RICE UNIVERSITY

Novel approaches for the production of fuels and chemicals in *Escherichia coli*

by

Clementina Dellomonaco

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Ramon Gonzalez
Associate Professor in Chemical &
Biomolecular Engineering and Bioengineering

Laura Segatori
T.N. Law Assistant Professor in
Chemical & Biomolecular Engineering

Pedro J. Alvarez
G.R. Brown Professor of Engineering
Chair, Department of Civil & Environmental
Engineering

HOUSTON, TEXAS
July 2011

Abstract

Novel approaches for the production of fuels and chemicals in *Escherichia coli*

by

Clementina Dellomonaco

Volatility of oil prices along with major concerns about climate change, oil supply security and depleting reserves have sparked renewed interest in the production of biofuels and biochemicals. While the carbohydrate portion of edible crops is currently used as the primary feedstock in the biological production of fuels and chemicals, the availability of fatty acid (FA)-rich feedstocks and recent progress in the development of oil-accumulating organisms have drawn the attention to FAs as an attractive alternative. However, microbial platforms to enable this were nearly absent. To this end, we engineered native and heterologous fermentative pathways in *E. coli* to enable the efficient synthesis of fuels and chemicals from FAs. The current *de facto* strategy for the synthesis of non-native products in model organisms is Heterologous Metabolic Engineering (HeME), which consists of recruiting foreign genes from native producers. However, the relative incompatibility of the heterologous pathways with the host metabolism may be considered a drawback. As an alternative approach, the HoME (Homologous Metabolic Engineering) strategy that we propose overcomes this limitation by harnessing the metabolic potential of the host strain. HoME aims at reconstructing heterologous pathways to enable biosynthesis of non-natural products by identifying and

assembling native functional surrogates. Implementation of both HeME and HoME strategies in the context of fuels and chemicals biosynthesis has usually been directed to the conversion of feedstocks constituents into a specific product. However, we demonstrated a novel metabolic platform based on a functional reversal of the fatty acid catabolic pathway (β-oxidation) as a means of synthesizing a wide array of products with various chain lengths and functionalities.

Acknowledgments

Several people have deeply (and somehow unintentionally) contributed to its realization. However, acknowledging each and every one of them would be nearly impossible.

Although this dissertation signifies the conclusion of this experience, I like to think this moment more as a transition and share with you what I will bring with me from Rice.

I will bring with me:

… (and I hope to continue enjoying) the motivation, respect, ethic, personability and friendship of my advisor. His determination and excitement have made our weekly, daily, and even hourly research brainstorming very stimulating.

… the person that for five years has always been a phone call or an email away and has provided me with the greatest support.

… the Texan Italianized brother (and sister) that I recently acquired and with whom I shared close to every single day of my life as a PhD student along with approximately 1,165 morning coffees ☺.

… the laughter of our office and lab; I hope to have contributed enough misphrased sentences so that my colleagues can keep making fun of me for long time.

... the members of all the families I felt belonging to: the IACC, the ICCC and the Italian Expo for the great entertainment and the fun we have shared. The Pasquali family for having nominated me for 5 years in a row special guest to their dinners ☺.

...all my family (mamma, papa', Ale e nonna Italia) and my family in-law (Anna, Robel e Yonni) for the unconditional support and love, even if from far away.

... my *best* half for being always there for me, so lovely and caring.

Contents

List of Figures

List of Tables

Extended abstract

Novel approaches for the production of fuels and chemicals in *Escherichia coli*

by

Clementina Dellomonaco

Volatility of oil prices along with major concerns about climate change, oil supply security and depleting reserves have sparked renewed interest in the production of biofuels and biochemicals. Although several biomass feedstocks are available for the production of fuels and chemicals, carbohydrates have been so far the option of choice for bioconversions. Presently, the design of microbial platforms to efficiently convert these building blocks to valuable products involves either manipulating metabolic pathways of native producers, or reconstructing synthetic pathways in organisms that are more industrially and genetically tractable but do not possess the metabolic trait sought for.

This research work improves upon existing technologies in that *i.* It demonstrates a novel microbial platform that efficiently utilize fatty acid-rich feedstocks to synthesize fuels and chemicals; *ii.* It establishes a new approach to construct efficient microbial cell factories by harnessing *E. coli* innate metabolic potential; and *iii.* It demonstrates a novel platform to produce biomolecules with different chain lengths and functionalities.

While the carbohydrate portion of edible crops is currently used as the primary feedstock in the biological production of fuels and chemicals, the availability of fatty acid

(FA)-rich feedstocks and recent progress in the development of oil-accumulating organisms have drawn the attention to FAs as an attractive alternative. However, microbial platforms to enable this were nearly absent prior to our work. Moreover, FAs are metabolized only under respiratory conditions, a metabolic mode that does not support the synthesis of fermentation products. In this work we engineered several native and heterologous fermentative pathways to function in *E. coli* under aerobic conditions, thus creating a respiro-fermentative metabolic mode that enables the efficient synthesis of fuels and chemicals from FAs. The yield of ethanol, acetate, and acetone in the engineered strains exceeded those reported in the literature for their production from sugars, and in the case of ethanol and acetate also surpassed the maximum theoretical values that can be achieved from lignocellulosic sugars. Butanol was produced at yields and titers between two- and three-fold higher than those reported for its production from sugars in previously engineered microorganisms. Moreover, this work demonstrates propionate production in *E. coli*, a compound previously thought to be synthesized only by Propionibacteria. Finally, the synthesis of isopropanol and succinate was also demonstrated. The work hereby reported represents the first effort towards engineering microorganisms for the conversion of FAs to the aforementioned products.

While several of the metabolic traits sought after in an ideal biocatalyst can be found in native microorganisms, these are often distinct from the organisms proved effective in industrial processes and the environmental conditions under which these capabilities are natively utilized often differ from the ones used in a bioreactor. Additionally, the same pathways can be controlled by complex regulatory cascades established by evolution. When the synthesis of the product of interest cannot be

efficiently achieved in the native organism, synthetic biology provides the means to overcome this hurdle by importing biosynthetic pathways in industrially tractable bacteria, thus consolidating in one cell any trait of interest. However, incompatibility of the imported pathway with the host metabolism is of significant concern. We propose an alternative strategy, Homologous Metabolic Engineering (HoME), as a means of synthesizing non-native products in industrial organisms without importing foreign pathways. We demonstrated the viability of this new approach by engineering and redirecting *E. coli* functionome toward metabolic modes that are typical of organisms genetically and metabolically dissimilar. Proof of concept was provided for the case of acetone-butanol-ethanol fermentation, peculiar of Clostridial spp., anaerobic dissimilation of glycerol to 1,3-PDO, a metabolic mode characteristic of *Klebsiella* and *Citrobacter* spp. and the yeast homoethanolic fermentation.

Harnessing the metabolic potential of model organisms provides the opportunity to produce a broad array of platform chemicals, thereby creating a consolidated framework for the production of advanced fuels and chemicals. By analogy with the petrochemical industry, a successful bio industry will build upon the platform-chemical approach, where each chemical intermediate is converted to a large number of chemical products. Here we demonstrate that a functional reversal of the fatty acid catabolic pathway (β-oxidation cycle) can be used as a metabolic platform for the synthesis of alcohols and carboxylic acids with various chain lengths and functionalities. The reversal of the β-oxidation cycle was engineered in *Escherichia coli* and used in combination with endogenous dehydrogenases and thioesterases to synthesize n-alcohols, fatty acids and 3-hydroxy-, 3-keto- and trans-Δ^2-carboxylic acids. The superior nature of the engineered

pathway was demonstrated by producing higher-chain linear n-alcohols ($C \geq 4$) and extracellular long-chain fatty acids ($C > 10$) at higher efficiency than previously reported.

Chapter 1

Introduction

Volatility of oil prices along with major concerns about climate change, oil supply security and depleting reserves have sparked renewed interest in the production of fuels and chemicals from renewable resources. Because of its abundance and renewable nature, biomass has the potential to offer diverse supplies of reliable, affordable, and environmentally sound bioproducts. Given the complexity of biomass in terms of chemical composition, a conventional bioprocess for the biosynthesis of fuels and chemicals entails several steps such as the use of different forms of biomass, feedstock deconstruction to obtain biomass constituents (e.g., monosaccharides, fatty acids, etc.) and their conversion to bioproducts (Figure 1-1). A major challenge presented by the establishment of a viable bioindustry consists in creating microbial cell factories that can perform efficient and economical bioconversions from different biomass building blocks to a set of biomolecules with various chain lengths and functionalities.

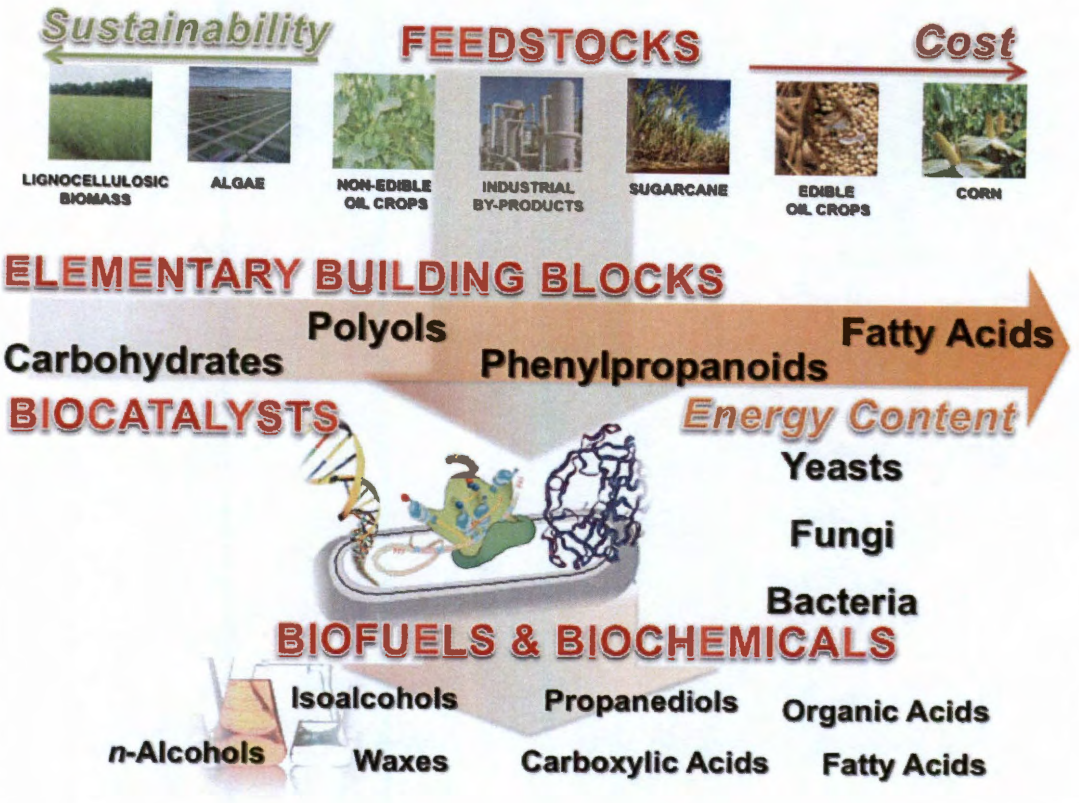

Figure 1-1 General bioprocess scheme for the production of fuels from renewable feedstocks The different feedstocks are listed arranged according to their environmental and economical sustainability. Feedstock deconstruction releases elementary building blocks such as pentoses, hexoses, polyols, fatty acids, etc. that are then microbially converted to biofuels (Dellomonaco et al., 2010a).

Because of its abundance and renewable nature, biomass has the potential to offer diverse supplies of reliable, affordable, and environmentally sound biomolecules. Many biomass feedstocks can be used for the production of renewable fuels and chemicals (Figure 1-1). These include agricultural lignocellulosic residues, edible and non-edible crops, and waste streams (e.g. bagasse from sugar manufacture, industrial by-products) (Figure 1-1).

Vast supplies of diverse renewable resources are therefore available for conversion to generic feedstock constituents (carbohydrates, polyols, fatty acids, etc.) that can be converted into valuable fuels. The availability of diverse generic feedstock building blocks provides many opportunities to employ microbial means for the conversion of a wide variety of carbon sources into valuable products. While a number of platforms that convert carbohydrates and polyols to biomolecules have already been established, microbial processes that can exploit fatty acid-rich feedstocks for the synthesis of fuels and chemicals are nearly absent. *Thus, the need for the establishment of a novel platform that can utilize these abundant renewable feedstocks for fuel and chemical production.*

Many native microorganisms possess biochemical pathways for the dissimilation of feedstock constituents and their conversion into a diverse set of products. However, these native microorganisms, which are usually derived from environmental isolates, often present very stringent requirements in terms of nutrients and environmental conditions, and therefore do not represent a viable option as industrial microbial platforms (Alper and Stephanopoulos, 2009). In addition, while these microorganisms could have the ability to convert biomass constituents to a desired bioproduct, the favorable traits can still be limited, due to complex regulatory cascades arisen as consequence of evolution. Genetic manipulation would be needed in this scenario in order to obtain an integrated and multifunctional microbial platform. However, these organisms typically suffer from a lack of genetic and metabolic tractability, needed to transform them into versatile metabolic platforms.

Metabolic engineering and the most recent synthetic biology are valuable technologies for enabling diversification of molecules biosynthesis by engineering native or *de novo* pathways, and thus combining favorable traits in industrial model organisms (Jarboe et al., 2010). As such, several engineering efforts have revolved around the implementation of these methodologies in the context of fuels and chemicals production (Dellomonaco et al., 2010a; Lee et al., 2008a). While there have been several rather successful attempts of engineering *de novo* pathways in model organisms by importing heterologous genes, the resulting engineered strains often underperform when compared to native organisms. This observation raises therefore the concern as to whether the attractive metabolic capabilities that we aim at importing are in reality controlled enabled by a multitude of genes and physiological factors in the native organisms (Alper and Stephanopoulos, 2009). Hence, compatibility of the imported pathway with the host metabolism becomes of concern. *The need for novel methodologies that can harness native functionome and consolidate favorable traits in model organisms becomes therefore crucial.*

Metabolic engineering and synthetic biology strategies have been to this point tailored for the synthesis and optimization of single products. However, by integrating synthetic competences from diverse metabolic and phylogenetic origins within an integrated biological system, it is possible to envisage envision the development of a single technological platform for synthesizing a wide variety of biorenewable molecules for use in the chemical and energy industry. By analogy with the current non-renewable petrochemical industry, a successful bio-based chemical industry will therefore build upon the platform-chemical approach, where each chemical intermediate is converted to a

large number of chemical products. *Harnessing the metabolic potential of model organisms provides therefore opportunity to produce a broad array of platform chemicals, thereby creating a consolidated framework for the production of advanced fuels and chemicals.*

The overall goal of this study is the development of novel approaches to engineer Escherichia coli for the production of fuels and chemicals. The innovative character of the research hereby presented consists in combining the design of novel metabolic platforms with the development and implementation of innovative metabolic engineering approaches for the efficient conversion of diverse biomass constituents to a set of biomolecules with various chain lengths and functionalities.

The development of oil-accumulating organisms and the abundance of non-edible fatty acid-rich feedstocks have recently ignited interest in the use of fatty acids as attractive alternative to carbohydrates for the production of fuels and chemicals. However, microbial platforms to enable this technology were nearly absent. FAs are catabolized only under respiratory conditions, a metabolic mode that hinders the synthesis of fermentation products. In this work, we established a respiro-fermentative metabolic mode in *E. coli* that enables the synthesis of several fuels and chemicals from fatty acids. Yields of ethanol, acetate, and acetone in the engineered strains surpassed those reported in the literature for their production from sugars, and in the case of ethanol and acetate exceeded the maximum theoretical value attainable from lignocellulosic sugars. The strain engineered for butanol production outperformed in yields and titers microorganisms previously engineered for the synthesis of butanol from sugars. Synthesis of succinate, isopropanol and propionate, a compound previously thought to be

synthesized only by Propionibacteria, was also demonstrated. This work represents the first effort towards engineering microorganisms for the conversion of FAs to fuels and chemicals.

While native bacteria possess a wide set of metabolic capabilities sought after in an ideal microbial platforms, genetical tractability may pose concerns when manipulation to improve on yields is required. An attractive alternative is the utilization of model organisms that are industrially and genetically amenable. However, model organisms may not possess the metabolic trait sought for, as inferred from phylogenetics. Metabolic engineering and synthetic biology have provided the means to overcome this limitation by importing biosynthetic pathways. However, incompatibility of the imported pathway with the host metabolism is a significant concern. The proposed alternative strategy, Homologous Metabolic Engineering (HoME), aims at harnessing the native metabolic potential of industrially amenable microorganisms. The viability of this new approach has been demonstrated by engineering and redirecting *E. coli* functional differentiation toward metabolic modes that are typical of organisms genetically or metabolically distinct. Proof of concept has been provided for the case of acetone-butanol-ethanol fermentation, peculiar of Clostridial spp., anaerobic dissimilation of glycerol to 1,3-PDO, a metabolic mode characteristic of *Klebsiella* and *Citrobacter* spp., and the yeast homoethanolic fermentation.

Harnessing the metabolic potential of model organisms provides the opportunity to create a consolidated framework for the production of a diverse array of chemicals and fuels. In analogy with the current petrochemical industry, a viable bioindustry will build upon the platform-chemical approach, where a wide array of products are synthesized

from a common substrate. In this regard, we propose the functional reversal of *E. coli* fatty acid catabolic pathway (*β*-oxidation) as a metabolic platform for the synthesis of alcohols and carboxylic acids with various chain lengths and functionalities. The reversal of the β-oxidation cycle has been engineered in *Escherichia coli* and used in combination with endogenous dehydrogenases and thioesterases to synthesize n-alcohols, fatty acids and 3-hydroxy-, 3-keto- and trans-Δ^2-carboxylic acids. The superior nature of the engineered pathway is demonstrated by the higher efficiencies attained for higher-chain linear n-alcohols and extracellular long-chain fatty acids biosynthesis.

This work has encompassed the following specific objectives, which are presented below in the context of each chapter.

1. Biological Production of biofuels and biochemicals from fatty acid-rich feedstocks

i. *Establish appropriate respiro-fermentative metabolic mode to enable the production of reduced chemicals and fuels from fatty acids.* Metabolism of free fatty acids (FAs), the main constituents of fats and oils, requires the presence of an external electron acceptor, which in turn could preclude the biosynthesis of fermentative reduced products. Genetic manipulations and adjustment of the availability and/or rate of consumption of terminal electron acceptors have been explored in the context of establishing a respiro-fermentative metabolic mode in which a balance between cell growth and synthesis of reduced products is simultaneously achieved.

ii. *Engineer the constitutive expression of the β-oxidation pathway for the utilization of different length FAs at high flux.* Under aerobic conditions, *E. coli*

can use fatty acids of various chain lengths as sole carbon and energy sources. However, long-chain fatty acids (LCFAs) of at least 12 carbon atoms are required for induction of the catabolic enzymes. The *fad* regulon is primarily responsible for the transport, acylation, and *β*-oxidation of medium- and long-chain fatty acids and the expression of these genes is specifically controlled by *fadR* gene product. In addition to enzymes of the *fad* system, growth of *E. coli* on short chain-fatty acids requires two degradative enzymes encoded by the *ato* operon and regulated by the *atoC* gene product. Genetic manipulation of the local regulator FadR and AtoC has been explored to attain constitutive expression of the β-oxidation pathway for the efficient utilization of FAs of different chain length.

iii. *Engineer native and synthetic pathways in Escherichia coli for the efficient production of fuels and chemicals from fatty acids.* Strategies implemented to this end include 1. Blocking the synthesis of byproducts that compete for both carbon and reducing equivalents; 2. Overexpressing native or heterologous pathways; 3. Manipulating global regulators and environmental conditions to ensure optimal functioning of the pathways.

2. Engineering functional differentiation in bacteria by harnessing the pluripotent capabilities of native hosts

i. *Establish proper sequence, structure and domain-organization homology search criteria and metrics for the identification of functional analogues.* Functional homology is inferred by implementing both phylogenetics and functional inference from secondary and tertiary protein structures. A

comparison of sequence, structure, and domain-organization & composition is performed between the enzymes mediating the heterologous pathways and those encoded in the *E. coli* genome. A set of functional surrogates is obtained after ranking the initial targets by appropriate homology metrics and analyzing them in the context of primary literature and information from genomic and metabolic databases.

ii. *In vivo assembly and functional characterization of native synthetic pathways.* The pathway identified in "*i.*" is assembled and functionally expressed in wild-type *E. coli* by direct overexpression, and fermentative pathways that consume common intermediates are blocked. Since the genes proposed as surrogates might be also subjected to several levels of regulation, expression can be also promoted by manipulating the corresponding regulators.

3. Engineered reversal of the β-oxidation cycle as a metabolic platform for the combinatorial synthesis of advanced fuels and chemicals

   i. *Establish a native metabolic platform for the combinatorial synthesis of fuels and chemicals.* The fatty acid biosynthesis pathway has been recently proposed as a platform for the production of a diverse set of chemicals. However, the use of acyl-ACP intermediates and malonyl-ACP as the 2-C donor during chain elongation in the fatty acid biosynthesis pathway limits its ATP efficiency, making it an energy-consuming pathway. Conversely, the β-oxidation pathway operates with coenzyme-A (CoA) thioester intermediates and directly uses acetyl-CoA for acyl-chain elongation (rather than first requiring ATP-dependent

activation to malonyl-CoA), characteristics that enable product synthesis at maximum carbon and energy efficiency.

ii. *Engineer pathways for the synthesis of carboxylic acids and alcohols from acyl-CoA intermediates generated in the functional reversal of the β-oxidation cycle.* Acyl-CoAs generated by the reversal of the β-oxidation cycle are converted to alcohols by a two-step pathway composed of aldehyde-forming acyl-CoA reductases & alcohol dehydrogenases. Carboxylic acids, on the other hand, are synthesized from acyl-CoAs via specific thioesterases.

iii. *Improve the efficiency of the engineered reversal of the β-oxidation cycle during the synthesis of carboxylic acids and fuels.* Ensuring an adequate supply of acetyl-CoA and reducing equivalents is necessary for the optimal functioning of the engineered reversal of the β-oxidation cycle. Fuels and chemicals biosynthesis is therefore promoted by blocking the synthesis of byproducts that compete for both carbon (AcCoA) and reducing equivalents. Global regulators and environmental conditions are manipulated to ensure optimal functioning of the surrogate enzymes that compose the native pathways.

Chapter 2

Production of biofuels and biochemicals from fatty acid-rich feedstocks

2.1. Motivation and background

Concerns about climate change and depletion and cost of petroleum resources have ignited interest in the establishment of a bio-based industry (Bevan and Franssen, 2006; Ragauskas et al., 2006; Stephanopoulos, 2007) and the conceptual model of a biorefinery has emerged (Kamm et al., 2006; Kamm and Kamm, 2004; Octave and Thomas, 2009).

Because of its abundance and renewable nature, biomass has the potential to offer diverse supplies of reliable, affordable, and environmentally sound biofuels to replace fossil fuels (Dellomonaco et al., 2010a). Many biomass feedstocks can be used for the production of biofuels. These include agricultural lignocellulosic residues, edible and non-edible crops, and waste streams (e.g. bagasse from sugar manufacture, industrial by-

products) (Dellomonaco et al., 2010a). Given their abundance in nature, starch (i.e. corn, wheat, barley, etc.) and sugar crops (i.e. cane, beet, etc.) are the primary feedstocks currently used in the biological production of fuels and chemicals (Dellomonaco et al., 2010a; Ragauskas et al., 2006; Rude and Schirmer, 2009). Although the use of non-edible lignocellulosic sugars has been proposed as an efficient and sustainable avenue to the aforementioned processes, availability of fatty acid (FA)-rich feedstocks and recent progress in the development of oil-accumulating organisms make FAs an attractive alternative.

Edible oil-rich crops such as rapeseed, sunflower, soybean and palm are currently available and widely used as feedstocks for chemical conversion to biodiesel (Carlsson, 2009), while oleaginous algae and non-edible FA-rich crops along with industrial by-products are receiving greater attention as longer-term alternatives. These non-edible FA-rich feedstocks are presently generated in large amounts and can be exploited for the biological production of fuels and chemicals (Durnin et al., 2009; Hu et al., 2008; Rosenberg et al., 2008; Schenk et al., 2008; Service, 2009). However, microbial platforms to enable this were almost absent prior to our work (Dellomonaco et al., 2010b).

2.1.1. Fatty acid degradation in *Escherichia coli*

Escherichia coli can utilize fatty acids of diverse chain lengths as sole carbon and energy source (Clark and Cronan, 2005), under respiratory conditions. Once activated (Figure 2-1), fatty acids can either be catabolized via the β-oxidation pathway or they serve as precursors for membrane phospholipid biosynthesis. The respiratory catabolic

pathway for fatty acids in *E. coli* is mediated by enzymes encoded within the *fad* operon, which are responsible for transport and activation of long-chain fatty acids, and their degradation into acetyl-CoAs (Figure 2-1). Transport of long chain and short chain fatty acids in *E. coli* are governed by distinct mechanisms. Long chain fatty acids are transported in the cytoplasm by an outer membrane protein, FadL (Black and Dirusso, 1994), and activated to coenzyme A thioesters by an acyl-CoA synthase, FadD (Black et al., 1992) (Figure 2-1). While eukaryotic systems possess multiple acyl-CoA synthases with different fatty acid chain-length specificities, *Escherichia coli* has a single acyl-CoA synthase, FadD, characterized by broader substrate specificity (Overath et al., 1969). This reaction consumes one ATP per molecule of free fatty acid activated. At each cycle of the β-oxidation enzymes encoded within the *fad* operon progressively shorten the fatty acid chain by two carbons while capturing the energy in the form of NADH and $FADH_2$. At the end of each cycle of four reactions, one acetyl-CoA two-carbon unit is released from the end of the fatty acid molecule, which then goes through another round of β-oxidation, continuing to oxidize and shorten even-chain fatty acids until their complete degradation to acetyl-CoA (Figure 2-1). Fatty acids with an odd number of carbons in the acyl chain are reduced to propionyl-CoA, which is then converted to succinyl-CoA that then enters the Krebs cycle.

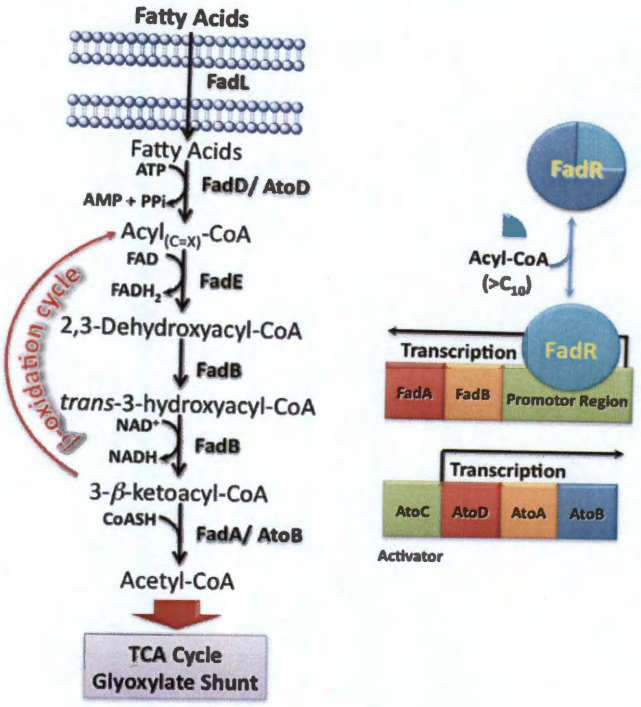

Figure 2-1 Biochemical pathway and regulation of fatty acids catabolism in *Escherichia coli*

FA degradation and biosynthesis pathways are regulated both locally and globally based on availability of fatty acids to sustain membrane lipid homeostasis and in order to quickly respond to environmental changes. Expression of *fad* genes is repressed by the global cyclic AMP-dependent catabolite repression system, as several other sources are preferred *E. coli* substrates over fatty acids. The second specific regulatory system is instead locally exerted by the FadR trascriptional factor, which represses the transcription of each of the genes essential for fatty acid transport, activation, and β-oxidation, (Campbell et al., 2003; Clark and Cronan, 2005; Nunn, 1986a) while activating transcription of genes encoding key enzymes in the biosynthesis of unsaturated fatty acids (Campbell and Cronan, 2001).

Growth of wild-type *E. coli* strains in the presence of long-chain ($\geq C_{12}$) fatty acids coordinately induces the fatty acid degradative (*fad*) enzymes (Klein et al., 1971)(Figure 2-1). On the contrary, fatty acids of medium (C_7 to C_{11}) or short (C_4 to C_6) chain lengths cannot induce synthesis of the *fad* enzymes in *Escherichia coli* (Black and Dirusso, 1994). Although wild-type *E. coli* strains can only utilize long-chain fatty acids as a sole carbon and energy source, MCFAs can serve as growth substrates for *fadR* mutants that constitutively express the *fad* enzymes (Clark and Cronan, 2005). Growth of *E. coli* on short-chain fatty acids requires instead expression of the *atoDAB* operon (Pauli and Overath, 1972). The *atoB* gene encodes β-ketoacyl-CoA thiolase (thiolase II), responsible for the thiolytic cleavage of short-chain fatty acyl-CoA and subsequent synthesis of acetyl-CoA. The *atoC* gene product positively regulates transcription of the aforementioned three structural genes of the *ato* operon (Jenkins and Nunn, 1987a)(Figure 2-1). Campbell and coworkers (Campbell et al., 2003) recently tested the ability of *E. coli* K12 to grow on fatty acids under conditions of anaerobic respiration. An anaerobic non-fermentative pathway for FA catabolism was identified in *E. coli* (Campbell et al., 2003), distinct from the aerobic one because it functions with fatty acids ($C_{8:0}$ and $C_{10:0}$) that cannot be utilized by wild-type *Escherichia coli* strains under aerobic conditions, and it does not seem to be under the regulation of FadR.

2.1.2. Feasibility analysis of the utilization of fatty acid-rich feedstocks

2.1.2.1. Currently available fatty acid-based feedstocks

Currently, edible oil-rich feedstocks are available and widely used as feedstocks for chemical conversion to biodiesel (Carlsson, 2009). Most seed oils are mainly composed of a wide range of FAs, the most abundant being palmitic ($C_{16:0}$) and oleic acid (C18:1) (Azocar et al., 2010; Gui et al., 2008). While an upward trend of global vegetable oils supply has been reported over the last 10 years, palm oil is currently the world largest edible oil produced (45.86 million metric tons per year; USDA Report February 2011, http://www.fas.usda.gov), due to the highest yield per hectare (Gui et al., 2008). Palm oil processing and refining generates as by-product palm fatty acid distillate (PFAD), a side stream with free FAs content up to 93% (palmitic acid being by far the most abundant fatty acid at 58.1%). Production of PFAD accounts for ~5% of palm oil total annual production and could already support synthesis of biochemicals such to replace close to 10% of the petrochemicals produced in US. Other side-streams of the meat and vegetable oil industry that contain mixtures of free FAs are also available. POME (palm oil mill effluent) is a high volume liquid waste of palm oil wet mill processing characterized by high concentrations of lipids, carbohydrates, proteins, nitrogenous compounds, and minerals (Hwang et al.). At present, about 2.5 t of POME are produced for every ton of oil extracted in an oil mill (Ahmad et al., 2005; Wu et al., 2009). Additionally, by-products of meat rendering, such as inedible animal tallow and yellow grease, are a viable alternative. They consist of a mixture of triglycerides, their major fatty acyl residues being oleic (45%), stearic (25%), and palmitic (25%). Inedible tallow is one of the least expensive animal fats (19 US cents/pound; www.thejacobsen.com, June 2010) and its

production in the United States (1.5 million metric tons in 2010; Sources: U.S. Census Bureau, M311K series for Fat and Oils: Production, Consumption, and Stocks.) greatly exceeds domestic consumption. In summary, the production of biochemicals from FA-rich industrial byproducts can become a realistic alternative in the foreseeable future.

2.1.2.2. Prospects for large-scale availability of FA-rich feedstocks

Lately oleaginous algae and non-edible crops are receiving greater attention as longer-term alternatives to edible oil crops. Algae accumulate oils mainly in the form of triglycerides ($\geq$ 80%), with a FAs profile rich in C_{16} and C_{18}, similar to plant seed oils (Meng et al., 2009). Algae grow rapidly compared to crops and many are exceedingly rich in oil, with the average lipid content up to 80% by weight of dry biomass (Rodolfi et al., 2009). Today, approximately 5,000-10,000 tons of algae are commercially produced primarily for the production of high-value, low-volume food supplements and nutraceuticals (Pienkos and Darzins, 2009). A techno-economic model developed at the National Renewable Energy Laboratory (NREL; www.nrel.gov), and based on open-pond cultivation of algae, has estimated the cost of production for algal lipids ranging from \$25/gal for low productivity (10g m^{-2} day^{-1} at 15% triglycerides) to \$2.50/gal for high productivity (50g m^{-2} day^{-1} at 50% triglycerides) (Pienkos and Darzins, 2009). Although at the present stage of technical development the cost-effective high-scale production of algal oils is not viable, recent progress in the optimization of these oil-accumulating organisms will make FAs available at the scale necessary to support the production of biofuels (Hu et al., 2008; Rosenberg et al., 2008; Schenk et al., 2008; Service, 2009). Non-edible oil crops like *Jatropha curcas L.*, *Ricinus communis*, *Pongamia pinnata*, *Hevea brasiliensis*, *Cerbera odollam* or *Cerbera manghas* are characterized by average

oil content of dry seeds up to 50%, corresponding to a productivity of 225-1590 kg oil/ha (Gui et al., 2008). The oil composition of non-edible oil crops is very similar to the more common edible vegetable oils, rich in palmitic, oleic, linoleic and stearic acid (Gui et al., 2008). Although non-edible oil crops are still at an early stage of development, they represent very promising platforms for the production of oils.

2.1.2.3. Product yield and feedstock cost: comparison of sugar- and fatty acid-based feedstocks

Availability of FA-rich feedstocks and recent progress in the development of oil-accumulating organisms (see Section 2.1.2.1) will undoubtedly make FAs available at the scale necessary to support the production of biofuels and are currently available to support significant biochemical production.

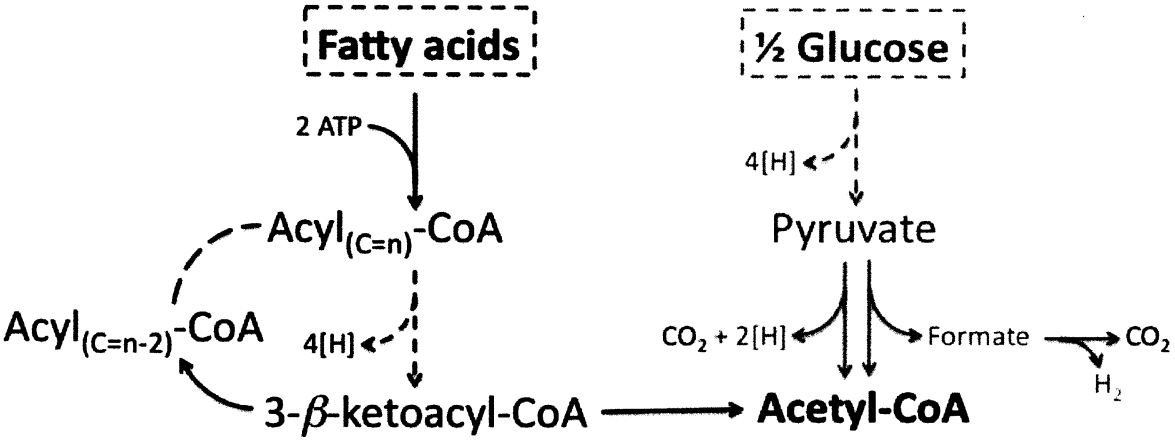

Figure 2-2 Comparison of catabolism of fatty acids and glucose to acetyl-CoA in Escherichia coli.

β-oxidation of fatty acids is represented on the left. Two ATP are released upon activation of every molecule of fatty acids. At each turn of the β-oxidation cycle the fatty acid chain is progressively shortened of 2C, with the release of one molecule of acetyl-CoA. The dashed line (without arrowheads) between the two Acyl-CoA names is meant to imply that the two names are just different instantiations of the same compound. Abbreviations: n, fatty acid chain length; Abbreviations: 2[H] = NADH = $FADH_2$ = H_2.

FAs are not only abundant but also offer several advantages when used for fuel and chemical production. For example, their metabolism to the key intermediate metabolite acetyl-CoA is very efficient as it results in 100% carbon recovery (Figure 2-2). Since many fuels and chemicals can be derived from acetyl-CoA, high yields can be realized if FAs are used as the carbon source (Figure 2-2). In contrast, sugar metabolism generates one molecule of carbon dioxide (or formate) per molecule of acetyl-CoA, limiting the yield of products derived from acetyl-CoA (Figure 2-2 and Table 2-2). The product yield advantage of FAs over sugars is also supported by the higher reduced nature of their carbon atoms. Table 2-2 provides a comparison of maximum theoretical yields, on both weight and carbon basis, for the production of biofuels and biochemicals from FAs and lignocellulosic sugars. Maximum theoretical yields have been calculated from stoichiometry based on the pathways shown in Figure 2-4 for the utilization of FAs and glucose, the synthesis of products, the TCA cycle, and the oxidative phosphorylation. The stoichiometric coefficients were obtained by conducting elemental balances on carbon, hydrogen and oxygen, and an ATP balance was included in the analysis.

As an example, when production of biofuels (e.g. ethanol and butanol) is considered, utilization of FAs (e.g. palmitic acid/$C_{16:0}$) as substrate returns product yields 2.7-fold (w/w), or 1.4-fold (C/C), higher than sugars (calculations provided for glucose but equally valid for other lignocellulosic sugars). Although the current price of feedstocks on a weight basis is comparable (lower than 20 USD cents/pound), data reported in Figure 2-3 shows that the price per carbon for glucose derived from corn is remarkably higher. Regardless of the basis used for calculations (i.e. weight or carbon

basis), when maximum theoretical yields and costs of FA and sugar feedstocks are accounted for, the advantages of using FAs are remarkable (Figure 2-3a).

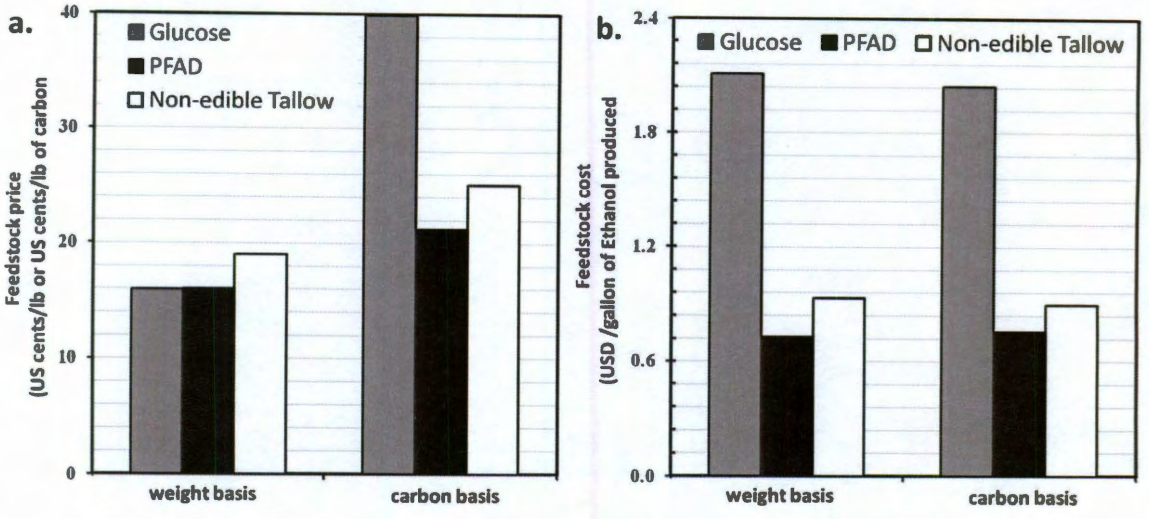

Figure 2-3 Price and cost comparison for the use of sugar- and fatty acid-based feedstocks in the production of biofuels and biochemicals
(a) Prices of raw sugar, palm fatty acid distillate (PFAD) and non-edible tallow are reported in US cents per pound and on a per-pound of carbon basis. Prices for complex mixtures of fatty acids (PFAD, non-edible tallow) have been calculated by weight averaging carbon content of the constituents according to reported composition. Source: The Jacobsen, www.thejacobsen.com; United States Department of Agriculture, Economic Research Service, www.ers.usda.gov.
(b) Feedstock cost per gallon of ethanol produced is reported in USD. Data shown have been calculated on weight and carbon basis using maximum theoretical yields shown in **Table 2-2** along with feedstock prices from panel (a).

2.1.3. Fuels and chemicals currently produced by microbial conversion of oils/fats

Despite the numerous aforementioned advantages, utilization of FA-rich feedstocks has only been explored in the context of chemical conversions, such as synthesis of surfactants or transesterification to FAMEs for use in biodiesel production (Biermann et al., 2011). The microbial conversion of FA-rich feedstocks has instead been

exploited only for the production of polyhydroxyalkanoates (Liu et al., 2011; Park et al., 2005; Park et al., 2002), with no report to date of organisms engineered for the conversion of FAs to fuels.

Polyhydroxyalkanoates (PHAs) form a class of natural polyesters that many organisms in the environment accumulate in the form of intracellular granules to store carbon and reducing equivalents. Metabolic pathways involved in the synthesis and degradation of fatty acids generate intermediates activated by acyl carrier protein (ACP) and coenzyme A (CoA), respectively. Such metabolites in *de novo* fatty acid biosynthesis and fatty acid β-oxidation pathways are possible precursors for PHA biosynthesis as they can be easily converted into (R)-3-hydroxyacyl-CoAs (R3HA-CoAs), the most favorable substrates for PHA synthase, the key enzyme responsible for polymerization of activated (R)-hydroxyacyl-monomers in PHA biosynthesis. However, when production of PHA from intermediates of the β-oxidation pathway is considered, an intrinsic limitation resides in having to use inhibited metabolic pathways. This limitation has been overcome by manipulating strains to make the β-oxidation pathway functional. The β-oxidation pathway has been further engineered to support PHA biosynthesis in two ways, one involving inhibition of key β-oxidation pathway enzymes, such as FadA and FadB (Langenbach et al., 1997; Liu et al., 2011; Vo et al., 2007), and the other involving amplification of enzymes that directly convert β-oxidation pathway intermediates to R3HA-CoAs, such as enoyl-CoAhydratase and 3-ketoacyl-ACP reductase (Li et al., 2007b; Li et al., 2010; Vo et al., 2008).

2.2. Methods

2.2.1. Strains, plasmids and genetic methods

Strains, plasmids, and primers used in this study are listed in Table 2-1. Wild-type K12 *E. coli* strain MG1655 (Kang et al., 2004b) was used as the host to implement metabolic engineering strategies. Gene knockouts were introduced in MG1655 and its derivatives by P1 phage transduction as described elsewhere (Miller, 1992), using as donors single gene knock out mutants from the National BioResource Project (NIG, Japan) (Baba et al., 2006). Details of the specific protocol used have been described before (Yazdani and Gonzalez, 2008). Strain *fadR atoC*(c), which exhibits constitutive expression of the *fad* regulon (due to *fadR*) and *ato* operon (due to *atoC*(c)), was obtained as follows. Strain MG1655 *fadR* was isolated as an spontaneous mutant able to grow on MOPS (morpholino-propanesulfonic acid) minimal medium (Neidhart et al., 1974) plates containing 0.2% (w/v) decanoic acid ($C_{10:0}$) as the sole carbon source. Strain *fadR atoC*(c) was then obtained by transducing *fadR* with a phage lysate from strain LS5218, which is a FadR AtoC(c) mutant (Jenkins and Nunn, 1987b; Spratt et al., 1981b). Successful transduction of the *atoC*(c) mutation into strain *fadR* was identified by growth on MOPS minimal medium plates containing 0.2% (w/v) hexanoic acid ($C_{6:0}$).

Standard recombinant DNA procedures were used for gene cloning, plasmid isolation, and electroporation. Manufacturer protocols and standard methods (Miller, 1992; Sambrook and Russell, 2001) were followed for DNA purification (Qiagen, CA, USA), restriction endonuclease digestion (New England Biolabs, MA, USA), and DNA amplification (Stratagene, CA, USA and Invitrogen, CA, USA). The strains were kept in

32.5% glycerol stocks at -80°C. Plates were prepared using Luria-Bertani (LB) medium containing 1.5% agar and appropriate antibiotics were included at the following concentrations: 100 μg/mL ampicillin, 50 μg/mL kanamycin, 34 μg/mL chloramphenicol and 12.5 μg/mL tetracycline. All strains created were confirmed by polymerase chain reaction using verification primers listed in Table 2-1 and appropriate phenotypic tests, if suitable.

The overexpression of an aerotolerant mutant of *E. coli* alcohol/acetaldehyde dehydrogenase (i.e. AdhE*) was facilitated by the construction of expression vector pTHadhE*. The *adhE* coding region was PCR amplified using genomic DNA of *E. coli* MG1655 as template and c-*adhE* primers (Table 2-1). The resulting PCR product was cloned into vector pUC19 (New England BioLabs). A QuikChange site-directed mutagenesis kit from Stratagene Inc. (La Jolla, CA) was then used to generate a Glu568Lys (E568K) mutation in the *adhE* coding sequence using primers m-*adhE* (Table 2-1). The E568K *adhE* mutant (*adhE**), encoding an aerobically active AdhE protein, was then subcloned into the vector pTrc His 2A from Invitrogen Corp. (Carlsbad, CA) using c-*adhE* primers (Table 2-1).

The butanol pathway from *C. acetobutylicum* ATCC824 was re-constructed in *E. coli* using plasmids pZS.crt.bcd.etfAB.hbd and pTH.atoB.adhE2. pZS.crt.bcd.etfAB.hbd was constructed by first amplifying the *crt-bcd-etfAB-hbd* operon from *C. acetobutylicum* ATCC824 genomic DNA with primers c-*crt.bcd.etfAB.hbd* (Table 2-1) and then ligating the resulting PCR product into the *Kpn*I and *Mlu*I sites of pZSKLM (Yazdani and Gonzalez, 2008). The pTrc.atoB.adhE2 plasmid was constructed by using overlap PCR to fuse the PCR amplified *E. coli atoB* gene with the PCR amplified *C. acetobutylicum*

adhE2 gene. The fused *atoB-adhE2* genes were then cloned into pTrc-His2A (Invitrogen, Carlsbad, CA) using the In-Fusion PCR cloning system from Clontech Laboratories, Inc. (Mountain View, CA). Details about primers used along with promoters and antibiotic markers are provided in Table 2-1.

The phosphotransacetylase (PTA)-acetate kinase (ACK) pathway for the synthesis of acetate from acetyl-CoA (Clark and Cronan, 2005) (Figure 2-4) was amplified through overexpression of the *ackApta* operon (pZS.ackA.pta). The *ackApta* operon was PCR amplified using genomic DNA from strain MG1655 and primers c-*ackA.pta* (Table 2-1). The resulting PCR product was cloned into pZSKLM (Yazdani and Gonzalez, 2008) via In-Fusion PCR cloning (Clontech Laboratories, Inc., Mountain View, CA).

The production of acetone was evaluated using plasmid pTrc.atoDAB.adc, which was constructed as follow. The *E. coli atoDA* genes were amplified using primers c-*atoDA* (Table 2-1) and the PCR product directionally cloned into the *BamH*I and *Kpn*I restriction sites of pTrc His 2A (Invitrogen Corp., Carlsbad, CA). The resulting plasmid was named pTrc.atoDA. The *C. acetobutylicum* ATCC824 *adc* gene was then cloned into pTrc.atoDA within the *Kpn*I and *Eco*RI restriction sites (c-*adc*, Table 2-1), yielding pTrc.atoDA.adc. Lastly, the *E. coli atoB* gene was amplified (c-*atoB*, Table 2-1) and subsequently cloned within the *Kpn*I restriction sites of pTrc.atoDA.adc, yielding the final vector pTrc.atoDAB.adc. Multiple colonies were screened by restriction analysis to assess the directionality of the *atoB* insertion.

To facilitate the conversion of acetone to isopropanol, the *sadh* gene from *C. acetobutylicum* was custom synthesized by GenScript Corp. (Piscataway, NJ), PCR

amplified with primers c-*sadh* (Table 2-1), and cloned into pZSKLM (Yazdani and Gonzalez, 2008) using the In-Fusion PCR cloning system from Clontech Laboratories, Inc. (Mountain View, CA). The same approach was used to clone *sadh* into pZS.ackA.pta, thus generating pZS.ackA.pta.sadh (Table 2-1).

The production of propionate from FAs was tested by cloning the *E. coli* genes *scpA*, *scpB*, and *scpC* into pTrc99a (Invitrogen, Carlsbad, CA). To this effect, the *scpAargKscpBscpC* operon from *E. coli* MG1655 was PCR amplified using primers c-*scpAargKscpBscpC* (Table 2-1) and cloned into pTrc99a, both digested with *Eco*RI and *Hind*III. PCR was used to amplify this construct with the exception of *argK* (primers c-*scpABC*, Table 2-1), adding a flanking *Xba*I site downstream of *scpA* and upstream of *scpB*. *Xba*I digestion and religation of this PCR product and cloning into pTrc-His2A rendered pTrc.scpABC.

2.2.2. Culture medium and cultivation conditions

MOPS minimal medium (Neidhart et al., 1974) supplemented with 0.5% palmitic acid (w/v) and 0.2% (w/v) Brij58 (Fluka Chemie AG, Buchs, Switzerland) was used. The oxygen transfer rate (k_La) was estimated from previous reports on gas-liquid mass transfer in shake-flask systems (Maier et al., 2004; Van Suijdam et al., 1978) and further confirmed by in-vessel measurements conducted by conventional dynamic gassing-out technique (Durnin et al., 2009).

Unless otherwise stated, all chemicals for culture media were obtained from Fisher Scientific (Pittsburgh, PA) and Sigma-Aldrich Co. (St. Louis, MO). MOPS minimum medium, as designed by Neidhart, supplemented with 0.5% palmitic acid (w/v)

and 0.2% (w/v) Brij58 (Fluka Chemie AG, Buchs, Switzerland) was used, unless otherwise stated. When required the medium was supplemented with 100 μg/mL ampicillin, 50 μg/mL kanamycin, 34 μg/mL chloramphenicol and 12.5 μg/mL tetracycline. Isopropyl-beta-D-thiogalactopyranoside (IPTG, 0.1 mM) and anhydrotetracycline (100 ng/ml) were used to induce gene expression from constructed plasmids.

