

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Constructing and engineering fatty acid metabolic pathways for the production of fuels and chemicals

Permalink

<https://escholarship.org/uc/item/5qj6t4m3>

Author

Steen, Eric James

Publication Date

2010

Peer reviewed|Thesis/dissertation

**Constructing and engineering fatty acid metabolic pathways for the
production of fuels and chemicals**

By

Eric James Steen

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Joint Doctor of Philosophy

with University of California, San Francisco

in

Bioengineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jay D. Keasling, Chair

Professor Adam P. Arkin

Professor Michelle C. Chang

Professor Hiroshi Nikaido

Professor Christopher A. Voigt

Fall 2010

©2010

Eric James Steen

All rights reserved

Table of Contents

List of Figures

List of Tables

Acknowledgements

Chapter 1: Microbial production of biofuels and chemicals

1.1. Motivation

1.2. Organization

1.3. References

Chapter 2: Engineering *Saccharomyces cerevisiae* for the production of n-butanol.

2.1. Introduction

2.2. Materials and Methods

2.3. Results and Discussion

2.3.1. Expression of n-butanol pathway isozymes from a range of organisms in *S. cerevisiae* results in n-butanol production

2.3.2. Analysis of intermediary metabolites

**2.3.3. Engineering *adhe2* for solubility; increase activity to butyryl-CoA
Metabolic byproducts of n-butanol producing yeast strains**

2.3.4. Removal of competing pathways

2.3.5. Mass Spectrometric Analysis of protein levels

2.3.6. Acetyl-CoA compartmentalization

2.4. Discussion and future directions

2.5 References

Chapter 3: Engineering *Escherichia coli* for the production of fatty acid derived chemicals and fuels.

3.1. Introduction

3.2. Materials and Methods

3.3. Results

3.3.1. Cytosolic expression of the native *E. coli* thioesterase, ‘*tesA*

3.3.2. Coexpression of alternative thioesterases and fatty acid modifying enzymes.

3.3.3. Coexpression of ethanol and fatty acid pathways.

3.3.4. Coexpression of fatty acid and long chain alcohol pathways.

3.3.5. Coexpression of FAEE production and hemicelluloses degradation pathways.

3.3.6. Cloning and overexpressing *fadD*

3.4. Discussion and future directions

3.5. References

Chapter 4. Optimization of fatty acid and fatty acid ethyl ester production in *Escherichia coli*.

4.1. Introduction

4.2. Materials and Methods

4.3. Results

4.3.1. 2-Liter fermentations studies of strain A2A

4.3.2. Byproduct formation in test tubes

4.3.3. Plasmid stability of strain A2A

4.3.4. Increasing antibiotic concentration for increasing plasmid maintenance

4.3.5. Optimization of induction and growth parameters of strain A2A.

4.3.6. Single plasmid productivity of FAEEs

4.3.7. Integration of thioesterase, ‘*tesA*

4.4. Discussion and future directions

4.5. References

Chapter 5. Engineering yeasts for production of fatty-acid-derived fuels and chemicals.

5.1. Introduction

5.2. Materials and methods

5.3. Preliminary Results

5.3.1. Overexpressing FAS3 and FAS1/2

5.3.2. Removal of neutral lipid storage

5.3.3 Removal of competing pathways (b-oxidation)

5.3.4. Removal of acyl-CoA binding protein ACB1

5.3.5. Expressing the wax ester synthase, *atfA*

5.3.6. Expressing the ATP:citrate lyase and malic enzyme

5.4. Discussion and future directions

5.5. References

Chapter 6. Diacid production from fatty acid biosynthesis.

6.1. Introduction

6.2. Materials and methods

6.3. Results

6.3.1 Production of C14 acid, hydroxyacid and diacids

6.4. Discussion

6.5. References

List of figures:

Figure 2-1: Clostridial ABE fermentation pathway.

Figure 2-2: The n-butanol biosynthetic pathway.

Figure 2-1: Representative plasmids used in this study.

Figure 2-2: n-Butanol production from engineered *S. cerevisiae*

Figure 2-3: n-Butanol pathway intermediates at 24 h

Figure 2-4: Protein analysis of strain ESY7 expressing Adhe2

Figure 2-5: Metabolic byproduct formation in strain ESY7.

Figure 2-6: Butanol evolution from ADH knockouts

Figure 2-7: Butanol evolution from ADH knockouts

Figure 2-8: Butanol production from ADH knockout strains.

Figure 2-9: Amorphadiene production from ACL-expressing strains

Figure 3-1: Engineered pathways for production of fatty acid-derived molecules from hemicelluloses or glucose and depiction of the synthetic operons used in this study

Figure 3-2: Total free fatty acid production and respective theoretical yield by engineered *E. coli* strains

Figure 3-3: Fatty acid chain length distribution in *E. coli* DH1 *fadD* knockout expressing *tesA*.

Figure 3-4: Engineered production of FAEEs and fatty alcohols with controlled chain length

Figure 3-5: Toward a single cell catalyst: Biodiesel (FAEE) production by various strains without exogenous ethanol supplementation.

Figure 3-6: Growth measurements from *E. coli* DH1 *fadE* knockout strain in various concentrations of methyl oleate (biodiesel).

Figure 3-7: Distribution of fatty esters in *fadE* knockout expressing '*tesA*, *fadD*, and *atfA*.

Figure 3-8: Consolidated bioprocessing: Growth and FAEE production by xylan-utilizing strains.

Figure 4-1: Free fatty acid and FAEE production from strain A2A.

Figure 4-2: Yield of FAEEs from strain A2A.

Figure 4-3: Production of fatty acid ethyl esters (FAEEs) and fatty acids (FAMES) under low IPTG induction conditions with strain A2A.

Figure 4-4: Glucose utilization by strain A2A under low IPTG.

Figure 4-5: Acetate formation by strain A2A under low IPTG.

Figure 4-6: Ethanol formation by strain A2A under low IPTG

Figure 4-7: Plasmid retention in strain A2A under low IPTG.

Figure 4-8: Colony forming units for strain A2A under low IPTG.

Figure 4-9: FAME and FAEE production in strain A2A with varying antibiotic doses.

Figure 4-10: FAME and FAEE production in strain A2A with varying IPTG concentrations.

Figure 4-11: FAME and FAEE production in strain A2A with varying IPTG concentrations.

Figure 4-12: FAME and FAEE production in strain A2A with varying ITPG and OD .

Figure 4-13: FAME and FAEE production in single plamsid strains.

Figure 4-14: FAME and FAEE production in single plamid strains with varying OD and IPTG.

Figure 4-15: Metabolite levels in single plasmid strains with varying IPTG and OD.

Figure 4-16: FAME production in L'*tesA* integrated strains.

Figure 5-1: Transcriptional regulation elements upstream of the FASs

Figure 5-2: FAS3 reaction scheme, including the transfer of CO₂ to acetyl-CoA in forming malonyl-CoA

Figure 5-3: Fatty acid elongation cycle, initiated with an acyl-ACP which grows by one unit of malonyl-ACP.

Figure 5-4: Proposed expression construct for FASs, under the regulation of the repressible GAL promoter, where all FAS1, 2, & 3 are represented by fasX.

Figure 5-5: The 3 FASs are cloned into three pESC vectors with different auxotrophic markers (leu, his, & ura).

Figure 5-6: Fatty acid production from strains overexpressing FAS1, 2, & 3.

Figure 5-7: Increased acetyl-CoA levels and final production of amorphadiene upon expressing an ATP citrate lyase

Figure 6-1: Diacid production in strains expressing a p450 and thioesterase.

List of tables

Table 2-1: Characteristics of short-chain alcohols

Table 2-2: Strains used in this study

Table 2-3: Predicted solubilities of heterologous butanol enzymes.

Table 2-4: Metabolite formation in strain ESY7

Table 2-5: Protein levels of various strains in n-butanol producing cultures.

Table 3-1: Strains and plasmids.

Table 3-1: Sequencing results and mutations in *fadD*.

Table 4-1: Plasmids and strains used in this study.

Table 4-2: Media recipe for fed-batch fermentation of strain A2A

Table 4-3: Fermentation characteristics of strain A2A.

Table 4-4: A2A operons and antibiotic concentrations

Table 4-5: Induction and OD parameters.

Acknowledgements

This work would not have been possible without my advisor, Jay D. Keasling, who has been instrumental to my development as a graduate student, synthetic biologist and metabolic engineer. Jay's energy seems endless - his management of a large lab in addition to his many other duties go on seamlessly to us, but undoubtedly take immense effort on his behalf. Not only does Jay have an amazing capacity for engaging his individual students scientifically, he's also always ready for a hearty laugh that's sure to brighten everyone's spirit. Jay's knack for identifying and helping to solve big picture problems that we can address with the tools of synthetic biology have positioned him at the forefront of the field and I am grateful to have worked with him.

Further the collaborative nature within the Keasling lab has promoted many successful endeavors that must be acknowledged. Greg Bokinsky has been responsible for engineering *E. coli* for catabolism of plant biomass; Mario Ouellet and Heather Szmidt have been fundamental to optimizing and further understanding of the FAEE pathway; Helcio Burd is responsible for the fermentation studies of A2A; Edward Baidoo performed the diacid metabolite analysis experiments; Rossana Chan and Nilu Prasad aided in construction of the n-butanol pathway. There are numerous others who have contributed and provided advice. For this I am thankful.

Abstract

Engineering metabolic pathways for the production of fuels and chemicals

By

Eric James Steen

Doctor of Philosophy in Bioengineering

University of California, Berkeley

Professor Jay D. Keasling, Chair

Fatty acid biosynthesis creates highly-reduced hydrocarbons dubbed nature's petroleum. The pathway is responsible for products mostly used to build the lipid-bilayer of cell membranes, but also are implicated in cell signaling, brain function, and energy storage. As a source of natural hydrocarbons, we sought to exploit fatty acid biosynthesis for microbial production of a number of molecules including butanol, higher chain fatty alcohols, fatty acid ethyl esters (biodiesel), wax esters and long chain diacids. We attempted to exploit the two main forms of fatty acid biosynthesis, type I and II, which are characterized by single polypeptides that encode multiple reactions or single polypeptides that encode single reactions, respectively. The different types of fatty acid biosynthesis are found in comparing prokaryotes and eukaryotes, thus we engineered both *Escherichia coli* and *Saccharomyces cerevisiae*. Our strategy for building and engineering a pathway is broadly applicable and consists of identifying requisite reactions and the enzymes that encode those reactions, expressing those enzymes in a single host, analyzing production levels of metabolites related to the pathway, generating a hypothesis that explains the system's behavior and re-perturbing it to test the hypothesis in an effort to more efficiently produce a molecule of interest. With this strategy we successfully construct and engineer fatty acid pathways in *E. coli* and *S. cerevisiae*.

Chapter 1: Microbial production of biofuels and chemicals

1.1. Motivation

Petroleum is in finite supply and its exploitation harms the environment.

Petroleum, literally rock oil, was first reported to be exploited 4000 years ago in the construction of Babylon, isolated from the Issus River, a tributary of the Euphrates. Additional accounts of its use for lighting and medicinal purposes exist in Ancient Persian tablets¹. Later, in the 1600-1800s, whale oil dominated the supply of hydrocarbon dense fuels and was used for lighting the Western world, but soon subsided due to other sources like the “prairie whale” (hogs)² and finally coal and kerosene. In 1853, the first oil well was drilled in Poland followed by further exploration in the Russian empire, the US and Saudia Arabia in the 20th century. Founded upon the abundant supply of black gold, a number of industries arose including the pharmaceutical, industrial chemicals, etc. These industries and human existence have become heavily dependent upon oil and its supply. In 1956 a geophysicist, M. King Hubbert, proposed a theory that simply states, because oil is finite, its rate of extraction or production from a given area will follow the derivative of a logistic curve, increasing to a maximum production rate and then decreasing until reserves are exhausted. Although the theory is a simple one, its implications and effects can be seen in the oil crises, wars, and steady increasing dependence upon foreign sources of oil and the predictions that have been made today indicating the world is nearing its peak production levels. Nearing peak production and an ever-increasing world population and demand for oil motivates development of a renewable, alternative sources.

The idea to obtain a renewable source of energy is not new and was spurred by the inability to obtain oil or raw energy in World War I and II. Later, in the 1970s, the oil crisis began, marked by substantial oil shortages for the major industrial nations, including the United States due to peak production output that was reached by Germany, the US, Iran, and Venezuela, which in turn forced nations to become more reliant upon foreign oil supplies. Acknowledgement and in some cases, the stark realization that oil is a finite resource has promoted the search for and development of alternative energy and petroleum sources. Another motivation stems from the many negative environmental effects, including release of otherwise sequestered carbon dioxide and sulfur and nitrogen-containing compounds that contribute to increasing levels of green house gases due to the utilization of petroleum-based fuels. Further, the mere exploration and extraction methods used to obtain oil harm the environment. In April of 2010, BP’s attempt to find oil thousands of feet below the ocean surface in the Gulf of Mexico has resulted in the biggest oil spill the US has faced. Although there are many

approaches being pursued, one promising approach is to convert simple sugars and biomass (plants or other organisms that derive its energy from sunlight) into biofuels and chemicals by metabolic engineering of microorganisms like *Saccharomyces cerevisiae* and *Escherichia coli*. Outside of the scope of this study, yet still important are considerations of where the biomass or microbial feedstock is derived from, because competition for arable land will ultimately compete with food supplies³.

The two commonly used biofuels are ethanol and biodiesel. In the US, 4.86 billion gallons of ethanol and 250 million gallons of biodiesel were produced in 2006^{4,5}. In comparison, the US consumed 7.6 billion barrels of petroleum in 2004⁶.

While these alternatives to petroleum-derived fuel may provide energy sustainability and decreased green house gas emissions, there are drawbacks. Ethanol in particular will be difficult to integrate into the existing energy infrastructure on a large scale because it is hygroscopic which results in increased potential for rusting and precludes its delivery by pipeline. Additionally, ethanol requires a significant amount of energy to be processed, especially in its distillation step, which accounts for over 50% of the total energy consumed (7.45 million kcal per 1000L EtOH)⁷ during its production⁸. Finally, ethanol has a low energy content compared to gasoline and therefore decreases the available miles traveled per volume of liquid. Current production standards for biodiesel are also non-optimal for similar reasons to ethanol. Biodiesel requires a significant amount of energy for growing oil crops like soybeans and converting the oil into fatty acid methyl esters (FAMES).

More attractive biofuel alternatives may be produced microbially using the tools of synthetic biology. The complex nature of fuels and industrial chemicals suggests that production of a wide range of molecules with different properties will be necessary to offset petroleum-dependence. Thus, we pursued production of a variety of fuels and chemicals that are produced through fatty acid – like chemistry and include n-butanol, fatty alcohols, fatty esters, biodiesel, and diacids. In contrast to producing ethanol and biodiesel, producing these longer chain fuels from microbes will eliminate a significant amount of energy required for processing (distillation, crushing, etc.), and eliminate the requirements for fertilizer and irrigation dependent upon feedstock. Fatty acid biosynthesis is a metabolic pathway that can produce the precursors to long chain aldehydes, alcohols, biodiesels and alkanes by reduction or esterification of free fatty acids. This pathway is attractive for biofuel production because fatty acids, which contain over two times the energy of carbohydrates or proteins, are commonly used as energy storage molecules in many organisms. Some yeasts have been shown to

accumulate fatty acids to 50% of their weight⁹⁻¹⁴, indicating that it is possible to shunt significant amounts of energy into building fatty acids. In contrast, simpler, prokaryotic organisms like *E. coli* only accumulate fatty acids to 5% of their weight¹⁵. Finally, many of these fuels are directly compatible with current transportation infrastructure and engines.

1.2. Organization

The following chapters elucidate different aspects of our efforts to construct and engineer metabolic pathways for the production of fuels and chemicals. Chapter 2 provides relevant aspects of engineering an n-butanol pathway in *S. cerevisiae* and some of the previous efforts and problems that have been encountered with the native butanol producer, *Clostridium*. In chapter 3, I discuss our efforts toward engineering the native fatty acid biosynthesis pathway in *E. coli*, where we produce a wide range of fatty-acid products. Chapter 4 investigates the next steps in optimizing production of fatty acid ethyl esters and elucidates the next, most promising areas for engineering. Chapter 5 discusses strategies employed for producing fatty acid fuels and chemicals in yeast. Chapter 6 explores production of diacids via the fatty acid pathway.

1.3. References

1. Wikipedia. Petroleum. www.wikipedia.com/petroleum (accessed 4/2010).
2. Dolin, E. in *Boston Globe* (2007).
3. Hill, J., Nelson, E., Tilman, D., Polasky, S. & Tiffany, D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proc Natl Acad Sci U S A* **103**, 11206-10 (2006).
4. (SECO)., S. E. C. O. (accessed 9/2007).
5. <http://www.biodiesel.org>, (accessed 4/2010).
6. <http://www.eia.doe.gov/oiaf/lleo/oil.html>. (accessed 4/2010).
7. Pimentel, D., Patzek, T. & Cecil, G. Ethanol production: energy, economic, and environmental losses. *Rev Environ Contam Toxicol* **189**, 25-41 (2007).
8. Chou H, S. E. Life cycle analysis of alternative biofuels produced from genetically engineered microbes. (2007).
9. Botham, P. A. & Ratledge, C. A biochemical explanation for lipid accumulation in *Candida* 107 and other oleaginous micro-organisms. *J Gen Microbiol* **114**, 361-75 (1979).
10. A Ykema, E. V., HJH Nijkamp, H Smit. Isolation and characterization of fatty acid auxotrophs from the oleaginous yeast *Apiotrichum curvatum*. *Appl Microbiol Biotechnol*, 76-84 (1989).
11. Gangar, A., Karande, A. A. & Rajasekharan, R. Isolation and localization of a cytosolic 10 S triacylglycerol biosynthetic multienzyme complex from oleaginous yeast. *J Biol Chem* **276**, 10290-8 (2001).

12. Gill, C. O., Hall, M. J. & Ratledge, C. Lipid accumulation in an oleaginous yeast (*Candida 107*) growing on glucose in single-stage continuous culture. *Appl Environ Microbiol* **33**, 231-9 (1977).
13. Hall, M. J. & Ratledge, C. Lipid Accumulation in an Oleaginous Yeast (*Candida 107*) Growing on Glucose Under Various Conditions in a One- and Two-Stage Continuous Culture. *Appl Environ Microbiol* **33**, 577-584 (1977).
14. M Hassan, P. B., L Granger, A Pareilleux, G Goma. Lipid production by an unsaturated fatty acid auxotroph of the oleaginous yeast *Apiotrichum curvatum* grown in a single-state continuous culture. *Appl Microbiol Biotechnol*, 483-488 (1993).
15. Atkinson B, M. F. Biochemical Engineering and Biotechnology Handbook. (1983).

Chapter 2: Engineering *Saccharomyces cerevisiae* for the production of n-butanol.

2.1. Introduction

Soaring energy costs and increased awareness of global warming have motivated production of renewable, biomass-derived fuels and chemicals. The reasons for producing alternatives to ethanol, the current biofuel standard, are numerous and clear: ethanol suffers from low energy density, it is hydroscopic, it cannot be piped, and it is costly to distill, an aspect that detracts from the total energy output of its production. Some of the properties of short-chain alcohols, targeted for use as gasoline replacements are shown in Table 2-1, modified from Rude et al, 2009¹. Ideally, biofuels will require minimal energy to separate from fermentation broths, be non-toxic to the host microorganism, and be efficiently produced from a variety of feedstocks². Compared to ethanol, n-butanol is more hydrophobic, has a higher energy density, can be transported through existing pipeline infrastructure, and can be mixed with gasoline at any ratio. Thus, n-butanol is a substantially better biofuel than ethanol. Further, as an industrial solvent, n-butanol commands a \$7-8.4B market with production between 10-12B lbs per year³.

Table 2-1: Characteristics of short-chain alcohols

Product	Pathway	Density (g/mL)	Metabolic Mass yield	gal product / ton glucose	enthalpy of combustion (MJ/kg)	enthalpy of combustion yield
Gasoline replacements						
ethanol	fermentative	0.79	51%	155	-29.7	97%
butanol	fermentative	0.81	41%	121	-36.1	95%
isobutanol	non-fermentative	0.81	41%	122	-35.9	94%
3-methyl-1-butanol	non-fermentative	0.81	33%	98	-37.7	80%

n-Butanol can be produced either chemically from petroleum or fermentatively in a variety of Clostridial species. Advances in biotechnology and increased petroleum costs have renewed interest in fermentative n-butanol production, however, Clostridia are not ideal because of the relative lack of genetic tools to manipulate their metabolism, their slow growth, their intolerance to n-butanol above 1–2% and oxygen, and their production of butyrate, acetone, and ethanol as byproducts. Thus, there is interest in producing n-butanol in a more suitable industrial organism.

Recently, two groups have re-constructed the n-butanol pathway from *Clostridia* and measured n-butanol production in *Escherichia coli* (~1 mM)^{4,5}. We chose *Saccharomyces cerevisiae* as a host for n-butanol production because it is a genetically tractable, well-characterized organism, the current industrial strain alcohol (ethanol) producer, and it has been previously manipulated to produce other heterologous metabolites⁶. Since n-butanol and ethanol only differ by two carbons, *S. cerevisiae* may be able to tolerate high concentrations of n-butanol by the same mechanisms it tolerates ethanol. Recently, *S. cerevisiae* has been demonstrated to have tolerance to n-butanol⁷. Here we demonstrate the engineering of *S. cerevisiae* for the production of n-butanol and provide insight for the next steps in engineering by detection of intermediate pathway metabolites.

S. cerevisiae was engineered with an n-butanol biosynthetic pathway, in which isozymes from a number of different organisms (*S. cerevisiae*, *E. coli*, *Clostridium beijerinckii*, and *Ralstonia eutropha*) were substituted for the *Clostridia* enzymes and their effect on n-butanol production was compared. By choosing the appropriate isozymes, we were able to improve production of n-butanol ten-fold to 2.5 mg/L. The most productive strains harbored the *C. beijerinckii* 3-hydroxybutyryl-CoA dehydrogenase, which uses NADH as a co-factor, rather than the *R. eutropha* isozyme, which uses NADPH, and the acetoacetyl-CoA transferase from *S. cerevisiae* or *E. coli* rather than that from *R. eutropha*. Surprisingly, expression of the genes encoding the butyryl-CoA dehydrogenase from *C. beijerinckii* (*bcd* and *etfAB*) did not improve butanol production significantly as previously reported in *E. coli*. Using metabolite analysis, we were able to determine which steps in the n-butanol biosynthetic pathway were the most problematic and ripe for future improvement.

2.2. Materials and methods

Chemicals

Ethyl acetate was purchased from Sigma-Aldrich (St. Louis, MO). Complete Supplement Mixtures for formulation of Synthetic Defined (SD) media were purchased from Qbiogene (Irvine, CA). All other media components were purchased from Sigma-Aldrich.

Strains and media

Clostridium beijerinckii NCIMB 8052 was purchased from ATCC, catalog number 51743. *E. coli* strains DH10B and DH5 α were used for bacterial transformation and plasmid amplification in the construction of the expression plasmids used in this study. The strains were cultivated at 37°C in Luria-Bertani medium with 100 mg L⁻¹

¹ampicillin. *S. cerevisiae* strain BY4742⁸, a derivative of S288C, was used as the parent strain for all yeast strains. This strain was grown in rich YPD medium at 30°C⁹. Engineered yeast strains were grown in SD medium⁹ with leucine, uracil, histidine, and/or methionine dropped out where appropriate. For induction of genes expressed from the *GAL1* and *GAL10* promoters, *S. cerevisiae* strains were grown in 2% galactose as the sole carbon source unless otherwise indicated.

Plasmid construction

C. beijerinckii genes were cloned from genomic DNA: *hbd* encodes 3-hydroxybutyryl-CoA dehydrogenase; *crt*, crotonase; *bcd*, butyryl-CoA dehydrogenase; and *etfA* & *etfB*, two-electron transferring flavoproteins A & B. *phaA* and *phaB* (*Ralstonia eutropha*), *adhe2* (*C. beijerinckii*), and *ccr* (*Streptomyces collinus*) were synthesized (Epoch Biolabs). All genes were PCR amplified with Phusion polymerase (New England Biolabs). Primers were designed to have 30-bp flanking regions homologous to the plasmid insertion regions, either the *gal1* or *gal10* promoter and the *CYC1*, *ADH1*, or *PGK1* terminator. Plasmid construction was carried out using the Sequence and Ligation Independent Cloning (SLIC) method, previously described¹⁰. The constructed plasmids were derived from pADS-AMO-CPR-opt-LEU2D plasmid¹¹ and pESC-HIS (Stratagene).

Yeast transformation and strain construction

Transformation of all *S. cerevisiae* strains was performed using the lithium acetate method¹². Strains ESY1-11 were constructed by the co-transformation of the indicated plasmids followed by selection on SD-LEU or SD-LEU-HIS plates as appropriate.

Yeast cultivation

All optical density measurements at 600 nm (OD₆₀₀) were taken using a Beckman DU-640 spectrophotometer. To measure n-butanol production, culture tubes containing 5 mL of SD (2% galactose) medium (with appropriate amino acid omissions as described above) were inoculated with the strains of interest. These inocula were grown at 30°C to an OD₆₀₀ between 1 and 2. Capped serum vials (100 mL) containing 50 mL SD medium were inoculated to an OD₆₀₀ 0.05 with these seed cultures in order to achieve a "semi" anaerobic condition. Samples were collected at 24, 72, 120, and 144 h and analyzed for metabolites as discussed below.

Metabolite detection

For metabolite analysis, cultures were sampled (10 mL) at 24, 72, 120, and 144 h. For n-butanol detection, 2 mL ethyl acetate containing n-pentanol (0.005% v/v), an internal standard, was added to the 10 mL sample and vortexed for 1 min. The ethyl acetate was then recovered and applied to a Thermo Trace Ultra gas chromatograph (GC) equipped with a Triplus AS autosampler and a TR-WAXMS column (Thermo Scientific). The samples were run on the GC with the following program: initial temperature, 40°C for 1.2 min, ramped to 130°C at 25°C/min, ramped to 220°C at 35°C/min. Final quantification analysis was achieved with Xcalibur software. For pathway intermediate analysis, a method previously established was employed¹³. Specifically, 10-mL samples were pelleted (9000 × g, 5 min, 4°C). The supernatant was aspirated, and the cells were suspended in 1 mL of 10% TCA containing propionyl-CoA (10 µM) as an internal standard. The cells were bead-beaten for 4 min. The supernatant was collected and neutralized with 2× volume of 1 M octylamine. Samples were then filtered and separated on a Zorbax 300SB-C1 8 column (Agilent; 2.1 mm i.d. × 10 cm length) using an Agilent 1100 series HPLC at a flow rate of 150 µL/min. The LC conditions used were adapted from Pitera *et al*¹⁴. Briefly, samples were run from initial conditions of $t = 0$ min in 95% 100 mM ammonium acetate (Buffer A), 5% 100 mM ammonium acetate:acetonitrile (70%:30%) (Buffer B); $t = 5$ min, 95% Buffer A, 5% Buffer B; $t = 12$ min, 80% Buffer A, 20% Buffer B; $t = 16$ min, 10% Buffer A, 90% Buffer B; $t = 25$ min, 10% Buffer A, 90% Buffer B; $t = 28$ min, 95% Buffer A, 5% Buffer B; $t = 45$ min, 95% Buffer A, 5% Buffer B. The LC system was interfaced to an Applied Biosystems Q TRAP 2000 LC/MS/MS via a Turbo Ionspray source operating in the positive ion mode (5500 V). The MS was operated in single-ion-monitoring (SIM) mode with a dwell time of 200 ms for each CoA metabolite of interest. Data were collected and analyzed with Analyst™ 1.4.2 (Applied Biosystems).

Organic acids, sugars, and ethanol were measured by sampling 1 mL of culture, centrifuging 18,000 × g, 5 min, and applying the supernatant to an Agilent 1100 series HPLC equipped with an Aminex HPX-87H ion exchange column (Biorad). The solvent, 4 mM H₂SO₄, flow rate was 0.6 mL/min and the column was maintained at 50°C. All metabolites were detected with an Agilent 1200 series DAD and RID detectors.

Protein Detection

Briefly, strains were grown in inducing conditions for 72h, cells were pelleted and then lysed via bead beating in 500mM triethylammonium bicarbonate (TEAB)

with 8M urea. The soluble protein fraction was reduced with tris-(2-carboxyethyl)phosphine (TCEP) and alkylated with iodoacetic acid. The sample was diluted to reduce the urea concentration and subsequently digested with mass spectrometric grade trypsin (Trypsin Gold, Promega Inc, Madison, WI; 50:1 sample:trypsin ratio) at 37°C for 16 hours¹⁵. The digestion was stopped by freezing at -20°C. The sample was stored at -20°C until analysis via reverse-phase LC-MS/MS. For LC-MS/MS analysis, 40 µL of digested protein were loaded on an Ultimate nano-HPLC system (Dionex-LC Packings, Sunnyvale, CA) coupled to a QSTAR Pulsar-i mass spectrometer (Applied Biosystems Inc. Foster City, CA) via a Switchos Micro Column Switching pump (Dionex-LC Packings) operating at 20 µL/min flow rate. The peptides were separated on a homemade nano-column (75 µm i.d. x 15 cm length) packed with Magic C18 packing material (5 µm bead size, 100Å pore size) at a flow rate of 200 nL/min. The LC conditions were as follows: $t=0$ min, 100% Buffer A (2% (v/v) acetonitrile, 0.1% (v/v) formic acid); $t=10$ mins, 100% Buffer A; $t=130$ mins, 65% Buffer A, 35% Buffer B (98% (v/v) acetonitrile, 0.1% (v/v) formic acid); $t=140$ mins, 20% Buffer A, 80% Buffer B; $t=160$ mins, 20% Buffer A, 80% Buffer B; $t=165$ mins, 100% Buffer A; $t=185$ mins, 100% Buffer A. The LC system was interfaced to the QSTAR mass analyzer via a nanospray source operating in the positive ion mode (2400 V). Data were collected with Analyst™ 1.1 (Applied Biosystems) and Information Dependent Acquisition (IDA; Applied Biosystems). The two most abundant multiply-charged ions from a 0.5 s MS survey scan (350-1600 amu) above a threshold of 50 counts were selected for IDA analysis. Selected ions were isolated in Q1 (resolution=LOW) and were fragmented with rolling collision energy. MS/MS scans were collected for 2 s over a mass range of 100-1600 amu. Parent ions and isotopes were excluded from subsequent IDA selection for 60 s following one repeat analysis.

A ProteinPilot 2.0 (Applied Biosystems) search was performed for each dataset against a *S. Cerevisiae* database consisting of all putative ORF sequences, exogenous protein sequences for the n-butanol biosynthesis pathway, and common contaminants (*e.g.*, trypsin, keratin.) Search parameters were matched to the sample preparation conditions (*i.e.*, cysteine modification by a carboxymethyl group from iodoacetic acid, digestion with trypsin, potential oxidation of methionine) with the Paragon algorithm set to report proteins identified at 95% confidence level (ProtScore = 1.3) from a thorough search with biological modifications included. A ProtScore of 2 corresponds to a peptide identified by the Paragon algorithm at the 99% confidence level and protein Protscores are the sum of all the identified peptide ProtScores for that specific protein.

2.3. Results and discussion

2.3.1. Expression of n-butanol pathway isozymes from a range of organisms in *S. cerevisiae* results in n-butanol production

In addition to the difficulty associated with engineering Clostridia strains, the pathway for producing butanol is not efficient as outline below in Figure 3-1, where the acidogenic phase first produces the acids acetate and butyrate through a phosphate-activated intermediate. Then, during the solventogenic phase, the acids are re-activated to a CoA intermediate that is reduced, requiring ATP and NADH as cofactors.

