Title
Exploring carbon-carbon bond formation in nature for the production of biorenewable fuels and chemicals

Author
Blaisse, Michael Blaisse

Publication Date
2016

Peer reviewed|Thesis/dissertation

Exploring carbon-carbon bond formation in nature for the production of biorenewable fuels and chemicals

by

Michael Robert Blaisse

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

Spring 2016

Exploring carbon-carbon bond formation in nature
for the production of biorenewable fuels and chemicals

by Michael Robert Blaisse

Abstract

Exploring carbon-carbon bond formation in nature
for the production of biorenewable fuels and chemicals

by

Michael Robert Blaisse

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Michelle C. Y. Chang, Chair

Microorganisms have enormous potential as catalysts for the conversion of renewable organic material into useful chemicals, such as fuels and polymers. But engineering metabolic pathways in microorganisms requires a greater understanding of the enzymes available to catalyze each reaction of a sequential biosynthetic scheme. Fundamental to the construction of organic molecules is the formation of carbon-carbon bonds. This dissertation describes our work to identify and characterize enzymes from non-canonical carbon metabolism pathways in nature. In particular, we focus on two fermentation pathways found in eukaryotic facultative anaerobes: branched acid fermentation in the roundworm *Ascaris suum* and wax ester fermentation in the single-celled algae *Euglena gracilis*. These pathways are based on thiolase enzymes, which catalyze carbon-carbon bond formation through Claisen condensation. Our investigation has led to a greater understanding of the factors that determine substrate and product selectivity in thiolases and thiolase-based pathways and has provided insight into the physiology of *E. gracilis*. We further demonstrate the use of *A. suum* enzymes to engineer pathways in bacteria for the production of α-methyl organic acids. Our results help to expand the synthetic capabilities of engineered microorganisms.

Table of Contents

Chapter 4: Wax fermentation in *Euglena gracilis*

Appendices

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of Abbreviations

ACP	acyl carrier protein
ADP	adenosine diphosphate
ATP	adenosine triphosphate
BME	β-mercaptoethanol
BSA	bovine serum albumin
BSTFA	N,O-bis(trimethylsilyl)trifluoroacetamide
Cb	carbenicillin
Cm	chloramphenicol
CoA	coenzyme A
DCE	dichloroethane
DMAP	4-dimethylaminopyridine
DMAPP	dimethylallyl pyrophosphate
DMSO	dimethylsulfoxide
dNTP	deoxynucleotide triphosphate
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray ionization-mass spetrometry
FAS	fatty acid synthase
FPLC	fast protein liquid chromatography
GC-MS	gas chromatography-mass spectrometry
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl β-D-1-thiogalactopyranoside
Km	kanamycin
LC-MS	liquid chromatography-mass spectrometry
MES	2-(N-morpholino)ethanesulfonic acid
MIP	mitochondrial intermediate peptidase
MPP	mitochondrial processing peptidase
MRM	multiple reaction monitoring
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate

NMR	nuclear magnetic resonance
PBS	phosphate buffered saline
PEG	polyethylene glycol
PHA	poly(hydroxyalkanoate)
PHB	poly(hydroxybutyrate)
PKS	polyketide synthase
PMSF	phenylmethane sulfonyl fluoride
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
TCA	trichloroacetic acid
TCEP	tris(2-carboxyethyl)phosphine
TEMED	N,N,N',N'-tetramethyl-ethane-1,2-diamine
TEV	tobacco etch virus
TMS	trimethylsilyl
Tris	tris(hydroxymethyl)aminomethane

Acknowledgments

The methods, data, observations, figures, etc. described in this dissertation are only a reflection of everything that the Ph.D. experience has entailed, and it would be impossible to distinguish between the contributions others have made toward the research results presented here from those toward my overall experience over the past six years.

Prof. Michelle Chang has of course been a big part of the story from the beginning, and I have been especially grateful for her time spent editing and offering feedback both on this work and other materials and presentations as I've come to the end of this journey. She has also managed to assemble a group of some of the most supportive, intelligent, generous labmates a grad student could ask for, and it has been a pleasure working with all of them. To risk singling out just one to mention, I particularly wish to acknowledge Dr. Matthew Davis, whose generosity and personal time spent helping the lab run ever more smoothly is possibly unmatched even among the M. Chang lab, which is saying a lot. Finally, Lissette Garcia has gone above and beyond in assisting me and the other members of the lab with administrative affairs, which has made our lives immensely easier.

There are few people for whom the description "This work would not have been possible without…" is a better fit than for my excellent undergraduate research assistant, Beverly Fu. Most Ph.D. students consider themselves lucky if the effort they spend training an undergrad gets returned in equal measure, but Beverly accomplished that in her first summer. It has been a great privelege working with her and I only hope I have been close to as good a mentor as she has been a student. Thanks so much to Effie Zhou for finding her! I have also had the opportunity to mentor a number of rotation or visiting students during my time in grad school. Thanks to Steven Strutt, Janice Chen, Lucas Harrington, and Dominik Brücher for all of their hard work. And of course, my protégés, Jase Gehring and Vivian Yu – it has been wonderful having them in the lab.

A number of purified proteins contributed by Dr. Jeffrey Hanson made the work presented here much easier to accomplish. In terms of the technical knowledge I have gained, I am indebted to Dr. Stefan Bauer, who taught me a great deal about analytic instrumentation early on in my Ph.D, and also to Dr. Mark Walker, who was very helpful in teaching me about RNAseq data processing. A big step forward in my project came from a helpful conversation with Tristan de Rond and his advice on chemical synthesis. A number of people have contributed to helping me learn X-ray protein crystallography, namely Dr. James Holton, Michael Lawson, and Dr. Ningkun Wang.

I would not have been able to accomplish the work presented in this dissertation, or write about it, if I had starved to death, and a perhaps disconcertingly large number of people need to be acknowledged for keeping me fed during the most stressful times. Thank you Vivian, Beatriz, and Oliver. And a special thanks to Hiro and Ringo, for keeping my blood pressure down this past month.

Finally, thank you to all of my friends and family for your love and support over the past six years. I know that many of you have shared my ups and downs as I've experienced them, and this success is just as much yours as mine.

Chapter 1: *Introduction*

1.1. Fermentation as a replacement for chemical synthesis

The rise of industrial chemical synthesis over the past century, and in particular the petrochemical industry, has transformed our society by offering access to thousands of materials, pharmaceuticals, and specialty compounds from organic commodity chemicals derived from oil and gas. Compared to traditional chemical synthesis, industrial fermentation using living organisms has unique advantages that position it to potentially replace or supplement chemical manufacture [1, 2], access renewable and/or recalcitrant feedstocks [3, 4], or access new compounds and synthetic precursors, such as chiral building blocks [5] and amino acids, [6] which may be difficult to make synthetically. Unlike chemical synthesis, in which intermediates must be synthesized stepwise in separate reactors, isolated, and transported to a new vessel, living organisms have evolved to carry out hundreds of reactions in a single reaction vessel, the cell. Finally, a cellular biocatalyst is self-renewing, evolvable, and highly amenable to the design-build-test cycle necessary for industrial optimization and scaling [7].

Aside from the historic use of fermented foods and beverages in a number of cultures around the world, industrial fermentation arose in the 20th century as a way of producing commodity chemicals during shortages of standard sources. Prominent examples include the production of chemical solvents [8], citric acid [9], and glutamic acid [10] by culturing *Clostridium acetobutylicum*, *Aspergillus niger*, and *Corynebacterium glutamicum*, respectively. Over time, concerns over climate change, energy security, and petroleum fuel costs have stimulated interested in the production of fuels [11, 12], monomers [13, 14], and other bulk chemicals [15] from renewable sources. Most microbes have evolved to efficiently consume simple sugars such as glucose and convert these raw materials into a variety of chemical compounds. The synthetic potential of biology can be realized now more than ever due to technical advances in our ability to construct and manipulate DNA *in vitro* and *in vivo* [16, 17]. Through expression of heterologous enzymes, whole metabolic pathways can be reconstituted in a genetically engineered host microbe, including novel chimeric pathways consisting of enzymes from various origins [18].

Compared to the production of high value compounds like pharmaceuticals, production of fuels and commodity chemicals through fermentation offers challenging demands for both titers and carbon yields [19]. The relatively high cost of glucose compared to petroleum requires that the carbon atoms contained in sugar feedstocks are channeled into desired product molecules with near-quantitative yields at highly productive rates. To achieve this aim, it is important to understand the interrelated nature of metabolic networks that distinguishes chemical synthesis in living systems from that of non-living systems. As illustrated in *Figure 1.1*, carbon metabolism, energy usage, and reducing power are intricately connected in the cell. While both oxidation and reduction reactions are ubiquitous, these processes are actually decoupled in metabolism through the use of redox cofactors. A common example is the nicotinamide cofactor NAD^+ and its phosphorylated derivative $NADP^+$, which are capable of "storing" reducing equivalents in the form of dihydropyridine moieties for use in separate reduction reactions. In most cases, these reducing equivalents are derived from the oxidation of a portion of the supplemented carbon. Metabolism of glucose through the glycolysis pathway produces pyruvic acid, NADH, and a small amount of energy in the form of ATP through substrate-level phosphorylation. In the absence of oxygen, reduced NADH must be re-oxidized for glucose consumption to continue. This cofactor recycling is typically achieved using the pyruvic acid byproduct as an oxidizing agent, resulting in the accumulation of lactic acid, as in bacteria or mammalian muscle tissue, or

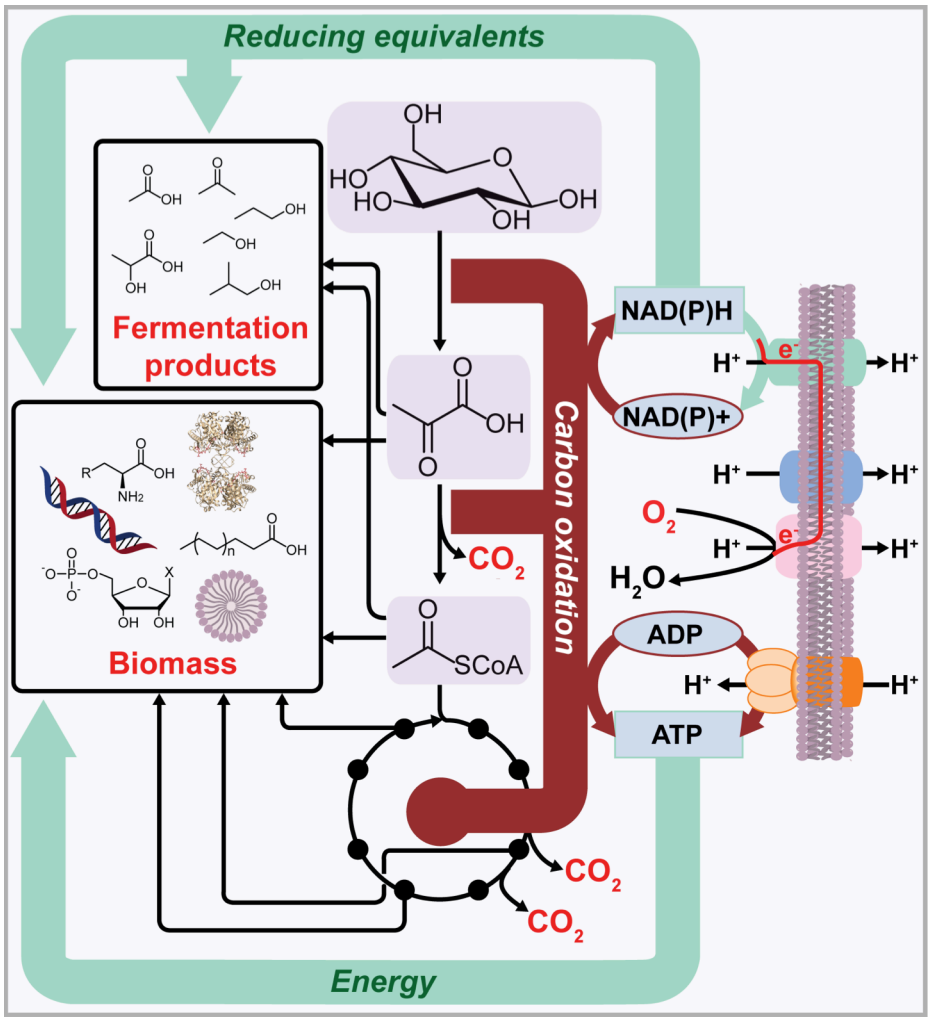


Figure 1.1. *Carbon, energy, and redox metabolism are interconnected in living systems. Oxidation and reduction processes are decoupled. Oxidation of carbonaceous material, such as glucose or similar sugars, forms reducing equivalents in the form of reduced nicotinamide (NADH, NADPH) and other cofactors and provides ATP through substrate-level phosphorylation. Energy, reducing equivalents, and organic material are required for production of cellular components, such as lipids. In the absence of inorganic oxidants, organic material is also used to recycle redox cofactors, resulting in accumulation of fermentation products. In the presence of oxygen or other inorganic oxidants, reducing equivalents can be oxidized through an electron transport chain, forming a trans-membrane potential that drives further energy production.*

through decarboxylative reduction to form ethanol, as in *Saccharomyces cerevisiae*. Because accumulation of organic compounds is needed as an electron sink for the organism in the absence of oxygen, anaerobic or near-anaerobic growth may be the optimal condition to achieve high-yield production of most commodity chemicals [20]. On the other hand, certain synthetic pathways may require energy investment in the form of ATP. In the presence of oxygen, reducing equivalents can be oxidized by a series of membrane-associated enzymes to produce a large amount of ATP through oxidative phosphorylation. While this process produces much more energy per glucose consumed, higher energy load may facilitate the accumulation of biomass in the form of cell structure components (*e.g.* proteins, membrane lipids, nucleic acids) rather than target chemicals. Sophisticated regulatory networks have evolved over millions of years to control the cell's metabolism according to its environment [21]. There is thus a tension between our need for moderate cell growth to sustain productive bio-catalysis and the cell's evolutionary imperative to maximize growth and survival. To summarize, carbon yield is inseparable from energy usage, redox balance, and cell growth in living organisms, and controlling synthetic chemistry in this context is a major technical problem.

The challenges outlined above of using synthetic biology for commodity chemical production are capable of being addressed through advances in enzyme engineering, pathway design, and whole organism optimization [22, 23]. Critical to these advancements are also synergistic solutions, with chemical and chemo-enzymatic processes used in conjunction with fermentation. Thoughtful judgment must be made as to what fuels and chemicals will be most promising in the short term as industrial fermentation targets.

1.2. The evolutionary origins of biochemical pathways differentiate them from synthetic schemes

Designing biosynthetic pathways is similar in many ways to the process of designing a chemical synthesis scheme. Over the past two centuries, chemists have discovered millions of reactions, in which particular functional groups and catalysts are found to produce particular products under a set of conditions [24]. With an extensive knowledge of the chemical transformations at one's disposal, an industrial chemist can plan a series of appropriate reactions to produce a desired chemical in high yield. Likewise, biochemists have discovered, characterized, and annotated hundreds of enzyme catalyzed reactions over the past century and are increasingly able to infer the activity of uncharacterized enzymes based on sequence data and to expand and develop new reactions through enzyme engineering [25]. These enzymes can be expressed in a genetically modified host to assemble a sequential scheme for the production of a desired compound in high yield.

Biological synthesis, however, differs from chemical synthesis in important ways. One major difference is seen at the level of synthetic scheme design: the metabolic pathway. Natural pathways typically feed off of central pathways (*e.g.* glycolysis, pentose phosphate pathway, Kreb's cycle) that assimilate glucose and convert it into ubiquitous two or three carbon building blocks that give rise to the diversity of metabolites in the cell. In order to compete for carbon with the rest of metabolism, a synthetic pathway needs to support high flux, with turnover through the pathway at a similar magnitude to the rate of glucose assimilation. High flux pathways in nature display a number of features that facilitate this, such as the use of an exergonic, irreversible step, which commits carbon atoms to continue through that pathway rather than another [26]. In contrast to chemical synthesis, where each reaction is carried out

independently and must therefore be individually high-yielding, a reaction in a biosynthetic pathway might have only moderately favorable or even moderately unfavorable thermodynamics, so long as there is one or more committed step and formation of the final product is favorable. Even if it is irreversible, a high flux pathway must also support the energetic and redox demands of the cell, or else proper cofactors and precursors will become limiting as carbon is diverted through pathways needed to maintain growth and survival [27].

A second important difference exists at the level of enzymes, which provide the repertoire of available biological reactions. The diversity of reactions developed by chemists has arisen from exploring a wide breadth of chemical functionality, including the use of rare elements, and employing a range of reaction conditions, such as varying solvent systems, temperatures, pH, and pressures. Through the judicious use of protecting groups, these reaction methodologies, once developed and optimized, are for the most part generally applicable towards molecules with the requisite features. In contrast, although a range of specialty reactions exist in nature, such as the installation of specialized functional groups [28] or unusual molecular assemblies [29, 30], these reactions may not meet the desired energetic and flux design demands, function in low-flux pathways in nature, and are not generalizable. Instead, the majority of metabolites arise from a surprisingly small number of catalytic reactions [31]. This redundancy arises from the evolutionary mechanism by which nature has solved problems in chemical synthesis [25]. Chemical function is re-appropriated for different substrates through mutations in enzymes leading to substrate promiscuity, followed by further mutations that optimize specificity for the new substrate. While this process can occasionally result in new reactions [32], this new chemistry is mechanistically related and limited by the narrow range of intracellular conditions – pH (5-8), solvent (water), and temperatures (those supporting liquid water) – conserved throughout life. In contrast to the diversity of carbon-carbon bond forming reactions used by synthetic chemists, for example, just a few major reactions are used by nature to assemble building blocks in the cell: aldol and Claisen condensations and prenyl transfer in isoprenoids.

1.3. Natural stategies for small molecule assembly

Considering the differences we have noted between chemical and biological synthesis, we believe that designing novel metabolic pathways will require a deeper understanding of the biosynthetic strategies available in nature. While a host of small molecules, many consisting of four or fewer carbons, can and have been made using engineered pathways that start from one or more major cellular metabolites [33], here we highlight common iterative pathways that account for the construction of most larger hydrocarbon structures in nature from central metabolic building blocks, as well as downstream modifications for tailoring molecular targets. These pathways could in principle be adapted for the biosynthesis of a variety of non-natural commodity chemicals. We discuss the advantages and challenges of each pathway in terms of design features mentioned above, including theoretical carbon yield, energetic and redox balance requirements, flux control elements, enzyme selectivity, and capacity for enzyme discovery or engineering.

Chain elongation with malonate derivatives: decarboxylative Claisen-condensation. Almost all membrane lipids in most organisms are synthesized by the large enzyme complex fatty acid synthase (FAS), which connects acetate units consecutively through decarboxylation-driven Claisen condensation of thioesters (*Figure 1.2*) [26]. In this process, acetyl-CoA, an acetate thioester of the ubiquitous thiol coenzyme A, is carboxylated through a biotin-dependent,

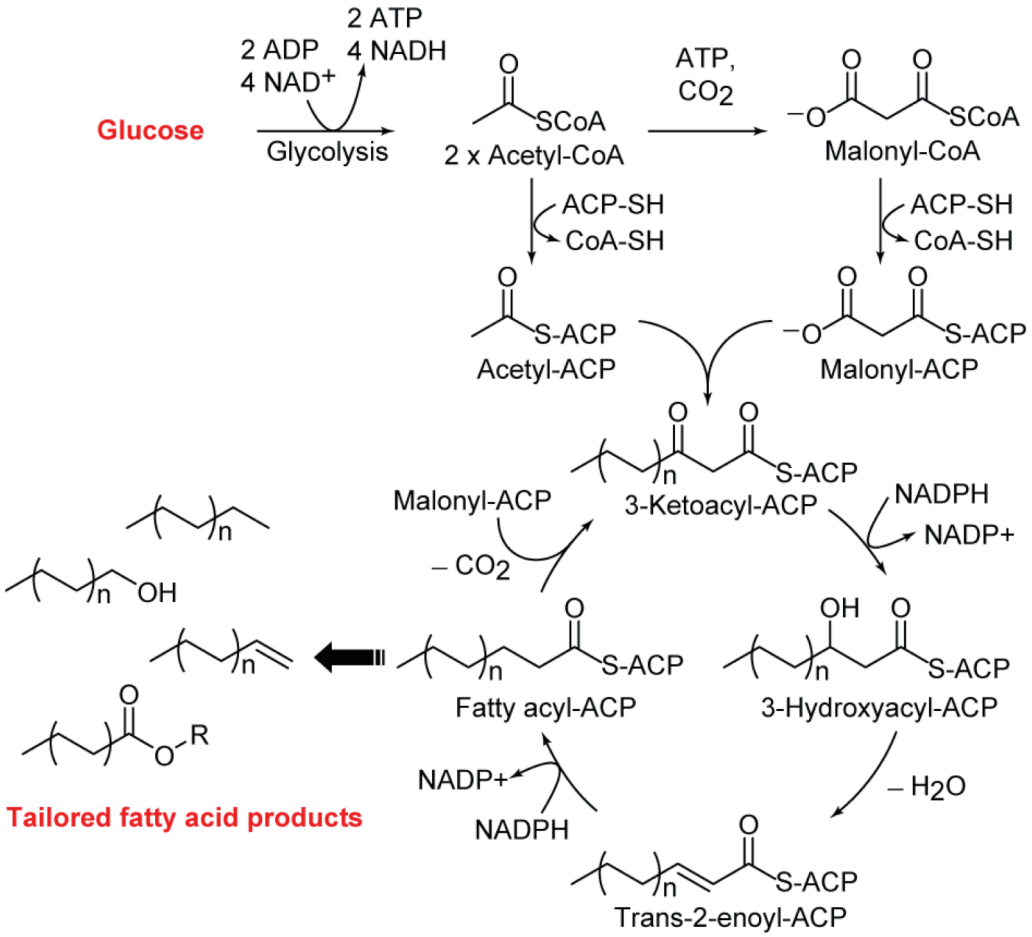


Figure 1.2. *Chain elongation by decarboxylative Claisen condensation of malonic acid derivates, as observed in fatty acid synthesis. Condensation, reduction, dehydration, and further reduction result in extension of the primary carbon chain by two atoms.*

ATP-driven process to form the extender unit malonyl-CoA. The malonyl moiety is transferred to a prosthetic group of the acyl carrier protein (ACP). Claisen condensation with a saturated acyl-ACP is catalyzed by a ketosynthase enzyme, releasing CO_2 and forming a 3-oxoacyl-ACP species, which is subsequently reduced with NADPH by a ketoreductase. Elimination of water by a dehydratase enzyme forms an enoyl-ACP, which is further reduced by an enoyl-ACP reductase to form a saturated acyl-ACP species whose chain length is now 2 carbons longer. This process continues until chain lengths of around 16 or 18 carbons are reached, depending on the organism, at which point the fatty acid is incorporated into lipids or triacylglycerol fat deposits through various organism-specific mechanisms. In animals and fungi, just one or two large polypeptide chains contain all of the enzymatic functions involved in fatty acid synthesis [34].

As a native high-flux pathway, fatty acid synthesis offers an attractive route towards lipid-based commodity chemicals [35, 36]. A major advantage is the decarboxylation of malonyl-CoA, which provides an initial irreversible step to commit carbon flux through the pathway. Some products that could be made include free fatty acids, which can be esterified either chemically or *in vivo* to make biodiesel or oxidatively decarboxylated to form terminal alkenes [37]. Fatty acyl-ACPs can also be reduced to fatty alcohols or aldehydes, the latter which can be deformylated to produce alkanes [38-41]. However, decoupling a pathway involved in lipid homeostasis from the actual growth and maintenance of the cell can be a significant challenge [42, 43]. As mentioned previously, producing commodity chemicals through a pathway that is intricately connected to cell growth, especially one that requires investment of ATP, poses unique challenges. Nevertheless, strategies for engineering flux control exist. Major elements to consider are the formation of malonyl-CoA by acetyl-CoA carboxylation, which is tightly regulated and inhibited through negative feedback [44]; the availability of the acetyl-CoA substrate; and the various termination reactions required to convert acyl-ACP thioesters to chemicals of interest. A common approach has been to overexpress a thioesterase enzyme, which hydrolyzes acyl-ACP to release free fatty acids [45]. This strategy has the dual result of lifting negative feedback by acyl thioester products and providing another irreversible step to pull flux through the pathway. The disadvantage is that certain molecular targets, such as alkanes, esters, and alcohols, require the energy-intensive reactivation of free fatty acids to acyl-CoA thioester, although methods for circumventing this reactivation have been demonstrated using various decarboxylating oxygenases to form odd-chain hydrocarbon products [46, 47]. To avoid futile cycles that waste energy, artificial regulatory networks have been engineered to modulate fatty acid ethyl ester (FAEE) production in *Escherichia coli* [48]. Such artificial networks may prove useful in solving flux control problems in other contexts as well. Other pathway bottlenecks, such as transcription-based regulation of enzyme expression, can be lifted by the overexpression of proteins under constitutive promoters [36].

An ongoing challenge is to increase the cellular acetyl-CoA pool, which in many organisms is regulated post-translationally in ways that are sensitive to redox environment in the cell and tied to cell growth, such that even overexpression of enzymes may reach a ceiling in the amount of flux improvement that can be achieved [49]. A few natural examples exist of anaerobic fermentation pathways based on acetyl-CoA, namely the fermentation of butanol by *C. acetobutylicum* [50] and the wax ester fermentation reported in the unicellular green algae *Euglena gracilis* [51]. In most organisms, however, acetyl-CoA production is much more closely associated with growth regimes involving biomass accumulation. The popular industrial microbe *S. cerevisiae* has further difficulties related to organelle compartmentalization of acetyl-CoA and maintains a very low acetyl-CoA pool for fatty acid synthesis in the cytosol [52, 53]. Approaches

to engineer a citrate lyase shunt into *S. cerevisiae* in resemblance to oil-accumulating yeasts have shown success [36], but further research in this area could significantly improve the prospects of synthetic pathways for all kinds of novel chemicals [54].

An important question is whether FAS could be adapted as a generalized chain extension strategy for producing novel commodity chemicals. This might involve engineering the enzymatic domains of the complex to accept alternative extender units, such as methylmalonyl-CoA or ethylmalonyl-CoA, or modulating chain extension to produce more medium length (8 or 10 carbon) structures. Synthetic regulatory elements have been demonstrated as a proof of concept to modulate chain lengths [55], and alternative FabH enzymes are known in various bacteria to initiate chain extension with starter units other than acetyl-CoA, including branched acyl-CoA compounds [56-58]. Beyond this, however, FAS engineering may be difficult since the complex has evolved throughout all the kingdoms of life to produce a rather homogenous pool of molecules. Polyketide synthases (PKSs) are a related group of multi-module biosynthetic complexes that are dedicated to producing particular secondary metabolites using the same malonate-condensing cycle as FAS [59]. PKSs site-specifically incorporate a variety of starter units and substituted malonate extender units and forego certain reduction or dehydration steps at key chain lengths to form specific natural products. In principle, then, PKSs might be highly adaptable, and characterized enzymatic modules from various PKSs could be mixed and matched to synthesize a range of custom chemicals [60]. This has not been straightforward to achieve in practice, however, as individual enzyme domains or modules may not function in isolation or when fused to non-native modules, and the principles that govern extender unit selection and enzyme processivity are complex and still being actively investigated. The low catalytic turnover of many PKSs might also make them less suitable for commodity chemical production, which requires high pathway flux to be feasible. On the other hand, the engineering of PKSs for commodity chemical production has been demonstrated for adipic acid [61] and for polyene hydrocarbons [62], the latter which can be hydrogenated to make alkane fuels. Further work in the area of PKS engineering could therefore lead to promising routes for custom commodity chemical production.

We conclude by noting that some malonate derivatives, such as methylmalonyl-CoA and ethylmalonyl-CoA, can be synthesized in the cell primarily through pathways besides ATP-dependent carboxylation [63, 64]. Therefore cellular energy investment is not necessarily required for biosynthetic approaches using decarboxylative Claisen condensation. Furthermore, even malonyl-CoA production could in principle be engineered using non-native pathways that conserve energy. The discovery of the 3-hydroxypropionate cycle in certain non-sulfur photosynthetic bacteria and thermophilic archae, for example, offers potentially useful enzymes towards this end [65].

Claisen condensation using thiolase-based pathways. Thiolase enzymes, which catalyze non-decarboxylative Claisen-condensation of acyl-CoA compounds, present an alternative to FAS and PKSs with different advantages and challenges while forming the same products (*Figure 1.3*). The main difference is that thiolases do not use malonic acid derivatives for chain extension but rather deprotonate acetyl-CoA directly to form an enolate nucleophile. Since CO_2 release does not accompany carbon-carbon bond formation, the Claisen-condensation is reversible and in fact considerably unfavorable energetically [66]. Indeed, thiolases most commonly function physiologically in the degradation of lipids and other hydrocarbon molecules

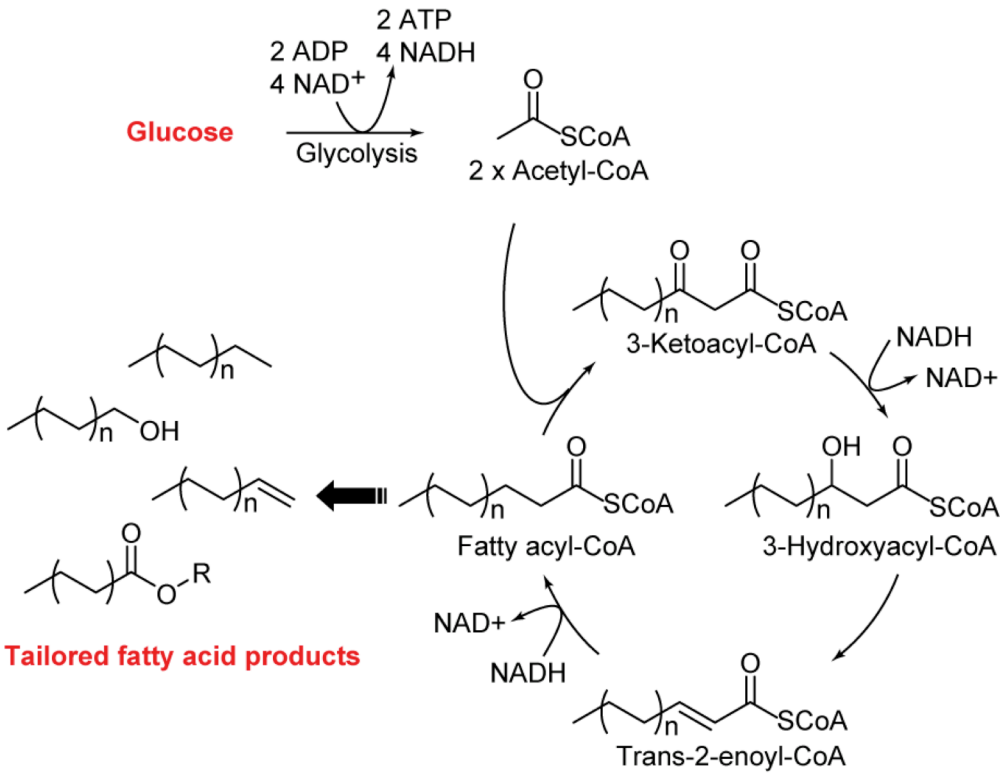


Figure 1.3. *Claisen condensation using thiolase-based pathways. This process is the reverse of the β-oxidation pathway and is mechanistically similar to fatty acid synthesis, but does not use malonic acid derivatives*

(β-oxidation), rather than biosynthesis. Prominent examples of biosynthetic thiolases exist, however, in both natural and engineered high flux pathways, including the synthesis of poly(3-hydroxybutyrate) (PHB) biopolymer in numerous bacterial species [67, 68], synthesis of *n*-butanol [69-71], and the initial step of isoprenoid production in animals and fungi [72]. In many of these examples, an irreversible enzymatic step increases flux and commits carbon to the pathway. PHB synthesis, for example, results in an insoluble plastic that is sequestered through phase separation, and synthetic butanol production includes the irreversible reduction of crotonyl-CoA with NADH [69, 71]. In another example, near-quantitative yields of butanol could be produced through a pathway consisting of butanol degradation operating in reverse [70]. Although this latter case did not contain any irreversible enzymatic steps, the extremely favorable energetics of enoyl-CoA reduction combined with the overall reducing power in the cell were evidently sufficient to ensure high titers.

A challenge of using thiolase-based elongation as a generalized paradigm for commodity chemical production is that the feasibility of sustaining high flux with multiple iterations is unclear. Although effective production of C_6-C_{10} alcohols and acids has been demonstrated using a thiolase-based pathway [70, 73], butanol titers in the literature have generally exceeded the best reported titers of longer chain products. It is possible that the unfavorable energetics of Claisen condensation are compounded after each elongation. Butyryl-CoA synthesis, for example, proceeds *via* the condensation of two acetyl-CoA, whose intracellular concentration is ~0.5-1 mM in the cell. But for subsequent condensations, the intracellular concentration of each successive saturated acyl-CoA is likely even lower than its previous intermediate. On the other hand, *in silico* system-level assessment of this pathway has shown that, in theory, favorable NADH-dependent enoyl-CoA reduction, combined with elimination of competing fermentation pathways and other genetic modifications depending on the final product of interest, should be sufficient to drive flux through multiple iterations of elongation [74]. At least one example exists in nature of a high-flux thiolase-based fermentation pathway to form longer chain products; *E. gracilis* is known to accumulate fatty wax monoesters under anaerobic conditions *via* the ATP-independent condensation of acetyl-CoA [51, 75-77].

Thiolase based pathways also share challenges mentioned for FAS and PKS, such as uncoupling acetyl-CoA production from biomass accumulation, as well as the question of whether the elongation strategy can be generalized to make novel compounds. Biosynthetic thiolases that condense two acetyl-CoA molecules are ubiquitous and well characterized [78], but other acyl-CoA substrates, particularly nucleophiles besides acetyl-CoA that may introduce substitutions at the α-position, are less represented. Recent work with thiolases and ketoreductases is expanding the utility and substrate scope of these enzymes [5, 79, 80]. Since thiolase pathways operate orthogonally to the ACP pool and do not require ATP, they might offer greater flexibility for carbon-carbon bond formation and allow higher carbon yields through anaerobic fermentation.

Iterative extension of 2-oxoacids. Another methodology that nature has developed for building up biomolecules is distinct from the Claisen-based pathways described above and involves the extension of 2-oxoacids (*Figure 1.4*) [81]. This four step pathway is conserved in a number of metabolic contexts, such as the Kreb's cycle; the biosynthesis of certain amino acids, including threonine, leucine, isoleucine, and lysine; and the synthesis of the archael cofactor, coenzyme B [82]. In this process, a 2-oxoacid undergoes aldol condensation with acetyl-CoA through a mechanism that includes hydrolysis of the thioester, resulting in a 3-

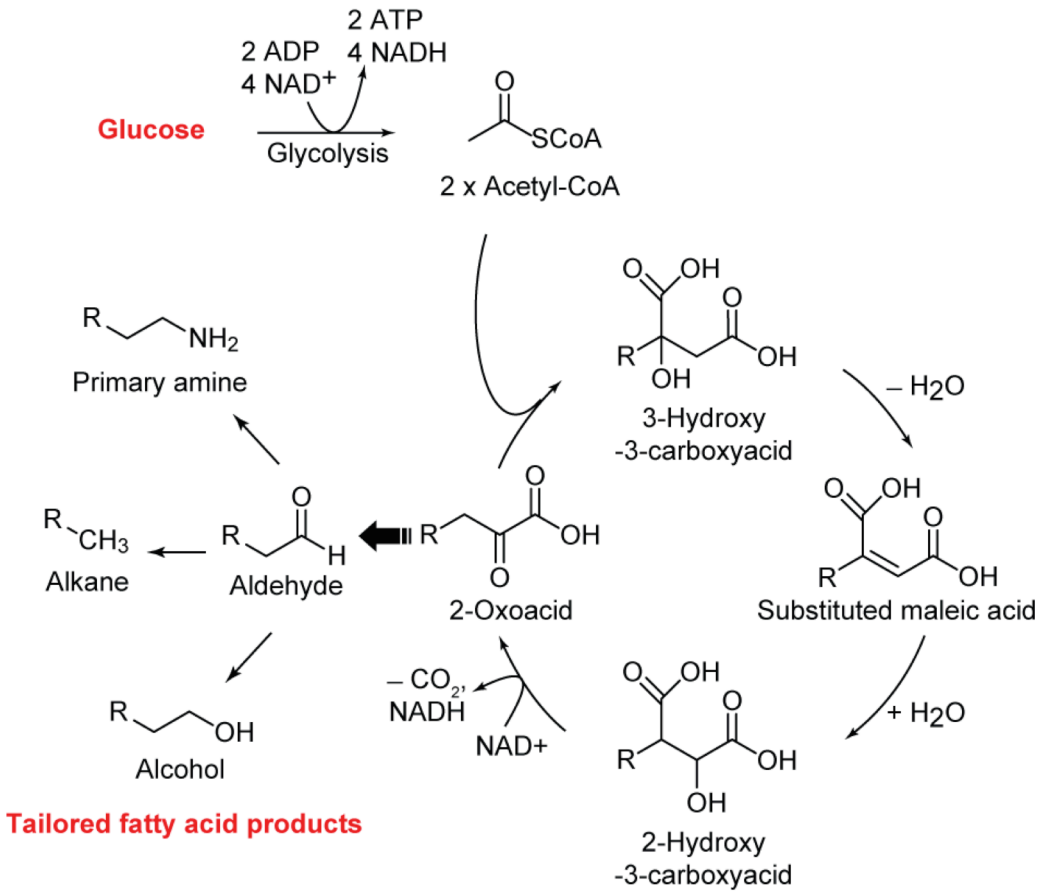


Figure 1.4. *Aldol condensation with 2-oxoacids. The resulting 3-hydroxy-3-carboxyacid is isomerized and oxidized to form a new 2-oxoacid extended by one carbon.*

hydroxy-3-carboxyacid, which is dehydrated and rehydrated, sometimes by a single enzyme, to form a 2-hydroxy-3-carboxyacid. Finally, oxidation-decarboxylation results in a 2-oxoacid that is one carbon longer than before.

The main advantages of this approach are the favorable energetics that facilitate high pathway flux and the utility of the 2-oxoacid motif for downstream molecular tailoring. Since thioester hydrolysis accompanies the aldol condensation, carbon-carbon bond formation is irreversible. The oxidation-decarboxylation that completes the cycle is also favorable. The resulting 2-oxoacids serve as a useful intermediate to access chemicals and fuels of interest. Irreversible decarboxylation of 2-oxoacids forms aldehydes, which are precursors to alcohols, alkanes, saturated acids, or terminal amines [83, 84].

As a general biosynthesis strategy, this method allows only the linear elongation of 2-oxoacid starter units. Because the α-carbon must be a quaternary carbon at the end of each elongation cycle, there is no opportunity for introducing extenders other than acetyl-CoA. The feasibility of expanding the reaction scope to alternative starter units has been illustrated and discussed [82]. Another drawback is the low carbon yield inherent in the pathway – only one third of the carbon atoms in each glucose molecule can be incorporated into product. This ratio might be improved, however, if carbon can be recaptured using specialty pathways or engineered strains [85].

Prenyl Transfer. Isoprenoids compromise the largest class of natural products and have incredible potential as commodity chemicals [86]. Their biosynthesis is wholly different from the pathways described above, because carbon-carbon bonds are formed not through aldol or Claisen-condensation but by stepwise electrophilic addition of prenyl cations to the double bond of a successive isoprenyl group (*Figure 1.5*). The common extender unit is dimethylallyl pyrophosphate (DMAPP), which is synthesized in the cell through one of two different pathways – the mevalonate pathway [87] or the deoxyxylulose-5-phosphate (DXP) pathway [88]. While the stoichiometry of the DXP pathway allows a higher stoichiometric yield of carbon (83%) compared to the mevalonate pathway (56%), it also requires a net investment of energy by the cell.

Carbon-carbon bond formation through prenyl transfer is not a generalizable synthetic scheme in the same sense as other elongation cycles described here, but the spectrum of compounds and hydrocarbon architectures that could be accessed by known examples of this natural product class is vast. One challenge of high flux isoprenoid biosynthesis has been integrating the upstream DMAPP synthesis pathway with downstream target-specific biosynthesis enzymes, which may be rate-limiting or require additional redox regeneration systems that are difficult to express heterologously. Although one strategy is to focus on augmenting DMAPP biosynthesis as much as possible to increase precursor supply for downstream steps, this approach may result in the accumulation of metabolites that are toxic to the cell [89]. Synthetic biology tools have therefore been an important part of engineering isoprenoid biosynthesis for chemical production [90, 91].

1.4. Pathway assembly and optimization

The biosynthetic pathways reviewed above are by no means intended as an exhaustive inventory of synthesis reactions available in biology, but rather an overview of some of the most common frameworks for molecular assembly encountered in metabolism that could provide

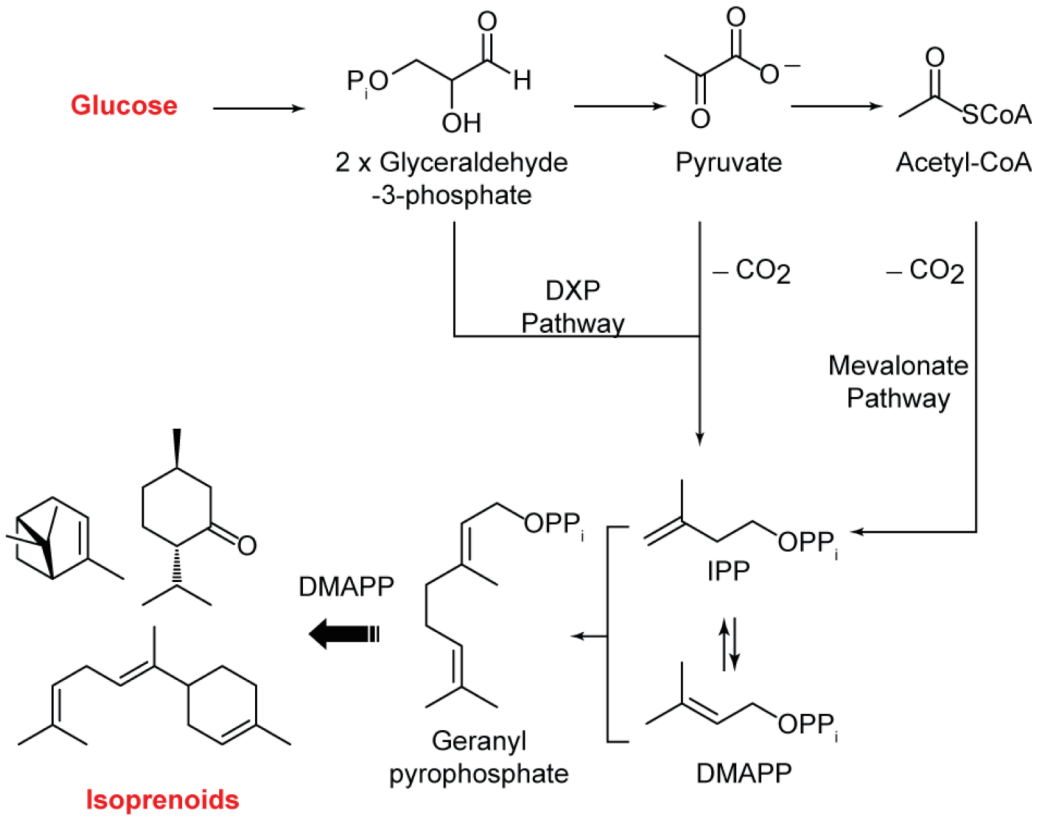


Figure 1.5. *Prenyl transfer, based on electrophilic addition to alkenes, is the basis for the biosynthesis of isoprenoids, a large class of natural products.*

general strategies for thinking about artificial pathway design. Another common framework is decarboxylative transketolation of pyruvate, which is observed in a number of branched amino acid biosynthesis pathways [83]. The Shikimate pathway offers a useful platform for the production of various aromatic commodity chemicals in living cells [92].

With a repertoire of biochemical reactions in hand, a metabolic pathway can be designed for the production of a desired target molecule, first by identifying the sequences of chemical transformations that could lead to the target beginning with some central metabolite(s), and then by considering the flux control and thermodynamic features of each candidate pathway in order to select the best prospects. Because the design space of biochemical reactions and precursor metabolites is more confined in biology compared to organic synthesis, automated computational methods for carrying out this analysis have found wider use in synthetic biology than in synthetic chemistry [93, 94], although many proof of concept examples are awaiting experimental confirmation. These automated methods can identify chimeric pathways that may not be immediately obvious and fill in gaps with enzymes from unrelated pathways or with enzymes whose function has not been verified but is anticipated based on partially characterized pathways, identification of catalytic amino acid sequence motifs, or protein structure-based modeling [95, 96].

Caution must be exercised when employing enzymes whose activity has not been experimentally demonstrated with the intended substrate in the synthetic pathway. Enzymatic substrate selectivity reflects the evolutionary function for that enzyme in the context of its native organism, and enzyme activity can vary substantially between two substrates that are even slightly different, such as acetyl-CoA vs. propionyl-CoA. On the other hand, enzymes can be "promiscuous", lacking selectivity if flexibility towards multiple substrates was more conducive to fitness in the species of origin [97]. Poor or promiscuous catalytic activity need not necessarily exclude a particular enzyme as a candidate for use in a high-flux synthetic pathway. If the enzymatic step is not irreversible or not a flux control point in the pathway, then enzyme promiscuity may not be a problem. Low catalytic efficiency could be also made up for by higher overexpression of that enzyme. Ultimately, pathway optimization will be desired to overcome bottlenecks, balance the expression of each pathway enzyme, and minimize the metabolic burden on the cell and the buildup of intermediates or side products. Strategies for pathway optimization have been extensively studied and reviewed [26, 98-101]. Nevertheless, the usefulness of poorly-suited enzymes in pathways for commodity chemical manufacture should not be overexaggerated. The identification of a new, more efficient or better suited enzyme is often the most reliable way to fix a pathway bottleneck.

In summary, the ability to construct biochemical pathways through a retro-synthetic approach analogous to chemical synthesis remains an as-yet-unachieved ideal, limited by the available biochemical toolbox of accessible enzymatic functions. Much progress has been made in the field of enzyme engineering, both computational methods [102, 103] and the development of new screening strategies that may be relevant to general metabolic enzymes [104], which tend to lack colorimetric assays or other sensitive, high-throughput screening methods. Altering the substrate specificity of an enzyme to confer the efficiency and selectivity that are essential for a highly productive pathway is still a difficult and resource-intensive endeavor, however, and we stress the need to choose enzyme classes in which investment would prove most fruitful. The enzymes involved in the elongation cycles we have highlighted in this section are excellent candidates for further study. The development of libraries of these enzymes (*e.g.* aldol and

Claisen condensation catalysts, ketoreductases, enoyl reductases) could allow generalizable strategies for commodity chemical pathway design. Product tailoring enzymes, such as thioesterases [105] and olefin and alkane synthases [106, 107] will also benefit from research and development since the efficiency and specificity of the final steps in a pathway are critical to its success. Finally, looking to nature for the discovery of novel enzymes and metabolic capabilities will continue to be one of the best ways to expand our biosynthetic portfolio.

1.5 Thesis organization

The goal of this work is to expand the potential of microbial fermentation for the renewable production of chemicals, especially commodity chemicals and fuels. Because biosynthesis involves formation of larger hydrocarbon architectures through the assembly of smaller metabolites, we focus in particular on strategies for the construction of carbon-carbon bonds. We have looked to nature for examples of unusual or under-explored carbon metabolism, which offer potential for enzyme and pathway discovery, and we have adopted a two-pronged approach in our investigation, including both *in vitro* biochemical characterization as well as *in vivo* pathway design and metabolic engineering. *Chapter 2* describes the discovery and characterization of enzymes involved in the branched acid fermentation of *Ascaris suum*, as well as the use of these enzymes for the production of α-methyl organic acids in living cells. Preliminary investigation of *in vivo* production of novel PHA polymers is also discussed. *Chapter 3* digs deeper into the structural basis for differences in substrate selectivity observed in *A. suum* thiolase enzymes through X-ray protein crystallography and mutagenesis studies. Finally, *Chapter 4* describes genetic sequencing, enzyme discovery, and physiological studies of *E. gracilis*, in order to gain insight into the fermentative biosynthesis of wax monoesters in this organism.

1.6 References

1. J. D. Keasling, Manufacturing molecules through metabolic engineering. *Science* **2010,** 330, (6009), 1355-1358.
2. J. Becker and C. Wittmann, Advanced biotechnology: Metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angew. Chem. Int. Ed.* **2015,** 54, (11), 3328-3350.
3. A. J. Wargacki, E. Leonard, M. N. Win, D. D. Regitsky, C. N. S. Santos, P. B. Kim, S. R. Cooper, R. M. Raisner, A. Herman, A. B. Sivitz, A. Lakshmanaswamy, Y. Kashiyama, D. Baker, and Y. Yoshikuni, An engineered microbial platform for direct biofuel production from brown macroalgae. *Science* **2012,** 335, (6066), 308-313.
4. M. C. Chang, Harnessing energy from plant biomass. *Curr. Opin. Chem. Biol.* **2007,** 11, (6), 677-684.
5. C. H. Martin, H. Dhamankar, H.-C. Tseng, M. J. Sheppard, C. R. Reisch, and K. L. J. Prather, A platform pathway for production of 3-hydroxyacids provides a biosynthetic route to 3-hydroxy-γ-butyrolactone. *Nature Communications* **2013,** 4, 1414.
6. J. Becker and C. Wittmann, Systems and synthetic metabolic engineering for amino acid production – the heartbeat of industrial strain development. *Curr. Opin. Biotechnol.* **2012,** 23, (5), 718-726.

7. R. I. Zeitoun, A. D. Garst, G. D. Degen, G. Pines, T. J. Mansell, T. Y. Glebes, N. R. Boyle, and R. T. Gill, Multiplexed tracking of combinatorial genomic mutations in engineered cell populations. *Nat. Biotechnol.* **2015,** 33, (6), 631-637.

8. D. R. W. D T Jones, Acetone-butanol fermentation revisited. *Microbiol. Rev.* **1986,** 50, (4), 484.

9. M. Papagianni, Advances in citric acid fermentation by aspergillus niger: Biochemical aspects, membrane transport and modeling. *Biotechnol. Adv.* **2007,** 25, (3), 244-263.

10. S. Kinoshita, Amino-acid production by the fermentation process. *Nature* **1972,** 240, (5378), 211-211.

11. S. Y. Lee, H. M. Kim, and S. Cheon, Metabolic engineering for the production of hydrocarbon fuels. *Curr. Opin. Biotechnol.* **2015,** 33, 15-22.

12. C. A. Rabinovitch-Deere, J. W. K. Oliver, G. M. Rodriguez, and S. Atsumi, Synthetic biology and metabolic engineering approaches to produce biofuels. *Chem. Rev.* **2013,** 113, (7), 4611-4632.

13. H. Chung, J. E. Yang, J. Y. Ha, T. U. Chae, J. H. Shin, M. Gustavsson, and S. Y. Lee, Bio-based production of monomers and polymers by metabolically engineered microorganisms. *Curr. Opin. Biotechnol.* **2015,** 36, 73-84.

14. S. Kind, S. Neubauer, J. Becker, M. Yamamoto, M. Völkert, G. v. Abendroth, O. Zelder, and C. Wittmann, From zero to hero – production of bio-based nylon from renewable resources using engineered *Corynebacterium glutamicum*. *Metab. Eng.* **2014,** 25, 113-123.

15. X. Sun, X. Shen, R. Jain, Y. Lin, J. Wang, J. Sun, J. Wang, Y. Yan, and Q. Yuan, Synthesis of chemicals by metabolic engineering of microbes. *Chem. Soc. Rev.* **2015,** 44, (11), 3760-3785.

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

17. T. Jakočiūnas, M. K. Jensen, and J. D. Keasling, CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* **2016,** 34, 44-59.

18. J. H. Shin, H. U. Kim, D. I. Kim, and S. Y. Lee, Production of bulk chemicals via novel metabolic pathways in microorganisms. *Biotechnol. Adv.* **2013,** 31, (6), 925-935.

19. M. Wen, B. B. Bond-Watts, and M. C. Chang, Production of advanced biofuels in engineered *E. coli*. *Curr. Opin. Chem. Biol.* **2013,** 17, (3), 472-479.

20. H. F. Cueto-Rojas, A. J. A. van Maris, S. A. Wahl, and J. J. Heijnen, Thermodynamics-based design of microbial cell factories for anaerobic product formation. *Trends Biotechnol.* **2015,** 33, (9), 534-546.

21. V. Chubukov, L. Gerosa, K. Kochanowski, and U. Sauer, Coordination of microbial metabolism. *Nature Reviews Microbiology* **2014,** 12, (5), 327-340.

22. J. W. Lee, D. Na, J. M. Park, J. Lee, S. Choi, and S. Y. Lee, Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* **2012,** 8, (6), 536-546.

23. C. Y. Ng, A. Khodayari, A. Chowdhury, and C. D. Maranas, Advances in *de novo* strain design using integrated systems and synthetic biology tools. *Curr. Opin. Chem. Biol.* **2015,** 28, 105-114.

24. B. A. Grzybowski, K. J. M. Bishop, B. Kowalczyk, and C. E. Wilmer, The wired universe of organic chemistry. *Nat. Chem.* **2009,** 1, (1), 31-36.

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

26. L. d'Espaux, D. Mendez-Perez, R. Li, and J. D. Keasling, Synthetic biology for microbial production of lipid-based biofuels. *Curr. Opin. Chem. Biol.* **2015,** 29, 58-65.

27. J. Sheng and X. Feng, Metabolic engineering of yeast to produce fatty acid-derived biofuels: Bottlenecks and solutions. *Front. Microbiol.* **2015,** 6, (17), 6.

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

29. A. Chatterjee, A. B. Hazra, S. Abdelwahed, D. G. Hilmey, and T. P. Begley, A "radical dance" in thiamin biosynthesis: Mechanistic analysis of the bacterial hydroxymethylpyrimidine phosphate synthase. *Angew. Chem.* **2010,** 122, (46), 8835-8838.

30. Y. Zhang, X. Zhu, A. T. Torelli, M. Lee, B. Dzikovski, R. M. Koralewski, E. Wang, J. Freed, C. Krebs, S. E. Ealick, and H. Lin, Diphthamide biosynthesis requires an organic radical generated by an iron–sulphur enzyme. *Nature* **2010,** 465, (7300), 891-896.

31. J. McMurry and T. P. Begley, *The organic chemistry of biological pathways*. Roberts and Co. Publishers: Englewood, Colo., 2005.

32. C. Pandya, J. D. Farelli, D. Dunaway-Mariano, and K. N. Allen, Enzyme promiscuity: Engine of evolutionary innovation. *J. Biol. Chem.* **2014,** 289, (44), 30229-30236.

33. Y.-S. Jang, B. Kim, J. H. Shin, Y. J. Choi, S. Choi, C. W. Song, J. Lee, H. G. Park, and S. Y. Lee, Bio-based production of C2-C6 platform chemicals. *Biotechnol. Bioeng.* **2012,** 109, (10), 2437-2459.

34. O. Tehlivets, K. Scheuringer, and S. D. Kohlwein, Fatty acid synthesis and elongation in yeast. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **2007,** 1771, (3), 255-270.

35. R. M. Lennen and B. F. Pfleger, Engineering *Escherichia coli* to synthesize free fatty acids. *Trends Biotechnol.* **2012,** 30, (12), 659-667.

36. W. Runguphan and J. D. Keasling, Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid-derived biofuels and chemicals. *Metab. Eng.* **2014,** 21, 103-113.

37. B. F. Pfleger, M. Gossing, and J. Nielsen, Metabolic engineering strategies for microbial synthesis of oleochemicals. *Metab. Eng.* **2015,** 29, 1-11.

38. Y. J. Choi and S. Y. Lee, Microbial production of short-chain alkanes. *Nature* **2013,** 502, (7472), 571-574.

39. A. Schirmer, M. A. Rude, X. Li, E. Popova, and S. B. del Cardayre, Microbial biosynthesis of alkanes. *Science* **2010,** 329, (5991), 559-562.

40. M. J. Sheppard, A. M. Kunjapur, and K. L. J. Prather, Modular and selective biosynthesis of gasoline-range alkanes. *Metab. Eng.* **2016,** 33, 28-40.

41. X. Song, H. Yu, and K. Zhu, Improving alkane synthesis in *Escherichia coli* via metabolic engineering. *Appl. Microbiol. Biotechnol.* **2015,** 100, (2), 757-767.

42. H. J. Janßen and A. Steinbüchel, Fatty acid synthesis in *Escherichia coli* and its applications towards the production of fatty acid based biofuels. *Biotechnology for Biofuels* **2014,** 7, (1), 1.

43. W. S. Teo, H. Ling, A.-Q. Yu, and M. W. Chang, Metabolic engineering of *Saccharomyces cerevisiae* for production of fatty acid short- and branched-chain alkyl esters biodiesel. *Biotechnology for Biofuels* **2015,** 8, (1), 1713.

44. X. Li, D. Guo, Y. Cheng, F. Zhu, Z. Deng, and T. Liu, Overproduction of fatty acids in engineered *Saccharomyces cerevisiae*. *Biotechnol. Bioeng.* **2014,** 111, (9), 1841-1852.

45. E. J. Steen, Y. Kang, G. Bokinsky, Z. Hu, A. Schirmer, A. McClure, S. B. del Cardayre, and J. D. Keasling, Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **2010,** 463, (7280), 559-562.

46. Y.-X. Cao, W.-H. Xiao, D. Liu, J.-L. Zhang, M.-Z. Ding, and Y.-J. Yuan, Biosynthesis of odd-chain fatty alcohols in *Escherichia coli*. *Metab. Eng.* **2015,** 29, 113-123.

47. T. P. Howard, S. Middelhaufe, K. Moore, C. Edner, D. M. Kolak, G. N. Taylor, D. A. Parker, R. Lee, N. Smirnoff, S. J. Aves, and J. Love, Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in *Escherichia coli*. *Proceedings of the National Academy of Sciences* **2013,** 110, (19), 7636-7641.

48. F. Zhang, J. M. Carothers, and J. D. Keasling, Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* **2012,** 30, (4), 354-359.

49. A. Krivoruchko, Y. Zhang, V. Siewers, Y. Chen, and J. Nielsen, Microbial acetyl-CoA metabolism and metabolic engineering. *Metab. Eng.* **2015,** 28, 28-42.

50. S. Kim, Y.-S. Jang, S.-C. Ha, J.-W. Ahn, E.-J. Kim, J. Hong Lim, C. Cho, Y. Shin Ryu, S. Kuk Lee, S. Y. Lee, and K.-J. Kim, Redox-switch regulatory mechanism of thiolase from *Clostridium acetobutylicum*. *Nature Communications* **2015,** 6, 8410.

51. H. Inui, K. Miyatake, Y. Nakano, and S. Kitaoka, Fatty acid synthesis in mitochondria of *Euglena gracilis*. *Eur. J. Biochem.* **1984,** 142, (1), 121-126.

52. J. Lian, T. Si, N. U. Nair, and H. Zhao, Design and construction of acetyl-CoA overproducing saccharomyces cerevisiae strains. *Metab. Eng.* **2014,** 24, 139-149.

53. J. Lian and H. Zhao, Recent advances in biosynthesis of fatty acids derived products in *Saccharomyces cerevisiae* via enhanced supply of precursor metabolites. *Journal of Industrial Microbiology & Biotechnology* **2014,** 42, (3), 437-451.

54. Y. Chen, L. Daviet, M. Schalk, V. Siewers, and J. Nielsen, Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metab. Eng.* **2013,** 15, 48-54.