Table 2-1 Strains, plasmids, and primers used in this study

Strain/plasmid Primer	Description/Genotype/ Sequence	Source
Strains		
MG1655	F- l- ilvG- rfb-50 rph-1	(Kang et al., 2004b)
Δ*adhE*	MG1655, Δ*adhE*::FRT-kan-FRT; deletion mutant for *adhE* in MG1655	This study
Δ*sdhB*	MG1655, Δ*sdhB*::FRT-kan-FRT; deletion mutant for *sdhB* in MG1655	This study
LS5218	*fadR*601 *atoC(*Con)2	(Spratt et al., 1981a)
fadR	MG1655 evolved for rapid growth onto agar plates with decanoic acid	This study
fadR atoC(c)	*fadR* transduced with the *atoC*(C) gene of LS5218	This study
fadRatoC(c)Δ*adhE*	*fadR atoC*(c), Δ*adhE*::FRT-kan-FRT; deletion mutant for *adhE* in *fadR atoC*(c)	This study
fadR atoC(c)Δ*sucA*	*fadRatoC*(c), Δ*sucA*::FRT-kan-FRT; deletion mutant for *sucA* in *fadR atoC*(c)	This study
fadR atoC(c)Δ*sdhB*	*fadRatoC*(c), Δ*sdhB*::FRT-kan-FRT; deletion mutant for *sdhB* in *fadR atoC*(c)	This study
fadR atoC(c)Δ*sucD*	*fadRatoC*(c), Δ*sucD*::FRT-kan-FRT; deletion mutant for *suc*D in *fadR atoC*(c)	This study
Plasmids		
pTH.adhE*	*E. coli adhE* mutant (E568K) under control of Ptrc (AmpR, oriR pBR322)	This study
pZS.crt.bcd.etfAB.hbd	*C. acetobutylicum* butyryl-CoA synthesis operon (*crt, bcd, etfAB, hbd*) under control of PLtetO-1 (tetR, oriR SC101*, cat)	This study
pTH.atoB.adhE2	*E. coli atoB* gene and *C. acetobutylicum adhE2* gene under control of Ptrc (AmpR, oriR pBR322)	This study
pTH.atoDAB.adc	*E. coli atoD*, *atoA* and *atoB* genes and *C. acetobutylicum adc* gene under control of Ptrc (AmpR, oriR pBR322)	This study
pZS.sadh	*C. acetobutylicum sadh* gene under control of PLtetO-1 (tetR, oriR SC101*, cat)	This study
pZS.ackA.pta	*E. coli ackApta* operon under control of PLtetO-1 (tetR, oriR SC101*, cat)	This study
pZS.ackA.pta.sadh	*E. coli ackApta* operon and *C. acetobutylicum sadh* gene under control of PLtetO-1 (tetR, oriR SC101*, cat)	This study
pTrc.scpABC	*E. coli scpA, scpB and scpC* genes under control of PLtetO-1 (AmpR, oriR pBR322)	This study
Primers [a]		
v-*adhE*	gtttaacattatcaggag; gtcaactaatccttaac	This study
v-*sucA*	cacatcactgtgcgtggtagtatcc; caggtcagggaccagaatatctacg	This study
v-*sdhB*	cttccgtaccgaaagccgtg; accacgcacagtgatgtgcg	This study
v-*sucD*	gacagcggcctgaatattattgcag; catcgcgataagcacaaaaaaggcc	This study
m-*adhE*	catccggaaactcacttcgaaaagctggcgctg; cagcgccagcttttcgaagtgagtttccgga	This study
c-*adhE*	cattaaagaggagaaaggtaccatggctgttactaatg; gatgcctctagcacgcgtttaagcggatttttttcg	This study
c-*ackA.pta*	cattaaagaggagaaaggtaccatgtcgagtaagttag; gatgcctctagcacgcgtttactgctgctgtgc	This study
c-*ackA.pta.sadh*	ctgcacagcagcagtaaacgcgtgaggaatgaaaggctttgcg; gatgcctctagcacgcgtttacagaatcaccaccgc	This study
c-*sadh*	cattaaagaggagaaaggtaccatgaaaggctttgcgatgctg; gatgcctctagcacgcgtttacagaatcaccaccgc	This study
c-*crt.bcd.etfAB.hbd*	gatggtaccatggaactaaacaatgtcatccttg; gatcacgcgtttattttgaataatcgtagaaacc	This study
c-*atoB*	gagatctgcagctggtaccatgaaaaattgtgtcatc; ctttttgatttgtaactttcatttaattcaaccgttcaatcacc	This study
c-*adhE2*	ggtgattgaacggttgaattaaatgaaagttacaaatcaaaaag; cgggcccaagcttcgaattcttaaaatgattttatatagatatcc	This study
c-*atoB.adhE2*	gagatctgcagctggtaccatgaaaaattgtgtcatc; cgggcccaagcttcgaattcttaaaatgattttatatagatatcc	This study
c-*scpABC*	gctctagaatgtcttatcagtatgttaagg; gctctagattaatcatgatgctggc	This study
c-*scpAargKscpBC*	cggaattcatgtctaacgtgcaggagtgg; gacaagcttttaacccagcatcgagccg	This study

c-*atoDA*	gggatccatgaaaacaaaattgatgac; gaggtacctcataaatcaccccgttg	This study
c-*adc*	cggtaccatgttaaaggatgaag; cggaattcttaagataatcatatataac	This study
c-*atoB*	acggtaccaggaggaaatgaaaaattgtgtcatcgtcagtgc; ccttcctcctttaattcaaccgttcaatcaccatcgc	This study

[a] "v" indicates the primer sequences (5' to 3') that were used for verification purposes during the creation of disruption mutants by phage transduction. "m" indicates primers used for site-directed mutagenesis of *adhE*. "c" indicates primers used for cloning purposes. The forward follows the reverse sequence in each case, separated by a semicolon. Genes or operons deleted or cloned are apparent from primer names.

Prior to use, cultures (stored as glycerol stocks at -80 °C) were streaked onto LB plates (with appropriate antibiotics if required) and incubated overnight at 37 °C. A single colony was used to inoculate 10 mL of LB broth in 15 ml test tubes (GeneMate, Kaysville, UT), which were incubated at 37 °C until an OD_{550} of ~0.5 was reached. An appropriate volume of this actively growing pre-culture was centrifuged and the pellet washed and used to inoculate 25 mL of medium in 50 ml shake flasks to an initial OD_{550} of 0.05. Fifty-milliliter shake flasks (Corning Glass Works, Corning, NY) with four baffles and plastic foam-stoppers were used for aerobic cultures. Flasks were incubated for 72 hours at 37°C in a C24 Rotary Incubator Shaker (New Brunswick Scientific, NJ). Samples were centrifuged to pellet cells while the aqueous supernatant was collected for metabolite analysis.

For ethanol and butanol production by high cell density cultures, 10 ml LB liquid medium in a test tube was inoculated with *E. coli* strains and incubated at 37°C in a rotator (Glas-Col Inc., Terre Haute, IN) until an OD_{550} of 0.7 was reached. This 10 ml culture was used to inoculate 50 ml of LB medium in a 125-ml conical flask (Corning Glass Works, Corning, NY) and incubated at 37°C in a C24 Rotary Incubator Shaker until an OD_{550} of 0.7 was reached. Cells were collected by centrifugation at 8,000×*g*, 4°C

for 20 min, and washed twice with MOPS minimal medium. The collected cells were resuspended in MOPS minimal medium to a final OD_{550} of 10. The cell suspension (25 ml) was transferred to a 50-ml baffled shake flask, and palmitic acid ($C_{16:0}$) was added to a final concentration of 0.5% (w/v). Additional palmitic acid (0.5% w/v) was added to the medium at indicated time intervals.

For acetate production by high cell density cultures, 12.5 ml LB liquid medium in a test tube was inoculated with *E. coli* strain and incubated at 37°C in a rotator (Glas-Col Inc., Terre Haute, IN) until an OD_{550} of 0.8 was reached. This 12.5 ml culture was used to inoculate 250 ml of LB medium in a 500-ml conical flask (Corning Glass Works, Corning, NY) and incubated at 37°C in a C24 Rotary Incubator Shaker until an OD_{550} of 0.8 was reached. Cells were collected by centrifugation at 8,000×*g*, 4°C for 20 min, and washed twice with MOPS minimal medium. The collected cells were resuspended in MOPS minimal medium to a final OD_{550} of 10. The cell suspension (50 ml) was transferred to a 500-ml fermenter in a SixFors multi-fermentation system (Infors HT, Bottmingen, Switzerland) with independent control of temperature (37 °C), pH (controlled at 7.0 with NAOH 10M), stirrer speed (500 rpm), and dissolved oxygen (not controlled) (64). Palmitic acid ($C_{16:0}$) was added to a final concentration of 0.5% (w/v). Additional palmitic acid (0.5% w/v) was added to the medium at indicated time intervals. Microaerobic conditions were maintained by sparging the medium with air at 0.05 L/min.

2.2.3. Enzyme activities

The activity of alcohol dehydrogenase (AdhE) was measured following the procedure reported by Kessler and coworkers (Kessler et al., 1991). *E. coli* cells from

anaerobic or aerobic cultures (OD_{550} ~ 0.7) were harvested by centrifugation (2 min, 10,000 x g), washed twice with a solution of NaCl (9 g/l) and stored as cell pellets at -20°C. For assays, cell pellets were resuspended in 0.2 ml of 0.1M MOPS-KOH and permeabilized by vortex mixing with chloroform. Alcohol dehydrogenase activity was assayed by measuring the change in absorbance at 340 nm and 30°C in a 1 ml reaction mixture containing 0.1 M MOPS-KOH buffer (pH 7.5), 6 mM DTT, 5 mM $MgSO_4$, 0.3 mM $Fe(NH_4)_2(SO_4)_2$, 0.4 mM NADH, 10 mM acetaldehyde, and 30 µl crude cell extract (Kessler et al., 1991). For the anaerobic assay, the above sample preparation and assay were carried out in a BACTRON I Anaerobic Chamber (Sheldon Manufacturing Inc., Cornelius, OR). Linearity of the reactions (protein concentration and time) was established for all preparations. All spectrophotometric measurements were conducted in a BioMate 5 spectrophotometer (Thermo Scientific, MA, USA). The nonenzymatic rates were subtracted from the observed initial reaction rates. Enzyme activities are reported as mmoles of substrate/minute/mg of cell protein and represent averages for at least three cell preparations.

2.2.4. Analytical methods

Cell growth was monitored by measuring total protein concentration using the method of Lowry (Lowry et al., 1951) and a predetermined correlation between cell dry weight (CDW) and total protein concentration. The CDW was determined by collecting 5 ml of a cell suspension on a membrane filter (diameter 47 mm; pore size 0.45 mm) (Sartorius, Gottingen, Germany) and applying suction. The filter was washed twice with distilled water, dried to constant weight, and desiccated before weighting.

The identity of all metabolic products was determined through 1D proton NMR spectroscopy as previously described (Murarka et al., 2008). Organic acids, ethanol, and isopropanol were quantified by HPLC as previously reported (Dharmadi and Gonzalez, 2005).

Butanol was quantified in a Varian CP-3800 gas chromatograph (Varian Associates, Inc., Palo Alto, CA) equipped with a flame ionization detector (FID) following a modification of the method reported by Atsumi and coworkers (Atsumi et al., 2008a). The separation of alcohol compounds was carried out using a VF-5ht column (15 m, 0.32 mm internal diameter, 0.10-μm film thickness; Varian Associates, Inc., Palo Alto, CA). The oven temperature was initially held at 40 °C for 2 min and raised with a gradient of 5 °C/min until 45 °C and held for 4 min. The temperature was then raised with a gradient of 15 °C/min until 230 °C and held for 4 min. Helium (Matheson Tri-Gas, Longmont, CO) was used as the carrier gas with 14 p.s.i inlet pressure. The injector and detector were maintained at 225 °C. A 0.5-μl sample was injected in splitless injection mode.

The analysis of FAs was carried out in a Varian CP-3800 gas chromatograph (Varian Associates, Inc., Palo Alto, CA) after hexane/ methyl tertiary butyl ether (MTBE) extraction and FAs transesterification with methanol (Lewis et al., 2000), according to the following method: 50°C held for 1 min, 30°C/min to 160°C, 15°C/min to 200°C, 200°C held for 1.5 min, 10°C/min to 225°C, 225°C held for 15 min.

Sample preparation (i.e. MTBE extraction and transesterification) was conducted as follows. One ml samples were transferred to 5 ml serum bottles (Supelco, Bellefonte,

PA) containing 2 ml of hexane/ MTBE (1:1). Eighty μl of 50% H_2SO_4 and 0.05 g NaCl were added for pH and ionic strength adjustment, respectively (Lalman and Bagley, 2004). Triplicate calibration standards for FAs analysis of 1, 2, 5, 10, 20, 30, 40, 50, and 100 mg/l were prepared in 1:1 mixture of chlorophorm/methanol using a 1,500 mg/l stock solution of each FA. The bottles were sealed with Teflon®-lined septa (Fisher Scientific Co., Fair Lawn, NJ), secured with caps, and shaken using an orbital shaker at 200 rpm for 15 min. The samples were then centrifuged for 5 min at 1,750 x *g* to separate the aqueous and organic layers. After centrifugation, the top organic layer was separated from the aqueous layer. A 0.5 ml aliquot of the top layer was transferred carefully using glass pipettes to 1.0 ml Supelco Reacti-Vials (Sigma-Aldrich, St. Louis, MO) and dried under a stream of nitrogen in an N-EVAP evaporator (Organomation Associates, Inc., Berlin, MA).

A fresh solution of the transesterification reaction mix (methanol:hydrochloric acid:chlorophorm (10:1:1 v/v/v, 3ml) was added to the dried lipid extract in Reacti-Vials (Sigma-Aldrich, St. Louis, MO), which were capped tightly and vortexed. These vials were placed in a heater block (AccuBlock Digital Dry Bath, LabNet, Woodbridge, NJ) and heated at 90°C for 15 min.

The transesterification reaction tubes were cooled to room temperature. Water (1ml) was added to each tube and the fatty acid methyl esters (FAMEs) were extracted with hexane and chloroform (4:1 v/v, 3 x 2 ml). The tubes were vortexed for 30 seconds and the upper organic phase collected with a Pasteur pipette. This extraction procedure was repeated 3 times.

The combined hexane:chloroform solution was evaporated under N_2 to dryness, and the dry residue redissolved in 60 μL of hexane, transferred to GC vials (Fisher Scientific Co., Fair Lawn, NJ) and capped under N_2. A 2.0 μl aliquot of FAMEs solution was injected (injection temperature 300°C) into the chromatograph. The carrier gas flow-rate was 2.0 ml/min, with a split ratio of 1:25.

Stearic ($C_{18:0}$), palmitic ($C_{16:0}$), myristic ($C_{14:0}$), lauric ($C_{12:0}$), capric ($C_{10:0}$), caprylic ($C_{8:0}$), and caproic ($C_{6:0}$) acids (Sigma Chemical Co., St. Louis, MO) were used to calibrate the gas chromatograph. The carrier gas used was helium (Matheson Tri-Gas, Longmont, CO). Hexane, chloroform, diethyl ether, and MTBE were HPLC grade (Sigma Chemical Co., St. Louis, MO). Sodium chloride and concentrated sulfuric acid were reagent grade (VWR International, West Chester, PA).

2.2.5. Calculation of fermentation parameters

Growth and product yields (gram of product per gram of substrate consumed) were calculated as the amount of cell mass or product synthesized per amount of palmitic acid ($C_{16:0}$) consumed. Production of caproic acid ($C_{6:0}$) and biomass were accounted for in product yield calculations. An average molecular weight for an *E. coli* cell of 24.7 g/C-mol, which corresponds to an average cell of a molecular formula $CH_{1.9}O_{0.5}N_{0.2}$ (Nielsen et al., 2003), was used in the above calculations.

2.3. Results

While proven advantageous, the high degree of reduction of carbon in FAs poses a metabolic challenge. The conversion of fatty acids to biomass in *E. coli* results in the generation of reducing equivalents (Figure 2-4) as the average degree of reduction per carbon in FAs is higher than in biomass (e.g. $\kappa_{\text{Palmitic acid/C16:0}} = 5.8$ while $\kappa_{\text{Biomass}} = 4.3$)(Nielsen et al., 2003). Under aerobic conditions the respiratory metabolism is active and the reducing equivalents are transferred to O_2, thus facilitating redox balance and enabling growth on fatty acids. On the contrary, growth on fatty acids is not possible under anaerobic conditions without an electron acceptor (i.e. fermentative conditions), because *E. coli* does not have the ability to ensure redox balance and ATP synthesis in the absence of an external electron acceptor (Section 2.1.1).

The balance between cell growth and generation of reduced chemicals and fuels can therefore be achieved in two ways: by genetic modifications aimed to reduce the efficiency of the respiratory metabolism and thus cell growth or by adjusting the availability and/or rate of consumption of terminal electron acceptors (establishing microrespiratory conditions).

The present study has focused on engineering a respiro-fermentative metabolic mode that supports the synthesis of fermentative products during respiratory metabolism of FAs. To this end, we metabolically engineered native and heterologous pathways for the efficient catabolism of FAs and the synthesis of fuels and chemicals in *E. coli* (Dellomonaco et al., 2010b). Biofuels, commodity chemicals, and polymer building blocks were chosen as model products to illustrate the feasibility of the proposed approach (Dellomonaco et al., 2010b).

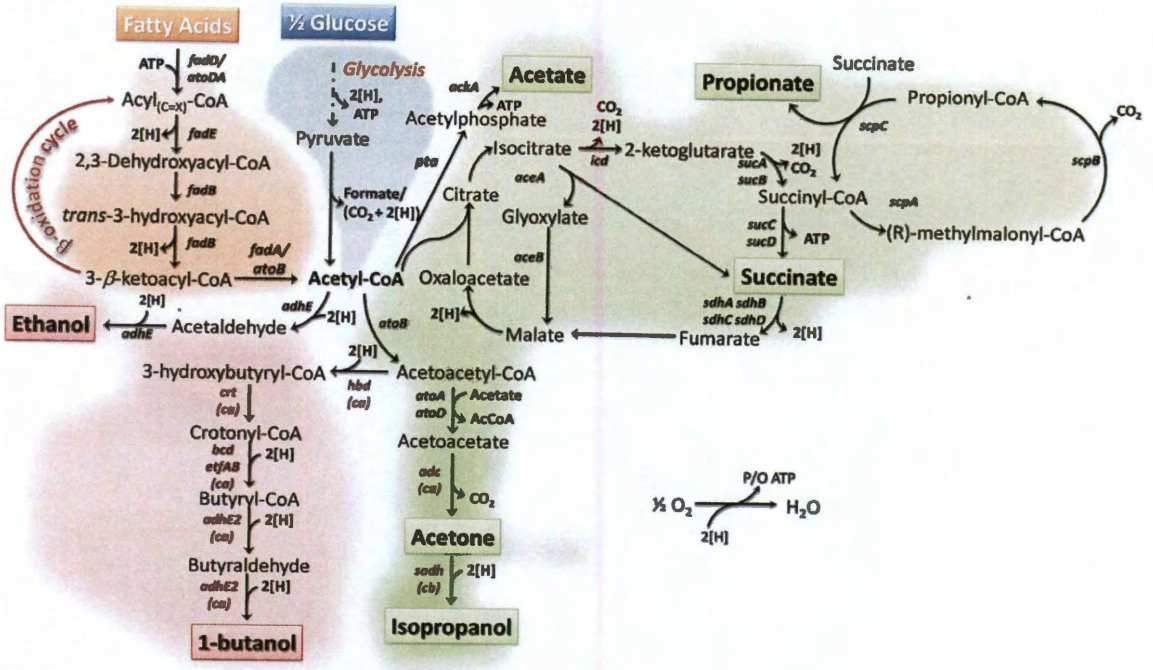

Figure 2-4 Pathways engineered in *Escherichia coli* for the conversion of fatty acids to fuels and chemicals

Fuels and chemicals are shown in red and green shade, respectively. Also shown are the catabolism of fatty acids via the β-oxidation pathway (orange shade) and glucose through the Embden-Meyerhof-Parnas pathway (blue shade). Relevant reactions are represented by the names of the genes coding for the enzymes (*E. coli* genes unless otherwise specified in parenthesis as follows: *C. acetobutylicum, ca; C. beijerinckii, cb*): *aceA*, isocitrate lyase; *aceB*, malate synthase A; *adc*, acetoacetate decarboxylase (*ca*); *ackA*, acetate kinase; *adh*, secondary alcohol dehydrogenase (*cb*); *adhE*, acetaldehyde/alcohol dehydrogenase; *adhE2*, secondary alcohol dehydrogenase (*ca*); *atoA*, *atoD*, acetyl-CoA:acetoacetyl-CoA transferase; *atoB*, acetyl-CoA acetyltransferase; *bcd*, butyryl-CoA dehydrogenase (*ca*); *crt*, crotonase (*ca*); *etfAB*, electron transfer flavoprotein (*ca*); *fadA*, 3-ketoacyl-CoA thiolase; *fadB*, enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase; *fadD*, acyl-CoA synthetase; *fadE*, acyl-CoA dehydrogenase; *hbd*, β-hydroxybutyryl-CoA dehydrogenase (*ca*); *icd*, isocitrate dehydrogenase; *pta*, phosphate acetyltransferase; *sdhABCD*, succinate dehydrogenase; *scpA*, methylmalonyl-CoA mutase; *scpB*, methylmalonyl-CoA decarboxylase; *scpC*, propionyl-CoA:succinate CoA transferase; *sucA*, 2-oxoglutarate dehydrogenase; *sucB*, dihydrolipoyltranssuccinylase; *sucCD*, succinyl-CoA synthetase. Maximum theoretical yields of products were calculated as shown in **Table 2-2** and reported for both glucose (*Gluc*, black) and palmitic acid ($C_{16:0}$, red). Abbreviations: 2[H] = NADH = $FADH_2$ = H_2; P/O, amount of ATP produced per oxygen consumed in the oxidative phosphorylation.

2.3.1. Engineering fermentative pathways in *Escherichia coli* for biofuels production from fatty acids

2.3.1.1. Ethanol

Although ethanol is only a minor product of fermentative metabolism in *E. coli* (Sawers and Clark, 2004), this bacterium has been engineered for the production of ethanol from sugars (Jarboe et al., 2007) and glycerol (Durnin et al., 2009; Yazdani and Gonzalez, 2008). The maximum theoretical yield for ethanol synthesis from glucose is 0.51 g ethanol/ g glucose or 0.67 on C-mole basis (Table 2-2). The low ethanol yield on glucose is due to the low metabolic efficiency of the pathway that converts this sugar into ethanol, which generates two molecules each ethanol and CO_2 from one molecule of glucose (Figure 2-4 and Table 2-2). Given their higher energy content and reduced state, the use of FAs represents a more efficient alternative, as each β-oxidation cycle yields one molecule of acetyl-CoA, which can be converted to ethanol without any carbon loss (Figure 2-4). When accounting for the fact that the last 2-C fragment of a FA molecule converted to acetyl-CoA does not generate reducing equivalents, a maximum theoretical yield of 1.38 (w/w) is obtained from a 16-C FA molecule (i.e. palmitic acid/$C_{16:0}$, the most abundant FA in oil-rich feedstocks; see Section 2.1.2.1) (Table 2-2). This value is 2.71-fold higher than the yield on glucose. On carbon basis, the ethanol yield on $C_{16:0}$ is 1.44-fold higher (Table 2-2). The above calculations, however, do not account for the 2 ATP equivalents required to drive the β-oxidation pathway, which when accounted for would slightly reduce the maximum theoretical (Table 2-2).

Table 2-2 Comparison of maximum theoretical yields for the production of biofuels and biochemicals from fatty acids and lignocellulosic sugars
Calculated yields are shown on both weight (w/w) and carbon (C/C) basis.

Pathway stoichiometry for the synthesis of the specified product from glucose ($C_6H_{12}O_6$) or palmitic acid ($C_{16}H_{32}O_2$) [a]	Maximum yield (w/w)/(C/C)
Biofuels	
Ethanol (C_2H_6O)	
$C_6H_{12}O_6 \rightarrow 2\ C_2H_6O + 2CO_2$	0.51/0.67
$C_{16}H_{32}O_2 \rightarrow 23/3\ C_2H_6O + 2/3\ CO_2$	1.38/0.96
$C_{16}H_{32}O_2 + 51/7\ H_2O \rightarrow 53/7\ C_2H_6O + 6/7\ CO_2 + 8/7\ [H]$; $8/7\ [H] + 2/7\ O_2 \rightarrow 4/7\ H_2O$	*1.36/0.95*
Butanol ($C_4H_{10}O$)	
$C_6H_{12}O_6 \rightarrow C_4H_{10}O + 2CO_2 + H_2O$	0.41/0.67
$C_{16}H_{32}O_2 + 7/2\ H_2O \rightarrow 53/14\ C_4H_{10}O + 6/7\ CO_2 + 8/7\ [H]$; $8/7\ [H] + 2/7\ O_2 \rightarrow 4/7\ H_2O$	*1.10/0.95*
Biochemicals	
Acetate ($C_2H_4O_2$)	
$C_6H_{12}O_6 + 2H_2O \rightarrow 3\ C_2H_4O_2$	1.00/1.00
$C_{16}H_{32}O_2 + 7\ H_2O + 7\ CO_2 \rightarrow 23/2\ C_2H_4O_2$	2.70/1.44
Acetone (C_3H_6O)	
$C_6H_{12}O_6 \rightarrow 3/2\ C_3H_6O + 3/2\ CO_2 + 3/2\ H_2O$	0.48/0.75
$C_{16}H_{32}O_2 + 5/4\ H_2O + 5/4\ CO_2 \rightarrow 23/4\ C_3H_6O$	1.30/1.08
Isopropanol (C_3H_8O)	
$C_6H_{12}O_6 \rightarrow 4/3\ C_3H_8O + 2\ CO_2 + 2/3\ H_2O$	0.44/0.67
$C_{16}H_{32}O_2 + 40/9\ H_2O \rightarrow 46/9\ C_3H_8O + 2/3\ CO_2$	1.20/0.96
Succinate ($C_4H_6O_4$)	
$C_6H_{12}O_6 + 6/7\ CO_2 \rightarrow 12/7\ C_4H_6O_4 + 6/7\ H_2O$	1.12/1.14
$C_{16}H_{32}O_2 + 152/17\ CO_2 + 86/17\ H_2O \rightarrow 106/17\ C_4H_6O_4 + 80/17\ [H]$; $80/17\ [H] + 20/17\ O_2 \rightarrow 40/17\ H_2O$	*2.87/1.56*
Propionate ($C_3H_6O_2$)	
$C_6H_{12}O_6 \rightarrow 12/7\ C_3H_6O_2 + 6/7\ CO_2 + 6/7\ H_2O$	0.70/0.86
$C_{16}H_{32}O_2 + 262/83\ CO_2 + 370/83\ H_2O \rightarrow 530/83\ C_3H_6O_2 + 216/83\ [H]$; $216/83\ [H] + 54/83\ O_2 \rightarrow 108/83\ H_2O$	*1.81/1.20*

[a] Stoichiometry is based on the pathways shown in Figure 2-4 for the utilization of FAs and glucose, the synthesis of products, the TCA cycle, and the oxidative phosphorylation. For the synthesis of biochemicals, CO_2 fixation via the Wood–Ljungdahl pathway (Ragsdale and Pierce, 2008) ($2CO_2$ + ATP + 8[H] → AcCoA) or the carboxylation of phosphoenolpyruvate (Sauer and Eikmanns, 2005) (PEP + CO_2 → OAA + ATP) were also considered (not shown in Figure 2-4). The stoichiometric coefficients were obtained by conducting elemental balances on carbon, hydrogen and oxygen. An ATP balance was also included in the analysis for the reactions shown in *italics*. All other reactions represent ATP-generating pathways. Every

acetyl-CoA oxidized through the TCA cycle generates 3NADH, 1 $FADH_2$ and one ATP equivalent. Eleven ATPs are generated from the oxidation of the NADH and $FADH_2$ produced in the TCA cycle (2 and 3 ATPs per $FADH_2$ and NADH, respectively) via coupling between electron transfer chain and oxidative phosphorylation.

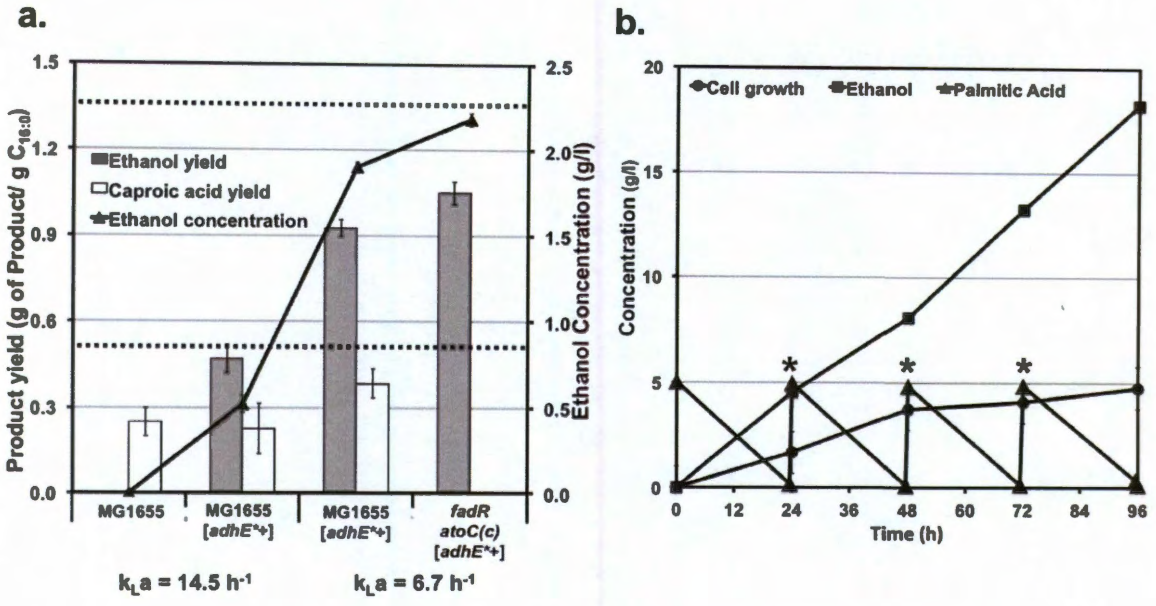

Figure 2-5 Engineering *Escherichia coli* for the production of ethanol from fatty acids
Values represent the means and error bars standard deviations for triplicate cultures. Gene overexpressions and deletions are indicated by "+" and "Δ", respectively, next to the corresponding genes or operons. Details about the pathways can be found in **Figure 2-4**.
(a) Ethanol concentration (line) and ethanol and caproic acid yields (bars) for 72-hour cultures of wild-type MG1655 and strains containing engineered pathways for ethanol production (oxygen-tolerant acetaldehyde/alcohol dehydrogenase, *adhE**) and efficient catabolism of FAs (*fadR atoC*(c)). k_La = volumetric oxygen transfer coefficient (h^{-1}).
(b) Fermentation profile for strain Δ*adhE fadR atoC*(c) [*adhE**+] in minimal medium with 5 g/L palmitic acid ($C_{16:0}$) and using a k_La = 6.7 h^{-1}. Additions of palmitic acid (5 g/L each) were made every 24 hours. ΔCDW, increase in cell dry weight with respect to initial value.

Despite the aforementioned theoretical advantages of FAs, no ethanol was produced by wild-type *E. coli* strain MG1655 during the respiratory metabolism of palmitic acid (Figure 2-5). Efforts to engineer homoethanol production in *E. coli* have focused on diverting carbon at the pyruvate node by manipulating native (pyruvate dehydrogenase) or heterologous (*Zymomonas mobilis*' pyruvate decarboxylase) pathways (Ohta et al.,

1991). These strategies, however, are not applicable to the production of ethanol from FAs because pyruvate is not an intermediate in the β-oxidation pathway (Figure 2-4). Instead, we focus on diverting carbon to ethanol synthesis at the acetyl-CoA node by manipulating the enzyme acetaldehyde/alcohol dehydrogenase (AdhE) (Figure 2-4). Since AdhE is oxygen sensitive and *adhE* expression is very low in the presence of oxygen (Holland-Staley et al., 2000), we hypothesized that the lack of ethanol production in MG1655 could be related to the low activity/expression of this enzyme under aerobic conditions.

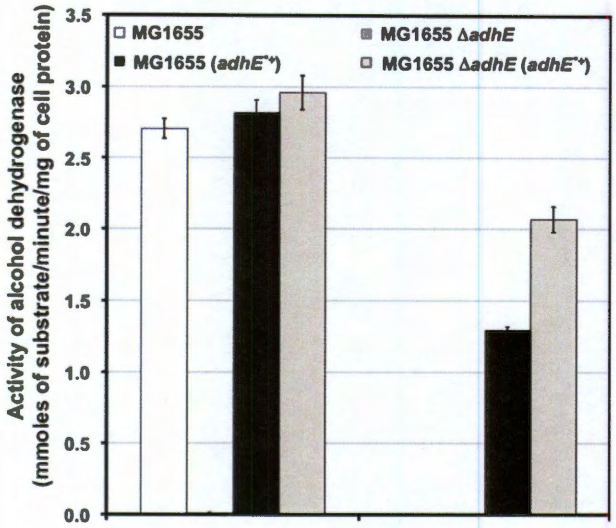

Figure 2-6 Activity of alcohol dehydrogenase (AdhE) under anaerobic and aerobic conditions

We then constructed an aerotolerant mutant of AdhE (AdhE*) and overexpressed it from the oxygen-independent *trc* promoter. AdhE* was obtained by site-directed mutagenesis of AdhE to substitute the residue at position 568 (Glu568Lys) (Holland-Staley et al., 2000). Enzyme assays of cells overexpressing AdhE* confirmed its high activity under both anaerobic and aerobic conditions (Figure 2-6). Overexpression of AdhE* in MG1655 resulted in the production of ethanol at a yield of 0.48 g /g $C_{16:0}$ (Figure 2-5a: strain MG1655 [*adhE**+]). To investigate a potential competition between

the ethanol and biosynthetic and respiratory pathways for reducing equivalents, we decreased the respiration rate by lowering the volumetric oxygen transfer coefficient (k_La) from 14.5 h^{-1} to 6.7 h^{-1}. This modification led to a 2-fold increase in ethanol yield (0.92 g ethanol/g $C_{16:0}$: Figure 2-5a).

While the ethanol yield was very high in strain MG1655 [*adhE**+], oxidation of palmitic acid was incomplete resulting in the accumulation of caproic acid ($C_{6:0}$) (Figure 2-5a). The metabolism of short-chain FAs (C_4 to C_6) in *E. coli* requires expression of both the *fad* regulon (regulated by FadR) and two degradative enzymes encoded by the *atoDAEB* operon (regulated by AtoC) (Clark and Cronan, 2005) (Figure 2-1). Previous studies have shown that mutants exhibiting constitutive expression of the *fad* regulon and *ato* operon (referred to here as *fadR* and *atoC*(c), respectively) catabolize both long- and short-chain FAs (Jenkins and Nunn, 1987b; Nunn, 1986b). However, such mutants have not been tested for their ability to completely oxidize FAs without the generation of short-chain FAs as by-products, which is the specific problem observed in strain MG1655 [*adhE**+].

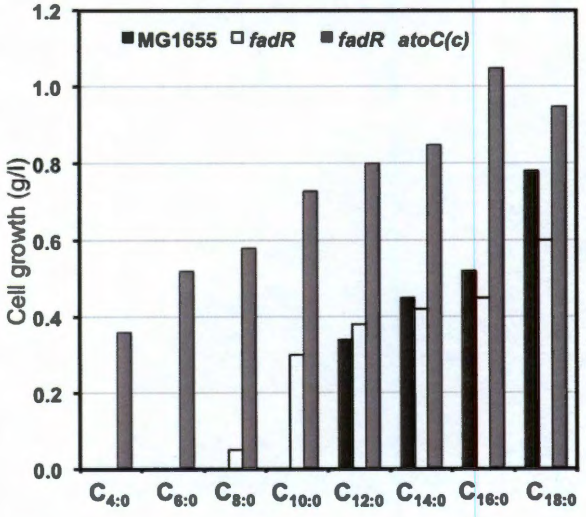

Figure 2-7 Growth comparison of wild-type *Escherichia coli* MG1655 and derivatives *fadR* and *fadR atoC*(c) on saturated fatty acids of different chain-lengths
Experiments were performed using MOPS minimal medium and baffled shake flasks for 96 hours, as reported in Section 2.2.2.

Therefore, we engineered the *fadR atoC*(c) phenotype in MG1655 (Figure 2-7) and evaluated its ability to prevent the generation of by-product caproic acid. Overexpression of AdhE* in strain *fadR atoC*(c) resulted in efficient metabolism of palmitic acid without the accumulation of caproic acid and also high ethanol yield (Figure 2-5a). It is noteworthy that the ethanol yield in this strain is twice the maximum theoretical from sugars (w/w basis) and already represents 77% of the maximum achievable from palmitic acid (Figure 2-5a). We also performed studies to illustrate the production of ethanol at higher titers. For this purpose, experiments were conducted using higher concentration of palmitic acid and higher cell density (Figure 2-5b). Under these conditions, strain *fadR atoC*(c) [*adhE**+] produced ethanol as the only fermentation product at concentrations close to 20 g/l, demonstrating that high yield and titers can be achieved.

2.3.1.2. Butanol

While microbial production of butanol traditionally utilizes *Clostridium acetobutylicum* (Maddox, 1989; Mermelstein et al., 1994), the limitations of this organism has triggered efforts to engineer *E. coli*, *Pseudomonas putida*, *Bacillus subtilis*, and *Saccharomyces cerevisiae* for butanol production, reaching titers and yields of up to 1.2 g/l and 6 % (w/w), respectively (Atsumi et al., 2008a; Inui et al., 2008; Nielsen et al., 2009; Steen et al., 2008).

Production of butanol from FAs would provide a 2.7-fold yield advantage over sugars (w/w), as can be seen in Table 2-2. Our strategy to produce butanol from FAs in *E. coli* is based on the engineering of a synthetic pathway composed of the genes required for

butanol synthesis from acetoacetyl-CoA in *C. acetobutylicum* (*crt*, *bcd*, *etfAB*, *hbd*, *adhE2*) in combination with the *E. coli* gene responsible for the conversion of acetyl-CoA to acetoacetyl-CoA (*atoB*) (Figure 2-4).

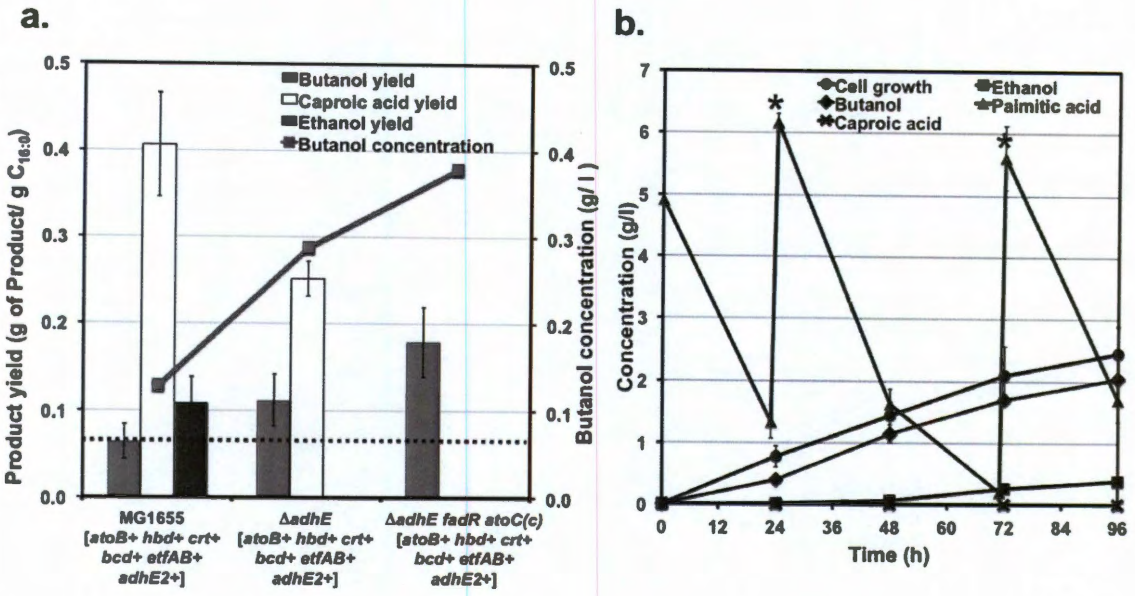

Figure 2-8 Engineering *Escherichia coli* for the production of butanol from fatty acids
Values represent the means and error bars standard deviations for triplicate cultures. Gene overexpressions and deletions are indicated by "+" and "Δ", respectively, next to the corresponding genes or operons. Details about the pathways can be found in **Figure 2-4**.
(a) Butanol concentration (line) and butanol, ethanol and caproic acid yields (bars) for 72-hour cultures of strains engineered by: i) expression of *C. acetobutylicum*'s pathway for the synthesis of butanol from acetoacetyl-CoA (*hbd*, *crt*, *bcd*, *etfAB*, and *adhE2*), ii) overexpression of *E. coli*'s acetyl-CoAacetyltransferase (*atoB*) for the conversion of acetyl-CoA to acetoacetyl-CoA, iii) elimination of *E. coli*'s native ethanol pathway (*adhE*), and iv) engineering the β-oxidation pathway for efficient catabolism of FAs (*fadR atoC*(c)). A $k_La = 6.7\ h^{-1}$ was used.
(b) Fermentation profile for strain Δ*adhE fadR atoC*(c) [*atoB*+ *hbd*+ *crt*+ *bcd*+ *etfAB*+ *adhE2*+] in minimal medium with 5 g/L palmitic acid ($C_{16:0}$) and a $k_La = 6.7\ h^{-1}$. Two additions of palmitic acid (5 g/L each) were made at 24 and 72 hours. ΔCDW, increase in cell dry weight respect to initial value.

When these pathways were expressed in *E. coli* cells metabolizing palmitic acid, we observed significant production of butanol (Figure 2-8a, strain MG1655 [*atoB*+ *hbd*+ *crt*+ *bcd*+ *etfAB*+ *adhE2*+]). We hypothesized that the relatively low yields and titers

observed originate in part from the competition between the ethanol and butanol pathways for acetyl-CoA and reducing equivalents (Figure 2-4). In agreement with this hypothesis, deletion of the native ethanol pathway (Δ*adhE*) resulted in a two-fold increase in butanol yield (Figure 2-8a, strain Δ*adhE* [*atoB*+ *hbd*+ *crt*+ *bcd*+ *etfAB*+ *adhE2*+]). As in ethanol-producing strains (see above), the engineering of the β-oxidation pathway for more efficient utilization of FAs (*fadR atoC*(c)) also led to an increase in product yield (Figure 2-8a, strain *fadR atoC*(c) [*atoB*+ *hbd*+ *crt*+ *bcd*+ *etfAB*+ *adhE2*+]). Overall, our approach resulted in butanol yields up to 3-fold higher than those currently reported in the literature (Atsumi et al., 2008a; Inui et al., 2008; Nielsen et al., 2009; Steen et al., 2008) and represent 16.4% of the maximum achievable from palmitic acid.

To further ascertain the potential of the engineered strains for butanol production, we performed experiments with strain Δ*adhE fadR atoC*(c) [*atoB*+ *hbd*+ *crt*+ *bcd*+ *etfAB*+ *adhE2*+] at higher palmitic acid concentration and higher cell density (Figure 2-8b). Butanol was produced as the main fermentation product at a concentration of 2.05 g/l, along with very small amounts of ethanol (Figure 2-8b). The synthesis of ethanol could be due to the ability of Clostridial secondary alcohol dehydrogenase to catalyze the conversion of acetyl-CoA to ethanol (Petersen et al., 1991; Youngleson et al., 1988). Even when the butanol pathway has yet to be optimized, the butanol yield and titer achieved by our strategy are superior to those reported with strains engineered to produce butanol from other carbon sources (Atsumi et al., 2008a; Inui et al., 2008; Nielsen et al., 2009; Steen et al., 2008).

2.3.2. Engineered strains for the production of biochemicals from fatty acids

FA-containing (industrial) by-product streams such as palm fatty acid distillate or palm oil mill effluent (Sumathi et al., 2008; Wu et al., 2009) can be exploited for biological production of chemicals (refer to Section 2.1.2.1 for details). In the next sections we present the engineering of *E. coli* to enable the respiro-fermentative production of commodity chemicals, organic solvents, and polymer building blocks from FAs.

2.3.2.1. Acetic Acid (Acetate)

Acetate is an important industrial chemical (Agreda and Zoeller, 1993) and a native product of sugar fermentation in *E. coli*, where it is primarily produced through the phosphotransacetylase (PTA)-acetate kinase (ACK) pathway (Clark and Cronan, 2005) (Figure 2-4). Although the maximum theoretical yield for acetate production from glucose is 1.00 (w/w) (Table 2-2), the biological process used in the commercial production of acetate from sugars (Cheryan et al., 1997) and that developed by engineering *E. coli* (Causey et al., 2003) are both based on a maximum theoretical yield of 0.67 g of acetate/g glucose. The synthesis of acetate from FAs contributes in CO_2 fixation and can be achieved at a much higher yield: e.g. a yield of 2.70 (w/w) can be realized on palmitic acid (Table 2-2).

Unlike metabolism of sugars, utilization of FAs by *E. coli* does not lead to the synthesis of acetate (Figure 2-9a). Since acetyl-CoA is generated in the β-oxidation of FAs, we reasoned that overexpression of the ACK-PTA pathway (Figure 2-4) should lead to significant production of acetate. Indeed, strain MG1655 [*ackA+ pta+*] produced

acetate at a yield of 0.65 g acetate/ g $C_{16:0}$ (Figure 2-9a). As in the case of other products discussed in previous sections, a strain with constitutive expression of the *fad* regulon and *ato* operon (i.e. strain *fadR atoC*(c) [*ackA*+ *pta*+]) not only utilized palmitic acid more efficiently but improved acetate production as well (Figure 2-9a), exhibiting yields that surpassed those reported for *E. coli* strains engineered to produce acetate from glucose (Causey et al., 2003).

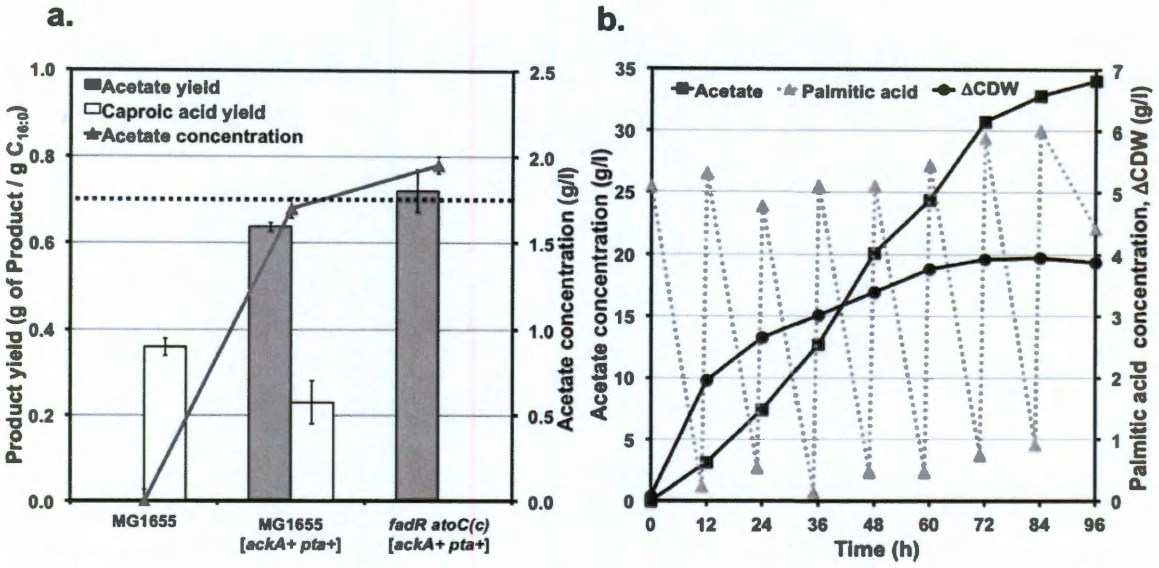

Figure 2-9 Production of acetate from FAs in engineered *Escherichia coli* strains
Details about the engineered pathways are shown in **Figure 2-4**. Gene overexpressions is indicated by "+"next to the corresponding gene. A k_La = 14.5 h^{-1} was used in all experiments. Values represent the means and error bars standard deviations for 72-hour triplicate cultures.
(a) Acetate concentration (line) and acetate and caproic acid yields (bars) in wild-type MG1655 and recombinant strains constructed by overexpressing the native phopshoacetyltransferase (*pta*)-acetate kinase (*ackA*) pathway for conversion of acetyl-CoA to acetate and modifying the β-oxidation pathway for efficient catabolism of FAs (*fadR atoC*(c)).
(b) Fermentation profile for strain *fadR atoC*(c)[*ackA*+*pta*+] in minimal medium with 5 g/L palmitic acid ($C_{16:0}$). Additions of palmitic acid (5 g/L each) were made every 12 hours. ΔCDW, increase in cell dry weight respect to initial value.

An experiment was conducted to assess the production of acetate at higher titers. As shown in Figure 2-9b, 35 g/L of acetate were produced at a yield of 1.10 g acetate/ g $C_{16:0}$

(40.7% of the maximum achievable from palmitic acid), demonstrating the feasibility of achieving high yields and titers. It is noteworthy that this yield is in excess of the maximum theoretical on glucose (Table 2-2) and about 2-fold higher than the best yield reported to date by an *E. coli* strain engineered to produce acetate (Causey et al., 2003).

2.3.2.2. Acetone

Acetone is a product of the acetone-butanol fermentation by *Clostridium* spp. (Maddox, 1989) and an industrially relevant chemical.

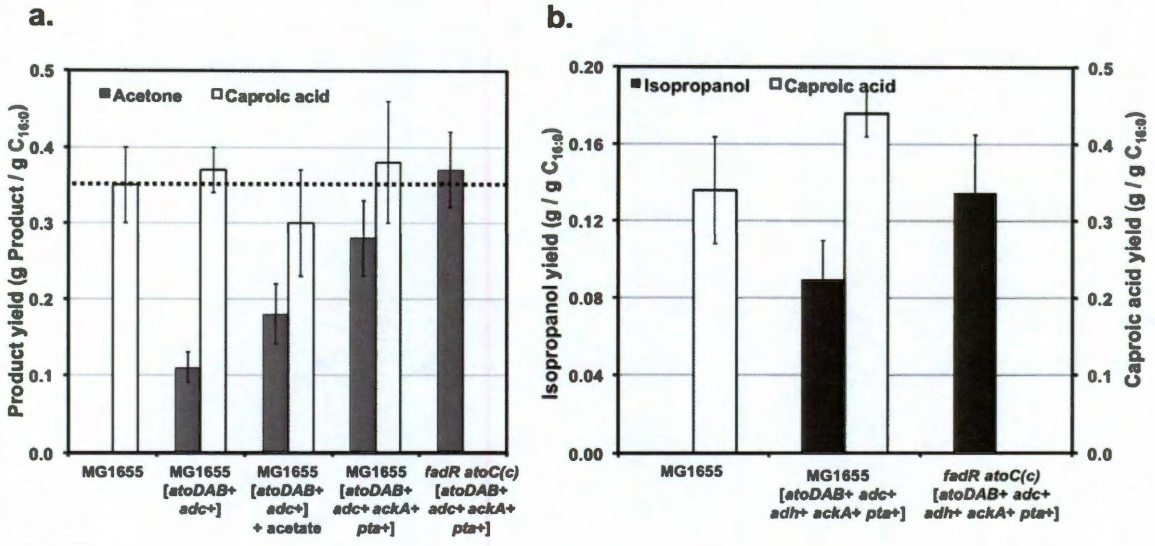

Figure 2-10 Production of acetone and isopropanol from fatty acids in engineered *Escherichia coli* strains

Details about the engineered pathways are shown in **Figure 2-4**. Gene overexpression is indicated by a "+" next to the corresponding genes. A $k_La = 14.5\ h^{-1}$ was used in all experiments. Values represent the means and error bars standard deviations for 72-hour triplicate cultures.

(a) Acetone and caproic acid yields in strains constructed by overexpressing *E. coli* acetyl-CoA and acetoacetyl-CoA transferases (*atoDAB*) and *C. acetobutylicum* acetoacetate decarboxylase (*adc*) along with amplification of the native acetate pathway (*ackApta*). These modifications were implemented in wild-type MG1655 and the *fadR atoC*(c) derivative.

(b) Isopropanol and caproic acid yields in strains engineered for acetone production and overexpressing a secondary alcohol dehydrogenase from *C. butyricum* (*adh*), which converts acetone to isopropanol.

