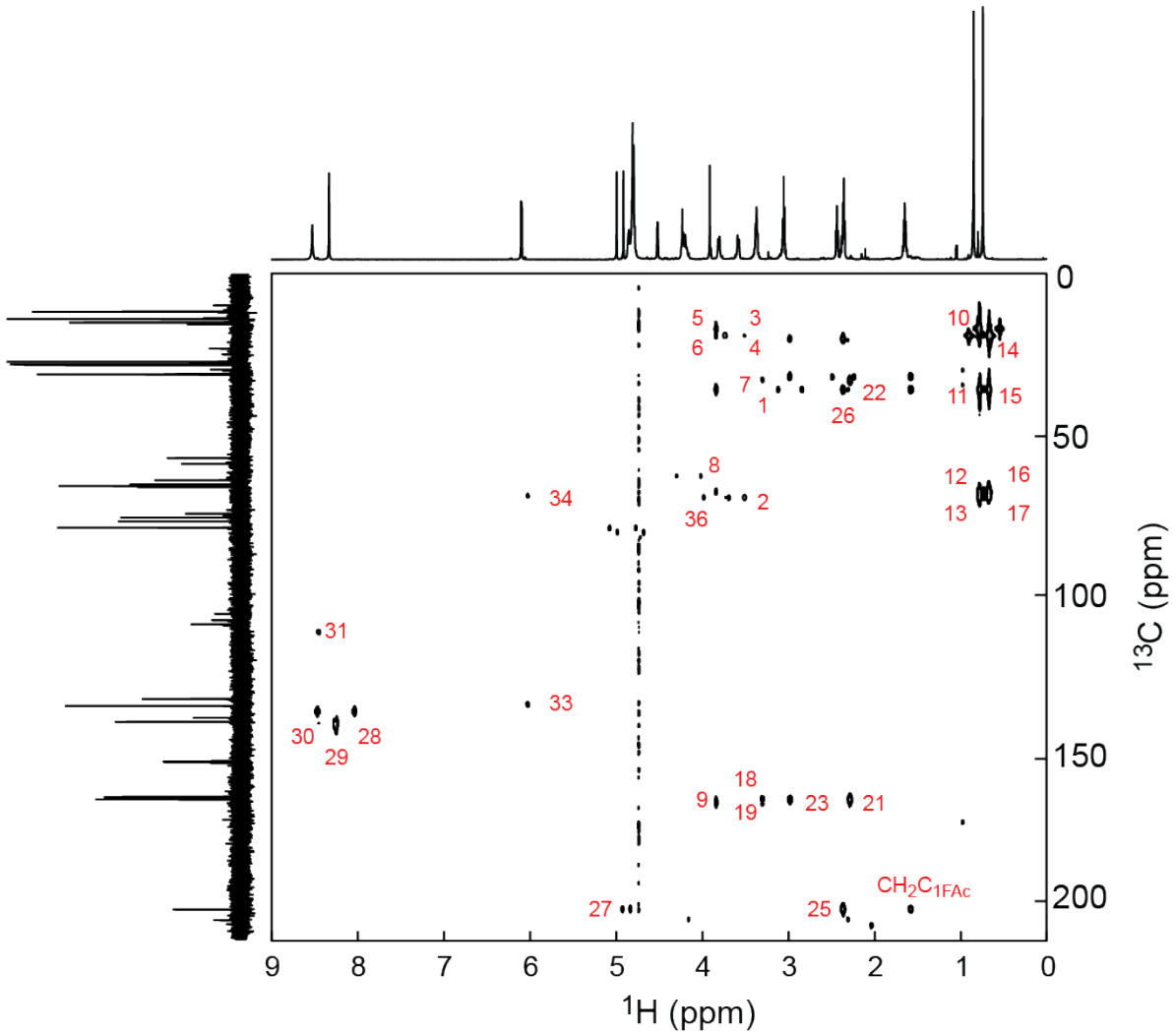


Figure 11. $^{1}H/^{13}C$ HMBC spectrum of acetyl-carba(dethia)-CoA in D_2O.

