Engineering Escherichia Coli Fatty Acid Metabolism for the Production of Biofuel Precursors

Engineering *Escherichia coli* fatty acid metabolism for the production of biofuel precursors

A dissertation presented

by

Tyler J. Ford

to

The Division of Medical Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological and Biomedical Sciences

Harvard University

Cambridge, Massachusetts

May 2015

Dissertation Advisor: Professor Pamela A. Silver Tyler J. Ford

Engineering *Escherichia coli* fatty acid metabolism for the production of biofuel precursors

Abstract

Medium chain fatty acids (MCFAs, 6-12 carbons) are potential precursors to biofuels with properties similar to gasoline and diesel fuel but are not native products of *Escherichia coli* fatty acid synthesis. Herein we engineer *E. coli* to produce, metabolize, and activate MCFAs for their future reduction into alcohols and alkanes (potential biofuels). We develop an *E. coli* strain with an octanoate (8-carbon MCFA) producing enzyme (a thioesterase), metabolic knockouts, and the capability to inducibly degrade an essential metabolic enzyme that would otherwise divert carbon flux away from octanoate. We show that this strain can produce octanoate at 12% theoretical yield. To determine limitations on octanoate catabolism that could prevent its conversion into an acyl-CoA thioester activated for later reduction into alcohols and alkanes, we evolve *E. coli* to grow on octanoic acid as sole carbon source. We show that our fastest growing evolved strain contains mutations that enhance the expression of acyl-CoA synthetase FadD. We then directly mutate the *fadD* gene and screen for mutations that enhance growth rate on octanoic acid. *In-vitro* assays show that the mutations we identify increase FadD activity on MCFAs. These results, homology modeling, and further mutagenesis lead us to hypothesize that our mutations enhance FadD activity by aiding product exit. This work develops a technique (inducible degradation of an essential metabolic enzyme) and generates *fadD* mutants that should be useful for the production of medium chain biofuels and other compounds.

Table of Contents

Chapter 1: Introduction

The current state of biofuels production

The United States has a vested interest in developing biofuels technologies that can compete in terms of cost and performance with traditional petroleum. Global warming caused by green house gas (GHG) emissions due to human activity is already beginning to have negative and costly impacts on the Earth's climate [1]. Large reductions in GHG emissions are required to prevent further drastic and dangerous alterations to the climate [1]. Energy use in the US transportation sector accounts for 27% of all US GHG emissions [2], and is second only to electricity among the sources of US GHG emissions. Motor gasoline accounts for more than 50% of GHG emissions in the transportation sector [3]. Fortunately, gasoline can be replaced with fuels produced by the microbial conversion of plant biomass into liquid hydrocarbons thereby reducing GHG emissions from the transportation sector. The most widely used gasoline replacement produced this way is ethanol. However, ethanol is not an optimal replacement for gasoline due to its current modes of production, distribution, and potential for use in conventional combustion engines. In the work that follows, we demonstrate new ways of altering *Escherichia coli* fatty acid metabolism to produce precursors to biofuels better suited to replace gasoline.

Ethanol produced from cornstarch is the US' most widely consumed biofuel (Figure 1.1). In the ethanol production process, farmers based largely in the northwest of the US transport corn from the farm to the ethanol mill where it is enzymatically hydrolyzed to glucose monomers, fermented by yeast, separated from the culture media, stored, and shipped. Only the starchy component of the corn (the corn kernel) is used for ethanol production. The corn stalk and leaves (corn stover) are usually left in the field, while the proteinacious components of the corn are separated out after fertilization, dried, and sold as animal feed (distillers dried grains

with solubles, DDGS). Most ethanol mills are dry mills dedicated to ethanol production, but more expensive mills, wet mills, can be used to produce additional co-products. Ethanol is produced from corn-starch at roughly 95% of its theoretical yield resulting in the production of about 2.8 gallons of ethanol per bushel of corn [4,5].

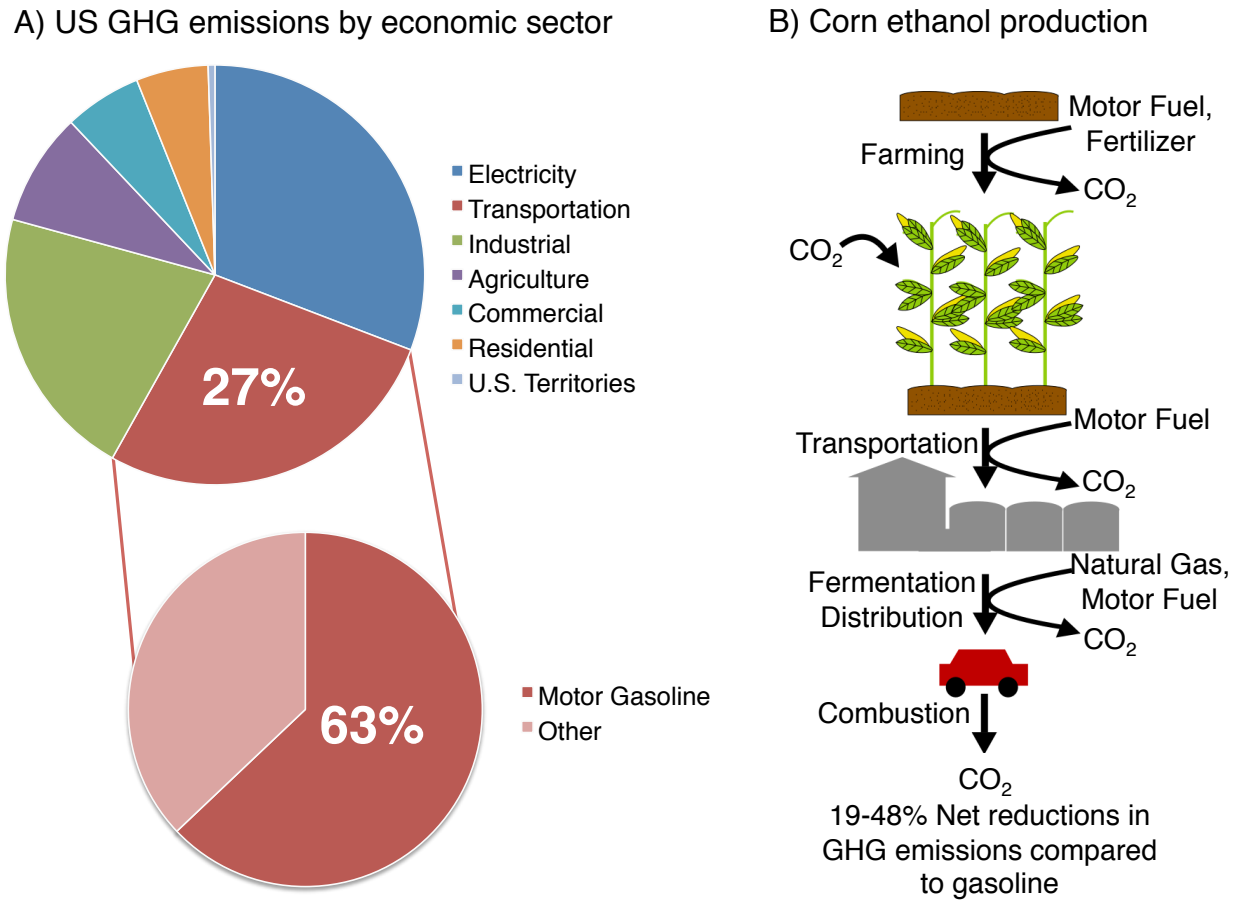

Figure 1.1. Potential for gasoline and ethanol replacement with advanced biofuels. A) United States GHG emissions by economic sector (Data from [2] and [3]). B) Ethanol production process with GHG sources and sinks indicated (Data from [6]).

The biological conversion process for cornstarch ethanol is efficient, but is not without issues. For instance, despite the fact that cornstarch ethanol is touted as a greener alternative to gasoline, there has been much debate over the actual level of GHG emissions reductions achieved from replacing gasoline with cornstarch ethanol. A recent report places GHG reductions in the range of 19-48% [6]. Taking a broad view of cornstarch ethanol production, it

seems that it should produce few GHG emissions; as corn grows it fixes CO_2 from the atmosphere to produce all of its component parts. The starchy component of the corn is then converted into ethanol releasing some of this stored up CO_2 in the process, and, later, CO_2 stored in ethanol is released when it is combusted in an automobile. Overall, no more CO_2 is released into the atmosphere than is fixed by the corn in the first place. Further scrutiny reveals that this view is overly simplistic. The cornstarch ethanol production process has many energy inputs: vehicles to till the land, vehicles to transport the corn, vehicles to transport the ethanol, fertilizer to provide the corn with nitrogen and other nutrients, and electricity and natural gas to run the ethanol mill. The energy for these processes is usually provided by conventional fossil fuels that, of course, release GHGs. On top of these obvious sources of GHG emissions, converting land into a corn farm can release GHGs stored in the soil and force the relocation of whatever was previously on this land resulting in further indirect GHG emissions. The precise magnitude of the GHG emissions caused by direct and indirect land use changes is not well understood but can drastically alter estimations of corn ethanol GHG mitigation potential [6,7]. Wang *et al.* 2012 estimate that the biggest single contributor to GHG emissions in the cornstarch ethanol production process is natural gas used to run the corn ethanol mill [6] which could potentially be reduced if alternative energy sources are used in the production process.

Ethanol produced from sugarcane and lignocellulosic biomass has much greater GHG benefits relative to that produced from cornstarch. Sugarcane is grown for ethanol production in Brazil and can reduce GHG emissions compared to gasoline by 40-62%. Lignocellulosic biomass feedstocks come from the non-edible components of many different plants (including corn stover), can be produced in many different climates, and can reduce GHG emissions compared to gasoline by 77-115% [6,8]. Ethanol production using these biomass sources has the

benefit that their byproducts can be used to produce electricity to run the ethanol plant and displace other non-renewable electricity sources like natural gas. Unfortunately, the US lacks the appropriate climate for sugarcane production and cellulosic biomass conversion processes are in their infancy [6].

In recognition of cornstarch ethanol's limited potential for GHG reductions, the US Renewable Fuel Standard (RFS), which mandates blending increasing volumes of renewable fuels (stratified by their GHG mitigation potential) into the US fuel supply through 2022, caps the total volume of cornstarch ethanol that can be used to meet the standard to 15 billion gallons [9]. With over 10 billion gallons of cornstarch ethanol blended into the US gasoline supply in 2014, and the total renewable fuel mandate at 20.5 billion gallons for 2015, the US Environmental Protection Agency (EPA) clearly expects production of cellulosic biofuels to increase (although the EPA can lower cellulosic requirements if there is simply not enough cellulosic ethanol available).

It is unclear whether future lignocellulosic biofuels should be produced biologically or non-biologically. An extensive 2011 report from the National Renewable Energy Laboratory (NREL) calculating the final price of mixed alcohols (but primarily ethanol) produced from woody forest material using a Fischer Trophs-based (non-biological) method estimated that such a biofuel could be produced at $3.11 per the volume with the same energy content as a gallon of gasoline (a gasoline equivalent gallon) [10]. A similar 2013 study on biological production of ethanol from corn stover (a corn-based lignocellulosic biomass source more amenable to biological processing) estimated production could be achieved at $3.27 per gasoline equivalent gallon [11]. Both processes generate their own heat and power and do not require additional fossil fuel consumption to provide electricity. However, the biological process produces more

excess electricity than the non-biological process. These NREL studies did not provide extensive analysis of GHG mitigation potential from each process, but analyses of similar processes by the International Energy Agency showed that, because of higher GHG emissions from producing chemicals required for fertilizer and feedstock pretreatment, the biological conversion process has greater GHG emissions than the non-biological [12]. The biological process is also estimated to use more water than the non-biological process [10,11]. However, the biological conversion process uses corn stover as a feedstock as opposed to forest materials and so may be more compatible with existing corn ethanol mills. These different trade-offs combined with the limited commercial use of either process leaves no clear choice for which of these processes should be further promoted but indicate that both processes should be pursued in parallel. The work done here can be used to improve the biological process.

Beyond problems associated with GHG reductions that may be remedied through the development of lignocellulosic feedstocks, there is a limit on how much ethanol the current transportation fuel infrastructure can handle. According to statistics from the US Energy Information Administration (EIA), ethanol blended into motor gasoline accounted for nearly 10% of all motor gasoline use in 2014 [13]. This is an important benchmark because fueling with ethanol blends higher than 10% could cause drivers to break their warrantees and potentially damage their vehicles [9]. The EPA additionally limits fuel ethanol blends to 15% in most passenger vehicles [9,14]. Flex fuel vehicles, vehicles specifically designed to run on up to 85% ethanol-gasoline blends (E85), are available, but these only account for 17.4 million [15] of the roughly 250 million cars on the road [16]. Additionally, E85 makes up only a small, <1%, fraction of all gasoline consumed in the US [17]. Larger scale adoption of flex fuel vehicles is possible, but realizing their potential will require a drastic expansion in the number of E85

fueling stations in the US (currently 2,576 E85 fueling stations compared to 160,000 gasoline stations [18,19]).

Hydrophilicity and corrosiveness are largely to blame for ethanol's incompatibility with most current vehicles and also make it difficult to transport ethanol in US oil pipelines [20,21]. Ethanol requires its own dedicated pipelines or modified gasoline pipelines. One such pipeline has been constructed in central Florida, and others have been proposed but will require substantial investment [22]. To avoid pipeline issues, most fuel ethanol is currently transported by truck, rail, or barge which themselves emit GHGs. Transport by pipeline could lower GHG emissions relative to these established methods [23,24].

With ethanol's limited potential to replace gasoline without drastic changes to infrastructure and the automobile fleet, there is clearly room to develop alternatives. The ideal biofuel would be hydrophobic, have an energy density similar to that of gasoline, have combustion properties similar to those of gasoline, and be capable of being produced in high yields. Gasoline itself is a complex mixture of medium chain (6-12 carbon) branched and unbranched alkanes, cyclic alkanes, and aromatics [25]. Any biologically produced alternative would likely be mixed with gasoline as a "drop-in" fuel with similar properties.

NREL recently performed a study analyzing the processes and economics behind using bacteria to convert lignocellulose-derived sugars into the 16-carbon fatty acid palmitate. Fatty acids are highly reduced, energy dense biological compounds. Their aliphatic tails are structurally similar to many of the compounds in gasoline and they can be enzymatically reduced into alcohols and alkanes with similar energy density and hydrophobicity to gasoline or diesel fuel [26-32]. In the NREL study, non-biological processing is used to convert palmitate into hexadecane and pentadecane as diesel fuel blend stocks. They assume that palmitate is produced

aerobically and that the sugar to palmitate conversion achieves 79% of its theoretical yield (such efficiencies are difficult to achieve as will be discussed below). NREL estimates that the final product could be sold at $5.10 per gasoline equivalent gallon. While this price is notably higher than those for the routes to ethanol discussed above, the authors point out that increasing the microbes' ability to process complex carbohydrates in the lignocellulosic feedstock, finding a microbe capable of anaerobic fatty acid production, and lowering enzymatic costs could substantially lower end product cost [33].

In addition to the cost reducing possibilities highlighted above, the lignocellulose to hydrocarbon biological conversion process would also benefit from the use of a microbe that could produce a "drop-in" fuel as opposed to palmitate. This could substantially lower final product costs by minimizing the need for product upgrading (ie. converting palmitate into the appropriately sized alkanes). While not the most costly component of the hydrocarbon production process from lignocellulosic biomass, in the NREL study, product recovery and upgrading contribute to 26 million of the 582.7 million in total capital investment [33]. The NREL study also focuses on the production of a diesel replacement derived from palmitate despite that fact that diesel fuel only accounts for about 25% of transportation fuel consumption [3,33]. A fuel more similar to gasoline would have greater potential to replace fossil fuels and lower GHG emissions.

In this work we engineer *Escherichia coli* to produce fatty acid-based precursors to gasoline-like biofuels in an effort to improve the lignocellulose to hydrocarbon conversion process. *E. coli* naturally produce long chain fatty acids (LCFAs, 13+ carbons), but we direct *E. coli* to produce medium chain fatty acids (MCFAs, 6-12 carbons). MCFAs are precursors to compounds already found in gasoline (medium chain alkanes) as well as medium chain alcohols

that are similar to gasoline in terms of energy density, hydrophobicity, and melting temperatures [32].

E. coli fatty acid metabolism

Many organisms produce LCFAs including palmitate as a normal part of metabolism and lipid synthesis. In *E. coli*, for instance, where fatty acid metabolism has been studied extensively [34-38], LCFAs are major components of cellular lipids that make up 9% of cell dry weight [39]. *E. coli* fatty acid production (Figure 1.2) begins with the 2-carbon compound, acetyl-CoA, produced through glycolysis. The enzyme acetyl-CoA carboxylase (ACC) adds an additional carbon to acetyl-CoA forming malonyl-CoA. FabD swaps the CoA group on malonyl-CoA with a small protein called acyl-carrier protein (ACP) producing malonyl-ACP. ACP shuttles growing fatty acids to the active sites of the fatty acid synthesis enzymes. Malonyl-ACP is decarboxylated and condensed with a second acetyl-CoA molecule through the action of β-ketoacyl-ACP synthase III (FabH) forming acetoacetyl-ACP. Acetoacetyl-ACP is then reduced (FabG), dehydrated (FabZ or FabA), reduced again (FabI), and condensed with a new molecule of malonyl-ACP via the actions of β-ketoacyl-ACP synthase I (FabB) and II (FabF). These reduction, dehydration, reduction, and condensation processes repeat adding 2 carbons onto the growing fatty acid with each round. The process generally stops when the nascent fatty acid is 16 or 18 carbons long at which point the fatty acid is transferred from ACP to glycerol phosphate via PlsB beginning the process of phospholipid synthesis [35,38].

The amount and types of fatty acids produced in *E. coli* are tightly regulated. High levels of acyl-ACPs, indicative of decreased incorporation into the membrane feedback inhibit acetyl-CoA carboxylase (the first enzyme in fatty acid synthesis). Fatty acid synthesis is thereby

balanced with the needs of membrane synthesis [40]. In addition to creating saturated fatty acids, the *E. coli* fatty acid synthesis pathway can produce mono-unsaturated fatty acids with a double

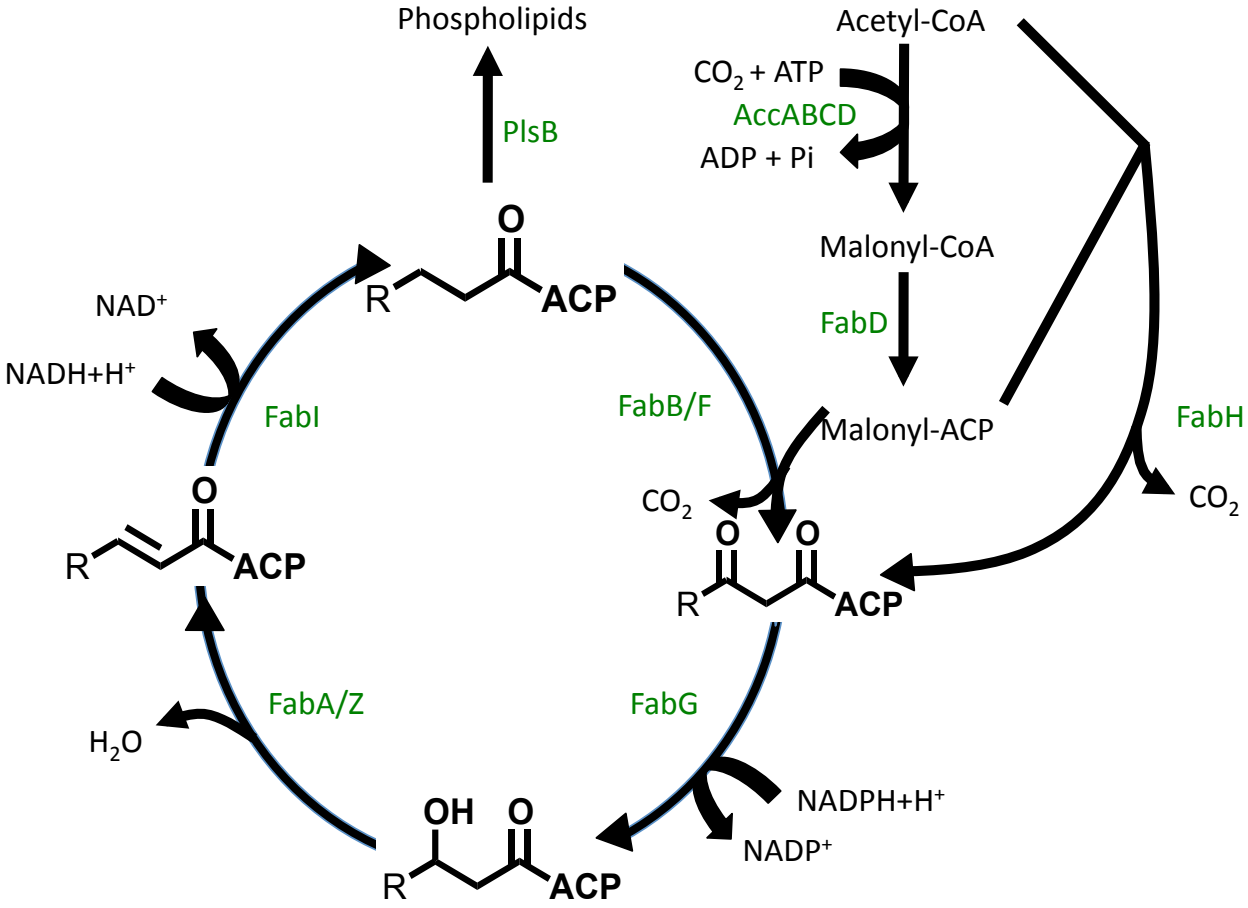

Figure 1.2. Overview of fatty acid synthesis in *E. coli*. See text for details.

bond between carbons 7 and 8. This double bond is introduced through the action of FabA which both dehydrates the nascent β-hydroxyacyl-ACP intermediate in fatty acid synthesis and isomerizes the resulting trans-2-decenoyl-ACP to cis-3-decenoyl-ACP, a precursor to unsaturated fatty acids. Importantly, cis-3-decenoyl-ACP must later be elongated by FabB making FabB essential for unsaturated fatty acid synthesis and growth [35]. The transcription factor FadR (which also functions to repress fatty acid degradation) up-regulates the expression of both FabB and FabA in the absence of fatty acyl-CoAs greater than 11 carbons long [36,41]. This both increases fatty acid synthesis in general and increases the proportion of unsaturated to

saturated fatty acids. As will be discussed later, enhanced FadR expression can have beneficial effects in strains engineered to produce LCFAs [42]. Finally, transcription factor FabR represses FabA and FabB in the presence of unsaturated LCFAs although the precise ligand sensed by FabR is in dispute [43-45].

FabF, the second β-ketoacyl-ACP synthase provides a level of non-essential regulation over fatty acid synthesis: thermal regulation. FabF increases its elongation of palmitoleic acid (16 carbons) to *cis*-vaccenic acid (18 carbons) when cells are grown at low temperatures thereby modulating the properties of the membrane. FabB can also catalyze the elongation of palmitoleic acid and *fabF* mutants are viable but lack the ability to upregulate *cis*-vaccenic acid levels at high temperature [35]. As it is non-essential, *fabF* can either be knocked out or engineered to have different activity without affecting *E. coli* viability.

Importantly, unlike fatty acid synthesis in eukaryotes, the separate fatty acid synthesis enzymes (FabD, FabH, FabB, FabF, FabA, FabZ, FabI, and FabG) are dissociated in *E. coli* making this a type II fatty acid synthesis system as opposed to the type I systems in eukaryotes. In type I fatty acid synthesis systems found in yeast and higher eukaryotes, the synthesis enzymes form individual domains of a large single protein. The dissociated nature of the *E. coli* type II fatty acid synthesis pathway makes it so individual enzymes in the pathway can be altered without necessarily impacting the other synthesis enzymes. It is unclear how altering any one domain in the eukaryotic complex would affect the function of the entire complex [46-48].

The process of fatty acid degradation, β-oxidation, limits the amount of free fatty acids found in *E. coli*. β-oxidation is essentially the reverse of fatty acid synthesis and results in the production of ATP, NADH, and acetyl-CoA that can be further metabolized by the tricaboxylic acid cycle [37,49]. The *fadD* and *fadE* genes (encoding the first and second enzymes in β-

oxidation respectively) are common targets for deletion in strains engineered to produce free fatty acids as their deletion prevents the degradation of the free fatty acid products. However, the acyl-CoA synthetase reaction catalyzed by FadD can play an important role in the activation of free fatty acids for later reduction into alcohols (potential biofuels) [26,27]. FadD catalyzes the conversion of LCFAs into fatty acyl-CoAs with the concomitant hydrolysis of ATP, but has poor activity on fatty acids 10 carbons and shorter making it difficult to use FadD to produce shorter biofuels [50-54]. FadR represses the transcription of β-oxidation genes including FadD in the absence of long chain fatty acyl-CoA thioesters. In the presence of the LCFAs, basal FadD activity is high enough to produce long chain fatty acyl-CoAs to bind to FadR causing its dissociation from cognate β-oxidation gene promoters and the activation of β-oxidation genes [49]. Overall, the various enzymes of the β-oxidation pathway and its regulation through FadR provide means to both enhance fatty acid synthesis and to produce fatty acid derived products as will be discussed below.

Engineering *E. coli* for the production of free fatty acids

Free fatty acids can be produced in *E. coli* by over expressing thioesterases that hydrolyze free fatty acids from ACP after their reduction by FabI [55-58] (Figure 1.2). Thioesterases are naturally used by plants and microorganisms to produce a variety oils and waxes and have many different chain length specificities. The endogenous *E. coli* thioesterase TesA (which is present in the cytoplasm only in a truncated form), has a broad chain length specificity and produces a mixture of fatty acids 12 to 18 carbons long [56]. Other thioesterases have much narrower specificities. For instance, FatB1 from *Cuphea palustris* has high specificity for the hydrolysis of 8-carbon octanoyl-ACP [57]. Because the fatty acid synthesis pathway itself

ends only with the production of long chain acyl-ACPs, the chain length specificities of *E coli* engineered to produce fatty acids are largely determined by the specific thioesterase used. Thioesterases have the added benefit that they lower the levels of long chain acyl-ACPs and relieve feedback inhibition on acetyl-CoA carboyxlase [59].

To compare fatty acid production levels between strains grown using different carbon sources, authors often report yields of fatty acids as the percent theoretically achievable given the organism's metabolism and the carbon source used. The theoretical yield takes into account the reactions of the metabolic pathway used to produce the compound of interest and the energy requirements of the pathway. For example, to produce octanoic acid from glucose, we assume that every molecule of glucose produces two acetyl-CoAs via glycolysis. It takes four acetyl-CoA molecules to produce one octanoic acid through fatty acid synthesis and so, at the theoretical maximum productivity of this pathway, we can produce one molecule of octanoic acid from two molecules of glucose. As for energy requirements, every glucose molecule is converted into two ATP molecules, four NADH (two from glycolysis and two from pyruvate dehydrogenase), and two acetyl-CoA. Three ATP are consumed during octanoic acid synthesis to convert each of three acetyl-CoA's into three malonyl-CoA's that are condensed with a single additional acetyl-CoA. Six NADPH are then required for the reduction reactions in fatty acid synthesis. Assuming that NADH and NADPH are interchangeable, this leaves two additional NADH and an additional ATP to be used for cellular growth. The octanoic acid production pathway therefore leaves ample energy to be used for other processes. The same is true for the production of LCFAs.

Although thioesterase expression can lead to substantial yields of free fatty acids, particularly LCFAs [26,55,56], much further engineering can be done to enhance production and

approach theoretical yields. As indicated above, to prevent fatty acid degradation, the first two genes of β-oxidation, *fadD* and *fadE,* are often knocked out [42,60]. Further enhancement of fatty acid production can be achieved by over-expressing acetyl-CoA carboxylase and other enzymes in fatty acid synthesis. Indeed, one of the greatest yields of LCFAs reported in the literature (73%) was achieved by over-expressing FadR and the long chain thioesterase, TesA. FadR over-expression resulted in the up-regulation of many genes involved in fatty acid synthesis as determined by microarray. Over expression of a subset of these genes individually failed to increase fatty acid production to a comparable level highlighting the benefit of coordinately up-regulating the entire fatty acid synthesis pathway [42].

Alcohols [26,32], alkanes [28-31], and fatty acid ethyl esters [26,60,61] can all be produced from free fatty acids as potential biofuels. Alcohols can be produced at low yields by the over expression of a single type of enzyme, a fatty acyl-CoA reductase [27]. As with thioesterases, acyl-CoA reductases have different chain length specificities resulting in different length alcohol products. These reductases act on acyl-CoAs and therefore require the activity of an acyl-CoA synthetase such as FadD to generate acyl-CoAs from free fatty acids. Recently discovered carboxylic acid reductases can instead directly reduce free fatty acids to produce aldehydes and alcohols [62,63]. Low yields of alkanes have been achieved using reductases that produce fatty aldehydes and aldehyde deformylating oxygenases that liberate the aldehyde carbonyl as formate thereby producing n-1 alkanes [28-31].

Steen *et al.* 2010 produced fatty acid ethyl esters and fatty alcohols in the range of 8-18 and 10-16 carbons respectively as potential diesel fuel replacements [26]. Ethyl esters were produced through the simultaneous over-expression of thioesterases with varying chain length specificities, FadD, an ethanol production pathway (or ethanol supplementation), and a wax-ester

synthase that condensed the ethanol with the fatty acyl-CoAs. Long chain alcohols were produced using a simpler pathway consisting of thioesterses, FadD, and an acyl-CoA reductase that reduced the acyl-CoAs produced by the thiosterases and FadD into alcohols. Production by these various pathways reached no more than 15% theoretical yield and primarily produced long-chain products, but this work provides an excellent demonstration of the many different products achievable through engineering fatty acid synthesis.

Most of the above work uses strains engineered for increased LCFA production. In chapter one of this work we engineer *E. coli* to instead produce the full range of MCFAs as precurors to biofuels that are more similar to gasoline than LCFA derived fuels. MCFAs provide a unique challenge to metabolic engineering efforts because, while the native *E. coli* fatty acid synthesis pathway ends in the production of LCFAs, MCFAs are not abundant in *E. coli* [64,65]. In addition, MCFAs are known to be toxic to *E. coli* meaning that their production is naturally selected against [66-68]. In chapter one, we test the affects of a variety of gene knockouts on the production of the 8-carbon MCFA, octanoate. These knockouts were chosen based on the potential of the encoded enzymes to siphon fatty acid precursors into a variety of other metabolic pathways. Our efforts more than double octanoate production over that of a strain only expressing the 8-carbon fatty acid selective thioesterase FatB1, but only achieve ~5% theoretical yield.

To enhance MCFA production further, we move beyond enzyme over-expression and gene deletions and instead dynamically modulate fatty acid synthesis. These efforts follow in the footsteps of Zhang *et al.* 2011 who increase fatty acid ethyl ester production to 28% theoretical yield using a system wherein the expression levels of ethanol producing enzymes and a wax-ester synthase are dynamically enhanced in the presence of acyl-CoAs [61]. This system

dynamically coordinates the production of the various components of the ethyl ester product such that energy is not wasted on the production of toxic ethanol before fatty acyl-CoA is available to condense with it [61].

In a similar dynamically regulated system, Xu *et al.* 2014 show that by activating the expression of *fabADGI* and *tesA* and inhibiting expression of *accABCD* (the genes for acetyl-CoA carboxylase) in the presence of high intracellular levels of malonyl-CoA, they could enhance the production of LCFAs by more than two fold compared to a strain without this control [69]. In their system, fatty acid synthesis is only activated when there is an abundance of fatty acid precursor (malonyl-CoA) while malonyl-CoA production is inhibited by excess malonyl-CoA. This presumably prevents the cells from wasting resources on enzyme and malonyl-CoA production when they are not needed and balances the malonyl-CoA pool preventing its leakage into other pathways or degradation [69].