Low product selectivity and titers are key factors that represent a driver for heterologous expression of this solventogenic pathway in industrial organisms such as *E. coli* (Bermejo et al., 1998). The acetone pathway involves the conversion of acetyl-CoA to acetone via reactions catalyzed by three different enzymes (Jones and Woods, 1986), as shown in Figure 2-4.

Conversion of sugars to acetone is also inefficient in terms of C-recovery, with a maximum theoretical yield on glucose limited to 0.48 (w/w) (Table 2-2). Utilization of palmitic acid would support a yield 2.71-fold higher than sugars (Table 2-2).

Since acetone is not a native product of *E. coli* metabolism, a synthetic pathway containing genes from *C. acetobutylicum* and *E. coli* was required to convert the acetyl-CoA produced by the β-oxidation of FAs to acetone (Figure 2-4). While significant acetone production was achieved with this strategy, the resulting yield was low (0.11 g acetone /g $C_{16:0}$: Figure 2-10a, strain MG1655 [*atoDAB*+ *adc*+]). We hypothesized that acetone production may be limited by the low levels of acetate, a metabolite required in the second step of the acetone pathway (i.e. the conversion of acetoacetyl-CoA to acetoacetate, Figure 2-4). In agreement with this hypothesis, supplementation of the growth medium with acetate increased the acetone yield by more than 60% (Figure 2-10a). Simultaneous overexpression of the acetate-producing ACK-PTA pathway, along with the synthetic acetone pathway, enabled the synthesis of acetone in the absence of added acetate at a yield of 0.28 g acetone/ g $C_{16:0}$ (Figure 2-10a, strain MG1655 [*atoDAB*+ *adc*+ *ackA*+ *pta*+]). When this strategy was implemented in a strain with constitutive expression of the *fad* regulon and *ato* operon (i.e. strain *fadR atoC*(c)), efficient utilization of palmitic acid along with further improvement in acetone

production was observed (Figure 2-10a). The acetone yield on this strain (0.37 g acetone/ g $C_{16:0}$) surpassed the yields previously reported using engineered *E. coli* (Bermejo et al., 1998; Hanai et al., 2007) and corresponds to 28.5% of the maximum achievable from palmitic acid.

2.3.2.3. Isopropanol

Isopropanol is a secondary alcohol with applications as both chemical intermediate and solvent (Logsdon and Loke, 2000). Isopropanol is produced in *C. beijerinckii* (Chen and Hiu, 1986; George et al., 1983) from acetyl-CoA via the acetone pathway (Figure 2-4), as part of a mixed product fermentation. Limitations in the utilization of native strains have increased interest to engineer the pathway in industrially relevant hosts (Hanai et al., 2007). Isopropanol production from glucose is limited to a yield of 0.44 (w/w) (Table 2-2). When compared to sugars, the use of FAs offers a yield 2.73-fold higher (Table 2-2). Since isopropanol is a non-native product of *E. coli* metabolism, we introduced the clostridial route, which produces this alcohol from acetyl-CoA, as shown in Figure 2-4. For this purpose, three different pathways were assembled in strain MG1655 (Figure 2-4): i) the *C. acetobutylicum* acetone pathway that converts acetyl-CoA to acetone (see Section 2.3.2.2), ii) a secondary alcohol dehydrogenase from *C. beijerinckii* that converts acetone to isopropanol, and iii) the *E. coli* pathway that converts acetyl-CoA to acetate. The resulting strain, MG1655 [*atoDAB*+ *adc*+ *ackA*+ *pta*+ *adh*+], produced isopropanol at a yield of 0.09 g/ g $C_{16:0}$ (Figure 2-10b). When these pathways were assembled in the *fadR atoC*(c) strain, efficient metabolism of FAs and further increase in isopropanol yield were observed (0.13 g of isopropanol/ g $C_{16:0}$, 10.8% of the maximum achievable from palmitic acid) (Figure 2-10b).

2.3.2.4. Succinic Acid (Succinate)

Succinate is expected to become a future platform chemical (Werpy and Petersen, 2004) and major efforts in recent years have focused on its production by microbial fermentation of sugars (Hong and Lee, 2002; Song and Lee, 2006).

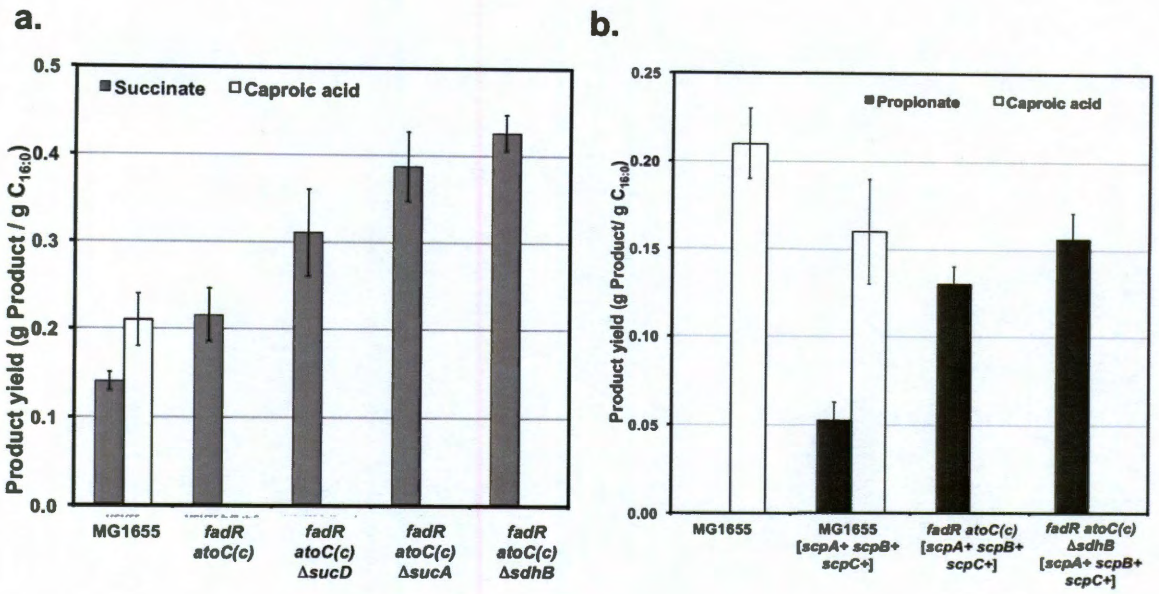

Figure 2-11 Production of succinate and propionate from fatty acids in engineered *Escherichia coli* strains.
Details about the engineered pathways are shown in **Figure 2-4**. Gene overexpressions and deletions are indicated by "+" and "Δ", respectively, next to the corresponding genes or operons. A k_La = 14.5 h^{-1} was used in all experiments. Values represent the means and error bars standard deviations for 72-hour triplicate cultures.
(a) Effect of *fadR atoC*(c) genotype and deletions of genes encoding TCA cycle enzymes on succinate and caproic acid yields.
(b) Propionate and caproic acid yields in strains overexpressing a metabolic cycle that catalyzes the decarboxylation of succinate to propionate (*scpA*, *scpB*, *scpC*). The effect of amplification of this cycle, along with the deletion of succinate dehydrogenase (Δ*sdhB*), was evaluated in wild-type MG1655 and *fadR atoC*(c) derivative.

Succinate is a minor product of sugar fermentation in *E. coli* (Clark, 1989) but it is not produced under aerobic conditions because this metabolite is an intermediate of the

TCA cycle (Figure 2-11). As in the case of other products discussed above, production of succinate from FAs would offer a significant yield advantage (Table 2-2). Metabolic engineering strategies previously reported for the production of succinate from sugars in *E. coli* (Lin et al., 2005; Song and Lee, 2006) cannot be implemented for its production from FAs because of the significant differences in the metabolic pathways. For example, while the synthesis of succinate from sugars is limited by the availability of both reducing equivalents and intermediate phosphoenolpyruvate (Hong and Lee, 2002), its production from FAs is a redox generating process that relies on the availability of acetyl-CoA (see Table 2-2 and Figure 2-4). Since *fadR* mutants are known to exhibit higher expression of the glyoxylate shunt enzymes (Gui et al., 1996), and succinate is generated in this pathway (Figure 2-4), we evaluated succinate production in strain *fadR atoC*(c) and observed a significant increase in yield (Figure 2-11a). Succinate production was further increased by minimizing its conversion to fumarate (Δ*sdhB* mutation) and decreasing the oxidation of acetyl-CoA to CO_2 via the TCA cycle (Δ*sucC*/Δ*sucD* mutations) (Figure 2-4 and Figure 2-11a). These engineering strategies resulted in up to 2.3-fold increase in succinate yield with respect to wild-type (Figure 2-11a), that represents 14.6% of the maximum achievable from palmitic acid.

2.3.2.5. Propionic acid (Propionate)

There is a high interest in producing propionate from sugars via fermentation with Propionibacteria but this process suffers from relatively low product yield and concentration (Himmi et al., 2000). While a propionate yield on glucose of 0.70 (w/w) can be realized the use of FAs would support yields approximately 2.59-fold higher

(Table 2-2). Three of the genes located in the *E. coli scpAargKscpBscpC* operon encode the enzymes of a postulated metabolic cycle that can catalyze the decarboxylation of succinate to propionate (Figure 2-4) (Haller et al., 2000). However, propionate has not been reported as a product of *E. coli* metabolism. By overexpressing the aforementioned genes, we achieved for the first time the production of propionate in this organism (Figure 2-11b). Since succinate is the precursor of propionate in this pathway, the same modifications that led to increased production of succinate (Figure 2-11a) also resulted in higher propionate yields (Figure 2-11b), with a maximum yield of 8.8% of the maximum achievable from palmitic acid.

2.4. Discussion

The use of FA-rich feedstocks for the production of biofuels and biochemicals is a promising avenue to establish biorefineries. In addition to their abundance, the metabolism of FAs is highly efficient as it yields 100% carbon recovery in the key intermediate metabolite acetyl-CoA from which most biofuels and biochemicals can be derived (Figure 2-2). This high metabolic efficiency, along with the high reduced state of carbon in FAs, could enable the production of fuels and chemicals at yields superior than those obtained with the use of lignocellulosic sugars (Table 2-2). However, metabolism of FAs requires the presence of an external electron acceptor, which in turn precluded the synthesis of fermentation products (Figure 2-5 and Figure 2-8). To overcome this hurdle, we engineered a respiro-fermentative metabolic mode that enables the efficient production of the desired fuels and chemicals in combination with adequate catabolism of FAs.

E. coli was chosen as model organism to illustrate the feasibility of this approach, which was demonstrated by engineering the synthesis of ethanol, butanol, acetate, acetone, isopropanol, succinate, and propionate (Figures 1-5, 1-8, 1-9, 1-10, and 1-11). This work represents the first effort towards engineering microorganisms for the conversion of FAs to the aforementioned products. The yield of ethanol (Figure 2-5), acetate (Figure 2-9), and acetone (Figure 2-10a) in the engineered strains exceeded those reported in the literature for their production from sugars (Bermejo et al., 1998; Causey et al., 2003; Hanai et al., 2007; Jarboe et al., 2007). In the case of ethanol and acetate the yields also surpassed the maximum theoretical values that can be achieved from lignocellulosic sugars. Butanol (Figure 2-8), on the other hand, was produced at yields and titers between two- and three-fold higher than those reported for its production from lignocellulosic sugars in engineered *E. coli*, *P. putida*, *B. subtilis*, and *S. cerevisiae* strains (Atsumi et al., 2008a; Inui et al., 2008; Nielsen et al., 2009; Steen et al., 2008). Our work (Dellomonaco et al., 2010b) also represents the first report of propionate production in *E. coli* (Figure 2-11b), a compound previously thought to be synthesized only by Propionibacteria. Finally, the synthesis of isopropanol and succinate was also demonstrated (Figure 2-10b and Figure 2-11a).

Taken together, the results reported here represent the foundation to establish a new approach for the production of biofuels and biochemicals from renewable feedstocks. The synthetic respiro-fermentative metabolic mode engineered in this work will also serve as the basis for the synthesis of other reduced products in the presence of external electron acceptors (e.g. oxygen), thus taking advantage of the most significant attributes of respiratory and fermentative metabolism.

Chapter 3

Engineering functional differentiation in bacteria by harnessing the pluripotent capabilities of native hosts

3.1. Motivation and background

The availability of diverse biomass feedstocks provides ample opportunity to employ microbial means for the conversion of a wide variety of carbon sources into valuable products. Exploitation of diverse carbon sources therefore entails solid knowledge of the different metabolic pathways associated with the metabolism of each carbon source in order to develop and implement a variety of engineering strategies for the efficient production of fuels and chemicals from each available source.

Many native microorganisms possess biochemical capabilities to utilize biomass constituents and convert them into fuels and chemicals (Alper and Stephanopoulos, 2009). However, native producers are often slow growers and they synthesize a range of byproducts

along with the molecule of interest. Strategies designed to engineer cells and their metabolic pathways have been extensively exploited over the past few decades in order to improve biosynthesis and process efficiencies (Bailey, 1991; Stephanopoulos, 2002). However, genetic manipulation of native organisms to convert them into efficient microbial platforms is often hindered by their genetic intractability, making desirable establishing expression of their entire biosynthetic gene clusters in a more suitable host. As a consequence, more recently, metabolic engineering and the emerging field of synthetic biology (Pleiss, 2006) have provided key tools aiding in the design, assembly, and implementation of various heterologous pathways in model industrial organisms (Clomburg and Gonzalez, 2010; Dellomonaco et al., 2010a; Jarboe et al., 2010), building upon the functional diversity among species, as inferred by phylogenetics and taxonomy.

3.1.1. Microbial functional diversity

Microbial functional diversity is at present highly investigated (Bernhardsson et al., 2011; Caetano-Anolles et al., 2009; Dalmolin, 2011; Hernandez-Montalvo et al., 2001; Koonin and Wolf, 2010; Moya et al., 2009; Pereto', 2011; Tripathi et al., 2007) given the impressive biosynthetic potential and the implausible plasticity of microorganisms to adapt to changing environment and nutritional conditions. Modern studies on microbial functional evolution seek to determine the underlying processes determining functional diversity and investigate the reasons why otherwise phenotypically identical bacterial genotypes result instead genetically distinct (Travisano, 2005), suggesting a need for the development of metrics to measure diversity that can go beyond phenetic considerations. Early taxonomic classification of microorganisms was based solely on phenotype, morphology and nutritional requirements of bacteria (Bergey, 1994; den Dooren de Jong, 1926). However, these criteria

of categorization failed to provide evolutionary relationships (Koonin, 2010). The criteria (morphological traits and phenotypes) the classification was based on were for example very subjective and ambiguous (Palleroni, 1983). For the case of bacterial species, this type of classification was very difficult as many of them share similar cell shape, size, arrangement of flagella, etc. Important criteria like staining reactions, growth-related properties, physiology, and serology were therefore introduced to ease the process of categorization. In 1990, Carl Woese addressed the shortcomings of the phenetic method by proposing a restructured scheme of biological classification based on comparison of ribosomal RNA sequences and established a sequence-based phylogenetic tree to relate different organisms (Woese et al., 1990). Evolutionary distance in the new system proposed by Woese does not account for time but is instead solely estimated by the differences identified when pairs of rRNA sequences are aligned (Olsen and Woese, 1993). He classified all the organisms as belonging to three domains: Eucarya (Eukaryotes), Bacteria and Archea (Figure 3-1).

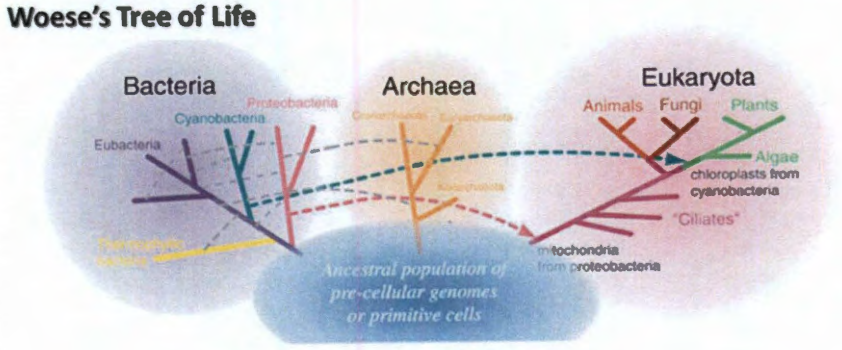

Figure 3-1 A schematic of the phylogenetic tree created by Carl Woese shows the three domains of life: Bacteria, Eucaria, and Archaea. Eukaryotes, Archeans and Bacteria all share a common ancestor (Woese et al., 1990).

Based on this cladistic classification, i.e. a categorization based on evolutionary pathways derived from homology of macromolecules (Priest and Austin, 1993), we expect a striking and compelling difference from a molecular and metabolic standpoint between Eucarya and Bacteria. However, consistent with the early appearance of bacteria on Earth (Rasmussen et al., 2008), prokaryotes (including Bacteria and Archea) and Eukaryotes contain some common pathways for the dissimilation of carbohydrates, energy generation and reducing power (Moat, 2002). Metabolic modes of energy generation in Eukaryotes (Helms, 1997), with energy being produced in the form of ATP, encompass alcohol fermentation (e.g. *Saccharomyces* sp.) (Vandijken and Scheffers, 1986), homolactic fermentation (e.g. muscle and red blood cells), aerobic respiration (e.g. molds, protozoa, and animals) or anaerobic respiration (parasitic helminths, fresh-water snails, lower marine organisms, etc.) (Tielens and Van Hellemond, 1998), and oxygenic photosynthesis (e.g. algae, plants, cyanobacteria). The aforementioned metabolic modes for energy generation exist also within Prokaryotes, in addition to alternative ways of producing energy that are instead typical of microbial metabolism (Herrmann et al., 2008) (Thauer et al., 1977) .

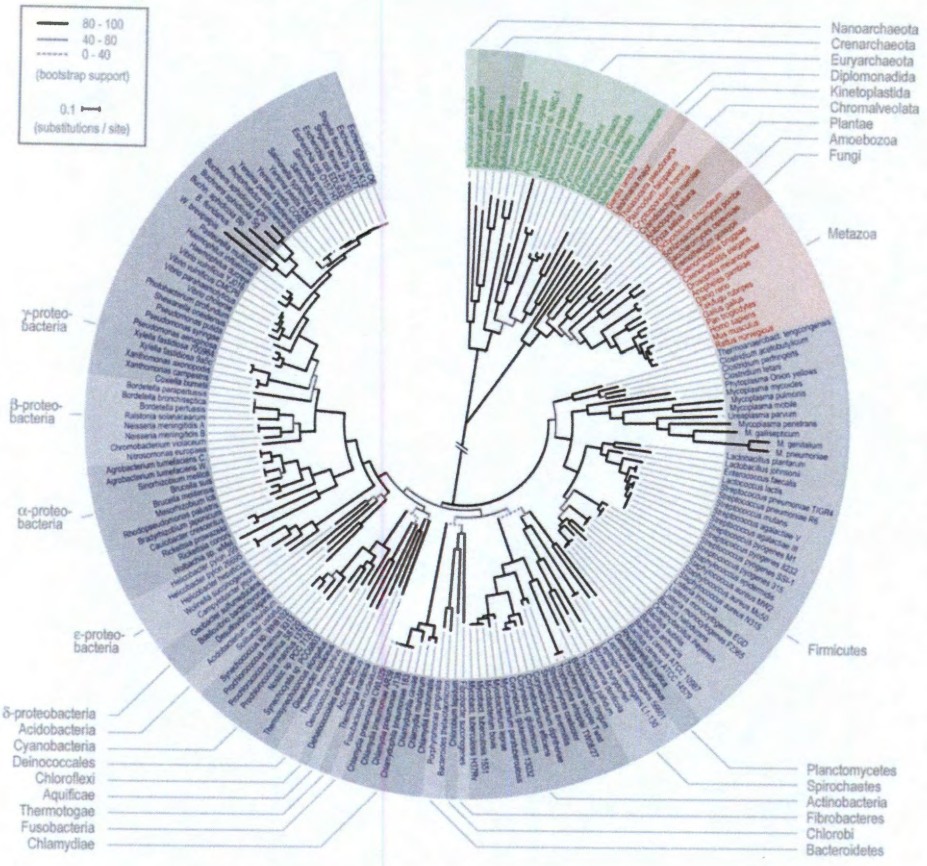

Figure 3-2 Phylogenetic tree of life of fully sequenced organisms
(Ciccarelli et al., 2006) Green section, Archaea; red, Eukaryota; blue, Bacteria.

For instance, while evolutionarily distant from a molecular/genetic standpoint in Woese's tree of life (Figure 3-2), bacteria can ferment sugars in the same way that brewing yeast (i.e., *Saccharomyces cerevesiae*) ferments sugars to produce ethanol and carbon dioxide. Ethanol is indeed the main product of the fermentative metabolism of both *Zymomonas* and *Saccharomyces* species, belonging to the phylum of Proteobacteria (Bacteria) and Ascomycota (Fungi), respectively. While dissimilation of sugars to pyruvate employs distinct metabolic pathways (Smolke, 2009), in both species ethanol is produced via a 2-step enzymatic process from pyruvate, the end product of glycolysis (Pronk et al., 1996; Rogers, 2007). The first reaction decarboxylates pyruvate to acetaldehyde and CO_2 and it is

mediated by pyruvate decarboxylase, a key-enzyme in the eukaryotic alcohol fermentation. In both species acetaldehyde is then reduced to ethanol via an alcohol dehydrogenase (Pronk et al., 1996; Rogers, 2007). Lactic acid fermentation is another noteworthy example of functional similarities in evolutionary distant species that takes place in several microorganisms and higher organisms, including plants and vertebrates (Smolke, 2009). Lactic acid fermentation occurs for example in the muscle cell under hypoxic conditions; while usually muscle cells in vertebrates get the energy needed for contraction from aerobic metabolism, during high activity the circulatory system is unable to meet the cells oxygen demand (Smolke, 2009). Hence, the breakdown of glycogen through glycogenolysis releases glucose in the form of glucose 6-phosphate that is promptly channeled in the glycolysis to produce ATP as energy source and pyruvate, in turn converted to lactic acid. The reaction mediating the conversion of pyruvate to lactic acid consumes the reducing equivalents generated through the glycolysis and, by regenerating NAD+, is accordingly essential to maintain redox balance and glycolytic flux. Although the notable phylogenetic distance with animal cells, *Lactobacillales* perform homolactic fermentation in the same way under conditions of excess glucose and limited oxygen (Kandler, 1983). Morphology/anatomy, nutritional requirements and genetics of bacteria might at this point appear to be relatively easy to ascertain compared to most eukaryotes, probably suggesting more conserved metabolic capabilities in the realm of Bacteria. In fact, while eukaryotes exhibit considerable structural diversity, no such a drastic distinction is identified among the prokaryotes. Prokaryotic cells are, unlike eukaryotic ones, not compartmentalized; therefore, incompatible metabolic processes cannot be separated within discrete compartments (Pace, 1997). As a consequence, the hurdle posed by the inability of confining cellular processes has hindered

structural evolution of prokaryotes, while pushing their metabolism toward the acquisition of a more diverse set of functions (Pace, 1997). Prokaryotes tolerate extreme conditions and flourish in any habitat suitable for life on earth. Hence, their diversity needs to be accounted with regards to the reactions they can mediate or the type of stress they can endure. Dramatic differences in metabolism are then readily identified within the domain of Bacteria.

3.1.2. Bacterial fermentative metabolism

Three physiological groups of bacteria can be distinguished based on their anaerobic metabolism: phototrophs, respirers (e.g. denitrifiers, sulfate reducers, methanogens and acetogens), and fermentative organisms (Ljungdahl, 2003). Respiration and fermentation are two alternative energy-generating metabolic modes that chemoheterotrophes (i.e. organisms that use organic compounds as carbon and energy source) utilize (Ljungdahl, 2003). Respiration entails the complete oxidation of carbohydrates (or any other organic compound) to CO_2 and H_2O; molecular O_2 and nitrate/fumarate are examples of compounds that serve as terminal electron acceptors under aerobic or anaerobic conditions, respectively. Conversely, fermentation is an energy-generating process that takes place anaerobically and in absence of an external electron acceptor. In fermentative metabolism, ATP is generated via substrate-level phosphorylation during the catabolic oxidation of the substrate. However, in order for this process to be sustained, the electrons generated in the catabolism of the substrate have to be consumed by a reductive process: this provides the basis for a coupling, which then enables redox balance/regeneration of NAD+ to keep the catabolic pathway (and hence substrate-level phosphorylation) functioning. The key issue in fermentative metabolism is therefore recycling of reduced equivalents to regenerate the oxidized form so that catabolism of the carbon source can proceed. Since the amount of reducing equivalents produced varies

according to the carbon and energy source, the composition of the mixture of fermentation products must also change accordingly (Ljungdahl, 2003). As an example, the phyla of Proteobacteria and Firmicutes, within the common domain of Bacteria, include fermentative organisms with remarkable differences in morphology, central metabolism and product composition of the fermentation mix. Based on the distinct mix of products synthesized out of the dissimilation of sugars, the fermentative metabolism mediated by Enterobacteria and Clostridia, belonging to the aforementioned phyla, is referred as mixed-acid and acetone-butanol-ethanol (ABE) fermentation, respectively (Ljungdahl, 2003) (refer to Section 3.1.2.2 and 3.1.2.3 for details).

3.1.2.1. Conversion of pyruvate to acetyl CoA under anaerobic conditions

The final product of the oxidation of any carbohydrate via glycolysis is pyruvate. Under anaerobic conditions and in the absence of external electron acceptors, pyruvate is the main intermediate metabolite for the synthesis of most fermentative products. Pyruvate is dissimilated through different pathways (Kessler, 1996; Lee et al., 2008b; Sawers and Clark, 2004)(Figure 3-3).

Lactic acid is produced by reducing pyruvate and occurs in a wide range of organisms when they grow under fermentative conditions (Figure 3-3), as aforementioned. In *Escherichia coli* two different enzymes mediate the conversion of pyruvate into acetyl-CoA: pyruvate formate lyase (PFL) and the pyruvate dehydrogenase complex (PDHC) (Böck and Sawers, 1996; Kessler, 1996) (Figure 3-3). PFL is known to be active under hypoxic or anoxic conditions, while it is widely accepted that PDHC mediates the aerobic dissimilation of pyruvate (Kessler, 1996). The PDHC role during fermentative glucose metabolism has

recently been elucidated as associated to the metabolic requirements of CO_2 (Murarka et al., 2010).

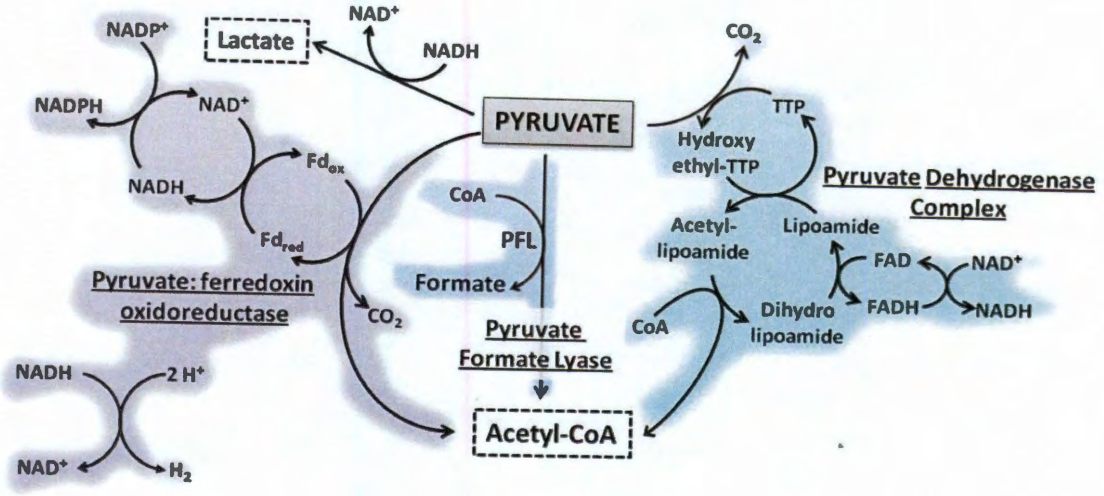

Figure 3-3 Metabolic pathways for anaerobic pyruvate dissimilation
Color shades correspond to different dissimilation routes: Pyruvate Dehydrogenase Complex (PDH, light blue); Pyruvate Formate Lyase (PFL, blue); Pyruvate: ferredoxin oxidoreductase (PFOR, purple). Abbreviations: TTP, Thiamine pyrophosphate; Fd_{red}, Reduced ferredoxin; Fd_{ox}, Oxidized ferredoxin.

A comparison of the PDHC and PFL reactions for the conversion of pyruvate into acetyl-CoA also reveals that the reaction catalyzed by PDHC produces NADH as opposed to the alternative pathway (catalyzed by PFL), which results in no net generation of reducing equivalents. An additional enzyme involved in the dissimilation of pyruvate in Enterobacteria is pyruvate oxidase (PoxB) (Kessler, 1996). PoxB catalyzes the decarboxylation of pyruvate to acetate and CO_2, a reaction that also generates reducing equivalents in the form of flavin adenine dinucleotide. The metabolism of pyruvate via PoxB is less efficient than the route via the PDH; however, the PoxB route is important for wild-type growth efficiency and responsible for a significant portion of pyruvate metabolism under aerobic conditions (Li et

al., 2007a). However, its metabolic role (if any) under fermentative conditions remains still uncertain.

An alternative route for anaerobic pyruvate dissimilation, present in several species within the Enterobacteriaceae family (Blaschkowski et al., 1982; Pomposiello and Demple, 2000; Wahl and Ormejohnson, 1987) and in Clostridia is the oxidative decarboxylation by pyruvate-ferrodoxin (Fd) oxidoreductase to form acetyl-CoA, CO_2, and reduced ferrodoxin (FdH_2) (Thauer et al., 1977) (Furdui and Ragsdale, 2000). Electrons produced from Fd can either be used in the reduction of protons to form H_2, or can be transferred to NAD(P) and thereby utilized for the production of alternative reduced fermentation products.

3.1.2.2. Mixed acid fermentation

Under anaerobic conditions and in the absence of alternative electron acceptors, members of the Enterobacteriaceae family (*Escherichia*, *Enterobacter*, *Salmonella*, *Klebsiella*, and *Shigella*) ferment sugars to a mixture of acetic, formic, lactic, and succinic acid, ethanol, molecular hydrogen and carbon dioxide, a fermentative metabolic mode known as mixed-acid fermentation (Sawers and Clark, 2004).

Hexoses such as glucose or fructose are dissimilated through the Embden-Meyerhof pathway to produce 2 NADH per C6 upon conversion to pyruvate. The fate of pyruvate is then either reduction to lactate by the action of lactate dehydrogenase, or cleavage to acetyl-CoA and formate by pyruvate:formate lyase, a key enzyme of mixed acid fermentation (Sawers and Clark, 2004), as described in the previous section.

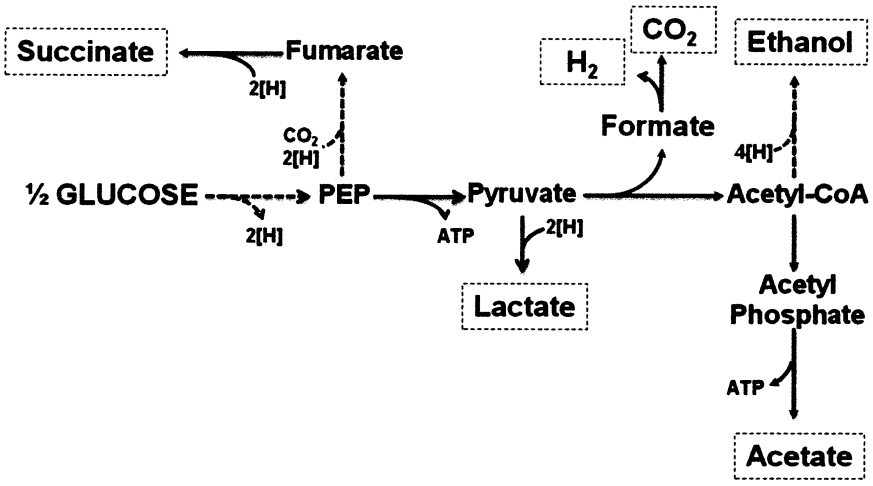

Figure 3-4 Mixed acid fermentation by *Escherichia coli*
Broken lines illustrate multiple steps. Products of the mixed-acid fermentation are boxed. Abbreviations: PEP, phosphoenolpyruvate; 2[H] = NADH = $FADH_2$ = H_2

Most commonly, part of the acetyl-CoA is cleaved to acetate via acetyl-P with generation of ATP. Concomitantly, and in order to attain redox balance, ethanol is produced from acetyl-CoA by means of a two-step process catalyzed by the bifunctional enzyme acetaldehyde/alcohol dehydrogenase that consumes two reducing equivalents (NADH). Another alternative is the conversion of pyruvate or posphoenolpyruvate (PEP) to a C-4 intermediate of the Krebs cycle via incorporation of the CO_2 produced by the reaction mediated by the formate: hydrogen lyase. The composition of the aforementioned compounds is therefore fine-tuned in order to produce energy in the form of ATP while preserving redox-balance (Sawers and Clark, 2004).

3.1.2.3. Acetone-butanol-ethanol (ABE) fermentation

Species such as *Clostridium acetobutylicum* belonging to the Clostridiaceae within the phylum of Firmicutes perform the so-called acetone-butanol-ethanol (ABE) fermentation,

a process divided into an acidogenic and a solventogenic phase (Jones and Woods, 1986; Lee et al., 2008b). At first butyrate is produced, thus leading to acidification of the medium. The low pH activates acetoacetate decarboxylase, which mediates the decarboxylation of acetoacetate to acetone while reducing the pool of acetoacetyl-CoA that can be reduced to butyryl-CoA. Taking up again the butyrate circumvents this redox imbalance. Butyrate is activated by CoA transfer and subsequently reduced to butanol via butyraldehyde, thus maintaining the redox balance (Lee et al., 2008b).

3.1.2.4. Homoethanolic fermentation

Many microorganisms, including bacteria and yeasts, can produce ethanol as the main fermentation product from carbohydrates (Figure 3-5). Glycolysis generates two moles of pyruvate per mole of glucose converted, along with two reducing equivalents and a net of two moles of ATP.

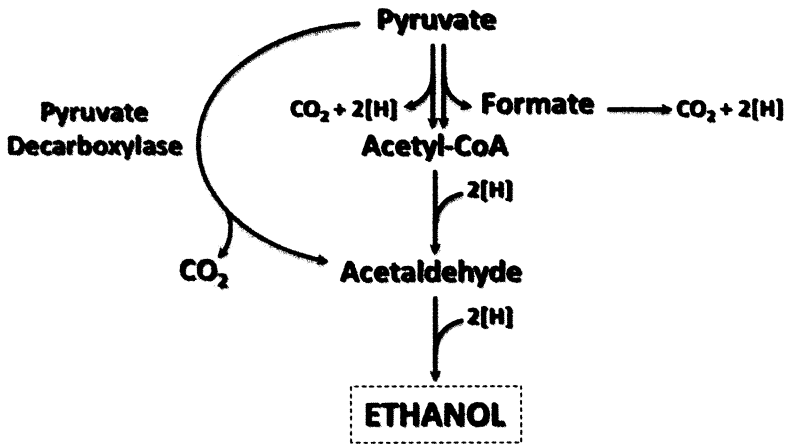

Figure 3-5 Pathway mediating the conversion of pyruvate to ethanol
The reaction reported in red is common to yeasts, plants, and bacteria (i.e., *Zymomonas mobilis*). Abbreviations: 2[H] = NADH = $FADH_2$ = H_2

The production of compounds more reduced than pyruvate (ethanol, lactate, etc.) serves as a mechanism to re-oxidize the reducing equivalents produced and it is therefore essential for glycolysis. In the only known homoethanol pathway that evolved in yeast, plants, and bacteria (i.e., *Z. mobilis*), pyruvate is decarboxylated to yield carbon dioxide and acetaldehyde by the nonoxidative pyruvate decarboxylase (Figure 3-5). The resulting acetaldehyde serves as the electron acceptor for NADH oxidation by the alcohol dehydrogenase during production of ethanol from pyruvate. Although a pyruvate decarboxylase-dependent pathway has been elucidated in *Zymomonas mobilis* (member of the Proteobacteria family) (Neale et al., 1987), a completely different ethanol fermentation pathway exists in Enterobacteriaceae (i.e. *E. coli*), in which pyruvate is converted to acetyl-CoA and formate by pyruvate formate-lyase. Acetyl-CoA is then converted to ethanol via a two-step pathway catalyzed by the bifunctional enzyme aldehyde-alcohol dehydrogenase that consumes two reducing equivalents (NADH) (Figure 3-5). Due to the requirement of two reducing equivalents per ethanol generated, the second acetyl-CoA from glycolysis of glucose is converted to acetate and an additional ATP. Thus, the native *E. coli* pathway for ethanol from acetyl-CoA cannot support homoethanol fermentation due to a redox imbalance.

3.1.2.5. Glycerol fermentation

The different characteristics of the anaerobic fermentative metabolism presented above for *Escherichia* sp. and *Clostridium* sp. seem to corroborate their significant phylogenetic distance as reported by Woese. Somewhat counterintuitively, however, substantial divergence in terms of metabolism can also be found within the same phylum. Enterobacteria of the genera *Klebsiella* (Homann et al., 1990) (Biebl et al., 1998),

Enterobacter (Barbirato et al., 1996), and *Citrobacter* (Homann et al., 1990) along with Lactobacilli (Schutz and Radler, 1984) and Clostridia (Forsberg, 1987) ferment glycerol to 1,3-propanediol (1,3-PDO). Glycerol fermentation branches out in two parallel pathways (Figure 3-6): oxidative and reductive (Booth, 2005; Yazdani and Gonzalez, 2007). The first steps of the oxidative pathway are glycerol dehydrogenation to dihydroxyacetone and its subsequent conversion to dihydroxyacetonephosphate. Succinic acid, acetic acid, lactic acid, 2,3-BDO and ethanol are then produced through the oxidative branch (Yazdani and Gonzalez, 2007)(Figure 3-6). The conversion of glycerol to 1,3-PDO is realized instead through the reductive branch and consists of a vitamin B12-mediated dehydration to 3-hydroxypropionaldehyde and its NADH-dependent reduction to 1,3-PDO (Biebl et al., 1999). The capability of fermenting glycerol had for long been thought peculiar of species able to synthesize 1,3-PDO. However, Gonzalez and coworkers have recently elucidated in *E. coli* an alternative mechanism of glycerol fermentation 1,3-PDO-independent (Gonzalez et al., 2008). Production of 1,2-PDO provides in this case an avenue to dispose of the excess reducing equivalents produced by the synthesis of biomass, thus facilitating redox balance (Gonzalez et al., 2008)(Figure 3-6). Concurrently, ethanol synthesis enables ATP-generation via substrate level phosphorylation by consuming the NADH generated in glycolysis, and therefore represents the major fermentation product.

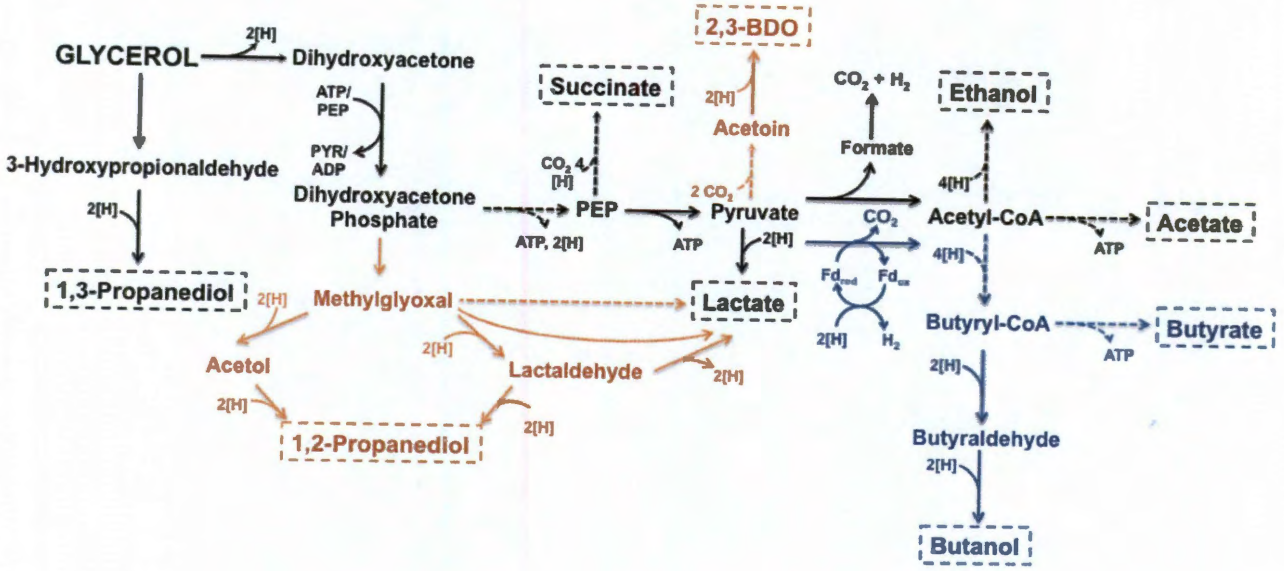

Figure 3-6 Biochemical pathways of anaerobic fermentative metabolism of glycerol
Reactions reported in orange and blue are specific to Enterobacteriaceae and Clostridia, respectively. Broken lines illustrate multiple steps. Abbreviations: PEP, phosphoenolpyruvate; PYR, pyruvate; Fd_{red}, Reduced ferredoxin; Fd_{ox}, Oxidized ferredoxin; 2[H] = NADH = $FADH_2$= H_2.

3.1.3. Functional differentiation of metabolic pathways

Previous examples underline a substantial divergence between molecular (genetic) and functional homology and evolution. Although the sequence-based phylogenetic (cladistic) classification proposed by Woese provides a good representation of genomes evolution, distribution of functions related to carbon and energy metabolism among different microorganisms does not necessarily follow the same classification pattern of rRNA (Pace, 1997). An intrinsic limitation of the phylogenetic analysis based on the comparison of the highly conserved small ssu-rRNA subunits consists, for example, in the fact that the

evolutionary history of any single gene may differ from the phylogenetic history of the whole organism from which the corresponding molecule was isolated. Furthermore, although comparative genomics based on orthologous genes has elucidated phylogenetic relationships among the known genomes, the mechanisms underlying their speciation remain elusive (Boussau and Daubin, 2010). Particularly, phylogenetics fails to account for the contribution of external forces such as horizontal gene transfer and pressure from environmental factors to genome evolution (Papp et al., 2009; Ragan and Beiko, 2009).

The broad variety of metabolic capabilities existent in present-day microorganisms can have emerged in several ways (Fani and Fondi, 2009). Enzymes mediating identical, parallel or reverse reactions may have for example arisen independently. Hence, the realization that molecular similarity is only a partial indicator of evolutionary functional relatedness opens the question of whether molecular divergence is ultimately correlated to functional divergence. So, the establishment of a metric alternative to sequence-homology for assessing functional similarities becomes impelling. The introduction of modern technologies of characterization of protein 3D structures like NMR and X-ray crystallography has elucidated that structures of protein domains are more conserved than their primary sequence (Jones, 2000). Functional homology becomes therefore tightly linked to structural domain homology (Loewenstein et al., 2009), with the domains evolving either from different ancestral proteins or from an evolutionary common ancestor (Buljan and Bateman, 2009). Protein domains can therefore be defined as functional and structural units of proteins interactions (Kummerfeld and Teichmann, 2009). Consequently, a number of novel trees (Figure 3-7) have been proposed which are based on the occurrence of this specific feature throughout the whole genome (Bro et al., 2006; Dagan et al., 2010; Forslund

et al., 2008; Fukami-Kobayashi et al., 2007; Gupta, 1998; Yang and Bourne, 2009).

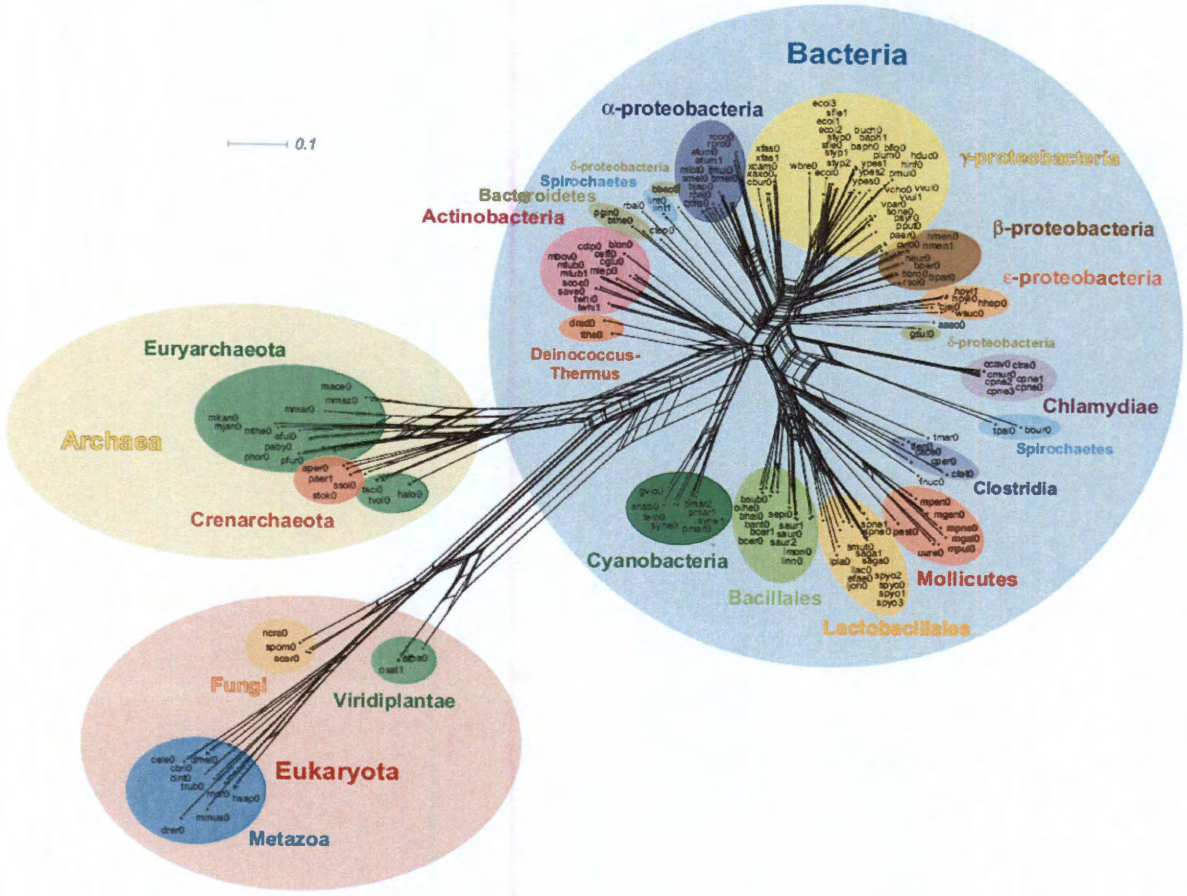

Figure 3-7 A network based on the domain organization repertoires.
The tree includes 17 archaeal, 136 bacterial, and 14 eukaryotic organisms (Fukami-Kobayashi et al., 2007). The organisms in the tree are color coded according to the NCBI taxonomy (Benson et al., 2011; Wheeler et al., 2008). The Pfam (Finn et al., 2010) domain organizations were used for constructing the domain organization repertoire of each organism.

Domains within protein structures are defined as structures spatially distinct that could conceivably fold and function independently (autonomously folding units) (Ponting and Russell, 2002). It has been fairly established that the number of naturally occurring protein folds is approximately of the order of 10^3-10^4 (Liu et al., 2004) and that the vast

majority of these folds is represented within the Protein Data Bank (Berman et al., 2003). The question remains open as to whether proteins functionally differentiated and converged toward a finite set of optimal conformations or whether we are depicting a snapshot of a protein structural ensemble in continuous evolution. The Class, Architecture, Topology, and Homologous (CATH) superfamily protein domain database (Orengo et al., 1997) along with the Structural Classification of Proteins (SCOP) database (Murzin et al., 1995) are currently available resources for the analysis of protein structures and their hierarchical classification families, superfamilies, folds and classes. Based on these available resources, Aguilar and coworkers recently proposed a series of trees based on the distribution of metabolic functions in the metabolome of microorganisms for which the whole genome sequence is known (Aguilar et al., 2004). The concept they propose aims at drawing relationships between species living in analogous or different niches; the topology of the trees for several metabolic classes represent the functional adaptation of pathways, through either gene mutation or horizontal transfer, upon selective pressure exerted by the habitat. In addition, the method developed by Aguilar and coworkers analyzes the sequential order of domains within the primary structures of proteins along with the domain distribution, thus providing more detailed aspects of evolution (Aguilar et al., 2004). Homology of protein domains can therefore suggest a functional evolutionary relationship, which remains detectable even when sequences have diverged beyond any recognizable similarity (Aguilar et al., 2004). Forslund and coworkers (Forslund et al., 2008) have further investigated protein evolution based on domain trees, analyzing, in particular, evolution of global protein architectures. This methodology, being based on non species-specific comparison of global protein architectures (Mulder, 2010), overcomes the limitations posed by previous methods of not being able to

account for horizontal gene transfer (HGT) (Forslund et al., 2008).

An alternative strategy to characterize bacterial functional evolution entails making use of current knowledge on prokaryotic genomes to derive phylogenies based on the whole genomic information of organisms rather than just a small number of genes (Henz et al., 2005). Phylogenomics, the science that lays at the intersection of evolution and genomics (Eisen and Fraser, 2003), utilizes the vast amount of genomic data available to generate more accurate functional predictions that consider the evolution of the *ensemble* rather than the individual genes. Phylogenomic trees have been constructed based on this concept and have partially covered the gap between phylogeny and functional evolution (Casci, 2011; Delsuc et al., 2005; Eisen and Fraser, 2003). Concomitantly, theoretical models involving networks evolution (Hirsh and Sharan, 2007; Yosef et al., 2009) have been developed given the availability of large-scale protein interaction (PIN) datasets (Fraser et al., 2002; Robertson and Lovell, 2009), with the scope of elucidating protein functions, predicting interactions and ultimately providing a more detailed evolutive profiling of metabolic networks.

3.1.4. Current metabolic engineering and synthetic biology approaches applied to the production of fuels and chemicals

While novel methodologies are developed to fill the knowledge gap between genetic and functional taxonomies, metabolic engineering and synthetic biology utilize the wealth of information generated in this process to construct optimal microbial cell factories to produce products of interests by combining favorable metabolic pathways from diverse genera. Not surprisingly, in fact, industrial microbiology has historically developed in demand to develop cheap and faster substitutes of chemical reactions and based on information provided by

philogenetics. Classical industrial microbiology, for example, has for long time aimed at developing industrial processes and synthesizing products of interest by exploiting the native refined metabolic capabilities of modern bacterial species, previously characterized by phenetics. In the early 20th century, for example, industrial fermentation processes harnessing the metabolism of Clostridia, as described in Section 3.1.2.3, were developed for the solvent production (acetone and n-butanol) (Awang et al., 1988). Clostridia are spore formers and obligate anaerobes that grow at slow rates, have complex nutritional requirements and produce a mixture of products, thus considerably decreasing product yields. In the attempt to address the limitations posed from an industrial perspective by such a complex metabolism, metabolic engineering strategies have been developed in the last few decades, which aimed at the improvement of n-butanol yield, titer, and productivity by means of genetic manipulation of Clostridial metabolism. However, these efforts were hindered by the lack of efficient genetic tools to manipulate Clostridia, along with their complex metabolism. In an effort to overcome these issues, the genes that enable the synthesis of n-butanol or acetone in native producers like Clostridia have been therefore imported into industrial organisms that are genetically and metabolically tractable such as *E. coli* (Atsumi et al., 2008a; Atsumi and Liao, 2008; Bermejo et al., 1998; Dellomonaco et al., 2010b; Inui et al., 2008), *Saccharomyces cerevisiae* (Steen et al., 2008), *Pseudomonas putida* (Nielsen et al., 2009), *Bacillus subtilis* (Nielsen et al., 2009), *Lactococcus lactis* (Nielsen et al., 2009), and *Lactobacillus* species (Berezina et al., 2010).

