METABOLIC ENGINEERING OF THE HYPERTHERMOPHILIC ARCHAEON *PYROCOCCUS FURIOSUS* FOR THE RENEWABLE PRODUCTION OF BIOFUELS AND COMMODITY CHEMICALS

by

MATTHEW WILLIAM KELLER

(Under the Direction of MICHAEL W. W. ADAMS)

ABSTRACT

The human race depends primarily on fossil fuels for the production of carbon based commodity chemicals and transportation fuels. Plant biomass is the leading feedstock in efforts to renewably produce liquid transportation fuels, but its use at large scales is inefficient and results in similar carbon emissions as traditional gasoline on an energy basis. Microorganisms, with their extremely diverse metabolic abilities, offer a wide range of alternative strategies for producing renewable fuels. One such strategy is to use a metabolically-engineered microbe to direct carbon dioxide into a carbon fixation cycle and directly into a fuel synthesis pathway. Our strategy is to engineer the hyperthermophile *Pyrococcus furiosus* (T_{opt} 100°C) to express the 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) cycle for carbon dioxide fixation, drive that pathway with molecular hydrogen via the native soluble hydrogenase, and direct the resulting acetyl-CoA to a bacterial butanol or ethanol fermentation pathway. The work here focuses on three parallel goals of the larger objective. The first is to demonstrate the first sub-pathway of the carbon fixation cycle, the second is to assemble a hybrid pathway for butanol production

from acetyl-CoA, and the third is to explore multiple gene donors for ethanol production from acetyl-CoA. The first sub-pathway of the carbon fixation cycle generates the key intermediate 3-HP, which is a valuable plastics precursor that is currently produced from petroleum. The insertion of the five genes encoding this three enzyme pathway into *P. furiosus* resulted in the accumulation of 50 mg/L 3-HP in the medium at 75°C. Under the controlled conditions of an *in vitro* assay, 3-HP production was demonstrated to be dependent upon H_2 and CO_2. Since no single butanol pathway sufficiently thermophilic for expression in *P. furiosus* was known, an artificial pathway was assembled from multiple gene donors. Concentrated cell suspensions were incubated at 60°C and butanol (70 mg/L) was produced from maltose. Unlike 3-HP and butanol, a small amount of ethanol (40 mg/L) is natively produced by *P. furiosus* at low growth temperatures. The bifunctional AdhE is capable of the sequential reduction of acetyl-CoA to ethanol. Since acetyl-CoA is the intended link between carbon fixation and fuel synthesis, AdhE was inserted into *P. furiosus* to demonstrate ethanol production from acetyl-CoA. Eight thermophilic bacteria were used as gene donors and a maximum of 200 mg/L ethanol was produced by recombinant *P. furiosus*. This demonstrates a new functional route of ethanol production from acetyl-CoA that is directly compatible with the overall strategy.

INDEX WORDS: *Pyrococcus furiosus*, electrofuels, biofuels, *Metallosphaera sedula*, carbon fixation, anaerobe, archaea, biotechnology, metabolic engineering, thermophile, *Thermoanaerobacter*

METABOLIC ENGINEERING OF THE HYPERTHERMOPHILIC ARCHAEON *PYROCOCCUS FURIOSUS* FOR THE RENEWABLE PRODUCTION OF BIOFUELS AND COMMODITY CHEMICALS

by

MATTHEW WILLIAM KELLER

B.S. College of Charleston, 2010

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2016

METABOLIC ENGINEERING OF THE HYPERTHERMOPHILIC ARCHAEON *PYROCOCCUS FURIOSUS* FOR THE RENEWABLE PRODUCTION OF BIOFUELS AND COMMODITY CHEMICALS

by

MATTHEW WILLIAM KELLER

Major Professor:	Michael W. W. Adams
Committee:	Claiborne V. Glover III Joy Doran-Peterson William N. Lanzilotta

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2016

DEDICATION

I dedicate this to my wife Jessica, and I thank her for her wonderful loving support.

ACKNOWLEDGMENTS

First I would like to thank my advisor Dr. Michael Adams for his support and guidance. I would also like to thank the Adams lab members and the ARPA-E team, especially Dr. Gerrit Schut and Dr. Gina Lipscomb, for helping to foster a fun, positive, and productive research experience.

This work was supported by the US Department of Energy as part of the Electrofuels Project of ARPA-E (grant number DE-AR0000081) and the National Science Foundation (grant number 1021RR166824). We thank Brian Vaccaro, Dennis Phillips and Zhirui Wang of the University of Georgia for running the NMR, ESI-MS and GC-MS analyses of 3-hydroxypropionate, respectively. The GC-MS facility is supported in part by the Center for Plant and Microbial Complex Carbohydrates funded by the US Department of Energy (DE-FG02-93ER-20097).

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

GOAL OF THIS WORK

The primary goal of this work is to establish a proof of concept for the renewable production of commodity chemicals and fuel targets within *P. furious*. This is accomplished by searching through the diverse metabolisms of thermophilic microorganisms in order to generate a biotechnologically useful pathway, and subsequently demonstrating that pathway in *P. furiosus*. This project, which began shortly after the development of a genetic system in *P. furiosus*, aims to engineer a carbon fixation cycle and a fuel synthesis pathway into *P. furiosus*. The ideal outcome would be autonomous hydrogen driven carbon fixation and alcohol synthesis occurring in a relatively dormant host (**Figure 1.1**). The first step in demonstrating the complete pathway for carbon fixation is demonstrating the first three steps; the carboxylation and hydrogen dependent reduction of acetyl-CoA to the key intermediate 3-HP which is also a valuable plastics precursor. Demonstrating the first portion of the 3-HP/4-HB pathway for carbon fixation is the first goal here. Demonstration of a butanol pathway alone within *P. furiosus* is the second goal, as butanol is a target molecule of the project. The challenge here is that no sufficiently thermophilic butanol pathway is known. Instead, a hybrid synthetic pathway from multiple gene donors is to be assembled into an artificial operon and expressed in *P. furiosus*. Finally, alternate strategies for the production of the simpler fuel target, ethanol, will be explored. Recently, a novel pathway for ethanol production has been discovered in *P.*

furiosus that depends on the native enzyme AOR operating in reverse to produce acetaldehyde from acetate which is reduced to ethanol by the highly active *Thermoanaerobacter* AdhA (1). Alternatively, *Thermoanaerobacter*, and many other bacteria, contain a bifunctional AdhE that is capable of the sequential reduction of acetyl-CoA to ethanol. Since acetyl-CoA is the intended link between carbon fixation and fuel synthesis, demonstrating activity and ethanol formation of AdhE alone is the third goal here. The ultimate objective is to combine these components to couple SH1, the complete 3-HP/4-HB cycle for carbon fixation, and an active ethanol or butanol synthesis pathway to demonstrate biofuel synthesis directly from CO_2 and hydrogen.

The current need and status of renewable liquid transportation fuels

The repercussions of fossil fuel use

Since the industrial revolution, fossil fuels have been exploited as a cheap and abundant source of energy (2). The result has been an explosion in the world's population (**Figure 1.2A**) and the average standard of living (3-7). However, their use is extremely problematic for two primary reasons: finite availability and negative environmental impacts.

Fossil fuels (coal, oil, and gas) occur as natural geological processes act upon ancient beds of biomass over the course of millions of years (8). Because they exist in finite quantities, and we are utilizing them at a much faster rate than they are naturally produced, fossil fuels suffer from the common economic problem of paucity. The finite nature of crude oil is typically discussed using the term "peak oil" which is the point in time which oil production irreversibly decreases and asymptotically approached zero (9). The timing of this phenomenon will depend on the amount of oil and the rate at which we utilize that oil. Total oil reserves for the planet

are estimated to be between 2 and 4 trillion barrels and are expected to peak between 2026 and 2047 (**Figure 1.3**) (10, 11). Crude oil alternatives such as unconventional oil and tar sand inject new vigor into an oil industry that is all too eager to cite that a 1919 U. S. Geological survey that concluded oil would run out in 9 years, and that new oil discoveries will sustain humanity for generations (12, 13). This is erroneous reasoning though as current oil production growth at 7% a year can consume a volume of oil equal to the entire planet in ~300 years. This means one way or another, oil production will decrease in the very near future. Moreover, cited alternatives such as Canadian tar sands and Venezuelan and Russian heavy oil are not suitable replacements for conventional oil as their processing uses up two out of every three barrels. The inescapable conclusion is that the energy needs of the current generation and their offspring will not be satisfied with fossil fuels (14, 15).

The second primary repercussion of fossil fuels is the negative environmental impact from their production and final use. While the environmental impacts of fossil fuels are hotly debated among popular media sources (16), the data behind these impacts and the discussion among scientists clearly describe an ongoing calamity that can and must be averted (17). A notable recent disaster in oil production is the 2010 explosion of the Deepwater Horizon exploratory drilling rig that killed 11 workers and spilled ~680,000 tons of oil. This disaster garnered much media attention and scientific investigation into the environmental impact of spilled oil from this incident and all incidents in recent history. From 1974 to 2010, 1,213 oil spills are responsible for the release of 9.8 million tons of oil (18). The cleanup of such a spill is also problematic as dispersants are used as an emergency measure to prevent shoreline

buildup of an ocean oil slick, but their effect on the microbial community and the ultimate breakdown of the hydrocarbons is unclear (19).

Spills, explosions, and other acute disasters garner much media attention and public outcry due to their shocking nature and ability to generate headlines; however, the primary problem with fossil fuel use, other than the fact that it is non-renewable, is the chronic problem of CO_2 emissions. The use of fossil fuels has been accompanied by a dramatic increase (40%) in atmospheric CO_2 (from 280 to 400 ppm, **Figure 1.2B**) (20). Carbon dioxide is a greenhouse gas (GHG) and its accumulation in our atmosphere is causing a warming trend known as global warming and broader impacts know as climate change. Climate change studies are extremely numerous and while they are not without dissension, it is abundantly clear that human activity is rapidly changing the composition of our atmosphere resulting in a warming trend and global consequences (17, 21, 22). Climate change is currently causing dry-season rain fall reductions that are comparable to the "dust bowl" era and a rise in sea-level due to thermal warming. Lower limit estimations for the sea level rise are 0.4-1.0 m if atmospheric CO_2 reaches 600 ppm and 0.6-1.9 m if 1000 ppm is reached. The additional contribution to the sea level rise from glacial and ice sheet melting is estimated to be several meters but is uncertain. Moreover, these effects will last for a millennium after humanity becomes carbon neutral (23). The effects of climate change are extensive, and the only way to mitigate their impact is to decrease and eliminate CO_2 emissions as rapidly as possible (24).

Climate change skeptics most often attempt to undermine this evidence by stating something to the effect of: this is, at least to some extent, part of the natural atmospheric

cycles that the Earth has been experiencing for billions of years. There are two fallacies in this argument: first is a misunderstanding of the magnitude of human versus natural GHG emissions, and second, is the misunderstanding that if something appears to be natural, then it is acceptable and no action should be taken because no harm will come. First is a misunderstanding of the magnitude of human GHG emissions that is clearly exercised by Australian ABC reporter Ian Pilmer who stated:

> *Human additions of CO_2 to the atmosphere must be taken into perspective. Over the past 250 years, humans have added just one part of CO_2 in 10,000 to the atmosphere. One volcanic cough can do this in a day.* (25)

The argument that volcanos emit more CO_2 than humans is simply false as humans emit roughly 135 times more CO_2 than all volcanic activity (26, 27). Also, an elegant comparison of human versus natural GHG emissions is now possible thanks to the 2010 eruption of the Icelandic volcano Eyjafjallajökull, which emitted a plume of 150,000 tons of CO_2 per day (28). This plume also caused the grounding of 60% of European air traffic (29, 30) which reduced CO_2 emissions by more than 200,000 tons of CO_2 per day. This major geological event is effectively made carbon negative by a partial disrupting in aviation which only accounts for 2.5-3% of anthropogenic CO_2 emissions (31). The second fallacy of climate skeptics is that if something appears to be natural, then it is acceptable and no action should be taken because no harm will come. By applying this faulty logic to climate change and the observation that the atmospheric CO_2 content has fluctuated in the past, climate skeptics conclude that the current observations about our environment are natural and nothing to worry about. This reasoning cannot be farther from the truth. Roughly 250 million years ago, a similar event occurred that, like today,

involved suddenly rising GHGs in the atmosphere. This event, known as the Permian-Triassic extinction event or simply the Great Dying, let to the extinction of up to 95% of life on Earth (32-35). Humans should make every effort to avoid replicating this natural event (36).

Alternatives to fossil fuels in transportation

Transportation related to fossil fuel use is estimated to account for 30-40% of anthropogenic CO_2 emissions (37). The largest single use of transportation energy (>50%) in the United States is gasoline consumption by light vehicles (38), and there are a variety of strategies to reduce emissions from personal vehicles.

Electricity

Plug-in electric vehicles (PEVs) are an alternative to gasoline powered internal combustion engines that utilize an on board battery that is charged from an external source to drive the wheels with an electric motor. PEVs are a rapidly growing market; for example, the Tesla Model 3® received over 325,000 pre-orders worth approximately $14B in the first week following its announcement (39). While some electric vehicles tout badges to the effect of "zero emissions", this does not take into account that two thirds of our electricity comes from the burning of coal and natural gas (40). Still, electric vehicles are an economical way to reduce emissions. For the sake of comparison, let us consider the Tesla Model 3® which can travel 215 miles on 66 kw-hrs of electricity and the Chevy Cruz® which can travel 462 miles (33mpg) on 14 gallons of gasoline (39, 41). At these values, the electricity needed to drive the Tesla Model 3® one mile will cost $0.05, and the generation of that electricity will have released 52 g of CO_2 (39, 42). Driving a gasoline powered Chevy Cruze® that same mile will cost $0.08 and directly

emit 270 g of CO_2 (**Figure 1.4**; based on California averages) (40, 41). Despite ongoing innovations by companies like Tesla®, many consumers suffer from range anxiety and are reluctant to transition. Hence the critical flaw of PEVs is they require batteries. The production of such batteries, as well as the photovoltaic panels to charge them, requires the large scale mining of lithium and rare earth elements. This problem is mitigated by the fact that the mining would be at a much smaller scale since it is not pursuing an energy source and the raw materials can be efficiently recycled (43, 44). Charging a battery requires more time than filling a tank with a liquid, and even the most advanced batteries have a relatively low specific energy. The Tesla Model S® offers an 85 kWh battery that weighs 1,200 lbs (39). This corresponds to a specific energy of 0.65 Mj/kg, compared to 48 Mj/kg for gasoline (38) (**Table 1.1**). Despite this, PEVs are a viable and popular solution for routine personal transportation needs such as commuting to and from work. The permanent batteries and in-home charging work well for this type of routine transportation but cannot match the versatility that comes with liquid fuels. In the long-term, PEVs will continue to grow and likely dominate ground transportation; however, liquid fuels will always be useful due to the ease of refueling and the range imparted by the high specific energy.

Hydrogen

Hydrogen is an appealing fuel because its oxidation does not release carbon dioxide and it has an extremely high specific energy of 120-142 Mj/kg (45). For vehicle applications, hydrogen can be combusted to drive a piston, but it is typically oxidized in a fuel cell to generate an electric current which is subsequently used to drive an electric motor. Hydrogen is

not an energy source but an energy carrier as it must be produced by some other energy source. Nearly all hydrogen is produced from natural gas which is a non-renewable resource and contributes to CO_2 emissions and subsequently climate change. There are a wide range of renewable strategies for hydrogen production including the electrolysis of water using renewable generated electricity, thermochemical and thermal/photo hybrid solar water splitting, photocatalytic materials, photobiological, and carbohydrate catabolism to hydrogen and fully oxidized carbon (46-49). Currently, solar driven electrolysis of water is the leading strategy for renewable hydrogen, and as of early 2016 there are 22 hydrogen filling stations operational in the continental United States (45). Despite a very high specific energy (120-142 Mj/kg), hydrogen has a very low physical density (0.025 kg/L at 5,100 psi) which means its energy density is quite low (3-3.5 Mj/L) (**Table 1.1**). The result is even a large tank capable of holding 80 L of compressed hydrogen would only be able to hold the energy equivalent of two gallons of gasoline (45). Thus hydrogen powered vehicles, along with PEVs, suffer from range limitations due to the on vehicle storage of the fuel. Advanced materials offer improvements in the storage density of hydrogen by reversibly converting molecular hydrogen to more compact hydrides and/or protons (50). Improvements in storage, along with the ease of refueling, make hydrogen a promising energy carrier for ground transportation.

Ethanol

Ethanol is a fuel that can be produced by the fermentation of renewable plant-based sugars. As a liquid, it has a high specific energy (26.8 Mj/kg) and a high energy density (21.2 Mj/L), which are desirable characteristics for a transportation fuel. The sugar-based ethanol

industry is highly developed, and in 2014 produced 24.5 billion gallons worldwide to account for 94% of all biofuel production. This industry is led by the United States and Brazil using corn starch and sugar cane as feedstocks to produce 14.3 billion gallons and 6.2 billion gallons respectively (45).

In the United States, sugar-based ethanol is ubiquitously found as a 10% blend with gasoline (E10) and, to a lesser extent, as an 85% blend (E85) that requires engine modification, which overall accounts for 9% of all gasoline sold by volume (45). With the market for ethanol saturated, efforts now have shifted to the use of alternate feedstocks. Currently, ethanol is produced only from the edible portion of certain plants. In addition to competing with food production, life cycle analysis of reveals that the production and use of corn starch ethanol releases roughly 70-100% as much carbon dioxide as the production and use of the energy equivalent amount of traditional gasoline (**Table 1.1**). This analysis describes the carbon intensity of an energy source/carrier using the unit grams of CO_2 equivalent per megajoule (gCO_2e/MJ). Traditional petroleum has a carbon intensity of 90-100 gCO_2e/MJ, where corn starch ethanol direct emissions total 40-90 gCO_2e/MJ (45). While there is a great deal of variability in the reporting of carbon intensities, and the inclusion of land use change into these analyses is quite controversial (+30 gCO_2e/MJ), corn starch ethanol appears to have only a minor effect on the release of GHG's and therefore only a minor value as a biofuel. Additionally, Chinese corn ethanol has a carbon intensity of 370 gCO_2e/MJ, making it far worse than gasoline (51).

In order to reduce both the competition with food production and the carbon intensity of the fuel, global efforts have pushed to use alternate feedstocks such as non-edible corn stover, or non-agricultural plants such as poplar and switchgrass (52). The United States Energy Independence and Security Act of 2007 (EISA 2007) and the European Parliament directive 2009/28/EC outline the strategies to accomplish these goals by, among other things, phasing in cellulosic and lignocellulosic ethanol (53, 54). The United States was far behind the targeted cellulosic ethanol production as the first (heavily subsidized) commercial scale cellulosic ethanol plants are only now beginning to open (52); however, in the years since the passing of EISA the targets for biofuel production have been greatly reduced from 250 million gallons of cellulosic biofuel to 6 million gallons (55). The US in now in compliance with its goals, but this was only made possible by lowering the bar.

This industry has led to important developments in the renewable production of bio-alcohols. The current target alcohol, ethanol, has some intrinsic drawbacks that mean an eventual shift to a higher alcohol may ultimately be necessary. Ethanol is hygroscopic, and the water vapor it absorbs is corrosive to the current infrastructure, particularly pipelines, that were designed to accommodate petroleum products (56, 57). Among other liquid transportation fuels, ethanol has a relatively low energy density, which results in E85 driven vehicles suffering from a 25-30% reduction in fuel economy (41). Despite these undesirable fuel characteristics, ethanol is the easiest liquid fuel to produce at high yields, and improving the feedstocks used to make ethanol is a desirable strategy since it can effectively and safely be used to dilute gasoline.

Advanced biofuels

Advanced biofuels are naturally-occurring compounds, usually hydrocarbons, which have physical properties that more closely mimic the physical properties of the fossil fuels they are designed to replace. Terpenes are a diverse group of hydrocarbons, based on units of isoprene, which are a potential precursor to advanced biofuels (58, 59). Terpene based biofuels include sabinene, which has been produced to a maximum yield of 82 mg/L in *Escherichia coli*, (60), sesquiterpene, which has been produced to a maximum yield of 80 mg/L in *Saccharomyces cerevisiae*, (61), and amorphadiene and kaurene, which have been produced to a maximum yield of 313 mg/L and 20 mg/L respectively in *E. coli* (62). These chemicals have complex structures and are difficult to synthesize. As a result, they have yet to be industrially produced as a biofuel. Interestingly, the related compound artemisnin is an effective antimalarial drug, and its precursor, artemisinic acid, has been produced to a maximum yield of 250 mg/L in *S. cerevisiae* (63). Linear alkanes are a simpler target molecule as they can be produced by derivatizing fatty acid precursors. The introduction of cyanobacterial acyl-ACP reductase and aldehyde decarbonylase into *E. coli* resulted in the accumulation of 40 mg/L alkanes (64). *Clostridium acetobutylicum* fatty acyl-CoA reductase and *Arabidopsis thaliana* fatty aldehyde decarbonylase were also used to engineer *E. coli* to produce 328 mg/L nonane, 137 mg/L dodecane, 65 mg/L tridecane, 43 mg/L 2-methyl-dodecane, and 9 mg/L tetradecane (65). The benefit of producing alkanes such as these is their physical properties are nearly or completely identical to the gasoline, diesel, and aviation fuel they aim to replace.

Butanol

Butanol is another alcohol that, like ethanol, can be renewably produced as a fermentation end product. It is a more desirable target molecule than ethanol because it has a higher specific energy (36 Mj/kg), it is not hygroscopic, and it can be used in higher blend ratios (up to 100%) without engine modification (41). Historically, butanol has been produced commercially via acetone-butanol-ethanol (ABE) fermentation in solventogenic strains of *Clostridium* (66, 67). This biological industry was halted in the mid-20th century due to an inability to compete with cheaper petroleum based processes (45). Now, with interest in butanol as a biofuel and advances in genetics, the study of these organisms and their pathways has been revived with vigor and urgency (67). While the majority of butanol production today, roughly 1 billion gallons worldwide per year, is dedicated to its use as an industrial solvent and chemical intermediate and uses petroleum as a feedstock (68), several companies such as DuPont, BP, and Gevo are involved in renewably producing and selling butanol. However due to low yields and the low price of fuel, this bio-butanol is primarily sold as a solvent to maintain profitability (45).

Electrofuels

Electrofuels is a program under the US Department of Energy's Advanced Research Project Agency – Energy (69). The goal of the electrofuels program is to increase the efficiency of renewable liquid transportation fuel production by eliminating the current dependence on plant based photosynthesis and carbohydrate intermediates (**Figure 1.5**). The paradigm shifting electrofuels strategy is based on exploiting microbial pathways found in nature for the

assimilation of carbon dioxide. Energy carriers such as electric current, hydrogen, formate, and carbon monoxide can be generated efficiently and renewably from solar photons and can be used to directly drive carbon fixation and fuel synthesis pathways (70-72). This strategy is essentially to build a fuel molecule directly and one carbon at a time is of stark contrast to the current sugar-based strategy.

The extremely thermophilic archaeon *Metallosphaera sedula* (T_{opt} 73°C) and the moderately thermophilic bacterium *Thermoanaerobacter ethanolicus* (T_{opt} 69°C) are candidates for gene donors for the production of electrofuels and commodity chemicals in the hyperthermophilic archaeon *P. furiosus* (T_{opt} 100°C). Thermophiles, and their enzymes, are desirable for biotechnological applications due to their intrinsic stability against denaturation (73). In an industrial fermentation setting, losses due to contamination and cooling of metabolic heat by refrigerated-glycol are major expenditures for mesophilic processes (74, 75). Thermophiles are potentially superior to mesophiles due to their reduced risk of contamination and feasible use of air cooling to dissipate metabolic energy, and are therefore targeted for use in the production of electrofuels.

Thermophilic Microorganisms

Having a temperature optimum is a canonical feature of all known forms of life. It is common knowledge that the average internal body temperature for a human is 37°C (98.6°F) with life threatening symptoms occurring just a few degrees outside this temperature (76). This phenomenon is also observed in the microbial world and is characterized as the optimal growth temperature (T_{opt}), or the temperature at which any particular microbial culture most rapidly

divides. While the exact nomenclature varies, 'psychrophilic' usually refers to an organism with a T_{opt} below 15°C; 'mesophilic' organisms have a T_{opt} between 15 and 45°C; and 'thermophilic' organisms grow optimally above 45°C. However, the temperature range for thermophily is very broad, extending to 122°C, the current upper temperature limit of life (77). Therefore, thermophiles have been further divided into moderate thermophiles (T_{opt} 45-70°C) and extreme thermophiles (T_{opt} ≥ 70°C, e.g. *Caldicellusiruptor bescii*); special cases of extreme thermophiles growing with a T_{opt} ≥ 80°C (e.g. *Pyrococcus furiosus*) are typically referred to as hyperthermophiles (78, 79) (**Figure 1.6**). While the lower limit of life is debated, psychrophiles are found well below 0°C (80, 81).

The likely first use of the term "thermophile" was by Miquel in 1888 and was used to describe an organism that apparently thrived at temperatures that would destroy all other known forms of life (82). This first thermophile to be described, *Bacillus thermophilus*, was abundant in air, rain water, and river water and grew optimally at 65-70°C (83). Since their discovery, these thermophilic microorganisms have been catalogued and characterized with great interest alongside other forms of life (**Figure 1.7**) (84, 85). Perhaps the most well-known application of a thermophilic enzyme is in the polymerase chain reaction (PCR) which exponentially amplifies a specific sequence of DNA. During this process, the enzyme DNA polymerase is repeatedly heated past the melting temperature of DNA and nearly to the boiling point of water. As such, this process required the use of extremely thermostable DNA polymerases, such as those from *Thermus aquaticus* (*Taq*, T_{opt} 80°C) and *P. furiosus* (*Pfu*, T_{opt} 100°C). The development of this powerful and ubiquitous molecular technique earned Dr. Kary

Mullis one half of the 1993 Nobel prize in chemistry (86). Other molecular applications of thermophilic enzymes include: thermophilic DNA ligases for ligase chain reactions, *T. aquaticus* serine protease PreTaq which is used in the purification of DNA or RNA, and *P. furiosus* protease S for digesting proteins for peptide sequencing (87). These enzymes are widely used in industrial processes due to both their unique biochemistry and intrinsic stability. Some of the many applications of thermophilic enzymes include: alcohol dehydrogenases for chiral synthesis, amylases and pullulanases for the production of high-glucose syrup, cellulases in pulp and paper processing as well as in laundry detergents, DNA polymerase for PCR, glucoamylase in starch conversion, glycosidase in the hydrolysis of lactose and the synthesis of certain detergents, laccases xylanases in textile bleaching and laundry detergents, ligases in ligase chain reaction, lipases in detergents, and isomerases in the production of high-fructose corn syrup (88, 89). While this list of applications is hardly exhaustive, it demonstrates the wide breadth of opportunities offered by thermophiles and their enzymes.

Extremophiles, organisms capable of surviving in an extreme environment, are not limited to tolerance of heat (thermophiles) and cold (psychrophiles), but also commonly include tolerance to extremes of high pH (alkaliphiles), low pH (acidophiles), salt (halophiles), pressure (barophilic), and radiation (90). The domain of life Archaea is famous for containing many of these extremophiles (91) and is important to understand since the candidate host *P. furiosus* and one candidate gene donor *M. sedula* are members of this domain.

Archaea

All known living organisms are currently divided into three domains as the Eukarya, Bacteria, and Archaea (92-95). These domains can be organized into a phylogenetic tree (**Figure 1.7**) by comparing the sequences of ribosomal RNA. The first Archaea studied were methanogens and were originally classified as "archaebacteria" as they were believed to be bacteria. They were later reclassified to initiate the domain Archaea (95, 96). Through phylogenetics, Archaea were found to be more closely related to eukaryotes (97, 98) than bacteria, and genome sequencing reveals them to be a "genetic mosaic" of the other two domains. The most striking similarity between archaea and eukaryotes is in information processing. While eukaryotic processes are generally more complex, the core machinery of DNA replication, RNA transcription, and protein translation are very similar across archaea and eukaryotes (99). Archaea and bacteria constitute the prokaryotes as they lack a central membrane bound nucleus, though this term is becoming obsolete. They both also lack organized membrane bound organelles and are exclusively single-celled.

While the information processing machinery of archaea and bacteria are quite distinct, many of the metabolic processes that maintain the cell are quite similar (99-101). Despite these many similarities to the other two domains, the Archaea are not simply a hybrid domain between the Bacteria and the Eukarya. A notable distinction is the composition of the archaeal membranes, which are branched isoprene chains with an ether linkage to an (*S*)-glycerol phosphate whereas bacterial and eukaryotic membranes are composed of fatty acid chains with an ester linkage to a (*R*)-glycerol phosphate (**Figure 1.8)**. This distinctive lipid composition

affords greater stability and may begin to explain why archaea are commonly found in extreme environments across the planet (91, 102).

The domain Archaea constitutes a wide variety of the microbial diversity found on Earth. One archaeon in particular, *P. furiosus*, is of keen interest as a candidate host organisms for the production of biofuels and commodity chemicals. In order to exploit this organism as a host, an in-depth understanding is required to be able to predict and test how its metabolism will interact with and respond to a heterologous pathway.

Pyrococcus furiosus

P. furiosus is a hyperthermophilic archaeon (**Figures 1.6 and 1.7**) that was isolated off the coast of Italy in marine sediments that were geothermally heated to 90-100°C. It is a strictly anaerobic heterotroph that grows between 70 and 103°C with an optimal growth temperature of 100°C and a doubling time of 37 minutes. It metabolizes carbohydrates such as maltose to primarily H_2, CO_2, and acetate at 95 °C. Acetoin is produced at suboptimal temperatures (70-80°C) and small amounts of alanine and ethanol are also produced (102-105). In addition to proteolytic growth, *P. furiosus* grows saccharolytically on a range of carbohydrates including starch, glycogen, maltose, cellobiose, and pyruvate. It cannot utilize polymers such as chitin or cellulose or monosaccharides like glucose, fructose, ribose, or galactose for growth (102, 106). While *P. furiosus* cannot grow on glucose, it can grow on the glucose disaccharides maltose or cellobiose. Once these disaccharides are transported into the cell and cleaved to glucose monomers, that glucose can be metabolized by the *P. furiosus* glycolytic pathway (**Figure 1.9**). Though currently available strains do not grow on chitin, an abundant carbohydrate in the

marine environment, the deletion of a single nucleotide at position 1006 in PF1234 repairs a frameshift mutation in the enzyme chitinase and completely restores its ability to grown on chitin (107). Carbohydrates are imported into the cell by an ABC transporter that is specific to the different types of sugar. In the case of cellobiose, CbtA, a high affinity transporter belonging to the Opp family of ABC transporters, is up-regulated along with four other downstream permeases (108). This is believed to be responsible for the import of the β-1,4 glucose disaccharide. Inside the cell, two β-glucosidases, CelA and CelB, hydrolyze the disaccharide into glucose (109). In the presence of maltose, expression of two highly similar operons, Mal-I and Mal-II, are upregulated and are believed to be responsible for the import of the α-1,4 glucose disaccharide (110-112). Once imported, a single 125 kDa α-glucosidase breaks down maltose into glucose (104, 109, 113).

P. furiosus metabolizes sugars via a modified Embden-Meyerhof glycolytic pathway (**Figure 1.9**) which converts one molecule of glucose to two molecules of pyruvate. These are further metabolized to two molecules of acetate, two molecules of CO_2, two equivalents of ATP, and eight electrons in the form of reduced ferredoxin (114-116). For the activation of glucose to fructose-1,6-phosphate, ADP is used as the phosphate donor to yield AMP and the phosphorylated carbohydrate (103). Glyceraldehyde-3-phosphate is oxidized directly to 3-phosphoglycerate by the enzyme glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) utilizing ferredoxin (Fd) as a single electron acceptor. This is a significant difference from the conventional glycolytic pathway as not only is the electron acceptor different, but direct production of 3-phosphoglycerate reduces the net ATP yield from substrate level

phosphorylation (117). Pyruvate is oxidatively decarboxylated to acetyl-CoA via pyruvate ferredoxin oxidoreductase (POR) which also utilizes ferredoxin as the electron acceptor (118).

The final critical step in the metabolism of sugars by *P. furiosus* is the conversion of acetyl-CoA to the end product acetate by acetyl-CoA synthetase (ACS). This step is critical to *P. furiosus* because it is where appositive net of ATP is finally produced yielding 2 ATP/glucose (119). This enzyme and its isoforms are also very relevant to the use of *P. furiosus* as a host because they can interfere with a heterologous pathway by degrading potential intermediates such as acetyl-CoA, propionyl-CoA, and succinyl-CoA. A detailed understanding of the specificities of ACS and its isoforms is therefore necessary to predict how the metabolism of *P. furiosus* will interact with a heterologous pathway.

Two isoforms of ACS, ACS1 and ACS2, have been purified from *P. furiosus* and both have an $\alpha_2\beta_2$ structure (119). The genome of *P. furiosus* contains genes encoding an additional three α subunits. The two β subunits and the five α subunits can form a possible ten isozymes. These possible combinations were produced by heterologously expressing the subunits in *E. coli*. The characterization of these isozymes revealed that the β subunit determined ATP vs GTP specificity, and the α-subunit determined acyl-CoA specificity. ACS isozymes that contain the β1 subunit, as with the native ACS1, prefer GTP/GDP as a substrate, and ACS isozymes that contain the β2 subunit, as with the native ACS2, prefer ATP/ADP as a substrate. *In vitro*, these enzymes are characterized in the reverse direction where the organic acids are supplied as the substrate. The specificity for these organic acids is determined by which α subunit is present. Acetate is a substrate for α1, α2, α4, and α5 containing isozymes. 3-Methylthiopropionate, which

corresponds to the sulfur contains amino acid methionine, is a substrate for α2, α3, α4, and α5 containing isozymes. The branched chain amino acid related substrates isobutyrate, isovalerate, and 2-methylbutyrate substrates are substrates for α1, α2, α3, and α5 containing isozymes. The aromatic amino acid related substrates phenylacetic acid and 4-hydroxyphenylacetic acid are substrates for α2, α4, and α5 containing isozymes. Indole-3-acetic acid, which is also related to an aromatic amino acid, is a substrate for α2 and α4 but not α5 containing isozymes. Imidazole-4-acetic acid is specific to α4 containing isozymes, and succinate is specific to α3 containing isozymes (120). The specificity and redundancy of these enzymes may be important as they can potentially be competitors with engineered pathways.

P. furiosus does have a simple respiratory complex to generate further ATP in the recycling of reduced electron carriers that are produced from the glycolytic pathway and pyruvate oxidation. Membrane bound hydrogenase (MBH) utilizes reduced ferredoxin (Fd_{red}) to generate molecular hydrogen (**Figure 1.9**). This energy conserving complex also forms an ion gradient which is coupled to a Na^+ ATP-synthase to yield an additional ~1.2 ATP/glucose (121, 122). While most molecular hydrogen produced is an end product, a portion of it can be recycled by soluble hydrogenase 1 (SH1) to generate NADPH for biosynthesis (123-125). This enzyme by itself is of keen biotechnological interest and its presence in *P. furiosus* heightens the desire to use this as a platform organism for metabolic engineering.

P. furiosus has been reported to produce trace amounts of labeled ethanol at 90°C in cell suspensions (103) and ~1 mM ethanol in cultures grown at 72°C (1). Target molecules for biofuel production are commonly simple alcohols such as ethanol and butanol, and the terminal

step in their biosynthesis involves an alcohol dehydrogenase (ADH). Alcohol dehydrogenases are an extremely broad group of redox enzymes that carry out the interconversion of aldehydes and ketones with their corresponding alcohols. Any strategy to engineer *P. furiosus* to produce alcohols such as ethanol, butanol, or 2,3-butanediol must take into account the specificity, activity, and expression of ADHs native to *P. furiosus*.

Alcohol dehydrogenases native to *P. furiosus*

The genome of *P. furiosus* contains genes that encode four ADHs that are identified by the order in which they appear in the genome: AdhA (PF0074), AdhB (PF0075), AdhC (PF0608), and AdhD (PF0991). Using microarray data (**Table 1.2**), *P. furiosus* AdhA and AdhB are expressed at very low levels at both 95°C and at 72°C whereas AdhC and AdhD are expressed at high and moderate levels respectively at both temperatures. Only AdhA and AdhD from *P. furiosus* have been characterized (126, 127).

AdhA, a short chain ADH, and AdhB, and iron-containing ADH, are adjacent in the genome of *P. furiosus* and in-between two glycosyl hydrolases, CelB and LamA. The genes for these two ADHs were cloned into *Escherichia coli* for production of the recombinant enzymes. AdhB, which oxidizes methanol and ethanol, rapidly lost activity during preparation due to the loss of its iron which prevented its biochemical characterization. Unlike AdhB, AdhA is extremely stable with a melting temperature of 102.7°C. AdhA is a short chain ADH that utilizes NADP(H) as a cofactor and preferentially oxidizes secondary alcohols with maximal activity (46 U/mg) on 2-pentanol. AdhA also had significant reductive activity on pyruvaldehyde (32 U/mg), but no other aldehyde or ketone tested had more than ~10% of this activity (126). The

oxidation of 2-pentanol yields methyl propyl ketone, and the reduction of pyruvaldehyde yield hydroxyacetone if the aldehyde is reduced or lactaldehyde if the ketone is reduced. However, the physiological roles that these reactions play in *P. furiosus* are not clear. Despite this, the ability of *P. furiosus* to oxidize alcohols with potential applications as a biofuel or commodity chemical such as 2-pentanol, must be considered in any future attempts to produce such alcohols.

P. furiosus AdhD, also known as threonine dehydrogenase (TDH), has been characterized following purification of the recombinant form from *E. coli*. AdhD is a secondary ADH and is specifically uses NAD(H) as an electron carrier. It is most active on L-threonine and 2,3-butanediol (10.3 and 9.7 U/mg respectively) in the oxidation reaction. Due to its instability, 2-aminoacetoacetate, which is the corresponding ketone to threonine, could not be tested. However, acetoin (3-hydroxy-2-butanone), which is the corresponding ketone to 2,3-butanediol, is stable and was found to be reduced at a V_{max} of 3.9 U/mg. The physiological role of AdhD is suggested to be threonine catabolism (127), but its other activities must be considered in any attempt to engineer *P. furiosus* to produce biotechnologically valuable propane and butane diols. Acetoin has recently been found to be a low temperature (70-80°C) end-product of *P. furiosus* metabolism (105). Overexpression of AdhD may result in the production of the commodity chemical 2,3-butanediol, if it is not already produced. Conversely, if AdhD favors the oxidation of 2,3-butanediol *in vivo* as it does *in vitro*, then it would likely have to be knocked out as part of a strategy to convert acetoin to 2,3-butanediol.

Thermococcus paralvinellae (f. *Thermococcus* strain ES1) contains an ADH that is homologous (>90% coverage and >30% identity) to both AdhB and AdhC from *P. furiosus* (128). This is an iron-containing ADH that forms a homotetramer with 1.0 ± 0.1 Fe/subunit. Kinetic analysis showed that it is a primary ADH specific for NADP(H) as an electron carrier. It is active on the primary alcohols ethanol (100%), 1-propanol (120.6%), 1-butanol (154.6%), 1-pentanol (184%), 1-hexanol (137.8%), 1-heptanol (94%), 1-octanol (68.8%), and 2-phenylethanol (40.2%). No activity was seen with methanol, glycerol, or the secondary alcohols 2-propanol, 2-butanol, or 2-pentanol. It is also has lower K_M values for acetaldehyde (1.0 mM) and NADPH (7 μM) than ethanol (10.4 mM) and NADP (60 μM) (129). Despite the lack of direct characterization of the two homologs in *P. furiosus* AdhB and AdhC, the great disparity in expression between them indicates AdhC as the dominant ethanol forming ADH in *P. furiosus*.

The native hydrogenases and ADHs in *P. furiosus* will likely play a role in a strategy to metabolically engineer *P. furiosus* to produce a biofuel or commodity chemical. However, the most important aspect of such a strategy is the genetic system itself. With a genetic system, not only can native genes such as the hydrogenases and AHDs be overexpressed or knocked out, but foreign genes can be introduced and exploited.

Genetics in *P. furiosus*

A major challenge in the development of a genetic system for *P. furiosus*, and other extremely thermophilic archaea, is the means to impose a selective pressure upon the organism in order to generate positive transformants. A common strategy for genetics in microorganisms is the use of antibiotics; however, antibiotics and their corresponding resistance markers are

commonly unstable at temperatures at which extremely thermophilic organisms grow. Simvastatin represents an important exception to the rule that antibiotics are only effective against mesophilic bacteria. It affects thermophilic archaea because it targets the production of archaeal membranes (130, 131), and displays no detectable degradation at 100°C in the absence of oxygen (132). These important characteristics have allowed simvastatin to play a major role in the development of genetic systems in the hyperthermophilic archaea *Thermococcus kodakarensis* (131) and *P. furiosus* (133). Nutritional selection, specifically uracil prototrophic selection, is the current strategy for the genetic modification of *P. furiosus* (**Figure 1.10**). Synthesis of uracil involves enzymes encoded by the *pyrE* (orotate phosphoribosyltransferase) and *pyrF* (orotidine-5'-phosphate decarboxylase) genes that, besides their role in uracil production, also convert the synthetic chemical 5-fluoroorotic acid (5-FOA) into the cytotoxic fluorodeoxyuridine (134, 135). Therefore, growth of strains with functional uracil pathways on media containing 5-FOA and uracil selects for natural mutants with disruptions in *pyrE* or *pyrF* (134, 136). The resultant uracil auxotroph, once purified, is now the parent strain for genetic manipulation via uracil prototrophic selection (137).

This powerful genetic system affords the opportunity to use *P. furiosus* as a host for metabolic engineering. The extremely thermophilic archaeon *M. sedula* and the moderately thermophilic bacteria *T. ethanolicus* are candidates for gene donors for the production of biofuels and commodity chemicals in the hyperthermophilic archaeon *P. furiosus*. While all of these organisms are thermophiles, there is an up to 30°C disparity between their optimal growth temperatures. Specifically, the host *P. furiosus* grows optimally at higher temperatures

than the gene donors. In order for these foreign genes to be active, *P. furiosus* must be incubated at suboptimal temperatures, and in order to exploit *P. furiosus* as a host, we must understand how it behaves at these suboptimal temperatures.

The effect of suboptimal temperatures on *P. furiosus* and its metabolism

P. furiosus has a very broad temperature range for growth – ranging from 70°C to 103°C with a T_{opt} from 90-100°C. Under optimal conditions, *P. furiosus* can have a doubling times as low as 37 minutes (102), but this increases to approximately 5 hours at 72°C. A cold-shock, where a mid-log phase culture is switched from a high to low temperature such as 95°C to 72°C, will cause a second lag phase before growth resumes at the slower rate (138) (**Figure 1.11**). Similarly, and predictably, the *in vitro* activity of *P. furiosus* enzymes decreases at suboptimal temperatures (**Figure 1.12**). This has been measured for the enzymes POR, ACSI, ACSII, aldehyde ferredoxin oxidoreductase (AOR), and SH1 (117, 119, 139, 140).

If *P. furiosus* is to serve as host to genes and pathways from archaea and bacteria with optimal growth temperatures 20-40°C below that of the itself, then the effect of these suboptimal temperatures must be considered. The reduced rate of metabolic activity and cell division means longer incubation times will be required for active enzyme expression and *in vivo* product formation that depends on the glycolytic pathway of *P. furiosus*.

The extremely thermophilic archaeon *M. sedula* and the thermophilic bacterium *T. ethanolicus* are candidates for gene donors for the production of biofuels and commodity chemicals in *P. furiosus*. *M. sedula* has a cycle for carbon fixation that may be an efficient route for carbon assimilation for biofuel synthesis. This cycle also contains intermediates such as 3-

HP, succinate and 4-HP which are valuable commodity chemicals (141). Even if the core genes are known for the *M. sedula* cycle for carbon fixation, we must still understand the metabolic context in which it operates, particularly accessory genes and branch points, if it is to be exploited in *P. furiosus*.

Metallosphaera sedula and a thermophilic pathway for carbon fixation

M. sedula is an extremely thermoacidophilic archaeon that was isolated from an aerobic acidic (pH ~2) runoff from a hot water pond (25-52°C) at Pisciarelli Solfatara (near Naples, Italy). Following enrichment at 65°C, isolate TH2 was purified by repeated serial dilutions in mineral medium containing pyrite and was named *Metallosphaera sedula*. This organism is an obligate aerobe that grows optimally between pH 1.0 and 4.5 at 75°C with a doubling time of 5.25 hours. Under heterotrophic conditions, it utilizes complex organic substrates such as beef extract, casamino acids, peptone, tryptone, and yeast extract, but not sugars. It is capable of growing lithoautrophically on single sulfide ores such as pyrite, chalcopyrite and sphalerite, on the synthetic sulfides CdS, SnS and ZnS, and on elemental sulfur where it produces sulfuric acid to a final concentration of 45 mM. *M. sedula* is named after its ability to mobilize metals. During growth on sulfide ore, it mobilized 100% of the copper, uranium, and zinc within 3 weeks (142). *M. sedula* was later found to grow most rapidly under mixotrophic conditions where the doubling time dropped to 3.7 hours compared to strictly heterotrophic or autotrophic growth that have doubling times of 5.0 and 11-13 hours respectively (143).

M. sedula and other autotrophic *Chrenarchaeota* lack key enzymes for the Calvin cycle, the reductive citric acid cycle, and the reductive acetyl-CoA cycle. *M. sedula* does, however,

have significant biotin dependent acetyl/propionyl-CoA carboxylase activity that can be detected in autotrophically grown *M. sedula* cell extracts. Acetyl-CoA carboxylase activity is unexpected in archaea since they do not contain fatty acids. Radiolabeling experiments in *M. sedula* extracts demonstrate the sequential conversion of acetyl-CoA to malonyl-CoA to 3-hydroxypropionate to propionyl-CoA to succinate, indicating the presence of a 3-hydroxypropionate cycle for carbon fixation (144). The key enzyme of this cycle is the acetyl/propionyl-CoA carboxylase, and it is active from 40-80°C with an optimal temperature of 75°C. The enzyme followed Michaelis–Menten kinetics with apparent K_M values of 0.06 mM for acetyl-CoA and 0.07 mM for propionyl-CoA. The K_M values for ATP and HCO_3^- were 0.04 mM and 0.3 mM respectively. The product of acetyl-CoA carboxylation is malonyl-CoA, and the product of propionyl-CoA carboxylation is methylmalonyl-CoA. The enzyme consists of three subunits and is suggested to have an $(\alpha\beta\gamma)_4$ composition (145). Extracts of *M. sedula* also contain NADPH-specific malonyl-CoA reductase activity that is highly upregulated and active under autotrophic conditions. This enzyme is conserved across autotrophic *Sulfolobales*, but it is distinct from *Chloroflexus aurantiacus* malonyl-CoA reductase. This indicates a partial convergence and overlap of the *Chloroflexus* and *Sulfolobaceae* 3-hydroxypropionate pathways (146, 147).

The pathway in *M. sedula* was found not only to be a distinct pathway from that in *Chloroflexus*, but it is a novel archaeal pathway for autotrophic carbon fixation known as the 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) pathway (**Figure 1.13**). The cycle is initiated by the carboxylation of acetyl-CoA to malonyl-CoA. Malonyl-CoA is reduced to malonate

semialdehyde via NADPH and coenzyme A is released. Malonate semialdehyde is reduced via NADPH to the key intermediate 3-hydroxypropionic acid (3-HP). 3-HP is activated with coenzyme A to form 3-HP-CoA by the hydrolysis of ATP directly to AMP. 3-HP-CoA is dehydrated to acryloyl-CoA which is subsequently reduced via NADPH to propionyl-CoA. The bifunctional carboxylase then carboxylates propionyl-CoA to methylmalonyl-CoA by again hydrolyzing ATP to ADP. At this point, the C2 unit acetyl-CoA has been carboxylated twice to a C4 unit. Methylmalonyl-CoA is rearranged by an epimerase and isomerase to succinyl-CoA before the bifunctional malonyl/succinyl-CoA reductase reduces it to succinic semialdehyde by consuming NADPH and releasing coenzyme A. Succinic semialdehyde is then reduced to the second key intermediate 4-hydroxybutyrate (4-HB) via NADPH. 4-HB is then activated with coenzyme A to form 4-HB-CoA by the hydrolysis of ATP directly to AMP. 4-HB is dehydrated to crotonyl-CoA then hydrated to 3-HB-CoA. 3-HB-CoA is then oxidized to acetoacetyl-CoA, generating NADH in the process. Acetoacetyl-CoA, the final C4 intermediate, is then hydrolyzed to two molecules of acetyl-CoA. In total, one molecule of acetyl-CoA, six equivalents of ATP, and eight electrons (ten from NADPH and two returned by NADH) are required to fix two molecules of fully oxidized carbon and produce two molecules of acetyl-CoA. (148-154)

M. sedula is a potential gene donor for the production of biofuels and commodity chemicals in *P. furiosus* due to the presence of the 3-HP/4-HB cycle for carbon fixation. This cycle contains the intermediates 3-HP, succinate, and 4-HB which are valuable commodity chemicals, and the cycle in its entirety can serve as an efficient means of carbon assimilation for biofuel synthesis. In order to engineer *P. furiosus* to produce a biofuel, an additional gene

donor is required. The biosynthesis of ethanol and butanol, which are common biofuel target molecules, requires an ADH. *T. ethanolicus* is a candidate gene donor for biofuel synthesis because it is an ethanologenic bacterium that contains highly active ADHs (155).

Thermoanaerobacter ethanolicus and a thermophilic pathway for fuel synthesis

Thermoanaerobacter ethanolicus strain JW200 was isolate from hot springs located in Yellowstone National Park and was proposed as the type strain of the new genus *Thermoanaerobacter*. *T. ethanolicus* is an extremely thermophilic anaerobic bacterium that ferments readily available carbohydrates to primarily ethanol and CO_2 and is capable of producing 0.5 g ethanol/L/hr at a yield of 1.37 mole ethanol produced per mole xylose consumed (156, 157). It has a broad growth ranges for pH (5.8-8.5) and temperature (37-78°C) with an optimal growth temperature of 69°C. Viable growth substrates include: glucose, fructose, mannose, galactose, ribose, xylose, lactose, sucrose, maltose, cellobiose, starch, and pyruvate; but not cellulose, raffinose, rhamnose, fucose, merythritol, m-inositol, xylitol, glycerol, mannitol, sorbitol, trehalose, melezitose, melibiose, niacinamide, or amygdalin. Yeast extract is required but not sufficient for growth and cannot be substituted with tryptone, casein hydrolysate, beef extract, or ashed yeast extract (156).

The high yield and rate of ethanol production makes *T. ethanolicus* and the ethanologenic pathways of *Thermoanaerobacter* targets for use in a bioconversion process (158). The bifunctional aldehyde/alcohol dehydrogenase (AdhE) is composed of two conserved domains (aldehyde dehydrogenase, ALDH, and ADH) and is involved in the two step reduction of acetyl-CoA to ethanol. This enzyme has been shown to be critical for ethanologenesis and is

present in many bacteria including *Lactococcus lactis*, *Clostridium acetobutylicum*, *Escherichia coli*, *Clostridium thermocellum*, and *Thermoanaerobacterium saccharolyticum* (159-162). In *T. ethanolicus*, the overexpression of AdhE resulted in a ~50% increase in both ALDH activity and ethanol yield (155). In the closely related *Thermoanaerobacter mathranii*, four ADHs were characterized: an NADP(H) dependent primary alcohol dehydrogenase (AdhA), a secondary alcohol dehydrogenase (AdhB), a butanol dehydrogenase (BdhA), and an NAD(H) dependent bifunctional aldehyde/alcohol dehydrogenase (AdhE) (163). The deletion of AdhE in *T. mathranii* resulted in deficient ALDH activity but maintained ADH activity. This mutation also eliminated ethanol production which indicates while there is some redundancy for the ADH step, the ALDH step is the sole responsibility of AdhE, and AdhE is required for ethanol production (164).

Tables and Figures

Table 1.1 The characteristics of energy carriers for transportation

Specific energy, energy density, and carbon intensity of gasoline, lithium-ion batteries, compressed hydrogen, ethanol and butanol. (38, 39, 45, 52)

Energy Carrier	Specific Energy (Mj/kg)	Energy Density (Mj/L)	Carbon Intensity (gCO_2e/Mj)
Gasoline	48	34.8	90-100
Li-ion battery	0.65	1	Electricity: 48 CA average
Hydrogen (compressed)	142	3.5	75-150
Ethanol	26.8	21.2	40-90
Butanol	36	29.2	N/A

Table 1.2 Relative expression of the ADHs native to *P. furiosus*

Microarray expression data for the four *P. furiosus* ADHs are compared to the highly and constitutively expressed Slp and the cold induced CipA (138). The raw signal intensity is an indicator of absolute expression and is used to calculate the fold change in expression between two conditions.

Microarray data	AdhA	AdhB	AdhC	AdhD	Slp	CipA
ORF	PF0074	PF0075	PF0608	PF0991	PF1399	PF0190
95°C Signal Intensity	1703	827	29461	13564	32451	8541
75°C Signal Intensity	1884	1292	27349	11087	41888	56051
Log_2ratio	0.565	N/A	-0.158	-0.125	0.325	2.882
Fold Change (75°C/95°C)	1.479	N/A	0.896	0.917	1.253	7.373
Absolute Expression	Low	Low	High	Moderate	High	Low→High
Regulated by temp	No	No	No	No	No	Yes

Figure 1.1 Goal of this work

The overall goal of this project is to engineer a carbon fixation cycle and a fuel synthesis pathway into *P. furiosus*. The ideal outcome would be autonomous hydrogen driven carbon fixation and alcohol synthesis occurring in a relatively dormant host.

Figure 1.1

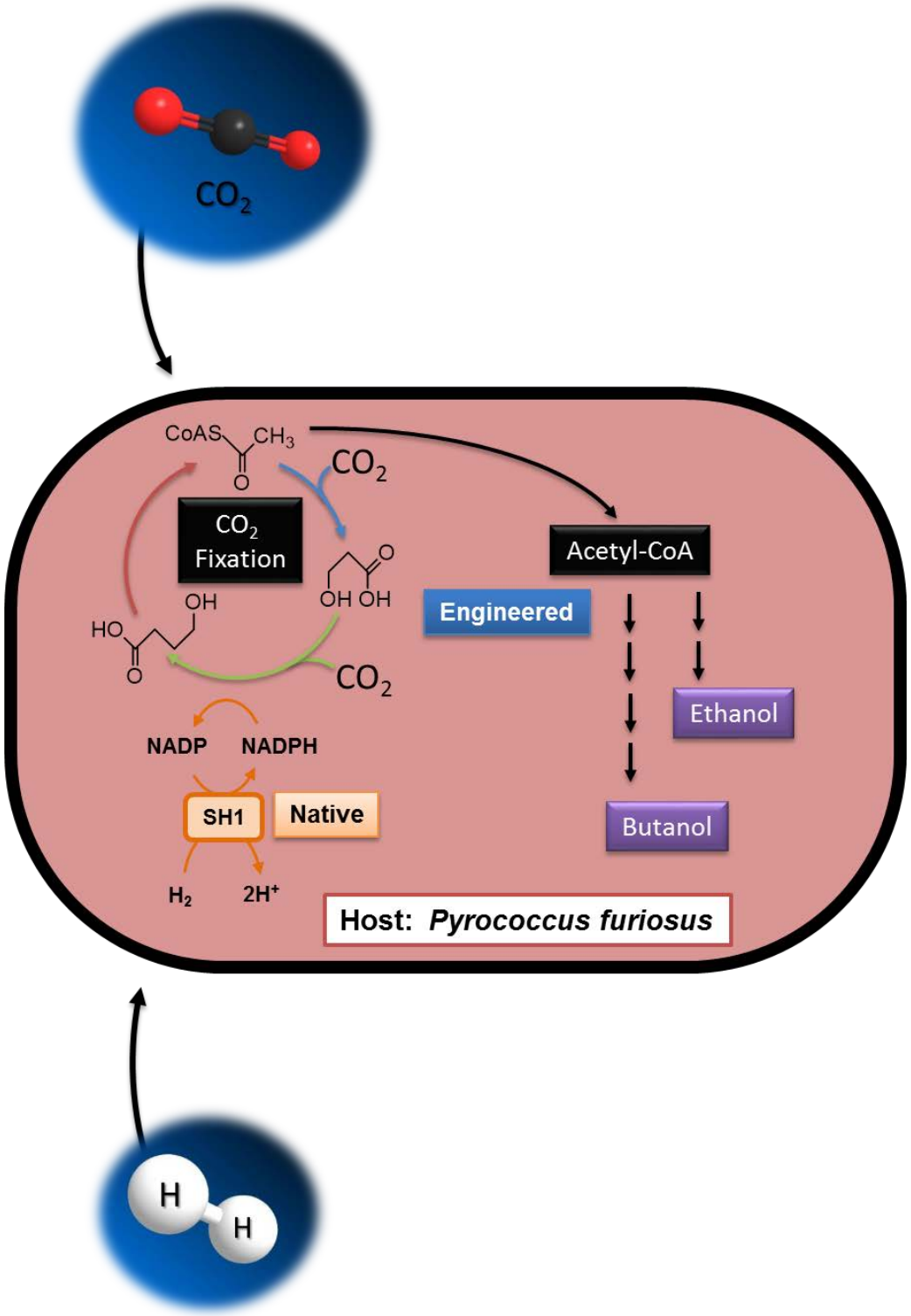
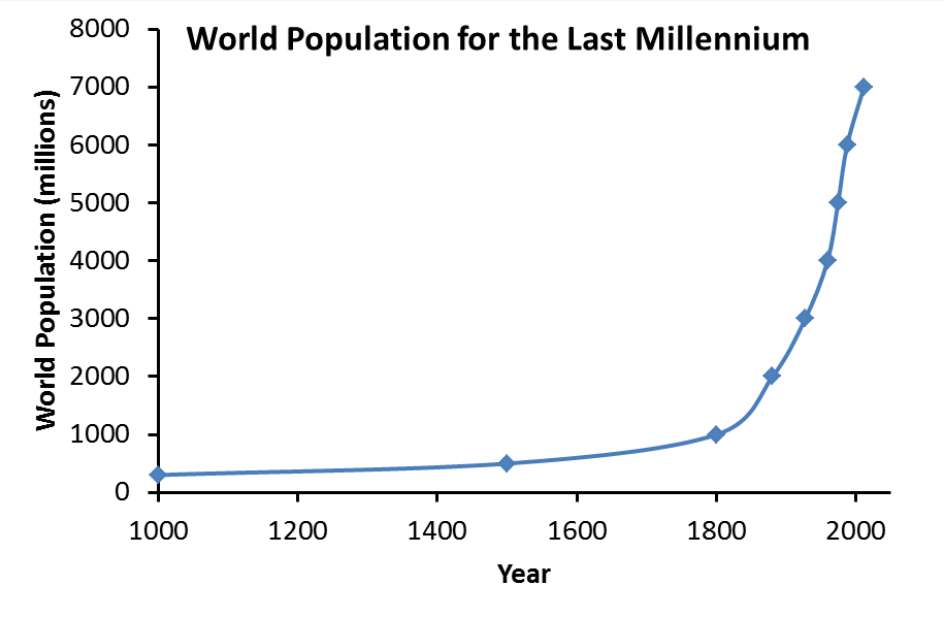

CO_2
CoAS
CH_3
O
CO_2
CO_2 Fixation
OH OH
O
HO
OH
O
CO_2
NADP
NADPH
SH1
Native
H_2
$2H^+$
Acetyl-CoA
Engineered
Ethanol
Butanol
Host: *Pyrococcus furiosus*
H
H

Figure 1.2 World population and atmospheric CO_2 for the past millennium

The world population (A) (3-7) and atmospheric CO_2 (B) (20) for the last 1,000 years. The recent explosion in population has been accompanied by a dramatic increase (40%) in atmospheric CO_2 from 280 to 400 ppm.

Figure 1.2

A


World Population for the Last Millennium
World Population (millions)
8000
7000
6000
5000
4000
3000
2000
1000
0
1000
1200
1400
1600
1800
2000
Year

B

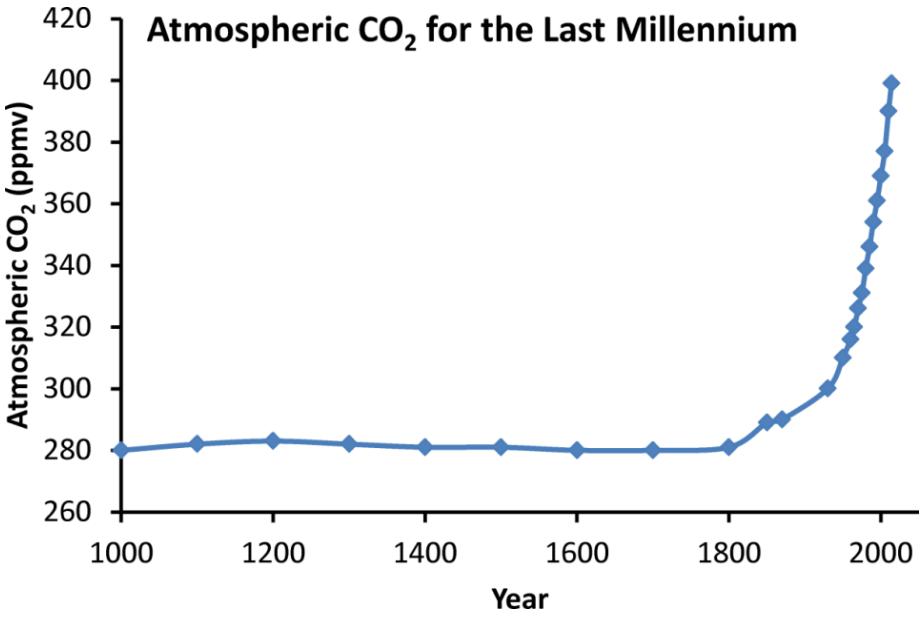

Atmospheric CO_2 for the Last Millennium
Atmospheric CO_2 (ppmv)
420
400
380
360
340
320
300
280
260
1000
1200
1400
1600
1800
2000
Year

Figure 1.3 Peak world oil production

The paucity of crude oil is typically discussed using the term “peak oil” which is the point in time which oil production irreversible decreases and asymptotically approached zero (9). Total oil reserves for the planet are estimated to be between 2 (red) and 4 (green) trillion barrels and production is expected to peak between 2026 and 2047 (10, 11).