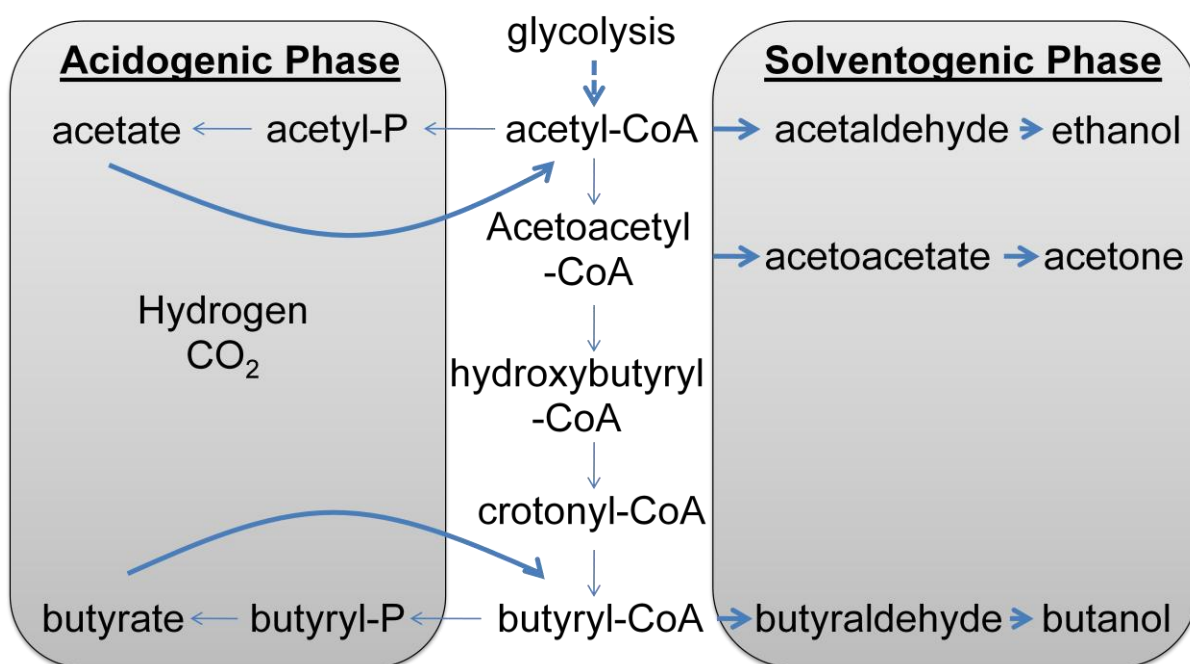


Figure 2-1: Clostridial ABE fermentation pathway.

Thus, we eliminated the acidogenic phase reactions and focused upon the solventogenic phase reactions in reconstructing a butanol pathway in *S. cerevisiae* in order to avoid excess cofactor utilization. The reactions and corresponding enzymes are outlined in Figure 3-2. As enzymes with the same catalytic function from different organisms have different catalytic activity, ability to be expressed, solubility, etc., we sought to test a number of candidates for each reaction in the pathway. The genes encoding these enzymes were cloned into two different plasmids (Figure 3-3 and Table 3-2) and transformed into *S. cerevisiae* BY4742.

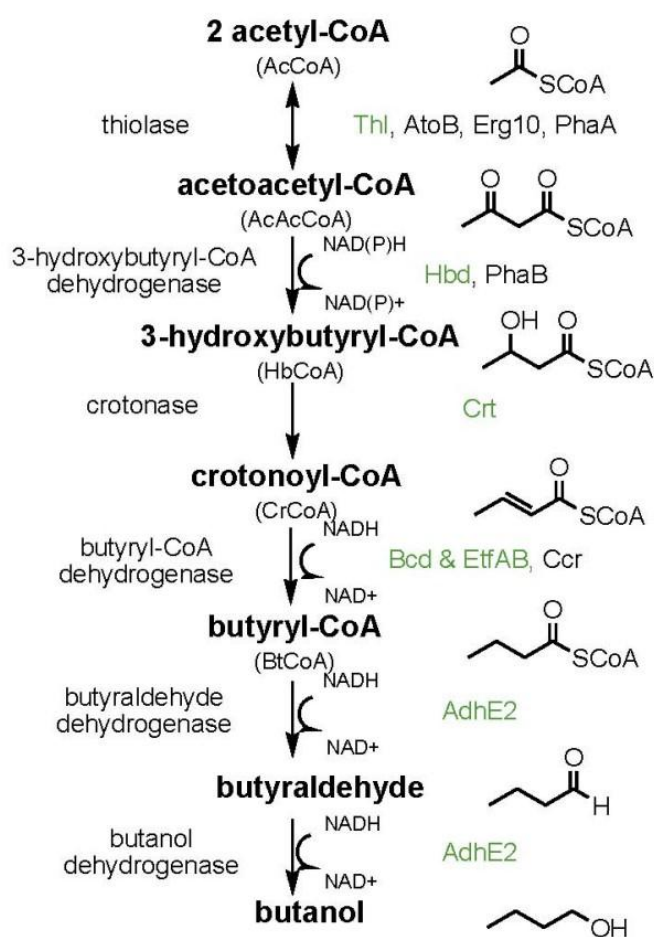


Figure 2-2: The n-butanol biosynthetic pathway. The enzymes in green are from *Clostridium beijerinckii*. Enzymes in black are from other organisms: *AtoB*, *Escherichia coli*; *Erg10*, *S. cerevisiae*; *PhaA*, *Ralstonia eutropha*; *PhaB*, *Ralstonia eutropha*; *Ccr*, *Streptomyces collinus*. Each enzyme candidate was evaluated in the pathway for n-butanol production (except *thl*, which is native to Clostridia).

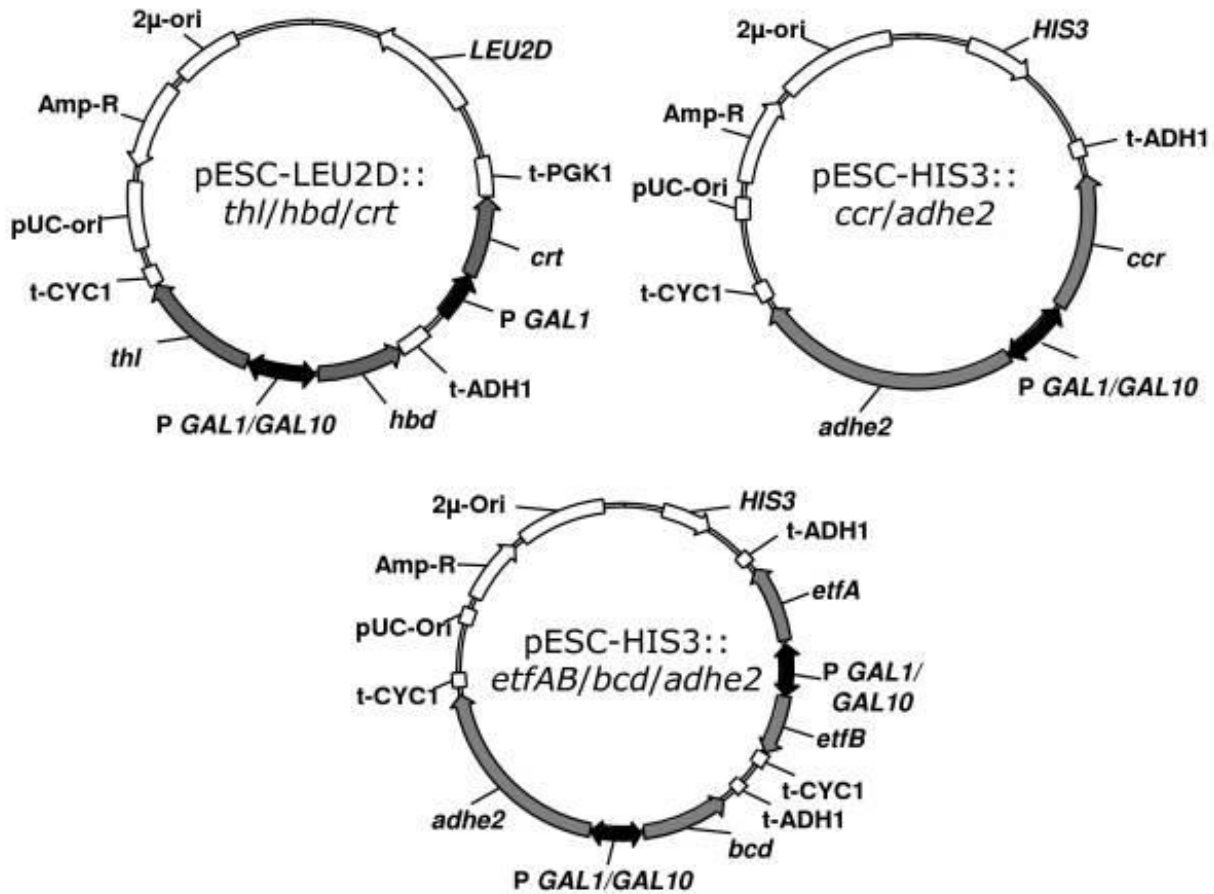


Figure 2-3: Representative plasmids used in this study. Plasmids were constructed by the SLIC method, previously described. They contain the 2μ origin of replication, *LEU2D* or *HIS3* genes for selection, the *GAL1* or *GAL10* promoters, and the *CYC1*, *ADH1*, or *PGK1* transcription terminators. The first three genes of the n-butanol pathway were placed on the pESC-LEU2D plasmid and the last two or four (in the case of the *etfAB*, *bcd* bearing strain) genes were placed on the pESC-HIS3 plasmid.

Table 2-2: Strains used in this study

Name	Genotype	Reference
BY4742	Mat α ; <i>HIS3Δ1</i> ; <i>LEU2Δ0</i> ; <i>LYS2Δ0</i> ; <i>URA3Δ0</i> ; <i>YDR242w::kanMX4</i>	⁸
ESY2	BY4742: pESC- <i>phaA-phaB-crt-LEU2D</i> + pESC- <i>ccr-adhe2-HIS</i>	This study
ESY3	BY4742: pESC- <i>atoB-phaB-crt-LEU2D</i> + pESC- <i>ccr-adhe2-HIS</i>	This study
ESY4	BY4742: pESC- <i>ERG10-phaB-crt-LEU2D</i> + pESC- <i>ccr-adhe2-HIS</i>	This study
ESY5	BY4742: pESC- <i>phaA-hbd-crt-LEU2D</i> + pESC- <i>ccr-adhe2-HIS</i>	This study
ESY6	BY4742: pESC- <i>atoB-hbd-crt-LEU2D</i> + pESC- <i>ccr-adhe2-HIS</i>	This study
ESY7	BY4742: pESC- <i>ERG10-hbd-crt-LEU2D</i> + pESC- <i>ccr-adhe2-HIS</i>	This study
ESY11	BY4742: pESC- <i>ERG10-hbd-crt-LEU2D</i> + pESC- <i>etfA-etfB-bcd-adhe2-HIS</i>	This study

The first set of strains, ESY2, ESY3, and ESY4, were engineered with enzymes from *Ralstonia eutropha* (*phaA* and *phaB*), *Streptomyces collinus* (*ccr*), *Clostridium beijerinckii* (*crt* and *adhe2*), *E. coli* (*atoB*), and *S. cerevisiae* (*ERG10*) and only varied in the first committed step, the thiolase (PhaA, AtoB, and ERG10). We tested the enzymes from *R. eutropha*, because these enzymes have previously been demonstrated to retain high activity for the production of polyhydroxyalkanoates in *E. coli*¹⁶. The two other thiolases were tested because ERG10 is the native thiolase in *S. cerevisiae*, and AtoB has been successfully used to overproduce acetoacetyl-CoA (AcAcCoA) in other metabolic pathways¹⁷. These strains produced different levels of n-butanol (Figure 3-4), the highest being ESY2 (1 mg/L), suggesting that PhaA is the best thiolase in this specific pathway configuration. From these strains, we constructed and tested strains with different enzyme candidates for their ability to increase n-butanol production.

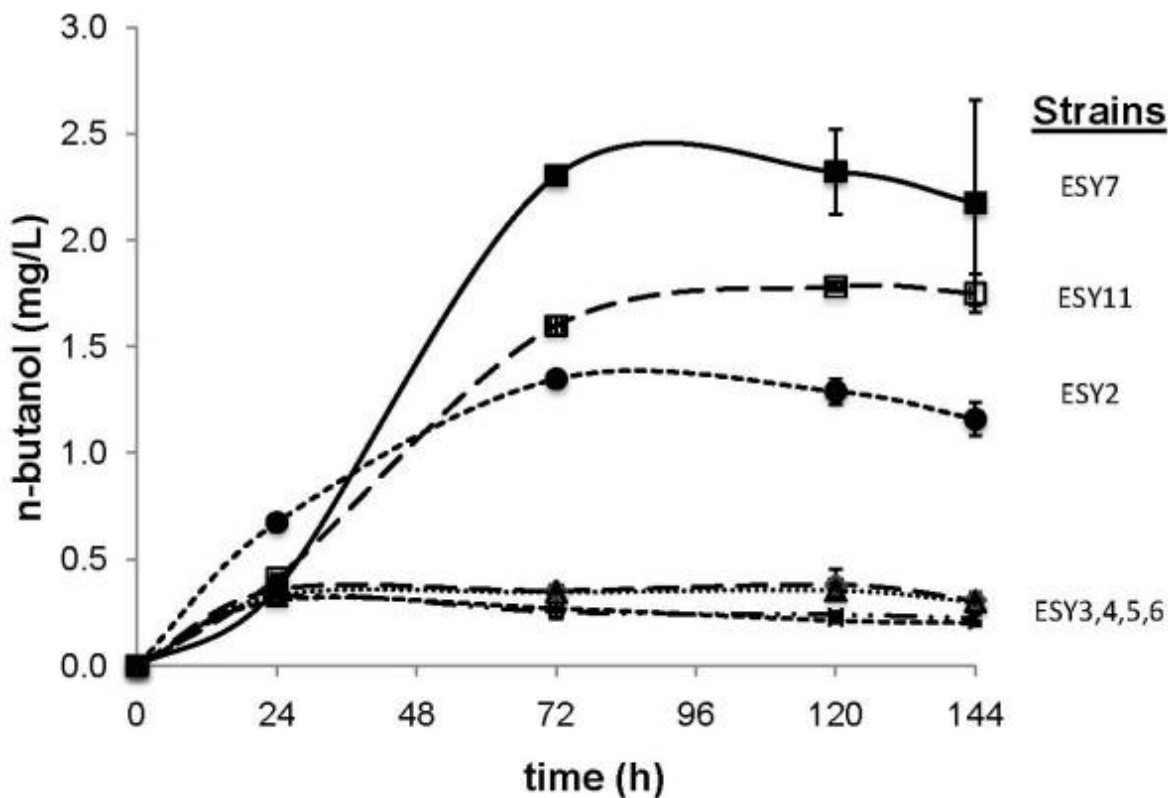


Figure 2-4: n-Butanol production from engineered *S. cerevisiae*. Symbols and strains: black squares, ESY7; empty squares, ESY11; black circles, ESY2; the rest of the samples all produced approximately the same amount of n-butanol and are indicated on the graph. Symbols and error bars represent the mean and standard deviation of triplicate cultures.

The second set of strains was constructed to compare n-butanol production using different isozymes for 3-hydroxybutyryl-CoA (HbCoA) dehydrogenase, which converts AcAcCoA into HbCoA. One of these isozymes uses NADPH (PhaB) as a cofactor, whereas the other isozyme uses NADH (Hbd) as a cofactor. As cells growing under fermentative conditions generally have excess NADH, Hbd might be preferred to PhaB¹⁸. Although PhaA appeared to be the best thiolase based on the results using the previous set of strains, it was unclear that PhaA would function as well with HbCoA dehydrogenase as it did with PhaB (given the context dependence of some pathway configurations¹⁹). Thus, we tested Hbd with all three thiolases. These strains, ESY5, ESY6, and ESY7, all produced detectable levels of n-butanol, while ESY7 (*ERG10, hbd*) doubled the production of n-butanol over the previously highest producing strain, ESY2 (*phaA, phaB*), reaching 2.5 mg/L. It is

difficult to attribute the increase in production to one specific enzyme, because ESY7 harbors a different thiolase (*ERG10*) and a different HbCoA dehydrogenase (*hbd*) that utilizes NADH, compared to ESY2. It is curious why the strain harboring PhaA is not as productive with Hbd as it is with PhaB. It may be that PhaA and PhaB have been optimized through evolution to work together to maximize production of polyhydroxybutyrate in *R. eutropha*. That the strain harboring *ERG10* and Hbd is the highest producer is not surprising given that *ERG10* is the native thiolase and Hbd uses NADH, which should be in ample supply under fermentation conditions.

The final strain (ESY11) was made to determine if an alternative butyryl-CoA (BtCoA) dehydrogenase, which was previously shown to greatly increase n-butanol production in *E. coli*¹⁸, would improve n-butanol production over that by ESY7. Surprisingly, expression of the *C. beijerinckii bcd* and *etfAB* (ESY11) did not improve n-butanol production significantly. Notably, ESY11 was the second highest n-butanol producer and compared to ESY7, only differed in the BtCoA dehydrogenase.

In summary, the best strain, ESY7, produced 2.5 mg/L n-butanol from 2% galactose as a carbon source. This strain overexpressed the native thiolase (*ERG10*) and an HbCoA dehydrogenase (Hbd) that utilized NADH, which the cell should have in plentiful supply under fermentative conditions. ESY7, which also harbored *ccr*, produced slightly more n-butanol than ESY11, which harbored *bcd* and *etfAB*.

2.3.2. Analysis of intermediary metabolites

In an attempt to explain the differences in n-butanol production by the various strains, we analyzed the pathway intermediates. Previously, this analysis has proven to be successful in interrogating and optimizing metabolic pathways¹⁴. We developed a single LC-MS method to monitor all of the metabolites in the n-butanol biosynthetic pathway simultaneously. Detection of all the metabolic standards was successful, while detection of intermediates in cell extracts was successful for all intermediates except AcAcCoA. Although we analyzed the metabolic pathway intermediate levels in all strains, we present the results for strains ESY4, ESY7, and ESY11 only. We chose ESY7 because it was the highest producer, ESY4 because it differed from ESY7 only in the HbCoA dehydrogenase (PhaB in ESY4 versus Hbd in ESY7) and yet displayed a 10-fold difference in n-butanol titers (Figure 3-5). We chose ESY11 because it was the second highest producer and differed from ESY7 only in the choice of BtCoA dehydrogenase (*Ccr* in ESY7 and *EtfAB/Bcd* in ESY11).

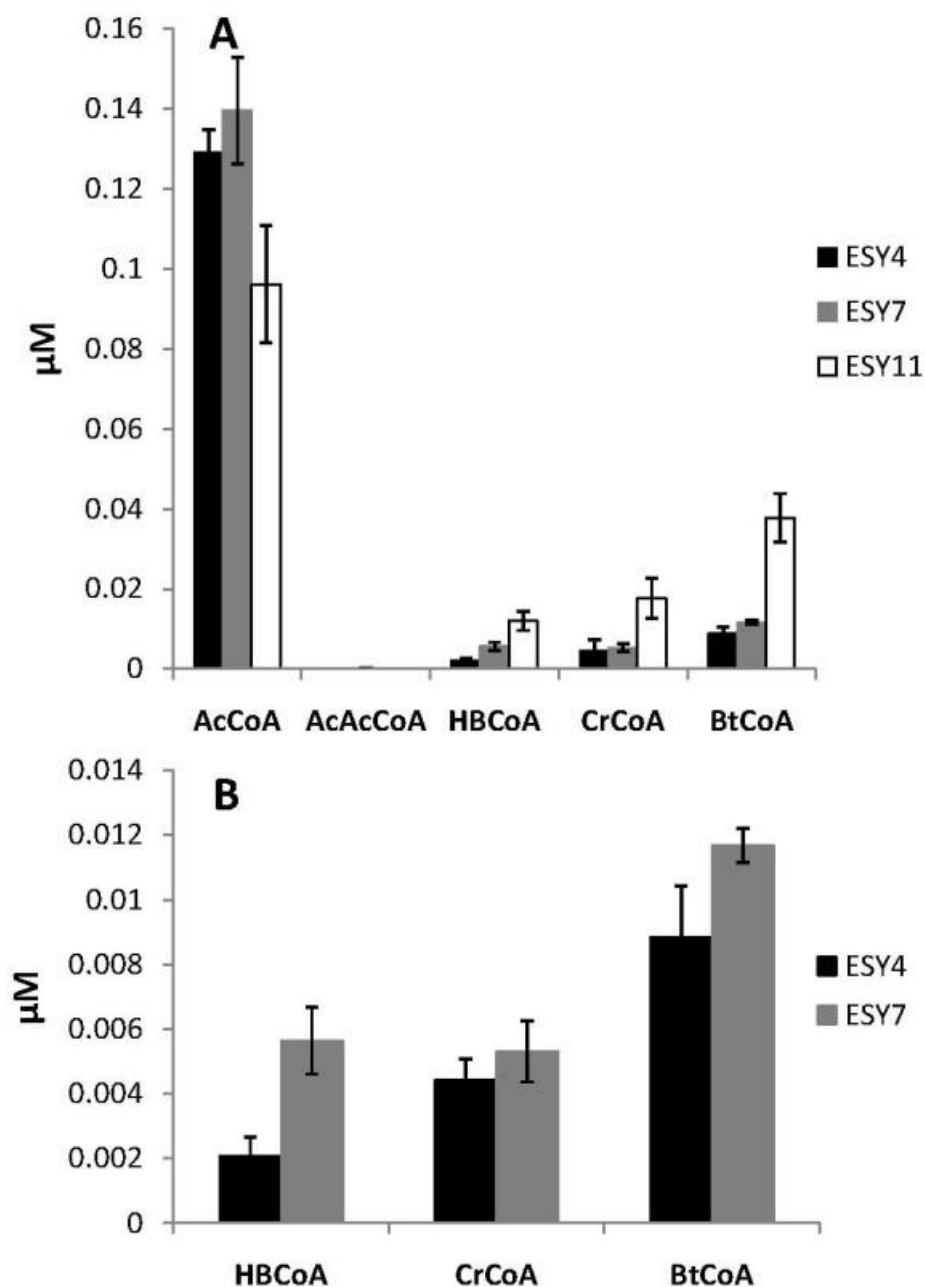


Figure 2-5: n-Butanol pathway intermediates at 24 h. Bars and strains: black bars, ESY4; gray bars, ESY7; white bars, ESY11. (A) All pathway intermediates in strains ESY4, 7 and 11. (B) HbCoA, CrCoA and BtCoA intermediates in strains ESY4 and ESY7. Levels of AcCoA were similar except for strains ESY11 (A). Levels of 3-hydroxybutyryl-CoA (HbCoA) and butyryl-CoA (BtCoA) were notably higher in ESY7 compared to ESY4, while crotonyl-CoA (CrCoA) was relatively similar in the two strains. Values and error bars represent the mean and standard deviation of triplicate cultures.

At 24 h, the levels of AcCoA in strains ESY4 and ESY7 were indistinguishable, while the levels of HbCoA and BtCoA were higher in strain ESY7 than in ESY4, suggesting that the Hbd enzymatic reaction had higher flux than the PhaB enzymatic reaction. ESY11, which differs from ESY7 only in the BtCoA dehydrogenase (Ccr in ESY7 versus EtfAB/Bcd in ESY11), accumulated BtCoA, which suggests a bottleneck at BtCoA. Increased expression of *adhe2* in ESY11 could alleviate the accumulation of BtCoA and improve the production of n-butanol, perhaps beyond that of ESY7. Further, it should be noted that Ccr is a bifunctional carboxylase and reductase that mainly produces ethylmalonyl-CoA [Erb Alber, PNAS, 2009]. Since the levels of the other pathway metabolites were essentially indistinguishable between ESY4 and ESY7, we conclude that production of HbCoA by 3-hydroxybutyryl-CoA dehydrogenase is the rate-limiting reaction that determines the n-butanol production in this context. Proteomics analysis showed that Hbd (in ESY7) and PhaB (in ESY4) were expressed at equivalent levels (data not shown). As such, the differences in the NADPH and NADH levels may explain the higher production of n-butanol by ESY7 relative to ESY4. Indeed, NADPH limitation has been demonstrated for other pathways²⁰. Additionally, there is slightly more BtCoA in ESY7 than in ESY4, which may ultimately correspond to higher n-butanol production by ESY7. Interestingly, BtCoA seems to accumulate in all strains and suggests further engineering of Adhe2 is necessary. In support of this claim, we analyzed the solubility of Adhe2 by western blot and found the majority of the protein in the insoluble fraction (Figure 2-6). The trends in the levels of pathway intermediates between strains ESY7 and ESY4 were maintained through 72 h. Furthermore, increased expression of *crt* may alleviate the higher levels of HbCoA and create a more balanced pathway.

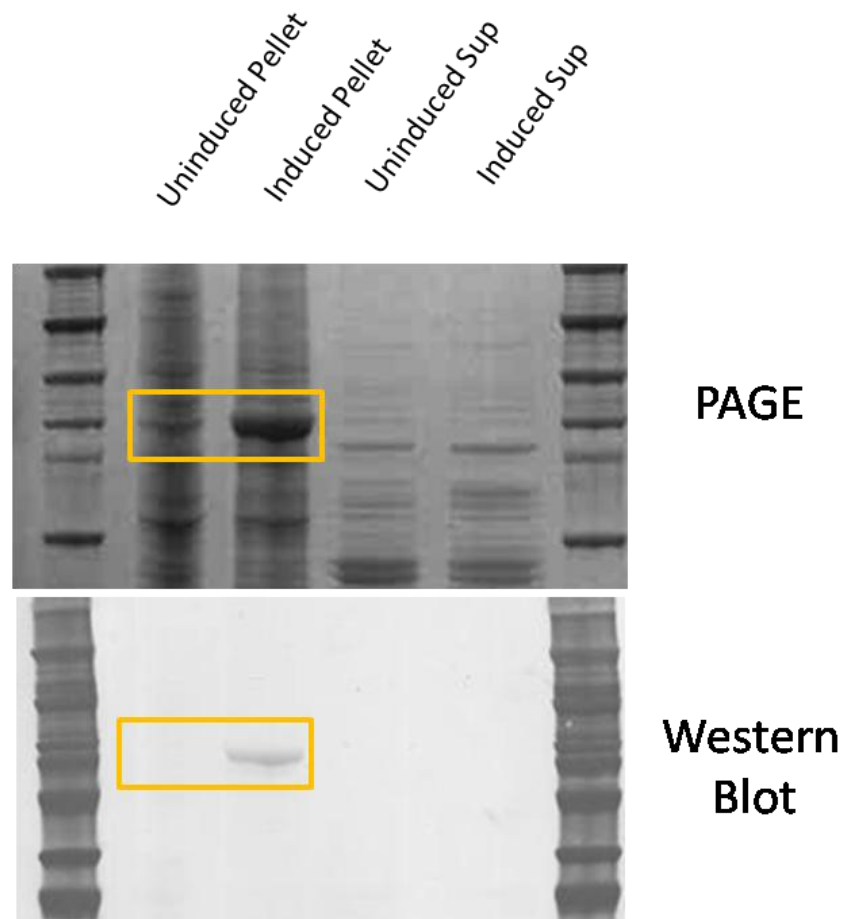


Figure 2-6: Protein analysis of strain ESY7 expressing Adhe2.

2.3.3. Engineering *adhe2* for solubility; increase activity to butyryl-CoA

It is useful to predict whether heterologous proteins will express in soluble form because solubility is related to functional protein and therefore flux of metabolic pathways. There exist a number of bioinformatics tools that make this prediction²¹⁻²³. Using two algorithms, one developed by Wilkinson et al, and another developed by Smialowski et al, the predicted solubility of various heterologous proteins in the butanol pathway are shown in Table 2-3.

Table 2-3: Predicted solubilities of heterologous butanol enzymes.

Protein	Predicted chance of insolubility (<i>E. coli</i>) (%)	PROSO predicted chance of insolubility (%)
ADHE2	71	94
HBD	52	53
CRT	65	58
PHAB	75	27
CCR	55	54
ATOB	72	53

The prediction that ADHE2 is insolubly expressed was verified by SDS-PAGE and western blot analysis as previously illustrated. Due to the insolubility of ADHE2 in *S. cerevisiae*, we hypothesized that a soluble form of the protein would aid in increasing the production of n-butanol. There are a number of approaches to increasing protein solubility. Some of the most common techniques rely upon fusing the insoluble protein to linker proteins, mutating protein surface hydrophobic residues to hydrophilic residues, co-expressing chaperone proteins or simply growing cells at lower temperatures²⁴. These strategies are reported to have mixed success and generally attempt to increase proper protein folding by slowing the kinetics of folding. We attempted fusing *adhe2* to the gene encoding maltose binding protein and saw no improvement in butanol production.

2.3.4. Metabolic byproducts of n-butanol producing yeast strains

An important goal of metabolic engineering is efficient production of target metabolites such that energy is not lost in the production of byproducts. Thus, we were interested in analyzing the levels of the common byproducts produced by yeast during butanol fermentation. Since yeast naturally produces a significant amount of ethanol and glycerol under high sugar concentrations, redirecting flux away from these products has been a desired goal, although little success has been achieved in completely eliminating ethanol production²⁵. Part of this difficulty lies in the number of alcohol dehydrogenase enzymes present in *S. cerevisiae*. By 72 h, 2% galactose was consumed, correlating with the cessation of both butanol and ethanol production. Additionally, other pathways in *S. cerevisiae* are hypothesized to metabolize some of the n-butanol intermediates and the product itself. In the cytosol, the major competing reactions that utilize AcCoA are the fatty acid and mevalonate pathways. The mevalonate pathway may also compete for AcAcCoA, therefore down-regulation of these pathways may be a successful strategy in directing carbon flux to butanol production. Furthermore, the other intermediates are participants in β -oxidation, although the pathway is confined to

the peroxisome. Finally, it is known that n-butanol can be oxidized in vitro by the native *S. cerevisiae* ADHs, although the activity is low compared to ethanol^{26,27}. In order to determine what metabolites were forming other than butanol we measured them via HPLC and report the values for various times below in Table 2-4 and Figure 2-7.

Table 2-4: Metabolite formation in strain ESY7

h	butanol (mM)	succinate (mM)	lactate (mM)	glycerol (mM)	acetate (mM)	ethanol (mM)
24	0.003	0.55	0	15.64	10.07	30.42
72	0.032	3.04	2.47	20.52	1.53	185.81
120	0.032	3.63	2.17	28.44	1.55	196.46

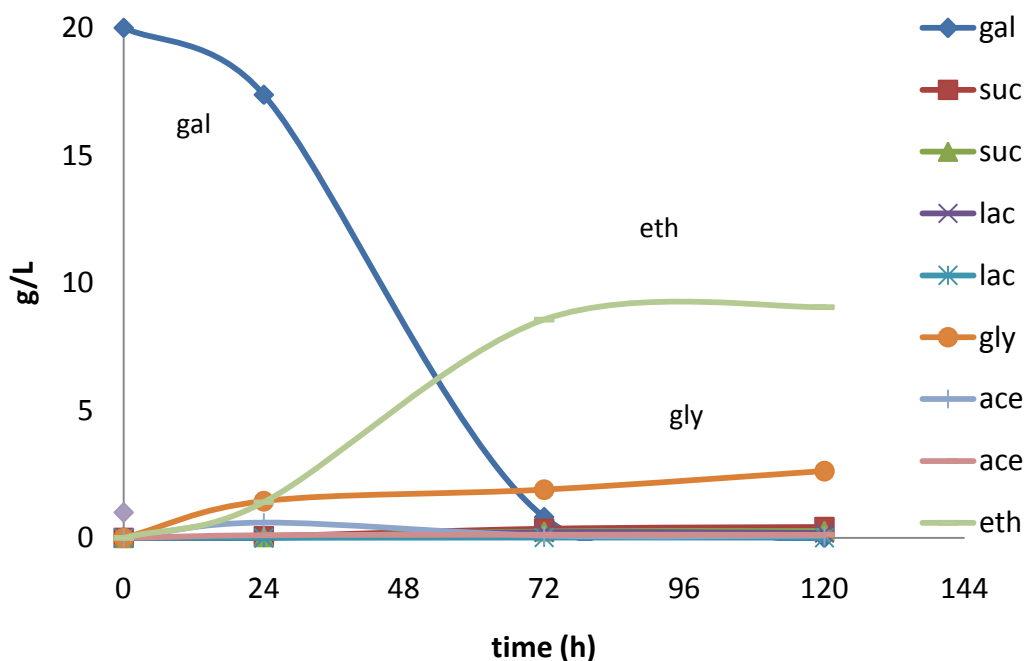


Figure 2-7: Metabolic byproduct formation in strain ESY7.