Similarly, the metabolism of glycerol in Clostridia and Enterobacteriaceae (Section 3.1.2.5) has been originally explored for the industrial production of 1,3-propanediol (Biebl et al., 1999). However, metabolic pathways from *Klebsiella*, *Citrobacter* or *Clostridia* have

recently been reconstructed in *E. coli* (Celinska, 2010; Laffend, 1997; Nakamura and Whited, 2003; Tong et al., 1991) or *S. cerevisiae* (Cameron et al., 1998; Celinska, 2010; Rao et al., 2008) because of their industrial amenability and efficient carbohydrate metabolism.

The strategies hereby reviewed all consist of importing and recombining specialized biosynthetic capabilities evolved in one *phyla* with other favorable traits of the bacterial host, an approach that we refer to as Heterologous Metabolic Engineering (HeME). HeME-based approaches have been used to engineer chemical and fuel production in the past and are currently viewed as the strategy of choice when, based on phylogenetic considerations, the host organism is not predicted to possess the desired metabolic function (Na et al., 2010; Alper and Stephanopoulos, 2009). Current metabolic engineering and synthetic biology approaches are therefore still based on high metabolic segregation models, therefore lagging behind when compared to more recent classification of organisms. As an alternative, the Homologous Metabolic Engineering (HoME) methodology that we propose builds on the different evolution patterns followed by genomes and functionomes (ensemble of metabolic capabilities of an organism) and it aims at exploiting the biosynthetic metabolic potential to synthesize non-native products via native pathways.

3.2. Methods

3.2.1. *In silico* analysis

All data reported are based upon sequence comparisons made by using the freely accessible genome data of *Escherichia coli* K-12 substr. MG1655 (http://ecocyc.org/ (Keseler et al., 2005)). Protein sequenced were searched through UniProT (www.uniprot.org

(Bairoch et al., 2005)) and subjected to similarity searches (BLAST) at the websites of the National Center for Biotechnology Information at the National Institutes of Health, Bethesda, Md., at www.ncbi.nlm.nih.gov, and of The Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom, at www.sanger.ac.uk/Projects/S_coelicolor. Default settings were used, however, without filtering low-complexity regions. Sequence alignments were conducted with CLUSTALW, applying predefined algorithms at www.ebi.ac.uk/clustalw of The European Bioinformatics Institute at The European Molecular Biology Laboratory (Thompson and He, 2006).

The three-dimensional structure and functionality of *E. coli* genes was predicted using homology models generated by I-TASSER, a free online resource available at http://zhang.bioinformatics.ku.edu/I-TASSER (Roy et al., 2010).

Protein family signatures were obtained from the EMBL European Bioinformatics Institute (EBI) at http://www.ebi.ac.uk/Tools/pfa/iprscan/ (Quevillon et al., 2005) and from the European molecular Biology Laboratory (EMBL) at smart.embl-heidelberg.de (Schultz et al., 1998).

3.2.2. Strains, plasmids and genetic methods

Strains, plasmids, and primers used in this study are listed in Table 3-1. Wild-type K12 *E. coli* strain MG1655 (Kang et al., 2004a) was used as the host to implement metabolic engineering strategies. Gene knockouts were introduced in MG1655 and its derivatives by P1 phage transduction as described elsewhere (Miller, 1992), using as donors single gene knock out mutants from the National BioResource Project (NIG, Japan) (Baba et al., 2006). Deletion mutants were created using a single-step gene inactivation method developed by

Datsenko and Wanner (Baba et al., 2006) which employs a gene disruption strategy. In the method a PCR insert is first created from a pHD4 template and primers with flanking regions homologous to the start and end sequences of the gene to be inactivated. This results in an insert with a kanamycin resistance gene in the center of the two FRT (FLP recognition sites) and outer homology regions. Wild type cells expressing λ Red recombinase were then transformed with this PCR insert, with recombination resulting in the gene sequence between the homology regions being replaced with the FRT::Km::FRT sequence. Subsequent plating on LB with kanamycin plates allowed for selecting of positive mutants. The Keio Collection mutants were obtained in this state, with the mutant expressing kanamycin resistance. When required, the elimination of the kanamycin resistance gene was accomplished by transforming the cells with pCP20, a temperature sensitive plasmid expressing FLP. FLP expressed from this plasmid cut the Km region from the FRT::Km::FRT site, leaving behind one FRT site. pCP20 was then removed by growing the cells at 42 ◦C. This simple procedure enabled the construction of numerous gene deletion mutants from wild type cells by simply varying the homologous regions.

Strain *fadR atoC(c)*, which exhibits constitutive expression of the *fad* regulon (due to *fadR*) and *ato* operon (due to *atoC(c)*), was obtained as follows. Strain MG1655 *fadR* was isolated as an spontaneous mutant able to grow on MOPS (morpholino-propanesulfonic acid) minimal medium (Neidhart et al., 1974) plates containing 0.2% (w/v) decanoic acid ($C_{10:0}$) as the sole carbon source. Strain *fadR atoC(c)* was then obtained by transducing *fadR* with a phage lysate from strain LS5218, which is a *fadR atoC(c)* mutant (Jenkins and Nunn, 1987a). Successful transduction of the *atoC(c)* mutation into strain *fadR* was identified by growth on MOPS minimal medium plates containing 0.2% (w/v) hexanoic acid ($C_{6:0}$).

Table 3-1 Strains, plasmids, and primers used in this study

Strain/plasmid Primer	Description/Genotype/ Sequence	Source
Strains		
MG1655	F- l- ilvG- rfb-50 rph-1	(Kang et al., 2004b)
ΔadhE	MG1655, *ΔadhE*::FRT-kan-FRT	This study
ΔadhE ΔackA-pta ΔldhA	MG1655, *ΔadhE*::FRT *ΔackA.pta*::FRT *ΔldhA*-FRT-kan-FRT	This study
ΔadhE ΔackA-pta ΔldhA fadR	MG1655 *fadR*, *ΔadhE*::FRT *ΔackA.pta*::FRT *ΔldhA*-FRT-kan-FRT	This study
fadR atoC(c) ΔadhE	MG1655 *fadR*, *atoC(c) ΔadhE*-FRT-kan-FRT	This study
Plasmids		
pTH.atoDA	*E. coli atoD* and *atoA* operons under control of Ptrc (AmpR, oriR pBR322)	This study
pTH.pflCD.yiaY	*E. coli pflCD* and *yiaY* operons under control of Ptrc (AmpR, oriR pBR322)	This study
pTH.pflEF.yiaY	*E. coli pflEF* and *yiaY* operons under control of Ptrc (AmpR, oriR pBR322)	This study
pTH.eutD.ackA	*E. coli eutD* and *ackA* genes under control of Ptrc (AmpR, oriR pBR322)	This study
pTH.eutD.tdcD	*E. coli eutD* and *tdcD* genes under control of Ptrc (AmpR, oriR pBR322)	This study
pTH.atoB	*E. coli atoB gene* under control of Ptrc (AmpR, oriR pBR322)	This study
pTHyqeF	*E. coli yqeF gene* under control of Ptrc (AmpR, oriR pBR322)	This study
pZS.fucO	*E. coli fucO* gene under control of PLtetO-1 (tetR, oriR SC101*, cat)	This study
pZS.yqeF	*E. coli yqeF* gene under control of PLtetO-1 (tetR, oriR SC101*, cat)	This study
pZS.KLM.gldA	*E. coli dhaKLM* and *gldA* operons under control of PLtetO-1 (tetR, oriR SC101*, cat)	This study

Standard recombinant DNA procedures were used for gene cloning, plasmid isolation, and electroporation. Manufacturer protocols and standard methods (Miller, 1992; Sambrook and Russell, 2001) were followed for DNA purification (Qiagen, CA, USA), restriction endonuclease digestion (New England Biolabs, MA, USA), and DNA amplification (Stratagene, CA, USA and Invitrogen, CA, USA). The strains were kept in 32.5% glycerol stocks at -80◦ C.

Plates were prepared using Luria-Bertani (LB) medium containing 1.5% agar and appropriate antibiotics were included at the following concentrations: 100 μg/mL ampicillin, 50 μg/mL kanamycin, 34 μg/mL chloramphenicol and 12.5 μg/mL tetracycline. All strains created were confirmed by polymerase chain reaction and appropriate phenotypic tests, if suitable.

3.2.3. Culture medium and cultivation conditions

All chemicals for culture media were obtained from Fisher Scientific (Pittsburgh, PA) and Sigma-Aldrich Co. (St. Louis, MO), unless otherwise stated. MOPS minimal medium (Neidhart et al., 1974) supplemented with either glucose or glycerol (Fluka Chemie AG, Buchs, Switzerland) was used. When required the medium was supplemented with 100 μg/mL ampicillin, 50 μg/mL kanamycin, 34 μg/mL chloramphenicol and 12.5 μg/mL tetracycline. Isopropyl-beta-D-thiogalactopyranoside (IPTG, 0.1 mM) and anhydrotetracycline (100 ng/ml) were used to induce gene expression from constructed plasmids. Prior to use, cultures (stored as glycerol stocks at -80 ◦C) were streaked onto LB plates (with appropriate antibiotics if required) and incubated overnight at 37 ◦C.

In the case of butanol, acetone and butyric acid five single colonies were used to inoculate 30 mL of MOPS minimal medium supplemented with 2% of carbon source (w/v) in 50 ml baffled shake flasks (Corning Glass Works, Corning, NY), which were incubated at 37 ◦C until an OD_{550} of ~0.5 was reached. An appropriate volume of this actively growing pre-culture was centrifuged and the pellet washed and used to inoculate 30 mL of medium in 50 ml shake flasks to an initial OD_{550} of 0.05. Fifty-milliliter shake flasks (Corning Glass Works, Corning, NY) with four baffles and plastic foam-stoppers were used for microaerobic

cultures. Flasks were incubated for 48 hours at 37 °C in a C24 Rotary Incubator Shaker (New Brunswick Scientific Co., New Brunswick, NJ). In the case of acetone, higher cell density in-flask fermentations were inoculated starting from precultures prepared as described above. When required, acetic acid at a concentration of 1 gl^{-1} was added to the medium after 12 hours of cultivation.

Experiments for the synthesis of 1,3-propanediol were conducted in 17 ml hungate tubes (Bellco Glass, Inc., Vineland, NJ). A single colony was used to inoculate 10 ml of MOPS minimal medium supplemented with 1% of glycerol in 17 ml hungate tubes (Bellco Glass, Inc., Vineland, NJ), which were incubated at 37 °C for 48 hours. In the case of anaerobic fermentations, a single colony was used to inoculate 10 mL of medium in each tube, which were then incubated (37°C) for 48 h in an Isotemp Incubator (Fisher Scientific, Pittsburgh, PA) and sparged with ultra-high purity argon (Matheson Tri-Gas, Houston, TX). To maintain sterile conditions, the inlet gas was passed through a 0.2 mm HEPA filter (Millipore, Billerica, CA) and the outlet gas line immersed in a 1 M CuSO4 solution. When appropriate, the media was supplemented with coenzyme B_{12} at a concentration of 5 mg l^{-1}.

Samples were centrifuged to pellet cells while the aqueous supernatant was collected for metabolite analysis.

3.2.4. Analytical methods

Cell growth was monitored by measuring the absorbance at a wavelength of 550 nm.

The identity of all metabolic products was determined through 1D proton NMR spectroscopy as previously described (Murarka et al., 2008). Organic acids, ethanol, acetone and 1,3-propanediol isopropanol were quantified by HPLC as previously reported (Dharmadi

and Gonzalez, 2005). Samples (culture supernatant) were analyzed with ion-exclusion HPLC using a Shimadzu Prominence SIL 20 system (Shimadzu Scientific Instruments, Inc., Columbia, MD) equipped with an HPX-87H organic acid column (Bio-Rad, Hercules, CA) with operating conditions to optimize peak separation (0.3 mL/min flowrate, 30 mM H_2SO_4 mobile phase, column temperature 42 °C) (Dharmadi and Gonzalez, 2005).

Quantification of butanol was conducted through gas chromatography (GC) in a Varian CP-3800 gas chromatograph (Varian Associates, Inc., Palo Alto, CA) equipped with a flame ionization detector (GC-FID). 1ml of supernatant of culture broth was extracted with 500 µl of GC standard grade n-hexane (Fluka, Steinheim, Germany) and 1-heptanol (Sigma-Aldrich, Lausanne, Switzerland) was used as internal standard. The separation of alcohol compounds was carried out using a HP-INNOWax (30 m, 0.32 mm internal diameter, 0.50-µm film thickness; Agilent Technologies, Palo Alto, CA). The oven temperature was initially held at 40°C for 1 min and then raised with a gradient of 30°C/min to 130°C and held for 4 min. The temperature was then raised with a gradient of 15°C/min to 230°C and held for 4 min. Helium (1 ml min^{-1}, Matheson Tri-Gas, Longmont, CO) was used as the carrier gas. The injector and detector were maintained at 250°C. A 1.0-µl sample was injected in splitless injection mode.

For butyrate quantification, 5 mL of culture was placed in a 15 mL Corning CentriStar centrifuge tube (Corning Inc., Lowell, MA) in an ice bath, and then centrifuged at 8,000 g for 5 minutes and the appropriate volume of supernatant recovered. 2 ml of supernatant were acidified with 200 µl of acetic acid, supplemented with 100 mg of octanoic acid as an internal standard, and extracted with 2 ml of n-hexane-chlorophorm (4:1, v/v) (Lalman and Bagley, 2004). The organic layer was nitrogen evaporated to near dryness,

redissolved in 1 ml of a mixture of cholophorm:methanol:hydrochloric acid [10:1:1, vol/vol/vol] (Dellomonaco et al., 2010b) and incubated in a sealed vial at 90°C for 1 h. Fatty acid methyl esters were extracted with 1 ml of n-hexane-chlorophorm (4:1, v/v) after addition of 500 µl of 0.9% (wt/vol) NaCl in water. Gas chromatography was performed with 1µl of the n-hexane:chlorophorm solution on a HP-INNOWax (30 m, 0.32 mm internal diameter, 0.50-µm film thickness; Agilent Technologies, Palo Alto, CA) according to the following method: 50°C held for 1 min, 30°C/min to 160°C, 15°C/min to 200°C, 200°C held for 1.5 min, 10°C/min to 225°C, and 225°C held for 15 min (Dellomonaco et al., 2010b).

3.3. Results

3.3.1. Proposed approach and methodology

The Heterologous Metabolic Engineering (HeME) approach is currently the *de facto* strategy implemented to equip organisms with metabolic capabilities for the synthesis of non-native products. The HeME methodology builds upon the differences in metabolism among bacteria belonging to different phyla, and therefore assumes that evolutionary distant species from a genomic standpoint will utilize a different toolset of biosynthetic capabilities. By doing so, HeME aims at combining favorable traits of evolutionary distant microbial species and thus fails at fully exploiting the biosynthetic capabilities of individual bacteria. Moreover, the compatibility of the heterologous pathway with the host metabolism might represent a concern. As an alternative approach, the Homologous Metabolic Engineering (HoME) strategy hereby proposed overcomes this limitation by investigating and harnessing functional relatedness of evolutionary distant species applied in this context to the production

of industrially relevant molecules. The homologous functionality of molecular evolutionary distant proteins encoded within different microbial genomes identified by the presence of common protein domains/folds suggests the existence of a common functional ancestor. Hence, adaptation of ancestral microbial metabolism and the consequent bacterial speciation to endure specific niches is very likely to have evolved from an un-regulated common genome encoding broad-specificity enzymes. These broad biochemical capabilities would have provided the metabolic plasticity essential for the successful adaptation to different niches.

Consequently, we hypothesize that along with the differentiation of species-specific metabolic pathways, modern bacteria have retained ancestral enzymes and pathways with lower affinities for "nonphysiological" substrates, and that these metabolic capabilities are effectively masked by the regulatory mechanisms that have evolved (Caetano-Anolles et al., 2009). Based on this and in analogy with embryonic stem cell differentiation processes in higher organisms, we propose the existence of a functionally pluripotent metabolic bacterial cell able to differentiate into a variety of specialized metabolic forms upon gene recruitment and relief of specific regulations.

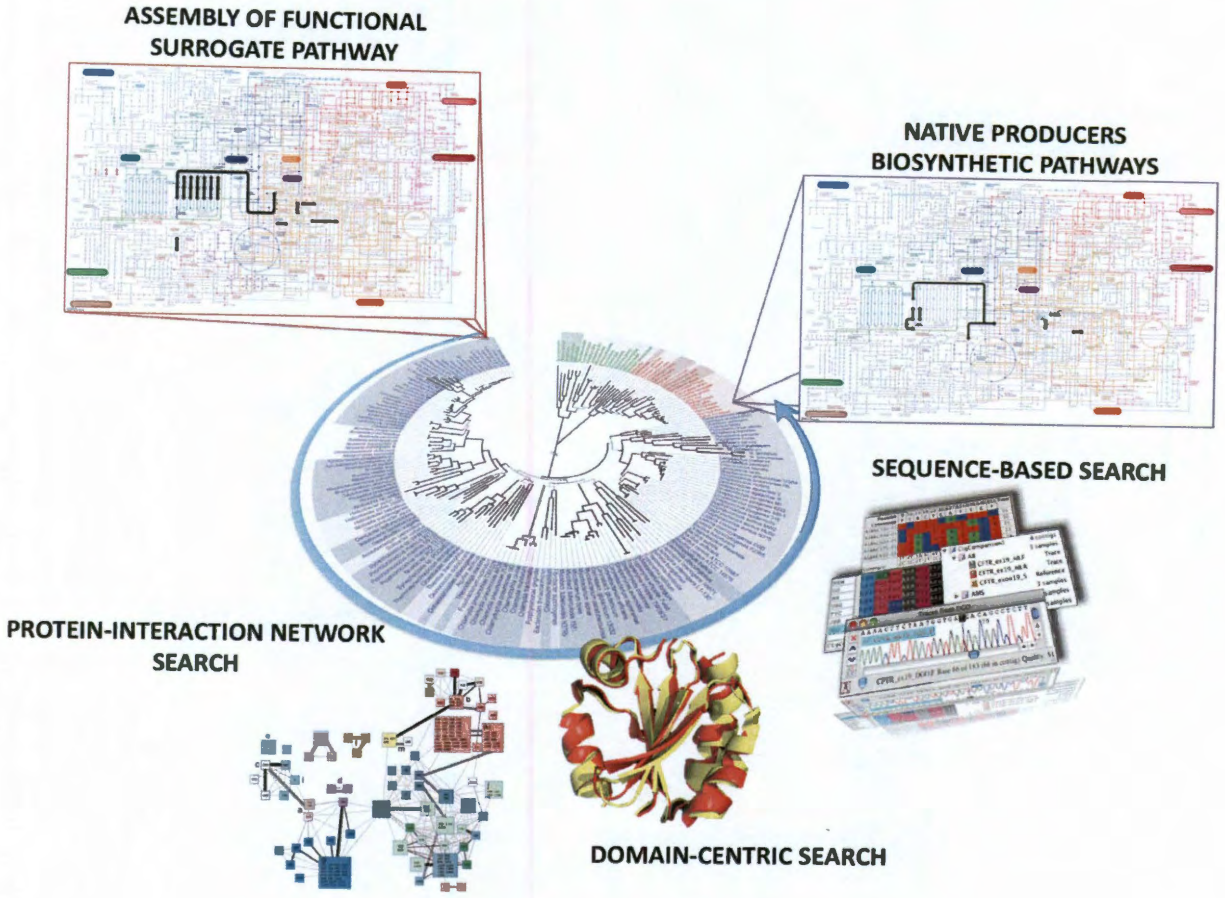

Figure 3-8 Workflow of the Homologous Metabolic Engineering (HoME) approach
Enzymes mediating biosynthetic pathways of the product of interest in native producers are identified and functional analogues are *homed* in the host organism by *i*. sequence-based search; *ii*. Protein domain-centric search; and *iii*. Protein-interaction network search. A functional surrogate pathway is reconstructed in the host strain, that now functions like the native producer. The blue arrow in the figure indicates re-tuning of the host metabolism.

Accordingly, we propose an alternative strategy to overcome the limitations posed by HeME that entails the recruitment and functional activation of native genes/proteins and their assembly into superior metabolic routes for the biosynthesis of products of interest, e.g. fuels and chemicals. Since no exogenous gene is recruited to establish the otherwise foreign pathway, we have termed this approach Homologous Metabolic Engineering (HoME).

3.3.1.1. Methodology

The proposed strategy aims at reconstructing heterologous pathways by *homing* and *assembling* native functional surrogates. *homing*: *i.* Protein-sequence comparative analysis (PSCA) using as query the protein known to efficiently mediate the reaction of interest against a library of protein sequences from the bacteria of choice by means of Basic Local Alignment Search Tool (pBLAST); *ii.* Domain-centric homology search of the aforementioned query using SMART (Quevillon et al., 2005; Schultz et al., 1998) and InterProScan (Quevillon et al., 2005) to identify protein similarities across significant evolutionary distances using sensitive protein domain profiles. Potential candidates are identified by appropriate homology metrics among the best-scoring alignments obtained via *i.* and *ii.* analyzing them in the context of primary literature and information from genomic and metabolic databases. *functional assembly*: Enzymes identified via *homing* will be assembled expressed into functional pathways in wild-type bacterial hosts, e.g. *E. coli* by directed gene expression. Since the genes proposed as surrogates might also be subjected to several levels of regulation, functional expression could also entail manipulating the corresponding regulators and blocking competing pathways for common metabolic intermediates. The effect of the expression levels of each enzyme in the pathway is also considered by using different promoters and RBSs and by evaluating polycistronic and individual overexpression of each gene/operon (i.e. individual promoters and RBSs).

In an effort to lay the foundations of HoME, we have already demonstrated the viability of engineering and redirecting the functional differentiation of bacteria by harnessing native biosynthetic capabilities. Clostridial ABE fermentation, glycerol fermentation to 1,3-PDO, and the yeast homoethanol fermentation were selected as case

studies as they represent metabolic capabilities considered unique to organisms that are genetically and metabolically distant from *E. coli*. For all three cases, HeME approaches have already been implemented to equip *E. coli* with the biosynthetic capability of interest (Atsumi et al., 2008a; Celinska, 2010; Dellomonaco et al., 2010b; Inui et al., 2008; Laffend, 1997; Nakamura and Whited, 2003; Nielsen et al., 2009; Shen and Liao, 2008; Tong et al., 1991).

3.3.2. Reconstruction of Acetone-Butanol-Ethanol (ABE) fermentative metabolism in *Escherichia coli* using native pathways

Industrial microbial fermentations exploiting the biphasic metabolism of Clostridia (as described in Section 3.1.2.3) have been among the first large scale bioprocesses to be developed for the production of solvents, i.e. n-butanol and acetone (Jones and Woods, 1986), and butyric acid. Despite the ability of Clostridial metabolism to utilize a broad range of carbon sources (monosaccharides, disaccharides, starches, inulin, pectin, whey, and xylan) (Nolling et al., 2001), limitations to their industrial use are posed by the tight coupling of solventogenesis to regulation of endospore formation (Durre, 2005) and the low product yields due to the inevitable formation of byproducts. Although several metabolic engineering efforts have been directed to engineering Clostridial metabolism to overcome these shortcomings, poor understanding of Clostridial differentiation processes and solventogenesis has ignited interest in the development of strains that would combine genetic tractability with efficient synthesis of solvents by means of Heterologous Metabolic Engineering (HeME) (Atsumi et al., 2008a; Berezina et al., 2010; Dellomonaco et al., 2010b; Fischer et al., 2010; Nielsen et al., 2009).

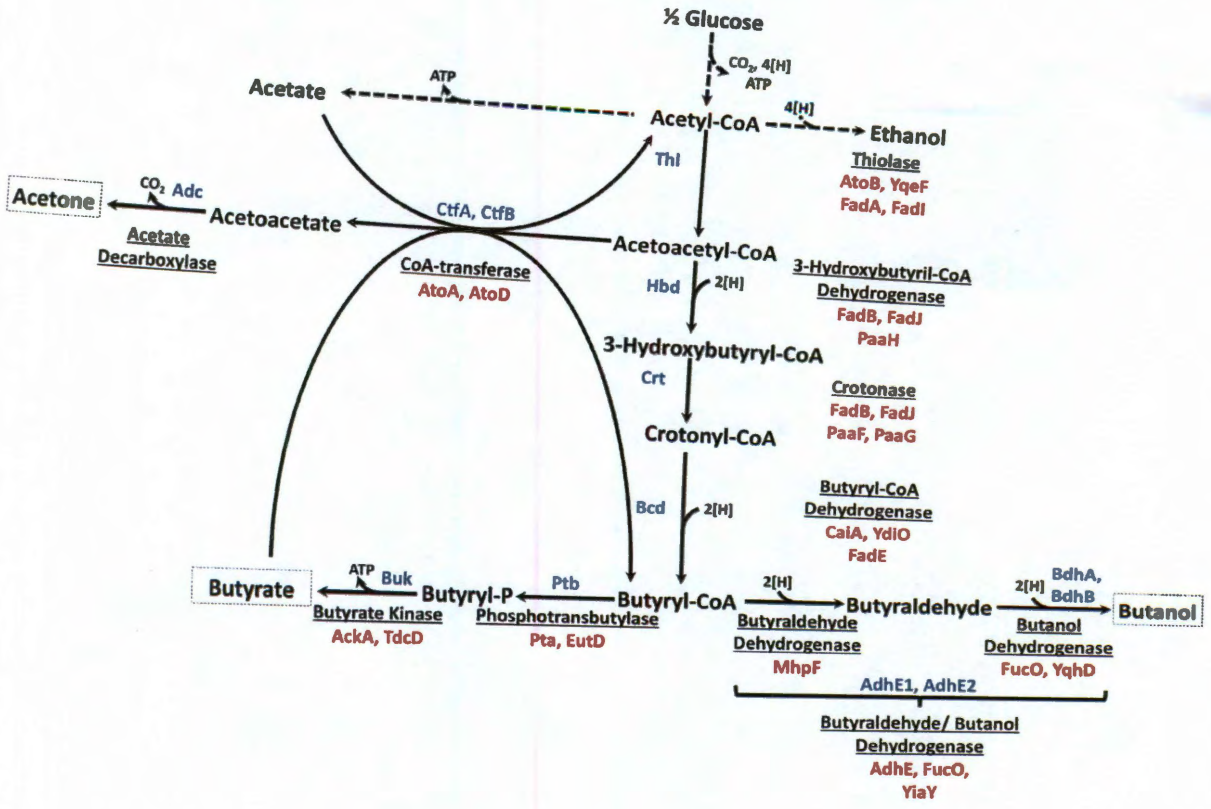

Figure 3-9 Acetone-Butanol-Ethanol (ABE) fermentative metabolism reconstructed with *Escherichia coli* native genes

Native Clostridial enzymes are reported in blue; *E. coli* set of functional surrogate enzymes are reported in red. Boxes indicate non-native *E. coli* products. Abbreviations: 2[H] = NADH = $FADH_2$ = H_2.

We selected *Escherichia coli* as organism of choice and implemented the Homologous Metabolic Engineering (HoME) methodology described above to engineer functional differentiation in *E. coli* for the synthesis of products typical of the Clostridial acetone-butanol-ethanol (ABE) fermentative metabolism, without recruiting any foreign genes.

3.3.2.1. n-butanol

n-butanol biosynthesis takes place in a six-step conversion starting from acetyl-CoA (Figure 3-9) (Jones and Woods, 1986; Lee et al., 2008b). A thiolase, encoded by *thl*, catalyzes the condensation of two molecules of acetyl-CoA to acetoacetyl-CoA (Stimherndon et al., 1995), which is subsequently converted to butyryl-CoA by three enzymes, products of the BCS operon: *crt*, *bcd*, *etfAB* and *hbd* coding respectively for crotonase, butyryl-CoA dehydrogenase, electron transfer proteins, and 3-hydroxybutyryl-CoA dehydrogenase. Lastly, butyryl-CoA is converted to n-butanol in a two-step process catalyzed by a bi-functional aldehyde/alcohol dehydrogenase (*adhE1*, *adhE2*). There is evidence that AdhE1 can also mediate the conversion of Acetyl-CoA to ethanol. The conversion of butyraldehyde to n-butanol is also catalyzed with high affinity and specificity by BdhA and BdhB, two different n-butanol dehydrogenases present in *C. acetobutylicum* (Welch et al., 1989).

Butanol and butyrate are both synthesized via a two-step reaction from butyryl-CoA, as shown in Figure 3-9. A set of *E. coli* enzymes that can act as surrogates of the Clostridial pathway mediating the conversion of acetyl-CoA to the common intermediate butyryl-CoA (Figure 3-10b) was identified via sequence (Zhang et al., 2010), structure (Altschul et al., 1997) (**Table 3-2**) and domain-centric (Quevillon et al., 2005; Schultz et al., 1998) (Table 3-3) comparisons between the Clostridial enzymes (reported in blue in 10) and those encoded in the *E. coli* genome, along with an analysis of the primary literature and information from genomic and metabolic databases. Five *E. coli* thiolases (AtoB, YqeF, FadA, FadI and PaaJ) exhibited sequence and structure similarity to the *C. acetobutylicum* acetyl-CoA acetyltransferase (ThlA) (Figure 3-10, **Table 3-2**, Table 3-3). Among all identified surrogates, AtoB and YqeF exhibited the highest similarity to the *C. acetobutylicum* acetyl-

CoA acetyltransferase, are classified as biosynthetic acetoacetyl-CoA thiolases (Haapalainen et al., 2006) and have higher specificity for short-chain acyl-CoA molecules (Feigenbaum and Schulz, 1975; Jenkins and Nunn, 1987a). It is noteworthy that YqeF is a predicted acyltransferase of unknown function.

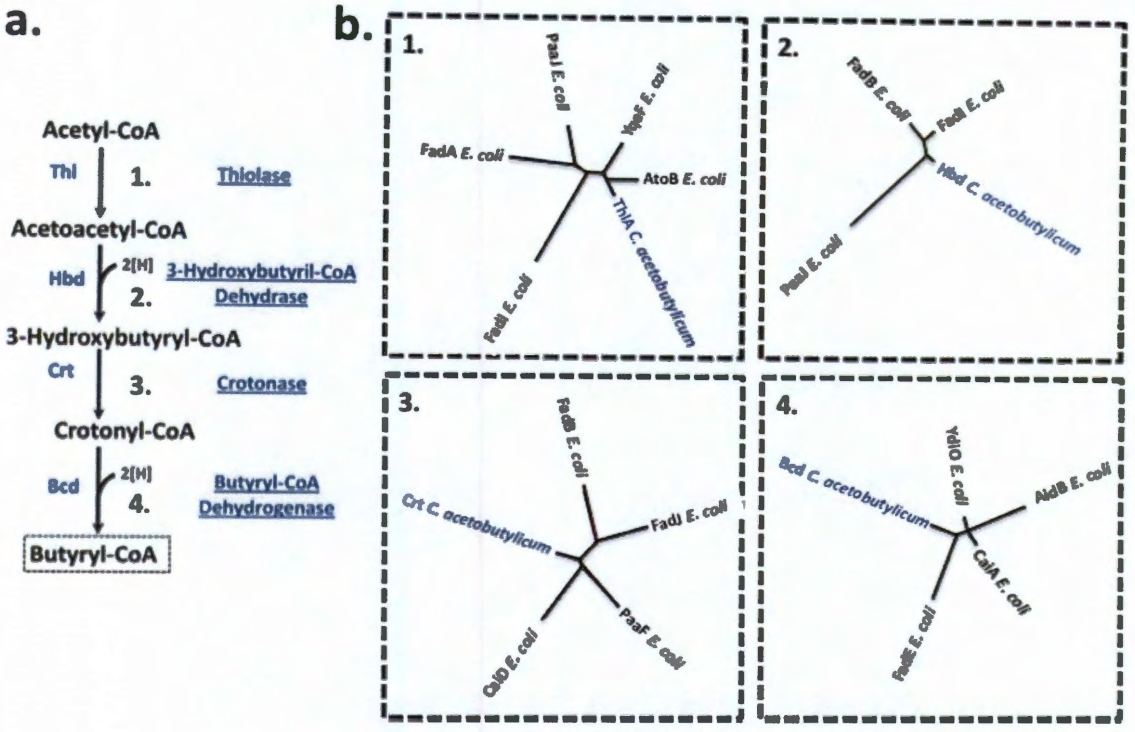

Figure 3-10 Identification of *E. coli* functional surrogates mediating the conversion of acetyl-CoA to butyryl-CoA.
a. Clostridial enzymatic steps (1. through 4.) mediating the conversion of acetyl-CoA into butyryl-CoA. Clostridial enzymes are reported in blue. Abbreviations: 2[H] = NADH = $FADH_2$ = H_2.
b. *E. coli* functional surrogates identified through sequence-, structure-, and domain-centric homology search are reported for each step, as indicated in panel a.. The length of the segments is function of the degree of homology between the query and the potential functional surrogates identified in *E. coli*, calculated as average of the E-values reported in Table 3-2 and Table 3-3.

Three *E. coli* enzymes were identified as potential surrogates of *C. acetobutylicum* 3-hydroxybutyryl-CoA dehydrogenase (Hbd) (Figure 3-10b, **Table 3-2**, Table 3-3): two hydroxyacyl-CoA dehydrogenases (FadB and FadJ), which are involved in the β-oxidation of

fatty acids (Campbell et al., 2003; Clark and Cronan, 2005), and a 3-hydroxyadipyl-CoA dehydrogenase (PaaH) that participates in the degradation of phenylacetate (Teufel et al., 2010). While involved in degradative pathways, these three enzymes catalyze reversible reactions and hence were considered viable substitutes of the Clostridial enzyme.

Four members of the enoyl-CoA hydratase family were chosen among the identified surrogates for the Clostridial 3-hydroxybutyryl-CoA dehydratase (Crt), given their high structural and sequence similarities (Figure 3-10, **Table 3-2**, Table 3-3): aerobic (FadB) and anaerobic (FadJ) enoyl-CoA hydratases that participate in the β-oxidation of fatty acids (Clark and Cronan, 2005), a 2,3-dehydroadipyl-CoA hydratase (PaaF) involved in pheylacetate degradation (Teufel et al., 2010), and a crotonobetainyl-CoA hydratase (CaiD) involved in carnitine degradation (Elssner et al., 2001). Along with FadA, FadB forms the fatty acid oxidation complex that also exhibits 3-hydroxyacyl-CoA epimerase and 3-hydroxyacyl-CoA dehydrogenase activities (Dirusso, 1990; Yang et al., 1988). FadJ (YfcX) is a paralog of FadB that bears 3-hydroxyacyl-CoA dehydrogenase activity and has been found strongly involved in the anaerobic degradation of long and medium-chain fatty acids in the presence of nitrate (Snell et al., 2002), and weakly involved in the aerobic degradation of long-chain fatty acids (Campbell et al., 2003). Despite the primary role of these enzymes in degradative pathways, the reversible nature of this reaction indicates that they can catalyze the dehydration of hydroxyacyl-CoAs.

Four potential surrogates were identified for the enzyme catalyzing the fourth step of the n-butanol synthesis pathway (butyryl-CoA dehydrogenase) (Figure 3-10, **Table 3-2**, Table 3-3): an acyl-CoA dehydrogenase (FadE), which is involved in the catabolism of fatty

acids (Campbell and Cronan, 2002; Clark and Cronan, 2005), a crotonobetainyl-CoA reductase (CaiA) that mediates carnitine degradation (Elssner et al., 2001), an isovaleryl-CoA dehydrogenase (AidB) involved in the adaptive DNA-repair response of *E. coli* to alkylating agents (Rohankhedkar et al., 2006), and a predicted acyl-CoA dehydrogenase (YdiO). The butyryl-CoA complex of Clostridia is associated to an electron transfer protein consisting of two subunits (EtfAB) and responsible for coupling the reduction of ferredoxin (Fd) to the reduction of crotonyl-CoA to butyryl-CoA (Bennett and Rudolph, 1995). While the sequence-based search did not return any match, the two subunits of the electron transfer protein, EtfA and EtfB, share homologous domains (Table 3-3) with YdiR and YdiQ, predicted to be subunits of a flavoprotein (Campbell et al., 2003).

The last two steps in the n-butanol pathway involve the sequential reduction of butyryl-CoA to butyraldehyde to n-butanol, which in Clostridia are catalyzed by butyraldehyde and butanol dehydrogenases (BdhA, BdhB, Adh1, Adh2) (Figure 3-9) (Lee et al., 2008b). The sequence and structural analysis suggested twelve potential *E. coli* surrogates: one bifunctional aldehyde/alcohol dehydrogenase (AdhE), two aldehyde dehydrogenases (MhpF and GapA) and five alcohol dehydrogenases (FucO, YiaY, EutG, YqhD, YbdH). (**Table 3-2**, Table 3-3). Among them, we selected FucO due to its higher structure and sequence similarity to the Clostridial enzymes and its kinetic properties (Sophos and Vasiliou, 2003), which in turn should favor the synthesis of n-butanol.

Many of the enzymes proposed above as *E. coli* functional surrogates for the first four steps of the Clostridial n-butanol pathway are involved in catabolic functions that mediate the degradation of fatty acids (FadB, FadJ, FadE, AtoB) (Clark and Cronan, 2005), phenylacetate

(PaaH, PaaF) (Teufel et al., 2010), and carnitine (CaiA, CaiD) (Eichler et al., 1994) at their expression is regulated both locally and globally (Figure 3-10).

These reactions can mediate the conversion of acetyl-CoA to butyryl-CoA. Since the same set of CoA thioesters intermediates are used in both the Clostridial pathway and the *E. coli* fatty acid degradation pathway, catalyzed by the products of the *fad* operon, we exploited the fatty acid catabolic route to generate butyryl-CoA, which can be acted upon by appropriate aldehyde/alcohol dehydrogenases to generate n-butanol. In an attempt to promote the conversion of acetyl-CoA to n-butanol exploiting the aforementioned steps we diverted carbon to butanol synthesis at the acetyl-CoA node by disrupting the *E. coli adhE* gene. To investigate further competition for the acetyl-CoA we also overexpressed *atoB* in the strain MG1655 Δ*adhE* but observed no butanol biosynthesis (Figure 3-11). However, expression of genes encoded within the *fad* operon is subjected to local regulation exerted by FadR (Section 2.1.1). In order to attain constitutive expression of the fatty acid catabolic pathway mutations were introduced in the FadR regulator (Dellomonaco et al., 2010b); pathways leading to synthesis of byproducts (ethanol, acetate, and lactate) were also blocked in strain MG1655 *fadR*. Butanol biosynthesis up to 277 mg l^{-1} (Figure 3-11) was observed when the *E. coli* putative acetyl-CoA acetyltransferase YqeF was overexpressed in combination with the predicted alcohol dehydrogenase FucO in strain MG1655 Δ*adhE* Δ*ackA.pta* Δ*frdA* Δ*ldhA* *fadR*. However, overexpression of FucO also contributed to ethanol biosynthesis up to 150 mg l^{-1} (Figure 3-11), which was the main byproduct along with pyruvic acid (210 mg l^{-1})

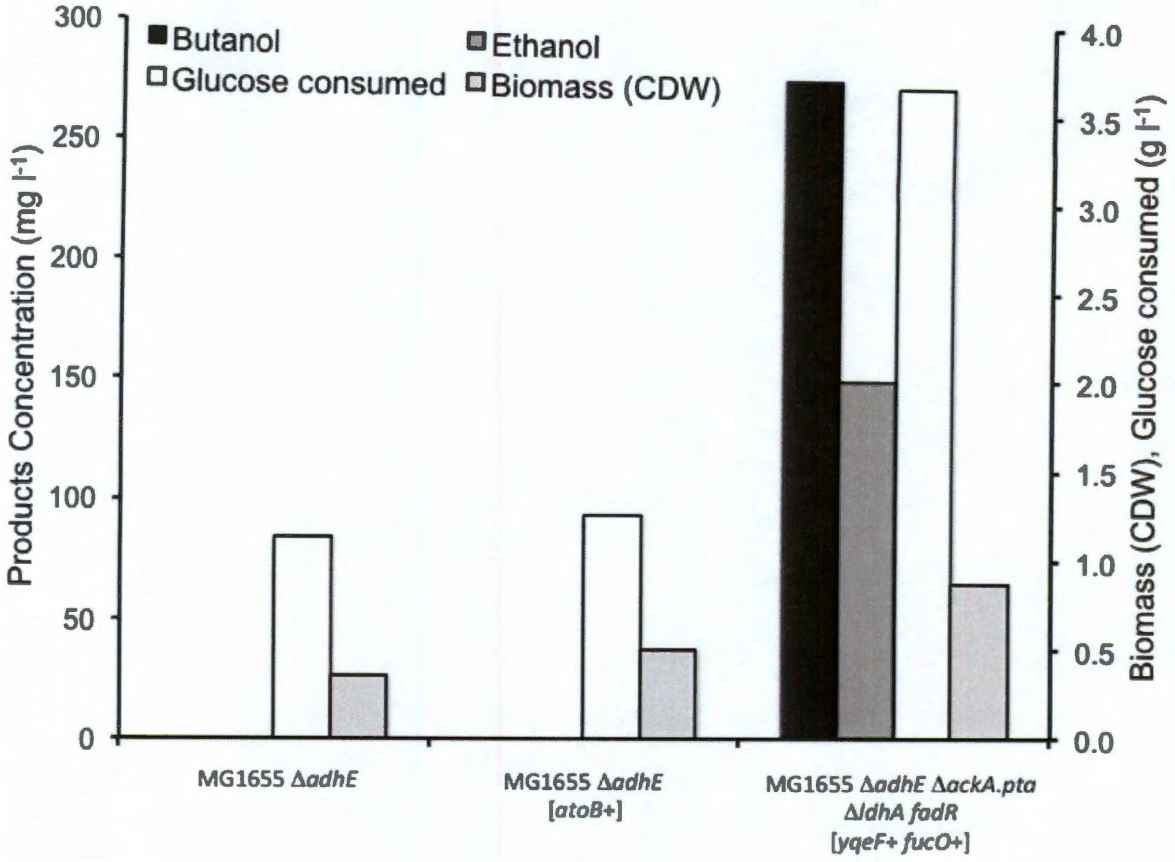

Figure 3-11 Synthesis of n-butanol using *Escherichia coli* native genes identified via sequence-based and domain-centric analyses

Gene overexpressions and deletions are indicated by "+" and "Δ", respectively, next to the corresponding genes. Experiments were run at 37 °C for 48 hours in shake flasks using glucose (2% w/v) minimal medium.

3.3.2.2. Butyrate

Butyrate is produced in Clostridia during the acidogenic phase; a two-step reaction catalyzed by the phosphotransbutylase (Ptb) and the butyrate kinase (Buk) mediates the conversion of butyryl-CoA to butyrate with the intermediate formation of butyryl-phosphate (Figure 3-12).

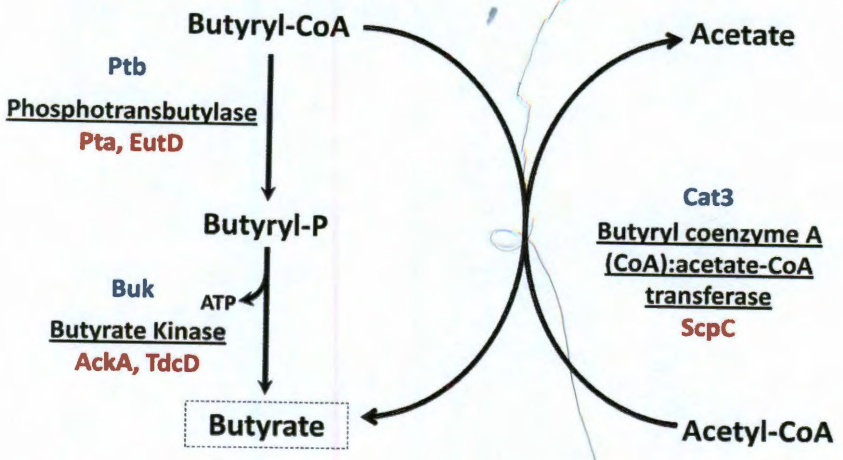

Figure 3-12 *Escherichia coli* surrogate pathway for the clostridial butyrate biosynthesis. Native Clostridial enzymes are reported in blue; *E. coli* set of functional surrogate enzymes are reported in red.

An avenue alternative to this has however been recently elucidated in *Clostridium kluyveri*, where butyryl CoA is exchanged with acetate to yield acetyl-CoA and butyrate (Seedorf et al., 2008) (Figure 3-12).

Homology search of functional homologues based on protein sequence comparison has identified EutD and Pta as possible candidates for the phosphotransbutylase reaction (Figure 3-12) (**Table 3-2**). The EutD and Pta are both *E. coli* proteins characterized by phosphate-acetyltransferase activity (Bologna et al., 2010; Campos-Bermudez et al., 2010) and responsible for the conversion of acetyl-CoA to acetyl-phosphate. EutD has higher specificity for acetyl-CoA than Pta and it is therefore hypothesized to be a more efficient phosphotransacetylase than Pta in catalyzing the reaction to acetyl-phosphate (Bologna et al., 2010; Campos-Bermudez et al., 2010).

The second step in the butyrate biosynthesis consists in the conversion of butyryl-phosphate to butyrate, which in Clostridia is catalyzed by the butyrate kinase (Buk) (Figure

3-12) (Lee et al., 2008b). The sequence and structural analysis suggested two potential *E. coli* surrogates: TdcD and AckA. The *E. coli* enzymes TdcD (YhaA) and AckA are characterized by both propionate and acetate kinase activity (Hesslinger et al., 1998), although a higher specificity for propionyl-phosphate has been measured for TdcD (Hesslinger et al., 1998), suggesting a longer-chain specificity.

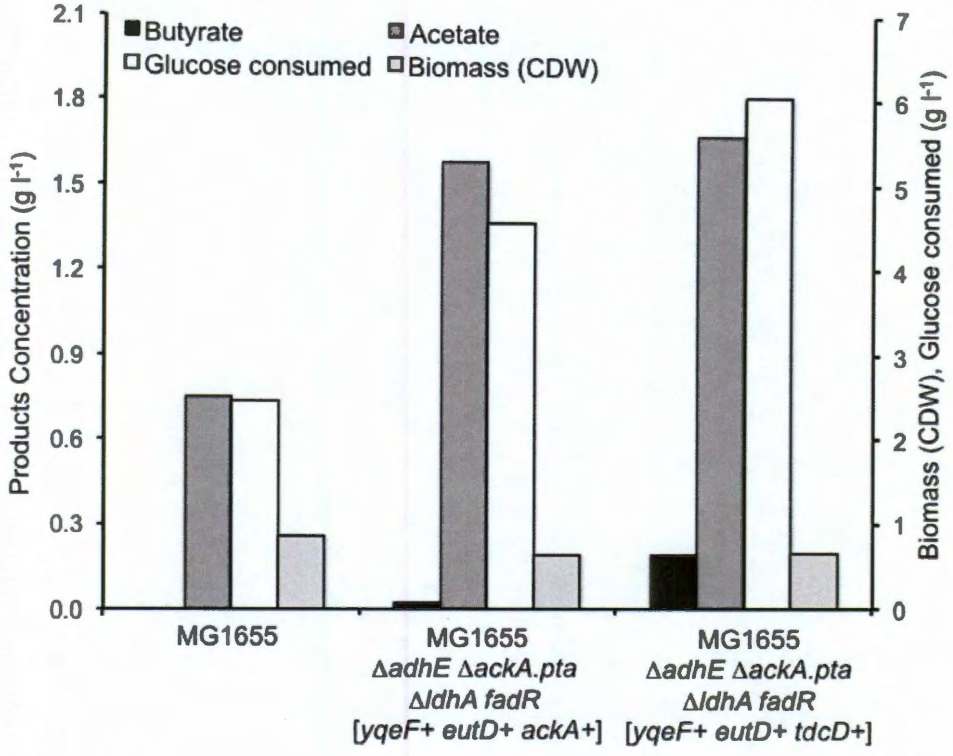

Figure 3-13 Butyrate biosynthesis using *Escherichia coli* functional surrogates identified via sequence-based and domain-centric analyses. Gene overexpressions and deletions are indicated by "+" and "Δ", respectively, next to the corresponding genes.

Since butyryl-CoA is a common intermediate of both n-butanol and butyrate, accumulation of this intermediate was achieved by implementing a metabolic engineering strategy analogous to what was reported in Section 3.3.2.1. Constitutive expression of the *fad*

regulon was achieved by introducing mutations in the local regulator FadR (Dellomonaco et al., 2010b).

The two-step pathway mediating the conversion of butyryl-CoA to butyrate was overexpressed in the engineered *E. coli* strain MG1655 Δ*adhE* Δ*ldhA* Δ*ackA.pta fadR*, in which competing pathways for the acetyl-CoA pool had been deleted, along with an alternative native *E. coli* thiolase (YqeF). Overexpression of EutD-AckA resulted in the production of acetate as the main fermentation product, along with the synthesis of 23 mg l^{-1} of butyrate (Figure 3-13). However, when TdcD was expressed in *E. coli* in combination with EutD, butyrate biosynthesis up to 185 mg l^{-1} was observed (Figure 3-13), in agreement with the higher specificity of TdcD for longer chain acyl-CoA phosphate intermediates (Hesslinger et al., 1998).

3.3.2.3. Acetone

Clostridial acetone biosynthesis from acetoacetyl-CoA is a two-step pathway catalyzed by the enzymes CoA-transferase, and acetoacetate decarboxylase, as reported in Figure 3-9. In *Clostridium acetobutylicum* the acetate and butyrate produced during the acidogenic phase is coupled to the production of acetone via the acetate(butyrate)-acetoacetate CoA-transferase complex (CoAT) (Andersch et al., 1983) as a detoxification mechanism to reduce the inhibitory effects of acids accumulation on cell growth. The Co-A transferase acts by converting the carboxylic acids to the respective CoA thioester intermediate at the expense of the CoA intermediate of another species of carboxylic acid (Figure 3-9). By doing so, the action of the Co-A transferase results energetically favorable since it transfers the energy from a thioester to an acid without expenditure of ATP

(Hartmanis and Gatenbeck, 1984). The CoA transferase complex appears to be a tetramer composed by two subunits, encoded by *ctfA* and *ctfB*, respectively (Cary et al., 1990).

The *E. coli* enzymes AtoD and AtoA were identified via sequence- and structure-based search as highly homologous (76% and 72%, respectively) to the Clostridial CoAT subunits CtfA and CtfB (**Table 3-2**). Both enzymes are encoded within the *E. coli atoDABE* operon and are involved in the catabolism of short-chain fatty acids (Jenkins and Nunn, 1987a). AtoD and AtoA are two subunits of a protein complex characterized by acetyl-CoA:acetoacetyl-CoA transferase activity that mediates the activation of short chain fatty acids to their respective thioesters (Jenkins and Nunn, 1987a). The *in silico* analysis also identified YdiF (Reed et al., 2003) as a potential functional surrogate. The YdiF protein from *Escherichia coli* belongs to a large sequence family of CoA transferases, present in bacteria to humans, which utilize oxoacids as acceptors (Rangarajan et al., 2005). YdiF is organized into tetramers, with each monomer having an open alpha/beta structure characteristic of Family I CoA transferases and it displays enzymatic activity with short-chain acyl-CoAs (Rangarajan et al., 2005).