Figure 1.3

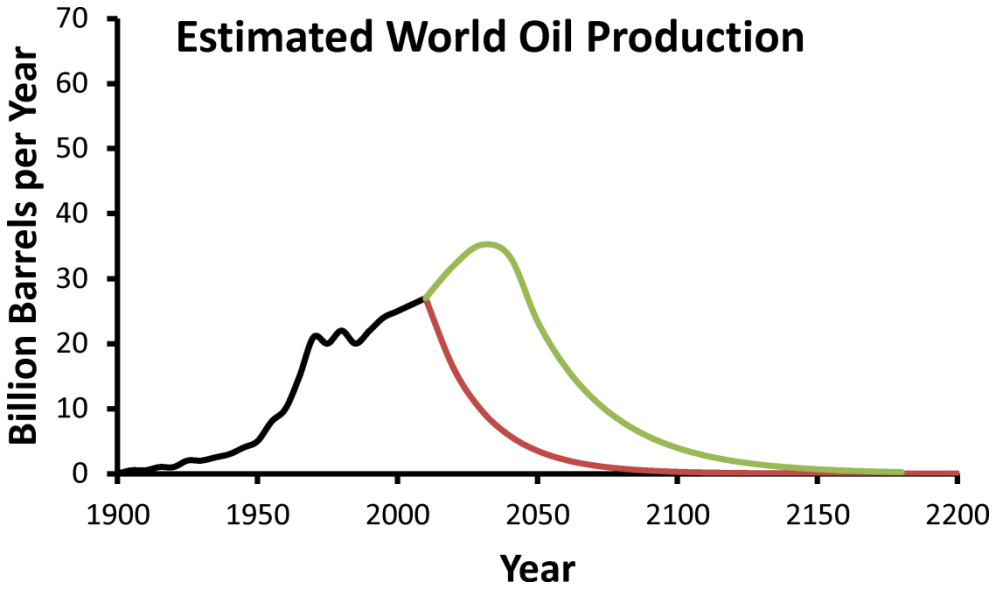

Estimated World Oil Production
Billion Barrels per Year
70
60
50
40
30
20
10
0
1900
1950
2000
2050
2100
2150
2200
Year

Figure 1.4 Cost and emissions of a gasoline versus plug in electric vehicle

Plug in electric vehicles are an economical way to reduce CO_2 emissions compared to conventional gasoline. The Chevy Cruz® is a modern and efficient gasoline powered vehicle that over one mile costs $0.08 to fuel and emits 270 g CO_2, based on California averages. The Tesla Model 3® is the PEV counterpart that over one mile costs $0.05 to fuel and emits 54 g CO_2, based on California averages. (39-41)

Figure 1.4

Fuel cost and CO_2 emissions of gasoline and electric vehicles over one mile

	Chevy Cruz®	Tesla Model 3®
		
Direct monetary cost of gasoline or electricity	$0.08	$0.03
Direct CO_2 emissions	270 grams	54 grams

Figure 1.5 Sugar-based ethanol versus electrofuels

Current strategies for biofuel production, namely corn ethanol, depend on plant based photosynthesis, carbohydrate intermediates, and sugar fermentation. A more efficient strategy, electrofuels, uses renewable sources of reducing power, such as electric current, hydrogen, formate, and carbon monoxide, to drive microbial carbon fixation and direct fuel synthesis. (70-72)

Figure 1.5

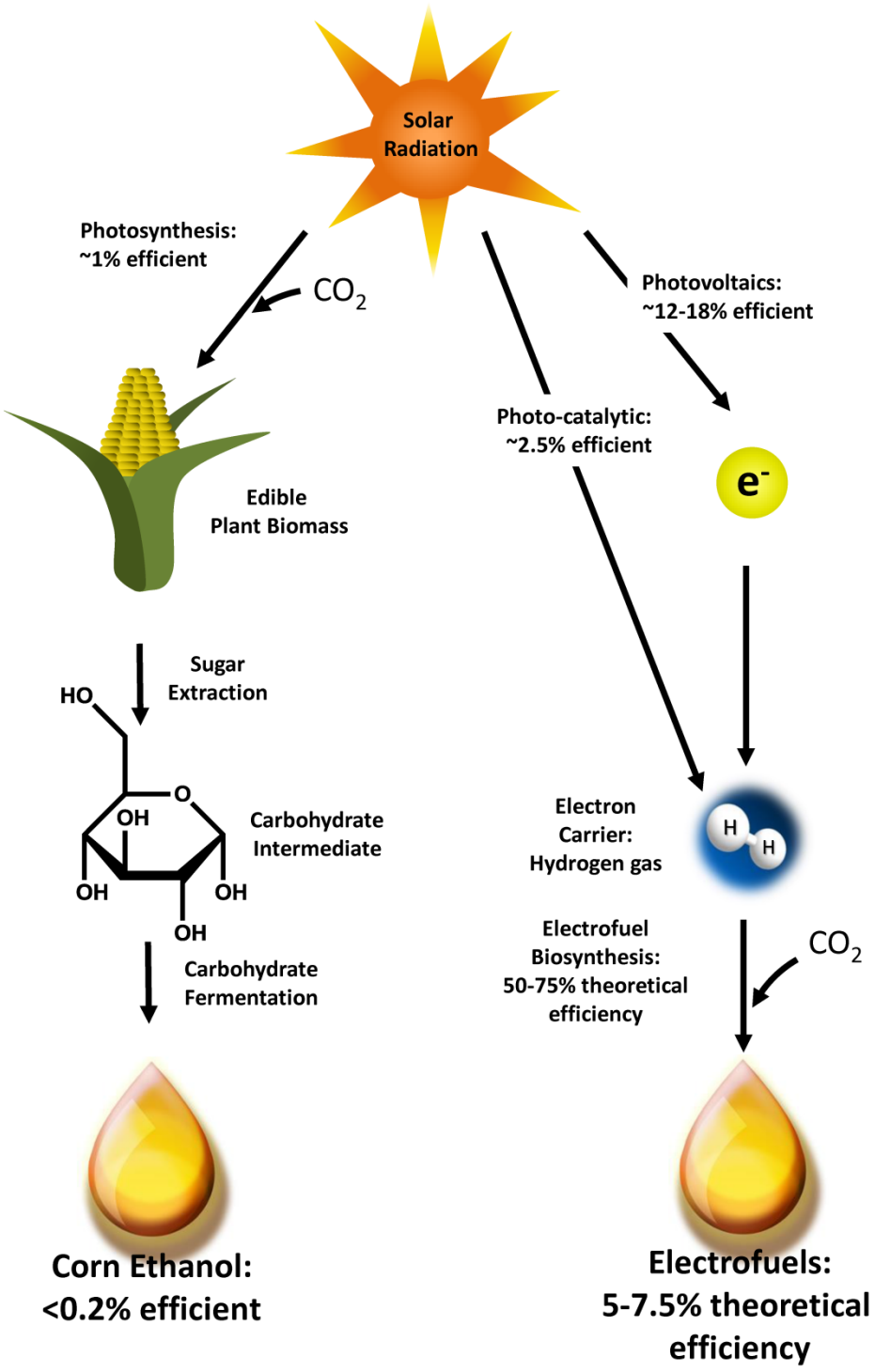

Solar Radiation
Photosynthesis: ~1% efficient
CO_2
Photovoltaics: ~12-18% efficient
Photo-catalytic: ~2.5% efficient
e^-
Edible Plant Biomass
Sugar Extraction
HO
O
OH
OH
OH
OH
Carbohydrate Intermediate
Carbohydrate Fermentation
Electron Carrier: Hydrogen gas
H
H
Electrofuel Biosynthesis: 50-75% theoretical efficiency
CO_2
Corn Ethanol: <0.2% efficient
Electrofuels: 5-7.5% theoretical efficiency

Figure 1.6 The classification of life based on growth temperature

Growth temperature is a canonical characteristic in the classification of microbes and life as a whole. While the exact nomenclature varies, 'psychrophilic' usually refers to an organism with a T_{opt} below 15°C; 'mesophilic' organisms have a T_{opt} between 15 and 45°C; and 'thermophilic' organisms grow optimally above 45°C. However, the temperature range for thermophily is very broad, extending to 122°C, the current upper temperature limit of life (77). Therefore, thermophiles have been further divided into moderate thermophiles (T_{opt} 45-70°C) and extreme thermophiles ($T_{opt} \geq 70$°C, e.g. *C. bescii* and *P. furiosus*); special cases of extreme thermophiles growing with a $T_{opt} \geq 80$°C (e.g. *P. furiosus*) are typically referred to as hyperthermophiles (78, 79).

Figure 1.6

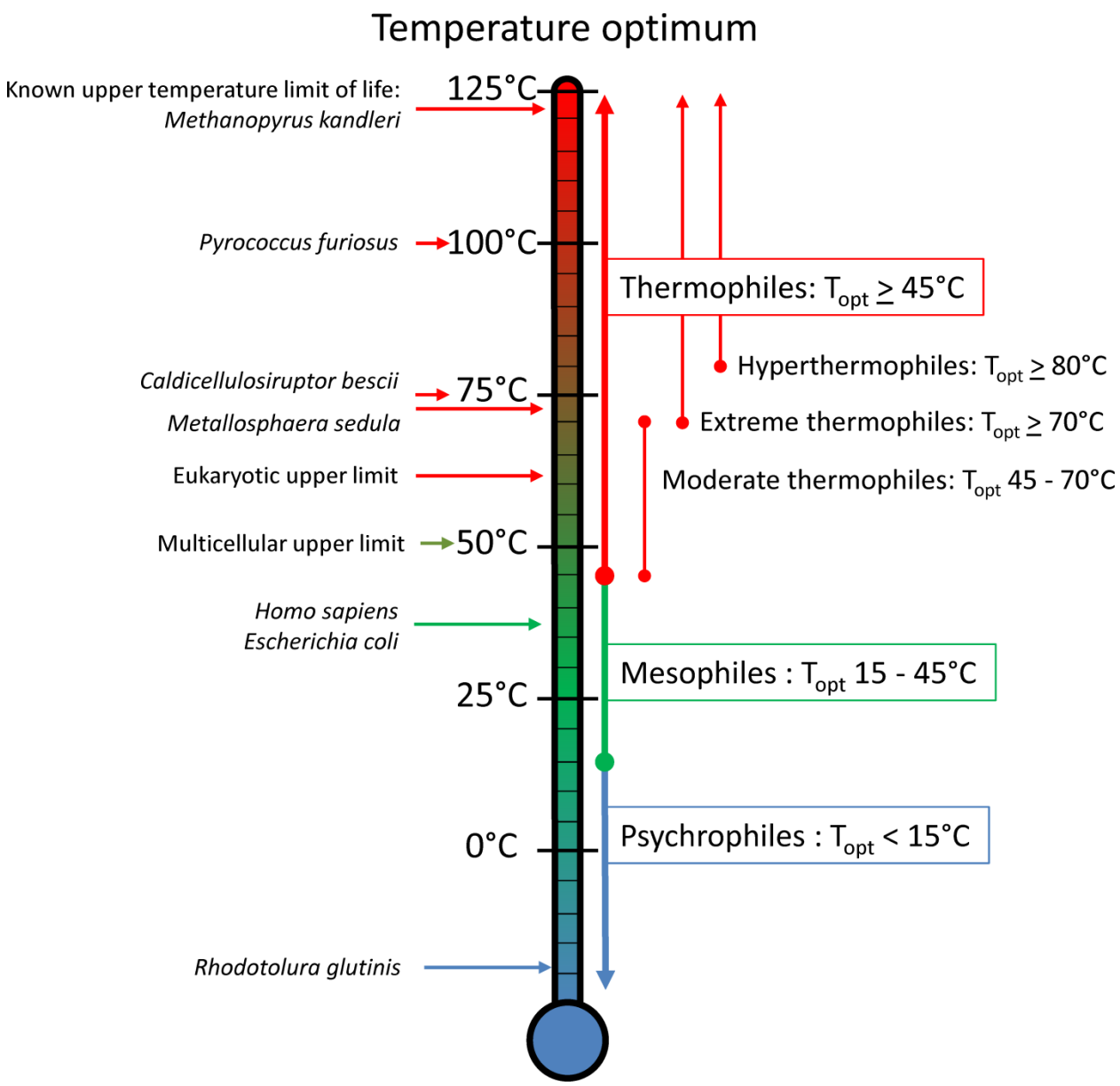

Temperature optimum
Known upper temperature limit of life:
Methanopyrus kandleri
125°C
Pyrococcus furiosus
100°C
Thermophiles: $T_{opt} \geq 45°C$
Hyperthermophiles: $T_{opt} \geq 80°C$
Caldicellulosiruptor bescii
75°C
Metallosphaera sedula
Extreme thermophiles: $T_{opt} \geq 70°C$
Eukaryotic upper limit
Moderate thermophiles: T_{opt} 45 - 70°C
Multicellular upper limit
50°C
Homo sapiens
Escherichia coli
Mesophiles : T_{opt} 15 - 45°C
25°C
0°C
Psychrophiles : $T_{opt} < 15°C$
Rhodotolura glutinis

Figure 1.7 The classification of life based on 16S rRNA sequence

The phylogenetic tree of life based on 16S rRNA sequences is displayed with hyperthermophilic lineages ($T_{opt} \geq 80°C$) highlighted in red. Modified from data presented in (99, 165)

Figure 1.7

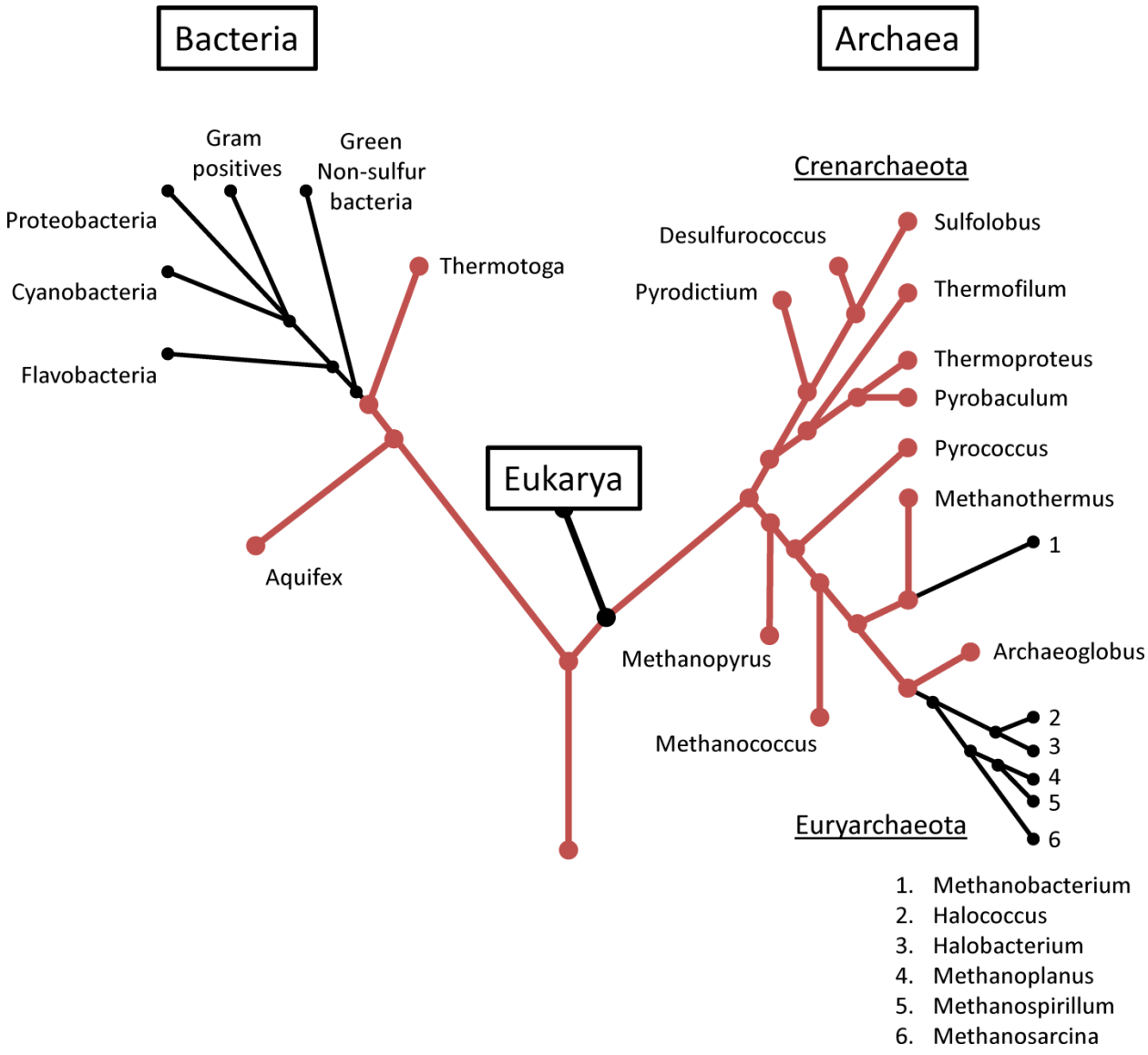

Bacteria
Archaea
Gram positives
Green Non-sulfur bacteria
Proteobacteria
Cyanobacteria
Flavobacteria
Thermotoga
Aquifex
Eukarya
Crenarchaeota
Desulfurococcus
Pyrodictium
Sulfolobus
Thermofilum
Thermoproteus
Pyrobaculum
Pyrococcus
Methanothermus
1
Archaeoglobus
Methanopyrus
Methanococcus
2
3
4
5
6
Euryarchaeota
1. Methanobacterium
2. Halococcus
3. Halobacterium
4. Methanoplanus
5. Methanospirillum
6. Methanosarcina

Figure 1.8 Archaeal versus bacterial membrane lipids

Archaeal membranes (A) are composed of branched isoprenes ether linked to an (*S*)-glycerol. Bacterial and eukaryotic membranes (B) are fatty acids with an ether linkage to a (*R*)-glycerol. Archaeal membranes can also be found to be a mono layer where the isoprene chain extends through the membrane to covalently connect two glycerol phosphate head together. (91, 102)

Figure 1.8

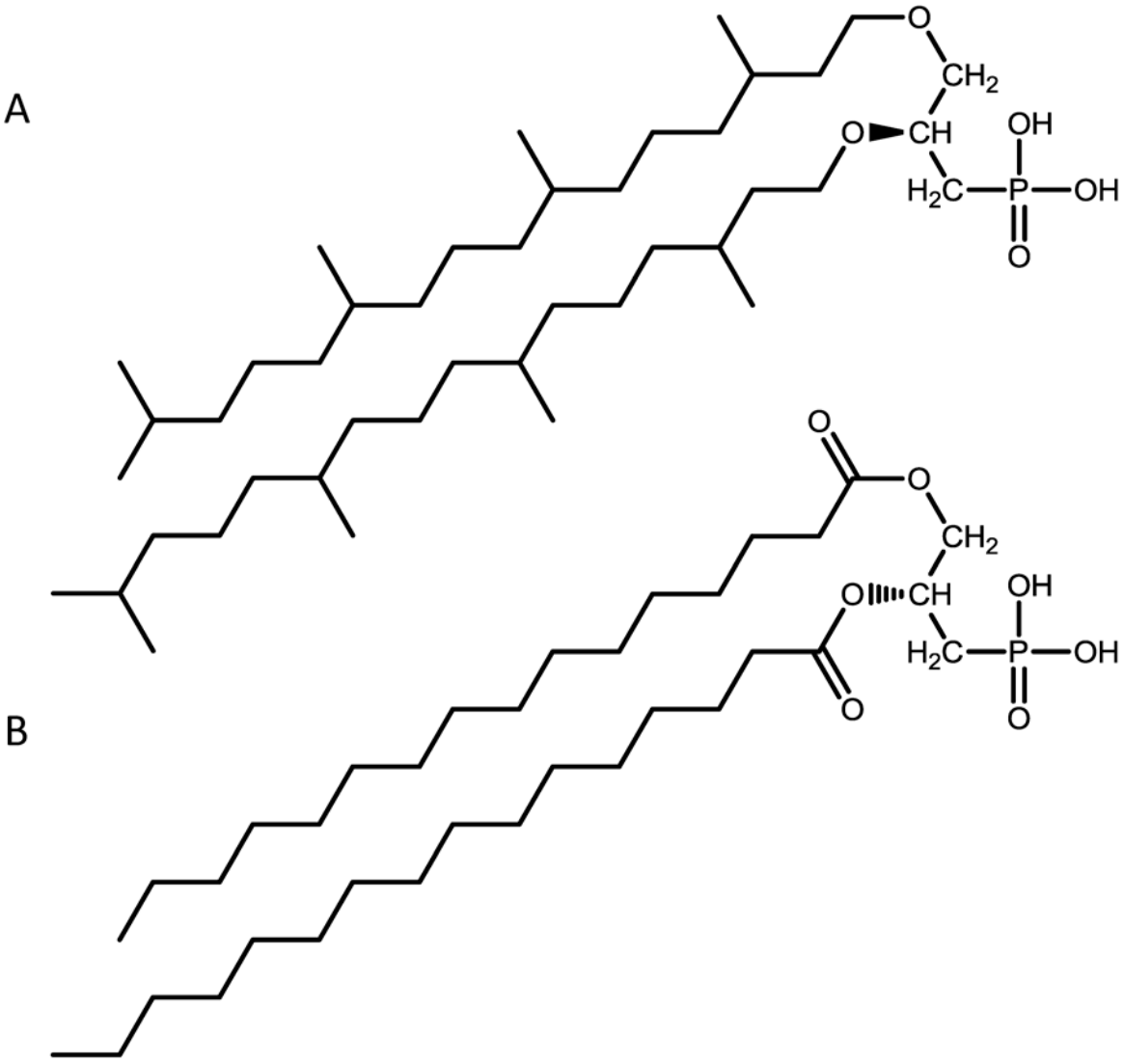


Figure 1.9 Carbohydrate metabolism in *P. furiosus*

The glycolytic pathway in *P. furiosus* converts one molecule of glucose to two molecules of pyruvate, which is subsequently converted to two molecules of acetate, and two molecules of CO_2. Ferredoxin (Fd) Is used to accept the eight electrons from glyceraldehyde-3-phosphate ferredoxin oxidoreductase (GAPOR) and pyruvate ferredoxin oxidoreductase (POR) (114-118). The pathway nets two ATP from substrate level phosphorylation after acetyl-CoA is converted to acetate via acetyl-CoA synthetase (ACS) (119). Membrane bound hydrogenase (MBH) utilizes reduced ferredoxin (Fd_{red}) to generate molecular hydrogen. This energy conserving complex also forms an ion gradient which is coupled to a Na^+ ATP-synthase to yield an additional ~1.2 ATP/glucose (121, 122). While most molecular hydrogen produced is an end product, a portion of it can be recycled by soluble hydrogenase 1 (SH1) to generate NADPH for biosynthesis (123-125).

Figure 1.9

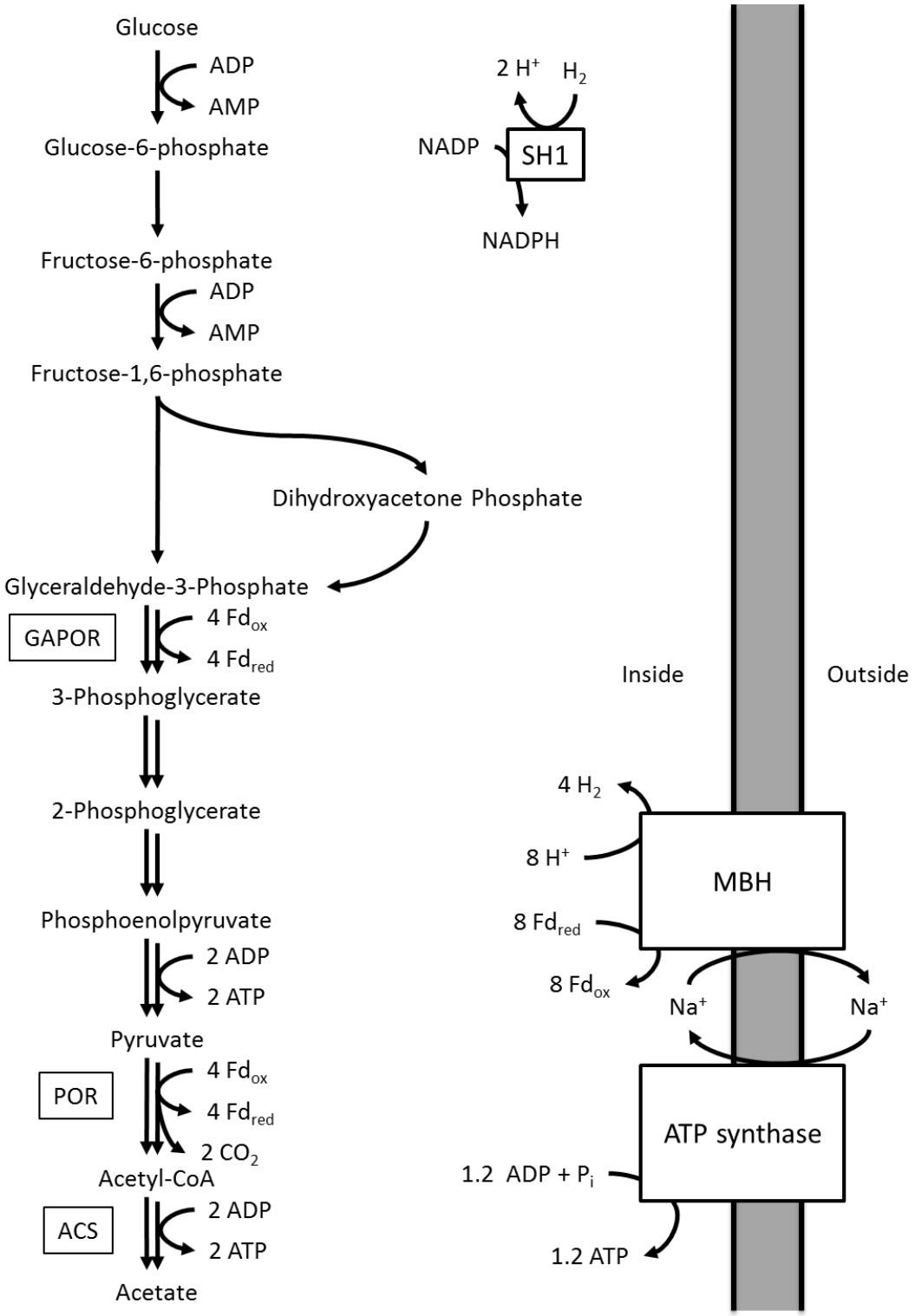

Glucose
ADP
AMP
Glucose-6-phosphate
Fructose-6-phosphate
ADP
AMP
Fructose-1,6-phosphate
Dihydroxyacetone Phosphate
Glyceraldehyde-3-Phosphate
GAPOR
4 Fdox
4 Fdred
3-Phosphoglycerate
2-Phosphoglycerate
Phosphoenolpyruvate
2 ADP
2 ATP
Pyruvate
POR
4 Fdox
4 Fdred
2 CO2
Acetyl-CoA
ACS
2 ADP
2 ATP
Acetate
2 H+ H2
NADP
SH1
NADPH
Inside
Outside
4 H2
8 H+
MBH
8 Fdred
8 Fdox
Na+
Na+
ATP synthase
1.2 ADP + Pi
1.2 ATP

Figure 1.10 The genetic system for *P. furiosus*

The strategy to generate markerless chromosomal deletions or insertions in *Pyrococcus furiosus* (137). A genetically tractable subpopulation is exposure to toxic 5-FOA, resulting in an acceptor strain that lacks *pyrF* function and the ability to produce uracil. A linear construct containing a functional copy of *pyrF* and flanking regions homologous to a region in the chromosome is introduced, and selected for by growth in (-)uracil media. Double homologous recombination occurs to delete or insert the target gene and is selected for via uracil prototrophy. Successful transformants can subsequently be passaged on 5-FOA to select for removal of the *pyrF* marker, recovering an acceptor strain for further manipulations (134, 135). A similar strategy is used in many of the other extreme thermophiles, especially *Thermococcus kodakarensis*. (131, 133)

Figure 1.10

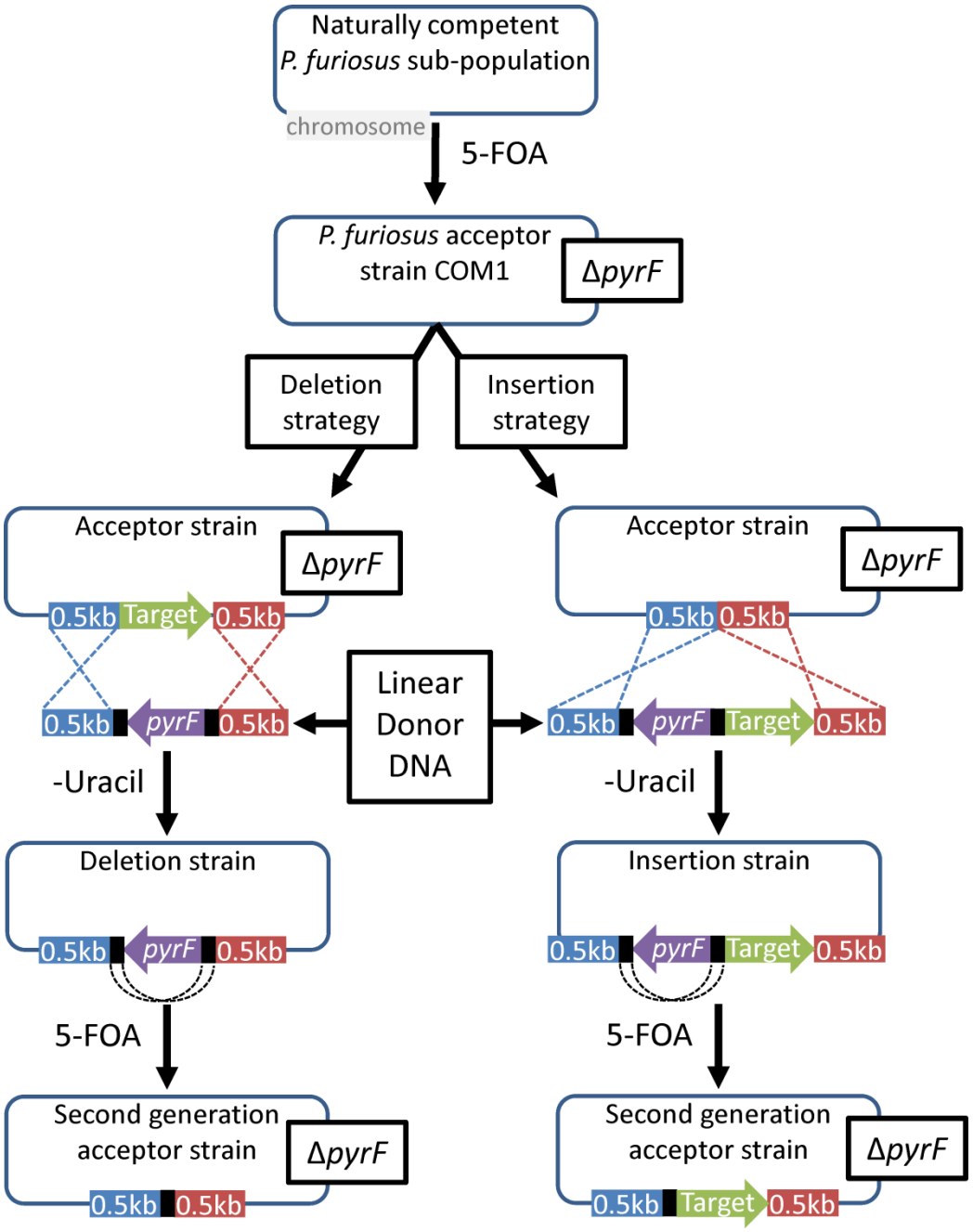

Naturally competent
P. furiosus sub-population
chromosome
5-FOA
P. furiosus acceptor
strain COM1
ΔpyrF
Deletion
strategy
Insertion
strategy
Acceptor strain
ΔpyrF
0.5kb
Target
0.5kb
0.5kb
pyrF
0.5kb
Linear
Donor
DNA
Acceptor strain
ΔpyrF
0.5kb
0.5kb
0.5kb
pyrF
Target
0.5kb
-Uracil
-Uracil
Deletion strain
0.5kb
pyrF
0.5kb
Insertion strain
0.5kb
pyrF
Target
0.5kb
5-FOA
5-FOA
Second generation
acceptor strain
ΔpyrF
0.5kb
0.5kb
Second generation
acceptor strain
ΔpyrF
0.5kb
Target
0.5kb

Figure 1.11 Growth of *P. furiosus* at optimal and sub-optimal temperatures

The effect of temperature (95°C in red and 72°C in blue) on *P. furiosus* growth. (A) At 95°C short lag phase precedes exponential growth with a doubling time as low as 37 minutes. (B) Interrupting growth at 95°C with a sudden shift to 72°C results in a second longer lag phase followed by exponential growth with a doubling time of approximately 5 hours. (C) Growth at 72°C begins with a long lag phase followed by exponential growth with a doubling time of approximately 5 hours. Modified from data presented in (138).

Figure 1.11

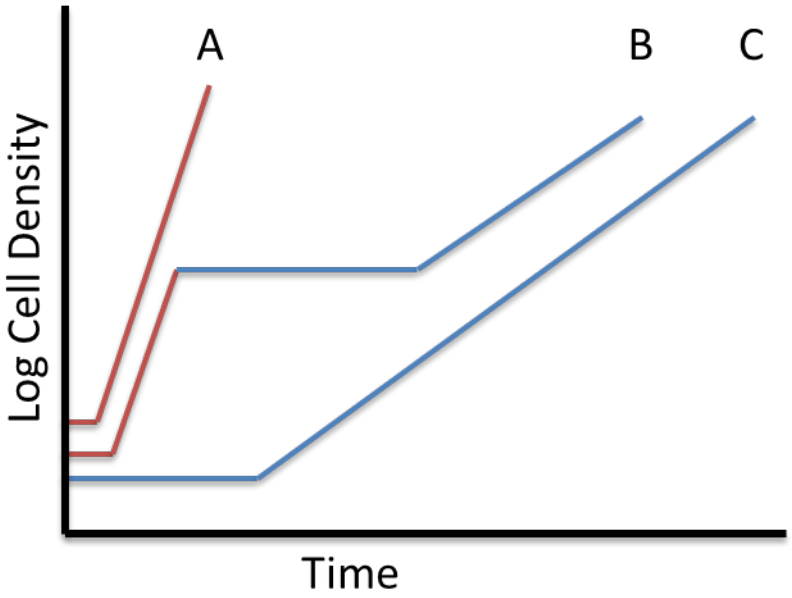


Figure 1.12 Enzymes of *P. furiosus* that have been characterized as suboptimal temperatures

The effect of suboptimal temperatures on the relative activity of ACSI (blue diamonds), ACSII (red diamonds), POR (green triangles), AOR (purple squares), and SH1 (orange circles). AOR activity was determined spectrophotometrically by the aldehyde-dependent reduction of methyl viologen. ACS I and ACS II activities were determined by measuring the rate of phosphate formed from acetate, ATP, and CoASH for ASI and from indoleacetate, GTP, and CoASH for ACSII. POR activity was determined by measuring the rate of acetyl-CoA formation by coupling with a malate dehydrogenase/citrate synthase assay. SH1 activity was determined spectrophotometrically by the hydrogen-dependent reduction of NADP. Modified from data presented in (117, 119, 139, 140)

Figure 1.12

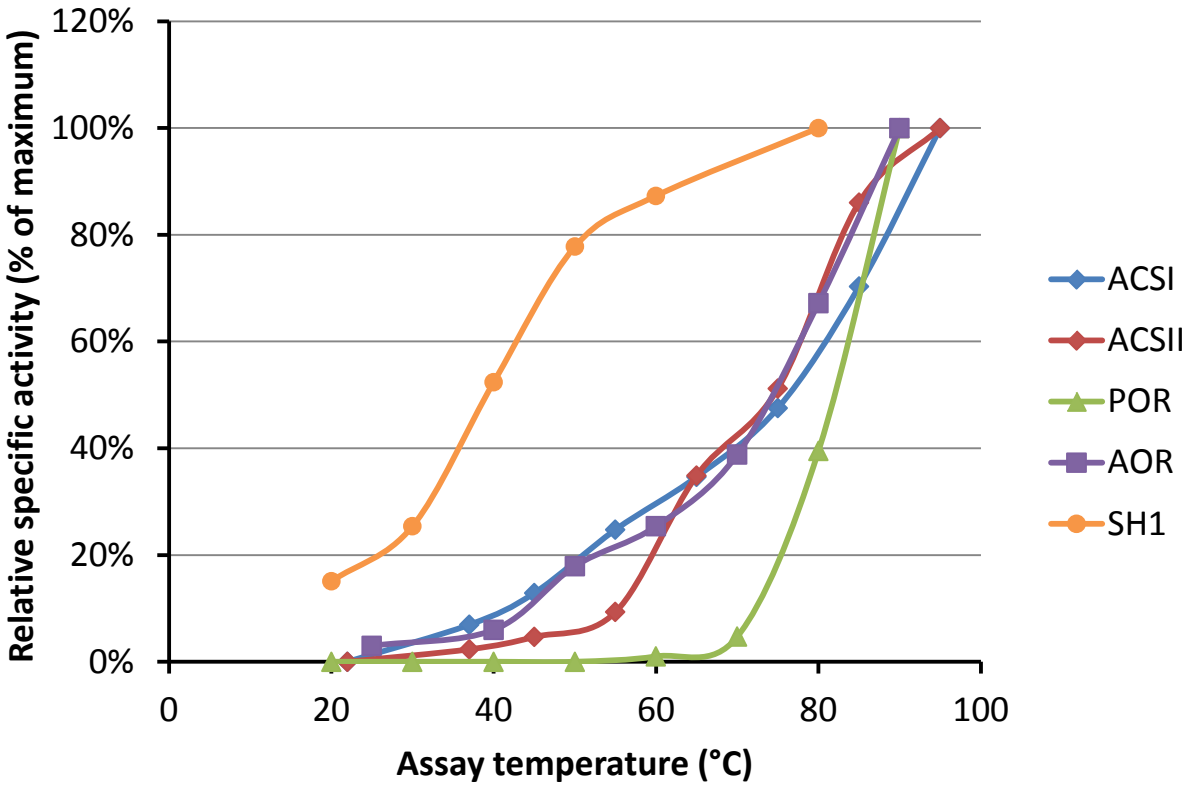

120%
100%
80%
60%
40%
20%
0%
Relative specific activity (% of maximum)
0
20
40
60
80
100
Assay temperature (°C)
ACSI
ACSII
POR
AOR
SH1

Figure 1.13 Carbon dioxide assimilation in *Metallosphaera sedula*

A

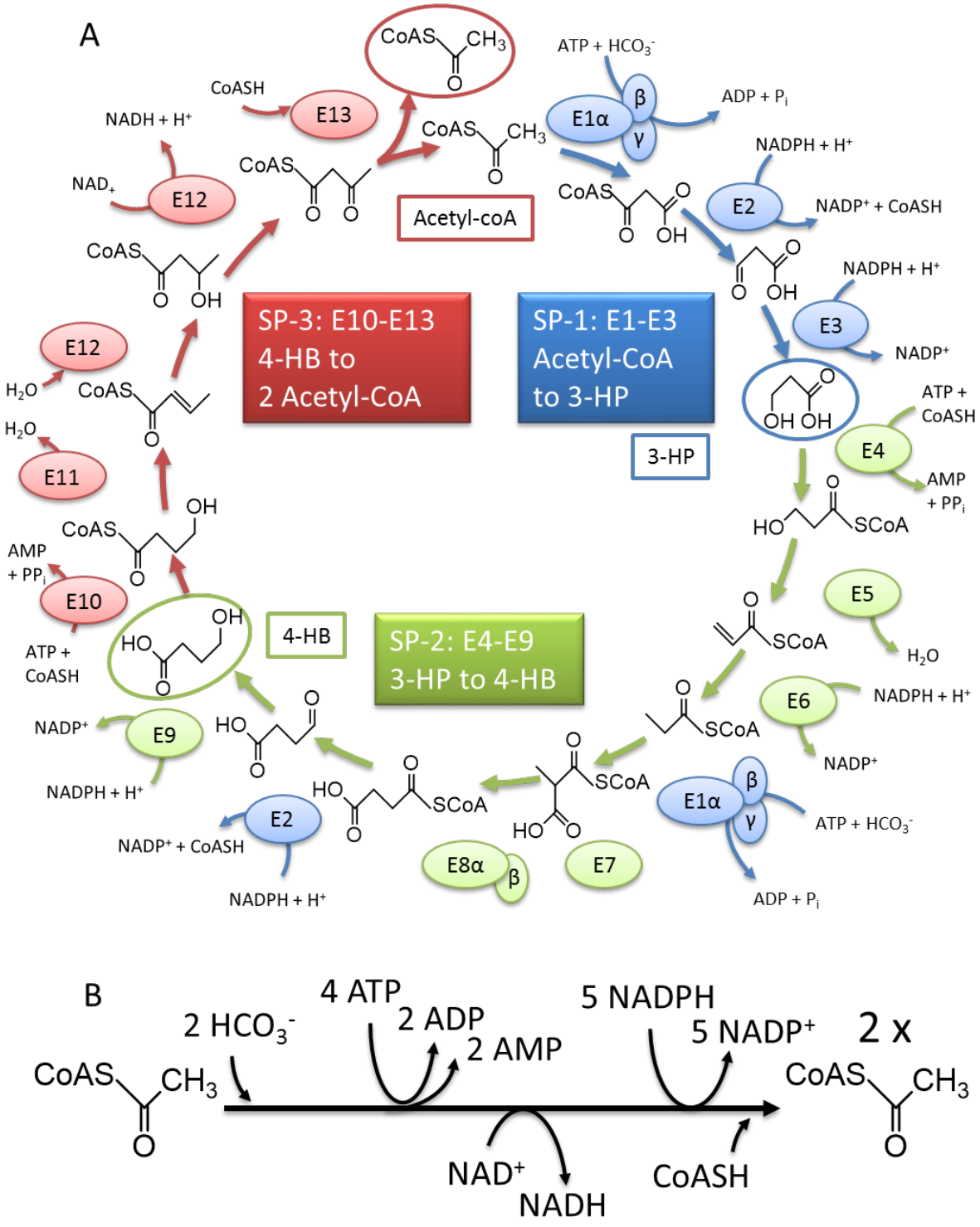
The 3-Hydroxypropionate/4-Hydroxybutyrate carbon fixation cycle in Metallosphaera sedula. E1αβγ Msed_0147, 1048, 1375: Acetyl/propionyl-CoA carboxylase, E2 Msed_0709: Malonyl/succinyl-CoA reductase, E3 Msed_1993: Malonate semialdehyde reductase, E4 Msed_1456: 3-Hydroxypropionyl-CoA synthase, E5 Msed_2001: 3-Hydroxypropionyl dehydratase, E6 Msed_1426: Acryloyl-CoA reductase, E7 Msed_0639: Methylmalonyl-CoA epimerase, E8αβ Msed_0638, 2055: Methylmalonyl-CoA mutase, E9 Msed_1424: Succinate semialdehyde reductase, E10 Msed_0406 or 0394: 4-Hydroxybutyrate-CoA ligase, E11 Msed_1321: 4-Hydroxybutyrl-CoA dehydratase, E12 Msed_0399: Crotonyl-CoA hydratase/ (S)-3-Hydroxybutyrl-CoA dehydrogenase, and E13 Msed_0656: Acetoacetyl-CoA β-ketothiolase.

B

The stoichiometry of the cycle. The net production of one molecule of Acetyl-CoA by this cycle requires carbon from two molecules of HCO_3^-, the energy equivalent of hydrolysis of six ATP to ADP, five reducing equivalents in the form of NADPH (one regenerated in the form of NADH), and one molecule of Coenzyme A.

(148-154)

Figure 1.13


A
CoAS CH3
O
CoASH
E13
NADH + H+
NAD+
E12
CoAS
O O
Acetyl-coA
ATP + HCO3-
E1α
β
γ
ADP + Pi
NADPH + H+
CoAS
O OH
E2
NADP+ + CoASH
SP-3: E10-E13
4-HB to
2 Acetyl-CoA
SP-1: E1-E3
Acetyl-CoA
to 3-HP
3-HP
E3
NADP+
OH OH
E4
ATP +
CoASH
AMP
+ PPi
HO SCoA
E5
H2O
SCoA
E6
NADPH + H+
NADP+
E1α
β
γ
ATP + HCO3-
ADP + Pi
E7
E8α
β
SCoA
HO
O
E2
NADP+ + CoASH
NADPH + H+
HO
O
E9
NADP+
NADPH + H+
4-HB
SP-2: E4-E9
3-HP to 4-HB
HO
OH
O
E10
AMP
+ PPi
ATP +
CoASH
CoAS
OH
O
E11
H2O
E12
H2O
CoAS
O
B
2 HCO3-
4 ATP
2 ADP
2 AMP
5 NADPH
5 NADP+
2 x
CoAS CH3
O
NAD+
NADH
CoASH
CoAS CH3
O

CHAPTER 2

EXPLOITING MICROBIAL HYPERTHERMOPHILICITY TO PRODUCE AN INDUSTRIAL CHEMICAL, USING HYDROGEN AND CARBON DIOXIDE

Keller MW†, Schut GJ†, Lipscomb GL, Menon AL, Iwuchukwu IJ, Leuko TT, Thorgersen MP, Nixon WJ, Hawkins AS, Kelly RM, Adams MWW (2013) Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. *Proc Natl Acad Sci* 110(15):5840-5.

†Authors contributed equally to this work

Reprinted here with permission of the publisher PNAS

Abstract

Microorganisms can be engineered to produce useful products, including chemicals and fuels from sugars derived from renewable feedstocks, such as plant biomass. An alternative method is to use low potential reducing power from nonbiomass sources, such as hydrogen gas or electricity, to reduce carbon dioxide directly into products. This approach circumvents the overall low efficiency of photosynthesis and the production of sugar intermediates. Although significant advances have been made in manipulating microorganisms to produce useful products from organic substrates, engineering them to use carbon dioxide and hydrogen gas has not been reported. Herein, we describe a unique temperature-dependent approach that confers on a microorganism (the archaeon *Pyrococcus furiosus*, which grows optimally on carbohydrates at 100°C) the capacity to use carbon dioxide, a reaction that it does not accomplish naturally. This was achieved by the heterologous expression of five genes of the carbon fixation cycle of the archaeon *Metallosphaera sedula*, which grows autotrophically at 73°C. The engineered *P. furiosus* strain is able to use hydrogen gas and incorporate carbon dioxide into 3-hydroxypropionic acid, one of the top 12 industrial chemical building blocks. The reaction can be accomplished by cell-free extracts and by whole cells of the recombinant *P. furiosus* strain. Moreover, it is carried out some 30°C below the optimal growth temperature of the organism in conditions that support only minimal growth but maintain sufficient metabolic activity to sustain the production of 3-hydroxypropionate. The approach described here can be expanded to produce important organic chemicals, all through biological activation of carbon dioxide.

Introduction

Metabolically-engineered microorganisms can be used to produce a variety of products, ranging from bulk chemicals and fuels to complex pharmaceutical molecules. The largest effort in biofuel production is currently based on renewable plant biomass (166-168). First generation biofuels include ethanol from corn fermentation and fatty acid methyl esters from oils and fats, while second generation biofuels utilize cellulosic biomass as feedstocks and can generate higher alcohols (59). An alternative method for the microbial production of both fuels and chemicals is to utilize low potential reducing power from sources, such as hydrogen gas, reduced metals, or electricity, to reduce carbon dioxide directly to useful products. This circumvents the overall low efficiency in generating both plant and algal photosynthetic products (169). Moreover, such electron sources can potentially be used to reduce carbon dioxide directly to produce liquid fuels or 'electrofuels' (71), or to produce industrial chemicals without a sugar intermediate. However, while significant advances have been made in manipulating microorganisms to produce various fuels from organic substrates (59, 170, 171), the engineering of microorganisms to utilize carbon dioxide and hydrogen gas has not been reported.

Herein, we have developed a novel temperature-dependent approach (172) to confer upon a microorganism that cannot naturally utilize carbon dioxide but grows on sugars optimally at 100°C, the capacity to utilize carbon dioxide near 70°C. Hydrogen gas is used as the reductant to incorporate the carbon of carbon dioxide to produce 3-hydroxypropionic acid (3-HP), which is one of the top twelve industrial chemical building blocks used in the production of

acrylic acid, acrylamide and 1,3-propanediol (173, 174). Furthermore, the metabolic burden of the engineered microorganism during chemical production from hydrogen and carbon dioxide is minimized by the strategic operation at temperatures that are suboptimal for its growth.

The hyperthermophilic archaeon *Pyrococcus furiosus* is an obligate heterotroph that grows optimally (T_{opt}) at 100°C by fermenting sugars to hydrogen, carbon dioxide and acetate (102). It cannot utilize carbon dioxide as its sole carbon source. A genetic system is available for *P. furiosus* based on a competent strain with a known sequence (175) that has allowed both homologous (176, 177) and heterologous over-expression of genes (172). A novel means of metabolic control was recently reported in *P. furiosus* that exploited the difference in the temperature dependence of the host's metabolism and the inserted foreign synthetic pathway (172). For example, expression in *P. furiosus* of the gene encoding lactate dehydrogenase from a moderately thermophilic bacterium (*Caldicellulosiruptor bescii,* T_{opt} of 78 °C) resulted in temperature-dependent lactate formation (172). Moreover, the engineered pathway was active near 70°C, conditions under which the host metabolism of *P. furiosus* is minimal, as it is nearly 30°C below its optimal temperature. Hence, the host will require minimal maintenance energy and, as a result, incur minimal metabolic burden, while the engineered pathway that it contains is optimally active. Herein, we have utilized this temperature-dependent strategy for optimal bioproduct generation by expressing in *P. furiosus* genes encoding carbon dioxide fixation and 3-HP synthesis from the thermoacidophilic archaeon *Metallosphaera sedula* (T_{opt} 73°C: (178)).

Results and Discussion

The genes that were incorporated into *P. furiosus* to enable it to utilize carbon dioxide are the first part of the 3-hydroxypropionate/4-hydroxybutyrate pathway of *M. sedula,* which consists of 13 enzymes in total (151). In one turn of the cycle, two molecules of carbon dioxide are added to one molecule of acetyl-CoA (C_2) to generate a second molecule of acetyl-CoA (**Figure 2.1c**). The cycle can be divided into three sub-pathways (SP1-SP3) where SP1 generates 3-hydroxypropionate (3-HP) from acetyl-CoA and carbon dioxide, SP2 generates 4-hydroxybutyrate (4-HB) from 3-HP and carbon dioxide, and SP3 converts 4-HB to two molecules of acetyl-CoA. The reducing equivalents and energy for the pathway are supplied by NADPH and ATP, respectively (**Figure 2.1d**). Notably, the 3-HP/4-HB pathway is purportedly more energetically efficient than carbon dioxide fixation by the ubiquitous Calvin cycle (141).

The first three enzymes of the Msed 3-HP/4-HB cycle comprise the SP1 pathway and together they produce 3-HP from carbon dioxide and acetyl-CoA (**Figure 2.1b**). The three enzymes are referred to here as: E1 (acetyl/propionyl-CoA carboxylase, encoded by Msed_0147, Msed_0148, Msed_1375), E2 (malonyl/succinyl-CoA reductase, Msed_0709) and E3 (malonate semialdehyde reductase, Msed_1993) (141, 145, 147). E1 carboxylates acetyl-CoA using bicarbonate and requires ATP. E2 breaks the CoA-thioester bond and with E3, reduces the carboxylate to an alcohol with NADPH as the electron donor. E1 and E2 are bi-functional and are also involved in the SP2 part of the cycle (**Figure 2.1c**). To demonstrate the concept, we expressed the *M. sedula* SP1 pathway in *P. furiosus* so that the organism could utilize carbon dioxide for the production of 3-HP, using hydrogen as the electron donor. Hydrogen is utilized

in *P. furiosus* by a native cytoplasmic hydrogenase (SHI) that reduces NADP to NADPH (179). SHI is extremely active, even at 70°C, and a *P. furiosus* strain engineered to over-express the enzyme was previously developed (176).

The five genes encoding the three enzymes (E1αβγ, E2, E3) of *M. sedula* SP1 were combined into a single synthetic operon with transcription driven by P_{slp}, a native, constitutive promoter of the highly expressed S-layer protein (PF1399) of *P. furiosus* (176). The *M. sedula* ribosomal binding sites (RBS) for E1(γ), E2 and E3 were replaced with RBSs for known highly-expressed *P. furiosus* proteins (**Figure 2.1a**). The *M. sedula* RBS for E1β was retained since the two genes, E1α and E1β, appear to be translationally-coupled. The SP1 operon was inserted into *P. furiosus* (strain COM1) at two genome regions. In *P. furiosus* strain PF506, the SP1 operon was inserted at the site of the *pdaD* marker (PF1623; **Figure 2.2**). The MW56 strain contained the SP1 operon between convergently-transcribed genes (PF0574 and PF0575: **Figures 2.3 and 2.4**) within a ~100-bp region having little to no transcriptional activity, according to a previous tiling array study of *P. furiosus* (180). The *P. furiosus* strains used here are summarized in **Table 2.1**.

The premise for the temperature-dependent strategy is that *P. furiosus* (T_{opt} 100°C) shows little growth and has very low metabolic activity (138) near the temperature at which the enzymes from *M. sedula* (T_{opt} 73°C) are expected to be optimally active. In the recombinant *P. furiosus* strains (PF506 and MW56), the SP1 operon was under the control of a temperature-independent, constitutive promoter (P_{slp}), hence the operon will be transcribed at both 100°C and 75°C. However, the resulting E1-E3 enzymes should be stable and active only near 75°C. *P.*

furiosus strains PF506 and MW56 were, therefore, grown at 98°C (to ~ 1 x 10^8 cells/ml) in closed static cultures and then transferred to 75°C (**Figure 2.5a**). There was no measurable activity of E1, E2 or E3 in cell-free extracts prior to the temperature change, but all three activities were present in cells after 16 hr at 75°C. Moreover, the specific activities were comparable to those measured in extracts of *M. sedula* cells grown autotrophically on hydrogen and carbon dioxide and to values reported by others (**Figures 2.5c and 2.6**) (151, 178). Indeed, when grown in a stirred, pH-controlled culture, the activity of the linked E2 + E3 enzymes in strain MW56 continued to increase over a 50 hr period, reaching over 8-fold greater than that measured in *M. sedula* (**Figure 2.7c**). When strain PF506 was grown at 95°C and then incubated for 16 hours at temperatures between 55° and 95°C, the maximum specific activity of the linked E2 + E3 enzymes was measured in cultures incubated at 70 and 75°C, with dramatically lower values at 65 and 80°C (**Figure 2.5b**). This clearly indicates that the *M. sedula* enzymes functioned optimally in *P. furiosus* at 70-75°C, especially since significant E2 + E3 activity could be measured at assay temperatures above 75°C using cell-free extracts prepared from cultures incubated at 70-75°C (**Figure 2.5d**). Moreover, the enzymes are very thermostable, with a half-life of approximately 60 min at 90°C (**Figure 2.8**). This suggests that the lack of enzyme activity of the *M. sedula* enzymes (and of 3-HP production) in cultures that were incubated at 80°C or higher is not due to the thermal instability of the *M. sedula* enzymes per se, but rather to the temperature sensitivity of the protein folding process during the synthesis of these enzymes, which is optimal in the 70-75°C range.

To determine the nature of the products of the SP1 pathway, recombinant *P. furious* strains PF506 and MW56 were grown at 95°C (to ~1 x 10^8 cells/ml) and then transferred to 70°C for 16 hours (**Figure 2.9**). In extracts of these cells, the specific activities of the E1, E2, and E3 enzymes were comparable to those measured in extracts of autotrophically-grown *M. sedula* cells (**Figure 2.6**). Two methods were used to detect 3-HP and to confirm its production by the SP1 pathway in the recombinant *P. furiosus* strains. In the presence of acetyl-CoA, $NaHCO_3$, and either NADPH or hydrogen gas as the electron donor, the 2-nitrophenylhydrazide-derivative (3-HP/HZ; m/z 224) was identified by electrospray ionization mass spectrometry (ESI-MS) in cell-free extracts of PF506, but was not detected in extracts of the parent *P. furiosus* strain (**Figure 2.10**). This was confirmed by gas chromatography-mass spectrometry (GC-MS) of the O-trimethylsilylate derivative of 3-HP (3HP/TMS), using malonyl-CoA and either NADPH or hydrogen gas as the electron donor (**Table 2.2**). The GC-MS also allowed quantitation of 3-HP/TMS and showed that approximately 150 µM 3-HP was produced from malonyl-CoA, after a 2 hr incubation at 72°C with extracts of PF506 containing NADP under hydrogen gas (**Table 2.2**).

For routine analysis of 3-HP, a method was developed to extract 3-HP/HZ and to separate and quantitate it by HPLC. As shown in Fig. 3A, this method was used to confirm 3-HP production from acetyl-CoA and carbon dioxide by the combined action of the enzymes E1, E2, and E3 in cell-free extracts. As expected, *P. furiosus* did not appear to further metabolize 3-HP, as the compound was stable when added to *P. furiosus* cultures. Moreover, the production of 3-HP from acetyl-CoA was dependent upon either $NaHCO_3$ or CO_2 as the C-1 carbon source and either NADPH or hydrogen gas (and NADP) as the electron donor (**Figure 2.7a**). The

incorporation of electrons from hydrogen gas and the carbon from carbon dioxide into a single desired product is essentially the paradigm for 'electrofuels' (71).

P. furiosus grows by fermenting sugars (such as the disaccharide maltose) to acetate, carbon dioxide and hydrogen, and can also utilize pyruvate as a carbon source (102). Acetyl-CoA and carbon dioxide are generated as the product of the pyruvate ferredoxin oxidoreductase (POR) reaction (**Figure 2.11**). The reduced ferredoxin is oxidized by a membrane-bound hydrogenase to generate hydrogen gas (122). Although growth is limited at 75°C (138), it was expected that when whole cells were incubated at 75°C with maltose or pyruvate, sufficient acetyl-CoA would be produced by the low metabolic activity of *P. furiosus* for the SP1 enzymes to produce 3-HP. This was confirmed by HPLC detection and quantitation of 3-HP as the 2-nitrophenylhydrazide derivative. For example, high cell density suspensions ($\geq 10^{10}$ cells/ml) of *P. furiosus* strains MW56 and PF506 produced up to 0.2 mM 3-HP after one hour incubation at 75°C in the presence of maltose, hydrogen gas, and $NaHCO_3$ (**Figures 2.12 and 2.13**), and 3-HP production was dependent upon the presence of maltose or pyruvate (**Table 2.3**). Moreover, recombinant *P. furiosus* strains PF506 and MW56, grown in static cultures to late-log phase (~1 x 10^8 cells/ml) at 98°C on maltose, produced up to 0.6 mM 3-HP (60 mg/l) when subsequently incubated at 72°C for up to 40 hours (**Figure 2.7B**). Furthermore, in a stirred, pH-controlled culture, strain MW56 produced 3-HP continuously over a 50 hr period at 72°C (**Figure 2.7C**). Overall, there appeared to be no significant difference between the two recombinant *P. furiosus* strains in terms of 3-HP production. This indicated that the genome region of the synthetic operon derived from *M. sedula* was not a determining factor. This bodes well for the

insertion of additional synthetic operons in *P. furiosus* to extend the results reported here to other industrial chemicals.

In summary, this work demonstrates the use of hydrogen as the electron donor for carbon dioxide fixation into a product of great utility in the chemical industry, namely 3-HP. Moreover, it is carried out by an engineered heterotrophic hyperthermophile some 30°C below the optimal growth temperature of the organism, conditions that support minimal growth, but sufficient metabolic activity is retained to sustain the production of 3-HP (71). The reaction can be accomplished by cell-free extracts, and also by whole cells in culture using sugar (maltose) as the source of the acetyl-CoA and ATP in a hydrogen- and carbon dioxide-dependent manner. The feasibility of using hydrogen gas as the source of reducing power (NADPH) for chemical synthesis, in this case 3-HP, is also of high significance given the availability of relatively inexpensive natural gas as a hydrogen source (181). It is important to note that the low metabolic activity of *P. furiosus* at 72°C was sufficient to provide the ATP needed for carbon dioxide fixation. These results are a significant step forward towards the overall goal of incorporating into *P. furiosus* the complete *M. sedula* 3-HP/4-HB pathway, in which two molecules of carbon dioxide are reduced to acetyl-CoA that can then be converted into a variety of valuable products including biofuels (71). Clearly, there will be a balance between using a fixed carbon source (sugar) via the low metabolic activity of the host to produce ATP and the high catalytic activity of the heterologous enzymes to generate the desired product. The hydrogen-dependent fixation of carbon dioxide has enormous potential for the production of a variety of chemicals and fuels through strategic use of established biosynthetic pathways

and exploiting the hyperthermophilicity of metabolically-engineered microbial hosts (59, 168, 171, 181).

Materials and Methods

Construction of a synthetic SP1 operon

PCR was performed using *P. furiosus* or *M. sedula* genomic DNA to generate the individual PCR products of the *P. furiosus* S-layer promoter (P_{slp}) and the five *M. sedula* SP1 genes, consisting of coupled E1αβ (Msed_0147-Msed_0148), E1γ (Msed_1375), E2 (Msed_0709) and E3 (Msed_1993). *P. furiosus* ribosomal binding sites, consisting of 11-14 bp of sequence upstream of highly-expressed proteins, were added in front of E1γ (5'-GGAGGTTTGAAG, sequence upstream from *porγ*, PF0791), E2 (5'-GGGAGGTGGAGCAT, sequence upstream from *slp*, PF1399), and E3 (5'-GGTGATATGCA, sequence upstream from *cipA*, PF0190). The primer sequences are given in (**Table 2.4**). SOE-PCR (splicing by overlap extension and PCR, (182)) was performed to combine the individual PCR products and generate the expression cassette for SP1 (**Figure 2.1A**).

Construction of vectors for insertion of the SP1 operon into *P. furiosus*

The SP1 expression cassette (**Figure 2.1B**) was cloned into pSPF300 (177), generating the plasmid pALM506-1, to be used for targeted insertion of the synthetic SP1 operon into the *P. furiosus* ΔpdaD strain (**Figure 2.2**). SOE-PCR (182) was used to combine ~0.5 kb flanking regions targeting homologous recombination in the intergenic space between convergent genes PF0574-PF0575, with a marker cassette, including restriction sites for cloning. The marker cassette for uracil prototrophic selection consisted of the *pyrF* gene driven the *gdh* promoter

region (P*gdh*, 157 bases upstream of PF1602) and terminated with 12 bases of the 3' UTR of the *hpyA1* gene (5'- aatctttttttag, PF1722). A 65-b sequence of the 3' end of the marker cassette (5'-ctaaaaaagattttatcttgagctccattctttcacctcctcgaaaatcttcttagcggcttccc) was repeated at the beginning of the cassette to serve as a homologous recombination region for selection of marker removal (183). Vector pGL007 targeting homologous recombination at the PF0574-PF0575 intergenic space was constructed by cloning the SOE-PCR product into pJHW006 (137) (**Figure 2.3**). The SP1 expression cassette was PCR-amplified from pALM506. A terminator sequence was added to the 3' end of the operon (5'- aatctttttttag, from the 3' UTR of PF1722), and the construct was cloned into the *Asc*I-*Not*I sites of pGL007 to make pGL010 (**Figure 2.4**), for targeted insertion of the SP1 operon at the PF0574-PF0575 intergenic space. Transformation of *P. furiosus* ΔpdaD strain was performed as previously described for COM1 (137) except that the defined medium contained maltose instead of cellobiose as the carbon source and was supplemented with 0.1% w/v casein hydrolysate. Transformation of COM1 was performed as previously described (28), except that linear plasmid DNA was used for transformation.

Growth of *P. furiosus*

Strains were cultured as previously described in a sea-water based medium containing 5 g/L maltose and 5 g/L yeast extract, 0.5 µg/L riboflavin, and 20 µM uracil or 4 mM agmatine as needed (28). Cultures were grown at 95°C until ~1x10^8 cells/mL and then cooled at 23°C until the temperature reached 70 to 75°C, which was maintained for up to 48 hours. For growth in a 20 L fermenter, the culture was sparged with 10% CO_2/90% N_2, stirred, and the pH was

maintained at 6.8 by addition of 10% $NaHCO_3$. Cell extracts prepared anaerobically as described previously (28) in 100 mM MOPS, pH 7.5, re-concentrated three-times with a 3 kDa centrifugation filter and stored at -80°C.

Growth of *M. sedula* for biochemical assays and product analysis

M. sedula (DSM 5348) was grown autotrophically at 70°C with micro-bubblers feeding 1 mL/min 80/20 H_2/CO_2 and 100 mL/min air in the defined medium, DSMZ 88, at pH 2.0 as previously described(148). To obtain cell-free extracts, frozen cell pellets were anaerobically suspended in 50 mM Tris HCl pH 8.0 containing 0.5 µg/mL DNase I and stirred for 1 hr in an anaerobic chamber. The cell extract was centrifuged at 100,000 x g for 1 hr and the supernatant was stored at -80°C.

E1, E2, and E3 assays

All reactions were carried out in sealed anaerobic cuvettes at 75°C containing 100 mM MOPS pH 7.5, 5 mM $MgCl_2$, 5 mM DTT. After addition of NADPH (to A_{340} ~ 1.0) and the relevant substrate (see below), NADPH oxidation was measured at 340 nm. The substrates for the E2, E2+E3 and E1+E2+E3 assays were succinyl-CoA, malonyl-CoA and acetyl CoA (each 1 mM) respectively. The latter assay also contained 1 mM ATP, and 10 mM $NaHCO_3$. E1 activity was measured by phosphate release. The assay contained 10 mM $NaHCO_3$, 1 mM ATP, and 1 mM acetyl-CoA. Samples (20 µL) were removed at 2-4 min, diluted with water (180 µl), and the BioVision (Mountain View, CA) phosphate assay reagent (20 µl) was added. The phosphate produced was calculated using a molar extinction coefficient of 90,000 $M^{-1}cm^{-1}$ at 650 nm.

Measurement of 3-hydroxypropionic acid (3-HP)

3-HP (H0297, 30%, w/v, in water) was obtained from TCI America (http://www.tciamerica.net/). By HPLC and ^{1}H NMR, it was 75% pure with the remaining 25% as 3,3'-oxydipropanoic acid. For GC-MS analysis, inositol was the internal standard. Samples were freeze-dried, incubated in 2 M trifluoroacetic acid at 80°C for 1 hour, dried under nitrogen, and per-*O*-trimethylsilylated by treatment with Tri-Sil (Pierce) at 80°C for 30 minutes. GC-MS analysis was performed on an AT 7890n GC interfaced to a 5975C MSD using a Grace EC-1 column (30 m x 0.25 mm). The exact mass of 3-HP-TMS is 162. Derivatization of 3-HP with 2-nitrophenyl hydrazine was carried out as described previously (184). The 3HP-hydrazide was extracted by adding 1.0 mL of 1 M KPO_4 buffer pH 7.0 and 1.5 mL of ether to 800 µL of the sample, centrifuging for 10 min at 6,000 x g to separate the phases, removing the top ether layer and evaporating. The dried sample was resuspended in 200 µL ethanol and 10-50 µL aliquots were analyzed by HPLC. The column and run conditions were as follows: column, Supelco LiChrosorb RP-8 (5 µm); solvent system, A 0.05%TFA, B 100% acetonitrile; gradient 0-20min, 0-100% B, 20-22min: 100 %B; flow rate: 1 mL/min; temperature: 30°C. For ESI-MS analysis, the dried derivative was dissolved in methanol and directly injected on a Perkin-Elmer API 1 plus in negative mode. The mass of the anionic 3-HP-hydrazide derivative is 224.

Production of 3-HP *in vitro* from malonyl-CoA by E2+E3 and from acetyl-CoA by E1+E2+E3

To the *P. furiosus* extract (1-2 mg/mL) in 100 mM MOPS pH 7.5, 5 mM $MgCl_2$, and 5 mM DTT, was added 1-2 mM malonyl-CoA (for E2+E3) or 10 mM $NaHCO_3$ (or 100% CO_2 in the gas phase), 2 mM ATP and 2 mM acetyl-CoA (for E1+E2+E3). The electron source was 2 mM NADPH

or 0.5 mM NADP with 20% H_2 in the headspace. Sealed anaerobic vials containing the reaction mixture were incubated at 75°C for up to 2 hr. Samples were derivatized with 2-nitrophenyl hydrazine and analyzed for 3-HP by HPLC as described above.