2.3.5. Removal of competing pathways

One of yeast's major aerobic fermentation byproducts is ethanol, which it produces through the Crabtree effect when high concentrations of sugar are present in the media. The Crabtree effect occurs because the rest of yeast metabolism kinetics are not as fast as ethanol production, which rapidly regenerates NAD⁺ that is required for glycolysis^{28,29}. We hypothesized that removal of the ethanol pathway may increase the production of butanol in yeast. Further, since butanol and ethanol are similar molecules only differing by two carbons, we hypothesized that butanol degradation by yeast could be possible. There are seven annotated alcohol dehydrogenases in yeast (ADH1-ADH7). ADH1-ADH5 utilize NADH as a cofactor and are involved in ethanol metabolism: ADH1, 3, 4, and 5 reduce acetaldehyde to ethanol, while ADH2 is more associated with oxidation of ethanol to acetaldehyde. ADH6 and ADH7 are broad substrate alcohol dehydrogenases thought to be involved in lignin degradation and utilized NADPH as a cofactor. We obtained knockouts from OpenBiosystems in the background strain, BY4742 and investigated whether exogenously added butanol could be catabolized by yeast in the various knockout strains growing on either 2% glucose or 2% galactose in culture tubes. A control condition was used where butanol evolution from yeast media was monitored to determine if butanol was naturally evaporating or actually being catabolized. Data are shown in Figures 2-8 and 2-9 below. These preliminary experiments indicate there is not a major difference between butanol evaporation from the media and catabolism. However, it does seem that removal of ADH2 may slightly decrease butanol catabolism if any is occurring. We did not further pursue these experiments because the initial results did not show major differences when the ADHs were removed. However, it could be that removal of all the ADHs will be necessary to see more significant differences due to their substrate promiscuity and complementation of one another.

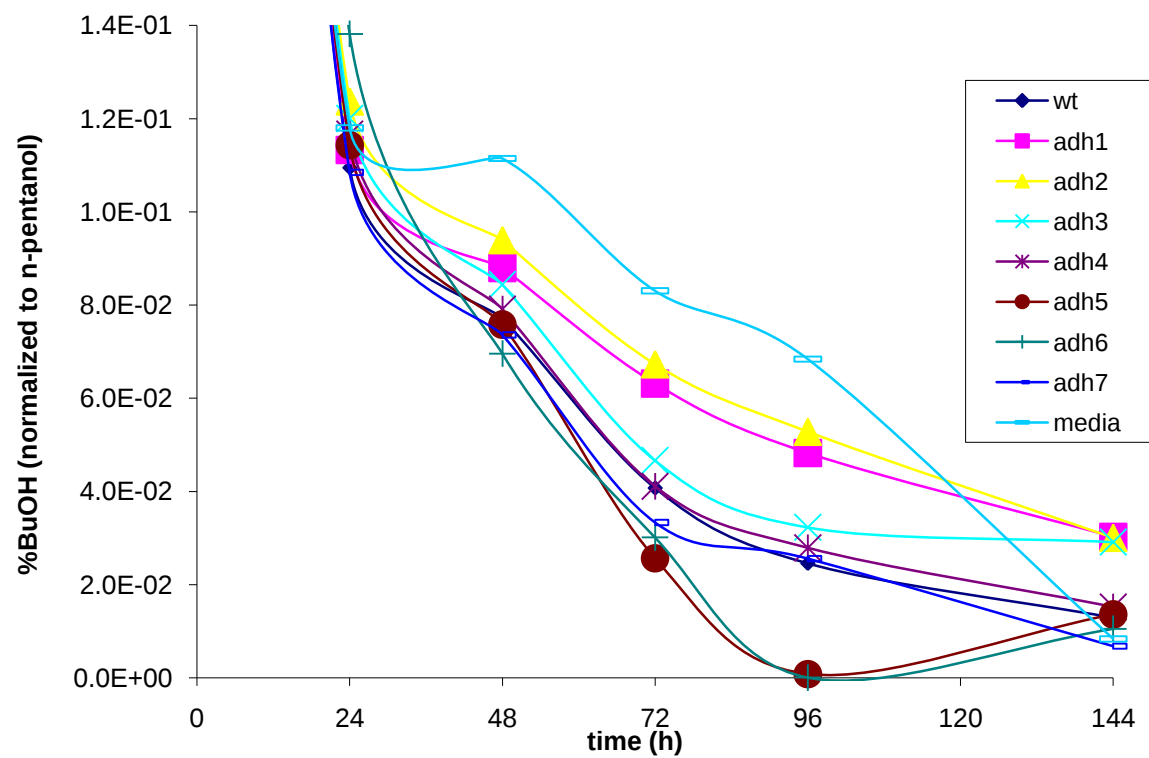


Figure 2-8: Butanol evolution from ADH knockouts. Strains with a single ADH removed (ADH1-7) were grown in 2% galactose with 0.5% exogenously added n-butanol.

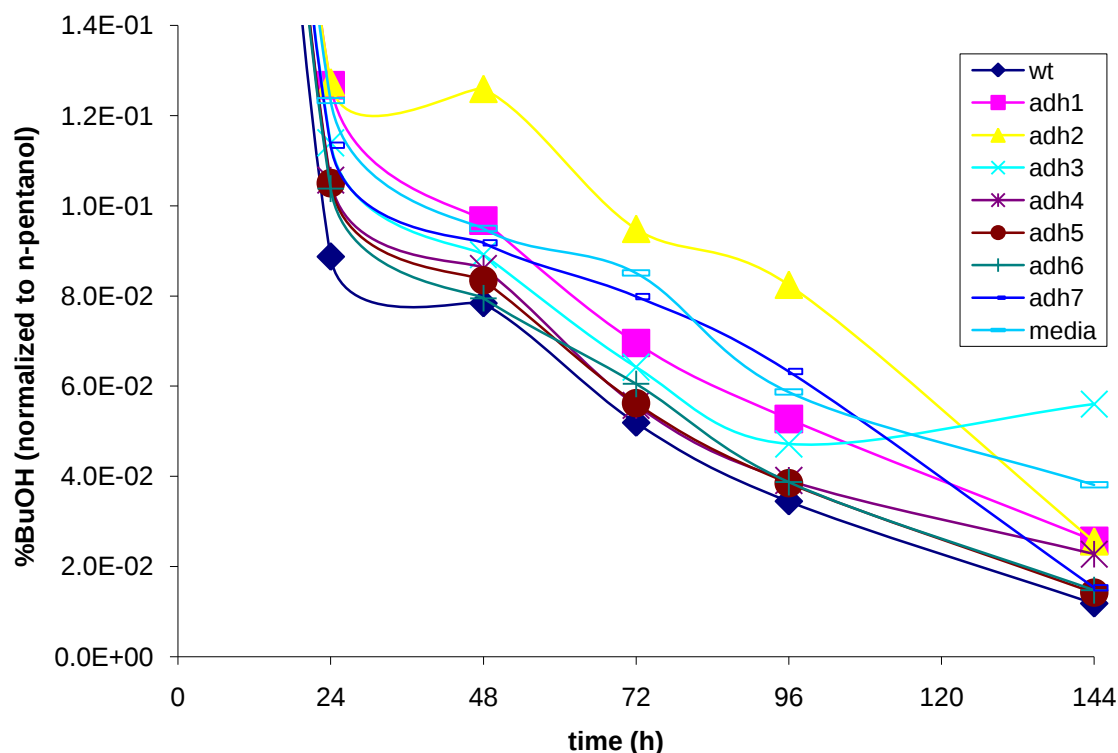


Figure 2-9: Butanol evolution from ADH knockouts. Strains with a single ADH removed (ADH1-7) were grown in 2% glucose with 0.5% exogenously added n-butanol.

Although removal of the ADHs had little effect on exogenously added butanol, we were interested in whether the same finding would be true of cultures that are producing butanol, thus we transformed the knockouts with the plasmids in strain ESY7 (pESC-Erg10-hbd-crt-leu2d + pESC-ccr-adhe2-his) (Figure 2-10 below). Comparisons of butanol production were made to the control strain, ESY7. Interestingly, all knockout strains except ADH5 and ADH7 exhibited lower levels of butanol production than ESY7 (~16 mg/L). Coupled with evidence of insolubility of Adhe2, this finding led us to believe that the native yeast ADHs may also involved in reducing butyryl-CoA to butanol.

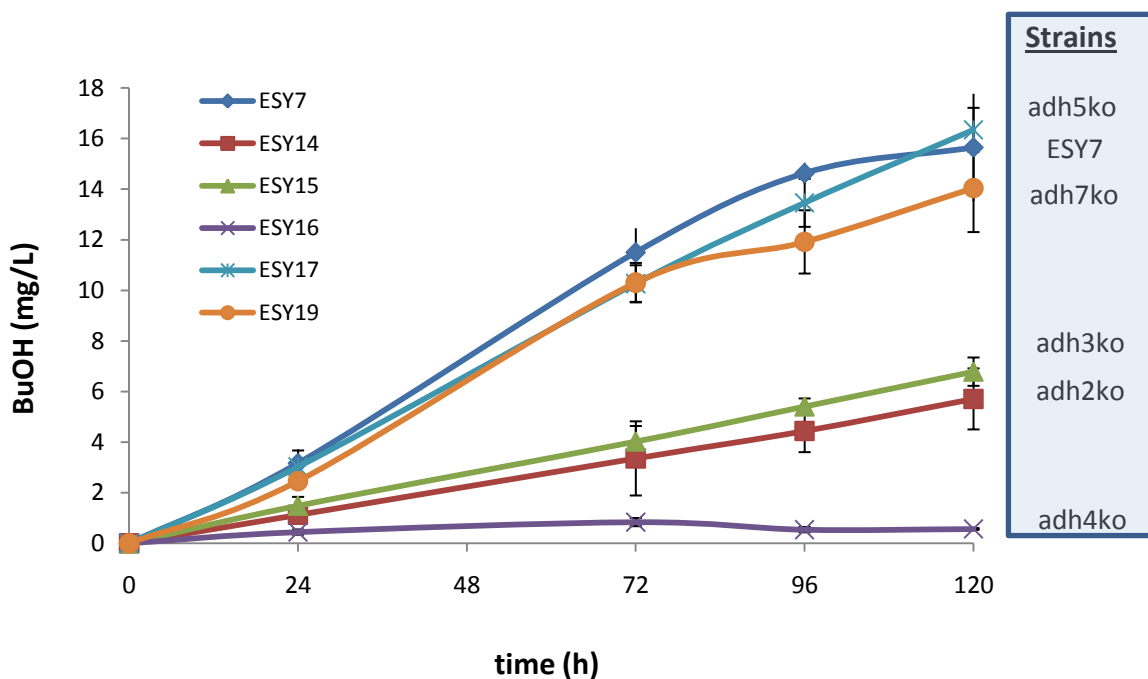


Figure 2-10: Butanol production from ADH knockout strains. N-butanol production was measured in strains with single ADH knockouts (ADH2, 3, 4, 5, and 7) harboring the plasmids in strain ESY7 that encode the n-butanol pathway. Values and error bars represent the mean and standard deviation of 3 experiments. ADH1 and ADH6 were unsuccessful in transforming the ESY7 plasmids.

2.3.6. Mass Spectrometric Analysis of protein levels

Data of representative strains are shown below in Table 2-5. While a new method, multiple reaction monitoring (MRM), of proteins by mass spectrometry is best implemented for making comparisons of the same protein across different levels of expression due to the biasing that may occur when comparing isozymes that may have different fragmentation patterns. Thus, in the strains depicted in the data set below, comparisons can only be made between HBD, CRT, BCD, and AAD, all of which are the same, whereas the THL is only the same for ESY3 and ESY6 (AtoB) and ESY4 and ESY7 (Erg10) constant. It seems there is a general expression problem for the BCD gene that is driven by the GAL10 promoter since the BCD is not detected. But based upon analysis of the intermediates in the pathway, a native *S. cerevisiae* enzyme exists or the BCD is expressed at low levels and does convert crotonyl-CoA into butyryl-CoA. The GAL10 promoter also drives expression of

the HBD gene, which is detected at low levels. Finally, the low detection of AAD is expected in part due to its insolubility. However, it is clear that the highest producer, ESY7 has detectable levels of soluble AAD, while ESY3 and EYS4 do not.

Table 2-5: Protein levels of various strains in n-butanol producing cultures.

	THL	HBD	CRT	BCD	AAD
ESY3	6.00	2.00	12.14	ND	ND
ESY4	15.55	ND	8.70	ND	ND
ESY6	8.04	1.19	10.60	ND	2.20
ESY7	6.46	ND	2.89	ND	1.35

2.3.7. Acetyl-CoA compartmentalization

Many industrially relevant products are derived in some way from acetyl – coenzyme A (CoA), including butanol, fatty acids, isoprenoids, etc. Thus the availability of acetyl-CoA is an important consideration in optimizing production from these metabolic pathways. Although prokaryotes like *E. coli* produce acetyl-CoA in the cytosol and do not have other organelles, many industrially-utilized hosts like *S. cerevisiae* are eukaryotes that sequester much of the acetyl-CoA in the mitochondrion, instead of the cytosol. This presents an engineering challenge because most of the metabolic pathways of interest are localized in the cytosol and cytosolic expression of heterologous and native enzymes is more elucidated. Thus, there are a few potential strategies for optimally directing acetyl-CoA into metabolic pathways in eukaryotes: re-targeting the pathway to the mitochondrial matrix, where the pyruvate dehydrogenase converts pyruvate into acetyl-CoA, relocating the pyruvate dehydrogenase to the cytosol, or transporting acetyl-CoA from the mitochondria into the cytosol directly or indirectly. Examples of mitochondrial targeting of heterologous enzymes in yeast exist^{30,31}, but direct transport of acetyl-CoA across the mitochondrial membrane is less clear, although may proceed through the acyl:carnitine shuttle. A natural mechanism for increasing cytosolic acetyl-CoA exists in oleaginous yeasts that rely on nitrogen limitation to allosterically inactivate the isocitrate dehydrogenase via nitrogen scavenging by AMP deaminase activity and cytosolic activity of an ATP:citrate lyase. Further to produce an abundance of lipids, cytosolic malic enzyme converts malate and NADP⁺ into pyruvate and CO₂ and NADPH, therefore providing the

requisite reducing equivalents for fatty acid biosynthesis³²⁻³⁷. *S. cerevisiae* lacks both cytosolic malic enzyme and ATP:citrate lyase (ACL) activity, thus our first attempts to reconstitute the oleaginous yeasts' mechanism for cytosolic acetyl-CoA production were focused on expressing a heterologous ACL. Based on kinetic data reported in the Brenda Enzyme database, we selected an ACL native to *Aspergillus nidulans* and had it codon optimized and synthesized by Epoch Biosciences for expression in *S. cerevisiae*. Expression of the ACL was done in the background strain EPY230 that has been previously optimized for amorphadiene production⁶ where 1.8% galactose and 0.2% glucose were used for growth and production. The resulting amorphadiene production, assayed by GC-MS as previously described⁶, is shown below, where ESY231 (expressing pESC-ACL-ura3) produces roughly 40% more amorphadiene than the control EPY230 strain. These production levels could be further increased by combining cytosolic expression of malic enzyme and limiting nitrogen. Further, the native isocitrate dehydrogenase in *S. cerevisiae* may not be subject to allosteric inhibition to the same extent as homologues in oleaginous yeasts, thus, it may be necessary to replace it as well, or target its degradation after biomass generation.

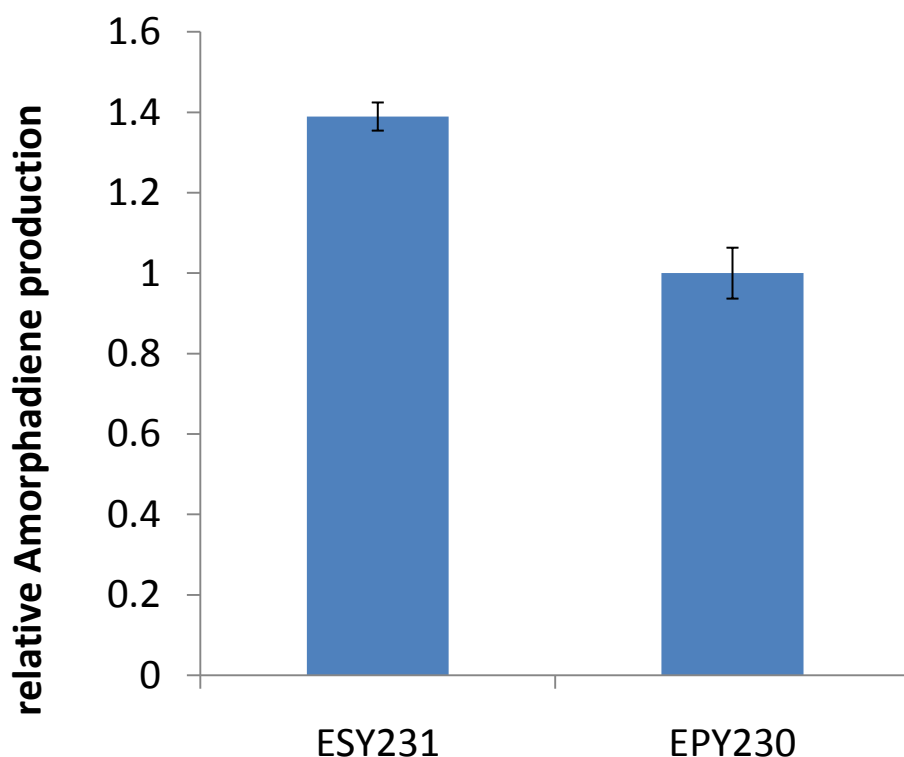


Figure 2-11: Amorphadiene production from ACL-expressing strains. EPY230 was transformed with pESC-ACL-ura3 and strains were assayed for

amorphadiene production. Values and error bars represent the mean and standard deviations of three experiments.

2.4. Conclusions

Here we provide the first demonstration of n-butanol production in *S. cerevisiae* to 2.5 mg/L and tested a variety of isozymes for different reactions in the metabolic pathway. There are a number of obstacles to overcome when expressing heterologous biosynthetic pathways, including enzyme choice, verification of gene expression, balancing of metabolic pathway intermediates and protein levels, and removal of competing pathways, all the while increasing production and maintaining host organism viability. We successfully demonstrated substitution of isozymes and analysis of pathway intermediates and determined potential engineering targets for increasing n-butanol biosynthesis. Comparison of our strain to the native n-butanol producers, Clostridia (~10 g/L), or the recently engineered *E. coli* strains (~0.5 g/L) provides a goal for future n-butanol titer. Given the results presented above and *S. cerevisiae*'s other attributes—inherent tolerance to solvents⁷, widespread use for industrial production of ethanol, and ability to withstand oxygen (as opposed to Clostridia)—*S. cerevisiae* may be an ideal host for industrial n-butanol production. While increases in product titer will certainly be necessary, increases of this magnitude and greater have precedence⁶.

2.5. References

1. Rude, M. A. & Schirmer, A. New microbial fuels: a biotech perspective. *Curr Opin Microbiol* **12**, 274-81 (2009).
2. Fortman, J. L. et al. Biofuel alternatives to ethanol: pumping the microbial well. *Trends Biotechnol* **26**, 375-81 (2008).
3. Lee, S. Y. et al. Fermentative butanol production by Clostridia. *Biotechnol Bioeng* **101**, 209-28 (2008).
4. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-9 (2008).
5. Inui, M. et al. Expression of Clostridium acetobutylicum butanol synthetic genes in Escherichia coli. *Appl Microbiol Biotechnol* **77**, 1305-16 (2008).
6. Ro, D. K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-3 (2006).
7. Fischer, C. R., Klein-Marcuschamer, D. & Stephanopoulos, G. Selection and optimization of microbial hosts for biofuels production. *Metab Eng* **10**, 295-304 (2008).
8. Brachmann, C. B. et al. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115-32 (1998).

9. Burke D, D. D., Stearns T. . Methods in yeast genetics: a Cold Spring Harbor Laboratory Course Manual. Plainview, NY. (2000).
10. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods* **4**, 251-6 (2007).
11. Ro, D. K. et al. Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnol* **8**, 83 (2008).
12. Gietz, R. RA guide to yeast genetics and molecular biology and cell biology. Part B. San Diego, CA: Academic Press, Inc., 87-96 (2002).
13. Shimazu, M., Vetcher, L., Galazzo, J. L., Licari, P. & Santi, D. V. A sensitive and robust method for quantification of intracellular short-chain coenzyme A esters. *Anal Biochem* **328**, 51-9 (2004).
14. Pitera, D. J., Paddon, C. J., Newman, J. D. & Keasling, J. D. Balancing a heterologous mevalonate pathway for improved isoprenoid production in Escherichia coli. *Metab Eng* **9**, 193-207 (2007).
15. Kinter M, S. N. *Protein Sequencing and Identification Using Tandem Mass Spectrometry* (John Wiley and Sons, 2000).
16. Lee, S. Y. & Lee, Y. Metabolic engineering of Escherichia coli for production of enantiomerically pure (R)-(--)-hydroxycarboxylic acids. *Appl Environ Microbiol* **69**, 3421-6 (2003).
17. Martin, V. J., Pitera, D. J., Withers, S. T., Newman, J. D. & Keasling, J. D. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat Biotechnol* **21**, 796-802 (2003).
18. Atsumi, S. et al. Metabolic engineering of Escherichia coli for 1-butanol production. *Metab Eng* **10**, 305-11 (2008).
19. Yoshikuni, Y., Dietrich, J. A., Nowroozi, F. F., Babbitt, P. C. & Keasling, J. D. Redesigning enzymes based on adaptive evolution for optimal function in synthetic metabolic pathways. *Chem Biol* **15**, 607-18 (2008).
20. Chin, J. W., Khankal, R., Monroe, C. A., Maranas, C. D. & Cirino, P. C. Analysis of NADPH supply during xylitol production by engineered Escherichia coli. *Biotechnol Bioeng* **102**, 209-20 (2009).
21. Wilkinson, D. L. & Harrison, R. G. Predicting the solubility of recombinant proteins in Escherichia coli. *Biotechnology (N Y)* **9**, 443-8 (1991).
22. Smialowski, P. et al. Protein solubility: sequence based prediction and experimental verification. *Bioinformatics* **23**, 2536-42 (2007).
23. Magnan, C. N., Randall, A. & Baldi, P. SOLpro: accurate sequence-based prediction of protein solubility. *Bioinformatics* **25**, 2200-7 (2009).
24. Trevino, S. R., Scholtz, J. M. & Pace, C. N. Measuring and increasing protein solubility. *J Pharm Sci* **97**, 4155-66 (2008).
25. Drewke, C., Thielen, J. & Ciriacy, M. Ethanol formation in adh0 mutants reveals the existence of a novel acetaldehyde-reducing activity in Saccharomyces cerevisiae. *J Bacteriol* **172**, 3909-17 (1990).
26. Larroy, C., Fernandez, M. R., Gonzalez, E., Pares, X. & Biosca, J. A. Characterization of the Saccharomyces cerevisiae YMR318C (ADH6) gene product as a broad specificity NADPH-dependent alcohol dehydrogenase: relevance in aldehyde reduction. *Biochem J* **361**, 163-72 (2002).

27. Leskovac, V., Trivic, S. & Pericin, D. The three zinc-containing alcohol dehydrogenases from baker's yeast, *Saccharomyces cerevisiae*. *FEMS Yeast Res* **2**, 481-94 (2002).
28. De Deken, R. H. The Crabtree effect: a regulatory system in yeast. *J Gen Microbiol* **44**, 149-56 (1966).
29. Postma, E., Verduyn, C., Scheffers, W. A. & Van Dijken, J. P. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **55**, 468-77 (1989).
30. Pena-Diaz, J. et al. Mitochondrial localization of the mevalonate pathway enzyme 3-Hydroxy-3-methyl-glutaryl-CoA reductase in the Trypanosomatidae. *Mol Biol Cell* **15**, 1356-63 (2004).
31. Haldi, M. & Guarente, L. N-terminal deletions of a mitochondrial signal sequence in yeast. Targeting information of delta-aminolevulinate synthase is encoded in non-overlapping regions. *J Biol Chem* **264**, 17107-12 (1989).
32. Ratledge, C. Single cell oils--have they a biotechnological future? *Trends Biotechnol* **11**, 278-84 (1993).
33. Ratledge, C. Microbial routes to lipids. *Biochem Soc Trans* **17**, 1139-41 (1989).
34. Hall, M. J. & Ratledge, C. Lipid Accumulation in an Oleaginous Yeast (*Candida 107*) Growing on Glucose Under Various Conditions in a One- and Two-Stage Continuous Culture. *Appl Environ Microbiol* **33**, 577-584 (1977).
35. Botham, P. A. & Ratledge, C. A biochemical explanation for lipid accumulation in *Candida 107* and other oleaginous micro-organisms. *J Gen Microbiol* **114**, 361-75 (1979).
36. Gill, C. O., Hall, M. J. & Ratledge, C. Lipid accumulation in an oleaginous yeast (*Candida 107*) growing on glucose in single-stage continuous culture. *Appl Environ Microbiol* **33**, 231-9 (1977).
37. Zhang Y, R. C., et al. Malic enzyme: the controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. *Microbiology* **153**, 2013-2025 (2007).

Chapter 3: Engineering *E. coli* for the production of fatty acid-derived chemicals and fuels.

3.1. Introduction

Increasing energy costs and environmental concerns have emphasized the need to sustainably produce renewable fuels and chemicals¹ and have focused upon shorter chain alcohols including ethanol², butanol³, isobutanol⁴, isopentanol⁵, etc. Although production of a range of specific fuels will be necessary to supplement the complex nature of petroleum-derived fuels, major efforts to this end have also focused on the microbial production of high-energy fuels through cost effective “consolidated bioprocesses”⁶. Fatty acids are composed of long alky chains and represent nature’s “petroleum,” being a primary metabolite used by cells for both chemical and energy storage functions. Some bacteria and eukaryotes are known to store fatty acid or fatty acid – like molecules, including neutral lipids, poly-hydroxyalkanoates and wax esters. These energy rich molecules are today isolated from plant and animal oils for a diverse set of products ranging from fuels to oleochemicals. A more scalable, controllable, and economic route to this important class of chemicals would be through the microbial conversion of renewable feedstocks, such as biomass-derived carbohydrates. Here we demonstrate the engineering of *E. coli* to produce structurally tailored fatty esters (biodiesel), fatty alcohols, and waxes directly from simple sugars and the further engineering of the biodiesel-producing cells to secrete hemicellulases, a step toward producing these compounds directly from hemicellulose, a major component of plant-derived biomass.

Fuels and chemicals have been produced from the fatty acids of plant and animal oils for over a century. Today these oils are the raw materials for a growing diversity of products including biodiesel, “renewable diesel,” surfactants, solvents, and lubricants. The increased demand and limited supply of these oils has resulted in competition with food, higher prices, questionable land use practices and environmental concerns associated with their production⁷. A sustainable alternative is to produce these products directly from abundant and cost effective renewable resources by fermentation. The well-studied industrial microorganism, *Escherichia coli*, is ideally suited for this purpose. *E. coli* is approximately 9.7% lipid, produces fatty acid metabolites at the commercial productivity of 0.2g/L/hr per gram of cell mass just to grow, can achieve product-dependent mass yields of 30-35%⁸, and is exceptionally amenable to genetic manipulation. Further, *E. coli* does not naturally store fatty acids for energy, a property that may be better suited for production and secretion of fatty acid-derived products. Combining this natural fatty acid

synthetic ability with new biochemical reactions realized through synthetic biology has provided a means to divert fatty acid metabolism directly toward fuel and chemical products of interest (Figure 1).

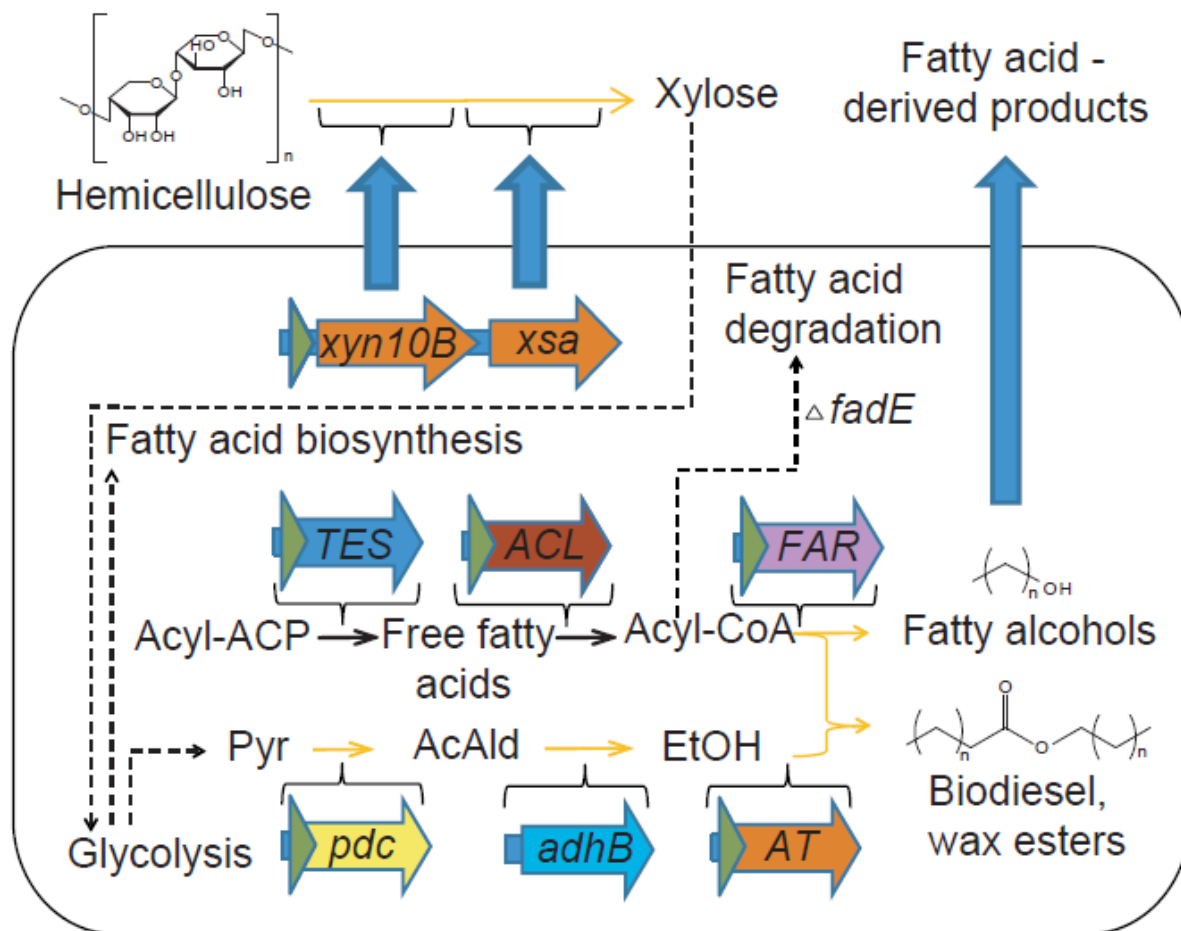


Figure 3-1: Engineered pathways for production of fatty acid-derived molecules from hemicelluloses or glucose and depiction of the synthetic operons used in this study. Flux through the *E. coli* fatty acid pathway (black lines) was increased to improve production of free fatty acids and acyl-CoAs by eliminating β -oxidation (knockouts are *fadE*), by overexpressing thioesterases (TES) and acyl-CoA ligases (ACL). Various products were produced from non-native pathways (orange lines) including biodiesel, alcohols and wax esters. Alcohols were produced directly from fatty acyl-CoAs by overexpressing fatty acyl-CoA reductases, (FAR); the esters were produced by expressing an acyltransferase (AT) in conjunction with an alcohol-forming pathway; biodiesel was produced by introduction of an ethanol pathway (*pdc* and *adhB*) and wax esters were produced from the fatty alcohol pathway (FAR). Finally, expressing and secreting xylanases (*xyn10B* and *xsa*) allowed for the utilization of

hemicellulose. Overexpressed genes or operons are indicated; green triangles represent the lacUV5 promoter. Abbreviations: pyruvate (pyr), acetaldehyde (AcAld), ethanol (EtOH).

Fatty acid biosynthesis. There are two types of fatty acid synthases, Type I, found in eukaryotes and Type II, found in prokaryotes. The main difference between the prokaryotic and eukaryotic is the enzyme composition – the eukaryotic system utilizes *three* multi-subunit genes comprising *eleven* catalytic domains while the prokaryotic system utilizes *eleven* genes for each of the *eleven* catalytic domains.

