8-2013

Application of Synthetic Biology for Increasing Anaerobic Biodiesel Production in Escherichia coli

Michael Christopher Wierzbicki
mwierzbi@utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Michael Christopher Wierzbicki entitled "Application of Synthetic Biology for Increasing Anaerobic Biodiesel Production in Escherichia coli." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Chemical Engineering.

Cong T. Trinh, Major Professor

We have read this thesis and recommend its acceptance:

Eric Boder, Barry Bruce

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Application of Synthetic Biology for Increasing Anaerobic Biodiesel Production in *Escherichia coli*

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Michael Christopher Wierzbicki

August 2013

Acknowledgements

I would like to thank Dr. Cong Trinh for employing me in his laboratory, giving me much-needed experience in research and for serving on my committee. Additionally I would like to specially thank Dr. Narayan Niraula, Donovan Layton, and Akshitha Yarrabothula for assisting in all aspects of this project. I would also like to thank Dr. Eric Boder and Dr. Barry Bruce for serving as the committee members of my thesis defense. Acknowledgement also goes to the Department of Chemical and Biomolecular Engineering for giving me the opportunity to pursue a master's degree.

Abstract

The ever-increasing demand for transportation biofuels requires new and novel approaches to solve the complexities associated with efficient biofuel production. Ethanol, the most common biofuel, has physical limitation associated with difficulty of separations and issues with water contamination and as such is not a long-term transportation fuel solution. (Lou & Singh, 2010; Wheals, Basso, Alves, & Amorim, 1999) Biodiesel is seen as a possible alternative to ethanol due to its hydrophobicity and also has comparable energy density and cetane number to its petroleum derived counterpart. (Kalscheuer, Stölting, & Steinbüchel, 2006) Because of feedstock limitations, biodiesel produced from vegetable oils is limited by the supply of vegetable oil crops which creates scaling issues and land usage concerns. (Kalscheuer et al., 2006) An alternate method for biodiesel (fatty acid ethyl esters, FAEEs) generation was proposed which would bypass the need for vegetable oils by utilizing the fatty acids and ethanol made in engineered *Escherichia coli.* (Kalscheuer et al., 2006) FAEE is not water soluble, so water contamination in fuel supplies seen with ethanol is not likely to cause damage to fuel infrastructure and has similar combustion properties to petroleum based diesel. (Lou & Singh, 2010)

The goal of this project is to apply metabolic engineering and synthetic biology principles to engineer *E. coli* for efficient anaerobic production of FAEE from fermentable sugars. The strain utilized for this project builds upon previous work in which elementary mode analysis was used to design an *E. coli* strain that minimizes fermentative by-products under anaerobic conditions. The engineered strain provided an optimized platform to employ plug-and-play modular principles for fully endogenous FAEE biosynthesis.

In order to produce the desired esters, two parallel pathways were introduced, the first produced fatty-acyl-CoA, and the second the alcohol of interest. These two molecules are then catalyzed by a wax-ester synthase to produce desired biodiesel. Pathway flux engineering principles were employed to balance the metabolic fluxes of the two pathways that complete for the common substrate, pyruvate. The results show the dynamic range of module fluxes that can be achieved by varying promoter strength, operon orientation, and plasmid copy number.

Contents

List of Tables

List of Figures

List of Equations

Chapter 1 Introduction

1.1Why We Are Interested In Developing Advanced Biofuel Technologies

Increasing energy costs and environmental concerns have emphasized the need to produce sustainable and renewable fuels and chemicals. To this end conversion of biomass to biofuels has been the subject of intense research since the 1970s. Recent concerns raised about climate change, global energy security, and questions about petroleum supply have brought methods for renewable and carbon neutral biological fuel production into prominence. (Fortman et al., 2008)

Research in transportation biofuels has generally taken two approaches. The first being incremental changes in improved efficiency that result in the narrowing the gap between real world efficiency and the theoretical max efficiency. An example of this incremental improvement approach being the achievement of near theoretical limit production of ethanol in engineered *E. coli*. (Trinh, Unrean, & Srienc, 2008a) But, as seen in the *E. coli* scenario, the theoretical limit is simply not sufficient to answer great biofuel search. The second path of transportation fuel advancement is the complete change of the approach used to answer the biofuel question. Alternatives to ethanol have been proposed, such as isobutanol (Atsumi, Wu, Eckl, et al., 2010; Smith & Liao, 2011), butanol (C R Shen & Liao, 2008), transesterification of vegetable oils and animal fats (Ma & Hanna, 1999), and fatty acid derived ethyl esters (FAEE) (Liu & Khosla, 2010; Steen et al., 2010; Zhang, Carothers, & Keasling, 2012). While isobutanol and butanol offer several advantages due to the physical properties, such as decreased water solubility, there are some serious and potentially unsurpassable obstacles; namely that the solubility of isobutanol and butanol in water is greater than *E. coli* ‘s toxicity threshold, which

results in cell growth arrest and cell death.(Atsumi, Wu, Machado, et al., 2010) Because of these limitations, biodiesel appears to posse favorable traits for further research.

1.2 How Biodiesel is Currently Made

Thus far, biodiesel production has come from two different methods. The first is the esterification of vegetable oils or animal fats. Fats and oils are generally not water soluble and consist of one mole of glycerol and three moles of fatty acids. (Shay, 1993) These compounds are generally referred to as triglycerides. Though this method has been implemented with success, with the US producing approximately 875 million gallons of biodiesel in 2012 according to the US Energy Information Administration, this represents only 1.5% of total diesel consumption in the US during the same time period. There are several obstacles in manufacture that have limited large-scale implementations. First, the composition of the fatty acids of interest varies widely depending on the feedstock used and as a result diesel quality can be the subject undesirable variations. (Ma & Hanna, 1999). The esterification reaction itself also requires the use of a catalyst such as a strong acid and at least a three to one molar ratio of alcohol to triglyceride to drive the reversible esterification reaction. (Ma & Hanna, 1999) Furthermore, the alcohol most commonly used is methanol, which is generally derived from methane, a non-renewable carbon source. (Periana et al., 1993) As a result, excess production costs are accumulated from wasted non-renewable alcohol. Beyond processing complications, the source of the triglycerides is subject to limitations due to the specific crops needed to generate the oils. (Kalscheuer et al., 2006)

1.3 Benefits and Drawbacks of FAEE

One promising biofuel is the biodiesel that combines ethanol and fatty-acyl-CoA to form FAEE entirely endogenously. FAEE is not water-soluble and as a result, high concentrations of the fuel can be produced with minimal to no inhibitory effect on the cells. (Zhang et al., 2012) To date, the Keasling group has laid the fundamental groundwork on the production of FAEE. Steen *et al* managed to reach approximately 7.5% of theoretical FAEE yield in 2010, (Steen et al., 2010) and Zhang *et al* further improved upon those results with the implementation of dynamic control of the ethanol pathway to delay ethanol production which resulted in reaching 28% of theoretical yield. (Zhang et al., 2012) The previous works demonstrated that there are several obstacles that need to be overcome before endogenously produced biodiesel will be able to reach its potential as an alternative to petroleum derived fuels. As demonstrated by Steen *et al.* and Zhang *et al.* the fatty acyl-CoA module of the FAEE pathway has a significantly slower reaction rate than that of the ethanol pathway. Since glucose is the sole feedstock for both pathways, the presence of excess ethanol directly decreases the maximal amount of biodiesel that can be produced. Therefore, the ideal case for optimal FAEE production is hypothesized to be production of fatty acyl-CoA in equimolar quantities as ethanol so that all substrates are fully utilized.

1.4 Biological Conversion Route for Producing Fatty Acid Ethyl Ester

Production of FAEE utilizes the endogenous pathways that uptake glucose as a substrate to generate fatty-acyl CoA via the fatty acid synthesis pathway, also referred to as the fatty acid elongation cycle. The fatty acid synthase pathway is a process in which a pair of carbons are added to a fatty acid until a thioesterase pulls the carbon chain out of the elongation cycle at a specific length determined by a thioesterase. Previous studies have demonstrated the ability to

modify the length of the produced fatty acids by swapping different thioesterases in place of the endogenous one, and predictably alter the ester carbon chain length. Efforts thus far, though, have only been able to generate shorter than endogenous length fatty acids, which are less desirable for biodiesel applications. (Steen et al., 2010) In parallel, ethanol is generated via an overexpressed pathway that converts pyruvate to ethanol. The fatty-acyl CoA molecule is then esterified with an ethanol molecule by an overexpressed wax ester synthase gene to form the FAEE biodiesel.

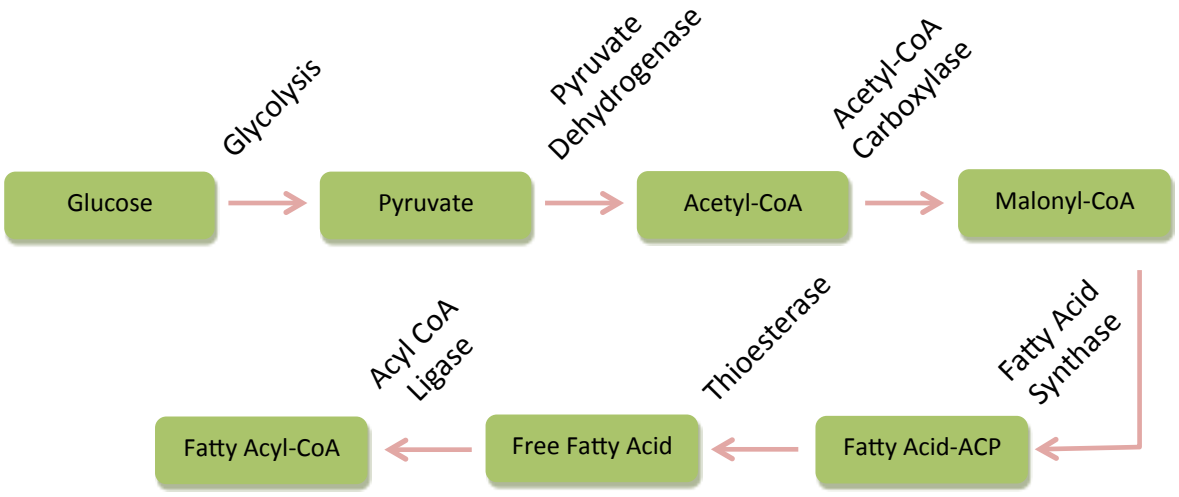

Figure 1 A simplified metabolic pathway to convert glucose into fatty acyl-CoA.

The metabolic steps that were overexpressed were the thioesterase (*tesA*) and acyl-CoA ligase (*fadD*).

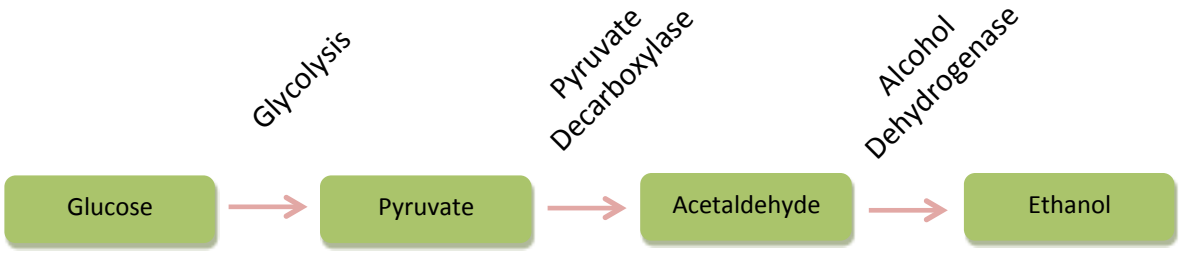

Figure 2 A simplified metabolic pathway to convert glucose into ethanol.

The metabolic steps that were overexpressed were pyruvate decarboxylase (*pdc*) and alcohol dehydrogenase (*adhB*). These genes were derived from *Zymomonas mobilis*.

Figure 1 and Figure 2 show a simplified representation of the two pathways of interest in this project that are required for the production of FAEE. Fatty Acyl-CoA and ethanol are joined together with a wax-ester synthase (AtfA) to form FAEE (see figure 3 for more details). The two pathways compete for the common precursor, pyruvate, and as a result optimal flux distribution is key for ensuring high yields of FAEE.

