Production of anteiso-branched fatty acids in *Escherichia coli*; next generation
biofuels with improved cold-flow properties

Robert W. Haushalter^{1,2}, Woncheol Kim^{1,2}, Ted A. Chavkin³, Lionadi The³, Megan E.
Garber³, Melissa Nhan^{1,2}, Christopher J. Petzold^{1,2}, Leonard Katz^{3,4}, and Jay D.
Keasling^{1,2,3,4,5}.

¹Joint BioEnergy Institute, 5885 Hollis Street, 4th Floor, Emeryville, CA 94608;
²Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley,
CA, 94720; ²QB3 Institute, University of California-Berkeley, 5885 Hollis Street, 4th
Floor, Emeryville, CA 94608; ⁴Synthetic Biology Engineering Research Center,
University of California, Berkeley, CA 94720; ⁵Department of Chemical &
Biomolecular Engineering, Department of Bioengineering, University of California,
Berkeley, CA 94720.

Corresponding author: Jay D. Keasling, Joint BioEnergy Institute, 5885 Hollis Street,
4th Floor, Emeryville, CA 94608.

Email: jdkeasling@lbl.gov

Phone: 510-495-2620

Abstract

Microbial fermentation is emerging as an increasingly important resource for the production of fatty acids to serve as precursors for renewable diesel as well as detergents, lubricants and other industrial chemicals, as an alternative to traditional sources of reduced carbon such as petroleum. A major disadvantage of fuels derived from biological sources is their undesirable physical properties such as high cloud and pour points, and high viscosity. Here we report the development of an *Escherichia coli* strain that efficiently produces anteiso-branched fatty acids, which can be converted into downstream products with lower cloud and pour points than the mixtures of compounds produced via the native metabolism of the cell. This work addresses a serious limitation that must be overcome in order to produce renewable biodiesel and oleochemicals that perform as well as their petroleum-based counterparts.

Keywords: biofuels, synthetic biology

Highlights

- The branching pattern of fatty acids is controlled by the abundance of acyl-CoAs.
- Generated an *E. coli* strain that produced 20.4% anteiso-branched fatty acids.
- Dynamic regulation of *bFabH2* expression improved total fatty acid titer.

1. Introduction

Concerns about climate change and the increasing global demand for oil have motivated extensive efforts to find alternative, renewable sources of reduced carbon not only for liquid fuels, but also for detergents, lubricants, plastics, and other chemicals that have historically been derived from petroleum. Biodiesel, a mixture of fatty acyl methyl- and ethyl esters (FAMEs and FAEEs), is a promising alternative to petroleum-derived diesel, with many advantages including reduced net emissions of greenhouse gases, sulfur oxides and particulates¹, as well as the potential for domestic production from feedstocks that do not compete with food production for land or resources². Currently, most biodiesel is produced via esterification of fatty acids from agricultural sources, e.g., soybean oil¹. Such production requires the use of fertilizers and pesticides and decreases the availability of the crops for food. Microbial production of fatty acid precursors from cellulosic biomass, on the other hand, would minimize these ethical, environmental, and economic issues. As a consequence, *E. coli* has been engineered to produce and accumulate significant levels of free fatty acids³, as well as to enzymatically convert these fatty acids to more valuable compounds such as ethyl esters, fatty alcohols, wax esters⁴, methyl esters⁵ and alkanes^{6,7}.

Since global consumption of biodiesel has increased, specific standards for its physical properties have been established by the American Society for Testing and Materials (ASTM)⁸ and other government agencies around the world. These standards are outlined according to the following criteria: cetane number, heat of combustion, cloud and pour points, oxidative stability, viscosity and lubricity⁸.

While comparable to petroleum-derived diesel in cetane number and heat of combustion, and superior in lubricity, a glaring drawback of biodiesel is its poor performance at low temperatures due to its high cloud and pour points and high viscosity⁸. Whether sourced from microbes, plants, or animals, fatty acids consist primarily of straight-chain compounds which, when esterified with methanol or ethanol, yield FAMES and FAEEs, respectively, that freeze or gel at relatively high temperatures⁹ (Fig. 1a). Increasing the proportion of unsaturated fatty acids in these fuel mixtures lowers their cloud and pour points^{9,10}, but also increases the oxidative susceptibility of the fuel, which ultimately leads to poor performance¹¹. Fatty acids that contain methyl branches also exhibit lower freezing points, but do not suffer from the increased reactivity seen with unsaturated fatty acids⁹. The position of the methyl branch in the carbon chain affects the freezing point significantly; methyl branches further from the end of the hydrocarbon chain result in more significant depressions in freezing point⁹. For example, anteiso-branched fatty acids (in which the methyl branch is two carbon atoms from the ω end of the molecule), and their corresponding methyl esters, display freezing points considerably lower than their straight-chain and iso-branched counterparts⁹ (Fig.1a, b). Because *E. coli* is naturally incapable of producing branched fatty acids, we set out to rewire the metabolism of established *E. coli* strains to produce anteiso-branched fatty acids.

