Engineering free fatty acid and free fatty acid derived chemical production in *Escherichia coli*

by

J. Tyler Youngquist

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Chemical and Biological Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2013

Date of final oral examination: 06/27/13

The dissertation is approved by the following members of the Final Oral Committee:
Brian F. Pfleger, Assistant Professor, Chemical and Biological Engineering
Regina M. Murphy, Smith-Bascom Professor, Chemical and Biological Engineering
Sean P. Palecek, Professor, Chemical and Biological Engineering
Eric V. Shusta, Professor, Chemical and Biological Engineering
James L. Steele, Professor, Food Science

Abstract

To alleviate the demand for products synthesized via finite petroleum sources, research into renewable alternatives is desirable to meet future needs. Current production of short chain alcohols, such as ethanol, provides one alternative to light transportation fuels. However, the heavy transportation sector requires a higher energy density liquid fuel and numerous chemical products (polymers, etc), areas that are not ideally served by short chain alcohols. Microbial production of these compounds is attractive owing to native synthesis of long, primarily hydrocarbon compounds (acyl chains) in the synthesis of the cellular membrane. Using biomass derived sugars in conjunction with minimal alterations to a native host's pathway provides a promising alternative route to producing substitutes for fuels and petrochemicals.

The aforementioned challenges are tackled in this thesis by applying synthetic biology principles to develop a metabolic engineering strategy in *Escherichia coli*. The fatty acid biosynthetic pathway is used as a basis for overproducing medium chain length fatty acids (10-14 carbon chain length), which are used as intermediates for the production of fuels and chemicals. A thioesterase (TE) hydrolyzes acyl-acyl carrier protein thioester bonds to alleviate feedback inhibition in the fatty acid biosynthesis pathway and produce a free fatty acid (FFA). Previous efforts in the lab have resulted in the selection of the TE to use in FFA production and the identification of stresses placed on *E. coli* during periods of high FFA production. However, minimal work in the field has been done to determine the ideal external production conditions for FFA biosynthesis.

For FFA derived fuels and chemicals to be competitive in the marketplace, it is necessary to produce them at high productivities close to the theoretical maximum yield. In this thesis, controlled reactor and media conditions were investigated as a strategy to find the optimum

production regime for FFA in an engineered strain of *E. coli*. To that end, an antibiotic free, FFA overproducing strain of *E. coli* was engineered and then cultivated in continuous culture to generate kinetic parameters for a model of FFA production. The data and model were used to identify growth regimes with high productivity and yield as well as to predict FFA production in alternative reactor conditions. Investigation into alternative nutrient limitation as a means to increase FFA yield also had a significant impact. The use of phosphate as a limiting nutrient tripled specific productivity and provided a means to produce FFA consistently during stationary phase. Using media and reactor optimization strategies, specific productivity reached as high as 0.065 g/gDCW/h and yield surpassed 40% of maximum theoretical values.

As production of low price, high volume chemicals such as fuels would require near maximum yield and high productivities, converting the initial FFA to a more valuable product would seem a logical option to achieve economic viability. Used in detergents, cosmetics, and emulsifiers, fatty alcohols represent over a 3 billion dollar market. Additionally, fatty alcohols with C_{10} to C_{14} chain lengths are sold at ~$1.00 per pound compared to diesel at ~$0.40 per pound, indicating a significant advantage in their production. However, previous efforts to produce fatty alcohols endogenously through the FFA biosynthetic route have been unable to reach 10% of theoretical maximum yield, leaving significant amounts of unconsumed FFAs. Expression of a TE, an acyl-CoA synthetase (ACS), and an acyl-CoA reductase (ACR) resulted in the production of medium chain length fatty alcohols. By balancing the expression of these genes and applying the optimal reactor and media conditions for FFA overproduction, yields of over 40% of theoretical maximum were achieved with over 26% of fed carbon going to fatty alcohol production. Additionally, residual FFA levels represented less than 5% of the combined titer between fatty alcohols and FFAs.

To close in on maximum theoretical yield, numerous strategies have been pursued, including: functional genomics to provide hypotheses for improvement areas, alterations of global transcriptional regulators, and introduction of heterologous pathways bypassing regulation of FFA biosynthesis. Elimination of flagella synthesis, a target from microarray studies, was successful in producing an 11% increase in yield. However, all strategies have been unsuccessful in achieving yields above 45% of maximum theoretical. Still, functional genomic analysis has provided a good picture of how FFA overproducing *E. coli* starved for phosphate alters gene expression in response to stresses. Additionally, the dataset has provided insight into how one could down regulate certain central metabolic pathways to shift carbon flux away from CO_2 and toward FFA biosynthesis.

Overall, this thesis demonstrates an effective strategy for determining reactor conditions for optimum FFA productivity and yield as well as for balancing enzymatic expression to achieve full conversion of FFA to fatty alcohols. Phosphate limitation was identified as a media condition to significantly enhance FFA production. A fatty alcohol yield more than two fold greater than the highest published value was achieved. Strategies to further increase productivity and yield, including identifying target pathways to reduce carbon flux to CO_2, were pursued. While some strategies provided minor improvements to yield, further work is necessary to find conditions and the engineered strain required to reach theoretical maximum yield.

Acknowledgements

When I first started graduate school I had minimal knowledge of the state of Wisconsin or molecular and synthetic biology. Throughout my many (5) years here I have gained knowledge in both and would like to thank all of those who have helped (and tolerated) me along the way.

I would like to thank all the collaborators from other labs and the GLBRC who have provided their time to help me in many areas, including but not limited to: sample processing, training on instruments and techniques, and providing input to experiments. (I hope to cover all who have helped me along the way but if I forget anyone, I apologize and hope you know I am grateful for your efforts) From the GLBRC experimental fermentation facility: Wes Marner, Ruwan Ranatunga, Will Bothfeld, Alex LaRue, Haibo Li, and Yaoping Zhang. All of you were instrumental in helping to run many of the chemostat experiments that took up weeks in your lab, and you were all willing to help out when asked. From the GLBRC metabolomics group (Robert Zinkel, Mick McGee, and Alan Higbee), you were all very helpful in answering my questions on analytical techniques and to process the multitudes of samples generated from my chemostat runs. I would also like to thank Camo Cotten and Professor Jennifer Reed for their help in setting up GAMS for running basic metabolic flux analysis and providing direction for targets of future metabolic engineering efforts. I would also like to thank the chemical and biological engineering department staff for all their help from processing forms to fixing computers.

To my undergrads (Zachariah Harris, Michael Luc, Thomas Raines, Josh Rose, Haley Schoenberger, Marty Schumacher), you guys have helped me immensely throughout my graduate career, even if your individual project was unsuccessful in its intended goals. Josh, thank you for putting up with lots of sampling; whether from chemostats or shake flask

production, you are the reason I actually got more than 4 hours of sleep during those later chemostats. Marty, thank you for running all those initial alcohol experiments, without your help there was no way I would have been able to identify alcohol production issues without the extra sets of experiments that you ran. Thomas, while not really my undergrad per se, you have been extremely productive in your semester here and have provided numerous help in getting the data for the alcohol paper and have really gotten that fermentation room up and running.

I would also like to thank every member of the Pfleger lab (Dan Agnew, Matt Begemann, Ryan Clark, Matt Copeland, Spencer Hoover, Travis Korosh, Andrew Markley, Daniel Mendez-Perez, Mark Politz, and Jackie Rand) for putting up with my presence over the years, and putting up with my monotoneness and random stupid sayings. Daniel deserves extra thanks as he had to put up with living in the same house for over 2 years.

I would like to thank my advisor, Brian Pfleger, for all his guidance throughout the years and for putting up with my verbose writing style in manuscripts. I appreciate all that you have done for me and consider myself lucky to have joined your lab.

To Michelle, thank you for putting up with my long hours during those chemostat experiments and for always being there for me when needed. I think I owe you over 473 ice cream cones by now :)

Table of Contents

List of Figures

List of Figures (continued)

List of Tables

Chapter 1: Fuels and petrochemicals: current and alternative routes to their synthesis

1.1. Motivation in the search for alternatives to petroleum products

Petroleum products are used in many facets of every day life from transportation to structural plastics to specialty chemicals used in cosmetics and consumer products. The finite supply of such a resource makes the search for alternative modes of production a worthwhile endeavor. Based on projections from the Energy Information Administration (EIA), oil prices are expected to continually increase, owing to increased worldwide demand, with a decreasing supply domestically. Production from oil shale in the United States only increases capacity by 3 million barrels per day over projected numbers (Figure 1.1). While Figure 1.1 does indicate worldwide consumption outstripping supply, with optimistic predictions it is possible the gap can be closed from a combination of shale gas, production from biofuels, liquids from coal, and natural gas. However, there would likely be additional impacts in the polymers and other products produced from an initial barrel of oil (Figure 1.1). Additionally, oil supply limitations would also push research toward finding alternative methods of production.

Supply and cost issues aside, it is also worthwhile to find ways to reduce greenhouse gas emissions attributed to the use and processing of fossil fuels. Atmospheric CO_2 levels have increased rapidly since the beginning of the industrial revolution and have now reached levels higher than those prior to the last ice age (Petit et al., 1999). While CO_2 levels and global temperatures have been higher in the distant past, such a rapid rise will have dramatic and likely devastating effects on the world's ecosystems. Along with affecting growing seasons and animal populations, increases in temperature could lead to increasing frequencies of violent weather and extreme climate events (Parmesan and Yohe, 2003). One factor among many in attempts to

reduce carbon dioxide levels is the development of renewable energy sources to replace the need for fossil fuels and petroleum products.

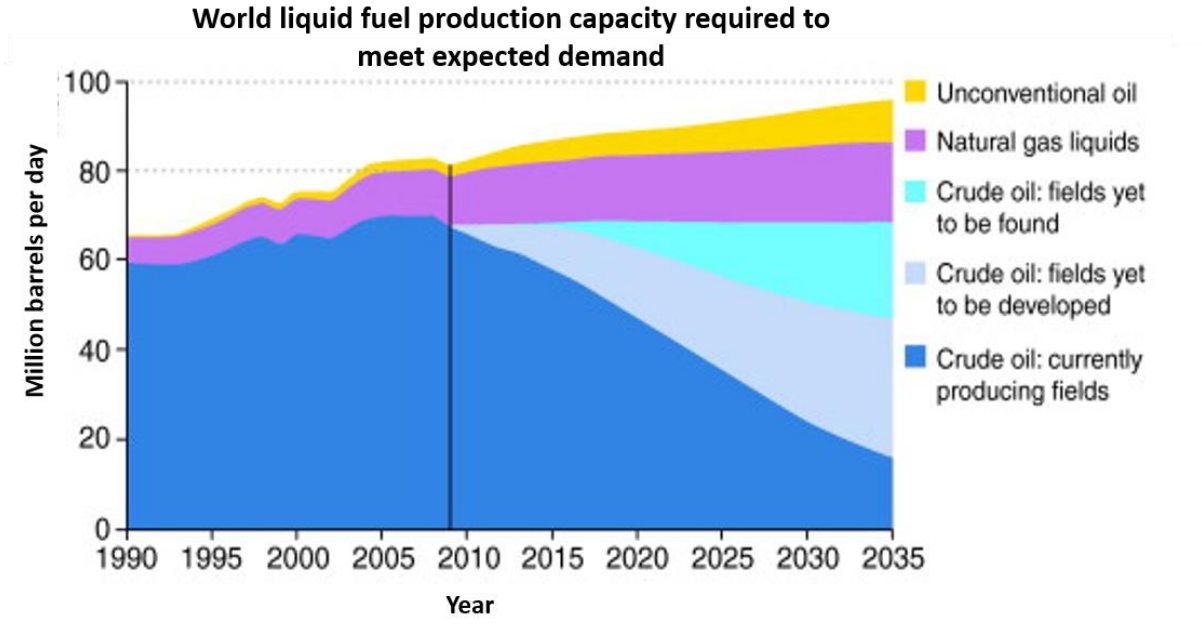


Figure 1.1: Worldwide production capacity from the 2009 EIA annual report. The straight line on the chart indicates production levels as of 2009 (EIA 2009). A gap of approximately 43 million barrells per day is projected to exist by 2035. Developing currently known oil fields could cut that gap in half. Production from oil shale could add as much as 3 million barrels per day to capacity (EIA Annual Energy Outlook 2013). Reproduced from Hughes and Rudolph, 2011.

As nearly 80% of a barrel of oil goes toward transportation fuels, e.g. gasoline, diesel, and jet (EIA), it is necessary to discuss the progress being made towards renewable production of fuels. First generation biofuel efforts focused on production of ethanol from sugarcane (sucrose) and corn starch as well as biodiesel from oils stored in plant seeds (including soybean, rapeseed, and palm). However, the competition between food (including animal feed) and fuel uses of crops spurred rigorous debate and remains an inherent problem for biofuels being deployed as significant replacements of fossil fuels (Banerjee et al., 2010). This issue has motivated the

development of technologies that enable fuel production from non-edible components of lignocellulostic biomass. Lignocellulosic biomass is comprised of three fractions: lignin, hemicellulose, and cellulose. Biomass is either directly converted to molecules of fuel value or broken down to release sugars that can be fermented to fuel compounds by metabolically engineered organisms. In the non fermentation route, strategies such as pyrolysis of biomass to yield a deoxygenated bio-oil, gasification via the partial combustion of biomass to produce syngas (upgraded to fuel levels via Fisher-Tropsch synthesis), liquefaction by catalytic thermal decomposition (eventually producing species that polymerize to a bio-oil), and acid hydrolysis of biomass (Alonso et al., 2010). Additionally, liquid phase reforming using acid hydrolysis pretreated biomass and bimetallic catalysis to generate levulinic acid which can be upgraded to compounds (gamma-valerolactone (GVL), dimethylfuran, or methyltetrahydrofuran) that can be directly blended with gasoline (Alonso et al., 2012). More recent data has also shown that GVL can be used in the generation of a concentration stream of glucose and xylose in a product stream that could be used in fermentation. Alternative routes use the action of commercially available enzymes to break down cellulose and hemicellulose into their sugar monomer or short oligomer components for use in further fermentation (Jin et al., 2010). However, in comparing these enzymatic steps against catalytic steps, they contribute $ 0.60 to $1.60 to the per gallon cost of ethanol, catalysis steps only add ~$0.01/gal (Klein-Marcuschamer et al., 2012). However, sugars produced from these enzymatic or catalysis steps (e.g. glucose, xylose, arabinose) can be used in feedstocks for metabolically engineered organisms to produce desired end product biofuels and chemicals.

An initially appealing alternative to lignocellulostic biofuels is the use of lipid accumulating or secreting algae to produce fuels and chemicals. As algae can accumulate up to

50% of their dry cell weight in neutral lipids or oil, significant research has focused on cultivating algae in large scale and harvesting algal biomass to isolate lipids that can be converted to fuels (Hu et al., 2008). Additionally, the ability to grow at high rates compared to plants, tolerate marginal lands (photobioreactors and open ponds on otherwise unusable land), utilize growth nutrients (nitrogen, phosphorous) from agricultural runoff and wastewater, and sequester carbon dioxide. Despite such advantages, numerous challenges still exist in the production of algal biofuels. Open ponds, while cheap to build and easy to operate, suffer from contamination issues and difficulty in maintaining a constant culture environment. Additionally, as light doesn't penetrate more than a few centimeters into a dense algal culture, scale up is almost entirely on the basis of surface area instead of volume. Development of photobioreactors seeks to relieve the drawbacks of open pond systems, but costs are still far to high to allow for economically viable production (Scott et al., 2010). Still, research continues into more effective route of algal biofuel production.

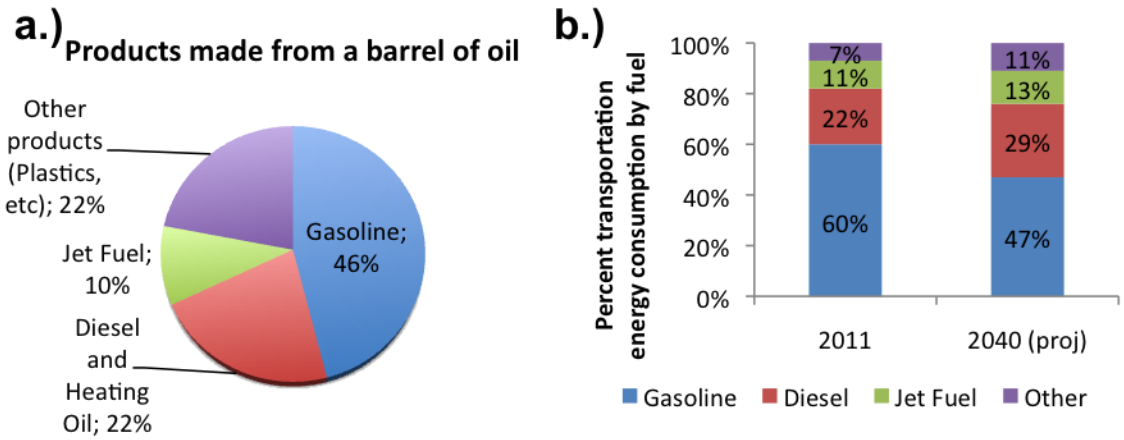


Figure 1.2: a.) The fraction of each type of product made from a barrel of oil and b.) 2011 and projected 2040 transportation energy consumption by fuel type. Source: Department of Energy and U.S. Energy Information Administration Annual Energy Outlook 2013 (www.eia.gov).

Most of the aforementioned lignocellulostic biofuels research has focused on the production of ethanol for use as an alternative or additive to gasoline (Huffer et al., 2012;

Vispute and Huber, 2008). While developments in *n*-butanol and isobutanol production are promising, these short chain alcohols do not have the energy density needed for ideal use in diesel or jet fuels (Peralta-Yahya et al., 2012). As diesel and jet fuels currently comprise a third of domestic transportation energy consumption and are expected to comprise over 40% by 2040 (Figure 1.2), it is imperative that alternative production methods of these fuels are studied. Medium and long chain hydrocarbons have such properties to be used as replacements, and, owing to their similarity to current fuel compounds, are more likely to be compatible with existing engine and storage infrastructure.

A few biochemical routes exist for the production of long carbon chain highly reduced molecules, including: isoprenoid biosynthesis, polyketide biosynthesis, and fatty acid biosynthesis. For the isoprene route either the mevalonate pathway or the deoxyxylulose 5-phosphate (DXP) pathway uses acetyl-CoA or pyruvate and glyceraldehyde-3-phosphate, respectively, as the precursor and produce isopentyl diphosphate (IPP) (Lange et al., 2000). IPP is used as the precursor towards forming a number of different long chain hydrocarbon products (via the condensation of the IPP or derived dimethylallyl disphospahte monomers), such as farnesyl disphosphate (used to generate farnesene) and eventually other compounds (Peralta-Yahya et al., 2012). Alternatively, long hydrocarbon chain compounds can be generated via polyketide synthases. Their often modular design allows for easy manipulation of the enzymatic activites required to generate different levels of reduction at each two carbon interval. However, many polyketide synthases are difficult to functionally express in a heterologous host and significant research still needs to be done in the area of the production of possible biofuel chemicals (Yuzawa et al., 2012). The third pathway is fatty acid biosynthesis where acetyl-CoA most often serves as a precursor for the build up of long chain aliphatic compounds that are used

in the production of cellular membrane components. Significant research in producing biofuel and value added chemical products from such a pathway has been done, and a summary can be found in Lennen at al. (Lennen and Pfleger, 2013). This research thesis deals with the compounds that can be derived through fatty acid biosynthesis (specifically free fatty acids (FFAs) themselves and fatty alcohols) and will be discussed in detail throughout the thesis.

1.2. Rationale for research approach

1.2.1. Synthetic Biology

The most common definition of synthetic biology is that it is the design and construction of biological devices and systems for useful purposes. These systems range from controlling expression of one or two genes via engineering a promoter to be controlled by a transcriptional regulator (Zhang et al., 2012), designing RNA parts that respond to signals (environmental or chemical) and regulate entire metabolic pathways (Win et al., 2009), or to genetically engineer a microbe to produce a chemical it has never encountered (Keasling, 2012). To achieve these outcomes, DNA sequences are designed to encode the enzymatic or regulatory components desired, and are then inserted into an organism of interest using self-replicating vectors (e.g. plasmids or chromosome). If a design is successful, the inserted DNA will confer a new ability too alter an existing ability in the organism. These genetic modification strategies allow one to insert a gene or metabolic pathway from another, less characterized organism, into a model organism such as *Escherichia coli* or *S*accharomyces *cerevisiae*, or an organism that naturally possesses a desirable trait (e.g. tolerant to a target product). For example, introduction of a pathway to produce artemisinic acid (a precursor to the anti-malarial drug artemisinin) was introduced into *S. cerevisiae*, significantly decreasing the cost of production and extraction of the

compound from its original plant host (Paddon et al., 2013; Ro et al., 2006). Such strategies have also been used in the production of entire classes of antibiotics, commodity chemicals, and possible biofuels (Keasling, 2012). As such, using these tools should provide the ability to continually improve titers of FFA and FFA derived products in *E. coli*.

1.2.2. Using E. coli as a host organism

When choosing an organism to work with, one needs to consider a few factors in the selection. The first factor is the tolerance of the organism to high levels of the desired product. Second, one must determine how well characterized an organism is to know if it is amenable to genetic manipulation, whether essential pathways, features, and regulation are present, and whether regulatory tools are available. Third, one needs to determine the ease of working with such an organism as a host, as organisms that require very specific conditions or have extremely slow growth rates would not initially seem good target organisms in producing biofuels and commodity chemicals. Looking at the first factor, one would assume to target organisms that already produce the target compound in their native metabolism, such as some oleaginous yeast species (e.g. *Rhodosporidium* sp.) that can accumulate intracellular lipids up to 70% of their dry biomass weight (Li et al., 2007). However, other factors such as native regulation or uncharacterized metabolic pathways make such strategies not necessarily ideal. In our own lab, significant work is being performed to establish a synthetic biology toolbox, including inducible promoters, surface display, and counter-selection, for an industrially attractive but poorly characterized cyanobacterial strain. Common model host organisms such as *E. coli*, *S. cerevisiae*, and *Bacillus subtilis* have been well studied and have numerous genetic tool sets. Using one would allow researcher easy access to a library of previous research on cellular metabolism, gene regulation, and protocols for genetic manipulation. Here, *E. coli* was selected as the organism of

interest, for its ability to consume the dominant components of lignocellulosic biomass (glucose, xylose, arabinose) and its rapid growth rate. Additionally, *E. coli* appears to tolerate growth in the presence of alkanes, FAEE, and fatty alcohols (Schirmer et al., 2010; Steen et al., 2010). Therefore the majority of this thesis will focus on the use of *E. coli* in the production of FFA and FFA derived compounds.

1.2.3. Kinetic modeling of FFA production and process optimization

Deterministic models of production optimization have been used to increase the yield and productivity of many compounds (antibiotics, lactic acid, ethanol, etc) under controlled conditions (Todorov et al., 2004; Zhu et al., 2010). Numerous parameters such as oxygen level, nutrient limitation, culture pH, and temperature impact product titer, yield and productivity. A continuous culture (chemostat) system allows control over an organism's growth rate and allows one to generate data to determine kinetic parameters in developing a model. Many models exist for non-genetically engineered *E. coli* strains (Orencio-Trejo et al., 2010; Valgepea et al., 2011). Often times in the case of secondary metabolites, yield frequently increases under conditions of slow growth (such as transition to stationary phase or slow feed rates in a fed batch culture) (Gramajo et al., 1993). Additional factors such as nutrient limitation have been used to increase production of various secondary metabolites, such as polyhydroxybutyrate under nitrogen limiting conditions (Johnson et al., 2010). The models generated from chemostat experiments can be used to design cultivation strategies for optimal production under different fermentation conditions.

1.3. Overview of Thesis

The remainder of the thesis explores the multitude of factors required to produce FFA derived chemicals at high levels from unrelated carbon sources. Using synthetic biology tools and process optimization will be discussed heavily. Chapter 2 comprises a review of fatty acid biosynthesis, metabolic engineering efforts to increase titer and yield of FFA derived products, process optimization methods, and nutrient limitation conditional effects. Chapter 3 covers the construction of a plasmid-free FFA overproducing strain and a model for FFA production based on cell growth. Chapter 4 explores production under alternative nutrient limitation and expands the FFA production model. Chapter 5 works to use the data gained from Chapters 3 and 4 to increase the production of fatty alcohols. Finally, Chapter 6 covers alternative strategies to increase the production of FFA biosynthesis and future directions for research.

1.3.1. Chapter 2

The objective of Chapter 2 is to serve as a reference to the reader to gain adequate background knowledge to fully understand the topics in the subsequent chapters. Chapter 2 serves to update the reader on fatty acid biosynthesis and its regulation. Additionally, there is significant discussion and literature review on the field of engineered FFA biosynthesis in *E. coli* as well as FFA derived products, with an emphasis on fatty alcohols. Background involving bioreactor process optimization will also be covered in detail, starting with the most basic deterministic models and discussing the effects of different limiting nutrients on cell metabolism.

1.3.2. Chapter 3

Chapter 3 describes original research on the development of a FFA production model based on cell growth. Previous work in the Pfleger lab produced FFA at a high level via the expression of an acyl-ACP thioesterase in *E. coli* on self-replicating plasmids that require

selective pressure for stable inheritance. In this chapter, we describe how the thioesterase expression cassette was integrated into three loci on the chromosome, thereby generating an antibiotic resistance free strain that produced FFA at comparable titers to plasmid containing strains. The plasmid-free strain was then used in continuous culture to determine parameters for a kinetic model of *E. coli* growth and FFA production. The collected data indicated that low growth rates resulted in the highest yield of FFA and intermediate growth (dilution) rates resulted in the highest FFA productivity. Additional data pointed strongly implicated moving experiments toward non-carbon limitation as a method to increase both FFA productivity and yield.

1.3.3. Chapter 4

The results from Chapter 3 lead directly to the work in Chapter 4. This chapter explores the use of alternative nutrient limitation on FFA production. Once the best limiting nutrient condition (phosphate) was identified, a set of chemostat experiments were run to update the model of FFA production under these conditions. The overall trend of yield and productivity followed that seen under carbon limitation. However, the yields and productivities observed were 2-3 fold higher under phosphate limitation. The updated FFA production model was then used to predict FFA production in a phosphate limited batch reactor, with the prediction being within 10% of the final observed FFA titer.

1.3.4. Chapter 5

Chapter 5 explores the possible steps in ensuring full conversion from FFA to fatty alcohol product. Using a three step enzyme system (thioesterase, acyl-CoA synthetase, acyl-CoA reductase), enzymes with specific activities were taken from a number of different organisms and the ones with the highest conversion ability were selected for use in production. After initial

selection expression levels at each genetic level were varied until minimal residual FFA remained and fatty alcohol titer and yield reached their highest levels. Then, knowledge gained from the FFA production model in Chapters 3 and 4 was used to cultivate the best strain in a bioreactor supplied with minimal, phosphate limitated media to ensure maximum fatty alcohol yield. The resulting fatty alcohol yield of 0.134 g fatty alcohol per gram glucose consumed is more than twice the highest recorded value to date in literature.

1.3.5. Chapter 6

Chapter 6 serves as an aggregate of experiments that motivate and support approaches to further improve FFA titers and yields. Based on the results in Chapter 5, data suggests that FFA production is the limiting step in fatty alcohol production (as yields are higher than for FFA alone). During the experiments described in Chapters 3 and 4, RNA samples were taken at each of the steady state conditions and analyzed using microarrays. Chapter 6 explores these large datasets, discusses preliminary follow-up experiments, and shows additional preliminary results motivated by other groups. Additionally, the effect of increasing expression of genes that regulate fatty acid biosynthesis and degradation will be explored. Lastly, the chapter discusses the introduction of a heterologous pathway to bypass the native FFA biosynthesis route.

1.3.6. Epilogue

The Epilogue serves to reinforce the major themes regarding FFA and fatty alcohol production in *E. coli*. It covers significant accomplishments from each section and their importance to the field. The section finishes by providing future research directions based on data throughout the thesis.

1.4. References

Alonso DM, Bond JQ, Dumesic J a. 2010. Catalytic conversion of biomass to biofuels. *Green Chemistry* **12**:1493.

Alonso DM, Wettstein SG, Dumesic J a. 2012. Bimetallic catalysts for upgrading of biomass to fuels and chemicals. *Chemical Society reviews* **41**:8075–98.

Banerjee S, Mulidar S, Sen R, Giri B, Satpute D, Chakrabarti T, Pandey RA. 2010. Commercializing lignocellulosic bioethanol : technology bottlenecks and possible remidies. *Biofuels, Bioproducts, and Biorefining* **4**:77–93.

Gramajo HC, Takano E, Bibb MJ. 1993. Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. *Molecular Microbiology* **7**:837–45.

Hu Q, Sommerfeld M, Jarvis E, Ghirardi M, Posewitz M, Seibert M, Darzins A. 2008. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal* **54**:621–39.

Hughes L, Rudolph J. 2011 Future world oil production: growth, plateau, or peak? *Current Opinion in Environmental Sustainability* **3**(4):225-234.

Huffer S, Roche CM, Blanch HW, Clark DS. 2012. *Escherichia coli* for biofuel production: bridging the gap from promise to practice. *Trends in Biotechnology* **30**:538–45.

Jin M, Lau MW, Balan V, Dale BE. 2010. Two-step SSCF to convert AFEX-treated switchgrass to ethanol using commercial enzymes and *Saccharomyces cerevisiae* 424A(LNH-ST). *Bioresource Technology* **101**:8171–8.

Johnson K, Kleerebezem R, Van Loosdrecht MCM. 2010. Influence of the C/N ratio on the performance of polyhydroxybutyrate (PHB) producing sequencing batch reactors at short SRTs. *Water Research* **44**:2141–52.

Keasling JD. 2012. Synthetic biology and the development of tools for metabolic engineering. *Metabolic Engineering* **14**:189–95.

Klein-Marcuschamer D, Oleskowicz-Popiel P, Simmons B a, Blanch HW. 2012. The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnology and Bioengineering* **109**:1083–7.

Lange BM, Rujan T, Martin W, Croteau R. 2000. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proceedings of the National Academy of Sciences of the United States of America* **97**:13172–7.

Lennen RM, Pfleger BF. 2013. Microbial production of fatty acid-derived fuels and chemicals. *Current Opinion in Biotechnology* **21**:1–10.

Li Y, Zhao Z (Kent), Bai F. 2007. High-density cultivation of oleaginous yeast Rhodosporidium toruloides Y4 in fed-batch culture. *Enzyme and Microbial Technology* **41**:312–317.

Orencio-Trejo M, Utrilla J, Fernández-Sandoval MT, Huerta-Beristain G, Gosset G, Martinez a. 2010. Engineering the *Escherichia coli* Fermentative Metabolism. *Advances in Biochemical Engineering/Biotechnology* **121**:71–107.

Paddon CJ, Westfall PJ, Pitera DJ, Benjamin K, Fisher K, McPhee D, Leavell MD, Tai a, Main a, Eng D, Polichuk DR, Teoh KH, Reed DW, Treynor T, Lenihan J, Fleck M, Bajad S, Dang G, Dengrove D, Diola D, Dorin G, Ellens KW, Fickes S, Galazzo J, Gaucher SP, Geistlinger T, Henry R, Hepp M, Horning T, Iqbal T, Jiang H, Kizer L, Lieu B, Melis D, Moss N, Regentin R, Secrest S, Tsuruta H, Vazquez R, Westblade LF, Xu L, Yu M, Zhang Y, Zhao L, Lievense J, Covello PS, Keasling JD, Reiling KK, Renninger NS, Newman JD. 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**:528–32.

Parmesan C, Yohe G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature* **421**:37–42.

Peralta-Yahya PP, Zhang F, Del Cardayre SB, Keasling JD. 2012. Microbial engineering for the production of advanced biofuels. *Nature* **488**:320–8.

Petit JR, Raynaud D, Basile I, Chappellaz J, Davisk M, Ritz C, Delmotte M, Legrand M, Lorius C, Pe L, Saltzmank E. 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**:429-436.

Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho K a, Eachus R a, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, Keasling JD. 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**:940–3.

Schirmer A, Rude M a, Li X, Popova E, Del Cardayre SB. 2010. Microbial biosynthesis of alkanes. *Science* **329**:559–562.

Scott S a, Davey MP, Dennis JS, Horst I, Howe CJ, Lea-Smith DJ, Smith AG. 2010. Biodiesel from algae: challenges and prospects. *Current opinion in biotechnology* **21**:277–86.

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

Todorov SD, Van Reenen CA, Dicks LMT. 2004. Optimization of bacteriocin production by *Lactobacillus plantarum* ST13BR, a strain isolated from barley beer. *The Journal of General and Applied Microbiology* **50**:149–57.

Valgepea K, Adamberg K, Vilu R. 2011. Decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting. *BMC Systems Biology* **5**:106.

Vispute TP, Huber GW. 2008. Breaking the chemical and engineering barriers to lignocellulosic biofuels. *International Sugar Journal* **110**:138–149.

Win MN, Liang JC, Smolke CD. 2009. Frameworks for programming biological function through RNA parts and devices. *Chemistry & Biology* **16**:298–310.

Yuzawa S, Kim W, Katz L, Keasling JD. 2012. Heterologous production of polyketides by modular type I polyketide synthases in *Escherichia coli*. *Current Opinion in Biotechnology* **23**:727–35.

Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD. 2012. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic Engineering* **14**:653–60.

Zhu X, Zhang W, Chen X, Wu H, Duan Y, Xu Z. 2010. Generation of high rapamycin producing strain via rational metabolic pathway-based mutagenesis and further titer improvement with fed-batch bioprocess optimization. *Biotechnology and Bioengineering* **107**:506–515.

Chapter 2: Review of the current literature

2.1. Introduction

The potential of producing renewable fuels and chemicals from simple sugars has motivated metabolic engineering research to make organisms produce chemicals that can act as drop in fuel replacements. As discussed in Chapter 1, short chain alcohols have inherent limitations and are not viable replacements for heavy transportation fuels. This fuel sector is best served by medium chain alkanes that are common to petroleum and refinery streams. In biology, fatty acids are the closest class of abundant compounds to these fuels. Thus, this chapter will review the current literature with a focus on the development of fuels and chemicals that can be produced via fatty acid metabolism.

While the production of fuels drives most of the research funding in this area, commercial companies such as Solazyme and Codexis have focused on developing microbial routes to specialty oils and oleochemicals. The economics of producing high-value, small(er) volume oleochemicals (e.g. comparing the wholesale price of fatty alcohols (~$1/lb) versus diesel (~$0.40/lb), ICIS and EIA 2012 pricing) from renewable sugars is significantly more favorable than for the case of producing biodiesels at today's diesel prices and volumes. Many biofuels companies (e.g. LS9) are simultaneously pursuing processes for producing chemicals and fuels with the hope that lessons learned from the chemical process development can lead to low-cost biodiesel production.

Fatty acid biosynthesis is not the only route to produce fuel-grade hydrocarbons. Polyketide synthesis, a pathway with many similarities to fatty acid biosynthesis, and isoprenoid biosynthesis has also garnered significant academic and industrial attention. However, one advantage of fatty acid metabolism, over these other pathways, is that nearly all forms of life

make fatty acids in large quantities for lipid synthesis. Therefore any organism, whether photosynthetic, solvent tolerant, cellulolytic, thermotolerant, halotolerant, or possessing another distinct physiological advantage, could be developed to produce fatty acid derived chemicals. Once a strain was engineered to produce fatty acids at high yields, it could be engineered to express one of a wide range of enzymatic pathways for converting fatty acids to fuel compounds, chemical building blocks, and materials (Figure 2.1). Section 2.3 will review this class of chemical targets as well as the enzymes required to produce them from intermediates in fatty acid metabolism. One of the major challenges to producing fuels and oleochemicals from sugars via fatty acid biosynthesis is the need to achieve theoretical yields in order to be economically competitive with petroleum-derived alternatives. As the yields for most free fatty acid (FFA) derived chemicals (fatty acid ethyl esters (FAEE), alkanes, long chain alcohols) reported in the literature are 40% of the theoretical maximum value or lower (Schirmer et al., 2010; Zhang et al., 2011; Chapter 5), significant research must continue to understand the physiological, metabolic,

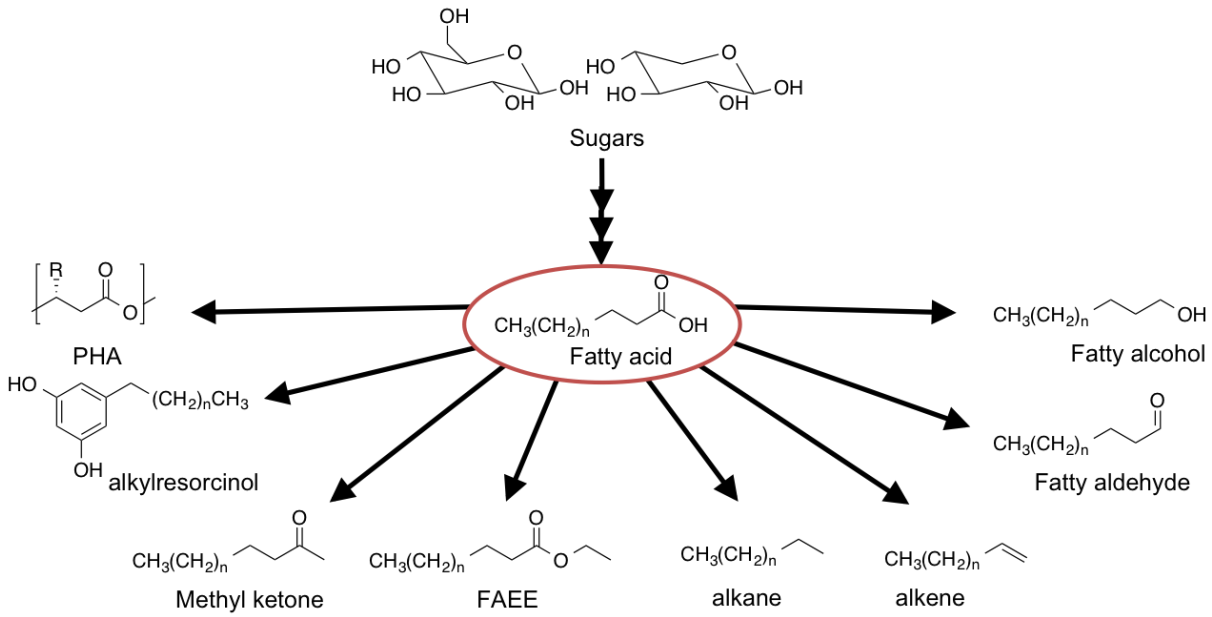


Figure 2.1: The variety of compounds that can be produced in vivo from endogenous free fatty acid production.

and regulatory barriers that limit current yields.

To identify these barriers one needs to both understand the native regulatory pathways controlling synthesis, previous efforts to increase fatty acid production, and external factors that may limit production via metabolic or physiological stresses. For these reasons, Chapter 2 serves as a source of background knowledge on all three areas focusing on native fatty acid biosynthesis and regulation as well as an alternative synthesis route used in *Corynebacterineae*, yeast, and mammals. Past and current efforts to increase FFA and FFA derived chemical production will also be discussed. Finally, the remainder of this chapter will discuss the external factors involved in process optimization and the impacts they may have on cellular metabolism and regulation that may shed light on production impacts from using such strategies.

2.2. Fatty acid biosynthesis

2.2.1. Standard biosynthesis in E. coli

Fatty acid biosynthesis is a major metabolic pathway in the synthesis of long chain hydrocarbon compounds in all organisms. The enzymes involved in *E. coli* have been extensively reviewed (Magnuson et al., 1993; Rock and Cronan, 1996), and can be seen in Figure 2.2. First, a carbon source is converted to acetyl-coenzyme-A (acetyl-CoA) through the appropriate central catabolic pathways. The most common example is the conversion of glucose to pyruvate via glycolysis with a subsequent conversion to acetyl-CoA catalyzed by pyruvate dehydrogenase. Acetyl-CoA is then converted to malonyl-CoA by acetyl-CoA carboxylase (ACC), a four protein complex encoded by *accABCD* in *E. coli*. In the majority of microbes, malonyl-CoA is then transferred to acyl-carrier protein (ACP) by an acyl transferase (AT) encoded by *fabD* in *E. coli*. Malonyl-ACP thioesters are the elongation unit used in the

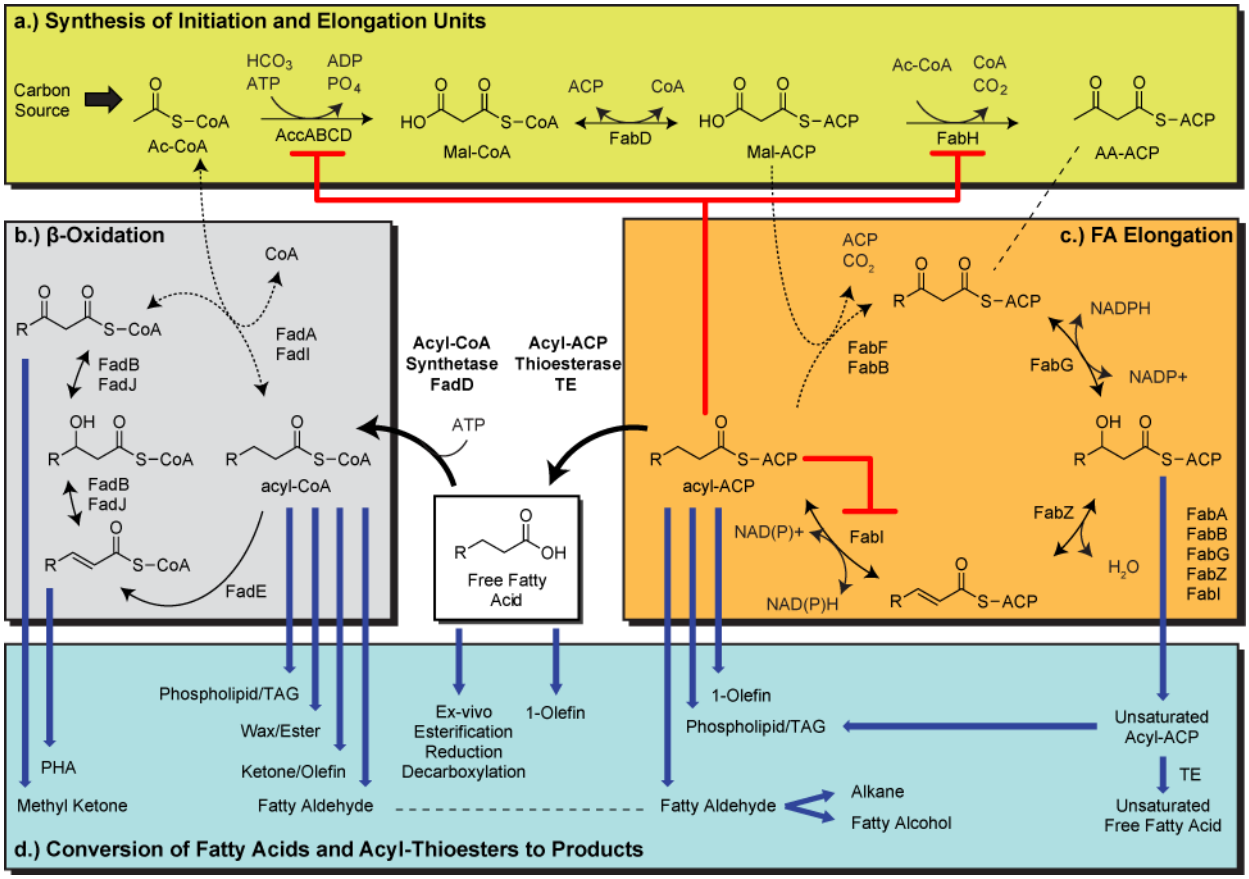


Figure 2.2: General mechanism of fatty acid biosynthesis. a.) Synthesis and elongation steps b.) β-oxidation, c.) fatty acid elongation are all shown, and d.) final products are shown. Red lines indicate the feedback inhibition from long chain acyl-ACPs and blue lines indicate possible final products. Figure modified from (Lennen, 2012).

elongation and reductive cycle of fatty acid biosynthesis. A malonyl-ACP is incorporated into a growing fatty acyl-ACP by a Clasien condensation catalyzed by a keto-acyl synthase (KS) or fatty acid synthase (FAS). *E. coli* contains three FAS isozymes encoded by *fabF* (primary), *fabH* (initiation), and *fabB* (unsaturated). The FAS reaction produces carbon dioxide and a keto-acyl group two carbons longer than the prior acyl-ACP (Magnuson et al., 1993). The keto-acyl group is reduced to an acyl group by the reductive actions of a keto-reductase (KR, encoded by *fabG*), a dehydratase (DH, encoded by *fabZ*), and an enoyl reductase (ER, encoded by *fabI*), allowing for further extension. The elongation/reduction cycle begins with the synthesis of acetoacetyl-ACP from malonyl-ACP and acetyl-CoA (Harder et al., 1974) via a reaction catalyzed by a KS

(Fab H, KS III). In an alternative pathway, malonyl-ACP can be decarboxylated by a different KS (FabB, KS I), forming acetyl-ACP which can then be condensed with malonyl-ACP forming the aceotacetyl-ACP group (Magnuson et al., 1993) (Figure 2.2a and Figure 2.3). Essential unsaturated fatty acids are made by generating a cis-double bond via dehydration of the 3-hydroxydecanoyl-ACP by a DH (FabA) in place of FabZ. The cis double bond cannot be reduced by ER (FabZ) and can only be elongated by a distinct KS (FabB). Termination of the elongation/reduction cycle occurs when the acyl group is transferred to a phospholipid, removed from ACP by a thioesterase, or modified by another metabolic enzyme (e.g. acyl-ACP reductases that produce fatty aldehydes).

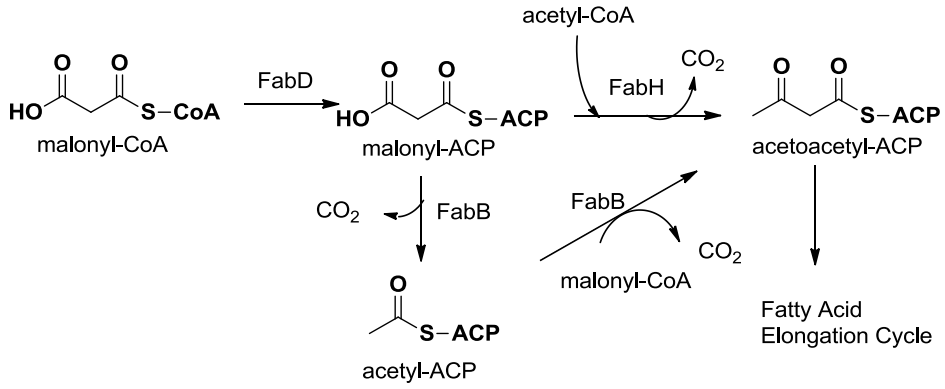


Figure 2.3: Type II fatty acid biosynthesis initiation pathway within *E. coli*. The fatty acid elongation cycle and origin of malonyl-CoA are seen in Figure 2.2.

Two types of fatty acid synthesis mechanisms exist in nature. Type I fatty acid synthesis systems contain each of the enzymatic domains (AT, KS, KR, DH, ER, ACP) on a single protein (Schweizer and Hofmann, 2004). In contrast, the enzymes of type II systems remain separate, discrete proteins in the cellular cytoplasm. Type II systems are common to *E. coli* and present in all bacteria, while type I systems are common to higher eukaryotes. However, type I systems have been found in some species of the order Actinomycetales (Fernandes and Kolattukudy,

1996; Radmacher et al., 2005; Stuible et al., 1997). Section 2.2.2 provides additional details of Type I FAS systems.

Regulation of fatty acid biosynthesis has been well characterized in *E. coli* (Rock and Cronan, 1996). The entire fatty acid biosynthesis pathway is feedback inhibited by the long-chain (C_{16}-C_{18}) acyl-ACP intermediates that are the substrates for lipid synthesis (Figure 2.2c) (Magnuson et al., 1993). Inhibition can be alleviated by high level expression of acyl-ACP thioesterases (Cho and Cronan, 1993), which deplete the concentrations of the long-chain acyl-ACPs. Another basic regulatory mechanism occurs during bacterial growth when the majority of acyl-ACPs are transferred into the phospholipid biosynthesis pathway (Yoshimura et al., 2007). The type II pathway is further optimized for phospholipid biosynthesis as FabI is membrane associated, putting it in close proximity to phospholipid synthesis enzymes. Additionally, the separate domain architecture of type II systems allows for acyl-ACP intermediates to be acted upon by enzymes involved with other cellular functions, such as in pathways for lipid A, homoserine lactone, and SpoT synthesis (Byers and Gong, 2007). Also, external regulatory mechanisms exist by the expression of the transcriptional regulators FadR and FabR. FadR is primarily involved in fatty acid breakdown (a.k.a. β-oxidation) and is activated by long chain acyl-CoAs (Rock and Cronan, 1996). Additionally, along with FabR, FadR regulates unsaturated fatty acid biosynthesis in *E. coli* (Zhang et al., 2002). When long chain acyl-ACPs are not present, FadR activates FabA and FabB, facilitating increased unsaturated fatty acid production. Conversely, FabR represses FabA and FabB, but the control of its own expression is not known (Fujita et al., 2007). Each of these regulatory mechanisms are integral to the correct functioning of the native fatty acid synthesis pathway of *E. coli*.

While transport of long chain fatty acids across the outer cell membrane into the cell has long been known to be facilitated by the transport protein FadL (Nunn, 1986), others protein products have been associated with fatty acid uptake (TolC and Prc) as knockouts of their coding genes led to defective growth on a fatty acid (oleate) as a sole carbon source (Azizan and Black, 1994). However, only recently have genes associated with fatty acid export been known. Work in our lab (Lennen et al., 2013) found that disruption of putative transport proteins exacerbated a dramatic loss of fitness as free fatty acid producing cells entered stationary phase. The outer membrane protein TolC as well as inner membrane proteins AcrAB, MdtEF, and EmrAB were all necessary to allow for full FFA titers and minimization of fitness loss.

Fatty acids are taken up and broken down in a cellular process known as β-oxidation (Figure 2.2b). First, extracellular fatty acids are transported across the outer membrane via FadL. Next, the fatty acids are activated to acyl-CoAs via the action of FadD, likely in the inner membrane (DiRusso et al., 1999). These acyl-CoAs are then catabolized in an iterative pathway composing four enzymatic reactions. An acyl-CoA dehydrogenase (FadE) converts the acyl-CoA to an enoyl-CoA, while a gene product with both enoyl-CoA hydratase and (3S)-hydroxylacyl-CoA activity (FadB) facilitates conversion to a ketoacyl-CoA. Finally, a ketoacyl-thiolase (FadA) generates an acetyl-CoA and an acyl-CoA two carbons shorter than the initial chain length, which is then used as a substrate for the next round of the cycle. Once the acyl-CoA chain length is short enough, *ato* gene products are also able to process short chain acyl-CoA and fatty acid products (Jenkins and Nunn, 1987). In addition to the genes mentioned in aerobic, saturated fatty acid breakdown, a 2,4-dienoyl-CoA reductase (FadH) acts as an analog for FadE for unsaturated fatty acid substrates. Also, the gene products of *fadK*, *fadI*, and *fadJ* have strong sequence homology to FadD, FadA, and FadB, respectively. These genes have also been shown

to be critical for fatty acid breakdown under anaerobic conditions in *E. coli* (Campbell et al., 2003). The primary regulator of these genes is the transcription factor FadR. Under conditions with no acyl-CoAs present, FadR represses the expression of genes involved in β-oxidation. FadR specific binding sites have been found in the operator regions of fadD, *fadE*, *fadBA*, *fadH*, *fadIJ*, and *fadL* (Feng and Cronan, 2012; Fujita et al., 2007). When acyl-CoAs are present they bind to FadR, triggering a conformational change in FadR and releasing it from its cognate DNA sequence (Cronan, 1997). With the loss of repression, the β-oxidation genes are expressed and acyl-CoAs are broken down.

2.2.2. Alternative mechanisms of fatty acid biosynthesis

An alternative to the fatty acid synthesis mechanism described above is the bacterial type I fatty acid synthase found in many orders of Actinomycetales. Type I systems have been heterologously expressed in *E. coli* (Stuible et al., 1997) but are not common to most bacteria. In contrast to type II synthesis, type I systems contain all the functional FAS domains on a single polypeptide chain, forming one single enzyme complex. The reactions take place in a similar fashion except with the acyl-ACP chain staying within the enzyme complex during extension. The type I synthase also contains a malonyl/palmitoyl transferase (MPT) domain which terminates elongation by transfering the acyl chain to a CoA releasing the acyl-CoA product (Schweizer and Hofmann, 2004). That is in stark contrast to the native fatty acid biosynthesis in *E. coli* which required further activation to convert the acyl-ACP product or FFA to an acyl-CoA. In addition to the differing domains, there is also a slight functional difference in the acyl-chain synthesis. As type I FASs make use of only one ACP per elongation reaction, an additional acyl group carrier would be required to allow for elongation. Such a task requires the transfer of the acyl-chain to the active cysteine of the KS domain. This frees up the 4’-phosphopantetheine

(PPT) moiety of ACP to accept a malonyl-CoA through mediation of the MPT domain. The malonyl-ACP and acyl-KS are then transferred to allow extension of the acyl-ACP in a similar manner to the general fatty acid biosynthesis pathway. Such differences in the fatty acid biosynthesis pathway indicate the possibility for avoiding some of the regulatory pathways in the type II system. However, owing to issues with expression of large heterologous proteins in *E. coli* (Chang et al., 2007), there may be significant obstacles to overcome in using a type I system in *E. coli*.

2.3. Genetic modifications to increase FFA derived product levels in *Escherichia coli*

2.3.1. FFA biosynthesis

Long chain acyl-groups, the precursors to hydrocarbons are produced natively in *E. coli* as ACP thioesters that are the substrates for synthesis of phospholipids. Acyl-ACPs cannot be transported across the cell membrane and would not be useful as fuels. Therefore to produce free compounds that could be purified from cell culture, the thioester bond between an acyl chain moiety and ACP must be broken. This reaction is catalyzed by a thioesterase and leads to the production of a free fatty acids (FFA). A FFA is defined as the free species created by thioesterase cleavage of an acyl-ACP. Compounds derived from FFAs are numerous and include fuels (alkanes, FAEEs), commodity chemicals (polymers, fatty alcohols), and high value products (resorcinols). FFAs have been targeted as model compounds because of the small number of modifications needed to produce them in genetically amenable hosts such as *E. coli*. In addition, FFAs are actively transported out of *E. coli* and can be catalytically converted to many products ex-vivo. While not useful as fuels, FFAs are used commercially in various consumer products and are therefore a desirable target product. Our goal is optimize the

production of fatty acid-derived compounds by first maximizing flux to FFA and then optimizing downstream metabolic pathways for conversion to higher value products (e.g. fatty alcohols – see Chapter 5).

Initial work in FFA overproduction in *E. coli* focused on characterization of thioesterases and determining the regulation of fatty acid biosynthesis. The first metabolic engineering strategy to improve FFA titer was employed by Lu et al. (Lu et al., 2008). Modifications to β-oxidation (Δ*fadD*) and fatty acid biosynthesis (overexpressing acetyl-coA carboxylase) were combined with plasmid based expression of a thioesterase (either the native 'TesA (TesA without a leader sequence) or a plant acyl-ACP thioesterase from *Cinnamomum camphorum*) to achieve an increase of approximately 20 fold of total FFA levels compared to the original strain without the thioesterase. A controlled fed batch fermentation produced a titer of 2.5 g/L FFA and yield of 0.048 g FFA per g glycerol consumed. While an improvement, the yield seen by Lu et al. was significantly lower than the maximum theoretical yield of between 0.35-0.39 g FFA per g carbon source consumed (values depend on FFA species and carbon source, with production of C_{12} FFA and glucose as a carbon source having a maximum yield of 0.35 g/g) (Lennen and Pfleger, 2012). Further studies have achieved similar yields with different FFA species using different thioesterases (Lennen et al., 2010; Li et al., 2012; Zhang et al., 2011). Other genetic modifications to β-oxidation have included using a Δ*fadE* rather than a Δ*fadD* strain of *E. coli* DH1, which led to a 70% increase in titer and increase in yield up to 0.06 g FFA / g glucose. However, the same deletion was not as effective in *E. coli* MG1655 (Hoover et al., 2012). Additionally, a Δ*fadL* strain expressing 'TesA was able to reach titers two fold higher than those seen by Lu et al., but it came at the expense of yield, which dropped to 0.044 g FFA per g glucose. However, that difference could be related to the differences in maximum FFA yield off

of glucose versus glycerol (Lennen and Pfleger, 2012). Recently, further improvements have been made to titer, yield, and productivity through genetic modifications of fatty acid metabolism, central catabolic metabolism, and regulatory proteins. FabZ overexpression was identified by flux balance analysis modeling and its expression with the thioesterase from *Ricinus communis* was able to produce a FFA titer of 1.7 g/L with a yield of 0.14 g FFA / g glucose (Ranganathan et al., 2012). The reversal of the β-oxidation pathway (via constitutive expression of *fadR* and *atoC* and mutations to *crp*) was able to produce up to 6.9 g/L FFA with a yield of up to 0.2 g FFA / g carbon source in a modified *E. coli* MG1655 base strain using the native thioesterase FadM (Dellomonaco et al., 2011). While the evidence suggested the ability to run β-oxidation in reverse was due to the activity of YdiO (catalyzing the reverse reaction of FadE), it was not overexpressed (nor *fadE* removed) in the study. Also, the process required the addition of pantothenate, possibly limiting its ability for commercialization (depending on pantothenate consumption). Additionally, overexpression of *fadR* was seen to vastly improve FFA production in *E. coli* DH1 Δ*fadE* (Zhang et al., 2012b). The observed yield reached 67% of theoretical (0.24 g FFA /g carbon source) and the titer reached 5.2 g/L using the thioesterase ‘TesA. Transcriptomics studies done in FadR overexpression strains identified upregulated genes in fatty acid biosynthesis, however, upregulation of those genes (*fabB*, *fabF*) in the absence of FadR did not produce even half of the titer increase seen with FadR (Zhang et al., 2012b). These results indicate a significance of FadR in relation to fatty acid biosynthesis beyond its effects on unsaturated species. Alternative modifications have included varying expression levels of genes involved in glycolysis, acetyl-CoA activation, and fatty acid synthesis on multiple plasmids and optimizing their expression (Xu et al., 2013). Using the *Cocos nucifera* fatty acyl-ACP thioesterase CnfatB2, the highest titer (8.6 g/L) and productivity (0.124 g/L/h) of FFA was

achieved. However, yields (0.08 g FFA/g carbon source) and specific productivity numbers (~0.007 g/gDCW/h) were lower than seen elsewhere (Chapter 4; Zhang et al., 2012).

2.3.2. Compounds for biofuels

While FFAs are compounds that can be made via simple modifications to cellular metabolic pathways and can be converted to versatile products, they do not have ideal fuel properties. For example, saturated fatty acids with a chain length longer than nine are solids at room temperature. Thus, focus has turned to producing drop in fuels from FFA, including: fatty acid methyl esters (FAME), fatty acid ethyl esters (FAEE), alkenes, and alkanes. Production of FAEEs in *E. coli* has required a gradual introduction of each pathway required for direct synthesis of the compound in vivo from an unrelated carbon source. Efficient endogenous ethanol production in *E. coli* was achieved expressing the heterologous pyruvate decarboxylase (Pdc) and alcohol dehydrogenase (AdhB) from *Zymomonas mobilis* (Ingram et al., 1987). Expressing the promiscuous was ester synthase/acyl-CoA:diacylglycerol acyltransferase (WS/DGAT; AtfA), a genetically engineered ethanol overproducing *E. coli* was able to produce up to 1.28 g/L FAEE when oleate was exogenously supplied to the medium (Kalscheuer et al., 2006). The work by Steen et al. expanded on this by overexpressing 'TesA and AtfA in an ethanol producing *E. coli* DH1Δ*fadE* strain and acheivin a titer of 0.4 g/L that increased to 0.67 g/L with the addition of a dodecane overlayer (Steen et al., 2010). Further improvement was achieved using a sensor regulator system based on the transcriptional regulator FadR. By introducing FadR regulated promoters onto plasmids containing *pdc*/*adhB* and *atfA*/*fadD*, respectively, a titer of 1.5 g/L and yield 28% of maximum theoretical was achieved (Zhang et al., 2012a).

In addition to FAEE, metabolic engineering interest has started to look toward alkane and alkene production as a viable drop in biofuel alternative. Work by Schirmer et al. (Schirmer et al., 2010) was able to demonstrate the production of alkanes in *E. coli* by using an acyl-ACP reductase (AAR) from *Synechococcus elongotas* and aldehyde decarbonylase (ADC) from *Nostoc punctiforme*. Titers reached 75 mg/L of mostly pentadecane and heptadecane. More recent work has replaced the AAR with a carboxylic acid reductase complex from *Photorhabdus lumisescens* and has allowed for more chain length control at the level of the thioesterase (Howard et al., 2013). Additionally, Howard et al. was able to display the ability to synthesize branched chain alkanes in *E. coli* by expressing an α-keto acid dehydrogenase complex and β-keto acyl-acyl carrier protein synthase III from *Bacillus subtilis* in addition to the existing system. However, total alkane titers remained below 10 mg/L (with significant residual aldehydes) indicating substantial room for improvement. In addition to alkanes, olefins have also been of recent interest owing to α-olefins found in cyanobacteria. A genetic study within our laboratory found the gene *ols* to be responsible for olefin production in *Synechococcus sp. PCC 7002* (Mendez-Perez et al., 2011). Ols has significant homology to type I polyketide synthases and experimental evidence points toward loading acyl substrates for one round of elongation and eventual decarboxylation. While overexpression has been successful in increasing olefin titers in the native organism, expression in *E. coli* has thus far failed to produce olefins. The cytochrome P450 enzyme from *Jeotgalicoccus*, OleT, has shown promise in *E. coli*. Expression led to unreported titers of both C_{15} and C_{17} alkenes (Rude et al., 2011). The central challenge of producing alkenes via cytochrome P450s or polyketide synthases has been getting functional expression of these enzymes in *E. coli,* as expression problems have been reported for similar enzymes in the past (Chang et al., 2007; Pfeifer and Khosla, 2001). While all mentioned

pathways show promise, alkane production appears the most promising owing to the expressed protein products being small, soluble proteins (between 200 and 350 amino acids) and only having to optimize the expression of one pathway rather than two simultaneous pathways (as in FAEE production). However, current production levels are far below needed to have an industrially relevant process for fuel production (Figure 2.4), so avenues to produce value added chemicals appear the most viable in the short term.