Fatty acid biosynthesis in *E. coli* has been investigated for many years and much has been learned about its reaction pathway and the enzymes encoding it, and the regulation of the enzymes. This understanding has been crucial to engineering the native pathway. Excellent reviews of fatty acid metabolism is discussed elsewhere^{9,10}, but will also be discussed herein. *E. coli* fatty acids are produced by the Type II fatty acid enzymes that are encoded by separate genes, in comparison to the Type I systems that exist mainly in higher order organisms and are encoded by single genes that have multiple enzymatic functions. Fatty acid biosynthesis is initiated by carboxylation of acetyl-CoA and bicarbonate to form malonyl-CoA by the heterotetrameric enzyme, acetyl-CoA carboxylase, which is encoded by *accA*, *accB*, *accC*, and *accD*. The malonate from malonyl-CoA is transferred to an acyl-carrier protein (ACP) by a malonyl-ACP transferase, encoded by *fabD*. This complex serves as the reaction center for the remaining condensation and reduction cycle. Acetyl-CoA is then condensed with malonyl-ACP, an unfavorable reaction that is driven by a coupled decarboxylation of malonate, therefore extending the acyl chain by two carbons with every condensation by the b-ketoacyl-ACP synthase, encoded by *fabH*. The b-ketoacyl-ACP is then dehydrated to form a trans-2-enoyl-ACP by a b-ketoacyl-ACP dehydratase encoded by *fabZ*. Alternatively, unsaturated fatty acid biosynthesis proceeds at this branch point by dehydration and isomerization of 3-hydroxyacyl-ACP by an enzyme encoded by *fabA*. The trans-2-enoyl-ACP is reduced to form an acyl-ACP using NAD(P)H as a cofactor by the enoyl-ACP reductase encoded by *fabI*. This condensation and reduction cycle continues until a final acyl-ACP product with the majority of fatty acids ranging in chain length from C16-C18 then directed to cellular components such as structural or storage lipids^{9,11}. The accumulation of fatty acyl-ACP feedback inhibits fatty acid biosynthesis by classic allosteric inhibition such that *E. coli* does not accumulate or produce more fatty acids than are necessary to grow

and divide. Specifically, acyl-ACPs have been shown to inhibit the acetyl-CoA carboxylase, the b-ketoacyl-ACP synthase, and the enoyl-ACP reductase¹²⁻¹⁴.

In addition to allosteric regulation, *E. coli* regulates fatty acid biosynthesis at the transcriptional level. Two main transcriptional activators responsible for fatty acid regulation are *fadR*¹⁵ and *fabR*. FadR, or **fatty acid degradation regulon**, is a dual activator and repressor that inhibits fatty acid degradation and activates unsaturated fatty acid biosynthesis. Inhibition of fatty acid degradation is exerted by inhibiting transcription of the b-oxidation genes, including *fadD*, *fadE*, *fadI*, *fadJ*, and *fadL*^{16,17}. FadR binding sites have been found 5' of *uspA*, a gene encoding a global stress response regulator, which is implicated in survival of *E. coli* in severe stress conditions and stationary phase. Further, FadR inhibits acetate metabolism and the glyoxylate shunt^{18,19}. Activation of unsaturated fatty acid biosynthesis occurs via activation of transcription of *fabA* and *fabB*. The mechanism of FadR – DNA binding is regulated by the presence of acyl-CoA, which binds FadR and causes dissociation from DNA^{20,21}.

FabR, or **fatty acid biosynthesis repressor**, represses the transcription of *fabA* and *fabB*, which encode genes responsible for unsaturated fatty acid biosynthesis²¹⁻²⁵. In a similar fashion to FadR, FabR binds acyl-CoA and acyl-ACPs to exert its repression. FabR controls membrane fluidity and the ratio of saturated to unsaturated fatty acids by differential binding and repression due to the acyl-thioester pool, where saturated acyl-thioesters compete with unsaturated acyl-thioesters and the lower FabR-saturated acyl-thioester DNA binding allows for *fabBA* transcription and an increase in unsaturated fatty acid biosynthesis²².

Two important discoveries were made in relation to de-regulating the feedback exerted by acyl-ACP in relation to fatty acyl-ACP thioesterases that cleave the fatty acid from the ACP. First, Cronan and colleagues demonstrated that expression of a cytoplasmic thioesterase results in hydrolysis of these acyl-ACPs, deregulation of fatty acid biosynthesis, and overproduction and secretion of significant levels of free fatty acids^{11,26}. Second, Davies and colleagues demonstrated that expression of heterologous thioesterases found in plants could redirect fatty acid biosynthesis for the production of fatty acids with different chain lengths^{27,28}. Thus, these two important findings suggested that type II fatty acid biosynthesis of *E. coli* could be exploited for the production of not only fatty acids, but also more valuable fatty acid-derived products.

3.2. Materials and methods

Strains and plasmids. *E. coli* strains DH1 and C41 (DE3) were utilized as the wild-type strain for all studies, except where indicated. Knockouts of *fadD* and

fadE, in DH1 were performed as previously described²⁹. A list of strains and plasmids constructed is given in Table 1 and a list of genes and corresponding accession numbers used is given in Table 2. Construction of plasmids was carried out with standard molecular biology methods and is described in detail in Methods.

Growth and production parameters. Strains were cultivated in baffled flasks at 37°C in M9 minimal medium supplemented with trace elements and the appropriate antibiotics (50 mg/L ampicillin, 20 mg/L chloramphenicol, 5 mg/L tetracycline, 100 mg/L carbenicillin, 100 mg/L spectinomycin, and 50 mg/L kanamycin). Pathway induction was achieved by addition of 1 mM IPTG at an optical density of 0.5 – 1 measured at a wavelength of 600 nm (OD₆₀₀). For the thioesterase expression studies, 2% ethanol was added and the post-induction temperature was 25°C for production of FAEEs. Xylan media was prepared by addition of 2% beechwood xylan and 0.2% glucose to M9 minimal media. Metabolites were identified and quantified by gas chromatography-mass spectrometry, further described in Methods.

Methods

Reagents. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and include fatty acid methyl ester standards, fatty acid ethyl ester standards, fatty aldehyde standards, fatty alcohol standards, and beechwood xylan.

Plasmid and strain construction. For the study in which various thioesterases were used, *E. coli* C41 (DE3) with a deletion of the acyl-CoA dehydrogenase (*ΔfadE*) was used as an expression strain. The *fadE* deletion strain of *E. coli* C41 (DE3) was constructed using the temperature sensitive plasmid pKO3 (a gift from George M. Church of Harvard Medical School, Harvard University, and Boston, MA). DNA sequences downstream and upstream of *fadE* were amplified with the downstream primers, LF_*NotI* (5'-CAACCAGCGGCCGCGCGACGAAGCTGCCGCTTC) and LR (5'-CCTACAAGTAAGGGGCTTTTCGTTATGAATAACGGAGCCGAAAGGCTC) and the upstream primers, RF1 (5'-CTTTCGGCTCCGTTATTCATAACGAAAAGCCCCTTACTTGTAGGAGG) and RR_*Bam* (5'-CCAGGATCCAGGTCGGATGCGGCGTGAAC), using genomic DNA from C41 (DE3) as the template. The two PCR products were mixed and joined together by PCR amplification with primers, LF_*NotI* and RR_*Bam* described above. This PCR product (~1.8 kb) was digested with *NotI* and *BamHI* and ligated between the *NotI* and *BamHI* sites of pKO3. The pKO3 derivative was used to construct the *fadE* in-frame deletion according to Link et al³⁰.

E. coli DH10B and DH5 α were used for bacterial transformation and plasmid amplification in the construction of the expression plasmids used in this study. Native *E. coli* genes were cloned from DH1. *mFAR1* (*Mus musculus*, GenBank Accession BC007178)³¹ was synthesized and codon optimized for *E. coli* expression (Epoch biolabs). *atfA* (*Acinetobacter* sp. strain ADP1) was synthesized (Epoch biolabs). *pdC* and *adhB* were cloned from *Z. mobilis* genomic DNA (ATCC 31821). *FAA2* was cloned from *Saccharomyces cerevisiae* (BY4742) genomic DNA. DNA encoding the catalytic domain of *xynB* from *C. stercorearium* NCIMB 11754 was synthesised and codon-optimized by DNA 2.0. *E. coli* OsmY gene was synthesized with the native sequence by DNA 2.0. Full-length *xsa* gene was codon-optimized and synthesized in-house³². A BglII site followed by a strong RBS was appended to the OsmY gene, which ended with a glycine-serine linker, followed by a BamHI site. *xynB* and *xsa* were cloned with a BglII site on their 5' end and BamHI-TAA-XhoI sites on their 3' ends. Gene fusions were constructed by digestion of the OsmY-carrying plasmid with BamHI/XhoI and ligation with the BglII/XhoI-cut inserts (*xynB* and *xsa*). The plasmid bearing the OsmY-Xyn10B fusion was subsequently re-cut with BamHI/XhoI, and OsmY-Xsa insert cut with BglII/XhoI was ligated. To generate biodiesel from xylan, the OsmY-XynB, OsmY-Xsa bicistronic gene was cut with EcoRI/BamHI and ligated in front of pKS104 cut with EcoRI/BglII. Prior to this the BglII site within *fadD* was removed with Quickchange mutagenesis. Other plasmids were constructed using the “Sequence and Ligation Independent Cloning” (SLIC) method³³ or standard methods³⁴. All genes were over-expressed under the control of the IPTG-inducible lacUV5 or trc promoters as indicated; LS9 vectors use the T7 promoters (pDuet vectors). Cloning and overexpression of *fadD* from *E. coli* was achieved only with a mutated version and is discussed further in the results section.

For the thioesterase studies, pDuet vectors were used for plasmid based-expression of wax synthase (*AtfA*), fatty acyl-CoA reductase (*Acr1*), acyl-CoA synthetase (*FadD*) and the thioesterases *tesA* and *FatA*. All other thioesterases were expressed from pMAL-c2x (New England Biolabs). The wax synthase gene (*atfA*) from *Acinetobacter baylyi* ADP1 was amplified with primer *adp1ws_NdeI* (5'-TCATATGGCGCCCATACATCCG) and *adp1ws_AvrII* (5'-TCCTAGGAGGGCTAATTTAGCCCTTTAGTT). After amplification, the PCR product was digested with *NdeI* and *AvrII* (underlined sites) and ligated with pCOLADuet-1 cut with *NdeI* and *AvrII* to produce pLS9-*atfA*. The *fadD* gene of *E. coli* was amplified with forward primer (5'-CCATGGTGAAGAAG-GTTTGGCTTAA) and reverse primer (5'-AAGCTTTCAGGCTTTATTGTCCAC), using genomic DNA of *E. coli* strain XL-Blue (Stratagene, La Jolla, CA) as a template. The PCR product was digested

with *Nco*I and *Hind*III and ligated with pCDFDuet-1 linearized with *Nco*I and *Hind*III. The resulting plasmid, pLS9-*fadD*, was digested with *Not*I and *Avr*II, and the fragment was ligated with the *Not*I/*Avr*II fragment from pLS9-*atfA* to generate plasmid pLS9-*atfA-fadD*. The *Nco*I and *Hind*III fragment from pLS9-*fadD* was also cloned into *Nco*I and *Hind*III digested pLS9-*acr1* generating plasmid pLS9-*acr1-fadD*.

Metabolite Analysis. **Free fatty acids** were extracted from 5 mL cultures by addition of 500 μ L HCl and 5 mL of ethyl acetate, spiked with 10 mg/L of methyl nonadecanoate as an internal standard. The culture tubes were vortexed for 15 s followed by shaking at 200 rpm for 20 min. The organic layer was separated and a second extraction was performed by addition of another 5 mL ethyl acetate to the culture tubes. The free fatty acids were then converted to methyl esters by addition of 200 μ L TMS-diazomethane, 10 μ L HCl, and 90 μ L MeOH³⁵. This reaction was allowed to proceed for 2 hr and then was applied to a Thermo Trace Ultra gas chromatograph (GC) equipped with a Triplus AS autosampler and a TR-WAXMS column (Thermo Scientific). The GC program was as follows: initial temperature of 40°C for 1.2 min, ramped to 220°C at 30°C/min and held for 3 min. Final quantification analysis was performed with Xcalibur software.

Fatty acid ethyl esters (FAEEs), fatty alcohols, and wax esters were extracted from cultures by addition of 10% (v/v) ethyl acetate, spiked with 10 mg/L methyl nonadecanoate, followed by shaking at 200 rpm for 20 min. Analysis of FAEEs was performed on an HP 6890 Series GC with an Agilent 5973 Network MSD equipped with a DB5 column (Thermo). The GC program was the same as for quantifying FAMES. Fatty alcohols and aldehydes were separated with a TR-Wax column (Agilent). The GC program was as follows: initial temperature of 70°C, held for 1 min, ramped to 240°C at 25°C/min and held for 3 min. BSTFA derivatization of fatty alcohols aided in their identification and separation. Where described, a 10% (v/v) dodecane overlay was added to cultures at the time of induction. 20 μ L of dodecane was sampled and diluted 25 times in 480 μ L ethyl acetate.

Ethanol was measured by sampling 1 mL of culture, centrifuging 18,000 $\times$ g, 5 min, and applying the supernatant to an Agilent 1100 series HPLC equipped with an Aminex HPX-87H ion exchange column (Biorad). The solvent, 4 mM H₂SO₄, flow rate was 0.6 mL/min and the column was maintained at 50°C. All metabolites were detected with an Agilent 1200 series DAD and RID detectors.

Table 3-1: Strains and plasmids.

Plasmids	Replication Origin	Overexpressed Genes	Resistance	Reference
pACYC184	p15a		Cam	36
pBBR1MCS-3	pBBR		Tet	36
pBR322	pBR322		Amp	36
pKS1	p15a	<i>P_{lacUV5}: 'tesA</i>	Cam	This study
pKS2	pBBR	<i>P_{lacUV5}: 'tesA</i>	Tet	This study
pKS5	p15a	<i>P_{lacUV5}: mFar1</i>	Cam	This study
pKS13	pBBR	<i>P_{lacUV5}: pdc, adhB</i>	Tet	This study
pKS17	pBBR	<i>P_{lacUV5}: pdc, adhB, atfA</i>	Tet	This study
pKS18	pBBR	<i>P_{lacUV5}: mFar1</i>	Tet	This study
pKS100	pBR322	<i>P_{TRC}: fadD</i>	Amp	This study
pKS101	pBR322	<i>P_{TRC}: FAA2</i>	Amp	This study
pKS104	ColE1	<i>P_{lacUV5}: fadD (M335I), atfA</i>	Amp	This study
pGB1	pBR322	<i>P_{Pro}: xyn10B, xsa</i>	Amp	This study
pGB2	pBR322	<i>P_{lacUV5}: fadD (M335I), xyn10B, xsa</i>	Amp	This study
pGB3	pBR322	<i>P_{Pro}: xynB</i>	Tet	This study
pETDuet-1				
pMAL c2X				
pCDFDuet-1				
pCOLADuet-1				
pLS9-' <i>tesA</i>	pETDuet-1	<i>P_{T7}: 'tesA</i>		This study
pLS9-ucfatB1	pMAL c2X	<i>P_{T7}: ucfatB1</i>		This study
pLS9-chfatB2	pMAL c2X	<i>P_{T7}: ucfatB2</i>		This study
pLS9-chfatB3	pMAL c2X	<i>P_{T7}: ucfatB3</i>		This study
pLS9-atfAtA3	pETDuet-1	<i>P_{T7}: atfAtA3</i>		This study
pLS9-acr1	pCDFDuet-1	<i>P_{T7}: acr1</i>		This study
pLS9-atfA	pCOLADuet-1	<i>P_{T7}: atfA</i>		This study
pLS9-acr1- <i>fadD</i>	pCDFDuet-1	<i>P_{T7}: acr1 fadD</i>		This study
pLS9-aftA- <i>fadD</i>	pCOLADuet-1	<i>P_{T7}: atfA fadD</i>		This study
Strains	Relevant Genotype	Reference		
LT	DH1: pKS1	This study		

LT-ΔfadD	DH1: Δ fadD, pKS1	This study
LT-ΔfadE	DH1: Δ fadE, pKS1	This study
DE-EtOH	DH1: Δ fadE, pKS13	This study
LAAP	DH1: Δ fadE, pKS1, pKS17	This study
faa2	DH1: Δ fadE, pKS1, pKS17, pKS101	This study
A1A	DH1: Δ fadE, pKS1, pKS13, pKS104	This study
A2A	DH1: Δ fadE, pKS1, pKS17, pKS104	This study
KS5	DH1: Δ fadE, pKS5	This study
KS12	DH1: Δ fadE, pKS2, pKS5, pKS104	This study
GB-X	BL21: pGB3	This study
GB-XX	BL21: pGB1	
HE-GH	DH1: Δ fadE, pKS1, pKS17, pGB2	This study
C41 (DE3)		³⁷
C41 (DE3, ΔfadE)	C41: Δ fadE	This study

Gene description	Host Organism	Accession Number
<i>tesA</i> , thioesterase without leader sequence	<i>E. coli</i>	AAC73596
<i>fadD</i> , CoA ligase	<i>E. coli</i>	NP416319
<i>faa2</i> , CoA ligase	<i>S. cerevisiae</i>	NP010931.1
<i>mFar1</i> , acyl CoA reductase	<i>M. musculus</i>	BC007178
<i>acr1</i> , acyl CoA reductase	<i>A. calcoaceticus</i>	AAC45217
<i>atfA</i> , wax ester synthase	<i>A. baylyi</i> ADP1	AAO17391
<i>pdc</i> , pyruvate decarboxylase	<i>Z. mobilis</i> ZM4	YP_163095.1
<i>adhB</i> , alcohol dehydrogenase	<i>Z. mobilis</i> ZM4	YP_163331
<i>xyn10B</i> , endoxylanase	<i>C.</i>	AJ508407

	<i>stercorarium</i>	
<i>xsa, xylanase</i>	<i>B. ovatus</i>	AAB08024
<i>ucfatB1, thioesterase</i>	<i>U. californica</i>	Q41635
<i>chfatB2, thioesterase</i>	<i>C. hookeriana</i>	Q39513
<i>chfatB3, thioesterase</i>	<i>C. hookeriana</i>	AAC49269
<i>atfAtA3, thioesterase</i>	<i>A. thaliana</i>	NP189147

3.3. Results

3.3.1. Cytosolic expression of the native *E. coli* thioesterase, *tesA*.

By cytosolic expression of a native *E. coli* thioesterase (*tesA*), normally localized to the periplasm, we demonstrate free fatty acid production of ~0.32 g/L, similar to previous findings^{11,38}, (Figure 2). *tesA* exhibits preference for C14 fatty acyl-ACPs, although a range of free fatty acids (C8 to C18) are observed when *tesA* is produced (Figure 3). The length of the fatty acid chain produced can be controlled by expressing alternate thioesterases from plants³⁹. In order to further improve production of free fatty acids, we eliminated the first two competing enzymes associated with β -oxidation, *FadD* and *FadE*, resulting in an additional three- to four-fold increase in titer to ~1.2 g/L. *tesA- Δ fadE* effected a 6% yield of fatty acids from 2% glucose in shake flasks, 14% of the theoretical limit.

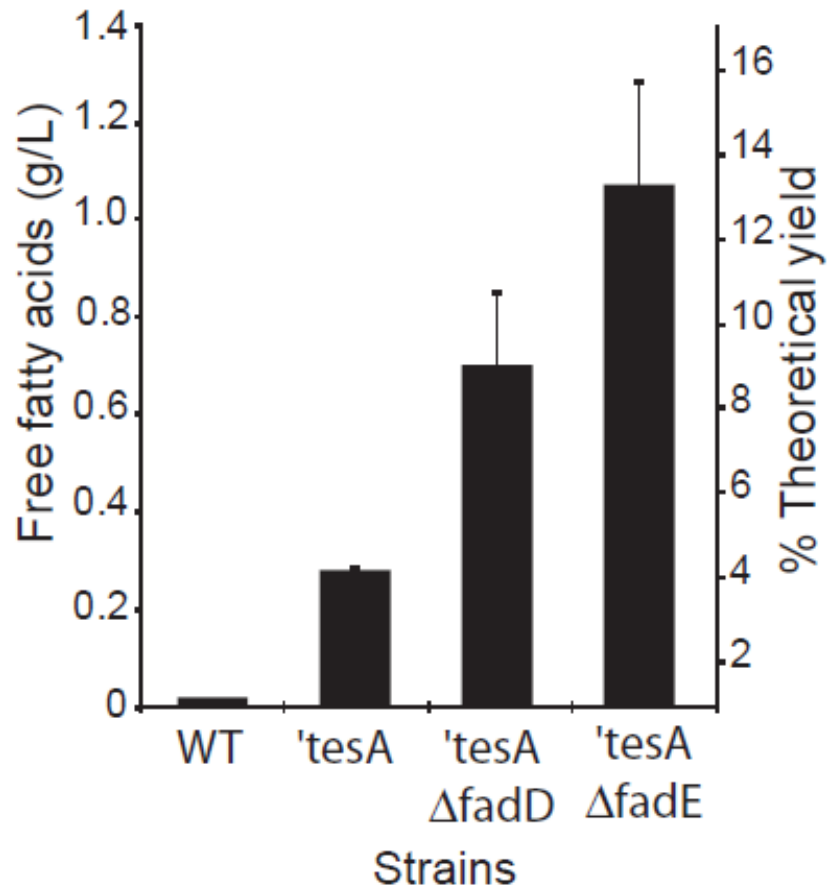


Figure 3-2: Total free fatty acid production and respective theoretical yield by engineered *E. coli* strains. Overexpressed and knocked out genes are indicated. WT: wild-type DH1; LT: '*tesA*'; LT- Δ *fadD*: LT Δ *fadD*, LT- Δ *fadE*: LT Δ *fadE*. Values and error bars represent the mean and standard deviation of triplicate experiments.

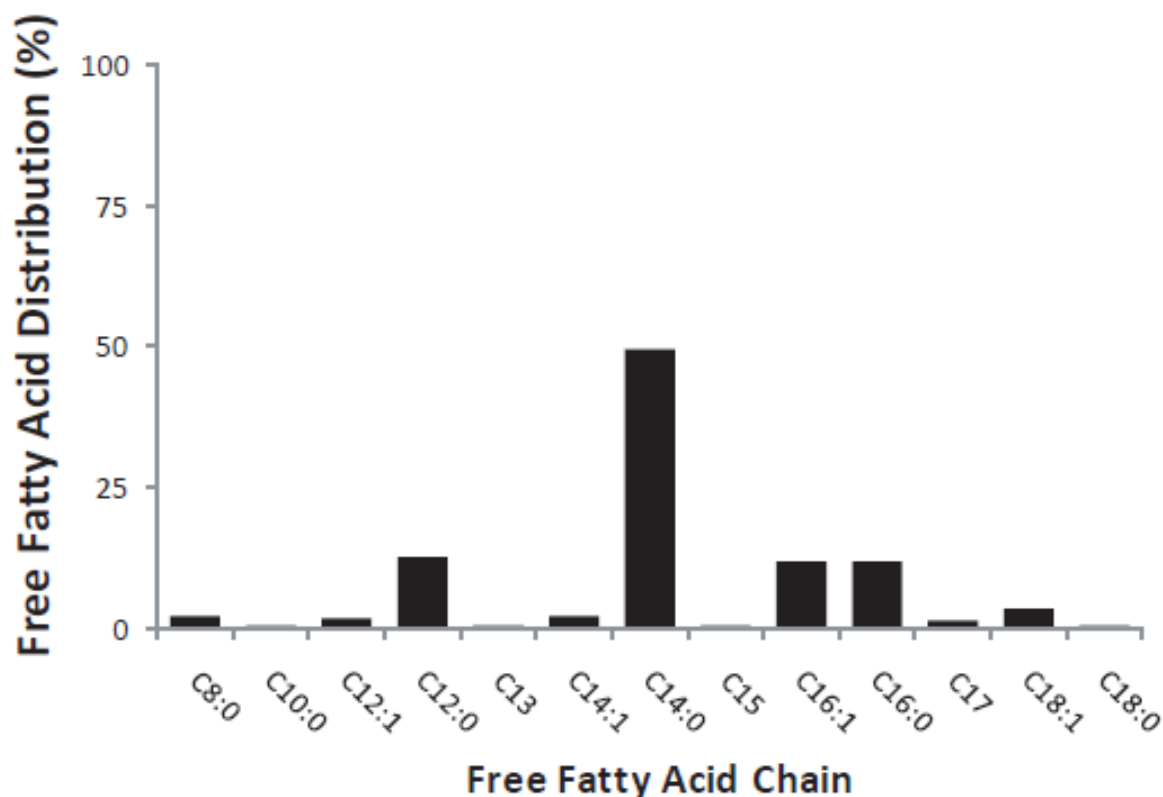


Figure 3-3: Fatty acid chain length distribution in *E. coli* DH1 *fadD* knockout expressing *tesA*.

3.3.2. Co-expression of alternate thioesterases and fatty acid modifying enzymes.

Although free fatty acids are valuable, directly-compatible renewable fuels and chemicals are in greater demand. Microbial production of biofuels has focused on gasoline supplements^{3,4}, however, diesel is globally in greatest demand, and has a growth rate three times that of the gasoline market⁴⁰. Biodiesel, the primary renewable alternative to diesel, is consumed at greater than 2 billion gal per year and is composed of fatty acid methyl and ethyl esters (FAMES and FAEEs) derived from the chemical transesterification of plant and animal oils⁷. Previously, *E. coli* was engineered to produce FAEE by esterifying exogenously-added fatty acids with endogenously-produced ethanol,⁴¹ however, fatty acids are not a commercially viable feedstock. Thus, we engineered *E. coli* to directly produce FAEE from glucose and ethanol. Expression of the gene encoding a wax-ester synthase (*atfA*) with expression of *fadD* and *tesA* and addition of ethanol (to 2%) resulted in production of approximately 400 mg/L of FAEEs in 48 hours, with the composition of FAEEs ranging from C12 to C18. Overexpression of *fadD*, the

first step for fatty acid *degradation*, and deletion of *fadE*, the second step in fatty acid degradation, dramatically increases FAEE production as well as other fatty acid products as described below (Figure 4a).

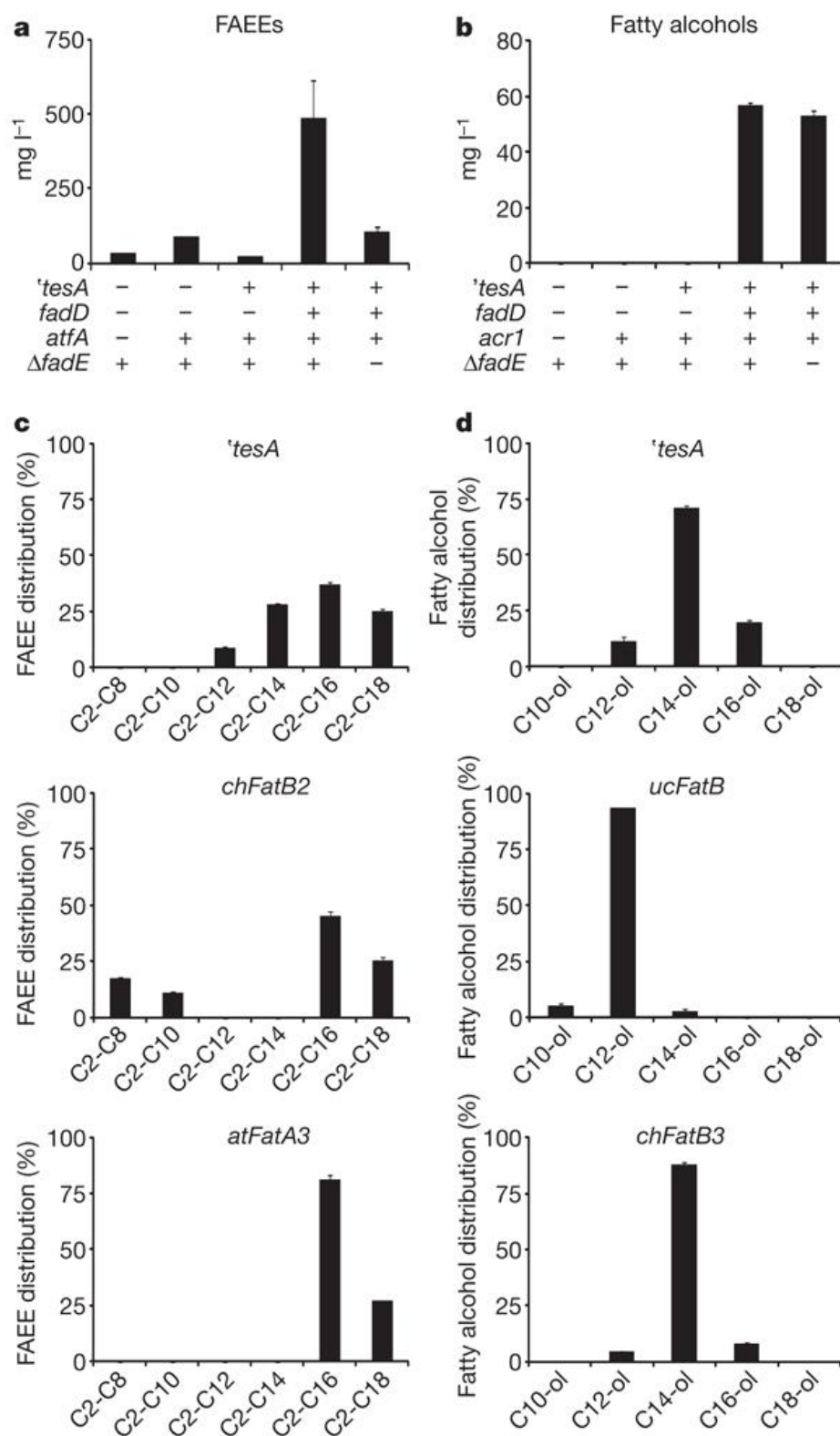


Figure 3-4: Engineered production of FAEs and fatty alcohols with controlled chain length. Production of FAEs (a) or fatty alcohols (b) by overexpressing 'tesA, fadD, and atfA or acr1 in C41(DE3) Δ fadE. c & d, The

chain length distribution of FAEEs fatty alcohols was varied by expressing thioesterases with different substrate specificities. Shown are the percent composition of each chain length FAEEs or fatty alcohols resulting from the expression of thioesterase genes “*tesA*, *chFatB2*, *atfAtA3*, *ucFatB*, and *chFatB3*, in concert with *fadD* and an ester synthase (*atfA*) or acyl-CoA reductase (*acr1*) in C41(DE3) Δ *fadE*. Values and error bars represent the mean and standard deviation of triplicate experiments.

In addition to esters, there is a large market for fatty alcohols, aldehydes, and wax esters, which are used predominantly in soaps, detergents, cosmetic additives, pheromones, and flavoring compounds, and potentially as biofuels; their value was approximately \$1500/ton (2004 ICIS pricing), and represent a ~ \$3B market⁴². Fatty alcohols are produced through hydrogenation of plant oil-derived FAMES or through synthesis from petrochemical precursors. Previous identification and expression of fatty alcohol-forming fatty acyl-CoA reductases, such as that encoded by *acr1* from *Acinetobacter calcoaceticus* BD413, have been described^{31,43,44}. Expressing *acr1* in place of *atfA* (used in the FAEE-producing strains) resulted in production of medium chain fatty alcohols up to ~60 mg/L (Figure 4b). While *FadD* improved production of all products, the lower level of alcohol as compared to FAEEs (400 mg/L) suggests *Acr1* may be limiting in this pathway.

Fatty acid chain length and saturation directly impact fuel and chemical properties, such as cetane number and melting point. Knothe et al. demonstrated that the poor performance of plant-derived biodiesel can be improved by tailoring its fatty ester composition⁴⁵. To engineer FAEE and fatty alcohol chain length composition, “*tesA* in the FAEE and fatty alcohol strains was replaced with plant genes encoding thioesterases previously shown to prefer different chain length fatty acyl-ACPs³⁹. This single genetic manipulation directly affected the composition of the fatty acids shunted into these pathways and resulted in controlled FAEE and fatty alcohol chain length distributions (Figure 4c, d). This genetic tool provides a means to easily tailor the composition, and hence the performance, of the fuel and chemical product being produced.

3.3.3. Co-expression of ethanol and fatty acid pathways

A benefit of microbial catalysis is the ability to genetically compile multiple, complex, biosynthetic pathways into a single cell, simplifying process and raw material requirements, and decreasing costs. To obviate the need to feed ethanol to produce FAEE, we engineered our FAEE-producing strain to co-produce ethanol

by expressing the *Zymomonas mobilis* genes *pdc* and *adhB*, which encode pyruvate decarboxylase and alcohol dehydrogenase. *E. coli* expressing *pdc* and *adhB* alone produced ~5 g/L ethanol after 24 h, as reported previously². Addition of the minimal FAEE biosynthetic pathway (‘‘*tesA* and *atfA*’) into the ethanol-producing strain resulted in low production of FAEEs to 37 mg/L (strain HE-LAAP) (Figure 5) as expected, since the native acyl-CoA ligase (*fadD*) activity was limiting. Overexpression of *FAA2*, an acyl-CoA ligase from *S. cerevisiae*, resulted in an approximately 2.5-fold increase in FAEE production to 96 mg/L (strain *faa2*). Another 2.5-fold increase to 233 mg/L was achieved by overexpression of an *E. coli*-derived *fadD* (strain A1A). Expressing an additional copy of *atfA* resulted in production of 427 mg/L FAEEs (strain A2A). Finally, by overlaying an organic phase of dodecane to potentially prevent FAEE evaporation, the engineered *E. coli* produced FAEE at 674 mg/L, which is 9.4% of the theoretical yield. FAEEs were not growth inhibitory to *E. coli* up to 100 g/L (Supplementary Figure 2). By combining the pathways for EtOH and FAEE biosynthesis, FAEE was produced efficiently from glucose as the sole carbon source.