The literature contains many further examples of dynamic regulation enhancing the production of a variety of compounds including lycopene [70], alpha-santalene [71], and amorphadiene [72]. In the second half of chapter one we extend this use of dynamic regulation to enhance the production of MCFAs in *E. coli*. Our strains use the ClpXP protease to degrade target proteins tagged with a mutant form of the SsrA tag recognized by ClpXP [73]. This tag causes binding to the ClpXP protease only in the presence of the SspB adapter protein which we place under the control of an IPTG inducible promoter [74]. Our goal through this regulation is to shut off otherwise essential metabolic pathways thereby preventing growth and diverting cellular resources into the production of MCFAs. A similar non-inducible system was previously used to enhance the production of L-tyrosine, but our work represents the first attempt to use inducible degradation to enhance fatty acid production. Our efforts successfully direct the *E. coli*

fatty acid synthesis pathway and enhance production of MCFAs produced by the 8-carbon specific thioesterase FatB1 from *Cuphea palustris* [57].

Converting MCFAs into activated CoA thioesters

MCFAs must be further modified to produce medium chain products like alcohols, alkanes, and ethyl esters. The *E. coli* acyl-CoA synthetase FadD, is commonly used to activate fatty acids by converting them into acyl-CoA thioesters that can later be reduced or converted into ethyl esters [26-28,60,61]. However, FadD has low activity on MCFAs. This limits FadD's usefulness for further conversion of MCFAs into downstream products.

In chapter 2 of this work we use the unbiased technique of adaptive evolution to generate strains of *E. coli* with enhanced growth rate on octanoic acid predicting that, as the FadD catalyzed conversion of fatty acids into fatty acyl-CoAs is the first step in *E. coli* β-oxidation, we would discover mutations that facilitate this process. We successfully generate strains with enhanced growth on octanoic acid, and whole genome sequencing reveals two mutations that up-regulate *fadD* expression. A final mutation in dihydroxyacetone kinase (*dhaM*) moderately enhances growth rate on many different carbon sources and does not have specific effects on MCFA metabolism. This work directs us to focus on engineering FadD directly to alter medium chain acyl-CoA production.

The precise mechanism determining FadD's substrate specificity is not well understood. Structures of a homologous long chain acyl-CoA synthetase from *Thermus thermophilus* and a medium chain acyl-CoA synthetase from *Homo sapiens* suggest that it is the length of the FadD fatty acid binding pocket that determines specificity [75,76]. There is however, no structure available for FadD. Therefore, in chapter three we use error prone polymerase chain reaction

(PCR) to generate FadD mutants and a growth-based screen to select for mutants with enhanced activity on MCFAs. Based off of FadD homology modeling and *in vitro* assays, we propose that the FadD mutations we discover enhance activity on MCFAs by making it easier for AMP produced in the acyl-CoA synthetase reaction to escape the FadD active site. These mutants should prove useful for the future production of MCFA derived products.

The overarching goal of this work was to direct *E. coli* fatty acid synthesis toward the production of medium chain length compounds and to generate tools and techniques that could be used for further metabolic engineering. The work in chapters one and three should be useful for a variety of applications. For instance, the essential gene inducible degradation technique used in chapter one should be applicable to any pathway in which the down-regulation of an essential enzyme is required to enhance yields. The broadened substrate specificity of the FadD mutants developed in chapter three should make them suitable for engineering the production of fatty acid-derived compounds of many different lengths. These mutants could also be tested on a variety of non-fatty acid substrates to determine if novel activity has been obtained. Finally, because adenylate forming enzymes like FadD have similar structures, the technique of aiding substrate exit to enhance activity on different length substrates may be applicable to these enzymes as well. Overall, the techniques and enzymes developed here will hopefully find broader application in further metabolic engineering of *E. coli* and other organisms for the production of a wide variety of compounds.

Chapter Summary

Chapter 1: Engineering E. coli *for the production of medium chain fatty acids*

This chapter summarizes work performed in concert with Joseph P. Torella to enhance MCFA production in *E. coli*. We show that *E. coli* can produce odd chain fatty acids when supplemented with propionate or expressing a propionate production pathway. Our strains produce all fatty acids 4-12 carbons long by expressing thioesterases with different chain length specificities. We additionally test a variety of metabolic knockouts for their effects on MCFA production and use an inducible degradation system to direct fatty acid synthesis toward MCFA production. We achieve 12% theoretical yield of the MCFA octanoate.

Chapter 2: Adaptive evolution to enhance E. coli *growth on octanoic acid*

E. coli grow poorly on MCFAs even in the presence of regulatory mutations that de-repress expression of β-oxidation genes [50,52,77,78]. Hypothesizing that this slow growth could limit *E. coli* conversion of MCFAs into downstream products like alcohols and alkanes, in this chapter we use adaptive evolution to generate strains with increased growth rate on the MCFA octanoic acid. Whole genome sequencing and genetic characterization of the acquired mutations reveal that mutations that enhance expression of *E. coli* acyl-CoA synthetase FadD are sufficient to enhance growth rate on octanoic acid.

Chapter 3: Enhancing E. coli *acyl-CoA synthease FadD activity on medium chain fatty acids*

FadD, which catalyzes the first step in *E. coli* β-oxidation, the activation of free fatty acids into acyl-CoA thioesters, has low activity on MCFAs. This low activity may limit the production of downstream products like alcohols and alkanes from MCFAs in *E. coli*. In this

chapter we use a growth-based screen to generate mutations in *fadD* that enhance FadD activity on MCFAs. We verify the activity of these FadD mutants via *in vitro* assays of partially purified His_6-tagged FadD and its mutants. Using homology models and our biochemical data we develop the hypothesis that our mutants' increased activity is due to enhanced product exit. We then design further mutations hypothesized to enhance product exit and show that they too enhance FadD activity on MCFAs.

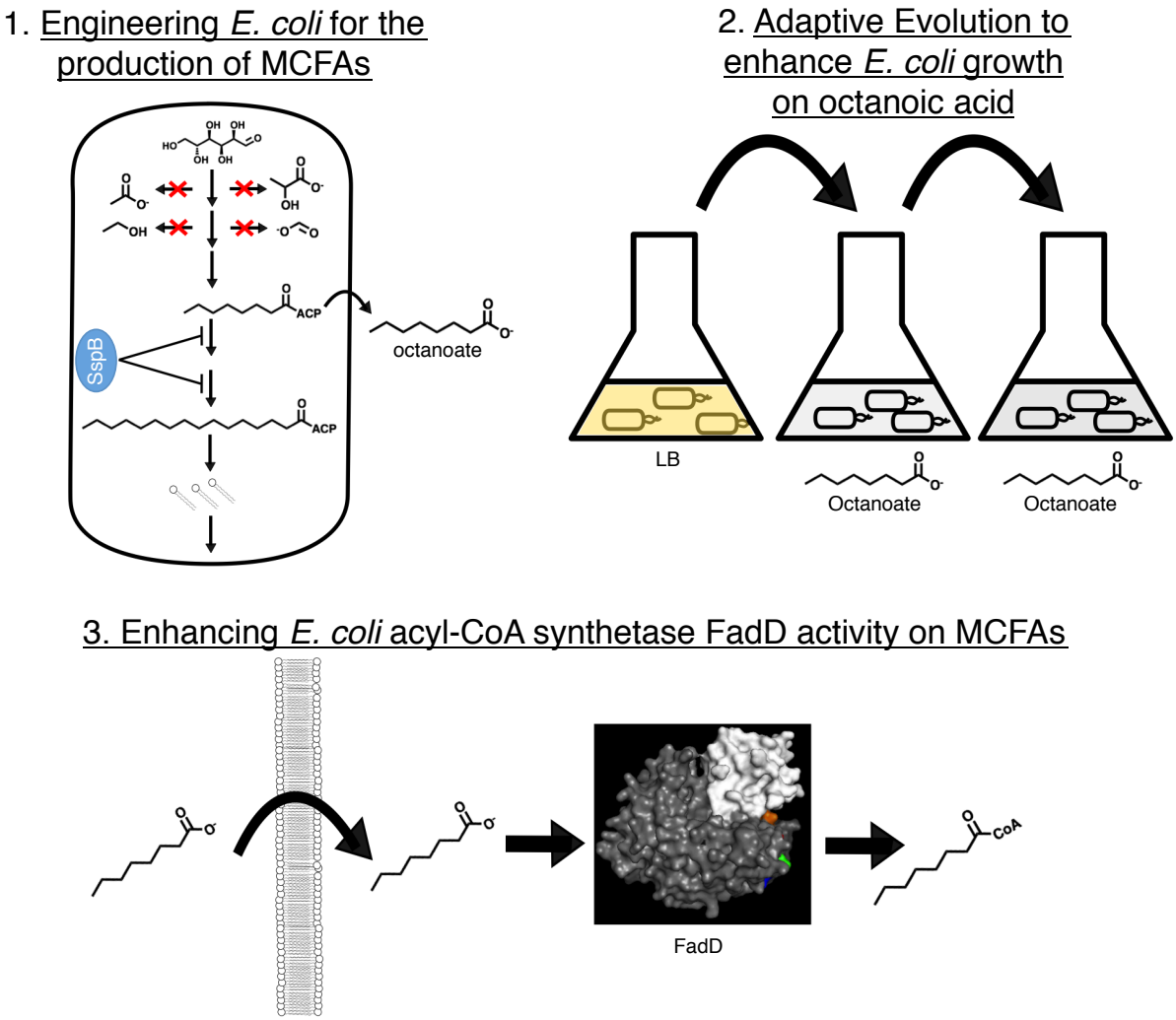

Figure 1.3. Chapter overview. See text for details

References

1. Intergovernmental Panel on Climate Change.: **CLIMATE CHANGE 2014 SYNTHESIS REPORT**. 2014.

2. United States. Environmental Protection Agency.: **National Greenhouse Gas Emissions Data DRAFT INVENTORY OF U.S. GREENHOUSE GAS EMISSIONS AND SINKS: 1990-2013 (February 2015)**. 2015.

3. United States. Environmental Protection Agency.: **Fast Facts U.S. Tranportation Sector Greenhouse Gas Emissions 1990-2012**. 2015.

4. Mueller S, Kwik J: **2012 Corn Ethanol: Emerging Plant Energy and Environmental Technologies**. 2013.

5. United States. Department of Energy.: **Ethanol Feedstocks, http://www.afdc.energy.gov/fuels/ethanol_feedstocks.html**. 2015.

6. Wang M, Han J, Dunn JB, Cai H, Elgowainy A: **Well-to-wheels energy use and greenhouse gas emissions of ethanol from corn, sugarcane and cellulosic biomass for US use**. *Environmental Research Letters* 2012, **7**.

7. Searchinger T, Heimlich R, Houghton RA, Dong F, Elobeid A, Fabiosa J, Tokgoz S, Hayes D, Yu TH: **Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change**. *Science* 2008, **319**:1238-1240.

8. United States. Department of Energy.: **U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry**. 2011.

9. Schnepf R, Yacobucci BD: **Renewable Fuel Standard (RFS): Overview and Issues**. 2013:35.

10. Dutta A, Talmadge M, Hensley J, Worley M, Dudgeon D, Barton D, Groenendijk P, Ferrari D, Stears B, Searcy EM, et al.: **Process design and economics for conversion of lignocellulosic biomass to ethanol Thermochemical Pathway by Indirect Gasification and Mixed Alcohol Synthesis**. *Contract* 2011, **303**:275-3000.

11. Humbird D, Davis R, Tao L, Kinchin C, Hsu D, Aden A, Schoen P, Lukas J, Olthof B, Worley M: **Process design and economics for biochemical conversion of lignocellulosic biomass to ethanol. National Renewable Energy Laboratory Technical Report NREL**. 2011.

12. International Energy Agency.: **Advanced Biofuels-GHG Emissions and Energy Balances: A REPORT TO IEA TASK 39**. 2013.

13. United states. Energy Information Administration.: **Short term energy outlook**. 2015.

14. United States. Environmental Protection Agency.: **E15 (a blend of gasoline and ethanol), http://www.epa.gov/oms/regs/fuels/additive/e15/**. 2014.

15. United States. Department of Energy.: **Alternative Fuels Data Center: Flexible Fuel Vehicles, http://www.afdc.energy.gov/vehicles/flexible_fuel.html**. 2014.

16. United States. Department of Transportation. Office of the Assistant Scretary for Research and Technology.: **Table 1-11: Number of U.S. Aircraft, Vehicles, Vessels, and Other Conveyances, http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_statistics/html/table_01_11.html**. 2015.

17. United States. Energy Information Administration.: **Renewable and Alternative Fuels Alternative Fuel Vehicle Data, http://www.eia.gov/renewable/afv/**. 2013.

18. United States. Department of Energy.: **Ethanol Fueling Station Locations, http://www.afdc.energy.gov/fuels/ethanol_locations.html**. 2015.

19. United States. Energy Information Administration.: **Access to Alternative transportation fuel stations varies across the lower 48 states, http://www.eia.gov/todayinenergy/detail.cfm?id=6050**. 2012.

20. Whims J: **Pipeline considerations for ethanol**. *Agricultural Marketing Resource Center, http://www.agmrc.org/media/cms/ksupipelineethl_8BA5CDF1FD179.pdf* 2002.

21. Hammel-Smith C, Fang J, Powders M, Aabakken J: **Issues associated with the use of higher ethanol blends**. *National Renewable Energy Laboratory* 2002.

22. United States. Department of Energy.: **Ethanol Production and Distribution, http://www.afdc.energy.gov/fuels/ethanol_production.html**.

23. Strogen B, Horvath A, Zilberman D: **Energy intensity, life-cycle greenhouse gas emissions, and economic assessment of liquid biofuel pipelines**. *Bioresource technology* 2013, **150**:476-485.

24. United States. Department of Energy.: **Report to Congress: Dedicated ethanol pipeline feasibility study**. *Energy Independence and Security Act of 2007* 2010, **243**:1-26.

25. Ritter S: **GASOLINE Most people don't give car fuel much thought, but it is quite a complex mixture of hydrocarbons**. In *Chemical and Engineering News*, edn February 21, 2005. Edited by: American Chemical Society; 2005:1. vol 83.]

26. Steen EJ, Kang YS, Bokinsky G, Hu ZH, Schirmer A, McClure A, del Cardayre SB, Keasling JD: **Microbial production of fatty-acid-derived fuels and chemicals from plant biomass**. *Nature* 2010, **463**:559-U182.

27. Doan TT, Carlsson AS, Hamberg M, Bulow L, Stymne S, Olsson P: **Functional expression of five Arabidopsis fatty acyl-CoA reductase genes in Escherichia coli**. *Journal of plant physiology* 2009, **166**:787-796.

28. Choi YJ, Lee SY: **Microbial production of short-chain alkanes**. *Nature* 2013, **502**:571-574.

29. Howard TP, Middelhaufe S, Moore K, Edner C, Kolak DM, Taylor GN, Parker DA, Lee R, Smirnoff N, Aves SJ, et al.: **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 of the United States of America* 2013, **110**:7636-7641.

30. Sachdeva G, Garg A, Godding D, Way JC, Silver PA: **In vivo co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner**. *Nucleic acids research* 2014, **42**:9493-9503.

31. Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB: **Microbial biosynthesis of alkanes**. *Science* 2010, **329**:559-562.

32. Akhtar MK, Dandapani H, Thiel K, Jones PR: **Microbial production of 1-octanol: A naturally excreted biofuel with diesel-like properties**. *Metabolic engineering communications* 2015, **2**:1-5.

33. Davis R, Tao L, Tan E, Biddy M, Beckham G, Scarlata C, Jacobson J, Cafferty K, Ross J, Lukas J: **Process design and economics for the conversion of lignocellulosic biomass to hydrocarbons: Dilute-acid and enzymatic Deconstruction of biomass to sugars and biological conversion of Sugars to Hydrocarbons**. Edited by: National Renewable Energy Laboratory (NREL), Golden, CO.; 2013.

34. Rock CO, Cronan JE: **Escherichia coli as a model for the regulation of dissociable (type II) fatty acid biosynthesis**. *Biochimica et biophysica acta* 1996, **1302**:1-16.

35. Magnuson K, Jackowski S, Rock CO, Cronan JE, Jr.: **Regulation of fatty acid biosynthesis in Escherichia coli**. *Microbiological reviews* 1993, **57**:522-542.

36. Fujita Y, Matsuoka H, Hirooka K: **Regulation of fatty acid metabolism in bacteria**. *Molecular microbiology* 2007, **66**:829-839.

37. Nunn WD: **A molecular view of fatty acid catabolism in Escherichia coli**. *Microbiological reviews* 1986, **50**:179-192.

38. White SW, Zheng J, Zhang YM, Rock: **The structural biology of type II fatty acid biosynthesis**. *Annual review of biochemistry* 2005, **74**:791-831.

39. Neidhardt FC, Ingraham JL, Low KB, Magasanik B, Schaechter M, Umbarger H: *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology. Volumes I and II*: American Society for Microbiology; 1987.

40. Davis MS, Cronan JE: **Inhibition of Escherichia coliAcetyl Coenzyme A Carboxylase by Acyl-Acyl Carrier Protein**. *Journal of bacteriology* 2001, **183**:1499-1503.

41. DiRusso C, Heimert TL, Metzger AK: **Characterization of FadR, a global transcriptional regulator of fatty acid metabolism in Escherichia coli. Interaction with the fadB promoter is prevented by long chain fatty acyl coenzyme A**. *Journal of Biological Chemistry* 1992, **267**:8685-8691.

42. Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD: **Enhancing fatty acid production by the expression of the regulatory transcription factor FadR**. *Metabolic engineering* 2012, **14**:653-660.

43. Zhang Y-M, Rock CO: **Transcriptional regulation in bacterial membrane lipid synthesis**. *Journal of lipid research* 2009, **50**:S115-S119.

44. Zhu K, Zhang YM, Rock CO: **Transcriptional regulation of membrane lipid homeostasis in Escherichia coli**. *The Journal of biological chemistry* 2009, **284**:34880-34888.

45. Feng Y, Cronan JE: **Complex binding of the FabR repressor of bacterial unsaturated fatty acid biosynthesis to its cognate promoters**. *Molecular microbiology* 2011, **80**:195-218.

46. Schweizer E, Hofmann J: **Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems**. *Microbiology and molecular biology reviews : MMBR* 2004, **68**:501-517, table of contents.

47. Maier T, Leibundgut M, Ban N: **The crystal structure of a mammalian fatty acid synthase**. *Science* 2008, **321**:1315-1322.

48. Lomakin IB, Xiong Y, Steitz TA: **The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together**. *Cell* 2007, **129**:319-332.

49. Black PN, DiRusso CC: **Molecular and biochemical analyses of fatty acid transport, metabolism, and gene regulation in Escherichia coli**. *Biochimica et biophysica acta* 1994, **1210**:123-145.

50. Salanitro JP, Wegener WS: **Growth of Escherichia coli on short-chain fatty acids: growth characteristics of mutants**. *Journal of bacteriology* 1971, **108**:885-892.

51. Overath P, Pauli G, Schairer HU: **Fatty acid degradation in Escherichia coli. An inducible acyl-CoA synthetase, the mapping of old-mutations, and the isolation of regulatory mutants**. *European journal of biochemistry / FEBS* 1969, **7**:559-574.

52. Campbell JW, Morgan-Kiss RM, Cronan JE, Jr.: **A new Escherichia coli metabolic competency: growth on fatty acids by a novel anaerobic beta-oxidation pathway**. *Molecular microbiology* 2003, **47**:793-805.

53. Kameda K, Nunn WD: **Purification and characterization of acyl coenzyme A synthetase from Escherichia coli**. *The Journal of biological chemistry* 1981, **256**:5702-5707.

54. Iram SH, Cronan JE: **The beta-oxidation systems of Escherichia coli and Salmonella enterica are not functionally equivalent**. *Journal of Bacteriology* 2006, **188**:599-608.

55. Voelker TA, Davies HM: **Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of a plant medium-chain acyl-acyl carrier protein thioesterase**. *Journal of bacteriology* 1994, **176**:7320-7327.

56. Cho H, Cronan JE, Jr.: **Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis**. *The Journal of biological chemistry* 1995, **270**:4216-4219.

57. Dehesh K, Edwards P, Hayes T, Cranmer AM, Fillatti J: **Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of Cuphea palustris seed oil**. *Plant physiology* 1996, **110**:203-210.

58. Jing F, Cantu DC, Tvaruzkova J, Chipman JP, Nikolau BJ, Yandeau-Nelson MD, Reilly PJ: **Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity**. *BMC biochemistry* 2011, **12**:44.

59. Jiang P, Cronan JE, Jr.: **Inhibition of fatty acid synthesis in Escherichia coli in the absence of phospholipid synthesis and release of inhibition by thioesterase action**. *Journal of bacteriology* 1994, **176**:2814-2821.

60. Handke P, Lynch SA, Gill RT: **Application and engineering of fatty acid biosynthesis in Escherichia coli for advanced fuels and chemicals**. *Metabolic engineering* 2011, **13**:28-37.

61. Zhang F, Carothers JM, Keasling JD: **Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids**. *Nature biotechnology* 2012, **30**:354-359.

62. Akhtar MK, Turner NJ, Jones PR: **Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**:87-92.

63. Akhtar MK, Dandapani H, Thiel K, Jones PR: **Microbial Production of 1-octanol: A naturally extreted biofuel with diesel-like properties**. *Metabolic engineering communications* 2015, **2**:1-5.

64. Jackowski S, Rock CO: **Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria**. *The Journal of biological chemistry* 1987, **262**:7927-7931.

65. Rock CO, Jackowski S: **Regulation of phospholipid synthesis in Escherichia coli. Composition of the acyl-acyl carrier protein pool in vivo**. *The Journal of biological chemistry* 1982, **257**:10759-10765.

66. Sherkhanov S, Korman TP, Bowie JU: **Improving the tolerance of Escherichia coli to medium-chain fatty acid production**. *Metabolic engineering* 2014, **25**:1-7.

67. Lennen RM, Pfleger BF: **Modulating membrane composition alters free fatty acid tolerance in Escherichia coli**. *PloS one* 2013, **8**:e54031.

68. Lennen RM, Kruziki MA, Kumar K, Zinkel RA, Burnum KE, Lipton MS, Hoover SW, Ranatunga DR, Wittkopp TM, Marner WD, 2nd, et al.: **Membrane stresses induced by overproduction of free fatty acids in Escherichia coli**. *Applied and environmental microbiology* 2011, **77**:8114-8128.

69. Xu P, Li LY, Zhang FM, Stephanopoulos G, Koffas M: **Improving fatty acids production by engineering dynamic pathway regulation and metabolic control**. *Proceedings of the National Academy of Sciences of the United States of America* 2014, **111**:11299-11304.

70. Farmer WR, Liao JC: **Improving lycopene production in Escherichia coli by engineering metabolic control**. *Nature biotechnology* 2000, **18**:533-537.

71. Scalcinati G, Knuf C, Partow S, Chen Y, Maury J, Schalk M, Daviet L, Nielsen J, Siewers V: **Dynamic control of gene expression in Saccharomyces cerevisiae engineered for the production of plant sesquitepene alpha-santalene in a fed-batch mode**. *Metabolic engineering* 2012, **14**:91-103.

72. Dahl RH, Zhang F, Alonso-Gutierrez J, Baidoo E, Batth TS, Redding-Johanson AM, Petzold CJ, Mukhopadhyay A, Lee TS, Adams PD, et al.: **Engineering dynamic pathway regulation using stress-response promoters**. *Nature biotechnology* 2013, **31**:1039-+.

73. Baker TA, Sauer RT: **ClpXP, an ATP-powered unfolding and protein-degradation machine**. *Biochimica et biophysica acta* 2012, **1823**:15-28.

74. McGinness KE, Baker TA, Sauer RT: **Engineering controllable protein degradation**. *Molecular cell* 2006, **22**:701-707.

75. Hisanaga Y, Ago H, Nakagawa N, Hamada K, Ida K, Yamamoto M, Hori T, Arii Y, Sugahara M, Kuramitsu S, et al.: **Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer**. *The Journal of biological chemistry* 2004, **279**:31717-31726.

76. Kochan G, Pilka ES, von Delft F, Oppermann U, Yue WW: **Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A**. *Journal of molecular biology* 2009, **388**:997-1008.

77. Maloy SR, Ginsburgh CL, Simons RW, Nunn WD: **Transport of long and medium chain fatty acids by Escherichia coli K12**. *The Journal of biological chemistry* 1981, **256**:3735-3742.

78. Nunn WD, Simons RW, Egan PA, Maloy SR: **Kinetics of the utilization of medium and long chain fatty acids by mutant of Escherichia coli defective in the fadL gene**. *The Journal of biological chemistry* 1979, **254**:9130-9134.

Chapter 2: Engineering *E. coli* for the production of medium chain fatty acids

Preface

This work was performed in concert Joseph P. Torella as well as two undergraduates, Amanda M. Chen, and Scott N. Kim, who primarily generated plasmid constructs and generated some of the strains used in this work. This chapter is adapted from our paper published in *Proceedings of the National Academy of Sciences of the United States of America*:

Torella, Joseph P., et al. "Tailored fatty acid synthesis via dynamic control of fatty acid elongation." *Proceedings of the National Academy of Sciences* 110.28 (2013): 11290-11295.

Abstract

Medium-chain fatty acids (MCFAs, generally 6-12 carbons, but here expanded to 4-13 carbons) are valuable as precursors to industrial chemicals and biofuels, but are not canonical products of microbial fatty acid synthesis. We engineered microbial production of the full range of even- and odd-chain-length MCFAs and found that MCFA production is limited by rapid, irreversible elongation of their acyl-ACP precursors. To address this limitation, we programmed an essential ketoacyl synthase to degrade in response to a chemical inducer, thereby slowing acyl-ACP elongation and redirecting flux from phospholipid synthesis to MCFA production. Our results show that induced protein degradation can be used to dynamically alter metabolic flux, and thereby increase the yield of a desired compound. The strategy reported here should be widely useful in a range of metabolic engineering applications in which essential enzymes divert flux away from a desired product, as well as in the production of polyketides, bioplastics, and other recursively-synthesized hydrocarbons for which chain length control is desired.

Introduction

Long chain (16-18 carbon) fatty acids are required to synthesize the essential lipids that maintain bacterial membrane integrity [1]. In-depth genetic and biochemical understanding of fatty acid synthesis in bacteria [2,3] has made it possible to engineer microbes for increased fatty acid production, and catalyzed efforts to industrialize the process [4]. As fatty acids can be transformed into a variety of biofuels and industrially useful oleo-chemicals, engineered microbes are a promising renewable alternative to petroleum-based chemical feedstocks [4-8].

Recent efforts have focused on engineering *E. coli* fatty acid synthesis for production of free fatty acids (Reviewed in [4]), alcohols [5], esters [6] and alkanes [7]. While most such studies have focused on long-chain (C14-C18) compounds due to their natural abundance, medium-chain fatty acids (MCFAs, 4-12 carbons), which are derived from low-abundance acyl-ACPs, are nevertheless valuable industrial chemicals [9-12], and their shorter chain lengths are associated with improved fuel quality [8,13]. Tailoring chain length specificity in fatty-acid-producing microbes is therefore an important challenge in the production of renewable chemicals and biofuels.

Chain length control is exerted at many steps throughout fatty acid synthesis (FAS). FAS begins with the condensation of acetyl-CoA and malonyl-ACP by the ketoacyl synthase (KAS) FabH to yield acetoacetyl-ACP, a four-carbon β-ketoacyl-ACP that is reduced to yield a four-carbon acyl-ACP. In subsequent rounds of fatty acid synthesis, this acyl-ACP is condensed with additional malonyl-ACP molecules by long-chain KAS enzymes FabB and FabF, then reduced to an acyl-ACP with two additional carbons. In this way, *E. coli* fatty acid synthesis builds acyl-ACPs two carbons at a time, in a recursive process yielding a range of even-chain-length acyl-ACPs [2]. Odd-chain acyl-ACPs can also be produced when cellular propionyl-CoA levels are

high, due to their incorporation in place of acetyl-CoA in the initial step of fatty acid synthesis [14-16]. Due in part to KAS chain-length specificity, acyl-ACP elongation ends with the production of long-chain (16-18 carbon) acyl-ACPs, which are used by the cell to synthesize membrane lipids (Reviewed in [2]).

Engineered fatty acid production in *E. coli* is typically achieved by expressing a cytosolic derivative of the *E. coli* thioesterase TesA [5], which hydrolyzes long-chain acyl-ACPs to yield long-chain free fatty acids (LCFAs) [17]. High yields are due to both the natural abundance of long-chain acyl-ACPs in the cell and thioesterase-mediated depletion of the long chain acyl-ACP pool, which feedback-inhibits upstream enzymes in fatty acid synthesis such as FabH [16,17]. While MCFAs can be produced by expressing thioesterases with substrate specificity for medium-chain acyl-ACPs [5,18-21], yields are generally lower than for LCFAs, likely because medium-chain acyl-ACPs are not abundant [22,23] and their depletion does not resolve an existing buildup of inhibitory long-chain acyl-ACPs.

In this work, we engineered *E. coli* to produce free fatty acids with all even- and odd-chain lengths from 4-13 carbons, and rationally modified fatty acid synthesis to favor the production of medium-chain acyl-ACPs. Treatment of MCFA-producing cells with the KAS inhibitor cerulenin [22] increased yields, demonstrating that MCFA production is limited by overly-rapid acyl-ACP elongation. To test whether this strategy could be implemented genetically to increase the production of a specific MCFA, octanoic acid, we replaced the KAS FabF with a mutant incapable of elongating beyond 8 carbons [24]. We then engineered FabB to degrade in response to a chemical inducer [25], such that elongation beyond C8 acyl-ACP could be slowed on-demand, and flux redirected from lipid synthesis to octanoate production. These interventions increased octanoate yield, and demonstrated the utility of inducible degradation to

degrade essential genes and redirect metabolic flux [26]. Our results demonstrate that altering the chain length specificity of microbial fatty acid synthesis requires concerted changes in both the thioesterase and fatty acid synthesis machinery itself, and suggest strategies for the production of intermediate-length products in other recursive biosynthetic systems [27,28].

Materials and Methods

Plasmid Construction and Transformation

N-terminally his-tagged *CpfatB1*, codon optimized for expression in *S. elongatus*, was PCR amplified from plasmid pDFS705 (constructed by Dave F. Savage) using primers TF0013 and TF0014, digested with SpeI and HindIII, and ligated into the XbaI and HindIII sites of plasmid pETDuet-1 (Merck) to generate pTJF010. It was also PCR amplified with primers TF0133 and TF0014, digested with NcoI and HindIII, and inserted into the NcoI and HindIII sites of pACYCDuet-1 (Merck) to generate pTJF038. To generate pJT255, *CpFatB1* was amplified with JT014 and JT015, digested with SpeI/BglII, and ligated into pWW308 (a gift from Weston Whitaker in the Dueber Lab carrying the aTc-inducible Pzt promoter) digested with NheI/BamHI. The resulting plasmid, pJT180, was then digested with HpaI/XhoI and the Pzt-CpFatB1 fragment ligated into pCDFDuet-1 digested with HpaI/XhoI. *UcfatB2* codon-optimized for expression in *S. elongatus* was PCR amplified from pDFS703 using primers JT0138 and JT0134, digested with XbaI and AflII, and inserted into the same sites of pETDuet-1 to generate pJT208.