The last step in the acetone pathway involves the decarboxylation of acetoacetyl-CoA to acetone, which in Clostridia is catalyzed by the acetoacetate decarboxylase (Adc). The homology search based on protein sequence, structure, and domain organization similarity, has in this case not identified any potential candidate (**Table 3-2**, Table 3-3). This result is further supported by a taxonomic analysis of phylogeny conducted to assess the distribution of enzymes with acetoacetate decarboxylase activity in the γ-Proteobacteria class, according to which this activity has not been detected to date in *E. coli* species, as reported in Figure

3-14. However, an alternative possibility would be that other known native *E. coli* decarboxylases could act on the acetoacetyl-CoA intermediate.

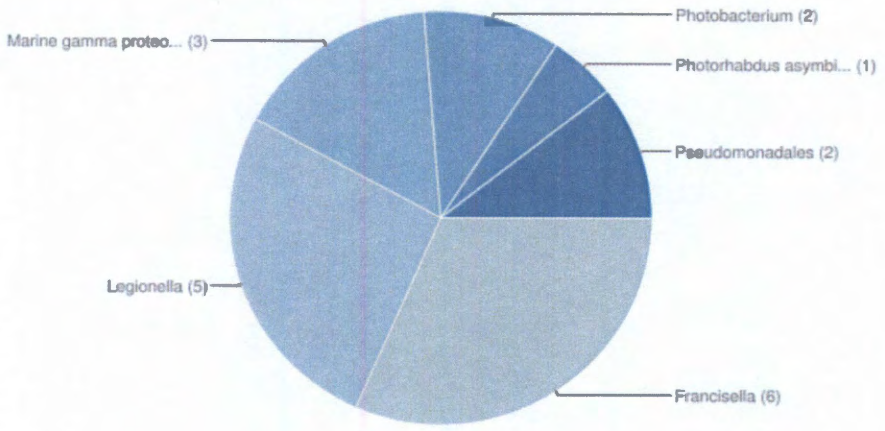

Figure 3-14 Taxonomic distribution of the acetoacetate decarboxylase in the class of γ-Proteobacteria based on currently annotated protein sequences
(UniProt, www.uniprot.org (Bairoch et al., 2005)

Han and coworkers (Han, 2011) have however recently demonstrated the feasibility of synthesizing acetone in Clostridial strains in which the *adc* gene was disrupted. They propose a non-enzymatic decarboxylative mechanism based on the hypothesis that the carbon–carbon bond breakage in acetoacetate that results in release of CO_2 would be facilitated by conditions that make the β-carbonyl group a better electron sink, like for example, the low pH environment created during acidogenesis (Han, 2011).

In an attempt to promote the conversion of acetoacetyl-CoA to acetone based on the considerations stated above, we blocked enzymes involved in the synthesis of by-products competing for the acetyl-CoA pool like ethanol (*adhE*), lactate (*ldhA*) and acetate (*ackA.pta*) and we overexpressed the *E. coli* acetyl-CoA:acetoacetyl-CoA transferase complex (AtoDA)

along with the thiolase identified above for the n-butanol synthesis (AtoB). However, no acetone was produced under these conditions (Figure 3 12a).

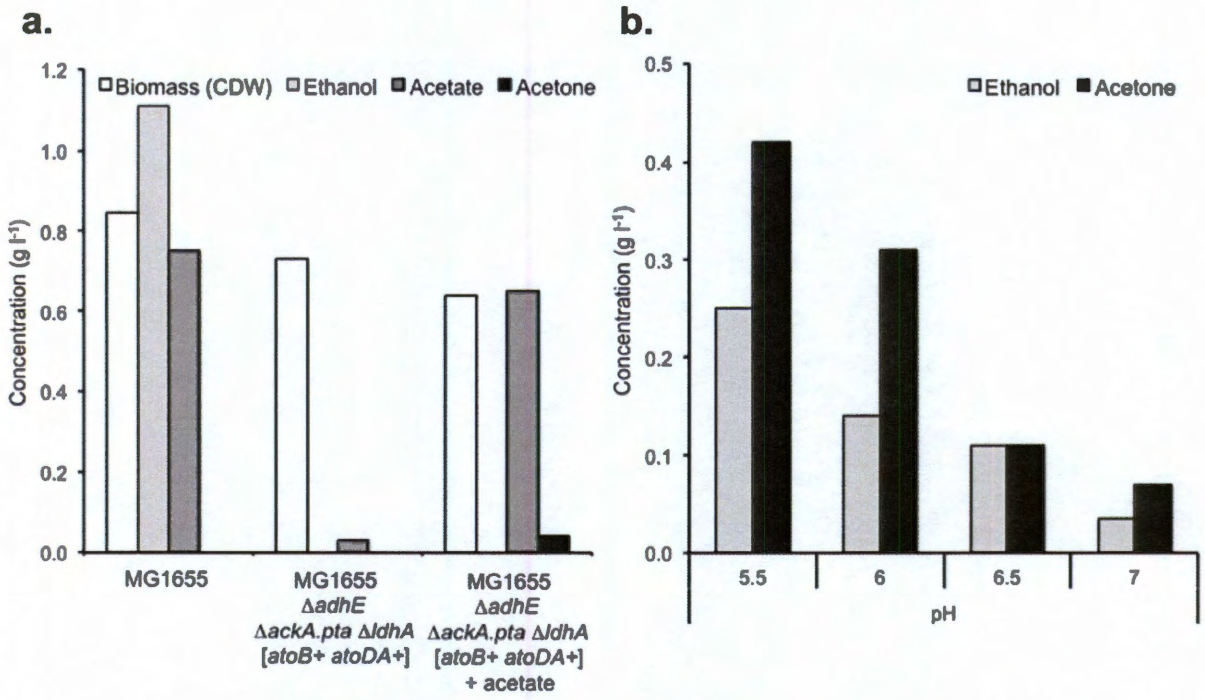

Figure 3-15 Biosynthesis of acetone using *Escherichia coli* functional surrogates

a. Synthesis of acetone using native genes identified via sequence-based and domain-centric analyses. Gene overexpressions and deletions are indicated by "+" and "Δ", respectively, next to the corresponding genes.

b. Acetone synthesis in strain MG1655 Δ*adhE* Δ*ldhA* Δ*ackA.pta* [*atoB*+ *atoDA*+] is reported for different pH. An initial OD_{550} of approximately 2 was used. Cultures were run at 37°C for 24h in MOPS minimal medium + glucose 20 gl^{-1} + acetate 1gl^{-1}.

Thus, we reasoned that the synthesis of acetone could have been limited by the low levels of acetate, a metabolite required in the second step of the acetone pathway from acetyl-CoA (Figure 3-9). In agreement with this hypothesis, supplementation of the growth medium with acetate resulted in acetone production up to 43 mg l-1 (Figure 3-15a). Acetone synthesis in Clostridia is synthesized, as mentioned above, during solventogenesis, a metabolic phase postulated to be induced by the low pH environment. Based on this observation, we

hypothesized that pH might influence activity of the acetone biosynthetic pathway and/or enhance the non-enzymatic decarboxylation of acetoacetate, as recently demonstrated by studies conducted *in vitro* by Han and coworkers (Han, 2011). A range of different pH was therefore explored (Figure 3-15) to simulate the low pH environment created by acidogenesis in Clostridia. In agreement with our hypothesis, acetone synthesis was enhanced by low pH conditions, returning titers up to 420 mg l^{-1} (Figure 3-15b). Although acetone production had been previously demonstrated in *E. coli* by heterologous expression of the Clostridial pathway (Bermejo et al., 1998; Dellomonaco et al., 2010b), this work represents the first report of acetone synthesis in *E. coli* using solely native genes.

Table 3-2 Identification of *Escherichia coli* surrogates for Acetone-Butanol-Ethanol metabolism via sequence-based comparative analysis

Source of gene sequences used to identify *E. coli* surrogates					Surrogates identified via I-TASSER(Roy et al., 2010)				Surrogates identified via protein pBLAST (Altschul et al., 1997)				
Organism	**Accession #**	**Protein**	**E.C. #**	**Function**	**E.C. #**	**TM-Score[a]**	**EC-Score[b]**	**Protein**	**E.C. #**	**Protein**	**E-value[c]**	**Identity**	**Similarity**
Identification of E. coli surrogates for the biosynthesis of ABE intermediates													
Clostridium acetobutylicum	P45359	ThlA	2.3.1.9	Acetyl-CoA C-acetyltransferase	2.3.1.9	0.9872	3.4125	AtoB	2.3.1.9	YqeF	7E-133	56%	76%
								YqeF	2.3.1.9	AtoB	5E-128	61%	74%
					2.3.1.16	0.9108	2.4499	FadA	2.3.1.174	PaaJ	6E-91	45%	63%
								FadI	2.3.1.16	FadA	1E-72	40%	60%
									2.3.1.16	FadI	5E-51	33%	52%
Clostridium acetobutylicum	P52041	Hbd	1.1.1.157	3-hydroxybutyryl-CoA dehydrogenase	1.1.1.35	0.9424	2.5506	FadB	1.1.1.157	PaaH	4E-67	44%	64%
								FadJ	1.1.1.35	FadJ	2E-43	37%	57%
									1.1.1.35	FadB	2E-40	36%	55%
Clostridium acetobutylicum	P52046	Crt	4.2.1.55	3-hydroxybutyryl-CoA dehydratase	4.2.1.17	0.9552	2.4493	FadB	4.2.1.17	PaaF	4E-47	37%	63%
								FadJ	4.2.1.-	CaiD	7E-39	35%	56%
								PaaF	4.2.1.17	FadB	9E-33	36%	61%
								PaaG	4.1.3.36	MenB	6E-30	31%	50%
									4.2.1.17	FadJ	4E-29	35%	59%
Clostridium acetobutylicum	P52042	Bcd	1.3.99.2	Butyryl-CoA dehydrogenase	1.3.99.10	0.9778	2.4998	AidB	1.3.99-	CaiA	9E-43	28%	47%
					1.3.99.3	0.9724	2.4449	FadE	n.d.	YdiO	2E-41	30%	48%
									1.3.99-	FadE	4E-15	30%	53%
Clostridium acetobutylicum	P52039	EtfA	1.5.5.1	Electron transfer flavoprotein subunit	-[d]	-[d]	-[d]	-[d]	1.5.5.1	YdiR	7E-31	30%	53%
									1.5.5.1	FixB	7E-26	31%	49%
Clostridium acetobutylicum	P52040	EtfB	1.5.5.1	Electron transfer flavoprotein subunit	-[d]	-[d]	-[d]	-[d]	1.5.5.1	YdiQ	2E-21	30%	48%
									1.5.5.1	FixA	1E-18	32%	51%
Identification of E. coli surrogates for solventogenesis													
BUTANOL													
Clostridium acetobutylicum	P33744	AdhE	1.1.1.1. 1.2.1.10	Aldehyde-alcohol dehydrogenase	-[d]	-[d]	-[d]	-[d]	1.1.1.1. 1.2.1.10	AdhE	0.0	56%	75%
Clostridium saccharobutylicum(Berezina et al., 2009)	P13604	Adh1	1.1.1.-	NADH-dependent butanol dehydrogenase	1.1.1.77	0.9404	2.2424	FucO	1.1.1.1.	AdhE	2E-91	42%	62%
					1.1.1.1.	0.8679	1.8069	YiaY	1.1.1.77	FucO	5E-66	35%	58%

								AdhE	1.1.1.-	YiaY	2E-65	37%	56%
								AdhP	1.1.-.-	EutG	4E-61	35%	54%
								FrmA	1.1.1.-	YqhD	6E-28	25%	46%
					1.1.1.6	0.7979	1.4449	GldA					
Pseudomonas sp. strain CF600(Powlowski et al., 1993)	Q52060	DmpF	1.2.1.10	Acetaldehyde dehydrogenase	1.2.1.12	0.7439	1.3969	GapA	1.2.1.10	MhpF	6E-146	79%	92%
Clostridium acetobutylicum	Q04944	BdhA	1.1.1.-	NADH-dependent butanol dehydrogenase	1.1.1.77	0.9072	1.7913	FucO	1.1.1.-	YqhD	3E-68	36%	55%
					1.1.1.1	0.8465	1.7242	YiaY	1.1.1.-	YiaY	1E-32	26%	46%
								AdhE	1.1.-.-	EutG	5E-27	27%	42%
								AdhP	1.1.1.77	FucO	7E-27	26%	46%
								FrmA	1.1.1.1.	AdhE	10E-24	24%	45%
					1.1.1.6	0.7761	1.4418	GldA					
Clostridium acetobutylicum	Q04945	BdhB	1.1.1.-	NADH-dependent butanol dehydrogenase	1.1.1.77	0.9039	1.7305	FucO	1.1.1.-	YqhD	4 E-69	39%	57%
					1.1.1.1	0.8443	1.7044	YiaY	1.1.1.-	YiaY	9E-30	26%	44%
								AdhE	1.1.1.77	FucO	4E-25	26%	44%
								AdhP	1.1.-.-	EutG	1E-22	27%	42%
								FrmA	1.1.1.1.	AdhE	7E-18	22%	44%
					1.1.1.6	0.7747	1.4652	GldA					

ACETONE

Clostridium acetobutylicum	P33752	CtfA	2.8.3.9	Butyrate-acetoacetate CoA-transferase subunit	2.8.3.8	0.9817	2.8995	AtoD	2.8.3.8	AtoD	5E-62	47%	76%
									2.8.3.8	YdiF	4E-09	21%	41%
Clostridium acetobutylicum	P23673	CtfB	2.8.3.9	Butyrate-acetoacetate CoA-transferase subunit	-[d]	-[d]	-[d]	-[d]	2.8.3.8	AtoA	2E-63	54%	72%

Identification of E. coli surrogates for acetogenesis

BUTYRATE

Clostridium acetobutylicum	P58255	Ptb	2.3.1.19	Phosphate butyryltransferase	2.3.1.8	0.9617	2.0729	Pta	2.3.1.8	EutD	3E-15	26%	45%
									2.3.1.8	Pta	6E-11	24%	44%
Clostridium acetobutylicum	Q45829	Buk	2.7.2.7	Butyrate kinase	2.7.2.1	0.9628	1.8007	AckA	2.7.2.15	TdcD	2E-09	25%	48%
					2.7.2.15	0.9637	1.6866	TdcD	2.7.2.1	AckA	3E-09	28%	50%
Clostridium kluyveri(Seedorf et al., 2008)	A5N390	Cat3	2.8.3.8	Butyryl-CoA:acetate CoA transferase	4.1.3.6	0.7651	1.4438	CitE	2.8.3.-	ScpC	4E-17	24%	44%

[a] The Template Modeling-score (TM-score) is defined to assess the topological similarity of protein structure pairs. Its value ranges between 0 and 1, and a higher score indicates better structural match. Statistically, a TM-score < 0.17 means a randomly selected protein pair with the gapless alignment taken from PDB (Roy et al., 2010).

[b] An EC-score > 1.1 is a good indicator of the functional similarity between the query and the identified enzyme analogs (Roy et al., 2010).

[c] The BLAST E-value measures the statistical significance threshold for reporting protein sequence matches against the organism genome database; the default threshold value is 1E-5, in which 1E-5 matches would be expected to occur by chance, according to the stochastic model of Karlin and Altschul (http://www.ncbi.nlm.nih.gov/BLAST/tutorial/).

[d] Result not significant because of TM-score<0.17 and/or EC-score<1.1 (Roy et al., 2010)

Table 3-3 Identification of *Escherichia coli* surrogates for Acetone-Butanol-Ethanol metabolism via domain/motif recognition-based comparative analysis

Source of gene sequences used to identify *E. coli* surrogates					Surrogates identified via domain organization and composition (SMART(Schultz et al., 1998), InterProScan (Quevillon et al., 2005))					
Organism	**Accession #**	**Protein**	**E.C. #**	**Function**	**Domain**	**Pfam (Finn et al., 2010) Accession #**	**Begin**	**End**	**E-value[a]**	**Protein**
Clostridium acetobutylicum	P45359	ThlA	2.3.1.9	Acetyl-CoA C-acetyltransferase	Thiolase, N-terminal domain	PF00108	1	262	8.20E-116	PaaJ
					Thiolase, C-terminal domain	PF02803	269	391	6.50E-52	FadA AtoB YqeF FabH FadA FadI
Clostridium acetobutylicum	P52041	Hbd	1.1.1.157	3-hydroxybutyryl-CoA dehydrogenase	3-hydroxyacyl-CoA dehydrogenase, NAD binding domain	PF02737	3	182	7.20E-66	PaaC
					3-hydroxyacyl-CoA dehydrogenase, C-terminal domain	PF00725	184	280	4.10E-35	FadB FadJ PaaC FadB FadJ
Clostridium acetobutylicum	P52046	Crt	4.2.1.55	3-hydroxybutyryl-CoA dehydratase	Enoyl-CoA hydratase/isomerase family	PF00378	15	184	1.40E-54	MmcD PaaF PaaG
Clostridium acetobutylicum	P52042	Bcd	1.3.99.2	Butyryl-CoA dehydrogenase	Acyl-CoA dehydrogenase, N-terminal domain	PF02771	6	118	3.20E-33	CaiA
					Acyl-CoA dehydrogenase, middle domain	PF02770	122	173	1.70E-15	CaiA
					Acyl-CoA dehydrogenase, C-terminal domain	PF00441	229	378	4.10E-46	CaiA
Clostridium acetobutylicum	P52039	EtfA	1.5.5.1	Electron transfer flavoprotein subunit	Electron transfer flavoprotein domain	PF01012	9	200	9.33E-61	YdiR
					Electron transfer flavoprotein FAD-binding domain	PF00766	210	295	3.90E-35	YdiR
Clostridium	P52040	EtfB	1.5.5.	Electron transfer	Electron transfer flavoprotein	PF01012	22	211	2.89E-66	YdiQ

acetobutylicum				flavoprotein subunit	domain					YdiR YgcR

Identification of E. coli surrogates for solventogenesis

BUTANOL

Clostridium acetobutylicum	P33744	AdhE	1.1.1.1 1.2.1.10	Aldehyde-alcohol dehydrogenase	Aldehyde dehydrogenase	PF00171	1	406	6.10E-15	AdhE
					Iron-containing alcohol dehydrogenase	PF00465	455	850	9.90E-91	AdhE
Clostridium saccharobutylicum(Berezina et al., 2009)	P13604	Adh1	1.1.1.-	NADH-dependent butanol dehydrogenase	Iron-containing alcohol dehydrogenase	PF00465	7	373	9E-114	Adh2 YbdH EutG YqhD
Pseudomonas sp. strain CF600(Powlowski et al., 1993)	Q52060	DmpF	1.2.1.10	Acetaldehyde dehydrogenase	Semialdehyde dehydrogenase, NAD binding domain	PF01118	6	124	1.32E-10	AdhE
					Prokaryotic acetaldehyde dehydrogenase, dimerisation	PF09290	132	285	1.50E-63	AdhE
Clostridium acetobutylicum(Walter et al., 1992)	Q04944	BdhA	1.1.1.-	NADH-dependent butanol dehydrogenase	Iron-containing alcohol dehydrogenase	PF00465	9	381	3.40E-98	Adh2 YbdH EutG YqhD
Clostridium acetobutylicum	Q04945	BdhB	1.1.1.-	NADH-dependent butanol dehydrogenase	Iron-containing alcohol dehydrogenase	PF00465	9	381	7.40E-95	Adh2 YbdH EutG YqhD

ACETONE

Clostridium acetobutylicum	P33752	CtfA	2.8.3.9	Butyrate-acetoacetate CoA-transferase subunit	Coenzyme A transferase	PF01144	5	216	1.69E-58	YdiF AtoD AtoA
Clostridium acetobutylicum	P23673	CtfB	2.8.3.9	Butyrate-acetoacetate CoA-transferase subunit	Coenzyme A transferase	PF01144	10	205	2.00E-51	YdiF AtoD AtoA
Clostridium acetobutylicum	P23670	Adc	4.1.1.4	Acetoacetate decarboxylase	Acetoacetate decarboxylase	PF06314	11	243	8.10E-72	-

Identification of E. coli surrogates for acetogenesis

BUTYRATE

Clostridium acetobutylicum	P58255	Ptb	2.3.1.19	Phosphate butyryltransferase	Phosphate acetyl/butaryl transferase	PF01515	56	295	3.70E-09	-
Clostridium acetobutylicum	Q45829	Buk	2.7.2.7	Butyrate kinase	Acetokinase family	PF00871	3	354	4.40E-80	TdcD
Clostridium kluyveri (Seedorf et al., 2008)	A5N390	Cat3	2.8.3.8	Butyryl-CoA:acetate CoA transferase	Acetyl-CoA hydrolase/transferase N-terminal domain	PF02550	6	192	1.40E-23	ScpC

[a] The E-value measures the statistical significance threshold for reporting protein domain matches against the organism proteome database; the default threshold value is 1E-5, in which 1E-5 matches would be expected to occur by chance

3.3.1. Homoethanolic fermentation

While *E. coli* has been extensively exploited for bioconversions due to its fast growth under both anaerobic and aerobic conditions, and the high glycolytic flux it can sustain, it is not regarded as optimal for ethanol production because of the low yield it can achieve when compared to yeast strains or *Zymomonas* spp. Ethanol yield is poor because under fermentative conditions *E. coli* also produces lactic, acetic, formic, and succinic acids (Section 3.1.2.2). Homoethanol fermentation in *E. coli* is hindered by redox imbalance. The pathway to ethanol starts from pyruvate, which is cleaved into acetyl-CoA and formic acid by pyruvate formate lyase (PFL) (Figure 3-16). Reduction of acetyl-CoA to ethanol proceeds in two steps through acetaldehyde as intermediate; the multienzyme protein AdhE plays the role of acetaldehyde dehydrogenase and alcohol dehydrogenase, each requiring one NADH (Figure 3-16). Thus, on a triose basis the pathway from pyruvate to ethanol consumes two NADH, while glycolysis to pyruvate only provides one NADH (in conversion of glyceraledehyde-3-P to 1,3-bisphosphoglycerate). Therefore, ethanol production is balanced by other more oxidized products such as acetic acid (no NADH consumed). To circumvent the redox limitation of the endogenous ethanol pathway, pyruvate decarboxylase (PDC) and alcohol dehydrogenase (ADH) enzymes from *Z. mobilis* have been expressed in *E. coli*, via a plasmid bearing an artificial *pet* (production of ethanol) operon containing the *pdc* and *adhB* genes (Ingram et al., 1998). The transformation has conferred homoethanol pathway, with ethanol accounting for 95% of the fermentation products. Also, redox balance is possible in the heterologous pathway because conversion of pyruvate to

acetaldehyde and CO_2 by PDC is non-oxidative, requiring only one NADH for the reduction of acetaldehyde to ethanol (Figure 3-16).

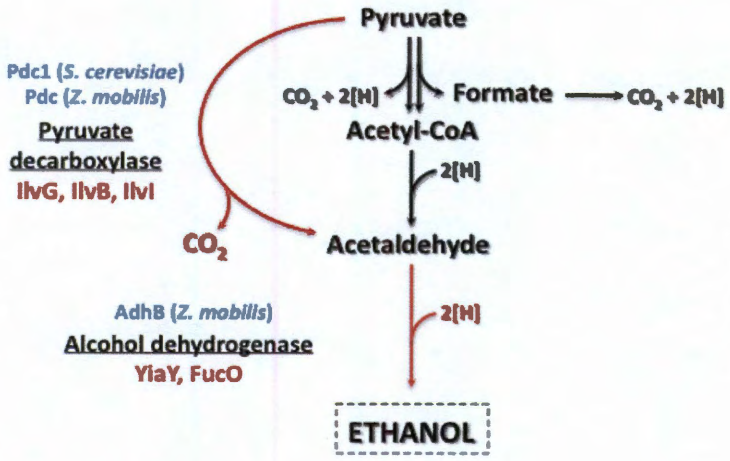

Figure 3-16 Homoethanolic fermentation engineered in *Escherichia coli*
Reactions engineered are reported in red. Enzymes from native producers (*S. cerevisiae, Z. mobilis*) are reported in blue, while *E. coli* potential functional surrogates are reported in red. Abbreviations: 2[H] = NADH = $FADH_2$ = H_2

The HoME methodology was implemented to identify functional surrogates of the non-oxidative pyruvate decarboxylation and hence establish a homoethanologenic metabolic mode in *E. coli*. Sequence-, structure- (Table 3-4), and domain-centric (Table 3-5) homology search for the *S. cerevisiae* and *Z. mobilis* pyruvate decarboxylase consistently identified three *E. coli* functional surrogates, IlvG, IlvB, and IlvI. The products of the genes *ilvG*, *ivlB* and *ilvI* have been identified as distinct isoforms of the acetolactate synthase (Vyazmensky et al., 2009), known to have different catalytic properties and used for the biosynthesis of the amino acids isoleucine and valine. IlvG is however not expressed in *E. coli* K-12 due to a frameshift mutation (Lawther et al., 1982). *In silico* analysis of the *dhaB* gene product from *Zymomonas mobilis* has

identified the hypothetical proteins Yiay and YqhD, along with the lactaldehyde reductase FucO as *E. coli* enzymes that could potentially catalyze the conversion of acetaldehyde to ethanol with high specificity and efficiency.

While a metabolic engineering strategy to establish a homoethanologenic fermentative metabolism in *E. coli* has been already devised, we are currently designing the experimental framework to support its viability.

Table 3-4 Identification of *Escherichia coli* surrogates for homo-ethanologenic metabolism via sequence-based comparative analysis

Source of gene sequences used to identify *E. coli* surrogates					Surrogates identified via I-TASSER(Roy et al., 2010)				Surrogates identified via protein pBLAST (Altschul et al., 1997)				
Organism	**Accession #**	**Protein**	**E.C. #**	**Function**	**E.C. #**	**TM-Score[a]**	**EC-Score[b]**	**Protein**	**E.C. #**	**Protein**	**E-value[c]**	**Identity**	**Similarity**
Identification of E. coli surrogates for the biosynthesis of ABE intermediates													
Saccharomyces cerevisiae	P06169	Pdc1	4.1.1.1	Pyruvate decarboxylase isozyme 1	-[d]	-[d]	-[d]	-[d]	2.2.1.6	IlvG	1E-15	22%	38%
									2.2.1.6	IlvB	2E-15	21%	41%
									2.2.1.6	IlvI	2E-14	22%	40%
Zymomonas mobilis	P06672	Pdc	4.1.1.1	Pyruvate decarboxylase	-[d]	-[d]	-[d]	-[d]	2.2.1.6	PoxB	1E-14	21%	41%
									2.2.1.6	IlvB	3E-13	25%	41%
									2.2.1.6	IlvG	4E-13	21%	38%
									2.2.1.6	IlvI	1E-12	21%	39%
Zymomonas mobilis	P06758	AdhB	1.1.1.1	Alcohol dehydrogenase	1.1.1.77	0.9624	2.6598	FucO	-	YiaY	6E-144	63%	77%
									1.1.1.77	FucO	2E-87	43%	63%
									1.1.-.-	EutG	4E-63	37%	55%
									1.1.1.1	AdhE	1E-52	33%	54%

[a] The Template Modeling-score (TM-score) is defined to assess the topological similarity of protein structure pairs. Its value ranges between 0 and 1, and a higher score indicates better structural match. Statistically, a TM-score < 0.17 means a randomly selected protein pair with the gapless alignment taken from PDB (Roy et al., 2010).

[b] An EC-score > 1.1 is a good indicator of the functional similarity between the query and the identified enzyme analogs (Roy et al., 2010).

[c] The BLAST E-value measures the statistical significance threshold for reporting protein sequence matches against the organism genome database; the default threshold value is 1E-5, in which 1E-5 matches would be expected to occur by chance, according to the stochastic model of Karlin and Altschul (http://www.ncbi.nlm.nih.gov/BLAST/tutorial/).

[d] Result not significant because of TM-score<0.17 and/or EC-score<1.1 (Roy et al., 2010)

Table 3-5 Identification of *Escherichia coli* surrogates for homo-ethanologenic metabolism via sequence-based comparative analysis

Source of gene sequences used to identify *E. coli* surrogates					Surrogates identified via domain organization and composition (SMART (Schultz et al., 1998), InterProScan (Quevillon et al., 2005))					
Organism	**Accession #**	**Protein**	**E.C. #**	**Function**	**Domain**	**Pfam (Finn et al., 2010) Accession #**	**Begin**	**End**	**E-value[a]**	**Protein**
Saccharomyces cerevisiae	P06169	Pdc1	4.1.1.1	Pyruvate decarboxylase isozyme 1	N-terminal TPP binding domain	PF02776	4	180	2.30E-31	IlvG
					TPP, central domain	PF00205	201	340	1.60E-33	IlvI PoxB IlvB IlvG IlvI PoxB IlvB
					C-terminal TPP binding domain	PF02775	387	527	1.30E-10	IlvG IlvI PoxB IlvB
Zymomonas mobilis	P06672	Pdc	4.1.1.1	Pyruvate decarboxylase	N-terminal TPP binding domain	PF02776	3	179	1.10E-29	IlvG
					TPP, central domain	PF00205	199	360	1.10E-14	IlvI PoxB IlvB IlvG IlvI PoxB IlvB
					C-terminal TPP binding domain	PF02775	387	537	2.10E-17	IlvG IlvI PoxB IlvB
Zymomonas mobilis	P06758	AdhB	1.1.1.1	Alcohol dehydrogenase	Iron-containing alcohol dehydrogenase	PF00465	10	375	1.10E-117	Adh2

3.3.2. Glycerol fermentation to 1,3-propanediol

Fermentative metabolism of glycerol has been investigated in great detail in several species of the *Enterobacteriaceae* family, such as *Citrobacter freundii* and *Klebsiella pneumonia* as its dissimilation in these organisms is strictly linked to their capacity to synthesize the highly reduced product 1,3-propanediol (1,3-PDO) (Bouvet et al., 1995). Although it has recently been demonstrated that *Escherichia coli* is able to utilize glycerol fermentatively (Gonzalez et al., 2008), its anaerobic dissimilation has been demonstrated associated to the production of two alternative reduced products, ethanol and 1,2-propanediol (Gonzalez et al., 2008).

In this study we have implemented the HoME methodology to the identification and functional expression of an *E. coli* homologous pathway for the synthesis of 1,3-PDO from glycerol. The conversion of glycerol to 1,3-PDO consists in the dehydration of glycerol dehydratase to form 3-hydroxypropionaldehyde, and its subsequent reduction to the major fermentation product 1,3-PDO by the NADH-linked 1,3-PDO dehydrogenase (1,3-PDO) (Figure 3-17). Glycerol dehydratase is an enzyme consisting of three subunits and catalyzes the conversion of glycerol to 3-hydroxypropionaldehyde by a free radical mediated reaction involving coenzyme B_{12} as an essential cofactor (Biebl et al., 1999). Glycerol dehydratase is encoded by three structural genes *dhaB1*, *dhaB3* and *dhaB4* in the case of *K. pneumoniae* or *dhaB*, *dhaC* and *dhaE* in the case of *C. freundii* (Homann et al., 1990).

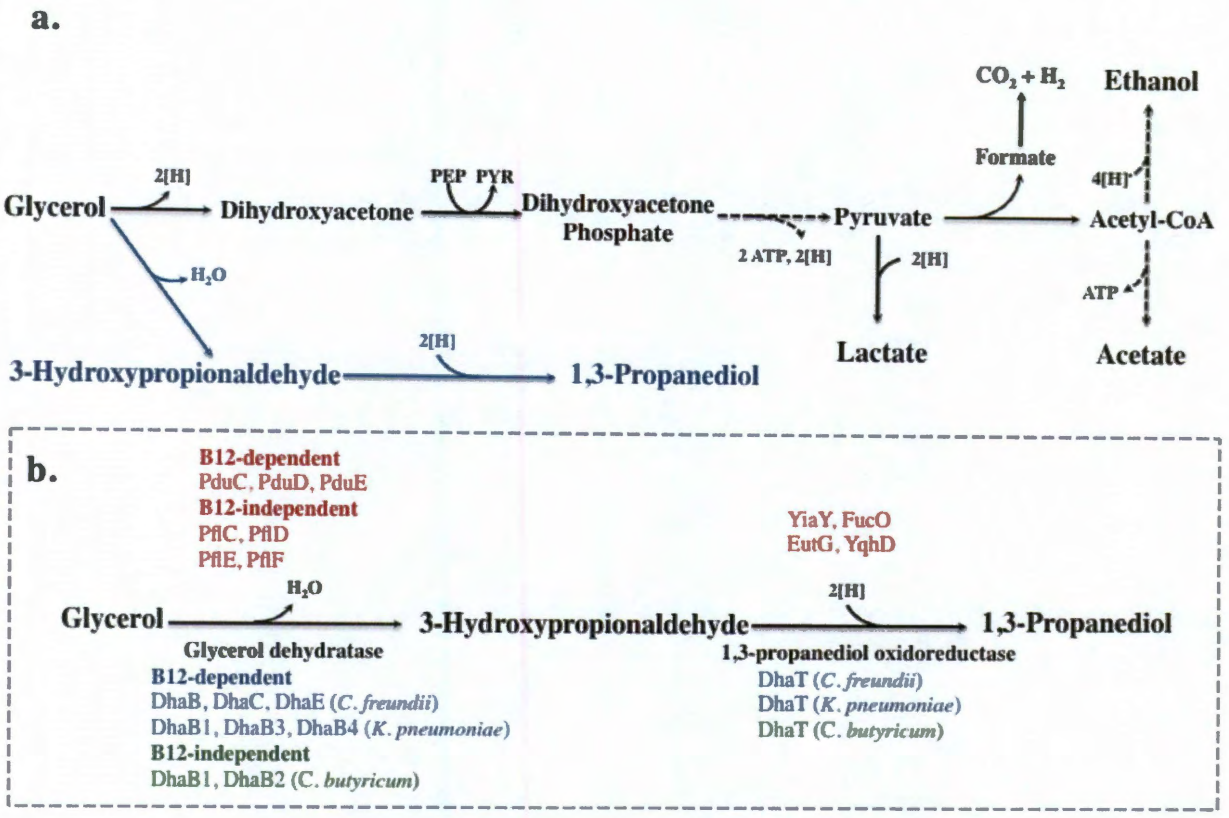

Figure 3-17 Metabolic pathway of the fermentative metabolism of glycerol
(a) Reactions reported in black are common to *E. coli*, *C. freundii*, *K. pneumonia*, *C. butyricum*. (b) Enzymes from native producers (*C. freundii*, *K. pneumoniae*, blue; *C. butyricum*, green) and *E. coli* potential functional surrogates (red) mediating the conversion of glycerol to 1,3-propanediol. Abbreviations: PEP, phosphoenolpyruvate; PYR, pyruvate; 2[H] = NADH = $FADH_2$ = H_2

The structure and enzymatic mechanism are similar for both *Klebsiella* and *Citrobacter* species (Homann et al., 1990). The B_{12}-dependence has important implications for the biosynthesis of 1,3-PDO, because the vitamin B_{12} bound to the enzyme is periodically rendered inactive by glycerol during catalysis. Inactivation by glycerol involves irreversible cleavage of the Co-C bond of coenzyme B_{12}, forming 5'-deoxyadenosine and an alkylcobalamin-like species (Homann et al., 1990). A reactivation

factor catalyzes the exchange of inactive with active vitamin B_{12} by facilitating the dissociation of the inactive vitamin from glycerol dehydratase. A third subunit is finally responsible for the reactivation of the inactivated vitamin (Homann et al., 1990). Similarity search for sequence and structure homologues of the glycerol dehydratase from *K. pneumoniae* and *C. freudii* against the *E. coli* genome have not returned any match (Table 3-6). However, the analysis of protein domain composition suggests PduC, PduD and PduE as potential candidates in *E. coli* (Table 3-7). These genes, part of the *pduABCDE* operon, have been characterized in *Salmonella enterica* in the context of propanediol catabolism (Bobik et al., 1997). Specifically, they are shown to be necessary and sufficient for propanediol dehydratase activity (Bobik et al., 1997). However, the *pduABCDE* operon has only been identified in pathogenic serological types of *E. coli*; hence, genes belonging to this operon have not been pursued in the context of this study.

The demand for vitamin B_{12} for efficient provision of active glycerol dehydratase supersedes generally the natural availability in the host cell and hence requires addition of an excessive amount of a high-cost molecule, vitamin B_{12}, to the culture medium. An alternative approach to overcome the vitamin B_{12} requirement has been recently provided by the *Institut National de la Recherche* in Paris (O'Brien et al., 2004) with the isolation of a coenzyme B12-independent glycerol dehydratase from *Clostridium butyricum.*

The sequence and structural analysis (Table 3-6), in agreement with the domain organization and composition analysis (Table 3-7), suggested PflD and PflF, and PflC and PflE as potential surrogates of the Clostridial enzymes DhaB1 and DhaB2, respectively. Both sets of enzymes have recently been predicted based on homology to function as pyruvate formate lyases (PFL) (Kim and Reed, 2010).

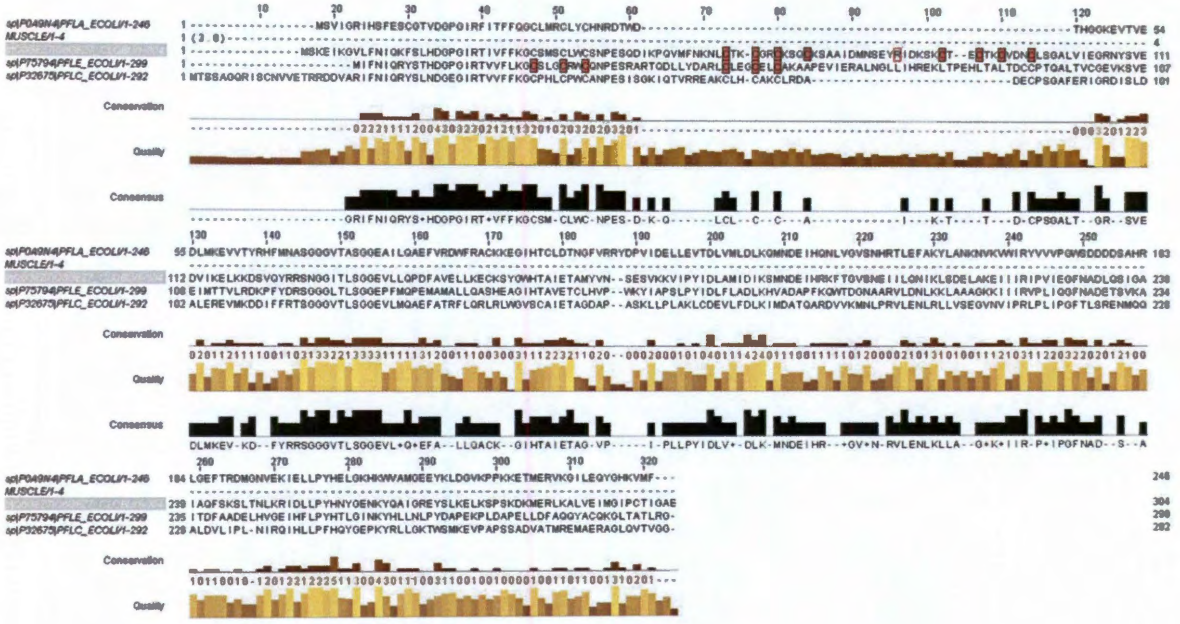

Figure 3-18 Schematic structure showing the domain organization and special sequence features of the *C. butyricum dhaB2*.

The ferredoxin-like binding domain is reported in red. Uniprot accession numbers: DhaB2 Cb (Q8GEZ7), PflA Ec (P0A9N4), PflC Ec (P32675), PflE Ec (P75794).The last two characters of the abbreviated names refer to the organism: *Cb*, *C. butyricum; Ec*, *E. coli*.

Alignment of the *dhaB1* gene product with the PflE and PflF returned high homology around the presumed active site cysteine (Catalytic domain) and the conserved glycine residue (Gly-radical domain), in agreement with what reported by Raynaud and coworkers (Raynaud et al., 2003). Sequence comparison of DhaB2 with PflC and PflE has detected presence of the first cysteine domain in both PflC and PflE. The ferredoxin-like binding motif was instead only present in PflE (although switched of few bases, as shown in Figure 3-18), making it a better candidate.

The second step in the conversion of glycerol to 1,3-PDO consists in the reduction of 3-hydroxypropionaldehyde to 1,3-PDO mediated by the NADH-linked 1,3-PDO

dehydrogenase. The products of the *dhaT* genes from *K. pneumonia*, *C. freudii* and *C. butyricum* are 1,3-propanediol dehydrogenases that share very high sequence homology and present comparable properties (Biebl et al., 1999; Saxena et al., 2009). Homology search based on sequence-, structure- (Table 3-6) and domain composition (Table 3-7) comparisons suggest three *E. coli* oxidoreductases as potential surrogates: YiaY, FucO and YqhD.

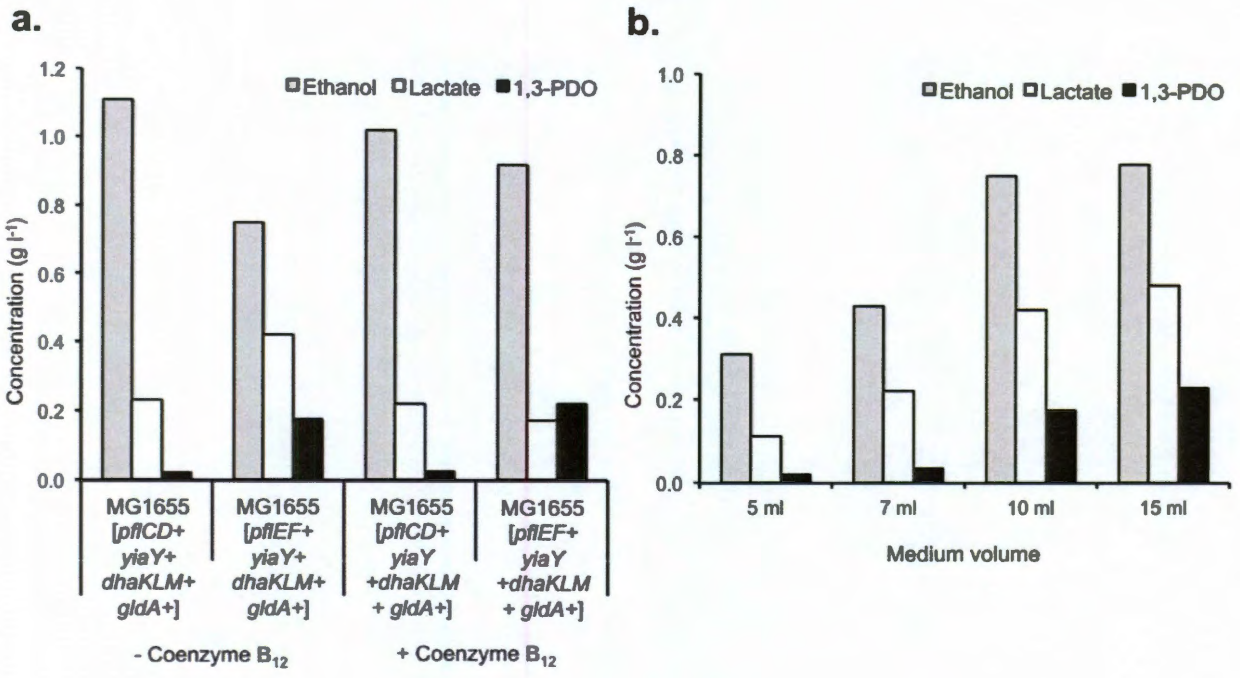

Figure 3-19 Fermentative dissimilation of glycerol to 1,3-propanediol using *Escherichia coli* native genes

a. Coenzyme B_{12} dependence has been explored for two alternative 1,3-PDO dehydratase complexes identified through sequence and domain-centric search.

b. Oxygen sensitivity of the 1,3-propanediol biosynthesis pathway is investigated in the *E. coli* strain MG1655 [*pflEF+ yiaY+ dhaKLM+ gldA+*].

Cultures were run in hungate tubes at 37°C for 48h in MOPS minimal medium + glycerol 10 gl^{-1}.

Based on the considerations reported above, we engineered *E. coli* MG1655 and explored the overexpression of the two alternative 1,3-PDO dehydrogenases (*pflCD*, *pflEF*) in combination with the YiaY 1,3-PDO dehydrogenase (Figure 3-19a). The glycerol dissimilation pathway was concurrently overexpressed to enhance cell growth and generate reducing equivalents needed for the synthesis of 1,3-PDO. While 1,3-PDO synthesis was observed for both MG1655 [*pflCD+ yiaY+ dhaKLM+ gldA+*] and MG1655 [*pflEF+ yiaY+ dhaKLM+ gldA+*], ethanol and lactic acid were main by-products of the fermentation (Figure 3-19a), possibly sustaining redox-balance by acting as alternative sinks for reducing power. Also, our results support the hypothesis that both PflCD and PflEF function as B_{12} independent 1,3-PDO dehydratases, as presence of coenzyme B_{12} in the medium did not seem to affect 1,3-PDO biosynthesis. Based on the high similarity between the *E. coli* surrogate enzyme complex *PflEF* and the native Clostridial dehydratase, we also investigated and demonstrated the oxygen sensitivity of the assembled pathway in MG1655 [*pflEF+ yiaY+ dhaKLM+ gldA+*] (Figure 3-19b). A ten-fold increase in the 1,3-PDO titer was in fact achieved under conditions of anaerobiosis compared to more aerobic ones (Figure 3-19b).

While the results hereby reported demonstrate for the first time biosynthesis of this reduced product in *E. coli* using solely native genes, they also suggest the existence of an alternative latent avenue for the anaerobic glycerol fermentation in *E. coli*. This study proves the presence in *E. coli* of metabolic capabilities considered peculiar of species genetically and metabolically distant, thus corroborating our hypothesis of the existence of functionally pluripotent bacterial cells.

Table 3-6 Identification of *Escherichia coli* surrogates for glycerol fermentation to 1,3-propanediol via sequence-based comparative analysis

Source of gene sequences used to identify *E. coli* surrogates					Surrogates identified via I-TASSER(Roy et al., 2010)				Surrogates identified via protein pBLAST (Altschul et al., 1997)				
Organism	**Accession #**	**Protein**	**E.C. #**	**Function**	**E.C. #**	**TM-Score[a]**	**EC-Score[b]**	**Protein**	**E.C. #**	**Protein**	**E-value[c]**	**Identity**	**Similarity**
B12-independent glycerol dehydratase													
Clostridium butyricum (Raynaud et al., 2003)	Q8GEZ8	DhaB1	4.2.1.30	Glycerol dehydratase	2.3.1.54	0.8266	1.8474	PflD	2.3.1.54	PflD	2E-179	41%	62%
								PflF	2.3.1.54	PflF	3E-153	37%	57%
									2.3.1.-	PflB	9E-50	28%	46%
									2.3.1.-	TdcE	4E-45	27%	47%
Clostridium butyricum (Raynaud et al., 2003)	Q8GEZ7	DhaB2	4.2.1.30	Glycerol dehydratase activator	1.97.1.4	0.7660	1.4907	PflA	1.97.1.4	PflE	2E-48	39%	57%
								PflC	1.97.1.4	PflC	9E-48	34%	54%
								PflE	1.97.1.4	PflA	1E-34	37%	60%
									-	YjjW	2E-23	29%	47%
1,3-propanediol oxidoreductases													
Citrobacter freundii (Daniel et al., 1995)	P45513	DhaT	1.1.1.202	1,3-propanediol oxidoreductase	1.1.1.77	0.9748	4.2711	FucO	1.1.1.1	YiaY	2E-100	48%	64%
					1.1.1.1.	0.8643	2.0342	AdhE	1.1.1.77	FucO	4E-82	43%	61%
								AdhP	1.1.-.-	EutG	5E-57	35%	54%
								FrmA	1.1.1.1	AdhE	7E-53	33%	52%
					1.1.1.6	0.8046	1.6012	GldA	1.1.1.-	YqhD	4E-21	25%	43%
									1.1.1.6	GldA	4E-04	22%	38%
Klebsiella pneumonia (Zhuge et al., 2010)	Q59477	DhaT	1.1.1.202	1,3-propanediol oxidoreductase	1.1.1.77	0.9530	2.6543	FucO	1.1.1.1	YiaY	2E-101	49%	63%
					1.1.1.1.	0.8661	2.0322	AdhE	1.1.1.77	FucO	1E-82	43%	61%
								AdhP	1.1.-.-	EutG	3E-57	35%	54%
								FrmA	1.1.1.1	AdhE	1E-53	34%	54%
					1.1.1.6	0.8056	1.5961	GldA	1.1.1.-	YqhD	3E-21	26%	43%
									1.1.1.6	GldA	3E-06	24%	40%
Clostridium butyricum (Raynaud et al., 2003)	Q8GEZ6	DhaT	1.1.1.202	1,3-propanediol oxidoreductase	1.1.1.77	0.9603	2.6526	FucO	1.1.1.1	YiaY	2E-104	49%	66%
					1.1.1.1	0.8693	2.0217	AdhE	1.1.1.77	FucO	6E-78	43%	61%
								AdhP	1.1.-.-	EutG	3E-58	35%	52%
								FrmA	1.1.1.1	AdhE	7E-49	32%	52%

	1.1.1.6	0.7779	1.6171	GldA	1.1.1.- 1.1.1.6	YqhD GldA	1E-17 1E-07	25% 24%	45% 40%

[a] The Template Modeling-score (TM-score) is defined to assess the topological similarity of protein structure pairs. Its value ranges between 0 and 1, and a higher score indicates better structural match. Statistically, a TM-score < 0.17 means a randomly selected protein pair with the gapless alignment taken from PDB (Roy et al., 2010).

[b] An EC-score > 1.1 is a good indicator of the functional similarity between the query and the identified enzyme analogs (Roy et al., 2010).

[c] The BLAST E-value measures the statistical significance threshold for reporting protein sequence matches against the organism genome database; the default threshold value is 1E-5, in which 1E-5 matches would be expected to occur by chance, according to the stochastic model of Karlin and Altschul (http://www.ncbi.nlm.nih.gov/BLAST/tutorial/).