55. J. P. Torella, T. J. Ford, S. N. Kim, A. M. Chen, J. C. Way, and P. A. Silver, Tailored fatty acid synthesis via dynamic control of fatty acid elongation. *Proceedings of the National Academy of Sciences* **2013,** 110, (28), 11290-11295.

56. R. W. Haushalter, W. Kim, T. A. Chavkin, L. The, M. E. Garber, M. Nhan, P. D. Adams, C. J. Petzold, L. Katz, and J. D. Keasling, Production of anteiso-branched fatty acids in *Escherichia coli*; Next generation biofuels with improved cold-flow properties. *Metab. Eng.* **2014,** 26, 111-118.

57. W. Jiang, Y. Jiang, G. J. Bentley, D. Liu, Y. Xiao, and F. Zhang, Enhanced production of branched-chain fatty acids by replacing β-ketoacyl-(acyl-carrier-protein) synthase III (FabH). *Biotechnol. Bioeng.* **2015,** 112, (8), 1613-1622.

58. H. Tao, D. Guo, Y. Zhang, Z. Deng, and T. Liu, Metabolic engineering of microbes for branched-chain biodiesel production with low-temperature property. *Biotechnology for Biofuels* **2015,** 8, (1), 11206.

59. J. L. Meier and M. D. Burkart, The chemical biology of modular biosynthetic enzymes. *Chem. Soc. Rev.* **2009,** 38, (7), 2012-2045.

60. H. G. Menzella, R. Reid, J. R. Carney, S. S. Chandran, S. J. Reisinger, K. G. Patel, D. A. Hopwood, and D. V. Santi, Combinatorial polyketide biosynthesis by *de novo* design and rearrangement of modular polyketide synthase genes. *Nat. Biotechnol.* **2005,** 23, (9), 1171-1176.

61. A. Hagen, S. Poust, T. d. Rond, J. L. Fortman, L. Katz, C. J. Petzold, and J. D. Keasling, Engineering a polyketide synthase for *in vitro* production of adipic acid. *ACS Synthetic Biology* **2016,** 5, (1), 21-27.

62. Q. Liu, K. Wu, Y. Cheng, L. Lu, E. Xiao, Y. Zhang, Z. Deng, and T. Liu, Engineering an iterative polyketide pathway in *Escherichia coli* results in single-form alkene and alkane overproduction. *Metab. Eng.* **2015,** 28, 82-90.

63. T. J. Erb, I. A. Berg, V. Brecht, M. Müller, G. Fuchs, and B. E. Alber, Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: The ethylmalonyl-CoA pathway. *Proceedings of the National Academy of Sciences* **2007,** 104, (25), 10631-10636.

64. A. A. Hunaiti and P. E. Kolattukudy, Source of methylmalonyl-coenzyme A for erythromycin synthesis: Methylmalonyl-coenzyme A mutase from *Streptomyces erythreus*. *Antimicrob. Agents Chemother.* **1984,** 25, (2), 173-178.

65. M. Hügler, C. Menendez, H. Schägger, and G. Fuchs, Malonyl-coenzyme A reductase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO(2) fixation. *J. Bacteriol.* **2002,** 184, (9), 2404-2410.

66. T. A. Leaf and F. Srienc, Metabolic modeling of polyhydroxybutyrate biosynthesis. *Biotechnol. Bioeng.* **1998,** 57, (5), 557-570.

67. J.-M. Jeon, C. J. Brigham, Y.-H. Kim, H.-J. Kim, D.-H. Yi, H. Kim, C. Rha, A. J. Sinskey, and Y.-H. Yang, Biosynthesis of poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) (P(HB-co-HHx)) from butyrate using engineered *Ralstonia eutropha*. *Appl. Microbiol. Biotechnol.* **2014,** 98, (12), 5461-5469.

68. Q. Wang, Y. Luan, X. Cheng, Q. Zhuang, and Q. Qi, Engineering of *Escherichia coli* for the biosynthesis of poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) from glucose. *Appl. Microbiol. Biotechnol.* **2015,** 99, (6), 2593-2602.

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

70. C. Dellomonaco, J. M. Clomburg, E. N. Miller, and R. Gonzalez, Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **2011,** 476, (7360), 355-359.

71. C. R. Shen, E. I. Lan, Y. Dekishima, A. Baez, K. M. Cho, and J. C. Liao, Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.* **2011,** 77, (9), 2905-2915.

72. L. Hiser, M. E. Basson, and J. Rine, Erg10 from *Saccharomyces cerevisiae* encodes acetoacetyl-CoA thiolase. *J. Biol. Chem.* **1994,** 269, (50), 31383-31389.

73. S. Kim, J. M. Clomburg, and R. Gonzalez, Synthesis of medium-chain length (C6–C10) fuels and chemicals via β-oxidation reversal in *Escherichia coli*. *Journal of Industrial Microbiology & Biotechnology* **2015,** 42, (3), 465-475.

74. A. Cintolesi, J. M. Clomburg, and R. Gonzalez, *In silico* assessment of the metabolic capabilities of an engineered functional reversal of the β-oxidation cycle for the synthesis of longer-chain (C≥4) products. *Metab. Eng.* **2014,** 23, 100-115.

75. H. Inui, K. Miyatake, Y. Nakano, and S. Kitaoka, Wax ester fermentation in *Euglena gracilis*. *FEBS Lett.* **1982,** 150, (1), 89-93.

76. M. Nakazawa, H. Andoh, K. Koyama, Y. Watanabe, T. Nakai, M. Ueda, T. Sakamoto, H. Inui, Y. Nakano, and K. Miyatake, Alteration of wax ester content and composition in *Euglena gracilis* with gene silencing of 3-ketoacyl-CoA thiolase isozymes. *Lipids* **2015,** 50, (5), 483-492.

77. S. Tucci, R. Vacula, J. Krajcovic, P. Proksch, and W. Martin, Variability of wax ester fermentation in natural and bleached *Euglena gracilis* strains in response to oxygen and the elongase inhibitor flufenacet. *J. Eukaryot. Microbiol.* **2010,** 57, (1), 63-69.

78. E.-J. Kim and K.-J. Kim, Crystal structure and biochemical characterization of PhaA from *Ralstonia eutropha*, a polyhydroxyalkanoate-producing bacterium. *Biochem. Biophys. Res. Commun.* **2014,** 452, (1), 124-129.

79. C. D. Fage and A. T. Keatinge-Clay, Crystal structure of beta-ketoacyl thiolase B (BktB) from *Ralstonia eutropha*. **2015**.

80. M. J. Sheppard, A. M. Kunjapur, S. J. Wenck, and K. L. J. Prather, Retro-biosynthetic screening of a modular pathway design achieves selective route for microbial synthesis of 4-methyl-pentanol. *Nature Communications* **2014,** 5, 5031.

81. Y. Tashiro, G. M. Rodriguez, and S. Atsumi, 2-keto acids based biosynthesis pathways for renewable fuels and chemicals. *Journal of Industrial Microbiology & Biotechnology* **2014,** 42, (3), 361-373.

82. E. A. Felnagle, A. Chaubey, E. L. Noey, K. N. Houk, and J. C. Liao, Engineering synthetic recursive pathways to generate non-natural small molecules. *Nat. Chem. Biol.* **2012,** 8, (6), 518-526.

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

84. P. Jambunathan and K. Zhang, Novel pathways and products from 2-keto acids. *Curr. Opin. Biotechnol.* **2014,** 29, 1-7.

85. J. C. Liao, L. Mi, S. Pontrelli, and S. Luo, Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nat. Rev. Microbiol.* **2016,** 14, (5), 288-304.

86. M. C. Y. Chang and J. D. Keasling, Production of isoprenoid pharmaceuticals by engineered microbes. *Nat. Chem. Biol.* **2006,** 2, (12), 674-681.

87. H. M. Miziorko, Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* **2011,** 505, (2), 131-143.

88. W. Eisenreich, A. Bacher, D. Arigoni, and F. Rohdich, Biosynthesis of isoprenoids via the non-mevalonate pathway. *Cellular and Molecular Life Sciences CMLS* 61, (12), 1401-1426.

89. P. K. Ajikumar, W.-H. Xiao, K. E. J. Tyo, Y. Wang, F. Simeon, E. Leonard, O. Mucha, T. H. Phon, B. Pfeifer, and G. Stephanopoulos, Isoprenoid pathway optimization for taxol precursor overproduction in escherichia coli. *Science* **2010,** 330, (6000), 70-74.

90. C. M. Immethun, A. G. Hoynes-O'Connor, A. Balassy, and T. S. Moon, Microbial production of isoprenoids enabled by synthetic biology. *Front. Microbiol.* **2013,** 4.

91. C. J. Paddon and J. D. Keasling, Semi-synthetic artemisinin: A model for the use of synthetic biology in pharmaceutical development. *Nature Reviews Microbiology* **2014,** 12, (5), 355-367.

92. S. Noda, T. Shirai, S. Oyama, and A. Kondo, Metabolic design of a platform *Escherichia coli* strain producing various chorismate derivatives. *Metab. Eng.* **2016,** 33, 119-129.

93. N. Hadadi and V. Hatzimanikatis, Design of computational retrobiosynthesis tools for the design of *de novo* synthetic pathways. *Curr. Opin. Chem. Biol.* **2015,** 28, 99-104.

94. H. Yim, R. Haselbeck, W. Niu, C. Pujol-Baxley, A. Burgard, J. Boldt, J. Khandurina, J. D. Trawick, R. E. Osterhout, R. Stephen, J. Estadilla, S. Teisan, H. B. Schreyer, S. Andrae, T. H. Yang, S. Y. Lee, M. J. Burk, and S. Van Dien, Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat. Chem. Biol.* **2011,** 7, (7), 445-452.

95. E. Brunk, M. Neri, I. Tavernelli, V. Hatzimanikatis, and U. Rothlisberger, Integrating computational methods to retrofit enzymes to synthetic pathways. *Biotechnol. Bioeng.* **2011,** 109, (2), 572-582.

96. W. S. Mak, S. Tran, R. Marcheschi, S. Bertolani, J. Thompson, D. Baker, J. C. Liao, and J. B. Siegel, Integrative genomic mining for enzyme function to enable engineering of a non-natural biosynthetic pathway. *Nature Communications* **2015,** 6, 10005.

97. G. I. Guzmán, J. Utrilla, S. Nurk, E. Brunk, J. M. Monk, A. Ebrahim, B. O. Palsson, and A. M. Feist, Model-driven discovery of underground metabolic functions in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2015,** 112, (3), 929-934.

98. J. L. Avalos, G. R. Fink, and G. Stephanopoulos, Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols. *Nat. Biotechnol.* **2013,** 31, (4), 335-341.

99. J. T. Boock, A. Gupta, and K. L. Prather, Screening and modular design for metabolic pathway optimization. *Curr. Opin. Biotechnol.* **2015,** 36, 189-198.

100. H. C. Tseng and K. L. J. Prather, Controlled biosynthesis of odd-chain fuels and chemicals via engineered modular metabolic pathways. *Proceedings of the National Academy of Sciences* **2012,** 109, (44), 17925-17930.

101. P. Xu, Q. Gu, W. Wang, L. Wong, A. G. W. Bower, C. H. Collins, and M. A. G. Koffas, Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications* **2013,** 4, 1409.

102. H. K. Privett, G. Kiss, T. M. Lee, R. Blomberg, R. A. Chica, L. M. Thomas, D. Hilvert, K. N. Houk, and S. L. Mayo, Iterative approach to computational enzyme design. *Proc. Natl. Acad. Sci. U.S.A.* **2012,** 109, (10), 3790-3795.

103. M. A. Smith, P. A. Romero, T. Wu, E. M. Brustad, and F. H. Arnold, Chimeragenesis of distantly-related proteins by noncontiguous recombination. *Protein Sci.* **2012,** 22, (2), 231-238.

104. I. Chen, B. M. Dorr, and D. R. Liu, A general strategy for the evolution of bond-forming enzymes using yeast display. *Proceedings of the National Academy of Sciences* **2011,** 108, (28), 11399-11404.

105. M. D. McMahon and K. L. J. Prather, Functional screening and *in vitro* analysis reveal thioesterases with enhanced substrate specificity profiles that improve short-chain fatty acid production in *Escherichia coli*. *Appl. Environ. Microbiol.* **2014,** 80, (3), 1042-1050.

106. L. J. Rajakovich, H. Nørgaard, D. M. Warui, W.-c. Chang, N. Li, S. J. Booker, C. Krebs, J. M. Bollinger Jr., and M.-E. Pandelia, Rapid reduction of the diferric-peroxyhemiacetal intermediate in aldehyde-deformylating oxygenase by a cyanobacterial ferredoxin: Evidence for a free-radical mechanism. *JACS* **2015,** 137, (36), 11695-11709.

107. Z. Rui, N. C. Harris, X. Zhu, W. Huang, and W. Zhang, Discovery of a family of desaturase-like enzymes for 1-alkene biosynthesis. *ACS Catalysis* **2015,** 5, (12), 7091-7094.

Chapter 2: *Discovery and engineering of pathways for production of α-branched organic acids*

2.1 Introduction

Living systems have unique advantages as catalysts for industrial chemical synthesis [1]. Microorganisms constitute highly efficient self-renewing reaction vessels, which carry out the telescoped conversion of glucose and other renewable feedstocks into the vast array of metabolites and natural products observed in nature. The infrastructure for microbial biosynthesis has already been developed for the production of chemicals such as chiral amino acids [2, 3] that would be difficult to make synthetically, and the technology to modify DNA *in vitro* and *in vivo*, and thus the enzymes encoded by this DNA, has rapidly improved in recent years [4-6]. Many target molecules have been proposed which could be derived biologically from central metabolic intermediates, and serve as drop-in replacements for chemicals currently obtained from petroleum [9, 10]. But significantly expanding the synthetic capability of biology will require a deeper understanding of the toolbox of reaction methodologies available in nature.

Chemists approach problems in synthesis using an analytical, retro-synthetic approach, identifying sites of disconnection in the carbon framework of a target molecule [11]. As such, synthetic chemists have placed great emphasis in the past century on the discovery of carbon-carbon bond forming reactions. This strategy has proved rewarding, resulting in a wide range of reactions, the diversity of which has allowed for the selective assembly of chemical precursors into complex structures. Compared to chemical synthesis, nature uses a surprisingly small set of chemical reactions and functional groups to construct a variety of natural products from the same few raw materials [12]. Most carbon-carbon bonds, for example, are synthesized using just two related reactions, aldol and Claisen condensation. A hypothetical coupling reaction of two compounds, R-X and R′-Y, can be used to illustrate this comparison. In chemical synthesis, the specificity of the R-R′ coupling arises from the unique chemistry of –X and –Y functional groups in the context of the reaction conditions, catalysts, or specialty reagents employed. In biology, X and Y are ubiquitous (*e.g.* a coenzyme A thioester) and the solvent system and reaction conditions inside the cell are similar throughout life. Reaction orthogonality arises from the binding specificity of enzymes, which have evolved to distinguish, or alternatively, to be permissive of, slight differences in R and R′, while conserving biosynthetic mechanisms.

Thiolases are an an important class of enzyme that catalyze reversible carbon-carbon bond formation in nature, through the Claisen condensation of acyl-CoA thioesters [13]. Thiolases can function physiologically in both catabolism, as in the β-oxidation of fatty acids, as well in biosynthesis. Prominent examples of biosynthetic functions of thiolases include the first steps in the production of PHA polymers in bacteria [14], the first step of terpenoid synthesis in the cytosol of animals and fungi [15, 16], and natural and synthetic pathways for the production of butanol [17-19]. In all these cases, the particular reaction catalyzed is the condensation of two acetyl-CoA to form 3-oxobutyryl-CoA. In principle, thiolase-catalyzed condensation followed by functional tailoring through ketoreduction, dehydration, etc. could be used to construct carbon-carbon bonds between any desired acyl-CoA substrates in the cell, but in practice the biosynthetic thiolases that have been characterized represent a narrow range of substrates. Especially underrepresented are thiolase-based pathways that can install a substituent at the α-position by using an acyl-CoA nucleophile other than acetyl-CoA. Even small substitutions to a molecule, such as the introduction of a methyl group, can offer substantial advantages in downstream properties and applications, such as the material properties of plastics [20] or the altered combustion properties of branched hydrocarbon fuels [21]. Expanding the scope of

thiolases in engineered pathways would provide a useful route to synthetic precursors and chiral building blocks, including branched hydroxyacids for novel synthetic polymers [22].

The roundworm *Ascaris suum* has been reported to accumulate short branched organic acids under anaerobic conditions [23, 24]. This pathway is a unique example in biology of a thiolase-based pathway that functions physiologically in biosynthesis and condenses both acetate and propionate units (*Figure 2.1*). De Mata *et al*. reported the purifications of more than one thiolase isozyme from muscle tissue mitochondria, one of which had higher activity with propionyl-CoA than acetyl-CoA [25, 26], as well as a ketoreductase which was more active on branched substrates [27]. Hydroxyacyl-CoA compounds formed in this way are dehydrated, providing a pool of enoyl-CoA compounds that serve as terminal electron acceptors in a rhodoquinone-dependent electron transport chain [28, 29]. The propionyl-CoA itself originates from the decarboxylation of succinyl-CoA formed from fumarate reduction, providing further ATP [30]. The result of this anaerobic respiration is the accumulation of a mixture of short organic acids, both linear and branched.

We became interested in studying this unusual thiolase-based pathway in order to exploit its potential for producing hydrocarbons with novel architectures and stereochemical purity (*Figure 2.1*). Using the available RNAseq transcriptome database of *A. suum*, we identified thiolase, ketoreductase, and dehydratase enzyme candidates, characterized them *in vitro*, and used them to engineer pathways for the production of α-methyl organic acids in *Escherichia coli*.

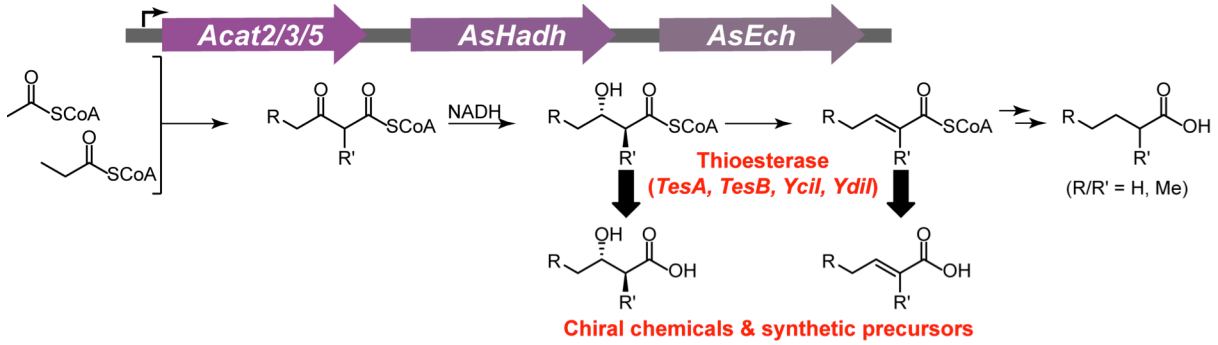


Figure 2.1. *Branched acid fermentation in Ascaris suum expands the biosynthetic utility of thiolase-based pathways. Propionyl-CoA can be condensed to form α-methyl organic acids.*

2.2 Materials and methods

Commercial materials. LB Broth Miller, LB Agar Miller, Terrific Broth (TB), glycerol, ammonium hydroxide, methylsulfoxide (DMSO), and chloramphenicol (Cm) were purchased from EMD Biosciences (Darmstadt, Germany). D-(+)-glucose was purchased from MP Biochemicals (Santa Ana, CA). Tris-(2-carboxyethyl)phosphine hydrochloride (TCEP) was purchased from Biosynth, Inc. (Itasca, IL). Imidazole was purchased from Acros Organics (Morris Plains, NJ). Isopropyl-β-D-thiogalactopyranoside (IPTG) was purchased from Santa Cruz Biotechnology (Dallas, TX). Sodium hydroxide was purchased from Avantor Performance Materials (Center Valley, PA). Deuterated water (D_2O) was purchased from Cambridge Isotope Laboratories (Tewksbury, MA). Carbenicillin sodium salt (Cb), kanamycin monosulfate (Km), phenylmethanesulfonyl fluoride (PMSF), Tris, sodium chloride, potassium chloride, potassium phosphate monobasic, potassium phosphate dibasic, dithiothreitol (DTT), 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), magnesium chloride hexahydrate, acetonitrile, diethyl ether, N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA), ethyl acetate, hydrochloric acid, sodium sulfate anhydrous, sodium bicarbonate, ammonium acetate, and methanol were purchased from Fisher Scientific (Pittsburgh, PA). Ethyl 2-methyl-3-oxopentanoate was purchased from Fragmenta (Monmouth Junction, NJ). Ethyl 2-methyl-3-oxohexanoate and ethyl 2-ethyl-3-oxohexanoate were purchased from FCH Group (Chernigiv, Ukraine). Ethyl 2-methylacetoacetate, ethyl 3-oxovalerate, ethyl 3-oxohexanoate, sodium 3-hydroxybutyrate, 2-methyl-3-hydroxybutyric acid, crotonic acid, 2-pentenoic acid, *trans*-2-methyl-2-butenoic acid (tiglic acid), *tran*-2-methyl-2-pentenoic acid, 4-(Dimethylamino)pyridine (DMAP), N,N-diisopropylcarbodiimide, hydroxylamine hydrochloride, 5,5′-Dithiobis(2-nitrobenzoic acid) (DTNB), adipic acid, 6-bromohexanoic acid, thiophenol, sodium borohydride, sodium propionate, 2-(N-Morpholino)ethanesulfonic acid hydrate (MES), calcium chloride, poly(ethylene glycol) av. Mn 3,350 (PEG 3350), coenzyme A sodium salt hydrate (CoA), acetyl-CoA, propionyl-CoA, butyryl-CoA, acetoacetyl-CoA, crotonyl-CoA, β-nicotinamide adenine dinucleotide reduced dipotassium salt (NADH), β-nicotinamide adenine dinucleotide 2'-phosphate reduced tetrasodium salt (NADPH), β-nicotinamide adenine dinucleotide hydrate (NAD^+), formic acid, trichloroacetic acid (TCA), 1-propanol, dichloroethane, β-mercaptoethanol (BME), lysozyme from chicken egg white, and bovine serum albumin (BSA) were purchased from Sigma-Aldrich (St. Louis, MO). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad Protein Assay reagent, ammonium persulfate, and N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED) were purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, Q5 DNA Polymerase, T4 DNA ligase, Phusion DNA polymerase, T5 exonuclease, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Pierce™ BCA Protein Assay Kit was purchased from Thermo Scientific (Waltham, MA) Oligonucleotides and gBlocks were purchased from Integrated DNA Technologies (Coralville, IA). DNA purification kits and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). cOmplete EDTA-free protease inhibitor were purchased from Roche Applied Science (Penzberg, Germany). TEV protease was purchased from the QB3 MacroLab at UC Berkeley. Amicon Ultra spin concentrators and $MultiScreen_{HTS}$ 0.22μm filter plates were purchased from Merck Millipore (Cork, Ireland). Mass spectral analyses were carried out at the College of Chemistry Mass Spectrometry Facility at UC Berkeley.

Bacterial strains. *E. coli* DH10B-T1^R was used for DNA construction and BL21(de3)-T1^R was used for heterologous production of proteins for purification. *E. coli* BAP-1 [31] was used for propionate-supplemented production of organic acids.

To prepare chemically competent cells from plasmid transformations, the appropriate *E. coli* strain was cultured from a single colony in 100 mL LB broth, harvested at OD_{600} of 0.4-0.5, and resuspended in 10 mL of sterile Transformation-Storage Solution (10% w/v PEG 3350, 5% v/v DMSO, 20 mM $MgCl_2$ in LB broth). 100 µL aliquots were flash frozen in liquid nitrogen and stored at -80°C for later use. Plasmid transformations were performed by adding one or more plasmid, 20 µL 5X KCM solution (0.5M KCl, 0.15M $CaCl_2$, 0.25M $MgCl_2$), and 80 µL sterile water to a thawed competent cell aliquot of the appropriate strain. After 10 min incubation on ice, cells were subject to heat shock at 42°C for 90 sec, diluted to 1 mL with LB broth, and incubated at 37°C for 1 hour before plating with sterile beads on LB agar containing appropriate antibiotics.

Gene and plasmid construction. Plasmid construction was carried out using standard molecular biology techniques and the Gibson protocol [5]. Descriptions of strains and plasmids are listed in *Appendix, Table S1A*. PCR amplifications were carried out with Q5 DNA polymerase, following manufacturer instructions. Primer sequences are listed in *Appendix, Table S1B*. Constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA). Synthetic genes were assembled using gBlock sequences (*Appendix, Table S1C*). gBlocks were resuspended at 10 ng/µl in 10 mM Tris-HCl, pH 8.5 and used directly for assembly of vectors.

pET16b-AsAcat1, pET16b-AsHadh, and pET16b-AsEch were constructed by inserting the corresponding gBlocks into pET16b digested with NdeI and BamHI. pSV272.1-AsAcat2, pSV272.1-AsAcat3, pSV272.1-AsAcat4, and pSV272.1-AsAcat5 were constructed by inserting the corresponding gBlocks into pSV272.1 digested with SfoI and BamHI, using the Gibson protocol. pET28a-CnHadh was constructed by inserting the corresponding gBlock into pET28a digested with NdeI and BamHI. pET23a vectors containing *A. suum* thiolases were constructed by PCR amplifying the gene from the corresponding pET16b or pSV272.1 vector using the corresponding primer set and inserting into pET23a digested with SfoI and XhoI.

pTrc99-A2-AsHadh-Ech vector was constructed by PCR amplifying Acat2 with primers Acat FP and Acat RP-Hadh, Hadh with primers Hadh FP-Acat2 and Hadh RP-Ech, and Ech with primers Ech FP and Ech RP, and inserting all three amplicons into pTrc99 digested with NcoI and HindIII. pTrc99-A2-AsHadh vector was constructed by PCR amplifying the Acat2-Hadh region from pTrc99-A2-AsHadh-Ech with primers Acat2 FP and Hadh RP and inserting into pTrc99 digested with NcoI and HindIII. Vectors containing Acat3 and Acat5 were constructed in an analogous fashion.

pTrc33 vectors containing TesA, TesB, YciA, and YdiI were constructed by PCR amplifying the thioesterase genes from the genome of *E. coli* using the corresponding primers (*e.g.* TesA FP and TesA RP for TesA) and a boiled *E. coli* colony as template. The amplicons were inserted into pTrc33 digested with SacI and HindIII. pTrc33-TesB2 was constructing by amplifying TesB gene with primers TesB2 FP and TesB RP and inserting into pTrc33.

To construct the A3AsH and A3CnH plasmid series, Acat3 gene and a ketoreductase gene were amplified and inserted into pTrc99a or pCWori vector, as described in *Appendix, Table S1D*.

Poly1-4 vectors were constructed by PCR amplifying Acat2 or Acat3 as well as *Rhodococcus ruber* PhaC or *Rhodococcus opacus* PhaC and inserting the appropriate combination into pTrc99a, as described in *Appendix, Table S1E*. For RruPhaC and RopPhaC genes, the templates used were pPol4.1 and pPol4.2 plasmids, respectively. DKD2 was constructed by PCR amplifying PhaJ, AsHadh, and AsEch using the corresponding primers (*e.g.* DKD PhaJ F and DKD PhaJ R) and inserting into pTrc33. pTrc33-PhaB was constructed by amplifying PhaB with primers pTrc33-PhaB Fwd and pTrc33-PhaB Rev and inserting into pTrc33. The template for PhaJ and PhaB amplification was pBT33-Bu3 plasmid.

Expression and purification of heterologously expressed enzymes *Appendix, Table S1F* lists amino acid sequences of purified proteins from expression vectors constructed in this study. Other proteins, including His_{10}-PhaA, His_{10}-PhaB, His_{10}-PhaJ, Strep-Crt, and His_{10}-Hbd, were provided by Dr. Jeffrey Hanson.

Expression of His_6-tagged thiolase enzymes. TB (1 L) containing 50 µg/mL Cb in a 2.8 L Fernbach baffled shake flask was inoculated to OD_{600} = 0.05 with an overnight culture of freshly transformed *E. coli* containing the pET23a overexpression plasmid. The cultures were grown at 37°C at 250 rpm to OD_{600} = 0.7 to 1.0, at which point cultures were cooled on ice for 20 min, followed by induction of protein expression with IPTG (0.2 mM) and overnight growth at 16°C. Cell pellets were harvested by centrifugation at 6,000 × *g* for 7 min at 4ºC and stored at -80ºC.

Expression of His_{10}-AsHadh, His_6-AsEch, and His_{10}-TesB. TB (0.5 L) containing 50 µg/mL Cb or Km, as appropriate, in a 2 L baffled shake flask was inoculated to OD_{600} = 0.1 with an overnight culture of freshly transformed *E. coli* containing the overexpression plasmid. The cultures were grown at 37°C at 250 rpm to OD_{600} = 0.8 to 0.9 at which protein expression was induced with IPTG (0.4 mM) for 3-4 hours at 37°C. Cell pellets were harvested by centrifugation at 6,000 × *g* for 7 min at 4ºC and stored at -80ºC.

Purification of Ascaris suum thiolases. Frozen cell pellets were thawed and resuspended at 5 mL per g cell paste with Buffer 2A (50 mM HEPES, 300 mM NaCl, 15% glycerol, pH 7.5) supplemented with TCEP (2 mM) and PMSF (1 mM). For Acat4, PMSF was substituted with cOmplete EDTA-free protease inhibitor. Lysozyme was added to a concentration of 0.2 mg/mL and the cell suspension was stirred at 4°C for 30-60 min. The cells were lysed by sonication using a Misonix Sonicator 3000, and lysate was centrifuged at 6,000 × *g* for 40 min at 4°C to separate soluble and insoluble fractions. The soluble lysate was loaded onto a Ni-NTA agarose column (Qiagen) by gravity flow or on an ÄKTApurifier FPLC (2 mL/min; GE Healthcare). The column was washed with 18 column volumes of Buffer 2B (Buffer A with 0.4 mM TCEP) followed by 12 column volumes of 84% Buffer 2B, 16% Buffer 2C (Buffer 2B with 0.25 M imidazole) before eluting with Buffer 2C. Fractions containing protein were pooled by A_{280nm} and concentrated using an Amicon Ultra spin concentrator.

Protein was dialyzed against Buffer 2B overnight with TEV protease [32], 1 mg/10 mg protein substrate, to remove the His_6 tag. The protein was passed three times over Ni-NTA agarose by gravity flow and the column washed with Buffer 2B until the eluate reached A_{280nm} below 0.05. The flow-through was concentrated and loaded onto a Superdex 200 16/60 PG (GE Healthcare) column equilibrated with Buffer 2D (20 mM HEPES, 100 mM NaCl, 10% glycerol, 0.4 mM TCEP, pH 6.8). Fractions containing active enzyme were concentrated between 10-20 mg/mL, frozen in liquid nitrogen, and stored at -80°C. Protein concentrations were estimated using ε_{280nm} calculated by ExPASY ProtParam [33] as follows: Acat1 (ε_{280nm} = 22920 M^{-1} cm^{-1}),

Acat2 (ε_{280nm} = 21430 M^{-1} cm^{-1}), Acat3 (ε_{280nm} = 24410 M^{-1} cm^{-1}), Acat4 (ε_{280nm} = 15930 M^{-1} cm^{-1}), Acat5 (ε_{280nm} = 24410 M^{-1} cm^{-1}).

Purification of His_{10}-AsHadh and His_6-CnHadh. Frozen cell pellets were thawed and resuspended at 5 mL per g cell paste in Buffer 2E (50 mM potassium phosphate, 300 mM NaCl, 10% glycerol, 20 mM imidazole, pH 8.0) supplemented with BME (20 mM) and PMSF (1 mM). Lysozyme was added to a concentration of 0.2 mg/mL and the cell suspension was stirred at 4°C for 60 min. The cells were lysed by sonication using a Misonix Sonicator 3000, and lysate was centrifuged at 18,000 × *g* for 25 min at 4°C to separate soluble and insoluble fractions. The soluble lysate was loaded onto a Ni-NTA agarose column (Qiagen) on a ÄKTApurifier FPLC (2 mL/min; GE Healthcare). The column was washed with 18 column volumes of Buffer E followed by 12 column volumes of 92% Buffer 2E, 8% Buffer 2F (50 mM potassium phosphate, 300 mM NaCl, 10% glycerol, 250 mM imidazole, pH 8.0) before eluting with Buffer 2F. Fractions containing protein were pooled by A_{280nm} and concentrated using an Amicon Ultra spin concentrator. The protein was loaded onto a Superdex 200 16/60 PG (GE Healthcare) column equilibrated with Buffer 2G (30 mM HEPES, 150 mM NaCl, 15% glycerol, pH 6.8, 0.4 mM TCEP). Fractions containing protein were concentrated, frozen in liquid nitrogen, and stored at -80°C. Protein concentration of His_{10}-AsHadh was estimated using ε_{280nm} calculated by ExPASY ProtParam [33]: 7450 M^{-1} cm^{-1}. Protein concentration of His_6-CnHadh was estimated using the Pierce™ BCA Protein Assay Kit with BSA as standard.

Purification of His_{10}-AsEch1 and His_{10}-TesB. Ni-NTA purification was carried out in the same way as AsHadh purification. After elution, fractions containing protein were pooled and concentrated using an Amicon Ultra spin concentrator, diluted with Buffer 2G, and concentrated again. Aliquots were frozen in liquid nitrogen and stored at -80°C. Final protein concentrations were estimated using ε_{280nm} calculated by ExPASY ProtParam [33] as follows: His_{10}-AsEch (ε_{280nm} = 20970 M^{-1} cm^{-1}), His_{10}-TesB (ε_{280nm} = 28420 M^{-1} cm^{-1}).

Preparation of acyl-CoA compounds. All acyl-CoA compounds were prepared by coupling free acid to thiophenol, followed by transthioesterification with CoA and HPLC purification.

Free 3-oxoacid compounds were prepared by hydrolyzing 1.5 mmol of the corresponding ethyl ester in 1.1 equivalents of 1 M aqueous NaOH overnight at room temperature. The reaction was neutralized with 6 M HCl to pH 7 and extracted three times with ethyl acetate. The aqueous layer was acidified to pH 1 and extracted three times with ethyl acetate. Combined organic extracts were dried with Na_2SO_4, and the solvent removed by rotary evaporation. 3-oxohexanoic acid and 3-oxopentanoic acid were observed to be white solids while the branched compounds were oils. The free acids obtained were used immediately in the subsequent coupling.

Crude 3-oxoacid products from above, or 1.5 mmol commercial tiglic acid, was dissolved in 10 mL ethyl acetate in a scintillation vial with a stir bar, to which 2 to 3 mg DMAP and 1.2 mmol thiophenol (122 µL) were added. The reaction was cooled in an ice bath before dropwise addition of 1.5 mmol N,N'-diisopropylcarbodiimide (235 µL). Formation of a white precipitant was observed after 5-10 min. The reaction was allowed to continue for 2 hours at 0°C and 2 hours at room temperature before the white solid was filtered and the filtrate extracted three times with 1 M $NaHCO_3$. The organic layer was dried with Na_2SO_4 and solvent was removed by rotary evaporation. The crude mixture was dissolved in 100 µL acetonitrile and transferred to a stirring solution of 55 µmol CoA in 2 mL 0.5 M $NaHCO_3$ in an ice bath. After one hour at 0°C,

the reaction was warmed to room temperature for one hour before quenching by addition of 50% formic acid to pH 3. The reaction was extracted twice with diethyl ether, and nitrogen was passed over the aqueous layer to remove residual ether.

Compounds were purified on an Agilent 1200 HPLC system using a Zorbax Eclipse XDB C-18 column (5 μm, 9.4 × 250 mm, Agilent) with a linear gradient over 24 min, starting at 4% acetonitrile and ending between 20 and 50% acetonitrile depending on the acyl-CoA product, with 100 mM ammonium acetate, pH 5.0, as the aqueous mobile phase (4 mL/min). Fractions were screened by ESI-MS and those containing pure 3-oxoacyl-CoA were lyophilized overnight, dissolved in water (1 ml), and desalted using the same HPLC column with a linear gradient from 0 to 15% acetonitrile over 16 min with water as the mobile phase (4 ml/min). Desalted 3-oxoacyl-CoA was lyophilized and redissolved in D_2O for NMR measurements. Complete or significant H-D exchange was observed at α-carbon atoms by ^1H NMR in D_2O. NMR samples were lyophilized, redissolved in H_2O, and quantified, and aliquots were stored at -80°C.

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

Characterization of synthesized compounds. Synthetic 3-oxoacyl-CoA compounds were analyzed by HPLC on an Agilent Eclipse XDB-C18 (3.5 μm, 3.0 × 150 mm) using a linear gradient from 4% to 70% acetonitrile in 100 mM ammonium acetate, pH 5.0 to methanol over 12 min at 0.6 mL/min with detection of CoA absorbance at 260 nm.

^1H NMR spectra were collected at 25ºC on the AV-600 spectrometer at the College of Chemistry NMR Facility at the University of California, Berkeley. ^{13}C NMR and ^{13}C-^1H HSQC spectra were collected on the Bruker Biospin 900 MHz spectrometer at the QB3 Central California 900 MHz NMR Facility. The ^{13}C NMR spectrum of 2-ethyl-3-oxohexanoyl-CoA was referenced to added 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS, 0 ppm) and subsequent spectra were referenced to added methanol (CH_3OH, 51.59 ppm). 100 μM EDTA was added to some samples to sharpen NMR signal. NMR assignments are based on ^{13}C-^1H HSQC spectra and previous literature. Full spectra are shown in *Appendix, Figures S1-S7*. High-resolution mass spectral analyses were carried out at the College of Chemistry Mass Spectrometry Facility.

2-Methyl-3-oxobutyryl coenzyme A. ^1H NMR (600 MHz, D_2O = 4.79 ppm): δ 8.65 (s, 1H, H_8), 8.40 (s, 1H, H_2), 6.20 (d, J = 5.2 Hz, 1H, $H_{1'}$), 4.90 – 4.83 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.65 – 4.60 (m, 1H, $H_{4'}$), 4.29 (qdd, J = 11.7, 5.0, 2.9 Hz, 2H, $H_{5'}$), 4.09 (q, 0H, J = 7.0 Hz, H_b), 4.05 (s, 1H, $H_{3''}$), 3.89 (dd, J = 9.7, 4.8 Hz, 1H, pro-R-$H_{1''}$), 3.62 (dd, J = 9.8, 4.7 Hz, 1H, pro-S-$H_{1''}$), 3.47 (t, J = 6.8 Hz, 2H, $H_{3''}$), 3.38 (t, J = 6.3 Hz, 2H, $H_{8''}$), 3.07 (t, J = 6.2 Hz, 2H, $H_{9''}$), 2.45 (t, J = 6.6 Hz, 2H, $H_{6''}$), 2.29 (s, 3H, H_d), 1.33 (d, J = 8.4 Hz, 3H, H_g), 0.96 (s, 3H, $H_{10''}$), 0.83 (s, 3H, $H_{11''}$). **^{13}C NMR** (226 MHz, D_2O, CH_3OH = 51.59 ppm)**:** δ 211.61 (C_c), 202.81 (C_a), 177.45 ($C_{4''}$), 176.71 ($C_{7''}$), 154.69 (C_6), 151.44 (C_4), 150.56 (C_2), 144.11 (C_8), 121.15 (C_5), 89.67 ($C_{1'}$), 86.26 (d, J = 10.2 Hz, $C_{4'}$), 76.86 (s, $C_{2'}$, $C_{3'}$, or $C_{3''}$), 76.68 (s, $C_{2'}$, $C_{3'}$, or $C_{3''}$), 74.59 (d, J = 6.0 Hz, $C_{1''}$), 67.90 ($C_{5'}$), 63.63 (t, J = 19.9 Hz, C_b), 41.05 ($C_{8''}$), 38.12 ($C_{5''}$ or $C_{6''}$), 38.03 ($C_{5''}$ or $C_{6''}$), 31.26 ($C_{9''}$ or C_d), 31.19 ($C_{9''}$ or C_d) 23.61 ($C_{10''}$), 20.88 ($C_{11''}$), 15.80 (C_g). **HR-ESI-MS $[M-H]^-$:** calculated for $C_{26}H_{41}N_7O_{18}P_3S$, *m/z* 864.1447, found *m/z* 864.1443.

3-Oxopentanoyl coenzyme A. **^{1}H NMR** (600 MHz, D_2O = 4.79 ppm)**:** δ 8.68 (s, 1H, H_8), 8.44 (s, 1H, H_2), 6.22 (d, J = 5.0 Hz, 1H, $H_{1'}$), 4.90 – 4.84 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.65 – 4.61 (m, 1H, $H_{4'}$), 4.35 – 4.24 (m, 2H, $H_{5'}$), 4.06 (s, 1H, $H_{3''}$), 3.90 (dd, J = 9.8, 4.8 Hz, 1H, pro-R-$H_{1''}$), 3.63 (dd, J = 9.8, 4.6 Hz, 1H, pro-S- $H_{1''}$), 3.48 (t, J = 6.7 Hz, 2H, $H_{5''}$), 3.38 (t, J = 6.3 Hz, 2H, $H_{8''}$), 3.07 (t, J = 6.2 Hz, 2H, $H_{9''}$), 2.65 (q, J = 7.2 Hz, 2H, H_d), 2.46 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.01 (t, J = 7.2 Hz, 3H, H_e), 0.97 (s, 3H, $H_{10''}$), 0.84 (s, 3H, $H_{11''}$). **^{13}C NMR** (226 MHz, D_2O, CH_3OH = 51.59 ppm)**:** δ 211.51 (C_c), 198.64 (C_a), 177.47 ($C_{4''}$), 176.76 ($C_{7''}$), 155.15 (C_6), 151.23 (C_4), 148.38 (C_2), 144.83 (C_8), 121.13 (C_5), 89.98 ($C_{1'}$), 86.34 (d, J = 8.3 Hz, $C_{4'}$), 76.92 – 76.76 (m, $C_{2'}$, $C_{3'}$, and $C_{3''}$), 74.62 (d, J = 5.8 Hz, $C_{1''}$), 67.86 ($C_{5'}$), 58.39 (dt, J = 59.3, 20.3 Hz, C_b), 41.09 ($C_{8''}$), 39.42 (C_d), 38.12 ($C_{5''}$ or $C_{6''}$), 38.03 ($C_{5''}$ or $C_{6''}$), 31.26 ($C_{9''}$), 23.64 ($C_{10''}$), 20.90 ($C_{11''}$), 9.56 (C_e). **HR-ESI-MS [M-H]$^-$:** calculated for $C_{26}H_{41}N_7O_{18}P_3S$, *m/z* 864.1447, found *m/z* 864.1435.

2-Methyl-3-oxopentanoyl coenzyme A. **^{1}H NMR** (600 MHz, D_2O = 4.79 ppm): δ 8.65 (s, 1H, H_8), 8.40 (s, 1H, H_2), 6.19 (d, J = 5.5 Hz, 1H, $H_{1'}$), 4.90 – 4.83 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.65 – 4.59 (m, 1H, $H_{4'}$), 4.36 – 4.23 (m, 2H, $H_{5'}$), 4.11 (q, 0H, J = 7.0 Hz, H_b), 4.05 (s, 1H, $H_{3''}$), 3.89 (dd, J = 9.8, 4.9 Hz, 1H, pro-R-$H_{1''}$), 3.63 (dd, J = 9.8, 4.7 Hz, 1H, pro-S-$H_{1''}$), 3.47 (t, J = 6.7 Hz, 2H, $H_{3''}$), 3.36 (t, J = 6.3 Hz, 2H, $H_{8''}$), 3.06 (t, J = 6.3 Hz, 2H, $H_{9''}$), 2.67 (qd, J = 7.2, 3.6 Hz, 2H, H_d), 2.44 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.32 (d, J = 7.8 Hz, 3H, H_g), 0.99 (t, J = 7.2 Hz, 3H, H_e), 0.96 (s, 3H, $H_{10''}$), 0.83 (s, 3H, $H_{11''}$). **^{13}C NMR** (226 MHz, D_2O, CH_3OH = 51.59 ppm)**:** δ 214.21 (C_c), 202.80 (C_a), 177.46 ($C_{4''}$), 176.70 ($C_{7''}$), 153.79 (C_6), 151.33 (C_4), 149.26 (C_2), 144.54 (C_8), 121.16 (C_5), 89.85 ($C_{1'}$), 86.32 (d, J = 8.6 Hz, $C_{4'}$), 76.91 – 76.71 (m, $C_{2'}$, $C_{3'}$, and $C_{3''}$), 74.61 (d, J = 6.3 Hz, $C_{1''}$), 67.87 ($C_{5'}$), 62.78 (t, J = 19.9 Hz, C_b), 41.07 ($C_{8''}$), 38.12 ($C_{5''}$ or $C_{6''}$), 38.04 ($C_{5''}$ or $C_{6''}$), 31.15 ($C_{9''}$), 23.63 ($C_{10''}$), 20.90 ($C_{11''}$), 15.98 (C_g), 9.74 (C_e). **HR-ESI-MS [M-H]$^-$:** calculated for $C_{27}H_{43}N_7O_{18}P_3S$, *m/z* 878.1604, found *m/z* 878.1589.

3-Oxohexanoyl coenzyme A. **^{1}H NMR** (600 MHz, D_2O = 4.79 ppm): δ 8.65 (s, 1H, H_8), 8.39 (s, 1H, H_2), 6.21 (d, J = 5.9 Hz, 1H, $H_{1'}$), 4.91 – 4.84 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.64 – 4.61 (m, 1H, $H_{4'}$), 4.33 – 4.23 (m, 2H, $H_{5'}$), 4.05 (s, 1H, $H_{3''}$), 3.88 (dd, J = 9.8, 4.9 Hz, 1H, pro-R-$H_{1''}$), 3.61 (dd, J = 9.8, 4.8 Hz, 1H, pro-S-$H_{1''}$), 3.47 (t, J = 6.5 Hz, 2H, $H_{3''}$), 3.37 (t, J = 6.3 Hz, 2H, $H_{8''}$), 3.07 (t, J = 6.3 Hz, 2H, $H_{9''}$), 2.59 (t, J = 7.2 Hz, 2H, H_d), 2.46 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.55 (h, J = 7.4 Hz, 2H, H_e), 0.93 (s, 3H, $H_{10''}$), 0.87 (t, J = 7.4 Hz, 3H, H_f), 0.82 (s, 3H, $H_{11''}$). **^{13}C NMR** (226 MHz, D_2O, CH_3OH = 51.59 ppm)**:** δ 211.15 (C_c), 198.56 (C_a), 177.47 ($C_{4''}$), 176.76 ($C_{7''}$), 155.11 (C_6), 151.62 (C_2 or C_4), 151.55 (C_2 or C_4), 143.95 (C_8), 121.24 (C_5), 89.58 ($C_{1'}$), 86.29 ($C_{4'}$), 76.91 ($C_{2'}$ or $C_{3'}$), 76.84 ($C_{3''}$), 76.65 ($C_{2'}$ or $C_{3'}$), 74.58 ($C_{1''}$), 67.92 ($C_{5'}$), 58.66 (m, C_b), 47.78 (C_d), 41.09 ($C_{8''}$), 38.12 ($C_{5''}$ or $C_{6''}$), 38.05 ($C_{5''}$ or $C_{6''}$), 31.26 ($C_{9''}$), 23.62 ($C_{10''}$), 20.86 ($C_{11''}$), 19.25 (C_e), 15.40 (C_f). **HR-ESI-MS [M-H]$^-$:** calculated for $C_{27}H_{43}N_7O_{18}P_3S$, *m/z* 878.1604, found *m/z* 878.1591.

2-Methyl-3-oxohexanoyl coenzyme A. **^{1}H NMR** (600 MHz, D_2O = 4.79 ppm): δ 8.61 (s, 1H, H_8), 8.33 (s, 1H, H_2), 6.22 (d, J = 6.5 Hz, 1H, $H_{1'}$), 4.91 – 4.84 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.64 – 4.59 (m, 1H, $H_{4'}$), 4.28 – 4.24 (m, 2H, $H_{5'}$), 4.04 (s, 1H, $H_{3''}$), 3.86 (dd, J = 9.7, 4.6 Hz, 1H, pro-R-$H_{1''}$), 3.59 (dd, J = 9.5, 4.2 Hz, 1H, pro-S-$H_{1''}$), 3.46 (t, J = 6.6 Hz, 2H, $H_{3''}$), 3.36 (t, J = 6.3 Hz, 2H, $H_{8''}$), 3.06 (t, J = 6.2 Hz, 2H, $H_{9''}$), 2.63 (td, J = 6.9, 3.9 Hz, 2H, H_d), 2.44 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.53 (h, J = 7.3 Hz, 2H, H_e), 1.31 (s, 3H, H_g), 0.93 (s, 3H, $H_{10''}$), 0.85 (t, J = 7.4 Hz, 3H, H_f), 0.79 (s, 3H, $H_{11''}$). **^{13}C NMR** (226 MHz, D_2O, CH_3OH = 51.59 ppm)**:** δ 213.71 (C_c), 202.69 (C_a), 177.45 ($C_{4''}$), 176.68 ($C_{7''}$), 157.53 (C_6), 154.43 (C_2), 151.95 (C_4), 142.92 (C_8), 121.33 (C_5), 89.10 ($C_{1'}$), 86.23 (d, J = 9.2 Hz, $C_{4'}$), 77.00 (d, J = 4.9 Hz, $C_{2'}$ or $C_{3'}$), 76.81 ($C_{3''}$), 76.50 (s, $C_{2'}$ or $C_{3'}$),

74.56 (d, J = 6.4 Hz, $C_{1''}$), 67.98 (d, J = 5.5 Hz, $C_{5'}$), 63.00 (t, J = 19.7 Hz, C_b), 46.44 (C_d), 41.05 ($C_{8''}$), 38.12 ($C_{5''}$ or $C_{6''}$), 38.05 ($C_{5''}$ or $C_{6''}$), 31.14 ($C_{9''}$), 23.57 ($C_{10''}$), 20.83 ($C_{11''}$), 19.30 (C_e), 15.87 (C_g), 15.40 (C_f). **HR-ESI-MS [M-H]$^-$:** calculated for $C_{28}H_{45}N_7O_{18}P_3S$, *m/z* 892.1760, found *m/z* 892.1751.

2-Ethyl-3-oxohexanoyl coenzyme A. **^{1}H NMR** (600 MHz, D_2O = 4.79 ppm): δ 8.69 (s, 1H, H_8), 8.35 (s, 1H, H_2), 6.22 (d, J = 6.3 Hz, 1H, $H_{1'}$), 4.93 – 4.83 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.62 (s, 1H, $H_{4'}$), 4.27 (s, 2H, $H_{5'}$), 4.05 (s, 1H, $H_{3''}$), 4.02 (t, J = 7.3 Hz, 0H, H_b), 3.87 (d, J = 9.5 Hz, 1H, pro-R-$H_{1''}$), 3.59 (d, J = 9.5 Hz, 1H, pro-S-$H_{1''}$), 3.46 (t, J = 6.6 Hz, 2H, $H_{3''}$), 3.36 (t, J = 6.3 Hz, 2H, $H_{8''}$), 3.07 (t, J = 6.3 Hz, 2H, $H_{9''}$), 2.71 – 2.56 (m, 2H, H_d), 2.43 (t, J = 6.7 Hz, 2H, $H_{6''}$), 1.83 (qd, J = 7.2, 4.7 Hz, 2H, H_g), 1.53 (h, J = 7.2 Hz, 2H, H_e), 0.93 (s, 3H, $H_{10''}$), 0.90 – 0.82 (m, 6H, H_f and H_h), 0.80 (s, 3H, $H_{11''}$). **^{13}C NMR** (226 MHz, D_2O, CH_3OH = 51.59 ppm, DSS = 57.08, 21.82, 17.69, 0 ppm)**:** δ 213.09 (C_c), 201.56 (C_a), 177.48 ($C_{4''}$), 176.68 ($C_{7''}$), 155.37 (C_6), 151.66 (C_2 or C_4), 151.34 (C_2 or C_4), 143.95 (C_8), 121.31 (C_5), 89.58 ($C_{1'}$), 86.34 (d, J = 7.8 Hz, $C_{4'}$), 76.96 (d, J = 4.9 Hz, $C_{2'}$ or $C_{3'}$), 76.84 ($C_{3''}$), 76.64 (d, J = 5.0 Hz, $C_{2'}$ or $C_{3'}$), 74.60 (d, J = 5.6 Hz, $C_{1''}$), 70.55 (t, J = 18.8 Hz, C_b), 67.93 (d, J = 6.1 Hz, $C_{5'}$), 47.30 (C_d), 41.10 ($C_{8''}$), 38.14 ($C_{5''}$ or $C_{6''}$), 38.08 ($C_{5''}$ or $C_{6''}$), 31.20 ($C_{9''}$), 25.47 (C_g), 23.63 ($C_{10''}$), 20.87 ($C_{11''}$), 19.26 (C_e), 15.42 (C_f or C_h), 13.63 (C_f or C_h). **HR-ESI-MS [M-H]$^-$:** calculated for $C_{29}H_{47}N_7O_{18}P_3S$, *m/z* 906.1917, found *m/z* 906.1906.

Tigloyl coenzyme A. **^{1}H NMR** (600 MHz, D_2O = 4.79 ppm): δ 8.68 (s, 1H, H_8), 8.44 (s, 1H, H_2), 6.90 (q, J = 6.2 Hz, 1H, H_c), 6.21 (d, J = 6.3 Hz, 1H, $H_{1'}$), 4.91 – 4.86 (m, 2H, $H_{2'}$ and $H_{3'}$), 4.64 – 4.61 (m, 1H, $H_{4'}$), 4.35 – 4.24 (m, 2H, $H_{5'}$), 4.05 (s, 1H, $H_{3''}$), 3.90 (dd, J = 9.8, 4.8 Hz, 1H, pro-R-$H_{1''}$), 3.64 (dd, J = 9.8, 4.6 Hz, 1H, pro-S-$H_{1''}$), 3.47 (t, J = 6.6 Hz, 2H, $H_{3''}$), 3.36 (t, J = 6.3 Hz, 2H, $H_{8''}$), 3.02 (t, J = 6.4 Hz, 2H, $H_{9''}$), 2.45 (t, J = 6.6 Hz, 2H, $H_{6''}$), 1.83 – 1.80 (m, 6H, H_d, H_e), 0.97 (s, 3H, $H_{10''}$), 0.84 (s, 3H, $H_{11''}$). **^{13}C NMR** (226 MHz, D_2O, CH_3OH = 51.59 ppm)**:** δ 199.68 (C_a), 177.45 ($C_{4''}$), 176.68 ($C_{7''}$), 152.57 (C_6), 151.17 (C_4), 147.50 (C_2), 145.18 (C_8), 141.61 (C_b), 139.06 (C_c), 121.18 (C_5), 90.18 ($C_{1'}$), 86.26 (dd, J = 8.7, 2.8 Hz, $C_{4'}$), 76.99 – 76.70 (m, $C_{2'}$, $C_{3'}$, and $C_{3''}$), 74.72 (d, J = 6.1 Hz, $C_{1''}$), 67.85 (d, J = 5.0 Hz, $C_{5'}$), 41.43 ($C_{8''}$), 38.17 ($C_{5''}$ or $C_{6''}$), 38.06 ($C_{5''}$ or $C_{6''}$), 30.66 ($C_{9''}$), 23.63 ($C_{10''}$), 20.96 ($C_{11''}$), 16.61 (C_d or C_e), 14.10 (C_d or C_e). **HR-ESI-MS [M-H]$^-$:** calculated for $C_{26}H_{41}N_7O_{17}P_3S$, *m/z* 848.1493, found *m/z* 848.1483.

Quantification of assay reagents. All measurements described here were done in triplicate. CoA concentration was measured by diluting the stock aliquot in PBS buffer containing 0.3 mM DTNB and measuring the absorbance at 412 nm after incubating for 10 min. An extinction coefficient for DTNB of 14.7 $mM^{-1}\,cm^{-1}$ was measured using a standard curve with BME in the same buffer. Saturated acyl-CoA concentrations were determined in water using the extinction coefficient listed by the supplier (16.4 $mM^{-1}\,cm^{-1}$ at 260 nm). 3-oxoacyl-CoA compounds were quantified by hydrolysis with freshly prepared hydroxylamine solution. Solid hydroxylamine hydrochloride was neutralized with 1 M NaOH to give 1 M hydroxylamine. 2.0 µL 3-oxoacyl-CoA compound was added to 4 µL hydroxylamine solution and incubated for 20 min before diluting with 1.50 mL of 0.3 mM DTNB in PBS and then quantifying as described for free CoA. NADH was quantified using an extinction coefficient of 6.22 $mM^{-1}\,cm^{-1}$ at 340 nm in water at pH 12, as suggested by the manufacturer. Tigloyl-CoA and crotonyl-CoA were quantified using a literature extinction coefficient of 22.6 $mM^{-1}\,cm^{-1}$ [34].

Enzyme assays. k_{cat} and K_M were determined by fitting to initial rate data using Origin (OriginLab, Northampton, MA) using the equation:

$$v_0 = \frac{k_{cat}[S]}{K_M + [S]}$$

where v_o is the initial rate and [S] is the substrate concentration, or using the substrate inhibition equation

$$v_0 = \frac{k_{cat}[S]}{K_M + [S] + \frac{[S]^2}{K_{I,substrate}}}$$

if the calculated $K_{I,substrate}$ was less than 100 times the calculated K_M. For substrates with large K_M values, it was not possible to measure rates in the saturating region. In this case, only the catalytic efficiency k_{cat}/K_M was calculated, by fitting the following equation to the pre-saturation kinetic data.

$$v_0 = \frac{\left(k_{cat}/K_M\right)[S]}{1 + \left([S]/K_M\right)}$$

Screening of Acat thiolases. Thiolysis activity was measured by monitoring the reaction of released CoA with DTNB at 412 nm using an extinction coefficient of 14.7 mM^{-1} cm^{-1}. Assays were performed at 25°C in a 1 cm cuvette in a total volume of 600 µL containing PBS, pH 7.4, 0.3 mM DTNB, 1 mM acyl-CoA, and Acat enzyme. After pre-incubation of the substrate with DTNB, the reaction was initiated by addition of thiolase enzyme and mixing rapidly by inversion. Rate measurements were corrected for background absorbance increase that was observed after enzyme addition even in the absence of acyl-CoA substrate.

Acat2, Acat3, and Acat5 thiolases, in direction of thiolysis. Thiolysis activity was measured by monitoring the initial formation of new thioester functional groups at 232 nm as previously reported [8]. Assays were performed at 25°C in a 2 mm cuvette in a total volume of 400 µL containing 50 mM potassium phosphate, pH 8.0, CoA (0.2 mM), varying amounts of 3-oxoacyl-CoA, and thiolase enzyme. The reaction was initiated by addition of thiolase enzyme and mixing rapidly by inversion. Assays with Acat5 also included 100 mM KCl.

AsHadh and CnHadh, ketoreduction. AsHadh and CnHadh activity was measured by monitoring the initial consumption of NADH at 340 nm. Assays were performed at 30°C in a 1 cm cuvette in a total volume of 800 µL containing 100 mM imidazole hydrochloride, pH 6.3, NADH (0.1 mM), varying amounts of 3-oxoacyl-CoA compound, and AsHadh or CnHadh enzyme. The reaction was initiated by addition of 3-oxoacyl-CoA compound and mixing rapidly by inversion.