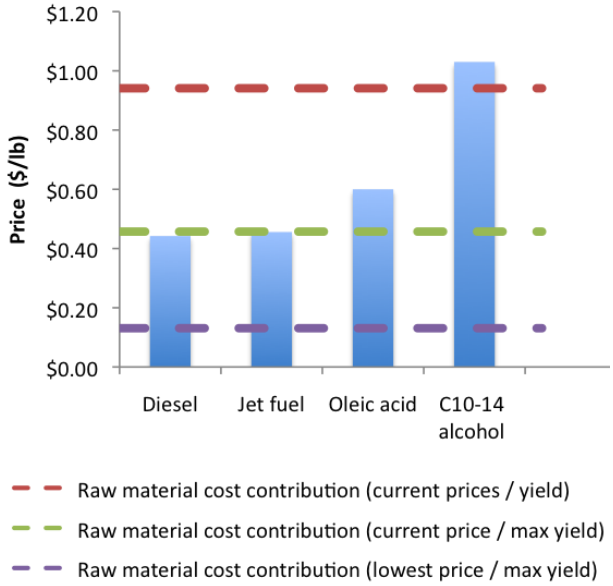


Figure 2.4: Current price (ICIS and EIA, blue bars) of specific FFA derived compounds. Dashed lines represent cost contribution of raw materials to final price at current sugar prices and current maximum yields (red), current sugar prices and maximum theoretical yields (green), and ideal sugar prices and maximum theoretical yields (purple). Ideal sugar price is based on the corn stover price of $40 per dry ton, within the range of $37 per dry ton and $45 per dry ton (United States Department of Agriculture, 2012; Maung, 2013), assuming full recovery of sugars from cellulose and hemicellulose fractions.

2.3.3. Fatty alcohol biosynthesis

While fatty alcohols do not have the ideal properties to exist as the sole component of a fuel, they are still valuable chemicals with a significant market in the production of detergents, emulsifiers, lubricants, and cosmetics (Mudge et al., 2008; Rupilius and Ahmad,

2006). Fatty alcohols have 2.5 times the value of diesel fuel and can be produced at an economically viable level with a much lower productivity and yield than fuel chemicals (Figure 2.4). Most industrial production uses plant oils or petrochemicals as the source. Fatty alcohols have been detected in numerous organisms, including: plants, animals, protists, and specific genera of bacteria (Doan et al., 2009; Hellenbrand et al., 2011; Rowland and Domergue, 2012). Fatty alcohols are synthesized via the action of an acyl-CoA or acyl-ACP reductase. In vitro tests showed a number of these enzymes only have activity for the conversion of the acyl-ACP or acyl-CoA to a fatty aldehyde (Reiser and Somerville, 1997; Teerawanichpan and Qiu, 2010). While observations of the fatty aldehydes *in vivo* would seem to support that finding (Schirmer et al., 2010; Steen et al., 2010), further explanation is needed to explain the fatty alcohol presence. It is thought that spontaneous oxidation or other enzymatic activity is responsible for the conversion to the alcohol. Expression of the hypothesized aldehyde reductase in *E. coli* (YjgB) has been shown to increase fatty alcohol levels in conjunction with a reductase that only catalyzes the conversion steps to an aldehyde product (Akhtar et al., 2013). In addition to reductases that convert from the acyl-ACP or acyl-CoA to the aldehyde, two other classes of these enzymes are known. A carboxylic acid reductase from *Mycobacterium marinum* has been shown to convert fatty acids directly to fatty aldehydes, and acyl-CoA reductases from certain plants and marine bacteria have been shown to catalyze both steps in the conversion acyl-CoAs to fatty alcohols (Akhtar et al., 2013; Hofvander et al., 2011; Rowland and Domergue, 2012; Willis et al., 2011). The carboxylic acid reductases also have the interesting characteristic of having 20% greater activity at higher pHs (7.5-8.0) compared to neutral pH (Akhtar et al., 2013). A summary of the routes to fatty alcohol production can be seen in Figure 2.5.

While some of these reductases have been expressed in *E. coli* (Doan et al., 2009; Reiser and Somerville, 1997; Willis et al., 2011), minimal work has been done in attempting to use them for production of high levels of fatty alcohols (a full summary can be seen in Table 2.1). Fatty alcohol titers of 60 mg/L from single activity bacterial acyl-CoA reductases (Steen et al., 2010), 10 mg/L from plant acyl-CoA reductases (Doan et al., 2009), 100 mg/L from acyl-ACP reductases (Schirmer et al., 2010), and 350 mg/L from carboxylic acid reductases have been observed (Akhtar et al., 2013), all with reported yields less than 0.03 g alcohol / g carbon source. In contrast, titers as higher than 5 g/L and yields as high as 0.26 g FFA / g carbon source have been reported for fatty acids (Zhang et al., 2012b). As there are more engineered conversion steps required in the fatty alcohol biosynthesis pathway compared to fatty acid biosynthesis, it is likely that balancing enzymatic expression of each step is required to achieve higher yields. The work by Steen et al. used overexpression of a thioesterase, acyl-coA ligase, and single activity acyl-CoA reductase in an *E. coli ΔfadE* strain, where it was shown the elimination of *fadE* was crucial to prevent competition for the acyl-CoA intermediate (Steen et al., 2010). Further engineering efforts were able to show that swapping thioesterases led to different chain length alcohol products (Steen et al., 2010). However, as there are reductases with different specificities toward specific acyl-chain lengths, one would need to express different enzymes at each step per chain length product to allow for higher yields. Additionally, it is likely that balancing expression of each step is necessary to avoid the accumulation of intermediates, especially owing to toxicity effects with a buildup of FFAs (Lennen et al., 2011).

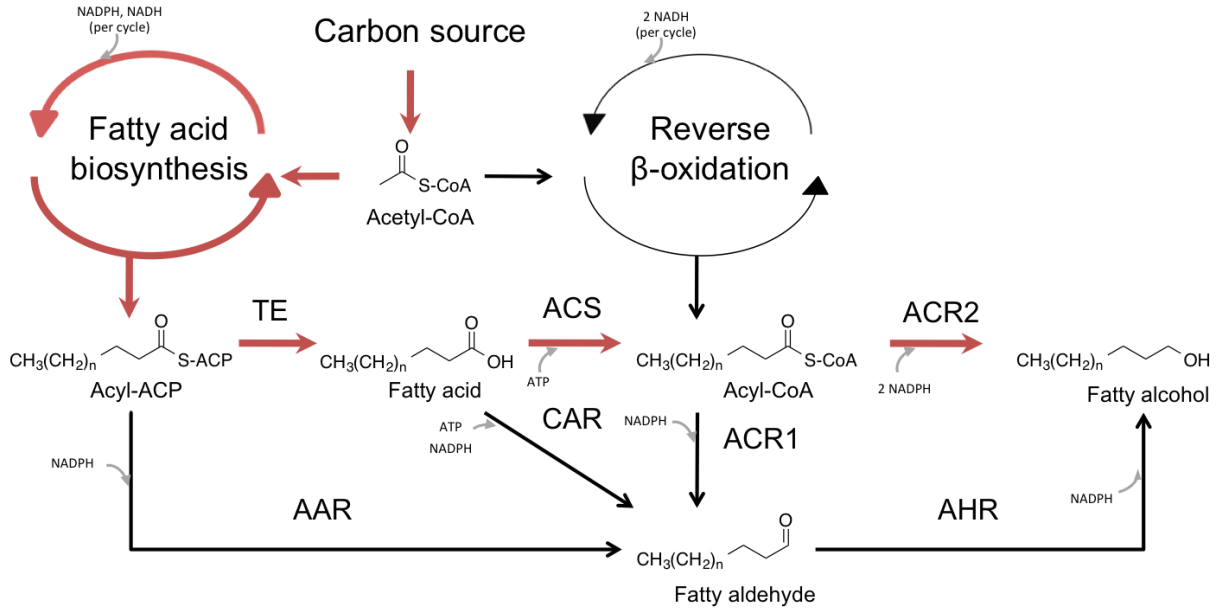


Figure 2.5: Summary of the different routes to bioproduction of fatty alcohols in *E. coli*. The route used in Chapter 5 is in red. Abbreviations are as follows: AAR (acyl-ACP reductase), AHR (aldehyde reductase), ACR1 (single activity acyl-CoA reductase), ACR2 (double activity acyl-CoA reductase), ACS (acyl-CoA synthetase), CAR (carboxylic acid reductase), and TE (thioesterase).

In addition to the fatty acid biosynthesis route, fatty alcohols have also been synthesized through reverse β-oxidation, with titers (primarily of C_8 and C_{10}) reaching 250 mg/L fatty alcohols with a yield of 0.049 g fatty alcohol per gram glucose consumed (other sources of carbon were available, but no consumption data was provided) (Dellomonaco et al., 2011). The reverse β-oxidation route allows for a higher maximum theoretical yield than the fatty acid biosynthesis route. In fatty acid biosynthesis, ATP is required in the synthesis of malonyl-ACP and in the activation of the fatty acid product, while reverse β-oxidation does not use these steps. However, the process requires the mutation of *crp*, *fadR*, and *atoC* and the disruption of *arcA*, all of which are either global regulators (*crp*, *fadR*, *arcA*) or are critical to the β-oxidation pathway (*atoC*). Overall, while fatty alcohol bioproduction shows much promise, it is still in the

developing stages and further research needs to be done in the balancing of gene expression and optimization of conditions.

Table 2.1: Reported titers and yields of fatty alcohols that are microbially synthesized.

Primary alcohol species	Organism	Titer (g/L) / Yield(% w/w)	Carbon source	Medium	Reference
C_{16}	*E. coli*	0.06 / 0.003	Glucose	Minimal defined	(Steen et al., 2010)
C_{16}	*E. coli*	0.14 / 0.005	Glucose	Minimal defined	(Schirmer et al., 2010)
C_{10}	*E. coli*	0.25 / 0.049*	Glucose*	Minimal defined	(Dellomonaco et al., 2011)
C_{12}, C_{14}	*E. coli*	0.45 / 0.01	Glucose	Minimal defined	(Zheng et al., 2012)
C_{16}, C_{18}	*E. coli*	0.10 / 0.002	Glucose	Minimal defined	(Zheng et al., 2012)
C_{12}, C_{14}	*E. coli*	0.36 / 0.043	Glucose	Minimal defined	(Akhtar et al., 2013)
C_{12}, C_{14}	*E. coli*	1.7 / 0.134	Glucose	Minimal defined	Chapter 5

*-Additional carbon source used, yield calculated from primary carbon source

2.3.4. Alternative high value products

In addition to the fatty alcohols mentioned previously, a number of other value added products can be synthesized via the fatty acid biosynthesis pathway. Hydroxylated fatty acids can be polymerized to form medium chain length polyhydroxyalkonotates (PHA). PHA is of commercial interest owing to its biodegradable properties and the possibility of it being produced renewably. Exogenous feeding of vegetable oils to engineered *E. coli* has generated polymerizable 3-hydroxyacyl-CoA monomers (Gao et al., 2011). Other work has demonstrated that disruption of β-oxidation (after FadE) allows the generation of a more uniform chain length monomer product, and chain length can be controlled at the level of exogenous feeding or

endogenous production via expression of a thioesterase (Agnew et al., 2012). In addition to the bio-plastics, a class of compounds known as alklylresorcinols can also be produced. Some specific versions of these compounds have antitumor (climacostol) properties or can be used as a component in resins (cardanol) (Stasiuk and Kozubek, 2010). These compounds are synthesized by the action of a type III polyketide synthase (chalcone family) that catalyses the addition of two or three malonyl-CoA units onto an acyl-CoA, creating a triketide or tetraketide. The triketide is then cyclized to a triketide pyrone and the tetraketide is cyclized to a resorcinol with a long alkyl chain (Funa et al., 2006). The type III polyketide synthesis from *Azotobacter vinelandii* (ArsB) and *B. subtilis* (BpsA) have shown the most promise owing to their broad substrate specificity and soluble expression in *E. coli* (Funa et al., 2006; Nakano et al., 2009). However, the only heterologous *in vivo* production within *E. coli* has used the type III polyketide synthase from *Pseudomonas fluorescens* (PhlD), which produces the simple resourcinol phloroglucinol via three malonyl-CoA precursors, to get resorcinol production as high as 1 g/L (Zha et al., 2009). While these higher value products show much promise, they are still very much in the exploratory research phase.

2.4. Using growth rate models to anticipate production

Since the development of the Monod growth model (Monod, 1949), researchers have continually looked at ways to create black box models to accurately predict biomass yields, product yields, and specific productivities. Early equations still used as the basis in research include the Herbert-Pirt equation (Herbert et al., 1956), where a nutrient balance is performed with the cell acting as the system boundary, and the Luedeking-Piret equation for growth and non growth associated productivity (Luedeking and Piret, 1959). While the Luedeking-Piret equation was initially used for batch production of lactic acid, it has since been applied to

numerous other compounds in both eukaryotes and prokaryotes, including: citric acid (Baei et al., 2008), gluconic acid (Znad et al., 2004), and aminopeptidase (Nandan et al., 2010). While empirical, the equation was able to express the relationship between the production of a compound, cellular growth, and biomass concentration:

$$q_p = \alpha\mu + \beta \qquad (2.1)$$

where q_p is the biomass specific productivity (in grams product per gram dry cell weight per hour) and μ is the growth rate (h^{-1}). The parameters α (gram product per gram dry cell weight) and β (gram product per gram dry cell weight per hour) are the growth and non-growth associated empirical constants, respectively.

There are certain target metabolites (or production under certain fermentation conditions) whose production follows the linear trend predicted by the Luedeking-Piret equation, however, there is a significant subset of compounds which has a production profile vastly different than would be predicted. Assumptions not accounting for diminished cell viability at non-optimum pH, effects at very high cell densities, different nutrient limitations, and toxicity effects of primary or secondary products limit the potential of the model. Rather than throwing out the model, modifications have been made to account for things such as product toxicity (Boonmee et al., 2003), pH levels (Rodríguez et al., 2006), and other growth inhibiting factors (Monteagudo et al., 1997). Additionally, further models have modified the growth associated term significantly, to a hyperbolic model

$$q_p = \alpha\frac{\mu}{\mu+\gamma} + \beta \qquad (2.2)$$

where productivity exhibits diminishing returns as the growth rate reaches its maximum level. Such models have been successful in predicting the production of lysine (Kiss and Stephanopoulos, 1992) and inulinase (Hensing et al., 1995). Further, more extensive models

have been proposed and a list can be seen in reviews by Heijnen (Heijnen, 2009) and Thierie (Thierie, 2013).

More recent production models have started to incorporate reaction elements that are often used in metabolic flux balance analysis models. Kinetic parameters for each reaction used in generating the product (Shimizu et al., 1999), ATP production and maintenance (Zeng, 1995), nucleic acid synthesis (Garcia-Ochoa et al., 2004), simultaneous substrate uptake (Maiti et al., 2010), oxygen availability (Fu and Mathews, 1999), and inhibitive effects of products on individual enzymes (Shinto et al., 2007) have led to increasing use of more structured kinetic models. While many of these structured models are better able to predict quick perturbations to media conditions, they are not usually any more effective than unstructured models in predicting production. The extra steps add additional empirical constants that are specific to a certain set of external conditions (Thilakavathi et al., 2007). However, as these structured models further combine with stoichiometric flux balance analysis models, eventually extensive knowledge gained in regulatory pathways could lead to an overall accurate model of cell function and production under all possible conditions and quick perturbations.

To generate the parameters for these production models, one needs to be able to control for either the concentration of the limiting substrate or growth. The most useful tool for allowing the control of either one of those parameters as an independent variable is the continuous culture or chemostat. The benefit of using a chemostat owes to the constant in and outflow of nutrients and biomass as can be seen in a material balance on the reactor:

$$\frac{dC_i}{dt} = F_{in}C_{i_0} - F_{out}C_i + q_i C_x V \qquad (2.3)$$

where the subscript i represents the specific compound of interest. Under steady state conditions the concentration of each compound is not changing in the reactor over time so the left side of

the equation is equal to zero, allowing one to calculate experimental values for biomass specific production or consumption rates for the controlled in and outflows as well as the concentrations. If one uses biomass as the q-rate, assuming a sterile feed stream, equation 2.4 collapses to

$$q_x = \frac{F_{out}}{V} = \mu = D \text{ (dilution rate)} \qquad (2.4)$$

where the growth rate is controlled based on the set flow rate out of the reactor. Thus, chemostats can be used to generate kinetic parameters to make accurate production predictions in alternative reactor set ups, and has been demonstrated for numerous organisms and products (Gron et al., 1996; Hensing et al., 1995; Nielsen et al., 1991; Thomassen et al., 2005). However, careful consideration needs to be taken to ensure significant mutant populations do not take over the culture in long running chemostats, as cell populations are constantly evolving in response to their environment (Ferenci, 2008; Lenski et al., 1991).

The kinetic parameters from chemostats and their use in models have been extensively applied to determining the best fermentation conditions for commodity chemicals and high value products. Improved specific productivity and yield values have been found for vancomycin in *Amycolatopsis orientalis* (McIntyre et al., 1996), succinate in *E. coli* (Lin et al., 2005), polymers in *Pseudomonas aeruginosa* (Kommedal et al., 2001), Actinorhodin in *Streptomyces coelicolor* (Melzoch et al., 1997), and other compounds (Boonmee et al., 2003; Curless et al., 1989; Thomassen et al., 2005). A significant amount of the studies have indicated maximum productivity at mid-level to near maximum (90%) dilution rates (Boonmee et al., 2003; Fu and Mathews, 1999; McIntyre et al., 1996; Melzoch et al., 1997) rather than at the maximum level. While toxicity is a factor, for *E. coli* and other organisms, inability in maintaining a steady state near maximum growth values has been observed (Majewski and Domach, 1990), affecting productivity near the maximum growth rate in a chemostat. In addition to growth rate effects,

altering oxygen levels and nutrient limitation conditions have also been used in optimizing fermentation conditions. Nitrogen, phosphate, and sulfur limitation have been used for the decrease in biomass yield on substrate and allow for greater carbon flux to products not containing the limiting nutrient (Hua et al., 2004; Kumar and Shimizu, 2011; Wanner and Egli, 1990). Further discussion on nutrient limitation strategies to increase productivity and yield is located in chapter 4. In addition, fermentation strategies varying specific levels of every essential nutrient source have been employed to determine best production, such as with succinate production in *Corynebacterium glutamicum* (Jeon et al., 2013). Overall, chemostat derived production models have proven useful in setting the basis for determining optimal production conditions.

2.5. External factors affecting metabolism and production

2.5.1. Carbon limitation

Bacterial cells take up carbon to generate energy and to synthesize the compounds required for cell growth. Carbon is close to half of the total biomass. The standard stoichiometric formula for *E. coli* dry biomass is $CH_{1.8}N_{0.25}O_{0.4}$ (Pickett et al., 1979), indicating a high percentage of carbon, nitrogen, and oxygen. Therefore, a higher proportion of carbon source is required for cell growth and maintenance than any other nutrient source, and significant research work has been done in a carbon limited media background compared to other nutrient limitation conditions.

When a growing culture is limited for carbon, initial exhaustion of the preferred carbon source leads to metabolic changes within the cell to transition to a stationary phase physiology. The stationary phase sigma factor, RpoS, is responsible for the general stress response and

regulates the transcription of hundreds of stationary phase genes. RpoS is strongly induced under conditions of carbon starvation, and is required for survival under conditions of long term nutrient starvation (Lange, 1991). RpoS is synthesized during both exponential and stationary phase albeit at 10 fold higher levels during exponential growth. The gene product of *sprE* (called SprE or RssB) directs RpoS for degradation by the protease ClpXP (Pratt and Silhavy, 1996). The ClpXP protease rapidly degrades RpoS and is active under exponential growth conditions (Schweder et al., 1996). However, SprE synthesis is inhibited under conditions of carbon starvation and ClpXP is unable to degrade RpoS. Thus, under conditions of carbon starvation, stable RpoS levels increase rapidly to facilitate the expression of stress response genes associated with stationary phase (Zgurskaya et al., 1997).

In addition to inducing RpoS expression, carbon limitation also induces the cellular stringent response (Figure 2.6). Expression of *relA* and *spoT* is associated with increases in increasing the cellular pool of guanosine tetraphospahte, ppGpp. However, only SpoT has been shown to be directly related to carbon starvation (Traxler et al., 2008). ppGpp adjusts the rate of protein synthesis by downregulating growth related gene transcription of tRNA and rRNA (Magnusson et al., 2005).

Cells entering carbon starvation have been shown to enter a quiescent phase marked by a lack of oxygen consumption and heat production (Ballesteros et al., 2001). The inability of the cell to maintain energy levels without having to break down its own structures or enter an environment with a new carbon source leads to the observed metabolism. Additionally, significant levels of protein oxidation have been seen within 48 hours of entering carbon starvation (Ballesteros et al., 2001). Cells under carbon limitation have been shown to have

stationary phase stress response proteins present at a level 20 fold greater than under log phase, with a half life around 40 fold longer (Mandel and Silhavy, 2005).

2.5.2. Nitrogen limitation

Nitrogen is a significant component of biomass composing around fifteen percent of cell dry weight (Chapter 3). Nitrogen is a component of nucleotides, amino acids, and proteins. Nitrogen limitation has been used to keep a non-growing cell metabolically active for producing products that lack nitrogen, such as polyhydroxyalkanoates (Hassan et al., 1996; Ratledge, 2002). Over long term nitrogen starvation, cell viability is significantly lower than other modes of nutrient starvation (e.g. carbon, phosphate) and indicators of metabolism (e.g. oxygen consumption, CO_2 evolution} consistently decrease (Ballesteros et al., 2001). Additionally, glucose uptake rate has been shown to decrease when cells are starved for nitrogen (Wanner and Egli, 1990).

Cells limited for nitrogen first attempt to scavenge nitrogen sources from the external environment by regulating nitrogen uptake into the cell. Initially, external nitrogen limitation is perceived via a decrease in the internal concentration of glutamine (Ninfa and Jiang, 2005). Low glutamate levels lead to the activation of nitrogen regulatory protein C (NtrC) controlled genes, ultimately slowing of cell growth. NtrC regulation leads to increased transcription of genes that code for a high affinity ammonia transporter (AmtB), glutamine synthase (GlnA), and specific amino acid permeases. In addition, other regulatory protein encoding genes (*ntrB*, *ntrC*, *glnK*, and *nac*) and sigma factor 54 (σ^{54} or σ^{N}) dependent genes are upregulated (Zimmer et al., 2000). Regulation is thought to lead to the slowing of cellular growth while increasing cell scavenging for external nitrogen sources.

Once all nitrogen sources are exhausted cell growth ceases (Mandel and Silhavy, 2005). Regulatory changes caused by carbon or nitrogen starvation are linked through the P_{II} family of proteins. P_{II} is a small nitrogen regulatory protein that has an important role in the regulation of nitrogen metablism by detecting levels of alphaketoglutarate (AKG) and ATP. The level of AKG links central carbon and energy metabolism with nitrogen assimilation, indicating one would expect internal metabolic activity shutdown after long periods of nitrogen starvation (Commichau et al., 2006). In *E. coli* entering nitrogen starvation, indicators of metabolism (e.g. oxygen consumption, CO_2 evolution) consistently decrease (Ballesteros et al., 2001). Additionally, nitrogen starvation induces the cellular stringent response, ppGpp (Figure 2.6), downregulating tRNA and rRNA transcription, leading towards decreasing metabolic activity (Magnusson et al., 2005; Traxler et al., 2008). Further interaction between carbon and nitrogen limitation responses has been observed as 60% of genes under control of σ^N are also controlled by σ^S (Dong et al., 2011). However, despite the decreasing metabolic activity, nitrogen starvation has been successfully used to gain an increased extracellular product yield on a given carbon source (Johansson et al., 2005; Marzan and Shimizu, 2011).

2.5.3. Phosphate limitation

Phosphorous in the form of the phosphate ion is critical to cell viability due to its presence in phospholipids, ATP, and nucleic acid backbones. The uptake of phosphate and regulation of metabolism by phosphate has been well studied (Hsieh and Wanner, 2010; Wanner and Egli, 1990). In *E. coli*, phosphate uptake and metabolism is controlled by the Pho regulon which is activated by extracellular phosphate levels. The regulon itself contains at least 31 genes in eight separate operons (*phnCDEFGHIJKLMNOP*, *phoA*, *phoBR*, *phoE*, *phoH*, *psiE*, *pstSCAB*, *phoU*, *ugpBAECQ*) (Hsieh and Wanner, 2010; Matsuzawa et al., 2010), with expression of close

to 400 proteins being correlated to phosphate starvation (VanBogelen et al., 1996). When extracellular phosphate is in excess (> 4 μM), phosphate is primarily taken up by the ABC transporter Pst, with some being taken up by the ion transporter PitA (Jackson et al., 2008). Under conditions of phosphate limitation, the Pst transporter is the primary route for phosphate uptake (Baek and Lee, 2007). To control cellular response to phosphate limitation, regulatory proteins PhoR and PhoB act as a two-component system. PhoR is an inner membrane histidine kinase sensor protein that responds to changes in phosphate levels by undergoing a conformational change that allows it to phosphorylate PhoB. However, recent studies have suggested that Pst may be the primary sensor of environmental phosphate and allow for the modification of PhoR in conjunction with PhoU in order to phosphorylate PhoB (Hsieh and Wanner, 2010). PhoU is thought to act as a chaperone owing to its structural similarity to chaperones in the Hsp70 family in *Aquifex aeolicus* and *Thermotoga maritima* and its cooperative action with PhoR (Lamarche et al., 2008). Regardless of how it is activated, phosphorylated PhoB acts as a transcriptional regulator, controlling expression of the genes in the Pho regulon. The crystal structure, mechanism of DNA binding, and DNA binding pattern for PhoB are all known (Arribas-Bosacoma et al., 2007; Yamane et al., 2008; Yang et al., 2012). If the cells return to a media condition of phosphate excess, increased amounts of PhoU and PstB are correlated with a conformational change in PhoR that leads to dephosphorylation of PhoB (Hsieh and Wanner, 2010). Effects of knocking out *phoB* and *phoR* are known in *E. coli*. Mutations and knockouts of *phoB* lead to cells being unable to synthesize alkaline phosphatase (product of *phoA*) or the high affinity phosphate binding protein (PstS) of the ABC transporter Pst (Yamada et al., 1989). Additionally, mutations in *phoB* appear to lead lower levels of polyphosphate accumulation compared to wild type strains (Ault-riché et al., 1998). In terms of

fermentation characteristics, strains with *phoB* knockouts produce higher biomass concentrations versus wild type *E. coli*, but lower specific glucose uptake rates under phosphate limiting conditions. However, *phoR* knockouts have been shown to decrease acetate production (*phoB* knockouts increase acetate levels under normal media conditions) (Marzan and Shimizu, 2011).

Phosphate limitation has similarities to carbon limitation in that upon entry into starvation, there is an increase in the stationary phase stress response sigma factor (RpoS) (Figure 2.6). While the promoter for *rpoS* is not thought to be affected by the Pho regulon, other interactions can explain the increase in RpoS levels. The mRNA encoding for RpoS has a long leader region with significant secondary structure that occludes the RBS, lowering translational efficiency (Brown and Elliott, 1997). Expression of specific sRNAs and the chaperone Hfq have been shown to relieve the inhibition by binding to the RpoS mRNA leader sequence in *E. coli* and similar organisms (Majdalani et al., 1998). Knockouts to the *pstSCAB* operon have identified an intergenic region between *ptsA* and *ptsB* as critical to RpoS expression under phosphate limiting condtions, indicating one sRNA responsible for increased RpoS translation efficiency could be there (Schurdell et al., 2007). Despite increase expression, these stress response proteins are less stable under phosphate and have significantly shorter half-lives (Mandel and Silhavy, 2005; Peterson et al., 2005). Additionally, unlike carbon limitation, phosphate limitation does not result in an inhibition of SprE activity, allowing the protease ClpXP to degrade RpoS.

An extension of unstable RpoS expression is seen with the short period of continued cell growth and long period of consistent metabolic activity seen after phosphate starvation (Ballesteros et al., 2001; Mandel and Silhavy, 2005). Additionally, higher long term viability has been observed with *Escherichia coli* under phosphate starvation versus carbon or nitrogen starvation (Ballesteros et al., 2001). Phosphate limited cells display increased acid tolerance as a

possible mechanism for survival. Both the glutamate and lysine decarboxylase complexes (GadABC and CadBA), which consume a proton, were found to be contributing factors in long term cell viability under phosphate starvation (Moreau, 2007). Additionally, long term viability in stationary phase has been shown to be related to the expression of PhnE permease, which primarily transports phosphonates and organic phosphate esters (Guillemet and Moreau, 2012). Despite the differences between regulation by phosphate and regulation by alternative nutrients (e.g. carbon, nitrogen), phosphate regulation is still interconnected with carbon and nitrogen regulatory systems by its effects on RpoS and global regulators (Marzan and Shimizu, 2011).

Phosphate limitation appears promising for the production of FFA derived fuels and chemicals in stationary phase. Metabolically active but non-growing cells allow for higher product yields as no initial substrate is directed to biomass formation and substrate uptake rates remains high. Limiting phosphate content in culture media has already been used to produce high value chemicals and proteins (Huber et al., 2011; Johansson et al., 2005; McIntyre et al., 1996). We will explore the effect of limiting nutrients on FFA production in Chapter 4.

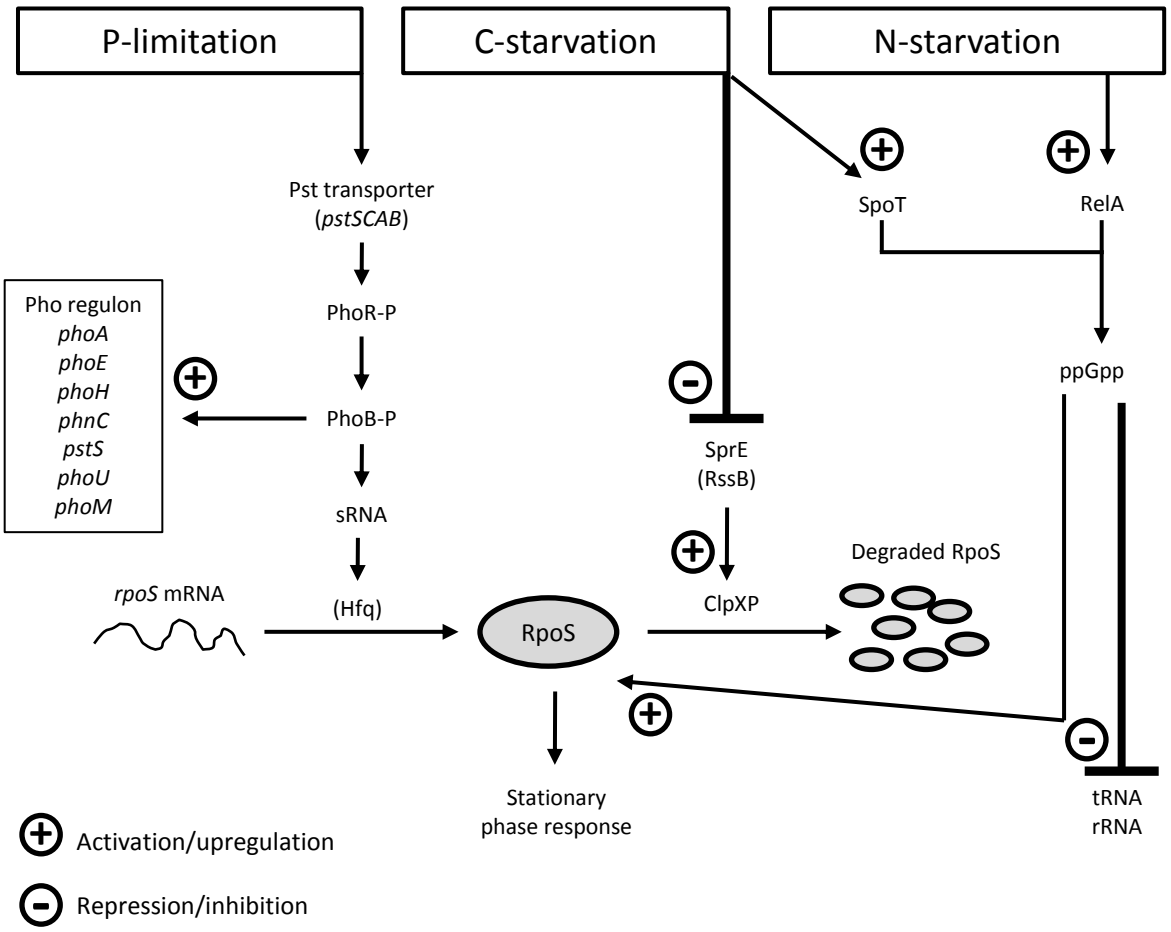


Figure 2.6: Different nutrient limitation conditions and their effects on the stationary phase response of *E. coli*. Phosphate limited cells alleviate inhibition of RpoS translation through a stabilizing sRNA and the chaperone Hfq. However, RpoS is degraded by the protease ClpXP. Under carbon limitation, ClpXP is inhibited, preventing RpoS degradation. Under both carbon and nitrogen limitation, ppGpp levels significantly increase, leading to a decrease in rRNA and tRNA synthesis.

2.6.4. Other external effects

Other limiting nutrients or external conditions also have significant impacts on cell function. Sulfate limitation affects synthesis of proteins, as the CysB master regulator affects cysteine and methionine metabolism (Gyaneshwar et al., 2005b; van der Ploeg et al., 2001). Genes upregulated under sulfate limitation also have significant overlap with those upregulated under nitrogen limitation (Gyaneshwar et al., 2005a). Fermentation parameters for *E. coli* under

sulfate limitation have also been observed similar to those under nitrogen limitation, where glucose uptake slowly decreases and ceases, and there is minimal change in dry cell weight after the exhaustion of the limiting nutrient (Wanner and Egli, 1990). Limiting oxygen availability in the external environment while still maintaining an aerobic to micro-aerobic environment can also impact metabolism. The global regulator system ArcA/B is upregulated when cell growth is limited for oxygen, leading to the production of mixed acid fermentation products (e.g. acetate, lactate) (Matsuoka and Shimizu, 2011). In addition to limiting growth nutrients, external factors such as pH and temperature can have an impact on cellular metabolism. Temperature shifts can cause increases in cellular expression of chaperones (GroEL, DnaK) to help stabilize protein folding under conditions of extreme heat stress (Matsuoka and Shimizu, 2011). The affect of pH stress activates the glutamate dependent acid response, which has the net effect of removing a proton from the cell to attempt to restore proton motive force lost under conditions of low pH (Matsuoka and Shimizu, 2011).

2.7. Conclusions

Microbial production of FFA derived compounds is a promising route to make fuels and chemicals. Our understanding of FFA metabolism, regulation, and cellular response to external stress has enabled our research studies to produce these compounds in engineered microbes. The remainder of the thesis will draw on much of what was discussed in this chapter to set up the experiments for modeling FFA production in *E. coli* and increasing production of FFA derived products.

Many challenges remain in synthesizing FFA derived chemicals. While there has been suggestive evidence regarding membrane stability in FFA overproducing *E. coli*, the true effect on cellular viability under high level production of these compounds remains relatively poorly

understood. Future work could move to organisms that better tolerate these hydrocarbon products, such as cyanobacteria that naturally produce olefins. Additionally, there currently exists a disconnect between high levels of FFA production (where yields have reached over two thirds of theoretical maximum) and FFA derived products (where for many compounds, yields are an order of magnitude lower). Better balancing of enzyme expression in these alternative product pathways could lead to increased production. Additionally, as with the acyl-CoA reductases and decarbonylases, new classes of these enzymes are continually being discovered, with different activities (such as the carboxylic acid reductase) and substrate specificities. It is possible that ideal enzymes for synthesizing each type of product have yet to be found, though efforts are underway to mutagenize acyl-CoA reductases for better selectivity toward specific substrates (Behrouzian et al., 2012). While many areas exist for improvement, Chapters 3 and 4 will focus on finding ideal production conditions for FFAs and developing a kinetic model. Chapter 5 will work toward balancing expression of genes to better convert FFAs to a more valuable fatty alcohol product.

2.8. References

Agnew DE, Stevermer AK, Youngquist JT, Pfleger BF. 2012. Engineering *Escherichia coli* for production of C_{12}-C_{14} polyhydroxyalkanoate from glucose. *Metabolic Engineering* **14**:705–13.

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

Arribas-Bosacoma R, Kim S-K, Ferrer-Orta C, Blanco AG, Pereira PJB, Gomis-Rüth FX, Wanner BL, Coll M, Solà M. 2007. The X-ray crystal structures of two constitutively active mutants of the *Escherichia coli* PhoB receiver domain give insights into activation. *Journal of Molecular Biology* **366**:626–41.

Ault-riché D, Fraley CD, Tzeng C, Ault-riche D, Kornberg A. 1998. Novel Assay Reveals Multiple Pathways Regulating Stress-Induced Accumulations of Inorganic Polyphosphate in *Escherichia coli*. *Journal of Bacteriology* **180**(7):1841-1847.

Azizan A, Black PN. 1994. Use of Transposon TnphoA To Identify Genes for Cell Envelope Proteins of *Escherichia coli* Required for Long-Chain Fatty Acid Transport : the Periplasmic Protein Tsp Potentiates Long-Chain Fatty Acid Transport. *Journal Of Bacteriology* **176**:6653–6662.

Baei MS, Mahmoudi M, Yunesi H. 2008. Short Communication A kinetic model for citric acid production from apple pomac by *Aspergillus niger*. *African Journal of Biotechnology* **7**:3487–3489.

Baek JH, Lee SY. 2007. Transcriptome Analysis of Phosphate Starvation Response in *Escherichia coli*. *Journal of Microbiology and Biotechnology* **17**:244–252.

Ballesteros M, Fredriksson Å, Henriksson J, Nyström T. 2001. Bacterial senescence: protein oxidation in non-proliferating cells is dictated by the accuracy of the ribosomes. *The European Molecular Biology Organization Journal* **20**:5280–5289.

Behrouzian B, Clark L, Zhu Y, Clay M, Karlshoej K. 2012. Fatty alcohol forming acyl reductases (FARS) and methods of use thereof. United States Patent Office 13/171138.

Boonmee M, Leksawasdi N, Bridge W, Rogers PL. 2003. Batch and continuous culture of Lactococcus lactis NZ133 : experimental data and model development. *Biochemical Engineering Journal* **14**:127–135.

Brown L, Elliott T. 1997. Mutations that increase expression of the *rpoS* gene and decrease its dependence on *hfq* function in *Salmonella typhimurium*. *Journal of Bacteriology* **179**:656–662.

Byers DM, Gong H. 2007. Acyl carrier protein: structure-function relationships in a conserved multifunctional protein family. *Biochemistry and cell biology Biochimie et biologie cellulaire* **85**:649–662.

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

Chang MCY, Eachus R a, Trieu W, Ro D-K, Keasling JD. 2007. Engineering Escherichia coli for production of functionalized terpenoids using plant P450s. *Nature chemical biology* **3**:274–7.

Cho H, Cronan JE. 1993. *Escherichia coli* thioesterase I, molecular cloning and sequencing of the structural gene and identification as a periplasmic enzyme. *The Journal of Biological Chemistry* **268**:9238–45.

Commichau FM, Forchhammer K, Stülke J. 2006. Regulatory links between carbon and nitrogen metabolism. *Current Opinion in Microbiology* **9**:167–172.

Cronan JE. 1997. In vivo evidence that acyl coenzyme A regulates DNA binding by the *Escherichia coli* FadR global transcription factor. *Journal of Bacteriology* **179**:1819–1823.

Curless CE, Forrer PD, Mann MB, Fenton DM. 1989. Chemostat Study of Kinetics of Human L ymphokine Synthesis in Recombinant *Escherichia Coli*. *Biotechnology and Bioengineering* **34**:415-421

Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R. 2011. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**:355–9.

DiRusso CC, Black PN, Weimar JD. 1999. Molecular inroads into the regulation and metabolism of fatty acids, lessons from bacteria. *Progress in lipid research* **38**:129–97.

Doan TTP, Carlsson AS, Hamberg M, Bülow L, Stymne S, Olsson P. 2009. Functional expression of five *Arabidopsis* fatty acyl-CoA reductase genes in *Escherichia coli*. *Journal of Plant Physiology* **166**:787–796.

Dong T, Yu R, Schellhorn H. 2011. Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in *Escherichia coli*. *Molecular microbiology* **79**:375–86.

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

Ferenci T. 2008. Bacterial physiology, regulation and mutational adaptation in a chemostat environment. *Advances In Microbial Physiology*. **53**(7):169-229.

Fernandes ND, Kolattukudy PE. 1996. Cloning, sequencing and characterization of a fatty acid synthase-encoding gene from *Mycobacterium tuberculosis* var. bovis BCG. *Gene* **170**:95–99.

Fu W, Mathews AP. 1999. Lactic acid production from lactose by *Lactobacillus plantarum* : kinetic model and effects of pH, substrate, and oxygen. *Biochemical Engineering Journal* **3**:163–170.

Fujita Y, Matsuoka H, Hirooka K. 2007. Regulation of fatty acid metabolism in bacteria. *Molecular Microbiology* **66**:829–39.

Funa N, Ozawa H, Hirata A, Horinouchi S. 2006. Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in *Azotobacter vinelandii*. *Proceedings of the National Academy of Sciences of the United States of America* **103**:6356–6361.

Gao X, Chen J-C, Wu Q, Chen G-Q. 2011. Polyhydroxyalkanoates as a source of chemicals, polymers, and biofuels. *Current opinion in biotechnology* **22**:768–74.

Garcia-Ochoa F, Santos VE, Alcon A. 2004. Chemical structured kinetic model for xanthan production. *Enzyme and Microbial Technology* **35**:284–292.

Gron S, Jochumsen KV, Biedermann K, Emborg C. 1996. Mathematical Modeling of Proteinase A Overproduction by *Saccharomyces cerevisiae*. *Annals Of The New York Academy Of Sciences* **782**:350–362.

Guillemet ML, Moreau PL. 2012. Activation of the cryptic PhnE permease promotes rapid adaptive evolution in a population of *Escherichia coli* K-12 starved for phosphate. *Journal of bacteriology* **194**:253–60.

Gyaneshwar P, Paliy O, McAuliffe J, Jones A, Jordan MI, Kustu S. 2005a. Lessons from *Escherichia coli* genes similarly regulated in response to nitrogen and sulfur limitation. *Proceedings of the National Academy of Sciences of the United States of America* **102**:3453–8.

Gyaneshwar P, Paliy O, McAuliffe J, Popham DL, Jordan MI, Kustu S. 2005b. Sulfur and Nitrogen Limitation in *Escherichia coli* K-12: Specific Homeostatic Responses. *Journal Of Bacteriology* **187**:1074–1090.

Harder ME, Ladenson RC, Schimmel SD, Silbert DF. 1974. Mutants of *Escherichia coli* with Temperature-sensitive Malonyl Coenzyme A-Acyl Carrier Protein Transacylase. *The Journal of Biological Chemistry* **249**:7468–7475.

Hassan M, Blanc PJ, Granger L-M, Pareilleux A, Goma G. 1996. Influence of nitrogen and iron limitations on lipid production by *Cryptococcus curvatus* grown in batch and fed-batch culture. *Process Biochemistry* **31**:355–361.

Heijnen J. 2009. Black box models for growth and product formation. In: Smolke, C, editor. *Handbook for metabolic pathway engineering: Fundamentals*. San Diego, CA: CRC Press.

Hellenbrand J, Biester E-M, Gruber J, Hamberg M, Frentzen M. 2011. Fatty acyl-CoA reductases of birds. *BMC Biochemistry* **12**:64.

Hensing MCM, Vrouwenvelder JS, Hellinga C, Vandijken JP, Pronk JT. 1995. Use of chemostat data for modelling extracellular-inulinase production by kluyveromyces marxianus in a high-cell-density fed-batch process. *Journal of Fermentation and Bioengineering* **79**:54–58.

Herbert D, Elsworth R, Telling R. 1956. The continuous culture of bacteria; a theoretical and experimental study. *Journal of General Microbiology* **14**:601–622.

Hofvander P, Doan TTP, Hamberg M. 2011. A prokaryotic acyl-CoA reductase performing reduction of fatty acyl-CoA to fatty alcohol. *FEBS Letters* **585**:3538–43.

Hoover SW, Youngquist JT, Angart PA, Withers ST, Lennen RM, Pfleger BF. 2012. Isolation of Improved Free Fatty Acid Overproducing Strains of *Escherichia coli* via Nile Red Based High-Throughput Screening. *Environmental Progress and Sustainable Energy* **31**:17–23.

Howard TP, Middelhaufe S, Moore K, Edner C, Kolak DM, Taylor GN, Parker DA, Lee R, Smirnoff N, Aves SJ, Love J. 2013. Synthesis of customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **110**:7636–7614.

Hsieh Y-J, Wanner BL. 2010. Global regulation by the seven-component Pi signaling system. *Current Opinion in Microbiology* **13**:198–203.

Hua Q, Yang C, Oshima T, Mori H, Shimizu K. 2004. Analysis of Gene Expression in *Escherichia coli* in Response to Changes of Growth-Limiting Nutrient in Chemostat Cultures. *Society* **70**:2354–2366.

Huber R, Roth S, Rahmen N, Büchs J. 2011. Utilizing high-throughput experimentation to enhance specific productivity of an *E. coli* T7 expression system by phosphate limitation. *BMC Biotechnology* **11**:22.

Ingram LO, Conway T, Clark DP, Sewell GW, Preston JF. 1987. Genetic engineering of ethanol production in *Escherichia coli*. *Applied and Environmental Microbiology* **53**:2420–5.

Jackson RJ, Binet MRB, Lee LJ, Ma R, Graham AI, McLeod CW, Poole RK. 2008. Expression of the PitA phosphate/metal transporter of *Escherichia coli* is responsive to zinc and inorganic phosphate levels. *FEMS microbiology letters* **289**:219–24.

Jenkins LS, Nunn WD. 1987. Genetic and molecular characterization of the genes involved in short-chain fatty acid degradation in *Escherichia coli*: the *ato* system. *Journal of Bacteriology* **169**:42–52.

Jeon J-M, Rajesh T, Song E, Lee H-W, Lee H-W, Yang Y-H. 2013. Media optimization of *Corynebacterium glutamicum* for succinate production under oxygen-deprived condition. *Journal of microbiology and biotechnology* **23**:211–7.

Johansson L, Lindskog A, Silfversparre G, Cimander C, Nielsen KF, Lidén G. 2005. Shikimic acid production by a modified strain of *E. coli* (W3110.shik1) under phosphate-limited and carbon-limited conditions. *Biotechnology and Bioengineering* **92**:541–52.

Kalscheuer R, Stölting T, Steinbüchel A. 2006. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology* **152**:2529–2536.

Kiss RD, Stephanopoulos G. 1992. Metabolic characterization of a L-lysine-producing strain by continuous culture. *Biotechnology and Bioengineering* **39**:565–574.

Kommedal R, Bakke R, Dockery J, Stoodley P. 2001. Modelling production of extracellular polymeric substances in a *Pseudomonas aeruginosa* chemostat culture. *Water Science and Technology* **43**:129–134.

Kumar R, Shimizu K. 2011. Transcriptional regulation of main metabolic pathways of cyoA, cydB, fnr, and fur gene knockout *Escherichia coli* in C-limited and N-limited aerobic continuous cultures. *Microbial Cell Factories* **10**:3.

Lamarche MG, Wanner BL, Crépin S, Harel J. 2008. The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis. *FEMS microbiology reviews* **32**:461–73.

Lange R. 1991. Identification of a central regulator of stationary-phase gene expression in *Escherichia coli*. *Molecular Microbiology* **5**:49–59.

Lennen RM, Braden DJ, West RA, Dumesic JA, Pfleger BF. 2010. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnology and Bioengineering* **106**:193–202.

Lennen RM, Kruziki M a, Kumar K, Zinkel R a, Burnum KE, Lipton MS, Hoover SW, Ranatunga DR, Wittkopp TM, Marner WD, Pfleger BF. 2011. Membrane stresses induced by overproduction of free fatty acids in *Escherichia coli*. *Applied and Environmental Microbiology* **77**:8114–28.

Lennen RM, Pfleger BF. 2012. Engineering *Escherichia coli* to synthesize free fatty acids. *Trends in Biotechnology* **30**:659–67.

Lennen RM, Politz MG, Kruziki M a, Pfleger BF. 2013. Identification of transport proteins involved in free fatty acid efflux in *Escherichia coli*. *Journal of Bacteriology* **195**:135–44.

Lenski RE, Rose MR, Simpson SC, Tadler SC, Url S, Scott C. 1991. Long term experimental evolution in *Escherichia coli*. I. Adaptation and divergence during 2000 generations.. *The American Naturalist* **138**:1315–1341.

Li M, Zhang X, Agrawal A, San K-Y. 2012. Effect of acetate formation pathway and long chain fatty acid CoA-ligase on the free fatty acid production in *E. coli* expressing acy-ACP thioesterase from *Ricinus communis*. *Metabolic engineering* **14**:380–7.

Lin H, Bennett GN, San K-Y. 2005. Chemostat culture characterization of *Escherichia coli* mutant strains metabolically engineered for aerobic succinate production: a study of the modified metabolic network based on metabolite profile, enzyme activity, and gene expression profile. *Metabolic engineering* **7**:337–52.

Lu X, Vora H, Khosla C. 2008. Overproduction of free fatty acids in *E. coli*: implications for biodiesel production. *Metabolic Engineering* **10**:333–339.

Luedeking R, Piret EL. 1959. A kinetic study of lactic acid fermentation. Batch process at controlled pH. *Journal of Biochemical and Microbiological Technology and Engineering* **1**:393–412.

Magnuson K, Jackowski S, Rock CO, Jr JEC. 1993. Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiological Reviews* **57**:522–542.

Magnusson LU, Farewell A, Nyström T. 2005. ppGpp: a global regulator in *Escherichia coli*. *Trends in microbiology* **13**:236–42.

Maiti SK, Singh KP, Lantz AE, Bhushan M, Wangikar PP. 2010. Substrate uptake, phosphorus repression, and effect of seed culture on glycopeptide antibiotic production: process model development and experimental validation. *Biotechnology and Bioengineering* **105**:109–120.

Majdalani N, Cunning C, Sledjeski D, Elliott T, Gottesman S. 1998. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proceedings of the National Academy of Sciences of the United States of America* **95**:12462–7.

Majewski RA, Domach MM. 1990. Chemostat-Cultivated *Escherichia coli* at High Dilution Rate : Multiple Steady States and Drift. *Biotechnology and Bioengineering* **36**:179–190.

Mandel MJ, Silhavy TJ. 2005. Starvation for Different Nutrients in Escherichia coli Results in Differential Modulation of RpoS Levels and Stability Starvation for Different Nutrients in *Escherichia coli* Results in Differential Modulation of RpoS Levels and Stability. *Journal of Bacteriology* **187**:434–442.

Marzan LW, Shimizu K. 2011. Metabolic regulation of *Escherichia coli* and its *phoB* and *phoR* genes knockout mutants under phosphate and nitrogen limitations as well as at acidic condition. *Microbial Cell Factories* **10**:39.

Matsuoka Y, Shimizu K. 2011. Metabolic regulation in *Escherichia coli* in response to culture environments via global regulators. *Biotechnology Journal* **6**:1–12.

Matsuzawa M, Katsuyama Y, Funa N, Horinouchi S. 2010. Alkylresorcylic acid synthesis by type III polyketide synthases from rice Oryza sativa. *Phytochemistry* **71**:1059–1067.

Maung T, Gustafson C. 2013. Economic Impact of Harvesting Corn Stover under Time Constraint: The Case of North Dakota. *Economics Research International* **2013**:1-13.

McIntyre JJ, Bull a T, Bunch a W. 1996. Vancomycin production in batch and continuous culture. *Biotechnology and Bioengineering* **49**:412–20.

Melzoch K, De Mattos MJT, Neijssel OM. 1997. Production of actinorhodin by *Streptomyces coelicolor* A3(2) grown in chemostat culture. *Biotechnology and Bioengineering* **54**:577–582

Mendez-Perez D, Begemann MB, Pfleger BF. 2011. Modular synthase-encoding gene involved in α-olefin biosynthesis in *Synechococcus sp. strain PCC 7002*. *Applied and Environmental Microbiology* **77**:4264–7.

Monod J. 1949. The Growth of Bacterial Cultures. *Annual Review of Microbiology* **3**:371–394.

Monteagudo M, Rodr L, Rinco J, Fuertes J. 1997. Kinetics of Lactic Acid Fermentation by *Lactobacillus delbrueckii* Grown on Beet Molasses. *Journal of Chemical Technology and Biotechnology* **68**:271-276.

Moreau PL. 2007. The Lysine Decarboxylase CadA Protects *Escherichia coli* Starved of Phosphate against Fermentation Acids. *Journal of Bacteriology* **189**:2249–2261.

Mudge SM, Belanger SE, Nielsen AM. 2008. Fatty Alcohols: Anthropogenic and Natural Occurence in the Environment. Cambridge, UK: The Royal Society of Chemistry.

Nakano C, Ozawa H, Akanuma G, Funa N, Horinouchi S. 2009. Biosynthesis of Aliphatic Polyketides by Type III Polyketide Synthase and Methyltransferase in *Bacillus subtilis*. *Journal Of Bacteriology* **191**:4916–4923.

Nandan A, Gaurav A, Pandey A, Nampoothiri KM. 2010. Arginine Specific Aminopeptidase from *Lactobacillus brevis* **53**:1443–1450.

Nielsen J, Nikolajsen K, Villadsen J. 1991. Structured modeling of a microbial system: II. Experimental verification of a structured lactic acid fermentation model. *Biotechnology and Bioengineering* **38**:11–23.

Ninfa AJ, Jiang P. 2005. PII signal transduction proteins: sensors of alpha-ketoglutarate that regulate nitrogen metabolism. *Current opinion in microbiology* **8**:168–73.

Nunn WD. 1986. A Molecular View of Fatty Acid Catabolism in *Escherichia coli*. *Microbiological Reviews* **50**:179–192.

Peterson CN, Mandel MJ, Silhavy TJ. 2005. *Escherichia coli* Starvation Diets : Essential Nutrients Weigh in Distinctly. *Journal of Bacteriology* **187**:7549–7553.

Pfeifer BA, Khosla C. 2001. Biosynthesis of Polyketides in Heterologous Hosts Biosynthesis of Polyketides in Heterologous Hosts. *Microbiology and Molecular Biology Reviews* **65**:106–118.

Pickett AM, Bazin M, Topiwala H. 1979. Growth and composition of *Escherichia coli* subjected to square-wave perturbations in nutrient supply—Effect of varying frequencies. *Biotechnology and Bioengineering* **21**:1043–1055.

van der Ploeg J, Eichhorn E, Leisinger T. 2001. Sulfonate-sulfur metabolism and its regulation in *Escherichia coli*. *Archives of Microbiology* **176**:1–8.

Pratt LA, Silhavy TJ. 1996. The response regulator SprE controls the stability of RpoS. *Proceedings of the National Academy of Sciences of the United States of America* **93**:2488–2492.

Radmacher E, Alderwick LJ, Besra GS, Brown AK, Gibson KJC, Sahm H, Eggeling L. 2005. Two functional FAS-I type fatty acid synthases in *Corynebacterium glutamicum*. *Microbiology* **151**:2421–2427.

Ranganathan S, Tee TW, Chowdhury A, Zomorrodi AR, Yoon JM, Fu Y, Shanks J V, Maranas CD. 2012. An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering* **14**:687–704.

Ratledge C. 2002. Regulation of lipid accumulation in oleaginous micro-organisms. *Biochemical Society Transactions* **30**:1047–50.

Reiser S, Somerville C. 1997. 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. *Journal of Bacteriology* **179**:2969–2975.

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

Rodríguez J, Kleerebezem R, Lema JM, Van Loosdrecht MCM. 2006. Modeling product formation in anaerobic mixed culture fermentations. *Biotechnology and Bioengineering* **93**:592–606.

Rowland O, Domergue F. 2012. Plant fatty acyl reductases: enzymes generating fatty alcohols for protective layers with potential for industrial applications. *Plant Science* **193-194**:28–38.

Rude M a, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A. 2011. Terminal olefin (1-alkene) biosynthesis by a novel p450 fatty acid decarboxylase from *Jeotgalicoccus* species. *Applied and Environmental Microbiology* **77**:1718–1727.

Rupilius W, Ahmad S. 2006. The Changing World of Oleochemicals. *Palm Oil Developments* **44**:15–28.

Schirmer A, Rude M a, Li X, Popova E, Del Cardayre SB. 2010. Microbial biosynthesis of alkanes. *Science* **329**:559–562.

Schurdell MS, Woodbury GM, McCleary WR. 2007. Genetic evidence suggests that the intergenic region between *pstA* and *pstB* plays a role in the regulation of *rpoS* translation during phosphate limitation. *Journal of bacteriology* **189**:1150–3.

Schweder T, Lee KH, Lomovskaya O, Matin A, Schweder T, Lee K, Lomovskaya O. 1996. Regulation of *Escherichia coli* starvation sigma factor (sigma s) by ClpXP protease. *Journal of Bacteriology* **178**:470-476.

Schweizer E, Hofmann J. 2004. Microbial Type I Fatty Acid Synthases (FAS): Major Players in a Network of Cellular FAS Systems. *Microbiology and Molecular Biology Reviews* **68**:501–517.

Shimizu H, Takiguchi N, Tanaka H, Shioya S. 1999. Metabolic Reaction Model for Lysine Production. *Metabolic Engineering* **308**:299–308.

Shinto H, Tashiro Y, Yamashita M, Kobayashi G, Sekiguchi T, Hanai T, Kuriya Y, Okamoto M, Sonomoto K. 2007. Kinetic modeling and sensitivity analysis of acetone-butanol-ethanol production. *Journal of Biotechnology* **131**:45–56.

Stasiuk M, Kozubek a. 2010. Biological activity of phenolic lipids. *Cellular and molecular life sciences CMLS* **67**:841–860.

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

Stuible HP, Meurer G, Schweizer E. 1997. Heterologous expression and biochemical characterization of two functionally different type I fatty acid synthases from *Brevibacterium ammoniagenes*. *The Federation of European Biochemical Societies Journal* **247**:268–273.

Teerawanichpan P, Qiu X. 2010. Fatty acyl-CoA reductase and wax synthase from *Euglena gracilis* in the biosynthesis of medium-chain wax esters. *Lipids* **45**:263–273.

Thierie J. 2013. Computing and interpreting specific production rates in a chemostat in steady state according to the Luedeking-Piret model. *Applied biochemistry and biotechnology* **169**:477–92.

Thilakavathi M, Basak T, Panda T. 2007. Modeling of enzyme production kinetics. *Applied Microbiology and Biotechnology* **73**:991–1007.

Thomassen YE, Verkleij AJ, Boonstra J, Verrips CT. 2005. Specific production rate of VHH antibody fragments by *Saccharomyces cerevisiae* is correlated with growth rate, independent of nutrient limitation. *Journal of Biotechnology* **118**:270–7.

Traxler MF, Summers SM, Nguyen H-T, Zacharia VM, Hightower GA, Smith JT, Conway T. 2008. The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Molecular Microbiology* **68**:1128–1148.

VanBogelen R a, Olson ER, Wanner BL, Neidhardt FC. 1996. Global analysis of proteins synthesized during phosphorus restriction in *Escherichia coli*. *Journal Of Bacteriology* **178**:4344–4366.

Wanner U, Egli T. 1990. Dynamics of microbial growth and cell composition in batch culture. *FEMS Microbiology Reviews* **75**:19–44.

Willis RM, Wahlen BD, Seefeldt LC, Barney BM. 2011. Characterization of a fatty acyl-CoA reductase from *Marinobacter aquaeolei* VT8: a bacterial enzyme catalyzing the reduction of fatty acyl-CoA to fatty alcohol. *Biochemistry* **50**:10550–8.

Xu P, Gu Q, Wang W, Wong L, Bower AGW, Collins CH, Koffas M a G. 2013. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications* **4**:1409.

Yamada M, Makino K, Amemura M, Shinagawa H, Nakata a. 1989. Regulation of the phosphate regulon of *Escherichia coli*: analysis of mutant phoB and phoR genes causing different phenotypes. *Journal of Bacteriology* **171**:5601–6.

Yamane T, Okamura H, Ikeguchi M, Nishimura Y, Kidera A. 2008. Water-mediated interactions between DNA and PhoB DNA-binding/transactivation domain: NMR-restrained molecular dynamics in explicit water environment. *Proteins* **71**:1970–83.

Yang C, Huang T-W, Wen S-Y, Chang C-Y, Tsai S-F, Wu W-F, Chang C-H. 2012. Genome-wide PhoB binding and gene expression profiles reveal the hierarchical gene regulatory network of phosphate starvation in *Escherichia coli*. *PloS one* **7**:e47314.

Yoshimura M, Oshima T, Ogasawara N. 2007. Involvement of the YneS/YgiH and PlsX proteins in phospholipid biosynthesis in both *Bacillus subtilis* and *Escherichia coli*. *BMC Microbiology* **7**:69.

Youngquist JT, Rose JP, Pfleger BF. 2013. Free fatty acid production in *Escherichia coli* under phosphate-limited conditions. *Applied Microbiology and Biotechnology* **97**(11):5149-5159.

Zeng a P. 1995. A kinetic model for product formation of microbial and mammalian cells. *Biotechnology and Bioengineering* **46**:314–324.

Zgurskaya HI, Keyhan M, Matin a. 1997. The sigma S level in starving *Escherichia coli* cells increases solely as a result of its increased stability, despite decreased synthesis. *Molecular microbiology* **24**:643–51.

Zha W, Rubin-Pitel SB, Shao Z, Zhao H. 2009. Improving cellular malonyl-CoA level in *Escherichia coli* via metabolic engineering. *Metabolic Engineering* **11**:192–198.

Zhang F, Carothers JM, Keasling JD. 2012a. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature biotechnology* **30**:354–9.

Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD. 2012b. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic Engineering* **14**:653–60.

Zhang F, Rodriguez S, Keasling JD. 2011. Metabolic engineering of microbial pathways for advanced biofuels production. *Current opinion in biotechnology* **22**:775–83.

Zhang Y-M, Marrakchi H, Rock CO. 2002. The FabR (YijC) transcription factor regulates unsaturated fatty acid biosynthesis in *Escherichia coli*. *The Journal of biological chemistry* **277**:15558–65.

Zheng YN, Li L, Liu Q, Yang J, Wang X, Liu W, Xu X, Liu H, Zhao G, Xian M. 2012. Optimization of fatty alcohol biosynthesis pathway for selectively enhanced production of C12/14 and C16/18 fatty alcohols in engineered *Escherichia coli*. *Microbial Cell Factories* **11**:65.

Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender R a, Kustu S. 2000. Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: Scavenging as a defense against nitrogen limitation. *Proceedings of the National Academy of Sciences of the United States of America* **97**:14674–14679.

Znad H, Markoš J, Baleš V. 2004. Production of gluconic acid from glucose by *Aspergillus niger*: growth and non-growth conditions. *Process Biochemistry* **39**:1341–1345.