Bacillus subtilis and many other bacteria are able to initiate fatty acid synthesis with branched acyl-CoA starter units, which leads to the production of lipids containing methyl branches in the final product¹² (Fig.1b). The plasma membrane of *B. subtilis* contains a mixture of straight-chain saturated, unsaturated,

iso-, and anteiso-branched lipids, and the cells vary the proportion of branched fatty acids to modulate membrane fluidity. When exposed to cold temperatures, *B. subtilis* increases the proportion of low-freezing anteiso-branched fatty acids relative to iso-branched and straight-chain fatty acids in its membrane¹³. Branched acyl-CoA starter units are generated through the breakdown of branched chain amino acids (BCAAs) (Fig.1b). Enzymes encoded in the *bkd* operon catabolize BCAAs into branched acyl-CoAs, and two specialized ketosynthases, bFabH1 and bFabH2, initiate fatty acid biosynthesis by condensing these branched starter units with malonyl-ACP¹². Production of these *B. subtilis* enzymes in *E. coli* has been reported as a means of producing iso-branched fatty acids and alkanes^{6,12}. However, the potential for improving the properties of fatty acids from *E. coli* will not be fully realized without producing anteiso-branched lipids⁹ (Fig. 1b).

Figure 1. Biosynthesis of straight-chain and branched fatty acids and derivitization to methyl esters. a) Straight-chain fatty acids are primed with acetyl-CoA, extended, and derivatized into FAMES with high freezing points. b) Branched chain amino acids are converted to branched acyl-CoAs, which prime the synthesis of branched fatty acids. Anteiso-branched FAMES possess freezing points significantly lower than their straight-chain and iso-branched counterparts. Enzymes that are overproduced in our engineered strains are highlighted in blue.

In this report, we demonstrate that the branching pattern can be tailored by varying the supply of BCAAs in the medium and by genetic manipulation of amino acid biosynthesis. Using these strategies, we constructed a strain of *E. coli* in which 20.4% of the free fatty acids produced are anteiso-branched, albeit at a lower total fatty acid titer than our starting fatty acid production strain. Dynamic regulation of *bFabH2* expression resulted in total fatty acid titers nearly as high as the strain expressing just *tesA* and *fadR* with an anteiso proportion of 9.9%, highlighting an apparent trade-off between anteiso proportion and total fatty acid titer. This work provides a means of improving the properties of microbial biodiesel and is an important step toward the development of biofuels that can perform at least as well as fuels from petroleum sources.

2. Materials and Methods

2.1. Construction of P_{lacUV5}-Driven Expression Plasmids

The genes *tesA* and *fadR* were PCR amplified and integrated downstream of the lacUV5 promoter and kanamycin resistance gene, respectively, in biobrick vector pB5K to produce the plasmid pPhatty. pET15b constructs of *bfabH1* and *bfabH2* were generous gifts from the lab of Prof. Charles O. Rock¹². The entire *bkd* operon was PCR amplified from the genomic DNA of *B. subtilis*. The resulting 4750-bp fragment was cloned into the multiple cloning site of a pET28a vector to yield pWK-J9. Using the pET15b-bFabH2 and pWK-J9 constructs as PCR templates, the genes encoding bFabH2 and the Bkd enzymes were PCR amplified and assembled using circular polymerase extension cloning (CPEC)¹⁴ to yield the vector pPhatty ω 2, in

which *tesA*, *bfabH2*, and the four genes of the *bkd* operon are expressed under control of P_{lacUV5} as a single transcript. The plasmid pPhatty.bkd was assembled in a similar manner.

The plasmids pAFC3, pAFC46, and pCS49 were generous gifts from the lab of Prof. James C. Liao¹⁶. pAFC3 contains the genes *ilvGM* from *Salmonella typhimurium* in an operon with *E. coli ilvCD*, and pAFC46 contains the gene *ilvA* from *Corynebacterium glutamicum*. A PCR product containing the genes *ilvGMCD* and another containing *ilvA* and the backbone of pAFC46 were assembled using Gibson cloning¹⁵ to yield the plasmid pIlvGMCD.

2.2. Design and construction of synthetic *ilvY*-regulated promoters for Dynamically Regulated bFabH2 Production

The *ilvY* gene and the upstream promoter region were PCR amplified from *E. coli* K-12 genomic DNA. We replaced the native ribosomal binding site (RBS) upstream of *ilvC* with the strong RBS found in vector pET28. This PCR fragment was integrated into a GFP expression plasmid using Gibson assembly to produce the plasmid $P_{ilvC.rbs}\text{-GFP}$.

The *ilvY* gene and the modified divergent *ilvYC* promoter, as well as the gene *bfabH2*, were PCR amplified and assembled with the pBbA5C plasmid backbone using Gibson assembly to produce the plasmid pIlvC.rbs.bfabH2.

2.3. Construction of *E. coli* Strains

Strain JBEI-3111 (MG1655 Δ *fadE*)¹⁷ was transformed with pPhatty to yield the fatty acid production strain JBEI-7800. JBEI-3111 was transformed with pPhatty ω 2 to yield strain JBEI-7801. *E. coli* CRS24¹⁶, in which the genes *metA*, *tdh*, *ilvB* and *ilvN* have been removed, was a generous gift from the lab of Prof. James C. Liao (Table 2). The gene *fadE* was deleted from this strain using P1 phage transduction¹⁸, using strain JW5020 from the Keio collection as the donor strain¹⁹. The resulting strain, JBEI-7180, was transformed with pPhatty ω 2, pCS49, and pIlvGMCD A to yield the anteiso-branched fatty acid production strain JBEI-7802.