Product analysis of E1+E2+E3 activities in whole cells

P. furiosus strains PF506 and MW56 were grown in 2 L cultures at 95°C for 10 hours until cell densities of 1 x 10^8 cells/mL and then cooled and incubated at 75 °C for 16 hours. Harvested cells were suspended to 5 x 10^{10} cells/mL in 100 mM MOPS pH 7.5 and base salts (28 g/L NaCl, 3.5 g/L $MgSO_4$ • 7 H_2O, 2.7 g/L $MgCl_2$ • 6 H_2O, 0.33 g/L KCl, 0.25 g/L NH_4Cl, 0.14 g/L $CaCl_2$ • $2H_2O$). The cell suspension was sealed in a serum vial, degassed with argon, and cysteine HCl (0.5 g/L), $NaHCO_3$ (10 mM) and either maltose (10 mM) or pyruvate (40 mM) were added. The vials were degassed and flushed with H_2 and incubated at 75°C for 60 minutes. Samples for 3-HP analysis were derivatized with 2-nitrophenyl hydrazine, using 1 mM p-hydroxyphenyl acetic acid as an internal standard, ether-extracted and analyzed by HPLC as described above.

Analysis of the *P. furiosus* culture medium for 3-HP

P. furiosus strains PF506, MW56 and COM1 were grown at 98°C in 50 mL cultures with maltose (10 mM) as the carbon source until a cell density of 8 x10^7 cells/mL was reached. The incubation temperature was then shifted to 72°C for up to 4 days. Sample (1 mL) were periodically removed, centrifuged (10,000 x g, 10 min) and to a 100 µl aliquot of the supernatant (the spent medium) 1 mM p-hydroxyphenyl acetic acid was added as an internal standard. The sample was derivatized with 2-nitrophenyl hydrazine, ether-extracted and analyzed by HPLC as described above.

Acknowledgements

We thank Brian Vaccaro, Dennis Phillips, and Zhirui Wang (University of Georgia) for running the NMR, ESI-MS, and GC-MS analyses of 3-HA, respectively. The GC-MS facility is supported in part by the Center for Plant and Microbial Complex Carbohydrates, funded by the US Department of Energy (DE-FG02-93ER-20097). This work was supported by the US Department of Energy as part of the Electrofuels Program of the Advanced Research Projects Agency - Energy (Grant DE-AR0000081).

Author Contributions

Author contributions: R.M.K. and M.W.W.A. designed research; M.W.K., G.J.S., G.L.L., A.L.M., I.J.I., T.T.L., M.P.T., W.J.N., and A.S.H. performed research; M.W.K., G.J.S., G.L.L., A.L.M., I.J.I., M.P.T., W.J.N., and M.W.W.A. analyzed data; and M.W.K., G.J.S., G.L.L., R.M.K., and M.W.W.A. wrote the paper.

Tables and Figures

Table 2.1 Strains used and constructed in the study of SP1

Strain	Parent	Genotype/Description	Reference
COM 1	DSM 3638	Δ*pyrF*	(137)
ΔpdaD	COM 1	Δ*pyrF* Δ*pdaD*::P_{gdh}*pyrF*	(177)
PF506	ΔpdaD	Δ*pyrF* Δ*pdaD*::*pdaD* P_{slp}^{-}E1αβγ-E2-E3	This work
MW56	COM 1	Δ*pyrF* P_{gdh}*pyrF* P_{slp}^{-}E1αβγ-E2-E3	This work

Table 2.2 GC-MS verification of 3-HP

GC-MS identification and quantitation of 3-HP produced from malonyl-CoA and NADPH or H_2 by cell-free extracts of *P. furiosus* strain PF506. The assays were carried out in a total volume of 1 mL containing 0.25 mg of cell-free extract under H_2 in a shaking water bath. The amount of 3-HP produced was determined after 2 hr at 72°C.

Vial	Added Electron Donor	Substrate	Theoretical 3-HP	3-HP/ Inositol peak area	Estimated 3-HP
1	2 mM NADPH	2 mM malonyl-CoA	1 mM	0.0288	0.2 mM
2	2 mM NADPH, H_2	2 mM malonyl-CoA	2 mM	0.0467	0.3 mM
3	1 mM NADP, H_2	2 mM malonyl-CoA	2 mM	0.0274	0.2 mM
4	1 mM NADP, H_2	None (control)	0	0.0064	0.05 mM
5	1 mM NADP, H_2	None (control)	2 mM	0.2839	2.0 mM

Table 2.3 3-HP production from pyruvate or maltose

3-HP production using maltose or pyruvate as the source of acetyl-CoA by whole cells of *P. furiosus* strains PF506 and MW56. The amount of 3-HP indicated was present in 1 mL of the cell suspension of *P. furiosus*.

MW56		**PF506**	
Pyruvate	**Maltose**	**Pyruvate**	**Maltose**
155 nmol	100 nmol	70 nmol	145 nmol

Table 2.4 Primers used in the construction of the synthetic SP1 operon

Primer target	Direction	5' to 3' sequence
P_{slp}	Forward	GAATCCCCGCGGCCCGGGCTGGCAGAATAGAA
	Reverse	GCAACCAAAACTCTACTAAAGGGTGGCATTTTTCTCCACCTCCCAATAATCTG
Msed_0147-0148	Forward	ATGCCACCCTTTAGTAGAGTTTTGG
	Reverse	GTTGCAGTCATCTTCAAACCTCCTTACTTTATCACCACTAGGATATCTCC
Msed1375	Forward	GTGATAAAGTAAGGAGGTTTGAAGATGACTGCAACTTTTGAAAAACCGGAT
	Reverse	CGTTCTCCTCATATGCTCCACCTCCCTTAGAGGGGTATATTTCCATGCTTC
Msed_0709	Forward	GGCAATGTCATATGAGGAGAACGCTAAAGGCCGCAATTC
	Reverse	CCTTTTCAGTCATTGCATATCACCTCATCTCTTGTCTATGTAGCCCTTC
Msed_1993	Forward	TAGACAAGAGATGAGGTGATATGCAATGACTGAAAAGGTATCTGTAGTTGGAG
	Reverse	CCAATGCATGCTTATTTTTCCCAAACTAGTTTGTATACCTTC

Figure 2.1 The first sub-pathway in the *Metallosphaera sedula* cycle for carbon dioxide assimilation

The expression of the first sub pathway (SP1) of the *M. sedula* 3-HP/4-HB cycle for carbon fixation. **(A)** The synthetic operon constructed to express the *M. sedula* genes encoding E1 (αβγ), E2 and E3 in *P. furiosus* under the control of the promoter for the S-layer protein gene (P_{slp}). This includes *P. furiosus* ribosomal binding sites (rbs) from highly-expressed genes encoding pyruvate ferredoxin oxidoreductase subunit γ (*porγ,* PF0971), the S-layer protein (*slp,* PF1399) and cold-induced protein A (*cipA,* PF0190). **(B)** The first three enzymes of the *M. sedula* 3-HP/4-HB cycle produce the key intermediate 3-hydroxypropionate (3-HP). E1 is acetyl/propionyl-CoA carboxylase (αβγ, encoded by Msed_0147, Msed_0148, Msed_1375): E2 is malonyl/succinyl-CoA reductase (Msed_0709) and E3 is malonate semialdehyde reductase (Msed_1993). NADPH is generated by *P. furiosus* soluble hydrogenase I (SH1), which reduces NADP with hydrogen gas. **(C)** The first three enzymes (E1-E3) are shown in context of the complete 3-HP/4-HP cycle for carbon dioxide fixation by *Metallosphaera sedula* showing the three sub-pathways, SP1 (blue), SP2 (green) and SP3 (red). **(D)** The horizontal scheme shows the amount of energy (ATP), reductant (NADPH), oxidant (NAD) and CoASH required to generate one mole of acetyl-CoA from two moles of carbon dioxide

Figure 2.1

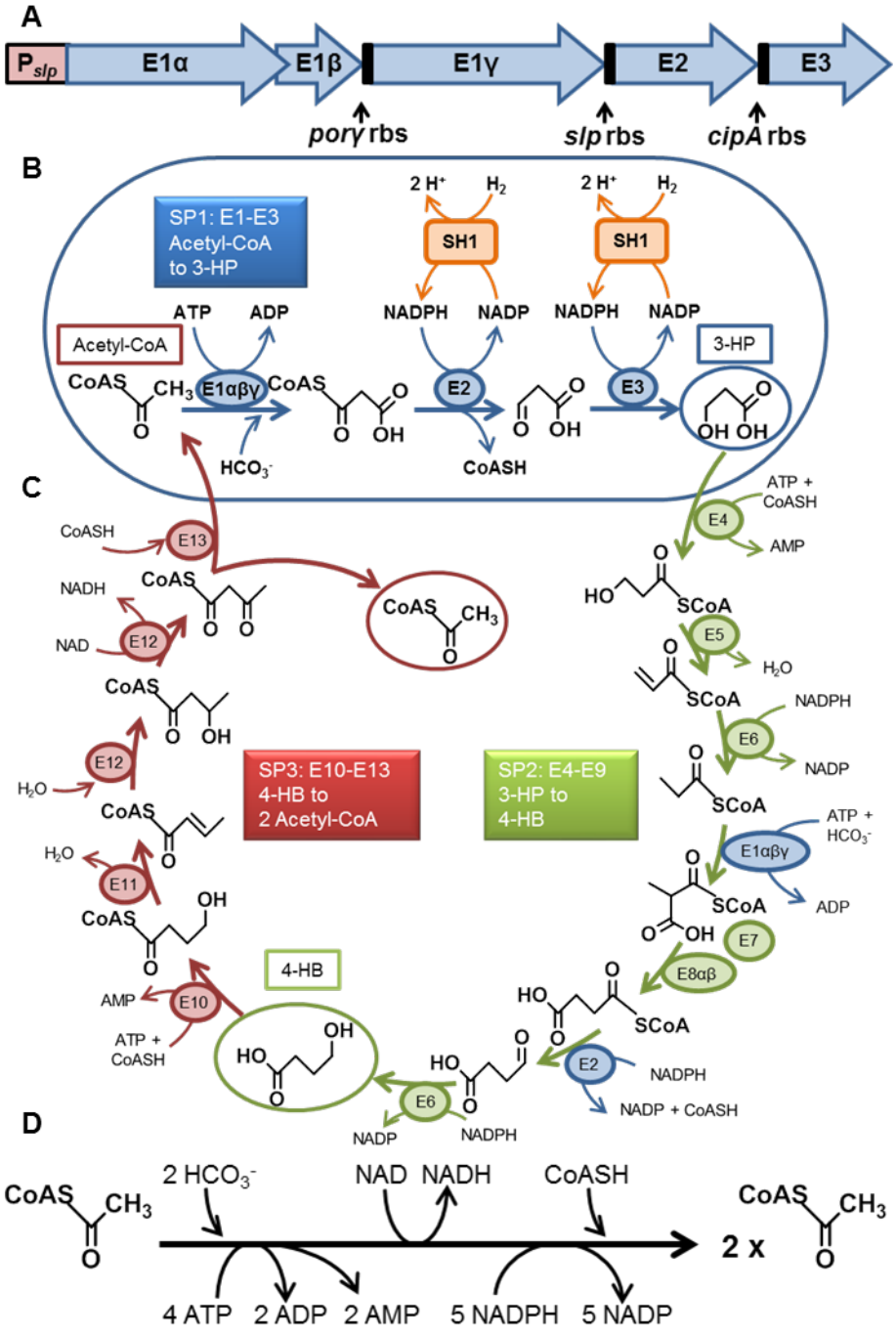

A
P slp
E1α
E1β
E1γ
E2
E3
porγ rbs
slp rbs
cipA rbs
B
SP1: E1-E3
Acetyl-CoA
to 3-HP
2 H+
H2
SH1
NADPH
NADP
ATP
ADP
Acetyl-CoA
E1αβγ
HCO3-
E2
CoASH
E3
3-HP
C
CoASH
E13
NADH
NAD
E12
H2O
E11
AMP
E10
ATP +
CoASH
4-HB
SP3: E10-E13
4-HB to
2 Acetyl-CoA
SP2: E4-E9
3-HP to
4-HB
E4
AMP
E5
E6
E7
E8αβ
ADP
D
2 HCO3-
NAD NADH
CoASH
2 x
4 ATP
2 ADP
2 AMP
5 NADPH
5 NADP

Figure 2.2 Plasmid map of pALM506-1 used to transform *P. furiosus* strain Δ*pdaD* to generate strain PF506

Figure 2.2

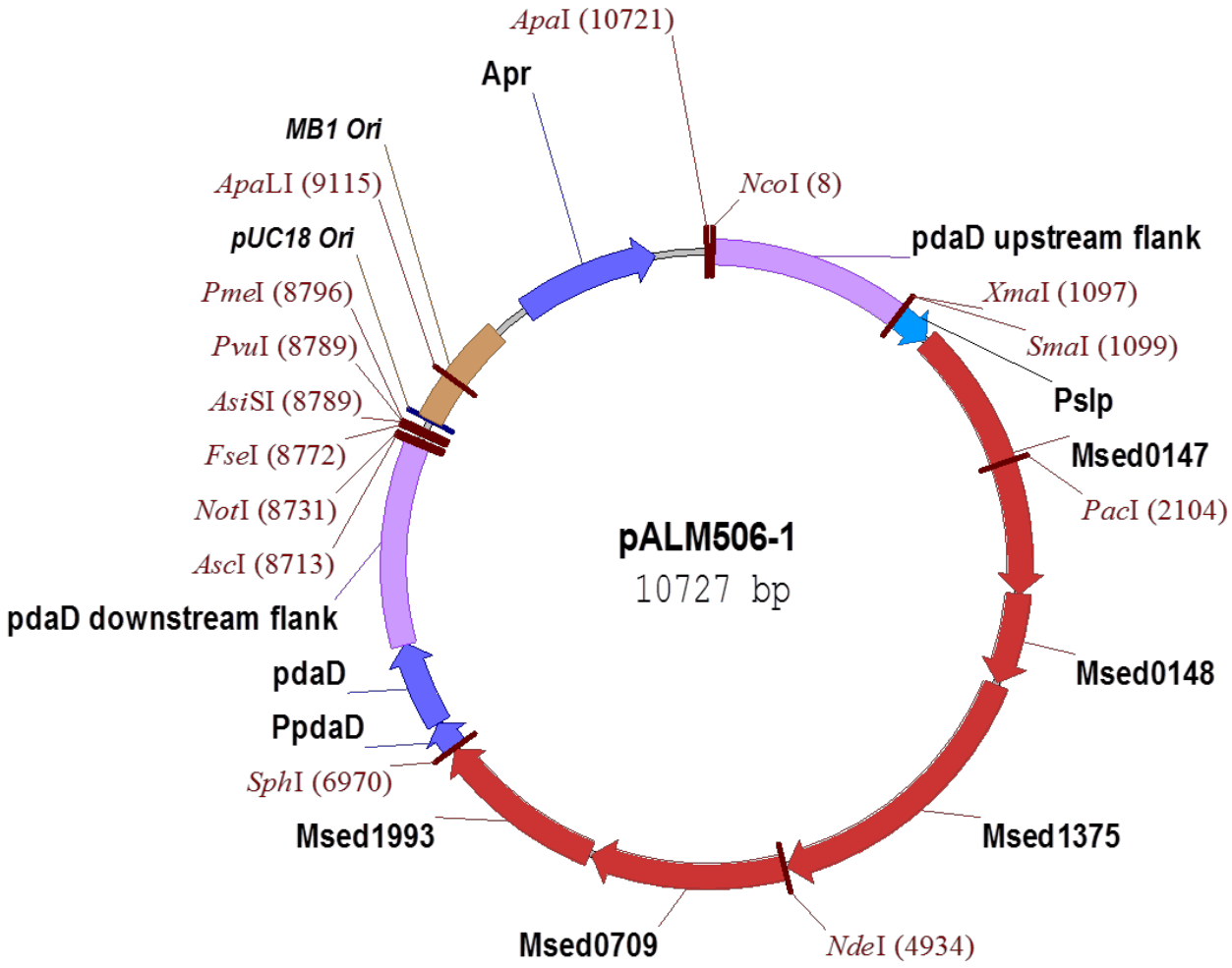

ApaI (10721)
Apr
MB1 Ori
ApaLI (9115)
NcoI (8)
pUC18 Ori
pdaD upstream flank
PmeI (8796)
XmaI (1097)
PvuI (8789)
SmaI (1099)
AsiSI (8789)
Pslp
FseI (8772)
Msed0147
NotI (8731)
PacI (2104)
pALM506-1
10727 bp
AscI (8713)
pdaD downstream flank
Msed0148
pdaD
PpdaD
SphI (6970)
Msed1993
Msed1375
Msed0709
NdeI (4934)

Figure 2.3 Plasmid map of pGL007 vector targeting the region between PF0574 and PF0575 in the *P. furiosus* genome

Figure 2.3

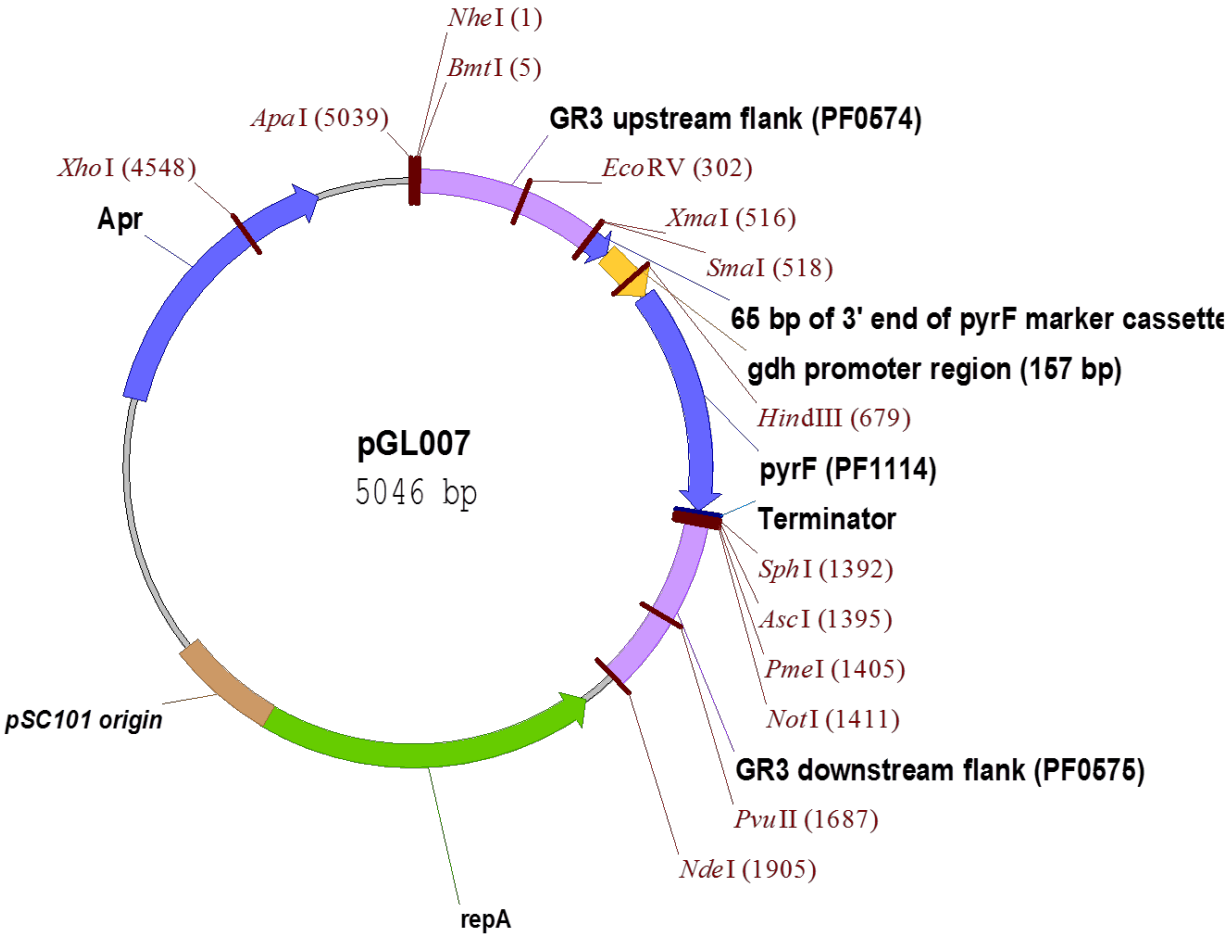

NheI (1)
BmtI (5)
ApaI (5039)
GR3 upstream flank (PF0574)
XhoI (4548)
EcoRV (302)
Apr
XmaI (516)
SmaI (518)
66 bp of 3' end of pyrF marker cassette
gdh promoter region (157 bp)
HindIII (679)
pGL007
5046 bp
pyrF (PF1114)
Terminator
SphI (1392)
AscI (1395)
PmeI (1405)
NotI (1411)
pSC101 origin
GR3 downstream flank (PF0575)
PvuII (1687)
NdeI (1905)
repA

Figure 2.4 Plasmid map of pGL010 used to transform *P. furiosus* COM1 to generate strain MW56

Figure 2.4

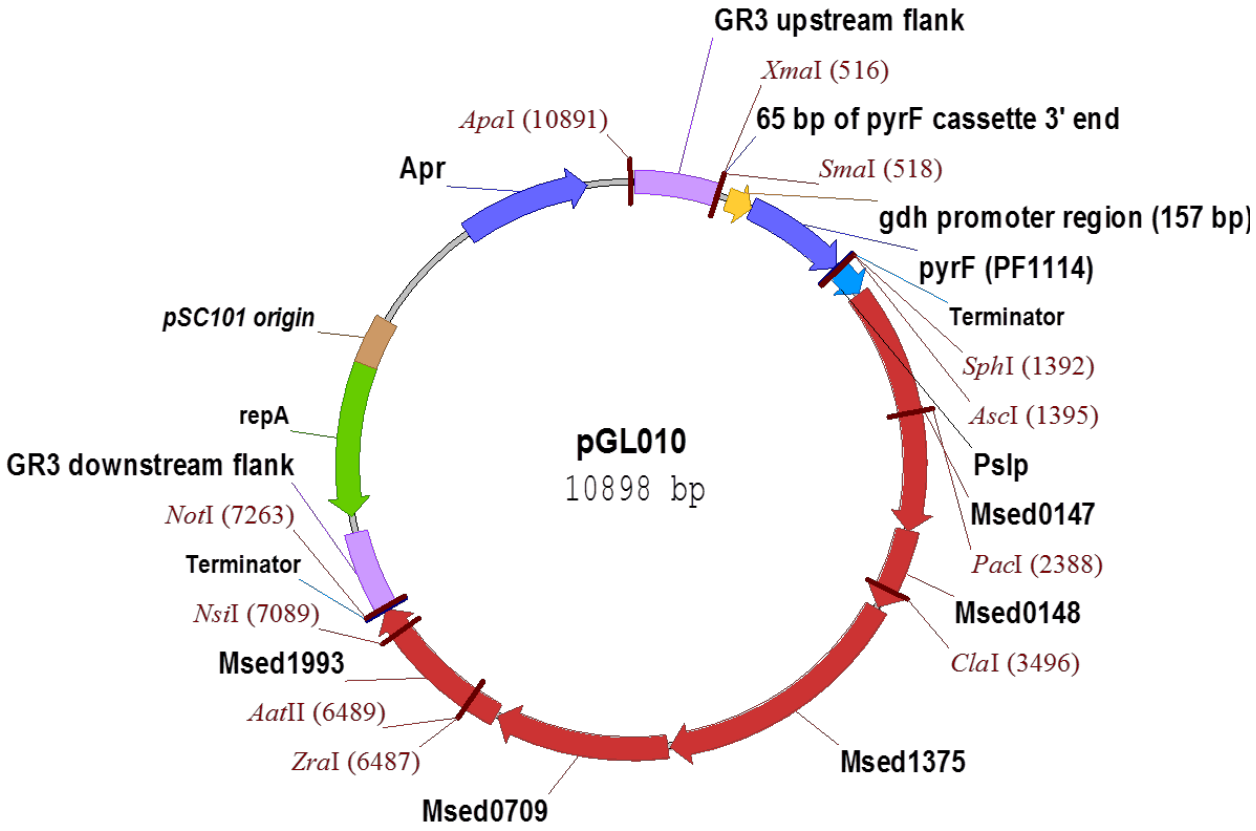

GR3 upstream flank
XmaI (516)
65 bp of pyrF cassette 3' end
ApaI (10891)
SmaI (518)
Apr
gdh promoter region (157 bp)
pyrF (PF1114)
Terminator
pSC101 origin
SphI (1392)
AscI (1395)
repA
pGL010
10898 bp
PsIp
GR3 downstream flank
Msed0147
NotI (7263)
PacI (2388)
Terminator
Msed0148
NsiI (7089)
ClaI (3496)
Msed1993
AatII (6489)
ZraI (6487)
Msed1375
Msed0709

Figure 2.5 Time and temperature-dependent production of the SP1 pathway enzymes in *P. furiosus* strain PF506

(A) Growth of triplicate cultures at 98°C (red circles) and temperature (black line) for the temperature shift from 98 to 75°C are shown. **(B)** Specific activity (µmol NADPH oxidized/min/mg) of the coupled activity of E2+E3 in cell-free extracts from cultures grown at 95°C to a high cell density of 1 x 10^8 cells/ml and then incubated for 18 hrs at the indicated temperature. **(C)** Activities of E1, E2+E3, and E1+E2+E3 after the temperature shift to 75°C for the indicated time period (see Fig. S4). The activities of a cell-free extract of autotrophically-grown *M. sedula* cells is also shown (labeled Msed). The specific activities are: E1+E2+E3 coupled assay with acetyl-CoA and bicarbonate (blue), E2+E3 coupled assay with malonyl-CoA (red), and E2 with succinyl-CoA (green) as substrates. **(D)** Temperature dependence of the coupled activity of E2+E3 (blue circles) in the cell-free extracts after induction at 72°C for 16 hr. The activity of *P. furiosus* glutamate dehydrogenase in the same cell-free extracts is also shown (red squares).

Figure 2.5

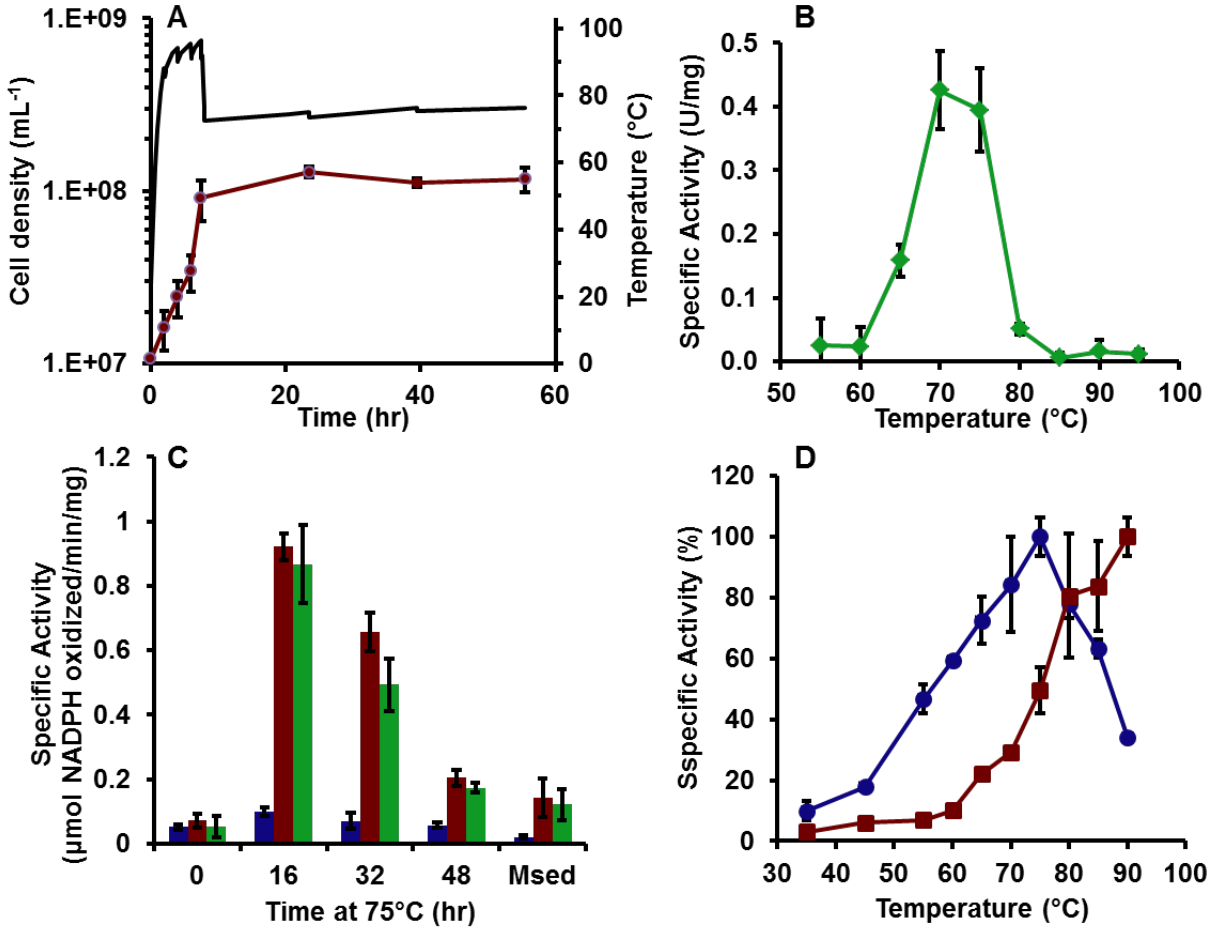

A
1.E+09
1.E+08
1.E+07
Cell density (mL-1)
100
80
60
40
20
0
Temperature (°C)
0
20
40
60
Time (hr)
B
0.5
0.4
0.3
0.2
0.1
0.0
Specific Activity (U/mg)
50
60
70
80
90
100
Temperature (°C)
C
1.2
1
0.8
0.6
0.4
0.2
0
Specific Activity
(µmol NADPH oxidized/min/mg)
0
16
32
48
Msed
Time at 75°C (hr)
D
120
100
80
60
40
20
0
Sspecific Activity (%)
30
40
50
60
70
80
90
100
Temperature (°C)

Figure 2.6 SP1 enzyme activities in *M. sedula* and *P. furiosus* strains

E1 (blue) and coupled E2+E3 (red) in cell-free extracts of the indicated *P. furiosus* strains after incubation at 70°C for 16 hr (**Fig. 2.9**), compared to that measured for the cell-extract of autotrophically-grown *M. sedula* cells (labeled Msed).

Figure 2.6

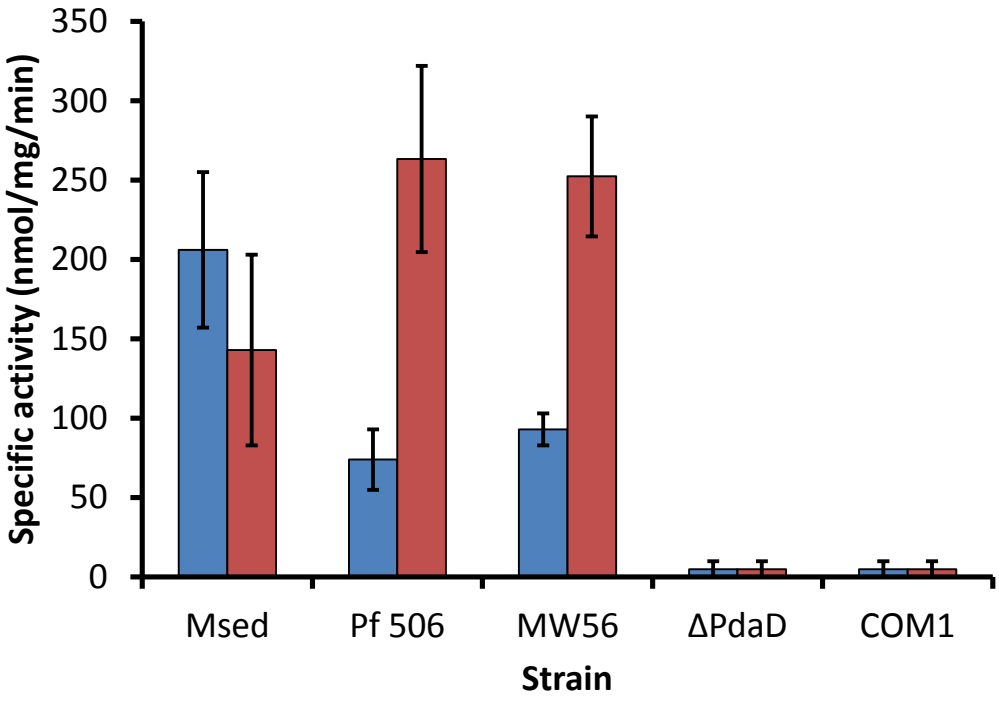

Specific activity (nmol/mg/min)
350
300
250
200
150
100
50
0
Msed
Pf 506
MW56
ΔPdaD
COM1
Strain

Figure 2.7 3-HP production by *P. furiosus*

Cells were grown at 95°C and then incubated at 72°C for 16 hr to produce the SP1 enzymes. **(A)** *In vitro* 3-HP production from acetyl-CoA performed in triplicate. The sources of the C1 carbon (CO_2 or HCO_3^-) and reducing equivalents (NADPH or NADP/H_2) are indicated. Rates are expressed as µmoles of 3-HP produced/min/mg. **(B)** *In vivo* 3-HP production by whole cells (static) using maltose as the source of acetyl-CoA in the presence of hydrogen gas and bicarbonate using cells grown in a 100 ml sealed bottles without pH control. The *P. furiosus* strains are MW56 (circles, blue) and COM1 (squares, red). **(C)** *In vivo* 3-HP production by whole cells (stirred) of MW56 using maltose as the source of acetyl-CoA (circles, blue) and E2+E3 specific activity of the cell-free extracts (diamonds, green) using cells grown in a 20 L fermenter with pH control (6.8).

Figure 2.7

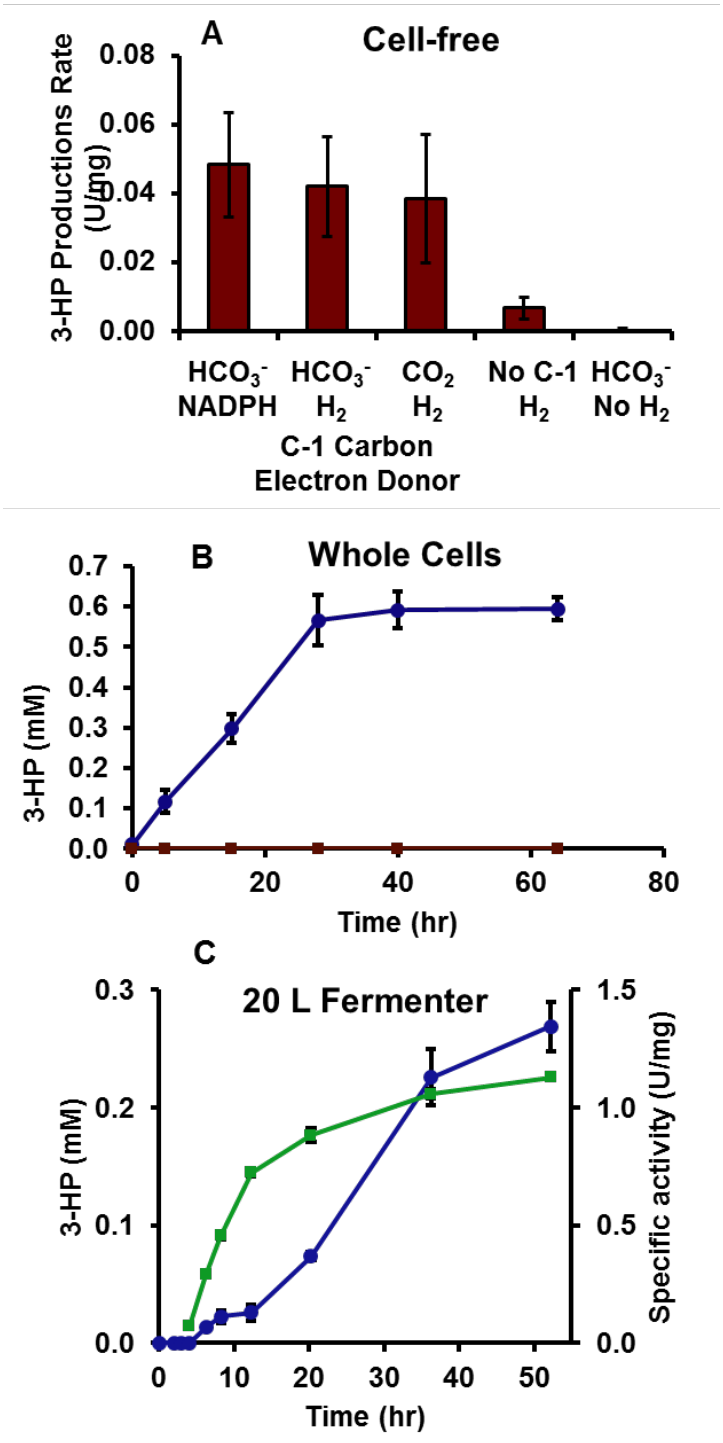

A
Cell-free
3-HP Productions Rate (U/mg)
0.08
0.06
0.04
0.02
0.00
HCO3- NADPH
HCO3- H2
CO2 H2
No C-1 H2
HCO3- No H2
C-1 Carbon Electron Donor
B
Whole Cells
3-HP (mM)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
20
40
60
80
Time (hr)
C
20 L Fermenter
3-HP (mM)
0.3
0.2
0.1
0.0
1.5
1.0
0.5
0.0
Specific activity (U/mg)
0
10
20
30
40
50
Time (hr)

Figure 2.8 The thermostability of E2 and E3 versus the endogenous *P. furiosus* glutamate dehydrogenase

Using an E2+E3 coupled assay at 75°C after incubation at 90°C for the indicated amount of time in cell-free extracts of *P. furiosus* strain PF506 (blue circles) and of the endogenous *P. furiosus* glutamate dehydrogenase (red squares). The specific activity of E2+E3 in PF506 (grown at 72°C) is about 2-fold higher than that measured in *M. sedula* and exhibits similar thermostability as the endogenous *P. furiosus* glutamate dehydrogenase. Activity is expressed as percent of maximum activity.

Figure 2.8

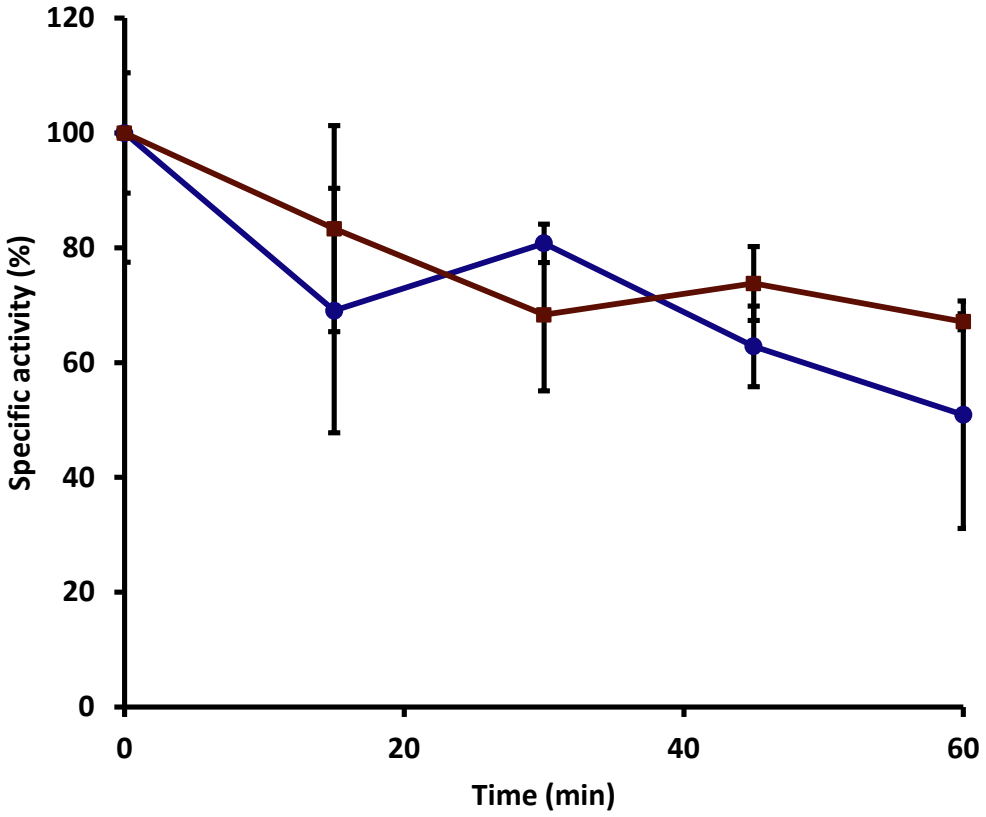


Figure 2.9 Growth of *P. furiosus* COM 1, MW56 and PF506 during the temperature shift from 98°C to 70°C

Cell densities of COM1 (blue diamonds), MW0056 (red squares), and PF506 (green triangles) are indicated. The 400 mL cultures were grown at 95°C for 9 hr and then allowed to cool at room temperature to 70°C before being placed in a 70°C incubator.

Figure 2.9

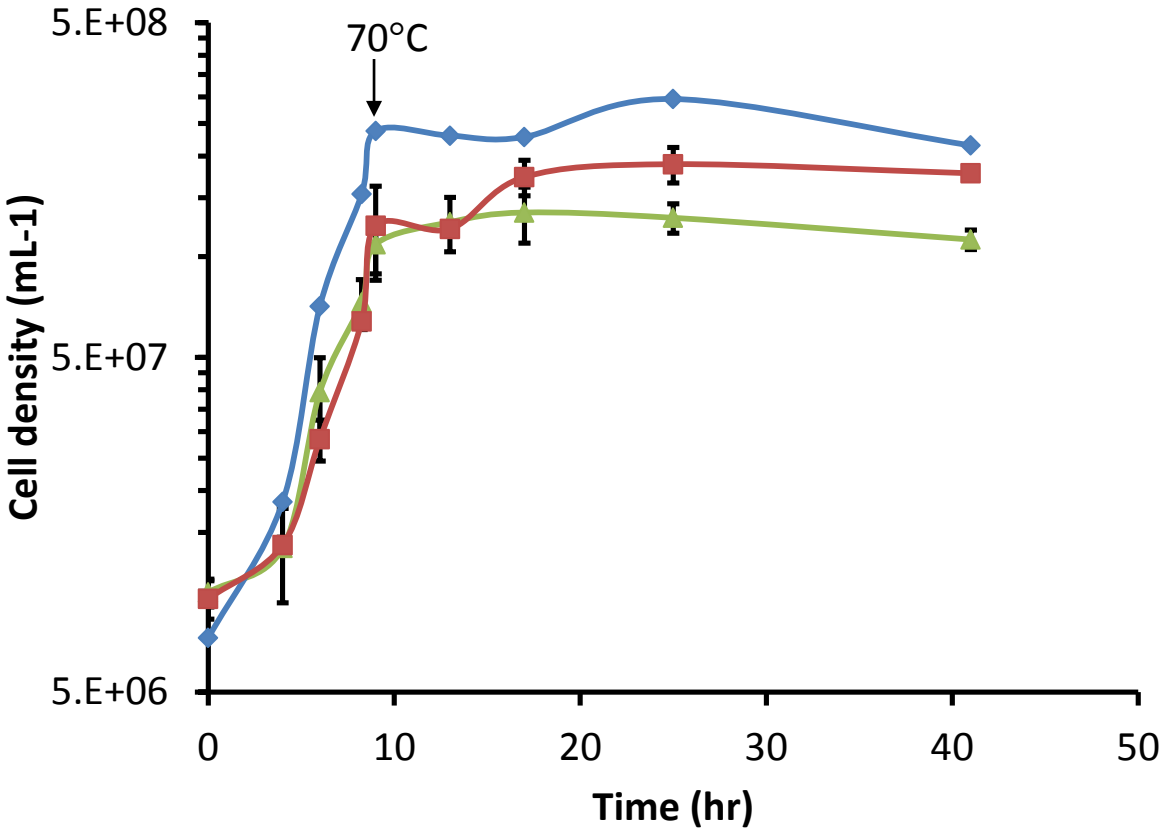


Figure 2.10 ESI-MS identification of 3-HP

ESI-MS identification of 3-HP produced from acetyl-CoA, CO_2 and H_2 (or NADPH) by cell-free extracts of *P. furiosus* strains ΔPdaD (A) and PF506 (B). The MS peak corresponding to the 3HP derivative (m/z 224, green circle) was present above background only in the recombinant PF506 strain

Figure 2.10

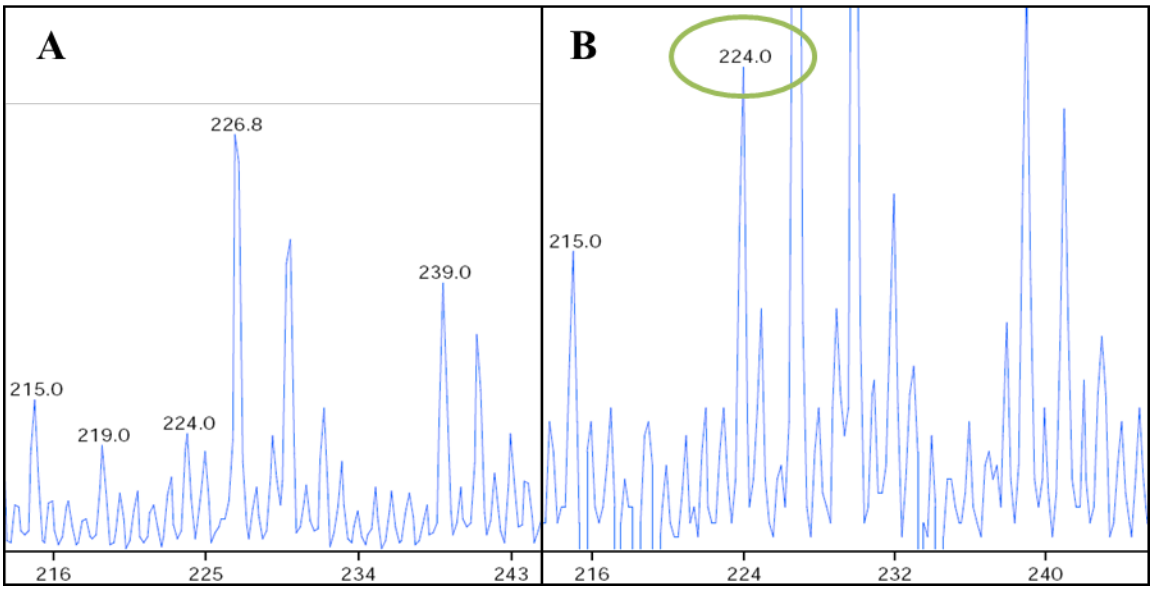

A
226.8
239.0
215.0
219.0
224.0
216
225
234
243
B
224.0
215.0
216
224
232
240

Figure 2.11 Carbohydrate metabolism by *P. furiosus*

Maltose and pyruvate metabolism by *P. furiosus,* and the key roles of pyruvate ferredoxin oxidoreductase (POR) in acetyl-CoA production and of the membrane-bound hydrogenase (MBH) in H_2 production.

Figure 2.11

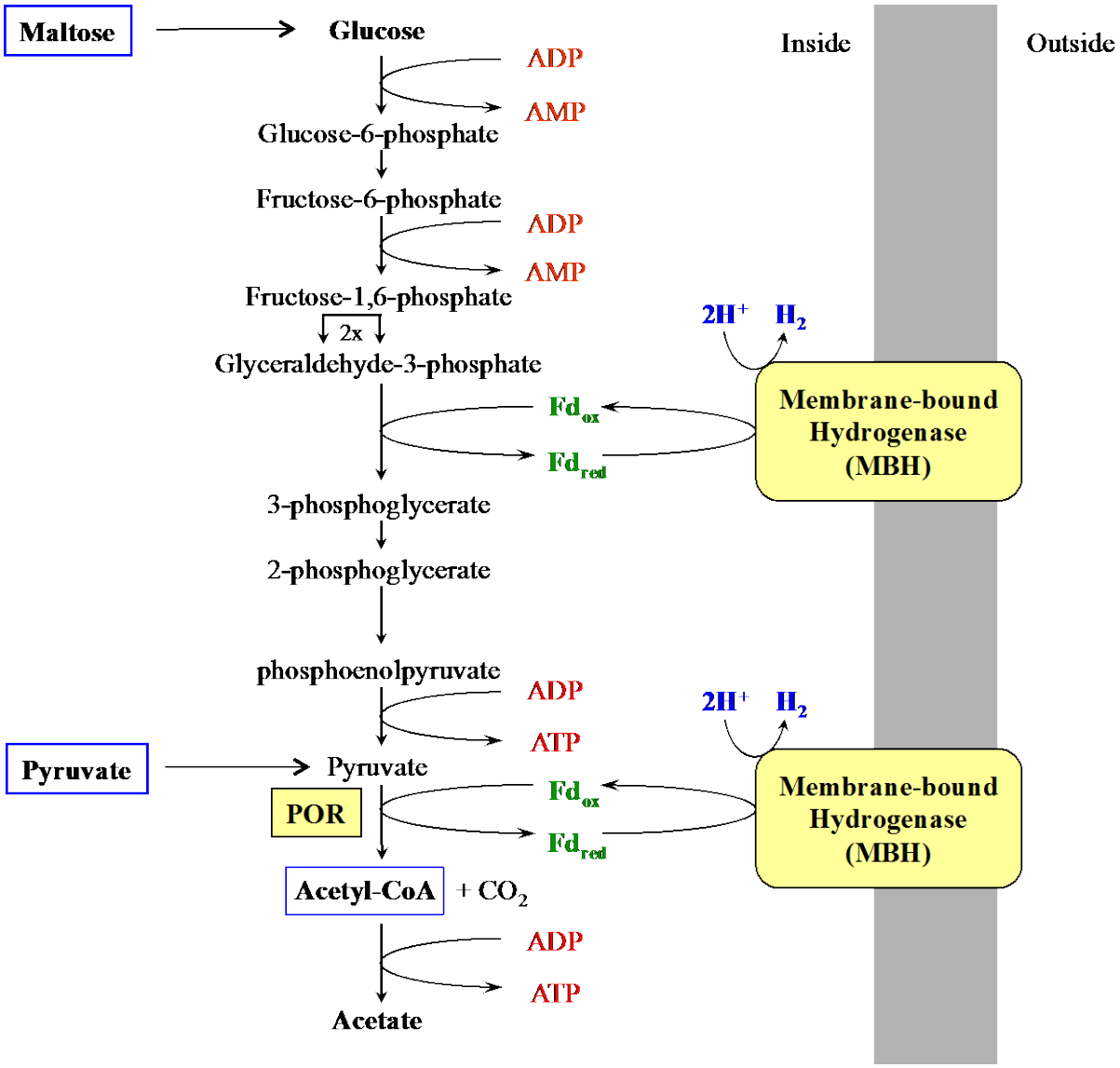

Maltose
Glucose
ADP
AMP
Glucose-6-phosphate
Fructose-6-phosphate
ADP
AMP
Fructose-1,6-phosphate
2x
Glyceraldehyde-3-phosphate
Fd ox
Fd red
3-phosphoglycerate
2-phosphoglycerate
phosphoenolpyruvate
ADP
ATP
Pyruvate
Pyruvate
POR
Fd ox
Fd red
Acetyl-CoA + CO2
ADP
ATP
Acetate
Inside
Outside
2H+ H2
Membrane-bound Hydrogenase (MBH)
2H+ H2
Membrane-bound Hydrogenase (MBH)

Figure 2.12 *In vivo* 3-HP production in strain MW56

In vivo production of 3-HP from maltose by whole cells of *P. furiosus* strain MW56 after 10 min (blue) and 60 min (red) compared to a 1 mM 3-HP standard (black). A black arrow indicates the position of the 3-HP peaks. A total of 135 µM of 3-HP was produced by a cell suspension of MW56 ($5x10^{10}$ cells/mL) after 60 min at 75°C.

Figure 2.12

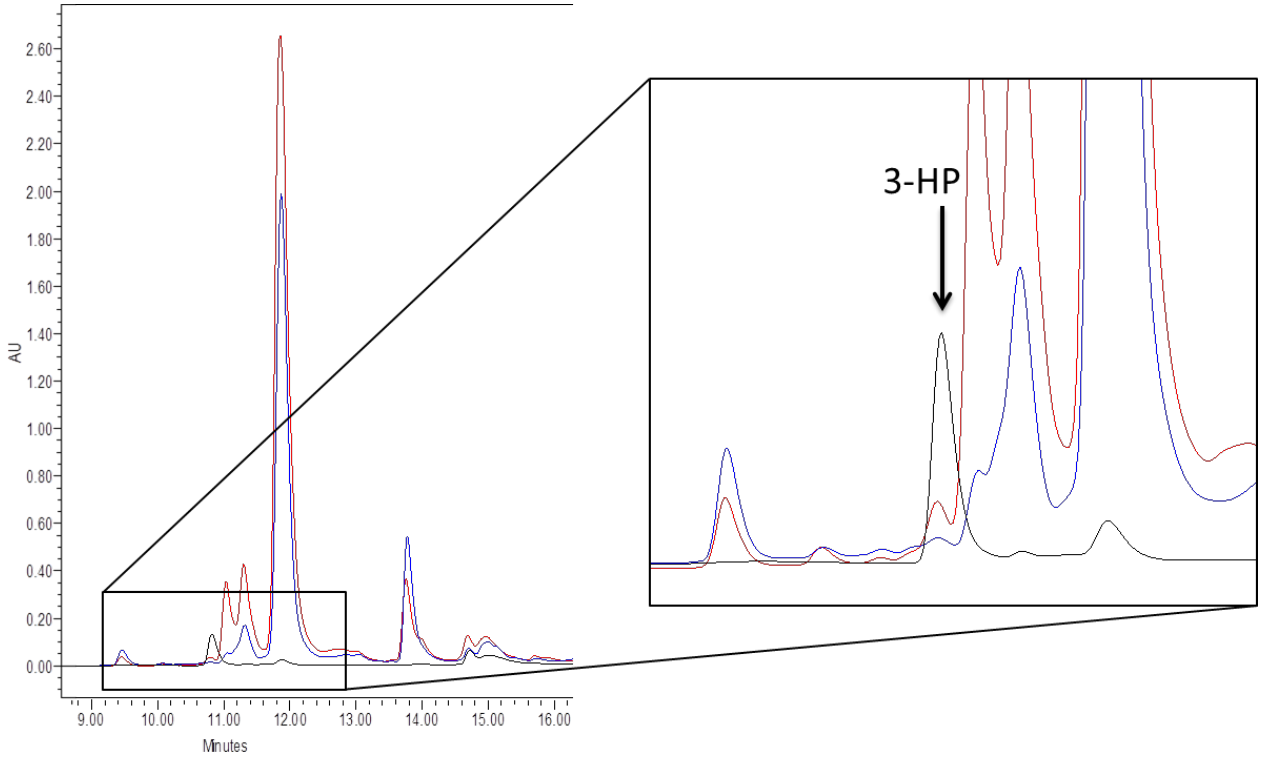
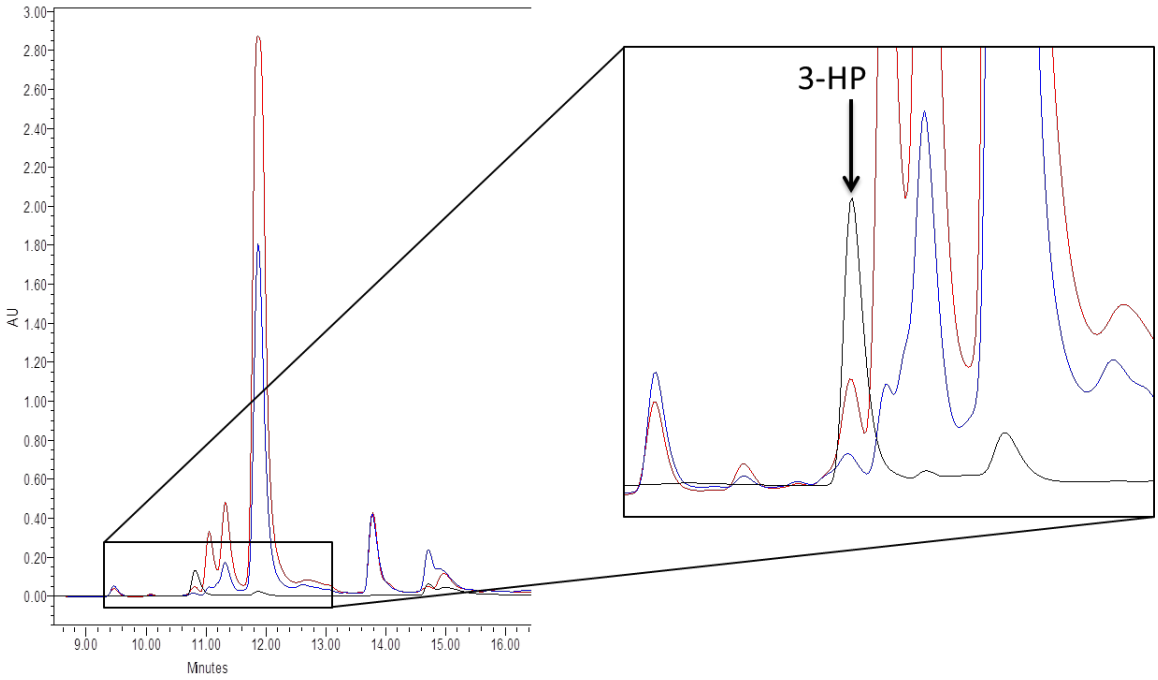

AU
2.60
2.40
2.20
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00
Minutes
3-HP

Figure 2.13 *In vivo* 3-HP production in strain PF506

In vivo production of 3-HP from maltose by whole cells of *P. furiosus* strain PF506 after 10 min (blue) and 60 min (red) compared to 1 mM 3-HP sample (black). A black arrow indicates the position of the 3-HP peaks. A total of 199 µM of 3-HP was produced by a cell suspension of PF506 ($5x10^{10}$ cells/mL) after 60 min at 75°C.

Figure 2.13


3-HP
AU
Minutes
3.00
2.80
2.60
2.40
2.20
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
9.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00

CHAPTER 3

A HYBRID SYNTHETIC PATHWAY FOR BUTANOL PRODUCTION BY A HYPERTHERMOPHILIC MICROBE

Keller MW, Lipscomb GL, Loder AJ, Schut GJ, Kelly RM, Adams MWW (2015) A hybrid synthetic pathway for butanol production by a hyperthermophilic microbe. *Metab Eng* 27:101-6.

Reprinted here with permission of the publisher Elsevier

Abstract

Biologically produced alcohols are of great current interest for renewable solvents and liquid transportation fuels. While bioethanol is now produced on a massive scale, butanol has superior fuel characteristics and an additional value as a solvent and chemical feedstock. Butanol production has been demonstrated at ambient temperatures in metabolically-engineered mesophilic organisms, but the ability to engineer a microbe for *in vivo* high-temperature production of commodity chemicals has several distinct advantages. These include reduced contamination risk, facilitated removal of volatile products, and a wide temperature range to modulate and balance both the engineered pathway and the host's metabolism. We describe a synthetic metabolic pathway assembled from genes obtained from three different sources for conversion of acetyl-CoA to 1-butanol, and 1-butanol generation from glucose was demonstrated near 70°C in a microorganism that grows optimally near 100°C. The module could also be used in thermophiles capable of degrading plant biomass.

Introduction

Butanol as an advanced biofuel has several advantages over ethanol. It has a greater energy density, can be used in higher blend ratios, and is more compatible with the current infrastructure for transportation (59, 185-187). For this reason, a significant amount of research has been directed toward microbial production of butanol and so-called 'advanced biofuels', by metabolically-engineered mesophilic organisms (188). High titers (30 g/L) of 1-butanol from glucose have been achieved by a heavily engineered strain of *Escherichia coli* with *in situ* product removal (170). Butanol and isobutanol have also been generated at ambient temperature using carbon dioxide as the carbon source in the mesophilic bacteria *Clostridium ljungdahlii* (189), *Synechoccoccus elongates* (190) and *Ralstonia eutropha* (191). While platform microorganisms that thrive at extreme temperatures offer potentially critical advantages for biofuel production (172, 192-194), few thermophilic organism has been engineered for butanol production: increased butanol in *Thermoanaerobacterium thermosaccharolyticum* (formerly *Clostridium thermosaccharolyticum*) at 55°C by overexpression of the native *bcs* operon (195), *Thermoanaerobacterium saccharolyticum* at 55°C by expression of the butanol pathway from the closely related *Thermoanaerobacterium thermosaccharolyticum* (196), and isobutanol production from *Geobacillus thermoglucosidasius* at 50°C by expressing and optimizing an isobutanol pathway from multiple gene donors (197).

Metabolic engineering of thermophilic microorganisms for the production of a variety of useful chemicals such as polymer feedstocks and liquid transportation fuels has traditionally been limited by the availability of genetically tractable model organisms. Within the last

decade, however, genetic capabilities have been developed for several that grow at 85°C and above (137, 198, 199). An advantage of using a thermophilic organism for metabolic engineering is the possibility of utilizing temperature as a way to modulate and balance both the engineered pathway and the host's metabolism. This novel concept has been demonstrated in the hyperthermophile *Pyrococcus furiosus,* which has an optimal growth temperature (T_{opt}) of 100°C, by the heterologous expression of lactate dehydrogenase from *Caldicellulosiruptor bescii* (T_{opt} 78°C) using a low temperature (72°C) shift to induce lactate formation (172). *P. furiosus* has also been engineered to utilize hydrogen and incorporate carbon dioxide into the valuable molecule 3-hydroxypropionate using a similar strategy for temperature-dependent product formation (192). Herein we demonstrate metabolic engineering of *P. furiosus* for the temperature-dependent production of the liquid fuel and chemical feedstock 1-butanol.

Materials and Methods

Strain Construction

The pathway genes for BuOH-1 (*thl*, TTE0549; *hbd*, TTE0548; *crt*, TTE0544; *ter*, Stherm_c16300; *bad*, Teth514_1942; *bdh*, Teth514_1935) and the genes for BuOH-2 (*thl*; *hbd*; *crt*; *ter*; *adhE*, Teth514_0627; *adhA*, Teth514_0564) were assembled into artificial operons (see **Figure 3.1b**). The *ter* gene was codon optimized for *P. furiosus* (Genewiz, Inc., South Plainfield, NJ) and the sequence is listed in **Table 3.1**. The operon was driven by the promoter of the gene encoding the gamma subunit of pyruvate oxidoreductase (175 bp immediately upstream of PF0971, including its ribosomal binding site) and terminated with 19 bases of the 3' UTR of the *hpyA1* gene (5'- aatctttttttagcactttt, locus PF1722). *P. furiosus* ribosomal binding sites (RBSs)

consisting of 12–14 bp of sequence upstream of genes encoding highly expressed proteins were placed upstream of each internal gene using primer tails as follows: the RBS for highly expressed S-layer protein (5'-ggaggtggagaaaa) was placed in front of *ter* and *hbd* and the RBS for the highly expressed PEP synthase (5'-ggaggtttgaag) was placed in front of and *bdh* or *adhA*, *thl* and *crt*. Gibson Assembly (200) was used to assemble the BuOH expression constructs into the vector pGL050 containing the *pyrF* genetic marker and flanking regions to delete ACSIα (PF1540) (201), generating pGL053 and pGL056 (for BuOH-1 and BuOH-2, respectively **Figure 3.2**). Plasmids were sequence-verified and pGL056 was found to contain a single point mutation in the RBS in front of *adhA*, which may be inconsequential. Plasmids were linearized and used to transform *P. furiosus* COM1 according to previously described methods (137). Transformants were colony-purified and strains were designated as follows: MW164 (pGL053 transformant) and MW196 (pGL056 transformant). These are herein referred to as the BuOH-1 and BuOH-2 strains, respectively (**Table 3.2**).

Growth of *P. furiosus*

Strains were cultured as previously described in a sea-water based medium containing 5 g/L maltose and 5 g/L yeast extract, 0.5 µg/L riboflavin, and 20 µM uracil (172, 202). Cultures were grown at 98°C to ~1x10^8 cells/mL and then transferred to a water bath calibrated to 50 to 70°C, which was maintained for up to 72 h. For growth in a 15 L fermenter, the culture was sparged with 10% CO_2/90% N_2, stirred, and the pH was maintained at 6.8 by addition of 10% $NaHCO_3$. The cultures were harvested by centrifugation at 6,000 x *g* for 10 min. Cell extracts were prepared anaerobically by sonicating the cell pellets in 100 mM MOPS, pH 7.5,

ultracentrifugation at 100,000 x *g* for 60 minutes, washing and re-concentrating three-times with a 10 kDa centrifugation filter and storing in sealed vials at -80°C.