Table 2.1. Strains Used in this Chapter

Strain Designation in Text	**Description**	**Source**
BL21*(DE3)	F- ompT hsdSB (rB-mB-) gal dcm rne131(DE3)	Invitrogen
S001	BL21*(DE3)*ΔfadD*	This study
S002	BL21*(DE3)*ΔfadE*	This study
S003	BL21*(DE3)*ΔfadE fabF**	This study
S004	BL21*(DE3)*ΔfadE fabBDeg*	This study
S005	BL21*(DE3)*ΔfadE fabF* fabBDeg*	This study
S006	BL21*(DE3)*ΔfadD Δpta ΔlacY fabF* fabBDeg*	This study
S007	BL21*(DE3)*ΔfadE strepfabBDeg*	This Study
Δpta	BL21*(DE3)*Δpta*	This Study
ΔpflB	BL21*(DE3)*ΔplfB*	This Study
ΔadhE	BL21*(DE3)*ΔadhE*	This Study
ΔldhA	BL21*(DE3)*ΔldhA*	This Study
ΔpoxB	BL21*(DE3)*ΔpoxB*	This Study
ΔrelA	BL21*(DE3)*ΔrelA*	This Study
ΔfadD	BL21*(DE3)*ΔfadD*	This Study
ΔfadΔpta	BL21*(DE3)*ΔfadDΔpta*	This Study
ΔfadDΔpflB	BL21*(DE3)*ΔfadDΔpflB*	This Study
ΔfadDΔadhE	BL21*(DE3)*ΔfadDΔadhE*	This Study
ΔfadDΔldha	BL21*(DE3)*ΔfadDΔldhA*	This Study
ΔfadDΔpoxB	BL21*(DE3)*ΔfadDΔpoxB*	This Study
ΔfadDΔrelA	BL21*(DE3)*ΔfadDΔrelA*	This Study
ΔfadDΔptaΔpflB	BL21*(DE3)*ΔfadDΔptaΔpflB*	This Study
ΔfadDΔptaΔadhE	BL21*(DE3)*ΔfadDΔptaΔadhE*	This Study
ΔfadDΔptaΔldhA	BL21*(DE3)*ΔfadDΔptaΔldhA*	This Study
ΔfadDΔptaΔpoxB	BL21*(DE3)*ΔfadDΔptaΔpoxB*	This Study
ΔfadDΔptaΔrelA	BL21*(DE3)*ΔfadDΔptaΔrelA*	This Study
TB10	Recombineering Strain MG1655, nadA::Tn10 λcI857 Δ(cro-bioA)	(46)

Table 2.2. Plasmids used in this chapter

Plasmid	**Description**	**Promoter**	**Resistance**	**Source**
sspBpET21b	*E.coli sspB* cloned into pET21b, obtained from Robert T. Sauer's Lab at MIT	T7lac	Amp	(26)

Table 2.2 (Continued).

pTJF010	*C. palustris fatb1* codon-optimized for *S. elongatus*, containing N-terminal 6x-his tag cloned into the XbaI and HindIII sites of pETDuet-1	T7lac	Amp	This Study
pTJF038	*C. palustris fatb1* codon-optimized for *S. elongatus*, containing N-terminal 6x-his tag cloned into the NcoI and HindIII sites of pACYCDuet-1	T7lac	Cam	This Study
pEET	*B. formatexigens* thioesterase EET61113, codon-optimized for *E. coli*, cloned into pUC57.	*lacZ*	Amp	(22)
pJT208	*U.californica fatB2* codon optimized for *S. elongatus*, cloned into pETDuet-1	T7lac	Amp	This Study
pJT255	*Cpfatb1* from pTJF010 cloned into an anhydrotetracycline-inducible vector	PLtetO	Sp	This Study
pJT261	SspBpET21b with T7lac promoter replaced with PLtetO	PLtetO	Amp	This Study
pDFS705	Syneccochocus expression vector containing codon-optimized, N-terminally his-tagged *C. palustris fatB1*	*Trc*	Cam	Gift from Dave Savage
pUC19_FabF	pUC19 containing *E. coli fabF*	*lacZ*	Amp	Gift from Drew MacKellar
pCP20	contains heat shock inducible flippase and is itself heat curable	λpR	Amp	(49)
pCOLAthrAfrBCilvAfr	pCOLADuet-1 expressing thrABC (thrA is feedback-resistant) and ilvA (feedback-resistant), obtained from Kristala Prather's lab at MIT.	T7lac	Kan	(16)

Table 2.3. Primers used in this chapter

Primer	**Description**	**Sequence**
TF0013	*CpfatB1* forward primer with RBS, 6x His-tag, 5' SpeI used to amplify *CpfatB1* from DFS705	CACACTAGTAGGAGGAAAAACATATGCATCACCATCATCATCACAGTTCGTTGTTGACCGCTATCACTAC

Table 2.3 (Continued).

TF0014	*CpfatB1* Rev primer with 3' XbaI and HindII restriction sites used to amplify *CpfatB1* from DFS705	CTGCAAGCTTTCTAGATTACGTTTTTCCGGTCG
TF0101	Forward primer for *fabBDeg_Kan*	ACCAACGCCACGCTGGTAATG
TF0102	Reverse primer for *fabBDeg_Kan*	GCGACGCTGGCGCGTCTAC
TF0111	sense *fabF I-F* mutatgenic primer	ATTGGCTCCGGGTTTGGCGGCCTCG
TF0113	antisense *fabF I-F* mutagenic primer	CGAGGCCGCCAAACCCGGAGCCAAT
TF0119	forward kanR primer containing 5' frt site and homology to the 3' end of *fabF*	GATCTAAGTCGACCTGCAGGGAAGTTCCTATTCTCTAGAAAGTATAGGAACTTCCTTTAAGAAGGAGATATACCATGAGCCATATTCAACGGGAAACGTCTTGC
TF0120	Reverse kanR primer containing 3' frt site and 40bp of homology downstream of *fabF*	AAGCTAAGAAAAAAGGCCCGCAAGCGGACCTTTTATAAGGGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCTTAGAAAAACTCATCGAGCATCAAATGAAACTGC
TF0121	Forward *fabF* primer containing 40bp homology upstream of the *fabF* I-F mutation for replacing *fabF* with *fabF**	AATATGGAATTGTCGCTGGCGTTCAGGCCATGCAGGATTC
TF0122	Reverse *fabF* primver for amplifying *fabF* from pUC19_*fabF*	CCTGCAGGTCGACTTAGATC
TF0133	*fatB1* forward primer with 6x His-tag and 5' NcoI site used to amplify *Cpfatb1* from DF705	CACCCATGGGCCATCACCATCATCATCACAGTTCGTTGTTGACCGCTATCACTACTG
TF0134	Forward *bla* primer containing 40bp of homology to the region upstream of the *fabF* I-F mutation	ATAACGGAAGAGAACGCAACCCGCATTGGTGCCGCAATTGATGAGTATTCAACATTTCCGTGTCG
TF0135	Reverse *bla* primer containing 40bp of homology downstream of *fabF*	AAGCTAAGAAAAAAGGCCCGCAAGCGGACCTTTTATAAGGGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCTTACCAATGCTTAATCAGTGAGGCAC
TF0222	Forward primer containing the N-terminal strep tag for *strepfabBDeg_kan*	ATTCGAAACTTACTCTATGTGCGACTTACAGAGGTATTGAATGTGGAGCCACCCGCAGTTCGAAAAAAAACGTGCAGTGATTACTGG

Table 2.3 (Continued).

JT0138	Forward primer for codon-optimized *UcfatB2.*	TCAGTCTAGACGCAATCTCACGCA AATCACGGTAATCCCTAAA TGACGAATCTCGAATGGAAACC
JT0134	Reverse primer for codon-optimized *UcfatB2*	GACTCTTAAGCTAGACCCGGGGTT CAGCTGG

Growth and Induction

Fatty acid production experiments were performed either in shake flasks or 96-well plates. For shake flask cultures, individual colonies were grown overnight in 5 mL LB in a 15 mL falcon tube at 37°C, then diluted 1:100 into 10 mL M9 + 0.5% glycerol (unless otherwise noted) in a 50 mL flask, grown at 30°C with 250 rpm shaking until an OD_{600} of 0.4 – 0.6 was reached, and induced with 1 mM IPTG. Cells were harvested 24 h or 48 h later, for even- and odd-chain production experiments respectively. For 96-well plates, individual colonies were inoculated into 1 mL LB with appropriate antibiotics in a 2 mL deep-well plate (Thermo Scientific), and shaken at 30°C, 1200 rpm on a Titramax 1000 platform shaker (Heidolph) for 12-18 hours. Cultures were then diluted 1:20 into M9 + 0.5% glucose and grown for 3.5 hours before induction with 1 mM IPTG (unless otherwise noted). Growth was allowed to proceed for 24 or 44 hours before harvesting and analysis.

Western Blotting

Strains S007 and S002 were streaked out on an LB amp plate, grown overnight, and single colonies were picked into 5 mL LB/amp. LB cultures were allowed to grow for 18 h at 30 °C and 300 µl diluted into 6 mL M9 0.5% glucose with 50 µg/mL ampicillin. M9 cultures were grown for 3.5 h at 30 °C and either induced (+ IPTG) or not (- IPTG) with 1 mM IPTG. 1 mL samples were taken just before induction and 2, 4, 6, and 8 h after induction. 100 µL of each 1

mL sample was used to measure OD_{595}. The remaining 900 µL was centrifuged at top speed for 10 min in a table top centrifuge, the cell pellet resuspended in 100 µL 3% SDS and boiled for 10 min at 100 °C. Samples volumes used in western blotting were normalized to OD_{595} and prepared in 1x Tris-Glycine SDS Sample buffer (Novex) with 1x NuPAGE® sample reducing agent (Novex). Samples were run in a Novex 4-20% tris-glycine gel for 1.5 h at 120 V, transferred to a nitrocellulose membrane using the iblot transfer system (Novex), and blotted with the Strep•Tag® II Antibody HRP Conjugate at a 1:4000 dilution in TBS-tween with 1% BSA (EMD Millipore). Band intensities were quantified in Image J [29].

P1 Transduction

Each Keio collection knockout strain contains an insertion in a given gene, replacing most of its coding sequence with a kanamycin cassette, bracketed by FRT sites. Phage generated from a given Keio strain was used to transduce the desired strain as previously described [30]. Successfully transduced strains were streaked out serially 3x on LB + 5 mM sodium citrate plates to eliminate phage contamination. These were then transformed with the temperature-sensitive FLP recombinase plasmid pCP20 [31] and plated at 30°C on LB-Amp plates. Re-streaking the transformants at 37°C on LB simultaneously cured the plasmid and induced expression of FLP recombinase, excising the kanamycin cassette as previously described [31] . Strains streaked out at 37°C were re-streaked on both LB + Kan and LB + Amp plates to test for retention of either the kanamycin cassette or the pCP20 plasmid; colonies with neither resistance marker were then PCR amplified at the appropriate locus, and the PCR product sequenced to ensure successful knockout or allelic replacement.

Generation of Ketoacyl-Synthase Mutant Strains

To generate the *fabF** mutant [24], *fabF* was cloned into the BamH1 and HindIII sites of cloning vector pUC19 to generate plasmid pUC19-FabF*. Site directed mutagenesis was performed with primers TF0111 and TF0113 using the QuikChange® Multi Site-Directed Mutagenesis Kit (Agilent Technologies Inc) according to the manufacturers' instructions. *fabf** was amplified using primers TF0121 and TF0122. kanR was then amplified using primers TF0119 and TF0120 which both FRT flanked kanR and added 20bp of homology to *fabF* to the 5' end of kanR and 40bp of homology to the genomic region downstream of *fabF* to the 3' end of kanR. The *fabF** and kanR constructs were then used in a third PCR reaction containing TF0121 and TF0120. The product of this reaction was the portion of *fabF** containing the I-F mutation linked to a frt flanked kanR.

Attempts to use lambda red mediated recombineering to insert this construct directly into the *E. coli* strain TB10 [32] as previously described [33] failed to result in strains containing the *fabF* I-F mutation. A region of *fabF* starting 48 bp upstream of the I-F mutation and ending 40bp downstream of the *fabF* gene was therefore first replaced with an ampicillin resistance gene (*bla*) amplified from plasmid pETDuet-1 using primers TF0134 and TF0135. *bla* was then replaced with the *fabF**_kanR cassette in TB10 and clones that could grow on kanamycin and not ampicillin were sequenced for the *fabF* I-F mutation. A strain positive for the mutation was then used to generate P1 phage and the phage used to insert the *fabF** mutation into the strains indicated in the text. Although sequencing results showed that the *fabF**_kanR construct was properly incorporated into the genome, the kanR cassette could not be removed using the pCP20 plasmid.

To generate *fabBDeg*, a DNA construct composed (from 5' to 3') of 40bp of the 3' end of *fabB*, an ssrA DAS+4 tag [25], an FRT-flanked kanR and 40bp downstream of the genomic copy of *fabB* was synthesized (Intergraded DNA Technologies). This construct was PCR amplified using primers TF0101 and TF0102. Purified PCR products were then used for lambda red mediated recombination in *E. coli* strain TB10 as previously described [32,33]. The genomic copy of *fabBDeg* with the downstream kanR was then P1 transduced into the *E. coli* strains indicated in the text. The FRT-flanked kanamycin cassette was then excised using plasmid pCP20 [31].

To construct Strain S007, a linear DNA fragment containing N-terminally strep-tagged FabBDeg with a C-terminal kanamycin resistance cassette was constructed by PCR of the DAS+4 tagged FabB from genomic DNA from S006 using primers TF0102 and TF0222. Once made, this construct was inserted into the TB10 genome and transferred into strain BL21*(DE3) *ΔfadE* using P1 transduction as previously described.

Fatty Acid Identification and Quantification using GC-MS

Fatty acids were extracted from 400 µL of culture by adding 50 µL 10% NaCl, 50 µL glacial acetic acid, 20 µL of 800 mg/L pentadecanoate (as internal standard) and 200 µL ethyl acetate, then vortexing for 20 s. These were centrifuged at 16,000 g for 10 min. Ethyl esters were generated by mixing 100 µL of the organic phase with 900 µL of a 30:1 mixture of EtOH and 37% HCl, and incubated at 55°C for 1 h (based on protocol in [34]). After cooling to room temperature, 500 µL dH_2O and 500 µL hexane were added. The mixture was vortexed for 10 s, and 150 µL taken from the top (hexane) layer for analysis via GCMS. Extracts were run on an Agilent GCMS 5975/7890 (Agilent Technologies Inc) using an HP-5MS (length: 30m, diameter:

either 0.25 or 0.50 mm, film 0.25um) column; the method ramped from 75 to 325°C at 30°C per min. Compound identities were determined via GC retention times (compared to known standards) and verified by mass spectra. Where quantification was necessary, standard curves of the quantified compounds were generated by extracting them from M9 and esterifying them as described. The area under the each peak of a single-ion trace selective for ethyl esters (88 m/z) was integrated and normalized to pentadecanoate before using it to quantify the desired species.

Estimation of lipid-incorporated fatty acids

The amount of fatty acid incorporated into lipids was estimated from OD_{600} at 24 h by assuming the following: an OD_{600} of 1.0 equals 0.3 gDCW/L [35]; 9.1% of *E. coli* dry cell weight is lipid [36]; fatty acids make up 71% of lipid mass (calculated as the % molecular weight of dipalmitoyl phosphatidylglycerol that comes from its palmitic acid moieties).

Results

Our strategy for engineering MCFA synthesis was to first demonstrate the production of MCFAs with all even- and odd-chain lengths from 4-13 carbons, then identify factors that limit MCFA yield. After finding that elongation rates were too rapid for optimal MCFA production, we genetically engineered fatty acid synthesis to slow elongation in response to a chemical inducer, and thereby direct fatty acid synthesis toward the production of a specific MCFA, octanoic acid. Finally, we tested a set of knock-out mutants for their ability to increase carbon flux into fatty acid synthesis and maximize yields in our engineered strain.

Comprehensive production of even- and odd-chain MCFAs by engineering E. coli

We engineered production of MCFAs with 4-13 carbons in *E. coli* by varying the expressed thioesterase and adding propionate to the culture medium. Three thioesterases capable of hydrolyzing a range of medium-chain acyl-ACPs were chosen: BfTES (EET61113 from *B. formatexigens* [21]), CpFatB1 [20] and UcFatB2 [18]. Strain S001 (BL21*(DE3) Δ*fadD*, Table 2.1) was transformed with one of three plasmids encoding codon-optimized versions of these thioesterases (pEET, pTJF010 and pJT208 respectively), grown in M9 + 0.5% glycerol with no added propionate, and induced with IPTG for 24 h (Materials and Methods). GCMS analysis of supernatant free fatty acids showed that thioesterase expression increased the total moles of fatty acid produced in each case (Figure 2.1), and that each even-chain fatty acid from 4-12 carbons was produced by at least one thioesterase, with BfTES, CpFatB1 and UcFatB2 producing primarily butanoic, octanoic, and dodecanoic acids respectively (Figure 2.2B, Figure 2.1). Propionate supplementation has previously been shown to cause odd-chain fatty acid production via the incorporation of propionyl-CoA into the initial step of fatty acid synthesis [14,19] (Figure 2.2A). Addition of 100 mM propionate to the culture medium caused all three thioesterases to produce odd-chain fatty acids, with each odd-chain fatty acid in the 5-13 carbon range produced by at least one thioesterase (Figure 2.2B). These experiments demonstrated that the full range of 4-13 carbon MCFAs could be produced by engineered *E. coli*.

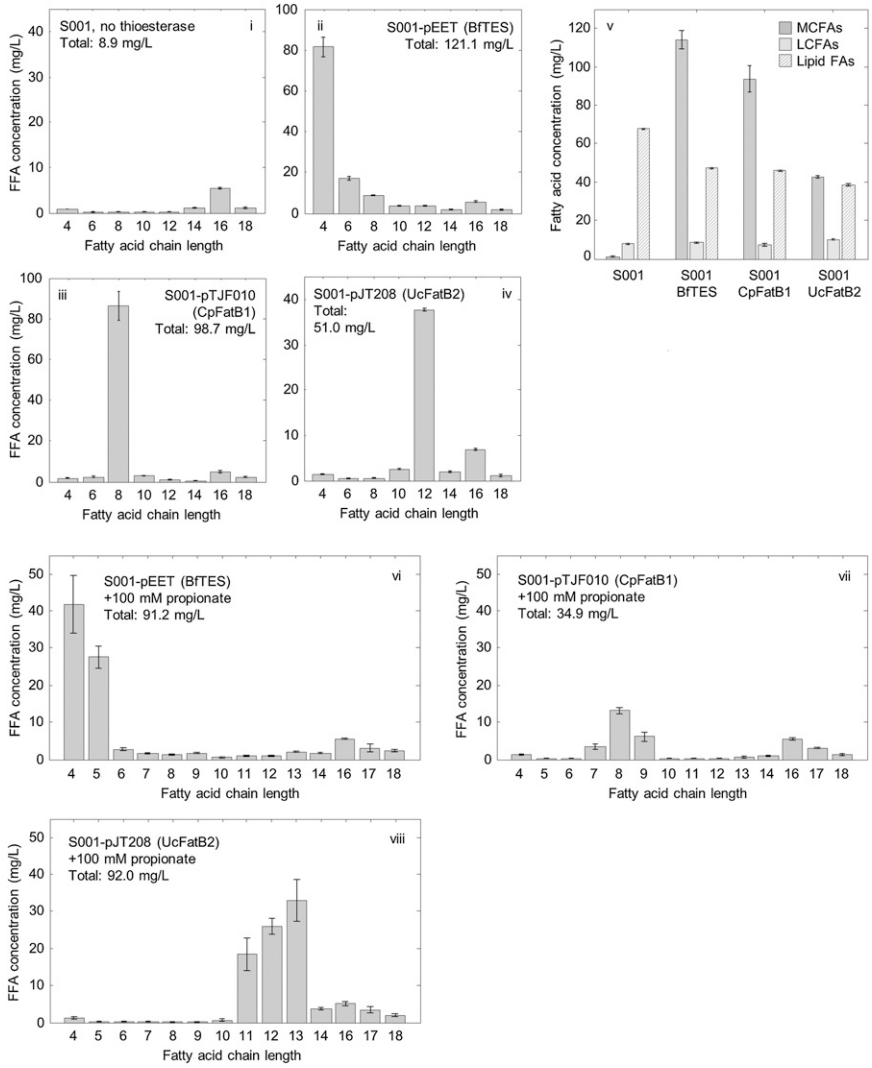

Figure 2.1. Fatty acid chain-length profile of medium-chain fatty acid (MCFA)-producing strains. Strain S001 alone (i) or bearing either pEET (BfTES, ii), pTJF010 (CpFatB1, iii), or pJT208 (UcFatB2, iv) was grown in M9 + 0.5% glycerol and induced with IPTG for 24 h (n = 3, error bars = SEM). Extraction and esterification of free fatty acids (FFA) with ethanol, as well as detection by GC-MS and quantification, were performed as described in Materials and Methods. The FFA chain- length profile is shown for (i) S001, (ii) S001-pEET, (iii) S001-pJTF010, and (iv) S001-pJT208; odd chains were produced at negligible quantities (<2 mg/L) and are not shown. Analysis of the fate of all produced fatty acids, including those incorporated into lipids, is shown in v. MCFAs refers to all FFAs from 4 to 12 carbons, but long-chain fatty acids (LCFAs) refers to all FFAs from 14 to 18 carbons. The concentration of lipid-incorporated fatty acids was estimated as described in Materials and Methods. Panel v. shows that thioesterase expression increases the total amount of carbon flowing into fatty acid synthesis, as the sum of MCFAs + LCFAs + Lipid-FAs increases. This result is primarily because of an increase in MCFA yield. Panels vi to viii show the FFA chain-length profile for S001- pEET (vi), S001-pTJF010 (vii), and S001-pJT208 (viii) during supplementation with 100 mM propionate.

Odd-chain fatty acids could be produced without the need for propionate supplementation by genetically encoding a propionyl-CoA production pathway. Building on recent work in which endogenous propionyl-CoA production was engineered to synthesize a series of valuable small molecules [15,37], we transformed strain S001-pJT208 with the plasmid pCOLA-thrA^{fr}BCilvA^{fr} [15], which produces propionyl-CoA primarily by increasing flux through the isoleucine biosynthetic pathway (Figure 2.2A). Strain S001-pJT208-pCOLA-thrA^{fr}BCilvA^{fr} produced the odd-chain fatty acids undecanoate and tridecanoate from glycerol as a sole carbon source (Figure 2.2C), with similar selectivity to S001-pJT208 in the presence of propionate (Figure 2.2B, iii). Thus, *E. coli* can be engineered to produce both even- and odd-chain fatty acids from a single carbon source, without the need for propionate supplementation.

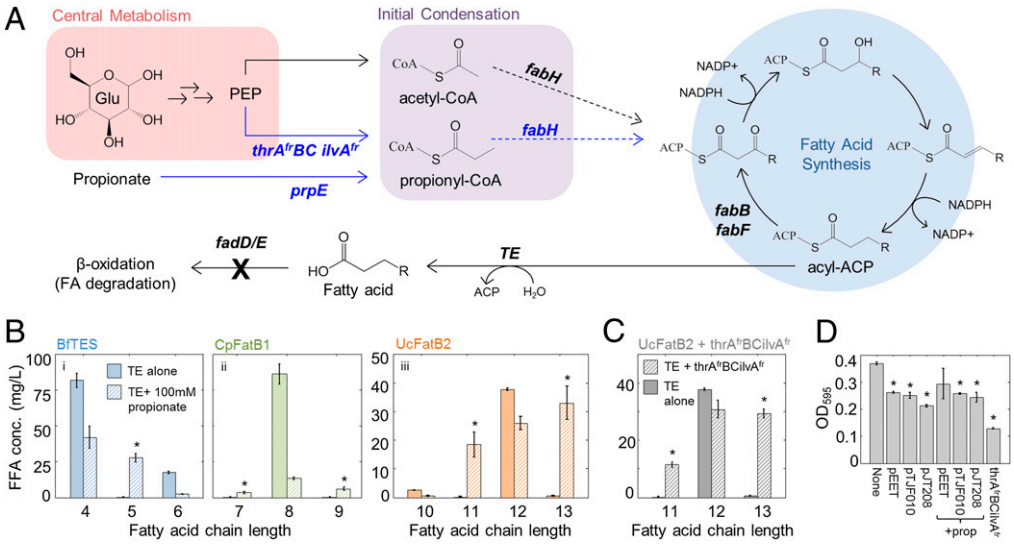

Figure 2.2. Engineering production of all even- and odd-length Medium-Chain Fatty Acids (MCFAs) in *E. coli*. (A) The ketoacyl synthase (KAS) FabH elongates acetyl-CoA to form a four-carbon β-ketoacyl ACP, which is reduced to C4 acyl-ACP. In subsequent rounds of fatty acid synthesis, the KAS's FabB and FabF elongate acyl-ACPs 2 carbons at a time to yield a range of even-chain-length fatty acyl-ACPs. Incorporation of propionyl-CoA in place of acetyl-CoA causes production of odd-chain acyl-ACPs; propionyl-CoA can be produced from propionate (*prpE*) or from expression of a genetic cassette that increases flux through the isoleucine pathway (*thrA^{fr}BC ilvA^{fr}*) [15]**.** Acyl-ACPs can be hydrolyzed to free fatty acids (FFAs) by an appropriate thioesterase (TE). Fatty acid degradation can be blocked by knocking

Figure 2.2 (Continued). out the β-oxidation enzymes *fadD* or *fadE*. (B) GCMS analysis of FFA production by strain S001 (BL21*(DE3) Δ*fadD*) containing plasmid pEET (BfTES) (i), pTJF010 (CpFatB1) (ii) or pJT208 (UcFatB2) (iii) in M9 +0.5% glycerol alone (solid bars) or supplemented with 100 mM propionate (striped bars), 24 hours after IPTG induction (N=3, error bars=SEM). FFAs shown for the no-propionate experiments accounted for 67% (i), 85% (ii) and 72% (iii) of total FFAs (full chain-length profiles shown in Figure 2.1). FFAs shown for the propionate experiments accounted for (i) 79%, (ii) 65% and (iii) 84% of the total. FFAs shown for the no-propionate odd-chain experiment accounted for 81% of total. (C) Production of odd-chain FFAs by S001-pJT208-pCOLAthrA^{fr}BCilvA^{fr} (N=3 error bars=SEM). Asterisks in (B) and (C) indicate significantly increased odd-chain production ($p<0.05$, one-tailed student's t-test). (D) Final OD_{595} of strains in this figure, with or without propionate supplementation (+prop) (N=3, error bars=SEM). Asterisks indicate significantly decreased OD_{595} compared to strain S001 alone ($p<0.05$, one-tailed student's t-test).

Wild-type fatty acid elongation rates are non-optimal for MCFA production

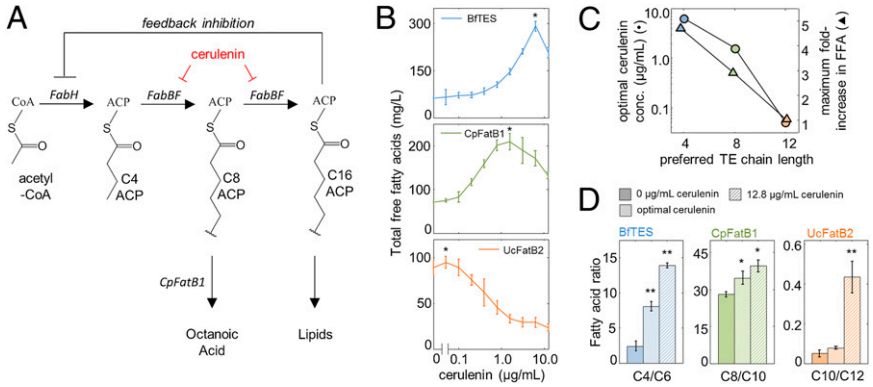

Figure 2.3. Chemical inhibition of fatty acid elongation can increase MCFA production. (A) Schematic diagram of fatty acid elongation. FabH performs the slow initial elongation step in fatty acid biosynthesis to generate C4-ACP; FabB and FabF more rapidly elongate the acyl-ACP 2 carbons at a time to produce 16-18 carbon acyl-ACPs, which are precursors to lipid synthesis and inhibit FabH more strongly than other acyl-ACPs [16,17]. Cerulenin inhibits FabB and FabF, but not FabH. (B) Total free fatty acids (FFA) produced by strain S001 expressing the indicated thioesterases over a range of cerulenin concentrations, 24 hours after IPTG induction, as measured by GCMS (N=3, error bars=SEM). Long-chain FFAs (C14-C18) accounted for ≤ 12 mg/L in any given measurement. Asterisks indicate the point at which mean fatty acid production is highest; the increase in production is significant for BfTES and CpFatB1 ($p<0.05$, one-tailed student's t-test), but not for UcFatB2. OD_{595} data are provided in Figure 2.5. (C) Filled circles: cerulenin concentration at which fatty acid production is maximal as a function of each thioesterase's preferred chain length. C4 for BfTES (blue); C8 for CpFatB1 (green); C12 for UcFatB2 (orange). Filled triangles indicate the maximum fold increase in FFA production over the no-cerulenin control. (D) Ratio of the two most abundant FFAs produced by each thioesterase (shorter/longer), as a function of cerulenin concentration (N=3, error bars=SEM). Single asterisks indicate that the given bar is significantly different from the no-cerulenin control

Figure 2.3 (continued). with $p < 0.05$ (two-tailed student's t-test). Double asterisks indicate that the given bar is significantly different from both other bars in the panel.

One challenge in MCFA production is the low concentration of medium-chain acyl-ACPs in growing cells [22,23]. Treatment of *E. coli* with cerulenin, an antibiotic that targets the KAS enzymes FabB and FabF, inhibits acyl-ACP elongation without inhibiting the initial condensing enzyme FabH (Figure 2.3A) and causes accumulation of medium-chain acyl-ACPs *in vivo* [22]. This accumulation is likely due to two factors: (i) decreased elongation of medium-chain acyl-ACPs to long-chain acyl-ACPs, and (ii) the resulting decrease in long-chain acyl-ACPs, which would relieve feedback inhibition of FabH [16,17] and increase flux into fatty acid synthesis (Figure 2.3A). We hypothesized that there should be an optimal level of cerulenin at which elongation is slow enough to increase the concentration of medium-chain acyl-ACPs and rate of MCFA production, but not so slow that elongation itself becomes rate-limiting.

Adding cerulenin to MCFA-producing cultures demonstrated that free fatty acid (FFA) yields could be increased by inhibiting elongation (Figure 2.3B). Yield increases came primarily from MCFAs rather than other fatty acids, as long-chain fatty acids were detected at < 12 mg/L in all samples, and biomass generally did not change with cerulenin concentration (Figure 2.4). Cerulenin also increased the MCFA production rate (Figure 2.4A). Assuming MCFA production rate increases with the concentration of its substrate acyl-ACPs, and assuming cerulenin does not directly affect thioesterase activity, this result is consistent with an increase in the cellular concentration of medium-chain acyl-ACPs. These data suggested that wild-type fatty acid elongation rates are not optimal for the production of MCFAs in our strains and demonstrated, counter-intuitively, that inhibiting fatty acid synthesis can increase the production of fatty acids.