AsHadh, oxidation. For determination of stereoselectiviy, AsHadh activity was measured in the oxidative direction by by monitoring the formation of NADH at 340 nm. 3*S*-hydroxybutyryl-CoA or 3*R*-hydroxybutyryl-CoA was formed *in situ* from crotonyl-CoA and either crotonase (Crt) or PhaJ enzyme. Assays were performed at 25°C in a standard 96-well plate in a total volume of 200 µL (path length 0.6 cm) containing Tris buffer, pH 8.0, NAD^+ (0.4 mM), varying amounts of crotonyl-CoA, 1 unit of either Crt or PhaJ, and AsHadh. The reaction was initiated by addition of crotonyl-CoA and mixing rapidly by pipette.

AsEch, hydration. AsEch activity was measured in the direction of hydration of enoyl-CoA by monitoring the decrease in absorbance at 263 nm as previously reported [17]. Assays were performed at 25°C in a UV-compatible 96-well plate in a total volume of 200 µL (path length 0.6 cm) containing 50 mM Tris buffer, pH 8.0, varying amounts of enoyl-CoA, and AsEch enzyme. The reaction was initiated by addition of crotonyl-CoA or tigloyl-CoA and mixing rapidly by pipette.

TesB, thioester hydrolysis. TesB activity was measured by monitoring the reaction of released CoA with DTNB at 412nm using an extinction coefficient of 14.7 mM^{-1} cm^{-1}. Hydroxyacyl-CoA compounds were produced *in situ* by incubating 1mM oxoacyl-CoA in PBS buffer with either 1.5 mM NADH and 20 µg AsHadh or 1.5 mM NADPH and 10 µg PhaB, to form 3*S*-hydroxyacyl-CoA or 3*R*-hydroxyacyl-CoA, respectively. After incubating at room temperature for 15 min, a portion of the reaction was diluted in 10% TCA in water, filtered, and tested for completion of the reaction using the ESI-MS fraction screening method described above. The remaining unquenched samples were used as 1 mM hydroxyacyl-CoA substrate stocks for TesB activity assay.

Hydrolysis assays were performed at 25°C in a standard 96-well plate in a total volume of 200 µL (path length 0.6 cm) containing PBS, pH 7.4, 0.3 mM DTNB, varying amounts of acyl-CoA substrates, and TesB. After pre-incubation of the substrate with DTNB, the reaction was initiated by addition of TesB and mixing rapidly by pipette.

Estimation of K_{eq} of hydration of enoyl-CoA substrates. K_{eq} values were calculated by incubating 250 µM crotonyl-CoA and 0.25 µg AsEch or 250 µM tigloyl-CoA and 1.5 µg AsEch in 80 mM potassium phosphate, pH 7.4, at 37°C for 30 min. The reaction was quenched with TCA to 10% final concentration and separated by HPLC using the method for acyl-CoA characterization described above, monitoring at A_{260nm}. Adjusting for differences in extinction coefficient (ε_{260nm} = 22.6 mM^{-1} cm^{-1} for enoyl-CoA species and ε_{260nm} = 16.4 mM^{-1} cm^{-1} for hydroxyacyl-CoA species), K_{eq} of hydration was calculated as:

$$K_{eq} = 1.378 \times \frac{[peak\ area]_{\text{hydroxyacyl-CoA}}}{[peak\ area]_{\text{enoyl-CoA}}}$$

Production of organic acids in shake flasks. Chemically competent *E. coli* were transformed with the appropriate plasmids. A single colony was used to inoculate an overnight culture grown in LB with 1% w/v glucose, 50 µg/mL Cb, and 30 µg/mL Cm at 37ºC in a rotary shaker (200 rpm) for 12-16 hours. These starter cultures were used to inoculate 30 mL of the same media in 250 mL baffled culture flasks (Kimble Glass; Chicago, IL) to OD_{600}=0.05. Cultures were grown at 37ºC in a rotary shaker (200 rpm) to OD_{600} = 0.3-0.5, at which point IPTG and sodium propionate were added to concentrations of 0.2 mM and 0.2% w/v, respectively. At this time the growth temperature was reduced to 30°C.

For MES buffered cultures, sterile filtered 1 M MES-KOH stock at pH 5.0, 6.0, or 7.0 was added to LB broth after autoclaving, to a final concentration of 100 mM. The pH of the buffered LB was found to be approximately 5.5, 6.1, and 6.7, respectively.

Quantification of enoic acid titers. After 2 days of growth, 200 µL of cell culture sample was acidified with 20 uL 6M HCl and extracted with 1mL ethyl acetate containing 100 mg L^{-1} 6-bromohexanoic acid as internal standard. 40 µL of the ethyl acetate layer was added to a GC vial with a glass insert containing 80 µL of a 1:1 mixture of ethyl acetate and BSTFA. Samples

were analyzed on a Trace GC Ultra (Thermo Scientific; Waltham, MA) using an HP-5ms column (0.25 mm × 30 m, 0.25 μM film thickness, Agilent). The oven program was as follows: 80°C for 2.5 min, ramp to 225°C at 30°C/min, ramp to 325°C at 50°/min, 325°C for 0.17 min. Enoic acids were quantified by mass spectrometery (DSQII single quadrupole, Thermo Scientific; Waltham, MA) using single ion monitoring of the following ions: from 3.8-4.4 min, m/z 99 and 143 (crotonic acid TMS ester); from 4.4-5.1 min, mz/ 83, 113, and 157 (pentenoic acid and tiglic acid TMS esters); from 5.1-6.5 min, m/z 97, 157, and 171 (2-methylpentenoic acid TMS ester); from 6.5-9.5 min, m/z 117, 137, and 139 (bromohexanoic acid TMS ester). Samples were quantified relative to a standard curve of 7.8125, 15.625, 31.25, 62.5, 125, 250, 500, and 1000 mg L^{-1} of each enoic acid. Standard curve was normalized for injection volume using the internal standard.

Quantification of hydroxyacid titers. Cell culture samples (1 mL) after 2 days of growth were cleared of biomass *via* centrifugation at 20817× *g* for 2 min with an Eppendorf 5417R Centrifuge (Hamburg, Germany). 10 μL cleared media was diluted in 190 μL water containing 0.5 mM adipic acid as internal standard. Samples were filtered through a 96-well MultiScreen$_{HTS}$ plate before injecting onto an Agilent 1290 HPLC equipped with an auto-sampler, Phenomenex (Torrence, CA) Rezex-ROA Organic Acid H^+ column (150 × 4.6 mm), and Carbo-H^+ Security Guard cartridge. 0.5% v/v formic acid was used as mobile phase (0.3 mL/min, column temperature 55°C), and hydroxyacids were quantified by mass spectrometry on an Agilent 6460 triple quadrupole MS with ESI source, operating in negation ion MRM transition mode with fragmentor voltage 70V. Between 5-8 min, the following transitions and collision energies were monitored: m/z 145.1→83.1, 10V (adipic acid, internal standard); m/z 131.1→73.2, 5V (2-methyl-3-hydroxypentanoic acid); m/z 117.1→73.2, 5V (2-methyl-3-hydroxybutyric acid); m/z 117.1→59.2, 5V (2-hydroxypentanoic acid); m/z 103.1→59.2, 5V (3-hydroxybutyric acid). Samples were quantified relative to a standard curve of 7.8125, 15.625, 31.25, 62.5, 125, 250, 500, and 1000 mg L^{-1} hydroxyacid. Standards of 3-hydroxypentanoate and 2-methyl-3-hydroxypentanoate were prepared by the reduction of their corresponding 2-methyl-3-oxo esters with sodium borohydride using the method of Kalaitzakis *et al*. [35] followed by saponification in 1 M NaOH. Standard curve was normalized for injection volume using the internal standard.

Quantification of propionate. Cell culture samples (1 mL) after 2 days of growth were cleared of biomass via centrifugation at 20817 × *g* for 2 min with an Eppendorf 5417R Centrifuge (Hamburg, Germany) and filtered before injecting onto an Agilent 1260 HPLC equipped with Bio-Rad Aminex HPX87H column, Cation H+ Cartridge Micro-Guard column, and refractive index detector. 10 mM H_2SO_4 was used as mobile phase (0.7 mL/min, column temperature 60°C) and propionate was quantified by peak area of the refractive index signal in comparison to a standard curve of 0.125, 0.25, 0.5, 1, and 2 g L^{-1} sodium propionate.

Production and quantification of PHA polymers. Cell culture for PHA production was carried out in the same way as hydroxyacid production with the appropriate plasmids, except LB media contained 2% w/v glucose rather than 1%.

For testing Poly1-4 plasmids, 10 mL cell culture samples after 1 day of growth were harvested by centrifugation. The cell pellet was resuspended in 1 mL conc. H_2SO_4, transferred to a glass vial, and heated at 90°C for 1 hour to depolymerize any polyesters before diluting 1:10 into ice-cold water. The aqueous sample was cleared of debris by centrifugation and then injected onto an Agilent 1260 HPLC equipped with Bio-Rad Aminex HPX87H column, Cation H+ Cartridge Micro-Guard column, with detection of enoic acid products at 240 nm.

For testing Poly1B, 7B, and 8B, 10 mL cell culture samples after 1 day of growth were harvested by centrifugation. The cell pellet was resuspended in 750 µL 20% v/v. conc. HCl in 1-propanol and transferred to a glass vial. 750 µL dichloroethane was added and the mixture was heated at 100°C for 2 hours. The dichloroethane layer was diluted in a sample vial in ethyl acetate and analyzed on a Trace GC Ultra using the same oven program described for enoic acid quantification. The MS was operated in scan mode between m/z 35 and 300.

Cell lysate assays for PHA production strains. 10 mL cell culture after 1 day of growth was harvest by centrifugation, and the cell pellet was resuspended in 4 mL of cold Buffer 2A supplemented with 1 mM TCEP and sonicated for 2 min. The lysate was cleared by centrifugation and the supernatant tested for the presence of active enzyme with the following assays. All assays were carried out at 25°C with 200 µL total volume in a 96-well plate (path length 0.6 cm). Protein concentration of soluble lysate was measured with Bio-Rad Protein Assay using BSA as a standard.

Acat2/Acat3, in direction of thiolysis. 50 mM potassium phosphate, pH 8.0, CoA (60 µM), 2-methyl-3-oxobutyryl-CoA (0.2 mM), and lysate. The reaction was initiated by addition of CoA and monitored at 232nm in a UV-compatible 96-well plate, $\varepsilon_{232nm} = 4500\ M^{-1}cm^{-1}$.

AsHadh, ketoreduction. 100 mM MES-KOH, pH 6.5, NADH (0.2 mM), 3-oxobutyryl-CoA (0.1 mM), and lysate. The reaction was initiated by addition of 3-oxobutyryl-CoA and monitored at 340nm, $\varepsilon_{340nm} = 6220\ M^{-1}cm^{-1}$.

AsEch, hydration. AsEch activity was measured by coupling the hydration of crotonyl-CoA to stereoselective oxidation of 3*S*-hydroxybutyryl-CoA, to distinguish from PhaJ activity. The assay contained 50 mM Tris, pH 7.6, crotonyl-CoA (0.1 mM), NAD^+ (0.2 mM), 1 unit of Hbd enzyme, and lysate. The reaction was initiated by addition of crotonyl-CoA and monitored at 340 nm, $\varepsilon_{340nm} = 6220\ M^{-1}cm^{-1}$.

PhaJ, hydration. PhaJ activity was measured by coupling the hydration of crotonyl-CoA to stereoselective oxidation of 3*R*-hydroxybutyryl-CoA, to distinguish from AsEch activity. The assay contained 50 mM Tris, pH 7.6, crotonyl-CoA (0.1 mM), $NADP^+$ (0.2 mM), 1 unit of PhaB enzyme, and lysate. The reaction was initiated by addition of crotonyl-CoA and monitored at 340nm, $\varepsilon_{340nm} = 6220\ M^{-1}cm^{-1}$.

PhaC, CoA release. PhaC activity was measured by release of CoA upon polymerization of *in situ* generated 3*R*-hydroxybutyryl-CoA. The assay contained 0.3 mM DTNB in PBS, crotonyl-CoA (0.1 mM), 1 unit of PhaJ enzyme, and lysate. The reaction was initiated by addition of crotonyl-CoA and monitored at 412 nm, $\varepsilon_{412nm} = 14.7\ mM^{-1}\ cm^{-1}$.

2.3 Results

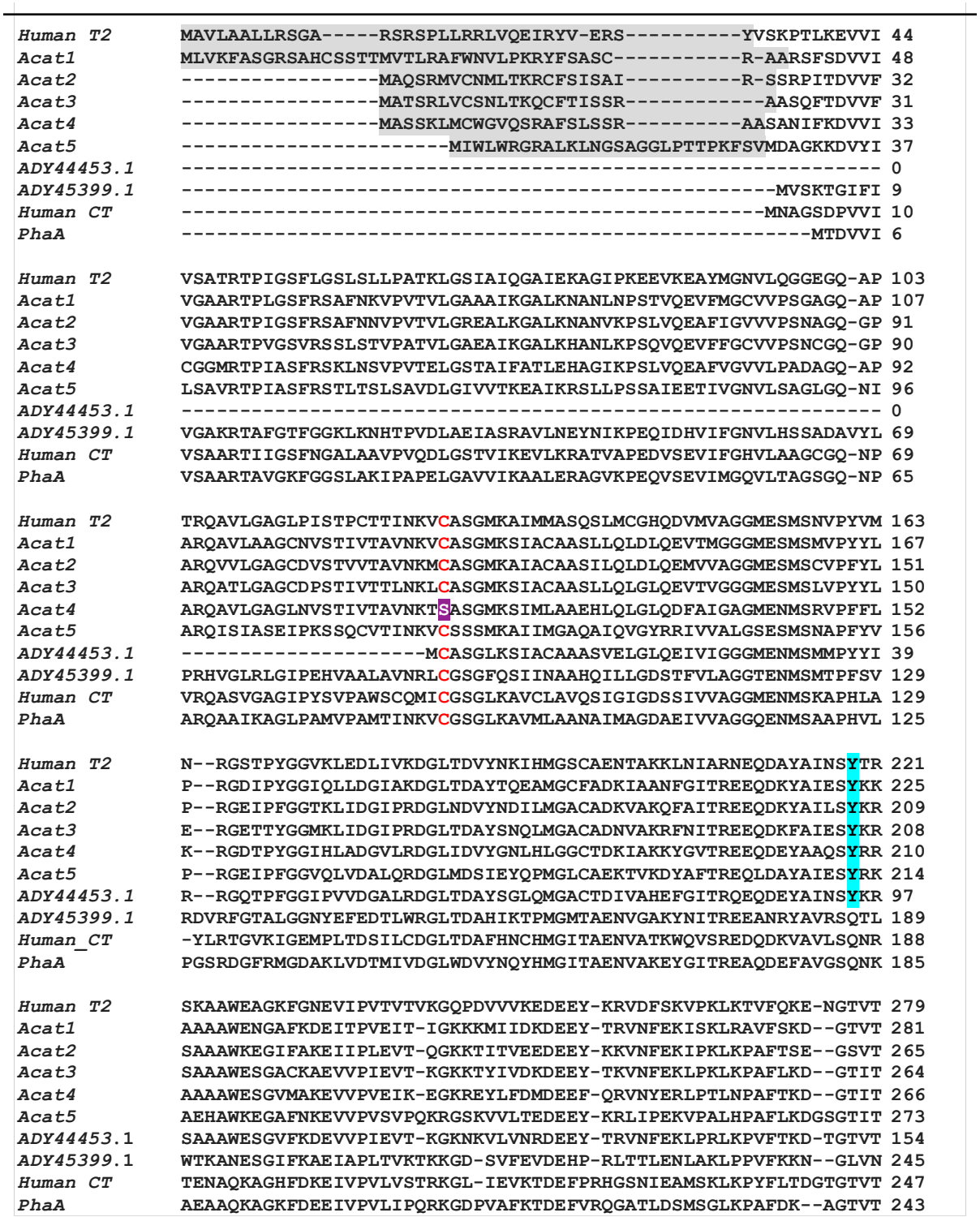
Identification and characterization of thiolase enzymes from *Ascaris suum*. Because branched acid production has been observed in the mitochondria of *A. suum* muscle tissue, we sought out *A. suum* thiolase candidates expected to be targeted to that organelle. The human mitochondrial thiolase T2 functions in the context of the degradation of isoleucine, a branched amino acid, and has previously been shown to catalyze the thiolysis of 2-methyl-3-oxobutyryl-CoA with high efficiency [8]. Using T2 (UniProtKB P24752) as a search query against the Transcriptome Shotgun Assembly database in GenBank [36, 37], we identified seven thiolase candidates (*Figure 2.2*), six of which had sequence identity greater than 50%. Accession

ADY44453.1 was removed from further study as it was missing the N-terminal region of other thiolases and might be a truncated contig. We reasoned that accession ADY45399.1 was likely a cytosolic isozyme due to its greater similarity to human cytosolic thiolase CT, lack of a mitochondrial targeting sequence, and lack of characteristic residues typically found in mitochondrial thiolase enzymes, including those implicated in the binding of structural K^+ and Cl^- ions [8]. The other five thiolases, which we named Acat1 through Acat5, were predicted to be targeted to the mitochondria by both MitoProt and TargetP prediction algorithms [38, 39], although both programs assigned only ~65% certainty for Acat5. Structural studies of human T2 by Haapalainen *et al.* have shown the importance of removing the mitochondrial targeting sequence at a specific residue for optimum heterologous expression [8]. By comparing aligned sequences of the candidate thiolases with the optimum cleavage site of T2 and invoking, when possible, either the "R-2" or "R-10" rules that have been observed to govern mitochondrial processing peptidase (MPP) and inner membrane peptidase (MIP) site specificities [40], we chose the most likely sequence cleavage site for each thiolase and synthesized codon-optimized genes for heterologous expression in *E. coli* (*Figure 2.3*).

Following His_6-tag removal, SEC purification revealed differences in the biological assemblies observed for different isozymes and their associated catalytic activities (*Figure 2.4AB*). While all five thiolases eluted with a peak at around 60 mL, Acat1 and Acat3 had an additional peak at 69 mL. Thiolase enzyme activity was only associated with the first peak, which could reasonably correspond to a tetrameric assembly as is typically found for biosynthetic thiolases in solution [8]. These fractions were concentrated and stored at -80°C for further assays with minimal activity loss. The second SEC peak in Acat1 and Acat3 may correspond to a lower-order assembly, such as a dimer. We also found that Acat3 stability was sensitive to concentration. A dilute Acat3 solution lost ~90% of activity after 72 hours at 4°C, while concentrated protein lost only ~20%. We re-analyzed a sample of the active SEC fractions from the first peak which had been diluted to < 0.1 mg/mL and allowed to stand for 72 hours at 4°C, and we found that this protein eluted almost exclusively at 69 mL (*Figure 2.4B*), suggesting that activity loss of Acat3 is related to dissociation of a higher-order assembly. Concentrating these fractions did not recover enzymatic activity.

Purified enzymes were screened for activity with acetyl-CoA, propionyl-CoA, and butyryl-CoA by CoA release with DTNB and compared to the canonical thiolase PhaA from *Cupriavidus necator*, a PHA-accumulating bacterial species [41, 42] (*Figure 2.5A*). Although butyryl-CoA activity could not be detected above background with any of the thiolases, Acat2 and Acat3 showed higher activity with propionyl-CoA than acetyl-CoA. Acat1 also displayed activity with propionyl-CoA condensation to a lesser degree, but Acat5 showed no activity with propionyl-CoA and resembled PhaA in its specificity for acetyl-CoA. In Acat4, a Ser is found in position 105 in place of the catalytically essential Cys, which is conserved in all other members of the thiolase superfamily (*Figure 2.2*) [43]. Acat4 showed no activity as a thiolase in the synthetic or thiolytic directions, and its function is unknown. Finally, since Acat1-5 all contained conserved residues that have been implicated in the binding of KCl in human T2 [8], we checked for the effect of KCl on thiolase activity (*Figure 2.5B*). Only Acat1 activity was augmented by KCl addition, by 2.3-fold compared to a 4-fold increase reported in T2. Acat2 and Acat5 showed no substantial difference with KCl, and Acat3 activity decreased by 47% with 100 mM KCl present.

Assays of thiolase activity in the synthetic direction are not adequate for characterizing mixed substrate condensation since multiple products can be formed, and the CoA release assay


Human T2 MAVLAALLRSGA-----RSRSPLLRRLVQEIRYV-ERS----------YVSKPTLKEVVI 44
Acat1 MLVKFASGRSAHCSSTTMVTLRAFWNVLPKRYFSASC----------R-AARSFSDVVI 48
Acat2 ----------MAQSRMVCNMLTKRCFSISAI----------R-SSRPITDVVF 32
Acat3 ----------MATSRLVCSNLTKQCFTISSR----------AASQFTDVVF 31
Acat4 ----------MASSKLMCWGVQSRAFSLSSR----------AASANIFKDVVI 33
Acat5 ----------MIWLWRGRALKLNGSAGGLPTTPKFSVMDAGKKDVYI 37
ADY44453.1 ---------- 0
ADY45399.1 ----------MVSKTGIFI 9
Human CT ----------MNAGSDPVVI 10
PhaA ----------MTDVVI 6
Human T2 VSATRTPIGSFLGSLSLLPATKLGSIAIQGAIEKAGIPKEEVKEAYMGNVLQGGEGQ-AP 103
Acat1 VGAARTPLGSFRSAFNKVPVTVLGAAAIKGALKNANLNPSTVQEVFMGCVVPSGAGQ-AP 107
Acat2 VGAARTPIGSFRSAFNNVPVTVLGREALKGALKNANVKPSLVQEAFIGVVVPSNAGQ-GP 91
Acat3 VGAARTPVGSVRSSLSTVPATVLGAEAIKGALKHANLKPSQVQEVFFGCVVPSNCGQ-GP 90
Acat4 CGGMRTPIASFRSKLNSVPVTELGSTAIFATLEHAGIKPSLVQEAFVGVVLPADAGQ-AP 92
Acat5 LSAVRTPIASFRSTLTSLSAVDLGIVVTKEAIKRSLLPSSAIEETIVGNVLSAGLGQ-NI 96
ADY44453.1 ---------- 0
ADY45399.1 VGAKRTAFGTFGGKLKNHTPVDLAEIASRAVLNEYNIKPEQIDHVIFGNVLHSSADAVYL 69
Human CT VSAARTIIGSFNGALAAVPVQDLGSTVIKEVLKRATVAPEDVSEVIFGHVLAAGCGQ-NP 69
PhaA VSAARTAVGKFGGSLAKIPAPELGAVVIKAALERAGVKPEQVSEVIMGQVLTAGSGQ-NP 65
Human T2 TRQAVLGAGLPISTPCTTINKVCASGMKAIMMASQSLMCGHQDVMVAGGMESMSNVPYVM 163
Acat1 ARQAVLAAGCNVSTIVTAVNKVCASGMKSIACAASLLQLDLQEVTMGGGMESMSMVPYYL 167
Acat2 ARQVVLGAGCDVSTVVTAVNKMCASGMKAIACAASILQLDLQEMVVAGGMESMSCVPFYL 151
Acat3 ARQATLGAGCDPSTIVTTLNKLCASGMKSIACAASLLQLGLQEVTVGGGMESMSLVPYYL 150
Acat4 ARQAVLGAGLNVSTIVTAVNKTSASGMKSIMLAAEHLQLGLQDFAIGAGMENMSRVPFFL 152
Acat5 ARQISIASEIPKSSQCVTINKVCSSSMKAIIMGAQAIQVGYRRIVVALGSESMSNAPFYV 156
ADY44453.1 ----------MCASGLKSIACAAASVELGLQEIVIGGGMENMSMMPYYI 39
ADY45399.1 PRHVGLRLGIPEHVAALAVNRLCGSGFQSIINAAHQILLGDSTFVLAGGTENMSMTPFSV 129
Human CT VRQASVGAGIPYSVPAWSCQMICGSGLKAVCLAVQSIGIGDSSIVVAGGMENMSKAPHLA 129
PhaA ARQAAIKAGLPAMVPAMTINKVCGSGLKAVMLAANAIMAGDAEIVVAGGQENMSAAPHVL 125
Human T2 N--RGSTPYGGVKLEDLIVKDGLTDVYNKIHMGSCAENTAKKLNIARNEQDAYAINSYTR 221
Acat1 P--RGDIPYGGIQLLDGIAKDGLTDAYTQEAMGCFADKIAANFGITREEQDKYAIESYKK 225
Acat2 P--RGEIPFGGTKLIDGIPRDGLNDVYNDILMGACADKVAKQFAITREEQDKYAILSYKR 209
Acat3 E--RGETTYGGMKLIDGIPRDGLTDAYSNQLMGACADNVAKRFNITREEQDKFAIESYKR 208
Acat4 K--RGDTPYGGIHLADGVLRDGLIDVYGNLHLGGCTDKIAKKYGVTREEQDEYAAQSYRR 210
Acat5 P--RGEIPFGGVQLVDALQRDGLMDSIEYQPMGLCAEKTVKDYAFTREQLDAYAIESYRK 214
ADY44453.1 R--RGQTPFGGIPVVDGALRDGLTDAYSGLQMGACTDIVAHEFGITRQEQDEYAINSYKR 97
ADY45399.1 RDVRFGTALGGNYEFEDTLWRGLTDAHIKTPMGMTAENVGAKYNITREEANRYAVRSQTL 189
Human_CT -YLRTGVKIGEMPLTDSILCDGLTDAFHNCHMGITAENVATKWQVSREDQDKVAVLSQNR 188
PhaA PGSRDGFRMGDAKLVDTMIVDGLWDVYNQYHMGITAENVAKEYGITREAQDEFAVGSQNK 185
Human T2 SKAAWEAGKFGNEVIPVTVTVKGQPDVVVKEDEEY-KRVDFSKVPKLKTVFQKE-NGTVT 279
Acat1 AAAAWENGAFKDEITPVEIT-IGKKKMIIDKDEEY-TRVNFEKISKLRAVFSKD--GTVT 281
Acat2 SAAAWKEGIFAKEIIPLEVT-QGKKTITVEEDEEY-KKVNFEKIPKLKPAFTSE--GSVT 265
Acat3 SAAAWESGACKAEVVPIEVT-KGKKTYIVDKDEEY-TKVNFEKLPKLKPAFLKD--GTIT 264
Acat4 AAAAWESGVMAKEVVPVEIK-EGKREYLFDMDEEF-QRVNYERLPTLNPAFTKD--GTIT 266
Acat5 AEHAWKEGAFNKEVVPVSVPQKRGSKVVLTEDEEY-KRLIPEKVPALHPAFLKDGSGTIT 273
ADY44453.1 SAAAWESGVFKDEVVPIEVT-KGKNKVLVNRDEEY-TRVNFEKLPRLKPVFTKD-TGTVT 154
ADY45399.1 WTKANESGIFKAEIAPLTVKTKKGD-SVFEVDEHP-RLTTLENLAKLPPVFKKN--GLVN 245
Human CT TENAQKAGHFDKEIVPVLVSTRKGL-IEVKTDEFPRHGSNIEAMSKLKPYFLTDGTGTVT 247
PhaA AEAAQKAGKFDEEIVPVLIPQRKGDPVAFKTDEFVRQGATLDSMSGLKPAFDK--AGTVT 243

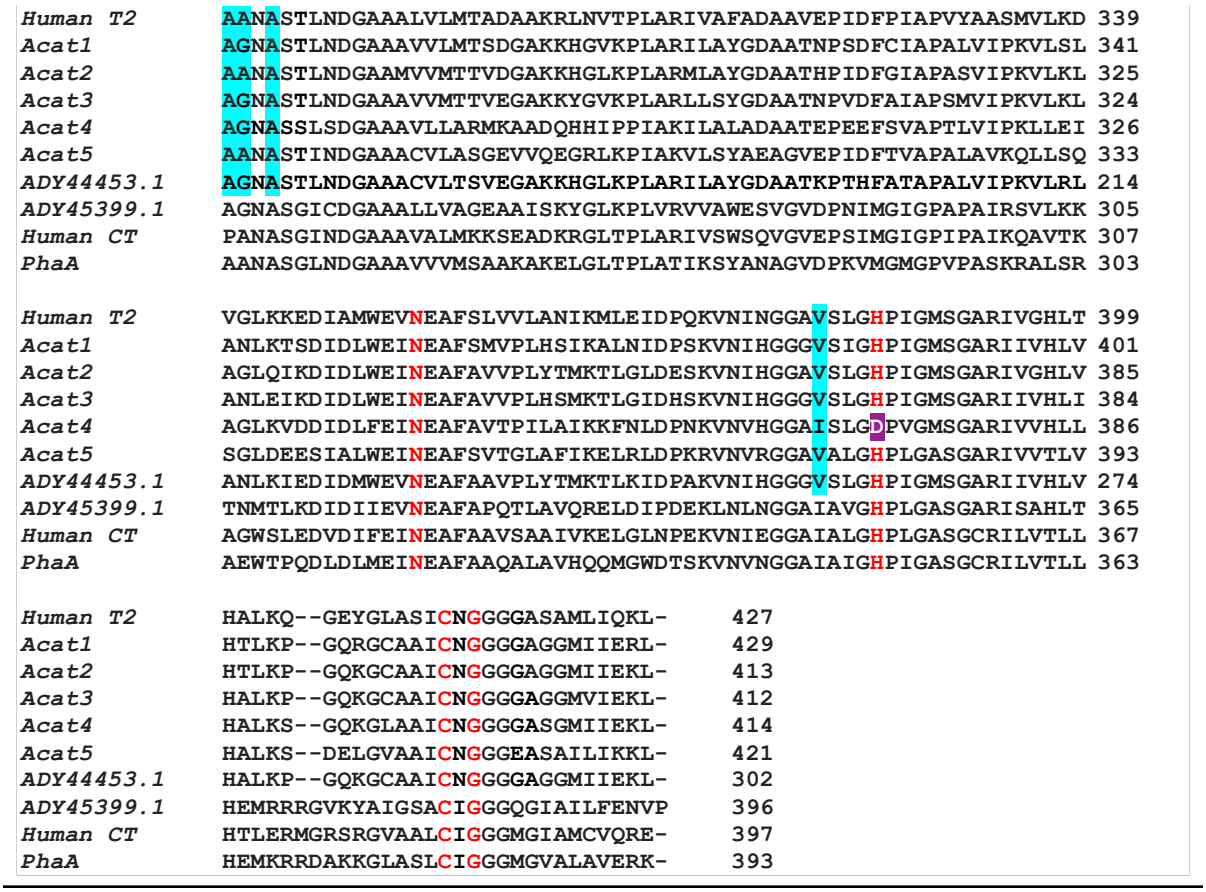


Figure 2.2. *Alignment of putative thiolases from Ascaris suum with human T2 (mitochondrial) and CT (cytosolic) thiolases and C. necator PhaA. Predicted cleaved mitochondrial targeting peptides are highlighted in* ***gray****. Conserved residues of the thiolase superfamily involved in covalent or acid-base catalysis or oxyanion stabilization are in* ***red, bold type****, with exceptional residues of Acat4 highlighted in* ***maroon****. Residues implicated in the binding of* K^+ *in T2 [8] are highlighted in* ***blue****.*

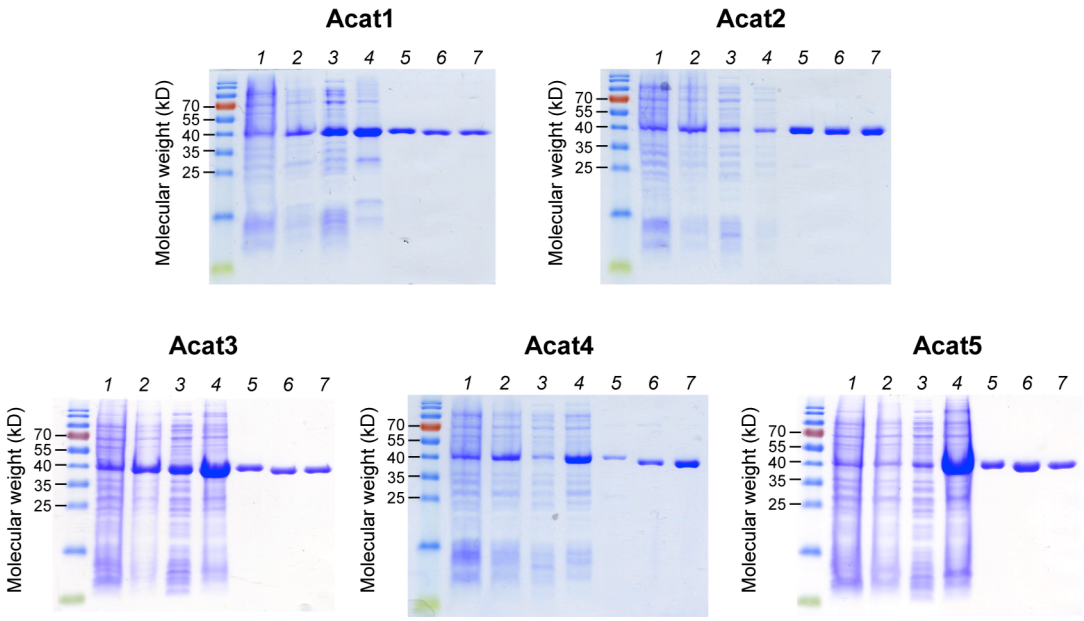


Figure 2.3 *SDS-PAGE of purification of A. suum thiolases Acat1 through Acat5: 1. Pre-induction, 2. Post-induction, 3. Soluble lysate, 4. Insoluble lysate, 5. Ni-NTA chromatography, 6. TEV cleavage, 7. Size exclution chromatography.*

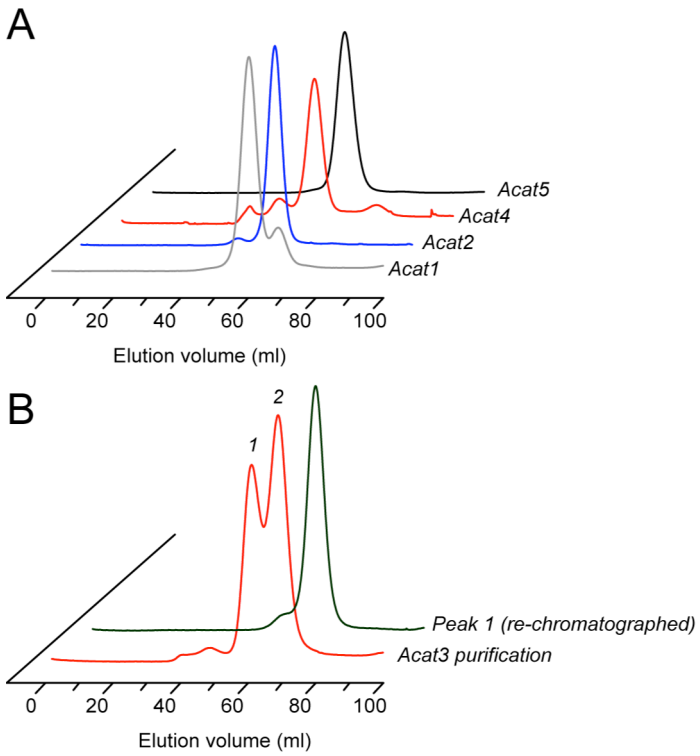


Figure 2.4. *Size exclusion chromatography of A. suum thiolases. (A) Acat1, 2, 4, and 5. (B) Acat3 eluted as two peaks, with only Peak 1 containing active enzyme. Diluted peak 1 fractions lost activity after 72 hours and eluted at a later time when re-chromatographed.*

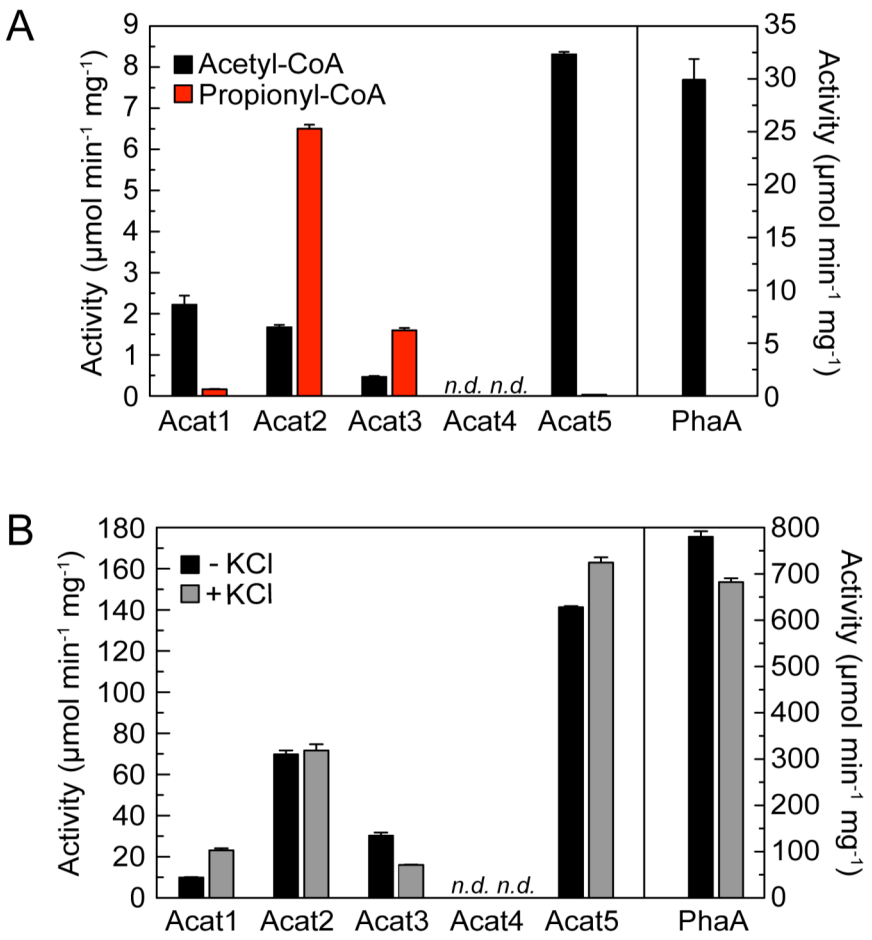


Figure 2.5. *Characterization of thiolases from A. suum. (A) Activity of Claisen-condensation with 1 mM acetyl-CoA or 1 mM propionyl-CoA. n.d. = not detectable. (B) Activity of thiolysis with 200 µM 3-oxobutyryl-CoA and 200 µM CoA, with or without addition of 100 mM KCl. n.d. = not detectable.*

with DTNB is further limited by its thermodynamics and the possible inactivation of enzyme by DTNB reagent. We selected thiolases Acat2, Acat3, and Acat5 for a more extensive kinetic characterization in the direction of thiolysis, using a set of chemically synthesized 3-oxoacyl-CoA substrates (*Figure 2.6, Table 2.1*). For Acat2 (*Figure 2.7*) and Acat3 (*Figure 2.8*), catalytic efficiencies with all 3-oxoacyl-CoA substrates formed from combinations of 2- or 3-carbon units were of similar magnitude, but the individual k_{cat} and K_M values were both higher for α-methyl substrates. This is a similar to human T2 thiolase, although the observed K_M values were significantly higher than those observed in T2 and other short chain thiolases [8, 44, 45]. These higher K_M values may reflect an evolved preference to release 3-oxoacyl-CoA product since biosynthesis is expected to be the physiologically relevant direction of the reaction in *A. suum* mitochondria. It must also be noted, however, that the K_M values we report are observed with the concentration of free CoA maintained well above saturation, at 200 μM. It has been hypothesized that *A. suum* maintains very low concentrations of CoA in the mitochondrial matrix [46].

Some substrates showed substrate inhibition behavior at higher concentrations. The initial rate data for these substrates fit poorly to the standard Michaelis-Menten equation but fit well to the modified form for substrate inhibition. This effect was most prominent in the 6-carbon chain substrates 3-oxohexanoyl-CoA and 2-methyl-3-oxohexanoyl-CoA, which also had lower catalytic efficiencies in general. Although this effect has not been reported for thiolases in particular to our knowledge, approximately 20% of enzymes show substrate inhibition behavior [47]. One explanation is the bi-bi ping-pong catalytic mechanism of thiolases [48], in which the enzyme is first acylated by 3-oxoacyl-CoA substrate, followed by transthioesterification with CoA. At high substrate concentrations, 3-oxoacyl-CoA may compete with free CoA, although this does not explain why substrate inhibition was especially potent with the largest substrates.

In contrast to Acat2 and Acat3, Acat5 had high catalytic efficiencies with linear 3-oxoacyl-CoA substrates but a catalytic defect of 100-fold in k_{cat}/K_M upon introduction of an α-methyl substituent to 3-oxobutyryl-CoA, mostly due to an increase in K_M (*Figure 2.9*). Extending the main chain by one carbon (2-methyl-3-oxopentanoyl-CoA) lowered k_{cat}/K_M by an additional 30-fold. This further catalytic defect is not simply an additive effect, because only a 4-fold decrease is observed when comparing k_{cat}/K_M of 3-oxopentanoyl-CoA to 3-oxobutyryl-CoA. Thus while Acat5 is capable of accommodating a 5-carbon linear substrate, the adjustments required in the active site must significantly aggravate any ability to accommodate an α-substituent.

Identification and characterization of ketoreductase AsHadh. Searching the Transcriptome Shotgun Assembly with human mitochondrial short chain hydroxyacyl-CoA dehydrogenase HADH2 [49] as query, we identified just one ketoreductase candidate in *A. suum*, AsHadh (ADY43626.1) . His_{10}-AsHadh was expressed and purified (*Figure 2.10*), and kinetic characterization of the enzyme (*Table 2.2, Figure 2.11A*) revealed a preference for α-methyl substrates 2-methyl-3-oxobutyryl-CoA and 2-methyl-3-oxopentanoyl-CoA compared to linear or α-ethyl substrates. This selectivity arose mostly from significantly lower K_M values (sub-micromolar) for the α-methyl substrates. This makes sense in the context of a thiolase-based biosynthetic pathway, because the concentration of 3-oxoacyl-CoA would be expected to remain low given that the thiolase reaction is highly endergonic in the direction of condensation [50].

Potent substrate inhibition was also observed for AsHadh. This phenomenon has been widely observed in NAD(P)H oxidoreductases, including lactate dehydrogenase [51], and can be accounted for by the catalytic mechanism. Kinetic studies of other ketoreductases [52] suggest an ordered-binding reaction pathway for this enzyme class. Nicotinamide cofactor binds first,

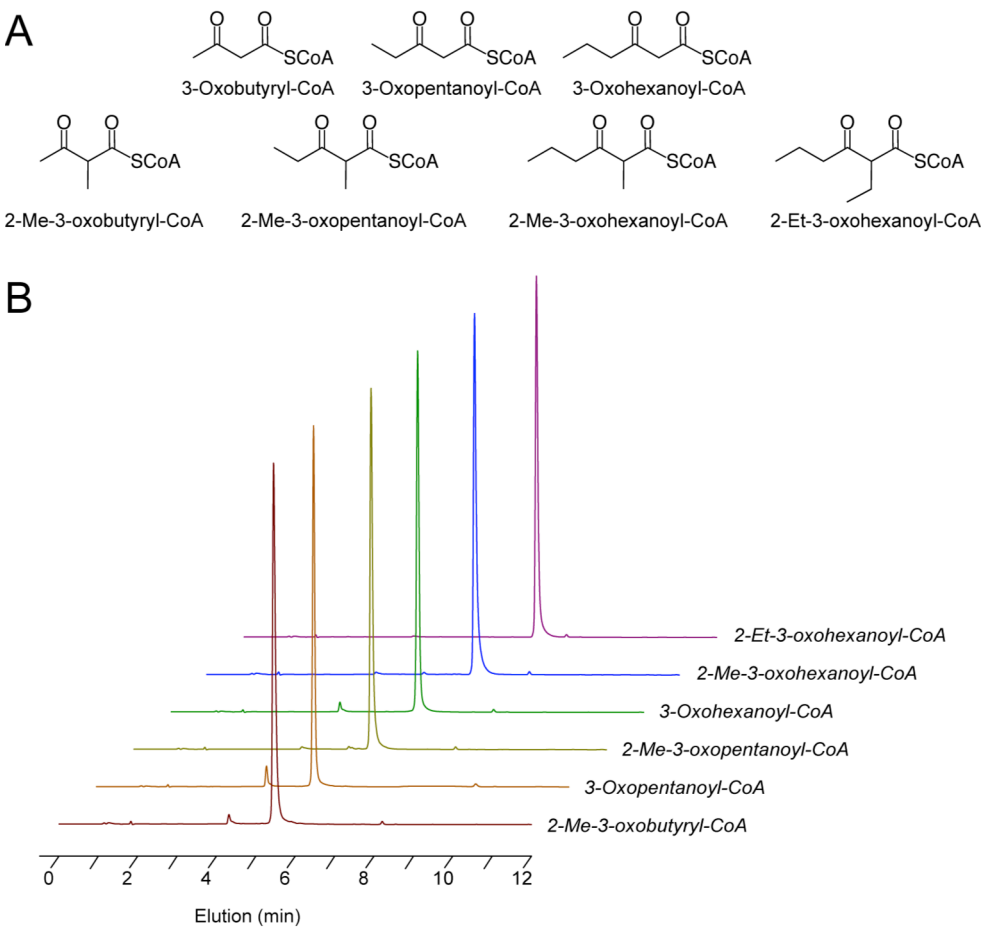


Figure 2.6. *Structures (A) and HPLC traces (B) of chemically synthesized 3-oxoacyl-CoA substrates for enzyme assays.*

Table 2.1. *Kinetics constants measured for thiolysis of 3-oxoacyl-CoA by Acat2, Acat3, and Acat5.*[a]

Enzyme	Substrate	Relative k_{cat}/K_M	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_M (μM)	$K_{I,\ substrate}$ (μM)
AsAcat2	3-Oxobutyryl-CoA	1	$(3.9 \pm 0.2) \times 10^6$	$(7.6 \pm 0.1) \times 10^1$	$(2.0 \pm 0.1) \times 10^1$	$(1.9 \pm 0.2) \times 10^3$
	3-Oxopentanoyl-CoA	0.8 ± 0.1	$(3.3 \pm 0.3) \times 10^6$	$(5.6 \pm 0.1) \times 10^1$	$(1.7 \pm 0.2) \times 10^1$	-
	3-Oxohexanoyl-CoA	0.17 ± 0.03	$(7 \pm 1) \times 10^5$	$(2.2 \pm 0.2) \times 10^1$	$(3.3 \pm 0.5) \times 10^1$	$(9 \pm 2) \times 10^2$
	2-Me-3-oxobutyryl-CoA	0.21 ± 0.03	$(8 \pm 1) \times 10^5$	$(1.9 \pm 0.1) \times 10^2$	$(2.3 \pm 0.3) \times 10^2$	-
	2-Me-3-oxopentanoyl-CoA	0.3 ± 0.06	$(1.2 \pm 0.2) \times 10^6$	$(1.9 \pm 0.09) \times 10^2$	$(1.6 \pm 0.3) \times 10^2$	-
	2-Me-3-oxohexanoyl-CoA	0.03 ± 0.01	$(1.3 \pm 0.5) \times 10^5$	$(3.5 \pm 0.7) \times 10^1$	$(2.7 \pm 0.8) \times 10^2$	$(4 \pm 1) \times 10^2$
AsAcat3	3-Oxobutyryl-CoA	1	$(4.2 \pm 0.3) \times 10^5$	$(1.39 \pm 0.02) \times 10^1$	$(3.3 \pm 0.2) \times 10^1$	-
	3-Oxopentanoyl-CoA	1.4 ± 0.2	$(5.9 \pm 0.6) \times 10^5$	$(1.00 \pm 0.02) \times 10^1$	$(1.7 \pm 0.2) \times 10^1$	-
	3-Oxohexanoyl-CoA	0.40 ± 0.04	$(1.7 \pm 0.1) \times 10^5$	5.0 ± 0.1	$(2.9 \pm 0.2) \times 10^1$	-
	2-Me-3-oxobutyryl-CoA	0.67 ± 0.09	$(2.8 \pm 0.3) \times 10^5$	$(3.9 \pm 0.1) \times 10^1$	$(1.4 \pm 0.2) \times 10^2$	-
	2-Me-3-oxopentanoyl-CoA	0.8 ± 0.1	$(3.3 \pm 0.3) \times 10^5$	$(4.6 \pm 0.1) \times 10^1$	$(1.4 \pm 0.1) \times 10^2$	-
	2-Me-3-oxohexanoyl-CoA	0.14 ± 0.05	$(6 \pm 2) \times 10^4$	$(1.2 \pm 0.2) \times 10^1$	$(2.0 \pm 0.5) \times 10^2$	$(3.3 \pm 0.9) \times 10^2$
AsAcat5	3-Oxobutyryl-CoA	1	$(5.7 \pm 0.5) \times 10^6$	$(1.00 \pm 0.02) \times 10^2$	$(1.7 \pm 0.1) \times 10^1$	-
	3-Oxopentanoyl-CoA	0.25 ± 0.04	$(1.5 \pm 0.2) \times 10^6$	$(7.6 \pm 0.5) \times 10^1$	$(5.2 \pm 0.6) \times 10^1$	$(8 \pm 1) \times 10^2$
	3-Oxohexanoyl-CoA	0.0027 ± 0.0008	$(1.6 \pm 0.4) \times 10^4$	1.9 ± 0.3	$(1.2 \pm 0.3) \times 10^2$	$(2.3 \pm 0.5) \times 10^2$
	2-Me-3-oxopentanoyl-CoA	0.009 ± 0.001	$(5.0 \pm 0.3) \times 10^4$	> 25	$(> 4) \times 10^2$	-
	2-Me-3-oxopentanoyl-CoA	0.00026 ± 0.00003	$(1.5 \pm 0.1) \times 10^3$	> 0.9	$(> 4) \times 10^2$	-

[a] Data are mean ± SE as determined by non-linear curve fitting. Entries with $K_{I,substrate}$ value were fitted to the substrate inhibition equation. Error in the relative k_{cat}/K_M was obtained from propagation of error from the individual k_{cat}/K_M. Error in k_{cat}/K_M was obtained from propagation of error from the individual kinetics terms, or obtained directly from curve fitting for data sets that did not reach substrate satuation, with lower bounds given for k_{cat} and K_M.

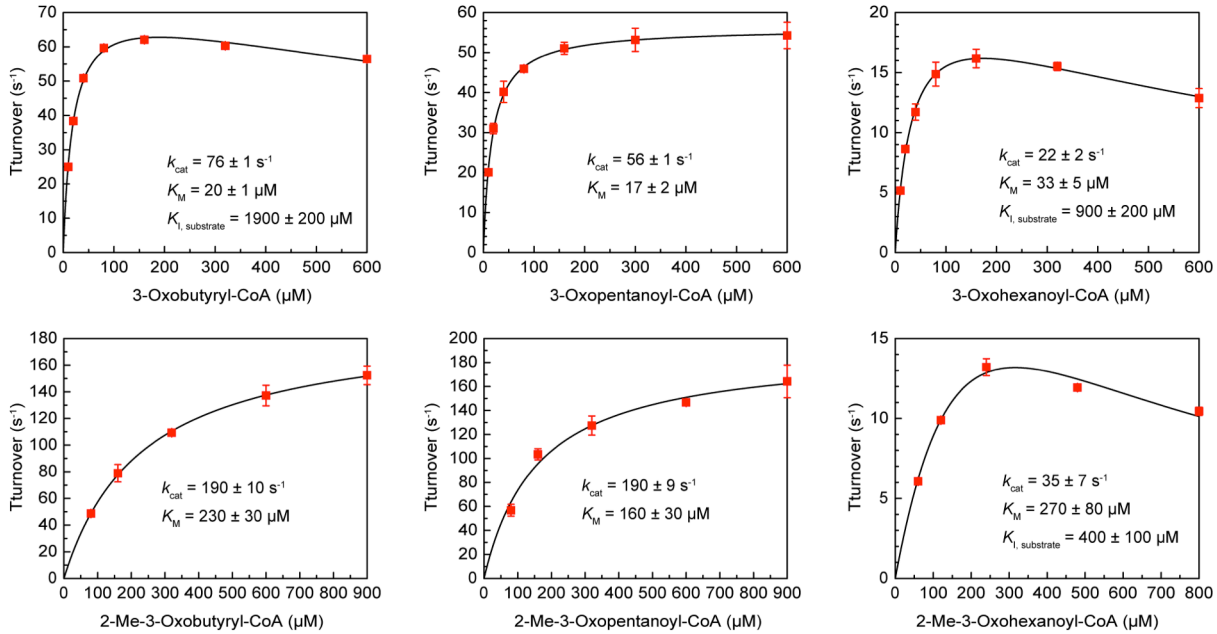


Figure 2.7. *Initial rate data for Acat2 kinetic characterization. Error bars represent mean ± standard error.*

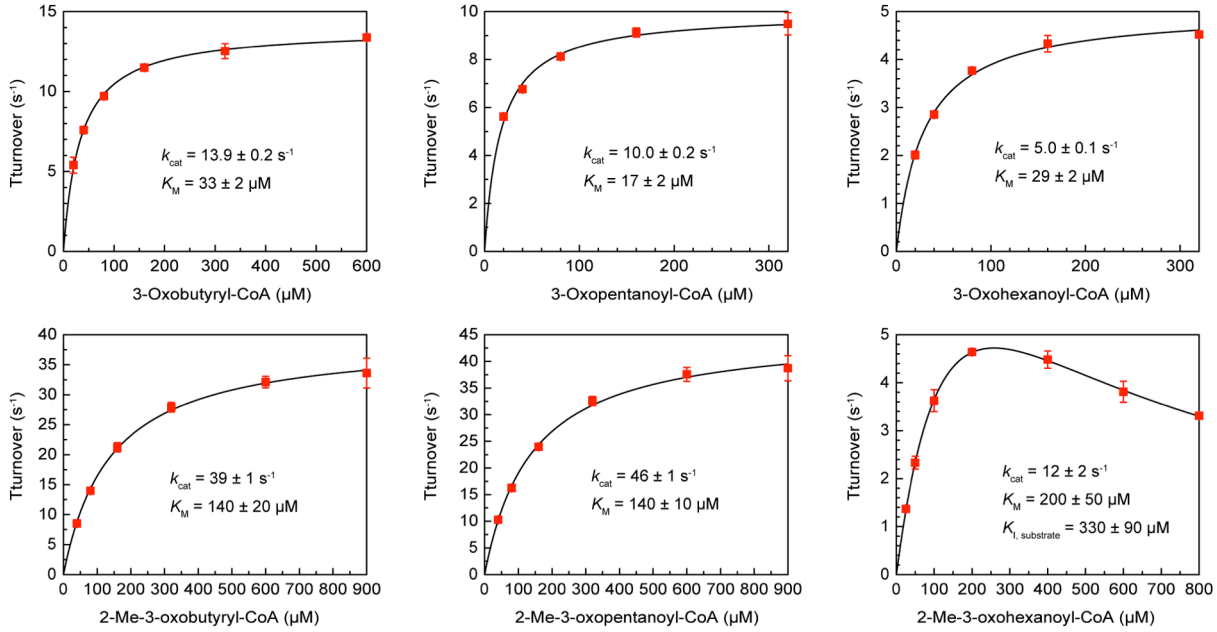


Figure 2.8. *Initial rate data for Acat3 kinetic characterization. Error bars represent mean ± standard error.*

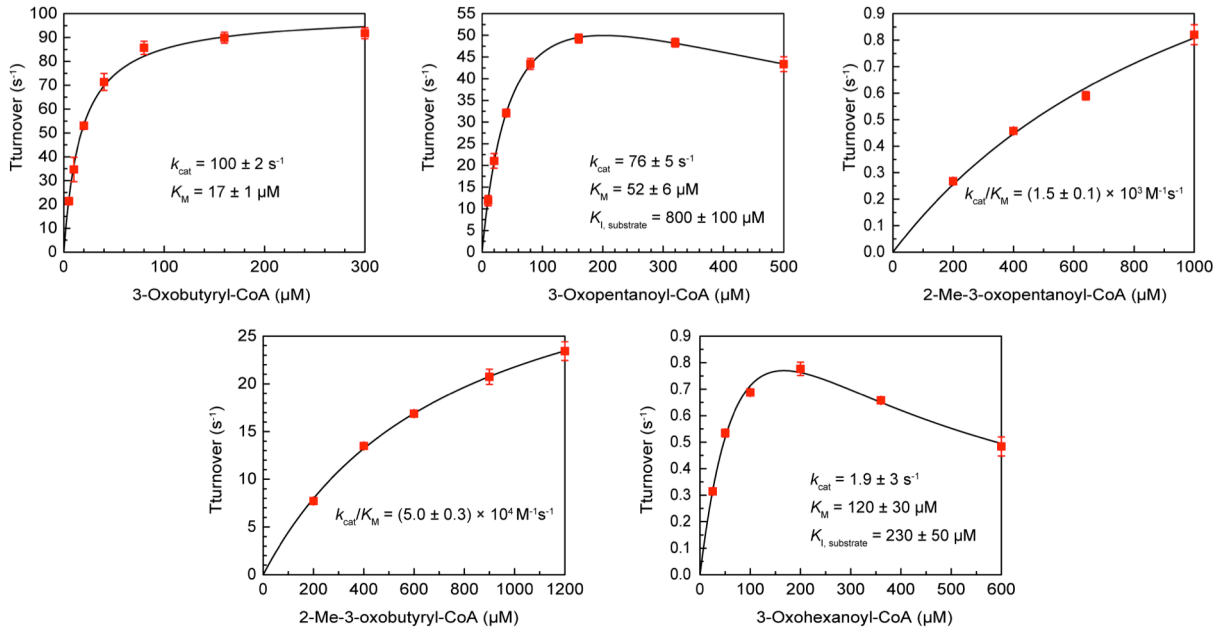


Figure 2.9. *Initial rate data for Acat5 kinetic characterization. Error bars represent mean ± standard error.*

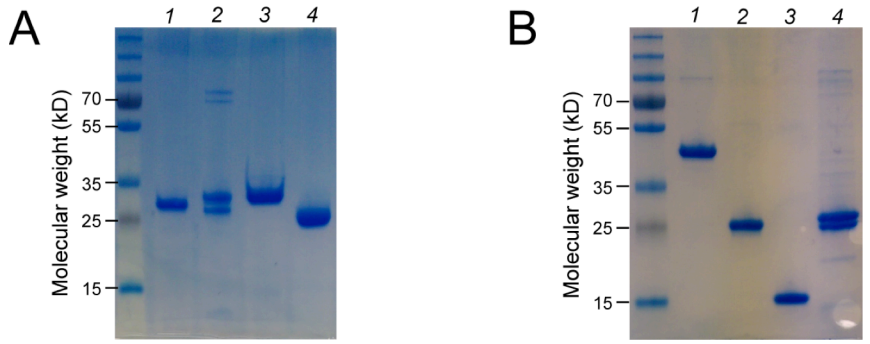


Figure 2.10. *SDS-PAGE of purified enzymes in this study. (A) Purified proteins used in synthetic pathways for organic acid production: 1. His_{10}-AsHadh, 2. His_{10}-AsEch, 3. His_{10}-TesB, 4. His_6-CnHadh. (B) Purified proteins used in coupled assays: 1. His_{10}-PhaA, 2. His_{10}-PhaB, 3. His_{10}-PhaJ, 4. Strep-Crt.*

Table 2.2. *Kinetics constants measured for reduction of 3-oxoacyl-CoA or oxidation of 3-hydroxyacyl-CoA[a] by AsHadh[b]*

Enzyme	Substrate	Relative k_{cat}/K_M	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_M (μM)	$K_{I,\ substrate}$ (μM)
AsHadh (reduction)	3-Oxobutyryl-CoA	1	$(2.1 \pm 0.6) \times 10^6$	$(2.9 \pm 0.3) \times 10^1$	$(1.3 \pm 0.4) \times 10^1$	$(5 \pm 2) \times 10^2$
	2-Me-3-oxobutyryl-CoA	20 ± 10	$(4 \pm 2) \times 10^7$	$(3.4 \pm 0.4) \times 10^1$	$(8 \pm 3) \times 10^{-1}$	$(3 \pm 1) \times 10^1$
	2-Me-3-oxopentanoyl-CoA	18 ± 9	$(4 \pm 2) \times 10^7$	$(3.4 \pm 0.4) \times 10^1$	$(9 \pm 3) \times 10^{-1}$	$(1.3 \pm 0.4) \times 10^1$
	2-Et-3-oxohexanoyl-CoA	1.0 ± 0.3	$(2.1 \pm 0.4) \times 10^6$	$(1.67 \pm 0.05) \times 10^1$	8 ± 1	-
AsHadh (oxidation)	Crotonyl-CoA + Crt	1	$(9 \pm 1) \times 10^4$	1.60 ± 0.08	$(1.8 \pm 0.3) \times 10^1$	-
	Crotonyl-CoA + PhaJ	0.024 ± 0.005	$(2.1 \pm 0.3) \times 10^3$	$(8.5 \pm 0.4) \times 10^{-2}$	$(4.0 \pm 0.6) \times 10^1$	-

[a] 3-hydroxyacyl-CoA compounds were formed *in situ* by hydration of crotonyl-CoA with Crt or PhaJ.