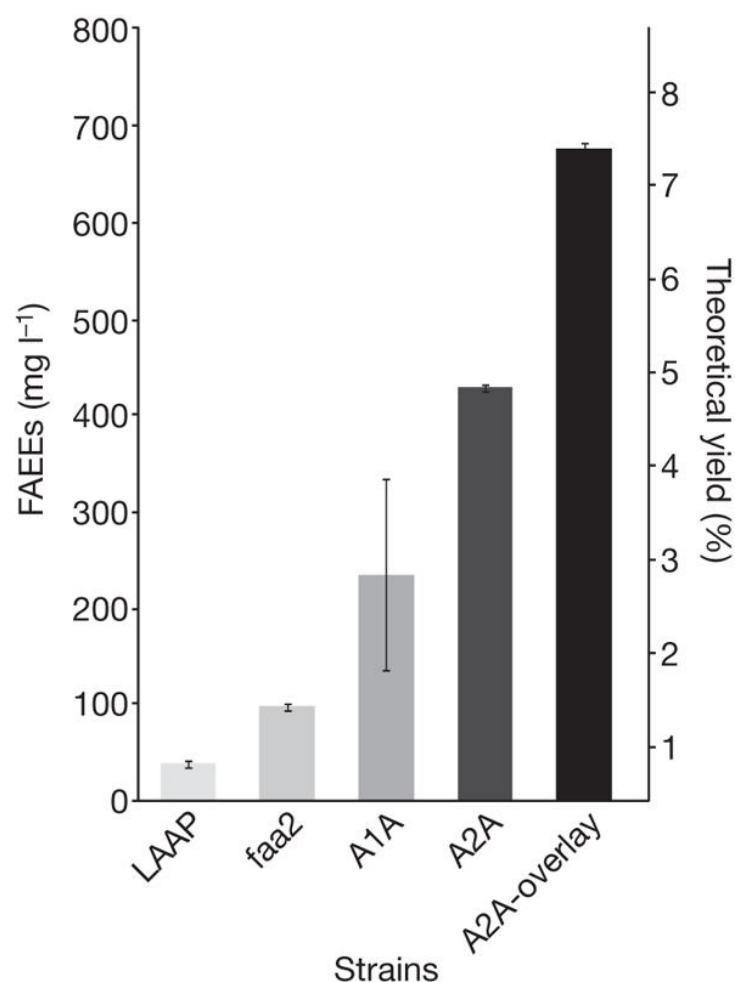


Figure 3-5: Toward a single cell catalyst: Biodiesel (FAEE) production by various strains without exogenous ethanol supplementation. HE-LAAP: DH1, $\Delta fadE$, $\Delta tesA$, $\Delta atfA$, Δpdc , $\Delta adhB$; *faa2*: HE-LAAP *FAA2*; A1A: HE-LAAP *fadD*; A2A: A1A harboring an additional copy of *atfA*. Values and error bars represent the mean and standard deviation of triplicate experiments.

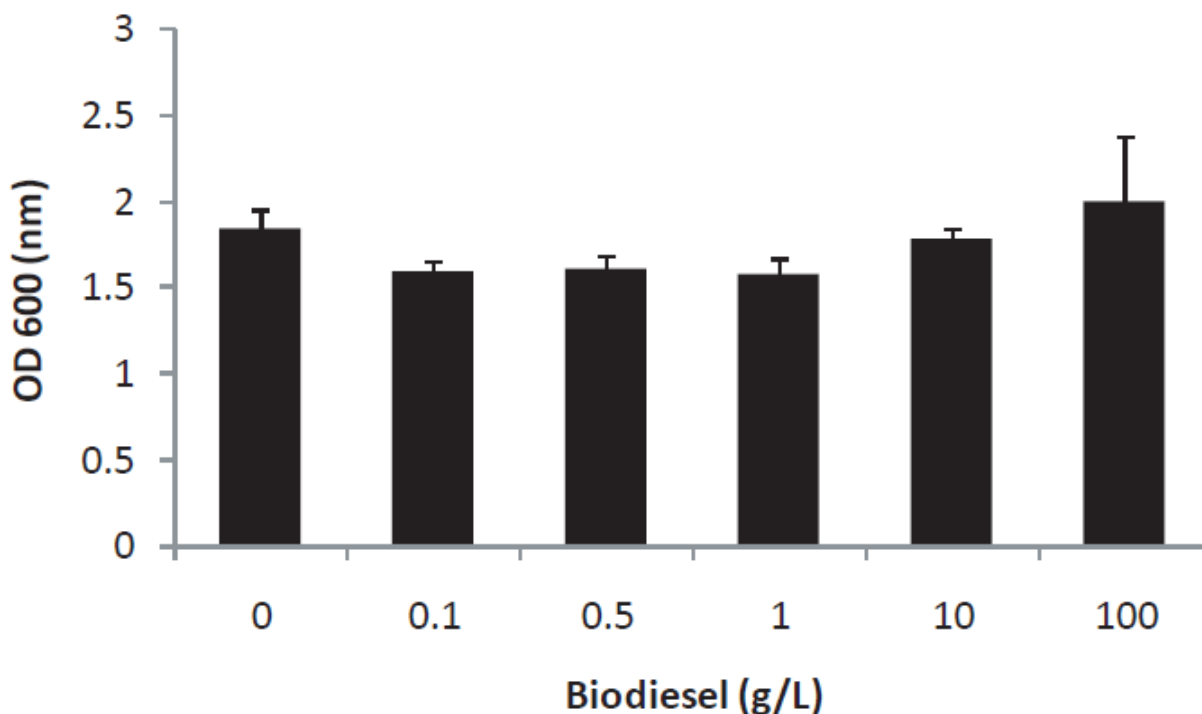


Figure 3-6: Growth measurements from *E. coli* DH1 *fadE* knockout strain in various concentrations of methyl oleate (biodiesel). Strains were grown in the given concentrations of biodiesel and OD 600 was measured after 24h of growth. Values and error bars represent the mean and standard deviation of triplicate cultures.

3.3.4. Coexpression of fatty acid and long chain alcohol pathway.

In a similar fashion, an *E. coli* strain (strain KS12) engineered to express *tesA*, *fadD*, *mFar1*, and *atfA* produced wax esters including tetradecanoate hexadecyl ester (C14:0-C16), hexadecanoate hexadecyl ester (C16:0-C16), 9-hexadecenoate hexadecyl ester (C16:1-C16), and 9-octadecenoate octadecyl ester (C18:1-C18) from glucose as the sole carbon source (Figure X).

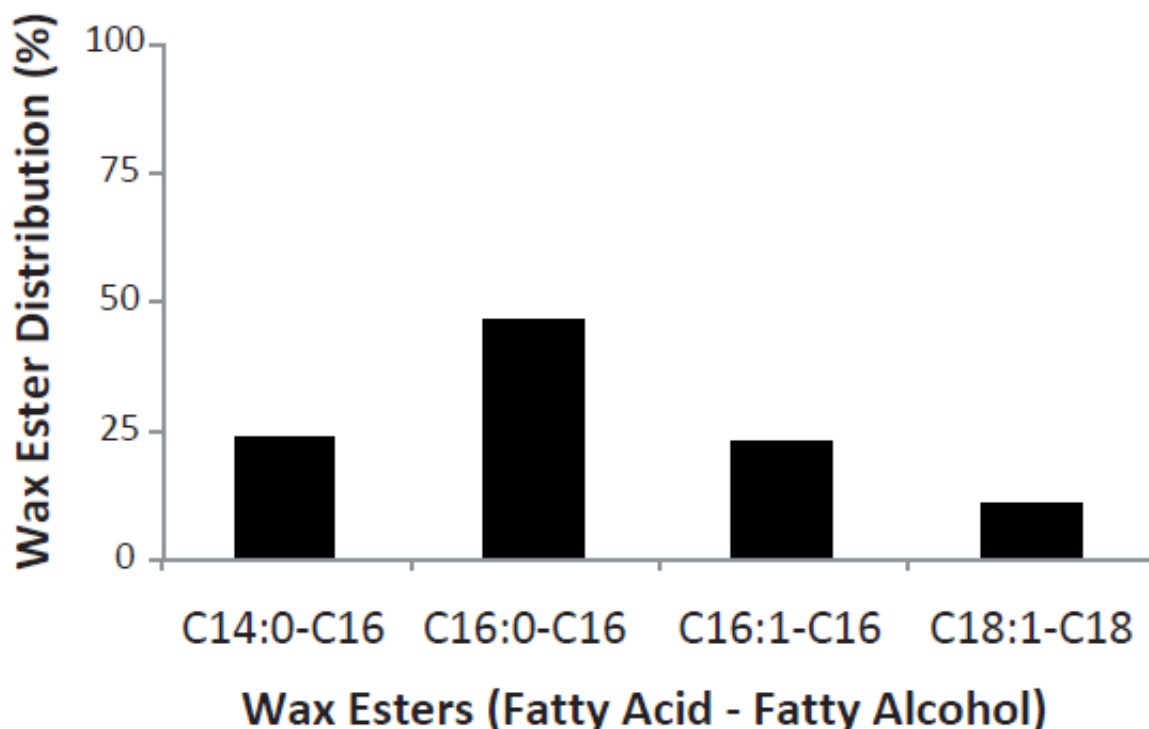


Figure 3-7: Distribution of fatty esters in *fadE* knockout expressing ‘*tesA*, *fadD*, and *atfA*.

3.3.5. Coexpression of FAEE production and hemicellulose degradation pathways.

Production of advanced fuels and chemicals such as FAEEs and fatty alcohols from sugar has promising advantage over first generation products, such as corn ethanol and plant oil-derived biodiesel. However, a primary goal in renewable energy is the production of these compounds directly from cellulosic plant biomass. Currently, biochemical processing of cellulosic biomass requires costly enzymes for sugar liberation. Consolidated bioprocessing, in which the biofuel-producing organism also produces glycosyl hydrolases, eliminates the need to add these enzymes and can improve economics⁶. To further simplify our FAEE process, we engineered into our FAEE-producing *E. coli* the genes encoding an endoxylanase catalytic domain (Xyn10B) from *Clostridium stercorarium*⁴⁶ and a xylanase (Xsa) from *Bacteroides ovatus*⁴⁷. In order to hydrolyze the hemicellulose in the growth medium into xylose, which could then be imported and catabolized using the native *E. coli* metabolism, the hemicellulases were fused

on their N-termini to the *E. coli* protein OsmY⁴⁸. Expression of OsmY-xylanase fusions enabled *E. coli* growth on hemicellulose without need for exogenously-added enzymes (Figure 5a). Addition of the biodiesel genes to the xylan-degrading strain resulted in production of 3.5 mg/L FAEEs from 0.2% glucose alone (Figure 5b). Addition of 2% xylan, the hemicellulose substrate, achieved a three-fold increase in FAEE production to 11.6 mg/L, compared to glucose alone, and demonstrated a step toward microbial biodiesel production from biomass in a consolidated process. Future engineering of the FAEE-producing *E. coli* to secrete various cellulases would enable production of FAEE's directly from both cellulose and hemicellulose without addition of expensive enzymes, further reducing the cost of cellulosic biodiesel.

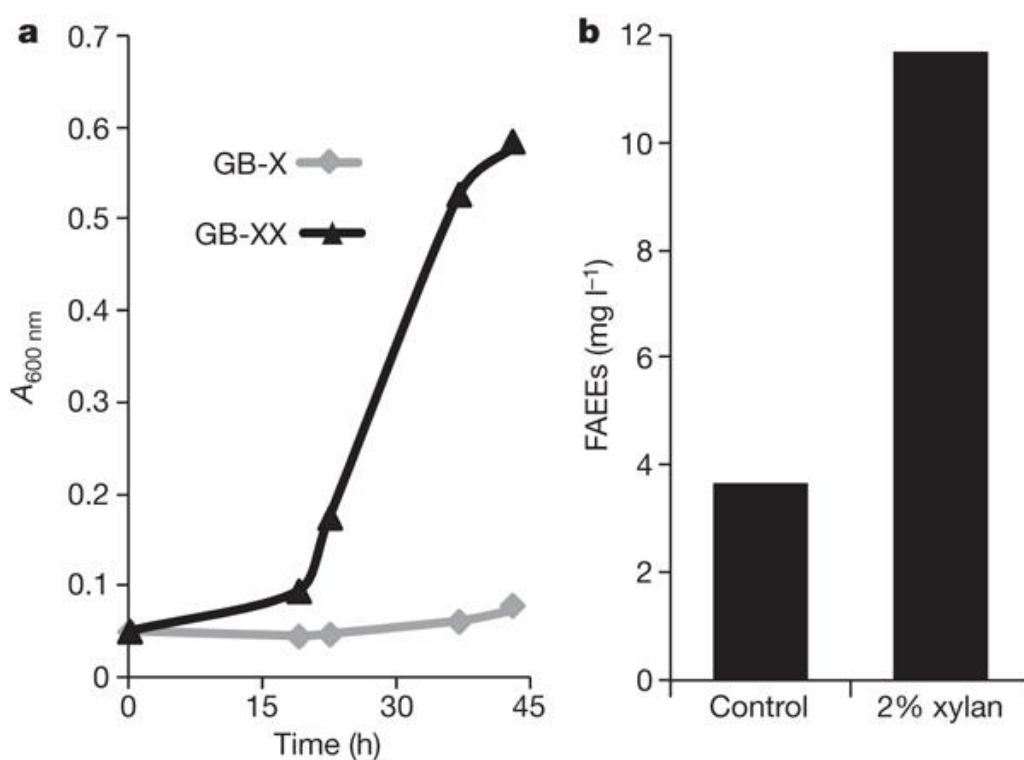


Figure 3-8: Consolidated bioprocessing: Growth and FAEE production by xylan-utilizing strains. **a**, Growth of xylan-utilizing strains on 0.2% xylan-containing minimal salts medium. Blue diamonds: GB-X, *E. coli* BL21 *xynB*; red triangles: GB-XX, GB-X expressing *xsa*. **b**, FAEE production from xylan. Control strain HE-GX (*E. coli* DH1 Δ *fadE* expressing *xynB*, *xsa*, '*tesA*, *atfA*, *pdc*, and *adhB*) grown in 0.2% glucose and no xylan; 2% xylan is HE-GX grown in 0.2% glucose and 2% xylan.

3.3.6. Cloning and overexpressing *fadD*

In our efforts to clone and overexpress *fadD* there were a number of unexpected difficulties in obtaining proper sequencing results and many attempts at cloning it into different expression vectors. This was unexpected and remains difficult to explain because *fadD* has already been cloned and overexpressed in *E. coli*⁴⁹, although overexpressing it in strain DH1 has not been reported. *FadD* has been previously characterized and contains an AMP/ATP motif (amino acids 213-224 and 356-361) and a fatty acyl-coA synthase motif (amino acids 431-455) that is potentially involved in fatty acid chain length specificity^{50,51}. We attempted to clone *fadD* into p15a-lacUV5-*'tesA*, pCDF-duet, pFIP (invertor plasmid by T. Ham), pTRC (colE1 ori), and always had multiple, random point mutations in *fadD*. The seemingly random distribution of point mutations prompted us to screen a number of clones by sequencing and lead to finding one clone with 3 DNA base pair mutations, one of which was silent and the other was repaired by quick-change PCR, leaving our final *fadD* construct containing a single amino acid change (M335I). We attempted to remove the other two mutations and were successful in removing one. We did not further pursue correcting the mutations after expression of the vector in conjunction with *L' tesA*, *pdc*, *adhB*, and *atfA* increased production of FAEEs. The difficulty of cloning *fadD* may suggest tight regulation of the gene's expression.

Table 3-1: Sequencing results and mutations in *fadD*.

Seq file	Nucleotide number	<i>fadD</i>	mutation
F2-1_2008-10-3	234	G	A
F1-1_2008-11	382	G	A
F4_2008-11	1283	T	C
I2-F4_2008-11	1282	A	G
I3-F3_2008-11	891	A	G
I3-F4_2008-11	1295	G	A
I4-F3_2008-11	118	C	A
I4-F3_2008-11	959	A	G
I6-F1_2008-11	90	G	A
I6-F1_2008-11	66	C	T
I6-F3_2008-11	842	T	G
I6-F3_2008-11	849	T	C
I8-F3_2008-11	210	G	A
I10-F1_2008-11	159	G	A
I10-F1_2008-11	274	T	C

I10-F1_2008-11	440	T	C
I10-F1_2008-11	480	A	G
C8-F4-2009-1-25	1005	G	A
C8-F4-2009-1-25	1113	T	C
C8-F3-2009-1-25	480	A	G

3.4. Discussion

In summary, this work demonstrates the practical utility of microbial fatty acid metabolism. By combining the energetically favorable TE-catalyzed hydrolysis of fatty acyl-ACP, to overproduce fatty acids and deregulate fatty acid biosynthesis, with fatty acyl-CoA synthase catalyzed reactivation of the fatty acid carboxylate group, fatty acid metabolism is efficiently diverted to fatty acyl-CoA, an important general substrate for the production of esters, alcohols, and other products. This engineering strategy supports yields of these products within an order of magnitude of that required for commercial production. Indeed, strain and process improvements of this magnitude have been achieved during scale-up of other bioprocesses⁵², and significant steps to this end have already been achieved for this process. A highly simplified calculation that only accounts for the cost of sugar and neglects fermentation, separation, etc costs indicates that at our current yields of 14%, a gallon of our FAEE is ~\$25. The further production of this biodiesel from hemicellulose demonstrates a necessary and promising achievement toward realizing a consolidated bioprocess. We believe these data can significantly contribute to the ultimate goal of producing scalable and cost effective advanced biofuels and renewable chemicals.

3.5. References

1. Fortman, J. L. et al. Biofuel alternatives to ethanol: pumping the microbial well. *Trends Biotechnol* **26**, 375-81 (2008).
2. Ingram, L. O., Conway, T., Clark, D. P., Sewell, G. W. & Preston, J. F. Genetic engineering of ethanol production in *Escherichia coli*. *Appl Environ Microbiol* **53**, 2420-5 (1987).
3. Steen, E. J. et al. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microb Cell Fact* **7**, 36 (2008).

4. Atsumi, S., Hanai, T. & Liao, J. C. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**, 86-9 (2008).
5. Chou H, K. J. (US, 2009).
6. Lynd, L. R., van Zyl, W. H., McBride, J. E. & Laser, M. Consolidated bioprocessing of cellulosic biomass: an update. *Curr Opin Biotechnol* **16**, 577-83 (2005).
7. Hill, J., Nelson, E., Tilman, D., Polasky, S. & Tiffany, D. Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proc Natl Acad Sci U S A* **103**, 11206-10 (2006).
8. Rude, M. A. & Schirmer, A. New microbial fuels: a biotech perspective. *Curr Opin Microbiol* **12**, 274-81 (2009).
9. Magnuson, K., Jackowski, S., Rock, C. O. & Cronan, J. E., Jr. Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiol Rev* **57**, 522-42 (1993).
10. Rock, C. O. & Cronan, J. E. *Escherichia coli* as a model for the regulation of dissociable (type II) fatty acid biosynthesis. *Biochim Biophys Acta* **1302**, 1-16 (1996).
11. Jiang, P. & Cronan, J. E., Jr. Inhibition of fatty acid synthesis in *Escherichia coli* in the absence of phospholipid synthesis and release of inhibition by thioesterase action. *J Bacteriol* **176**, 2814-21 (1994).
12. Heath, R. J. & Rock, C. O. Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in *Escherichia coli*. *J Biol Chem* **271**, 1833-6 (1996).
13. Heath, R. J. & Rock, C. O. Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in *Escherichia coli*. *J Biol Chem* **271**, 10996-1000 (1996).
14. Davis, M. S. & Cronan, J. E., Jr. Inhibition of *Escherichia coli* acetyl coenzyme A carboxylase by acyl-acyl carrier protein. *J Bacteriol* **183**, 1499-503 (2001).
15. DiRusso, C. C. Nucleotide sequence of the *fadR* gene, a multifunctional regulator of fatty acid metabolism in *Escherichia coli*. *Nucleic Acids Res* **16**, 7995-8009 (1988).
16. Simons, R. W., Hughes, K. T. & Nunn, W. D. Regulation of fatty acid degradation in *Escherichia coli*: dominance studies with strains merodiploid in gene *fadR*. *J Bacteriol* **143**, 726-30 (1980).
17. Simons, R. W., Egan, P. A., Chute, H. T. & Nunn, W. D. Regulation of fatty acid degradation in *Escherichia coli*: isolation and characterization of strains bearing insertion and temperature-sensitive mutations in gene *fadR*. *J Bacteriol* **142**, 621-32 (1980).
18. Maloy, S. R. & Nunn, W. D. Genetic regulation of the glyoxylate shunt in *Escherichia coli* K-12. *J Bacteriol* **149**, 173-80 (1982).
19. Maloy, S. R. & Nunn, W. D. Role of gene *fadR* in *Escherichia coli* acetate metabolism. *J Bacteriol* **148**, 83-90 (1981).
20. Xu, Y., Heath, R. J., Li, Z., Rock, C. O. & White, S. W. The FadR-DNA complex. Transcriptional control of fatty acid metabolism in *Escherichia coli*. *J Biol Chem* **276**, 17373-9 (2001).
21. Cronan, J. E., Jr. In vivo evidence that acyl coenzyme A regulates DNA binding by the *Escherichia coli* FadR global transcription factor. *J Bacteriol* **179**, 1819-23 (1997).
22. Zhu, K., Zhang, Y. M. & Rock, C. O. Transcriptional regulation of membrane lipid homeostasis in *Escherichia coli*. *J Biol Chem* **284**, 34880-8 (2009).
23. Marrakchi, H., Zhang, Y. M. & Rock, C. O. Mechanistic diversity and regulation of Type II fatty acid synthesis. *Biochem Soc Trans* **30**, 1050-5 (2002).

24. Zhang, Y. M., Marrakchi, H. & Rock, C. O. The FabR (YijC) transcription factor regulates unsaturated fatty acid biosynthesis in *Escherichia coli*. *J Biol Chem* **277**, 15558-65 (2002).
25. Cronan, J. E., Jr. & Gelmann, E. P. Physical properties of membrane lipids: biological relevance and regulation. *Bacteriol Rev* **39**, 232-56 (1975).
26. Cho, H. & Cronan, J. E., Jr. Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis. *J Biol Chem* **270**, 4216-9 (1995).
27. Voelker, T. A. & Davies, H. M. 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. *J Bacteriol* **176**, 7320-7 (1994).
28. Voelker, T. A. et al. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. *Science* **257**, 72-4 (1992).
29. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-5 (2000).
30. Link, A. J., Phillips, D. & Church, G. M. Methods for generating precise deletions and insertions in the genome of wild-type *Escherichia coli*: application to open reading frame characterization. *J Bacteriol* **179**, 6228-37 (1997).
31. Cheng, J. B. & Russell, D. W. Mammalian wax biosynthesis. I. Identification of two fatty acyl-Coenzyme A reductases with different substrate specificities and tissue distributions. *J Biol Chem* **279**, 37789-97 (2004).
32. Hoover, D. M. & Lubkowski, J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res* **30**, e43 (2002).
33. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods* **4**, 251-6 (2007).
34. Sambrook, J., Fritsch, E.F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press* **1,2,3** (1989).
35. Aldai, N., Osoro, K., Barron, L. J. & Najera, A. I. Gas-liquid chromatographic method for analysing complex mixtures of fatty acids including conjugated linoleic acids (cis9trans11 and trans10cis12 isomers) and long-chain (n-3 or n-6) polyunsaturated fatty acids. Application to the intramuscular fat of beef meat. *J Chromatogr A* **1110**, 133-9 (2006).
36. Chang, A. C. & Cohen, S. N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**, 1141-56 (1978).
37. Miroux B, W. J. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol.* **260**, 289-298 (1996).
38. Lu, X., Vora, H. & Khosla, C. Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. *Metab Eng* **10**, 333-9 (2008).
39. Dehesh, K., Jones, A., Knutzon, D. S. & Voelker, T. A. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. *Plant J* **9**, 167-72 (1996).
40. Oil Market Report. *International Energy Agency* (January 2008).
41. Kalscheuer, R., Stolting, T. & Steinbuchel, A. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology* **152**, 2529-36 (2006).

42. Ahmad, W. R. a. S. The changing world of oleochemicals. *Palm Oil Developments*, 15-28 (2006).
43. Metz, J. G., Pollard, M. R., Anderson, L., Hayes, T. R. & Lassner, M. W. Purification of a jojoba embryo fatty acyl-coenzyme A reductase and expression of its cDNA in high erucic acid rapeseed. *Plant Physiol* **122**, 635-44 (2000).
44. Reiser, S. & Somerville, C. Isolation of mutants of *Acinetobacter calcoaceticus* deficient in wax ester synthesis and complementation of one mutation with a gene encoding a fatty acyl coenzyme A reductase. *J Bacteriol* **179**, 2969-75 (1997).
45. Knothe, G. "Designer" Biodiesel: Optimizing Fatty Ester Composition to Improve Fuel Properties†. *Energy & Fuels* **22**, 1358-1364 (2008).
46. Adelsberger, H., Hertel, C., Glawischnig, E., Zverlov, V. V. & Schwarz, W. H. Enzyme system of *Clostridium stercorarium* for hydrolysis of arabinoxylan: reconstitution of the in vivo system from recombinant enzymes. *Microbiology* **150**, 2257-66 (2004).
47. Whitehead, T. R. & Hespell, R. B. The genes for three xylan-degrading activities from *Bacteroides ovatus* are clustered in a 3.8-kilobase region. *J Bacteriol* **172**, 2408-12 (1990).
48. Qian, Z. G., Xia, X. X., Choi, J. H. & Lee, S. Y. Proteome-based identification of fusion partner for high-level extracellular production of recombinant proteins in *Escherichia coli*. *Biotechnol Bioeng* **101**, 587-601 (2008).
49. Black, P. N., DiRusso, C. C., Metzger, A. K. & Heimert, T. L. Cloning, sequencing, and expression of the *fadD* gene of *Escherichia coli* encoding acyl coenzyme A synthetase. *J Biol Chem* **267**, 25513-20 (1992).
50. Weimar, J. D., DiRusso, C. C., Delio, R. & Black, P. N. Functional role of fatty acyl-coenzyme A synthetase in the transmembrane movement and activation of exogenous long-chain fatty acids. Amino acid residues within the ATP/AMP signature motif of *Escherichia coli* *FadD* are required for enzyme activity and fatty acid transport. *J Biol Chem* **277**, 29369-76 (2002).
51. Black, P. N., Zhang, Q., Weimar, J. D. & DiRusso, C. C. Mutational analysis of a fatty acyl-coenzyme A synthetase signature motif identifies seven amino acid residues that modulate fatty acid substrate specificity. *J Biol Chem* **272**, 4896-903 (1997).
52. Tsuruta, H. et al. High-level production of amorpho-4,11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS One* **4**, e4489 (2009).

4. Optimization of fatty acid and fatty acid ethyl ester production in *E. coli*.

4.1. Introduction

Our previous efforts in developing a base fatty acid pathway that could be exploited for producing a wide variety of products was successful, but the titers of the products remained below the theoretical maximum yields and had slow productivities in shake flasks and test tubes. Thus our efforts focused on understanding what limitations existed in our pathway and strains and whether our strains could be scaled to more controlled fermentations. There are many approaches to pathway optimization and can be categorized in two ways: methods that are rational and stem from knowledge derived from experiments aimed at answering a hypothesis (adding extra copies of genes in the pathway to check for bottlenecked reactions) or allow an observation that leads to a rational change (transcriptomics, metabolomics, and fluxomics) and methods that rely on screens and selections to explore vast experimental space in order to identify or select a better producer; the following describes our efforts in pursuing the more rational approach. Herein, we describe fermentation scale capability of our strain A2A, formation of byproducts, plasmid stability, mitigating stability by addition of antibiotics, optimizing growth and production parameters, cloning the pathway onto a single plasmid, and integrating the thioesterase.

4.2. Materials and Methods

Strains and plasmids. *E. coli* strains DH1, MG1655, and MDS42recA- were utilized where indicated. Knockouts of *fadE*, in DH1 were performed as previously described^{1,2}. A list of strains and plasmids constructed is given in Table 4-1 and a list of genes and corresponding accession numbers used is given in Table 3-2. Construction of plasmids was carried out with standard molecular biology methods and is described in detail in Methods.

Table 4-1: Plasmids and strains used in this study.

Plasmids	Replication Origin	Overexpressed Genes	Resistance	Reference
pKS1	p15a	<i>P_{lacUV5}: 'tesA</i>	Cam	²
pKS17	pBBR	<i>P_{lacUV5}: pdc, adhB, atfA</i>	Tet	²
pKS104	ColE1	<i>P_{lacUV5}: fadD (M335I), atfA</i>	Amp	²
pES115	ColE1	<i>fadD(M335I), atfA, pdc, adhB, 'tesA, 3 operon</i>	Kan	This study

pES116	ColE1	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, 2 operon</i>	Kan	This study
pES117	ColE1	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, ''tesA</i>	Kan	This study
pES118	ColE1	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, pdc, adhB</i>	Kan	This study
pES119	ColE1	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, fadD</i>	Kan	This study
pES120	ColE1	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, atfA</i>	Kan	This study
pES121	p15a	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, 3 operon</i>	Kan	This study
pES122	p15a	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, 2 operon</i>	Kan	This study
pES123	p15a	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, ''tesA</i>	Kan	This study
pES124	p15a	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, pdc, adhB</i>	Kan	This study
pES125	p15a	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, fadD</i>	Kan	This study
pES126	p15a	<i>fadD(M335I),atfA, pdc, adhB, 'tesA, atfA</i>	Kan	This study

Strains	Relevant Genotype	Reference
A2A	DH1: <i>ΔfadE</i> , pKS1, pKS17, pKS104	2

Growth and production parameters. Strains were cultivated in 50 mL test-tubes (5 mL culture volume) at 37°C in M9 minimal medium supplemented with trace elements and the appropriate antibiotics, unless otherwise indicated (50 mg/L ampicillin, 20 mg/L chloramphenicol, 5 mg/L tetracycline, 100 mg/L carbenicillin, and 50 mg/L kanamycin). Strains were passed overnight in M9 MOPS minimal media, OD was recorded the following day (~16h) and the strains were centrifuged and resuspended at a defined OD (usually 1 unless otherwise indicated) in fresh medium containing 0-1 mM IPTG. Metabolites were identified and quantified by gas chromatography-mass spectrometry, further described in Methods.

Reagents. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) and include fatty acid methyl ester standards, fatty acid ethyl ester standards, and ethyl acetate.

Plasmid and strain construction. *E. coli* DH10B and DH5 α were used for bacterial transformation and plasmid amplification in the construction of the expression plasmids used in this study

Metabolite Analysis. Free fatty acids and ethyl esters were sampled by removing 500 μ L of culture to a 1.5 mL eppendorf centrifuge tube. 50 μ L HCl and 500 μ L of ethyl acetate, spiked with 10 mg/L of methyl nonadecanoate as an internal standard, were added to the culture sample, which was vortexed 10 s at max speed. Samples were then centrifuged at max speed for 1 min and then 400 μ L of the ethyl acetate layer was recovered into a GC vial. To this, 40 μ L of 9:1 methanol:HCl was added, followed by addition of 50-100 μ L of TMS-diazomethane (2M in hexanes). The vial was capped, mixed and vented until the reaction completed. This was modified from previous work^{2,3}. This reaction was allowed to proceed for 2 hr and then was applied to a Thermo Trace Ultra gas chromatograph (GC) equipped with a Triplus AS autosampler and a TR-5/DB-5, 10m $\times$ 0.32 $\times$ 0.1 column (Thermo Scientific). The GC program was as follows: inlet temperature of 250°C, initial temperature of 45°C for 2.25 min, ramped to 300°C at 40°C/min. Final quantification analysis was performed with Xcalibur software.

Glucose, Acetate, Succinate, Lactate, and Ethanol were measured by sampling 0.5 mL of culture, centrifuging 18,000 $\times$ g for 5 min, and applying the supernatant to an Agilent 1100 series HPLC equipped with an Aminex HPX-87H ion exchange column (Biorad). The solvent, 4 mM H₂SO₄, flow rate was 0.6 mL/min and the column was maintained at 50°C. All metabolites were detected with an Agilent 1200 series DAD and RID detectors.