1.5 Importance of Anaerobic Production for Biofuels

To date, production of fatty acid ethyl in esters *E. coli* has been limited to aerobic production. (Steen et al., 2010) Aerobic production suffers from two hindrances that make aerobic bio-production undesirable for industrial scale manufacture, as described in a review of anaerobic industrial fermentative by Stark and Pohler. (Stark & Pohler, 1950) First, aerobic processes need a sterile source of oxygen. Sterilization of oxygen, at rates required for industrial scale bio-production poses a significant challenge since any contaminant can result in the total loss of a production run. Advances in air purification to meet these demands have moved beyond

the steel wool and carbon described by Stark, with the implementation of advanced polymer-based filtering technologies, but challenges still remain. Secondly, aerobic bio-production is susceptible to challenges associated with equal and consistent dissolved oxygen concentrations. Inconsistencies in dissolved oxygen distribution can result in differing cellular states, and as a result in unpredictable product formation. (Trinh, Li, Blanch, & Clark, 2011)

With these limitations in mind, the appeal of anaerobic industrial scale production of desirable chemicals is clear. Further justification for anaerobic production is based on the metabolic pathways of the cell and the methods available for engineering cells for increased efficiency.

1.6 Metabolic Engineering and the Role of Metabolic Network Modeling

Though there are metabolic modeling techniques in literature, discussion in this report will be limited to elementary mode analysis. The strain used for the characterizations in this study was created based on knockouts predicted by elementary mode analysis to optimize two characteristics. First, by-product forming pathways were eliminated so that the carbon substrate is utilized for only biomass formation (life sustaining pathways) and ethanol production. Secondly, knockouts were performed that would ensures a strain will perform consistently for a given set of conditions and is theoretically incapable of a yield lower than maximum, a critical characteristic for industrial applications. As mentioned before, TCS083, the base strain utilized in this study, was developed for efficient production of ethanol. Since fatty acids were considered as part of the biomass formation, no negative effects in fatty acid production are expected and so this strain was considered a promising strain for continued development. TCS083 was further specialized with the elimination of *fadE*, the second gene in the fatty acid

degradation pathway. Knocking-out *fadE* prevents acyl CoA from being catabolized and has been reported to increase yields by three to four fold. (Steen et al., 2010)

1.7 Project Scope

The primary goal of the project was to apply engineering principles to a model system that had real world benefits. This meant that a logical and well-defined system needed be implemented that had a toolbox of techniques would allow this project to achieve the increased FAEE production efficiency goal. Synthetic biology is seen as a powerful tool to allow engineering of microbes in a logical and predictable fashion. Many attempts at engineering microbes to date have been a result of mixing and matching genetic sequences and seeing what happens. This approach has two main drawbacks. First, the results are often unpredictable, and secondly there is no clear path towards reaching a certain goal due to the lack of the ability to fine tune pathways. Synthetic biology applies many of the same principles employed by the Industrial Revolution to mass-produce goods, but to a microbial frame of mind. Where standardized and interchangeable car parts made the cars common, so to do "parts," "modules," and standardization principles in synthetic biology hope to make engineered microbes part of the modern industrial landscape.

1.8 Project Contributions

This project was performed as a collaborative effort of a research team consisting of Dr. Cong Trinh, Dr. Narayan Niraula, Akshitha Yarrabothula, and myself. My contribution to this project was to focus on strain characterizations for optimal ester production.

Chapter 2 Methods

2.1 Strains and Plasmids

As previously mentioned, the primary strain used in this study is a modified version of TCS083, a strain generated for optimal anaerobic production of ethanol. (Trinh et al., 2008a) TCS083 was further modified by knocking out *fadE*, the first step in the fatty acid degradation pathway. TCS083 was derived from the wild type *E. coli* strain MG1655, and so MG1655 was used as a control to determine the effectiveness of the final strain towards enhancing anaerobic FAEE production. Below is a table of the knockouts that differentiate the wild type from the strain used in this

Table 1 Genotype of TCS083 Δ*fadE*

Gene Eliminated	Enzymatic Function
zwf	Glucose-6-phosphate-1-dehydrogenase
ndh	NADH dehydrogenase II
sfcA/maeB	NAD/NADP-dependent malate enzyme
ldhA	D-lactate dehydrogenase
frdA	Fumarate reductase
poxB	Pyruvate oxidase
pta	Phosphate acetyl-transferase
fadE	Acyl-CoA dehydrogenase

TCS083 Δ*fadE* was further modified by using Novagen's λDE3 lysogenization kit, catalogue number 69734-3 to express the bacteriophage T7 polymerase. This permitted the use of a T7 promoter. The T7 expression system is, best known for its ability to express high

quantities of proteins in the BL21 (λDE3) *E. coli* strain. The T7 promoter has the distinct advantage of being transcriptionally orthogonal to the native system, and is thus theoretically isolated from the physiological state of the bacteria's transcriptional machinery (Conway, Sewell, Osman, & Ingram, 1987). The expected result should be a pathway that is less subject to variations. (Rao, 2012)

Several plasmids were generated for this study using the BGL brick assembly method. The method also provides a convenient and systematic way to design operons in a way that allows two-gene insertion into a backbone repeatedly. Below is a table that lists the plasmids used in this study as well as the functions they serve. Additional information about the plasmids composition can be found in the supplementary information.

Table 2 List of plasmids used in this study..

Plasmid Name	Product	Genotype
pCT13	Isobutanol	pCOLA p_{T7}::*alsS*::*ilvC*::*ilvD*::p_{T7}::*kivd*::*adhE*; Kan^R
pCT24	Ethanol	pETite C-His p_{T7}::*pdc*::*adhB*; Kan^R
pAY1-pAY4	Ethanol	pETite C-His p_{T7}::*pdc*::*adhB*; Kan^R
pNN14-pNN18	Ethanol	pETite C-His plac::*pdc*::*adhB*; Kan^R
pButanol	Butanol	pETite C-His plac::*adhE*; Kan^R
pCT63	Fatty Acyl-CoA + AtfA	pETite C-His p_{T7}::*tesA*::*fadD*::*atfA*; Kan^R
pCT69 LC	pCT63 -leaderless version of *tesA*	pETite C-His p_{T7}::'*tesA*::*fadD*::*atfA*; Amp^R
pCT69 MLC	pCT69 but higher copy number	pETite C-His p_{T7}::'*tesA*::*fadD*::*atfA*; Amp^R
pCT74-pCT77	FAEE in a single plasmid	pETite C-His p_{lac}::*pdc*::*adhB*:: *plac*::'*tesA*::*fadD*::*atfA*; Kan^R

2.2 Strain Characterization

Culture Media

A modified M9 salt medium was used containing M9 salts (see table 3), 1mM magnesium sulfate, 10mM calcium chloride, 1mL/L trace metals solution as described in Trinh *et al*., 2011, 20g/L glucose, and 5g/L of yeast extract and appropriate antibiotics. Unless otherwise specified, this medium was used in all characterizations.

Table 3 Composition of 5x M9 salt solution.

Mass	Chemical
64g	Na_2HPO_4-$7H_2O$
15g	KH_2PO_4
2.5g	NaCl
5g	NH_4Cl

Cultivation

Seed cultures were grown overnight in aerobic test tubes in the modified M9 medium with appropriate antibiotics. Seed cultures were then sub-cultured into a 500mL baffled aerobic flask with 100mL of the modified M9 medium to an OD (measured at 600nm) of 0.1. Once the OD reached 0.8, IPTG was added to induce transcription of the T7 promoter system. Once the OD reached 1.0, the cells were transferred into two 50mL centrifuge tubes and centrifuged at 11,000RPM at 37^{o}C for 7 minutes. The supernatant was then discarded and the cell pellet was resuspended to an OD of 10 using fresh modified M9 media with appropriate antibiotics. 10mL of the resuspended culture was transferred into a 20mL test tube with additional IPTG. A rubber

stopper was applied onto the test tube and a needle with 3mL syringe packed with cotton was inserted into the on top of the rubber stopper to allow for venting of gases generated during fermentation.

High Throughput Testing Platform

A testing method needed be established to rapidly evaluate the libraries of constructs created. Two criteria needed to be satisfied for a testing method to be acceptable. First, the method would need to be able to create a pseudo-steady state environment that could mimic an in-vivo enzymatic assay. This means that biomass should not change or change only negligibly throughout the experiment, and as a result the starting cell count should be sufficiently large enough to perform the enzymatic steps. This was achieved with what is referred to as the "High OD" method, in which a starting OD of 10 is used (preparation described in the previous section). Secondly, in order to be industrially relevant, a bioprocess must be performed anaerobically, as aerobic bioprocesses are both capital and labor intensive and as a result is cost prohibitive for bulk chemical production. Additionally the production of these FAEE has yet to be demonstrated under anaerobic conditions, making FAEE an interesting candidate for this study.

2.3 Analytical Method

Samples extracted for analysis were stored at -20^{o}C in sterile 15mL centrifuge tubes until the samples were needed for analysis. Cell density was analyzed by measuring optical density of cell suspension at 600nm using a spectrophotometer (Thermo Scientific Spectronic 200+).

Samples for HPLC were thawed at room temperature and then centrifuged at 15,000RPM for 10 min. The supernatant was then carefully removed with a needle and syringe to not disturb

the cell pellet. The supernatant was then passed through a 0.2 μm PTFE filter that attached via luer lock to the end of the syringe into a 96 HPLC well plate to remove any remaining contaminates.

The glucose and ethanol concentrations were determined by high performance liquid chromatography system equipped with UV-VIS (Shimadzu UV:SPD-20AV), RID detector (Shimadzu RID-10A) and a BioRad Aminex HPX 87-H column. Sulfuric acid (10 nM) was used as a solvent that was run through the column isocratically at 0.6mL/min and 50^{o}C. Fatty acid esters were extracted from cultures by lysing with 6M HCl and extracting with ethyl acetate containing 5mg/L of pentadecanoic acid ethyl ester as an internal standard. A 1-μL portion of the extracted ethyl acetate fraction was injected into the GC-MS using splitless injection (1 μL at 280°C); helium was used as a carrier gas (1 μL/min). The mass transfer line and ion source were at 250°C and 200°C, respectively. All the esters were detected with electron ionization in scan mode (50 to 650 m/z) and selected ion monitoring mode at m/z 88 and parent ion of each compound (for quantitative analysis). The column temperature was initially held at 160°C for 1 min, then increase by 20°C for 7 min, and then held at 300°C for 2 min.

2.4 Result Reporting

Results were compared in terms of milligrams of ester per gram glucose as determined by GCMS and HPLC sample analysis. With the exception to one construct set, all or nearly all glucose was consumed in the 24 hour time point. The total amount of esters formed was then divided by the amount of glucose consumed to determine the yield. Additionally, the ester length distribution was analyzed to determine what, if any effect, the constructs had on the esters produced. Ester composition is significant since fuel quality is directly linked to the ester chain length.

Chapter 3 Engineering the FAEE-Producing Pathway for Enhanced FAEE Production in *E. coli*

3.1 Experimental Test Pathway

The model system chosen to evaluate the tools made available by synthetic biology and metabolic engineering had to meet three criteria. First the system must be sensitive to genetic change, meaning that applied changes would result in a phenotypic change. Secondly, the system had to be divisible into clear modules, and as a result, the convergence of parallel processes would be ideal. Thirdly, the results needed to be easily analyzed and have low noise. Following these guidelines, the fatty acid ethyl ester pathway was chosen as an excellent candidate.