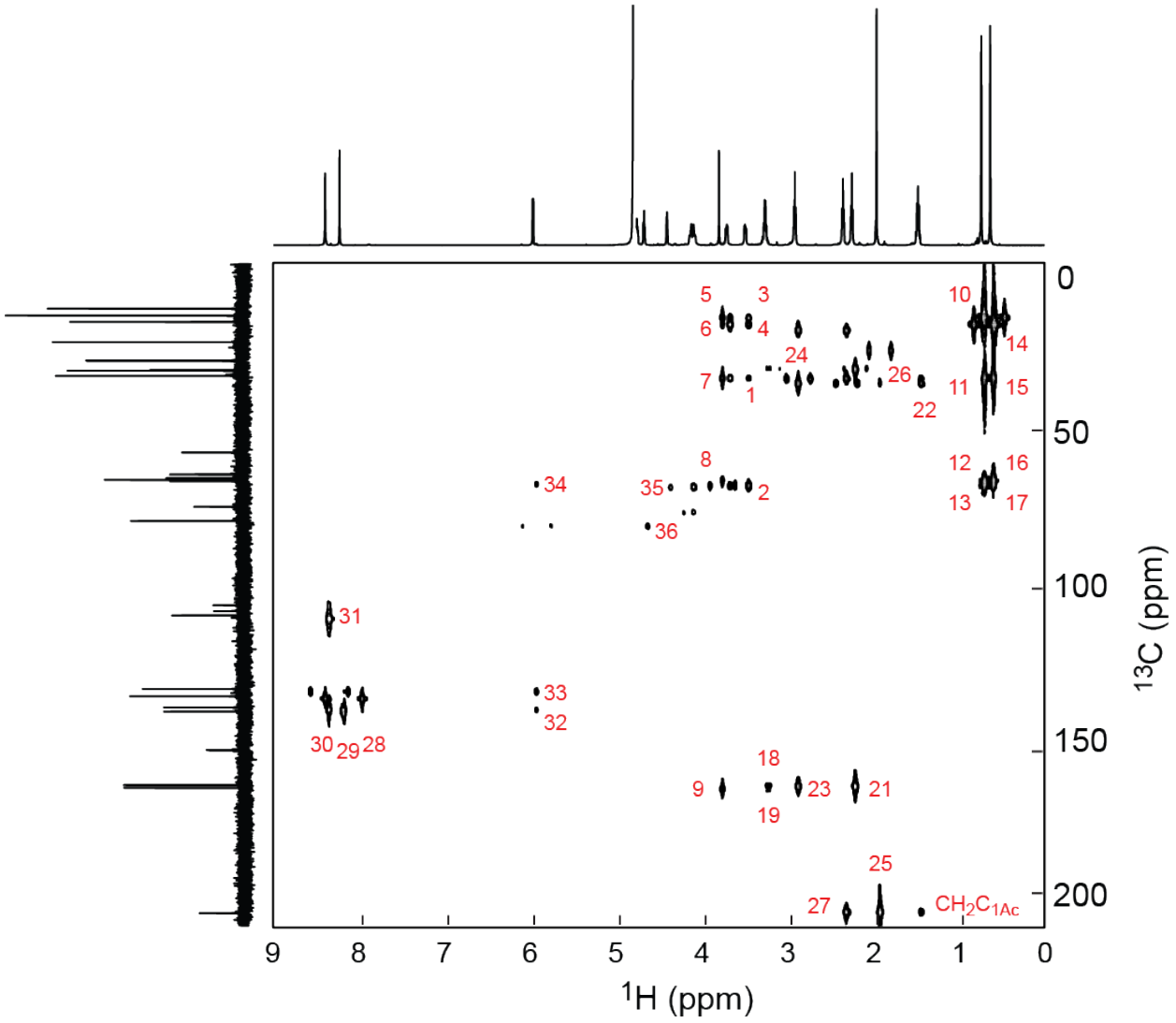


Figure 12. $^{1}H/^{13}C$ HMBC spectrum of *(S)*-2-fluoropropionyl-CoA in D_2O.