We replaced *thrL*, the native attenuator of the *thrABC* operon, with P_{Lac01} in strain JBEI-7180 using the lambda red system²⁰ to generate strain JBEI-7761, in which expression of the chromosomal *thrABC* operon is IPTG-inducible. This strain was transformed with pPhatty.bkd, pIlvGMCD A, and p.IlvC.rbs.bFabH2 to produce strain JBEI-7803. A negative control strain JBEI-7804, lacking *bfabH2*, was generated by transforming JBEI-7761 with pPhatty.bkd, pIlvGMCD A, and pBb.A5C.RFP.

2.4. Strain Growth

E. coli strains were grown from single colonies overnight in LB medium containing the appropriate antibiotics to maintain selection for the plasmid(s) present in the host. Overnight cultures were then diluted 1:100 in EZ-rich medium (with the appropriate antibiotics) for fatty acid production and analysis. These cultures were

grown at 37 °C until they reached an OD₆₀₀ of 0.8-1.0, at which point expression of enzymes was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). The cells were incubated with shaking at 30 °C for an additional 48 hours before they were harvested for fatty acid analysis.

2.5. Fatty Acid Analysis

Fatty acids were extracted and derivatized to methyl esters using the method described by Hamid, et al.²¹ After 48 hours of incubation after induction, 100 μL aliquots of cells were taken and extracted by the addition of 5 μL of 40% tetrabutylammonium hydroxide followed immediately by 100 μL of 200 mM methyl iodide in dichloromethane containing 100 mg/L methyl tricosanoate as an internal standard. The mixtures were shaken for 30 minutes using a vortex mixer, and then centrifuged to promote phase separation. The dichloromethane layer was removed, transferred into a GC vial with glass insert, and allowed to evaporate to dryness. The extracted methyl esters were then resuspended in pure dichloromethane prior to analysis by gas chromatography mass spectrometry (GC-MS).

The extracts were analyzed on an Agilent 6890 series GC equipped with a DB5.625 chiral column (30m x 0.250 mm) and an Agilent 5973 network mass selective detector. After splitless injection, the temperature was maintained at 40°C for 3 minutes, and then raised to 250°C at a rate of 20°C/min and maintained for 15 minutes. We identified the compounds of interest by comparison with authentic standards (12-methyltetradecanoic acid and 13-methyltetradecanoic acid were purchased from Sigma-Aldrich) and by comparison of the retention times and mass

spectra with the National Institute of Standards and Technology (NIST) database. Quantification of FAMES was performed by comparing the total ion chromatograph integrations of FAME peaks with that of the methyl tricosanoate internal standard, which was added into our samples at a concentration of 100 mg/L.

2.6. Targeted Proteomic Analysis

2.6.1. Sample Preparation for Targeted Proteomic Analysis

Cells for proteomic analysis were pelleted by centrifugation (8000g for 5 minutes), the supernatant was decanted and protein was extracted by chloroform/methanol (2/1 ratio) precipitation. The protein samples were re-suspended in 100 mM ammonium bicarbonate buffer and the concentration of the lysed samples was measured using the DC Protein Assay Kit (Bio-Rad, Hercules, CA) with bovine serum albumin used as a standard. A total of 100 µg protein from each sample was digested with trypsin for targeted proteomic analysis. Protein reduction was performed by the addition of 5 mM tris 2-(carboxyethyl)phosphine (TCEP) followed by incubation at room temperature (RT) for 30 minutes. Iodoacetamide (4 mM final concentration) was added and the samples were incubated for 30 minutes in the dark at RT, then trypsin was added (1:50 trypsin:total protein ratio) and the samples were digested overnight at 37 °C.

2.6.2. Liquid Chromatography-Mass Spectrometry

Peptides were analyzed using an Agilent 1260 liquid chromatography system coupled to an Agilent 6460QQQ mass spectrometer. The peptide samples (10 µg) were separated on an Ascentis Express Peptide ES-C18 column (2.7 µm particle size, 160 Å pore size, 5 cm

length x 2.1 mm i.d., coupled to a 5 mm x 2.1 mm i.d. guard column with similar particle and pore size; Sigma-Aldrich, St. Louis, MO), with the system operating in normal flow mode at 0.400 mL/min (column compartment = 60 °C). A 13-minute method with the following gradient was used: starting from 95% Buffer A (0.1% formic acid), 5% Buffer B (98% acetonitrile, 0.1 % formic acid) the Buffer B composition was increased to 35% over 5.70 minutes, followed by an increase to 90% B in 0.3 minutes, where it was held for 2 minutes. Buffer B was then decreased to 5% in 30 seconds where it was held for 4.5 minutes to re-equilibrate the column. The peptides were ionized by an Agilent Jet Stream ESI source operating in positive-ion mode with the following source parameters: Gas Temperature = 350°C, Gas Flow = 10 L/min, Nebulizer Pressure = 45 psi, Sheath Gas Temperature = 400°C, Sheath Gas Flow = 11 L/min, Nozzle Voltage = 1000 V, Capillary Voltage = 5000 V. The data were acquired using Agilent MassHunter, version B.06.00. Data were analyzed and quantified via Skyline version 2.5. Each SRM transition was designed *in silico* by Skyline (version 2.5) and manually reviewed prior to use for quantification. From the data set, individual peptide peak areas were summed for each protein and biological replicates were averaged and the error bars represent the standard deviation. Data were plotted using TIBCO Spotfire platform, a data visualization tool.