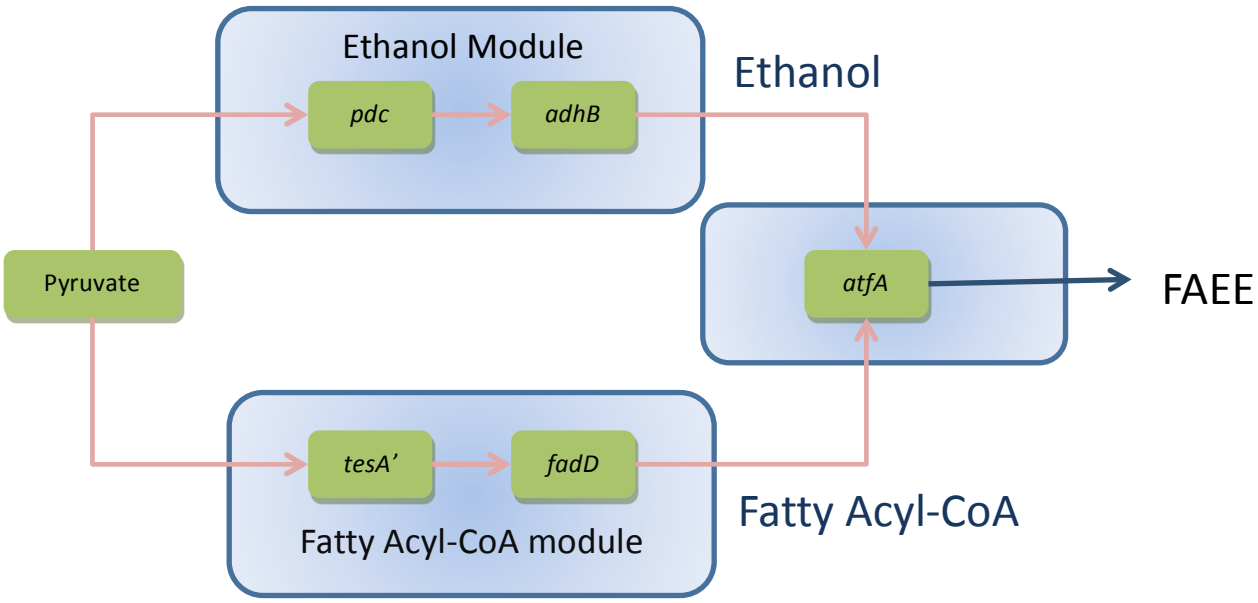

Figure 3 A simplified pathway for the production of fatty acid ethyl esters.

The fatty acid ethyl ester construct consists of two parallel pathways that draw from pyruvate and join together to form the product ester.

The FAEE construct involves two parallel pathways that utilize the common precursor, pyruvate, and join together at the end to form the desired ester. Modification of the enzymatic quantities of each pathway should result in an altered phenotype. The methods used in this report for alteration of enzymatic quantities and enzymatic ratios are: variation of promoter strengths used, augmenting the copy number of suspected rate limiting pathways, using different operon orientations, and the implementation of a multi-strain system. The altered phenotypes were easily detectable by the change in the yield of ester formation as a result of change in the ethanol to fatty-acyl CoA pathway ratios.

3.2 Base Case Study

A base construct was created to establish a baseline for comparison by Dr. Cong Trinh. This construct consisted of an ethanol producing plasmid (pCT24) that utilizes genes from *Z. mobilis*, which have been demonstrated to be very efficient for ethanol production (Trinh, Unrean, & Srienc, 2008b). This construct utilized a bacteriophage T7 promoter. The endogenous ribosomal binding site (Shine-Delgarno sequence) was preserved in all constructs. The second plasmid consisted of a leaderless version of *tesA* (‘*tesA*), *fadD*, and *atfA* (pCT69) and was created by Dr. Narayan Niraula. *‘TesA* acts to pull the fatty-acyl ACP from the elongation cycle and to catalyze the formation of free fatty acids.(Ikeda & Richardson, 1986) *FadD* then acts to transform the free fatty acid into fatty-acyl-CoA. Finally AtfA combines the ethanol precursor with fatty-acyl-CoA and forms the target FAEE. Both constructs were built into a low copy number plasmid backbone with a copy number of approximately 20 copies per cell. (Tsui, 2006) It is worth noting at this point, both *‘tesA* and *atfA* have a fairly large degree of promiscuity, and as a result can form a range of products as opposed to one single product.

As reported in literature, *'tesA* has the highest specificity towards fatty-acyl-ACP of length C14, with C12, C16, and C18 being possible alternate products in order of typical abundance greatest to least. (Steen et al., 2010) Results from our studies described in greater detail in section 3.9 have indicated that the prominent carbon length tends to be C14 and C12, with C16 and C18 having relatively low abundance in anaerobic conditions, which is not unexpected given the ATP limitations placed on the otherwise ATP intensive fatty acid elongation process. (Lu, Vora, & Khosla, 2008; Steen et al., 2010)

3.3 Effect of Promoter Swapping on FAEE Production

With the baseline pCT24 + pCT69 dual plasmid construct established, the first sets of variables were introduced. As discussed previously, decreasing the amount of excess ethanol was hypothesized to indirectly result in an increase in the amount of fatty acids produced, and as a result increase ester yield. In order to achieve this goal, methods common to synthetic biology were employed. One of the manifestations of the synthetic biology revolution is the development of the "Parts Registry." The parts registry is a depository of open-source, user tested and submitted genetic parts and devices. From this depository, the Anderson library of promoters was utilized with the aim to decrease the transcriptional rate of the ethanol pathway. The Anderson library was created by inserting point mutations to a consensus sequence of the lac promoter with the lac repressor site removed. The newly created mutants were then tested using fluorescent proteins and flow cytometry for analysis. The strength of the promoters was reported relative to the original consensus sequence (relative strength).(Atsumi, Wu, Machado, et al., 2010) Since there is no reported characterization of the T7 promoter that can be used to directly compare the T7 promoter to the constitutive lac promoters of the Anderson library, a range of promoters strengths were selected from the consensus promoter (J23100), 0.7 relative strength

(J23101) and, 0.51 relative strength (J23108) and an accidental, previously uncharacterized mutant was accidently created and used as well (Unknown). These new ethanol constructs were designed by Akshitha Yarrabothula and constructed by Dr. Narayan Niraula,

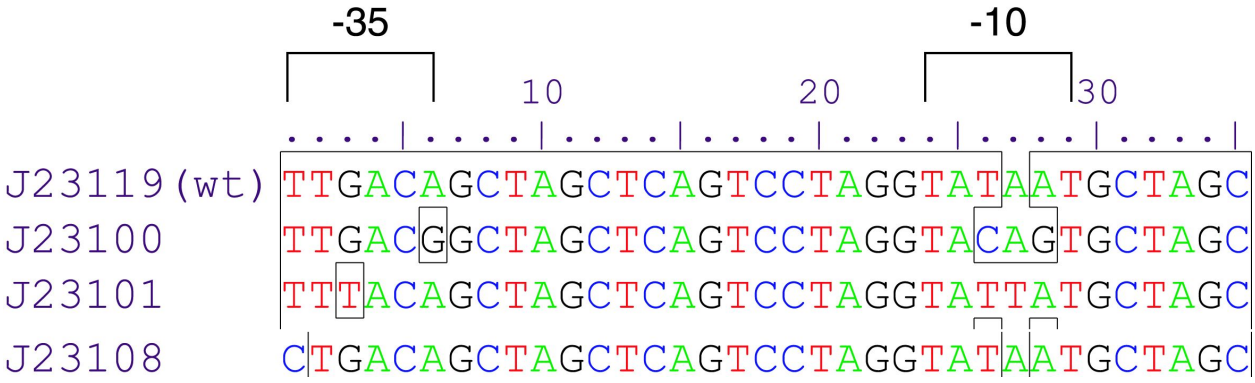

Figure 4 DNA sequence alignment of variant lac promoters.

The above chart depicts the DNA sequence alignment of the three promoters utilized from the Anderson library along with the wild type lac promoter with the lac repression site removed. The library was generated by inserting point mutations into the otherwise conserved -10 and -35 regions of the promoter.

Though the Anderson Library promoters were characterized using fluorescent proteins and flow cytometry, the differences in strain and application from one system to another as well as translational effects could create variations that effect results. In order to attempt to reduce variability in the experiments, a low copy number plasmid (~20) was utilized to help control consistent fatty acyl-CoA level between various strains. This strategy has the implicit drawback in that the total FAEE production will be decreased as a result of limited AtfA, a known rate-limiting enzyme. But the control of expression was necessary to evaluate the promoter effects.

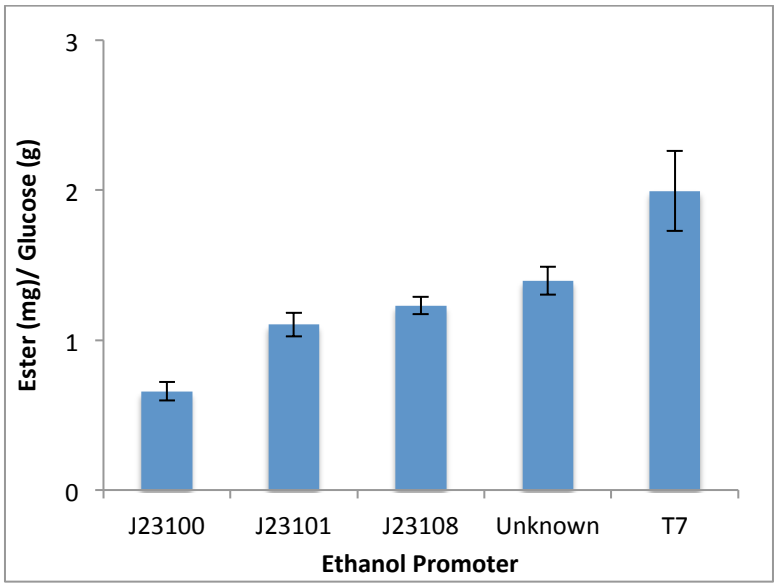

Figure 5 **The baseline ester yield experiment.**

These results were generated using the high OD method and the experiment was run in triplicate. TCS083 Δ*fadE* (DE3) with the baseline pCT69 plasmid was co-transformed with plasmids housing the ethanol construct under transcription control of the T7 prompter and promoters from the Anderson Library.

Promoter Variation Results

The replacement of the T7 promoter from the baseline construct with the select promoters from the Anderson library resulted in changes to the ester yield. These yields were predicted by the hypothesis that altering the carbon flux distribution between the parallel ethanol and fatty acyl-CoA. Relative to the baseline, the strongest promoter J23100 resulted in a three-fold decrease in ester yield. J23101 resulted in a 1.81-fold decrease, and J23108 resulted in a 1.62-fold decrease. The unknown mutant promoter resulted in a 1.43-fold decrease in ester yield. Error bars represent one standard deviation from the average and the experiment was performed

in triplicates using the high OD method described above. Glucose was totally consumed by the 24 hour sample point used for yield analysis. Ethanol yield was on average 51% for all promoters in this experiment. Cells were induced as described in section 2.2 Strain Characterization.

Promoter Variation Discussion

The results indicate that varying promoter strength, and as a result varying the transcriptional rate is a plausible method for altering ester yield for this system. As can be seen in the above figure, there is a three-fold difference between the strongest constitutive promoter and the T7 promoter, with the remaining promoters occupying between a 1.81 and 1.43-fold decrease. Error bars of one standard deviation from triplicate experiments indicate no overlapping regions between the tested promoters, indicating that the observed changes are real. The standard deviation covers the rage of +/- 5-10% of total esters formed can be attributed to natural variations associated with the copy number of the plasmid. The copy number dictates the physical copies of the plasmid, but fluctuations are typical from cell to cell and cellular physiological state. As a result, it was feared that the wide range of variation could overshadow changes generated by promoter swapping, though this appears to not be the case. Even with the observed noise, these results indicate that the high OD testing method and FAEE platform is a suitable candidate for testing synthetic biology principles.

Promoter Strength in an in vivo Context

The hypothesis for this experiment states that the stronger the ethanol promoter the less ester yield is expected. This hypothesis hold true for the promoters borrowed from the Anderson library, and so based on this finding; there is evidence to support the claim that the T7 promoter

is the “weakest” of the group. This result seems not consistent with information from companies that sell expression systems that claim very high expression rates, though there are no papers comparing the T7 promoter to the constitutive promoters of the Anderson library. The results are plausible if the definition of promoter strength is defined in an in-vivo context. If the promoter strength is analyzed in terms of affinity between the polymerase and the promoter sequence, then the claim that the T7 promoter and polymerase combination are the strongest may be valid. But this is based on the assumption that the only factor that determines transcriptional rate is the coupling affinity.

$$R_t = f(F_{affinity}),$$

Equation 1 A general equation that describes traditional understandings of what contributes to transcriptional rate.

Where Rt is the transcriptional rate, and $F_{affinity}$ is the coupling affinity.