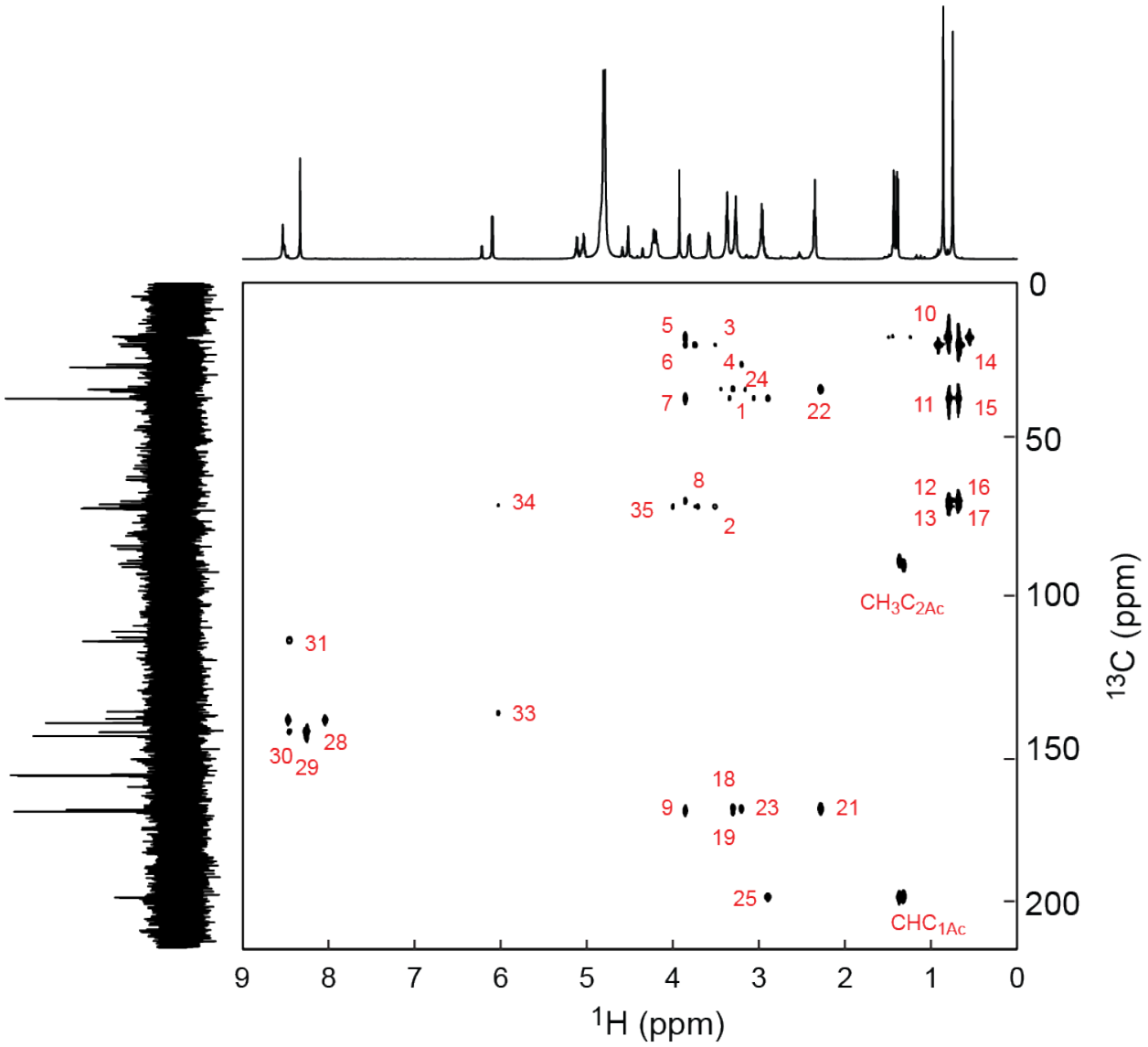