RNA extraction and quantitative RT-PCR analyses

RNA was extracted from 10 mL samples of culture using the Direct-zol™ RNA MiniPrep kit (Zymo Research). Genomic DNA was digested using TURBO DNase (Ambion) and further purified by phenol:chloroform extraction and precipitation. Synthesis of cDNA was performed with 1 µg purified RNA using the Affinity Script QPCR cDNA synthesis kit (Agilent). The Brilliant III SYBR® Green QPCR Master Mix (Agilent) was used for quantitative RT-PCR experiments using an MX3000P instrument (Stratagene). Primers were designed to amplify a ~150 b product within the target gene, including PF0971 (*P. furiosus* PORγ control gene), *bad*, *bdh*, *adhE*, *adhA*, codon-optimized *ter*, *thl*, *hbd*, and *crt*.

NADH dependent enzyme assays

All reactions were carried out in sealed anaerobic cuvettes at 60°C (unless otherwise indicated) containing 25 mM MOPS pH 7.5. Extract was added to a final concentration of 0.2 mg/mL and NADH to A_{340} ~ 1.0. After addition of the relevant substrate, NADH oxidation was measured at 340 nm. The relevant substrates for the Thl+Hbd, Hbd, Ter, Bad, and Bdh assays are acetyl-CoA, acetoacetyl-CoA, crotonyl-CoA, butyryl-CoA, and butyraldehyde respectively.

Butanol dependent enzyme assays

All reactions were carried out in sealed anaerobic vials at 60°C (unless otherwise indicated) containing 100 mM MOPS pH 7.5 and 5 mM $MgCl_2$. Extract was added to a final concentration of 2.0 mg/mL. NADH and the indicated substrates were added to a butanol theoretical yield of 1 mM. Acetyl-CoA was added to assay the entire synthetic pathway, butyryl-CoA was added to assay Bad+Bdh, and butyraldehyde was added to assay Bdh. Samples were acidified to 8% formic acid and analyzed via gas chromatography flame ionization detection.

***In vivo* butanol production**

P. furiosus strains COM1, MW164, and MW196 were grown in 1 L cultures in sealed bottles at 98°C until the cell density reached $8x10^7$ cells/mL. They were then incubated at 60°C for 72 h for enzyme expression. After the enzyme expression, the cultures were harvested and suspended in sea-water based salts with the addition of 100 mM MOPS (pH 7.5) and 50 g/L maltose. Cells from each 1 L culture were suspended in 5 mL so that the cell density was concentrated to ~ 200x the original culture density. The suspensions were then incubated at 60°C for product formation. Aliquots of 250 µL were centrifuged, and the supernatants were acidified to pH 2 using formic acid and analyzed via gas chromatography flame ionization detection. Maltose consumption was estimated by the production of acetate, ethanol, and butanol in 1:4, 1:4, and 1:2 stoichiometric ratios respectively.

Results and Discussion

Design of the hybrid synthetic pathways

P. furiosus is a genetically tractable marine anaerobe that has an optimal growth temperature (T_{opt}) of 100°C and ferments carbohydrates and peptides (102, 137). The highest temperature at which engineered butanol formation has been reported is 55°C (195, 196). To increase this temperature limit for butanol production in *P. furiosus*, genes were selected from three different organisms to assemble an artificial pathway that combines steps from a butyrate-forming fermentation pathway with aldehyde and alcohol dehydrogenases (**Figure 3.1a**). The first three enzymes of the pathway convert acetyl-CoA to crotonyl-CoA and are derived from the fermentative bacterium *Thermoanaerobacter tengcongensis*, (T_{opt} 75°C) (203). The reduction of crotonyl-CoA to butyryl-CoA is carried out by the enzyme trans-2-enoyl-CoA reductase (Ter). *T. tengcongensis* contains a homolog of this enzyme, but it is a complex electron-bifurcating enzyme that requires two equivalents of NADH and produces reduced ferredoxin in addition to reducing crotonyl-CoA (170, 204). We, therefore, chose the enzyme Ter from *Spirochaeta thermophila*, a fermentative bacterium that grows optimally at 65°C (205), as it requires only one equivalent of NADH as substrate and has the added benefit of being irreversible (170). For the conversion of butyryl-CoA to butanol we chose to follow two strategies. For one pathway, butyraldehyde dehydrogenase (Bad) and butanol dehydrogenase (Bdh) were selected from *Thermoanaerobacter* sp. X514 (T_{opt} 65°C) (206). These are both previously uncharacterized from this organism but together should reduce the CoA derivative to the alcohol, equivalent to the bi-functional enzyme alcohol dehydrogenase E (AdhE) (207).

Bad and Bdh have homology to the two domains of AdhE, which are an aldehyde dehydrogenase and an iron containing alcohol dehydrogenase. The pathway containing Bad and Bdh is designated as the BuOH-1 pathway. For comparison, we also constructed a pathway, termed BuOH-2, containing the bi-functional AdhE along with alcohol dehydrogenase A (AdhA), both from *Thermoanaerobacter* sp. X514 (**Figure 3.1a**).

The pathway genes were assembled into synthetic operons, driven by the strong constitutive promoter of the gene encoding the γ subunit of pyruvate oxidoreductase (POR) of *P. furiosus* (**Figure 3.1b**). Ribosomal binding sites of highly expressed *P. furiosus* proteins were included between each gene for optimal protein production, along with a terminator sequence from *P. furiosus* at the end of the operon. A codon-optimized version of the Ter gene was used because the GC content of *S. thermophila* (61.9%) was significantly higher than *P. furiosus* (40.8%).

Modifying native *P. furiosus* metabolism

The butanol pathway expression constructs were inserted into the *P. furiosus* chromosome to simultaneously delete the alpha subunit of acetyl-CoA synthetase I (ACSI), generating the BuOH-1 and BuOH-2 strains (**Table 3.2**). ACSI is one of two major enzymes responsible for conversion of acetyl-CoA to acetate (119); therefore, the BuOH strains are designed to redirect acetyl-CoA away from acetate production and into the butanol pathway. While the effect of this knockout on butanol production was not determined here, a knockout of ACSIα was previously shown to improve production of 3-hydroxypropionate in an engineered

P. furiosus strain (201). This is an analogous strategy as both the butanol and 3-hydroxypropionate pathways utilize acetyl-CoA as their primary substrate.

Expression strategy for the fuel pathway

We performed quantitative PCR to determine expression levels of the BuOH-1 and BuOH-2 pathways at 98, 65 and 60°C (**Figure 3.3**). The expression levels of the genes in the synthetic operons are similar between the two strains, and they are within an order of magnitude of the native host gene encoding the gamma subunit of the glycolytic enzyme pyruvate ferredoxin oxidoreductase, whose promoter was used to drive expression of the pathways.

Expression of active heterologous enzymes that yield a functional butanol pathway in *P. furiosus* is dependent on growth temperature as the genes are derived from organisms with lower optimal growth temperatures than the host. At 98°C they are all expected to be inactive, hence growth of the *P. furiosus* BuOH-1 strain at this temperature results in wild type metabolism and the production of acetate (via other ACS homologs), carbon dioxide and hydrogen from sugar (the disaccharide, maltose). A shift to a lower temperature would lead to the production of active forms of the butanol pathway enzymes that together should divert acetyl-CoA to butanol. The net result is one molecule of glucose yielding two molecules of carbon dioxide and one molecule of butanol (**Figure 3.4**).

***In vitro* characterization of BuOH-1 and BuOH-2**

The butanol pathways were assembled from three different gene donors with optimal growth temperatures ranging from 60-75°C. To determine if active forms of the heterologously-

expressed enzymes were produced in *P. furiosus,* the BuOH-1 strain was grown to mid-log phase at 98°C, near the optimal temperature of the host, and then shifted to temperatures ranging from 50°C to 70°C. The cytosolic extracts were assayed to determine if butanol could be produced *in vitro* at 60°C, the lowest optimal growth temperature of the three gene donors. As shown in **Figure 3.5a**, butanol was produced from extracts of cells shifted to all temperatures tested for BuOH-1, but activity was extremely low in BuOH-2. The highest specific activity for butanol formation was in the 60 to 65°C range. To compare C4 specificity in the BuOH-1 versus BuOH-2 strain (**Figure 3.1b**), cytosolic extracts from both strains were prepared following heterologous expression at 60°C. Along with parental strain COM1, cell-extracts were prepared and were assayed at 60°C for NADH dependent activity on crotonyl-CoA (Ter), butyryl-CoA (Bad), and butyraldehyde (Bdh) (**Figure 3.5b**). While the Ter activities for both recombinant strains are nearly identical, extracts of the strain BuOH-1 shows higher activity on the C-4 substrates, particularly in the final step that converts butyraldehyde to butanol. Hence, in the following we focus on *P. furiosus* strain BuOH-1 for more in depth characterization.

Specific activities were determined for individual enzymes or sets of enzymes in the BuOH-1 strain at various assay temperatures using cell extracts prepared from cells shifted to 60°C (**Figure 3.6a-f**). In *P. furiosus* extracts, the Thl and Hbd enzymes from *T. tengcongensis* (**Figure 3.6a-b**) exhibited maximal activity at or above 80°C, which is not surprising as this is the gene donor with the highest optimal growth temperature (75°C). Temperatures above 80°C were not assayed because of the instability of the substrate. The Ter enzyme showed maximal activity at 60°C (**Figure 3.6c**), consistent with the optimal growth temperature of *S. thermophila*

(60-65°C). The Bad and Bdh enzymes exhibited broader temperature ranges for activity, and background Bdh activity was observed from the *P. furiosus* parent strain at the higher assay temperatures (**Figure 3.6d-e**). For the complete pathway, the *in vitro* conversion of acetyl-CoA and NADH to butanol was optimal at 60°C (**Figure 3.6f**), and this is most likely limited by the optimal temperature for activity of either the Ter or Bdh enzyme.

***In vivo* demonstration of BuOH-1 and BuOH-2**

To demonstrate *in vivo* butanol production in *P. furiosus* and to further compare the use of Bad and Bdh in the BuOH-1 strain versus AdhE and AdhA in the BuOH-2 strain, concentrated cell suspensions ($\geq 10^{10}$ cells/mL) of both strains, along with the parental strain COM1, were prepared in the presence of maltose and incubated at 60°C. The cell suspensions were analyzed for butanol, acetate, and ethanol at various time points up to 48 h (**Figure 3.7**). In strain BuOH-1 extracellular butanol (**Figure 3.7a**) production reached a maximum of ~70 mg/L after 48 h while consuming approximately 1500 mg of maltose which corresponds to 5% of the maximal theoretical yield (**Figure 3.7b**). The strain BuOH-2 had much lower productivity and efficiency as it produced 15 mg/L (**Figure 3.7a**) after 48 hrs which corresponds to 1% of the maximal theoretical yield (**Figure 3.7b**). Additionally, the native *P. furiosus* products ethanol and acetate were analyzed (**Figure 3.7c-d**). While no difference was observed in acetate formation, in the AdhE and AdhA containing strain BuOH-2, ethanol production was tripled from 500 mg/L to 1500 mg/L. The strain BuOH-1 showed little to no increase in ethanol production compared to the parental strain COM-1. This, along with the *in vitro* data presented in figure 2, demonstrates that the use of Bad and Bdh for butanol production is superior to the use of AdhE and AdhA.

Conclusion

We have demonstrated butanol production in a hyperthermophilic archaeon, using an artificial pathway assembled from three gene donors having moderate differences in growth temperature optima. Furthermore, the BuOH-1 pathway can utilize acetyl-CoA from the host metabolism, which is operating at close to 40°C below its optimal temperature. A fuel module such as this can be combined with a more advanced strategy in which H_2 and CO_2 can be used to supply acetyl-CoA in place of host sugar metabolism (71, 192). The module may also be used in other genetically tractable thermophilic microbes, in particular, those that have slightly lower growth temperature optima than *P. furiosus* and may be more suited for expression of a pathway at 60°C (208), or ones that are capable of degrading plant biomass (209). Such a strategy would couple the artificial fuel pathway to the native metabolism of the host and would have applications for consolidated bioprocessing of renewable biomass to a readily useable liquid transportation fuel.

Acknowledgements

This work was supported by the US Department of Energy as part of the Electrofuels project of ARPA-E (grant number DE-AR0000081).

Author Contributions

R.M.K. and M.W.W.A. conceived of and managed the research; A.J.L. and G.L.L. performed operon design and construction, DNA cloning, and transformation of *P. furiosus*; M.W.K. and G.J.S. performed physiological and enzymatic analyses; M.W.K., G.L.L., R.M.K. and M.W.W.A. wrote the manuscript.

Tables and Figures

Table 3.1 Sequence details for the GC and codon optimized *ter* gene

Sequence of the optimized *ter* gene (47.6% GC) from the gene donor *Spirochaeta thermophila* (61.9% GC)
atggttattaagcctatgataaggaacaacatttgcatgaacgcacaccctgagggatgcaggcttttcgtagacaggcagatagactat gtaaggcagaggggtcctcttagcagcggaccaaagaacgtacttgttgtaggtagcagcggaggatacggacttgctactaggatagc agctggattcggagcaggagcagcaacaattggtgtttacttcgagaaggagccaagcgaaaagaggactggaactccaggatggtac gcagcaaggcacttccttaagagggcagaggcagcaggacttaaggttcacgacataaacggagatgcattcagcttcgaggttaagg agcaggttgttgaccttattaagaaggagtacggaactatagacctttcgtttacagtcttgcaagcggtgttaggactgaccctgttagc ggagtaacatacaggagcgcattaaagcctataggaaggacatatactagcaaggcattagacccacttaaaagggaggtaatacaag ctactattgagccagcaagcgaggaggagatagaggctacagttaaggtaatgggtggagaggactggcaactttgggtagaggctctt aaggagggaggagttcttgctgagggagctgttacagtagcatacagctacattggaccagagttaacaaggcctatatacagagaag gtacaataggtaaagctaaggaacaccttgaggcaacagcaaagacacttgacgaggcacttgcaagcgttggtggaaaggcttacgt tagcgttaacaaggctgttgtaactagggcaagcgcagtaataccagcagttcctctttaccttagcattcttttcacagttatgaaggcta agggtattcacgagggttgcatagaacaaatgtacagaatgcttgcagagaggctttatgcaggaggtcctgttccagtagacgacgag ggtaggataaggatggacgactgggagatgagggaggacgttcagagggaagtagcagaggtttgggagaggataacaacagagac acttgacagggaggcagacgttgacggttaccttacagactttcttaacttacatggattcggttttaaggagatagactacgagaagga ggttgacccaagggagggatga

Table 3.2 Strains used and constructed in the study of butanol production in *P. furiosus*

Strain	Alias	Parent	Genotype/Description	Reference
MW002	COM1	DSM 3638	Δ*pyrF*	(137)
MW164	BuOH-1	MW002	Δ*pyrF* P_{gdh} *pyrF* P_{porG} *bad bdh ter thl hbd crt* Δ*acsA1*	This work
MW196	BuOH-2	MW002	Δ*pyrF* P_{gdh} *pyrF* P_{porG} *adhE adhA ter thl hbd crt* Δ*acsA1*	This work

Figure 3.1 The hybrid synthetic butanol pathways

(**a**) Diagram of the BuOH-1 and BuOH-2 pathways which use acetyl-CoA generated from the host's glycolytic pathway along with NADH to produce butanol. The pathways differ following the production of butyryl-CoA by the enzyme Ter. (**b**) Schematic diagram of the six gene synthetic operons for expressing the BuOH-1 and BuOH-2 pathways, under control of the *P. furiosus* PORγ promoter (bent arrow).

Figure 3.1

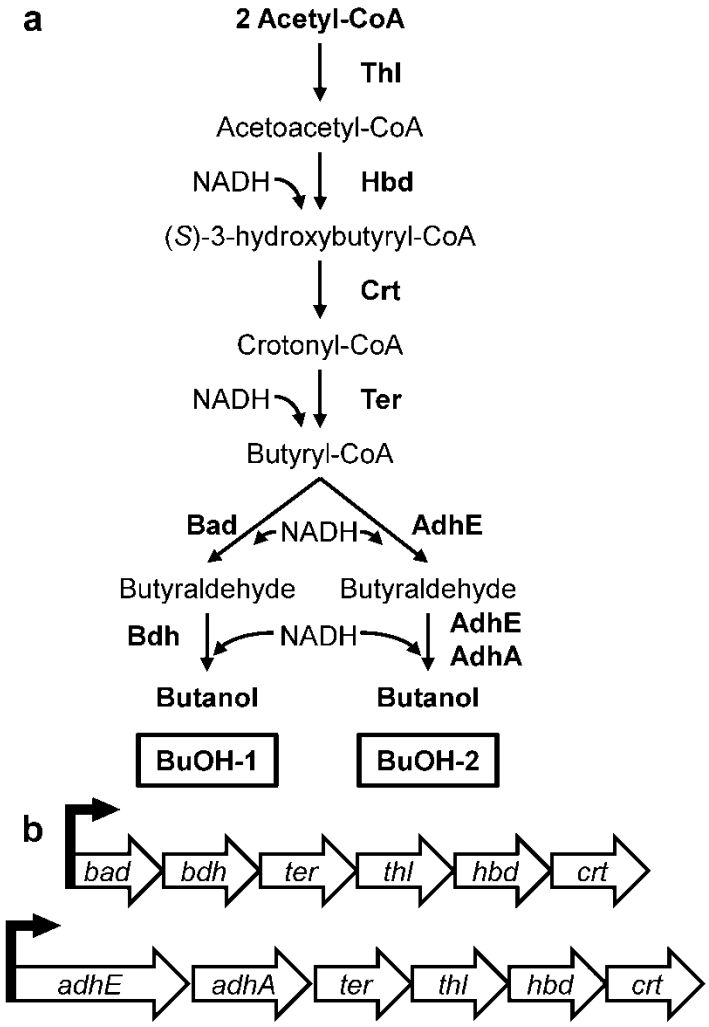

a
2 Acetyl-CoA
Thl
Acetoacetyl-CoA
NADH
Hbd
(S)-3-hydroxybutyryl-CoA
Crt
Crotonyl-CoA
NADH
Ter
Butyryl-CoA
Bad
NADH
AdhE
Butyraldehyde
Butyraldehyde
Bdh
NADH
AdhE
AdhA
Butanol
Butanol
BuOH-1
BuOH-2
b
bad
bdh
ter
thl
hbd
crt
adhE
adhA
ter
thl
hbd
crt

Figure 3.2 Plasmids used to generate the BuOH-1 and BuOH-2 strains

Plasmids used to generate BuOH-1 (pGL056) and BuOH-2 (pGL053) strains, with BuOH pathway genes (red), promoter (yellow), *P. furiosus* RBS's (purple), terminator (orange), P_{gdh}*pyrF* selectable marker for *P. furiosus* transformation (blue), homologous recombination regions for insertion into *P. furiosus* chromosome to simultaneously delete the gene encoding ACSIα (lavender), and *E. coli* plasmid components including pSC101 origin and apramycin resistance cassette (grey). Plasmids were linearized prior to transformation of *P. furiosus*.

Figure 3.2

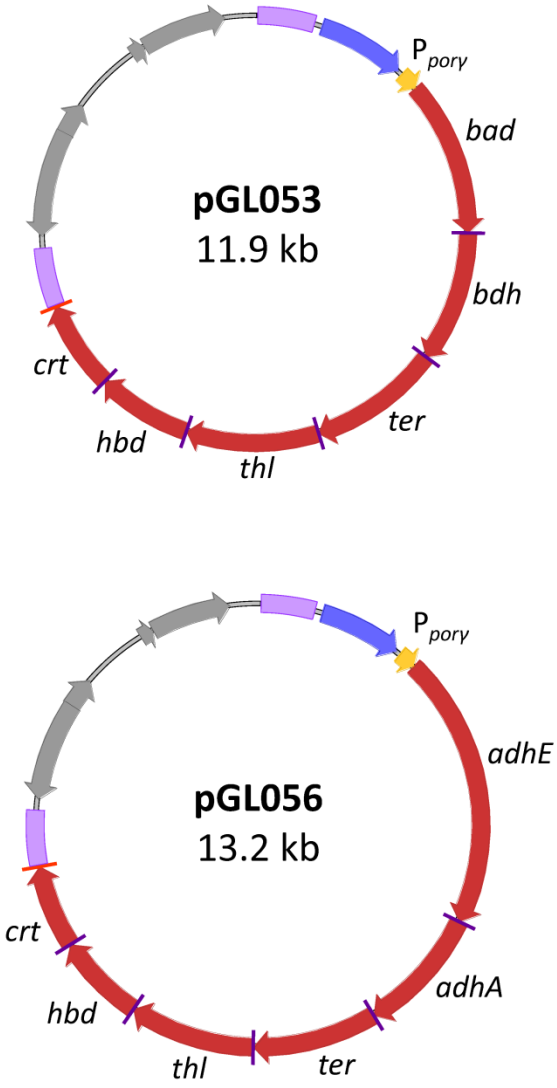

P porγ
bad
bdh
ter
thl
hbd
crt
pGL053
11.9 kb
P porγ
adhE
adhA
ter
thl
hbd
crt
pGL056
13.2 kb

Figure 3.3 Quantitative PCR of BuOH-1 and BuOH-2 genes

Quantitative PCR data showing expression levels of the BuOH pathway genes at various growth temperatures, relative to the expression level of the control gene encoding the gamma subunit of the glycolytic enzyme pyruvate ferredoxin oxidoreductase.

Figure 3.3

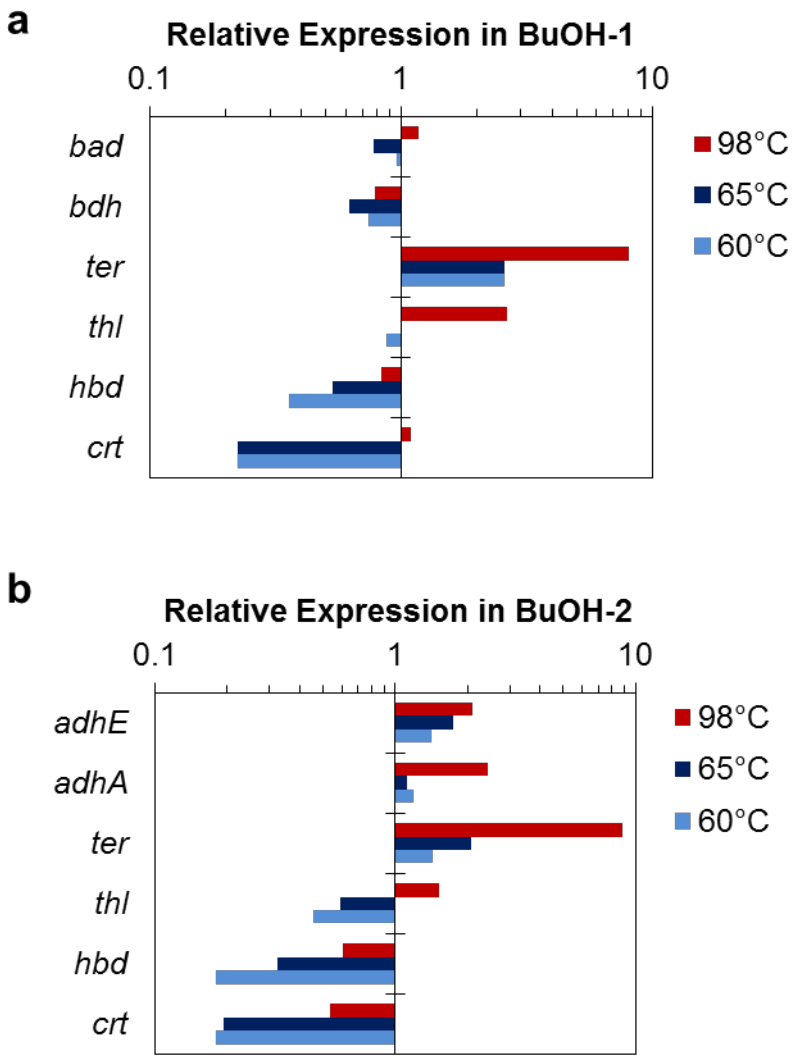

a
Relative Expression in BuOH-1
0.1
1
10
bad
bdh
ter
thl
hbd
crt
98°C
65°C
60°C
b
Relative Expression in BuOH-2
0.1
1
10
adhE
adhA
ter
thl
hbd
crt
98°C
65°C
60°C

Figure 3.4 Carbohydrate metabolism of *P. furiosus* to butanol

Maltose metabolism by *P. furiosus* is displayed, showing the key roles of glyceraldehyde-3-phosphate oxidoreductase (GAPOR), pyruvate ferredoxin oxidoreductase (POR), and acetyl-CoA synthetase (ACS) in acetyl-CoA production and of the membrane-bound hydrogenase (MBH) in H_2 production. Upon a shift to lower temperature, the artificial pathway will utilize cellular acetyl-CoA for butanol production.

Figure 3.4

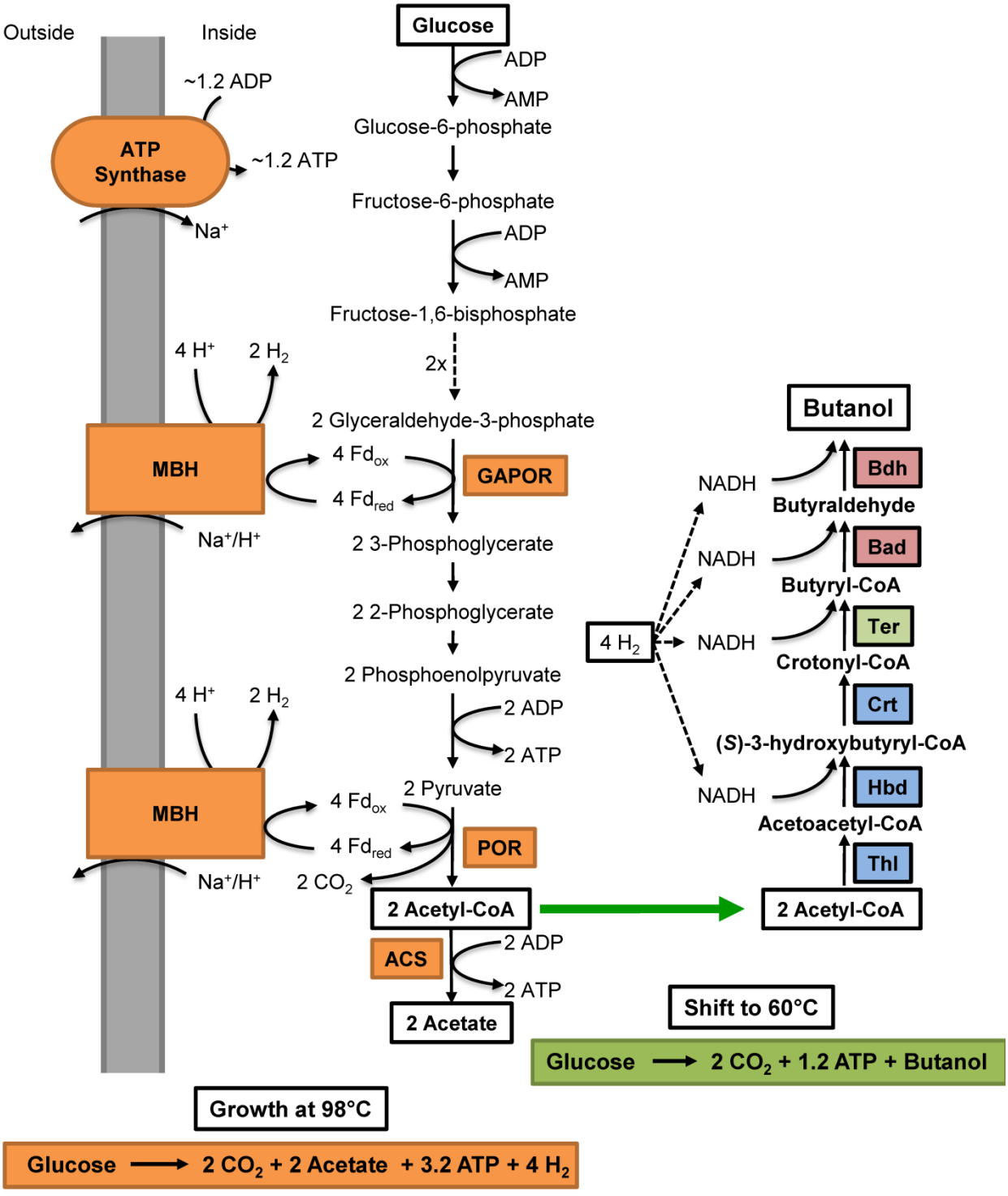

Outside
Inside
~1.2 ADP
ATP Synthase
~1.2 ATP
Na+
4 H+
2 H2
MBH
Na+/H+
4 Fdox
4 Fdred
Glucose
ADP
AMP
Glucose-6-phosphate
Fructose-6-phosphate
ADP
AMP
Fructose-1,6-bisphosphate
2x
2 Glyceraldehyde-3-phosphate
GAPOR
2 3-Phosphoglycerate
2 2-Phosphoglycerate
2 Phosphoenolpyruvate
2 ADP
2 ATP
2 Pyruvate
POR
2 CO2
2 Acetyl-CoA
ACS
2 ADP
2 ATP
2 Acetate
Growth at 98°C
Glucose → 2 CO2 + 2 Acetate + 3.2 ATP + 4 H2
4 H2
NADH
Butanol
Bdh
Butyraldehyde
Bad
Butyryl-CoA
Ter
Crotonyl-CoA
Crt
(S)-3-hydroxybutyryl-CoA
Hbd
Acetoacetyl-CoA
Thl
2 Acetyl-CoA
Shift to 60°C
Glucose → 2 CO2 + 1.2 ATP + Butanol

Figure 3.5 The optimal temperature for the expression of the BuOH-1 pathway and a comparison of the BuOH-1 and BuOH-2 pathways

(**a**) The rate of butanol formation from acetyl-CoA and NADH (nmol butanol/min/mg) measured in cytosolic extracts prepared from BuOH-1 (blue) and BuOH-2 (red) grown to $1x10^8$ cells/mL at 98°C and shifted to the indicated temperature and incubated for 72 hours. (**b**) The NADH dependent activity (µmol/min/mg) on crotonyl-CoA (Ter), butyryl-CoA (Bad), and butyraldehyde (Bdh) are compared in the strains BuOH-1 (blue), BuOH-2 (red), and the parental strain COM1 (orange).

Figure 3.5

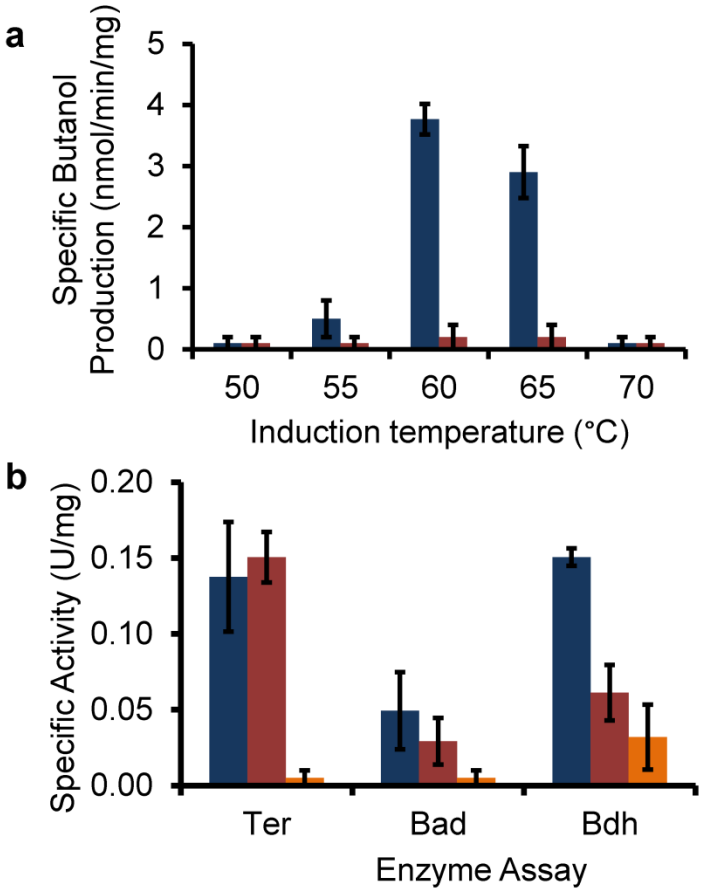

a
Specific Butanol Production (nmol/min/mg)
5
4
3
2
1
0
50
55
60
65
70
Induction temperature (°C)
b
Specific Activity (U/mg)
0.20
0.15
0.10
0.05
0.00
Ter
Bad
Bdh
Enzyme Assay

Figure 3.6 Optimal temperatures for activities of the BuOH-1 pathway enzymes

(**a-e**) Enzyme activities for pathway enzymes and reactions (shown in **Figure 3.1a**) in the BuOH-1 strain (blue) compared to the COM1 control strain (orange), assayed at the indicated temperatures. The strains were grown at 98°C to $1x10^8$ cells/mL and incubated at 60°C for 72 h in a 15L pH controlled fermenter. The individual enzyme rates (µmol NADH oxidation/min/mg) were measured by the addition of acetyl-CoA for Thl+Hbd (**a**), acetoacetyl-CoA for Hbd (**b**), crotonyl-CoA for Ter (**c**), butyryl-CoA for Bad (**d**), and butyraldehyde for Bdh (**e**). (**f**) Specific butanol production from acetyl-CoA and NADH (nmol butanol/min/mg) for the complete pathway was measured in cytosolic extracts at the indicated temperatures.

Figure 3.6

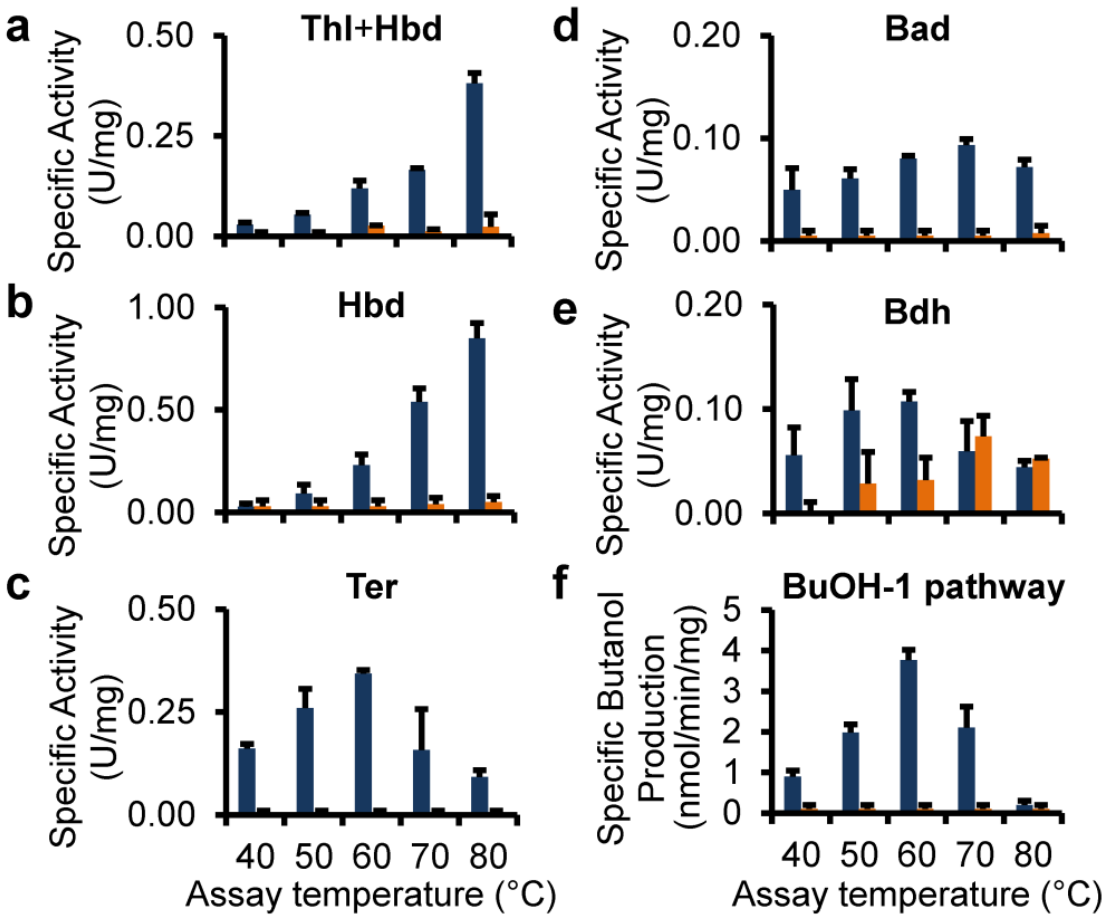
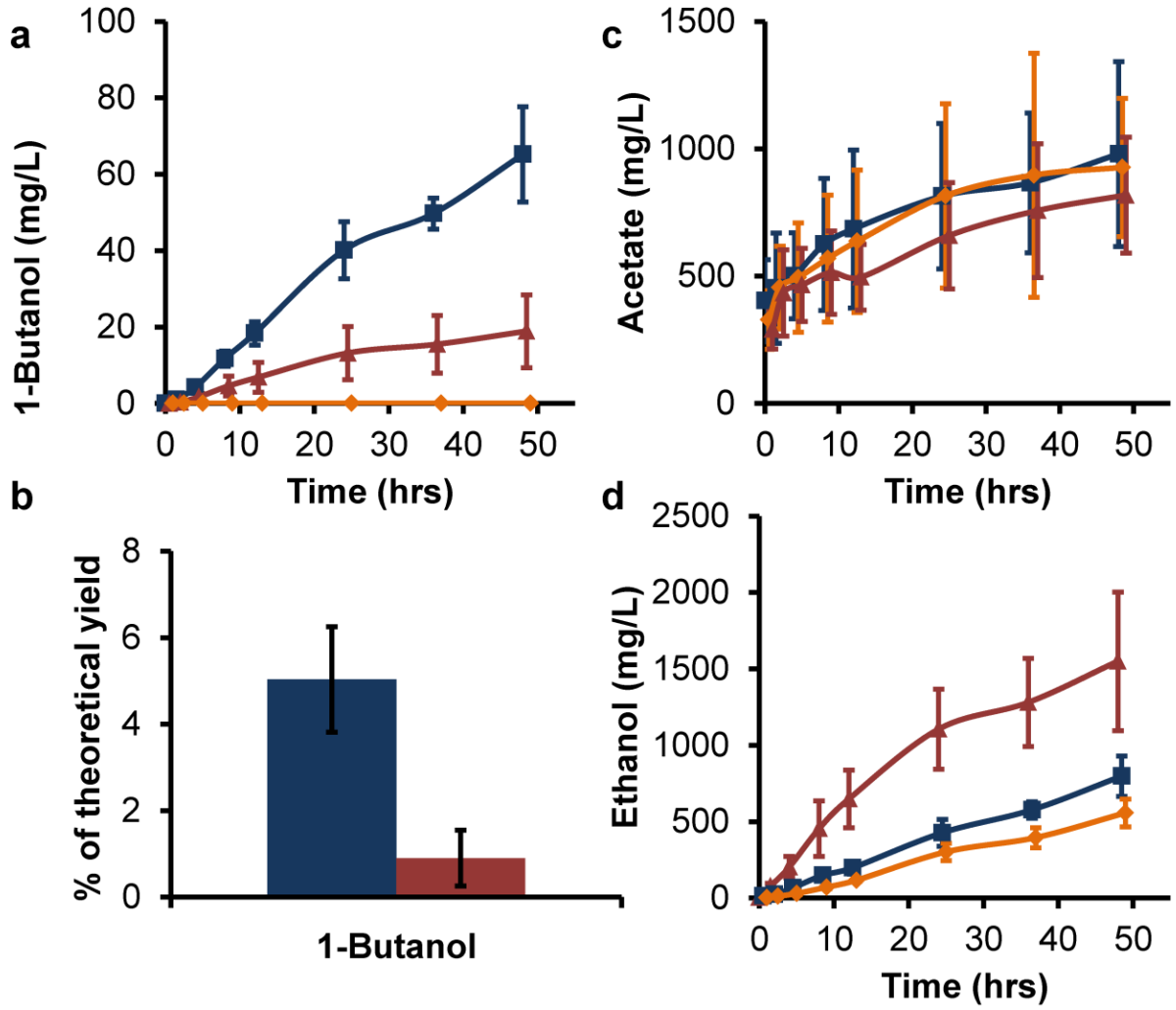

a
Thl+Hbd
Specific Activity (U/mg)
0.50
0.25
0.00
b
Hbd
1.00
0.50
0.00
c
Ter
0.50
0.25
0.00
40 50 60 70 80
Assay temperature (°C)
d
Bad
0.20
0.10
0.00
e
Bdh
0.20
0.10
0.00
f
BuOH-1 pathway
Specific Butanol Production (nmol/min/mg)
5
4
3
2
1
0
40 50 60 70 80
Assay temperature (°C)

Figure 3.7 *In vivo* butanol, acetate, and ethanol production from maltose in BuOH-1, BuOH-2, and COM1 strains in concentrated cell suspensions

In vivo butanol (**a-b**), acetate (**c**), and ethanol (**d**) production from maltose in BuOH-1 (blue), BuOH-2 (red), and COM1 (orange) strains in concentrated cell suspensions. The strains were grown at 98°C and then incubated at 60°C for 72 h after which a 200x concentrated cell suspension ($\geq 10^{10}$ cells/mL) was prepared in the presence of maltose and incubated at 60°C to assess product formation.

Figure 3.7


a
1-Butanol (mg/L)
100
80
60
40
20
0
0 10 20 30 40 50
Time (hrs)
c
Acetate (mg/L)
1500
1000
500
0
0 10 20 30 40 50
Time (hrs)
b
% of theoretical yield
8
6
4
2
0
1-Butanol
d
Ethanol (mg/L)
2500
2000
1500
1000
500
0
0 10 20 30 40 50
Time (hrs)

CHAPTER 4

ETHANOL PRODUCTION FROM THE HYPERTHERMOPHILIC ARCHAEON *PYROCOCCUS FURIOSUS* BY EXPRESSION OF BACTERIAL BIFUNCTIONAL ALCOHOL DEHYDROGENASES

Keller MW†, Lipscomb GL†, Nguyen, DM, Crowley AT, Schut GJ, Scott IM, Kelly RM, Adams MWW (2016) Alternative routes for ethanol production in the hyperthermophile *Pyrococcus furiosus*. *Microbial biotechnology* to be submitted

†Authors contributed equally to this work

Summary

Ethanol is an important target for the renewable production of liquid transportation fuels. It can be produced biologically from two metabolic starting points: from pyruvate, via pyruvate decarboxylase, or from acetyl-CoA, by alcohol dehydrogenase E (AdhE). Thermophilic bacteria utilize AdhE, which is a bifunctional enzyme that contains both acetaldehyde dehydrogenase and alcohol dehydrogenase activities. Many of these organisms also contain a separate alcohol dehydrogenase (AdhA) that generates ethanol from acetaldehyde, although the role of AdhA in ethanol production is typically not clear. Since acetyl-CoA is a key central metabolite that can be generated from a wide range of substrates, AdhE can serve as a single gene fuel module to produce ethanol through primary metabolic pathways. The focus here is on the hyperthermophilic archaeon *Pyrococcus furiosus,* which grows by fermenting sugar to acetate, CO_2 and H_2. Previously, by the heterologous expression of *adhA* from a thermophilic bacterium, *P. furiosus* was shown to produce ethanol by a novel mechanism from acetate, mediated by AdhA and the native enzyme aldehyde oxidoreductase (AOR). In this study, the AOR gene was deleted from *P. furiosus* in order to evaluate ethanol production directly from acetyl-CoA by heterologous expression of the *adhE* gene from eight thermophilic bacteria. Only AdhEs, from two *Thermoanaerobacter* strains, showed significant activity in cell-free extracts of recombinant *P. furiosus* and supported ethanol production *in vivo.* In the AOR deletion background, the highest amount of ethanol (61% theoretical yield) was produced when *adhE* and *adhA* from *Thermoanaerobacter* were co-expressed.

Introduction

Currently, ethanol is by far the most common biofuel, and there are currently three metabolic starting points for biological ethanol production: pyruvate, acetyl-CoA and acetate. These metabolic intermediates are converted to acetaldehyde which can then be converted to ethanol via an alcohol dehydrogenase for the terminal step. The most common pathway for ethanol production in mesophilic microorganisms involves pyruvate decarboxylation to acetaldehyde and subsequent reduction to ethanol, via pyruvate decarboxylase (PDC) and an alcohol dehydrogenase (ADH) (210). PDC is widely found in eukaryotes and various bacteria, including obligate aerobes, such as species in the genera *Acetobacter* (211) and *Gluconacetobacter* (212), facultative anaerobes, some cyanobacteria, such as *Microcystis aeruginosa* (213), and nitrogen-fixing acidophiles such as *Beijerinckia indica* (214). There are no PDC homologs in the Archaea. However, pyruvate ferrerdoxin oxidorductase (POR)from the archaeon *Pyrococcus furiosus*, which uses the pyruvate produced by glycolysis, also has pyruvate decarboxylase activity as a side reaction *in vitro*, a transformation that likely occurs under physiological conditions, as shown in **Figure 4.1** (139).

Most ethanologenic bacteria that lack PDC utilize the two step reduction of acetyl-CoA to ethanol by the bifunctional aldehyde dehydrogenase (ALDH)/alcohol dehydrogenase (ADH), also known as alcohol dehydrogenase E (AdhE), an enzyme first described in *E. coli* (215). Acetyl-CoA is an attractive substrate for formation of fuels such as ethanol since it is a central metabolite (216) and can serve as a link between fuel production and a variety of substrate utilization pathways. The bifunctional AdhE is widely found across Bacteria and Eukarya, from

the thermophilic strictly anaerobic bacterium *Thermoanaerobacter ethanolicus* (217) to the green alga *Chlamydomonas reinhardtii* (218).

With the growing interest in thermophilic routes for biofuel production (194, 219), AdhE has played a central role in increasing the maximum temperature for metabolically-engineered ethanol production, since thermostable versions of pyruvate decarboxylase have not been discovered (220). PDC from *Zymomonas mobilis* has been used at a maximum temperature of 52°C via heterologous expression in the moderate thermophile *Geobacillus thermoglucosidasius*, which grows optimally at 62°C (221). In contrast, bacteria with either native or heterologous AdhE genes have been engineered for ethanol production at higher temperatures. For example, extensive mutational and overexpression studies have improved ethanol tolerance and production in *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum* at 55°C (222). In *G. thermoglucosidasius*, the deletion of lactate dehydrogenase (LDH), formate dehydrogenase (FDH), and overexpression of pyruvate dehydrogenase (PDH) directed carbon and electron flow to improve ethanol production at 60°C (193). Directed evolution in this same organism revealed that the disruption of adenine phosphoribosyltransferase also improves ethanol production (223). The overexpression of *Clostridium thermocellum* AdhE in a *Caldicellulosiruptor bescii* strain with a disrupted LDH resulted in ethanol production from plant biomass at 65°C (224). *Thermoanaerobacter mathranii* lactate dehydrogenase (LDH) was knocked out and its AdhE was overexpressed resulting in an improvement in ethanol production at 70°C (163, 164). *C. bescii* has also been

engineered with the AdhE from *Thermoanaerobacter pseudethanolicus* 39E, which grows optimally at 65°C (225), and this supported ethanol production at 75°C (226).

Curiously, and for reasons that are not clear, no members of the Archaea, thermophilic or otherwise, are known to contain either PDC or AdhE. However, a pathway for ethanol production from acetate was recently demonstrated in the thermophilic archaeon *P. furiosus* at 72°C by the heterologous expression of AdhA from *Thermoanaerobacter* sp. strain X514 (*T.* X514), which grows optimally at 60°C (1). As shown in **Figure 4.1**, the acetaldehyde substrate for AdhA is generated by the ferredoxin-dependent aldehyde oxidoreductase (AOR) of *P. furiosus.* This enzyme is thought to detoxify through oxidation the acetaldehyde produced as the side reaction of POR mentioned above (139). However, in the presence of AdhA, AOR catalyzes the reverse reaction, acetate reduction, and drives the AdhA-catalyzed reduction of acetaldehyde to ethanol. In *P. furiosus,* the acetate is generated from acetyl-CoA by acetyl-CoA synthetase (ACS), which also generates ATP (**Figure 4.1**).

Hence, in *P. furiosus,* a combination of ACS, AOR and AdhA, which will be referred to here as the AAA pathway, catalyzes the same overall reaction as AdhE, the conversion of acetyl-CoA to ethanol (**Figure 4.1**). However, there are some important differences. The three-step AAA pathway generates ATP and requires reduced ferredoxin (Fd_{red} for AOR) and NADPH (for AdhA) as electron donors, while the single step AdhE does not conserve energy and requires only NADH as the reductant. *P. furiosus* therefore affords a model system with which to compare the efficacy of these two very different pathways for acetyl-CoA to ethanol conversion.

In the original report of the AAA pathway (1), the *adhE* from *T.* X514 was heterologously expressed in *P. furiosus*, but interpreting the results are complicated by the presence of, and potential interference by, the endogenous AOR, which might compete for any acetaldehyde that was formed as an intermediate from the AdhE reaction. Hence, in the present study, a strain of *P. furiosus* was utilized in which the gene encoding AOR was deleted. Herein, AdhE genes from eight thermophilic bacteria for expression in a Δ*aor* strain of *P. furiosus.* The AdhE activities in cell extracts were examined as was the ability of the recombinant strains to produce ethanol. In addition, since the AdhE genes are from bacteria with optimal growth temperatures significantly lower than that of the *P. furiosus*, which grows optimally near 100°C, a temperature shift strategy was used for their expression. This strategy is similar to those previously used in heterologous pathway expression in *P. furiosus* demonstrating lactate production (172), partial CO_2 fixation and 3-hydroxypropionate production (192), ethanol production (1), butanol production (227), native acetoin formation (105), and energy-conserving carbon monoxide oxidation (228). The results here demonstrate *in vitro* AdhE-dependent catalytic activity and *in vivo* ethanol production for two out of eight thermophilic AdhE homologs expressed in *P. furiosus*.

Results

Selection of AdhE source organisms

P. furiosus is capable of growth at temperatures as low as 72°C, almost 30° below its optimal growth temperature (138). At temperatures below 72°C, *P. furiosus* metabolic processes continue, as shown in previous work in strains engineered for butanol formation which was demonstrated at temperatures as low as 60°C (227). In choosing AdhE-encoding genes for expression in *P. furiosus*, candidate organisms were selected based on growth temperature. The AdhE gene from *Clostridium thermocellum*, which grows optimally at 60°C, has been heterologously expressed in a thermophilic bacterium producing up to 15 mM ethanol at 65°C (224). We therefore used the *C. thermocellum* AdhE gene to perform a BLAST search (229) against the NCBI database, including the search term "therm*" to limit hits to thermophilic organisms (by definition, organisms that have optimal growth temperatures at 45°C and above). Of the approximately 60 resulting hits, a total of eight bacteria that had optimal growth temperatures of 60°C or higher were selected, including two representatives each of *Geobacillus* and *Thermoanaerobacter* species, along with *C. thermocellum* (**Table 4.1**). There were no AdhE homologs in any archaeon, regardless of growth temperatures. The eight AdhE enzymes had 46-63% sequence identity (67-80% similarity) to *E. coil* AdhE, with various degrees of sequence homology within the group (**Table 4.2**). Many of these organisms naturally produce ethanol, with *T. ethanolicus* having both the highest growth temperature optimum and the highest reported ethanol yield (**Figure 4.2**). The AdhE enzymes from some of these

organisms have been characterized *in vitro*, including recombinant forms of *Tx*AdhE (230) *Ct*AdhE (230) and *Te*AdhE (155), as well as natively purified *Gt*AdhE (207).

Strain construction and validation

There is a robust genetic system available for *P. furiosus*. This utilizes a single counter-selectable marker (*pyrF*) in a naturally competent genetic background strain having a deletion of the *pyrF* gene (COM1), allowing multiple deletions and/or insertions to be made in a single strain (137, 183). To facilitate analysis of ethanol produced exclusively as a result of heterologous expression of *adhE*, a background strain lacking *aor* was constructed in COM1 to eliminate ethanol production from acetaldehyde derived from acetate (1). Deletion of the AOR gene was previously demonstrated to reduce ethanol production in an AdhA-expressing strain to background levels (1). The deletion of AOR did not affect growth on maltose and resulted in an 80% reduction in acetaldehyde dependent benzyl viologen (BV) reduction (**Figure 4.3**). The remaining 20% of activity can be attributed to other AOR-family enzymes that also display some activity with acetaldehyde (231-233). The Δ*aor* genetic background was used for expression of AdhE constructs placed under control of the promoter of the highly expressed S-layer protein gene (P_{slp}) (192) and inserted between genes PF0738 and PF0739, a location referred to here as genome region 4. Eight of these contained the AdhE from each organism listed in **Table 1**. A strain containing the *T. ethanolicus* AdhE in a synthetic operon with AdhA was also constructed. The AdhA gene was amplified from *T.* X514; it should be noted that the amino acid sequence of *T*xAdhA is identical to that of *Te*AdhA. The previously published strain MW611, containing a deletion of *aor* in the AdhA expression strain (1), was also used for comparison.

The control strains used in growth experiments were the COM1c2 strain containing *pyrF* added back to COM1 to restore uracil prototrophy (234), and the ΔAORc strain containing a marker-replaced deletion of *aor*. The strains used as controls in the activity assays were the uracil auxotrophic versions of these strains, COM1 and ΔAOR. All *P. furiosus* strains used and constructed in this study are listed in **Table 4.3**.

In vivo ethanol production of AdhE-containing strains

As the AdhE enzymes were obtained from organisms with optimal growth temperatures from 60 to 70°C, and *P. furiosus* grows optimally at near 100°C, a strategy involving a temperature shift was used for growth to measure *in vivo* ethanol production. All strains were grown at 95°C to mid-log phase ($1x10^8$ cells mL^{-1}), and cultures were then shifted to 65°C and incubated for up to 90 h for ethanol production (**Figure 4.4A**). A low concentration of ethanol (≤1 mM) was consistently observed in the control strains COM1c2 and ΔAOR, consistent with previous results (1).

Of the bacterial strains expressing the various heterologous AdhE genes, only those containing AdhE and AdhA from *Thermoanaerobacter* sp. produced ethanol above background (**Figure 4.4A**). The Te-AdhEA strain containing both AdhE and AdhA produced the most ethanol (4.2 mM), followed by Te-AdhE (2.6 mM), AdhA (1.8 mM), and Tx-AdhE (1.5 mM). Acetoin has recently been reported as a major fermentation end product of *P. furiosus* when grown at 70-80°C (105); however, under these conditions at 65°C, no acetoin was detected. Therefore, ethanol and acetate are the only major carbon end products from glucose under these conditions. For these four strains, the amount of ethanol produced per glucose consumed was

increased from the background level (~0.3-0.4) to 1.2, 1.0, 0.8, and 0.7, respectively (**Figure 4.4B**).

Construction and analysis of strains expressing mutated AdhE genes

It was considered that nucleotide specificity of the ADH enzymes might play a role in ethanol yield *in vivo*, as AdhE is NADH-dependent. Under physiological conditions (where [NADPH]>[NADP]), the reduction potential of NADPH (approx. – 380 mV) is lower than that of NADH (~ - 280 mV, [NADH]<[NAD]; (235)). In *P. furiosus*, NADPH can be generated from fermentatively-produced H_2 via soluble hydrogenase I (124). There are a number of reports that identify mutations in AdhE that cause a change in nicotinamide nucleotide specificity. For example, a glycine to aspartic acid mutation in the NADH binding site of the ADH domain of AdhE from *Thermoanaerobacterium saccharolyticum* (G544D in strains LL1049 and LL1194) results in a switch from NADH to NADPH without any dramatic reduction in enzyme activity. Interestingly, this switch in specificity is observed for both ALDH and ADH activities even though the mutation is located in the ADH domain (222). The amino acid sequence of *T. saccharolyticum* AdhE is 86% and 87% identical (93% and 94% similar) to *Te*AdhE and *Tx*AdhE, respectively. Therefore, these mutations were introduced into both *Thermoanaerobacter* AdhE genes, and they were each expressed in a Δ*aor P. furiosus* background. However, these mutations did not result in NADPH-dependent ALDH or ADH activities in cell extracts; rather, they increased the NADH-dependent ALDH activities but abolished the NADH-dependent ADH activities (**Table 4.4**). Rather than improving ethanol yield, the mutations had the effect of reducing ethanol production to the background level (**Figure 4.5**).

In vitro ALDH and ADH activities in *P. furiosus* AdhE-containing strains

In order to determine the levels of activity of the various heterologously-expressed AdhE and AdhA enzymes in *P. furiosus*, cultures were grown to mid-log phase at 95°C and then switched to 65°C for 40 h. Cell-free extracts were assayed anaerobically at 65°C for ALDH activity (NADH or NADPH-dependent reduction of acetyl-CoA) and ADH activity (NADH or NADPH-dependent reduction of acetaldehyde).

AdhE is a bifunctional enzyme, containing both ALDH and ADH activities, both of which have been reported to be NADH-dependent in previously characterized AdhE enzymes (161). NADH-dependent ALDH activity was found in strains expressing the wild-type or mutated *Thermoanaerobacter* AdhE enzymes, and these values were comparable to those measured in cell extracts of the source organisms (**Table 4.4**) and were about an order of magnitude lower than those measured for purified recombinant *Te*AdhE (155). No NADPH-dependent ALDH activity was detected. For Tx-AdhE, Te-AdhE and Te-AdhEA strains, which contained wild-type AdhE enzymes, extracts contained NADH-specific ADH activity. Strains AdhA and Te-AdhEA also displayed high NADPH-dependent ADH activity due to the AdhA, which has been previously shown to be highly active and NADPH-specific (1). In the strains containing mutated *Thermoanaerobacter* AdhE enzymes (Tx-AdhE* and Te-AdhE*), the ADH activity was abolished. For the remaining AdhE-containing strains, no ALDH activity and very little if any ADH activity was detected (**Table 4.4**). For *P. furiosus* genetic background strains COM1 and ΔAOR, no activity (the limit of detection ~0.05 U/mg) could be detected for the following reactions: the reduction of acetyl-CoA via NAD(P)H, the reduction of acetaldehyde via NAD(P)H, or the

oxidation of ethanol via NADH or BV. However, the cell extracts of both strains contained 0.2 ± 0.05 U/mg of NADP-dependent ethanol oxidation at 80°C.

Discussion

P. furiosus metabolizes carbohydrates at temperatures near 100°C to primarily H_2, CO_2 and acetate (102, 104). Small amounts of ethanol are also produced natively at sub-optimal growth temperatures (1). There are three possible routes of either native or engineered ethanol production stemming from three different sources of acetaldehyde as an intermediate: pyruvate, acetyl-CoA and acetate (**Figure 4.1**). Previous work demonstrated that the native AOR enzyme can function in concert with a heterologously-expressed AdhA enzyme to drive ethanol production from acetate (1), referred to here as the AAA pathway. AOR is a complex metalloenzyme containing a tungsten cofactor (236), and this may limit its use to organisms with a tungsten cofactor biosynthesis pathway, such as has been shown in the thermophilic bacterium *Caldicellulosiruptor bescii* (237). Conversely, AdhE is a single gene fuel pathway that is readily transferable to alternate hosts and requires only a single electron donor (NADH). We show here that *P. furiosus* has the capability for ethanol production from acetyl-CoA catalyzed by AdhE (**Figure 4.1**).

We chose to express various AdhE enzymes in a *P. furiosus* background stain lacking AOR to examine ethanol production from acetyl-CoA alone. Eight AdhE genes from thermophilic source organisms were expressed in *P. furiosus*. Surprisingly, only two of these AdhE enzymes were functional despite the substantial sequence identity (47-97%, with 67-98% similarity) among all of the AdhE proteins (**Table 4.4**). Strains containing both

Thermoanaerobacter AdhE enzymes (strains Te-AdhE and Tx-AdhE) produced ethanol *in vivo* (**Figure 4.2**). As might be expected, the recombinant *P. furiosus* strains that produced significant amounts of ethanol were those that contained measurable NADH-dependent ALDH activity (**Table 4.4** and **Figure 4.4**). Expressing AdhA along with AdhE from *T. ethanolicus* (strain Te-AdhEA) resulted in the highest concentration of ethanol measured (1.2 mole ethanol produced per mole glucose consumed). The improvement in ethanol yield with Te-AdhEA compared to Te-AdhE indicates that the highly active AdhA augments the AdhE pathway by providing an NADPH specific route to ethanol from acetaldehyde. Specifically, the contribution of AdhA to this AdhE strain results in an increase of 1.6 mM in ethanol production (an additional 0.2 ethanol per glucose). Interestingly, the control strain expressing AdhA alone, with AOR deleted, resulted in similar ethanol production as the Tx-AdhE strain (an increase of ~0.7 mM over the background concentration), suggesting that AdhA is highly efficient for scavenging acetaldehyde. While it is clear that AdhA can produce ethanol from natively-generated acetaldehyde, it appears that AdhA also directly augments the AdhE pathway by utilizing acetaldehyde produced as an intermediate by AdhE, as is the case with ethanologenesis in *Thermoanaerobacter mathranii* (163). the ethanol-boosting effect of AdhA on AdhE is further supported by *in vitro* results showing that purified *Tx*AdhE has a high K_M for acetaldehyde (22,000 μM ± 4,800) (230). In a previous study, a similar effect was seen in an AOR-containing background where, at 72°C, a strain containing *Tx*AdhE produced 2 mM ethanol and a strain with both *Tx*AdhE and AdhA produced 10 mM ethanol (1). However, most of the ethanol in those strains was likely derived from acetate (via AOR), rather than from acetyl-CoA. We previously reported that growth of *P. furiosus* at suboptimal temperatures from 70-80°C

resulted in the production of acetoin, in addition to acetate, H_2 and CO_2 (105). However, we did not observe acetoin production in our strains at 65°C. Therefore, assuming that ethanol or acetate are the main carbon end-products of glycolysis, strain Te-AdhEA produced ethanol from glucose at 61% theoretical yield (**Figure 4.4B**).

The selected expression temperature (65°C) for the heterologous *P. furiosus* strains was within the temperature ranges of all of the source bacteria (Fig. 1). Yet, only the two *Thermoanaerobacter* AdhE enzymes were functional in *P. furiosus.* Although these were from the upper and lower end of the optimal growth temperatures (**Table 4.1**), the two enzymes share 97% sequence identity (**Table 4.2**). Clearly, growth temperature of the source organisms is not a good indicator of bio-catalytic function *in vivo* in *P. furiosus*. For example, the AdhA from *T.* X514, which has an optimal growth temperature of 60°C (238, 239), functions in the AOR-AdhA pathway at temperatures up to near 80°C in *P. furiosus* (1). Other factors that could influence heterologous enzyme function in *P. furiosus* include codon usage. Although the source organisms have similar GC content to that of *P. furiosus* (40.8%; **Table 4.1**), slight differences in codon usage, particularly in higher GC-content organisms could affect enzyme expression. In addition, all of the bacteria were isolated from terrestrial locations while *P. furiosus* is a marine organism that is known to accumulate a range of intracellular solutes (240); this factor might affect the functionality of the heterologously-produced enzymes.

The specific activities of the various AdhE enzymes in cell extracts of the source bacteria also does not necessarily indicate whether they will function in *P. furiosus*. In the case of the *Tx*AdhE, the *in vitro* activity in the cell extract of the *P. furiosus* strain Tx-AdhE was similar to

that in *T.* X514 cell extract. However, this was not the case for *C. thermocellum*; we show that it contains high AdhE activity, but no activity could be detected in cell extracts of *P. furiosus* strain Ct-AdhE (**Table 4.4**), nor was any ethanol produced by this strain *in vivo* (above background). Natural ethanol production by the host organisms also does not indicate whether the AdhE will function in *P. furiosus* since both *C. thermocellum* and *G. thermoglucosidasius* produce significant amounts of ethanol, but no ethanol above background was observed when their enzymes were expressed in *P. furiosus* (**Figure 4.4**). The metabolic context of the host organism is clearly important, since even though heterologous expression of *C. thermocellum* AdhE in *C. bescii* resulted in ~10 mM ethanol produced at 65°C (224), this AdhE did not lead to increased ethanol production in *P. furiosus* at that temperature. In Clostridiales, AdhE often appears in a conserved unit along with a kinase and phosphatase (241). Co-expression of these enzymes may be necessary for any post-translational modification that may be required for the activation of AdhE. Furthermore, AdhE has been reported to form large multimeric complexes termed spirosomes, which are thought to enhance stability or facilitate channeling of the acetaldehyde intermediate (161, 207, 242), and expression of these enzymes outside their native context may disrupt formation of these structures.