Chapter 3: Kinetic modeling of free fatty acid production in *Escherichia coli* based on continuous cultivation of a plasmid free strain[1]

3.1. Introduction

The overall theme of this thesis research is to produce renewable alternatives to petrochemicals. Our strategy was to engineer metabolic pathways to overproduce free fatty acids (FFAs), and then convert them to a range of chemical products via enzymatic and non-enzymatic reactions. Chapter 5 explores one of these alternative products, fatty alcohols. This chapter explores the construction of a plasmid-free, FFA producing strain and the development of an experimental apparatus and computational models to explore the connection between growth rate, nutrients, and FFA production. Chapter 4 will discuss studies of nutrient limitation that expand upon the themes uncovered herein. Chapter 3 implements a metabolic engineering strategy developed in our lab and by others (Lu et al., 2008; Steen et al., 2010) to produce FFA from renewable feedstocks. Using our approach, we have achieved yields approaching 33% of the theoretical without incorporating antibiotic-requiring genetic elements. The strategy relies on expression of an acyl-ACP thioesterase for the production of FFA and elimination of multiple steps of the fatty acid β-oxidation pathway to prevent the consumption of the fatty acid product. It is important to note that this work was made possible due to the efforts of the Pfleger lab involving the overproduction of FFA in *E. coli* (Lennen et al., 2010).

The finite nature of petroleum reserves and growing environmental concern regarding greenhouse gas emissions are motivating the development of alternative transportation fuels for gasoline, diesel, and jet engines. (Du et al., 2011). While many energy solutions, e.g. ethanol, H_2, electricity, are viable for personal transportation, the heavy transportation and aviation markets

[1] Portions of this chapter was published were published in Youngquist et al., *Biotechnol Bioeng*, (2012)

require specific properties that are best provided by medium chain hydrocarbons. Of the many types of hydrocarbons, fatty acid derived compounds are attractive because fatty acids are produced in high amounts, for synthesis of membranes, in most microbial species (Handke et al., 2011; Lennen et al., 2012; Lennen and Pfleger, 2013). By coupling the ability of an engineered microbe to consume biomass derived sugars with metabolic pathways to produce and convert fatty acids to biofuel compounds at high yield, sustainable processes for synthesizing liquid transportation fuels could be developed. Recent reports have demonstrated that fatty acids can be enzymatically converted to biofuel products such as alkanes (Schirmer et al., 2010), fatty alcohols (Steen et al., 2010), fatty acid methyl esters (e.g. biodiesel) (Kalscheuer et al., 2006; Steen et al., 2010), and olefins (Beller et al., 2010; Mendez-Perez et al., 2011; Rude et al., 2011), that could be purified from culture. Unfortunately, each of these pathways has been shown to have low yield on fatty acids and will require significant protein and metabolic engineering. Alternatively, FFAs can be purified from culture and catalytically decarboxylated to yield alkanes (Lennen et al., 2010). To be economically viable, both strategies will require strains to approach theoretical yields of fatty acids (or reduced products) on glucose. As such, cultivation strategies that achieve the highest performance for a given strain need to be designed and tested.

Process optimization has been used to increase the yield and productivity of certain compounds (antibiotics, lactic acid, *etc*) under controlled conditions (Khanna and Srivastava, 2005; Todorov et al., 2004; Zhu et al., 2010). Process parameters, such as temperature, pH, oxygen levels, nutrient limitation, and particularly growth rate, impact product titer, yield, and productivity. In order to get an accurate indication of what parameters are important to production, it is necessary to have a controlled system where only one variable is manipulated at a time. A continuous culture (chemostat) allows control over growth rate of an organism and

development of a kinetic model of a cell culture that produces a desired metabolite. In the case of anabolic (i.e. secondary) metabolites, production rates frequently increase under conditions of slow growth, e.g. transition and stationary phase of batch cultures, or slow dilution rates in fed-batch and continuous culture (Gramajo et al., 1993). Kinetic models facilitate the design of non-batch cultivation strategies and can be used to determine optimal conditions to maximize strain performance. Models of non-genetically engineered strains of *E. coli* are available (Fru et al., 2011; Goochee et al., 1985; Nielsen and Villadsen, 1992; Orencio-Trejo et al., 2010; Valgepea et al., 2011), providing both guidance and comparison for evaluating biofuel producing strains.

In this chapter, we present a kinetic model for a FFA producing strain of *E. coli*. In past work, we demonstrated that overexpression of the C_{12} specific acyl-ACP thioesterase from *Umbellularia californica* (BTE) in *E. coli* leads to the production of C_{12} and C_{14} FFA in high titers (Hoover et al., 2011; Lennen et al., 2010). However, expression of this enzyme must be optimized, at an intermediate level, to maximize FFA titer (Hoover et al., 2011; Lennen et al., 2010) and avoid toxicities generated by high rates of FFA production (Lennen et al., 2011). To facilitate chemostat experiments, a plasmid-free strain of *E. coli* harboring three BTE expression cassettes on the chromosome was constructed. Using chemostats, the effect of two parameters, growth rate and carbon to nitrogen ratio, on FFA production was studied. Measurements of CO_2, FFA, biomass, residual glucose, and fermentation products were used to perform a carbon mass balance that was closed with 2-12% carbon unaccounted. The collected data was used to determine model parameters such as biomass yield, maintenance energy requirement, and biomass specific FFA productivity as a function of dilution rate. The model was subsequently used to determine conditions that maximize FFA productivity and yield.

3.2. Results and Discussion

3.2.1. Construction and evaluation of a plasmid-free FFA overproducing E. coli

In past work, plasmid-based, heterologous expression of a codon-optimized gene encoding an acyl-ACP thioesterase (BTE) from *Umbellularia californica* in *E. coli* resulted in production of medium-chain ($C_{12:0}$, $C_{12:1}$, $C_{14:0}$) FFA (Hoover et al., 2011; Lennen et al., 2010). To study FFA production in the absence of plasmids, an expression cassette consisting of the BTE gene under the control of the strong, inducible P_{TRC} promoter (Brosius et al., 1985) was constructed (Figure 1) and integrated onto the *E. coli* chromosome in place of *fadD* (strain TY01), *fadD* and *fadE* (TY03), and *fadD*, *fadE*, and *fadAB* (TY05). These insertion sites were selected owing to both the involvement of these genes in fatty acid beta-oxidation (Fujita et. al., 2007) and previous data showing increased titers when these genes are not present (Hoover et. al., 2011). A corresponding set of control strains was made by integrating a non-functional variant, BTE-H204A, in place of the same genes to generate strains TY02, TY04, and TY06. For comparison, *E. coli* RL08 (K12 MG1655 Δ*araBAD* Δ*fadD*) was transformed with either pTRC99A-BTE or pTRC99A-BTE-H204A. Maximum induction of BTE from high-copy (>25) plasmids (e.g. pTRC99A, pUC19) using either the P_{TRC} or P_{BAD} promoters results in reduced FFA yields and cell viability (Hoover et al., 2011; Lennen et al., 2010; Lennen et al., 2011). The highest FFA titers are achieved when cultures harboring pTRC99A-BTE are induced with 50 µM IPTG (Hoover et al. 2011). In contrast, the maximum FFA titers were achieved from integrant strains (TY01-TY06) when induced with 1 mM IPTG (a saturating concentration). The FFA titers of *E. coli* TY01-TY06 and the two plasmid containing strains grown in MOPS minimal media supplemented with 0.4% glucose (Figure 3.1) and LB + 0.4% glucose (Figure 3.8) were compared under optimal induction conditions. The triple integrant, TY05, produced titers that

were equivalent to the plasmid containing strain whereas the single integrant, TY01, and double integrant, TY03, produced approximately one third and two-thirds of the FFA titer observed in the plasmid based system. As expected, all negative controls produced negligible amounts of C_{12}, $C_{12:1}$, and C_{14} fatty acids. While FFA titers trended similarly in each medium, titers were consistently higher in the rich LB medium compared to the MOPS minimal medium.

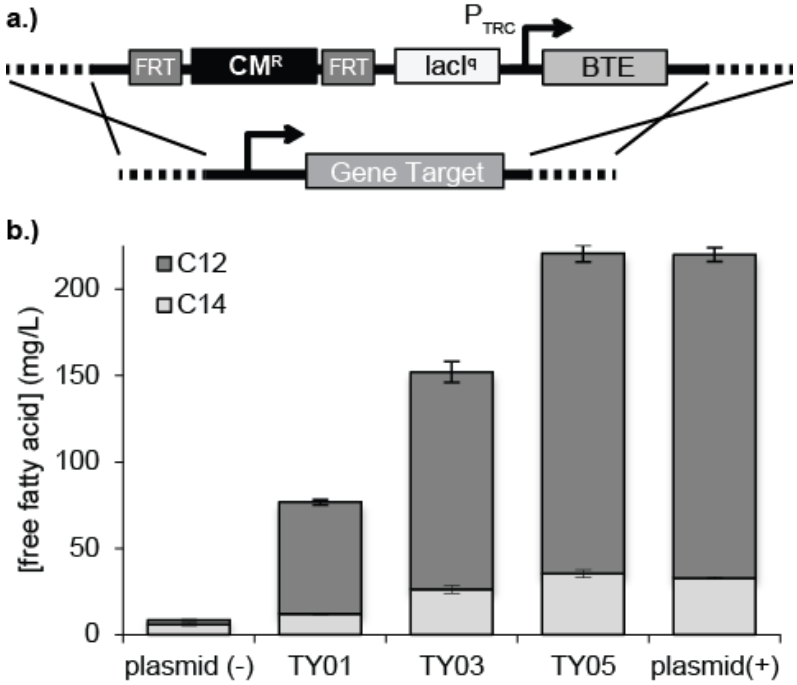


Figure 3.1: a.) Schematic of the BTE integration cassette consisting of homologous recombination target sequences (dotted lines), flippase recombinase target (FRT) sites, chloramphenicol resistance cassette (CM^R), IPTG responsive regulatory protein ($lacI^Q$), P_{TRC}, and BTE. b.) Titer of FFA species at 24 hours after induction for strains grown in MOPS minimal media + 0.4% glucose. Plasmid(+) represents pTRC99A-BTE and plasmid(-) represents pTRC99A-BTE-H204A.

3.2.2. Continuous cultivation of a FFA producing E. coli at various growth rates

The parameters of a kinetic model describing cell growth, glucose uptake, and FFA production in *E. coli* TY05 were determined from data collected from chemostat experiments. A culture of *E. coli* TY05 was fed MOPS-based, minimal glucose medium containing 1 mM IPTG for induction of the chromosomal copies of BTE. The medium was carbon limited because

additional nitrogen was introduced in the form of ammonium hydroxide, which was used to maintain pH and minimize foaming. Three sets of continuous culturing experiments were performed. In each, the bioreactor was inoculated with a culture grown from a single colony of *E. coli* TY05 and operated in batch mode for 9 hours. At this point, fresh media was fed continuously and the reactor was operated as a chemostat as described below in the materials and methods. For each experimental variable, a steady state was established prior to collecting biomass, supernatant, and foam samples over a 12-24 hour period.

Table 3.1: Steady state concentrations of assayed metabolites from continuous culture of TY05 in MOPS minimal media + 0.4% glucose at various dilution rates

Dilution rate h^{-1}	**Biomass g/L**	**Glucose g/L**	**C12 FA mg/L**	**C14 FA mg/L**	**CO_2 ppm**	**Acetate** g/L**	**Carbon recovery %**
0.05±0.002	1.11±0.05	*	221±6	56±4	1,600±100	*	98±4
0.10±0.001	1.21±0.07	*	172±5	41±5	2,900±200	*	93±4
0.15±0.002	1.29±0.06	*	137±8	39±2	4,100±200	*	91±3
0.20±0.001	1.47±0.07	*	127±2	34±1	5,500±100	*	95±2
0.25±0.001	1.54±0.09	*	#	#	6,600±100	*	88±3
0.31±0.002	1.64±0.08	*	91±7	26±4	7,500±300	*	94±3
0.35±0.002	1.67±0.09	0.02±0.01	67±8	12±1	7,700±100	0.20±0.01	94±3
0.50±0.001	1.24±0.07	1.23±0.02	35±2	5.4±0.5	6,300±100	0.16±0.01	92±3

*Below limit of detection (0.1 mM for glucose, 1 mM for acetate),
#Not measured
** Other fermentation products were not observed under any condition.
C12 FA is the combination of dodecanoic acid and dodecenoic acid
C14 FA is the combination of tetradecanoic acid and tetradecenoic acid

In the first experiment, *E. coli* TY05 was continuously cultured under different dilution rates with a feed glucose concentration of 4 g/L (0.4%). For each dilution rate, the concentrations of biomass (calculated from OD_{600}), residual glucose, FFA, CO_2, O_2 and fermentation products (e.g. acetate, formate, ethanol, succinate, pyruvate, lactate, glycerol, and xylitol) were measured (Table 3.1). Biomass concentration and CO_2 evolution increased with dilution rate, with the exception of the highest dilution rate (0.5 h^{-1}). Glucose was not observed above the instrumental limit of detection until the highest dilution rate tested. Absolute FFA titers decreased with

dilution rate. Of the monitored fermentation products, only acetate was observed and only when dilution rates were greater than 0.35 h^{-1}. Using mass balances (equation 1) for biomass, glucose, and FFA, the rates of production for each species (q_i) were calculated (Figure 3.2ac).

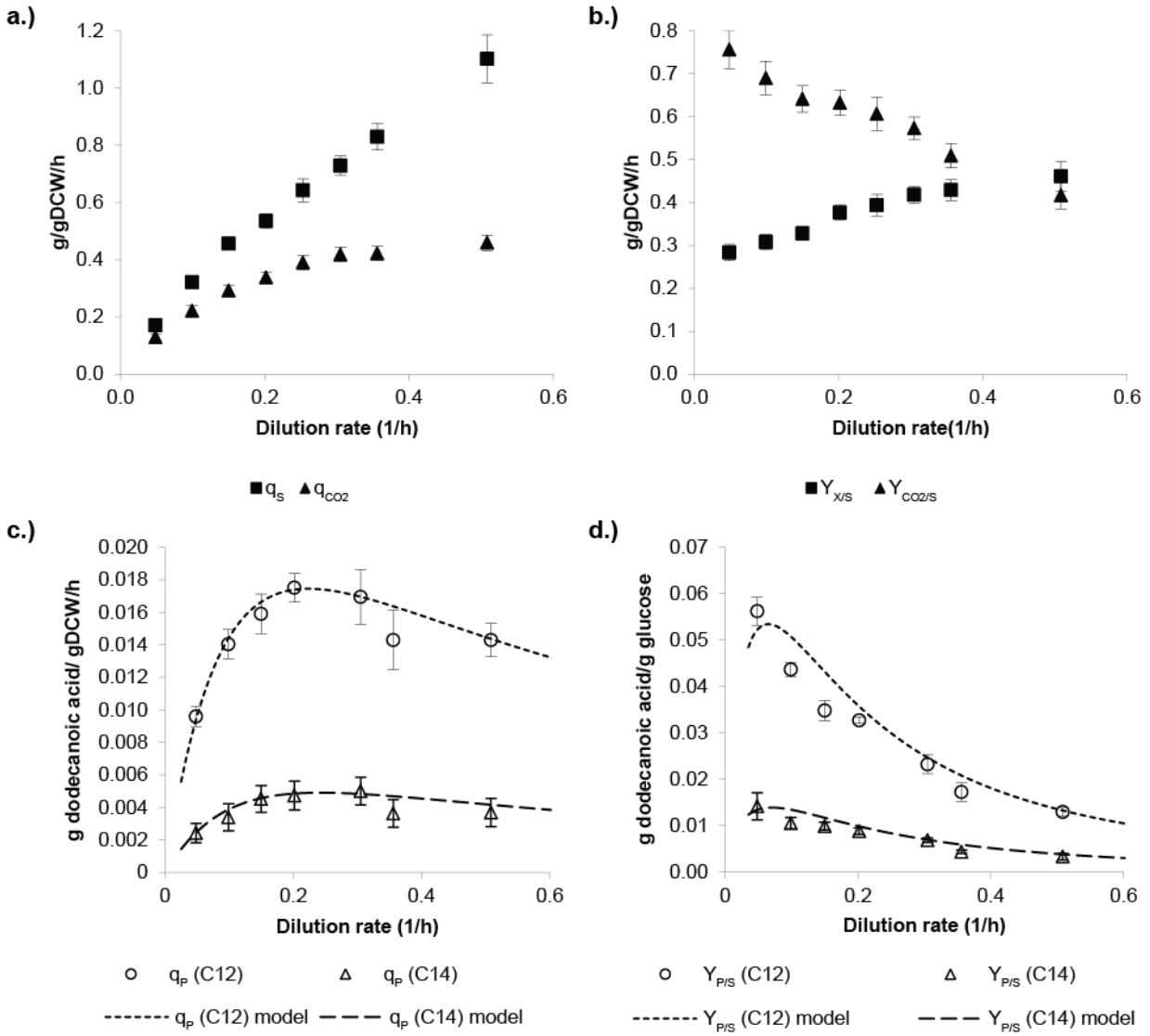


Figure 3.2: Productivity (A,C) and yield (B,D) values calculated for each dilution rate tested. A.) Specific glucose uptake rate and CO_2 productivity calculated mass balances around the vessel. B.) Yield of biomass and CO_2 on glucose. C.) Fatty acid productivity and D.) fatty acid yield for C_{12} and C_{14} species. Experimental values are represented by open symbols, while the model prediction is represented by the dotted (C_{12}) or dashed (C_{14}) line. Error bars represent standard deviations of 3-5 samples taken over a residence time at each steady state.

At all dilution rates, the yield of FFA on glucose differed from yields of batch cultures grown in 50 mL shake flasks (0.04 g C_{12}/g glucose). The yield of product on substrate (Y_{ps}) increased with decreasing dilution rate, with the highest yield (0.071 g FFA/g glucose) observed at the lowest dilution rate (0.05 h^{-1}) tested. The productivity of FFA was maximized between dilution rates of 0.2 and 0.25 h^{-1} (Figure 3.2c) at a value of 0.023 g FFA/(g DCW*hr). Biomass (Y_{XS}) yield increased with dilution rate and all values fall within reported ranges for *E. coli* (Paalme et al. 1995; Pickett et al. 1979). Carbon dioxide yield decreased with increasing dilution rate. The carbon content of cell pellets was 47±1 wt.%. The compositions of other compounds were 7% hydrogen, 13.7% nitrogen, and 26.3% oxygen, giving an empirical formula for TY05 as $CH_{1.79}N_{0.25}O_{0.42}$, consistent with published formulas of *E. coli* (Pickett et al. 1979). Carbon balances on the system accounted for 91 to 98% of carbon inputs under all tested growth conditions. The breakdown of carbon flux can be seen in Figure 3.3. Under all dilution rates, CO_2 and biomass represented the largest fraction of carbon.

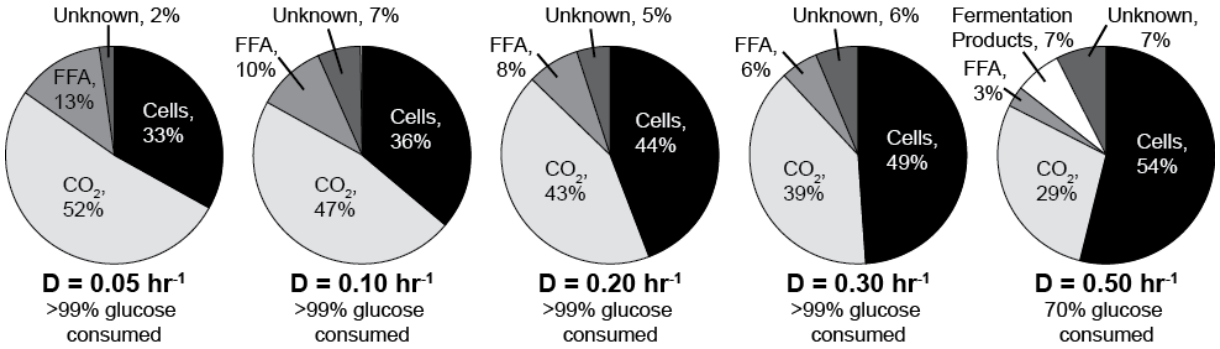


Figure 3.3: Graphical representation of the carbon mass balance. Values represent the fraction of carbon found in various products. Mass balances for dilution rates of 0.05, 0.1, 0.2, 0.3, and 0.5 h-1 are shown. Fermentation Products refers to the total concentrations of acetate, succinate, formate, and pyruvate observed in supernatant analyses.

A kinetic model was generated from the chemostat data collected for *E. coli* TY05. As seen in Figure 3.2c the FFA productivity, q_p, of *E. coli* TY05 was not linear with respect to dilution rate (equal to growth rate) and therefore five common empirical relationships were

examined (Heijnen 2009; Hensing et al. 1995). The fit via regression of the best relationship (equation 3), is shown in Figure 3.2c.

$$q_p = \frac{\alpha\mu}{\beta + \mu + \gamma\mu^2} \quad (3)$$

Relationships displaying a maximum in q_p at an intermediate dilution rate have been found for the production of anabolic compounds such as Penicillin-G (van Gulik et al., 2000), lysine (Kiss and Stephanopoulos, 1992), and other secondary metabolites (McIntyre et al., 1996). Here, the values of β were found to be 0.17 h^{-1} and 0.19 h^{-1} for the C_{12} and C_{14} species respectively, while the γ value was found to be 3.4 h for both the C_{12} and C_{14} relationship. Values for α were found to be 0.044 gC_{12}/gDCW/h for C_{12} and 0.013 gC_{14}/gDCW/h for C_{14}. To determine Y_{XS} and maintenance energy, each of the experimental q_p values were factored out of equation 2, leaving a modified q_s term as a linear function of growth. The resulting plot and linear fit can be seen in Figure 3.4. The yield of biomass on substrate (Y_{xs}) was determined to be 0.51 g DCW/g glucose which agrees with published values for *E. coli* (Paalme et al., 1995). The maintenance energy term, m_s, was determined to be 0.07 g glucose h^{-1}. This value is significantly greater than published values for *E. coli* (Heijnen and Roels, 1981), suggesting that reducing maintenance energy consumption could improve FFA yields. A model of the apparent FFA yield was calculated and plotted in Figure 3.2d.

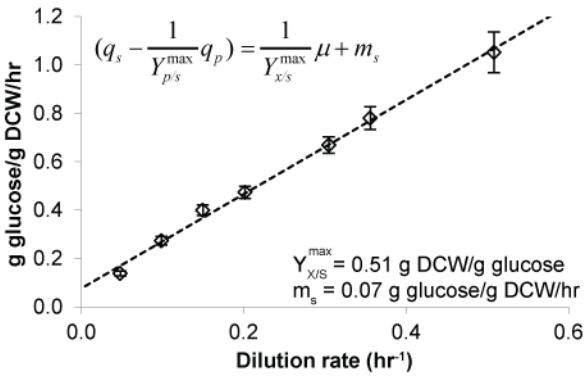


Figure 3.4: The maximum yield of biomass on glucose ($Y_{x/s}^{max}$, inverse slope) and maintenance energy (m_s, intercept) were calculated from a plot of modified glucose consumption (with product formation removed) versus dilution rate.

3.2.3. Variation of elemental feed composition

In a second set of chemostat experiments, the effect of feed nutrient composition on FFA production was tested. Chemostat cultures of TY05 were run under carbon limitation using glucose concentrations of 0.2, 0.4, and 1% (w/v) in MOPS minimal media containing excess nitrogen (7x N), phosphorous (7x P), sulfur (7x S), and magnesium (2x Mg^{2+}) relative to the standard MOPS recipe (Neidhard et al.). These recipes generate C/N mass ratios of 0.86, 1.72, and 3.44 respectively. A C/N mass ratio above 7.1 is required for a culture to become nitrogen limited (Neidhard et al.). For each C/N ratio, samples for biomass, glucose, FFA, and fermentation products were taken five times over a period of 16 h (1.5 residence times) after a steady state (as defined above) was established (Table 3.2). Residual glucose concentrations were below the 0.1 mM limit of detection (not shown). Biomass concentration and CO_2 output increased with increasing glucose feed concentration. Of the fermentation products assayed, none were observed under any C/N ratio. FFA titer increased with increasing C/N feed ratio. The data collected captured 88-98% of the carbon fed to the system (Table 3.2).

Table 3.2: Steady state concentrations of assayed metabolites from continuous culture of TY05 in modified MOPS/glucose media at a dilution rate of 0.1 h^{-1}

C/N ratio	Glucose **	Biomass	C12 FA	C14 FA	CO_2	Carbon recovery
g/g	g/L	g/L	mg/L	mg/L	ppm	%
0.85±0.01	2	0.62±0.03	60±5	16±2	1,360± 24	89±2
1.72±0.02	4	1.19±0.03	141±8	32±4	2,730±110	88±2
3.0±0.1	*4*	*1.21±0.07*	*172±5*	*41±5*	*2,900±200*	*93±4*
3.47±0.08	10	7.4±0.1	499±15	121±5	7,500±200	98±2

**Residual glucose concentrations were all below the limit of detection (1 mM)
*** Other fermentation products were not observed under any condition.
C12 FA is the combination of dodecanoic acid and dodecenoic acid
C14 FA is the combination of tetradecanoic acid and tetradecenoic acid
Note: row 3 (C/N 3.0±0.1) is copied from Table 3.1

The data presented in Table 3.2 was used to calculate productivities and yield for biomass, glucose, FFA, and fermentation products (Figure 3.5). Specific glucose uptake rates varied and specific CO_2 productivity, q_{CO2} increased slightly with increasing C/N ratio (Figure 3.5a). In contrast, the specific productivity of FFA increased to its maximum value at a C/N ratio of 3.5 (Figure 3.5c). Across all tested C/N ratios, the yield of biomass on glucose was consistent at a value of 0.31 ± 0.01 (g/g) (Figure 3.5b), the value observed at dilution rate of 0.1 h^{-1} in Figure 3.2, while the carbon dioxide yield was also consistent at 0.69 ± 0.03 (g/g). In contrast, the yield of FFA increased with increasing C/N ratios indicating changes in metabolic flux under different nutrient conditions (Figure 3.5d). Additionally the production of fermentation products was not observed under any fermentation conditions, indicating that excess carbon was not being directed to metabolic overflow production. It appears therefore, that FFA yield is impacted by carbon to nitrogen ratio, with regimes closer to nitrogen limitation allowing for increased yield of FFA on glucose. Similar results have been reported for secondary metabolites and polyketides such as the production of vancomycin (McIntyre et al., 1996).

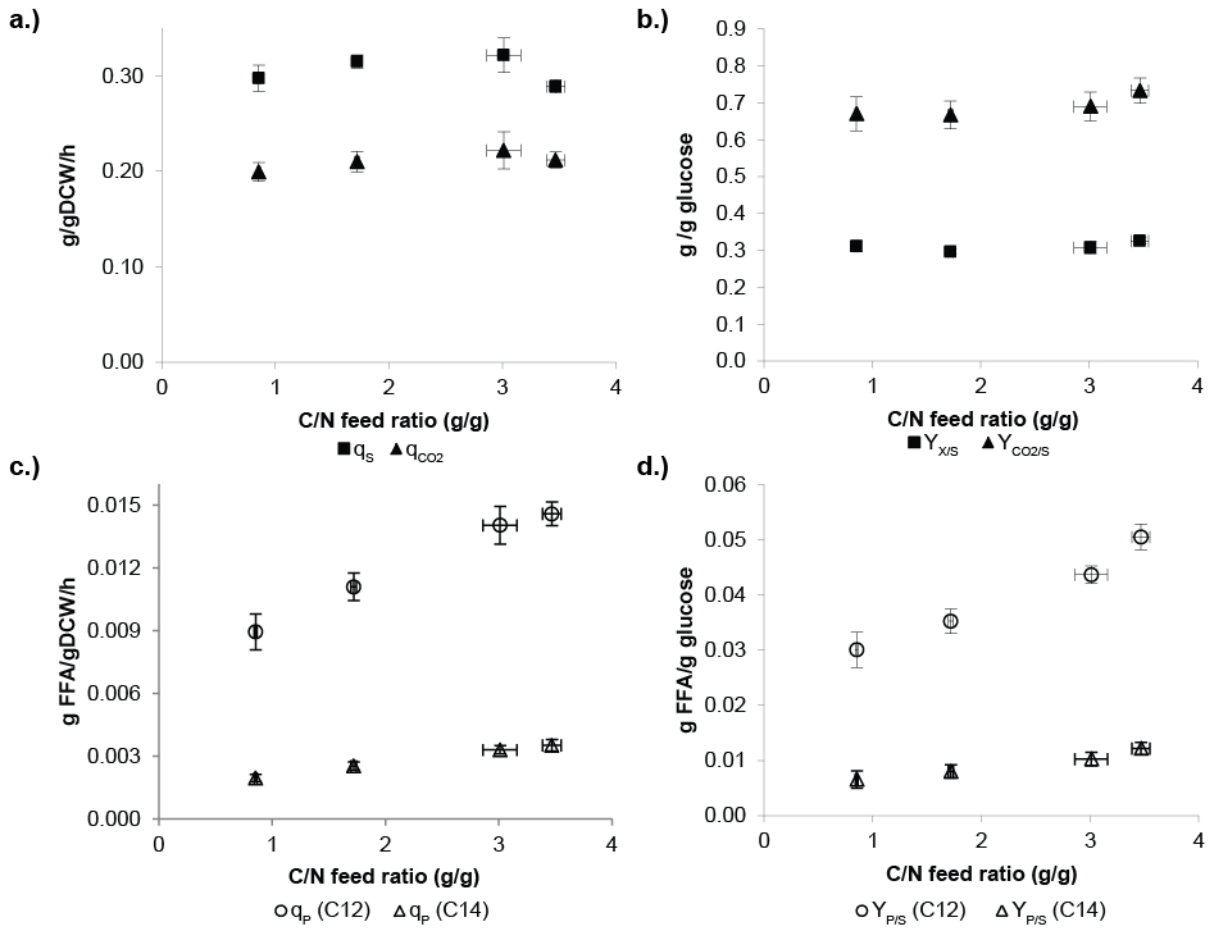


Figure 3.5: Productivity (A,C) and yield (B,D) values calculated for each C/N ratio tested. A.) Specific glucose uptake rate and CO_2 productivity calculated from mass balances around the vessel. B.) Yield of biomass and CO_2 on glucose. C.) Fatty acid productivity and D.) Fatty acid yield for C_{12} and C_{14} species. Error bars represent standard deviations of 5 samples taken over 16 hours for each steady state. The x-error bars represent error introduced from addition of base (10% NH_4OH).

3.3. Conclusions

A strain of *E. coli* harboring no antibiotic resistance markers was constructed and shown to produce C_{12} and C_{14} FFA species at levels comparable to prior studies that utilized plasmid based expression of a thioesterase. The strain was stably maintained over two weeks of continuous culture. A kinetic model was established, parameters estimated, and optimal continuous cultivation conditions predicted. Maximum yields were obtained when cells were

maintained at low growth rates while maximum productivities were obtained when cells were maintained at intermediate dilution rates. The optimal conditions generated a 37% increase in FFA yield over that observed from a batch culture using the same media.

The yield and productivity of FFA from *E. coli* TY05 were strongly dependent on the ratio of carbon to nitrogen in the media. Increasing the C/N feed ratio led to a further 15% increase in yield when TY05 was grown in continuous culture. Continuous culture studies of other secondary metabolites have shown significant production differences between different nutrient limitation conditions in the past (McIntyre et al., 1996), motivating the use of fed-batch and two-phase production strategies. In TY05, both yield and productivity increased as the C/N ratio approached non-carbon limitations, strongly motivating extension of the current model into non-carbon limitations. Understanding the impact of non-carbon limitation on metabolism in FFA producing *E. coli* will improve our modeling efforts and enable the design of fed-batch strategies for maximizing FFA yield. Linking flux balance models (Feist et al., 2007) to the data presented here could generate new hypotheses and enhance the current kinetic model. Similarly, incorporation of product inhibition and death terms may also improve the model's accuracy. Recent work has shown that overproduction of FFA in *E. coli* has a negative impact on cell physiology and viability (Lennen et al., 2011). These observations may explain the elevated maintenance energy requirements of *E. coli* TY05. These approaches and others will be the focus of future studies.

3.4. Materials and Methods

3.4.1. Chemicals, enzymes, and other materials

All enzymes were purchased from New England Biolabs (Ipswich, MA). DNA purification materials were purchased from Qiagen (Venlo, Netherlands). Chemicals were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Hampton, NH) unless otherwise specified. Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA).

3.4.2. Strain and plasmid construction

All strains and plasmids used in this study are listed in Appendix 3. A DNA cassette consisting of the chloramphenicol resistance marker flanked by flippase recombinase target (FRT) sites, constitutive promoter, *lacI*^{Q} gene, P_{TRC} promoter, and BTE gene was constructed (Figure 3.1) as a template for λ Red mediated recombination (Datsenko and Wanner, 2000). The template was constructed by amplifying the *cat* cassette flanked by FRT sites from pKD3 using primers (Appendix 4) that introduced SphI sites at the 5' and 3' ends of the PCR product, and cloning into the SphI site upstream of the *lacI*^{Q} region in plasmids pTrc99A-BTE and pTrc99A-BTE-H204A. Clones were selected that contained the *cat* cassette in the same orientation for plasmids containing both BTE and BTE-H204A. The integration cassette was amplified by PCR using primers (Appendix 4) containing 40 base-pairs flanking the target chromosomal sequence in 5' tails. In each case, the resulting DNA was purified by gel-extraction and transformed into *E. coli* DY330, a strain containing a chromosomal copy of the gene encoding λ Red recombinase (Yu et al., 2000) to generate recombinants. Mutant strains were selected on LB-agar supplemented with 34 μg/mL chloramphenicol. Each integration cassette was transferred from the corresponding DY330 integrants to recipient *E. coli* strains using P1 phage transduction

methods (Miller, 1972). The Cm^R resistance marker, located between two FRT sites on the integration cassette, was removed by transforming cells with pCP20 (Datsenko and Wanner, 2000). Strains harboring pCP20 express flip recombinase which catalyzes excision of DNA sequences located between two FRT sites. Subsequently, pCP20 was cured from each strain using methods described by Yazdani (Yazdani and Gonzalez, 2008). The process outlined above was used to iteratively generate strains harboring one (TY01, TY02; integrated into *fadD*), two (TY03, TY04 integrated into *fadD, fadE*), or three (TY05, TY06 integrated into *fadD, fadE, fadAB*) copies of either the BTE or BTE-H204A expression cassette. Confirmation of each integration was performed by colony PCR and sequencing (Figure 3.6).

3.4.3. Batch shake flask studies

FFA production from TY01, TY03, and TY05 was determined from cells grown in 250 mL shake flasks with 50 mL MOPS minimal medium (Neidhard et al., 1974), supplemented with 0.4% glucose. Single colonies from verified freezer stocks were inoculated into 5 mL of LB and incubated for eight hours at 37°C with shaking. The cultures grown in LB were back-diluted 1:100 into 5 mL of MOPS minimal with 0.4% glucose and incubated overnight at 37 °C with shaking. Production cultures were inoculated to an OD_{600} of 0.04 and incubated at 37°C until each reached and OD_{600} of 0.2, when IPTG (1 mM final concentration) was added to induce BTE. For comparison, *E. coli* RL08 was transformed with pTRC99A-BTE or pTRC99A-BTE-H204A and cultured as described above except for the addition of 50 µg/mL ampicillin and 50 µM IPTG, the optimal level determined in past work (Hoover et al., 2011). After 24 hours of incubation at 37 °C, 2.5 mL of cell culture was extracted with 5 mL of 1:1 chloroform:methanol to quantify fatty acids. Extracts were processed, derivatized to methyl-esters, and analyzed by

GC/MS (three technical replicates) as described previously (Lennen et al., 2010). An identical study of TY01-TY06 production in LB supplemented with 0.4% glucose was also performed.

3.4.4. Chemostat studies

Chemostat experiments were performed in a 3L stirred bioreactor (Applikon Biotechnology Inc, Schiedam, Netherlands) using a 1.5 L working volume. Vessel temperature was maintained at 37 °C using a cooling water jacket and an electric heat blanket (Applikon, model number M3414). Temperature, pH, and dissolved oxygen (DO_2) were monitored using specific probes (Applikon). The parameters were recorded every 3 minutes and controlled using provided BioXpert software (Applikon). The vessel pH was maintained at 7.00±0.05 by addition of 10% NH_4OH solution. Agitation was provided by a single impeller with the stirrer speed set to 500 rpm. Stirrer speed was occasionally increased to ensure the DO_2 content did not decrease below 40% saturation in order to maintain an aerobic environment (Becker et al., 1997; Tseng et al., 1996). Air inflow was maintained at 1.5 L/min using a Smart-Trak® digital mass flow controller (Sierra Instruments, Monterey, CA). Another Smart-Track® controller operated as a purge and recorded the gas outflow rate. Off-gas composition was continuously measured by a MGA iSCAN mass spectrometer (Hamilton Sundstrand, Windsor Locks, CT) relative to a house air reference. Liquid media inflow and waste outflow rates were controlled by displacement pumps (400 sci-Q, Watson-Marlow, Wilmington, MA) which could vary flow rates between 0.04 and 0.72 kg/h. Foam overflow from the reactor was collected in sterile 1L bottles, that were changed prior to each sampling point. The masses of the sterile media feed reservoir, the bioreactor, foam overflow, and waste reservoirs were continuously monitored using digital balances (model number EA35EDE-I, Sartorius AG, Goettingen, Germany) which sent outputs to LabVIEW (National Instruments, Austin, TX) every minute.

Chemostat experiments were inoculated using the following propagation schedule. Five mL cultures of TY05 grown in LB at 37ºC for 10 h from single colonies were used to inoculate 50 mL of MOPS minimal media supplemented with 0.4% glucose. The 50 mL culture was grown at 37ºC overnight and was used as the inoculum for the chemostat. For the dilution rate experiments, bioreactors containing MOPS minimal media supplemented with 0.4% glucose were inoculated to an OD_{600} of 0.04. Induction with 1mM IPTG occurred when the OD_{600} reached 0.2. The reactor was operated in batch mode for 9 hours, at which point flow of fresh media was initiated. Outflow was controlled by maintaining a constant liquid level through the placement of the outlet line at a height corresponding to a 1.5 L working volume. The OD_{600} of the culture was monitored periodically (18 samples per 24 h) during batch and continuous operation. For all tested conditions steady-state was assumed after a full residence time of constant OD_{600}, CO_2 output, and dO_2 values were observed (Diderich et al., 1999). Once a steady-state was achieved, culture samples were collected through an outlet port for fatty acids, sugars, and organic acids using a sterile syringe. For the dilution rate experiments, two separate cultivations using the same initial strain were performed over two-week time periods. The first chemostat run tested step changes of dilution rates starting at 0.05 h^{-1} and increasing in 0.05 h^{-1} increments up to 0.5 h^{-1} (~80% μ_{max}). The second chemostat tested dilution rates of 0.05, 0.1, 0.15, 0.2, 0.3 h^{-1} in a random order to ensure that the cultures did not experience any hysteresis. The order of steady states was 0.1, 0.3, 0.15, 0.05, 0.2, 0.1, 0.2, 0.05, 0.3, 0.15 h^{-1}. For the C/N mass ratio experiments, the starter cultures followed the same propagation scheme as the dilution rate experiments. Once all samples were taken for a given C/N ratio, sterile media feed was changed to MOPS minimal media supplemented with a different glucose concentration. The

order of the C/N mass ratios was 1.72, 0.85, and 3.47, corresponding to initial glucose concentrations of 4, 2, and 10 g/L.

3.4.5. Cell density and elemental composition

A linear correlation between OD_{600} and dry cell weight (DCW) was determined using 30 mL of culture taken for each steady state condition. Washed cell pellets were lyophilized (Labconco Corp, Kansas City, MO) until constant mass was achieved. After each chemostat run, the remaining biomass was dried as above and sent to Elemental Analysis Inc (Lexington, KY) to determine weight percent of carbon, hydrogen, oxygen, and nitrogen in the biomass.

3.4.6. Sugar and fermentation product quantification

Four supernatant sugar samples of culture were harvested (four times per steady state) out of the reactor into an ice-cold 15 mL conical tube. The resulting sample was immediately centrifuged for 5 min at 8000 x g and 4°C and passed through a 0.2 µm filter (Millipore, Billerica, MA) into individual microcentrifuge tubes for storage at -80˚C prior to analysis. Analysis was performed using an Agilent 1260 infinity HPLC system with a quaternary pump, chilled (4°C) autosampler, vacuum degasser, and refractive index detector (Agilent Technologies, Inc., Palo Alto, CA). An Aminex HPX-87H with Cation-H guard column (Biorad, Inv. Hercules C) size 300 x 7.8 mm was used. The isocratic separation used a mobile phase of 0.02 N H_2SO_4 with a flow rate of 0.5 mL/min to separate a 50 µL injection volume over 28 min. Instrument control, data collection, and analysis was performed using Chem Station version B04.03 software (Agilent). Standards containing glucose, xylose, pyruvate, xylitol, succinate, lactate, glycerol, formate, acetate, and ethanol were used to quantify each compound.

3.4.7. Fatty acid samples

At five times per steady state, 2.5 mL culture samples from the bioreactor and foam overflow were taken in technical replicate and stored in 1:1 chloroform:methanol at -80˚C until derivitization to methyl esters and analysis by GC/MS as described above (Lennen et al., 2010). The overall FFA titer was calculated by factoring in both values from the foam overflow and reactor vessel. Foam overflow was collected in sterile 1 L bottles and the mass was recorded in order to determine a steady state rate of foam production. Based on analysis of FFA in the extracellular media and cell pellet, $C_{12:0}$, $C_{12:1}$, $C_{14:0}$, and $C_{14:1}$ species were counted as FFA.

3.4.8. Kinetic model

A mass balance around the reactor, given by equation (1) was used to calculate specific productivity rates for substance i (biomass [X], glucose [S], or product [P]).

$$\frac{d(VC_i)}{dt} = 0 = C_{i_0}F_{in} - C_iF_{out} + q_iC_xV \quad (1)$$

In equation (1), F is the volumetric flow rate into or out of the bioreactor, C_i is the specific concentration of substance i and q_i is the productivity of substance i. At steady state, $d(VC_i)/dt$ is equal to zero and the specific productivity is directly calculated from equation (1). As the concentration of cells in the feed (C_{x0}) is zero, the growth rate (μ) is equal to outflow rate divided by reactor volume, which is termed the dilution rate (D). The specific glucose uptake rate (q_S) was defined as the amount of glucose consumed per hour per mass of dry biomass. FFA productivity (q_P) was defined as the amount of FFA produced per hour per mass of dry biomass. From the calculated q rates, the yield coefficients (Y_{is}) were calculated by dividing a productivity of interest (q_i, e.g. FFA or biomass) by q_s. In the kinetic model, the specific productivities were related by the Herbert-Pirt equation (2) (Herbert et al., 1956).

$$q_s = \frac{1}{Y_{xs}^{\max}}\mu + \frac{1}{Y_{ps}^{\max}}q_p + m_s \quad (2)$$

Parameters for the kinetic model were determined from data by least-squares regression.

3.4.9. Confirmation of strain construction

Each strain was confirmed by colony PCR using primers listed in Appendix 4. Specifically, screening for TY01 and TY02 used the *fadD* knockout colony pcr primers, screening for TY03 and TY04 used both the *fadD* and *fadE* knockout colony pcr primers, and screening for TY05 and TY06 used the *fadD*, *fadE* and *fadAB* knockout primers to make sure that each specific integration was present. Colony PCR was performed after both the initial integration and after the removal of the chloramphenicol resistance marker by flip recombinase. Additional screening was performed after the knockout step using the internal ptrcseq primer (Appendix 4) and the corresponding reverse knockout primer. Controls used previously made integration strains (e.g. using TY03 as positive control for *fadD* and *fadE* integrations in screening TY05) as well as negative controls using *E. coli* MG1655. An example of both screens in the identification of TY05 (model strain) can be seen in Figure 3.6a. After screening the corresponding cleaned up strains using the outer primers, the colony PCR product was purified and sequenced using the internal sequencing primers Ptrc_seq, BTE_seqR, and $lacI^Q$_seqR (Appendix 4). The expected sequences were found in each strain. A set of multiplex primers were also designed (BTEseqF, TY05fadABR, TY05fadDR, fadEKO_colPCR_fwd; Appendix 4) to better facilitate identification of the three integration sites. Screening using the multiplex primers yielded DNA band sizes of 300 (insertion into *fadE* locus), 700 (insertion into *fadAB* locus), and 1200 (insertion into *fadD* locus), and results can be seen in Figure 3.6b.

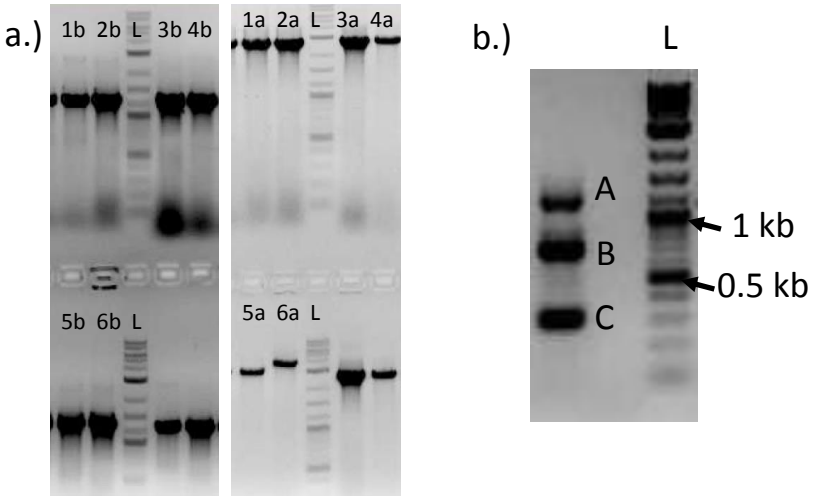


Figure 3.6: a.) Colony pcr results for the three areas of integration in TY05. Results for both the entire (a, 2.8kb) and internal-external primer (b, 1.3kb) colony pcr results are shown. 1a and b represent the *fadD* region, 4a and b represent the *fadE* region, and 5a and 5b represent the *fadAB* region. Respective controls are represented by 2a and 2b for *fadD*, 3a and 3b for *fadE*, and 6a and 6b for *fadAB*. All represent positive controls except for 6a which is from a control strain that has the original *fadAB* region intact. b.) Multiplex primer run for TY05, bands for *fadE* insertion (300 bases), *fadAB* insertion (700 bases), and *fadD* (1100 bases) insertion can be seen. The L column represents a 2-log DNA ladder.

3.4.10. Analysis of free fatty acids

To gather an accurate depiction of fatty acid titer within the foam overflow, a new foam overflow bottle was added at the minimum amount of time before a sampling (determined by the rate, which varied based on dilution rate or C/N ratio) in order to accumulate at least 5 mL of foam overflow. Effects of time on titer in the foam overflow bottle were also investigated by allowing more foam to accumulate in the bottle after the sampling for the timepoint, then taking another fatty acid sample after ~4 hours. In all cases such titer fell within the standard deviation of the titer seen previously. The percent of total flow going to the foam overflow and percent going to waste from reactor were used to weight the titers obtained from the foam overflow bottle and reactor vessel, respectively.

During the first continuous culture run, the percent C_{12} FFA in the supernatant versus in the cell pellet was determined. 10 mL of culture was taken and centrifuged at for 5 minutes at 6000 x g and 4˚C, at which point two technical replicates of 2.5 mL of supernatant were taken for derivatization and analysis by GC/MS. The remaining supernatant was removed and the cell pellet was resuspended in 10 mL of 1X PBS. After two subsequent centrifugation, supernatant removal, and resuspension steps, technical replicates of 2.5 mL of the resulting suspension were taken for derivatization and analysis by GC/MS to determine the fraction of dodecanoic acid in the cell pellet. Analysis of the fraction of each fatty acid species in the supernatant versus cell pellet was performed for each growth condition. The fraction of each species in the supernatant can be seen in Figure 3.7b while the corresponding composition of the fatty acid species in the pellet is shown in Figure 3.7a. 91% of the C_{12} acid titer present in the reactor was found in the supernatant, while 69% of the C_{14} acid titer was found in the supernatant. In contrast, more than 85% of each of the C_{16} and C_{18} fatty acid species were found in the cell pellet. To find get a more conservative assumption of free C_{12} FFA in the foam, the cellular fatty acid composition seen in Figure 3.7a was applied using the C_{16} and C_{18} levels found in the foam overflow. The resulting calculation yielded levels of free dodecanoic acid between 99.7% and 99.9% of their original values. When applying that the contribution of dodecanoic acid in the foam was at least 70% of the total titer in all cases, the levels of free dodecanoic acid would always be above 97% of the total measured titer. A similar relationship exists for tetradecanoic acid, with the foam contributing ~90% of tetradecanoic acid levels, with 99.2-99.4% of that seen in the foam and at least 95% of the overall titer existing as FFAs. Therefore, only minimal amounts of C_{12} and C_{14} FFA remains intracellularly and not in a free form.

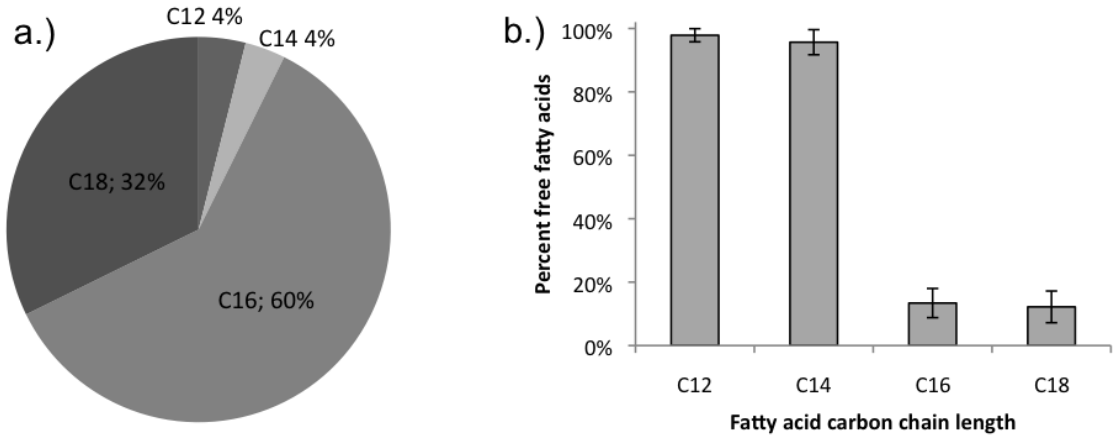


Figure 3.7: a.) Percent composition of fatty acid species within the cell pellet for C_{12}, C_{14}, C_{16}, and C_{18} fatty acid species and b.) percentage of each fatty acid species that was found in the supernatant.

3.4.11. Analysis of TY01-TY06 production in rich media

FFA production from TY01-TY06 was determined from cells grown in 250 mL shake flasks with 50 mL LB, supplemented with 0.4% glucose. Single colonies from verified freezer stocks were inoculated into 5 mL of LB and incubated overnight at 37°C with shaking. Production cultures were inoculated to an OD_{600} of 0.02 and incubated at 37°C until each reached and OD_{600} of 0.2, at which point it was induced with 1 mM IPTG. For comparison, *E. coli* RL08 was transformed with pTRC99A-BTE or pTRC99A-BTE-H204A and cultured as described above except for the addition of 50 µg/mL ampicillin and 50 µM IPTG, the optimal level determined in past work (Hoover et al., 2011). After 24 hours of incubation at 37 °C, 2.5 mL of cell culture was extracted with acidified methanol to quantify fatty acids. Extracts were processed, derivatized to methyl-esters, and analyzed by GC/MS (three technical replicates) as described previously (Lennen et al., 2010). Results can be seen in Figure 3.8.

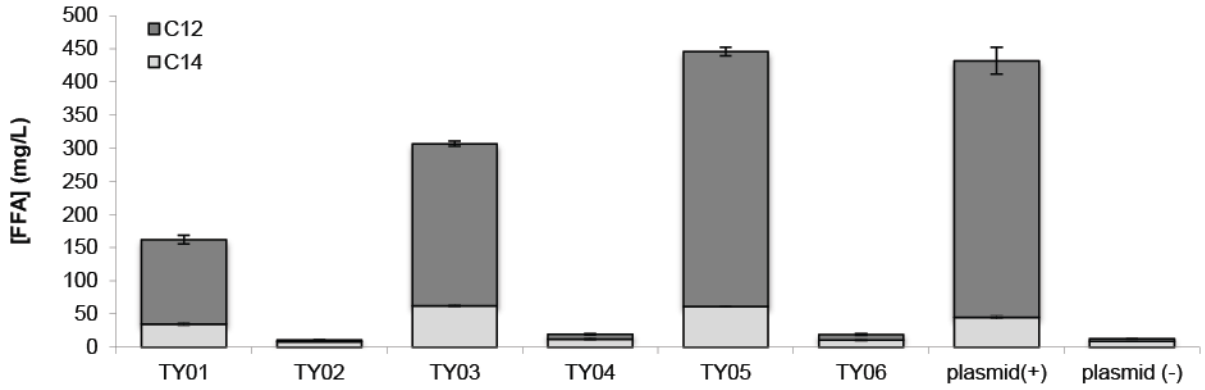


Figure 3.8: Production of C_{12} and C_{14} FFAs from each successive P_{TRC}-BTE (and P_{TRC}-BTE-H204A) integration compared to BTE expressed on a plasmid. Plasmid(+) represents pTRC99A-BTE and plasmid(-) represents pTRC99A-BTE-H204A. Production levels 24 hours after induction with IPTG in LB + 0.4% glucose.

3.5. References

Amann E, Ochs B, Abel KJ. 1988. Tightly regulated *tac* promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**(2):301-315.

Becker S, Vlad D, Schuster S, Pfeiffer P, Unden G. 1997. Regulatory O-2 tensions for the synthesis of fermentation products in *Escherichia coli* and relation to aerobic respiration. *Archives of Microbiology* **168**(4):290-296.

Beller HR, Goh E-B, Keasling JD. 2010. Genes Involved in Long-Chain Alkene Biosynthesis in *Micrococcus luteus*. *Applied and Environmental Microbiology* **76**(4):1212-1223.

Brosius J, Erfle M, Storella J. 1985. Spacing of the -10 and -35 regions in the tac promoter - effect on its in vivo activity. *Journal of Biological Chemistry* **260**(6):3539-3541.

Cherepanov PP, Wackernagel W. 1995. Gene disruption in *Escherichia-coli* - tcr and km(r) cassettes with the option of FLP-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**(1):9-14.

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

Diderich JA, Schepper M, van Hoek P, Luttik MAH, van Dijken JP, Pronk JT, Klaassen P, Boelens HFM, de Mattos RJT, van Dam K and others. 1999. Glucose uptake kinetics and transcription of HXT genes chemostat cultures of Saccharomyces cerevisiae. *Journal of Biological Chemistry* **274**(22):15350-15359.

Du J, Shao ZY, Zhao HM. 2011. Engineering microbial factories for synthesis of value-added products. *Journal of Industrial Microbiology & Biotechnology* **38**(8):873-890.

Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BO. 2007. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology* **3**.

Fru EC, Ofiteru ID, Lavric V, Graham DW. 2011. Non-linear population dynamics in chemostats associated with live-dead cell cycling in *Escherichia coli* strain K12-MG1655. *Applied Microbiology and Biotechnology* **89**(3):791-798.

Fujita Y, Matsuoka, H, Hirooka, K. 2007. Regulation of fatty acid metabolism in bacteria. *Biotechnology and Bioenginering* **66**(4):829-839.

Goochee CF, Hatch RT, Cadman TW. 1985. The response of a nitrogen-limited chemostat culture of *Escherichia coli* B/r to pH and dilution rate steps. *Biotechnology and Bioengineering* **27**(4):439-446.

Gramajo HC, Takano E, Bibb MJ. 1993. Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. *Molecular Microbiology* **7**(6):837-845.

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

Heijnen JJ. 2009. Black Box Models for Growth and Product Formation. In: Smolke CD, editor. Handbook for Metabolic Pathway Engineering: Fundamentals. San Diego, CA: CRC Press.

Heijnen JJ, Roels JA. 1981. A macroscopic model describing yield and maintenance relationships in aerobic fermentation processes. *Biotechnology and Bioengineering* **23**(4):739-763.

Hensing MCM, Vrouwenvelder JS, Hellinga C, Vandijken JP, Pronk JT. 1995. Use of chemostat data for modeling extracellular inulinase production by *Kluyveromyces marxianus* in a high-cell-density fed-batch process. *Journal of Fermentation and Bioengineering* **79**(1):54-58.

Herbert D, Elsworth R, Telling RC. 1956. The continuous culture of bacteria; a theoretical and experimental study. *Journal of general microbiology* **14**(3):601-22.

Hoover SW, Marner WD, Brownson AK, Lennen RM, Wittkopp TM, Yoshitani J, Zulkifly S, Graham LE, Chaston SD, McMahon KD and others. 2011. Bacterial production of free fatty acids from freshwater macroalgal cellulose. *Applied Microbiology and Biotechnology* **91**(2):435-446.

Hoover SW, Youngquist JT, Angart PA, Withers ST, Lennen RM, Pfleger BF. 2012. Isolation of improved free fatty acid overproducing strains of *Escherichia coli* via nile red based high-throughput screening. *Environmental Progress and Sustainable Energy* **31**:17–23.

Kalscheuer R, Stolting T, Steinbuchel A. 2006. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology* **152**:2529-2536.

Khanna S, Srivastava AK. 2005. Statistical media optimization studies for growth and PHB production by *Ralstonia eutropha*. *Process Biochemistry* **40**(6):2173-2182.

Kiss RD, Stephanopoulos G. 1992. Metabolic characterization of a L-lysine-producing strain by continuous culture. *Biotechnology and Bioengineering* **39**(5):565-574.

Lennen RM, Braden DJ, West RM, Dumesic JA, Pfleger BF. 2010. A Process for Microbial Hydrocarbon Synthesis: Overproduction of Fatty Acids in *Escherichia coli* and Catalytic Conversion to Alkanes. *Biotechnology and Bioengineering* **106**(2):193-202.

Lennen RM, Kruziki MA, Kumar K, Zinkel RA, Burnum KE, Lipton MS, Hoover SW, Ranatunga DR, Wittkopp TM, Marner WD, 2nd and others. 2011. Membrane stresses induced by overproduction of free fatty acids in *Escherichia coli*. *Applied and Environmental Microbiology* **77**:8114–28.

Lennen, RM., Pfleger, BF., 2012. Engineering *Escherichia coli* to synthesize free fatty acids. *Trends in Biotechnology* **30**, 659-67.

Lennen, RM. Pfleger, BF., 2013. Microbial production of fatty acid-derived fuels and chemicals. *Current Opinion in Biotechnology* **21**:1-10.

McIntyre JJ, Bull AT, Bunch AW. 1996. Vancomycin production in batch and continuous culture. *Biotechnology and Bioengineering* **49**(4):412-420.

Mendez-Perez D, Begemann MB, Pfleger BF. 2011. Modular synthase-encoding gene involved in alpha-olefin biosynthesis in *Synechococcus sp. strain PCC 7002*. *Applied and Environmental Microbiology* **77**(12):4264-7.

Miller JH. 1972. Experiments in Molecular Genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Neidhard FC, Bloch PL, Smith DF. 1974. Culture medium for enterobacteria. *Journal of Bacteriology* **119**(3):736-747.

Nielsen J, Villadsen J. 1992. Modeling of microbial kinetics. *Chemical Engineering Science* **47**(17-18):4225-4270.

Orencio-Trejo M, Utrilla J, Fernandez-Sandoval MT, Huerta-Beristain G, Gosset G, Martinez A. 2010. Engineering the *Escherichia coli* Fermentative Metabolism. *Advances in Biochemical Engineering/Biotechnology* **121**:71-107.

Paalme T, Kahru A, Elken R, Vanatalu K, Tiisma K, Vilu R. 1995. The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat). *Journal of Microbiological Methods* **24**(2):145-153.

Pickett AM, Bazin MJ, Topiwala HH. 1979. Growth and composition of *Escherichia coli* subjected to square-wave perturbations in nutrient supply - effect of varying frequencies. *Biotechnology and Bioengineering* **21**(6):1043-1055.

Rude MA, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A. 2011. Terminal Olefin (1-Alkene) Biosynthesis by a Novel P450 Fatty Acid Decarboxylase from Jeotgalicoccus Species. *Applied and Environmental Microbiology* **77**(5):1718-1727.

Schirmer A, Rude MA, Li XZ, Popova E, del Cardayre SB. 2010. Microbial Biosynthesis of Alkanes. *Science* **329**(5991):559-562.

Steen EJ, Kang YS, Bokinsky G, Hu ZH, Schirmer A, McClure A, del Cardayre SB, Keasling JD. 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**(7280):559-U182.

Todorov SD, van Reenen CA, Dicks LMT. 2004. Optimization of bacteriocin production by *Lactobacillus plantarum* ST13BR, a strain isolated from barley beer. *Journal of General and Applied Microbiology* **50**(3):149-157.

Tseng CP, Albrecht J, Gunsalus RP. 1996. Effect of microaerophilic cell growth conditions on expression of the aerobic (*cyoABCDE* and *cydAB*) and anaerobic (*narGHJI*, *frdABCD*, and *dmsABC*) respiratory pathway genes in *Escherichia coli*. Journal of Bacteriology 178(4):1094-1098.

Valgepea K, Adamberg K, Vilu R. 2011. Decrease of energy spilling in *Escherichia coli* continuous cultures with rising specific growth rate and carbon wasting. *BMC Systems Biology* **5**.

van Gulik WM, de Laat W, Vinke JL, Heijnen JJ. 2000. Application of metabolic flux analysis for the identification of metabolic bottlenecks in the biosynthesis of penicillin-G. *Biotechnology and Bioengineering* **68**(6):602-618.

Yazdani SS, Gonzalez R. 2008. Engineering *Escherichia coli* for the efficient conversion of glycerol to ethanol and co-products. *Metabolic Engineering* **10**(6):340-351.

Yu DG, Ellis HM, Lee EC, Jenkins NA, Copeland NG, Court DL. 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **97**(11):5978-5983.

Zhu X, Zhang W, Chen X, Wu H, Duan Y, Xu Z. 2010. Generation of High Rapamycin Producing Strain via Rational Metabolic Pathway-Based Mutagenesis and Further Titer Improvement With Fed-Batch Bioprocess Optimization. *Biotechnology and Bioengineering* **107**(3):506-515.

Chapter 4: Free fatty acid production in *Escherichia coli* under phosphate limited conditions[2]

4.1. Introduction

Chapter 3 described a plasmid-free, fatty acid producing strain of *E. coli*, and efforts to use classical biochemical engineering to model its growth and productivity. A central challenge to producing free fatty acid (FFA) derived chemicals is optimizing the flux of carbon through fatty acid metabolism to optimize product yields. After analyzing the kinetic model of FFA production, we hypothesized that moving media conditions away from carbon limitation could provide a significant advantage to FFA yields. FFA productivity and yields all increased as the carbon to nitrogen (C:N) or carbon to phosphorous (C:P) ratio in the media decreased. Chapter 4 describes efforts to extend the model to fermentation conditions where carbon is no longer the limiting nutrient and experiments that validate strain performance predicted by the model. Additionally, Chapter 4 explores the effect of different fermentation parameters and their effect on FFA yield and productivity.