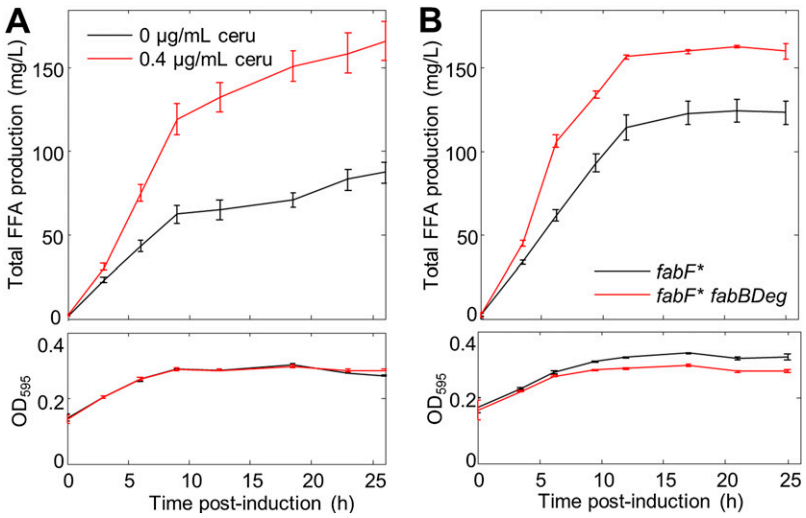

Figure 2.4. Time courses of free fatty acid (FFA) production with and without KAS inhibition. (A) Strain S001-pTJF010 (CpFatB1) was grown in M9 + 0.5% glucose and induced with 1 mM IPTG as previously described, with or without the addition of 0.4 µg/mL cerulenin. At the indicated time points (0 h = time of induction), cultures were tested for OD_{595} and for total FFAs via the Roche Half-Micro test (N=3, error bars=SEM). At all time points other than t=0, samples with added cerulenin had greater FFA production than those without cerulenin ($p < 0.05$, one-tailed student's t-test). The initial rate of FFA production (first four data points) was also greater in the samples containing cerulenin: 6.7 ± 0.6 $mgL^{-1}h^{-1}$ for 0 µg/mL cerulenin versus 13.2 ± 2.9 $mgL^{-1}h^{-1}$ for 0.4 µg/mL cerulenin. (B) Time course (performed as in A) of strains S003 (BL21*(DE3) Δ*fadE fabF**) and S005 (BL21*(DE3) Δ*fadE fabF* fabBDeg*) both expressing SspBpET21b and pTJF038. At all time points other than t=0, fatty acids produced by S005 were significantly greater than those produced by S003 ($p < 0.05$, one-tailed student's t-test). The initial rate of fatty acid production (first four data points) was also greater: 9.6 ± 0.7 $mgL^{-1}h^{-1}$ for S003 versus 14.5 ± 2.4 $mgL^{-1}h^{-1}$ for S005.

The maximal increase in FFA yield was greater for shorter-chain-length-specific thioesterases, as was the amount of cerulenin needed to achieve it (Figure 2.3C); this is consistent with a model in which the distribution of acyl-ACPs is a function of elongation rate, with slower elongation rates yielding shorter acyl-ACP pools and enhancing production of shorter fatty acids. GCMS analysis offered further evidence for this model. First, cerulenin treatment caused all three thioesterases to shift their production toward shorter products (Figure 2.3D), which is most easily explained by an increase in the fraction of shorter acyl-ACPs in the cell. Second, although cerulenin decreased total and C12-specific FFA yields in the UcFatB2

strain, it increased the yield of C10 FFAs (Figure 2.6), suggesting differential effects of cerulenin on shorter and longer acyl-ACP pools.

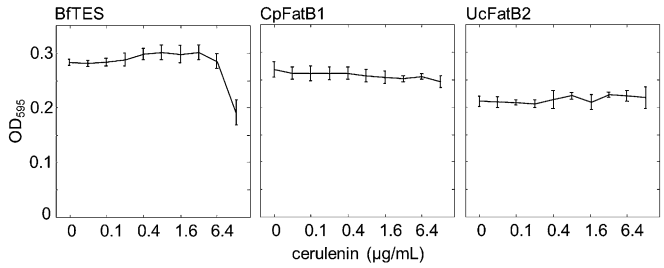

Figure 2.5. Final OD_{595} measurements from cerulenin titration experiments in Figure 2.3B. Experiments were performed as described in the caption of Figure 2.3B. OD_{595} was measured for each strain at each concentration of cerulenin tested (N=3, error bars=SEM) for the same samples used to generate Figure 2.3B. BfTES, CpFatB1 and UcFatB2 indicate strains S001-pEET, S001-pTJF010 and S001-pJT208 respectively.

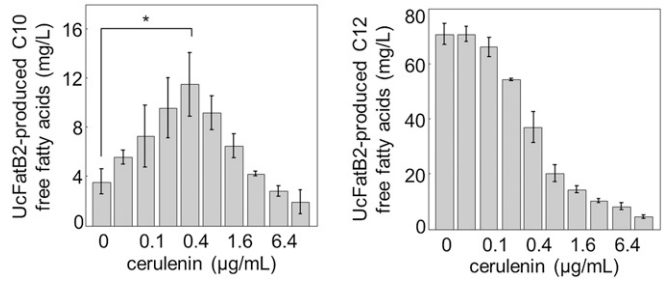

Figure 2.6. Production of C10 and C12 free fatty acids (FFAs) by UcFatB2 as a function of cerulenin concentration. Strain S001-pJT208 was grown in M9 + 0.5% glucose and induced with 1 mM IPTG and different concentrations of cerulenin for 24 hours (as described for Figure 2.3B in the text). Shown are the concentrations of C10 and C12 FFAs as a function of cerulenin concentration, as determined by GCMS analysis (N=3, error bars=SEM). Whereas C10 increases significantly from 0 to 0.4 0.4 µg/mL cerulenin (asterisk; $p<0.05$ by one-tailed student's t-test), C12 decreases consistently with increasing concentrations of cerulenin.

These results suggest that BfTES and CpFatB1's FFA yields increase in response to cerulenin treatment due to a shift of the acyl-ACP distribution toward shorter (4-10 carbon) acyl-ACPs, which are more effectively hydrolyzed by these thioesterases; FFA production by UcFatB2, however (which prefers C12 acyl-ACPs), may fail to increase for the same reason. These apparent shifts in the acyl-ACP distribution are consistent with the expectation that cerulenin would simultaneously increase the rate at which acyl-ACPs are initiated by FabH and

decrease the rate at which they elongate, and with previous results demonstrating that medium-chain acyl-ACPs accumulate in response to cerulenin treatment [22]. In all, these results suggest that inhibiting fatty acid elongation increases the concentration of medium-chain acyl-ACPs with ≤ 10 carbons, a useful strategy for increasing MCFA yield.

Targeted genetic production of medium chain fatty acids via inducible degradation

Having shown that inhibiting elongation can increase MCFA yield, we attempted to engineer production of a specific MCFA, octanoic acid, by implementing two genetic interventions: (i) replacing FabF with a mutant, FabF*, incapable of elongating beyond C8 due to the insertion of a bulky phenylalanine into its fatty acid binding pocket [24], and (ii) engineering FabB with a C-terminal SsrA DAS+4 tag, enabling inducible degradation by the *E. coli* ClpXP system via overexpression of the adaptor protein SspB [25,38] (Figure 2.7A). The inducible degradation system was necessary because *fabB* is required for unsaturated fatty acid synthesis, and is essential for growth on minimal media [39]. We hypothesized that FabF* would allow high rates of octanoyl-ACP production, while inducing degradation of FabB would slow elongation beyond C8 acyl-ACP, similar to the effect of cerulenin. This should cause octanoyl-ACP to accumulate and increase fatty acid yields in strains expressing CpFatB1, which primarily produces octanoic acid (Figure 2.1, Figure 2.2B).

Expressing CpFatB1 (pTJF038) and SspB (SspBpET21b) [25] in strain S003 (BL21*(DE3)*ΔfadE fabF**) did not result in significantly increased FFA yields over those of S002 (BL21*(DE3)*ΔfadE*) (Figure 2.7C), suggesting that the octanoyl-ACP generated by FabF* was not accumulating sufficiently to increase octanoate production, likely due to elongation by FabB. We therefore appended an SsrA DAS+4 degradation tag [25] to the endogenous *fabB* in both S002 and S003, generating strains S004 (BL21*(DE3) *ΔfadE fabBDeg*) and S005

(BL21*(DE3) *ΔfadE fabF* fabBDeg*) respectively. SspB-induced degradation of FabBDeg was demonstrated via western blot of strain S007, a derivative of S004 in which FabBDeg is appended with an N-terminal strep tag to facilitate detection (Figure 2.7B, Figure 2.8).

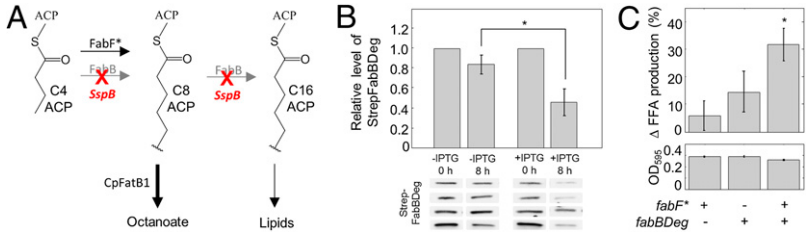

Figure 2.7. Engineering ketoacyl synthases for increased octanoic acid production. (A) Schematic indicating the elongation reactions carried out by ketoacyl synthases FabB and FabF. Wild-type FabB and FabF elongate acyl-ACPs with ≥ 4 carbons up to a length of 16-18 carbons. In our engineered system, FabBDeg can elongate up to 16-18 carbons but is degraded by the *E. coli* ClpXP system upon SspB expression. FabF* can only elongate up to 8 carbons. (B) StrepFabBDeg degradation in strain S007-SspBpET21b. Strain S007-SspBpET21b was induced (+IPTG, 1 mM) or not (-IPTG) at time 0 and StrepFabBDeg detected via western blotting and densitometry at 0 and 8 h (N=4, error bars=SEM). Loading was normalized to OD_{600}, and reported values are normalized to StrepFabBDeg levels at t = 0 h. StrepFabBDeg bands are shown below the bar graph; original blots are shown in Figure 2.8. * Indicates a significant decrease in StrepFabBDeg between the -/+ IPTG samples at 8 h ($p<0.05$, one-tailed student's t-test). (C) Free fatty acid (FFA) production (as determined by the Free Fatty Acids, Half Micro Test (Roche)) and final culture OD_{595} 44 h after induction, for strains expressing CpFatB1 and SspB in S002 (BL21*(DE3) *ΔfadE*) with the indicated modifications (N=24; error bars=SEM). FFA production is represented as the % increase over FFA production in the parent strain (S002). *Indicates a significant increase in FFA production compared to S002 ($p<0.05$ by one-tailed student's t-test).

Inducing SspB and CpFatB1 expression in strain S005, which contains both *fabF** and *fabBDeg*, resulted in a significant, 32% increase in FFA yield to 118 mg/L and a 12% decrease in final OD_{595} compared to parent strain S002 (Figure 2.7C). This suggests that FabB degradation shunts fatty acid synthesis to FabF* and increases the pool of octanoyl-ACP for hydrolysis by CpFatB1. Consistent with this, time courses showed that strain S005 (*fabF* fabBDeg)* had a greater rate of FFA production than strain S003 (*fabF** alone) (Figure 2.4B). Importantly, because *fabBDeg* alone (S004) does not significantly alter yield or OD_{595} compared with the parent strain (S002), it is unlikely that the yields observed in the *fabF* fabBDeg* strain

(S005) are due to some general effect on growth or fatty acid synthesis caused by SspB synthesis or FabB degradation.

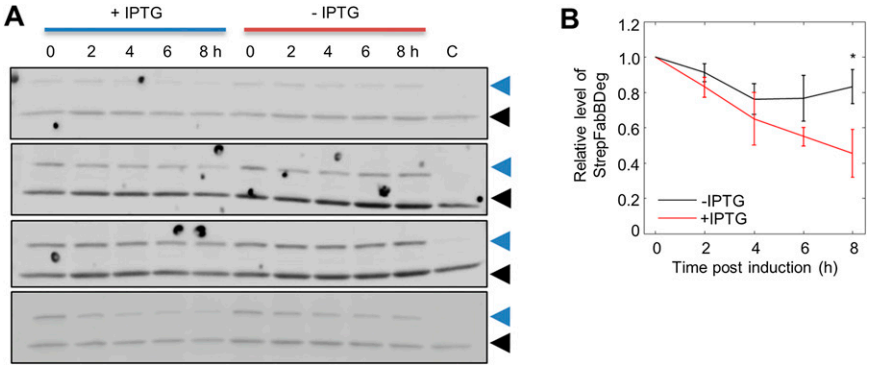

Figure 2.8. Western blotting of StrepFabBDeg following SspB induction. Strain S007 (BL21*(DE3) *ΔfadE strepfabBDeg*) + SspBpET21b was grown in M9 +0.5% glucose and either induced (+IPTG) or not induced (-IPTG) with 1 mM IPTG. StrepFabBDeg levels were monitored over time as described in Materials and Methods (numbers indicate the time in hours post-induction); loading was normalized to OD_{600}. Blue triangles indicate StrepFabBDeg while black triangles indicate an unidentified background band. The band corresponding to StrepFabBDeg ran at the expected size (45 kDa) and did not appear in control strain S002 (lane C on the gel). Strain S002 encodes wild-type *fabB* without a strep or degradation tag, and was loaded at t=0 without IPTG induction. (B) Quantification of the western blot in (A), normalized to StrepFabBDeg levels at t = 0 h, shows that StrepFabBDeg levels drop about 54% of their original level over 8h of induction. * indicates significantly lower StrepFabBDeg levels compared to -IPTG at 8h ($p < 0.05$, one-tailed student's t-test).

Optimization of the engineered strain

We further optimized octanoic acid production in our engineered strain by: (i) screening a series of knockout mutations for their ability to increase yield, and (ii) titrating the level of SspB induction to optimize FabB degradation and free fatty acid production by CpFatB1.

We expressed CpFatB1 in BL21*(DE3) derivatives with knockouts in fatty acid degradation (*fadD, fadE*), fermentation of undesirable products (*pta*, *poxB*, *adhE*, *ldhA*, *pflB*) [3] and the stringent response (*relA*) which inhibits lipid synthesis during starvation [40], and measured the total free fatty acid (FFA) production (Figure 2.9) and growth (Figure 2.10). As strains with *fadD* and *pta* showed the greatest improvement in FFA among the single-knockout

strains (Figure 2.9A, green), we constructed all possible double mutants in a *ΔfadD* background (Figure 2.9A, blue), as well as all possible triple mutants in a *ΔfadD Δpta* background (Figure 2.9A, orange). The strain with the highest yield (*ΔfadD Δpta ΔldhA*) showed little increase in FFA production beyond that of the *ΔfadD Δpta* double knockout. *ΔfadD* and *Δpta* were therefore used to further improve FFA yields in our strains engineered with *fabF** and *fabBDeg*.

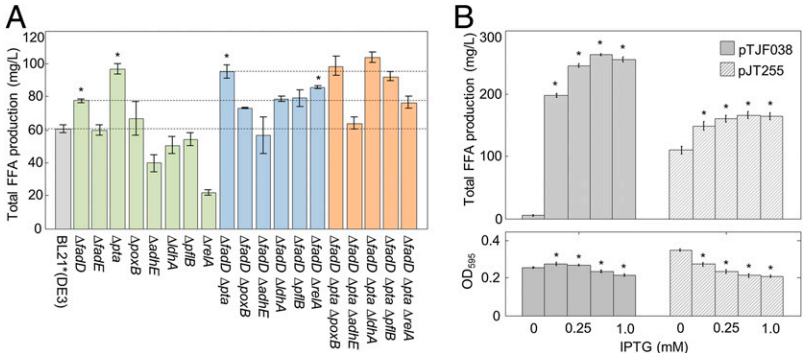

Figure 2.9. Optimizing octanoic acid yields via metabolic knockouts and SspB titration. (A) Total free fatty acid (FFA) production in BL21*(DE3) knockout strains. Each strain was grown in M9 + 0.5% glucose, CpFatB1 was induced from pTJF010 with 1mM IPTG for 44h, and total FFA measured with the Free Fatty Acids, Half Micro Test (Roche) (N=3, error bars=SEM). Green, blue, and orange bars indicate single knockouts, double knockouts and triple knockouts respectively. Dashed lines indicate fatty acid production in the parent strains BL21*(DE3), BL21*(DE3) Δ*fadD*, and BL21*(DE3) Δ*fadD* Δ*pta*. * indicates FFA production significantly greater than the parent strain ($p<0.05$ by one-tailed student's t-test). OD_{595} for these strains is shown in Figure 2.10. (B) FFA production and OD_{595} as a function of IPTG concentration (0, 0.125, 0.25, 0.5 or 1.0 mM). CpFatB1 and SspB expressed from pTJF038 and SspBpET21b respectively were co-titrated in S006 (BL21*(DE3)*ΔfadD Δpta ΔlacY fabF* fabBDeg*) (solid grey bars) with IPTG for 24h in M9 + 0.5% glucose (N=18, error bars=SEM). SspB alone was titrated with IPTG in S006 expressing SspB from SspBpET21b and CpFatB1 from aTc-inducible pJT255 induced with 400 ng/mL aTc (striped bars, N=18, error bars=SEM). * indicates significantly different FFA production, or altered OD_{595}, versus the 0 mM IPTG control ($p<0.05$ by two-tailed student's t-test).

Our improved strain S006 (BL21*(DE3) *ΔfadD Δpta ΔlacY fabf* fabBDeg*), which included a *ΔlacY* mutation to allow titratable IPTG induction [41], increased FFA production ~2-fold over S005 to 263mg/L (Figure 2.9B, solid bars) at the optimal level of IPTG induction of SspBpET21b and pTJF038 (CpFatB1). GCMS analysis showed that free fatty acid production by S006 was highly selective for octanoic acid (92.1± 0.2%) (Figure 2.11), with a >

130 mg/L increase in octanoate production over the non-optimized parent strain (S002). We estimated from OD_{595} data that S006 produced only 15 mg/L less lipid-bound long-chain fatty acid than S002 (Materials and Methods); our genetic interventions therefore increase the total moles of fatty acid produced primarily by increasing MCFA production.

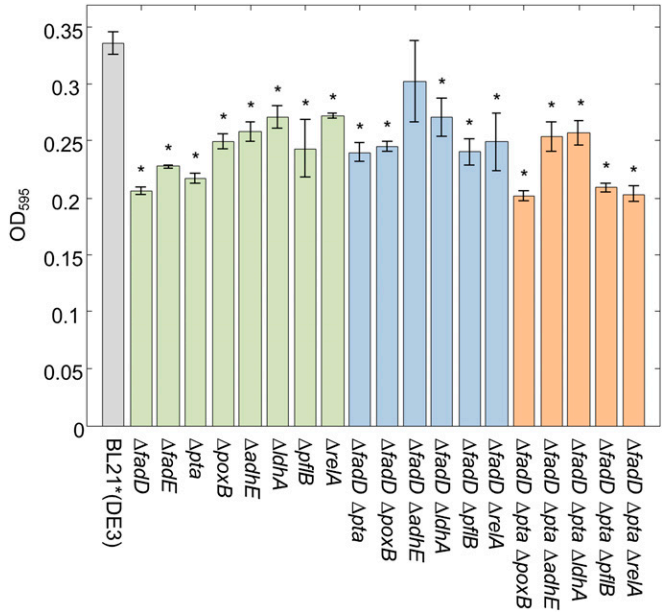

Figure 2.10. Biomass of knockout collection expressing CpFatB1. Strain BL21*(DE3) pTJF010 (CpFatB1) was constructed with the indicated mutations, grown in M9 + 0.5% glucose and induced. OD_{595} was measured 44 h after induction. This data was taken from the same samples tested for free fatty acids in Figure 2.9A. Green, blue and orange bars indicate single, double or triple knockout strains respectively. Asterisks indicate OD595 values significantly different from that of BL21*(DE3) pTJF010 ($p < 0.05$ by two-tailed student's t-test).

As IPTG induction controlled both CpFatB1 and SspB expression, it was unclear whether the increase in yield was due to altering CpFatB1 expression, altering degradation, or a combination of the two. Generating a strain in which we could induce CpFatB1 and SspB separately showed that titration of SspB could increase FFA yields compared to an un-induced control. Strain S006 was transformed with an anhydrotetracycline (aTc)-inducible, CpFatB1-expressing plasmid (pJT255) and an IPTG-inducible, SspB-expressing plasmid (SspBpET21b).

This strain was induced with 400 ng/mL aTc, and IPTG titrated from 0 to 1mM (Figure 2.9B, striped bars). FFA yields increased significantly at lower doses of IPTG (110 to 166 mg/L, from 0.125-0.5 mM IPTG) while OD_{595} decreased, but plateaued beyond this point. Our highest yield in this strain was 166 mg/L. Although this is lower than the yields achieved with TJF038 and SspBpET21b, these results nevertheless demonstrate that optimizing inducible degradation of FabB can increase yields over those achieved through more traditional knock out strategies, and that it is important to tune degradation of FabB to achieve optimal production while limiting effects on cell growth.

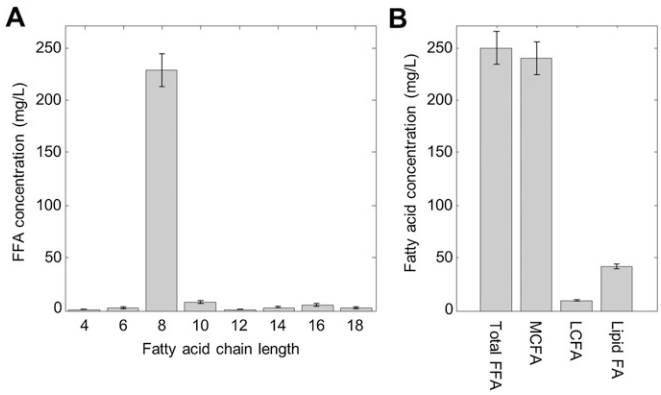

Figure 2.11. Fatty acid profile of optimized strain from Figure 2.9B. Strain S006-pTJF038-SspBpET21b was grown in M9 + 0.5% glucose and induced with 0.5 mM IPTG for 24 h. (A) Chain-length profile of free fatty acids. Octanoic acid represents 92% of total free fatty acid production in this strain. (B) Concentrations of total free fatty acids produced, free medium-chain fatty acids (MCFAs, C4-C12), free long-chain fatty acids (LCFAs, C14-C18), and lipid-bound fatty acids (Lipid FA). The latter was estimated from OD_{595} (Figure 2.9B) as described in Material and Methods.

Discussion

Medium-chain fatty acids (MCFAs) and their derivatives are important precursors to high-quality fuel and industrial molecules [3,8,9]. In this work, we engineered the production of a range of MCFAs, and showed that their yields increase in response to chemical inhibition of fatty acid elongation. By engineering ketoacyl synthases to selectively block long-chain acyl-

ACP elongation, we demonstrated that we can substantially increase the yield of a particular MCFA, octanoic acid. In principle, this strategy can be adapted for the production of any MCFA in the 4-13 carbon range.

In addition to even-chain MCFA production, we were able to produce odd-chain MCFAs, which are industrially useful as plasticizers, as herbicides, and in the fragrance industry [11,12,42]. Odd-chain fatty acid production has previously been achieved, but required propionate supplementation [14,19]. Endogenous propionyl-CoA production has also been achieved [15,37], but not in conjunction with fatty acid production. We synthesized these strategies to engineer production of odd-chain MCFAs from a single carbon source in the absence of propionate. Importantly, this was achieved using thioesterases known only to produce even-chain fatty acids, and some thioesterases produced odd chains better than others. Among the wide range of known even-chain thioesterases [21], some may therefore be capable of effectively or even selectively producing odd-chain fatty acids.

An important finding of our work was that application of the ketoacyl synthase (KAS) inhibitor cerulenin increased the yield of some MCFAs, suggesting that the optimal rate of acyl-ACP elongation for MCFA production is slower than the wild-type elongation rate. Our results suggest that this yield increase is due to accumulation of medium-chain acyl-ACPs, which may be a consequence of KAS inhibition both (i) slowing the loss of desired medium-chain acyl-ACPs to elongation, and (ii) indirectly increasing the rate of new acyl-ACP synthesis by FabH. While the goal of many metabolic engineering projects is to increase flux through a naturally-occurring pathway, our results suggest that in recursive systems like fatty acid synthesis, intermediate-length products may be produced more efficiently by slowing the native pathway.

We demonstrated that we could genetically engineer KAS's to inhibit unwanted elongation of medium-chain acyl-ACPs and thereby increase the yield of a specific MCFA, octanoic acid. While a large body of work has been amassed on attempts to increase flux to fatty acid precursors [8,43,44], to prevent fatty acid degradation [5,43], and to optimize the thioesterase [5,8] (reviewed in [3,4]) there have been few attempts to engineer the fatty acid synthesis machinery itself. Reasons for this may include the toxicity of FabF overexpression [45] and the essentiality of FabB on minimal medium [39]. Nevertheless, we were able to mimic the effect of cerulenin supplementation by coupling expression of the octanoyl-ACP-specific FabF* [24] with inducible degradation of FabB. These interventions presumably slowed the elongation of octanoyl-ACP produced by FabF*, thereby increasing octanoyl-ACP levels and octanoate production by CpFatB1. Through these means we achieved 12% theoretical yield of octanoate (Figure 2.9B). This strategy should be generally applicable to other MCFAs, though the optimal degradation rate and specificity of FabF is likely to be chain-length dependent.

Degradation tags have previously been used to decrease the basal level of metabolic enzymes [46] and to dynamically degrade regulatory proteins [47]. We engineered our strains to inducibly degrade an essential enzyme and thereby redirect metabolic flux on demand. Inducible degradation is an attractive metabolic engineering strategy because it allows an essential gene's activity to be decreased or knocked out entirely, without blocking the strain's ability to grow. Inducible degradation of FabB allowed us to achieve our highest yields. This approach should be generally applicable to all essential metabolic genes in *E. coli*, opening up additional opportunities for flux modulation.

This study presents a strategy to increase fatty acid production that combines ketoacyl synthase manipulation and more traditional metabolic engineering. It demonstrates the utility of

engineering the fatty acid synthesis machinery itself, and not just the thioesterase or other enzymes that feed into the pathway, to enhance the yield of specific fatty acids. Given the interest in producing diverse fatty acid-derived compounds for use as biofuels and industrial chemicals, it is likely that the engineering of KAS's and the other enzymes of fatty acid synthesis will become increasingly important in microbial fatty acid engineering. We expect the strategy of dynamically inhibiting essential enzymes to be broadly useful as a tool for metabolic engineering and synthetic biology. In particular, dynamically inhibiting polymer elongation may be a useful approach for increasing production or tuning the chain length of other valuable molecules synthesized by recursive biosynthetic pathways, such as iteratively synthesized polyketides, polysaccharides and bioplastics.

References

1. Cronan JE: **Bacterial membrane lipids: where do we stand?** *Annual review of microbiology* 2003, **57**:203-224.

2. Magnuson K, Jackowski S, Rock CO, Cronan JE, Jr.: **Regulation of fatty acid biosynthesis in Escherichia coli**. *Microbiological reviews* 1993, **57**:522-542.

3. Handke P, Lynch SA, Gill RT: **Application and engineering of fatty acid biosynthesis in Escherichia coli for advanced fuels and chemicals**. *Metab Eng* 2011, **13**:28-37.

4. Lennen RM, Pfleger BF: **Engineering Escherichia coli to synthesize free fatty acids**. *Trends in biotechnology* 2012, **30**:659-667.

5. Steen EJ, Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., Del Cardayre, S. B., Keasling, J. D.: **Microbial production of fatty-acid-derived fuels and chemicals from plant biomass**. *Nature* 2010, **463**:559-562.

6. Zhang F, Carothers JM, Keasling JD: **Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids**. *Nature biotechnology* 2012, **30**:354-359.

7. Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB: **Microbial biosynthesis of alkanes**. *Science* 2010, **329**:559-562.

8. Liu T, Vora H, Khosla C: **Quantitative analysis and engineering of fatty acid biosynthesis in E. coli**. *Metabolic engineering* 2010, **12**:378-386.

9. Ohlrogge JB: **Design of New Plant Products: Engineering of Fatty Acid Metabolism**. *Plant Physiol* 1994, **104**:821-826.

10. Yang ST: *Bioprocessing for value-added products from renewable resources: new technologies and applications*: Elsevier Science; 2006.

11. Braude GL: **Preparation of polymeric plasticizers from tall oil fatty acids**. Edited by: Google Patents; 1967.

12. Susumu O, Mieko I, Yoshio T: **Synthesis of various kinds of esters by four microbial lipases**. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism* 1979, **575**:156-165.

13. Knothe G: **Improving biodiesel fuel properties by modifying fatty ester composition**. *Energy & Environmental Science* 2009, **2**:759-766.

14. Ingram LO, Chevalier LS, Gabba EJ, Ley KD, Winters K: **Propionate-induced synthesis of odd-chain-length fatty acids by Escherichia coli**. *Journal of bacteriology* 1977, **131**:1023-1025.

15. Tseng HC, Prather KL: **Controlled biosynthesis of odd-chain fuels and chemicals via engineered modular metabolic pathways**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**:17925-17930.

16. Heath RJ, Rock CO: **Inhibition of-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in Escherichia coli**. *Journal of Biological Chemistry* 1996, **271**:10996-11000.

17. Jiang P, Cronan JE, Jr.: **Inhibition of fatty acid synthesis in Escherichia coli in the absence of phospholipid synthesis and release of inhibition by thioesterase action**. *Journal of bacteriology* 1994, **176**:2814-2821.

18. Voelker TA, Davies, H.M.: **Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of plant medium-chain acyl-acyl carrier protein thioesterase**. *Journal of bacteriology* 1994, **176**:7320-7327.

19. Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R: **Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals**. *Nature* 2011, **476**:355-359.

20. Dehesh K, Edwards, P., Hayes, T., Cranmer, a M., & Fillatti, J. : **Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of Cuphea palustris seed oil**. *Plant physiology* 1996, **110**:203-210.

21. Jing F, Cantu DC, Tvaruzkova J, Chipman JP, Nikolau BJ, Yandeau-Nelson MD, Reilly PJ: **Phylogenetic and experimental characterization of an acyl-ACP thioesterase family reveals significant diversity in enzymatic specificity and activity**. *BMC biochemistry* 2011, **12**:44.

22. Jackowski S, Rock CO: **Acetoacetyl-acyl carrier protein synthase, a potential regulator of fatty acid biosynthesis in bacteria**. *The Journal of biological chemistry* 1987, **262**:7927-7931.

23. Rock CO, Jackowski S: **Regulation of phospholipid synthesis in Escherichia coli. Composition of the acyl-acyl carrier protein pool in vivo**. *The Journal of biological chemistry* 1982, **257**:10759-10765.

24. Val D, Banu, G., Seshadri, K., Lindqvist, Y., & Dehesh, K. : **Re-engineering ketoacyl synthase specificity**. *Structure* 2000, **8**:565-566.

25. McGinness KE, Baker, T. A., Sauer, R. T.: **Engineering controllable protein degradation**. *Molecular cell* 2006, **22**:701-707.

26. Way JC, Davis JH: **Methods and Molecules for Yield Improvement Involving Metabolic Engineering**. Edited by: Google Patents; 2010.

27. Staunton J, Weissman KJ: **Polyketide biosynthesis: a millennium review**. *Nat Prod Rep* 2001, **18**:380-416.

28. Suriyamongkol P, Weselake R, Narine S, Moloney M, Shah S: **Biotechnological approaches for the production of polyhydroxyalkanoates in microorganisms and plants - a review**. *Biotechnol Adv* 2007, **25**:148-175.

29. Schneider CA, Rasband WS, Eliceiri KW: **NIH Image to ImageJ: 25 years of image analysis**. *Nature methods* 2012, **9**:671-675.

30. Thomason LC, Costantino N, Court DL: **E. coli genome manipulation by P1 transduction**. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* 2007, **Chapter 1**:Unit 1 17.

31. Cherepanov PP, Wackernagel W: **Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant**. *Gene* 1995, **158**:9-14.

32. Johnson JE LL, Hale CA, de Boer PA.: **ZipA is required for targeting of DMinC/DicB, but not DMinC/MinD, complexes to septal ring assemblies in Escherichia coli.** *Journal of bacteriology* 2004, **186**:2418-2419.