Table 1. Plasmids used in this study

Plasmids:	Relevant Features:	Reference:
pPhatty	BBR1 ori, Kan ^r , P _{LacUV5} - <i>tesA</i> (<i>E. coli</i>); <i>fadR</i> downstream of Kan ^r	This study
pPhatty ω 2	BBR1 ori, Kan ^r , P _{LacUV5} - <i>tesA</i> (<i>E. coli</i>), <i>bfabH2</i> , <i>ipdV</i> , <i>bkdAA</i> , <i>bkdAB</i> , <i>bkdB</i> (<i>B. subtilis</i>); <i>fadR</i> downstream of Kan ^r	This study
pCS49	SC101 ori, Spec ^r , P _{LacO1} - <i>thrABC</i> (<i>E. coli</i>)	Cann and Liao. ¹⁶
pIlvGMCD	colE1 ori, Amp ^r , P _{LacO1} - <i>ilvGM</i> (<i>S. typhimurium</i>), <i>ilvCD</i> (<i>E. coli</i>), <i>ilvA</i> (<i>C. glutamicum</i>)	This study
pPhatty.Bkd	BBR1 ori, Kan ^r , P _{LacUV5} - <i>tesA</i> (<i>E. coli</i>), <i>ipdV</i> , <i>bkdAA</i> , <i>bkdAB</i> , <i>bkdB</i> (<i>B. subtilis</i>); <i>fadR</i> downstream of Kan ^r	This Study
pIlvC.bfabH2	P15a ori, Cm ^r , P _{ilvC} - <i>bfabH2</i>	This Study
pIlvC.rbs.bfabH2	P15a ori, Cm ^r , P _{ilvC.rbs} - <i>bfabH2</i> , P _{ilvC.rbs} - <i>ilvY</i>	This Study
pBbA5C-RFP	P15a ori, Cm ^r , P _{LacUV5} -RFP	This study

Table 2. Strains used in this study

Strains:	Relevant Genotype:	Reference:
MG1655	F- lambda- <i>ilvG</i> - <i>rfb</i> -50 <i>rph</i> -1	Blattner, et al. ²²
JBEI-3111	MG1655 Δ <i>fadE</i>	Bokinsky, et al. ¹⁷
CRS24	BW25113 F' Δ <i>metA</i> Δ <i>tdh</i> Δ <i>ilvB</i> Δ <i>ilvI</i>	Cann and Liao. ¹⁶
JBEI-7180	CRS24 Δ <i>fadE</i>	This study
JBEI-7761	CRS24 Δ <i>fadE</i> , <i>thrL</i> ::P _{LacO1}	This Study
JBEI-7800	JBEI-3111 bearing plasmid pPhatty	This Study
JBEI-7801	JBEI-3111 bearing plasmid pPhatty ω 2	This Study
JBEI-7802	JBEI-7180 bearing plasmids pPhatty ω 2, pIlvGMCD, and pCS49	This Study
JBEI-7803	JBEI-7761 bearing plasmids pPhatty.Bkd, pIlvGMCD, and pIlvC.rbs.bFabH2	This Study
JBEI-7804	JBEI-7761 bearing plasmids pPhatty.Bkd, pIlvGMCD, and pBbA5C-RFP	This Study

3. Results and Discussion

3.1. Controlling fatty acid branching pattern with amino acid feeding

We began by determining how the availability of branched chain amino acids affects the proportion of branched fatty acids produced. Enzymes encoded in the *bkd* operon (*IpdV*, *BkdAA*, *BkdAB*, *BkdB*) convert the amino acids valine and leucine to isobutyryl- (2-methylpropyl-) and isovaleryl-CoA (3-methylbutyryl-CoA), respectively (Fig.1b). Fatty acid biosynthetic enzymes extend these precursors into iso-branched fatty acids. Isoleucine is catabolized by the *bkd* pathway to 2-methylbutyryl-CoA, which primes the biosynthesis of anteiso-branched fatty acids (Fig. 1b). The enzyme *FabH* catalyzes the first carbon-carbon bond formation in bacterial fatty acid synthesis, in which an acyl-CoA is condensed with malonyl-ACP. *B. subtilis* possesses two isoforms of *FabH* that are specialized in their ability to use branched acyl-CoAs as substrates¹². We hypothesized that we could selectively produce anteiso-branched fatty acids by expressing *bFabH2* and the *bkd* operon in *E. coli*, and by supplying the cells with exogenous isoleucine.