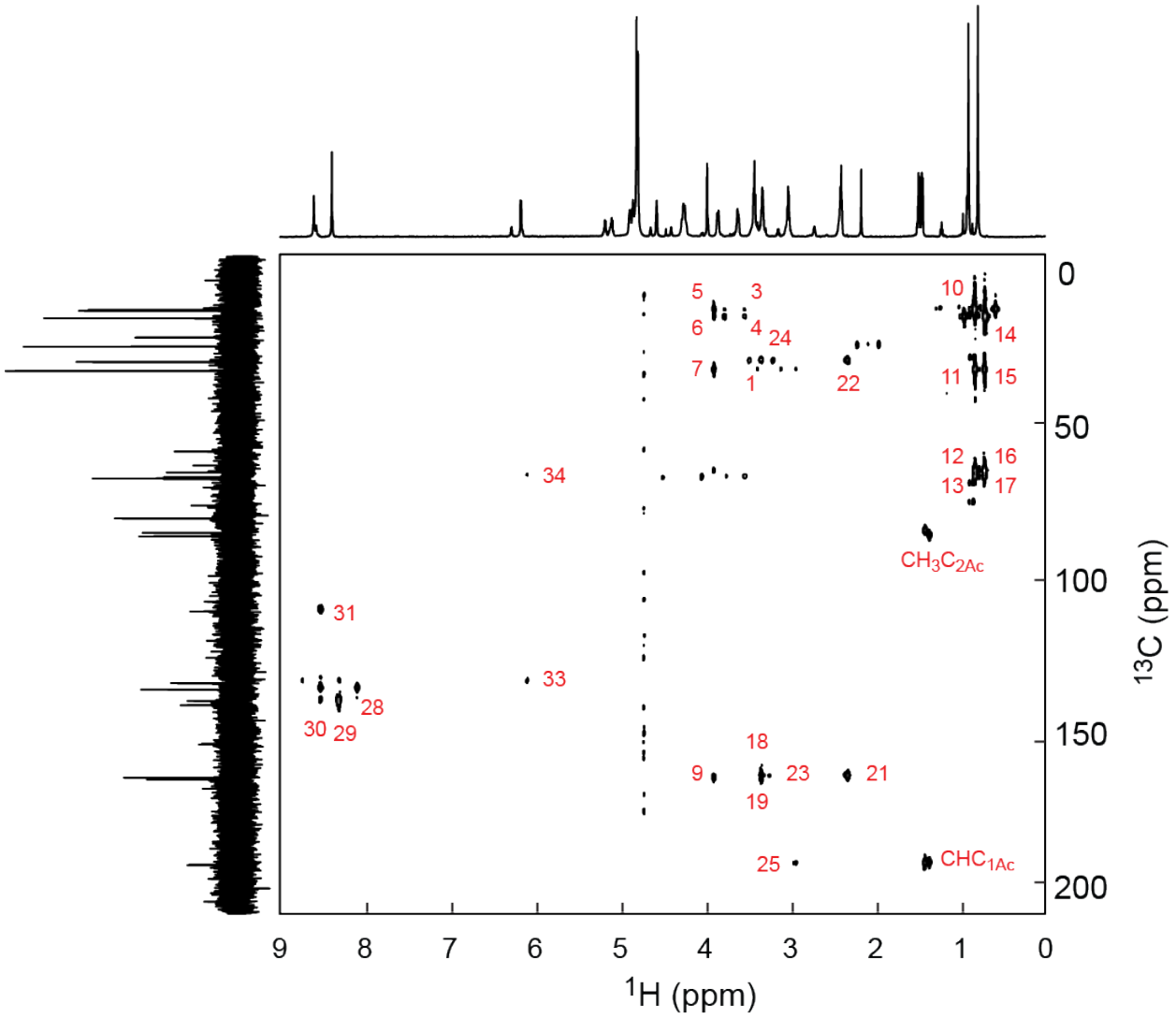