Fermentation parameters. Adapted glycerol stocks of A2A were made and stored at -80°C. One, 1mL stock was added to 30 mL M9 MOPS minimal media with appropriate antibiotic concentrations as previously described and the culture was grown overnight to and OD of 2.0. The 30-mL culture was then used as inoculum for a beginning batch phase culture of 1.47 L. The media of the batch culture is described below in Table 4-2. The batch phase was grown for 60 h at which point feeding started and was controlled by the respiratory quotient and maintained near 1.2%. The feed rate varied between 0.05 – 0.20 mL / min and the fermentation was ended at 120h. The fermentor was purchased from Sartorius and was model Biostat C.

Table 4-2: Media recipe for fed-batch fermentation of strain A2A

Batch Phase	
Component	g/L
Glucose	65
M9 Salts	2.3
NH ₄ Cl	15
CaCl ₂	.2
Mineral solution	2
FeSO ₄	.084
MgSO ₄	.7
Thiamine	.009
Anti-foam	20 (mL)
Fed Batch Solution	
	g/L
Glucose	200
M9 Salts	5
NH ₄ Cl	40
CaCl ₂	1
Mineral solution	3
FeSO ₄	.084
MgSO ₄	3
Thiamine	.018
Anti-foam	30 (mL)

4.3. Results

4.3.1. 2-Liter fermentation studies of strain A2A

Strain A2A, the highest producing fatty acid ethyl ester strain engineered in our previous studies, was grown in a 2-L fed batch fermentation in order to determine productivity, titer, and yields and address the scalability of our process. Three of the most important properties of a sound industrial microbial fermentation process are encompassed in yield, titer, and productivity. Yield is the amount of sugar or other feedstock required per g of product formed; titer is the maximum, final product concentration; and productivity is the rate at which product is formed. The production time-course illustrates production over a 120-h time and demonstrates

FAEEs production to ~6 g/L, while free fatty acids, measured as FAMES remain below 0.5 g/L (Figure 4-1).

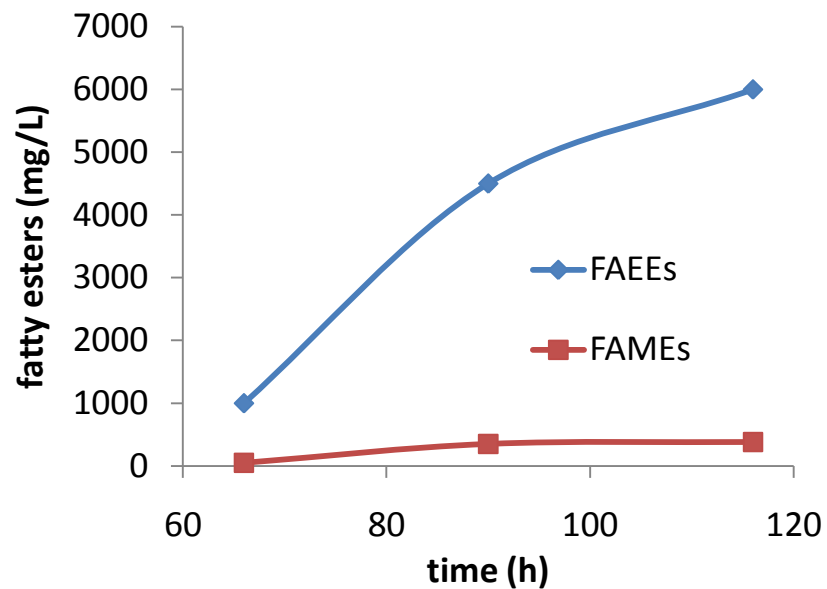


Figure 4-1: Free fatty acid and FAEE production from strain A2A.

The yield of FAEEs per gram of glucose was also plotted over 120 hours and indicates a peak in yield of 7% (or roughly 20% of the theoretical maximum yield). It is also noteworthy that the yield drops after 90 hours to 2.5% (Figure 4-2).

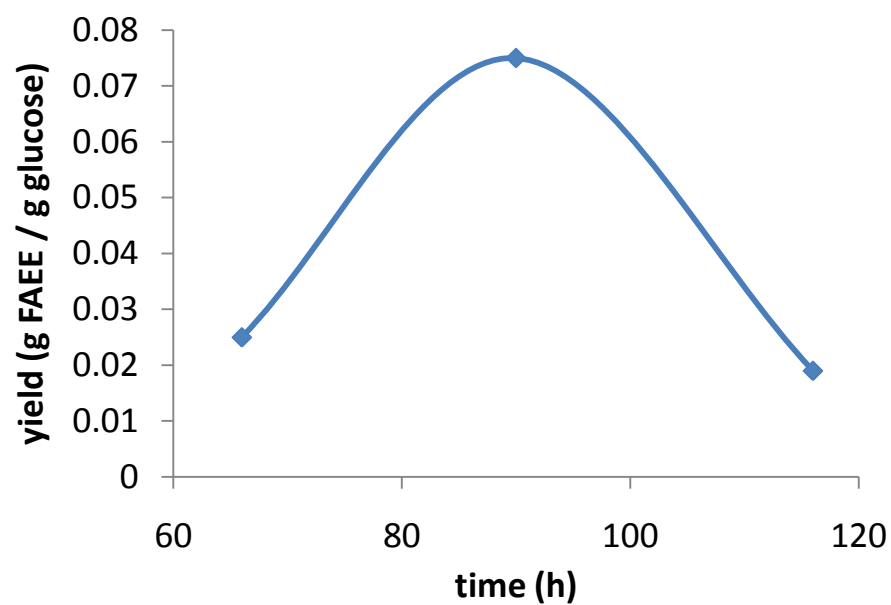


Figure 4-2: Yield of FAEEs from strain A2A.

Productivity can be calculated from the production time-course and is roughly 0.15 g/L/h as shown in Table 4-3 in addition to a summary of the other parameters.

Table 4-3: Fermentation characteristics of strain A2A.

Fermentation Parameter	Achieved	Goal
Productivity (mg/L/h)	150	>1000
Titer (g/L)	6	>100
Yield (g FAEE / g glucose)	.08	.33
Cost (\$ / gallon)	25	2.50

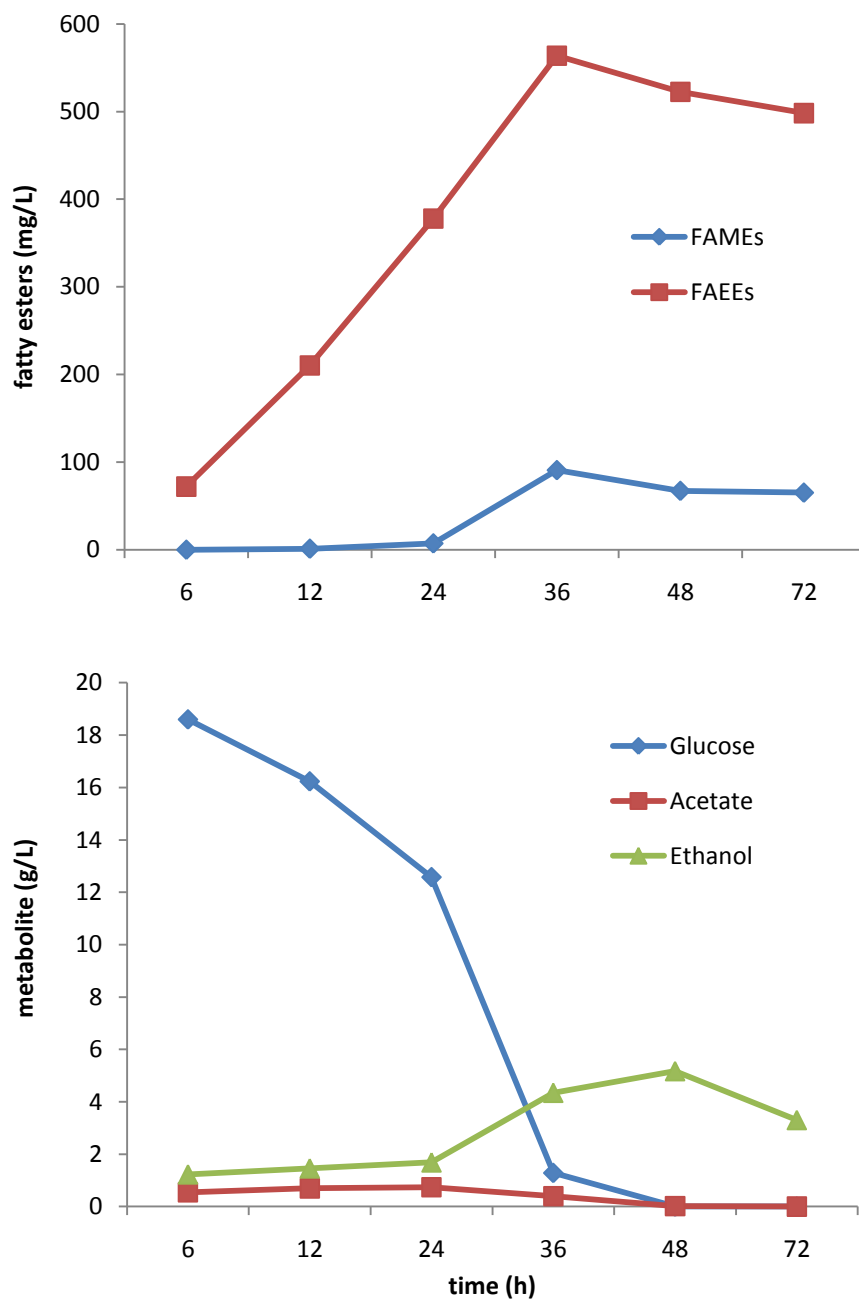


Figure 4-3: Metabolite measurement from strain A2A under low IPTG-induction conditions. Metabolites measured include FAEs, FAMES, glucose, acetate and ethanol.

4.3.2. Byproduct formation in test tubes.

In optimizing the FAEE pathway, it is useful to determine into what products the feedstock is converted and how fast the feedstock is consumed. These products detract from the total amount of carbon that can be converted into FAEEs and therefore removal of these products may result in higher yields and feedstock

consumption rates can indicate if a given process is carbon limited. Typical byproducts of *E. coli* mixed acid fermentation are acetate, lactate, succinate, ethanol, and carbon dioxide. Additional byproducts for us that indicate an unbalanced pathway in addition to ethanol are free fatty acids. We measured these products and found production of FAEEs to 550 mg/L and FAMEs to 100 mg/L (Figure 4-3).

A2A consumes 2% glucose at a constant rate until 24 h at which point, the rate of consumption increases, until 48 h when glucose is depleted (Figure 4-3). Production of acetate rises to a maximum of 0.7 g/L at 24h and then is steadily assimilated through 48h at which point it becomes nearly undetectable. Ethanol production steadily increases up to 24 h, reaching ~1.5 g/L, rapidly increases to ~5 g/L at 48 h, and then decreases to 72 h. Lactate and succinate levels were detected and less than 0.01 g/L (data not shown).

Plasmid stability of strain A2A

Plasmid stability was analyzed in order to identify potential losses in productivity either in test tubes or fermentor. The A2A strain harbors three plasmids that each encode the genes required for the FAEE pathway. Each plasmid also encodes resistance to one of three antibiotics: tetracycline (Tc), chloramphenicol (Cm), and carbenicillin (Cb). The pathway genes each plasmid encodes in addition to the antibiotic concentrations and origin of replications can be found in Table 4-4.

Table 4-4: A2A operons and antibiotic concentrations

Plasmid (antibiotic and origin of replication)	Genes (driven by lacUV5 promoter)	[Antibiotic] ug/mL
pKS17 (Tc, pBBR)	<i>pdhB-adhB-atfA</i>	5
pKS1 (Cm, p15A)	<i>tesA</i>	50
pKS104 (Cb, colE1)	<i>fadD-atfA</i>	100

The plasmids were retained to different degrees. Plasmid pKS17 was retained the best and exhibited ~100% retention over 72 h, plasmid pKS1 exhibited retention between 50%-80% over 72 h, and plasmid pKS104 exhibited the worst retention, dropping to 50% at 24 h and 0% at 48 h (Figure 4-4). Finally, the combination of all three plasmids dropped to 20% by 24 h and 0% at 48 h.

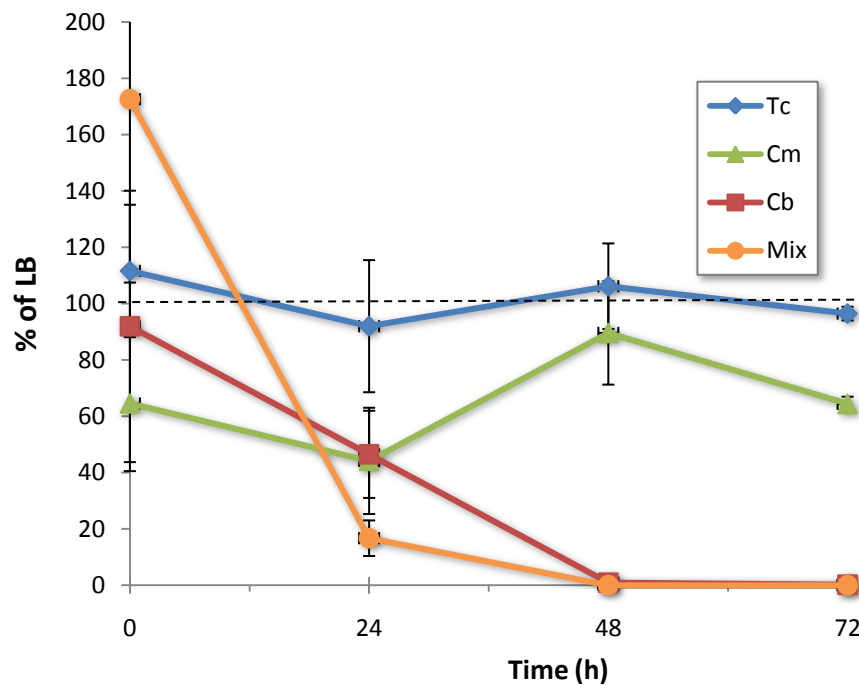


Figure 4-4: Plasmid retention in strain A2A under low IPTG.

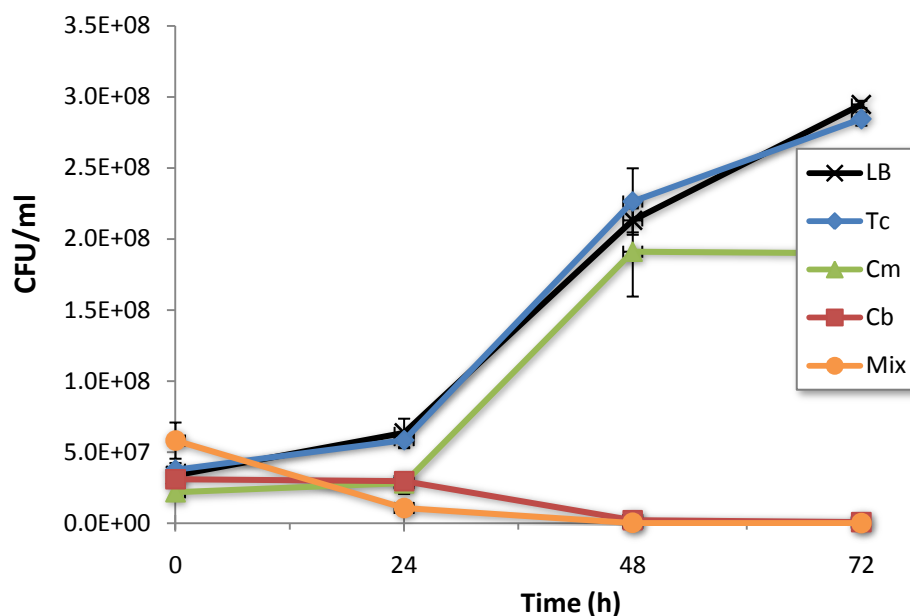


Figure 4-5: Colony forming units for strain A2A under low IPTG.

4.3.3. Increasing antibiotic concentration for increasing plasmid maintenance

Our results that investigated plasmid retention clearly show loss of the ColE1 plasmid whose retention is selected based upon resistance to carbenicillin and potential loss of the p15a plasmid, whose retention is selected for based upon resistance to chloramphenicol. Thus, we hypothesized that increasing antibiotic concentration may result in increased plasmid retention and therefore increased fatty acid ethyl ester production levels. There are two approaches to increasing the antibiotics that we tried. In one case, we added 1X, 2X, or 3X the usual concentrations of Cm and Cb to the beginning of the induced culture and in another, we added additional, 1X concentrated antibiotics at 12 h. Results are shown below in Figure XX. At 24 h, addition of more concentrated antibiotic does not seem to have much of an effect, with all three cultures producing roughly 700 mg/L FAEs and 120 mg/L of FAMES. At 48 h however, addition of 2X or 3X the antibiotics results in a bump in production of FAEs to ~1 g/L, compared to the 1X, that produces ~750 mg/L. Further, adding more antibiotic after 16 h also increases production to ~900 mg/L compared to the standard concentration at 48 h. Interestingly, at 72 h less FAEs were recovered in all cases than the 48 h time

point, but the trend is similar, where 2X, 3X, or additional antibiotics increases FAEEs compared to the standard condition.

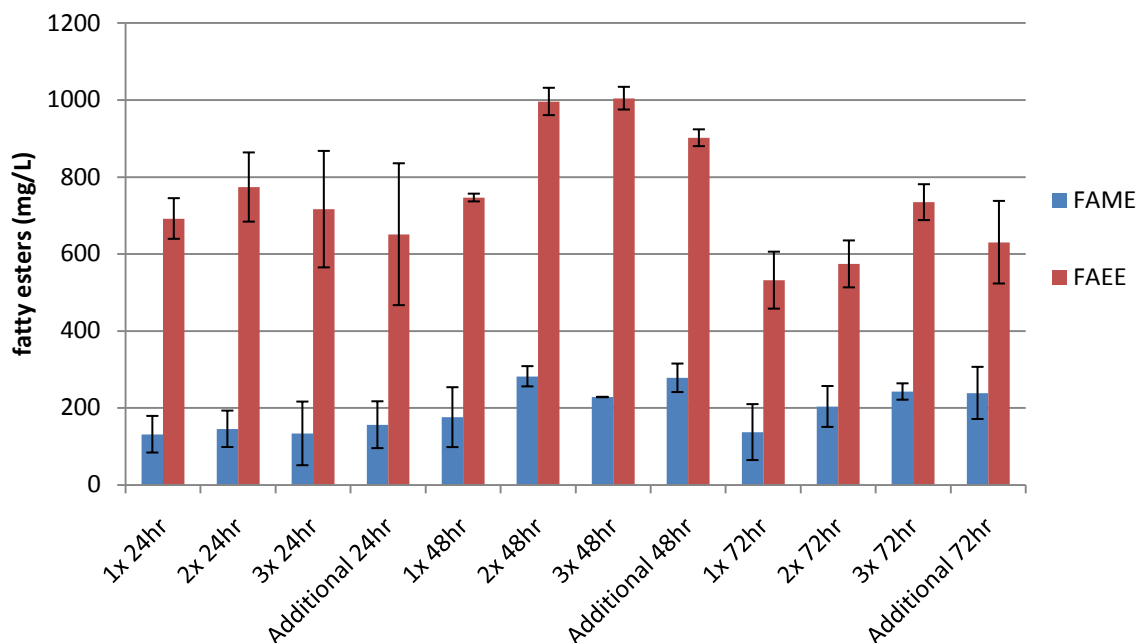


Figure 4-6: FAME and FAEE production in strain A2A with varying antibiotic doses.

4.3.4. Optimization of induction and growth parameters of strain A2A.

In an effort to increase production of FAEEs from the strain A2A and establish an easier, more reproducible way of inducing cultures, we tested a number of IPTG induction concentrations and time of induction. Previously, pathway induction was done with high concentrations of IPTG (1 mM) at late log phase (OD=0.8-1.0). This method of induction proved to be extremely laborious when cultures did not grow at the same rates, thus, we began adapting cultures overnight and then resuspending enough cells to OD=1.0 in fresh media with a given amount of IPTG. We titrated IPTG from 0 to 1 mM and found a range of FAEE production from 100 mg/L to 1.2 g/L, with free fatty acid production remaining mostly constant near 200 mg/L. The correlation between IPTG and FAEE production is clear, demonstrating that lower concentrations of IPTG (10 μ M) results in higher production of FAEEs, while raising the inducer concentration steadily decreases production. Further, optimal production lies between 0 and 50 μ M IPTG (Figure

4-12). In order to have greater resolution of the optimal IPTG concentration, the experiment was repeated (Figure 4-13). However, this experiment resulted in less than half the production at the optimal IPTG concentration (~500 mg/L), approximately the same levels of free fatty acid production (~200 mg/L) and the optimal concentration was not the same as the previous experiment. Additionally, the trend of inducer concentration and production is reversed; increasing inducer concentration results in higher production of FAEs. These differences can only be explained by the differences in growth phase or OD at the time of induction. The first experiment cultures were at OD = 1.2, whereas the second experiment cultures were at OD = 0.9. Thus, a slight difference in growth phase may be responsible for large differences in both production and optimal induction parameters.

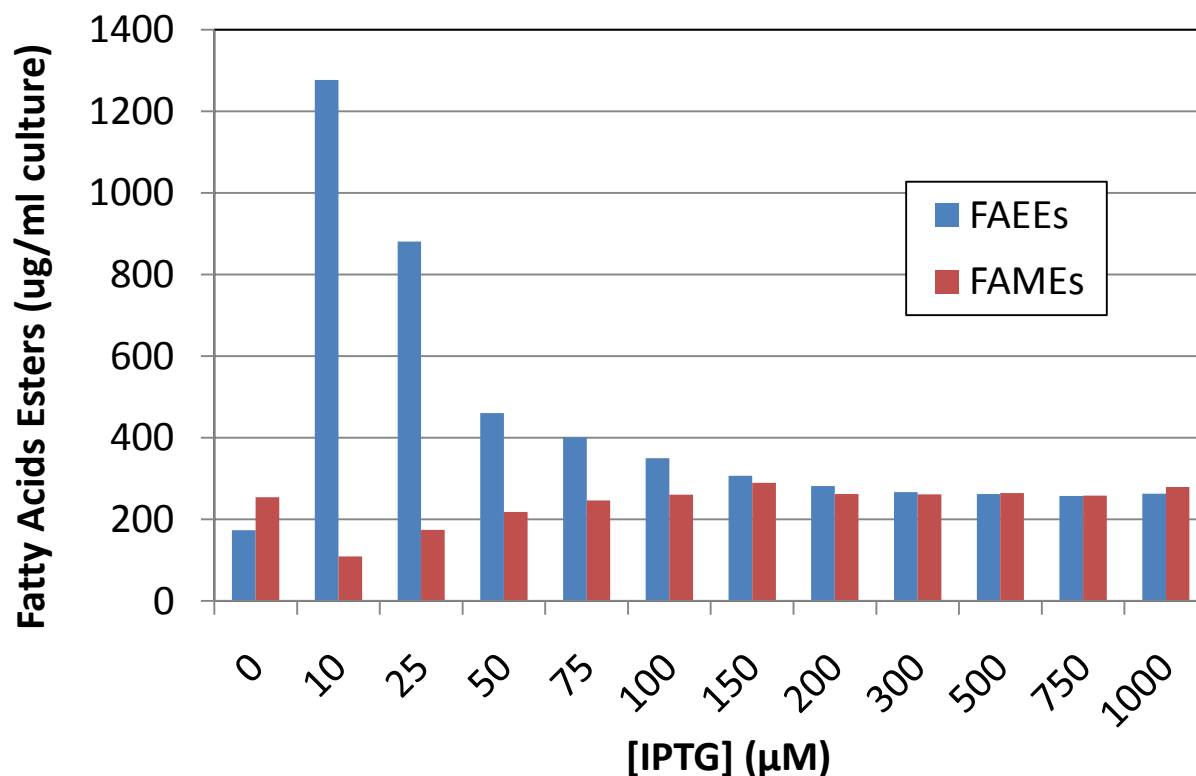


Figure 4-7: FAME and FAEE production in strain A2A with varying IPTG concentrations.

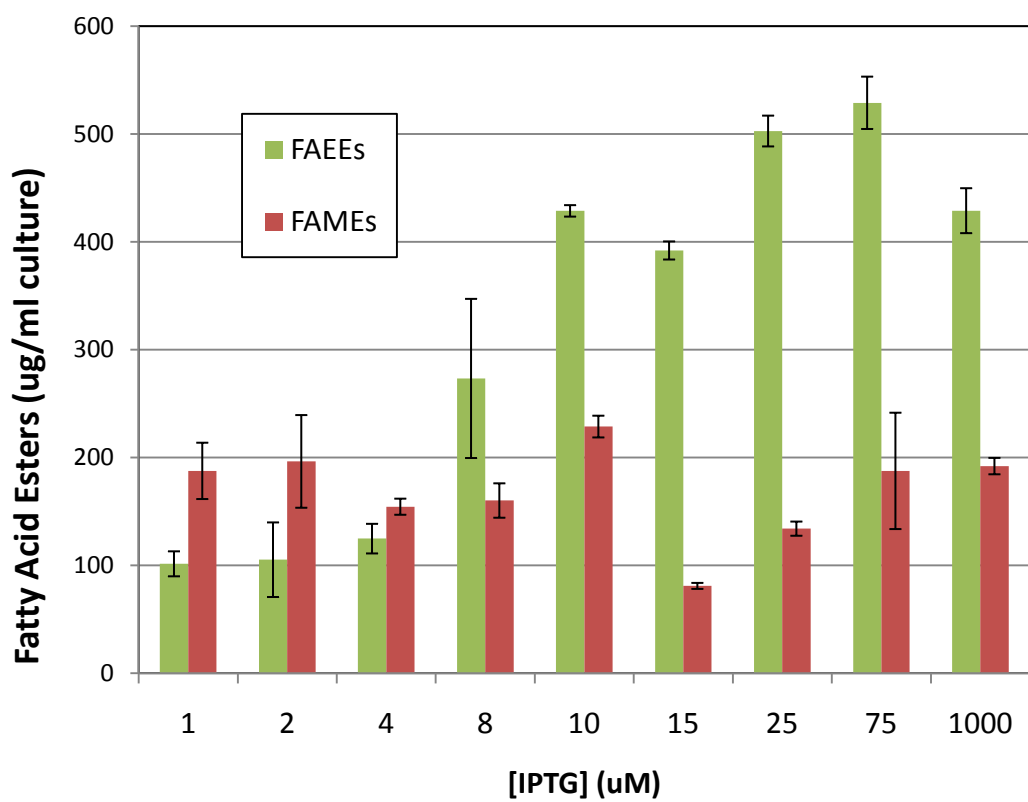


Figure 4-8: FAME and FAEE production in strain A2A with varying IPTG concentrations.

These discrepancies led us to investigate how the differences in both growth phase and inducer concentrations could affect FAEE production. We tested two different growth phases (OD = 0.866 or 1.5, corresponding to mid-exponential and late exponential phase) and two different inducer concentrations (10 μ M and 1 mM IPTG). Cultures were adapted overnight and resuspended at OD = 1.0 in fresh medium with IPTG when they had reached a given growth phase. Production of FAEEs ranged from 150 mg/L to 800 mg/L and free fatty acids ranged from 100 mg/L to 350 mg/L. Earlier compared to later growth phases with the same inducer concentration resulted in higher production of FAEEs, reaching 800 mg/L compared to 550 mg/L for low inducer concentration and reaching 250 mg/L compared to 150 mg/L for high inducer concentration. Lower inducer concentrations result in higher production for early growth phase, whereas higher inducer concentrations result in higher production for later growth phase, but overall, lower inducer concentration and early growth phase results in higher production of FAEEs.

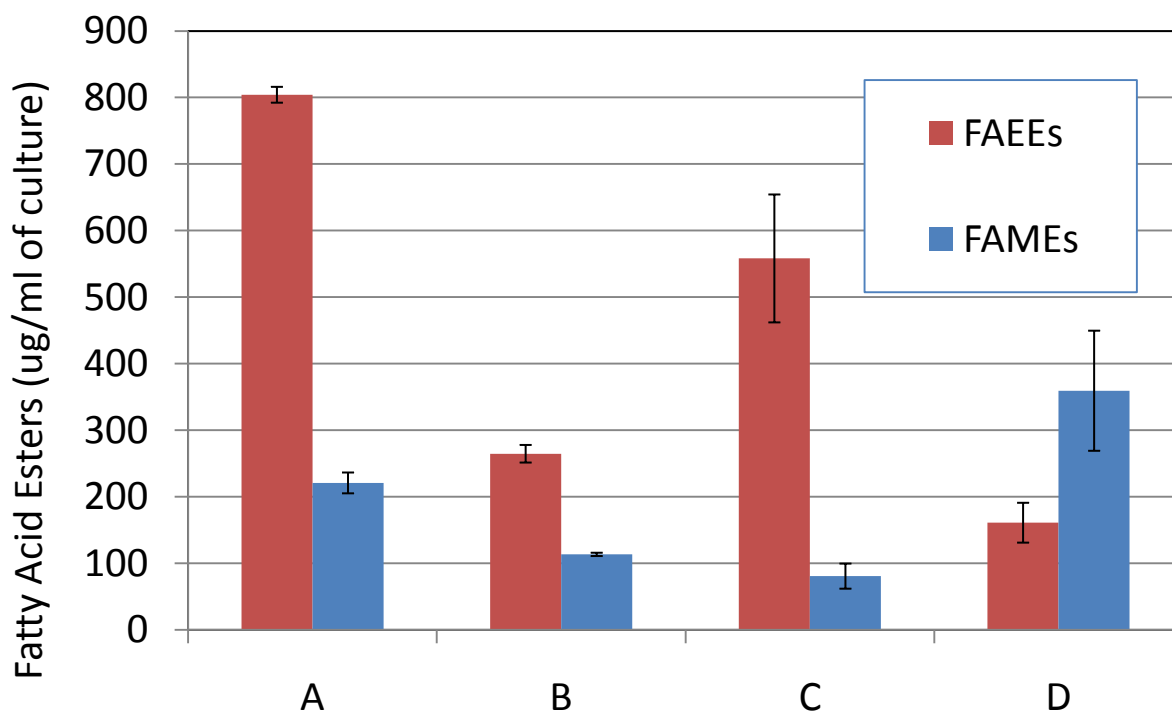


Figure 4-9: FAME and FAEE production in strain A2A with varying IPTG and OD .

Table 4-5: Induction and OD parameters.

Sample ID	OD (preculture)	[IPTG]
A	Low (0.866)	low (10 uM)
B	Low	High (1 mM)
C	High (1.5)	Low
D	High	High

4.3.5. Single plasmid productivity of FAEEs

Multiple strategies exist for optimizing metabolic pathways and often focus on removal of toxic intermediates that impair growth and productivity, and optimizing expression by codon-optimization, tuning promoter strength, ribosome-binding site strength⁴, or gene copy number⁵. Since the three plasmid system is clearly unstable, and especially the plasmid pKS104, we sought to clone the entire pathway onto a single plasmid (high and medium copy) vector with three operons intact, using kanamycin resistance for plasmid selection. Additionally, this would allow adding an extra copy of each gene to interrogate whether one of the 5 genes is limiting. Finally, since acetate was accumulating and potentially slowing metabolism and flux to FAEEs we tested these single plasmid constructs (ES115-ES126) in the acetate pathway knockout background, strain PE (DH1, Δ *ackA*, Δ *poxB*, Δ *pta*, Δ *fadE*) in addition to the HE background (DH1, Δ *fadE*). Surprisingly, production of FAEEs in most strains dropped compared to the A2A control (~400 mg/L), except for HE-ES119 with an extra copy of *fadD*, and PE-ES116, the dual-operon plasmid, which achieved roughly the same level of production. However, two strains, HE-ES123 harboring an extra copy of *tesA* and PE-ES115 with a three-operon plasmid, showed high production of free fatty acids to 1 and 1.2 g/L respectively (Figure 4-15).