We therefore combined the necessary *B. subtilis* genes described above with modifications that allow for increased fatty acid accumulation in *E. coli*. Specifically, we overexpressed two *E. coli* genes to promote free fatty acid accumulation: *tesA*, a cytosolic ACP-thioesterase that cleaves acyl chains from acyl-ACP, thereby deregulating fatty acid synthesis and releasing free fatty acids into the cytosol^{3,4}, and *fadR*, a transcriptional regulator that controls the expression of several *E. coli* fatty acid synthase genes²³. When *TesA* and *FadR* are overproduced in a $\Delta fadE$ host, which renders the cell incapable of degrading fatty acids through β -oxidation, very

high titers of free fatty acids have been reported^{23,24}. We employed plasmid pPhatty, in which *tesA* expression is driven by a lacUV5 promoter, and *fadR* is expressed constitutively in an operon 3' of the kanamycin resistance gene (Table 1). Cells harboring pPhatty alone produced only straight-chain, even-carbon fatty acids (Fig. 2). We transformed JBEI-3111 with pPhatty ω 2, generating strain JBEI-7801, and cultured it in EZ-rich medium supplemented with different branched amino acids. Without addition of excess amino acids, JBEI-7801 produced a mixture of iso- and anteiso-branched fatty acids in proportions of 11.2% and 2.1%, respectively, of total fatty acids (Fig. 2b). The cells also produced 8.0% odd-chain fatty acids, the majority of which was pentadecanoic acid (Fig. 2). When 0.5 mM isoleucine was added to the growth medium, the proportion of anteiso-branched fatty acids markedly increased to 8.8%, the most abundant of which was the C15 anteiso-branched fatty acid 12-methyltetradecanoic acid (Fig.2), while the proportion of iso-branched fatty acids decreased to 7.1%. Conversely, feeding 0.5 mM leucine resulted in 14.0% iso-branched and 2.6% anteiso-branched fatty acids. Taken together, these data indicate that the branching pattern can be manipulated through the supply of branched chain amino acids. Although the proportion of anteiso-branched fatty acids was increased in these feeding experiments, the total yield of fatty acids decreased significantly in cells expressing *bfabH2*, in comparison to cells expressing only *tesA* and *fadR*. It should be noted that no significant amount of unsaturated branched fatty acids were detected, suggesting that one or both of the key enzymes in unsaturated fatty acid production (FabA and FabB) cannot accommodate branched substrates.

Figure 2. Feeding branched chain amino acids to cells expressing *bfabH2* and the *bkd* operon changes the fatty acid distribution. a) Strain JBEI-7800, overexpressing only *tesA* and *fadR*, produces only straight-chain fatty acids with an even number of carbons at a titer of 1922 mg/L. b) Strain JBEI-7801, which additionally expresses *bfabH2* and *bkd* genes, produces a diverse mixture of fatty acids including iso- and anteiso-branched fatty acids at a titer of 1549 mg/L. c) Feeding isoleucine to JBEI-7801 results in an increase in anteiso-branched fatty acids. d) Conversely, feeding feeding leucine results in an increase in iso-branched fatty acids. These data indicate that the fatty acid branching pattern is controlled through the availability of acyl-CoA precursors. The error bars indicate the standard deviations from triplicate experiments.

3.2 Controlling fatty acid branching pattern through manipulation of amino acid biosynthesis

As amino acids are relatively expensive compounds, feeding isoleucine is not an economically viable solution for steering the cell's metabolism specifically towards desirable anteiso-branched fatty acids. We therefore set out to construct an *E. coli* strain that would produce anteiso-branched precursors directly from glucose. From the amino acid feeding experiments, we surmised that supply of acyl-CoA precursors is the primary determinant of fatty acid branching pattern, and that cells engineered to specifically produce the necessary 2-methylbutyryl-CoA precursor directly from glucose should accumulate anteiso-branched fatty acids without the need to supply isoleucine in the medium.

E. coli strains producing high titers of isoleucine, as well as 2-methyl-1-butanol, a potential gasoline replacement derived from isoleucine precursors have been reported^{25,16}. In the latter study, overexpression of heterologous genes involved in the biosynthesis of isoleucine and *E. coli* genes that catalyze formation of its metabolic precursor threonine, coupled with several gene deletions that prevent metabolic flux towards other amino acids, resulted in the identification of a strain that selectively produced significant quantities of 2-methyl-1-butanol¹⁶. We adapted this system by first deleting *fadE* from *E. coli* CRS24, a strain in which the genes *metA*, *tdh*, *ilvB* and *ilvN* had been removed¹⁶, yielding strain JBEI-7180 (Table 2). We then consolidated the *S. typhimurium* genes *ilvGMCD* and the *C. glutamicum* gene *ilvA* onto a single plasmid, designated pIlvGMCD_A, and transformed JBEI-7180 with this

plasmid along with pPhatty ω 2 and pCS49¹⁶, the latter of which encodes the enzymes ThrABC. When the resulting strain, JBEI-7802, was analyzed for its fatty acid profile, we found that anteiso-branched fatty acids comprised 20.4% of the free fatty acids in the cell (Fig.3). JBEI-7802 also produced significant amounts of odd-carbon fatty acids, which were not produced in significant amounts in the absence of the bFabH and Bkd enzymes. We attribute this production to the IlvA-catalyzed breakdown of threonine to 2-ketobutyrate, which is subsequently converted to propionate or propionyl-CoA by the native pyruvate dehydrogenase complex, the Bkd enzymes, or PoxB. Propionyl-CoA can serve as a substrate for both *E. coli* and *B. subtilis* FabH enzymes to prime odd-carbon fatty acid biosynthesis¹² (Fig.3). While we observed a large proportion of desired anteiso-branched fatty acids in these experiments, the titer of total fatty acids was still much less than that observed when expressing *tesA* and *fadR* alone.