Many groups have applied metabolic engineering strategies to improve the titer and yield of FFAs in microorganisms such as *E. coli* (reviewed in Lennen and Pfleger, 2012). Basic strategies that utilize a chain length specific thioesterase to cleave intermediates in fatty acid biosynthesis have produced yields around 0.06 g fatty acid / g consumed carbon source (Steen et al., 2010; Lu et al., 2008; Zheng et al., 2012). Supplemental metabolic engineering strategies targeting central metabolism have been used to increase fatty acid yield to 0.14 g fatty acid/g carbon source consumed (Ranganathan et al. 2012). The highest reported yield of fatty acid, 73% of the theoretical limit for *E. coli*, was achieved by overexpressing FadR, a regulator of fatty acid metabolism, in a thioesterase expressing strain (Zhang et al. 2012). While these yields are

[2] Portions of this chapter were published in Youngquist et al., *Applied Microbiology and Biotechnology*, 2013.

promising, all used carbon as the limiting nutrient in their media formulation. No study has examined nutrient limitation as a means to approach the theoretical limit of FFA yield in *E. coli*.

The yield of FFA on glucose is maximized under conditions where the flux of carbon to biomass is minimized. Therefore, advanced cultivation strategies such as fed batch are needed to achieve the optimal balance of cell growth and product synthesis. In past work, a kinetic model of FFA production in *E. coli* under carbon limitation predicted increased yield and biomass-specific productivity as media formulations moved toward non-carbon limitation (Chapter 3). This prediction is consistent with efforts to use nitrogen starvation as a trigger of lipid accumulation in oleaginous microbes (Ratledge and Wynn, 2002; Hassan et al., 1996). The effects of different essential elements limiting cellular *E. coli* growth has been well-studied with significant research on growth effects (Wanner and Egli, 1990), differences in stationary phase (Peterson et al., 2005), and transcriptional effects (Ferenci, 2008; Hua et al., 2004; Matsuoka and Shimizu 2011). Cultures of *E. coli* grown under nitrogen limiting conditions exhibit decreased yield of biomass on glucose (Hua et al., 2004; Kumar and Shimizu, 2011), and increased yield of desired products, such as poly-hydroxybutyrate (Johnson et al., 2010). *E. coli* cultivated under phosphate limiting conditions also exhibit a decreased biomass yield on various carbon sources (Johansson et al., 2005; Marzan and Shimizu, 2011). In contrast to nitrogen limited cultures, the metabolic activity of these cells remains high in stationary phase, with a low level of protein oxidation (Ballesteros et al., 2001; Gerard et al., 1999). Metabolically active but non-growing cells are attractive for the production of chemicals as no initial substrate is directed to biomass formation and substrate uptake rates remains high. Thus, phosphate limitation has been used in *E. coli* to increase the production of shikimic acid (Johansson et al., 2005) and recombinant proteins (Huber et al., 2011). Additionally, use of phosphate limitation has led to increased

production of vancomycin in *Amycolatopsis orientalis* (McIntyre et al., 1996) and green fluorescent protein in *Hansenula polymorpha* (Kottmeier et al., 2010). However, phosphate limitation has not been specifically studied in the production of FFAs from *E. coli*. For these reasons, phosphate limitation is an attractive strategy for controlling *E. coli* growth while maintaining a high rate of glucose utilization and FFA synthesis.

Here, we studied the impact of cultivation variables (growth rate, nutrient limitation) on FFA production and biomass accumulation in a FFA producing *E. coli*. In shake flask experiments, cells cultivated in phosphate limited media generated higher FFA titers and yields. Next, a series of chemostat experiments were performed to quantitatively model the impact of cultivation variables. The concentrations of FFA, biomass, CO_2, residual glucose, and fermentation products in chemostat effluents were measured in order to close a carbon mass balance around the bioreactor. The data collected was used to update the kinetic model from Chapter 3 to account for FFA production as a function of nutrient limitation and dilution rate and find conditions that maximize FFA yield and biomass specific productivity. The model was used to accurately predict the biomass and FFA yield in batch cultivation.

4.2. Results

4.2.1. Initial evaluation of non-carbon limiting conditions

In Chapter 3, a kinetic model was developed to predict the performance of a FFA producing strain of *E. coli* (TY05) under carbon limitation. This study concluded that FFA yields were highest at low growth rates and suggested that moving away from carbon limiting conditions would lead to greater FFA yield and specific productivity. To initially determine whether to pursue nitrogen or phosphate as the non-carbon, limiting nutrient, *E. coli* TY05 was

cultivated in batch, shake flask experiments using MOPS minimal medium with glucose as the sole carbon source, ammonium chloride as the sole nitrogen source, and potassium phosphate as the sole phosphate source. Both nitrogen and phosphate limitation led to higher FFA yields on glucose than the same cells cultivated under carbon limitation (Figure 4.1). Phosphate limitation resulted in the highest yield. Additionally, the phosphate limited cultures consumed significantly more glucose than the nitrogen limited cultures, even though their final OD_{600} values were equivalent. This result is consistent with the observation that glucose uptake significantly slows after nitrogen starvation (Ballesteros et al., 2001).

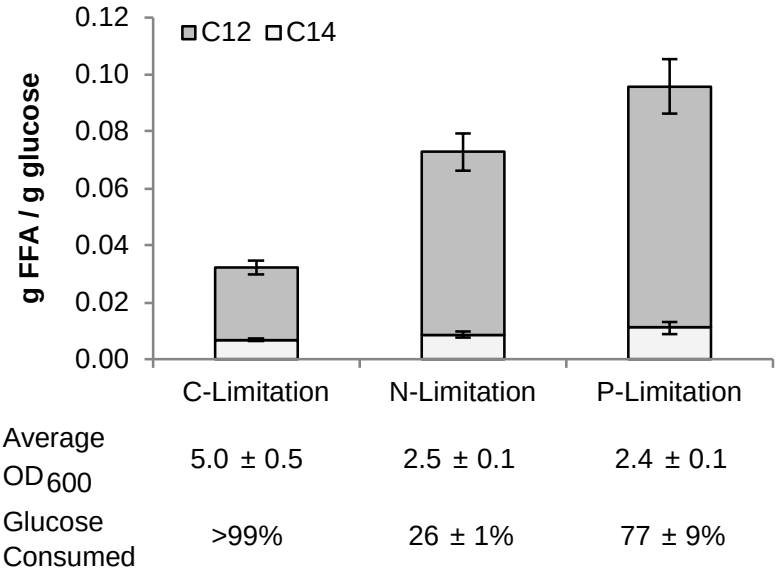


Figure 4.1: Yield of FFA on glucose from *E. coli* TY05 grown in batch cultures under various nutrient limitation conditions (carbon-, nitrogen-, or phosphorous-limiting).

4.2.2. Continuous cultivation of a FFA producing E. coli at various P-limited growth rates

To expand our model of FFA production in *E. coli* TY05, data was collected from chemostat experiments under phosphorus limiting conditions. A culture of *E. coli* TY05 was fed MOPS minimal glucose medium containing 1 mM IPTG for induction of BTE. The medium was phosphorus limited as it contained 0.37 mM phosphate instead of the typical 1.32 mM. During

the first phosphorus limitation growth rate experiment, *E. coli* TY05 was continuously cultured using different dilution rates for a feed containing 10 g/L (1.0%) glucose. Under each regime, the biomass concentration (calculated from OD_{600}), residual glucose, FFA titer, CO_2 evolution rate (calculated from off-gas composition and air flow rate), and the titer of alternative fermentation products (acetate, formate, ethanol, succinate, pyruvate, and lactate) were measured **(**Table 4.1**)**. Similar to carbon limitation, residual glucose levels increased and FFA titer decreased with increasing dilution rate. Of the alternative fermentation products assayed, only acetate was present in a measurable quantity, with concentrations decreasing with increasing dilution rate. In contrast to carbon limited experiments, extracellular acetate appeared under all tested dilution rates. Using mass balances for biomass, glucose, and FFA, the production rates (q-rates) for each were calculated **(**Figure 4.2 for FFA**)**. These rates were used to calculate FFA and biomass yields as a function of dilution rate. Additional data for *E. coli* TY06 was taken for comparison at dilution rates of 0.05, 0.1, 0.2, and 0.3 h^{-1} and results can be seen in Table 4.2.

Unlike carbon-limited conditions, significant residual glucose was found in the bioreactor effluent for each of the phosphate-limited dilution rates. Glucose consumption rates were equal for the lowest three dilutions rates in contrast to the monotonically increasing trend observed for carbon limitation. Above a dilution rate of 0.1 h^{-1}, glucose uptake rates increased with a slope similar that that observed for carbon limitation. Carbon dioxide specific productivity was static at low (< 0.1 h^{-1}) and high (> 0.3 h^{-1}) dilution rates. Biomass yields under phosphorus limitation (Figure 4.2c) increased with increasing dilution rate and followed a similar trend to the yield under carbon limitation except with values around one third less for high dilution rates. Larger differences were observed for low dilution rates. The yield of carbon dioxide from glucose

decreased with increasing dilution rate under both carbon and phosphorus limiting conditions, following nearly the same curve.

Table 4.1: Steady state concentrations of assayed metabolites from continuous culture of TY05 in MOPS minimal media with 1.0% glucose and 370 µM phosphate at various dilution rates

Dilution rate h^{-1}	Biomass g/L	Glucose g/L	C_{12} FA mg/L	C_{14} FA mg/L	CO_2 g/h	Acetate* g/L	Carbon recovery %
0.050±0.001	0.62±0.04	2.8±0.2	510±38	74±3	0.38±0.03	0.26±0.02	80
0.075±0.001	0.69±0.03	4.5±0.1	430±37	61±5	0.42±0.01	0.16±0.01	84
0.100±0.001	0.84±0.05	5.3±0.1	403±32	58±3	0.50±0.02	0.13±0.01	94
0.151±0.001	0.69±0.03	6.7±0.1	242±13	43±3	0.45±0.01	0.13±0.01	92
0.202±0.001	0.68±0.02	6.9±0.6	186±13	35±1	0.60±0.02	0.11±0.01	93
0.256±0.001	0.59±0.03	7.8±0.1	130±5	26±1	0.48±0.01	0.09±0.01	91
0.304±0.002	0.54±0.01	8.0±0.4	98±6	20±1	0.52±0.05	0.11±0.01	94
0.411±0.001	0.41±0.02	8.72±0.2	52±4	11±2	0.39±0.01	0.063±0.002	93

* Other fermentation products were not observed under any condition.
C_{12} FA is the combination of dodecanoic acid and dodecenoic acid
C_{14} FA is the combination of tetradecanoic acid and tetradecenoic acid

Table 4.2: Steady state concentrations of assayed metabolites from continuous culture of TY06 in MOPS minimal media with 1.0% glucose and 370 µM phosphate at various dilution rates

Dilution rate h^{-1}	Biomass g/L	Glucose g/L	C_{12} FA mg/L	C_{14} FA mg/L	CO_2 g/h	Acetate* g/L	Carbon recovery %
0.050±0.001	0.84±0.04	4.9±0.1	ND	ND	0.36±0.02	0.53±0.01	99
0.100±0.001	0.85±0.04	6.6±0.1	ND	ND	0.39±0.02	0.48±0.01	100
0.201±0.001	0.73±0.03	7.6±0.1	ND	ND	0.42±0.02	0.26±0.01	92
0.298±0.001	0.58±0.02	8.4±0.1	ND	ND	0.44±0.02	0.20±0.01	93

* Other fermentation products were not observed under any condition.
C_{12} FA is the combination of dodecanoic acid and dodecenoic acid
C_{14} FA is the combination of tetradecanoic acid and tetradecenoic acid
ND refers to not detected or below minimum standard of standard curve

Similar to carbon limited cultivations, maximum FFA biomass specific productivity occurred between dilution rates of 0.2 and 0.25 h^{-1} **(**Figure 4.2a**)**. However, FFA specific productivities were consistently three times larger under phosphate limitation, reaching a maximum value (for all FFA species combined) of 0.068 g FFA/(gDCW*hr). Unlike the trends

observed under carbon limitation, the yield of FFAs on glucose increased from the lowest dilution rate (0.05 h^{-1}) to a maximum value of 0.1 g FFA/g glucose at a dilution rate of 0.1 h^{-1} **(**Figure 4.2b**)**. FFA yields decreased beyond this maximum but remained significantly greater than the corresponding value obtained under carbon limitation.

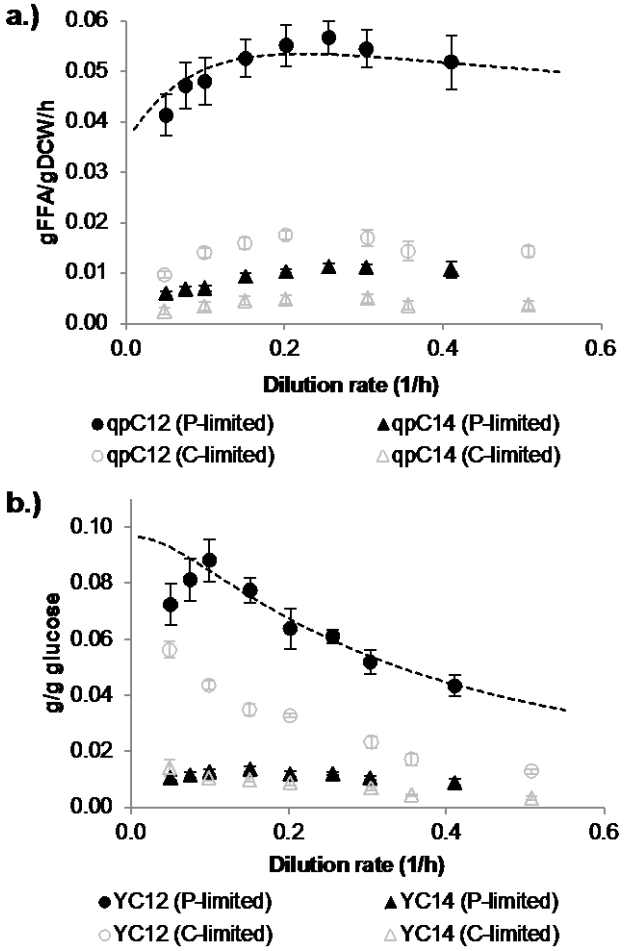


Figure 4.2: FFA productivity (A) and yield (B) values calculated for each dilution rate tested. Shown are values for C_{12} and C_{14} species under both nutrient limitation conditions. Experimental values from phosphorus limiting conditions are represented by filled symbols, closed symbols represent previous data from chemostats run under carbon limitation (Youngquist et al. 2012), and the production model is represented by the dashed line. Error bars represent standard deviations of 4-5 samples taken over a residence time at each steady state.

Using the biomass specific productivities in Figures. 4.2 and those calculated from Tables 4.1 and 4.2, a carbon mass balance was performed for each dilution rate. For most dilution rates, over 90% of the fed carbon was found in residual glucose, biomass, FFA, CO_2, and assayed fermentation products. At the two smallest dilution rates, only 80% of the carbon was accounted for. For these dilution rates, extracellular protein content was assayed using a Bradford reagent and elevated protein levels were not detected (data not shown). Flow cytometric analysis of cells stained with SYTOX green (permeability assay) showed only a population of intact cells (Chapter 6, Figure 6.2) unlike stationary phase samples of similar *E. coli* strains expressing BTE (Lennen et al., 2011). The breakdown of the carbon products for selected dilution rates can be seen in Figure 4.3a. Under all but the highest dilution rate, CO_2 represented the largest fraction of carbon followed by biomass and FFA.

To compare how the carbon flow in TY05 differed from a non-producing strain, *E. coli* TY06 (Chapter 3) was grown in chemostat culture at the dilution rates of 0.05, 0.1, 0.2, and 0.3 h^{-1}. The negative control strain, TY06, followed a similar carbon distribution trend to TY05, except with insignificant carbon flow to FFA and increases to biomass, CO_2, and acetate (Figure 4.3b). Carbon flow to CO_2 remained the largest fraction, though amount to biomass increased close to 10% at the lower dilution rates. In contrast to TY05, TY06 did not consume glucose as quickly under the same growth conditions. Additionally, the significant amount of unaccounted for carbon was not seen in the TY06 negative control strain (Figure 4.3b) at lower dilution rates. This result could indicate that the FFA titer in TY05 was underestimated as TY06 was not found to produce significant amounts of any other product at those dilution rates.

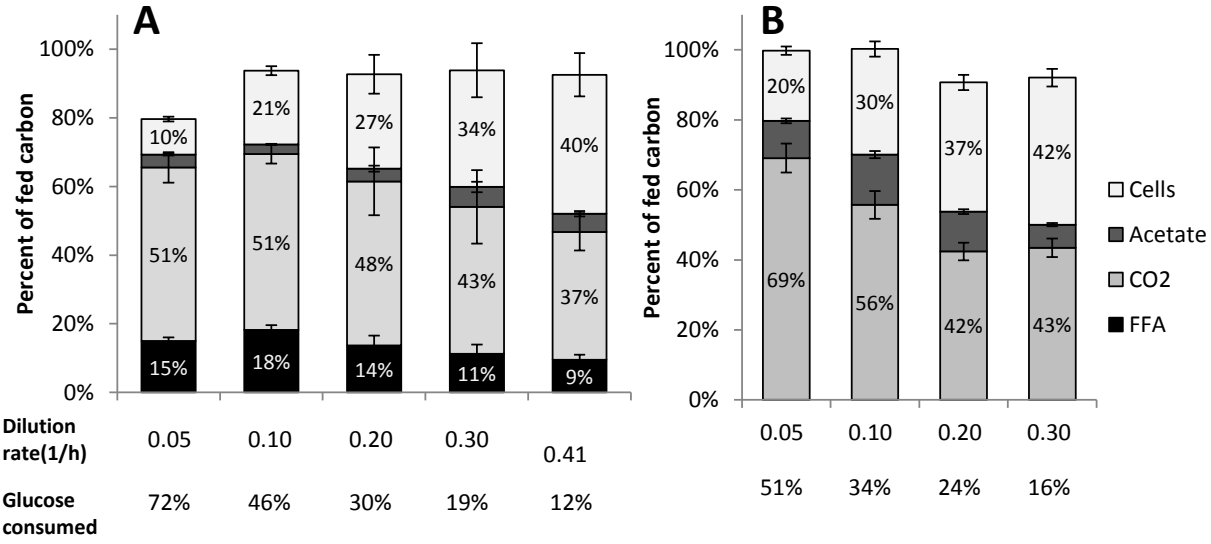


Figure 4.3: Graphical representation of the carbon mass balance for *E. coli* TY05 (A) and TY06 (B) grown under phosphate limitation. Mass balances for the dilution rates of 0.05, 0.1, 0.2, 0.3, and 0.4 h^{-1} are shown. Percent glucose consumption is also indicated.

4.2.3. Analysis of the stability of E. coli TY05

The production of FFA introduces many physical stresses and is therefore detrimental to *E. coli* physiology (Lennen et al., 2011). In order for long, FFA production runs to occur, production strains must be genetically stable. To examine the stability of TY05, we compared the FFA production phenotype of the original strain to the phenotype of cultures started from freezer stocks prepared from various points in the 240 hr chemostat cultivation. The resulting titers were statistically equivalent (Figure 4.4) indicating that the strain did not lose the FFA production trait over 240 hr.

In addition to the phenotypic comparison, the genomes of the stock TY05 and the cells collected at the end of the chemostat experiment were sequenced using high-throughput DNA pyrosequencing. Relative to the published sequence of MG1655, the stock TY05 strain contained the expected genetic manipulations that replaced beta-oxidation genes with BTE expression cassettes. In addition, 22 point mutations were identified (Appendix 7). The sequence of TY05

(post chemostat) contained the same mutations as the stock culture and an additional 15 point mutations. None of the additional mutations is predicted to change the sequence of any proteins. Mutations identified in 3 genes were found to be common mutations in MG1655 stock cultures based on data from the National Center for Biotechnology Information database also showed mutations in the strain. These results suggest that TY05 is genetically and phenotypically stable over 240 hrs of continuous cultivation.

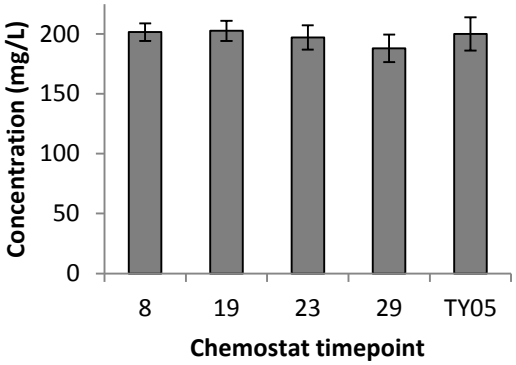


Figure 4.4: C_{12} FFA production titer from triplicate cultures started from freezer stocks from different timepoint during chemostat runs. This particular run used the growth rates came from the phosphate limited chemostat testing dilution rates of 0.25 (8), 0.075 (19), 0.15 (23), and 0.4 (29) h^{-1}.

4.2.4 Kinetic model of FFA production under phosphate limitation

A kinetic model was updated using the chemostat data collected for *E. coli* TY05 under phosphorus limiting conditions. As seen in Figure 4.3a**,** the biomass specific productivity under phosphorus limitation was non-linear with respect to dilution rate, with a similar relationship to that seen under carbon limitation. While an equation with a maximum in q_p at an intermediate dilution rate could be fit to the data, addition of a non-growth associated constant (Equation 4.5)

to the equation previously determined for FFA production under carbon limitation (Chapter 3) provided a better fit over both conditions. The fact that *E. coli* continues to import carbon and remain metabolically active in stationary phase under phosphorus limiting conditions (Ballesteros et al., 2001) supports the use of a non-growth associated term for the production of FFA under non-carbon limitation. The fact that none of the other model parameters were altered further supported this hypothesis. The optimal value of δ was 0.036 g_{C12}/gDCW/h, while values for α, β, and γ were found to remain at 0.044 gC_{12}/gDCW/h, 0.17 h^{-1}, and 3.4 h respectively.

$$q_p = \frac{\alpha\mu}{\beta + \mu + \gamma\mu^2} + \delta \tag{4.5}$$

To see if there was a similar effect to the yield of biomass on substrate (Y_{xs}) and maintenance energy (m_s) terms, the experimental q_p terms were factored out of the equation to generate a modified glucose consumption (q_s) term as a linear function of growth (Figure 4.5). In generating the modified glucose consumption term, the experimental biomass specific productivities of each component are divided by its maximum theoretical yield and then subtracted from the experimentally observed specific substrate uptake rates. The resulting modified glucose uptake rate versus growth has a linear relationship where maximum biomass yield (inverse slope) and maintenance energy (y-axis intercept) can be determined by a linear fit. The maximum yield of biomass on glucose (Y_{xs}) was found to be 0.54 g/g, in line with published values (Paalme et al., 1995), while m_s was found to be 0.28 g glucose/gDCW/h, significantly higher than the value for FFA producing *E. coli* under carbon limitation (Chapter 3), for wild-type *E. coli* K-12 (Heijnen and Roels, 1981), and for TY06 under phosphate limitation (0.12 g/gDCW/h). The resulting modified q_s equation was used in conjunction with the q_p equation to generate an updated model for yield of FFA on glucose under limiting phosphate conditions (Figure 4.2b). The model matched FFA yields for dilutions rates above 0.1 h^{-1}.

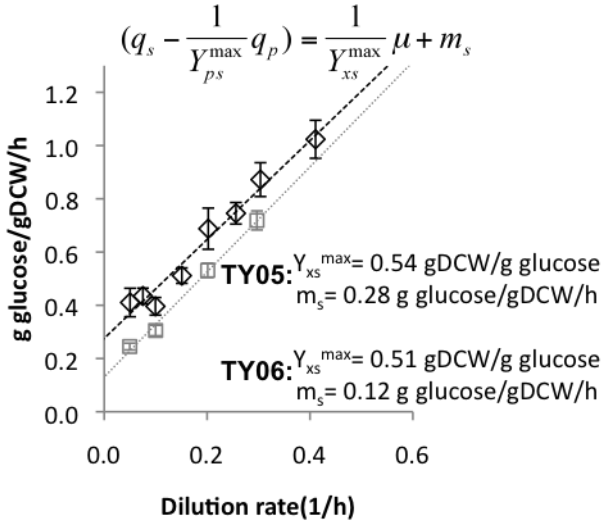


Figure 4.5: The maximum yield of biomass on glucose ($Y_{x/s}^{max}$, inverse slope) and maintenance energy (m_s, intercept) for TY05 (open diamonds) and TY06 (open squares) were calculated from a plot of modified glucose consumption (with product formation removed) versus dilution rate. The observed biomass specific productivity of each component is divided by its maximum theoretical yield and then subtracted from the observed specific substrate uptake rate giving a modified specific substrate uptake rate affected by only the biomass and maintenance terms.

4.2.5. Effect of nutrient ratios on FFA specific productivity

Next, the effect of the media carbon to phosphorus ratio on FFA biomass specific productivity and yield was tested. A set of chemostat experiments was performed holding a constant dilution rate (0.1 h^{-1}) while varying either the concentration of phosphate or glucose in the feed. Experiments were performed using glucose concentrations of 0.4, 0.7, 1.0 and 2.0% (w/v) in MOPS minimal media containing 370 µM phosphate and using a glucose concentration of 1.0% (w/v) in MOPS minimal media containing either 185, 370, or 530 µM phosphate. These concentrations correspond to C:P (mole carbon atom per mole phosphate ion) ratios ranging from 360 to 1800. A C:P molar ratio below 120 is required to enter limiting carbon conditions (Neidhardt et al., 1974). For each C:P ratio, samples for biomass, glucose, FFA, and alternative fermentation products were taken five times over a period of 16 hours (1.5 residence times) after

a steady state (as defined above) was established. Residual glucose in the reactor effluent was found for all conditions except the lowest ratio (0.4% glucose in MOPS minimal media containing 370 μM phosphate). For phosphate limited conditions, biomass, FFA, and CO_2 effluent concentrations increased with increased phosphate content in the media feed while glucose variations had no impact. Acetate was produced under all tested conditions except the lowest C:P ratio. When mass balances were calculated, 82-98% of carbon fed to the system was accounted under all steady states (Table 4.3).

Table 4.3: Steady state concentrations of assayed metabolites from continuous culture of TY05 in modified MOPS minimal media at a dilution rate of 0.1 h^{-1}

C:P ratio	Variation (C or P)	Biomass g/L	Glucose g/L	C_{12} FA mg/L	C_{14} FA mg/L	CO_2 g/h	Acetate* g/L	Carbon recovery %
1800	C	0.79±0.03	14.5±1.4	371±37	62±5	0.51±0.01	0.14±0.01	82
1800	P	0.45±0.03	7.6±0.1	212±8	32±2	0.26±0.01	0.10±0.01	89
900	C	0.80±0.03	5.5±0.5	356±21	59±3	0.51±0.02	0.14±0.03	87
900	P	0.84±0.05	5.3±0.1	403±32	58±3	0.50±0.02	0.13±0.01	94
630	C	0.75±0.04	2.3±0.1	369±28	59±6	0.47±0.01	0.16±0.01	87
630	P	1.15±0.03	3.0±0.3	661±46	92±5	0.70±0.04	0.12±0.01	90
360	C	0.74±0.03	0	357±32	64±5	0.37±0.02	0	86

* Other fermentation products were not observed under any condition.
C_{12} FA is the combination of dodecanoic acid and dodecenoic acid
C_{14} FA is the combination of tetradecanoic acid and tetradecenoic acid

Data from Table 4.3 was used to calculate q-rates and yields for all measured compounds. In general, trends in q-rates were independent of whether glucose (C) or phosphate (P) was varied. Specific FFA and CO_2 productivities remained nearly constant under all tested conditions (Figure 4.6). The average $q_{p\text{-}C12}$ value was 0.049±0.004 gC_{12} FFA/(gDCW*hr) with small variations over the range of C:P ratios tested. The average q_{CO2} value was 0.41±0.03 gCO_2/(gDCW*hr) with the only significant deviation occurring at the lowest C:P ratio (Figure 4.6b). At this low C:P ratio (360) and a low dilution rate, it is likely the culture is bordering on

co-limitation where effects of carbon limitation begin to appear. Therefore, no C:P ratios below 360 were tested.

Unlike the other productivities, glucose uptake rates differed significantly between glucose and phosphate variations. As expected, glucose uptake rates decreased with decreasing phosphate concentrations (i.e. increasing C:P ratio), consistent with decreased biomass concentrations (Table 4.3). When glucose concentrations were varied, glucose uptake rates averaged 0.58±0.05 g glucose/(gDCW*hr) with the only significant deviation occurring at the lowest C:P ratio (Figure 4.6a). Despite the differences in glucose uptake, the biomass specific productivity of FFA (Figure 4.6c) was not affected and can be assumed to remain constant at a constant dilution rate under phosphate limitation.

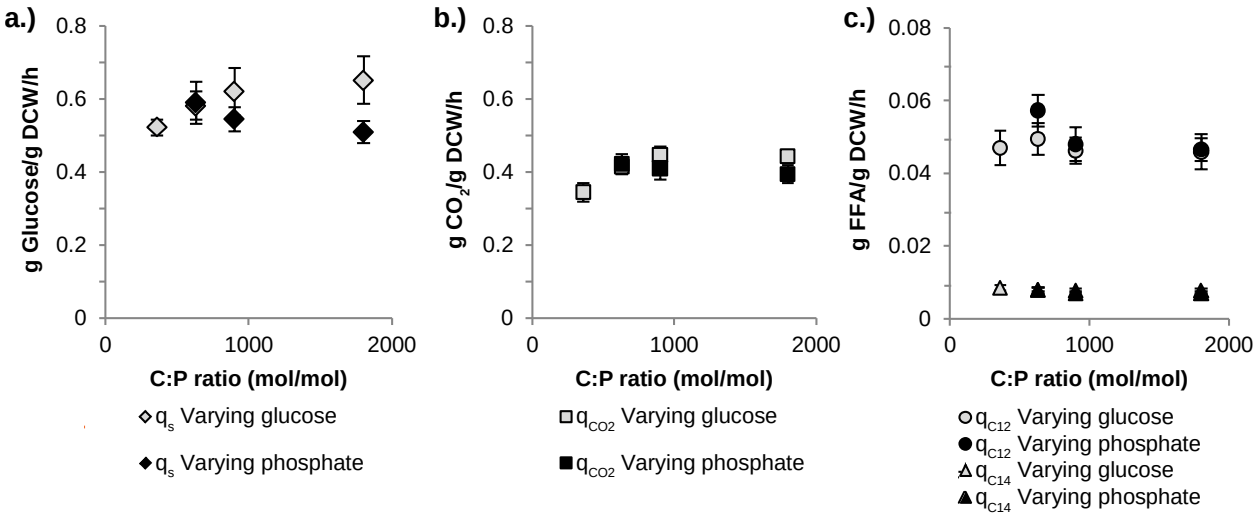


Figure 4.6: The effect of C:P ratio while varying glucose or phosphate feed concentrations on (A) q_s, (B) q_{CO2}, and both (B) q_{C12} and q_{C14}. Error bars represent standard deviations of 5 samples taken over 16 hours for each steady state.

4.2.6. Model validation for alternative fermentation condition

As a test of the FFA production model, three 48 h batch cultures were performed in a bioreactor using MOPS minimal media containing 370 µM phosphate and 0.7% glucose (Figure

4.7). Off-gas composition was recorded continuously throughout to indirectly monitor metabolic activity in stationary phase. Sugar and optical density samples were taken throughout the experiment. FFA samples were taken at 3, 6, 9, 18, 21, 24, 30, and 48 hours after induction. However, after the 9 hour sample a significant amount of foam accumulated within the reactor and did not exit into a foam overflow trap bottle. For this reason, an accurate FFA titer could not be taken between 9 and 18 hours, when the remaining foam in the reactor was collapsed using the addition of antifoam (Antifoam 204, Sigma Aldrich) and the concentration of FFAs collected in the foam trap bottle was determined.

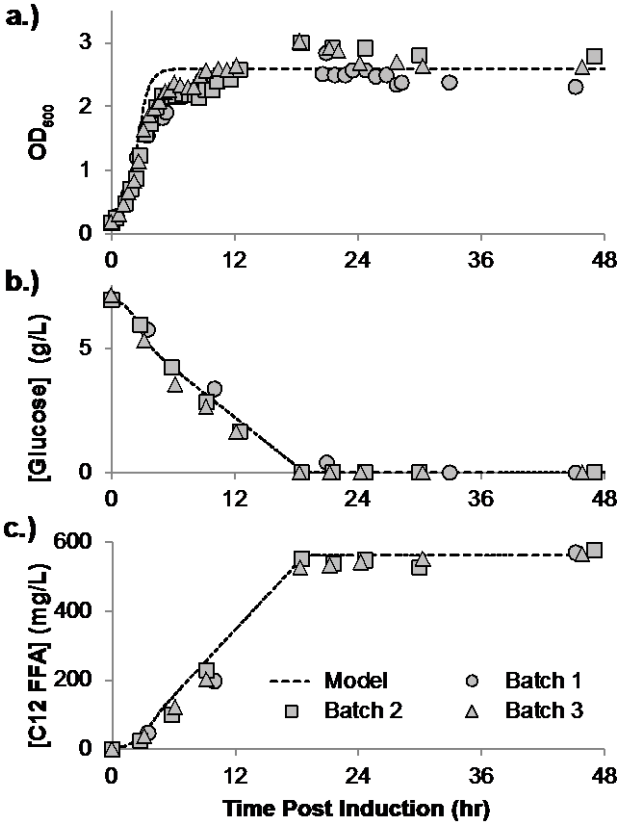


Figure 4.7: Model prediction (dotted lines) and actual production (filled points) during three controlled batch fermentations under phosphate limited conditions. Optical density (A), glucose concentration (B), and fatty acid concentration (C) are represented.

The kinetic model used both q_s and q_p equations as a function of μ. For the Monod growth model, a K_s of 2 μM was estimated from the non-zero, residual phosphate data points

(taken at high dilution rates). The model predicted FFA titer was insensitive to the value of K_s; in part because the majority of production came after the phosphate concentration had dropped below C_{in}^{max}. Most FFA production occurred once C_{in} had dropped to the value of C_a, where q_p became a constant until all glucose was consumed.

The consumption of glucose and cell density predicted by the model were consistent with the experimental data from the controlled batch fermentation. The measured FFA concentrations were also consistent with predictions up to 9 h and after 18 h. Unfortunately, the inability to collect accurate FFA titers between 9 and 18 h prevented us from confirming whether the FFA production ceased as abruptly as the model and CO_2 data suggested. The model accurately estimated the final titer of FFA. Around 0.5 g/L acetate was observed at the 24 hour timepoint (immediately prior to full glucose consumption), almost all of which was consumed by the 48 h timepoint. As cells are still metabolically active during the second 24 hour period before the complete consumption of acetate, some may go towards producing FFAs, a feature not captured in the model.

4.3. Discussion

The experimental and modeling results suggest several areas where improvements in fatty acid production could be made by additional metabolic engineering. First, elimination of one acetate production pathway that also produces carbon dioxide (via a knockout of *poxB*) could provide a boost as acetate production currently accounts for 3-6% of the carbon consumed in *E. coli* TY05. Adding a *poxB* deletion to strains grown under phosphate limitation could give a minor boost in FFA yield and specific productivity. Second, the observed maintenance energy of 0.28 g glucose/gDCW/h under phosphate limitation is already four times higher than that seen under carbon limitation and much higher than reported values for wild type *E. coli* (Button 1985;

Heijnen and Roels, 1981). Higher maintenance energies have been reported for bacteria under excess carbon conditions (Tempest and Neijssel, 1984; Zeng and Deckwer, 1995), however, this phenomenon appears to be the result of energy spilling (e.g. futile cycles) to consume the excess carbon (Russell, 2007). Spurious reactions involved in energy spilling could be identified in a transcriptomic and proteomic analysis of the strain. Further analysis into the cause of the elevated maintenance energy could provide insight into ways to further increase FFA yields. A similar FFA-producing *E. coli* strain was found to be subject to elevated membrane stress and was less viable in comparison to a non-producing control strain in stationary phase (Lennen et al., 2011). *E. coli* could be expending energy in response to these stresses, thereby increasing the maintenance energy at the expense of FFA production. Lastly, under phosphate-limited growth, a large percentage of the fed glucose remains unconsumed. One strategy for increasing glucose consumption is to evolve strains under non-carbon limiting conditions in a very slow growth or stationary phase to select for cells with higher metabolic activity and increased specific glucose uptake rates. Such a strategy was successfully implemented to increase specific glucose uptake rate under nitrogen limited conditions (Sonderegger et al., 2005). As phosphate starved cells have greater metabolic activity in stationary phase than nitrogen starved cells (Ballesteros et al., 2001), there may not be as significant of an improvement in glucose uptake rate. However, any effort to increase the rate of glucose uptake under non-carbon limitations could increase the rate of FFA production and FFA yield if excess carbon flux can be directed through fatty acid biosynthesis. Future efforts will focus on these and other modifications to engineer a strain that will come closer to producing FFAs at their theoretical limit (0.35 g FFA/g glucose) (Lennen and Pfleger, 2013).

In summary, phosphate-limitation was shown to increase FFA production from an engineered strain of *E. coli*. Using data collected from chemostat experiments, a kinetic model of FFA production in *E. coli* was updated to include production under phosphate limiting conditions. Maximum yields were obtained at low dilution rates and maximum productivities were obtained at lower to intermediate dilution rates. FFA yields and productivities were not impacted by the ratio of carbon to phosphate as long as the media remained phosphate limited. The model was used to successfully predict the production of a controlled batch culture under phosphorus limitation. The optimal conditions identified generated a 40% increase in yield and 300% increase in productivity versus the optimal conditions observed under carbon limitation.

4.4. Materials and Methods

4.4.1. Bacterial strains

All experiments were performed using *E. coli* TY05 (Chapter 3) or *E. coli* TY06. Both contain three chromosomal copies (replacing *fadD*, *fadE*, and *fadAB*) of an expression cassette for either the acyl-ACP thioesterase (BTE) from *Umbellularia californica* (TY05) or a non functional mutant version of the thioesterase (BTE-H204A, where the active site histidine was been altered to an alanine, TY06) under the control of the P_{trc} promoter . When BTE expression is induced with IPTG, *E. coli* TY05 produces $C_{12:0}$, $C_{12:1}$, $C_{14:0}$ and $C_{14:1}$ FFA.

4.4.2. Chemicals, enzymes, and other materials

Enzymes were purchased from New England Biolabs (Ipswich, MA). DNA purification materials were purchased from Qiagen (Venlo, Netherlands). Chemicals were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Hampton, NH) unless otherwise specified.

4.4.3. Batch shake flask studies

FFA production under different nutrient limitation conditions was determined by cultivating *E. coli* TY05 in 250 mL shake flasks containing 50 mL MOPS minimal medium (Neidhardt et al., 1974) supplemented with 1.0% glucose, 1.32 mM dibasic potassium phosphate, 0.276 mM potassium sulfate, and 9.5 mM ammonium chloride with variations depending on nutrient limitation. For nitrogen limitation, the total amount of ammonium chloride in the media was 4.75 mM. To test carbon limitation, the medium was supplemented with 0.4% glucose. Under phosphate limitation, the medium contained 0.240 mM phosphate. Media conditions for phosphate and nitrogen limitation were determined based on the original formulation of MOPS minimal media (Neidhardt et al., 1974) and are consistent with past phosphate limitation experiments in MOPS minimal media (Baek and Lee, 2006; Gerard et al., 1999). Concentrations significantly under the C:P ratio limit of 120 moles of carbon (molC) per mole of phosphate (molP) were used throughout.

For all studies, single colonies from freezer stocks were inoculated into 5 mL of LB and incubated overnight at 37°C with shaking. Shake flask production cultures were inoculated to an OD_{600} of 0.04 from the overnight cultures and incubated at 37°C until each reached OD_{600} 0.2, when IPTG (1 mM final concentration) was added to induce BTE. After 40 hours of incubation at 37°C, samples for fatty acid methyl-ester (FAME) analysis, glucose consumption, and cell density were taken, processed, and quantified.

4.4.4. Chemostat studies

Chemostat experiments were performed in a 3L stirred bioreactor (Applikon Biotechnology Inc, Schiedam, Netherlands) using a 1.5 L working volume. Temperature, pH, airflow, dissolved oxygen, off-gas composition, liquid flow, feed and waste masses, and stirrer

speed were controlled/recorded as described in Chapter 3. Chemostat experiments were inoculated using the following propagation schedule. Cultures (5 mL) of TY05 grown in LB at 37ºC for 10 h were used to inoculate 50 mL of MOPS minimal media supplemented with 1.0% glucose. The 50 mL culture was grown at 37ºC overnight and was used as the inoculum for the chemostat. For dilution rate experiments, modified MOPS minimal media (containing 370 µM phosphate) supplemented with 1.0% glucose, 0.276 mM potassium sulfate, and 9.5 mM ammonium chloride was inoculated to an OD_{600} of 0.04. Induction with 1 mM IPTG occurred when the OD_{600} reached 0.2. The reactor was operated in batch mode for 9 hours, at which point flow of fresh media was initiated. Outflow was controlled by maintaining a constant liquid level through the placement of the outlet line at a height corresponding to a 1.5 L working volume. The OD_{600} of the culture was monitored periodically (18 samples per 24 h) during batch and continuous operation. Steady state determination and sampling procedures were performed as described in Chapter 3. During the dilution rate variation experiments, two chemostat experiments were performed, with flow regimes tested of 0.10, 0.05, 0.3, 0.2 and 0.25, 0.075, 0.15, and 0.4 h^{-1} respectively. As such, growth rates ranging from 7% to 60% of μ_{max} were covered (μ_{max} was found to be 0.69 h^{-1} under phosphate limitation).

For C:P ratio experiments, starter cultures followed the same propagation scheme as the dilution rate experiments. Once all samples were taken for a given C:P ratio, the sterile media feed was changed to MOPS minimal media supplemented with a different glucose or phosphate concentration while maintaining a constant dilution rate of 0.1 h^{-1}. Two separate chemostat experiments were performed. The first C:P ratio experiment held the phosphate concentration in the media constant at 370 µM while testing glucose feed concentrations of 7, 4, 20, and 10 g/L respectively. The second C:P ratio experiment held glucose concentrations constant at 1.0%

while varying feed phosphate concentrations of 370, 185, and 530 μM. These steady states correspond to C:P ratios of 1800, 900, 630, and 360 molC/molP.

For experiments with TY06, starter cultures followed the same propagation scheme as with TY05. TY06 experiments were run in similar fashion to the TY05 dilution rate experiments, using modified MOPS minimal media (containing 370 μM phosphate) supplemented with 1.0% glucose, 0.276 mM potassium sulfate, and 9.5 mM ammonium chloride. The flow regimes of 0.05, 0.1, 0.2, and 0.3 h^{-1} were tested. As such, growth rates ranging from 7% to 41% of μ_{max} were covered.

4.4.5. Batch bioreactor studies

Controlled batch cultivation experiments were performed using the same equipment as the chemostat experiments. Starter cultures were prepared using the same propagation scheme as the chemostat experiments. A 1.5 L working volume of MOPS minimal media (containing 370 μM phosphate) supplemented with 0.7% glucose, 0.276 mM potassium sulfate, and 9.5 mM ammonium chloride was used. Cultures were induced with 1mM IPTG at OD_{600} 0.2. Fatty acid, cell density (OD_{600}), and supernatant samples were taken periodically over 48 hours after induction. pH was maintained at 7.00±0.05 via the addition of 10% (v/v) NH_4OH base.

4.4.6. Product analysis

Culture samples for analyzing fatty acids via GC/MS were taken and derivatized to methyl esters as described previously (Lennen et al., 2010; Chapter 3). All $C_{12:0}$, $C_{12:1}$, $C_{14:0}$, and $C_{14:1}$ were considered FFAs as results from Chapter 3 indicate that less than 10% of these species are found in membrane fractions. Supernatant samples for HPLC quantification of sugars, alcohols, and short chain carboxylic acids were collected and processed as described in Chapter

3. Supernatant samples were also analyzed for phosphate content using a Malachite Green Phosphate Assay Kit (BioAssay Systems, Hayward, CA) following manufacturer's instructions. Additionally, protein levels were determined using a Bradford assay (Bradford 1976). Concentrations of bovine serum albumin were used to generate a standard curve. To each well of a microtiter plate, 293 µL Bradford reagent and 7 µL sample or standard were added in technical triplicate. Absorbance at 595 nm was measured using a microtiter plate reader (Tecan, Männedorf, Switzerland). A linear correlation between OD_{600} and dry cell weight (DCW) was determined as described previously (Chapter 3).

4.4.7. Kinetic modeling

In order to determine how growth rate changed with the limiting substrate (phosphate) concentration, a Monod model (Monod, 1949) was initially used for high concentrations of extracellular phosphate (equation 4.1). The value of μ_{max} was estimated from the initial batch growth curve (prior to switching to chemostat operation). Levels of residual, extracellular phosphate were below the limit of detection for all but two conditions. Values for K_{PO4} were determined from steady state phosphate levels at different dilution rates using equation 1. Monod models are not accurate after the incorporation of all extracellular phosphate because growth continues after this point in a batch reactor despite the extracellular concentration of limiting substrate going to zero (Baek and Lee, 2006; Lubke et al., 1995). Modifications to the Monod model to address the consumption of intracellular polyphosphate, first described by Nyholm (Nyholm 1976), were added in the form of equation 4.2. Here, C_a is the minimum internal concentration of phosphate required for a viable *E. coli* cell, C_{in} is the internal concentration of phosphate, and C_{in}^{max} is the concentration of internal phosphate when there is sufficient extracellular phosphate in the media. C_{in}^{max} was assumed to be the value at which the *Pho*

regulon (phosphate starvation response) is at full expression, 0.4 µM (Wanner, 1993; Hsieh and Wanner, 2010). A value of 0.28 µM was used for C_a, as literature values of intracellular phosphate and polyphosphate concentrations under phosphate limitation (Fagerbakke et al., 1996; Sharfstein et al., 1996) indicated similar levels. Once C_{PO4} had dropped below C_{in}^{max}, it was assumed that C_{PO4} equaled C_{in} and µ was determined by equation 4.2.

$$\mu = \mu_{\max} \frac{C_{\mathrm{PO4}}}{K_{\mathrm{PO4}} + C_{\mathrm{PO4}}} \qquad \text{when } C_{\mathrm{PO4}} > C_{in}^{\max} \tag{4.1}$$

$$\mu = \mu_{\max} \frac{C_{in} - C_a}{C_{in}^{\max} - C_a} \qquad \text{when } C_{\mathrm{PO4}} \le C_{in}^{\max} \tag{4.2}$$

While the standard Herbert-Pirt equation was used to determine glucose uptake (equation 4.3), a simplified Herbert-Pirt equation (equation 4.4) (Herbert et al., 1956) was used to determine the phosphate uptake rate into the cell. In the Herbert-Pirt equation q_s and q_p represent the biomass specific substrate uptake rate and biomass specific productivity, respectively. The product-associated (bq_p) and maintenance-associated (m_s) phosphate consumption terms, which are significant under carbon-limitation, were eliminated because phosphate is not consumed to generate extracellular products or for cellular maintenance (Mallette et al., 1964). The remaining constant (a_{PO4}) is the inverse of the maximum biomass yield on phosphate which was determined by a linear plot of q_{PO4} versus growth rate, with (a_{PO4}) being the slope. Once the relationships between q_s for glucose (equation 4.3) and q_p versus growth were determined (equation 5), a model was generated to predict FFA production in *E. coli* under all growth conditions for phosphate limitation.

$$q_s = a\mu + bq_p + m_s \tag{4.3}$$

$$q_{\mathrm{PO4}} = a_{\mathrm{PO4}}\mu \tag{4.4}$$

Equations (4.1), (4.2), (4.3), (4.4), and (4.5) were solved by Euler method implemented in MATLAB (MathWorks, Inc. Natick, MA) to predict FFA production under given reactor configuration, media composition, and growth rate.

4.4.8. Chemostat strain stability testing

Freezer stocks of TY05 grown in chemostat culture were taken every day for each continuous culture run. To determine if the strain was phenotypically different from the strain used at the start of the experiment, FFA titers were measured for cultures grown from individual colonies isolated from each freezer stock and compared to the baseline TY05 control. Individual colonies were used to inoculate 5 mL of LB overnight at 37°C with shaking. Overnight cultures were used to innoculate 250 mL shake flasks containing 50 mL MOPS minimal medium (Neidhardt et al. 1974) supplemented with 0.4% glucose, 1.32 mM dibasic potassium phosphate, Shake flask production cultures were inoculated to an OD_{600} of 0.04 from the overnight cultures and incubated at 37°C until each reached OD_{600} 0.2, when IPTG (1 mM final concentration) was added to induce BTE. After 40 hours of incubation at 37°C, samples for fatty acid methyl-ester (FAME) analysis, glucose consumption, and cell density were taken, processed, and quantified.

In addition to phenotypic testing, the genotype of TY05 prior to and after 170 hours of continuous culture were compared by whole genome re-sequencing. Genomic DNA (gDNA) was isolated from cultures of *E. coli* TY05 grown from single colonies of the freezer stocks made prior to and after the chemostat studies using a Wizard® gDNA purification kit (Promega, Madison, WI). The resulting gDNA was sequenced by the Department of Energy Joint Genome Institute (JGI) (Walnut Creek, CA). Deviations from the base sequence were identified using the Integrative Genomics Viewer software (Broad Institute, Cambridge, MA).

4.5. References:

Agnew DE, Stevermer AK, Youngquist JT, Pfleger BF. 2012. Engineering *Escherichia coli* for production of C12-C14 polyhydroxyalkanoate from glucose. *Metabolic Engineering* 14 (6):705-713.

Baek JH, Lee SY. 2006. Novel gene members in the Pho regulon of *Escherichia coli*. *FEMS Microbiology Letters* **264**(1):104-109.

Ballesteros M, Fredriksson A, Henriksson J, Nystrom T. 2001. Bacterial senescence: protein oxidation in non-proliferating cells is dictated by the accuracy of the ribosomes. *EMBO Journal* **20**(18):5280-5289.

Bradford MM. 1976. Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Analytical Biochemistry* **72**(1-2):248-254

Button DK. 1985. Kinetics of nutrient-limited transport and microbial-growth. *Microbiological Reviews* **49**(3):270-297

Fagerbakke KM, Heldal M, Norland S. 1996. Content of carbon, nitrogen, oxygen, sulfur and phosphorus in native aquatic and cultured bacteria. *Aquatic Microbial Ecology* **10**(1):15-27.

Ferenci T. 2008. Bacterial physiology, regulation and mutational adaptation in a chemostat environment. *Advances in Microbial Physiology* **53**:169-229.

Gerard F, Dri AM, Moreau PL. 1999. Role of *Escherichia coli* RpoS, LexA and H-NS global regulators in metabolism and survival under aerobic, phosphate-starvation conditions. *Microbiology* **145**:1547-1562

Goh E-B, Baidoo EEK, Keasling JD, Beller HR. 2012. Engineering of Bacterial Methyl Ketone Synthesis for Biofuels. *Applied and Environmental Microbiology* **78**(1).

Hassan M, Blanc PJ, Granger LM, Pareilleux A, Goma G. 1996. Influence of nitrogen and iron limitations on lipid production by *Cryptococcus curvatus* grown in batch and fed-batch culture. *Process Biochemistry* **31**(4).

Heijnen JJ, Roels JA. 1981. A macroscopic model describing yield and maintenance relationships in aerobic fermentation processes. *Biotechnology and Bioengineering* **23**(4):739-763.

Herbert D, Elsworth R, Telling RC. 1956. The continuous culture of bacteria; a theoretical and experimental study. *Journal of general microbiology* **14**(3):601-622

Hsieh Y-J, Wanner BL. 2010. Global regulation by the seven-component Pi signaling system. *Current Opinion in Microbiology* **13**(2).

Hua Q, Yang C, Oshima T, Mori H, Shimizu K. 2004. Analysis of gene expression in *Escherichia coli* in response to changes of growth-limiting nutrient in chemostat cultures. *Applied and Environmental Microbiology* **70**(4):2354-2366.

Huber R, Roth S, Rahmen N, Buechs J. 2011. Utilizing high-throughput experimentation to enhance specific productivity of an *E. coli* T7 expression system by phosphate limitation. *BMC Biotechnology* **11**.

James OO, Maity S, Mesubi MA, Usman LA, Ajanaku KO, Siyanbola TO, Sahu S, Chaubey R. 2012. A review on conversion of triglycerides to on-specification diesel fuels without additional inputs. *International Journal of Energy Research* **36**(6).

Johansson L, Lindskog A, Silfversparre G, Cimander C, Nielsen KF, Liden G. 2005. Shikimic acid production by a modified strain of *E. coli* (W3110.shik1) under phosphate-limited and carbon-limited conditions. *Biotechnology and Bioengineering* **92**(5):541-552.

Johnson K, Kleerebezem R, van Loosdrecht MCM. 2010. Influence of the C/N ratio on the performance of polyhydroxybutyrate (PHB) producing sequencing batch reactors at short SRTs. *Water Research* **44**(7):2141-2152.

Kalscheuer R, Stolting T, Steinbuchel A. 2006. Microdiesel: *Escherichia coli* engineered for fuel production. *Microbiology* **152**:2529-2536.

Knothe G. 2010. Biodiesel and renewable diesel: A comparison. *Progress in Energy and Combustion Science* **36**(3):364-373.

Kottmeier K, Mueller C, Huber R, Buechs J. 2010. Increased product formation induced by a directed secondary substrate limitation in a batch *Hansenula polymorpha* culture. *Applied Microbiology and Biotechnology* **86**(1):93-101.

Kumar R, Shimizu K. 2011. Transcriptional regulation of main metabolic pathways of *cyoA*, *cydB*, *fnr*, and *fur* gene knockout *Escherichia coli* in C-limited and N-limited aerobic continuous cultures. *Microbial Cell Factories* **10**.

Lennen RM, Braden DJ, West RM, Dumesic JA, Pfleger BF. 2010. A Process for Microbial Hydrocarbon Synthesis: Overproduction of Fatty Acids in *Escherichia coli* and Catalytic Conversion to Alkanes. *Biotechnology and Bioengineering* **106**(2):193-202.

Lennen RM, Kruziki MA, Kumar K, Zinkel RA, Burnum KE, Lipton MS, Hoover SW, Ranatunga DR, Wittkopp TM, Marner WD, II, Pfleger BF. 2011. Membrane Stresses Induced by Overproduction of Free Fatty Acids in *Escherichia coli*. *Applied and Environmental Microbiology* **77**(22):8114-8128.

Lennen RM, Pfleger BF. 2012. Engineering *Escherichia coli* to synthesize free fatty acids. *Trends in Biotechnology* **30**(12):659-667.

Lu XF, Vora H, Khosla C. 2008. Overproduction of free fatty acids in *E. coli*: Implications for biodiesel production. *Metabolic Engineering* **10**(6):333-339.

Lubke C, Boidol W, Petri T. 1995. Analysis and optimization of recombinant protein-production in *Escherichia coli* using the inducible phoA promoter of the *Escherichia coli* alkaline-phosphatase. *Enzyme and Microbial Technology* **17**(10):923-928.

Mallette MF, Cowan CI, Campbell JJ. 1964. Growth and Survival of *Escherichia coli* in Medium Limited in Phosphate. *Journal of Bacteriology* **87**:779-785.

Marzan LW, Shimizu K. 2011. Metabolic regulation of *Escherichia coli* and its *phoB* and *phoR* genes knockout mutants under phosphate and nitrogen limitations as well as at acidic condition. *Microbial Cell Factories* **10**.

Matsuoka Y, Shimizu K. 2011. Metabolic regulation in *Escherichia coli* in response to culture environments via global regulators. *Biotechnology Journal* **6**(11):1330-1341.

McIntyre JJ, Bull AT, Bunch AW. 1996. Vancomycin production in batch and continuous culture. *Biotechnology and Bioengineering* **49**(4):412-420.

Mendez-Perez D, Begemann MB, Pfleger BF. 2011. Modular Synthase-Encoding Gene Involved in alpha-Olefin Biosynthesis in *Synechococcus sp* Strain PCC 7002. *Applied and Environmental Microbiology* **77**(12):4264-4267.

Monod J. 1949. The Growth of Bacterial Cultures. *Annual Review of Microbiology* **3**(1):371-394.

Nawabi P, Bauer S, Kyrpides N, Lykidis A. 2011. Engineering *Escherichia coli* for biodiesel production utilizing a bacterial fatty acid methyltransferase. *Applied and Environmental Microbiology* **77**(22):8052-8061.

Neidhardt FC, Bloch PL, Smith DF. 1974. Culture medium for enterobacteria. *Journal of Bacteriology* **119**(3):736-747.

Nyholm N. 1976. Mathematical-model for microbial-growth under limitation by conservative substrates. *Biotechnology and Bioengineering* **18**(8):1043-1056.

Paalme T, Kahru A, Elken R, Vanatalu K, Tiisma K, Vilu R. 1995. The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat). *Journal of Microbiological Methods* **24**(2):145-153.

Peralta-Yahya PP, Zhang F, del Cardayre SB, Keasling JD. 2012. Microbial engineering for the production of advanced biofuels. *Nature* **488**(7411):320-328.

Peterson CN, Mandel MJ, Silhavy TJ. 2005. *Escherichia coli* starvation diets: Essential nutrients weigh in distinctly. *Journal of Bacteriology* **187**(22):7549-7553.

Ranganathan S, Tee TW, Chowdhury A, Zomorrodi AR, Yoon JM, Fu Y, Shanks JV, Maranas CD. 2012. An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering* **14**(6):687-704.

Ratledge C, Wynn JP. 2002. The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms. *Advances in Applied Microbiology* **51**:1-51.

Rude MA, Baron TS, Brubaker S, Alibhai M, Del Cardayre SB, Schirmer A. 2011. Terminal Olefin (1-Alkene) Biosynthesis by a Novel P450 Fatty Acid Decarboxylase from *Jeotgalicoccus* Species. *Applied and Environmental Microbiology* **77**(5):1718-1727.

Russell JB. 2007. The energy spilling reactions of bacteria and other organisms. *Journal of Molecular Microbiology and Biotechnology* **13**(1-3):1-11.

Schirmer A, Rude MA, Li XZ, Popova E, del Cardayre SB. 2010. Microbial Biosynthesis of Alkanes. Science **329**(5991):559-562.

Sharfstein ST, VanDien SJ, Keasling JD. 1996. Modulation of the phosphate-starvation response in *Escherichia coli* by genetic manipulation of the polyphosphate pathways. *Biotechnology and Bioengineering* **51**(4):434-438.

Sonderegger M, Schumperli M, Sauer U. 2005. Selection of quiescent *Escherichia coli* with high metabolic activity. *Metabolic Engineering* **7**(1):4-9.

Steen EJ, Kang YS, Bokinsky G, Hu ZH, Schirmer A, McClure A, del Cardayre SB, Keasling JD. 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**(7280):559-U182.

Tempest DW, Neijssel OM. 1984. The status of YATP and maintenance energy as a biologically interpretable phenomena. *Annual Review of Microbiology* **38**:459-486.

Wanner BL. 1993. Gene regulation by phosphate in enteric bacteria. *Journal of Cellular Biochemistry* **51**(1):47-54.

Wanner U, Egli T. 1990. Dynamics of microbial-growth and cell composition in batch culture. *FEMS Microbiology Reviews* **75**(1):19-44.

Youngquist JT, Lennen RM, Ranatunga DR, Bothfeld WH, Marner WD, 2nd, Pfleger BF. 2012. Kinetic modeling of free fatty acid production in *Escherichia coli* based on continuous cultivation of a plasmid free strain. *Biotechnology and Bioengineering* **109**(6):1518-1527.

Zeng AP, Deckwer WD. 1995. A kinetic model for substrate and energy consumption of microbial growth under substrate sufficient conditions. *Biotechnology Progress* **11**(1):71-79.

Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling J. 2012. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic Engineering* **14**(6):653-660.

Zheng Y, Li L, Liu Q, Qin W, Yang J, Cao Y, Jiang X, Zhao G, Xian M. 2012. Boosting the free fatty acid synthesis of *Escherichia coli* by expression of a cytosolic *Acinetobacter baylyi* thioesterase. *Biotechnology for Biofuels* **5**.

Chapter 5: Production of medium chain length fatty alcohols in *Escherichia coli*[3]

5.1. Introduction

The major goal of this thesis is to address the challenge of finding renewable alternatives to petrochemicals via free fatty acid (FFA) derived products. Chapters 3 and 4 have addressed the production of the FFA precursor. Chapter 5 will focus on the conversion of FFA to a more valuable end product. As bacterial hosts are well suited to produce precise chemical products and chemicals with good fuel properties need not be specific compounds, initial production of FFA derived chemicals in a commercial setting would likely focus on specific higher value products.