33. Datsenko KA, Wanner BL: **One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products**. *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**:6640-6645.

34. Ichihara K, Fukubayashi Y: **Preparation of fatty acid methyl esters for gas-liquid chromatography**. *J Lipid Res* 2010, **51**:635-640.

35. Panula-Perala J, Siurkus J, Vasala A, Wilmanowski R, Casteleijn MG, Neubauer P: **Enzyme controlled glucose auto-delivery for high cell density cultivations in microplates and shake flasks**. *Microb Cell Fact* 2008, **7**:31.

36. Neidhardt FC, Ingraham JL, Low KB, Magasanik B, Schaechter M, Umbarger H: *Escherichia coli and Salmonella typhimurium. Cellular and molecular biology. Volumes I and II*: American Society for Microbiology; 1987.

37. Tseng HC, Harwell CL, Martin CH, Prather KL: **Biosynthesis of chiral 3-hydroxyvalerate from single propionate-unrelated carbon sources in metabolically engineered E. coli**. *Microb Cell Fact* 2010, **9**:96.

38. Farrell CM, Grossman AD, Sauer RT: **Cytoplasmic degradation of ssrA-tagged proteins**. *Molecular microbiology* 2005, **57**:1750-1761.

39. Cronan JE, Jr., Birge CH, Vagelos PR: **Evidence for two genes specifically involved in unsaturated fatty acid biosynthesis in Escherichia coli**. *Journal of bacteriology* 1969, **100**:601-604.

40. Heath RJ, Jackowski S, Rock CO: **Guanosine tetraphosphate inhibition of fatty acid and phospholipid synthesis in Escherichia coli is relieved by overexpression of glycerol-**

3-phosphate acyltransferase (plsB). *Journal of Biological Chemistry* 1994, **269**:26584-26590.

41. Jensen PR WH, Michelsen O: **The use of lac-type promoters in control analysis**. *European Journal of Biochemistry* 1993, **211**:181-191.

42. Miller TW, Felsot A, Racke K: **Natural herbicides and amendments for organic weed control**. In *Crop protection products for organic agriculture. ACS Symposium Series*: ACS Publications: 2006:174-175.

43. Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD: **Enhancing fatty acid production by the expression of the regulatory transcription factor FadR**. *Metabolic engineering* 2012, **14**:653-660.

44. Davis MS, Solbiati J, Cronan JE, Jr.: **Overproduction of acetyl-CoA carboxylase activity increases the rate of fatty acid biosynthesis in Escherichia coli**. *The Journal of biological chemistry* 2000, **275**:28593-28598.

45. Subrahmanyam S, Cronan JE, Jr.: **Overproduction of a functional fatty acid biosynthetic enzyme blocks fatty acid synthesis in Escherichia coli**. *Journal of bacteriology* 1998, **180**:4596-4602.

46. Doroshenko VG, Shakulov RS, Kazakova SM, Kivero AD, Yampolskaya TA, Mashko SV: **Construction of an L-phenylalanine-producing tyrosine-prototrophic Escherichia coli strain using tyrA ssrA-like tagged alleles**. *Biotechnology letters* 2010, **32**:1117-1121.

47. Huang D, Holtz WJ, Maharbiz MM: **A genetic bistable switch utilizing nonlinear protein degradation**. *J Biol Eng* 2012, **6**:9.

Chapter 3: Adaptive evolution to enhance *E. coli* growth on octanoic acid

Abstract

Medium chain fatty acid (MCFA) catabolism shares the same first steps as the conversion of MCFAs into alcohols and alkanes. *E. coli* grow poorly on MCFAs even in the presence of regulatory mutations that enhance the expression of the fatty acid degradation (β-oxidation) machinery. Herein, we evolve *E. coli* for rapid growth on the MCFA, octanoate, with the goal of discovering mutations that enhance MCFA catabolism and that may be useful for the production of MCFA derived biofuels. We identify two mutations that enhance the expression of acyl-CoA synthetase FadD, the first enzyme in β-oxidation, and a third mutation that positively impacts growth on a variety of carbon sources and is unlikely to be directly related to fatty acid catabolism. This work points to FadD activity as an important limitation on MCFA catabolism and guides future efforts to alter the activity of this enzyme.

Introduction

E. coli grow slowly on the 8-carbon fatty acid octanoate. Octanoate catabolism shares the same first steps as pathways for its conversion into the potential next generation biofuels heptane and octanol (Figure 3.1) [1-5]. The β-oxidation repressor, FadR represses transcription of fatty acid catabolic enzymes, but fails to bind its cognate promoters in the presence of fatty acyl-CoA's 12 carbons and longer resulting in the induction of β-oxidation genes by these long chain fatty acyl-CoAs [6,7]. Fatty acyl-CoAs are produced by acyl-CoA synthetase FadD in the first step of β-oxidation. FadD has low activity on octanoate [8,9], but even if it is made, octanoyl-CoA fails to de-repress β-oxidation. Furthermore, *E. coli* grow slowly on octanoate even in the presence of mutations in *fadR* that up-regulate β-oxidation in the absence of long chain fatty acyl-CoAs [10-13]. This slow growth phenotype is likely a consequence of octanoate's slow

conversion into octanoyl-CoA due, in part, to FadD's low activity toward octanoate, but is otherwise poorly understood and may also be caused by MCFA toxicity [8,9,14-16].

Predicting that the metabolic limitations that prevent fast growth on octanoate could also prevent its conversion into downstream products and that any mutations that enhance growth rate on octanoate could also aid in the production of downstream derivatives, we evolved *E. coli* on octanoate minimal medium [17]. We hypothesized that mutations acquired through adaptive evolution on octanoate could accelerate *E.coli* growth and expedite the octanoate to octanoyl-CoA conversion process by enhancing octanoate uptake, preventing toxicity, and/or increasing the activity of β-oxidation enzymes. We successfully generated six strains that grow rapidly on octanoate and sequenced the genome of the fastest growing strain. This strain had a large insertion in the β-oxidation repressor *fadR*, a mutation in the promoter of *fadD* downstream of the transcription start site, and a point mutation causing an H430Q substitution in DhaM, a component of dihydroxyacetone kinase [18].

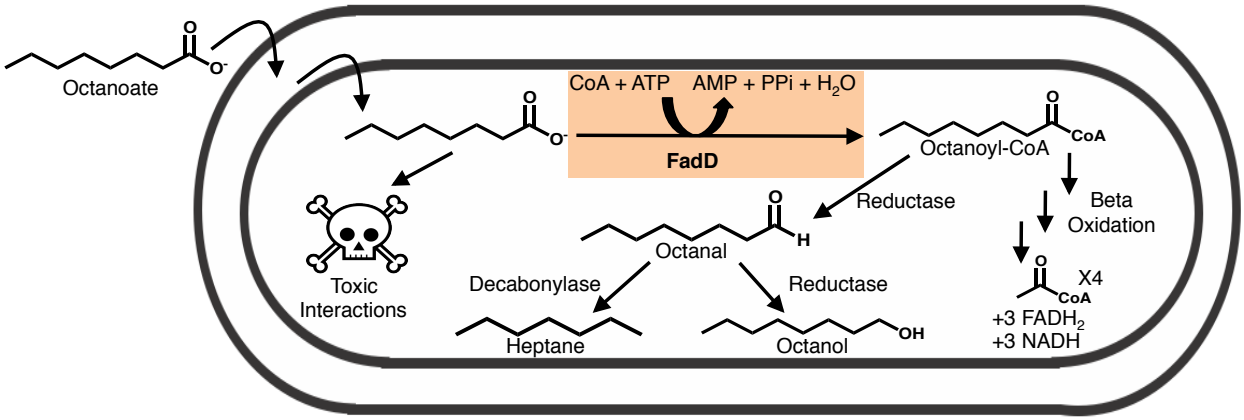

Figure 3.1. Potential pathways for octanoate catabolism and commodity production in *E.coli*. Octanoate can cross the *E.coli* outer and inner membranes by simple diffusion [12] and causes cellular toxicity through poorly understood mechanisms [14-16], but may also be catabolized through β-oxidation starting with its conversion into octanoyl-CoA by the long chain fatty acyl-CoA synthetase, FadD (highlighted in orange). After this activation step, octanoyl-CoA may be iteratively degraded to acetyl-CoA with the concomitant production of reducing equivalents by other enzymes in β-oxidation. Fatty acyl-CoAs can be converted into aldehydes by acyl-CoA reductases [1-3,5,19] and there are likely reductases that can convert octanoyl-CoA into octanal which could later be further decrabonylated or reduced into heptane or octanol respectively [4,5].

Materials and Methods

E.coli adaptive evolution on octanoate minimal medium

E.coli strain MG1655 was streaked from a glycerol stock on LB plates and grown overnight. The next day, six individual colonies were picked into LB and grown overnight. After overnight growth in LB, 3 cultures were diluted 1:50 into 5 mL of a 50:50 mixture of LB and 1.25 g/L myristate minimal medium (M9 minimal medium with 1.25 g/L myristic acid and 2.1% Triton X-100 [Sigma]), the 3 remaining cultures were diluted 1:50 into 5 mL cultures containing a 50:50 mixture of LB and 3 g/L myristate minimal medium (M9 minimal medium with 3 g/L myristic acid and 2.1% Triton X-100 [Sigma]). Once cultures reached an OD_{600} of ~2.0, they were diluted 1:10 into a 90:10 mixture of myristate:octanoate minimal medium (either 1.25 or 3 g/L myristate and 2 g/L octanoate minimal medium: M9 minimal medium with 2 g/L octanoate and 0.2% Nonidet^{TM} P 40 Substitute [Sigma]) and the process repeated iteratively 10 times increasing the proportion of octanoate minimal medium until all cultures were growing in 100% octanoate minimal medium (containing 2 g/L octanoate). Once cultures reached an OD_{600} of ~2.0, they were diluted 1:50 into 50 mL octanoate minimal medium (now with 1 g/L octanoate) and the process repeated 10 more times with cultures diluted 1:50 into octanoate miminal medium on each dilution. Once the 10th dilutions in 1 g/L octanoate minimal medium grew to an OD_{600} of ~2.0, glycerol stocks were made and stored at -80 °C for later analysis (strains E1-E6, Table 3.1).

To measure the growth rate of each strain generated by adaptive evolution, each strain E1-E6 glycerol stock was streaked on an LB plate and 6 independent colonies from each plate as well as 6 colonies from an MG1655 control plate were picked into 1 mL LB each in a separate well of a 96 well deep well plate (Nunc). Cultures were allowed to grow overnight (~18 h) and

then back diluted 1:50 into 1 mL M9 Octanoate (1 g/L octanoate). The OD_{590} of each culture was monitored throughout growth in octanoate minimal medium using a Victor 3v Multilabel Plate Reader (Perkin Elemer). Doublings and doubling/h were determined by diving all OD_{590} values by the OD_{590} recorded 1h after dilution, calculating the log in base 2 of this value, and plotting this against hours of growth. The slope of the linear portion of this curve was recorded as the doublings per hour and slopes were only accepted from linear regressions with R^2 >0.9. Strains M5 and M6 were grown in 5 mL of media in 14 mL test tubes to avoid effects from evaporation but were monitored as indicated above. Growth rates on glucose (1.7 g/L), glycerol (1.7 g/L), and sodium acetate (2.3 g/L) were measured similarly.

Strain generation

Table 3.1. Strains used in chapter 2

Name	**Genotype**	**Description**	**Reference**
TB10	MG1655 *nadA::Tn10* λcI857 Δ(*cro-bioA*)	Recombineering strain	[20]
MG1655	F- lambda- *ilvG- rfb*-50 *rph*-1	wt K-12 *E.coli* strain	Lab Stock
E1	N/A	MG1655 derivative evolved to grow on octanoate minimal medium	This Study
E2	N/A	MG1655 derivative evolved to grow on octanoate minimal medium	This Study
E3	N/A	MG1655 derivative evolved to grow on octanoate minimal medium	This Study
E4	N/A	MG1655 derivative evolved to grow on octanoate minimal medium	This Study
E5	MG1655 *dhaM* H430Q *pfadD* fadR::IS2*	MG1655 derivative evolved to grow on octanoate minimal medium	This Study
E6	N/A	MG1655 derivative evolved to grow on octanoate minimal medium	This Study
E7	E5 *dhaM* H430Q-*kan*	E5 with kanamycin resistance gene downstream of *dhaM*(H430Q)	This Study
E8	E5 *dhaM-kan*	E5 with a kanamycin resistance gene downstream of wt *dhaM*	This Study

Table 3.1 (Continued).

E9	E5 *dhaM* H430Q-*kan fadR*	E5 with a kanamycin resistance cassette inserted downstream of the mutated *dhaM* gene and the *fadR* gene reverted to wt	This Study
E10	E5 *pfadD-kan*	E5 with a kanamycin resistance cassette inserted downstream of the *fadD* gene with promoter reverted to wt	This Study
E11	E5 *pfadD*-kan*	E5 with a kanamycin resistance cassette inserted downstream of the *fadD* gene with mutated promoter	This Study
E12	E5 *dhaM-kan fadR*	E5 with a kanamycin resistance cassette inserted downstream of the *dhaM* gene and both *dhaM* and *fadR* reverted to wt	This Study
E13	E5 *pfadD*-fadD-His-kan*	E5 with a genomically integrated, C-terminally his-tagged *fadD* linked to a kanamycin resistance cassette downstream of *fadD*	This Study
E14	E5 *pfadD-fadD-His-kan*	E5 with a genomically integrated, C-terminally his-tagged *fadD* linked to a kanamycin resistance cassette downstream of *fadD* and the *fadD* promoter reverted to wt	This Study
M1	MG1655 *dhaM-kan*	MG1655 with a kanamycin resistance cassette downstream of *dhaM*	This Study
M2	MG1655 *dhaM-kan fadR::IS2*	MG1655 with a kanamycin resistance cassette downstream of *dhaM* and the *fadR::IS2* mutation from E5	This Study
M3	MG1655 *dhaM* H430Q-*kan*	MG1655 with the *dhaM* H430Q mutation from E5 linked to a downstream kanamycin resistance cassette	This Study
M4	MG1655 *pfadD-kan*	MG1655 with wt *fadD* linked to a downstream kanamycin resistance cassette	This Study
M5	MG1655 *pfadD*-kan*	MG1655 with the promoter mutation in *fadD* from E5 linked to a kanamycin resistance cassette downstream of *fadD*	This Study

Table 3.2. Primers used in Chapter 2

Name	Abbreviation	Description	Sequence
TF0260	Kan dhaM Link Primer For	Forward primer for adding homology to the region downstream of *dhaM* to the 5' end of the kanamycin resistance cassette from pFabB_Deg_kan (Chapter 1)	GCTGGACGTTAAAAC GCAACGTTTCAACCG TCAGGGTTAACTACG CAGACGCAAGCTAAG
TF0261	Kan dhaM Link Primer Rev	Reverse primer for adding homology to the region downstream of *dhaM* to the 3' end of the kanamycin resistance cassette from pFabB_Deg_kan (Chapter 1)	TTGCCGGATGACATC AGAACGATGCCATCC GAACAGTGGCTACTC CGACCTACTGCGAAG
TF0276	dhaM For	Forward primer for amplifying and sequencing *dhaM*	ATGGTAAACCTGGTC ATAG
TF0277	dhaM Rev	Reverse primer for amplifying and sequencing *dhaM*	TTAACCCTGACGGTTG
TF0284	dhaM upstream	Forward primer for amplifying and sequencing *dhaM*	GTTATCTCGGTGAAC GCAG
TF0285	dhaM downstream	Reverse primer for amplifying and sequencing *dhaM*	CTGTTTATTAGCCAGC CAGC
TF0294	FadR Check For	Forward primer for amplifying and sequencing *fadR*	TATCAGCGTAGTTAG CCCTC
TF0295	FadR Check Rev	Reverse primer for amplifying and sequencing *fadR*	TGATGTGATGCTCGA ACAG
TF0359	FadD Kan Rev	Reverse primer for adding homology to the region downstream of *fadD* 3' end of the kanamycin resistance cassette from pFabB_Deg_kan (Chapter 1)	AAACGCCGGATTAAC CGGCGTCTGACGACT GACTTAACGCTACTCC GACCTACTGCGAAG
TF0393	FadD His_kan_For	Forward primer for adding homology to the region downstream of *fadD* and a his-tag to the 5' end of the kanamycin resistance cassette from pFabB_Deg_kan (Chapter 1)	ACGTGACGAAGCGCG CGGCAAAGTGGACAA TAAAGCCGGTTCTTCT CATCACCATCATCATC ACTGACTACGCAGAC GCAAGCTAAG

All strains used in this study are listed in Table 3.1. Strains E1-E6 were generated by adaptive evolution as indicated above. Strains E7-E12 and M1-M5 were generated by P1 transduction of E5 or MG1655 (respectively) (see below) with the appropriate kanamycin resistance cassette (kan)-linked genes from TB10. These TB10 strains were generated by lambda

red mediated recombination (see below) with linear constructs containing the kan-linked genes amplified using the primers indicated in Table 3.2.

2.3 Western Blotting

Strains E13 and E14 expressing a C-terminally His_6-tagged genomically integrated FadD with and without the *pfadD** mutation were grown on M9 octanoate minimal medium as described above. 1 mL samples were taken from early exponential phase cultures and boiled in 3% SDS. Total protein concentration was normalized by A_{280} and samples were western blotted with an HRP conjugated anti His_6 antibody (ab1187, Abcam) diluted 1:10000 in TBS-tween with 1% BSA. Blots were stripped with Restore™ Western Blot Stripping Buffer (Thermo Scientific), and re-blotted with an HRP conjugated antibody to GAPDH (ab85760, Abcam) diluted 1:5000 in TBS-tween with 1% BSA. These were then blotted with an HRP-linked anti-goat secondary antibody diluted 1:1000 in TBS-tween with 1% BSA. Relative FadD band intensities were quantified in Image J [21].

TB10 lambda red mediated recombination

TB10 recombination was performed as previously described [22].

P1 transduction

P1 Transductions were performed as previously described [23].

E.coli genomic DNA isolation for whole genome sequencing

E.coli cultures for genomic DNA isolation were started from three single colonies of strain E5 picked into 5 mL LB and grown overnight at 37° C. 2 mL of each overnight culture was then transferred to a 2 mL microcentrifuge tube and spun at top speed in an Eppendorf Centrifuge 5424 (Eppendorf) for 1.5 min. The pellet was then resuspended in 467 µl TE containing 33.3 µg/mL RNase A, 30 µl of 10% sodium dodecyl sulfate, and 3 µl of 20 mg/mL proteinase K. These were then mixed by inversion and the tube incubated at 37° C for 1 h or until the solution was clear. 500 µl of phenol:chloroform:isoamyl alchohol (25:24:1) was then added and the solution vortexed until the phases completely mixed. The entire solution was then transferred to a 1.5 mL Phase Lock Gel Heavy tube (5 Prime) and the sample spun at top speed in an Eppendorf Centrifuge 5424 (Eppendorf) for 5 min. The upper, aqueous phase was then pipetted into a new 1.5 mL microcentrifuge tube, 50 µl of 5 M NaCl added, and the sample mixed by inversion. 600 µl of isopropanol was then added to the sample and mixed by inversion. The sample was then incubated for 30 min at -20° C to precipitate out the gDNA, the sample spun at top speed in an Eppendorf Centrifuge 5424 (Eppendorf) for 10 min, and the resultant gDNA pellet washed with 1 mL 75% ethanol. The pellet was then air dried and, once dry, the gDNA pellet suspended in 200 µl TE. gDNA concentration and purity were then measured via A260 and A260/A280 on a Nandrop® ND-1000 spectrophotometer (Nanadrop). gDNA with A260/A280 >1.8 was considered pure.

E.coli whole genome sequencing and analysis (performed by Wade Hicks, PhD)

Sequencing libraries were prepared from purified genomic DNA using the TruSeq kit Low Sample Protocol (Illumina). Standard Illumina adaptors were added to each library to allow for multiplexing. Pooled libraries were sequenced on the Illumina MiSeq desktop sequencing

platform using a version 2 paired-end 300-cycle reagent cartridge. Raw sequencing reads were aligned to the annotated MG1655 genome (Genebank accession NC_000913.2) and single nucleotide polymorphisms and structural variations were called using the SeqMan NGen (DNAStar) XNG assembler (Version 4.0.0.116). Summary statistics for whole genome sequencing can be found in Table 3.3.

Table 3.3. Summary Statistics for E5 whole genome sequencing

Isolate	Contig Length without gaps (base pairs)	Number of Sequences	Median Coverage
E5 1	4639675	528712	16.79
E5 2	4639675	605012	19.28
E5 3	4639675	815664	25.92

Results

Adaptive evolution yields E.coli strains with high growth rates on octanoate

Six *E.coli* strains with improved growth rate on octanoate were generated by adaptive evolution using the scheme in Figure 3.2A. In this scheme, separate inoculates of *E.coli* strain MG1655 were grown for more than 90 generations in octanoate-containing medium. After adaptive evolution, we measured each strain's growth on octanoate and sequenced the genome of the fastest growing strain, E5 (Materials and Methods) (Figure 3.2B). The three mutations identified in E5 are described in Table 3.4. Two of the three mutations were in or near genes known to affect fatty acid catabolism. These were a ~1.2 kb insertion in the β-oxidation repressor *fadR* (*fadR::IS2*) [7,24] and a G-A transition 22 bp downstream of the *fadD* transciption start site (*pfadD**). The last mutation causes an H430Q substitution in DhaM (*dhaM* H430Q), a component of dihydroxy acetone kinase [18].

Table 3.4. Mutations found in strain E5

Gene	Gene Function	Mutation Type	Designation
fadR	β-oxidation repressing transcription factor	2kb insertion	*fadR::IS2*
fadD	Fatty acyl-CoA synthetase, first step in β-oxidation	G-A transition 21bp downstream of transcription start site	*pfadD**
dhaM	Component of dihydroxyacetone kinase	C-A transversion causing H430-Q430	*dhaM*_H430Q

Each of the mutations generated by adaptive evolution enhances growth on octanoate

The E5 *fadR::IS2* and *pfadD** mutations increase growth on octanoate while the *dham* H430Q mutation causes E5 to maintain a higher OD_{590} in stationary phase. To determine how these mutations contribute to E5's growth on octanoate, they were each individually reverted to wt in E5 and mutated in MG1655 (Figure 3.2C). Reverting *fadR::IS2* to wt (Figure 3.2Ci) increases the lag time before exponential growth but does not alter exponential growth rate.

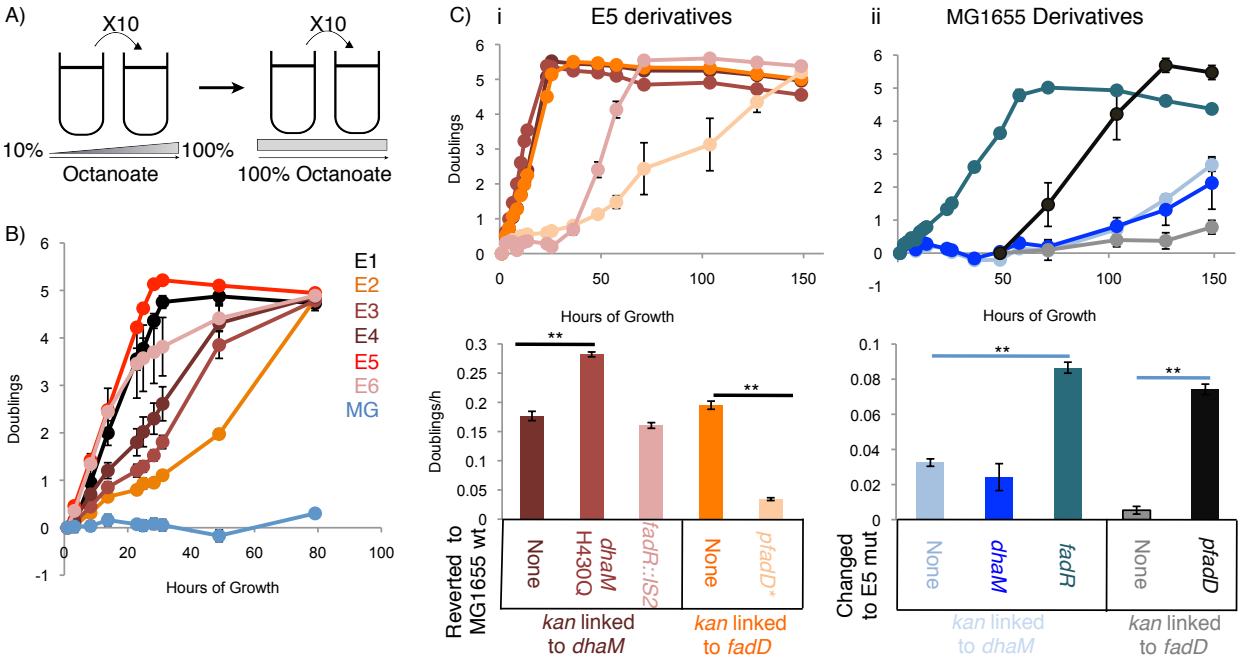

Figure 3.2. *E.coli* adaptive evolution on octanoate points toward FadD as an important bottleneck in octanoate metabolism. A) Schematic representation of *E.coli* adaptive evolution on octanoate (Materials and Methods). B) Growth of each of 6 independently evolved *E.coli* cultures in octanoate minimal medium after all 20 rounds of evolution indicated in (A).

Figure 3.2 (Continued). MG=MG1655. C) The effects of each individual mutation found in strain E5 (Table 3.4) on growth in octanoate minimal medium were determined by (i) replacing the mutations with their *kan*-linked wt counter parts with the *kan* cassette in the indicated location, (ii) by introducing each *kan*-linked mutation individually into the MG1655 parent strain with the *kan* cassette in the indicated location, and measuring the growth of each new strain on octanoate. n=3-6, error bars indicate standard error, and ** indicates $p<0.05$ by two-sided students T-Test between the indicated strains. All growth rates are the slopes of lines generated by linear regression during exponential growth. All lines had $R^2 >0.9$ except for those generated for MG1655 *dhaM* (H430Q)-*kan* and MG1655 *pfadD-kan* which had highly variable growth.

In contrast, reverting *pfadD** to wt (Figure 3.2Ci) decreases the growth rate on octanoate, but does not decrease the lag time before exponential growth begins. Reverting *dhaM* H430Q to wt counter-intuitively increases growth rate, but also moderately decreases the stationary phase OD_{590} (~5-10% decrease, Figure 3.2Ci). When added to MG1655, the *fadR::IS2* and *pfadD** mutations increased growth rate on octanoate. The *fadR::IS2* mutation additionally removed the lag between dilution and the start of exponential growth, while the *pfadD** mutation decreased this lag by ~20h. In contrast, the *dhaM* (H430Q) mutation had little effect on its own with no statistically significant differences in growth rate over the period measured (Figure 3.2Cii).

The *pfadD** mutation enhances *E. coli* growth rate on octanoate minimal medium by increasing FadD expression. FadD was C-terminally His_6-tagged, linked to a kanamycin resistance cassette and integrated into the genome of strain E5 generating strains E10 and E11 with and without the *pfadD** mutation respectively. As with the strains in Figure 3.2, the strain with the *pfadD** mutation (E10) had a higher growth rate compared to the strain without the mutation (E11) (Figure 3.3A). Anti-His western blots performed using samples made from early exponential phase cultures of these two strains in octanoate minimal medium showed that the *pfadD** mutation increases expression of His_6-tagged FadD. These results confirm FadD as a key limitation on octanoate metabolism.

The *dhaM* (H430Q) mutation enhances *E. coli* growth rate on the carbon sources glucose, glycerol, and acetate. Because dihydroxyacetone kinase, encoded by *dhaM*, has no known link to fatty acid metabolism, we tested whether or not this mutation generally enhances growth on a variety of carbon sources by growing strains with and without the mutation on glucose, glycerol, and acetate (Figure 3.4). When the *dhaM* (H430Q) mutation was reverted to wild type in evolved strain E5, growth rate decreased on all carbon sources tested, and when it was introduced into MG1655, growth rate increased on all carbon sources. These results show that *dhaM* mutation causes a small, general growth effect and is not likely directly related to octanoate metabolism.

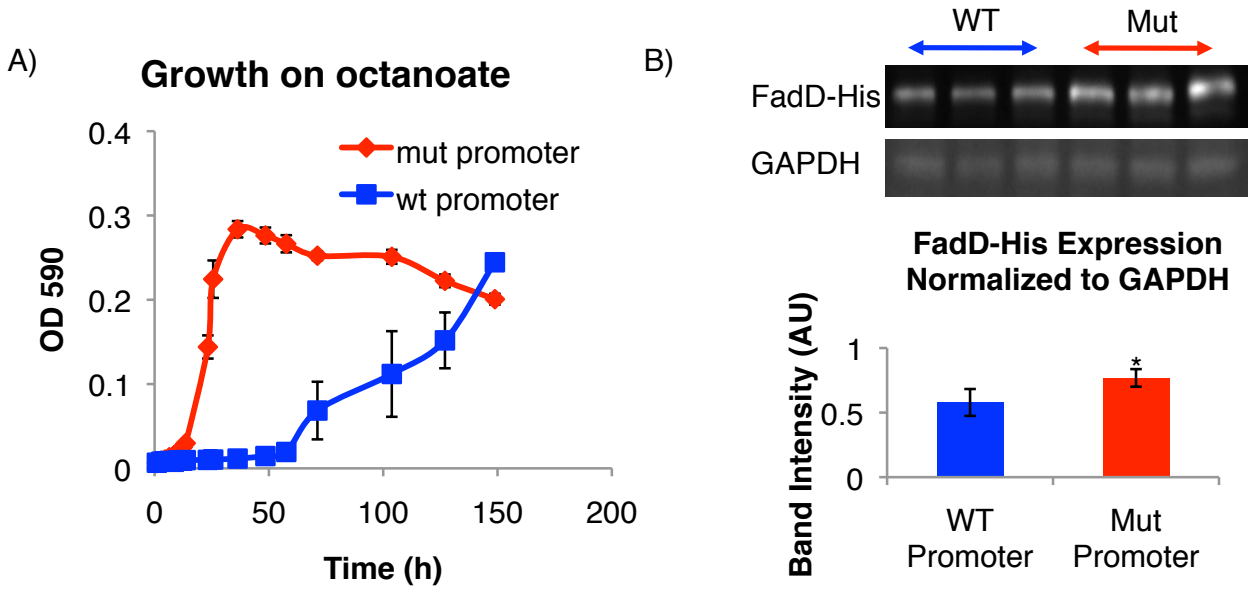

Figure 3.3. A mutation in the *fadD* promoter generated by adaptive evolution enhances evolved strain (E5) growth on octanoate minimal medium and increases FadD expression. A) Growth of evolved strain E5 containing a genomically integrated His_6-tagged FadD with and without the *pfadD** mutation (n=6, error bars indicate standard deviation). B) Western blot with anti-His and anti-GAPDH antibodies using SDS-PAGE samples from evolved strain E5 with (mut) and without (wt) the *pfadD** mutation taken during early exponential growth on octanoate. C) Quantification of FadD bands in B normalized to GAPDH. n=3, error bars indicate standard deviation, and * indicates mutant expression is statistically significantly different from wild-type with $p<0.1$ by two sided students T-Test.