[b] Data are mean ± SE as determined by non-linear curve fitting. Entries with $K_{I,substrate}$ value were fitted to the substrate inhibition equation. Error in the relative k_{cat}/K_M was obtained from propagation of error from the individual k_{cat}/K_M, and error in k_{cat}/K_M was obtained from propagation of error from the individual kinetics terms.

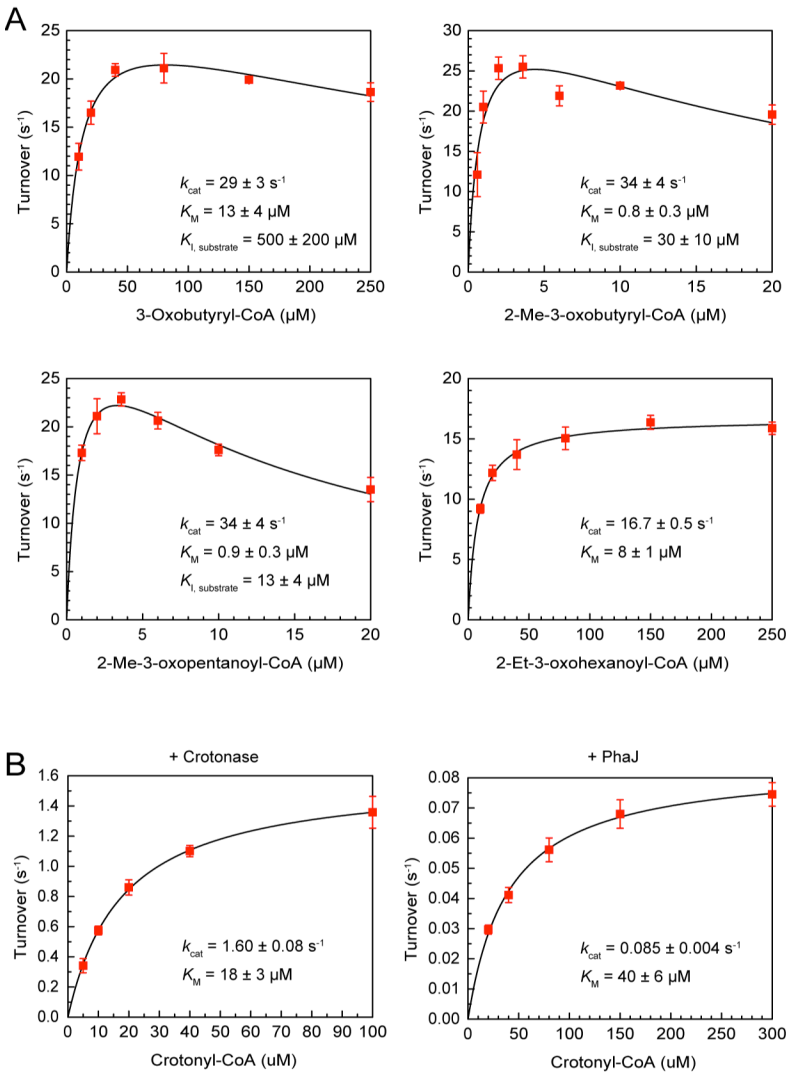


Figure 2.11. *Initial rate data for AsHadh kinetic characterization. Error bars represent mean ± standard error. (A) AsHadh reduction of 3-oxoacy-CoA substrates. (B) AsHadh oxidation of in situ 3-hydroxyacyl-CoA substrates.*

followed by a conformational shift before binding of the 3-oxoacyl-CoA substrate. After 3-hydroxyacyl-CoA product leaves the active site, the enzyme must return to the initial conformation state for cofactor exchange. If substrate concentration is high, it is feasible that new substrate may bind before NAD^+ cofactor has exchanged with NADH. It is unlikely that this inhibition plays a major physiological role in *A. suum*, however, since steady-state 3-oxoacyl-CoA concentrations are expected to remain low as stated previously.

We also investigated the stereoselectivity of AsHadh at the 3-hydroxy position by assaying in the direction of oxidation of 3*R*-hydroxybutyryl-CoA or 3*S*-hydroxybutyryl-CoA with NAD^+ (*Table 2.2, Figure 2.11B*). Chiral 3-hydroxybutyryl-CoA substrates could be formed *in situ* by the stereospecific hydration of crotonyl-CoA using one of two different hydratase enzymes, crotonase or PhaJ [17]. We found that AsHadh had 40-fold higher catalytic efficiency when 3-hydroxybutyryl-CoA was formed from crotonase compared to PhaJ, corresponding to a preference for the *3S* enantiomer over *3R*. This assay did not reveal the stereoselectivity of AsHadh for the α-methyl substituent, which is potentially set by the thiolase in the context of the pathway, but could potentially racemize due to the low pK_a of the 3-oxoacyl-CoA alpha proton and is not definitely set until after the ketoreduction step. In the ^{1}H-NMR analysis of our synthetic 3-oxoacyl-CoA substrates, for example, we generally saw an attenuated or absent signal for the α-proton due to deuterium exchange with D_2O solvent, indicating the ease of non-enzymatic racemization. Nevertheless, characterization of murine HADH2 [53] and the fact that the subsequent dehydration is expected to proceed by a *syn*-elimination mechanism [54] lead us to predict a (*2S,3S*) configuration.

Production of branched enoic acids in *E. coli*. We next set out to use *A. suum* enzymes to produce branched organic acids, such as enoic and hydroxyacids, in propionate-supplemented bacterial culture. First, an *A. suum* dehydratase, AsEch (ADY46836.1), was identified based on its similarity to the human mitochondrial hydratase ECHS1 [55]. Although two candidate dehydratases were identified with high similarity ($E < 10^{-100}$) to the human homolog, only one was predicted by MitoProt to be targeted to the mitochondria and possessed an "R-10" MIP protease signal sequence. We constructed plasmids for expression of Acat2, 3, or 5 thiolase, AsHadh ketoreductase, and AsEch dehydratase from a pTrc99a vector in *E. coli* for enoic acid production. For hydroxyacid production, similar plasmids were constructed containing only a thiolase and AsHadh. To promote high expression of pathway enzymes, we also designed ribosome binding sites using the RBS Calculator utility developed by Espah Borujeni *et al*. [56], which uses free energy calculations of RNA secondary structure and binding to the Shine-Delgarno sequence to modulate translation initiation rates (*Appendix, Table S1A*). Finally, to release free organic acids from acyl-CoA intermediates *in vivo*, we tested a panel of four native *E. coli* thioesterases [57] – TesA, TesB, YciA, or YdiI – overexpressed from a pTrc33 vector.

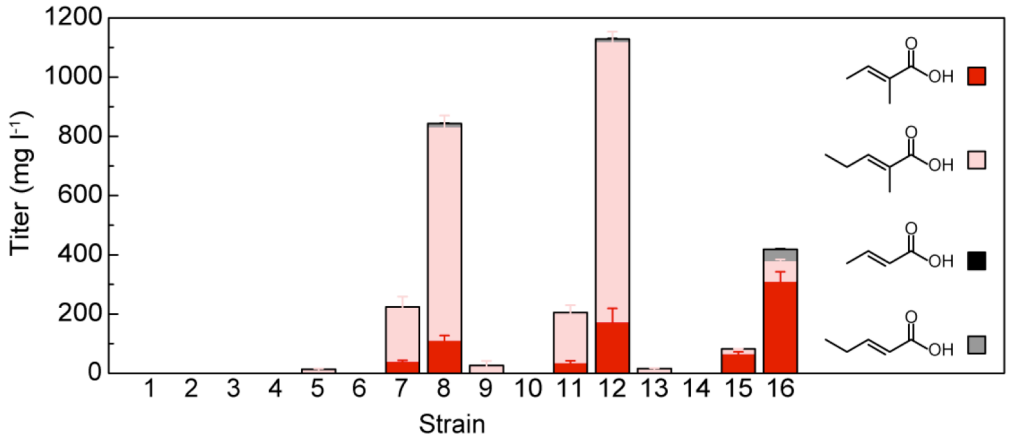
Dual-plasmid systems were transformed into *E. coli* BAP1 for enoic acid production. *E. coli* BAP1 is a modified BL21(de3) strain originally engineered by Pfeifer *et al.* to increase intracellular concentration of propionyl-CoA for the biosynthesis of the polyketide 6-deoxyerythronolide B [31]. Native *E. coli* consumes propionate through the methyl citrate pathway encoded by the *prpRBCDE* operon, beginning with the activation of propionate to propionyl-CoA by the ATP-dependent enzyme PrpE [58]. In contrast, *E. coli* BAP1 has *prpE* under the control of an inducible T7 promoter and lacks the *prpRBCD* genes for propionyl-CoA utilization, thus promoting propionyl-CoA accumulation from supplemented propionate. *Figure 2.12* shows the results of enoic acid production in strains **1-16**. Titer was heavily dependent on

the thioesterase used for enoyl-CoA hydrolysis. YciA and YdiI were both competent to produce enoic acids, with the highest titers obtained using YdiI. This result is in agreement with the thioesterase screening study of McMahon *et al*., who found *E. coli* YdiI to possess the highest activity with enoyl-CoA substrates [57]. Strain **12** containing Acat3, AsHadh, and AsEch in combination with YdiI produced a total of 1,120 ± 60 mg L^{-1} branched enoic acid products tiglic acid and 2-methylpentenoic acid, resulting in the incorporation of 87% of supplemented propionate into these products.

Interestingly, branched enoic acids were produced almost exclusively in all enoic acid-producing strains, including those containing Acat5, the thiolase that we found to have poor activity with branched substrates *in vitro*. This may reflect the importance of the ketoreductase AsHadh, which we found to be selective for 2-methyl-3-oxoacyl-CoA substrates, in determining product profiles. Alternatively, the dehydratase AsEch may be highly selective for 2-methyl-3-hydroxyacyl-CoA substrates. To test this, we purified His_6-AsEch and tested the enzyme with crotonyl-CoA or synthetic tigloyl-CoA substrates (*Figure 2.13AB*). While AsEch was permissive of the 2-methyl substrate tigloyl-CoA, we measured 15-fold higher catalytic efficiency with the linear crotonyl-CoA substrate. AsEch selectivity therefore does not account for the selective production of branched enoic acids in *E. coli*.

Another factor that may affect enoic acid product distribution is the thermodynamic difference in dehydration of linear or branched hydroxyacyl-CoA substrates. Component contribution methods of estimating reaction thermodynamics in biological contexts have been developed by Flamholz *et al.* and can be accessed using the eQuilibrator web interface [7]. Estimates for thiolase-based condensation pathways using this tool suggest the possibility of differences in pathway energetics for several enzymatic steps, including condensation and ketoreduction, but most especially for dehydration of α-methyl vs. non-methylated substrates (*Figure 2.13CD*). To verify this experimentally, we measured K_{eq} values for the hydration of crotonyl-CoA and tigloyl-CoA by incubating substrates with AsEch enzyme and analyzing by HPLC. Whereas K_{eq} of crotonyl-CoA hydration was ~5.1, the measured K_{eq} for tigloyl-CoA hydration was only ~0.25. The unfavorable hydration of α-methyl substrates is likely related to the chemical stability of the tri-substituted alkene compared to a di-substituted alkene, and this factor cannot be excluded as a contributor to the branched enoic acid selectivity observed *in vivo*.

Production of branched hydroxyacids in *E. coli.* In order to construct a pathway for the production of hydroxyacids, a combination of thiolases and thioesterases were assembled with AsHadh and tested in strains **17-32** (*Figure 2.14*). Even in strains containing the thioesterase, TesB, which has been demonstrated in multiple studies to be effective for producing short linear hydroxyacids in *E. coli* [59], productivity was considerably lower than corresponding enoic acids. We thus set out to optimize flux through branched hydroxyacid synthesis by modulating pathway enzyme expression levels with different gene promoter and RBS sequences. The most significant mprovements in titers were achieved by using a TesB expression construct with an alternative RBS (strains **33, 34**). Titration of upstream enzymes Acat3 and AsHadh (strains **35-42**) led only to a slight improvement in selective production of α-methyl vs. linear products (strain **42**). A titration of Acat3 in conjunction with CnHadh (WP_011614598), a homolog of AsHadh (52% sequence identity) from *C. necator*, showed that these strains (**43-50**) were capable of producing higher titers of branched hydroxyacids, up to 140 ± 20 mg L^{-1}, but this improved productivity was accompanied by a loss of selectivity, with an additional 500 ± 100 mg L^{-1} of linear 3-hydroxybutyrate and 3-hydroxypentanoate products formed as well. CnHadh



Strain	Plasmid 1	Plasmid 2	Strain	Plasmid 1	Plasmid 2
1	pTrc99	pTrc33-TesA	**9**	pTrc99-A3-AsHadh-Ech	pTrc33-TesA
2	pTrc99	pTrc33-TesB	**10**	pTrc99-A3-AsHadh-Ech	pTrc33-TesB
3	pTrc99	pTrc33-YciA	**11**	pTrc99-A3-AsHadh-Ech	pTrc33-YciA
4	pTrc99	pTrc33-YdiI	**12**	pTrc99-A3-AsHadh-Ech	pTrc33-YdiI
5	pTrc99-A2-AsHadh-Ech	pTrc33-TesA	**13**	pTrc99-A5-AsHadh-Ech	pTrc33-TesA
6	pTrc99-A2-AsHadh-Ech	pTrc33-TesB	**14**	pTrc99-A5-AsHadh-Ech	pTrc33-TesB
7	pTrc99-A2-AsHadh-Ech	pTrc33-YciA	**15**	pTrc99-A5-AsHadh-Ech	pTrc33-YciA
8	pTrc99-A2-AsHadh-Ech	pTrc33-YdiI	**16**	pTrc99-A5-AsHadh-Ech	pTrc33-YdiI

Figure 2.12. *Production of enoic acids in E. coli BAP1 strains containing A. suum thiolase, ketoreductase, and dehydratase, and a thioesterase. Error bars represent mean ± SD of biological replicates (n = 3).*

A

Enzyme	Substrate	Relative k_{cat}/K_M	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_M (μM)
AsEch (hydration)	Crotonyl-CoA	1	$(2.6 \pm 0.1) \times 10^7$	> 1000	> 30
	Tigloyl-CoA	0.066 ± 0.005	$(1.7 \pm 0.1) \times 10^6$	> 100	> 40

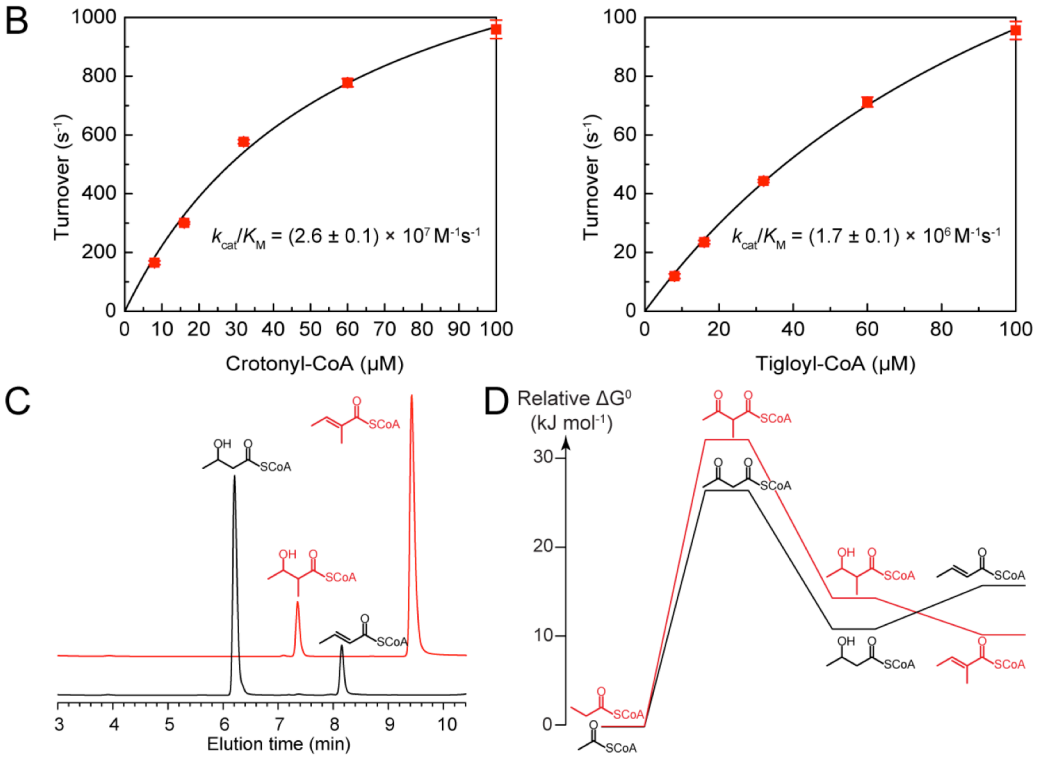
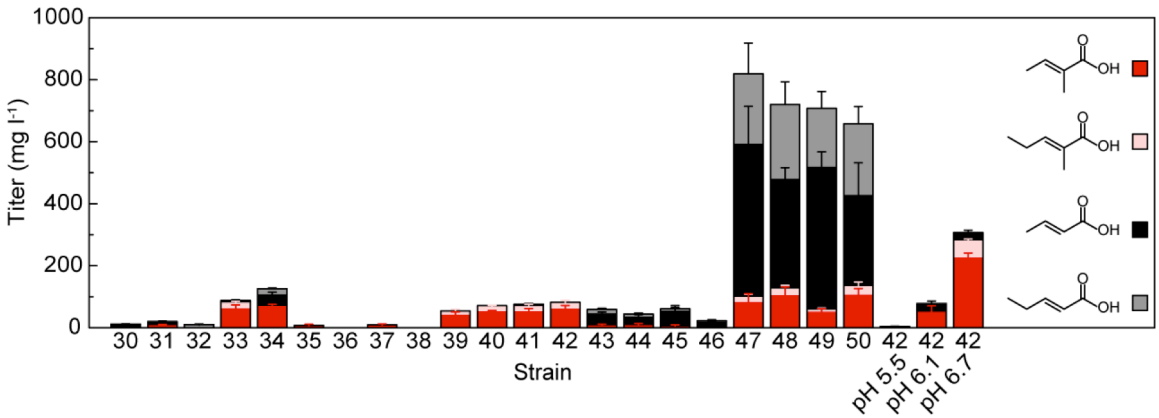


Figure 2.13. *Kinetic and thermodynamic effects on branched enoic acid synthesis. (A) k_{cat}/K_M measured for hydration of crotonyl-CoA and tigloyl-CoA by AsEch. Data are mean ± SE as determined by non-linear curve fitting, with lower bounds given for k_{cat} and K_M. Error in the relative k_{cat}/K_M was obtained from propagation of error from the individual k_{cat}/K_M. (B) Initial rate data for AsEch kinetic characterization. Error bars represent mean ± standard error. (C) HPLC trace of crotonyl-CoA and tigloyl-CoA at equilibrium with their respective hydration products. (D) Predicted relative ΔG° values of branched and linear pathway intermediates, estimated by additive contribution of each enzymatic step. ΔG° values for thiolase condensation and ketoreduction were estimated using the eQuilibrator web utility [7] at pH 7.4, ionic strength 0.2 M. ΔG° of dehydration was calculated from experimentally determined K_{eq} values using $\Delta G° = -RT\ln K_{eq}$.*



Strain	Plasmid 1	Plasmid 2	Strain	Plasmid 1	Plasmid 2
17	pTrc99	pTrc33-TesA	**34**	pTrc99-A5-AsHadh	pTrc33-TesB2
18	pTrc99	pTrc33-TesB	**35**	pTrc99-A3AsH00	pTrc33-TesB2
19	pTrc99	pTrc33-YciA	**36**	pCWori-A3AsH00	pTrc33-TesB2
20	pTrc99	pTrc33-YdiI	**37**	pTrc99-A3AsH01	pTrc33-TesB2
21	pTrc99-A2-AsHadh	pTrc33-TesA	**38**	pCWori-A3AsH01	pTrc33-TesB2
22	pTrc99-A2-AsHadh	pTrc33-TesB	**39**	pTrc99-A3AsH10	pTrc33-TesB2
23	pTrc99-A2-AsHadh	pTrc33-YciA	**40**	pCWori-A3AsH10	pTrc33-TesB2
24	pTrc99-A2-AsHadh	pTrc33-YdiI	**41**	pTrc99-A3AsH11	pTrc33-TesB2
25	pTrc99-A3-AsHadh	pTrc33-TesA	**42**	pCWori-A3AsH11	pTrc33-TesB2
26	pTrc99-A3-AsHadh	pTrc33-TesB	**43**	pTrc99-A3CnH00	pTrc33-TesB2
27	pTrc99-A3-AsHadh	pTrc33-YciA	**44**	pCWori-A3CnH00	pTrc33-TesB2
28	pTrc99-A3-AsHadh	pTrc33-YdiI	**45**	pTrc99-A3CnH01	pTrc33-TesB2
29	pTrc99-A5-AsHadh	pTrc33-TesA	**46**	pCWori-A3CnH01	pTrc33-TesB2
30	pTrc99-A5-AsHadh	pTrc33-TesB	**47**	pTrc99-A3CnH10	pTrc33-TesB2
31	pTrc99-A5-AsHadh	pTrc33-YciA	**48**	pCWori-A3CnH10	pTrc33-TesB2
32	pTrc99-A5-AsHadh	pTrc33-YdiI	**49**	pTrc99-A3CnH11	pTrc33-TesB2
33	pTrc99-A3-AsHadh	pTrc33-TesB2	**50**	pCWori-A3CnH11	pTrc33-TesB2

Figure 2.14. *Production of hydroxyacids acids in E. coli BAP1 strains. Error bars represent mean ± SD of biological replicates (n = 3). Strains* ***17-32*** *contain combinations of A. suum thiolases and E. coli thioesterases with AsHadh. Strains* ***33*** *and* ***34*** *contain TesB under an alternate RBS (TesB2). Strains* ***35-42*** *titrate expression of Acat3 and AsHadh. Strains* ***43-50*** *titrate expression of Acat3 and CnHadh. The final three entries were cultured in MES-buffered media.*

A

Enzyme	Substrate	Relative k_{cat}/K_M	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_M (μM)	$K_{I,\ substrate}$ (μ
CnHadh	3-Oxobutyryl-CoA	1	(1.3 ± 0.3) × 10^7	(5.0 ± 0.5) × 10^2	(3.8 ± 0.7) × 10^1	(3.3 ± 0.8) ×
	2-Me-3-oxobutyryl-CoA	4 ± 2	(6 ± 2) × 10^7	(2.8 ± 0.4) × 10^2	5 ± 1	(2.6 ± 0.6) ×

B

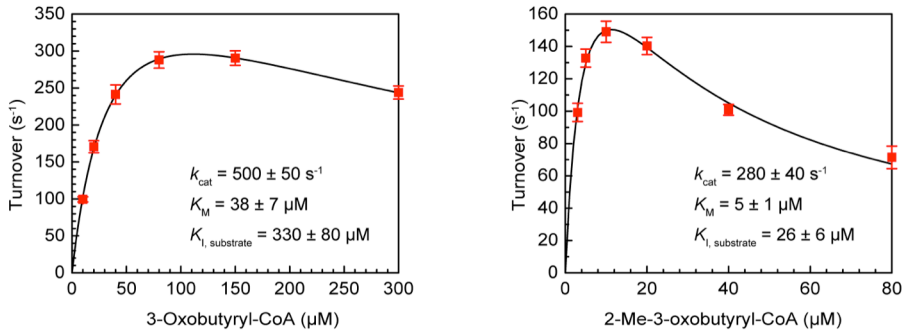


Figure 2.15. *Kinetic characterization of CnHadh. (A) Kinetics constants measured for reduction of 3-oxoacyl-CoA by AsHadh. Data are mean ± SE as determined by non-linear curve fitting to the substrate inhibition equation. Error in the relative k_{cat}/K_M was obtained from propagation of error from the individual k_{cat}/K_M, and error in k_{cat}/K_M was obtained from propagation of error from the individual kinetics terms. (B) Initial rate data. Error bars represent mean ± standard error.*

is predicted to be associated with the degradation of branched amino acids [60]. *In vitro* biochemical characterization of CnHadh showed that it did indeed accept α-methyl substituted substrates with high catalytic efficiency, although with only a 4-fold preference for the branched substrate (*Figure 2.15*), compared to 20-fold for AsHadh.

Because the improvements in titer observed in strains **34** and **35** compared to strains **26** and **30** suggest a critical role for the thioesterase TesB in controlling flux through the pathway (*Figure 2.14*), we examined the variation in expression of TesB. Although pTrc33-TesB2, which correlated with higher hydroxyacid production, was predicted to have stronger TesB expression than pTrc33-TesB based on its RBS site, SDS-PAGE showed that soluble TesB expression was actually lower from this vector (*Figure 2.16*). This shows that successful hydroxyacid production relies on a delicate balance of thioesterase expression and that overly high expression of TesB can derail production of hydroxyacids. We also tested strains lacking any plasmid-expressed thioesterase (**51**, **53**; *Figure 2.17A*) and found that these strains produced a modest level of hydroxyacids that exceeded those strains containing the high-expressing pTrc33-TesB plasmid. Analysis of spent media indicated that a substantial portion of the propionate remaining after 48 hours growth, which excluded off-pathway propionate utilization as an explanation for observed titers (*Figure 2.17B***)**. To further explore the origin of this behavior, we expressed and purified TesB for *in vitro* biochemical characterization. By testing its hydrolytic activity against acetyl-CoA, propionyl-CoA, and various 3-hydroxyacyl-CoA substrates, we found that propionyl-CoA and 2-methyl-3*S*-hydroxybutyryl CoA performed similarly as TesB substrates with very little catalytic defect (3.7-fold) compared to the control substrate (3*S*-hydroxybutyryl-CoA; *Figure 2.18*). Taken together, the data are consistent with a model that the pathway is limited by competition between hydrolysis of the propionyl-CoA substrate and the product by the thioesterase.

The absence of significant off-pathway reactions gives rise to the possibility of increasing throughput through the pathway by optimizing growth conditions. For example, pH is an important factor that affects bacterial growth, and standard bacterial culture media, such as LB

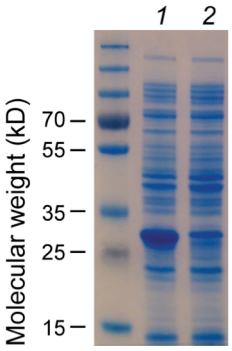
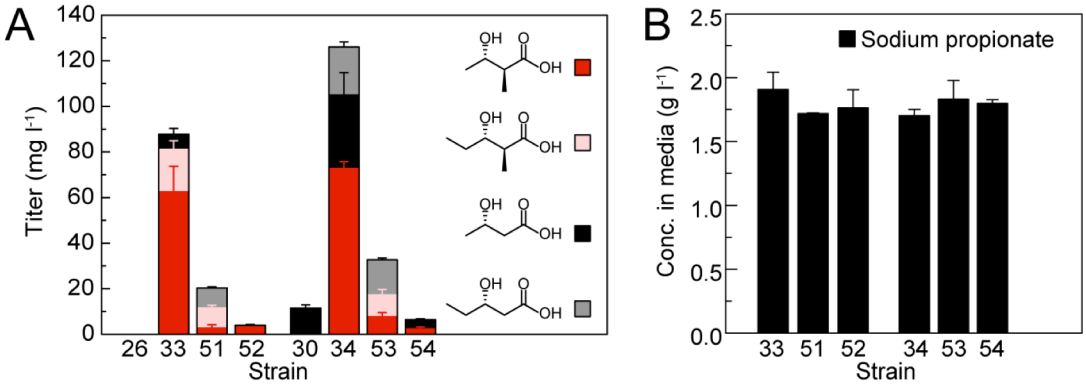


Figure 2.16. *Soluble lysate of E. coli BAP1 expressing 1. pTrc33-TesB and 2. pTrc33-TesB2.*



Strain	***Plasmid 1***	***Plasmid 2***	***Strain***	***Plasmid 1***	***Plasmid 2***
26	pTrc99-A3-AsHadh	pTrc33-TesB	**30**	pTrc99-A5-AsHadh	pTrc33-TesB
33	pTrc99-A3-AsHadh	pTrc33-TesB2	**34**	pTrc99-A5-AsHadh	pTrc33-TesB2
51	pTrc99-A3-AsHadh	pTrc33	**53**	pTrc99-A5-AsHadh	pTrc33
52	pTrc99-A3	pTrc33-TesB2	**54**	pTrc99-A5	pTrc33-TesB2

Figure 2.17. *Elevated TesB expression is detrimental to product titer. (A) Production of hydroxyacids in E. coli BAP1 strains* ***26****,* ***33****,* ***51****,* ***52****,* ***30****,* ***34****,* ***53****, and* ***54****. Error bars represent mean ± SD of biological replicates (n = 3). (B) Propionate remaining in media after 2 days of culture. Initial media contained 2 g* L^{-1}*.*

A

Enzyme	Substrate	Relative k_{cat}/K_M	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_M (μM)
TesB	Acetyl-CoA	0.19 ± 0.01	(2.7 ± 0.2) × 10^3	(> 8) × 10^{-1}	(> 2) × 10^2
	Propionyl-CoA	1	(1.44 ± 0.04) × 10^4	> 4	(> 2) × 10^2
	2-Me-3S-hydroxybutyryl-CoA	1.0 ± 0.1	(1.4 ± 0.2) × 10^4	2.0 ± 0.1	(1.4 ± 0.2) × 10^2
	3R-Hydroxybutyryl-CoA	1.2 ± 0.2	(1.8 ± 0.2) × 10^4	4.0 ± 0.2	(2.2 ± 0.3) × 10^2
	3S-Hydroxybutyryl-CoA	4 ± 1	(5 ± 1) × 10^4	(1.8 ± 0.2) × 10^1	(3.3 ± 0.7) × 10^2

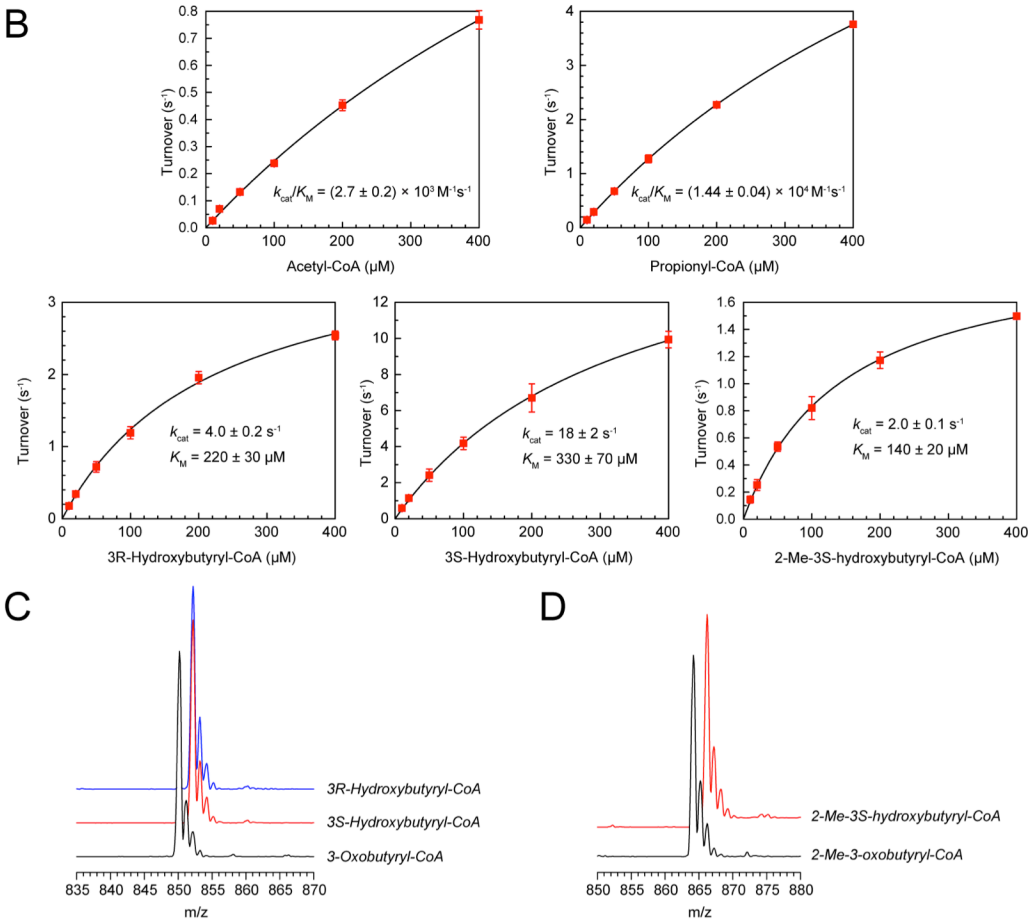


Figure 2.18. *Kinetic characterization shows the promiscuity of TesB. (A) Kinetics constants measured for hydrolysis of acyl-CoA substrates by TesB. [a]Data are mean ± SE as determined by non-linear curve fitting. Error in the relative k_{cat}/K_M was obtained from propagation of error from the individual k_{cat}/K_M. Error in k_{cat}/K_M was obtained from propagation of error from the individual kinetics terms, or directly from curve fitting for data sets that did not reach substrate satuation, with lower bounds given for k_{cat} and K_M. (B) Initial rate data for TesB kinetic characterization. Error bars represent mean ± standard error. (C), (D) Mass spectra of 3-hydroxyacyl-CoA substrates for TesB assays, prepared in situ by addition of AsHadh (for 3S stereochemistry) or PhaB (for 3R stereochemistry) and NAD(P)H to (C) 3-oxobutyryl-CoA or (D) 2-Me-3-oxobutyryl-CoA.*

broth, have a low capacity for buffering against acidification that occurs during growth under conditions of insufficient oxygenation or excessive glucose [61]. By supplementing our standard production media (LB broth, 1% glucose) with 100 mM MES buffer, we could significantly increase hydroxyacid titers with strain **42** in a clearly pH dependent manner, with up to 280 ± 10 mg L^{-1} branched acid products produced in media at pH 6.7 (*Figure 2.14*).

Pathway design for the production of α-methyl PHAs *E. coli*. 3-Hydroxy acids serve as monomers for the production of commercial polyesters, such as polylactide (PLA). While biologically-sourced PHAs are unsubstituted, α-branched versions offer potential for altering their thermal, mechanical, and degradation behavior [20]. The PhaC polymerase that accepts these substrates is highly stereoselective for *3R*-hydroxyacyl-CoA [62], as produced by the PhaB ketoreductase from *C. necator*. However, stereochemical characterization of AsHadh demonstrated its preference for the *3S*-hydroxy enantiomer, which is consistent with what has been reported for murine HADH2 [53]. In the absence of a known ketoreductase that both permits α-methyl substituents and confers 3*R*-hydroxy stereochemistry, we designed a synthetic pathway using *A. suum* enzymes which potentially bypasses stereochemistry issues by invoking a two-step racemization through the use of complementary dehydratases, AsEch and PhaJ (*Figure 2.19*). In this pathway, 2-methyl-3*S*-hydroxyacyl-CoA formed from Acat2 or Acat3 and AsHadh is dehydrated by AsEch to form a branched enoyl-CoA. We hypothesized that this enoyl-CoA could then be rehydrated to form branched 3*S*-hydroxyacyl-CoAs *in vivo*, which serve as subsrates for PhaC polymerase. The equilibrium of 3*R*- and 3*S*-hydroxyacyl-CoA should in principle be shifted towards the 3*R*-monomer as this product is depleted through irreversible polymerization and sequestration in a solid phase [17]. Although no PhaC isozymes have been demonstrated biochemically to permit α-methyl substituents with comparable efficiency to non-branched hydroxyacids, *in vivo* studies of biopolymer production by *R. ruber* and *R. opacus* showed that these organisms incorporate hydroxypivalic acid, a 2,2,-dimethyl substrate, into PHAs [63].

We constructed a series of plasmids, Poly1, 2, 3, and 4, containing an *A. suum* thiolase and *Rhodococcus sp.* PhaC, in various combinations (*Table 2.3*). We also designed plasmid DKD2, containing AsHadh, AsEch, and PhaJ, and tested these two-plasmid systems for the production of branched PHAs, but were unable to detect any polymer formation by HPLC. To troubleshoot this system, we measured the activity of pathway enzymes in cell culture lysate (*Table 2.4*). Activity was detected for pathway enzymes in all of the strains, with the exception of the strain containing pTrc99, which lacked any thiolase and also had no detectable thiolase activity. Ketoreductase and dehydratase activities were of a similar magnitude to that which has been described for high flux biosynthesis of butanol *in vivo* [17]. Thiolase activity, however, was relatively low, so we sought to increase Acat2 expression with a new plasmid, Poly1B. Cell lysate assay with a strain expressing this construct did indeed show a 2.4-fold increase in thiolase activity of 0.95 U mg^{-1}. We also constructed plasmids Poly7B and Poly8B containing NphT7, a condensing enzyme using malonyl-CoA, in place of a thiolase [64]. When we tested Poly1B, 7B, and 8B, however, none of these constructs allowed the production of PHA in conjuction with DKD plasmid. In conjuction with pTrc33-PhaB, which contains the canonical 3R-specific ketoreductase PhaB, only Poly7B allowed PHA production as detected by propanolysis and GC-MS, with the primary monomer component being 3-hydroxypentanote and no branched monomers detected (*Figure 2.20*).

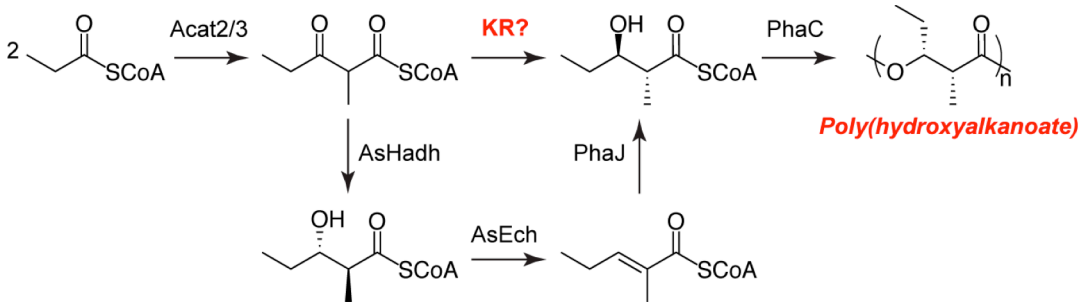


Figure 2.19. *Proposed pathway for the production of α-methyl PHAs in E. coli using A. suum enzymes, Rhodococcus sp. PhaC. In the absence of a ketoreductase (KR) with the requisite features, an in vivo racemization system based on dual hydratase enzymes is required to match the stereoselectivity of PhaC.*

Table 2.3. *Plasmids used for PHA production in E. coli.*

Plasmid 1	***Condensing enzyme***	***PhaC***
Poly1	Acat2	*R. ruber*
Poly2	Acat3	*R. ruber*
Poly3	Acat2	*R. opacus*
Poly4	Acat3	*R. opacus*
Poly1B	Acat2[a]	*R. ruber*
Poly7B	NphT7	*R. ruber*
Poly8B	NphT7	*R. ruber*

Plasmid 2	***Ketoreductase***	***Dehydratase***
DKD2	AsHadh	PhaJ + AsEch
pTrc33-PhaB	PhaB	none

[a] Acat2 gene in Poly1B contains a higher expressing RBS.

Table 2.4. *Enzymes activities in cell lysate of PHA-producing strains.*[a]

Plasmid system	Acat2/3	AsHadh	AsEch	PhaJ
DKD2 + Poly1	0.39 U mg^{-1}	0.13 U mg^{-1}	122 U mg^{-1}	135 U mg^{-1}
DKD2 + Poly2	0.33 U mg^{-1}	0.18 U mg^{-1}	96 U mg^{-1}	114 U mg^{-1}
DKD2 + Poly3	0.54 U mg^{-1}	0.13 U mg^{-1}	95 U mg^{-1}	121 U mg^{-1}
DKD2 + Poly4	0.30 U mg^{-1}	0.20 U mg^{-1}	125 U mg^{-1}	125 U mg^{-1}
DKD2 + pTrc99	*n.d.*	0.25 U mg^{-1}	122 U mg^{-1}	134 U mg^{-1}

[a] Activities are a single replicate. n.d. = not detectable. 1 U mg^{-1} = 1 µmol substrate min^{-1} per mg total protein.

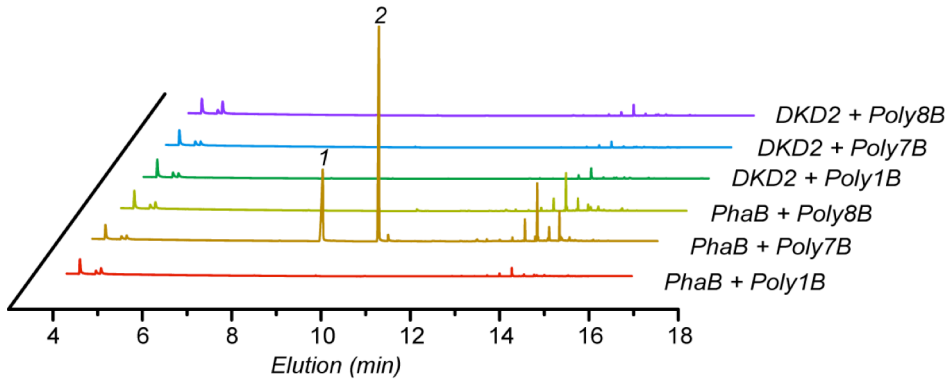


Figure 2.20. *GC-MS traces of propanolized PHA produced by strains containing plasmids Poly1B, Poly7B, and Poly8B. 1. 3-hydroxybutyrate, propyl ester. 2. 3-hydroxypentanoate, propyl ester.*

PhaC enzymes often have unusually solubility features, as they natively associate with granules of PHA and other proteins called phasins involved in chaperoning biopolymer synthesis [65]. Soluble isolation of polymerases are therefore challenging, as are methods for assaying them *in vitro*. We were unable to detect PhaC enzyme activity in cell lysate of any strains tested. But since this was also the case for the PHA-producing strain containing Poly7B, whether functional PhaC polymerase was expressed or not remains inconclusive. Finally, we measured the activity of purified PhaJ enzyme with tigloyl-CoA and found that PhaJ displayed ~10,000-fold lower activity with a branched substrate (~0.06 s^{-1} with tigloyl-CoA vs. ~600 s^{-1} with crotonyl-CoA). The feasibility of the proposed scheme in *Figure 2.19* would therefore depend on whether a PhaJ ortholog could be identified with the same stereoselectivity but permissiveness towards α-methyl enoyl-CoA substrates.

2.4 Discussion

The branched acid fermentation of *A. suum* expands the utility of thiolase-based pathways for biosynthesis and illustrates the effectiveness of identifying novel and non-canonical carbon metabolism in nature as a source of innovation in synthetic biology. Thiolase pathways operate directly on acyl-CoA substrates in the cell and therefore orthogonally to the ACP pool of FAS or PKS complexes, which may be tightly coupled to cellular lipid or energy homeostasis [66] or highly evolved for a specific molecular product. Also of note is the recently reported OleA enzyme class, which forms part of the thiolase superfamily and initiates olefin biosynthesis by irreversibly condensing long-chain acyl-CoA compounds to form branched 3-oxoacid products [67, 68].

In contrast to other reported thiolase-based biosynthesis strategies [17, 19, 59, 69, 70], the *A. suum* pathway allows for more than one possible nucleophile and therefore introduces ambiguity in the directionality of carbon-carbon bond formation. We have demonstrated that the choice of ketoreductase plays an important role in dictating the distribution of condensation end products formed *in vivo*. The role of the ketoreductaseis seen in the production of branched enoic acids using Acat5, a thiolase that functioned poorly with propionyl-CoA condensation *in vitro* but nonetheless allowed production of branched enoic acids almost exclusively. While thermodynamics likely also affect enoic acid biosynthesis, high selectivity for branched products is observed for hydroxyacids as well. Comparing the titers of 2-methyl-3-hydroxybutyric acid and 3-hydroxypentanoic acid with strain **42** (*Figure 2.14*), for example, highlights the role of AsHadh ketoreductase selectivity. Since both of these acids are formed from the condensation of one acetyl-CoA and one propionyl-CoA, the partition between them should be unaffected by the relative abundance of propionyl-CoA vs. acetyl-CoA in the cell.

The selective production of branched hydroxyacids with AsHadh – which displays a 20-fold catalytic preference for α-methyl compounds *in vitro* – is significantly diminished with CnHadh (strain **50**; *Figure 2.14*), despite the 4-fold branched substrate preference measured *in vitro* with that homolog. CnHadh is expected to function catabolically in isoleucine degradation as a 2-methyl-3-hydroxybutyryl-CoA dehydrogenase. Mammalian mitochondrial HADH2, to which AsHadh is similar, also functions in isoleucine degradation, but is promiscuous and serves important functions as a steroid dehydrogenase and a binding protein of amyloid-beta peptides as well [71]. AsHadh may be specially suited for a biosynthetic function in *Ascaris suum*. Its very low K_M values for 2-methyl-3-oxoacyl-CoA substrates suggest this.

The ThYme (Thioester-Active EnzYmes) database established by Cantu *et al*. [72], AsHadh and HADH2 fall in the same superfamily of ketoacyl reductases (KR1) as PhaB, despite their differences in stereochemistry and branched substrate preference. Although murine HADH2 and PhaB share little sequence identity (32%), alignments of PDB 1E3W [71] and 4N5M [73] entries show remarkable structural similarity, with conserved Ser-Tyr-Lys catalytic triad aligning in the active site. Therefore the structural basis for differing stereoselectivity remains to be elucidated. Sequence motifs characteristic of Type A and Type B ketoreductases have been described in the polyketide synthase literature [74], but these motifs are not observed in HADH2, AsHadh, or PhaB. A 3-oxoacyl-CoA reductase that both confers *3R*-hydroxy stereochemistry and is selective for α-methyl substrates would be highly desirable in order to produce novel methyl-substituted PHA polymers *in vivo*. The existence of such a ketoreductase is anticipated by the report of PHA containing 3-hydroxy-2-methylpentanoate produced in activated sludge [75], but none has been identified.

We have demonstrated a biological route for the production of branched organic acids starting from glucose and propionate, including 2-methyl enoic acids (1,120 ± 60 mg L^{-1}, 87% propionate conversion) and chiral 2-methyl-3-hydroxyacids (280 ± 10 mg L^{-1}, 13% propionate conversion). Chemoenzymatic methods have previously been employed for the stereoselective synthesis of 2-methyl-3-hydroxypentanoate as a precursor to sitophilate, an insect pheromone [35], but a fermentation platform allows for the production of small molecules directly from cheaply available feedstocks. Although our present system requires the supplementation of sodium propionate, propionyl-CoA synthesis from central metabolites could be engineered through other pathways [70].

2.5 References

1. X. Sun, X. Shen, R. Jain, Y. Lin, J. Wang, J. Sun, J. Wang, Y. Yan, and Q. Yuan, Synthesis of chemicals by metabolic engineering of microbes. *Chem. Soc. Rev.* **2015,** 44, (11), 3760-3785.
2. J. Becker and C. Wittmann, Systems and synthetic metabolic engineering for amino acid production - the heartbeat of industrial strain development. *Curr. Opin. Biotechnol.* **2012,** 23, (5), 718-26.
3. J. Becker and C. Wittmann, Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. *Angew. Chem. Int. Ed. Engl.* **2015,** 54, (11), 3328-50.
4. C. Engler, R. Gruetzner, R. Kandzia, and S. Marillonnet, Golden Gate Shuffling: A one-pot DNA shuffling method based on Type IIS restriction enzymes. *PLoS One* **2009,** 4, (5), e5553.
5. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, and H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **2009,** 6, (5), 343-345.
6. T. Jakočiūnas, M. K. Jensen, and J. D. Keasling, CRISPR/Cas9 advances engineering of microbial cell factories. *Metab. Eng.* **2016,** 34, 44-59.
7. A. Flamholz, E. Noor, A. Bar-Even, and R. Milo, eQuilibrator--the biochemical thermodynamics calculator. *Nucleic Acids Res.* **2012,** 40, (Database issue), D770-5.

8. A. M. Haapalainen, G. Merilainen, P. L. Pirila, N. Kondo, T. Fukao, and R. K. Wierenga, Crystallographic and kinetic studies of human mitochondrial acetoacetyl-CoA thiolase: the importance of potassium and chloride ions for its structure and function. *Biochemistry* **2007,** 46, (14), 4305-21.

9. H. Chung, J. E. Yang, J. Y. Ha, T. U. Chae, J. H. Shin, M. Gustavsson, and S. Y. Lee, Bio-based production of monomers and polymers by metabolically engineered microorganisms. *Curr. Opin. Biotechnol.* **2015,** 36, 73-84.

10. J. W. Lee, D. Na, J. M. Park, J. Lee, S. Choi, and S. Y. Lee, Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nat. Chem. Biol.* **2012,** 8, (6), 536-546.

11. E. J. Corey and X.-M. Cheng, *The Logic of Chemical Synthesis*.; J. Wiley: New York, 1989.

12. J. McMurry and T. P. Begley, *The Organic Chemistry of Biological Pathways*. Roberts and Co. Publishers: Englewood, Colo., 2005.

13. A. R. Fox, G. Soto, M. Mozzicafreddo, A. N. Garcia, M. Cuccioloni, M. Angeletti, J. C. Salerno, and N. D. Ayub, Understanding the function of bacterial and eukaryotic thiolases II by integrating evolutionary and functional approaches. *Gene* **2014,** 533, (1), 5-10.

14. A. Steinbüchel and S. Hein, Biochemical and molecular basis of microbial synthesis of polyhydroxyalkanoates in microorganisms. *Adv. Biochem. Eng. Biotechnol.* **2001,** 71, 81-123.

15. L. Hiser, M. E. Basson, and J. Rine, ERG10 from Saccharomyces cerevisiae encodes acetoacetyl-CoA thiolase. *J. Biol. Chem.* **1994,** 269, (50), 31383-31389.

16. H. M. Miziorko, Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* **2011,** 505, (2), 131-143.

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

18. S. Kim, Y.-S. Jang, S.-C. Ha, J.-W. Ahn, E.-J. Kim, J. Hong Lim, C. Cho, Y. Shin Ryu, S. Kuk Lee, S. Y. Lee, and K.-J. Kim, Redox-switch regulatory mechanism of thiolase from Clostridium acetobutylicum. *Nature Communications* **2015,** 6, 8410.

19. C. R. Shen, E. I. Lan, Y. Dekishima, A. Baez, K. M. Cho, and J. C. Liao, Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.* **2011,** 77, (9), 2905-2915.

20. D. K. Schneiderman and M. A. Hillmyer, Aliphatic polyester block polymer design. *Macromolecules* **2016,** 49, (7), 2419-2428.

21. R. W. Haushalter, W. Kim, T. A. Chavkin, L. The, M. E. Garber, M. Nhan, P. D. Adams, C. J. Petzold, L. Katz, and J. D. Keasling, Production of anteiso-branched fatty acids in *Escherichia coli*; next generation biofuels with improved cold-flow properties. *Metab. Eng.* **2014,** 26, 111-118.

22. E. W. Dunn and G. W. Coates, Carbonylative polymerization of propylene oxide: A multicatalytic approach to the synthesis of poly(3-hydroxybutyrate). *J. Am. Chem. Soc.* **2010,** 132, (33), 11412-3.

23. R. Komuniecki, T. Campbell, and N. Rubin, Anaerobic metabolism in *Ascaris suum*: acyl CoA intermediates in isolated mitochondria synthesizing 2-methyl branched-chain fatty acids. *Mol. Biochem. Parasitol.* **1987,** 24, (2), 147-154.

24. R. Komuniecki, P. R. Komuniecki, and H. J. Saz, Pathway of formation of branched-chain volatile fatty acids in *Ascaris* mitochondria. *J. Parasitol.* **1981,** 67, (6), 841-6.

25. Z. S. de Mata, J. Arevalo, and H. J. Saz, Propionyl-CoA condensing enzyme from *Ascaris* muscle mitochondria. *Arch. Biochem. Biophys.* **1991,** 285, (1), 166-171.

26. Z. S. de Mata, R. Lizardo, F. Diaz, and H. J. Saz, Propionyl-CoA condensing enzyme from *Ascaris* muscle mitochondria. *Arch. Biochem. Biophys.* **1991,** 285, (1), 158-165.

27. Z. S. de Mata, H. J. Saz, and D. J. Pasto, 2-methylacetoacetate reductase and possible propionyl coenzyme A condensing enzyme activity in branched chain volatile fatty acid synthesis by *Ascaris lumbricoides*. *J. Biol. Chem.* **1977,** 252, (12), 4215-24.

28. K. Kita and S. Takamiya, Electron-transfer complexes in *Ascaris* mitochondria. *Adv. Parasitol.* **2002,** 51, 95-131.

29. R. Komuniecki, S. Fekete, and J. Thissen-Parra, Purification and characterization of the 2-methyl branched-chain Acyl-CoA dehydrogenase, an enzyme involved in NADH-dependent enoyl-CoA reduction in anaerobic mitochondria of the nematode, *Ascaris suum*. *J. Biol. Chem.* **1985,** 260, (8), 4770-4777.

30. H. J. Saz and B. S. deBruyn, 2-Methylbutyryl-CoA: succinate acyl-CoA transferase activity and function in *Ascaris suum* muscle. *Comp. Biochem. Physiol. Biochem. Mol. Biol.* **1994,** 108, (4), 513-9.

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

32. S. van den Berg, P.-A. Löfdahl, T. Härd, and H. Berglund, Improved solubility of TEV protease by directed evolution. *J. Biotechnol.* **2006,** 121, (3), 291-298.

33. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Ed. J. M. Walker, Humana Press: 2005; 571-607

34. T. Elssner, L. Hennig, H. Frauendorf, D. Haferburg, and H. P. Kleber, Isolation, identification, and synthesis of gamma-butyrobetainyl-CoA and crotonobetainyl-CoA, compounds involved in carnitine metabolism of *E. coli*. *Biochemistry* **2000,** 39, (35), 10761-9.

35. D. Kalaitzakis, S. Kambourakis, D. J. Rozzell, and I. Smonou, Stereoselective chemoenzymatic synthesis of sitophilate: a natural pheromone. *Tetrahedron: Asymmetry* **2007,** 18, (20), 2418-2426.

36. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **1990,** 215, (3), 403-10.

37. D. A. Benson, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers, GenBank. *Nucleic Acids Res.* **2015,** 43, (Database issue), D30-5.

38. M. G. Claros and P. Vincens, Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.* **1996,** 241, (3), 779-786.

39. H. Nielsen, J. Engelbrecht, S. Brunak, and G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **1997,** 10, (1), 1-6.

40. O. Gakh, P. Cavadini, and G. Isaya, Mitochondrial processing peptidases. *Biochim. Biophys. Acta* **2002,** 1592, (1), 63-77.

41. E.-J. Kim and K.-J. Kim, Crystal structure and biochemical characterization of PhaA from *Ralstonia eutropha*, a polyhydroxyalkanoate-producing bacterium. *Biochem. Biophys. Res. Commun.* **2014,** 452, (1), 124-129.

42. S. Sato, H. Maruyama, T. Fujiki, and K. Matsumoto, Regulation of 3-hydroxyhexanoate composition in PHBH synthesized by recombinant *Cupriavidus necator* H16 from plant oil by using butyrate as a co-substrate. *J. Biosci. Bioeng.* **2015,** 120, (3), 246-251.

43. C. Jiang, S. Y. Kim, and D. Y. Suh, Divergent evolution of the thiolase superfamily and chalcone synthase family. *Mol. Phylogenet. Evol.* **2008,** 49, (3), 691-701.

44. P. Kursula, H. Sikkilä, T. Fukao, N. Kondo, and R. K. Wierenga, High resolution crystal structures of human cytosolic thiolase (CT): A comparison of the active sites of human CT, bacterial thiolase, and bacterial KAS I. *J. Mol. Biol.* **2005,** 347, (1), 189-201.

45. B. Middleton, The kinetic mechanism and properties of the cytoplasmic acetoacetyl-coenzyme A thiolase from rat liver. *Biochem. J* **1974,** 139, (1), 109-121.

46. J. Thissen, S. Desai, P. McCartney, and R. Komuniecki, Improved purification of the pyruvate dehydrogenase complex from *Ascaris suum* body wall muscle and characterization of PDHa kinase activity. *Mol. Biochem. Parasitol.* **1986,** 21, (2), 129-138.

47. M. C. Reed, A. Lieb, and H. F. Nijhout, The biological significance of substrate inhibition: A mechanism with diverse functions. *Bioessays* **2010,** 32, (5), 422-429.

48. Y. Modis and R. K. Wierenga, A biosynthetic thiolase in complex with a reaction intermediate: the crystal structure provides new insights into the catalytic mechanism. *Structure* **1999,** 7, (10), 1279-1290.

49. R. Ofman, J. P. N. Ruiter, M. Feenstra, M. Duran, B. T. Poll-The, J. Zschocke, R. Ensenauer, W. Lehnert, J. O. Sass, W. Sperl, and R. J. A. Wanders, 2-Methyl-3-hydroxybutyryl-CoA dehydrogenase deficiency is caused by mutations in the HADH2 gene. *Am. J. Hum. Genet.* **2003,** 72, (5), 1300-1307.

50. T. A. Leaf and F. Srienc, Metabolic modeling of polyhydroxybutyrate biosynthesis. *Biotechnol. Bioeng.* **1998,** 57, (5), 557-570.

51. J. Soler, D. de Arriaga, Q. Cadenas, and E. Cadenas, Substrate inhibition of lactate dehydrogenase from *Phycomyces blakesleeanus*: NADH dependence. *Exp. Mycol.* **1981,** 5, (4), 357-362.

52. W. C. Cooper, Y. Jin, and T. M. Penning, Elucidation of a complete kinetic mechanism for a mammalian hydroxysteroid dehydrogenase (HSD) and identification of all enzyme forms on the reaction coordinate: the example of rat liver 3alpha-HSD (AKR1C9). *J. Biol. Chem.* **2007,** 282, (46), 33484-93.

53. M. J. Luo, L. F. Mao, and H. Schulz, Short-chain 3-hydroxy-2-methylacyl-CoA dehydrogenase from rat liver: purification and characterization of a novel enzyme of isoleucine metabolism. *Arch. Biochem. Biophys.* **1995,** 321, (1), 214-220.

54. P. Willadsen and H. Eggerer, Substrate stereochemistry of the enoyl-CoA hydratase reaction. *Eur. J. Biochem.* **1975,** 54, (1), 247-52.