In recent years, significant effort has focused on producing hydrophobic compounds via fatty acid biosynthesis for use as liquid transportation fuels or commodity chemicals (Lennen and Pfleger, 2013). Aliphatic compounds such as fatty alcohols also have applications as detergents, emulsifiers, lubricants, and cosmetics. While fatty alcohols normally make up about 3-5 percent of the final formulation of these products, some such as solid anti-perspirants contain up to 25% fatty alcohols (Mudge et al., 2008). As of 2006, over 1.3 million tons of fatty alcohols were used worldwide each year (Mudge et al., 2008). As a whole, the industry represents over a 3 billion dollar market (Rupilius and Ahmad, 2006). Currently, fatty alcohols are produced either through processing natural fats and oils (oleochemicals) or from petrochemicals (e.g. crude oil, natural gas). In the oleochemical route, fatty acids or fatty acid methyl esters are released from triglycerides and hydrogenated to form fatty alcohols (Matheson, 1996). In one common petrochemical route, paraffins are separated from kerosene, then converted to olefins, before being converted to fatty alcohols. As both processes require either modifications to biodiesel or

[3] Portions of this chapter have been submitted for publication and are currently under review

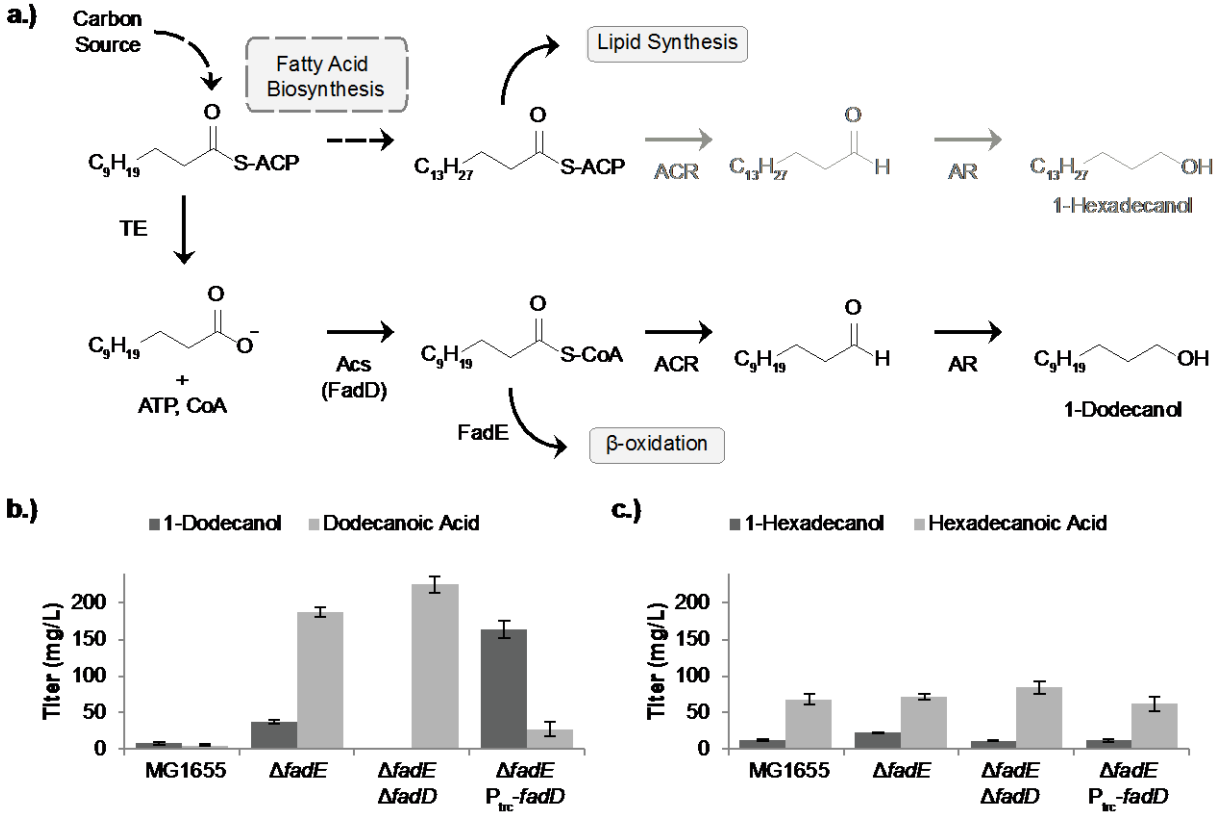


Figure 5.1: (Fatty Alcohol Biosynthesis) a.)Schematic of metabolic pathways that lead to fatty alcohols. Fatty acid biosynthesis generates acyl-acyl-carrier proteins (acyl-ACP) that are the substrates for lipid synthesis, thioesterases (TE) and acyl-CoA reductases (ACR). The fatty aldehydes produced by ACR can be reduced to primary alcohols by aldehyde reductases (AR). Expression of ACR/AR pairs leads to the formation of fatty alcohols that match the predominant acyl-ACP species (i.e. 16 carbons in *E. coli*). Alternatively, medium chain length alcohols can be produced by using an acyl-ACP thioesterase to produce a smaller fatty acid. Free fatty acids are then converted to acyl-CoA thioesters, by acyl-CoA synthetases (AS) and subsequently reduced by ACR to aldehydes and by AR to alcohols. b.) Conversion of exogenously fed dodecanoic acid to 1-dodecanol by *E. coli* strains harboring ptrc99a-MAACR (MAACR contains both ACR and AR activities). To produce 1-dodecanol, β-oxidation must be blocked (Δ*fadE*) and ACS (FadD) activity must be increased from native levels. c.) Expression of MAACR also results in equivalent production of 1-hexadecanol in each of strain.

petrochemical fuel stocks, microbial production of fatty alcohols from an unrelated carbon source is a promising alternative.

Fatty alcohols can be generated by microorganisms endogenously (Figure 5.1a) via reduction of fatty aldehydes that are made via reduction of acyl-thioesters (coenzyme A or acyl-carrier protein) (Reiser and Somerville, 1997). Alternatively, fatty acids have been shown to be

directly converted to fatty aldehydes via the action of a carboxylic acid reductase (Akhtar et al., 2013). Genes encoding long chain acyl-CoA reductase activity have been isolated from many organisms including bacteria (Reiser and Somerville, 1997), insects (Liénard et al., 2010), birds (Hellenbrand et al., 2011), mammals (Cheng and Russell, 2004), and protists (Teerawanichpan and Qiu, 2010). Many of these enzymes are used to synthesize fatty alcohols as precursors to wax esters. Three classes of reductases have been expressed in *E. coli* to produce C_{12-14} alcohols – reductases from soil bacteria (Reiser and Somerville, 1997; Steen et al., 2010), reductases from plants such as *Arabidopsis* or *Simmondsia* (Doan et al., 2009; Rowland and Domergue, 2012), and reductases found in marine bacteria (Willis et al., 2011; Hofvander et al., 2011). These classes differ in their ability to catalyze multiple reactions and in their substrate preference. Reductases similar to those found in *Acinetobacter* contain only the domain to catalyze conversion of acyl-thioesters to fatty aldehydes. Conversely, reductases from plants can catalyze both reductions, but generally do not have broad substrate specificity, preferring the dominant long acyl chains found in lipids. Reductases from marine bacteria catalyze both reductions and are active on a wide range of chain lengths.

While fatty acids have been produced with yields of greater than 0.2 g fatty acid per gram carbon source consumed (Dellomonaco et al., 2011; Zhang et al., 2012), the highest reported yields of fatty alcohols have been at least five fold lower. The work of Steen et al (Steen et al., 2010) demonstrated that fatty alcohols can be produced through overexpression of a thioesterase (*'tesA*), an acyl-CoA ligase (*fadD*), and an acyl-CoA reductase (*acr1*), with titers of around 60 mg/L fatty alcohol and yields of less than 0.005 g fatty alcohol / g carbon source. Further metabolic engineering and fermentation efforts have increased the titer to ~450 mg/L, but with no significant improvement in yield (Zheng et al., 2012). Alternative strategies have led to

slightly higher fatty alcohol yields from a defined carbon source. Expressing a carboxylic acid reductase and aldehyde reductase, titers have reached ~350 mg/L with a yield of 0.04 g fatty alcohol / g carbon source (Akhtar et al., 2013). Through the reversal of the fatty acid ß-oxidation, yields have been achieved between 0.04 and 0.055 g fatty alcohol / g carbon source consumed (Dellomonaco et al., 2011). However, no published strategy has yet to focus on the individual tailoring of expression for the multiple enzymes required in medium chain length fatty alcohol production.

In this chapter, we present a strategy to improve the yield of 1-dodecanol and 1-tetradecanol from an unrelated carbon source (e.g. glucose) in *E. coli*. We characterized three acyl-CoA ligase and three acyl-CoA reductase enzymes in conjunction with the thioesterase BTE from *Umbellularia californica* to determine the impact on fatty acid consumption and fatty alcohol synthesis. Additionally, we altered individual expression levels of the thioesterase, acyl-CoA ligase, and acyl-CoA reductase to best convert sugar substrates into fatty alcohol products. Our final strain overexpressed BTE, native FadD from *E. coli*, and the acyl-CoA reductase (MAACR) from *Marinobacter aquaeolei* VT8. In a bioreactor, a titer of over 1.65 g/L fatty alcohol (1.55 g/L C_{12-14} alcohol) and a yield of over 0.13 g fatty alcohol / g consumed glucose (0.12 g C_{12-14} fatty alcohol / g consumed glucose) was achieved.

5.2. Results

5.2.1. Establishing production of 1-dodecanol in E. coli

Fatty alcohol production was established in *E. coli* by heterologous expression of enzymes that catalyze reduction of acyl-thioesters and the resulting fatty aldehydes. In addition, two modifications of the *E. coli* MG1655 chromosome were necessary to produce 1-dodecenol

from exogenously fed dodecanoic acid (Figure 5.1b). First, β-oxidation was blocked to prevent consumption of the exogenously fed FFA. Small amounts of 1-dodecenol was produced when this objective was accomplished by deleting *fadE* (encoding enoyl-CoA reductase). Second, acyl-CoA synthetase (FadD) activity was essential for converting the exogenous lauric acid to the corresponding acyl-CoA thioester, a substrate for the heterologously expressed acyl-CoA/ACP reductase (i.e. MAACR from *Marinobacter aquaeolei* VT8 in Figure 6.1b). Unexpectedly, the Δ*fadE* strain only converted 18% of the lauric acid fed to the culture. The fractional conversion of lauric acid to 1-dodecenol was increased when the levels of FadD were elevated by replacing the native P_{fadD} with the strong, IPTG inducible P_{trc} promoter. In all strains, small amounts of 1-hexadecanol (15-20% of the endogenous hexadecanoic acid content) were produced, possibly indicating low activity of MAACR towards C_{16} acyl-ACPs.

5.2.2. Impact of various acyl-CoA synthetase on consumption of lauric acid

Given the dependence of 1-dodecenol conversion on acyl-CoA synthetase activity, the impact of three candidate synthetases on dodecenol production was examined by determining the rates of lauric acid consumption in *E. coli* MHS04 (Δ*fadD,* Δ*fadR*). Deletion of *fadR* removed repression of enzymes involved in β-oxidation (Dirusso et al., 1992) and increased the likelihood that acyl-CoA synthesis was the rate limiting step in lauric acid consumption. FadD and two alternative acyl-CoA synthetases were cloned into a medium copy plasmid and expressed from the P_{trc} promoter. The second acyl-CoA synthetase gene, *fadD6* from *M. Tuberculosis* was chosen because it has high activity toward C_{12} fatty acids and is soluble even when highly expressed (Arora et al., 2005). A third acyl-CoA synthetase, PP_0763 from *Pseudomonas putida*, was selected because of its ability to activate C_{12} fatty acids and enhance mcl-PHA production (Agnew et al., 2012; Wang et al., 2012). While each of the CoA synthetases conferred the ability

to consume 250 mg/L lauric acid within 12 hours (Figure 5.2), the strain expressing *fadD* was able to consume over 90% of the fed FFA within 8 hours. Each of the other ligases took at least 11 hours to reach the same mark. When the same acyl-CoA synthetases were expressed in a *fadR*+ strain (RL08), each consumed lauric acid at statistically equivalent rates (data not shown). This suggested that that acyl-CoA synthesis was not the rate limiting step in β-oxidation in these cells.

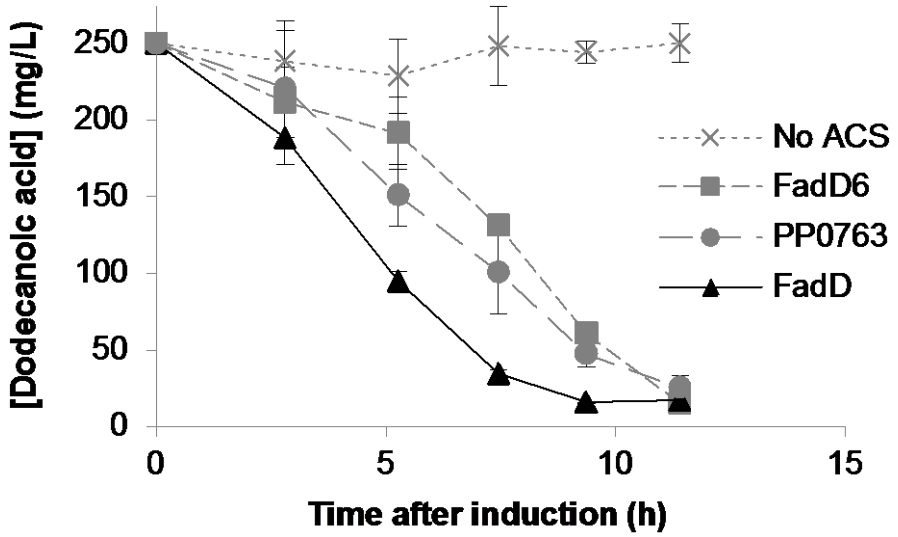


Figure 5.2: *E. coli* MHS04 (Δ*fadR* Δ*fadD*) harboring one of four acyl-CoA synthetase expression plasmids (medium copy, P_{trc}) was fed dodecanoic acid. The rate of consumption was fastest for the strain expressing FadD. The control strain carried the empty pACYCtrc plasmid. The error bars represent standard deviations from biological triplicate shake flask cultures.

5.2.3. Selection of acyl-CoA reductase

Once *fadD* was selected as the acyl-CoA ligase, it became necessary to determine the acyl-CoA reductase best suited for the conversion of C_{12} acyl-CoAs into fatty alcohols. Genes coding for three different types of acyl-CoA reductases (*acr1* from *Acinetobacter calcoaceticus* (Reiser and Somerville, 1997), *far6* from *Arabidopsis thaliana* (Doan et al. 2009), and MAACR from *Marinobacter aquaeolei* VT8 (Willis et al., 2011)) were tested to see which allowed for the

highest conversion of FFAs to fatty alcohols. Heterologous expression of a codon-optimized variant of *far6* failed to produce 1-dodecenol when cultures were fed dodecanoic acid (data not shown). Conversely, heterologous expression of both *acr1* and MAACR resulted in conversion of exogenous dodecanoic acid to 1-dodecanol. In these experiments, acyl-CoA reductases were expressed from medium copy plasmids harboring the IPTG inducible P_{trc} promoter in strain MHS01 (*ΔfadE* Φ[P_{trc}-*fadD*]). MAACR facilitated the fastest conversion of dodecanoic acid to 1-dodecanol (Figure 5.3a), with 80% of the initial fed dodecanoic acid had been converted to dodecanol within 12 hours after induction. One advantage of MAACR is its ability to also reduce dodecanaldehyde, by-passing endogenous aldehyde reductase activity and minimizing production of potentially toxic intermediates. However, here the dominance of MAACR appears to come from its higher activity on dodecanoic acid (Figure 5.3b).

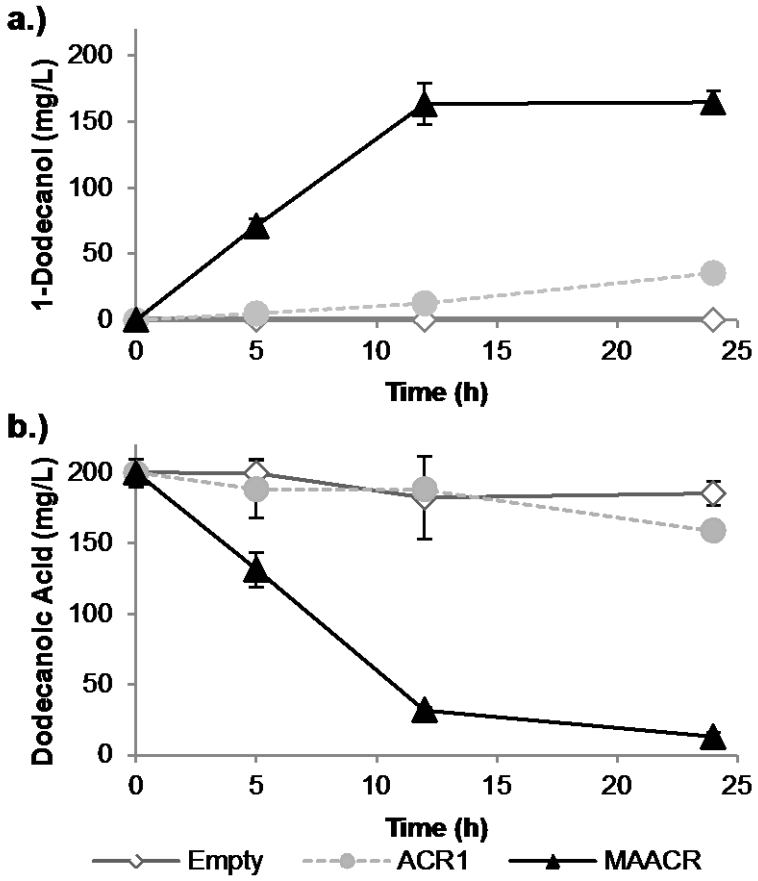


Figure 5.3: Comparison of Acyl-CoA Reductases. a.) The titer of 1-dodecanol produced by *E. coli* MHS01 (*ΔfadE* Φ[P_{trc}-*fadD*]) when dodecanoic acid Each strain harbored one of three plasmids – pTRC99A, pTRC99A-ACR1, or pTRC99A-MAACR. The error bars represent standard deviations from biological triplicate shake flask cultures.

5.2.4. Balancing reduction with endogenous fatty acid production

In order to use unrelated feedstocks, the medium chain length thioesterase (BTE) from *Umbellularia californica* (Voelker and Davies, 1994) was heterologously expressed in *E. coli* to endogenously produce C_{12} and C_{14} FFAs for subsequent conversion to the corresponding alcohols. A family of FFA producing strains were constructed by inserting a DNA cassette containing BTE under the control of the IPTG inducible P_{trc} promoter into various genomic loci (*fadE*, *fadAB*, and *ackA-pta*). Increasing BTE copy number (up to 3 copies) has been shown to increase FFA titers (Chapter 3). In effort to balance the expression of the downstream reductive

activities with fatty acid production, the acyl-CoA reductase from *M. aquaeolei*, MAACR, was cloned onto a series of plasmids (origins of replication: pBBr1, pACYC, pBR322, and pUC) that are maintained at different copy numbers (copy number: 4, 15, 25, and 100 respectively) (Lennen et al., 2010). Each MAACR plasmid was expressed in either MHS03 (1x BTE) or TY30 (2x BTE) to identify the optimal level of gene expression for each activity. Each strain contained elevated acyl-CoA synthetase activity in the form of a P_{trc}-*fadD* chromosomal cassette.

Strains expressing MAACR from the low copy number pBBr1 origin plasmid produced the most fatty alcohols (Figure 5.4), while strains with the high copy number pUC origin plasmid produced the least. Additionally, the strains containing the pUC origin plasmid displayed significantly impaired growth compared to the other strains (data not shown), suggesting a high metabolic burden associated with over-expression of MAACR. Surprisingly, there was small difference in final fatty alcohol titer between the same plasmid expressed in either the MHS03 or TY30 strain.

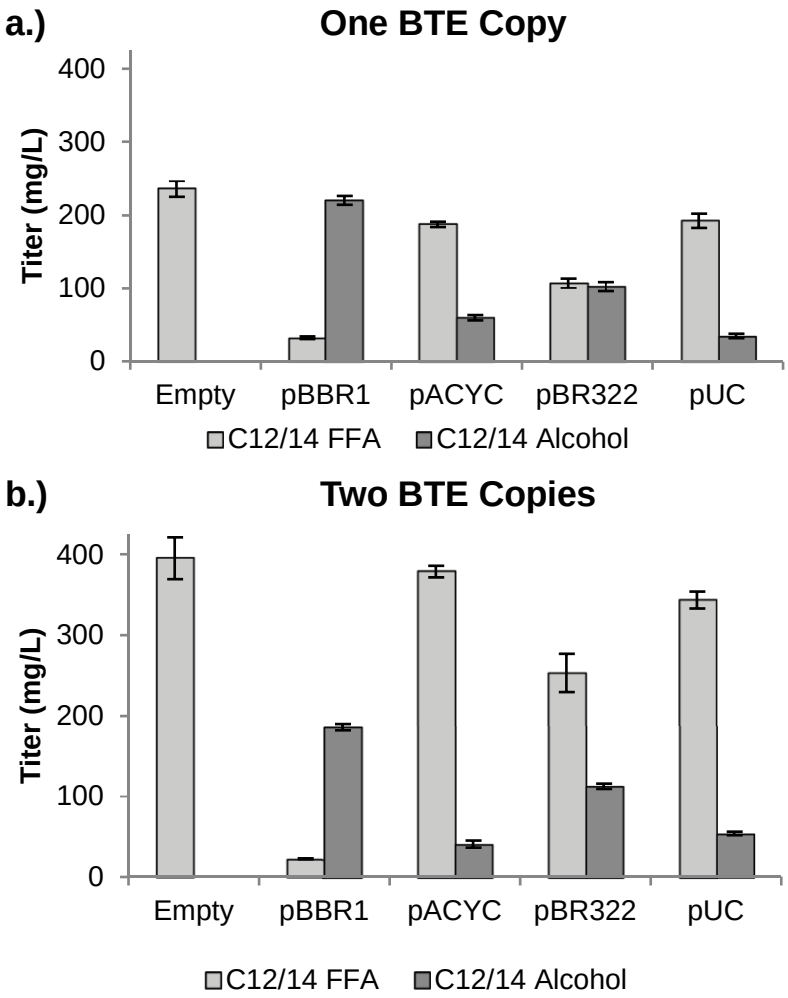


Figure 5.4: Optimization of MAACR expression. The combined titers of 1-dodecanol and 1-tetradecanol as well as residual dodecanoic and tetradecanoic acid were measured for a.) *E. coli* MHS03 (*ΔfadE::trcBTE* Φ(P_{Trc}-*fadD*)) and b.) *E. coli* TY30 (*ΔfadE::trcBTE ΔfadAB::trcBTE* Φ(P_{Trc}-*fadD*)) harboring MAACR on four different copy number vectors. The Empty plasmid was ptrc99a. The plasmid copy number increases from left to right. The highest alcohol titers were achieved when MAACR was expressed on a low copy vector, independent of the number of copies of BTE. Error bars represent standard deviation of biological triplicate shake flask cultures.

To optimize the level of acyl-CoA synthetase activity, a family of strains was constructed to vary *fadD* expression. *E. coli* TY33 (Δ*fadD*), TY31 (native *fadD*), and TY30 (ΦP_{trc}-*fadD*), were transformed with pBTRK-MAACR and either pACYCtrc or pACYC-fadD. Expression of *fadD* was quantified by qPCR using RNA samples isolated at an OD_{600} of 0.8. The *fadD* promoter replacement resulted in the maximum production of both 1-dodecanol and 1-

tetradecanol (Figure 5.5). The promoter replacement increased *fadD* levels by 46 ± 10 fold while expression from a medium copy plasmid increased expression by 1000 ± 200 fold relative to *fadD* under its native promoter on the chromosome.

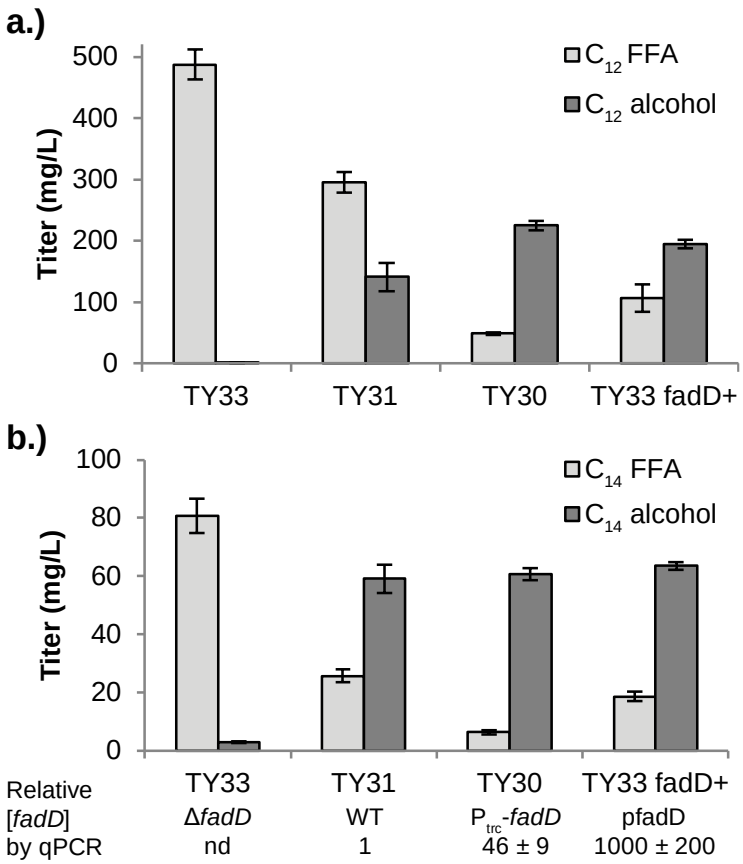


Figure 5.5. Optimization of FadD expression. a.) Dodeccanol and b.) tetradecanol production as a function of the relative expression level of acyl-CoA synthetase (*fadD*) in *E. coli* strains harboring pBTRK-MAACR compared to native expression (TY31). Error bars represent standard deviation from biological triplicate shake flask cultures.

5.2.5. Endogenous production of dodecanol and tetradecanol from glucose

To determine fatty alcohol yield, strains MHS03 (1x BTE), TY30 (2x BTE), and TY34 (3x BTE) each containing pBTRK-MAACR were cultivated in MOPS minimal media using glucose as a carbon source in controlled bioreactors. To simulate a fed batch, a bolus of 2 g glucose was added on five separate occasions. After 120 hours the final fatty alcohol titer was

280, 470, and 1185 mg/L for the MHS03, TY30, and TY34 versions, respectively, with over 90% coming from 1-dodecanol and 1-tetradecanol (Figure 5.6). Based on the amount of glucose consumed by these cultures, the resulting yields were 0.031, 0.040, and 0.097 g fatty alcohol per g glucose consumed for the MHS03, TY30, and TY34 strains expressing MAACR, respectively.

In each experiment, a slight white sludgy material (assumed to be fatty alcohol) was deposited on the bioreactor wall. This material prevented an accurate timecourse of fatty alcohol production from being taken. To by-pass the problem, 20 mL of dodecane was added to the fermentation 6-h after induction. Three replicates of TY34 containing pBTRK-MAACR were run in controlled bioreactor fermentations with the dodecane emulsion. The addition of dodecane allowed for an accurate timecourse of fatty alcohol production to be taken and increased the final fatty alcohol titers to 1.65 g/L (0.134 g alcohol/g glucose consumed, Figure 5.6a). Samples were taken to monitor biomass, CO_2, acetate, FFAs, and other excreted metabolites (Figures 5.6b, 5.6c, 5.6d, and 5.6f). Analysis of these samples led to a carbon balance accounting for 86% of the carbon, with elevated levels of acetate and CO_2 being produced.

Separate samples for the supernatant and cell pellet were taken at the last time point of each bioreactor run to determine if the addition of dodecane allowed for an increased transport of fatty alcohols to the extracellular medium. Less than 5% of the FFAs and fatty alcohols were found in the cell pellet fraction from cultures grown in co-culture (Figure 5.6c). In contrast approximately 60% of the fatty alcohol species were found in the cell pellet fraction.

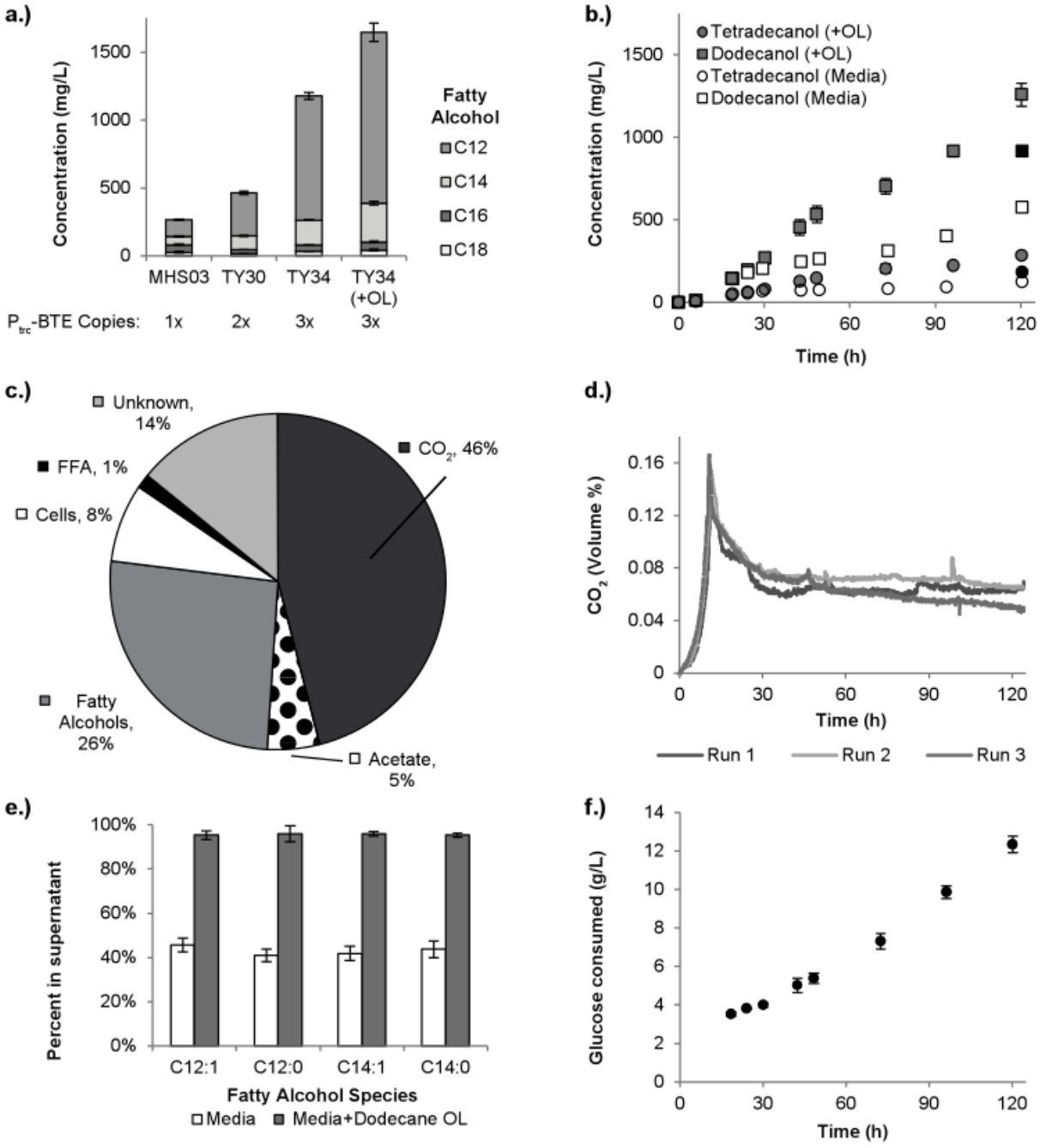


Figure 5.6: Production of fatty alcohols in co-culture a.)Final observed fatty alcohol titer breakdown in strains MHS03 (1 copy of BTE), TY30 (2 copies of BTE), and TY34 (3 copies of BTE) harboring pBTRK-MAACR after being run in a stirred bioreactor. OL refers to the presence of a dodecane overlayer added during fermentations b.) The titer of fatty alcohol produced from *E. coli* TY34 pBTRK-MAACR fed-batch cultivations is plotted as a function of time. c.) The relative quantity of metabolic products, as percentage of fed carbon, shows large percentages going to CO_2, acetate, biomass, and fatty alcohols. d.) The off-gas [CO_2] suggests that a metabolic steady state is achieved ~30 h post induction. e.) Co-culturing *E. coli* with dodecane increases the amount of fatty alcohol found outside the cell. f.) Glucose consumption was nearly linear over the fed-batch portion of the culture.

5.3. Discussion

5.3.1. Selection of acyl-CoA reductases

The selection of an acyl-CoA reductase influences fatty alcohol production in multiple ways. Biosynthesis of specific chain length fatty alcohols requires cleavage and reduction of the corresponding acyl-thioester (-CoA or -ACP) to a fatty aldehyde. The distribution of chain lengths for most fatty alcohol producers matches the strain's fatty acid profile, indicating that the reductase activity/affinity is not strong (or at least weaker than that of fatty acid elongation) for shorter chain substrates. Conversely, thioesterases are known to have high activity on a wide range of acyl-thioester chain lengths depending on the specific enzyme. The disadvantage of utilizing thioesterases for fatty alcohol production is the need to reactivate the acyl-chain for reduction. If acyl-ACP reductases could be engineered to have stronger activities towards specific acyl-ACPs, higher yields could be achieved. Similar efforts to engineer chain length specificity in thioesterases has been reported in the patent literature (Yuan et al., 1999) and could guide acyl-ACP reductase engineering.

Here, expression of the dual-activity acyl-CoA reductase, MAACR, led to the highest fatty alcohol productivity. It is likely that substrate channeling between the acyl-ACP and fatty aldehyde reduction domains prevented release of the reactive, potentially toxic, aldehyde intermediate. If novel, high-activity acyl-CoA reductases are identified, fusion (or incorporation into a complex via a protein scaffold) of aldehyde reductases could have similar benefits (Dueber et al. 2009). Alternatively, separate enzymes could be targeted to microcompartments to sequester the aldehydes from the cytoplasm. Many bacteria use this strategy to avoid the toxicity of aldehyde intermediates and/or increase the local concentration of substrates when enzymes have weak activity (Sampson and Bobik 2008; Frank et al., 2013). While this strategy is

promising, the microcomparments would need to be engineered to transport the substrates (e.g. acyl-ACP) and products.

5.3.2. Balancing expression of the acyl-CoA synthetase and acyl-CoA reductase

One of the metabolic engineering objectives in this study was to tailor the expression levels of the acyl-CoA synthetase and acyl-CoA reductase to balance the overall conversion between fatty acid and fatty alcohol. The optimal levels of acyl-CoA synthetase and acyl-CoA reductase that maintain balanced activity could be determined from knowledge of the *in vivo* kinetic parameters (k_{cat}, K_m), if known. Based on *in vitro* experiments, the specific activity and K_m for *fadD* conversion of lauric acid to lauryl-CoA are 2,630 nmol/min/mg protein and 1.6 µM, respectively (Kamedas and Nunn, 1981). For the conversion of lauryl-CoA to the aldehyde intermediate, the specific activity of MAACR *in vitro* is 34 nmol/min/mg enzyme and the K_m is 4 µM (the specific activity for the second step, aldehyde to alcohol, is two orders of magnitude higher) (Willis et al., 2011). These values suggest that the k_{cat}/K_m ratio is about 100 fold higher for the acyl-CoA synthetase step than the reductase step. Optimal production of fatty alcohols occurred with MAACR on a low copy (~5:1 ratio to genomic DNA) plasmid using the same promoter as the chromosomal P_{trc}-*fadD* cassette (Lennen et al., 2010). Overexpression of MAACR decreased fatty alcohol titer, placing an upper limit on acyl-CoA reductase activity. This observation could be attributed to metabolic burden of protein overexpression or improper folding of MAACR expressed at a high level. Prior studies have shown that soluble overexpression of MAACR is problematic without addition of an N-terminal maltose binding protein (Willis et al., 2011), which was used in this study. Figure 5.5 shows that a 45 fold decrease in *fadD* transcript levels, between that controlled by P_{trc} promoter and the native promoter, resulted in only a 50% decrease in fatty alcohol titer. This result suggests that further

strain optimization could be achieved by decreasing *fadD* expression. This strategy is consistent with the optimal ratio of MAACR and FadD levels predicted by their relative *in vitro* kinetics.

5.3.3. Improving fatty alcohol yield

Implementation of the metabolic engineering strategy described above generated a strain that produced the highest reported yield (0.134 g/g) and titer of fatty alcohols (1.65 g/L fatty alcohols, with 77% and 17% being C_{12} and C_{14} species, respectively) from an unrelated carbon source. Previous studies that leveraged native fatty acid biosynthesis pathways produced fatty alcohol titers of up to ~450 mg/L C_{12-14} fatty alcohols with yields less than 0.01 g fatty alcohol / g carbon source (Steen et al., 2010; Zheng et al., 2012). Alternative pathways in *E. coli* have yielded up to ~350 mg/L fatty alcohols and yields up to 0.05 g fatty alcohol / g carbon source (Akhtar et al., 2013; Dellomonaco et al., 2011). Based on theoretical yields, *E. coli* is capable of producing 0.32 g 1-dodecanol per g glucose fed. As current yields are much less than theoretical, further optimization and metabolic engineering efforts are needed to improve yield. However, current yields of combined C_{12-14} fatty acids and fatty alcohols are similar to that seen in a corresponding FFA producing strain (Chapter 4), indicating that efforts should focus on redirecting carbon flux toward fatty acid biosynthesis (Lennen and Pfleger, 2012) rather than the conversion to fatty alcohol. The carbon balance on the bioreactor (Figure 5.6) indicates that a significant amount of fed carbon is going to carbon dioxide production. Therefore, decreasing flux to CO_2 production could lead to improved fatty alcohol titers. Similarly, a small percentage of carbon flux ended in the secretion of acetate (Figure 5.6). Given that TY34 is *ΔackAΔpta*, it is likely that the observed acetate was generated by the pyruvate oxidation pathway that concurrently generates proton motive force, as it is coupled to the electron transport chain (Abdel-Hamid et al., 2001). Deletion of *poxB* would eliminate acetate production through this

pathway and potentially increase fatty alcohol yields (Zha et al., 2009; Xu et al., 2013). Other studies have successfully achieved higher yields of fatty acids by overexpressing genes *fabZ* (Ranganathan et al., 2012) or *fadR* (Zhang et al., 2012). Similar manipulations employed in conjunction with this metabolic engineering strategy could lead to fatty alcohol production at yields closer to the theoretical limit.

The high yield and titer reported above was achieved by cultivating strain *E. coli* TY34 pBTRK-MAACR in a 1-L working volume using a fed-batch strategy. One interesting observation from this experiment was the consistent production of fatty alcohols and consumption of glucose (Figure 5.6) during a prolonged stationary phase (>96 hours). Such a result may be expected due to the unique qualities of phosphate starvation, in which metabolic activity in the cell remains high despite no further cell growth (Ballesteros et al., 2001). During stationary phase a specific productivity of 0.016 g fatty alcohol/gDCW/h was observed and a glucose consumption rate of 0.11 g glucose/gDCW/h. A better understanding of metabolism and regulation under these conditions will help guide efforts to maintain the stability of producing strains and maximize the time strains can spend in the production phase.

5.4. Conclusions

E. coli was engineered to produce 1-dodecenol and 1-tetradecanol from an unrelated carbon source. Cultivation of the strain in a bioreactor with 10% dodecane achieved the highest reported titer (1.65 g/L) and yield (0.134 g fatty alcohol / g glucose) from a minimal glucose based media to date. The key steps to optimize this fatty alcohol producing strain were selection of FadD from *E. coli* as the acyl-CoA synthetase and MAACR from *M. aquaeolei* VT8 as the acyl-CoA reductase. In addition, overexpression of these two enzymes was found to be detrimental to fatty alcohol productivity. The optimal expression levels were found by replacing

the native P_{fadD} promoter with a stronger inducible promoter (P_{trc}) and expressing MAACR from a low copy vector. The yields observed were nearly equivalent to the yield of FFAs in past work (Chapter 4), suggesting that the strain may be capable of higher yields if FFA production could be increased. Future work will aim to reduce the flux to other products that compete with flux to fatty alcohols.

5.5. Materials and Methods

5.5.1. Bacterial strains and chromosome engineering

All bacterial strains used in this study are listed in Appendix 3. Single gene deletions were transferred P1 transduction of phage lysates from the collection of single gene knockouts from the National BioResource Project (NIG, Japan) (Baba et al., 2006). Chromosomal integration of a BTE expression cassette (acyl-ACP thioesterase from *Umbellularia californica* under the control of the IPTG inducible P_{trc} promoter) was performed as described previously (Chapter 3). All deletions and insertions were verified by colony PCR.

5.5.2. Reagents and media

Enzymes were purchased from New England Biolabs (Ipswich, MA). Nucleic acid purification materials were purchased from Qiagen (Venlo, Netherlands), Promega (Madison, WI), or Thermo Scientific (Waltham, MA). Chemicals were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Hampton, NH) unless otherwise specified. Oligonucleotides (sequences are listed in Appendix 4) were purchased from Integrated DNA Technologies (Coralville, IA). For all growth experiments, single colonies obtained from freezer stocks were used to inoculate 5 mL LB starter cultures grown overnight prior to the inoculation of experimental cultures. All shake flask growth experiments were performed at 30°C in a rotary

shaker (250 rpm). Cultures were supplemented with appropriate antibiotics (100 µg mL^{-1} ampicillin and/or 50 µg mL^{-1} kanamycin and/or 34 µg mL^{-1} chloramphenicol) where necessary.

5.5.3. Plasmid construction

All plasmids used in this study are listed in Appendix 3. *E. coli* acyl-CoA synthetase *fadD* was amplified by PCR from genomic DNA isolated from *E. coli* MG1655. Codon optimized versions of the acyl-CoA synthetase *fadD6*, acyl-CoA reductase *acr1*, and acyl-CoA reductase *far6* were custom synthesized by Life Technologies (Carlsbad, CA). *P. putida* KT2440 genomic DNA was used as a template to PCR amplify PP_0763. MAACR was amplified by PCR from a plasmid containing the *Marinobacter aquaeolei* acyl-CoA reductase generously donated by Dr. Brett Barney (University of Minnesota). Base plasmids pBTRKtrc, pACYCtrc, and pUCtrc were constructed by generating PCR products using primers 1 and 2 to amplify the antibiotic resistance marker and origin of replication from plasmids pBAD35, pBAD33, and pBAD34 (Lennen et al., 2010) respectively. Primers 3 and 4 were then used to amplify the multi-cloning site, P_{trc} promoter, and $lacI^q$ region from pTrc99a. The PCR products were combined using the Gibson assembly method (Gibson et al., 2009). To construct each of the individual expression plasmids, pBTRKtrc, pACYCtrc, pTrc99a, and pUCtrc were amplified with primers 13 and 14, MAACR was amplified with primers 5 and 6. These PCR products were then combined using the Gibson assembly method to generate the constructs listed in Appendix 3. For the codon optimized genes *acr1* and *far6*, pTrc99a and the vectors containing the genes were digested with *Kpn I* and *Hind III*, ligated with an analogous digest of pTrc99a using T4 DNA ligase. The same procedure was used with the codon optimized *fadD6* and pACYCtrc. For the other CoA synthetases, the PCR products from *fadD* (7 and 8) and PP_0763 (11 and 12) were

digested with *Kpn I* and *Xba I*, and ligated with digested pACYCtrc. All constructs were confirmed by DNA sequencing.

5.5.4. Culturing conditions

For experiments where dodecanoic acid was supplied exogenously (Figures 5.1, 5.2, and 5.3), each strain was cultured in 50 mL LB starting with an inoculum at optical density (OD_{600}) of 0.02. At OD_{600} 0.2, cultures were induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG) and supplemented with either 40 (for alcohol production studies) or 50 μL (for dodecanoic acid consumption studies) of a 250 mg/mL solution of dodecanoic acid in ethanol (initial [dodecanoic acid] = 200 or 250 mg/L). After induction, cultures were incubated at 30°C with shaking and 2.5 mL culture samples were taken at either 2.5, 7, 9, and 11 or 4, 8, 12, and 24 hours after induction for dodecanoic acid consumption and alcohol production studies, respectively. Culture samples were processed for FAME analysis as described previously (Agnew et al. 2012).

For fatty alcohol production experiments (Figures 5.4 and 5.5), each strain was inoculated to an OD_{600} of 0.02 in 50 mL LB + 0.4% glycerol and induced with 1 mM IPTG at an OD_{600} of 0.2. Following induction, cultures were incubated with shaking at 30°C for 48 hours. Culture samples of 2.5 mL were taken at 24 and 48 hours for fatty alcohol and FAME analysis. To determine the fraction of fatty alcohol associated with cells, an additional 10 mL sample from the 48 hour timepoint was centrifuged at 4000 x g for 10 min and the resulting cell pellet was resuspended to 10 mL in 1x PBS. After repeating the process, 2.5 mL of the resuspended cell pellet was taken for fatty alcohol and FAME analysis.

Bioreactor experiments were performed in a 3-L stirred bioreactor (Applikon Biotechnology, Inc., Schiedam, Netherlands), using a 1 L working volume. Temperature was

maintained at 30°C using a heat blanket (Applikon, model number M3414) and cooling water. Reactor temperature, pH and dissolved oxygen (DO_2) were monitored using specific probes (Applikon). Carbon dioxide and oxygen off-gas levels were monitored using a Blusens BlueInOne Ferm (Blusens, Herten, Germany). Reactor pH was maintained at 7.00 ± 0.01 by the addition of 10% (v/v) NH_4OH or 1 M HCl solutions. Agitation was provided by a single impeller with the stir speed set between 240-320 rpm. Stirrer speed was varied to ensure the DO_2 content did not decrease below 40% saturation in order to maintain an aerobic environment (Becker et al., 1997; Tseng et al.,1996). The air inflow rate was maintained at 1.0 L/min.

Bioreactor experiments were performed using a phosphate limited MOPS minimal media recipe (Chapter 4). Cultures were inoculated to an OD_{600} of 0.04 using a culture of *E. coli* MHS03, TY30, or TY34 containing pBTRK-MAACR grown to an OD600 >2 in MOPS minimal media (Neidhardt et al., 1974) supplemented with 0.7% glucose overnight. Bioreactor starting media was MOPS minimal media supplemented with 0.7% glucose, 0.276 mM potassium sulfate, and 9.5 mM ammonium chloride but containing only containing 370 µM K_2HPO_4. Cultures were induced with 1 mM IPTG at OD_{600} 0.2. Each experiment was performed using a discontinuous fed-batch where a bolus of 2 g glucose (10 mL of a 20% (w/v) glucose solution) was added at 18, 24, 30, 42, and 48 hours post-induction. In three experiments, 20 mL dodecane was added to the culture 6 hours after induction to provide a sink for fatty alcohols. For all experiments, CO_2 off-gas levels and pH were measured continuously and culture samples (10 mL) were taken periodically prior to glucose additions to determine OD_{600}, [glucose], [acetate], [fatty alcohol], and [fatty acid].

5.5.5. Fatty acid and fatty alcohol extraction and characterization

FAME analysis was performed on 2.5 mL of culture, supernatant, or resuspended washed cell pellet as described previously (Lennen et al., 2010). Analysis of fatty alcohols followed the same procedure except 20 µL of 10 mg/mL pentadecanol in ethanol was added to the chloroform methanol mix as an internal standard in addition to the fatty acid internal standards.

5.5.6. Quantitative-PCR

For RNA samples, 1 mL of culture at an OD_{600} of 0.8 was centrifuged at 8000x g for 3 minutes at 4ºC. The supernatant was quickly removed and then the cell pellet was snap frozen in a dry ice ethanol bath for 5 minutes before storing the samples at -80ºC until further processing. RNA was isolated using an RNeasy mini kit (QIAgen). Residual DNA was digested using the Ambion DNA-*free*™ Kit (Applied Biosystems). The corresponding cDNA was synthesized using the GoScript™ Reverse Transcription System (Promega) following manufacturer's instructions. To run the qPCR, the Maxima SYBR Green/Fluorescein qPCR Master Mix (Thermo Scientific) was used. Primers were designed for amplifying both a 100 bp region of *fadD* and a 100 bp region of *rrsA* to act as a reference for normalization of samples (Kobayashi et al., 2006).

5.6. References

Abdel-Hamid a M, Attwood MM, Guest JR. 2001. Pyruvate oxidase contributes to the aerobic growth efficiency of *Escherichia coli*. *Microbiology* **147**:1483–1498.

Agnew DE, Stevermer AK, Youngquist JT, Pfleger BF. 2012. Engineering *Escherichia coli* for production of C12-C14 polyhydroxyalkanoate from glucose. *Metabolic Engineering* **14**: 705–713.

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

Amann E, Ochs B, Abel K, 1988. Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**:301–315.

Arora P, Vats A, Saxena P, Mohanty D, Gokhale RS, 2005. Promiscuous fatty acyl CoA ligases produce acyl-CoA and acyl-SNAC precursors for polyketide biosynthesis. *Journal of the American Chemical Society* **127**:9388–9389.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko K a, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**:1–11.

Ballesteros M, Fredriksson A, Henriksson J, Nystrom T. 2001. Bacterial senescence: protein oxidation in non-proliferating cells is dictated by the accuracy of the ribosomes. *EMBO Journal* **20**(18):5280-5289.

Becker S, Vlad D, Schuster S, Pfeiffer P, Unden G. 1997. Regulatory O-2 tensions for the synthesis of fermentation products in *Escherichia coli* and relation to aerobic respiration. *Archives of Microbiology* **168**(4):290-296.

Cheng JB, Russell DW. 2004. Mammalian wax biosynthesis. I. Identification of two fatty acyl-Coenzyme A reductases with different substrate specificities and tissue distributions. The *Journal of Biological Chemistry* **279**:37789–37797.

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

Dellomonaco C, Fava F, Gonzalez R. 2010. The path to next generation biofuels: successes and challenges in the era of synthetic biology. *Microbial Cell Factories* **9**:3.

Dirusso CC, Heimert TL, Metzger, AK. 1992. Characterization of FadR, a Global Transcriptional Regulator of Fatty Acid Metabolism in *Escherichia coli*. *The Journal of Biological Chemistry* **267**:8685–8691.

Doan TTP, Carlsson AS, Hamberg M, Bülow L, Stymne S, Olsson P. 2009. Functional expression of five *Arabidopsis* fatty acyl-CoA reductase genes in *Escherichia coli*. *Journal of Plant Physiology* **166**:787–796.

Dueber JE, Wu GC, Malmirchegini GR, Moon TS, Petzold CJ, Ullal AV, Prather KLJ, Keasling JD. 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology* **27**:753–759.

Frank S, Lawrence AD, Prentice MB, Warren MJ. 2013. Bacterial microcompartments moving into a synthetic biological world. *Journal of Biotechnology* **163**:273–279.

Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, Smith HO. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**:343–345.

Hellenbrand J, Biester EM, Gruber J, Hamberg M, Frentzen M. 2011. Fatty acyl-CoA reductases of birds. *BMC Biochemistry* **12**:64.

Hofvander P, Doan TTP, Hamberg M. 2011. A prokaryotic acyl-CoA reductase performing reduction of fatty acyl-CoA to fatty alcohol. *FEBS Letters* **585**:3538–43.

Kamedas K, Nunn WD. 1981. Purification and Characterization of Acyl Coenzyme A Synthetase from *Escherichia coli*. *The Journal of Biological Chemistry* **256**:5702–5707.

Keasling JD. 2012. Synthetic biology and the development of tools for metabolic engineering. *Metabolic Engineering* **14**:189–95.

Kobayashi A, Hirakawa H, Hirata T, Nishino K, Yamaguchi A. 2006. Growth phase-dependent expression of drug exporters in *Escherichia coli* and its contribution to drug tolerance. *Journal of Bacteriology* **188**:5693–703.

Lennen RM, Braden DJ, West RA, Dumesic JA, Pfleger BF. 2010. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnology and Bioengineering* **106**:193–202.

Lennen RM, Pfleger BF. 2012. Engineering *Escherichia coli* to synthesize free fatty acids. *Trends in Biotechnology* **30**:659–67.

Lennen RM, Pfleger BF. 2013. Microbial production of fatty acid-derived fuels and chemicals. *Current Opinion in Biotechnology* **21**:1–10.

Liénard MA, Hagström AK, Lassance JM, Löfstedt C. 2010. Evolution of multicomponent pheromone signals in small ermine moths involves a single fatty-acyl reductase gene. *Proceedings of the National Academy of Sciences of the United States of America* **107**:10955–60.

Matheson KL. 1996. Surfactants raw materials: classification, synthesis, and uses. In: Spitz, L. (Ed.), Soaps and Detergents: A Theoretical and Practical Review, AOCS, Champaign, IL, pp. 288–303.

Mudge SM, Belanger SE, Nielsen AM. 2008. Fatty Alcohols: Anthropogenic and Natural Occurrence in the Environment. The Royal Society of Chemistry, Cambridge, UK.

Neidhardt FC, Bloch PL, Smith DF. 1974. Culture medium for enterobacteria. *Journal of Bacteriology* **119**(3):736-747.

Ranganathan S, Tee TW, Chowdhury A, Zomorrodi AR, Yoon JM, Fu Y, Shanks J V, Maranas CD. 2012. An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering* **14**:687–704.

Reiser S, Somerville C. 1997. 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. *Journal of Bacteriology* **179**:2969–2975.

Rowland O, Domergue F. 2012. Plant fatty acyl reductases: enzymes generating fatty alcohols for protective layers with potential for industrial applications. *Plant Science* **193-194**:28–38.

Rupilius W, Ahmad S. 2006. The Changing World of Oleochemicals. *Palm Oil Developments* **44**:15–28.

Sampson EM, Bobik TA. 2008. Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate. *Journal of Bacteriology* **190**:2966–71.

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

Teerawanichpan P, Qiu X. 2010. Fatty acyl-CoA reductase and wax synthase from *Euglena gracilis* in the biosynthesis of medium-chain wax esters. *Lipids* **45**:263–273.

Tseng C, Albrecht J, Gunsalus RP. 1996. Effect of Microaerophilic Cell Growth Conditions on Expression of the Aerobic (*cyoABCDE* and *cydAB*) and Anaerobic Pathway Genes in *Escherichia coli*. *Microbiology* **178**:1094–1098.

Voelker TA, Davies HM, 1994. Alteration of the specificity and regulation of Alteration of the Specificity and Regulation of Fatty Acid Synthesis of *Escherichia coli* by Expression of a Plant Medium- Chain Acyl-Acyl Carrier Protein Thioesterase. *Journal of Bacteriology* **176**(23):7320-7327.

Wang Q, Tappel RC, Zhu C, Nomura CT. 2012. Development of a new strategy for production of medium-chain-length polyhydroxyalkanoates by recombinant *Escherichia coli* via

inexpensive non-fatty acid feedstocks. *Applied and Environmental Microbiology* **78**:519–527.

Willis RM, Wahlen BD, Seefeldt LC, Barney BM. 2011. Characterization of a fatty acyl-CoA reductase from *Marinobacter aquaeolei* VT8: a bacterial enzyme catalyzing the reduction of fatty acyl-CoA to fatty alcohol. *Biochemistry* **50**:10550–8.

Xu P, Gu Q, Wang W, Wong L, Bower AGW, Collins CH, Koffas M a G. 2013. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications* **4**:1409.

Youngquist JT, Lennen RM, Ranatunga DR, Bothfeld WH, Marner WD, 2nd, Pfleger BF. 2012. Kinetic modeling of free fatty acid production in *Escherichia coli* based on continuous cultivation of a plasmid free strain. *Biotechnology and Bioengineering* **109**(6):1518-1527.

Youngquist JT, Rose JP, Pfleger BF. 2013. Free fatty acid production in *Escherichia coli* under phosphate-limited conditions. *Applied Microbiology and Biotechnology* **97**(11):5149-5159.

Yu DG, Ellis HM, Lee EC, Jenkins NA, Copeland NG, Court DL. 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **97**(11):5978-5983.

Yuan L, Dehesh K, Kridl J, Knauf V. 1999. Engineering plant thioesterases for altered substrate specificity. US5955329.

Zha W, Rubin-Pitel SB, Shao Z, Zhao H. 2009. Improving cellular malonyl-CoA level in *Escherichia coli* via metabolic engineering. *Metabolic Engineering* **11**:192–198.

Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD. 2012. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic Engineering* **14**:653–60.

Zheng YN, Li L, Liu Q, Yang J, Wang X, Liu W, Xu X, Liu H, Zhao G, Xian M. 2012. Optimization of fatty alcohol biosynthesis pathway for selectively enhanced production of C12/14 and C16/18 fatty alcohols in engineered *Escherichia coli*. *Microbial Cell Factories* **11**:65.

Chapter 6: Additional strategies and future directions aimed at increasing free fatty acid derived chemical production

6.1. Introduction

The previous chapters of the thesis have covered basic strategies to increase titer and yield of free fatty acids (FFAs) and FFA derived products in *E. coli*. Chapter 2 gave an overview of the many routes to different FFA-derived products and the types of reactor and media conditions used to optimize production. The work in Chapter 3 was used as the basis for a FFA production model in *E. coli* and was expanded to cover different nutrient limitation conditions in Chapter 4. Chapter 5 combined the knowledge of the FFA production model and genetic engineering strategies to increase titer and yield of fatty alcohols. Despite the improvement seen in Chapter 5 versus other published studies, yields remained below 50% of maximum theoretical values for both FFA and fatty alcohol production. Chapter 6 will focus on additional strategies that aimed to improve the yield of FFA and FFA-derived products from simple sugars.

To achieve the highest yields and titers, carbon flux to FFA biosynthesis must be maximized. Many approaches have been attempted to achieve the goal of increasing product yield on a given carbon source, including: transposon mutagenesis (Hoover et al., 2012), targeted gene knockouts (Ranganathan et al., 2012), altering transcriptional regulator levels (Zhang et al., 2012), and alternative pathway engineering (Martin et al., 2003). The overall challenge is to circumvent native regulation that controls lipid synthesis (Cronan, 2003; Heath and Rock, 1996) and balance the catalytic activity of each enzyme involved in fatty acid biosynthesis. As detailed in chapter 2, long-chain acyl-ACPs regulate fatty acid biosynthesis by inhibiting enzymes involved in the pathway. Attempts to increase FFA levels by increasing expression of enzymes that compete for acetyl-CoA (acetyl-CoA carboxylase) and twelve carbon acyl-ACPs (BTE)

have been unsuccessful in achieving significant increases in FFA titer (Hoover et al., 2011; Lennen et al., 2010). That leaves modification of alternative pathways or reactor conditions as strategies to increase production. As the latter was discussed in previous chapters, Chapter 6 will discuss the former.

Metabolic flux modeling and microarray analysis have been successful in increasing titers and production of other compounds though the identification of pathway modifications (Park et al., 2007; Ranganathan et al., 2012; Santos et al., 2012). In *E. coli*, gene expression differences between a FFA producer and non-producer identified a number of stress responses that were upregulated under FFA production (Lennen et al., 2011). However, those upregulated pathways may be necessary for cellular survival under the stresses presented by FFA production. The differences in gene expression under conditions tested in Chapters 3 and 4 provide a unique dataset comparing cells under similar FFA stresses. FFA producing phosphate limited cells have a higher maintenance energy than phosphate limited non producing cells or carbon limited FFA producing cells. Thus, identified increases in expression could identify energy spilling pathways to eliminate and allow for further carbon flux to FFA production.

Additionally, alterations to transcriptional regulator expression and introduction of heterologous pathways could be used to increase production. Modifying regulatory genes such as *fadR* and *crp* has been used to increase the production of FFAs (Dellomonaco et al., 2011; Zhang et al., 2012). However, alterations to global regulators similar to *crp* can have negative impacts on cellular fitness (Perrenoud and Sauer, 2005). Alternatively, introducing a different FFA synthesis pathway could bypass many regulatory elements. A similar strategy introduced the non-native mevalonate pathway to increase the production of isoprenoids in *E. coli* (Martin et al.,

2003). However, it is not always clear if each enzyme in a pathway will have a similar activity in a heterologous organism compared to its native host.

Chapter 6 summarizes experiments that were conducted to identify strategies for increasing FFA and FFA-derived product titer and yield and preliminary attempts to implement the strategies in *E. coli*. The strategies considered include targeted gene-knockouts identified by microarray analysis, altering expression of transcriptional regulators, and introducing alternative pathways. When each strategy was implemented, a modest increase in FFA titer and yield was achieved. Unfortunately, the observed titers and yields remain far below the theoretical maximum values.

6.2. Results and discussion

6.2.1. Whole genome transcriptional analysis

Prior work in our laboratory (Lennen et al., 2011) has shown that production of FFAs has a dramatic effect on global gene expression. These studies were conducted in batch growth studies to identify targets for improving FFA production and strain stability. Here, we performed similar experiments to determine the effects of fatty acid overproduction on global gene expression as a function of growth rate and media composition. To identify additional targets for strain improvement, a transcriptomic analysis was performed via microarray analysis on a series of carbon (Chapter 3) and phosphate (Chapter 4) limited chemostats. *E. coli* TY05 was cultured in various media compositions designed to generate phosphate or carbon limitation. Experiments were performed at a dilution rate of 0.1 h^{-1} in continuous culture and biomass samples were collected at 4 timepoints per steady state for transcriptome analysis. Total RNA was isolated from each sample and the resulting cDNA was hybridized against custom *E. coli* oligo arrays.

Data was analyzed using ArrayStar software (DNASTAR, Madison, WI) and reported below as linear fold changes between mean intensities of carbon and phosphate limited TY05 chemostat cultures.

To validate the collected data we examined the expression values for *fadA*, *fadB*, *fadD*, and *fadE*. These genes had been disrupted in *E. coli* TY05 and should therefore not be expressed in our samples. As expected, these transcripts showed extremely low expression values that were not statistically different than background (data not shown). The data was further verified by the observance of a significant increase in expression for the genes in the Pho regulon (selected genes in Table 6.1, full set in Appendix 5) under phosphate limitation. The Pho regulon is known to be upregulated under phosphate limitation. After normalization and statistical analysis, 602 genes were found to be significantly up- or down-regulated (fold change greater than 2 with a P value of less than 0.05) under phosphate limitation. To expedite the search for hypotheses, we focused on evaluating pathways (multiple genes that lead to the production of a particular compound or generate ATP) that were up or down regulated as a whole. Selected up- and down-regulated genes can be seen in Table 6.1. The full dataset is presented in Appendix 5.

Table 6.1: Selected significantly up- and down-regulated genes in E. coli TY05 grown under phosphate limitation versus carbon limitation.

Category/ Locus	Gene name	Fold Change[a]	P-value	Description
Pho regulon				
b0383	*phoA*	59.7	2.5×10^{-7}	bacterial alkaline phosphatase
b0399	*phoB*	19.9	6.3×10^{-6}	DNA-binding response regulator in two-component regulatory system with PhoR
Acetate production				
b0871	*poxB*	3.7	4.5×10^{-3}	pyruvate oxidase
Acid stress				
b3517	*gadA*	11.8	3.2×10^{-2}	glutamate decarboxylase A
b1493	*gadB*	11.3	3.0×10^{-2}	glutamate decarboxylase B
b1492	*gadC*	10.3	1.9×10^{-2}	glutamate:gamma-aminobutyric acid antiporter
b3512	*gadE*	20.0	1.6×10^{-2}	DNA-binding transcriptional activator
b3509	*hdeB*	4.9	8.1×10^{-3}	acid resistance protein
b3511	*hdeD*	6.8	3.6×10^{-2}	acid-resistance membrane protein
b0186	*ldcC*	3.1	2.7×10^{-2}	lysine decarboxylase 2
b3506	*slp*	7.2	2.8×10^{-2}	starvation lipoprotein
b3508	*yhiD*	7.9	3.7×10^{-2}	Mg(2+) transport, acid resistance
Amino acid degradation[b]				
b2903	*gcvP*	4.63	3.6×10^{-6}	glycine decarboxylase
b2905	*gcvT*	4.3	3.1×10^{-5}	aminomethyltransferase
b3616	*tdh*	2.2	2.4×10^{-2}	threonine dehydrogenase
b2965	*speC*	3.9	1.0×10^{-2}	ornitine decarboxylase
Fatty acid metabolism				
b0954	*fabA*	-3.0	4.8×10^{-2}	beta-hydroxydecanoyl thioester dehydrase
b2323	*fabB*	-2.1	3.4×10^{-2}	3-oxoacyl-[acyl-carrier-protein] synthase I
b2342	*fadI*	-11.1	3.0×10^{-2}	beta-ketoacyl-CoA thiolase, anaerobic, subunit
b2341	*fadJ*	-8.8	3.6×10^{-2}	fused enoyl-CoA hydratase and epimerase and isomerase -!- 3-hydroxyacyl-CoA dehydrogenase
b2344	*fadL*	-14.9	3.2×10^{-2}	long-chain fatty acid outer membrane transporter

a. Fold change is calculated based on linear difference in average intensities between TY05 under phosphate versus carbon limitation.

b. Only genes that were significantly up or down regulated (absolute value of fold change greater than 2, P value less than 0.05) in the indicated pathways are listed in this table

Table 6.1: (continued) Selected significantly up- and down-regulated genes in E. coli TY05 grown under phosphate limitation versus carbon limitation.