The highest amount of ethanol produced by any of the *P. furiosus* strains was 1.22 mole of ethanol per mole of glucose consumed, which is a 61% theoretical yield from glucose. This may be limited by cofactor specificity, particularly considering the additive effect of the NADPH-dependent AdhA on ethanol production. AdhE is NADH-specific, but *P. furiosus* does not utilize a NADH-producing glyceraldehyde-3-phosphate (GAP) dehydrogenase in glycolysis; rather its

GAP oxidoreductase uses ferredoxin as the electron acceptor (243). In order to analyze the effect of nucleotide cofactor specificity of the heterologous *Thermoanaerobacter* AdhE enzymes on ethanol production, a single mutation was made in the ADH domain NADH-binding site (**Figure 4.5A**). This mutation was previously reported to cause a complete shift in nucleotide cofactor specificity in *T. saccharolyticum* AdhE from NADH to NADPH (222). However, cell extracts of the *P. furiosus* strains displayed no NADPH-dependent ADH or ALDH activities (**Table 4.4**). Curiously, the NADH-dependent ALDH specific activities increased in the mutated AdhE* strains, but the NADH-dependent ADH activities were abolished (**Table 4.4**), and the ethanol produced in the resulting strains was reduced to background levels (**Figure 4.5B**). The glycine to aspartic acid mutation has also been investigated in the AdhE from *Lachnoclostridium phytofermentans* (f. *Clostridium phytofermentans*) (244), which bears 50% sequence identity to AdhE from *T. saccharolyticum* and 51% sequence identity to both *T. ethanolicus* and *T.* X514 AdhE, but the amino acid change resulted in a completely inactive enzyme where no NAD(P)(H)-dependent ALDH or ADH activity could be detected in either direction (245). These results suggest that the 86 and 87% sequence identities (93 and 94% similarity) of the *Thermoanaerobacter* AdhE proteins with *T. saccharolyticum* AdhE are not sufficient to effect a change in specificity, or the native context of *T. saccharolyticum* is important for the mutation to have the cofactor-switching effect, perhaps through post-translational modification.

At sub-optimal temperatures *P. furiosus* produces a small amount (~1 mM) of ethanol (1), which we also observed in this study (**Figure 4.4**). *P. furiosus* contains four genes annotated

as alcohol dehydrogenase, referred to here as *Pf*AdhA (PF0074), AdhB (PF0075), AdhC (PF0608), and AdhD (PF0991), and some of these have been characterized (126, 127). These ADHs are likely responsible for reducing natively produced acetaldehyde to ethanol (**Figure 4.1**). Interestingly, deletion of the AOR gene in *P. furiosus* strain COM1 did not abolish the ethanol that was produced at 65°C. Both the parent and deletion strains produced 0.3 mole ethanol per mole of acetate, or 23% of the theoretical yield from glucose (**Figure 4.4B**). Therefore although AOR plays little if any role in natural *P. furiosus* ethanol production, cell extracts of the AOR deletion strain contained ~20% of the dye-linked acetaldehyde oxidoreductase activity of the parent strain (**Figure 4.3**), indicating that other enzymes are present that could reduce acetate to acetaldehyde. These include the other members of the tungsten-containing AOR family that are present in *P. furiosus,* including formaldehyde ferredoxin oxidoreductase (FOR) (231) and two other partially characterized tungsten-containing oxidoreductases (WOR4 and WOR5) (232, 233). In addition, acetaldehyde is produced as a side reaction of pyruvate ferredoxin oxidoreductase (139), which converts pyruvate to acetyl-CoA (**Figure 4.1**). For example, in *in vitro* assays at 80°C (pH 8.0), purified POR non-oxidatively decarboxylates pyruvate to acetaldehyde at a rate that is approximately 40% of the oxidative decarboxylation of pyruvate to acetyl-CoA, and similar activity ratios were reported for the POR from a related species (139, 246). Furthermore, in the absence of *aor*, expression of AdhA alone results in a ~0.7 mM increase in ethanol production, confirming the presence of other acetaldehyde-producing enzymes in *P. furiosus*.

Although ethanol production from acetyl-CoA was demonstrated in *P. furiosus* in this study, the highest ethanol yield (from strain Te-AdhEA) is still lower than that of the previously reported AAA pathway in *P. furiosus* which functions via native enzymes ACS and AOR along with heterologously expressed AdhA. Ethanol production in strain MW608, expressing AdhA and containing in-tact AOR, resulted in ~1.8 mole ethanol per mole glucose which is approximately 90% theoretical yield, assuming ethanol and acetate are the only two carbon end-products of glycolysis (1). A distinct advantage of the AAA pathway is the ATP-generating step converting acetyl CoA to acetate. Furthermore, reducing equivalents for the pathway ultimately are derived from ferredoxin generated during glycolysis. Acetate to acetaldehyde conversion by AOR is ferredoxin-dependent, and AdhA utilizes NADPH which can be regenerated by SHI via uptake of H_2 generated by the ferredoxin-dependent membrane bound hydrogenase (124). The AAA pathway is therefore well integrated with *P. furiosus* glycolysis in terms of energy conservation and redox balance. Conversely, the AdhE pathway relies solely on NADH (unless AdhA is also expressed), and it is unclear if and how the pools of NADH and ferredoxin are interconverted in *P. furiosus*, although possible enzyme candidates are two so-called ferredoxin-NAD(P)H oxidoreductases (FNOR) (247). Furthermore, AdhE diverts acetyl-CoA away from the energy conserving ACS reaction.

At first glance, the AdhE enzymes appear to function in a very predictable manner, but it is becoming clear that the characteristics of these enzymes are very intricate, and engineering pathways that contain these enzymes to function in a heterologous host is not a straight-

forward task. A better understanding of how these enzymes work in their respective native contexts would benefit future efforts to utilize them for applied purposes.

Experimental Procedures

Selection of AdhE gene source organisms

The *C. thermocellum* AdhE gene was used to perform a BLAST search (229) against the NCBI database including the search term "therm*". Out of the approximately 60 resulting hits, seven organisms in addition to *C. thermocellum* were selected based on their growth temperature profile (**Table 4.1**). Organisms or genomic DNA were obtained from the Leibniz Institute DSMZ German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany) for the following strains: *T. ethanolicus* JW200 (DSM 2246), *G. thermoglucosidasius* NBRC 107763 (DSM 2542), *T. celere* (DSM 8682), *A. flavithermus* WK1 (DSM 21510), and *D. kuznetsovii* (DSM 6115). *T.* sp. strain X514 (ATCC BAA-938) was obtained from the American Type Culture Collection (Manassas, VA, USA). *G. stearothermophilus* NUB3621 (BGSC* 9A5) was obtained from the Bacillus Genetic Stock Center (Columbus, OH, USA). The *C. thermocellum* AdhE gene was PCR-amplified from *C. bescii* heterologous expression strain JWCB032 (224).

Plasmid construction and DNA manipulation

A linear transformation construct was prepared for deletion of the *P. furiosus* AOR gene (PF0346). Flanking regions of ~0.5 kb upstream and downstream of *aor* were PCR-amplified from *P. furiosus* genomic DNA and combined by splice-overlap extension PCR (SOE-PCR, (182)) with a PCR-amplified pop-out P_{gdh}*pyrF* marker cassette (192) positioned between the flanking regions.

A vector was constructed for the insertion of heterologous expression cassettes at a location identified here as "Genome Region 4", between convergent genes PF0738.1n and

PF0739. The insertion site for Genome Region 4 is between nucleotides 734567-734568 in the *P. furiosus* DSM 3638 genome (AE009950). This position has little to no transcriptional activity as determined from analysis of tiling array data (180). SOE-PCR was used to combine ~0.5 kb flanking regions to Genome Region 4 with the intervening pop-out marker cassette. This fragment was cloned into pJHW006 (137) using NheI and NdeI restriction sites to make plasmid pGL008. AdhE gene sequences of *G. thermoglucosidasius*, *G. stearothermophilus*, *T. celere*, *A. flavithermus*, *D. kuznetsovii, C. thermocellum* were PCR-amplified and combined with the *slp* promoter (P_{slp}) (192) and the pGL008 plasmid backbone via Gibson Assembly (200, 248), (NEB) to make plasmids pGL085, pGL086, pGL087, pGL088, pGL089, and pGL107 respectively. These plasmid constructs were sequence-verified. The *Te*AdhE, *Tx*AdhE and *Tx*AdhA genes were PCR-amplified and combined with flanking regions and the pop-out marker cassette via SOE-PCR, and the PCR constructs were used for transformation of *P. furiosus*. PCR constructs for construction of the mutant AdhE strains were amplified from genomic DNA of Tx-AdhE and Te-AdhE strains, each in two segments, using primers to introduce the desired mutation in the AdhE gene (G532D for Te*adhE* and G557D for Tx*adhE*). The two segments for each were joined by SOE-PCR with the overlap region containing the introduced mutation.

***P. furiosus* strain construction**

The Δ*aor* PCR construct was used to transform *P. furiosus* COM1 (Δ*pyrF*) as previously described (137). Transformant colonies were screened by PCR using primers amplifying from outside the flanking regions used for homologous recombination. The strain was colony purified on solid defined medium lacking uracil (137), the *aor* deletion was sequence-verified, and the

strain was designated MW616. In order to construct a markerless Δ*aor* genetic acceptor strain, loss of the pop-out marker cassette was selected for on solid defined medium containing 2.5 mM 5-FOA and 20 µM uracil as previously described (137, 192). Isolates were screened by PCR to verify marker loss and were further purified on solid defined medium with 20 µM uracil. A markerless Δ*aor* strain was designated MW252. MW252 was transformed as previously described (137) with linearized plasmids pGL085, pGL086, pGL087, pGL088, pGL089, and pGL107 as well as PCR constructs containing P_{slp}Te*adhE*, P_{slp}Tx*adhE*, P_{slp}Te*adhEA*, P_{slp}Te*adhE*(G532D), and P_{slp}Tx*adhE*(G557D). Transformants were screened by PCR with primers outside the homologous flanking regions to Genome Region 4, and strains were colony purified on solid defined medium lacking uracil. For strains that were transformed with PCR constructs, the inserted region was sequence-verified. Final strains were designated as indicated in Table 2.

Growth of archaeal and bacterial strains

P. furiosus strains were cultured anaerobically in a sea-water based medium as previously described (227) containing 5 g L^{-1} maltose, 5 g L^{-1} yeast extract and 0.5 µg L^{-1} riboflavin (referred to as YM5 medium). Cultures were grown at 95°C until they reached $1x10^8$ cells mL^{-1} at which point the temperature was decreased to 65°C. For metabolite analysis, 50 mL cultures were grown in 150 mL serum bottles. Samples were taken at the time of the temperature switch and 24, 56, and 90 h after the switch. For enzymatic analysis, 250 mL cultures were grown in 1 L anaerobic bottles and the cultures were harvested after 40 h by centrifugation at 6,000 x g for 10 min. For the bacterial wild-type strains of *T.* sp. X514 and *T. ethanolicus*, 250 mL cultures in 1 L anaerobic bottles were grown in *C. bescii* media (249) with

the addition of 5 g L^{-1} xylose. Cultures were incubated at 65°C for 16-24 h to 1x10^8 cells/mL before harvesting by centrifugation at 6,000 x g for 10 min. *C. thermocellum* cell pellets were obtained from David Mulder and Paul King at the National Renewable Energy Laboratory (NREL).

In vivo ethanol and acetate analysis

Strains were grown in triplicate as described above for *in vivo* ethanol quantitation. A deep-well 96-well plate was used to collect the 1 mL samples. These plates were centrifuged at 4,000 x g to pellet the cells. A multi-channel pipette was used to transfer 200 µL of supernatant to vial inserts containing 5 µL of 88% formic acid. The inserts were assembled into vials and capped. The samples were then analyzed via GC-FID with an identical method as previously described (227).

Cell extract preparation

Cell extracts of *P. furiosus* strains and bacterial wild-type strains were prepared under strict anaerobic conditions in an anaerobic chamber (Coy Laboratories). Cell pellets were suspended in 0.5-1.25 mL 100 mM MOPS, pH 7.0, containing 5 µM $FeCl_2$ (222). Lysozyme was added to the suspended bacterial cells, and these were incubated at 25°C for 20 min. The cell lysate was then centrifuged at 10,000 x g for 10 min, and the supernatant was transferred to a vial which was sealed and stored at 4°C. The protein concentration of the extracts was measured using Bradford reagent (Bio-Rad).

***In vitro* analysis of ALDH, ADH, and AOR activities**

All reactions were carried out in sealed anaerobic cuvettes using 25 mM MOPS, pH 7.0 containing cell extract at a final concentration of 0.2 mg mL^{-1}. ADH and ALDH activities were measured as 65°C using NADPH or NADH at a concentration of 0.2 mM. After the addition of either 0.5 mM acetyl-CoA for ALDH or 10 mM acetaldehyde for ADH, the rate of NAD(P)H oxidation was measured at 340 nm and calculated using a molar absorptivity of 6,220 M^{-1} cm^{-1}. Native *P. furiosus* ADH activity was measured at 80°C in 100 mM EPPS, pH 8.0, using 0.5 mM NADP and 50 mM ethanol. AOR activity was also measured at 80°C in 100 mM EPPS, pH 8.0, using the redox dye benzyl viologen (BV, 1 mM) as the electron acceptor. The AOR reaction was initiated by the addition of 10 mM acetaldehyde, and the rate of BV reduction was measured at 600 nm and calculated using a molar absorptivity of 10,000 M^{-1} cm^{-1}.

Acknowledgments

We thank David Mulder and Paul King for supplying cells of *C. thermocellum* and Daehwan Chung and Janet Westpheling for supplying *C. bescii* strain JWCB032. This work was supported by the U.S. National Science Foundation (CBET-1264052 and CBET-1264053).

Tables and Figures

Table 4.1 Gene donors for AdhE

Eight thermophilic bacteria with optimal growth temperature ranging from 60°C to 69°C were used as gene donors for AdhE insertion into *P. furiosus*.

Abbr.	Organism	Growth optimum (range) (°C)	Genome GC content (%)	AdhE GI accession number	AdhE sequence identity to Ct (%)
Ct	*Clostridium thermocellum* (ATCC 27405)	60 (28-69)	39	489614598	100
Tx	*Thermoanaerobacter* sp. X514	60 (50-70)	34.5	166853526	52
Gt	*Geobacillus thermoglucosidasius* NBRC 107763	62 (42-69)	43.7	617767229	51
Dk	*Desulfotomaculum kuznetsovii* DSM 6115	62.5 (50-85)	54.9	333978712	47
Af	*Anoxybacillus flavithermus* WK1	62.5 (40-70)	41.8	431999657	51
Gs	*Geobacillus stearothermophilus* NUB3621	65 (48-73)	44.4	612116738	51
Tc	*Thermobrachium celere* DSM 8682	66 (43-75)	31.3	517490057	71
Te	*Thermoanaerobacter ethanolicus* JW200	69 (37-78)	34.1	489965432	52

Table 4.2 Sequence identity of the AdhEs

The amino acid sequence identity (%) of the eight AdhEs of this study in comparison to each other.

	Te	Tx	Gs	Gt	Af	Ct	Tc	Dk
Te	100	97	67	68	68	52	54	65
Tx	97	100	68	68	68	52	54	65
Gs	67	68	100	97	88	51	53	63
Gt	64	68	97	100	89	51	54	63
Af	68	68	88	89	100	51	56	64
Ct	52	52	51	51	51	100	71	47
Tc	54	54	53	54	56	71	100	49
Dk	65	65	63	63	64	47	49	100

Table 4.3 *P. furiosus* strains used and constructed in this study

Strain	Alias	Genotype	Reference
Genetic background and control strains			
MW002	COM1	Δ*pyrF*	(137)
MW252	ΔAOR	Δ*pyrF* Δ*aor*	This work
MW004	COM1c2	Δ*pyrF*::*pyrF*	(234)
MW616	ΔAORc	Δ*pyrF* Δ*aor*::P_{gdh} *pyrF*	This work
AdhE expression strains			
MW627	Ct-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} Ct*adhE* P_{gdh}*pyrF*	This work
MW332	TX-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} TX*adhE* P_{gdh}*pyrF*	This work
MW371	Gt-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} Gt*adhE* P_{gdh}*pyrF*	This work
MW364	Dk-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} Dk*adhE* P_{gdh}*pyrF*	This work
MW373	Af-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} Af*adhE* P_{gdh}*pyrF*	This work
MW357	Gs-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} Gs*adhE* P_{gdh}*pyrF*	This work
MW372	Tc-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} Tc*adhE* P_{gdh}*pyrF*	This work
MW392	Te-AdhE	Δ*pyrF* Δ*aor* GR4::P_{slp} Te*adhE* P_{gdh}*pyrF*	This work
MW325	Te-AdhEA	Δ*pyrF* Δ*aor* GR4::P_{slp} Te*adhE* TX*adhA* P_{gdh}*pyrF*	This work
Mutant AdhE expression strains			
MW623	TX-AdhE*	Δ*pyrF* Δ*aor* GR4::P_{slp} TX*adhE*(G557D) P_{gdh}*pyrF*	This work
MW625	Te-AdhE*	Δ*pyrF* Δ*aor* GR4::P_{slp} Te*adhE*(G532D) P_{gdh}*pyrF*	This work

[a] GR3 represents insertion at genome region 3 (between genes PF0574 and PF0575) and GR4 represents insertion at genome region 4 (between genes PF PF0738.1n and PF0739). Organism abbreviations are listed in Table 1.

Table 4.4 ALDH and ADH specific activities in *P. furiosus* strains and bacterial crude extracts

Strain	ALDH activity[a] (acetyl-CoA)		ADH activity[a] (acetaldehyde)	
	NADH	NADPH	NADH	NADPH
Control strains				
COM1	<0.05	<0.05	<0.05	<0.05[b]
ΔAOR	<0.05	<0.05	<0.05	<0.05[b]
AdhA	<0.05	<0.05	<0.05	0.94
P. furiosus AdhE expression strains				
Te-AdhEA	0.58	<0.05	0.48	1.66
Te-AdhE	0.29	<0.05	1.00	0.07
Tx-AdhE	0.17	<0.05	0.53	<0.05
Gt-AdhE	<0.05	<0.05	<0.05	<0.05
Dk-AdhE	<0.05	<0.05	0.05	<0.05
Af-AdhE	<0.05	<0.05	0.05	<0.05
Gs-AdhE	<0.05	<0.05	<0.05	<0.05
Tc-AdhE	<0.05	<0.05	0.06	<0.05
Ct-AdhE	<0.05	<0.05	<0.05	<0.05
Mutant AdhE expression strains				
Te-AdhE*	0.44	<0.05	<0.05	<0.05
Tx-AdhE*	0.72	<0.05	<0.05	<0.05
Activity in the native organism				
C. thermocellum	4.96	0.06	19.42	<0.05
T. sp. strain X514	0.18	0.02	1.09	3.38
T. ethanolicus	0.70	<0.05	2.82	6.96

[a] Specific activities are listed as µmol NAD(P)H oxidized min^{-1} mg^{-1}.
[b] ADH activity in the forward reaction was not detected; however, the reverse assay resulted in 0.2 U mg^{-1} of ethanol reduction via NADP

Figure 4.1 Pathways for native and engineered ethanol production in *P. furiosus*

Glycolysis to pyruvate (0.5 mol glucose) yields no ATP and produces 2 Fd_{red} (net 2 e^-). Ethanol production from: **A)** pyruvate via the pyruvate decarboxylase activity of pyruvate ferredoxin oxidoreductase (POR) and an alcohol dehydrogenase (ADH) results in no net ATP, produces 2 Fd_{red} and uses 1 NAD(P)H (net 0 e^-); **B)** acetyl-CoA via AdhE yields no net ATP, produces 4 Fd_{red} and uses 2 NADH (net 0 e^-); **C)** acetate via aldehyde ferredoxin oxidoreductase (AOR, here also representing other oxidoreductases with aldehyde oxidizing activity), and AdhA results in 1 net ATP via acetyl-CoA synthase (ACS), produces 2 Fd_{red} and uses 1 NADPH (net 0 e^-). *P. furiosus* enzymes are shown in orange, AdhA is shown in red, and AdhE is shown in blue. Proposed physiological reactions are indicated with solid arrows and engineered pathways in recombinant strains are indicated with dashed arrows.

Figure 4.1

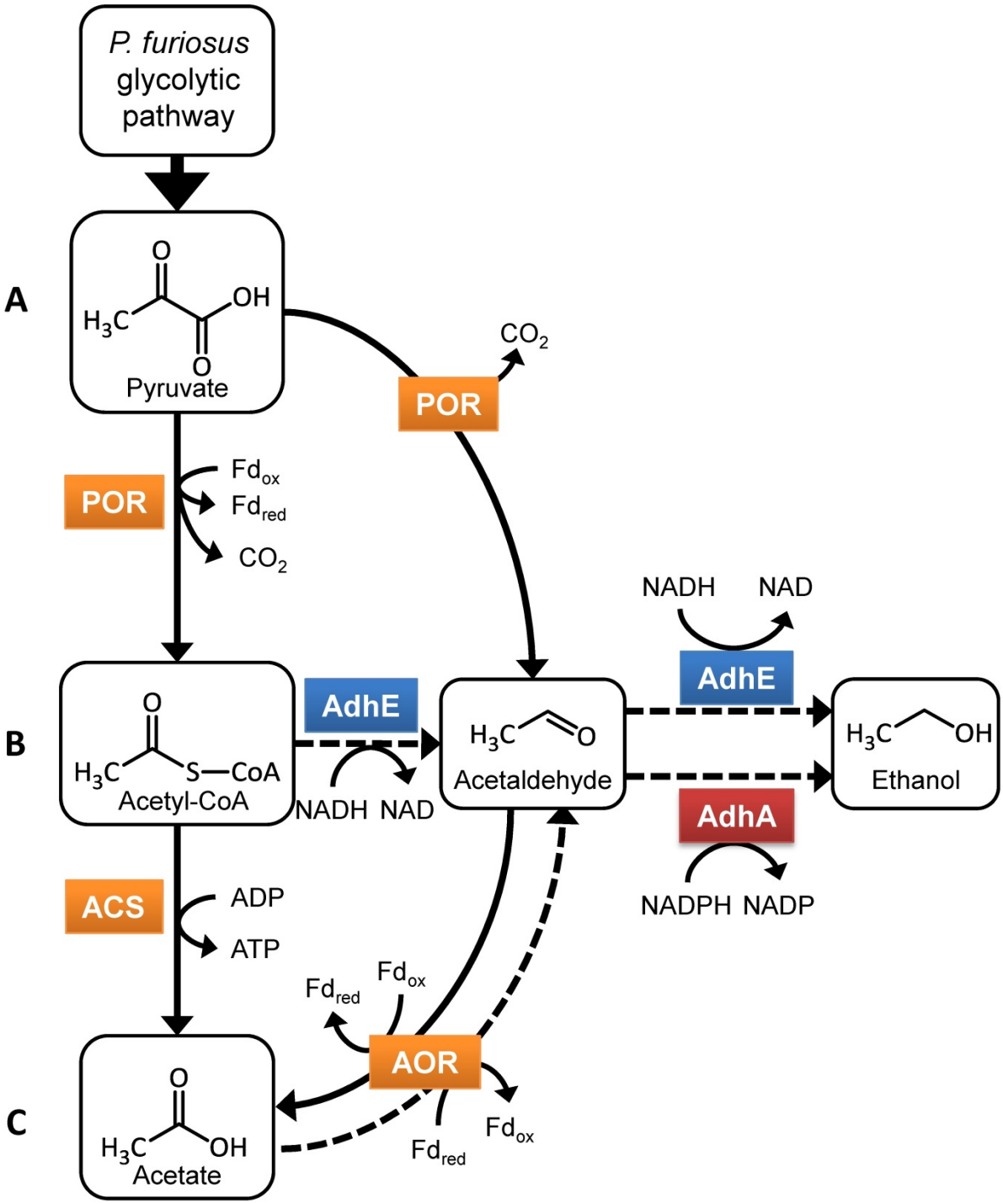

P. furiosus glycolytic pathway
A
Pyruvate
POR
CO2
POR
Fdox
Fdred
CO2
NADH NAD
AdhE
B
Acetyl-CoA
AdhE
NADH NAD
Acetaldehyde
Ethanol
AdhA
NADPH NADP
ACS
ADP
ATP
Fdred Fdox
AOR
C
Acetate
Fdox
Fdred

Figure 4.2 Ethanol yield versus growth temperature of selected source organisms for AdhE genes

Organism abbreviations are listed in Table 1. Optimal growth temperature is indicated by an open circle with a line indicating the maximum growth temperature. Ethanol yields were calculated from previously published data (155, 162, 193, 239, 250-253). Dk is reported to grown on but not produce ethanol (254).

Figure 4.2

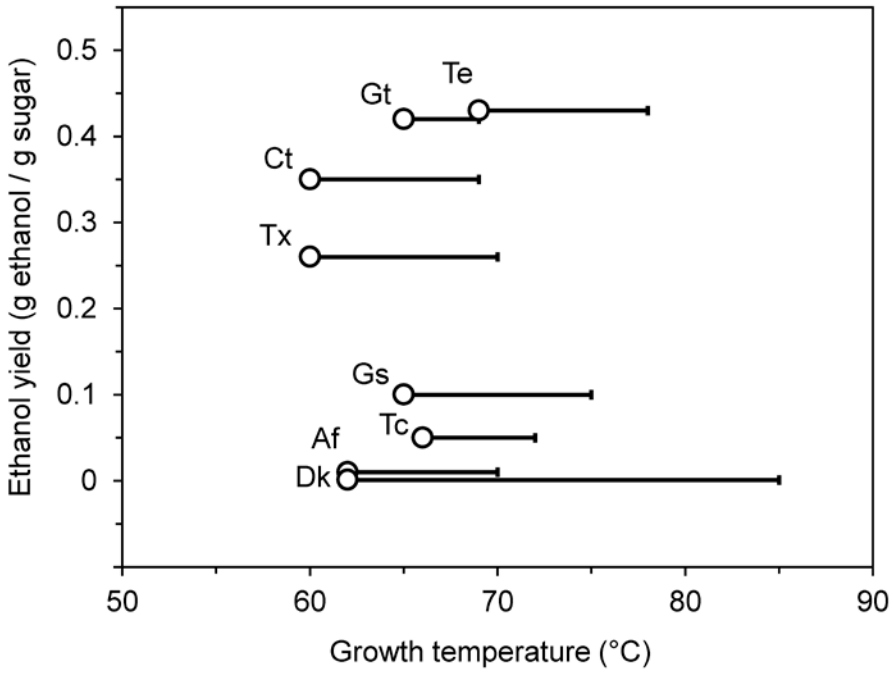

Ethanol yield (g ethanol / g sugar)
0.5
0.4
0.3
0.2
0.1
0
Te
Gt
Ct
Tx
Gs
Tc
Af
Dk
50
60
70
80
90
Growth temperature (°C)

Figure 4.3 Deletion of *aor* does not affect growth of *P. furiosus* and abolishes 80% of aldehyde oxidoreductase activity

A. MW004 (Δ*pyrF*::*pyrF*, black line with closed circles) and MW616 (Δ*pyrF* Δ*aor*::P_{gdh}*pyrF*, dashed line with open circles) strains were grown at 72°C for 72 h and then at 65°C up to 8 d in YM medium containing a total of 2 g L^{-1} yeast extract. Error bars represent S.D., *n* = 3.

B. Aldehyde oxidoreductase (AOR) activity in cell extracts of COM1 (Δ*pyrF*) and ΔAOR (Δ*pyrF* Δ*aor*) at 80°C.

Figure 4.3

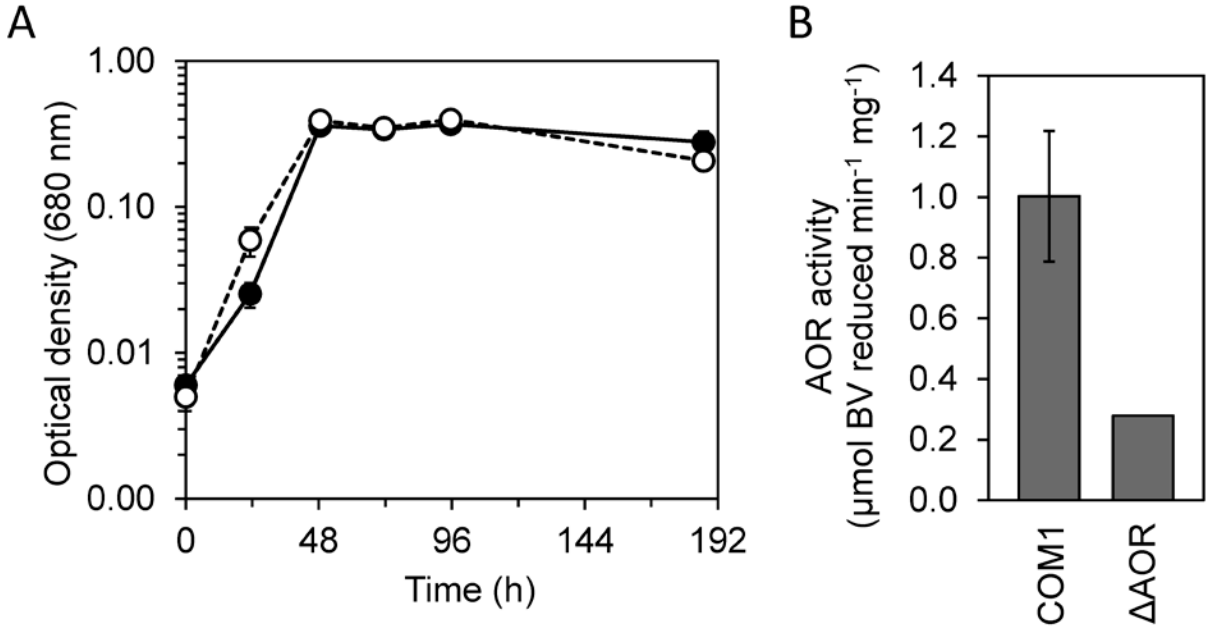

A
1.00
0.10
0.01
0.00
Optical density (680 nm)
0
48
96
144
192
Time (h)
B
AOR activity
(µmol BV reduced min-1 mg-1)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
COM1
ΔAOR

Figure 4.4 Ethanol production by AdhE-containing strains of *P. furiosus*

All strains, aside from COM1c2, have AOR deleted. *P. furiosus* strains and two control strains (ΔAORc and COM1c2) were grown in closed serum bottles with 50 mL YM5 medium to mid-log phase at 95°C and then incubated at 65°C for 90 h. **A)** Ethanol production over time. **B)** Ethanol produced per glucose consumed calculated from ethanol and acetate values for the 90 h time point (shown in A), ordered from highest to lowest. Ethanol and acetate are assumed to be the only carbon end-products of glycolysis. Error bars represent S.D., *n* = 3.

Figure 4.4

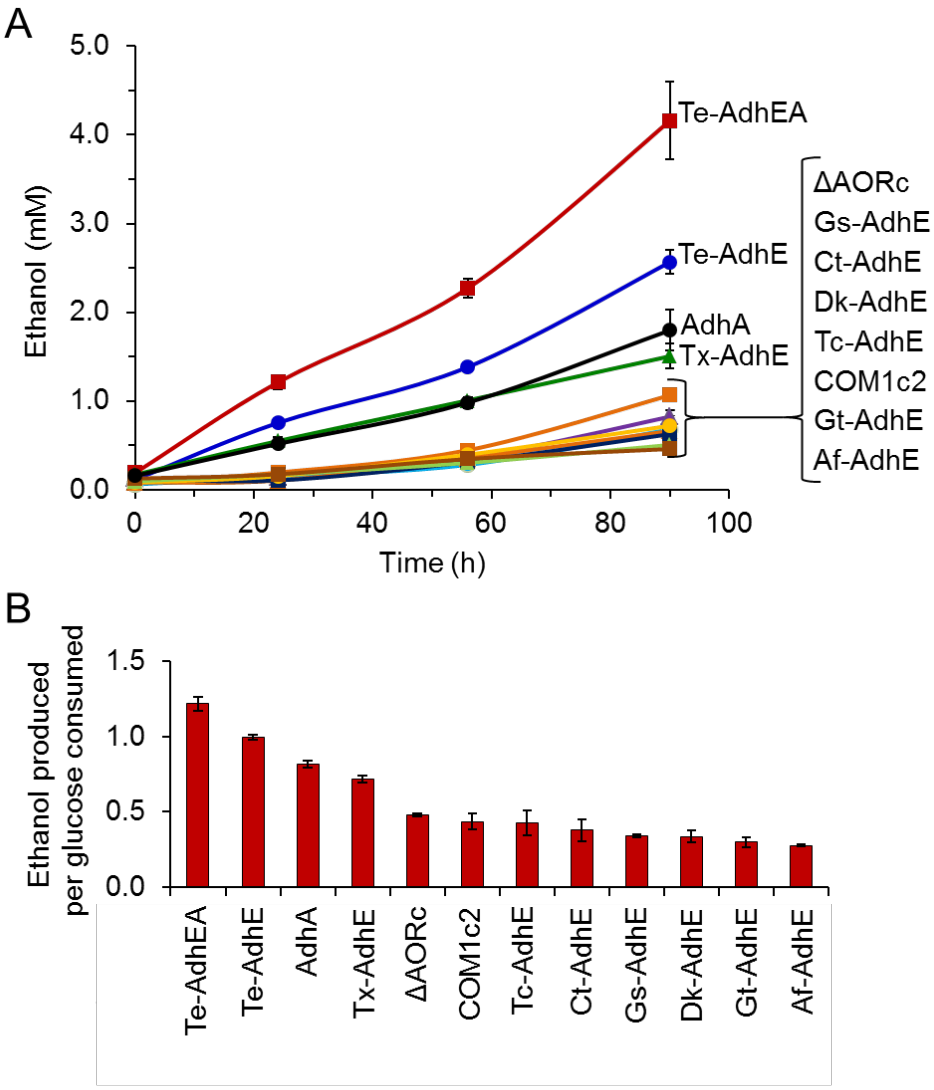
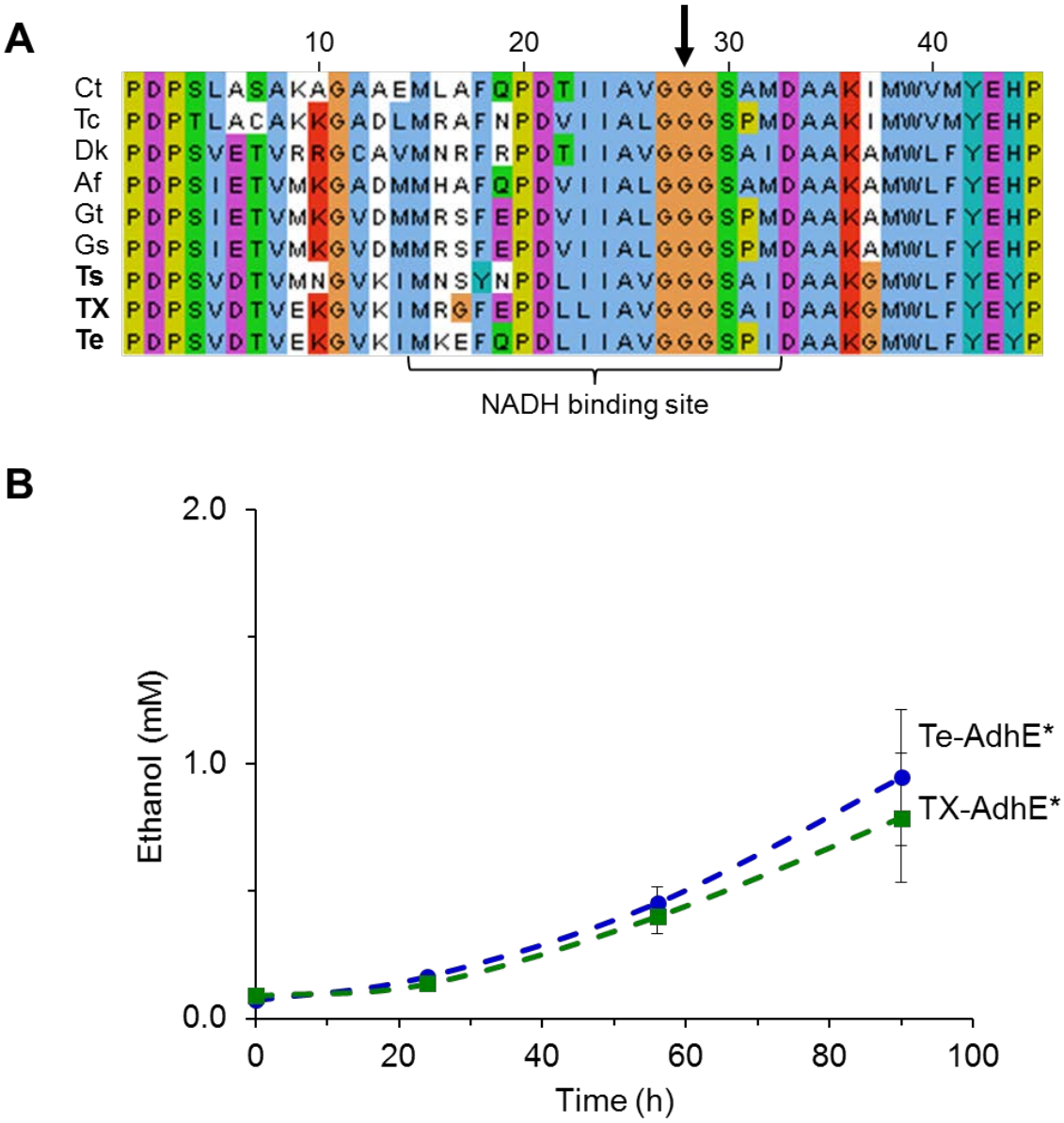

A
Ethanol (mM)
Time (h)
Te-AdhEA
Te-AdhE
AdhA
Tx-AdhE
ΔAORc
Gs-AdhE
Ct-AdhE
Dk-AdhE
Tc-AdhE
COM1c2
Gt-AdhE
Af-AdhE
B
Ethanol produced per glucose consumed
Te-AdhEA
Te-AdhE
AdhA
Tx-AdhE
ΔAORc
COM1c2
Tc-AdhE
Ct-AdhE
Gs-AdhE
Dk-AdhE
Gt-AdhE
Af-AdhE

Figure 4.5 Mutations to change the nucleotide specificity of AdhE

A. Multiple alignment of ADH domain NADH binding site of selected AdhEs with *T. saccharolyticum* AdhE. Alignment performed using Clustal Omega (255) and visualized with Jalview (256). Location of NADH binding site is indicated with a bracket and the position of the G→D mutation (G544D in *T. saccharolyticum*) is indicated with an arrow. Organism abbreviations are as follows: Ct, *Clostridium thermocellum*; Tc, *Thermobrachium celere* DSM 8682; Dk, *Desulfotomaculum kuznetsovii* DSM 6115; Af, *Anoxybacillus flavithermus* WK1; Gt, *Geobacillus thermoglucosidasius* NBRC 107763; Gs, *Geobacillus stearothermophilus* NUB3621; Ts, *T. saccharolyticum*; TX, *Thermoanaerobacter* sp. X514; Te, *Thermoanaerobacter ethanolicus* JW200. **B.** *In vivo* ethanol production by *P. furiosus* strains containing the mutated *Thermoanaerobacter* AdhEs was indistinguishable from background. Error bars represent S.D., *n* = 3.

Figure 4.5


A
10
20
30
40
Ct PDPSLASAKAGAAEMLAFQPDTIIAVGGGSAMDAAKIMWVMYEHP
Tc PDPTLACAKKGADLMRAFNPDVIIALGGGSPMDAAKIMWVMYEHP
Dk PDPSVETVRRGCAVMNRFRPDTIIAVGGGSAIDAAKAMWLFYEHP
Af PDPSIETVMKGADMMHAFQPDVIIALGGGSAMDAAKAMWLFYEHP
Gt PDPSIETVMKGVDMMRSFEPDVIIALGGGSPMDAAKAMWLFYEHP
Gs PDPSIETVMKGVDMMRSFEPDVIIALGGGSPMDAAKAMWLFYEHP
Ts PDPSVDTVMNGVKIMNSYNPDLIIAVGGGSAIDAAKGMWLFYEYP
TX PDPSVDTVEKGVKIMRGFEPDLLIAVGGGSAIDAAKGMWLFYEYP
Te PDPSVDTVEKGVKIMKEFQPDLIIAVGGGSPIDAAKGMWLFYEYP
NADH binding site
B
2.0
1.0
0.0
Ethanol (mM)
0
20
40
60
80
100
Time (h)
Te-AdhE*
TX-AdhE*

CHAPTER 5

EXPRESSION OF THE COMPLETE 3-HP/4-HB CYCLE FOR CARBON FIXATION IN *PYROCOCCUS FURIOSUS*

Keller MW†, Lipscomb GL†, Thorgersen MP†, Schut GJ†, WJ, Hawkins AS, Kelly RM, Adams MWW (2016) Expression of the complete 3-HP/4-HB cycle for carbon fixation in *Pyrococcus furiosus*. *Metab eng.* To be submitted

†Authors contributed equally to this work

Abstract

An important aspect of a biofuel strategy is the substrate utilization portion of that strategy. The ideal substrate or feedstock is not only cheap and abundant but can be efficiently utilized and renewably replenished. Sugar from corn is currently the leading biofuel feedstock, but its used is an inefficient process that also competes with food production. Emerging strategies aim to utilize the non-edible recalcitrant portion of plant biomass as a feedstock. These strategies make improvements for the problem of direct competition with food; however, land usage is still problematic, and efficiencies are only marginally improved. A very different strategy, known as electrofuels, is to abandon plant based photosynthesis and carbohydrate intermediates altogether. In this strategy, microbial carbon dioxide assimilation pathways are directly coupled to fuel synthesis pathways. Our strategy is to engineer the hyperthermophile *Pyrococcus furiosus*, which grows optimally at 100°C, to express the 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) cycle for carbon dioxide fixation, drive that pathway with molecular hydrogen via the native soluble hydrogenase, and direct the acetyl-CoA produced to a bacterial butanol or ethanol fermentation pathway. The 16 genes encoding the *M. sedula* 3-HP/4-HB cycle were transferred to *P. furiosus* in four iterative transformations, and their expression was verified by qPCR and enzymatic analysis. While the entire pathway could be demonstrated to function *in vitro*, it could not be demonstrated *in vivo*, which will likely require further metabolic engineering to overcome bottlenecks and remove competing pathways.

Introduction

Sugar-based ethanol is by far the most abundant biofuel produced today, but it is made by an inefficient process with a high carbon intensity that also competes with food production (45). In order to overcome these problems, an alternative strategy is to exploit microbial carbon fixation pathways and direct that carbon into a fuel synthesis pathway. This strategy, which is essentially to build a fuel molecule directly and one carbon at a time, is in stark contrast to the current sugar-based strategy. Within a microbe that is metabolically engineered to fix carbon dioxide and synthesize a fuel molecule, energy carriers such as electric current, hydrogen, formate, and carbon monoxide, which can be generated efficiently and renewably from solar photons, can be used to directly drive carbon dioxide fixation and fuel synthesis pathways (70-72). The use of molecular hydrogen as an energy carrier is possible with *Pyrococcus furiosus* as it contains a soluble hydrogenase (SH1) that can produce NADPH from H_2 (123, 124). However, *P. furiosus* does not contain a carbon dioxide fixation pathway. Therefore, the overall goal of this work was to produce the 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) cycle for carbon dioxide fixation from *Metallosphaera sedula* (**Figure 5.1**) in *P. furiosus*. This pathway consists of 13 enzymes encoded by 16 genes and produces acetyl-CoA as a product and importantly uses NADPH as its sole source of reductant (148-154), which is compatible with the *P. furiosus* H_2 uptake system SH1 (123, 124). The first three enzymes (E1-3) convert acetyl-CoA to 3-HP and constitute sub-pathway 1 (SP1) (192). The next six enzymes (E4-9), along with bifunctional activities of E1 and E2, convert 3-HP to 4-HB and constitute sub-pathway 2 (SP2). The final four enzymes (E10-13) convert one molecule of 4-HB to two molecules of acetyl-CoA and constitute sub-pathway 3 (SP3). Here, the 3-HP/4-HB cycle for carbon dioxide fixation will

be expressed in *P. furiosus* and evaluated for enzymatic activity of individual and coupled pathway enzymes, *in vitro* product formation via sub-pathways, and *in vivo* CO_2 fixation by H_2/CO_2 dependent product formation and ^{13}C labeling.

Results

Extensive genetic manipulation was required to incorporate the complete 3-HP/4-HB cycle for carbon dioxide fixation into *P. furiosus*. As shown in **Figure 5.1**, the overall pathway consists of 13 enzymes encoded by 16 genes. The first sub-pathway is composed of five genes, the products of which catalyze the first three steps of the *M. sedula* 3-HP/4-HB pathway for carbon fixation. In *M. sedula*, these five genes do not appear in an operon. For expression in *P. furiosus* strain MW56, they were assembled into an artificial operon and inserted into the chromosome of *P. furiosus* (192). In order to insert the remaining 11 genes, three additional operons were assembled: SP2A, SP2B, and SP3 (**Figure 5.2**). For certain recombinant strains of *P. furiosus* such as MW76, the SP1 operon also contains the accessory genes encoding biotin protein ligase (BL) and carbonic anhydrase (CA) of *M. sedula*. These operons also needed to be inserted into the *P. furiosus* chromosome for expression. The targets for these insertions, known as genome regions (GR), were identified via whole genome tiling array (180), and seven were selected on the basis of large transcriptionally inactive regions in between two convergent genes (GR1-7; **Table 5.1**). Three regions between parallel genes were also selected (GR8-10; **Table 5.1**). The intention of this design is to avoid disruption of any native gene or its promoter and as of the approximately 400 unique *P. furiosus* strains generated to date, no phenotypic issue has arisen due to the insertion of foreign genes into any particular genome region.

To construct *P. furiosus* strain MW98, which expresses the entire *M. sedula* cycle for carbon dioxide fixation, four iterative transformations (**Figure 5.3**) were performed to insert S-layer promoter driven (192) artificial operons SP2A containing E4, E5, and E6 into GR1, SP2B containing E7, E8α, E8β, and E9 into GR2, SP1 containing BL, CA, E1α, E1β, E1γ, E2, and E3 into GR3, and SP3 containing E10, E11, E12, and E13 into GR4 (**Figure 5.2**). The expression of these 18 foreign genes in *P. furiosus* strain MW98 was verified by qPCR analysis (**Figure 5.4**). Strains with a single insertion operon, regardless of the content or location of that operon, showed no phenotype and grew at an identical rate to the parental strain COM1. However, the insertion of a second artificial operon, regardless of the content or location, resulted in a growth defect (**Figure 5.5**). The effect was not continuously dose dependent but was discrete in that all strains with one or no artificial operon grew at identical rates, and all strains with two or more artificial operons grew at identical rates. This growth defect could not be correlated to the content or location of any inserted operon, rather only the presence of two or more operons. Possible explanations of this defect are the burden of protein expression and the dilution of transcription factors across the multiple S-layer promoters.

Additional pathways such as ethanol and butanol synthesis pathways (Keller and Lipscomb, unpublished data) (1, 227, 230) and an energy conserving carbon monoxide dehydrogenase (CODH) (228) were also inserted in the recombinant *P. furiosus* strains (**Table 5.2**). The ethanol and butanol pathway containing strains were designed to allow for the detection of H_2 and CO_2 dependent ethanol or butanol rather than acetate, which is a major end-product of *P. furiosus* native metabolism. To enhance the conversation of acetyl-CoA to ethanol or butanol rather than acetate, a gene encoding the enzyme that competes with the

pathway, acetyl-CoA synthase (ACS), was deleted in these strains. CODH was incorporated to generate additional ATP, which the cycle requires six equivalents for every complete turn. Counter-selection of MW98 resulted in a full cycle containing Δ*pyrF* acceptor strain, but initial attempts to perform a fifth transformation using this strain were unsuccessful. Therefore, larger operons were inserted into *P. furiosus* to increase the amount of genetic material transferred in the same number of transformations. The Δ*pyrF* acceptor strain generated from MW78 (SP1+SP2), which is the parental strain of MW98 (SP1+SP2+SP3), was used as the parental strain to express the complete 3-HP/4-HB pathway along with a) an ethanol pathway, b) a butanol pathway, c) deletion in ACS, and d) additional homologs of E10 in SP3. It was found that incubating plates at 80°C rather than 90°C increased the plating efficiency to allow for a fifth transformation of *P. furiosus*. In the case of MW353, E7+E8+E9 was inserted at GR2, E4+E5+E6 was inserted at genome region 1, BL+CA+E1+E2+E3 was inserted at GR3, CODH was inserted at GR5, and Teth-AdhE+E10+E11+E12+E13 was inserted at GR4. The result is a *P. furiosus* strain whose chromosome is roughly 2% foreign DNA (40 kb / 2 Mb).

P. furiosus strains MW56 expressing SP1, MW76 expressing SP1+SP2, MW47 and MW76 expressing SP2 and SP1+SP2 respectively, and MW98 expressing SP1+SP2+SP3, were used for enzymatic and product analysis (**Table 5.3** and **Table 5.4**). By employing a similar strategy as was utilized in the production of 3-HP by SP1 (192), the remaining enzymes of the *M. sedula* 3-HP/4-HB pathway for carbon fixation could be characterized. The recombinant *P. furiosus* strains were grown at 95°C until the cell density reached mid log phase (~2-5x10^7 cells/mL rather than 1x10^8 cells/mL for wild-type) and shifted to 75°C for 16 hours to allow for the expression of active enzymes. In these strains, the specific activities of 13 reactions were

measured and found to be comparable or superior to that found in cell-extracts from autotrophically grown *M. sedula*. These reactions are listed in **Table 5.3** and illustrated around the cycle in **Figure 5.6**. Together, these enzyme assays span the entire cycle except for E10 and E11.

Detection of key intermediates was also performed in order to demonstrate the complete cycle *in vitro*. In MW56 (SP1) extract, 3-HP could be detected from acetyl-CoA. In MW78 (SP1+SP2), 4-HB could be detected from succinic semialdehyde. In MW98 (SP1+SP2+SP3) extract, acetate could be detected from propionate, acrylate, and 3-HP (**Table 5.4**). Though only detected in small concentrations above the control, the 3-HP dependent production of acetate indicates SP2 and SP3 are functional *in vitro* (**Figure 5.7**). Additionally, SP2 depends on the bi-functional SP1 enzymes, E1 and E2, so the production of acetate from 3-HP required the coupling of 12 out of the 13 enzymes of the entire pathway. Acetate was detected as a product by coupling the production of acetyl-CoA to the *P. furiosus* enzyme ACS.

Despite such promising results *in vitro* showing that the complete pathway is active, neither the complete cycle nor utilization of the product of one sub-pathway by a second sub-pathway has been demonstrated *in vivo*. Numerous of attempts have failed to demonstrate *in vivo* the H_2 and CO_2 dependent production of acetate in the full cycle strain MW98, butanol production in the strain MW203 which contains the full cycle and the BuOH-1 pathway (227), ethanol production in strains that contain an ADH (MW220, MW222, and MW353), or any accumulation of 4-HB in strain MW78 which contains the first two sub-pathways. *M. sedula* contains several candidates for E10, and its identity was disputed at the initial stages of this project (257, 258). In *P. furiosus* strain MW98, the activity of E10 (~0.2 U/mg, by measuring the

ATP and 4-HB dependent disappearance of CoASH), was similar to that measured in the parental strain COM1, presumably due to the reverse activity of ACS. Alternate candidates for E10 from *M. sedula* were evaluated in *P. furiosus*. Four strains expressed four different candidates for E10 (E10a Msed_0406, E10b* Msed_1353 *W424G, E10c Msed_1394, and E10d Msed_1422) along with the complete cycle: MW98, MW100, MW102, and MW375 respectively. These strains, along with the parental strain, were evaluated for succinic semialdehyde dependent acetate production (E9+**E10**+E11+E12+E13+ACS). MW98, which contained the original E10 candidate, performed the worst while MW102 and MW375, which contain E10c and E10d respectively, performed the best (**Figure 5.8**). Future strategies should make use of E10c or E10d for an optimally functioning sub-pathway 3.

The incorporation of CODH along with E10d prompted a new round of *in vivo* attempts at $CO/H_2/CO_2$ acetate (MW375) or ethanol (MW377) production. In order to increase sensitivity, ^{13}C labeled carbon monoxide was supplied to concentrated cell suspensions of MW375 and MW377 and after ~24 hours at 72°C, spent media were analyzed by mass spectrometry for ^{13}C labeled acetate or ethanol. Unfortunately, no enrichment of ^{13}C labeled acetate or ethanol was detected. Again, despite the promising *in vitro* results, it appears that the *M. sedula* cycle as expressed in *P. furiosus* is at least partially incorrect, incomplete, or interfered with by the native metabolism of *P. furiosus*.

The second sub-pathway (SP2) is also suspected to be problematic since 4-HB could not be detected *in vivo* in strain MW78 (SP1+SP2). The conversion of acetyl-CoA to 4-HB by SP1+SP2 requires 11 steps to be performed by 9 enzymes (**Figure 5.1**). E1 was identified as a

bottleneck in SP2 by incubating MW98 with 3-HP (and the other necessary substrates: ATP, CoASH, NADPH, $NaHCO_3$, NAD, ADP, and P_i) *in vitro* in an effort to convert that 3-HP to acetate by SP2 and SP3. The 3-HP was readily taken up by the pathway, but the major product was propionate rather than acetate (**Table 5.4** and **Figure 5.7**). This indicates that the propionyl-CoA produced by E4+E5+E6 was hydrolyzed to propionate by ACS rather than carboxylated to methylmalonyl-CoA by E1.

Discussion

The negative results of ^{13}C incorporation using SP1+SP2+SP3 in strain MW98 indicate an impasse somewhere in the cycle. Careful analysis must be performed to both validate the identity of the cycle enzymes and to identify and overcome any bottlenecks. *In vitro*, this has already been done and indicates the carboxylation of propionyl-CoA to methylmalonyl-CoA (**Figure 5.1**) as a bottleneck. The consequence of this bottleneck is propionyl-CoA is hydrolyzed to propionate by ACS and the flux of carbon through the pathway is greatly diminished (**Figure 5.7**). The acetyl/propionyl-CoA carboxylase (E1) which, in addition to having the lowest specific activity of any of the cycle enzymes in *P. furiosus*, is required at two steps of the pathway (**Figure 5.1**). Multiple copies of the carboxylase may therefore be required.

Competing pathways in *P. furiosus* may need to be deleted. ACS1α has been deleted in strains that also contain SP1 which resulted in a roughly 3-fold decrease of *in vitro* hydrolysis of acetyl-CoA and propionyl-CoA to acetate and propionate respectively and a 3-fold increase of 3-HP production (201). This enzyme competes directly with E1 for the substrates acetyl-CoA and propionyl-CoA by hydrolyzing them to the free acid form and preventing their carboxylation

(**Figure 5.1**). In strain MW98, which contains the full cycle and ACS, this effect can be seen where 3-HP and acrylate are readily incorporated in the cycle *in vitro*, but they result in large amounts of propionate accumulation (**Figure 5.7**). In fact, they are nearly completely and exclusively converted to propionate (**Table 5.4**) which indicates SP2A (E4+E5+E5) operates very efficiently, but E1 carboxylation of propionyl-CoA to methylmalonyl-CoA is a severe bottleneck, and the accumulating propionyl-CoA is hydrolyzed to propionate by ACS (**Table 5.4**). The deletion of ACS1α should be evaluated where 3-HP and all co-substrates are incubated *in vitro* in a ΔACS1α+SP1+SP2+SP3 strain (such as MW244, MW222, or MW203). Ideally, the accumulation of propionate should be reduced, compared to what has been shown in **Table 5.4** and **Figure 5.7,** and the accumulation acetate, ethanol, or butanol in strains MW244, MW222, and MW203 respectively should increase. An isozyme of the alpha subunit of ACS, ACS3α, confers specificity to succinyl-CoA as a substrate which also appears in the cycle as an intermediate (120). The deletion of this ACS3α and removal of this competing pathway is also likely necessary.

Additionally, the pathway requires a large amount of ATP. Recently, an energy-conserving carbon monoxide dehydrogenase (CODH) complex has been heterologously expressed in *P. furiosus* (228). In the presence of CO, this complex has been shown to provide energy for growth on fermentable substrates in the absence of an electron acceptor (such as S^0) and cause the uptake of acetate, which is presumably due to an abundance of ATP and the *in vivo* reversal of ACS. Combining the CODH driven ATP production with the 3-HP/4-HB cycle could rescue the growth defect and supply ATP to improve the cycle once it is demonstrated to be whole and functional to any degree.

Materials and Methods

Strain construction

PCR was performed using genomic DNA to generate the individual PCR products of the *P. furiosus* S-layer promoter (P_{slp}), the phosphoenolpyruvate synthase promoter (P_{pep}), and the membrane bound hydrogenase promoter (P_{mbh}), or the *M. sedula* genes. The pathway genes were assembled into the artificial operons. The SP1 operon expressed E1αβ (Msed_0147-Msed_0148), E1γ (Msed_1375), E2 (Msed_0709) and E3 (Msed_1993) under the control of P_{slp} and in some instances BL (Msed_2010) and CA (Msed_0390) under the control of P_{pep}. The SP2A operon contained genes encoding E4 (Msed_1456), E5 (Msed_2001), and E6 (Msed_1426) under the control of P_{slp}. The SP2B operon contained genes encoding E7 (Msed_0639), E8α (Msed_0638), E8β (Msed_2055), and E9 (Msed_1424) under the control of P_{slp}. The SP3 operon contained genes encoding E10a (Msed_0406), E11 (Msed_1321), E12a (Msed_0399), and E13 (Msed_0656) under the control of P_{slp} (148-154). The artificial operons also contained native *P. furiosus* ribosomal binding sites (192) and were incorporated into the chromosome as previously described (137). Alternative gene candidates identified by homology include E10b* (Msed_1353 *W424G), E10c (Msed_1394), and E10d (Msed_1422), and E12b (Msed_1423). In some strains ACSIα (PF1540) was deleted (**Table 5.2**) (201). The BuOH-1 pathway (227) expressing Bad (Teth514_1942) and Bdh (Teth514_1935) from *Thermoanaerobacter* sp. strain X514, Ter (STHERM_c16300; The *ter* gene was codon-optimized for *P. furiosus* Genewiz, Inc., South Plainfield, NJ) from *Spirochaeta thermophila*, Thl (TTE0549), Hbd (TTE0548), and Crt (TTE0544) from *Caldanaerobacter subterraneus* subsp. *tengcongensis* (227) was co-expressed

with the complete cycle and the ACSIα deletion (MW203). Alcohol dehydrogenase A (AdhA) from *T.* X514 (Teth514_0564) was co-expressed with the complete cycle in strain MW220 (1). Alcohol dehydrogenase E (AdhE) from *T.* X514 (Teth514_0627) was co-expressed with the complete cycle and the ACSIα deletion in strain MW222 (1). AdhE from *Thermoanaerobacter ethanolicus* was co-expressed with the complete cycle, the ACSIα deletion, and the energy conserving CODH from *Thermococcus onnurineus* (TON_1017-TON_1031) in strain MW353, and the full cycle and CODH were co-expressed in strain MW351 (228).

Growth and expression of *P. furiosus* strains containing the 3-HP/4-HB cycle

P. furiosus recombinant strains were grown in a media containing 5 g/L yeast extract and maltose at 95°C until the cell density reached mid log phase(~2-5x10^7 cells/mL) and shifted to 75°C for 16 hours (192). For strains expressing CODH, the headspace was 100% CO, maltose was omitted, and yeast extract was reduced to 1 g/L (228). For growth in a 20 L fermenter, the initial growth temperature was 90°C, and the culture was sparged with 10% CO_2/90% N_2, stirred, and the pH was maintained at 6.8 by addition of 10% $NaHCO_3$. Cell extracts were prepared anaerobically as described previously for the preparation of SP1 containing extracts (137) in 100 mM MOPS, pH 7.5, re-concentrated three-times with a 3 kDa centrifugation filter and stored at -80 °C. (192)

M. sedula (DSM 5348) was grown autotrophically at 70 °C with micro-bubblers feeding 1 mL/min 80/20 H_2/CO_2 and 100 mL/min air in the defined medium, DSMZ 88, at pH 2.0 as previously described (148). To obtain cell-free extracts, frozen cell pellets were anaerobically suspended in 50 mM Tris HCl pH 8.0 containing 0.5 µg/mL DNase1 and stirred for 1 hr in an

anaerobic chamber. The cell extract was centrifuged at 100,000 x g for 1 hr and the supernatant was stored at -80 °C. (142, 192)

***In vitro* analysis of the 3-HP/4-HB cycle**

All enzyme assays were carried out anaerobically at 75°C. (**Table 5.3**) All NADPH dependent reactions were carried out in sealed cuvettes containing 50 mM MOPS pH 7.5, 5 mM $MgCl_2$, 5 mM DTT and 0.2 mg/mL cell free extract. After addition of 0.2 mM NADPH (A_{340} ~ 1.0) and the relevant substrate (see below), NADPH oxidation was measured at 340 nm. The substrates for the E2, E2+E3, E7+E8+E2+E9, and E9 assays were succinyl-CoA, malonyl-CoA, methylmalonyl-CoA, and succinic semialdehyde (each 1 mM) respectively. In the case of E12, which performs the two steps hydration of crotonyl-CoA to 3-HB-CoA and subsequent oxidation to acetoacetyl-CoA, 1 mM NAD was added and A_{340} was monitored following the addition of crotonyl-CoA for the dehydration and oxidation or 3-HB-Coa for just the oxidation (1 mM each). For E4+E5+E6 and E4+E6, 3-HP and acrylate (1 mM each) were added respectively along with 1 mM ATP and CoASH, and NADPH oxidation was monitored at 340 nm. For E1+E2+E3 and E1+E7+E8+E2+E9, acetyl-CoA or propionyl-CoA was added (1 mM each), respectively, and NADPH oxidation was monitored at 340 nm. These carboxylation reactions also required the presence of 1 mM ATP, and 10 mM $NaHCO_3$.

E1 activity was measured by phosphate release. The assay contained 10 mM $NaHCO_3$, 1 mM ATP, and 1 mM acetyl-CoA. Samples (20 µL) were removed at 2-4 min, diluted with water (180 µl), and the BioVision (Mountain View, CA) phosphate assay reagent (20 µl) was added. The phosphate produced was calculated using a standard curve of concentration versus

absorption at 650 nm. E4 was similarly measured by the phosphate release using 1 mM each of 3-HP, CoASH, and ATP. E13 was measured using a similar assay except that the acetoacetyl-CoA (1 mM) dependent disappearance of CoASH (1 mM) was measured using Ellman's reagent (DTNB) (259).

For *in vitro* product detection (**Table 5.4**), 2.0 mg/L of a cell-free extract was used in sealed vials containing 100 mM MOPS pH 7.5, 5 mM $MgCl_2$, and 5 mM DTT. Co-substrates were added as necessary for the cycle and include: 5 mM ATP, 4 mM CoASH, 10 mM $NaHCO_3$, 4 mM NAD, and 4 mM ADP + P_i. The assays were incubated at 75°C for 6 minutes, 10 minutes, and 60 minutes for the assays using strains MW56, MW78, and MW98 respectively. 3-HP, 4-HB, and acetate were derivatized with 2-nitrophenyl hydrazine and analyzed via HPLC as previously described (192).

***In vivo* analysis of H_2/CO_2 incorporation into acetate or ethanol**

P. furiosus strains were grown at 95 or 90°C to mid log phase as described above. The cultures were then shifted to 75-65°C for up to a week and the headspaces were replaced with either 100% H_2, 50:50 H_2:CO_2, or 100% CO. Cell suspensions were also used where *P. furiosus* strains were grow at 95 or 90°C to mid log phase and shifted to 75°C for 16 hours to allow for enzyme expression. The cultures were harvested and suspended in sea-water based salts (202) with the addition of 100 mM MOPS (pH 7.5) and 0-50 g/L maltose under 100% H_2, 50:50 H_2:CO_2, or 100% CO. For the labeling experiment, $NaH^{13}CO_3$ was added to a final concentration of 10 g/L and the spent media was analyzed via GC-MS. The suspensions were then incubated at 75-

65°C for up to a week. Media samples were analyzed via GC-FID with an identical method as previously described (227).

Acknowledgements

This work was supported by the US Department of Energy as part of the Electrofuels Project of ARPA-E (grant number DE-AR0000081) and the National Science Foundation (grant number 1021RR166824).

Tables and Figures

Table 5.1 Genome region targets for gene insertion

Genome regions (GR) targeted for the insertion of genes or operons, heterologous or homologous, into the chromosome (180).