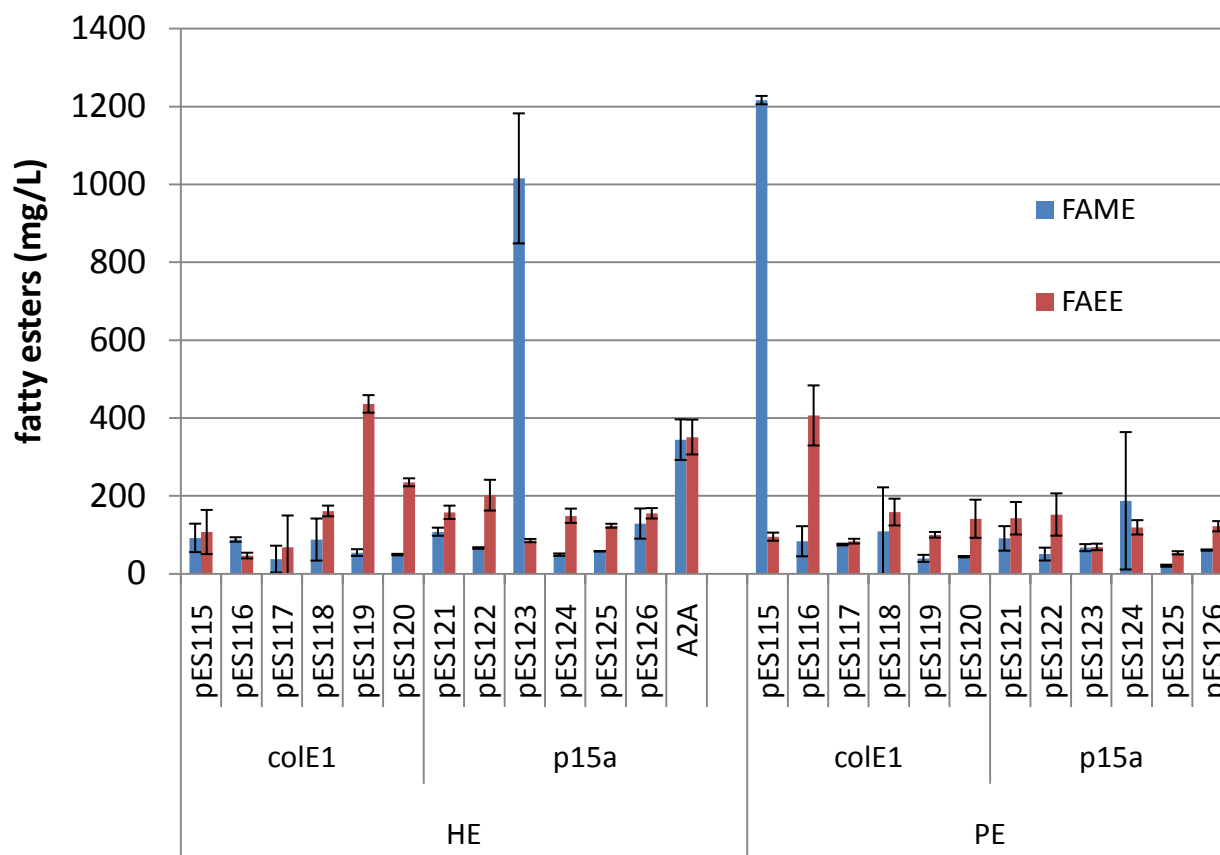


Figure 4-10: FAME and FAEE production in single plasmid strains.

Due to our previous efforts in optimizing growth phase and inducer concentration in strain A2A, we performed a similar experiment, varying growth phase, taking into account the higher final ODs that the single plasmid strains reached (OD ~ 1.0, 2.5, and 4; A2A ODs ~ 0.5, 0.9, 1.6) and inducer concentrations (IPTG = 25 μ M or 1 mM). A range of FAEE and free fatty acid production was achieved, from 25 mg/L to 350 mg/L and 10 mg/L to 500 mg/L respectively. Generally, low inducer concentrations resulted in no major change in either FAEE or free fatty acid production and produced more FAEEs compared to high inducer concentrations. Further, the levels of FAEE production for 120 and A2A under low inducer concentrations were approximately the same, reaching ~310 mg/L, while the levels of FAEE production for 119 only reached 100-150 mg/L. In contrast, production from cultures induced with high concentrations of IPTG resulted in much less FAEE production, in some cases dropping by more than 3-fold to less than 100

mg/L. Further, HE120 exhibited an increasing trend of free fatty acid production from 250 mg/L to 500 mg/L as the induction growth phase was increased, while HE119 and A2A free fatty acid production levels did not change.

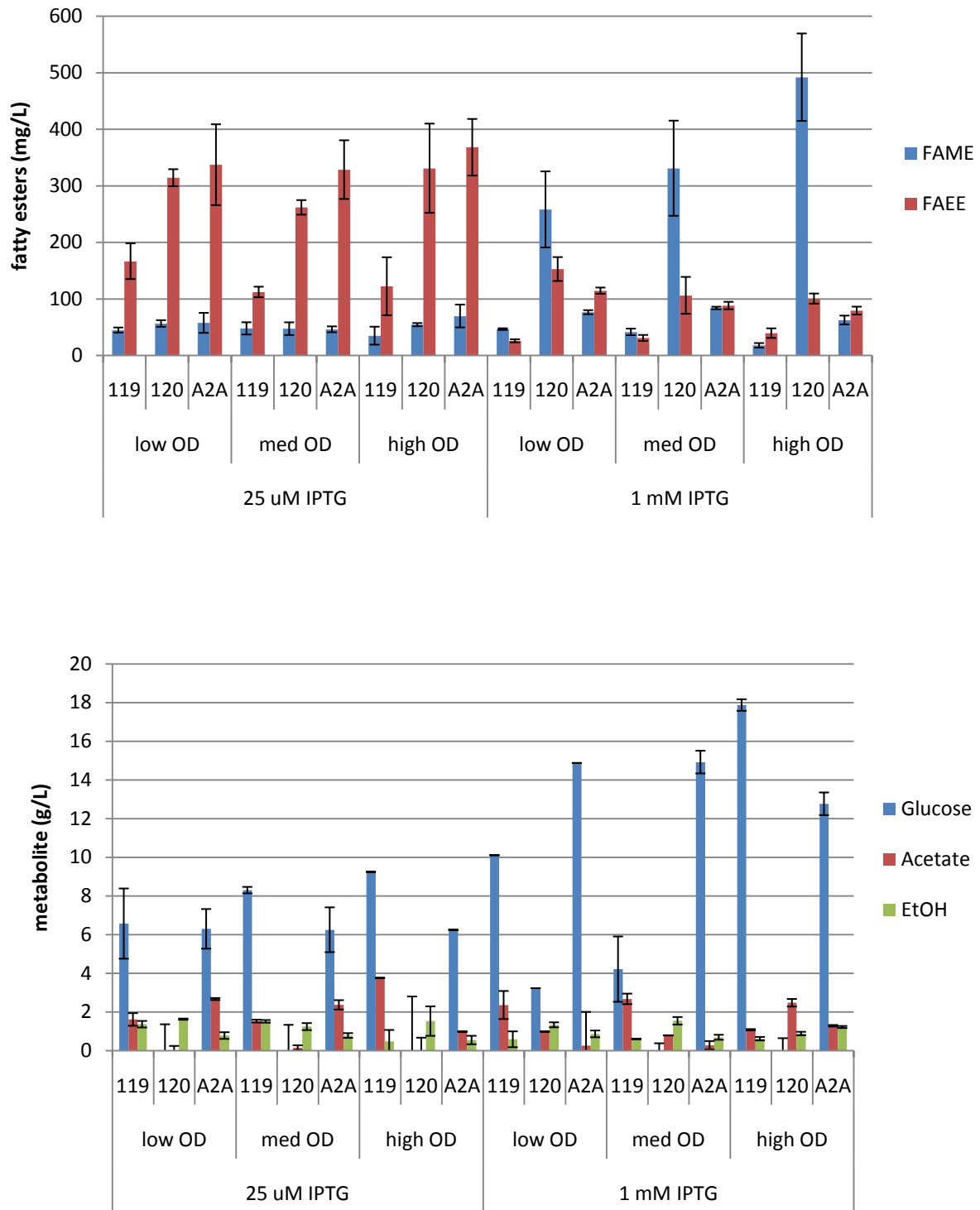


Figure 4-11: Top. FAME and FAEE production in single plamid strains with varying OD and IPTG. Bottom. Metabolite measurement of glucose, acetate, and ethanol (EtOH). Values represent the mean and standard deviation of triplicate samples.

In addition to measuring fatty acid production levels, we measured the typical byproducts of mixed-acid fermentations and glucose and detected glucose, acetate and ethanol (Figure 4-17). Similar to the fatty acid production data, we found similar levels of metabolites for each strain across the low, medium and high OD phases. Specifically, 119 and A2A had similar levels of glucose, acetate and ethanol, with levels near 6 g/L, 2 g/L, and 1 g/L respectively. However, the levels of acetate increased slightly to ~4 g/L for strain 119 in the high OD condition. In contrast, strain 120 consumed all the glucose and accumulates little acetate and ethanol. In the high inducer condition, strain A2A maintains roughly the same levels of metabolites across the low, medium, and high OD conditions. However, levels of glucose in strains 119 and 120 vary across the OD conditions. Strain 119 consumed 10 g/L, 16 g/L and 2 g/L of glucose when induced at low, medium, and high ODs, and strain 120 consumed 17 g/L, 20 g/L, and 20 g/L when induced at low, medium, and high ODs. Levels of acetate and ethanol remained roughly the same.

4.3.6. Integration of thioesterase, *'tesA*

Two other studies were published in 2010 that elucidated some of the difficulty of engineering the fatty acid pathway to produce free fatty acids^{6,7}, and some of their work may corroborate or help to explain our findings. First, Lennen et al. demonstrated that titration of a plant thioesterase (BTE) by expressing it from single, low, medium and high copy number plasmids resulted in an optimal level of expression that was not the highest level of expression⁷, suggesting that higher levels of thioesterase somehow inhibits growth and fatty acid production. The second development by Liu et al. was the demonstration of a cell free fatty acid biosynthesis system in which various enzymes of the pathway can be titrated and their effect determined. In the study, it was found that titration of the native *E. coli* thioesterase, *'tesA*, resulted in a dramatic peak for free fatty acid production, where a narrow window of thioesterase concentration must be achieved in order to reach maximal levels of fatty acid production⁶. Coupled with our datasets that have been difficult to reproduce and demonstrate very stringent parameters for production, we sought to integrate the thioesterase into the chromosome in an attempt to remove any instability associated with plasmid copy number. Further, we wanted to investigate a number of other wild-type *E. coli* strains other than DH1, which has a growth deficiency and potentially some auxotrophies, including one for thiamine.

'*tesA*' was integrated into the *fadE* gene in the *E. coli* strains BW25113¹, DH1, and MDS42⁸ and the strains were tested for free fatty acid production over a range of inducer concentrations (Fig 4-18). Surprisingly, production was roughly constant across all integrated strains and the DH1 *fadE* knockout strain carrying '*tesA*' on a plasmid. The integrated strains resulted in production of ~200 mg/L, while the plasmid-based strain (DH1: Δ *fadE* + pKS1) resulted in production to ~ 600 mg/L.

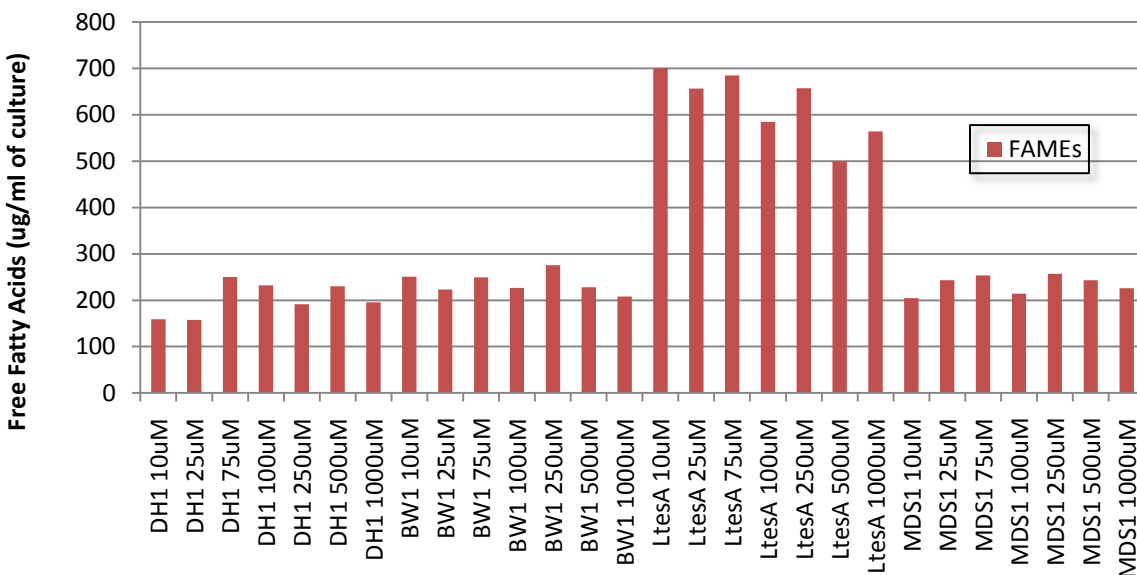


Figure 4-12: FAME production in '*tesA*' integrated strains.

4.4. Discussion

In an effort to diagnose FAEE pathway bottlenecks and generate a more stable strain than the three-plasmid A2A strain, we cloned the five pathway genes into a single plasmid system and added additional copies of each gene in the pathway. However in doing this we seemingly generated more instability into our system and had difficulty achieving the high production levels we had achieved in A2A. In a couple cases we produced nearly double the levels of A2A (1.2 g/L FAEE) and in others we produced less (200 mg/L). In understanding this instability we optimized production parameters and found that low concentrations of IPTG (10-75 uM) usually resulted in higher levels of FAEE production. We found that three plasmids were not retained at maximal levels beyond 12 hours, especially the high copy, ColE1 vector, which suggests plasmid loss may contribute to production instability. Thus, we investigated whether higher concentrations or additional

doses of antibiotics could result in higher, sustained production and indeed found this to be true. Other data suggested a potential problem with the thioesterase and requirement for precise expression levels, thus we also investigated whether integration of the thioesterase could provide a more stable system for free fatty acid production. Interestingly, the system did not seem to be IPTG-titratable, although *lacI* was present. This result corresponds to some of the leaky expression we have seen in the system and one fermentation result where production of FAEs reached 6 g/L and a yield of 8% without addition of IPTG.

With these results in mind, the next logical steps to take are attempting fermentations with higher concentrations of antibiotics for the 3-plasmid system, integrating the entire pathway or the most problematic portions of the pathway (*colE1-fadD-atfA-ampR*) into the chromosome. Efforts toward this end are currently underway and are promising directions for more stable production of FAEs.

4.5. References

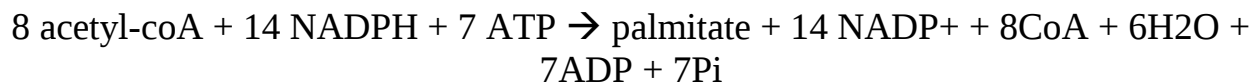
1. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-5 (2000).
2. Steen, E. J. et al. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-62.
3. Aldai, N., Osoro, K., Barron, L. J. & Najera, A. I. Gas-liquid chromatographic method for analysing complex mixtures of fatty acids including conjugated linoleic acids (cis9trans11 and trans10cis12 isomers) and long-chain (n-3 or n-6) polyunsaturated fatty acids. Application to the intramuscular fat of beef meat. *J Chromatogr A* **1110**, 133-9 (2006).
4. Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27**, 946-50 (2009).
5. Anthony, J. R. et al. Optimization of the mevalonate-based isoprenoid biosynthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metab Eng* **11**, 13-9 (2009).
6. Liu, T., Vora, H. & Khosla, C. Quantitative analysis and engineering of fatty acid biosynthesis in *E. coli*. *Metab Eng*.
7. Lennen, R. M., Braden, D. J., West, R. A., Dumesic, J. A. & Pfleger, B. F. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnol Bioeng* **106**, 193-202.
8. Posfai, G. et al. Emergent properties of reduced-genome *Escherichia coli*. *Science* **312**, 1044-6 (2006).

Chapter 5: Engineering yeasts for production of fatty-acid-derived fuels and chemical

Although engineering fatty acid biosynthesis in *E. coli* has been successful to a point, it is desirable to engineer the same properties in yeast due to its potentially superior qualities as an industrial host. Further, production of fatty acid ethyl esters in brewer's yeast may enable use of the native ethanol fermentation pathway. However, fatty acid biosynthesis in yeast is fundamentally different from the analogous pathway in *E. coli* and thus presents different challenges and potentially advantages for engineering. The most salient difference is use of the type I pathway, defined by single proteins that have multiple enzymatic activities that encompass the entire fatty acid pathway. This property is true of all eukaryotes and only a few bacteria. Even within type I fatty acid systems there are slight differences between organisms. For example mammalian type I systems contain a thioesterase and therefore release free fatty acids, whereas the yeast system ends with an acyl-transferase that transfers the final C16-C18 fatty acid bound to the fatty acid synthase to a CoA molecule. Further, in yeast, the acyl-CoAs are sequestered by the acyl-CoA binding protein (ACB1), compared to the *E. coli* system that maintains fatty acid thioesters covalently bound to acyl-carrier protein (ACP). Further, the eukaryotic system may provide advantages because some oleaginous yeasts have been demonstrated to accumulate fatty acids beyond 70% (w/v) and thus have a great capacity for production. However, there are some potential drawbacks of employing yeast for fatty acid biosynthesis. Because yeasts have a natural capacity to accumulate fatty acids in neutral lipids, there is a physical constraint imposed by the size of the cell. One strategy may be to express the fatty acid enzymes in a heterologous system like *E. coli* or remove yeast's fat storage capacity and express transporters that enable the fatty acid products to be secreted. Further, mechanisms of regulation exist that maintain stringent control of the pathway. Finally, acetyl-CoA, the precursor to not only fatty acid biosynthesis, but also other metabolic pathways of interest, is mostly sequestered in the mitochondria, which presents a challenge for engineering. Herein, we begin to explore the effects of overexpressing the FASs, removing β -oxidation, eliminating the ACB1 protein, and recapitulating the mechanism oleaginous yeasts use to accumulate fatty acids in *S. cerevisiae*.

Fatty acid biosynthesis in *S. cerevisiae* begins with the formation of malonyl-coA via carboxylation of acetyl-coA catalyzed by FAS3 (Figure 5-2). The elongation cycle begins by condensation of malonyl-acyl carrier protein (ACP) and acyl-ACPs to form a β -ketoacyl-ACP catalyzed by FAS2. The elongation cycle proceeds by reduction and subsequent removal of water catalyzed by an NADPH-

dependent β -ketoacyl-ACP reductase and β -hydroxyacyl-ACP dehydrase with FAS2 and FAS1. A final reduction occurs by enoyl-ACP reductase and forms acyl-ACP catalyzed by FAS1. The requirements for carbon, ATP, and NADPH can be derived from the stoichiometry of palmitate synthesis (C16:0):



To date, production of fatty acids has focused on synthesizing a cocoa butter equivalent² or fatty acid-derived hormones, which may be therapeutically beneficial. Both *S. cerevisiae* and *E. coli* have been utilized for fatty acid-based production platforms. The most abundant fatty acids found in *S. cerevisiae* are palmitic acid (C16:0), palmitoleic acid (C16:1), stearic acid (C18:0), and oleic acid (C18:1)³. However, a range of fatty acids have been identified, from C14 to C26⁴. This range of fatty acids provides the precursors to fuels that are directly compatible with current engine technology and could be directly integrated into a crude oil fraction or biodiesel after reduction of the carboxylic group.

Similar to the regulation in *E. coli* that is exerted by allosteric inhibition of the fatty acid synthases by acyl-ACP, yeast FAS3, the acetyl-CoA carboxylase, is inhibited by acyl-CoAs with $K_i = 5.5 \text{ nM}$ ⁵⁻⁷. Yeast fatty acyl-CoAs, the product of fatty acid biosynthesis, are sequestered by the acyl-CoA binding protein, ACB1, that has an affinity of $\sim 5 \text{ nM}$ for the acyl-CoAs⁸. In the same study, overexpression of *ACB1* resulted in a nearly 2-fold increase in the acyl-CoA pool to $6.54 \text{ nmol} / 5 \times 10^9 \text{ cells}$. Additionally, removal of the *ACB1* gene resulted in a 2-fold increase in acyl-CoA pools, where only the C18:0-CoA levels increased⁹. Acyl-CoAs are also thought to be involved in regulating transcription and metabolism^{10,11}. Thus, removal of these mechanisms of allosteric inhibition will be important for maximizing fatty acid biosynthesis in yeast. Beyond allosteric modulation of fatty acid biosynthesis, FAS3 is subject to AMP-dependent phosphorylation which reduces activity by at least 50% *in vivo*¹². Although the phosphorylation sites of higher eukaryotic ACCs have been identified, the sites in yeast remain unidentified. Disruption of *snf1*, the mammalian AMP-dependent kinase, results in sustained FAS3 activity, although the effect of FA synthesis was not directly measured. Therefore, identification and subsequent removal of the phosphorylation domains may be necessary in order to maintain maximal activity.

FASs are coordinately regulated¹³ at the level of transcription by two elements: constitutive upstream activation site (UAS) and inositol / choline responsive elements (ICRE) (Figure 4)^{14, 15}. Transcription of FASs is repressed in the presence of high concentrations of the phospholipid precursors inositol and choline¹⁶. When inositol is absent, the transcription factors INO2 and INO4 are transcribed and phosphatidic acid binds Opi1p in the ER membrane. Ino2p and Ino4p dimerize and bind to the ICRE, promoting transcription of the FASs. During stationary phase, phosphatidic acid levels drop, Opi1p is released and inhibits binding of Ino2p to the ICRE, therefore decreasing FAS transcription¹⁷.

The three fatty acid synthases (FASs) native to *S. cerevisiae* are responsible for conversion of the fatty acid precursor, acetyl-CoA, into acyl-ACPs. Previous over-expression of the individual genes FAS1, FAS2, and FAS3 has been achieved, although the co-over-expression has only been completed for FAS1 and FAS2. GC-MS analysis indicated that co-overexpression of FAS1 and FAS2 on high-copy vectors with constitutive promoters resulted in increased medium chain fatty acids (MCFAs), which contain 4-12 carbons, in comparison to wild-type^{18,19}. However, this result does not clarify whether the total fatty acid production is higher or there is a shift in the fatty acid profile towards MCFAs. These findings demonstrate that the FASs can be overexpressed and result in phenotypic change.

In order to overcome native transcriptional regulation, it is helpful to understand it. FASs are coordinately regulated¹³ at the level of transcription by two elements: constitutive upstream activation site (UAS) and inositol / choline responsive elements (ICRE) (Figure 5-1)^{14,15}. Transcription of FASs is repressed in the presence of high concentration of the phospholipid precursors inositol and choline¹⁶. When inositol is absent, the transcription factors INO2 and INO4 are transcribed and phosphatidic acid binds Opi1p in the ER membrane. Ino2p and Ino4p dimerize and bind to the ICRE, promoting transcription of the FASs. During stationary phase, phosphatidic acid levels drop, Opi1p is released and inhibits binding of ino2p to the ICRE, therefore decreasing FAS transcription¹⁷.

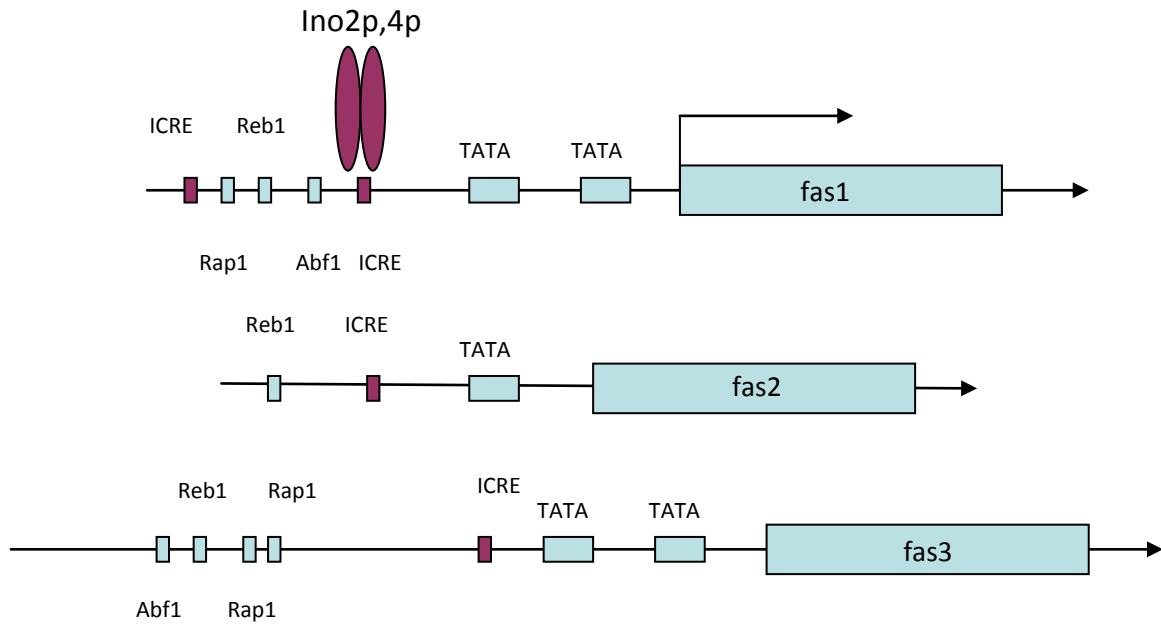


Figure 5-1: Transcriptional regulation elements upstream of the FASs. Abf1, Reb1, and Rap1 account for basal, constitutive transcription. ICRE accounts for regulation based on the inositol / choline response, where high levels of inositol repress transcription and low levels derepress transcription by dimerization and binding of Ino2p & 4p (modified from ¹).

FAS3 catalyzes the first committed step in fatty acid biosynthesis and is encoded by a 6.7-Kb gene and contains two enzymatic domains: biotin carboxylase, and biotin carboxyltransferase. Figure 5-2 illustrates the initiation reaction for fatty acid biosynthesis for which FAS3 is responsible, where biotin carboxyl carrier protein (BCCP) – biotin carboxylates acetyl-CoA to form malonyl-CoA.

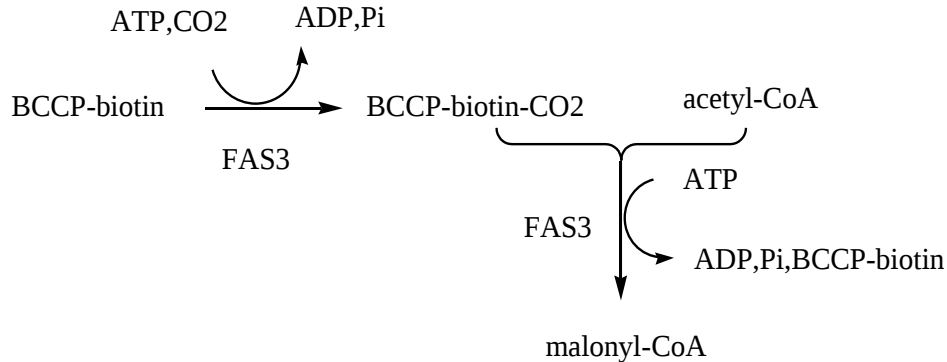


Figure 5-2: FAS3 reaction scheme, including the transfer of CO₂ to acetyl-CoA in forming malonyl-CoA

FAS2 is encoded by a 5.7-Kb gene and contains four domains: an acyl-carrier protein, beta-ketoacyl reductase, beta-ketoacyl synthase, and phosphopantetheinyl transferase (PPT). FAS1 is encoded by a 6.2-Kb gene and contains five domains: acetyltransacylase, dehydratase, enoyl reductase, malonyl transacylase, and palmitoyl transacylase. FAS1 and FAS2 complex to form a heterododecamer, containing six each of FAS1 and FAS2 subunits ¹⁹. Figure 5-3 illustrates the elongation reactions performed by FAS1 & 2.

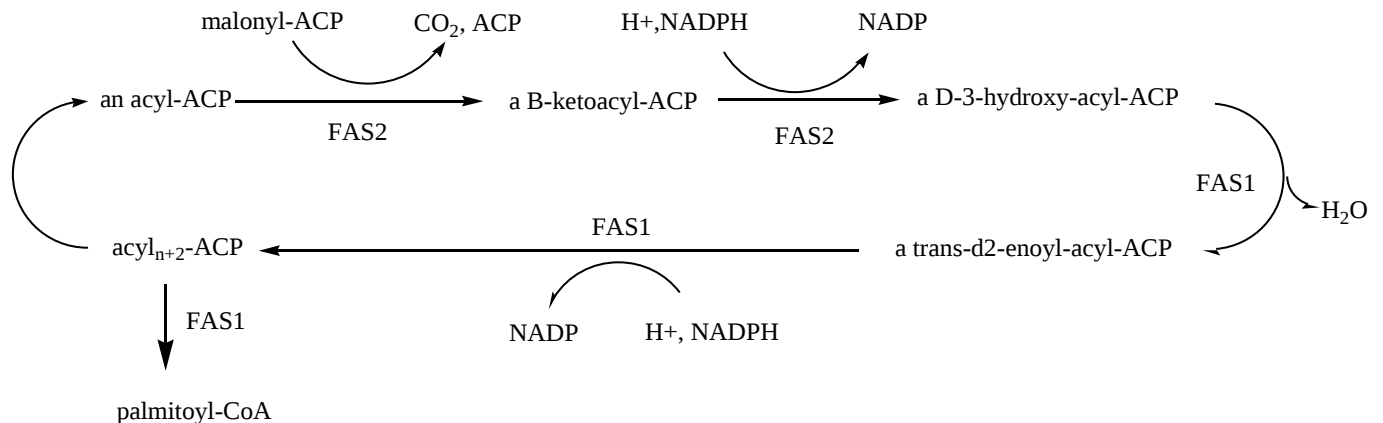


Figure 5-3: Fatty acid elongation cycle, initiated with an acyl-ACP which grows by one unit of malonyl-ACP.

Here, we describe engineering the *S. cerevisiae* type I fatty acid biosynthesis pathway by removing native transcriptional regulation exerted on FAS1, 2 and 3,

expressing an ATP:citrate lyase to increase cytosolic acetyl-CoA, and necessary deletions of the acyl-CoA binding protein (ACB1) and β -oxidation in order to produce fatty acids or fatty acid-derived products like fatty acid ethyl esters.

5.2. Methods

Chemicals

Ethyl acetate was purchased from Sigma-Aldrich (St. Louis, MO). Complete Supplement Mixtures for formulation of Synthetic Defined (SD) media were purchased from Qbiogene (Irvine, CA). All other media components were purchased from Sigma-Aldrich.

Strains and media

Clostridium beijerinckii NCIMB 8052 was purchased from ATCC, catalog number 51743. *E. coli* strains DH10B and DH5 α were used for bacterial transformation and plasmid amplification in the construction of the expression plasmids used in this study. The strains were cultivated at 37°C in Luria-Bertani medium with 100 mg L⁻¹ ampicillin. *S. cerevisiae* strain BY4742²⁰, a derivative of S288C, was used as the parent strain for all yeast strains. This strain was grown in rich YPD medium at 30°C²¹. Engineered yeast strains were grown in SD medium²¹ with leucine, uracil, histidine, and/or methionine dropped out where appropriate. For induction of genes expressed from the *GAL1* and *GAL10* promoters, *S. cerevisiae* strains were grown in 2% galactose as the sole carbon source unless otherwise indicated.

Plasmid construction

C. beijerinckii genes were cloned from genomic DNA: *hbd* encodes 3-hydroxybutyryl-CoA dehydrogenase; *crt*, crotonase; *bcd*, butyryl-CoA dehydrogenase; and *etfA* & *etfB*, two-electron transferring flavoproteins A & B. *phaA* and *phaB* (*Ralstonia eutropha*), *adhe2* (*C. beijerinckii*), and *ccr* (*Streptomyces collinus*) were synthesized (Epoch Biolabs). All genes were PCR amplified with Phusion polymerase (New England Biolabs). Primers were designed to have 30-bp flanking regions homologous to the plasmid insertion regions, either the *gal1* or *gal10* promoter and the *CYC1*, *ADH1*, or *PGK1* terminator. Plasmid construction was carried out using the Sequence and Ligation Independent Cloning (SLIC) method, previously described²². The constructed plasmids were derived from pADS-AMO-CPR-opt-LEU2D plasmid²³ and pESC-HIS (Stratagene).

Yeast transformation and strain construction

Transformation of all *S. cerevisiae* strains was performed using the lithium acetate method²⁴. Strains ESY1-11 were constructed by the co-transformation of the indicated plasmids followed by selection on SD-LEU or SD-LEU-HIS plates as appropriate.