Category/ Locus	Gene name	Fold Change[a]	P-value	Description
Multi-drug efflux[b]				
b2074	*mdtA*	3.0	1.2×10^{-3}	multidrug efflux system
b3513	*mdtE*	19.6	1.8×10^{-2}	mutlidrug resistance efflux transporter
b3514	*mdtF*	8.4	2.7×10^{-2}	mutlidrug resistance efflux transporter
b0842	*cmr*	4.4	3.1×10^{-4}	multidrug efflux system protein
Pentose phosphate[b]				
b1852	*zwf*	2.4	8.0×10^{-4}	glucose-6-phosphate dehydrogenase
b0767	*pgl*	2.0	4.8×10^{-2}	6-phosphogluconolactonase
b2029	*gnd*	3.0	2.2×10^{-4}	gluconate-6-phosphate dehydrogenase, decarboxylating
b2465	*tktB*	4.4	3.4×10^{-2}	transketolase 2, thiamin-binding
b0008	*talB*	2.2	8.0×10^{-4}	transaldolase B
Glycolysis[b]				
b4025	*pgi*	2.2	1.4×10^{-3}	glucosephosphate isomerase
b1723	*pfkB*	3.28	2.3×10^{-4}	6-phosphatefructokinase II
b2097	*fbaB*	2.1	4.0×10^{-2}	fructose bisphosphate aldolase class I
b3612	*gpmI*	4.2	2.6×10^{-7}	phosphoglyceromutase III, cofactor independent
b1676	*pykF*	3.1	2.6×10^{-4}	pyruvate kinase I
TCA cycle[b]				
b0723	*sdhA*	-2.4	2.9×10^{-2}	succinate dehydrogenase, flavoprotein subunit
b0721	*sdhC*	-3.7	9.6×10^{-4}	succinate dehydrogenase, membrane subunit
b0722	*sdhD*	-2.8	3.2×10^{-2}	succinate dehydrogenase, membrane subunit
b1612	*fumA*	-2.9	2.2×10^{-2}	fumarate hydratase
b1611	*fumC*	-3.2	1.2×10^{-2}	fumarate hydratase
b2210	*mqo*	17.4	3.1×10^{-4}	malate dehydrogenase
b4015	*aceA*	-3.2	1.2×10^{-4}	isocitrate lyase
b4014	*aceB*	-3.3	2.0×10^{-5}	malate synthase A

a. Fold change is calculated based on linear difference in average intensities between TY05 under phosphate versus carbon limitation.
b. Only genes that were significantly up or down regulated (absolute value of fold change greater than 2, P value less than 0.05) in the indicated pathways are listed in this table

Glycolysis and pentose phosphate pathways were both observed to be upregulated under phosphate limitation, which was expected based on increased biomass specific glucose uptake rates (Figure 6.3). Another factor behind pentose phosphate pathway upregulation could owe to

the production of NADPH cofactor required for enzymatic activity in fatty acid biosynthesis (Rock and Jackowski, 2002). Surprisingly, a majority of genes in the TCA cycle were not upregulated, many, including succinate dehydrogenase, were downregulated.

Despite the significant increase in FFA biomass specific productivity under phosphate versus carbon limiting conditions, the differences in transcript level of fatty acid biosynthesis genes, such as *accA, accB, accC, accD, fabD, fabF, fabG, fabH fabI,* and *fabZ* (1.2, 1.2, 1.3, 1.0, 1.1, -1.5, -1.8, 1.2, 1.2, and -1.4 fold changes), were minimal and statistically insignificant. Genes encoding enzymes involved with unsaturated fatty acid biosynthesis were down-regulated under phosphate limitation (*fabA* and *fabB,* Table 6.1), consistent with prior results seen by our lab (Lennen et al., 2011). FabR represses both *fabA* and *fabB* (Zhu et al., 2009), however, it was not observed to be upregulated under phosphate limitation. As hypothesized by Lennen, such a decrease in expression could owe to the pool of saturated ACPs being depleted by the activity of BTE, leaving a higher proportion of unsaturated ACPs to ultimately make up the cell membrane. An increase in unsaturated content could lead to vastly different properties of the cell membrane and affect integrity (Lennen and Pfleger, 2013). The lower expression levels could be a way to combat such an effect by lowering the flux toward unsaturated fatty acid biosynthesis.

Numerous amino acid degradation pathways were found to be upregulated under phosphate limiting conditions (Table 6.1; glycine, threonine, glutamate, lysine). Threonine and glycine degradation pathways lead to the generation of NADH and NADPH along with CO_2, indicating a cell need for further reducing power under phosphate limitation. The glycine and threonine cleavage systems are induced by extracellular threonine and glycine, and negatively regulated in the presence of purines (Ghrist and Stauffer, 1995). The glycine cleavage system (Figure 6.1) is also significant as it is used to donate one carbon units in the conversion of

glycine to serine (while at the same time generating NADH and nitrogen in the form of ammonia), which can be degraded to pyruvate (Han et al., 2002; Kikuchi et al., 2008). However, as serine degradation genes are not upregulated under phosphate limitation, it is possible that the N5,N10-methylenetetrahydrofolate generated during glycine cleavage is converted to N5,N10-methenyltetrahydrofolate (to eventually be converted to a purine) and generates an NADPH (Kikuchi et al., 2008). As FFA biosynthesis requires NADPH, it would make sense that more flux would go toward pathways needing to synthesize NADPH, and it is possible that some of these amino acid degradation pathways are upregulated for the generation of reducing power in synthesizing FFAs.

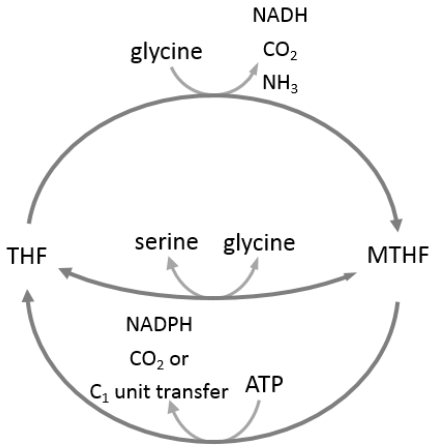


Figure 6.1: The glycine cleavage system and recycling of tetrhydrofolate (THF) and N5,N10-methylenetetrahydrofolate (MTHF). When the glycine cleavage system genes are overexpressed the cell breaks down glycine to produce NADH (also generating MTHF). The MTHF can be recycled back to THF in the conversion of glycine to serine. Additionally, alternative pathways that generate NADPH (along with consuming ATP) can also help restore THF. Arrows connecting THF and MTHF represent multiple enzymatic steps in a metabolic pathway.

Another possible reason behind the upregulation of amino acid degradation genes could have to do with restoring proton motive force, as the decarboxylation has the added effect of consuming an intracellular proton (Zhao and Houry, 2010). The high expression of the phage shock and significant upregulation of acid stress response genes suggest a loss of either

membrane integrity and/or proton motive force (Castanié-Cornet et al., 2007; Kobayashi et al., 2006). Removal of amino acid catabolic genes has been shown to lead to a significant reduction in growth, possibly owing to fewer options to regain proton motive force (Chattopadhyay et al., 2009). As experimental evidence suggested chemostat cultured cells were intact (Figure 6.2), full loss of membrane integrity is unlikely (based on the decarboxylases upregulated and the upregulation of acid stress response genes). Maintenance of the proton motive force is necessary to allow for ATP synthesis via the electron transport chain. Thus, it could be possible that FFA synthesis affects cellular pH and dissipates proton motive force as suggested by microarray data in this report and others (Lennen et al., 2011).

Interestingly, not all genes observed by Lennen et al. to be up-regulated under FFA producing conditions relative to wild type were observed to be differentially regulated under higher FFA producing conditions when cells were grown under phosphate limitation. For example, the genes encoding the phage shock proteins (*pspABCDEG*) were not observed to be significantly up-regulated (1.0-1.4 fold change) under phosphate limitation. However, normalized intensity values of *pspA, pspB, pspC,* and *pspD* transcrpts were found to be similar to genes in the pentose phosphate pathway, indicating at least moderate expression. Additionally, the genes *marA,* rob, and *soxS* either had insignificant changes in expression (1.2 and -1.4 fold changes for *marA* and *rob*) or significant downregulation (-4.3 fold change with a P-value of 0.023 for *soxS*). Also of note were the low expression levels of each regulator under phosphate limitation (5 fold less than the phage shock transcript levels). However, these differences may be the result of different expression in stationary phase versus controlled slow growth. Additionally, phosphate limited cells stain intact (when analyzed by SYTOX staining, Figure 6.2) during slow growth versus approximately 35% in stationary phase (Lennen and Pfleger, 2013).

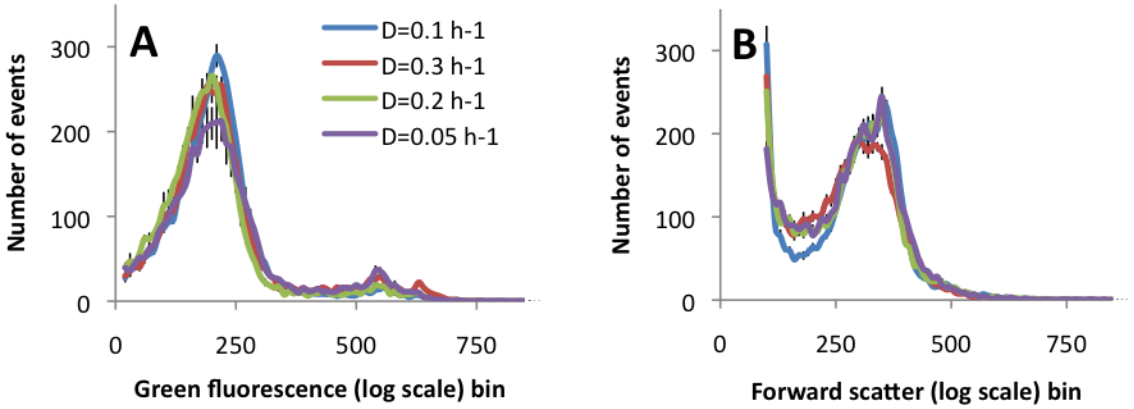


Figure 6.2: SYTOX staining study A. Average green fluorescence (485nm excitation, 525 nm emission) histograms of SYTOX green stained cells and B. averaged log scale forward scatter for E. coli TY05 cells taken from phosphate limited chemostat cultures at dilution rates of 0.05, 0.1, 0.2, and 0.3 h-1. Non intact cells show log scale fluorescence values of greater than 500 (Lennen et al., 2011), as SYTOX is a green nucleic acid dye that only can penetrate cells when the inner membrane is not intact. Intact cells show a lower fluorescence value that is associated with staining of the cell surface (Roth et al., 1997).

In contrast, several genes demonstrated differential expression consistent with our prior transcriptomic results (Lennen et al., 2011) and those of other FFA producing strains (Zhang et al., 2012). Genes encoding for the multidrug efflux pumps AcrAB, Cmr, MdtABC, MdtEF, and TolC were all increased in expression under phosphate limitation. Additionally, various membrane stress and acid shock proteins also increased in expression under phosphate limitation, including *hdeAB-yhiD* and genes in the glutamate dependent acid resistance. Surprisingly, *rpoE* (σ_E/σ^{24}, the extreme membrane/heat stress sigma factor) and *cadAB* were not observed to be up-regulated, despite *rpoE* expression being significantly associated with the cellular acid stress response (Hommais et al., 2004).

The increase in expression of *mdtEF* (>8 fold) and *tolC* (~2 fold) under phosphate limitation could possibly explain the higher specific productivities. The inner membrane efflux pumps AcrAB, MdtEF, and EmrAB are likely involved in FFA export along with the outer membrane pump TolC (Lennen et al., 2013). Knocking out all three inner membrane efflux

pumps essentially eliminated FFA production and complementation of a knockout of one pump with overexpression of one of the other three leads to a restoration of FFA tolerance (Lennen et al., 2013). Thus, it is conceivable that the increased expression of *mdtEF* and *tolC* is contributing to increased FFA export allowing for a higher biomass specific FFA production rate without affecting the amount of intact cells.

In addition to a comparison of gene expression under phosphate versus carbon limitation, we analyzed our datasets to examine the effect of growth rate on gene expression. Microarray analysis was performed on samples taken during TY05 steady states under phosphate limitation at the growth rates of 0.05, 0.1, 0.2, and 0.3 h^{-1}. Samples were processed as described above. After statistical analysis, 684 genes were found to have significant change between the multiple growth rates tested, with over 400 significant differences between the highest and lowest tested growth rates. A brief table highlighting interesting sets of genes is shown in Table 6.2. A full data table for all statistically differentially expressed genes can be found in Appendix 6.

Table 6.2: Selected significantly up and down regulated genes in TY05 under high versus low growth rates under phosphate limitation

Category/ Locus	Gene name	Fold Change[a]	P-value	Description
Flagellar synthesis (subset)				
b1922	*fliA*	19.6	$1.6x10^{-3}$	RNA polymerase, sigma 28 (F) factor
b1923	*fliC*	68.3	$7.7x10^{-4}$	flagellar filament structural protein
b1076	*flgE*	58.0	$1.8x10^{-3}$	flagellar hook protein
b1946	*fliN*	33.0	$1.3x10^{-3}$	flagellar motor switching and energizing component
Chemotaxis				
b1888	*cheA*	41.20	$7.6\ x10^{-4}$	fused chemotactic sensory histidine kinase in two-component regulatory system with CheB and CheY: sensory histidine kinase -!- signal sensing protein
b1883	*cheB*	15.94	$2.5x10^{-3}$	fused chemotaxis regulator -!- protein-glutamate methylesterase in two-component regulatory system with CheA
b1884	*cheR*	17.51	$3.6x10^{-4}$	chemotaxis regulator, protein-glutamate methyltransferase
b1882	*cheY*	20.32	$2.3\ x10^{-3}$	chemotaxis regulator transmitting signal to flagellar motor component
b1881	*cheZ*	12.35	$1.7\ x10^{-3}$	chemotaxis regulator, protein phosphatase for CheY
Multidrug efflux				
b2685	*emrA*	2.9	$5.9x10^{-3}$	multidrug efflux system
b2686	*emrB*	2.7	$5.5x10^{-4}$	multidrug efflux system
Sulfur transport (subset)				
b0936	*ssuA*	9.86	$3.1\ x10^{-2}$	alkanesulfonate transporter subunit -!- periplasmic-binding component of ABC superfamily
b0935	*ssuD*	5.18	$1.3\ x10^{-2}$	alkanesulfonate monooxygenase, FMNH(2)-dependent
b0366	*tauB*	3.48	$1.3\ x10^{-3}$	taurine transporter subunit -!- ATP-binding component of ABC superfamily
b3917	*sbp*	3.95	$2.4\ x10^{-3}$	sulfate transporter subunit -!- periplasmic-binding component of ABC superfamily

a. Fold change is calculated based on linear difference in average intensities between TY05 under phosphate limitation at a dilution rate of 0.3 h^{-1} to 0.05 h^{-1}.

The data set showed significant up-regulation in all flagella synthesis genes, with increased expression from 10 to 70 fold at the highest growth rate. Such a result may have been expected due to data showing high expression of flagella synthesis genes during mid log growth (Chilcott and Hughes, 2000). Additional genes involved in flagella operation, such as *cheAB* and *cheYZ* involved in chemotaxis also had significantly higher expression under the highest tested growth rate (Table 6.2, Appendix 6). The 27 most upregulated transcripts and 46 of the first 50 are related to flagella assembly/operation or chemotaxis. Of the genes with fold change differences greater than 5 (61), only genes associated with the Rac prophage (9 genes) and sulfonate utilization (3 genes). The increased transcript level of sulfonate utilization genes (*ssuEAD*) is odd as expression is related to consumption of aliphatic sulfonates as sulfur sources (Eichhorn, 1999). While *E. coli* has been shown to use MOPS as a sulfur source under conditions of sulfate starvation (Neidhardt et al., 1974), the chemostat was not run under limiting sulfur conditions. Also, at the dilution rate most likely to have co-limitation conditions (0.05 h^{-1}), sulfonate utilization genes had expression levels similar to background. Despite this, other sulfur starvation proteins were observed to have increased expression, including a taurine transporter (*tauABC*, 3.5 fold increase) (Van der Ploeg et al., 1996), indicating the possibility of sulfur co-limitation at the higher growth rate. The genes coding for the multidrug efflux pump EmrAB were observed to have higher expression, which was surprising owing to the lack of difference in expression between carbon and phosphate limitation at the same growth rate. In addition to these genes, the ribosome modulation factor (*rmf*) was seen downregulated under higher growth conditions. Such a result is expected as expression has been shown to increase in growth regimes that are in a transitional region between exponential growth and stationary phase (such as 0.05 h^{-1}) (Wada et al., 1995). However, the majority of expressed genes detected in the microarray

study showed no clear pattern with respect to increasing growth rate. Despite this, it appears as though flagella synthesis/operation and chemotaxis show a clear pattern between higher growth and expression under phosphate limitation. Additionally, all significantly up and down regulated genes can be found in Appendix 6.

6.2.2. Using transcriptomic study to identify targets for modification

From analysis of the data in chapters 4, 5 and 6, we hypothesized that a significant amount of carbon is being converted to CO_2, around double that of what would be expected if the cell was maximizing FFA production (see Chapter 4). In *E. coli*, glycolysis and the TCA cycle are likely sources of CO_2 production if the cell requires extra ATP for maintenance when producing elevated levels of FFA. In addition, over 70 enzymatic reactions encoded in the *E. coli* genome also produce CO_2. Many of these were significantly expressed in TY05 and up-regulated when cells were grown in phosphate limited media. We therefore hypothesized that deletion of up-regulated CO_2 producing enzymes, deletion of unnecessary ATP sinks, and disruption of alternative carbon sinks could direct extra carbon flux to FFA. We therefore used the functional genomics data collected above to prioritize a list of knockout targets that could increase flux to FFA.

Considering the demand for NADPH in *E. coli* TY05 (for producing FFA), forcing flux through the NADPH generating pathways instead of glycolysis could improve FFA yields. For example, the pentose phosphate pathway (PPP) is a catabolic pathway that produces anabolic sugar-phosphates (precursors to nucleotides and aromatic amino acids), intermediates in glycolysis, and reducing power from glucose (Figure 6.3). The pathway converts three molecules of glucose-6-phosphate to two molecules of fructose-6-phosphate, one molecule of glyceraldehyde-3-phosphate, three molecules of CO_2, and six NADPH. The cell uses the PPP to

synthesize NADPH needed in anabolic pathways such as fatty acid biosynthesis. The cell is likely already partially using this strategy. From Table 6.1, genes in the PPP are up-regulated 2-4 fold under phosphate limitation versus carbon limitation.

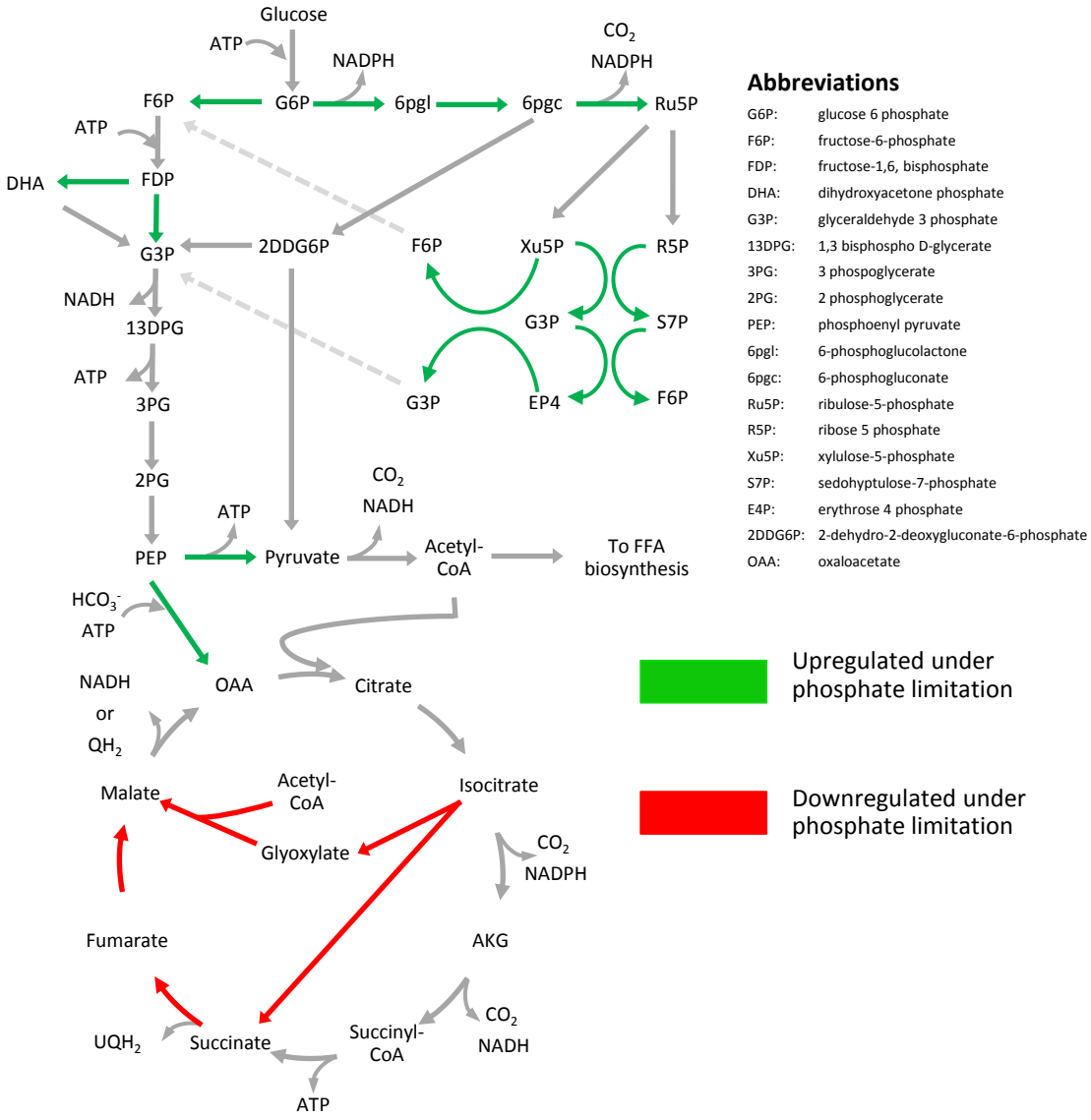


Figure 6.3: Combined pathway of glycolysis, PPP, ED, and the TCA cycle. ATP, CO_2 and reducing equivalents that are consumed or generated are shown as side products for each reaction where they are required/generated. Dashed arrows do not indicate a reaction but rather serve to show where those compounds are used later in the pathway. Arrows in green indicate that transcripts coding for enzymes in those steps were upregulated under phosphate limitation, while red arrows indicate downregulation.

The disadvantage of the PPP relative to glycolysis is the loss of three carbons as CO_2 via the activity of 6-phosphogluconate dehydrogenase (Gnd) on 6-phosphogluconate. Alternatively, the 6-phosphogluconate can be acted upon by phosphogluconate dehydratase (Edd), which would send the compound through the Entner-Doudoroff (ED) pathway. The ED pathway is an alternative to the top of glycolysis, which produces one NADPH and one NADH (where glycolysis generates 2 NADH per glucose) at the expense of generating one less ATP. Directing flux through the ED pathway could help generate the needed NADPH without wasting carbon as CO_2. Thus, knocking out *gnd* from TY05 could be a possible mechanism to reduce CO_2 as it would allow for NADPH production and eliminate a source of CO_2 production at the expense of an ATP. Similarly, blocking the isomerization of glucose-6-phosphate to fructose-6-phosphate would force all carbon flux through either the PPP or ED pathway.

Another pathway identified under phosphate limiting conditions was the glutamate dependent acid resistance pathway (Figure 6.4). The pathway is typically expressed when a cell is exposed to extreme acid conditions and has the effect of increasing the proton motive force while consuming a proton while converting glutamate to 4-aminobutyrate (γ-aminobutyric acid, GABA) and CO_2 (Castanié-Cornet et al., 2007). However, the system requires extracellular glutamate to function properly (Ma et al., 2003). Up-regulation of pathway has also been observed when cells are exposed to osmotic stress, heat stress, and anaerobic phosphate starvation (Moen et al., 2009; Moreau, 2007). In our functional genomics experiments, the genes for the decarboxylases (*gadA* and *gadB*), transporter (*gadC*), and transcriptional regulators (*gadX* and *gadE*) were upregulated 10-20 fold under phosphate limitation in TY05 versus carbon limitation. Given that the cells were not grown in the presence of extracellular glutamate and no extracellular GABA was detected, GABA must have been recycled. This can be achieved by

using α-ketoglutarate (AKG) to convert GABA to glutamate and succinate via reactions catalyzed by PuuE and Sad. Overall, this set of reactions would be equivalent to steps in the TCA cycle (AKG to succinate) with the loss of one ATP (normally made by succinyl-CoA synthetase (SucCD)). We hypothesized that eliminating the glutamate dependent acid resistance pathway could reduce the maintenance energy of our FFA producing cells. Expression of *gadA* and *gadBC* requires the regulator GadE (Ma et al., 2003), so we decided to knock it out of TY05. GadE has been shown to regulate other genes outside of the glutamate dependent acid resistance pathway (Sayed et al., 2007), so direct disruption may be a better option.

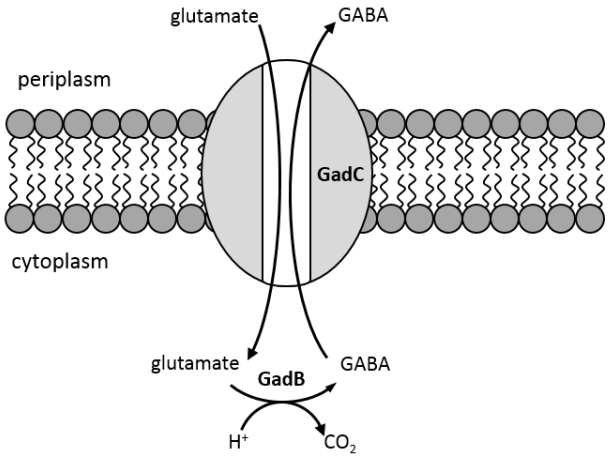


Figure 6.4: The glutamate dependent acid resistance pathway. In the cytosol, GadB (or GadA) converts glutamate to GABA consuming a proton and giving off CO_2. GABA export to the periplasm is then coupled to glutamate import through GadC (inner membrance transporter), which has the net effect of exporting a proton out of the cell. Enzymes are represented by bolded names.

A third attractive target for reducing ATP demand was flagella synthesis and operation. When FFAs were produced at the highest biomass specific rates, the genes involved in flagella synthesis and operation were significantly upregulated (give specific values for certain genes and refer to the appropriate table). As flagella synthesis and operation is very energy and biomass intensive (Zhao et al., 2007), eliminating the ability of the cell to express these genes could allow for more cellular resources to be diverted to FFA production. While eliminating flagella function

prevents cells from scavenging nutrients (Liu et al., 2005), such a concern would not be as significant in well mixed bioreactor. As FliA (σ^{28}) is the master regulator of flagella synthesis (Berg, 2003), we hypothesized that a disruption of *fliA* would eliminate flagella synthesis. Prior research has indicated significantly reduced motility in a *fliA* knockout strain of *E. coli* (Wood et al., 2006). Alternatively, disrupting the gene encoding the primary filament of the flagellum (FliC) could prevent significant energy expenditure. This hypothesis is supported by the fact that *E. coli* mutants lacking *fliC* display a non-motile phenotype (Copeland et al., 2010). Other regulatory elements of flagella exist (FlgM, FlhC, and FlhD), however, overlapping regulation would require multiple knockouts to eliminate cellular motility (Berg, 2003).

A fourth target for improving FFA production is the removal of the acetate synthesis pathways. When *E. coli* TY05 was grown in chemostat cultures under phosphate limitation, significant levels of acetate (3-4% of total carbon processed) were observed. Acetate production is used to generate additional ATP during rapid cell growth and is reincorporated when other carbon sources have been exhausted. Our FFA producing cells require acetyl-CoA for production of malonyl-CoA via acetyl-CoA carboxylase (AccABCD). This key reaction, and regulatory point, competes with acetate production pathways for available acetyl-CoA. FFA production studies in our lab as well as research in other groups has shown that incorporation of an *ackA* (encoding acetate kinase) or *ackA-pta* (encoding acetate kinase and phosphate acetyltransferase) knockout to a FFA producing strain has a minimal impact on FFA titer and yield (Li et al., 2012). In addition to AckA-Pta, pyruvate oxidase (PoxB) can produce acetate and CO_2 from pyruvate in competition with synthesis of acetyl-CoA. The *poxB* transcript level was increased over four fold under phosphate limiting conditions versus carbon limiting conditions (Table 6.1). Therefore, it may be the dominant route for producing the acetate in our FFA producing cells.

6.2.3. Production results from targeted knockouts

Based on directions from the microarray study knockouts to the *fliA*, *fliC*, *gadA*, *gadB*, *gadC*, *gadE*, *gnd*, and *poxB* loci in *E. coli* TY05 were made using P1 transduction. These strains were tested in triplicate shake flask production runs using phosphate limited modified MOPS minimal media. As each strain harbored no plasmids, no antibiotic was added to the media. Strains were cultivated for 48 hours after induction with IPTG, at which point they were harvested and the FFA titer was analyzed as described previously.

The only gene deletions that increased FFA yield on glucose were *fliA* and *fliC* (~11%)(Figure 6.5), with only TY05Δ*fliC* being statistically significant with high confidence (P value – 0.035). Conversely, knocking out genes involved in the glutamate dependent acid resistance system significantly decreased yield by 25-35%. Strains with knockouts of *gnd* and *poxB* showed no difference in final lauric acid yield on glucose compared to the TY05 control. Lauric acid titer was also reduced by 10-15% as TY05Δ*gnd* and TY05Δ*poxB* consumed only 85% of the fed glucose. All other strains consumed over 95% of fed glucose. Acetate production was around 0.1 g/L in most tested strains, with TY05Δ*gadB*, TY05Δ*gnd*, and TY05Δ*poxB* producing 0.02-0.03 g/L.

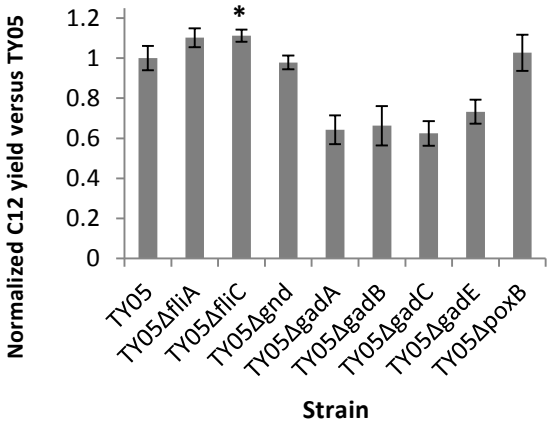


Figure 6.5: Targeted knockout production results. Normalized yield values of C_{12} FFA on glucose in each of the TY05 knockout strains. The asterisk indicates where the difference was statistically significant (P value less than 0.05).

It is possible that the significant decrease in yield with the *gadE* knockout owes to its regulation of many genes outside those involved in glutamate dependent acid response (Sayed et al., 2007). Also, it is worth noting that initial experiments were performed in a strain containing both acetate production pathways and in shake flasks without pH control. Running the experiment using TY34Δ*gadE* or in a pH controlled bioreactor could alter the negative impact on yield seen in TY05Δ*gadE*. Additionally, there could be polar effects on expression of the downstream transporter MdtEF, identified as a FFA exporter (Lennen et al., 2013), owing to its proximity to *gadE*. Evidence linking the three genes to a single transcription unit, and data indicating that GadE activates expression of *mdtE* and *mdtF* have been reported (Nishino et al., 2008; Sayed et al., 2007). Such regulation could be a possible explanation of the lower FFA titers observed in *gadE* deletions. Individual knockouts of *gadA* and *gadB* likely showed no increase in FFA titer as each is capable of compensating for the loss of the other (De Biase et al., 1999). Additionally, a *gadA* disruption could have polar effects on *gadX*, which encodes for a transcriptional regulator affecting multiple genes (Nishino et al., 2008; Tramonti et al., 2002).

The lower FFA titer observed in the Δ*gadC* mutant could be related to fact that GadA and GadB do not need GadC to decarboxylate glutamate (Capitani et al., 2003). The marginal effect seen with removal of *poxB* and *gnd* in TY05 could indicate a minimal shift in metabolic flux (pathways not previously consuming much carbon) or that flux has shifted to alternative pathways that still produce similar amounts of CO_2 (or direct carbon away from FFA biosynthesis). The increase in FFA yield due to the *fliC* knockout is likely due to retasking resources that are no longer being used on the synthesis and operation of the bacterial flagella to FFA synthesis or cellular maintenance. As the flagella in *E. coli* are powered by proton pumps and very energy intensive (Gauger et al., 2007; Gosink and Häse, 2000), their removal should allow for further ATP synthesis or maintenance of the cellular proton motive force. It is also likely that only a small increase in production occurred as the genes were only observed to be highly expressed under medium to high growth rates (Table 6.2) and are likely only using up a small percentage of the overall cellular resources, limiting the maximum benefit in FFA production that can be achieved from their removal.

Analysis of the functional genomics data provides several additional hypotheses for methods of improving fatty acid titers and yields. As fatty acid biosynthesis requires NADPH, it could be beneficial to disrupt *pgi* in a strain that is already Δ*gnd*. The resulting strain would then go through both PPP and ED to generate an NADPH not generated in glycolysis. Metabolic flux analysis has even identified a *pgi* knockout as significant to increasing FFA yields along with a disruption to a transaldolase (encoded by *talA* or *talB*) with the goal of forcing flux through the ED pathway (Ranganathan et al., 2012). Ranganathan also suggested downregulation of the TCA cycle to maintain a high pool of acetyl-CoA for fatty acid biosynthesis, and was able to achieve higher yields by disrupting a gene encoding for succinyl-CoA sythetase (*sucC*). However, a

possible alternative strategy to knocking out genes in central metabolic pathways would be to find a method to better control knockdown of gene expression. Metabolic flux analysis has predicted knockdowns of genes encoding for phosphoglycerate mutase (*gpmA* and *gpmM*), folate metabolism (which is suggested to be active based on the possible glycine and threonine degradation seen in the microarray data), and the entire TCA cycle. If one could express some sort of sequence specific effector for the native promoters of certain central metabolic genes, far more control over metabolic flux could be achieved. In addition to disrupting a gene or downregulating a pathway, upregulation of a regulatory gene such as *fadR* (Zhang et al., 2012) or genes coding for pyruvate dehydrogenase (e.g. *aceE*) have also been suggested (Xu et al., 2013). Overall, the microarray data, initial set of knockouts, and work by others in the field has provided a number of directions to take in reaching the theoretical maximum yield of FFAs on a given carbon source.

6.2.4. Gene overexpression to increase production (fadR overexpression)

Recent research has shown that overexpression of the transcriptional regulator *fadR* increases production of C_{14}-C_{16} FFAs in *E. coli* DH1 expressing the TesA' thioesterase (Zhang et al., 2012). We therefore wanted to examine the effect of FadR overexpression on our C_{12} FFA producing strains. Three strains, TY03, TY05, and TY25, harboring two, three, and four copies of BTE on the chromosome were transformed with either a plasmid expressing *fadR* under the IPTG inducible trc promoter or the corresponding empty vector for use as a control. The quadruple BTE integrant (TY25) was examined in case BTE expression limited in FFA production in the triple integrant when *fadR* was overexpressed. Cultures of each strain were grown in phosphate limiting, MOPS minimal media containing 25 μg/mL kanamycin with 0.7% glucose. To remain phosphate limited, the phosphate concentration in the media was reduced to

0.24 mM instead of the normal 1.32 mM. After 40 hours, the C_{12} FFA titers of the *fadR* overexpression strain and the control strain were not statistically different, regardless of the BTE copy number (Figure 6.6A). As per previous fermentations, each TY25 strain showed significant variation in fatty acid titer, which could possibly indicate instability in the strain (due to four copies of the same gene on the chromosome) or possible toxicity effects due to too much BTE expression (Hoover et al., 2011; Lennen et al., 2010). However, there were increases in C_{14}, C_{16}, and C_{18} (Figure 6.6B) for the strains overexpressing *fadR*. While the percentage increase was greatest in the C_{18} fatty acid species, each titer rose by approximately 20 mg/L with *fadR* overexpression (C_{14}, C_{16}, and C_{18}) so a significant increase in overall FFA titer was not observed when *fadR* was co-expressed with BTE.

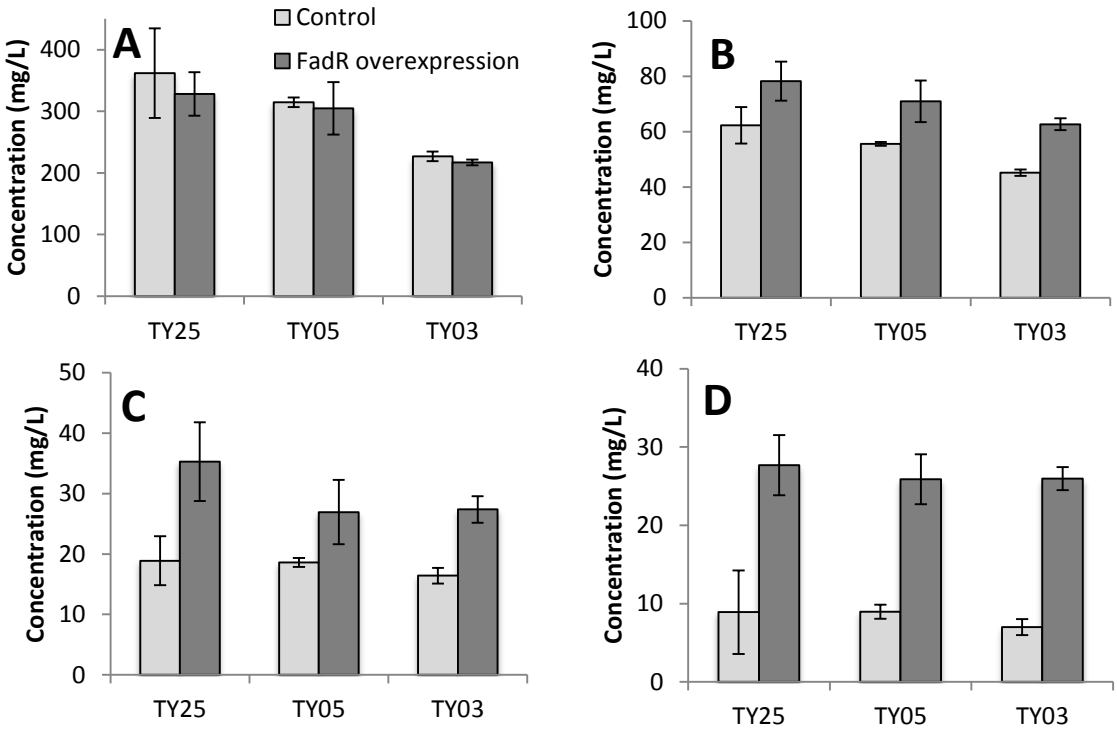


Figure 6.6: Initial FadR expression study. C_{12} (A), C_{14} (B), C_{16} (C), and $C_{18:1}$ (D) titer in TY03, TY05, and TY25 strains expressing either pBTRK-fadR (overexpression) or pBTRK (control). Error bars represent standard deviations from triplicate cultures.

To determine if BTE activity (limited by only four gene copies) was limiting the impact of FadR co-expression, BTE was cloned onto a medium copy (6-15 per cell) p15a origin plasmid under the control of the same trc promoter (pACYC-BTE). We were generously donated a compatible vector containing *fadR* on an arabinose inducible plasmid (pE8-fadR) from Dr. Fuzhong Zhang. *E. coli* MG1655Δ*ara*Δ*fadE* was transformed with combinations of pACYC-BTE, pE8-fadR, and/or their respective empty vectors (negative controls). The same sets of plasmids were also transformed into *E. coli* DH1Δ*fadE* (graciously donated by Dr. Fuzhong Zhang, Washington University in St. Louis). To test production, each strain/plasmid combination was grown in 50 mL of modified M9 minimal media containing 2.0% glucose (Zhang et al., 2012), 50 μg/mL ampicillin, and 34 μg/mL chloramphenicol. After induction at an OD_{600} of 0.8, the cultivation continued for 72 hours before harvesting for fatty acid and sugar samples. Unlike the initial *fadR* overexpression test, both C_{12} and C_{14} FFA titer were significantly increased (1.5-2-fold) in the presence of extra FadR (Figure 6.7A). However, the *fadR* overexpressing strains also consumed more glucose, leading to a modest increase in yield (Figure 6.7B). Additionally, while the relative titer was improved in plasmid-based BTE strains co-expressing *fadR*, the absolute titer was still significantly lower than our optimal production strains (such as TY05 which produces two-fold higher C_{12} FFA ~ 0.8 g/L). These results made us suspect that the high titers observed by Zhang et al. in *fadR* co-expressing cells may be a thioesterase or strain dependent effect (Zhang et al., 2012).

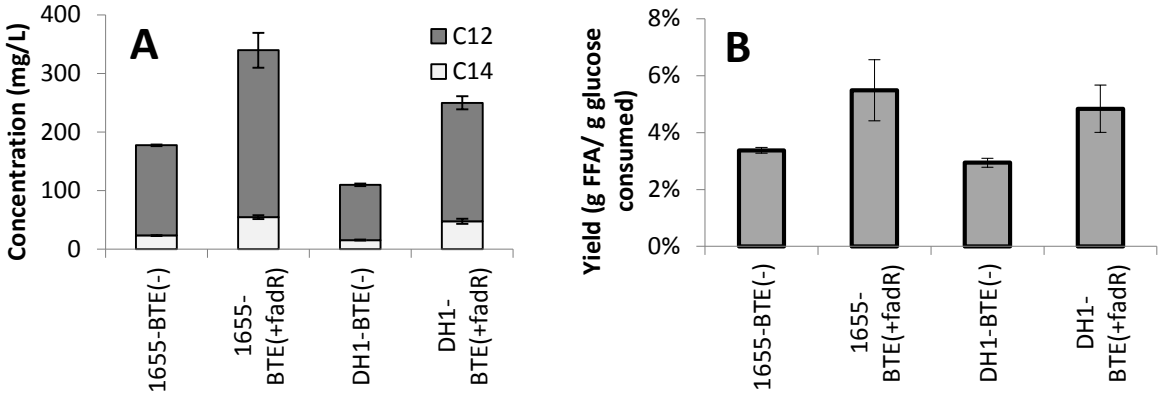


Figure 6.7: FadR expression in DH1 and MG1655 based FFA overproducers. C_{12} and C_{14} FFA titer (A) and total FFA yield (B) for co-expression of *fadR* with BTE in either a DH1Δ*fadE* or MG1655Δ*ara*Δ*fadE*.

To test the effect of thioesterase or cultivation condition dependent effects on FFA titer in *fadR* overexpression, the plasmid pKS1 (Steen et al., 2010) was acquired through the Addgene repository (Addgene, Cambridge, MA). It was transformed into DH1Δ*fadE* along with either pE8-fadR or an empty vector control. Production studies were performed in 50 mL of modified M9 minimal media containing 2.0% glucose (Zhang et al., 2012), 50 μg/mL ampicillin, and 34 μg/mL chloramphenicol, with similar induction conditions. Additionally, a second set of cultures were grown in 5 mL tubes containing the same media mix and conditions to identically replicate the conditions given described by Zhang. After 72 hours, sugar and FFA samples were taken, processed, and analyzed.

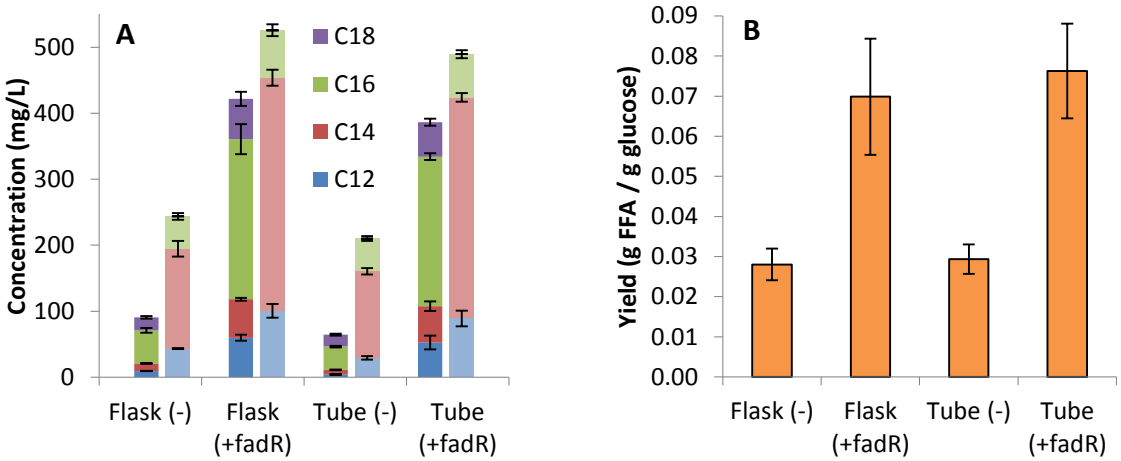


Figure 6.8: FFA production in a 'TesA and FadR expressing DH1 strain. The breakdown of FFA titer in DH1Δ*fadE* strains (A) with saturated species in a lighter shade and saturated in a darker shade is shown. Total FFA yield on glucose is displayed in (B).

While the overall titer was still significantly lower than that observed by Zhang, there were significant impacts from overexpression of *fadR*. When one breaks down the overall titer into its individual fatty acid species, there is a significant increase in all unsaturated FFA species (Figure 6.8A). For example in shake flask cultures, the strain overexpressing *fadR* has a 5 fold increase in both the $C_{14:1}$ (60 vs. 12 mg/L) and $C_{16:1}$ (250 vs. 50 mg/L) species. However, despite these differences in both saturated and unsaturated titer, the yield of the strains expressing FadR off of a plasmid only had a 2.5-3 fold increase in yield (Figure 6.8B), owing to the greater glucose uptake in those strains. Yields were consistent in either shake flasks or tubes around 0.07 g FFA /g glucose, much lower than the 0.24-0.26 g FFA /g glucose reported in the original paper (Zhang et al., 2012). Based on observations, it is likely that FadR overexpression contributes to a significant increase in unsaturated FFA biosynthesis. Such a result would be consistent with observations that FadR activates expression of *fabA* and *fabB*, whose gene products are involved in unsaturated fatty acid biosynthesis (Fujita et al., 2007). Experiments that overexpressed FabB

rather than FadR led to a similar two fold increase in fatty acid yield (Zhang et al., 2012), indicating there may be an additional effect besides trying to reestablish the optimal balance between saturated and unsaturated fatty acid species. However, metabolic flux balance analysis predictions (Ranganathan et al., 2012) and experiments (San et al., 2011) have only shown a 20% improvement in FFA yield, which is more consistent with the experimental data seen here.

6.2.5. Heterologous pathway introduction

With the exception of reverse β-oxidation (Dellomonaco et al., 2011), all other production studies in Chapters 3, 4, 5, and 6 have looked at FFA and FFA derived product overproduction using native type II fatty acid biosynthesis, a highly regulated metabolic pathway that is used to produce essential metabolites (lipids) within *E. coli* (Fujita et al., 2007; Rock and Cronan, 1996). A successful strategy for circumventing native regulation is heterologous expression of a separate pathway that leads to the same products. For example, the production of isoprenoids in *E. coli* is greatly facilitated by by-passing the native DXP pathway via heterologous expression of a foreign mevalonate pathway (Martin et al., 2003). Here, native fatty acid biosynthesis could be by-passed by expressing a type I fatty acid synthase that is common in yeasts, mammals, and some bacteria. A type I system fuses all of the required enzymatic activities needed to produce a fatty acid into a single protein. Type I FAS enzymes are therefore highly integrated and channel intermediates to the next step in the pathway without the need to leave the proximity of the last active site. Type I FAS often terminate by thiol-transfer to coenzyme A, the substrates for many downstream products of fatty acids. Using a type I FAS would also allow direct production of acyl-CoAs for the production of fatty alcohols, removing the need for a TE or ACS (Figure 6.9). We chose to study two type I fatty acid synthases from

Corynebacterium glutamicum (FAS-A and FAS-B) because they have been shown to express in *E. coli* (Stuible et al., 1997).

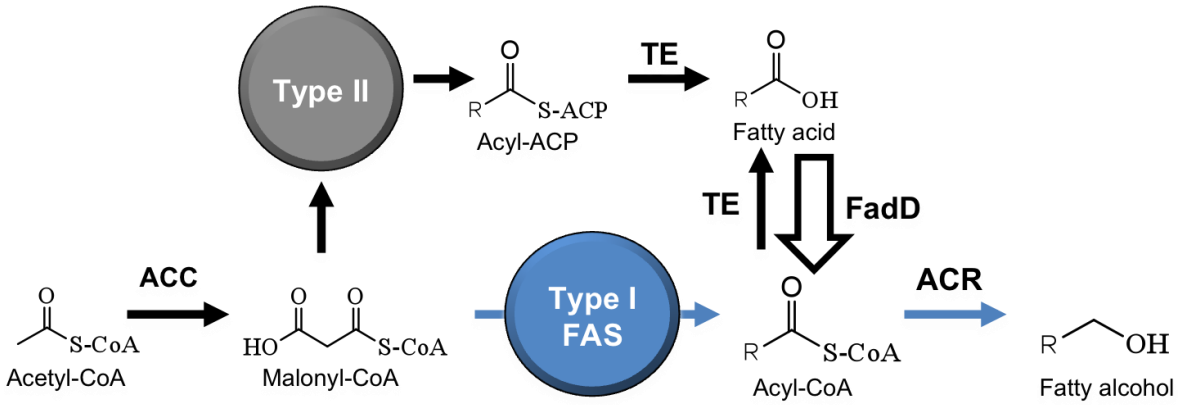


Figure 6.9: Illustration of how a type I fatty acid synthase (FAS) would lead to the production of fatty alcohols in E. coli. ACC represents acetyl-CoA carboxylase, TE represents a thioesterase, and ACR represents an acyl-CoA reductase.

A codon optimized FAS-A and native FAS-B gene were cloned into the medium copy vector pBAD18. Protein expression was tested using *E. coli* BAP1, a strain that expresses enzymes capable of adding essential post-translational modifications to Type I FAS. Strains expressing one of the type I synthases or the empty plasmid were grown for 24 hours in LB + 0.4% glycerol. FAS-B expressing cells produced slightly (~10%) more C_{16} fatty acids than the control (empty FAS-B vector) (Figure 6.10). The analysis used at the time, could not discern whether the C_{16} acyl-chains were free, bound in lipids, or in a CoA thioester. To determine if the FAS-B was functionally expressed, it was co-expressed with the acyl-CoA reductase from *Acinetobacter calcoaceticus* (Acr1) in *E. coli* DE.. Acr1 only acts on acyl-CoAs (unlike other reductases) and the gene coding for the long chain acyl-CoA synthetase (*fadD*) was knocked out. Therefore any observed alcohols would necessarily have been produced by the heterologous FAS-B. Cultures were grown for over 24 hours in LB + 0.4% glycerol. Culture extracts contained no detectable levels of fatty alcohols or fatty aldehydes (aldehydes are produced from

Acr1), whereas positive control cells (*E. coli* producing acyl-CoA's via FadD and expressing Acr1) were able to produce quantities of up to 50 mg/L.

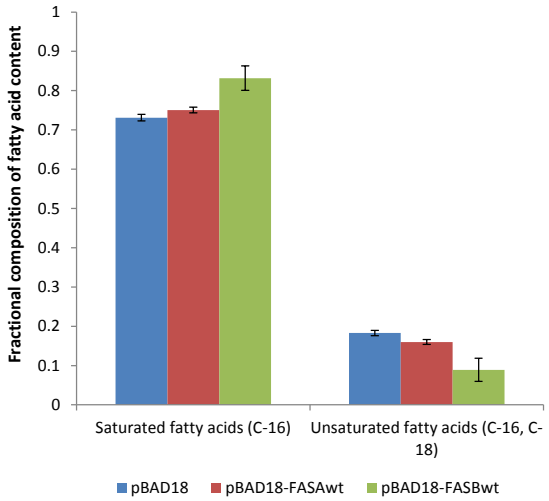


Figure 6.10: Type I FAS expression effect on fatty acid profile. Fractional composition of fatty acid content from *E. coli* BAP1 harboring pBAD18, pBAD18-FASA, and pBAD18-FASBwt.

The lack of fatty alcohol production indicated issues with either transcription, translation, or protein folding. To enhance expression, FAS-A and FAS-B were cloned into pBP18, a high-copy vector, under the control of the strong, IPTG inducible trc promoter. This vector also contained a translational coupling system developed in our lab to troubleshoot translation of heterologous genes (Mendez-Perez et al., 2012). Briefly, the coupling system links full-length translation of a gene of interest with translation of a resistance marker or fluorescent protein. In this system, resistant or fluorescent phenotypes can only be obtained if the gene of interest is fully translated into protein. Both type I FASs were expressed along with an empty vector control and cell pellets were lysed to analyze for FAS proteins. Despite being his tagged, no purified type I FAS protein was identified (Figure 6.11). However, significantly large bands slightly above 250 kDa were observed in both the FAS-A and FAS-B insoluble fractions, and an additional fainter band was observed at the same size in the FAS-B cell lysate. These results

indicate that FAS-A and FAS-B are not being highly expressed as soluble, well-folded proteins. Thus, prior to attempting to use either type I FAS for production of FFA derived compounds, the issues with functional expression need to be addressed.

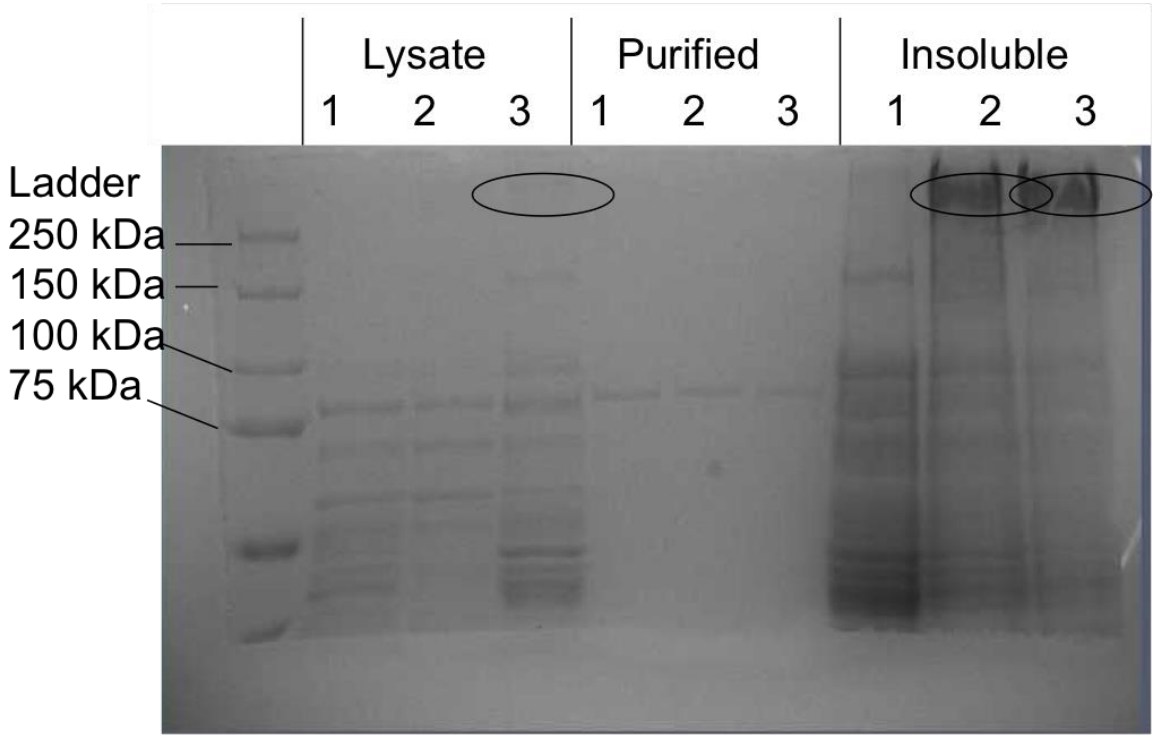


Figure 6.11: Type I FAS protein expression. Analysis of protein expression in *E. coli* BAP1 containing pBP18 (1), pBP18-FASA (2), and pBP18-FASB(3). All show expression in the crude cell lysate, his-tag purified, and insoluble protein fractions. The protein ladder with a standard as high as 250 kDa is on the left most lane.

6.3. Conclusions

A microarray study was performed comparing gene expression between FFA producing *E. coli* under carbon and phosphate limiting conditions. Upregulated non-essential pathways that generated CO_2 were identified, and used as a guide to modify *E. coli* TY05. Preliminary results indicated that eliminating flagella synthesis pathways increased FFA yield by ~10%. Multiple

knockouts or downregulation of central metabolic pathways are likely needed before a significant impact in yield is achieved.

Additionally, modifications to transcriptional regulators and introduction of a heterologous production pathway were investigated. Overexpressing FadR in a FFA producing strain of *E. coli* led to minor increases in FFA yield and significant increases in unsaturated FFA levels. Based on our findings and the number of experiments run, it is possible some other factor besides FadR led to higher titers observed elsewhere. Introduction of the Type I FAS was ultimately unsuccessful due to functional expression issues. Expression of the individual domains in a translational coupling system could identify problems in translation leading to insoluble protein formation (Mendez-Perez et al., 2012). Going forward, efforts to modulate expression or external conditions could lead to enough functional expression to pursue further work.

Overall, Chapter 6 provides significant direction moving forward. The large microarray dataset provides many directions for metabolic modification. Recent advances in the control of gene expression could allow for strict control over central metabolic pathways (Politz et al., 2013). Thus, future work will have more tools to alter intracellular carbon fluxes in an effort to maximize FFA production.

6.4. Materials and Methods

6.4.1. Strains, plasmids, enzymes, media, and bacterial cultivations

Single gene deletions were transferred by P1 transduction of phage lysates from the collection of single gene knockouts obtained from the National BioResource Project (NIG, Japan) (Baba et al., 2006). Enzymes were purchased from New England Biolabs (Ipswich, MA).

Nucleic acid purification materials were purchased from Qiagen (Venlo, Netherlands), Promega (Madison, WI), or Thermo Scientific (Waltham, MA). Chemicals were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Hampton, NH) unless otherwise specified.

6.4.2. RNA sample preparation and microarray analysis

Culture conditions for RNA sample analysis have been described in chapters 3 and 4. Carbon (Chapter 3) and phosphate (Chapter 4) limited chemostats that were run at a constant dilution rate of 0.1 h^{-1} were used as the conditions for RNA sample collection. Conditions with various carbon to phosphate ratios were tested (4 ratios for each limiting nutrient condition). Additionally, phosphate limited chemostats run at different dilution rates (0.05, 0.1, 0.2, and 0.3 h^{-1}) were run to gather further RNA samples. Samples for RNA analysis were taken at 4 different time points per steady state condition.

RNA was isolated and prepared for microarray experiments using the approach described by Lennen (Lennen et al., 2011). Briefly, 10 OD-mL of culture was added to 1.25 mL of a solution of 5% phenol in absolute ethanol that had been sitting on ice. The solution underwent centrifugation at 5000 x g for 5 minutes at 4°C, after which the supernatant was removed and the remaining pellets were snap frozen in a dry ice ethanol bath before storage at -80°C. Total RNA was isolated using a Qiagen RNeasy minikit following manufacturer's instructions. Total RNA quality was assessed using an Agilent BioAnalyzer (Agilent Technologies, Santa Clara, CA) using the Prokaryote Total RNA Nano series II chip set.

10 μg of RNA was used to synthesize cDNA for microarray hybridization. 1 μg was labeled with Cy3 based on the NimbleGen array user guide instructions. The labeled cDNA samples were hybridized to Roche NimbleGen (Madison, WI) *Escherichia coli* K-12 microarray slides containing four arrays per slide and 72,000 probes per array. A GenePix 4000B scanner

(Molecular Devices, Sunnyvale, CA) was used to scan arrays. The photomultiplier tube gain was set to achieve 10^{-5} normalized counts at a saturated intensity level. Images were then processed and .pair files were generated based on the intensities.

For analysis, the intensity data files were imported into ArrayStar software (DNASTAR, Madison, WI) and processed as described by Lennen (Lennen et al., 2011). The processed files were used to determine mean intensities across sample technical replicates, sample biological replicates, and the limitation condition as a whole. The reported values are linear changes in expression, with the statistical P-value calculated based on a modified t-test. Ratios with a change greater than 4 fold or pathways where a majority of the genes had a change greater than 3 fold and with P-values less than or equal to 0.05 were considered significant.

6.4.3. Batch shake flask cultivation of strain knockouts

The TY05 knockouts and studies with *fadR* overexpression in TY03, TY05, TY25, and TY34 were cultivated in 250 mL shake flaks containing 50 mL of MOPS minimal medium (Neidhardt et al., 1974) supplemented with 0.7% glucose, 0.276 mM potassium sulfate, and 9.5 mM ammonium chloride, but only containing 0.24 mM dibasic potassium phosphate. All cultures were grown in triplicate.

For all studies single colonies from fresh freezer stocks were used to inoculate 5 mL LB allowing for growth of at least 8 hours at 37°C. In cultures to be used in *fadR* overexpression a second cultivation step was used to condition them to the final media. These cultures were grown in the 5 mL modified MOPS minimal media spiked with 1 mM phosphate source and grown for at least 12 hours at 37°C with shaking. The initial outgrowth was used to inoculate the shake flask production cultures to an OD_{600} of 0.04. Cultures were incubated with shaking at 37°C (or 30°C if using a TY34 base strain) until each reached an OD of 0.2, at which point the cultures

were induced with 1 mM IPTG. After 40 hours of incubation, samples for fatty acid-methyl-ester (FAME) analysis, glucose consumption, and cell density were taken, processed, and quantified.

For FFA production experiments testing the effects of *fadR* overexpression, strains DH1 Δ*fadE* or MG1655 Δ*ara*Δ*fadE* containing a combination of BTE or *'tesA* expressed on p15a origin chloramphenicol resistant plasmids and pE8-fadR or ptrc99a were used in the cultivation. Strains were cultivated in a modified M9 medium containing 2.0% glucose as described in Zhang et al. (Zhang et al., 2012). The propagation scheme started with growing single colonies from fresh freezer stocks in 5 mL LB allowing for growth of at least 8 hours at 37°C. These cultures were used to inoculate 5 mL of the previously described media in a 1:100 dilution to condition the cells, allowing for growth overnight. The overnight cultures were then used to inoculate either or both 5 mL and 50 mL cultures for production runs to an OD_{600} of 0.04. Cultures were incubated with shaking at 37°C and induced with 1 mM IPTG when each OD_{600} value reached 0.8. Growth continued for 72 hours before harvesting for analysis for fatty acid methyl-ester (FAME) analysis, glucose consumption, and cell density.