Genome Region	Length (bases)	Surrounding genes	Intergenic distance (bases)
1	168	PF0265 - PF0266	813
2	351	PF0435 - PF0436	782
3	108	PF0574 - PF0575	458
4	111	PF0738 - PF0739	1372
5	153	PF1232 - PF1233	646
6	263	PF1308 - PF1309	790
7	86	PF1420 - PF1421	396
8	174	PF0537 - PF0538	631
9	181	PF0646 - PF0647	665
10	111	PF2056 - PF2057	456

Table 5.2 Strains used and constructed in this study

P. furiosus strain	Expression construct(s)	GR	Purpose
MW004 (COM1c2)	NA	NA	Control
MW56	SP1	3	3-HP production
MW76	BL-CA-SP1	3	Effect of SP1+BLCA; 3-HP production
MW45	SP2A	1	SP2A pathway analyses
MW43	SP2B	2	SP2B pathway analyses
MW48	SP2B + SP2A	2, 1	SP2A+B pathway analyses
MW52	SP1 + SP2B	3, 2	SP1 and SP2B pathway analyses
MW54	SP1 + SP2A + SP2B	2, 1 3	SP1 + SP2 pathway analysis
MW78	SP2B + SP2A + BL-CA-SP1	2, 1 3	Effect of BL and CA on SP1 + SP2 pathways
MW80	SP2B + SP2A + BL-CA + SP1	2, 1 5, 3	BLCA+ SP1 + SP2 pathway analysis
MW84	SP1 + SP2A + SP2B + E10aE11	3, 1 2, 4	SP1 + SP2 + partial SP3 pathway analysis
MW100	SP2B + SP2A + BL-CA-SP1 + SP3(b)	2, 1 3, 4	SP3 pathway analyses for E10b
MW102	SP2B + SP2A + BL-CA-SP1 + SP3(c)	2, 1 3 + 4	SP3 pathway analyses for E10c
MW375	SP2B + SP2A + BL-CA-SP1 + SP3(d) + CODH	2, 1 3, 4 5	SP3 pathway analyses for E10c ^{13}CO conversion to ^{13}C-acetate
MW377	SP2B + SP2A + BL-CA-SP1 + $AdhA_{TX514}$-SP3(d) + CODH	2, 1 3 4 5	^{13}CO conversion to ^{13}C-ethanol

MW98	**SP2B** + **SP2A** + **BL-CA-SP1** + **SP3**(a)	2, 1 3, 4	SP123(E10a)BLCA analyses; **acetyl-CoA production**
MW125	**SP2B** + **SP2A** + **BL-CA-SP1,** ΔACSIα	2, 1 3 ΔPF1540	Effect of ACSIα on **4-HB production**
MW244	**SP2B** + **SP2A** + **BL-CA-SP1,** ΔACSIα + **SP3(E10a)**	2, 1 3 ΔPF1540 4	Effect of ACSIα on **acetyl CoA production**
MW220	**SP2B** + **SP2A** + **BL-CA-SP1** + **SP3**(a)-**EtOH-1**	2, 1, 3, 4	**EtOH production** from 3-HP/4-HB cycle (H_2/CO_2)
MW222	**SP2B** + **SP2A** + **BL-CA-SP1** + ΔACSIα **+ SP3**(a)- **EtOH-2**	2,1 3, ΔPF1540 4	**EtOH production** from 3-HP/4-HB cycle (H_2/CO_2)
MW203	**BL-CA-SP1** + ΔACSIα + **BuOH-1 +** **SP2ASP2BSP3**(a)	3, ΔPF1540 PF1540 1	**BuOH production** from 3-HP/4-HB cycle (H_2/CO_2) Application of SP2+SP3 super operon
MW351	**SP2B** + **SP2A** + **BL-CA-SP1** + **CODH** + **+ SP3**(a)	2, 1 3, 5 4	Acetyl-CoA production and autotrophic growth from 3-HP/4-HB cycle (H_2/CO_2/CO)
MW353	**SP2B** + **SP2A** + **BL-CA-SP1** + **CODH** + **+ SP3**(a)- **EtOH-3**	2, 1 3, 5 4	**EtOH production** from 3-HP/4-HB cycle (H_2/CO_2/CO)
Key	**SP1:** P_{slp}-E1αβγ-E2-E3 **BL-CA-SP1:** P_{pep}-BL-CA P_{slp}-E1αβγ-E2-E3 **SP2A:** P_{slp}-E4-E5-E6 **SP2B:** P_{slp}-E7-E8αβ-E9 **SP3(a/b/c/d)** (with E10a, b_{mut}, c, or d): P_{slp}-E10a/b_{mut}/c/d-E11-E12-E13		**EtOH-1:** P_{por}-$AdhA_{TX514}$ **EtOH-2:** P_{por}-$AdhE_{TX514}$ **EtOH-3:** P_{por}-$AdhE_{Teth}$ **BuOH-1:** P_{por}-Thl-Hbd-Crt_{Cteng}-Ter_{Stherm}-Bad-Bdh_{TX514} **CODH:** P_{mbh}-$CODH_{Ton}$ **ACSIα:** PF1540

Table 5.3 Enzymatic analysis of the *M. sedula* 3-HP/4-HB cycle for carbon fixation in *P. furiosus*

Rxn	Enzymes	*P. furiosus* strain	Substrates		*P. furiosus*	*M. sedula*
			Reaction initiated with	Analytes	Specific Activity (U/mg)	
A	E1+E2+E3	MW56 (SP1)	Acetyl-CoA	NADPH	0.07	0.02
B	E1	MW56 (SP1)	Acetyl-CoA	P_i	0.074	0.206
C	E2+E3	MW56 (SP1)	Malonyl-CoA	NADPH	0.24	0.08
D	E4	MW76 (SP1+SP2)	3-HP	P_i	0.06	0.08
E	E4+E6	MW76 (SP1+SP2)	Acrylate	NADPH	0.16	0.01
F	E4+E5+E6	MW76 (SP1+SP2)	3-HP	NADPH	0.17	0.01
G	E1+E7+E8 +E2+E9	MW76 (SP1+SP2)	Propionyl -CoA	NADPH	0.05	0.08
H	E7+E8+ E2+E9	MW76 (SP1+SP2)	Methylmalonyl -CoA	NADPH	0.4	0.16
I	E2	MW56 (SP1)	Succinyl-CoA	NADPH	0.13	0.08
J	E9	MW47 (SP2)	succinic semialdehyde	NADPH	8.6	0.48
K	E12a+E12b	MW98 (SP1+SP2+SP3)	Crotonyl -CoA	NADH	0.33	0.21
L	E12b	MW98 (SP1+SP2+SP3)	3-HB-CoA	NADH	0.19	0.14
M	E13	MW98 (SP1+SP2+SP3)	Acetoacetyl -CoA	CoASH	0.09	0.04

Table 5.4 *In vitro* product formation of the *M. sedula* 3-HP/4-HB cycle for carbon fixation in *P. furiosus*

Cell free extracts (2.0 mg/mL) of the indicated strains were incubated at 75°C in a sealed anaerobic vial with the indicated substrates and co-substrates with 5 mM $MgCl_2$ and 5 mM DTT in 100 mM MOPS pH 7.5. Cell free extracts of MW56 (SP1) were incubated for 6 minutes and produced 3-HP from acetyl-CoA. Cell free extracts of MW78 (SP12) were incubated for 10 minutes and produced 4-HB from succinic semialdehyde. Cell free extracts of MW98 (SP123) were incubated for 60 minutes and produced 3-HP from propionate, acrylate, and 3-HP.

Strain	Substrate	Co-Substrates	Enzymes	Detected
MW56 (SP1)	2 mM acetyl-CoA	10 mM $NaHCO_3$ 5 mM MgATP 5 mM NADPH	E1+E2+E3	0.5 mM 3-HP 1.5 mM acetate
MW78 (SP12)	1 mM succinic semialdehyde	2 mM NADPH	E9	1 mM 4-HB
MW98 (SP123)	1 mM propionate	5 mM ATP 4 mM CoASH 10 mM $NaHCO_3$ 4 mM NADPH 4 mM NAD 4 mM ADP + P_i	E4+E1+E7+E8+E2+E9+ E10+E11+E12+E13+ACS	0.1 mM acetate 0.9 mM propionate
	1 mM acrylate	5 mM ATP 4 mM CoASH 4 mM NADPH 10 mM $NaHCO_3$ 4 mM NAD 4 mM ADP + P_i	E4+E6+E1+E7+E8+E2+E9+ E10+E11+E12+E13+ACS	0.1 mM acetate 0.9 mM propionate trace acrylate
	1 mM 3-HP	5 mM ATP 4 mM CoASH 4 mM NADPH 10 mM $NaHCO_3$ 4 mM NAD 4 mM ADP + P_i	E4+E5+E6+E1+E7+E8+E2+E9+ E10+E11+E12+E13+ACS	0.1 mM acetate 0.7 mM propionate 0.2 mM 3-HP

Figure 5.1 Carbon dioxide assimilation in *Metallosphaera sedula*

A

The 3-Hydroxypropionate/4-Hydroxybutyrate carbon fixation cycle in Metallosphaera sedula (148-154). E1αβγ Msed_0147, 1048, 1375: Acetyl/propionyl-CoA carboxylase, E2 Msed_0709: Malonyl/succinyl-CoA reductase, E3 Msed_1993: Malonate semialdehyde reductase, E4 Msed_1456: 3-Hydroxypropionyl-CoA synthase, E5 Msed_2001: 3-Hydroxypropionyl dehydratase, E6 Msed_1426: Acryloyl-CoA reductase, E7 Msed_0639: Methylmalonyl-CoA epimerase, E8αβ Msed_0638, 2055: Methylmalonyl-CoA mutase, E9 Msed_1424: Succinate semialdehyde reductase, E10(a, b*, c, or d) Msed_0406, Msed_1353 *W424G, Msed_1394, or Msed_1422: 4-Hydroxybutyrate-CoA ligase, E11 Msed_1321: 4-Hydroxybutyrl-CoA dehydratase, E12(a or b) Msed_0399 or Msed_1423): Crotonyl-CoA hydratase/ (S)-3-Hydroxybutyrl-CoA dehydrogenase, and E13 Msed_0656: Acetoacetyl-CoA β-ketothiolase.

B

The stoichiometry of the cycle. The net production of one molecule of Acetyl-CoA by this cycle requires carbon from two molecules of HCO_3^-, the energy equivalent of hydrolysis of six ATP to ADP, five reducing equivalents in the form of NADPH (one regenerated in the form of NADH), and one molecule of Coenzyme A.

Figure 5.1

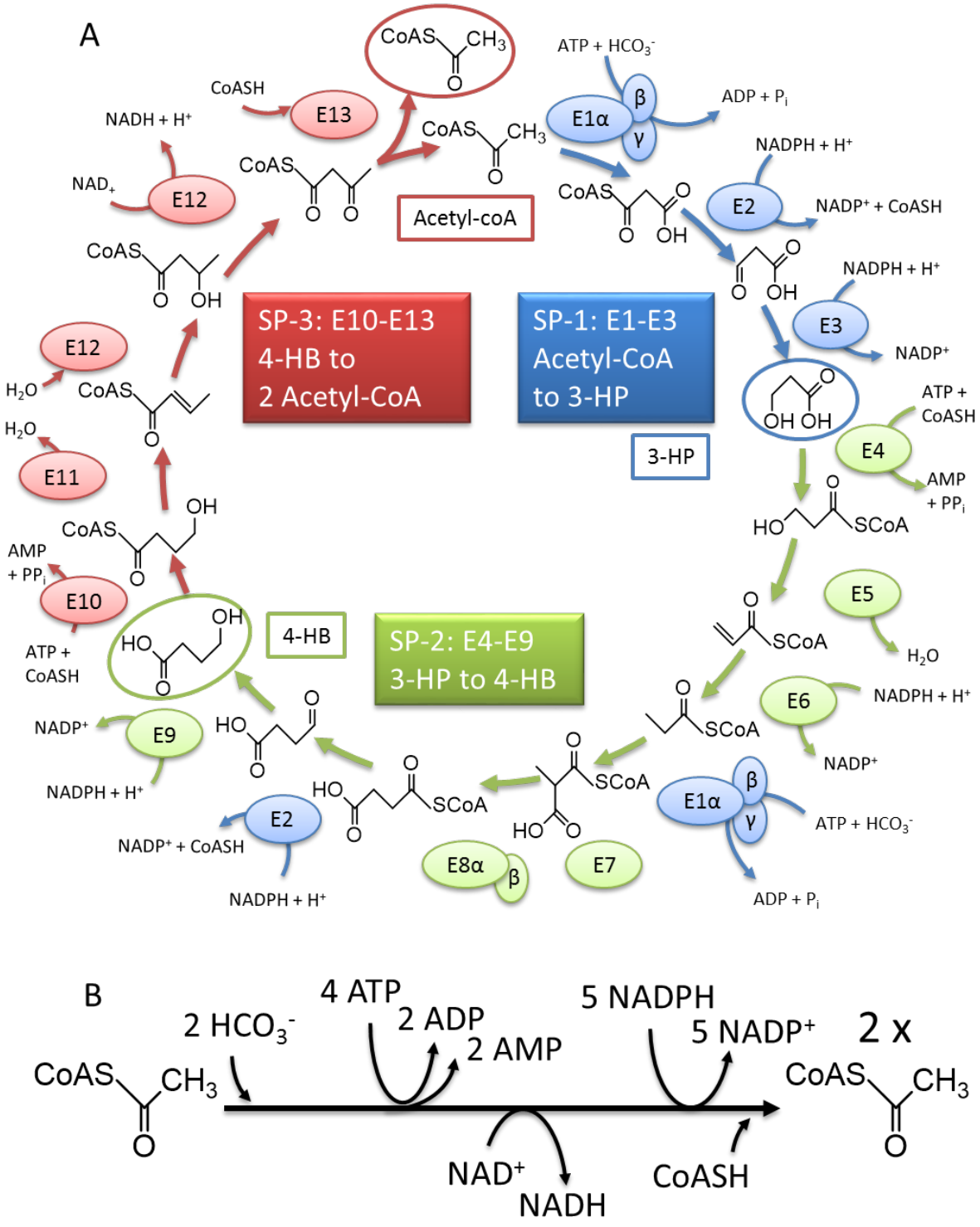

A
SP-3: E10-E13
4-HB to
2 Acetyl-CoA
SP-1: E1-E3
Acetyl-CoA
to 3-HP
SP-2: E4-E9
3-HP to 4-HB
Acetyl-coA
3-HP
4-HB
E1α
β
γ
E2
E3
E4
E5
E6
E7
E8α
E9
E10
E11
E12
E13
ATP + HCO3-
ADP + Pi
NADPH + H+
NADP+ + CoASH
NADP+
ATP +
CoASH
AMP
+ PPi
H2O
NADH + H+
NAD+
CoASH
B
2 HCO3-
4 ATP
2 ADP
2 AMP
5 NADPH
5 NADP+
2 x
NAD+
NADH
CoASH

Figure 5.2 Construction of *P. furiosus* strain MW98 to express the complete *M. sedula* 3-HP/4-HB cycle for carbon dioxide assimilation

The complete *M. sedula* 3-HP/4-HB cycle was assembled into four artificial operons for expression in *P. furiosus* strain MW98.

Figure 5.2

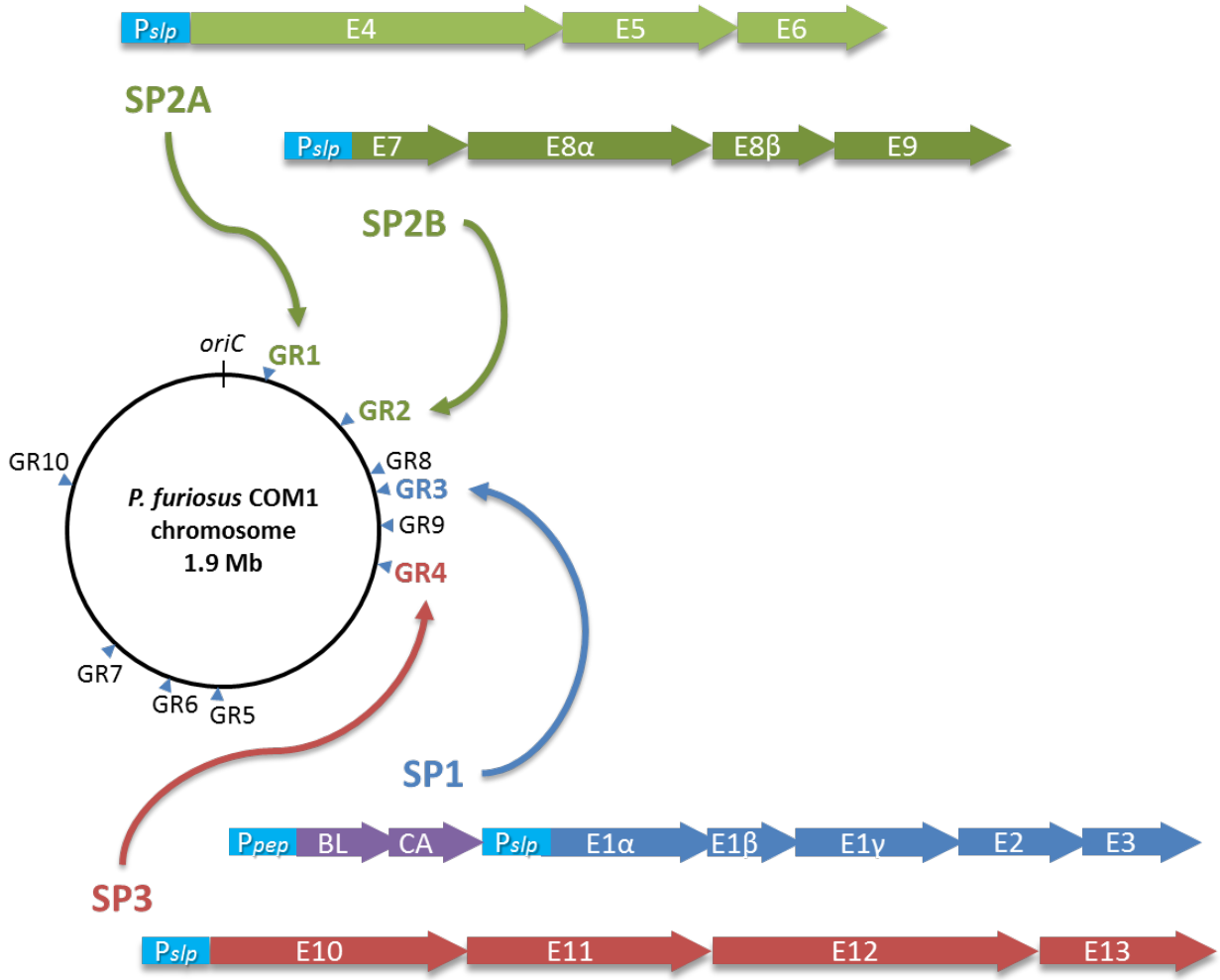

Pslp
E4
E5
E6
SP2A
Pslp
E7
E8α
E8β
E9
SP2B
oriC
GR1
GR2
GR10
GR8
GR3
GR9
GR4
P. furiosus COM1
chromosome
1.9 Mb
GR7
GR6
GR5
SP1
Ppep
BL
CA
Pslp
E1α
E1β
E1γ
E2
E3
SP3
Pslp
E10
E11
E12
E13

Figure 5.3 The *pyrF* marker is used for iterative selections in the absence of uracil and counterselections in the presence of 5-FOA

The expression construct and *pyrF* marker are incorporated into the Δ*pyrF* acceptor strain and purified by plating and sub-culturing in media lacking uracil. The *pyrF* marker is subsequently removed along homologous regions 65 bases in length in the presence of 5-FOA. The resulting strain still contains the first expression construct as is now a Δ*pyrF* acceptor strain for subsequent genetic manipulations. (137)

Figure 5.3

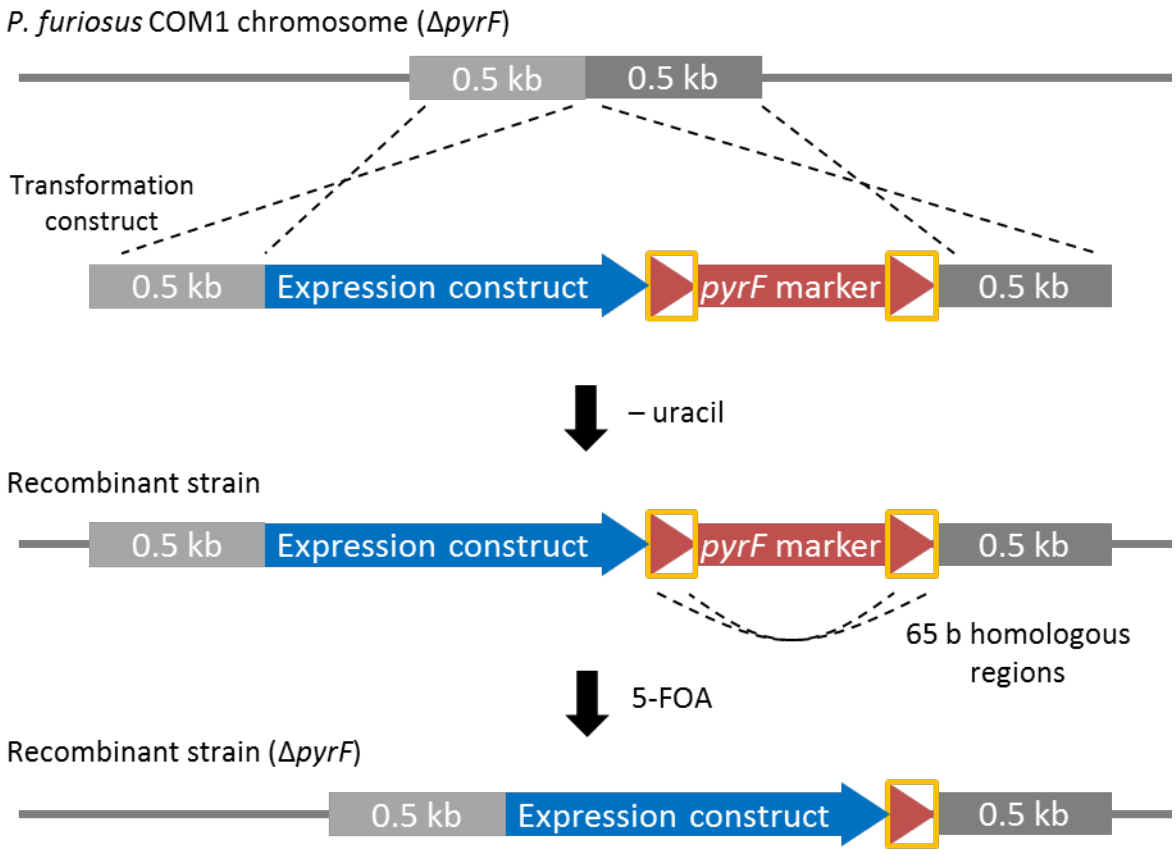

P. furiosus COM1 chromosome (ΔpyrF)
0.5 kb
0.5 kb
Transformation construct
0.5 kb
Expression construct
pyrF marker
0.5 kb
– uracil
Recombinant strain
0.5 kb
Expression construct
pyrF marker
0.5 kb
65 b homologous regions
5-FOA
Recombinant strain (ΔpyrF)
0.5 kb
Expression construct
0.5 kb

Figure 5.4 Quantitative PCR of the 3-HP/4-HP genes in *P. furiosus* strain MW98

The expression of 18 foreign genes in *P. furiosus* strain MW98 was analyzed by qPCR following growth at 90°C to mid log phase (red) and 16 hours following a temperature shift to 72°C (blue).

Figure 5.4

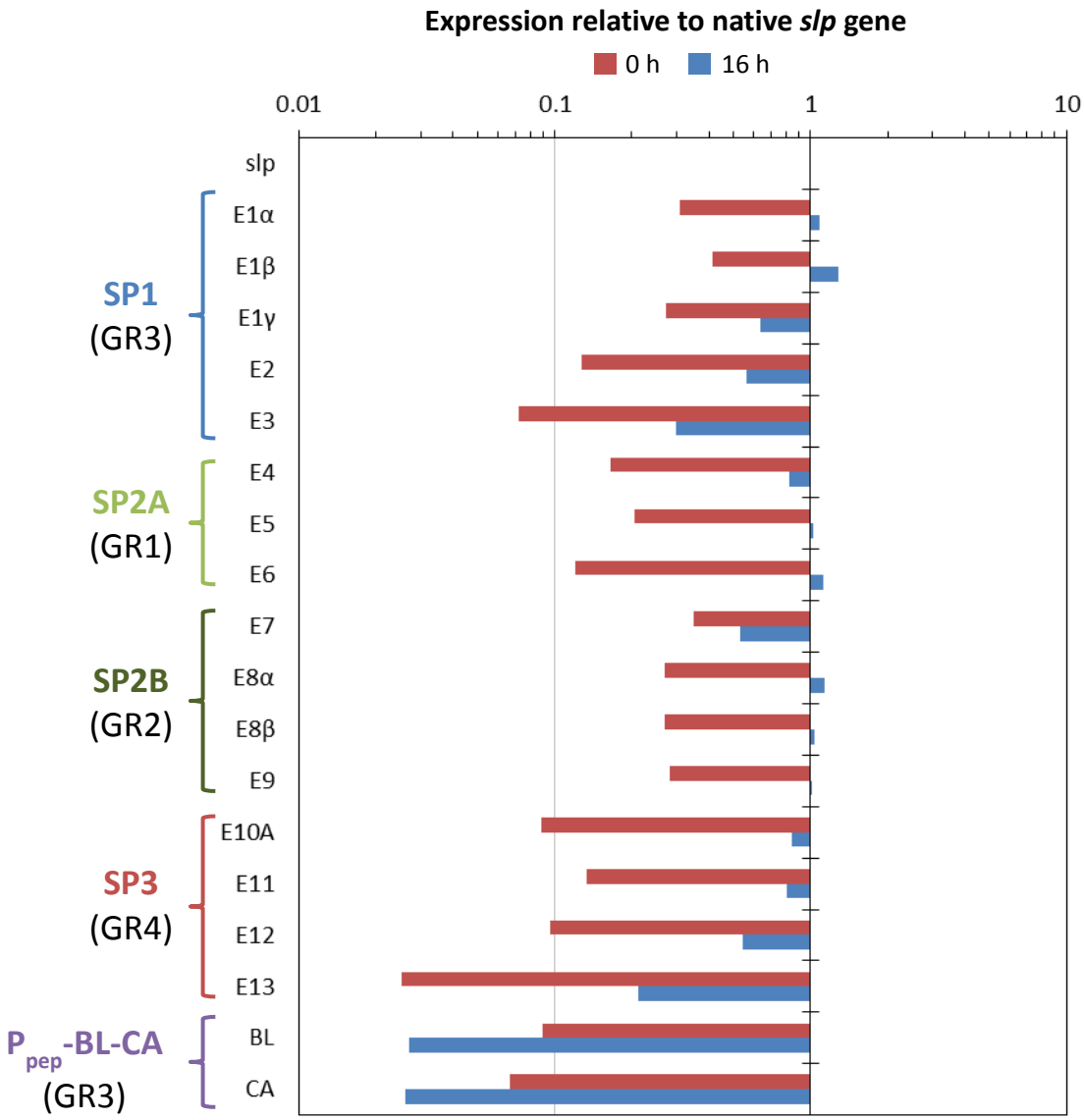
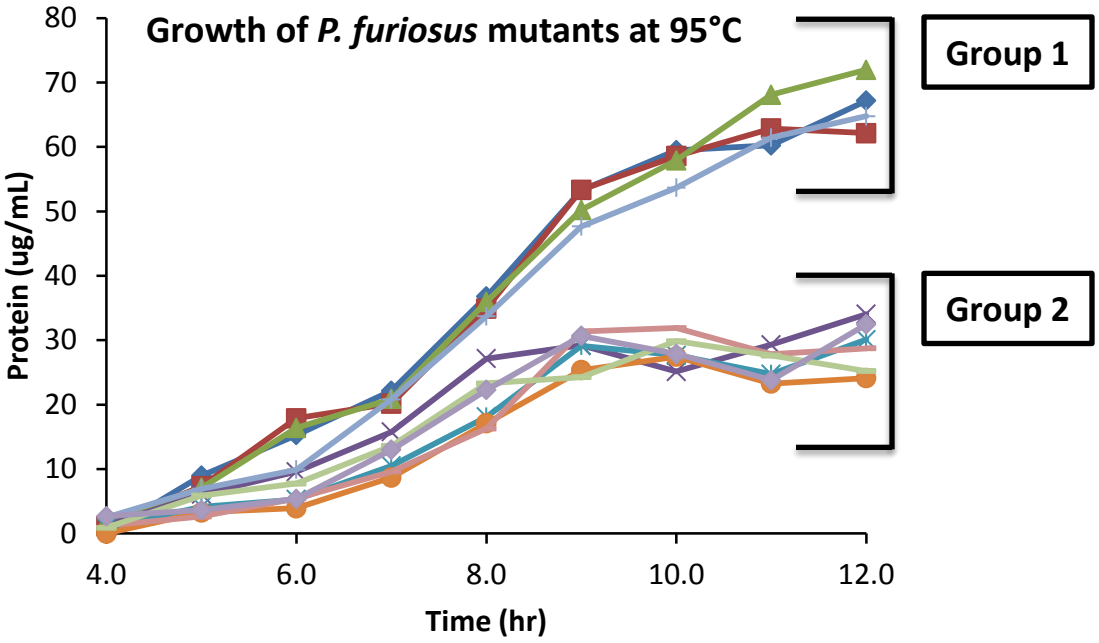

Expression relative to native *slp* gene
0 h
16 h
0.01
0.1
1
10
slp
E1α
E1β
E1γ
E2
E3
E4
E5
E6
E7
E8α
E8β
E9
E10A
E11
E12
E13
BL
CA
SP1
(GR3)
SP2A
(GR1)
SP2B
(GR2)
SP3
(GR4)
P pep -BL-CA
(GR3)

Figure 5.5 Growth of 3-HP/4-HB cycle containing *P. furiosus* strains

P. furiosus strains containing two or more S-layer promoter driven artificial operons have a defect in growth.

Figure 5.5



Group 1
MW004: COM1c2
MW56: SP1
MW45: SP2A
MW43: SP2B

Group 2
MW48: SP2A+SP2B
MW52: SP1+SP2B
MW54: SP1+SP2A+SP2B
MW78: SP1+SP2A+SP2B
MW80: BLCA+SP1+SP2A+SP2B
MW84: SP1+SP2A+SP2B+E10aE11

Figure 5.6 *In vitro* analysis of the 3-HP/4-HB enzymes in *P. furiosus* (red) and in autotrophically grown *M. sedula* (blue) cell free extracts

A) The *M. sedula* 3-HP/4-HB cycle for carbon dioxide assimilation is labelled with the enzymatic assays that have been performed.

B) Enzymatic analysis of the 3-HP/4-HB enzymes in *P. furiosus* (red) and in autotrophically grown *M. sedula* (blue) cell free extracts

Figure 5.6

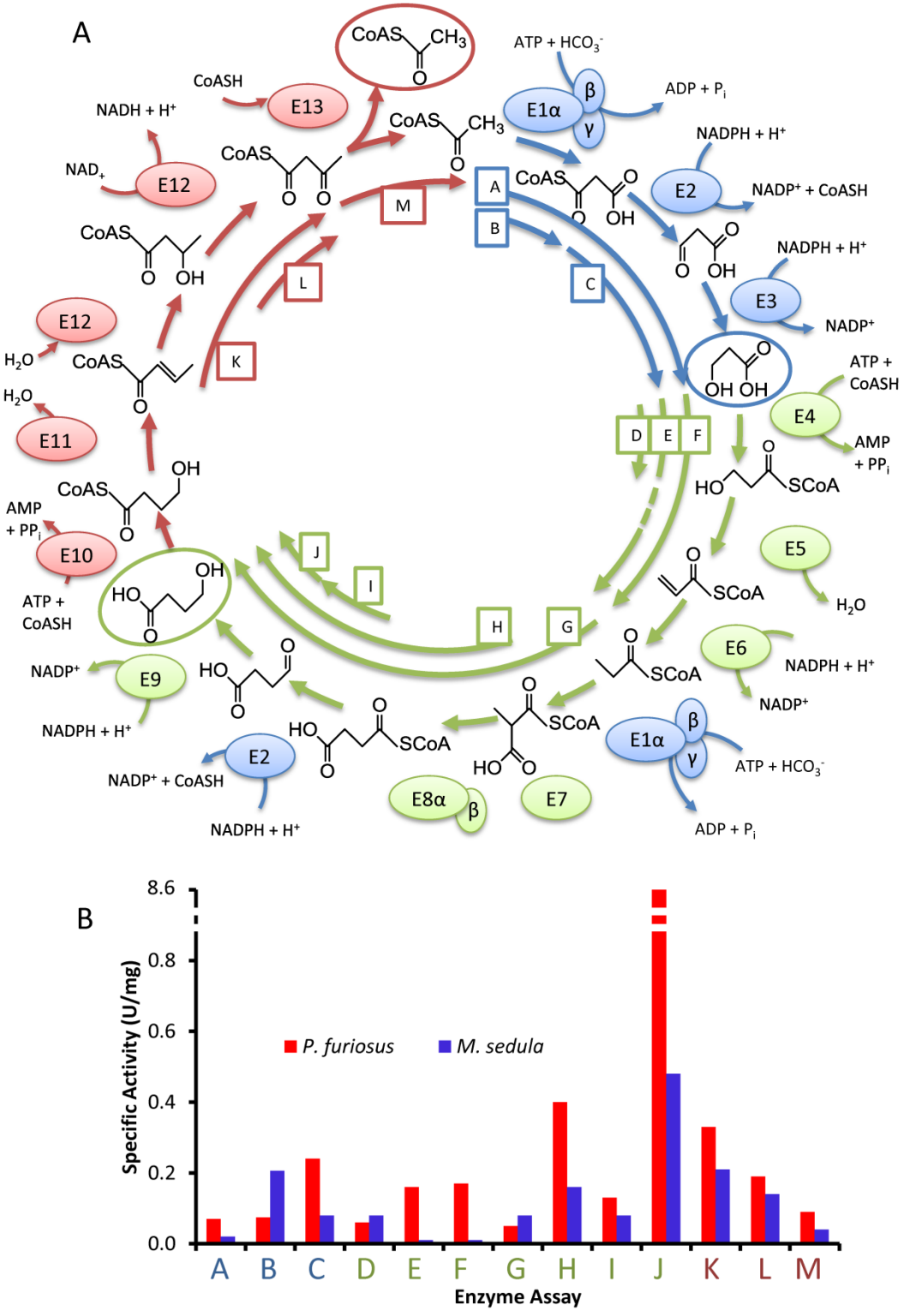

A
CoAS CH3
O
CoASH
E13
NADH + H+
NAD+
E12
CoAS
O O
CoAS CH3
O
M
CoAS
O OH
L
E12
H2O
CoAS
K
H2O
E11
CoAS
OH
O
AMP
+ PPi
E10
ATP +
CoASH
OH
HO
O
J
I
H
G
NADP+
E9
NADPH + H+
HO
O
HO
O
SCoA
O
E2
NADP+ + CoASH
NADPH + H+
E8α
β
SCoA
HO
O
E7
E1α
β
γ
ATP + HCO3-
ADP + Pi
ATP + HCO3-
E1α
β
γ
ADP + Pi
NADPH + H+
E2
NADP+ + CoASH
A
B
CoAS
O OH
C
O
O OH
NADPH + H+
E3
NADP+
OH OH
D
E
F
ATP +
CoASH
E4
AMP
+ PPi
HO
O
SCoA
O
SCoA
E5
H2O
E6
NADPH + H+
NADP+
B
8.6
0.8
0.6
0.4
0.2
0.0
Specific Activity (U/mg)
P. furiosus
M. sedula
A B C D E F G H I J K L M
Enzyme Assay

Figure 5.7 *In vitro* product formation in *P. furiosus* strain MW98

In vitro detection of acetate from cycle intermediates in *P. furiosus* strain MW98. All assays contained the accessory substrates ATP, HCO_3^-, NADPH, CoASH, and NAD. Acetate can be detected in greater amounts when cell free extracts are also incubated with 3-HP, acrylate, or propionate compared to the control which contains only the accessory substrates.

Figure 5.7

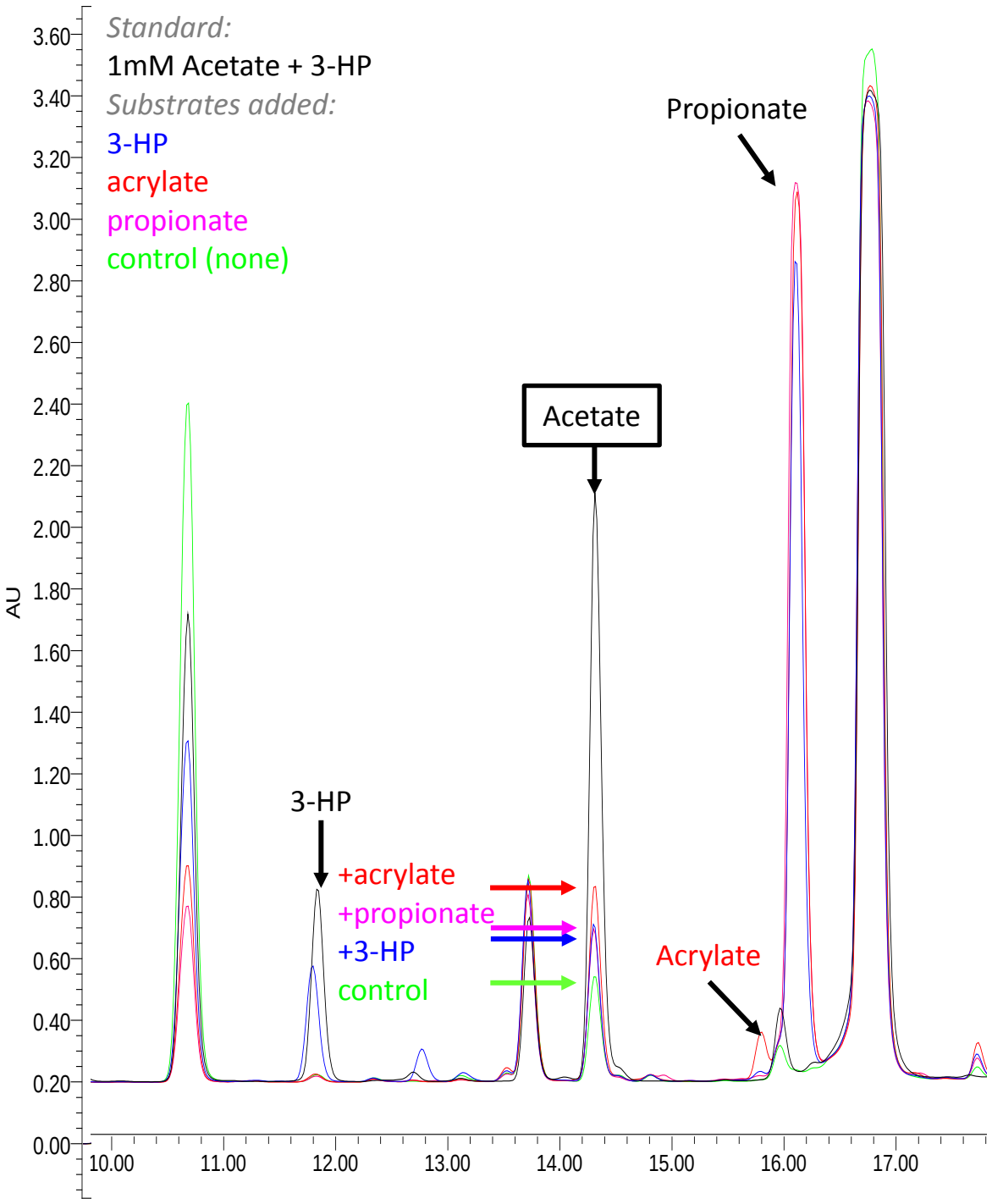
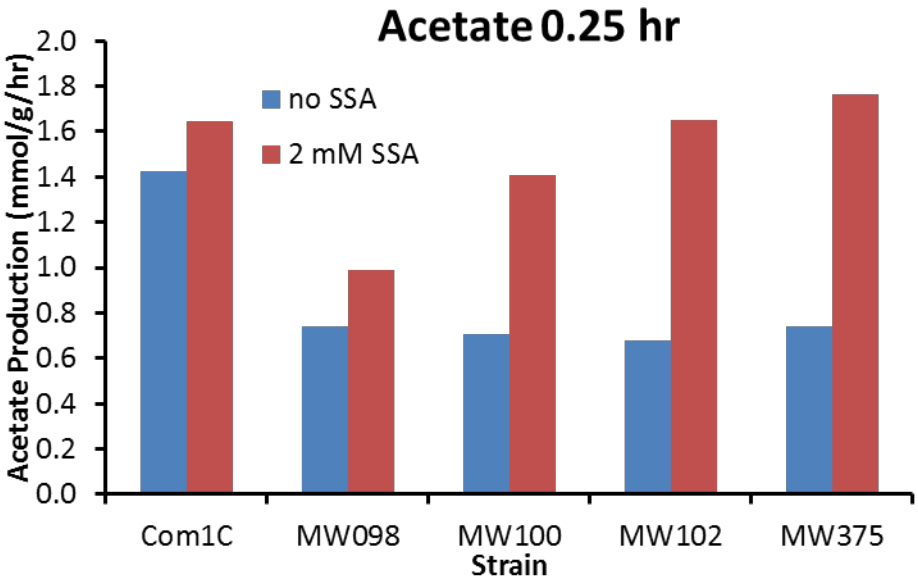

Standard:
1mM Acetate + 3-HP
Substrates added:
3-HP
acrylate
propionate
control (none)
Propionate
Acetate
3-HP
+acrylate
+propionate
+3-HP
control
Acrylate
AU
3.60
3.40
3.20
3.00
2.80
2.60
2.40
2.20
2.00
1.80
1.60
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
10.00
11.00
12.00
13.00
14.00
15.00
16.00
17.00

Figure 5.8 *In vitro* acetate production from succinic semialdehyde for four candidates for E10

Cell free extracts of the indicated strains were incubated at 75°C with and without succinic semialdehyde and analyzed for acetate production. The co-substrates NADPH, ATP, CoASH, NAD, ADP, and P_i were added for the coupled reaction of E9+**E10**+E11+E12+E13+ACS.

Figure 5.8


Acetate 0.25 hr
Acetate Production (mmol/g/hr)
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
no SSA
2 mM SSA
Com1C
MW098
MW100
MW102
MW375
Strain

CHAPTER 6

DISCUSSION

Here, *P. furiosus* has been successfully engineered to produced 3-hydroxypropionate (3-HP), butanol, and ethanol from acetyl-CoA (**Figure 6.1**). Further studies will be needed to improve our understanding of how these pathways work within the context of *P. furiosus* in order to improve product yield and better understand the metabolism of *P. furiosus*.

The production of 3-HP by *P. furiosus* was intended to be a stepping stone toward engineering a complete carbon dioxide fixation pathway into *P. furiosus*. 3-HP itself, however, is a valuable target molecule due to its ability to displace a large petroleum based market. In addition to direct polymerization, 3-HP can easily be dehydrated to acrylate. Currently acrylate is produced from the partial oxidation of propene, which is derived from petroleum (174). The renewable production of 3-HP, and subsequently acrylate, could have a positive impact on an extremely broad range of market opportunities all of an extremely large market size. In order to improve the production of 3-HP, a competing pathway was blocked (**Figure 6.2**). This involved acetyl-CoA synthetase (ACS), which is the terminal step in the conversion of glucose to acetate by *P. furiosus* where the high energy thioester bond acetyl-CoA is hydrolyzed in order to produce ATP. The deletion of ACSIα resulted in an extended lag phase, decreased acetate production, and a decrease in the acetyl-CoA dependent CoA release activity in cell free extracts. In strains containing the first three enzymes of the *M. sedula* 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) cycle for carbon dioxide fixation known as sub-pathway 1 (SP1),

the deletion of ACSIα resulted in a 3x fold increase in 3-HP production under identical conditions (201).

Accessory genes to SP1 were also considered. Carbonic anhydrase (CA, Msed_0390) and biotin protein ligase (BL, Msed_2010) were expressed along with SP1 in *P. furiosus* strain MW76. CA was added to catalyze the interconversion of CO_2 with bicarbonate, since bicarbonate is suspected to be the substrate for the carboxylase, and BL was added to aid in the assembly of biotin onto the β-subunit of the carboxylase. While the original study demonstrated the requirement of gaseous H_2 and CO_2 *in vitro*, the effect of their availability *in vivo* was not explored (192). Therefore, mass transfer was considered and sparging and agitation were optimized. An agitation rate of 400 rpm and gas flow rate of 50 mL/min N_2/CO_2 was found to be optimum and resulted in a 15x improvement over the original yield to 0.28 g/L 3-HP. Increasing the agitation and gas flow with a small pore micro-bubbler increased the rate of 3-HP production to 11 mg/L/hr, but left the final tier approximately the same. This titer of 3-HP appears to be accompanied with a sudden loss of viability which is suspected to be the result of toxicity due to intracellular 3-HP or other intermediates such as malonate semialdehyde since COM1 in un affected by even 20 g/L extracellular 3-HP (260). Since no 3-HP exporters are known to exist, developing such an exporter, which is no trivial task, is likely essential to further improve 3-HP productivity and yield.

Additionally, the sub-pathway that produces 3-HP, as well as the entire cycle, requires electrons in the form of NADPH, which can be replenished from H_2 via soluble hydrogenase 1 (SH1). This compatibility was of deliberate design since SH1 is native to *P. furiosus* and its homologous overexpression has already been demonstrated (176). Overexpressing SH1 within

the 3-HP or full cycle strains would result in faster NADPH replenishment from H_2. This process assumes a ready and renewable supply of hydrogen gas, which is currently produced primarily from natural gas. Fortunately, there are a multitude of options for the renewable production of hydrogen gas. The most straightforward method is the photovoltaic powered electrolysis of water which is capable of capturing up to 15% of the incident solar radiation and storing that energy in the chemical bond of molecular hydrogen (46). Hydrogen can also be produced biologically. *In vivo* fermentation of carbohydrates has a maximum theoretical yield of 4 H_2/glucose. Alternatively, *in vitro* enzymatic degradation of starch to H_2 and CO_2 yields up to 12 H_2/glucose (48). The need to transport and store the hydrogen would be minimal since it would be generated on site for immediate use as an energy carrier. In a similar fashion, H_2 has been used as an energy carrier in order to link the production of electricity from a solar panel to the synthesis of a biofuel in the bacterium *Ralstonia eutropha* (261).

Butanol is a target molecule not only because of its desirable fuel characteristics, but its wide range of applications as a commodity chemical. A new challenge arose in the effort to produce butanol in *P. furiosus* in that there was not a sufficiently thermophilic pathway to utilize. Instead, a hybrid synthetic pathway was assembled from three thermophilic bacteria with gene candidate identified by homology (227). The *P. furiosus* butanol strains BuOH-1 and BuOH-2 were also constructed in a ΔACSIα background to direct acetyl-CoA into the pathway rather than to acetate (201). Still, the production of butanol *in vivo* could only occur at sufficient rates at 40°C below the optimal growth temperature of the host and in concentration cell suspensions (227). In order to better understand the kinetic limitations of this pathway and in an effort to overcome them, the pathway enzymes were individually expressed and purified

in *E. coli* and assembled *in vitro*. The first four steps of the pathway are responsible for the C4 synthesis as they convert acetyl-CoA to butyryl-CoA. The last two steps are the ALDH+ADH conversion of the CoA ester to the alcohol (**Figure 6.3**). When all the enzymes are combined *in vitro* to equal activities, the alcohol formation segment competes with the C4 formation segment for acetyl-CoA and ethanol is produced as a product along with butanol. This is due to poor specificity of C4 over C2 substrates by the ALDH+ADH portion of the pathway, and the greater availability of acetyl-CoA over butyryl-CoA in the assay. To compensate for this, the C4 synthesis portion of the pathway can be added in excess compared to the ALDH+ADH portion. The result is a greater butanol/ethanol yield due to the greater availability of butyryl-CoA. The downside is a decreased rate of alcohol formation since the ALDH+ADH portion of the pathway is now kinetically limiting (230). Clearly, the limiting factor is this strategy is the butyraldehyde/butanol dehydrogenases, and alternate gene donors, such as the related and more thermophilic organism *Thermoanaerobacter ethanolicus*, and/or directed evolution to improve the kinetics and specificity of these enzymes could improve the yield of butanol in *P. furiosus*.

Unlike 3-HP and butanol, small amounts of ethanol are produced natively by *P. furiosus* (1, 103). There are two routes for acetaldehyde production and multiple alcohol dehydrogenases (ADHs) in the genome that could convert this acetaldehyde to ethanol in *P. furiosus* (**Figure 6.4**). Pyruvate ferredoxin oxidoreductase (POR) from *P. furiosus* and *Thermococcus guaymasensis* have been shown to non-oxidatively decarboxylate pyruvate to acetaldehyde. This side reaction resembles the activity of pyruvate decarboxylase, an enzyme that is not found in extreme thermophiles. The direct evidence for this reaction is limited to *in*

vitro analysis under a non-physiological pH extreme (139, 246). Acetaldehyde has also been shown to be produced from acetate via aldehyde ferredoxin oxidoreductase (AOR) (1). The reduction of acetate to acetaldehyde is unexpected as the standard reduction potential of this reaction is -0.60 volts. The electron donor of this reaction is ferredoxin which has a standard reduction potential of -0.42 volts. Despite the use of this strong reductant, which is available from the strongly reducing anaerobic metabolism of the host, this reaction alone is still quite unfavorable. In *P. furiosus*, this reaction is made possible when the AdhA from *Thermoanaerobacter* was heterologously expressed. In addition to having favorable kinetics, AdhA utilized the electron donor NADPH which has a standard reduction potential of -0.32 volts. Under physiological conditions, the reduction potential of NADPH is closer to -0.37 volts (235). This, along with the modest reduction potential of acetaldehyde to ethanol of -0.20 volts balances the reduction potential of the entire reaction.

The deletion of AOR in the presence of AdhA, as expected, disrupts the AOR+AdhA pathway in *P. furiosus* and results in severely reduced ethanol production (1). However, the deletion of *aor* in a COM1 background does not disrupt native ethanol production. The opposite is actually true as ethanol production increases ~50% in Δ*aor* versus COM1 (Keller and Lipscomb, unpublished data). This indicates that in a native context, AOR performs the previously proposed physiological role of detoxifying aldehydes to carboxylic acids (117, 236, 262). When deleted, the increased availability of acetaldehyde results in slightly increased ethanol production. This indicates AOR is not the dominant native route for native acetaldehyde production, and AOR only generates acetaldehyde in the presence of a highly

active ADH like AdhA. Additionally, the deletion of AOR results in an 80% reduction of dye-linked acetaldehyde oxidoreductase activity (Keller and Lipscomb, unpublished data). Even with the deletion of AOR, it remains possible that acetate is reduced to acetaldehyde by other members of the tungsten-containing AOR family including formaldehyde ferredoxin oxidoreductase (FOR) (231) and two other partially characterized tungsten-containing oxidoreductases (WOR 4 and WOR5) (232, 233). Directly demonstrating that POR is the route for native acetaldehyde production would be difficult as knocking out this central glycolytic enzyme is not expected to be possible. Further alternatives can be ruled out by knocking out FOR, WOR1, and WOR2 which have overlapping substrate specificities, specifically acetaldehyde, with AOR (231-233). By also knocking out these alternative oxidoreductases, acetaldehyde production from acetate would be completely eliminated and any acetaldehyde, and subsequently ethanol, produced would be solely due to the non-oxidative decarboxylation of pyruvate by POR.

P. furiosus contains four ADHs which are identified by the order in which they appear in the genome: AdhA (PF0074), AdhB (PF0075), AdhC (PF0608), and AdhD (PF0991). Only two ADHs from *P. furiosus*, AdhA and AdhD/TDH, have been characterized and they are both secondary ADHs (126, 127). AdhB and AdhC are candidates for the ethanol-producing enzyme due to their homology to a characterized primary ADH in *Thermococcus paralvinella* (f. *Thermococcus* sp. strain ES1) (128, 129). Microarray data, and more recently qPCR analysis of *P. furiosus* (**Figure 6.5**), reveal that *adhC* is expressed at a much higher level than *adhB*. The knockout of AdhC in COM1 resulted in a substantial decrease in ethanol production (0.1 mole

ethanol per mole glucose for Δ*adhC* vs 0.4 mole ethanol per mole glucose for COM1) which indeed confirms that AdhC is the dominate ethanol-producing enzyme in *P. furiosus* (Keller and Lipscomb, unpublished data). It would be very interesting to overexpress AdhC in the knockout of *P. furiosus*. First, it will confirm its role in ethanol production by demonstrating the rescue of the phenotype. Second, it should enhance ethanol production ~20°C above its boiling point. And finally it will afford an opportunity to further study the roles of POR and AOR in acetaldehyde production. Since *P. furiosus* AdhC and *Thermoanaerobacter* AdhA are both NADPH dependent primary alcohol dehydrogenases, the overexpression of AdhC is expected to have a similar result as the overexpression of AdhA where the initiation of the AOR route for acetaldehyde is accompanied by a dramatic increase in ethanol production (**Figure 6.4**). In addition, if AdhC has a broad substrate specificity, the conversion of exogenous acids to their corresponding alcohols should occur and should be disrupted with the deletion of AOR (1). If the deletion of AOR (as well as FOR, WOR1, and WOR2) result in COM1 levels of ethanol production despite the overexpression of AdhC, this would further indicate that POR is the limiting factor in COM1 ethanol production (**Figure 6.4**).

The current economic and environmental need for renewable fuels and commodity chemicals is undeniable. The energy and products that define modern society are frighteningly dependent upon and dominated by fossil fuels. It is critical that we find alternatives to support civilization as fossil fuel supplies dwindle and environmental impacts grow in severity. Improving the technology behind renewable processes is critical to driving this transition, and the vastly diverse metabolism of microorganisms is rich with opportunity for discovering better

technology. The science presented here, initiated by the Electrofuels program of the U.S. Department of Energy, represent a step in the right direction. Superior fuel targets will ease the transition for consumers as they are compatible with current vehicles and infrastructure. Efficient routes for production not only decrease the carbon footprint of the fuel, but reduce land use, water use, and competition with food production. Additionally, the intermediates of these strategies can be a resource for commercialization. If 3-HP can be economically produced in a similar fashion as presented here, it can not only be a renewable means to produce this important commodity chemical, but a precious revenue stream during the scale up process of a hypothetical Electrofuels company. As this project, and the myriad of others, continue to develop, it is inevitable that the technology created will make fossil fuels obsolete.

Figures

Figure 6.1 Goal of this work

P. furiosus has been engineered to produce 3-HP (192), butanol (227), and ethanol from acetyl-CoA (Chapter 4).

Figure 6.1

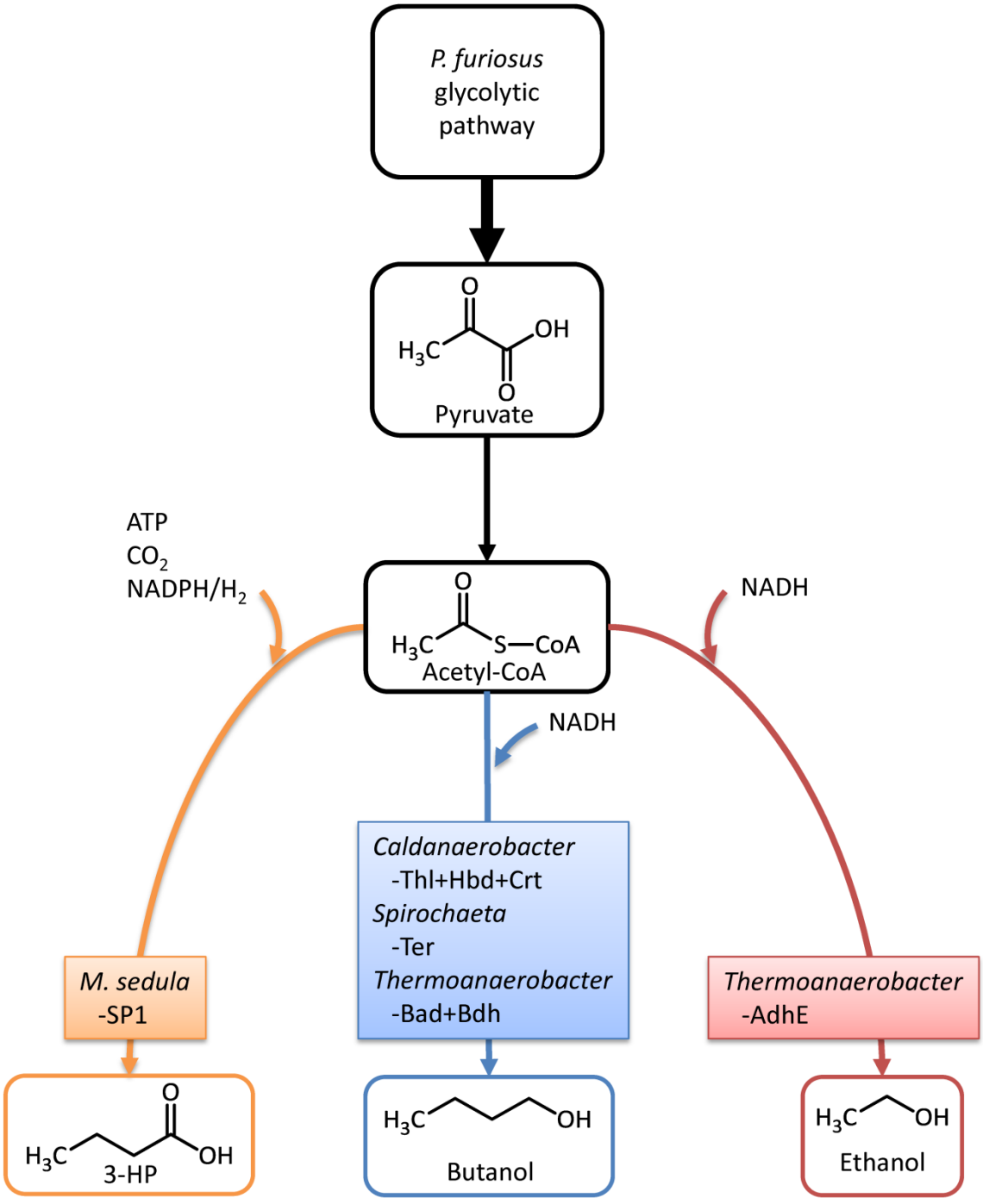

P. furiosus
glycolytic
pathway
Pyruvate
ATP
CO_2
$NADPH/H_2$
NADH
Acetyl-CoA
NADH
Caldanaerobacter
-Thl+Hbd+Crt
Spirochaeta
-Ter
Thermoanaerobacter
-Bad+Bdh
M. sedula
-SP1
Thermoanaerobacter
-AdhE
3-HP
Butanol
Ethanol

Figure 6.2 Deletion of a competing pathway improves 3-HP production

M. sedula SP1 has to compete with *P. furiosus* ACS for acetyl-CoA, which is typically converted to acetate by to generate ATP. The deletion of ACS, specifically ACSIα, increase the availability of acetyl-CoA and improves 3-HP yield (201).

Figure 6.2

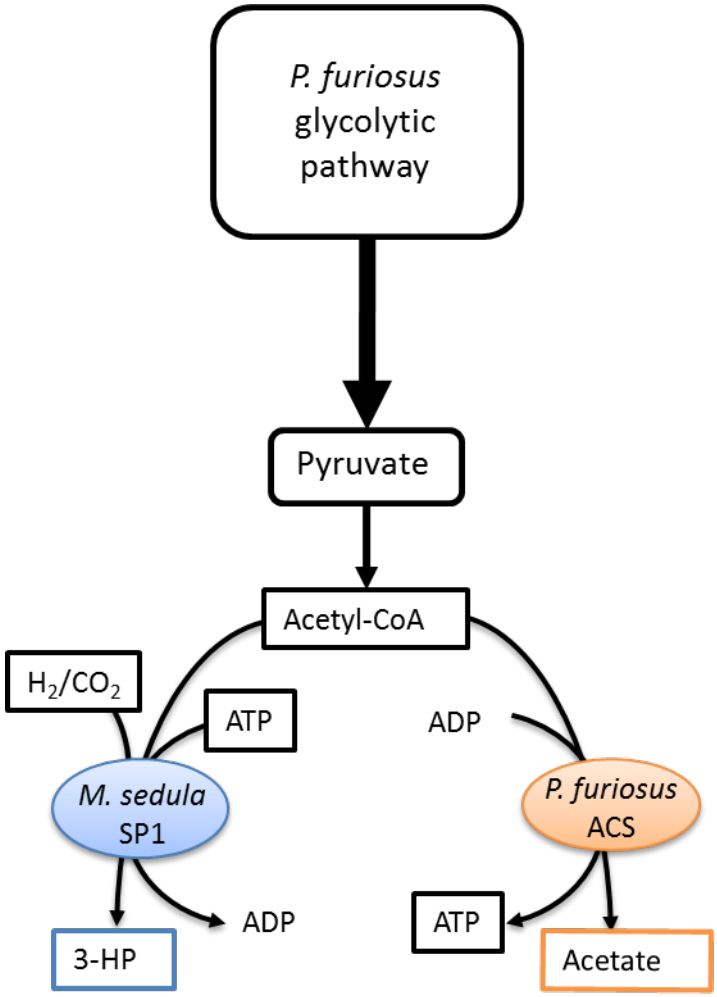

P. furiosus glycolytic pathway
Pyruvate
Acetyl-CoA
H_2/CO_2
ATP
ADP
M. sedula SP1
P. furiosus ACS
ADP
ATP
3-HP
Acetate

Figure 6.3 Kinetic modelling of the BuOH-1 pathway

The alcohol formation segment of the BuOH-1 pathway competes with the C4 formation segment for acetyl-CoA. Providing an abundance the of C4 formation segment results in a more favorable ratio of butanol/ethanol but limits the total productivity of the pathway (230).

Figure 6.3

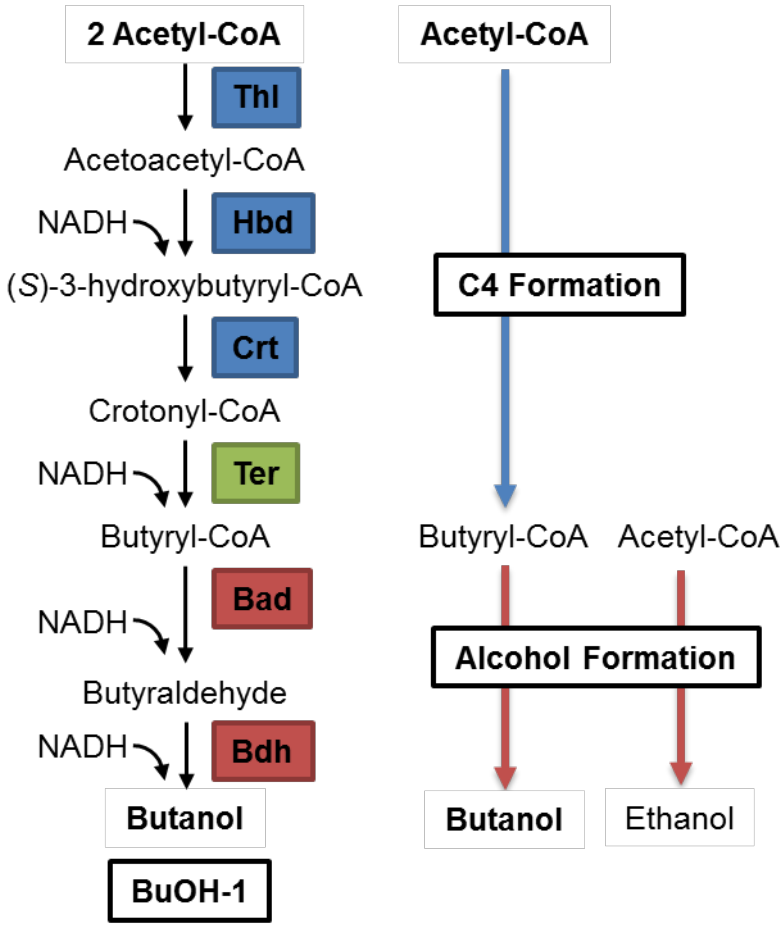

2 Acetyl-CoA
Thl
Acetoacetyl-CoA
NADH
Hbd
(S)-3-hydroxybutyryl-CoA
Crt
Crotonyl-CoA
NADH
Ter
Butyryl-CoA
NADH
Bad
Butyraldehyde
NADH
Bdh
Butanol
BuOH-1
Acetyl-CoA
C4 Formation
Butyryl-CoA
Acetyl-CoA
Alcohol Formation
Butanol
Ethanol

Figure 6.4 Three routes for ethanol formation in *P. furiosus*

Native (solid arrows) and engineered (dashed arrows) routes for ethanol production in *P. furiosus*. Glycolysis to pyruvate (0.5 mol glucose) yields no ATP and produces 2 Fd_{red} (net 2 e^-). A) Ethanol production from pyruvate via the pyruvate decarboxylase activity of pyruvate ferredoxin oxidoreductase (POR) and *P. furiosus* AdhC results in no net ATP, produces 2 Fd_{red} and uses 1 NADPH (net 2 e^-). B) Ethanol production from acetyl-CoA via AdhE yields no net ATP, produces 4 Fd_{red} and uses 1 NADPH (net 0 e^-). C) Ethanol production from acetate via aldehyde ferredoxin oxidoreductase (AOR, here also representing other oxidoreductases with aldehyde oxidizing activity), and an ADH results in 1 net ATP via acetyl-CoA synthase (ACS), produces 2 Fd_{red} and uses 1 NADPH (net 0 e^-). *P. furiosus* enzymes are shown in orange, AdhA is shown in red, and AdhE is shown in blue. Physiological *P. furiosus* reactions are shown with a solid arrow, and engineered reactions in recombinant strains are shown with a dashed arrow.