Since this project compares both endogenous and orthogonal promoter system combinations, an alternate definition must be employed to account for the discrepancy in the number of free polymerases available for coupling. Since the T7 polymerase genes is transcriptionally controlled by a lac promoter with the repressor site retained, it is probable that the number of free polymerase itself is rate limiting. Whereas the endogenous polymerase is likely more abundant and so the probability of a polymerase docking and completing transcription are greater simply because the event happens much more frequently and counter acts any relative coupling deficiencies. The resulting expression is modified to account for this proposed phenomenon.

$$R_t = f(F_{affinity}, N_{polymerase}),$$

Equation 2 A proposed equation that reflects the effects observed in different expression systems.

Where Rt is the transcriptional rate, and $F_{affinity}$ is the coupling affinity, and $N_{polymerase}$ represents the quantity of polymerase available per promoter.

3.4 Effect of Copy Number Augmentation

In order to further address the flux discrepancy between the ethanol module and fatty- acyl CoA + *atfA* module (pCT69), the copy number of pCT69 was increased by removing the repressor of primer (ROP) site from the plasmid. The ROP site acts to decrease the rate of replication of the origin of replication of the plasmid backbone. Though the pETite system is not available without the ROP from the manufacturer, literature has shown that removal of the ROP site from the pBR322 origin of replication used in the pETite backbone should yield a 3 to 5-fold increase in plasmid copy number per cell. (Kelly et al., 2009) One possible concern in this approach arises from the plasmid maintenance requirements of maintaining both a "high" and "low" copy number origin of replication within the same cell. The concern stems from the possible shortage of endogenous machinery required to replicate the plasmids in a consistent and predictable manner. The construction of the high copy number pCT69 was performed by Dr. Narayan Niraula using a design created by Donovan Layton.

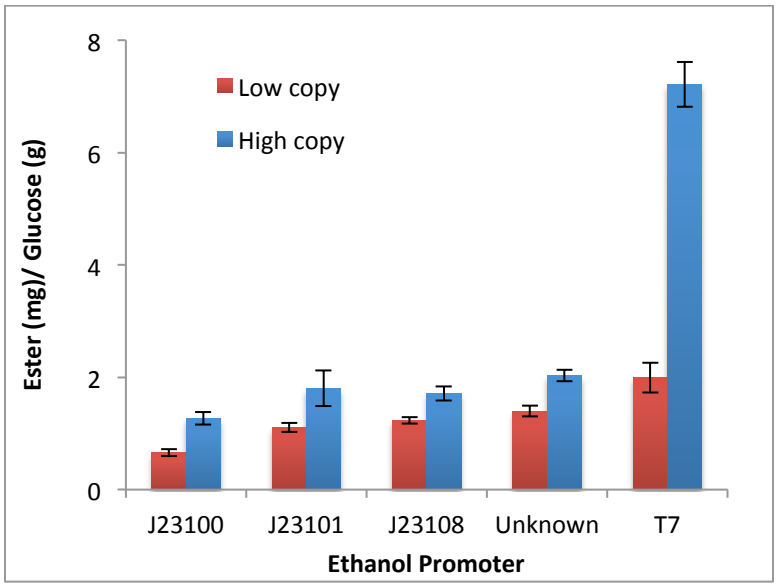

Figure 6 Effect of plasmid copy number on FAEE production.

All promoters were tested as well in combination with the low and high copy number. These results were generated using the high OD method and the experiment was run in triplicate.

Copy Number Augmentation Results

As shown in Figure 6, the yields in esters per gram glucose per OD increase between 1.4 to 3.6 fold in comparison to the low copy equivalent. The greatest change in ester yield comes from the T7 promoter construct. Promoters J23101, J23108, and the unknown mutant promoter all displayed approximately the same yields, though all increased over the baseline. Error bars represent one standard deviation from the average and the experiment was performed in triplicates using the high OD method described above. Glucose was totally consumed by the 24 hour sample point used for yield analysis. Ethanol yield was on average 48% for all promoters in this experiment. Cells were induced as described in section 2.2 Strain Characterization.

Copy Number Augmentation Discussion

Increases in yield were observed in every one of the tested constructs though the yield increase was on the low side of the reported increase in copy number, but the shortcoming can be attributed to the concerns stated above with respect to plasmid replication stress. A differentiation between the engineered endogenous promoters is somewhat ambiguous, but there is a distinct separation between the T7 promoter and the constitutive, endogenous promoters. Based on the results observed here, a truly high copy number plasmid using a different origin of replication with a copy number >500 copies per cell may result in an increase in yield of up to 10 fold.

3.5 Effect of Operon Orientation on FAEE Production

An alternate method for increasing the "effective" copy number was devised that would address the concern associated with copy number related cell stress. Typical single plasmid multi operon constructions consist of each gene set being flanked by a promoter and terminator. This motif is the result of the effect observed when the further from the promoter a polymerase travels, the higher the likelihood of the polymerase prematurely ending transcription before reaching the further downstream genes. This is the main premise of codon optimization, since foreign codons have a significant negative impact on the transcriptional efficiency.(Tsui, 2006) The main advantage, then, to using a promoter and terminator flanking a gene set of two or more genes is that each gene has roughly the same level of expression across all of the operons by reducing the number of events likely to stall transcription. These characteristics, though, are not ideal for the system of study and so a hypothesis was formed that could result in a significantly increase expression rate of the downstream operon. This strategy was named the "run through" method, as the second operon would be transcribed twice as often as the first operon as a result

of the second operon being under the control of both the first and second promoter. This was achieved by elimination of the terminator between the first and second operon in a single plasmid system. The construct was generated by Dr. Narayan Niraula using Dr. Cong Trinh's design.

This method was expected to yield approximately twice of the base line construct, but with additional benefits possibly resulting from decreased plasmid replication stress, and reduced stress resulting from using only one antibiotic. One uncertainty arose from the way the endogenous polymerase would interact with the orthogonal T7 promoter.

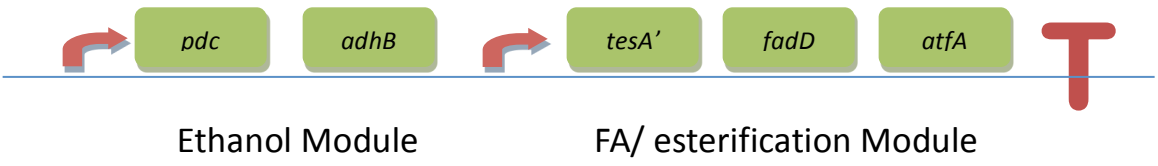

Figure 7 A representation of the single plasmid system orientation.

The implementation of a single plasmid as opposed to the dual plasmid system was expected to reduce variability and gave a good foundation to test the run through hypothesis.

The baseline construct, as well as the varying strength endogenous promoter constructs, were joined together into a single, low copy plasmid (Figure 7). Note that a single terminator was used to terminate transcription for both the first and second operons.

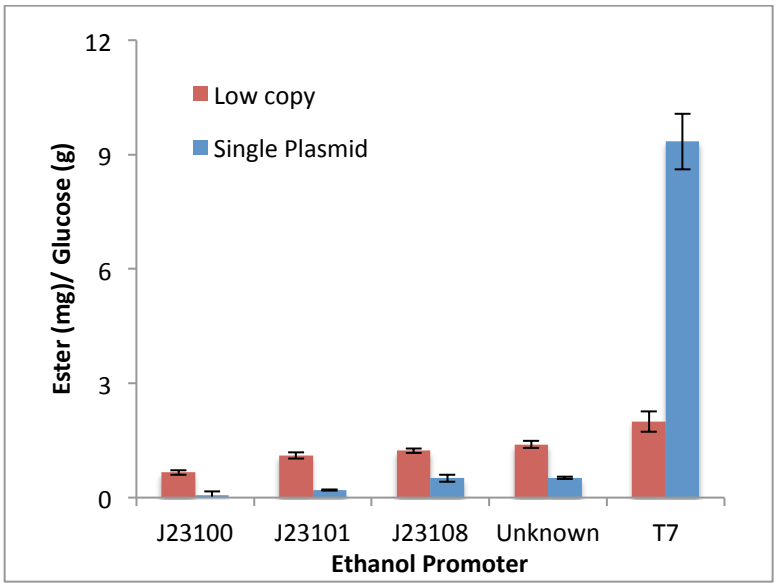

Figure 8 Effects of the single plasmid system in comparison to the baseline dual plasmid systems on FAEE production.

Joining the dual plasmids from the baseline study into a single plasmid was expected to generate an increase in yield due to a decrease in plasmid and antibiotic stress. These results were generated using the high OD method and the experiment was run in triplicate.

Single Plasmid Results

The single plasmid system generated a 4.7-fold increase in ester yield in the single plasmid system that utilized T7 promoters for both operons. In the mixed promoter system, when one endogenous and one orthogonal promoter was used, a significant decrease in ester yield was observed. Error bars represent one standard deviation from the average and the experiment was performed in triplicates using the high OD method described above. Glucose was totally consumed by the 24-hour sample point used for yield analysis in the dual T7 promoter system. Ethanol yield was on average 0.47g ETOH/g glucose for this particular case. For the mixed

promoter system, ethanol yield averaged 33% with only 4g of glucose being consumed by the 24 hour sampling point. Cells were induced as described in section 2.2 Strain Characterization.

Single Plasmid Discussion

Surprisingly there was a nearly five-fold increase over the base line for the dual T7 promoter case, which supports the "run through" effect. The hypothesis states that without a terminator separating the first and second operon, the second operon would benefit from an increase in expression rate as a multiple of the number of promoters preceding it. The run through effect in this experiment manifested itself by increasing the expression rate of the fatty-acyl CoA operon relative to the ethanol operon, and as a result produced a greater ester field.

Significance of Single Plasmid Run Through Effect

These findings are significant from a gene and pathway network regulation standpoint because an entirely new set of tools is now available to the synthetic biology community. First, a tight expression coupling of the first and second operon is achieved by virtue of using a single plasmid. Effects of differing copy number inherent to dual plasmid systems are eliminated by use of a single plasmid. But where-as previous studies have had to reduce an operon expression by using weaker promoters to achieve desired expression ratios, and as a result compromising the maximal expression rate, this method allows scaling greater than the strongest promoter. The drawback of the traditional motif is particularly problematic in constructs where genes of low enzymatic activity are required for product synthesis, where higher expression directly results in increase yield. With the operon motif utilized above, the maximal expression rate for each operon is a function of not only promoter strength, but also operon location relative to the first operon.

One can envision a scenario where genes are organized from most enzymatically active to least enzymatically active, with promoters being finely tuned to the needs of each gene's enzymatic activity. Implementation of the run through method offers the advantage over current methods in that the enzyme of highest activity can begin with a strong promoter and subsequent genes have expression rates even greater than previously possible as a result of the "run through" multiplier effect.

Furthermore, the traditional motif for gene expression, which depends strictly on the influence of only one promoter, can be expanded upon by possible incorporating two or more promoters per single operon. The hypothesis being that transcriptional rate can be expanded beyond a single promoter which is not simply a function of coupling strength between the polymerase and promoter sequence and number of available polymerases, but also the number of docking sites per operon. This grouping of promoters has been named a promoter cluster. The expression rate of an operon downstream of a promoter element can be defined as:

$$R_t = f(F_{affinity}, N_{polymerase,} N_{promoters}),$$

Equation 3 A proposed equation accounting for the run through effect.

Where R_t is the transcriptional rate, and $F_{affinity}$ is the coupling affinity, and $N_{polymerase}$ represents the quantity of polymerase available per promoter, and $N_{promoters}$ represents the number of stream promoters.

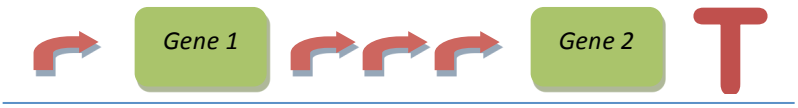

Figure 9 A proposed construct for high expression of the second operon beyond the limitations present in current expression motifs.

Diagramed in Figure 9 is a possible solution to greatly increasing the expression rate of a gene (gene 2), which has relatively low enzymatic activity without having to use techniques that reduces the expression rate of an upstream gene (gene 1) to achieve the desired expression ratio.