Three cultures of *E. coli* BAP1 containing the plasmids pBAD18, pBAD18-FASAwt, and pBAD18-FASBwt, respectively, were grown in 50 mL shake flasks in Luria-Bertani (LB) medium overnight. The resulting cultures were used to inoculate 50 mL cultures of LB and 0.4% glycerol to an OD_{600} of 0.025. After an OD_{600} of 0.2 was reached the cultures were induced with 0.5% L-arabinose and 0.2% IPTG. Samples were collected (three for each culture) after 24 hours of growth and prepared for GC/MS analysis as described previously.

6.4.4. Product analysis

Culture samples for analyzing fatty acids via GC/MS were taken and derivatized to methyl esters as described previously (Lennen, 2010 et al.; Chapter 3). All $C_{12:0}$, $C_{12:1}$, $C_{14:0}$, and

$C_{14:1}$ were considered FFAs as past results indicate that less than 10% of these species are found in membrane fractions (Chapter 3). Supernatant samples for HPLC quantification of sugars, alcohols, and short chain carboxylic acids were collected and processed as described previously (Chapter 3). Culture samples for SYTOX flow cytometry analysis were taken and analyzed following procedures described elsewhere (Lennen et al., 2011). SYTOX is a green nucleic acid dye that only can penetrate cells when the inner membrane is not intact. Intact cells show a lower fluorescence value that is associated with staining of the cell surface (Roth, 1997).

6.4.5. Protein preparation

A of *E. coli* BAP1 containing pBP18-FASBwt or pBP18-FASA was grown in LB containing 50 μg/mL ampicillin and used to inoculate 50 mL of LB + 0.4% glycerol containing appropriate antibiotics. The cultures were grown with shaking at 30ºC. Induction with 1 mM IPTG occurred at an OD of 0.4 and the culture was harvested 6 hours after induction. The culture was centrifuged for 20 minutes at 4000g at 4ºC, the supernatant removed, and the cell pellet stored at -80ºC until processing. Cells were lysed using BugBuster from Novagen (EMD Chemicals, Germany) following manufacturer's instructions. For insoluble fraction analysis, remaining insoluble pellet was reuspended in 700 μL BugBuster lysis buffer, centrifuged at 16000 g for 10 minutes, and then had the supernatant removed. The process was repeated another time before resolubilizing the remaining pellet in 1 mL 8 M urea. For the affinity column purification, 700 μL cell lysate was added to a Pierce purification column (Thermo Scientific) with 50 μL Ni-NTA agarose resin (QIAgen) already added. The lysate and resin were mixed via gentle rocking for 60 minutes at 4ºC. The bottom end of the columns were removed and they were centrifuged at 1000g for 30 seconds and the flowthrough was removed. Through multiple additions and spins, 5 times intial lysate volme of wash buffer was added and allowed to flow

through the column via centrifugation. The final 2 spins added 50 μL elution buffer where the flowthrough was collected to obtain the purified protein. When the cell lysate, insoluble, and purified fraction were obtained, they were loaded to the same amount of protein (based on Bradford Assay results), after which 4 μL of 5x SDS loading buffer, 1 μL mercaproethanol, and BugBuster lysis buffer (enough to get the total mix to 20 μL) were added. The mix was heated for 5 minutes at 90ºC and then 10 μL was loaded into each well of the protein gel. The protein gel itself was a 6% gel (which contained 7.9 mL water, 3 mL of 30% acrylamide/Bis solution (BioRAD), 3.8 mL 1.5 M Tris (pH 8.8), 0.15 mL 10% SDS, 0.15 mL 10% ammonium persulfate, and 12 μL TEMED in the resolving gel and 3.4 mL water, 0.83 mL 30% acrylamide, 0.63 mL 1.0 M Tris (pH 6.8), 0.05 mL 10% SDS, 0.05 mL 10% ammonium persulfate, and 5 μL TEMED in the stacking gel), which was run at 100 volts until the visible band for the fatty acid synthases passed into the resolving gel. After the gel was run, it was stained in coomassie stain (~30 mL of a 1L stock solution containing 1 g coomassie R250, 100 mL glacial acetic acid, 400 mL methanol, and 500 mL water) and set on a gel rocker overnight at 4ºC. The coomassie stain was removed and a water wash was added to remove any lose stain. After the water wash, the gel was placed in a bath of destain solution (20% methanol, 10% acetic acid) until the background of the gel was clear and protein bands were easily visible.

6.5. References

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko K a, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology* **2**:1–11.

Berg HC. 2003. The rotary motor of bacterial flagella. *Annual Review of Biochemistry* **72**:19–54.

De Biase D, Tramonti A, Bossa F, Visca P. 1999. The response to stationary-phase stress conditions in *Escherichia coli*: role and regulation of the glutamic acid decarboxylase system. *Molecular Microbiology* **32**:1198–211.

Capitani G, De Biase D, Aurizi C, Gut H, Bossa F, Grütter MG. 2003. Crystal structure and functional analysis of *Escherichia coli* glutamate decarboxylase. *The EMBO Journal* **22**:4027–37.

Castanié-Cornet M-P, Treffandier H, Francez-Charlot A, Gutierrez C, Cam K. 2007. The glutamate-dependent acid resistance system in *Escherichia coli*: essential and dual role of the His-Asp phosphorelay RcsCDB/AF. *Microbiology (Reading, England)* **153**:238–46.

Chattopadhyay MK, Tabor CW, Tabor H. 2009. Polyamines are not required for aerobic growth of Escherichia coli: preparation of a strain with deletions in all of the genes for polyamine biosynthesis. *Journal of Bacteriology* **191**:5549–52.

Chilcott GS, Hughes KT. 2000. Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar typhimurium and *Escherichia coli*. *Microbiology and molecular biology reviews : MMBR* **64**:694–708.

Copeland MF, Flickinger ST, Tuson HH, Weibel DB. 2010. Studying the dynamics of flagella in multicellular communities of *Escherichia coli* by using biarsenical dyes. *Applied and Environmental Microbiology* **76**:1241–50.

Cronan JE. 2003. Bacterial membrane lipids: where do we stand? *Annual Review of Microbiology* **57**:203–24.

Dellomonaco C, Clomburg JM, Miller EN, Gonzalez R. 2011. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature* **476**:355–9.

Eichhorn E. 1999. Characterization of a Two-component Alkanesulfonate Monooxygenase from *Escherichia coli*. *Journal of Biological Chemistry* **274**:26639–26646.

Fujita Y, Matsuoka H, Hirooka K. 2007. Regulation of fatty acid metabolism in bacteria. *Molecular Microbiology* **66**:829–39.

Gauger EJ, Leatham MP, Mercado-Lubo R, Laux DC, Conway T, Cohen PS. 2007. Role of motility and the *flhDC* Operon in *Escherichia coli* MG1655 colonization of the mouse intestine. *Infection and Immunity* **75**:3315–24.

Ghrist AC, Stauffer G V. 1995. Characterization of the *Escherichia coli gcvR* Gene Encoding a Negative Regulator of gcv Expression. *Journal of Bacteriology* **177**.

Gosink KK, Häse CC. 2000. Requirements for conversion of the Na(+)-driven flagellar motor of Vibrio cholerae to the H(+)-driven motor of *Escherichia coli*. *Journal of Bacteriology* **182**:4234–40.

Han L, Doverskog M, Enfors S-O, Häggström L. 2002. Effect of glycine on the cell yield and growth rate of *Escherichia coli*: evidence for cell-density-dependent glycine degradation as determined by (13)C NMR spectroscopy. *Journal of Biotechnology* **92**:237–49.

Heath RJ, Rock CO. 1996. Regulation of Fatty Acid Elongation and Initiation by Acyl-Acyl Carrier Protein in *Escherichia coli*. *Annual Review of Microbiology* **271**:1833–1836.

Hommais F, Krin E, Coppee J-Y, Lacroix C, Yeramian E, Danchin A, Bertin P. 2004. GadE (YhiE): a novel activator involved in the response to acid environment in *Escherichia coli*. *Microbiology* **150**:61–72.

Hoover SW, Marner WD, Brownson AK, Lennen RM, Wittkopp TM, Yoshitani J, Zulkifly S, Graham LE, Chaston SD, McMahon KD, Pfleger BF. 2011. Bacterial production of free fatty acids from freshwater macroalgal cellulose. *Applied Microbiology and Biotechnology* **91**:435–46.

Hoover SW, Youngquist JT, Angart PA, Withers ST, Lennen RM, Pfleger BF. 2012. Isolation of Improved Free Fatty Acid Overproducing Strains of *Escherichia coli* via Nile Red Based High-Throughput Screening. *Environmental Progress and Sustainable Energy* **31**:17–23.

Kikuchi G, Motokawa Y, Yoshida T, Hiraga K. 2008. Review Glycine cleavage system : reaction mechanism , physiological significance , and hyperglycinemia. *Proceedings of the Japan Academy. Series B. Physical and Biological Sciences* **84**:246–263.

Kobayashi A, Hirakawa H, Hirata T, Nishino K, Yamaguchi A. 2006. Growth phase-dependent expression of drug exporters in *Escherichia coli* and its contribution to drug tolerance. *Journal of Bacteriology* **188**:5693–703.

Lennen RM, Braden DJ, West RA, Dumesic JA, Pfleger BF. 2010. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnology and Bioengineering* **106**:193–202.

Lennen RM, Kruziki M a, Kumar K, Zinkel R a, Burnum KE, Lipton MS, Hoover SW, Ranatunga DR, Wittkopp TM, Marner WD, Pfleger BF. 2011. Membrane stresses induced by overproduction of free fatty acids in *Escherichia coli*. *Applied and Environmental Microbiology* **77**:8114–28.

Lennen RM, Pfleger BF. 2013. Modulating Membrane Composition Alters Free Fatty Acid Tolerance in *Escherichia coli*. *PloS one* **8**:e54031.

Lennen RM, Politz MG, Kruziki M a, Pfleger BF. 2013. Identification of transport proteins involved in free fatty acid efflux in *Escherichia coli*. *Journal of Bacteriology* **195**:135–44.

Li M, Zhang X, Agrawal A, San K-Y. 2012. Effect of acetate formation pathway and long chain fatty acid CoA-ligase on the free fatty acid production in *E. coli* expressing acy-ACP thioesterase from *Ricinus communis*. *Metabolic Engineering* **14**:380–7.

Liu M, Durfee T, Cabrera JE, Zhao K, Jin DJ, Blattner FR. 2005. Global transcriptional programs reveal a carbon source foraging strategy by Escherichia coli. *The Journal of Biological Chemistry* **280**:15921–15927.

Ma Z, Gong S, Richard H, Tucker DL, Conway T, Foster JW. 2003. GadE (YhiE) activates glutamate decarboxylase-dependent acid resistance in *Escherichia coli* K-12. *Molecular Microbiology* **49**:1309–1320.

Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD. 2003. Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nature Biotechnology* **21**:796–802.

Mendez-Perez D, Gunasekaran S, Orler VJ, Pfleger BF. 2012. A translation-coupling DNA cassette for monitoring protein translation in *Escherichia coli*. *Metabolic Engineering* **14**:298–305.

Moen B, Janbu AO, Langsrud S, Langsrud Ø, Hobman L, Constantinidou C, Kohler A, Rudi K. 2009. Global responses of *Escherichia coli* to adverse conditions determined by microarrays and FT-IR spectroscopy. *Canadian Journal of Microbiology* **728**:714–728.

Moreau PL. 2007. The Lysine Decarboxylase CadA Protects *Escherichia coli* Starved of Phosphate against Fermentation Acids. *Journal of Bacteriology* **189**:2249–2261.

Neidhardt FC, Bloch PL, Smith DF. 1974. Culture Medium for Enterobacteria. *Journal of Bacteriology* **119**:736–747.

Nishino K, Senda Y, Yamaguchi A. 2008. The AraC-family regulator GadX enhances multidrug resistance in *Escherichia coli* by activating expression of *mdtEF* multidrug efflux genes. *Journal of Infection and Chemotherapy* **14**:23–29.

Park JH, Lee KH, Kim TY, Lee SY. 2007. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proceedings of the National Academy of Sciences of the United States of America* **104**:7797–802.

Perrenoud A, Sauer U. 2005. Impact of Global Transcriptional Regulation by ArcA , ArcB , Cra , Crp , Cya , Fnr , and Mlc on Glucose Catabolism in *Escherichia coli*. *Journal of Bacteriology* **187**:3171–3179.

van der Ploeg JR, Weiss MA, Saller E, Nashimoto H, Saito N, Kertesz MA, Leisinger T. 1996. Identification of sulfate starvation-regulated genes in *Escherichia coli* : a gene cluster involved in the utilization of taurine as a sulfur source. *Journal of Bacteriology* **178**:5438–5446.

Politz MC, Copeland MF, Pfleger BF. 2013. Artificial repressors for controlling gene expression in bacteria. *Chemical communications (Cambridge, England)* **49**:4325–7.

Ranganathan S, Tee TW, Chowdhury A, Zomorrodi AR, Yoon JM, Fu Y, Shanks J V, Maranas CD. 2012. An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metabolic Engineering* **14**:687–704.

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

Rock CO, Jackowski S. 2002. Forty years of bacterial fatty acid synthesis. *Biochemical and Biophysical Research Communications* **292**:1155–1166.

San KY, Li M, Zhang X. 2011. Bacteria and method for synthesizing fatty acids. WO2011116279.

Santos CNS, Xiao W, Stephanopoulos G. 2012. Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **2012**:13538–13543.

Sayed AK, Odom C, Foster JW. 2007. The *Escherichia coli* AraC-family regulators GadX and GadW activate *gadE*, the central activator of glutamate-dependent acid resistance. *Microbiology (Reading, England)* **153**:2584–92.

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

Stuible HP, Meurer G, Schweizer E. 1997. Heterologous expression and biochemical characterization of two functionally different type I fatty acid synthases from *Brevibacterium ammoniagenes*. *The Federation of European Biochemical Societies Journal* **247**:268–273.

Tramonti A, Visca P, Canio M De, Biase D De, Falconi M. 2002. Functional Characterization and Regulation of gadX , a Gene Encoding an AraC / XylS-Like Transcriptional Activator of the *Escherichia coli* Glutamic Acid Decarboxylase System. *Journal of Bacteriology* **184**:2603–2613.

Wada A, Igarashi K, Yoshimura S, Aimoto S, Ishihama A. 1995. Ribosome Modulation Factor: Stationary growth phase-specific inhibitor of ribosome functions from *Escherichia coli*. *Biochemical and Biophysical Research Communications* **214**:410–417.

Wood TK, González Barrios AF, Herzberg M, Lee J. 2006. Motility influences biofilm architecture in *Escherichia coli*. *Applied Microbiology and Biotechnology* **72**:361–7.

Xu P, Gu Q, Wang W, Wong L, Bower AGW, Collins CH, Koffas M a G. 2013. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications* **4**:1409.

Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD. 2012. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic Engineering* **14**:653–60.

Zhao B, Houry WA. 2010. Acid stress response in enteropathogenic gammaproteobacteria : an aptitude for survival. *Biochemistry and cell biology Biochimie et biologie cellulaire* **88**:301–314.

Zhao K, Liu M, Burgess RR. 2007. Adaptation in bacterial flagellar and motility systems: from regulon members to "foraging"-like behavior in *E. coli*. *Nucleic Acids Research* **35**:4441–52.

Zhu K, Zhang Y-M, Rock CO. 2009. Transcriptional regulation of membrane lipid homeostasis in *Escherichia coli*. *The Journal of Biological Chemistry* **284**:34880–8.

Epilogue

The production of FFA derived fuels and chemicals remains a rich area of study. The thesis focused on increasing the yield and titer of FFA and fatty alcohols. The motivation for my thesis work is the need to find viable alternatives to petroleum fuels and other petrochemicals. As renewable fuels and chemicals require a premium price relative to their petroleum-derived counterparts, further research is necessary to make costs competitive. This thesis attempted to develop a better understanding of fatty acid metabolism using classical biochemical engineering approaches and to increase the titer and yield of FFA and FFA derived chemicals produced from *E. coli* using modern metabolic engineering approaches. Chapter 2 provided a background on FFA biosynthesis and control and a review of what others had done in the field. Additionally, process optimization and modes of nutrient limitation were also reviewed as ideas that could likely lead to increased production. Using those process modifications, in Chapter 3 a model was generated to better understand FFA production under a multitude of growth conditions. A basic trend of FFA production was discovered pushing production regimes toward low growth rates to achieve the highest yield, and to provide a basis for the next chapters. In Chapter 4, the model was expanded to cover alternative limiting nutrient conditions, establishing phosphate limitation as a superior media formulation to carbon limitation. The model was applied in fed batch and was able to predict FFA production within 10%. Chapter 5 then applied metabolic engineering and fermentation techniques to increase the titer and yield of fatty alcohols. Under a phosphate limited reactor system, fatty alcohols were synthesized at a yield of 0.134 g fatty alcohol per gram glucose (a rate more than 2 fold higher any value reported in literature). The main outcome of the work was to advance knowledge of FFA and FFA derived chemical production and possibly pave the way for other researchers to expand on such work.

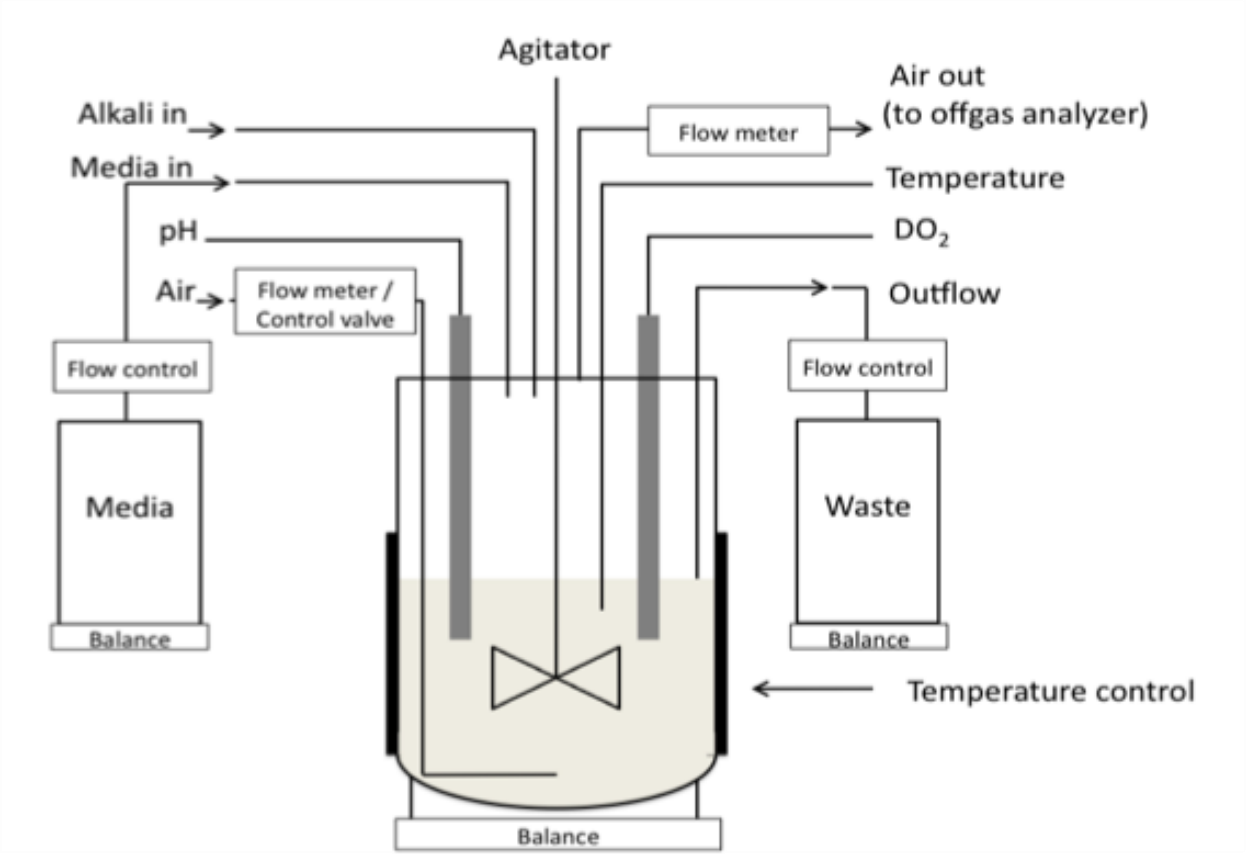
Finally, Chapter 6 focused on current work and future directions for continued research. This chapter discussed using trancriptomic microarrays to identify pathways to modify to reduce carbon dioxide flux to allow for increased carbon flux to the desired product. However, the section was limited in the pathways it was able to modify without significantly altering cellular growth. Recent advances in control of gene expression show promise in being able to broaden the context of modifying genes in a metabolic network. Also, introducing alternative pathways was discussed in Chapter 6. Despite the lack of functionality from initial studies, advances in heterologous protein expression will likely make such a strategy viable in the near future. The thesis contained limited discussion on the downstream separation and purification of the FFA derived chemicals. Based on results from Chapter 5, a dodecane overlayer is able to facilitate the transport of fatty alcohols to the supernatant facilitating a likely easier extraction procedure downstream. Despite the lack of discussion in some areas, it is hoped that the work presented in this thesis has provided a significant contribution to the scientific community and will be the used in some form for future research.

Appendices

Appendix 1: Abbreviations

Abbreviation	Term
AAR	acyl-ACP reductase
ACC	acetyl-CoA carboxylase
ACS	acyl-CoA synthetase
ADC	aldehyde decarbonylase
AHR	aldehyde reductase
ACP	acyl-carrier protein
ACR	acyl-CoA reductase
AKG	α-ketoglutarate
AT	acyl transferase
BTE	acyl-ACP specific thioesterase from *Umbellularia californica*
CAR	carboxylic acid reductase
CoA	coenzyme A
DCW	dry cell weight
DH	dehydrogenase
DO_2	dissolved oxygen
ED	Entner-Doudoroff
ER	enoyl reductase
FAEE	fatty acid ethyl ester
FAME	fatty acid methyl ester
FAS	fatty acid synthase
FFA	free fatty acid
GABA	γ-aminobutyric acid
IPTG	isopropyl β-D-1-thiogalactopyranoside
KR	keto reductase
KS	keto synthase
LB	lysogeny broth
MAACR	acyl-coA reductase from *Marinobacter aquaeolei*
MOPS	3-(N-morpholino)propanesulfonic acid
MPT	malonyl/palmitoyl transferase
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
PHA	polyhydroxyalkonoate
PPP	pentose phosphate pathway
PPT	phosphopantetheine
q_{CO2}	biomass specific carbon dioxide production rate
q_p	biomass specific extracellular product production rate
q_s	biomass specific substrate uptake rate
RBS	ribosome binding site
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
TCA	tricarboxylic acid
TE	thioesterase

Appendix 2: Chemostat set up



Monotoring/Sampling	**Controlling**
DO_2	Temperature
Offgas (CO_2, O_2, etc)	pH
Gas in and outflow	Agitation rate
Acid/Base consumption	Media composition
Mass of reactor, media storage, and waste	Liquid and gas inflow rates, liquid outflow rate
Temperature	
OD_{600}	
Biomass	
Sugars and short chain acids	
Fatty acids	

Appendix 3: Strains and Plasmids used in this study

Strain/Plasmid	Relevant Genotype/Property	Source or Reference
Strains		
E. coli K-12 MG1655	F^- λ^- *ilvG*$^-$ *rfb*-50 *rph*-1	ECGSC
E. coli DH10B	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lacZ*ΔM15 Δ*lacX*74 *recA*1 *endA*1 *araD*139 Δ(*ara, leu*)7697 *galU galK* λ^- *rpsL nupG*	Invitrogen
E. coli DH5α	F^- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*) U169 *recA1 endA1 hsdR17* (r_k-, m_k+) *pho*A *sup*E44 λ^- *thi*$^-$1 *gyr*A96 *rel*A1	Invitrogen
E. coli DH1	*endA1 recA1 gyrA96 thi-1 glnV44 relA1 hsdR17*(r_k^- m_k^+) λ^-	ECGSC
E. coli DY330	F^- λ^- rph-1 INV(*rrnD, rrnE*) Δ*lacU169 gal490 pgl*Δ*8* λcI857 Δ(*cro-bioA*)	(Yu et al., 2000)
Pseudomonas putida KT2440	Source for PP_0763	ATCC 47054™
DH1E	DH1Δ*fadE*	(Steen et al., 2010)
MHS01	MG1655 Δ*araBAD* Δ*fadE* Φ(P_{Trc}-*fadD*)	This work
MHS02	MG1655 Δ*araBAD* Δ*fadE*::trcBTE	This work
MHS03	MG1655 Δ*araBAD* Δ*fadE*::trcBTE Φ(P_{Trc}-*fadD*)	This work
MHS04	MG1655 Δ*araBAD* Δ*fadR* Δ*fadD*	This work
DE	MG1655 Δ*araBAD* Δ*fadE* Δ*fadD*	This work
E	MG1655 Δ*araBAD* Δ*fadE*	(Agnew et al., 2012)
RL08	MG1655 Δ*araBAD* Δ*fadD*	(Lennen et al., 2010)
TY01d	DY330 Δ*fadD*::trcBTE*cat*	This work
TY01c	K-12 MG1655 Δ*fadD*::trcBTE*cat*	This work
TY01	K-12 MG1655 Δ*fadD*::trcBTE	This work
TY02d	DY330 Δ*fadD*::trcBTEH204A*cat*	This work
TY02c	K-12 MG1655 Δ*fadD*::trcBTEH204A*cat*	This work
TY02	K-12 MG1655 Δ*fadD*::trcBTEH204A	This work
TY03d	DY330 Δ*fadE*::trcBTE*cat*	This work
TY03c	TY01 Δ*fadE*::trcBTE*cat*	This work
TY03	TY01 Δ*fadE*::trcBTE	This work
TY04d	DY330 Δ*fadE*::trcBTEH204A*cat*	This work
TY04c	TY02 Δ*fadE*::trcBTEH204A*cat*	This work
TY04	TY02 Δ*fadE*::trcBTEH204A	This work
TY05d	DY330 Δ*fadAB*::trcBTE*cat*	This work
TY05c	TY03 Δ*fadAB*::trcBTE*cat*	This work
TY05	TY03 Δ*fadAB*::trcBTE	This work
TY06d	DY330 Δ*fadAB*::trcBTEH204A*cat*	This work
TY06c	TY04 Δ*fadAB*::trcBTEH204A*cat*	This work
TY06	TY04 Δ*fadAB*::trcBTEH204A	This work
TY06d	DY330 Δ*fadAB*::trcBTEH204A*cat*	This work

Appendix 3: (continued)

Strain/Plasmid	Relevant Genotype/Property	Source or Reference
Strains		
TY19	MG1655 Δ*araBAD* Δ*fadR* Δ*fadE*::trcBTE	This work
TY23d	DY330 ΔackApta::trcBTEcm	This work
TY24d	DY330 ΔackApta::trcH204Acm	This work
TY23	MG1655ΔfadD::trcBTEΔfadE::trcBTE ΔackApta::trcBTE	This work
TY24	MG1655ΔfadD::trcH204AΔfadE::trcH204A ΔackApta::trcH204A	This work
TY25	MG1655ΔfadD::trcBTEΔfadE::trcBTE ΔfadAB::trcBTEΔackApta::trcBTE	This work
TY27	MG1655 Δ*araBAD* Δ*fadD* Δ*fadE*::trcBTE	This work
TY30	MG1655 Δ*araBAD* Δ*fadE*::trcBTE Δ*fadAB*::trcBTE Φ(P_{Trc}-*fadD*)	This work
TY31	MG1655 Δ*araBAD* Δ*fadE*::trcBTE Δ*fadAB*::trcBTE	This work
TY32	MG1655 Δ*araBAD* Δ*fadR* Δ*fadE*::trcBTE Δ*fadAB*::trcBTE	This work
TY33	MG1655 Δ*araBAD* Δ*fadD* Δ*fadE*::trcBTE Δ*fadAB*::trcBTE	This work
TY34	MG1655 Δ*araBAD* Δ*fadE*::trcBTE Δ*fadAB*::trcBTE Δ*ackApta*::trcBTE Φ(P_{Trc}-*fadD*)	This work
MHS05	TY05Δ*gadA*	This work
MHS06	TY05Δ*gadB*	This work
MHS07	TY05Δ*gadC*	This work
MHS08	TY05Δ*gadE*	This work
MHS09	TY05Δ*fliA*	This work
MHS10	TY05Δ*fliC*	This work
MHS11	TY05Δ*poxB*	This work
MHS12	TY05Δ*gnd*	This work
Plasmids		
pCP20	Carries yeast FLP recombinase under constituative promoter, pSC101 origin, λ*c*I857, λp_R Rep^{ts}, Amp^R, Cm^R	(Cherepanov and Wackernagel 1995)
pKD3	Template plasmid R6K gamma origin, Amp^R, Cm^R	(Datsenko and Wanner 2000)
pBTRKtrc	P_{trc} promoter, pBBR1 origin, Kan^R	This work
pUCtrc	P_{trc} promoter, pUC origin, Amp^R	This work
pACYCtrc	P_{trc} promoter, pACYC origin, Cm^R	This work
pBAD18	P_{BAD} promoter, pBR322 origin, Amp^R	(Guzman et al., 1995)
pBAD33	P_{BAD} promoter, pACYC origin, Cm^R	(Guzman et al., 1995)
pBAD35	pBBR1-MCS origin, containing *araC*-P_{BAD}-MCS-rrnB fragment from pBAD18, Kan^R	(Lennen, 2010)

Appendix 3: (continued)

Strain/Plasmid	Relevant Genotype/Property	Source or Reference
Plasmids		
pACYC-fadD6	pACYCtrc carrying *fadD6* (*M. tuberculosis*) under P_{trc} control, Cm^R	This work
pACYC-fadD	pACYCtrc carrying *fadD* under P_{trc} control, Cm^R	This work
pTrc99A	P_{Trc} promoter, pBR322 origin, Amp^R	(Amann et al., 1988)
pACYC-fadD6	pACYCtrc carrying *fadD6* (*M. tuberculosis*) under P_{trc} control, Cm^R	This work
pTRC99A-BTE	pTRC99A carrying BTE under P_{trc} control, Amp^R	(Hoover et al., 2011)
pTRC99A-BTE-H204A	pTRC99A carrying BTEH204A under P_{trc} control, Amp^R	(Hoover et al., 2011)
pBTE-int	pTRC99A-BTE carrying *cat* under constitutive control flanked by FRT sites and containing $lacI^Q$, Amp^R, Cm^R	This work
pBTE-H204A-int	pTRC99A-BTE-H204A carrying *cat* under constitutive control flanked by FRT sites and containing $lacI^Q$, Amp^R, Cm^R	This work
ACR1	pTrc99A carrying *acr1* from *Acinetobacter calcoaceticus* under P_{trc} control, Amp^R	This work
FAR6	pTrc99A carrying *far6* from *Arabidopsis thaliana* under P_{trc} control, Amp^R	This work
ptrc99a-MAACR	pTrc99A carrying MAACR from *Marinobacter aquaeolei* under Ptrc control and fused to a maltose binding protein, Amp^R	This work
pBTRK-MAACR	pBTRKtrc carrying MAACR	This work
pACYC-MAACR	pACYCtrc carrying MAACR	This work
pUCtrc-MAACR	pUCtrc carrying MAACR	This work
pBTRK-fadR	pBTRKtrc carrying *fadR*	This work
pE8-fadR	pBR322 origin plasmid carrying *fadR* under P_{BAD} control, Amp^R	(Zhang et al., 2012)
pKS1	p15a origin plasmid carrying ‘tesA under lacUV control, Cm^R	(Steen et al., 2010)
pACYC-BTE	pACYCtrc carrying BTE	This work
pACYC-‘tesA	pACYCtrc carrying *‘tesA*	This work
pBAD18-FASBwt	pBAD18 carrying wildtype *fasB* from *Corynebacterium glutamicum* under P_{BAD} control, Amp^R	This work
pBAD18-FASA	pBAD18 carrying codon optimized *fasA* from *Corynebacterium glutamicum* under P_{BAD} control, Amp^R	This work

Appendix 4: Oligonucleotide primers used

Primer Name	**Sequence (5' to 3')**
1. Forward rrnb	gaaaggttttgcaccattcgatggtgtcggtgcctaatgagtgagctaac
2. Reverse before lacI	gaaaggttttgcaccattcgatggtgtcggtgcctaatgagtgagctaac
3. Forward before lacI	atcgaatggtgcaaaacctttc
4. Reverse from rrnb to get MCS, ptrc, and lacI	gaaacgcaaaaaggccatcc
5. Gibson MAACR fwd (MAACR gib F)	acacaggaaacagaccatcaccaacaaggaccatagc
6. Gibson MAACR rev (MAACR gib R)	tcatccgccaaaacagcttatcagtgatggtgatgatgg
7. fadD fwd	gaaaagagctcggtaccaggaggtataagaattgaagaaggtttggcttaacc
8. fadD rev	gaaaagtcgactctagattatcaggctttattgtccactttgc
9. BTEack-pta_int_F	atgttaatcataaatgtcggtgtcatcatgcgctacgctcggcatgcgttcctattccgaagttcc
10. BTEack-pta_int_R	agcgcaaagctgcggatgatgacgagattactgctgctgttacatccgccaaaacagccaag
11. PP0763 fwd	gagaaagagctcggtaccaggaggtaaaataatgttgcagacacgcatcatc
12. PP0763 rev	gaaaagcctgcaggtctagattagtgatggtgatggtgatgcaacgtggaaaggaacgc
13. rev from start of MCS (ptrc gib R)	atggtctgtttcctgtgtg
14. fwd from end of MCS (ptrc gib F)	gctgttttggcggatgag
cat-cassette_fwd	ggggcatgcgtgtaggctggagctgcttc
cat-cassette_rev	cgggcatgcgttcctattccgaagttcctattctc
fadDKO_colPCR_fwd	acggcatgtatatcatttggg
fadDKO_colPCR_rev	ctttagtgggcgtcaaaaaaaac
fadEKO_colPCR_fwd	cgattgatggtaaaacggtgttgtt
fadEKO_colPCR_rev	ctgaagtgcggataaaaacagcaa
fadABKO_colPCR_fwd	ggagtgaataagtaacgcatcc
fadABKO_colPCR_rev	gctgtcgcgtcttatcgtgc
BTEint_fadD_fwd	ggttgcgatgacgacgaacacgcattttagaggtgaagaaggcatgcgttcctattccgaagttcc
BTEint_fadD_rev	cgttagatttcggtaactcatcacgaaactccaccagttacatccgccaaaacagccaag
BTEint_fadE_fwd	acaagtggtcagacctcctacaagtaaggggcttttcgttggcatgcgttcctattccgaagttcc
BTEint_fadE_rev	cgcggcttcaactttccgcactttctccggcaactttaccttacatccgccaaaacagccaag
BTEint_fadAB_fwd	tctggtacgaccagatcaccttgcggattcaggagactgaggcatgcgttcctattccgaagttcc

*Regions of homology to the *E. coli* K-12 chromosome are underlined

Appendix 4: (continued)

Primer Name	Sequence (5' to 3')
BTEint_fadAB_rev	aacccgctcaaacaccgtcgcaataccctgacccagaccgttacatccgccaaaacagcc aag
Ptrc_seq	gcataattcgtgtcgctcaag
BTE_seqR	gatggaggtagaacgatctg
$lacI^Q$_seqR	tggcaaccagcatcgcagtg
tesA_gib_trc_F	tcacacaggaaacagaccattaggaggtaaaataatggcggacacgttattgattctgg
tesA_gib_trc_R	tctctcatccgccaaaacagcttattatgagtcatgatttactaaaggctgc
BTE_gib_trc_F	tcacacaggaaacagaccattaggaggtaaaataatgactctagagtggaaaccg
BTE_gib_trc_R	tctctcatccgccaaaacagcttattaaacacgaggttccgcc
BTEseqF	accatctgctgcaactggaag
TY05fadABR	ggtattgcaggatcacggtac
TY05fadDR	atgtccatcaggttaagcatcg
TY05ackR	acaatcaataccagccagtcacc
gadEKOcPCRf	agtgcgtgatggataaatctg
gadEKOcPCRr	gacatggtgaggtcaacgac
gndKOcPCRf	gcaactttgatcgaatttcatcag
gndKOcPCRr	gttgttcgatggtcaacgcaac
fliAKOcPCRf	ctgcctggaatttacctgtaacc
fliAKOcPCRr	ccattttcatcaagcagcgtg
pgiKOcPCRf	actgcgctaagggtttacactc
pgiKOcPCRr	gattaacctcacggtatgatttcc
fliCKOcPCRf	actcccagcgatgaaatacttg
CG5 FASB 700 Seq F	ggatgtggccattcctttcc
CG5 FASB 1450 Seq F	cccatacgttgcattcaagcc
CG5 FASB 2150 Seq F	gcattgaagctcgtctcaacg
CG5 FASB 2900 Seq F	gcacttcttgctggcgttgc
CG5 FASB 3650 Seq F	gcatggttcagctcaacgacg
CGM FASB 4350 Seq F	cgttgatggccagcagtacg
CGM FASB 5100 Seq F	cgctgccaatgccatcatgg
CGM FASB 5800 Seq F	gctggctgaagcttgtcacc
CGM FASB 6570 Seq F	cgaagtcaaggctgctttcg
CGM FASB 7350 Seq F	cgacatctccttctccgtaacc
CG3 FASB 8100 Seq F	gcatccgttgctgctgaactgg
CG 7600 XmSp F	cggcccgggtttaatgcatgcagtccctacgcaagctg
C. Glut XmaI F	ccgcccgggattaaggtgagtatttcttcactgacaccgc
C. Glut 4100 MfeI R	ccgtctagaaaaatcaattgcgtttgcagccttgaggcg
C. Glut 4000 MfeI F	cggcccgggtttaatcaattgtcgacggctccctctacgc
C. Glut 7600 SphI R	ccgtctagaaaaatgcatgccaccgataccggtgc
C. Glut 7600 SphI F	cggcccgggtttaatgcagtccctacgcaagctgttcgtc
C. Glut XbaI R	ccgtctagattagccctggtagatgccgtcagc

*Regions of homology to the *E. coli* K-12 chromosome are underlined

Appendix 4: (continued)

Primer Name	Sequence (5' to 3')
fadIJkoF	caggtcagaccactttatttatttttttacaggggagtgtgaagcggcatgcgttcctattcc
fadIJkoR	ttgcaggtcagttgcagttgttttccaaaaactttccccagtgtaggctggagctgcttc
fadABkoF	tctggtacgaccagatcaccttgcggattcaggagactgagaagcggcatgcgttcctattcc
fadABkoR	aacccgctcaaacaccgtcgcaataccctgacccagaccggtgtaggctggagctgcttc
KeiokanIntF	cagtcatagccgaatagcct
KeiokanIntR	cggtgccctgaatgaactgc
qPCR_fadD_F	atcatagtggtagcatcgg
qPCR_fadD_R	atcacctgcggtcctttgac
qPCR_fadD_F2	atgctgggttactggcag
qPCR_fadD_R2	gcaggaatccttcttcatcc
qPCR_rrsA_F	ggtcttgacatccacggaag
qPCR_rrsA_R	gctcgttgcgggacttaacc
gadAKO_cpcr_F	agctttcctgcacagtctg
gadAKO_cpcr_R	ggatgataatgtgtgatgttctacg
gadBKO_cpcr_F	cgtgataattcaggagacacag
gadBKO_cpcr_R	ggcaaagaatccaagtaatgtg
poxBKO_cpcr_F	gctgctgtaagacaaaagtgg
poxBKO_cpcr_R	gcagagcattaacggtagg

*Regions of homology to the *E. coli* K-12 chromosome are underlined

Appendix 5: Full microarray comparison data between carbon and phosphate limited *E. coli* TY05 in chemostat culture at a dilution rate of 0.1 h^{-1}. (* indicates part of Pho regulon)

Locus	Gene name	Fold Change	P-value	Description
b4106*	*phnC*	637.83	7.86E-08	phosphonate/organophosphate ester transporter subunit -!- ATP-binding component of ABC superfamily
b4105*	*phnD*	574.69	9.30E-07	phosphonate/organophosphate ester transporter subunit -!- periplasmic binding component of ABC superfamily
b3615*	*yibD*	374.64	1.44E-07	predicted glycosyl transferase
b0241*	*phoE*	346.47	1.45E-04	outer membrane phosphoporin protein E
b4104*	*phnE*	233.79	7.98E-07	phosphonate/organophosphate ester transporter subunit -!- membrane component of ABC superfamily
b4099*	*phnI*	208.16	3.39E-05	carbon-phosphorus lyase complex subunit
b4098	*phnJ*	174.47	1.08E-06	carbon-phosphorus lyase complex subunit
b4100*	*phnH*	162.96	1.75E-05	carbon-phosphorus lyase complex subunit
b4097*	*phnK*	156.24	4.17E-06	carbon-phosphorus lyase complex subunit
b4102*	*phnF*	150.79	1.01E-05	predicted DNA-binding transcriptional regulator of phosphonate uptake and biodegradation
b4101*	*phnG*	126.42	7.27E-05	carbon-phosphorus lyase complex subunit
b4096*	*phnL*	102.16	8.11E-06	carbon-phosphorus lyase complex subunit
b4095*	*phnM*	98.04	3.86E-06	carbon-phosphorus lyase complex subunit
b3452*	*ugpA*	80.99	5.97E-07	glycerol-3-phosphate transporter subunit -!- membrane component of ABC superfamily
b4094*	*phnN*	80.07	4.55E-06	ribose 1,5-bisphosphokinase
b3451*	*ugpE*	59.81	1.15E-07	glycerol-3-phosphate transporter subunit -!- membrane component of ABC superfamily
b0383*	*phoA*	59.71	2.60E-07	bacterial alkaline phosphatase
b4030*	*yjbA*	56.38	6.68E-04	predicted phosphate starvation inducible protein
b3450*	*ugpC*	39.53	2.67E-08	glycerol-3-phosphate transporter subunit -!- ATP-binding component of ABC superfamily
b0384*	*psiF*	32.65	1.44E-05	conserved protein
b4093*	*phnO*	24.03	1.75E-05	predicted acyltransferase with acyl-CoA N-acyltransferase domain
b2028	*ugd*	20.97	3.30E-05	UDP-glucose 6-dehydrogenase
b3726*	*pstA*	20.66	5.70E-05	phosphate transporter subunit -!- membrane component of ABC superfamily
b3512	*gadE*	20.02	1.60E-02	DNA-binding transcriptional activator
b0399*	*phoB*	19.90	6.30E-06	DNA-binding response regulator in two-component regulatory system with PhoR (or CreC)
b3513	*mdtE*	19.58	3.23E-02	multidrug resistance efflux transporter
b3453*	*ugpB*	19.22	2.38E-07	glycerol-3-phosphate transporter subunit -!- periplasmic-binding component of ABC superfamily
b3449*	*ugpQ*	18.05	1.08E-05	glycerophosphodiester phosphodiesterase, cytosolic
b3727*	*pstC*	17.57	1.77E-06	phosphate transporter subunit -!- membrane component of ABC superfamily
b2210	*mqo*	17.41	3.14E-04	malate dehydrogenase, FAD/NAD(P)-binding domain
b0697	*kdpB*	17.11	2.82E-02	potassium translocating ATPase, subunit B
b1038	*csgF*	16.99	4.30E-03	predicted transport protein
b2048	*cpsG*	15.31	5.65E-03	phosphomannomutase
b0696	*kdpC*	14.54	2.85E-02	potassium translocating ATPase, subunit C
b1788	*yoaI*	14.45	7.09E-03	predicted protein
b1982*	*amn*	13.92	4.35E-07	AMP nucleosidase
b3728*	*pstS*	13.27	5.87E-07	phosphate transporter subunit -!- periplasmic-binding component of ABC superfamily
b1039	*csgE*	12.79	3.00E-03	predicted transport protein
b1020*	*phoH*	12.28	1.22E-04	conserved protein with nucleoside triphosphate hydrolase domain
b1040	*csgD*	12.19	3.50E-04	DNA-binding transcriptional activator in two-component regulatory system
b3517	*gadA*	11.81	3.28E-02	glutamate decarboxylase A, PLP-dependent

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b2045	*wcaK*	11.81	6.42E-04	predicted pyruvyl transferase
b0698	*kdpA*	11.81	2.35E-02	potassium translocating ATPase, subunit A
b0978	*appC*	11.77	9.50E-03	cytochrome bd-II oxidase, subunit I
b3725*	*pstB*	11.39	5.78E-08	phosphate transporter subunit -!- ATP-binding component of ABC superfamily
b0400*	*phoR*	11.38	1.16E-04	sensory histidine kinase in two-component regulatory system with PhoB
b1493	*gadB*	11.25	3.98E-02	glutamate decarboxylase B, PLP-dependent
b3724*	*phoU*	10.51	6.66E-07	negative regulator of PhoR/PhoB two-component regulator
b1528	*ydeA*	10.35	3.19E-05	predicted arabinose transporter
b1492	*gadC*	10.31	1.88E-02	predicted glutamate:gamma-aminobutyric acid antiporter
b0897	*ycaC*	8.63	8.37E-04	predicted hydrolase
b0979	*appB*	8.62	1.84E-02	cytochrome bd-II oxidase, subunit II
b2541	*hcaB*	8.40	1.80E-02	2,3-dihydroxy-2,3-dihydrophenylpropionate dehydrogenase
b3514	*mdtF*	8.39	2.78E-02	multidrug transporter, RpoS-dependent
b3614	*yibQ*	8.32	7.20E-04	predicted polysaccharide deacetylase
b0336	*codB*	8.30	8.16E-03	cytosine transporter
b4111	*proP*	8.23	2.73E-04	proline/glycine betaine transporter
b2542	*hcaD*	8.13	2.08E-03	phenylpropionate dioxygenase, ferredoxin reductase subunit
b2540	*hcaC*	8.12	1.20E-02	3-phenylpropionate dioxygenase, predicted ferredoxin subunit
b2047	*wcaJ*	8.10	3.22E-02	predicted UDP-glucose lipid carrier transferase
b3945	*gldA*	8.04	7.93E-08	glycerol dehydrogenase, NAD
b3508	*yhiD*	7.85	4.70E-02	predicted Mg(2+) transport ATPase inner membrane protein
b0337	*codA*	7.75	8.42E-04	cytosine deaminase
b4092*	*phnP*	7.46	7.07E-04	carbon-phosphorus lyase complex accessory protein
b1588	*ynfF*	7.36	3.76E-04	oxidoreductase subunit
b3506	*slp*	7.23	2.83E-02	outer membrane lipoprotein
b4244	*pyrI*	7.16	1.18E-03	aspartate carbamoyltransferase, regulatory subunit
b0207	*dkgB*	7.16	2.82E-04	2,5-diketo-D-gluconate reductase B
b3511	*hdeD*	6.80	3.59E-02	acid-resistance membrane protein
b2539	*hcaF*	6.79	1.59E-02	3-phenylpropionate dioxygenase, small (beta) subunit
b4326	*yjiD*	6.75	1.26E-02	DNA replication/recombination/repair protein
b0982	*etp*	6.66	2.28E-03	phosphotyrosine-protein phosphatase
b2044	*wcaL*	6.58	1.44E-03	predicted glycosyl transferase
b1109	*ndh*	6.38	2.82E-05	respiratory NADH dehydrogenase 2/cupric reductase
b1490	*yddV*	6.32	7.50E-04	predicted diguanylate cyclase
b2538	*hcaE*	6.28	1.49E-02	3-phenylpropionate dioxygenase, large (alpha) subunit
b4269	*yjgB*	6.28	4.16E-04	predicted alcohol dehydrogenase, Zn-dependent and NAD(P)-binding
b1202	*ycgV*	6.20	1.44E-04	predicted adhesin
b1489	*dos*	6.15	4.24E-05	cAMP phosphodiesterase, heme-regulated
b1062	*pyrC*	6.13	3.46E-05	dihydro-orotase
b3519	*treF*	6.02	3.79E-02	cytoplasmic trehalase
b4245	*pyrB*	5.80	9.06E-04	aspartate carbamoyltransferase, catalytic subunit
b3516	*gadX*	5.64	3.36E-02	DNA-binding transcriptional dual regulator
b1905	*ftn*	5.55	2.39E-02	ferritin iron storage protein (cytoplasmic)
b1846	*yebE*	5.50	1.60E-04	conserved protein
b2000	*flu*	5.43	1.36E-02	CP4-44 prophage; antigen 43 (Ag43) phase-variable biofilm formation autotransporter
b0571	*cusR*	5.41	1.03E-04	DNA-binding response regulator in two-component regulatory system with CusS
b2616	*recN*	5.39	1.93E-02	recombination and repair protein

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b2924	*mscS*	5.34	8.90E-05	mechanosensitive channel
b3364	*tsgA*	5.34	2.77E-02	predicted transporter
b2988	*gss*	5.29	1.80E-04	fused glutathionylspermidine amidase -!- glutathionylspermidine synthetase
b4217	*ytfK*	5.27	2.92E-04	conserved protein
b2508	*guaB*	5.21	4.63E-09	IMP dehydrogenase
b1983	*yeeN*	5.13	5.70E-04	conserved protein
b0124	*gcd*	5.03	5.42E-06	glucose dehydrogenase
b1021	*ycdP*	4.93	1.31E-03	predicted inner membrane protein
b1962	*yedJ*	4.93	2.93E-04	predicted phosphohydrolase
b2034	*wbbI*	4.89	4.35E-02	conserved protein
b0908	*aroA*	4.79	6.38E-04	5-enolpyruvylshikimate-3-phosphate synthetase
b0068	*tbpA*	4.78	5.34E-08	thiamin transporter subunit -!- periplasmic-binding component of ABC superfamily
b0980	*appA*	4.66	7.89E-03	phosphoanhydride phosphorylase
b2903	*gcvP*	4.63	3.64E-06	glycine decarboxylase, PLP-dependent, subunit (protein P) of glycine cleavage complex
b0869	*ybjT*	4.55	1.65E-03	conserved protein with NAD(P)-binding Rossmann-fold domain
b2301	*yfcF*	4.53	2.79E-04	predicted enzyme
b2561	*yfhH*	4.47	1.13E-02	predicted DNA-binding transcriptional regulator
b2465	*tktB*	4.43	3.47E-02	transketolase 2, thiamin-binding
b3553	*tiaE*	4.42	6.31E-05	2-keto-D-gluconate reductase (glyoxalate reductase) (2-ketoaldonate reductase)
b0885	*aat*	4.39	9.97E-05	leucyl/phenylalanyl-tRNA-protein transferase
b0842	*cmr*	4.36	3.14E-04	multidrug efflux system protein
b0003	*thrB*	4.35	7.16E-05	homoserine kinase
b1341	*ydaM*	4.31	2.02E-04	predicted diguanylate cyclase, GGDEF domain signalling protein
b2905	*gcvT*	4.27	3.08E-05	aminomethyltransferase, tetrahydrofolate-dependent, subunit (T protein) of glycine cleavage complex
b3160	*yhbW*	4.27	8.37E-04	predicted enzyme
b1657	*ydhP*	4.25	5.81E-04	predicted transporter
b3292	*zntR*	4.22	2.20E-03	DNA-binding transcriptional activator in response to Zn(II)
b0783	*moaC*	4.21	6.19E-04	molybdopterin biosynthesis, protein C
b3612	*gpmI*	4.21	2.63E-07	phosphoglycero mutase III, cofactor-independent
b4006	*purH*	4.19	5.41E-03	fused IMP cyclohydrolase -!- phosphoribosylaminoimidazolecarboxamide formyltransferase
b0401	*brnQ*	4.12	3.60E-03	predicted branched chain amino acid transporter (LIV-II)
b3351	*kefG*	4.05	1.40E-03	component of potassium effux complex with KefB
b0945	*pyrD*	4.03	3.62E-03	dihydro-orotate oxidase, FMN-linked
b0004	*thrC*	3.99	7.73E-04	threonine synthase
b3515	*gadW*	3.95	3.58E-02	DNA-binding transcriptional activator
b3866	*yihI*	3.91	2.54E-04	conserved protein
b2963	*mltC*	3.91	5.12E-04	membrane-bound lytic murein transglycosylase C
b2965	*speC*	3.89	1.05E-02	ornithine decarboxylase, constitutive
b2507	*guaA*	3.89	2.99E-07	GMP synthetase (glutamine aminotransferase)
b4377	*yjjU*	3.87	1.86E-02	predicted esterase
b3028	*mdaB*	3.85	3.69E-03	NADPH quinone reductase
b1059	*solA*	3.82	2.32E-05	N-methyltryptophan oxidase, FAD-binding
b0759	*galE*	3.80	3.10E-03	UDP-galactose-4-epimerase
b0907	*serC*	3.76	5.56E-05	3-phosphoserine/phosphohydroxythreonine aminotransferase
b0546	*ybcM*	3.74	8.24E-03	DLP12 prophage; predicted DNA-binding transcriptional regulator
b2249	*yfaY*	3.72	8.26E-05	conserved protein

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b0871	*poxB*	3.67	4.55E-03	pyruvate dehydrogenase (pyruvate oxidase), thiamin-dependent, FAD-binding
b2328	*mepA*	3.65	2.33E-04	murein DD-endopeptidase
b2904	*gcvH*	3.63	1.28E-05	glycine cleavage complex lipoylprotein
b2820	*recB*	3.62	2.37E-04	exonuclease V (RecBCD complex), beta subunit
b0104	*guaC*	3.60	1.21E-04	GMP reductase
b2302	*yfcG*	3.59	4.35E-02	predicted glutathione S-transferase
b1037	*csgG*	3.55	4.22E-02	outer membrane lipoprotein
b2086	*yegS*	3.55	2.09E-02	conserved protein
b0315	*yahA*	3.54	1.63E-03	predicted DNA-binding transcriptional regulator
b4005	*purD*	3.49	1.40E-02	phosphoribosylglycinamide synthetase phosphoribosylamine-glycine ligase
b0844	*ybjI*	3.46	4.87E-03	conserved protein
b2501	*ppk*	3.43	2.03E-04	polyphosphate kinase, component of RNA degradosome
b2351	*yfdH*	3.39	8.67E-03	CPS-53 (KpLE1) prophage; bactoprenol glucosyl transferase
b2581	*yfiF*	3.38	1.28E-05	predicted methyltransferase
b0640	*holA*	3.37	6.82E-04	DNA polymerase III, delta subunit
b4263	*yjgR*	3.34	3.74E-03	predicted ATPase
b2562	*yfhL*	3.32	2.99E-02	predicted 4Fe-4S cluster-containing protein
b0758	*galT*	3.32	1.48E-02	galactose-1-phosphate uridylyltransferase
b0002	*thrA*	3.32	1.27E-06	fused aspartokinase I -!- homoserine dehydrogenase I
b1464	*yddE*	3.29	9.26E-05	conserved protein
b1723	*pfkB*	3.28	2.31E-04	6-phosphofructokinase II
b0789	*ybhO*	3.27	1.56E-02	cardiolipin synthase 2
b3956	*ppc*	3.26	8.62E-03	phosphoenolpyruvate carboxylase
b2190	*yejO*	3.24	4.33E-03	predicted autotransporter outer membrane protein
b3826	*yigL*	3.24	1.62E-03	predicted hydrolase
b0784	*moaD*	3.22	2.52E-03	molybdopterin synthase, small subunit
b4136	*dipZ*	3.22	4.29E-04	fused thiol:disulfide interchange protein: activator of DsbC -!- conserved protein
b2498	*upp*	3.18	3.70E-02	uracil phosphoribosyltransferase
b1676	*pykF*	3.15	2.68E-04	pyruvate kinase I
b1286	*rnb*	3.14	8.87E-07	ribonuclease II
b3407	*yhgF*	3.14	1.30E-06	predicted transcriptional accessory protein
b3537	*bcsF*	3.12	1.44E-03	predicted protein
b2036	*glf*	3.12	4.53E-02	UDP-galactopyranose mutase, FAD/NAD(P)-binding
b1327	*ycjY*	3.11	3.33E-03	predicted hydrolase
b0186	*ldcC*	3.09	2.74E-02	lysine decarboxylase 2, constitutive
b0870	*ltaE*	3.09	6.63E-04	L-allo-threonine aldolase, PLP-dependent
b0052	*pdxA*	3.09	8.27E-04	4-hydroxy-L-threonine phosphate dehydrogenase, NAD-dependent
b0067	*thiP*	3.07	4.42E-02	fused thiamin transporter subunits of ABC superfamily: membrane components
b4179	*rnr*	3.07	3.09E-03	exoribonuclease R, RNase R
b1198	*dhaH*	3.06	1.21E-04	fused predicted dihydroxyacetone-specific PTS enzymes: HPr component -!- EI component
b2916	*argP*	3.04	7.53E-05	DNA-binding transcriptional activator, replication initiation inhibitor
b1693	*aroD*	3.03	8.72E-06	3-dehydroquinate dehydratase
b1287	*yciW*	3.03	5.79E-03	predicted oxidoreductase
b1709	*btuD*	3.02	6.18E-04	vitamin B12 transporter subunit : ATP-binding component of ABC superfamily
b2074	*mdtA*	3.01	1.18E-03	multidrug efflux system, subunit A
b1927	*amyA*	3.01	2.13E-02	cytoplasmic alpha-amylase

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b2029	*gnd*	2.99	2.17E-04	gluconate-6-phosphate dehydrogenase, decarboxylating
b1963	*yedR*	2.97	9.43E-03	predicted inner membrane protein
b2039	*rfbA*	2.97	1.32E-02	glucose-1-phosphate thymidylyltransferase
b2159	*nfo*	2.94	4.23E-03	endonuclease IV with intrinsic 3'-5' exonuclease activity
b1761	*gdhA*	2.94	2.94E-03	glutamate dehydrogenase, NADP-specific
b1380	*ldhA*	2.91	1.43E-04	fermentative D-lactate dehydrogenase, NAD-dependent
b4220	*ytfM*	2.91	1.46E-03	predicted outer membrane protein and surface antigen
b0239	*frsA*	2.91	7.57E-04	hydrolase, binds to enzyme IIA(Glc)
b2418	*pdxK*	2.90	1.29E-05	pyridoxal-pyridoxamine kinase/hydroxymethylpyrimidine kinase
b0639	*nadD*	2.89	1.79E-03	nicotinic acid mononucleotide adenylyltransferase, NAD(P)-dependent
b0304	*ykgC*	2.89	4.61E-02	predicted oxidoreductase with FAD/NAD(P)-binding domain and dimerization domain
b0694	*kdpE*	2.88	9.43E-03	DNA-binding response regulator in two-component regulatory system with KdpD
b0019	*nhaA*	2.87	2.68E-03	sodium-proton antiporter
b2953	*yggU*	2.87	5.62E-03	conserved protein
b3011	*yqhD*	2.86	1.75E-03	alcohol dehydrogenase, NAD(P)-dependent
b2265	*menF*	2.84	1.56E-02	isochorismate synthase 2
b3344	*yheM*	2.84	3.68E-05	predicted intraceullular sulfur oxidation protein
b2840	*ygeA*	2.83	3.13E-03	predicted racemase
b3920	*yiiQ*	2.82	9.44E-03	conserved protein
b3503	*arsC*	2.82	5.18E-03	arsenate reductase
b1668	*ydhS*	2.81	1.33E-02	conserved protein with FAD/NAD(P)-binding domain
b3350	*kefB*	2.81	6.02E-04	potassium:proton antiporter
b0754	*aroG*	2.80	1.17E-04	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase, phenylalanine repressible
b3729	*glmS*	2.80	1.14E-03	L-glutamine:D-fructose-6-phosphate aminotransferase
b0790	*ybhP*	2.78	4.65E-02	predicted DNase
b1114	*mfd*	2.78	3.52E-05	transcription-repair coupling factor
b1135	*ymfC*	2.76	8.67E-04	23S rRNA pseudouridine synthase
b0674	*asnB*	2.74	3.97E-03	asparagine synthetase B
b1687	*ydiJ*	2.74	7.38E-05	predicted FAD-linked oxidoreductase
b0695	*kdpD*	2.73	8.33E-03	fused sensory histidine kinase in two-component regulatory system with KdpE: signal sensing protein -!- sensory histidine kinase
b1538	*dcp*	2.73	4.28E-05	dipeptidyl carboxypeptidase II
b3867	*hemN*	2.73	7.19E-06	coproporphyrinogen III oxidase, SAM and NAD(P)H dependent, oxygen-independent
b0999	*cbpM*	2.72	2.04E-02	modulator of CbpA co-chaperone
b2330	*prmB*	2.71	6.36E-06	N5-glutamine methyltransferase
b3574	*yiaJ*	2.71	4.97E-02	predicted DNA-binding transcriptional repressor
b3054	*ygiF*	2.71	3.41E-04	predicted adenylate cyclase
b3440	*yhhX*	2.69	1.42E-06	predicted oxidoreductase with NAD(P)-binding Rossmann-fold domain
b3250	*mreC*	2.67	1.67E-03	cell wall structural complex MreBCD transmembrane component MreC
b1000	*cbpA*	2.66	2.26E-02	curved DNA-binding protein, DnaJ homologue that functions as a co-chaperone of DnaK
b1781	*yeaE*	2.66	9.85E-06	predicted oxidoreductase
b2509	*xseA*	2.65	6.82E-05	exonuclease VII, large subunit
b3053	*glnE*	2.64	1.54E-05	fused deadenylyltransferase -!- adenylyltransferase for glutamine synthetase
b0951	*pqiB*	2.64	6.26E-03	paraquat-inducible protein B
b0604	*dsbG*	2.62	3.96E-03	periplasmic disulfide isomerase/thiol-disulphide oxidase
b3996	*nudC*	2.59	1.12E-03	NADH pyrophosphatase
b0855	*potG*	2.58	5.04E-03	putrescine transporter subunit: ATP-binding component of ABC superfamily