Figure 3. Engineering *E. coli* to overproduce 2-methylbutyryl-CoA directly from glucose. Carbon flux is directed towards production of threonine by overproducing the enzymes ThrABC, and removal of the enzymes MetA and Tdh. Threonine is converted to 2-ketobutyrate through the action of IlvA. Pyruvate and 2-ketobutyrate are condensed by IlvGM and converted to 2-keto-3-methylvalerate by IlvCD. The enzymes encoded in the *bkd* operon then convert 2-keto-3-methylvalerate to 2-methylbutyryl-CoA. bFabH2 then uses this metabolite to prime anteiso-branched fatty acid synthesis. Overproduced enzymes are colored in blue; enzymes whose genes have been deleted from the chromosome are shown in red.

3.3. Dynamic regulation of bFabH2 expression to improve total fatty acid titer

3.3.1. Construction of IlvY-regulated promoters for dynamic regulation of *bfabH2*

In all cases where we drove *bfabH2* expression with P_{LacUV5}, we observed a lower total fatty acid titer compared to strains lacking *bfabH2*, indicating that high-level expression of this gene is detrimental to fatty acid synthesis. FabH catalyzes only the first of 5-8 chain extensions in fatty acid synthesis; it is possible, therefore, that if the concentration of FabH is too high, it could compete with other FAS enzymes to bind ACP, and thus inhibit the overall process. Indeed, when the *E. coli* fatty acid synthase was reconstituted *in vitro*, the rate of fatty acid synthesis decreased when FabH was added at concentrations higher than the other enzymes in the pathway²⁶. We hypothesized that if we attenuate *bfabH2* expression such that it is only expressed when its substrate is present, this would alleviate the inhibition of fatty acid synthesis seen at high expression levels.

There is no known transcriptional regulator that responds to 2-methylbutyryl-CoA. However, in *E. coli*, a transcriptional regulator that responds to acetohydroxy acids, IlvY, controls the expression of the *ilvC* gene²⁷. These compounds are intermediates in the synthesis of isoleucine and valine; they are produced by IlvGM and serve as substrates for IlvC²⁸. We hypothesized that if we could control *bfabH2* expression with a promoter regulated by IlvY, then the bFabH2 would only be produced when there is carbon flux flowing through the ilvGMCD pathway towards 2-methylbutyryl-CoA.

IlvC and *ilvY* are natively expressed divergently from the same promoter, in which acetohydroxy acid binding to IlvY results in increased expression of *ilvC* and decreased expression of *ilvY*²⁸. We cloned the *ilvY* gene and the divergent promoter that drives *ilvY/ilvC* expression and placed them on plasmid pBbB5K GFP such that the promoter divergently expresses *ilvY* and *gfp*. We found that GFP was not expressed in the presence of IPTG in a strain carrying this plasmid alone, but was highly expressed when the genes *thrABC* and *ilvGMCD*A were simultaneously expressed (Fig. S1b). Additionally, we found that GFP fluorescence increases steadily after IPTG induction for several hours when expressed from the IlvY-regulated promoter, possibly because of the gradual accumulation of the acetohydroxy acids upon induction of *thrABC* and *ilvGMCD*A. After confirming that the IlvY-regulated promoter only expresses genes under its control when *thrABC* and *IlvGMCD*A are co-expressed, we cloned *bfabH2* 3' of this promoter.

3.3.2. Dynamic regulation of *bfabH2* to recover total fatty acid titers

In order to isolate *bFabH2* on its own plasmid to optimize its regulation, we consolidated all the other heterologous genes onto two plasmids to maintain a three-plasmid strain. We cloned a new vector pPhatty. Bkd, in which the *bfabH2* gene is removed from pPhatty ω 2, and only *tesA* and the four genes of *bkd* operon are expressed with P_{LacUV5} (Table 1). As described earlier, the genes *thrABC* must be overexpressed to supply threonine, the substrate of IlvA, for conversion to 2-ketobutyrate and ultimately to 2-methylbutyryl-CoA (Figure 3). These genes are native to *E. coli*, so we chose to overexpress the genomic copy of these genes. We

therefore replaced the native attenuator *thrL* with P_{lacO1} using the lambda red system²⁰. In the resulting strain, JBEI-7761, the expression of the *thrABC* operon is IPTG-inducible, while the genes *fadE*, *ilvB*, *ilvI* *metA*, and *tdh* have been deleted. JBEI-7761 was transformed with pPhatty.bkd, pIlvGMCD, and pIlvC.rbs.bFabH to yield strain JBEI-7803. In this anteiso-branched fatty acid production strain, the expression of all the genes in the pathway to 2-methylbutyryl-CoA are IPTG-inducible, and the expression of bFabH2 is driven by the *ilvY*-regulated promoter. As a control to compare with a strain lacking bFabH2, we additionally generated strain JBEI-7804, which is strain JBEI-7761 harboring pPhatty.bkd, pIlvGMCD, and pBbA5C-RFP.