Run Through Effect as a Repressor System

Equally surprising was the near deactivation of the second operon in the mixed promoter system as demonstrated by overlaying the baseline system with the single plasmid system (Figure 8). Repression of the second operon, characterized by the decrease of ester production, followed closely the strength of the ethanol promoter from a relative strength of 1 to 0.51. Without more evenly spread data points it is not possible able to establish a range for which this correlation holds true. Though based on the extrapolation of the line of best fit, a promoter with a relative strength of 0.354 may be the point in which the up-stream operon has no to negligible repression effect in the mix promoter system.

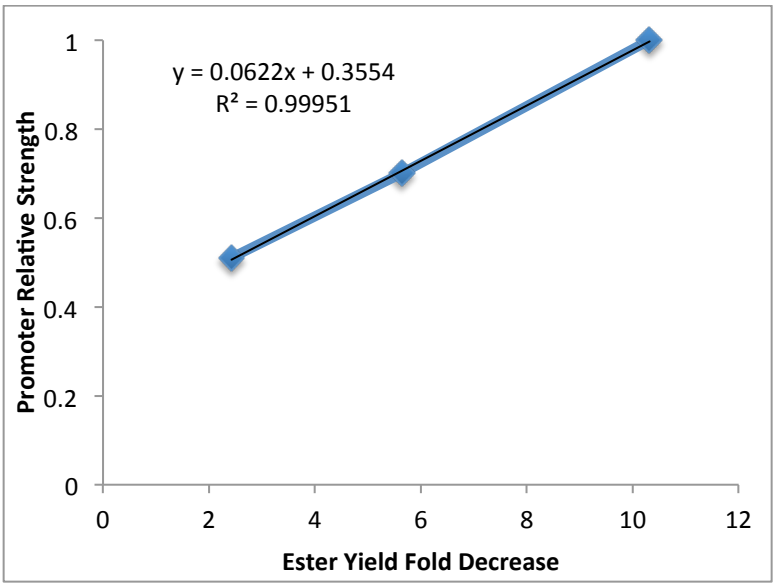

Figure 10 The relation between promoter strength and repression effect in the single plasmid mixed promoter system.

These results were generated using the high OD method and the experiment was run in triplicate.

The initial explanation for this effect was based on the terminator like activity of the second promoter and so some sort of secondary structure was expected to be formed by the promoter region that is polymerase specific. This would perhaps explain why when two T7 promoters were used, the polymerase passed unhindered, but when the endogenous polymerase encountered the T7 promoter, it stalled as if a terminator was present. This hypothesis, though, would not be consistent with the observation of the promoter strength to repression correlation observed in Figure 10. Additionally, no literature found supports the idea of secondary structures being present in prokaryotic promoters. Additional experimentation, such as the utilization of red fluorescent protein and green fluorescent protein in series, will need to be performed to obtain direct observation of the hypothesized phenomena.

The mix promoter repression effect does not provide an advantageous effect for ester production in the orientation used in this experiment, but it does provide for a method to establish inverse expression relation between the first and second operon by implementing a repressor site into the first promoter.

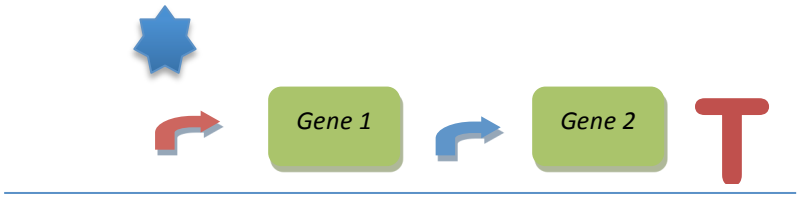

Figure 11 A proposed construct to demonstrate an inverse gene expression system.

Figure 11 provides a possible solution that takes advantage of the mixed polymerase system to express one operon or the other operon. The first promoter features a repressor site (such as lacI) so that when the first operon is activated by repressor removal, then the second operon is inactive. Conversely, when the first operon is repressed, then the second operon is active.

3.6 Effect of Alternate Ethanol Pathway on FAEE Production

To this point, the maximal total ester production has been ~145mg/L from 20g/L glucose, with ethanol production reaching near theoretical limits. Plasmid pCT24 utilizes genes that were originally selected by Dr. Ingram because of the extremely high enzymatic rates. An alternate and less enzymatically active ethanol production pathway was proposed to help drive

flux towards the fatty-acyl CoA module. The endogenous *adhE* gene was inserted into the low copy plasmid and promotion was controlled by the T7 promoter and library of promoters discussed above. These plasmids were then co-transformed with the high copy number pCT69. The dual plasmid system was then characterized in the same manner as the above experiments.

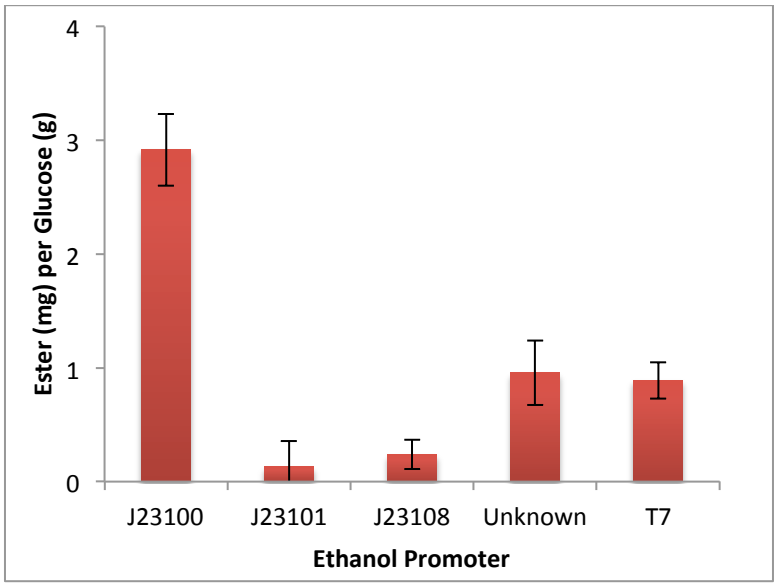

Figure 12 Effect of an alternate ethanol pathway on FAEE production.
These results were generated using the high OD method and the experiment was run in triplicate.

Alternate Ethanol Pathway Results

The trend observed in the "slow" ethanol pathway coupled with the high copy fatty acyl-CoA module is not as consistent as that observed of the "fast" ethanol pathway. The strongest

endogenous promoter had the highest yield, while the second and third strongest followed no particular trend. Besides the promoter J23100, all variations resulted in a decreased yield as compared to their "fast" ethanol pathway counterparts. Error bars represent one standard deviation from the average and the experiment was performed in triplicates using the high OD method described above. Glucose was not totally consumed by the 24-hour sample point used for yield analysis and was not consistent between experiments within this set. Ethanol yield was on average 0.45 g ETOH/g glucose for all promoters in this experiment. Cells were induced as described in section 2.2 Strain Characterization.

Alternate Ethanol Pathway Discussion

These results show some promise in that the strongest ethanol promoter, where the greatest amount of ester was produced, the ethanol yield decreased from the near theoretical (~ 0.5 g ETOH/ g glucose) observed in the above experiments to an ethanol yield of 0.36 g ETOH/ g glucose. This indicates that the overexpression of *adhE* did in fact have an impact on the ethanol production yield as hypothesized. Unfortunately, the ester yield was not as high as observed in the "fast" ethanol pathway. What this indicates is that an equi-molar balance between the ethanol and fatty acyl-CoA pathways may not in fact yield the maximal ester yield due to system complexities not yet understood. One hypothesis for explaining the above observation is that there is a minimal amount of ethanol required to drive the equilibrium of AtfA into ester formation. This is a troubling hypothesis in that the theoretical maximal yield is not possible to obtain since there must always be a sufficiently large excess of ethanol to drive the enzyme equilibrium. From the above observation and those from the high copy number experiments, there is room for argument that AtfA is insufficient for increasing product

formation rates and hindering the maximal yield obtainable. An alternate wax-ester synthase or modification of AftA for better performance should be considered for further work.

3.7 Effect of the Engineered Strain on FAEE production

To establish the increase yield as a result in using an engineered strain, the pCT69 with removed ROP plasmid and pCT24 were co-transformed into a wild type strain MG1655 that was transfected with the DE3 expression system and the engineered strain. Briefly, the engineered strain was devised by implementing a metabolic engineering technique known as elementary mode analysis for optimal ethanol production under anaerobic conditions. The strain was further modified for this study by knocking-out the *fadE* gene. While this strain was expected to generate higher yields, verification was necessary.

Engineered Strain Results

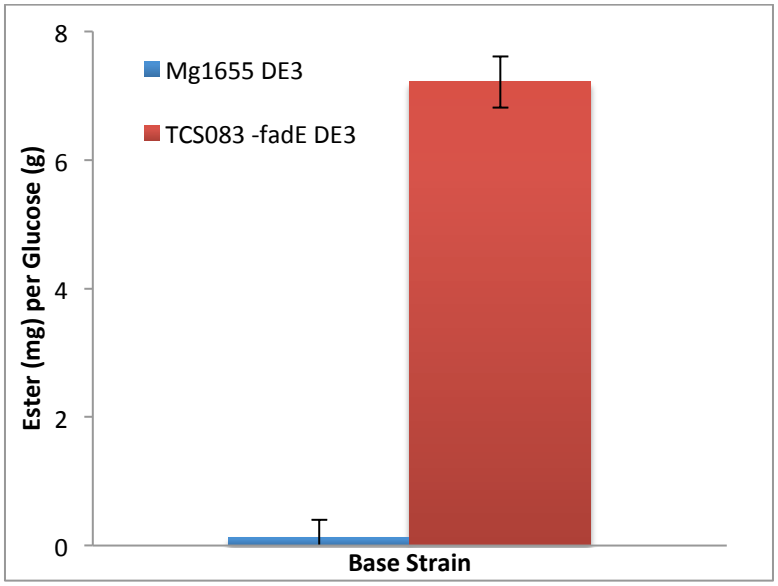

Figure 13 Effect of strain engineering in comparison to the wild type on FAEE production. These results were generated using the high OD method and the experiment was run in triplicate.

As shown above, a 56-fold increase was observed by utilizing the engineered strain in comparison to the parent strain with the same construct combination. Error bars represent one standard deviation from the average and the experiment was performed in triplicates using the high OD method described above. Glucose was totally consumed by the 24 hour sample point used for yield analysis. Ethanol yield was 0.08g ETOH/ g glucose for the wild type strain and 55% for the engineered strain. Cells were induced with IPTG as described in section 2.2 Strain Characterization.

Engineered Strain Discussion

Metabolic engineering has shown in the past to be able to provide a logical method to analyze a metabolic network and predict which set of genes to knock out to obtain desired results. Here we utilized strain TCS083 in comparison to the parent wild type strain MG1655. In the test conditions examined in this experiment, we see the engineered strain drastically outperform the wild type strain as expected.

3.8 Effect of Utilization A Mixed Strain Culture on FAEE Production

Thus far the synthetic biology toolbox utilized in this project has been limited to adjusting transcription rates within a single organism for a single carbon source. Innovations in the field of metabolic engineering have developed methods to custom tailor microbes for specific carbon sources to better address the mixed sugars typically found in industrial feedstocks. Here we demonstrate the use of a strain that can co-utilize 5-carbon sugar, xylose, and 6-carbon sugar, glucose, housing the low copy fatty acyl-CoA plasmid (TCS083 Δ*fadE* (DE3)), with a strain engineered to only utilize 5-carbon xylose to produce ethanol (CT1101 DE3). (Trinh et al., 2008a) The high OD method was modified such that an OD of 5 of each of the strains were mixed in equal volumes in a broth of 10g/L glucose and 10g/L xylose. The strain cocktail should be directly comparable to the baseline established initially.