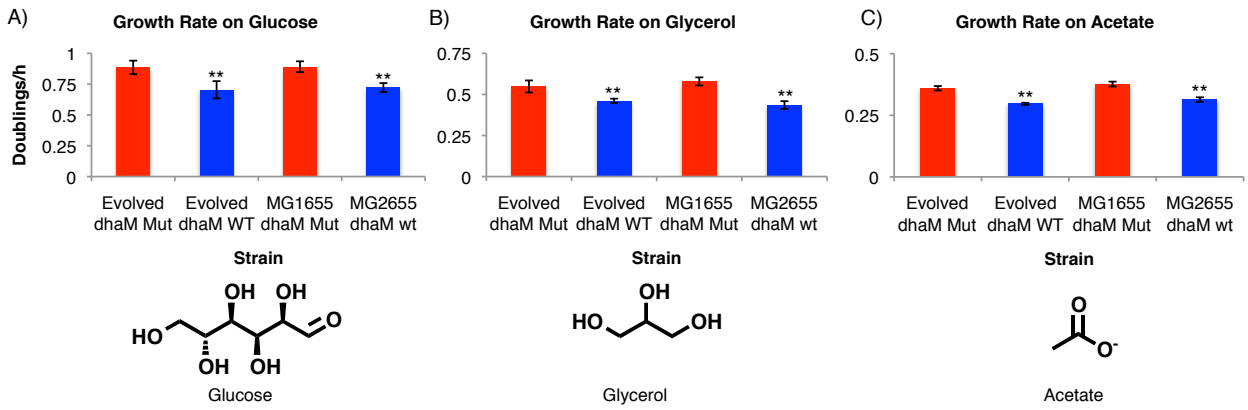

Figure 3.4. The *dhaM** mutation enhances growth rate on a variety of carbon sources. Strains E8 (evolved *dhaM* mut), E9 (evolved *dhaM* wt*),* M3 (MG1655 *dhaM* mut), and M1 (MG1655 *dhaM* wt) were grown overnight at room-temperature in minimal medium containing (A) glucose, (B) glycerol, and (C) acetate as sole carbon source. Growth rates were determined as indicated in materials and methods. n=4, error bars indicate standard deviation, and ** indicates $p<0.05$ compared to background matched strains containing the *dhaM* mutation by two sided students T-Test.

Discussion

The goal of this work was to discover mutations that enhance *E. coli* octanoate catabolism and therefore efficiently generate medium chain acyl-CoAs that can be converted into reduced products like alcohols and alkanes. We used this unbiased method hypothesizing that any mutations discovered would 1) enhance octanoate import, 2) limit octanoate toxicity, or 3) enhance the activity of β-oxidation enzymes on MCFAs. Of these different possibilities, two of our mutations (*fadR::IS2,* and *pfadD**) fall into the third category, whereas the final mutation (*dham* (H430Q)) falls into an unanticipated fourth category of mutations that non-specifically enhance growth.

The effect of the *fadR::IS2* mutation is consistent with the known effects of loss of function mutations in *fadR*. FadR is a transcription factor that represses the expression of β-oxidation genes including *fadD* [7]. Long chain fatty acyl-CoAs produced by basally expressed FadD are known to bind to FadR and inhibit its binding to cognate promoters, resulting in de-

repression; however, FadR does not bind to octanoyl-CoAs with high affinity [25]. *fadR* mutations increase the expression of β-oxidation genes in the absence of long chain fatty acyl-CoAs thereby giving *fadR* mutants the ability to grow on MCFAs like octanoate [10,11,26]. Given that our *fadR::IS2* mutants have a roughly 2 kb insertion in the *fadR* gene and that the effects of this mutation mimic those of *fadR* loss of function mutations, the *fadR::IS2* mutation is likely a loss of function that enhances the expression of *fadD* and other genes in β-oxidation.

The *pfadD** mutation enhances *fadD* expression but does not lie in any known transcription factor binding sites. This mutation consists of a single G-A transition 22 bp downstream of the *fadD* transcription start site. Binding sites for the transcription factors FadR, arcA, and CRP lie upstream of the *fadD* transcription start site, but there are no annotated binding sites downstream [27-29]. Additionally, this mutation is 30 bp upstream of the *fadD* ribosome binding site [27]. It is possible that this mutation indirectly impacts transcription factor binding or lies in an un-annotated transcription factor binding site. It may also indirectly alter ribosome binding through RNA secondary structure [30]. Regardless, the enhanced expression of *fadD* due to the *pfadD** mutation increases acyl-CoA synthetase activity enough to achieve a substantial (>5 fold) increase in growth rate on octanoate. This result agrees well with work in chapter three showing that plasmid-based *fadD* expression is sufficient to enhance growth on octanoate in a *ΔfadR* background.

The *dhaM* (H430Q) mutation has no obvious link to *fadD* or fatty acid metabolism. DhaM is a phosphotransferase system-like protein and component of dihydroxy acetone kinase [18]. H430 is part of a phosphate relay in DhaM that ultimately transfers a phosphate group from phosphoenol pyruvate to dihydroxy acetone and additionally affects transcription through transcription factor *dhaR* [18,31]. However, how mutating this residue causes the observed

effects on stationary phase growth is unknown. This mutation was likely beneficial to strain E5 only because the adaptive evolution was performed in batch culture and not because of any particular link to octanoate metabolism. Our cultures were allowed to reach stationary phase before back dilution and therefore any mutations that increased *E.coli*'s ability to thrive in stationary phase, like the *dhaM* H430Q mutation, were selected for under these conditions.

Although the mechanism behind the effects of the *dhaM (H430Q)* mutation were not determined in this work, this mutation may prove useful in its ability to increase growth rate on a variety of carbon sources. The faster growth rate afforded by this mutation may enhance production rates in *E. coli* engineered to produce compounds such as fatty acids. We tested whether or not this mutation could increase fatty acid yields but found no benefit of this mutation in an otherwise wild-type MG1655 background (data not shown). Future work should test the effects of this mutation under a variety of different production settings.

The *fadR::IS2* and *pfadD** mutations generated in this work were expected *apriori* given *fadR*'s well studied regulatory role in fatty acid catabolism [7] and FadD's low activity on MCFAs [8,9]. The identification of these mutations motivates the work in chapter three where we directly focus on mutating FadD to enhance its activity on MCFAs. Future studies could additionally sequence the genomes of the five other evolved strains generated in this work and may identify novel mutations benefiting fatty acid catabolism.

References

1. Doan TT, Carlsson AS, Hamberg M, Bulow L, Stymne S, Olsson P: **Functional expression of five Arabidopsis fatty acyl-CoA reductase genes in Escherichia coli**. *Journal of plant physiology* 2009, **166**:787-796.

2. Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD: **Microbial production of fatty-acid-derived fuels and chemicals from plant biomass**. *Nature* 2010, **463**:559-562.

3. Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB: **Microbial biosynthesis of alkanes**. *Science* 2010, **329**:559-562.

4. Kallio P, Pasztor A, Thiel K, Akhtar MK, Jones PR: **An engineered pathway for the biosynthesis of renewable propane**. *Nature communications* 2014, **5**:4731.

5. Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R: **Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals**. *Nature* 2011, **476**:355-359.

6. DiRusso C, Heimert TL, Metzger AK: **Characterization of FadR, a global transcriptional regulator of fatty acid metabolism in Escherichia coli. Interaction with the fadB promoter is prevented by long chain fatty acyl coenzyme A**. *Journal of Biological Chemistry* 1992, **267**:8685-8691.

7. Fujita Y, Matsuoka H, Hirooka K: **Regulation of fatty acid metabolism in bacteria**. *Molecular microbiology* 2007, **66**:829-839.

8. Kameda K, Nunn WD: **Purification and characterization of acyl coenzyme A synthetase from Escherichia coli**. *The Journal of biological chemistry* 1981, **256**:5702-5707.

9. Morgan-Kiss RM, Cronan JE: **The Escherichia coli fadK (ydiD) gene encodes an anerobically regulated short chain acyl-CoA synthetase**. *The Journal of biological chemistry* 2004, **279**:37324-37333.

10. Salanitro JP, Wegener WS: **Growth of Escherichia coli on short-chain fatty acids: growth characteristics of mutants**. *Journal of bacteriology* 1971, **108**:885-892.

11. Campbell JW, Morgan-Kiss RM, Cronan JE, Jr.: **A new Escherichia coli metabolic competency: growth on fatty acids by a novel anaerobic beta-oxidation pathway**. *Molecular microbiology* 2003, **47**:793-805.

12. Maloy SR, Ginsburgh CL, Simons RW, Nunn WD: **Transport of long and medium chain fatty acids by Escherichia coli K12**. *The Journal of biological chemistry* 1981, **256**:3735-3742.

13. Nunn WD, Simons RW, Egan PA, Maloy SR: **Kinetics of the utilization of medium and long chain fatty acids by mutant of Escherichia coli defective in the fadL gene**. *The Journal of biological chemistry* 1979, **254**:9130-9134.

14. Lennen RM, Kruziki MA, Kumar K, Zinkel RA, Burnum KE, Lipton MS, Hoover SW, Ranatunga DR, Wittkopp TM, Marner WD, 2nd, et al.: **Membrane stresses induced by overproduction of free fatty acids in Escherichia coli**. *Applied and environmental microbiology* 2011, **77**:8114-8128.

15. Lennen RM, Pfleger BF: **Modulating membrane composition alters free fatty acid tolerance in Escherichia coli**. *PloS one* 2013, **8**:e54031.

16. Sherkhanov S, Korman TP, Bowie JU: **Improving the tolerance of Escherichia coli to medium-chain fatty acid production**. *Metabolic engineering* 2014, **25**:1-7.

17. Dragosits M, Mattanovich D: **Adaptive laboratory evolution -- principles and applications for biotechnology**. *Microbial cell factories* 2013, **12**:64.

18. Gutknecht R, Beutler R, Garcia-Alles LF, Baumann U, Erni B: **The dihydroxyacetone kinase of Escherichia coli utilizes a phosphoprotein instead of ATP as phosphoryl donor**. *The EMBO journal* 2001, **20**:2480-2486.

19. Sachdeva G, Garg A, Godding D, Way JC, Silver PA: **In vivo co-localization of enzymes on RNA scaffolds increases metabolic production in a geometrically dependent manner**. *Nucleic acids research* 2014, **42**:9493-9503.

20. Johnson JE, Lackner LL, Hale CA, de Boer PA: **ZipA is required for targeting of DMinC/DicB, but not DMinC/MinD, complexes to septal ring assemblies in Escherichia coli**. *Journal of bacteriology* 2004, **186**:2418-2429.

21. Schneider CA, Rasband WS, Eliceiri KW: **NIH Image to ImageJ: 25 years of image analysis**. *Nature methods* 2012, **9**:671-675.

22. Datsenko KA, Wanner BL: **One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products**. *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**:6640-6645.

23. Thomason LC, Costantino N, Court DL: **E. coli genome manipulation by P1 transduction**. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* 2007, **Chapter 1**:Unit 1 17.

24. Nunn WD: **A molecular view of fatty acid catabolism in Escherichia coli**. *Microbiological reviews* 1986, **50**:179-192.

25. DiRusso CC, Tsvetnitsky V, Hojrup P, Knudsen J: **Fatty acyl-CoA binding domain of the transcription factor FadR. Characterization by deletion, affinity labeling, and isothermal titration calorimetry**. *The Journal of biological chemistry* 1998, **273**:33652-33659.

26. Overath P, Pauli G, Schairer HU: **Fatty acid degradation in Escherichia coli. An inducible acyl-CoA synthetase, the mapping of old-mutations, and the isolation of regulatory mutants**. *European journal of biochemistry / FEBS* 1969, **7**:559-574.

27. Black PN, DiRusso CC, Metzger AK, Heimert TL: **Cloning, sequencing, and expression of the fadD gene of Escherichia coli encoding acyl coenzyme A synthetase**. *The Journal of biological chemistry* 1992, **267**:25513-25520.

28. Cho BK, Knight EM, Palsson BO: **Transcriptional regulation of the fad regulon genes of Escherichia coli by ArcA**. *Microbiology* 2006, **152**:2207-2219.

29. Feng Y, Cronan JE: **Crosstalk of Escherichia coli FadR with global regulators in expression of fatty acid transport genes**. *PloS one* 2012, **7**:e46275.

30. de Smit MH, van Duin J: **Secondary structure of the ribosome binding site determines translational efficiency: a quantitative analysis**. *Proceedings of the National Academy of Sciences of the United States of America* 1990, **87**:7668-7672.

31. Bachler C, Schneider P, Bahler P, Lustig A, Erni B: **Escherichia coli dihydroxyacetone kinase controls gene expression by binding to transcription factor DhaR**. *The EMBO journal* 2005, **24**:283-293.

Chapter 4: Enhancing *E. coli* acyl-CoA synthetase FadD activity on medium chain fatty acids

Abstract

FadD catalyses the first step in *E. coli* β-oxidation, the activation of free fatty acids into acyl-CoA thioesters. This activation makes fatty acids competent for catabolism and reduction into derivatives like alcohols and alkanes. Alcohols and alkanes derived from medium chain fatty acids (MCFAs, 6-12 carbons) are potential biofuels; however, FadD has low activity on MCFAs. Herein, we generate mutations in *fadD* that enhance its acyl-CoA synthetase activity on MCFAs. Homology modeling reveals that these mutations cluster on a face of FadD from which the co-product, AMP, is expected to exit. Using FadD homology models, we design additional FadD mutations that enhance *E. coli* growth rate on octanoate and provide evidence for a model wherein FadD activity on octanoate can be enhanced by aiding product exit. These studies provide FadD mutants useful for producing MCFA derivatives and a rationale to alter the substrate specificity of adenylating enzymes.

Introduction

Medium chain fatty acids (MCFAs, 6-12 carbons) are important precursors to fuel-like compounds and industrial chemicals [1,2]. *E. coli* have been engineered to produce MCFAs using a variety of techniques, but only a few labs have successfully demonstrated their conversion into fuel-like compounds such as alcohols and alkanes [3-8]. Conversion of MCFAs into aldehydes, alcohols, and alkanes, is facilitated by the activation of the MCFA carboxylic acid head group into a stronger electrophile. Biologically, this can be achieved by converting the carboxyl group into an acyl-CoA thioester. The acyl-CoA synthetase FadD catalyses this conversion in *E. coli* aerobic β-oxidation and has been used to activate long chain fatty acids (LCFAs, 13+ carbons) for their later reduction into fuel-like compounds [9-12]. However, FadD

has low activity on fatty acids 10 carbons and shorter resulting in slow *E. coli* growth rates on these fatty acids even in the presence of mutations de-repressing *fadD* and other genes involved in β-oxidation [13-17]. *Salmonella enterica*, which has a FadD very similar to that of *E. coli*, grows more quickly than *E. coli* on octanoate, but this is due to changes in *fadD* regulation and the activity of downstream β-oxidation enzymes and not to changes in FadD enzymatic activity [17].

The mechanisms behind FadD substrate specificity are not well understood. This protein belongs to a class of adenylate-forming enzymes for which numerous crystal structures have been solved, including an LCFA-specific, FadD homolog from *Thermus thermophilus* [18-23] and an MCFA-specific homolog from *Homo sapiens* with butyryl-CoA and AMP in the active site [24]. Analyzing the structure of *Thermus thermophilus* acyl-CoA sysnthease co-crystalized with myristoyl-AMP, Hisanaga and coworkers hypothesized that, because the myristoyl-AMP rests in a tunnel well-suited to accommodate its long, hydrophobic tail, it is the length of this tunnel that determines substrate specificity [18]. A similar mechanism has been used to explain the medium chain specificity of the human homolog [24], but it has not been shown whether decreasing or increasing the size of this tunnel experimentally can alter substrate specificity. Black et al., 1997 constructed mutations in a conserved fatty acyl-CoA synthetase (FACS) motif in FadD. These had subtle effects on FadD selectivity, but only one showed an absolute increase in activity on decanoate [25]. The FACS motif is adjacent to a region of FadD involved in fatty acid binding, but no further mutagenesis studies of this region have led to increased FadD activity on MCFAs shorter than 10 carbons [26].

Herein, we specifically enhanced FadD activity on MCFAs shorter that 10 carbons using a strategy incorporating *fadD* mutagenesis by error prone PCR and a growth-based screen for

acyl-CoA synthetase activity. We hypothesized that FadD mutants that enhance *E. coli* growth rate on octanoate would have increased activity on MCFAs because FadD catalyzes the first step in fatty acid catabolism. We generated FadD mutants that confer increased growth rate on the MCFAs hexanoate (6-carbons), octanoate (8-carbons), and moderately on decanoate (10-carbons), but not palmitate (16-carbons) or oleate (18-carbons). *In vitro* assays of partially purified wild-type FadD and mutant variants showed that they possess increased activity on octanoate and decanoate, but not oleate. Homology modeling revealed that the isolated mutations cluster around a proposed AMP exit channel from the FadD active site [18,24], and mutations designed to widen this exit channel confer increased growth rate on octanoate. These FadD mutants can aid in the production of biofuels and industrial chemicals and provide a rationale to engineer other adenylate-forming enzymes important for processes ranging from lignin processing [20] to antibiotic production [22].

Materials and Methods

Plasmid Construction

All plasmid constructs and primers are listed in Tables 4.1 and 4.2 respectively and were generated using standard cloning techniques. Briefly, primers TF0093 and TF0086 amplified wild-type *fadD* from MG1655 gDNA. Wild-type *fadD* was then digested with NcoI and HindIII before ligation into pETDuet-1 (Novagen®) forming TJF032. *fadD* mutants generated by error prone PCR (see below) were amplified similarly and cloned into pETDuet-1(Novagen®). His-tagged constructs were generated via PCR amplification from these constructs using primers TF0093 and TF0278.

Table 4.1. Plasmids Used in this Chapter

Name	Description	Resistance	Reference
pETDuet-1	Cloning Vector with IPTG inducible T7 polymerase promoter	Amp	Novagen ®
pFabB_deg_kan	Vector containing a degradation-tagged *fabB* linked to a frt-flanked kanamycin resistance cassette	Amp	[5]
TJF032	pETDuet-1 FadD	Amp	This Study
TJF040	pETDuet-1 FadD Q338R	Amp	This Study
TJF065	pETDuet-1 FadD H376R	Amp	This Study
TJF066	pETDuet-1 FadD V4W5, F4L5	Amp	This Study
TJF067	pETDuet-1 FadD F447S	Amp	This Study
TJF068	pETDuet-1 FadD V451A	Amp	This Study
TJF069	pETDuet-1 FadD D372G	Amp	This Study
TJF070	pETDuet-1 FadD Y9H	Amp	This Study
TJF059	pETDuet-1 FadD His	Amp	This Study
TJF072	pETDuet-1 FadD Q338 His	Amp	This Study
TJF073	pETDuet-1 FadD H376R His	Amp	This Study
TJF074	pETDuet-1 FadD V4W5, F4L5 His	Amp	This Study
TJF075	pETDuet-1 FadD F447S His	Amp	This Study
TJF076	pETDuet-1 FadD V451A His	Amp	This Study
TJF077	pETDuet-1 FadD D372G His	Amp	This Study
TJF078	pETDuet-1 FadD Y9H His	Amp	This Study
TJF080	pETDuet-1 FadD S379R	Amp	This Study
TJF081	pETDuet-1 FadD F447D	Amp	This Study
TJF082	pETDuet-1 FadD F447R	Amp	This Study
TJF083	pETDuet-1 FadD F447R S379R	Amp	This Study
TJF084	pETDuet-1 FadD F447D S379R	Amp	This Study
TJF085	pETDuet-1 FadD R554G	Amp	This Study
TJF086	pETDuet-1 FadD D551G	Amp	This Study
TJF087	pETDuet-1 FadD Q338G	Amp	This Study
TJF088	pETDuet-1 FadD R449G	Amp	This Study
TJF089	pETDuet-1 FadD V451G	Amp	This Study
TJF090	pETDuet-1 FadD H376G	Amp	This Study
TJF091	pETDuet-1 FadD S379G	Amp	This Study
TJF092	pETDuet-1 FadD Q339G	Amp	This Study
TJF093	pETDuet-1 FadD L459G	Amp	This Study
TJF094	pETDuet-1 FadD Y371G	Amp	This Study

Table 4.2. Primers Used in this Chapter

Name	Short Descriptor	Full Description	Sequence
TF0017	pETDuet -1 Seq-3	Forward primer for sequencing genes cloned into pETDuet-1	CTCGATCCCGCGAA ATTAATACG

Table 4.2 (Continued).

TF0018	pETDuet-1 Seq4	Reverse primer for sequencing genes cloned into pETDuet-1	CTTAAGCATTATGC GGCCGCAAG
TF0086	FadD Rev	Reverse primer used for cloning all non-his-tagged *fadD* variants into pETDuet-1 using NcoI and HindIII restriction sties	CTGCAAGCTTGGAT CCTCATTCATTCAG TGATGATGATGGTG ATGAGAAGAACCT CAGGCTTTATTGTC CAC
TF0093	FadD For 3	Forward primer used for cloning all *fadD* variants into pETDuet-1 using NcoI and HindIII restriction sites	CAGGACCATGGCA TTGAAGAAGGTTTG GCTTAAC85
TF0278	FadD Rev His	Reverse primer used with TF0093 to add a C-terminal His-tag to all appropriate *fadD* variants and clone into pETDuet-1 using NcoI and HindIII restriction sites	CTGCAAGCTTGGAT CCTCATTCATTCAG TGATGATGATGGTG ATGAGAAGAACCG GCTTTATTGTCCAC
TF0305	FadD S379-R379 For	Primer for generating S379R mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	ATATTGATTATCAT AGTGGACGTATCG GTTTGCCGGTGCCG
TF0306	FadD S379-R379 Rev	Primer for generating S379R mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CGGCACCGGCAAA CCGATACGTCCACT ATGATAATCAATAT
TF0307	FadD F447-D447 For	Primer for generating F447D mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GCGGTGATGGATG AAGAAGGAGATCT GCGCATTGTCGATC GTAAA
TF0308	FadD F447-D447 Rev	Primer for generating F447D mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	TTTACGATCGACAA TGCGCAGATCTCCT TCTTCATCCATCAC CGC
TF0313	FadD F447-R447 For	Primer for generating F447R mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CGCGGTGATGGAT GAAGAAGGACGCC TGCGCATTGTC
TF0314	FadD F447-R447 Rev	Primer for generating F447R mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GACAATGCGCAGG CGTCCTTCTTCATC CATCACCGCG
TF0319	Q339-G339 For	Primer for generating Q339G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	ACGCTCTGCCACCA CGCCCTGCACTGGC ATCCC

Table 4.2 (Continued).

TF0320	Q339-G339 Rev	Primer for generating Q339G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GGGATGCCAGTGC AGGGCGTGGTGGC AGAGCGT
TF0321	S379-G379 For	Primer for generating S379G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CCGGCAAACCGAT GCCTCCACTATGAT AATCAATATC
TF0322	S379-G379 Rev	Primer for generating S379G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GATATTGATTATCA TAGTGGAGGCATC GGTTTGCCGG
TF0323	H376-G376 For	Primer for generating H376G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	AACCGATGCTTCCA CTGCCATAATCAAT ATCATATGGGTTAA CGCTGACC
TF0324	H376-G376 Rev	Primer for generating H376G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GGTCAGCGTTAACC CATATGATATTGAT TATGGCAGTGGAA GCATCGGTT
TF0325	V451-G451 For	Primer for generating V451G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GAAGGATTCCTGC GCATTGGCGATCGT AAAAAAGACATG
TF0326	V451-G451 Rev	Primer for generating V451G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CATGTCTTTTTTAC GATCGCCAATGCG CAGGAATCCTTC
TF0327	R449-G449 For	Primer for generating R449G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	TTTTACGATCGACA ATGCCCAGGAATC CTTCTTCATC
TF0328	R449-G449 Rev	Primer for generating R449G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GATGAAGAAGGAT TCCTGGGCATTGTC GATCGTAAAA
TF0331	Q338-G338 For	Primer for generating Q338G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CTCTGCCACCACTT GGCCCACTGGCATC CCTCC
TF0332	Q338-G338 Rev	Primer for generating Q338G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GGAGGGATGCCAG TGGGCCAAGTGGT GGCAGAG

Table 4.2 (Continued).

TF0333	D551-G551 For	Primer for generating D551G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GCCGCGCGCTTCGC CACGTAATTCTCGT C
TF0334	D551-G551 Rev	Primer for generating D551G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GACGAGAATTACG TGGCGAAGCGCGC GGC
TF0335	R554-G554 For	Primer for generating R554G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CCACTTTGCCGCCC GCTTCGTCACG
TF0336	R554-G554 Rev	Primer for generating R554G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CGTGACGAAGCGG GCGGCAAAGTGG
TF0337	L459-G459 For	Primer for generating L459G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GGATAGACGTTAA AACCGGAAACGCC AATCATGTCTTTTT TACGATCGAC
TF0338	L459-G459 Rev	Primer for generating L459G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GTCGATCGTAAAA AAGACATGATTGG CGTTTCCGGTTTTA ACGTCTATCC
TF0339	Y371-G371 For	Primer for generating Y371G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	CTTCCACTATGATA ATCAATATCGCCTG GGTTAACGCTGACC AGCGGC
TF0340	Y371-G371 Rev	Primer for generating Y371G mutation in *fadD* using Quick Change II site directed mutagenesis kit (Agilent)	GCCGCTGGTCAGC GTTAACCCAGGCG ATATTGATTATCAT AGTGGAAG
TF0361	Amp For	Primer for sequencing *bla* in pETDuet-1	AATTTCTGGCGGCA CGATG
TF0362	Amp Rev	Primer for sequencing *bla* in pETDuet-1	GAACGAAAACTCA CGTTAAG
TF0363	pBR322 Ori For	primer for sequencing the pBR322 Ori in pETDuet-1	GCGATAAGTCGTGT CTTAC
TF0364	LacI For	primer for sequencing *lacI* in pETDuet-1	ACCGGAAGGAGCT GACTGG
TF0365	LacI Rev	primer for sequencing *lacI* in pETDuet-1	CTCCTTGCATGCAC CATTC

Table 4.3: Strains Used in this Chapter

Name	Genotype	Description	Reference
JW1176-1	*Δ(araD-araB)567 ΔlacZ4787*(::rrnB-3) λ^- *ΔfadR776::kan rph-1 Δ(rhaD-rhaB)568 hsdR514*	Keio collection knock out strain with a kanamycin resistance cassette replacing *fadR*	[27]
BL21*(DE3) *ΔfadD*	F- *ompT hsdS* ($r_B^-m_B^-$) *gal dcm rne131* (DE3) *fadD::kan*	BL21*(DE3) with a knock out in the *fadD* gene generated by transducing the kanamycin resistsance cassette from the Keio collection *fadD::kan* strain into BL21*(DE3) and flipping out the resistance cassette via expression of Flp recombinase	[5,27]

Error Prone PCR and fadD mutant screening

Error prone PCR mixtures contained 90 µl Go-Taq Green 2X Master Mix (Promega) mixed with ~150 ng of TJF032 template, 0.5 µM each of primers TF0093 and TF0086, 40 µM $MnCl_2$ and H_2O to 180 µl. The resulting PCR products were digested with NcoI and HindIII and ligated into pETDuet-1 (Novagen®). Ligation products were transformed into BW25113 *ΔfadR* with a separate TJF032 control and plated on octanoate minimal medium (M9 + 1g/L octanoate, 0.2% Nonidet^{TM} P 40 Substitute [Sigma] 15% agar), containing 50 µg/mL ampicillin (Amp). Transformants grew for 3 days at 37° C. Colonies larger than those on the TJF032 plate were restreaked on octanoate minimal plates along with TJF032 transformant controls, allowed to grow for 3 further days, and restreaked a second time. Transformant colonies larger than TJF032 transformant colonies after this third streak were picked into 5 mL LB/Amp, grown overnight, and miniprepped. Miniprepped constructs were sequenced using primers TF0017 and TF0018, re-transformed into JW1176-1, and transformants plated on LB/Amp. 6 colonies from each transformation as well as 6 colonies from a TJF032 control transformation were then picked into 1 mL LB/Amp each in a 96 well deep well plate (Nunc). Cultures were grown overnight (~18 h)

and then diluted 1:50 into 1 mL M9 octanoate containing 50 μg/mL Amp. The OD_{590} of each culture was monitored throughout growth in octanoate minimal medium using a Victor 3v Multilabel Plate Reader (Perkin Elemer). Doublings and doublings/h were determined by diving all OD_{590} values by the OD_{590} recorded 1h after dilution, calculating the log in base 2 of this value, and plotting this against hours of growth. The slope of the linear portion of this curve (R^2>0.9) was recorded as the doublings per hour. Growth rates on hexanoate (0.90 g/L), decanoate (0.80 g/L), palmitate (0.74 g/L), and oleate (0.73 g/L) were determined similarly. Palmitate and oleate minimal media had 0.4% NP40, 0.2% ethanol, and 1% Triton X-100 to solubilize the fatty acids.

Western Blotting

Strain JW1176-1 expressing the appropriate C-terminally His_6-tagged FadD variant was grown on M9 octanoate minimal medium as described above. 1 mL samples were taken from early exponential phase cultures and boiled in 3% SDS. Total protein concentration was normalized by A_{280} and samples were western blotted with an HRP conjugated antibody to the His_6 tag (ab1187, abcam) diluted 1:10000 in TBS-tween with 1% BSA. Relative band intensities of the FadD variants in their linear range (as determined by serial dilutions) were quantified in Image J [28].

Site Directed Mutagenesis

Site directed mutants were constructed using the QuikChange II Site-Directed Mutagenesis Kit (Agilent) using the primers indicated in Table 4.2 and plasmid TJF032 as template per the manufacturers' instructions. Successfully generated constructs were sequenced

and transformed into JW1176-1 and growth rates in octanoate minimal medium measured as indicated for the error prone PCR mutants above.

Partial Purification of C-terminally His_6-tagged FadD

For all purifications, the appropriate C-terminally His_6-tagged FadD variants were purified from BL21*(DE3) *ΔfadD*. Fresh transformation mixtures were diluted directly into 5 mL LB containing 50 μg/mL Amp. Cultures were grown overnight at 37° C with shaking at 250 rpm and back diluted 1:500 into 250 mL LB/Amp in 1 L Erlenmeyer flasks at 21° C with shaking at 250 RPM. After 13 h of growth at 21° C (OD_{600} ~0.2), cultures were induced with 0.1 mM IPTG and incubated for 9 further hours. Cells were then harvested by centrifuging at 4000 x g for 10 min at 4° C in a J6-M1 centrifuge (Beckman) (all buffers and incubations for the remainder of the procedure were at 4° C). The cell supernatant was then poured off and the pellet resuspended in 10 mL lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8) with 1 mg/mL Lysozyme (Sigma), 0.125 mg/mL DNase I (Sigma), 1 μg/mL pepstatin, and 1 protease inhibitor cocktail tablet for general use (Sigma). The resuspended pellet was then sonicated in a 550 Sonic Dismembrator (Fisher Scientific). Samples were centrifuged at 14,000 x g for 30 min in an Avanti J-301 centrifuge (Beckman). After centrifugation, the cell lysate (supernatant) was transferred to a new tube, the pellet discarded, and 5 μl of the lysate added to 5 μl 2X Tris-Glycine SDS sample buffer (Life Technologies) and stored at room temperature for later analysis by SDS-PAGE. FadD-His in the lysate was then bound to 200 μl NiNTA beads (Qiagen). Beads were washed twiced with 4 mL of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM immidazole, pH 8.0) and eluted twice in 1 mL of elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM Immidazole). 5 μl of flow through, wash, and elution samples were

taken as above to monitor the purification. Eluate was then transferred to an Amicon ® Ultra-15 Centrifugal Filter Ultracel ® with 30 kDa molecular weight cut off (Millipore). Samples were centrifuged at 4000 rpm for 15 min at in a bench top Centrifuge 5810 R (Eppendorf). The flow through was discarded, 12 mL buffer C (20 mM Tris-HCl, 150 mM NaCl, pH 8.0) added to the column, and the process repeated twice. Samples were centrifuged similarly a final time, resuspended in 2 mL buffer C, TCEP added to a final concentration of 5 mM, and stored at 4° C overnight for kinetic analysis the next day or glycerol added to a final concentration of 20% and the samples flash frozen in liquid nitrogen and stored at -80° C. All samples in 1X Tris-Glycine sample buffer were then visualized by SDS-PAGE and coomassie stained to ensure proper purification.