We compared anteiso fatty acid production in these strains to our previous strains expressing *bfabH2* with P_{lacUV5} (Figure 4). As expected, strain JBEI-7804, expressing RFP in place of *bfabH2*, produced only straight-chain fatty acids at a titer of 1634 mg/L, a value which approaches that seen in strain JBEI-7800 expressing just *tesA* and *fadR*. Roughly 3% of the fatty acids produced by this strain had an odd number of carbons; this was expected since the overexpression of *thrABC* and *ilvA* will still lead to propionyl-CoA production by way of 2-ketobutyrate. Strain JBEI-7803, with dynamically regulated bFabH2, produced 1532 mg/L fatty acids, 9.9% of which were anteiso-branched. While we didn't reach the titer of 1922 mg/L seen in strain JBEI-7800, dynamic regulation of bFabH2 production resulted in a significantly higher titer of fatty acids compared to strain JBEI-7802, which had *bfabH2* under the control of P_{lacUV5} , but with a lower proportion of anteiso branched fatty acids. To determine the relative abundance of bFabH2 in strains JBEI-7802 and

JBEI-7803, we performed a quantitative targeted proteomic analysis of the two strains. We found that bFabH2 is much more abundant in JBEI-7802 than in JBEI-7803, while the relative amounts of several control enzymes were similar in the two strains (Figure S2). These results support our hypothesis that high production of bFabH2 results in increased branched fatty acid proportion along with decreased total fatty acid titer.

Figure 4. Total fatty acid production and distributions of engineered strains. The average total titer is indicated in black, and the proportion of anteiso-branched fatty acids is shown in red for each strain. Error bars show the standard deviation from triplicate samples.

4. Conclusions

We have demonstrated here methods for manipulating the fatty acid profile of engineered *E. coli* and leveraging these methods to produce anteiso-branched fatty acids that will reduce the freezing point and increase the fluidity of lipids extracted from the cells. We found that high production of a key enzyme, bFabH2, results in higher proportions of branched fatty acids at the cost of total fatty acid titer. While our dynamic regulation strategy illustrated this trade-off, there are several possible avenues to improve the system further. We chose to overexpress the gene *fadR* because it has been shown to increase the expression of several fatty acid synthase genes, including *fabB* and *fabF*, which results in improved fatty acid titers²⁹. However, it has also been shown that high *fadR* expression also results in higher expression of the *fabH* operon, thus increasing production of the endogenous FabH, which could compete with bFabH2 and result in a lower proportion of branched fatty acids³⁰. Therefore, further optimization of *fadR* expression, or modifying the native promoters controlling fatty acid synthase expression, could help to produce more ideal total fatty acid titers and branched distributions. Because overexpressing *fadR* leads to increased expression of *fabB* and *fabA*, we see a high proportion of unsaturated fatty acids in our strains. Unsaturated fatty acids decrease the fluidity of the membrane, an essential physiological role that can also be performed by branched fatty acids. Introduction of the branched fatty acid pathway could obviate the need for unsaturated fatty acids; removal of the unsaturated pathway would allow for a higher proportion of branched species and

lower the oxidative susceptibility of the resulting fatty acid mixture. Experiments to investigate this possibility are ongoing.

As noted earlier, we observe a significant amount of odd-chain fatty acids, resulting from extension of propionyl-CoA by fatty acid synthase. It is likely that the enzyme PoxB is decarboxylating 2-ketobutyrate to form propionate, thus diverting metabolic flux away from branched fatty acids towards odd-chain fatty acids (Figure 3). Deletion of *poxB* could therefore divert flux back towards branched acyl-CoA precursors and improve branched fatty acid proportions. This is currently being investigated. Finally, our targeted proteomic experiments indicated incomplete lipoylation of BkdB, a covalent modification necessary for activity; supplementation with lipoic acid could help increase the pool of branched acyl-CoAs if the activity of BkdB is limiting.

This work addresses a major shortcoming of fuels derived from biological sources. Currently, biodiesel must be blended with low-freezing additives and/or petroleum-derived diesel in order to perform properly in cold climates⁸. Petroleum diesel owes its superior cold-flow performance to the complexity of its composition; it is a diverse mixture of hydrocarbons with a wide variety of chain lengths and branching patterns. The strategies outlined here greatly increase the diversity of fatty acids produced in *E. coli*, and this increased diversity will translate into improved physical properties in fuel mixtures derived from these fatty acids. Lower freezing points and higher fluidity are desirable not only in liquid fuels, but also in detergents, cosmetics, lubricants and other commodities that contain lipids. If biodiesel from microbes is to achieve its potential, utilizing the strategies and techniques described here will

decrease the need for blending with petroleum-derived diesel and expensive additives that increase fluidity in order to overcome performance limitations in low temperature environments.