55. G. Agnihotri and H. W. Liu, Enoyl-CoA hydratase. reaction, mechanism, and inhibition. *Bioorg. Med. Chem.* **2003,** 11, (1), 9-20.

56. A. Espah Borujeni, A. S. Channarasappa, and H. M. Salis, Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.* **2014,** 42, (4), 2646-59.

57. M. D. McMahon and K. L. J. Prather, Functional screening and *in vitro* analysis reveal thioesterases with enhanced substrate specificity profiles that improve short-chain fatty acid production in *Escherichia coli*. *Appl. Environ. Microbiol.* **2014,** 80, (3), 1042-1050.

58. M. Brock, C. Maerker, A. Schutz, U. Volker, and W. Buckel, Oxidation of propionate to pyruvate in *Escherichia coli*. Involvement of methylcitrate dehydratase and aconitase. *Eur. J. Biochem.* **2002,** 269, (24), 6184-94.

59. C. H. Martin, H. Dhamankar, H.-C. Tseng, M. J. Sheppard, C. R. Reisch, and K. L. J. Prather, A platform pathway for production of 3-hydroxyacids provides a biosynthetic route to 3-hydroxy-γ-butyrolactone. *Nature Communications* **2013,** 4, 1414.

60. A. E. Kazakov, D. A. Rodionov, E. Alm, A. P. Arkin, I. Dubchak, and M. S. Gelfand, Comparative genomics of regulation of fatty acid and branched-chain amino acid utilization in proteobacteria. *J. Bacteriol.* **2008,** 191, (1), 52-64.

61. G. L. Rosano and E. A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **2014,** 5, 172.

62. W. Yuan, Y. Jia, J. Tian, K. D. Snell, U. Müh, A. J. Sinskey, R. H. Lambalot, C. T. Walsh, and J. Stubbe, Class I and III polyhydroxyalkanoate synthases from *Ralstonia eutropha* and *Allochromatium vinosum*: Characterization and substrate specificity studies. *Arch. Biochem. Biophys.* **2001,** 394, (1), 87-98.

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

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

65. M. Cho, C. J. Brigham, A. J. Sinskey, and J. Stubbe, Purification of polyhydroxybutyrate synthase from its native organism, *Ralstonia eutropha*: implications in the initiation and elongation of polymer formation *in vivo*. *Biochemistry* **2012,** 51, (11), 2276-2288.

66. P. P. Peralta-Yahya, F. Zhang, S. B. del Cardayre, and J. D. Keasling, Microbial engineering for the production of advanced biofuels. *Nature* **2012,** 488, (7411), 320-328.

67. S. A. Bonnett, K. Papireddy, S. Higgins, S. del Cardayre, and K. A. Reynolds, Functional characterization of an NADPH dependent 2-alkyl-3-ketoalkanoic acid reductase involved in olefin biosynthesis in *Stenotrophomonas maltophilia*. *Biochemistry* **2011,** 50, (44), 9633-40.

68. B. R. Goblirsch, J. A. Frias, L. P. Wackett, and C. M. Wilmot, Crystal structures of *Xanthomonas campestris* OleA reveal features that promote head-to-head condensation of two long-chain fatty acids. *Biochemistry* **2012,** 51, (20), 4138-4146.

69. C. Dellomonaco, J. M. Clomburg, E. N. Miller, and R. Gonzalez, Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **2011,** 476, (7360), 355-9.

70. H. C. Tseng and K. L. J. Prather, Controlled biosynthesis of odd-chain fuels and chemicals via engineered modular metabolic pathways. *Proceedings of the National Academy of Sciences* **2012,** 109, (44), 17925-17930.

71. A. J. Powell, J. A. Read, M. J. Banfield, F. Gunn-Moore, S. D. Yan, J. Lustbader, A. R. Stern, D. M. Stern, and R. L. Brady, Recognition of structurally diverse substrates by type II 3-hydroxyacyl-CoA dehydrogenase (HADH II)/Amyloid-β binding alcohol dehydrogenase (ABAD). *J. Mol. Biol.* **2000,** 303, (2), 311-327.

72. D. C. Cantu, T. Dai, Z. S. Beversdorf, and P. J. Reilly, Structural classification and properties of ketoacyl reductases, hydroxyacyl dehydratases and enoyl reductases. *Protein Eng. Des. Sel.* **2012,** 25, (12), 803-11.

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

74. A. Baerga-Ortiz, B. Popovic, A. P. Siskos, H. M. O'Hare, D. Spiteller, M. G. Williams, N. Campillo, J. B. Spencer, and P. F. Leadlay, Directed mutagenesis alters the stereochemistry of catalysis by isolated ketoreductase domains from the erythromycin polyketide synthase. *Chem. Biol.* **2006,** 13, (3), 277-285.

75. A. Michinaka, J. Arou, M. Onuki, H. Satoh, and T. Mino, Analysis of polyhydroxyalkanoate (PHA) synthase gene in activated sludge that produces PHA containing 3-hydroxy-2-methylvalerate. *Biotechnol. Bioeng.* **2007,** 96, (5), 871-880.

Chapter 3: *Structural and biochemical characterization of A. suum thiolases: Insight into the basis of substrate selectivity*

3.1 Introduction

Carbon-carbon bonds form the foundation of organic molecules, and the biochemistry of their formation and degradation is fundamental to understanding the structural diversity observed in metabolites in nature. Thiolases are an important class of enzyme that catalyze the reversible formation of carbon-carbon bonds, through Claisen-condensation of thioesters (*Figure 3.1*). Thiolases have critical functions throughout life in the catabolism of fatty acids and amino acids [1], the biosynthesis of steroid compounds in animals and fungi [2, 3], and the biosynthesis of carbon storage polyesters in bacterial species [4]. Thiolases have also been important tools for the design of engineered metabolic pathways for the production of advanced fuels [5-8], polymers [9], and commodity chemicals [10]. We have engineered pathways in *E. coli* for the production of α-methyl organic acids using thiolases and other enzymes from *Ascaris suum*, a branched acid-accumulating roundworm (*Chapter 2*). Obtaining greater insight into the structural and biochemical basis of substrate selectivity in thiolases would help further expand the application of these enzymes in biotechnology

The structure and mechanism of thiolases have been investigated extensively, beginning with studies of porcine heart thiolase [11-13], and especially focusing on biosynthetic thiolases in bacteria [14-19] and the several isozymes of thiolases in humans [20-23]. Thiolase enzymes are a prominent member of the larger thiolase superfamily, which broadly catalyzes Claisen and aldol condensations and includes the ketosynthase enzymes of fatty acid and polyketide synthesis systems, as well as HMG-CoA synthase [1]. The structure of this superfamily is highly conserved, consisting of the βαβαβαββ fold, in which a β-sheet is layered between three α-helices. This fold is repeated at both the N- and C-termini, and in thiolases the N-terminal fold is interrupted by a series of loops and helices involved in substrate binding and quaternary interactions. Catalytic features are also highly conserved in the thiolase superfamily and include a Cys at the Nβ3-Nα3 bend that forms an acyl-enzyme adduct with the substrate, as well as two oxyanion holes, OAH1 and OAH2, that stabilize enolate or tetrahedral intermediates formed in the catalytic cycle [1]. OAH1 is formed by His or Asn residues or coordinated waters at the Cβ2-Cα2 and Cβ3-Cα3 hairpins. In thiolases, for example, OAH1 consists of His348 and water coordinated to N319 (*Figure 3.2*) [24]. OAH2 is formed by the backbone amide nitrogens of both the catalytic Cys and a Gly or other residue at the Nβ4-Nβ5 hairpin [1]. The distinguishing feature of thiolase enzymes, compared to other members of the superfamily, is an additional catalytic Cys at the Nβ4-Nβ5 hairpin, which serves as a general base to form the enolate nucleophile in the direction of condensation, or a general acid in the direction of thiolysis [25-27].

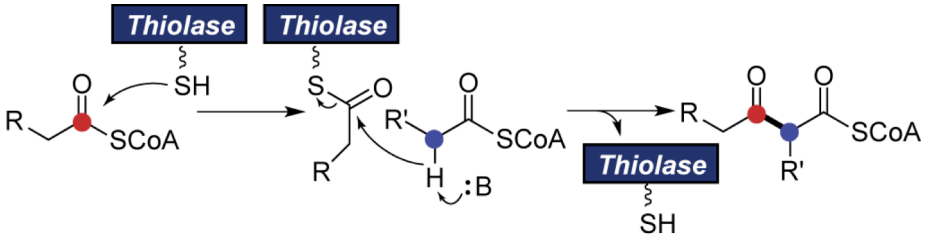


Figure 3.1. *Claisen condensation catalyzed by thiolase enzyme.*

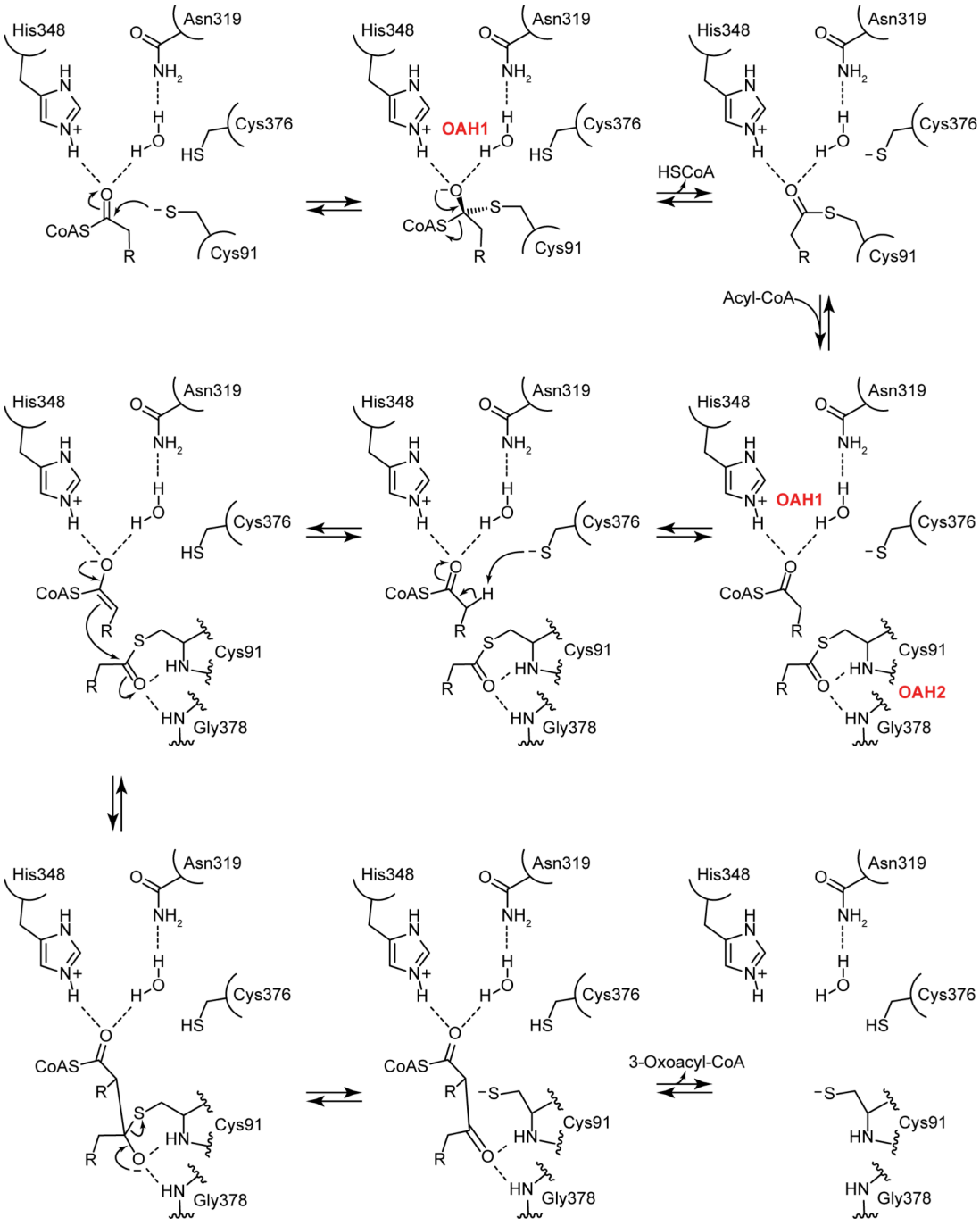


Figure 3.2. *Catalytic mechanism of Claisen condensation in thiolases. Residue numbers are for Acat2. Mechanism of thiolysis follows the same steps in reverse.*

The catalytic mechanism of thiolases is now well understood due to several decades of biochemical and structural work, particularly with the biosynthetic thiolase PhaA of *Zoogloea ramigera* [16-18, 25, 26], and is shown in *Figure 3.2*. It consists of two phases: acyl-enzyme formation and Claisen condensation. Activation by His348, as well as the placement of Cys91 at the N-terminal end of the Nα3 helical dipole, result in an electrostatically neutral active site with deprotonated Cys91 upon substrate binding [17, 22, 24]. Cys91 reacts with the acyl-CoA substrate to form a tetrahedral intermediate stabilized by OAH1, which decomposes to release CoA, forming an acyl-Cys91 adduct. The CoA leaving group abstracts the proton of Cys376 upon exiting, thus retaining an electrostatically neutral active site [17]. Importantly, the acyl-enzyme adduct must rotate to shift the carbonyl from OAH1 to OAH2 for the subsequent Claisen-condensation, but this shift does not appear to happen until the binding of the second acyl-CoA substrate [24]. In the second stage of catalysis, Cys376 deprotonates the acyl-CoA, forming an enolate stabilized by OAH1, which then attacks the acyl-Cys91 thioester to form a tetrahedral intermediate stabilized by OAH2. This resolves to release Cys91 and form the 3-oxoacyl-CoA product. The reaction pathway of thiolysis consists of the same steps in the opposite direction. A notable aspect of the mechanism is that OAH1 must stabilize two different types of oxyanions during the catalytic cycle, a tetrahedral intermediate and an enolate [17].

In addition to the catalytic mechanism, much insight has been gained into the factors that affect the acyl-CoA substrate selectivity or promiscuity of various subclasses of thiolase enzymes [4, 14, 18, 20, 21, 24]. Thiolases have often been roughly categorized as either Thiolase I or Thiolase II, with Thiolase I (EC:2.3.1.16) consisting of degradative thiolases with broad activity over a range of 3-oxoacyl-CoA carbon chain lengths, and Thiolase II (EC:2.3.1.9) consisting of short-chain thiolases that synthesize or degrade 3-oxobutyryl-CoA. More accurately, phylogenetic analysis by Pereto *et al*. has shown that thiolases consist of six clusters, with related thiolases in eukaryotes sharing similar subcellular localizations and functions [28]. Because prokaryotic thiolases cluster among their eukaryotic relatives rather than forming a separate clade, understanding the properties of the major eukaryotic clusters is a useful paradigm for considering thiolase function and selectivity more broadly. Structures of several human thiolase isozymes have been solved, including cytosolic thiolase CT [22] and mitochondrial thiolases T1 [21] and T2 [20]. CT and T2 are both short chain thiolases. In contrast to CT, however, and in accord with its role in the degradation of the branched amino acid isoleucine, T2 is permissive of the branched substrate 2-methyl-3-oxobutyryl-CoA, whose products are acetyl-CoA and propionyl-CoA. Haapalaiten *et al*. identified a Phe residue in the active site of T2 that is highly conserved in this class of thiolase, but is found as Met in CT and T1 thiolases [20]. Computational docking of 2-methyl-3-oxobutyryl-CoA suggested that the α-methyl substituent is directed towards this conserved Phe residue, leading to the hypothesis that this residue is the key factor which leads to permissiveness of α-branched substrates in thiolases.

In our recent studies of thiolase enzymes from *A. suum*, we noted that the thiolases Acat1 through Acat5 exhibited marked differences in selectivity towards the condensation of acetyl-CoA vs. propionyl-CoA, despite their similarity to each other and to T2. These differences were corroborated by further kinetic characterization in the direction of thiolysis. Specifically, Acat2 and Acat3 were permissive of 2-methyl-3-oxoacyl-CoA substrates, whereas Acat5 resembles CT and PhaA in its selectivity for 3-oxobutyryl-CoA. Yet Acat5 shares the distinguishing features of the T2 thiolase class, including the conserved Phe292 residue implicated in branched substrate permissiveness. We therefore set out to gain a better understanding of the structural basis of selectivity in *A. suum* thiolases, which could help inform future efforts in novel thiolase

discovery or engineering. Here we report structures of Acat2 and Acat5, two *A. suum* thiolase isozymes with differing substrate selectivity, as well as a structure of active site mutant Acat2-C91S with propionyl-CoA bound in the active site. We also report mutagenesis studies that implicate residues of the active site covering loop as an important parameter affecting substrate specificity.

3.2 Materials and methods

Commercial materials. LB Broth Miller, LB Agar Miller, 2 × YT broth, and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). TCEP was purchased from Biosynth, Inc. (Itasca, IL). Imidazole was purchased from Acros Organics (Morris Plains, NJ). IPTG was purchased from Santa Cruz Biotechnology (Dallas, TX). Carbenicillin sodium salt (Cb), sodium chloride, potassium chloride, potassium phosphate monobasic, potassium phosphate dibasic, HEPES, and benzonase nuclease were purchased from Fisher Scientific (Pittsburgh, PA). Ethyl 2-methyl-3-oxopentanoate was purchased from Fragmenta (Monmouth Junction, NJ). Ethyl 2-methylacetoacetate, DTNB, coenzyme A sodium salt hydrate (CoA), propionyl-CoA, acetoacetyl-CoA, BME, and lysozyme from chicken egg white were purchased from Sigma-Aldrich (St. Louis, MO). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate, ammonium persulfate, and TEMED were purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, Q5 DNA Polymerase, T4 DNA ligase, Phusion DNA polymerase, T5 exonuclease, and Taq DNA ligase were purchased from New England Biolabs (Ipswich, MA). Deoxynucleotides (dNTPs) were purchased from Invitrogen (Carlsbad, CA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides and gBlocks were purchased from Integrated DNA Technologies (Coralville, IA). DNA purification kits, Ni-NTA agarose, and NeXtal DWBlock crystallography suites were purchased from Qiagen (Valencia, CA). PD-10 desalting columsn were purchased from GE Healthcare (Pittsburgh, PA). cOmplete EDTA-free protease inhibitor were purchased from Roche Applied Science (Penzberg, Germany). TEV protease was purchased from the QB3 MacroLab at UC Berkeley. Amicon Ultra centrifugal concentrators were purchased from Merck Millipore (Cork, Ireland). Reservoir strips and ViewDrop II covers were purchased from TTP Labtech (Melbourn, United Kingdom). 2-methyl-2,4-pentanediol (MDP), Paratone® N, and Intelli-plate 96 sitting drop reservoir plates were purchased from Hampton Research (Aliso Viejo, CA).

Bacterial strains. *E. coli* DH10B-T1^R was used for DNA construction and BL21(de3)-T1^R was used for heterologous production of proteins for purification.

Gene and plasmid construction. Plasmids and PCR primers are listed in *Appendix, Table S2*. Plasmid construction was carried out using standard molecular biology techniques and the Gibson protocol [29]. PCR amplifications were carried out with Q5 DNA polymerase, following manufacturer instructions. Constructs were verified by sequencing (Quintara Biosciences; Berkeley, CA). Acat2-C91S was constructed by amplifying Acat2 gene in two segments, using primers AsAcat2 23a F1 and Acat2-C91S RP for the first segment and primers Acat2-C91S FP and AsAcat2 23a R1 for the second segment. The two segments were inserted into pET23a digested with SfoI and BamHI. Other Acat2 and Acat5 mutant constructs were constructed using a similar strategy.

Expression and purification of thiolase enzymes. Acat2 and Acat5 used for X-ray crystallography were expressed and purified as described in *Chapter 2*. Acat2-C91S mutant used for X-ray crystallography was expressed and purified using the same protocol.

Expression of His_6-tagged thiolases for kinetic characterization. Thiolases used for enzyme kinetics were expressed and purified as follows. 2 × YT broth (1 L) containing carbenecillin (50 μg/mL) in a 2.5 L Ultra Yield flask (Thomas Instrument Company, Oceanside, CA) was inoculated to OD_{600} = 0.05 with an overnight TB culture of freshly transformed *E. coli* containing the pET23a overexpression plasmid. The cultures were grown at 37°C at 250 rpm to OD_{600} = 0.7 to 1.0 at which point cultures were cooled on ice for 15 min, followed by induction of protein expression with IPTG (0.4 mM) and overnight growth at 16°C. Cell pellets were harvested by centrifugation at 6,000 × *g* for 7 min, and then resuspended in 5 mL per gram of wet cell pellet of Buffer 3A (50 mM potassium phosphate, 300 mM NaCl, 5% glycerol, 10 mM imidazole, pH 8.0) supplemented with 10 mM BME, 1 mg/mL lysozyme, and 0.04 μl/mL benzonase at 4ºC. The resuspended pellet was stored at -80ºC until purification.

Purification of His_6-tagged thiolase enzymes for kinetic characterization. Frozen, resuspended cell pellets were thawed in tepid water and then kept on ice. cOmplete EDTA-free protease inhibitor cocktail (Roche) was added, and the cell suspension was homogenized by ten passes with a glass-Teflon homogenizer and lysed by passage through a French pressure cell at 14000 psi. An additional 0.04 μl/mL benzonase was added, and the lysate was cleared by centrifugation for 30 min at 14000 × *g* to remove insoluble debris. The supernatant was passed over Ni-NTA resin pre-equilibrated with Buffer 3A, and washed with Buffer 3A supplemented with 5mM BME followed by Buffer 3B (50 mM potassium phosphate, 300 mM NaCl, 5% glycerol, 34 mM imidazole, pH 8.0) supplemented with 5mM BME. Protein was eluted with Buffer 3C (50 mM potassium phosphate, 300 mM NaCl, 5% glycerol, 250 mM imidazole, pH 8.0), and fractions containing protein were pooled and concentrated using an Amicon Ultra spin concentrator (10 kDa MWCO, Millipore), diluted with Buffer 3D (30 mM HEPES, 150 mM NaCl, 15% glycerol, pH 6.8, 0.4 mM TCEP), and concentrated again. Aliquots were frozen in liquid nitrogen and stored at -80°C. Final protein concentrations were estimated using the $\varepsilon_{280\ nm}$ calculated by ExPASY ProtParam [30] as follows: His_6-Acat2 ($\varepsilon_{280\ nm}$ = 22920 $M^{-1}\ cm^{-1}$), His_6-Acat5 ($\varepsilon_{280\ nm}$ = 25900 $M^{-1}\ cm^{-1}$), His_6-Acat2-Loop5 ($\varepsilon_{280\ nm}$ = 22920 $M^{-1}\ cm^{-1}$), His_6-Acat2-Y153I ($\varepsilon_{280\ nm}$ = 21430 $M^{-1}\ cm^{-1}$), His_6-Acat5-Loop2 ($\varepsilon_{280\ nm}$ = 25900 $M^{-1}\ cm^{-1}$), His_6-Acat5-I154Y ($\varepsilon_{280\ nm}$ = 27390 $M^{-1}\ cm^{-1}$).

Preparation of 3-oxoacyl-CoA compounds. 3-oxoacyl-CoA compounds were prepared as reported in *Chapter 2*.

Crystallization and structure determination. Crystals of Acat2 were obtained using the sitting drop vapor diffusion method by combining 200 nL each of protein solution (10 mg/mL Acat2, 5 mM 2-methyl-3-oxopentanoyl-CoA) and reservoir solution (0.2 M tri-potassium citrate, 20% w/v PEG 3350). Crystals were observed within two days, and a single crystal was obtained by adding 400 nL of cryogenic solution (50% v/v MDP solution containing 5 mM 2-methyl-3-oxopentanoyl-CoA) directly to the crystallization droplet, removing a single crystal within 1 min and flash freezing in liquid nitrogen. The electron density map of the solved structure did not contain density for a bound substrate or for acylation of Cys91.

Crystals of Acat5 were obtained using the hanging drop vapor diffusion method by combining 800 nl each of protein solution (12 mg/mL Acat5, 2 mM 2-oxobutyryl-CoA) and

reservoir solution (2 M sodium malonate, pH 6.5). Crystals were observed within 24 h, and a single crystal was obtained by adding Paratone® N oil to the droplet, transferring a single crystal to the oil layer, and removing residual water around the crystal before flash freezing in liquid nitrogen. The electron density map of the solved structure contained density in the binding pocket and at the Cys92 residue that was best modeled by free CoA and oxidized Cys92.

To obtain crystals of Acat2-C91S, protein was incubated with 2-methyl-3-oxopentanoyl-CoA in attempt to acylate Ser91. Molecular weight determination by LC-TOF suggested successful conversion to the acylated form, before the protein was desalted with a PD-10 column to remove acyl-CoA and concentrated to 10 mg/mL for crystallography. Crystals were obtained using hanging drop vapor diffusion method by combining 200 nL of protein solution (10 mg/ml Acat2-C91S, 4 mM propionyl-CoA) and reservoir solution (0.2 M ammonium nitrate, 20% w/v PEG 3350). Crystals were observed within two days, and a single crystal was transferred to a droplet of 1:1 reservoir solution and cryogenic solution (50% v/v MDP solution with 2 mM propioyl-CoA) for less than 1 min before freezing in liquid nitrogen. The electron density map of the solved structure contained density for propionyl-CoA and density near Ser91, but the latter density was disconnected from Ser91 and best modeled as an adjacent nitrate ion.

Data were collected at beamline 8.3.1 at the Advanced Light Source (Lawrence Berkeley National Laboratory). Data were processed with XDS [31] and scaled and merged with Aimless [32]. Molecular replacement using Phaser [33] was used for phase determination. For Acat2 and Acat5, the molecular replacement model was one monomer of T2 thiolase (PDB 2IB8) with side chains pruned using Sculptor [34] based on sequence alignment with Acat2 or Acat5. The phase_and_build utility of the Phenix software package [35] was used for phase improvement and chain building, and any missing parts of the model were manually built using Coot [36]. For Acat2-C91S, the solved structure of wild-type Acat2 was used for molecular replacement, and the resulting model was used directly for subsequent refinement with Phenix.refine. LigandFit [37] was used to model CoA and propionyl-CoA ligands when sufficient electron density was observed to create a model. Potassium ions and nitrate and acetate ligands were manually modeled with Coot where anomalous un-modeled density was observed. Iterative cycles using Phenix.refine and manual refinement in Coot were used to generate the final model, with Phenix.refine strategies based on XYZ coordinates, real-space refinement, atomic ccupancies, and individual *B*-factors. Since the resolutions exceeded 1.6Å for all three structures, the ReadySet utility was used to place riding hydrogens atoms in a late stage of refinement to improve the model. Chimera [38] was used for visualization of final structures.

Enzyme assays. Assay reagents were quantified as reported in *Chapter 2*. k_{cat} and K_M were determined by fitting to initial rate data using Origin (OriginLab, Northampton, MA) using the equation:

$$v_0 = \frac{k_{cat}[S]}{K_M + [S]}$$

where v_o is the initial rate and [*S*] is the substrate concentration, or using the substrate inhibition equation

$$v_0 = \frac{k_{cat}[S]}{K_M + [S] + \frac{[S]^2}{K_{I,substrate}}}$$

if the calculated $K_{I,substrate}$ was less than 100 times the calculated K_M. For substrates with large K_M values, it was not possible to measure rates in the saturating region. In this case, only the catalytic efficiency k_{cat}/K_M was calculated, by fitting the following equation to the pre-saturation kinetic data.

$$v_0 = \frac{\left(k_{cat}/K_M\right)[S]}{1 + \left([S]/K_M\right)}$$

Acat2, Acat3, and Acat5 thiolases, in direction of thiolysis. Thiolysis activity was measured by monitoring the initial formation of new thioester functional groups at 232 nm. Assays were performed at 25°C in a UV-compatible 96-well plate in a total volume of 200 µl containing 50 mM potassium phosphate, pH 8.0, coenzyme A (100 µM), thiolase enzyme, and 10, 20, 40, 80, 160, 320, or 600 µM 3-oxoacyl-CoA. The reaction was initiated by addition of thiolase enzyme and mixing rapidly by pipette. Assays with Acat5 also included 100 mM KCl.

3.3 Results

Overall structures and dimer interface of Acat2 and Acat5. We obtained crystal structures for Acat2, Acat5, and Acat2-C91S by molecular replacement to < 1.6Å using a modified search model of human T2 thiolase structure (PDB ID 2IB8) (*Table 3.1*). The residue numbering systems are not identical for the two isozymes due to the presence of gaps in the Acat2 sequence when aligned with Acat5 (*Figure 3.3*), as well as our choice to number residues based on the sequence following the expected N-terminal mitochondrial signal peptide. We thus adopted the convention in this study of listing the Acat2 residue first when referring to conserved or corresponding residues in both proteins.

Wild type Acat2 contained two monomers per asymmetric unit, which were derived from adjacent physiological tetramer assemblies in the crystal. Only a single monomer was contained in the asymmetric unit of Acat5. Nevertheless, both Acat2 and Acat5 contain characteristic homotetramer assembly loop domains, as observed among all thiolases of the T1, CT, and T2 clusters. Membrane-bound long-chain thiolases, such as peroxisomal thiolase, are dimers [39]. We previously reported that two other thiolase isozymes from *A. suum*, Acat1 and Acat3, did not maintain a stable assembly in solution, with a lower order species also present (possibly a dimer) that was not catalytically competent (*Chapter 2*). Given that the critical residues for catalysis and substrate binding are found on the same polypeptide chain in thiolases, quaternary structure may be important for catalysis for reasons not fully understand.

Higher-order models could be constructed from copies of the adjacent unit cells (*Figure 3.4A*), revealing the expected dimer of dimers architecture. Acat2 and Acat5 monomers show high similarity when superimposed (C_α average rmsd = 0.92 Å). Due to a small difference in the angle of the two subunits with respect to the tetramer center in Acat2 (56.9º) vs. Acat5 (58.8º) and in the orientation of opposing dimer pairs with respect to the long axis (42.1º vs. 40.7º), the rmsd for the second subunit within the same dimer increases to 1.89 Å when the tetramer is aligned with a single monomer. The greatest differences between single subunits is found between the N- and C-terminal domains of the protein, in the Nα3 helix and small loop leading to Cβ1, as well as in the active site covering loop and subsequent Lα2 helix (*Figure 3.4B*). The Nα3 helix is at the end of the complex and is distal to the active site, substrate binding pocket,

Table 3.1. *Data collection and refinement statistics for Acat2, Acat5, and Acat2-C91S.*

	Acat2	Acat5	Acat2-C91S
		Data collection	
space group	$P22_12_1$	$I222$	$P22_12_1$
cell dimensions			
a, *b*, *c* (Å)	71.57, 88.12, 142.26	71.66, 111.78, 112.76	71.89, 88.62, 88.62
wavelength (Å)	1.1159	1.1159	1.1159
resolution (Å)[a]	75 – 1.53 (1.585 – 1.53)	79 – 1.60 (1.657 – 1.60)	75 – 1.50 (1.554 – 1.50)
R_{merge} (%)[a]	6.7 (38.7)	8.0 (28.6)	7.9 (42.6)
completeness (%)[a]	99.6 (96.2)	99.6 (96.4)	99.9 (100)
$<I/\sigma I>$[a]	11.50 (2.49)	15.67 (3.08)	9.03 (2.60)
no. of unique reflections[a]	135412 (12914)	59736 (5698)	144700 (14328)
redundancy[a]	4.0 (3.0)	10.1 (4.2)	4.0 (4.0)
B-factor from Wilson plot (Å^2)	12.27	17.55	12.10
		Refinement statistics	
resolution (Å)	75 – 1.53	79 – 1.60	75 – 1.50
no. of reflections			
in the working set	128707	56623	137598
in the test set	6703	3071	7086
R_{work}/R_{free}	0.139/0.160	0.146/0.173	0.149/0.173
no. of non-hydrogen atoms	6526	3269	6641
protein	5656	2842	5642
ligands and ions	2	52	120
water	868	375	879
rmsd			
bond lengths (Å)	0.009	0.010	0.010
bond angles (deg)	1.29	1.33	1.37
average B-factor	16.10	24.40	16.90
protein	14.40	22.70	15.10
ligands and ions	10.20	54.80	18.60
water	27.70	33.10	28.20
Ramachandran plot			
favored (%)	98	98	99
outliers (%)	0	0.53	0

[a]Values in parentheses are for the highest-resolution shell.

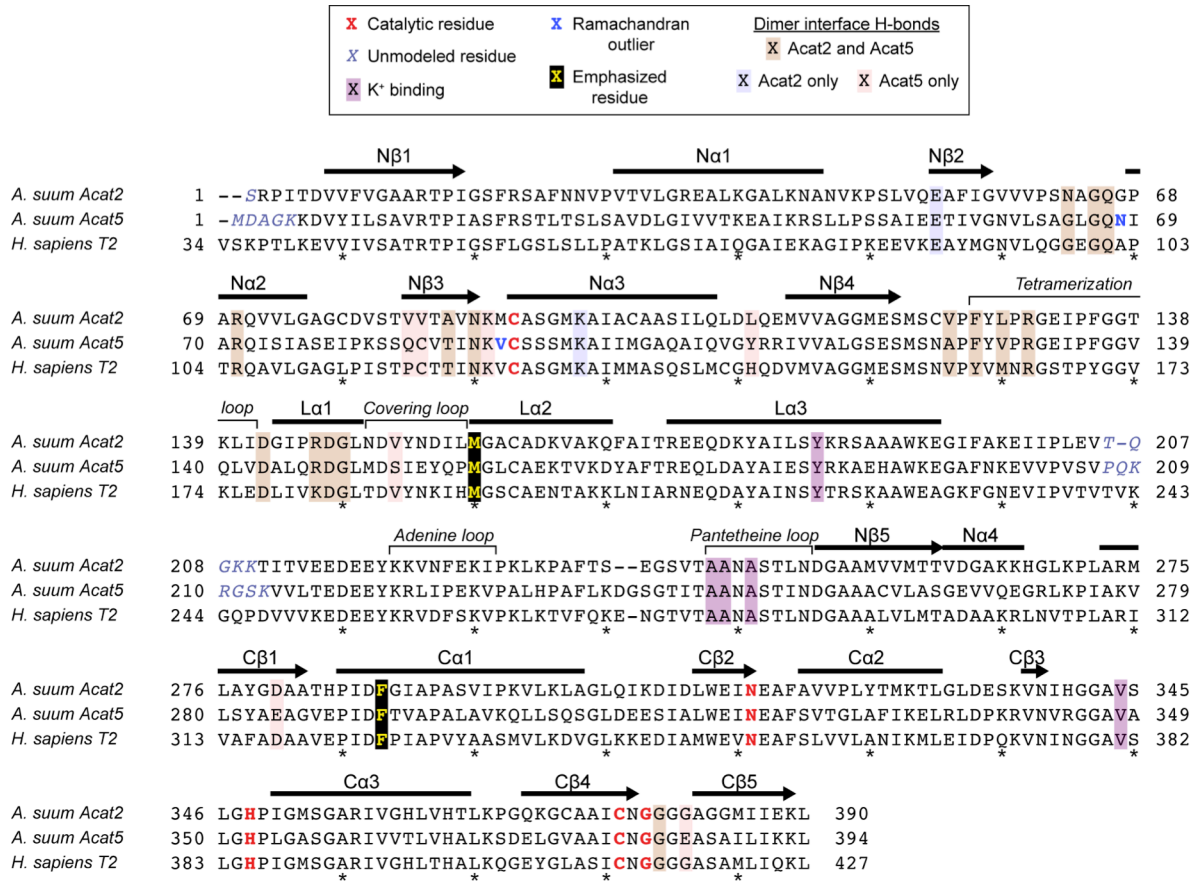


Figure 3.3. *Sequence alignment of Acat2, Acat3, and human T2 thiolases. Catalytic, unmodeled, and other emphasized residues are indicated based on the key. Residues involved in polar interactions at the dimer interface are also highlighted.*

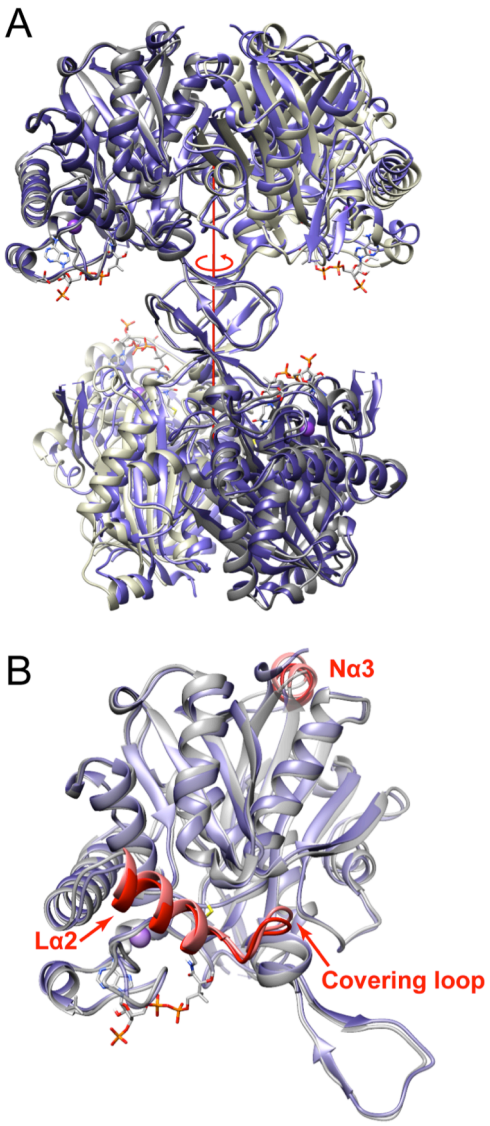


Figure 3.4. *Overall structures of Acat2 and Acat5. (A) Tetramer models of Acat2 (dark slate blue) and Acat5 (gray), aligned by the upper left subunit. (B) Single subunit of Acat2 (slate blue) and Acat5 (gray), with regions of highest rmsd highlighted in red for Acat2 and pink for Acat5.*

and inter-subunit contacts. The covering loop and Lα2, however, form a part of the substrate binding cavity, across from the pantetheine loop and catalytic residues, and shield the active site from solvent. Superposition of wild type Acat2 with the structure of active site mutant Acat2-C91S containing bound substrate showed essentially no change in the position of Cα atoms, with rmsd < 0.25 Å.

As in most thiolases, the opposing dimer pairs of Acat2 and Acat5 contain few contacts other than those within the tetramerization loops. Haapalaiten *et al.* reported hydrogen bonding in human T2 between the side-chain of a Tyr of the tetramerization loop and the pantetheine moiety of CoA substrate in the opposing dimer, mediated by a water molecule [20], but this interaction is absent in Acat2/Acat5, as the Tyr is replaced by Phe135/136. A role has also been proposed for the cationic loop domain in recruiting and binding substrate on the opposite dimer through interactions with the charged phosphates of CoA. This region contains between one and three Lys or Arg residues and is often highly disordered. We were unable to model residues in this loop for Acat2 or Acat5 with any certainty due to insufficient density and opted to allow a break in the chain between Thr206-Lys210 for Acat2 and Pro207-Lys213 for Acat5.

In contrast to opposing dimer pairs, the lateral interface between subunits of a dimer contain a number of polar interactions (*Table 3.2*). Overall, Acat5 has more polar interactions between dimer subunits than Acat2 or many other thiolases, with an additional five hydrogen bonds per monomer. As illustrated in *Figure 3.5A*, these additional interactions complete a collection of hydrogen bonds that line the entire upper half of the lateral interface. Kursula *et al*. have suggested that increased rigidity in PhaA from inter-subunit hydrogen bonds may contribute both

Table 3.2. *Contact distances for hydrogen and ionic bonds between dimer subunits.*

Enzyme	Donor atom (Acat2/Acat5)	Acceptor atom (Acat2/Acat5)	Distance (Å) (Acat2/Acat5)
Acat2 & Acat5	N(Asn63/Gly64)	OD2(Asp147/148)	2.8/2.9
	N(Gly65/66)	OD1(Asp147/148)	2.9
	NE2(Gln66/67)	O(Gly148/149)	2.8
	NE(Arg70/71)	O(Gly379/383)	2.8
	NH2(Arg79/71)	O(Gly379/383)	3.2/2.9
	N(Ala86/Thr87)	O(Asn88/89)	3.2/2.9
	N(Asn88/89)	O(Ala86/Thr87)	2.9/2.8
	N(Phe126/127)	O(Leu128/Val129)	2.9/2.8
	N(Leu128/Val129)	O(Phe126/127)	2.9/3.0
	NE(Arg130/131)	O(Val124/Ala125)	2.9
	NH1(Arg130/131)	OD2(Asp142/Asp143)	3.0
	NH2(Arg130/131)	O(Val124/Ala125)	3.0/2.9
	NH2(Arg130/131)	OD1(Asp142/Asp143)	2.5/2.7
	NE(Arg146/147)	O(Asn63/Gly64)	2.9/3.1
	NH2(Arg146/147)	O(Asn64/Gly64)	3.3/2.8
Acat2 only	NZ(Lys96)	OE2(Glu53)	2.8
	NE(Arg146)	OD1(Asn63)	3.1
Acat5 only	N(Gln67)	OG(Ser153)	2.8
	NE2(Gln84)	OE1(Glu385)	3.2
	N(Cys85)	OE2(Glu385)	2.9
	NZ(Lys90)	O(Cys85)	3.0
	OH(Tyr110)	OE1(Glu284)	2.5

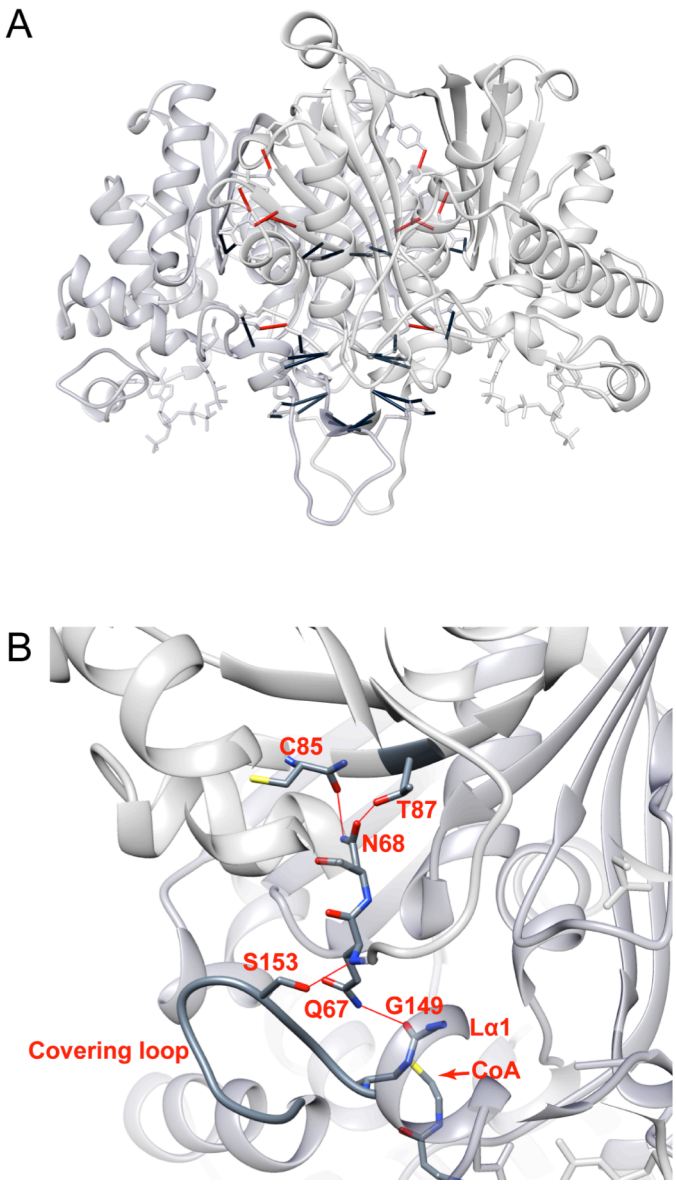


Figure 3.5. *Hydrogen bonding in Acat5. (A) Acat5 dimer with hydrogen bonds and ionic contacts between subunits indicated by solid lines. Black lines incidate contacts found in both Acat2 and Acat5. Red lines indicate contacts particular to Acat5. (B) Hydrogen bonding between the covering loop and opposing subunit in Acat5. Hydrogen bonds are indicated by solid red lines.*

to the high efficiency of PhaA and also to its more limited selectivity [24]. Of particular interest is a hydrogen bond in Acat5 between OG(Ser153) of the covering loop and N(Gln67) of the opposing subunit (*Figure 3.5B*). Gln67 is conserved among CT and T2 thiolases, but not among the T1 thiolases, which have quite different sequences in pre-Nα2 region. It has been suggested that Gly67 plays a role, along with the large helix size of Lα1, in limiting the ω chain length selectivity of the 3-oxoacyl-CoA substrate by protruding slightly further into the cavity than the corresponding Tyr residue in T1 [21]. While the hydrogen bond between NE2(Gln67) and O(Gly149) of Lα1 is conserved, the additional bond with Ser153 is unique to Acat5 and could help draw the opposing subunit even closer to the covering loop and active site cavity.

Our Acat5 structure contained two residues with unfavorable Ramachandran statistics – Asn68 (φ = 31°; ψ = 82°), the residue immediately following Gln67, and Val91 (φ = 45°; ψ = -129°), which precedes the catalytic Cys92. These residues were well defined in the electron density and both have been found to have unfavorable φ and ψ values in other reported thiolase structures [22, 24]. Asn68 is not found in most T2 thiolases, and is more characteristic of CT-type. In Acat5 the side chain of Asn68 bends upward to form two hydrogen bonds with Cys85 and Thr87 of the Nβ3 strand of the same monomer (*Figure 3.5B*).

Mode of substrate binding. Despite the fact that the wild type Acat2 and Acat5 proteins were crystallized with 2-methyl-3-oxovaleryl-CoA and 3-oxobutyryl-CoA, respectively, present in the protein solution, the wild type Acat2 structure did not contain electron density in the substrate binding pocket, and Acat5 was best modeled bound with free CoA product, which likely arose from the hydrolysis of 3-oxobutyryl-CoA in solution. We were able to obtain a structure of an Acat2-C91S mutant bound with propionyl-CoA, however, and this ligand is well defined by the electron density map. The CoA portion of the propionyl-CoA substrate aligns extremely well with the free CoA bound to human T2 reported by Haapalainen *et al.* (PDB ID 2IBW) (*Figure 3.6A*) [20]. The mode of binding is substantially similar, with Ser247/251 providing direct hydrogen bonds with the N1 and N2 atoms of the pantetheine arm, and all other polar interactions with pantetheine mediated by water molecules that remain present even in the apo structure. Whereas the O3 atom of CoA had no notable interactions with T2, in Acat2-C91S this atom forms a hydrogen bond with the amide nitrogen of Gly159/160. Otherwise, hydrophobic residues including Leu149/150, Phe235/237, Leu249/I253, Phe319/323, and Ile350/Leu354 line the pantetheine binding cavity. In other reported thiolase structures, though not T2, the O2 atom of CoA forms a hydrogen bond mediated through a water with a conserved His157, but in Acat2 and Acat5 this residue is Leu157/Pro158. Instead Leu157 in Acat2 further adds to the hydrophobicity in the pocket.

CoA ligand in Acat5 is only loosely bound, as seen by the much higher *B*-factors compared with propionyl-CoA, particularly in the central diphosphate and pantoic acid sections of the pantetheine arm (*Figure 3.6B*). The pantetheine group does not sit as deep in the binding pocket in Acat5 as in Acat2, suggesting a transitory state in which CoA product is leaving the active site. The adenine moiety is nevertheless well defined in both propionyl-CoA and CoA ligands, consistent with the observation by Haapalaiten *et al* that T2-type thiolases have additional features that provide tight binding of this moiety [20], including hydrogen bonds with Asn224/Ile226, the hydroxyl of Tyr184/185, and a highly coordinated water. One difference is that the N6A amine is not close enough to the water to form an additional interaction seen in human T2. A potassium ion is bound in wild type Acat2 in the same way reported for T2, coordinated by several residues of the pantetheine loop as well as Tyr184/185 and the

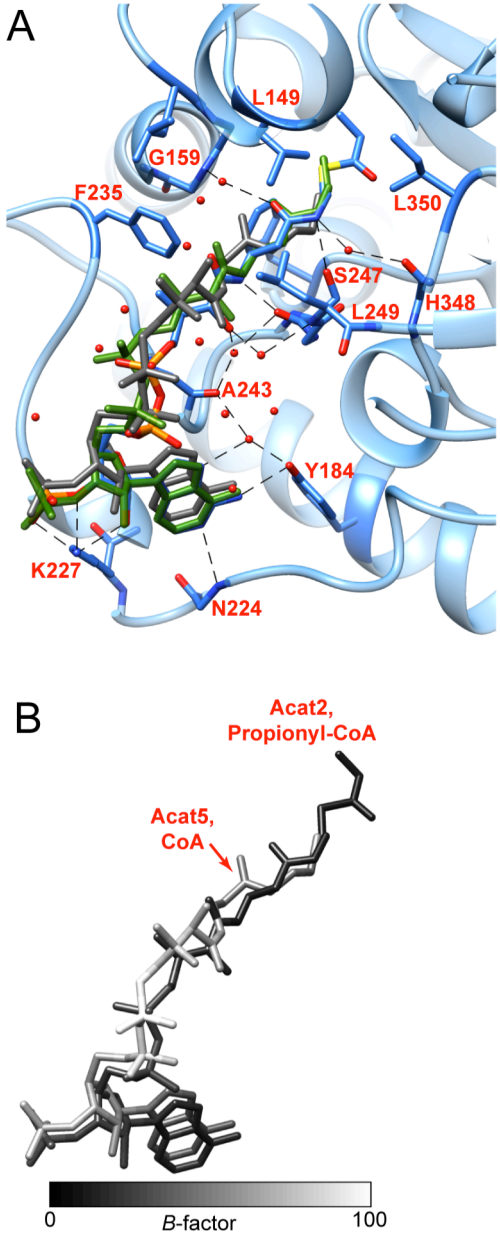


Figure 3.6. *Binding modes of substrates in Acat2 and Acat5. Residue numbering refers to Acat2. (A) Substrate binding pocket of Acat2-C91S. Protein and propionyl-CoA are in blue, with ribbon in a lighter shade. Also shown are CoA ligands from Acat5 (gray) and human T2 (dark green) (PDB ID 2IBW) [20]. (B) Propionyl-CoA from Acat2-C91S structure and CoA from Acat5 colored by B-factor.*

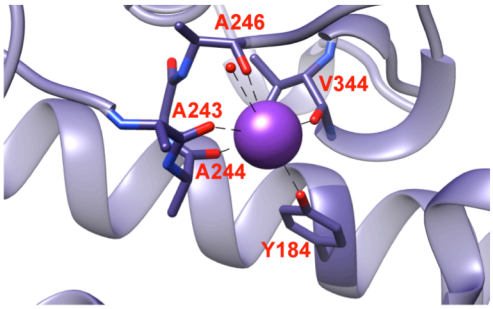


Figure 3.7. *Coordination of K^+ ion in apo-Acat2.*

coordinated water (*Figure 3.7*). We have previously observed that, unlike T2, Acat2 and Acat5 do not display an increase in activity in the presence of KCl. The Acat2-C91S and Acat5 structures did not contain potassium in the crystallization solution, and no potassium ion was observed based on the electron density. In Acat2-C91S, the ion is simply replaced by an additional water. Consistent with human T2, the positions of other waters and side chain atoms in the adenine and pantetheine loops remain the same between various structures, with the exception of Lys227/229, which rotates out towards the solvent in the unbound Acat2 structure but is turned inward in Acat2-C91S and Acat5 to form polar and ionic interactions with the ribose 3′-phosphate and 2′- and 3′-hydroxyl groups. We did not observe density for chloride ion in our Acat2 and Acat5 structures.

The positions and oxidation states of Cys91 in Acat2 (Ser91 in Acat2-C91S) and Cys92 in Acat5 are analogous to what was reported for human T2 [20]. Specifically, a slight rotation of Cys91(Ser91) towards His348 is seen between apo Acat2 compared to either bound Acat2-C91S or bound Acat5, and also Acat5 Cys92 was best modeled as an oxidized cysteine sulfoxide, with the oxygen seated in OAH2. In Acat2-C91S, the carbonyl of propionyl-CoA is seated in OAH1 formed by His348 and Wat26, the latter hydrogen bonded to Asn319 (*Figure 3.8A*). Since OAH1 is involved in the initial formation of the acyl-enzyme intermediate as well as in enolate stabilization in the Claisen condensation step, this position of propionyl-CoA in the structure appears to be consistent with either intermediate. In the first case, propionyl-CoA would be the electrophile and contribute to extending the linear chain length of the 3-oxoacyl-CoA product. But in the second case, propionyl-CoA is the nucleophile, and the terminal methyl group installs an α-methyl substituent in the product. Cys376 is well positioned to abstract the *pro*-S α-proton of propionyl-CoA, and the plane of this resulting nucleophile is directly above a nitrate ion from the crystallization solution that is bound in OAH2. In contrast, the geometry of Ser91 with respect to the plane of the propionyl-CoA thioester is not suitable for nucleophilic attack (*Figure 3.8B*). In the initial binding of propionyl-CoA or other acyl-CoA at the start of a catalytic cycle, the thioester would be expected to be rotated so that the plane is perpendicular to OS(Cys91). The substrate bound Acat2-C91S structure therefore appears to represent the intermediate preceding Claisen condensation, with propionyl-CoA positioned as the nucleophile and the nitrate ion occupying OAH2 acting as a substitute for acyl-enzyme.

Like CT and T2 thiolases, Acat2 and Acat5 work poorly with acyl-CoA substrates of four or more carbons (to give products with main chain length of 6 or more). When the first substrate binds, the acyl chain would point downwards towards Met90/Val91, Gln66/67, and the end of Lα1, and continue on to a binding pocket formed by Ile350/L354 and Val59/60 and terminating

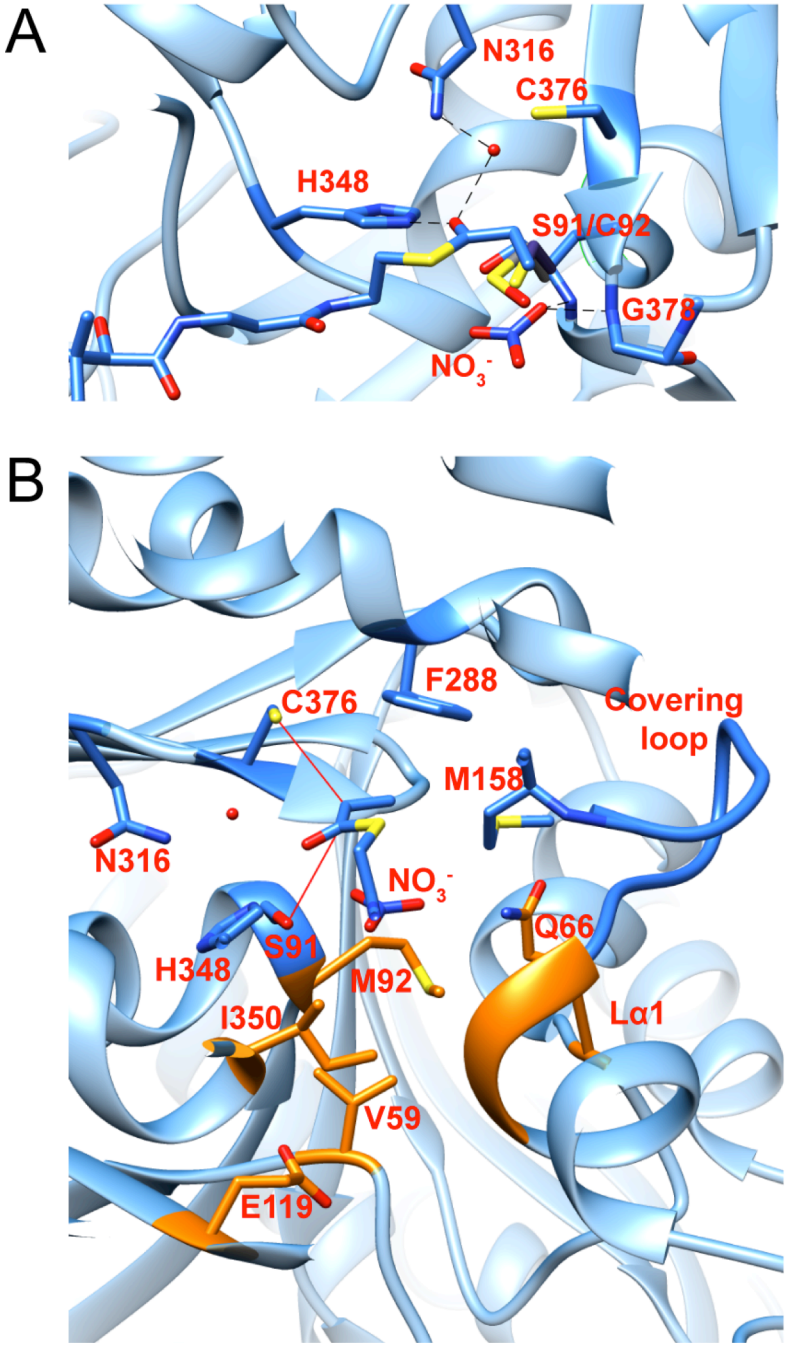


Figure 3.8. *Binding of propionyl-CoA in the active site. (A) Acat2-C91S and propionyl-CoA shown in blue, with ribbon in a lighter shade. For conserved residues, numbering refers to Acat2. Side chains of Cys91/92 in apo Acat2 and Acat5 are also shown. Acat5-Cys92 is oxidized to the sulfenic acid. The thioester carbonyl of propionyl-CoA occupies OAH1 formed by His348 and coordinated water. Nitrate ion occupies OAH2 formed by amide nitrogens of Ser91 and Gly378. Dashed black lines show hydrogen bond contacts. (B) Active site and substrate binding pocket of Acat2-C91S. Red lines indicate geometry of carbonyl and α-carbons of propionyl-CoA with respect to Ser91 and Cys376. Conserved residues of the long chain acyl-CoA binding pocket described by [21] et al. are highlighted in orange.*

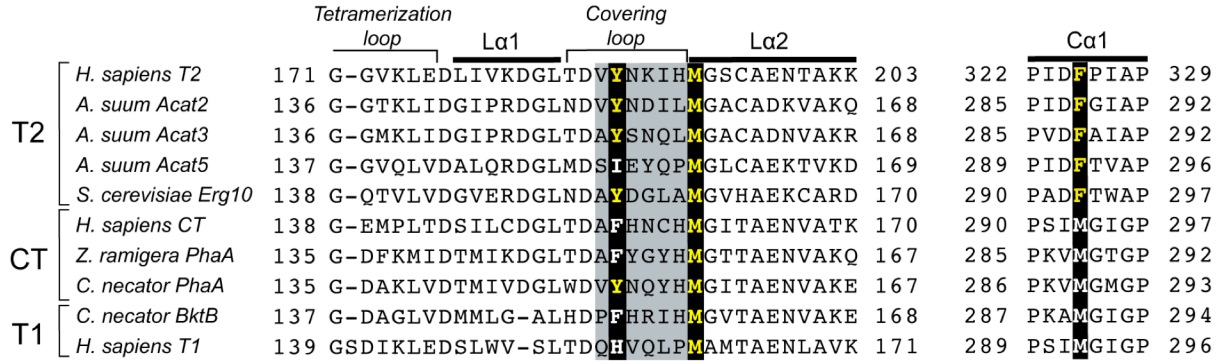


Figure 3.9. *Alignment of eukaryotic and prokaryotic thiolases. Residues of the covering loop targeted for mutagenesis are highlighted in gray. Tyr153, Met158, and Phe288 are shaded in black, and conserved residues in these positions are highlighted in yellow.*

at Glu119/120 as described by Kiema *et al*. for human T1, a long chain thiolase [21] (*Figure 3.8B*). This cavity, however, is blocked off in Acat2 and 5, primarily by the Lα1 helix, which is larger due to an additional Glu residue in CT and T2 thiolases that is not found in T1, and partially by Q66. Fage *et al.* highlighted the importance of the different Lα1 helix length in the selectivity of BktB, a T1-like bacterial enzyme [14]. In eukaryotic T1 thiolases, the fact that Lα1 is shorter can easily be obscured by the fact that these enzymes contain an additional Ser or Leu near the end of the tetramerization loop as well as the missing Glu residue. A structure-guided alignment helps to clarify the cause for long chain permissiveness in T1 (*Figure 3.9*). Since these insertion/deletion events are only 11 residues apart from each other and the Lα1 region is highly variable, standard sequence alignments tend to omit these gaps and create the impression that Lα1 is the same length in T1, T2, and CT-like thiolases.