Table 3-7 Identification of *Escherichia coli* surrogates for glycerol fermentation to 1,3-propanediol via domain/motif recognition-based comparative analysis

Source of gene sequences used to identify *E. coli* surrogates					Surrogates identified via domain organization and composition (SMART(Schultz et al., 1998), InterProScan (Quevillon et al., 2005))					
Organism	**Accession #**	**Protein**	**E.C. #**	**Function**	**Domain**	**Pfam (Finn et al., 2010) Accession #**	**Begin**	**End**	**E-value[a]**	**Protein**
B12-dependent glycerol dehydratase										
Citrobacter freundii (Seifert et al., 2001)	P45514	DhaB	4.2.1.30	Glycerol dehydratase large subunit	Dehydratase large subunit	PF02286	2	555	0.00E+00	PduC
Citrobacter freundii (Seifert et al., 2001)	Q59273	DhaC	4.2.1.30	Glycerol dehydratase 21kDa subunit	Dehydratase medium subunit	PF02288	15	189	1.10E-84	PduD
Citrobacter freundii (Seifert et al., 2001)	Q59274	DhaE	4.2.1.30	Glycerol dehydratase 16kDa subunit	Dehydratase small subunit	PF02287	2	138	1.00E-60	PduE
Klebsiella pneumonia (Kajiura et al., 2001)	Q59476	DhaB1	4.2.1.30	Glycerol dehydrase alpha subunit	Dehydratase large subunit	PF02286	2	555	0.00E+00	PduC
Klebsiella pneumonia (Kajiura et al., 2001)	Q59475	DhaB3	4.2.1.30	Glycerol dehydratase gamma subunit	Dehydratase small subunit	PF02287	2	138	7.50E-63	PduE
Klebsiella pneumonia (Kajiura et al., 2001)	Q59474	DhaB4	4.2.1.30	Glycerol dehydratase large subunit	Diol dehydratase reactivase ATPase-like domain	PF08841	276	607	3.30E-158	PduG
B12-independent glycerol dehydratase										
Clostridium butyricum (Raynaud et al., 2003)	Q8GEZ8	DhaB1	4.2.1.30	Glycerol dehydratase	Pyruvate formate lyase	PF02901	8	649	4.10E-260	PflB
					Glycine radical	PF01228	665	769	2.80E-28	PflD TdcE PflF
Clostridium butyricum (Raynaud et al., 2003)	Q8GEZ7	DhaB2	4.2.1.30	Glycerol dehydratase activator	Radical SAM superfamily	PF04055	26	238	9.90E-24	AslB ThiH HemN PflC YjeK YjjW YggW PflE YdeM YgiQ
Clostridium butyricum	Q8GEZ6	DhaT	1.1.1.202	1,3-propanediol	Iron-containing alcohol	PF00465	9	377	5.90E-116	Adh2

(Raynaud et al., 2003)				oxidoreductase	dehydrogenase					YbdH EutG YqhD

1,3-propanediol oxidoreductases

Citrobacter freundii (Daniel et al., 1995)	P45513	DhaT	1.1.1.202	1,3-propanediol oxidoreductase	Iron-containing alcohol dehydrogenase	PF00465	11	379	4.00E-122	Adh2 YbdH EutG YqhD
Klebsiella pneumonia (Zhuge et al., 2010)	Q59477	DhaT	1.1.1.202	1,3-propanediol oxidoreductase	Iron-containing alcohol dehydrogenase	PF00465	11	379	3.10E-123	Adh2 YbdH EutG YqhD
Clostridium butyricum (Raynaud et al., 2003)	Q8GEZ6	DhaT	1.1.1.202	1,3-propanediol oxidoreductase	Iron-containing alcohol dehydrogenase	PF00465	9	377	5.90E-116	Adh2 YbdH EutG YqhD

[a] The E-value measures the statistical significance threshold for reporting protein domain matches against the organism proteome database; the default threshold value is 1E-5, in which 1E-5 matches would be expected to occur by chance (Quevillon et al., 2005; Schultz et al., 1998)

3.4. Discussion

Current employed metabolic engineering methodologies build on phenetic observations and on the premise that bacteria that are genetically distant from a phylogenetic standpoint are also metabolically distinct. Filling the gap of knowledge between functional and molecular differentiation becomes therefore fundamental also to the development of novel approaches to design and construct optimal microbial cell factories.

While earlier industrial microbiology efforts aimed at exploiting the metabolic capabilities of native producer organisms, synthetic biology and metabolic engineering utilize the information generated by structural and functional genomics to identify efficient pathways for the biosynthesis of products of interest in native producers and reconstruct them in genetically tractable and industrially amenable bacteria, like for example the model organisms *E. coli* and *S. cerevisiae*. Implementation of this method, which we term Heterologous Metabolic Engineering (HeME), thus fails at exploiting the full biosynthetic potential of the host. Furthermore, incompatibility of the heterologous pathway with the host metabolism is of significant concern, as in many cases it results in suboptimal operation of the engineered pathway. As an alternative, we propose Homologous Metabolic Engineering (HoME) as a novel methodology that aims at exploiting functional relatedness of metabolic pathways in genetically and metabolically distant species. The implementation of HoME entails the recruitment and functional activation of native genes and/or pathways and their assembly into superior metabolic

routes for the biosynthesis of products of interest, e.g. fuels and chemicals. The proposed strategy proceeds by identifying functional homologues for the heterologous pathway by protein-sequence comparative analysis and domain-centric homology search. Pathways identified are assembled and directly expressed in *E. coli*; functional expression also involves manipulating the corresponding regulators and blocking competing pathways for common metabolic intermediates.

In an effort to lay the foundations of this new methodology, we selected the Clostridial acetone-butanol-ethanol (ABE) fermentation, the anaerobic glycerol fermentation to 1,3-Propanediol and the yeast homoethanolic fermentation as examples of metabolic modes peculiar of organisms that are molecularly and metabolically distant from *E. coli*.

E. coli functional surrogates of the enzymatic steps responsible for mediating pathways of the ABE fermentation in Clostridia were identified through sequence-, structure-, and domain-centric homology search (Section 3.3.2). Their assembly in *E. coli* strains engineered to block competing pathways for common intermediates, resulted in the biosynthesis of butanol, acetone and butyric acid, products peculiar of the Clostridial metabolism. Contrarily to what previously assumed, our results demonstrate the existence of latent metabolic potential in *E. coli* that can enable synthesis of non-native products in the absence of foreign genes.

The homoethanologenic metabolism of *Saccharomyces cerevisiae* and *Zymomonas mobilis* has also been explored (Section 3.3.1). A metabolic engineering strategy has been devised based on the indication of the presence of potential *E. coli*

functional surrogates provided by the homology search. However, experimental evidence cannot be provided at this moment.

Proof of concept has also been provided for the case of anaerobic glycerol fermentation to 1,3-PDO (Section 3.3.2). While this bioconversion was thought to be characteristic of *Klebsiella*, *Citrobacter* and *Clostridial* spp., our results hypothesize the existence of a latent similar pathway, therefore corroborating the hypothesis of a common metabolic ancestor.

Chapter 4

Engineered reversal of the β-oxidation cycle as a metabolic platform for the synthesis of advanced fuels and chemicals

4.1. Background and motivation

Metabolic engineering and synthetic biology strategies have for the most part been tailored for the synthesis and optimization of single biofuel or biochemical molecules. However, by integrating synthetic competences from diverse metabolic and phylogenetic origins within an integrated biological system, it is possible to envisage the development of a single technological platform for synthesizing an array of biorenewable molecules for use in the chemical and energy industry. By analogy with the current non-renewable petrochemical industry, a successful bio-based chemical industry will

therefore build upon the platform-chemical approach, where each chemical intermediate is converted to a large number of chemical products. Harnessing the metabolic potential of model organisms provides therefore opportunity to produce a broad array of platform chemicals, thereby creating a consolidated framework for the production of advanced fuels and chemicals. Polyketide and fatty acid biosynthesis are among the biochemical pathways presently explored in the context of the combinatorial biosynthesis of fuels and chemicals (Hertweck, 2009; Nikolau et al., 2008; Steen et al., 2010).

4.1.1. Metabolic platforms currently available

Advanced (long-chain) fuels and chemicals are generated from short-chain metabolic intermediates through pathways that require carbon-chain elongation. The condensation reactions mediating this carbon-carbon bond formation can be catalyzed by enzymes from the thiolase superfamily, including β-ketoacyl-acyl-carrier protein (ACP) synthases, polyketide synthases, 3-hydroxy-3-methylglutaryl-CoA synthases, and biosynthetic thiolases (Haapalainen et al., 2006).

Pathways involving these enzymes have been exploited for fuel and chemical production, with fatty acid biosynthesis (β-ketoacyl-ACP synthases) attracting the most attention in recent years (Handke et al., 2010; Lennen et al., 2010; Steen et al., 2010).

Polyketide biosynthesis takes place through the iterative generation of alkyl chain lengths with increasing chain length, ranging from four up to thirty carbon atoms. Fatty acids along with a diverse set of fungal and bacterial secondary metabolites belong to this class of compounds. Polyketides are synthesized by sequential reactions mediated by a

set of large protein complexes called polyketide synthases (PKSs). The biosynthesis occurs in a stepwise manner from acyl-CoA intermediates with 2-, 3- or 4- carbons, and their activated malonyl derivatives (Hertweck, 2009). Polyketide biosynthesis and fatty acid biosynthesis share striking similarities, although they are considered a metabolic branch point between primary and secondary metabolism. In both fatty acid synthesis and polyketide synthesis, the intermediates are covalently bound to ACP, or Acyl Carrier Protein. However, while the 2C donor for fatty acids growth is Acyl-CoA or Malonyl-CoA, poyketide synthases can use multiple primers including Acetyl-CoA, Propionyl-CoA, Isobutyrl-CoA, etc.. In both fatty acid synthesis and polyketide synthesis the CoA carriers are then exchanged with ACP before being incorporated into the growing molecule. Fatty acids and polyketide synthesis are however quite dissimilar in the mechanism of chain growth: unlike fatty acid biosynthesis in which each chain elongation step is followed by a fixed sequence of ketoreduction, dehydration and enoyl reduction, intermediates of polyketide biosynthesis undergo different functional group modifications (Hertweck, 2009). Nevertheless, in both pathways, when a defined chain length is reached, the thioester-bound substrates are released from the enzyme complex and hydrolyzed by thioesterases. The free carboxylic acids can be sequentially reduced to the corresponding aldehydes and primary alcohols, and, in turn, the alcohols can be acylated to form esters. Based on the thioesterases specificity it is therefore possible to generate a variety of products with diverse chain length and functionalities.

4.1.2. Functional reversal of the β-oxidation as a combinatorial platform for the production of fuels and chemicals

Degradative thiolases, which are part of the thiolase superfamily and naturally function in the β-oxidation of fatty acids (Clark and Cronan, 2005; Poirier et al., 2006), can also operate in the synthetic direction and thus enable carbon-chain elongation.

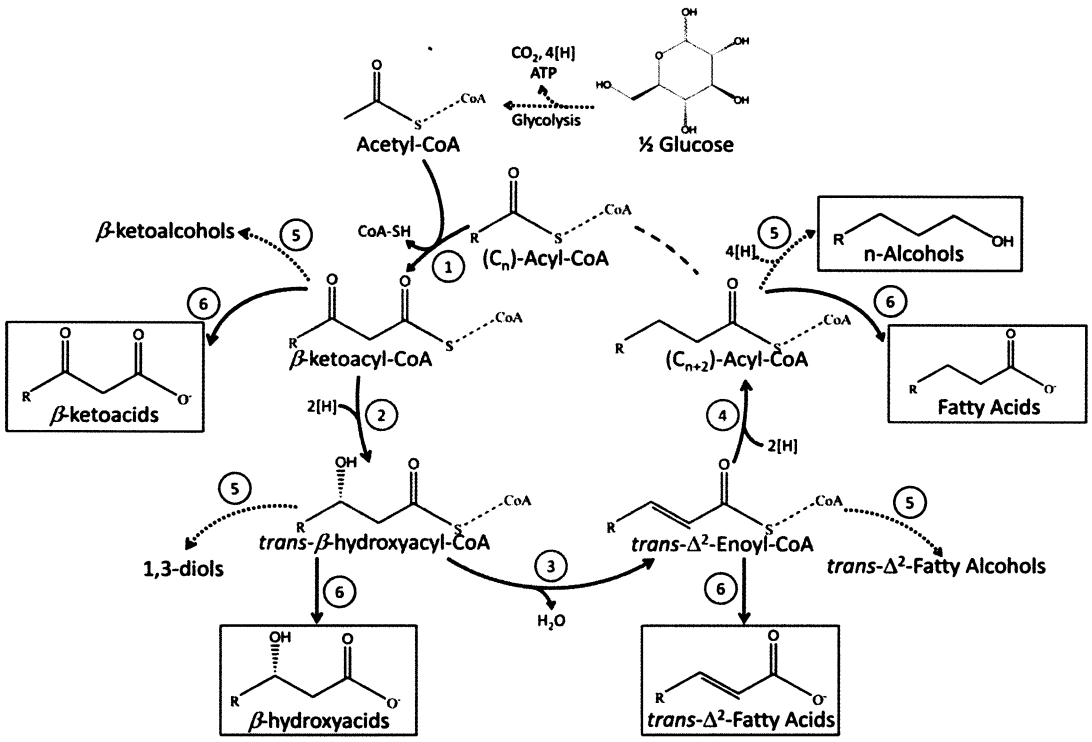

Figure 4-1 Proposed metabolic platform for the combinatorial synthesis of advanced fuels and chemicals

The functional reversal of the β-oxidation cycle engineered in this work (shown in blue) is composed of the following enzymes (gene names in parentheses): ① thiolase (*yqeF*, *fadA*); ② hydroxyacyl-CoA dehydrogenase (*fadB*); ③ enoyl-CoA hydratase (*fadB*); ④ enoyl-CoA reductase (*ydiO*). Intermediates in the engineered pathway can be converted to a functionally diverse set of molecules using aldehyde-forming acyl-CoA reductases & alcohol dehydrogenases (⑤) and acyl-CoA thioesterases (⑥), as indicated. Products whose synthesis was demonstrated in this study are shown in boxes. Abbreviations: R, side chain attached to the acyl-CoA group of the primer or starter molecule (e.g. R=H for acetyl-CoA and R=CH_3 for propionyl-CoA); 2[H], NADH/NADPH/Ferredoxin. Dotted lines indicate multiple steps while dashed lines without arrowheads connect identical metabolites of different chain length.

Here we demonstrate that a functional reversal of the β-oxidation cycle can be used as a metabolic platform for the combinatorial synthesis of alcohols and carboxylic acids with various chain lengths and functionalities (Figure 4-1).

This pathway operates with coenzyme-A (CoA) thioester intermediates and directly uses acetyl-CoA for acyl-chain elongation (rather than first requiring ATP-dependent activation to malonyl-CoA), characteristics that enable product synthesis at maximum carbon and energy efficiency. The reversal of the β-oxidation cycle was engineered in *Escherichia coli* and used in combination with endogenous dehydrogenases and thioesterases to synthesize n-alcohols, fatty acids and β-hydroxy-, β-keto- and trans-Δ^2-carboxylic acids. The superior nature of the engineered pathway was demonstrated by producing higher-chain linear n-alcohols ($C \geq 4$) and extracellular long-chain fatty acids ($C > 10$) at higher efficiency than previously reported (Atsumi et al., 2008b; Bond-Watts et al., 2011; Lennen et al., 2010; Shen and Liao, 2008; Steen et al., 2010). The ubiquitous nature of β-oxidation, aldehyde/alcohol dehydrogenase, and thioesterase enzymes has the potential to enable the efficient synthesis of these products in other industrial organisms.

4.2. Methods

4.2.1. Strains, plasmids and genetic methods

Wild-type K12 *Escherichia coli* strain MG1655 was used as the host to implement metabolic engineering strategies. Gene knockouts were introduced in MG1655 and its derivatives by P1 phage transduction as previously described (Yazdani and Gonzalez, 2008). The native cAMP receptor protein (CRP) was replaced with a

cAMP-independent CRP (CRP*) (Cirino et al., 2006) by introducing the *crp** gene into *crp* deletion mutants via P1 phage transduction using a lysate from strain ET25 (Eppler and Boos, 1999), followed by selection on (1%) glycerol minimal media plates with tetracycline. All resulting strains are listed in Table 4-1.

4.2.1.1. Genetic Methods

Standard recombinant DNA procedures were used for gene cloning, plasmid isolation, and electroporation (Sambrook and Russell, 2001). Manufacturer protocols and standard methods were followed for DNA purification (Qiagen, CA, USA), restriction endonuclease digestion (New England Biolabs, MA, USA), and DNA amplification (Stratagene, CA, USA and Invitrogen, CA, USA). For plasmid construction, genes were amplified from MG1655 genomic DNA using primers designed to create 15 bp of homology on each end of the gene insert for subsequent recombination into the desired plasmid.

Plasmids were linearized using restriction endonuclease digestion, then recombined with the appropriate gene(s) using an In-Fusion Dry-Down PCR Cloning Kit (Clontech, Mountain View, CA, USA) and subsequently used to transform chemically competent Fusion Blue cells (Clontech, Mountain View, CA, USA). Transformants that grew on LB plates containing the appropriate antibiotic were struck for isolation, and then subjected to preliminary screening by PCR. Colonies passing preliminary inspection were then individually grown for plasmid purification.

Table 4-1 Strains and plasmids used in this study

Strain/Plasmid	Description/Genotype	Source
Strains		
MG1655	F- λ- ilvG- rfb-50 rph-1	(Kang et al., 2004a)
fadR atoC(c)	MG1655 *fadR atoC(con)*	(Dellomonaco et al., 2010b)
RB01	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp**	This study
RB02	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA*	This study
RB02 Δ*yqhD*	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA* Δ*yqhD*	This study
RB02 Δ*eutE* Δ*yqhD*	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA* Δ*eutE* Δ*yqhD*	This study
RB02 Δ*yqeF*	MG1655 f*adR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA* Δ*yqeF*	This study
RB02 Δ*fadB*	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA* Δ*fadB*	This study
RB02 Δ*ydiO*	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA* Δ*ydiO*	This study
RB02 Δ*fucO*	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA* Δ*fucO*	This study
RB03	MG1655 *fadR atoC(con)* Δ*arcA* Δ*crp*::*crp** Δ*adhE* Δ*pta* Δ*frdA* Δ*fucO* Δ*yqhD* Δ*fadD*	This study
Δ*adhE* Δ*pta* Δ*frdA*	MG1655 Δ*adhE* Δ*pta* Δ*frdA*	This study
Δ*adhE* Δ*pta* Δ*frdA* Δ*fadD*	MG1655 Δ*adhE* Δ*pta* Δ*frdA* Δ*fadD*	This study
Plasmids		
pTrcHis2A	pTrcHis2A (pBR322-derived), oriR pMB1, *lacI*^{q}, *bla*	Invitrogen (Carlsbad, CA)
pTH *fadBA*	*E. coli fadBA* genes under *trc* promoter and *lacI*^{q} control in pTrcHis2A	This study
pTH *fadBA.fadM*	*E. coli fadBA, fadM* genes under *trc* promoter and *lacI*^{q} control in pTrcHis2A	This study
pTH *fadM*	*E. coli fadM* gene under *trc* promoter and *lacI*^{q} control in pTrcHis2A	This study
pTH *yqeF*	*E. coli yqeF* gene under *trc* promoter and *lacI*^{q} control in pTrcHis2A	This study
pZS blank	oriR pSC101*, tetR, *cat,* contains $P_{LtetO-1}$	(Yazdani and Gonzalez, 2008)
pZS *atoB*	*E. coli atoB* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *betA*	*E. coli betA* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *eutG*	*E. coli eutG* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *fadB*	*E. coli fadB* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *fadM*	*E. coli fadM* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *fucO*	*E. coli fucO* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *tesA*	*E. coli tesA* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *tesB*	*E. coli tesB* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *yciA*	*E. coli yciA* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *ydiO*	*E. coli ydiO* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *yiaY*	*E. coli yiaY* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study
pZS *yqhD*	*E. coli yqhD* gene under control of $P_{LtetO-1}$(tetR, oriR SC101*, cat)	This study

Purified plasmids were confirmed to have the appropriate insert both by PCR as well as restriction endonuclease digest verification. Plasmids in each case include the

plasmid promoter, a ribosomal binding site for each gene, MG1655 gene(s), and a plasmid terminator. Resulting plasmids are listed in Supplementary

Table 4-2 Comparison of *crp*, *fadR*, and *atoSC* loci of engineered strain RB02 (*fadR atoC(c)* Δ*arcA* Δ*crp*::*crp Δ*adhE* Δ*pta* Δ*frdA*) and wild-type MG1655**

Gene locus	Accession # for RB02 sequence	Mutations/insertions in RB02 sequence	Comments
crp	BankIt1445305 Seq1 JF781281	I113L, T128I, A145T	Mutations collectively reduce dependence on cAMP for activation of catabolic genes by CRP, as previously described (Eppler and Boos, 1999; Khankal et al., 2009).
fadR	BankIt1446148 Seq1 JF793627	IS5 insertion between bp 395-396 of *fadR* gene	Inactivation of *fadR* by IS5 insertion. Characteristic phenotype of *fadR* inactivation previously confirmed in strain *fadR atoC(c)* (Dellomonaco et al., 2010b)
atoSC	BankIt1445920 Seq1 JF793626	I129S	*atoC* tranduced from LS5218 (constitutive *atoC* expression(Spratt et al., 1981a)). Characteristic phenotype of constitutive expression of *ato* operon confirmed in strain *fadR atoC(c)* (Dellomonaco et al., 2010b).

4.2.1.2 Sequencing of regulators

Chromosomal DNA from strain RB02 served as the template for PCR using Phusion High-Fidelity DNA polymerase (Finnzymes, Espoo, Finland). PCR products, containing the gene(s) of interest as well as approximately 500 base pairs upstream of each gene, were purified using a QIAquick PCR Purification Kit (Qiagen, Valencia, CA) and sent to Lone Star Labs (Houston, TX) for sequencing. Sequencing results were

analyzed using Sequence Scanner v1.0 (Applied Biosystems, Foster City, CA) and Serial Cloner v2.1 (Serial Basics, http://serialbasics.free.fr/Home/Home.html). Accession numbers for genes sequenced along with the mutations identified are reported in Table 4-2.

4.2.2. Culture medium and cultivation conditions

Chemicals were obtained from Fisher Scientific (Pittsburg, PA) and Sigma-Aldrich Co. (St. Louis, MO).

The minimal medium designed by Neidhardt et al. (Neidhart et al., 1974) with Na_2HPO_4 in place of K_2HPO_4 and supplemented with 20 g/L glucose, 40 g/L calcium bicarbonate, 100 μM $FeSO_4$, 5 mM calcium pantothenate, 3.96 mM Na_2HPO_4, 5 mM $(NH_4)_2SO_4$, and 30 mM NH_4Cl was used. Fermentations conducted in the SixFors multi-fermentation system also included 1 mM betaine. Ampicillin (100 μg/mL), kanamycin (50 μg/mL), chloramphenicol (34 μg/mL), anhydrotetracycline (100 ng/mL), and isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.1 mM) were included when appropriate. A single colony was used to inoculate 30 mL of medium in 50 mL baffled flasks and incubated at 37° C in a rotary shaker (200 rpm) for 24 hours. An appropriate amount of pellet from this pre-culture was used to inoculate 30 mL of fresh medium in 50 mL baffled flasks to a target initial optical density at 550 nm of 0.05. These cultures were then incubated at 30° C in a rotary shaker (200 rpm) for 24 hours, unless otherwise stated. The same inoculation procedure was used for additional fermentations conducted in a SixFors multi-fermentation system (Infors HT, Bottmingen, Switzerland) with an air

flowrate of 1.5 N L/hr and control of dissolved oxygen, pH, and temperature at specified values.

4.2.3. Analytical methods

4.2.3.1. Metabolite Identification

The identity of metabolic products was determined through one-dimensional (1D) proton nuclear magnetic resonance (NMR) spectroscopy. 60 μL of D_2O and 1 μL of 600 mM NMR internal standard TSP [3-(trimethylsilyl) propionic acid-D4, sodium salt] were added to 540 μL of the sample (culture supernatant). The resulting solution was then transferred to a 5 mm-NMR tube, and 1D proton NMR spectroscopy was performed at 25°C in a Varian 500-MHz Inova spectrometer equipped with a Penta probe (Varian, Inc., Palo Alto, CA) using the following parameters: 8,000-Hz sweep width, 2.8-s acquisition time, 256 acquisitions, 6.3-μs pulse width, 1.2-s pulse repetition delay, and presaturation for 2 s. The resulting spectrum was analyzed using FELIX 2001 software (Accelrys Software Inc., Burlington, MA). Peaks were identified by their chemical shifts and J-coupling values, which were obtained in separate experiments in which samples were spiked with metabolite standards (2 mM final concentration).

Identification of n-alcohols was conducted through gas chromatography-mass spectroscopy (GC-MS) following a modification of the method reported by Atsumi and coworkers (Atsumi et al., 2008a). The analysis was performed on an Agilent 6890 GC/5973 MS (Agilent Technologies, Palo Alto, CA) instrument with a HP-5ms capillary column (30 m × 0.25 mm × 0.25 μm). 1ml of supernatant of culture broth was extracted

with 500 μl of GC standard grade n-hexane (Fluka, Steinheim, Germany) and 1-heptanol (Sigma-Aldrich, Lausanne, Switzerland) was used as internal standard. 0.5 μl of the extracted sample was injected using a 20:1 split at 250 °C. The oven temperature was initially held at 75°C for 2 min and then raised with a gradient of 5°C/min to 280°C and held for 2 min. Helium (Matheson Tri-Gas, Longmont, CO) was used as the carrier gas with a 14-lb/in^2 inlet pressure. The injector and detector were maintained at 255°C.

Identification of fatty acids was performed on a Shimadzu Auto-System GC 2010 (Shimadzu, Japan) equipped with a DB-5MS capillary column (30 m × 0.25 mm × 0.25 μm) and coupled with a QP-2010 mass detector. The following method was used: an initial temperature of 50°C was held for 2 min and then ramped to 220 °C at 4 °C per min and held for 10 min^2. Extraction and derivatization procedures are described in Section 4.2.3.2.

4.2.3.2. Metabolite Quantification

The quantification of glucose, organic acids, ethanol, and butanol was conducted by high-performance liquid chromatography (HPLC). Samples (culture supernatant) were analyzed with ion-exclusion HPLC using a Shimadzu Prominence SIL 20 system (Shimadzu Scientific Instruments, Inc., Columbia, MD) equipped with an HPX-87H organic acid column (Bio-Rad, Hercules, CA) with operating conditions to optimize peak separation (0.3 mL/min flowrate, 30 mM H_2SO_4 mobile phase, column temperature 42 °C)(Dharmadi and Gonzalez, 2005).

Quantification of longer chain (C > 4) n-alcohols was conducted through gas chromatography (GC) in a Varian CP-3800 gas chromatograph (Varian Associates, Inc., Palo Alto, CA) equipped with a flame ionization detector (GC-FID). Sample extraction procedure was as described above in Section 4.3.3.1. The separation of alcohol compounds was carried out using a FactorFour™ - VF-5ht column (15 m, 0.32-μm internal diameter, 0.10-μm film thickness; Varian Associates, Inc., Palo Alto, CA). The oven temperature was initially held at 40°C for 1 min and then raised with a gradient of 30°C/min to 130°C and held for 4 min. The temperature was then raised with a gradient of 15°C/min to 230°C and held for 4 min. Helium (1 ml min^{-1}, Matheson Tri-Gas, Longmont, CO) was used as the carrier gas. The injector and detector were maintained at 250°C. A 0.5-μl sample was injected in splitless injection mode.

For total free fatty acids quantification, 5 mL of culture was placed in a 15 mL Corning CentriStar centrifuge tube (Corning Inc., Lowell, MA) in an ice bath and sonicated using a Branson Sonifier 450 (Emerson Industrial Automation, Danbury, CT) for 5 minutes at a duty cycle of 30% and a power setting of 1. The tubes were then centrifuged at 12,000 g for 5 minutes and the appropriate volume of supernatant recovered. For extracellular free fatty acid analysis, the supernatant from non-sonicated culture samples was used instead. 2 ml of supernatant were acidified with 200 μl of acetic acid, supplemented with 100 mg of octanoic acid as an internal standard, and extracted with 2 ml of n-hexane-chlorophorm (4:1, v/v) (Lalman and Bagley, 2004). The organic layer was nitrogen evaporated to near dryness, redissolved in 1 ml of a mixture of cholophorm:methanol:hydrochloric acid [10:1:1, vol/vol/vol] (Dellomonaco et al., 2010b)

and incubated in a sealed vial at 90°C for 1 h. Fatty acid methyl esters were extracted with 1 ml of n-hexane-chlorophorm (4:1, v/v) after addition of 500 µl of 0.9% (wt/vol) NaCl in water. Gas-liquid chromatography was performed with 1µl of the n-hexane:chlorophorm solution on a Varian FactorFour™ -VF-5ht capillary column (Varian Associates, Inc., Palo Alto, CA, 15 m length, 0.32 µm diameter, and 0.1 µm film thickness) according to the following method: 50°C held for 1 min, 30°C/min to 160°C, 15°C/min to 200°C, 200°C held for 1.5 min, 10°C/min to 225°C, and 225°C held for 15 min(Dellomonaco et al., 2010b).

4.2.3.3. Enzyme Assays

For measurement of enzymatic activities, cells from 24 hour shake flask cultures were washed twice with 9 g/L sodium chloride under anaerobic conditions and stored at -80° C until use. Cell extracts for all assays were prepared as follows under anaerobic conditions. 40 units of $OD_{550\ nm}$ were re-suspended in 1 mL of 100 mM Tris–HCl buffer (pH 7.0) with 1mM DTT. After cellular disruption using a Disruptor Genie (Scientific Industries, Inc., Bohemia, NY), cellular debris was removed by centrifugation (13,000 x *g*, 4 °C, 10 min) and the supernatant used as cell extract. Absorbance changes for all assays were monitored in a Biomate 5 spectrophotometer (Thermo Scientific, MA, USA). The linearity of reactions (protein concentration and time) was established for all assays and the non-enzymatic rates were subtracted from the observed initial reaction rates. Enzymatic activities are reported as µmol of substrate per minute per mg of cell protein and represent averages for at least three cell preparations. Protein concentration was

measured using the Bradford Assay Reagent (Thermo Scientific, MA, USA) with BSA as a standard.

Acetyl-CoA acetyltransferase (THL) activity was determined using acetoacetyl-CoA and CoA as substrates, and the decrease in acetoacetyl-CoA concentration was measured at 303 nm (Wiesenborn et al., 1988). β-Hydroxybutyryl-CoA dehydrogenase activity was measured at 340 nm by monitoring the decrease in NADH concentration resulting from β-hydroxybutyryl-CoA formation from acetoacetyl-CoA (Hartmanis and Gatenbeck, 1984). Crotonase activity was measured by monitoring the decrease in crotonyl-CoA concentration at 263 nm, which results from β-hydroxybutyryl-CoA formation from crotonyl-CoA (Hartmanis and Gatenbeck, 1984). Butyryl-CoA dehydrogenase activity was assayed in the direction of crotonyl-CoA reduction by monitoring the ferricenium ion at 300 nm, which acts as an electron donor (Lehman et al., 1990). In addition, assays in which the ferricenium ion was replaced with 0.2 mM NAD(P)H and the absorbance measured at 340 nm were also run. Butyraldehyde dehydrogenase activity was assayed in the direction of butyraldehyde oxidation by monitoring $NAD(P)^+$ reduction at 340 nm (Berezina et al., 2010). To measure butanol dehydrogenase activity, the decrease in NAD(P)H concentration resulting from butanol formation from butyraldehyde is monitored at 340 nm under anaerobic conditions at 30 °C (Durre et al., 1987).

4.3. Results

Engineering the β-oxidation cycle as a metabolic platform for the combinatorial synthesis of advanced fuels and chemicals requires the reverse operation of this pathway in the absence of its natural substrate (i.e. fatty acids) and presence of a non-fatty acid carbon source (e.g. glucose). Constitutive expression of the *fad* and *ato* regulons, which encode the β-oxidation system in *Escherichia coli* (Clark and Cronan, 2005), in the absence of fatty acids was achieved by introducing previously reported *fadR* and *atoC(c)* mutations (Dellomonaco et al., 2010b). The *fadR* and *atoC(c)* mutations are due to an IS5 insertion and a point mutation (I129S), respectively (Table 4-2). Several operons encoding β-oxidation cycle enzymes are also activated by the cyclic-AMP (cAMP) receptor protein (CRP)-cAMP complex and therefore subjected to carbon catabolite repression in the presence of glucose (Deutscher, 2008). This regulatory mechanism was circumvented by replacing the native *crp* gene with a cAMP-independent mutant (*crp**) that contains three point mutations previously reported to confer a catabolite-derepressed phenotype (Eppler and Boos, 1999) (Table 4-2). Finally, since anaerobic/microaerobic conditions used in the production of fuels and chemicals would lead to ArcA-mediated repression of most operons encoding the β-oxidation cycle (Cho et al., 2006), the *arcA* gene was deleted.

Table 4-3 Butanol synthesis, glucose utilization, and cell growth in strain RB02 and its derivatives

Strain	Butanol produced		Glucose utilized	Cell growth
	Yield (g/g)	Concentration (g/L)	(g/L)	(g/L)
RB02 [*yqeF*+ *fucO*+]	0.182	1.85	10.19	0.72
Reaction ①: YqeF (predicted acyltransferase)				
RB02 Δ*yqeF* [*fucO*+]	0.019	0.10	5.19	0.32
RB02 [*fucO*+]	0.063	0.23	3.66	0.41
RB02 Δ*yqeF* [*fucO*+ *yqeF*+]	0.159	1.11	7.00	0.43
Reactions ② and ③: FadB (3-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase)				
RB02 Δ*fadB* [*yqeF*+ *fucO*+]	ND[b]	ND[b]	2.96	0.19
RB02 Δ*fadB* [*yqeF*+ *fadB*+]	0.157	0.92	5.86	0.52
RB02 Δ*fadB* [*yqeF*+]	ND[b]	ND[b]	1.84	0.11
Reaction ④: YdiO (predicted acyl-CoA dehydrogenase/enoyl-CoA reductase)				
RB02 Δ*ydiO* [*yqeF*+ *fucO*+]	0.038	0.12	3.14	0.18
RB02 Δ*ydiO* [*yqeF*+ *ydiO*+]	0.163	0.97	5.94	0.57
RB02 Δ*ydiO* [*yqeF*+]	0.018	0.103	1.66	0.13
Reaction ⑤: MhpF (aldehyde-forming acyl-CoA reductase) and FucO (n-butanol dehydrogenase)				
RB02 Δ*fucO* [*yqeF*+]	0.015	0.04	2.75	0.38
RB02 Δ*fucO* [*yqeF*+ *fucO*+]	0.151	0.85	5.64	0.51
RB02 [*yqeF*+]	0.088	0.35	4.00	0.68
RB02 Δ*mhpF* [*yqeF*+ *fucO*+]	ND[b]	ND[b]	1.16	0.16

[a] THL: thiolase; HBD: hydroxy-acyl-CoA dehydrogenase; CRT: crotonase; BDH: butanol dehydrogenase.
[b] ND, not detectable. Minimum detection level for butanol was 5.84 mg l^{-1}.
[c] Experiments were run for 24 hours in shake flasks using glucose (1% w/v) minimal medium at 30 °C.

The genetic manipulations outlined above were implemented in wild-type *E. coli* MG1655 (resulting strain named RB01: MG1655 *fadR atoC(c) crp** Δ*arc*A) and in a strain in which the consumption of acetyl-CoA by other fermentative pathways was minimized (resulting strain named RB02: MG1655 *fadR atoC(c) crp** Δ*arc*A Δ*adhE*

Δ*pta* Δ*frd*A). Despite these modifications, none of the products that can be derived from the β-oxidation cycle intermediates (Figure 4-1) were synthesized (Table 4-3). A limitation on the expression of β-oxidation enzymes was ruled out after confirming the presence of these activities in strain RB02 (Table 4-4).

Table 4-4 Activities of β-oxidation and butanol dehydrogenase enzymes in wild-type and engineered strains

Strain	Enzyme activity (μmol/mg protein/min) ± standard deviation			
	THL[a]	HBD[a]	CRT[a]	BDH[a]
MG1655	ND[b]	0.002±0.000	ND[b]	0.014±0.001
RB02 (*fadR atoC(c) crp** Δ*arcA* Δ*pta* Δ*adhE* Δ*frdA*)	0.310±0.079	0.304±0.032	0.339±0.049	0.004±0.002
RB02 [yqeF+ fucO+]	0.498±0.036	0.292±0.013	0.334±0.017	0.298±0.020

[a]THL: thiolase; HBD: hydroxy-acyl-CoA dehydrogenase; CRT: crotonase; BDH: butanol dehydrogenase.
[b] ND, not detectable. Minimum detectable activity was 0.001 μmol/mg protein/min.

The expression of appropriate termination enzymes is also vital to achieve the conversion of the CoA-thioester intermediates in the reversal of the β-oxidation cycle to the desired product(s) (Figure 4-1). We first targeted the synthesis of n-butanol, as this product can be generated in a one-turn reversal of the β-oxidation cycle. Unlike the significant activity observed for key β-oxidation enzymes (Table 4-4), the level of n-butanol dehydrogenase in strain RB02 was very low (Table 4-4), probably limiting n-butanol synthesis. To address this issue, two endogenous aldehyde/alcohol dehydrogenases with high sequence and structure similarity to the Clostridial butyraldehyde/butanol dehydrogenase were expressed in strain RB02: i.e. L-1,2-

propanediol oxidoreductase (FucO) and an aldehyde/alcohol dehydrogenase (YqhD) (Table 4-5). These modifications enabled a functional one-turn reversal of the β-oxidation cycle, as evidenced by the synthesis of n-butanol in strains RB02 [*yqhD*+] and RB02 [*fucO*+] (Figure 4-2). Despite the potential for YqhD to catalyze the conversion of butyraldehyde to n-butanol (Jarboe, 2010), overexpression of FucO led to higher concentration of n-butanol and lower concentration of the major fermentation by-product, ethanol (Figure 4-2a).

Since the initiation of the reversal of the β-oxidation cycle (i.e. conversion of acetyl-CoA to acetoacetyl-CoA) is another key step in the engineered pathway, we evaluated the effect of overexpressing *atoB*-encoded acetyl-CoA acetyltransferase, which has high specificity for short-chain acyl-CoA molecules (Feigenbaum and Schulz, 1975; Jenkins and Nunn, 1987a), and a predicted acyltransferase (YqeF) with high sequence similarity to AtoB. These modifications also supported the functioning of a reversal of the β-oxidation cycle, as evidenced by the synthesis of n-butanol in strains RB02 [*atoB*+] and RB02 [*yqeF*+] (Figure 4-2a).

Table 4-5 *In silico* identification of *Escherichia coli* surrogates for higher-chain (C ≥ 4) aldehyde-forming acyl-CoA reductases and aldehyde/alcohol dehydrogenases (reaction ⑤ in Figure 4-1). Genes shown in bold were tested in this study.

Source of gene sequences used to identify *E. coli* surrogates					Surrogates identified via I-TASSER (Roy et al., 2010)				Surrogates identified via protein BLAST(Altschul et al., 1997)				
Organism	**Accession #**	**Protein**	**E.C. #**	**Function**	**E.C. #**	**TM-Score[a]**	**EC-Score[b]**	**Protein**	**E.C. #**	**Protein**	**E-value[c]**	**Identity**	**Similarity**
Identification of E. coli surrogates for higher-chain (C ≥ 4) fatty aldehyde-forming acyl-CoA reductase													
Clostridium saccharobutylicum (Berezina et al., 2009)	P13604	Adh1	1.1.1.-	NADH-dependent butyraldehyde/ butanol dehydrogenase	1.1.1.77	0.9404	2.2424	**FucO**	1.1.1.1.	AdhE	2.0E-91	42%	62%
					1.1.1.1.	0.8679	1.8069	YiaY	1.1.1.77	**FucO**	5.0E-66	35%	58%
								AdhE	1.1.1.-	YiaY	2.0E-65	37%	56%
								AdhP	1.1.-.-	EutG	4.0E-61	35%	54%
								FrmA	1.1.1.-	**YqhD**	6.0E-28	25%	46%
					1.1.1.6	0.7979	1.4449	GldA					
Pseudomonas sp. strain CF600 (Powlowski et al., 1993)	Q52060	DmpF	1.2.1.10	Acetaldehyde dehydrogenase	1.2.1.12	0.7439	1.3969	GapA	1.2.1.10	**MhpF**	6.0E-146	79%	92%
Acinetobacter sp. Strain M-1 (Liu et al., 2009)	Q8RR58	AcrM	1.2.1.50	Acyl coenzyme A reductase			<1.1[b]		1.-.-.-	UcpA	5.0E-20	31%	49%
									1.-.-.-	YbbO	4.0E-19	30%	50%
									1.1.1.-	YdfG	6.0E-18	27%	47%
									1.1.1.69	IdnO	9.0E-18	29%	49%
									1.1.1.100	FabG	9.0E-17	31%	51%
Acinetobacter calcoaceticus (Reiser and Somerville, 1997)	P94129	Acr1	1.2.1.2	Fatty acyl-CoA reductase			<1.1[b]		1.1.1.100	FabG	2.0E-18	31%	53%
									1.1.1.69	IdnO	3.0E-17	29%	48%
									1.1.1.-	YdfG	1.0E-16	27%	47%
									1.-.-.-	YbbO	1.0E-16	28%	44%
									1.-.-.-	UcpA	2.0E-16	31%	52%
Identification of E. coli surrogates for higher-chain (C ≥ 4) aldehyde/alcohol dehydrogenases													
Clostridium acetobutylicum (Walter et al., 1992)	Q04944	BdhA	1.1.1.-	NADH-dependent butanol dehydrogenase	1.1.1.77	0.9072	1.7913	**FucO**	1.1.1.-	**YqhD**	3.0E-68	36%	55%
					1.1.1.1	0.8465	1.7242	YiaY	1.1.1.-	YiaY	1.0E-32	26%	46%
								AdhE	1.1.-.-	EutG	5.0E-27	27%	42%

							AdhP	1.1.1.77	FucO	7.0E-27	26%	46%
							FrmA	1.1.1.1.	AdhE	1.0E-24	24%	45%
					1.1.1.6	0.7761	1.4418	GldA					
Clostridium acetobutylicum (Walter et al., 1992)	Q04945	BdhB	1.1.1.-	NADH-dependent butanol dehydrogenase	1.1.1.77	0.9039	1.7305	**FucO**	1.1.1.-	**YqhD**	4.0E-69	39%	57%
					1.1.1.1	0.8443	1.7044	YiaY	1.1.1.-	YiaY	9.0E-30	26%	44%
								AdhE	1.1.1.77	FucO	4.0E-25	26%	44%
								AdhP	1.1.-.-	EutG	1.0E-22	27%	42%
								FrmA	1.1.1.1.	AdhE	7.0E-18	22%	44%
					1.1.1.6	0.7747	1.4652	GldA					
Geobacillus thermodenitrificans (Liu et al., 2009)	A4IS49	GTNG_2878	1.1.1.-	Long-chain alkyl alcohol dehydrogenase	1.1.1.77	0.9106	1.776	**FucO**	1.1.1.1	YqhD	6.0E-75	41%	59%
					1.1.1.1	0.8484	1.7205	YiaY	1.1.1.1	**YiaY**	7.0E-26	25%	42%
								AdhE	1.1.1.77	FucO	2.0E-24	26%	44%
								AdhP	1.1.1.1.	AdhE	2.0E-20	26%	42%
								FrmA	1.1.-.-	**EutG**	5.0E-20	25%	40%
					1.1.1.6	0.7781	1.4072	GldA					
Geobacillus thermodenitrificans (Liu et al., 2009)	A4IP64	GTNG_1754	1.1.1.-	Alcohol Dehydrogenase	1.1.1.202	0.9565	2.5069	**YqhD**	1.1.-.-	**EutG**	3.0E-57	39%	57%
					1.1.1.77	0.9312	2.1532	FucO	1.1.1.1	**YiaY**	3.0E-54	33%	52%
					1.1.1.1.	0.8512	1.9177	**YiaY**	1.1.1.77	FucO	2.0E-52	34%	53%
								AdhE	1.1.1.1.	AdhE	3.0E-43	34%	52%
								AdhP	1.1.1.1	YqhD	1.0E-18	28%	46%
								FrmA					
					1.1.1.6	0.7715	1.5259	GldA					
Pseudomonas oleovorans (Vanbeilen et al., 1992)	Q00593	AlkJ	1.1.99.-	Alcohol Dehydrogenase					1.1.99.1	**BetA**	2.0E-104	40%	59%
Thermococcus sp. ES (Ying et al., 2009)	C1IWT4	Adh	1.1.1.1.	Iron alcohol dehydrogenase	1.1.1.202	0.9301	2.2696	YqhD	1.1.1.1	**YiaY**	7.0E-40	33%	50%
					1.1.1.77	0.9041	1.9642	FucO	1.1.1.1	YqhD	4.0E-30	30%	46%
					1.1.1.1.	0.8352	1.8286	**YiaY**	1.1.-.-	**EutG**	3.0E-27	30%	46%
								AdhE	1.1.1.1.	AdhE	9.0E-27	29%	48%
								AdhP					
								FrmA					
					1.1.1.6	0.7800	1.3793	GldA					
Thermococcus hydrothermalis (Antoine et al., 1999)	Y14015		1.1.1.-	Alcohol Dehydrogenase					1.1.1.1	**YiaY**	4.0E-37	31%	48%

								1.1.-.-	**EutG**	2.0E-34	31%	47%
								1.1.1.1.	AdhE	1.0E-30	30%	50%
								1.1.1.77	FucO	4.0E-30	30%	45%
								1.1.1.1	YqhD	5.0E-19	30%	44%
Sulfolobus tokodaii (Yanai et al., 2009)	Q976Y8	ST0053	Hypothetical alcohol dehydrogenase	1.1.1.1.	0.9655	1.8286	**YiaY**	1.1.1.1.	AdhP	4.0E-33	31%	50%
							AdhE	1.-.-.-	YydjJ	8.0E-26	30%	50%
							AdhP	1.-.-.-	YphC	2.0E-21	31%	47%
							FrmA	1.-.-.-	YahK	2.0E-21	29%	45%
								1.1.1.-	RspB	6.0E-21	29%	49%
								1.1.1.103	Tdh	9.0E-20	27%	47%
								1.-.-.-	YjjN	3.0E-19	26%	45%
								1.1.-.-	GatD	1.0E-18	28%	44%

[a] The Template Modeling-score (TM-score) is defined to assess the topological similarity of protein structure pairs. Its value ranges between 0 and 1, and a higher score indicates better structural match. Statistically, a TM-score < 0.17 means a randomly selected protein pair with the gapless alignment taken from PDB (Roy et al., 2010).

[b] An EC-score > 1.1 is a good indicator of the functional similarity between the query and the identified enzyme analogs (Roy et al., 2010).

[c] The BLAST E-value measures the statistical significance threshold for reporting protein sequence matches against the organism genome database; the default threshold value is 1E-5, in which 1E-5 matches would be expected to occur by chance, according to the stochastic model of Karlin and Altschul (http://www.ncbi.nlm.nih.gov/BLAST/tutorial/).

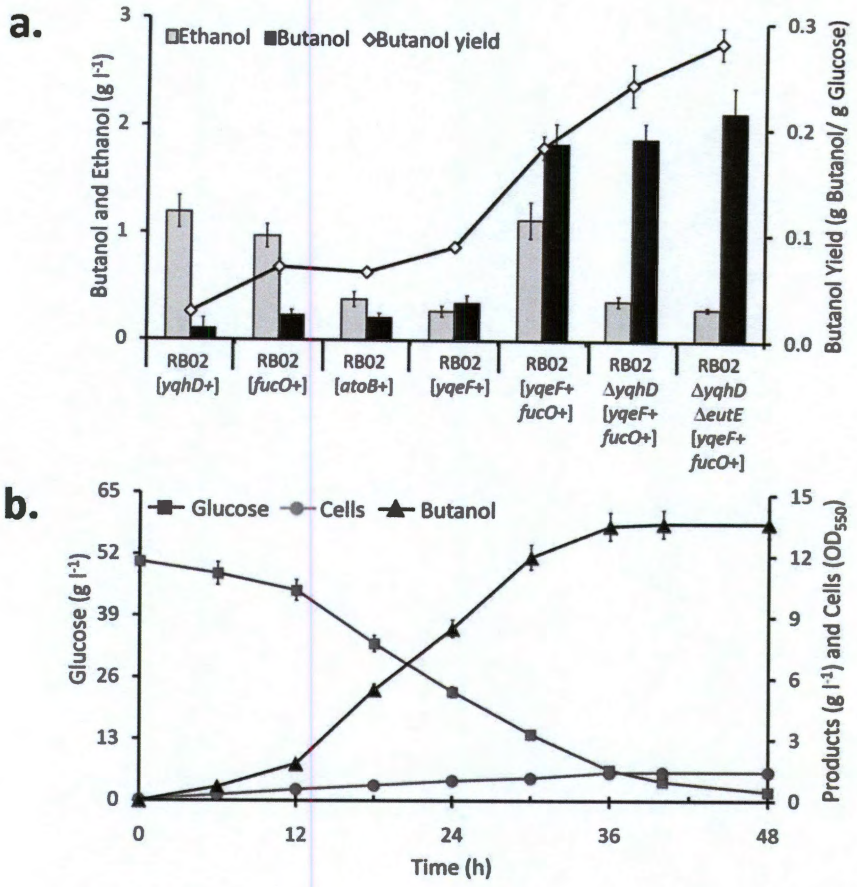

Figure 4-2 Engineered one-turn reversal of the β-oxidation cycle for the synthesis of n-butanol

(a) Effect of gene overexpressions and knockouts (indicated underneath *x* axis) on the synthesis of n-butanol and ethanol in strain RB02 (*fadR atoC(c) crp* ΔarcA ΔadhE Δpta ΔfrdA*). Experiments were performed at 30 °C for 24 hours in shake flasks using glucose (1% w/v) minimal medium. The n-butanol yield was calculated as g n-butanol/g total glucose consumed. (b) Kinetics of n-butanol production by strain RB02 Δ*yqhD* Δ*eutE* [*yqeF+fucO+*]. Cells were cultivated in fermentors containing minimal medium supplemented with 5 % (w/v) glucose. The dissolved oxygen was controlled at 5% of saturation, temperature at 30 °C, and pH at 7.

Overexpression of YqeF, whose function in *E. coli* metabolism is currently unknown, gave better results. The simultaneous overexpression of *yqeF* and *fucO* in strain RB02 yielded high n-butanol titer (1.9 g/L), n-butanol yield (0.19 g n-butanol/g total glucose consumed) and n-butanol-to-ethanol ratio (>5:1) (Figure 4-2a). No n-

butanol synthesis was observed upon overexpression of FucO and YqeF in wild-type MG1655 (Table 4-3), demonstrating the need for an active reversal of the β-oxidation cycle.

Further reduction in the synthesis of by-product ethanol was realized through deletion of *eutE* (encoding an aldehyde dehydrogenases with high sequence similarity to AdhE) and *yqhD* (YqhD overexpression favored the synthesis of ethanol over n-butanol: Figure 4-2a). The resulting strain (RB02 Δ*yqhD* Δ*eutE* [*yqeF*+*fucO*+]) synthesized 2.2 g/L of n-butanol in 24 hours at a yield of 0.28 g n-butanol/ g total glucose consumed (Figure 4-2a). When grown in a bioreactor using a higher initial concentration of glucose, this strain produced n-butanol at high titer (~ 14 g/L), yield (0.33 g n-butanol/ g total glucose consumed) and rate (~ 2 g n-butanol/g cell dry weight/h) (Figure 4-2b). This performance, which was achieved in the absence of rich nutrients and without importing foreign genes, is several folds better than reported for other organisms engineered for n-butanol production (Table 4-6) and also surpasses the n-butanol yield and specific productivity reported for native producers (Lee et al., 2008b). The reversal of the β-oxidation cycle engineered in this strain operated at a maximum carbon flux of 73.4 mmol acetyl-CoA/g cell dry weight/h (12-18 hours in Fig. 2b), which exceeds the flux reported in the literature for native or engineered fermentative pathways (Gonzalez et al., 2002; Papagianni et al., 2007).