Figure 6.4

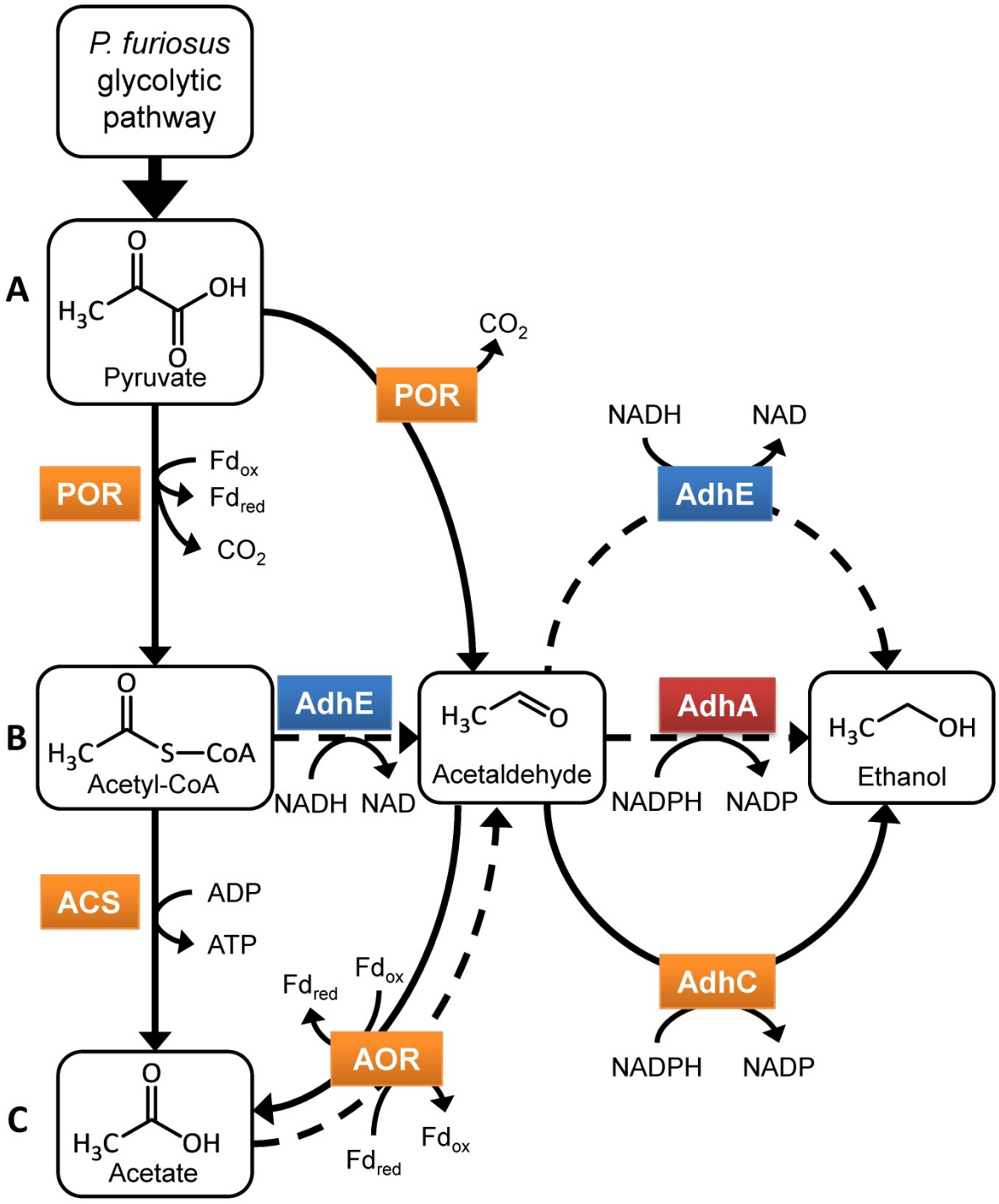
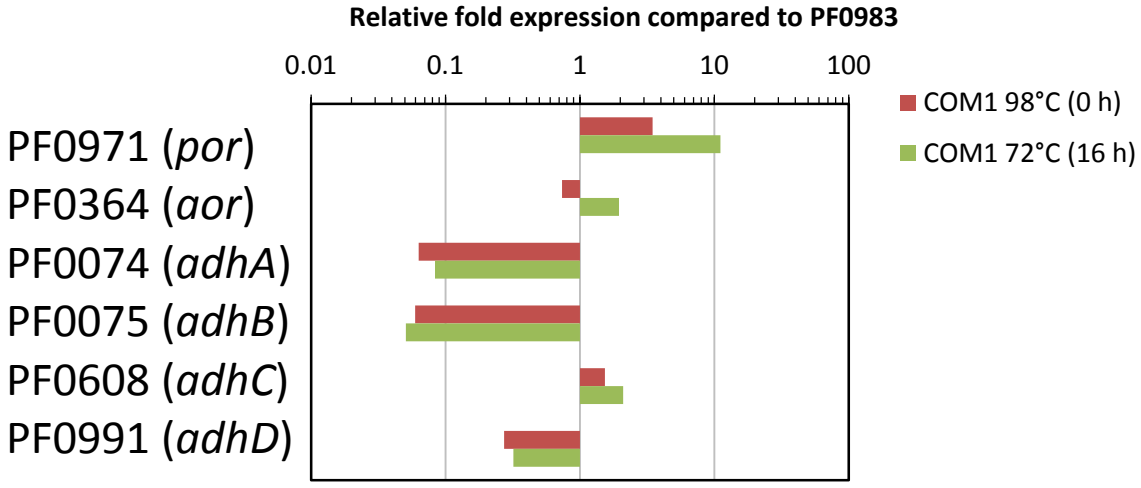

P. furiosus glycolytic pathway
A
Pyruvate
POR
CO2
Fdox
Fdred
CO2
NADH
NAD
AdhE
B
Acetyl-CoA
AdhE
NADH
NAD
Acetaldehyde
AdhA
NADPH
NADP
Ethanol
ACS
ADP
ATP
Fdred
Fdox
AdhC
NADPH
NADP
AOR
C
Acetate
Fdred
Fdox

Figure 6.5 Quantitative PCR of the four *P. furiosus* alcohol dehydrogenases

The expression levels of the four ADHs, POR, and AOR were analyzed via qPCR during growth at 98°C (red) and following a temperature shift to 72°C for 16 hours (Lipscomb, unpublished data).

Figure 6.5


Relative fold expression compared to PF0983
0.01
0.1
1
10
100
PF0971 (*por*)
PF0364 (*aor*)
PF0074 (*adhA*)
PF0075 (*adhB*)
PF0608 (*adhC*)
PF0991 (*adhD*)
COM1 98°C (0 h)
COM1 72°C (16 h)

REFERENCES

1. Basen M*, et al.* (2014) Single gene insertion drives bioalcohol production by a thermophilic archaeon. *Proc Natl Acad Sci* 111(49):17618-17623.
2. Rifkin J (2008) The third industrial revolution. 3(7):26-27.
3. Haub C (1995) How many people have ever lived on earth? *Population today* 23(2):4-5.
4. Kremer M (1993) Population growth and technological change: one million B.C. to 1990. *Q J Econ* 108(3):681-716.
5. Keinan A & Clark AG (2012) Recent explosive human population growth has resulted in an excess of rare genetic variants. *Science* 336(6082):740-743.
6. Easterlin RA (2000) The worldwide standard of living since 1800. *J Econ Perspect*:7-26.
7. www.census.gov (2016).
8. http://energy.gov/fe/office-fossil-energy (2016).
9. Wood HH, Long GR, & Morehouse DF (2004) Long-term world oil supply scenarios. (U. S. Energy Information Administration).
10. Shafiee S & Topal E (2009) When will fossil fuel reserves be diminished? *Energ Policy* 37(1):181-189.
11. Greene DL, Hopson JL, & Li J (2006) Have we run out of oil yet? Oil peaking analysis from an optimist's perspective. *Energy Policy* 34(5):515-531.
12. Madureira NL (2014) Oil reserves and peak oil. *Key Concepts in Energy*, (Springer), pp 101-130.
13. Maugeri L (2004) Oil: never cry wolf—why the petroleum age is far from over. *Science* 304(5674):1114-1115.
14. Grant L (2005) When will the oil run out? *Science* 309(5731):52-54.
15. Ehrenfeld D (2005) Recalculating future oil reserves. *Science* 309(5731):54-56.
16. http://www.climate-skeptic.com/ (2016).
17. Oreskes N (2004) The scientific consensus on climate change. *Science* 306(5702):1686-1686.
18. Eckle P, Burgherr P, & Michaux E (2012) Risk of large oil spills: a statistical analysis in the aftermath of deepwater horizon. *Environ Sci Technol* 46(23):13002-13008.
19. Kleindienst S, Paul JH, & Joye SB (2015) Using dispersants after oil spills: impacts on the composition and activity of microbial communities. *Nat Rev Microbiol* 13(6):388-396.
20. Ahn J*, et al.* (2012) Atmospheric CO2 over the last 1000 years: A high-resolution record from the West Antarctic Ice Sheet (WAIS) Divide ice core. *Global Biogeochem Cy* 26(2):GB2027.
21. Karl TR & Trenberth KE (2003) Modern global climate change. *Science* 302(5651):1719-1723.
22. Pielke RA (2005) Consensus about climate change? *Science* 308(5724):952-954.
23. Solomon S, Plattner G-K, Knutti R, & Friedlingstein P (2009) Irreversible climate change due to carbon dioxide emissions. *Proc Natl Acad Sci* 106(6):1704-1709.
24. Hansen J*, et al.* (2013) Assessing “dangerous climate change”: required reduction of carbon emissions to protect young people, future generations and nature. *PLoS ONE* 8(12):e81648.

25. http://www.theguardian.com/environment/georgemonbiot/2009/dec/16/ian-plimer-versus-george-monbiot (2009).
26. http://hvo.wr.usgs.gov/volcanowatch/archive/2007/07_02_15.html (2007).
27. Gerlach T (2011) Volcanic versus anthropogenic carbon dioxide. *Eos, Transactions American Geophysical Union* 92(24):201-202.
28. Allard P, Burton M, Oskarsson N, Michel A, & Polacci M (2011) Magmatic gas composition and fluxes during the 2010 Eyjafjallajökull explosive eruption: implications for degassing magma volumes and volatile sources. *Geophys. Res. Abstr*.
29. Ulfarsson G & Unger E (2011) Impacts and responses of Icelandic aviation to the 2010 Eyjafjallajökull volcanic eruption: case study. *Transport Res Rec* (2214):144-151.
30. Sultana J (2012) *Eyjafjallajökull - What have we learnt*.
31. Scheelhaase JD & Grimme WG (2007) Emissions trading for international aviation—an estimation of the economic impact on selected European airlines. *J Air Transp Manag* 13(5):253-263.
32. Sahney S & Benton MJ (2008) Recovery from the most profound mass extinction of all time. *P Roy Soc Lond B Bio* 275(1636):759-765.
33. Kamo SL*, et al.* (2003) Rapid eruption of Siberian flood-volcanic rocks and evidence for coincidence with the Permian–Triassic boundary and mass extinction at 251 Ma. *Earth and Planetary Science Letters* 214(1):75-91.
34. Hallam A & Wignall PB (1997) *Mass extinctions and their aftermath* (Oxford University Press).
35. Jablonski D (1986) Causes and consequences of mass extinctions: a comparative approach. *Dynamics of extinction*:183-229.
36. Barnosky AD*, et al.* (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471(7336):51-57.
37. https://www3.epa.gov/climatechange/ghgemissions/gases/co2.html (2016).
38. http://cta.ornl.gov/cta/factsheets.shtml (2016).
39. https://www.teslamotors.com/models (2016).
40. http://www.eia.gov/electricity/annual/ (2016).
41. http://www.fueleconomy.gov/ (2016).
42. Warner ES & Heath GA (2012) Life cycle greenhouse gas emissions of nuclear electricity generation. *J Ind Ecol* 16:S73-S92.
43. Scrosati B, Hassoun J, & Sun Y-K (2011) Lithium-ion batteries. A look into the future. *Energ Environ Sci* 4(9):3287-3295.
44. Alonso E*, et al.* (2012) Evaluating rare earth element availability: A case with revolutionary demand from clean technologies. *Environ Sci Technol* 46(6):3406-3414.
45. www.afdc.energy.gov/ (2016).
46. Rajeshwar K, McConnell RD, & Licht S (2008) *Solar hydrogen generation* (Springer).
47. Woodward J, Orr M, Cordray K, & Greenbaum E (2000) Biotechnology: enzymatic production of biohydrogen. *Nature* 405(6790):1014-1015.
48. Zhang Y-HP, Evans BR, Mielenz JR, Hopkins RC, & Adams MW (2007) High-yield hydrogen production from starch and water by a synthetic enzymatic pathway. *PLoS One* 2(5):e456.

49. Turner JA (2004) Sustainable hydrogen production. *Science* 305(5686):972-974.
50. Hull JF*, et al.* (2012) Reversible hydrogen storage using CO_2 and a proton-switchable iridium catalyst in aqueous media under mild temperatures and pressures. *Nat Chem* 4(5):383-388.
51. Yang Q & Chen GQ (2013) Greenhouse gas emissions of corn–ethanol production in China. *Ecol Model* 252:176-184.
52. www.eia.gov (2016).
53. https://www.epa.gov/laws-regulations/summary-energy-independence-and-security-act (2015).
54. http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32009L0028&from=EN (2009).
55. https://www.afpm.org/uploadedFiles/Content/Policy_Positions/Fuels/Related_Documents/RFS-Summary-2015.pdf (2015).
56. Strogen B, Horvath A, & McKone TE (2012) Fuel miles and the blend wall: Costs and emissions from ethanol distribution in the United States. *Environ Sci Technol* 46(10):5285-5293.
57. Kang S, Önal H, Ouyang Y, Scheffran J, & Tursun ÜD (2010) Optimizing the biofuels infrastructure: transportation networks and biorefinery locations in Illinois. *Handbook of bioenergy economics and policy*, ed Zilberman D (Springer, New York), pp 151-173.
58. Peralta-Yahya PP*, et al.* (2011) Identification and microbial production of a terpene-based advanced biofuel. *Nature Communications* 2:483.
59. Peralta-Yahya PP & Keasling JD (2010) Advanced biofuel production in microbes. *Biotechnol J* 5(2):147-162.
60. Zhang H*, et al.* (2014) Microbial production of sabinene--a new terpene-based precursor of advanced biofuel. *Microb Cell Fact* 13(1):20.
61. Takahashi S*, et al.* (2007) Metabolic engineering of sesquiterpene metabolism in yeast. *Biotechnol Bioeng* 97(1):170-181.
62. Wang J-F, Xiong Z-Q, Li S-Y, & Wang Y (2013) Enhancing isoprenoid production through systematically assembling and modulating efflux pumps in Escherichia coli. *Applied microbiology and biotechnology* 97(18):8057-8067.
63. Ye VM & Bhatia SK (2012) Metabolic engineering for the production of clinically important molecules: Omega-3 fatty acids, artemisinin, and taxol. *Biotechnol J* 7(1):20-33.
64. Schirmer A, Rude MA, Li X, Popova E, & del Cardayre SB (2010) Microbial biosynthesis of alkanes. *Science* 329(5991):559-562.
65. Choi YJ & Lee SY (2013) Microbial production of short-chain alkanes. *Nature* 502(7472):571-574.
66. Maddox IS (1989) The acetone-butanol-ethanol fermentation: recent progress in technology. *Biotechnol Genet Eng* 7(1):189-220.
67. Jones DT & Woods DR (1986) Acetone-butanol fermentation revisited. *Microbiol Rev* 50(4):484.
68. http://www.inchem.org/ (2016).

69. http://arpa-e.energy.gov/ (2016).
70. Conrado RJ, Haynes CA, Haendler BE, & Toone EJ (2013) Electrofuels: a new paradigm for renewable fuels. *Advanced Biofuels and Bioproducts*, ed Lee JW (Springer, New York), pp 1037-1064.
71. Hawkins AB*, et al.* (2011) Extremely thermophilic routes to microbial electrofuels. *ACS Catalysis* 9:1043-1050.
72. Keller M*, et al.* (2014) Production of lignofuels and electrofuels by extremely thermophilic microbes. *Biofuels* 5(5):499-515.
73. Cowan D (1997) Thermophilic proteins: stability and function in aqueous and organic solvents. *Comp Biochem Phys A* 118(3):429-438.
74. Lonsane BK*, et al.* (1992) Scale-up strategies for solid state fermentation systems. *Process Biochem* 27(5):259-273.
75. Wu YY*, et al.* (2013) Energy efficiency opportunities in wineries for retrofit and new construction projects. in *Industrial Energy Technology Conference 2013* (Energy Systems Laboratory, New Orleans, LA).
76. Houdas Y & Ring E (1982) *Human body temperature* (Springer Science & Business Media, New York).
77. Takai K*, et al.* (2008) Cell proliferation at 122°C and isotopically heavy CH_4 production by a hyperthermophilic methanogen under high-pressure cultivation. *Proc Natl Acad Sci* 105(31):10949-10954.
78. Morita RY (1975) Psychrophilic bacteria. *Bacteriol Rev* 39(2):144.
79. Kristjansson JK & Stetter KO (1991) *Thermophilic bacteria* (CRC press).
80. Clarke A*, et al.* (2013) A low temperature limit for life on Earth. *PLoS ONE* 8(6):e66207.
81. Reid IN*, et al.* (2006) Terrestrial models for extraterrestrial life: methanogens and halophiles at Martian temperatures. *Int J Astrobiol* 5(02):89-97.
82. Miquel P (1888) Monographie d'un Bacille vivant au-dela de 70°C. *Annales de Microgr* 1:4-10.
83. Macfadyen A & Blaxall FR (1896) Thermophilic bacteria. *J Pathol Bacteriol* 3(1):87-99.
84. Donk PJ (1920) A highly resistant thermophilic organism. *J Bacteriol* 5(4):373-374.
85. Brock TD & Freeze H (1969) *Thermus aquaticus* gen. n. and sp. n., a nonsporulating extreme thermophile. *J Bacteriol* 98(1):289-297.
86. www.nobelprize.org (2016).
87. Vieille C & Zeikus GJ (2001) Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol R* 65(1):1-43.
88. Bruins M, Janssen AM, & Boom R (2001) Thermozymes and their applications. *Appl Biochem Biotech* 90(2):155-186.
89. Mojsov K, Janevski A, Andronikov D, & Zezova S (2014) Hyperthermophilic enzymes with industrial applications. *IJIRSE* 2(12):806-810.
90. Pikuta EV, Hoover RB, & Tang J (2007) Microbial extremophiles at the limits of life. *Crit Rev Microbiol* 33(3):183-209.
91. van de Vossenberg JL, Driessen AJ, & Konings WN (1998) The essence of being extremophilic: the role of the unique archaeal membrane lipids. *Extremophiles* 2(3):163-170.

92. Leigh JA, Albers S-V, Atomi H, & Allers T (2011) Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol Rev* 35(4):577-608.
93. Winker S & Woese C (1991) A definition of the domains Archaea, Bacteria and Eucarya in terms of small subunit ribosomal RNA characteristics. *Syst Appl Microbiol* 14(4):305-310.
94. Blöchl E*, et al.* (1995) Isolation, taxonomy and phylogeny of hyperthermophilic microorganisms. *World J Microb Biot* 11(1):9-16.
95. Garrett RA & Klenk H-P (2008) *Archaea: evolution, physiology, and molecular biology* (John Wiley & Sons).
96. Woese CR & Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci* 74(11):5088-5090.
97. Rivera MC, Jain R, Moore JE, & Lake JA (1998) Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci* 95(11):6239-6244.
98. Yutin N, Makarova KS, Mekhedov SL, Wolf YI, & Koonin EV (2008) The deep archaeal roots of eukaryotes. *Mol Biol Evol* 25(8):1619-1630.
99. Olsen GJ & Woese CR (1997) Archaeal genomics: an overview. *Cell* 89(7):991-994.
100. Olsen GJ & Woese CR (1996) Lessons from an Archaeal genome: what are we learning from *Methanococcus jannaschii*? *Trends Genet* 12(10):377-379.
101. Koonin EV, Mushegian AR, Galperin MY, & Walker DR (1997) Comparison of archaeal and bacterial genomes: computer analysis of protein sequences predicts novel functions and suggests a chimeric origin for the archaea. *Mol Microbiol* 25(4):619-637.
102. Fiala G & Stetter K (1986) *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. *Arch. Microbiol.* 145(1):56-61.
103. Kengen SW*, et al.* (1994) Evidence for the operation of a novel Embden-Meyerhof pathway that involves ADP-dependent kinases during sugar fermentation by *Pyrococcus furiosus*. *J Biol Chem* 269(26):17537-17541.
104. Kengen SM, Stams AJ, & Vos Wd (1996) Sugar metabolism of hyperthermophiles. *FEMS microbiol Rev* 18(2):119-138.
105. Nguyen DM*, et al.* (2015) Temperature-dependent acetoin production by *Pyrococcus furiosus* is catalyzed by a biosynthetic acetolactate synthase and its deletion improves ethanol production. *Metab Eng* 34:71-79.
106. Stetter KO (1999) Extremophiles and their adaptation to hot environments. *FEBS letters* 452(1):22-25.
107. Kreuzer M, Schmutzler K, Waege I, Thomm M, & Hausner W (2013) Genetic engineering of *Pyrococcus furiosus* to use chitin as a carbon source. *BMC Biotechnol* 13(1):9.
108. Koning SM, Elferink MG, Konings WN, & Driessen AJ (2001) Cellobiose uptake in the hyperthermophilic archaeon *Pyrococcus furiosus* is mediated by an inducible, high-affinity ABC transporter. *J Bacteriol* 183(17):4979-4984.
109. Voorhorst W, Rik I, Luesink E, & Devos W (1995) Charaterization of the CelB gene coding for beta-glucosidase from the hyperthermohilic archaeon *Pyrococcus furiosus* and its

expression and site-directed mutagenesis in *Escherichia coli*. *J Bacteriol* 177(24):7105-7111.
110. Xavier KB, Peist R, Kossmann M, Boos W, & Santos H (1999) Maltose metabolism in the hyperthermophilic archaeon *Thermococcus litoralis*: Purification and characterization of key enzymes. *J Bacteriol* 181(11):3358-3367.
111. DiRuggiero J*, et al.* (2000) Evidence of recent lateral gene transfer among hyperthermophilic archaea. *Mol Microbiol* 38(4):684-693.
112. Schut GJ, Brehm SD, Datta S, & Adams MWW (2003) Whole-genome DNA microarray analysis of a hyperthermophile and an archaeon: *Pyrococcus furiosus* grown on carbohydrates or peptides. *J Bacteriol* 185(13):3935-3947.
113. Bauer MW, Bylina EJ, Swanson RV, & Kelly RM (1996) Comparison of a β-Glucosidase and a β-Mannosidase from the Hyperthermophilic Archaeon *Pyrococcus furiosus*. Purification, characterization, gene cloning, and sequence analysis. *J Biol Chem* 271(39):23749-23755.
114. Verhees C*, et al.* (2003) The unique features of glycolytic pathways in Archaea. *Biochem. J* 375:231-246.
115. Ronimus RS & Morgan HW (2003) Distribution and phylogenies of enzymes of the Embden-Meyerhof-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. *Archaea* 1(3):199-221.
116. Sakuraba H & Ohshima T (2002) Novel energy metabolism in anaerobic hyperthermophilic archaea: a modified Embden-Meyerhof pathway. *J Biosci Bioeng* 93(5):441-448.
117. Mukund S & Adams MWW (1991) The novel tungsten-iron-sulfur protein of the hyperthermophilic archaebacterium, *Pyrococcus furiosus*, is an aldehyde ferredoxin oxidoreductase - evidence for its participation in a unique glycolytic pathway. *J Biol Chem* 266(22):14208-14216.
118. Blamey JM & Adams MWW (1993) Purification and characterization of pyruvate ferredoxin oxidoreductase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *BBA-Protein Struct M* 1161(1):19-27.
119. Mai X & Adams M (1996) Purification and characterization of two reversible and ADP-dependent acetyl coenzyme A synthetases from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* 178(20):5897-5903.
120. Scott JW, Poole FL, & Adams MWW (2014) Characterization of ten heterotetrameric NDP-dependent acyl-CoA synthetases of the hyperthermophilic archaeon *Pyrococcus furiosus*. *Archaea* 2014.
121. Pisa KY, Huber H, Thomm M, & Müller V (2007) A sodium ion-dependent A1AO ATP synthase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Febs J* 274(15):3928-3938.
122. Sapra R, Bagramyan K, & Adams MWW (2003) A simple energy-conserving system: Proton reduction coupled to proton translocation. *Proc Natl Acad Sci* 100(13):7545-7550.

123. van Haaster DJ, Silva PJ, Hagedoorn P-L, Jongejan JA, & Hagen WR (2008) Reinvestigation of the steady-state kinetics and physiological function of the soluble NiFe-hydrogenase I of *Pyrococcus furiosus*. *J Bacteriol* 190(5):1584-1587.
124. Mertens R, Greiner L, van den Ban ECD, Haaker HBCM, & Liese A (2003) Practical applications of hydrogenase I from *Pyrococcus furiosus* for NADPH generation and regeneration. *J Mol Catal B-Enzym* 24–25:39-52.
125. Silva PJ*, et al.* (2000) Enzymes of hydrogen metabolism in *Pyrococcus furiosus*. *Eur J Biochem* 267(22):6541-6551.
126. van der Oost J*, et al.* (2001) Genetic and biochemical characterization of a short-chain alcohol dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur J Biochem* 268(10):3062-3068.
127. Machielsen R & van der Oost J (2006) Production and characterization of a thermostable L-threonine dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *FEBS Journal* 273(12):2722-2729.
128. Hensley SA, Jung J-H, Park C-S, & Holden JF (2014) *Thermococcus paralvinellae* sp. nov. and *Thermococcus cleftensis* sp. nov. of hyperthermophilic heterotrophs from deep-sea hydrothermal vents. *Int J Syst Evol Micr* 64(11):3655-3659.
129. Ying X, Grunden AM, Nie L, Adams MW, & Ma K (2009) Molecular characterization of the recombinant iron-containing alcohol dehydrogenase from the hyperthermophilic Archaeon, *Thermococcus* strain ES1. *Extremophiles* 13(2):299-311.
130. Lam WL & Doolittle WF (1989) Shuttle vectors for the archaebacterium *Halobacterium volcanii*. *Proc Natl Acad Sci* 86(14):5478-5482.
131. Matsumi R, Manabe K, Fukui T, Atomi H, & Imanaka T (2007) Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance. *J Bacteriol* 189(7):2683-2691.
132. Simões RG*, et al.* (2014) Thermal stability of simvastatin under different atmospheres. *J Pharm Sci* 103(1):241-248.
133. Waege I, Schmid G, Thumann S, Thomm M, & Hausner W (2010) Shuttle vector-based transformation system for *Pyrococcus furiosus*. *Appl Environ Microb* 76(10):3308-3313.
134. Krooth R, Hsiao W-L, & Potvin B (1979) Resistance to 5-fluoroorotic acid and pyrimidine auxotrophy: A new bidirectional selective system for mammalian cells. *Somat Cell Genet* 5(5):551-569.
135. Jund R & Lacroute F (1970) Genetic and physiological aspects of resistance to 5-fluoropyrimidines in *Saccharomyces cerevisiae*. *J Bacteriol* 102(3):607-615.
136. Worsham P & Goldman W (1988) Selection and characterization of ura5 mutants of *Histoplasma capsulatum*. *Mol Gen Genet* 214(2):348-352.
137. Lipscomb GL*, et al.* (2011) Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. *Appl Environ Microb* 77(7):2232-2238.
138. Weinberg MV, Schut GJ, Brehm S, Datta S, & Adams MW (2005) Cold shock of a hyperthermophilic archaeon: *Pyrococcus furiosus* exhibits multiple responses to a

suboptimal growth temperature with a key role for membrane-bound glycoproteins. *J Bacteriol* 187:336-348.
139. Ma K, Hutchins A, Sung S-JS, & Adams MWW (1997) Pyruvate ferredoxin oxidoreductase from the hyperthermophilic archaeon, *Pyrococcus furiosus*, functions as a CoA-dependent pyruvate decarboxylase. *Proc Natl Acad Sci* 94(18):9608-9613.
140. Hopkins RC (2012) Heterologous and homologous expression and characterization of a [Nife]-hydrogenase from the hyperthermophilic archaeon, *Pyrococcus furiosus*. PhD (University of Georgia, Athens, GA).
141. Berg IA, Kockelkorn D, Buckel W, & Fuchs G (2007) A 3-hydroxypropionate/4-hydroxybutyrate autotrophic carbon dioxide assimilation pathway in Archaea. *Science* 318(5857):1782-1786.
142. Huber G, Spinnler C, Gambacorta A, & Stetter KO (1989) *Metallosphaera sedula* gen, and sp. nov. represents a new genus of aerobic, metal-mobilizing, thermoacidophilic archaebacteria. *Syst Appl Microbiol* 12(1):38-47.
143. Auernik KS & Kelly RM (2010) Physiological versatility of the extremely thermoacidophilic archaeon *Metallosphaera sedula* supported by transcriptomic analysis of heterotrophic, autotrophic, and mixotrophic growth. *Appl Environ Microb* 76(3):931-935.
144. Menendez C*, et al.* (1999) Presence of acetyl coenzyme A (CoA) carboxylase and propionyl-CoA carboxylase in autotrophic *Crenarchaeota* and indication for operation of a 3-hydroxypropionate cycle in autotrophic carbon fixation. *J Bacteriol* 181(4):1088-1098.
145. Hügler M, Krieger RS, Jahn M, & Fuchs G (2003) Characterization of acetyl-CoA/propionyl-CoA carboxylase in *Metallosphaera sedula*. *Eur J Biochem* 270(4):736-744.
146. Hügler M, Menendez C, Schägger H, & Fuchs G (2002) Malonyl-coenzyme A reductase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO2 fixation. *J Bacteriol* 184(9):2404-2410.
147. Alber B*, et al.* (2006) Malonyl-coenzyme A reductase in the modified 3-hydroxypropionate cycle for autotrophic carbon fixation in archaeal *Metallosphaera* and *Sulfolobus* spp. *J Bacteriol* 188(24):8551-8559.
148. Han Y, Hawkins AS, Adams MWW, & Kelly RM (2012) Epimerase (Msed_0639) and mutase (Msed_0638 and Msed_2055) convert (S)-methylmalonyl-coenzyme A (CoA) to succinyl-CoA in the *Metallosphaera sedula* 3-hydroxypropionate/4-hydroxybutyrate cycle. *Appl Environ Microb* 78(17):6194-6202.
149. Sato T & Atomi H (2011) Novel metabolic pathways in Archaea. *Curr Opin Microbiol* 14(3):307-314.
150. Estelmann S*, et al.* (2011) Labeling and Enzyme Studies of the Central Carbon Metabolism in *Metallosphaera sedula*. *J Bacteriol* 193(5):1191-1200.
151. Ramos-Vera WH, Weiss M, Strittmatter E, Kockelkorn D, & Fuchs G (2011) Identification of missing genes and enzymes for autotrophic carbon fixation in *Crenarchaeota*. *J Bacteriol* 193(5):1201-1211.

152. Berg IA, Ramos-Vera WH, Petri A, Huber H, & Fuchs G (2010) Study of the distribution of autotrophic CO2 fixation cycles in *Crenarchaeota*. *Microbiology* 156(1):256-269.
153. Kockelkorn D & Fuchs G (2009) Malonic semialdehyde reductase, succinic semialdehyde reductase, and succinyl-coenzyme A reductase from *Metallosphaera sedula*: enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in *Sulfolobales*. *J Bacteriol* 191(20):6352-6362.
154. Alber BE, Kung JW, & Fuchs G (2008) 3-hydroxypropionyl-coenzyme A synthetase from *Metallosphaera sedula*, an enzyme involved in autotrophic CO2 fixation. *J Bacteriol* 190(4):1383-1389.
155. Peng H, Wu G, & Shao W (2008) The aldehyde/alcohol dehydrogenase (AdhE) in relation to the ethanol formation in *Thermoanaerobacter ethanolicus* JW200. *Anaerobe* 14(2):125-127.
156. Wiegel J & Ljungdahl LG (1981) *Thermoanaerobacter ethanolicus* gen. nov., spec. nov., a new extreme thermophilic anaerobic bacterium *Arch. Microbiol.* 128(4):343-348.
157. Lacis LS & Lawford HG (1988) Ethanol production from xylose by *Thermoanaerobacter ethanolicus* in batch and continuous culture. *Arch. Microbiol.* 150(1):48-55.
158. Hild H, Stuckey D, & Leak D (2003) Effect of nutrient limitation on product formation during continuous fermentation of xylose with *Thermoanaerobacter ethanolicus* JW200 Fe (7). *Appl Microbiol Biotechnol* 60(6):679-686.
159. Arnau J, Jørgensen F, Madsen SM, Vrang A, & Israelsen H (1998) Cloning of the *Lactococcus lactis adhE* gene, encoding a multifunctional alcohol dehydrogenase, by complementation of a fermentative mutant of *Escherichia coli*. *J Bacteriol* 180(12):3049-3055.
160. Fontaine L*, et al.* (2002) Molecular characterization and transcriptional analysis of adhE2, the gene encoding the NADH-dependent aldehyde/alcohol dehydrogenase responsible for butanol production in alcohologenic cultures of *Clostridium acetobutylicum* ATCC 824. *J Bacteriol* 184(3):821-830.
161. 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(1):59-63.
162. Lo J, Zheng T, Hon S, Olson DG, & Lynd LR (2015) The bifunctional alcohol and aldehyde dehydrogenase gene, *adhE*, is necessary for ethanol production in *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*. *J Bacteriol* 197(8):1386-1393.
163. Yao S & Mikkelsen MJ (2010) Identification and overexpression of a bifunctional aldehyde/alcohol dehydrogenase responsible for ethanol production in *Thermoanaerobacter mathranii*. *J Mol Microb Biotech* 19(3):123-133.
164. Yao S & Mikkelsen M (2010) Metabolic engineering to improve ethanol production in Thermoanaerobacter mathranii. *Appl Microbiol Biotechnol* 88(1):199-208.
165. Stetter KO (1996) Hyperthermophilic procaryotes. *FEMS Microbiol Rev* 18(2):149-158.
166. Somerville C, Youngs H, Taylor C, Davis SC, & Long SP (2010) Feedstocks for lignocellulosic biofuels. *Science* 329(5993):790-792.

167. Olson DG, McBride JE, Shaw AJ, & Lynd LR (2012) Recent progress in consolidated bioprocessing. *Curr Opin Microbiol* 23(3):396-405.
168. Steen EJ*, et al.* (2010) Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* 463(7280):559-562.
169. Wackett LP (2011) Engineering microbes to produce biofuels. *Curr Opin Biotech* 22(3):388-393.
170. Shen CR*, et al.* (2011) Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl Environ Microb* 77(9):2905-2915.
171. Connor MR & Liao JC (2009) Microbial production of advanced transportation fuels in non-natural hosts. *Curr Opin Biotech* 20(3):307-315.
172. Basen M, Sun J, & Adams MW (2012) Engineering a hyperthermophilic archaeon for temperature-dependent product formation. *mBio* 3(2):e00053-00012.
173. Paster M, Pellegrino JL, & Carole TM (2004) Industrial bioproducts: today and tomorrow.US Department of Energy and Energetics Inc., Columbia, MD.
174. Werpy T & Petersen G (2004) Top value added chemicals from biomass: volume 1—Results of screening for potential candidates from sugars and synthesis gas. *Dept. of Energy,* 102004-1992, http://www.osti.gov/bridge.
175. Bridger SL, Lancaster WA, Poole FL, 2nd, Schut GJ, & Adams MW (2012) Genome sequencing of a genetically tractable *Pyrococcus furiosus* strain reveals a highly dynamic genome. *J Bacteriol* 194(15):4097-4106.
176. Chandrayan SK*, et al.* (2012) Engineering hyperthermophilic archaeon *Pyrococcus furiosus* to overproduce its cytoplasmic [NiFe]-hydrogenase. *J Biol Chem* 287(5):3257-3264.
177. Hopkins RC*, et al.* (2011) Homologous expression of a subcomplex of *Pyrococcus furiosus* hydrogenase that interacts with pyruvate ferredoxin oxidoreductase. *PLoS One* 6(10):e26569.
178. Berg IA*, et al.* (2010) Autotrophic carbon fixation in archaea. *Nat Rev Micro* 8(6):447-460.
179. Ma K & Adams MWW (2001) Hydrogenases I and II from *Pyrococcus furiosus*. *Method Enzymol* Volume 331:208-216.
180. Yoon SH*, et al.* (2011) Parallel evolution of transcriptome architecture during genome reorganization. *Genome Res* 21(11):1892-1904.
181. Kreysa G (2009) Climate protection by an alternative use of methane--the carbon moratorium. *Chem Sus Chem* 2(1):49-55.
182. Horton RM, Hunt HD, Ho SN, Pullen JK, & Pease LR (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77(1):61-68.
183. Farkas J*, et al.* (2012) Recombinogenic properties of *Pyrococcus furiosus* Strain COM1 enable rapid selection of targeted mutants. *Appl Environ Microb* 78(13):4669-4676.
184. Miwa H (2000) High-performance liquid chromatographic determination of mono-, poly- and hydroxycarboxylic acids in foods and beverages as their 2-nitrophenylhydrazides. *Journal of chromatography. A* 881(1-2):365-385.

185. Atsumi S & Liao JC (2008) Metabolic engineering for advanced biofuels production from *Escherichia coli*. *Curr Opin Biotech* 19(5):414-419.
186. Zhang F, Rodriguez S, & Keasling JD (2011) Metabolic engineering of microbial pathways for advanced biofuels production. *Curr Opin Biotech* 22(6):775-783.
187. Peralta-Yahya PP, Zhang F, del Cardayre SB, & Keasling JD (2012) Microbial engineering for the production of advanced biofuels. *Nature* 488(7411):320-328.
188. Lan EI & Liao JC (2013) Microbial synthesis of n-butanol, isobutanol, and other higher alcohols from diverse resources. *Bioresource Technol* 135(0):339-349.
189. Köpke M*, et al.* (2010) *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc Natl Acad Sci* 107(29):13087-13092.
190. Lan EI & Liao JC (2012) ATP drives direct photosynthetic production of 1-butanol in cyanobacteria. *Proc Natl Acad Sci* 109(16):6018-6023.
191. Li H*, et al.* (2012) Integrated electromicrobial conversion of CO_2 to higher alcohols. *Science* 335(6076):1596.
192. Keller MW*, et al.* (2013) Exploiting microbial hyperthermophilicity to produce an industrial chemical, using hydrogen and carbon dioxide. *Proc Natl Acad Sci* 110(15):5840-5845.
193. Cripps RE*, et al.* (2009) Metabolic engineering of *Geobacillus thermoglucosidasius* for high yield ethanol production. *Metab Eng* 11(6):398-408.
194. Abdel-Banat BA, Hoshida H, Ano A, Nonklang S, & Akada R (2010) High-temperature fermentation: how can processes for ethanol production at high temperatures become superior to the traditional process using mesophilic yeast? *Appl Microbiol Biotechnol* 85(4):861-867.
195. Bhandiwad A, Guseva A, & Lynd L (2013) Metabolic engineering of *Thermoanaerobacterium thermosaccharolyticum* for increased n-butanol production. *Adv Microbial* 3:46-51.
196. Bhandiwad A*, et al.* (2014) Metabolic engineering of *Thermoanaerobacterium saccharolyticum* for n-butanol production. *Metab Eng* 21(0):17-25.
197. Lin PP*, et al.* (2014) Isobutanol production at elevated temperatures in thermophilic *Geobacillus thermoglucosidasius*. *Metab Eng* 24(0):1-8.
198. Sato T, Fukui T, Atomi H, & Imanaka T (2003) Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J Bacteriol* 185(1):210-220.
199. Kim YJ*, et al.* (2010) Formate-driven growth coupled with H_2 production. *Nature* 467(7313):352-355.
200. Merryman C & Gibson DG (2012) Methods and applications for assembling large DNA constructs. *Metab Eng* 14(3):196-204.
201. Thorgersen MP, Lipscomb GL, Schut GJ, Kelly RM, & Adams MWW (2014) Deletion of acetyl-CoA synthetases I and II increases production of 3-hydroxypropionate by the metabolically-engineered hyperthermophile *Pyrococcus furiosus*. *Metab Eng* 22(0):83-88.

202. Adams MWW*, et al.* (2001) Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* 183(2):716-724.
203. Xue Y, Xu Y, Liu Y, Ma Y, & Zhou P (2001) *Thermoanaerobacter tengcongensis* sp. nov., a novel anaerobic, saccharolytic, thermophilic bacterium isolated from a hot spring in Tengcong, China. *Int J syst Evol Micr* 51(4):1335-1341.
204. Li F*, et al.* (2008) Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*. *J Bacteriol* 190(3):843-850.
205. Aksenova HY, Rainey FA, Janssen PH, Zavarzin GA, & Morgan HW (1992) *Spirochaeta thermophila* sp. nov., an obligately anaerobic, polysaccharolytic, extremely thermophilic bacterium. *Int J syst Evol Bacteriol* 42(1):175-177.
206. Roh Y*, et al.* (2002) Isolation and characterization of metal-reducing *Thermoanaerobacter* strains from deep subsurface environments of the Piceance Basin, Colorado. *Appl Environ Microb* 68(12):6013-6020.
207. Extance J*, et al.* (2013) Structure of a bifunctional alcohol dehydrogenase involved in bioethanol generation in *Geobacillus thermoglucosidasius*. *Acta Cryst D* 69(10):2104-2115.
208. Atomi H, Fukui T, Kanai T, Morikawa M, & Imanaka T (2004) Description of *Thermococcus kodakaraensis* sp. nov., a well studied hyperthermophilic archaeon previously reported as *Pyrococcus* sp. KOD1. *Archaea* 1(4):263-267.
209. Basen M*, et al.* (2014) Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium *Caldicellulosiruptor bescii*. *Bioresource Technol* 152(0):384-392.
210. Hoppner T & Doelle H (1983) Purification and kinetic characteristics of pyruvate decarboxylase and ethanol dehydrogenase from *Zymomonas mobilis* in relation to ethanol production. *Eur J Appl Microbiol* 17(3):152-157.
211. Cleenwerck I, Vandemeulebroecke K, Janssens D, & Swings J (2002) Re-examination of the genus Acetobacter, with descriptions of *Acetobacter cerevisiae* sp. nov. and *Acetobacter malorum* sp. nov. *Int J Syst Evol Bacteriol* 52(5):1551-1558.
212. Andrés-Barrao C*, et al.* (2011) Genome sequences of the high-acetic acid-resistant bacteria *Gluconacetobacter europaeus* LMG 18890T and *G. europaeus* LMG 18494 (Reference Strains), *G. europaeus* 5P3, and *Gluconacetobacter oboediens* 174Bp2 (Isolated from Vinegar). *J Bacteriol* 193(10):2670-2671.
213. Kaneko T*, et al.* (2007) Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. *DNA Res* 14(6):247-256.
214. Tamas I*, et al.* (2010) Complete genome sequence of *Beijerinckia indica* subsp. indica. *J Bacteriol* 192(17):4532-4533.
215. Goodlove PE, Cunningham PR, Parker J, & Clark DP (1989) Cloning and sequence analysis of the fermentative alcohol-dehydrogenase-encoding gene of *Escherichia coli*. *Gene* 85(1):209-214.
216. Nikolau BJ, Oliver DJ, Schnable PS, & Wurtele ES (2000) Molecular biology of acetyl-CoA metabolism. *Biochem Soc T* 28(6):591-593.

217. Lee Y-E, Jain MK, Lee C, & Zeikus JG (1993) Taxonomic distinction of saccharolytic thermophilic anaerobes: description of *Thermoanaerobacterium xylanolyticum* gen. nov., sp. nov., and *Thermoanaerobacterium saccharolyticum* gen. nov., sp. nov.; reclassification of *Thermoanaerobium brockii*, *Clostridium*. *Int J Syst Bacteriol* 43(1):41-51.
218. Shrager J*, et al.* (2003) *Chlamydomonas reinhardtii* genome project. A guide to the generation and use of the cDNA information. *Plant Physiol* 131(2):401-408.
219. Taylor MP*, et al.* (2009) Thermophilic ethanologenesis: future prospects for second-generation bioethanol production. *Trends Biotechnol* 27(7):398-405.
220. Olson DG, Sparling R, & Lynd LR (2015) Ethanol production by engineered thermophiles. *Curr Opin Biotech* 33:130-141.
221. Thompson AH, Studholme DJ, Green EM, & Leak DJ (2008) Heterologous expression of pyruvate decarboxylase in *Geobacillus thermoglucosidasius*. *Biotechnology Letters* 30(8):1359-1365.
222. Zheng T*, et al.* (2015) Cofactor specificity of the bifunctional alcohol and aldehyde dehydrogenase (AdhE) in wild-type and mutants of *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*. *J Bacteriol* 197(15).
223. Zhou J, Wu K, & Rao C (2016) Evolutionary engineering of *Geobacillus thermoglucosidasius* for improved ethanol production. *Biotechnol Bioeng*.
224. Chung D, Cha M, Guss AM, & Westpheling J (2014) Direct conversion of plant biomass to ethanol by engineered *Caldicellulosiruptor bescii*. *Proc Natl Acad Sci* 111(24):8931-8936.
225. Onyenwoke RU, Kevbrin VV, Lysenko AM, & Wiegel J (2007) *Thermoanaerobacter pseudethanolicus* sp. nov., a thermophilic heterotrophic anaerobe from Yellowstone National Park. *Int J Syst Evol Micr* 57(10):2191-2193.
226. Chung D*, et al.* (2015) Cellulosic ethanol production via consolidated bioprocessing at 75 C by engineered *Caldicellulosiruptor bescii*. *Biotechnol Biofuel* 8(1).
227. Keller MW*, et al.* (2015) A hybrid synthetic pathway for butanol production by a hyperthermophilic microbe. *Metab Eng* 27:101-106.
228. Schut GJ, Lipscomb GL, Nguyen DM, Kelly RM, & Adams MW (2016) Heterologous production of an energy-conserving carbon monoxide dehydrogenase complex in the hyperthermophile *Pyrococcus furiosus*. *Front Microbiol* 7.
229. Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403-410.
230. Loder AJ*, et al.* (2015) Alcohol selectivity in a synthetic thermophilic n-butanol pathway is driven by biocatalytic and thermostability characteristics of constituent enzymes. *Appl Environ Microb* 81(20):7187-7200.
231. Roy R*, et al.* (1999) Purification and molecular characterization of the tungsten-containing formaldehyde ferredoxin oxidoreductase from the hyperthermophilic archaeon *Pyrococcus furiosus*: the third of a putative five-member tungstoenzyme family. *J Bacteriol* 181(4):1171-1180.
232. Roy R & Adams MW (2002) Characterization of a fourth tungsten-containing enzyme from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Bacteriol* 184(24):6952-6956.

233. Bevers LE, Bol E, Hagedoorn P-L, & Hagen WR (2005) WOR5, a novel tungsten-containing aldehyde oxidoreductase from *Pyrococcus furiosus* with a broad substrate specificity. *J Bacteriol* 187(20):7056-7061.
234. Thorgersen MP, Stirrett K, Scott RA, & Adams MWW (2012) Mechanism of oxygen detoxification by the surprisingly oxygen-tolerant hyperthermophilic archaeon, *Pyrococcus furiosus*. *Proc Natl Acad Sci* 109(45):18547-18552.
235. Huang H, Wang S, Moll J, & Thauer RK (2012) Electron bifurcation involved in the energy metabolism of the acetogenic bacterium *Moorella thermoacetica* growing on glucose or H_2 plus CO_2. *J Bacteriol* 194(14):3689-3699.
236. Chan M, Mukund S, Kletzin A, Adams M, & Rees D (1995) Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* 267(5203):1463-1469.
237. Scott IM*, et al.* (2015) A new class of tungsten-containing oxidoreductase in the genus of the plant biomass-degrading, thermophilic bacteria *Caldicellulosiruptor*. *Appl Environ Microb*.
238. Pei JJ*, et al.* (2010) *Thermoanaerobacter* spp. control ethanol pathway via transcriptional regulation and versatility of key enzymes. *Metab Eng* 12(5):420-428.
239. Hemme CL*, et al.* (2011) Correlation of genomic and physiological traits of *Thermoanaerobacter* species with biofuel yields. *Appl Environ Microb* 77(22):7998-8008.
240. Santos H & Da Costa MS (2002) Compatible solutes of organisms that live in hot saline environments. *Environ Microbiol* 4(9):501-509.
241. Zheng Y, Kahnt J, Kwon IH, Mackie RI, & Thauer RK (2014) Hydrogen Formation and Its Regulation in *Ruminococcus albus*: Involvement of an Electron-Bifurcating [FeFe]-Hydrogenase, of a Non-Electron-Bifurcating [FeFe]-Hydrogenase, and of a Putative Hydrogen-Sensing [FeFe]-Hydrogenase. *J Bacteriol* 196(22):3840-3852.
242. Bruchhaus I & Tannich E (1994) Purification and molecular characterization of the NAD+-dependent acetaldehyde/alcohol dehydrogenase from *Entamoeba histolytica*. *Biochem. J.* 303(3):743-748.
243. Mukund S & Adams MWW (1995) Glyceraldehyde-3-phosphate ferredoxin oxidoreductase, a novel tungsten-containing enzyme with a potential glycolytic role in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J Biol Chem* 270(15):8389-8392.
244. Yutin N & Galperin MY (2013) A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. *Environ Microbiol* 15(10):2631-2641.
245. Tolonen AC*, et al.* (2015) Physiology, Genomics, and Pathway Engineering of an Ethanol-Tolerant Strain of *Clostridium phytofermentans*. *Appl Environ Microb* 81(16):5440-5448.
246. Eram MS, Oduaran E, & Ma K (2014) The bifunctional pyruvate decarboxylase/pyruvate ferredoxin oxidoreductase from *Thermococcus guaymasensis*. *Archaea* 2014.
247. Ma K & Adams M (2001) Ferredoxin: NADP oxidoreductase from *Pyrococcus furiosus*. *Method Enzymol* 334:40-45.
248. Gibson DG*, et al.* (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343-345.

249. Yang SJ*, et al.* (2010) Classification of '*Anaerocellum thermophilum*' strain DSM 6725 as *Caldicellulosiruptor bescii* sp. nov. *Int J syst Evol Bacteriol* 60:2011-2015.
250. Bryant FO, Wiegel J, & Ljungdahl LG (1992) Comparison of alcohol dehydrogenases from wild-type and mutant Strain, JW200 Fe 4, of *Thermoanaerobacter ethanolicus*. *Appl Microbiol Biotechnol* 37(4):490-495.
251. Lamed R & Zeikus JG (1980) Ethanol production by thermophilic bacteria: relationship between fermentation product yields of and catabolic enzyme activities in *Clostridium thermocellum* and *Thermoanaerobium brockii*. *J Bacteriol* 144(2):569-578.
252. Dai J*, et al.* (2011) A new subspecies of *Anoxybacillus flavithermus* ssp. *yunnanensis* ssp. nov. with very high ethanol tolerance. *FEMS Microbiol Lett* 320(1):72-78.
253. Engle M*, et al.* (1996) *Thermobrachium celere* gen. nov., sp. nov., a rapidly growing thermophilic, alkalitolerant, and proteolytic obligate anaerobe. *Int J Syst Evol Bacteriol* 46(4):1025-1033.
254. Visser M*, et al.* (2013) Genome analysis of *Desulfotomaculum kuznetsovii* strain 17(T) reveals a physiological similarity with *Pelotomaculum thermopropionicum* strain SI(T). *Stand Genomic Sci* 8(1):69-87.
255. Sievers F*, et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539-539.
256. Clamp M, Cuff J, Searle SM, & Barton GJ (2004) The jalview java alignment editor. *Bioinformatics* 20(3):426-427.
257. Hawkins AS, Han Y, Bennett RK, Adams MWW, & Kelly RM (2013) Role of 4-Hydroxybutyrate-CoA Synthetase in the CO2 Fixation Cycle in Thermoacidophilic Archaea. *J Biol Chem* 288(6):4012-4022.
258. Hawkins AB, Adams MWW, & Kelly RM (2014) Conversion of 4-Hydroxybutyrate to Acetyl-CoA and its Anapleurosis in the *Metallosphaera sedula* 3-Hydroxypropionate/4-Hydroxybutyrate Carbon Fixation Pathway. *Appl Environ Microb*.
259. Riddles PW, Blakeley RL, & Zerner B (1979) Ellman's reagent: 5, 5'-dithiobis (2-nitrobenzoic acid)—a reexamination. *Anal Biochem* 94(1):75-81.
260. Hawkins AB*, et al.* (2015) Bioprocessing analysis of *Pyrococcus furiosus* strains engineered for CO_2-based 3-hydroxypropionate production. *Biotechnol Bioeng*:n/a-n/a.
261. Torella JP*, et al.* (2015) Efficient solar-to-fuels production from a hybrid microbial–water-splitting catalyst system. *Proc Natl Acad Sci* 112(8):2337-2342.
262. Heider J, Ma K, & Adams MWW (1995) Purification, characterization, and metabolic function of tungsten-containing aldehyde ferredoxin oxidoreductase from the hyperthermophilic and proteolytic archaeon *Thermococcus* Strain ES-1. *J Bacteriol* 177(16):4757-4764.
263. www.eia.gov (2015).
264. Berner RA (2003) The long-term carbon cycle, fossil fuels and atmospheric composition. *Nature* 426(6964):323-326.
265. Dien BS, Bothast RJ, Nichols NN, & Cotta MA (2002) The US corn ethanol industry: An overview of current technology and future prospects. *Int Sugar J* 104(1241):204-211.
266. https://www.fueleconomy.gov/ (2015).

267. Escobar JC*, et al.* (2009) Biofuels: environment, technology and food security. *Renew Sust Energ Rev* 13(6):1275-1287.
268. Schmer MR*, et al.* (2014) Energy potential and greenhouse gas emissions from bioenergy cropping systems on marginally productive cropland. *PLoS ONE* 9(3):e89501.
269. Congress U (2007) Energy independence and security act of 2007. *Public Law* 2.
270. Svetlitchnyi V*, et al.* (2013) Single-step ethanol production from lignocellulose using novel extremely thermophilic bacteria. *Biotechnol Biofuel* 6(1):31.
271. Himmel ME*, et al.* (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. *science* 315(5813):804-807.
272. Field CB, Behrenfeld MJ, Randerson JT, & Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281(5374):237-240.
273. Leopoldina GNAoS (2012) Bioenergy – Chances and limits. Halle (Saale).
274. Demain AL (2009) Biosolutions to the energy problem. *J Ind Microbiol Biot* 36(3):319-332.
275. Cowan D (1991) Enzymes from thermophilic archaebacteria: current and future applications in biotechnology. *Biochem Soc Symp*, pp 149-169.
276. Gupta R, Gigras P, Mohapatra H, Goswami VK, & Chauhan B (2003) Microbial α-amylases: a biotechnological perspective. *Process Biochem* 38(11):1599-1616.
277. Alvira P, Tomas-Pejo E, Ballesteros M, & Negro MJ (2010) Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. *Bioresource Technol* 101(13):4851-4861.
278. Yang B & Wyman CE (2008) Pretreatment: the key to unlocking low-cost cellulosic ethanol. *Biofuel Bioprod Bior* 2(1):26-40.
279. Lynd LR, Weimer PJ, van Zyl WH, & Pretorius IS (2002) Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol Mol Biol R* 66(3):506-577.
280. Fontes CMGA & Gilbert HJ (2010) Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Ann Rev Biochem* 79(1):655-681.
281. Argyros DA*, et al.* (2011) High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. *Appl Environ Microb* 77(23):8288-8294.
282. Rainey FA*, et al.* (1994) Description of *Caldicellulosiruptor saccharolyticus* gen. nov., sp. nov: an obligately anaerobic, extremely thermophilic, cellulolytic bacterium, . *FEMS Microbiol Lett* 120(3):263-266.
283. Hudson RC, Schofield LR, Coolbear T, Daniel RM, & Morgan HW (1991) Purification and properties of an aryl beta-xylosidase from a cellulolytic extreme thermophile expressed in Escherichia coli. *Biochem. J.* 273(Pt 3): 645-650.
284. Gibbs MD, Elinder, A. U., Reeves, R. A. & Bergquist, P. L. (1996) Sequencing, cloning and expression of a beta-1,4-mannanase gene, manA, from the extremely thermophilic anaerobic bacterium, Caldicellulosiruptor Rt8B.4. *FEMS Microbiol Lett* 141:37-43.
285. Gibbs MD*, et al.* (2000) Multidomain and multifunctional glycosyl hydrolases from the extreme thermophile *Caldicellulosiruptor* isolate Tok7B. *Curr Microbiol* 40(5): 333-340.
286. Zverlov V, Mahr S, Riedel K, & Bronnenmeier K (1998) Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile '*Anaerocellum*

thermophilum' with separate glycosyl hydrolase family 9 and 48 catalytic domains. *Microbiology* 144 (Pt 2):457-465.
287. Blumer-Schuette SE, Kataeva I, Westpheling J, Adams MWW, & Kelly RM (2008) Extremely thermophilic microorganisms for biomass conversion: status and prospects. *Curr Opin Microbiol* 19(3):210-217.
288. van de Werken HJG*, et al.* (2008) Hydrogenomics of the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *Appl Environ Microb* 74(21):6720-6729.
289. Blumer-Schuette S*, et al.* (2013) Thermophilic lignocellulose deconstruction. *FEMS Microbiol Rev* Epub ahead of print:1-56.
290. Ng TK, Benbassat A, & Zeikus JG (1981) Ethanol production by thermophilic bacteria - Fermentation of cellulosic substrates by cocultures of *Clostridium thermocellum* and *Clostridium thermohydrosulfuricum*. *Appl Environ Microb* 41(6):1337-1343.
291. Shaw AJ, Hogsett DA, & Lynd LR (2010) Natural competence in *Thermoanaerobacter* and *Thermoanaerobacterium* species. *Appl Environ Microb* 76(14):4713-4719.
292. Blumer-Schuette SE, Lewis DL, & Kelly RM (2010) Phylogenetic, microbiological, and glycoside hydrolase diversities within the extremely thermophilic, plant biomass-degrading Genus *Caldicellulosiruptor*. *Appl Environ Microb* 76(24):8084-8092.
293. Kataeva I*, et al.* (2013) Carbohydrate and lignin are simultaneously solubilized from unpretreated switchgrass by microbial action at high temperature *Energy & Environmental Science* 6:2186-2195.
294. Zurawski JV, Blumer-Schuette S, Conway JM, & Kelly RM (2014) The extremely thermophilic genus *Caldicellulosiruptor*: physiological and genomic characteristics for complex carbohydrate conversion to molecular hydrogen. *Advances in Photosynthesis and Respiration Including Bioenergy and Related Processes. Volume 38: Microbial BioEnergy: Hydrogen Production*, eds Zannoni D & De Philippis R (Springer).
295. Blumer-Schuette SE*, et al.* (2012) *Caldicellulosiruptor* core and pangenomes reveal determinants for noncellulosomal thermophilic deconstruction of plant biomass. *J Bacteriol* 194(15):4015-4028.
296. Dam P*, et al.* (2011) Insights into plant biomass conversion from the genome of the anaerobic thermophilic bacterium Caldicellulosiruptor bescii DSM 6725. *Nucleic Acids Res* 39(8):3240–3254.
297. Bielen AA*, et al.* (2010) Pyrophosphate as a central energy carrier in the hydrogen-producing extremely thermophilic *Caldicellulosiruptor saccharolyticus*. *FEMS Microbiol Lett* 307:48-54.
298. Izquierdo J, Sizova MV, & Lynd LR (2010) Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. *Appl Environ Microb* 76(11):3545-3553.
299. Lochner A*, et al.* (2011) Use of label-free quantitative proteomics to distinguish the secreted cellulolytic systems of *Caldicellulosiruptor bescii* and *Caldicellulosiruptor obsidiansis*. *Appl Environ Microb* 77(12):4042–4054.
300. Brunecky R*, et al.* (2013) Revealing nature's cellulase diversity: the digestion mechanism of *Caldicellulosiruptor bescii* CelA. *Science* 342(6165):1513-1516.

301. Ozdemir I, Blumer-Schuette SE, & Kelly RM (2012) S-Layer homology domain proteins *Csac*_0678 and *Csac*_2722 are implicated in plant polysaccharide deconstruction by the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. *Appl Environ Microb* 78(3):768-777.
302. Lochner A*, et al.* (2011) Label-free quantitative proteomics for the extremely thermophilic bacterium *Caldicellulosiruptor obsidiansis* reveal distinct abundance patterns upon growth on cellobiose, crystalline cellulose, and switchgrass. *J Proteome Res* 10:5302–5314.
303. Yang SJ*, et al.* (2009) Efficient degradation of lignocellulosic plant biomass, without pretreatment, by the thermophilic anaerobe "*Anaerocellum thermophilum*" DSM 6725. *Appl Environ Microb* 75(14):4762-4769.
304. Benner R, Maccubbin AE, & Hodson RE (1984) Anaerobic biodegradation of the lignin and polysaccharide components of lignocellulose and synthetic lignin by sediment microflora. *Appl Environ Microb* 47(5):998-1004.
305. Ko J-J*, et al.* (2009) Biodegradation of high molecular weight lignin under sulfate reducing conditions: lignin degradability and degradation by-products. *Bioresource Technol* 100(4):1622-1627.
306. DeAngelis KM*, et al.* (2012) Anaerobic decomposition of switchgrass by tropical soil-derived feedstock-adapted consortia. *Mbio* 3(1).
307. Wu Y-R & He J (2013) Characterization of anaerobic consortia coupled lignin depolymerization with biomethane generation. *Bioresource Technol* 139:5-12.
308. Dien BS, Cotta MA, & Jeffries TW (2003) Bacteria engineered for fuel ethanol production: current status. *Appl Microbiol Biotechnol* 63(3):258-266.
309. Jin MJ, Gunawan C, Balan V, & Dale BE (2012) Consolidated bioprocessing (CBP) of AFEX (TM)-pretreated corn stover for ethanol production using *Clostridium phytofermentans* at a high solids loading. *Biotechnol Bioeng* 109(8):1929-1936.
310. Willquist K, Zeidan AA, & van Niel EW (2010) Physiological characteristics of the extreme thermophile *Caldicellulosiruptor saccharolyticus*: an efficient hydrogen cell factory. *Microbial Cell Fact* 9.
311. Carere CR*, et al.* (2012) Linking genome content to biofuel production yields: a meta-analysis of major catabolic pathways among select H_2 and ethanol-producing bacteria. *Bmc Microbiology* 12:295.
312. Schut GJ & Adams MWW (2009) The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. *J Bacteriol* 191(13):4451-4457.
313. Chung D, Cha M, Farkas J, & Westpheling J (2013) Construction of a stable replicating shuttle vector for *Caldicellulosiruptor* species: use for extending genetic methodologies to other members of this genus. *PLos One* 8(5):e62881.
314. Chung DW, Farkas J, & Westpheling J (2013) Overcoming restriction as a barrier to DNA transformation in *Caldicellulosiruptor* species results in efficient marker replacement. *Biotechnology for Biofuels* 6.

315. Cha M, Chung DW, Elkins JG, Guss AM, & Westpheling J (2013) Metabolic engineering of *Caldicellulosiruptor bescii* yields increased hydrogen production from lignocellulosic biomass. *Biotechnology for Biofuels* 6.
316. Chang TH & Yao S (2011) Thermophilic, lignocellulolytic bacteria for ethanol production: current state and perspectives. *Appl Microbiol Biotechnol* 92(1):13-27.
317. Demain AL, Newcomb M, & Wu JHD (2005) Cellulase, clostridia, and ethanol. *Microbiol Mol Biol R* 69(1):124-154.
318. Yi ZL, Su XY, Revindran V, Mackie RI, & Cann I (2013) Molecular and biochemical analyses of CbCel9A/Cel48A, a highly secreted multi-modular cellulase by *Caldicellulosiruptor bescii* during growth on crystalline cellulose. *PLos One* 8(12).
319. Sun Q, Madan B, Tsai SL, DeLisa MP, & Chen W (2014) Creation of artificial cellulosomes on DNA scaffolds by zinc finger protein-guided assembly for efficient cellulose hydrolysis. *Chem Commun* 50(12):1423-1425.
320. Vazana Y, Morais S, Barak Y, Lamed R, & Bayer EA (2012) Designer cellulosomes for enhanced hydrolysis of cellulosic substrates. *Cellulases* 510:429-452.
321. van Niel EWJ, Claassen PAM, & Stams AJM (2003) Substrate and product inhibition of hydrogen production by the extreme thermophile, *Caldicellulosiruptor saccharolyticus*. *Biotechnol Bioeng* 81(3):255-262.
322. Yee KL*, et al.* (2012) Evaluation of the bioconversion of genetically modified switchgrass using simultaneous saccharification and fermentation and a consolidated bioprocessing approach. *Biotechnol Biofuel* 5:81.
323. Biswas R, Prabhu S, Lynd LR, & Guss AM (2014) Increase in Ethanol Yield via Elimination of Lactate Production in an Ethanol-Tolerant Mutant of *Clostridium thermocellum*. *PLoS ONE* 9(2):e86389.
324. Blankenship RE*, et al.* (2011) Comparing Photosynthetic and Photovoltaic Efficiencies and Recognizing the Potential for Improvement. *Science* 332(6031):805-809.
325. Hawkins AS, McTernan PM, Lian H, Kelly RM, & Adams MWW (2013) Biological conversion of carbon dioxide and hydrogen into liquid fuels and industrial chemicals. *Curr Opin Biotech* 24(3):376-384.
326. Nocera DG (2012) The Artificial Leaf. *Accounts Chem Res* 45(5):767-776.
327. Udupa K, Subramanian G, & Udupa H (1971) The electrolytic reduction of carbon dioxide to formic acid. *Electrochim Acta* 16(9):1593-1598.
328. Berg IA (2011) Ecological Aspects of the Distribution of Different Autotrophic CO2 Fixation Pathways. *Appl Environ Microb* 77(6):1925-1936.
329. Erb TJ (2011) Carboxylases in Natural and Synthetic Microbial Pathways. *Appl Environ Microb* 77(24):8466-8477.
330. Bar-Even A, Noor E, & Milo R (2012) A survey of carbon fixation pathways through a quantitative lens. *J Exp Bot* 63(6):2325-2342.
331. Bar-Even A, Noor E, Lewis NE, & Milo R (2010) Design and analysis of synthetic carbon fixation pathways. *Proc Natl Acad Sci* 107(19):8889-8894.
332. Boyle NR & Morgan JA (2011) Computation of metabolic fluxes and efficiencies for biological carbon dioxide fixation. *Metab Eng* 13(2):150-158.