Branched chain selectivity and role of the covering loop. The terminal methyl group of propionyl-CoA in Acat2-C91S fits precisely between the conserved residues, Phe288/292 and Met158/159 (*Figure 3.10A*). Haapalaiten *et al.* had predicted using computational docking that 2S-methyl-3-oxobutyryl-CoA would bind with the branching methyl pointing more directly towards the Phe residue ring face in T2 [20]. After enolization of propionyl-CoA followed by re-tetrahedralization of the α-carbon through Claisen condensation, the α-methyl in the 3-oxoacyl-CoA product should bend a bit closer towards the Phe residue, in accord with the prediction for T2. No change in the Phe288 side chain is observed between the apo-Acat2 and holo-Acat2-C91S structures, but in Acat5 the ring is rotated by 47º towards the substrate, decreasing the space available for a methyl substituent. This rotation appears to be due to the presence of Thr293, which is found as a Gly or Pro in Acat2, Acat3, and human T2. The hydroxyl group of Thr293 is directed towards the ring face of Phe292, and possibly forms a weak hydrogen bonding interaction with the π-system of the ring. The normal distance of OG(Thr293) to the plane of the ring is 3.5Å, and the closest point is just inside the ring from CZ(Phe292). Erg10 from *Saccharomyces cerevisiae* also contains a Phe-Thr motif at this position in the protein. The phylogenetic analysis by Pereto *et al*. classifies Erg10 with T2 thiolases [28], although Erg10 is cytosolic and its physiological function is analogous to human CT [2, 3]. Whether Erg10 is permissive of branched 3-oxoacyl-CoA substrates has not been reported.

Met158 undergoes a conformational shift upon binding of propionyl-CoA, turning up towards the top of the binding pocket formed by the active site covering loop (*Figure 3.10*). In

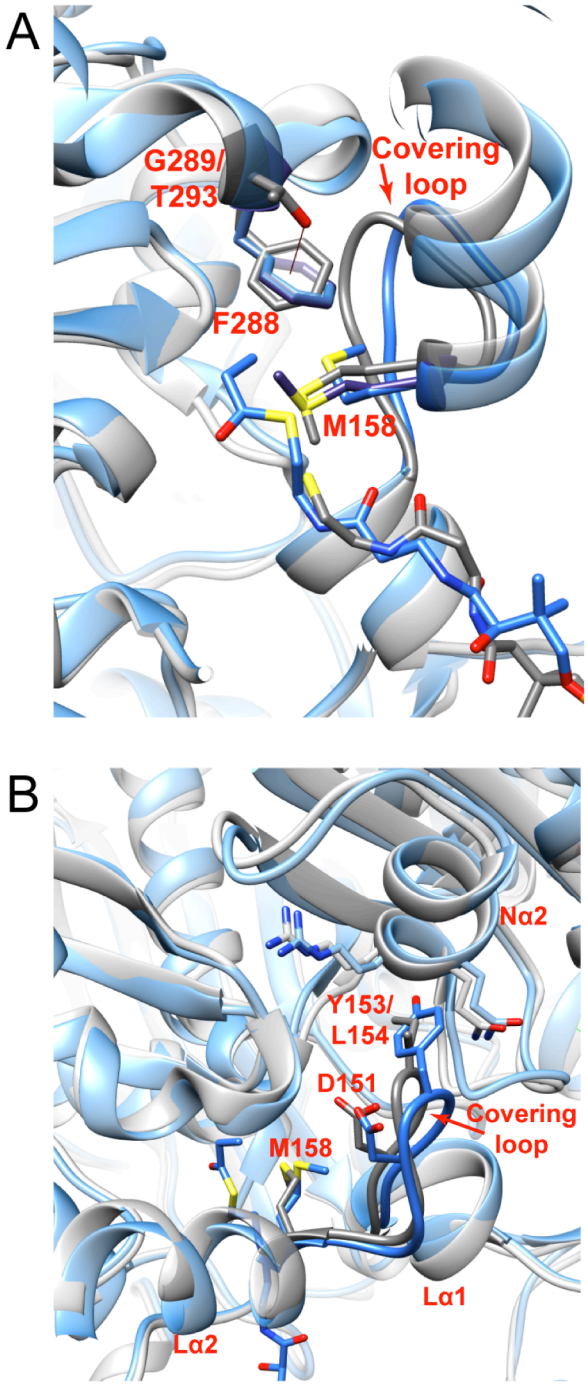


Figure 3.10. *Structural features affecting substrate selectivity in Acat2 and Acat5. Acat2-C91S and Acat5 protein and ligands are shown in blue and gray, respectively, with ribbons in a lighter shade. For conserved residues, numbering refers to Acat2. (A) The terminal methyl of propionyl-CoA in Acat2-C91S is directed towards Met158 and Phe288. Acat5-Phe289 is rotated towards Thr293. Met158 undergoes a conformation change upon binding. Met158 side chain of apo-Acat2 is in dark slate blue. (B) View of the binding pocket from above Lα2 and the covering loop. Tyr153 is in close proximity to Nα2.*

the structures of both apo-Acat2 and holo-Acat2-C91S, the movement of Met158 is well defined with no evidence of secondary conformations based on the electron density (*B*-factors of 12.4 Å and 10.2 Å, respectively). While this movement was also observed in T2 thiolase upon binding of free CoA, the position of Met193 is fluxional between both conformations in the apo T2 structure. The position of Met 159 in Acat5 is also not well-defined, with atomic *B*-factors ranging from 21Å to 36Å moving from the CB to CE atoms. This fluctionality is consistent with the understanding of the Acat5 structure as transitory with only loosely bound CoA. Structures of CT and T1 thiolases, including the entire series of *Z. ramigera* PhaA structures reported by Kursula *et al*. that covers all the intermediates of the catalytic cycle [24], contain the corresponding Met158 residue in the confirmation turned up towards the covering loop, as in the bound T2 state. These thiolases contain an additional Met residue in place of Phe288/292 (*Figure 3.7*). The chemical environment may be somewhat different in these active sites, as Meriläinen *et al*. have found that sulfur-sulfur interactions are important for tight binding of CoA substrates in PhaA [40].

Because the movement of Met158 brings it in closer proximity to the highly sequence-variable covering loop that forms the top of the binding pocket, we were interested in gaining a better understanding of the role of this region in governing substrate selectivity. The closest residues to Met158 in the substrate bound conformation are Asp151 of the covering loop and Gln66 of the neighboring subunit (*Figure 3.10B*), both of which are highly conserved. Looking to a sequence alignment of the covering loop between several representative thiolases, we identified Acat2 Tyr153 as a residue shared by thiolases permissive of α-methyl substrates that is not found in the most of the thiolases selective for linear substrates (*Figure 3.9*). Tyr is a bulky residue, and the polar phenolic hydroxyl of Tyr153 places it within 3.2Å of N(Gln71) of the Nα2 helix on the opposite subunit, but not with an orientation favorable for formation of a hydrogen bond. We hypothesized that the presence of Tyr153 might facilitate the adoption of a slightly more spacious binding pocket in Acat2, allowing greater permissiveness of α-methyl substrates. Conversely, in Acat5 this residue is Ile154, a smaller amino acid that could also more easily engage in hydrophobic interactions with the CG methylenes of Arg71 and Gln72. Coupled with the additional hydrogen bond present between Ser153 and Gln67 in Acat5, Ile154 may facilitate adoption of a less spacious binding pocket in Acat5.

To test this hypothesis, we cloned and purified Acat2-Y153I and Acat5-I154Y mutants and measured kinetic parameters with 3-oxobutyryl-CoA, 2-methyl-3-oxobutyry-CoA, and 2-methyl-3-oxopentanonoyl-CoA substrates. We also tested Acat2-Loop5 and Acat5-Loop2, mutants in which the six covering loop residues located between the conserved Asp151/152 and Met158/159 were swapped between Acat2 and Acat5 (*Figure 3.11*). The results are displayed in *Table 3.3* and *Figure 3.12*. Wild type Acat5 shows defects in catalytic efficiency of 91-fold and 2900-fold with the branched substrates 2-methyl-3-oxobutyryl-CoA and 2-methyl-3-pentanoyl-CoA. Exchanging six residues of the covering loop of Acat5 with those of Acat2 lowers the efficiency with 3-oxobutyryl-CoA by 6.4-fold, but increases the efficiency with the two branched substrates by 6.3-fold and 3.3-fold. This results in a 40-fold increase in the relative efficiency with 2-methyl-3-oxobutyryl-CoA compared to the linear substrate. Characterization of Acat5-I154Y shows that a substantial portion (11-fold) of this 40-fold increase in relative permissiveness arises from this particular point mutation. In contrast, characterization of Acat2-Loop5 and Acat2-Y153I showed that Acat2 was relatively agnostic to these mutations in terms of their affect on the enzyme's relative permissiveness towards α-methyl substrates.

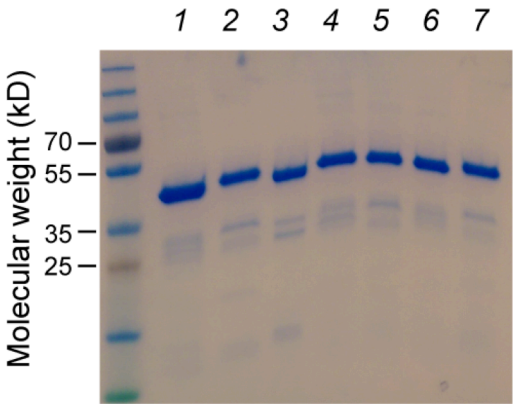


Figure 3.11. *SDS-PAGE of purified proteins. 1. His_6-Acat2, 2. His_6-Acat5, 3. Acat2-C91S, 4. His_6-Acat2-Loop5, 5. His_6-Acat5-Loop2, 6. His_6-Acat2-Y153I, 7. His_6-Acat5-I154Y.*

Table 3.3. *Kinetics constants measured for thiolysis of 3-oxoacyl-CoA by wild type and mutants of Acat2 and Acat5.*[a]

Enzyme	Substrate	Relative k_{cat}/K_M	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_M (μM)	$K_{I,\ substrate}$ (μM)
Acat2	3-Oxobutyryl-CoA	1	$(3.2 \pm 0.5) \times 10^6$	$(3.40 \pm 0.09) \times 10^1$	$(1.1 \pm 0.2) \times 10^1$	-
	2-Me-3-oxobutyryl-CoA	0.6 ± 0.1	$(1.8 \pm 0.2) \times 10^6$	$(7.8 \pm 0.2) \times 10^1$	$(4.2 \pm 0.4) \times 10^1$	-
	2-Me-3-oxopentanoyl-CoA	0.6 ± 0.1	$(2.1 \pm 0.2) \times 10^6$	$(1.01 \pm 0.02) \times 10^2$	$(4.8 \pm 0.4) \times 10^1$	-
Acat2 -Loop5	3-Oxobutyryl-CoA	1	$(3.4 \pm 0.6) \times 10^5$	$(2.2 \pm 0.2) \times 10^1$	$(7 \pm 1) \times 10^1$	$(1.0 \pm 0.3) \times 10^3$
	2-Me-3-oxobutyryl-CoA	0.5 ± 0.1	$(1.9 \pm 0.1) \times 10^5$	$(3.34 \pm 0.07) \times 10^1$	$(1.8 \pm 0.1) \times 10^2$	-
	2-Me-3-oxopentanoyl-CoA	1.6 ± 0.3	$(5.4 \pm 0.6) \times 10^5$	$(4.0 \pm 0.1) \times 10^1$	$(7.5 \pm 0.8) \times 10^1$	-
Acat2 -Y153I	3-Oxobutyryl-CoA	1	$(1.6 \pm 0.2) \times 10^6$	$(3.7 \pm 0.2) \times 10^1$	$(2.3 \pm 0.3) \times 10^1$	$(1.8 \pm 0.4) \times 10^3$
	2-Me-3-oxobutyryl-CoA	0.48 ± 0.07	$(7.7 \pm 0.7) \times 10^5$	$(5.3 \pm 0.2) \times 10^1$	$(7.0 \pm 0.6) \times 10^1$	-
	2-Me-3-oxopentanoyl-CoA	0.8 ± 0.1	$(1.3 \pm 0.1) \times 10^6$	$(8.0 \pm 0.2) \times 10^1$	$(6.0 \pm 0.6) \times 10^1$	-
Enzyme	Substrate	Relative k_{cat}/K_M	k_{cat}/K_M ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_M (μM)	$K_{I,\ substrate}$ (μM)
Acat5	3-Oxobutyryl-CoA	1	$(6.3 \pm 0.9) \times 10^6$	$(7.3 \pm 0.2) \times 10^1$	$(1.2 \pm 0.2) \times 10^1$	-
	2-Me-3-oxobutyryl-CoA	0.011 ± 0.002	$(7.0 \pm 0.3) \times 10^4$	> 20	$(> 2) \times 10^2$	-
	2-Me-3-oxopentanoyl-CoA	0.00034 ± 0.00006	$(2.2 \pm 0.2) \times 10^3$	> 0.7	$(> 3) \times 10^2$	-
Acat5 -Loop2	3-Oxobutyryl-CoA	1	$(9.9 \pm 0.8) \times 10^5$	$(2.55 \pm 0.05) \times 10^1$	$(2.6 \pm 0.2) \times 10^1$	-
	2-Me-3-oxobutyryl-CoA	0.44 ± 0.06	$(4.4 \pm 0.4) \times 10^5$	$(9.7 \pm 0.4) \times 10^1$	$(2.2 \pm 0.2) \times 10^2$	-
	2-Me-3-oxopentanoyl-CoA	0.0073 ± 0.0006	$(7.3 \pm 0.2) \times 10^3$	> 2.5	$(> 3) \times 10^2$	-
Acat2 -I154Y	3-Oxobutyryl-CoA	1	$(1.9 \pm 0.3) \times 10^6$	$(6.4 \pm 0.3) \times 10^1$	$(3.4 \pm 0.4) \times 10^1$	$(1.2 \pm 0.2) \times 10^3$
	2-Me-3-oxobutyryl-CoA	0.12 ± 0.02	$(2.2 \pm 0.2) \times 10^5$	> 50	$(> 2) \times 10^2$	-
	2-Me-3-oxopentanoyl-CoA	0.0027 ± 0.0004	$(5.0 \pm 0.1) \times 10^3$	> 2.5	$(> 3) \times 10^2$	-

[a] Data are mean ± SE as determined by non-linear curve fitting. Entries with $K_{I,substrate}$ value were fitted to the substrate inhibition equation. Error in the relative k_{cat}/K_M was obtained from propagation of error from the individual k_{cat}/K_M. Error in k_{cat}/K_M was obtained from propagation of error from the individual kinetics terms, or obtained directly from curve fitting for data sets that did not reach substrate satuation, with lower bounds given for k_{cat} and K_M.

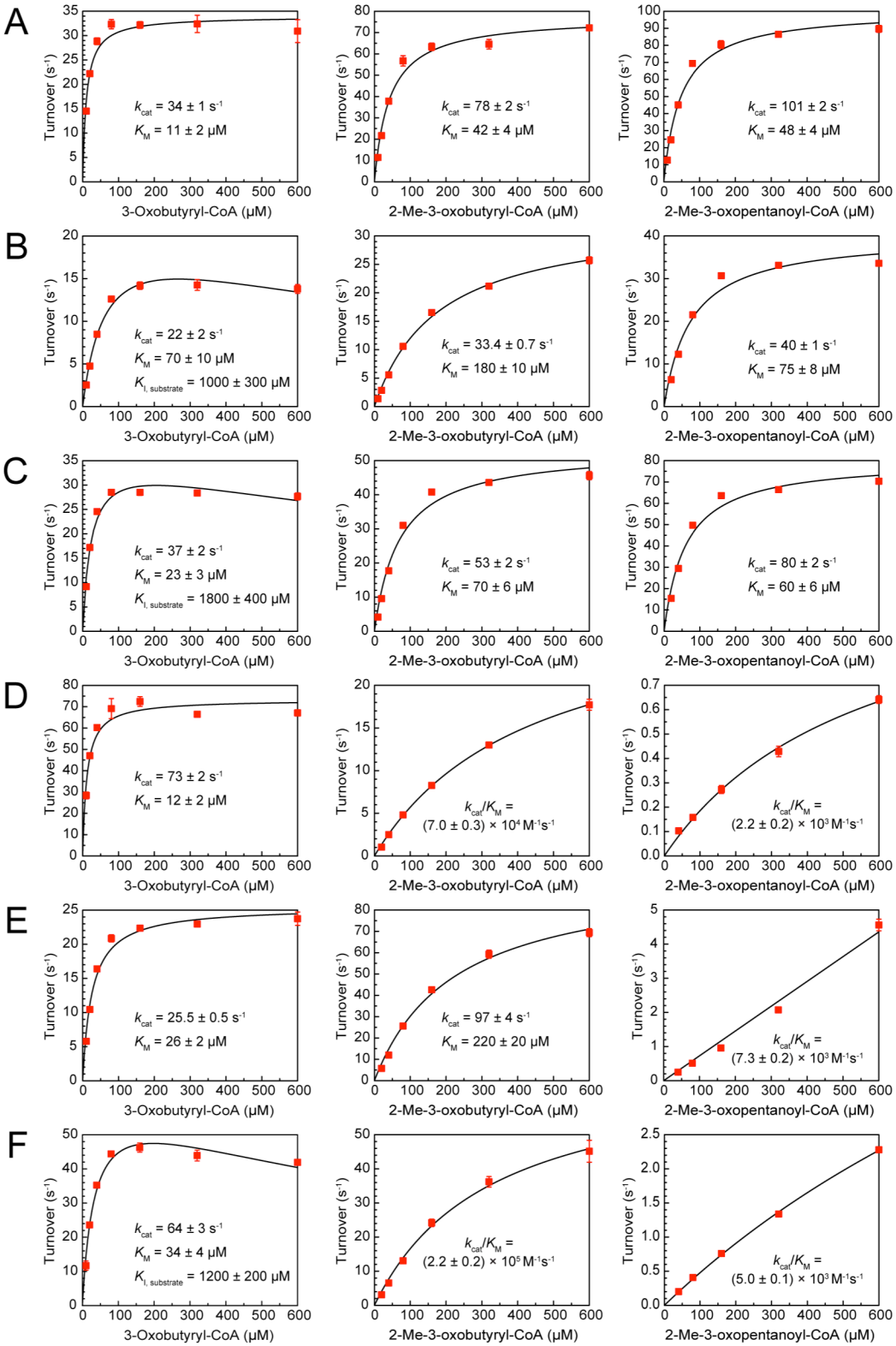


Figure 3.12. *Initial rate data for wild type and mutant Acat2 and Acat5 kinetic characterization. Error bars represent mean ± standard error. (A) Acat2. (B) Acat2-Loop2. (C) Acat2-Y153I. (D) Acat5. (E) Acat5-Loop2. (F) Acat5-I154Y.*

3.3 Discussion

We have presented structures of two thiolase enzymes from *A. suum*, which demonstrate differing substrate selectivity despite sharing characteristic features of the T2 thiolase group. Comparison of holo and apo structures to each other and to human T2 show that the substrate binding modes are highly similar in these enzymes, including conformational changes in the side chains of Lys227/229, Cys91/92, and Met158/159 that occur upon binding. Our propionyl-CoA bound structure contained the substrate situated in a position that represents the pre-Claisen intermediate of the reaction pathway, and showed that the branching methyl group is directed between Phe288/292 and Met158/159, near the active site covering loop. This positioning would result in *S* stereochemistry for the α-methyl substituent of the product. Although Haapalainen *et al.* point out that this methyl group is easily racemized due to the low pK_a of the α-proton in a 3-oxo thioester species [20], a 2*S* configuration would match the expected stereochemstry for the 2-methyl-3-oxoacyl-CoA product formed by ketoreductase AsHadh.

An interesting weak OH-π hydrogen bond between Thr293 and Phe292 facilitated a rotation of this ring in the Acat5 structure, decreasing the space in the binding pocket. Nevertheless, we found that exchange of the Acat5 covering loop with that of Acat2 was sufficient to increase the relative permissiveness of Acat5 with 2-methyl-3-oxobutyryl-CoA by 40-fold, with only a modest decrease in the enzyme's efficiency for the natural substrate. Mutagenesis studies showed that the sequence of this loop, and in particular Tyr153/Ile154, plays an important role in governing substrate selectivity in thiolases, although clearly other factors are involved since the selectivity of Acat2 was not modified by these changes. Tyr153 is mostly absent in thiolases that do not permit a branched substrate, but *Cupriavidus necator* PhaA presents an exception and does contain Tyr at that position (*Figure 3.9*). Differences in the beginning of the Cα1 helix are also likely important for selectivity, including the replacement of Phe288/292 with Met in CT-like thiolases and replacement of the hydrophobic Ile286 conserved in T2, which comes in close contact with the covering loop and Lα2 in those isozymes.

Acat5-Loop2 presents a fascinating example of a thiolase that is permissive of 2-methyl-3-oxobutyryl-CoA but is fairly selective against the slightly longer substrate. In Acat5-Loop2, extending the main chain by one carbon (2-methyl-3-oxopentanoyl-CoA) confers a catalytic defect of 60-fold in comparison to its shorter branched counterpart. This is even higher than the 32-fold defect observed in wild type Acat5 for this chain extension. We have previously showed that this double kinetic defect for 2-methyl-3-oxopentanoyl-CoA is a synergistic effect in Acat5, because the efficiency with linear 3-oxopentanoyl-CoA is only 4-fold lower than with 3-oxobutyryl-CoA. Since the hydrophobic tunnel for binding longer chain acyl-CoA substrates is obstructed in CT- and T2-type thiolases (*Figure 3.8B*), the terminal methyl of a propionylated enzyme intermediate might be rotated more towards the top of the active site near Met159 and Phe292 rather than down towards the tunnel. This would provide even less space in the binding pocket for a subsequent propionyl-CoA nucleophile.

The factors that affect subtle differences in substrate selectivity in thiolases are complex, but these results offer insight that could help guide future thiolase discovery or engineering projects. The mode of branched substrate accommodation in thiolases presents challenges, however, for expanding thiolase promiscuity. Fage *et al.* demonstrated that mutating Met290 to Ala in BktB conferred the ability to act on an α-methyl substrate, but the efficiency of the enzyme was significantly diminished [14], which may be related to the importance of sulfur-sulfur interactions in facilitating the binding of CoA [40]. And while hydrogen bonding networks and

other interactions that help firmly anchor secondary structures in the vicinity of the active site may contribute to high catalytic efficiencies, the rigidity may also decrease space in the binding cavity. Nevertheless, our studies of metabolic pathway engineering using *A. suum* enzymes showed that even wild type Acat5 was sufficient to produce moderate titers of branched enoic acids in *E. coli* (*Chapter 2*). Thus, in conjunction with the appropriate downstream enzymes, only a moderate degree of catalytic competency with an unnatural substrate may be required in thiolases to sustain high flux synthetic pathways.

3.5 References

1. A. M. Haapalainen, G. Meriläinen, and R. K. Wierenga, The thiolase superfamily: condensing enzymes with diverse reaction specificities. *Trends Biochem. Sci* **2006,** 31, (1), 64-71.
2. L. Hiser, M. E. Basson, and J. Rine, ERG10 from Saccharomyces cerevisiae encodes acetoacetyl-CoA thiolase. *J. Biol. Chem.* **1994,** 269, (50), 31383-31389.
3. H. M. Miziorko, Enzymes of the mevalonate pathway of isoprenoid biosynthesis. *Arch. Biochem. Biophys.* **2011,** 505, (2), 131-143.
4. S. Masamune, C. T. Walsh, A. J. Sinskey, and O. P. Peoples, Poly-(R)-3-hydroxybutyrate (PHB) biosynthesis: mechanistic studies on the biological Claisen condensation catalyzed by β-ketoacyl thiolase. *Pure Appl. Chem.* **1989,** 61, (3).
5. B. B. Bond-Watts, R. J. Bellerose, and M. C. Y. Chang, Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat. Chem. Biol.* **2011,** 7, (4), 222-227.
6. C. Dellomonaco, J. M. Clomburg, E. N. Miller, and R. Gonzalez, Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **2011,** 476, (7360), 355-359.
7. C. R. Shen, E. I. Lan, Y. Dekishima, A. Baez, K. M. Cho, and J. C. Liao, Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.* **2011,** 77, (9), 2905-15.
8. M. J. Sheppard, A. M. Kunjapur, S. J. Wenck, and K. L. J. Prather, Retro-biosynthetic screening of a modular pathway design achieves selective route for microbial synthesis of 4-methyl-pentanol. *Nat Commun* **2014,** 5.
9. Q. Zhuang, Q. Wang, Q. Liang, and Q. Qi, Synthesis of polyhydroxyalkanoates from glucose that contain medium-chain-length monomers via the reversed fatty acid beta-oxidation cycle in *Escherichia coli*. *Metab. Eng.* **2014,** 24, 78-86.
10. C. H. Martin, H. Dhamankar, H. C. Tseng, M. J. Sheppard, C. R. Reisch, and K. L. Prather, A platform pathway for production of 3-hydroxyacids provides a biosynthetic route to 3-hydroxy-gamma-butyrolactone. *Nat Commun* **2013,** 4, 1414.
11. U. Gehring, C. Riepertinger, and F. Lynen, [Purification and crystallization of thiolase; study of its action mechanism.] *Eur. J. Biochem.* **1968,** 6, (2), 264-280.
12. J. I. Harris and U. Gehring, The active site cysteines of thiolase. *Eur. J. Biochem.* **1970,** 16, (3), 492-498.

13. J. I. Harris and U. Gehring, The subunit structure of thiolase. *Eur. J. Biochem.* **1970,** 16, (3), 487-491.

14. C. D. Fage, J. L. Meinke, and A. T. Keatinge-Clay, Coenzyme A-free activity, crystal structure, and rational engineering of a promiscuous β-ketoacyl thiolase from *Ralstonia eutropha*. *J. Mol. Catal. B: Enzym.* **2015,** 121, 113-121.

15. E.-J. Kim and K.-J. Kim, Crystal structure and biochemical characterization of PhaA from *Ralstonia eutropha*, a polyhydroxyalkanoate-producing bacterium. *Biochem. Biophys. Res. Commun.* **2014,** 452, (1), 124-129.

16. S. Masamune, M. A. J. Palmer, R. Gamboni, S. Thompson, J. T. Davis, S. F. Williams, O. P. Peoples, A. J. Sinskey, and C. T. Walsh, Bio-claisen condensation catalyzed by thiolase from *Zoogloea ramigera*. Active site cysteine residues. *JACS* **1989,** 111, (5), 1879-1881.

17. G. Meriläinen, V. Poikela, P. Kursula, and R. K. Wierenga, The thiolase reaction mechanism: The importance of Asn316 and His348 for stabilizing the enolate intermediate of the claisen condensation. *Biochemistry* **2009,** 48, (46), 11011-11025.

18. Y. Modis and R. K. Wierenga, A biosynthetic thiolase in complex with a reaction intermediate: the crystal structure provides new insights into the catalytic mechanism. *Structure* **1999,** 7, (10), 1279-1290.

19. S. Thompson, F. Mayerl, O. P. Peoples, S. Masamune, A. J. Sinskey, and C. T. Walsh, Mechanistic studies on beta-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, (14), 5735-5742.

20. A. M. Haapalainen, G. Meriläinen, P. L. Pirilä, N. Kondo, T. Fukao, and R. K. Wierenga, Crystallographic and kinetic studies of human mitochondrial acetoacetyl-CoA thiolase: The importance of potassium and chloride ions for its structure and function. *Biochemistry* **2007,** 46, (14), 4305-4321.

21. T.-R. Kiema, R. K. Harijan, M. Strozyk, T. Fukao, S. E. H. Alexson, and R. K. Wierenga, The crystal structure of human mitochondrial 3-ketoacyl-CoA thiolase (T1): insight into the reaction mechanism of its thiolase and thioesterase activities. *Acta Crystallographica Section D Biological Crystallography* **2014,** 70, (12), 3212-3225.

22. P. Kursula, H. Sikkilä, T. Fukao, N. Kondo, and R. K. Wierenga, High resolution crystal structures of human cytosolic thiolase (CT): A comparison of the active sites of human CT, bacterial thiolase, and bacterial KAS I. *J. Mol. Biol.* **2005,** 347, (1), 189-201.

23. B. Middleton, The oxoacyl-coenzyme A thiolases of animal tissues. *Biochem. J* **1973,** 132, (4), 717-730.

24. P. Kursula, J. Ojala, A.-M. Lambeir, and R. K. Wierenga, The catalytic cycle of biosynthetic thiolase: A conformational journey of an acetyl group through four binding modes and two oxyanion holes. *Biochemistry* **2002,** 41, (52), 15543-15556.

25. Y. Modis and R. K. Wierenga, Crystallographic analysis of the reaction pathway of *Zoogloea ramigera* biosynthetic thiolase. *J. Mol. Biol.* **2000,** 297, (5), 1171-1182.

26. M. A. Palmer, E. Differding, R. Gamboni, S. F. Williams, O. P. Peoples, C. T. Walsh, A. J. Sinskey, and S. Masamune, Biosynthetic thiolase from *Zoogloea ramigera*. Evidence for a mechanism involving Cys-378 as the active site base. *J. Biol. Chem.* **1991,** 266, (13), 8369-8375.

27. S. F. Williams, M. A. Palmer, O. P. Peoples, C. T. Walsh, A. J. Sinskey, and S. Masamune, Biosynthetic thiolase from *Zoogloea ramigera*. Mutagenesis of the putative active-site base Cys-378 to Ser-378 changes the partitioning of the acetyl S-enzyme intermediate. *J. Biol. Chem.* **1992,** 267, (23), 16041-16043.

28. J. Peretó, P. López-García, and D. Moreira, Phylogenetic analysis of eukaryotic thiolases suggests multiple proteobacterial origins. *J. Mol. Evol.* **2005,** 61, (1), 65-74.

29. 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, (5), 343-5.

30. E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Ed. J. M. Walker, Humana Press: 2005; 571-607

31. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010,** 66, (Pt 2), 125-32.

32. P. R. Evans and G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **2013,** 69, (Pt 7), 1204-14.

33. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read, Phaser crystallographic software. *J. Appl. Crystallogr.* **2007,** 40, (Pt 4), 658-674.

34. G. Bunkoczi and R. J. Read, Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr. D Biol. Crystallogr.* **2011,** 67, (Pt 4), 303-12.

35. P. D. Adams, R. W. Grosse-Kunstleve, 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 Biol. Crystallogr.* **2002,** 58, (Pt 11), 1948-54.

36. P. Emsley and K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004,** 60, (Pt 12 Pt 1), 2126-32.

37. M. Montes, M. A. Miteva, and B. O. Villoutreix, Structure-based virtual ligand screening with LigandFit: pose prediction and enrichment of compound collections. *Proteins* **2007,** 68, (3), 712-25.

38. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004,** 25, (13), 1605-12.

39. M. Mathieu, Y. Modis, J. P. Zeelen, C. K. Engel, R. A. Abagyan, A. Ahlberg, B. Rasmussen, V. S. Lamzin, W. H. Kunau, and R. K. Wierenga, The 1.8 Å crystal structure of the dimeric peroxisomal 3-ketoacyl-CoA thiolase of *Saccharomyces cerevisiae*: implications for substrate binding and reaction mechanism. *J. Mol. Biol.* **1997,** 273, (3), 714-728.

40. G. Meriläinen, W. Schmitz, R. K. Wierenga, and P. Kursula, The sulfur atoms of the substrate CoA and the catalytic cysteine are required for a productive mode of substrate binding in bacterial biosynthetic thiolase, a thioester-dependent enzyme. *FEBS J.* **2008,** 275, (24), 6136-6148.

Chapter 4: *Wax fermentation in Euglena gracilis*

4.1. Introduction

Living systems have the potential to address challenges in sustainable chemical production. Industrial bioprocesses are already widely used for the manufacturer of high value molecules such as pharmaceuticals, bioactive peptides, and antibodies, but the production of low-cost commodity chemicals poses especially stringent requirements for product titers and yields in order to be economically feasible. These requirements are not straightforward to achieve, because cellular metabolism is controlled by layers of complex regulation and is directed towards maximizing the growth and survival of the organism, which may or may not support the concerted accumulation of a single chemical product. The science and strategy of metabolic engineering therefore requires consideration of when and how cellular fitness may be complemented, or even enhanced, by accumulation of a target molecule.

The microbial production of ethanol and of lactic acid provide two examples of metabolic pathways which permit high yield conversion and accumulation of a single product. Ethanol production has theoretical yield of 66% from glucose based on the stoichiometry of conversion. This means that up to 66% of the atoms of carbon contained in the glucose feed can be retained and converted into ethanol. The remaining 34% of carbon must necessarily be oxidized to carbon dioxide to provide a net reaction that is balanced for oxidation state. (*Figure 4.1A*) Lactic acid has an equivalent oxidation state to glucose and therefore has a theoretical yield of 100%. Critically, both ethanol and lactate can and have been produced at high titer and with yields close to their theoretical limits [1, 2]. The success of these processes is rooted in the biochemical "logic" of natural systems. Oxidation of organic matter provides cellular energy. In the absence of oxygen or of other inorganic oxidizing agents, this same organic matter must serve as terminal electron acceptors for oxidative processes (*Figure 4.1B*). Accumulation of reduced organic matter is thus concomitant with energy production under anaerobic conditions.

This paradigm can be extended to the biosynthesis of other small molecules, such as isobutanol [3], especially molecules which can be biosynthesized from pyruvic acid. But generalizing this approach to larger hydrocarbon architectures, such as fatty alcohol fuels and related products, is not trivial, because the assembly of larger molecules primarily relies on pathways that draw from acetyl-CoA. As reviewed in *Chapter 1*, acetyl-CoA is a critical nexus of metabolic regulation and energy homeostasis, and pathways based on its condensation face thermodynamic and flux control challenges. The production of butanol provides one example of successful high yield production of a higher-order biofuel from acetyl-CoA [4-6], but sustaining multiple iterations of condensation to produce lipid-like chemicals has proven difficult. Engineering fermentation pathways based on acetyl-CoA remains an ongoing and fundamental conceptual challenge in metabolic engineering.

Euglena gracilis is a unicellular, eukaryotic green algae, which stands out among well-characterized organisms for an unprecedented metabolic competency: the anaerobic fermentation of fatty acid products, namely wax monoesters [7]. It has been known for several decades that this omnitrophic organism is able to rapidly convert its polysaccharide storages into wax esters upon introduction to an anaerobic environment, with a net gain in ATP [8]. This process has been shown to proceed by a thiolase-based pathway that resembles the reverse of β-oxidation (*Figure 4.1B*) [9]. The dependence of the pathway on thiolase enzymes was further confirmed by a recent study, in which wax production was attenuated by RNAi knockdowns of particular thiolase genes [10]. The origin of acetyl-CoA in this process is an unusual ferridoxin-dependent pyruvate:$NADP^+$ oxidoreductase, which is oxygen-senstitive and only active under anaerobic

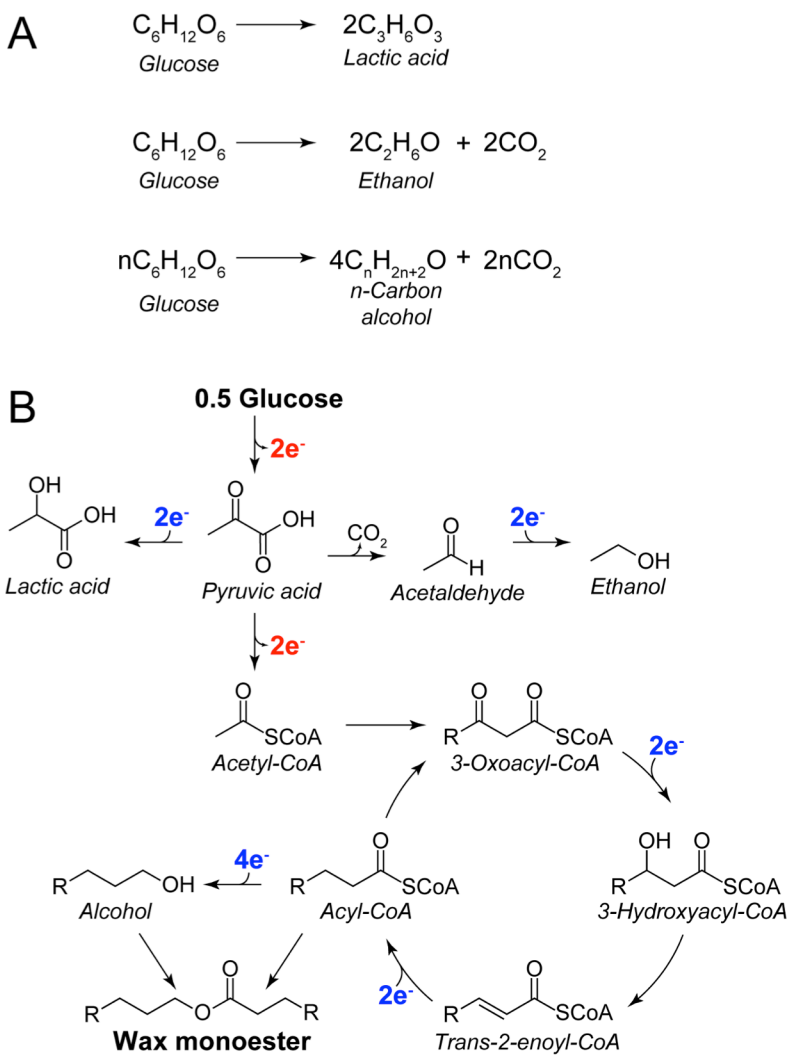


Figure 4.1. *Redox-balanced fermentation strategies in nature. (A) Balanced reactions showing net conversion of glucose into fermentation products pyruvate, ethanol, or other n-carbon chain alcohols. 100% conversion of the carbon of glucose can theoretically be converted to lactic acid. Alcohols are capable of being produced with a 66% stoichiometric conversion of glucose. (B) Pathways of fermentation from glucose. Glucose oxidation to pyruvate produces reducing equivalents, which must be used to reduce organic material in the absence of oxygen. Pyruvate can serve as electron acceptor to form lactate and ethanol. In E. gracilis, pyruvate is further oxidized to acetyl-CoA, the substrate for production of wax monoesters through a reductive thiolase-based pathway.*

conditions [11]. The reduction of an enoyl-CoA to form a saturated acyl-CoA, completing the four-step elongation cycle, was found to be catalyzed by one of three different enoyl reductase enzymes, each exhibiting a preference for a particular chain length substrate – long, medium, or short [9]. The short-chain enoyl reductase was isolated and characterized by Inui *et al*. and found to have unusual kinetic characteristics and a biological assembly that has not been reported in any known class of enoyl reductase [12]. The enzyme consisted of a heterodimer of two small subunits, 15kDa and 25kDa, both of which were required for activity. The enzyme contained a flavin cofactor but used NADH as the reducing agent and exhibited a strong preference for crotonyl-CoA as substrate. An unusual biphasic kinetic profile was observed as a function of crotonyl-CoA concentration, with reaction rate reaching an initial saturation, followed by an additional regime of rate increase before final substrate saturation. The behavior was described as representing negative cooperativity with regards to substrate binding. Similar behavior has been reported for the enzyme CTP:glycerol-3-phosphate cytidylyltransferase [13], but this enzyme bears no similarity to an enoyl reductase.

In 2005, Hoffmeister *et al*. reported the isolation of a *trans*-2-enoyl-CoA reductase EgTer from a large-scare purification from *E. gracilis* whole cell lysate [14]. Determination of the gene sequence coding for this protein revealed that the precursor protein contains an extremely long leader sequence of approximately 100 residues, which is cleaved to afford the active enzyme. Although the authors identified EgTer with the enoyl reductase isolated by Inui *et al*. [12] due to its short-chain selectivity and preference for NADH cofactor, the enzyme bore little resemblance based on its molecular weight, biological assembly, and lack of flavin cofactor. A homolog of EgTer, TdTer, was identified in the bacterial species *Treponema denticola*, which was found to have similar enzyme function with improved catalytic activity [15]. TdTer has been critical to the development of high flux synthetic pathways for the production of butanol biofuel [4].

The utility of TdTer for metabolic engineering applications motivates further investigation into the enzymes involves in wax ester fermentation in *E. gracilis*. And the potential of this organism as a genetic resource is not limited to just wax metabolism, as at least three other distinct fatty acid synthesis systems have been reported in *Euglena* [16-18], as well as a host of other interesting pathways and enzymes [19], making it a fascinating study of non-canonical carbon metabolism. A comprehensive genetic database of *Euglena* is not readily available, however, and transcriptome shotgun assemblies have only very recently been reported in the literature [19, 20]. We therefore set out to assemble a transcriptome database for *E. gracilis* and, in conjunction with biochemical and physiological studies, gain insight into the biochemistry of wax fermentation.

4.2. Materials and methods

Commercial materials. Luria-Bertani (LB) Broth Miller, LB Agar Miller, and 2xYT broth, glacial acetic acid, boric acid, and glycerol were purchased from EMD Biosciences (Darmstadt, Germany). D-(+)-glucose and DNase-free, RNase-free water were purchased from MP Biochemicals (Santa Ana, CA). Imidazole was purchased from Acros Organics (Morris Plains, NJ). Isopropyl-β-D-thiogalactopyranoside (IPTG) and *trans*-2-dodecenoic acid were purchased from Santa Cruz Biotechnology (Dallas, TX). Carbenicillin sodium salt (Cb), kanamycin sulfate (Km), potassium phosphate monobasic, potassium phosphate dibasic, ammonium sulfate, magnesium sulfate heptahydrate, calcium carbonate, ammonium iron sulfate hexahydrate, cupric sulfate heptahydrate, EDTA, Tris base, and aerosol resistant pipet tips were purchased from

Fisher Scientific (Pittsburgh, PA). Coenzyme A sodium salt hydrate (CoA), crotonyl-CoA, manganese sulfate monohydrate, sodium molybdate dehydrate, cobalt(II) sulfate pentahydrate, thiamine, cyanocobalamin were purchased from Sigma-Aldrich (St. Louis, MO). Acrylamide/Bis-acrylamide (30%, 37.5:1), electrophoresis grade sodium dodecyl sulfate (SDS), Bio-Rad protein assay dye reagent concentrate, ammonium persulfate, N,N,N',N'-tetramethyl-ethane-1,2-diamine (TEMED), and goat anti-rabbit IgG (H + L)-HRP conjugate were purchased from Bio-Rad Laboratories (Hercules, CA). Restriction enzymes, Q5 DNA Polymerase, T4 DNA ligase (2×10^6 U/mL), Phusion DNA polymerase, T5 exonuclease, Taq DNA ligase, Buffer #2, T4 PNK, Klenow Large Fragment, Klenow Fragment (3'→5' exo-), and T4 DNA Polymerase were purchased from New England Biolabs (Ipswich, MA). PageRuler™ Plus prestained protein ladder was purchased from Fermentas (Glen Burnie, Maryland). Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). Anaerobic Hungate culture tubes were purchased from Chemglass (Vineland, NJ). DNA purification kits, RNeasy RNA purification kit, QIAshredder homogenization columns, and Ni-NTA agarose were purchased from Qiagen (Valencia, CA). cOmplete EDTA-free protease inhibitor were purchased from Roche Applied Science (Penzberg, Germany). TURBO DNase, Ambion® Poly(A)Purist™ kit, Ambion® Fragmentation Reagent, 200 ng/µL random hexamers, SuperScriptII® Reverse Transcriptase, 10 mM dNTP Mix, RNase H, and *E. coli* DNA Polymerase I were purchased from Thermo Fisher Scientific (Waltham, MA). Primary rabbit serum antibody was purchased from ProSci Inc. (Poway, CA). Western Lightning® Plus-ECL reagent was purchased from PerkinElmer (Waltham, MA).

Bacterial strains. *E. coli* DH10B-T1^R was used for DNA construction and BL21(de3)-T1^R was used for heterologous production of proteins for purification.

***E. gracilis* strains and cell culture.** *Euglena gracilis* strain Z (ATCC 12716) and the etiolated mutant (ATCC 30285) were obtained from the American Type Culture Collection. Cultures were grown in 10 mL of the media of Hoffeister *et al*. [21] and maintained by diluting 1:50 into fresh media approximately every two weeks. Green *E. gracilis* was cultured adjacent to an east-facing window. The etiolated strain was cultured in the dark.

Preparation of cDNA libraries for Illumina® RNA sequencing. For RNA isolation and cDNA library synthesis, a 10 mL culture of green *Euglena gracilis* Z was grown to $OD_{750} = 0.3$, then transferred to an anaerobic Hungate culture tube, sealed, and sparged with argon gas filtered through a 0.2 µm filter for 15 min ($-O_2$ sample). The tube was kept in the dark for 60 hours prior to harvesting. An additional 10 mL culture ($+O_2$) was maintained under ambient oxygenation and lighting until harvesting. Samples $+O_2$ and $-O_2$ were henceforth treated identically for cDNA library preparation. RNase free water, reagents, pipet tips, and tubes were used until the completion of first strand cDNA synthesis. RNA samples during library synthesis, as well as the final cDNA libraries, were analyzed on an Experion™ Automated Electrophoresis System (Bio-Rad Laboratories)

Total RNA isolation. 10 mL *Euglena* culture was harvested by centrifugation at 2200g for 10 min. The supernatant was discarded, and the pellets were quickly resuspended in 450 µl Buffer RLT from an RNeasy RNA isolation kit (Qiagen) supplemented with 1% v/v BME. The cell suspension was transferred to a 1.7 mL tube, vortexed at 1500 rpm for 1-2 min, and applied to a QIAshredder column and centrifuged for 2 min at 18000g. The supernatant was decanted into a fresh 1.7 mL tube and ½ volume of ethanol was added. The mixture was applied to an RNeasy column (Qiagen) and washed and eluted according to the manufacturer instructions.

10 µg total RNA was treated with 1 µL TURBO DNase in a 50 µl reaction including 5 µL 10X TURBO DNase buffer. The reaction was incubated at 37º for 30 min. The reaction was diluted with 50 µL H_2O, 350 µL Buffer RLT, and 250 µL ethanol and transferred to an RNeasy column (Qiagen) for RNA cleanup following the manufacturer instructions.

mRNA isolation. mRNA was isolated from total RNA using the Ambion® Poly(A)Purist™ kit, following the manufacturer instructions. Following elution with the supplied elution buffer, a second round of purification was carried out using the same kit. The final mRNA was eluted in 200 µL THE RNA Storage Solution. 20 µL ammonium acetate and 1 µL glycogen solution (both supplied by the manufacturer) as well as 550 µL cold ethanol were added, and the sample was flash frozen in liquid nitrogen, thawed to just melting, and centrifuged at 21000g at 4ºC for 20 min. The supernatant was carefully removed by pipet, and the pellet was washed with 500 µL cold 70% v/v ethanol. The precipitate was centrifuged again for 10 min and the supernatant carefully removed before drying the pellet for several minutes on a 42º heating block. 10 µL water was added and the mRNA pellet was dissolved by vortexing.

mRNA fragmentation. To 9 µL mRNA was added 1 µL of Ambion® Fragmentation Reagent in a 200 µL tube. The RNA was heated on a thermocycler block preheated to 70ºC for exactly 2 min before rapid addition of 1µL Stop Solution and cooling on ice. 9 µL water, 2 µL ammonium acetate, 1 µL glycogen, and 55 µL cold ethanol were added, and the fragmented mRNA was precipitated using the procedure described above, resuspending in 12 µL water.

First strand cDNA synthesis. First strand cDNA was synthesized using SuperScriptII® Reverse Transcriptase. To 11 µL fragmented RNA was added 1 µL of 200 ng/µL random hexamers. The sample was incubation at 65ºC for 5 min and immediately placed on ice. 4 µL 5X First Strand Buffer, 2 µL DTT, and 1 µL 10 mM dNTP mix were added, and the reaction was incubated at 25ºC for 2 min before addition of 1 µL SuperScriptII®. The reaction was incubated at 25ºC for 10 min, 42ºC for 50 min, and 70ºC for 15 min.

Second strand cDNA synthesis. To the previous reaction was added directy 61 µL water, 3µL 10 mM dNTP mix, 10 µL Buffer #2, 1 µL RNase H, and 5 µL *E. coli* DNA Polymerase I. The reaction was kept on ice during addition of components, and then incubated at 16ºC for 2.5 hours. 500 µL Buffer PB was added and DNA was purified with a QIAquick PCR Purification kit following the manufacturer instructions.

End Repair. 10 µL 10X T4 Ligase buffer 4 µL 10 mM dNTPs, 5 µL T4 PNK, 1 µL Klenow Large Fragment, 2µL T4 DNA Polymerase, and water were added to double stranded cDNA to a total volume of 100 µL. The reaction was incubated at 20ºC for 30 min, and DNA was purified with a QIAquick PCR Purification kit.

A-tailing. 1 µL 10 mM dATP, 5 µL Buffer #2, 3 µL Klenow Fragment (3'→5' exo-), and water were added to blunt end cDNA to a final volume of 50 µL. The reaction was incubated at 37ºC for 30 min, and DNA was purified with a MinElute PCR Purification kit.

Preparation of Illumina® adapter mix. PE Adapter 1 and PE Adapter 2 were resuspended in Buffer EB (Qiagen) and combined in a single tube to give a concentration of 10 µM for each adapter. The adapter mix was heated at 95ºC for 3 min and slowly cooled to room temperature.

Adapter ligation. To 25 µL A-primed cDNAwas added 3 µL 10X T4 Ligase buffer, 1 µL 10 µM Illumina® adapter mix, and 1 µL T4 DNA Ligase. The reaction was incubated at room temperature for 30 min.

Agarose gel size selection. Adapter-ligated cDNA library was loaded into a 2% w/v agarose electrophoresis gel in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) run at 120 V for 50 min. 2 mm thick bands were cut out along the length of the gel and stored in 1.7 ml plastic tubes at -20ºC labeled A+ through X+ and A- through X-, corresponding to the $+O_2$ and $-O_2$ samples, respectively. Adapter ligated cDNA diffused much more slowly through the agarose gel compared to an O'GeneRuler 100 bp Plus DNA Ladder. For example, a band corresponding to 800 bp in the ladder actually had an average fragment size of 400 bp as measured by Experion™ automated electrophoresis.

The cDNA libraries in the T+ and T- gel samples were extracted using a QIAquick Gel Extraction kit, following the manufacturer instructions, except the gel band was dissolved in Buffer QG at room temperature rather than 55ºC. DNA was eluted with 50 µL Buffer EB.

Library PCR amplification. To 36.5 µL of cDNA library was added 1 µL each of PE PCR Primer 1.0 and PE PCR Primer 2.0, 1 µL 10 mM dNTPs, 10 µL 5X Phusion HF Buffer, and 0.5 µL Phusion DNA Polymerase. 10 rounds of PCR were carried out in a thermocycler using the following protocol:

1) 2 min at 98ºC
2) 20 sec at 98ºC
3) 30 sec at 65ºC
4) 30 sec at 72ºC
5) Return to 2, 9X
6) 5 min at 72ºC

The amplified libraries were each purified using a MinElute Purification kit and eluted with 15 µL Buffer EB to give libraries T+ and T-.

RNAseq and transcriptome assembly. T+ and T- libraries were quantified and sequenced at the University of California, Davis Genome Center in separate lanes on an Illumina® MiSeq system to give paired end reads with 250 cycles each, with 13-15 million paired end reads for each library. The reads were trimmed for adapter contamination and deteriorating quality score. Reads from the T+ and T- libraries were combined for transcriptome assembly using the Trinity software package [22].

Construction of protein expression plasmids. Construction of *E. coli* plasmids for heterologous expression of *E. gracilis* genes was carried out using the Gibson protocol [23], by PCR amplifying the gene from *E. gracilis* cDNA using the primers listed in *Appendix, Table S3* and inserting into the appropriate vector.

Expression of His-tagged *E. gracilis* enzymes. EgTer and EgMECR1 through 4 were expressed and purified following the procedure described in *Chapter 3* for Acat2 and Acat5 mutants, except that EgMECR2 was expressed from a pSV272.1 vector and required kanamycin (Km) in place of carbenicillin (Cb) for cell culture.

Western blot. Western blot was carried out using primary rabbit serum antibody purchased from ProSci Inc. For EgTer detection, antibodies were raised against purified EgTer protein as antigen. For SSADH detection, the antigen was synthetic peptide NH_2-APVYDQFVDQFAQRMGA-OH, which was identified within transcriptome contig 43096. Goat anti-rabbit IgG (H + L)-HRP conjugate was used for secondary antibody, and the blot was imaged by chemiluminescence using Western Lightning® Plus-ECL reagent.

Enzyme assays. *Trans*-2-enoyl-CoA reductase assays and synthesis of 2-dodecenoyl-CoA substrate were performed using the methods of Bond-Watts, *et al*. [24].

4.3. Results

Culturing and transcriptome sequencing of *E.* gracilis. We cultured *E. gracilis* using the defined media of Hoffmeister *et al*. [21]. We were unable to preserve viable cells cryogenically for long-term storage, so aliquots of the algae were sub-cultured approximately every two weeks. A long lag phase was observed if *E. gracilis* were inoculated using a culture at stationary phase for a long time, but aliquots sub-cultured within two weeks resumed logarithmic growth after a short time. Green *E. gracilis* typically reached a maximum OD_{750} of approximately 6.0 in this media.

E. gracilis converts its carbohydrate deposits into wax esters when introduced to anaerobic conditions. We were therefore interested in obtaining RNA for transcriptome sequencing from both aerobically and anaerobically grown cultures, in order to compare gene expression levels in each condition, as well as to ensure that a transcriptome assembly would be as comprehensive as possible and contain all transcripts relevant to wax fermentation. We initially tried culturing *E. gracilis* anaerobically starting from inoculation using media that was sparged with argon. We did not observe proliferation of cell culture under these conditions, but aliquots of the culture were viable and grew at normal rates once restored to aerobic conditions. This result is similar to that observed by Carre *et al*., in which *E. gracilis* were found to halt growth upon introduction to anoxia but remained viable for several weeks if oxygen was restored [25]. It is contrary to the result of Tucci *et al*., however, who were capable of culturing green *E. gracilis* in the dark in anaerobic conditions [26]. Because we were unable to culture *E. gracilis* anaerobically, we grew culture to a cell density sufficient for RNA isolation and then sparged the culture with sterile argon 60 hours prior to harvesting.

High quality total RNA was obtained from $+O_2$ and $-O_2$ *E. gracilis* samples that matched the unusual ribosomal RNA pattern reported by Schnare and Gray [27] (*Figure 4.2*). *E. gracilis* lacks a 28S rRNA component of the ribosome, and instead contains a series of smaller rRNAs that assemble in *trans* to carry out the equivalent function. After mRNA isolation and chemical fragmentation of RNA, we synthesized adapter ligated cDNA libraries for paired-end massively parallel shotgun sequencing. The size distribution of the libraries had averages of approximately 600 bp. This number includes both the insert size and approximately 60bp on each end from the PCR amplification primers. We obtained 13-15 million paired-end sequence reads from each library, and combined the two sets of reads for transcriptome assembly using the Trinity software package [22]. The resulting assembly contained 161396 contigs representing isoforms of 85805 genes. We used the RSEM utility [28] to quantify read coverage of transcripts and found that a large percentage of the contigs had very low coverage. Local alignment-based searching with BLAST showed that many of the exceptionally low coverage contigs matched genomic DNA of *Rhodococcus erythropolis* or of common expression plasmids. We therefore curated the transcriptome by filtering isoforms that represented less than 1 TPM or accounted for less than 1% of transcript variants for the given gene, reducing the number of contigs in the transcriptome to 56788.