AMP Production Assay

Kinetic assays coupling the FadD catalyzed production of acyl-CoAs and AMP from oleate and octanoate to the oxidation of NADH were monitored spectrophotometrically via measuring absorbance at 340 nm in a Synergy NEO HTS Multi Mode microplate reader (BioTek) [16]. Reactions were carried out at 30° C in 100 µl total of freshly prepared 20 mM Tris-HCl pH 7.5, 2.5 mM ATP, 8 mM $MgCl_2$, 2 mM EDTA, 0.1% Triton X-100, 0.5 mM CoA, 0.2 mM NADH, 0.3 mM phosphoenolpyruvate (PEP), 48 U Myokinase from Chicken Muscle (Sigma), 96 U Pyruvate Kinase From Rabbit Muscle (Sigma), 48 U of Lactic Dehydrogenase (Sigma), 0.2 µg Ni-NTA purified FadD, and the appropriate amount of fatty acid from 1000 X stock solutions in ethanol (Oleate concentrations: 2.66-170 µM, Octanoate concentrations: 15.0-964 µM). Reactions were initiated with the addition of CoA and absorbance at 340 nM was measured every 30 s for 10 minutes. To ensure that the reactions were limited by FadD and not

by the coupled enzymes, prior to measuring the activities of other purified enzymes, kinetics of wild-type FadD with oleate as substrate were determined with 0.4, 0.2, and 0.1 μg FadD. If the oleate V_{max} increased proportionally with the amount of enzyme, it was assumed the coupling enzymes were not limiting.

Acyl-CoA production assay

Acyl-CoA production assays directly measured the production of 14-C fatty acyl-CoAs based on fatty acyl-CoA partitioning into an aqueous phase vs. organic phase after the CoA synthetase reaction [16]. Assay mixtures contained 1.6 μg of Ni-NTA purified FadD (or the appropriate mutant), 0.05 M Tris-HCl pH 8.0, 0.01 M $MgCl_2$, 0.01% Triton X-100, 10 mM ATP, and 0.3 mM DTT in 1 mL. Radiolabled fatty acids were included at a final concentration of 0-1.5 mM for octanoate, 100 μM for oleate, and 50 μM for decanoate. Reactions were initiated via the addition of 200 μM CoA and 0.25 mL periodically transferred to separate tubes containing 1.25 mL stopping buffer (40:10:1 isopropanol:n-heptane:1M H_2SO_4) to terminate the reaction. The terminated mixtures were then extracted 3 times with 1 mL n-heptane and radioactivity in 200 μl of the remaining aqueous phase measured by liquid scintillation counting in an LS6500 Multi-purpose Scintillation Counter (Beckman Coulter). Counts determined in this way were plotted over time and standards containing known amounts of fatty acid were used to determine the counts/nmol fatty acid. Slopes of the counts v time plots were then converted to nmol fatty acid/time giving the acyl-CoA production rate.

K_m and V_{max} determination

Once enzymatic rates were determined for each FadD preparation, rates were plotted against the concentration of fatty acid substrate used in each reaction and curves were fit to the Michaelis Menten equation using the nlinfit fuction in MATLAB Release 2010b (The Mathworks, Inc., Natick, Massachusetts, United States).

$$V = V_{max}[x]/(K_m + [x])$$

Where x is the concentration of fatty acid and V is the rate of acyl-CoA production determined as indicated above. Only curves with R^2 values >0.9 were accepted.

To normalize kinetic assay results for protein purity, prior to running either assay, wild-type FadD and its variants were visualized by SDS-PAGE and Coomassie staining. The full-length FadD bands were then quantified in ImageJ [28]. Wild-type FadD band intensities were used to adjust all mutant protein concentrations used for rate determinations by the relative intensity of each full-length mutant band to the intensity of the full-length wild-type FadD band.

TSS Competent Cell Preparation and Transformation

All transformations were performed according to the TSS competent cell protocol described previously [29].

Homology Modeling

FadD homology models were generated using The SWISS-MODEL Homology modeling server [30-32] and the *Thermus thermophilus* structure as the template, the I-TASSER server

[33-35], and (iii) SAM-T08 [36-43]. Models were visualized in Mac Pymol (The PyMOL Molecular Graphics System, Version 1.7rc1 Schrödinger, LLC.) and Swiss-PdbViewer [44].

Results

Mutations generated in the FadD coding sequence increase E. coli growth rate on MCFAs but not LCFAs

fadD mutants generated by error prone PCR confer increased *E. coli* growth rate on octanoate. We generated mutations in the *fadD* coding sequence using error prone PCR and screened mutants for their ability to increase *E. coli* growth rate on octanoate (Figure 4.1, Materials and Methods). Plasmids from strains with increased growth rate over controls were isolated and sequenced. In total, seven FadD single mutants conferred increased growth rate on octanoate (Figure 4.1B).

The FadD mutants generated by error-prone PCR do not increase FadD expression. To ensure that the FadD mutants do not increase growth rate by simply enhancing FadD expression, wild-type FadD and the FadD mutants were C-terminally His_6-tagged, growth was measured (Figure 4.1B), and SDS-PAGE samples were prepared at early exponential phase (~26 h of growth in octanoate minimal medium). Samples were normalized for total protein by A_{280} and western blotted with an anti-his antibody (Materials and Methods). While increases in growth rate were very consistent ($p<0.05$ in all cases), there was no significant difference in FadD expression between the wild-type and mutant variants (Figure 4.2C).

FadD mutants increase growth rate on the MCFAs hexanoate, octanoate, and moderately on decanoate, but do not increase growth rate on the LCFAs palmitate and oleate. To determine whether the effects of these mutations, selected on octanoate, were specific to octanoate or were

more broadly effective on fatty acids of different chain lengths, we measured their effects on growth rate in hexanoate (C6), decanoate (C10), palmitate (C16), and oleate (C18) minimal medium. The mutants had strong effects on hexanoate and octanoate medium, but only marginally increased growth rate on decanoate and failed to alter growth rate on palmitate, or oleate (Figure 4.2).

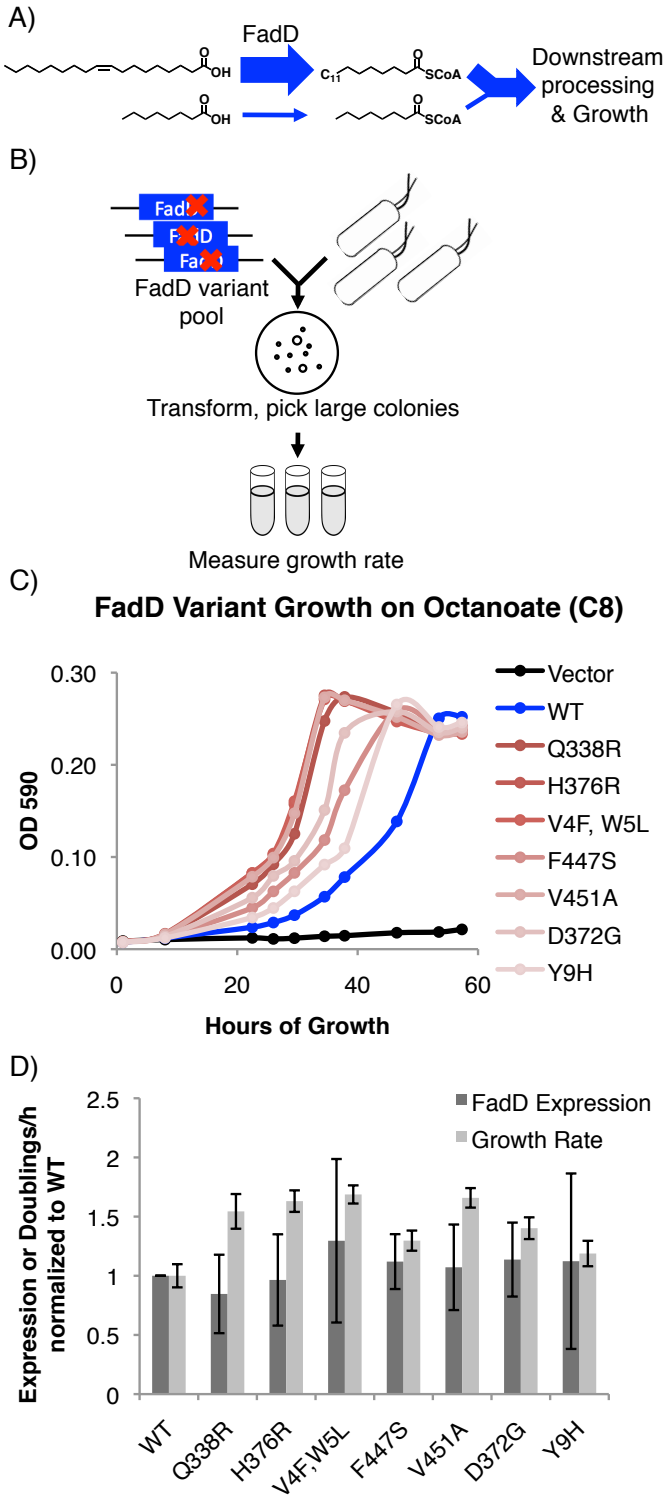

Figure 4.1. FadD mutants generated by error prone PCR increase *E. coli ΔfadR* growth rate on octanoate without increasing FadD expression. A) FadD catalyzses the first step in *E. coli* growth on fatty acids but has low activity on fatty acids shorter than 10 carbons. B) Error prone PCR and FadD screening scheme (Materials and Methods). C) Growth of *E. coli ΔfadR* expressing the indicated C-terminally His_6-tagged FadD mutants generated by error prone PCR from vector

Figure 4.1 (Continued). pETDuet-1 on octanoate. D) Relative increase in FadD expression (dark gray) and growth rate (light gray) conferred by His_6-tagged FadD mutants on octanoate compared to wild-type FadD. n=5 for FadD expression and 6 for growth rate; error bars indicate standard deviation. All increases in growth rate have $p<0.05$ by two sided students T-test while all changes in expression have $p>0.3$. FadD expression was measured using anti-his western blot samples normalized to total protein content by A_{280} (Materials and Methods).

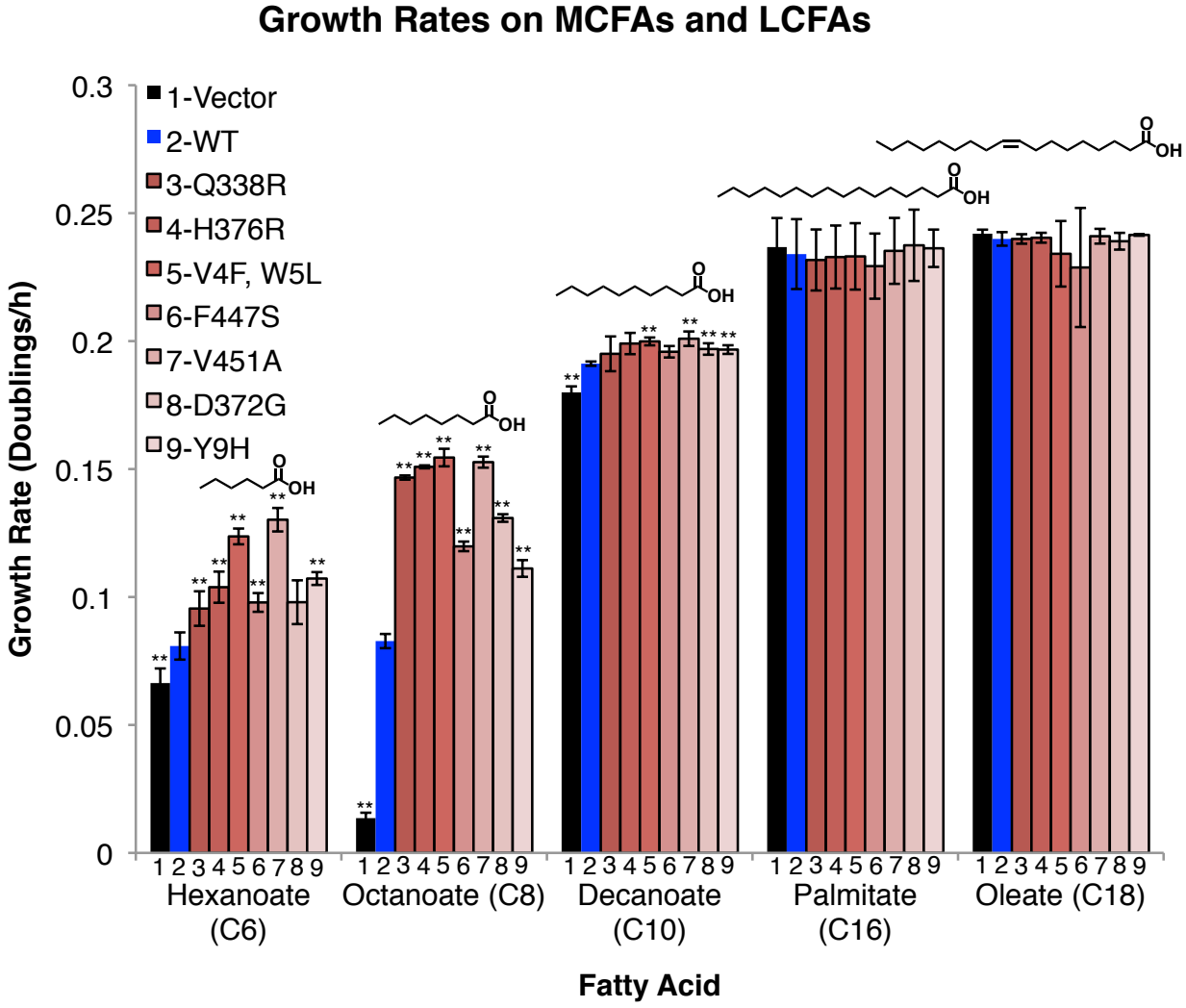

Figure 4.2. FadD mutants enhance the growth rate of *E.coli ΔfadR* on the MCFAs hexanoate, octanoate, and decanoate, but not on palmitate and oleate. *E.coli ΔfadR* transformed with empty pETDuet-1 (black) C-terminally His_6-tagged wild-type *fadD* (blue) or the indicated C-terminally His_6-tagged *fadD* mutants (Red) were grown on minimal medium containing the indicated fatty acid as the sole carbon source. Growth rates were measured by linear regression of the normalized $\log_2$(OD590) during exponential phase. n=3, errors bars indicate standard deviation, and ** indicates $p<0.05$ compared to wild-type by two sided students T-test.

FadD mutant proteins have increased activity on octanoate and decanoate, but not oleate

In-vitro assays measuring AMP production by Ni-NTA purified His_6-tagged FadD mutants (Figure 4.3) showed that they have increased activity on MCFAs but not LCFAs. The assay coupled AMP production in the acyl-CoA synthetase reaction to the oxidation of NADH which was monitored spectrophotometrically (Materials and Methods) [16] (Figure 4.4). The V_{max} values of the mutants were higher than those of wild-type FadD on octanoate, but were generally lower on oleate (with the exception of mutant H376R). There were no significant changes in the K_m toward octanoate for each of the mutants, although the mutant H376R showed an increased K_m toward oleate while Y9H had a decreased K_m toward oleate (Figure 4.5A and B). These results indicate that, while the FadD mutations increase the rate of the acyl-CoA synthetase reaction, they do not generally enhance FadD affinity for octanoate.

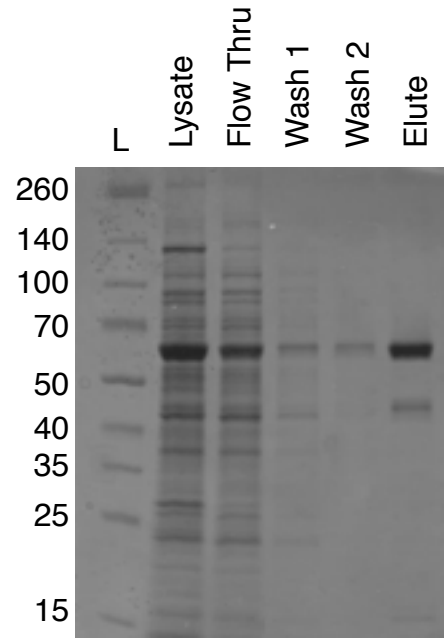

Figure 4.3. Sample partial purification of C-terminally His_6-tagged FadD. Lanes contain: (L) Spectra Multicolor Broad Range Protein Ladder (Thermo Scientific) with the molecular weight (kDa) of each band in the ladder next to the gel, (Lystate) lysate fraction from Ni-NTA partial purification, (Flow Thru), flow thru fraction from the Ni-NTA partial purification, (Wash 1 and 2) wash fractions from the Ni-NTA partial purification, and (Elute) pooled elutant fractions from the Ni-NTA partial purification of wild-type FadD (Materials and Methods).

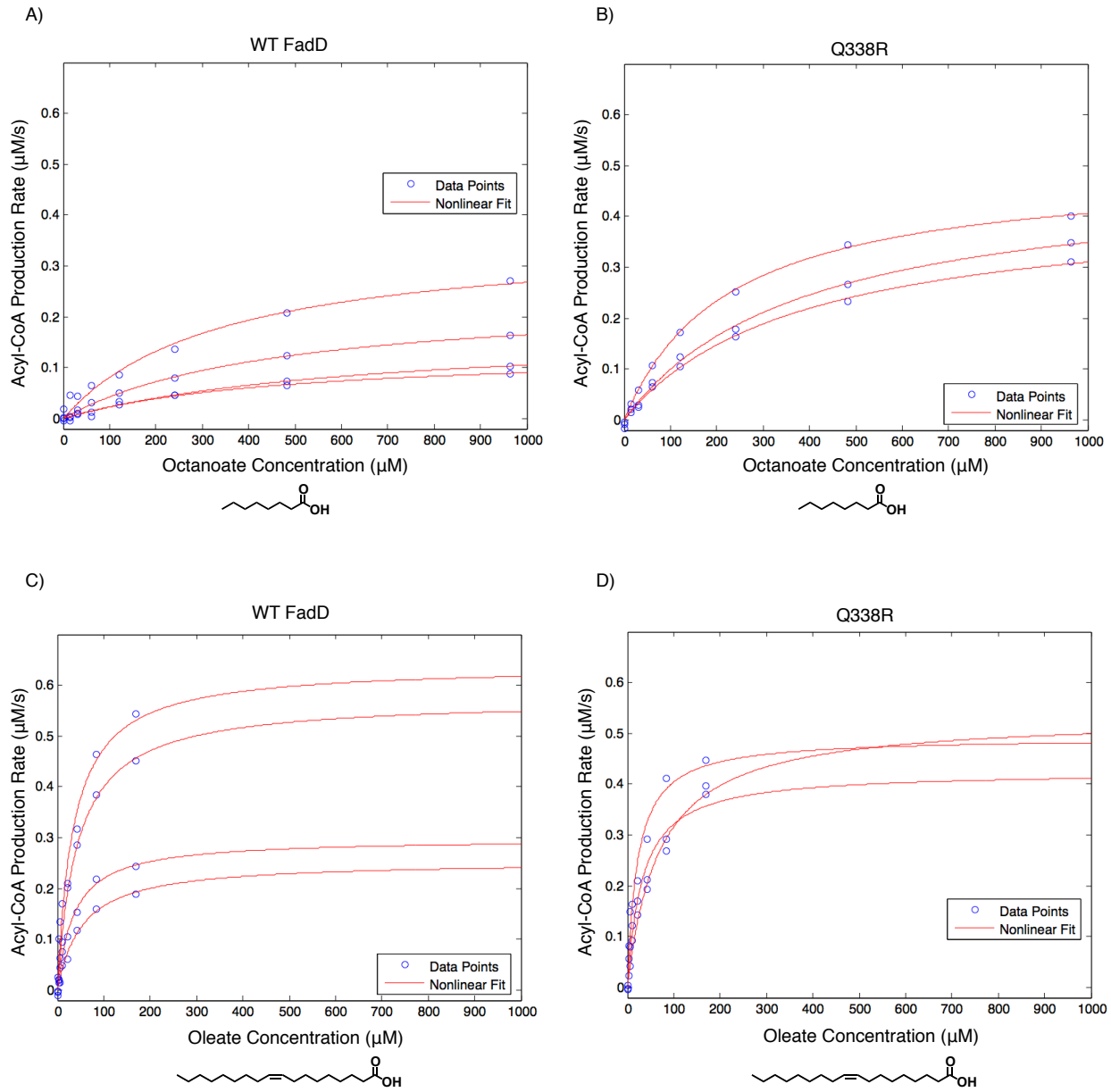

Figure 4.4. Sample rate vs. substrate concentration curves for the AMP production assay. Blue circles indicate rates determined from independent purifications and red lines are nonlinear fits to the Michaelis Menten equation for wild-type FadD and mutant Q338R using octanoate (A,B) and oleate (C,D) as substrates.

Calculated catalytic efficiencies (K_{cat}/K_m) show that all mutants except Q338R were less efficient than wild-type when oleate was used as a substrate, but four of the mutants (Q338R, V451A, D372G and Y9H) were more efficient than wild-type when using octanoate. The remainder of the mutants had lower or equivalent catalytic efficiency on octanoate indicating that

decreases in affinity toward octanoate (higher K_m) outweighed or matched increases in overall activity (higher K_{cat}).

A second *in-vitro* assay directly measuring acyl-CoA production showed that the mutants have increased activity on decanoate and octanoate but not oleate. Rates determined using decanoate and oleate as substrates at concentrations roughly 10 times their published K_m values [16] in the acyl-CoA production assay (Material and Methods) showed that, while most of the mutants have increased activity on decanoate, none have significant increases in activity on oleate and two have decreased activity (Figure 4.5C and D). This is consistent with the data from the AMP production assays.

FadD mutant proteins had higher maximal activity on octanoic acid in acyl-CoA production assays, consistent with the AMP production assays (data not shown). However, the rates determined in these assays had high background and poor fits to the Michaelis-Menten curve. Presumably the high background activity was due to the higher solubility of octanoate as compared to oleate or decanoate, which both gave lower background and more consistent measurements.

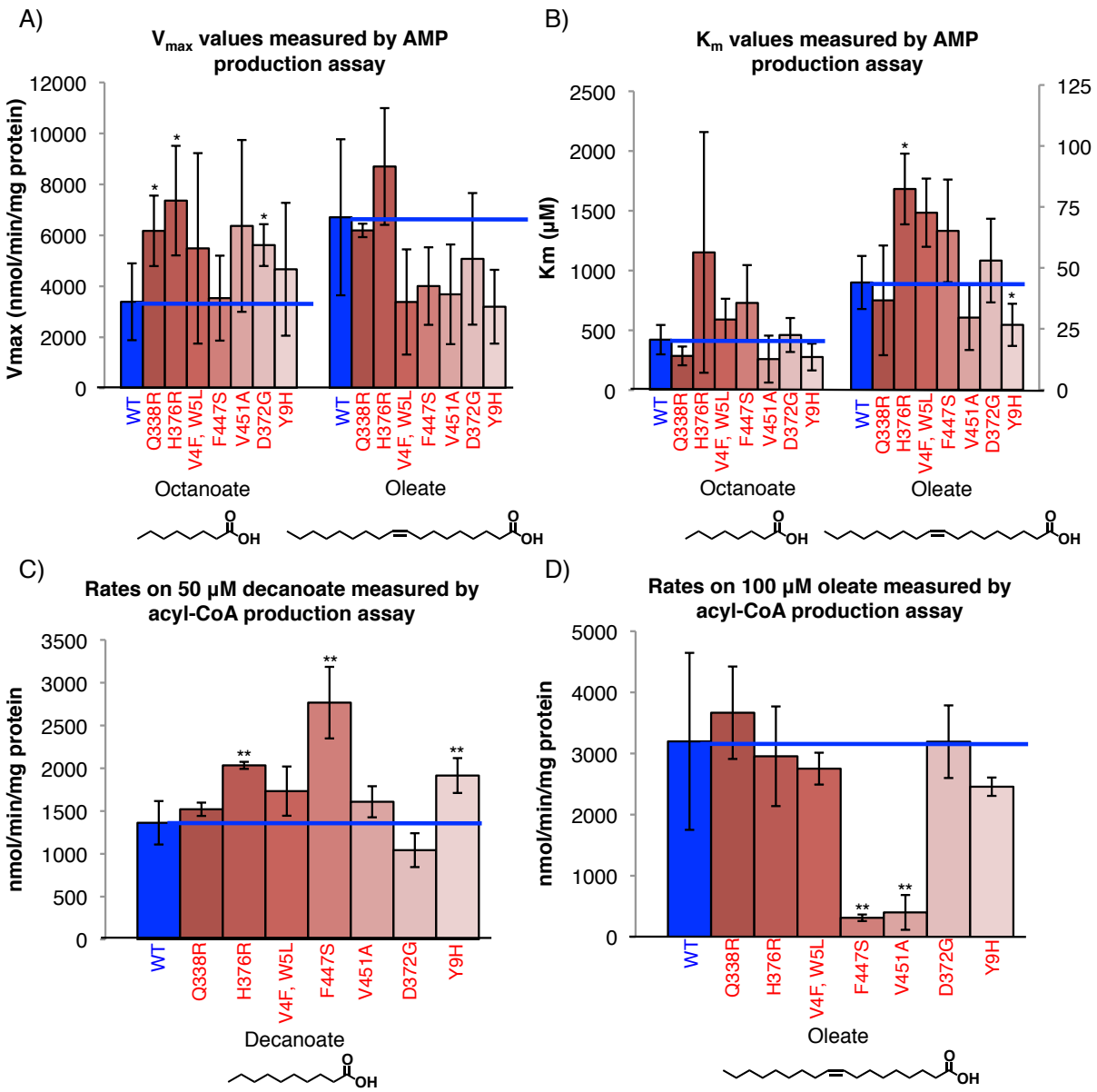

Figure 4.5. FadD mutants have increased activity on MCFAs but unaltered affinity for MCFAs. A) His_6-tagged wild-type FadD and the indicated mutants were partially purified via Ni-NTA purification (Materials and Methods) and steady state activity on the indicated fatty acid measured spectrophotometrically using the AMP production assay (Materials and Methods)[16]. V_{max} (A) and K_m (B) values for the indicated substrates. n=3-4 independent purifications, error bars indicate standard deviation, * indicates $p<0.1$ compared to wild-type by two sided students T-test. C) and D) Steady state rate of acyl-CoA production using 1.6 μg of Ni-NTA purified, C-terminally His_6-tagged wild-type FadD and the indicated mutants with decanoate (C) and oleate (D) as substrates at concentrations roughly 10 times the literature reported K_m values (Materials and Methods) [16]. n=3 independent measurements from one or two purifications, error bars indicate standard deviation, and ** indicates $p<0.05$ by two sided student's T-test compared to wild-type FadD.

Table 4.4. Catalytic efficiency (K_{cat}/K_m) of wild-type FadD and the indicated mutants as measured by AMP production assay

FadD Variant	Octanoate K_{cat}/K_m [a]	Oleate K_{cat}/K_m [a]
WT	0.10 ± 0.07	1.70 ± 0.94
Q338R	0.25 ± 0.13	1.98 ± 1.07
H376R	0.10 ± 0.06	1.16 ± 0.54
V4F, W5L	0.10 ± 0.05	0.46 ± 0.20
F447S	0.05 ± 0.02	0.74 ± 0.50
V451A	0.35 ± 0.28	1.46 ± 1.01
D372G	0.14 ± 0.05	1.09 ± 0.65
Y9H	0.20 ± 0.07	1.38 ± 0.80

[a]: $(M^{-1}*s^{-1} * 10^5)$
Values are indicated ± standard deviation

Site directed FadD mutants designed to open a proposed AMP exit channel increase E.coli ΔfadR growth rate on octanoate

FadD homology models generated using the SWISS-MODEL Homology modeling server [30-32] and the *Thermus thermophilus* structure as the template, the I-TASSER server [33-35], and SAM-T08 [36-43], show that several of the FadD mutations cluster around a possible ATP/AMP entrance/exit channel (Figure 4.6A, Figure 4.7). All models have features similar to those of known adenylating enzymes as well as the acyl-CoA synthetase from *Thermus thermophilus* [18-23]. These include a small, globular C-terminal domain (white), a large, globular N-terminal domain (grey), and an active site (annotated by the alignment in [18]) situated between the two domains. Comparing these homology models to the structure of the

Thermus thermophilus acyl-CoA synthetase shows that several of our FadD mutations cluster on a face of the protein from which ATP and AMP are proposed to enter and exit the active site [18]. Hisinaga *et al*., 2004 inferred that ATP binding precedes and enhances fatty acid binding, so enhancement of ATP binding would likely decrease the K_m for the fatty acid. Given that our mutants fail to decrease K_m, but do increase V_{max} toward octanoate, we hypothesize that they could facilitate AMP exit from the active site by opening this face of the protein.

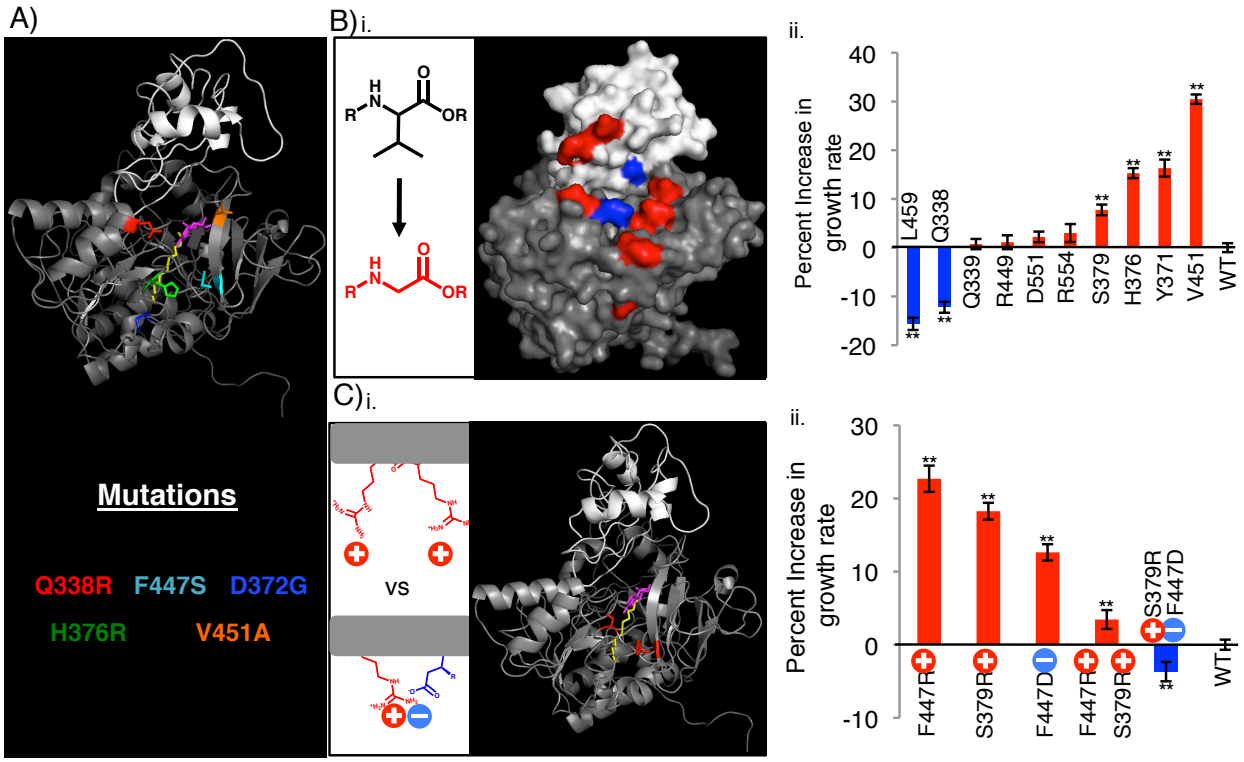

Figure 4.6. Rationally designed, site directed FadD mutants increase *E. coli ΔfadR* growth rate on octanoate when compared to wild-type FadD. A) FadD homology model generated using The SWISS-MODEL Homology modeling server [30-32] and the *Thermus thermophilus* structure as the template. The model was visualized in PyMOL with large N-terminal domain in gray, smaller C-terminal domain in white, and myristoyl-AMP (overlayed from the *Thermus thermophillus* structure) in yellow (myristoyl group) and magenta (AMP) [18]. Residues whose mutation results in increased growth rate on octanoate are color-coded according to the identity of the mutation (text below model, Y9H and V4F W5L are excluded from the model). B) (i) Surface representation of the FadD homology model with residues mutated to glycine in (ii) shown in blue (mutations that decrease growth rate compared to wild-type) and red (mutations that increase growth rate compared to wild-type). (ii) Percent increase in exponential growth rate compared to wild-type FadD caused by mutating the residues on the X-axis to glycine. C) (i) Cartoon representation of FadD homology model with residues mutated in (ii) in red. (ii) Percent increase in exponential growth rate compared to wild-type FadD caused by the FadD mutations

Figure 4.6 (Continued). depicted on the X-axis. n=13-18, error bars indicate standard error in all cases, ** indicates growth rate significantly different from wild-type with $p<0.05$ by two-sided students T-test.