333. Fast AG & Papoutsakis ET (2012) Stoichiometric and energetic analyses of non-photosynthetic CO2-fixation pathways to support synthetic biology strategies for production of fuels and chemicals. *Curr Opin Chem Eng* 1(4):380-395.
334. Lee HS*, et al.* (2008) The complete genome sequence of *Thermococcus onnurineus* NA1 reveals a mixed heterotrophic and carboxydotrophic metabolism. *J Bacteriol* 190(22):7491-7499.
335. Lipscomb GL*, et al.* (2014) Engineering hydrogen gas production from formate in a hyperthermophile by heterologous production of an 18-subunit membrane-bound complex. *J Biol Chem* 289(5):2873-2879.
336. Glasemacher J, Bock A-K, Schmid R, & Schönheit P (1997) Purification and properties of acetyl-CoA synthetase (ADP-forming), an Archaeal enzyme of acetate formation and ATP synthesis, from the hyperthermophile *Pyrococcus furiosus*. *Eur J Biochem* 244(2):561-567.
337. Jones DT, Shirley M, Wu X, & Keis S (2000) Bacteriophage Infections in the Industrial Acetone Butanol(AB) Fermentation Process. *J Mol Microb Biotech* 2(1):21-26.
338. Gleick PH (1993) Water and conflict: Fresh water resources and international security. *Int Security* 18(1):79-112.
339. Small C & Nicholls RJ (2003) A global analysis of human settlement in coastal zones. *J Coastal Res*:584-599.
340. Ferber D (2004) Synthetic biology. Microbes made to order. *Science* 303(5655):158-161.
341. Benner SA & Sismour AM (2005) Synthetic biology. *Nat Rev Genet* 6(7):533-543.
342. Andrianantoandro E, Basu S, Karig DK, & Weiss R (2006) Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* 2(1).
343. www.afdc.energy.gov/ (2015).
344. poet-dsm.com (2015).

APPENDIX A

PRODUCTION OF LIGNOFUELS AND ELECTROFUELS BY EXTREMELY THERMOPHILIC MICROBES

Keller MW, Loder AJ, Basen M, Izquierdo J, Kelly RM, Adams MWW. (2015) Production of lignofuels and electrofuels by extremely thermophilic microbes. *Biofuels* 5(5): 499-515.
Reprinted here with permission of the publisher Taylor & Francis

Abstract

Extreme thermophiles are microorganisms that grow optimally at elevated temperatures (≥ 70°C). They could play an important role in the emerging renewable energy landscape by exploiting thermophily to produce liquid transportation fuels. For example, Caldicellulosiruptor species can grow on unpretreated plant biomass near 80°C utilizing novel multi-domain glycoside hydrolases. Through metabolic engineering, advanced biofuels compatible with existing infrastructure liquid biofuels, so-called lignofuels, could be produced to establish consolidated bioprocessing at high temperatures. In another case, a new paradigm, electrofuels, addresses the inefficiency of biofuel production through the direct synthesis of advanced fuels from carbon dioxide using hydrogen gas as the electron carrier. This requires coupling of biological electron utilization to carbon dioxide fixation and ultimately to fuel synthesis. Using a hyperthermophilic host *Pyrococcus furiosus* and synthetic metabolic pathways comprised of genes from less thermophilic sources, temperature-regulated biosynthesis of industrial organic chemicals and liquid fuel molecules are possible. Herein, we review recent progress towards the synthesis of lignofuels and electrofuels by extremely thermophilic microorganisms.

Introduction

For the past century, the unabashed utilization of fossil fuels has resulted in an explosion in both the population and standard of living across much of the planet. The result has been the most rapid advancement humanity has seen (6). However, this has been accompanied by an insatiable demand for energy that in the very near future will not be met through the use of fossil fuels (10). Their finite nature will reach a critical point at which the rate of extraction will irreversibly begin to decrease and asymptotically approach zero. For the production of crude oil, this phenomenon is widely known as 'peak oil' (263). Peak oil is believed to have occurred in the United Kingdom in 1999, while in the United States it is predicted to occur in 2020 when production reaches 13 million barrels per day (263). Although the timing of peak oil on a global scale is controversial, even the most optimistic estimates place it within the lifespan of the current generation and their immediate offspring (263). In addition, the use of fossil fuels has been accompanied by a dramatic increase (40%) in atmospheric CO_2 (from 280 to 400 ppm) (20). Transportation related to fossil fuel use is estimated to account for 30-40% of anthropogenic CO_2 emissions (263). Even more controversial than projections for fossil fuel consumption is the relationship between atmospheric CO_2 concentrations and global warming. Nevertheless, while the extent to which atmospheric CO_2 and climate change are anthropogenic can be debated (264), without a renewable energy supply, humanity will not survive regardless of CO_2 and climate.

Developing renewable technologies capable of supporting future civilization is clearly of paramount importance. To this end, biofuels offer important opportunities. Of the renewable

liquid transportation fuels aimed at replacing the current need for gasoline, diesel and aviation fuel, ethanol currently constitutes 94% of the world's biofuel production (263). The highly developed bioethanol industry (265) is led by the United States and Brazil, using corn starch and sugar cane as primary feedstocks, respectively (263). Bioethanol in the United States is ubiquitously found as a 10% blend in gasoline (E10) and, to a lesser extent, as an 85% blend (E85) that requires engine modification (266). As a result, approximately 9% of gasoline used in the United States contains bioethanol (263). However, the use of ethanol as a fuel faces serious limitations. In addition to the federally mandated blend wall at 10%, vehicles utilizing E85 will experience a 25-30% reduction in fuel economy, due to the lower energy density of ethanol compared to gasoline (266). Moreover, ethanol is not compatible with the current liquid fuel infrastructure, particularly pipelines, because of its hygroscopic and corrosive nature (56, 57). The result of these limitations is that the U.S. ethanol market is currently saturated and production has reached a plateau (263). Consequently, efforts in this area are now turning to so-called 'advanced biofuels' that have superior fuel characteristics compared to ethanol. Advanced biofuels include higher alcohols, terpenoids and alkanes, all of which have higher energy density, lower hygroscopicity, and greater infrastructure compatibly (59, 185-187). Despite these advantages, current strategies for advanced biofuel production are still largely based on sugar from corn starch and sugar cane. This creates an inevitable conflict between energy and food production (267), in addition to an extremely low carbon efficiency since the majority of the carbon is re-released as CO_2 during the production of the fuel (268). Resolving the issues of food vs fuel, infrastructure compatible fuels, as well as the efficiency of vehicles and buildings is a primary goal of U.S. and European policymakers, as outlined in the Energy

Independence and Security Act of 2007 (EISA 2007) and the European Parliament directive 2009/28/EC on the promotion of the use of energy from renewable sources (53, 54).

Biomass as a source of carbon for biofuels

The strategy outlined in EISA 2007 calls for the production of cellulosic ethanol, which in comparison to current starch-based methods, relies on utilizing the substantial, but not easily accessible, cellulosic- and hemicellulosic-based carbohydrate components of plants (269). Cellulosic biofuels have, however, proven to be difficult to produce, as evidenced by the fact that current production is far below the EISA 2007 targets (263). The primary hurdle to producing cellulosic biofuels is that plant biomass is extremely resistant to degradation. This property, referred to as 'recalcitrance', inhibits access to the fermentable sugars comprising lignocellulose (209, 270, 271).

Nevertheless, conversion of lignocellulosic biomass to liquid fuels is a promising alternative to fossil fuels in the decades to come. The global net primary production of land plants is estimated to be ~65 Gt carbon per year (272). Although net primary production and energy consumption are not equally distributed around the planet, the light energy annually, if converted to chemical energy by land plants, exceeds the world's energy demand by a factor of three to four (273). While it is not realistic to completely exploit this energy store, these estimates nonetheless demonstrate the enormous potential of lignocellulosic biofuels, or what will be termed here 'lignofuels'.

The major constituents of plant biomass are cellulose (40−50%), a glucose (C_6) polymer), hemicellulose (23−38%), a polymer of xylose (C_5) and a variety of other sugars, and lignin

(19−38%), a complex polymer of aromatic units (274), with carbohydrates comprising about ~60% of the chemical energy stored in the plant. The conversion of lignocellulose to lignofuels involves fermentation of sugars, requiring oxygen-free (anaerobic) conditions. An anaerobic biochemical process for lignin degradation has not been reported, nevertheless, microbial fermentations could convert the chemical energy stored in the cellulose and hemicellulose into fuels. Hence, the conversion of lignocellulose would still be about three to four times more energy efficient than the conversion of corn starch to ethanol (274).

Despite their promise for addressing the 'energy gap' anticipated in the coming decades, there are significant challenges that face biofuels if they are to become a significant part of the overall liquid transportation fuel budget. As mentioned above, they should be produced from non-food, plant biomass derived materials (cellulose and hemicellulose). Furthermore, to improve process economics, these plant biomass feedstocks should require minimal (or preferably no) chemical/physical pretreatment, prior to the fermentation step. Yet another possibility for future biofuels production is to forego the use of plant biomass altogether and to directly convert point sources of CO_2 directly into liquid transportation fuels. Direct biological conversion of unpretreated plant biomass and direct conversion of CO_2 to liquid fuels are attractive possibilities that have recently been gaining attention. In particular, approaches for converting unpretreated plant biomass and CO_2 directly to biofuels are being considered using extremely thermophilic microorganisms, which grow optimally (T_{opt}) above 70°C. Here, we explore the very recent progress that has been made using extreme thermophiles to address current and future biofuels needs.

Biotechnological application of thermophiles

Growth temperature is a canonical characteristic in the classification of microbes. While the exact nomenclature varies, 'psychrophilic' usually refers to an organism with a T_{opt} below 15°C; 'mesophilic' organisms have a T_{opt} between 15 and 45°C; and 'thermophilic' organisms grow optimally above 45°C. However, the temperature range for thermophily is very broad, extending to 122°C, the current upper temperature limit of life (77). Therefore, thermophiles have been further divided into moderate thermophiles (T_{opt} 45-70°C) and extreme thermophiles ($T_{opt} \geq 70$°C, e.g. *C. bescii* and *P. furiosus*); special cases of extreme thermophiles growing with a $T_{opt} \geq 80$°C (e.g. *P. furiosus*) are typically referred to as hyperthermophiles (78, 79).

Due to their inherent stability against denaturation, thermophilic enzymes have been utilized in many biotechnological applications, such as polymerases for DNA manipulation, dehydrogenases and esterases for chemical synthesis, and amylases for starch liquefaction (275, 276). A case in point is a thermostable DNA polymerase used in the Polymerase Chain Reaction (PCR) for DNA amplification, a technique essential for the development of modern biotechnology. PCR requires a DNA polymerase capable of surviving repeated cycles of heating to temperatures approaching the normal boiling point of water. Thermostability of proteins also correlates positively with their resistance to other denaturing conditions, such as detergents and organic solvents (73). These properties allow thermophilic enzymes to play a role in various chemical processes under conditions that otherwise would be too harsh for conventional enzymes from more conventional mesophilic microbes. While thermophiles have

been considered for biofuel production for some time, it is only very recently with the emergence of molecular genetics tools that efforts with these microbes have greatly intensified. As described in the following, bioenergy applications that exploit thermophily have now been envisioned and demonstrated, and hold great promise in addressing current energy supply challenges.

LIGNOFUELS: Utilizing recalcitrant plant biomass for fuel production

Recalcitrance of plant biomass is the major obstacle for their efficient and economic conversion to fuels. It arises from both the crystalline nature of the cellulose microfibrils in lignocellulose and the inert polyaromatic nature of lignin. This problem has been traditionally addressed by harsh thermochemical and physical pre-treatments of plant biomass, followed by enzymatic digestion of the liberated cellulose and hemicellulose polymers (277, 278), ultimately yielding oligomeric sugars available for ethanol fermentation by yeast and bacteria. Consolidated bioprocessing (CBP) has been suggested as an alternative, more economical process (279). CBP involves the use of cellulolytic microorganisms as comprehensive biocatalysts, combining the steps of enzyme production, enzymatic hydrolysis and fermentation. However, at present, no single microorganism or consortium of microorganisms can produce ethanol or another biofuel from unpretreated lignocellulose at commercially relevant yields and titers. Indeed, the ability to degrade crystalline cellulose is a relatively rare ability among microorganisms in general.

Thermophilic microbes able to breakdown and utilize lignocellulosic substrates and convert them to other products belong to a phylogenetic group of microbes known as the

Firmicutes, mostly from the genera *Clostridium* (such as *Cl. thermocellum*, T_{opt} 55-60°C), *Thermoanaerobacter* (T_{opt} 60-75°C), *Thermoanaerobacterium* (T_{opt} 55-70°C), and *Caldicellulosiruptor* (T_{opt} 65-78°C). Among these naturally cellulolytic species, *Cl. thermocellum* is the most studied. In particular, it contains a large membrane-bound multienzyme complex known as the cellulosome that functions to attach to and degrade crystalline cellulose (280). However, while *Cl. thermocellum* efficiently degrades cellulose at ~60°C, it does not utilize C5 sugars derived from hemicellulose nor does it produce a sufficient amount or yield of ethanol, despite extensive efforts to genetically engineer its fermentation pathways (167). A co-culture of a genetically engineered strain of *Cl. thermocellum* and a strain of the ethanologenic *Thermoanaerobacterium saccharolyticum*, however, produced almost 40 g L^{-1} ethanol from 92 g L^{-1} crystalline cellulose (281). While this is a promising step towards CBP, it remains to be seen if the use of co-cultures of microbes is effective on lignocellulosic substrates (rather than crystalline cellulose).

The genus *Caldicellulosiruptor*, a clade of extremely thermophilic, gram positive, anaerobic, and asporogenous bacteria, contains species that grow on a broad range of lignocellulosic materials as growth substrates. The first known *Caldicellulosiruptor* species was isolated from a terrestrial hot spring in the Rotorua region of New Zealand's North Island in 1987 and named *Caldocellum saccharolyticum* (282). This bacterium was of great interest since it was able to produce an array of glycoside hydrolases (GHs) capable of degrading β-linked complex polysaccharides at temperatures at or above its optimal growth temperature of approximately 75°C (283-286). Despite the fact that several other related species were isolated

from terrestrial hot springs around the world over the subsequent two decades, there was limited interest in their applications towards the production of renewable chemicals (287, 288). However, the recent push for biofuels from lignocellulosic substrates has heightened interest in *Caldicellulosiruptor* species (270, 289). These bacteria thrive in terrestrial hot springs by using plant biomass as growth substrates in the form of fallen trees and branches or from runoff from adjacent grassy areas. *Caldicellulosiruptor* species are prolific cellulose degraders, and they do so at the highest temperatures known for this process (270, 290, 291).

Diversity of Caldicellulosiruptor species

Caldicellulosiruptor species are globally distributed throughout North America, Iceland, Russia, Japan, and New Zealand and have optimum growth temperatures between 70°C and 78°C (292). Not only are these species able to grow on pure crystalline cellulose and hemicellulose, but also on industrially-relevant loadings of unpretreated biomass (200 g/L), digesting up to 85% of insoluble unpretreated switchgrass, as in the case of *C. bescii(293)*. *Caldicellulosiruptor* species ferment the sugars liberated from the biomass substrate to molecular hydrogen, lactate, acetate, and small amounts of alcohol (294). These characteristics make *Caldicellulosiruptor* species attractive candidates for development into CBP microorganisms (287) **(Figure 6.1)**. To date, eight genomes of *Caldicellulosiruptor* species have been sequenced, revealing a core genome comprised of 1543 genes shared by all eight species that is non-cellulolytic (295). The 'pan-genome', which is the collection of all genes in all eight species (4009 genes), was found to be open, indicating that sequencing of new species should reveal novel genes (295). Our recent review (289) catalogs the biochemically characterized

carbohydrate-metabolizing enzymes (CAZymes) from *Caldicellulosiruptor* species. Within the genus, cellulolytic capability varies widely, with 5 species, namely *C. bescii*, *C. kronotskyensis*, *C. obsidiansis*, *C. lactoaceticus*, and *C. saccharolyticus* having higher cellulolytic ability than the other species with sequenced genomes (295). Because of their high cellulolytic activity, the focus here is on this core group of five species, in particular the more extensively studied *C. bescii*, *C. obsidiansis* and *C. saccharolyticus*.

Carbohydrate-Active EnZymes (CAZymes) in *Caldicellusiruptor*

The ability of lignocellulolytic organisms to effectively breakdown plant biomass is conferred by the diversity and specificity of carbohydrate-active enzymes (CAZymes) they can produce. Of these, glycoside hydrolases (GHs) are of particular importance since they target the most challenging and recalcitrant portions of plant biomass, i.e. cellulose and hemicellulose. *Caldicellulosiruptor* species characteristically have a large inventory of GHs in their genomes that generate a variety of pentose and hexose monosaccharides from lignocellulose, which are metabolized simultaneously without regulation by carbon catabolite repression (284, 287, 288). When looking at the five most highly cellulolytic *Caldicellulosiruptor* species, they all possess a combination of three extracellular GHs present across these genomes: GH9-CBM3-CBM3-CBM3-GH48, GH74-CBM3-CBM3-GH48, and GH9-CBM3-CBM3-CBM3-GH5, where CBM stands for Carbohydrate-Binding Module. A key features of these *Caldicellulosiruptor* CAZymes is not only are they multi-modular, but also that they combine very different types of active sites, which also makes them multi-functional (**Table 6.1**). Comparisons with other *Caldicellulosiruptor* species as well as other genera (295, 296) indicate that the prevalence of

CAZymes with multi-modular architecture is unique among cellulolytic bacteria (289). In addition, the presence of GH48 and CBM3 domains appear to be essential for crystalline cellulose degradation, compared to the weakly cellulolytic species where GH5 and GH9 enzymes are not associated with GH48 domains (295, 297). Most cellulolytic organisms contain only one GH48 enzyme, whereas *Caldicellulosiruptor* species contain multiple copies with different functionalities, conveyed by the other GH domains linked to them through multi-modular configurations. The only other instance for multiple GH48 domains is found in cellulolytic thermophiles (T_{opt} 55-60^{o}C), such as *Clostridium* (*Cl.*) *thermocellum*, *Cl. clariflavum* and *Cl. straminisolvens*, which possess two copies of GH48, one cellulosomal and one secreted as a free enzyme (298). Secretome analyses of both *C. bescii* and *C. obsidiansis* have revealed that the most abundant multi-functional, multi-domain cellulases contained GH5, GH9, GH10, GH43, GH44, GH48, and GH74 domains, as well as highly conserved CBM3 domains (299). Combinations of these 7 GH families are found throughout all the *Caldicellulosiruptor* genomes, but especially in the most cellulolytic species (**Table 6.1**). In addition to GH domains, multi-modular CAZymes in *Caldicellulosiruptor* species can also contain catalytic domains from the PL (polysaccharide lyase) and CE (carbohydrate esterase) families, allowing for a broader spectrum of catalytic synergism from individual enzymes.

CelA, a primary cellulase in *Caldicellulosiruptor* species

CelA is one of the largest, most thermostable, and most active GH enzymes yet identified for crystalline cellulose and it is ubiquitous in the most highly cellulolytic *Caldicellulosiruptor* species (See **Table 7.1**, GH9-CBM3-CBM3-CBM3-GH48). CelA is the most highly secreted cellulase when *C. bescii* grows on plant biomass (299). Recently, it has been reported that this cellulase is capable of excavating through the middle of the cellulose microfibrils creating cavities approximately 15-20 nm wide and 15-30 nm long, whereas previously characterized cellulases degrade the polymer from the ends of crystalline cellulose microfibrils (300). This hydrolysis mechanism promotes the formation of new cellulose fiber chain ends that can be synergistically digested by other cellulases produced by *Caldicellulosiruptor* species. Compared to thermophilic model organisms, such as *Clostridium thermocellum*, that rely on the membrane-bound multi-complex cellulosome (280), *Caldicellulosiruptor* species, therefore, represent a new paradigm for microbial cellulose degradation by way of these complex, multi-domain secreted cellulases.

Mechanisms of adhesion to plant biomass

Compared to other thermophilic lignocellulosic organisms where the cellulosome plays a key role in the attachment to plant biomass, *Caldicellulosiruptor* species have a number of diverse mechanisms that allow cells to maintain proximity with the lignocellulose substrate. *Caldicellulosiruptor* species have a broad range of S-layer homology (SLH) proteins that contain domains that have affinity to the outer para-crystalline protein layer in these organisms, as well as complex arrangements of carbohydrate binding domains and glycoside hydrolases. One key

enzyme in this category was found in the core pan-genome and contains GH5 and CBM28 domains. This enzyme plays a key role in adherence to plant biomass by C*. saccharolyticus* (301), *C. bescii* (296) and *C. obsidiansis* (302). Beyond this SLH protein shared by the core genome, many others have been identified in the *Caldicellulosiruptor* species, with large multi-modular domains. In addition to SLH domains, other classes of proteins have been associated with adherence to cellulose include a type 4 pilus (T4P) in *C. bescii* (296) and *C. obsidiansis* (302), and a two putative adhesins (295). It, therefore, seems very clear that the breadth of mechanisms involved in surface-microbe interactions in the *Caldicellulosiruptor* species has only begun to be uncovered and merits further exploration.

Caldicellulosiruptor plant biomass conversion

Among the few microorganisms capable of deconstructing plant biomass, *Caldicellulosiruptor* species possess the highest temperature optima, approaching 80°C (249), and they exploit the heat for accelerated lignocellulose degradation (293). Moreover, *C. bescii* degrades plant biomass without thermochemical pretreatment (293, 303), an enormous technological and economic advantage that has not been demonstrated for other microorganisms. In addition, concentrations of acid-pretreated switchgrass higher than 20 g L^{-1} caused a significant growth inhibition, likely due to substances released by the pretreatment process (209). After growth of three successive cultures of *C. bescii* on unpretreated switchgrass, where the residual plant biomass remaining after growth of one culture served as the carbon and energy source for the next, approximately 80% of the switchgrass was solubilized and able to pass through a 20 µm filter (293). Surprisingly, the remaining 20% was

not a 'ball of lignin', as might be expected. Rather, the cellulose/hemicellulose/lignin ratio was similar to that of the unprocessed biomass. These data suggest that *C. bescii* deconstructs switchgrass by an 'onion-peeling' mechanism. Since the amount of biomass carbohydrate that was degraded matched the carbon products that were generated by *C. bescii* (293), even using high loads (209), lignin is clearly not serving as a carbon or energy source for the organism. It would appear that degradation of the polymeric carbohydrate components by *C. bescii* also releases lignin complexes (< 20 μm), thereby exposing additional carbohydrate and enabling continued deconstruction of the biomass. Microbial enrichment cultures that degrade lignin under anaerobic conditions have been previously described (304-307), however, neither a biochemical pathway for the anaerobic degradation of lignin nor an enzyme involved in lignin activation without oxygen has been characterized. Such a process or enzyme, however, would be of significant interest within the fields of biochemistry and biotechnology. It is not known if *C. bescii* produces enzymes that can degrade components of the lignin directly, or if lignin release is a consequence of the hydrolysis of plant carbohydrates.

For biomass to biofuel conversion on an industrial scale, a titer of at least 4% (40 g L^{-1}) ethanol has to be achieved (308, 309), which is equivalent to about 110 g L^{-1} untreated switchgrass (209, 293). While *Caldicellulosiruptor* species do not yet generate ethanol using this substrate concentration, it is technologically promising that concentrations of switchgrass of more than 100 g L^{-1} are readily used as a carbon source and do not inhibit growth of *C. bescii*. This also suggests that acid-pretreatment is potentially dispensable and, perhaps, counterproductive for CBP based on this microbe (209). In pH-controlled media-optimized

fermentations using 50 g L^{-1} substrate loads, *C. bescii* completely degraded ~30 g L^{-1} crystalline cellulose and ~10 g L^{-1} of unpretreated switchgrass. All substrate carbon could be accounted for in the products, based on the amount of substrate degraded, with acetate and CO_2 being the major end products (209).

Unlike *Cl. thermocellum*, *Caldicellulosiruptor* species possess a complete pentose phosphate pathway, so that they can use C_5 sugars derived from hemicellulose in addition to C_6 sugars (295). Major fermentation end products are acetate and hydrogen, with only traces of ethanol produced, although very recently a *Caldicellulosiruptor* species has been isolated that produced as much as 72 mM ethanol from cellulose although it has yet to be fully characterized (270). Lactate production is also observed, but this can be avoided if the hydrogen partial pressure is kept low (209). While *Caldicellulosiruptor* species are not yet optimized for the production of liquid biofuels, they are very effective in channeling reductant towards H_2. In fact, molecular hydrogen production in *C. saccharolyticus* has been shown to approach the Thauer limit of 4 moles H_2 per mol glucose (310). In sugar fermentation, *Caldicellulosiruptor* species produce both ferredoxin (by the pyruvate ferredoxin oxidoreductase reaction) and NAD(P)H (in glycolysis) (294, 311), and both can obviously be re-oxidized by hydrogen production. *Caldicellulosiruptor* genomes encode for a bifurcating hydrogenase (312) that presumably accounts for the majority of hydrogen production.

Current limitations and perspectives

The recent development of a genetic system in *C. bescii* (313, 314) now enables gene deletions (315) and potentially expression of foreign genes and pathways. These developing tools will be very helpful in elucidating the role of different CAZy proteins in biomass degradation, and also make possible metabolic engineering of *Caldicellulosiruptor* species for efficient degradation of lignocellulosic biomass for liquid biofuel production (**Figure 6.1).** However, pathways for the production of liquid biofuels are rare in thermophilic microbes. *Cl. thermocellum* (T_{opt} 55°C) and different *Thermoanaerobacter* species (T_{opt} 60−75°C) can produce ethanol as the major end product of sugar fermentation (164, 206, 238, 311, 316). These obligately anaerobic and thermophilic ethanologens use a different pathway than that used in facultative anaerobic mesophilic bacteria and yeasts, wherein pyruvate decarboxylase produces acetaldehyde from pyruvate. The acetaldehyde is then further reduced to ethanol by a primary alcohol dehydrogenase. In contrast, thermophilic anaerobes oxidize pyruvate to acetyl-CoA, a reaction catalyzed by pyruvate ferredoxin oxidoreductase. A key enzyme for ethanol production is a bi-functional aldehyde/alcohol dehydrogenase, referred to as AdhE (163, 238) that catalyzes a two-step reduction of acetyl-CoA to ethanol.

Since the growth temperatures of *Caldicellulosiruptor* and *Thermoanaerobacter* species are similar (311), expression of enzymes from ethanologenic *Thermoanaerobacter* species, in particular alcohol dehydrogenases, is a promising engineering strategy for ethanol production in *Caldicellulosiruptor* species. As an alternative strategy, *Thermoanaerobacter* species could be grown in co-culture with *Caldicellulosiruptor* species. In that scenario, the latter would

breakdown the biomass and provide a certain fraction of the cellulose and hemicellulose derived sugars to the non-cellulolytic but ethanologenic *Thermoanaerobacter* species. This approach has been used with *Cl. thermocellum,* which cannot utilize C_5 sugars, growing together with non-cellulolytic *Thermoanaerobacterium* or *Thermoanaerobacter* species on defined C_6 and C_5 sugar substrates and resulting in 2- to 4-fold increases in ethanol titer (281, 290, 317). Currently, there is only one example of a *Thermoanaerobacter-Caldicellulosiruptor* co-culture, which showed a higher ethanol yield from cellulose fermentation than the pure *Caldicellulosiruptor* culture, ranging between 175% to 210% higher ethanol yields on the mol % basis (270). A possible third strategy is to identify all essential enzymes for plant biomass breakdown in *Caldicellulosiruptor*, and then genetically engineer an ethanologenic *Thermoanaerobacter* or *Thermoanaerobacterium* species to utilize cellulose and hemicellulose. Given the recent development of genetic tools in these organisms, this may be possible in the not too distant future (167, 291). An advantage is that the heterologous expression of secreted extracellular enzymes (e.g., CelA (318)) might be easier to achieve than heterologous production of complex cellulosomes (319, 320).

Another obstacle to overcome is the efficient deconstruction of lignocellulose. *C. bescii* completely degrades up to 60% of 50 g L^{-1} crystalline cellulose in pH-controlled fermentations. However, conversion of soluble sugars (cellulose, glucose) is incomplete, with sugars accumulating in the medium (209). This shows, on the one hand, that the glycoside hydrolases are relatively insensitive to inhibition by end products, but on the other hand, that *C. bescii* itself is sensitive. Specifically, the fermentation end products acetate and lactate accumulate in

the medium to concentrations above 150 mM, inhibiting growth and further sugar fermentation (209, 321). This inhibition is mainly due to ionic strength rather than osmotic pressure (209) and, of course, is likely overcome by engineering a strain that produces a neutral end product, such as ethanol or butanol.

Less is known about why *Caldicellulosiruptor* species stop degrading unpretreated switchgrass. Interestingly, independent of the initial switchgrass concentration, the same percentage (~30%) of unpretreated switchgrass is degraded by *C. bescii* (209, 293). The spent media from pH-controlled fermentations contain a significant fraction of unidentified sugars that represent ~20% of the product carbon. *C. bescii* did not grow on the so-called 'spent medium', which is that remaining at the end of a switchgrass fermentation, while a related microorganism, *Thermoanaerobacter mathranii,* did grow on the same spent medium. This suggests a species-specific inhibitor may be released into the medium during growth on switchgrass. While it has been shown that *C. bescii* tolerates high concentrations of unpretreated plant biomass, it will be crucial to elucidate and ameliorate the inhibitory mechanism to efficiently convert the lignocellulosic biomass. A likely source for inhibitory compounds is phenolic lignin monomers that are released during lignocellulose deconstruction by *C. bescii* (293, 322). However, it remains unclear why the inhibition is concentration-independent.

To date, research into the production of fuels in thermophiles has focused on ethanol, just as it has in mesophiles. Much has been done with thermophilic bacteria to increase ethanol productivity, particularly in *Thermoanaerobacter*, *Caldicellulosiruptor*, and *Clostridium* species

(164, 270, 323). However, there is little research thus far on engineering thermophiles to produce advanced biofuels, and there are no reports using extreme thermophiles or hyperthermophiles. There is one report of advanced biofuel production in a moderate thermophile (55°C), in which *Thermoanaerobacterium saccharolyticum* was engineered to produce *n*-butanol from xylose at 1 g/L titers and 26% theoretical yield (195). The reason for the dearth of advanced biofuels produced by thermophiles is the lack of metabolic pathways to produce such molecules in these organisms. Discovering or engineering these pathways merits further investigation and may require environmental sampling and sequencing of hot springs and/or combining known pathways into new configurations for fuel synthesis**.**

It is clear that *Caldicellulosiruptor* species are promising candidate organisms for CBP, particularly as they deconstruct high concentrations of unpretreated plant biomass at temperatures above 70°C. They take advantage of an arsenal of extracellular and cell surface attached CAZymes (295), some of which are among the most efficient cellulolytic enzymes known (300). The recent development of a genetic system (313-315) will lead to a better understanding of essential enzymes for plant biomass breakdown, and how these bacteria could be engineered to efficiently convert high concentrations of plant biomass to biofuels.

ELECTROFUELS: using CO_2 directly as a carbon source for fuel production

The use of plant biomass for biofuel production depends on the photosynthetic capture of solar energy, which operates at approximately 1% energy efficiency on an annual basis (**Figure 6.2**). The resulting low efficiency of this process poses a significant problem in moving to the necessary industrial scales (70, 324). A new non-photosynthetic paradigm for renewable fuels, 'electrofuels', exploits a variety of autotrophic pathways for the direct conversion of fully oxidized inorganic carbon into infrastructure-compatible fuels. This occurs via the incorporation of low potential electrons from inorganic energy forms such as hydrogen gas or an electrical current. This strategy was designed explicitly to avoid the inefficient photosynthetic production of carbohydrate intermediates in favor of a more direct and efficient approach (70, 71, 192, 325). In the electrofuels strategy, the energy supply, such as hydrogen gas, is an energy carrier and not the actual energy source. At present, the ultimate energy sources for these electrofuels are still fossil fuel, but in the not-too-distant future electrofuel production would be driven by solar energy. Hence, rather than using the solar energy to store electrons in carbohydrates via plant-based photosynthesis, the electrofuels would be generated using a variety of recent technological advances to capture solar energy. For example, photovoltaic-driven electrolysis of water to hydrogen (with 10% energy efficiency) has been shown to be roughly 10-fold more efficient than plant-based photosynthesis and 3-fold more efficient than microalgae based photosynthesis (3% energy efficiency) (324). Additionally, wireless systems are available, such as the 'artificial leaf', that catalyze the direct solar driven splitting of water using earth-abundant materials and near-neutral conditions, and do so at 2.5% energy efficiency (326). In addition to hydrogen gas, formate is an efficient energy carrier electrochemically produced

from carbon dioxide with high efficiency (327). Moreover, it is highly soluble, which is critical for efficient mass transfer in an aqueous reaction, and its electrons are readily accessible via biological means via formate dehydrogenase-type enzymes. Formate and hydrogen have similar energy contents. For example, the hydrogen electrode and the formate/CO_2 redox couples have the same reduction potential (E_o' = -420 mV, pH 7.0).

Biological CO_2 fixation

Assimilation of inorganic carbon in the form of CO_2 is key to the concept of electrofuels. There are currently six pathways of CO_2 fixation known to occur in nature, four of which are found in extreme thermophiles: the reductive tricarboxylic acid (rTCA) cycle, the reductive acetyl-CoA or Wood-Ljungdahl (W/L) pathway, the 3-hydroxypropionate/4-hydroxybutyrate (3-HP/4-HB) cycle, and the dicarboxylate/4-hydroxybutyrate (DC/4-HB) cycle. The reductive pentose phosphate cycle and 3-hydroxypropionate bicycle occur in moderate thermophiles, but not in extreme thermophiles growing above 70°C. The nature of all six carbon-fixation pathways has recently been reviewed extensively (178, 328-331) and so the focus herein is on the impact of pathway features for electrofuels production.

Each of the CO_2 fixation pathways found in extreme thermophiles produces acetyl-CoA as its product, but they differ in the identity of electron donors, oxygen tolerance and energetics. The rTCA, W/L, and DC/4-HB cycles use NAD(P)H and ferredoxin as electron donors, while the 3-HP/4-HB cycle uses only NADPH (71). The rTCA and 3-HP/4-HB cycles are oxygen-tolerant and known to operate in aerobes, despite the use of oxygen-sensitive ferredoxin in the rTCA cycle (323, 328). In contrast, the W/L and DC/4-HB pathways are found only in anaerobes,

likely due to the oxygen sensitivity of the key enzymes CO dehydrogenase/acetyl-CoA synthase and pyruvate synthase, respectively. The thermodynamic and kinetic characteristics of the carbon-fixation cycles have been evaluated due to their impact on electrofuel and biofuel production and on global carbon cycling. However, different conclusions can be reached as to which one is “best”, depending on the parameter being considered. For example, maximizing biomass production in the context of photorespiration results in the rTCA cycle having the highest energetic efficiency (332). A recent review (333) suggests that the W/L pathway is the best option when viewed from the perspective of energy efficiency in utilizing CO_2 and H_2. However, as concentrations of gaseous substrates decrease with increasing process temperature, the faster kinetics exhibited by the 3-HP/4-HB and DC/4-HB cycles may become important (330). For biotechnological applications, the 3-HP/4-HB cycle would be more desirable than the DC/4-HB cycle since, in contrast to the 3-HP/HB cycle (148, 151, 257, 258), most but not all of the genes for the DC/4-HB cycle have been identified (151).

The characteristics of carbon fixation pathways obviously impact their suitability for use in electrofuels production. While there are a variety of options for mesophilic microbes, robust genetics tools are not available for thermophilic autotrophs that naturally fix CO_2 directly. Hence, rather than engineering into a CO_2-fixing, autotrophic thermophile a pathway to produce a biofuel, efforts have focused on the recombinant expression of a CO_2-fixation pathway in a thermophilic heterotrophic host. However, given the complexity of the CO_2-fixing pathways, this presents a major challenge for electrofuels production at high temperature. In addition, the pathway must also integrate within and be compatible with the host’s

metabolism. This includes providing the appropriate electron donor(s) to the pathway, accounting for oxygen tolerance, and supplying the appropriate amount of ATP. This compatibility, while necessary, may also pose an issue as the host's metabolism can also interfere with the recombinant pathway leading to lowered efficiency. For this reason, extreme thermophiles are an attractive option since decreasing the operating temperature could allow for the optimal operation of a recombinant pathway in a background of minimized host metabolism (71, 192, 325). In fact, as described below, this strategy has already been employed in which a metabolic pathway from a thermophile has been expressed in a hyperthermophilic host. Upon a temperature shift from the temperature of the hyperthermophilic host (90-100°C) to the recombinant thermophilic pathway (60-80°C), the metabolic activity and potential interference of the host decreases, while the activity of the recombinant pathway becomes optimal.

Progress towards the production of electrofuels in a hyperthermophilic host

The hyperthermophile *P. furiosus* is an attractive option for metabolic engineering via genetic modification. It is a marine anaerobe that is capable of growing between 70 and 103°C (T_{opt} 100°C) by fermenting carbohydrates to hydrogen gas, carbon dioxide, and acetate (102). A naturally competent variant was discovered during attempts to introduce DNA in *P. furiosus* (137). This variant was used to construct a markerless deletion of the gene *pyrF* for uracil biosynthesis, and the resulting strain, designated COM1, became the basis for the development of an extremely robust genetic system (137). The wide temperature range of *P. furiosus* and the availability of a genetic system have enabled the temperature-dependent production in the

organism of the enzyme lactate dehydrogenase from, coincidentally, *Caldicellulosiruptor bescii* (T_{opt} 78°C). The gene encoding this NADH-dependent enzyme was expressed in *P. furiosus* and this resulted in a temperature dependent shift of the metabolic end product. At a high growth temperature (98°C), the strain produced the wild-type end products carbon dioxide and acetate; however, upon a temperature shift to 72°C, there was a shift in the metabolism toward lactate production (172).

One of the first genetic manipulations of *P. furiosus* resulted in the homologous overexpression of its soluble NADP-dependent hydrogenase I (SHI), resulting in a 10-fold increase over that normally produced (176). This enzyme is of particular biotechnological interest as it catalyzes the reversible oxidation of hydrogen gas and reduction of NADP to NADPH (123, 179), a reaction that could be used to drive an NADPH-dependent biosynthetic pathway with reductant from hydrogen gas. Hydrogen can be added to the electrofuel-producing system in a variety of ways. The simplest method, i.e. gas-sparging, faces the challenge of low solubility and mass transfer. An alternative method would be to supply a highly soluble electron carrier, such as formate, and produce the substrate hydrogen gas within the cells themselves. This has also been demonstrated in *P. furiosus* by the heterologous expression of an 18-subunit membrane-bound formate-hydrogen lyase (FHL) from an organism with a T_{opt} of 80°C (334, 335). The resulting *P. furiosus* strain is capable of converting high concentrations of formate to hydrogen and carbon dioxide, which can potentially be used directly as substrates for electrofuels synthesis (335).

As previously discussed, the electrofuels strategy is defined by directly coupling the use of low potential reductant to the assimilation of fully oxidized inorganic carbon (CO_2). In *P. furiosus*, the native hydrogen uptake system (SHI) can be coupled to the 3-HP/4-HB cycle for carbon fixation via the compatible electron carrier NADPH (123, 148, 151, 179, 257, 258). The strategy for electrofuels production in *P. furiosus* is based on the heterologous expression of the 3-HP/4-HB carbon fixation pathway from the archaeon *Metallosphaera sedula,* which grows optimally at 73°C (**Figure 6.3**). To demonstrate this as proof of principle, the five genes encoding the first three enzymes of the pathway were expressed in *P. furiosus.* In combination these catalyze the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA and subsequent two reductive steps to the cycle-intermediate 3-hydroxypropionate. The result was hydrogen- and carbon dioxide-dependent formation of 3-hydroxypropionate and validation of the electrofuels concept in *P. furiosus* (192). This process was further improved upon by studying strategic deletions aimed at directing the flow of acetyl-CoA toward the heterologous pathway and away from the metabolism of the host (201). The host normally converts acetyl-CoA to acetate while simultaneously producing ATP via the enzyme acetyl coenzyme A synthase (ACS) (119, 336). A single deletion in this pathway resulted in a 3-fold increase in 3-hydroxyproprionate production in comparison to the original strain (201).

This strategy for the recombinant expression of the *M. sedula* 3-HP/4-HB pathway in *P. furiosus* involved dividing the cycle into three sub-pathways and transferring them one at a time. The initial strain expressed the first three enzymes and produced the key intermediate 3-hydroxypropionate (192). While this is a small step towards making a fuel molecule, 3-

hydroxypropionate itself is a molecule of significant biotechnological interest. In addition to being a commercial chemical produced at commodity-scale itself, 3-hydroxypropionate can readily be converted to the commodity chemicals 1,3-propanediol, acrylate, methyl acrylate, and acrylamide. This progress represents a biological and renewable means to generate a molecule (3-hydroxypropionate) that constitutes a large and petroleum dependent chemical market. The resulting 3-hydroxypropionate is listed among the Department of Energy's list of top value added chemicals from biomass (174). Significant progress along these lines illustrates the potential for commercial bioprocesses based on the electrofuels paradigm.

Large-scale bioprocessing at elevated temperatures

Utilizing a hyperthermophile at a large-scale raises the question of how to provide sufficient energy to heat a microbial fermentation. However, this may prove to be a significant advantage of a hyperthermophilic platform organism. Large-scale fermentations of mesophilic organisms require a substantial and constant cooling effort due to the energy released from active metabolism (74). This is typically carried out by circulating refrigerated glycol and represents the major energy consuming process for a large scale fermentation (75). On the other hand, for a hyperthermophile at the appropriately large scale (hundreds of thousands to millions of liters), this energy can serve to maintain the temperature. This would result in the required heat exchange being limited to a one-time heating expenditure rather than a constant cooling requirement for the entire duration of the process. Moreover, a hyperthermophilic process requiring a temperature drop of 30°C or more will be more easily accomplished due to the large temperature differential between the fermentation and the ultimate heat acceptor,

the ambient air. This is due to the rate of heat flow being directly proportional to the respective temperature difference. In the case of a hyperthermophilic fermentation versus one at ambient temperature, heat will flow very quickly from the 80-100°C fermenter to the ~25°C air. For a mesophilic process at 37°C, the smaller temperature differential is insufficient for rapid heat transfer, and refrigerated glycol is required for effective cooling. The temperature control of a hyperthermophilic process could therefore represent an enormous benefit, with refrigerated glycol replaced with active air circulation as the primary means of cooling.

The use of metabolic activity for heating and circulated air for cooling would allow for a strategy that cycles at different temperatures (**Figure 6.4**) between that of a hyperthermophilic host (90-100°C) and a recombinant thermophilic pathway (60-80°C). The first stage of such a process would be an initial heating of the growth medium from approximately room temperature to the growth temperature of the hyperthermophilic host (90-100°C). This represents a single large energy investment, but also an opportunity to lower the risk of contamination. Mesophilic fermentations typically require sterilization at 121°C prior to inoculation and subsequent sterilization cycles to minimize contamination, a large energy investment given the scale of operation. In addition to cooling, losses due to contamination represent a substantial cost to current industrial fermentations (337). This is typically due to wild-type mesophilic organisms naturally inhabiting the area where the fermentation plant is located. Such mesophilic contaminants would be eradicated during an initial growth stage at 90-100°C. Thermophiles are still prevalent in the environment, so a complete lack of sterile technique might still lead to contamination. Utilizing a marine hyperthermophile, such as *P.*

furiosus, affords the additional advantage of not requiring fresh water. Utilizing primarily sea water would alleviate serious concerns over the limited nature of fresh water (338) and would geographically localize the process to the coast, which is where human population density is greatest (339).

Once the hyperthermophilic host has grown to a sufficient density, the fermentation would then be cooled to a temperature optimal for the heterologous pathway (60-80°C) via active air circulation. During this phase, the genes encoding the thermophilic enzymes would be expressed and the liquid biofuel would be produced. This shift to a lower temperature has the key advantage of reducing the metabolism of the host. This is critical to an industrial process as it reduces the host maintenance energy and interference from the host's metabolism. These advantages are analogous to the goals of the emerging field of synthetic biology (340-342). Rather than creating artificial organisms, this goal can be accomplished by a simple shift in temperature of a hyperthermophilic host containing a designed thermophilic process. The advantages of this temperature shift would be difficult to exploit in a mesophilic host, as it would require psychrophilic gene donors and would be limited by the freezing temperature of the growth medium. Finally, the fermentation process would be cycled back to a high temperature (90-100°C) for regeneration of the catalyst (re-growth of the strain) and simultaneous distillation of volatile fuel molecules (343) in preparation for another cycle of production formation (60-80°C) (**Figure 6.4**). These heating and cooling cycles would be accomplished using metabolically generated heat and active air circulation, potentially requiring only small additional energy inputs.

Initial pilot projects for large-scale hyperthermophilic fuel production would be of similar size to current cellulosic ethanol pilot plants, which produce up to 25×10^6 gal/year (344) and require fermentation volumes on the order of $10^5 – 10^6$ L. This scale can take advantage of heating through active metabolism as discussed above. Second-generation plants would likely be even larger to take further advantage of economies of scale.

Future Perspectives

While still in early stages, there are a number of promising directions for the use of thermophiles in solving problems with producing liquid transportation fuels and industrial chemicals. Certainly, other developments will positively impact this situation. Microbially-enabled consolidated bioprocessing of recalcitrant plant biomass for lignofuel production is an important opportunity, but other strategies are being investigated to make lignofuels relevant to a renewable future. For example, the engineering of plant biomass to reduce recalcitrance and advanced strategies for pre-treatment and carbohydrate extraction may prove vital to the production of lignocellulose-derived fuels. By improving carbon efficiencies and utilizing non-competing crops such as switchgrass or food crop waste such as corn stover, lignofuels are sure to play a role in the near and distant future of sustainable energy.

The development of electrofuels as an advanced strategy for biofuels has been a part of the expanding diversity of biotechnological research of carbon fixation pathways. Where plant-based biofuels rely only on the Calvin-Benson cycle and the enzyme Rubisco, the research on electrofuels has explored a variety of alternate carbon fixation pathways that nature has to offer. For example the 3-HP bicycle, the 3-HP/4-HB cycle, and the W/L pathway have also been

exploited in electrofuels strategies. These strategies have seen varying degrees of success. Some, such as efforts by OPX Biotechnologies focusing on engineering bacteria for efficient fuel production (www.opxbio.com/), have attracted private sector investment. Others on liquid fuel from renewable electricity and bacteria have led to the formation of new companies. These success stories are part of the electrofuels program, which is funded by the Advanced Research Project Agency – Energy (ARPA-E: http://arpa-e.energy.gov/). ARPA-E is aimed at advancing high-risk and high-reward technologies to improve the U.S. economic prosperity, national security, and environmental wellbeing. This strategy continues to be a success as ARPA-E projects have attracted more than $625 million in private funding to date. This agency will continue to drive advancements such as metabolic engineering and photovoltaics as a part of the upcoming revolution of renewable energy.

Tables and Figures

Table 7.1 Distribution of multimodular enzymes containing multiple GH domains across the *Caldicellulosiruptor* genomes

Multidomain architecture	Gene loci	Catalytic domain activity
GH43-CBM22-**GH43**-CBM6	Cbes_0182, Calkro_2388, Csac_2411, Calow_0121	xylanase, arabinanase
GH10-CBM3-CBM3-**GH48**	Cbes_1857, Calkro_0853	endo-1,4-β-xylanase, cellobiohydrolase
GH5-CBM3-CBM3-**GH44**	Cbes_1859, Calkro_0851, Csac_1077	mannanase, endoglucanase
GH74-CBM3-CBM3-(CBM3)-**GH48**	Cbes_1860, Calkro_0861, Csac_1085, COB47_1664, Calla_0015	xyloglucanase, cellobiohydrolase
GH9-CBM3-CBM3-CBM3-**GH5**	Cbes_1865, Calkro_0855, COB47_1669, Csac_1079	endoglucanase, mannanase
GH5-CBM3-CBM3-CBM3-**GH5**	Cbes_1866, Calkro_0854	mannanase, cellulase
GH9-CBM3-CBM3-CBM3-**GH48**	Cbes_1867, Calkro_0850, Csac_1076, COB47_1673	endoglucanase, cellobiohydrolase
GH10-CBM3-CBM3-**GH5**	COB47_1671	endo-1,4-β-xyalanse, endoglucanase
GH10-CBM3-**GH5**	Csac_1078	endo-1,4-β-xyalanse, endoglucanase
PL-CBM3-CBM3-CBM3-**GH44**	Calla_0017	rhamnogalacturonan lyase, endoglucanase
CBM22-CBM22-**GH10**-CBM3-CBM3-CBM3-**GH43**-CBM6	AAD30363	endo-1,4-β-xylanase, arabinanase

The abbreviations are GH, glycoside hydrolase; CBM, carbohydrate binding domain; PL, polysaccharide lyase. The gene loci are shown for *C. bescii* (Cbes), *C. kronotskyensis* (Calkro), *C. saccharolyticus* (Csac), *C. owensensis* (Calow), *C. obsidiansis* (COB47), *C. lactoaceticus* (Calla), and *Caldicellulosiruptor* sp. Tok7B.1 (accession number only)

Figure 7.1 The proposed strategy for lignofuel production in *Caldicellulosiruptor* sp.

Native extracellular and cell surface enzymes deconstruct untreated plant biomass at elevated temperature. They release oligomeric C_5 and C_6 sugars that are taken up and fermented to produce reductant (reduced ferredoxin and NAD(P)H) and acetyl-CoA. While in non-engineered strains, reductant oxidation is coupled to hydrogen production, reductant will be used in engineered strains to reduce acetyl-CoA to advanced biofuels or lignofuels.

Figure 7.1

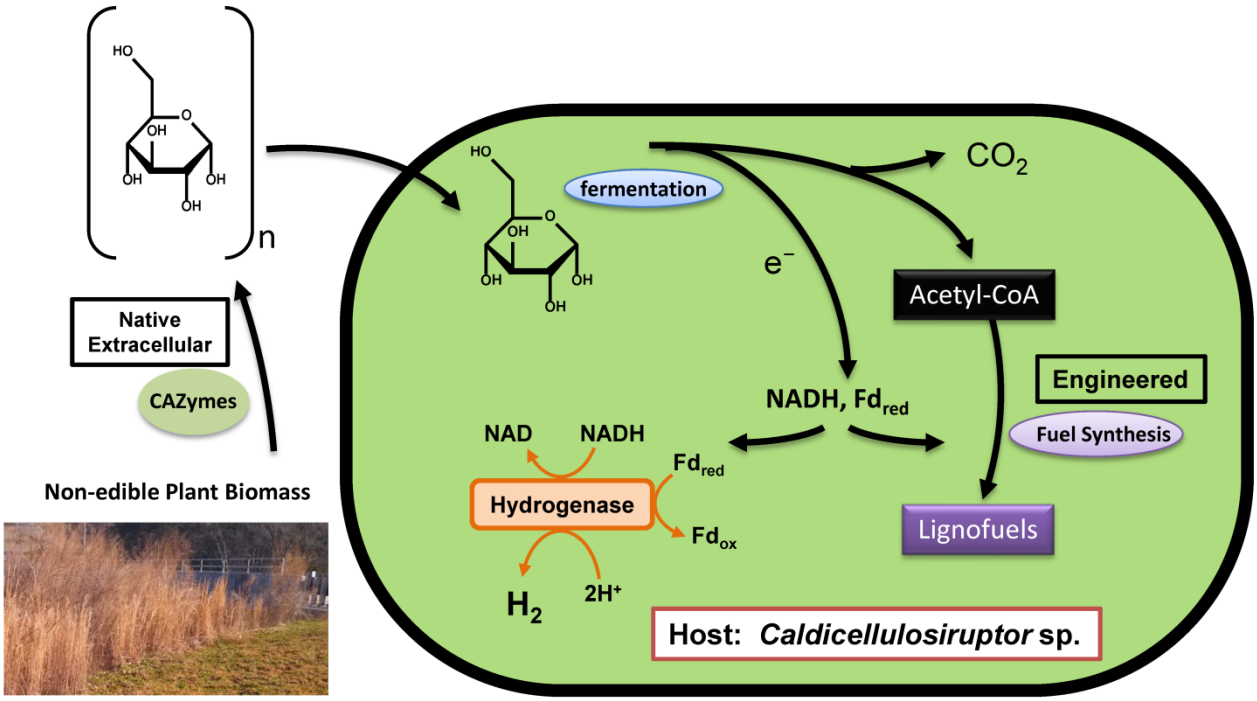

HO
O
OH
OH
OH
OH
n
Native Extracellular
CAZymes
Non-edible Plant Biomass
fermentation
CO_2
e^-
Acetyl-CoA
Engineered
Fuel Synthesis
NADH, Fd_{red}
NAD
NADH
Fd_{red}
Hydrogenase
Fd_{ox}
Lignofuels
H_2
$2H^+$
Host: *Caldicellulosiruptor* sp.

Figure 7.2 The current biofuels strategy versus lignofuels and electrofuels

The current biofuel strategy depends on the plant based capture of solar energy in the form of carbohydrate intermediates. This process has a measured overall energy efficiency of 0.2% in the conversion of photons to fuel. Lignofuels will utilize non-edible plant biomass that doesn't compete with food production and greatly increases the carbon efficiency. Electrofuels is based on the photovoltaic capture of solar energy. This energy is coupled to a carbon fixation and fuel synthesis pathway via an inorganic electron carrier such as hydrogen gas. This process has a potential energy efficiency of up to 7.5%, depending on the strategy for fuel synthesis.

Figure 7.2

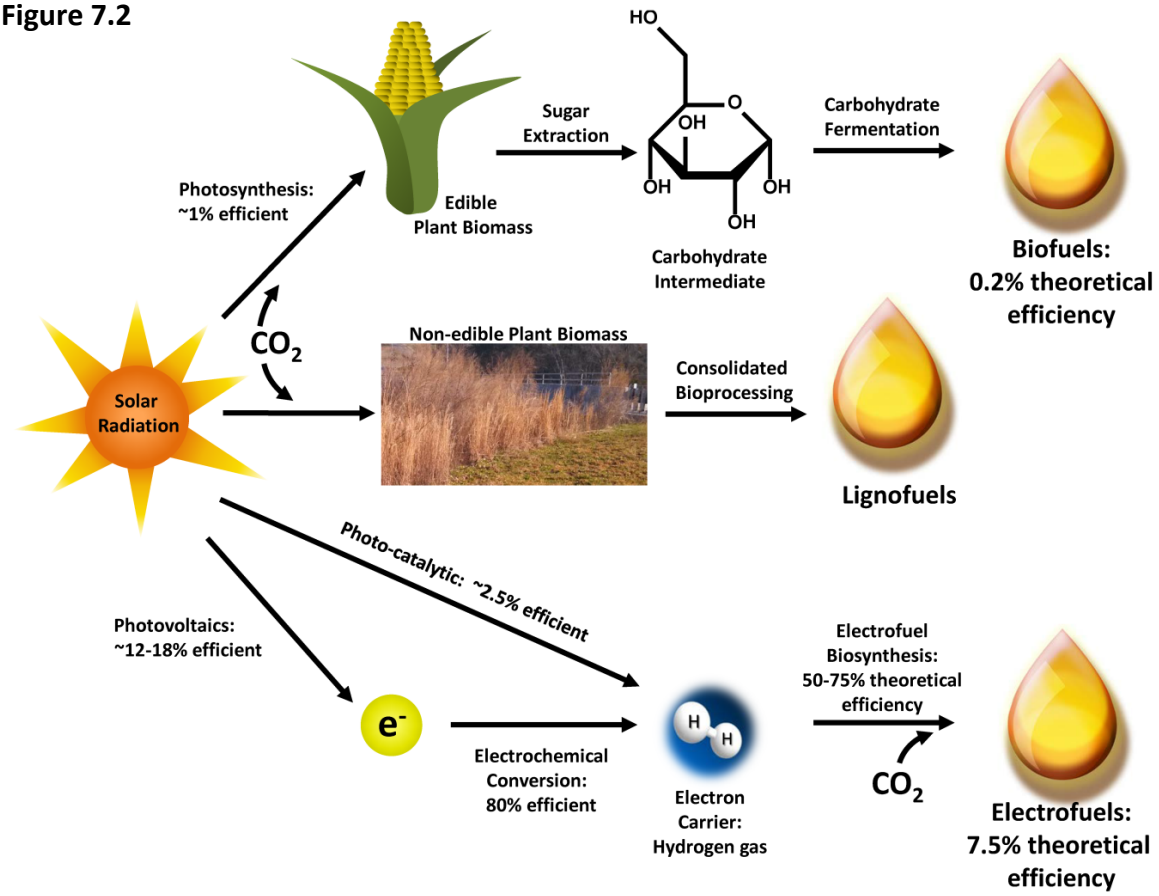

Photosynthesis:
~1% efficient
Edible
Plant Biomass
Sugar
Extraction
HO
O
OH
OH
OH
OH
Carbohydrate
Intermediate
Carbohydrate
Fermentation
Biofuels:
0.2% theoretical
efficiency
Solar
Radiation
CO_2
Non-edible Plant Biomass
Consolidated
Bioprocessing
Lignofuels
Photo-catalytic: ~2.5% efficient
Photovoltaics:
~12-18% efficient
e^-
Electrochemical
Conversion:
80% efficient
H
H
Electron
Carrier:
Hydrogen gas
Electrofuel
Biosynthesis:
50-75% theoretical
efficiency
CO_2
Electrofuels:
7.5% theoretical
efficiency

Figure 7.3 The proposed strategy for electrofuel production in *Pyrococcus furiosus*

The native hydrogen uptake system, SH1, will be coupled to the engineered 3-HP/4-HB carbon fixation pathway from *M. sedula* via the compatible electron carrier NADPH. The generated acetyl-CoA will then be fed into an advanced biofuel synthesis pathway. Key intermediates for the carbon fixation pathway are: acetyl-CoA, 3-hydroxypropionate (3-HP), and 4-hydroxybutyrate (4-HB).

Figure 7.3

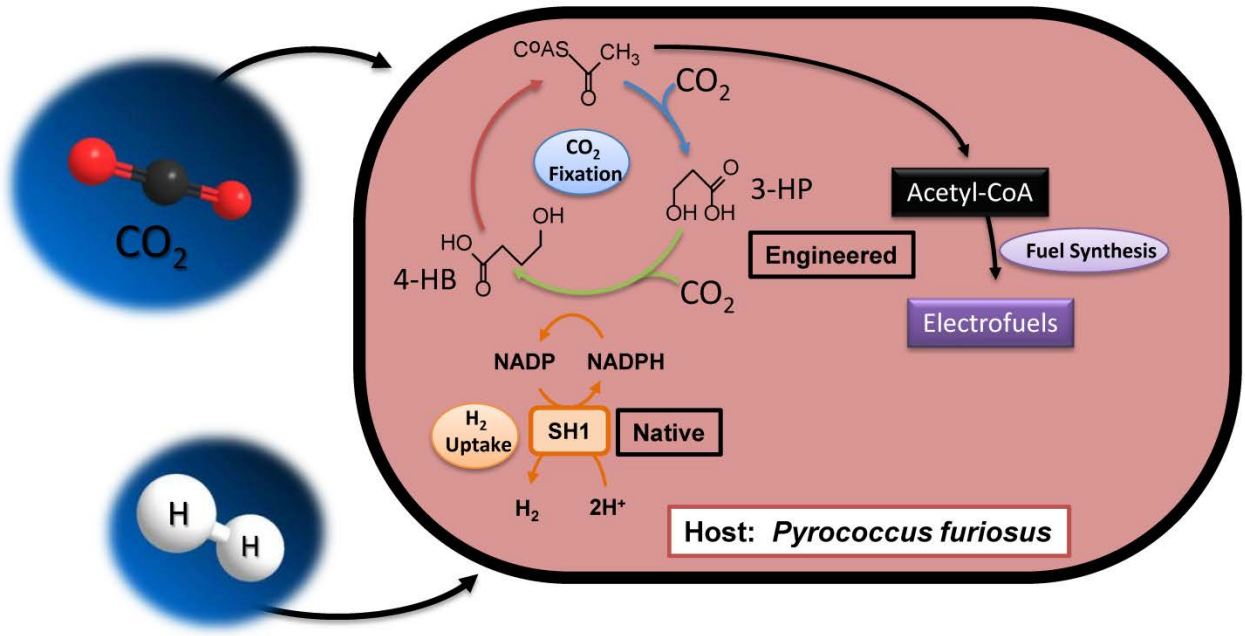
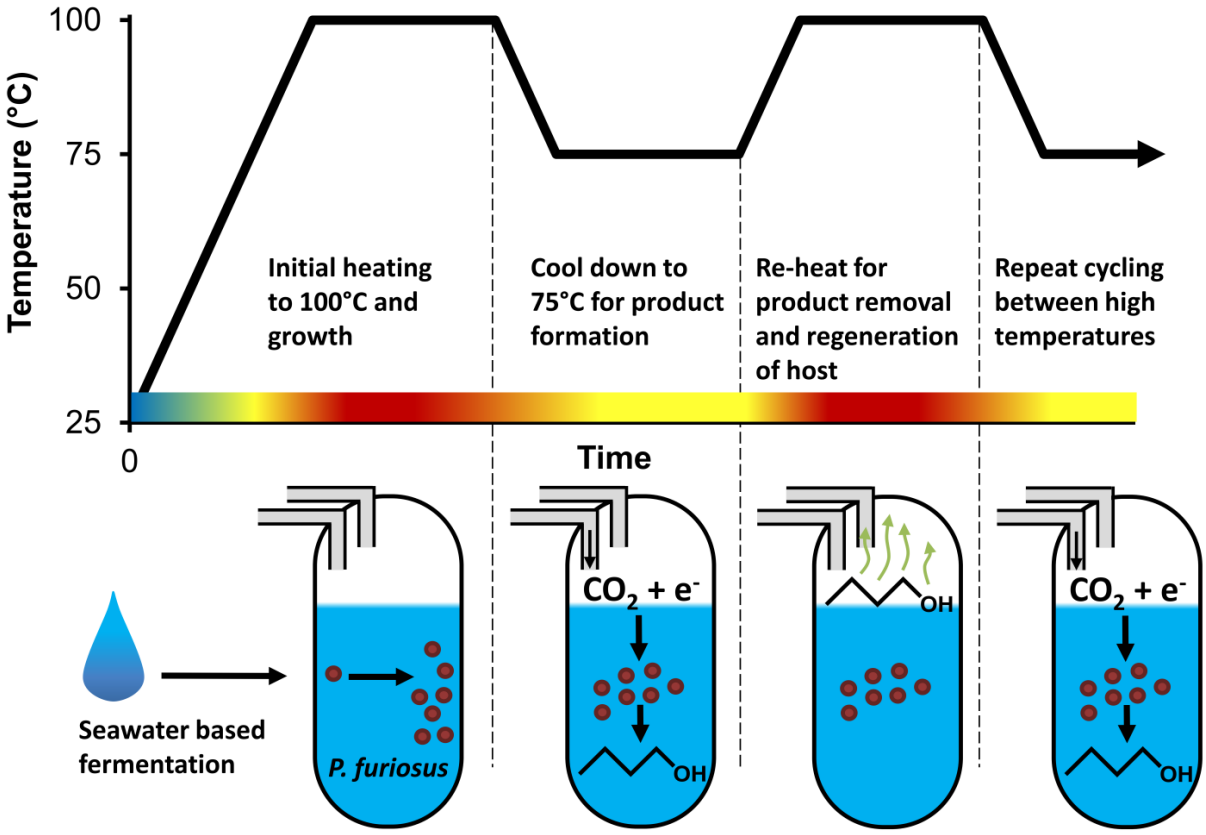

CO2
H
H
COAS
CH3
CO2
CO2 Fixation
3-HP
OH OH
HO
OH
4-HB
CO2
Acetyl-CoA
Engineered
Fuel Synthesis
Electrofuels
NADP
NADPH
H2 Uptake
SH1
Native
H2
2H+
Host: Pyrococcus furiosus

Figure 7.4 The temperature cycling strategy for the industrial use of a hyperthermophilic host (90-100°) and a recombinant thermophilic pathway (60-80°)

The first heating stage would allow for the initial growth of the host followed by a temperature shift for product formation. Afterwards, the temperature would be raised again for product removal and the regeneration of the host. This cycling would be repeated as long as product formation is occurring.

Figure 7.4


100
75
50
25
0
Temperature (°C)
Time
Initial heating to 100°C and growth
Cool down to 75°C for product formation
Re-heat for product removal and regeneration of host
Repeat cycling between high temperatures
Seawater based fermentation
P. furiosus
CO2 + e-
OH
CO2 + e-
OH
OH