Acknowledgements

We thank Prof. Charles O. Rock for generously providing bFabH constructs and Prof. James C. Liao for providing strain CRS24 and plasmids pCS49, pAFC3, and pAFC46. We acknowledge Dr. Eric Steen for producing strain JBEI-3111, Dr. Dan Groff for generating plasmid pPhatty, and Vivian Macnguyen, Meghana Rao, Michael Bonilla, and Andrew Brandon for technical assistance. This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

References

- (1) Hill J, Nelson E, Tilman D, Polasky S, Tiffany D (2006) Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proc. Natl. Acad. Sci.* 103(30):11206-11210
- (2) Blanch HW, Adams PD, Andrews-Cramer KM, Frommer WB, Simmons BA, Keasling JD (2008) Addressing the Need for Alternative Transportation Fuels: The Joint BioEnergy Institute. *ACS. Chem. Biol.* 3(1):17-20
- (3) Lu X, Vora H, Khosla C (2008) Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. *Metab. Eng.* 10(6):333-339
- (4) Steen E J, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, del Cardayre SB, Keasling JD (2010) Microbial production of fatty acid derived fuels and chemicals from plant biomass. *Nature.* 463:559-563
- (5) Nawabi P, Bauer S, Kyrpides N, Lykidis A, (2011) Engineering Escherichia Coli for biodiesel production utilizing bacterial fatty acid methyltransferase. *Appl. Environ Microbiol.* 77(22):8052-8061
- (6) Howard TP, Middlehauf S, Moore K, Edner C, Kolak DM, Taylor GN, Parker DA, Lee R, Smirnoff N, Aves SJ, Love J (2013) Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in Escherichia coli. *Proc Natl. Acad. Sci.* 110(19): 7636-7641
- (7) Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB (2010) Microbial biosynthesis of alkanes. *Science.* 329(5991): 559-562
- (8) Knothe G (2005) Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. *Fuel Process. Technol.* 86:1059-1070

- (9) Knothe G, Dunn RO (2009) A comprehensive evaluation of melting points of fatty acids and esters determined by differential scanning calorimetry. *J. Am. Oil Chem. Soc.* 86(9):843-856
- (10) Cao Y, Yang J, Xian M, Xu X, Liu W (2010) Increasing unsaturated fatty acid contents in *Escherichia coli* by coexpression of three different genes. *Appl. Microbiol. Biotechnol.* 87:271-280
- (11) Knothe G (2009) Improving fuel properties by modifying fatty ester composition. *Energy Environ. Sci.* 2(7):1039
- (12) Choi KH, Heath RJ, Rock CO (2000) β -ketoacyl-acyl carrier protein synthase III (fabH) is a determining factor in branched-chain fatty acid biosynthesis. *J. Bacteriol.* 182(2):365-370
- (13) Klein W, Weber MH, Marahiel MA (1999) Cold shock response of *Bacillus subtilis*: isoleucine-dependent switch in the fatty acid branching pattern for membrane adaptation to low temperatures. *J. Bacteriol.* 181(17): 5341-5349
- (14) Quan J, Tian J (2011) Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries. *Nat. Protoc.* 6(2): 242-251
- (15) Gibson DG, Young L, Chuang RY, Venter JC, Hutchinson CA 3rd, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods.* 6(5): 343-345
- (16) Cann AF, Liao JC (2008) Production of 2-methyl-1-butanol in engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 81(1):89-98
- (17) Bokinsky G, Peralta-Yahya PP, George A, Holmes BM, Steen EJ, Dietrich J, Lee TS, Tullman-Ercek D, Voigt CA, Simmons BA, Keasling JD. (2011) Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. *Proc Natl Acad Sci.* 108(50):19949-54.
- (18) Thomason LC, Costantino N, Court DL (2007) *E.coli* genome manipulation by P1 Transduction. *Curr. Protoc. Mol. Biol.* Chapter1: Unit 1.17.
- (19) Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* 2:2006.0008.
- (20) Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Nat. Acad. Sci.* 97(12):6640-6645.
- (21) Hamid ME, Minnikin DE, Goodfellow M (1993) A simple chemical test to distinguish mycobacteria from other mycolic-acid-containing actinomycetes. *J. Gen. Microbiol.* 139(9): 2203-2213
- (22) Blattner FR, Plunkett G 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. *Science.* 277(5331):1453-62
- (23) Zhang F, Carothers JM, Keasling JD (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nat. Biotechnol.* 30(4):354-35
- (24) Lennen RM, Pfleger BF (2012) Engineering free fatty acid production in *Escherichia coli*. *Trends Biotechnol.* 30(12):659-667

- (25) Hashiguchi K, Takesada H, Suzuki E, Matsui H (1999) Construction of an L-isoleucine overproducing strain of *Escherichia coli* K-12. *Biosci. Biotechnol. Biochem.* 63(4):672-679
- (26) Yu X, Liu T, Zhu F, Khosla C (2011) In vitro reconstitution and steady-state analysis of the fatty acid synthase from *Escherichia coli*. *Proc Natl. Acad. Sci.* 108(46): 18643-18648.
- (27) Rhee KY, Senear DF, Hatfield GW (1998) Activation of gene expression by a ligand-induced conformational change of a protein DNA complex. *J. Biol. Chem.* 273: 11257-11266.
- (28) Rhee KY, Opel M, Ito E, Hung SP, Arfin SM, Hatfield GW (1999) Transcriptional coupling between the divergent promoters of a prototypic LysR-type regulatory system, the *ilvYC* operon of *Escherichia coli*. *Proc Natl. Acad. Sci.* 96(25): 14294-14299
- (29) Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD (2012) Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metab. Eng.* 14(6):653-660
- (30) My L, Rekoske B, Lemke JJ, Viala JP, Gourse RL, Bouveret E (2013) Transcription of the *Escherichia coli* fatty acid synthesis operon *fabHDG* is directly activated by FadR and inhibited by ppGpp. *J. Bacteriol.* 195(16):3784-3795.