Figure 13. $^{1}H/^{13}C$ HMBC spectrum of *(R)*-2-fluoropropionyl-CoA in D_2O.



Appendix 4: *Protein purification gels*

Figure 1. Purification gel for His-tagged pantothenate kinase (PanK). Purification is described in section 4.2. 1, pre-induction; 2, post-induction; 3, soluble fraction; 4, streptomycin sulfate supernatant; 5, Ni-NTA eluate.

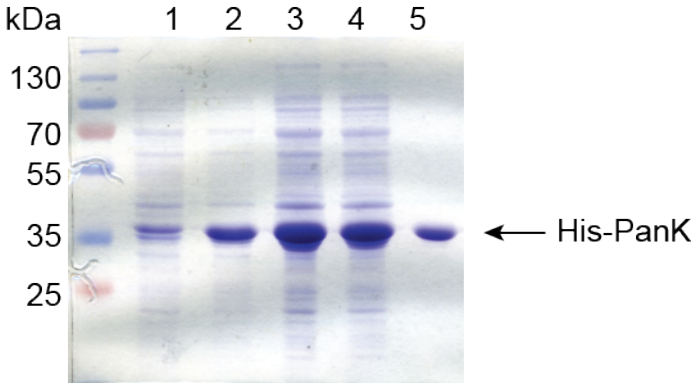


Figure 2. Purification gel for His-tagged phosphopantethiene adenylyltransferase (PPAT). Purification is described in section 4.2. 1, pre-induction; 2, post-induction; 3, soluble fraction; 4, streptomycin sulfate supernatant; 5, Ni-NTA eluate.

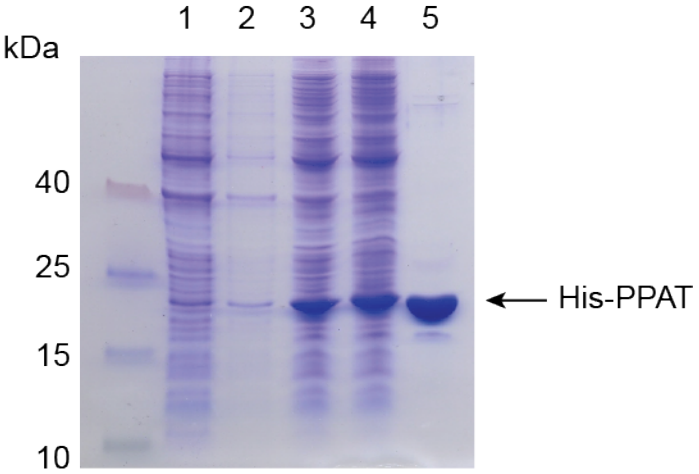


Figure 3. Purification gel for His-tagged dephospho-CoA kinase (DPCK). Purification is described in section 4.2. 1, pre-induction; 2, post-induction; 3, soluble fraction; 4, streptomycin sulfate supernatant; 5, Ni-NTA eluate.

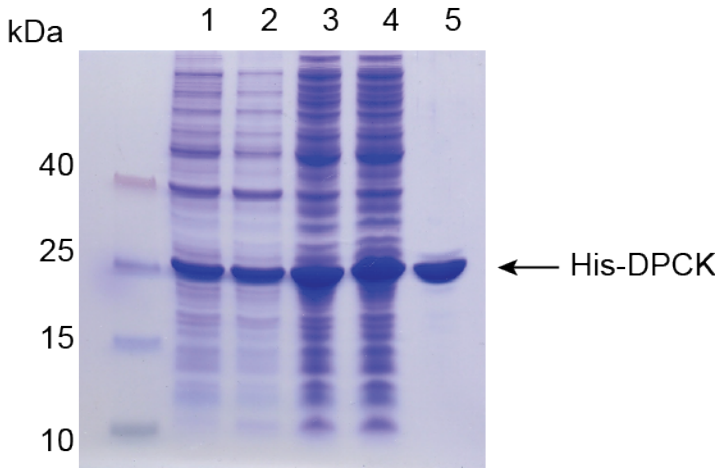


Figure 4. Purification gel for Strep-tagged taurine dioxygenase (TauD). 1, soluble fraction; 2, PEI supernatant; 3, Strep column flowthrough; 4, Strep column eluate.