Yeast cultivation

All optical density measurements at 600 nm (OD₆₀₀) were taken using a Beckman DU-640 spectrophotometer. To measure n-butanol production, culture tubes containing 5 mL of SD (2% galactose) medium (with appropriate amino acid omissions as described above) were inoculated with the strains of interest. These inocula were grown at 30°C to an OD₆₀₀ between 1 and 2. Capped serum vials (100 mL) containing 50 mL SD medium were inoculated to an OD₆₀₀ 0.05 with these seed cultures in order to achieve a "semi" anaerobic condition. Samples were collected at 24, 72, 120, and 144 h and analyzed for metabolites as discussed below.

5.3. Preliminary Results

5.3.1. Overexpressing FAS3 & FAS1/2

As an initial attempt to increase fatty acid biosynthesis, we overexpressed the FASs using the common high-copy pESC vectors²⁵ (Figure 5-4). Since the native FASs demonstrate lower activity during stationary phase due to ICRES, it seems plausible that expression of the FASs using a plasmid based system may alleviate repression and result in increased fatty acid production.

The three FASs were cloned into three different pESC vectors, each carrying a different auxotrophic marker. Strains were analyzed by GC-MS for differences in the C16 and C18 saturated and unsaturated fatty acids. Generally, there were not major gains in production levels in the induced and uninduced cultures, suggesting that the other mechanisms of regulation are likely more important. However, a slight gain in production may have been achieved by overexpressing FAS3 and may be supplying more malonyl-CoA for fatty acid production.

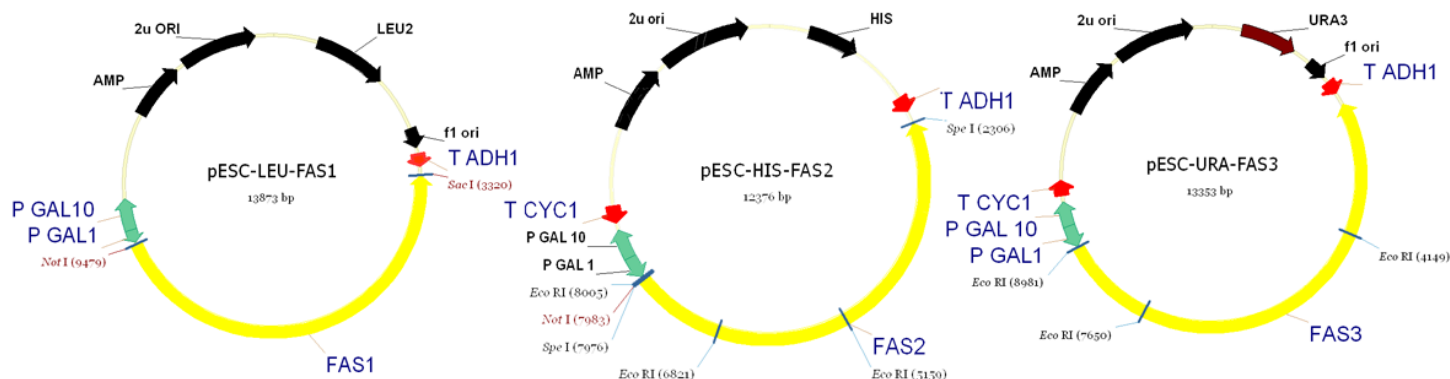


Figure 5-4: The 3 FASs are cloned into three pESC vectors with different auxotrophic markers (leu, his, & ura). Transcription of the genes are driven by the GAL promoter and terminated by the ADH1 terminator. 30 copies of the pESC vector are retained within each yeast cell due to the 2-micron origin. The vectors also contain the appropriate markers for replication in *E. coli*.

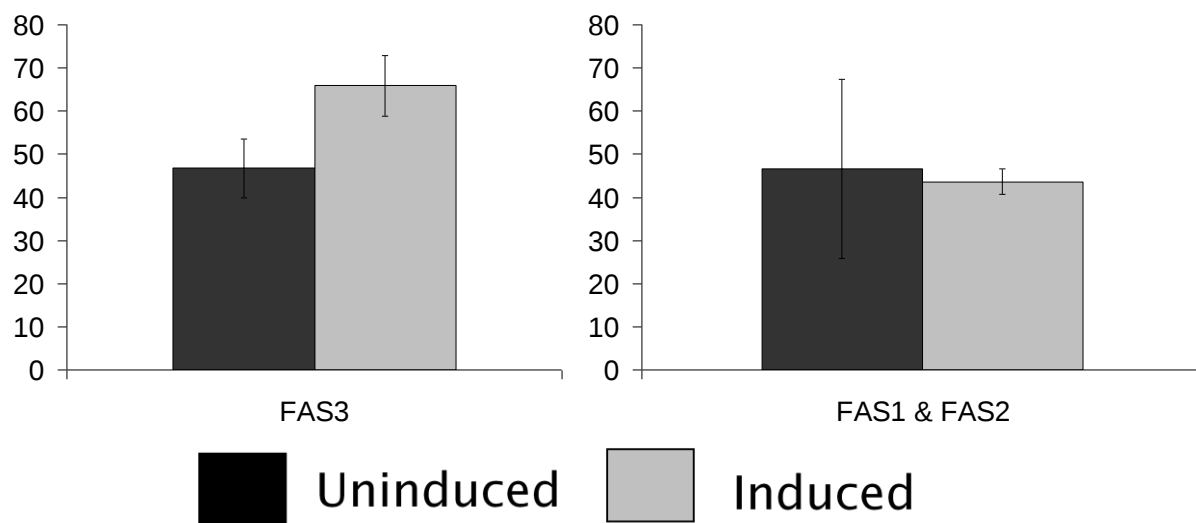


Figure 5-5: Fatty acid production from strains overexpressing FAS1, 2, & 3. Values and error bars are the mean and standard deviation of triplicate experiments and represent relative values.

5.3.2. Removal of neutral lipid storage

Yeasts are known to accumulate neutral lipids to different extents depending upon species. The two major classes of neutral lipids are steryl esters (STEs) and triacylglycerols (TAGs), which *S. cerevisiae* accumulates to approximately 10% (w/w). STEs biosynthesis requires the genes *ARE1* and *ARE2*; TAG biosynthesis requires the genes *DGA1* and *LRO1*. These genes have been disrupted and yeast remains viable²⁶.

5.3.3. Removal of competing pathways (β -oxidation)

In order to prevent degradation of fatty acids to acetyl-CoA and ATP it will be necessary to remove the beta-oxidation pathway, most of which is localized to the peroxisome in *S. cerevisiae*. We removed the gene, *POX1*, which encodes an acyl-CoA oxidase, replacing it with *NAT1*, a gene that confers resistance to nourseothricin.

5.3.4. Removal of the acyl-CoA binding protein ACB1

In order to enzymatically access the pool of acyl-CoAs, we hypothesized it would be necessary to remove the acyl-CoA binding protein. Thus, we deleted the gene *ACB1*, replacing it with the *KAN* gene, which confers resistance to G418, in the H1246 background, provided by Stymne and colleagues. This strain exhibited initial impaired growth, but seems to recover after multiple passes as previously reported for a single knockout strain of *ACB1*⁹.

5.3.5. Expressing the wax ester synthase gene, *atfA*

In order to produce FAEEs or other fatty acid esters from *S. cerevisiae*'s native supply of both fatty acids and short chain alcohols (derived from either the ethanol pathway or the fusel alcohol pathways), we expressed *atfA* in the H1246 background and our newly constructed H1246 with *ACB1* and *POX1* removed. We assayed for production of fatty acid esters and free fatty acids and did not see accumulation of fatty acid esters after 120 h (data not shown). However, this result does not correspond with previous expression of *atfA* in H1246, where fatty acid esters were detected²⁷, which suggests an expression problem or timing of analysis. Further, previous expression of *atfA* in yeast was coupled with feeding oleic acid, thus yeast's native regulation may still be impairing appreciable levels of acyl-CoAs to accumulate and provide enough substrate for the wax ester synthase.

5.3.6. Expressing ATP:citrate lyase and malic enzyme

We were interested in exploiting the mechanism that oleaginous yeasts use to accumulate significant amounts of lipids and shunt acetyl-CoA into the cytosol

during nitrogen limitation²⁸⁻³³. These organisms are able to accumulate lipid at high levels upon nitrogen limitation due to a regulation mechanism that results in inactivation of the TCA enzyme isocitrate dehydrogenase due to decreased levels of AMP, resulting from a nitrogen scavenging mechanism, where a deaminase cleaves to produce NH_3 . Isocitrate levels increase, backing up the TCA cycle, resulting in high levels of citrate, which rapidly equilibrates across the mitochondrion into the cytosol by the citrate-malate shuttle. ATP-citrate lyase (ACL) then cleaves citrate, forming acetyl-CoA and oxaloacetate. ACL is absent in *S. cerevisiae*, so engineering a similar source of increased acetyl-CoA would require heterologous expression of an ACL. Further, malic enzyme is localized to the mitochondrion. Upon expression of an ACL from *A. nidulans* from the pESC-his plasmid (Stratagene) in a strain engineered for amorphaadiene production (an isoprenoid compound) (ESY231), we achieved a 38% increase in production compared to the control (EPY230) in 2% galactose (Figure 5-7). Further expression of a heterologous isocitrate dehydrogenase and elimination of glycogen accumulation will aid in increasing cytosolic acetyl-CoA pools upon nitrogen limitation.

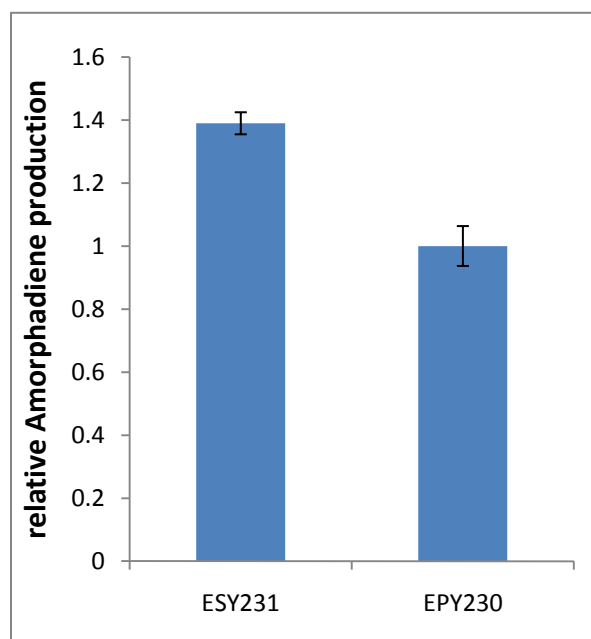


Figure 5-6: Increased acetyl-CoA levels and final production of amorphaadiene upon expressing an ATP citrate lyase

5.4. Discussion

Attempts at removing the transcriptional regulation of the FASs in *S. cerevisiae* did not have a major effect on fatty acid production levels. This is due to the more stringent allosteric control exerted by the acyl-CoAs on the FASs. Thus, we removed the acyl-CoA binding protein that had been previously described as a strain that exhibits slow growth and then recovers to a wild-type phenotype. We reasoned that this removal and resulting higher levels of acyl-CoAs may provide selective pressure for yeast to adapt in a way that fatty acid biosynthesis continues and is relieved of the allosteric inhibition. Although this may be the case, initial attempts at removing β -oxidation, neutral lipid accumulation and expressing the wax ester synthase *atfA* in order to make FAEEs were not successful. This may be due to a low level of acyl-CoAs due to clearance by endogenous hydrolases. In contrast, our attempts to increase cytosolic flux through acetyl-CoA by expressing the ACL from *A. nidulans* did serve to increase amorphadiene production compared to the control strain, suggesting that it serves to increase cytosolic acetyl-CoA levels. Although production of FAEEs was not initially detected, this result is not congruent with previous reports in *S. cerevisiae*. Thus, troubleshooting should be pursued to verify its expression and confirm its activity. Further, only one final timepoint was taken to measure both free fatty acids and the ethyl esters. Finally, nitrogen-limited conditions should be pursued in order to slow the TCA cycle and increase the available cytosolic citrate as the substrate for the ACL.

5.5. References

1. Veen, M., Lang, C. Production of lipid compounds in the yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **63**, 635-46 (2004).
2. Ratledge, C. Single cell oils--have they a biotechnological future? *Trends Biotechnol* **11**, 278-84 (1993).
3. Daum, G., Lees, N. D., Bard, M. & Dickson, R. Biochemistry, cell biology and molecular biology of lipids of *Saccharomyces cerevisiae*. *Yeast* **14**, 1471-510 (1998).
4. Schneider, R. & Kohlwein, S. D. Organelle structure, function, and inheritance in yeast: a role for fatty acid synthesis? *Cell* **88**, 431-4 (1997).
5. Nikawa J, T. T., Ogiwara H, Shiba T, Numa S. Inhibitory effects of long-chain acyl coenzyme A analogues on rat liver acetyl coenzyme A carboxylase.: *FEBS Lett.* **102**, 223-6 (1979).
6. Ogiwara H, T. T., Nikawa J, Numa S. Inhibition of rat-liver acetyl-coenzyme-A carboxylase by palmitoyl-coenzyme A. *Eur J Biochem* **89**, 33-41 (1978).
7. Sumper, M. & Trauble, H. Membranes as acceptors for palmitoyl CoA in fatty acid biosynthesis. *FEBS Lett* **30**, 29-34 (1973).

8. J Knudsen, N. J. F., H Skøtt, R Hummel, C Børsting, T M Rose, J S Andersen, P Højrup, P Roepstorff, and K Kristiansen. Yeast acyl-CoA-binding protein: acyl-CoA-binding affinity and effect on intracellular acyl-CoA pool size. *Biochem J.* **302**, 479–485 (1994).
9. Schjerling, C. K. et al. Disruption of the gene encoding the acyl-CoA-binding protein (ACB1) perturbs acyl-CoA metabolism in *Saccharomyces cerevisiae*. *J Biol Chem* **271**, 22514-21 (1996).
10. Knudsen, J., Neergaard, T. B., Gaigg, B., Jensen, M. V. & Hansen, J. K. Role of acyl-CoA binding protein in acyl-CoA metabolism and acyl-CoA-mediated cell signaling. *J Nutr* **130**, 294S-298S (2000).
11. Faergeman, N. J. & Knudsen, J. Role of long-chain fatty acyl-CoA esters in the regulation of metabolism and in cell signalling. *Biochem J* **323 (Pt 1)**, 1-12 (1997).
12. Woods, A. et al. Yeast SNF1 is functionally related to mammalian AMP-activated protein kinase and regulates acetyl-CoA carboxylase in vivo. *J Biol Chem* **269**, 19509-15 (1994).
13. Chirala, S. S. Coordinated regulation and inositol-mediated and fatty acid-mediated repression of fatty acid synthase genes in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **89**, 10232-6 (1992).
14. Schuller, H. J., Schutz, A., Knab, S., Hoffmann, B. & Schweizer, E. Importance of general regulatory factors Rap1p, Abf1p and Reb1p for the activation of yeast fatty acid synthase genes FAS1 and FAS2. *Eur J Biochem* **225**, 213-22 (1994).
15. Chirala, S. S., Zhong, Q., Huang, W. & al-Feel, W. Analysis of FAS3/ACC regulatory region of *Saccharomyces cerevisiae*: identification of a functional UASINO and sequences responsible for fatty acid mediated repression. *Nucleic Acids Res* **22**, 412-8 (1994).
16. Griac, P. & Henry, S. A. The yeast inositol-sensitive upstream activating sequence, UASINO, responds to nitrogen availability. *Nucleic Acids Res* **27**, 2043-50 (1999).
17. Chen, M., Hancock, L. C. & Lopes, J. M. Transcriptional regulation of yeast phospholipid biosynthetic genes. *Biochim Biophys Acta* **1771**, 310-21 (2007).
18. Furukawa, K., Yamada, T., Mizoguchi, H. & Hara, S. Increased ethyl caproate production by inositol limitation in *Saccharomyces cerevisiae*. *J Biosci Bioeng* **95**, 448-54 (2003).
19. Lomakin, I. B., Xiong, Y. & Steitz, T. A. The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. *Cell* **129**, 319-32 (2007).
20. Brachmann, C. B. et al. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast* **14**, 115-32 (1998).
21. Burke D, D. D., Stearns T. . Methods in yeast genetics: a Cold Spring Harbor Laboratory Course Manual. Plainview, NY. (2000).
22. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods* **4**, 251-6 (2007).
23. Ro, D. K. et al. Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnol* **8**, 83 (2008).
24. Gietz, R. RA guide to yeast genetics and molecular biology and cell biology. Part B. San Diego, CA: Academic Press, Inc., 87-96 (2002).
25. Futcher, A. B. & Cox, B. S. Copy number and the stability of 2-micron circle-based artificial plasmids of *Saccharomyces cerevisiae*. *J Bacteriol* **157**, 283-90 (1984).
26. Sandager, L. et al. Storage lipid synthesis is non-essential in yeast. *J Biol Chem* **277**, 6478-82 (2002).
27. Kalscheuer, R., Luftmann, H. & Steinbuchel, A. Synthesis of novel lipids in *Saccharomyces cerevisiae* by heterologous expression of an unspecific bacterial acyltransferase. *Appl Environ Microbiol* **70**, 7119-25 (2004).

28. Hall, M. J. & Ratledge, C. Lipid Accumulation in an Oleaginous Yeast (Candida 107) Growing on Glucose Under Various Conditions in a One- and Two-Stage Continuous Culture. *Appl Environ Microbiol* **33**, 577-584 (1977).
29. A Ykema, E. V., HJH Nijkamp, H Smit. Isolation and characterization of fatty acid auxotrophs from the oleaginous yeast *Apiotrichum curvatum*. *Appl Microbiol Biotechnol*, 76-84 (1989).
30. M Hassan, P. B., L Granger, A Pareilleux, G Goma. Lipid production by an unsaturated fatty acid auxotroph of the oleaginous yeast *Apiotrichum curvatum* grown in a single-state continuous culture. *Appl Microbiol Biotechnol*, 483-488 (1993).
31. Gangar, A., Karande, A. A. & Rajasekharan, R. Isolation and localization of a cytosolic 10 S triacylglycerol biosynthetic multienzyme complex from oleaginous yeast. *J Biol Chem* **276**, 10290-8 (2001).
32. Botham, P. A. & Ratledge, C. A biochemical explanation for lipid accumulation in Candida 107 and other oleaginous micro-organisms. *J Gen Microbiol* **114**, 361-75 (1979).
33. Gill, C. O., Hall, M. J. & Ratledge, C. Lipid accumulation in an oleaginous yeast (Candida 107) growing on glucose in single-stage continuous culture. *Appl Environ Microbiol* **33**, 231-9 (1977).

Chapter 6: Diacid production from fatty acid biosynthesis

6.1. Introduction

Specific production of a wide variety of valuable, fatty acid (FA) derived compounds can be achieved in microbial hosts. These molecules could be used for alternative fuels sources or chemicals in a wide variety of processes and are relevant to the medical / health field as a source of (omega-3 fatty acids) or the industrial chemicals polymer field as a source of diacids, diamines, etc. We demonstrate production of diacids from simple sugar or biomass by production of fatty acids and subsequent omega oxidation using existing or engineered P450 oxidases. Aliphatic diacids are used in a wide range of applications used to make polyesters, resins, polyamides, nylon, fuel additives, lubricants, rapid drying paints, varnishes, etc. C12 diacid is the most abundantly produced, greater than 40MM lb/yr and sells for roughly \$2/lb¹.

A suite of diacids ranging in chain length from C4 to C22 can be produced microbially by providing the requisite fatty acid of any given chain length in addition to a P450 that oxidizes the fatty acid at the omega carbon with further oxidases / dehydrogenases that oxidize the alcohol to an acid. Prior art has demonstrated that yeasts (*Candida tropicalis*) can successfully convert fatty acids and alkanes into diacids by their innate omega oxidation pathway after removal of the β -oxidation pathway². The industrial success of this report however is greatly disadvantaged by its reliance upon exogenous alkane, FA, or hydroxy-FA supplementation and the inability of the microbe to produce enough substrate for conversion into diacid (or other product). Additionally, the report does not demonstrate a way of controlling the final diacid product formed other than by controlling what chain length alkane / FA is fed. Thus, a way of both producing FA substrate, being able to control the chain length of that substrate and being able to oxidize the omega carbon to a carboxylate group would provide an economical, competitive route to diacids, diamines, etc. We hypothesized that our previous work demonstrating that different thioesterases could cleave specific chain length fatty acids could be employed for production of specific chain length diacids as well.

6.2 Methods

Two gene expression vectors were utilized to express the thioesterase, *l'tesA*, and the P450 hydroxylase. The thioesterase was expressed on a low-copy plasmid with pBBR origin of replication and tetracycline resistance, pKS2, and either P450-Bm3

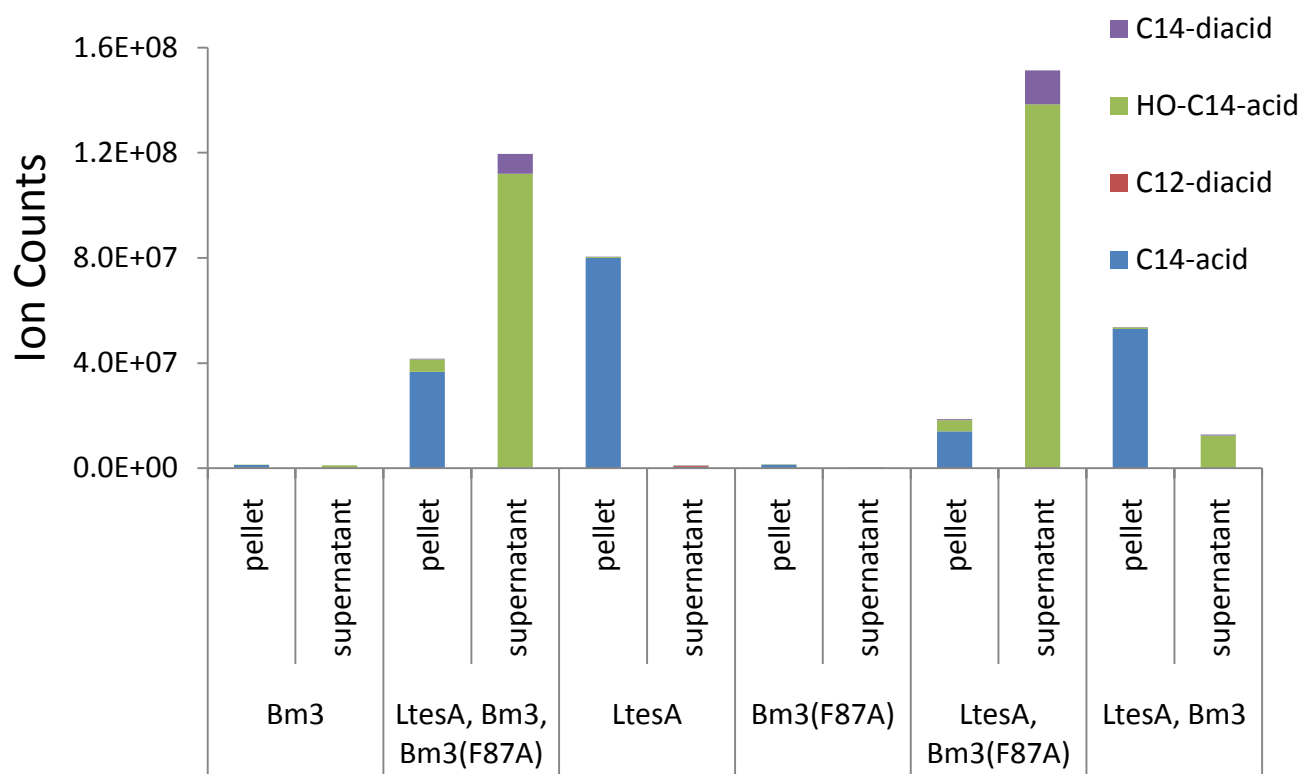
or P450-Bm3(F87A) was expressed on the medium copy p15 origin of replication plasmid bearing ampicillin resistance. Strain DH1-*fadDKO* was transformed with both pKS2 and either p450-Bm3 or p450-Bm3(F87A). These cultures were grown in LB medium overnight (16 h), and then resuspended to OD 1.0 in fresh LB with appropriate antibiotics (100 ug/mL carbenicillin, 10 ug/mL tetracycline) and 1 mM IPTG. Co-cultures (pKS2+p450-Bm3 and pKS2+p450-Bm3(F87A)) were resuspended to OD 0.5 such that the total final OD was 1.0. Cultures were grown at 30°C for 72 h post induction. For assaying production, 1 mL of culture was centrifuged, 30 s, 20,000 x g and the supernatant was separated from the pellet. The supernatant was directly filtered on a 30,000 MW cut-off filter (Microcon) and the pellet was extracted with 0.5 mL of MeOH by vortexing 30 s and then centrifuging, 30 s, 20,000 x g. Finally, 200 uL of the MeOH supernatant was recovered into 200 uL of ddH₂O, which was filtered on a 30,000 MW cut-off filter. Both samples were separated by liquid chromatography and analyzed by mass spectroscopy to identify the parent ion. Tetradecanoic acid was identified with mass = 227.195; hydroxytetradecanoic acid was identified with mass = 243.198; 1,14 tetradecanoic diacid was identified with mass = 257.179. Chemical standards were made up to 20 µM, in methanol and water (1:1, v/v). The separation of fatty acids was conducted on a ZIC-HILIC column (250 mm length, 2.1 mm internal diameter, and 3.5 µm particle size; from Merck SeQuant, and distributed via The Nest Group, Inc., MA., USA) using an Agilent Technologies 1200 Series HPLC system (Agilent Technologies, CA, USA). An injection volume of 4 µL was used throughout. The auto-sample tray was maintained at 4°C by an Agilent FC/ALS Thermostat. The column compartment was set to 50°C. Analytes were eluted isocratically with a mobile phase composition of 50 mM ammonium acetate, in water, and acetonitrile (3.6:6.4, v/v). A flow rate of 0.1 mL/min was used throughout.

The HPLC system was coupled to an Agilent Technologies 6210 time-of-flight mass spectrometer (LC-TOF MS), by a 1/3 post-column split. Contact between both instrument set-ups was established by a LAN card in order to trigger the MS into operation upon the initiation of a run cycle from the MassHunter workstation (Agilent Technologies, CA, USA). Electrospray ionization (ESI) was conducted in the negative ion mode and a capillary voltage of - 3500 V was utilized. MS experiments were carried out in full scan mode, at 0.85 spectra/ second and a cycle time of 1.176 seconds, for the detection of $[M - H]^-$ ions. The instrument was tuned for a range of 50 - 1700 m/z. Prior to LC-TOF MS analysis, the TOF MS was calibrated via an ESI-L-low concentration tuning mix (Agilent Technologies, CA, USA). Data acquisition and processing were performed by the MassHunter software package (Agilent Technologies, CA, USA).

6.3. Results

6.3.1. Production of C14 acid, hydroxyacid and diacids.

Expression of the gene encoding the thioesterase, *tesA* from the pBBR1 vector in conjunction with a P450 hydroxylase from the pCWori vector results in production of free fatty acids and hydroxylation of them. We utilized two hydroxylases, P450-Bm3 and a mutant, P450-Bm3(F87A) to perform the terminal, omega hydroxylation and hypothesized that further oxidation may proceed due to previous reports³ or native alcohol and aldehyde dehydrogenases in *E. coli*. Expression of P450-Bm3 or P450-Bm3(F87A) alone resulted in production of hydroxytetradecanoic acid, although our detection did not differentiate between hydroxylation positions of omega -1, -2, -3, yet expression did not result in production of detectable levels of the diacid. Coupling expression of the P450 hydroxylases with expression of a thioesterase resulted in production of detectable levels of both the hydroxytetradecanoic acid and tetradecanoic diacid. The best producer of tetradecanoic diacid in this case was expression of the mutated p450-Bm3(F87A), which greatly increases production compared to p450-Bm3. The co-culture of L '*tesA*, p450-Bm3 and L '*tesA*, p450-Bm3(F87A) did not result in the highest levels of diacid production. The products were detected in both the pellet and supernatant fractions, although to different degrees (Figure 6-1).



6.4. Discussion

Modifying fatty acids to make other, more valuable chemicals with different properties is a compelling strategy to generate renewable biochemicals. Herein we focused on proof-of-principle experiments to determine if coupling high-level production of free fatty acids with expression of hydroxylases could result in production of diacids, a valuable polymer precursor. Diacids have been produced naturally by yeasts in the past, but have relied on a petroleum based feedstock or fatty acids themselves^{2,4}. Our approach sought to produce diacids directly from a sugar source. This was achieved by co-expressing a thioesterase that was necessary to produce high levels of free fatty acids as previously shown⁵⁻⁹ and a P450 hydroxylase.

Our results indicate that sole expression of a hydroxylase does not result in production of diacids, but there are detectable levels of hydroxyacids. However,

additional expression of the thioesterase does result in diacid production, suggesting that the native fatty acid pool limits the availability of substrate for the hydroxylase. Further, higher levels of both diacid and hydroxyacid production were achieved for the mutant hydroxylase P450-Bm3(F87A), suggesting that this mutated form is a better enzyme choice for catalyzing the hydroxylation and potentially subsequent oxidations compared to the wild-type P450-Bm3. Because this hydroxylase has been reported to catalyze hydroxylation at multiple carbons close to the omega-carbon and can be modified by the F87A mutation¹⁰, it may not be the optimal enzyme choice as w-1, w-2 and w-3 hydroxylations will ultimately be lost from the available substrate to generate diacids. Better suited P450 hydroxylases may be derived from the mammalian class 4 types that are reported to mainly hydroxylate the omega position¹¹⁻¹³. Finally, it is unclear if the P450 hydroxylase is completing the entire hydroxylation and oxidations to produce diacids or whether a native *E. coli* alcohol and aldehyde dehydrogenase may be responsible. Future in vitro biochemical studies may elucidate these questions and provide insight into the next steps for engineering a more efficient conversion of the hydroxyacid into the diacid, the clear bottleneck from these initial experiments.

6.5. References

1. Mobley, D. Biosynthesis of long-chain dicarboxylic acid monomers from renewable resources. *US Department of Energy Report* (1999).
2. Picataggio, S. et al. Metabolic engineering of *Candida tropicalis* for the production of long-chain dicarboxylic acids. *Biotechnology (N Y)* **10**, 894-8 (1992).
3. Dietrich, J. A. et al. A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450(BM3). *ACS Chem Biol* **4**, 261-7 (2009).
4. Liu, W. H. & Kuo, C. K. Production of dodecanedioic acid from n-dodecane by yeasts. *Zhonghua Min Guo Wei Sheng Wu Ji Mian Yi Xue Za Zhi* **22**, 242-8 (1989).
5. Steen, E. J. et al. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**, 559-62.
6. Dehesh, K., Jones, A., Knutzon, D. S. & Voelker, T. A. Production of high levels of 8:0 and 10:0 fatty acids in transgenic canola by overexpression of Ch FatB2, a thioesterase cDNA from *Cuphea hookeriana*. *Plant J* **9**, 167-72 (1996).
7. Voelker, T. A., Jones, A., Cranmer, A. M., Davies, H. M. & Knutzon, D. S. Broad-range and binary-range acyl-acyl-carrier protein thioesterases suggest an alternative mechanism for medium-chain production in seeds. *Plant Physiol* **114**, 669-77 (1997).
8. Voelker, T. A. & Davies, H. M. 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. *J Bacteriol* **176**, 7320-7 (1994).
9. Voelker, T. A. et al. Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants. *Science* **257**, 72-4 (1992).

10. Oliver, C. F. et al. A single mutation in cytochrome P450 BM3 changes substrate orientation in a catalytic intermediate and the regiospecificity of hydroxylation. *Biochemistry* **36**, 1567-72 (1997).
11. CaJacob, C. A., Chan, W. K., Shephard, E. & Ortiz de Montellano, P. R. The catalytic site of rat hepatic lauric acid omega-hydroxylase. Protein versus prosthetic heme alkylation in the omega-hydroxylation of acetylenic fatty acids. *J Biol Chem* **263**, 18640-9 (1988).
12. Alterman, M. A., Chaurasia, C. S., Lu, P., Hardwick, J. P. & Hanzlik, R. P. Fatty acid discrimination and omega-hydroxylation by cytochrome P450 4A1 and a cytochrome P4504A1/NADPH-P450 reductase fusion protein. *Arch Biochem Biophys* **320**, 289-96 (1995).
13. Bambal, R. B. & Hanzlik, R. P. Effects of steric bulk and conformational rigidity on fatty acid omega hydroxylation by a cytochrome P450 4A1 fusion protein. *Arch Biochem Biophys* **334**, 59-66 (1996).