We quantified the expression levels of transcriptome entries in aerobic and anaerobic samples and searched for contigs whose expression level differed significantly between these

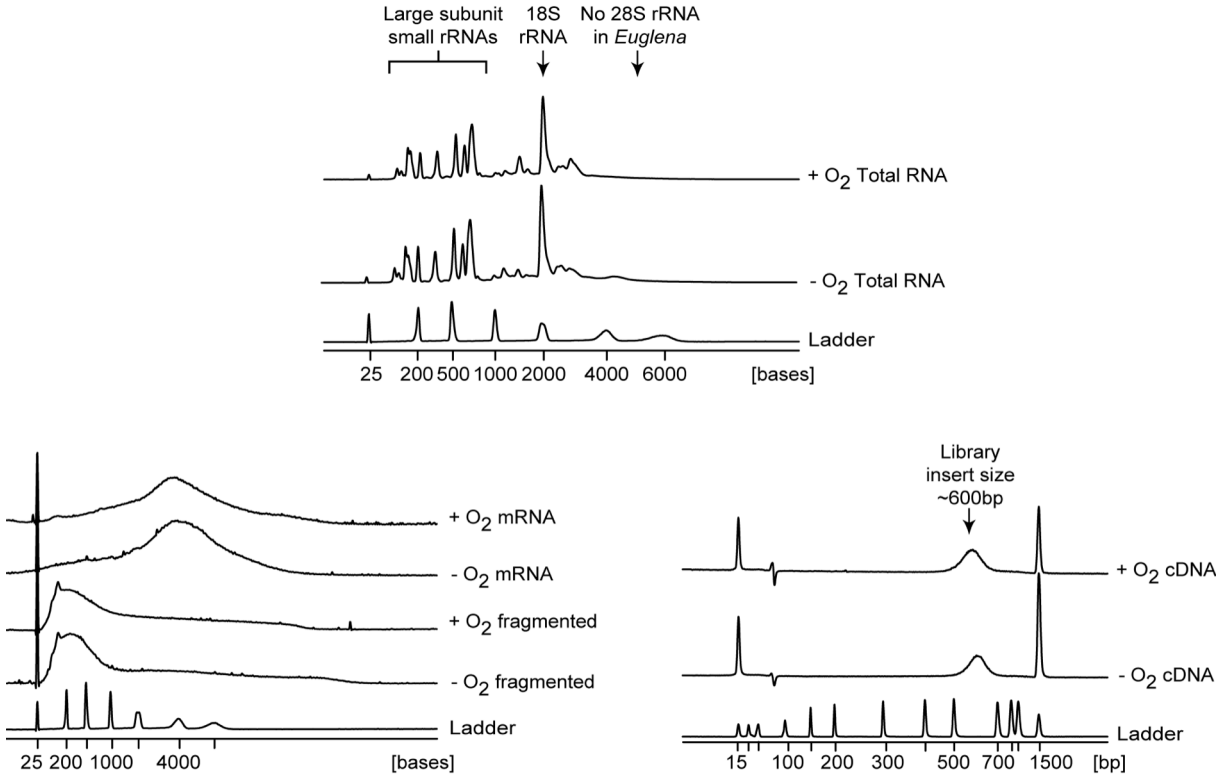


Figure 4.2. *Capillary electrophoresis traces of E. gracilis total RNA, mRNA, and cDNA libraries.*

Table 4.1. *Predicted function of proteins with greater than 2-fold increase in expression between $+O_2$ and $-O_2$ samples.*

Predicted protein based on closest BLAST hit[a]	E value	TPM, $+O_2$	TPM, $-O_2$	Fold increase
no similarity	(> 1E-5)	25	107	4.2
RNA polymerase beta chain [Euglena gracilis]	0.0	452	1633	3.6
ribosomal protein S29 [Chlamydomonas reinhardtii]	9E-16	87	236	2.7
pathogenesis-related protein 1C precursor [Geminocystis sp. NIES-3709]	4E-12	97	260	2.7
photosystem I assembly protein Ycf4 [Euglena gracilis]	1E-123	155	410	2.6
no similarity	(> 1E-5)	109	263	2.4
60s ribosomal protein l44 [Ceraceosorus bombacis]	7E-34	125	300	2.4
FMN-dependent NADH-azoreductase [Dechloromonas denitrificans]	3E-85	57	135	2.4
60S ribosomal protein L37a, putative [Bodo saltans]	2E-39	161	380	2.4
putative ribosomal protein L30 [Sipunculus nudus]	3E-53	207	487	2.3
glycine cleavage system regulatory protein [Sorangium cellulosum]	7E-13	53	123	2.3
bax inhibitor 1-like [Nicotiana tomentosiformis]	7E-40	112	255	2.3
no similarity	(> 1E-5)	91	206	2.3
no similarity	(> 1E-5)	70	159	2.3
ribosomal protein S28e [Perkinsela sp. CCAP 1560/4]	3E-23	118	262	2.2
large subunit ribosomal protein L38e [Perkinsela sp. CCAP 1560/4]	8E-25	244	536	2.2
40S ribosomal protein S5 [Zea mays]	5E-100	1339	2893	2.2
no similarity	(> 1E-5)	53	114	2.1
putative 60S ribosomal protein L10a [Leishmania major strain Friedlin]	3E-82	1328	2809	2.1
no similarity	(> 1E-5)	92	193	2.1
type I polyketide synthase [Methylocystis rosea]	3E-145	109	226	2.1
Mg-protoporyphyrin IX chelatase [Euglena gracilis]	0.0	671	1392	2.1
N-alpha-acetyltransferase 20 isoform X1 [Melopsittacus undulatus]	3E-80	52	108	2.1
40S ribosomal protein S12 [Diaphorina citri]	6E-44	751	1529	2.0
ribosomal protein S13p/S18e, putative [Acanthamoeba castellanii str. Neff]	1E-70	983	1995	2.0
no similarity	(> 1E-5)	54	108	2.0

[a] Based on closest BLAST search result. No similarity – E value greater than 1E-5.

conditions, but we found that expression levels showed very little variation. Of the 1419 contigs that each represented at least 0.01% of transcripts (TPM > 100) in the $-O_2$ sample, only 26 show an increase in expression of over 2-fold compared to the $+O_2$ sample. (*Table 4.1*). Many of these were themselves low-expressing or did not contain any significant similarity to known proteins. The most obvious trend is a number of ribosome protein subunits. Yoshida *et al*. recently reported a transcriptome assembly for *E. gracilis* and also observed that expression levels did not vary significantly between aerobically or anaerobically grown samples [20]. Hoffmeister *et al*. found by 2D SDS-PAGE that only a few changes in the proteome of *E. gracilis* appear to occur upon introduction to anoxia [21]. Inui *et al*., in their initial report of wax fermentation in *E. gracilis*, observed that conversion of carbohydrate storages into wax esters took place in just 15 min after placing *E. gracilis* in an oxygen-free environment [8]. In summary, *E. gracilis* does not appear to control metabolism significantly at the transcription and translation levels. Enzyme-level responses to changes in redox homeostasis may be a more important factor in controlling metabolic flux.

Transcriptome mining for gene candidates. We next searched the transcriptome for genes coding for enzymes related to fatty acid metabolism using a range of search queries covering various animals, plants, and bacteria. *E. gracilis* mRNA contains a small *trans*-spliced region at the 5' end that includes the sequence TTTTTTCG [29], and this sequence could be used to identify the reading frame for a given contig, as the start codon was often the first ATG following the TTTTTTCG sequence or the last CAT preceding a CGAAAAAA sequence for negative reading frames. Our analysis was guided by the fact that *E. gracilis* has been reported to contain at least four separate fatty acid synthesis systems. The organism contains a canonical type I cytosolic FAS complex typically found in animals, but also contains a type II FAS in the chloroplast, as is typically found in plants [16]. A difference is that in canonical plant FAS II, there are three forms of the ketosynthase, KASI, KASII, and KASIII, with slightly different functions and chain length specificities, whereas the FAS II complex purified from *E. gracilis* by Hendren and Bloch contains only one 3-oxoacyl-ACP synthase enzyme responsible for both initiation and chain elongation [17]. A third FAS system contained in *E. gracilis* is a CoA based fatty acid elongation system found in the endoplasmic reticulum, which is capable of extending acyl-CoA chain length starting all the way from acetyl-CoA to make fatty acids *de novo*. This is similar to the *de novo* fatty acid synthesis of *Trypanosoma brucei*, which is capable of chain elongation starting from butyryl-CoA [30]. Finally, the wax synthesis pathway of the mitochondria appears to consist of a pathway resembling the reverse of fatty acid degradation, but with NADH-dependent enoyl-CoA reductases [9]. Its exact composition is therefore unknown, but we also included fatty acid degradation enzymes as queries in searching the transcriptome.

Table 4.2 shows the contigs that we identified and their expected function based on BLAST searching, as well as other evidence. We also report indicators of the predicted subcellular localization of each transcript product based on signal peptide analysis with TargetP [31] and Phobius [32], although it is not clear how reflective of protozoan biochemistry the learning sets of these algorithms are. Phobius also predicts transmembrane topology, which is important because chloroplast-targeting pre-sequences in *Euglena* are unusual and contain one or more transmembrane helices. Durnford and Gray identified two types of chloroplast targeted proteins in *E. gracilis*, Class I and Class II [33]. Class I proteins contain two transmembrane helices at the N-terminus, whereas Class II contain a signaling peptide followed by a single transmembrane

Table 4.2. *Candidate enzymes related to fatty acid metabolism.*

Name (if any)	Contig number	Predicted function[a]	E value of closest hit	N-terminal TM helices[b]	Signal peptide[c]	TargetP[d]	RC[d]
	43140	Type I FAS	0.0	-	Yes	M	5
	43684	Acetyl-CoA carboxylase	0.0	2	No	C	5
	44060	Acyl carrier protein	1E-16	2	No	S	4
	35674	Malonyl-CoA:ACP transacylase	1E-77	2	No	Other	2
	36383	3-Oxoacyl-ACP synthase	0.0	2	No	Other	3
	36803	3-Oxoacyl-ACP reductase	1E-112	1	Yes	S	2
	31299	3-Hydroxyacyl-ACP dehydratase	8E-85	1	Yes	S	2
EgTer	41303	Trans-2-enoyl-CoA reductase	0.0	1	Yes	S	4
	16742	Acyl carrier protein	2E-18	-	No	M	1
	37640	Malonyl-CoA:ACP acyltransferase	1E-72	-	No	Other	2
	41129	3-Oxoacyl-ACP reductase	8E-56	-	Yes	M	2
	29789	3-Oxoacyl-ACP reductase	2E-117	-	No	M	4
EgMECR1	37290	Trans-2-enoyl-CoA reductase	2E-69	-	No	M	5
EgMECR2	36244	DnaJ family protein	6E-15	-	No	Other	4
		Trans-2-enoyl-CoA reductase	2E-75				
EgMECR3	31798	Trans-2-enoyl-CoA reductase	4E-38	-	No	M	3
EgMECR4	24588	Putative trans-2-enoyl-CoA reductase	3E-19	-	No	M	4
	43687	Fatty acid elongase	6E-96	-	No	other	2
	30562	Fatty acid elongase	4E-74	-	No	other	2
	31898	Fatty acid elongase	3E-78	-	No	other	4
	23437	Fatty acid elongase	6E-45	-	No	other	1
	33195	Fatty acid elongase	1E-37	-	No	other	4
	30599	Polyunsatured fatty acid elongase	3E-70	-	No	other	3
	12502	Gns1/Sur4 family elongase	5E-64	-	No	S	4
	12559	H. triquetra B-ketoacyl-CoA reductase	0.0	-		S	2
	33735	Very-long-chain 3-oxoacyl-CoA reductase	2E-37	Truncated	Truncated	Truncated	
	19994	Conserved unknown protein	1E-161	-	No	other	3
		Phs1-like (tyrosine phosphatase)	2E-140				

[a] Based on closest BLAST search results, predicted localization, and previous literature.
[b] N-terminal transmembrane helices beyond those expected for the predicted enzyme class.
[c] Based on Phobius [32].
[d] Based on TargetP [31]. RC, Reliability class. 1 is strongest, 5 is weakest.

Table 4.2 (continued). *Candidate enzymes related to fatty acid metabolism.*

Name (if any)	Contig number	Predicted function[a]	E value, closest hit	N-terminal TM helices[b]	Signal peptide[c]	TargetP[d]	RC[d]
EgKAT1	24825	Thiolase, wax fermentation	3E-178	-	No	M	5
EgKAT2	23015	Thiolase, wax fermentation	1E-179	-	Yes	M	4
EgKAT3	23011	Thiolase, wax fermentation	0.0	-	No	Other	4
EgKAT4	36502	Thiolase	0.0	-	No	M	5
EgKAT5	38651	3-Oxoacyl-CoA thiolase, peroxisomal	0.0	-	No	M	5
EgKAT6	27742	Thiolase, Erg10-like		-	Yes	M	3
	43165	Trifunctional enzyme, subunit beta	6E-133	-	No	M	2
	42618	Peroxisomal bifunctional enzyme	0.0	-	No	M	4
	39402	Trifunctional enzyme, subunit alpha	0.0	-	No	Other	5
	28179	L-3-Hydroxyacyl-CoA dehydrogenase	0.0	-	No	M	2
	43988	L-3-Hydroxyacyl-CoA dehydrogenase	8E-98	-	No	M	2
	32047	3-Hydroxyacyl-CoA dehydrogenase, type-2	4E-100	-	Yes	M	2
	30943	Enoyl-CoA hydratase	3E-105	-	No	M	4
	21065	Short-chain enoyl-CoA hydratase	1E-86		No	M	3
	39698	Short-chain acyl-CoA dehydrognenase	0.0	-	No	M	1
	42192	Short-chain acyl-CoA dehydrogenase	0.0	-	No	M	2
	37246	Isobutyryl-CoA dehydrogenase	3E-162	-	No	M	1
	27465	Isovaleryl-CoA dehydrogenase	2E-95	-	No	M	2
	36719	Acyl-Coa dehydrogenase	0.0	-	No	S	5
	35534	Acyl-CoA dehydrogenase	2E-141	-	No	Other	4
	37519	Acyl-CoA oxidase, peroxisomal	2E-144	-	No	Other	3
	42734	Cytochrome b5-like	0.0	-	No	Other	1
		Putative acyl-CoA dehydrogenase	0.0				

[a] Based on closest BLAST search results, predicted localization, and previous literature.
[b] N-terminal transmembrane helices beyond those expected for the predicted enzyme class.
[c] Based on Phobius [32].
[d] Based on TargetP [31]. RC, Reliability class. 1 is strongest, 5 is weakest.

helix. In either case, the chloroplast targeting peptides are rather long. Mitochondrial targeting sequences in *E. gracilis* and the related Trypanosomatids are diverse and more difficult to identify, though are generally somewhat richer in arginine and hydrophobic residues [34]. We identified contig 43140 as the FAS I complex. We also identified exactly one putative acetyl-CoA carboxylase complex (43684), 3-oxoacyl-ACP synthase (36383), and 3-hydroxyacyl-ACP dehydratase (31299) in *E. gracilis.* All three of these proteins had long N-terminal peptides preceding the start of the functional enzyme domain, and Phobius revealed that these contained either two transmembrane helices or a signal peptide and one transmembrane helix, consistent with a *E. gracilis* chloroplast targeting sequence. Two acyl carrier proteins, two malonyl-CoA:ACP transacylases, and three 3-oxoacyl-ACP reductase candidates were identified, but only one of each type contained a putative chloroplast targeting sequence.

The final enzyme required for a complete chloroplast FAS II system is an enoyl-ACP reductase, but we did not find any enoyl-ACP reductase candidates by searching the *Arabidopsis thaliana* homolog, or bacterial enoyl-ACP reductases FabI, InhA, FabK, or FabL. We also did not obtain any hits for the crotonyl-CoA reductase/carboxylase, Ccr. Querying with a MECR-type metazoan mitochondrial enoyl-CoA reductase identified four potential MECR-like enoyl-reductase candidates, which we named EgMECR1-4. None of these contained predicted chloroplast targeting sequences. The only other enoyl-ACP reductase search queries that returned any hit in our *E. gracilis* transcriptome were FabV-type proteins, which include TdTer, and the only result was EgTer itself. As reported by Hoffmeister *et al.*, EgTer does contain a long N-terminal pre-sequence that is cleaved to provide the final active enzyme [21]. We found that this pre-sequence strongly resembled those of the chloroplast targeting enzymes of the other enzymes predicted to be part of the FAS II system. This suggests that EgTer may actually be a component of the chloroplast FAS II system, either in addition to or instead of being involved in wax ester production in the mitochondria.

We attempted to test this hypothesis by determining the subcellular localization of EgTer through organelle isolation and Western blot. The isolation of intact mitochondria from *E. gracilis* is challenging, however, as the mitochondrial structure is highly variable and can often adopt the form of an interconnected reticulum-type network [35]. As an alternative, we obtained an etiolated strain of *E. gracilis* that lacks chloroplasts, and we carried out Western blot of whole

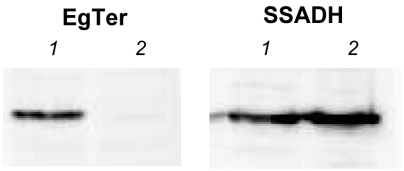


Figure 4.3. *Western blot of E. gracilis whole cell lysate with primary antibody against EgTer or SSADH. 1. Green E. gracilis. 2. Etiolated E. gracilis.*

cell lysate in both the green and etiolated strains using antibodies raised to either EgTer or to a synthetic peptide from contig 43096, which is predicted to be a mitochondrially targeted succinate semialdehyde dehydrogenase (SSADH). SSADH enzyme is commonly used as a marker of mitochondria in *E. gracilis.* The result showed that SSADH was detected in both green and etiolated strains, whereas EgTer was not detected in whole cell lysate of the etiolated strain (*Figure 4.3*). This result contradicts the idea that EgTer is a component of wax fermentation.

We identified seven candidate fatty acid elongases in *E. gracilis* that could be part of its ER-bound fatty acid synthesis system. This system would be expected to also include a membrane-bound 3-oxoacyl-CoA reductase, 3-hydroxyacyl-CoA dehydratase, and enoyl-CoA reductase. In the very-long chain elongation system of *Saccharomyces cerevisiae*, these enzymes correspond to Ifa38, Phs1, and Tsc13, respectively [36-38]. Phs1 is similar to tyrosine phosphatase enyzmes. In *E. gracilis*, we found two Ifa38-like candidates. Contig 12559 was actually an exact duplication of a predicted membrane-bound 3-oxoacyl-CoA reductase from *Heterocapsa triquetra*. Contig 33735 was similar to very-long chain 3-oxoacyl-CoA reductases, but appeared to be truncated at the N-terminus. The only result for Phs1 was a large protein that also contained a highly conserved protein domain of unknown function at the N-terminus. No hits were obtained using the enoyl-CoA reductase Tsc13 as query.

Because the wax fermentation pathway is expected to resemble the reverse of β-oxidation, we also surveyed the transcriptome for β-oxidation related enzymes. Seven candidate thiolases were identified, and six of these corresponded to KAT1 through KAT6 found by Nakazawa *et al*. [10]. To better predict the function of these thiolases, we aligned the sequences with a number of other eukaryotic thiolases (*Figure 4.4*). Sequence gaps characteristic of particular thiolase classes, particularly in the covering loop region and near the C-terminus, were the easiest way to compare sequences. Contig 43165 resembled the β-subunit of the mitochondrial trifunctional enzyme, because of the additional 17 residues at the C-terminus. KAT5 resembled human peroxisomal thiolase AB, with a sequence gap corresponding to its lack of a tetramerization loop. KAT6 was highly similar to T2-type thiolases such as human T2 or *S. cerevisiae* Erg10. KAT1 through KAT4 did not correspond closely to any of the other thiolases. They contained a longer sequence in the covering loop compared to other thiolases but did not share the 17 additional residues found in TFE-β thiolases. Nakazawa *et al*. found through RNAi knockdown experiments that KAT1, KAT2, and KAT3 are involved in wax fermentation.

Other β-oxidation enzymes were identified, including two standalone L-3-hydroxyacyl-CoA dehydrogenases that matched the sequences of the dehydrogenases reported by Winkler *et al*. for the β-oxidation complex of *E. gracilis* [39]. We also list candidate acyl-CoA dehydrogenase enzymes in *Figure 4.4* , even though NADH-dependent reductases are expected to carry out the function of enoyl reduction in the wax fermentation pathway.

Characterization of MECR enzymes. Other than EgTer, EgMECR1 through EgMECR4 were the only candidate enoyl-CoA reductases found in our transcriptome. EgMECR1 and EgMECR2 had the highest similarity to annotated mitochondrial enoyl-CoA reductases, although EgMECR2 had an unusual architecture that included a DnaJ-like chaperone domain fused at the N-terminus. EgMECR3 and EgMECR4 had less certainty in their functional assignment as enoyl reductases, and EgMECR4 was slightly smaller, with a molecular eight of only 29kDa. We cloned these genes into *E. coli* expression vectors to purify and test the enzymes for enoyl-CoA reductase activity (*Figure 4.5*). Soluble expression of EgMECR2 was only possible by fusing EgMECR2 to maltose binding protein at the N-terminus. A significant amount of EgMECR2 passed through the Ni-NTA resin with the flow-through fraction and subsequent washes, due to poor binding of the His_6-MBP-EgMECR2 fusion, but sufficient pure protein for subsequent experiments was obtained upon elution. Of the four MECR-type enzymes, however, only EgMECR1 displayed any activity as an enoyl-CoA reductase (*Table 4.3*). The enzyme preferred NADPH as substrate, and had activity with both crotonyl-CoA and dodecenoyl-CoA.

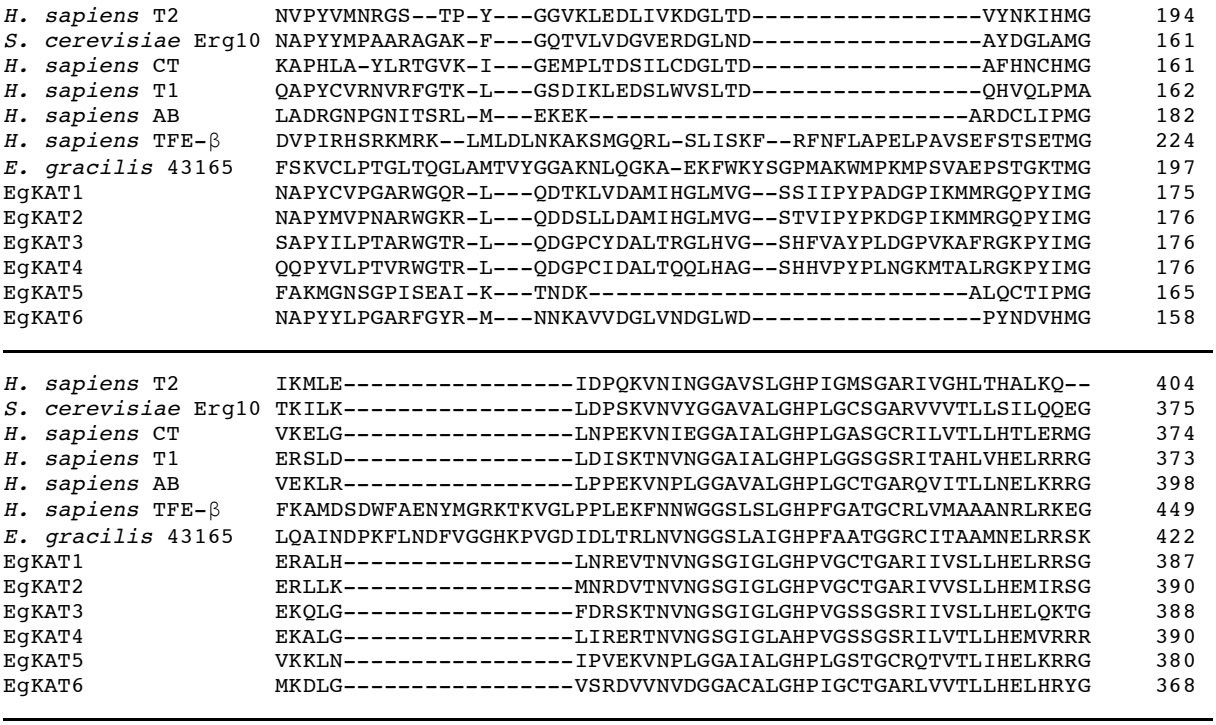


Figure 4.4. *Alignment of two regions of thiolase sequences from E. gracilis with other eukaryotic thiolases. Sequence gaps serve as an indication of thiolase class and function.*

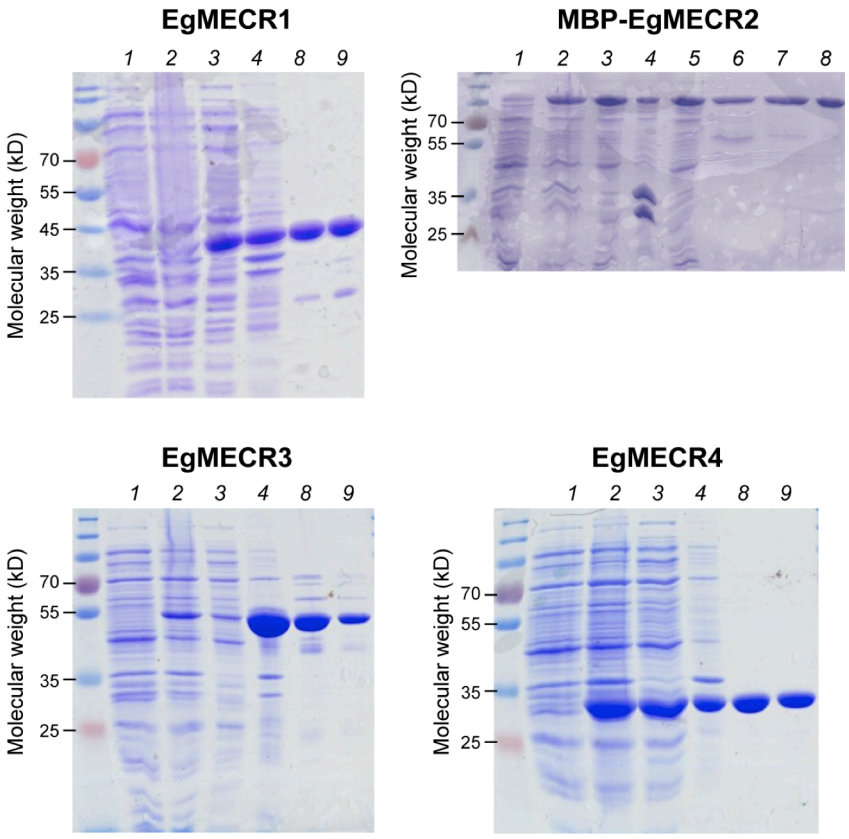


Figure 4.5. *SDS-PAGE of EgMECR protein purifications. 1. Pre-induction, 2. Post-induction, 3. Soluble lysate, 4. Insoluble lysate, 5. Ni-NTA flow-through, 6. First wash, 7. Second Wash, 8. Ni-NTA elution, 9. G-25 resin buffer exchange.*

Table 4.3. *Activity of EgMECR1.*[a]

Substrate	Cofactor	Activity (U/mg)
Crotonyl-CoA (100 μM)	NADH	1.7
Crotonyl-CoA (100 μM)	NADPH	32
Dodecenoyl-CoA (10 μM)	NADPH	4

[a]U/mg – μmol min^{-1} mg^{-1} protein.

4.4. Discussion

We were unable to increase cell density of *E. gracilis* under anaerobic conditions, in contrast to Tucci *et al*. [26]. A recent study by Furuhashi *et al*. found that the presence of CO_2 or sodium bicarbonate was necessary for wax production in anoxic *E. gracilis [40]*, although the physiological basis for this observation is not clear. Tucci *et al*. and others have reported that the proportion of odd-chain waxes increases under anaerobic conditions, in which propionyl-CoA is formed through succinate, which comes from the carboxylation and reduction of pyruvate. Even though wax production from glucose has a net release of CO_2, it could be that sufficient CO_2 is initially required to maintain a high enough rate of the pyruvate carboxylation step for propionyl-CoA production. The media of Tucci *et al*. is the same media used by us and contains 2 mM carbonate in the form of calcium carbonate. For Furuhashi *et al*., 1 mM sodium bicarbonate was insufficient to recover wax production but 10 mM was sufficient. The discrepancy between these observations remains to be resolved.

The localization of EgTer, whether to mitochondria or chloroplasts, has still not been conclusively determined. Hoffmeister *et al*. originally isolated EgTer from whole *E. gracilis* cell lysate and separately detected enoyl-CoA reductase activity in isolated mitochondria but not isolated chloroplast fractions [14]. Our genetic and biochemical evidence points towards a role for EgTer in the Type II FAS system found in *E. gracilis* chloroplasts, and we did not detect EgTer in an etiolated strain by Western blot. Although purified EgTer is clearly very active with enoyl-CoA substrates, this does not exclude its potential ability to act as an enoyl-ACP reductase. EgTer and TdTer are similar to FabV, an enoyl-ACP reductase in *Vibrio cholera* [41]. Like *E. gracilis, T. denticola* contains all the enzymes necessary for a complete type II FAS system, except that TdTer is its only candidate enoyl-ACP reductase. Vick *et al.* have found that FabI, an *E. coli* enoyl-ACP reductase, can substitute for EgTer in the efficient production of butyrate using a reverse β-oxidation pathway [42]. Although both EgTer and TdTer were originally reported as functioning better on short-chain enoyl-CoA substrates than long-chain, Bond-Watts *et al*. reported higher catalytic efficiency with dodecenoyl-CoA than crotonyl-CoA for TdTer that was initially obscured due to potent product inhibition [24]. Thus it is feasible that EgTer functions as an enoyl-ACP reductase that also has activity with enoyl-CoA substrates. We cannot rule out the possibility that it does indeed have a role in wax fermentation or is targeted to the mitochondria in at least some types of growth conditions.

Of the other candidate enoyl reductase enzymes we identified in *E. gracilis*, only EgMECR1 showed enoyl reductase activity. EgMECR1 does not resemble the enoyl-CoA reductases reported by Inui *et al*., as it preferred NADPH over NADH as cofactor [9]. We did not identify any enzymes that resemble the heterodimeric architecture reported by Inui *et al.* for short-chain enoyl-CoA reductase [12]. EgMECR4 comes the closest, as it is a smaller protein of only 29kDa. If a mitochondrial targeting sequence were cleaved at the N-terminus, it is feasible that MECR4 could correspond to the 25kDa subunit of the heterodimeric short-chain enoyl reductase. If that is the case, some other 15kDa would also be necessary for activity. The possibility that acyl-CoA dehydrogenase enzymes function as reductases in wax fermentation has also not been ruled out. An electron transport chain based pathway may be at work, similar to the branched acid anaerobic respiration of *Ascaris suum*. Further work will be needed to determine all of the enzymes involved in the unique wax ester metabolism of *E. gracilis.*

4.5. References

1. T. Ghaffar, M. Irshad, Z. Anwar, T. Aqil, Z. Zulifqar, A. Tariq, M. Kamran, N. Ehsan, and S. Mehmood, Recent trends in lactic acid biotechnology: A brief review on production to purification. *Journal of Radiation Research and Applied Sciences* **2014,** 7, (2), 222-229.
2. Y. Lin and S. Tanaka, Ethanol fermentation from biomass resources: current state and prospects. *Appl. Microbiol. Biotechnol.* **2006,** 69, (6), 627-42.
3. A. Baez, K. M. Cho, and J. C. Liao, High-flux isobutanol production using engineered *Escherichia coli*: a bioreactor study with in situ product removal. *Appl. Microbiol. Biotechnol.* **2011,** 90, (5), 1681-90.
4. B. B. Bond-Watts, R. J. Bellerose, and M. C. Y. Chang, Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat. Chem. Biol.* **2011,** 7, (4), 222-227.
5. C. Dellomonaco, J. M. Clomburg, E. N. Miller, and R. Gonzalez, Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **2011,** 476, (7360), 355-359.
6. C. R. Shen, E. I. Lan, Y. Dekishima, A. Baez, K. M. Cho, and J. C. Liao, Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl. Environ. Microbiol.* **2011,** 77, (9), 2905-15.
7. A. Atteia, R. van Lis, A. G. Tielens, and W. F. Martin, Anaerobic energy metabolism in unicellular photosynthetic eukaryotes. *Biochim. Biophys. Acta* **2013,** 1827, (2), 210-23.
8. H. Inui, K. Miyatake, Y. Nakano, and S. Kitaoka, Wax ester fermentation in *Euglena gracilis*. *FEBS Lett.* **1982,** 150, (1), 89-93.
9. H. Inui, K. Miyatake, Y. Nakano, and S. Kitaoka, Fatty acid synthesis in mitochondria of *Euglena gracilis*. *Eur. J. Biochem.* **1984,** 142, (1), 121-6.
10. M. Nakazawa, H. Andoh, K. Koyama, Y. Watanabe, T. Nakai, M. Ueda, T. Sakamoto, H. Inui, Y. Nakano, and K. Miyatake, Alteration of wax ester content and composition in *Euglena gracilis* with gene silencing of 3-ketoacyl-CoA thiolase isozymes. *Lipids* **2015,** 50, (5), 483-92.
11. H. Inui, K. Ono, K. Miyatake, Y. Nakano, and S. Kitaoka, Purification and characterization of pyruvate:NADP+ oxidoreductase in *Euglena gracilis*. *J. Biol. Chem.* **1987,** 262, (19), 9130-5.
12. H. Inui, K. Miyatake, Y. Nakano, and S. Kitaoka, Purification and some properties of short chain-length specific trans-2-enoyl-CoA reductase in mitochondria of *Euglena gracilis*. *J. Biochem.* **1986,** 100, (4), 995-1000.
13. S. Sanker, H. A. Campbell, and C. Kent, Negative cooperativity of substrate binding but not enzyme activity in wild-type and mutant forms of CTP:glycerol-3-phosphate cytidylyltransferase. *J. Biol. Chem.* **2001,** 276, (41), 37922-8.

14. M. Hoffmeister, M. Piotrowski, U. Nowitzki, and W. Martin, Mitochondrial trans-2-enoyl-CoA reductase of wax ester fermentation from *Euglena gracilis* defines a new family of enzymes involved in lipid synthesis. *J. Biol. Chem.* **2005,** 280, (6), 4329-38.

15. S. Tucci and W. Martin, A novel prokaryotic trans-2-enoyl-CoA reductase from the spirochete *Treponema denticola*. *FEBS Lett.* **2007,** 581, (8), 1561-6.

16. J. Delo, M. L. Ernst-Fonberg, and K. Bloch, Fatty acid synthetases from *Euglena gracilis*. *Arch. Biochem. Biophys.* **1971,** 143, (2), 384-391.

17. R. W. Hendren and K. Bloch, Fatty acid synthetases from *Euglena gracilis*. Separation of component activities of the ACP-dependent fatty acid synthetase and partial purification of the beta-ketoacyl-ACP synthetase. *J. Biol. Chem.* **1980,** 255, (4), 1504-8.

18. A. A. Khan and P. E. Kolattukudy, A microsomal fatty acid synthetase coupled to acyl-CoA reductase in *Euglena gracilis*. *Arch. Biochem. Biophys.* **1973,** 158, (1), 411-20.

19. E. C. O'Neill, M. Trick, L. Hill, M. Rejzek, R. G. Dusi, C. J. Hamilton, P. V. Zimba, B. Henrissat, and R. A. Field, The transcriptome of *Euglena gracilis* reveals unexpected metabolic capabilities for carbohydrate and natural product biochemistry. *Mol. Biosyst.* **2015,** 11, (10), 2808-20.

20. Y. Yoshida, T. Tomiyama, T. Maruta, M. Tomita, T. Ishikawa, and K. Arakawa, De novo assembly and comparative transcriptome analysis of *Euglena gracilis* in response to anaerobic conditions. *BMC Genomics* **2016,** 17, (1), 182.

21. M. Hoffmeister, A. van der Klei, C. Rotte, K. W. van Grinsven, J. J. van Hellemond, K. Henze, A. G. Tielens, and W. Martin, *Euglena gracilis* rhodoquinone:ubiquinone ratio and mitochondrial proteome differ under aerobic and anaerobic conditions. *J. Biol. Chem.* **2004,** 279, (21), 22422-9.

22. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011,** 29, (7), 644-52.

23. 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, (5), 343-5.

24. B. B. Bond-Watts, A. M. Weeks, and M. C. Chang, Biochemical and structural characterization of the trans-enoyl-CoA reductase from *Treponema denticola*. *Biochemistry* **2012,** 51, (34), 6827-37.

25. I. Carre, J.-L. Bomsel, and R. Calvayrac, Decay of cytochromes and appearance of a cyanide-insensitive electron transfer pathway in *Euglena gracilis* grown in anoxia. *Plant Sci.* **1988,** 54, (3), 193-202.

26. S. Tucci, R. Vacula, J. Krajcovic, P. Proksch, and W. Martin, Variability of wax ester fermentation in natural and bleached *Euglena gracilis* strains in response to oxygen and the elongase inhibitor flufenacet. *J. Eukaryot. Microbiol.* **2010,** 57, (1), 63-9.

27. M. N. Schnare and M. W. Gray, Sixteen discrete RNA components in the cytoplasmic ribosome of *Euglena gracilis*. *J. Mol. Biol.* **1990,** 215, (1), 73-83.

28. B. Li and C. N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **2011,** 12, 323.

29. L. H. Tessier, M. Keller, R. L. Chan, R. Fournier, J. H. Weil, and P. Imbault, Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by trans-splicing in *Euglena*. *EMBO J.* **1991,** 10, (9), 2621-5.

30. S. H. Lee, J. L. Stephens, K. S. Paul, and P. T. Englund, Fatty acid synthesis by elongases in trypanosomes. *Cell* **2006,** 126, (4), 691-9.

31. O. Emanuelsson, H. Nielsen, S. Brunak, and G. von Heijne, Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **2000,** 300, (4), 1005-16.

32. L. Kall, A. Krogh, and E. L. Sonnhammer, Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res.* **2007,** 35, (Web Server issue), W429-32.

33. D. G. Durnford and M. W. Gray, Analysis of *Euglena gracilis* plastid-targeted proteins reveals different classes of transit sequences. *Eukaryot. Cell* **2006,** 5, (12), 2079-91.

34. K. Krnacova, M. Vesteg, V. Hampl, C. Vlcek, and A. Horvath, *Euglena gracilis* and Trypanosomatids possess common patterns in predicted mitochondrial targeting presequences. *J. Mol. Evol.* **2012,** 75, (3-4), 119-29.

35. M. Pellegrini, Three-dimensional reconstruction of organelles in *Euglena gracilis* Z. I. Qualitative and quantitative changes of chloroplasts and mitochondrial reticulum in synchronous photoautotrophic culture. *J. Cell Sci.* **1980,** 43, 137-66.

36. G. Han, K. Gable, S. D. Kohlwein, F. Beaudoin, J. A. Napier, and T. M. Dunn, The *Saccharomyces cerevisiae* YBR159w gene encodes the 3-ketoreductase of the microsomal fatty acid elongase. *J. Biol. Chem.* **2002,** 277, (38), 35440-9.

37. A. Kihara, H. Sakuraba, M. Ikeda, A. Denpoh, and Y. Igarashi, Membrane topology and essential amino acid residues of Phs1, a 3-hydroxyacyl-CoA dehydratase involved in very long-chain fatty acid elongation. *J. Biol. Chem.* **2008,** 283, (17), 11199-209.

38. S. D. Kohlwein, S. Eder, C. S. Oh, C. E. Martin, K. Gable, D. Bacikova, and T. Dunn, Tsc13p is required for fatty acid elongation and localizes to a novel structure at the nuclear-vacuolar interface in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **2001,** 21, (1), 109-25.

39. U. Winkler, W. Saftel, and H. Stabenau, A new type of a multifunctional beta-oxidation enzyme in euglena. *Plant Physiol.* **2003,** 131, (2), 753-62.

40. T. Furuhashi, T. Ogawa, R. Nakai, M. Nakazawa, A. Okazawa, A. Padermschoke, K. Nishio, M. Y. Hirai, M. Arita, and D. Ohta, Wax ester and lipophilic compound profiling of *Euglena gracilis* by gas chromatography-mass spectrometry: toward understanding of wax ester fermentation under hypoxia. *Metabolomics* **2015,** 11, (1), 175-183.

41. R. P. Massengo-Tiasse and J. E. Cronan, Vibrio cholerae FabV defines a new class of enoyl-acyl carrier protein reductase. *J. Biol. Chem.* **2008,** 283, (3), 1308-16.

42. J. E. Vick, J. M. Clomburg, M. D. Blankschien, A. Chou, S. Kim, and R. Gonzalez, *Escherichia coli* enoyl-acyl carrier protein reductase (FabI) supports efficient operation of a functional reversal of beta-oxidation cycle. *Appl. Environ. Microbiol.* **2015,** 81, (4), 1406-16.

Appendix: *DNA sequences and NMR spectra*

Table S1. Base strains and plasmids (A), oligonucleotides (B), gBlock maps (C), scheme for construction of RBS/promoter screen plasmids (D), scheme for construction of PHA production vectors (E), and amino acid sequences of purified proteins (F) used in *Chapter 2*.

A. Strains and plasmids. Number in parentheses after some genes represents RBS strength predicted by the RBS Calculator [1].

Strain	*Genotype*	*Source*
BAP1	F– ompT gal dcm lon hsdSB(rB- mB-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) ΔprpRBCDE (sfp (T7), prpE (T7))]	[2]
DH10B	F- endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ-	Invitrogen
BL21(de3)	F^- *omp*T *gal dcm lon hsd*S_B(r_B^- m_B^-) λ(DE3 [*lac*I *lac*UV5-T7 *gene 1 ind*1 *sam*7 *nin*5])	Novagen

Plasmid	*Description*	*Source*
pET16b-AsAcat1	*His_{10}-Xa-Acat1* (T7), *lacI*, Cb^r, pBR322	This study
pSV272.1-AsAcat2	*His_6-MBP-Acat2* (T7), *lacI*, Km^r, pBR322	This study
pSV272.1-AsAcat3	*His_6-MBP-Acat3* (T7), *lacI*, Km^r, pBR322	This study
pSV272.1-AsAcat4	*His_6-MBP-Acat4* (T7), *lacI*, Km^r, pBR322	This study
pSV272.1-AsAcat5	*His_6-MBP-Acat5* (T7), *lacI*, Km^r, pBR322	This study
pET23a-AsAcat1	*His_6-TEV-Acat1* (T7), Cb^r, pBR322	This study
pET23a-AsAcat2	*His_6-TEV-Acat2* (T7), Cb^r, pBR322	This study
pET23a-AsAcat3	*His_6-TEV-Acat3* (T7), Cb^r, pBR322	This study
pET23a-AsAcat4	*His_6-TEV-Acat4* (T7), Cb^r, pBR322	This study
pET23a-AsAcat5	*His_6-TEV-Acat5* (T7), Cb^r, pBR322	This study
pET16b-AsHadh	*His_{10}-Xa-AsHadh* (T7), *lacI*, Cb^r, pBR322	This study
pET16b-AsEch	*His_{10}- Xa-AsEch* (T7), *lacI*, Cb^r, pBR322	This study
pET16b-TesB	*His_{10}- Xa-TesB* (T7), *lacI*, Cb^r, pBR322	This study
pET28a-CnHadh	*His_6-CnHadh* (T7), *lacI*, Km^r, pBR322	This study
pTrc99-A2-AsHadh-Ech	*Acat2*(33k)-*AsHadh*(200k)- *AsEch*(28k) (Trc), *lacI*, Cb^r, pBR322	This study
pTrc99-A3-AsHadh-Ech	*Acat3*(31k)-*AsHadh*(190k)- *AsEch*(28k) (Trc), *lacI*, Cb^r, pBR322	This study
pTrc99-A5-AsHadh-Ech	*Acat5*(26k)-*AsHadh*(200k)- *AsEch*(28k) (Trc), *lacI*, Cb^r, pBR322	This study
pTrc99-A2-AsHadh	*Acat2*(33k)-*AsHadh*(200k) (Trc), *lacI*, Cb^r, pBR322	This study
pTrc99-A3-AsHadh	*Acat3*(31k)-*AsHadh*(190k) (Trc), *lacI*, Cb^r, pBR322	This study
pTrc99-A5-AsHadh	*Acat5*(26k)-*AsHadh*(200k) (Trc), *lacI*, Cb^r, pBR322	This study
pTrc33-TesA	*TesA*(9.7k) (Trc), Cm^r, p15a	This study
pTrc33-TesB	*TesB*(9.2k) (Trc), Cm^r, p15a	This study
pTrc33-YciA	*YciA*(8.4k) (Trc), Cm^r, p15a	This study
pTrc33-YdiI	*YdiI*(6.3k) (Trc), Cm^r, p15a	This study
pTrc33-TesB2	*TesB*(110k) (Trc), Cm^r, p15a	This study
pTrc99-A3AsH00	*Acat3*(11k)-*AsHadh*(10k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3AsH00	*Acat3*(11k)-*AsHadh*(10k) (TacTac), *lacI*, Cb^r, ColE1	This study
pTrc99-A3AsH01	*Acat3*(11k)-*AsHadh*(100k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3AsH01	*Acat3*(11k)-*AsHadh*(100k) (TacTac), *lacI*, Cb^r, ColE1	This study
pTrc99-A3AsH10	*Acat3*(120k)-*AsHadh*(10k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3AsH10	*Acat3*(120k)-*AsHadh*(10k) (TacTac), *lacI*, Cb^r, ColE1	This study
pTrc99-A3AsH11	*Acat3*(120k)-*AsHadh*(100k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3AsH11	*Acat3*(120k)-*AsHadh*(100k) (TacTac), *lacI*, Cb^r, ColE1	This study
pTrc99-A3CnH00	*Acat3*(11k)-*CnHadh*(10k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3CnH00	*Acat3*(11k)-*CnHadh*(10k) (TacTac), *lacI*, Cb^r, ColE1	This study
pTrc99-A3CnH01	*Acat3*(11k)-*CnHadh*(100k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3CnH01	*Acat3*(11k)-*CnHadh*(100k) (TacTac), *lacI*, Cb^r, ColE1	This study
pTrc99-A3CnH10	*Acat3*(120k)-*CnHadh*(18k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3CnH10	*Acat3*(120k)-*CnHadh*(18k) (TacTac), *lacI*, Cb^r, ColE1	This study
pTrc99-A3CnH11	*Acat3*(120k)-*CnHadh*(120k) (Trc), *lacI*, Cb^r, pBR322	This study
pCWori-A3CnH11	*Acat3*(120k)-*CnHadh*(120k) (TacTac), *lacI*, Cb^r, ColE1	This study

pPol4.1	*NphT7-PhaB-RruPhaC* (Trc), *lacI*, Cb^r, pBR322	Dr. Benjamin Thuronyi
pPol4.2	*NphT7-PhaB-RopPhaC* (Trc), *lacI*, Cb^r, pBR322	Dr. Benjamin Thuronyi
pBT33-Bu3	*PhaA-PhaB* (Ara), *PhaJ* (Trc), Cm^r, p15a	[3]
Poly1	*Acat2*(9.1k)-*RruPhaC* (Trc), *lacI*, Cb^r, pBR322	This study
Poly2	*Acat3-RruPhaC* (Trc), *lacI*, Cb^r, pBR322	This study
Poly3	*Acat2-RopPhaC* (Trc), *lacI*, Cb^r, pBR322	This study
Poly4	*Acat3-RopPhaC* (Trc), *lacI*, Cb^r, pBR322	This study
Poly1B	*Acat2*(87k)-*RruPhaC* (Trc), *lacI*, Cb^r, pBR322	This study
Poly7B	*NphT7-RruPhaC* (Trc), *lacI*, Cb^r, pBR322	This study
Poly8B	*NphT7-RopPhaC* (Trc), *lacI*, Cb^r, pBR322	This study
DKD2	*PhaJ-AsHadh-AsEch* (Trc), Cm^r, p15a	This study
pTrc33-PhaB	*PhaB* (Trc), Cm^r, p15a	This study

B. Oligonucleotide sequences

Name	*Sequence*
AsAcat1 23a F1	catcatgagaatctctacttccagggtaccggccgtagcttctctgacgtggtgattg
AsAcat1 23a R1	ccggatctcagtggtggtggtggtggtgctcgattacaggcgttcaataatcataccgcc
AsAcat2 23a F1	atcatgagaatctctacttccagggtaccggcagccgcccgatcaccgac
AsAcat2 23a R1	tctcagtggtggtggtggtggtgctcgattacagcttctcaataatcataccacctgcac
AsAcat3 23a F1	atcatgagaatctctacttccagggtaccggcgcgagccagttcaccgacg
AsAcat3 23a R1	atctcagtggtggtggtggtggtgctcgattacaatttctcaatcaccataccgcctgc
AsAcat4 23a F1	agaatctctacttccagggtaccggcagcgcaaatatttttaaggatgttgtgatttgtg
AsAcat4 23a R1	ctcagtggtggtggtggtggtgctcgattacagcttctcaataatcataccagacgcac
AsAcat5 23a F1	gagaatctctacttccagggtaccggcatggatgcaggcaaaaaagacgtttacatcc
AsAcat5 23a R1	ctcagtggtggtggtggtggtgctcgattacagtttttttaatcagaatagcggaagcctc
Acat2 FP	tggaattgtgagcggataacaatttcacacaggaaacagataagcatcgaagagataaggagcgcaaaagggatgagccgcccgatcacc
Acat2 RP-Hadh	gacatagaggacctccttaaatttttttttacagcttctcaataatcataccacctgcacc
Acat2 RP-PhaB	gtcctcctaaatattgctaaaacttgatttacagcttctcaataatcataccacctgcac
Acat3 FP	attgtgagcggataacaatttcacacaggaaacagtttaaagtaaattactatcttaaggaattaagtaatggcgagccagttcaccgac
Acat3 RP-Hadh	aacctccttaatatttgggttgtactaggcttacaatttctcaatcaccataccgcctgc
Acat3 RP-PhaB	gagtcctcctaaatattgctaaaacttgatttacaatttctcaatcaccataccgcctgc
Acat5 FP	agcggataacaatttcacacaggaaacagtgaagaaggacgcgattactaagagagatagaaaatggacgcaggcaaaaaaagacgtttac
Acat5 RP-Hadh	acatagaggacctccttaaatttttttttacagtttttaatcagaatagcggaagcctcg
Acat5 RP-PhaB	cctcctaaatattgctaaaacttgatttacagtttttaatcagaatagcggaagcctcg
Hadh FP-Acat2	tgtaaaaaaaatttaaggaggtcctctatgtccgcactgcgctccacgaaaggtctggtc
Hadh FP-Acat3	aagcctagtacaacccaaatattaaggaggtttatcatgtccgcactgcgctcc
Hadh FP-Acat5	ttctgattaaaaaactgtaaaaaaaatttaaggaggtcctctatgtccgcactgcgctcc
Hadh RP-Ech	tcattgtttctttttcctgtgttttggttgcgagtgtggttacgccggcatacgcagagc
Hadh RP	gctgaaaatcttctctcatccgccaaaacagccaagcttttacgccggcatacgcagagc
Ech FP	tggatggtgctctgcgtatgccggcgtaaccacactcgcaaccaaaacacaggaaaaagaaacaatgacggttaaactgatggcgtccac
Ech RP	aatcttctctcatccgccaaaacagccaagcttttaggaggaggtccatttaggggtacg
TesA FP	tgtgagcggataacaatttcacacgagctcaaaaaattattccgaagaaggagcaaaagtaagatggcagcggacacgttattgattctg
TesA RP	tctcatccgccaaaacagccaagcttttatgagtcatgatttactaaaggctgcaactgc
TesB FP	gcgtcattgttagctaatgatccccccataacctaaggagaattaaatagatgagtcaggcgctaaaaaatttactgacattgttaaatc
TesB2 FP	gagcggataacaatttcacacgagctccgttcggggcataaataagggaggaaattcatgagtcaggcgctaaaaaatttactgacattg
TesB RP	cttctctcatccgccaaaacagccaagcttttaattgtgattacgcatcacccctcctg
YciA FP	aattgtgagcggataacaatttcacacaattatcttaaatttttagtaagctatctctatacatgtctacaacacataacgtccctcagg
YciA RP	gaaaatcttctctcatccgccaaaacagccaagcttttactcaacaggtaaggcgcgagg
YdiI FP	gcggataacaatttcacacgagctcagagccttatcacaatcctataagagaacacgaaatgatatggaaacggaaaatcaccctggaag
YdiI RP	gaaaatcttctctcatccgccaaaacagccaagctttcacaaaatggcggtcgtcaatcg

Acat3 F0	aattgtgagcggataacaatttcacacaggaaacatttagggaatttagagttctcggaataaggtatatatggcgagccagtt caccg
Acat3 F1	attgtgagcggataacaatttcacacaggaaacaaagagaagagttattaatctctaggaggtaagtatggcgagccagttc accgac
Acat3 AsHR0	tatccctccctactatttaagggtgtatattgattacaatttctcaatcaccataccgcctgc
Acat3 AsHR1	gacatagatattcctccttataacggtttattacaatttctcaatcaccataccgcctgc
Acat3 CnHR0	actcctctaaatatgaatttataagtgttctttacaatttctcaatcaccataccgcctg
Acat3 CnHR1	ttcctcctctttttcttataaatagaaaggttacaatttctcaatcaccataccgcctgc
AsHF0	tcaatatacacccttaaatagtagggagggataatgtccgcactgcgctcc
AsHF1	tgtaataaaccgttataaggaggaatatctatgtccgcactgcgctcc
AsHadh CWR	tcagatctgctcatgtttgacagcttatcatcgatttacgccggcatacgcagag
CnHF0	acacttataaattcatatttagaggagtttttatgcaaattcagggcaacgtttttattg
CnHF1	ctttctatttataagaaaaagaggaggaatcaatgcaaattcagggcaacgtttttattg
CnHadh CWR	agatctgctcatgtttgacagcttatcatcgatttattttggctgcatacgaatggcacc
CnHadh 99R	aatcttctctcatccgccaaaacagccaagcttttatttggctgcatacgaatggcacc
Poly1,3 Fwd	cggataacaatttcacacaggaaacagaccaggagatatataatgagccgcccgatcacc
Poly1 R	tccagcattaattatacctccggactagtttacagcttctcaataatcataccacctgca
Poly1 F	aagctgtaaactagtccggaggtataattaatgctggatcacgtgcataaaaaactgaag
Poly1,2 Rev	agcttgcatgcctgcaggtcgactggatccttagctaaaaacgtatgtaccaggggcatc
Poly2,4 Fwd	ggataacaatttcacacaggaaacagaccaggagatatataatggcgagccagttcaccg
Poly2 R	taattatacctccggactagtttacaatttctcaatcaccataccgcctg
Poly2 F	gagaaattgtaaactagtccggaggtataattaatgctggatcacgtgcataaaaaactg
Poly3,4 Rev	agcaagcttgcatgcctgcaggtcgactggatccttaggagtgatgcacatatgtgcccg
Poly3 R	ctgacattaattatacctccggactagtttacagcttctcaataatcataccacctgcac
Poly3 F	gctgtaaactagtccggaggtataattaatgtcagataccccgctgccg
Poly4 R	taattatacctccggactagtttacaatttctcaatcaccataccgcctg
Poly4 F	gattgagaaattgtaaactagtccggaggtataattaatgtcagataccccgctgccg
Poly1B Fwd	gctcgtataatgtgtggaattgtgagcggataacgaattctaaatcccttcaagtcctaaggagcgttataaatgagccgcccg atcacc
Poly7,8 Fwd	aattgtgagcggataacaatttcacacaggaaacagaccaggagatatataatgaccgacgttcgttttcgtatcattg
Poly7B R	cattaattatacctccggactagtgaattcttaccactcgatcagcgcgaag
Poly7B F	taagaattcactagtccggaggtataattaatgctggatcacgtgcataaaaaactgaag
Poly8B R	attaattatacctccggactagtgaattcttaccactcgatcagcgcgaag
Poly8B F	ggtaagaattcactagtccggaggtataattaatgtcagataccccgctgccg
DKD PhaJ F	aatcatccggctcgtataatgtgtggaattgtgagcggataacgagctcaggactagaatatacaactaaggaggtaacaaa tgagcgcgcagagcctg
DKD PhaJ R	agtgcggacatagaggacctccttaaatttttttctagattacggcagtttgaccacggc
DKD Hadh F	tcaaactgccgtaatctagaaaaaaatttaaggaggtcctctatgtccgcactgcgctcc
DKD Hadh R	tttatttcctaaattgtatattgtatatatctctggatccttacgccggcatacgcagag
DKD Ech F	agagatatatacaatatacaatttaggaaataaaatgacggttaaactgatggcgtccac
DKD Ech R	aatcttctctcatccgccaaaacagccaagcttttaggaggaggtccatttaggggtacg
pTrc33-PhaB Fwd	ttgtgagcggataacaatttcacacgaattcaaggagatatataatgacccagcgcatcgcttac
pTrc33-Pha Rev	caggctgaaaatcttctctcatccgccaaaacagccttagcccatgtgcaggccacc

C. gBlock assembly. Regions of overlap for Gibson assembly are highlighted in gray.

C1. gBlock assembly map for synthetic Acat1 (GenBank: ADY43117.1) construction

CATCATCATC ATCATCATCA TCATCATCAC AGCAGCGGCC ATATCGAAGG TCGTCATATG
CTGGTGAAGT TCGCGTCCGG TCGTAGCGCA CATTGTAGCA GCACGACGAT GGTCACGTTG
CGTGCCTTCT GGAACGTTCT GCCGAAGCGC TATTTTAGCG CATCTTGCCG TGCGGCACGT
AGCTTCTCTG ACGTGGTGAT TGTCGGTGCG GCGCGTACCC CGCTGGGCTC TTTCCGTAGC
GCCTTCAATA AGGTCCCGGT TACCGTCCTG GGCGCTGCTG CTATCAAAGG TGCGCTGAAA
AACGCCAACC TGAATCCGTC TACGGTGCAG GAAGTGTTTA TGGGTTGTGT CGTGCCGTCC
GGTGCTGGCC AGGCCCCGGC GCGTCAGGCG GTCCTGGCAG CAGGTTGTAA TGTGAGCACG
ATCGTCACCG CTGTCAACAA GGTTTGTGCG TCCGGTATGA AAAGCATTGC ATGTGCAGCG
AGCTTGTTGC AGTTGGACCT GCAAGAAGTT ACGATGGGTG GTGGTATGGA GAGCATGAGC
ATGGTGCCGT ATTATCTGCC GCGCGGTGAC ATTCCGTACG GCGGTATCCA ACTGTTGGAC
GGTATTGCGA AAGACGGTCT GACCGATGCG TACACCCAGG AAGCGATGGG TTGCTTTGCA

GATAAGATTG CGGCAAATTT CGGTATTACG CGCGAGGAAC AGGACAAGTA CGCGATCGAA
TCTTACAAAA AGGCTGCCGC AGCATGGGAA AACGGCGCGT TCAAGGACGA GATTACCCCG
GTTGAGATTA CCATCGGCAA GAAGAAGATG ATTATCGACA AAGATGAAGA GTACACCCGT
GTTAACTTTG AGAAAATCAG CAAGCTGCGT GCGGTTTTTT CCAAGGACGG CACCGTGACG
GCGGGCAACG CGAGCACGCT GAATGATGGC GCCGCAGCGG TTGTCTTGAT GACCTCCGAT
GGTGCGAAAA AGCACGGTGT GAAACCGCTG GCCCGTATCT TGGCGTACGG TGATGCAGCG
ACCAACCCGA GCGACTTCTG TATTGCCCCG GCACTGGTGA TCCCGAAAGT GCTGAGCCTG
GCGAACCTGA AGACCAGCGA TATTGATCTG TGGGAAATCA ATGAGGCGTT CAGCATGGTC
CCGCTGCACA GCATTAAGGC ATTGAATATT GACCCGAGCA AGGTCAATAT CCACGGTGGT
GGTGTCAGCA TCGGCCACCC GATTGGCATG AGCGGTGCTC GTATCATCGT GCACCTGGTT
CATACCCTGA AGCCGGGTCA ACGTGGTTGC GCAGCTATCT GCAATGGCGG CGGTGGCGCG
GGCGGTATGA TTATTGAACG CCTGTAAGGA TCCGGCTGCT AACAAAGCCC GAAAGGAAGC
TGAGTTGG

C2. gBlock assembly map for synthetic Acat2 (GenBank: ADY43074.1) construction

CAACAACCTC GGGATCGAGG AAAACCTGTA TTTTCAGGGC AGCCGCCCGA TCACCGACGT
GGTTTTCGTG GGTGCGGCGC GTACCCCGAT TGGTAGCTTC CGCAGCGCCT TTAATAACGT
CCCGGTGACC GTTCTGGGCC GTGAAGCACT GAAAGGTGCG CTGAAAAATG CCAACGTTAA
GCCGAGCCTG GTGCAGGAGG CGTTCATCGG TGTTGTCGTT CCGAGCAATG CAGGCCAGGG
TCCGGCCCGT CAAGTTGTCC TGGGTGCCGG TTGCGACGTT AGCACGGTTG TCACCGCCGT
CAATAAGATG TGCGCAAGCG GTATGAAGGC GATTGCGTGC GCAGCGTCTA TTCTGCAACT
GGACCTGCAA GAAATGGTTG TTGCAGGCGG TATGGAGAGC ATGAGCTGTG TCCCGTTCTA
CTTGCCGCGC GGTGAGATTC CGTTCGGTGG CACGAAACTG ATCGACGGCA TCCCGCGTGA
CGGCCTGAAC GATGTTTACA ACGACATTCT GATGGGTGCG TGCGCGGACA AAGTTGCTAA
GCAATTTGCA ATTACCCGCG AAGAGCAGGA CAAGTACGCG ATCCTGAGCT ACAAGCGTTC
TGCGGCGGCT TGGAAAGAGG GTATCTTTGC CAAAGAGATT ATCCCGCTGG AAGTGACCCA
GGGTAAGAAA ACCATTACCG TGGAAGAAGA TGAAGAGTAC AAGAAAGTTA ATTTTGAGAA
AATCCCGAAG TTGAAACCGG CGTTTACCAG CGAAGGCAGC GTTACGGCAG CCAACGCAAG
CACCCTGAAT GATGGTGCTG CAATGGTGGT GATGACGACC GTGGATGGCG CGAAAAAGCA
TGGCTTGAAG CCGCTGGCGC GTATGCTGGC GTATGGCGAT GCCGCTACGC ACCCGATCGA
CTTTGGTATC GCGCCGGCGA GCGTCATCCC GAAGGTGCTG AAACTGGCGG GTTTGCAGAT
CAAAGACATT GACCTGTGGG AAATTAACGA GGCCTTCGCC GTTGTTCCGC TGTATACCAT
GAAAACGCTG GGTTTGGATG AAAGCAAAGT CAACATTCAC GGTGGCGCAG TGAGCCTGGG
TCATCCGATC GGTATGAGCG GTGCGCGTAT CGTTGGTCAC CTGGTGCATA CCCTGAAGCC
GGGTCAGAAG GGTTGCGCAG CGATTTGTAA CGGTGGCGGC GGTGCAGGTG GTATGATTAT
TGAGAAGCTG TAAGGATCCC CGAATTCGAG CTCCGTCGAC AAGCTTGCGG CCG