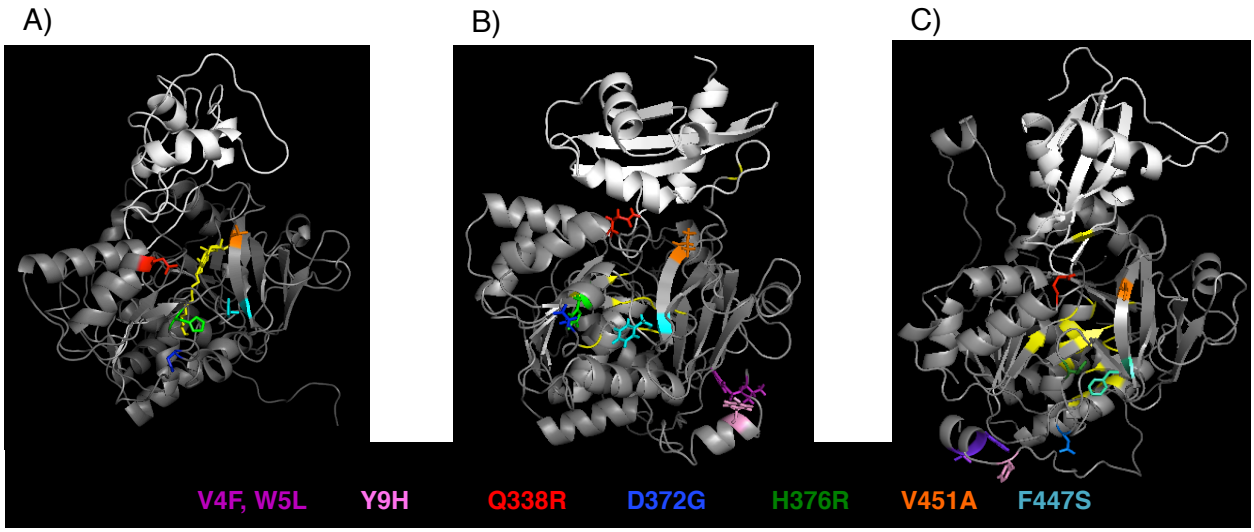

Figure 4.7. FadD Homology Modeling. Three separate FadD homology models were generated using (i) The SWISS-MODEL Homology modeling server [30-32] and the *Thermus thermophilus* structure as the template, (ii) the I-TASSER server [33-35], and (iii) SAM-T08 [36-43]. Models were visualized in PyMOL with large N-terminal domain in gray, smaller C-terminal domain in white, and fatty acid binding pocket or myristoyl-AMP (overlayed from the *Thermus Thermophillus* structure) in yellow [18,43]. Residues whose mutation results in increased growth rate on octanoate are color-coded according to the identity of the mutation (text below models).

FadD mutations designed to facilitate AMP exit from the FadD active site increased growth rate on octanoate (Figure 4.6). To test the hypothesis that opening the FadD AMP exit channel could facilitate product exit and increase FadD activity on MCFAs, we removed amino acid side chains surrounding the channel by mutating their associated residues to glycine, and measured the resultant mutants' growth on octanoate. Eight out of ten of these mutations increased the average growth rate of *E.coli ΔfadR* (JW1176-1) compared to wild-type. Two of these ten mutations decreased growth rate (Figure 4.6B).

Further mutations designed to electrostatically repel structurally adjacent amino acids (S379, F447) and thereby destabilize the closed confirmation of FadD and aid AMP exit

enhanced growth rate on octanoate. Mutations designed to electrostatically attract these same amino acids decreased growth rate on octanoate. The mutations were made separately and in combination and their effects on the growth rate of *E.coli ΔfadR* (JW1176-1) on octanoate minimal medium were measured. Specifically, residues S379 and F447, which are adjacent to each other in the FadD homology model (Figure 4.6Ci), were each mutated to arginine and aspartate singly and in combination. Each individual mutation and the double mutant designed to repel these residues and destabilize the closed confirmation of FadD (S379R, F447R) enhanced growth rate on octanoate. In contrast, the double mutant designed to form a salt bridge between these residues and stabilize the closed confirmation of FadD (S379R, F447D) decreased growth rate on octanoate.

4. Discussion

This work shows that *E. coli* FadD activity limits the conversion of medium chain fatty acids (MCFAs; 6-12 carbons) to acyl-CoA thioesters and provides a set of FadD mutants that will be useful in expediting this conversion. Medium chain acyl-CoA thioesters are important precursors to medium chain fatty alcohols and alkanes that could be used as next generation biofuels. While means to reduce MCFAs directly into alcohols using a carboxylic acid reductase (CAR) have been reported [7,45,46], CARs contain both thioester forming activity and thioester reducing activity. The FadD mutants created in this work provide a means to alter only the thioester formation portion of this process and may prove useful in tuning pathway flux or product composition.

We identified mutations in the *E. coli* K12 *fadD* gene by constructing a library of altered genes via PCR mutagenesis, transformation, and screening for enhanced growth on octanoic acid

(Figure 4.1). The mutant genes significantly increased the host growth rate on hexanoate and octanoate, somewhat on decanoate, and not at all on palmitate or oleate (Figure 4.2). Kinetic assays indicated that the FadD mutant proteins have an increased V_{max} toward octanoate, without significant effects on K_m. These results suggest that these mutations increase activity without enhancing substrate binding. Given that our FadD mutants were screened on 6.9 mM octanoate minimal medium, a concentration far in excess of the wild-type FadD K_m values determined here (422 μM for the coupled assay and 268 μM for the radioactivity assay), it is perhaps unsurprising that mutations conferring higher affinity for octanoate were not discovered.

The mechanism of FadD is complex and involves multiple substrate-binding and product exit steps through different channels in the protein. Hisanaga *et al.* solved a structure of a FadD homologue from *Thermus thermophilus*, with and without an AMP-fatty acid intermediate [18]. Based on these structures and prior biochemistry, they proposed that, as the FadD protein is a non-integral membrane-associated protein, the fatty acid enters from the membrane through a narrow channel in the back of the protein, while ATP enters through a distinct, large channel. ATP and the fatty acid bind first and form the AMP-fatty acid intermediate, releasing pyrophosphate. At this point, a flexible C-terminal domain, clamps onto the AMP-fatty acid intermediate to prevent its escape and position it for nucleophilic attack by CoA, which then binds and attacks the phosphoester bond, generating AMP and fatty acyl-CoA. Kochan *et al*., 2009 determined the structures of a human medium-chain Acyl-CoA Synthetase with ATP and butyryl-CoA/AMP in the active site. The pantotheine group of CoA enters by a third channel in the protein, distinct from the ATP and fatty acid entry sites [24].

When mapped onto a homology model of FadD, our mutations are nowhere near the binding sites of either the fatty acid or CoA, but some border on the ATP/AMP channel and

amino acids that may directly or indirectly affect the interaction of the flexible C-terminal domain with the rest of the protein; these include Val 451, which may make direct contact and Asp 372, His 376, Gln 338, and Phe 447, which may indirectly affect the structure of the AMP exit channel or the interaction of this region of the protein with the C-terminal domain. Additionally, none of our mutations fall in a region (residues 422-430) involved in fatty acid binding as shown by affinity labeling experiments [26]. We hypothesize that when CoA bonds to a long-chain fatty acid, the AMP is sterically pushed from the active site by this product. When CoA bonds to an MCFA on the other hand, it may move within the active site so that this steric push is less pronounced. The effect of these mutations might be to ease the transition to an open state and enhance AMP exit, which would result in the observed increase in V_{max}.

FadD mutants should be useful for the production of MCFA-derived chemicals. We have previously produced >1.5 mM octanoate in *E. coli* engineered for MCFA production [5]. If similar yields can be achieved in a strain further engineered to convert octanoyl-CoA into downstream products, enhancing the V_{max} for acyl-CoA production would be more beneficial than improving K_m because 1.5 mM octanoate is in roughly five fold excess of the mutant K_m for octanoate. In any well-engineered metabolic pathway, MCFAs are likely to be abundant intermediates. Assuming FadD activity limits further downstream conversions, our most effective mutants could potentially double the rate of MCFA to downstream product conversion. The faster MCFA to acyl-CoA conversion could limit unwanted side reactions, enhance MCFA uptake [47], prevent known MCFA toxicity [48-50], and enhance downstream product yields. Future work screening FadD mutants on lower MCFA concentrations could produce FadD mutants with decreased K_m toward MCFAs.

This work adds to our growing knowledge on structural determinants of FadD substrate specificity. One of the mutations discovered here (V451A) falls in the previously characterized FACS motif [25]. Our results confirm that this mutation increases FadD activity on decanoate (as shown previously) and further show that this mutation enhances activity on octanoate and hexanoate. The remainder of our mutations falls outside this motif. This agrees well with our kinetic data that the mutant proteins have unchanged K_m values and thus likely do not enhance MCFA binding.

Some of our mutants have unaltered activity on oleate (Q338R and H376R) while others have decreased activity toward oleate (F447S and V451A). These two types of mutants could prove useful for different reasons. For example, future work requiring a mixture of acyl-CoA lengths would benefit from the mutants with high activity on both MCFAs and LCFAs, while work specifically producing only medium chain products would benefit from the mutants with decreased activity on LCFAs and increased activity on MCFAs.

More broadly, there are many adenylate-forming enzymes like FadD that have similar structures and functions but modify different substrates [19-23]. The method of increasing activity on a similar but smaller substrate by aiding product exit may be applicable to other adenylate-forming enzymes with similar structures. This approach could open possibilities for engineering the degradation and modification of a variety of substrates important for applications ranging from lignin processing [20] to antibiotic production [22].

References

1. Handke P, Lynch SA, Gill RT: **Application and engineering of fatty acid biosynthesis in Escherichia coli for advanced fuels and chemicals**. *Metabolic engineering* 2011, **13**:28-37.

2. Knothe G: **Improving biodiesel fuel properties by modifying fatty ester composition**. *Energy and Environmental Science* 2009, **2**:8.

3. Voelker TA, Davies HM: **Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of a plant medium-chain acyl-acyl carrier protein thioesterase**. *Journal of bacteriology* 1994, **176**:7320-7327.

4. Dehesh K, Edwards P, Hayes T, Cranmer AM, Fillatti J: **Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of Cuphea palustris seed oil**. *Plant physiology* 1996, **110**:203-210.

5. Torella JP, Ford TJ, Kim SN, Chen AM, Way JC, Silver PA: **Tailored fatty acid synthesis via dynamic control of fatty acid elongation**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**:11290-11295.

6. Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R: **Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals**. *Nature* 2011, **476**:355-359.

7. Akhtar MK, Dandapani, H., Thiel, K., & Jones, P. R.: **Microbial Production of 1-octanol: A naturally excreted biofuel with diesel-like properties**. *Metabolic Engineering Communications* 2015, **2**:1-5.

8. Choi YJ, Lee SY: **Microbial production of short-chain alkanes**. *Nature* 2013, **502**:571-574.

9. Black PN, DiRusso CC, Metzger AK, Heimert TL: **Cloning, sequencing, and expression of the fadD gene of Escherichia coli encoding acyl coenzyme A synthetase**. *The Journal of biological chemistry* 1992, **267**:25513-25520.

10. Zhang F, Carothers JM, Keasling JD: **Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids**. *Nature biotechnology* 2012, **30**:354-359.

11. Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, Del Cardayre SB, Keasling JD: **Microbial production of fatty-acid-derived fuels and chemicals from plant biomass**. *Nature* 2010, **463**:559-562.

12. Doan TT, Carlsson AS, Hamberg M, Bulow L, Stymne S, Olsson P: **Functional expression of five Arabidopsis fatty acyl-CoA reductase genes in Escherichia coli**. *Journal of plant physiology* 2009, **166**:787-796.

13. Salanitro JP, Wegener WS: **Growth of Escherichia coli on short-chain fatty acids: growth characteristics of mutants**. *Journal of bacteriology* 1971, **108**:885-892.

14. Overath P, Pauli G, Schairer HU: **Fatty acid degradation in Escherichia coli. An inducible acyl-CoA synthetase, the mapping of old-mutations, and the isolation of regulatory mutants**. *European journal of biochemistry / FEBS* 1969, **7**:559-574.

15. Campbell JW, Morgan-Kiss RM, Cronan JE, Jr.: **A new Escherichia coli metabolic competency: growth on fatty acids by a novel anaerobic beta-oxidation pathway**. *Molecular microbiology* 2003, **47**:793-805.

16. Kameda K, Nunn WD: **Purification and characterization of acyl coenzyme A synthetase from Escherichia coli**. *The Journal of biological chemistry* 1981, **256**:5702-5707.

17. Iram SH, Cronan JE: **The beta-oxidation systems of Escherichia coli and Salmonella enterica are not functionally equivalent**. *Journal of Bacteriology* 2006, **188**:599-608.

18. Hisanaga Y, Ago H, Nakagawa N, Hamada K, Ida K, Yamamoto M, Hori T, Arii Y, Sugahara M, Kuramitsu S, et al.: **Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer**. *The Journal of biological chemistry* 2004, **279**:31717-31726.

19. Gulick AM, Starai VJ, Horswill AR, Homick KM, Escalante-Semerena JC: **The 1.75 A crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A**. *Biochemistry* 2003, **42**:2866-2873.

20. Hu Y, Gai Y, Yin L, Wang X, Feng C, Feng L, Li D, Jiang XN, Wang DC: **Crystal structures of a Populus tomentosa 4-coumarate:CoA ligase shed light on its enzymatic mechanisms**. *The Plant cell* 2010, **22**:3093-3104.

21. Conti E, Franks NP, Brick P: **Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes**. *Structure* 1996, **4**:287-298.

22. Conti E, Stachelhaus T, Marahiel MA, Brick P: **Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S**. *The EMBO journal* 1997, **16**:4174-4183.

23. Gulick AM: **Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase**. *ACS chemical biology* 2009, **4**:811-827.

24. Kochan G, Pilka ES, von Delft F, Oppermann U, Yue WW: **Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A**. *Journal of molecular biology* 2009, **388**:997-1008.

25. Black PN, Zhang Q, Weimar JD, DiRusso CC: **Mutational analysis of a fatty acyl-coenzyme A synthetase signature motif identifies seven amino acid residues that modulate fatty acid substrate specificity**. *The Journal of biological chemistry* 1997, **272**:4896-4903.

26. Black PN, DiRusso CC, Sherin D, MacColl R, Knudsen J, Weimar JD: **Affinity labeling fatty acyl-CoA synthetase with 9-p-azidophenoxy nonanoic acid and the identification of the fatty acid-binding site**. *The Journal of biological chemistry* 2000, **275**:38547-38553.

27. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H: **Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection**. *Molecular systems biology* 2006, **2**:2006 0008.

28. Schneider CA, Rasband WS, Eliceiri KW: **NIH Image to ImageJ: 25 years of image analysis**. *Nature methods* 2012, **9**:671-675.

29. Chung CT, Niemela SL, Miller RH: **One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution**. *Proceedings of the National Academy of Sciences of the United States of America* 1989, **86**:2172-2175.

30. Benkert P, Biasini M, Schwede T: **Toward the estimation of the absolute quality of individual protein structure models**. *Bioinformatics* 2011, **27**:343-350.

31. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L, et al.: **SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information**. *Nucleic Acids Research* 2014, **42**:W252-W258.

32. Arnold K, Bordoli L, Kopp J, Schwede T: **The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling**. *Bioinformatics* 2006, **22**:195-201.

33. Yang JY, Yan RX, Roy A, Xu D, Poisson J, Zhang Y: **The I-TASSER Suite: protein structure and function prediction**. *Nature Methods* 2015, **12**:7-8.

34. Roy A, Kucukural A, Zhang Y: **I-TASSER: a unified platform for automated protein structure and function prediction**. *Nature Protocols* 2010, **5**:725-738.

35. Zhang Y: **I-TASSER server for protein 3D structure prediction**. *Bmc Bioinformatics* 2008, **9**.

36. Karchin R, Cline M, Karplus K: **Evaluation of local structure alphabets based on residue burial**. *Proteins* 2004, **55**:508-518.

37. Karchin R, Cline M, Mandel-Gutfreund Y, Karplus K: **Hidden Markov models that use predicted local structure for fold recognition: alphabets of backbone geometry**. *Proteins* 2003, **51**:504-514.

38. Karplus K: **SAM-T08, HMM-based protein structure prediction**. *Nucleic acids research* 2009, **37**:W492-497.

39. Karplus K, Hu B: **Evaluation of protein multiple alignments by SAM-T99 using the BAliBASE multiple alignment test set**. *Bioinformatics* 2001, **17**:713-720.

40. Karplus K, Karchin R, Barrett C, Tu S, Cline M, Diekhans M, Grate L, Casper J, Hughey R: **What is the value added by human intervention in protein structure prediction?** *Proteins* 2001, **Suppl 5**:86-91.

41. Karplus K, Karchin R, Draper J, Casper J, Mandel-Gutfreund Y, Diekhans M, Hughey R: **Combining local-structure, fold-recognition, and new fold methods for protein structure prediction**. *Proteins* 2003, **53 Suppl 6**:491-496.

42. Karplus K, Katzman S, Shackleford G, Koeva M, Draper J, Barnes B, Soriano M, Hughey R: **SAM-T04: what is new in protein-structure prediction for CASP6**. *Proteins* 2005, **61 Suppl 7**:135-142.

43. Shackelford G, Karplus K: **Contact prediction using mutual information and neural nets**. *Proteins* 2007, **69 Suppl 8**:159-164.

44. Guex N, Peitsch MC: **SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling**. *Electrophoresis* 1997, **18**:2714-2723.

45. Akhtar MK, Turner NJ, Jones PR: **Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**:87-92.

46. Kallio P, Pasztor A, Thiel K, Akhtar MK, Jones PR: **An engineered pathway for the biosynthesis of renewable propane**. *Nature communications* 2014, **5**:4731.

47. Maloy SR, Ginsburgh CL, Simons RW, Nunn WD: **Transport of long and medium chain fatty acids by Escherichia coli K12**. *The Journal of biological chemistry* 1981, **256**:3735-3742.

48. Sherkhanov S, Korman TP, Bowie JU: **Improving the tolerance of Escherichia coli to medium-chain fatty acid production**. *Metabolic engineering* 2014, **25**:1-7.

49. Lennen RM, Pfleger BF: **Modulating membrane composition alters free fatty acid tolerance in Escherichia coli**. *PloS one* 2013, **8**:e54031.

50. Lennen RM, Kruziki MA, Kumar K, Zinkel RA, Burnum KE, Lipton MS, Hoover SW, Ranatunga DR, Wittkopp TM, Marner WD, 2nd, et al.: **Membrane stresses induced by overproduction of free fatty acids in Escherichia coli**. *Applied and environmental microbiology* 2011, **77**:8114-8128.

Chapter 5: Conclusion

The goals of this work were 1) to engineer *Escherichia coli* to produce precursors to biofuels more compatible with our liquid fuel infrastructure than ethanol and 2) to develop and demonstrate techniques and tools that can be used in a wide variety of synthetic biology applications. In reference to the first goal, we successfully produced medium chain fatty acids (MCFAs) in *E. coli*, but our results should be considered within the context of current biofuel engineering efforts and cost estimates of biofuels production. With reference to the second goal, further implementation of the techniques and tools we've developed will be the true test of their usefulness.

In the work discussed in chapter two, we engineered *E. coli* to produce MCFAs as precursors to medium chain biofuels that are more energy dense, more hydrophobic, and more compatible with our liquid fuel infrastructure than ethanol [1]. We achieved 12% theoretical yield of the 8-carbon fatty acid, octanoate from glucose as sole carbon source. To our knowledge, this was the highest yield of MCFAs produced using *E. coli* fatty acid synthesis at the time of publication, but it should be noted that long chain fatty acids (LCFAs) have been produced at 73% theoretical yield from glucose [2]. We clearly fall short of this fatty acid production benchmark. This high LCFA yield was achieved through the over-expression of the transcription factor FadR. Future studies should combine the MCFA engineering performed here with over-expression of FadR to attempt to achieve similar yields. Higher MCFA yields may also be achieved by combining the inducible degradation strategy used here with strategies that automatically switch engineered strains into production mode in the presence of adequate precursors [3,4]. This could be applied in our system by making degradation inducible by malonyl-CoA. Finally, high yields of MCFA derived products have additionally been achieved by abandoning fatty acid synthesis altogether and instead producing fatty acids and their

derivatives by reversing β-oxidation [5]. Fatty acid synthesis and the reversal of β-oxidation therefore offer competing means of achieving biofuels production, but further development is required to understand the benefits and trade-offs of each when implemented outside the lab.

While these techniques are all effective at lab scale (in shake flasks, test-tubes, or in small plates), it is unclear how effective they will be at larger scale. One potential problem with all highly engineered biological devices like the strains developed here is stability. Given that MCFA production directs cellular resources that could be used for growth into MCFAs, in a bioreactor with many bacteria, there will likely be selective pressure for mutants with decreased MCFA production. In our strains in particular, we specifically down-regulate a process essential for growth, the production of unsaturated fatty acids, through the degradation of FabB. If these strains were used in a large bioreactor, there could be selective pressure for mutants that fail to degrade FabB resulting in lower MCFA yields. In addition, it has been shown that biological circuits engineered at lab scale do not always function at larger scales without further engineering meaning that degradation and protein expression might require additional tuning at large scale to achieve high yields [6]. Lab scale demonstrations like those in this work are useful for showing that a production technique is possible, but future work should focus on MCFA production at larger scales.

While we did not produce any medium chain alcohols or alkanes as potential biofuels in this work, we did enhance the production of their acyl-CoA precursors. We attempted to produce the 8-carbon alcohol, octanol, in some of the strains we produced in chapter two by over-expressing acyl-CoA synthetase FadD and acyl-CoA reductases that could potentially convert octanoyl-CoA into octanol, but failed to achieve detectable octanol production. Preliminary efforts to produce octanol in our engineered strains by instead expressing a carboxylic acid

reductase that combines the activation and reducing activities of FadD and the acyl-CoA reductases into a single enzyme [1,7,8] have shown promising results. Nonetheless, the failure to produce octanol through wild type FadD provided the motivation behind the work in chapters three and four. The FadD mutants we generated in chapter four should prove useful for producing medium chain acyl-CoAs that can later be converted into alcohols and alkanes. Producing these compounds at high yields will likely require tuning the expression of the various enzymes involved.

To understand whether or not the strains developed here will be useful for commercial biofuels production, it is important to think about the role this work could play in the biofuels production process at large. As discussed in the introduction, NREL extensively modeled the production of hydrocarbons from lignocellulosic biomass in bacteria and estimated that a diesel fuel replacement derived from the LCFA palmitate could be sold at $5.10 per gasoline equivalent gallon if produced at 79% theoretical yield [9]. This assumes a higher theoretical yield than has been achieved for MCFA production meaning that our work can play a role in developing strains that can meet this model. However, the authors of the NREL study point out that the most important ways to lower cost will be to engineer microbes that can reliably convert lignin into biofuel product, to lower the cost and improve the performance of the enzymes that convert biomass feedstocks into simple sugars, and to lower capital costs by developing strains that can produce hydrocarbons under anaerobic conditions. We do not address any of these problems, but part of the importance of our work comes from its focus on the production of precursors to biofuels more like gasoline. The NREL study models the production of palmitate which has potential as a precursor to replacements for diesel fuels but these only make up about 25% percent of transportation fuels while gasoline makes up about 63% [10]. Our work can therefore

expand the GHG mitigation potential of biofuels made in a similar process by enabling the replacement of more than double the amount of fuel that could be replaced in the current design. Nonetheless, future studies may better benefit biofuels producers by focusing on the areas highlighted in the NREL study.

Given the need to achieve higher yields, the strains developed here for the production of MCFAs are unlikely to find direct application at commercial or industrial scale. However, the inducible degradation technique used in chapter two and the knowledge we gained from the development of the FadD mutants in chapter four may prove useful for the engineering of a variety of metabolic pathways. The inducible degradation technique is not limited to the degradation of FabB; there are many pathways that cannot simply be knocked out because doing so would prevent growth, but whose down-regulation might still benefit overall yields. For example, the citric acid cycle is required for the production of amino acids and other metabolites, but uses acetyl-CoA that could otherwise be used in fatty acid synthesis. We made attempts to slow the citric acid cycle by degradation tagging citrate synthase and phophoenolpyruvate carboxylase, two enzymes required to generate citric acid cycle intermediates, but our strains were hampered by poor growth. This indicated that the degradation tags may have impaired the function of the enzymes or that there was a high level of basal degradation. Nonetheless, with further tuning of degradation and trouble shooting the system with different degradation tags, this general strategy may prove useful in the future. Any pathways that use acetyl-CoA or its precursors as building blocks (eg. the isoprenoid pathways [11]) would potentially benefit from this technique.

The FadD mutants we generated in chapter four may prove directly useful for the production of MCFA derived biofuels, but may also guide further engineering of adenylate

forming enzymes. Studies on the crystal structures of FadD homologs generally focus on the fatty acid binding tunnel as the key determinant of substrate specificity of these enzymes [12,13], but our work shows that it is likely possible to modulate activity without affecting this tunnel and instead aid product exit by opening the enzyme. Considering other adenylate forming enzymes with substrates ranging from acetate [14] to coumaric acid [15] have similar structures [16], it is possible that the technique of aiding product exit by opening the enzyme could prove useful for engineering these enzymes as well.

Overall, these studies were successful at achieving their goals. Much work needs to be done before the systems and enzymes developed here can be used at commercial and/or industrial scale. This is true not only for the work presented here, but for many developments in synthetic biology that demonstrate successful engineering on the laboratory scale. Future efforts should be dedicated to proving the effectiveness of these systems in real-world settings. Only then can we really judge whether our attempts to engineer organisms with the tools and techniques of synthetic biology have been fruitful.

Acknowledgements

I would like to thank my Advisor, Pamela Silver, and her husband, Jeff Way, for supporting me throughout this work. I also thank all of my collaborators with special thanks to Joseph Torella who both trained me and was essential to the completion of the work in chapter two. Many thanks to all of my lab mates, especially Jake Wintermute, Abhishek Garg, and Jessica Polka who were all superb bay mates. I'd additionally like to thank my friends and family for their love and support throughout this work. Finally, I'd like to thank my Dissertation advisory committee members, Fred Winston, Jim Collins, and Roy Kishony as well as my exam committee members Deborah Hung, Neel Joshi, and Ahmad Khalil. This work was conducted with support from the Advanced Research Projects Agency-Energy 'Electrofuels' Collaborative Agreement DE-AR0000079, the National Science Foundation Graduate Research fellowship, and the Ruth L. Kirschtein National research Service Award program of Harvard Catalyst, The Harvard Clinical and Translational Science Center Award UL1 RR 025758 and financial contributions from Harvard University and its affiliated academic health care centers. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic health care centers, the National Center for Research Resources or the National Institutes of Health. This material is based upon work supported by the National Science Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

1. Akhtar MK, Dandapani H, Thiel K, Jones PR: **Microbial production of 1-octanol: A naturally excreted biofuel with diesel-like properties**. *Metabolic engineering communications* 2015, **2**:1-5.

2. Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD: **Enhancing fatty acid production by the expression of the regulatory transcription factor FadR**. *Metabolic engineering* 2012, **14**:653-660.

3. Zhang F, Carothers JM, Keasling JD: **Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids**. *Nature biotechnology* 2012, **30**:354-359.

4. Xu P, Li L, Zhang F, Stephanopoulos G, Koffas M: **Improving fatty acids production by engineering dynamic pathway regulation and metabolic control**. *Proceedings of the National Academy of Sciences of the United States of America* 2014, **111**:11299-11304.

5. Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R: **Engineered reversal of the beta-oxidation cycle for the synthesis of fuels and chemicals**. *Nature* 2011, **476**:355-359.

6. Moser F, Broers NJ, Hartmans S, Tamsir A, Kerkman R, Roubos JA, Bovenberg R, Voigt CA: **Genetic circuit performance under conditions relevant for industrial bioreactors**. *ACS synthetic biology* 2012, **1**:555-564.

7. Akhtar MK, Turner NJ, Jones PR: **Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**:87-92.

8. Kallio P, Pasztor A, Thiel K, Akhtar MK, Jones PR: **An engineered pathway for the biosynthesis of renewable propane**. *Nature communications* 2014, **5**:4731.

9. Davis R, Tao L, Tan E, Biddy M, Beckham G, Scarlata C, Jacobson J, Cafferty K, Ross J, Lukas J: **Process design and economics for the conversion of lignocellulosic biomass to hydrocarbons: Dilute-acid and enzymatic Deconstruction of biomass to sugars and biological conversion of Sugars to Hydrocarbons**. Edited by: National Renewable Energy Laboratory (NREL), Golden, CO.; 2013.

10. United States. Environmental Protection Agency.: **Fast Facts U.S. Tranportation Sector Greenhouse Gas Emissions 1990-2012**. 2015.

11. Immethun CM, Hoynes-O'Connor AG, Balassy A, Moon TS: **Microbial production of isoprenoids enabled by synthetic biology**. *Frontiers in microbiology* 2013, **4**:75.

12. Kochan G, Pilka ES, von Delft F, Oppermann U, Yue WW: **Structural snapshots for the conformation-dependent catalysis by human medium-chain acyl-coenzyme A synthetase ACSM2A**. *Journal of molecular biology* 2009, **388**:997-1008.

13. Hisanaga Y, Ago H, Nakagawa N, Hamada K, Ida K, Yamamoto M, Hori T, Arii Y, Sugahara M, Kuramitsu S: **Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer**. *Journal of Biological Chemistry* 2004, **279**:31717-31726.

14. Gulick AM, Starai VJ, Horswill AR, Homick KM, Escalante-Semerena JC: **The 1.75 √Ö crystal structure of acetyl-CoA synthetase bound to adenosine-5'-propylphosphate and coenzyme A**. *Biochemistry* 2003, **42**:2866-2873.

15. Hu Y, Gai Y, Yin L, Wang X, Feng C, Feng L, Li D, Jiang XN, Wang DC: **Crystal structures of a Populus tomentosa 4-coumarate:CoA ligase shed light on its enzymatic mechanisms**. *The Plant cell* 2010, **22**:3093-3104.

16. Schmelz S, Naismith JH: **Adenylate-forming enzymes**. *Current opinion in structural biology* 2009, **19**:666-671.