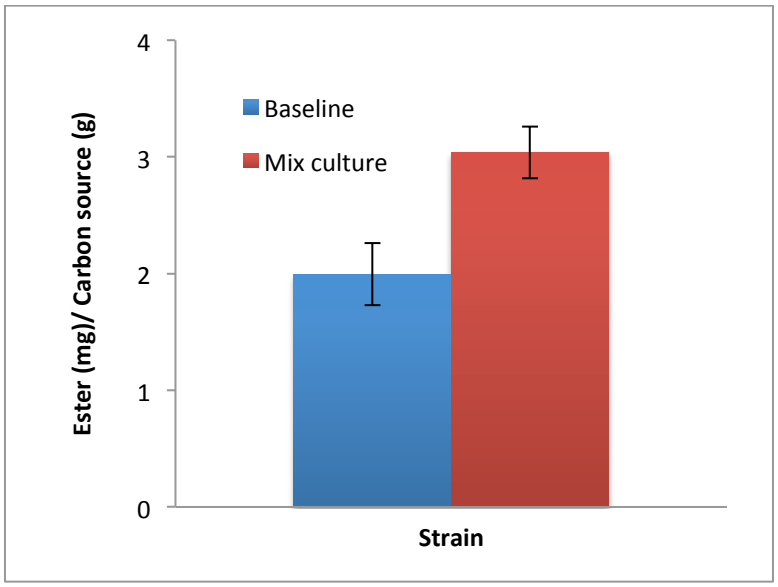

Figure 14 Effect of single and mixed cell cultures on FAEE production.

These results were generated using the high OD method and the experiment was run in triplicate.

Mixed Culture Results

A 1.5-fold increase in ester yield per gram carbon source over the baseline was observed by using the mixed culture. Error bars represent one standard deviation from the average and the experiment was performed in triplicates using the high OD method described above. M9 media with 10g/L glucose, 10g/L xylose, and 5g/L yeast extract was used for this. Cells were induced as described in section 2.2 Strain Characterization.

Mixed Culture Discussion

These results indicate that carbon substrate targeting is a viable technique to reducing the challenges associated with trying to drive flux from one pathway to another within a single organism. An additional advantage of the multi strain method stems from the ability of the

desired pathway to act independently of other product formation critical pathways. This permits the use of strains that are specifically engineered to perform a single task, such as produce ethanol only, without the need for compromise that could arise from production of multiple substrates, such as those observed in the production of FAEE. Using a microbe cocktail, such as the one demonstrated in this experiment may have significant implications for industrial applications. Biofuel production is often subject of varying carbon sources and as a result carbon compositions distribution. Expansion of the method described above would permit not only fine-tuning the strains used, but also the ratio of strains used to optimize the carbon utilization.

3.9 Effect of Varying Plasmid Copy Number on Chain Length Distribution of FAEE

One concern that arises from increasing the expression rate of the fatty acyl-CoA operon is the possibility of altering the carbon chain distribution. The precursor of the overexpressed gene *'tesA* is the product of the fatty acid elongation cycle. The elongation process adds two carbons per cycle onto a carbon chain until the carbon chain is "pulled out" of the cycle. Since one of the techniques employed to increase ester yields was to increase the copy number of the fatty acyl-CoA operon, it was hypothesized that overexpression could effectively increases the "pull" on the carbon chain out of the elongation cycle and result in premature esterification before the full chain elongation process was completed. This effect, coupled with the promiscuity of *'tesA*, may result in a shift towards shorter carbon lengths that are shorter due to an increased propensity to pull carbon chains out before they are fully matured into the preferred 14 carbon length. The average distribution of both the high and low copy number plasmids coupled with the "fast" ethanol pathway were compared to determine if there was a statistically relevant shift.

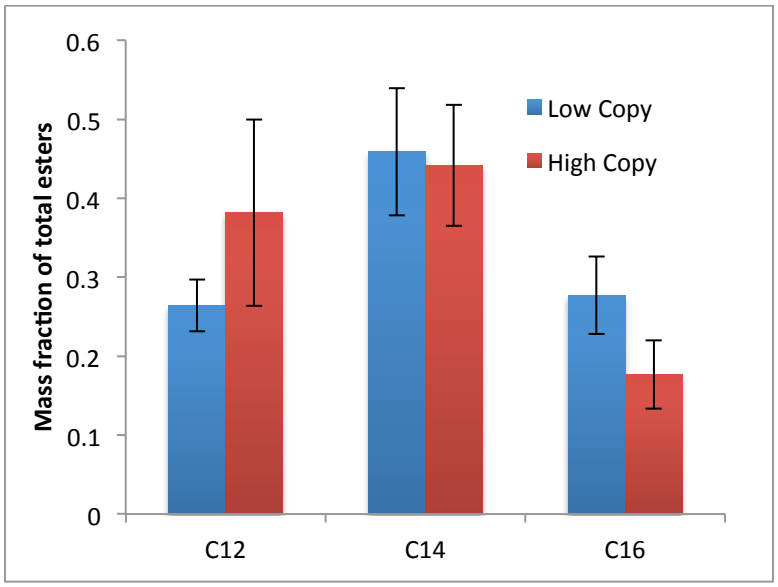

Figure 15 Effect of the fatty-acyl CoA + atfA plasmid copy number on FAEE moiety distribution.

These results were generated using the high OD method and the experiment was run in triplicate.

Chain Length Distribution Results

Figure 15 shows that the C14 carbon length distribution remains constant, though the C12 and C16 carbon length distributions appear to be on average different. C16 production, which is the product of a greater number of elongation cycles appears to have decreased in the high copy number strain. C12 production, which is a product of fewer number of elongation cycles, appears to have increased with the high copy number strain. Error bars represent one standard deviation from the average and the experiment was performed in triplicates using the high OD method described above.

Chain Length Distribution Discussion.

These perceived differences, though, are within one standard deviation and so a larger sample set would be required before any statistically significant claims could be made. If these findings prove to be a real phenomenon, a troubling forecast for even higher FAEE production is observed. The consequences of this observation being true have implications for the quality of fuel being produced. Biodiesel combustion properties are generally described by a cetane number. The greater the cetane number, the better the combustion properties are. As the base chain fatty acid that joins with the alcohol shifts to shorter chain lengths, the cetane number decreases and the fuel quality decreases. If the above suggested methods for increasing ester yields are implemented, then there is a risk for creating fuels that are not suitable for use. As a result, a new biodiesel is being proposed to counteract the shortening of the ester length. The use of an alternate alcohol base could restore the cetane number.

3.10 Generating Biodiesel with Enhanced Combustion Properties

Many studies have been performed on calculating the cetane number for different alcohol and fatty acid combinations, but often these studies do not test every combination at once for one reason or another. Unfortunately these tests are subject to testing protocol specifics and comparison between one study to another is not possible. Therefore absolute values of the cetane number obtained from these studies are less relevant than the trends that will be analyzed from the data gathered.

Table 4 Summary of cetane numbers for esters of different fatty acid and alcohol moieties.

Though the same pure esters were being tested, each testing protocol generated different absolute values, though the trends in cetane number remained constant. The source for each table is located in the top left corner of each table.

Moser, 2009	Alcohols Esterified with					**4.A**
Fatty Acid Length	Methyl	Ethyl	Isopropyl	Propyl	Isobutyl	Butyl
C12	67					
C16	86	93				
C18	101	97				92

Knothe, Matheaus, Ryan, 2002	Alcohols Esterified with					**4.B**
Fatty Acid Length	Methyl	Ethyl	Isopropyl	Propyl	Isobutyl	Butyl
C16	85.9	93.1	82.6	85	93.6	91.9
C18	101	97.7	96.5	90.9	99.3	92.5

Klopfenstein,1985	Alcohols Esterified with					**4.C**
Fatty Acid Length	Methyl	Ethyl	Isopropyl	Propyl	Isobutyl	Butyl
8	33.6					39.6
10	47.2	51.2	46.6	52.9		54.6
12	61.4					
14	66.2	66.9				
16	74.5					
18	86.9					

No literature was found that tested every ester and fatty acid combination of interest to this study. However, by plotting the trends of fatty acid base length versus cetane number and also alcohol length, a general trend can be determined to provide direction for improving fuel properties. For reference, petroleum based diesel has a reported cetane number in the mid 40's to mid 50's depending on testing method used. (Klopfenstein, 1985; Knothe, Matheaus, & Ryan, 2002; Moser, 2009)

Effect of Ester Composition on Cetane Number.

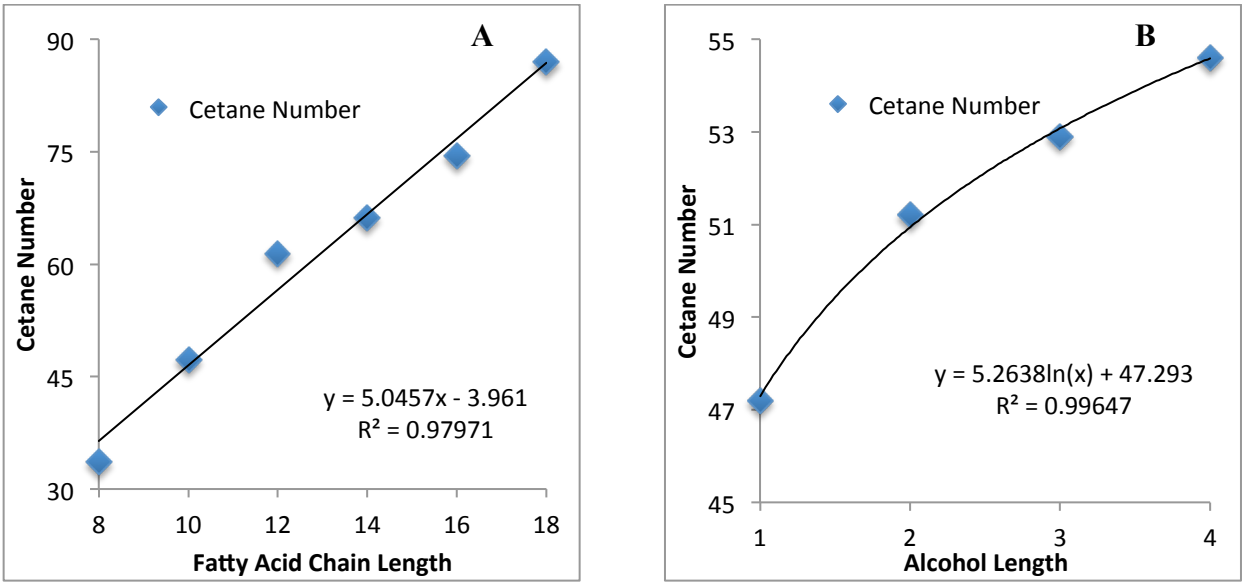

Figure 16 Plotting the effect of fatty acid and alcohol carbon chain length on the cetane number.

Panel A shows methyl ester cetane numbers of various fatty acid chain lengths. Panel B shows the cetane numbers of methyl, ethyl, propyl, and butyl decyl esters.

From the figures using data generated by Klopfenstein, the relation of methyl esters of differing fatty acid chain length can be clearly observed. The slope in the range of fatty acids of interest (C8 to C18) shows a 5:1 increase in cetane number per increase in carbon chain length for methyl esters. Additionally, data of the relation between linear alcohol chain length and cetane number closely follows a positive logarithmic trend in the range of alcohols of interest to this study.

Effect of Ester Composition on Cetane Number Discussion.

From the trends observed above, there is evidence to support that increasing alcohol chain length is a plausible method for combating the effects of decreasing fatty acid chain length. To date several longer linear chain and branch chained alcohols have been produced in *E. coli*. Propanol and butanol (Klopfenstein, 1985; Knothe et al., 2003; Moser et al., 2013), isopropanol, (C R Shen & Liao, 2008) and isobutanol (Inokuma, Liao, Okamoto, & Hanai, 2010; Lee et al., 2012) production have all been demonstrated endogenously in *E. coli* and *AtfA* has been demonstrated to be accepting of a wide array of alcohols, with greatest activity observed in the medium chain length alcohols (C14 to C18) using enzyme assays. (Stoveken, Kalscheuer, Malkus, Reichelt, & Steinbuchel, 2005)

Establishing feasibility.

A doping approach was devised to determine if *atfA* had sufficient activity for the relatively short alcohols thus far produced endogenously in *E. coli*. First the fatty acyl-CoA + *atfA* plasmid was expressed in the strain BL21 (λDE3) and was grown using standard aerobic characterization methods in double standard concentration LB to eliminate any variable associated with media and oxygen limitations. The media was doped with 5 g/L of different alcohols to determine the range of alcohols that could be used to replace the ethanol pathway.