Table 4-6 Summary of organisms engineered to produce higher-chain (C ≥ 4) linear n-alcohols and long-chain (C ≥ 10) fatty acids

Product	Engineered Host	Titer (g/L)/ Yield (% w/w)[a]	Culture Time (hr)	Carbon Source	Medium/ Cultivation	Reference
		Higher-chain (C ≥ 4) linear n-alcohols				
n-butanol (C4)	*E. coli*	0.55/2.8	24	Glycerol	Rich/Batch	(Atsumi et al., 2008a)
	E. coli	0.82/3.3	100	Glucose	Rich/Batch	(Atsumi et al., 2008b; Shen and Liao, 2008)
	E. coli	1.2/6.1	60	Glucose	Rich-MM/HCD[b]	(Inui et al., 2008)
	E. coli	0.58/11.6	48	Glycerol	Rich/Batch	(Nielsen et al., 2009)
	S. cerevisiae	0.002/0.0001	72	Galactose	Rich/Batch	(Steen et al., 2008)
	B. subtilis	0.024/0.48	72	Glycerol	Rich/Batch	(Nielsen et al., 2009)
	P. putida	0.120/2.4	72	Glycerol	Rich/Batch	(Nielsen et al., 2009)
	L. lactis	0.028/--	--	Glucose	Rich/Batch	(Liu et al., 2010)
	L. buchneri	0.066/--	--	Glucose	Rich/Batch	(Liu et al., 2010)
	L. brevis	0.3/1.5	60	Glucose	Rich/Batch	(Berezina et al., 2010)
	E. coli	2.05/13.4	96	Palmitic Acid	MM/Fed-Batch	(Dellomonaco et al., 2010b)
	E. coli	4.65/28	72	Glucose	Rich/Batch	(Bond-Watts et al., 2011)
	E. coli	14.5/32.8	36	Glucose	MM/Batch	This work
n-hexanol (C6)	*E. coli*	0.04/0.2	40	Glucose	Rich/Batch	(Zhang et al., 2008)
Fatty alcohols (Distribution of C10, C12, C14, and C16)	*E. coli*	0.06/0.3	--	Glucose	MM/Batch	(Steen et al., 2010)
Higher-chain n-alcohols (C6 to C10)	*E. coli*	0.42/8.3	48	Glucose	MM/Batch	(Dellomonaco, 2011b)
		Long-chain (C ≥ 10) fatty acids				
Fatty Acids (Predominantly C14)	*E. coli*	1.2/6	--	Glucose	MM/Batch	(Steen et al., 2010)[c]
Fatty Acids (Wide Distribution)	*E. coli*	2.5/--	22	Glycerol	MM/Fed-Batch/HCD	(Lu et al., 2008)[c]
Fatty Acids (Predominantly C12)	*E. coli*	0.81/16	29	Glycerol	Rich/Batch	(Lennen et al., 2010)[c]
Fatty Acids (Predominantly C16, C18)	*E. coli*	6.6/28	60	Glucose	MM/Batch	(Dellomonaco, 2011b)

[a] For products with carbon length distributions, titer represents the sum of products of all chain length produced. Yield assumes all the carbon source was consumed when carbon source consumption data not given in reference.

[b] Two-phase, high cell density (HCD) culture grown first in rich medium and then incubated in minimal medium (MM).

[c] Titers reported refer to total (i.e. sum of intracellular and extracellular) free fatty acids.

A characterization of the engineered reversal of the β-oxidation cycle was conducted to establish the identity of the enzymes catalyzing each individual step and their role on n-butanol synthesis (Table 4-4). Enzyme activity measurements showed high level of expression of key enzymes involved in the postulated pathway in strain RB02 [*yqeF+fucO+*] and negligible activity in wild type MG1655 (Table 4-4). Gene knockout and gene complementation experiments along with quantification of fermentation products (Table 1b) demonstrated that the primary enzymes involved in the synthesis of n-butanol through the engineered reversal of the β-oxidation pathway are (activities shown in parentheses along with reaction numbers, as per Figure 4-1): YqeF (predicted acyltransferase: Reaction ①), FadB (3-hydroxyacyl-CoA dehydrogenase and enoyl-CoA hydratase: Reactions ② and ③, respectively), YdiO (predicted acyl-CoA dehydrogenase/enoyl-CoA reductase: Reaction ④), and MhpF and FucO (aldehyde-forming acyl-CoA reductase and n-butanol dehydrogenase, respectively: Reaction ⑤).

It is noteworthy that YdiO is proposed here to catalyze the reduction of enoyl-CoA to acyl-CoA (Reaction ④). The reverse of this reaction is catalyzed by FadE and is the only irreversible step in the catabolic operation of the β-oxidation cycle (Clark and Cronan, 2005). In agreement with our proposal, deletion of *ydiO* in strain RB02 [*yqeF+fucO+*] completely abolished n-butanol synthesis (Table 4-3) while a *fadE* deletion did not have a significant effect (data not shown). Although *ydiO* was previously proposed to encode an acyl-CoA dehydrogenase that would replace FadE during the anaerobic catabolism of fatty acids (Campbell and Cronan, 2002), a sequence comparison

between YdiO and proteins encoded in the *E. coli* genome does not reveal a significant similarity to FadE (E-value 0.001 and 81% coverage: Table 4-7).

Table 4-7 Homology analysis and functional annotation of *Escherichia coli ydi* genes

Current annotation[a]		Sequence-based homologues identified via protein BLAST (Altschul et al., 1997)					Functional homologues identified via I-TASSER (Roy et al., 2010)		
Enzyme name	Function	Enzyme	Function	E-value[b]	Coverage	Similarity	Function	TM-Score[c]	EC-Score[d]
YdiO	Predicted acyl-CoA dehydrogenase	CaiA	Crotonobetainyl-CoA reductase	1.0E-102	99%	65%	Butyryl-CoA dehydrogenase	0.9584	2.2964
		AidB	Isovaleryl-CoA dehydrogenase	2.0E-14	64%	47%	Acyl-CoA dehydrogenase	0.9618	2.1730
		FadE	Acyl-CoA dehydrogenase	0.001	81%	37%			
YdiQ	Putative electron transfer flavoprotein subunit	FixA	probable flavoprotein subunit required for anaerobic carnitine metabolism	3.0E-68	99%	71%	Adenosine kinase		<1.1
YdiR	Putative electron transfer flavoprotein subunit	FixB	probable flavoprotein subunit required for anaerobic carnitine metabolism	6.0E-75	100%	64%			
YdiS	Predicted oxidoreductase with FAD/NAD(P)-binding domain	FixC	flavoprotein (electron transport)	5.0E-152	100%	78%	Electron-transferring-flavoprotein dehydrogenase	0.9404	1.8378
		YgcN	predicted oxidoreductase with FAD/NAD(P)-binding domain	9.0E-93	99%	62%			
YdiT	Ferredoxin-like protein	FixX	putative ferredoxin possibly involved in anaerobic carnitine metabolism	3.0E-33	96%	62%	Electron-transferring-flavoprotein dehydrogenase	0.9006	1.5362
		YgcO	predicted 4Fe-4S cluster-containing protein	1.0E-18	89%	62%			

[a] As annotated in Ecocyc(Keseler et al., 2009). Also reported by Campbell, J.W. and coworkers (Campbell et al., 2003).

[b] The BLAST E-value measures the statistical significance threshold for reporting protein sequence matches against the organism genome database; the default threshold value is 1E-5, in which 1E-5 matches would be expected to occur by chance, according to the stochastic model of Karlin and Altschul (http://www.ncbi.nlm.nih.gov/BLAST/tutorial/).

[c] The Template Modeling-score (TM-score) is defined to assess the topological similarity of protein structure pairs. Its value ranges between 0 and 1, and a higher score indicates better structural match. Statistically, a TM-score < 0.17 means a randomly selected protein pair with the gapless alignment taken from PDB (Roy et al., 2010).

[d] An EC-score > 1.1 is a good indicator of the functional similarity between the query and the identified enzyme analogs(Roy et al., 2010).

In contrast, YdiO shares high homology with crotonobetainyl-CoA reductase (CaiA; E-value 1.0E-112 and 99% coverage: Table 4-7). CaiA catalyzes the reduction of crotonobetainyl-CoA to γ-butyrobetainyl-CoA (Eichler et al., 1994), a reaction very similar to that catalyzed by YdiO in the engineered reversal of the β-oxidation cycle. Moreover, the operon *fixABCX* is required for the transfer of electrons to CaiA (Eichler et al., 1995; Walt and Kahn, 2002) and encodes flavoproteins and ferredoxin with high sequence similarity to YdiQRST (Table 4-7).

Table 4-8 Thermodynamic analysis of the engineered reversal of the β-oxidation cycle

Reaction number and enzyme name (*gene*)	Standard ΔG_r (Min ΔG_r, Max ΔG_r) [kcal/mol] [a]
① **Thiolase** (*yqeF*, *fadA*) 2 Acetyl-CoA → Acetoacetyl-CoA + CoA-SH	7.1 (-1.9, 16.1)
② **Hydroxyacyl-CoA dehydrogenase** (*fadB*) Acetoacetyl-CoA + NADH + H^+ → 3-hydroxybutyryl-CoA + NAD^+	-3.7 (-12.7, 5.3)
③ **Enoyl-CoA hydratase** (*fadB*) 3-hydroxybutyryl-CoA → Crotonyl-CoA + H_2O	2.1 (-2.4, 6.6)
④ **Acyl-CoA dehydrogenase (coupled to ubiquinone,** *fadE*)[b] Crotonyl-CoA + UQH_2 → Butyryl-CoA + UQ	5.7 (-3.3, 14.7)
④ **Enoyl-CoA reductase (coupled to ferredoxin,** *ydiO-ydiQRST*)[b] Crotonyl-CoA + Fd^{2-} → Butyryl-CoA + Fd	-16.5 (-25.5, -7.5)
Operation of β-oxidation reversal coupled to ubiquinone	11.2
Operation of β-oxidation reversal coupled to ferredoxin	-11.0

[a] Standard ΔG of formation values estimated using the group contribution method(Jankowski et al., 2008) and used to calculate the standard ΔG of reaction(Feist et al., 2007). Minimum and maximum ΔG_r values calculated assuming standard conditions (298.15 K, pH 7) with minimum and maximum metabolite concentrations set to 0.00001 M and 0.02 M, respectively(Feist et al., 2007). Listed ΔG_r values are in good agreement with experimentally measured/calculated ΔG_r values(Seubert and Podack, 1973; Thauer et al., 1977).

[b] Calculation of ΔG_r for enoyl-CoA reductase used standard reduction potentials from Thauer *et al*(Thauer et al., 1977).

This analysis suggests that ferredoxin and flavoproteins encoded by *ydiQRST* are involved in the transfer of electrons to YdiO during the reduction of enoyl-CoA to acyl-CoA. Indeed, deletion of the *ydiQRST* operon completely abolished n-butanol production (data not shown). Thermodynamic calculations of the engineered reversal of the β-oxidation cycle indicate that this pathway can function effectively in the biosynthetic direction as long as the reduction of enoyl-CoA to acyl-CoA is mediated by YdiO-YdiQRST (Table 4-8).

4.3.1. Establishing a platform for the synthesis of advanced fuels and chemicals

The engineered reversal of the β-oxidation cycle generates a diverse set of CoA thioester intermediates of different chain lengths, which in turn can be converted to the corresponding alcohols and carboxylic acids (Figure 4-1). The class of alcohol and carboxylic acid synthesized can be controlled by choosing a thiolase of appropriate chain-length specificity (reaction ①, Figure 4-1) and/or a termination enzyme with specificity towards the CoA-thioester intermediate of desired chain length and functionality (reactions ⑤ and ⑥, Figure 4-1). To illustrate product synthesis from intermediates other than acyl-CoA, we used thioesterase I (TesA) (Cho and Cronan, 1993) and thioesterase II (TesB) (Nie et al., 2008) as termination enzymes. As shown in Figure 4-3 and reported in Table 4-9, small amounts of β-hydroxybutyric, β-ketobutyric, and trans-2-butenoic acids were produced when these thioesterases were overexpressed in strain RB02.

Table 4-9 Cell growth, glucose utilization, product synthesis, and carbon recovery for derivatives of strain RB02 engineered for the synthesis of 4-C carboxylic acids

Strain[a]	Concentration[b] (g/L)					% C-recovery[d]
	Cells	Glucose utilized	β-keto-C4:0	β-hydroxy-C4:0	*trans*-2-butenoic acid	
RB02	0.29	1.76	ND[c]	ND	ND[c]	94.09
RB02 [*tesA*+]	0.62	4.25	ND[c]	0.032	0.010	95.30
RB02 [*tesB*+]	0.57	4.30	0.024	ND[c]	ND[c]	89.88
RB02 [*yqeF*+ *tesA*+]	0.56	4.95	ND[c]	0.045	0.012	91.08
RB02 Δ*ydiO* [*yqeF*+ *tesA*+]	0.42	3.34	ND[c]	0.140	0.171	89.23
RB03 [*yqeF*+ *tesB*+]	0.65	5.16	0.110	ND[c]	ND[c]	97.17
RB03 Δ*fadB* [*yqeF*+ *tesB*+]	0.62	5.03	0.450	ND[c]	ND[c]	95.53

[a] Data represent averages for three samples taken from shake flask cultures grown on 2 % (w/v) glucose minimal medium.

[b] All genotypes are shown in Table 4-1.

[c] ND, not detectable. Minimum detection levels are: β-keto-C4:0, 4.09 mg l^{-1}; β-hydroxy-C4:0, 3.03 mg l^{-1} *trans*-2-butenoic acid, 9.40 mg l^{-1}.

[d] Carbon recovery was calculated as the ratio of total moles of carbon in products per moles of carbon in total glucose consumed and expressed on percentage basis.

The level of these products was significantly increased by simultaneous overexpression of thioesterase and *yqeF*-encoded short-chain acyltransferase (Figure 4-3). Further increases in titer were realized upon deletion of *fadB* (~ 500 mg/L β-ketobutyric acid) and *ydiO* (~ 150 mg/L and 200 mg/L of β-hydroxybutyric and trans-2-butenoic acids, respectively) (Figure 4-3, Table 4-9).

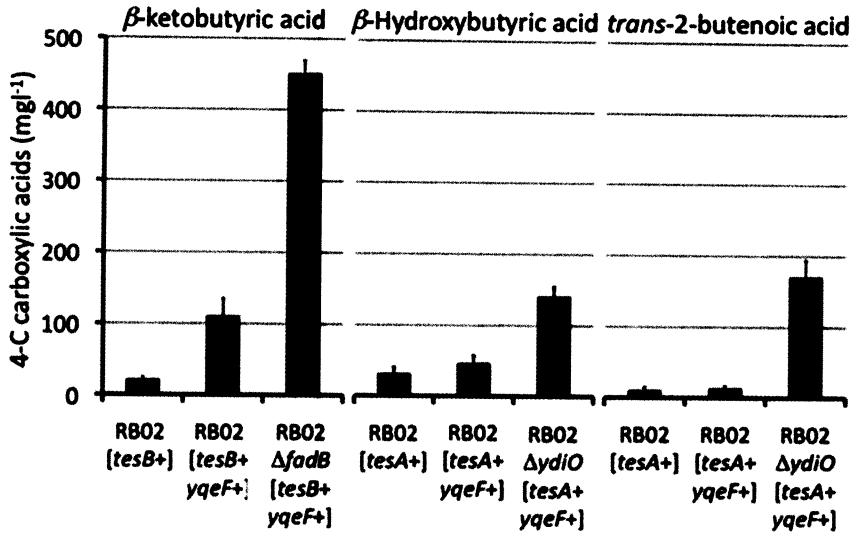

Figure 4-3 Engineered one-turn reversal of the β-oxidation cycle for the synthesis of short-chain carboxylic acids

Synthesis of β-ketobutyric (left panel), β-hydroxybutyric (center panel), and trans-2-butenoic (right panel) acids upon overexpression of thioesterases I (TesA) and II (TesB) in strains RB02, RB02 Δ*fadB*, and RB02 Δ*ydiO*. Experiments were run at 37 °C for 48 hours in shake flasks using glucose (1% w/v) minimal medium.

The operation of multiple cycles of the engineered reversal of the β-oxidation pathway, and hence the synthesis of CoA-thioester intermediates (and products) of longer chain length, can be facilitated by the overexpression of FadA, a 3-ketoacyl-CoA thiolase that is part of the β-oxidation complex (FadBA) and which possesses broad chain length specificity (Yang et al., 1990). This modification, in combination with the use of appropriate termination enzymes, should support the synthesis of longer chain (C > 4) carboxylic acids and alcohols. Indeed, overexpression of the β-oxidation complex (FadBA) in conjuction with thioesterases (TesA, TesB, FadM (Feng and Cronan, 2009) or YciA (Zhuang et al., 2008)) in strain RB03 (RB02 Δ*yqhD* Δ*fucO* Δ*fadD*) resulted in the accumulation of long-chain fatty acids in the extracellular medium (Figure 4-4). The

fadD knockout in strain RB03 blocks the catabolic operation of the β-oxidation cycle, thus preventing re-utilization of the synthesized fatty acids.

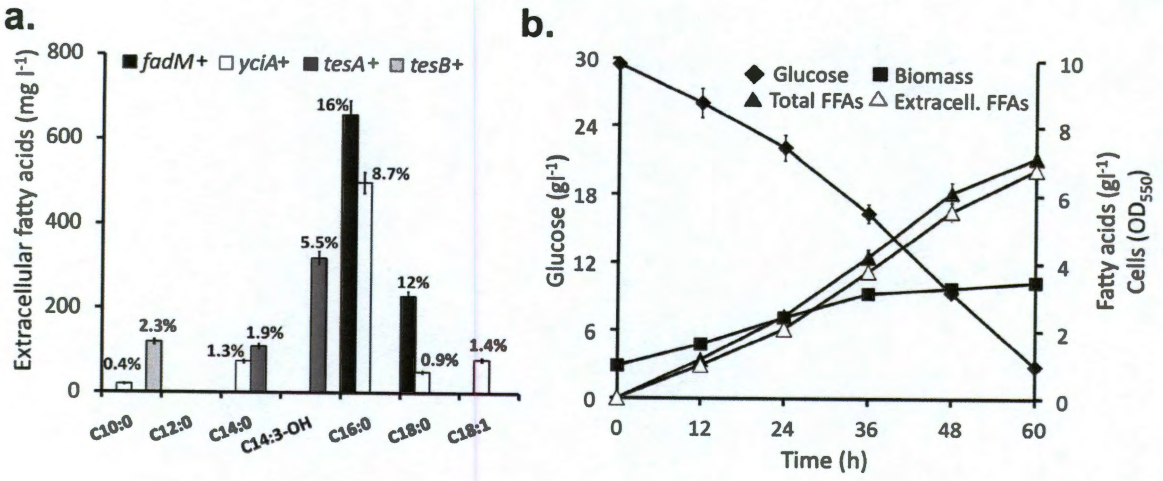

Figure 4-4 The engineered reversal of the β-oxidation cycle as a platform for the synthesis of higher-chain (C > 4) carboxylic acids

(a) Accumulation of long-chain (C > 10) free fatty acids in the extracellular medium of strain RB03 [*fadBA*+] upon overexpression of different thioesterases (FadM, YciA, TesA, TesB). The yield of each extracellular free fatty acid (g free fatty acid/g total glucose consumed × 100) is shown above the bars. Experiments were run at 37 °C for 48 hours in shake flasks using glucose (2% w/v) minimal medium.

(b) Kinetics of fatty acid synthesis by strain RB03 [*fadBA.fadM*+]. Cells were cultivated in fermentors containing minimal medium supplemented with 3 % (w/v) glucose. The dissolved oxygen was controlled at 2% of saturation, temperature at 37 °C, and pH at 7.

The choice of thioesterase allowed control over both length and functionality of the fatty acid side chain. For example, C16 and C18 saturated fatty acids were the only products when FadM was overexpressed while YciA and TesA overexpression supported the synthesis of β-hydroxy ($C_{14:3}OH$) and unsaturated ($C_{18:1}$) fatty acids, respectively (Figure 4-4a). When grown in a bioreactor using a higher initial concentration of glucose, strain RB03 [*fadBA.fadM*+] produced long-chain extracellular fatty acids at high titer (~ 7 g/L) and yield (0.28 g fatty acids/g total glucose consumed) using a mineral salts medium without rich nutrients (Figure 4-4b). These results are several folds better than

reported previously using an engineered fatty acid biosynthesis pathway (Steen et al., 2010) (Table 4-6). Unlike previous reports in which a large fraction of the produced fatty acids remained inside the cells (Handke et al., 2010; Lennen et al., 2010; Steen et al., 2010), more than 95% of the fatty acids synthesized by the engineered strains were secreted into the medium (Figure 4-4b).

Table 4-10 Synthesis of long-chain (C > 10) saturated fatty acids by RB03 (RB02 Δ*yqhD* Δ*fucO* Δ*fadD*) derivatives

Strain[a]	Concentration[b] (g/L)								
		Extracellular FFAs (Free Fatty Acids)							
	Glucose utilized	C10:0	C12:0	C14:0	C16:0	C18:0	Total FFAs[d]	Total FFAs/ CDW	% C-recovery[e]
MG1655	12.88	ND[c]	ND	ND	ND	ND	ND	/	93.05
RB03	4.30	ND	ND	ND	ND	ND	ND	/	90.64
RB03 [*fadBA*+]	2.24	ND	ND	ND	ND	ND	0.090	0.175	91.55
RB03 [*fadBA*+ *tesA*+]	5.82	ND	ND	0.110	ND	ND	-[f]	-	92.02
RB03 [*fadBA*+ *tesB*+]	5.16	0.120	ND	ND	ND	ND	-[f]	-	94.78
RB03 [*fadBA*+ *yciA*+]	2.87	0.050	ND	0.080	0.450	0.100	-[f]	-	91.94
RB03 [*fadBA*+ *fadM*+]	3.07	ND	ND	ND	0.700	0.180	-[f]	-	98.14
RB03 [*fadM*+][f]	3.30	ND	ND	0.020	0.150	0.080	0.435	0.790	92.86
RB03 [*fadBA.fadM*+]	5.91	ND	ND	ND	0.740	0.500	1.370	2.035	97.19
MG1655 Δ*adhE* Δ*pta* Δ*frdA* Δ*fadD*	5.04	ND	ND	ND	ND	ND	ND	/	86.44
MG1655 Δ*adhE* Δ*pta* Δ*frdA* Δ*fadD* [*fadM*+]	5.07	ND	ND	ND	ND	ND	0.070	0.092	94.15
MG1655 Δ*adhE* Δ*pta* Δ*frdA* Δ*fadD* [*fadM*+][f]	5.23	ND	ND	ND	ND	ND	0.261	0.279	96.17

[a] Data represent averages for three samples taken from shake flask cultures grown on 2 % (w/v) glucose minimal medium.

[b] All genotypes are shown in Supplementary Table 7.

[c] ND, not detectable. Minimum detection levels are: C10:0, 21.76 mg l^{-1}; C12:0, 20.45 mg l^{-1}; C14:0, 27.12 mg l^{-1}; C16:0, 20.17 mg l^{-1}; C18:0, 16.42 mg l^{-1}.

[d] FFAs, Free Fatty Acids

[e] Carbon recovery was calculated as the ratio of total moles of carbon in products per moles of carbon in total glucose consumed and expressed on percentage basis.

[f] *fadM* was overexpressed from medium-copy vector pTrcHis2A (Invitrogen, Carlsbad, CA).

No production of extracellular free fatty acids was observed upon overexpression of FadM in strain MG1655 Δ*adhE* Δ*pta* Δ*frd*A Δ*fadD* (Table 4-10), demonstrating the requirement of an active reversal of the β-oxidation cycle. Measurements of total free fatty acids in the whole broth (i.e. extracellular + intracellular) in strain RB03 [*fadBA.fadM*+] and the corresponding controls showed that the engineered reversal contributed to the synthesis of 90-95% of the total free fatty acids (Table 4-10).

Table 4-11 Synthesis of higher chain (C > 4) n-alcohols by derivatives of strain RB03 (RB02 Δ*yqhD* Δ*fucO* Δ*fadD*) derivatives

Strain[a]	Concentration[b] (g/L)					% C-recovery [d]
	Cells	Glucose utilized	n-C6-OH	n-C8-OH	n-C10-OH	
MG1655 Δ*adhE* Δ*pta* Δ*frdA* Δ*fadD* [*betA*+]	0.65	4.16	ND[c]	ND[c]	ND[c]	91.54
RB03 [*fadBA*+]	0.71	5.27	ND[c]	ND[c]	ND[c]	90.38
RB03 [*fadBA*+ *yiaY*+]	0.73	5.06	0.170	0.080	0.170	91.97
RB03 [*fadBA*+ *eutG*+]	0.52	4.00	0.170	0.070	0.010	90.36
RB03 [*fadBA*+ *betA*+]	0.72	3.98	0.210	0.100	0.020	91.16

[a] Data represent averages for three samples taken from shake flask cultures grown on 2 % (w/v) glucose minimal medium.
[b] All genotypes are shown in Supplementary Table 7.
[c] ND, not detectable. Minimum detection levels are: C10:0, 21.76 mg l^{-1}; C12:0, 20.45 mg l^{-1}; C14:0, 27.12 mg l^{-1}; C16:0, 20.17 mg l^{-1}; C18:0, 16.42 mg l^{-1}.
[d] FFAs, Free Fatty Acids
[e] Carbon recovery was calculated as the ratio of total moles of carbon in products per moles of carbon in total glucose consumed and expressed on percentage basis.
[f] *fadM* was overexpressed from medium-copy vector pTrcHis2A (Invitrogen, Carlsbad, CA).

The synthesis of longer-chain (C >4) n-alcohols was also demonstrated by overexpressing the appropriate termination enzymes. To this end, we identified native enzymes that could serve as potential surrogates for the aldehyde-forming acyl-CoA reductases and alcohol dehydrogenases present in organisms that synthesize higher-chain

linear n-alcohols (Table 4-5). The product titer (0.33 g/L) and yield (8.3 % w/w) achieved upon overexpression of YiaY (Figure 4-5) were several folds higher than previously reported (Steen et al., 2010; Zhang et al., 2008) (Table 4-11). Further improvements in the synthesis of these alcohols can be realized by expressing the heterologous aldehyde-forming acyl-CoA reductases and alcohol dehydrogenases listed in Table 4-5.

The ability of the engineered pathway to produce odd-chain alcohols was demonstrated by supplementing the medium with propionate as the precursor of propionyl-CoA (R = CH_3 in Figure 4-1).

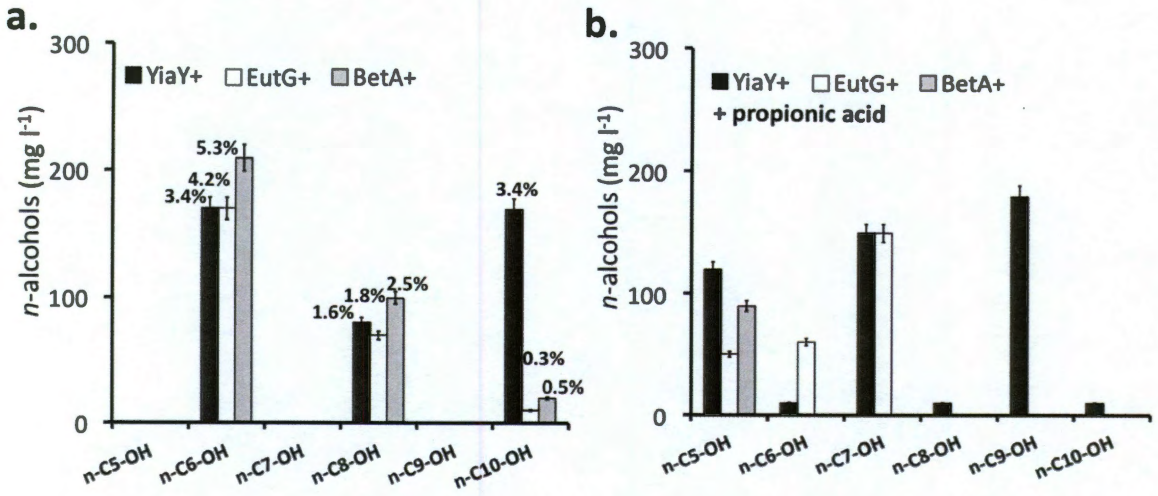

Figure 4-5 The engineered reversal of the β-oxidation cycle as a platform for the synthesis of higher-chain (C > 4) n-alcohols

(a) Synthesis of n-alcohols in strain RB03 [fadBA+] upon overexpression of alcohol dehydrogenases (YiaY, BetA, and EutG). The yield of each n-alcohol (g n-alcohol/g total glucose consumed × 100) is shown above the bars. Experiments were run at 37 °C for 48 hours in shake flasks using glucose (2% w/v) minimal medium.

(b) Effect of alcohol dehydrogenase overexpression (YiaY, BetA, and EutG) on the chain-length distribution of n-alcohols synthesized by strain RB03 [fadBA+] in the presence of 0.5 g/L propionate.

A clear shift in the distribution of n-alcohols was observed: odd-chain alcohols 1-pentanol, 1-heptanol, and 1-nonanol appeared as fermentation products and the synthesis of even-chain alcohols significantly decreased (Figure 4-5). BetA overexpression resulted in the synthesis of n-pentanol as the only higher-chain alcohol (Figure 4-5d), demonstrating the ability to control chain-length distribution.

4.4. Discussion

The functional reversal of the β-oxidation cycle engineered in this work represents a novel and efficient platform for the synthesis of advanced fuels and chemicals.

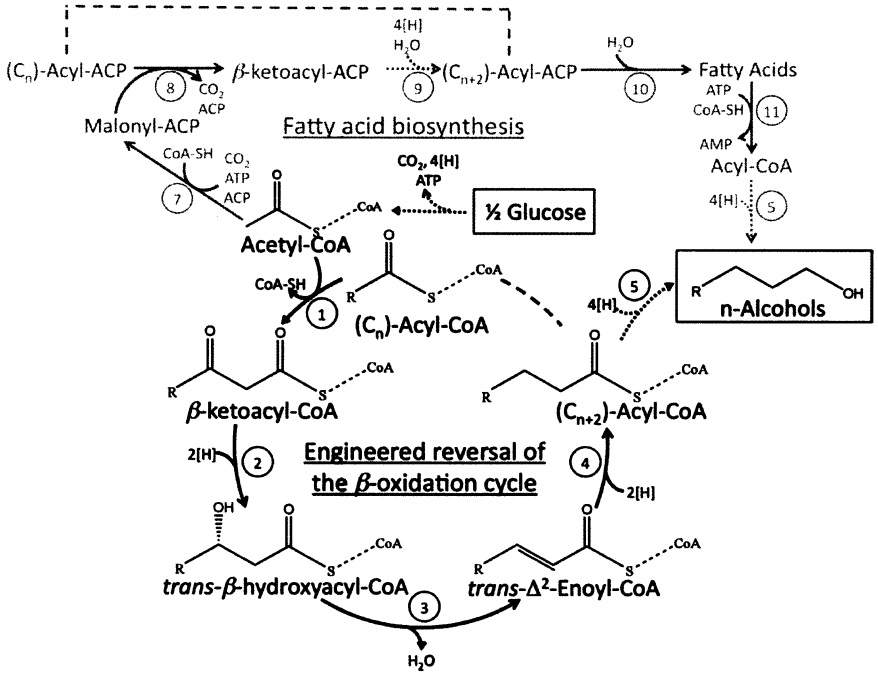

Figure 4-6 Comparison of n-alcohols synthesis via the fatty acid biosynthesis pathway and the engineered reversal of the β-oxidation cycle

Reactions ①-⑤ are as indicated in Figure 4-1. ⑦, acetyl-CoA carboxylase & malonyl-CoA ACP transacylase; ⑧, β-ketoacyl-acyl-carrier protein synthase; ⑨, 3-ketoacyl reductase, enoyl reductase, and 3-hydroxyacyl dehydratase; ⑩ & ⑪correspond to acyl-ACP conversion to free acids (⑩), and acylation (⑪).

Its superior nature is easily illustrated in the following balanced equation for the synthesis of n-alcohols from glucose: $n/4\ C_6H_{12}O_6 \rightarrow C_nH_{n+2}O + n/2\ CO_2 + (n/2 - 1)\ H_2O + n/2\ ATP$, with *n* being the chain length of the n-alcohol (Figure 4-1). As can be seen in this balance, the engineered pathway has the potential to achieve the maximum theoretical yield of n-alcohols on glucose (66.7 %, C-mole basis) and generates 1 ATP per each two-carbon unit incorporated into the n-alcohol molecule. This ATP yield is equivalent to that of efficient homo-fermentative pathways found in nature such as ethanol (Osman et al., 1987) and lactic acid (Papagianni et al., 2007) fermentations. The high carbon and energy efficiency of the engineered reversal of the β-oxidation cycle is possible because it directly uses acetyl-CoA as the donor of two-carbon units during chain elongation (as opposed to first requiring ATP-dependent activation to malonyl-CoA) and it functions with acyl-CoA intermediates, which are the precursors of n-alcohols and other important products (Figure 4-1).

The synthesis of n-alcohols through alternative metabolic routes, such as the fatty acid biosynthesis (Steen et al., 2010) and keto-acid (Zhang et al., 2008) pathways, is less efficient. For example, the use of the fatty acid biosynthesis pathway results in the net consumption of 1 ATP per molecule of n-alcohol synthesized (Figure 4-6).

The fatty acid biosynthesis pathway uses acyl-ACP intermediates and involves a β-ketoacyl-acyl-carrier protein synthase along with a 3-ketoacyl reductase, an enoyl reductase, and a 3-hydroxyacyl dehydratase (Figure 4-6)(Chan and Vogel, 2010; Cronan and Thomas, 2009). Production of n-alcohols from these acyl-ACP intermediates requires their conversion to free acids and acylation before their reduction to alcohols can be achieved (Steen et al., 2010). The use of acyl-ACP intermediates and malonyl-ACP as the

2-C donor during chain elongation in the fatty acid biosynthesis pathway limits therefore its ATP efficiency, making this an ATP-consuming pathway, as shown in the following balanced equation for n-alcohol synthesis from glucose: $n/4\ C_6H_{12}O_6 + ATP \rightarrow C_nH_{n+2}O + n/2\ CO_2 + (n/2 - 1)\ H_2O$, with n being the chain length of the n-alcohol.

The keto-acid pathway (Zhang et al., 2008) is also less efficient than the reversal of the β-oxidation cycle: e.g. the maximum theoretical yield of n-hexanol, the highest-chain linear n-alcohol reported with the keto-acid pathway, is only 50% C-mole (*2 Glucose* → *n-hexanol* + ATP + $2[H]$ + $6\ CO_2$).

The results presented here demonstrate that the engineered reversal of the β-oxidation cycle is a metabolic platform that can support the combinatorial synthesis of advanced fuels and chemicals at high carbon and energy efficiency. While we focused on the engineering of *E. coli*, the ubiquitous nature of β-oxidation, aldehyde/alcohol dehydrogenase, and thioesterase enzymes should enable the use of similar metabolic engineering strategies to achieve efficient synthesis of n-alcohols and carboxylic acids in other industrial organisms. Although our work (Dellomonaco, 2011b) only used native termination enzymes, further exploiting the potential of the engineered reversal of the β-oxidation cycle will likely require the use of heterologous thioesterases, CoA-thioester reductases and aldehyde/alcohol dehydrogenases.

Conclusions and future directions

This research study revolved around the development of novel approaches in the context of fuel and chemical production in *Escherichia coli*. Microbial bioconversions are a promising avenue for the development of viable processes for the generation of fuels and chemicals from sustainable resources. In order to become cost and energy effective, these processes must utilize organisms that can be optimized to efficiently produce candidate molecules from a variety of feedstocks. While foregoing metabolic engineering efforts had targeted the design of cell factories that could efficiently utilize carbohydrate feedstocks, this work has demonstrated a novel metabolic platform for the conversion of fatty-acid rich feedstocks, previously utilized only for chemical conversions, to fuels and chemicals.

In order to become a viable means for producing biofuels, microbial processes must utilize microorganisms that can be readily manipulated to maximize the production of candidate chemicals and fuels in large-scale industrial applications. In this regard, synthetic biology has provided an avenue to reconstruct *de novo* pathways in bacteria that are not known to possess the metabolic trait of interest by importing foreign pathways. The present study proposes a novel methodology in the panorama of the design of

microbial factories, alternative to current available platforms that entail either the utilization of native organisms or the heterologous expression in model organisms. Homologous Metabolic Engineering aims at exploiting innate metabolic potential of industrial organisms, and provides an opportunity to re-tune the functional differentiation of bacteria. Given the functional plasticity of bacteria, exploiting the native biosynthetic potential is the strategy of choice in this work, and is implemented to establish a novel platform that utilize a functional reversal of the β-oxidation cycle to produce a wide array of molecules with different chain lengths and functionalities.

Overall, the study hereby presented has aimed at *i.* exploring alternative renewable feedstocks and developing microbial platform to convert them to valuable products; *ii.* devising and implementing a new methodology to construct efficient metabolic platforms: *iii.* exploring the innate potential of *E. coli* to establish a platform enabling biosynthesis of a diverse array of products. The relevance of this work is therefore better realized when considered in the context of the bioindustry workflow (Figure 1-1) (Dellomonaco et al., 2010b).

In **Chapter 2** we explored the microbial utilization of fatty-acid rich feedstocks, exploited until recently mainly for chemical conversions. The use of FA-rich feedstocks for the production of biofuels and biochemicals was established as a promising alternative avenue to establish biorefineries. Since metabolism of FAs requires the presence of an external electron acceptor, that would in turn preclude the synthesis of fermentation products, we engineered a respiro-fermentative metabolic mode that could integrate the efficient production of the desired products in combination with adequate catabolism of FAs. The high metabolic efficiency of the catabolic pathway, along with

the high-reduced state of carbon in FAs, enabled the production of fuels and chemicals at yields superior than those obtained with the use of lignocellulosic sugars. For example, the yield of ethanol and acetate in the engineered strains surpassed the maximum theoretical values that can be achieved from lignocellulosic sugars (Dellomonaco et al., 2010b). Butanol, on the other hand, was produced at yields and titers between two- and three-fold higher than those reported for its production from lignocellulosic sugars in engineered *E. coli*, *P. putida*, *B. subtilis*, and *S. cerevisiae* strains (Atsumi et al., 2008a; Inui et al., 2008; Nielsen et al., 2009; Steen et al., 2008). Our work (Dellomonaco et al., 2010b) also represented the first report of propionate production in *E. coli*, a compound previously thought to be synthesized only by Propionibacteria.

The results obtained in this study have provided a novel avenue for a more efficient and productive use of fatty-acid feedstocks as a source for the generation of reduced chemicals and fuels. While *Escherichia coli* has been the model organism considered in this study, several other strains have been identified as possessing characteristics required for fatty acids catabolism. Yeasts, such as *Saccharomyces cerevisiae* and some oleaginous species, are able to grow on fatty acids as sole carbon and energy source (Trotter, 2001). A strategy similar to the one implemented in this study can therefore be extended to other model organisms, like for example *S. cerevisiae*. However, the design of the metabolic engineering strategy will have in this case to account for the complication posed by the compartmentalization of biochemical processes in yeasts.

Chapter 3 provides a novel methodology (Homologous Metabolic Engineering, HoME) (Dellomonaco, 2011a) to design optimal biocatalysts for the conversion of feedstock constituents to molecules of interest. This new approach builds upon the innate

plasticity of bacterial metabolism, and thus aims at awakening and engineering latent metabolic potential of functionally pluripotent bacteria to produce fuels and chemicals. The differentiation of *E. coli* function-ome has been engineered and re-tuned from a mixed-acid fermentative metabolism to a Clostridial acetone-butanol-ethanol metabolism. *E. coli* metabolism has also been engineered to dissimilate glycerol to 1,3-propanediol (Dellomonaco, 2011a), a metabolic capability present in the γ-Proteobacteria but never demonstrated to date in *E. coli* using native genes. While these results further demonstrate that phylogenic distance of species does not necessarily entail distinct biochemical capabilities, they also support our hypothesis of the existence of a functional pluripotent bacterial cell (Dellomonaco, 2011a). Furthermore, they provide a proof of concept demonstration of HoME as a viable alternative approach to Heterologous Metabolic Engineering strategies currently employed in biocatalyst development.

In **Chapter 4** a metabolic platform for the synthesis of a wide variety of molecules was demonstrated in *E. coli* by engineering a functional reversal of the fatty acids catabolic pathway (β-oxidation) (Dellomonaco, 2011b). The reversal of the β-oxidation cycle was engineered in *E. coli* and used in combination with endogenous dehydrogenases and thioesterases to synthesize n-alcohols, fatty acids and carboxylic acids. The superior nature of the engineered pathway was demonstrated by producing higher-chain linear n-alcohols ($C \geq 4$) and extracellular long-chain fatty acids ($C > 10$) at higher efficiency than previously reported.

The ubiquitous nature of β-oxidation, aldehyde/alcohol dehydrogenase, and thioesterase enzymes has the potential to enable the efficient synthesis of these products in other industrial organisms. Furthermore, the flexibility of the pathway allows its

implementation for the production of a wide array of products. Acyl-CoA intermediates generated by the engineered reversal of the β-oxidation can be converted to fuels and chemicals of different chain length by selecting and overexpressing termination enzymes (aldehyde/dehydrogenase, thioesterase, etc.) with appropriate specificity.

Since our results have demonstrated that a functional reversal of the β-oxidation cycle can be used as an efficient platform to produce acyl-CoA-derived products, synthesis of products derived from these intermediates can be envisioned. As an example, acetone production has already been demonstrated in Chapter 3 using native *E. coli* genes, and production of long chain fatty acids in Chapter 4 has proved the efficiency of this platform in generating longer chain Co-A derivatives. It is therefore possible, in principle, to produce methylketones with higher chain length by integrating both strategies.

Bibliography

Agreda, V. H., & Zoeller, J. R., 1993. Acetic acid and its derivatives. CRC Press, Boca Raton, FL.

Aguilar, D., Aviles, F.X., Querol, E., & Sternberg, M.J., 2004. Analysis of phenetic trees based on metabolic capabilites across the three domains of life. *Journal of Molecular Biology*. **340**, 491-512.

Ahmad, A.L., Ismail, S., & Bhatia, S., 2005. Membrane treatment for palm oil mill effluent: effect of transmembrane pressure and crossflow velocity. *Desalination*. **179**, 245-255.

Alper, H., & Stephanopoulos, G., 2009. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential. *Nature Reviews Microbiology*. **7**, 715-723.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., & Lipman D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*. **25**, 3389-3402.

Andersch, W., Bahl, W., & Gottschalk, G. 1983. Level of enzymes involved in acetate, butyrate, acetone and butanol formation by *Clostridium acetobutylicum*. *European Journal of Applied Microbiology and Biotechnology*. **18**, 327-332.

Antoine, E., Rolland, J.L., Raffin, J.P., & Dietrich, J. 1999. Cloning and over-expression in *Escherichia coli* of the gene encoding NADPH group III alcohol dehydrogenase from *Thermococcus hydrothermalis* - Characterization and comparison of the native and the recombinant enzymes. *European Journal of Biochemistry*. **264**, 880-889.

Aristidou, A., & Penttila, M., 2000. Metabolic engineering applications to renewable resource utilization. *Current Opinion in Biotechnology*. **11**, 187-198.

Atsumi, S., Cann, A.F., Connor, M.R., Shen, C.R., Smith, K.M., Brynildsen, M.P., Chou, K.J., Hanai, T., & Liao, J.C., 2008a. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metabolic Engineering*. **10**, 305-311.

Atsumi, S., Hanai, T., & Liao, J.C. 2008b. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature*. **451**, 86-89.

Atsumi, S., & Liao, J. C., 2008. Metabolic engineering for advanced biofuels production from *Escherichia coli*. *Current Opinion in Biotechnology*. **19**, 414-419.

Awang, G.M., Jones, G.A., & Ingledew, W.M., 1988. The acetone-butanol-ethanol fermentation. *Critical Reviews in Microbiology*. **15**, S33-S67.

Azócar, L., Ciudad, G., Heipieper, H.J., Navia, R., 2010. Biotechnological processes for biodiesel production using alternative oils. *Applied Microbiology and Biotechnology*. **88**, 621-636.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., & Mori, H., 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*. **2**, doi:10.1038/msb4100050.

Bailey, J. E., 1991. Toward a science of metabolic engineering. *Science*. **252**, 1668-1675.

Bairoch, A., Apweiler, R., Wu, C.H., Barker, W.C., Boeckmann, B., Ferro, S., Gasteiger, S., Huang, H., Lopez, R., Magrane, M., Martin, M.J., Natale, D.A., O'Donovan, C., Redaschi, N., & Yeh, L.S.L. 2005. The universal protein resource (UniProt). *Nucleic Acids Research*. **33**, D154-D159.

Barbirato, F., Grivet, J.P., Soucaille, P., & Bories, A. 1996. 3-hydroxypropionaldehyde, an inhibitory metabolite of glycerol fermentation to 1,3-propanediol by enterobacterial species. *Applied and Environmental Microbiology.* **62**, 1448-1451.

Bennett, G. N., & Rudolph, F. B., 1995. The central metabolic pathway from acetyl-CoA to butyryl-CoA in *Clostridium acetobutylicum. Fems Microbiology Reviews.* **17**, 241-249.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., & Sayers, E.W. 2011. GenBank. *Nucleic Acids Research.* **39**, D32-D37.

Berezina, O.V., Brandt, A., Yarotsky, S., Schwarz, W.H., & Zverlov, V.V., 2009. Isolation of a new butanol-producing *Clostridium* strain: High level of hemicellulosic activity and structure of solventogenesis genes of a new *Clostridium saccharobutylicum* isolate. *Systematic and Applied Microbiology.* **32**, 449-459.

Berezina, O.V., Zakharova, N.V., Brandt, A., Yarotsky, S.V., Schwarz, W.H., & Zverlov, V.V., 2010. Reconstructing the clostridial n-butanol metabolic pathway in *Lactobacillus brevis. Applied Microbiology and Biotechnology.* **87**, 635-646.

Bergey, D. H., & Holt, J.G., 1994. *Bergey's manual of determinative bacteriology*, Williams & Wilkins, Baltimore.

Berman, H. M., Westbrook, J., & Zardecki, C. 2003. The protein data bank. In. Chasman, D. (Ed.), *The protein data bank. Protein Structure: Determination, analysis, and applications for drug discovery*, CRC Press.

Bermejo, L.L., Welker, N.E., & Papoutsakis, E.T. 1998. Expression of *Clostridium acetobutylicum* ATCC 824 genes in *Escherichia coli* for acetone production and acetate detoxification. *Applied and Environmental Microbiology*. **64**, 1079-1085.

Bernhardsson, S., Gerlee, P. & Lizana, L. 2011. Structural correlations in bacterial metabolic networks. *Bmc Evolutionary Biology*. **11**, doi:10.1186/1471-2148-11-20.

Bevan, M. W., & Franssen, M. C. R., 2006. Investing in green and white biotech. *Nature Biotechnology*. **24**, 765-767.

Biebl, H., Menzel, K., Zeng, A.P., & Deckwer, W.D. 1999. Microbial production of 1,3-propanediol. *Applied Microbiology and Biotechnology*. **52**, 289-297.

Biebl, H., Zeng, A.P., Menzel, K., & Deckwer, W.D., 1998. Fermentation of glycerol to 1,3-propanediol and 2,3-butanediol by *Klebsiella pneumoniae*. *Applied Microbiology and Biotechnology*. **50**, 24-29.

Biermann, U., Bornscheuer, U., Meier, M.A., Metzger, J.O., & Schäfer, H.J., 2011. Oils and fats as renewable raw materials in Chemistry. *Angewandte Chemie-International Edition*. **50**, 3854-3871.

Black, P. N., & Dirusso, C. C., 1994. Molecular and biochemical analyses of fatty acid transport, metabolism, and gene regulation in *Escherichia coli*. *Biochimica Et Biophysica Acta-Lipids and Lipid Metabolism*. **1210**, 123-145.

Black, P.N., DiRusso, C.C., Metzger, A.K., Heimert, T.L., 1992. Cloning, sequencing, and expression of the *fadD* gene of *Escherichia coli* encoding acyl coenzyme A synthetase. *Journal of Biological Chemistry*. **267**, 25513-25520.

Blaschkowski, H.P., Neuer, G., Ludwig-Festl, M., Knappe, J., 1982. Routes of flavodoxin and ferredoxin reduction in *Escherichia coli*. CoA-acylating pyruvate: flavodoxin and NADPH: flavodoxin oxidoreductases participating in the activation of pyruvate formate-lyase. *European Journal of Biochemistry*. **123**, 563-569.

Bobik, T.A., Xu, Y., Jeter, R.M., Otto, K.E., Roth, J.R., 1997. Propanediol utilization genes (*pdu*) of *Salmonella typhimurium*: Three genes for the propanediol dehydratase. *Journal of Bacteriology*. **179**, 6633-6639.

Bologna, F.P., Campos-Bermudez, V.A., Saavedra, D.D., Andreo, C.S., Drincovich, M.F., 2010. Characterization of *Escherichia coli* EutD: A Phosphotransacetylase of the Ethanolamine Operon. *Journal of Microbiology*. **48**, 629-636.

Bond-Watts, B.B., Bellerose, R.J., & Chang, M.C.Y. 2011. Enzyme mechanism as a kinetic control element for designing synthetic biofuel pathways. *Nature Chemical Biology*. **7**, 222-227.

Booth, I. R., 2005. Glycerol and methylglyoxal metabolism. In. Niedhardt, F. C., et al., (Eds.) *Escherichia coli and Salmonella: cellular and molecular biology*. American Society for Microbiology, Washington, D.C.

Boussau, B., & Daubin, V., 2010. Genomes as documents of evolutionary history. *Trends in Ecology & Evolution*. **25**, 224-232.

Bouvet, O.M., Lenormand, P., Ageron, E., & Grimont, P.A., 1995. Taxonomic diversity of anaerobic glycerol dissimilation in the *Enterobacteriaceae*. *Research in Microbiology*. **146**, 279-290.

Brennan, L., & Owende, P., 2010. Biofuels from microalgae - A review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable & Sustainable Energy Reviews*. **14**, 557-577.

Broa, C., Regenberga, B., Försterb, J., & Nielsen, J. 2006. In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metabolic Engineering*. **8**, 102-111.

Buljan, M., & Bateman, A., 2009. The evolution of protein domain families. *Biochemical Society Transactions*. **37**, 751-755.

Böck, A., & Sawers, G., 1996. Fermentation. In. Niedhardt, F. C., et al., (Eds.), *Escherichia coli and Salmonella: Cellular and molecular biology*. American Society for Microbiology, Washington, D.C.

Caetano-Anollés, G., Yafremava, L.S., Gee, H., Caetano-Anollés, D., Kim, H.S., & Mittenthal, J.E., 2009. The origin and evolution of modern metabolism. *International Journal of Biochemistry & Cell Biology*. **41**, 285-297.

Cameron, D.C., Altaras, N.E., Hoffman, M.L., & Shaw, A.J., 1998. Metabolic engineering of propanediol pathways. *Biotechnology Progress*. **14**, 116-125.

Campbell, J. W., & Cronan, J. E., 2001. Escherichia coli FadR positively regulates transcription of the *fabB* fatty acid biosynthetic gene. *Journal of Bacteriology*. **183**, 5982-5990.

Campbell, J. W., & Cronan, J. E., 2002. The enigmatic *Escherichia coli fadE* gene is *yafH*. *Journal of Bacteriology*. **184**, 3759-3764.

Campbell, J.W., Morgan-Kiss, R.M., & Cronan, J.E. Jr., 2003. A new *Escherichia coli* metabolic competency: growth on fatty acids by a novel anaerobic beta-oxidation pathway. *Molecular Microbiology*. **47**, 793-805.

Campos-Bermudez, V.A., Bologna, F.P., Andreo, C.S., Drincovich, M.F., 2010. Functional dissection of *Escherichia coli* phosphotransacetylase structural domains and analysis of key compounds involved in activity regulation. *Febs Journal*. **277**, 1957-1966.

Carlsson, A. S., 2009. Plant oils as feedstock alternatives to petroleum - A short survey of potential oil crop platforms. *Biochimie*. **91**, 665-670.

Cary, J.W., Petersen, D.J., Papoutsakis, E.T., & Bennett, G.N. 1990. Cloning and expression of *Clostridium acetobutylicum* ATCC 824 acetoacetyl-coenzyme A:acetate/butyrate:coenzyme A-transferase in *Escherichia coli*. *Applied and Environmental Microbiology*. **56**, 1576-1583.

Casci, T., 2011. Phylogenomics: improving our family tree. *Nature Reviews Genetics*. 12, doi:10.1038/nrg2996.

Causey, T.B., Zhou, S., Shanmugam, K.T., & Ingram, L.O. 2003. Engineering the metabolism of *Escherichia coli* W3110 for the conversion of sugar to redox-neutral and oxidized products: Homoacetate production. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 825-832.

Celinska, E., 2010. Debottlenecking the 1,3-propanediol pathway by metabolic engineering. *Biotechnology Advances*. **28**, 519-530.

Chan, D. I., & Vogel, H. J., 2010. Current understanding of fatty acid biosynthesis and the acyl carrier protein. *Biochemical Journal*. **430**, 559-559.