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b2261	*menC*	2.58	1.11E-02	o-succinylbenzoyl-CoA synthase
b2989	*yghU*	2.57	8.24E-06	predicted S-transferase
b0049	*apaH*	2.57	9.01E-04	diadenosine tetraphosphatase
b1117	*lolD*	2.55	1.83E-03	outer membrane-specific lipoprotein transporter subunit -!- ATP-binding component of ABC superfamily
b0711	*ybgJ*	2.54	1.85E-03	predicted enzyme subunit
b1466	*narW*	2.54	3.26E-02	nitrate reductase 2 (NRZ), delta subunit (assembly subunit)
b3197	*kdsD*	2.54	1.19E-03	D-arabinose 5-phosphate isomerase
b1134	*ymfB*	2.53	1.90E-02	bifunctional thiamin pyrimidine pyrophosphate hydrolase and thiamin pyrophosphate hydrolase
b2042	*galF*	2.52	3.49E-03	predicted subunit with GalU
b3997	*hemE*	2.51	5.65E-03	uroporphyrinogen decarboxylase
b1664	*ydhQ*	2.51	3.46E-04	conserved protein
b3249	*mreD*	2.51	1.78E-02	cell wall structural complex MreBCD transmembrane component MreD
b3433	*asd*	2.50	3.10E-03	aspartate-semialdehyde dehydrogenase, NAD(P)-binding
b3428	*glgP*	2.50	6.56E-04	glycogen phosphorylase
b2829	*ptsP*	2.50	2.30E-03	fused PTS enzyme: PEP-protein phosphotransferase (enzyme I) -!- GAF domain containing protein
b3345	*yheN*	2.50	3.12E-04	predicted intracellular sulfur oxidation protein
b2962	*yggX*	2.50	5.33E-03	protein that protects iron-sulfur proteins against oxidative damage
b0160	*dgt*	2.50	7.85E-03	deoxyguanosine triphosphate triphosphohydrolase
b0243	*proA*	2.48	4.40E-04	gamma-glutamylphosphate reductase
b0638	*cobC*	2.48	1.57E-02	predicted alpha-ribazole-5'-P phosphatase
b3205	*yhbJ*	2.48	2.65E-04	predicted protein with nucleoside triphosphate hydrolase domain
b1812	*pabB*	2.47	1.48E-04	aminodeoxychorismate synthase, subunit I
b1703	*ydiA*	2.47	7.57E-04	conserved protein
b0606	*ahpF*	2.47	3.88E-05	alkyl hydroperoxide reductase, F52a subunit, FAD/NAD(P)-binding
b4129	*lysU*	2.47	1.01E-03	lysine tRNA synthetase, inducible
b3697	*yidA*	2.47	1.98E-03	predicted hydrolase
b3523	*yhjE*	2.46	2.89E-02	predicted transporter
b4386	*lplA*	2.46	1.02E-02	lipoate-protein ligase A
b1525	*yneI*	2.45	2.15E-04	predicted aldehyde dehydrogenase
b0231	*dinB*	2.45	1.24E-02	DNA polymerase IV
b2819	*recD*	2.44	8.86E-03	exonuclease V (RecBCD complex), alpha chain
b0757	*galK*	2.44	2.48E-02	galactokinase
b2494	*yfgC*	2.44	4.49E-03	predicted peptidase
b2152	*yeiB*	2.43	4.71E-03	conserved inner membrane protein
b0033	*carB*	2.42	1.73E-02	carbamoyl-phosphate synthase large subunit
b1438	*ydcQ*	2.40	1.31E-02	predicted DNA-binding transcriptional regulator
b2209	*eco*	2.40	8.45E-03	ecotin, a serine protease inhibitor
b4054	*tyrB*	2.40	7.31E-03	tyrosine aminotransferase, tyrosine-repressible, PLP-dependent
b1590	*ynfH*	2.40	5.77E-03	oxidoreductase, membrane subunit
b1845	*ptrB*	2.40	4.63E-03	protease II
b2593	*yfiH*	2.39	1.03E-02	conserved protein
b1232	*purU*	2.38	1.78E-03	formyltetrahydrofolate hydrolase
b4135	*yjdC*	2.38	1.26E-02	predicted transcriptional regulator
b0393	*rdgC*	2.38	6.48E-03	DNA-binding protein, non-specific
b0187	*yaeR*	2.38	8.72E-03	predicted lyase
b1782	*mipA*	2.38	3.91E-03	scaffolding protein for murein synthesizing machinery
b0109	*nadC*	2.38	1.43E-03	quinolinate phosphoribosyltransferase

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b2502	*ppx*	2.37	8.57E-04	exopolyphosphatase
b0047	*kefC*	2.37	2.92E-02	potassium:proton antiporter
b0188	*tilS*	2.36	5.37E-03	tRNA(Ile)-lysidine synthetase
b3385	*gph*	2.36	1.11E-03	phosphoglycolate phosphatase
b4571	*wbbL*	2.34	1.31E-02	lipopolysaccharide biosynthesis protein (pseudogene)
b2294	*yfbU*	2.34	4.16E-03	conserved protein
b2570	*rseC*	2.34	2.03E-03	RseC protein involved in reduction of the SoxR iron-sulfur cluster
b1852	*zwf*	2.33	8.24E-04	glucose-6-phosphate dehydrogenase
b1840	*yebZ*	2.31	1.39E-02	predicted inner membrane protein
b4264	*idnR*	2.30	2.88E-02	DNA-binding transcriptional repressor, 5-gluconate-binding
b1849	*purT*	2.30	3.23E-02	phosphoribosylglycinamide formyltransferase 2
b4392	*slt*	2.30	3.63E-03	lytic murein transglycosylase, soluble
b0886	*cydC*	2.30	2.28E-03	fused cysteine transporter subunits of ABC superfamily: membrane component -!- ATP-binding component
b2671	*ygaC*	2.30	3.20E-02	predicted protein
b2197	*ccmE*	2.29	6.10E-03	periplasmic heme chaperone
b2605	*yfiB*	2.29	4.86E-03	predicted outer membrane lipoprotein
b3822	*recQ*	2.29	3.17E-03	ATP-dependent DNA helicase
b1850	*eda*	2.28	4.86E-03	multifunctional 2-keto-3-deoxygluconate 6-phosphate aldolase and 2-keto-4-hydroxyglutarate aldolase and oxaloacetate decarboxylase
b0766	*ybhA*	2.28	6.86E-05	predicted hydrolase
b2321	*flk*	2.27	5.87E-04	predicted flagella assembly protein
b0632	*dacA*	2.27	4.76E-04	D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 5)
b0678	*nagB*	2.27	8.40E-04	glucosamine-6-phosphate deaminase
b3432	*glgB*	2.26	9.54E-04	1,4-alpha-glucan branching enzyme
b3998	*nfi*	2.26	9.89E-03	endonuclease V
b3621	*rfaC*	2.25	1.18E-02	ADP-heptose:LPS heptosyl transferase I
b3500	*gor*	2.25	3.20E-04	glutathione oxidoreductase
b0123	*cueO*	2.25	3.05E-02	multicopper oxidase (laccase)
b0772	*ybhC*	2.25	1.28E-03	predicted pectinesterase
b0452	*tesB*	2.24	5.18E-03	acyl-CoA thioesterase II
b0868	*ybjS*	2.24	1.35E-03	predicted NAD(P)H-binding oxidoreductase with NAD(P)-binding Rossmann-fold domain
b0962	*helD*	2.24	1.86E-04	DNA helicase IV
b4025	*pgi*	2.24	1.45E-03	glucosephosphate isomerase
b2360	*yfdQ*	2.23	7.20E-03	CPS-53 (KpLE1) prophage; predicted protein
b2913	*serA*	2.23	5.63E-03	D-3-phosphoglycerate dehydrogenase
b2153	*folE*	2.23	3.09E-04	GTP cyclohydrolase I
b4327	*yjiE*	2.23	2.82E-03	predicted DNA-binding transcriptional regulator
b0819	*ybiS*	2.22	2.89E-02	conserved protein
b0121	*speE*	2.22	4.91E-03	spermidine synthase (putrescine aminopropyltransferase)
b2893	*dsbC*	2.22	6.70E-03	protein disulfide isomerase II
b0802	*ybiJ*	2.22	2.99E-03	predicted protein
b3232	*yhcM*	2.22	1.16E-02	conserved protein with nucleoside triphosphate hydrolase domain
b1131	*purB*	2.22	1.88E-03	adenylosuccinate lyase
b3592	*yibF*	2.21	3.46E-03	predicted glutathione S-transferase
b0967	*yccW*	2.21	1.36E-04	predicted methyltransferase
b3812	*yigB*	2.21	8.20E-03	predicted hydrolase
b4137	*cutA*	2.21	1.75E-03	copper binding protein, copper sensitivity
b3651	*trmH*	2.20	1.64E-02	tRNA (Guanosine-2'-O-)-methyltransferase

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b2067	*yegE*	2.19	3.37E-02	predicted diguanylate cyclase, GGDEF domain signalling protein
b1636	*pdxY*	2.18	8.30E-03	pyridoxal kinase 2/pyridoxine kinase
b3973	*birA*	2.18	5.10E-03	bifunctional biotin-[acetylCoA carboxylase] holoenzyme synthetase and DNA-binding transcriptional repressor, bio-5'-AMP-binding
b3182	*dacB*	2.18	1.23E-03	D-alanyl-D-alanine carboxypeptidase
b2381	*ypdB*	2.18	1.24E-02	predicted response regulator in two-component system withYpdA
b1740	*nadE*	2.18	1.50E-03	NAD synthetase, NH3/glutamine-dependent
b0825	*fsaA*	2.17	9.64E-03	fructose-6-phosphate aldolase 1
b2447	*yffP*	2.16	4.23E-02	CPZ-55 prophage; predicted protein
b1261	*trpB*	2.16	1.83E-02	tryptophan synthase, beta subunit
b0854	*potF*	2.16	3.18E-03	putrescine transporter subunit: periplasmic-binding component of ABC superfamily
b2264	*menD*	2.16	3.72E-03	bifunctional 2-oxoglutarate decarboxylase and SHCHC synthase
b1344	*ydaO*	2.15	3.32E-02	predicted C32 tRNA thiolase
b3034	*nudF*	2.15	1.80E-02	ADP-ribose pyrophosphatase
b2892	*recJ*	2.15	2.94E-02	ssDNA exonuclease, 5' --> 3'-specific
b3616	*tdh*	2.14	2.42E-02	threonine 3-dehydrogenase, NAD(P)-binding
b2701	*mltB*	2.14	8.38E-03	membrane-bound lytic murein transglycosylase B
b0195	*yaeB*	2.14	9.98E-03	conserved protein
b2822	*recC*	2.14	5.24E-03	exonuclease V (RecBCD complex), gamma chain
b3248	*yhdE*	2.13	7.98E-03	conserved protein
b0712	*ybgK*	2.13	1.14E-04	predicted enzyme subunit
b0051	*ksgA*	2.13	4.08E-03	S-adenosylmethionine-6-N',N'-adenosyl (rRNA) dimethyltransferase
b3591	*selA*	2.13	1.03E-03	selenocysteine synthase
b4390	*nadR*	2.13	6.27E-03	bifunctional DNA-binding transcriptional repressor and NMN adenylyltransferase
b2592	*clpB*	2.12	2.81E-02	protein disaggregation chaperone
b2471	*yffB*	2.12	7.15E-03	conserved protein
b1412	*azoR*	2.12	7.95E-03	NADH-azoreductase, FMN-dependent
b2753	*iap*	2.12	7.19E-03	aminopeptidase in alkaline phosphatase isozyme conversion
b3640	*dut*	2.12	5.64E-03	deoxyuridinetriphosphatase
b3204	*ptsN*	2.12	1.30E-02	sugar-specific enzyme IIA component of PTS
b2198	*ccmD*	2.12	4.71E-02	cytochrome c biogenesis protein
b1199	*dhaL*	2.10	2.01E-02	dihydroxyacetone kinase, C-terminal domain
b3497	*yhiQ*	2.10	4.19E-03	predicted SAM-dependent methyltransferase
b3052	*rfaE*	2.09	2.81E-03	fused heptose 7-phosphate kinase -!- heptose 1-phosphate adenyltransferase
b2907	*ubiH*	2.09	4.66E-02	2-octaprenyl-6-methoxyphenol hydroxylase, FAD/NAD(P)-binding
b0032	*carA*	2.09	1.66E-02	carbamoyl phosphate synthetase small subunit, glutamine amidotransferase
b2622	*intA*	2.09	2.10E-03	CP4-57 prophage; integrase
b1321	*ycjX*	2.09	3.60E-03	conserved protein with nucleoside triphosphate hydrolase domain
b2263	*yfbB*	2.08	5.95E-03	predicted peptidase
b1961	*dcm*	2.08	1.76E-02	DNA cytosine methylase
b2041	*rfbB*	2.08	3.86E-03	dTDP-glucose 4,6 dehydratase, NAD(P)-binding
b0641	*rlpB*	2.08	1.52E-02	minor lipoprotein
b1800	*yeaU*	2.07	5.92E-03	predicted dehydrogenase
b0492	*ybbN*	2.07	9.39E-03	predicted thioredoxin domain-containing protein
b1971	*yedY*	2.07	4.07E-03	predicted reductase
b4374	*yjjG*	2.07	1.12E-02	predicted hydrolase
b2733	*mutS*	2.06	2.63E-02	methyl-directed mismatch repair protein
b2300	*yfcE*	2.06	1.83E-02	predicted phosphatase

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b0477	*gsk*	2.06	1.20E-02	inosine/guanosine kinase
b2097	*fbaB*	2.06	4.06E-02	fructose-bisphosphate aldolase class I
b4383	*deoB*	2.06	2.93E-02	phosphopentomutase
b3911	*cpxA*	2.06	3.49E-02	sensory histidine kinase in two-component regulatory system with CpxR
b1727	*yniC*	2.05	2.49E-02	predicted hydrolase
b2827	*thyA*	2.05	4.37E-02	thymidylate synthetase
b3212	*gltB*	2.05	4.72E-02	glutamate synthase, large subunit
b2935	*tktA*	2.05	3.50E-03	transketolase 1, thiamin-binding
b3888	*yiiD*	2.05	8.44E-03	predicted acetyltransferase
b1279	*yciS*	-2.06	2.29E-02	conserved inner membrane protein
b2323	*fabB*	-2.09	3.42E-02	3-oxoacyl-[acyl-carrier-protein] synthase I
b0956	*ycbG*	-2.10	1.79E-02	conserved protein
b0811	*glnH*	-2.13	5.09E-04	glutamine transporter subunit -!- periplasmic binding component of ABC superfamily
b1623	*add*	-2.15	2.40E-02	adenosine deaminase
b4297	*yjhG*	-2.20	3.16E-03	KpLE2 phage-like element; predicted dehydratase
b1406	*ydbC*	-2.21	2.07E-02	predicted oxidoreductase, NAD(P)-binding
b3163	*nlpI*	-2.23	2.29E-02	conserved protein
b2587	*kgtP*	-2.26	6.53E-06	alpha-ketoglutarate transporter
b3885	*yihX*	-2.29	2.62E-02	predicted hydrolase
b4213	*cpdB*	-2.30	4.34E-02	2':3'-cyclic-nucleotide 2'-phosphodiesterase
b4301	*sgcE*	-2.33	4.52E-03	KpLE2 phage-like element; predicted epimerase
b3636	*rpmG*	-2.53	2.34E-02	50S ribosomal subunit protein L33
b0156	*yadR*	-2.55	3.10E-02	conserved protein
b1516	*lsrB*	-2.59	4.23E-02	AI2 transporter -!- periplasmic-binding component of ABC superfamily
b4040	*ubiA*	-2.67	3.51E-02	p-hydroxybenzoate octaprenyltransferase
b2288	*nuoA*	-2.68	2.64E-04	NADH:ubiquinone oxidoreductase, membrane subunit A
b0146	*sfsA*	-2.70	3.52E-02	predicted DNA-binding transcriptional regulator
b4021	*pepE*	-2.77	2.42E-02	(alpha)-aspartyl dipeptidase
b0722	*sdhD*	-2.83	3.17E-02	succinate dehydrogenase, membrane subunit, binds cytochrome b556
b0145	*dksA*	-2.86	1.74E-02	DNA-binding transcriptional regulator of rRNA transcription, DnaK suppressor protein
b2013	*yeeE*	-2.86	1.18E-02	predicted inner membrane protein
b3995	*rsd*	-2.88	3.52E-02	stationary phase protein, binds sigma 70 RNA polymerase subunit
b2095	*gatZ*	-2.89	1.23E-07	D-tagatose 1,6-bisphosphate aldolase 2, subunit
b1237	*hns*	-2.91	4.47E-02	global DNA-binding transcriptional dual regulator H-NS
b1612	*fumA*	-2.94	2.15E-02	fumarate hydratase (fumarase A), aerobic Class I
b3066	*dnaG*	-2.97	2.41E-03	DNA primase
b1891	*flhC*	-2.97	1.99E-02	DNA-binding transcriptional dual regulator with FlhD
b0954	*fabA*	-3.00	4.80E-02	beta-hydroxygecanoyl thioester dehydratase
b2096	*gatY*	-3.12	1.21E-06	D-tagatose 1,6-bisphosphate aldolase 2, catalytic subunit
b4015	*aceA*	-3.16	1.22E-04	isocitrate lyase
b0389	*yaiA*	-3.16	3.77E-02	predicted protein
b1853	*yebK*	-3.19	3.14E-02	predicted DNA-binding transcriptional regulator
b1611	*fumC*	-3.20	1.27E-02	fumarate hydratase (fumarase C),aerobic Class II
b4014	*aceB*	-3.25	1.96E-05	malate synthase A
b1818	*manY*	-3.32	4.84E-05	mannose-specific enzyme IIC component of PTS
b1823	*cspC*	-3.33	1.18E-03	stress protein, member of the CspA-family
b0963	*mgsA*	-3.37	2.76E-03	methylglyoxal synthase
b3739	*atpI*	-3.43	6.91E-03	ATP synthase, membrane-bound accesory subunit

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b3603	*lldP*	-3.44	2.23E-02	L-lactate permease
b3002	*yqhA*	-3.44	1.24E-03	conserved inner membrane protein
b2309	*hisJ*	-3.60	1.59E-02	histidine/lysine/arginine/ornithine transporter subunit -!- periplasmic-binding component of ABC superfamily
b4488	*ilvG*	-3.64	7.73E-06	acetolactate synthase II, large subunit (pseudogene)
b0679	*nagE*	-3.69	1.92E-02	fused N-acetyl glucosamine specific PTS enzymes: IIC component -!- IIB component -!- IIA component
b1824	*yobF*	-3.70	3.44E-03	predicted protein
b0721	*sdhC*	-3.74	9.59E-04	succinate dehydrogenase, membrane subunit, binds cytochrome b556
b1819	*manZ*	-3.75	1.18E-04	mannose-specific enzyme IID component of PTS
b1594	*dgsA*	-3.76	6.82E-03	DNA-binding transcriptional repressor
b4529	*ydbJ*	-3.84	3.13E-03	predicted protein
b1332	*ynaJ*	-3.96	2.19E-02	predicted inner membrane protein
b2182	*bcr*	-3.96	3.24E-02	bicyclomycin/multidrug efflux system
b3669	*uhpA*	-3.99	9.20E-05	DNA-binding response regulator in two-component regulatory system wtih UhpB
b2310	*argT*	-4.00	5.76E-04	lysine/arginine/ornithine transporter subunit -!- periplasmic-binding component of ABC superfamily
b0440	*hupB*	-4.06	3.41E-03	HU, DNA-binding transcriptional regulator, beta subunit
b1423	*ydcJ*	-4.08	1.36E-03	conserved protein
b3092	*uxaC*	-4.19	4.08E-02	uronate isomerase
b4062	*soxS*	-4.26	2.38E-02	DNA-binding transcriptional dual regulator
b3769	*ilvM*	-4.33	1.46E-02	acetolactate synthase II, small subunit
b3750	*rbsC*	-4.44	1.57E-05	D-ribose transporter subunit -!- membrane component of ABC superfamily
b3337	*bfd*	-4.51	6.23E-03	bacterioferritin-associated ferredoxin
b3752	*rbsK*	-4.58	4.57E-02	ribokinase
b0411	*tsx*	-4.67	1.56E-02	nucleoside channel, receptor of phage T6 and colicin K
b3006	*exbB*	-4.67	4.49E-02	membrane spanning protein in TonB-ExbB-ExbD complex
b2012	*yeeD*	-4.70	7.55E-03	conserved protein
b3260	*dusB*	-4.70	5.99E-03	tRNA-dihydrouridine synthase B
b2805	*fucR*	-4.80	1.82E-02	DNA-binding transcriptional activator
b4016	*aceK*	-4.88	4.66E-02	isocitrate dehydrogenase kinase/phosphatase
b1929	*yedE*	-4.91	4.26E-03	predicted inner membrane protein
b2597	*yfiA*	-4.99	3.97E-03	cold shock protein associated with 30S ribosomal subunit
b1427	*rimL*	-4.99	1.04E-02	ribosomal-protein-L7/L12-serine acetyltransferase
b4124	*dcuR*	-5.22	4.37E-02	DNA-binding response regulator in two-component regulatory system with DcuS
b3926	*glpK*	-5.28	1.07E-07	glycerol kinase
b4322	*uxuA*	-5.72	4.06E-02	mannonate hydrolase
b1738	*chbB*	-6.10	9.91E-04	N,N'-diacetylchitobiose-specific enzyme IIB component of PTS
b1901	*araF*	-6.28	1.03E-02	L-arabinose transporter subunit -!- periplasmic-binding component of ABC superfamily
b3427	*yzgL*	-6.28	6.76E-03	predicted protein
b1817	*manX*	-6.51	5.51E-05	fused mannose-specific PTS enzymes: IIA component -!- IIB component
b4266	*idnO*	-6.59	4.50E-03	5-keto-D-gluconate-5-reductase
b3261	*fis*	-6.62	1.57E-03	global DNA-binding transcriptional dual regulator
b0651	*rihA*	-6.74	1.19E-03	ribonucleoside hydrolase 1
b0459	*maa*	-6.74	5.26E-03	maltose O-acetyltransferase
b4227	*ytfQ*	-6.79	2.02E-03	predicted sugar transporter subunit: periplasmic-binding component of ABC superfamily
b0553	*nmpC*	-6.94	1.13E-02	DLP12 prophage; truncated outer membrane porin (pseudogene)
b3671	*ilvB*	-7.16	5.61E-03	acetolactate synthase I, large subunit

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b3128	*garD*	-7.20	2.77E-02	(D)-galactarate dehydrogenase
b3081	*fadH*	-7.47	4.63E-02	2,4-dienoyl-CoA reductase, NADH and FMN-linked
b2247	*yfaW*	-7.55	1.07E-02	predicted enolase
b0344	*lacZ*	-7.70	6.33E-07	beta-D-galactosidase
b3418	*malT*	-7.82	9.45E-04	fused conserved protein -!- DNA-binding transcriptional activator, maltotriose-ATP-binding
b3570	*bax*	-8.23	3.47E-02	conserved protein
b2975	*yghK*	-8.33	4.37E-02	glycolate transporter
b3927	*glpF*	-8.62	3.88E-13	glycerol facilitator
b2341	*yfcX*	-8.80	3.66E-02	fused enoyl-CoA hydratase and epimerase and isomerase -!- 3-hydroxyacyl-CoA dehydrogenase
b2094	*gatA*	-8.81	1.75E-06	galactitol-specific enzyme IIA component of PTS
b3751	*rbsB*	-9.22	1.54E-06	D-ribose transporter subunit -!- periplasmic-binding compoent of ABC superfamily
b4267	*idnD*	-9.39	1.91E-03	L-idonate 5-dehydrogenase, NAD-binding
b4303	*sgcQ*	-9.59	1.62E-03	KpLE2 phage-like element; predicted nucleoside triphosphatase
b4367	*fhuF*	-9.69	2.30E-02	ferric iron reductase involved in ferric hydroximate transport
b3925	*glpX*	-9.86	7.18E-03	fructose 1,6-bisphosphatase II
b3566	*xylF*	-10.06	5.88E-03	D-xylose transporter subunit -!- periplasmic-binding component of ABC superfamily
b2844	*yqeF*	-10.07	6.76E-03	predicted acyltransferase
b2775	*yqcE*	-10.62	3.07E-03	predicted transporter
b3403	*pck*	-10.75	1.87E-05	phosphoenolpyruvate carboxykinase
b4216	*ytfJ*	-10.81	1.10E-02	predicted transcriptional regulator
b4512	*ybdD*	-10.87	3.96E-02	conserved protein
b2342	*yfcY*	-11.12	3.06E-02	beta-ketoacyl-CoA thiolase, anaerobic, subunit
b4139	*aspA*	-11.41	2.49E-02	aspartate ammonia-lyase
b1376	*uspF*	-11.46	9.87E-04	stress-induced protein, ATP-binding protein
b1930	*yedF*	-12.01	3.98E-02	conserved protein
b0791	*ybhQ*	-12.31	6.82E-03	predicted inner membrane protein
b0343	*lacY*	-13.02	1.74E-04	lactose/galactose transporter
b1384	*feaR*	-13.02	1.04E-02	DNA-binding transcriptional dual regulator
b1415	*aldA*	-13.16	1.72E-06	aldehyde dehydrogenase A, NAD-linked
b2537	*hcaR*	-13.49	6.80E-03	DNA-binding transcriptional activator of 3-phenylpropionic acid catabolism
b1498	*ydeN*	-13.67	1.01E-02	conserved protein
b2243	*glpC*	-13.71	2.83E-03	sn-glycerol-3-phosphate dehydrogenase (anaerobic), small subunit
b2092	*gatC*	-14.18	2.85E-05	galactitol-specific enzyme IIC component of PTS
b4304	*sgcC*	-14.40	2.20E-03	KpLE2 phage-like element; predicted phosphotransferase enzyme IIC component
b4305	*sgcX*	-14.48	4.25E-03	KpLE2 phage-like element; predicted endoglucanase with Zn-dependent exopeptidase domain
b2344	*fadL*	-14.94	1.80E-02	long-chain fatty acid outer membrane transporter
b0880	*cspD*	-15.82	1.22E-03	cold shock protein homolog
b3001	*yghZ*	-17.12	1.58E-03	aldo-keto reductase
b4069	*acs*	-18.29	9.51E-06	acetyl-CoA synthetase
b3670	*ilvN*	-18.68	1.25E-02	acetolactate synthase I, small subunit
b3672	*ivbL*	-19.04	3.14E-02	ilvB operon leader peptide
b4188	*yjfN*	-20.29	1.98E-02	predicted protein
b4189	*yjfO*	-21.01	2.88E-03	conserved protein
b2093	*gatB*	-21.03	5.49E-05	galactitol-specific enzyme IIB component of PTS
b2426	*ucpA*	-21.63	2.64E-02	predicted oxidoredutase, sulfate metabolism protein
b0610	*rnk*	-22.08	2.83E-02	regulator of nucleoside diphosphate kinase

Appendix 5: (continued)

Locus	Gene name	Fold Change	P-value	Description
b0342	*lacA*	-22.35	7.68E-05	thiogalactoside acetyltransferase
b4565	*sgcB*	-23.11	7.40E-03	predicted enzyme IIB component of PTS
b1205	*ychH*	-25.89	5.22E-05	predicted inner membrane protein
b0598	*cstA*	-26.86	2.98E-03	carbon starvation protein
b2242	*glpB*	-27.46	7.33E-04	sn-glycerol-3-phosphate dehydrogenase (anaerobic), membrane anchor subunit
b4067	*actP*	-29.05	6.36E-04	acetate transporter
b2241	*glpA*	-35.62	3.58E-03	sn-glycerol-3-phosphate dehydrogenase (anaerobic), large subunit, FAD/NAD(P)-binding
b1002	*agp*	-35.69	1.97E-07	glucose-1-phosphatase/inositol phosphatase
b3748	*rbsD*	-36.42	9.23E-06	predicted cytoplasmic sugar-binding protein
b2239	*glpQ*	-39.20	1.46E-03	periplasmic glycerophosphodiester phosphodiesterase
b4068	*yjcH*	-46.78	5.98E-03	conserved inner membrane protein involved in acetate transport
b3528	*dctA*	-61.16	1.80E-02	C4-dicarboxylic acid, orotate and citrate transporter
b2240	*glpT*	-62.23	3.39E-04	sn-glycerol-3-phosphate transporter
b2091	*gatD*	-89.23	2.26E-03	galactitol-1-phosphate dehydrogenase, Zn-dependent and NAD(P)-binding
b3666	*uhpT*	-108.56	1.29E-02	hexose phosphate transporter
b3708	*tnaA*	-162.32	1.57E-02	tryptophanase/L-cysteine desulfhydrase, PLP-dependent

Appendix 6: Statistically significant (change > 2, P < 0.05) microarray comparison data between phosphate limited *E. coli* TY05 in chemostat culture at dilution rates of 0.05 h^{-1} and 0.3 h^{-1}.

Locus	Gene name	Fold Change	P-value	Description
b1886	*tar*	115.81	8.32E-04	methyl-accepting chemotaxis protein II
b1923	*fliC*	68.27	7.71E-04	flagellar filament structural protein (flagellin)
b1076	*flgE*	58.02	1.88E-03	flagellar hook protein
b1885	*tap*	56.44	6.87E-04	methyl-accepting protein IV
b1924	*fliD*	55.00	1.33E-03	flagellar filament capping protein
b1073	*flgB*	52.20	5.73E-04	flagellar component of cell-proximal portion of basal-body rod
b1888	*cheA*	41.20	7.58E-04	fused chemotactic sensory histidine kinase in two-component regulatory system with CheB and CheY: sensory histidine kinase -!- signal sensing protein
b1075	*flgD*	39.40	1.12E-03	flagellar hook assembly protein
b1925	*fliS*	36.65	2.04E-03	flagellar protein potentiates polymerization
b1077	*flgF*	35.12	7.26E-04	flagellar component of cell-proximal portion of basal-body rod
b1074	*flgC*	34.83	3.40E-04	flagellar component of cell-proximal portion of basal-body rod
b1890	*motA*	34.65	5.98E-05	proton conductor component of flagella motor
b1194	*ycgR*	33.56	1.44E-03	protein involved in flagellar function
b1887	*cheW*	33.29	3.51E-03	purine-binding chemotaxis protein
b3525	*yhjH*	33.05	1.77E-03	EAL domain containing protein involved in flagellar function
b1946	*fliN*	32.97	1.28E-03	flagellar motor switching and energizing component
b1889	*motB*	29.82	4.21E-03	protein that enables flagellar motor rotation
b1944	*fliL*	27.67	4.73E-04	flagellar biosynthesis protein
b1082	*flgK*	27.08	1.27E-04	flagellar hook-filament junction protein 1
b4355	*tsr*	24.36	2.93E-03	methyl-accepting chemotaxis protein I, serine sensor receptor
b1079	*flgH*	24.09	2.86E-04	flagellar protein of basal-body outer-membrane L ring
b1080	*flgI*	23.16	2.87E-03	predicted flagellar basal body protein
b1921	*fliZ*	22.14	3.62E-04	predicted regulator of FliA activity
b1940	*fliH*	21.68	4.32E-04	flagellar biosynthesis protein
b1882	*cheY*	20.32	2.32E-03	chemotaxis regulator transmitting signal to flagellar motor component
b1922	*fliA*	19.59	1.66E-03	RNA polymerase, sigma 28 (sigma F) factor
b1884	*cheR*	17.51	3.58E-04	chemotaxis regulator, protein-glutamate methyltransferase
b1351	*racC*	17.25	2.09E-03	Rac prophage; predicted protein
b1939	*fliG*	16.60	1.13E-03	flagellar motor switching and energizing component
b1945	*fliM*	16.43	1.17E-03	flagellar motor switching and energizing component
b1883	*cheB*	15.94	2.47E-03	fused chemotaxis regulator -!- protein-glutamate methylesterase in two-component regulatory system with CheA
b1926	*fliT*	15.89	3.74E-04	predicted chaperone
b1078	*flgG*	15.57	8.85E-04	flagellar component of cell-distal portion of basal-body rod
b4526	*ydaE*	14.68	8.11E-03	Rac prophage; conserved protein
b1072	*flgA*	13.93	5.69E-04	assembly protein for flagellar basal-body periplasmic P ring
b1083	*flgL*	13.06	4.30E-04	flagellar hook-filament junction protein
b1881	*cheZ*	12.35	1.74E-03	chemotaxis regulator, protein phosphatase for CheY
b1938	*fliF*	11.67	3.05E-03	flagellar basal-body MS-ring and collar protein
b1742	*ydjR*	11.48	1.22E-03	conserved protein
b1081	*flgJ*	10.72	6.69E-04	muramidase
b1070	*flgN*	10.65	6.98E-03	export chaperone for FlgK and FlgL
b0936	*ssuA*	9.86	3.14E-02	alkanesulfonate transporter subunit -!- periplasmic-binding component of ABC superfamily
b1943	*fliK*	9.81	1.23E-03	flagellar hook-length control protein
b1942	*fliJ*	9.80	1.77E-02	flagellar protein

Appendix 6: (continued)

Locus	Gene name	Fold Change	P-value	Description
b1947	*fliO*	9.73	8.57E-04	flagellar biosynthesis protein
b1071	*flgM*	8.60	6.95E-04	anti-sigma factor for FliA (sigma 28)
b1355	*ydaG*	7.27	2.61E-05	Rac prophage; predicted protein
b1941	*fliI*	7.24	2.99E-02	flagellum-specific ATP synthase
b1880	*flhB*	6.82	4.85E-03	predicted flagellar export pore protein
b2252	*ais*	6.55	6.67E-03	conserved protein
b1349	*recT*	6.30	2.51E-03	Rac prophage; recombination and repair protein
b4254	*argI*	6.25	9.27E-04	ornithine carbamoyltransferase 1
b4110	*yjcZ*	5.97	2.94E-03	conserved protein
b1350	*recE*	5.85	5.35E-04	Rac prophage; exonuclease VIII, 5' -> 3' specific dsDNA exonuclease
b0937	*ssuE*	5.70	1.29E-03	NAD(P)H-dependent FMN reductase
b1566	*flxA*	5.64	3.10E-03	Qin prophage; predicted protein
b1949	*fliQ*	5.47	5.59E-04	flagellar biosynthesis protein
b1348	*lar*	5.47	4.54E-02	Rac prophage; restriction alleviation protein
b4527	*ydaF*	5.21	4.04E-03	Rac prophage; predicted protein
b0935	*ssuD*	5.18	1.29E-02	alkanesulfonate monooxygenase, FMNH(2)-dependent
b1879	*flhA*	5.11	1.77E-04	predicted flagellar export pore protein
b2093	*gatB*	4.99	1.15E-02	galactitol-specific enzyme IIB component of PTS
b2156	*lysP*	4.94	2.09E-04	lysine transporter
b1421	*trg*	4.93	2.96E-03	methyl-accepting chemotaxis protein III, ribose and galactose sensor receptor
b4114	*eptA*	4.92	2.92E-04	predicted metal dependent hydrolase
b1044	*ymdA*	4.89	4.75E-03	predicted protein
b1358	*ydaT*	4.87	4.57E-03	Rac prophage; predicted protein
b1357	*ydaS*	4.65	4.54E-02	Rac prophage; predicted DNA-binding transcriptional regulator
b1936	*intG*	4.56	2.39E-03	predicted defective phage integrase (pseudogene)
b1937	*fliE*	4.53	2.00E-02	flagellar basal-body component
b2175	*spr*	4.25	1.63E-02	predicted peptidase, outer membrane lipoprotein
b2255	*yfbG*	4.16	5.24E-05	fused UDP-L-Ara4N formyltransferase -!- UDP-GlcA C-4'-decarboxylase
b1346	*ydaQ*	4.12	1.95E-02	Rac prophage; conserved protein
b3917	*sbp*	3.95	2.42E-03	sulfate transporter subunit -!- periplasmic-binding component of ABC superfamily
b0763	*modA*	3.90	3.89E-03	molybdate transporter subunit -!- periplasmic-binding component of ABC superfamily
b3072	*aer*	3.80	8.58E-04	fused signal transducer for aerotaxis sensory component -!- methyl accepting chemotaxis component
b0336	*codB*	3.65	7.15E-03	cytosine transporter
b1124	*potC*	3.58	1.96E-02	polyamine transporter subunit -!- membrane component of ABC superfamily
b2703	*srlE*	3.58	6.93E-04	glucitol/sorbitol-specific enzyme IIB component of PTS
b1125	*potB*	3.58	1.25E-04	polyamine transporter subunit -!- membrane component of ABC superfamily
b1359	*ydaU*	3.56	7.63E-03	Rac prophage; conserved protein
b2014	*yeeF*	3.48	2.32E-03	predicted amino-acid transporter
b0366	*tauB*	3.48	1.25E-03	taurine transporter subunit -!- ATP-binding component of ABC superfamily
b0365	*tauA*	3.48	1.46E-04	taurine transporter subunit -!- periplasmic-binding component of ABC superfamily
b2240	*glpT*	3.40	3.02E-04	sn-glycerol-3-phosphate transporter
b1817	*manX*	3.39	1.57E-05	fused mannose-specific PTS enzymes: IIA component -!- IIB component
b1904	*yecR*	3.37	6.64E-03	predicted protein
b0342	*lacA*	3.33	5.71E-03	thiogalactoside acetyltransferase
b2669	*stpA*	3.31	7.71E-03	DNA binding protein, nucleoid-associated
b2253	*yfbE*	3.29	1.17E-05	uridine 5'-(beta-1-threo-pentapyranosyl-4-ulose diphosphate) aminotransferase, PLP-dependent

Appendix 6: (continued)

Locus	Gene name	Fold Change	P-value	Description
b2092	*gatC*	3.27	7.24E-05	galactitol-specific enzyme IIC component of PTS
b2678	*proW*	3.27	8.68E-05	glycine betaine transporter subunit -!- membrane component of ABC superfamily
b3958	*argC*	3.27	8.76E-04	N-acetyl-gamma-glutamylphosphate reductase, NAD(P)-binding
b2677	*proV*	3.24	4.50E-03	glycine betaine transporter subunit -!- ATP-binding compoent of ABC superfamily
b2750	*cysC*	3.22	3.69E-02	adenosine 5'-phosphosulfate kinase
b1948	*fliP*	3.22	1.21E-02	flagellar biosynthesis protein
b0211	*mltD*	3.18	9.80E-04	predicted membrane-bound lytic murein transglycosylase D
b1892	*flhD*	3.16	1.19E-02	DNA-binding transcriptional dual regulator with FlhC
b1216	*chaA*	3.16	6.41E-03	calcium/sodium:proton antiporter
b2927	*epd*	3.13	4.58E-02	D-erythrose 4-phosphate dehydrogenase
b0934	*ssuC*	3.12	4.58E-04	alkanesulfonate transporter subunit -!- membrane component of ABC superfamily
b2256	*yfbH*	3.09	3.40E-04	conserved protein
b2818	*argA*	3.08	1.43E-05	fused acetylglutamate kinase homolog (inactive) -!- amino acid N-acetyltransferase
b3959	*argB*	3.07	5.42E-04	acetylglutamate kinase
b2679	*proX*	3.07	2.77E-03	glycine betaine transporter subunit -!- periplasmic-binding component of ABC superfamily
b4034	*malE*	3.06	6.64E-03	maltose transporter subunit -!- periplasmic-binding component of ABC superfamily
b3788	*rffG*	3.02	3.91E-03	dTDP-glucose 4,6-dehydratase
b2497	*uraA*	2.94	1.14E-03	uracil transporter
b3774	*ilvC*	2.94	2.08E-05	ketol-acid reductoisomerase, NAD(P)-binding
b2174	*yeiU*	2.93	1.81E-03	undecaprenyl pyrophosphate phosphatase
b0273	*argF*	2.91	3.70E-04	CP4-6 prophage; ornithine carbamoyltransferase 2, chain F
b0961	*yccF*	2.88	4.18E-04	conserved inner membrane protein
b0556	*rzpD*	2.87	2.45E-02	DLP12 prophage; predicted murein endopeptidase
b3261	*fis*	2.86	1.52E-04	global DNA-binding transcriptional dual regulator
b2685	*emrA*	2.86	5.93E-03	multidrug efflux system
b4557	*yidD*	2.83	2.50E-03	predicted protein
b1878	*flhE*	2.81	1.93E-03	conserved protein
b3769	*ilvM*	2.81	1.37E-04	acetolactate synthase II, small subunit
b3260	*dusB*	2.76	2.72E-03	tRNA-dihydrouridine synthase B
b2951	*yggS*	2.76	5.49E-03	predicted enzyme
b0287	*yagU*	2.74	6.14E-03	conserved inner membrane protein
b1907	*tyrP*	2.74	7.56E-03	tyrosine transporter
b0557	*borD*	2.74	1.05E-02	DLP12 prophage; predicted lipoprotein
b2939	*yqgB*	2.74	1.62E-03	predicted protein
b1126	*potA*	2.72	8.13E-03	polyamine transporter subunit -!- ATP-binding component of ABC superfamily
b0777	*bioC*	2.71	3.33E-02	predicted methltransferase, enzyme of biotin synthesis
b2601	*aroF*	2.70	3.17E-03	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase, tyrosine-repressible
b1625	*ydgT*	2.69	1.32E-02	predicted regulator
b0764	*modB*	2.66	2.22E-03	molybdate transporter subunit -!- membrane component of ABC superfamily
b2254	*yfbF*	2.66	7.43E-03	undecaprenyl phosphate-L-Ara4FN transferase
b3785	*wzzE*	2.65	9.80E-05	Entobacterial Common Antigen (ECA) polysaccharide chain length modulation protein
b1891	*flhC*	2.63	7.64E-04	DNA-binding transcriptional dual regulator with FlhD
b3456	*livM*	2.63	3.80E-04	leucine/isoleucine/valine transporter subunit -!- membrane component of ABC superfamily
b2686	*emrB*	2.62	5.48E-04	multidrug efflux system protein

Appendix 6: (continued)

Locus	Gene name	Fold Change	P-value	Description
b4016	*aceK*	2.60	5.40E-03	isocitrate dehydrogenase kinase/phosphatase
b3570	*bax*	2.60	1.86E-02	conserved protein
b3654	*yicE*	2.59	6.87E-05	predicted transporter
b2094	*gatA*	2.58	6.95E-04	galactitol-specific enzyme IIA component of PTS
b3687	*ibpA*	2.57	4.43E-02	heat shock chaperone
b3955	*yijP*	2.56	1.04E-03	conserved inner membrane protein
b4064	*yjcD*	2.55	5.66E-03	predicted permease
b0776	*bioF*	2.54	4.14E-03	8-amino-7-oxononanoate synthase
b1345	*intR*	2.53	3.01E-02	Rac prophage; integrase
b0553	*nmpC*	2.52	1.68E-03	DLP12 prophage; truncated outer membrane porin (pseudogene)
b2574	*nadB*	2.49	1.24E-05	quinolinate synthase, L-aspartate oxidase (B protein) subunit
b1191	*cvrA*	2.48	2.09E-02	predicted cation/proton antiporter
b4036	*lamB*	2.48	1.71E-04	maltose outer membrane porin (maltoporin)
b3258	*panF*	2.47	6.94E-03	pantothenate:sodium symporter
b2702	*srlA*	2.43	2.54E-03	glucitol/sorbitol-specific enzyme IIC component of PTS
b1453	*ansP*	2.43	6.86E-03	L-asparagine transporter
b0905	*ycaO*	2.42	3.73E-03	conserved protein
b3493	*pitA*	2.41	1.89E-04	phosphate transporter, low-affinity
b2498	*upp*	2.39	2.21E-04	uracil phosphoribosyltransferase
b1101	*ptsG*	2.38	7.23E-03	fused glucose-specific PTS enzymes: IIBcomponent -!- IIC component
b0914	*msbA*	2.38	4.48E-03	fused lipid transporter subunits of ABC superfamily: membrane component -!- ATP-binding component
b3457	*livH*	2.38	1.70E-03	leucine/isoleucine/valine transporter subunit -!- membrane component of ABC superfamily
b1360	*ydaV*	2.38	4.46E-04	Rac prophage; predicted DNA replication protein
b3528	*dctA*	2.37	4.72E-03	C4-dicarboxylic acid, orotate and citrate transporter
b4014	*aceB*	2.36	5.92E-05	malate synthase A
b3670	*ilvN*	2.34	1.51E-04	acetolactate synthase I, small subunit
b0343	*lacY*	2.33	2.61E-03	lactose/galactose transporter
b0523	*purE*	2.33	1.33E-03	N5-carboxyaminoimidazole ribonucleotide mutase
b1819	*manZ*	2.32	2.77E-04	mannose-specific enzyme IID component of PTS
b3787	*rffD*	2.32	5.37E-03	UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase
b3704	*rnpA*	2.29	1.91E-03	protein C5 component of RNase P
b1818	*manY*	2.28	7.64E-03	mannose-specific enzyme IIC component of PTS
b1054	*lpxL*	2.27	4.08E-04	lauryl-acyl carrier protein (ACP)-dependent acyltransferase
b3225	*nanA*	2.27	3.10E-03	N-acetylneuraminate lyase
b1626	*ydgK*	2.27	6.12E-03	conserved inner membrane protein
b0433	*ampG*	2.26	1.33E-02	muropeptide transporter
b4024	*lysC*	2.26	2.88E-04	aspartokinase III
b3960	*argH*	2.26	3.33E-03	argininosuccinate lyase
b0344	*lacZ*	2.25	3.53E-04	beta-D-galactosidase
b3166	*truB*	2.23	2.39E-04	tRNA pseudouridine synthase
b1987	*cbl*	2.22	7.06E-04	DNA-binding transcriptional activator of cysteine biosynthesis
b0182	*lpxB*	2.21	2.97E-03	tetraacyldisaccharide-1-P synthase
b1627	*rsxA*	2.21	3.70E-02	predicted inner membrane subunit
b3418	*malT*	2.19	4.90E-03	fused conserved protein -!- DNA-binding transcriptional activator, maltotriose-ATP-binding
b4488	*ilvG*	2.19	3.03E-03	acetolactate synthase II, large subunit (pseudogene)
b1424	*mdoD*	2.19	3.46E-03	glucan biosynthesis protein, periplasmic
b3172	*argG*	2.19	1.06E-04	argininosuccinate synthetase
b1435	*ydcP*	2.18	6.36E-03	predicted peptidase

Appendix 6: (continued)

Locus	Gene name	Fold Change	P-value	Description
b3784	*rfe*	2.18	1.65E-02	UDP-GlcNAc:undecaprenylphosphate GlcNAc-1-phosphate transferase
b0479	*fsr*	2.18	2.02E-03	predicted fosmidomycin efflux system
b0945	*pyrD*	2.18	1.35E-02	dihydro-orotate oxidase, FMN-linked
b2277	*nuoM*	2.17	7.02E-03	NADH:ubiquinone oxidoreductase, membrane subunit M
b3750	*rbsC*	2.17	1.28E-03	D-ribose transporter subunit -!- membrane component of ABC superfamily
b2344	*fadL*	2.15	1.02E-02	long-chain fatty acid outer membrane transporter
b0112	*aroP*	2.12	1.26E-03	aromatic amino acid transporter
b3705	*yidC*	2.12	2.77E-03	cytoplasmic insertase into membrane protein, Sec system
b3789	*rffH*	2.11	1.12E-02	glucose-1-phosphate thymidylyltransferase
b3116	*tdcC*	2.11	9.11E-03	L-threonine/L-serine transporter
b3455	*livG*	2.11	4.47E-04	leucine/isoleucine/valine transporter subunit -!- ATP-binding component of ABC superfamily
b3786	*rffE*	2.10	4.08E-02	UDP-N-acetyl glucosamine-2-epimerase
b0860	*artJ*	2.09	8.34E-04	arginine transporter subunit -!- periplasmic-binding component of ABC superfamily
b0845	*ybjJ*	2.09	6.34E-03	predicted transporter
b1929	*yedE*	2.09	9.58E-04	predicted inner membrane protein
b4312	*fimB*	2.08	1.85E-02	tyrosine recombinase/inversion of on/off regulator of fimA
b1602	*pntB*	2.07	1.89E-03	pyridine nucleotide transhydrogenase, beta subunit
b3214	*gltF*	2.07	5.61E-03	periplasmic protein
b0411	*tsx*	2.06	4.02E-03	nucleoside channel, receptor of phage T6 and colicin K
b2600	*tyrA*	2.06	1.20E-03	fused chorismate mutase T -!- prephenate dehydrogenase
b3749	*rbsA*	2.06	5.31E-03	fused D-ribose transporter subunits of ABC superfamily: ATP-binding components
b2923	*argO*	2.06	4.57E-03	arginine transporter
b3821	*pldA*	2.05	7.26E-03	outer membrane phospholipase A
b4113	*basR*	2.05	2.58E-03	DNA-binding response regulator in two-component regulatory system with BasS
b1851	*edd*	2.05	7.06E-04	6-phosphogluconate dehydratase
b1466	*narW*	-2.05	2.34E-03	nitrate reductase 2 (NRZ), delta subunit (assembly subunit)
b1562	*hokD*	-2.05	4.44E-03	Qin prophage; small toxic polypeptide
b2662	*gabT*	-2.06	1.09E-02	4-aminobutyrate aminotransferase, PLP-dependent
b2893	*dsbC*	-2.06	2.37E-03	protein disulfide isomerase II
b1183	*umuD*	-2.07	7.96E-03	DNA polymerase V, subunit D
b3824	*rhtB*	-2.07	2.91E-03	neutral amino-acid efflux system
b0628	*lipA*	-2.09	5.37E-03	lipoate synthase
b3513	*mdtE*	-2.09	6.90E-03	multidrug resistance efflux transporter
b1563	*relE*	-2.10	1.54E-04	Qin prophage; toxin of the RelE-RelB toxin-antitoxin system
b0209	*yafD*	-2.10	2.23E-02	conserved protein
b2080	*yegP*	-2.10	2.02E-03	predicted protein
b1468	*narZ*	-2.10	5.71E-03	nitrate reductase 2 (NRZ), alpha subunit
b3021	*ygiT*	-2.11	2.73E-02	predicted DNA-binding transcriptional regulator
b1681	*sufD*	-2.12	1.93E-04	component of SufBCD complex
b3003	*yghA*	-2.12	1.08E-02	predicted glutathionylspermidine synthase, with NAD(P)-binding Rossmann-fold domain
b1896	*otsA*	-2.13	1.37E-03	trehalose-6-phosphate synthase
b0231	*dinB*	-2.13	1.23E-02	DNA polymerase IV
b4056	*yjbQ*	-2.13	8.87E-03	conserved protein
b1806	*yeaY*	-2.14	1.44E-05	predicted lipoprotein
b3688	*yidQ*	-2.14	1.56E-02	conserved outer membrane protein
b1469	*narU*	-2.15	8.42E-03	nitrate/nitrite transporter

Appendix 6: (continued)

Locus	Gene name	Fold Change	P-value	Description
b3138	*agaB*	-2.15	1.35E-02	N-acetylgalactosamine-specific enzyme IIB component of PTS
b2448	*yffQ*	-2.15	3.22E-02	CPZ-55 prophage; predicted protein
b1178	*ycgK*	-2.15	1.59E-02	predicted protein
b3511	*hdeD*	-2.17	5.74E-05	acid-resistance membrane protein
b2504	*yfgG*	-2.18	1.53E-02	predicted protein
b2809	*ygdI*	-2.18	3.05E-02	predicted protein
b2847	*yqeI*	-2.19	2.25E-03	predicted transcriptional regulator
b0226	*dinJ*	-2.19	6.08E-04	predicted antitoxin of YafQ-DinJ toxin-antitoxin system
b4062	*soxS*	-2.19	1.73E-02	DNA-binding transcriptional dual regulator
b4126	*yjdI*	-2.20	1.12E-03	conserved protein
b0284	*yagR*	-2.20	8.71E-03	predicted oxidoreductase with molybdenum-binding domain
b0871	*poxB*	-2.20	4.22E-03	pyruvate dehydrogenase (pyruvate oxidase), thiamin-dependent, FAD-binding
b1040	*csgD*	-2.21	4.56E-04	DNA-binding transcriptional activator in two-component regulatory system
b1657	*ydhP*	-2.21	4.70E-05	predicted transporter
b1110	*ycfJ*	-2.22	3.19E-03	predicted protein
b4331	*kptA*	-2.22	1.21E-02	2'-phosphotransferase
b0759	*galE*	-2.23	4.38E-05	UDP-galactose-4-epimerase
b4468	*glcE*	-2.23	1.98E-02	glycolate oxidase FAD binding subunit
b0545	*ybcL*	-2.23	3.49E-04	DLP12 prophage; predicted kinase inhibitor
b1495	*yddB*	-2.24	2.35E-02	predicted porin protein
b1668	*ydhS*	-2.26	7.55E-04	conserved protein with FAD/NAD(P)-binding domain
b4269	*yjgB*	-2.27	2.05E-03	predicted alcohol dehydrogenase, Zn-dependent and NAD(P)-binding
b4060	*yjcB*	-2.28	3.20E-02	predicted inner membrane protein
b3494	*yhiO*	-2.28	1.92E-02	predicted universal stress (ethanol tolerance) protein B
b1683	*sufB*	-2.28	2.19E-03	component of SufBCD complex
b3519	*treF*	-2.28	5.13E-03	cytoplasmic trehalase
b2006	*yeeW*	-2.28	5.69E-04	CP4-44 prophage; predicted protein
b3470	*yhhP*	-2.29	4.20E-02	conserved protein required for cell growth
b4552	*yrhC*	-2.29	1.34E-02	predicted protein fragment (pseudogene)
b4400	*creD*	-2.29	5.76E-03	inner membrane protein
b3589	*yiaY*	-2.30	2.22E-02	predicted Fe-containing alcohol dehydrogenase
b0758	*galT*	-2.32	2.89E-04	galactose-1-phosphate uridylyltransferase
b0753	*ybgS*	-2.33	1.78E-03	conserved protein
b0757	*galK*	-2.33	2.74E-04	galactokinase
b0837	*yliI*	-2.35	1.36E-02	predicted dehydrogenase
b1004	*wrbA*	-2.36	6.90E-05	predicted flavoprotein in Trp regulation
b1725	*yniA*	-2.37	1.40E-04	predicted phosphotransferase/kinase
b3336	*bfr*	-2.37	1.79E-04	bacterioferritin, iron storage and detoxification protein
b1837	*yebW*	-2.40	8.60E-03	predicted protein
b3012	*dkgA*	-2.41	2.08E-04	2,5-diketo-D-gluconate reductase A
b0631	*ybeD*	-2.45	4.06E-04	conserved protein
b1003	*yccJ*	-2.46	1.17E-02	predicted protein
b1743	*spy*	-2.47	4.38E-02	envelope stress induced periplasmic protein
b0812	*dps*	-2.48	4.82E-05	Fe-binding and storage protein
b2005	*yeeV*	-2.49	3.22E-02	CP4-44 prophage; toxin of the YeeV-YeeU toxin-antitoxin system
b2665	*ygaU*	-2.49	1.77E-03	predicted protein
b1398	*paaK*	-2.50	1.11E-03	phenylacetyl-CoA ligase
b0802	*ybiJ*	-2.54	6.92E-03	predicted protein

Appendix 6: (continued)

Locus	Gene name	Fold Change	P-value	Description
b4485	*ytfR*	-2.55	1.63E-02	predicted sugar transporter subunit: ATP-binding component of ABC superfamily
b2465	*tktB*	-2.56	1.09E-03	transketolase 2, thiamin-binding
b2137	*yohF*	-2.57	3.19E-03	predicted oxidoreductase with NAD(P)-binding Rossmann-fold domain
b2390	*ypeC*	-2.57	1.10E-03	conserved protein
b1783	*yeaG*	-2.58	3.69E-05	conserved protein with nucleoside triphosphate hydrolase domain
b2541	*hcaB*	-2.59	1.15E-03	2,3-dihydroxy-2,3-dihydrophenylpropionate dehydrogenase
b1970	*yedX*	-2.62	1.59E-03	conserved protein
b1953	*yodD*	-2.62	3.61E-03	predicted protein
b1684	*sufA*	-2.64	1.17E-02	Fe-S cluster assembly protein
b3512	*gadE*	-2.65	8.96E-04	DNA-binding transcriptional activator
b3594	*yibA*	-2.67	1.49E-02	lyase containing HEAT-repeat
b2542	*hcaD*	-2.69	1.84E-04	phenylpropionate dioxygenase, ferredoxin reductase subunit
b4554	*yibT*	-2.70	1.39E-02	predicted protein
b1955	*yedP*	-2.70	1.45E-04	conserved protein
b2979	*glcD*	-2.70	6.64E-03	glycolate oxidase subunit, FAD-linked
b3073	*ygjG*	-2.73	7.01E-04	putrescine:2-oxoglutaric acid aminotransferase, PLP-dependent
b4217	*ytfK*	-2.73	6.02E-04	conserved protein
b1846	*yebE*	-2.77	2.02E-03	conserved protein
b3614	*yibQ*	-2.79	6.03E-03	predicted polysaccharide deacetylase
b1252	*tonB*	-2.79	4.68E-05	membrane spanning protein in TonB-ExbB-ExbD complex
b3350	*kefB*	-2.80	3.83E-04	potassium:proton antiporter
b2540	*hcaC*	-2.82	1.34E-04	3-phenylpropionate dioxygenase, predicted ferredoxin subunit
b2464	*talA*	-2.82	2.70E-03	transaldolase A
b2538	*hcaE*	-2.83	8.78E-06	3-phenylpropionate dioxygenase, large (alpha) subunit
b0980	*appA*	-2.85	1.00E-02	phosphoanhydride phosphorylase
b2539	*hcaF*	-2.86	4.48E-04	3-phenylpropionate dioxygenase, small (beta) subunit
b4551	*yheV*	-2.87	1.40E-04	predicted protein
b4484	*cpxP*	-2.88	6.13E-04	periplasmic protein combats stress
b0546	*ybcM*	-2.90	4.14E-03	DLP12 prophage; predicted DNA-binding transcriptional regulator
b0220	*ivy*	-2.91	7.54E-03	inhibitor of vertebrate C-lysozyme
b2127	*mlrA*	-2.91	7.77E-03	DNA-binding transcriptional regulator
b3945	*gldA*	-2.92	2.64E-03	glycerol dehydrogenase, NAD
b2777	*ygcF*	-2.96	7.79E-03	conserved protein
b4326	*yjiD*	-2.97	1.59E-03	DNA replication/recombination/repair protein
b0627	*tatE*	-2.97	2.25E-03	TatABCE protein translocation system subunit
b3029	*ygiN*	-3.01	7.28E-03	quinol monooxygenase
b2209	*eco*	-3.02	1.64E-03	ecotin, a serine protease inhibitor
b0897	*ycaC*	-3.03	8.83E-05	predicted hydrolase
b0978	*appC*	-3.20	1.87E-03	cytochrome bd-II oxidase, subunit I
b1257	*yciE*	-3.22	3.54E-03	conserved protein
b1597	*asr*	-3.24	3.17E-03	acid shock-inducible periplasmic protein
b1259	*yciG*	-3.25	2.50E-03	predicted protein
b3022	*ygiU*	-3.28	7.35E-03	predicted cyanide hydratase
b1973	*yodA*	-3.38	8.01E-03	conserved metal-binding protein
b4518	*ymdF*	-3.38	5.74E-04	conserved protein
b2660	*ygaF*	-3.41	4.26E-03	predicted enzyme
b0979	*appB*	-3.43	6.50E-04	cytochrome bd-II oxidase, subunit II
b1258	*yciF*	-3.46	9.22E-04	conserved protein
b1596	*ynfM*	-3.50	1.74E-03	predicted transporter

Appendix 6: (continued)

Locus	Gene name	Fold Change	P-value	Description
b1596	*ynfM*	-3.50	1.74E-03	predicted transporter
b1188	*ycgB*	-3.50	8.68E-04	conserved protein
b3351	*kefG*	-3.51	1.10E-03	component of potassium effux complex with KefB
b4313	*fimE*	-3.52	1.89E-02	tyrosine recombinase/inversion of on/off regulator of fimA
b2466	*ypfG*	-3.60	1.77E-03	predicted protein
b2135	*yohC*	-3.62	6.73E-04	predicted inner membrane protein
b3028	*mdaB*	-3.68	1.85E-03	NADPH quinone reductase
b0806	*ybiM*	-3.73	9.47E-03	predicted protein
b4107	*phnB*	-3.97	2.22E-03	conserved protein
b0207	*dkgB*	-4.05	1.52E-03	2,5-diketo-D-gluconate reductase B
b1788	*yoaI*	-4.06	2.18E-02	predicted protein
b2659	*ygaT*	-4.20	3.20E-04	predicted protein
b3555	*yiaG*	-4.23	1.05E-03	predicted transcriptional regulator
b2378	*ddg*	-4.40	2.34E-02	palmitoleoyl-acyl carrier protein (ACP)-dependent acyltransferase
b3207	*yrbL*	-4.80	6.99E-03	predicted protein
b0953	*rmf*	-5.18	3.30E-03	ribosome modulation factor
b1836	*yebV*	-5.37	9.03E-04	predicted protein
b0836	*yliH*	-5.64	2.20E-03	conserved protein
b4030	*yjbA*	-6.16	1.23E-03	predicted phosphate starvation inducible protein

Appendix 7: Mutations identified from sequencing TY05 before and during a chemostat culture fermentation

Locus	Gene name	Coding[1] mutations in JTY01[4]	Coding[1] mutations in TY05	Non-coding[2] mutations JTY01[4]	Non-Coding[2] mutations in TY05	Known mutation[3] (Y/N)
b0214	*rnhA*	0	0	1	1	N
b0215	*dnaQ*	0	0	1	1	N
b0239	*frsA*	0	0	2	0	N
b0240	*crl*	2	0	0	0	Y
b0242	*proB*	0	0	2	0	N
b0519	*ylbE_2*	0	0	2	1	Y
b0522	*purK*	0	0	2	1	N
b1493	*gadB*	0	0	1	1	N
b1495	*yddB*	1	1	0	0	Y
b1496	*yddA*	0	0	1	1	N
b1804	*rnd*	0	0	1	0	N
b1806	*yeaY*	0	0	1	0	N
b2087	*gatR_1*	0	0	1	1	N
b2092	*gatC*	1	1	0	0	Y
b2093	*gatB*	0	0	1	1	N
b2979	*glcD*	0	0	1	1	N
b2983	*yghQ*	1	1	0	0	N
b2984	*yghR*	0	0	1	1	N
b3423	*glpR*	1	1	0	0	Y
b3424	*glpG*	0	0	1	1	N
b3776	*yifO*	0	0	1	1	Y
b3778	*rep*	0	0	1	1	N
b3838	*tatB*	0	0	2	1	N
b3842	*rfaH*	0	0	2	1	N
b3843	*ubiD*	0	0	1	1	N
b4073	*nrfD*	0	0	2	1	N
b4475	*rtcA*	0	0	1	1	N
b4507	*ylbE_1*	2	1	0	0	Y

1. Coding mutations refer to mutations identified in the gene itself
2. Non coding mutations indicate mutations identified within 5000 bases up or downstream of the coding region
3. Known mutations indicate common mutations identified in MG1655 stock strains (NCBI)
4. JTY01 refers to the sequenced strain obtained from a freezer stock of culture from 170 hours into a chemostat culture fermentation.

References for Appendices

Agnew DE, Stevermer AK, Youngquist JT, Pfleger BF. 2012. Engineering *Escherichia coli* for production of C_{12}-C_{14} polyhydroxyalkanoate from glucose. *Metabolic Engineering* **14**:705–13.

Amann E, Ochs B, Abel K. 1988. Tightly regulated tuc promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**:301–315.

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

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

Guzman LM, Belin D, Carson MJ, Beckwith J, Guzman L, Belin D, Carson MJ. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter . *Journal of Bacteriology* **177**:4121–4130.

Hoover SW, Marner WD, Brownson AK, Lennen RM, Wittkopp TM, Yoshitani J, Zulkifly S, Graham LE, Chaston SD, McMahon KD, Pfleger BF. 2011. Bacterial production of free fatty acids from freshwater macroalgal cellulose. *Applied Microbiology and Biotechnology* **91**:435–46.

Lennen RM, Braden DJ, West RA, Dumesic JA, Pfleger BF. 2010. A process for microbial hydrocarbon synthesis: Overproduction of fatty acids in *Escherichia coli* and catalytic conversion to alkanes. *Biotechnology and Bioengineering* **106**:193–202.

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

Yu D, Ellis HM, Lee E-C, Jenkins N a, Copeland NG, Court DL. 2000. An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **97**:5978–5983.

Zhang F, Ouellet M, Batth TS, Adams PD, Petzold CJ, Mukhopadhyay A, Keasling JD. 2012. Enhancing fatty acid production by the expression of the regulatory transcription factor FadR. *Metabolic Engineering* **14**:653–60.