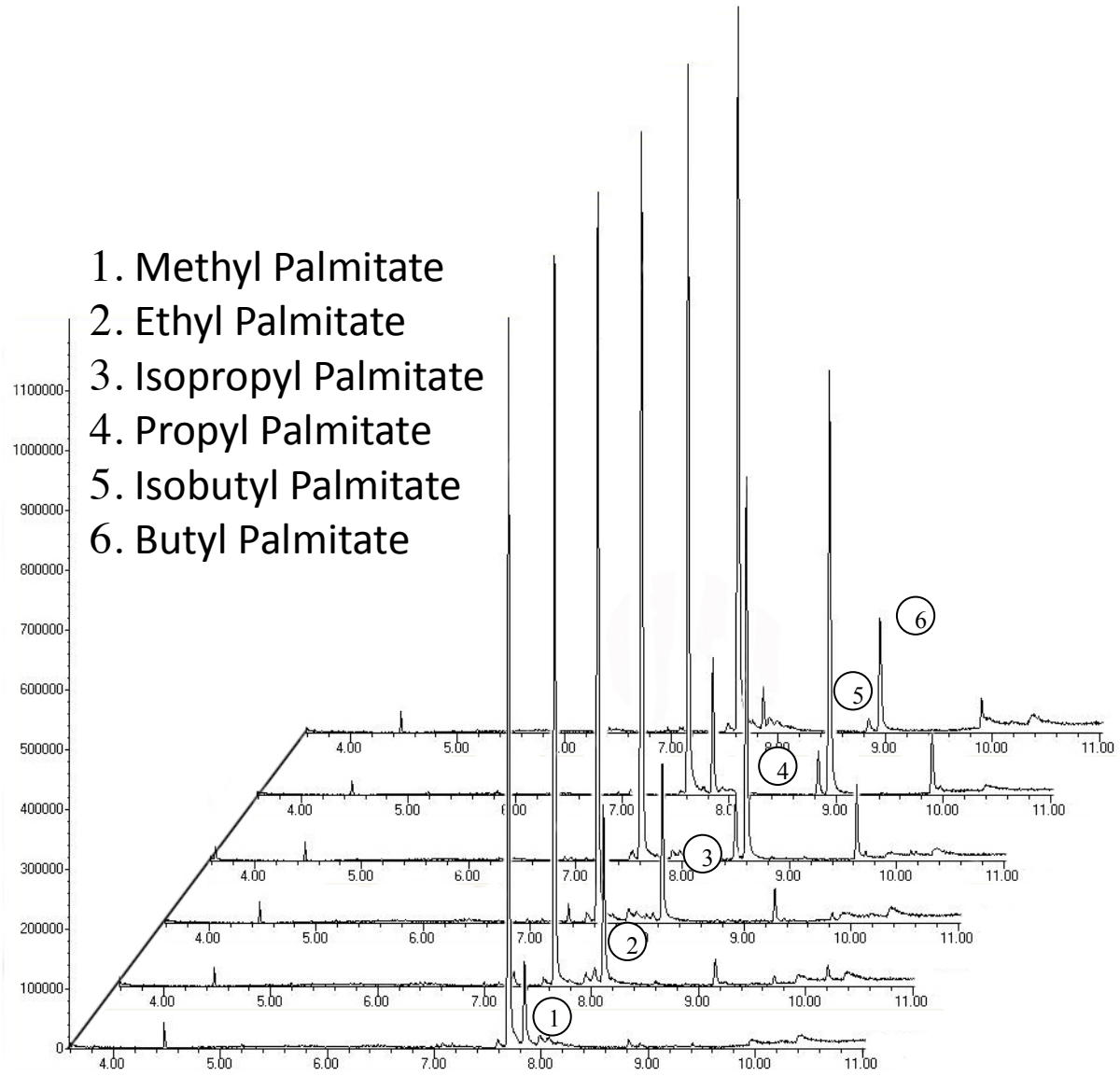

Figure 17 Overlaid chromatograms of different esters produced by the doping experiment.

Figure 17 depicts an overlay of esters formed from the doping experiment. The peak at ~7.7 minutes is C15 ethyl ester internal standard, the subsequent peaks are the novel esters formed. Based on these results, it was determined that the esterification gene, *atfA*, is sufficiently

broad in substrate uptake to be suitable for further experimentation in vivo. This is the first known demonstration of ester production with alcohol doping other than ethanol using endogenously produced fatty acids.

Alternate Alcohol Production

The isobutanol producing plasmid, pCT13, and butanol producing plasmid pButanol High Copy were expressed individually in TCS083 with the *fadE* knockout. The experiment was performed in duplicates using the high OD characterization method described above.

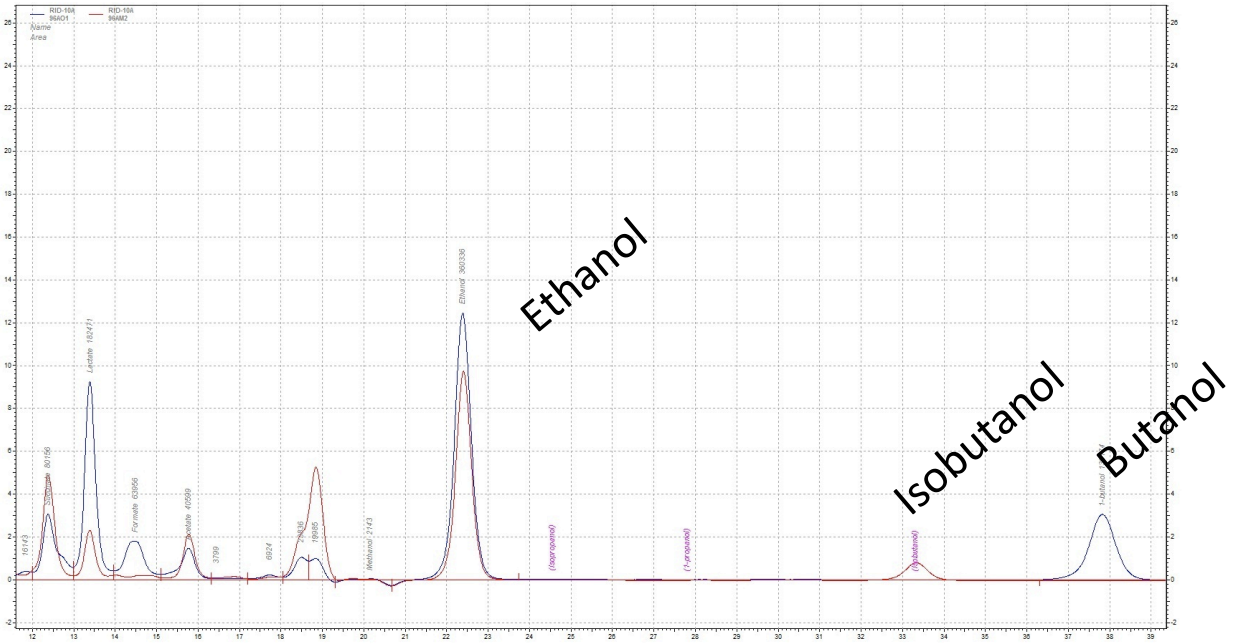

Figure 18 Overlaid HPLC chromatograms demonstrate the production of isobutanol and butanol by expressing pCT13 and pButanol respectively in TCS83 Δ*fadE* (DE3)

Alternate Alcohol Production Discussion

Production of isobutanol and butanol was demonstrated using TCS083 with the *fadE* knockout as expected. pCT13 is the exact plasmid demonstrated in Dr. Cong Trinh's prior work (Trinh et al., 2011) and pBUTANOL was created by Donovan Layton based on a previously demonstrated butanol construct. (Claire R Shen et al., 2011). In the above diagram it is worth noting that in both isobutanol and butanol production studies, ethanol was the primary alcohol produced at a rate of ~3:1w/w for isobutanol and ~4.6:1w/w for butanol. This result can be partially be explained by the native ethanol pathway still being present, and the knockout of *adhE* may attenuate this issue. Knocking out *adhE* is not likely to resolve this issue entirely as it is likely that the alcohol dehydrogenases employed by pCT13 and pBUTANOL have some amount of ethanol production capability due to less than ideal substrate specificity. Further studies in an *adhE* knockout strain would need to be performed to verify this conclusion.

Having co-production of ethanol and isobutanol or butanol is likely to produce a mixture of ethyl and butyl/isobutyl esters. As demonstrated above in the preliminary alcohol doping study, *atfA* was able to readily accept any of the tested alcohol as a substrate for ester formation. This characteristic presents a unique opportunity to fine-tune the physical properties of the biodiesel produced. Other characteristics of biodiesel that have not yet been discussed but are important for transportation fuel utilization performance such as cloud point, melting point, oxidative stability, and viscosity are also a function of alcohol and fatty acid base chain length. (Klopfenstein, 1985)

Endogenous Production of Different Base Alcohol Esters.

Attempts to endogenously produce fatty acid isobutyl and butyl esters has failed in TCS083 Δ*fadE* (DE3) knockout in the M9 defined media and high OD method used in the previous ester production studies. Isobutanol and butanol were produced as expected, though at low concentration (1.3 and 0.65 g/L respectively), it may be possible that the observations from the "slow" ethanol pathway are collaborative with these results and a minimal alcohol concentration is required for *atfA* to perform fatty acid esterification.

In an attempt to eliminate any possible limitations associated with media and anaerobic growth conditions. A wild type strain BL21 (λDE3), was transformed with an alternate version of pCT69 that retains the leading sequence on *tesA* and pCT13. This strain was then grown in double standard concentration LB with 40g/L glucose in an aerobic baffle flask with samples collected at 24, 36, 48, and 96 hours. The selection of these growing conditions was based on concerns of inadequate micronutrients available in the M9 media utilized in the ethanol derived ester formation studies above. The experimental conditions present in this experiment do not represent an optimal approach for industrial applications since BL21 is known to produce very large amounts of acetate during aerobic growth which reduce maximal ester yields.(Knothe et al., 2003)

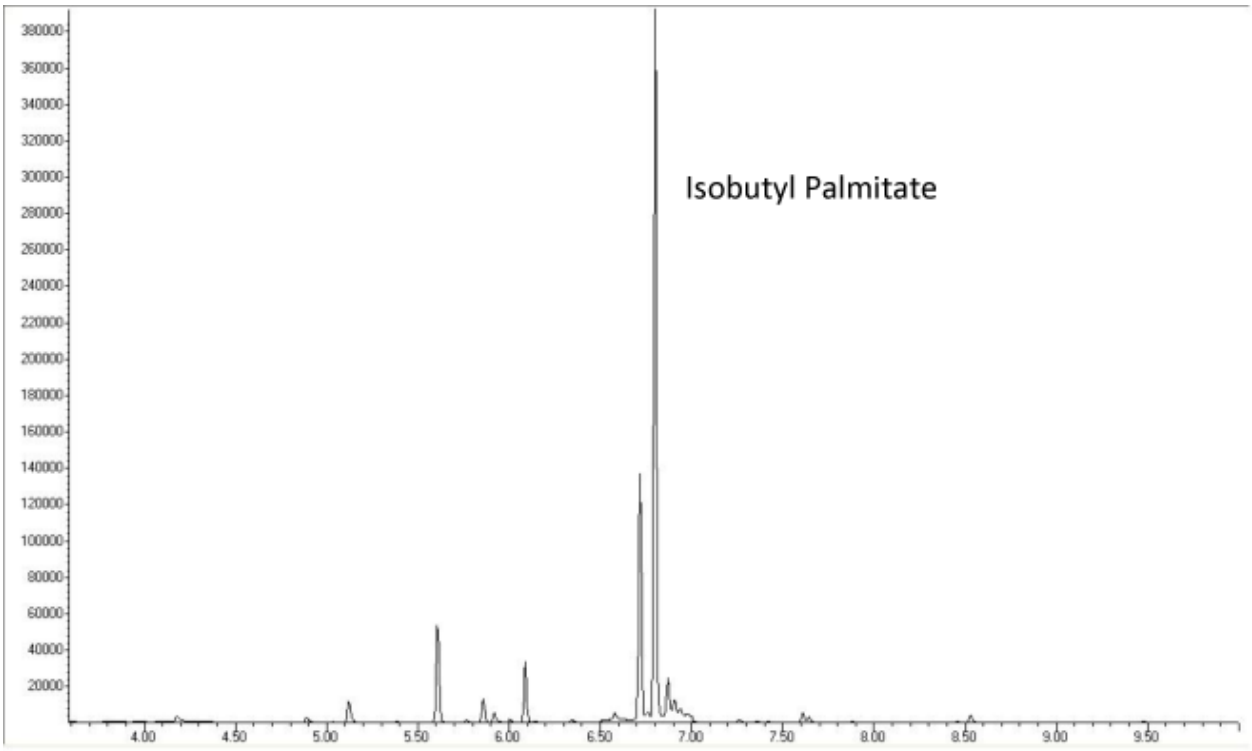

Figure 19 GCMS chromatogram demonstrates the production of endogenous isobutyl palmitate.