Chen, J. S., & Hiu, S. F., 1986. Acetone-butanol-isopropanol production by *Clostridium beijerinckii* (synonym, *Clostridium butylicum*). *Biotechnology Letters*. **8**, 371-376.

Cheryana, M., Parekha, S., Shaha, M., & Witjitra, K. 1997. Production of acetic acid by *Clostridium thermoaceticum*. *Advances in Applied Microbiology*, **43**, 1-33.

Cho, B.K., Knight, E.M., & Palsson, B.Ø., 2006. Transcriptional regulation of the *fad* regulon genes of *Escherichia coli* by ArcA. *Microbiology-Sgm*. **152**, 2207-2219.

Cho, H. S., & Cronan, J. E., 1993. *Escherichia coli* thioesterase I, molecular cloning and sequencing of the structural gene and identification as a periplasmic enzyme. *Journal of Biological Chemistry*. **268**, 9238-9245.

Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, C., & Bork, P. 2006. Toward automatic reconstruction of a highly resolved tree of life. *Science*. **311**, 1283-1287.

Cirino, P.C., Chin, J.W., & Ingram, L.O., 2006. Engineering *Escherichia coli* for xylitol production from glucose-xylose mixtures. *Biotechnology and Bioengineering*. **95**, 1167-1176.

Clark, D. P., 1989. The fermentation pathways of *Escherichia coli*. *FEMS Microbiol Reviews*. **5**, 223-34.

Clark, D. P., Cronan, J. E., 2005. Two-carbon compounds and fatty acids as carbon sources. In. Niedhardt, F. C., et al., (Eds.), *Escherichia coli and Salmonella: Cellular and molecular biology*. American Society for Microbiology, Washington, D.C.

Clomburg, J. M., & Gonzalez, R., 2010. Biofuel production in *Escherichia coli*: The role of metabolic engineering and synthetic biology. *Applied Microbiology and Biotechnology*. **86**, 419-434.

Cronan, J. E., & Thomas, J., 2009. Bacterial fatty acid synthesis and its relationships with polyketide synthetic pathways. *Methods in Enzymology*. **459**, 395-433.

Dagan, T., et al., 2010. Genome Networks Root the Tree of Life between Prokaryotic Domains. Genome Biology and Evolution. 2, 379-392.

Dalmolin, R. J. S., Castro, M.A.A., Filho, J.L.R., Souza, L.H.T., de Almeida, R.M.C., & Moreira, J.F.C., 2011. Evolutionary plasticity determination by orthologous groups distribution. *Biology Direct*. **6**, doi:10.1186/1745-6150-6-22.

Daniel, R., Boenigk, R., & Gottschalk, G. 1995. Purification of 1,3-propanediol dehydrogenase from *Citrobacter freundii* and cloning, sequencing, and overexpression of the corresponding gene in *Escherichia coli*. *Journal of Bacteriology*. **177**, 2151-2156.

Dellomonaco, C., Blankschien, M., Clomburg, J.M., & Gonzalez, R., 2011a. Engineering functional differentiation of bacteria by harnessing the pluripotent capabilities of native hosts. *In preparation.*

Dellomonaco, C., Clomburg, J.M., Miller, E.N., & Gonzalez, R., 2011b. Engineered reversal of the β-oxidation cycle for the synthesis of fuels and chemicals. *Nature*, doi: 10.1038/nature10333.

Dellomonaco, C., Fava, F., & Gonzalez, R. 2010a. The path to next generation biofuels: successes and challenges in the era of synthetic biology. Microbial Cell Factories. **9**, doi:10.1186/1475-2859-9-3.

Dellomonaco, C., Rivera, C., Campbell, P., & Gonzalez, R. 2010b. Engineered respiro-fermentative metabolism for the production of biofuels and biochemicals from fatty acid-rich feedstocks. *Applied and Environmental Microbiology*. **76**, 5067-5078.

Delsuc, F., Brinkmann, H., & Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics*. **6**, 361-375.

den Dooren de Jong, L., *Bijdrage tot de kennis van het mineralsatieproces*. (Doctoral Dissertation) Rotterdam, The Netherlands, 1926.

Deutscher, J., 2008. The mechanisms of carbon catabolite repression in bacteria. *Current Opinion in Microbiology*. **11**, 87-93.

Dharmadi, Y., & Gonzalez, R., 2005. A better global resolution function and a novel iterative stochastic search method for optimization of high-performance liquid chromatographic separation. *Journal of Chromatography A*. **1070**, 89-101.

Dirusso, C. C., 1990. Primary sequence of the *Escherichia coli fadBA* operon, encoding the fatty acid-oxidizing multienzyme complex, indicates a high degree of homology to eucaryotic enzymes. *Journal of Bacteriology*. **172**, 6459-6468.

Dismukes, G.C., Carrieri, D., Bennette, N., Ananyev, G.M., & Posewitz, M.C., 2008. Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Current Opinion in Biotechnology*. **19**, 235-240.

Durnin, G., Clomburg, J., Yeates, Z., Alvarez, P.J., Zygourakis, K., Campbell, P., & Gonzalez, R., 2009. Understanding and harnessing the microaerobic metabolism of glycerol in *Escherichia coli*. *Biotechnology and Bioengineering*. **103**, 148-161.

Durre, P., 2005. *Handbook on Clostridia*. CRC Press, Boca Raton, FL, USA.

Dürre, P., Kuhn, A., Gottwald, M., & Gottschalk, G. 1987. Enzymatic investigations on butanol dehydrogenase and butyraldehyde dehydrogenase in extracts of *Clostridium acetobutylicum*. *Applied Microbiology and Biotechnology*. **26**, 268-272.

Durrett, T.P., Benning, C., & Ohlrogge, J., 2008. Plant triacylglycerols as feedstocks for the production of biofuels. *Plant Journal*. **54**, 593-607.

Eichler, K., Bourgis, F., Buchet, A., Kleber, H.P., & Mandrand-Berthelot, M.A. 1994. Molecular characterization of the cai operon necessary for carnitine metabolism in *Escherichia coli*. *Molecular Microbiology*. **13**, 775-786.

Eichler, K., Buchet, A., Bourgis, F., Kleber, H.P., & Mandrand-Berthelot, M.A., 1995. The *fix Escherichia coli* region contains four genes related to carnitine metabolism. *Journal of Basic Microbiology*. **35**, 217-227.

Eisen, J. A., & Fraser, C. M., 2003. Phylogenomics: Intersection of evolution and genomics. *Science*. **300**, 1706-1707.

Elssner, T., Engemann, C., Baumgart, K., & Kleber, H.P., 2001. Involvement of coenzyme A esters and two new enzymes, an enoyl-CoA hydratase and a CoA-transferase, in the hydration of crotonobetaine to L-carnitine by *Escherichia coli*. *Biochemistry*. **40**, 11140-11148.

Eppler, T., & Boos, W., 1999. Glycerol-3-phosphate-mediated repression of *malT* in *Escherichia coli* does not require metabolism, depends on enzyme IIA(Glc) and is mediated by cAMP levels. *Molecular Microbiology*. **33**, 1221-1231.

Fani, R., & Fondi, M., 2009. Origin and evolution of metabolic pathways. *Physics of Life Reviews*. **6**, 23-52.

Feigenbaum, J., & Schulz, H., 1975. Thiolases of *Escherichia coli*: purification and chain length specificities. *Journal of Bacteriology*. **122**, 407-411.

Feist, A.M., Henry, C.S., Reed, J.L., Krummenacker, M., Joyce, A.R., Karp, P.D., Broadbelt, L.J., Hatzimanikatis, V., Palsson, B.Ø., 2007. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology*. 3, doi:10.1038/msb4100155.

Feng, Y., & Cronan, J. E., 2009. A new member of the *Escherichia coli fad* regulon: Transcriptional regulation of *fadM* (*ybaW*). *Journal of Bacteriology*. **191**, 6320-6328.

Finn, R.D., Tate, J., Mistry, J., Coggill, P.C., Sammut, S.J., Hotz, H.R., Ceric, G., Forslund, K., Eddy, S.R., Sonnhammer, E.L.L., & Bateman, A. 2010. The Pfam protein families database. *Nucleic Acids Research*. **38**, D211-D222.

Fischer, C.R., Tseng, H.C., Tai, M., Prather, K.L., Stephanopoulos, G., 2010. Assessment of heterologous butyrate and butanol pathway activity by measurement of intracellular pathway intermediates in recombinant *Escherichia coli*. *Applied Microbiology and Biotechnology*. **88**, 265-275.

Forsberg, C. W., 1987. Production of 1,3-Propanediol from glycerol by *Clostridium acetobutylicum* and other Clostridium species. *Applied and Environmental Microbiology*. **53**, 639-643.

Forslund, K., Henricson, A., Hollich, V., & Sonnhammer, E.L.L. 2008. Domain tree-based analysis of protein architecture evolution. *Molecular Biology and Evolution*. **25**, 254-264.

Fraser, H.B., Hirsh, A.E., Steinmetz, L.M., Scharfe, C., & Feldman, M.W., 2002. Evolutionary rate in the protein interaction network. *Science*. **296**, 750-752.

Fukami-Kobayashi, K., Minezaki, Y., Tateno, Y., & Nishikawa, K. 2007. A tree of life based on protein domain organizations. *Molecular Biology and Evolution*. **24**, 1181-1189.

Fukuda, H., Kondo, A., & Noda, H., 2001. Biodiesel fuel production by transesterification of oils. *Journal of Bioscience and Bioengineering*. **92**, 405-416.

Furdui, C., & Ragsdale, S. W., 2000. The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood-Ljungdahl pathway. *Journal of Biological Chemistry*. **275**, 28494-28499.

George, H.A., Johnson, J.L., Moore, W.E., Holdeman, L.V., & Chen, J.S., 1983. Acetone, isopropanol, and butanol production by *Clostridium beijerinckii* (syn. *Clostridium butylicum*) and *Clostridium aurantibutyricum*. *Applied and Environmental Microbiology*. **45**, 1160-1163.

Gonzalez, R., Murarka, A., Dharmadi, Y., & Yazdani, S.S., 2008. A new model for the anaerobic fermentation of glycerol in enteric bacteria: Trunk and auxiliary pathways in *Escherichia coli*. *Metabolic Engineering*. **10**, 234-245.

Gonzalez, R., Tao, H., Shanmugam, K.T., York, S.W., & Ingram, L.O., 2002. Global gene expression differences associated with changes in glycolytic flux and growth rate in *Escherichia coli* during the fermentation of glucose and xylose. *Biotechnology Progress*. **18**, 6-20.

Gruez, A., Roig-Zamboni, V., Grisel, S., Salomoni, A., Valencia, C., Campanacci, V., Tegoni, M., & Cambillau, C., 2004. Crystal structure and kinetics identify

Escherichia coli YdcW gene product as a medium-chain aldehyde dehydrogenase. *Journal of Molecular Biology*. **343**, 29-41.

Gui, L., Sunnarborg, A., & LaPorte, D.C., 1996. Regulated expression of a repressor protein: FadR activates *iclR*. *Journal of Bacteriology*. **178**, 4704-4709.

Guia, M.M., Lee, K.T., & Bhatia, S. 2008. Feasibility of edible oil vs. non-edible oil vs. waste edible oil as biodiesel feedstock. *Energy*. **33**, 1646-1653.

Gupta, R. S., 1998. Protein phylogenies and signature sequences: A reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukaryotes. *Microbiology and Molecular Biology Reviews*. **62**, 1435-1491.

Haapalainen, A.M., Meriläinen, G., & Wierenga, R.K., 2006. The thiolase superfamily: condensing enzymes with diverse reaction specificities. *Trends in Biochemical Sciences*. **31**, 64-71.

Haller, T., Buckel, T., Rétey, J., & Gerlt, J.A., 2000. Discovering new enzymes and metabolic pathways: Conversion of succinate to propionate by *Escherichia coli*. *Biochemistry*. **39**, 4622-4629.

Han, B., Gopalan, V., & Ezeji, T.H., 2011. Acetone production in solventogenic *Clostridium* species: new insights from non-enzymatic decarboxylation of acetoacetate. *Applied Microbiology and Biotechnology*, doi: 10.1007/s00253-011-3276-5.

Hanai, T., Atsumi, S., & Liao, J.C. 2007. Engineered synthetic pathway for isopropanol production in *Escherichia coli*. *Applied and Environmental Microbiology*. **73**, 7814-7818.

Handke, P., Lynch, S.A., & Gill, R.T. 2010. Application and engineering of fatty acid biosynthesis in *Escherichia coli* for advanced fuels and chemicals. *Metabolic Engineering*. **13**, 28-37.

Hartmanis, M. G. N., & Gatenbeck, S., 1984. Intermediary metabolism in *Clostridium acetobutylicum* - Levels of enzymes involved in the formation of acetate and butyrate. *Applied and Environmental Microbiology*. **47**, 1277-1283.

Helms, D. R., Helms, C.W., Kosinski, R.J., & Cummings, J.C., 1997. *Energetics, Fermentation and Respiration*. Macmillan.

Henz, S.R., Huson, D.H., Auch, A.F., Nieselt-Struwe, K., & Schuster, S.C., 2005. Whole-genome prokaryotic phylogeny. *Bioinformatics*. **21**, 2329-2335.

Hernández-Montalvo, V., Valle, F., Bolivar, F., & Gosset, G., 2001. Characterization of sugar mixtures utilization by an *Escherichia coli* mutant devoid of the phosphotransferase system. *Applied Microbiology and Biotechnology*. **57**, 186-191.

Herrmann, G., Jayamani, E., Mai, G., & Buckel, W., 2008. Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. *Journal of Bacteriology*. **190**, 784-791.

Hertweck, C., 2009. The biosynthetic logic of polyketide diversity. *Angewandte Chemie-International Edition*. **48**, 4688-4716.

Hesslinger, C., Fairhurst, S.A., & Sawers, G., 1998. Novel keto acid formate-lyase and propionate kinase enzymes are components of an anaerobic pathway in *Escherichia coli* that degrades L-threonine to propionate. *Molecular Microbiology*. **27**, 477-492.

Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W., & Foust, T.D., 2007. Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science*. **315**, 804-807.

Himmi, E.H., Bories, A., Boussaid, A., & Hassani, L., 2000. Propionic acid fermentation of glycerol and glucose by *Propionibacterium acidipropionici* and *Propionibacterium freudenreichii* ssp *shermanii*. *Applied Microbiology and Biotechnology*. **53**, 435-440.

Hirsh, E., & Sharan, R., 2007. Identification of conserved protein complexes based on a model of protein network evolution. *Bioinformatics*. **23**, E170-E176.

Holland-Staley, C.A., Lee, K., Clark, D.P., & Cunningham, P.R., 2000. Aerobic activity of *Escherichia coli* alcohol dehydrogenase is determined by a single amino acid. *Journal of Bacteriology*. **182**, 6049-6054.

Homann, T., Tag, C., Biebl, H., Deckwer, W.D., & Schink, B., 1990. Fermentation of glycerol to 1,3-propanediol by *Klebsiella* and *Citrobacter* strain. *Applied Microbiology and Biotechnology*. **33**, 121-126.

Hong, S. H., & Lee, S. Y., 2002. Importance of redox balance on the production of succinic acid by metabolically engineered *Escherichia coli*. *Applied Microbiology and Biotechnology*. **58**, 286-290.

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. *Plant Journal*. **54**, 621-639.

Hwang, T.K., Ong, S.M., Seow, C.C., & Tan, H.K., Chemical composition of palm oil mill effluents. *Planter*. **54**, 749-756.

Ingram, L.O., Gomez, P.F., Lai, X., Moniruzzaman, M., Wood, B.E., Yomano, L.P., York, S.W., 1998. Metabolic engineering of bacteria for ethanol production. *Biotechnology and Bioengineering*. **58**, 204-214.

Inui, M., Suda, M., Kimura, S., Yasuda, K., Suzuki, H., Toda, H., Yamamoto, S., Okino, S., Suzuki, N., Yukawa, H., 2008. Expression of *Clostridium acetobutylicum* butanol synthetic genes in *Escherichia coli*. *Applied Microbiology and Biotechnology*. 77, 1305-1316.

Jankowski, M.D., Henry, C.S., Broadbelt, L.J., & Hatzimanikatis, V., 2008. Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophysical Journal*. **95**, 1487-1499.

Jarboe, L., 2010. YqhD: a broad-substrate range aldehyde reductase with various applications in production of biorenewable fuels and chemicals. *Applied Microbiology and Biotechnology*, **89**, 249-257.

Jarboe, L.R., Grabar, T.B., Yomano, L.P., Shanmugan, K.T., & Ingram, L.O., 2007. Development of ethanologenic bacteria. *Biofuels*. **108**, 237-261.

Jarboe, L.R., Zhang, X., Wang, X., Moore, J.C., Shanmugam, K.T., & Ingram, L.O., 2010. Metabolic engineering for production of biorenewable fuels and chemicals: Contributions of synthetic biology. *Journal of Biomedicine and Biotechnology*, doi: 10.1155/2010/76104.

Jenkins, L. S., & Nunn, W. D., 1987a. 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.

Jones, D. T., 2000. Protein structure prediction in the postgenomic era. *Current Opinion in Structural Biology*. **10**, 371-379.

Jones, D. T., & Woods, D. R., 1986. Acetone-butanol fermentation revisited. *Microbiological Reviews*. **50**, 484-524.

Kajiura, H., Mori, K., Tobimatsu, T., & Toraya, T., 2001. Characterization and mechanism of action of a reactivating factor for adenosylcobalamin-dependent glycerol dehydratase. *Journal of Biological Chemistry*. **276**, 36514-36519.

Kamm, B., Gruber, P.R., & Kamm, M., 2006. *Biorefineries – Industrial processes and products: Status quo and future directions*. Wiley-VCH Verlag GmbH, Weinheim.

Kamm, B., & Kamm, M., 2004. Principles of biorefineries. *Applied Microbiology and Biotechnology*. **64**, 137-145.

Kandler, O., 1983. Carbohydrate metabolism in lactic acid bacteria. *Antonie Van Leeuwenhoek Journal of Microbiology*. **49**, 209-224.

Kang, Y., Durfee, T., Glasner, J.D., Qiu, Y., Frisch, D., Winterberg, K.M., & Blattner, F.R., 2004a. Systematic mutagenesis of the *Escherichia coli* genome. *Journal of Bacteriology*. **186**, 4921-4930.

Keseler, I.M., Bonavides-Martínez, C., Collado-Vides, J., Gama-Castro, S., Gunsalus, R.P., Johnson, D.A., Krummenacker, M., Nolan, L.M., Paley, S., Paulsen, I.T., Peralta-Gil, M., Santos-Zavaleta, A., Shearer, A.G., Karp, P.D., 2009. EcoCyc: A comprehensive view of *Escherichia coli* biology. *Nucleic Acids Research*. **37**, D464-D470.

Kessler, D., & Knappe, J., 1996. Anaerobic dissimilation of pyruvate. In: Neidhardt, F. C., Curtiss III, R., Ingraham, J.L., Lin, E.C.C., Low, K.B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M., & Umbarger, H.E., (Eds.), *Escherichia coli and Salmonella: Cellular and Molecular Biology*. American Society for Microbiology Press, Washington, DC.

Kessler, D., Leibrecht, I., & Knappe, J., 1991. Pyruvate-formate-lyase-deactivase and acetyl-CoA reductase activities of *Escherichia coli* reside on a polymeric protein particle encoded by *adhE*. *Febs Letters*. **281**, 59-63.

Khankal, R., Chin, J.W., Ghosh, D., & Cirino, P.C., 2009. Transcriptional effects of CRP* expression in *Escherichia coli*. *Journal of Biological Engineering*. **3**, 13, doi:10.1186/1754-1611-3-13.

Kim, J., & Reed, J. L., 2010. OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Systems Biology*. **4**, 53, doi:10.1186/1752-0509-4-53.

Klein, K., Steinberg, R., Fiethen, B., & Overath, P., 1971. Fatty acid degradation in Escherichia coli. An inducible system for the uptake of fatty acids and further characterization of old mutants. *European Journal of Biochemistry*. **19**, 442-450.

Koonin, E. V., 2010. The two empires and three domains of life in the postgenomic age. *Nature Education*. 3, 27.

Koonin, E. V., & Wolf, Y. I., 2010. Constraints and plasticity in genome and molecular-phenome evolution. *Nature Reviews Genetics*. **11**, 487-498.

Kummerfeld, S. K., & Teichmann, S. A., 2009. Protein domain organisation: adding order. *BMC Bioinformatics*. **10**, doi:10.1186/1471-2105-10-39.

Laffend, L. A., Nagarajan, V., & Nakamura, C.E., 1997. Bioconversion of a fermentable carbon source to 1, 3-propanediol by a single microorganism. *United States Patent 7169588*, Washington, DC: U.S. Patent and Trademark Office.

Lalman, J. A., & Bagley, D. M., 2004. Extracting long-chain fatty acids from a fermentation medium. *Journal of the American Oil Chemists Society*. **81**, 105-110.

Langenbach, S., Rehm, B,H, & Steinbüchel, A., 1997. Functional expression of the PHA synthase gene PhaC1 from *Pseudomonas aeruginosa* in *Escherichia coli* results in poly(3-hydroxyalkanoate) synthesis. *Fems Microbiology Letters*. **150**, 303-309.

Lawther, R.P., Calhoun, D.H., Gray, J., Adams, C.W., Hauser, C.A., & Hatfield, G.W., 1982. DNA sequence fine-structure analysis of ilvG (IlvG+) mutations of *Escherichia coli* K-12. *Journal of Bacteriology*. **149**, 294-298.

Lee, S.K., Chou, H., Ham, T.S., Lee, T.S., & Keasling, J.D., 2008a. Metabolic engineering of microorganisms for biofuels production: from bugs to synthetic biology to fuels. *Current Opinion in Biotechnology*. **19**, 556-563.

Lee, S.Y., Park, J.H., Jang, S.H., Nielsen, L.K., Kim, J., & Jung, K.S., 2008b. Fermentative butanol production by Clostridia. *Biotechnology and Bioengineering*. **101**, 209-228.

Lehman, T.C., Hale, D.E., Bhala, A., & Thorpe, C., 1990. An acyl-Coenzyme A dehydrogenase assay utilizing ferricenium ion. *Analytical Biochemistry*. **186**, 280-284.

Lennen, R.M., Braden, D.J., West, R.A., Dumesic, J.A., & Pfleger, B.F., 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.

Lewis, T., Nichols, P.D., & McMeekin, T.A., 2000. Evaluation of extraction methods for recovery of fatty acids from lipid-producing microheterotrophs. *Journal of Microbiological Methods.* **43**, 107-116.

Li, M., Yao, S., & Shimizu, K. 2007a. Effect of *poxB* gene knockout on metabolism in *Escherichia coli* based on growth characteristics and enzyme activities. *World Journal of Microbiology & Biotechnology.* **23**, 573-580.

Li, R., Zhang, H., & Qi, Q. 2007b. The production of polyhydroxyalkanoates in recombinant *Escherichia coli. Bioresource Technology.* **98**, 2313-2320.

Li, Z.J., Shi, Z.Y., Jian, J., Guo, Y.Y., Wu, Q., & Chen, G.Q., 2010. Production of poly(3-hydroxybutyrate-co-4-hydroxybutyrate) from unrelated carbon sources by metabolically engineered *Escherichia coli. Metabolic Engineering.* **12**, 352-359.

Lin, H., Bennett, G.N., & San, K.Y., 2005. Genetic reconstruction of the aerobic central metabolism in *Escherichia coli* for the absolute aerobic production of succinate. *Biotechnology and Bioengineering.* **89**, 148-156.

Liu, Q., Luo, G., Zhou, X.R., & Chen, G.Q., 2011. Biosynthesis of poly(3-hydroxydecanoate) and 3-hydroxydodecanoate dominating polyhydroxyalkanoates by beta-oxidation pathway inhibited *Pseudomonas putida. Metabolic Engineering.* **13**, 11-17.

Liu, S., Bischoff, K.M., Qureshi, N., Hughes, S.R., & Rich, J.O., 2010. Functional expression of the thiolase gene *thl* from *Clostridium beijerinckii* P260 in *Lactococcus lactis* and *Lactobacillus buchneri. New Biotechnology.* **27**, 283-288.

Liu, X., Dong, Y., Zhang, J., Zhang, A., Wang, L., & Feng, L., 2009. Two novel metal-independent long-chain alkyl alcohol dehydrogenases from *Geobacillus thermodenitrificans* NG80-2. *Microbiology-Sgm.* **155**, 2078-2085.

Liu, X., Fan, K., & Wang, W., 2004. The number of protein folds and their distribution over families in nature. *Proteins-Structure Function and Genetics.* **54**, 491-499.

Ljungdahl, L. G., 2003. Biochemistry and physiology of *anaerobic bacteria.* Springer.

Loewenstein, Y., Raimondo, D., Redfern, O.C., Watson, J., Frishman, D., Linial, M., Orengo, C., Thornton, J., & Tramontano, A., 2009. Protein function annotation by homology-based inference. *Genome Biology.* **10**, doi:10.1186/gb-2009-10-2-207.

Logsdon, J. E., & Loke, R. A., 2000. Isopropyl Alcohol. *Kirk-Othmer Encyclopedia of Chemical Technology.* John Wiley & Sons, New York, NY.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., & Randall, R.J., 1951. Protein measurement with the Folin phenol reagent. *Journal of Biological Chemistry.* **193**, 265-275.

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

Maddox, I. S., 1989. The acetone-butanol-ethanol fermentation: recent progress in technology. *Biotechnology & Genetic Engineering Reviews.* **7**, 189-220.

Maier, U., Losen, M., & Büchs, J., 2004. Advances in understanding and modeling the gas-liquid mass transfer in shake flasks. *Biochemical Engineering Journal.* **17,** 155-167.

Menga, X., Yanga, J., Xua, X., Zhanga, L., Nieb, Q., & Xian, M. 2009. Biodiesel production from oleaginous microorganisms. *Renewable Energy.* **34**, 1-5.

Mermelstein, L.D., Welker, N.E., Petersen, D.J., Bennett, G.N., & Papoutsakis, E.T., 1994. Genetic and metabolic engineering of *Clostridium acetobutylicum* ATCC 824. *Recombinant DNA Technology II.* **721**, 54-68.

Miller, J. H., 1992. *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria.* Cold Spring Harbor Laboratory Press, Plainview, NY.

Moat, A. G., Foster, J.W., & Spector, M.P., 2002. Central pathways of carbohydrate metabolism. In. *Microbial Physiology.* Wiley-Liss, Inc.

Moya, A., Gil, R., Latorre, A., Peretó, J., Pilar Garcillán-Barcia, M., & de la Cruz, F., 2009. Toward minimal bacterial cells: evolution vs. design. *Fems Microbiology Reviews.* **33**, 225-235.

Mulder, N. J., 2010. Protein domain architectures. In. Carugo, O. & Eisenhaber, F. (Eds.) *Data mining techniques for the life sciences*, Springer-Verlag New York, LLC.

Murarka, A., Clomburg, J.M., Moran, S., Shanks, J.V., & Gonzalez, R. 2010. Metabolic analysis of wild-type *Escherichia coli* and a pyruvate dehydrogenase complex (PDHC)-deficient derivative reveals the role of PDHC in the fermentative metabolism of glucose. *Journal of Biological Chemistry.* **285**, 31548-31558.

Murarka, A., Dharmadi, Y., Yazdani, S.S., & Gonzalez, R., 2008. Fermentative utilization of glycerol by *Escherichia coli* and its implications for the production of fuels and chemicals. *Applied and Environmental Microbiology.* **74**, 1124-1135.

Murzin, A.G., Brenner, S.E., Hubbard, T., & Chothia, C., 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *Journal of Molecular Biology.* **247**, 537-540.

Nakamura, C. E., & Whited, G. M., 2003. Metabolic engineering for the microbial production of 1,3-propanediol. *Current Opinion in Biotechnology*. **14**, 454-459.

Neale, A.D., Scopes, R.K., Wettenhall, R.E., & Hoogenraad, N.J., 1987. Pyruvate decarboxylase of *Zymomonas mobilis*: isolation, properties, and genetic expression in *Escherichia coli*. *Journal of Bacteriology*. **169**, 1024-1028.

Neidhardt, F.C., Bloch, P.L., & Smith, D.F., 1974. Culture media for enterobacteria. *Journal of Bacteriology*. **119**, 736-747.

Nie, L., Ren, Y., Janakiraman, A., Smith, S., & Schulz, H., 2008. A novel paradigm of fatty acid beta-oxidation exemplified by the thioesterase-dependent partial degradation of conjugated linoleic acid that fully supports growth of *Escherichia coli*. *Biochemistry*. **47**, 9618-9626.

Nielsen, D.R., Leonard, E., Yoon, S.H., Tseng, H.C., Yuan, C., & Prather, K.L., 2009. Engineering alternative butanol production platforms in heterologous bacteria. *Metabolic Engineering*, **11**, 262-273.

Nielsen, J.O., Villadsen, J., & Lidén, G., 2003. *Bioreaction Engineering Principles*. Kluwer Academic/Plenum Publishers, New York, NY.

Nikolau, B.J., Perera, M.A., Brachova, L., & Shanks, B., 2008. Platform biochemicals for a biorenewable chemical industry. *Plant Journal*. **54**, 536-545.

Nölling, J., Breton, G., Omelchenko, M.V., Makarova, K.S., Zeng, Q., Gibson, R., Lee, H.M., Dubois, J., Qiu, D., Hitti, J., Wolf, Y.I., Tatusov, R.L., Sabathe, F., Doucette-Stamm, L., Soucaille, P., Daly, M.J., Bennett, G.N., Koonin, E.V., & Smith, D.R., 2001. Genome sequence and comparative analysis of the solvent-

producing bacterium *Clostridium acetobutylicum*. *Journal of Bacteriology*. **183**, 4823-4838.

Nunn, W. D., 1986a. A molecular view of fatty acid catabolism in *Escherichia coli*. *Microbiological Reviews*. **50**, 179-192.

O'Brien, J.R., Raynaud, C., Croux, C., Girbal, L., Soucaille, P., & Lanzilotta, W.N., 2004. Insight into the mechanism of the B-12-independent glycerol dehydratase from *Clostridium butyricum*: Preliminary biochemical and structural characterization. *Biochemistry*. **43**, 4635-4645.

Octave, S., & Thomas, D., 2009. Biorefinery: Toward an industrial metabolism. *Biochimie*. **91**, 659-664.

Ohta, K., Beall, D.S., Mejia, J.P., Shanmugam, K.T., & Ingram, L.O., 1991. Genetic improvement of *Escherichia coli* for ethanol production: chromosomal integration of *Zymomonas mobilis* genes encoding pyruvate decarboxylase and alcohol dehydrogenase II. *Applied and Environmental Microbiology*. **57**, 893-900.

Olsen, G. J., & Woese, C. R., 1993. Ribosomal RNA: a key to phylogeny. *Faseb Journal*. **7**, 113-123.

Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M,B., & Thornton, J.M., 1997. CATH - a hierarchic classification of protein domain structures. *Structure*. **5**, 1093-1108.

Osman, Y.A., Conway, T., Bonetti, S.J., & Ingram, L.O., 1987. Glycolytic flux in *Zymomonas mobilis*: enzyme and metabolite levels during batch fermentation. *Journal of Bacteriology*. **169**, 3726-3736.

Overath, P., Pauli, G., & Schairer, H.U., 1969. Fatty acid degradation in Escherichia coli. An inducible acyl-CoA synthetase, the mapping of old-mutations, and the isolation of regulatory mutants. *European Journal of Biochemistry*. **7**, 559-574.

Pace, N. R., 1997. A molecular view of microbial diversity and the biosphere. *Science*. **276**, 734-740.

Palleroni, N. J., 1983. The taxonomy of Bacteria. *Bioscience*. **33**, 370-377.

Papagianni, M., Avramidis, N., & Filiousis, G. 2007. Glycolysis and the regulation of glucose transport in *Lactococcus lactis* spp. *lactis* in batch and fed-batch culture. *Microbial Cell Factories*. **6**, doi: 10.1186/1475-2859-6-16.

Papp, B., Teusink, B., Notebaart, R.A., 2009. A critical view of metabolic network adaptations. *Hfsp Journal*. **3**, 24-35.

Park, S.J., Choi, J.I., & Lee, S.Y., 2005. Short-chain-length polyhydroxyalkanoates: Synthesis in metabolically engineered *Escherichia coli* and medical applications. *Journal of Microbiology and Biotechnology*. **15**, 206-215.

Park, S.J., Park, J.P., & Lee, S.Y. 2002. Metabolic engineering of *Escherichia coli* for the production of medium-chain-length polyhydroxyalkanoates rich in specific monomers. *FEMS Microbiology Letters*. **214**, 217-222.

Pauli, G., & Overath, P., 1972. *ato* operon: a highly inducible system for acetoacetate and butyrate degradation in *Escherichia coli*. *European Journal of Biochemistry*. **29**, 553-562.

Pereto', J., 2011. Origin and evolution of metabolisms. In: Gargaud, M., López-Garcìa, P., & Martin, H., (Eds.), *Origins and Evolution of Life: An Astrobiological Perspective*. Cambridge University Press.

Pérez, J.M., Arenas, F.A., Pradenas, G.A., Sandoval, J.M., & Vásquez, C.C., 2008. *Escherichia coli* YqhD exhibits aldehyde reductase activity and protects from the harmful effect of lipid peroxidation-derived aldehydes. *Journal of Biological Chemistry*. **283**, 7346-7353.

Petersen, D.J., Welch, R.W., Rudolph, F.B., & Bennett, G.N., 1991. Molecular cloning of an alcohol (butanol) dehydrogenase gene cluster from *Clostridium acetobutylicum* ATCC 824. *Journal of Bacteriology*. **173**, 1831-1834.

Pienkos, P. T., & Darzins, A., 2009. The promise and challenges of microalgal-derived biofuels. *Biofuels, Bioproducts and Biorefining*. **3**, 431-440.

Pleiss, J., 2006. The promise of synthetic biology. *Applied Microbiology and Biotechnology*. **73**, 735-739.

Poirier, Y., Antonenkov, V.D., Glumoff, T., & Hiltunen, J.K., 2006. Peroxisomal beta-oxidation - A metabolic pathway with multiple functions. *Biochimica Et Biophysica Acta-Molecular Cell Research*. **1763**, 1413-1426.

Pomposiello, P. J., & Demple, B., 2000. Identification of SoxS-regulated genes in *Salmonella enterica serovar Typhimurium*. *Journal of Bacteriology*. **182**, 23-29.

Ponting, C. P., & Russell, R. R., 2002. The natural history of protein domains. *Annual Review of Biophysics and Biomolecular Structure*. **31**, 45-71.

Powlowski, J., Sahlman, L., & Shingler, V., 1993. Purification and properties of the physically associated meta-cleavage pathway enzymes 4-hydroxy-2-ketovalerate aldolase and aldehyde dehydrogenase (acylating) from *Pseudomonas* sp. strain CF600. *Journal of Bacteriology*. **175**, 377-385.

Priest, F. G., & Austin, B., 1993. *Modern Bacterial Taxonomy*. Springer.

Pronk, J.T., Yde Steensma, H., & Van Dijken, J.P., 1996. Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast*. **12**, 1607-1633.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., & Lopez, R., 2005. InterProScan: protein domains identifier. *Nucleic Acids Research*. **33**, W116-W120.

Ragan, M. A., & Beiko, R. G., 2009. Lateral genetic transfer: open issues. *Philosophical Transactions of the Royal Society B-Biological Sciences*. **364**, 2241-2251.

Ragauskas, A.J., Williams, C.K., Davison, B.H., Britovsek, G., Cairney, J., Eckert, C.A., Frederick, W.J. Jr, Hallett, J.P., Leak, D.J., Liotta, C.L., Mielenz, J.R., Murphy, R., Templer, R., & Tschaplinski, T., 2006. The path forward for biofuels and biomaterials. *Science*. **311**, 484-489.

Ragsdale, S. W., & Pierce, E., 2008. Acetogenesis and the Wood-Ljungdahl pathway of CO2 fixation. *Biochimica Et Biophysica Acta-Proteins and Proteomics*. **1784**, 1873-1898.

Rangarajan, E.S., Li, Y., Ajamian, E., Iannuzzi, P., Kernaghan, S.D., Fraser, M.E., Cygler, M., & Matte, A., 2005. Crystallographic trapping of the glutamyl-CoA thioester intermediate of Family ICoA transferases. *Journal of Biological Chemistry*. **280,** 42919-42928.

Rao, Z., Ma, Z., Shen, W., Fang, H., Zhuge, J., & Wang, X., 2008. Engineered *Saccharomyces cerevisiae* that produces 1,3-propanediol from d-glucose. *Journal of Applied Microbiology*. **105**, 1768-1776.

Rasmussen, B., Fletcher, I.R., Brocks, J.J., & Kilburn, M.R., 2008. Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature*. **455,** 1101-1104.

Raynaud, C., Sarçabal, P., Meynial-Salles, I., Croux, C., & Soucaille, P., 2003. Molecular characterization of the 1,3-propanediol (1,3-PD) operon of *Clostridium butyricum*. *Proceedings of the National Academy of Sciences of the United States of America*. **100**, 5010-5015.

Reed, J.L., Vo, T.D., Schilling, C.H., & Palsson, B.O., 2003. An expanded genome-scale model of *Escherichia coli* K-12. *Genome Biology*. **4**, doi:10.1186/gb-2003-4-9-r54.

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.

Robertson, D. L., & Lovell, S. C., 2009. Evolution in protein interaction networks: co-evolution, rewiring and the role of duplication. *Biochemical Society Transactions*. **37**, 768-771.

Rodolfi, L., Chini Zittelli, G., Bassi, N., Padovani, G., Biondi, N., Bonini, G., Tredici, M.R., 2009. Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnology and Bioengineering*. **102**, 100-112.

Rogers, P. L., Lee, K.J., Skotnicki, M.L., & Tribe, D.E., 2007. Ethanol production by *Zymomonas mobilis*. *Advances in Biochemical Engineering/Biotechnology - Biofuels*. **108**, 263-288.

Rohankhedkar, M.S., Mulrooney, S.B., Wedemeyer, W.J., & Hausinger, R.P., 2006. The AidB component of the *Escherichia coli* adaptive response to alkylating agents is a flavin-containing, DNA-binding protein. *Journal of Bacteriology*. **188**, 223-230.

Rosenberg, J.N., Oyler, G.A., Wilkinson, L., & Betenbaugh, M.J., 2008. A green light for engineered algae: redirecting metabolism to fuel a biotechnology revolution. *Current Opinion in Biotechnology*. **19**, 430-436.

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

Rude, M. A., & Schirmer, A., 2009. New microbial fuels: a biotech perspective. *Current Opinion in Microbiology*. **12**, 274-281.

Sambrook, J., & Russell, D. W., 2001. Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press.

Sauer, U., & Eikmanns, B. J., 2005. The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. *Fems Microbiology Reviews*. **29**, 765-794.

Sawers, R. G., & Clark, D. P., 2004. Fermentative Pyruvate and Acetyl-Coenzyme A Metabolism. In. *Escherichia coli and Salmonella: cellular and molecular biology*. American Society for Microbiology, Washington, D.C.

Saxena, R.K., Anand, P., Saran, S., & Isar, J., 2009. Microbial production of 1,3-propanediol: Recent developments and emerging opportunities. *Biotechnology Advances*. **27**, 895-913.

Schenk, P.M., Thomas-Hall, S.R., Stephens, E., Marx, U.C., Mussgnug, J.H., Posten, C., Kruse, O., & Hankamer, B. 2008. Second Generation Biofuels: High-Efficiency Microalgae for Biodiesel Production. *BioEnergy Research.* **1**, 1939-1234.

Schultz, J., Milpetz, F., Bork, P., & Ponting, C.P., 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences of the United States of America.* **95**, 5857-5864.

Schutz, H., & Radler, F., 1984. Anaerobic reduction of glycerol to propanediol-1.3 by *Lactobacillus brevis* and *Lactobacillus buchneri. Systematic and Applied Microbiology.* **5**, 169-178.

Seedorf , H., Fricke, W.F., Veith, B., Brüggemann, H., Liesegang, H., Strittmatter, A., Miethke, M., Buckel, W., Hinderberger, J., Li, F., Hagemeier, H., Thauer, R.K., & Gottschalk, G. 2008. The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. *Proceedings of the National Academy of Sciences of the United States of America.* **105**, 2128-2133.

Seifert, C., Bowien, S., Gottschalk, G., & Daniel, R., 2001. Identification and expression of the genes and purification and characterization of the gene products involved in reactivation of coenzyme B-12-dependent glycerol dehydratase of *Citrobacter freundii. European Journal of Biochemistry.* **268**, 2369-2378.

Service, R. F., 2009. BIOFUELS ExxonMobil fuels Venter's efforts to run vehicles on algae-based oil. *Science.* **325**, 379-379.

Seubert, W., & Podack, E. R., 1973. Mechanisms and physiological roles of fatty acid chain elongation in microsomes and mitochondria. *Molecular and Cellular Biochemistry.* **1**, 29-40.

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**, 312-320.

Smolke, C. D., *The metabolic pathway engineering handbook.* CRC Press, 2009.

Snell, K.D., Feng, F., Zhong. L., Martin, D., & Madison, L.L., 2002. YfcX enables medium-chain-length poly(3-hydroxyalkanoate) formation from fatty acids in recombinant *Escherichia coli fadB* strains. *Journal of Bacteriology.* **184**, 5696-5705.

Song, H., & Lee, S. Y., 2006. Production of succinic acid by bacterial fermentation. *Enzyme and Microbial Technology.* **39**, 352-361.

Sophos, N. A., & Vasiliou, V., 2003. Aldehyde dehydrogenase gene superfamily: the 2002 update. *Chemico-Biological Interactions.* **143**, 5-22.

Spratt, S.K., Ginsburgh, C.L., & Nunn, W.D., 1981a. Isolation and genetic characterization of *Escherichia coli* mutants defective in propionate metabolism. *Journal of Bacteriology.* **146,** 1166-1169.

Spratt, S.K., Ginsburgh, C.L., & Nunn, W.D., 1981b. Isolation and genetic characterization of *Escherichia coli* mutants defective in propionate metabolism. *Journal of Bacteriology.* **146**, 1166-1169.

Steen, E.J., Chan, R., Prasad, N., Myers, S., Petzold, C.J., Redding, A., Ouellet, M., & Keasling, J.D., 2008. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microbial Cell Factories.* 7, doi:10.1186/1475-2859-7-36.

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.

Stephanopoulos, G., 2002. Metabolic engineering: Perspective of a chemical engineer. *Aiche Journal*. **48**, 920-926.

Stephanopoulos, G., 2007. Challenges in engineering microbes for biofuels production. *Science*. **315**, 801-804.

Stim-Herndon KP, Petersen DJ, Bennett GN, 1995. Characterization of an acetyl-CoA C-acetyltransferase (thiolase) gene from *Clostridium acetobutylicum* ATCC 824. *Gene*. **154**, 81-85.

Sumathia, S., Chaia, S.P., & Mohamed, A.R., 2008. Utilization of oil palm as a source of renewable energy in Malaysia. *Renewable & Sustainable Energy Reviews*. **12**, 2404-2421.

Teufel, R., Mascaraque, V., Ismail, W., Voss, M., Perera, J., Eisenreich, W., Haehnel, W., & Fuchs, G. 2010. Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *Proceedings of the National Academy of Sciences of the United States of America*. **107**, 14390-14395.

Thauer, R.K., Jungermann, K., & Decker, K., 1977. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriological Reviews*. **41**, 100-180.

Thompson, J. C., & He, B. B., 2006. Characterization of crude glycerol from biodiesel production from multiple feedstocks. *Applied Engineering in Agriculture*. **22**, 261-265.

Tielens, A. G. M., & Van Hellemond, J. J., 1998. The electron transport chain in anaerobically functioning eukaryotes. *Biochimica Et Biophysica Acta-Bioenergetics*. **1365**, 71-78.

Tong, I.Y., Liao, H.H., & Cameron, D.C., 1991. 1,3-Propanediol production by *Escherichia coli* expressing genes from the *Klebsiella pneumoniae dha* regulon. *Applied and Environmental Microbiology*. **57**, 3541-3546.

Travisano, M., *Microbial Evolution*. eLS, 2005.

Tripathi, C.K., Tripathi, D., Praveen, V., & Bihari, V., 2007. Microbial diversity - Biotechnological and industrial perspectives. *Indian Journal of Experimental Biology*. **45**, 326-332.

Trotter, P. J., 2001. The genetics of fatty acid metabolism in *Saccharomyces cerevisiae*. *Annual Review of Nutrition*. **21**, 97-119.

Van Suijdam, J.C., Kossen, N.W.F., & Joha, A.C., 1978. Model for oxygen transfer in a shake flask. *Biotechnology and Bioengineering*. **20**, 1695-1710.

van Beilen, J.B., Eggink, G., Enequist, H., Bos, R., & Witholt, B., 1992. DNA sequence determination and functional characterization of the OCT-plasmid-encoded *alkJKL* genes of *Pseudomonas oleovorans*. *Molecular Microbiology*. **6**, 3121-3136.

Vandijken, J. P., & Scheffers, W. A., 1986. Redox balances in the metabolism of sugars by yeasts. *Fems Microbiology Reviews*. **32**, 199-224.

Vo, M.T., Lee, K.W., Jung, Y.M., & Lee, Y.H., 2008. Comparative effect of overexpressed *phaJ* and *fabG* genes supplementing (R)-3-hydroxyalkanoate

monomer units on biosynthesis of mcl-polyhydroxyalkanoate in *Pseudomonas putida* KCTC1639. *Journal of Bioscience and Bioengineering*. **106**, 95-98.

Vo, M.T., Lee, K.W., Kim, T.K., & Lee, Y.H., 2007. Utilization of *fadA* knockout mutant *Pseudomonas putida* for overproduction of medium chain-length-polyhydroxyalkanoate. *Biotechnology Letters*. **29**, 1915-1920.

Vyazmensky, M., Zherdev, Y., Slutzker, A., Belenky, I., Kryukov, O., Barak, Z., & Chipman, D.M., 2009. Interactions between large and small subunits of different acetohydroxyacid synthase isozymes of *Escherichia coli*. *Biochemistry*. **48,** 8731-8737.

Wahl, R. C., & Ormejohnson, W. H., 1987. Clostridial pyruvate oxidoreductase and the pyruvate-oxidizing enzyme specific to nitrogen fixation in *Klebsiella pneumoniae* are similar enzymes. *Journal of Biological Chemistry*. **262**, 10489-10496.

Walt, A., & Kahn, M. L., 2002. The *fixA* and *fixB* genes are necessary for anaerobic carnitine reduction in *Escherichia coli*. *Journal of Bacteriology*. **184**, 4044-4047.

Walter, K.A., Bennett, G.N., & Papoutsakis, E.T., 1992. Molecular characterization of two Clostridium acetobutylicum ATCC 824 butanol dehydrogenase isozyme genes. *Journal of Bacteriology*. **174**, 7149-7158.

Welch, R.W., Rudolph, F.B., & Papoutsakis, E.T., 1989. Purification and characterization of the NADH-dependent butanol dehydrogenase from *Clostridium acetobutylicum* (ATCC 824). *Archives of Biochemistry and Biophysics*. **273**, 309-318.

Werpy, T., & Petersen, G., Volume I: *Results of Screening for Potential Candidates from Sugars and Synthesis Gas. Top Value Added Chemicals from Biomass*. U.S.

Department of Energy, Office of Energy Efficiency and Renewable Energy, Washington, D.C., 2004.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., Dicuccio, M., Edgar, R., Federhen, S., Feolom M., Geerm L.Y., Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D.J., Madden, T.L., Maglott, D.R., Miller, V., Ostell, J., Pruitt, K.D., Schuler, G.D., Shumway, M., Sequeira, E., Sherry, S.T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R.L., Tatusova, T.A., Wagner, L., & Yaschenko, E., 2008. Database resources of the national center for biotechnology information. *Nucleic Acids Research*. **36**, D13-D21.

Wiesenborn, D.P., Rudolph, F.B., & Papoutsakis, E.T., 1988. Thiolase from *Clostridium acetobutylicum* ATCC824 and its role in the synthesis of acids and solvents. *Applied and Environmental Microbiology*. **54**, 2717-2722.

Woese, C.R., Kandler, O., & Wheelis, M.L., 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America*. **87**, 4576-4579.

Wu, T.Y., Mohammad, A.W., Jahim, J.M., & Anuar, N., 2009. A holistic approach to managing palm oil mill effluent (POME): Biotechnological advances in the sustainable reuse of POME. *Biotechnology Advances*. **27**, 40-52.

Yanai, H., Doi, K., & Ohshima, T., 2009. *Sulfolobus tokodaii* ST0053 produces a novel thermostable, NAD-dependent medium-chain alcohol dehydrogenase. *Applied and Environmental Microbiology*. **75**, 1758-1763.

Yang, S., & Bourne, P. E., 2009. The evolutionary history of protein domains viewed by species phylogeny. *Plos One.* **4**, doi:10.1371/journal.pone.0008378.

Yang, S.Y., Li, J.M., He, X.Y., & Cosloy, S.D., & Schulz, H., 1988. Evidence that the *fadB* gene of the *fadAB* operon of *Escherichia coli* encodes 3-hydroxyacyl-coenzyme A (CoA) epimerase, delta 3-cis-delta 2-trans-enoyl-CoA isomerase, and enoyl-CoA hydratase in addition to 3-hydroxyacyl-CoA dehydrogenase. *Journal of Bacteriology.* **170**, 2543-2548.

Yang, S.Y., He Yang, X.Y., Healy-Louie, G., Schulz, H., & Elzinga, M., 1990. Nucleotide sequence of the *fadA* gene. Primary structure of 3-ketoacyl-coenzyme A thiolase from *Escherichia coli* and the structural organization of the *fadAB* operon. *Journal of Biological Chemistry.* **265**, 10424-10429.

Yazdani, S. S., & Gonzalez, R., 2007. Anaerobic fermentation of glycerol: a path to economic viability for the biofuels industry. *Current Opinion in Biotechnology.* **18**, 213-219.

Yazdani, S. S., & Gonzalez, R., 2008. Engineering *Escherichia coli* for the efficient conversion of glycerol to ethanol and co-products. *Metabolic Engineering.* **10**, 340-351.

Ying, X., Grunden, A.M., Nie, L., Adams, M.W., & Ma, K., 2009. Molecular characterization of the recombinant iron-containing alcohol dehydrogenase from the hyperthermophilic Archaeon, *Thermococcus* strain ES1. *Extremophiles.* **13**, 299-311.

Yosef, N., Kupiec, M., Ruppin, E., & Sharan, R., 2009. A complex-centric view of protein network evolution. *Nucleic Acids Research.* **37**, doi: 10.1093/nar/gkp414.

Youngleson, J.S., Santangelo, J.D., Jones, D.T., & Woods, D.R. 1988. Cloning and expression of a *Clostridium acetobutylicum* alcohol dehydrogenase gene in *Escherichia coli*. *Applied and Environmental Microbiology*. **54**, 676-682.

Zhang, K., Sawaya, M.R., Eisenberg, D.S., & Liao, J.C., 2008. Expanding metabolism for biosynthesis of nonnatural alcohols. *Proceedings of the National Academy of Sciences of the United States of America*. **105**, 20653-20658.

Zhuang, Z., Song, F., Zhao, H., Li, L., Cao, J., Eisenstein, E., Herzberg, O., & Dunaway-Mariano, D., 2008. Divergence of function in the hot dog fold enzyme superfamily: The bacterial thioesterase YciA. *Biochemistry*. **47**, 2789-2796.

Zhuge, B., Zhang, C., Fang, H., Zhuge, J., & Permaul, K., 2010. Expression of 1,3-propanediol oxidoreductase and its isoenzyme in *Klebsiella pneumoniae* for bioconversion of glycerol into 1,3-propanediol. *Applied Microbiology and Biotechnology*. **87**, 2177-2184.