C3. gBlock assembly map for synthetic Acat3 (GenBank: ADY44833.1) construction

CAACAACCTC GGGATCGAGG AAAACCTGTA TTTTCAGGGC GCGAGCCAGT TCACCGACGT
TGTGTTTGTG GGTGCGGCGC GTACCCCGGT TGGCTCTGTT CGTTCCTCCC TGAGCACCGT
GCCGGCGACG GTCCTGGGTG CCGAAGCGAT CAAGGGCGCA CTGAAGCACG CAAATCTGAA
ACCGAGCCAG GTCCAGGAAG TCTTTTTCGG TTGCGTGGTT CCGAGCAATT GTGGCCAGGG
TCCGGCTCGC CAAGCGACGC TGGGTGCTGG CTGCGATCCG AGCACGATTG TGACGACCCT
GAATAAGCTG TGTGCAAGCG GTATGAAAAG CATCGCTTGC GCGGCGAGCC TGTTGCAATT
GGGTCTGCAA GAGGTGACCG TTGGTGGTGG TATGGAGAGC ATGAGCCTGG TCCCGTACTA
TCTGGAGCGC GGTGAAACGA CCTACGGTGG TATGAAACTG ATTGATGGTA TCCCGCGTGA
CGGTTTGACC GATGCGTACA GCAACCAACT GATGGGTGCG TGTGCGGATA ATGTCGCGAA
ACGTTTTAAC ATCACCCGCG AAGAACAAGA CAAGTTCGCA ATTGAGAGCT ACAAGCGCAG
CGCGGCAGCA TGGGAAAGCG GCGCGTGCAA GGCGGAGGTG GTGCCGATCG AAGTGACCAA
AGGTAAAAAA ACGTATATCG TTGATAAAGA TGAGGAATAC ACCAAGGTGA ATTTCGAGAA

ACTGCCGAAA CTGAAACCGG CATTTCTGAA GGATGGTACG ATTACCGCTG GTAACGCGAG
CACGCTGAAC GACGGTGCGG CAGCGGTCGT TATGACCACG GTGGAGGGTG CAAAAAAATA
CGGTGTTAAG CCGCTGGCGC GTCTGCTGAG CTATGGTGAC GCCGCAACGA ACCCGGTGGA
TTTCGCGATT GCTCCGAGCA TGGTTATTCC GAAGGTGTTG AAGCTGGCGA ATCTGGAGAT
TAAAGACATT GATCTGTGGG AGATCAACGA AGCATTCGCC GTCGTTCCGC TGCACAGCAT
GAAGACCCTG GGTATCGATC ATAGCAAAGT CAATATCCAC GGTGGTGGTG TGTCCCTGGG
CCACCCGATC GGCATGAGCG GTGCTCGTAT CATCGTTCAC CTGATCCACG CATTGAAACC
GGGTCAAAAG GGCTGTGCGG CAATTTGCAA TGGCGGCGGC GGTGCAGGCG GTATGGTGAT
TGAGAAATTG TAAGGATCCC CGAATTCGAG CTCCGTCGAC AAGCTTGCGG CCG

C4. gBlock assembly map for synthetic Acat4 (GenBank: ADY43579.1) construction

CAACCTCGGG ATCGAGGAAA ACCTGTATTT TCAGGGCAGC GCAAATATTT TTAAGGATGT
TGTGATTTGT GGCGGTATGC GTACCCCGAT TGCGAGCTTC CGTAGCAAGC TGAACAGCGT
TCCGGTTACG GAACTGGGTA GCACGGCTAT TTTCGCAACC CTGGAACACG CTGGTATCAA
ACCGTCCCTG GTCCAGGAAG CGTTCGTGGG CGTGGTTCTG CCGGCGGACG CCGGTCAGGC
CCCGGCACGC CAAGCTGTTC TGGGTGCCGG TCTGAATGTT TCCACCATCG TGACGGCGGT
TAATAAAACC AGCGCGAGCG GTATGAAGAG CATTATGCTG GCAGCAGAGC ATCTGCAACT
GGGTCTGCAA GATTTCGCTA TTGGTGCCGG TATGGAAAAT ATGTCTCGTG TGCCGTTTTT
TCTGAAACGT GGCGACACGC CGTATGGTGG CATTCATTTG GCGGACGGCG TTCTGCGTGA
TGGTCTGATC GACGTGTACG GTAATCTGCA CCTGGGTGGC TGCACCGATA AGATTGCTAA
GAAGTACGGT GTGACCCGTG AGGAACAGGA CGAATACGCT GCGCAGTCCT ACCGTCGTGC
TGCGGCAGCT TGGGAAAGCG GTGTGATGGC GAAAGAGGTT GTGCCGGTCG AAATCAAGGA
GGGTAAACGT GAGTATTTGT TTGATATGGA TGAAGAGTTT CAGCGTGTCA ATTACGAACG
TCTGCCGACG CTGAATCCGG CCTTCACGAA GGACGGTACG ATCACGGCAG GTAATGCAAG
CAGCCTGAGC GACGGCGCAG CGGCGGTCTT GCTGGCCCGC ATGAAGGCCG CAGATCAGCA
CCATATCCCG CCGATTGCCA AAATTCTGGC ACTGGCGGAC GCAGCAACCG AACCGGAGGA
GTTTTCCGTC GCCCCGACGC TGGTGATCCC GAAGCTGCTG GAGATTGCAG GTTTGAAGGT
CGATGATATT GATCTGTTTG AGATCAATGA GGCATTTGCC GTTACGCCGA TCCTGGCTAT
TAAGAAGTTT AATCTGGACC CGAACAAAGT TAACGTGCAT GGTGGTGCGA TCAGCCTGGG
TGACCCGGTT GGTATGAGCG GTGCGCGTAT TGTCGTCCAT CTGCTGCACG CACTGAAAAG
CGGTCAGAAA GGCCTGGCAG CGATCTGTAA CGGCGGCGGT GGTGCGTCTG GTATGATTAT
TGAGAAGCTG TAAGGATCCC CGAATTCGAG CTCCGTCGAC AAGCTT

C5. gBlock assembly map for synthetic Acat5 (GenBank: ADY44742.1) construction

CAACCTCGGG ATCGAGGAAA ACCTGTATTT TCAGGGCATG GATGCAGGCA AAAAAGACGT
TTACATCCTG AGCGCAGTTC GTACCCCGAT CGCGAGCTTT CGTAGCACCC TGACCAGCCT
GTCCGCGGTG GACCTGGGTA TTGTGGTCAC CAAAGAAGCG ATTAAGCGTA GCCTGCTGCC
GTCCAGCGCC ATCGAAGAAA CCATCGTCGG CAACGTTCTG TCTGCCGGTC TGGGTCAGAA
CATCGCACGC CAGATTTCCA TTGCCAGCGA GATCCCGAAG AGCAGCCAGT GTGTCACCAT
CAATAAAGTG TGCAGCAGCT CCATGAAAGC AATTATTATG GGTGCTCAGG CGATTCAGGT
GGGCTACCGT CGTATCGTGG TTGCTCTGGG CTCCGAGAGC ATGAGCAATG CACCGTTTTA
TGTGCCGCGT GGCGAGATTC CGTTTGGTGG CGTGCAACTG GTTGATGCCC TGCAACGTGA
TGGTCTGATG GACAGCATTG AGTACCAGCC GATGGGTCTG TGCGCAGAAA AAACCGTTAA
GGATTATGCT TTTACCCGCG AACAACTGGA TGCGTATGCG ATCGAGTCTT ATCGTAAGGC
AGAACATGCG TGGAAAGAGG GCGCTTTCAA TAAGGAAGTT GTTCCGGTTA GCGTGCCGCA
AAAGCGTGGT AGCAAGGTGG TCCTGACCGA GGACGAGGAG TACAAGCGTT TGATCCCGGA
AAAGGTTCCG GCTCTGCATC CGGCGTTTCT GAAAGACGGC AGCGGTACGA TCACCGCGGC
CAATGCCAGC ACGATTAATG ACGGTGCCGC GGCCTGCGTC CTGGCCAGCG GCGAGGTGGT
GCAGGAGGGT CGTCTGAAGC CGATCGCCAA AGTGCTGAGC TACGCTGAGG CAGGCGTGGA
ACCGATCGAT TTTACGGTTG CTCCGGCCCT GGCCGTTAAA CAACTGCTGT CCCAGAGCGG

TCTGGATGAA GAGAGCATCG CGCTGTGGGA GATTAACGAG GCGTTTAGCG TCACCGGCCT
GGCGTTCATT AAGGAGCTGC GCCTGGACCC GAAACGTGTG AATGTTCGTG GTGGTGCTGT
CGCGTTGGGC CATCCGTTGG GCGCGAGCGG CGCTCGCATC GTCGTTACGC TGGTGCACGC
GCTGAAAAGC GACGAACTGG GCGTCGCTGC GATTTGCAAC GGTGGCGGCG AGGCTTCCGC
TATTCTGATT AAAAAACTGT AAGGATCCCC GAATTCGAGC TCCGTCGACA AGCTTGCGGC

C6. gBlock assembly map for synthetic AsHadh (GenBank: ADY43626.1) construction

CATCATCATC ACAGCAGCGG CCATATCGAA GGTCGTCATA TGTCCGCACT GCGCTCCACG
AAAGGTCTGG TCGCGCTGGT GACGGGTGGT GCGAGCGGCC TGGGCCGTGG TGCCGCTGAA
AATCTGCTGA AACACGGCGC AAAAGTGGCT ATTCTGGATC TGCCGAGCAG CGCGGGCGCA
GAGGTTGCGA AAGAGCTGGG TGGTGATTGC ATTTTTACCC CGGCAAGCGT CACGGCAGCA
AGCGAAGTTA AGAGCGCCCT GGCGGATGTT AAGAAAAAAT TTGGTCGTTT GGACGTTGCG
GTGAACTGCG CGGGTATTGC GTACAGCTTT AAACTGTTCA ACGTGAAGAA GAAAAAACTG
TGCGATCTGG AAAGCGTGCG TAAAACGCTG GATGTCAACG TCATGGGCTA CTTCACGGTT
GCGGCTCACG CAGCGGAGCT GTTCGCGGAG AATGAGAAAG ACGAGATGGG CCAACGCGGT
GTTATTATTA ATACCGCAAG CATTGCAGCG TTCGACGGTC AAGCCGGCCA AAGCGCGTAT
AGCGCCTCCA AGGGCGCGAT TGTTGGCATG ACCCTGCCGC TGGCCCGCGA CTTCGCCGAC
GATGGCATCC GCGTTGTGAC CATCGCTCCG GGCATCTTTG ACACCCCGAT GATGGCGTCC
TTCCCGGACA AAGTTCGTAA CTTCCTGATT GGTTTGGTGC CGAATCCGAA GCGTTTCGGT
GTGCCGGAGG AATACGGCGC TCTGGTTCGT CATATCATCG AAAACCGCTA CCTGAACGGC
GAGGTTATTC GTCTGGATGG TGCTCTGCGT ATGCCGGCGT AAGGATCCGG CTGCTAACAA
AGCCCGAAAG GAAGCTGAGT TG

C7. gBlock assembly map for synthetic AsEch (GenBank: ADY46836.1) construction

CATCATCATC ACAGCAGCGG CCATATCGAA GGTCGTCATA TGACGGTTAA ACTGATGGCG
TCCACTACGA CTGCTGCCCC GGCAGCGATG GAAATGATTA AAGCCGAAAA AGCGGGCGAA
AAAAACAATG TAGGCCTGAT CCACCTGAAT CGCCCGAAAG CTCTGAACGC GCTGTGTGAT
CAGCTGATGT CCGAACTGAG CATTGCTCTG AAAGAATTCG ATAATGACGA AAGCATCGGT
GCGATCGTAA TCACTGGTAG CGAACGTAGC TTTGCAGCTG GCGCAGACAT CAAAGAAATG
CAGAACAAAC AATTCCCGGA AGTATATATG AACAAATTCC TGGAAAGCTG GTCCGCTGTG
TCTCGTATCA GCAAGCCGAT CATTGCGGCA GTTAACGGTT TCGCCCTGGG CGGCGGCTGC
GAACTGGCTA TGATGTGCGA CATTATTTAC GCGGGTGATA AAGCTCAGTT CGCTCAGCCA
GAGATCAATA TCGGTACCAT CCCGGGTGGC GGTGGTACCC AGCGCTGGCC GCGTGCTGCA
GGTAAATCTC TGGCAATGGA AATCTGTCTG ACCGGCAACC GTATGTCTGC TCAGGAAGCT
AAAGAATGTG GCCTGATTTC TAAAGTTTTC CCTGCTGATC AGGTTGTTAA AGAAGCGATT
AAAACCGCAG AAAAAATCGC GGAACAGAGC CCGGTTATCG TAGCAATGAT GAAAGAAGCG
GTAAACTCTG CTTACGAAAT GACCCTGCAG GAGGGCCTGC GCTTTGAAAA GCGTCTGTTC
CATCAGACCT TCGCAACTAA TGACCGTAAA GAGGGTATGA CCGCATTTGC GGAAAAACGT
ACCCCTAAAT GGACCTCCTC CTAAGGATCC GGCTGCTAAC AAAGCCCGAA AGGAAGCTGA
GTTG

C8. gBlock assembly map for synthetic CnHadh (GenBank: WP_011614598) construction

AGCAGCGGCC TGGTGCCGCG CGGCAGCCAT ATGCAAATTC AGGGCAACGT TTTTATTGTT
ACGGGGGGTG CTAGTGGACT GGGTGCTGGC ACGGCGCGTA TGCTGGCAGA AGCTGGTGCG
CGCGTGGTCA TTGCTGACCT GAATGAGAGC GCCGGTCAGG CTTTAGCCAC CGAATTGGGG
GGTCAATTTG TTCGTTGCGA TGTTACCTCT GAGGCCGACG GTCAGGCAGT TGTTGCGGCG
GCGCAAGCGC TCGGACGCTT ATCAGGCATT GTCAACTGTG CGGGCATCGC AACAGCCAAT
AAAACCGTGG GTAAGAATGG TCCCCATCCA TTAGAAGCTT TTGATAAAAC GATTCGCATT
AACCTCGTAG GCACTTTTAA CATGATCCGG TTAGCCGCTG CAGCGATGGT TCAAAATACC

CCCGACAGTG AAGGCGAACG CGGCGTCATT ATCAACACGG CCTCCGTTGC GGCCTTCGAC
GGGCAGATTG GGCAAGCCGC TTATGCCGCC AGCAAAGGTG GTGTAGTCGC CATGACTCTG
GCCATCGCAC GTGATCTGAG CCGCGACGGC GTGCGCTGTA TGACAATTGC ACCAGGGATT
TTTGAAACCC CGATGCTCCT GGGAATGCCG CAGGAAGTCC AGGATGCCCT CGGGCGCATG
GTCCCATTCC CGCCGCGTCT TGGCCGCCCC GCCGAATATG CGAAACTGGC GCGTAGCATT
ATCGAAAACA CGATGCTGAA TGGAGAAGTG ATCCGCCTGG ACGGTGCCAT TCGTATGCAG
CCAAAATAAG CTTAAGAATT CGAGCTCCGT CGACAAGCTT GCGGC

D. Scheme for construction of RBS/promoter screen.

Plasmid	*Base vector*	*Acat3 F. primer*	*Acat3 R. primer*	*Reductase*	*Reductase F. primer*	*Reductase R. primer*
pTrc99-A3AsH00	pTrc99a	Acat3 F0	Acat3 AsHR0	AsHadh	AsHF0	Hadh RP
pCWori-A3AsH00	pCWori	Acat3 F0	Acat3 AsHR0	AsHadh	AsHF0	AsHadh CWR
pTrc99-A3AsH01	pTrc99a	Acat3 F0	Acat3 AsHR1	AsHadh	AsHF1	Hadh RP
pCWori-A3AsH01	pCWori	Acat3 F0	Acat3 AsHR1	AsHadh	AsHF1	AsHadh CWR
pTrc99-A3AsH10	pTrc99a	Acat3 F1	Acat3 AsHR0	AsHadh	AsHF0	Hadh RP
pCWori-A3AsH10	pCWori	Acat3 F1	Acat3 AsHR0	AsHadh	AsHF0	AsHadh CWR
pTrc99-A3AsH11	pTrc99a	Acat3 F1	Acat3 AsHR1	AsHadh	AsHF1	Hadh RP
pCWori-A3AsH11	pCWori	Acat3 F1	Acat3 AsHR1	AsHadh	AsHF1	AsHadh CWR
pTrc99-A3CnH00	pTrc99a	Acat3 F0	Acat3 CnHR0	CnHadh	CnHF0	CnHadh 99R
pCWori-A3CnH00	pCWori	Acat3 F0	Acat3 CnHR0	CnHadh	CnHF0	CnHadh CWR
pTrc99-A3CnH01	pTrc99a	Acat3 F0	Acat3 CnHR1	CnHadh	CnHF1	CnHadh 99R
pCWori-A3CnH01	pCWori	Acat3 F0	Acat3 CnHR1	CnHadh	CnHF1	CnHadh CWR
pTrc99-A3CnH10	pTrc99a	Acat3 F1	Acat3 CnHR0	CnHadh	CnHF0	CnHadh 99R
pCWori-A3CnH10	pCWori	Acat3 F1	Acat3 CnHR0	CnHadh	CnHF0	CnHadh CWR
pTrc99-A3CnH11	pTrc99a	Acat3 F1	Acat3 CnHR1	CnHadh	CnHF1	CnHadh 99R
pCWori-A3CnH11	pCWori	Acat3 F1	Acat3 CnHR1	CnHadh	CnHF1	CnHadh CWR

E. Scheme for construction of PHA production vectors.

Plasmid	*Condensing enzyme*	*Acat/NphT7 F. primer*	*Acat/NphT7 R. primer*	*PhaC*	*PhaC F. primer*	*PhaC R. primer*
Poly1	Acat2	Poly1,3 Fwd	Poly1 R	*R. ruber*	Poly1 F	Poly1,2 Rev
Poly2	Acat3	Poly2,4 Fwd	Poly2 R	*R. ruber*	Poly2 F	Poly1,2 Rev
Poly3	Acat2	Poly1,3 Fwd	Poly3 R	*R. opacus*	Poly3 F	Poly3,4 Rev
Poly4	Acat3	Poly2,4 Fwd	Poly4 R	*R. opacus*	Poly4 F	Poly3,4 Rev
Poly1B	Acat2	Poly1B Fwd	Poly1B R	*R. ruber*	Poly1 F	Poly1,2 Rev
Poly7B	NphT7	Poly7,8 Fwd	Poly7B R	*R. ruber*	Poly7B F	Poly1,2 Rev
Poly8B	NphT7	Poly7,8 Fwd	Poly8B R	*R. opacus*	Poly8B F	Poly3,4 Rev

F. Amino acid sequences of purified proteins.

F1. Acat1 after purification and TEV cleavage

GTGRSFSDVVIVGAARTPLGSFRSAFNKVPVTVLGAAAIKGALKNANLNPSTVQEVFMGCVVPSGAGQ
APARQAVLAAGCNVSTIVTAVNKVCASGMKSIACAASLLQLDLQEVTMGGGMESMSMVPYYLPRGDIP
YGGIQLLDGIAKDGLTDAYTQEAMGCFADKIAANFGITREEQDKYAIESYKKAAAAWENGAFKDEITP
VEITIGKKKMIIDKDEEYTRVNFEKISKLRAVFSKDGTVTAGNASTLNDGAAAVVLMTSDGAKKHGVK
PLARILAYGDAATNPSDFCIAPALVIPKVLSLANLKTSDIDLWEINEAFSMVPLHSIKALNIDPSKVN
IHGGGVSIGHPIGMSGARIIVHLVHTLKPGQRGCAAICNGGGGAGGMIIERL

F2. Acat2 after purification and TEV cleavage

GTGSRPITDVVFVGAARTPIGSFRSAFNNVPVTVLGREALKGALKNANVKPSLVQEAFIGVVVPSNAG
QGPARQVVLGAGCDVSTVVTAVNKMCASGMKAIACAASILQLDLQEMVVAGGMESMSCVPFYLPRGEI
PFGGTKLIDGIPRDGLNDVYNDILMGACADKVAKQFAITREEQDKYAILSYKRSAAAWKEGIFAKEII
PLEVTQGKKTITVEEDEEYKKVNFEKIPKLKPAFTSEGSVTAANASTLNDGAAMVVMTTVDGAKKHGL
KPLARMLAYGDAATHPIDFGIAPASVIPKVLKLAGLQIKDIDLWEINEAFAVVPLYTMKTLGLDESKV
NIHGGAVSLGHPIGMSGARIVGHLVHTLKPGQKGCAAICNGGGGAGGMIIEKL

F3. Acat3 after purification and TEV cleavage

GTGASQFTDVVFVGAARTPVGSVRSSLSTVPATVLGAEAIKGALKHANLKPSQVQEVFFGCVVPSNCG
QGPARQATLGAGCDPSTIVTTLNKLCASGMKSIACAASLLQLGLQEVTVGGGMESMSLVPYYLERGET
TYGGMKLIDGIPRDGLTDAYSNQLMGACADNVAKRFNITREEQDKFAIESYKRSAAAWESGACKAEVV
PIEVTKGKKTYIVDKDEEYTKVNFEKLPKLKPAFLKDGTITAGNASTLNDGAAAVVMTTVEGAKKYGV
KPLARLLSYGDAATNPVDFAIAPSMVIPKVLKLANLEIKDIDLWEINEAFAVVPLHSMKTLGIDHSKV
NIHGGGVSLGHPIGMSGARIIVHLIHALKPGQKGCAAICNGGGGAGGMVIEKL

F4. Acat4 after purification and TEV cleavage

GTGSANIFKDVVICGGMRTPIASFRSKLNSVPVTELGSTAIFATLEHAGIKPSLVQEAFVGVVLPADA
GQAPARQAVLGAGLNVSTIVTAVNKTSASGMKSIMLAAEHLQLGLQDFAIGAGMENMSRVPFFLKRGD
TPYGGIHLADGVLRDGLIDVYGNLHLGGCTDKIAKKYGVTREEQDEYAAQSYRRAAAAWESGVMAKEV
VPVEIKEGKREYLFDMDEEFQRVNYERLPTLNPAFTKDGTITAGNASSLSDGAAAVLLARMKAADQHH
IPPIAKILALADAATEPEEFSVAPTLVIPKLLEIAGLKVDDIDLFEINEAFAVTPILAIKKFNLDPNK
VNVHGGAISLGDPVGMSGARIVVHLLHALKSGQKGLAAICNGGGGASGMIIEKL

F5. Acat5 after purification and TEV cleavage

GTGMDAGKKDVYILSAVRTPIASFRSTLTSLSAVDLGIVVTKEAIKRSLLPSSAIEETIVGNVLSAGL
GQNIARQISIASEIPKSSQCVTINKVCSSSMKAIIMGAQAIQVGYRRIVVALGSESMSNAPFYVPRGE
IPFGGVQLVDALQRDGLMDSIEYQPMGLCAEKTVKDYAFTREQLDAYAIESYRKAEHAWKEGAFNKEV
VPVSVPQKRGSKVVLTEDEEYKRLIPEKVPALHPAFLKDGSGTITAANASTINDGAAACVLASGEVVQ
EGRLKPIAKVLSYAEAGVEPIDFTVAPALAVKQLLSQSGLDEESIALWEINEAFSVTGLAFIKELRLD
PKRVNVRGGAVALGHPLGASGARIVVTLVHALKSDELGVAAICNGGGEASAILIKKL

F6. His_{10}-Hadh after purification

MGHHHHHHHHHHSSGHIEGRHMSALRSTKGLVALVTGGASGLGRGAAENLLKHGAKVAILDLPSSAGA
EVAKELGGDCIFTPASVTAASEVKSALADVKKKFGRLDVAVNCAGIAYSFKLFNVKKKKLCDLESVRK
TLDVNVMGYFTVAAHAAELFAENEKDEMGQRGVIINTASIAAFDGQAGQSAYSASKGAIVGMTLPLAR
DFADDGIRVVTIAPGIFDTPMMASFPDKVRNFLIGLVPNPKRFGVPEEYGALVRHIIENRYLNGEVIR
LDGALRMPA

F7. His_{10}-AsEch after purification

MGHHHHHHHHHHSSGHIEGRHMTVKLMASTTTAAPAAMEMIKAEKAGEKNNVGLIHLNRPKALNALCD
QLMSELSIALKEFDNDESIGAIVITGSERSFAAGADIKEMQNKQFPEVYMNKFLESWSAVSRISKPII
AAVNGFALGGGCELAMMCDIIYAGDKAQFAQPEINIGTIPGGGGTQRWPRAAGKSLAMEICLTGNRMS
AQEAKECGLISKVFPADQVVKEAIKTAEKIAEQSPVIVAMMKEAVNSAYEMTLQEGLRFEKRLFHQTF
ATNDRKEGMTAFAEKRTPKWTSS

F8. His_{10}-TesB after purification

MGHHHHHHHHHHSSGHIEGRHMSQALKNLLTLLNLEKIEEGLFRGQSEDLGLRQVFGGQVVGQALYAA
KETVPEERLVHSFHSYFLRPGDSKKPIIYDVETLRDGNSFSARRVAAIQNGKPIFYMTASFQAPEAGF
EHQKTMPSAPAPDGLPSETQIAQSLAHLLPPVLKDKFICDRPLEVRPVEFHNPLKGHVAEPHRQVWIR
ANGSVPDDLRVHQYLLGYASDLNFLPVALQPHGIGFLEPGIQIATIDHSMWFHRPFNLNEWLLYSVES
TSASSARGFVRGEFYTQDGVLVASTVQEGVMRNHN

F9. His_6-CnHadh after purification

MGSSHHHHHHSSGLVPRGSHMQIQGNVFIVTGGASGLGAGTARMLAEAGARVVIADLNESAGQALATELGGQFVR
CDVTSEADGQAVVAAAQALGRLSGIVNCAGIATANKTVGKNGPHPLEAFDKTIRINLVGTFNMIRLAAAAMVQNT
PDSEGERGVIINTASVAAFDGQIGQAAYAASKGGVVAMTLAIARDLSRDGVRCMTIAPGIFETPMLLGMPQEVQD
ALGRMVPFPPRLGRPAEYAKLARSIIENTMLNGEVIRLDGAIRMQPK

Table S2. Base strains and plasmids (A) and oligonucleotides (B) used in *Chapter 3*.

A. Strains and plasmids

Strain	*Genotype*	*Source*
DH10B	F- endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ-	Invitrogen
BL21(de3)	F^- *omp*T *gal dcm lon hsd*S_B(r_B^- m_B^-) λ(DE3 [*lac*I *lac*UV5-T7 *gene 1 ind*1 *sam*7 *nin*5])	Novagen

Plasmid	*Description*	*Source*
pET23a-AsAcat2	*His_6-TEV-Acat2* (T7), Cb^r, pBR322	*Chapter 2*
pET23a-Acat2-C91S	*His_6-TEV-Acat2(C91S)* (T7), Cb^r, pBR322	This study
pET23a-Acat2-Loop5	*His_6-TEV-Acat2(Loop5)* (T7), Cb^r, pBR322	This study
pET23a-Acat2-Y153I	*His_6-TEV-Acat2(Y153I)* (T7), Cb^r, pBR322	This study
pET23a-AsAcat5	*His_6-TEV-Acat5* (T7), Cb^r, pBR322	*Chapter 2*
pET23a-Acat5-Loop2	*His_6-TEV-Acat5(Loop2)* (T7), Cb^r, pBR322	This study
pET23a-Acat5-I154Y	*His_6-TEV-Acat5(I154Y)* (T7), Cb^r, pBR322	This study

B. Oligonucleotide sequences

Name	*Sequence*
AsAcat2 23a F1	atcatgagaatctctacttccagggtaccggcagccgcccgatcaccgac
AsAcat2 23a R1	tctcagtggtggtggtggtggtgctcgattacagcttctcaataatcataccacctgcac
AsAcat5 23a F1	gagaatctctacttccagggtaccggcatggatgcaggcaaaaaagacgtttacatcc
AsAcat5 23a R1	ctcagtggtggtggtggtggtgctcgattacagtttttttaatcagaatagcggaagcctc
Acat2-C91S FP	gccgtcaataagatgtctgcaagcggtatgaag
Acat2-C91S RP	cttcataccgcttgcagacatcttattgacggc
Acat2-Loop5 FP	gacggcctgaacgatagcattgagtaccagccgatgggtgcgtgcgcg
Acat2-Loop5 RP	cgcgcacgcacccatcggctggtactcaatgctatcgttcaggccgtc
Acat2-Y153I FP	ggcctgaacgatgttattaacgacattctgatg
Acat2-Y153I RP	catcagaatgtcgttaataacatcgttcaggcc
Acat5-I154Y FP	ggtctgatggacagctatgagtaccagccgatg
Acat5-I154Y RP	catcggctggtactcatagctgtccatcagacc
Acat5-Loop2 FP	gatggtctgatggacgtttacaacgacattctgatgggtctgtgcgcag
Acat5-Loop2 RP	ctgcgcacagacccatcagaatgtcgttgtaaacgtccatcagaccatc

Table S3. Base *E. coli* strains and plasmids (A) and oligonucleotides (B) used in *Chapter 4*.

A. *E. coli* strains and plasmids

Strain	*Genotype*	*Source*
DH10B	F- endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ-	Invitrogen
BL21(de3)	F^- *omp*T *gal dcm lon hsd*S_B(r_B^- m_B^-) λ(DE3 [*lac*I *lac*UV5-T7 *gene 1 ind*1 *sam*7 *nin*5])	Novagen

Plasmid	*Description*	*Source*
pET16b-EgTer	*His_{10}-Xa-EgTer* (T7), *lacI*, Cb^r, pBR322	*Dr. Brooks Bond-Watts*
pET16b-EgMECR1	*His_{10}-Xa-EgMECR1* (T7), *lacI*, Cb^r, pBR322	This study
pSV272.1-EgMECR2	*His_6-MBP-EgMECR2* (T7), *lacI*, Km^r, pBR322	This study
pET16b-EgMECR3	*His_{10}-Xa-EgMECR3* (T7), *lacI*, Cb^r, pBR322	This study
pET16b-EgMECR4	*His_{10}-Xa-EgMECR4* (T7), *lacI*, Cb^r, pBR322	This study

B. Oligonucleotide sequences

Name	*Sequence*
EgMECR1 F1	catcacagcagcggccatatcgaaggtcgtcatatgtctctgccggcgaagcg
EgMECR1 R1	agcttcctttcgggctttgttagcagccggatccttacagggtgaagagaaccttgccgc
EgMECR2 MBP F	cgggatcgaggaaaacctgtattttcagggcatgaccacgggtcatctcgagtattacac
EgMECR2 MBP R	gcttgtcgacggagctcgaattcggggatccttaatgttcagcgccttgcagggtc
EgMECR3 F1	tcatcacagcagcggccatatcgaaggtcgtcatatgtcgcgcaatgttctgacaaaggg
EgMECR3 R1	cttcctttcgggctttgttagcagccggatccttaacttgcaaactcggacaccatcttc
EgMECR4 F1	atcacagcagcggccatatcgaaggtcgtcatatggcccttgccaaaggctg
EgMECR4 R1	gcttcctttcgggctttgttagcagccggatccttagctgagggtcaggatgtgcac
PE Adapter 1[a,b]	P-gatcggaagagcggttcagcaggaatgccgag
PE Adapter 2[b]	acactctttccctacacgacgctcttccgatct
PE PCR Primer 1.0[b]	aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatct
PE PCR Primer 2.0[b]	caagcagaagacggcatacgagatcggtctcggcattcctgctgaaccgctcttccgatct

[a] Adapter was phosphorylated at 5′ position and HPLC purified.

[b]

Figure S1. 2-Methyl-3-oxobutyryl-CoA. (A) Structure. (B) ^{1}H NMR. (C) ^{13}C NMR.

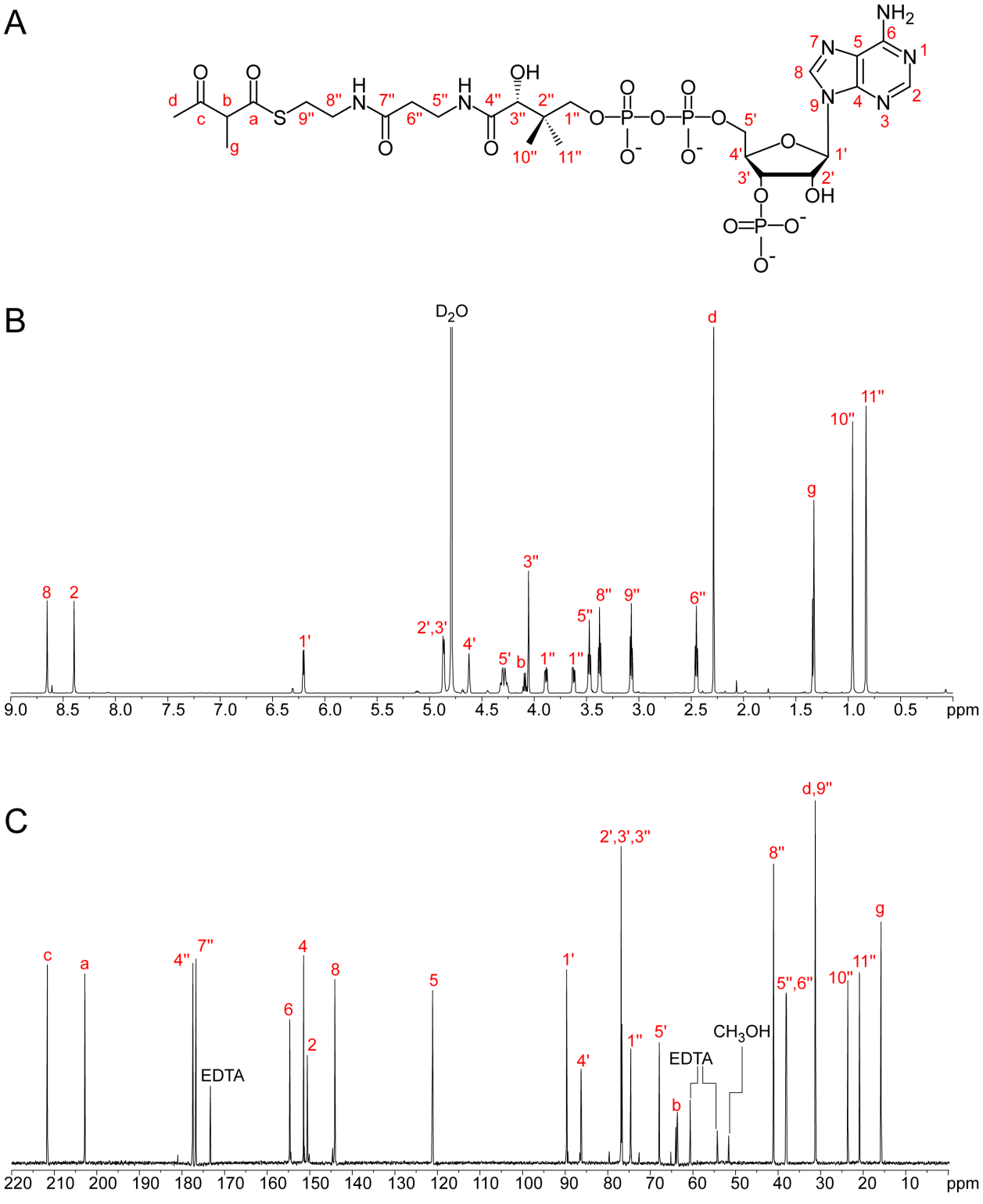


Figure S2. 3-Oxopentanoyl-CoA. (A) Structure. (B) ^{1}H NMR. (C) ^{13}C NMR.

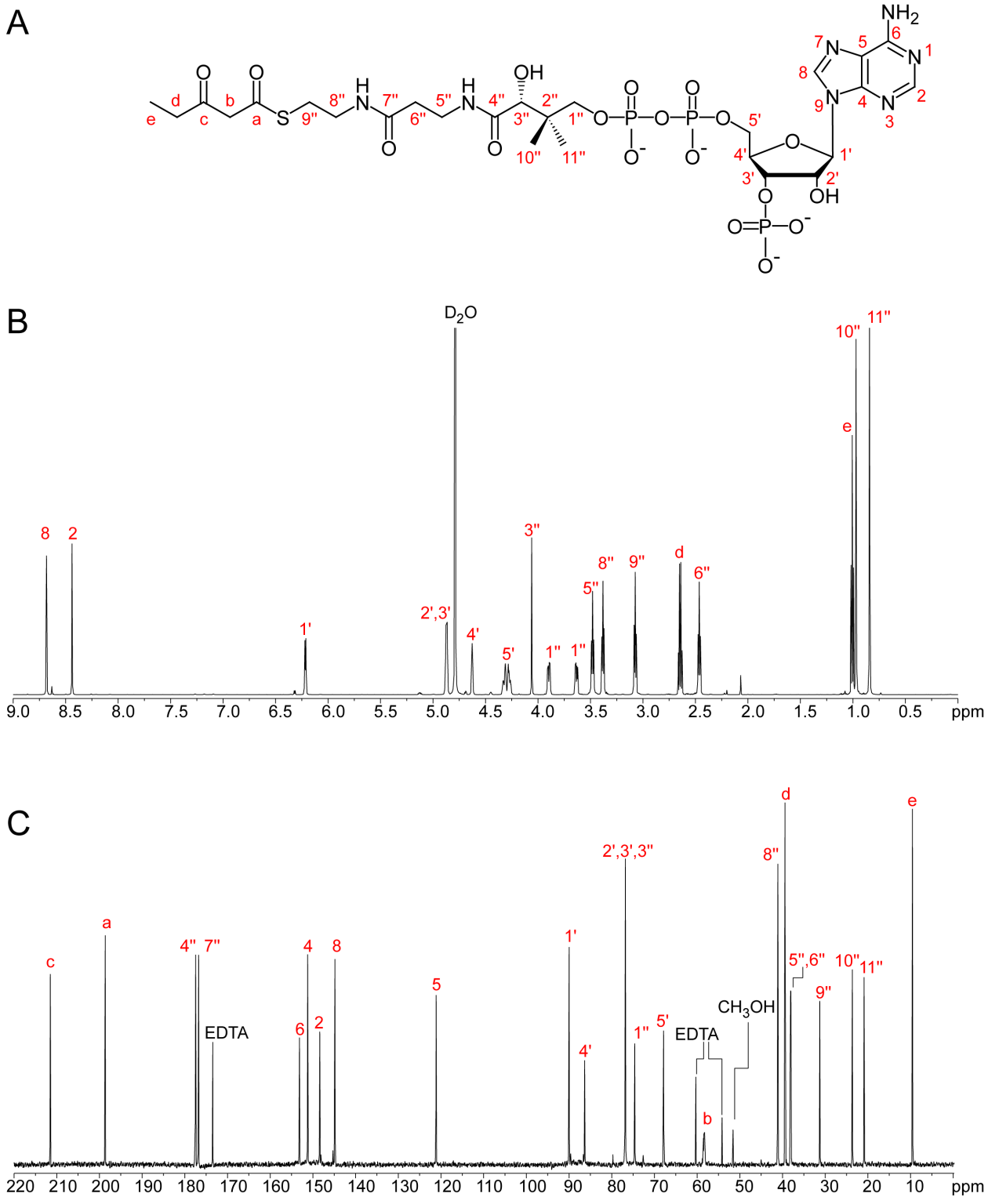


Figure S3. 2-Methyl-3-oxopentanoyl-CoA. (A) Structure. (B) ^{1}H NMR. (C) ^{13}C NMR.

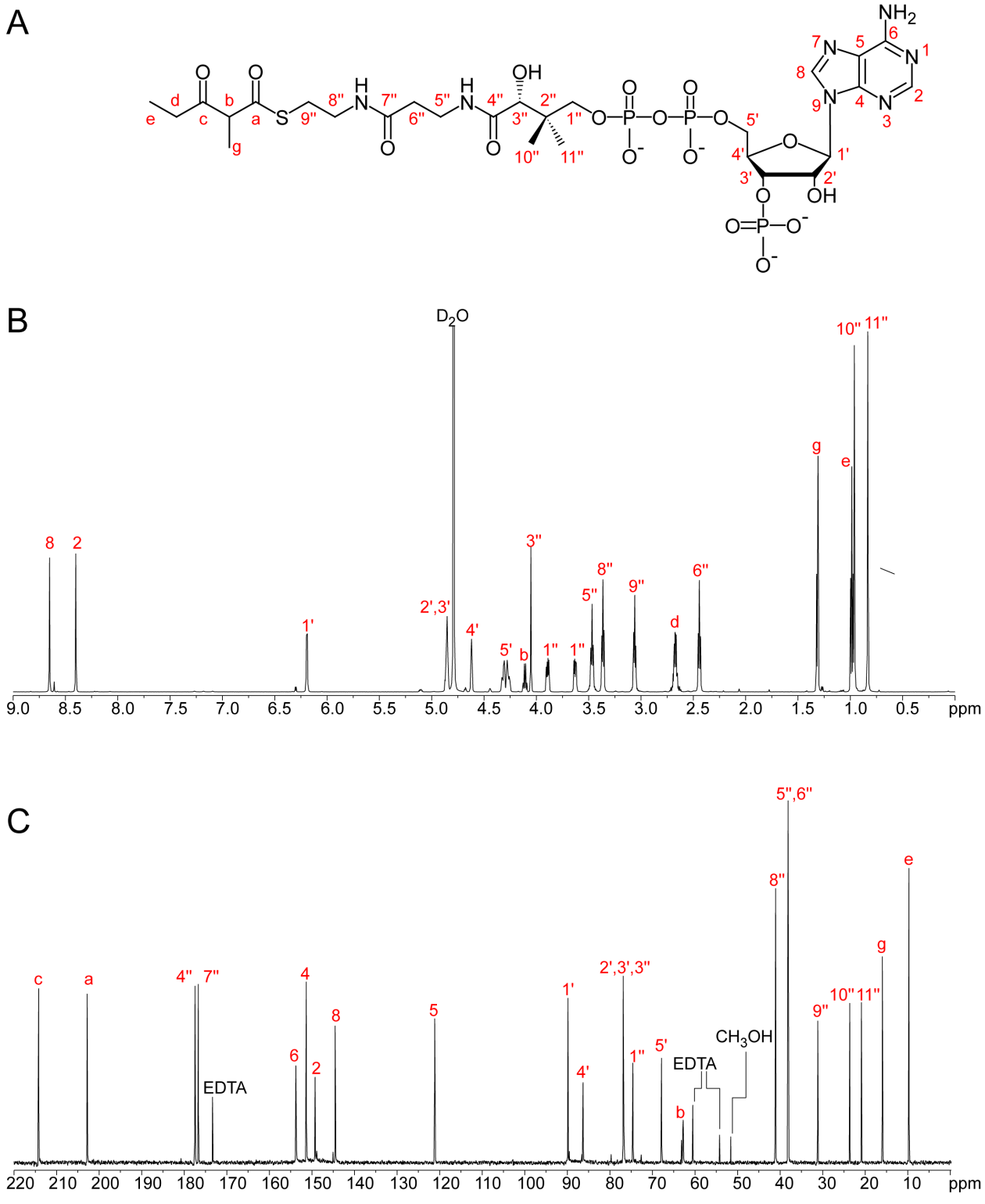


Figure S4. 3-Oxohexanoyl-CoA. (A) Structure. (B) ^{1}H NMR. (C) ^{13}C NMR.

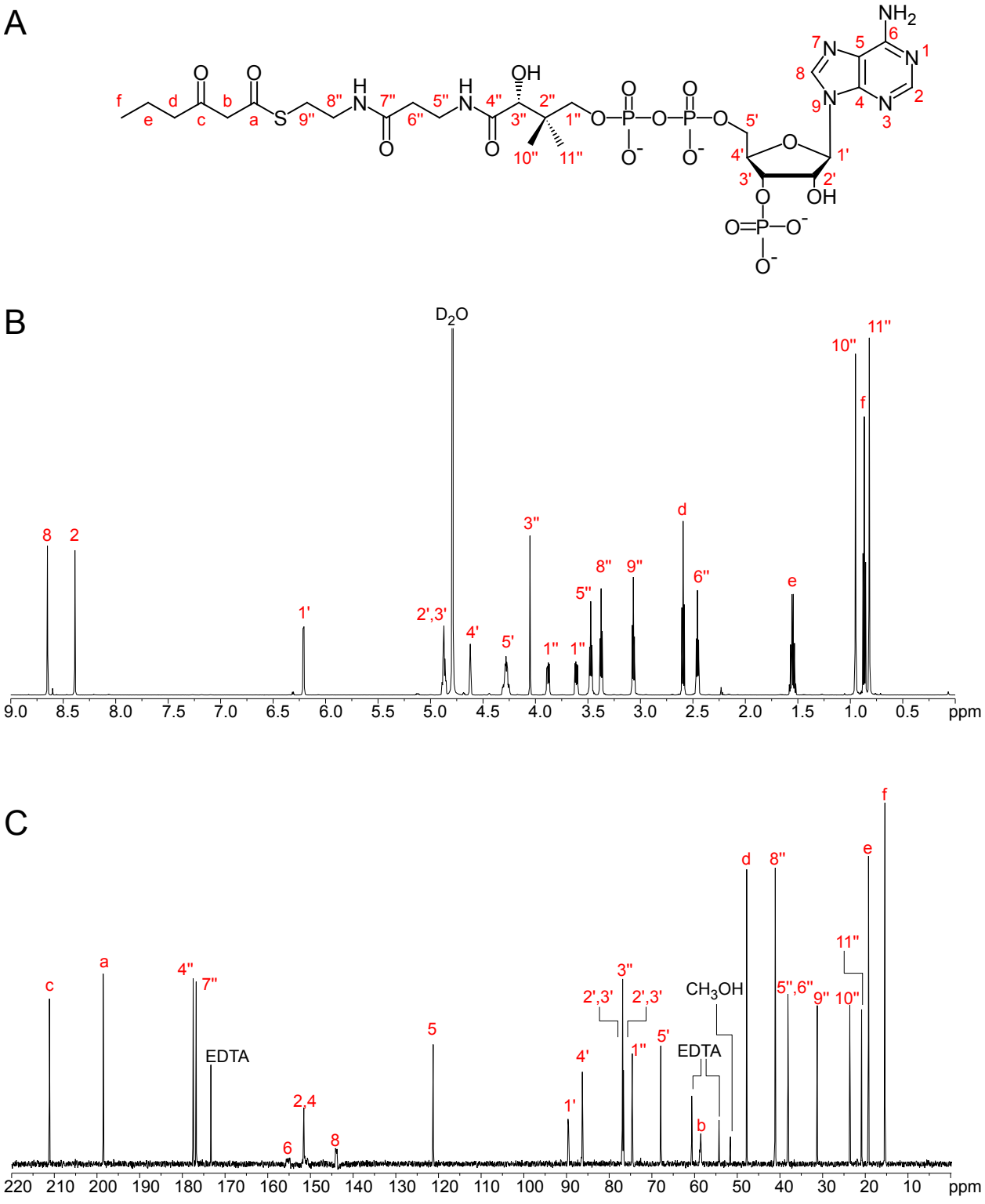


Figure S5. 2-Methyl-3-oxohexanoyl-CoA. (A) Structure. (B) ^{1}H NMR. (C) ^{13}C NMR.

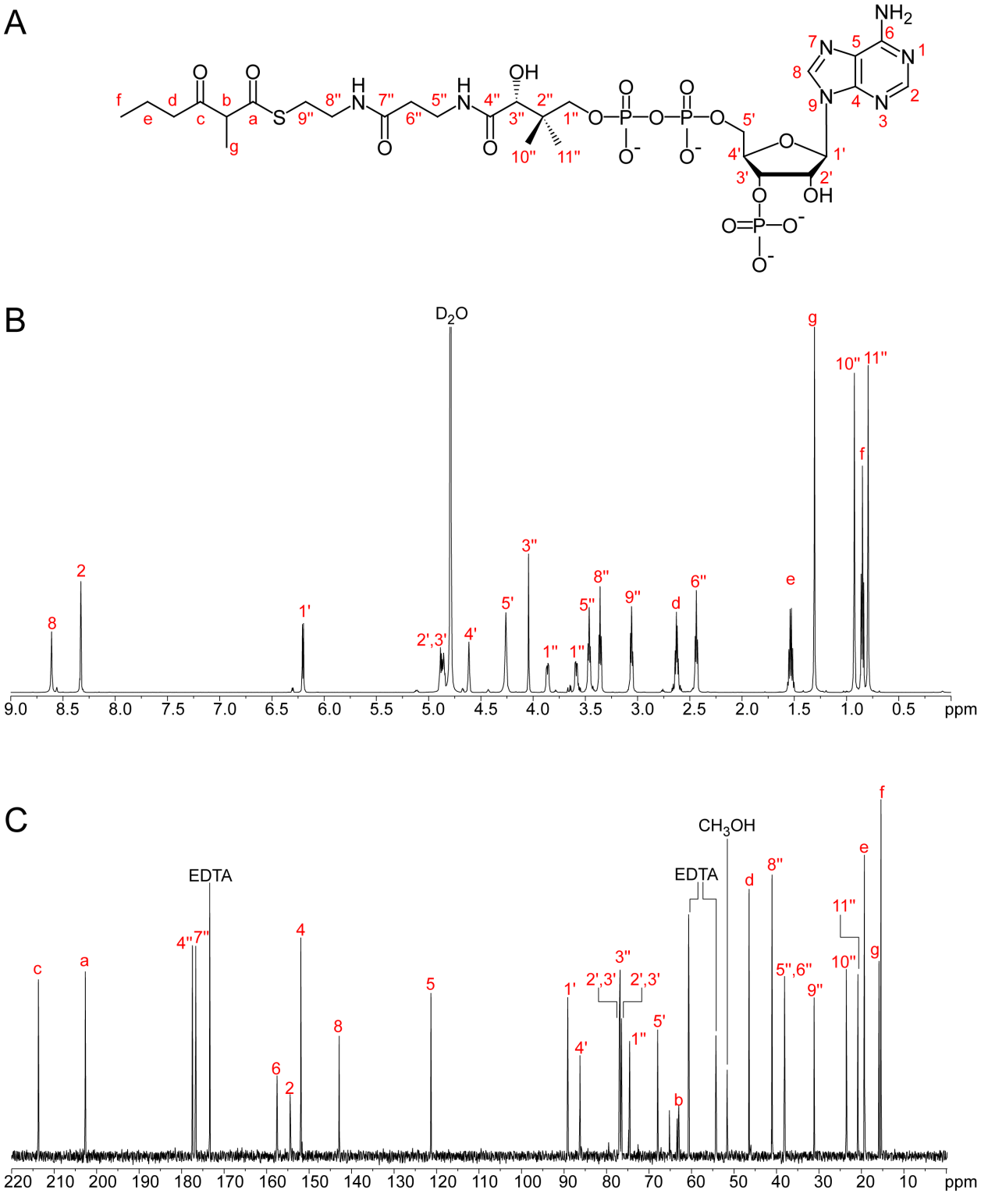


Figure S6. 2-Ethyl-3-oxohexanoyl-CoA. (A) Structure. (B) ^{1}H NMR. (C) ^{13}C NMR.

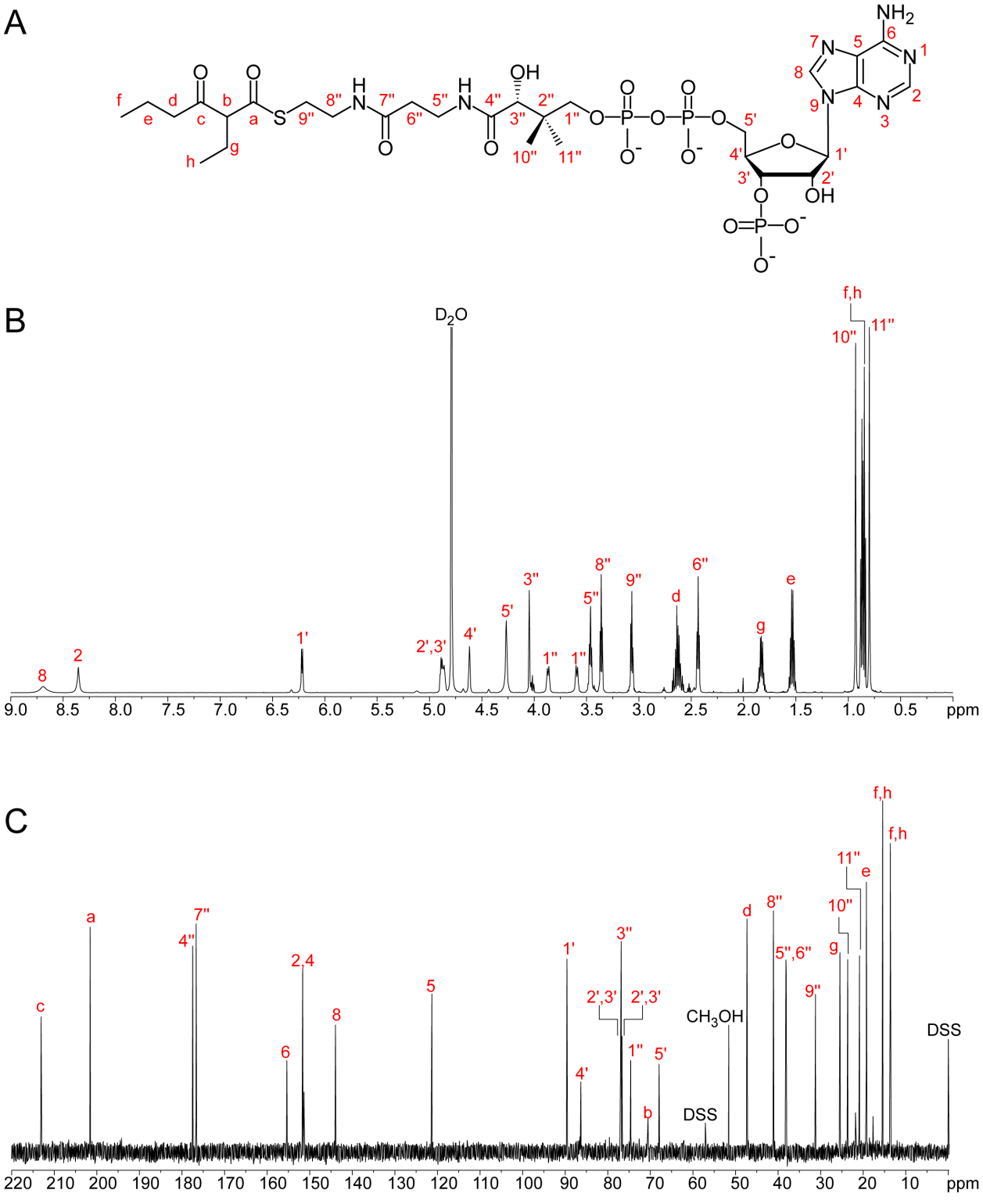
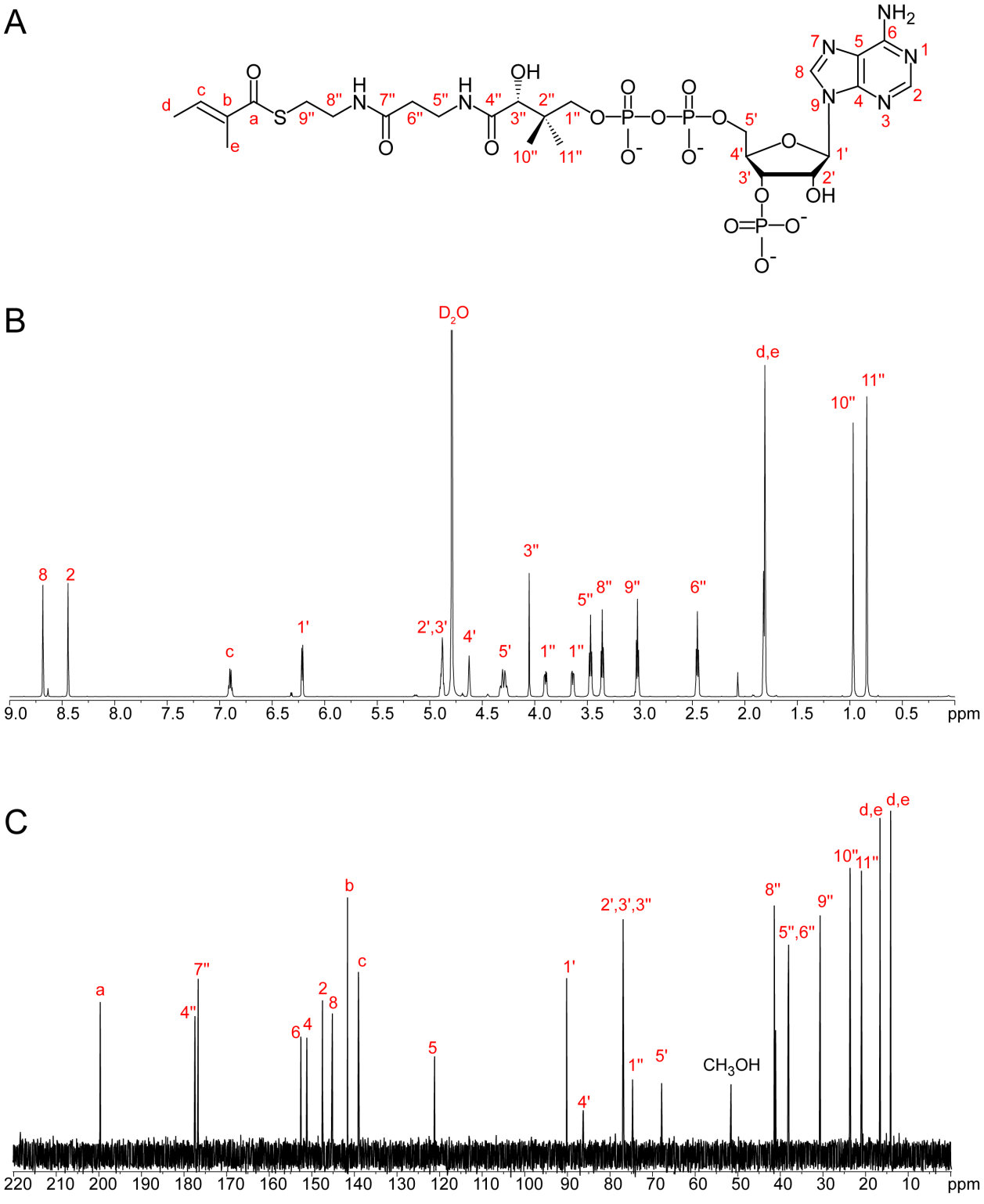


Figure S7. Tigloyl-CoA. (A) Structure. (B) ^{1}H NMR. (C) ^{13}C NMR.



References

1. A. Espah Borujeni, A. S. Channarasappa, and H. M. Salis, Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at upstream standby sites. *Nucleic Acids Res.* **2014,** 42, (4), 2646-59.
2. B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. E. Cane, and C. Khosla, Biosynthesis of complex polyketides in a metabolically engineered strain of E. coli. *Science* **2001,** 291, (5509), 1790-1792.
3. B. B. Bond-Watts, R. J. Bellerose, and M. C. Chang, Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nat. Chem. Biol.* **2011,** 7, (4), 222-7.