BL21 (DE3) with pCT63 + pCT13 using double standard concentration LB with 40g/L glucose grown aerobically was used to demonstrate the first endogenous production of isobutyl palmitate.

The results of the BL21 (DE3) characterization show a significant production isobutyl palmitate. Total ester production was reached 55.6 mg/L with isobutyl palmitate (C16 base fatty acid) contributing to 84% of total esters formed.

This non-optimal characterization shows that fully endogenous isobutyl ester formation is possible and offers a path for remedying decreased biodiesel cetane number observed in the optimization methods employed above. Further work with butanol as well as other alcohols is in progress to identify the most promising alcohol source for further optimizations.

Chapter 4 Conclusions and Future Work

In this study, several methods were demonstrated for increasing ester yield by way of implementing synthetic biology techniques. The use of standardized promoters created a technique to accurately control operon expression in a predictable way. The only deviation from the expected was the unexpectedly low strength of the T7 promoter system investigated in this study. By removing the ROP from the low copy number backbone, the effective transcriptional rate of pCT69 increased in line with what literature has stated. These two techniques created an 11-fold range of yield based on manipulating these two variables. Further development is possible by exploiting promoters from the Anderson library that have a relative strength of 0.01 or lower in some cases. Furthermore, development of a library of T7 promoters with point mutations could have advantageous effects in future studies where endogenous promoters become susceptible to unpredictable regulation associated with cellular stress.

This study also demonstrated the potential and power hidden in novel operon orientations. The run through effect generated a 5-fold increase in ester yield as a result of extra opportunities for transcription. While a 5-fold increase in yield is significant, the most exciting development stems from the possibility of avoiding down regulation of operons as a method for transcription control. The emergence of downstream augmentation in transcription as a function of number of promoters means that the strongest promoter is no longer the ceiling for transcriptional rates. Additionally, the repressor effect of the "run through" orientation creates a method for controlling two operons with an inverse expression relationship. Equally exciting is that this inverse relationship is tunable based on the strength of the upstream promoter, so the magnitude of the effect can be fine tuned as well.

Implementation of the overexpressed *adhE* gene (slow ethanol pathway) resulted in unpredictable results and even more troubling the possibility of the existence of a minimal concentration of ethanol required to drive the esterification process. This poses a major obstacle if this observation is true, first because the theoretical yield can never be reached since an excess of ethanol is required. Secondly, expansion into other alcohols, where alcohol yields tend to be low, may not be feasible. The key focus of further development should be in identifying an *atfA* alternative, or to engineer *atfA* in a manner that increases short chain alcohol activity to reduce to excess ethanol requirement.

The study also presents a look into the possible obstacles that may result from development of higher yields and rates of ester production. While the data lacks statistical certainty, there appears to be a trending to shorter chain esters with an increase in pCT69 expression. This creates a troubling forecast in that shorter chain esters in general have poorer cetane number. An alternate alcohol and fatty acid combination has been proposed that should reduce the effect of the shift in cetane number predicted with higher yields and rates of ester production. Unfortunately, completely endogenous production of butyl and isobutyl esters has not been demonstrated anaerobically in this study, though aerobic production of isobutyl esters was demonstrated. The roadblock is hypothesized to be the result of AtfA's excess alcohol requirements and the poor productivity of the isobutanol and butanol constructs utilized in this study. This finding further reinforces the need to find or engineer and atfA alternative if truly high ester yields are to be reached in the future.

References

Atsumi, S., Li, Z., & Liao, J. C. (2009). Acetolactate synthase from Bacillus subtilis serves as a 2-ketoisovalerate decarboxylase for isobutanol biosynthesis in Escherichia coli. *Applied and environmental microbiology*, *75*(19), 6306–11. doi:10.1128/AEM.01160-09

Atsumi, S., Wu, T., Machado, I. M. P., Huang, W., Chen, P., Pellegrini, M., & Liao, J. C. (2010). Evolution , genomic analysis , and reconstruction of isobutanol tolerance in Escherichia coli. *Molecular Systems Biology*, *6*(449), 1–11. doi:10.1038/msb.2010.98

Atsumi, S., Wu, T.-Y., Eckl, E.-M., Hawkins, S. D., Buelter, T., & Liao, J. C. (2010). Engineering the isobutanol biosynthetic pathway in Escherichia coli by comparison of three aldehyde reductase/alcohol dehydrogenase genes. *Applied microbiology and biotechnology*, *85*(3), 651–7. doi:10.1007/s00253-009-2085-6

Conway, T., Sewell, G. W., Osman, Y. a, & Ingram, L. O. (1987). Cloning and sequencing of the alcohol dehydrogenase II gene from Zymomonas mobilis. *Journal of bacteriology*, *169*(6), 2591–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=212129&tool=pmcentrez&rendertype=abstract

Fortman, J. L., Chhabra, S., Mukhopadhyay, A., Chou, H., Lee, T. S., Steen, E., & Keasling, J. D. (2008). Biofuel alternatives to ethanol: pumping the microbial well. *Trends in biotechnology*, *26*(7), 375–81. doi:10.1016/j.tibtech.2008.03.008

Ikeda, R. A., & Richardson, C. C. (1986). promoter during, *83*(June), 3614–3618.

Inokuma, K., Liao, J. C., Okamoto, M., & Hanai, T. (2010). Improvement of isopropanol production by metabolically engineered Escherichia coli using gas stripping. *Journal of bioscience and bioengineering*, *110*(6), 696–701. doi:10.1016/j.jbiosc.2010.07.010

Kalscheuer, R., Stölting, T., & Steinbüchel, A. (2006). Microdiesel: Escherichia coli engineered for fuel production. *Microbiology (Reading, England)*, *152*(Pt 9), 2529–36. doi:10.1099/mic.0.29028-0

Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., … Endy, D. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *Journal of biological engineering*, *3*, 4. doi:10.1186/1754-1611-3-4

Klopfenstein, W. E. (1985). Effect of molecular weights of fatty acid esters on cetane numbers as diesel fuels. *Journal of the American Oil Chemists Society*, *62*(6), 1029–1031. doi:10.1007/BF02935708

Knothe, G., Matheaus, A. C., & Ryan, T. W. (2003). Cetane numbers of branched and straight-chain fatty esters determined in an ignition quality tester☆. *Fuel*, *82*(8), 971–975. doi:10.1016/S0016-2361(02)00382-4

Lee, J., Jang, Y.-S., Choi, S. J., Im, J. A., Song, H., Cho, J. H., … Lee, S. Y. (2012). Metabolic Engineering of Clostridium acetobutylicum ATCC 824 for Isopropanol-Butanol-Ethanol Fermentation. *Applied and environmental microbiology*, *78*(5), 1416–23. doi:10.1128/AEM.06382-11

Liu, T., & Khosla, C. (2010). Genetic engineering of Escherichia coli for biofuel production. *Annual review of genetics*, *44*, 53–69. doi:10.1146/annurev-genet-102209-163440

Lou, X., & Singh, P. M. (2010). Role of water, acetic acid and chloride on corrosion and pitting behaviour of carbon steel in fuel-grade ethanol. *Corrosion Science*, *52*(7), 2303–2315. doi:10.1016/j.corsci.2010.03.034

Lu, X., Vora, H., & Khosla, C. (2008). Overproduction of free fatty acids in E. coli: implications for biodiesel production. *Metabolic engineering*, *10*(6), 333–9. doi:10.1016/j.ymben.2008.08.006

Ma, F., & Hanna, M. A. (1999). Biodiesel production : a review, *70*, 1–15.

Moser, B. R., In, S., Cellular, V., Plant, D. B., & Jun, B. M. (2013). Biodiesel properties , and feedstocks, *45*(3), 229–266.

Periana, R. A., Taube, D. J., Evitt, E. R., Löffler, D. G., Paul, R., Voss, G., … Wentrcek, P. R. (1993). A Mercury-Catalyzed , High-Yield System for the Oxidation of Methane to Methanol, *259*(5093), 340–343.

Rao, C. V. (2012). Expanding the synthetic biology toolbox: engineering orthogonal regulators of gene expression. *Current opinion in biotechnology*, 1–6. doi:10.1016/j.copbio.2011.12.015

Shay, E. G. (1993). Diesel fuel from vegetable oils: Status and opportunities. *Biomass and Bioenergy*, *4*(4), 227–242. doi:10.1016/0961-9534(93)90080-N

Shen, C R, & Liao, J. C. (2008). Metabolic engineering of Escherichia coli for 1-butanol and 1-propanol production via the keto-acid pathways. *Metabolic engineering*, *10*(6), 312–20. doi:10.1016/j.ymben.2008.08.001

Shen, Claire R, Lan, E. I., Dekishima, Y., Baez, A., Cho, K. M., & Liao, J. C. (2011). Driving forces enable high-titer anaerobic 1-butanol synthesis in Escherichia coli. *Applied and environmental microbiology*, *77*(9), 2905–15. doi:10.1128/AEM.03034-10

Smith, K. M., & Liao, J. C. (2011). An evolutionary strategy for isobutanol production strain development in Escherichia coli. *Metabolic engineering*, *13*(6), 674–81. doi:10.1016/j.ymben.2011.08.004

Stark, W. H. (Vickers-V. P. E. C., & Pohler, G. M. (The V. C. & S. C. (1950). Sterile Air for Industrial Fermentations, (September), 1789–1792.

Steen, E. J., Kang, Y., Bokinsky, G., Hu, Z., Schirmer, A., McClure, A., … Keasling, J. D. (2010). Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature*, *463*(7280), 559–62. doi:10.1038/nature08721

Stoveken, T., Kalscheuer, R., Malkus, U., Reichelt, R., & Steinbuchel, A. (2005). The Wax Ester Synthase / Acyl Coenzyme A : Diacylglycerol Acyltransferase from Acinetobacter sp . Strain ADP1 : Characterization of a Novel Type of Acyltransferase, *187*(4), 1369–1376. doi:10.1128/JB.187.4.1369

Trinh, C. T., Li, J., Blanch, H. W., & Clark, D. S. (2011). Redesigning Escherichia coli metabolism for anaerobic production of isobutanol. *Applied and environmental microbiology*, *77*(14), 4894–904. doi:10.1128/AEM.00382-11

Trinh, C. T., Unrean, P., & Srienc, F. (2008a). Minimal Escherichia coli cell for the most efficient production of ethanol from hexoses and pentoses. *Applied and environmental microbiology*, *74*(12), 3634–43. doi:10.1128/AEM.02708-07

Trinh, C. T., Unrean, P., & Srienc, F. (2008b). Minimal Escherichia coli cell for the most efficient production of ethanol from hexoses and pentoses. *Applied and environmental microbiology*, *74*(12), 3634–43. doi:10.1128/AEM.02708-07

Tsui, E. V. A. (2006). Determination of Exclusion Effects of Potentially rop Deficient Mutant pBR322 during Co-transformation with the Wild-type Plasmid, *10*(December), 23–26.

Wheals, a E., Basso, L. C., Alves, D. M., & Amorim, H. V. (1999). Fuel ethanol after 25 years. *Trends in biotechnology*, *17*(12), 482–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10557161

Zhang, F., Carothers, J. M., & Keasling, J. D. (2012). Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnology*, (February), 1–7. doi:10.1038/nbt.2149

Vita

Michael Wierzbicki was born in New York City, New York, to the parents of Maria and Krzysztof Wierzbicki. He does not have any brothers nor sisters. He attended John Marshall and C. C. Wells elementary schools, G. W. Carver Middle school and continued to Thomas Dale High School in Chester, Virginia. After graduation, he attended Virginia Commonwealth University where he obtained a bachelor in Chemical and Life Science Engineering in May 2011. He accepted a graduate research assistantship at The University of Tennessee, Knoxville, in the Chemical and Biomolecular Engineering. Michael graduated with a Master of Science degree in August 2013.